



# 6σ

## **MINITAB (v15-v19)**

### **Exercices de base (niveau LMD)**

**Vincent ISOZ V25.0 Révision 58 (2021-12-06), {oUUID 1.676}**

MINITAB® and all other trademarks and logos for the Company's products and services are the exclusive property of Minitab Inc. All other marks referenced remain the property of their respective owners. See [minitab.com](http://minitab.com) for more information.



# SPCC

# TABLE DES MATIÈRES

1.	Introduction.....	8
1.1.	À propos du rédacteur .....	11
1.2.	Avertissements .....	12
1.3.	Normes et validation .....	13
1.4.	Votre avis nous intéresse!.....	16
1.5.	Liens Internet.....	17
1.6.	Fichiers de données .....	18
1.7.	Bibliographie.....	19
1.8.	Dons .....	23
1.9.	Médiagraphie.....	24
2.	Dont's .....	25
3.	Historique des nouveautés .....	27
4.	Divers (paramètres/configuration) .....	32
4.1.	Déterminer la version .....	32
4.2.	Changer la langue de l'interface .....	33
4.3.	Désactiver la création automatique de fichiers de Backup.....	34
4.4.	Inclure test d'Anderson-Darling dans la droite de Henry .....	35
4.5.	Largueur de la fenêtre de session .....	37
4.6.	Obtenir les commandes de lignes des actions .....	37
4.7.	Afficher de la date/heure, nom du fichier projet/feuille sur les graphes .....	39
4.8.	Symboles graphiques.....	41
4.9.	Enregistrer seulement les feuilles Minitab pour ouverture ultérieure dans R .....	42
4.10.	Afficher les grilles par défaut dans les graphiques.....	43
4.11.	Définir les métadonnées des fichiers et des feuilles .....	44
5.	Project Manager .....	48
5.1.	Ajouter/Renommer des feuilles.....	48
5.2.	Joindre des fichiers externes.....	51
5.3.	Rapport/Notes internes (ReportPad).....	55
6.	Saisie de données/Manipulations des données .....	56
6.1.	Afficher/Masquer des colonnes.....	57
6.2.	Rechercher/Remplacer .....	58
6.3.	Trier des données.....	59
6.4.	Scinder des données .....	62
6.5.	Coder des valeurs .....	65
7.	Manipulations élémentaires .....	69
7.1.	Exercice 1.: Calculs arithmétiques élémentaires .....	69
7.2.	Exercice 2.: Valeurs propres.....	73
7.3.	Exercice 3.: Valeurs propres.....	75
7.4.	Exercice 4.: Importer/Exporter des données *.csv .....	79
7.5.	Exercice 5.: Importer/Exporter des données *.accdB via ODBC .....	86
7.6.	Exercice 6.: Empiler/Déempiler des données .....	91
7.7.	Exercice 7.: Générer une variable aléatoire et faire un histogramme avec ou sans ajustement 99	
7.8.	Exercice 8.: Formatage conditionnel.....	106
8.	Graphiques .....	109
8.1.	Exercice 9.: Graphique à barres.....	109
8.2.	Exercice 10.: Graphique en secteurs .....	112
8.3.	Exercice 11.: Diagramme à points.....	115
8.4.	Exercice 12.: Récapitulatif graphique de statistiques .....	121

8.5.	Exercice 13.: Diagramme de valeurs individuelles .....	124
8.6.	Exercice 14.: Histogrammes simples ou multiples avec ajustements.....	128
8.7.	Exercice 15.: Graphique nuage de points avec droite de régression .....	134
8.8.	Exercice 16.: Générer un graphique CDF empirique .....	137
8.9.	Exercice 17.: Générer un graphique à bulles .....	141
8.10.	Exercice 18.: Graphique matriciel diagonal .....	144
8.11.	Exercice 19.: Graphique dans les marges.....	148
8.12.	Exercice 20.: Générer un graphique à points 3D.....	151
8.13.	Exercice 21.: Graphique boîte à moustaches (BoxPlot) .....	155
8.14.	Exercice 22.: Combiner des graphiques (lattice).....	160
8.15.	Exercice 23.: Diagramme quantile-quantile (Q-Q plot/qq-plot) .....	167
8.16.	Exercice 24.: Générer le graphique de la loi de densité de probabilité .....	171
8.17.	Exercice 25.: Analyse de Pareto .....	177
8.18.	Exercice 26.: Schéma branche et feuilles .....	179
8.19.	Exercice 27.: Diagramme de surface en 3D .....	184
8.20.	Exercice 28.: Graphique de contours.....	187
8.21.	Exercice 29.: Mode de balayage.....	191
9.	Probabilités .....	194
9.1.	Exercice 30.: Calculer une probabilité cumulée .....	194
9.2.	Exercice 31.: Visualiser la probabilité cumulée (unilatéral ou bilatéral) .....	197
9.3.	Exercice 32.: Convergence de la loi Binomiale en loi Normale.....	200
10.	Statistiques paramétriques.....	205
10.1.	Exercice 33.: Statistiques descriptives .....	205
10.2.	Exercice 34.: Kurtosis et Skewness.....	209
10.3.	Exercice 35.: Calculer la probabilité cumulée et la probabilité inverse .....	216
10.4.	Exercice 36.: Mettre des données à plat (table de fréquence) .....	220
10.5.	Exercice 37.: Mettre des données à plat (table de contingence).....	223
10.6.	Exercice 38.: Test d'indépendance du Khi-deux d'une table de contingence .....	226
10.7.	Exercice: Test du Khi-deux de la différence .....	233
10.8.	Exercice 39.: Test exact de Fisher .....	235
10.9.	Exercice 40.: V de Cramér (mesure d'association).....	239
10.10.	Exercice 41.: Test de Mantel-Haenszel-Cochran .....	243
10.11.	Exercice 42.: Test d'ajustement du Khi-deux .....	252
10.12.	Exercice 43.: Ajustement d'une loi de Poisson par le Khi-deux.....	256
10.13.	Exercice 44.: Test de Grubbs (test des valeurs aberrantes de Grubbs).....	260
10.14.	Exercice 45.: Test de Dixon (test des valeurs aberrantes de Dixon) .....	263
10.15.	Exercice 46.: Effectif de l'échantillon pour l'estimation pour loi Normale .....	266
10.16.	Exercice 47.: Puissance d'un test Z à 1 échantillon en bilatéral .....	270
10.17.	Exercice 48.: Taille d'échantillon (effectif) d'un test Z en bilatéral.....	275
10.18.	Exercice 49.: Différence (résolution) d'un test Z en bilatéral.....	278
10.19.	Exercice 50.: Intervalle de confiance de la moyenne (test Z à un 1 échantillon) .....	280
10.20.	Exercice 51.: Test Z de différence de la moyenne (test Z à un 1 échantillon) en bilatéral 282	
10.21.	Exercice 52.: Test Z de différence de la moyenne (test Z à un 1 échantillon) en unilatéral gauche 287	
10.22.	Exercice 53.: Puissance d'un test t à 1 échantillon en bilatéral.....	293
10.23.	Exercice 54.: Taille d'échantillon (effectif) d'un test t à 1 échantillon en bilatéral .....	296
10.24.	Exercice 55.: Intervalle de confiance de la moyenne (test t à un 1 échantillon) .....	300
10.25.	Exercice 56.: Test t-Student de différence de la moyenne bilatéral d'un échantillon.....	305
10.26.	Exercice 57.: Test t-Student homoscédastique bilatéral d'égalité de la moyenne .....	308
10.27.	Exercice 58.: Test t-Student pour données appariées .....	311
10.28.	Exercice 59.: Test t-Student hétéroscédastique bilatéral d'égalité de la moyenne (test t de Welch) 315	
10.29.	Exercice 60.: TOST (Two One-Sided Test) de Student .....	317

10.30.	Exercice 61.: Taille d'échantillon (effectif) d'un test p (proportion) à 1 échantillon en bilatéral	322
10.31.	Exercice 62.: Intervalle de confiance de la proportion	326
10.32.	Exercice 63.: Taille d'échantillon (effectif) d'un test p (proportion) à 2 échantillons en bilatéral	329
10.33.	Exercice 64.: Comparaison de proportions sur une même population (test binomial exact)	332
10.34.	Exercice 65.: Comparaison de proportions sur 2 échantillons indépendants	334
10.35.	Exercice 66.: Test de Poisson à un échantillon unilatéral/bilatéral	337
10.36.	Exercice 67.: Test de Poisson à deux échantillons bilatéral	343
10.37.	Exercice 68.: Intervalle de confiance de l'écart-type (test du khi-2 de la variance)	346
10.38.	Exercice 69.: Test de Fisher d'égalité des variances	349
10.39.	Exercice 70.: Test de Levene d'égalité de deux variances	355
10.40.	Exercice 71.: Test de Levene et Bartlett d'égalité des variances d'une ANOVA	360
10.41.	Exercice 72.: Test de Dunnett de comparaison des moyennes à une valeur de contrôle	363
10.42.	Exercice 73.: Test de normalité de Kolmogorov-Smirnov	367
10.43.	Exercice 74.: Test de normalité de Ryan-Joiner (ie Shapiro-Wilk)	371
10.44.	Exercice 75.: Test de normalité d'Anderson-Darling (ie Agostino-Stephens)	374
11.	Statistiques non paramétriques	382
11.1.	Exercice 76.: Test de Friedman	382
11.2.	Exercice 77.: Test du signe binomial (dixit: test de la médiane)	385
11.3.	Exercice 78.: Intervalle de confiance de la médiane (via test du signe)	388
11.4.	Exercice 79.: Test de Mood (test des médianes)	392
11.5.	Exercice 80.: Test de la somme des rangs de (Wilcoxon)Mann-Withney pour deux échantillons indépendants	400
11.6.	Exercice 81.: Test de la somme des rangs signés de Wilcoxon pour 1 échantillon	405
11.7.	Exercice 82.: Test de la somme des rangs signés de Wilcoxon pour 2 échantillons appariés	408
12.	ANOVA paramétriques ou non paramétriques (balancées!)	411
12.1.	Exercice 83.: Puissance du test de l'ANOVA (WP)	411
12.2.	Exercice 84.: Analyses graphiques des interactions et effets principaux	416
12.2.1.	Méthodes classiques directement liées aux options ANOVA de Minitab	416
12.2.2.	Méthode liées aux options Qualité de Minitab	428
12.3.	Exercice 85.: Analyse de la variance à un facteur fixe (ANOVA-1) désempilé	433
12.4.	Exercice 86.: Analyse de la variance à un facteur fixe (ANOVA-1) empilé	436
12.5.	Exercice 87.: Test de Tukey	438
12.6.	Exercice 88.: Analyse de la variance à deux facteurs fixes (ANOVA-2) sans répétitions	443
12.6.1.	Minitab 16 et antérieur	443
12.6.2.	Minitab 17	447
12.7.	Exercice 89.: Analyse de la variance carré latin	451
12.8.	Exercice 90.: Analyse de la variance à deux facteurs fixes (ANOVA-2) avec répétitions (et sans interactions)	454
12.9.	Exercice 91.: Analyse de la variance Imbriquée/Hiérarchisée (nested ANOVA)	458
12.10.	Exercice 92.: Split-plot ANOVA (ANOVA en parcelles)	463
12.11.	Exercice 93.: Test de Kruskal-Wallis (ANOVA non paramétrique)	468
12.12.	Exercice 94.: ANCOVA (Analyse de la Covariance)	471
12.13.	Exercice 95.: MANOVA (multivariate ANOVA)	476
13.	Régressions et interpolations	478
13.1.	Exercice 96.: Effectuer une régression linéaire simple (modèle linéaire Gaussien)	478
13.1.1.	Résidus, résidus normalisés, résidus studentisés	484
13.1.2.	Test de l'homoscédasticité des résidus	486
13.1.3.	Effets de levier des points (valeurs influentes)	491
13.2.	Exercice 97.: Régression linéaire simple forcée à l'origine	493

13.3.	Exercice 98.: Effectuer une régression linéaire (nouvelles options v16/v17/v18) .....	497
13.3.1.	Intervalles de confiance des coefficients.....	505
13.3.2.	Inadéquation de l'ajustement (lack-of-fit) et erreur pure.....	507
13.4.	Exercice 99.: Facteur d'Inflation de la Variance (VIF)/colinéarité .....	509
13.5.	Exercice 100.: Régression (linéaire) par moindres carrés partiels (régression linéaire PLS univariée: PLS1).....	515
13.6.	Exercice 101.: Régression polynomiale cubique ou quadratique.....	520
13.7.	Exercice 102.: Effectuer une régression pas à pas (ascendante et descendante) .....	526
13.8.	Exercice 103.: Régression pas à pas combinatoire (meilleur sous-ensemble) .....	533
13.9.	Exercice 104.: Étude de linéarité de l'instrumentation et biais de mesure .....	536
13.10.	Exercice 105.: Effectuer un calcul de la corrélation $R$ (de Pearson) simple.....	542
13.11.	Exercice 106.: Matrice de corrélation $R$ de Pearson.....	546
13.12.	Exercice 107.: Effectuer un calcul de la covariance.....	550
13.13.	Exercice 108.: Effectuer une régression linéaire multiple (modèle linéaire Gaussien)...	553
13.14.	Exercice 109.: Déterminer les corrélations partielles .....	556
13.15.	Exercice 110.: Effectuer une régression non linéaire .....	561
13.16.	Exercice 111.: Régression logistique (logit) à variable catégorielle (qualitative) binaire (dixit: scoring).....	568
13.17.	Exercice 112.: Odds Ratio et intervalle de confiance.....	572
13.18.	Exercice 113.: Courbe ROC pour régression logistique (logit) à variable catégorielle (qualitative) binaire .....	575
13.19.	Exercice 114.: Régression logistique multinomiale .....	581
13.20.	Exercice 115.: Régression orthogonale (régression de Deming) .....	585
13.21.	Exercice 116.: Régression de Poisson log-linéaire (GLM Poisson).....	589
13.22.	Exercice 117.: Effectuer une autocorrélation et une autocorrélation partielle .....	599
13.23.	Exercice 118.: Régression pondérée.....	604
14.	Cartes de contrôle (CC).....	613
14.1.	Exercice 119.: CC par attributs $p$ (proportion).....	613
14.2.	Exercice 120.: CC par attributs de type $np$ (comptage par proportion).....	616
14.3.	Exercice 121.: CC par attributs de type $C$ .....	619
14.4.	Exercice 122.: CC par attributs de type $U$ .....	622
14.5.	Exercice 123.: CC par mesures de Levey-Jennings .....	625
14.6.	Exercice 124.: CC par mesures $S$ ( $S$ barre- $S$ ).....	629
14.7.	Exercice 125.: CC par mesures $X$ barre- $S$ simple .....	632
14.8.	Exercice 126.: CC par mesures $X$ barre- $S$ barre combinées.....	636
14.9.	Exercice 127.: CC par mesures $X$ barre- $S$ barre combinées groupées par dates .....	639
14.10.	Exercice 128.: CC par mesures $X$ barre- $S_p$ avec variance groupées et facteur de correction 643	
14.11.	Exercice 129.: CC par mesures $R$ barre- $R$ barre simple.....	646
14.12.	Exercice 130.: CC par mesures $X$ barre- $R$ barre simple.....	649
14.13.	Exercice 131.: CC par mesures $X$ barre- $R$ barre combinées .....	653
14.14.	Exercice 132.: CC par mesures individuelles par l'étendue mobile .....	656
14.14.1.	Couper une carte de contrôle quelconque .....	659
14.15.	Exercice 133.: CC par mesures individuelles $X$ barre par l'étendue mobile .....	661
14.16.	Exercice 134.: CC par mesures individuelles $X$ barre-EM barre combinées .....	664
14.17.	Exercice 135.: CC individuelle Moyenne Mobile (MA) avec limites basées sur l'étendue mobile 667	
14.18.	Exercice 136.: CC CUSUM V-Masque avec échantillons .....	670
14.19.	Exercice 137.: CC EWMA avec échantillons .....	674
14.20.	Exercice 138.: CC EWMA individuelle.....	677
14.21.	Exercice 139.: CC combinée $X$ Barre- $R$ et $R$ Barre- $R$ avec deux phases (avant/après)..	680
14.22.	Exercice 140.: CC $T$ (aux fréquences) .....	684
14.23.	Exercice 141.: CC $G$ .....	687
14.24.	Exercice 142.: CC $P'$ de Laney .....	689
14.25.	Exercice 143.: CC $U'$ de Laney .....	695

14.26.	Exercice 144.: Carte de contrôle de Hotelling (T2).....	701
15.	Outils de la qualité.....	712
15.1.	Exercice 145.: Test de normalité et Analyse de capabilité (Normale) .....	712
15.2.	Exercice 146.: Analyse de capabilité de Poisson (événements rares) .....	717
15.3.	Exercice 147.: Analyse de capabilité de Poisson avant/après .....	721
15.4.	Exercice 148.: Transformations de Box-Cox .....	726
15.5.	Exercice 149.: Transformations de Johnson (Yeo-Johnson) .....	734
15.6.	Exercice 150.: Analyse de capabilité court terme et long temps et niveau de qualité.....	738
15.7.	Exercice 151.: Intervalle de confiance de la capabilité .....	742
15.8.	Exercice 152.: Analyse SixPack.....	745
15.9.	Exercice 153.: Étude de l'instrumentation de type I .....	750
15.10.	Exercice 154.: Étude R&R (reproductibilité/répétabilité) croisée pour données continues 757	
15.11.	Exercice 155.: Créer une feuille de travail pour l'analyse R&R.....	768
15.12.	Exercice 156.: Étude R&R (reproductibilité/répétabilité) emboîtée pour données continues 771	
	Ajout d'une limite LSL/USL .....	778
15.13.	Exercice 157.: Étude R&R (reproductibilité/répétabilité) étendue pour données continues 781	
	R&R emboîtée avec facteurs aléatoires.....	781
	R&R emboîtée avec facteurs fixes .....	785
15.14.	Exercice 158.: Cartes d'essais de l'instrumentation .....	788
15.15.	Exercice 159.: Étude R&R (reproductibilité/répétabilité) pour données par attributs ....	791
15.16.	Exercice 160.: Kappa de Cohen (analyse d'accord d'attribut) .....	800
15.17.	Exercice 161.: Plan d'échantillonnage par mesures.....	804
15.18.	Exercice 162.: Plan d'échantillonnage par attribut .....	807
	Plan de contrôle de tolérance 0 (C=0/"Accept on None").....	810
15.19.	Exercice 163.: Intervalle de tolérance paramétrique pour loi Normale.....	813
16.	Statistiques multivariées (data mining/machine learning/fouille de données).....	816
16.1.	Exercice 164.: Étude de la cohérence avec l'alpha de Cronbach.....	816
16.2.	Exercice 165.: Clustering dendrogramme .....	820
16.3.	Exercice 166.: K-means (k-moyennes en groupes) .....	824
16.4.	Exercice 167.: Analyse Linéaire Discriminante .....	830
16.5.	Exercice 168.: Analyse en composantes principales (ACP) paramétrique.....	834
16.6.	Exercice 169.: Analyse factorielle exploratoire (méthode ACP sans rotation!).....	845
16.7.	Exercice 170.: Analyse factorielle exploratoire (méthode ACP avec rotation VariMax!) 860	
16.8.	Exercice 171.: Analyse factorielle exploratoire (méthode ACP avec rotation OrthoMax!) 864	
17.	Plans d'expériences.....	867
17.1.	Exercice 172.: Construction de plans d'expériences.....	870
17.1.1.	Construction d'un plan factoriel complet .....	870
17.1.2.	Construction d'un plan factoriel fractionnaire à 2 niveaux .....	874
17.1.3.	Construction de plans factoriels optimaux D/A/G .....	888
17.1.4.	Construction d'un plan de Taguchi .....	894
17.1.5.	Création d'un plan de criblage définitif (Minitab >18) .....	902
17.2.	Exercice 173.: Comparaison d'un plan d'expérience identique en factoriel complet général et factoriel à 2 niveaux .....	905
17.3.	Exercice 174.: Analyse de la variance multifactorielle .....	914
17.4.	Exercice 175.: Test de Fisher de la courbure .....	926
17.5.	Exercice 176.: Analyse du modèle d'un plan factoriel complet général.....	935
17.5.1.	Analyse sans blocs.....	938
17.5.2.	Analyse avec blocs.....	945
17.5.3.	Analyse avec données normalisées.....	948

17.6.	Exercice 177.: Analyse graphique des effets principaux et interactions du modèle d'un plan factoriel complet avec répétitions .....	954
	Diagrammes factoriels (interactions et facteurs principaux) .....	955
	Effets et Diagrammes de Pareto des Effets (PEF).....	958
17.7.	Exercice 178.: Lenth's PSE pour un plan factoriel à 2 niveaux sans répétitions .....	964
17.8.	Exercice 179.: Modèle linéaire généralisé d'un plan d'expérience .....	970
17.9.	Exercice 180.: Analyse d'un plan factoriel fractionnaire avec choix des alias .....	977
17.9.1.	Choix automatique des alias (confusions) .....	977
17.9.2.	Choix manuel des alias (confusions) .....	980
17.9.3.	Analyse du plan d'expérience aliasé (avec confusions) .....	984
17.9.4.	Analyse du plan d'expérience désaliasé .....	989
17.10.	Exercice 181.: Plan factoriel avec points centraux .....	995
17.11.	Exercice 182.: Analyse du modèle d'un plan factoriel personnalisé avec covariable ....	999
17.12.	Exercice 183.: Analyse du modèle d'un plan de Plackett-Burman.....	1006
17.13.	Exercice 184.: Analyse graphique avancée de base de plans factoriels .....	1015
17.14.	Exercice 185.: Analyse de la variabilité (WP).....	1027
17.15.	Exercice 186.: Analyse conjointe .....	1046
17.16.	Exercice 187.: Analyse du bruit (plan de Taguchi) .....	1052
17.17.	Exercice 188.: Plan d'expérience de Taguchi .....	1064
17.18.	Exercice 189.: Plan en parcelles divisées (split-plot) avec facteurs difficiles à changer	1081
17.19.	Exercice 190.: Plan de criblage définitif (definitive screening design – DSD) (WP) ...	1108
17.20.	Exercice 191.: Plan de réponse de surface composite face centrée (modèle du second degré ou "modèle quadratique").....	1127
17.21.	Exercice 192.: Plan de mélange centré du simplexe (sans variable de processus).....	1159
17.22.	Exercice 193.: Plan de mélange centré du simplexe avec variables de processus.....	1170
18.	Fiabilité/Survie .....	1186
18.1.	Exercice 194.: Analyse de données censurées à droite selon modèle non paramétrique de Kaplan-Meier .....	1186
18.2.	Exercice 195.: Ajustement de données censurées à droite à une distribution de Weibull à deux paramètres.....	1193
18.3.	Exercice 196.: Prédiction de garantie basé sur le maximum de vraisemblance d'une loi exponentielle .....	1200
18.4.	Exercice 197.: Plan de démonstration modèle cumulatif binomial pour taille d'échantillon nécessaire avec temps de test contraint .....	1203
18.5.	Exercice 198.: Plan de démonstration modèle cumulatif binomial pour temps cumulé de test nécessaire avec taille d'échantillon connue.....	1206
18.6.	Exercice 199.: Plan de démonstration modèle exponentiel-chi temps cumulé de test nécessaire .....	1209
19.	Séries temporelles .....	1216
19.1.	Exercice 200.: Générer un graphique de série chronologique (temporelle) .....	1216
19.1.1.	Extraire des données d'un graphique à points quelconque .....	1218
19.2.	Exercice 201.: Générer un graphique d'analyse de tendance.....	1223
19.3.	Exercice 202.: Générer un graphique de moyennes mobiles (MM3).....	1226
19.4.	Exercice 203.: Générer un graphique de lissage exponentiel simple .....	1230
19.5.	Exercice 204.: Lissage exponentiel double selon Holt.....	1234
19.6.	Exercice 205.: Lissage exponentiel triple selon Winters.....	1239
19.7.	Exercice 206.: ARIMA(0,1,1) et ARIMA(0,2,2) .....	1242
19.8.	Exercice 207.: Test de Durbin-Watson .....	1248
19.9.	Exercice 208.: Test des suites (test de Wald-Wolfowitz).....	1252
20.	Macros/Scripting .....	1255
20.1.	Exercice 209.: Convertir des actions en macro .....	1255
20.2.	Exercice 210.: Intégration d'une macro téléchargée .....	1264
20.3.	Exercice 211.: Macro automation MS Office Access-Minitab-MS Office Word .....	1271
	Conclusion.....	1285

---

21. Index..... 1286

ÉCHANTILLON



# 1. Introduction

«Portions of information contained in this publication/book are printed with permission of Minitab Inc. All such material remains the exclusive property and copyright of Minitab Inc. All rights reserved»

This information may have been translated for your convenience from the original and official English language version, which can be found at [www.minitab.com](http://www.minitab.com), or as embedded in Minitab Statistical Software. Minitab retains all rights therein, and Minitab disclaims any and all responsibility for any reliance by you upon the translated version, which you use at your own risk. In the event of any discrepancy the English language version shall control.

MINITAB® and all other trademarks and logos for the Company's products and services are the exclusive property of Minitab Inc. All other marks referenced remain the property of their respective owners. See [minitab.com](http://minitab.com) for more information.

Minitab® Statistical Software est un logiciel d'exploration de données créé en 1972 à l'origine à l'Université de Pennsylvania par Barbara Ryan, Thomas Ryan Junior et Brian Joiner destiné à la base aux étudiants mais qui a été rapidement adopté pour l'analyse statistique de données dans les industries et sociétés de services principalement dans le cadre de l'application des techniques de la méthode DMAIC Six Sigma:

- |                        |  |
|------------------------|--|
| D/M (Définir/Mesurer)  | <ul style="list-style-type: none"><li>• Diagrammes d'Ishikawa</li><li>• Analyse de Pareto</li><li>• Études R&amp;R</li><li>• Cartes de contrôle</li><li>• Analyse de capacité</li><li>• Graphiques</li></ul>   |
| A (Analyser)           | <ul style="list-style-type: none"><li>• Études R&amp;R</li><li>• Tests d'hypothèses</li><li>• ANOVA</li><li>• Corrélation</li><li>• Régressions uni/multivariées linéaires ou non linéaires</li><li>• Analyse de capacité</li><li>• Graphiques</li><li>• Plans d'expérience</li><li>• Arbres de classification</li></ul> |
| I (Improve: Améliorer) | <ul style="list-style-type: none"><li>• ANOVA</li><li>• Régressions uni/multivariées linéaires ou non linéaires</li><li>• Analyse de séries temporelles</li><li>• Optimisation de réponse</li><li>• Analyse de capacité et de survie</li><li>• Tests d'hypothèses</li><li>• Études R&amp;R</li></ul>                     |
| C (Contrôler)          | <ul style="list-style-type: none"><li>• Cartes de contrôles</li><li>• Tests d'hypothèses</li><li>• Fiabilité</li></ul>   |

Il faut savoir qu'à la base Minitab s'appelait Omnitab et avait été développé par le NIST (National Institute of Standards and Technology).

Il implémente une série de méthodes de fouille de données issues du domaine de la statistique exploratoire, de l'analyse de données, de l'inférence statistique, des plans d'expérience pour les besoins de métiers suivants:

- Master Black Belt Six Sigma
- Ingénieur procédés/processus
- Business Analyst
- Ingénieurs qualités (qualiticiens)
- Ingénieurs en gestion de projets
- Dirigeants de départements
- Doctorants

Le choix de Minitab® Statistical Software dans une organisation peut se justifier principalement selon moi (Vincent ISOZ) par le fait que:

- Il est devenu la référence dès les débuts du déploiement du Six Sigma dans le monde.
- La quantité de sites internet et d'ouvrages qui lui sont consacrés permettent de ne pas faire appel à tout bout de champ à des consultants ou à un support onéreux (bien que le support Minitab soit gratuit à ma connaissance).
- Il est la référence pour les tops manager et hauts potentiels dans les organisations (associé à @Risk, MS Excel et autres...).
- Le support après-vente et la disponibilité de ce même support est d'un niveau exceptionnel en comparaison à R, SAS ou SPSS (qui sont difficilement joignables si l'on n'est pas employé d'une grande structure).

Enfin rappelons que ce logiciel se destine à des ingénieurs et universitaires donc son rôle n'est pas de faire de l'esthétique mais de l'efficace!

Le lecteur remarquera que le début du présent document rassemble uniquement des points qui sont aussi faisables avec MS Excel avec plus ou moins de facilité (parfois plus rapidement, parfois pas...). Le but est d'abord de montrer les différences entre les deux logiciels comme le font de nombreux ouvrages avant de passer à des cas très spécifiques à Minitab® Statistical Software.

*Un ordinateur est un outil incomparable entre les mains de celui qui sait. Sous les doigts du Créatin, c'est un revolver manip par un aveugle au milieu de la foule* Chester Himes

*Remarques:*

R1. Les exercices sont quasiment les mêmes que ceux effectués dans les cours de statistique théorique, maîtrise statistique des procédés, d'analyse décisionnelle avec MS Excel ou lors du cursus de gestion de projets (dans lesquels les calculs sont faits à la main).

R2. Je ne présente ici normalement que des outils utilisant des concepts mathématiques dont j'expose la démonstration mathématique détaillée (démonstrations disponibles dans mon livre sur les Mathématiques Appliquées) dans mes cours et qui sont utilisées dans le cadre de mes activités de consultant. Si cela vient à ne pas être le cas, le titre du sujet est suivi de l'abréviation *WP* qui est l'abréviation de: Without Proof.

R3. J'ai rédigé ce document uniquement pour le fun afin de valider les résultats obtenus à la main et surtout pour jouer avec les théorèmes mathématiques que j'ai étudiés en détail dans le cadre de mes études (démonstrations mathématiques disponibles dans mon ouvrage sur les Mathématiques Appliquées).

R4. Suite à des questions de la part de lecteurs: Non je ne suis pas rétribué par Minitab® Statistical Software pour leur faire de la pub... Des logiciels comme:

XLStat, SPSS, R (voir mon PDF sur ce logiciel), SAS, PSPP, Gauss, MATLAB, Statistica, Stata, Medcalc, StatsDirect, SigmaXL, NumXL, JMP, Weibull++, Design-Expert, PlanExpert, UNISTAT, QIMacros (Quality Improvement Macros), EasyFitXL, NemrodW, qs-STAT, NeoLiCy, etc.

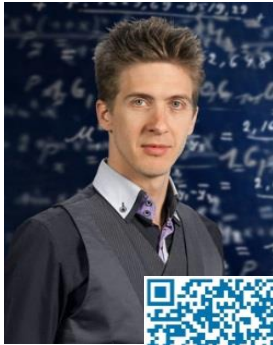
font à peu près pareil relativement aux sujets couverts dans ce document ou juste sur certains sujets particuliers. Il fallait juste que je fasse un choix... (je ne peux pas passer mon temps à écrire des supports sur tous les types de logiciels!) et celui-ci s'est porté sur le logiciel utilisé par la majorité de mes clients. Cependant si des passionnés veulent reproduire le contenu du présent livre avec leur logiciel de statistique favori qu'ils n'hésitent pas! Ce serait même fort intéressant de comparer les résultats!! Si j'ai le temps j'écrirai le même contenu mais avec SPSS.

R5. Oui de nombreux logiciels sont beaucoup plus puissant que Minitab® Statistical Software à ce jour. Mais ces derniers sont pour les statisticiens seniors ayant des doctorats alors que Minitab® Statistical Software s'adresse plutôt à des personnes possédant un niveau Bachelor à Master et ayant très peu de temps pour apprendre un nouveau logiciel.

Pour terminer, nous tenons à remercier ici les quelques collègues et clients qui ont bien voulu nous faire part de leurs remarques pour améliorer le contenu de ce livre électronique. Il est cependant certain qu'il est encore perfectible sur de nombreux points et qu'il va encore évoluer puisqu'il y a encore une petite dizaine de sujets dont les démonstrations mathématiques sont en cours de rédaction pour la prochaine édition du livre de Vincent ISOZ sur les Mathématiques Appliquées et nous montrerons alors comment les utiliser aussi avec Minitab® Statistical Software.

Si vous souhaitez être informé des nouvelles versions **majeures** de ce document n'hésitez pas à m'écrire un mail dans ce sens: [isoz@sciences.ch](mailto:isoz@sciences.ch).

## 1.1. À propos du rédacteur



Domicilié à ce jour à Lausanne (Suisse)

Formation: Ingénieur Physicien HES  
Équivalence Internationale: Bachelor of Science  
Fonction: Consultat scientifique senior

Année de naissance: 1978

Courriel: [isoz@sciences.ch](mailto:isoz@sciences.ch)

Actuellement, je travaille en tant que consultant spécialiste dans le conseil et la formation et techniques quantitatives pour les hauts-potentiels. Mes tâches consistent à faire du conseil (français, anglais et allemand) en mathématiques appliquées dans le tutorat d'analystes quantitatifs (niveau Bac+5 à Bac+7). Accessoirement, je suis le rédacteur de plusieurs livres électroniques dans les domaines suivants:

- maîtrise statistique des processus/procédés (méthodes paramétriques et non paramétriques)
- modélisation prévisionnelle/décisionnelle avancée (arbres de décisions, chaînes de Markov)
- recherche opérationnelle (simplexe, algorithmes génétiques, algorithme GRG)
- data mining (réseaux de neurones, ACP, AFC, régressions, scoring, clustering, etc.)
- modélisation du risque en gestion de projets et finance d'entreprise (monte-carlo, etc.)
- gestion de projets (modèles et best practices théoriques EFQM+Six Sigma, MS Project)
- ISO 9001:2008, 5807:1985, 10015:1999, 31000+31010:2009, 8258:1991, 10017:2003, etc.
- Adobe Photoshop et Illustrator
- 12 applications de la suite Microsoft Office System (Project, Visio, SharePoint, Access, etc.)

À ce jour interventions dans plus de ~1'500 entreprises dont 10 du *Fortune 500* selon listing 2009 et 3 universités et écoles d'ingénieurs suisses dans des cours de modélisation de bases de données et simulations stochastiques du risque. Formation de plusieurs dirigeants de multinationales en one to one.

Accessoirement j'interviens pour des formations sur des logiciels comme MS Project, MS Visio, MS Access et une vingtaine d'autres dont je délègue l'organisation à des entreprises spécialisées dans la formation continue en bureautique (niveau licence universitaire et en-dessous).

Enfin, je conseille aussi vivement à toute personne souhaitant vraiment maîtriser le sujet de lire mon e-book sur les *Mathématiques Appliquées* (~5'900 pages) ou de préférence la version anglophone (>7'000 pages).

## 1.2. Avertissements

Le contenu du présent support est élaboré par un processus de développement par lequel des experts en statistiques parviennent à un consensus. Ce processus qui rassemble des participants bénévoles recherche également les points de vue de personnes intéressées par le sujet de cet ouvrage. En tant que responsable du présent support, j'assure l'administration du processus et je fixe les règles qui permettent de promouvoir l'équité dans l'approche d'un consensus. Je me charge également de rédiger les textes, parfois de les tester/évaluer ou de vérifier indépendamment l'exactitude/solidité ou l'exhaustivité des informations présentées.

Je décline toute responsabilité en cas de dommages corporels, matériels ou autres de quelque nature que ce soit, particuliers, indirects, accessoires ou compensatoires, résultant de la publication, de l'application ou de la confiance accordée au contenu du présent support. Je n'émet aucune garantie expresse ou implicite quant à l'exactitude ou à l'exhaustivité de toute information publiée dans le présent support, et ne garantit aucunement que les informations contenues dans cet ouvrage satisfassent un quelconque objectif ou besoin spécifique du lecteur. Je ne garantis pas non plus les performances de produits ou de services d'un fabricant ou d'un vendeur par la seule vertu du contenu du présent support.

En publiant des textes, il n'est pas dans l'intention principale du présent support de fournir des services de spécialistes ou autres au nom de toute personne physique ou morale ni pour mon compte, ni d'effectuer toute tâche devant être accomplie par toute personne physique ou morale au bénéfice d'un tiers. Toute personne utilisant le présent support devrait s'appuyer sur son propre jugement indépendant ou, lorsque cela s'avère approprié, faire appel aux conseils d'un spécialiste compétent afin de déterminer comment exercer une prudence raisonnable en toute circonstance. Les informations et les normes concernant le sujet couvert par le présent support peuvent être disponibles auprès d'autres sources que le lecteur pourra souhaiter consulter en quête de points de vue ou d'informations supplémentaires qui ne seraient pas couverts par le contenu du présent site Internet.

Je ne dispose (malheureusement...) d'aucun pouvoir dans le but de faire respecter la conformité au contenu du présent ouvrage, et je ne m'engage nullement à surveiller ni à faire respecter une telle conformité. Je n'exerce (à ce jour...) aucune activité de certification, de test ni d'inspection de produits, de conceptions ou d'installations à fins de santé ou de sécurité des personnes et des biens. Toute certification ou autre déclaration de conformité en matière d'informations ayant trait à la santé ou à la sécurité des personnes et des biens, mentionnée dans le présent support, ne peut aucunement être attribuée au contenu du présent support et demeure sous l'unique responsabilité de l'organisme de certification ou du déclarant concerné.

Et pour rappel:

- Tous les modèles sont faux
- Certains modèles sont meilleurs que d'autres
- Le modèle le meilleur ne peut jamais être connu avec certitude!
- Plus simple est le modèle, mieux c'est!

### 1.3. Normes et validation

Rappelons conformément à ce que nous avons vu dans le cours théorique qu'il est indispensable pour le chercheur/statisticien/ingénieur professionnel de se baser sur les normes suivantes (dans l'ordre des plus utilisées) pour son travail et tous les outils dont le présent support fait l'objet:

- **ISO 8000**  
Qualité des données
- **ISO 9241-11**  
Mon logiciel est-il utilisable?
- **ISO 31:2006**  
Système international d'unités
- **ISO 3534-1:1999**  
Vocabulaire et symbole des statistiques
- **ISO 2602:1980**  
Interprétation statistique de résultats d'essais - Estimation de la moyenne - Intervalle de confiance
- **ISO 3301:1975**  
Interprétation statistique des données - Comparaison de deux moyennes dans le cas d'observations appariées
- **ISO 5479:1997**  
Interprétation statistique des données - Tests pour les écarts à la distribution normale
- **ISO 3494:1976**  
Interprétation statistique des données -- Efficacité des tests portant sur des moyennes et des variances
- **ISO 11453:1996**  
Interprétation statistique des données - Tests et intervalles de confiance portant sur les proportions
- **ISO 16269-4:2010**  
Interprétation statistique des données Détection et traitement des valeurs aberrantes
- **ISO 16269-6:2005**  
Interprétation statistique des données - Détermination des intervalles statistiques de tolérance
- **ISO 16269-8:2004**  
Interprétation statistique des données - Détermination des intervalles de prédiction
- **ISO/TR 18532:2009**  
Lignes directrices pour l'application des méthodes statistiques à la qualité et à la normalisation industrielle
- **ISO 3534-3:1999**  
Plans d'expérience (ou AFNOR NF X 06-080 + NF X 06-081)
- **ISO 8285:1991**  
Cartes de contrôle de Shewhart

- **ISO 17025:2005**  
Exigences générales concernant la compétence des laboratoires d'étalonnages et d'essais
- **ISO 10017:2003**  
Lignes directrices pour les techniques statistiques relatives à l'ISO 9001:2000
- **ISO 13300:2006**  
Guide général à l'attention du personnel des laboratoires d'analyse sensorielle
- **ISO 12888:2011**  
Illustrations choisies d'études de répétabilité et de reproductibilité par calibre
- **ISO 31010:2009**  
Techniques d'évaluations des risques
- **ISO 3951:2006**  
Règles d'échantillonnage pour les contrôles par mesures
- **ISO 11095:1996**  
Étalonnage linéaire utilisant des matériaux de référence
- **ISO 22514-2:2013**  
Indices de capabilité
- **ISO 5725**  
Précision et fiabilité des mesures en laboratoire (on y retrouve le test C de Cochran et aussi celui de Dixon)

Bref en gros en 2014 les normes relatives à l'analyse statistique peuvent se résumer avec la cartographie disponible sur le lien suivant:

<https://fr.scribd.com/doc/263063855/AFNOR-Cartographie-Normes-Statistiques>

**À propos de la validation de l'utilisation du logiciel R par la FDA (Food and Drug administration)** il est peut-être utile d'informer le lecteur que ceci n'est pas le rôle de la FDA mais de l'acheteur de s'assurer du bon fonctionnement de la solution informatique choisie (conformément à la norme **ISO 9241-11** vue dans le cours de gestion de projets pour tout logiciel ou machine). Par ailleurs, signalons l'article du code des régulations Titre 21 partie 820.70(i):

*“When computers or automated data processing systems are used as part of production or the quality system, the manufacturer shall validate computer software for its intended use according to an established protocol. All software changes shall be validated before approval and issuance. These validation activities and results shall be documented.”*

et aussi (Final Guidance for Industry and FDA Staff, January 11, 2002):

*“The device manufacturer is responsible for ensuring that the product development methodologies used by the off-the-shelf (OTS) software developer are appropriate and sufficient for the device manufacturer's intended use of that OTS software. For OTS software and equipment, the device manufacturer may or may not have access to the vendor's software validation documentation. If the vendor can provide information about their system requirements, software*

*requirements, validation process, and the results of their validation, the medical device manufacturer can use that information as a beginning point for their required validation documentation.”*

ÉCHANTILLON



## 1.4. Votre avis nous intéresse!

En tant que lecteur de ce document, vous êtes le critique et le commentateur le plus important. Votre opinion compte et il est très intéressant de savoir ce qui est bien, ce qui peut être mieux et les sujets que vous souhaiteriez voir être traités.

Vous pouvez m'envoyer un e-mail pour partager ce que vous avez aimé ou détesté dans le présent document afin d'en assurer une amélioration continue.

Si vous souhaitez compléter le présent support avec un sujet qui vous tient à coeur et pour lequel vous avez la démonstration mathématique n'hésitez pas à me contacter. J'intégrerai le sujet en précisant votre nom et prénom.

*Notez que malheureusement, je ne peux pas répondre gratuitement à des questions techniques d'ingénierie ou de problématique d'entreprise par e-mail pour des raisons professionnelles évidentes.*

E-mail: [isoz@sciences.ch](mailto:isoz@sciences.ch)

## 1.5. Liens Internet

<http://www.minitab.com>

<http://www.minitab.com/en-TW/support/macros/>

<http://www.itl.nist.gov/div898/handbook/>

Forums:

<http://www.talkstats.com>

<http://statistiques.forumpro.fr>

<http://www.les-mathematiques.net>

## 1.6. Fichiers de données

Les fichiers de données utilisés dans ce livre, c'est-à-dire:

- 119 fichiers (\*.mpj, \*.csv, \*.accdb, \*.mac, etc)
- Pour un total 4.61 MB de de données

sont mis à disposition, avec mises à jour gratuites perpétuelles, uniquement aux apprenants ayant suivi nos formations correspondantes à ce livre ou ayant payé (ou fait un don) d'une somme équivalente de **499\$**.

Les universités, écoles d'ingénieurs et les sociétés privées doivent acheter une seule fois le package de fichiers pour une utilisation illimitée de leurs étudiants ou employés, mais limité à une seule localisation géographique.

Pour obtenir plus d'informations sur le processus d'achat, vous pouvez contacter envoyer un courriel à l'adresse suivante: [isoz@sciences.ch](mailto:isoz@sciences.ch)

## 1.7. Bibliographie

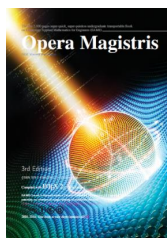
Voici la liste de livres d'une qualité pédagogique et de rigueur extraordinaires que j'ai eu la chance d'avoir entre les mains et dont je recommande l'acquisition. J'en ai lu beaucoup d'autres mais qui sont tellement mauvais qu'ils ne valent pas la peine d'être mentionnés:

Le lecteur aura donc compris que je recommande très fortement de compléter la lecture du présent e-book (non exhaustif sur le domaine de la gestion de projets) par la liste de lecture ci-dessous.



*Éléments de mathématiques appliquées* / 4'888 pages / Éditions Sciences.ch / Vincent ISOZ / 3<sup>ème</sup> édition  
ISBN: 978283999327

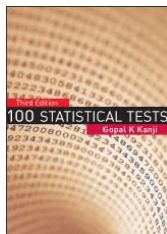
Commentaire: Livre rédigé par les soins de votre serviteur... Il contient les démonstrations mathématiques détaillées de tous les outils présentés dans ce présent support et pas que...



*Opera Magistris* / 5'576 pages / Editions Sciences.ch / Vincent ISOZ / 3<sup>rd</sup> édition  
ISBN: 978283999327

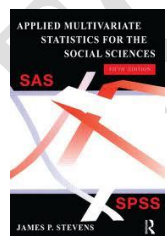
Commentaire: Équivalent du livre en français ci-dessus mais en anglais, bien plus complet et de qualité graphique considérablement meilleure!

**Attention!!! Pendant les séminaires et conférences de moins ou égal à 3 jours je prends oralement et délibérément des raccourcis théoriques dangereux voire parfois partiellement faux délibérément à cause du manque de temps pour aller dans les détails.**



*100 Statistical tests* / 257 pages / Éditions SAGE / Gopal K. KANJI  
ISBN: 101412923751

Commentaire: Un "must have" listant 100 tests statistiques avec les hypothèses d'utilisation, les limites, les formules (sans démonstration) et un exemple à chaque fois.



*Applied Multivariate Statistics for the social sciences* / 663 pages / Éditions Routledge / James P. STEVENS  
ISBN: 9780805859010

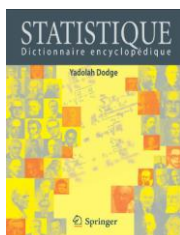
Commentaire: Livre un peu vieillissant mais contient des exemples très détaillés avec étapes de calcul de test multivariés complexes, ce qui est difficile à trouver dans la littérature. Par contre, il n'y a aucune démonstration détaillée.



*Business Statistics* / 892 pages / Éditions Routledge / Ronald M. WEIERS / 7ème édition

ISBN: 053845217X

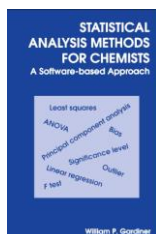
Commentaire: Le livre !!! à mettre entre les mains à ceux qui commencent à étudier les statistiques et qui cherchent un ouvrage sans démonstrations mathématiques mais présentant de façon très claire et pédagogique des tests jusqu'au niveau du Master et accompagnés à chaque fois de façon détaillée de Microsoft Excel et Minitab.



*STATISTIQUES: Dictionnaire encyclopédique* / 614 pages / Éditions Springer / Yadolah DODGE

ISBN: 9782287720932

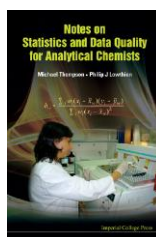
Commentaire: Un petit dictionnaire très utile contenant un peu tout de façon très condensée sur tout le domaine des statistiques. Très pratique pour retrouver des définitions qui génèrent des débats en entreprise.



*Statistical analysis methods for chemists* / 390 pages / Éditions The Royal Society of Chemistry / William P. GARDNIER

ISBN: 085404549X

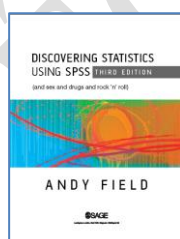
Commentaire: Un livre vieillissant sans démonstrations mathématiques mais qui intéressera les chimistes utilisant Minitab et qui veulent apprendre une petite partie du logiciel avec des exemples familiers à leur domaine d'activité.



*Statistical analysis methods for chemis* / 260 pages / Éditions Imperial College Presse / Michael THOMPSON, Philip J. LOWTHIAN

ISBN: 1848166168

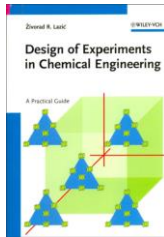
Commentaire: Uniquement théorique, mais sans démonstrations mathématiques ni exemples ou exercices il constitue un bon complément à l'ouvrage proposé juste précédemment qui était aussi destiné aux chimistes.



*Discovering Statistics using SPSS* / 854 pages / Éditions Sage / Andy Field / 3ème édition

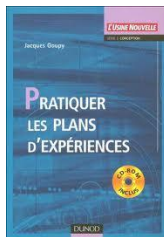
ISBN: 9781847879066

Commentaire: Comme je le dis toujours au début des cours: dans un contexte scientifique avoir seulement un logiciel d'analyse est insuffisant ne serait-ce que pour se couvrir contre les bugs et les variantes d'algorithmes. Dès lors, un deuxième outil complémentaire et bien supérieur à Minitab pour la partie statistique est SPSS! L'auteur de ce livre est remarquable! Outre le fait qu'il ait aussi écrit un ouvrage de la même qualité sur R est a beaucoup d'humour, un sens aigu de la pédagogie et sait particulièrement bien présenter des outils informatiques avec des exemples provenant de données réelles! Un must have absolu!



*Design of Experiments in Chemical Engineering* / 523 pages / Éditions Wiley / Zivorad R. Lazic  
ISBN: 3-527-31142-4

Commentaire: Au niveau des démonstrations mathématiques, c'est... pourri comme pour le 99% des livres sur le sujet mais sinon il y d'excellent exemples et certaines explications vulgarisées qui valent le détour! Cela reste une lecture à faire!



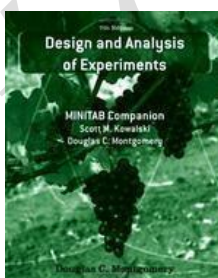
*Pratiquer les plans d'expériences* / 551 pages / Éditions Dunod / Jacques Goupy  
ISBN: 2100042173

Commentaire: Comme je le dis toujours au début des cours: dans un contexte scientifique avoir seulement un logiciel d'analyse est insuffisant ne serait-ce que pour ce couvrir contre les bugs et les variantes d'algorithmes. Dès lors, un deuxième outil complémentaire à Minitab pour les plans d'expérience est JMP! Et ce livre est d'un niveau de qualité remarquable tant par le niveau de détail des exemples industriels qui y sont donnés étape par étape que par la clarté des explications et de la présence parfois de quelques démonstrations mathématiques accompagnées de nombreuses figures. De plus, l'auteur est connu dans tout le monde francophone est un consultant dont la renommée n'est plus à faire.



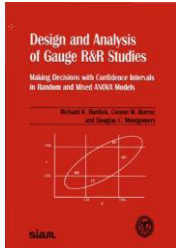
*A primer on experiments with mixture* / 368 pages / Éditions Wiley / John Cornell  
ISBN: 0470643382

Commentaire: Il est très difficile de trouver de la littérature détaillée et rigoureuse sur le sujet des plans de mélange ("mixture design" en anglais) mais John Cornell réussit ici un coup de maître. Son ouvrage est complet, rigoureux, détaillé, fourni d'exemple pertinents! On regrettera peut-être juste le fait qu'il ne fasse pas le parallèle avec un logiciel informatique de son choix mais bon... Achat recommandé!



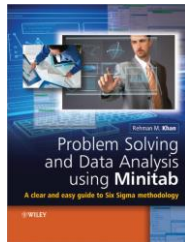
*Design and Analysis of Experiments* / 128 pages / Éditions Wiley / Douglas C. Montgomery, Scott M. Kowalski (Author)  
ISBN: 0470169907

Commentaire: Une sorte de condensé version papier de l'aide intégrée à Minitab mais avec des exemples pertinents sortant du cadre standard de Minitab. Pas de démonstration mathématiques, pas d'interprétation fouille des résultats et du contrôle des hypothèses mais intéressant à avoir lu au moins une fois.



*Design and Analysis of Gauge R&R Studies* / 220 pages / Éditions SIAM (Society for Industrial and Applied Mathematics) / Richard K. Burdick, Connie M. Borror and Douglas C. Montgomery (Author)  
ISBN: 0898715881

Commentaire: Un livre très très complet sur le sujet et qui est probablement ce qui se fait de mieux dans le domaine et qui va plus loin que ce que proposent la totalité des logiciels existants à ce jour et qui nous sont connus. Par contre ne vous attendez pas à avoir les démonstrations mathématiques...

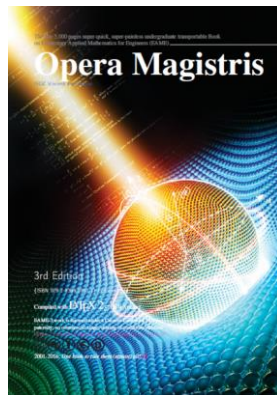


*Problem Solving and Data Analysis using Minitab, A clear and easy guide to Six Sigma Methodology* / 480 pages / Éditions Wiley / Rehman M. Khan  
ISBN: 9781118307571

Commentaire: Un vraiment très bon petit livre qui couvre les aspects les plus importants de Minitab de façon claire et concise. Ceux qui ne sont pas intéressés par les formules ou l'aspect mathématique apprécieront très probablement ce petit bijou de pédagogie!

## 1.8. Dons

Ce livre fait partie des livres compagnons du compendium de mathématiques théoriques Opera Magistris d'environ 6'900 pages qui est lui aussi gratuit.



Afin de poursuivre la rédaction de ce compendium et continuer à faire évoluer ce livre gratuit sur Minitab vos dons sont la bienvenue et nous vous en remercions d'avance.

Vous pouvez faire un don via Tipeee:

<https://www.tipeee.com/elements-of-applied-mathematics>

**tipeee**  
**.com**



## 1.9. Médiagraphie

Ceux qui souhaitent une approche plus "parlante" de Minitab peuvent consulter en ligne les 2 micro-formations d'introduction à Minitab (environ 6 heures à ce jour) faites avec mon partenaire [Video2Brain](#) (cliquez sur les images pour aller à la page web correspondante):



## 2. Dont's

Questions fréquentes que Minitab® Statistical Software R15 ne répond pas à ce jour simplement (sans macros ou prétraitement de l'information):

1. Minitab est actuellement (version 15) limité à 10'000'000 de lignes
2. Il n'y a actuellement aucune manière simple autre que de supprimer les lignes pour éliminer les outliers (valeurs aberrantes) lors d'une analyse...
3. Depuis la version 17 il n'est plus possible d'effacer simplement le contenu de la fenêtre de session
4. La régression exponentielle n'est pas disponible dans les versions 15 et antérieures (elle l'est maintenant dans la version 16 indirectement avec la méthode de Gauss-Newton) et de façon globale il manque les modèles de régression linéaire généralisés (~~Poisson~~, Gamma, Binomial, NB2, NBC, etc.)
5. ~~Pas de génération de plans d'expérience D-optimaux, A-optimaux ou E-optimaux~~ (cela est possible maintenant depuis la version 16).
6. ~~Pas de génération de plans d'expériences DSD~~ (Definitive Screening Design) (cela est possible depuis la version 18)
7. Ne permet pas d'analyser des plans d'expériences avec des facteurs ayant des effets sigmoïdes
8. Pas de solveur intégré pour chercher l'optimum de certaines fonctions non linéaires
9. Les tests Z à deux échantillons (ou le calcul de la puissance correspondante), Grubbs, Dixon et McNemar sont toujours disponibles que sous forme de macro à ce jour (par ailleurs il manque beaucoup de tests à Minitab)
10. L'analyse de série temporelle ne traite pas à ce jour la famille ARCH et GARCH.
11. Pas d'intervalle de confiance pour le Risque Relatif.
12. Pas de Régression de Passin & Bablok (alternative intéressante à la régression orthogonale de Deming)
13. Pas de ROC ou Lift Curve
14. Le double échantillonnage n'est disponible que sous forme de macro.
15. Le logiciel ne fait pas de simulations de Monte-Carlo dans le sens où on l'entend dans MS Excel.
16. Pas de régression linéaire de Cox disponible à ce jour (modèle de hasard proportionnel).
17. Il n'est pas possible à ce jour de supprimer les outliers d'une régression en même temps que la régression est effectuée!

18. Impossibilité d'afficher la table complète des tableaux de Taguchi ni de leur table d'interactions (il faut aller sur le site web ou avoir un ouvrage de référence!)
19. On ne peut pas faire des régressions pas-à-pas avec les transformations de Box-Cox activées en même temps

Pour voir d'autres choses que Minitab ne fait pas... voir mon livre sur le logiciel Open Source qui se nomme R!

ÉCHANTILLON

### 3. Historique des nouveautés

La liste originale en anglais vient d'ici:

<https://www.additive-net.de/en/software/produkte/minitab/minitab/versionshistorie>

#### Minitab 16

- Gestion du multilingue (choix de la langue dans les options du logiciel)
- Nouvel assistant:
  - Laissez le menu Assistant de Minitab vous guider tout au long de votre analyse.
  - Suivez un arbre de décision interactif pour déterminer la bonne approche.
  - Sélectionnez facilement l'outil approprié et passez en revue les directives pour vous assurer que votre analyse est réussie.
  - Gagnez du temps dans la réalisation de votre analyse grâce à une interface simplifiée.
  - Utilisez les interprétations générées par Minitab pour comprendre et répondre à vos résultats.
  - Présentez vos conclusions à l'aide de rapports détaillés et complets
- Test d'hypothèse pour la variance ou l'écart type
- Régression Générale
  - Spécification facile des termes d'interaction et polynomiaux.
  - Inclusion à la fois des variables prédictives continues et catégorielles.
  - Calcul des intervalles de confiance pour les coefficients de régression.
  - Transformation des variables de réponse à l'aide de la transformation de Box-Cox.
- Ajustement d'un modèle non linéaire à l'aide d'un catalogue de fonctions d'attente ou spécifiez votre propre fonction
- Disponibilité de la régression orthogonale
- Régression logistique binaire
  - Calcule des intervalles de confiance pour les probabilités d'événement.
  - Enregistrement des probabilités d'événements dans la feuille de calcul pour toutes les observations ou pour des observations uniques.
- Régression logistique ordinale et nominale
  - Probabilités d'événement stockées dans la feuille de calcul pour toutes les observations.
- Régression de Poisson pour les variables  $Y$  qui décrivent les décomptes
- Modèles non-linéaires généralisés
  - Interface simplifiée pour spécifier les termes du modèle
  - Interface simplifiée pour le calcul des valeurs prédites
  - Sélection automatique du modèle pour identifier les variables importantes
  - Les nouveaux graphiques incluent des tracés de contour et de surface
  - Transformation de Box-Cox
  - Optimiseur de réponse pour trouver les paramètres optimaux
  - Option de résultats pour afficher l'ensemble complet de coefficients pour tous les niveaux de facteur
  - Méthode de Fisher et sortie graphique pour les comparaisons multiples
- Moindres carrés partiels

- Spécification facilitée des termes d'interaction et polynomiaux.
- Inclusion à la fois des variables prédictives continues et catégorielles.
- Calcule des valeurs prédites pour les nouvelles observations des modèles linéaires généralisés
- ANOVA à une voie
  - Boîte de dialogue unique pour les données empilées ou déempilées
  - Méthode de Welch pour les variances inégales
  - Graphique fourni pour plusieurs comparaisons
- Plans d'expériences
  - Créez et analysez des conceptions avec des facteurs difficiles à modifier (split-plot designs)
  - Affichage facilité pour les noms de facteurs plus longs dans la boîte de dialogue *Termes*.
  - Affichage de la somme des carrés pour les termes individuels.
- Carte de contrôles
  - Création de diagrammes P' et U' pour surveiller les défectueux et les défauts avec de grandes tailles de sous-groupes.
  - Surveillez les événements rares avec les graphiques G et T.
  - Créez des graphiques individuels et I-MR à l'aide de l'estimation de Nelson pour l'écart type.
- Analyse de capabilité
  - PPM inclus sur le Sixpack™
- Analyse de la stabilité et de la durée de conservation de produits
- Possibilité d'afficher les Benchmark Z (niveau sigma) sur le Sixpack™
- Calcul des intervalles de tolérance pour les distributions normales et non normales
- Analyses Gages/R&R
  - Créer un diagramme Gage (Gage chart) avec des données manquantes (études déséquilibrées)
  - Calcul des intervalles de confiance pour les analyses Gage croisé et imbriqué R&R
- Analyse d'accord d'attribut
  - Affichage d'un tableau récapitulatif des désaccords lorsque des valeurs standard connues sont fournies.
- Analyse en composantes principales
  - Affichage des distances de Mahalanobis à l'aide d'un graphique des valeurs aberrantes.
  - Enregistrement des distances de Mahalanobis dans la feuille de travail.
- Puissance et tailles d'échantillons
  - Taille de l'échantillon pour l'estimation
  - Test de Student apparié
  - Taux de Poisson à 1 et 2 échantillons
  - 1 et 2 variances
  - Conception factorielle complète générale
- Nouveaux tutoriaux dans l'aide
- Export de la sortie directement vers Microsoft PowerPoint ou Word.

## Minitab 17

- Assistant
  - Améliorations apportées à l'interface et à la sortie

- Fiches de rapport et rapports de diagnostic pour l'analyse graphique
- Régression multiple plus optimisation des réponses
- DOE pour les plans de criblage et de modélisation
- Nouveaux menus pour les cartes de contrôle avant/après et l'analyse des capacités
- Cartes de contrôle et analyse de capacité pour des tailles de sous-groupes égales ou inégales
- Les cartes de contrôle et l'analyse des capacités incluent la vérification des cartes Laney P' et U'
- Tests d'hypothèse à 2 échantillons pour les données empilées ou non empilées
- Augmentation du nombre de niveaux pour le test d'ajustement du chi carré
- La sortie de la session gère les symboles grecs
- Affichage des équations du modèle pour les colonnes de réponse de régression, GLM et DOE
- Test de Bonett pour les distributions continues
- Coefficient de corrélation du rang de Spearman
- Tests de Grubbs et Dixon pour détecter les valeurs aberrantes
- Régression linéaire
  - De nouveaux graphiques incluent les effets principaux, l'interaction, les tracés de contour et de surface
  - Optimiseur de réponse pour trouver les paramètres optimaux
- Régression logistique binaire
  - Nouveau graphique pour représenter graphiquement les probabilités d'événements
  - De nouveaux graphiques incluent les effets principaux, l'interaction, les contours et les tracés de surface
  - Optimiseur de réponse pour trouver les paramètres optimaux
- Test de McNemar pour les proportions appariées
- Tests d'équivalences
  - Tests pour les plans croisés à 1 échantillon, 2 échantillons, appariés et 2x2
  - Analyse de puissance et de taille d'échantillon
- Graphiques
  - Nouveau graphique à bulles
  - Palette de couleurs, polices modernisées, etc.
- Fonctions
  - Nouvelles fonctions pour arrondir les colonnes de date/heure à l'année, au trimestre, au mois le plus proche, etc.

### Minitab 18.1

- L'équation de régression orthogonale et le taux de variance d'erreur sont maintenant affichés sur le tracé de droite d'ajustement.
- Le calcul de l'odds ratio prend en compte dorénavant les variables catégorielles
- Le Modèle à effets mixtes fournit le maximum de vraisemblance restreint et le maximum de vraisemblance pour ajuster des modèles linéaires avec des facteurs fixes et aléatoires.
- Modèles linéaires généralisés
  - Inclusion des intervalles de confiance pour les variances ainsi que les prédictions et les comparaisons multiples pour les facteurs aléatoires.
  - Inclusion d'optimisation des covariables avec ou sans contraintes

- Analyse des études avec des données manquantes (études non équilibrées).
- Analyse des études qui ont plus de deux facteurs.
- Inclure les termes qui sont croisés ou imbriqués.
- Incluez des facteurs fixes ou aléatoires.

## Minitab 19

- Version 64 bits disponible pour Microsoft Windows
- Nouveau format de fichier avec de nouvelles extensions de fichier
- Minitab pour MacOS (fonctionnalités légèrement limitées)
- Nouveau navigateur
- Nouvelle fenêtre de sortie comprenant des événements et des visualisations
- Basculement rapide entre les feuilles de calcul par onglets
- En plus des sorties graphiques et des cartes de contrôle, le résumé graphique, les diagrammes de Pareto, l'analyse des capacités et le tracé de probabilité normale peuvent maintenant être mis à jour par un simple clic
- Régressions
  - Nouveau modèle de sélection pas à pas basé sur AICc (critère d'information d'Akaike)
  - Nouveau modèle de sélection pas à pas basé sur BIC (critère d'information bayésien)
- Plans d'expériences
  - Possibilité d'analyser les données des variables de réponse binaires (par exemple bon/mauvais)
    - Peut être utilisé avec des plans factoriels, de criblage et de surface de réponse
    - l'optimisation de la réponse peut contenir des variables de réponse continues et binaires
- Analyse de capacité
  - Intervalles de confiance pour plus de paramètres
  - Méthode optimisée pour le calcul de la limite inférieure pour Z.Bench
  - Nouvelle option globale pour afficher PPM ou % et CI
- Test de rééchantillonnage/bootstrap/randomisation
  - Calculer des intervalles de confiance
  - Possibilité d'effectuer des tests d'hypothèse pour les statistiques sans supposer une distribution
- Intégration de Python

## Minitab 19.2

- Lorsque deux limites de spécification sont entrées dans la commande d'Etude de Stabilité, les utilisateurs ont désormais la possibilité de calculer la durée de conservation d'un médicament avec une limite de spécification et d'utiliser l'autre limite de spécification pour l'affichage uniquement.
- Les tables de corrélations fournissent les intervalles de confiance pour les corrélations de Pearson et de Spearman
- Pour aider les utilisateurs à visualiser clairement les résultats, la Régression inclut désormais un diagramme de Pareto facultatif pour montrer les effets significatifs.

## Minitab 20

- Version de Minitab disponible en tant que page Web sur le Cloud
- Intégration de R
- Organisation des résultats dans des "Rapports"
- Sélection multiple possible dans le navigateur
- Sélection de modèles pour la CART
- Validation de modèles de régression logistique binaire et de Poisson
- Graphiques
  - Nouveau graphique de coordonnées parallèles
  - Nouveau graphique de Nuage de points groupé
  - Nouveau graphique de carte de chaleur
  - Possibilité d'annoter les graphiques

### Minitab 21

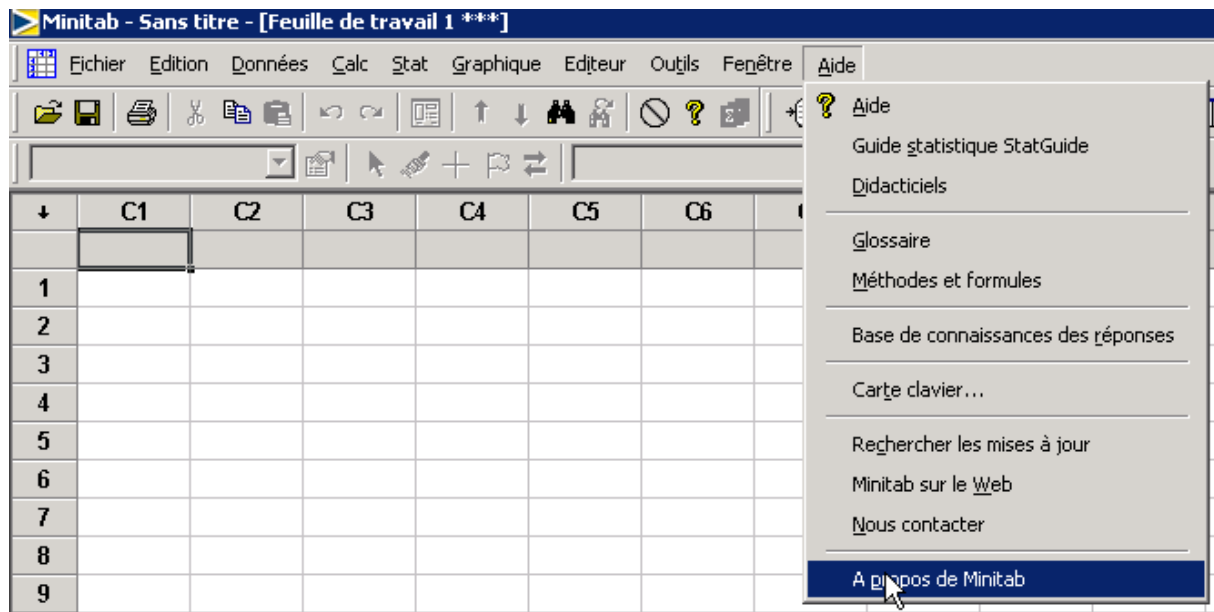
- Régression de Cox disponible (modèle à hasards proportionnels)
- Nouvel outil AutoML avec:
  - CART
  - TreeNet
  - Forêts aléatoires
  - Régression



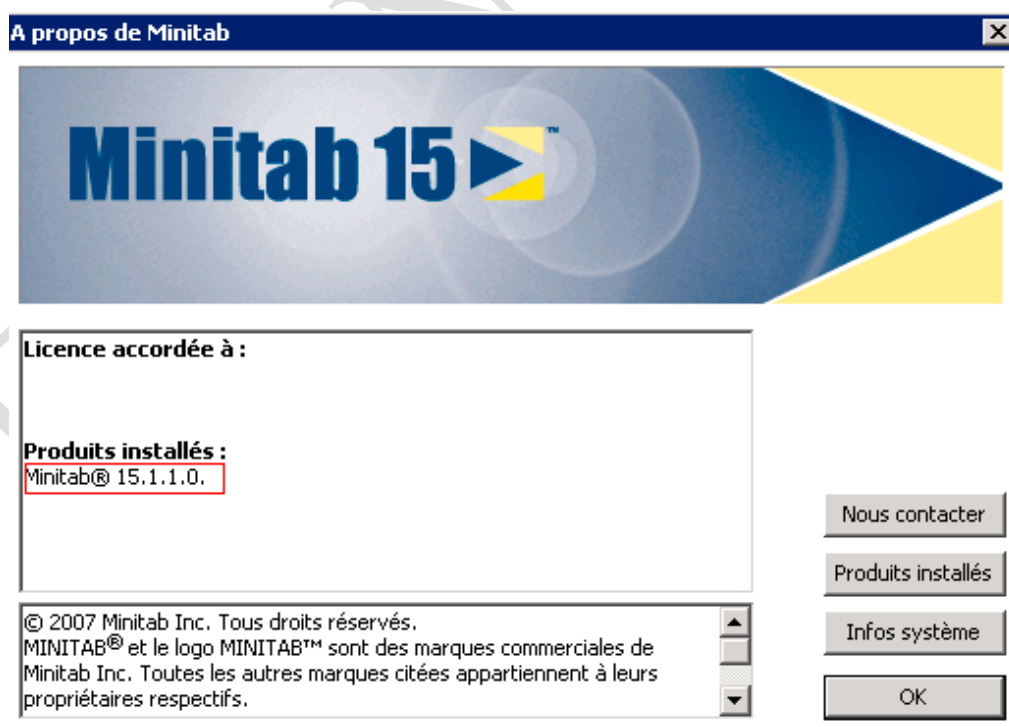
## 4. Divers (paramètres/configuration)

### 4.1. Déterminer la version

Lorsque vous citez le logiciel utilisé dans vos articles il ne faudra pas oublier de spécifier la version exacte utilisée du logiciel. Pour cela, allez dans le menu **Aide/À propos de Minitab**:



et vous aurez alors (ici la capture d'écran de la version 15):

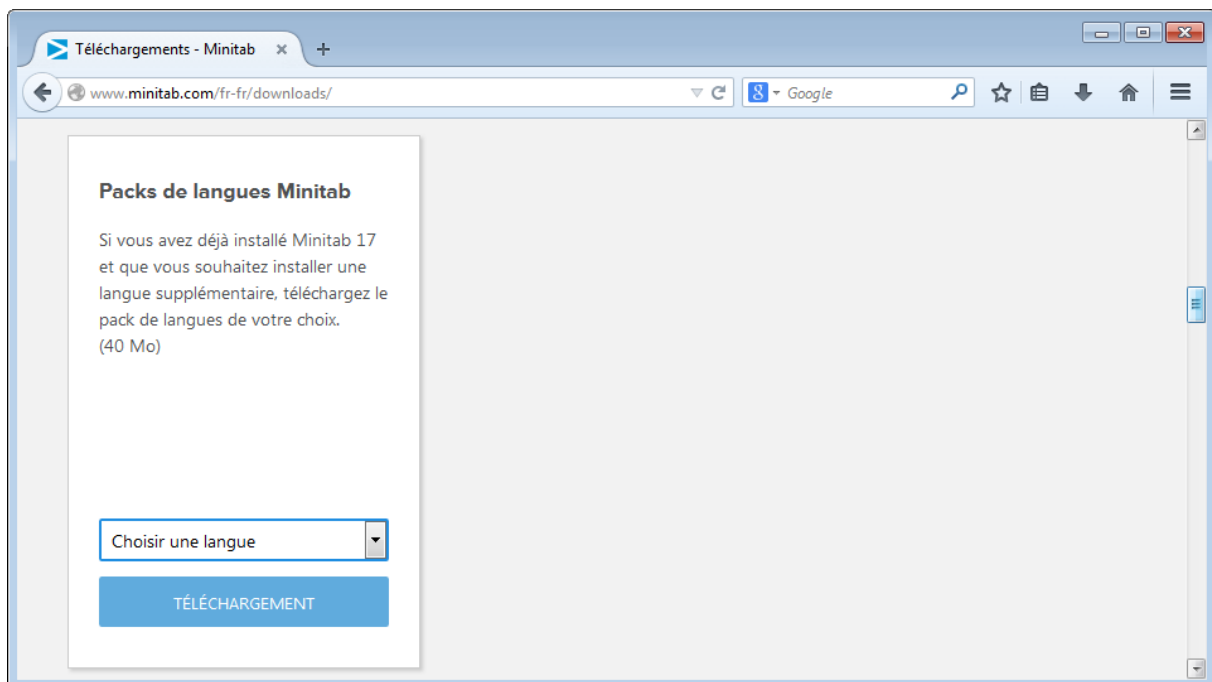


N'oubliez pas aussi qu'il est recommandé normalement d'utiliser une version majeure qu'à partir de sa deuxième version mineure (laissez les autres découvrir les bugs et s'énerver dessus!).

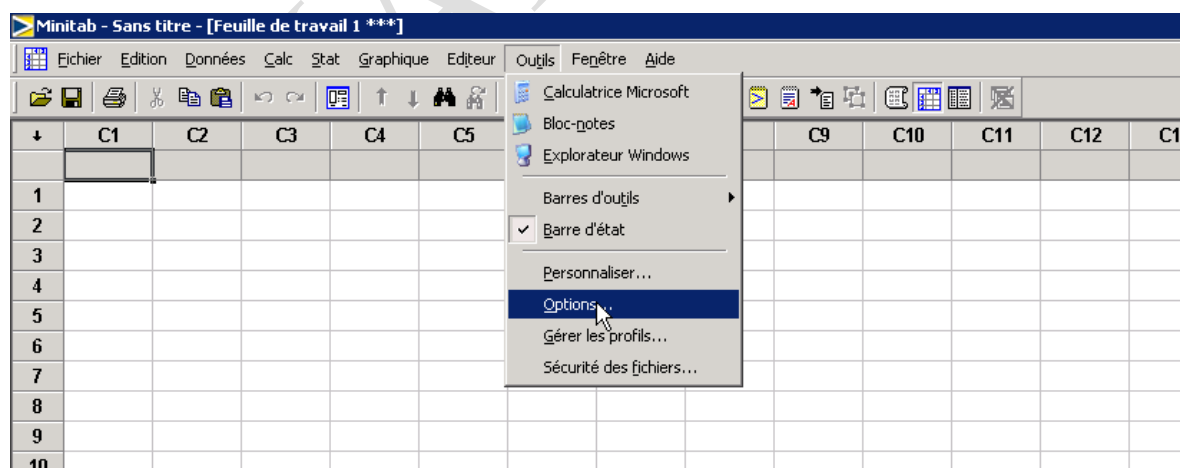
## 4.2. Changer la langue de l'interface

Un certain nombre d'entreprises font l'erreur d'acheter la version française de Minitab mais heureusement il y a la possibilité de basculer de langue à tout moment dans les dernières versions.

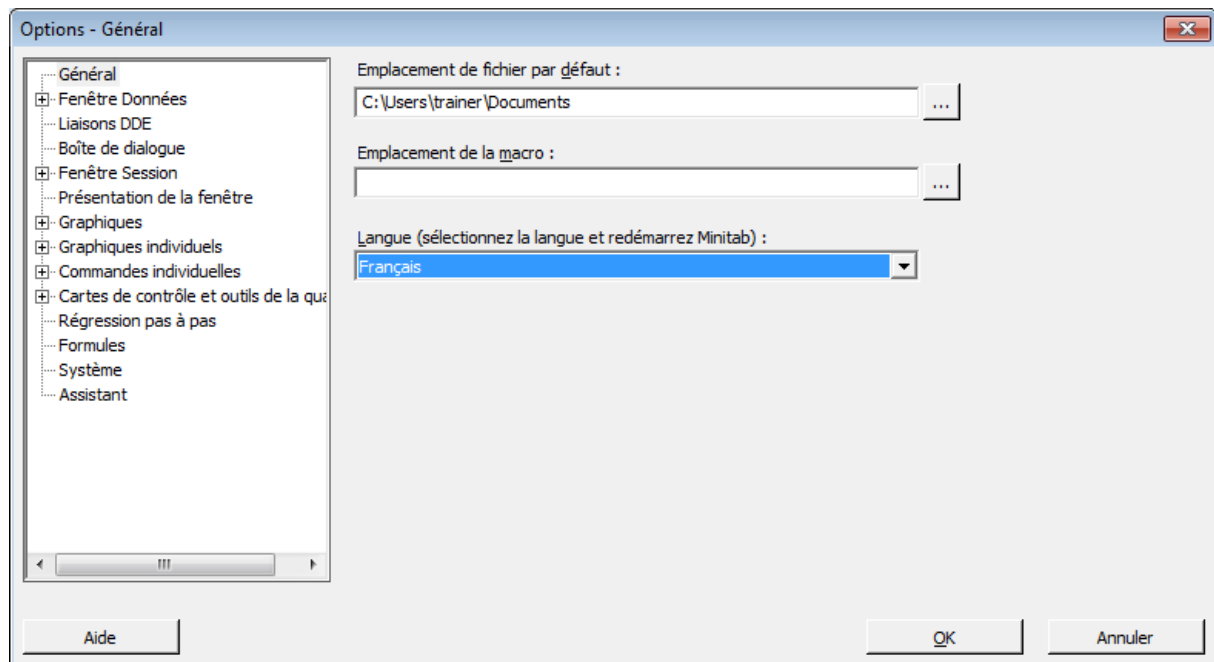
Pour ce faire vous pouvez vous rendre sur le site Internet de Minitab ou contacter leur service clientèle et télécharger et installer le pack de langue:



Ensuite, vous allez dans le menu **Outils/Options...**:



Pour dans la section **Général** y changer la langue:

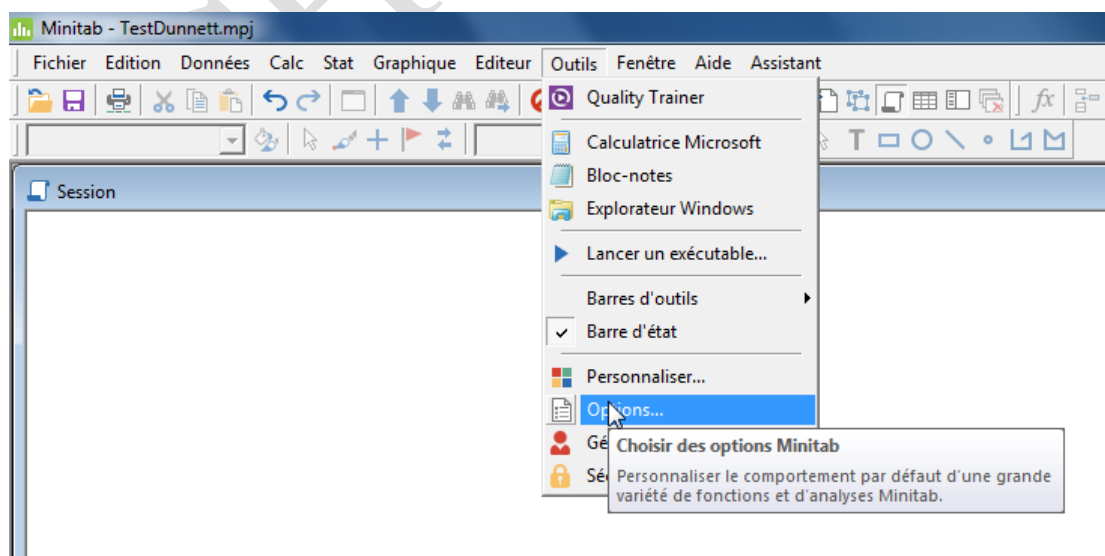


### 4.3. Désactiver la création automatique de fichiers de Backup

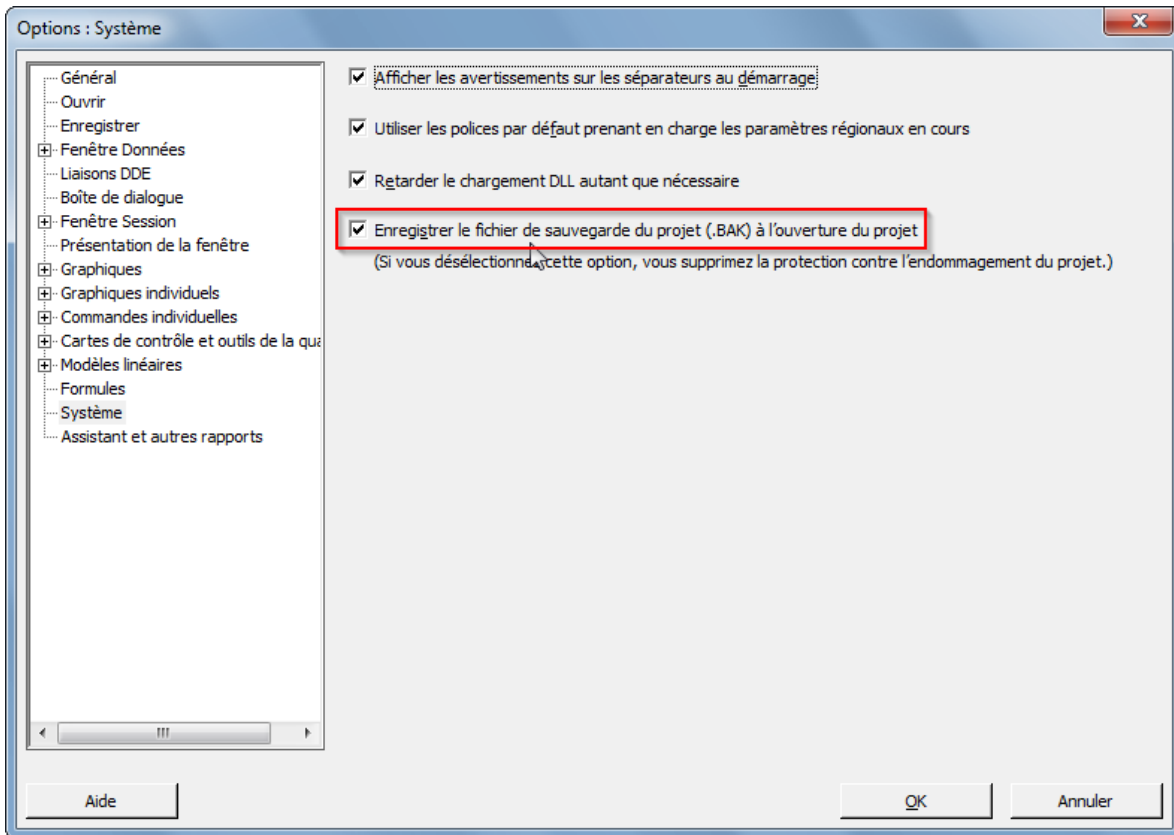
Par défaut à l'installation de Minitab, à chaque création/ouverture de fichier ce dernier crée une copie de sauvegarde \*.BAK des fichiers originaux. Cela peut être énervant surtout dans deux situations particulières:

1. Vous êtes professeurs/consultant et n'avez jamais besoin de ces back-up pour enseigner
2. Votre entreprise utilise une GED avec un système de versionning

Pour désactiver il suffit alors d'aller dans les options du logiciel via **Outils/Options...**:

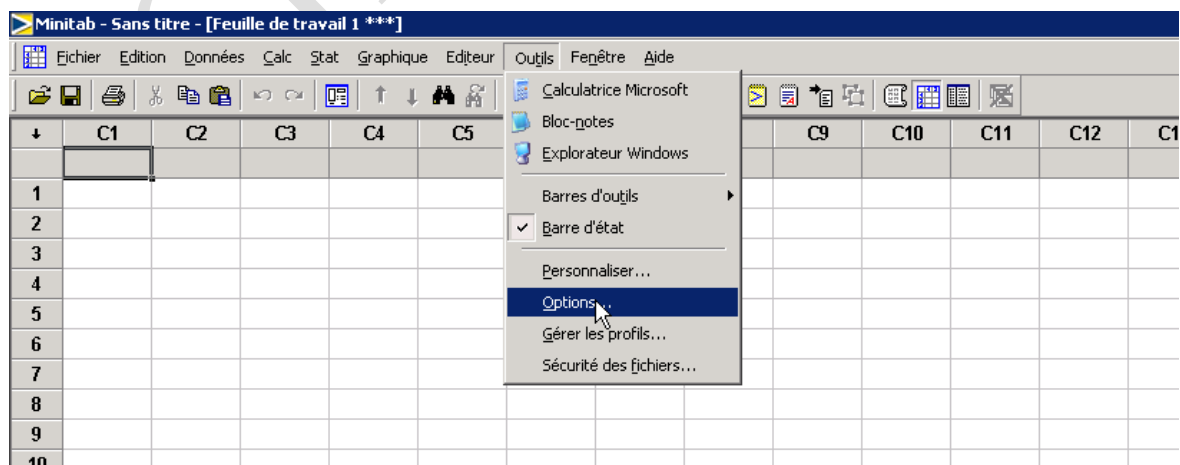


Et ensuite d'aller dans la section **Système** et décocher l'option **Enregistrer le fichier de sauvegarde du projet (.BAK) à l'ouverture du projet**:

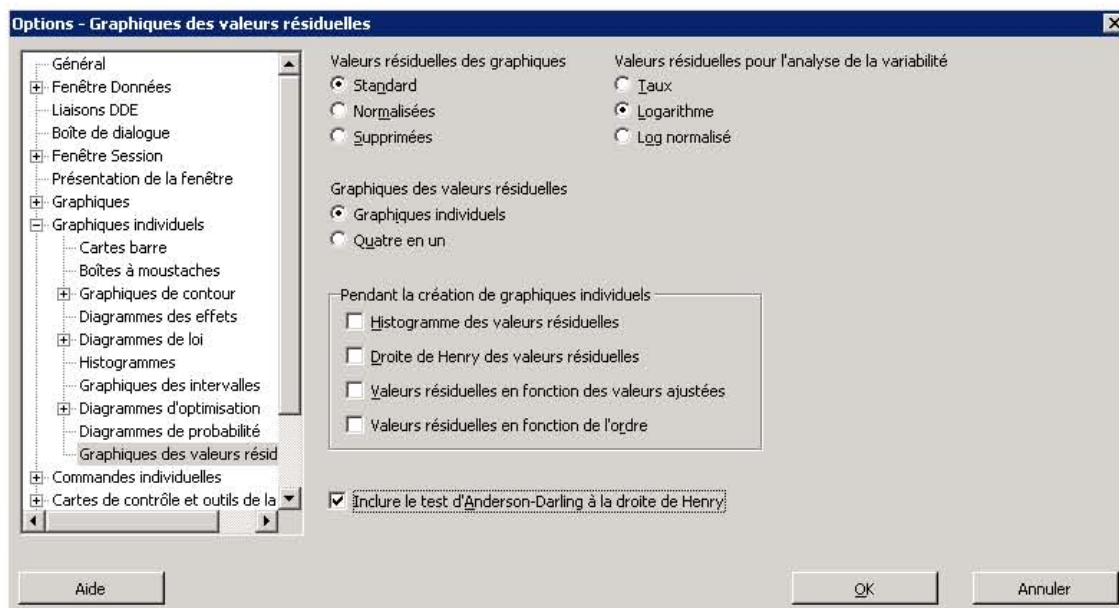


#### 4.4. Inclure test d'Anderson-Darling dans la droite de Henry

Le paramétrage le plus demandé en cours pour Minitab est celui concernant le nombre multiple d'outils affichant la droite de Henry mais cependant sans un test d'ajustement (ce qui est entre nous aberrant). Pour qu'à l'avenir toute droite de Henry vous affiche un test d'ajustement (de type Anderson-Darling dans le cadre de Minitab), il vous suffit d'aller dans le menu **Outils/Options...**:

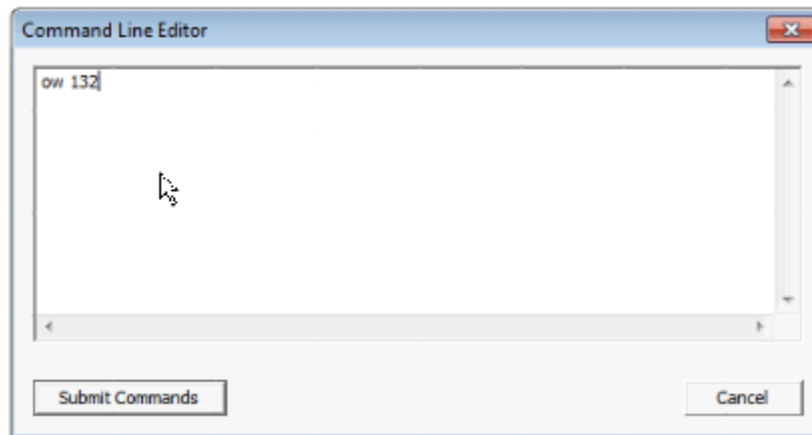


et ensuite aller dans **Graphiques individuels/Graphiques des valeurs résiduelles** et cocher **Inclure le test d'Anderson-Darling à la droite de Henry**:



## 4.5. Largueur de la fenêtre de session

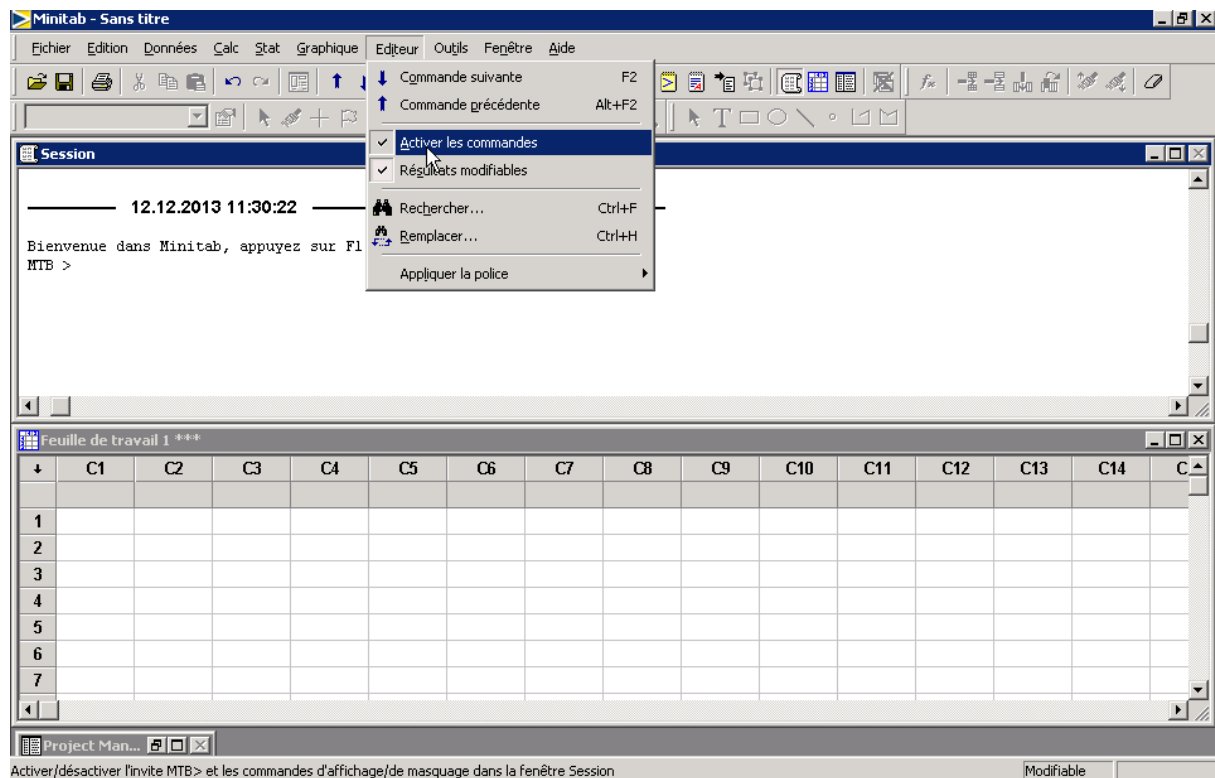
Autre paramètre parfois ennuyant dans Minitab... c'est que lors de certaines analyses statistiques, Minitab passe les résultats à la ligne dans la fenêtre de session. Alors pour changer cette largeur de sortie par défaut, il suffit d'ouvrir la fenêtre de commande (menu **Edition/Editeur de ligne de commande**) et de taper:



où *ow* signifie *Output Width*.

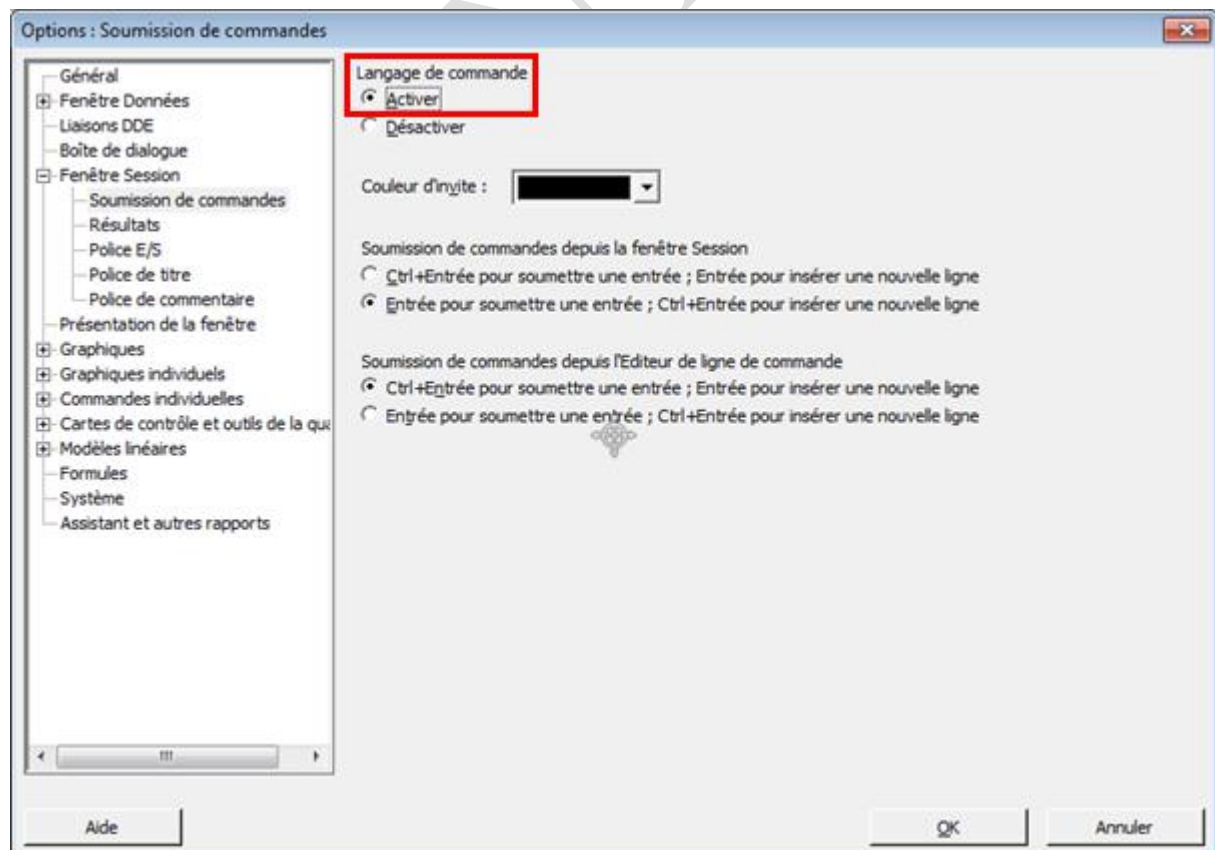
## 4.6. Obtenir les commandes de lignes des actions

Si vous souhaitez comme avec le logiciel R ou SAS communiquer dans vos articles/rapports les commandes Minitab utilisées par rapport aux données brutes (ce qui est normalement obligatoire), plutôt que de mettre des captures d'écran des différentes démarches activez simplement l'option suivante **Editeur/Activer les commandes**:



vous aurez ensuite chaque commande faite avec la souris qui sera écrit sous forme de code dans la fenêtre de session.

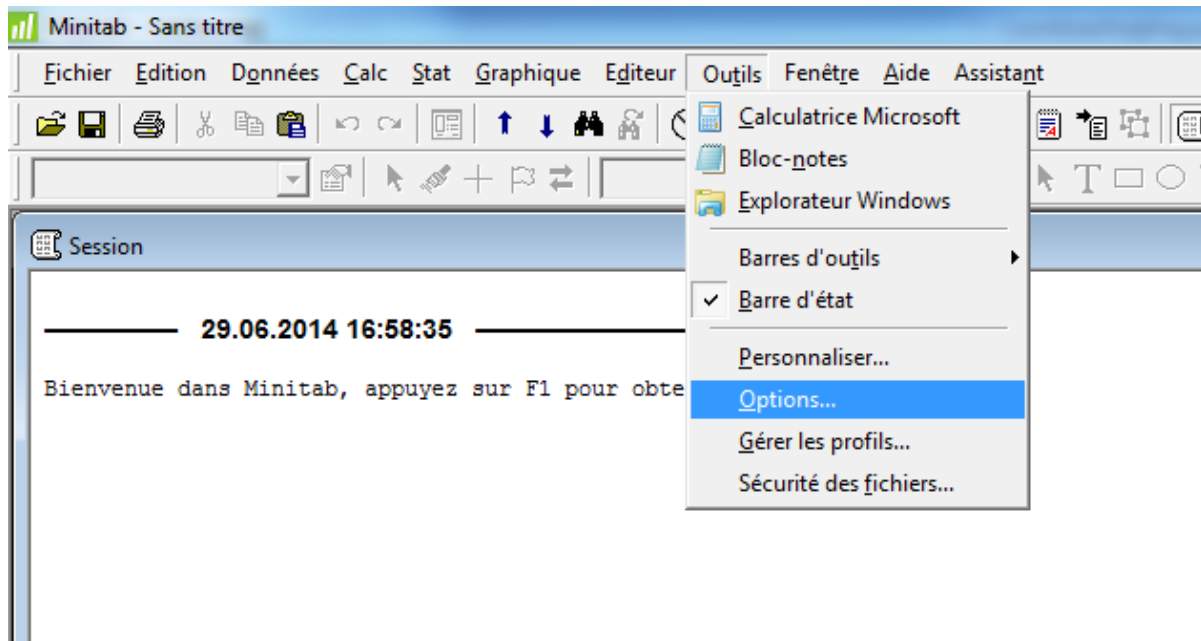
Pour Minitab 16 et 17 vous devrez passer par les options du logiciel:



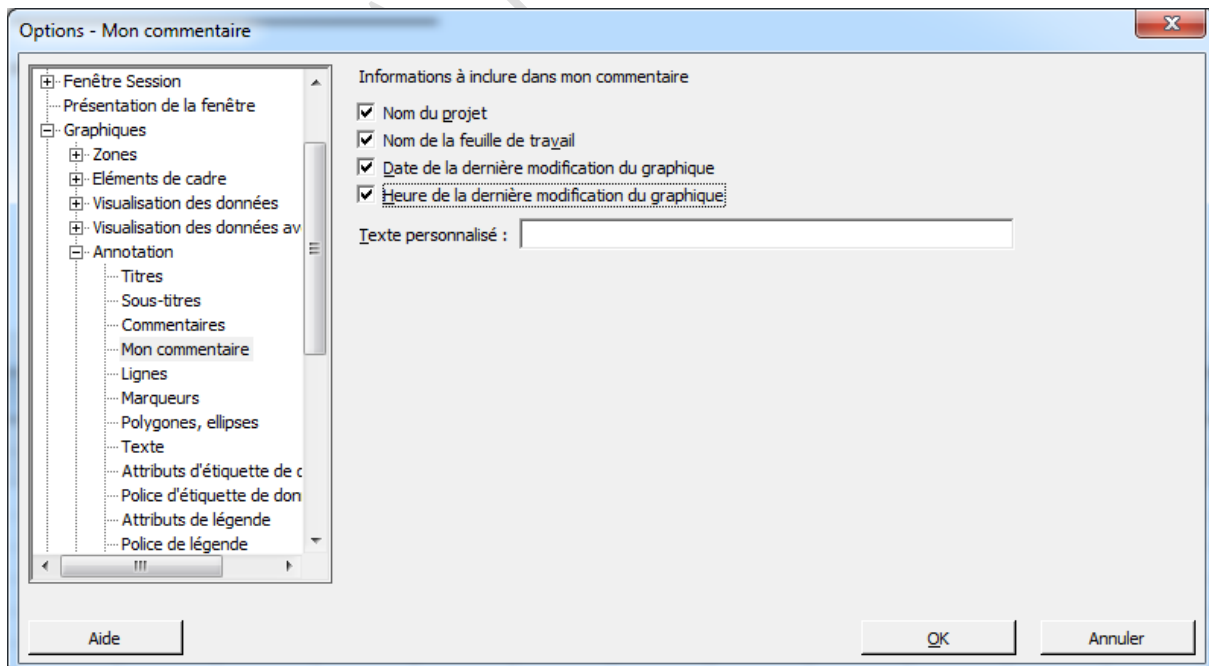
## 4.7. Afficher de la date/heure, nom du fichier projet/feuille sur les graphes

Dans les rapports il est souvent exigé pour des raisons de traçabilité d'afficher la date/heure, le nom du fichier projet/feuille sur chacun des graphiques.

Pour cela nous allons dans les options du logiciel via **Outils/Options...**:

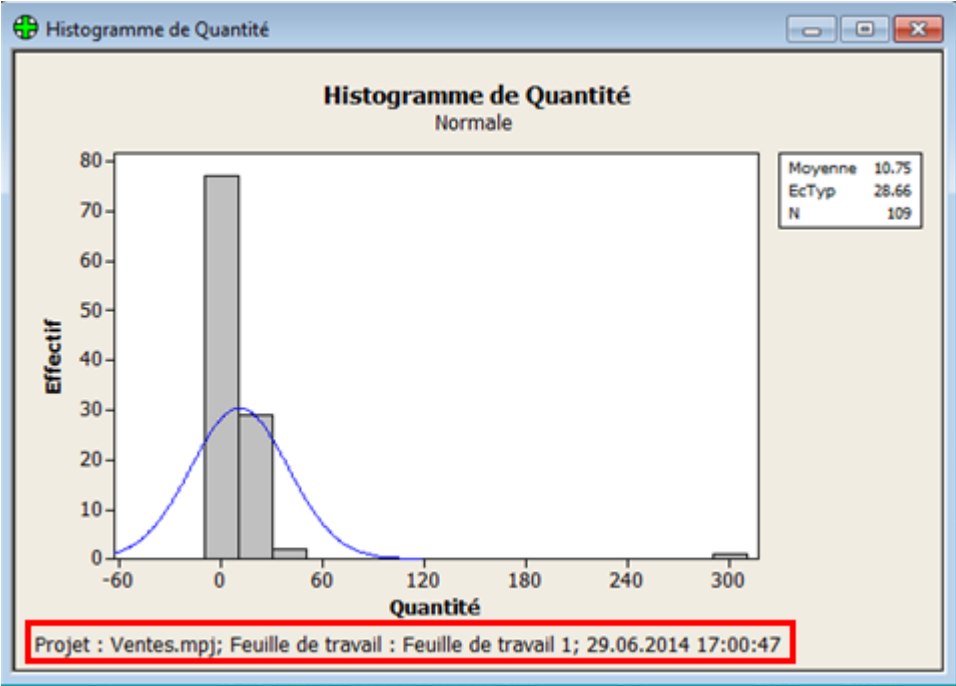


et dans la section **Graphiques/Annotation/Mon commentaire**, nous pouvons tout cocher:



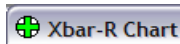
Ce qui donnera typiquement sur les graphes:





## 4.8. Symboles graphiques

Les graphiques ont tous dans l'extrémité supérieure gauche un petit logo qui varie en fonction du contexte. Pour ceux qui se demandent ce qu'ils représentent eh bien voilà un résumé de ces symboles d'état de graphiques:



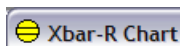
Graph à jour



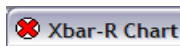
Graph à jour et ne pouvant pas être mis à jour si les données sources sont changées



Les données source du graphique ont été changées et le graphique ne peut être mis à jour.

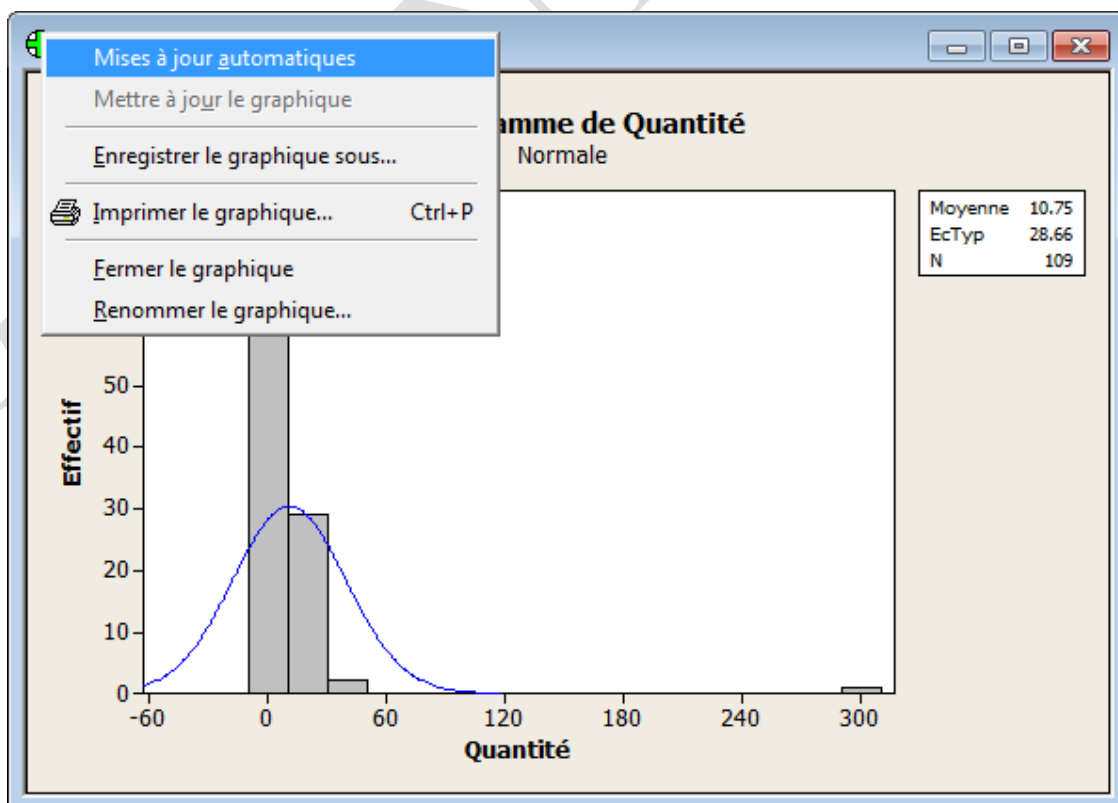


Les données sources du graphique ont été changées et le graphique peut être mis à jour



Les données sources du graphique ont été changées et certains paramètres font que celui-ci ne peut être mis à jour

Pour mettre à jour un graphique (qui peut être mis à jour) la manip classique consiste à faire un clic droit sur l'icône en haut à gauche.



## 4.9. Enregistrer seulement les feuilles Minitab pour ouverture ultérieure dans R

Il n'est pas judicieux d'enregistrer toujours les fichiers Minitab au format \*.mpj. Effectivement dès que votre organisation arrêtera d'acheter le logiciel vous serez bloqué avec un format propriétaire dont l'accès au contenu est plus que problématique.

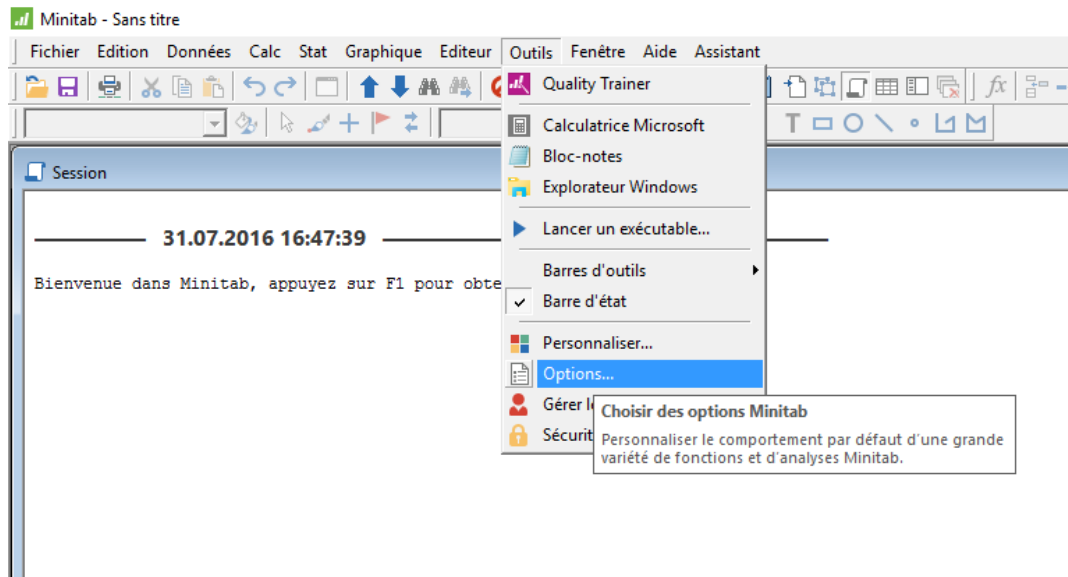
Donc à chaque fois que vous créez un fichier \*.mpj qui contient une seule feuille soit:

1. Vous mettez les données aussi dans un fichier Microsoft Excel \*.xlsx
2. Vous exportez (encore mieux) dans un fichier \*.csv
3. ~~Vous enregistrez au format portable \*.mtp~~ (fonctionnalité dépréciée)

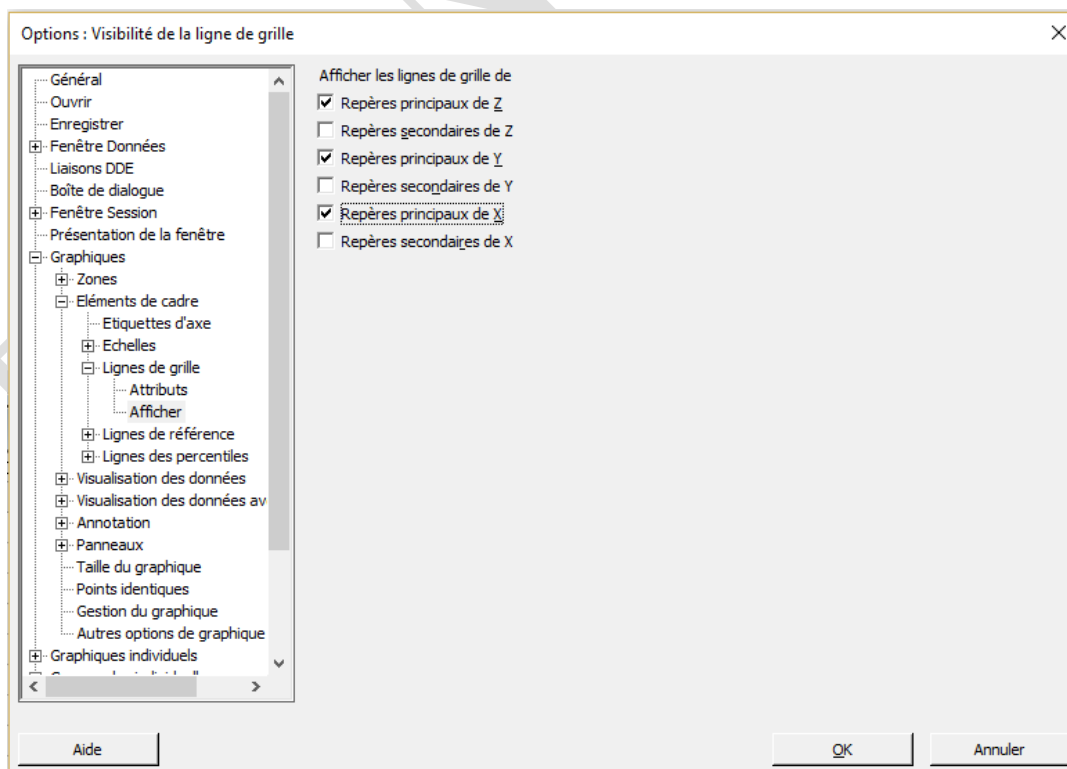
## 4.10. Afficher les grilles par défaut dans les graphiques

Il arrive fréquemment que les graphiques Minitab soient peu lisibles si on n'active pas la grille de fond des graphiques. Nous allons voir ici comment activer la grille pour tous les graphiques par défaut.

Il faut aller dans **Outils/Options...**:



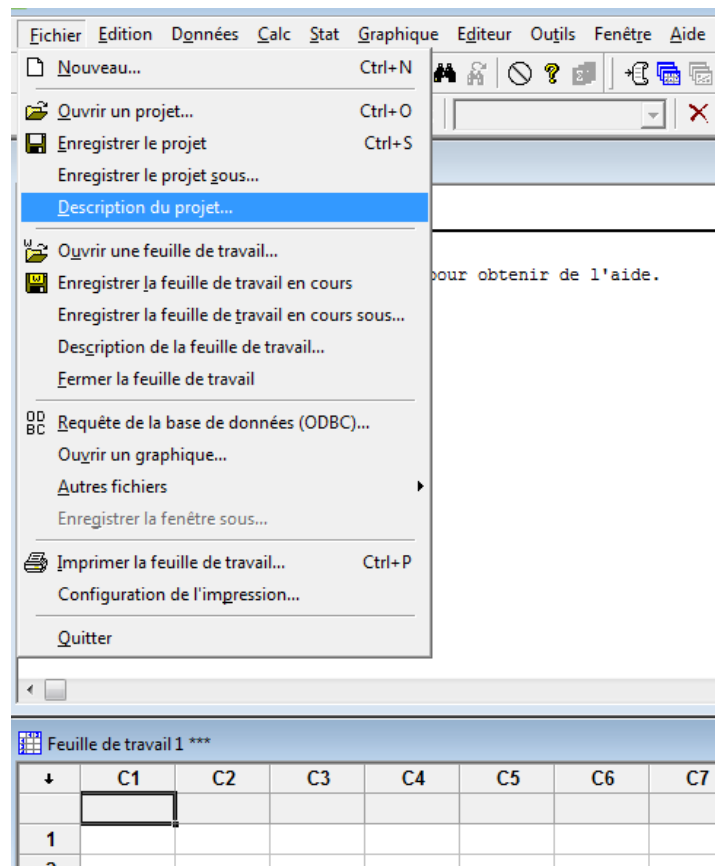
Ensuite dans la boîte de dialogue d'options il suffit d'aller dans **Graphiques/Eléments de cadre/Lignes de grille/Afficher...**:



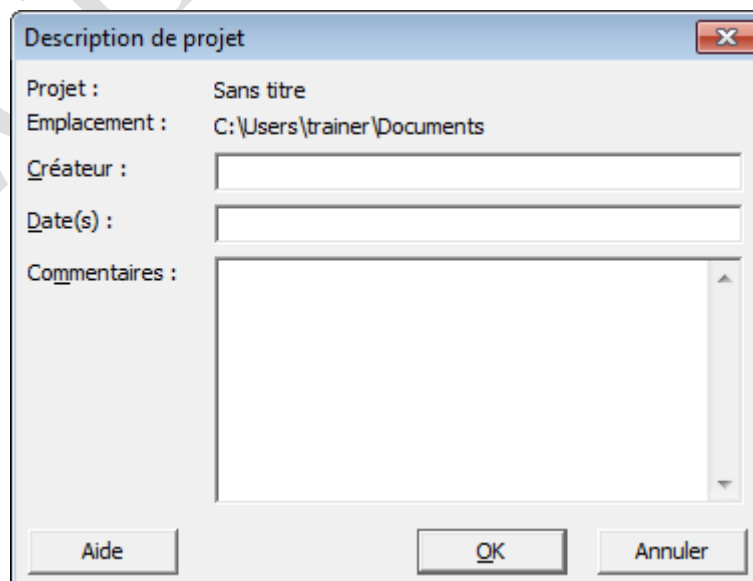
## 4.11. Définir les métadonnées des fichiers et des feuilles

Dans le cadre du respect de certaines normes documentaires internationales, nous avons l'obligation scientifique de saisir les métadonnées des fichiers de traitement statistiques.

Voyons d'abord comment définir les métadonnées d'un fichier en lui-même en allant dans le menu **Fichier/Description du projet...**:

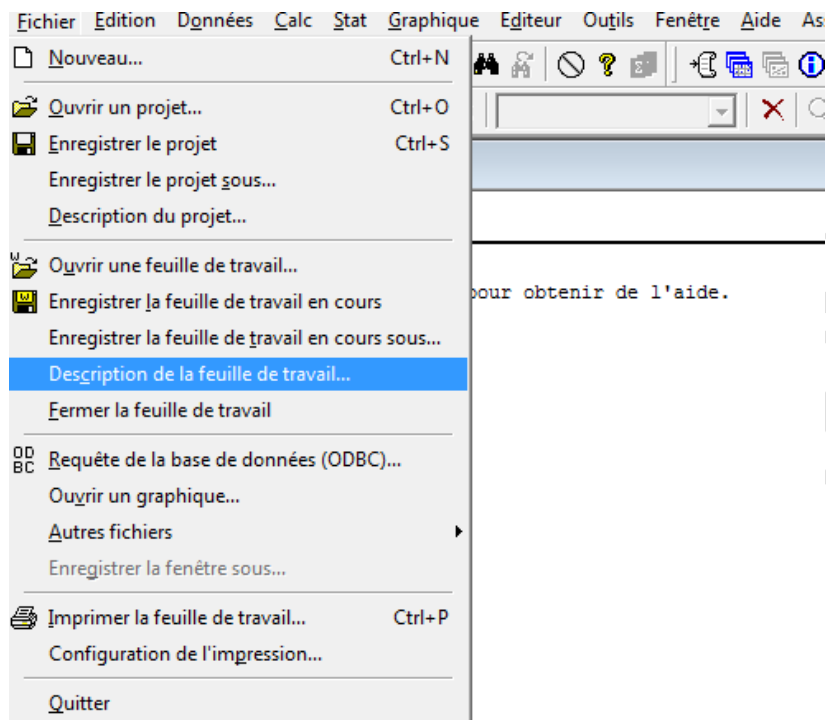


Ce qui nous amène à:

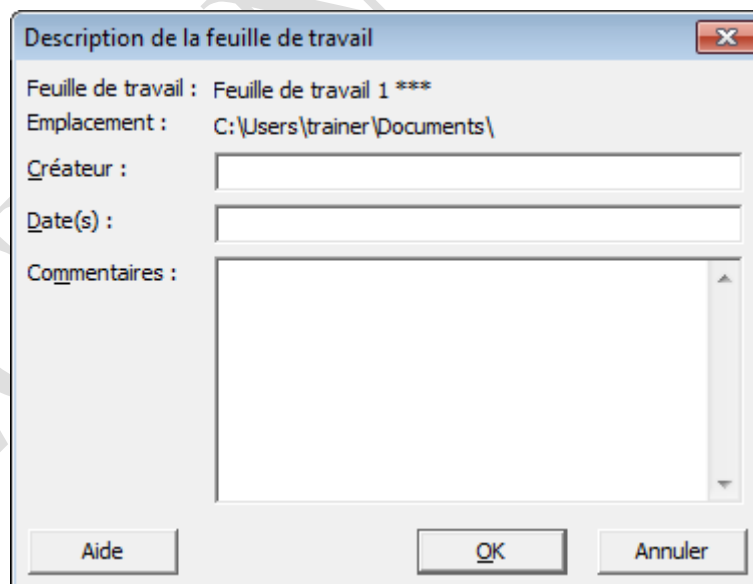


Les métadonnées sont bien pauvres... mais bon...

Voyons maintenant au niveau des feuilles de données (au même titre que R et SPSS). Pour cela nous allons dans le menu **Fichier/Description de la feuille de travail...**:

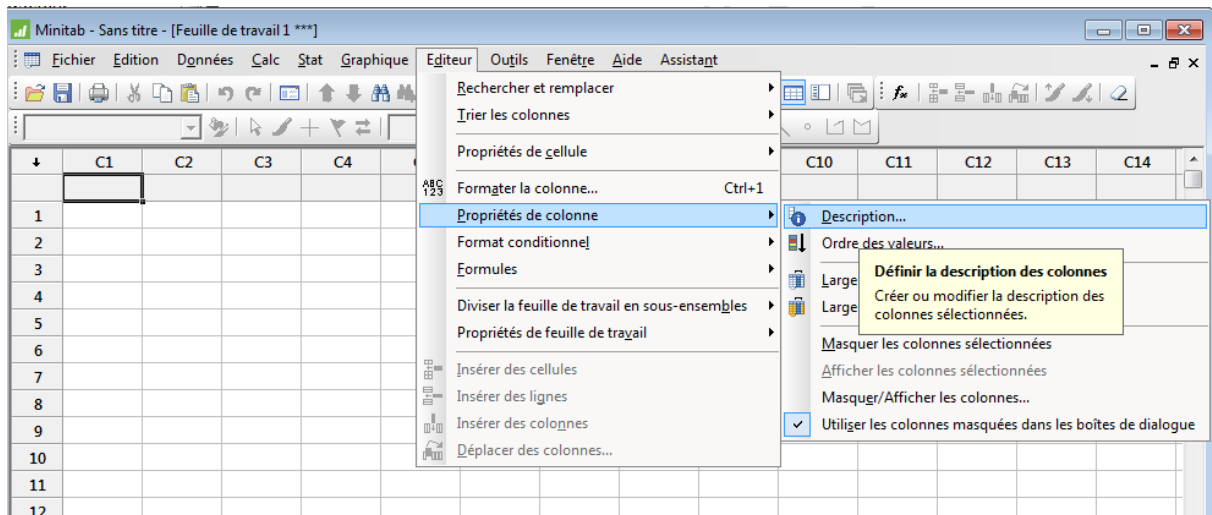


Ce qui nous amène à:

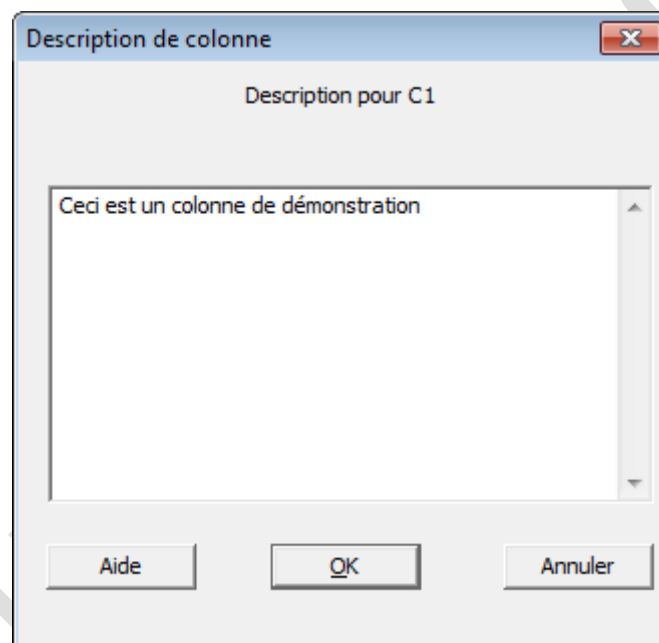


Les métadonnées sont là aussi bien pauvres...

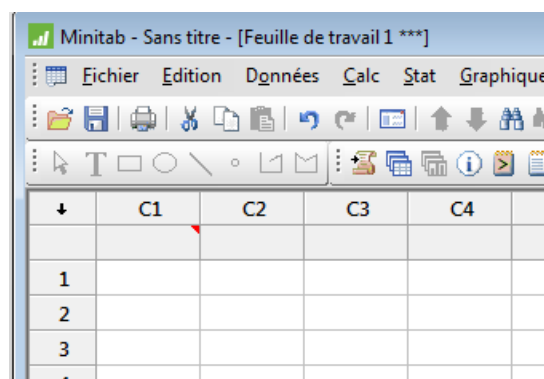
Nous pouvons aussi saisir une description au niveau des colonnes:



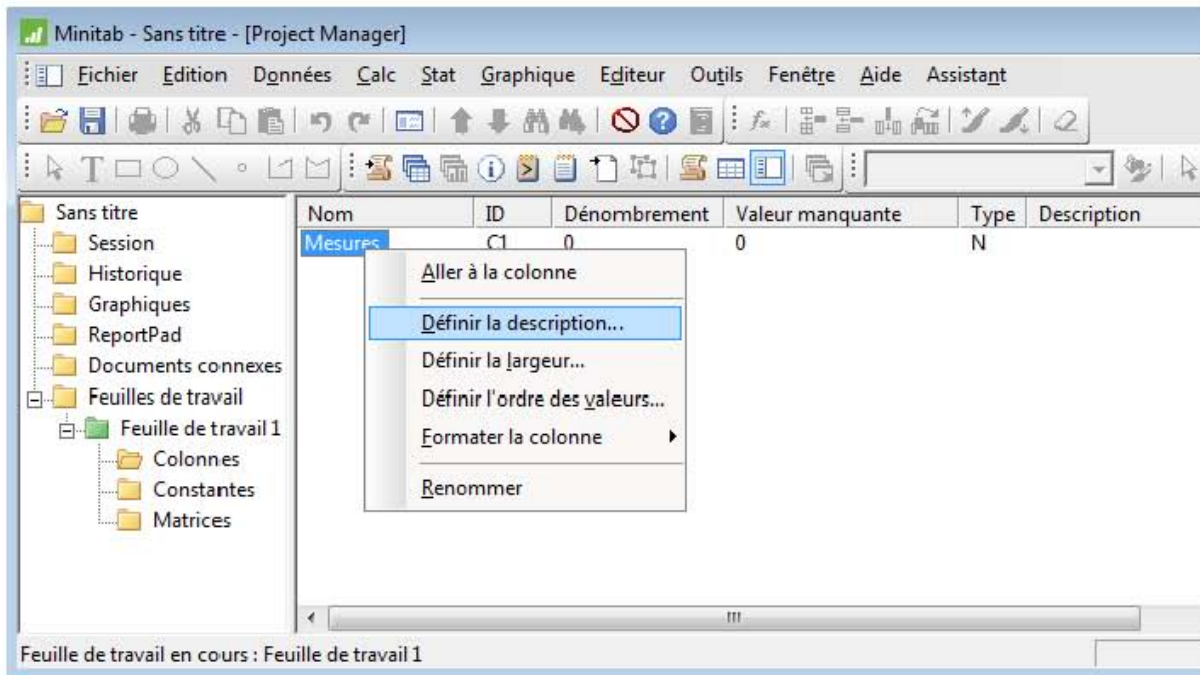
Une fois la description saisie:



il y aura comme dans Microsoft Excel un petit indicateur pour la colonne (il en est de même pour les commentaires sur les feuilles):



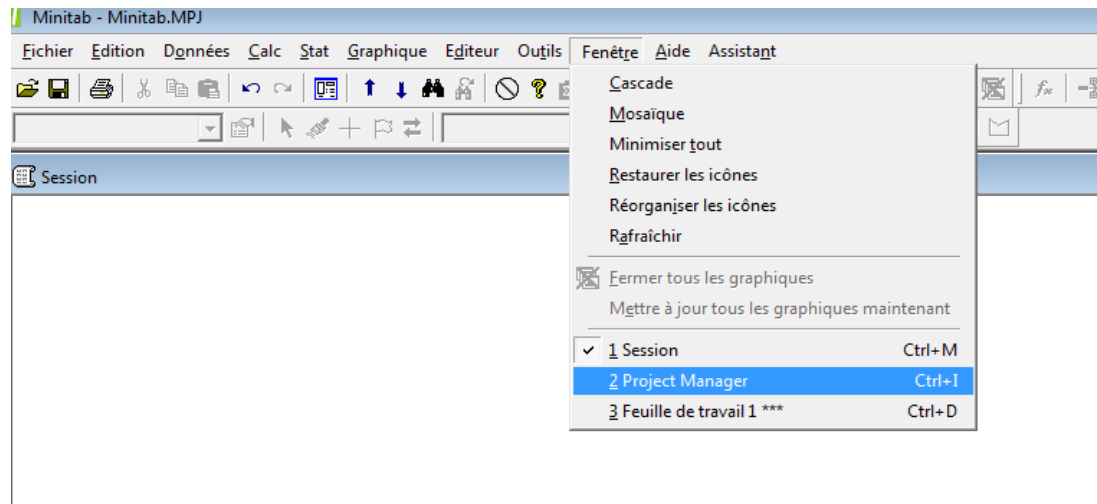
Nous pouvons aussi créer/consulter/modifier les commentaires et métadonnées via la fenêtre du project manager pour le fichier, les feuilles ET les colonnes:





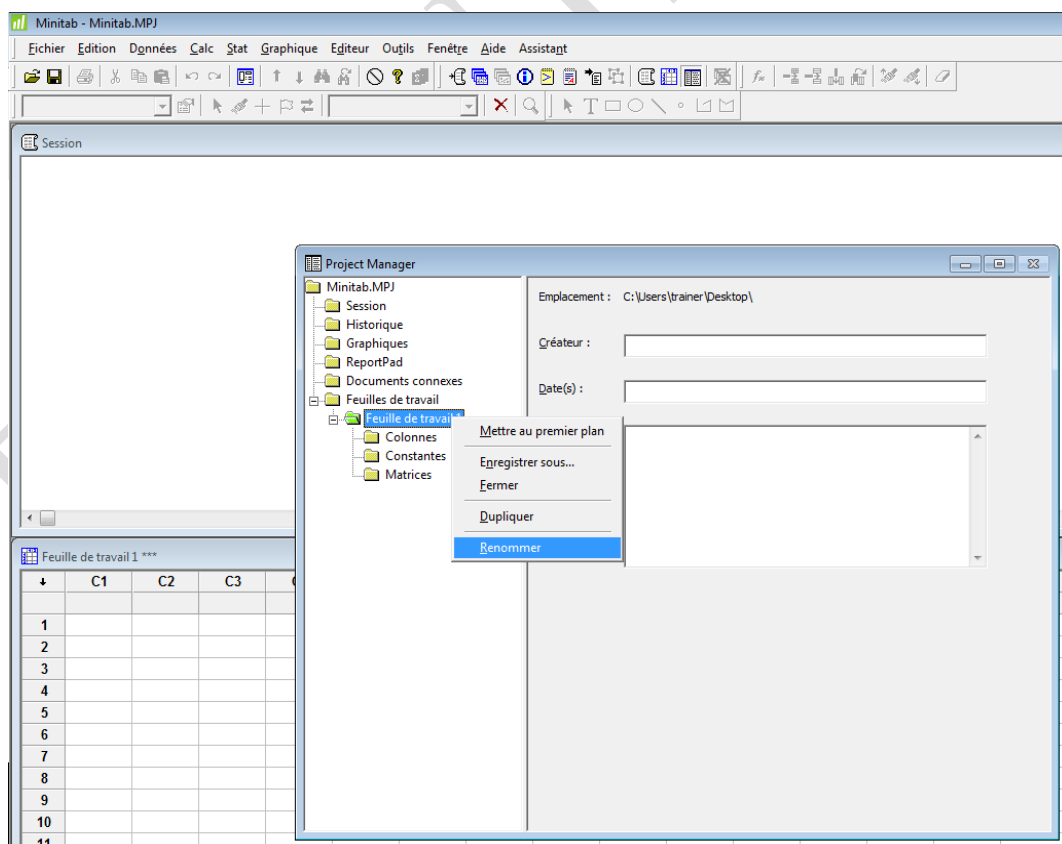
## 5. Project Manager

Tout le contenu MINITAB se gère et s'organise dans une fenêtre que l'on appelle le Project Manager. Pour y accéder, il nous suffit de cliquer sur **Fenêtre/Project Manager**:

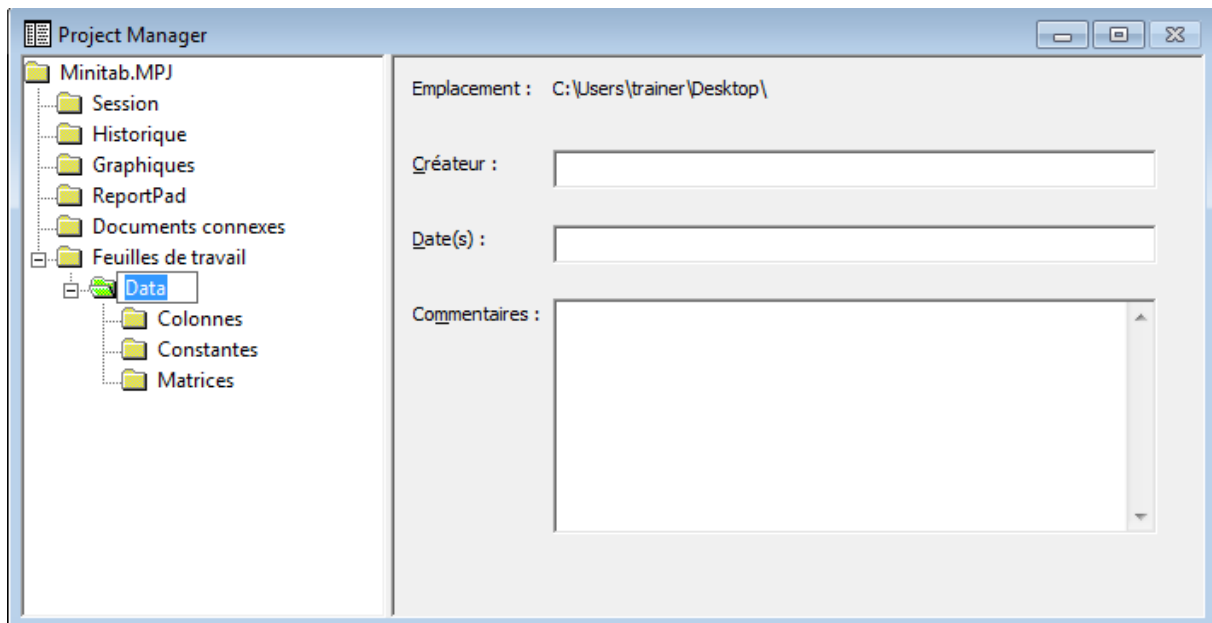


### 5.1. Ajouter/Renommer des feuilles

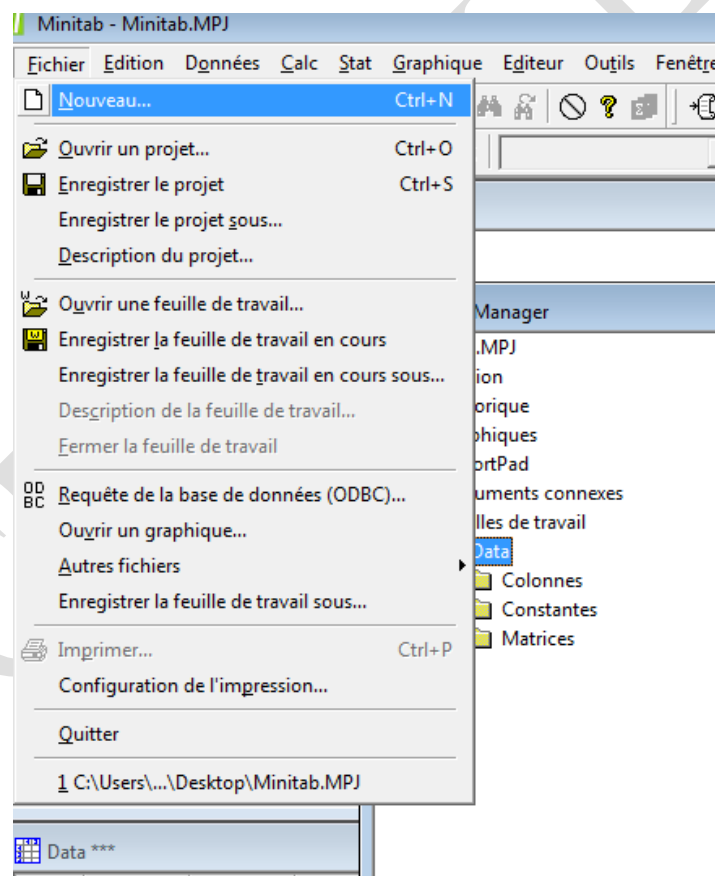
La première chose à savoir faire est de renommer les feuilles. Pour cela il suffit de faire un clic droit dessus et de choisir **Renommer** (observez cependant aussi **les autres options!**):



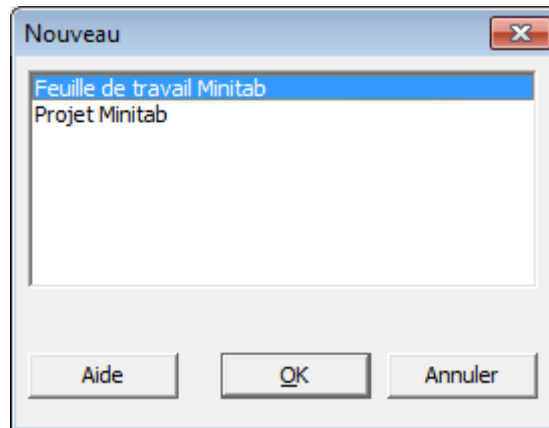
Et ensuite il vous suffit de saisir ce que vous voulez (**remarquez les métadonnées de la feuille à droite!**):



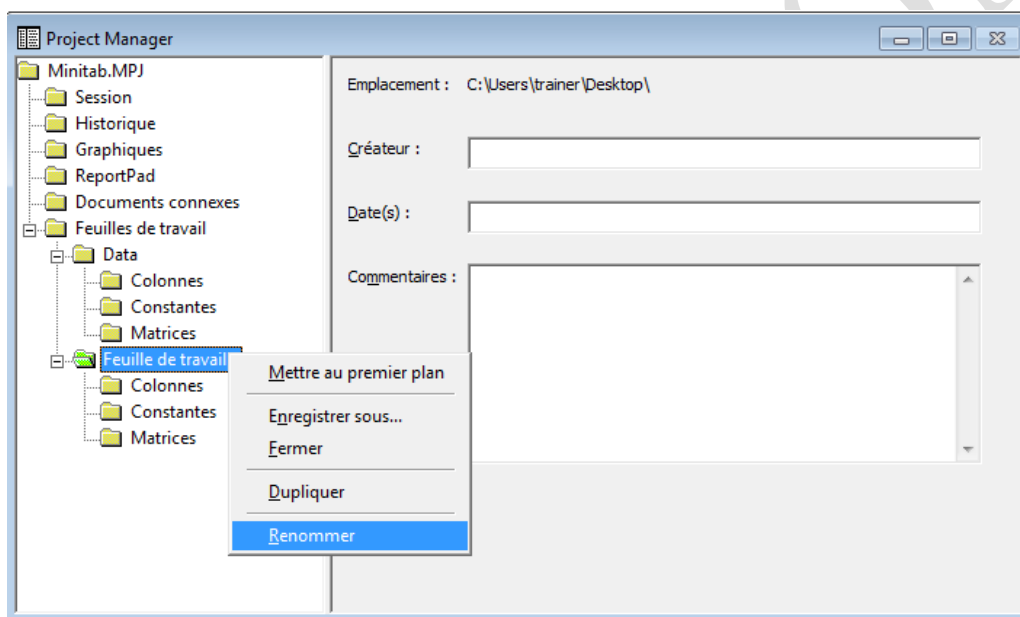
Si nous créons une nouvelle feuille de travail en allant dans **Fichier/Nouveau...**:



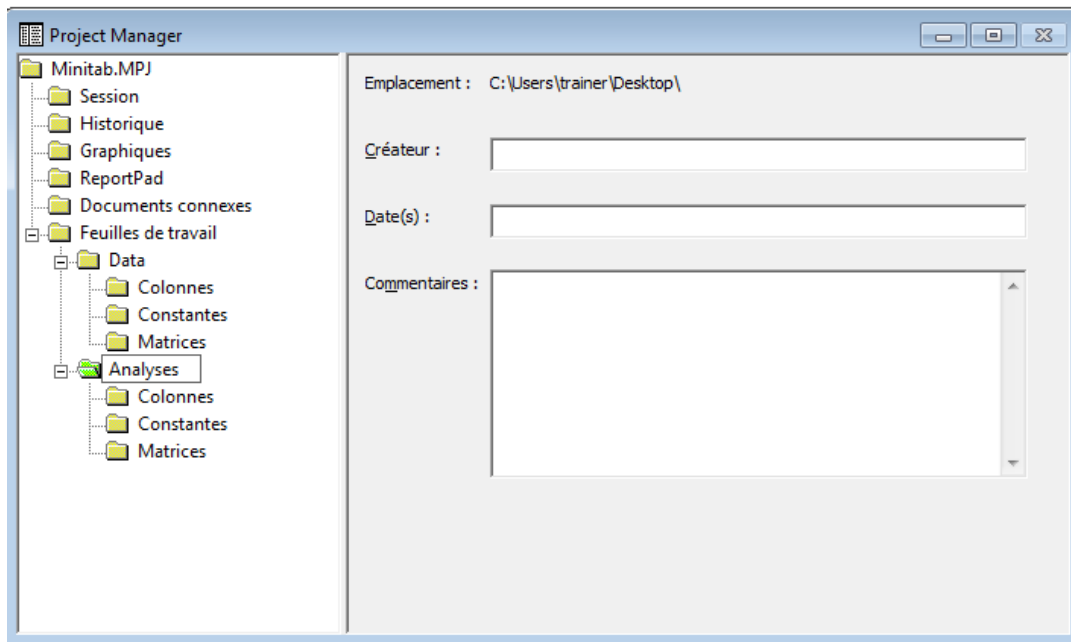
et en choisissant **Feuille de travail Minitab**:



Cette dernière s'ajoute bien évidemment au gestionnaire de projet et nous pouvons aussi renommer une feuille:



et là encore il suffit de saisir ce que l'on veut:



Voici sinon une image résumant l'aspect d'une feuille sous Minitab:

"T" for text column      "D" for date/time column      Numeric column

↓      ↓      ↓

+	C1-T	C2-D	C3	C4
	Physician	Appointment	Wait Time	Total time
1	Jones	1/5 15:00	28	43
2	Mayer	1/5 15:00	37	57
3	Crandall	1/5 15:00	22	34
4	Moran	1/6 8:00	15	29
5	Jones	1/6 8:00	3	20
6	Mayer	1/6 8:30	10	32

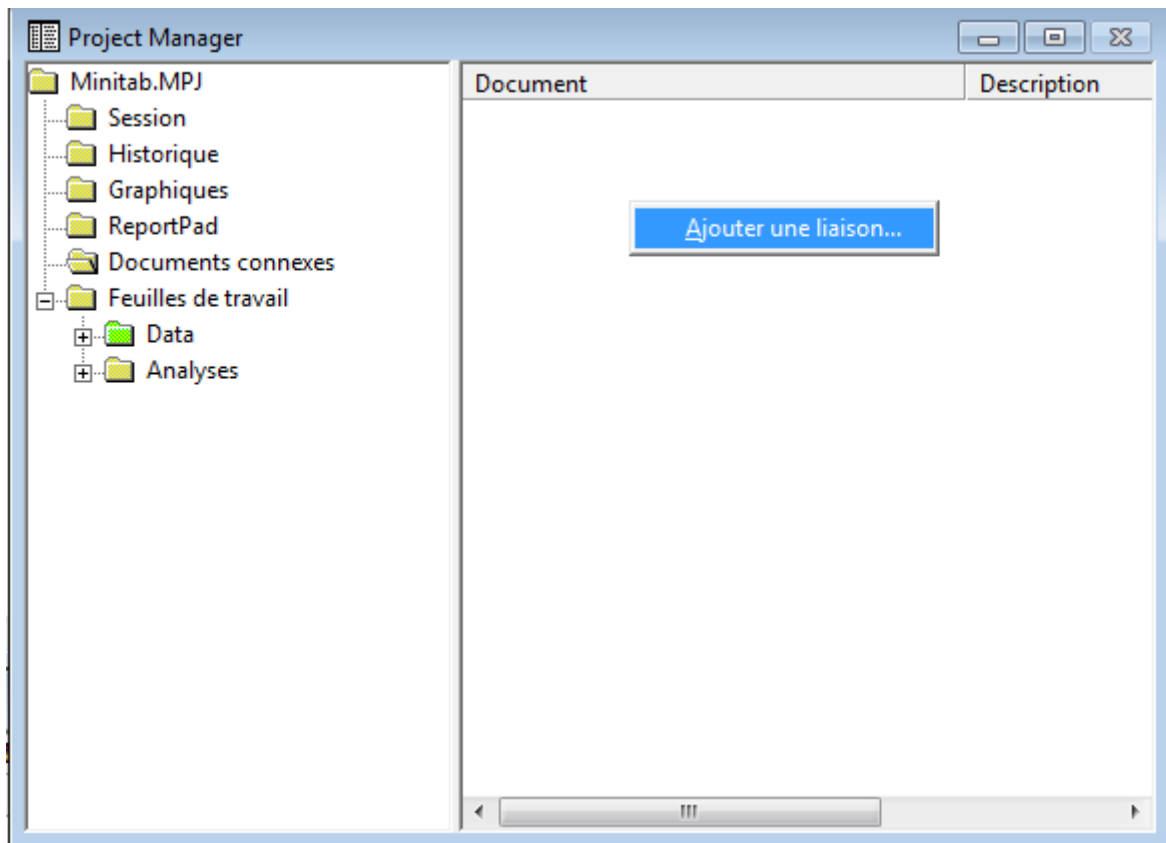
Data entry direction      Formula indicator

Column description indicator

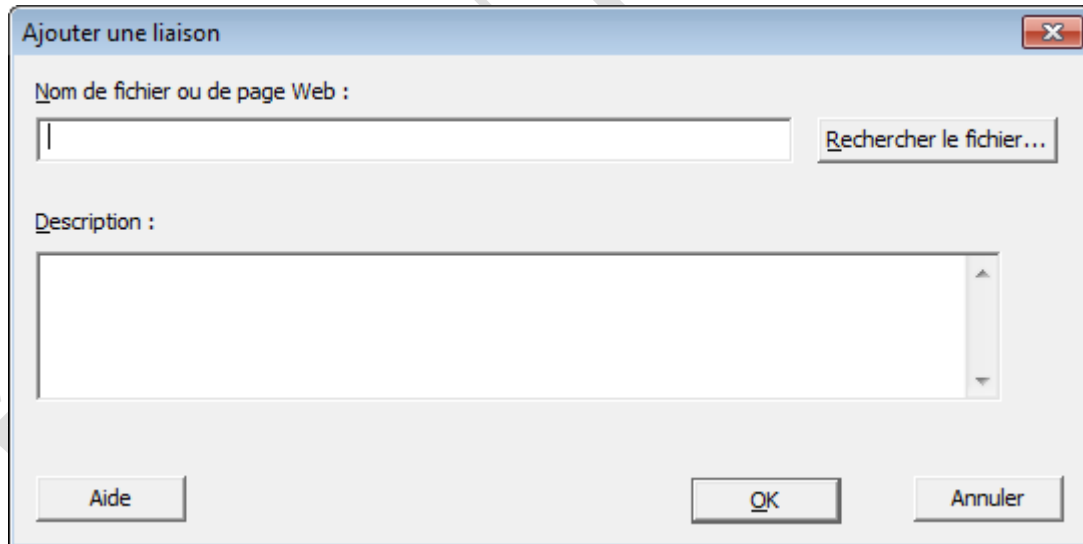
Les \*\*\* (étoiles) signifiant que la feuille est... la feuille actuellement active (...)!

## 5.2. Joindre des fichiers externes

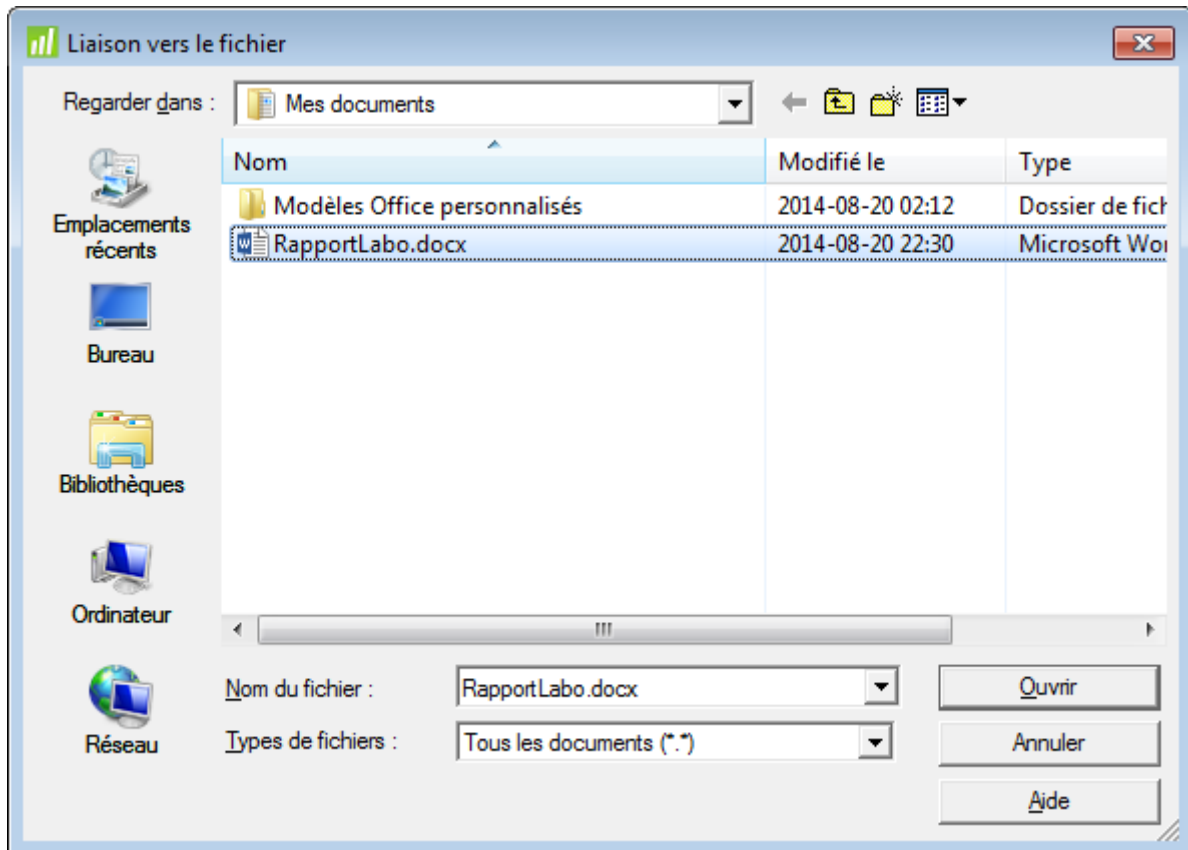
Chose très pratique dans Minitab, nous pouvons associer la liste des fichiers connexes à nos analyses statistiques. Pour ce faire, nous, nous allons dans la section **Documents connexes** du **Project Manager** et faisons un clic droit dans partie vide pour sélectionner l'option **Ajouter une liaison...**:



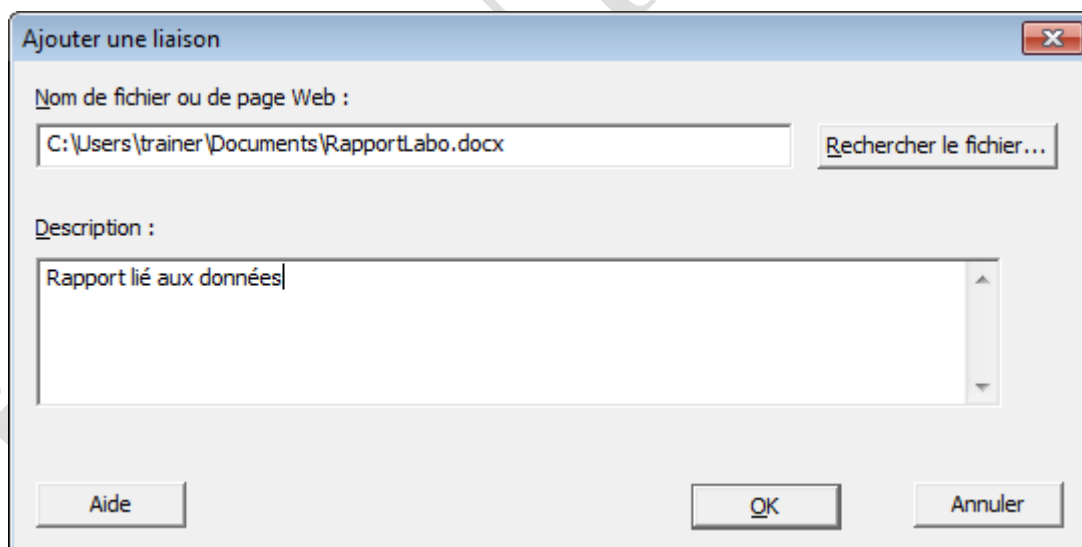
Apparaît alors:



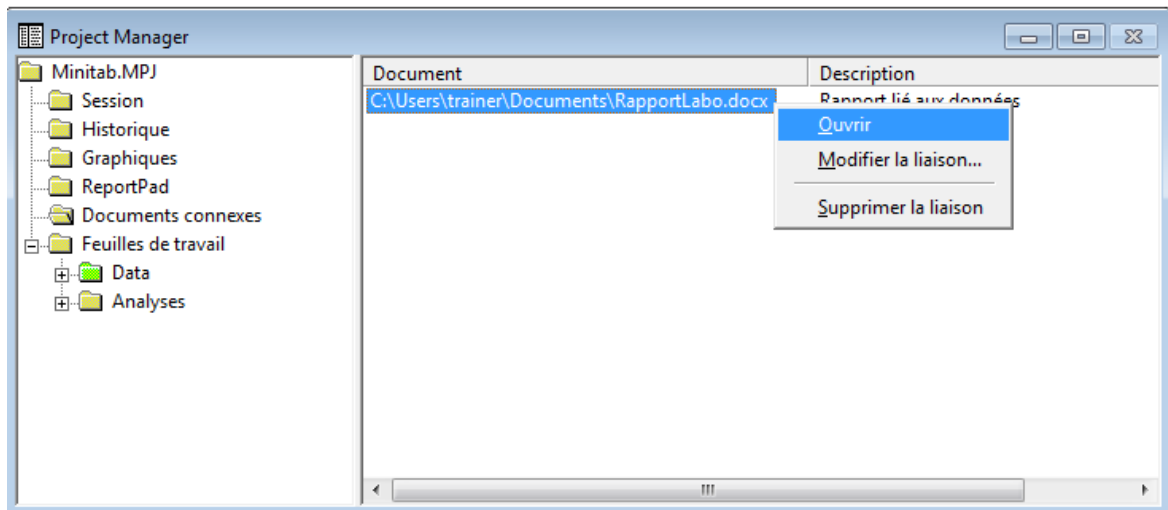
Nous cliquons sur **Rechercher le fichier...**:



et sélectionnons le fichier désiré et validons pour revenir à:



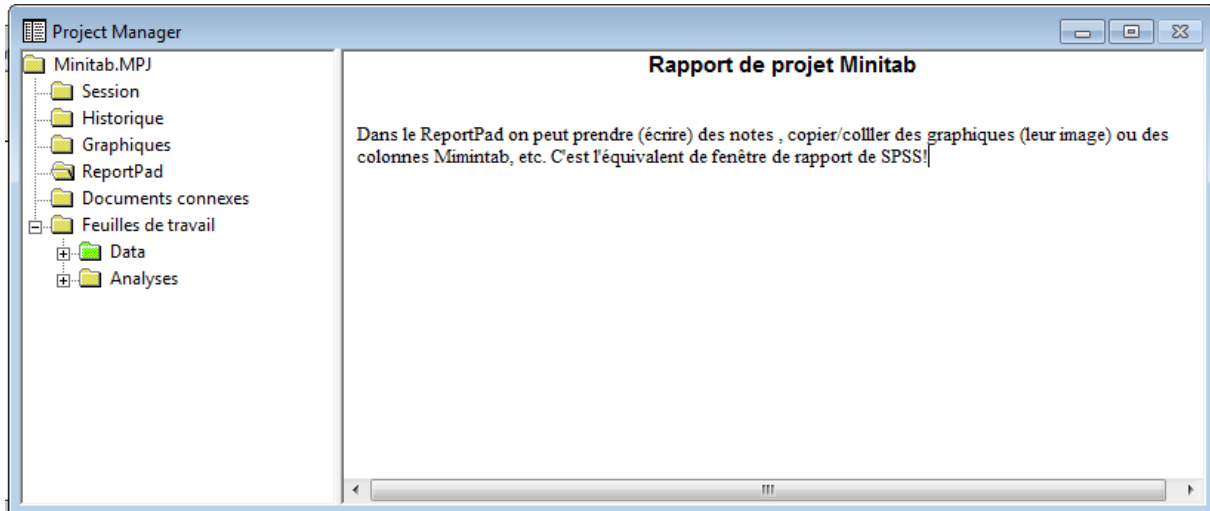
Tout en oubliant pas de saisir une **Description**. Ce qui donne après validation par **OK**:



et comme vous pouvez le constater, nous pouvons donc rapidement ouvrir la source ou modifier sa liaison.

### 5.3. Rapport/Notes internes (ReportPad)

Comme SPSS, Minitab possède dans le Project Manager une section **ReportPad** permettant aux analystes d'y prendre des notes diverses et variées et aussi d'y coller des images de graphiques:



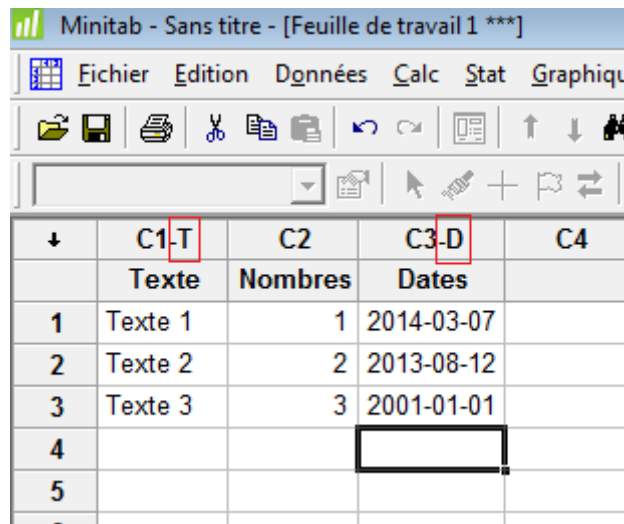
Rien de bien complexe à utiliser mais très pratique!



## 6. Saisie de données/Manipulations des données

La saisie de données dans Minitab se déroule comme dans n'importe quel tableur (mais attention Minitab n'est pas un tableur car on ne peut pas écrire de relations mathématiques dans les cellules!).

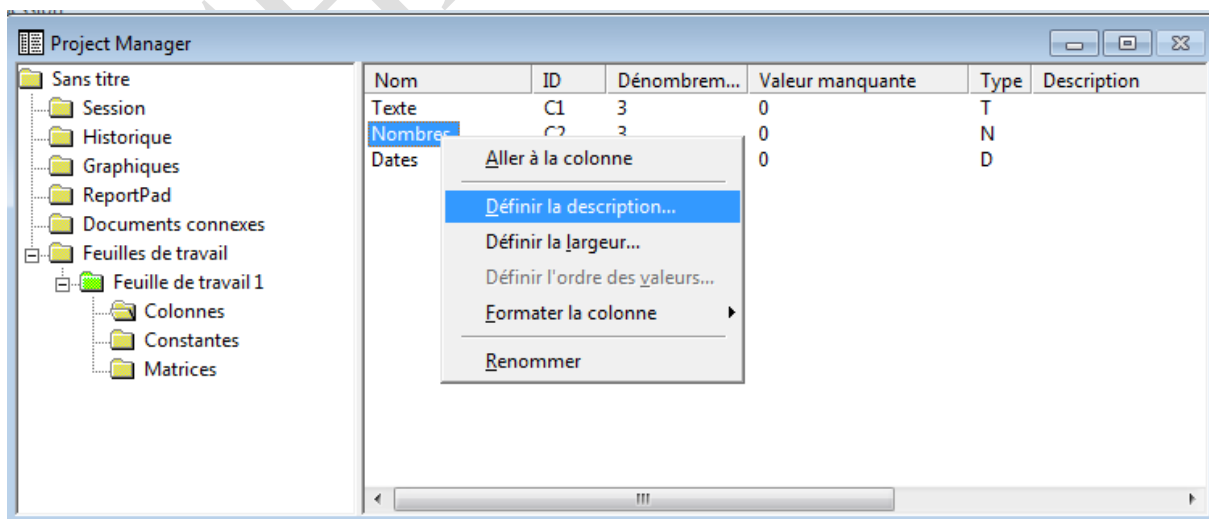
Par exemple si vous saisissez les données et les titres de colonnes suivantes comme exercice vous ne devriez avoir aucune difficulté:



	C1-T	C2	C3-D	C4
	Texte	Nombres	Dates	
1	Texte 1	1	2014-03-07	
2	Texte 2	2	2013-08-12	
3	Texte 3	3	2001-01-01	
4				
5				

Remarquez cependant après la saisie au niveau des colonnes le "-T" qui signifie que la colonne est dorénavant une colonne dédiée à des textes et que le "-D" à des dates. C'est ce que nous appelons le "typage de colonnes".

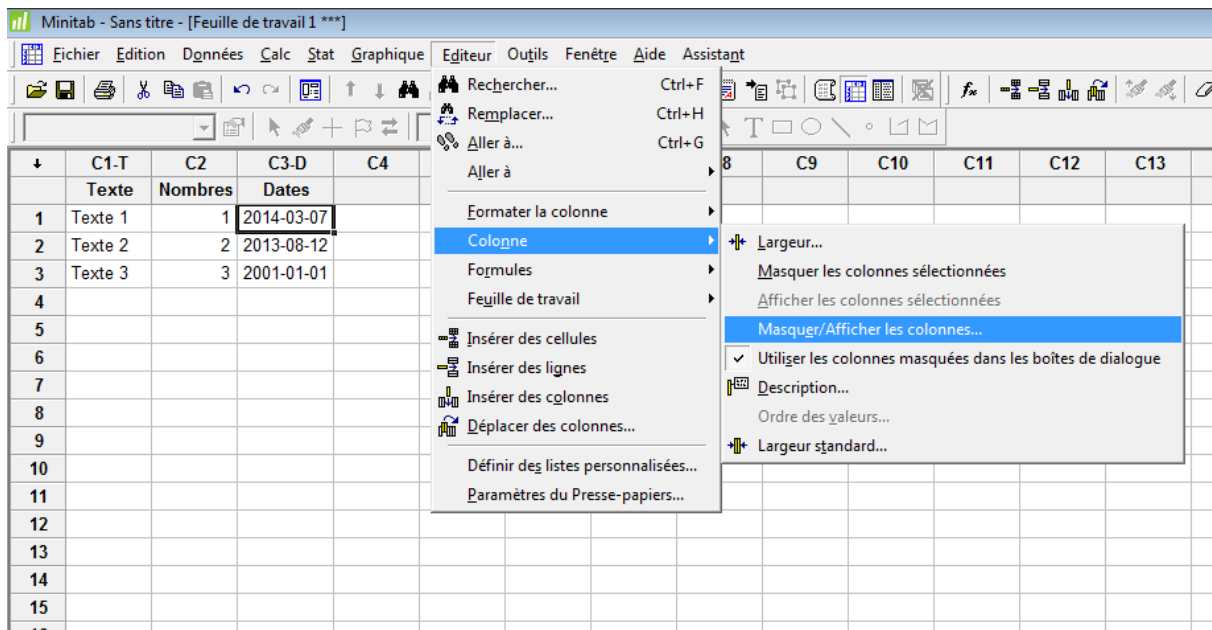
Si nous allons dans le Project Manager nous retrouvons les colonnes et leurs propriétés ainsi que la possibilité de définir une description (très utile!):



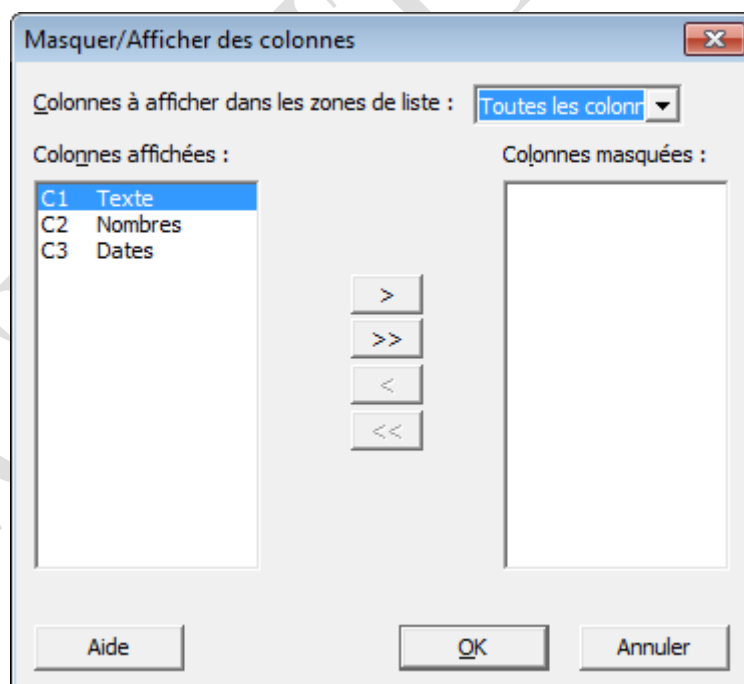
Nom	ID	Dénombre...	Valeur manquante	Type	Description
Texte	C1	3	0	T	
Nombres	C2	3	0	N	
Dates	C3	3	0	D	

## 6.1. Afficher/Masquer des colonnes

Nous pouvons aussi comme dans n'importe quel tableur afficher ou masquer les colonnes en allant dans le menu **Editeur/Masquer/Afficher les colonnes...**:

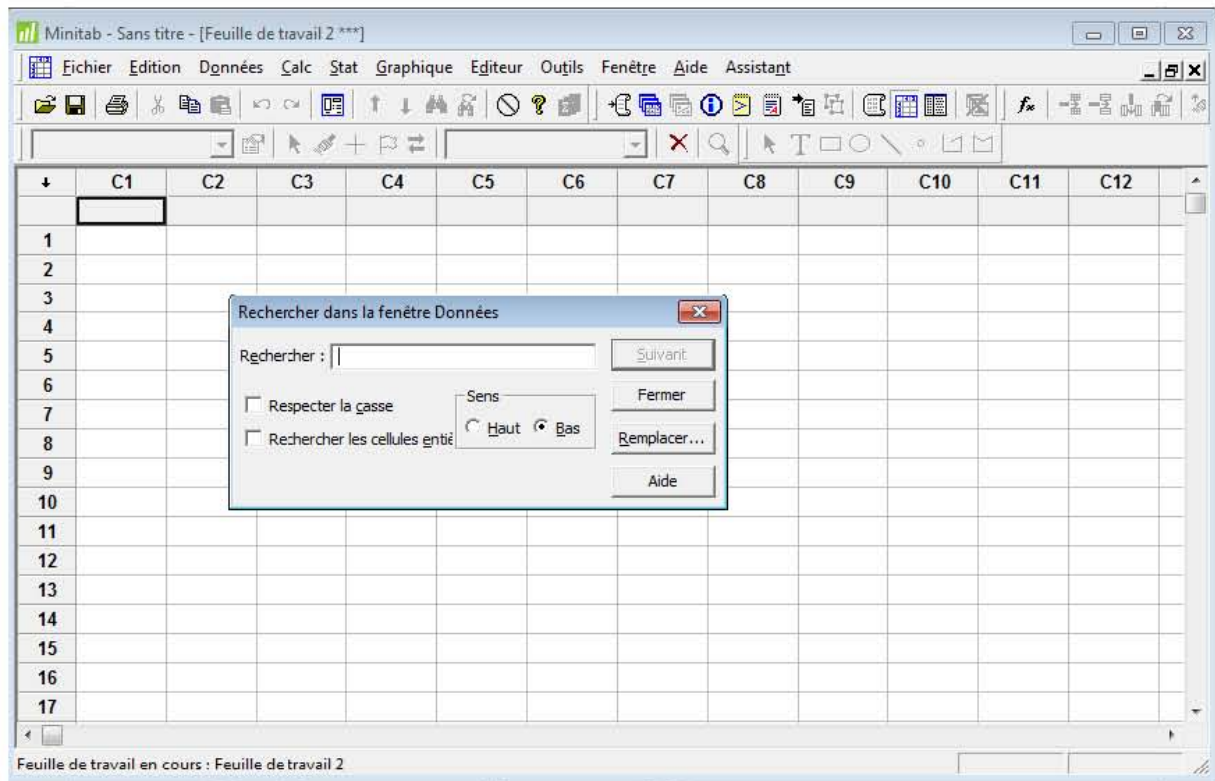


Ce qui nous montre une boîte de dialogue triviale à manipuler:



## 6.2. Rechercher/Remplacer

Bon rien d'extraordinaire ici... le classique Ctrl+F fonctionne aussi pour faire des rechercher/remplacer:

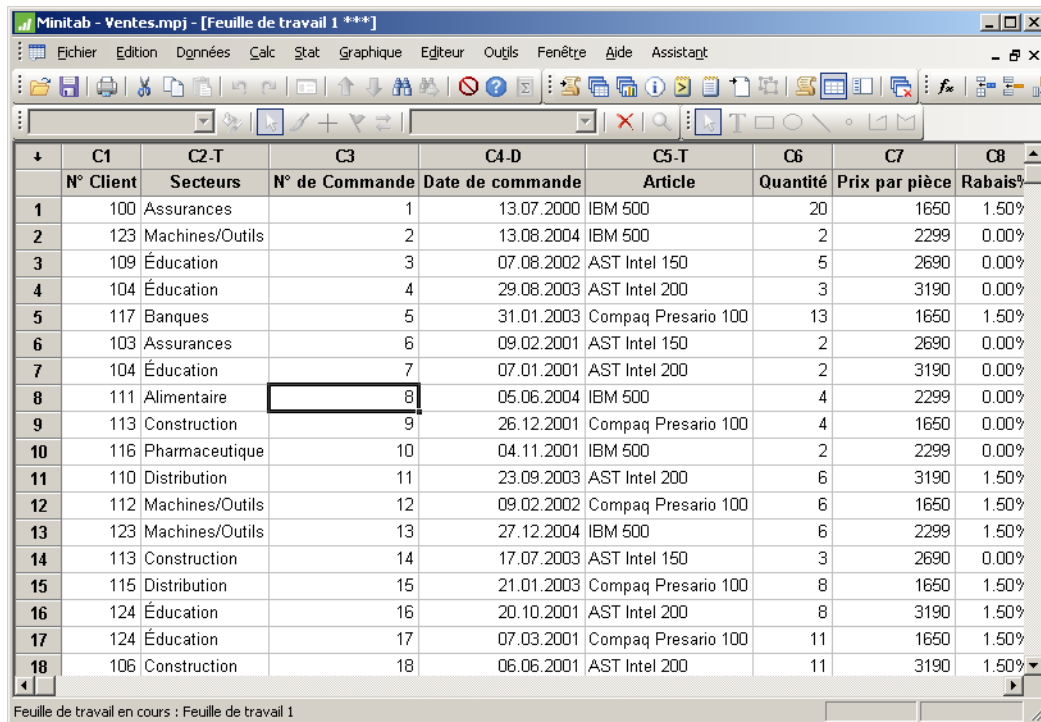


### 6.3. Trier des données

La moindre des choses que l'on puisse dire c'est que Minitab ne fait pas les choses simples au niveau du tri des données...

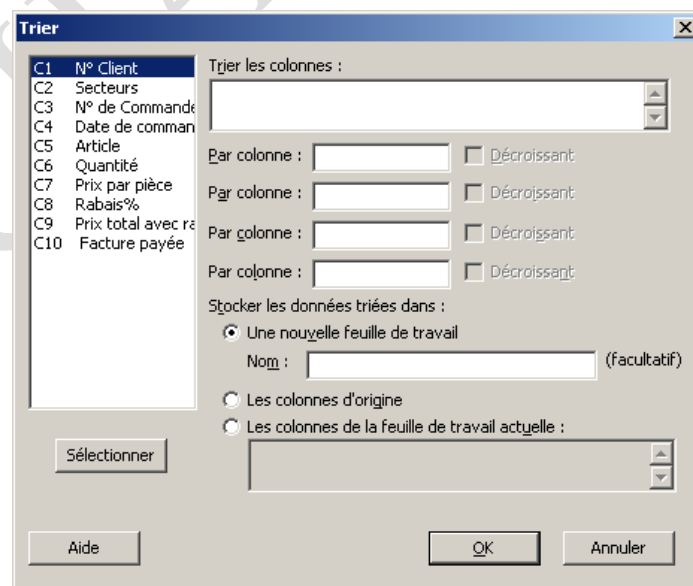
**Remarque:** Le tri ne peut pas être annulé :-)

Voyons cela avec un exemple en prenant comme base de travail notre fichier habituel:



	C1	C2-T	C3	C4-D	C5-T	C6	C7	C8
	N° Client	Secteurs	N° de Commande	Date de commande	Article	Quantité	Prix par pièce	Rabais%
1	100	Assurances	1	13.07.2000	IBM 500	20	1650	1.50%
2	123	Machines/Outils	2	13.08.2004	IBM 500	2	2299	0.00%
3	109	Éducation	3	07.08.2002	AST Intel 150	5	2690	0.00%
4	104	Éducation	4	29.08.2003	AST Intel 200	3	3190	0.00%
5	117	Banques	5	31.01.2003	Compaq Presario 100	13	1650	1.50%
6	103	Assurances	6	09.02.2001	AST Intel 150	2	2690	0.00%
7	104	Éducation	7	07.01.2001	AST Intel 200	2	3190	0.00%
8	111	Alimentaire	8	05.06.2004	IBM 500	4	2299	0.00%
9	113	Construction	9	26.12.2001	Compaq Presario 100	4	1650	0.00%
10	116	Pharmaceutique	10	04.11.2001	IBM 500	2	2299	0.00%
11	110	Distribution	11	23.09.2003	AST Intel 200	6	3190	1.50%
12	112	Machines/Outils	12	09.02.2002	Compaq Presario 100	6	1650	1.50%
13	123	Machines/Outils	13	27.12.2004	IBM 500	6	2299	1.50%
14	113	Construction	14	17.07.2003	AST Intel 150	3	2690	0.00%
15	115	Distribution	15	21.01.2003	Compaq Presario 100	8	1650	1.50%
16	124	Éducation	16	20.10.2001	AST Intel 200	8	3190	1.50%
17	124	Éducation	17	07.03.2001	Compaq Presario 100	11	1650	1.50%
18	106	Construction	18	06.06.2001	AST Intel 200	11	3190	1.50%

Ensuite, nous allons dans le menu **Données/Trier...** pour avoir:



**Trier**

Trier les colonnes :

Par colonne :   **Décroissant**

Par colonne :   **Décroissant**

Par colonne :   **Décroissant**

Par colonne :   **Décroissant**

Stocker les données triées dans :

Une nouvelle feuille de travail

    Nom :  (facultatif)

Les colonnes d'origine

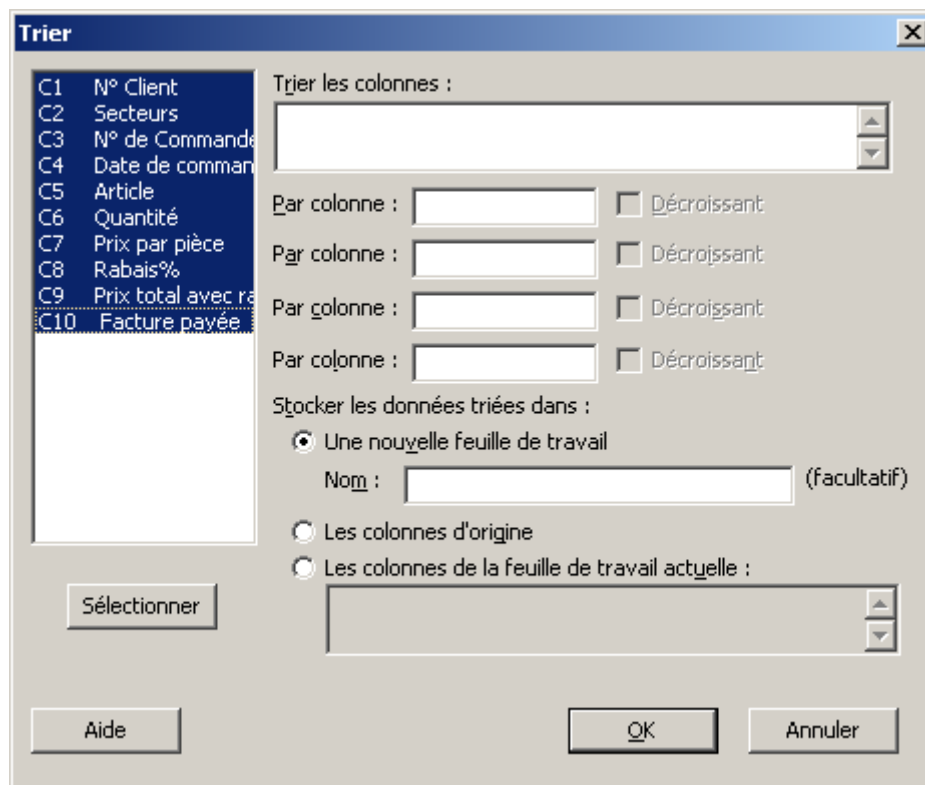
Les colonnes de la feuille de travail actuelle :

Sélectionner

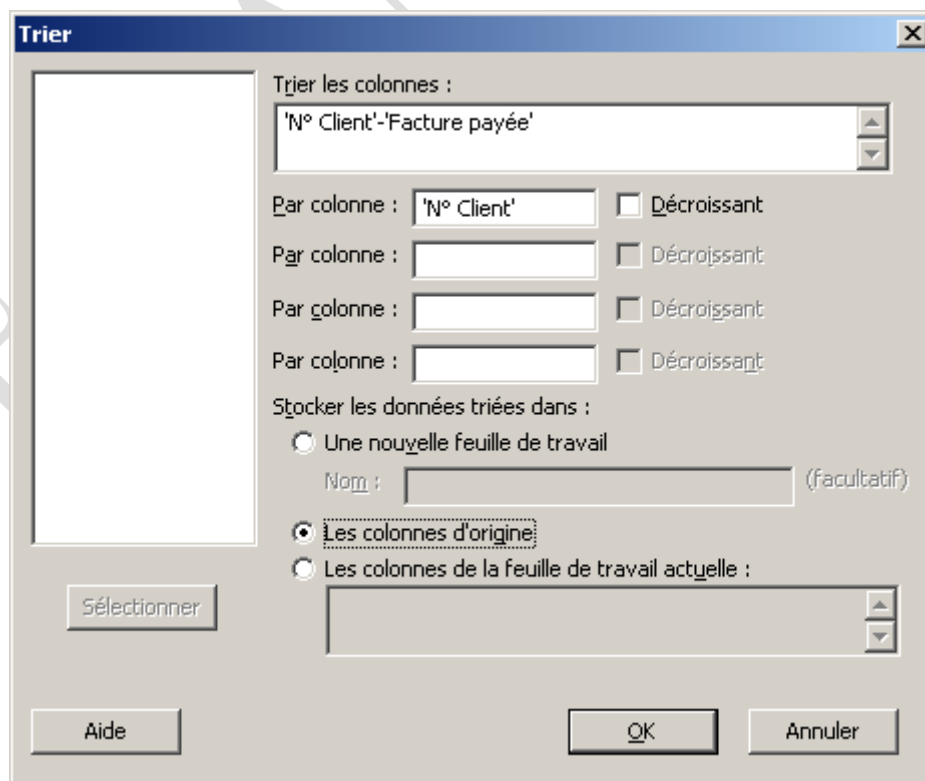
Aide      OK      Annuler

Par défaut Minitab propose donc de faire le futur tri dans un nouvelle feuille.

Nous ce que nous voulons c'est trier toutes les colonnes par rapport à une colonne spécifique tout en restant dans la même feuille. Pour cela nous sélectionnons tous les champs sur la gauche:



Nous cliquons sur le bouton **Sélectionner** ce qui va les ajouter dans le champ **Trier les colonnes** et comme première clé de tri nous prenons les numéros de clients:



et nous obtenons alors:

Minitab - Ventes.mpj - [Feuille de travail 1 \*\*\*]

Échelle: 100%

	C1	C2-T	C3	C4-D	C5-T	C6	C7	C8
	N° Client	Secteurs	N° de Commande	Date de commande	Article	Quantité	Prix par pièce	Rabais%
1	100	Assurances	1	13.07.2000	IBM 500	20	1650	1.50%
2	100	Assurances	23	13.04.2003	IBM 500	3	2299	0.00%
3	100	Assurances	45	10.10.2003	AST Intel 200	4	3190	0.00%
4	100	Assurances	50	10.03.2004	AST Intel 150	13	2690	1.50%
5	101	Construction	19	24.08.2004	Compaq Presario 100	14	1650	1.50%
6	101	Construction	35	16.11.2000	AST Intel 200	6	3190	1.50%
7	101	Construction	41	05.08.2002	AST Intel 200	1	3190	0.00%
8	101	Construction	69	12.10.2002	Compaq Presario 100	13	1650	1.50%
9	101	Construction	71	18.04.2000	IBM 500	5	2299	0.00%
10	101	Construction	76	02.05.2002	AST Intel 200	10	3190	1.50%
11	101	Construction	100	09.11.2002	Compaq Presario 100	5	1650	0.00%
12	102	Machines/Outils	52	06.02.2001	Compaq Presario 100	5	1650	0.00%
13	102	Machines/Outils	90	16.06.2000	IBM 500	7	2299	1.50%
14	102	Machines/Outils	107	12.12.2001	AST Intel 200	4	3190	0.00%
15	103	Assurances	6	09.02.2001	AST Intel 150	2	2690	0.00%
16	103	Assurances	29	10.12.2004	Compaq Presario 100	6	1650	1.50%
17	103	Assurances	55	19.02.2002	IBM 500	7	2299	1.50%
18	103	Assurances	70	06.10.2001	AST Intel 150	12	2690	1.50%

Feuille de travail en cours : Feuille de travail 1

## 6.4. Scinder des données

Comme il n'y a pas de filtres dans Minitab il faut savoir comme dans SPSS scinder un fichier. Considérons notre fichier de ventes habituel:

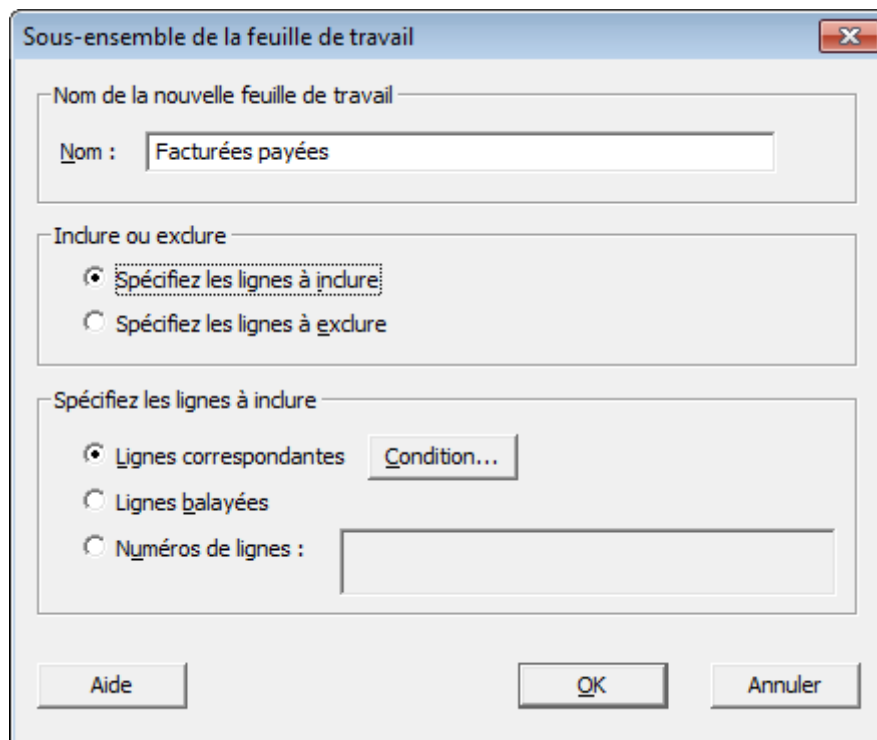
	C1	C2-T	C3	C4-D	C5-T	C6	C7	C8	C9	C10-T	C11	C12	C13
	N° Client	Activité	N° de Commande	Date de commande	Article	Quantité	Prix par pièce	Rabais	Prix total avec rabais	Facture payée			
1	100	Assurances	1	2000-01-03	Compaq Presario 100	12	1650	1.50%	19503.0	Oui			
2	123	Machines/Outils	2	2000-01-03	IBM 500	2	2299	0.00%	4598.0	Oui			
3	109	Éducation	3	2000-01-03	AST Intel 150	5	2690	0.00%	13450.0	Oui			
4	104	Éducation	4	2000-01-03	AST Intel 200	3	3190	0.00%	9570.0	Oui			
5	117	Banques	5	2000-01-04	Compaq Presario 100	13	1650	1.50%	21128.3	Oui			
6	103	Assurances	6	2000-01-04	AST Intel 150	2	2690	0.00%	5380.0	Oui			
7	104	Éducation	7	2000-01-04	AST Intel 200	2	3190	0.00%	6380.0	Oui			
8	111	Alimentaire	8	2000-01-04	IBM 500	4	2299	0.00%	9196.0	Oui			
9	113	Construction	9	2000-01-04	Compaq Presario 100	4	1650	0.00%	6600.0	Oui			
10	116	Pharmaceutique	10	2000-01-04	IBM 500	2	2299	0.00%	4598.0	Oui			
11	110	Distribution	11	2000-01-05	AST Intel 200	6	3190	1.50%	18852.9	Oui			
12	112	Machines/Outils	12	2000-01-05	Compaq Presario 100	6	1650	1.50%	9751.5	Oui			
13	123	Machines/Outils	13	2000-01-05	IBM 500	6	2299	1.50%	13587.1	Oui			
14	113	Construction	14	2000-01-05	AST Intel 150	3	2690	0.00%	8070.0	Oui			
15	115	Distribution	15	2000-01-05	Compaq Presario 100	8	1650	1.50%	13002.0	Oui			
16	124	Éducation	16	2000-01-05	AST Intel 200	8	3190	1.50%	25137.2	Oui			
17	124	Éducation	17	2000-01-05	Compaq Presario 100	11	1650	1.50%	17877.8	Oui			
18	106	Construction	18	2000-01-05	AST Intel 200	11	3190	1.50%	34563.7	Oui			
19	101	Construction	19	2000-01-05	Compaq Presario 100	14	1650	1.50%	22753.5	Non			
20	116	Pharmaceutique	20	2000-01-06	IBM 500	7	2299	1.50%	15851.6	Non			
21	112	Machines/Outils	21	2000-01-06	AST Intel 150	6	2690	1.50%	15897.9	Oui			
22	125	Construction	22	2000-01-06	Compaq Presario 100	23	1650	3.00%	36811.5	Oui			
23	100	Assurances	23	2000-01-06	IBM 500	3	2299	0.00%	6897.0	Oui			
24	125	Construction	24	2000-01-06	AST Intel 200	2	3190	0.00%	6380.0	Oui			
25	104	Éducation	25	2000-01-07	AST Intel 150	12	2690	1.50%	31795.8	Oui			
26	126	Machines/Outils	26	2000-01-07	AST Intel 150	24	2690	3.00%	62623.2	Oui			
27	121	Pharmaceutique	27	2000-01-07	IBM 500	8	2299	1.50%	18116.1	Non			
28	114	Distribution	28	2000-01-07	AST Intel 200	9	3190	1.50%	28279.3	Oui			
29	103	Assurances	29	2000-01-07	Compaq Presario 100	6	1650	1.50%	9751.5	Oui			
30	125	Construction	30	2000-01-07	AST Intel 200	4	3190	0.00%	12760.0	Oui			
31	120	Banques	31	2000-01-10	AST Intel 150	2	2690	0.00%	5380.0	Oui			
32	111	Alimentaire	32	2000-01-10	Compaq Presario 100	16	1650	3.00%	25608.0	Oui			
33	118	Éducation	33	2000-01-10	IBM 500	3	2299	0.00%	6897.0	Oui			
34	127	Alimentaire	34	2000-01-10	IBM 500	7	2299	1.50%	15851.6	Oui			
35	101	Construction	35	2000-01-11	AST Intel 200	6	3190	1.50%	18852.9	Oui			
36	118	Éducation	36	2000-01-11	Compaq Presario 100	5	1650	0.00%	8250.0	Oui			
37	119	Distribution	37	2000-01-11	AST Intel 200	23	3190	3.00%	71168.9	Non			
38	106	Construction	38	2000-01-11	IBM 500	4	2299	0.00%	9196.0	Oui			

Nous allons dans le menu **Diviser la feuille de travail en sous-ensembles...**:

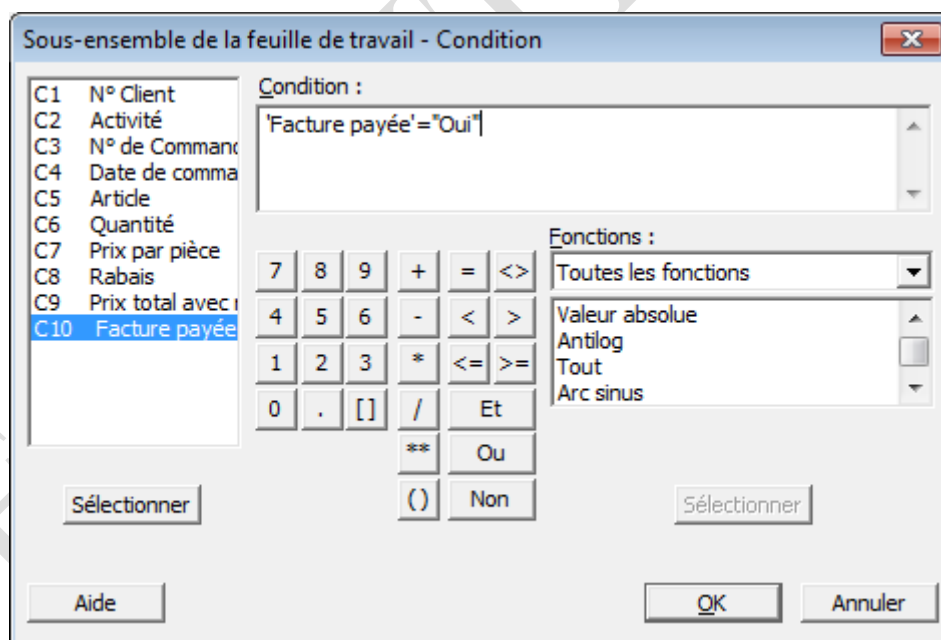
The screenshot shows the Minitab 'Données' menu with the following options:

- Diviser la feuille de travail en sous-ensembles... (highlighted)
- Fractionner la feuille de travail...
- Fusionner les feuilles de travail...
- Copier
- Désempiler les colonnes...
- Empiler
- Transposer les colonnes...
- Trier...
- Rang...
- Supprimer les lignes...
- Effacer les variables...
- Coder
- Changer le type des données
- Extraire de date/heure
- Concaténer...
- Afficher les données...

Il vient alors:

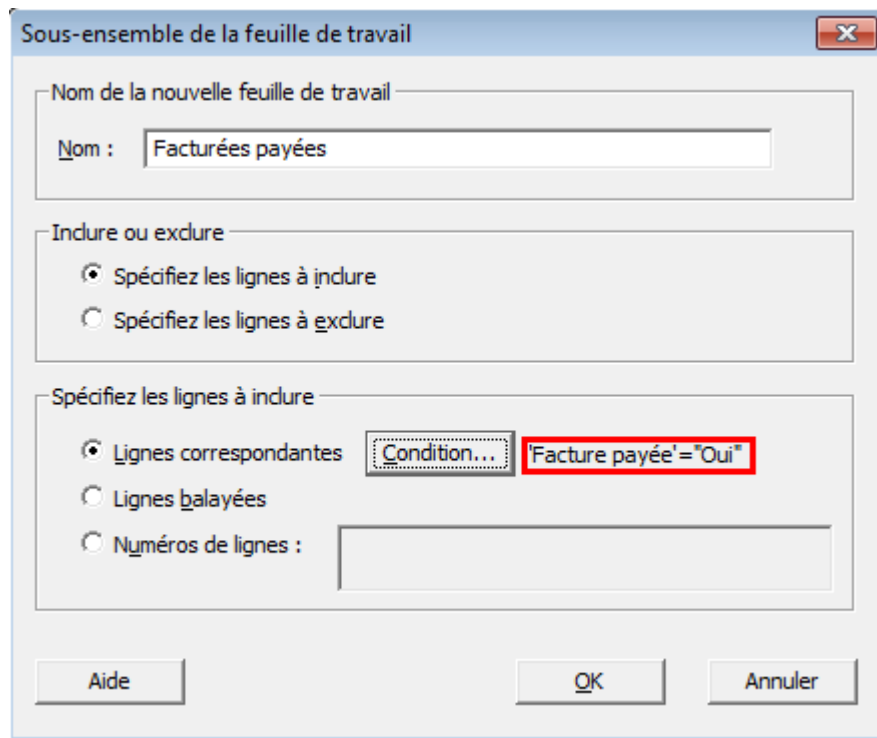


Nous cliquons sur **Condition...** pour y mettre:



Nous validons par **OK** ce qui nous ramène à:





Et si nous validons, nous obtenons bien une feuille avec uniquement des données ayant les factures à l'état "Oui":

	C1	C2-T	C3	C4-D	C5-T	C6	C7	C8	C9	C10-T
	N° Client	Activité	N° de Commande	Date de commande	Article	Quantité	Prix par pièce	Rabais	Prix total avec rabais	Facture payée
1	100	Assurances	1	2000-01-03	Compaq Presario 100	12	1650	1.50%	19503.0	Oui
2	123	Machines/Outils	2	2000-01-03	IBM 500	2	2299	0.00%	4598.0	Oui
3	109	Éducation	3	2000-01-03	AST Intel 150	5	2690	0.00%	13450.0	Oui
4	104	Éducation	4	2000-01-03	AST Intel 200	3	3190	0.00%	9570.0	Oui
5	117	Banques	5	2000-01-04	Compaq Presario 100	13	1650	1.50%	21128.3	Oui
6	103	Assurances	6	2000-01-04	AST Intel 150	2	2690	0.00%	5380.0	Oui
7	104	Éducation	7	2000-01-04	AST Intel 200	2	3190	0.00%	6380.0	Oui
8	111	Alimentaire	8	2000-01-04	IBM 500	4	2299	0.00%	9196.0	Oui
9	113	Construction	9	2000-01-04	Compaq Presario 100	4	1650	0.00%	6600.0	Oui
10	116	Pharmaceutique	10	2000-01-04	IBM 500	2	2299	0.00%	4598.0	Oui
11	110	Distribution	11	2000-01-05	AST Intel 200	6	3190	1.50%	18852.9	Oui
12	112	Machines/Outils	12	2000-01-05	Compaq Presario 100	6	1650	1.50%	9751.5	Oui
13	123	Machines/Outils	13	2000-01-05	IBM 500	6	2299	1.50%	13587.1	Oui
14	113	Construction	14	2000-01-05	AST Intel 150	3	2690	0.00%	8070.0	Oui
15	115	Distribution	15	2000-01-05	Compaq Presario 100	8	1650	1.50%	13002.0	Oui
16	124	Éducation	16	2000-01-05	AST Intel 200	8	3190	1.50%	25137.2	Oui
17	124	Éducation	17	2000-01-05	Compaq Presario 100	11	1650	1.50%	17877.8	Oui
18	106	Construction	18	2000-01-05	AST Intel 200	11	3190	1.50%	34563.7	Oui
19	112	Machines/Outils	21	2000-01-06	AST Intel 150	6	2690	1.50%	15897.9	Oui
20	125	Construction	22	2000-01-06	Compaq Presario 100	23	1650	3.00%	36811.5	Oui
21	100	Assurances	23	2000-01-06	IBM 500	3	2299	0.00%	6897.0	Oui
22	125	Construction	24	2000-01-06	AST Intel 200	2	3190	0.00%	6380.0	Oui
23	104	Éducation	25	2000-01-07	AST Intel 150	12	2690	1.50%	31795.8	Oui
24	126	Machines/Outils	26	2000-01-07	AST Intel 150	24	2690	3.00%	62623.2	Oui
25	114	Distribution	28	2000-01-07	AST Intel 200	9	3190	1.50%	28279.3	Oui
26	103	Assurances	29	2000-01-07	Compaq Presario 100	6	1650	1.50%	9751.5	Oui
27	125	Construction	30	2000-01-07	AST Intel 200	4	3190	0.00%	12760.0	Oui
28	120	Banques	31	2000-01-10	AST Intel 150	2	2690	0.00%	5380.0	Oui
29	111	Alimentaire	32	2000-01-10	Compaq Presario 100	16	1650	3.00%	25608.0	Oui
30	118	Éducation	33	2000-01-10	IBM 500	3	2299	0.00%	6897.0	Oui

## 6.5. Coder des valeurs

Une manipulation courante est de recoder des variables ordinales ou nominales. Certes, nous pouvons le faire avec le Rechercher/Remplacer mais quand nous savons que cela est répétitif, c'est laborieux!

Toujours avec notre fichier de ventes original, pour un exemple empirique, nous allons dans le menu **Données/Coder/Texte en numérique...**:

	C1	C5-T	C6	C7	C8	C9	C10-T
	N° Client	Article	Quantité	Prix par pièce	Rabais	Prix total avec rabais	Facture payée
1	100	Compaq Presario 100	12	1650	1.50%	19503.0	Oui
2	123	IBM 500	2	2299	0.00%	4598.0	Oui
3	109	AST Intel 150	5	2690	0.00%	13450.0	Oui
4	104	AST Intel 200	3	3190	0.00%	9570.0	Oui
5	117	Compaq Presario 100	13	1650	1.50%	21128.3	Oui
6	103	AST Intel 150	2	2690	0.00%	5380.0	Oui
7	104	AST Intel 200	2	3190	0.00%	6380.0	Oui
8	111	IBM 500	4	2299	0.00%	9196.0	Oui
9	113			1650	0.00%	6600.0	Oui
10	116			2299	0.00%	4598.0	Oui
11	110			3190	1.50%	18852.9	Oui
12	112			1650	1.50%	9751.5	Oui
13	123			2299	1.50%	13587.1	Oui
14	113			2690	0.00%	8070.0	Oui
15	115	Distribution	15	2000-0	1.50%	13002.0	Oui
16	124	Éducation	16	2000-0	1.50%	25137.2	Oui
17	124	Éducation	17	2000-0	1.50%	17877.8	Oui
18	106	Construction	18	2000-0	1.50%	34563.7	Oui
19	101	Construction	19	2000-0	1.50%	22753.5	Non
20	116	Pharmaceutique	20	2000-0	1.50%	15851.6	Non
21	112	Machines/Outils	21	2000-01-06	1.50%	15897.9	Oui
22	125	Construction	22	2000-01-06	3.00%	36811.5	Oui
23	100	Assurances	23	2000-01-06	0.00%	6897.0	Oui
24	125	Construction	24	2000-01-06	0.00%	6380.0	Oui
25	104	Éducation	25	2000-01-07	1.50%	31795.8	Oui
26	126	Machines/Outils	26	2000-01-07	3.00%	62623.2	Oui
27	121	Pharmaceutique	27	2000-01-07	1.50%	18116.1	Non
28	114	Distribution	28	2000-01-07	1.50%	28279.3	Oui

Il vient alors:

**Coder - Texte-numérique**

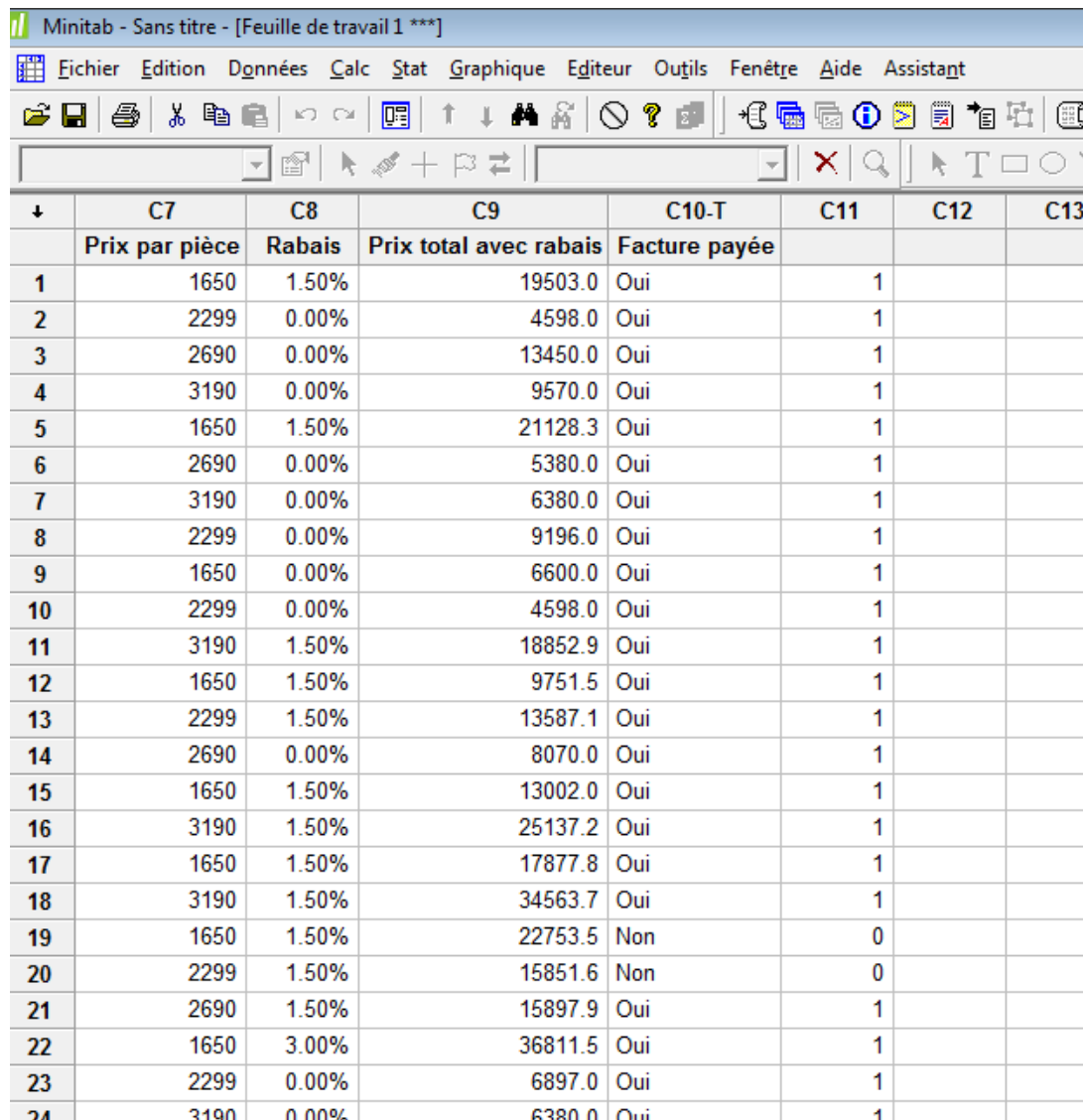
Coder les données depuis les colonnes :

Stocker les données codées dans les colonnes :

Valeurs initiales (par exemple, rouge "bleu clair") :      Nouvelles :

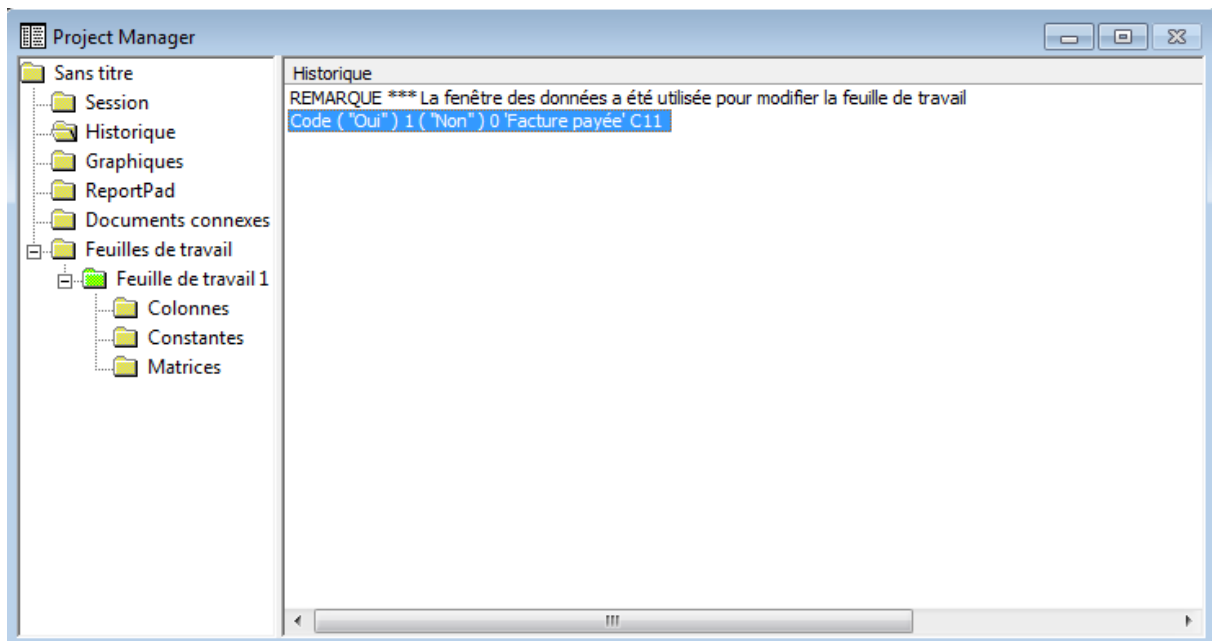
<input type="text" value="Oui"/>	<input type="text" value="1"/>
<input type="text" value="Non"/>	<input type="text" value="0"/>
<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>
<input type="text"/>	<input type="text"/>

Ce qui donnera:

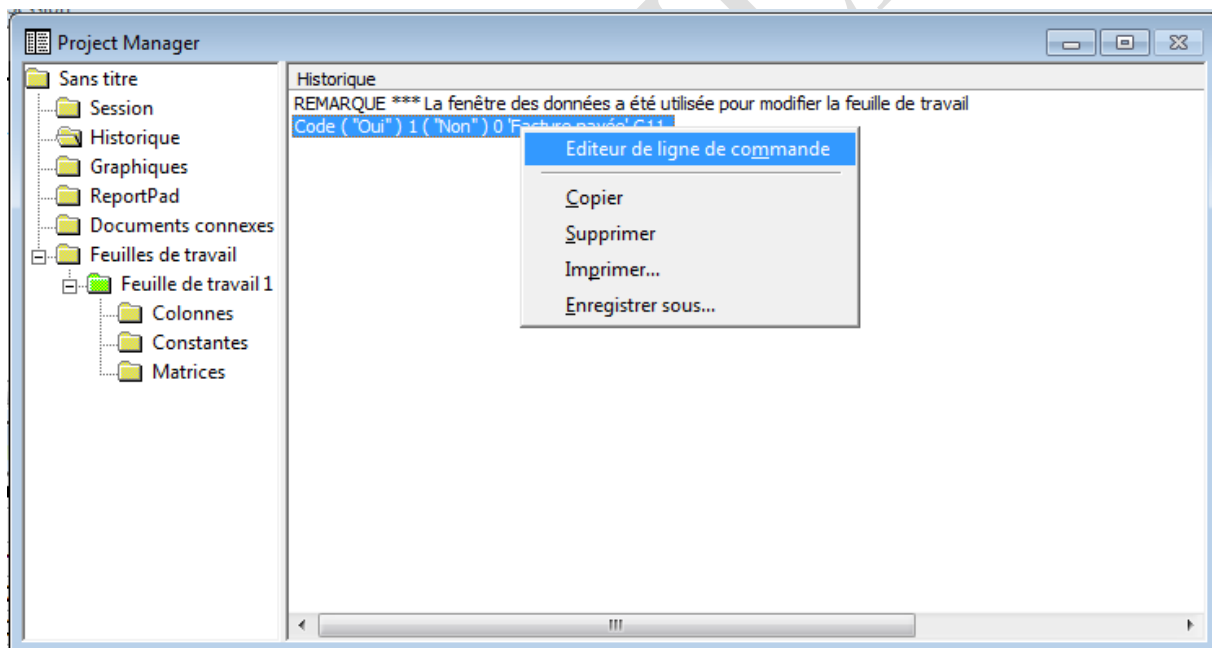


	C7	C8	C9	C10-T	C11	C12	C13
	Prix par pièce	Rabais	Prix total avec rabais	Facture payée			
1	1650	1.50%	19503.0	Oui	1		
2	2299	0.00%	4598.0	Oui	1		
3	2690	0.00%	13450.0	Oui	1		
4	3190	0.00%	9570.0	Oui	1		
5	1650	1.50%	21128.3	Oui	1		
6	2690	0.00%	5380.0	Oui	1		
7	3190	0.00%	6380.0	Oui	1		
8	2299	0.00%	9196.0	Oui	1		
9	1650	0.00%	6600.0	Oui	1		
10	2299	0.00%	4598.0	Oui	1		
11	3190	1.50%	18852.9	Oui	1		
12	1650	1.50%	9751.5	Oui	1		
13	2299	1.50%	13587.1	Oui	1		
14	2690	0.00%	8070.0	Oui	1		
15	1650	1.50%	13002.0	Oui	1		
16	3190	1.50%	25137.2	Oui	1		
17	1650	1.50%	17877.8	Oui	1		
18	3190	1.50%	34563.7	Oui	1		
19	1650	1.50%	22753.5	Non	0		
20	2299	1.50%	15851.6	Non	0		
21	2690	1.50%	15897.9	Oui	1		
22	1650	3.00%	36811.5	Oui	1		
23	2299	0.00%	6897.0	Oui	1		
24	3190	0.00%	6380.0	Oui	1		

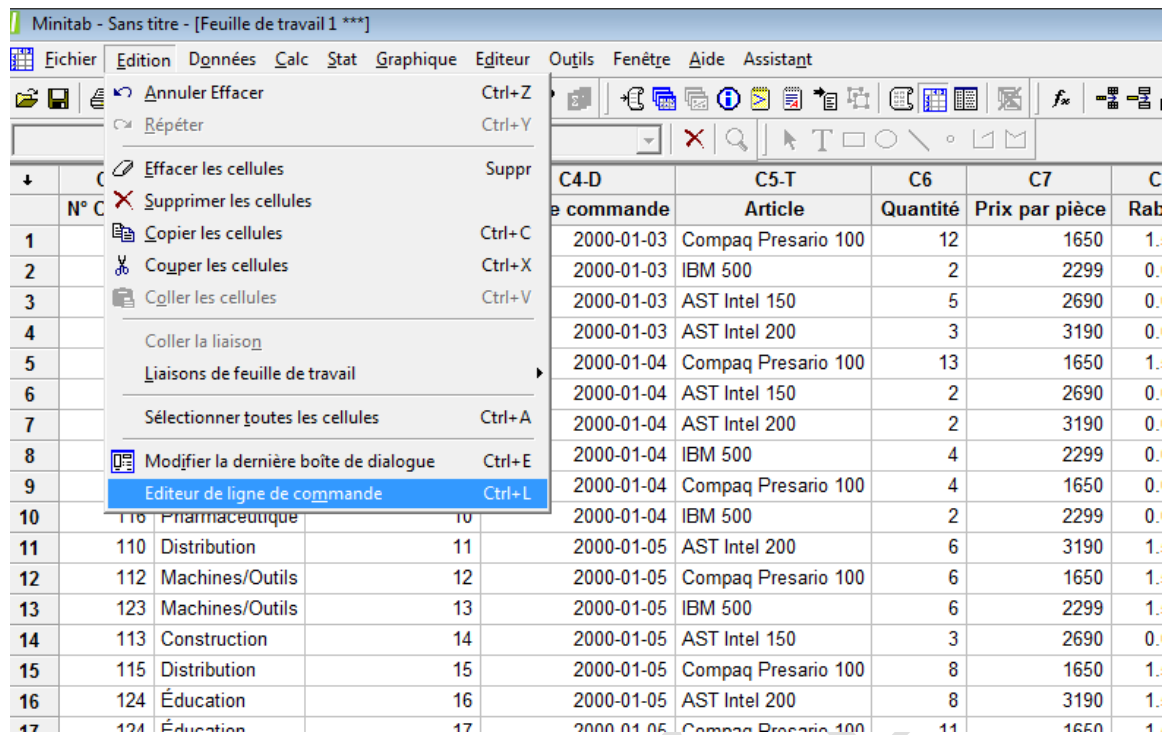
L'utilité de cette approche c'est que nous pouvons vite la reproduire à l'avenir dans des cas plus complexes. Effectivement, si nous ouvrons le Project Manager et allons dans l'**Historique** nous retrouvons le code y relatif:



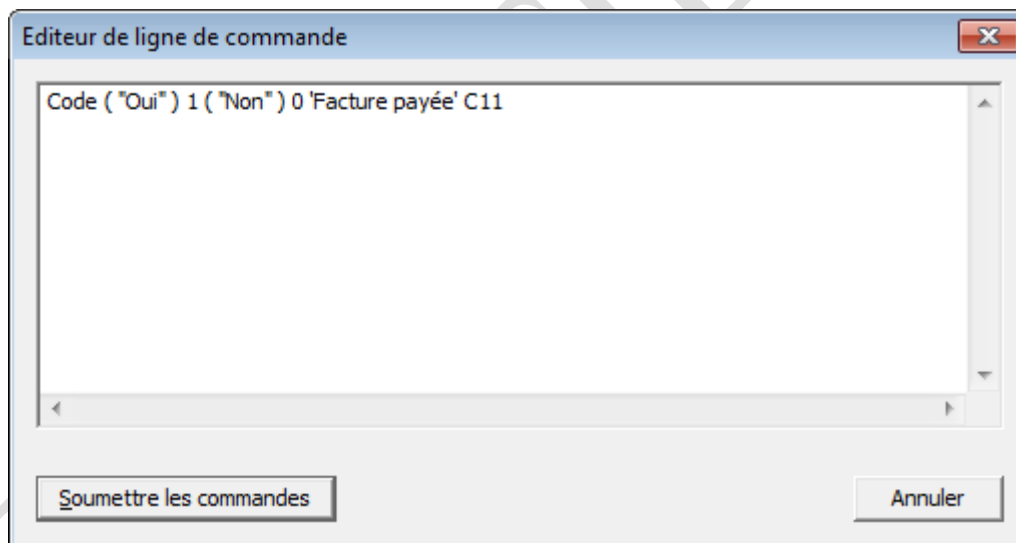
Si nous faisons un clic droit dessus, nous pouvons le **Copier** pour le stocker quelque part (dans le **ReportPad** typiquement):



Et plus tard, nous pourrons simplement aller dans l'**Editeur de ligne de commande**:



Pour y coller le code et rapidement le ré-exécuter:

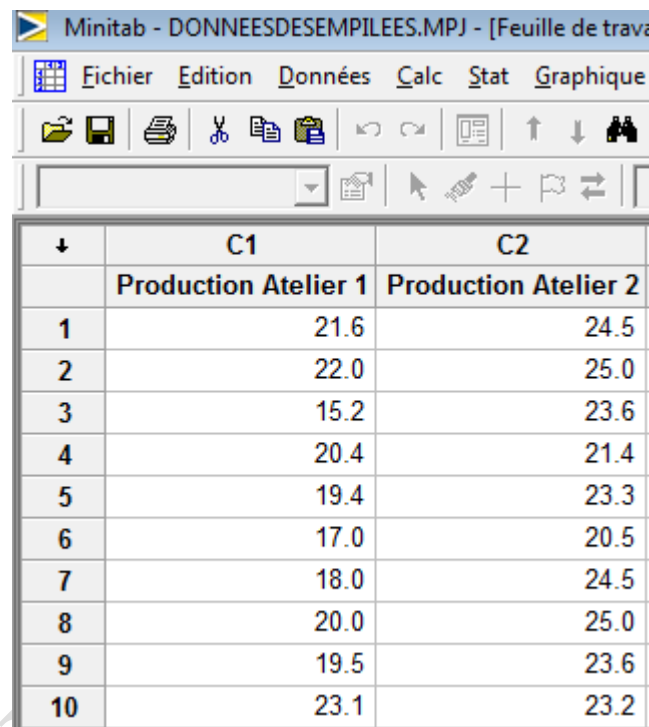


## 7. Manipulations élémentaires

### 7.1. Exercice 1.: Calculs arithmétiques élémentaires

Minitab® Statistical Software 15.1.1.0

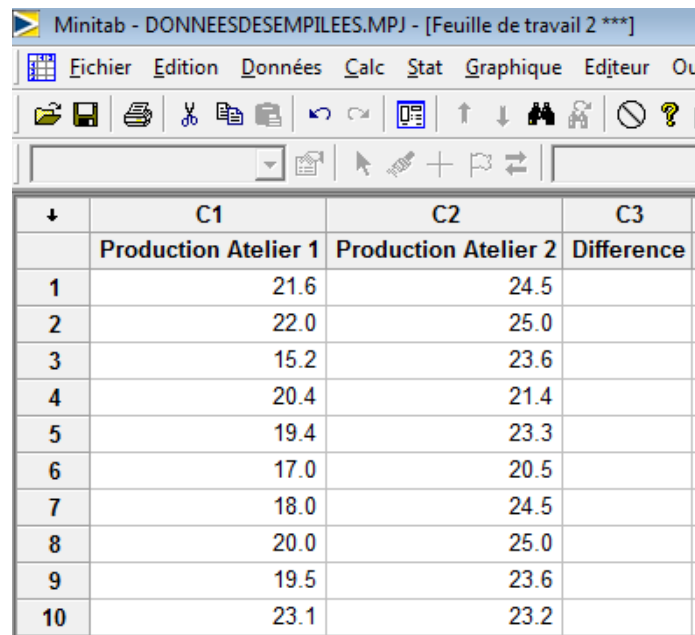
Il peut arriver dans Minitab® Statistical Software que nous ayons parfois besoin de faire des calculs arithmétiques élémentaires. Voyons un exemple. Ouvrez le fichier *CalculSimple.mpj*:



The screenshot shows the Minitab software interface with a data table. The table has two columns, C1 and C2, and 10 rows of data. The column headers are 'Production Atelier 1' and 'Production Atelier 2'. The data values are as follows:

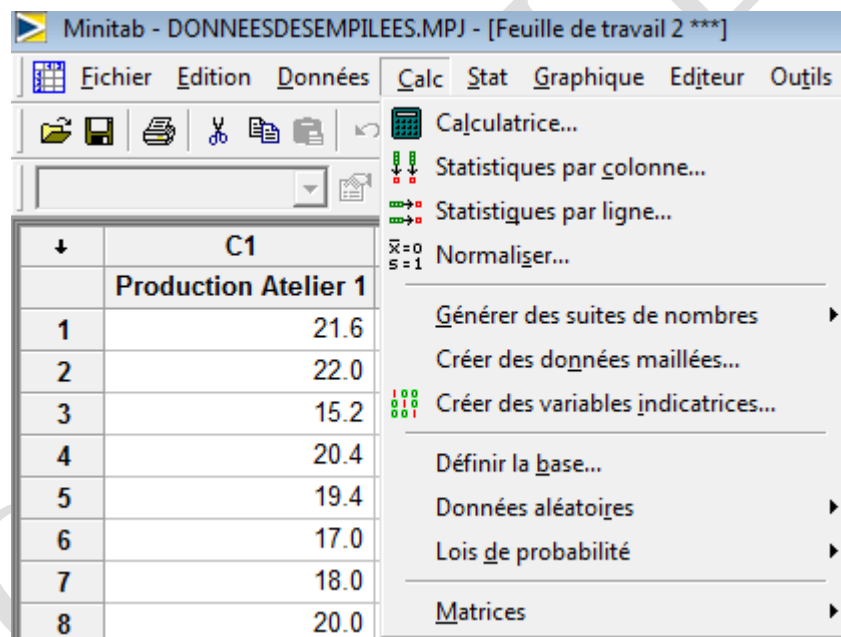
	C1	C2
	Production Atelier 1	Production Atelier 2
1	21.6	24.5
2	22.0	25.0
3	15.2	23.6
4	20.4	21.4
5	19.4	23.3
6	17.0	20.5
7	18.0	24.5
8	20.0	25.0
9	19.5	23.6
10	23.1	23.2

Nous souhaiterions simplement une colonne qui contient la différence des deux. Pour se faire, nous créons d'abord une nouvelle colonne servant à cet effet:



	C1	C2	C3
	Production Atelier 1	Production Atelier 2	Difference
1	21.6	24.5	
2	22.0	25.0	
3	15.2	23.6	
4	20.4	21.4	
5	19.4	23.3	
6	17.0	20.5	
7	18.0	24.5	
8	20.0	25.0	
9	19.5	23.6	
10	23.1	23.2	

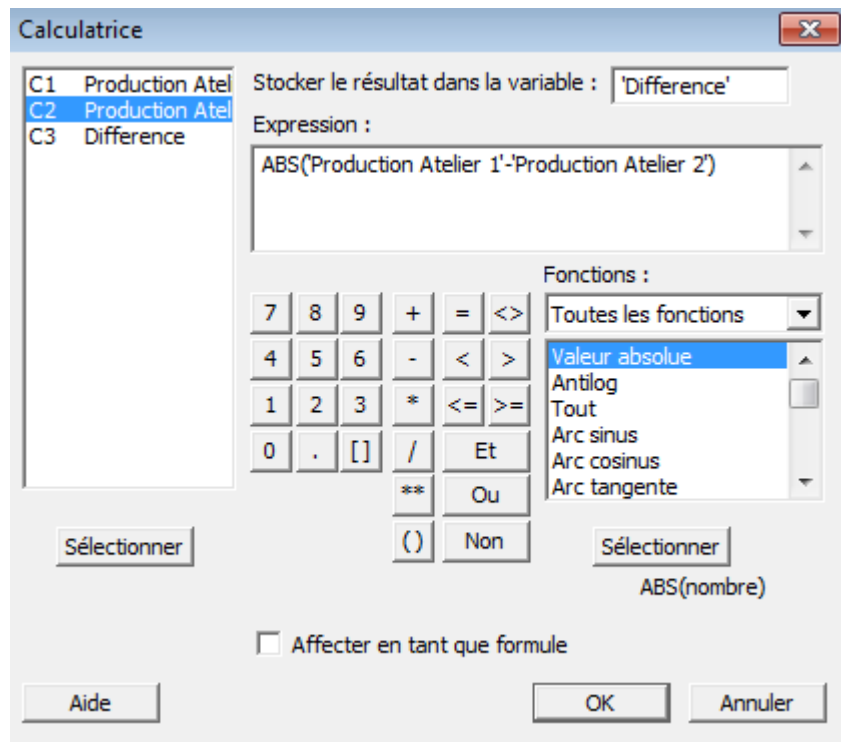
Ensuite, nous allons dans le menu **Calc/Calculatrice...**:



	C1
	Production Atelier 1
1	21.6
2	22.0
3	15.2
4	20.4
5	19.4
6	17.0
7	18.0
8	20.0

- Calculatrice...
- Statistiques par colonne...
- Statistiques par ligne...
- Normaliser...
- Générer des suites de nombres
- Créer des données maillées...
- Créer des variables indicatrices...
- Définir la base...
- Données aléatoires
- Lois de probabilité
- Matrices

pour avoir et écrire:



Nous validons par **OK** pour obtenir:

↓	C1	C2	C3
	Production Atelier 1	Production Atelier 2	Difference
1	21.6	24.5	2.9
2	22.0	25.0	3.0
3	15.2	23.6	8.4
4	20.4	21.4	1.0
5	19.4	23.3	3.9
6	17.0	20.5	3.5
7	18.0	24.5	6.5
8	20.0	25.0	5.0
9	19.5	23.6	4.1
10	23.1	23.2	0.1

Si vous changez une des valeurs des colonnes C1 ou C2, vous verrez que la Différence ne se recalcule pas. Ceci est normal (dans le domaine des statistiques les valeurs mesurées ne sont pas censées changer...). Mais si vous souhaitez quand même une formule dynamique, il vous suffira dans la boîte de la calculatrice, de cocher **Affecter en tant que formule**. La colonne apparaîtra alors avec un symbole spécial dans le coin supérieur droit:



C3
Difference
2.9
3.0
8.4
1.0
3.9
3.5
6.5
5.0
4.1
0.1

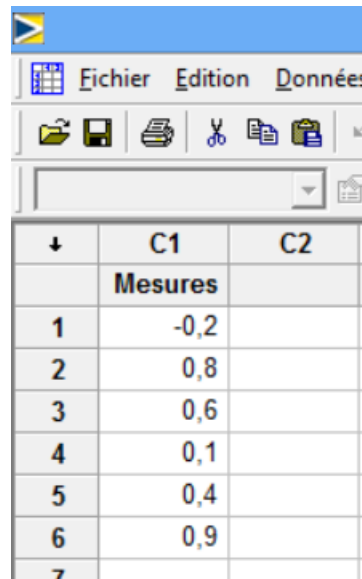
ÉCHANTILLON

## 7.2. Exercice 2.: Valeurs propres

Minitab® Statistical Software 15.1.1.0

Dans certaines situations (régressions, plans d'expériences et autres...) il peut être désirable de normaliser (centre-réduire) des données.

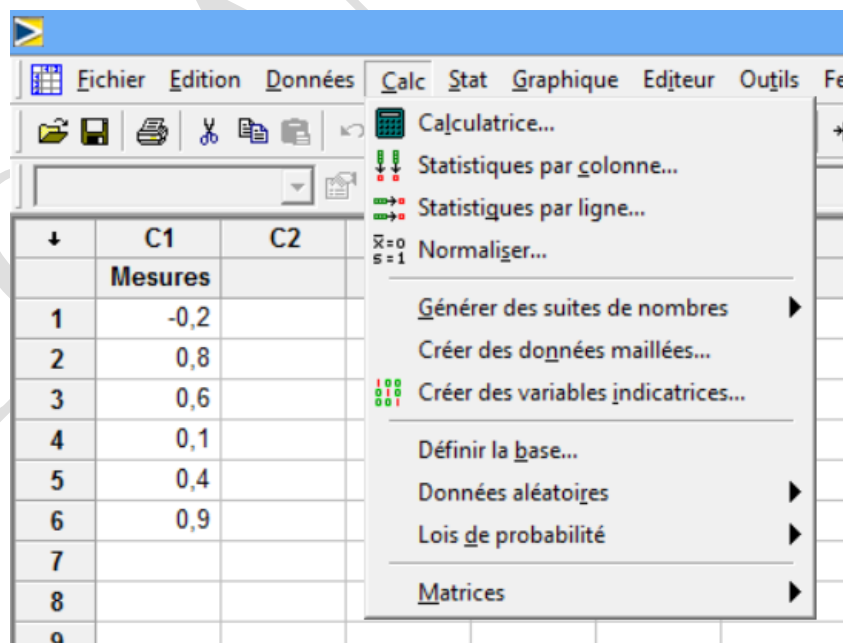
Pour voir comment faire cela dans Minitab, considérons les données suivantes:



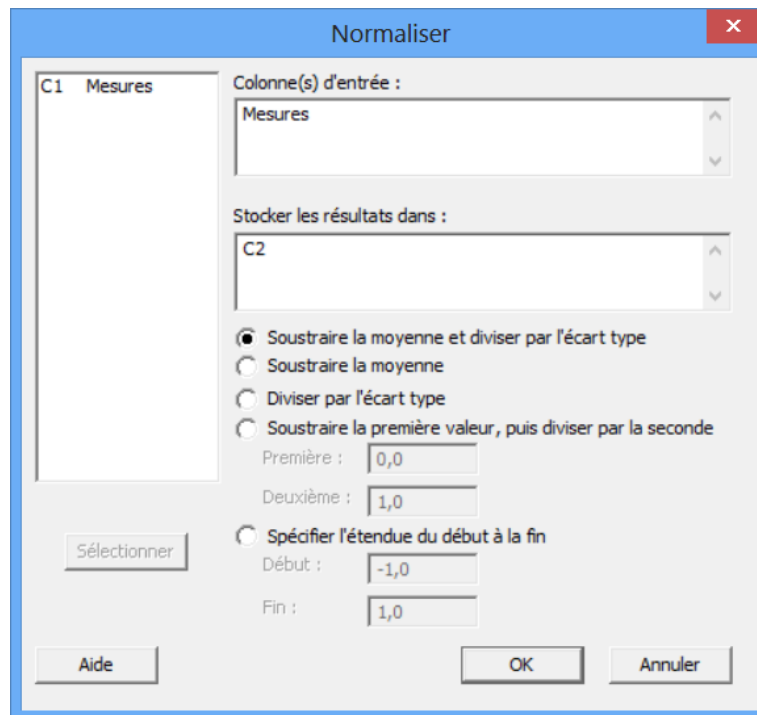
A screenshot of the Minitab data table showing two columns, C1 and C2. The data is as follows:

	C1	C2
	Mesures	
1	-0,2	
2	0,8	
3	0,6	
4	0,1	
5	0,4	
6	0,9	
7		

Ensuite, nous allons dans le menu **Calc/Normaliser...**:



où nous mettons:



Ce qui donne:

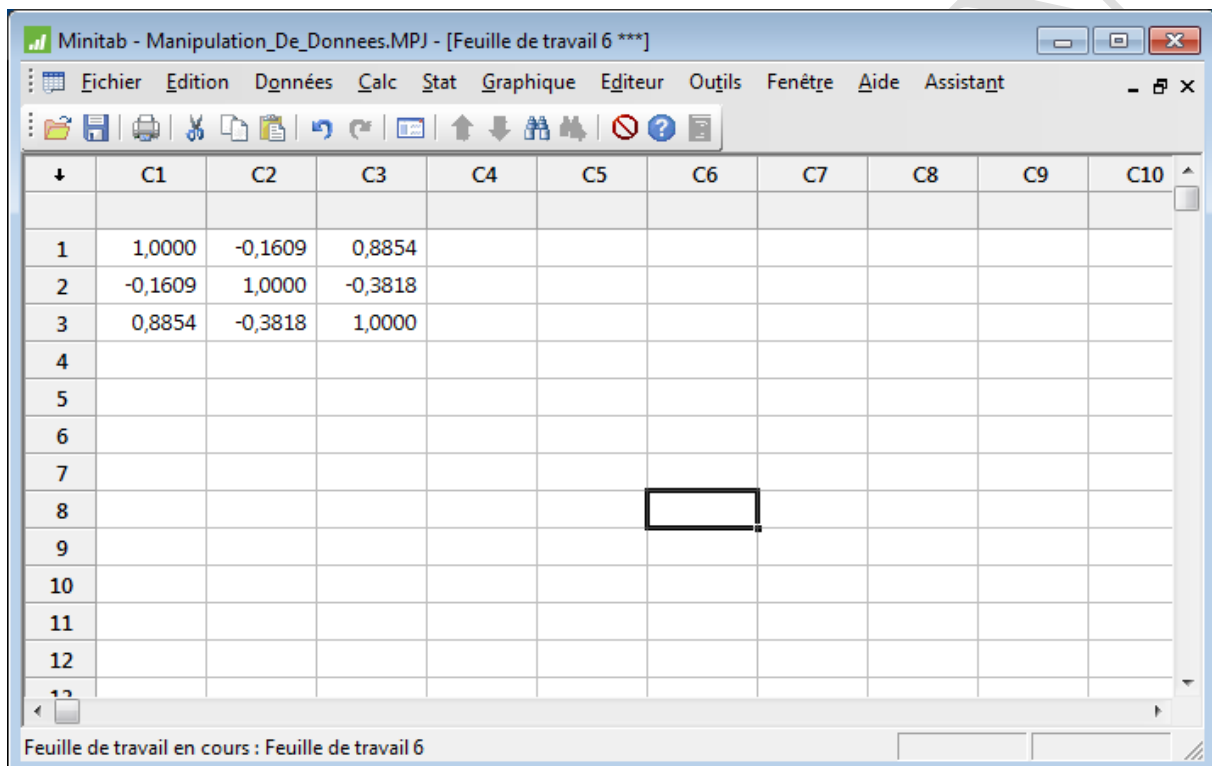
	C1	C2	C3
	<b>Mesures</b>		
1	-0,2	-1,49834	
2	0,8	0,86746	
3	0,6	0,39430	
4	0,1	-0,78860	
5	0,4	-0,07886	
6	0,9	1,10404	
7			

### 7.3. Exercice 3.: Valeurs propres

Minitab® Statistical Software 17.1.2

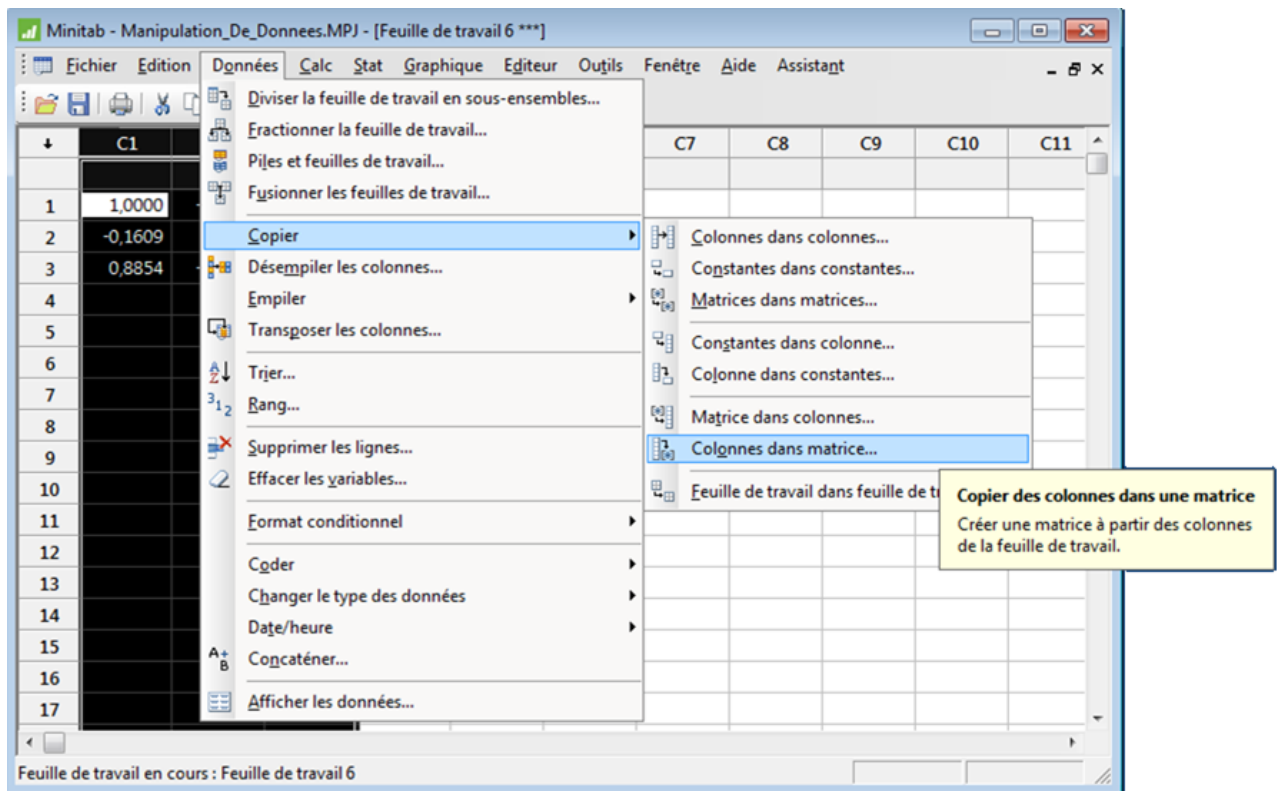
Avec Microsoft Excel nous savons qu'il est pénible de calculer des valeurs propres ce qui est pourtant essentiel dans de nombreux domaines des statistiques. Nous allons voir que ceci est beaucoup plus simple avec Minitab.

D'abord nous saisissons les composantes de la matrice de corrélation telle que calculée à la main dans le cours théorique lors de notre étude de l'analyse en composantes principales:

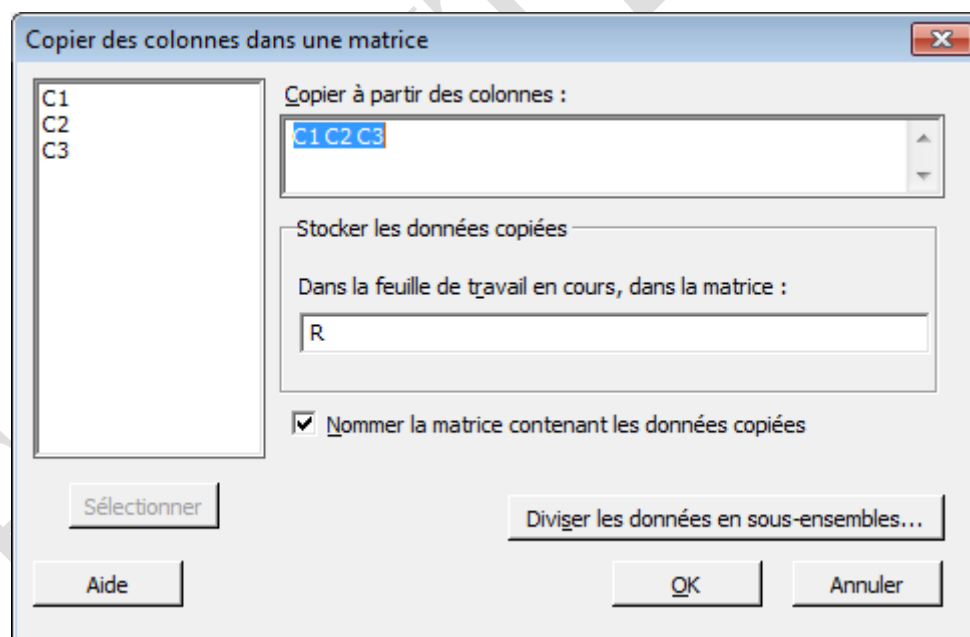


	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
1	1,0000	-0,1609	0,8854							
2	-0,1609	1,0000	-0,3818							
3	0,8854	-0,3818	1,0000							
4										
5										
6										
7										
8										
9										
10										
11										
12										
13										

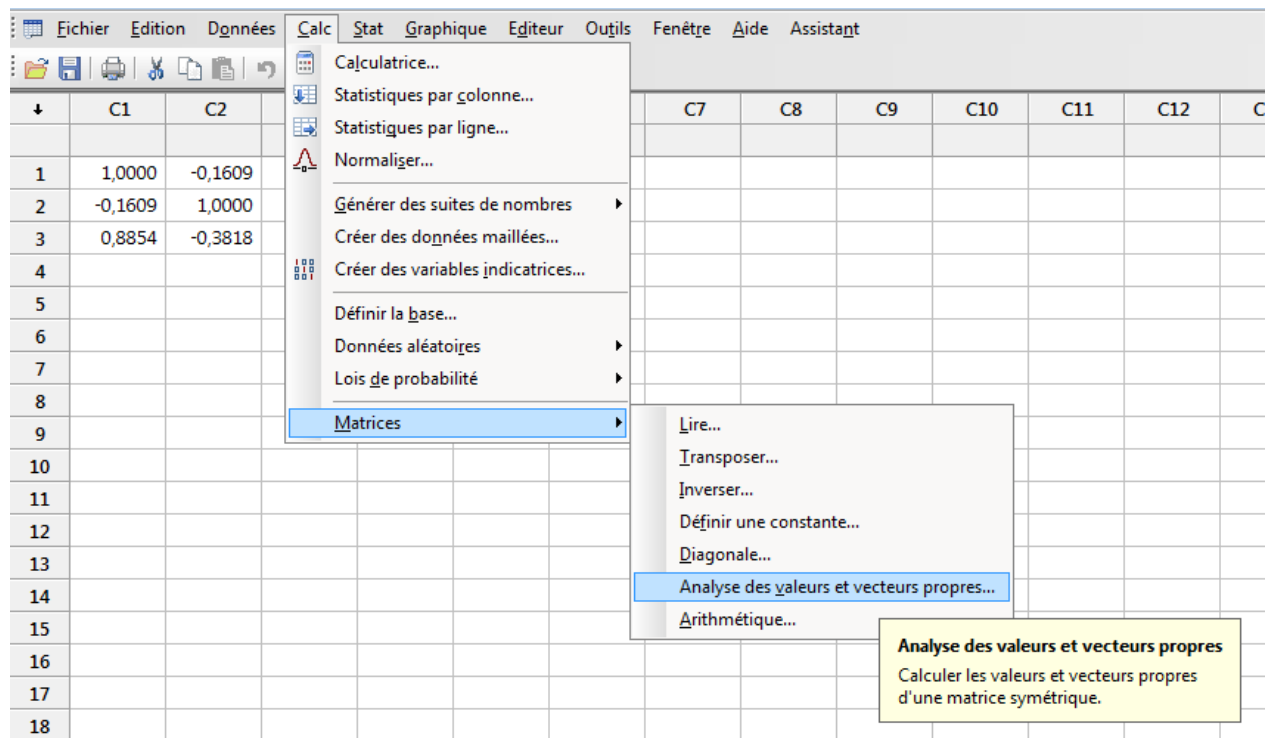
Une fois ceci fait, nous sélectionnons les trois colonnes et nous allons dans le menu **Données/Copier/Colonnes dans matrice...**:



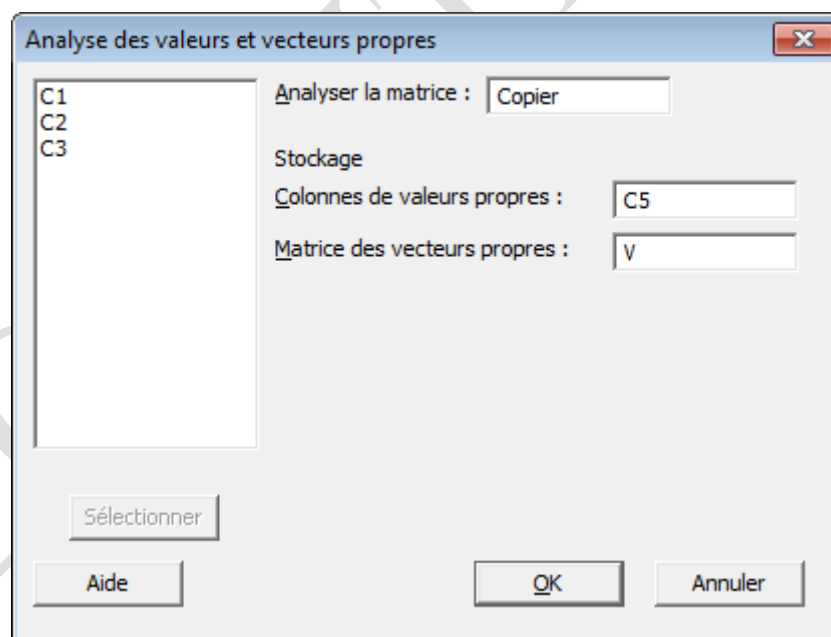
Dans la boîte de dialogue qui apparaît, nous mettons:



Nous validons par **OK** et ensuite nous allons dans le menu **Calc/Matrices/Analyse des valeurs et vecteurs propres**:



Dans la boîte de dialogue qui apparaît nous sélectionnons notre matrice (qui apparaîtra alors sous le nom *Copier*) et choisissons la colonne de stockage et éventuellement (optionnel) un nom pour la matrice des vecteurs propres:



Nous obtenons alors:

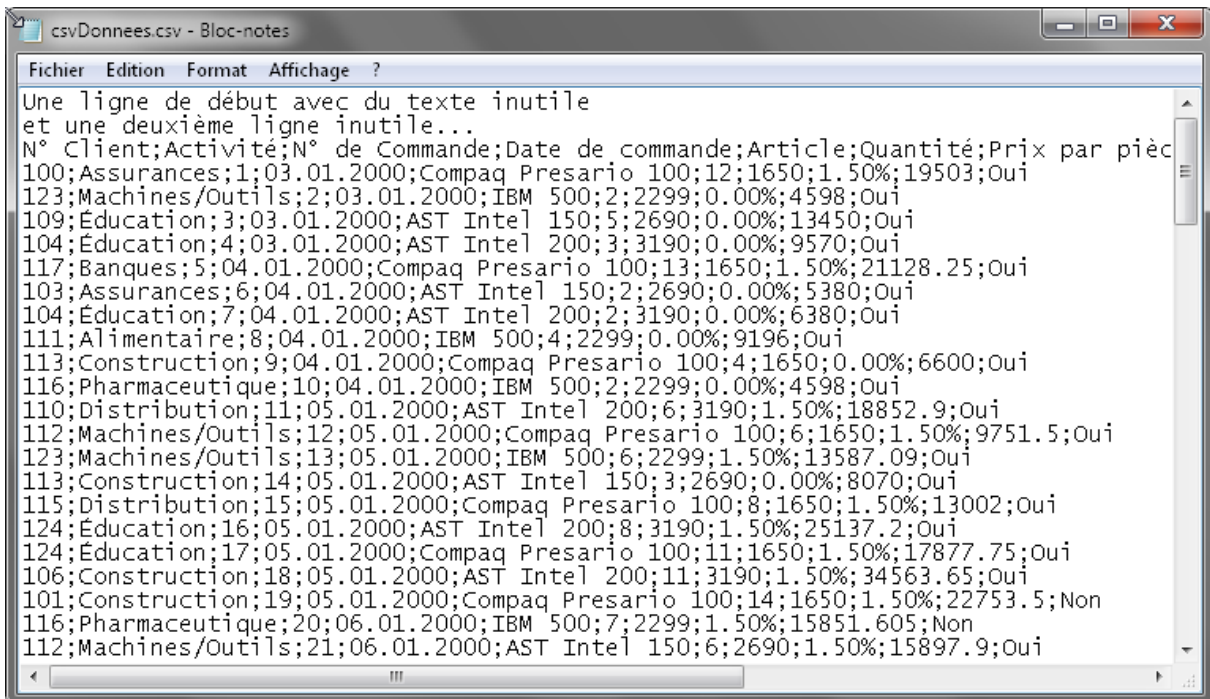
	C1	C2	C3	C4	C5	C6	C7	C8	C9
1	1,0000	-0,1609	0,8854		2,03014				
2	-0,1609	1,0000	-0,3818		0,88455				
3	0,8854	-0,3818	1,0000		0,08531				
4									
5									
6									
7									
8									
9									
10									
11									

Ce qui est bien conforme!

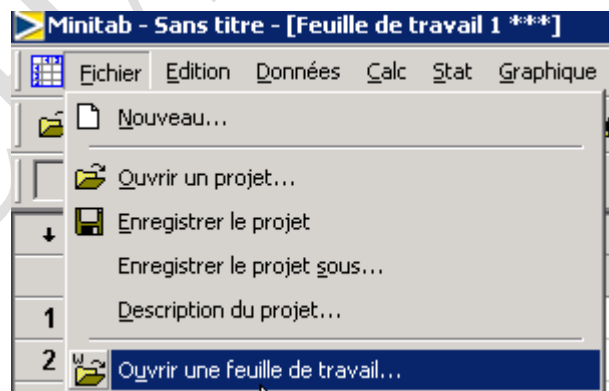
## 7.4. Exercice 4.: Importer/Exporter des données \*.csv

Minitab® Statistical Software 15.1.1.0

Même si jusqu'à aujourd'hui je n'ai jamais vu qui que ce soit utiliser l'outil d'import de Minitab mais passer toujours d'abord par Microsoft Excel pour faire un copier/coller ensuite, voici un exemple d'import de données avec le fichier suivant (qui contient exprès un piège):

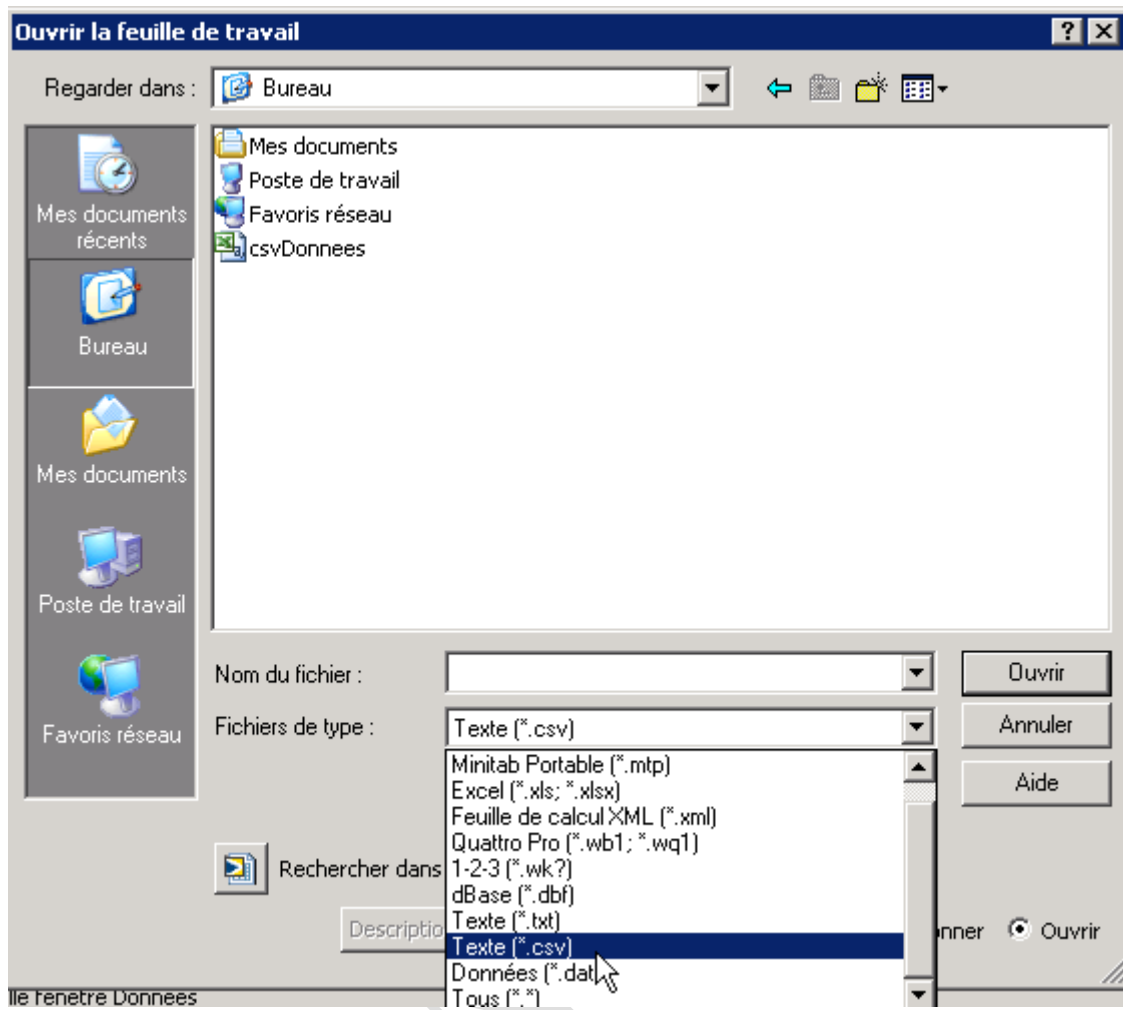


Donc dans Minitab on va **Fichier/Ouvrir une feuille de travail...**:

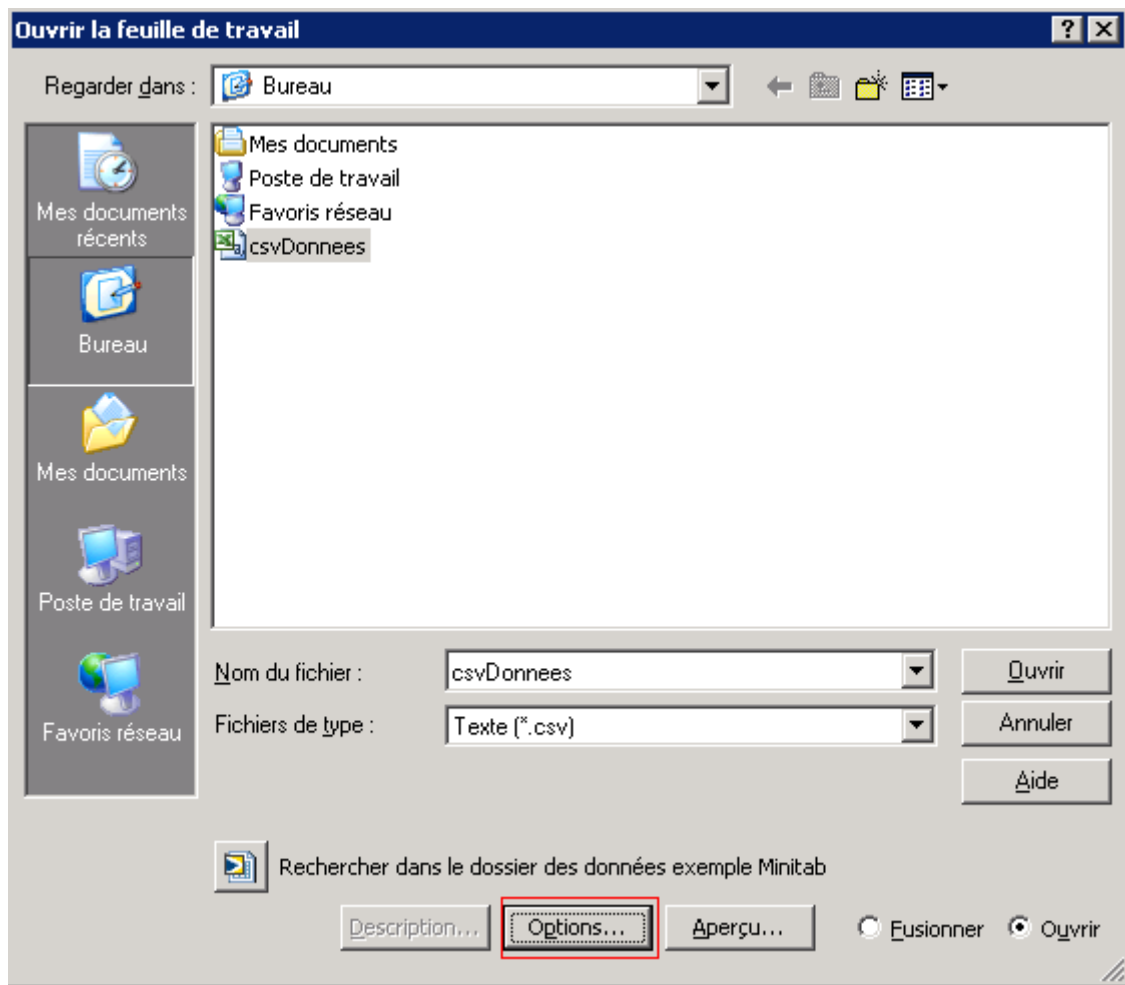


on sélectionne ensuite le format adéquat:

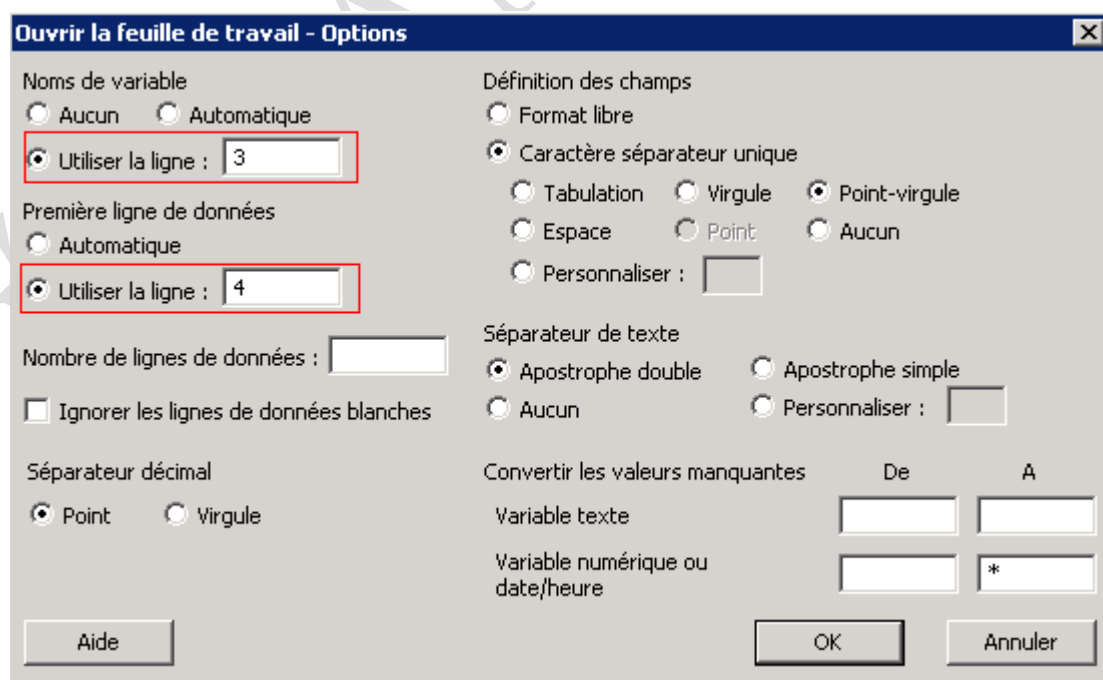




Ensuite, on clique sur **Options** à cause du piège dans le fichier \*.csv:



et on prend:

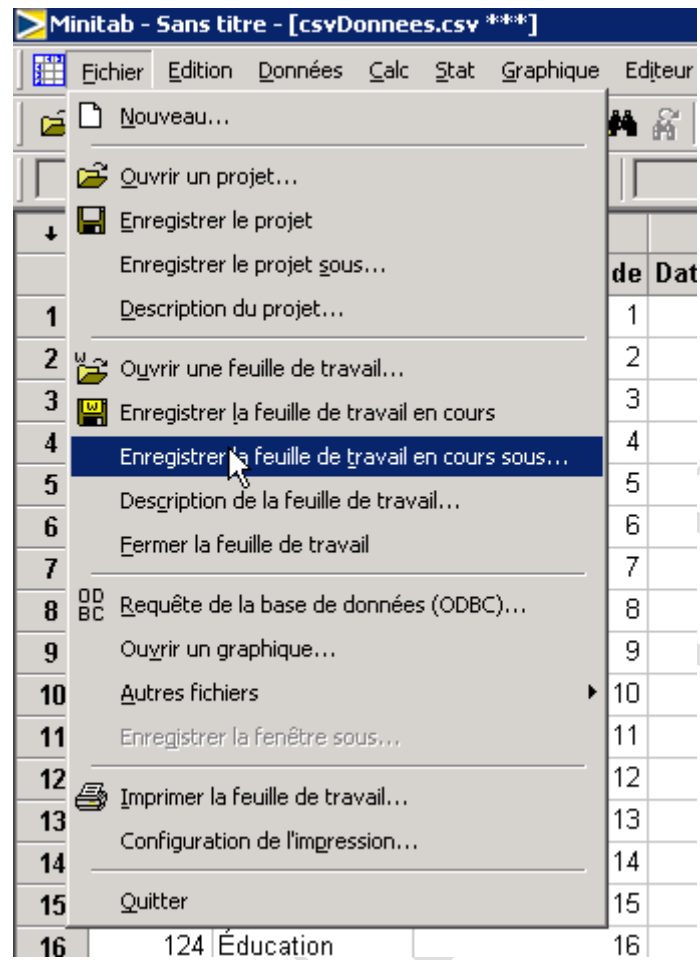


on valide le tout:

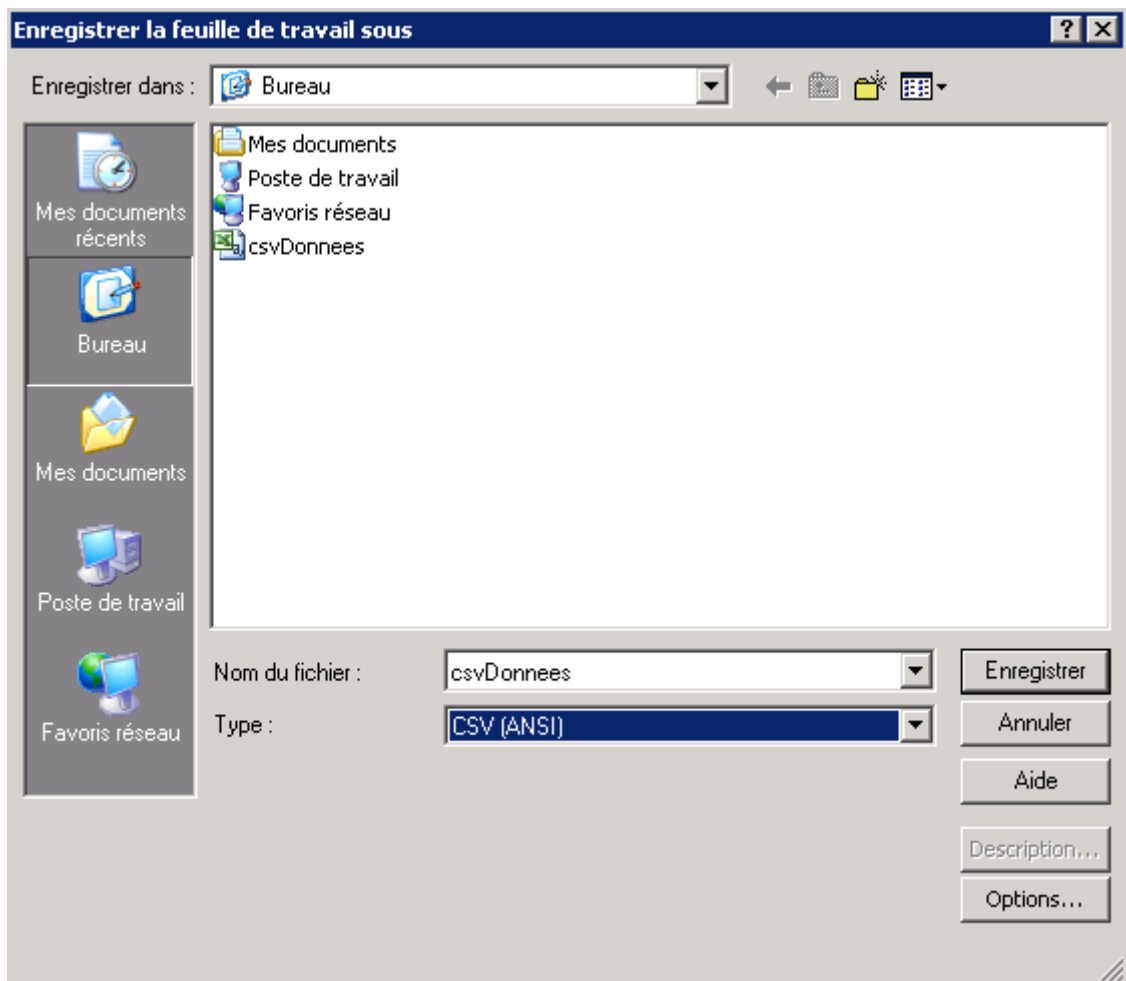
	C1	C2-T	C3	C4-D	C5-T	C6	C7	C8	C9
	N° Client	Activité	N° de Commande	Date de commande	Article	Quantité	Prix par pièce	Rabais%	Prix total avec rabais
1	100	Assurances	1	03.01.2000	Compaq Presario 100	12	1650	1.50%	19503.0
2	123	Machines/Outils	2	03.01.2000	IBM 500	2	2299	0.00%	4598.0
3	109	Éducation	3	03.01.2000	AST Intel 150	5	2690	0.00%	13450.0
4	104	Éducation	4	03.01.2000	AST Intel 200	3	3190	0.00%	9570.0
5	117	Banques	5	04.01.2000	Compaq Presario 100	13	1650	1.50%	21128.3
6	103	Assurances	6	04.01.2000	AST Intel 150	2	2690	0.00%	5380.0
7	104	Éducation	7	04.01.2000	AST Intel 200	2	3190	0.00%	6380.0
8	111	Alimentaire	8	04.01.2000	IBM 500	4	2299	0.00%	9196.0
9	113	Construction	9	04.01.2000	Compaq Presario 100	4	1650	0.00%	6600.0
10	116	Pharmaceutique	10	04.01.2000	IBM 500	2	2299	0.00%	4598.0
11	110	Distribution	11	05.01.2000	AST Intel 200	6	3190	1.50%	18852.9
12	112	Machines/Outils	12	05.01.2000	Compaq Presario 100	6	1650	1.50%	9751.5
13	123	Machines/Outils	13	05.01.2000	IBM 500	6	2299	1.50%	13587.1
14	113	Construction	14	05.01.2000	AST Intel 150	3	2690	0.00%	8070.0
15	115	Distribution	15	05.01.2000	Compaq Presario 100	8	1650	1.50%	13002.0
16	124	Éducation	16	05.01.2000	AST Intel 200	8	3190	1.50%	25137.2
17	124	Éducation	17	05.01.2000	Compaq Presario 100	11	1650	1.50%	17877.8
18	106	Construction	18	05.01.2000	AST Intel 200	11	3190	1.50%	34563.7
19	101	Construction	19	05.01.2000	Compaq Presario 100	14	1650	1.50%	22753.5
20	116	Pharmaceutique	20	06.01.2000	IBM 500	7	2299	1.50%	15851.6
21	112	Machines/Outils	21	06.01.2000	AST Intel 150	6	2690	1.50%	15897.9

et voilà...

Ensuite, pour export, nous allons dans **Fichier/Enregistrer la feuille de travail en cours sous...**



on reprend du CSV ANSI:



et voilà le résultat:

csyDonnees - Bloc-notes  
 Fichier Edition Format Affichage ?


N° Client	Activité	N° de commande	Date de commande	Article	Quantité
1100	Assurances	03.01.2000	Compaq Presario 100	12	1650
1123	Machines/outils	03.01.2000	IBM 500 2	0.00%	4598
1109	Education	03.01.2000	AST Intel 150	2690	0.00%
1104	Education	03.01.2000	AST Intel 200	3190	0.00%
1117	Banques 5	04.01.2000	Compaq Presario 100	13	1650
1103	Assurances	04.01.2000	AST Intel 150	2690	0.00%
1104	Assurances	04.01.2000	AST Intel 200	2	3190
1111	Alimentaire	04.01.2000	IBM 500 4	2299	0.00%
1113	Construction	04.01.2000	Compaq Presario 100	4	1650
1116	Pharmaceutique	04.01.2000	IBM 500 2	0.00%	4598
1110	Distribution	05.01.2000	AST Intel 200	6	1.50%
1113	Machines/outils	05.01.2000	Compaq Presario 100	3190	1.50%
1115	Distribution	05.01.2000	AST Intel 150	2690	0.00%
1123	Machines/outils	05.01.2000	Compaq Presario 100	8	1.50%
1124	Education	05.01.2000	AST Intel 200	3190	1.50%
1106	Construction	05.01.2000	Compaq Presario 100	11	1650
1101	Construction	05.01.2000	AST Intel 200	11	1.50%
1112	Pharmaceutique	06.01.2000	Compaq Presario 100	14	1650
1112	Machines/outils	06.01.2000	AST Intel 150	2690	1.50%
1125	Construction	06.01.2000	IBM 500 7	2299	1.50%
1100	Assurances	06.01.2000	Compaq Presario 100	23	1650
1125	Assurances	06.01.2000	IBM 500 3	0.00%	6897
1125	Construction	06.01.2000	AST Intel 200	2	3190
1104	Education	07.01.2000	AST Intel 150	12	2690
1126	Machines/outils	07.01.2000	AST Intel 150	24	2690
1121	Pharmaceutique	07.01.2000	IBM 500 8	2299	1.50%
1103	Distribution	07.01.2000	AST Intel 200	9	3190
1103	Assurances	07.01.2000	Compaq Presario 100	4	1.50%
1125	Construction	07.01.2000	AST Intel 200	2690	0.00%
1120	Banques 31	10.01.2000	IBM 500 2	16	5380
1111	Alimentaire	10.01.2000	Compaq Presario 100	0.00%	1650
1118	Education	10.01.2000	IBM 500 3	6897	1.50%
1127	Alimentaire	11.01.2000	IBM 500 7	2299	1.50%
1101	Construction	11.01.2000	AST Intel 200	6	1885

## 7.5. Exercice 5.: Importer/Exporter des données \*.accdb via ODBC

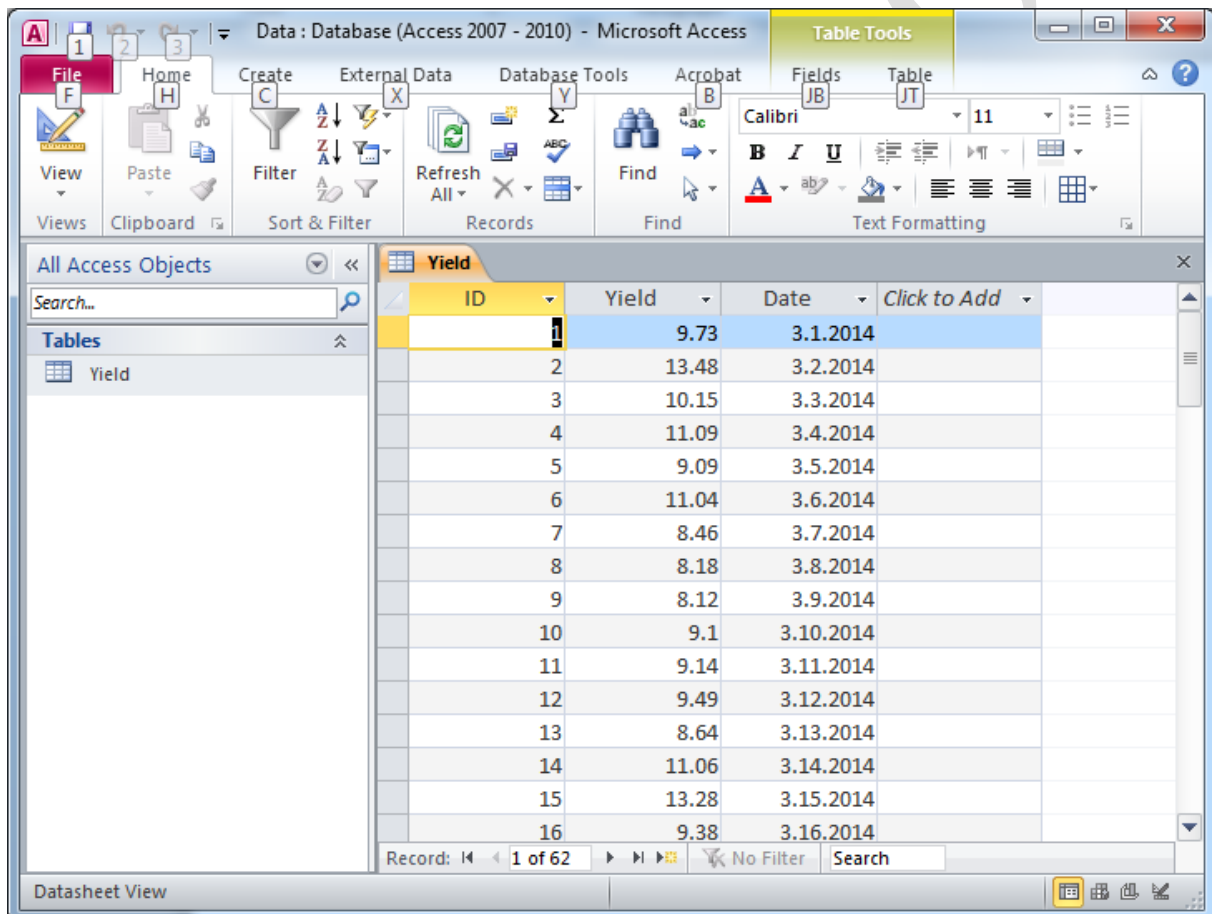
Minitab® Statistical Software 16.1

Un grand classique dans les entreprises consiste à importer des données d'une base Microsoft Office Access 2010.

Nous avons pour cela le fichier de base de données suivant à la racine du disque C:\

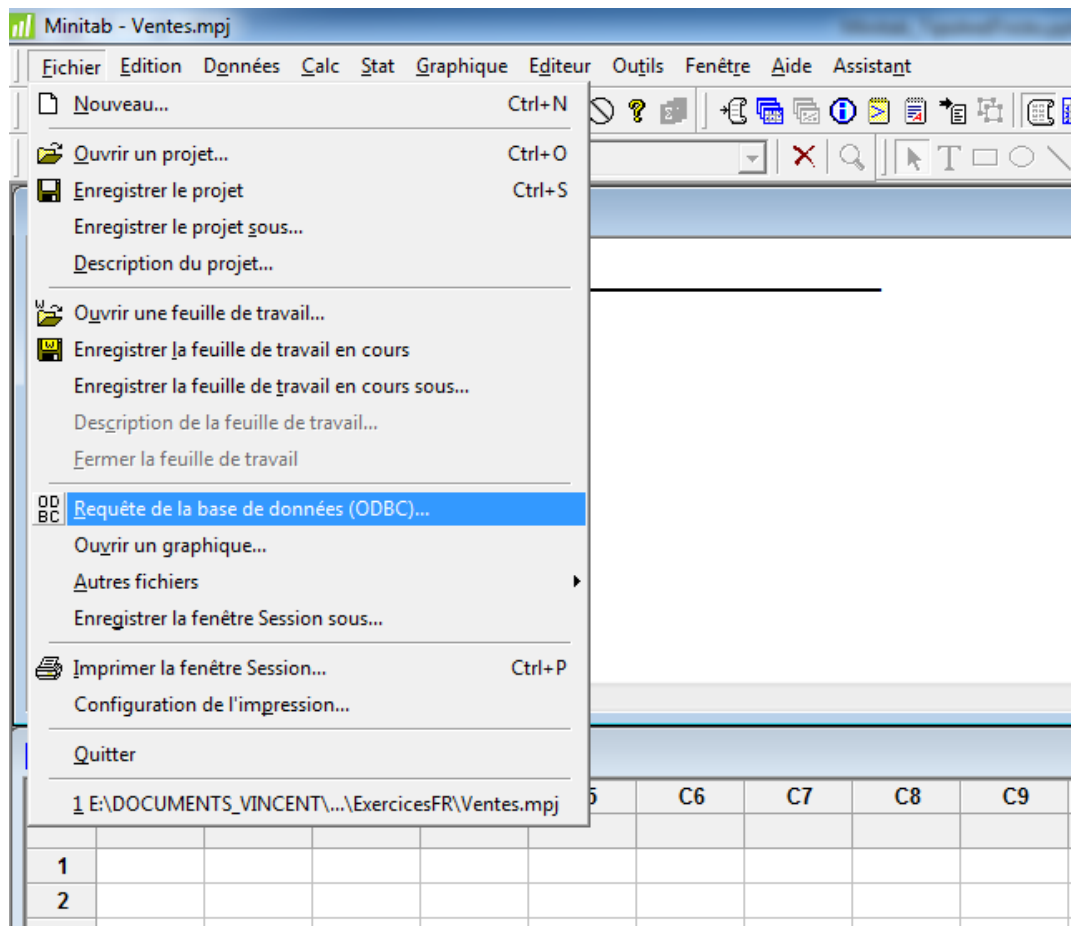
 Data.accdb

Son contenu est on ne peut plus standard:

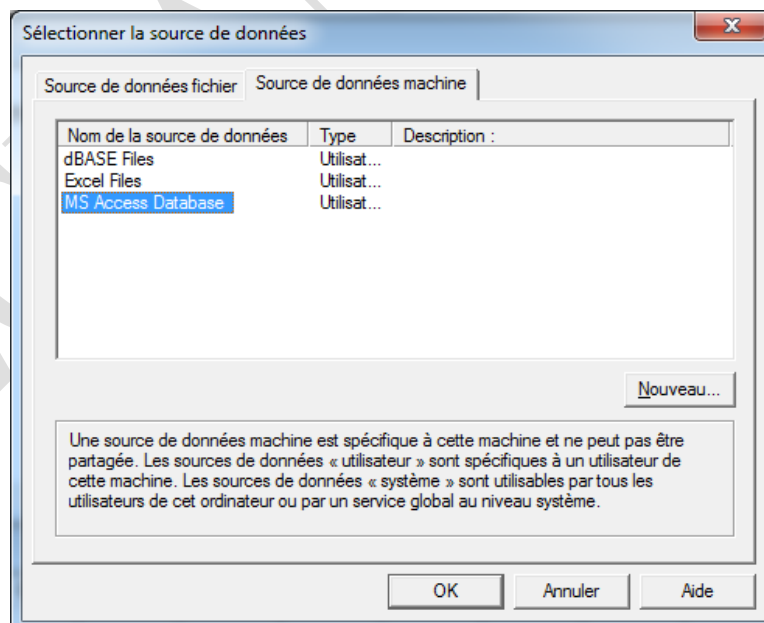


ID	Yield	Date	Click to Add
1	9.73	3.1.2014	
2	13.48	3.2.2014	
3	10.15	3.3.2014	
4	11.09	3.4.2014	
5	9.09	3.5.2014	
6	11.04	3.6.2014	
7	8.46	3.7.2014	
8	8.18	3.8.2014	
9	8.12	3.9.2014	
10	9.1	3.10.2014	
11	9.14	3.11.2014	
12	9.49	3.12.2014	
13	8.64	3.13.2014	
14	11.06	3.14.2014	
15	13.28	3.15.2014	
16	9.38	3.16.2014	

Ensuite pour importer dans Minitab (**et non lier!**) nous allons dans le menu **Fichier/Requête de la base de données (ODBC)...**:

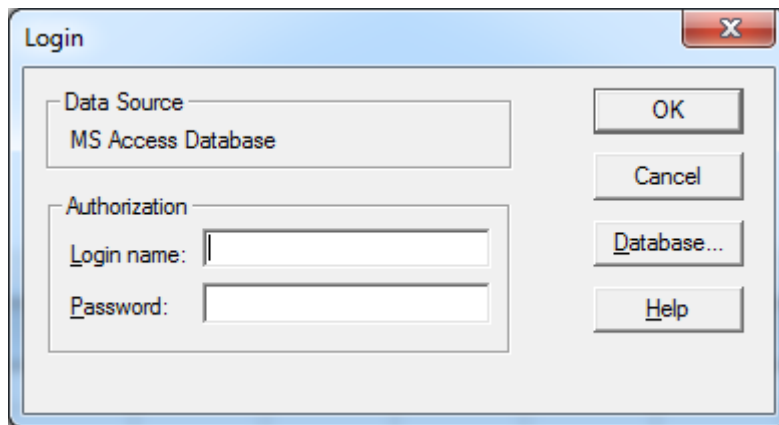


Apparaît la boîte de dialogue suivante où nous sélectionnons **MS Access Database** et nous cliquons sur **OK**:

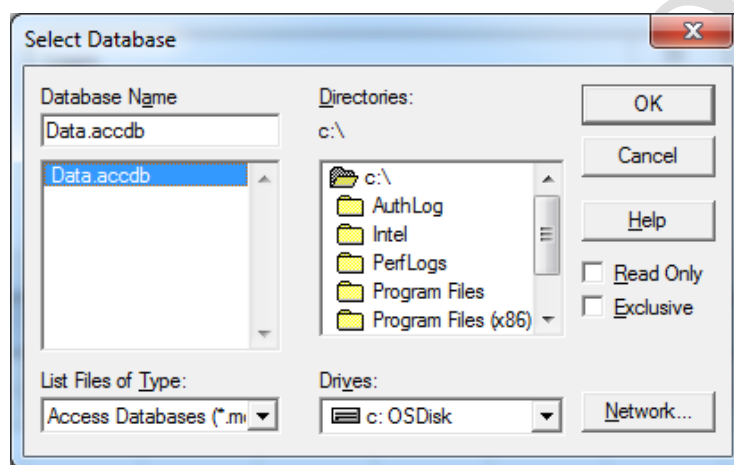


Apparaît alors:

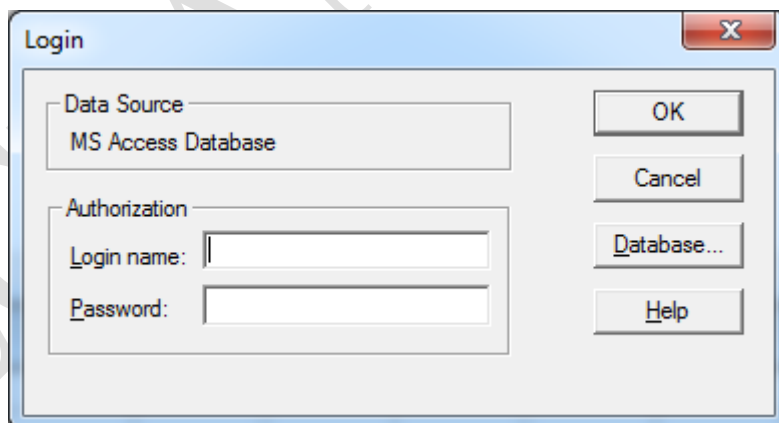




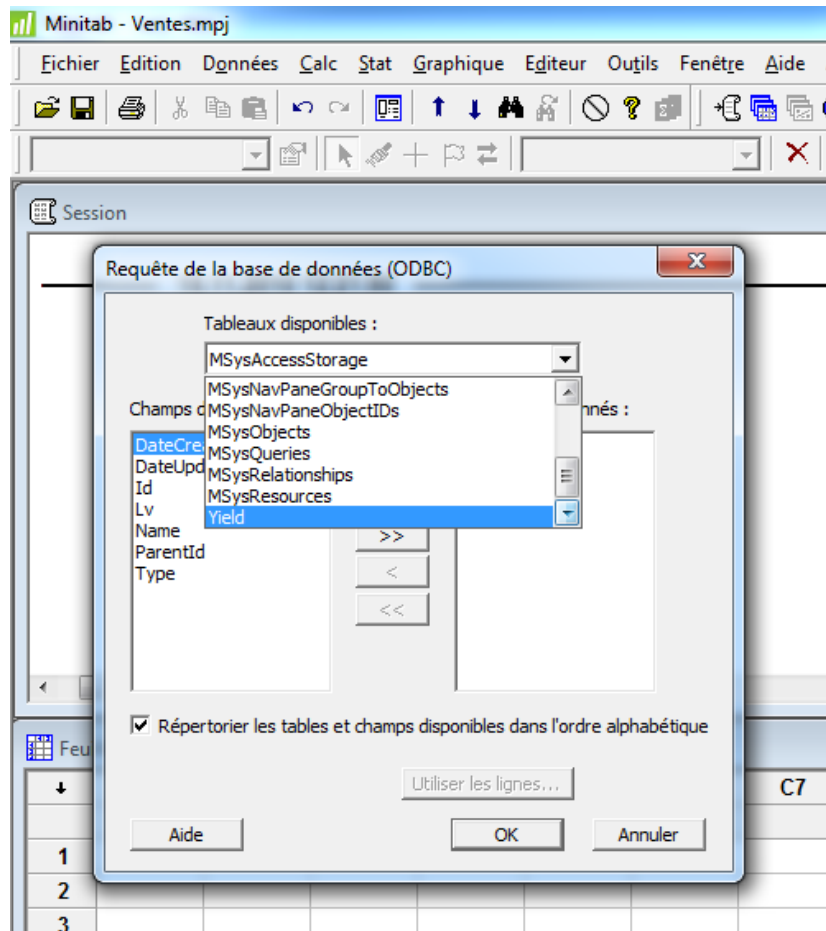
Nous cliquons sur **Database...** et allons chercher le fichier de base de données:



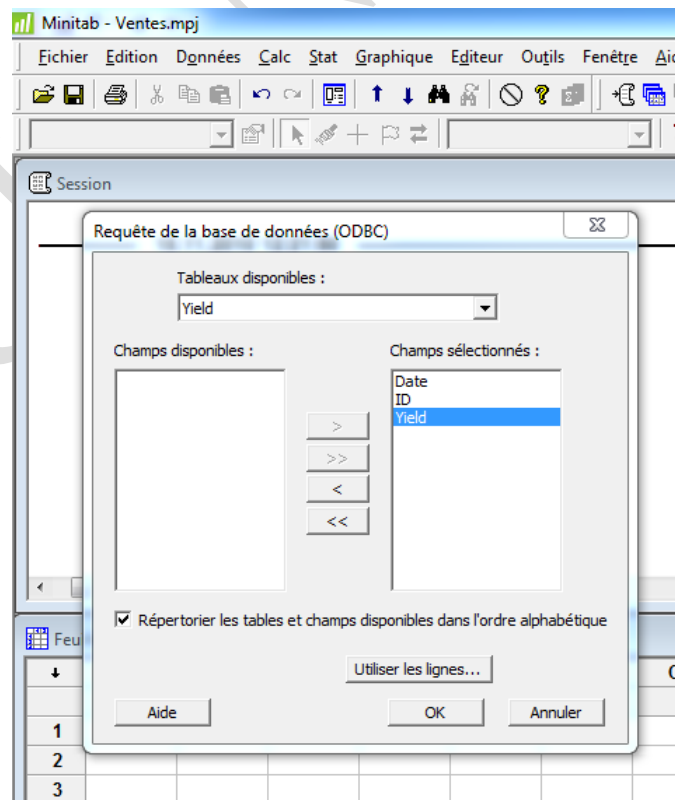
Nous validons par **OK** ce qui nous ramène à:



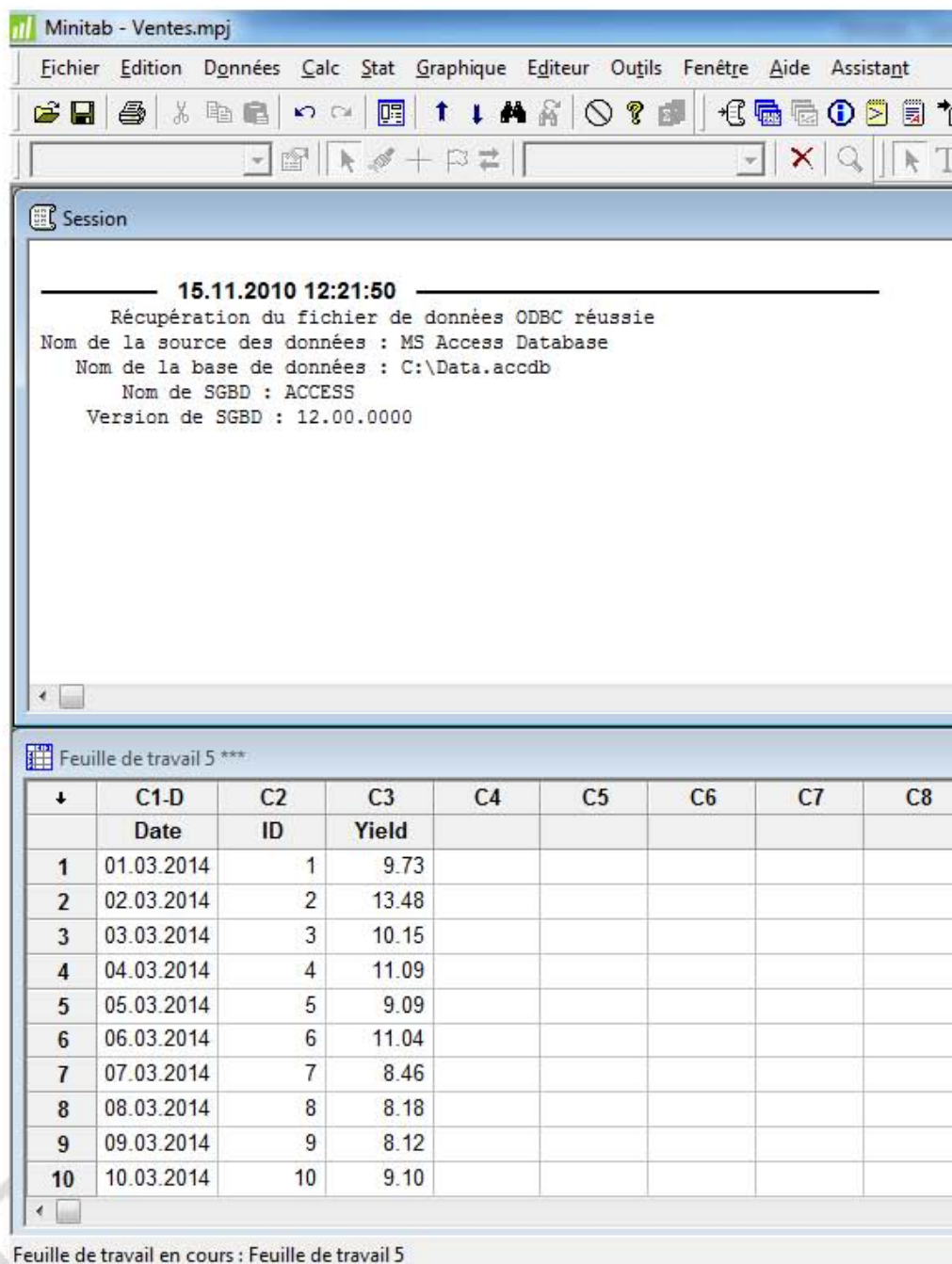
Quand nous validons par **OK**, cela nous amène à la boîte de dialogue suivante où nous pouvons choisir la table dans liste déroulante:



Une fois la table choisie, nous choisissons les colonnes:



Et si nous validons par **OK** les données sont alors importées:



The screenshot shows the Minitab interface with a session window and a worksheet. The session window displays the following text:

```
----- 15.11.2010 12:21:50 -----  
Récupération du fichier de données ODBC réussie  
Nom de la source des données : MS Access Database  
Nom de la base de données : C:\Data.accdb  
Nom de SGBD : ACCESS  
Version de SGBD : 12.00.0000
```

The worksheet, titled "Feuille de travail 5 \*\*\*", contains the following data:

	C1-D	C2	C3	C4	C5	C6	C7	C8
	Date	ID	Yield					
1	01.03.2014	1	9.73					
2	02.03.2014	2	13.48					
3	03.03.2014	3	10.15					
4	04.03.2014	4	11.09					
5	05.03.2014	5	9.09					
6	06.03.2014	6	11.04					
7	07.03.2014	7	8.46					
8	08.03.2014	8	8.18					
9	09.03.2014	9	8.12					
10	10.03.2014	10	9.10					

Feuille de travail en cours : Feuille de travail 5

## 7.6. Exercice 6.: Empiler/Désempiler des données

Minitab® Statistical Software 15.1.1.0

---

Il arrivera souvent dans la pratique que l'utilisateur copie/colle des données provenant de Microsoft Excel sous la forme suivante:

A	B
Ecart	Ecart
45%	15%
40%	15%
35%	15%
35%	10%
30%	10%
30%	10%
25%	5%
25%	5%
25%	0%
20%	-5%
20%	-5%
20%	-10%
20%	

ou avec encore plus de colonnes. Suivant les outils que nous verrons par la suite, il est nécessaire de mettre les données les unes sous les autres. Certes, un simple copier/coller suffit mais Minitab® Statistical Software a aussi un outil spécialement prévu à cet effet.

Ouvrez le fichier *DonneesDesempilees.mpj*:

	C1	C2	C3
	Ecarts1	Ecarts2	
1	0.45	0.15	
2	0.40	0.15	
3	0.35	0.15	
4	0.30	0.10	
5	0.30	0.10	
6	0.30	0.10	
7	0.25	0.05	
8	0.25	0.05	
9	0.25	0.00	
10	0.20	-0.05	
11	0.20	-0.10	
12	0.20		
13	0.20		
14			
15			
16			
17			
18			
19			

Feuille de travail en cours : Feuille de travail 1

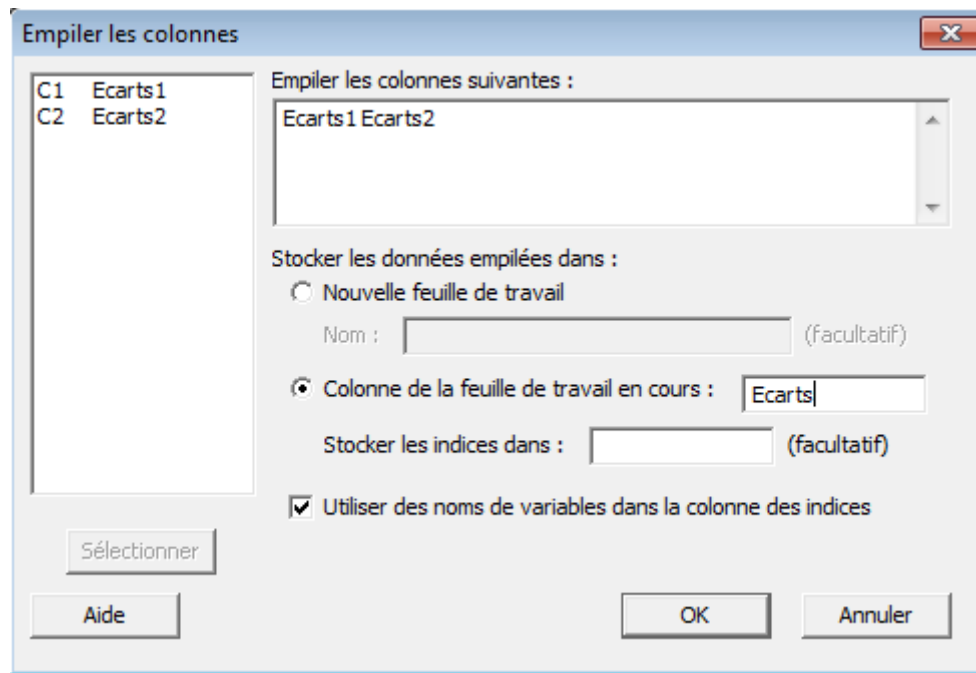
et allez dans le menu **Données/Empiler/Colonnes...**:

Minitab - DonneesDesempilees.MPJ - [Feuille de travail 1 \*\*\*]

Fichier Edition **Données** Calc Stat Graphique Editeur Outils Fenêtre Aide

- Diviser la feuille de travail en sous-ensembles...
- Fractionner la feuille de travail...
- Fusionner les feuilles de travail...
- Copier
- Déempiler les colonnes...
- Empiler**
  - Colonnes...
  - Blocs de colonnes...
  - Lignes...
- Transposer les colonnes...
- Trier...
- Rang...
- Supprimer les lignes...
- Effacer les variables...
- Coder
- Changer le type des données
- Extraire de date/heure
- Concaténer...
- Afficher les données...

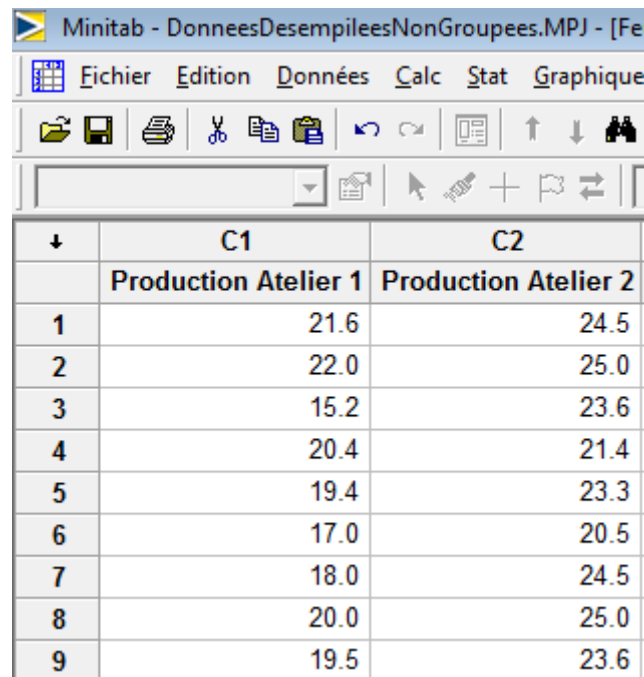
Il vous suffit alors de prendre:



et de valider par **OK** pour obtenir au final:

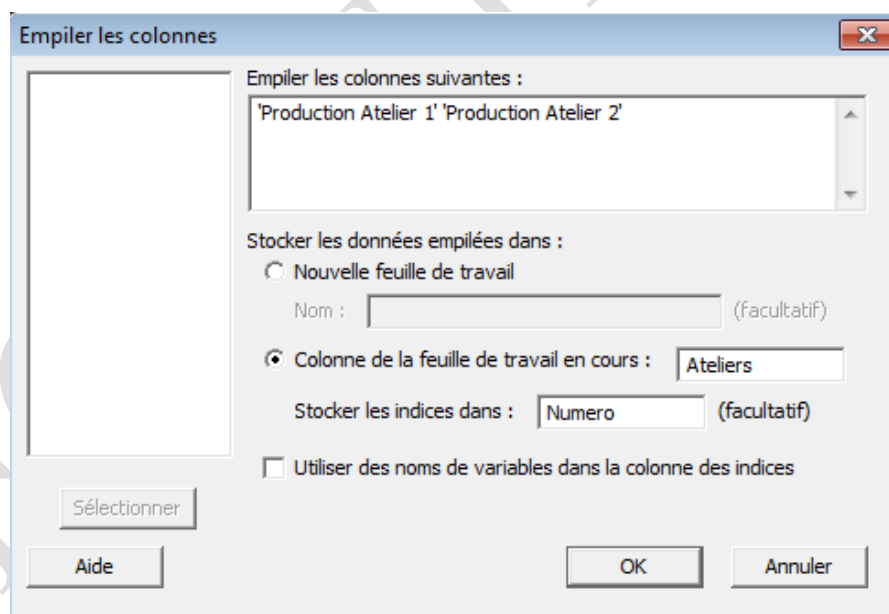
	C1	C2	C3
	Ecart1	Ecart2	Ecart3
1	0.45	0.15	0.45
2	0.40	0.15	0.40
3	0.35	0.15	0.35
4	0.30	0.10	0.30
5	0.30	0.10	0.30
6	0.30	0.10	0.30
7	0.25	0.05	0.25
8	0.25	0.05	0.25
9	0.25	0.00	0.25
10	0.20	-0.05	0.20
11	0.20	-0.10	0.20
12	0.20		0.20
13	0.20		0.20
14			0.15
15			0.15
16			0.15
17			0.10
18			0.10
19			0.10

Ouvrez maintenant le fichier *DonneesDesempileesNonGroupees.mpj*:

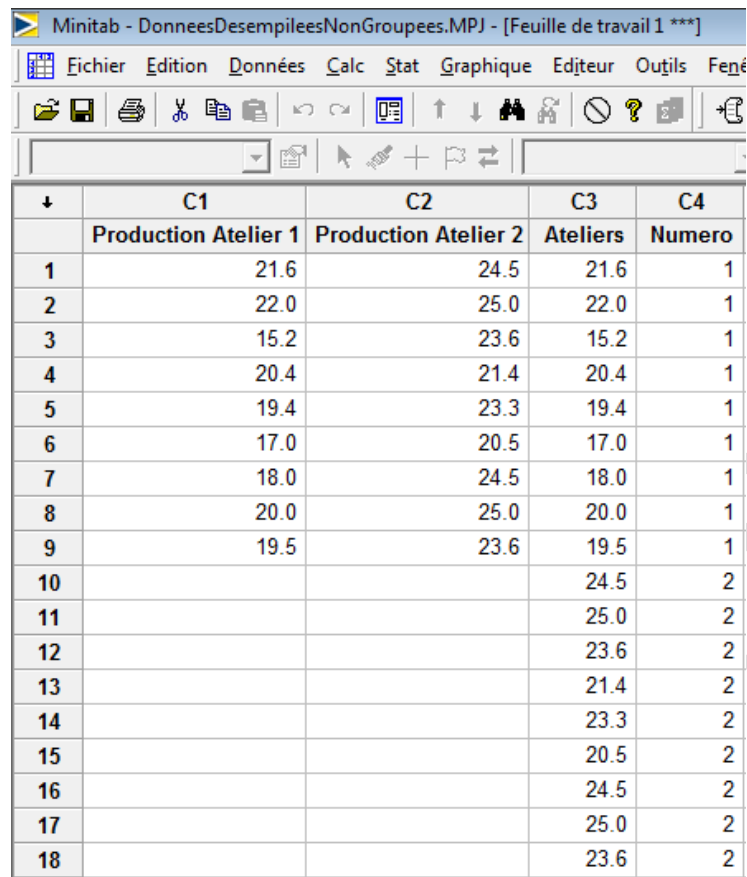


	C1	C2
	Production Atelier 1	Production Atelier 2
1	21.6	24.5
2	22.0	25.0
3	15.2	23.6
4	20.4	21.4
5	19.4	23.3
6	17.0	20.5
7	18.0	24.5
8	20.0	25.0
9	19.5	23.6

Nous souhaiterions faire la même chose qu'avant mais avec une petite subtilité nécessaire (disons plutôt "préférée" par certains utilisateurs). Il s'agit d'empiler les données mais avec une variable qualitative de l'appartenance de groupe. La méthode est la même qu'avant mais à la différence que dans la boîte de dialogue de regroupement nous définissons **Stocker les indices dans:**

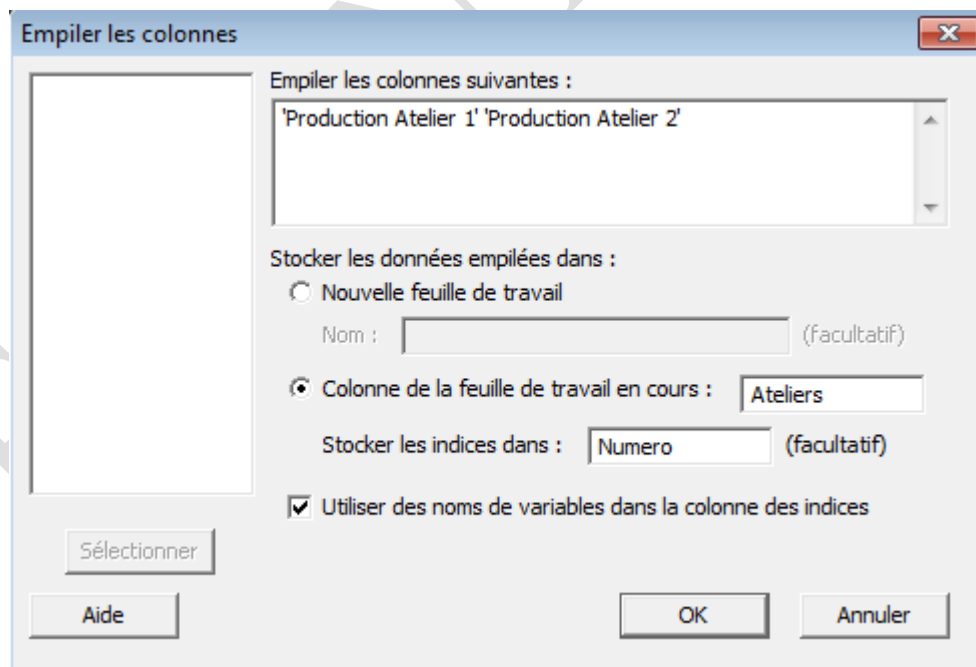


Pour obtenir:



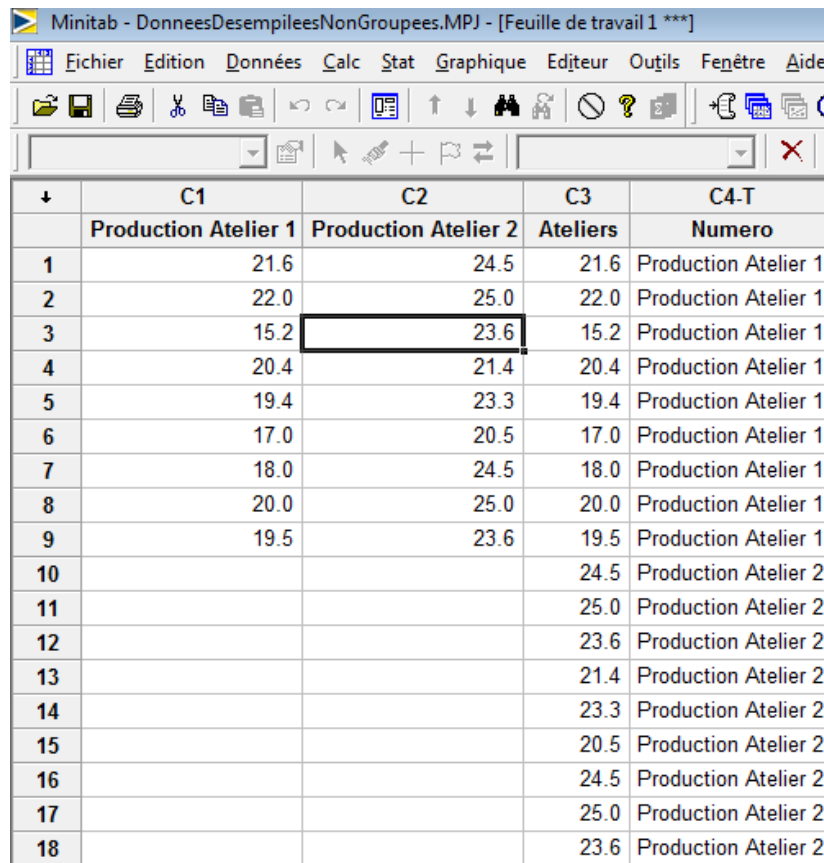
	C1	C2	C3	C4
	Production Atelier 1	Production Atelier 2	Ateliers	Numero
1	21.6	24.5	21.6	1
2	22.0	25.0	22.0	1
3	15.2	23.6	15.2	1
4	20.4	21.4	20.4	1
5	19.4	23.3	19.4	1
6	17.0	20.5	17.0	1
7	18.0	24.5	18.0	1
8	20.0	25.0	20.0	1
9	19.5	23.6	19.5	1
10			24.5	2
11			25.0	2
12			23.6	2
13			21.4	2
14			23.3	2
15			20.5	2
16			24.5	2
17			25.0	2
18			23.6	2

ou (ce qui est moins courant mais plus lisible dans les rapports de la fenêtre de session):



où nous avons recoché **Utiliser des noms de variables dans la colonne d'indice**. Ce qui donne:

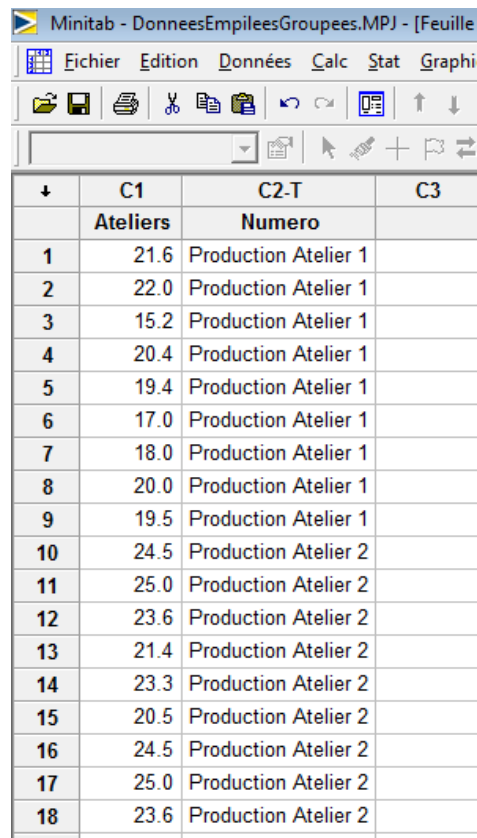




↓	C1	C2	C3	C4-T
	Production Atelier 1	Production Atelier 2	Ateliers	Numero
1	21.6	24.5	21.6	Production Atelier 1
2	22.0	25.0	22.0	Production Atelier 1
3	15.2	23.6	15.2	Production Atelier 1
4	20.4	21.4	20.4	Production Atelier 1
5	19.4	23.3	19.4	Production Atelier 1
6	17.0	20.5	17.0	Production Atelier 1
7	18.0	24.5	18.0	Production Atelier 1
8	20.0	25.0	20.0	Production Atelier 1
9	19.5	23.6	19.5	Production Atelier 1
10			24.5	Production Atelier 2
11			25.0	Production Atelier 2
12			23.6	Production Atelier 2
13			21.4	Production Atelier 2
14			23.3	Production Atelier 2
15			20.5	Production Atelier 2
16			24.5	Production Atelier 2
17			25.0	Production Atelier 2
18			23.6	Production Atelier 2

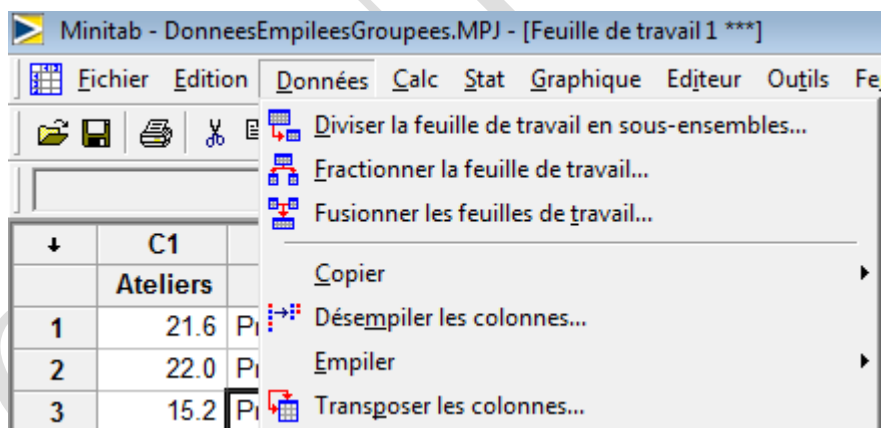
Maintenant, en imaginant que nous avons importé via copier/coller de Microsoft Excel le tableau précédent, supposons que nous souhaitons faire le contraire: désempiler des données.

Pour cela, ouvrez le fichier *DonneesEmpileesGroupees.mpj*:

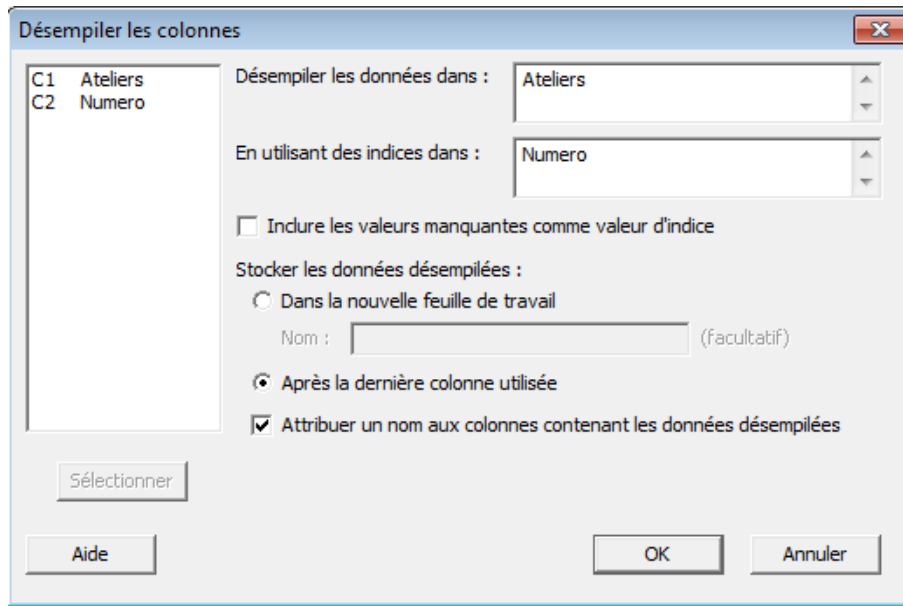


	C1	C2-T	C3
	Ateliers	Numero	
1	21.6	Production Atelier 1	
2	22.0	Production Atelier 1	
3	15.2	Production Atelier 1	
4	20.4	Production Atelier 1	
5	19.4	Production Atelier 1	
6	17.0	Production Atelier 1	
7	18.0	Production Atelier 1	
8	20.0	Production Atelier 1	
9	19.5	Production Atelier 1	
10	24.5	Production Atelier 2	
11	25.0	Production Atelier 2	
12	23.6	Production Atelier 2	
13	21.4	Production Atelier 2	
14	23.3	Production Atelier 2	
15	20.5	Production Atelier 2	
16	24.5	Production Atelier 2	
17	25.0	Production Atelier 2	
18	23.6	Production Atelier 2	

et allons dans le menu **Données/Désempiler les colonnes...**:



Il vient:



si nous validons par **OK** nous obtenons:

	C1	C2-T	C3	C4
	Ateliers	Numero	Ateliers_Production Atelier 1	Ateliers_Production Atelier 2
1	21.6	Production Atelier 1	21.6	24.5
2	22.0	Production Atelier 1	22.0	25.0
3	15.2	Production Atelier 1	15.2	23.6
4	20.4	Production Atelier 1	20.4	21.4
5	19.4	Production Atelier 1	19.4	23.3
6	17.0	Production Atelier 1	17.0	20.5
7	18.0	Production Atelier 1	18.0	24.5
8	20.0	Production Atelier 1	20.0	25.0
9	19.5	Production Atelier 1	19.5	23.6
10	24.5	Production Atelier 2		
11	25.0	Production Atelier 2		
12	23.6	Production Atelier 2		
13	21.4	Production Atelier 2		
14	23.3	Production Atelier 2		
15	20.5	Production Atelier 2		
16	24.5	Production Atelier 2		
17	25.0	Production Atelier 2		
18	23.6	Production Atelier 2		

on repassera pour le choix des étiquettes de colonne...

## 7.7. Exercice 7.: Générer une variable aléatoire et faire un histogramme avec ou sans ajustement

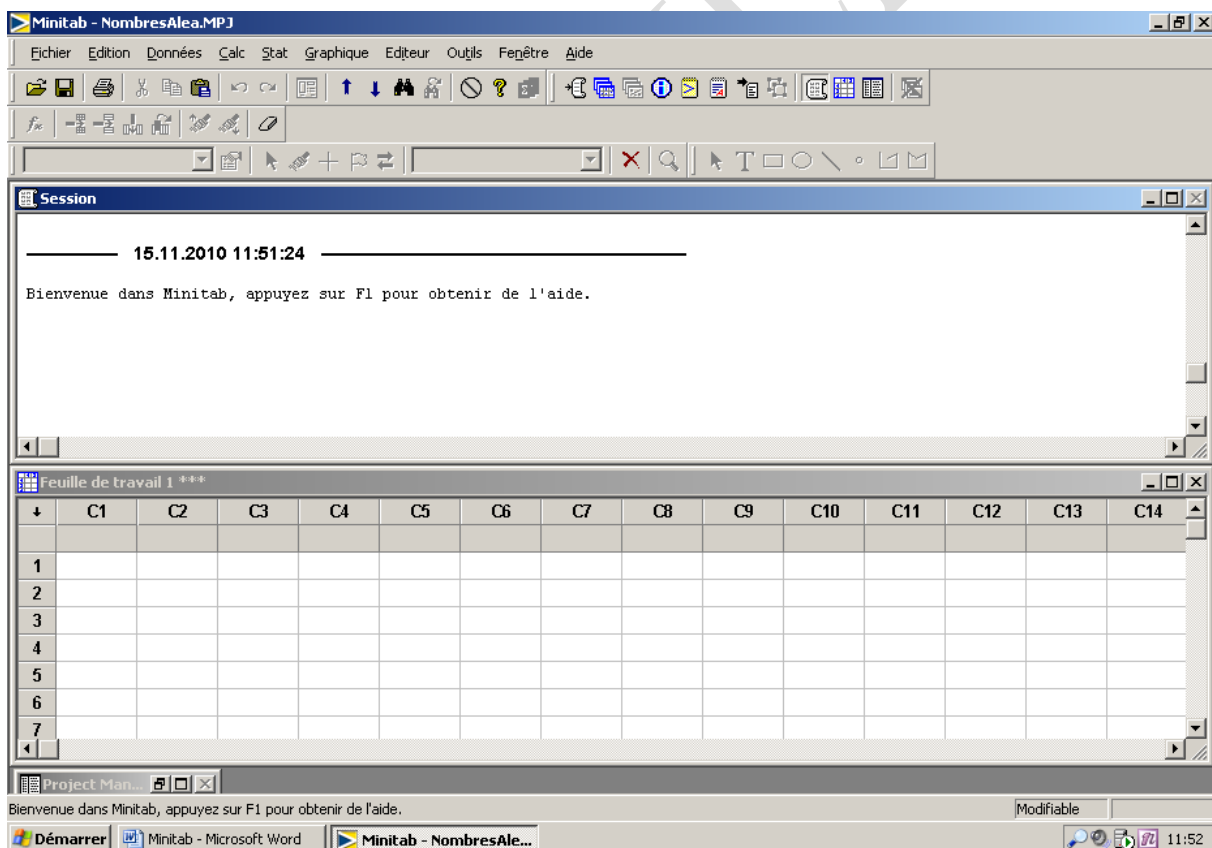
Minitab® Statistical Software 15.1.1.0

Il arrive fréquemment que l'on souhaite comparer des données mesurées dont nous faisons l'hypothèse qu'elle suive une loi de probabilité donnée, avec des vraies variables aléatoires issues de cette loi. Ne serait-ce que pour faire une représentation graphique de l'histogramme de chacun et de pouvoir les comparer.

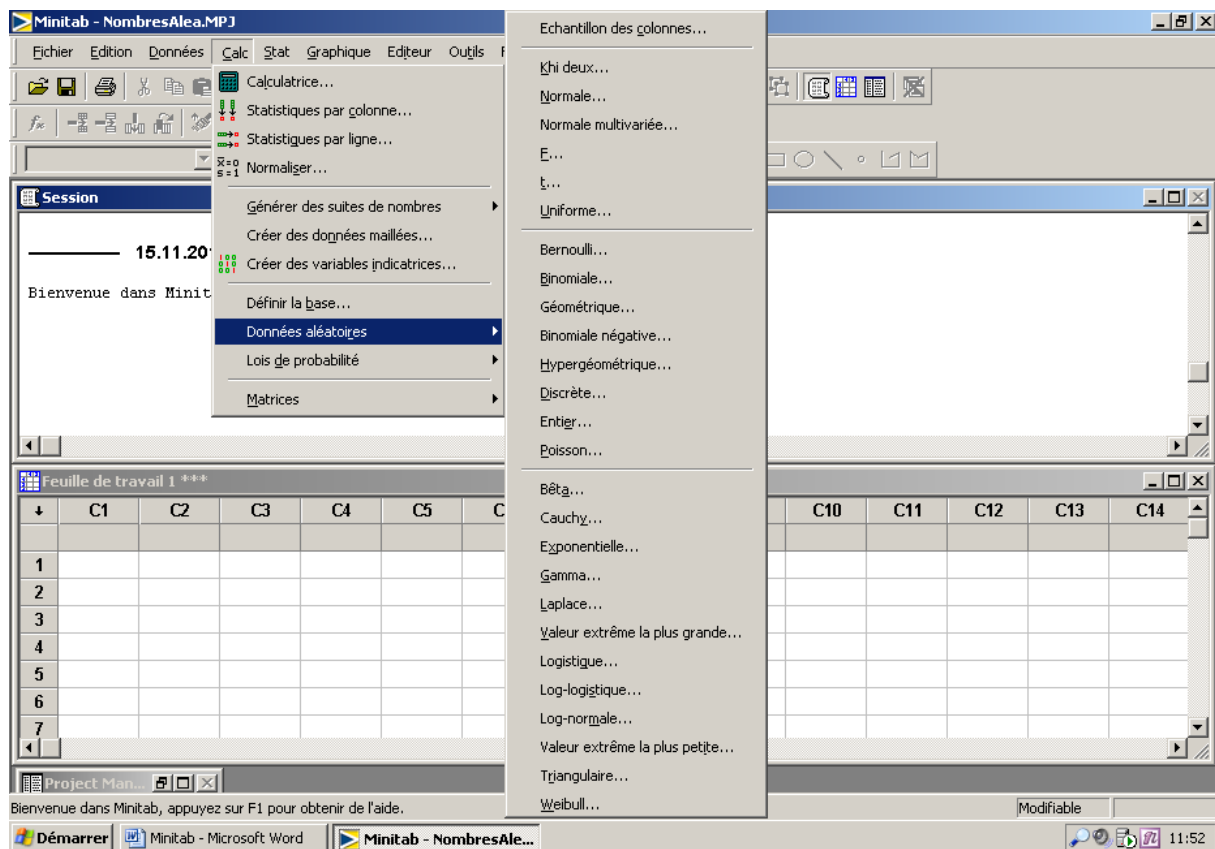
La génération de variables aléatoires est également utile dans le cadre de simulations de Monte-Carlo après avoir copier/coller les données dans Microsoft Excel.

Comme il s'agit d'une manipulation simple à effectuer dans Minitab® Statistical Software, nous allons donc commencer par ce premier point.

Ouvrez donc Minitab® Statistical Software à vide et enregistrez le fichier projet sous le nom *NombreAlea.mpj*.

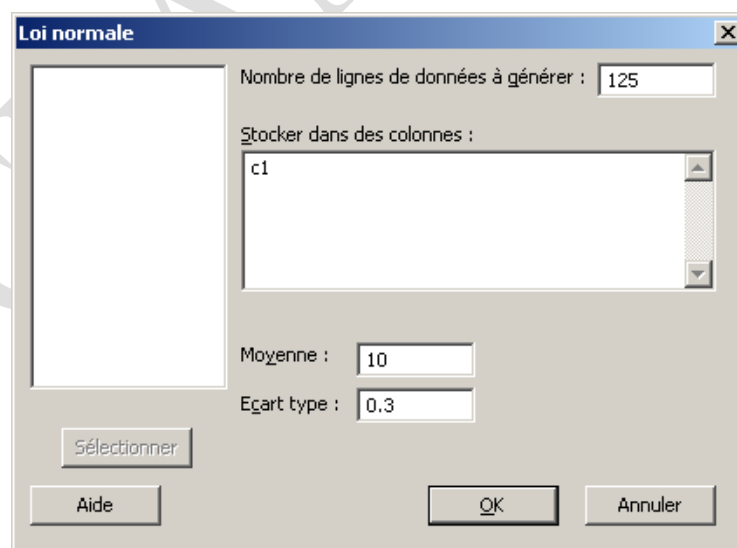


Ensuite allez dans le menu **Calc/Données aléatoires**:

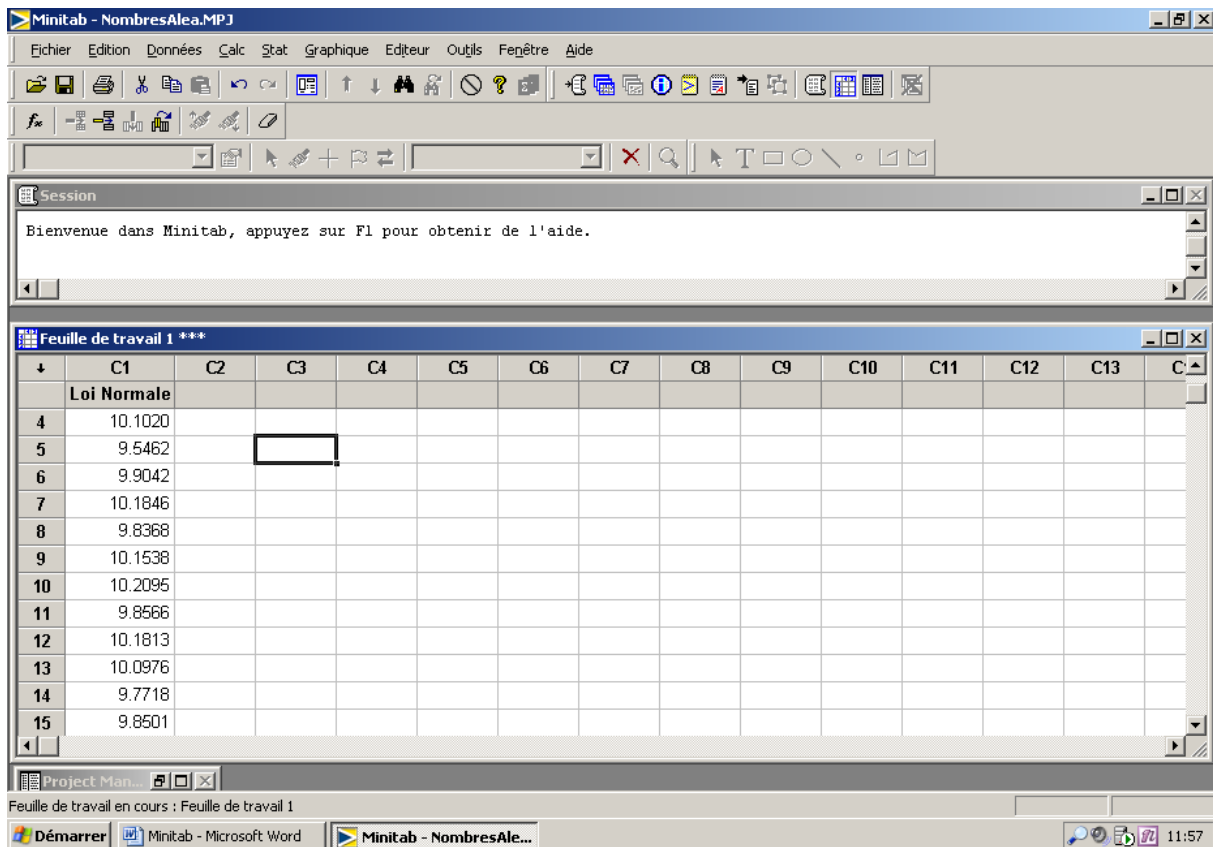


Vous y verrez la possibilité de générer de nombreuses variables aléatoires (beaucoup plus que ne le propose l'utilitaire d'analyse de Microsoft Excel).

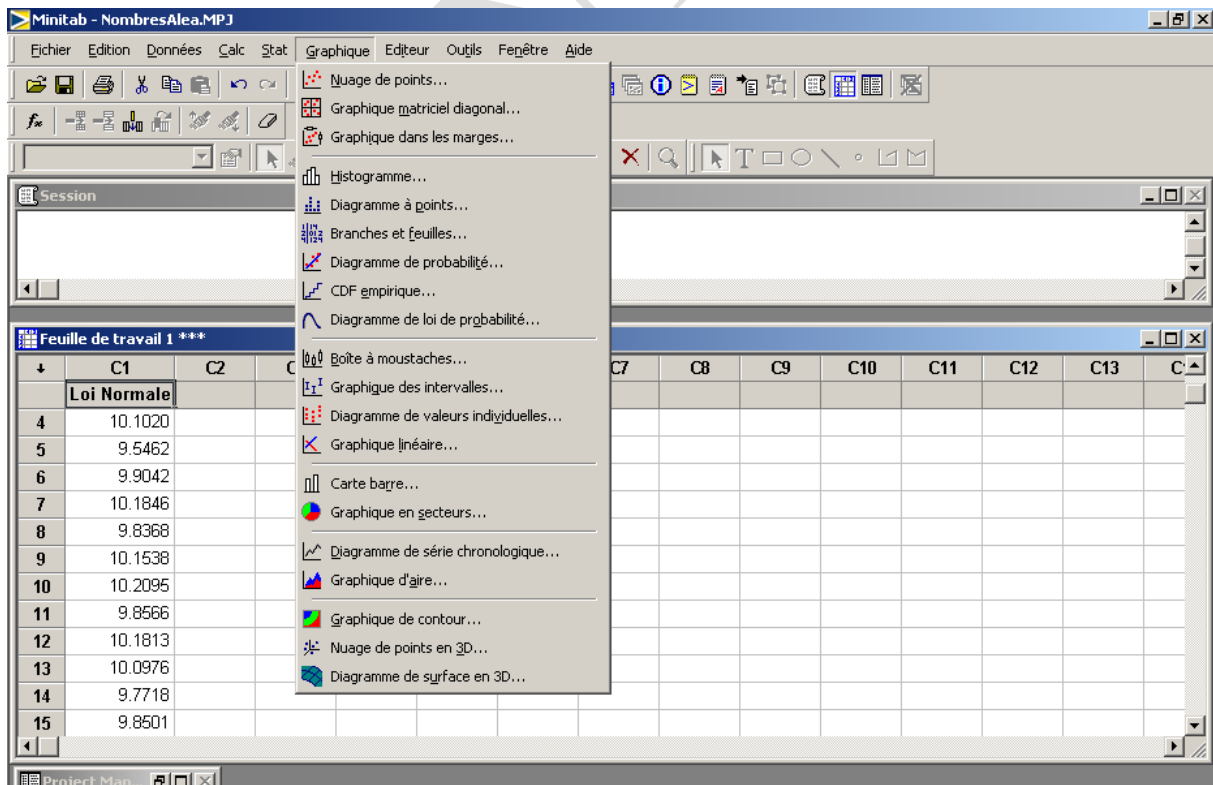
Prenons dans l'ordre des cas les plus courants dans l'industrie. Commençons par la loi Normale:



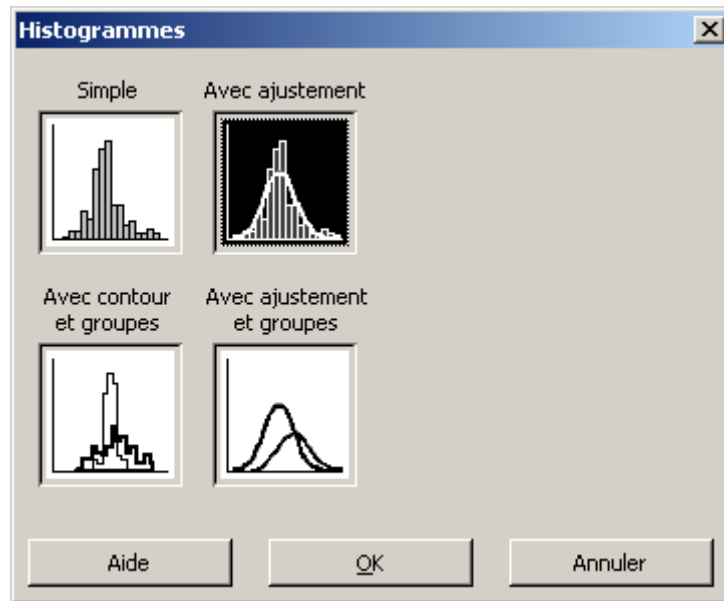
Nous aurons alors des données distribuées selon une loi normale de moyenne 10 et d'écart-type de 0.3:



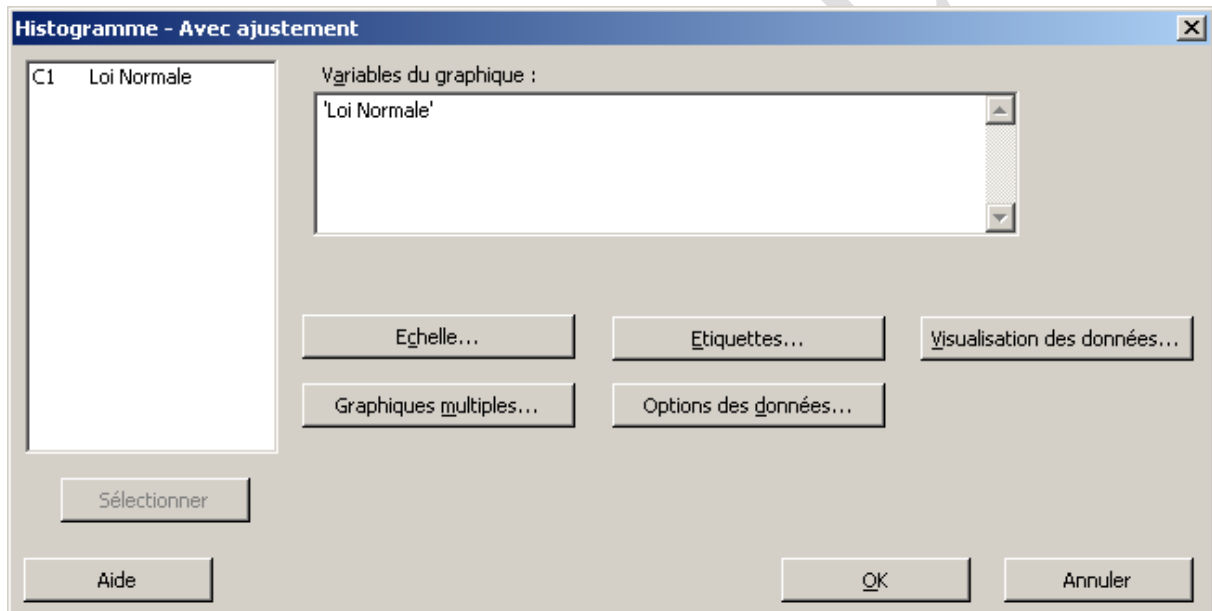
Nous pouvons tout de suite en faire un graphique en allant dans le menu **Graphique/Histogramme...**:



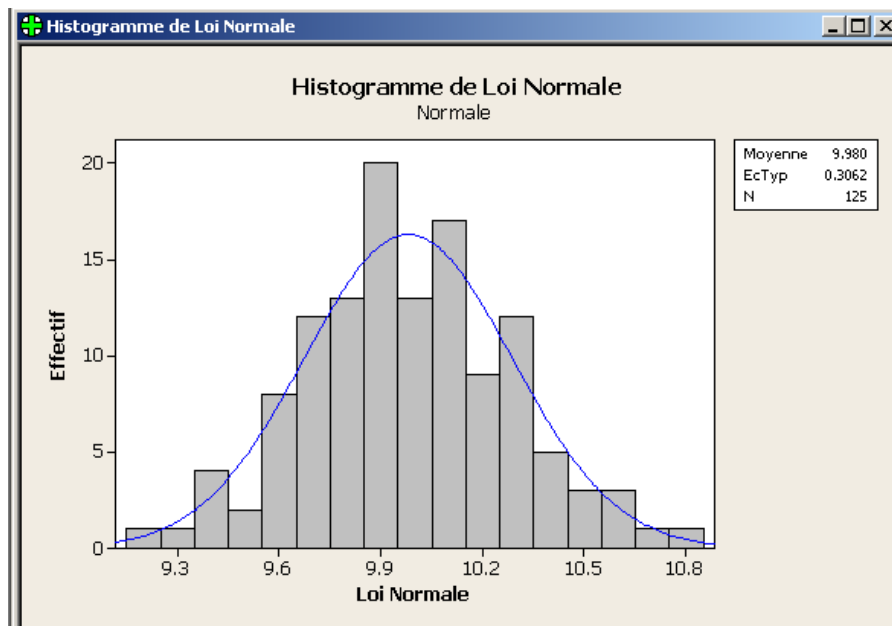
Ce qui fait apparaître:



Prenez l'option **Avec ajustement**:

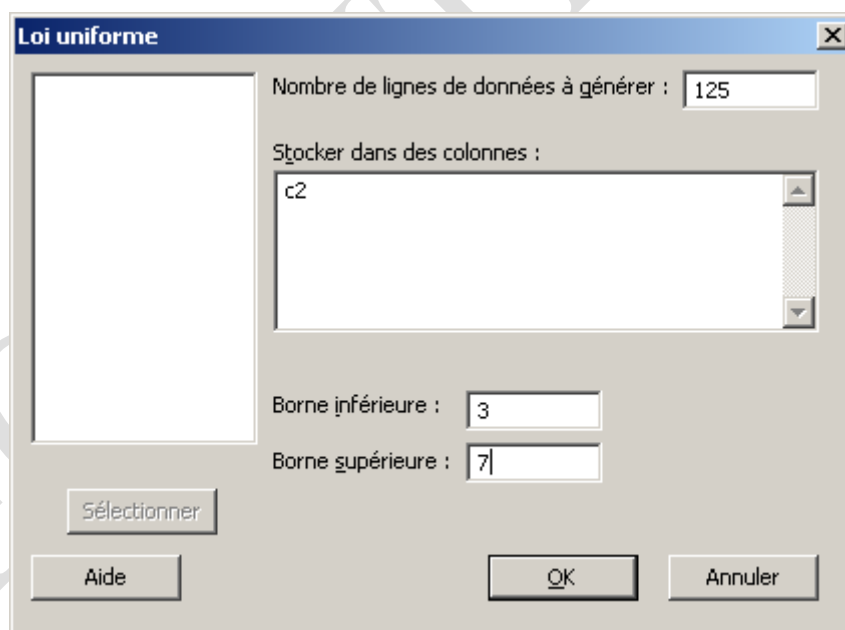


Et validez par **OK**:



Remarque: Contrairement à Microsoft Excel et R il semblerait qu'avec Minitab on ne puisse pas dans le cas précédent avoir un axe avec les fréquences et en même temps celui avec la densité.

Refaisons un autre exemple du même type avec la loi Uniforme (très utilisée en gestion de projets) cette fois:



Loi uniforme

Nombre de lignes de données à générer : 125

Stocker dans des colonnes :  
c2

Borne inférieure : 3

Borne supérieure : 7

Sélectionner

Aide

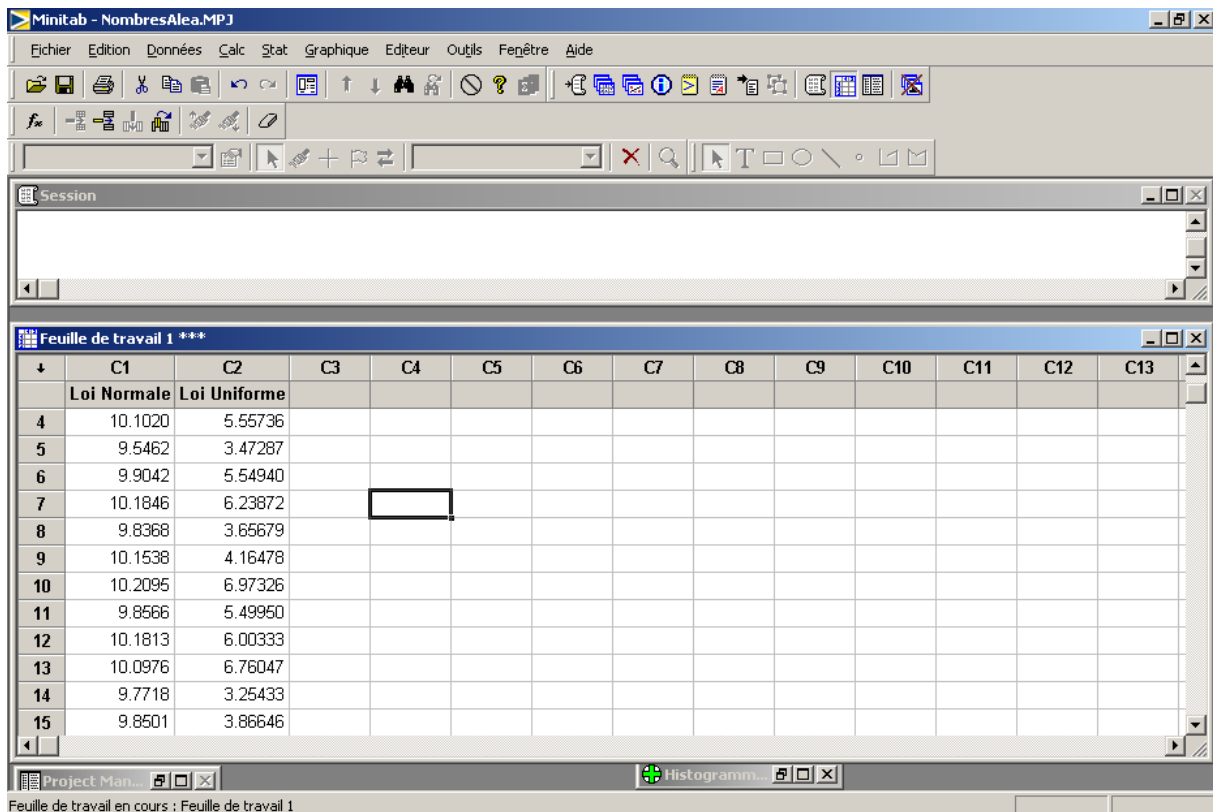
OK

Annuler

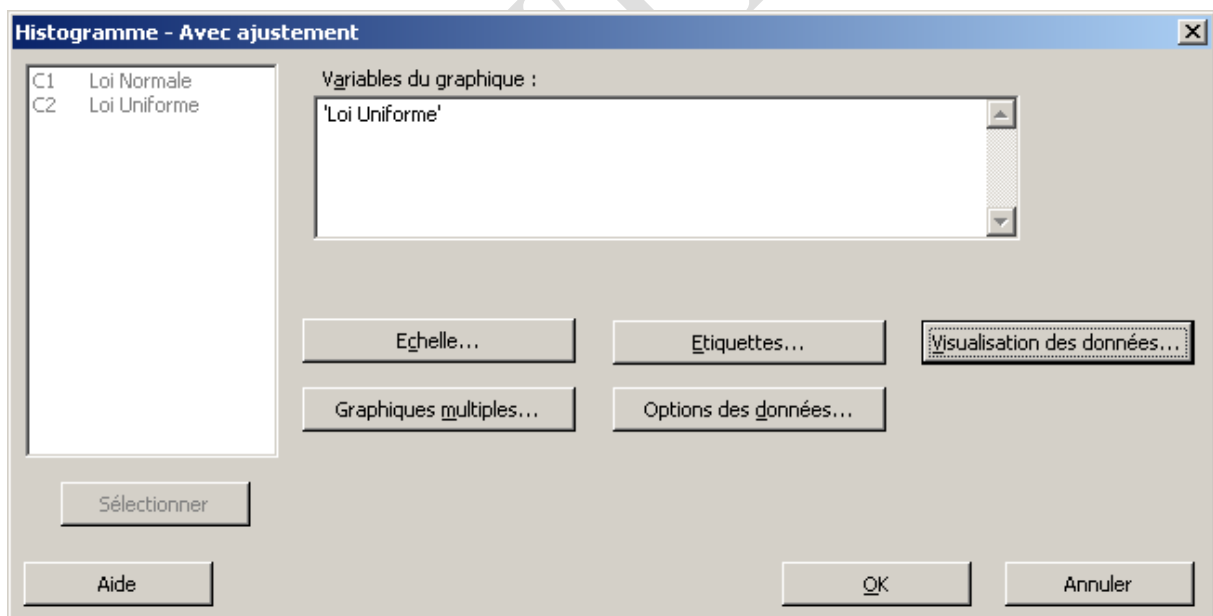
Detailed description: A dialog box for generating a uniform distribution. It has a title bar 'Loi uniforme'. The 'Nombre de lignes de données à générer' is set to 125. The 'Stocker dans des colonnes' list contains 'c2'. The 'Borne inférieure' is 3 and the 'Borne supérieure' is 7. There are buttons for 'Sélectionner', 'Aide', 'OK', and 'Annuler'.

Nous aurons alors:

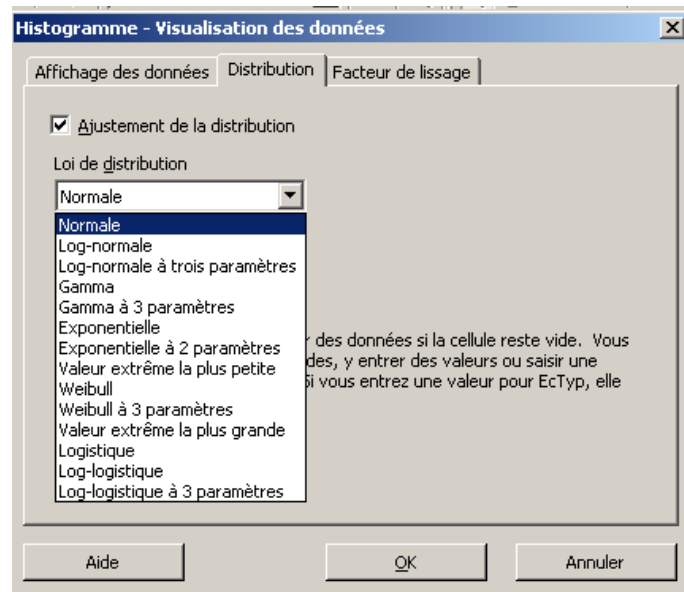




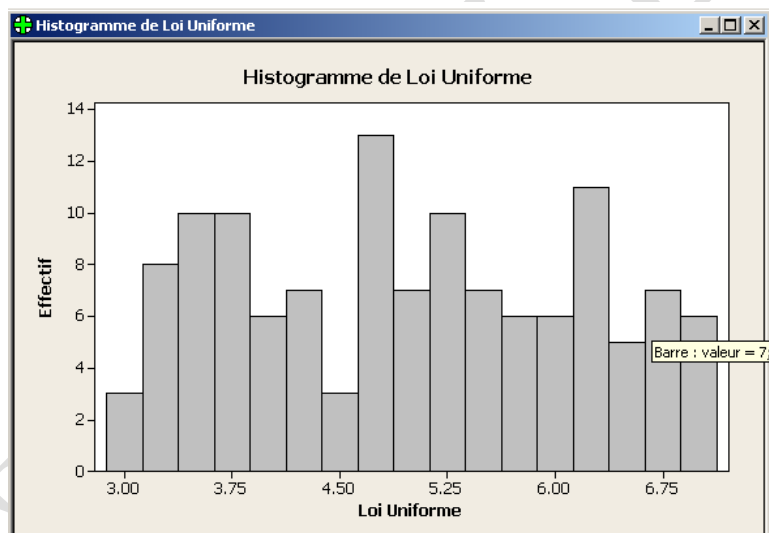
Si nous en faisons aussi un histogramme ajusté:



Il faudra cette fois-ci cliquer sur le bouton **Visualisation des données** pour choisir le type d'ajustement:



Mais nous voyons que Minitab® Statistical Software 15.1.1.0 ne propose pas d'ajustement pour une loi uniforme. Nous ferons alors un histogramme sans ajustement:



Et nous pouvons continuer ainsi de suite encore longtemps pour les plus utilisées conjointement avec Microsoft Excel comme: la loi de Poisson, la loi Triangulaire et la loi de Weibull.

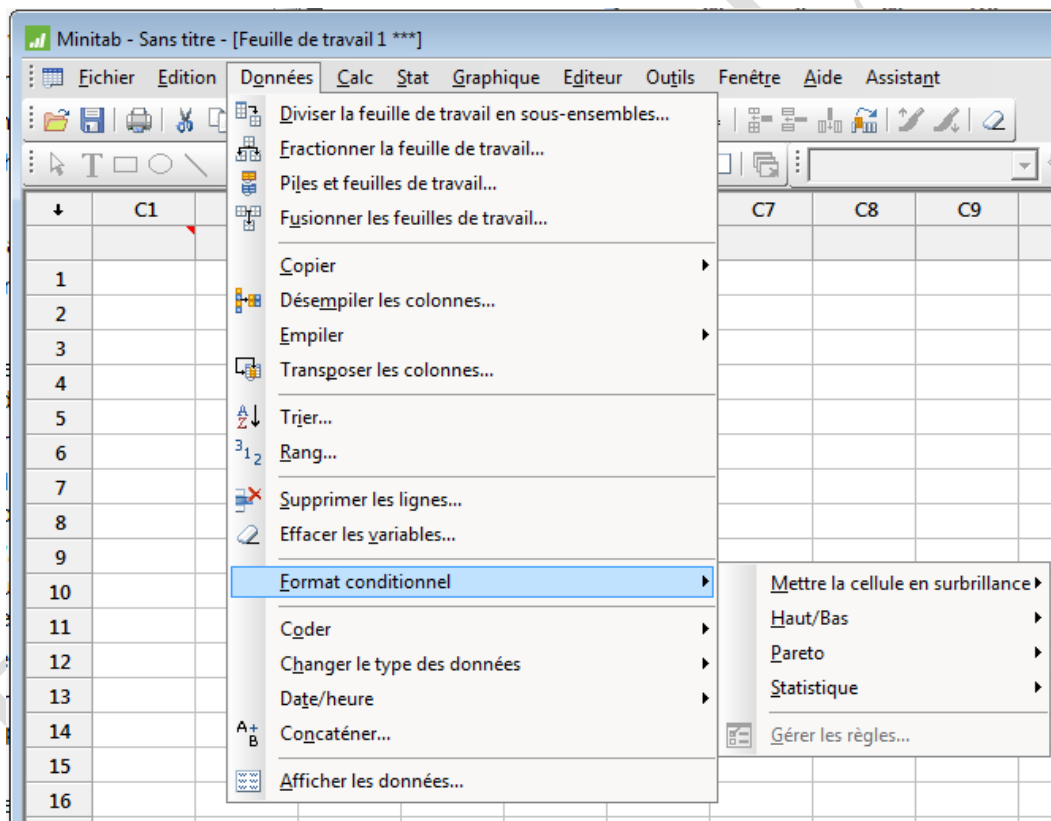
## 7.8. Exercice 8.: Formatage conditionnel

Minitab® Statistical Software 17.2.1

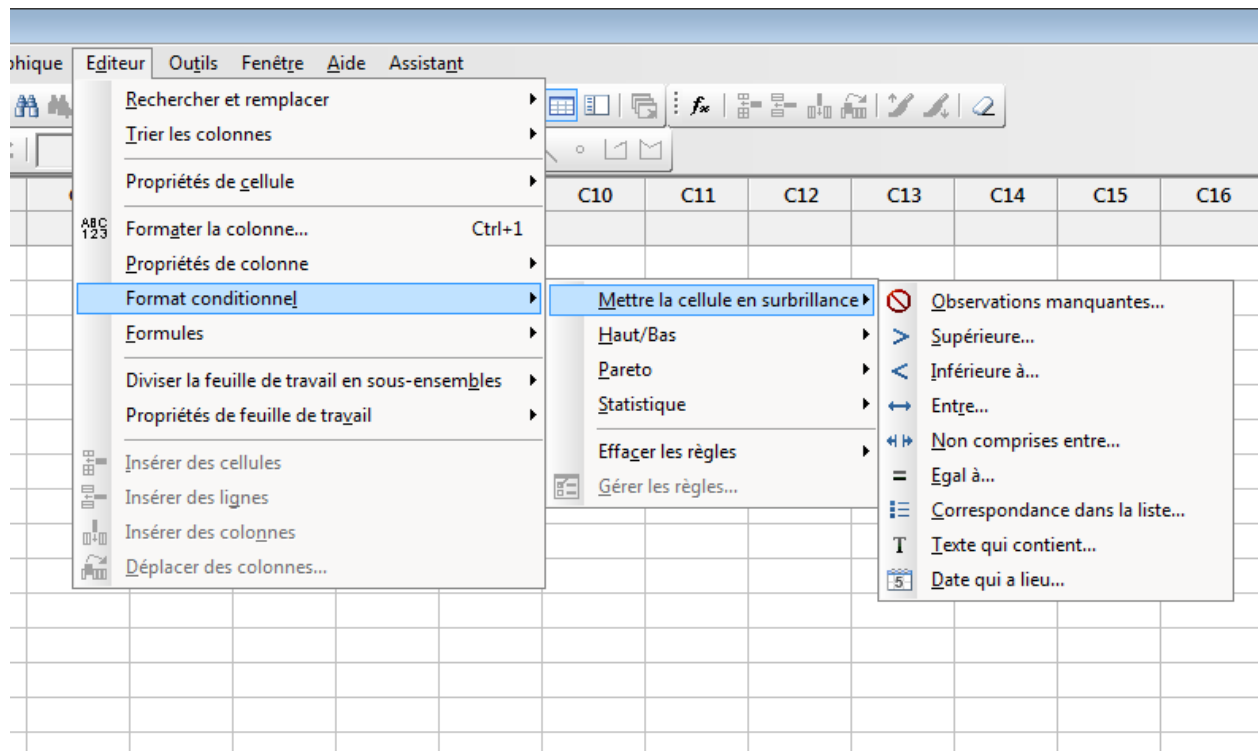
Nous allons voir ici une nouveauté de Minitab 17 très utile lorsque nous traitons de données avec de nombreuses colonnes et permettant d'identifier les machines, les dates et heures, les clients ou les expérimentateurs dont les données mesurées ne sont qu'une parmi d'autres colonnes et que certaines de ces données ayant des propriétés particulières sont bien visibles dans les graphiques mais difficiles par la suite à retrouver dans la liste de données elle-même!

Commençons par les éléments qui sont les mêmes que dans Microsoft Excel et qui ne nécessitent aucun exemple (puisque Microsoft Excel est censé être maîtrisé avec d'étudier ou d'utiliser Minitab pour rappel).

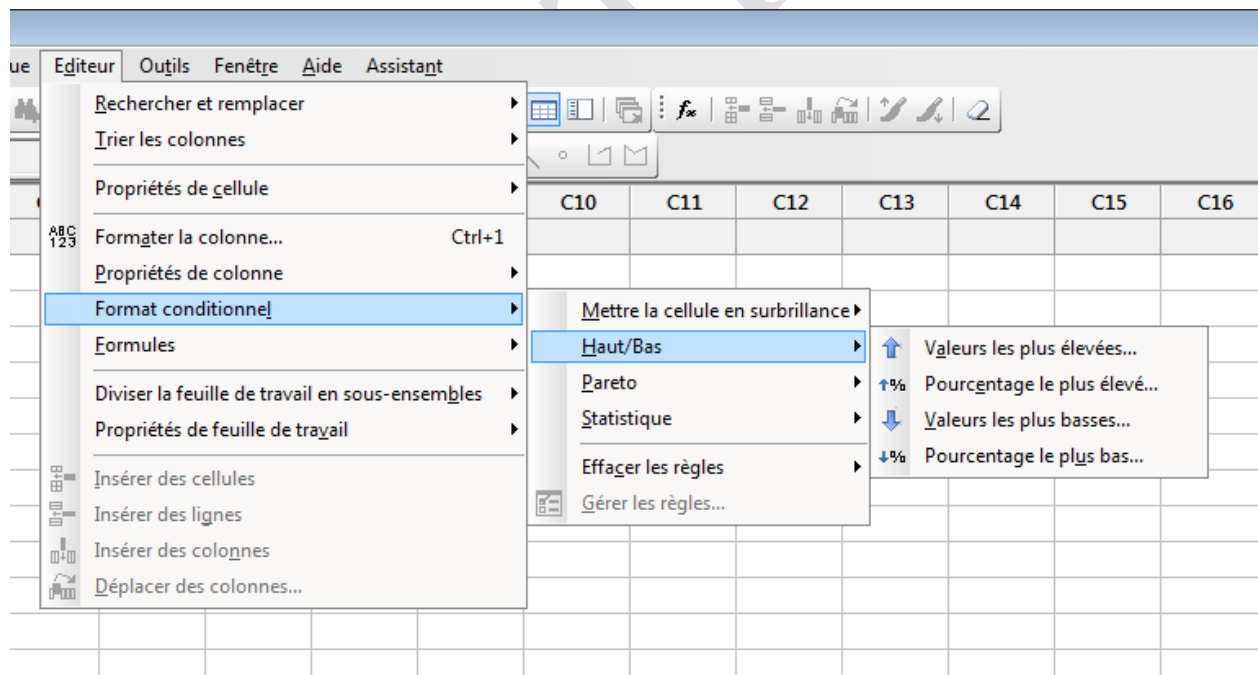
Nous passerons par le menu **Données/Format conditionnel** (oui il existe plusieurs manières d'y accéder):



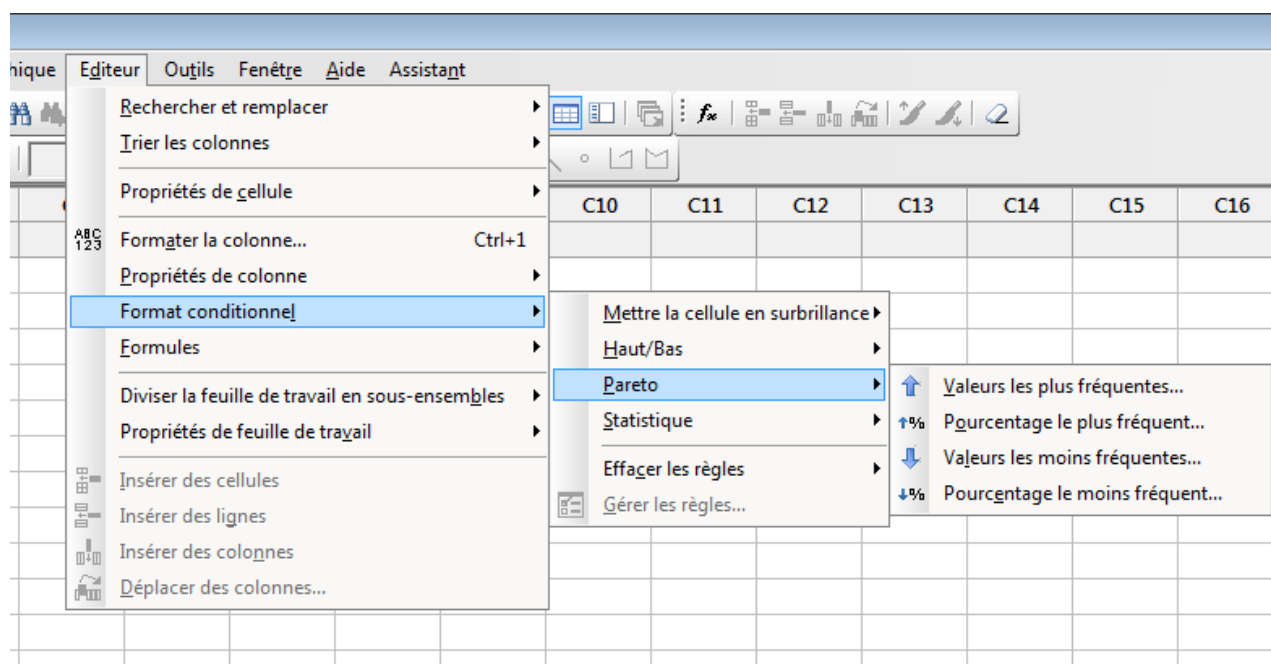
Ensuite nous avons le cas des surbrillances triviales qui ne nécessitent pas d'exemples:



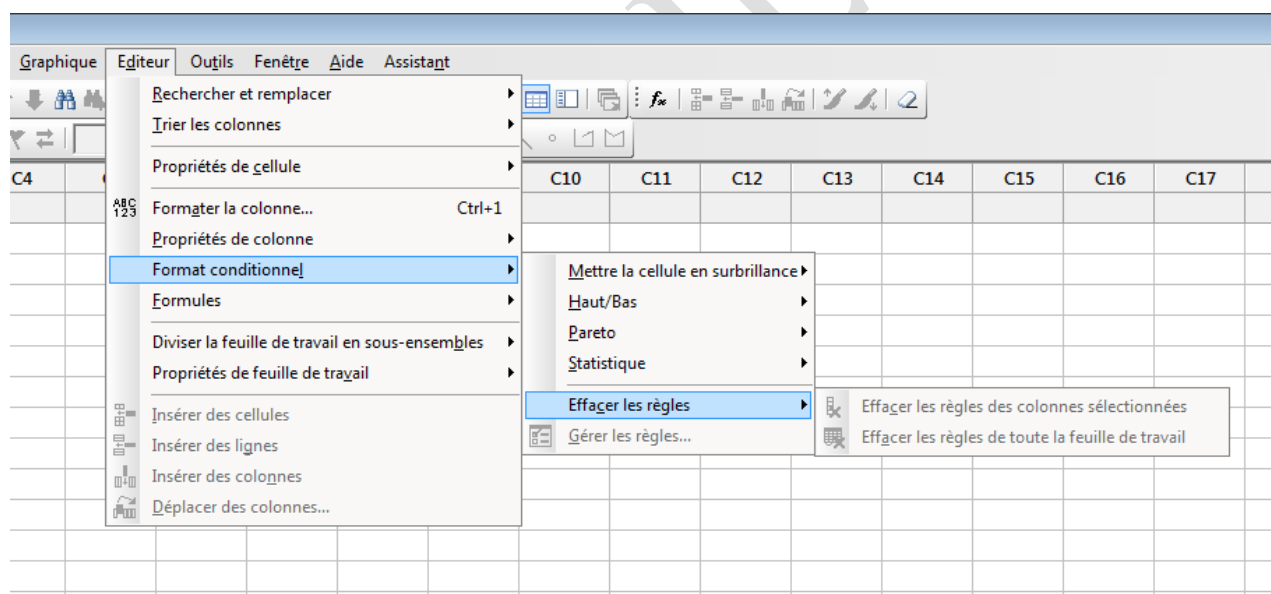
En encore des cas **Haut/Bas** là aussi identique à Microsoft Excel et ne nécessitant pas non plus d'exemples:



Et aussi le cas **Pareto** identique à Microsoft Excel et ne nécessitant pas non plus d'exemples:



Et enfin les options **Effacer les règles** et **Gérer les règles** qui sont basées aussi sur la même philosophie que Microsoft Excel.



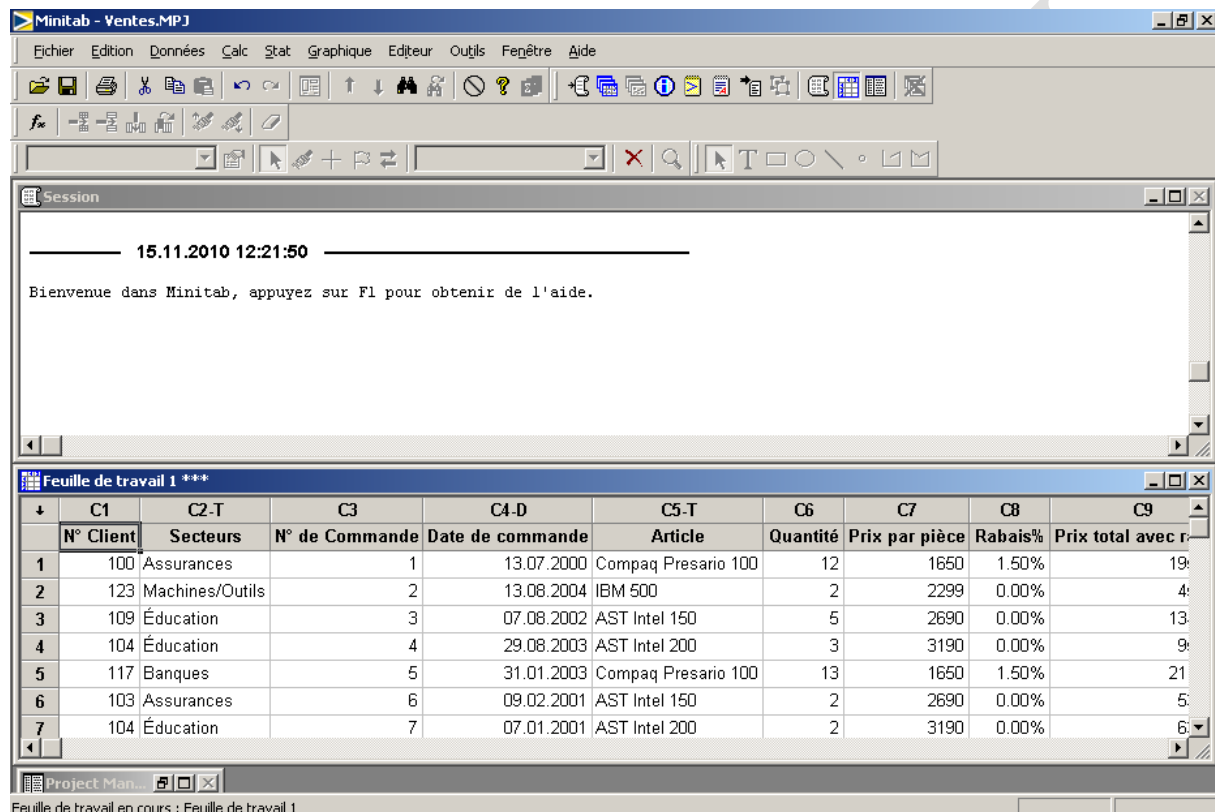
Il ne reste donc que la partie Statistique qui est particulière (j'écrirai quelque chose là-dessus quand j'aurai un peu plus de temps).

## 8. Graphiques

### 8.1. Exercice 9.: Graphique à barres

Minitab® Statistical Software 15.1.1.0

Ouvrez le fichier projet *Ventes.mpj*:

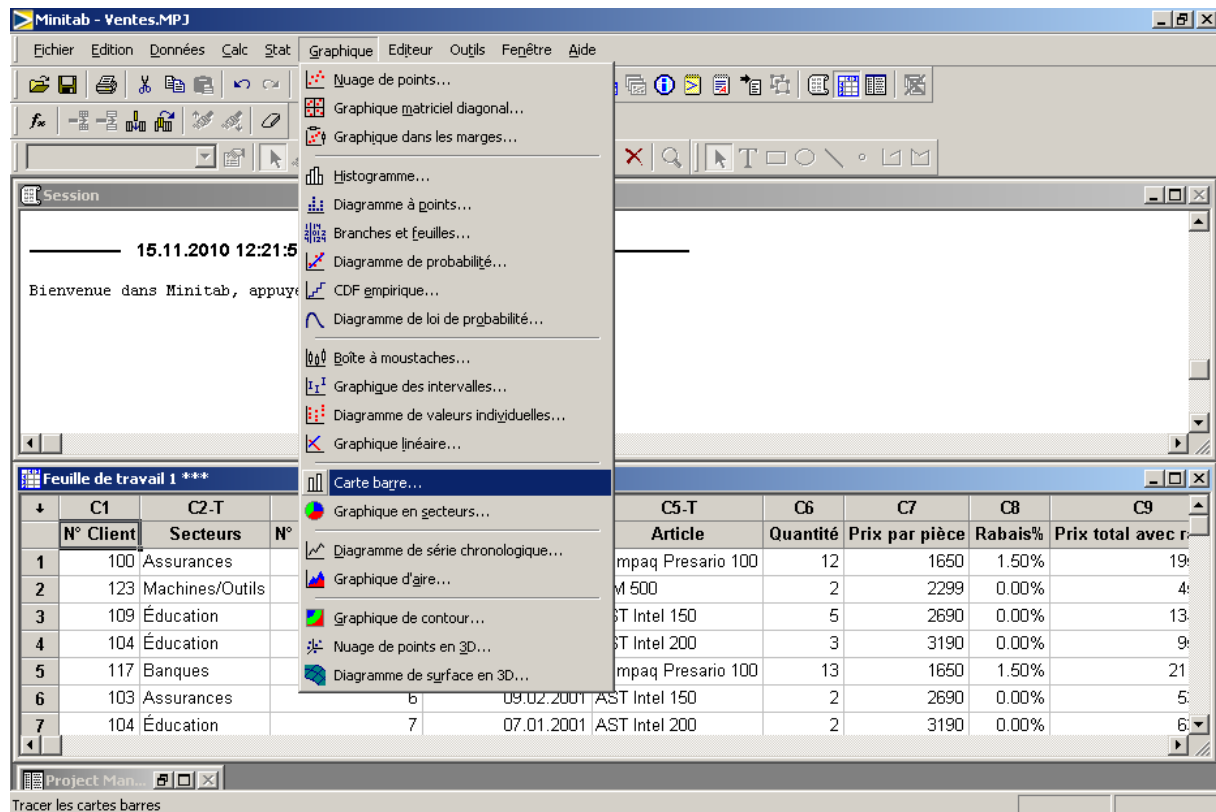


The screenshot shows the Minitab Statistical Software interface. The top window is the 'Session' window, which displays the date and time '15.11.2010 12:21:50' and a welcome message: 'Bienvenue dans Minitab, appuyez sur F1 pour obtenir de l'aide.' Below this is the 'Feuille de travail 1' window, which contains a data table with 10 columns (C1 to C9) and 7 rows of data.

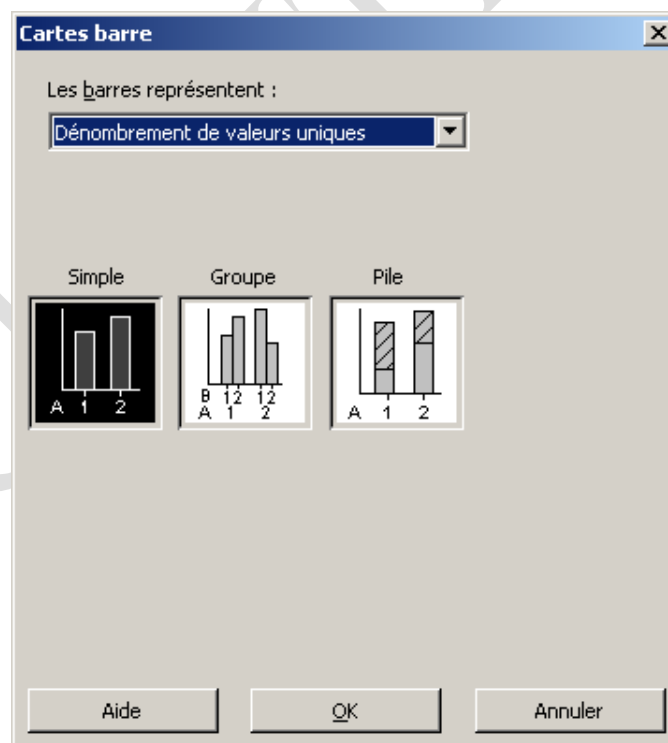
	C1	C2-T	C3	C4-D	C5-T	C6	C7	C8	C9
	N° Client	Secteurs	N° de Commande	Date de commande	Article	Quantité	Prix par pièce	Rabais%	Prix total avec r
1	100	Assurances	1	13.07.2000	Compaq Presario 100	12	1650	1.50%	19:
2	123	Machines/Outils	2	13.08.2004	IBM 500	2	2299	0.00%	4:
3	109	Éducation	3	07.08.2002	AST Intel 150	5	2690	0.00%	13:
4	104	Éducation	4	29.08.2003	AST Intel 200	3	3190	0.00%	9:
5	117	Banques	5	31.01.2003	Compaq Presario 100	13	1650	1.50%	21:
6	103	Assurances	6	09.02.2001	AST Intel 150	2	2690	0.00%	5:
7	104	Éducation	7	07.01.2001	AST Intel 200	2	3190	0.00%	6:

Nous souhaiterions effectuer quelques graphiques classiques de ce tableau mais que Microsoft Excel par exemple ne saurait pas effectuer sans un traitement préalable ou sans passer par un tableau croisé dynamique.

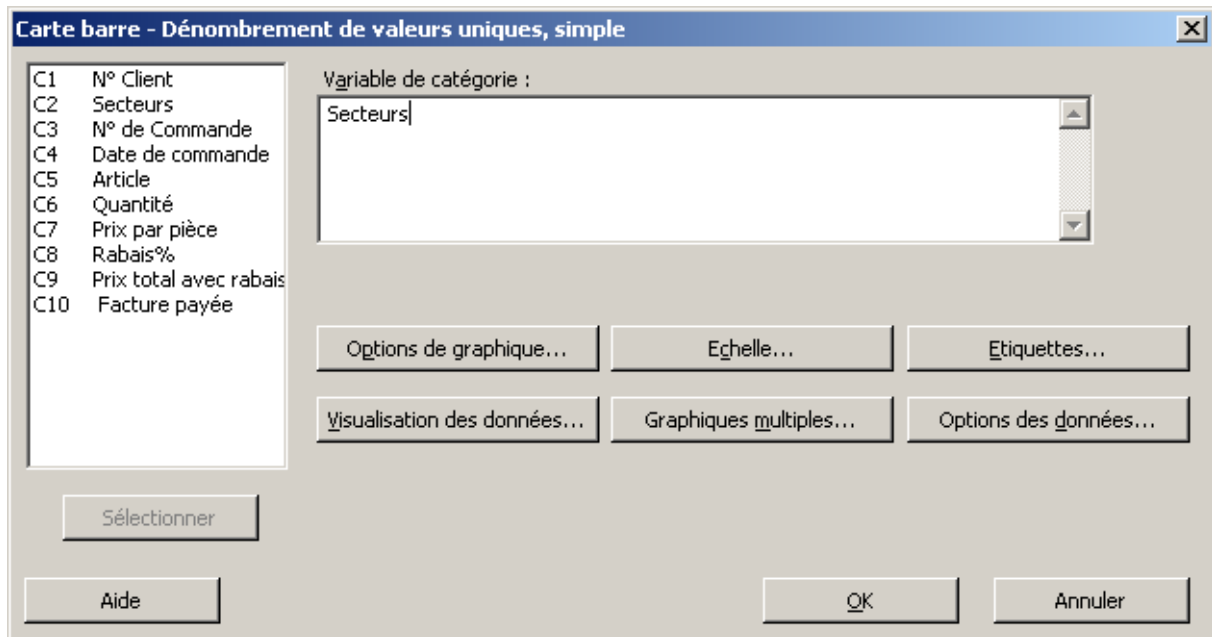
Faisons un premier graphique de type carte à barre représentant le nombre d'occurrence des secteurs dans tout le tableau. Pour ceci, allons dans le menu **Graphique/Carte barre...**:



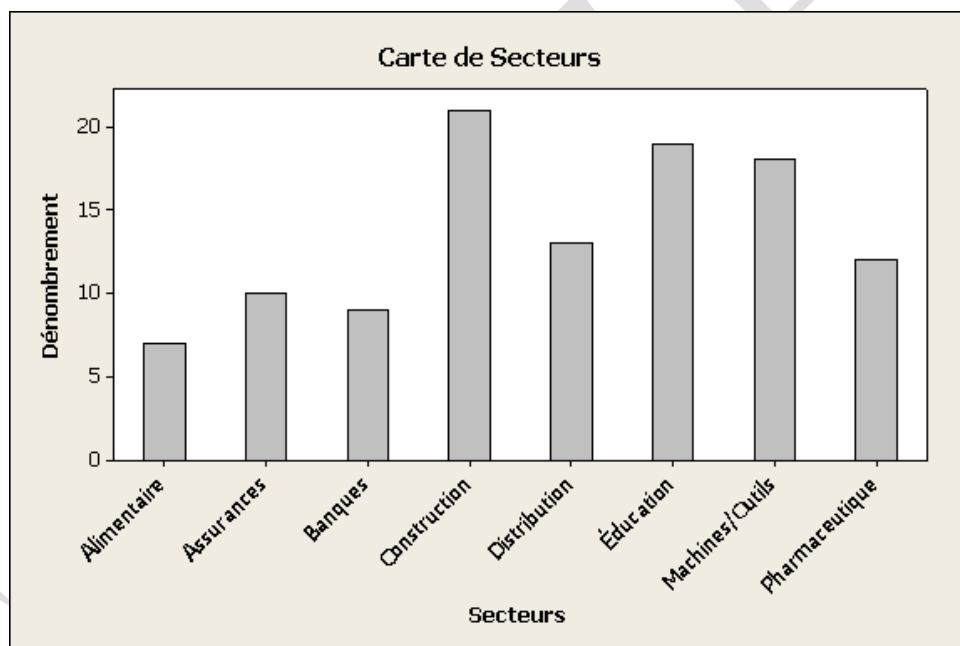
Nous prenons donc un **Dénombrement de valeurs uniques** de type simple:



Nous validons par **OK**:



Et nous revalidons par **OK**:





## 8.2. Exercice 10.: Graphique en secteurs

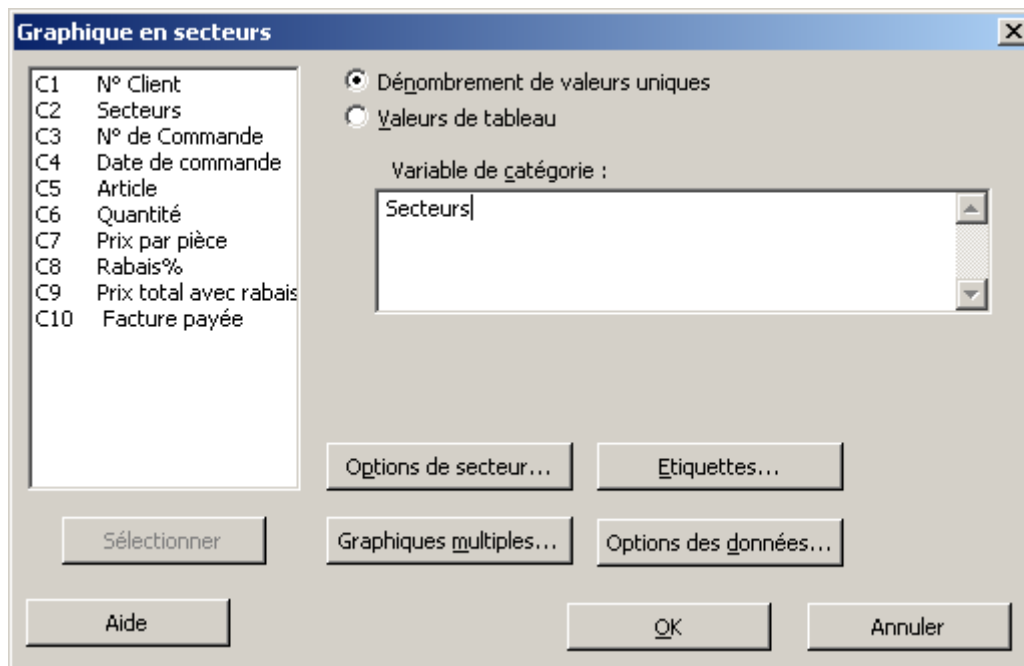
Minitab® Statistical Software 15.1.1.0

Faisons maintenant la même analyse toujours avec le même jeu de données avec un graphique en secteurs. Nous allons donc dans le menu **Graphique/Graphique en secteurs...**:

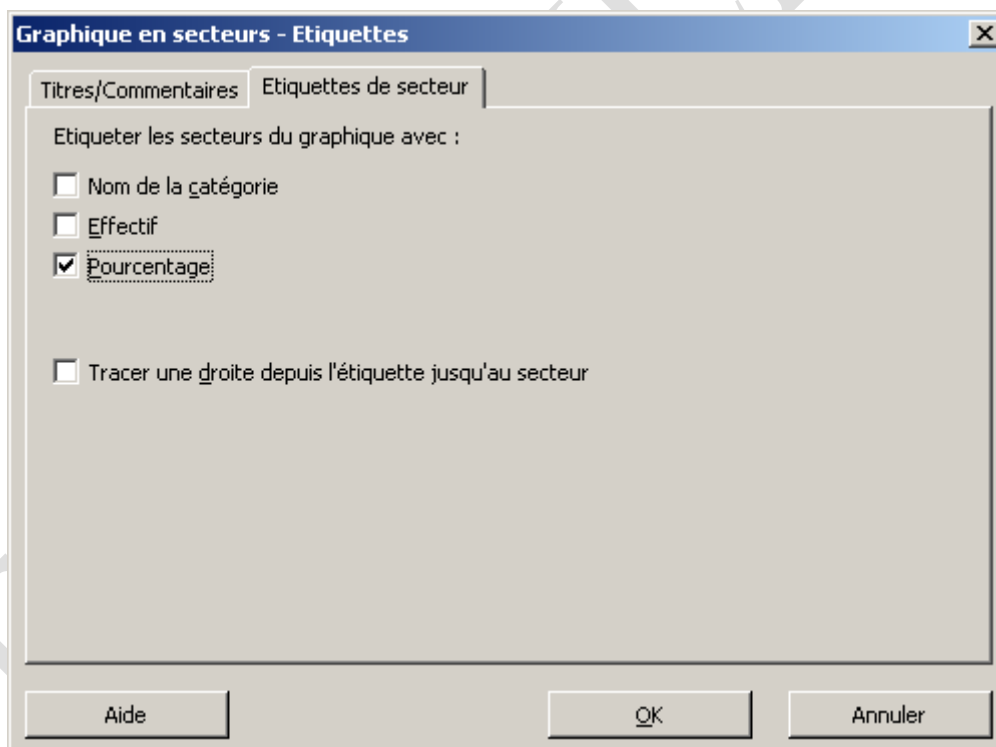
The screenshot shows the Minitab Statistical Software 15.1.1.0 interface. The 'Graphique' menu is open, and 'Graphique en secteurs...' is highlighted. The worksheet window displays a data table with the following content:

	C1	C2-T		C5-T	C6	C7	C8	C9
	N° Client	Secteurs	N°	Article	Quantité	Prix par pièce	Rabais%	Prix total avec r
1	100	Assurances		mpaq Presario 100	12	1650	1.50%	19:
2	123	Machines/Outils		M 500	2	2299	0.00%	4:
3	109	Éducation		ST Intel 150	5	2690	0.00%	13:
4	104	Éducation		ST Intel 200	3	3190	0.00%	9:
5	117	Banques		mpaq Presario 100	13	1650	1.50%	21:
6	103	Assurances	6	09.02.2001 AST Intel 150	2	2690	0.00%	5:
7	104	Éducation	7	07.01.2001 AST Intel 200	2	3190	0.00%	6:

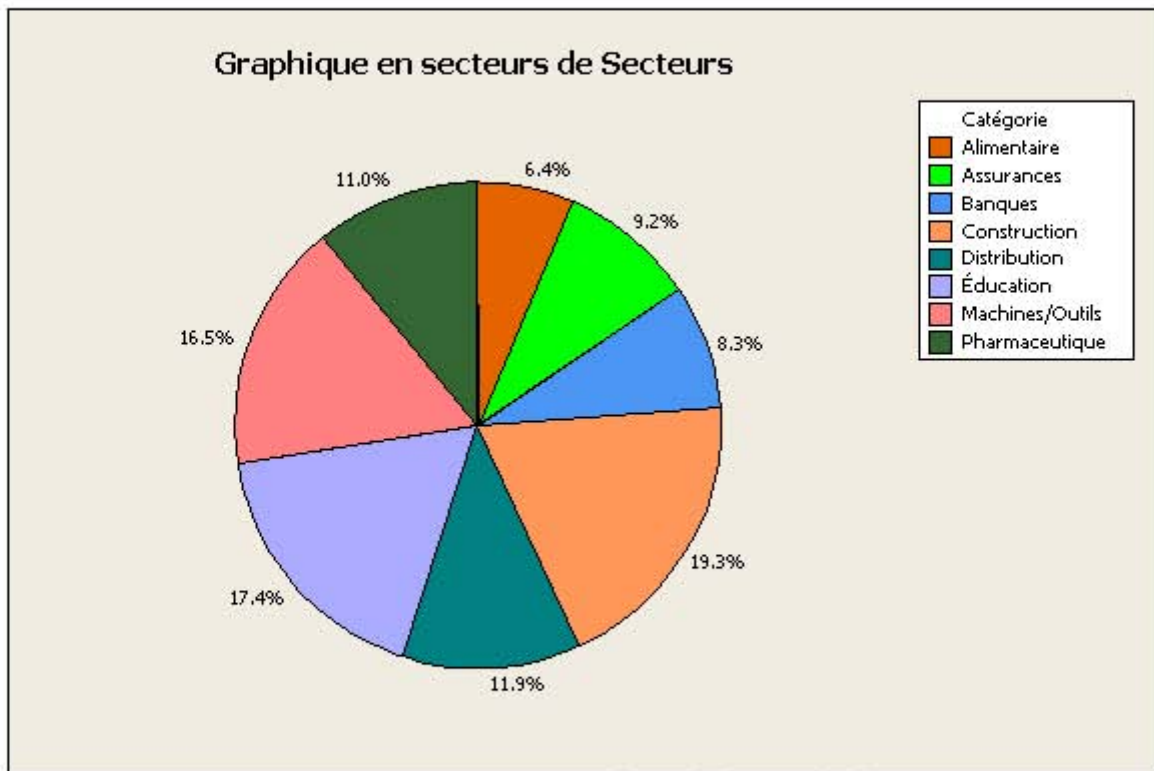
Nous avons alors:



Nous allons dans le bouton **Etiquettes**:



Pour cocher **Pourcentages**. Nous validons deux fois par **OK** pour obtenir au final:



### 8.3. Exercice 11.: Diagramme à points

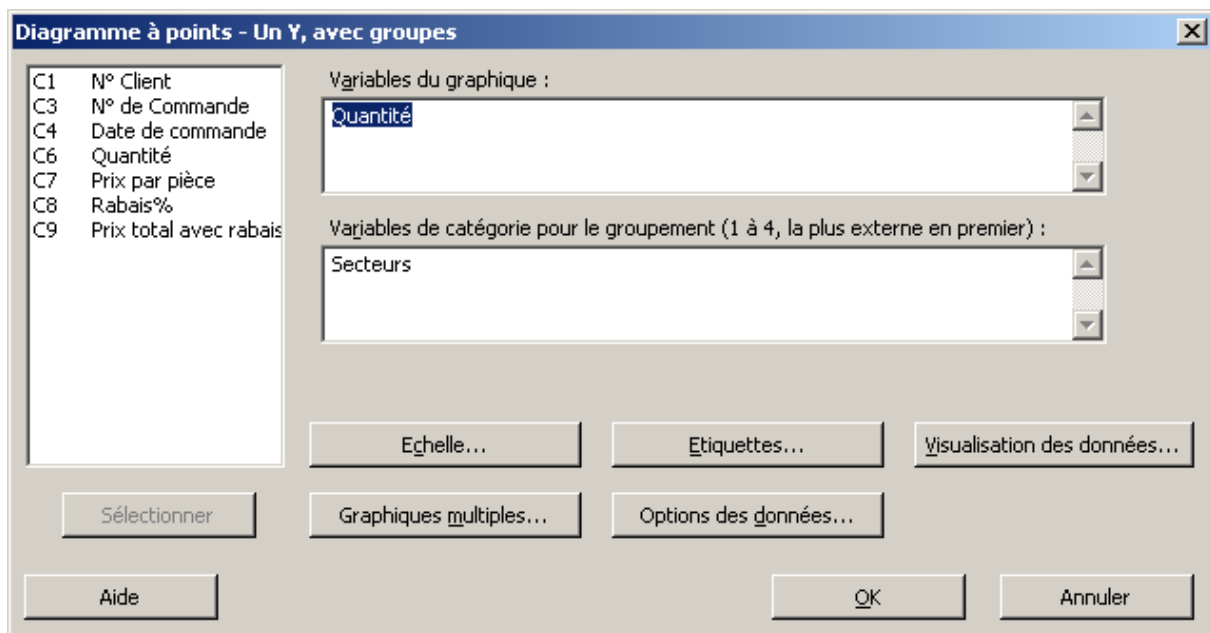
Minitab® Statistical Software 15.1.1.0

Faisons maintenant un diagramme à points et ce toujours avec le même jeu de données:

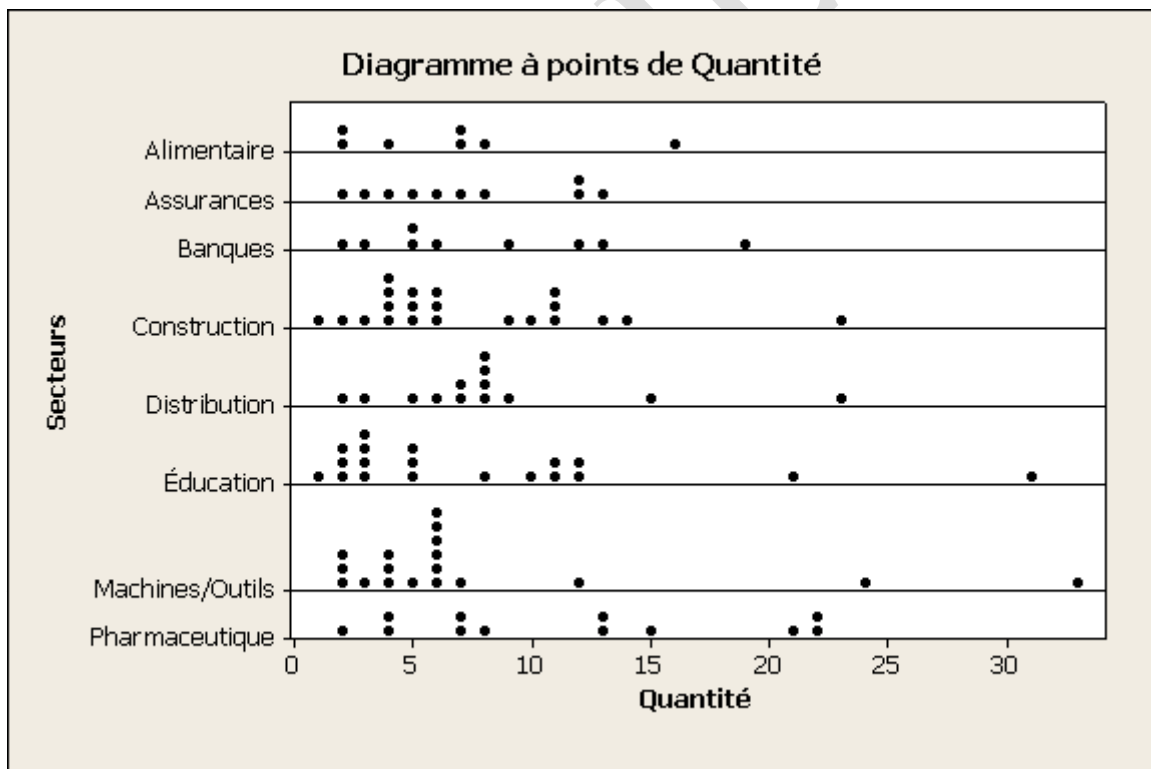
	C1	C2-T		C5-T	C6	C7	C8	C9
	N° Client	Secteurs	N°	Article	Quantité	Prix par pièce	Rabais%	Prix total avec r
1	100	Assurances		Compaq Presario 100	12	1650	1.50%	19
2	123	Machines/Outils		M 500	2	2299	0.00%	4
3	109	Éducation		AST Intel 150	5	2690	0.00%	13
4	104	Éducation		AST Intel 200	3	3190	0.00%	9
5	117	Banques	5	31.01.2003 Compaq Presario 100	13	1650	1.50%	21
6	103	Assurances	6	09.02.2001 AST Intel 150	2	2690	0.00%	5
7	104	Éducation	7	07.01.2001 AST Intel 200	2	3190	0.00%	6

De type:

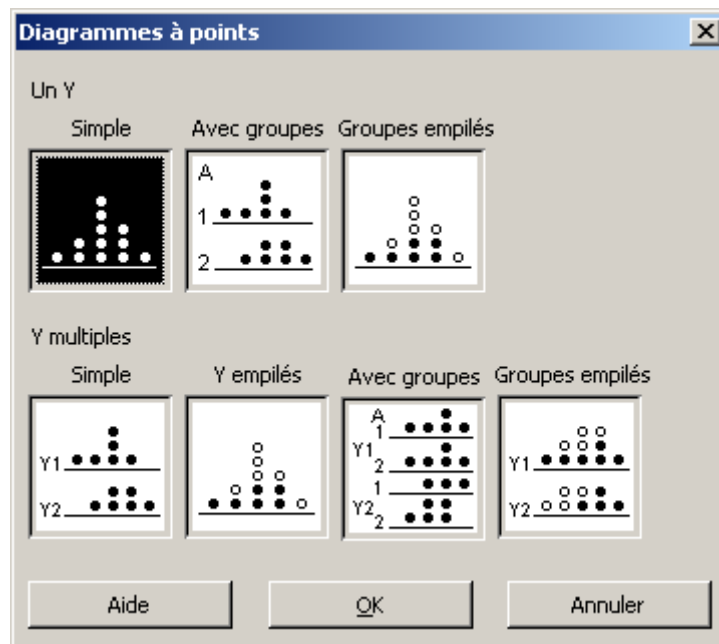
Avec:



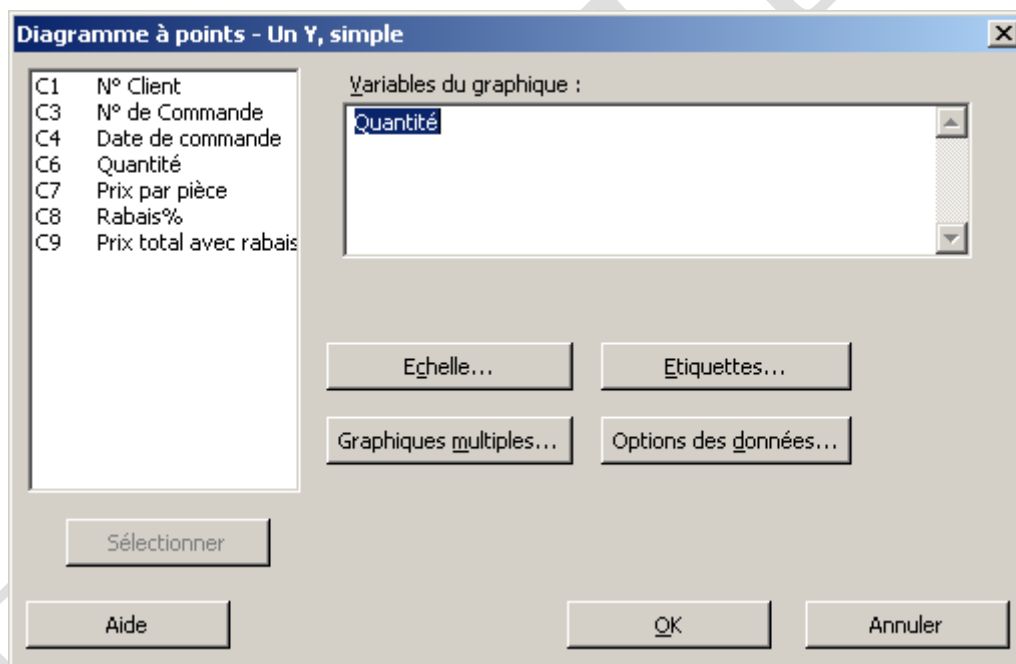
Et validons par **OK**:



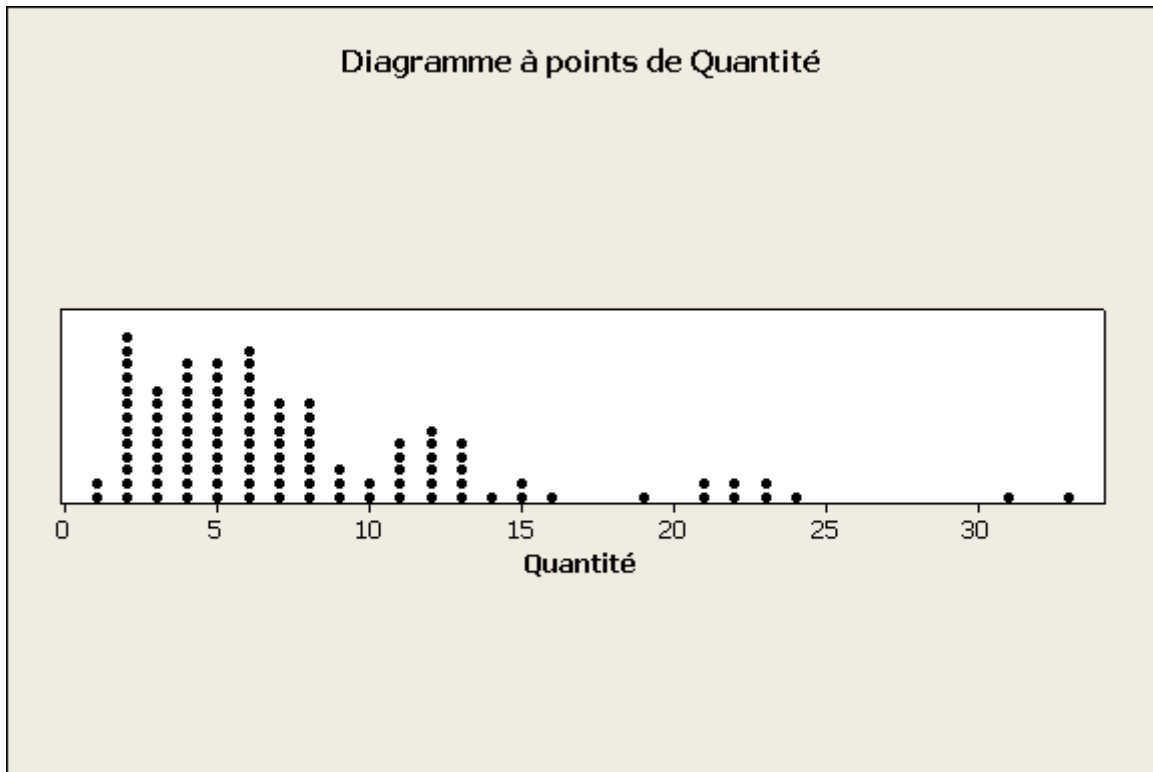
Nous voyons bien les distributions. Refaisons le même mais cette fois-ci sans groupes:



Nous validons par **OK** et faisons l'analyse sur la Quantité:



Pour obtenir:



Chaque unité d'abscisse à ainsi son analyse fréquentielle. Evidemment nous pouvons obtenir la même chose que le graphique précédent en jouant un peu avec les cartes à barre:

Minitab - Ventres.MPJ

Graphique

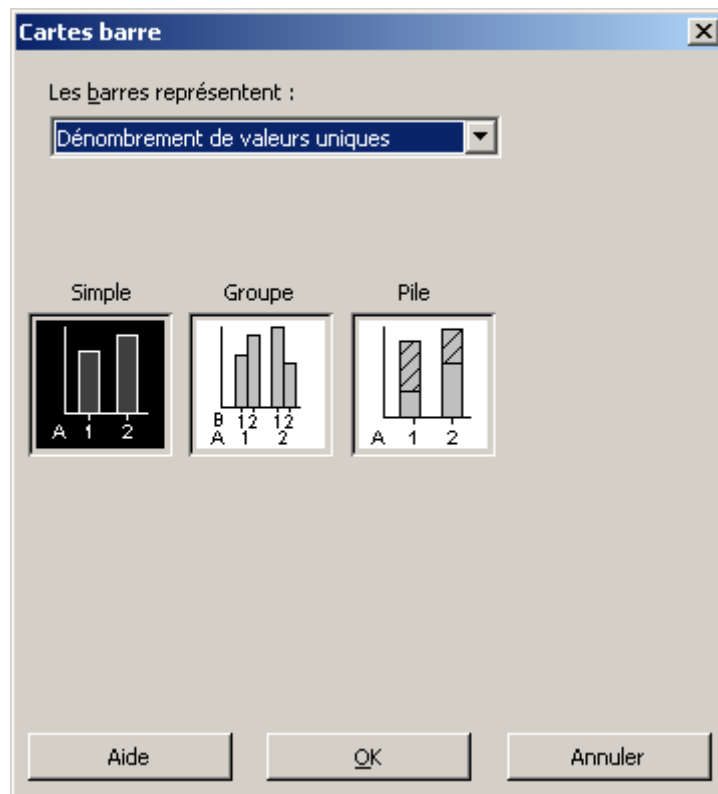
- Nuage de points...
- Graphique matriciel diagonal...
- Graphique dans les marges...
- Histogramme...
- Diagramme à points...
- Branches et feuilles...
- Diagramme de probabilité...
- CDF empirique...
- Diagramme de loi de probabilité...
- Boîte à moustaches...
- Graphique des intervalles...
- Diagramme de valeurs individuelles...
- Graphique linéaire...
- Carte barre...**
- Graphique en secteurs...
- Diagramme de série chronologique...
- Graphique d'aire...
- Graphique de contour...
- Nuage de points en 3D...
- Diagramme de surface en 3D...

	C1	C2-T		C5-T	C6	C7	C8	C9
	N° Client	Secteurs	N°	Article	Quantité	Prix par pièce	Rabais%	Prix total avec r
1	100	Assurances		Compaq Presario 100	12	1650	1.50%	19
2	123	Machines/Outils		M 500	2	2299	0.00%	4
3	109	Éducation		AST Intel 150	5	2690	0.00%	13
4	104	Éducation		AST Intel 200	3	3190	0.00%	9
5	117	Banques	5	31.01.2003 Compaq Presario 100	13	1650	1.50%	21
6	103	Assurances	6	09.02.2001 AST Intel 150	2	2690	0.00%	5
7	104	Éducation	7	07.01.2001 AST Intel 200	2	3190	0.00%	6

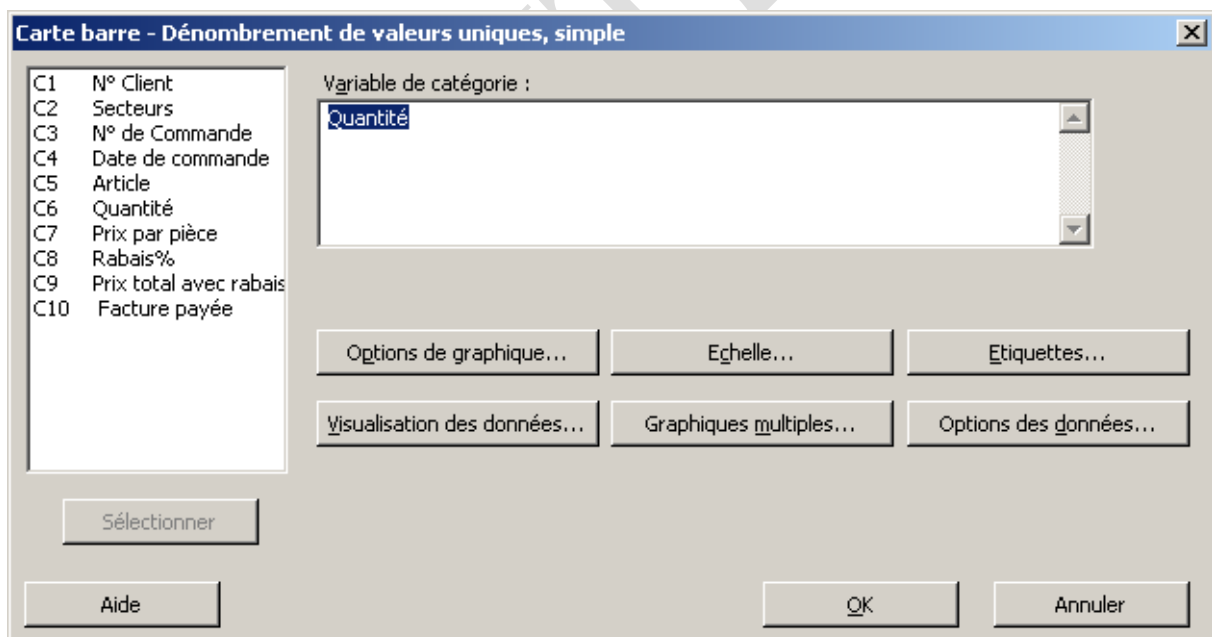
Tracer les cartes barres

Démarrer Minitab - Microsoft Word Minitab - Ventres.MPJ 14:00

En prenant:

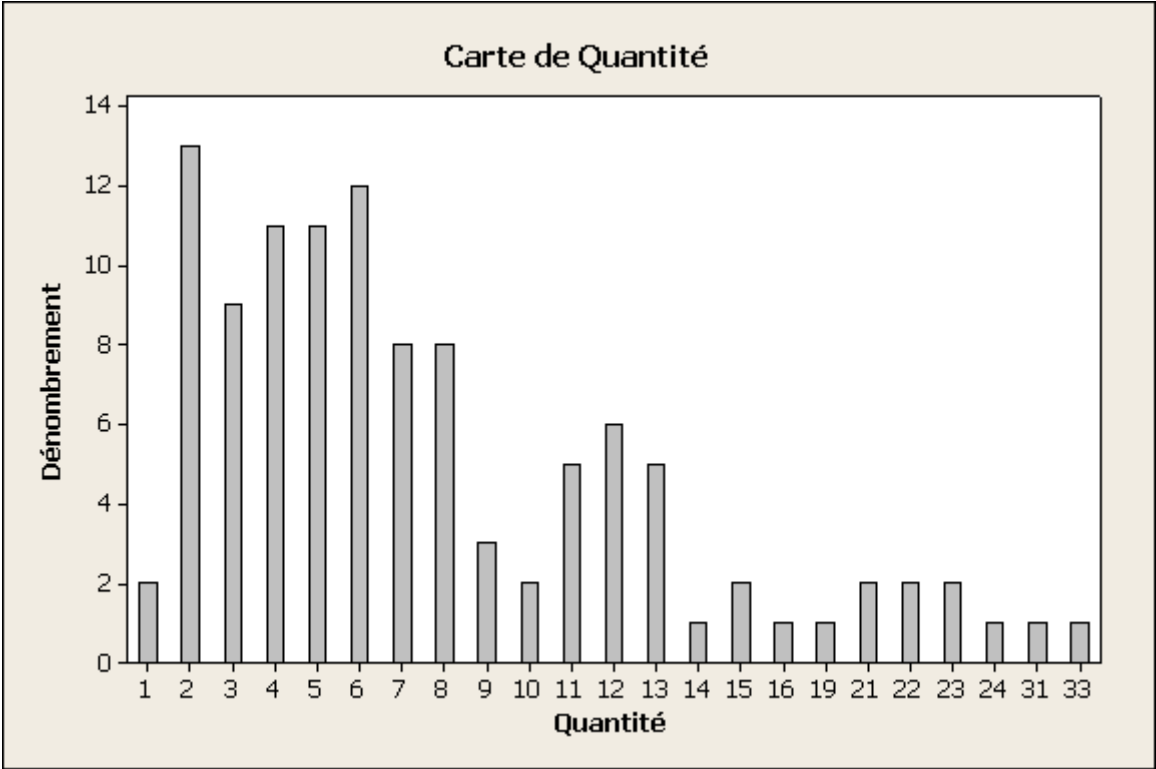


En prenant les *Quantités*:



Et en validant par **OK**:





Bref ce n'est qu'un choix esthétique en fin de compte...

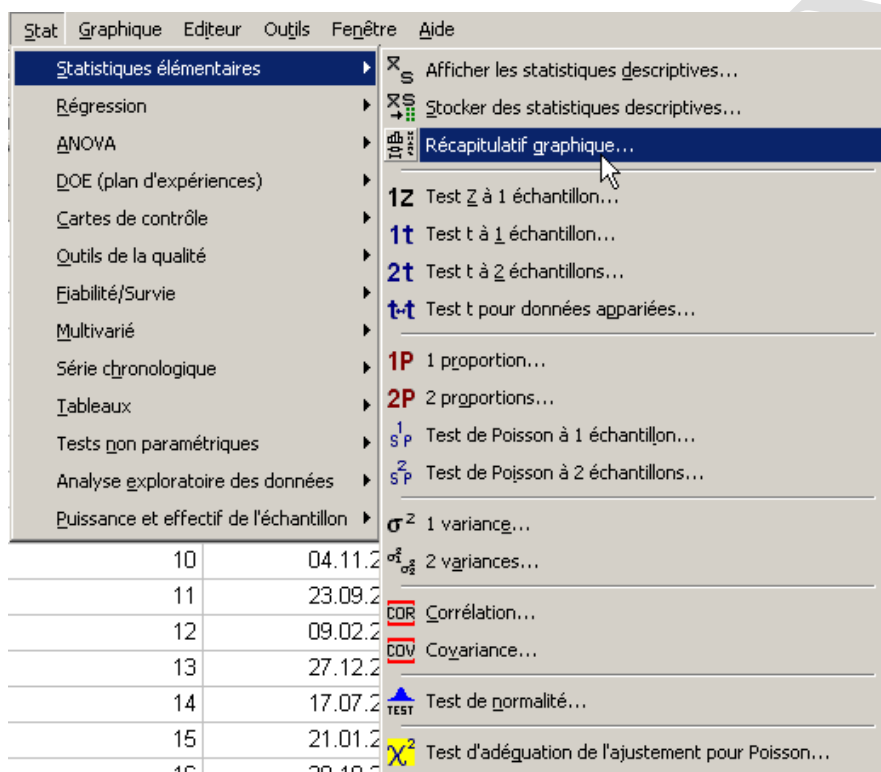
ÉCHANTILLON

## 8.4. Exercice 12.: Récapitulatif graphique de statistiques

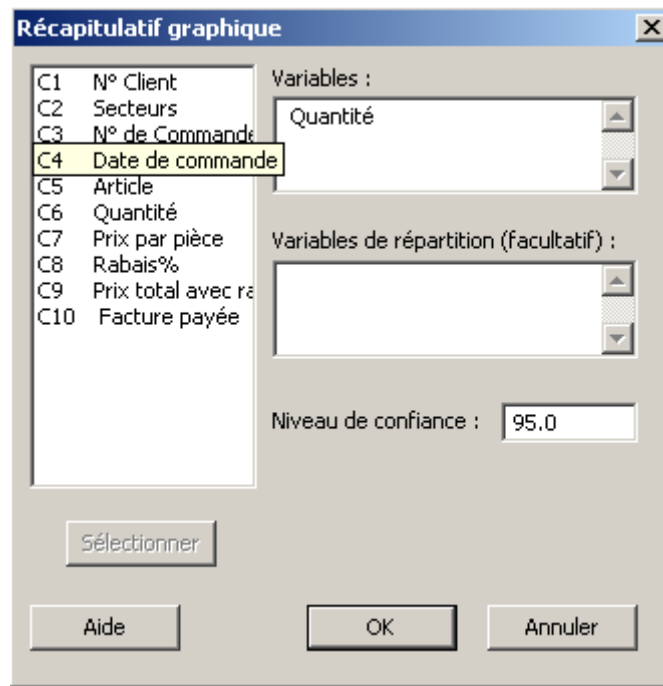
Minitab® Statistical Software 15.1.1.0

Enfin, voici mon graphique préféré toujours avec le même jeu de données (même s'il est relativement facile à faire aussi avec Microsoft Excel, cela reste quand même mon graphique préféré...).

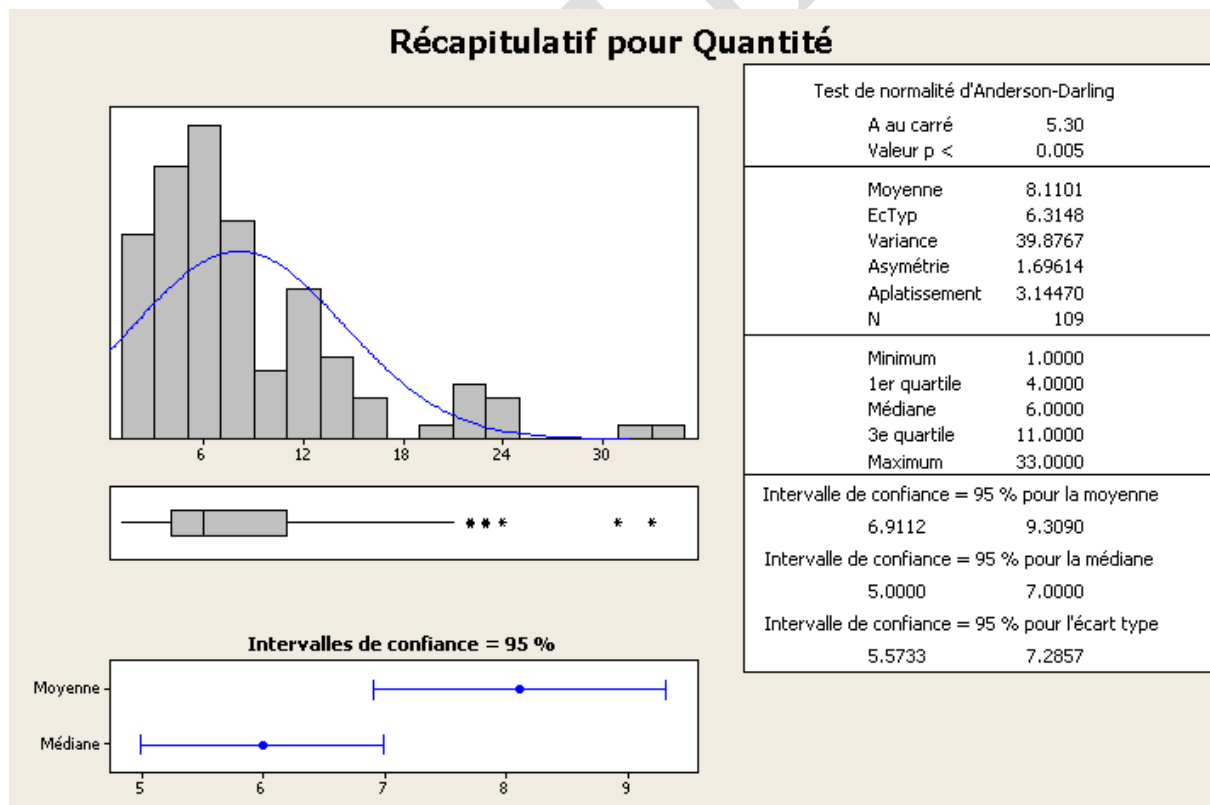
Donc toujours avec le même fichier, nous allons dans le menu **Statistiques élémentaires/Récapitulatif graphique...**:



Nous avons alors:

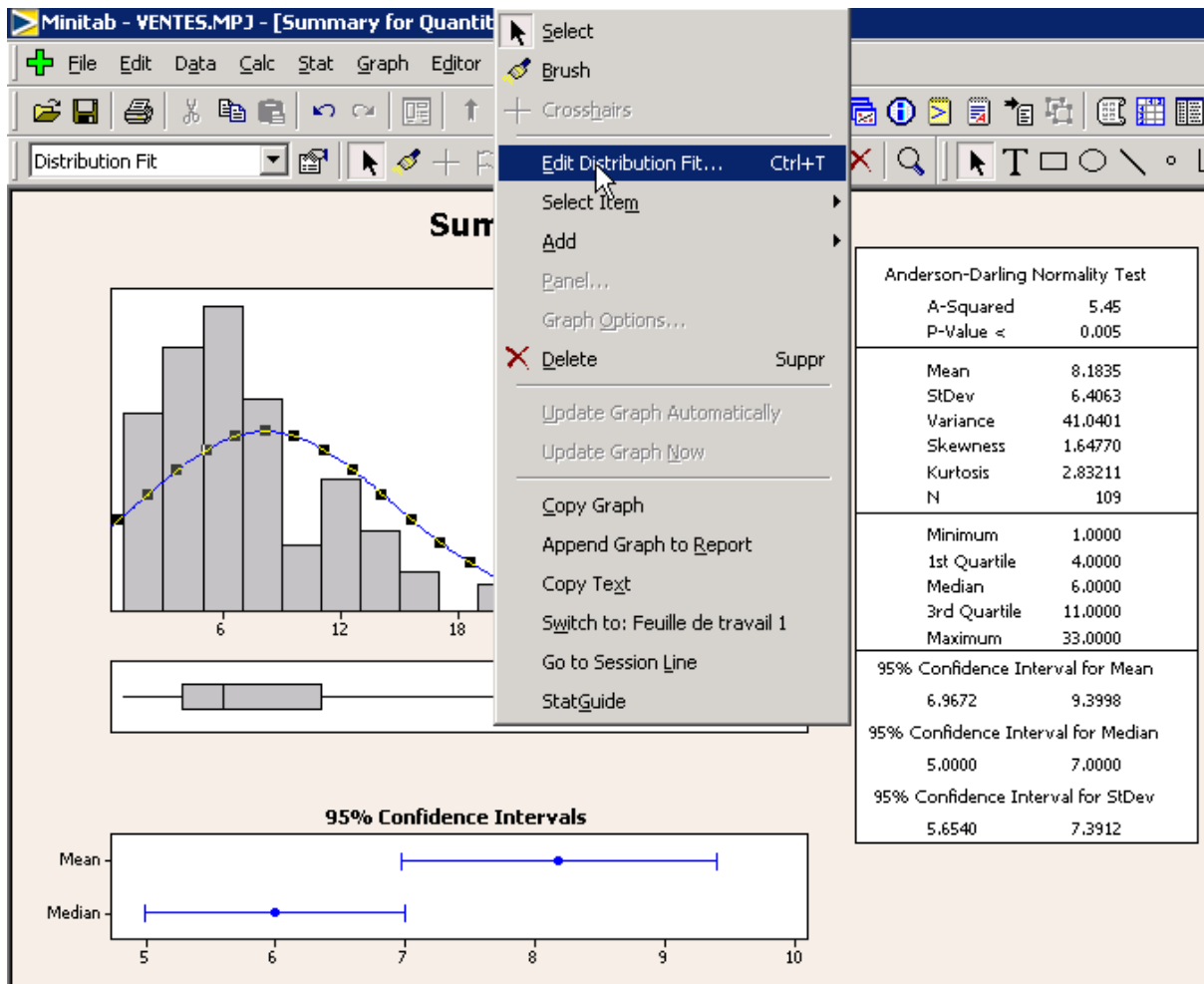


Le champ **Variables de répartition** vous permet de créer des graphiques à la volée en fonction des valeurs nominales de la variable de répartition. Nous n'allons pas l'utiliser ici. Nous validons par **OK** pour obtenir:



Rien ne nous empêche de faire un histogramme mais nous avons déjà vu cela plus haut.

Pour changer l'ajustement de la distribution ci-dessus cela ne peut se faire lors de la création du graphique mais seulement après:

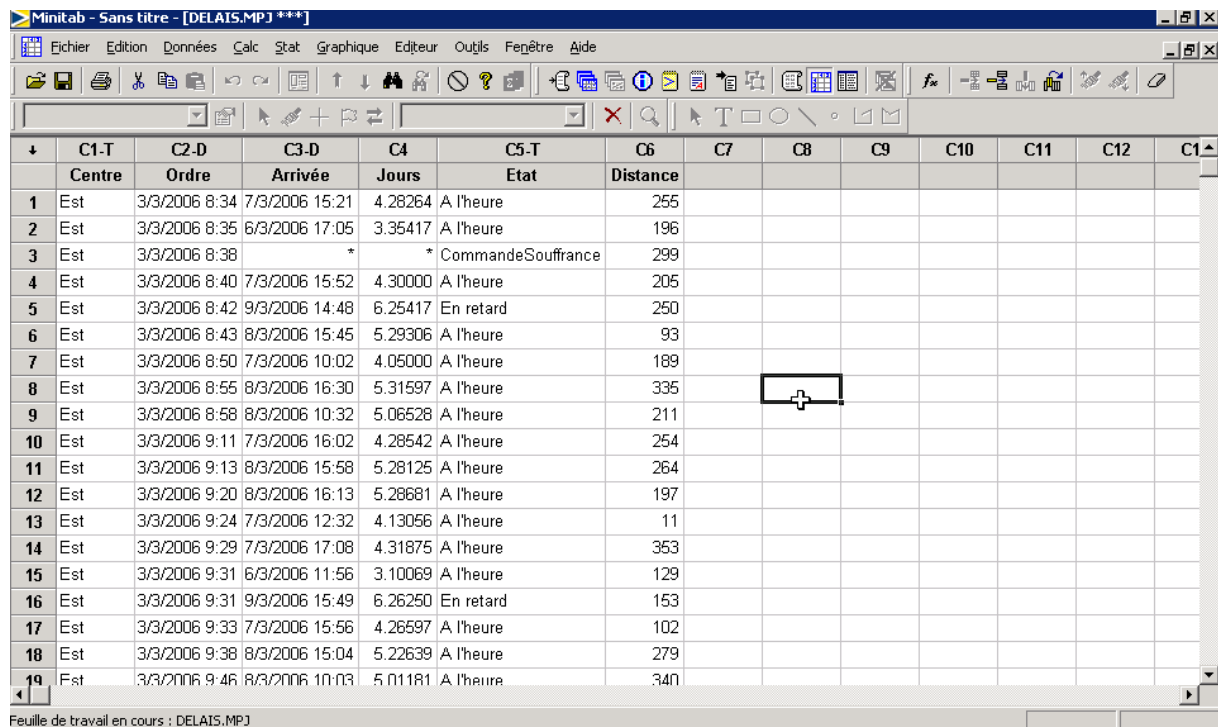


**Le problème c'est que si vous poursuivez... certes le type de distribution sera changé mais la A<sup>2</sup> et la p-value ne seront pas recalculées!!!**

## 8.5. Exercice 13.: Diagramme de valeurs individuelles

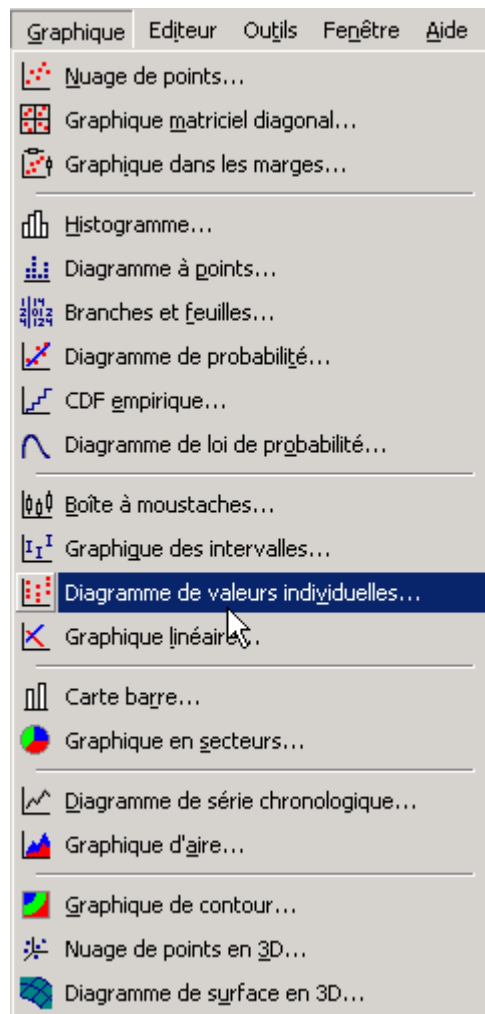
Minitab® Statistical Software 15.1.1.0

Pour la suite, nous allons reprendre à l'identique un exemple fourni par l'aide de Minitab® Statistical Software (oui comme ce n'est pas des maths je n'ai pas envie de recréer la roue) en ouvrant le fichier *Delais.mpj*:

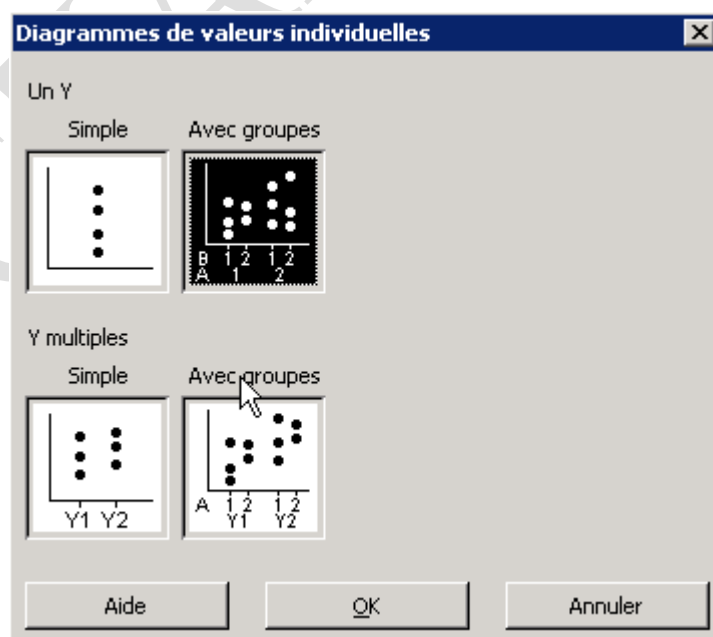


	C1-T	C2-D	C3-D	C4	C5-T	C6	C7	C8	C9	C10	C11	C12	C13
	Centre	Ordre	Arrivée	Jours	Etat	Distance							
1	Est	3/3/2006 8:34	7/3/2006 15:21	4.28264	A l'heure	255							
2	Est	3/3/2006 8:35	6/3/2006 17:05	3.35417	A l'heure	196							
3	Est	3/3/2006 8:38	*	* CommandeSouffrance		299							
4	Est	3/3/2006 8:40	7/3/2006 15:52	4.30000	A l'heure	205							
5	Est	3/3/2006 8:42	9/3/2006 14:48	6.25417	En retard	250							
6	Est	3/3/2006 8:43	8/3/2006 15:45	5.29306	A l'heure	93							
7	Est	3/3/2006 8:50	7/3/2006 10:02	4.05000	A l'heure	189							
8	Est	3/3/2006 8:55	8/3/2006 16:30	5.31597	A l'heure	335							
9	Est	3/3/2006 8:58	8/3/2006 10:32	5.06528	A l'heure	211							
10	Est	3/3/2006 9:11	7/3/2006 16:02	4.28542	A l'heure	254							
11	Est	3/3/2006 9:13	8/3/2006 15:58	5.28125	A l'heure	264							
12	Est	3/3/2006 9:20	8/3/2006 16:13	5.28681	A l'heure	197							
13	Est	3/3/2006 9:24	7/3/2006 12:32	4.13056	A l'heure	11							
14	Est	3/3/2006 9:29	7/3/2006 17:08	4.31875	A l'heure	353							
15	Est	3/3/2006 9:31	6/3/2006 11:56	3.10069	A l'heure	129							
16	Est	3/3/2006 9:31	9/3/2006 15:49	6.26250	En retard	153							
17	Est	3/3/2006 9:33	7/3/2006 15:56	4.26597	A l'heure	102							
18	Est	3/3/2006 9:38	8/3/2006 15:04	5.22639	A l'heure	279							
19	Est	3/3/2006 9:46	8/3/2006 10:03	5.01181	A l'heure	340							

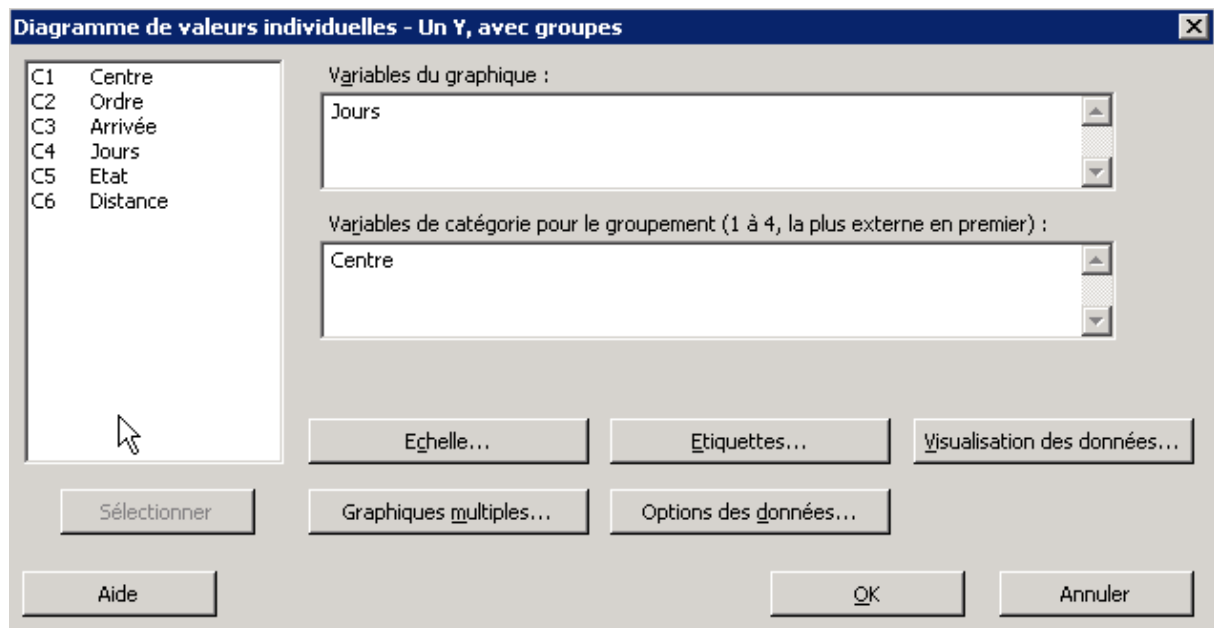
et allez dans le menu **Graphique/Diagramme de valeurs individuelles...**:



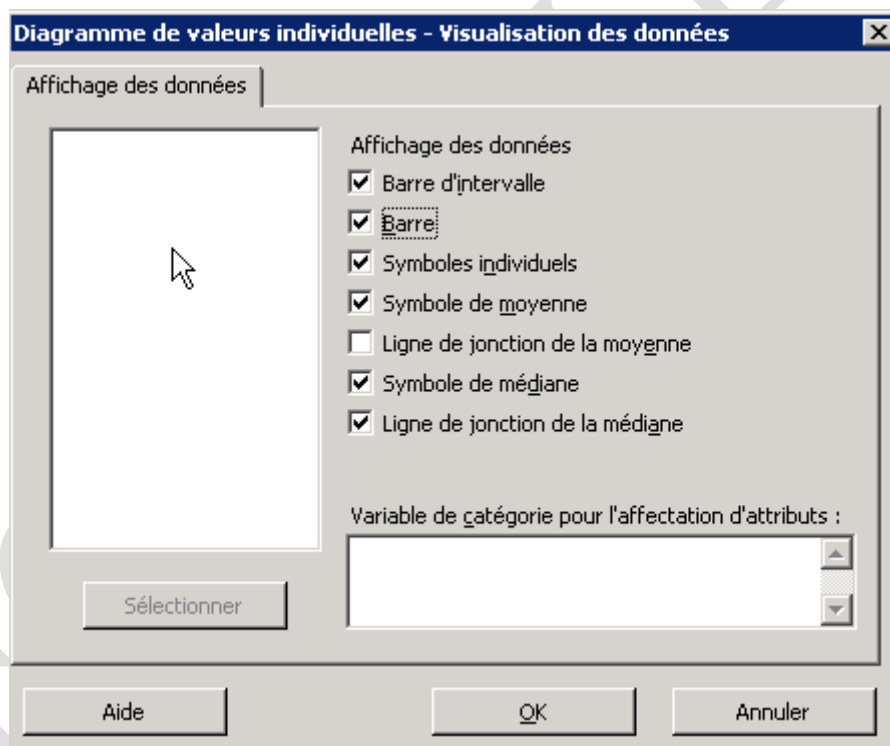
Prenez Avec groupes:



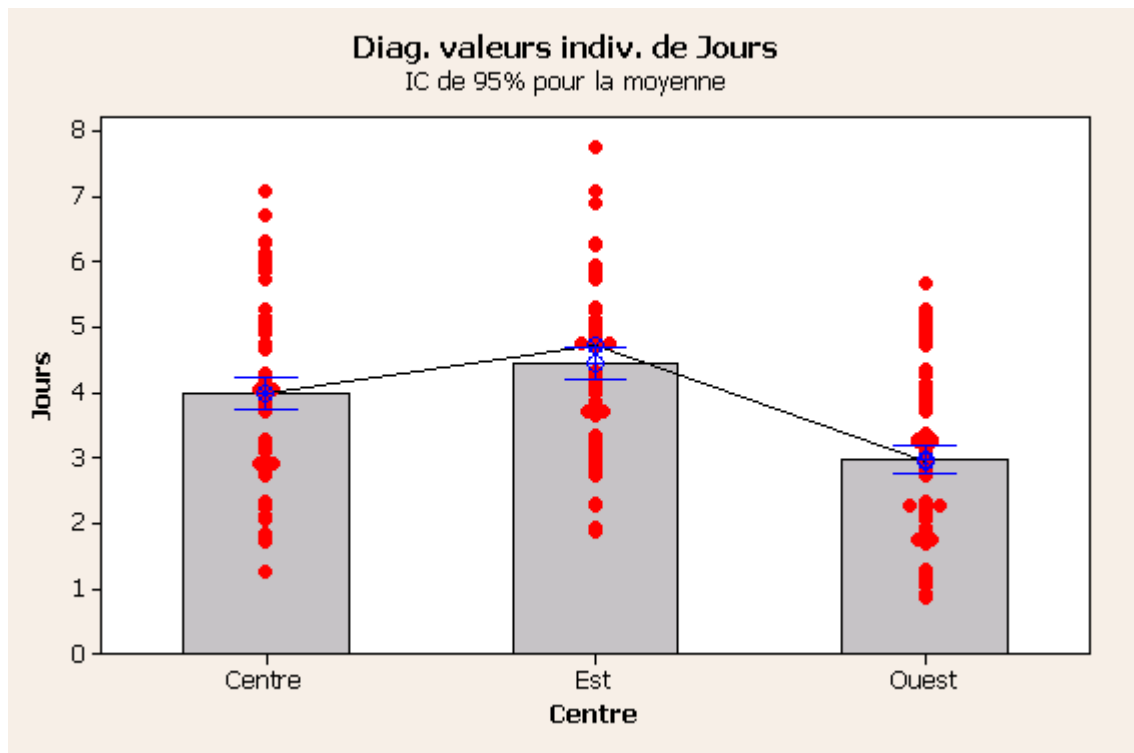
Valiez par **OK** et choisissez les variables comme indiqué ci-dessous:



Cliquez sur le bouton **Visualisation des données...** et cochez la même chose que ci-dessous:



Validez par **OK** deux fois:



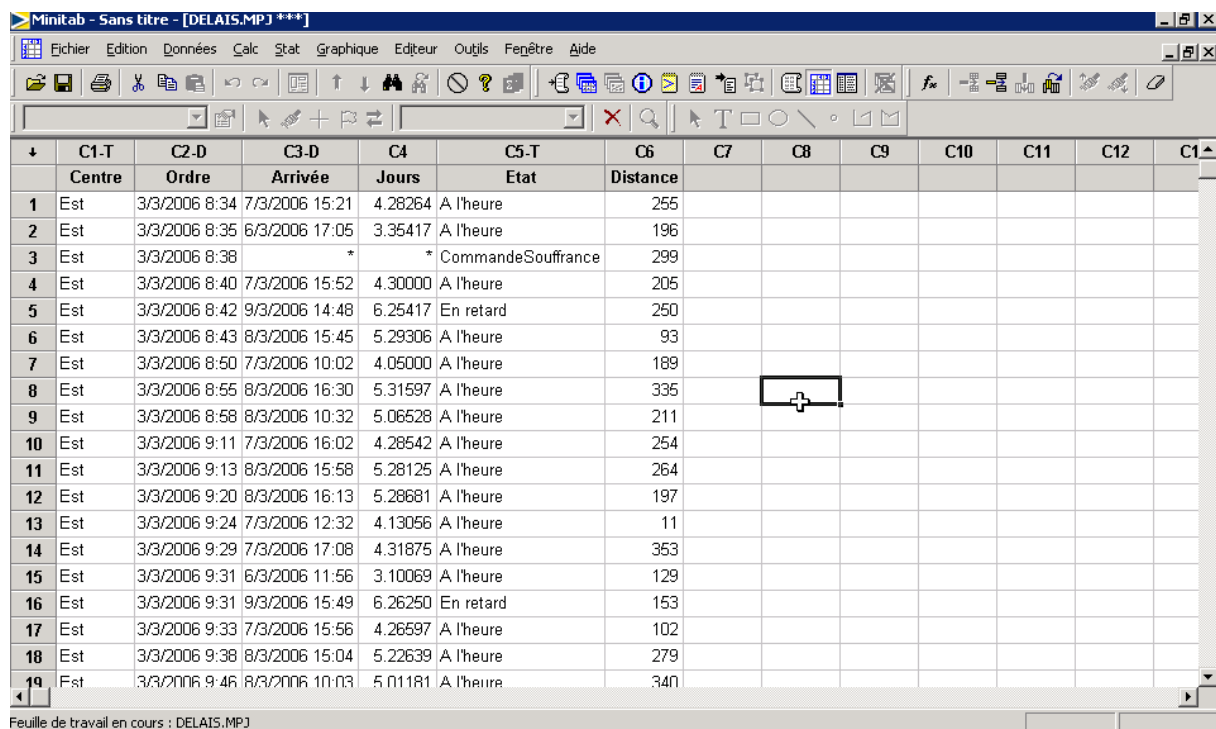
Nous avons donc un résultat qu'il est difficilement à obtenir avec un tableur. Surtout la partie concernant les points qui sont regroupés sur une même verticale sur une variable catégorielle de type texte (le reste étant facilement faisable avec un tableur).



## 8.6. Exercice 14.: Histogrammes simples ou multiples avec ajustements

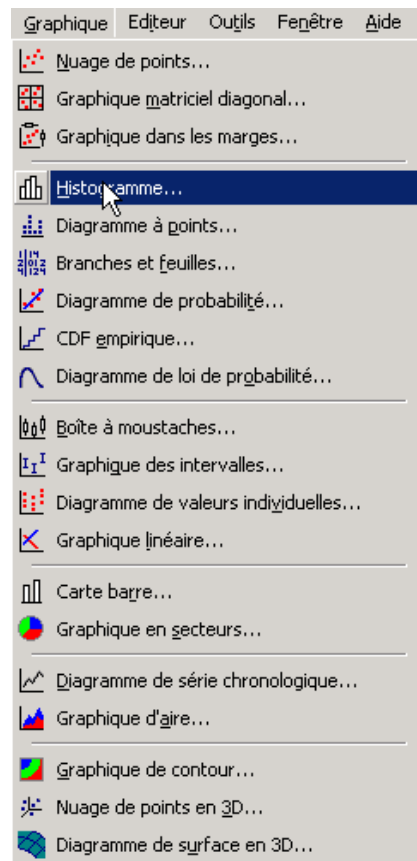
Minitab® Statistical Software 15.1.1.0

Toujours avec le même jeu de données:

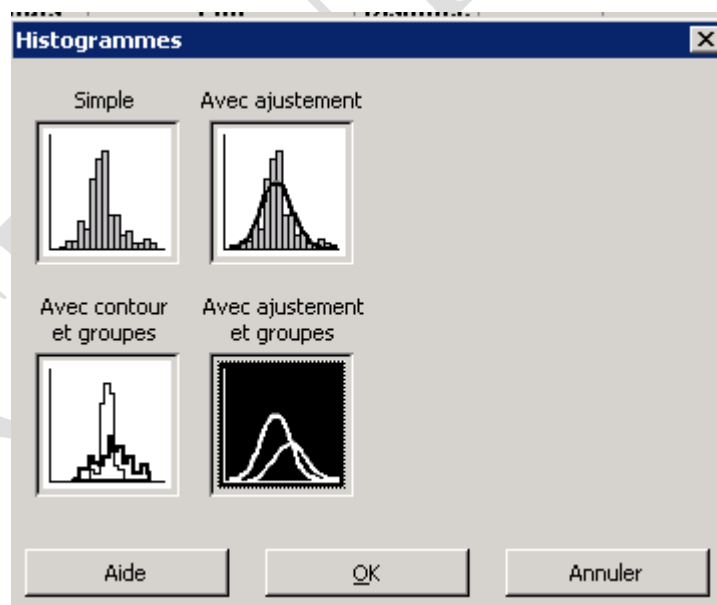


	C1-T	C2-D	C3-D	C4	C5-T	C6	C7	C8	C9	C10	C11	C12	C13
	Centre	Ordre	Arrivée	Jours	Etat	Distance							
1	Est	3/3/2006 8:34	7/3/2006 15:21	4.28264	A l'heure	255							
2	Est	3/3/2006 8:36	6/3/2006 17:05	3.35417	A l'heure	196							
3	Est	3/3/2006 8:38	*	*	CommandeSouffrance	299							
4	Est	3/3/2006 8:40	7/3/2006 15:52	4.30000	A l'heure	205							
5	Est	3/3/2006 8:42	9/3/2006 14:48	6.25417	En retard	250							
6	Est	3/3/2006 8:43	8/3/2006 15:45	5.29306	A l'heure	93							
7	Est	3/3/2006 8:50	7/3/2006 10:02	4.05000	A l'heure	189							
8	Est	3/3/2006 8:55	8/3/2006 16:30	5.31597	A l'heure	335							
9	Est	3/3/2006 8:58	8/3/2006 10:32	5.06528	A l'heure	211							
10	Est	3/3/2006 9:11	7/3/2006 16:02	4.28542	A l'heure	254							
11	Est	3/3/2006 9:13	8/3/2006 15:58	5.28125	A l'heure	264							
12	Est	3/3/2006 9:20	8/3/2006 16:13	5.28681	A l'heure	197							
13	Est	3/3/2006 9:24	7/3/2006 12:32	4.13056	A l'heure	11							
14	Est	3/3/2006 9:29	7/3/2006 17:08	4.31875	A l'heure	353							
15	Est	3/3/2006 9:31	6/3/2006 11:56	3.10069	A l'heure	129							
16	Est	3/3/2006 9:31	9/3/2006 15:49	6.26250	En retard	153							
17	Est	3/3/2006 9:33	7/3/2006 15:56	4.26597	A l'heure	102							
18	Est	3/3/2006 9:38	8/3/2006 15:04	5.22639	A l'heure	279							
19	Est	3/3/2006 9:46	8/3/2006 10:03	5.01181	A l'heure	340							

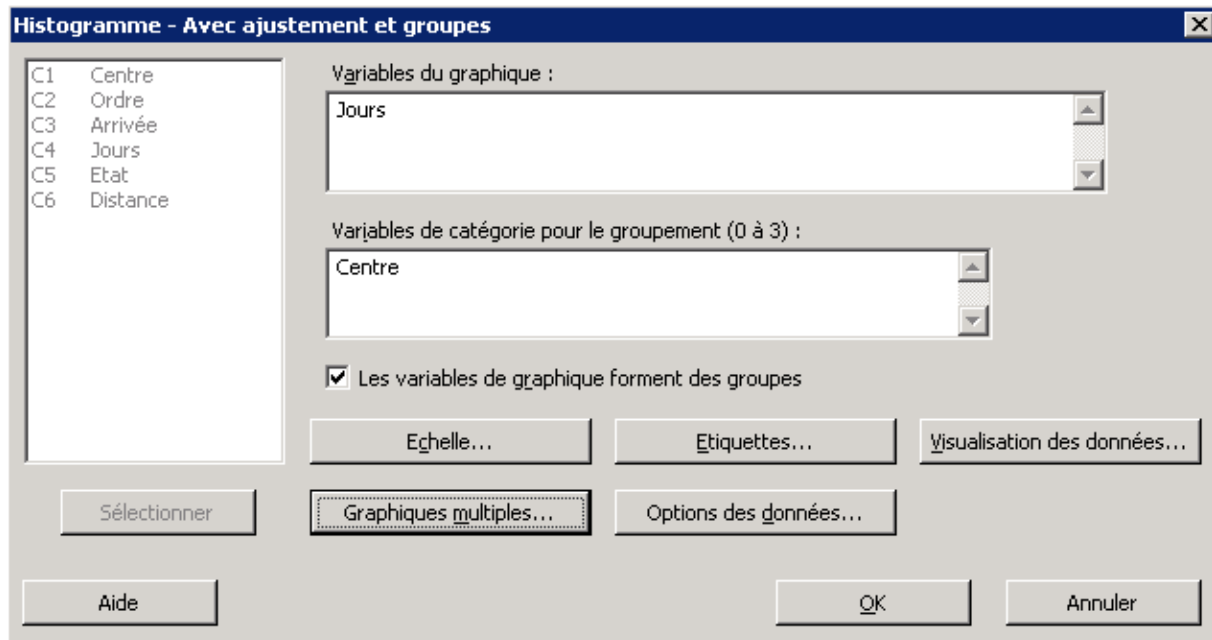
Nous allons dans le menu **Graph/Histogramme...** à nouveau:



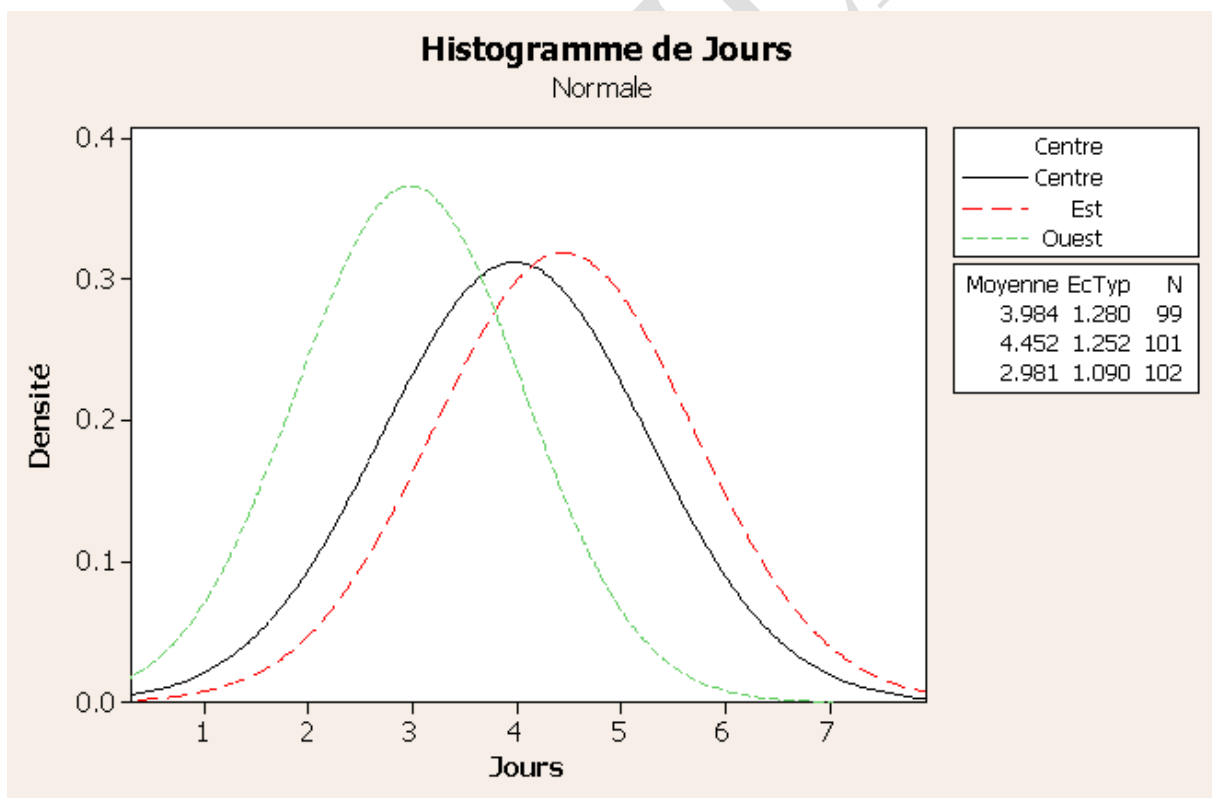
Et sélectionnez **Avec ajustement et groupes**:



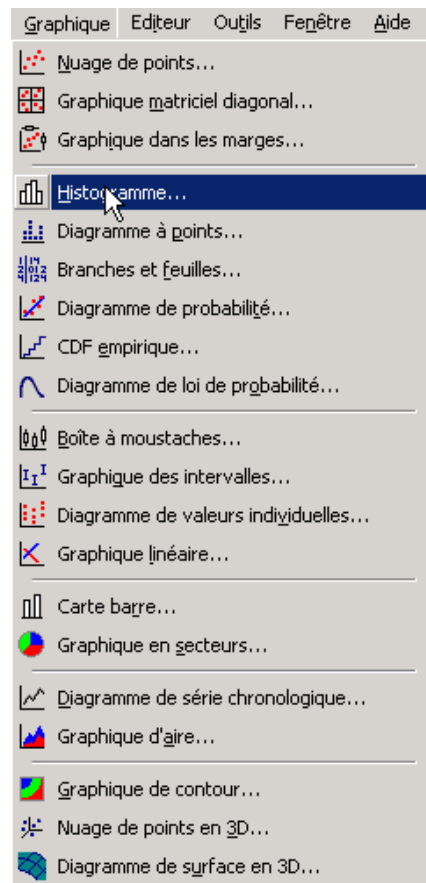
Validez par **OK** et saisissez les paramètres comme visible ci-dessous:



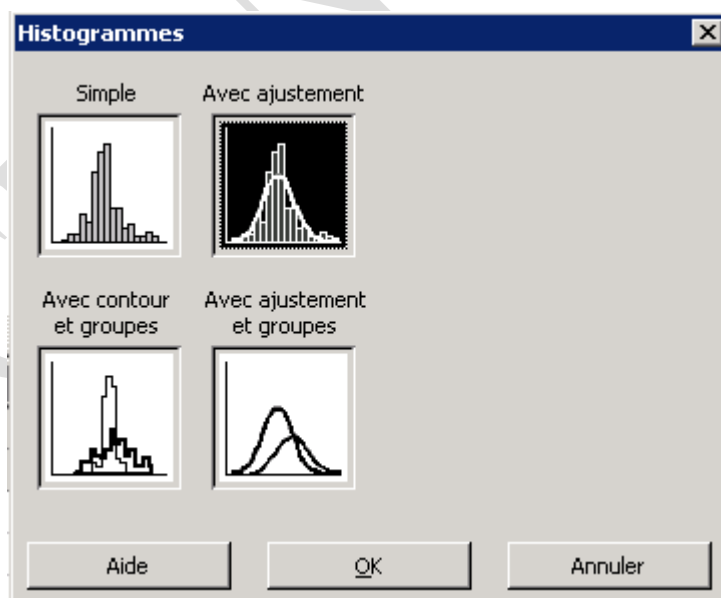
Nous avons alors (en laissant Minitab® Statistical Software choisir la loi par défaut qui est bien évidemment une loi de Gauss-Laplace...):



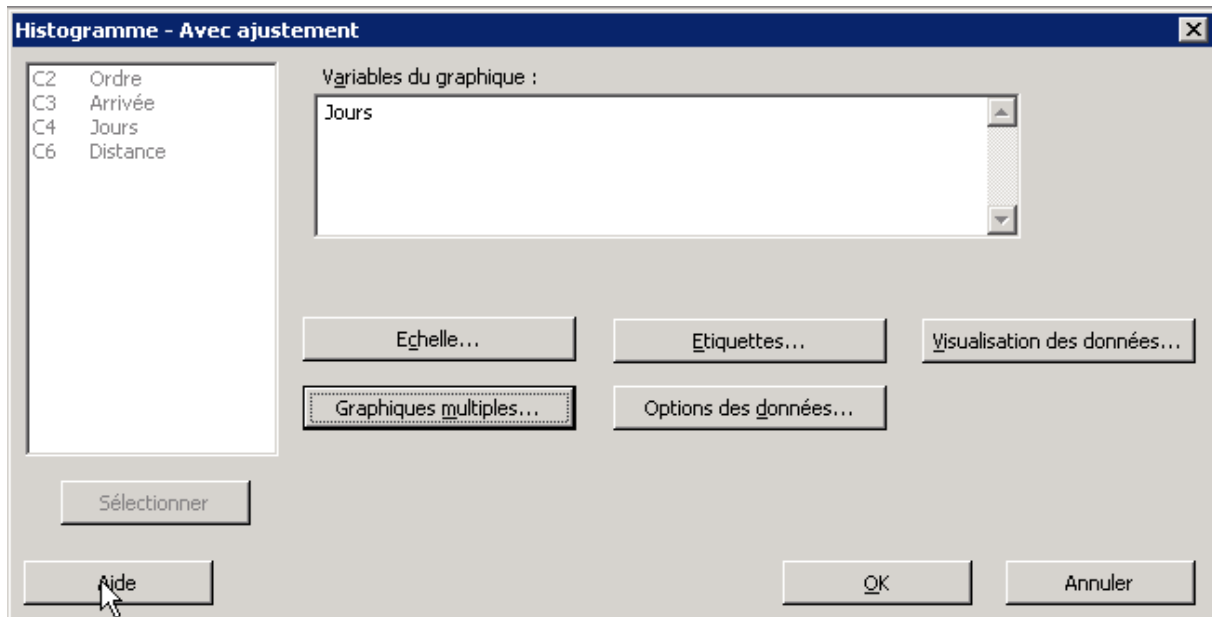
Allons un tout petit peu plus loin. Allons à nouveau dans le menu **Graph/Histogramme...**:



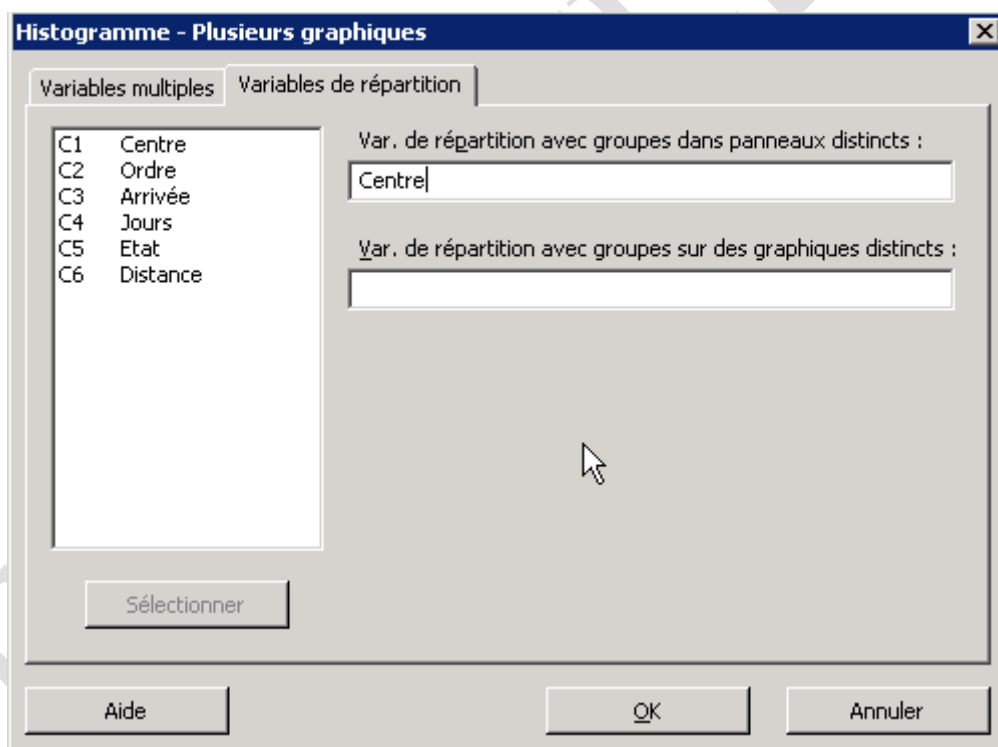
Et prenez **Avec ajustement** (pour l'instant cela ressemble à l'exemple de l'exercice précédent):



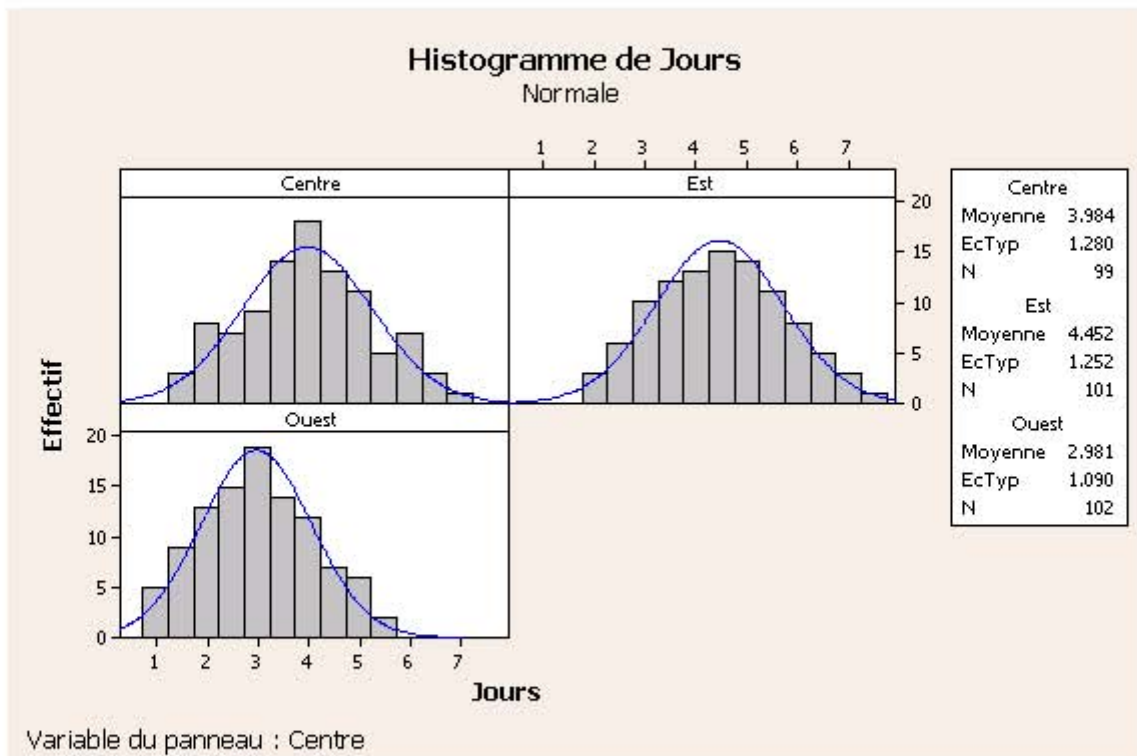
Saisissez les variables adéquates:



et maintenant, la différence c'est que nous allons cliquer sur le bouton **Graphiques multiples...**:



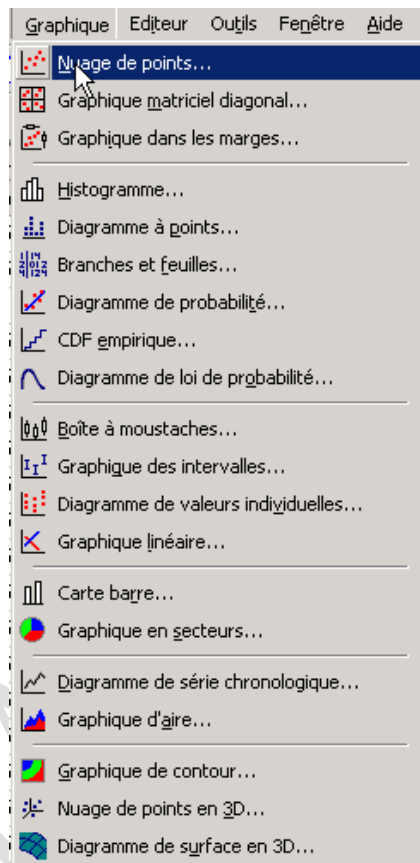
Nous avons alors (toujours en laissant Minitab® Statistical Software avec la loi d'ajustement par défaut qui est la loi de Gauss-Laplace):



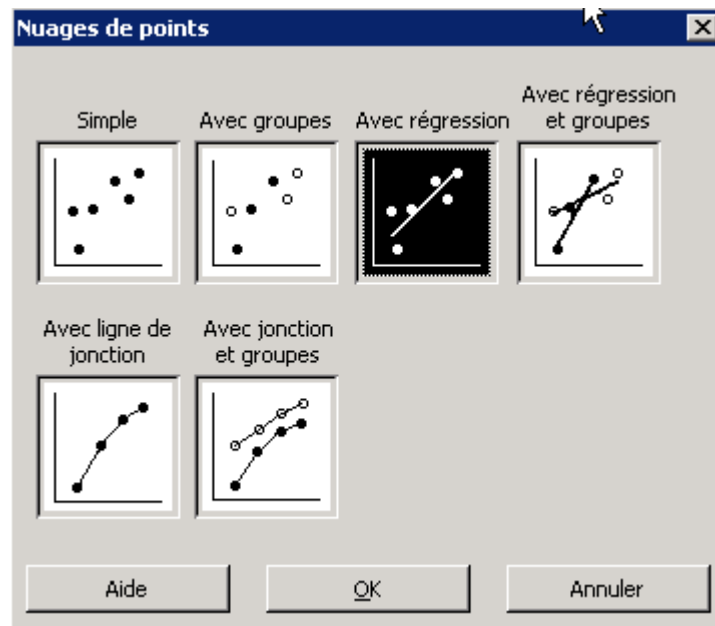
## 8.7. Exercice 15.: Graphique nuage de points avec droite de régression

Minitab® Statistical Software 15.1.1.0

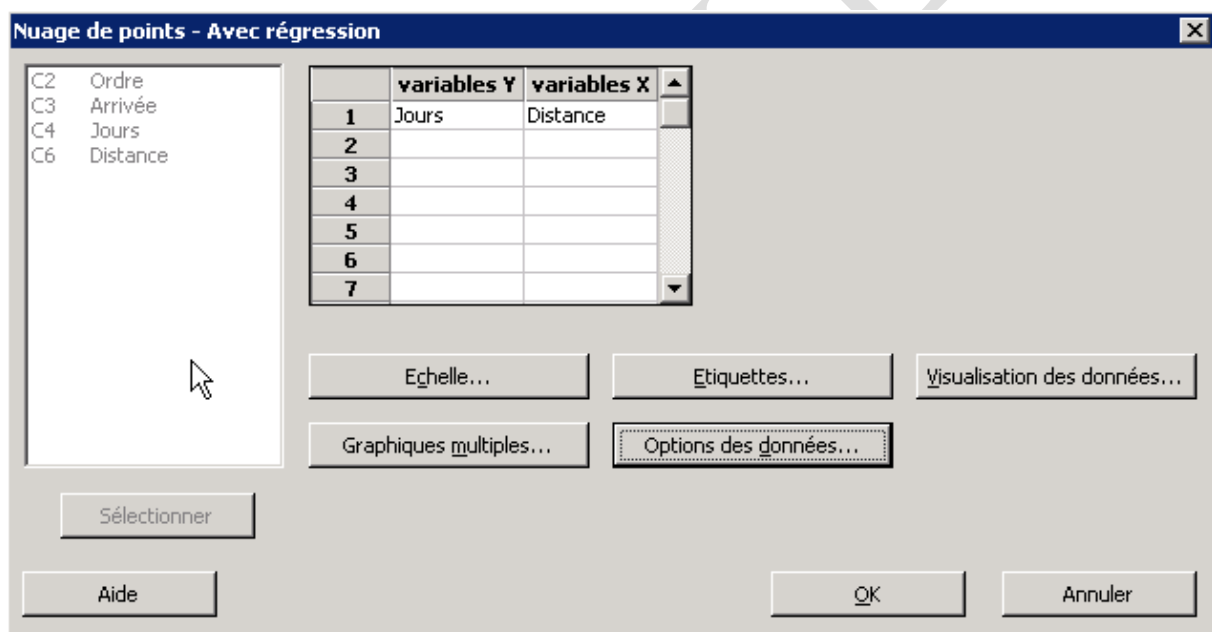
Maintenant, faisons un peu de même avec un autre type de graphique que nous avons déjà vu plus haut mais dans un contexte très simpliste. Allez dans le menu **Graph/Nuage de points...**:



Cliquez sur **Avec régression**:

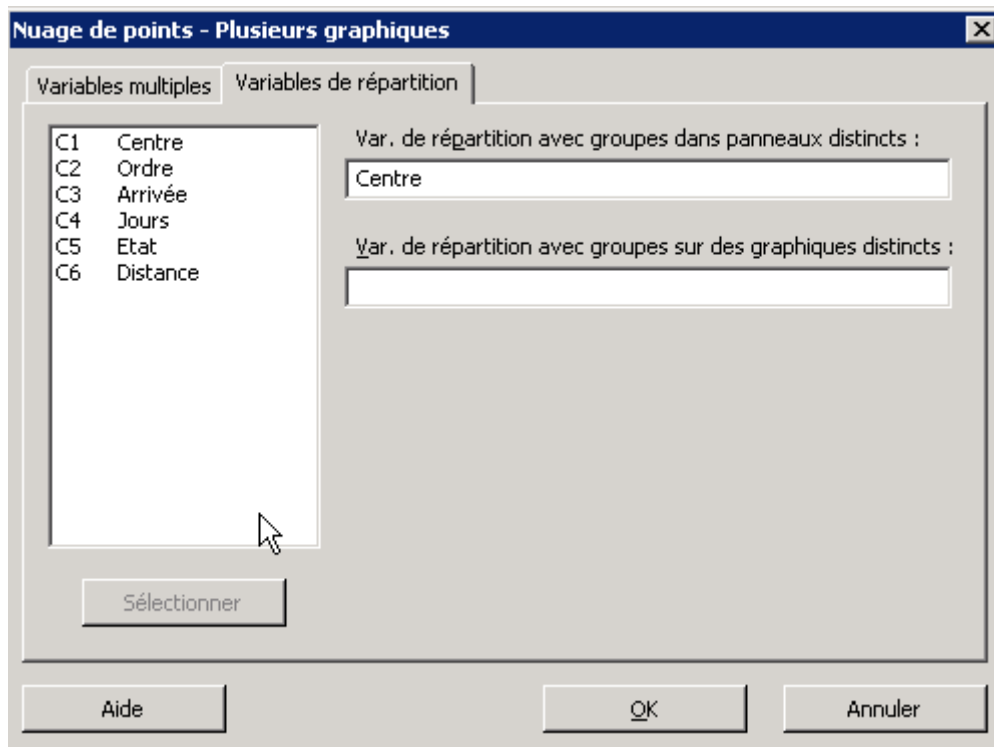


Validez par **OK** et saisissez les paramètres comme visibles ci-dessous:

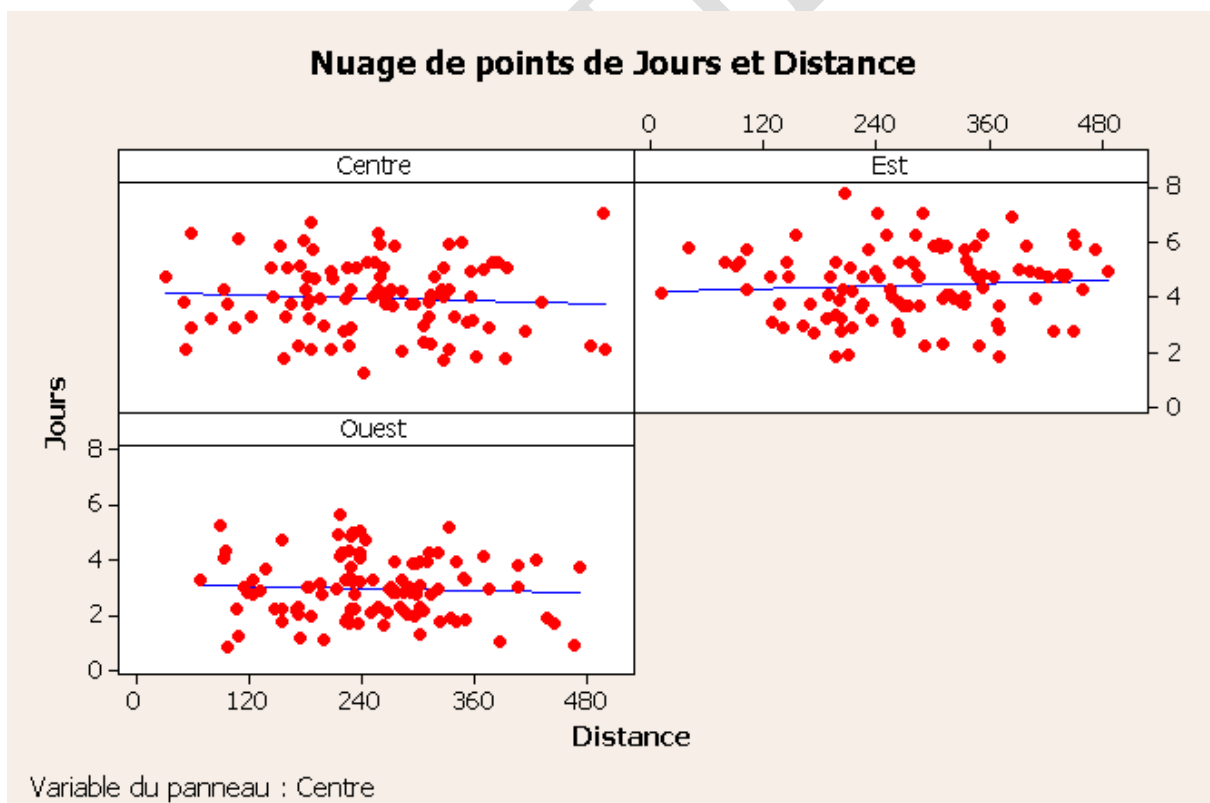


Cliquez sur le bouton **Graphiques multiples...** et mettez la variable de répartition *Centre*:





Validez par **OK**:

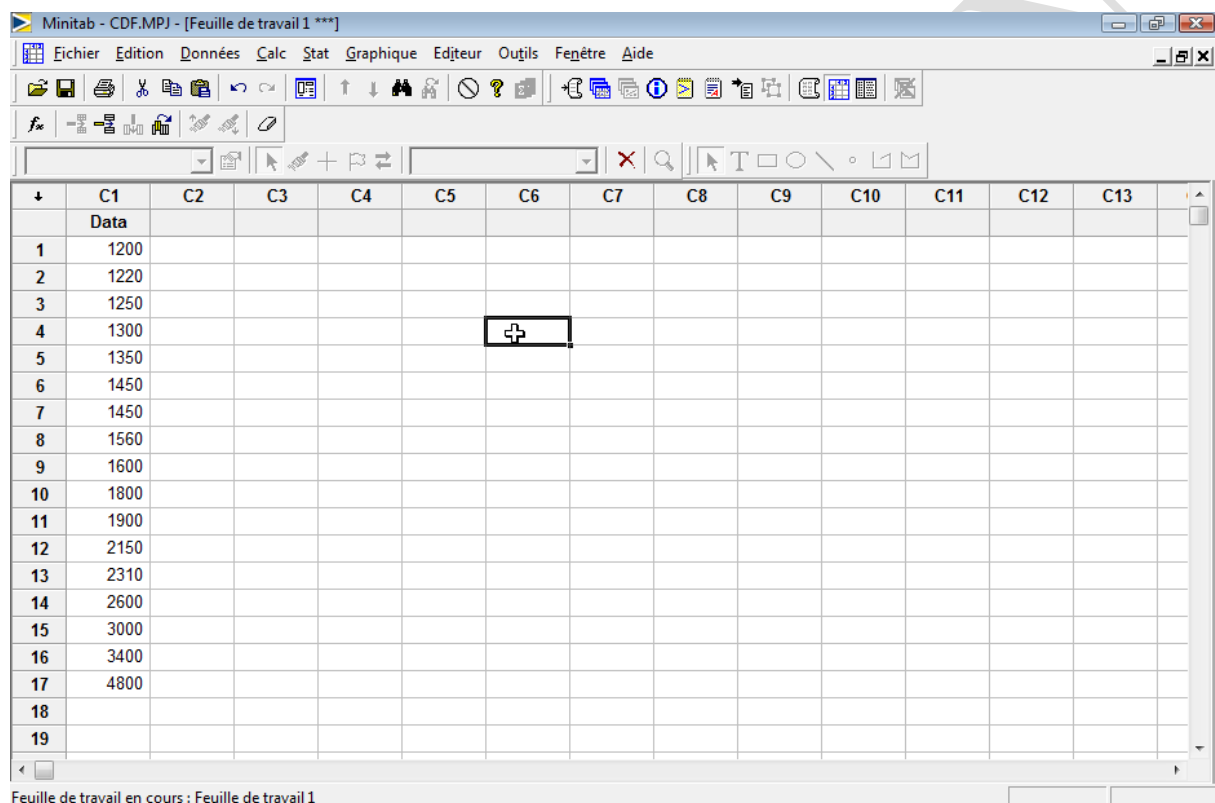


## 8.8. Exercice 16.: Générer un graphique CDF empirique

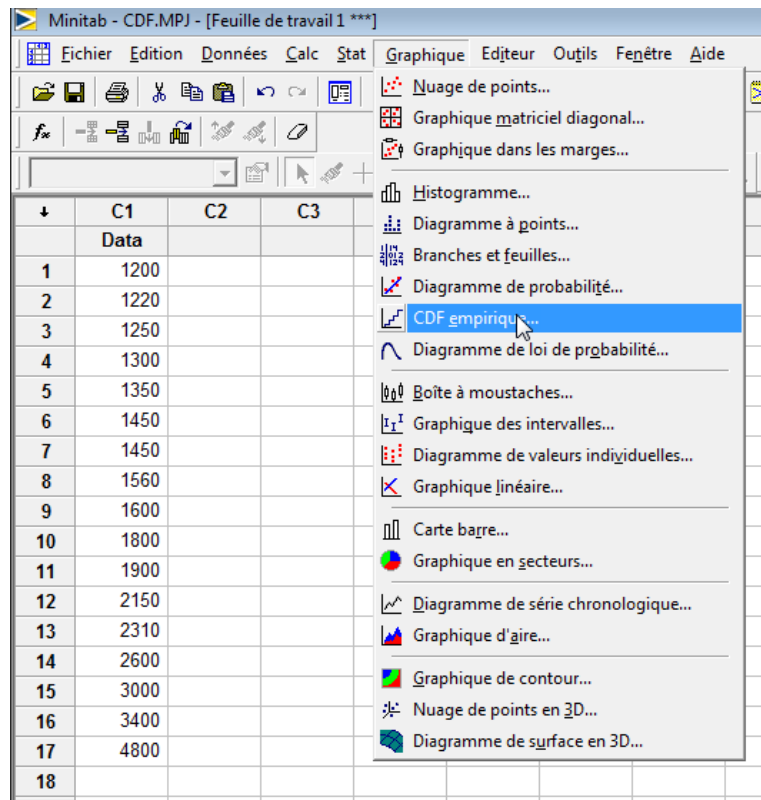
Minitab® Statistical Software 15.1.1.0

Nous avons vu dans le cours de statistique théorique comme il était malaisé de faire un graphique de la distribution empirique de probabilité avec Microsoft Excel. Eh bien nous allons voir qu'avec Minitab® Statistical Software c'est heureusement facile et rapide.

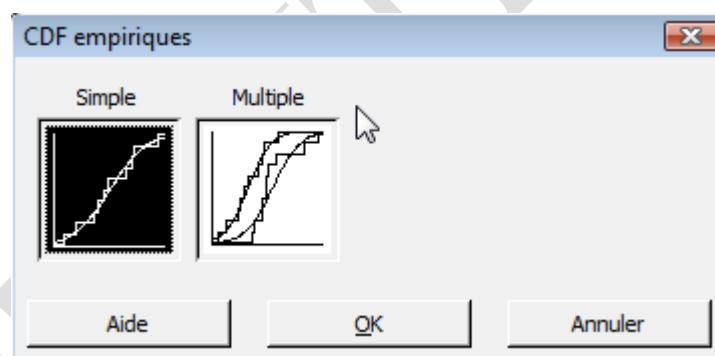
Donc nous partons des mêmes données que dans le cours de statistique théorique. Ouvrez donc le fichier *CDF.mpj*:



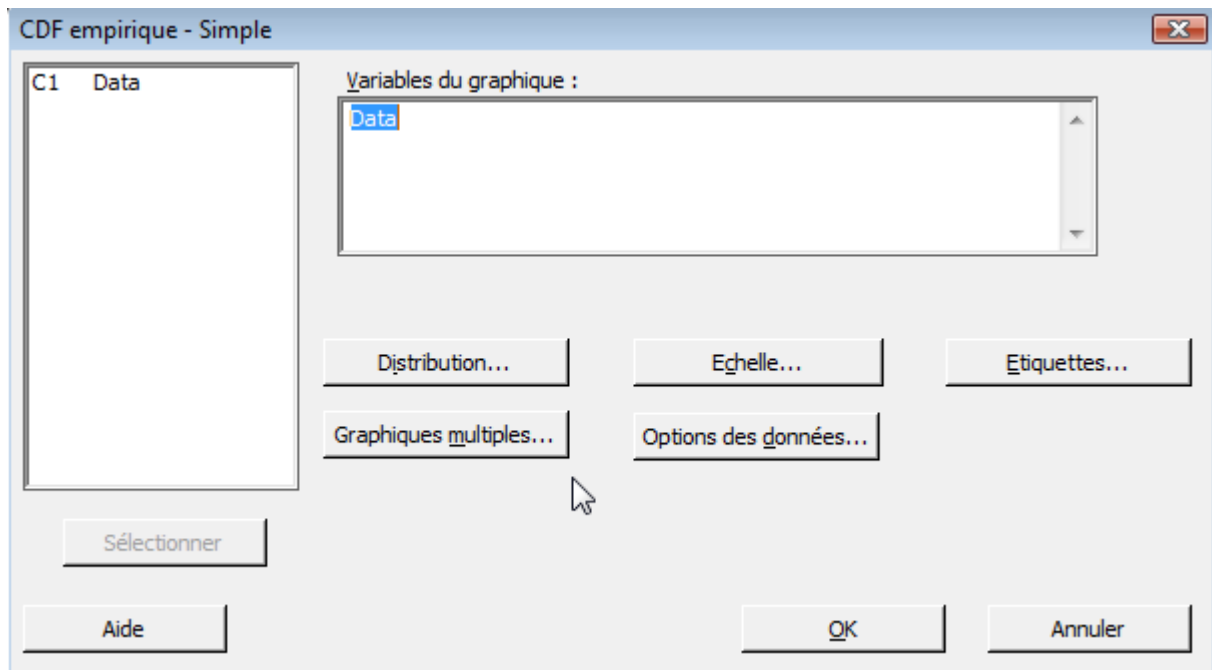
et nous allons dans le menu **Graphique/CDF empirique...**:



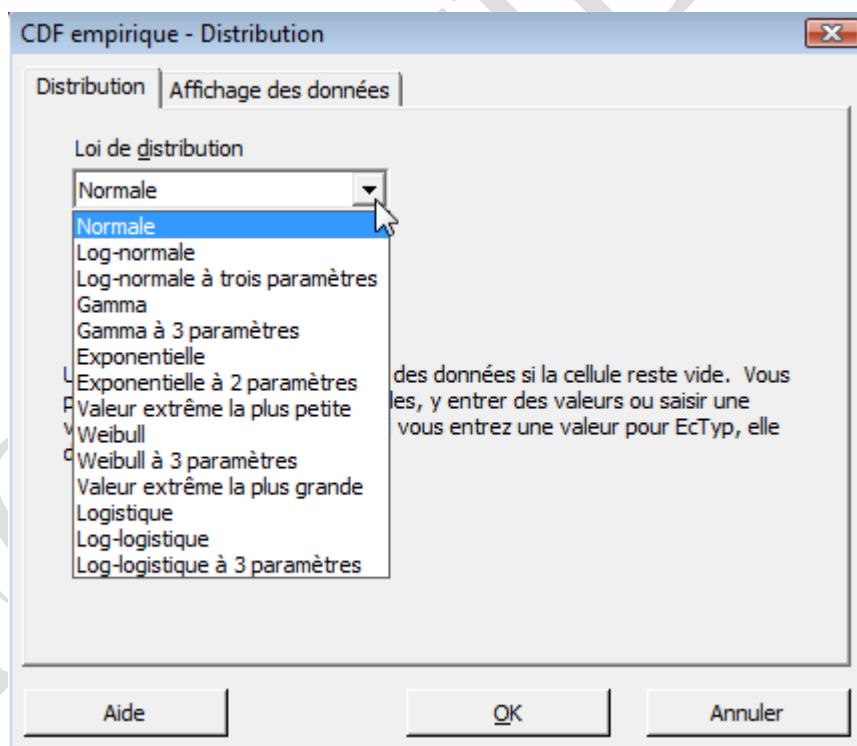
Nous avons alors:



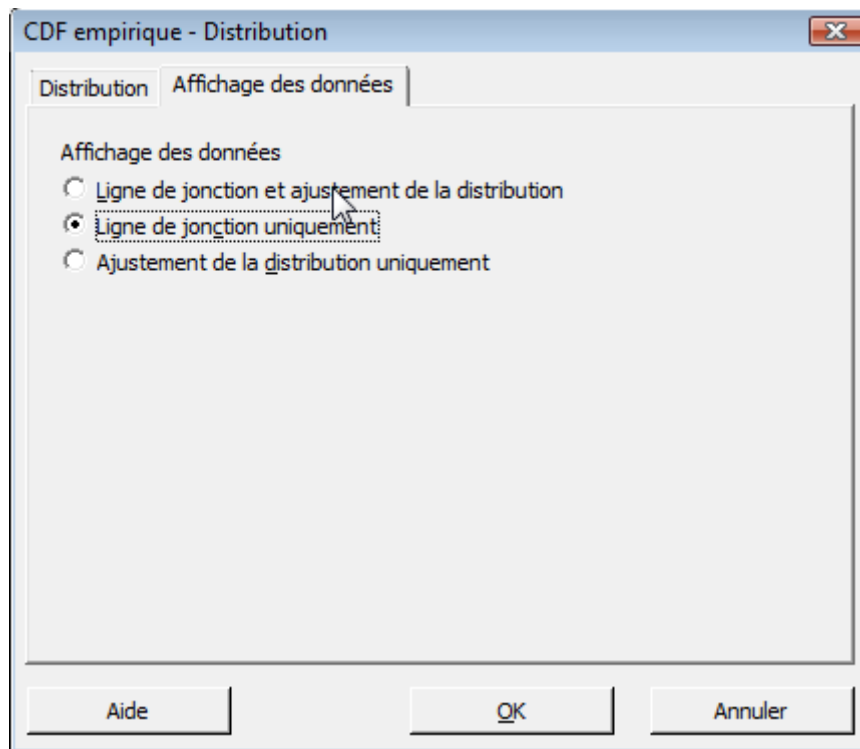
nous prenons le **Simple** et validons par **OK**:



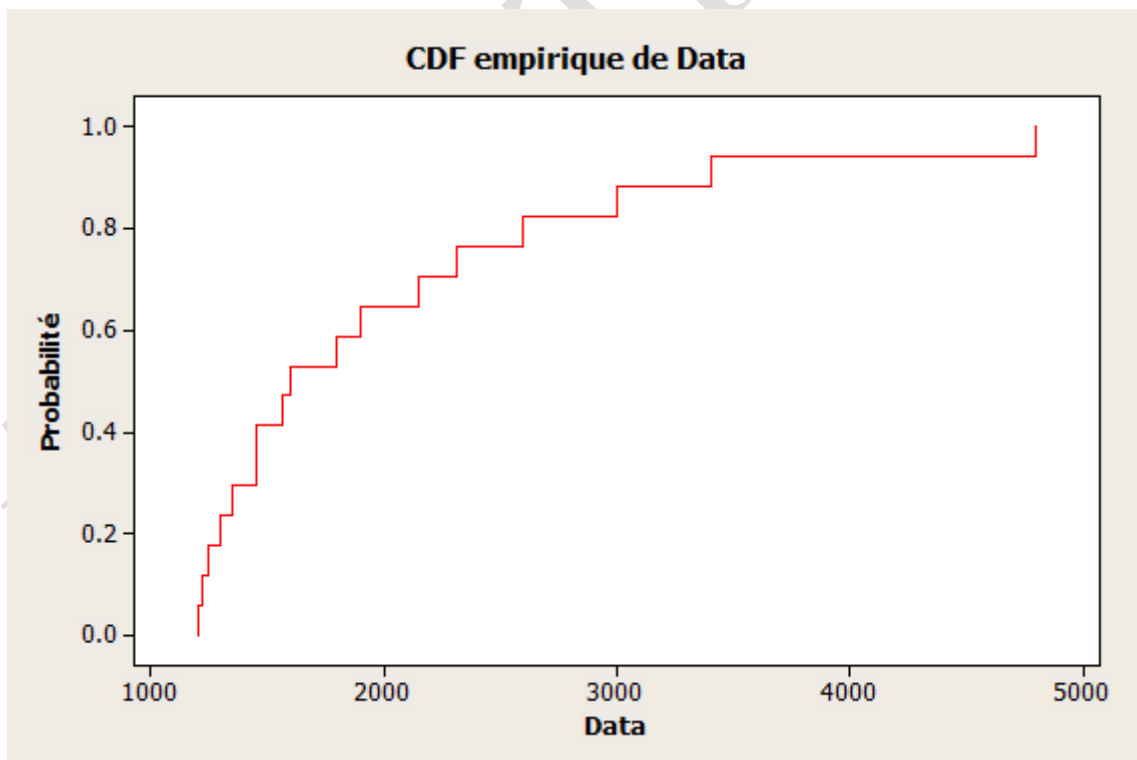
Nous pouvons comparer la CDF empirique à une distribution de notre choix en cliquant sur le bouton **Distribution...**:



Mais si nous ne savons pas à quoi la comparer (faute d'avoir fait un test d'ajustement au préalable), nous pouvons aller dans l'onglet **Affichage des données**:



et donc cliquer sur **Ligne de jonction uniquement**. Nous validons par **OK** deux fois et nous obtenons:

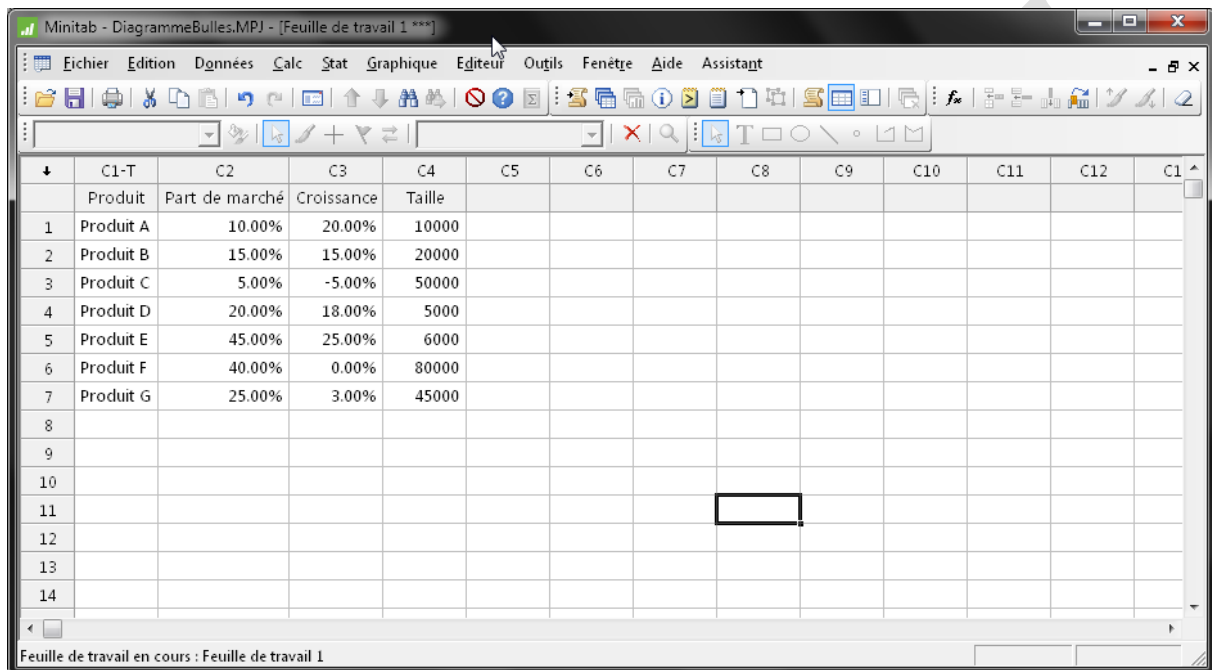


## 8.9. Exercice 17.: Générer un graphique à bulles

Minitab® Statistical Software 17.1.0

Rien de spécial ici à part qu'il s'agit d'une nouveauté de la version 17. Nous allons refaire ici le diagramme à bulle fait dans le cours Microsoft Excel et R.

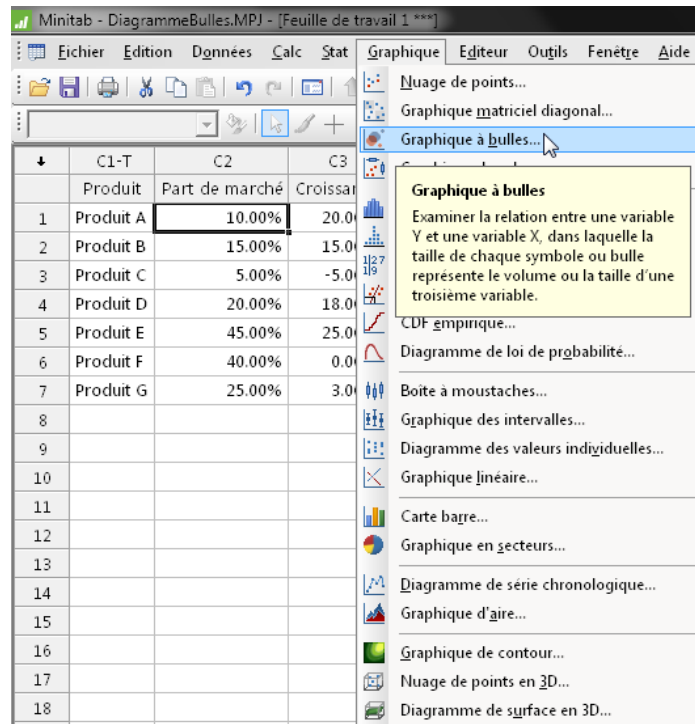
Nous partons donc des données suivantes:



The screenshot shows the Minitab interface with a data table. The table has the following data:

	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C1
	Produit	Part de marché	Croissance	Taille									
1	Produit A	10.00%	20.00%	10000									
2	Produit B	15.00%	15.00%	20000									
3	Produit C	5.00%	-5.00%	50000									
4	Produit D	20.00%	18.00%	5000									
5	Produit E	45.00%	25.00%	6000									
6	Produit F	40.00%	0.00%	80000									
7	Produit G	25.00%	3.00%	45000									
8													
9													
10													
11													
12													
13													
14													

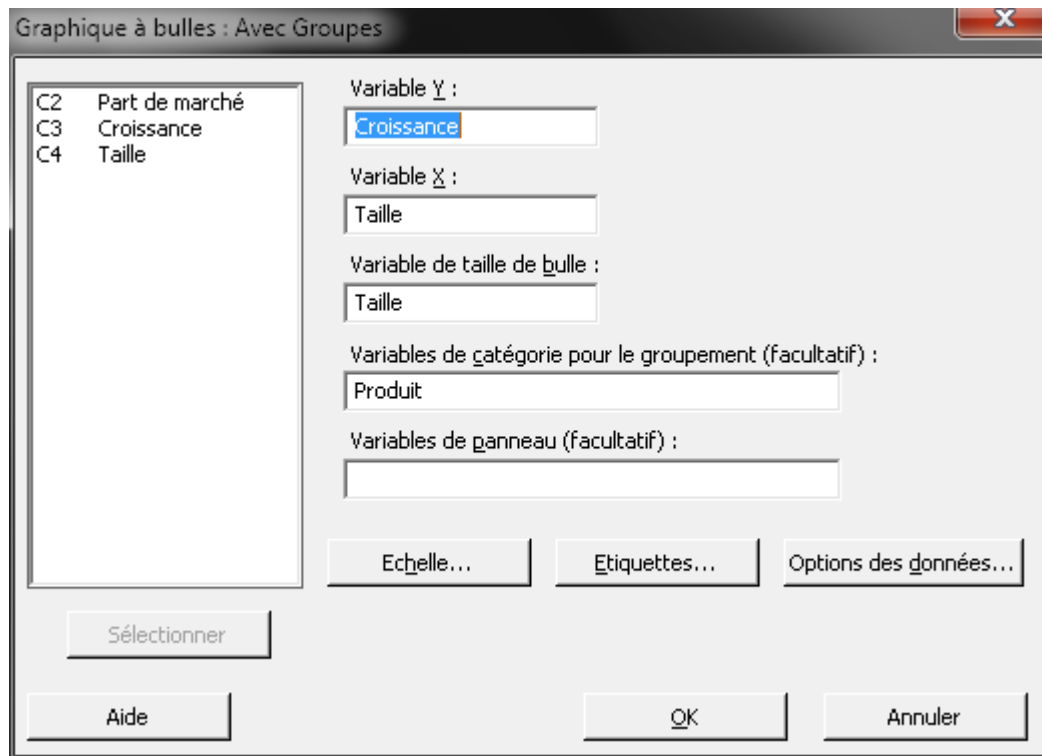
Et nous allons dans le menu **Stat/Graphique à bulles...**:



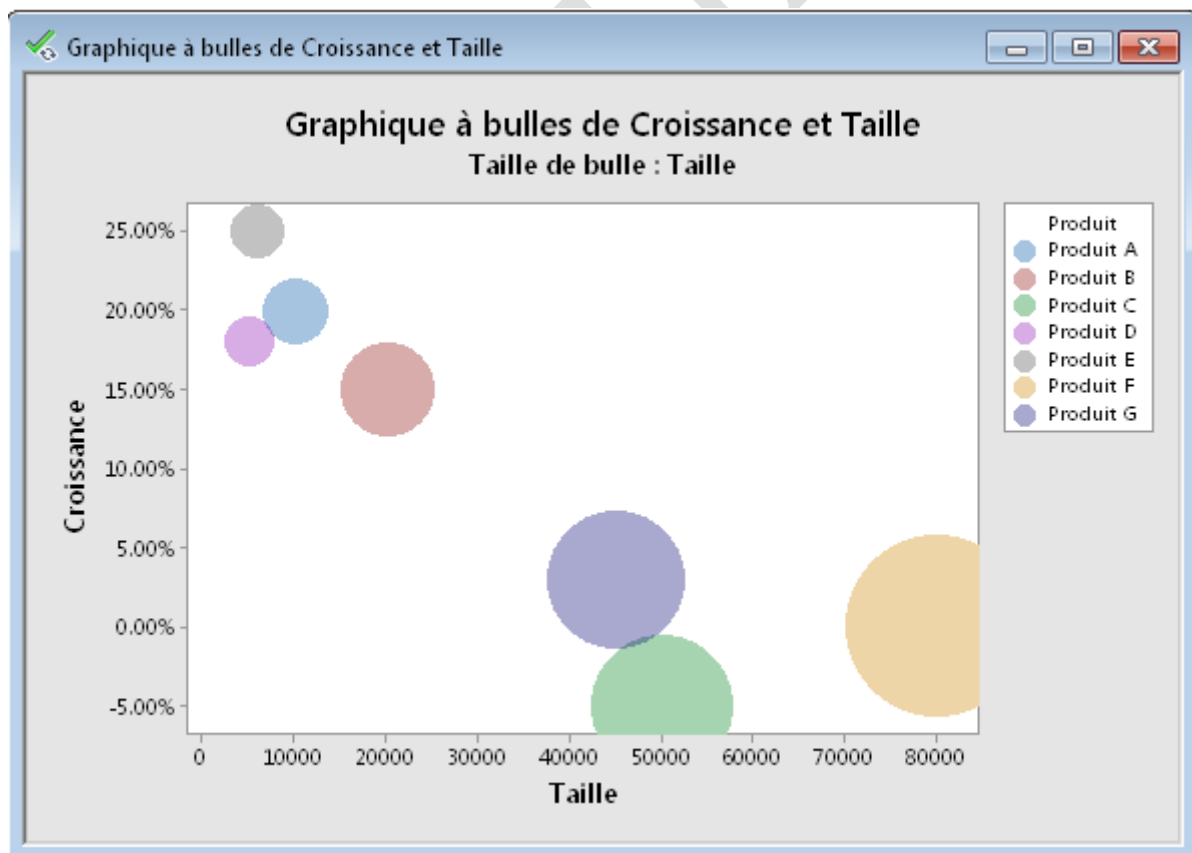
Nous prenons l'option **Avec groupes**:



Nous prenons alors:



et nous validons pour obtenir:

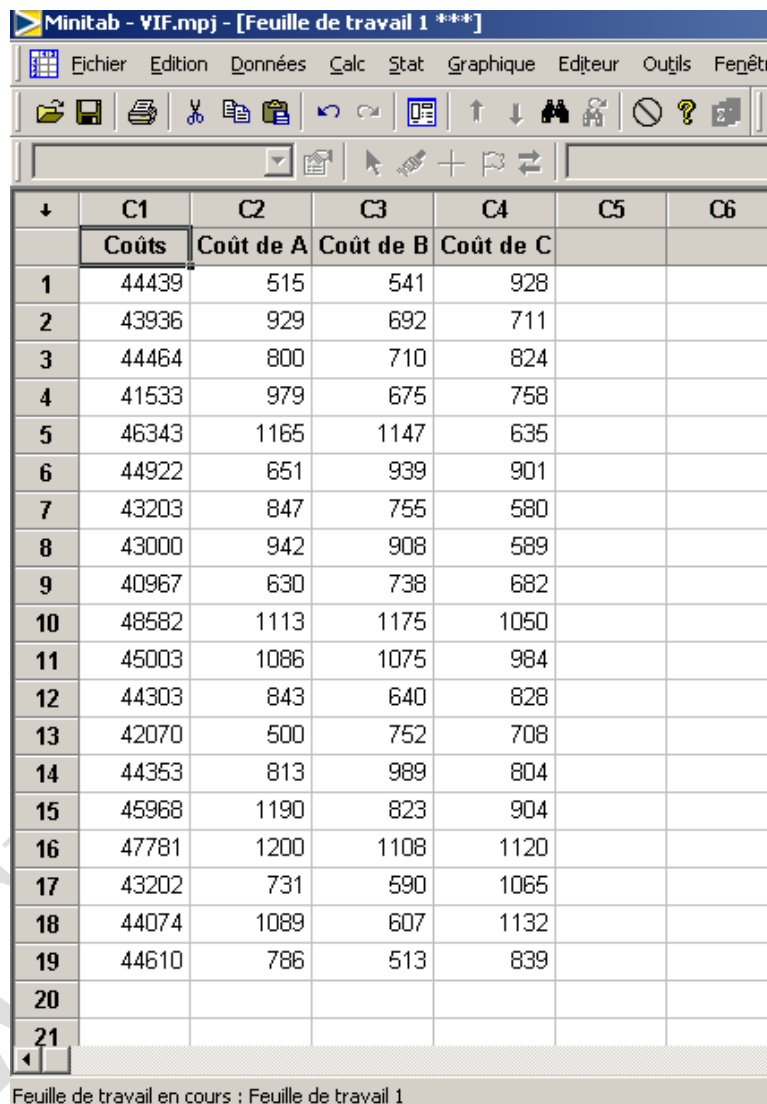




## 8.10. Exercice 18.: Graphique matriciel diagonal

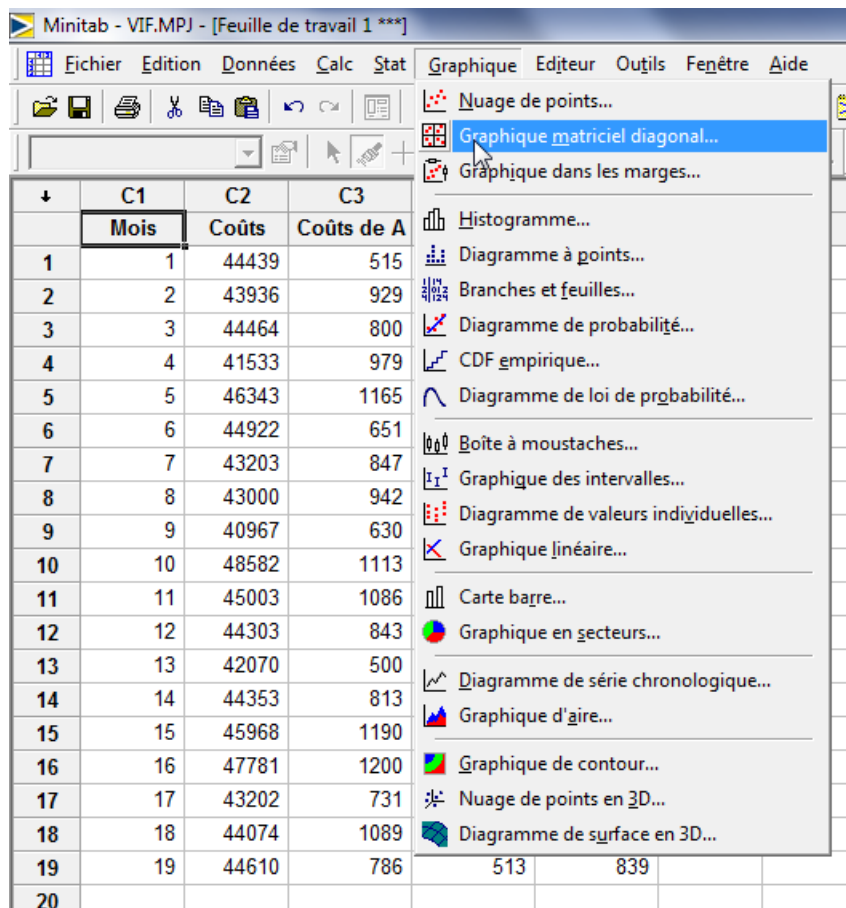
Minitab® Statistical Software 15.1.1

Le graphique matriciel diagonal peut être utilisé pour juger qualitativement (nous verrons comment juger cela quantitativement plus tard!) de la multi-colinéarité de variables de régression. Voyons comment il marche en partant du jeu de données suivant:

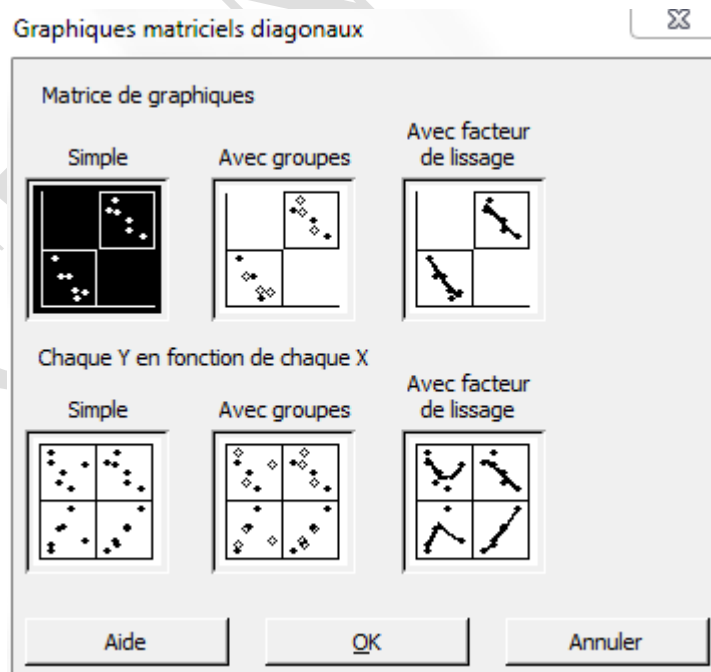


	C1	C2	C3	C4	C5	C6
	Coûts	Coût de A	Coût de B	Coût de C		
1	44439	515	541	928		
2	43936	929	692	711		
3	44464	800	710	824		
4	41533	979	675	758		
5	46343	1165	1147	635		
6	44922	651	939	901		
7	43203	847	755	580		
8	43000	942	908	589		
9	40967	630	738	682		
10	48582	1113	1175	1050		
11	45003	1086	1075	984		
12	44303	843	640	828		
13	42070	500	752	708		
14	44353	813	989	804		
15	45968	1190	823	904		
16	47781	1200	1108	1120		
17	43202	731	590	1065		
18	44074	1089	607	1132		
19	44610	786	513	839		
20						
21						

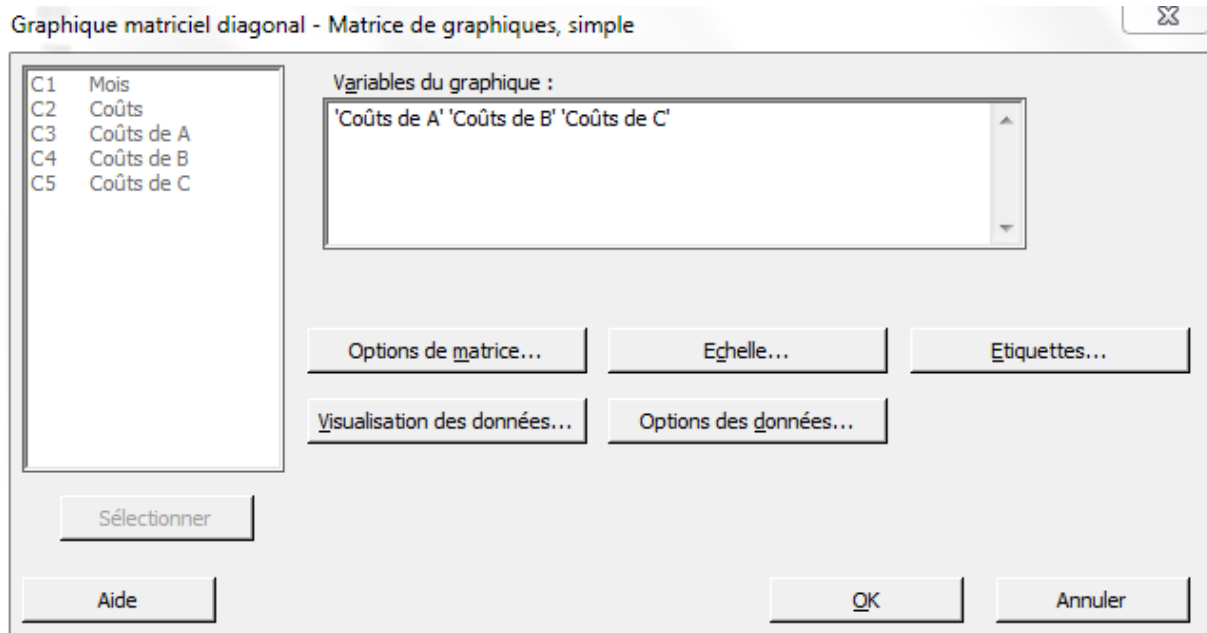
Nous allons ensuite dans le menu **Graphique/Graphique matriciel diagonal...**:



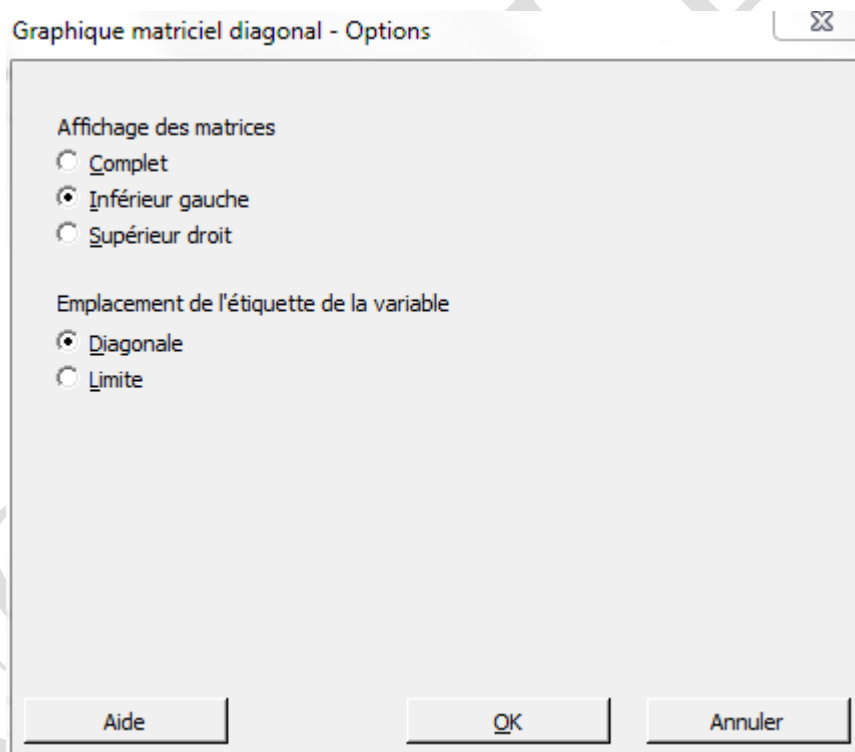
Ensuite nous prenons le **Simple**:



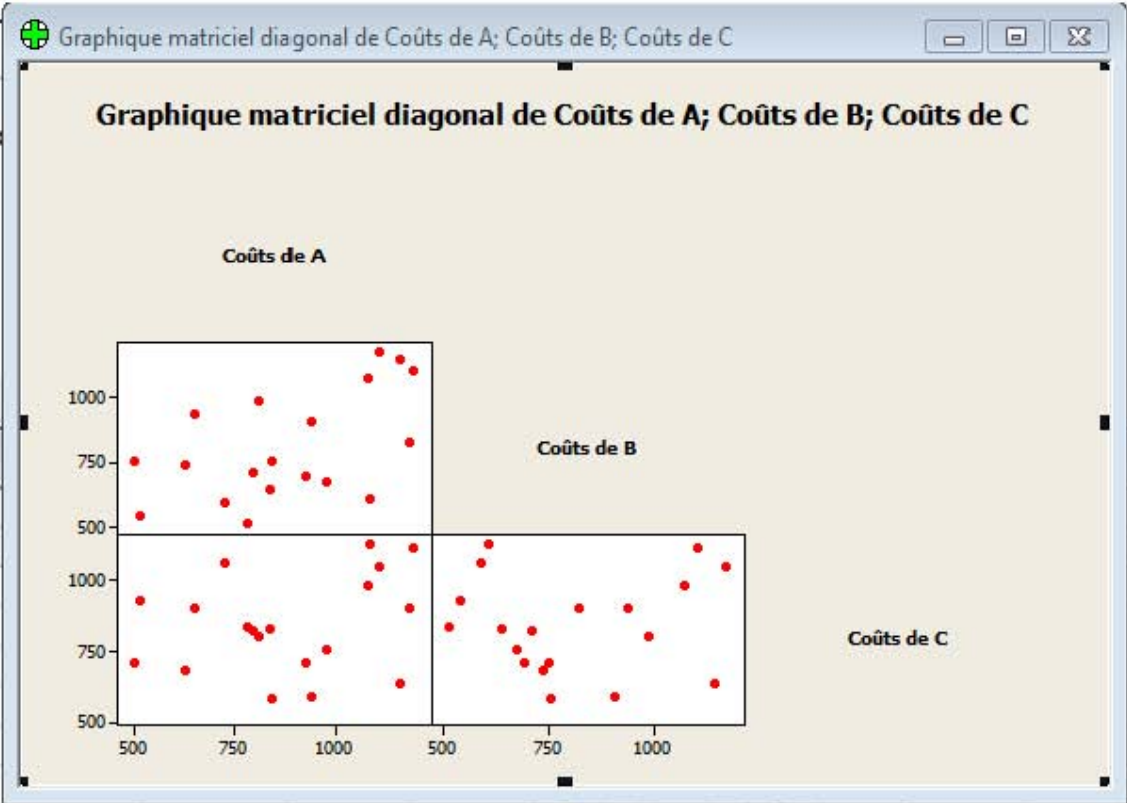
Dans la boîte de dialogue qui apparaît, nous mettons:



Ensuite, nous cliquons sur le bouton **Options de matrice...** pour prendre **Inférieur gauche**:



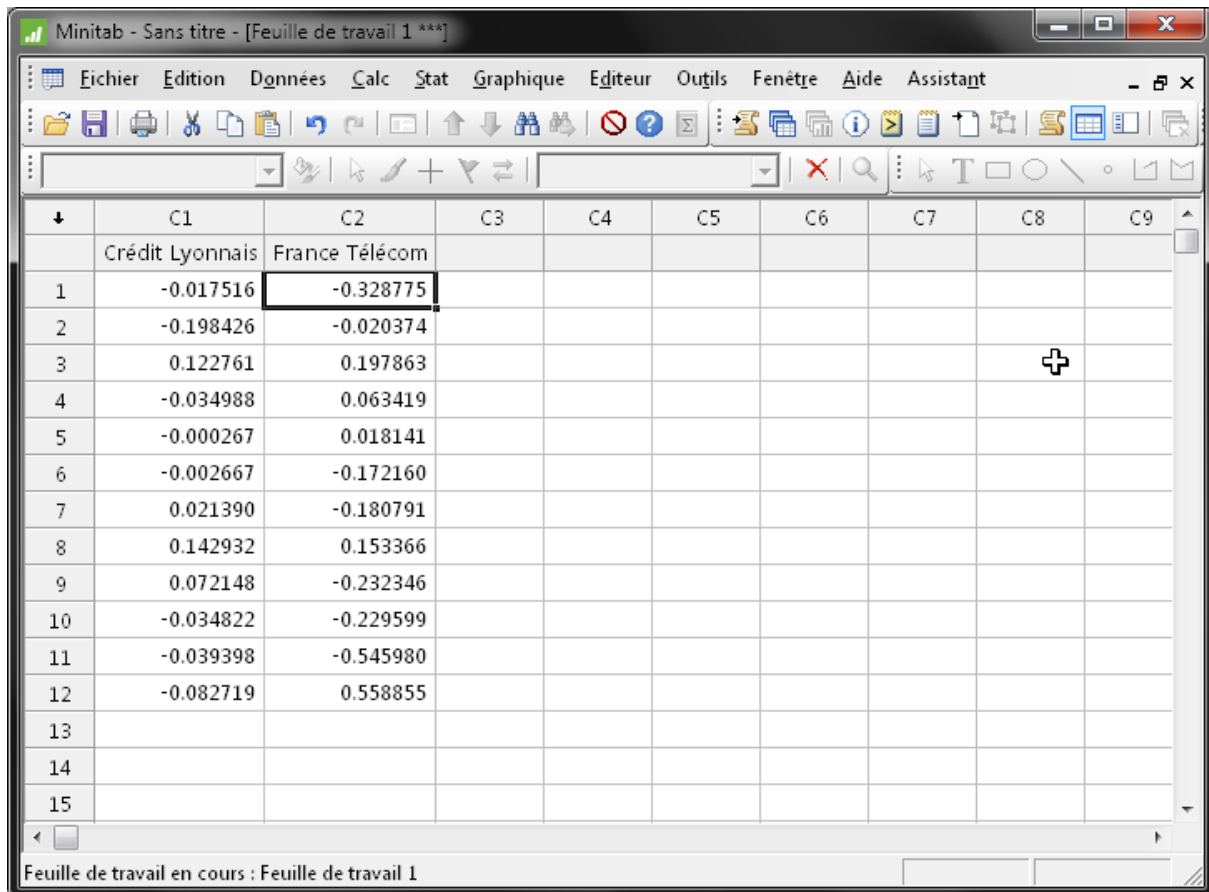
Ce qui donne:



## 8.11. Exercice 19.: Graphique dans les marges

Minitab® Statistical Software 17.1.0

Une fonctionnalité qui existe déjà depuis bien longtemps dans Minitab mais qui n'existe pas par défaut dans Microsoft Excel. Si nous le faisons ici c'est juste pour comparer avec le logiciel R en utilisant les mêmes données:



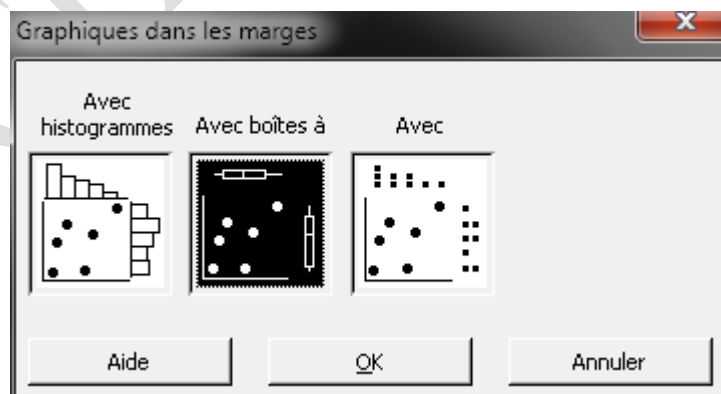
The screenshot shows the Minitab software interface with a worksheet titled 'Minitab - Sans titre - [Feuille de travail 1 \*\*\*]'. The menu bar includes 'Fichier', 'Edition', 'Données', 'Calc', 'Stat', 'Graphique', 'Editeur', 'Outils', 'Fenêtre', 'Aide', and 'Assistant'. The toolbar contains various icons for file operations, editing, and data manipulation. The worksheet grid has columns labeled C1 through C9 and rows numbered 1 through 15. The data is as follows:

	C1	C2	C3	C4	C5	C6	C7	C8	C9
	Crédit Lyonnais	France Télécom							
1	-0.017516	-0.328775							
2	-0.198426	-0.020374							
3	0.122761	0.197863							
4	-0.034988	0.063419							
5	-0.000267	0.018141							
6	-0.002667	-0.172160							
7	0.021390	-0.180791							
8	0.142932	0.153366							
9	0.072148	-0.232346							
10	-0.034822	-0.229599							
11	-0.039398	-0.545980							
12	-0.082719	0.558855							
13									
14									
15									

Nous allons dans le menu **Graphique/Graphique dans les marges...**:

	C1	C2
	Crédit Lyonnais	France Télécom
1	-0.017516	-0.328775
2	-0.198426	-0.020374
3	0.122761	0.197863
4	-0.034988	0.063419
5	-0.000267	0.018141
6	-0.002667	-0.172160
7	0.021390	-0.180791
8	0.142932	0.153366
9	0.072148	-0.232346
10	-0.034822	-0.229599
11	-0.039398	-0.545980
12	-0.082719	0.558855
13		
14		
15		
16		
17		
18		
19		

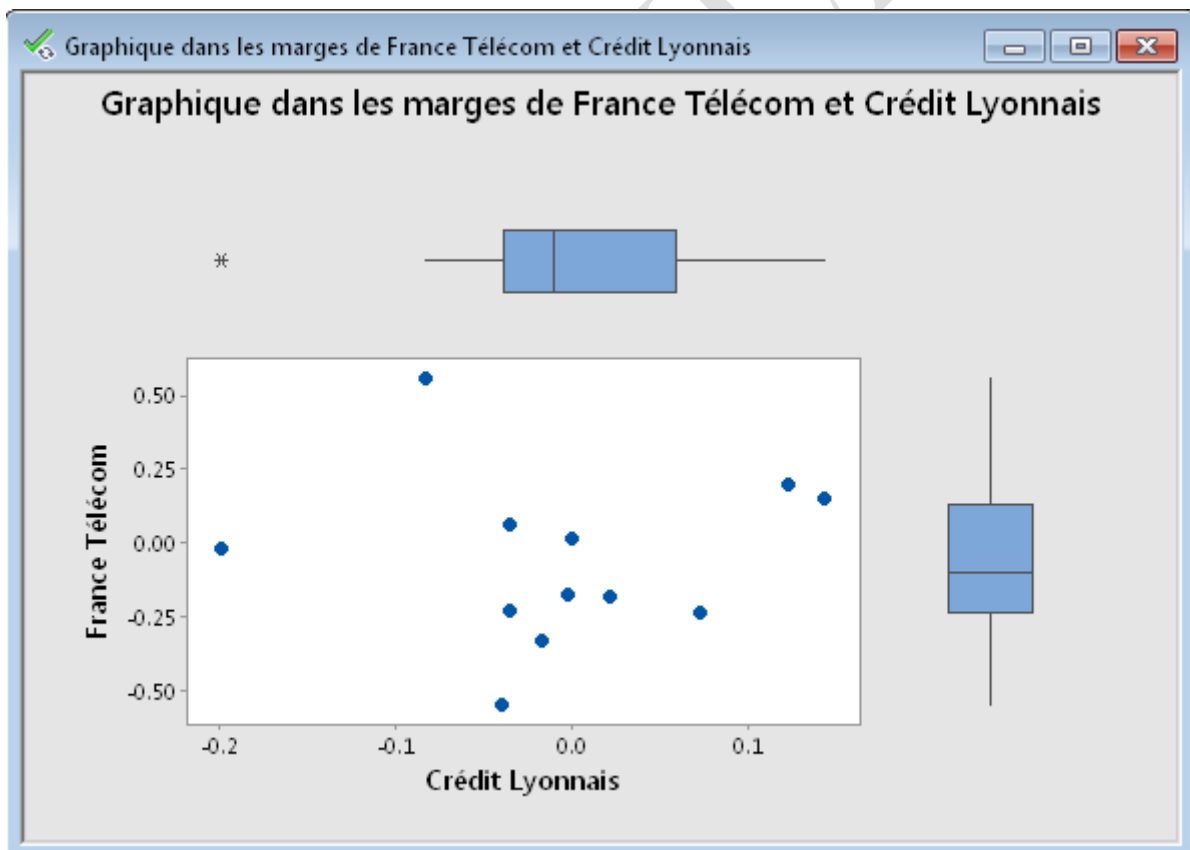
et nous prenons l'option du milieu:



Ensuite, nous prenons:



Ce qui nous donne:

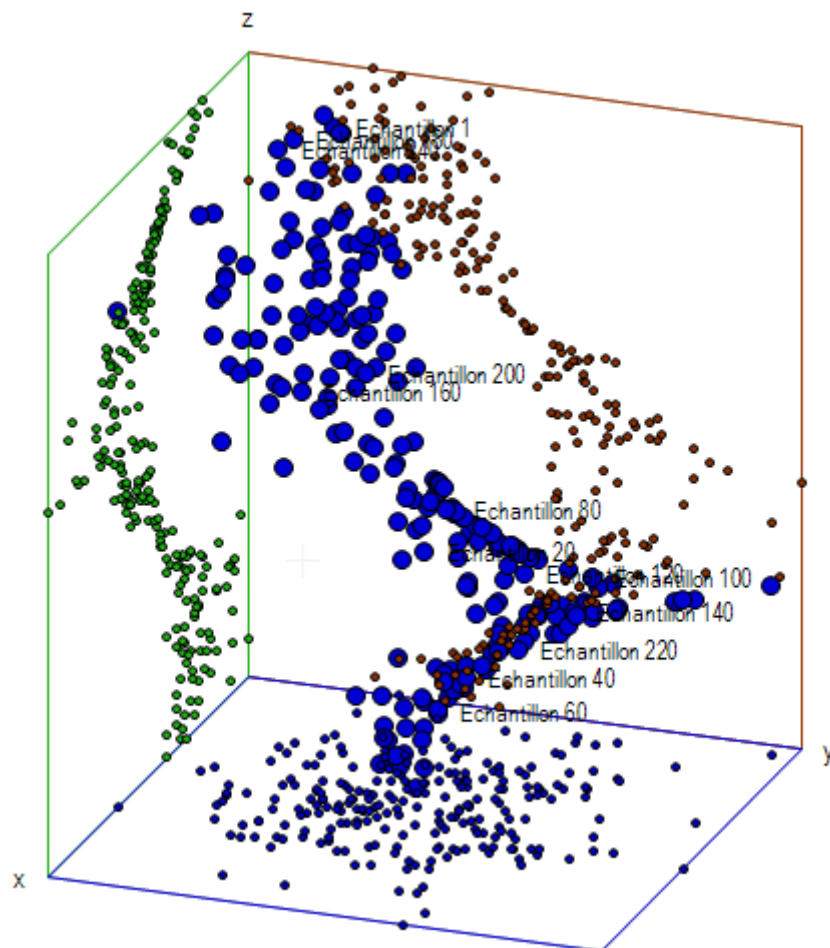


Bref comme nous le voyons le résultat n'est pas spécialement esthétiquement mieux que dans R.

## 8.12. Exercice 20.: Générer un graphique à points 3D

Minitab® Statistical Software 15.1.1.0

Nous allons reprendre les données vues dans un cours Microsoft Excel (où nous avons vu comment faire un graphique 3D à partir d'un code VBA). Nous y avons obtenu:



Nous souhaitons voir ce que Minitab® Statistical Software nous propose. Nous ouvrons pour cela le fichier *3DScatterPlot.mpj*:



The screenshot shows the Minitab interface with a 3D scatter plot window titled "3DScatterPlot.MPJ". The main window displays a session log with the timestamp "22.05.2011 14:58:03". Below the session log is a data table with columns C1 through C14. The data is as follows:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	X	Y	Z											
1	2.93564	2.88500	3.00900											
2	4.81348	4.97150	-0.08465											
3	4.79703	4.46382	0.25901											
4	3.99052	3.53974	2.33107											
5	2.05557	3.06640	-2.88904											
6	3.39610	2.68965	2.79445											
7	2.97684	4.88072	-1.69820											
8	4.18250	3.64112	1.82984											
9	4.43010	4.04133	0.68906											
10	4.36754	4.47288	-0.12948											

The status bar at the bottom indicates "Feuille de travail en cours : Feuille de travail 2".

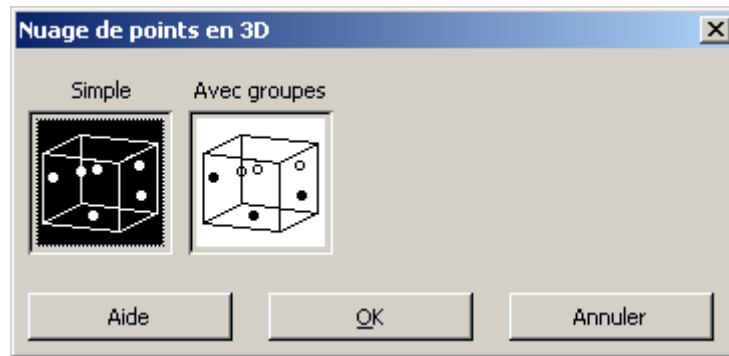
et nous allons dans le menu **Graphique/Nuage de points en 3D...**:

The screenshot shows the Minitab interface with the "Graphique" menu open. The menu options are:

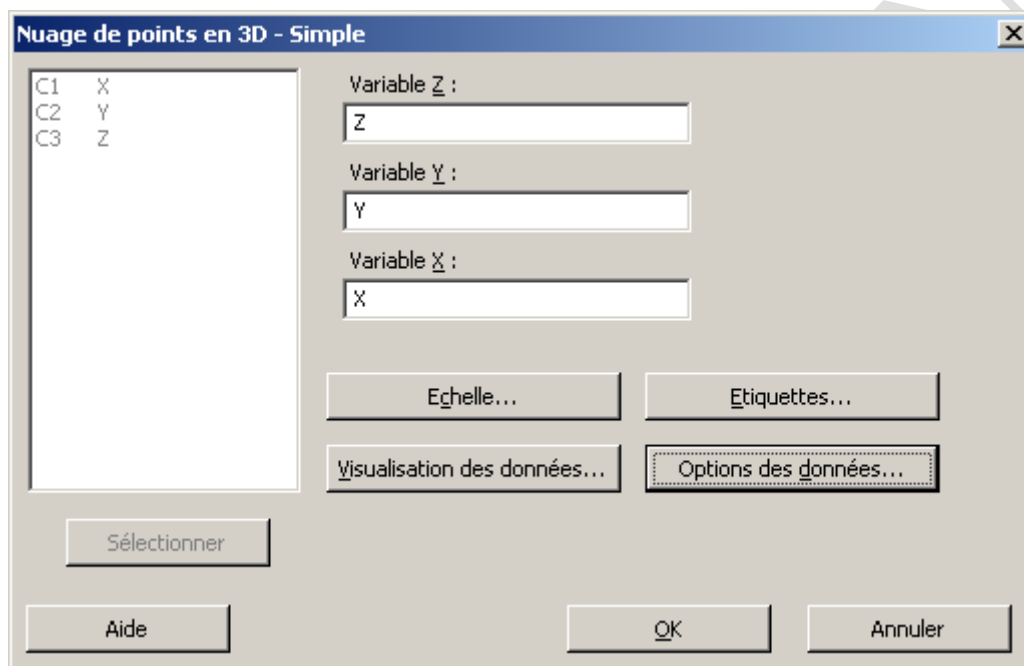
- Nuage de points...
- Graphique matriciel diagonal...
- Graphique dans les marges...
- Histogramme...
- Diagramme à points...
- Branches et feuilles...
- Diagramme de probabilité...
- CDF empirique...
- Diagramme de loi de probabilité...
- Boîte à moustaches...
- Graphique des intervalles...
- Diagramme de valeurs individuelles...
- Graphique linéaire...
- Carte barge...
- Graphique en secteurs...
- Diagramme de série chronologique...
- Graphique d'aire...
- Graphique de contour...
- Nuage de points en 3D...**
- Diagramme de surface en 3D...

The status bar at the bottom indicates "Tracer des nuages de points en 3D".

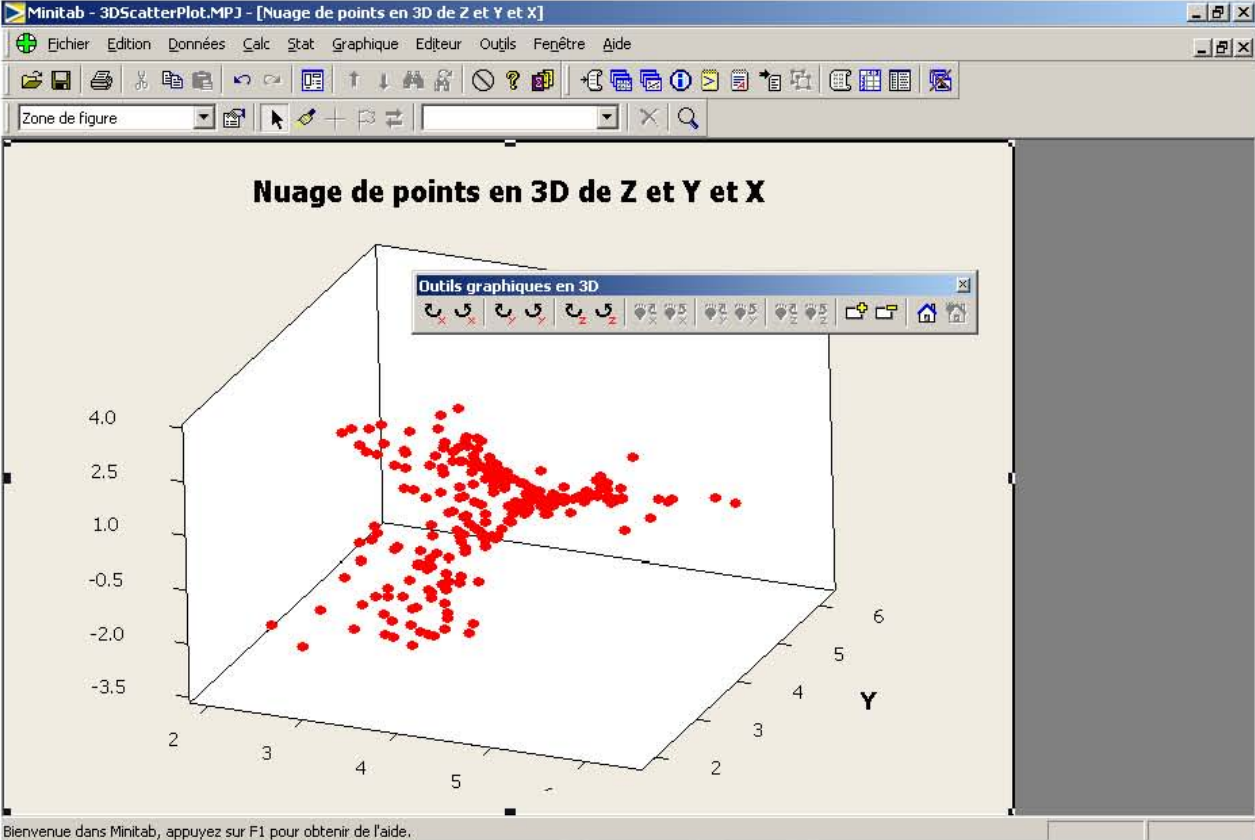
Nous avons alors:



Nous prenons la version **Simple**:



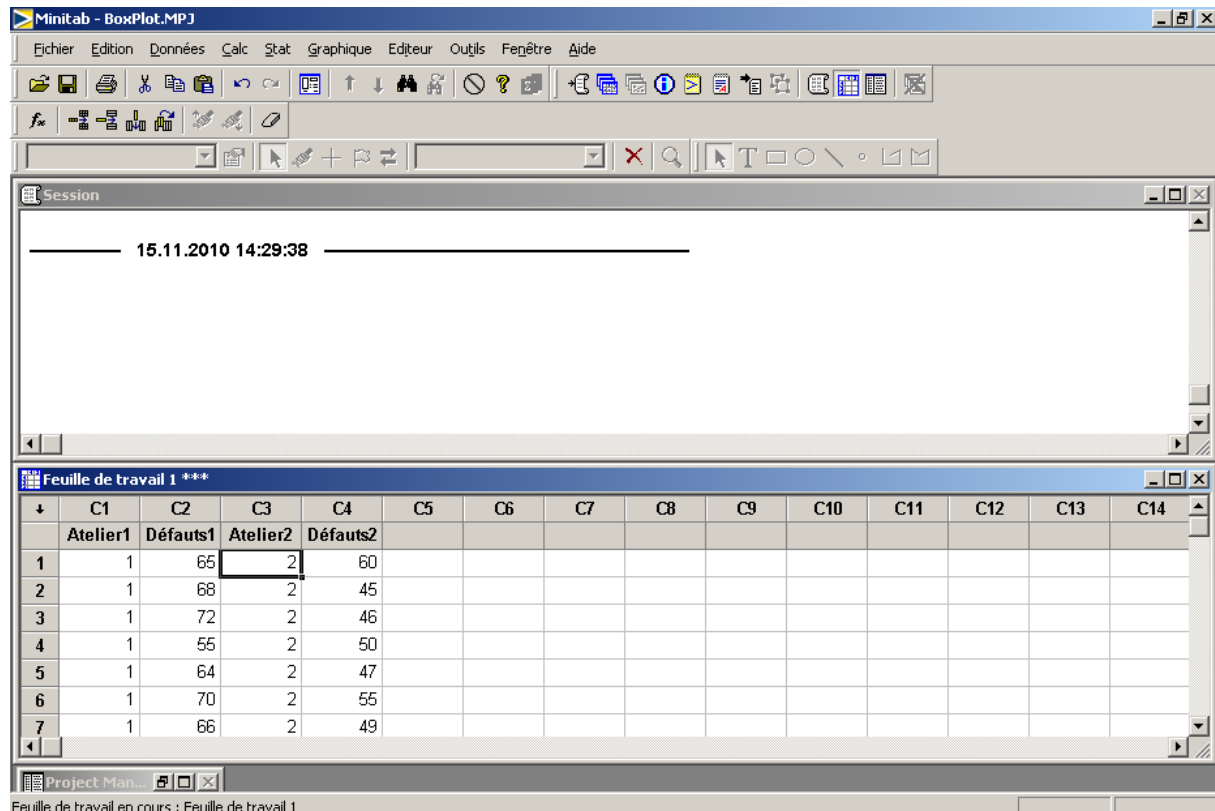
et nous validons par **OK** pour obtenir un résultat loin d'être génial (en plus la rotation dans l'espace est à améliorer...):



## 8.13. Exercice 21.: Graphique boîte à moustaches (BoxPlot)

Minitab® Statistical Software 15.1.1.0

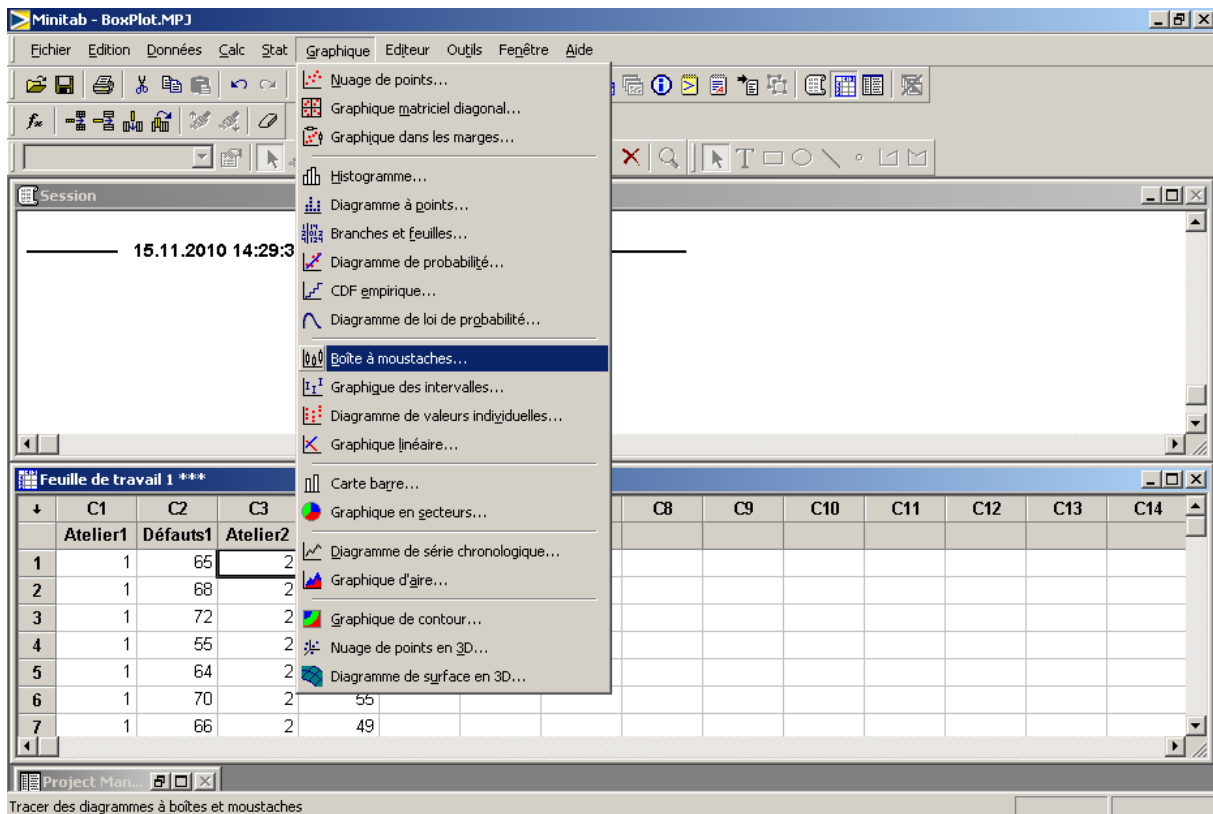
Ouvrez le fichier projet *BoxPlot.mpj*:



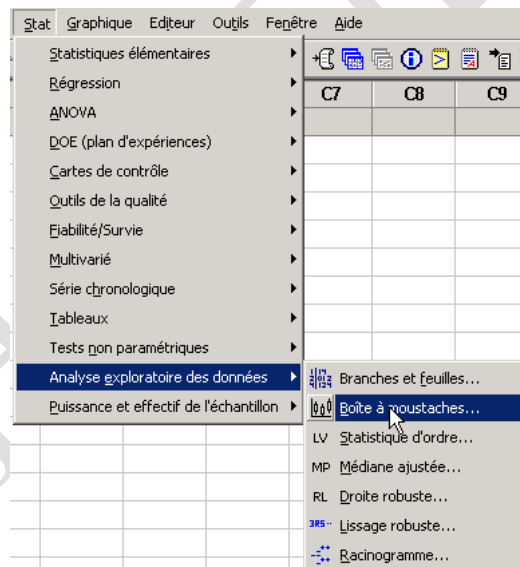
The screenshot displays the Minitab Statistical Software interface. The main window is titled "Minitab - BoxPlot.MPJ" and shows a session window with the date and time "15.11.2010 14:29:38". Below the session window is a spreadsheet titled "Feuille de travail 1 \*\*\*" with the following data:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Atelier1	Défauts1	Atelier2	Défauts2										
1	1	65	2	60										
2	1	68	2	45										
3	1	72	2	46										
4	1	55	2	50										
5	1	64	2	47										
6	1	70	2	55										
7	1	66	2	49										

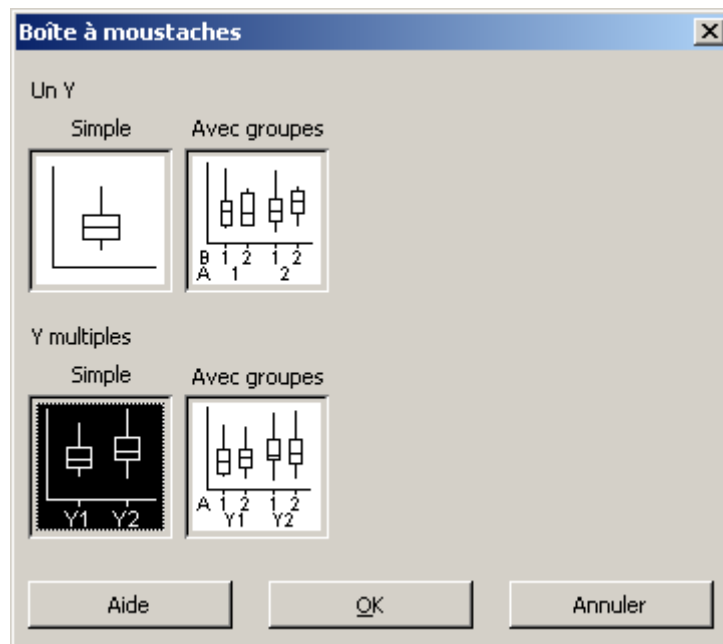
Il s'agit donc des mêmes données que dans les cours de qualité et statistique de Microsoft Excel. Mais cette fois-ci nous souhaiterions faire les box-plot avec Minitab® Statistical Software. Nous allons alors dans le menu **Graph/Boîte à moustaches** ou **Stat/Analyse exploratoire de données/Boîte à moustaches** (puisque les deux chemins mènent au même outil!):



Tracer des diagrammes à boîtes et moustaches

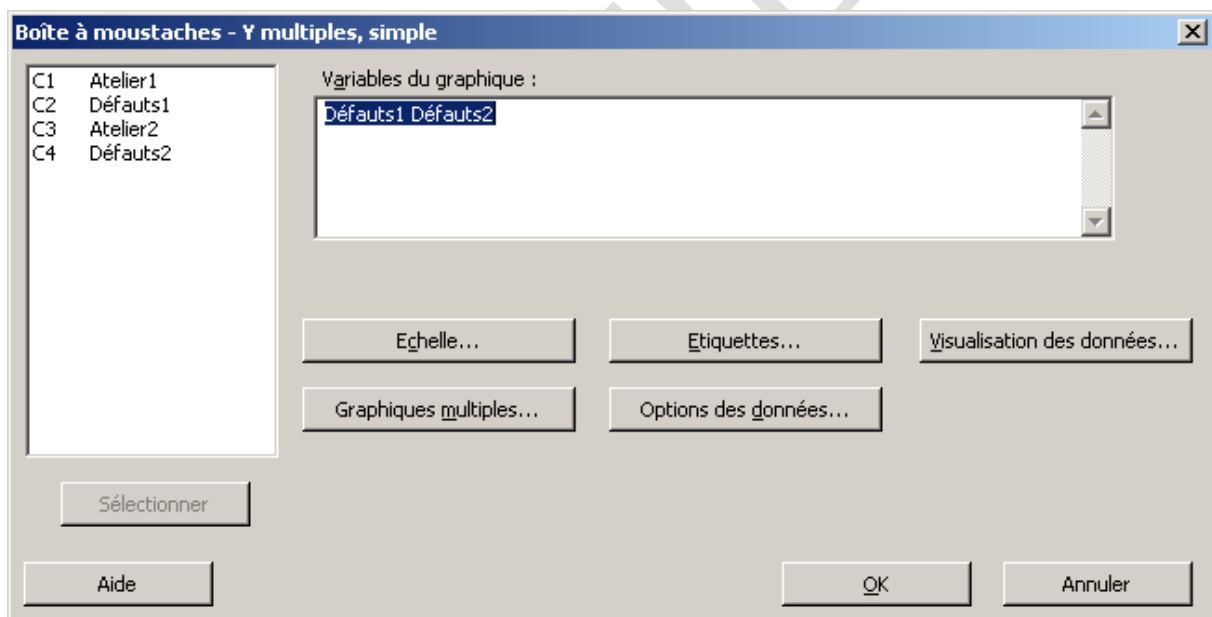


Vient alors:

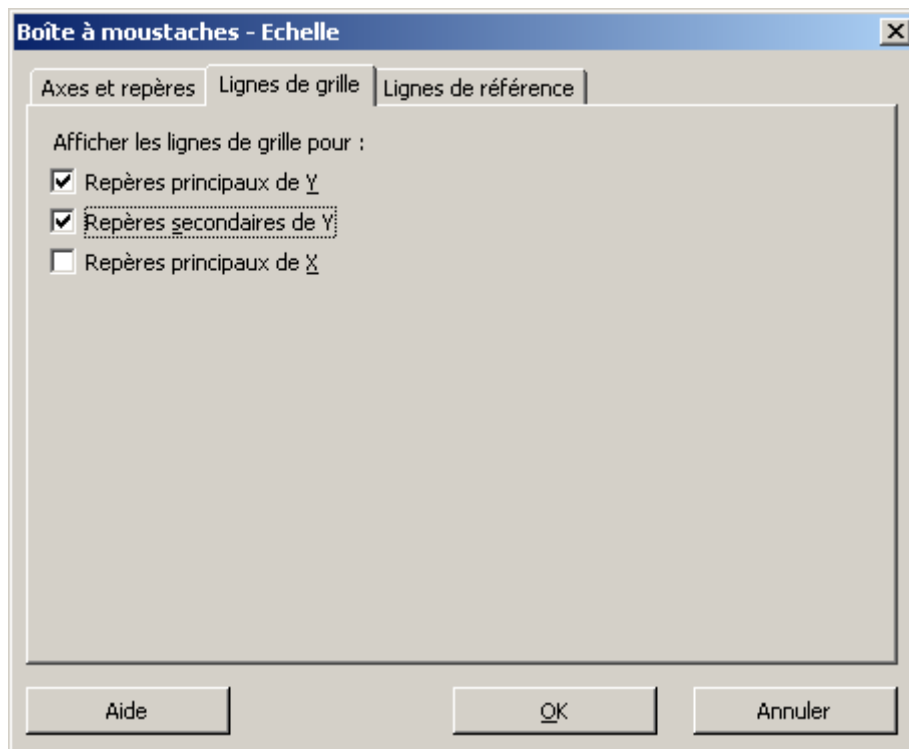


Nous prenons l'option **Y multiples Simple** et validons par **OK**:

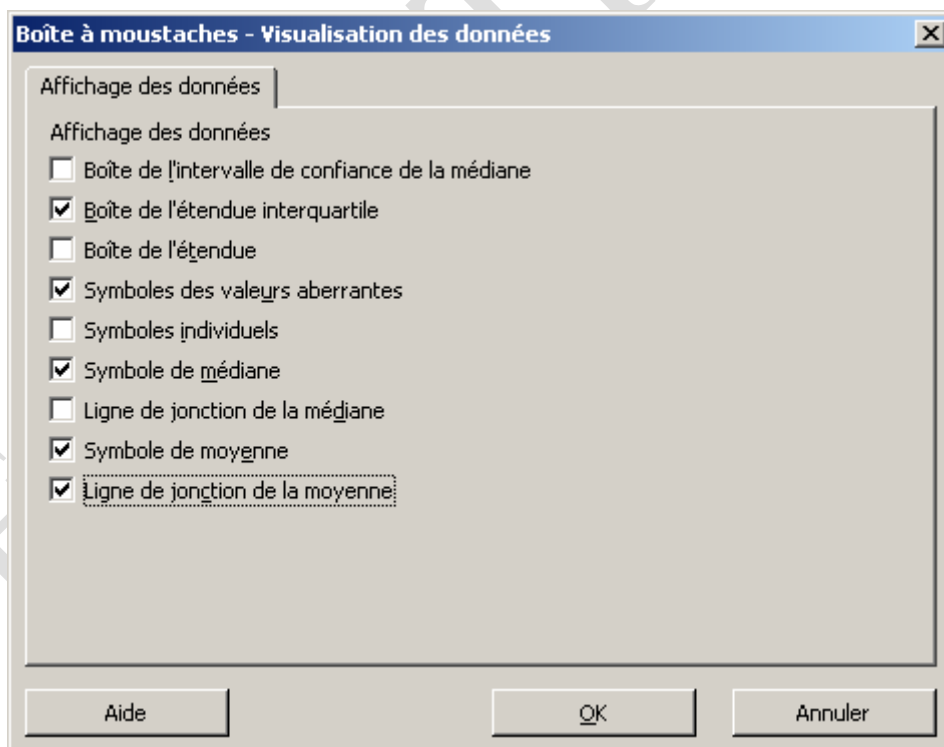
Nous prenons ensuite:



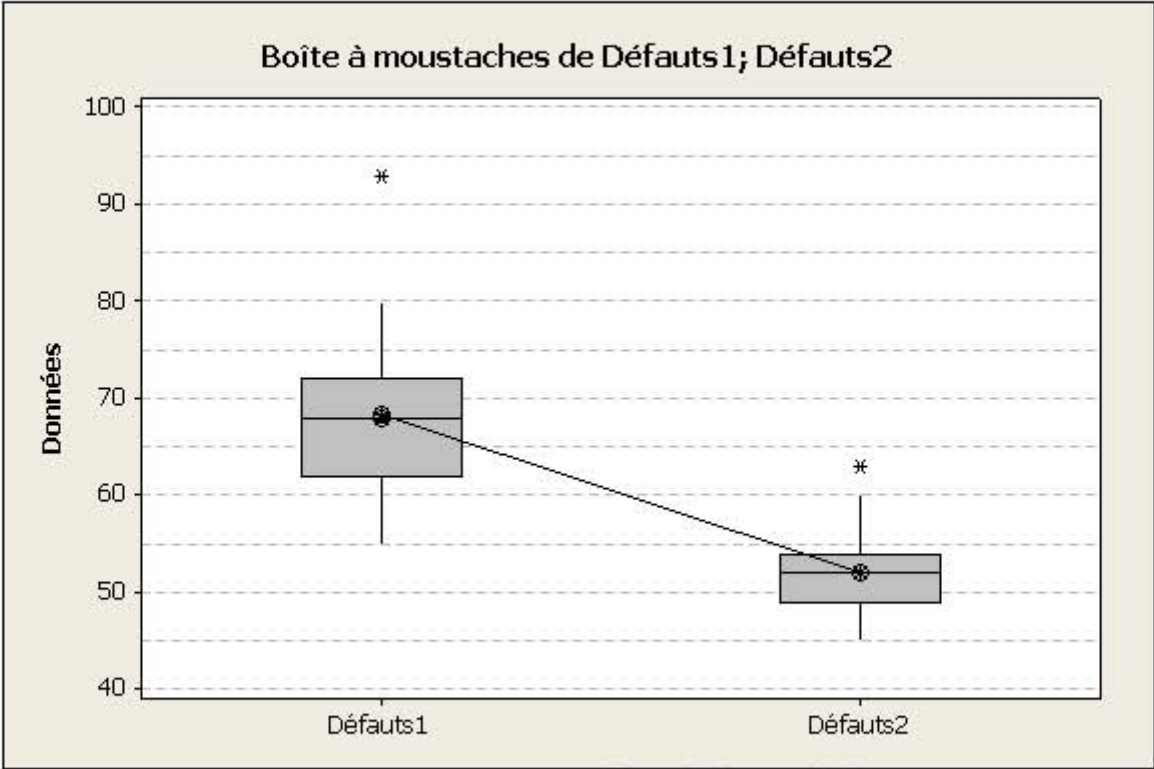
Nous allons dans le bouton **Echelle** pour activer la grille de **Y**:



Nous validons par **OK**. Nous allons ensuite dans le bouton **Visualisation de données** pour y cocher:



Et nous validons deux fois par **OK** pour obtenir au final:





## 8.14. Exercice 22.: Combiner des graphiques (lattice)

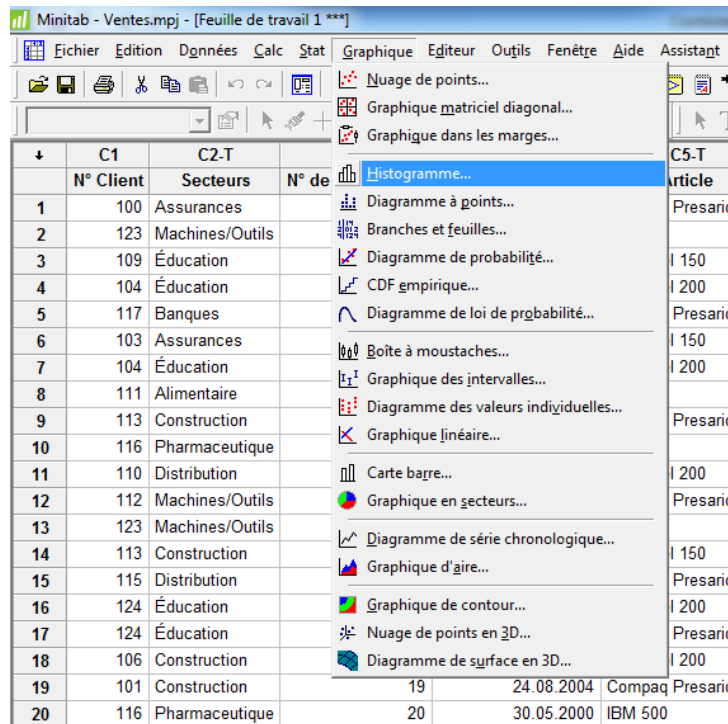
Minitab® Statistical Software 16.1.

Au même titre que dans R et MATLAB nous désirons souvent regrouper des graphiques.

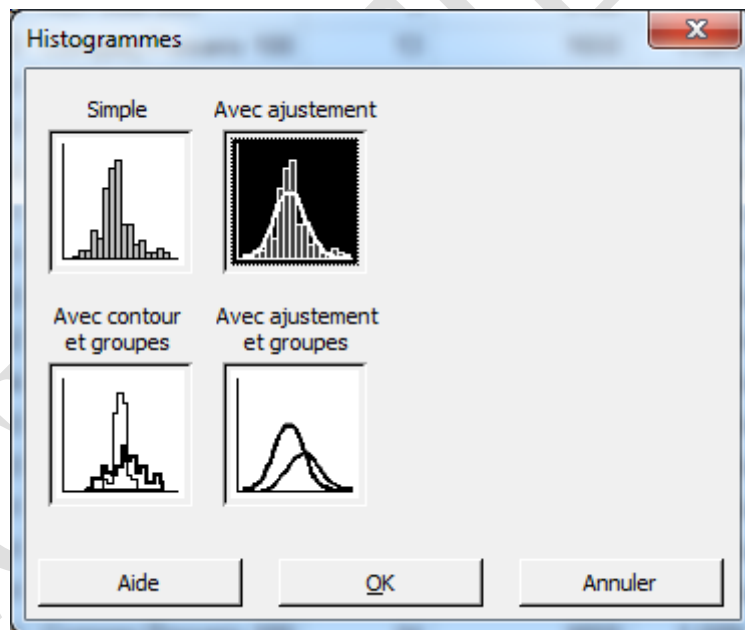
Pour ce faire, il faut d'abord avoir des graphiques (peu importe lesquels). Par exemple construisons un histogramme des quantités à partir du fichier *Ventes.mpj*:

	C1	C2-T	C3	C4-D	C5-T	C6	C7	C8	C9	C10-T	C11	C12	C13
	N° Client	Secteurs	N° de Commande	Date de commande	Article	Quantité	Prix par pièce	Rabais%	Prix total avec rabais	Facture payée			
1	100	Assurances	1	13.07.2000	Compaq Presario 100	12	1650	1.50%	19503.0	Oui			
2	123	Machines/Outils	2	13.08.2004	IBM 500	2	2299	0.00%	4598.0	Oui			
3	109	Éducation	3	07.08.2002	AST Intel 150	5	2690	0.00%	13450.0	Oui			
4	104	Éducation	4	29.08.2003	AST Intel 200	3	3190	0.00%	9570.0	Oui			
5	117	Banques	5	31.01.2003	Compaq Presario 100	13	1650	1.50%	21128.3	Oui			
6	103	Assurances	6	09.02.2001	AST Intel 150	2	2690	0.00%	5380.0	Oui			
7	104	Éducation	7	07.01.2001	AST Intel 200	2	3190	0.00%	6380.0	Oui			
8	111	Alimentaire	8	05.06.2004	IBM 500	4	2299	0.00%	9196.0	Oui			
9	113	Construction	9	26.12.2001	Compaq Presario 100	4	1650	0.00%	6600.0	Oui			
10	116	Pharmaceutique	10	04.11.2001	IBM 500	2	2299	0.00%	4598.0	Oui			
11	110	Distribution	11	23.09.2003	AST Intel 200	6	3190	1.50%	18852.9	Oui			
12	112	Machines/Outils	12	09.02.2002	Compaq Presario 100	6	1650	1.50%	9751.5	Oui			
13	123	Machines/Outils	13	27.12.2004	IBM 500	6	2299	1.50%	13587.1	Oui			
14	113	Construction	14	17.07.2003	AST Intel 150	3	2690	0.00%	8070.0	Oui			
15	115	Distribution	15	21.01.2003	Compaq Presario 100	8	1650	1.50%	13002.0	Oui			
16	124	Éducation	16	20.10.2001	AST Intel 200	8	3190	1.50%	25137.2	Oui			
17	124	Éducation	17	07.03.2001	Compaq Presario 100	11	1650	1.50%	17877.8	Oui			
18	106	Construction	18	06.06.2001	AST Intel 200	11	3190	1.50%	34563.7	Oui			
19	101	Construction	19	24.08.2004	Compaq Presario 100	14	1650	1.50%	22753.5	Non			
20	116	Pharmaceutique	20	30.05.2000	IBM 500	7	2299	1.50%	15851.6	Non			
21	112	Machines/Outils	21	23.06.2001	AST Intel 150	6	2690	1.50%	15897.9	Oui			
22	125	Construction	22	29.12.2002	Compaq Presario 100	23	1650	3.00%	36811.5	Oui			
23	100	Assurances	23	13.04.2003	IBM 500	3	2299	0.00%	6897.0	Oui			
24	125	Construction	24	21.09.2001	AST Intel 200	2	3190	0.00%	6380.0	Oui			
25	104	Éducation	25	29.06.2002	AST Intel 150	12	2690	1.50%	31795.8	Oui			
26	126	Machines/Outils	26	15.04.2000	AST Intel 150	24	2690	3.00%	62623.2	Oui			
27	121	Pharmaceutique	27	07.05.2002	IBM 500	8	2299	1.50%	18116.1	Non			

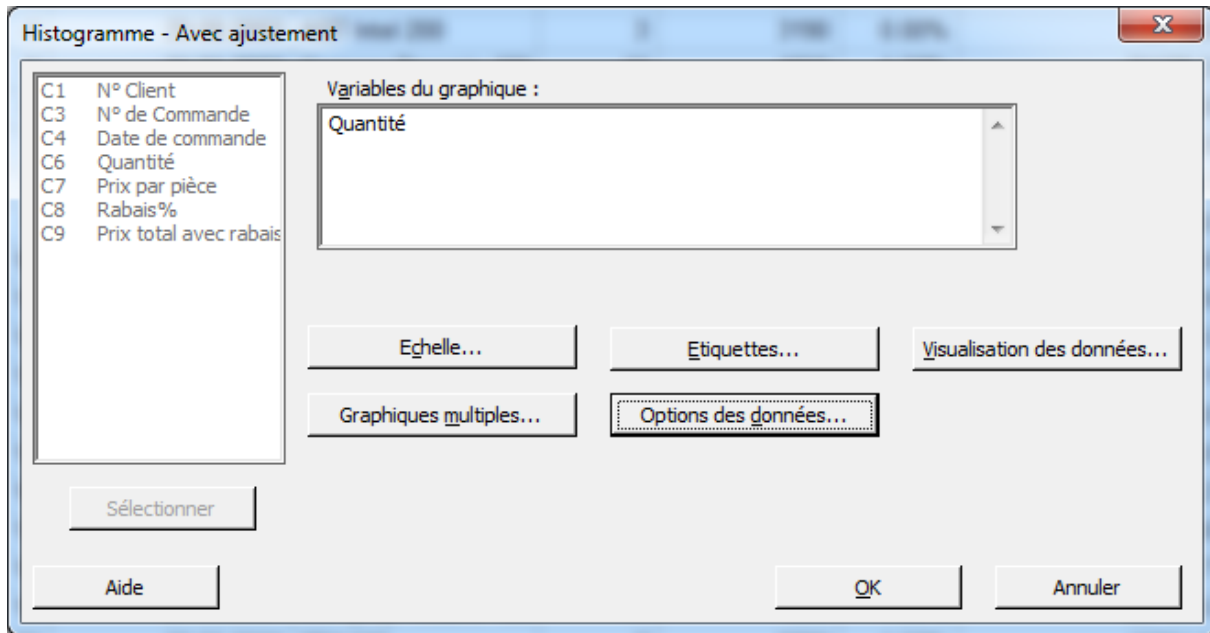
Nous allons donc dans **Graphique/Histogramme...**:



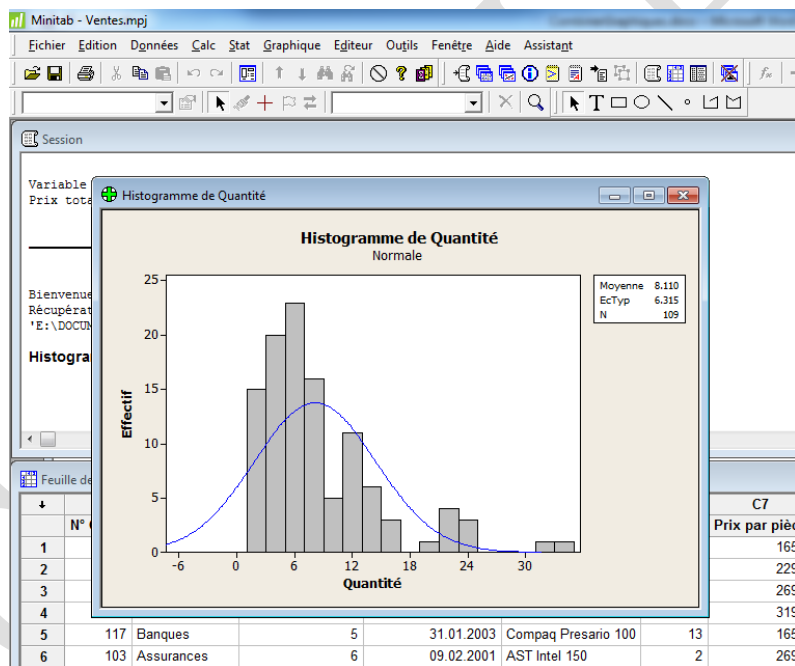
Nous prenons **Avec ajustement**:



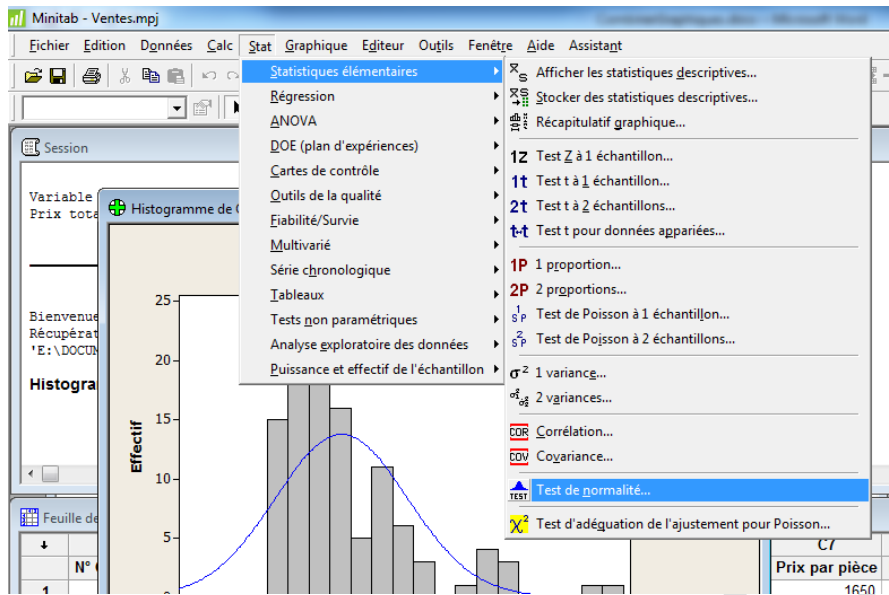
nous choisissons de faire cela sur la colonne *Quantité*:



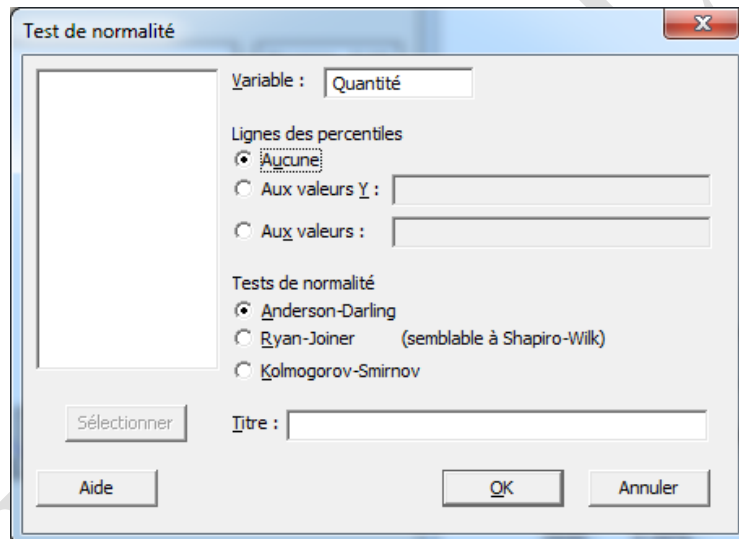
Ce qui donne classique le graphique suivant:



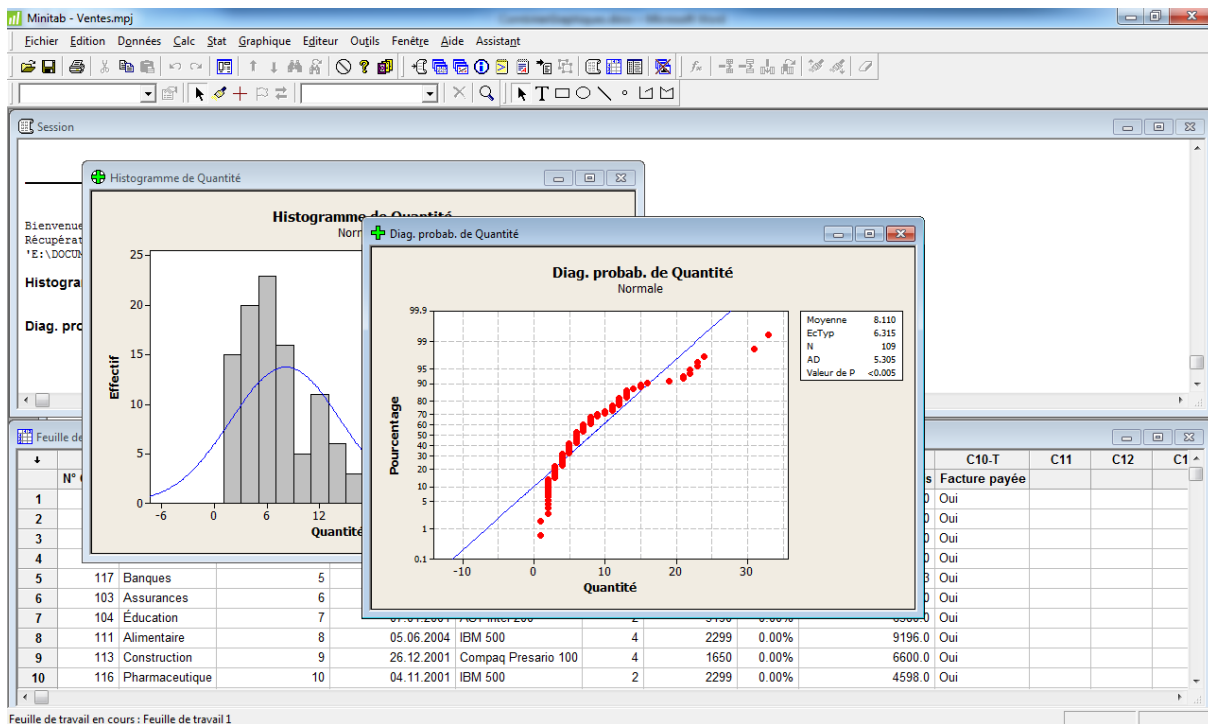
Nous faisons direct derrière un autre graphique en prenant un peu au hasard le graphique qui se génère d'un test de Normalité en allant **Stat/Statistiques élémentaires/Test de normalité...**



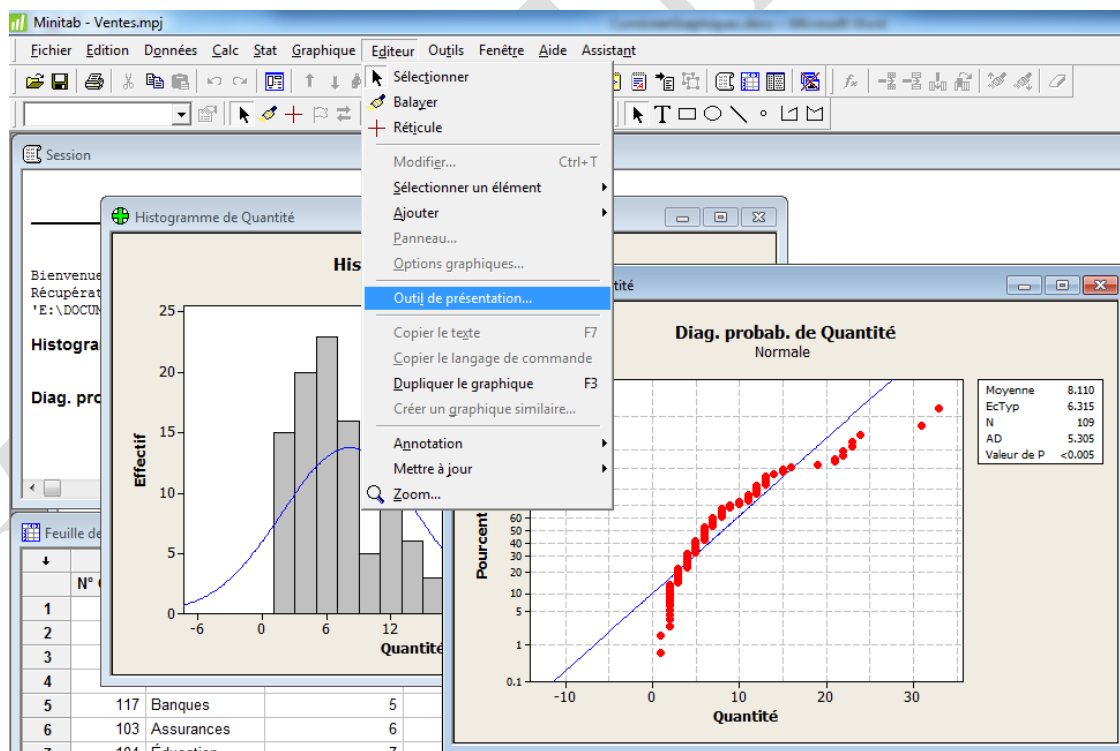
Nous prenons au hasard aussi sur la *Quantité*:



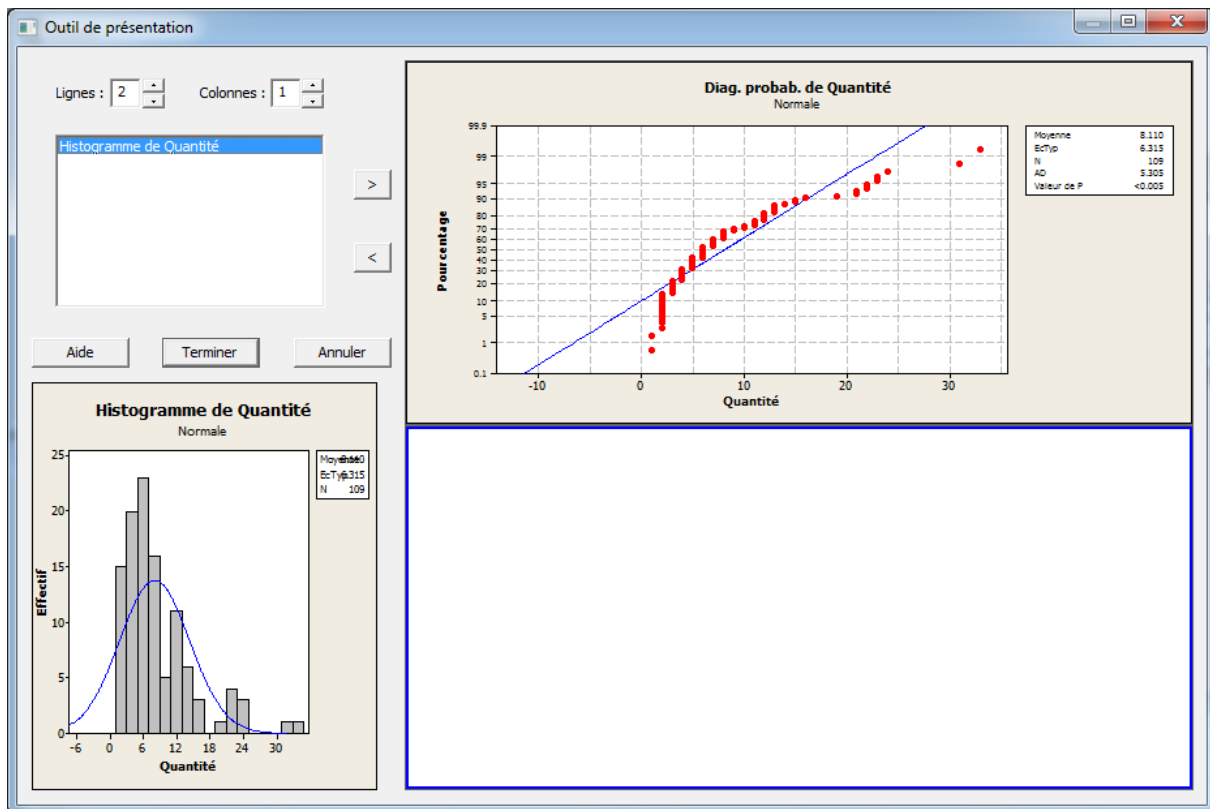
Cela donne donc deux graphiques indépendants:



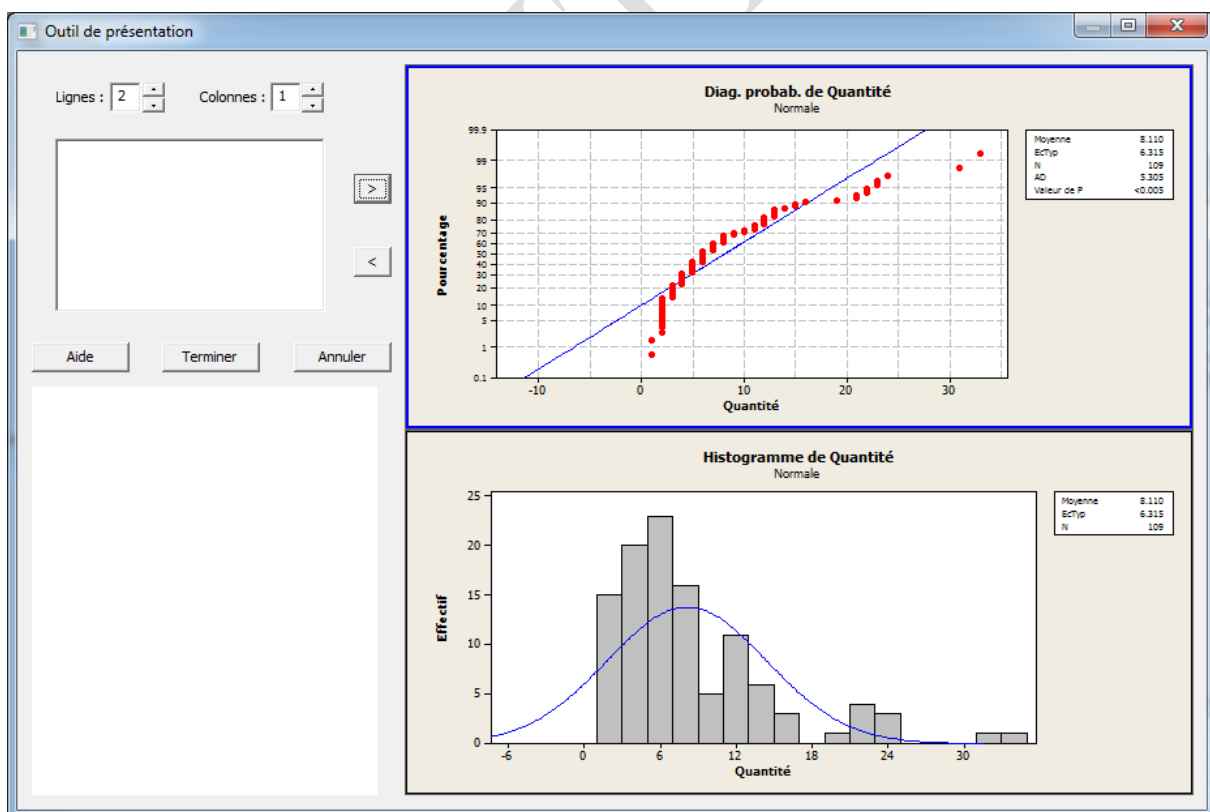
Pour les combiner dans une fenêtre unique, nous allons dans le menu **Editeur/Outil de présentation...**:



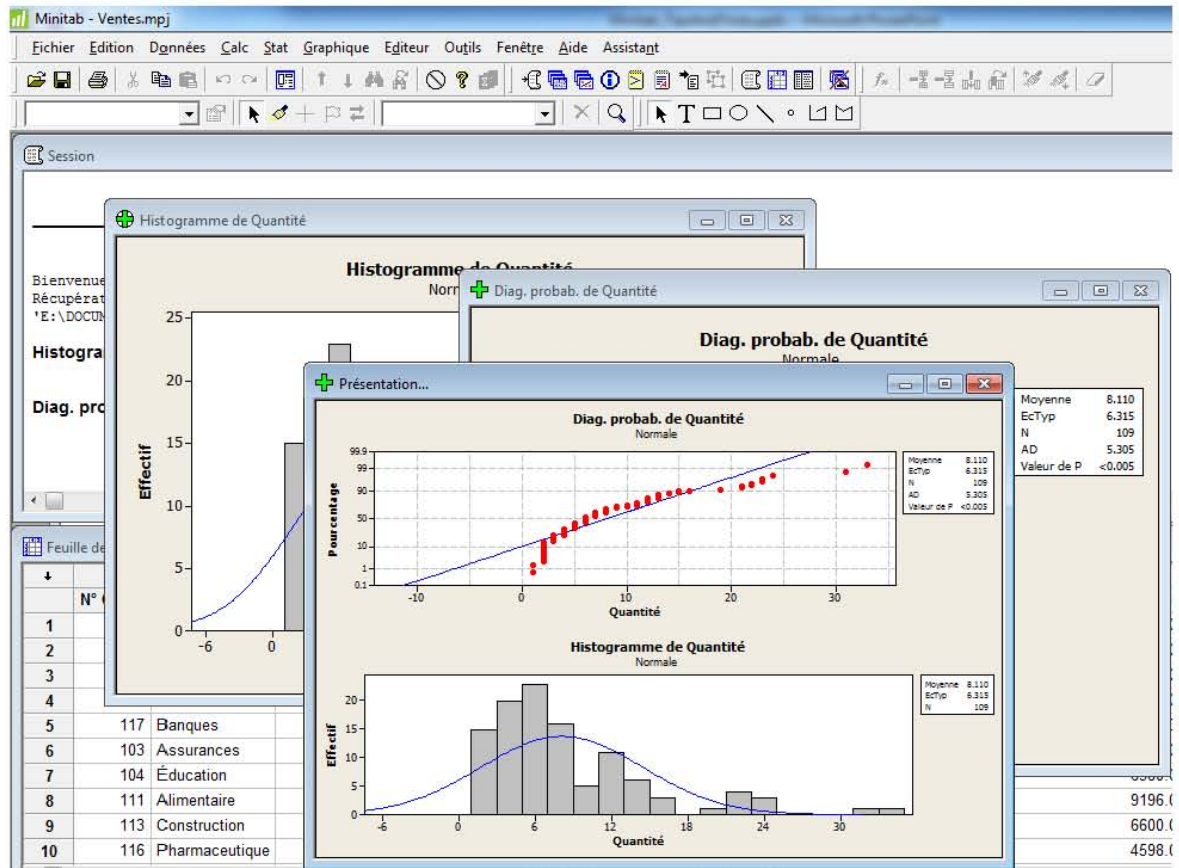
Apparaît alors la boîte de dialogue suivante:



où nous pouvons choisir le nombre de colonnes et de lignes de la lattice (ou subplots):



En cliquant sur **Terminer** nous obtenons:



## 8.15. Exercice 23.: Diagramme quantile-quantile (Q-Q plot/qq-plot)

Minitab® Statistical Software 15.1.1.0

Le but ici va être de tracer un ce que de nombreux utilisateurs de Minitab appellent un Q-Q plots alors qu'en réalité Minitab ne saif fait qu'un demi Q-Q Plot.

Voyons un exemple avec des données quelconques donc nous souhaiterions contrôler visuellement l'adéquation avec une loi Normale et en même temps avoir un test de Normalité d'Anderson-Darling.

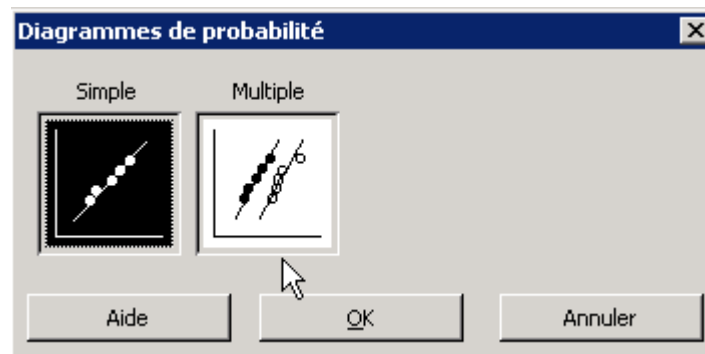
Nous allons dans le menu **Diagramme de probabilité...**:

	C1	C2	C3
	Frc_Inj_Max	Frc_Inj_Moy	
1	13.2800	25.8500	
2	14.1200	25.4500	
3	14.2500	25.3800	
4	14.4400	25.2600	
5	14.7000	25.2500	
6	14.8000	25.2200	
7	14.9000	25.0600	
8	14.9000	25.0100	
9	14.9000	24.8100	
10	15.0000	24.7500	
11	15.1000	24.7400	
12	15.1000	24.7000	
13	15.1000	24.6996	
14	15.2000	24.6700	
15	15.2000	24.6500	
16	15.2570	24.5800	
17	15.3000	24.5700	
18	15.3000	24.5600	
19	15.3700	24.5100	
20	15.3720	24.5100	
21	15.4000	24.4000	
22	15.5000	24.2955	

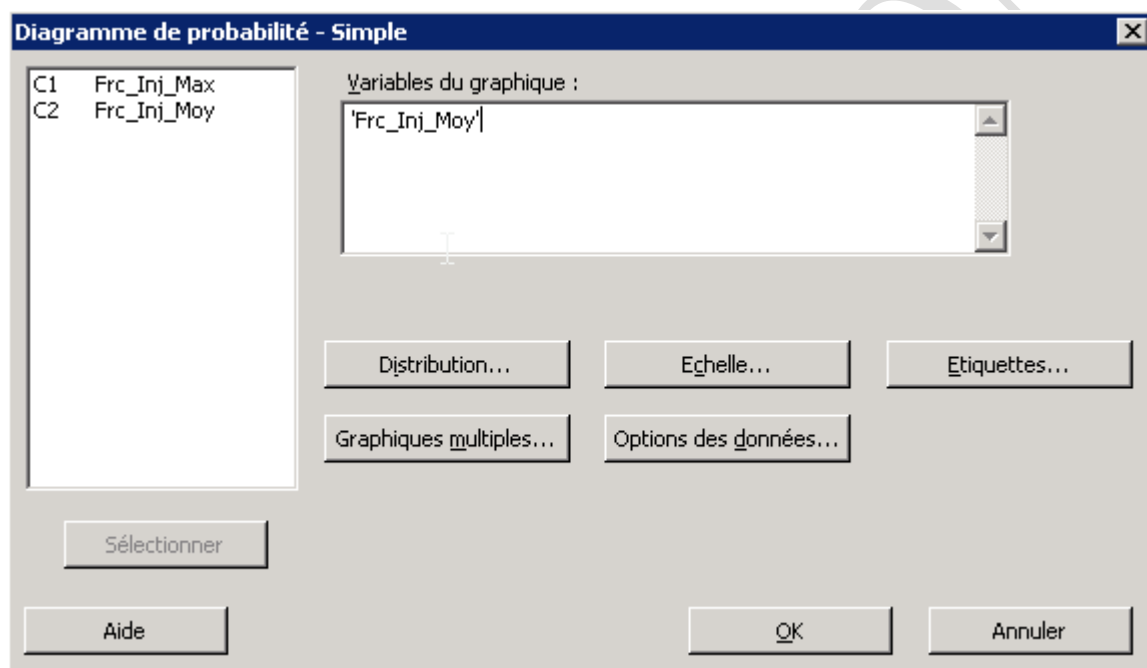
Tracer des droites de Henry, des graphiques de probabilité log-normale, des graphiques de probabilité de Weibull ou des di



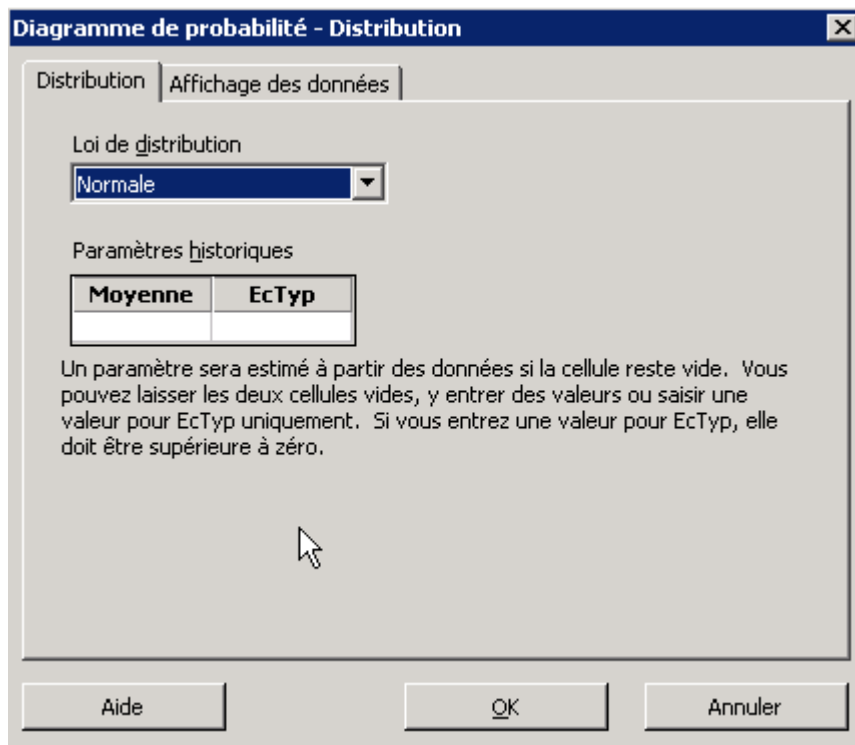
Ensuite, nous choisissons **Simple** et validons par **OK**:



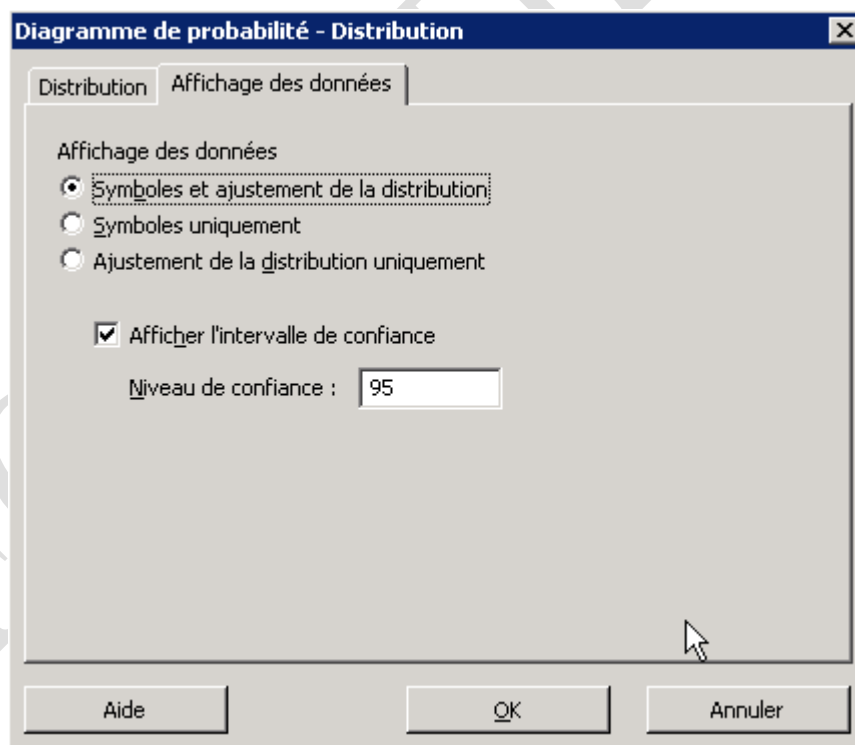
et nous prenons la variable d'intérêt:



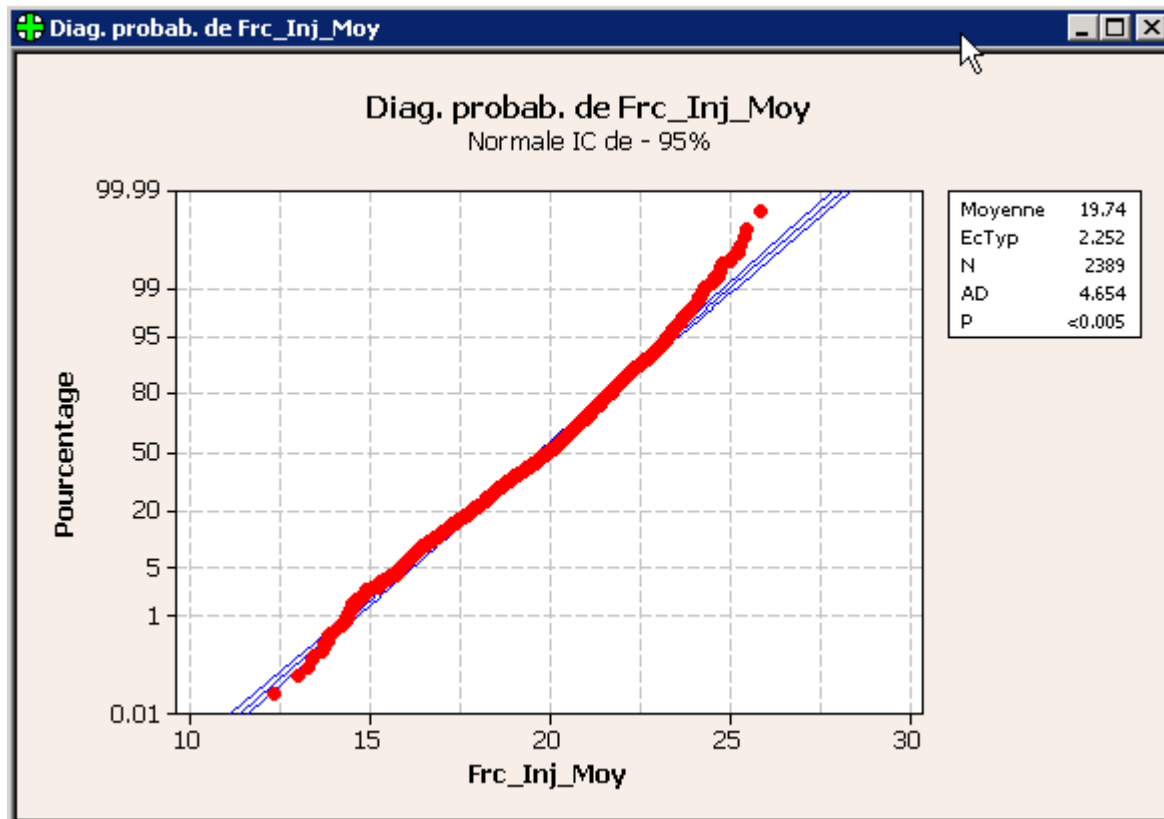
dans le bouton **Distribution...**, nous prenons **Normale**:



Ensuite dans l'onglet **Affichage des données** nous prenons:



Nous validons le tout par **OK** pour obtenir:



et voilà... rien de spécial à dire à part que ce n'est pas un vrai Q-Q plot contrairement à ce que beaucoup disent (mais c'est un détail).

## 8.16. Exercice 24.: Générer le graphique de la loi de densité de probabilité

Minitab® Statistical Software 15.1.1.0

Il arrive souvent que l'on calcule les paramètres lois de probabilités afin de prendre des décisions relativement aux données de la fonction de répartition.

Voyons comment procéder par exemple pour une loi Normale d'espérance 25 et d'écart-type de 3.

Nous allons dans le menu **Graphique/Diagramme de loi de probabilité....**:

The screenshot shows the Minitab interface with the 'Graphique' menu open. The 'Diagramme de loi de probabilité...' option is highlighted. The main window displays a normal distribution function and a data table.

Session

Normale avec moyenne = 25 et écart-type = 3

$$P(X \leq x) = \Phi\left(\frac{x - 25}{3}\right)$$

0.66	26.2374
------	---------

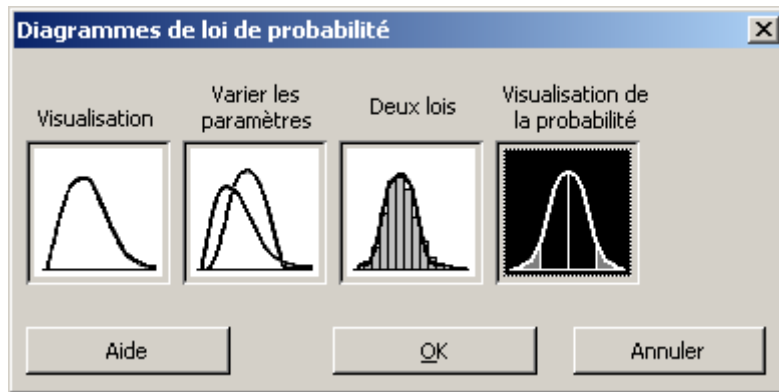
Diagramme de répartition

	C1	C2	C3
1	0.66		
2			
3			
4			
5			
6			
7			

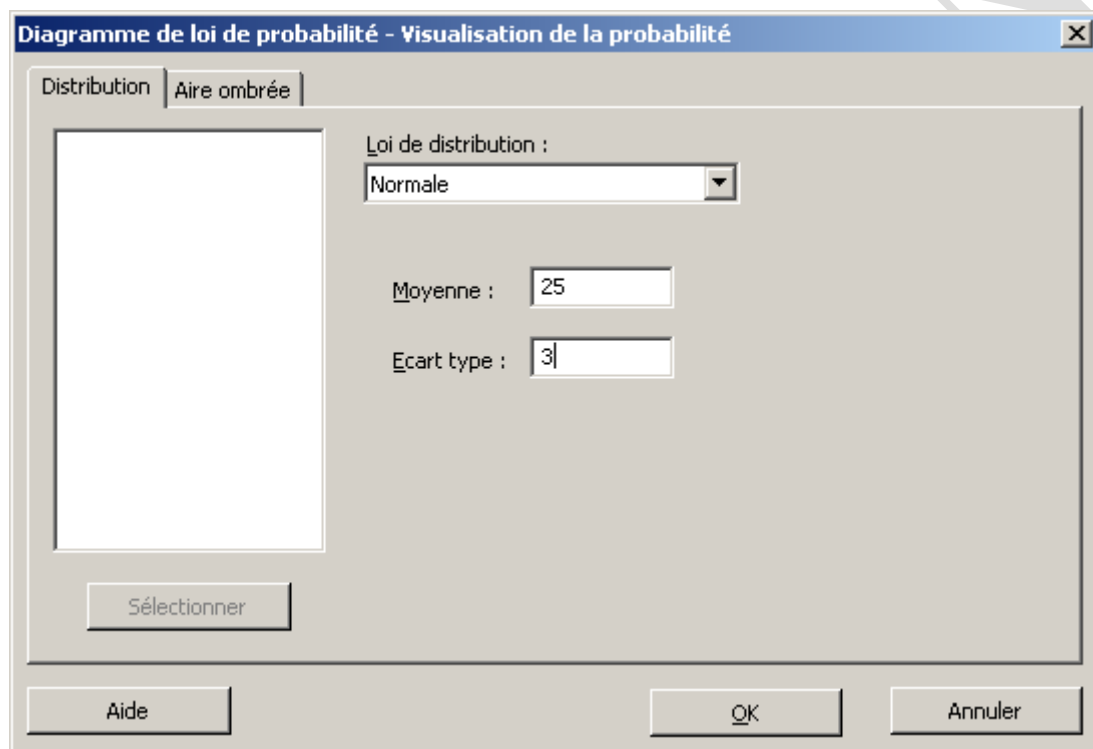
Tracer des diagrammes de loi de probabilité

Démarrer | Minitab - Microsoft Word | Minitab - BoxPlot.MPJ | 15:18

Nous prenons une option qui est longue à faire avec Microsoft Excel (à titre de comparaison c'est plus long dans Excel mais au moins on peut en superposer autant que l'on veut dans Excel ce qui est très utile dans la pratique et qui manque à Minitab):

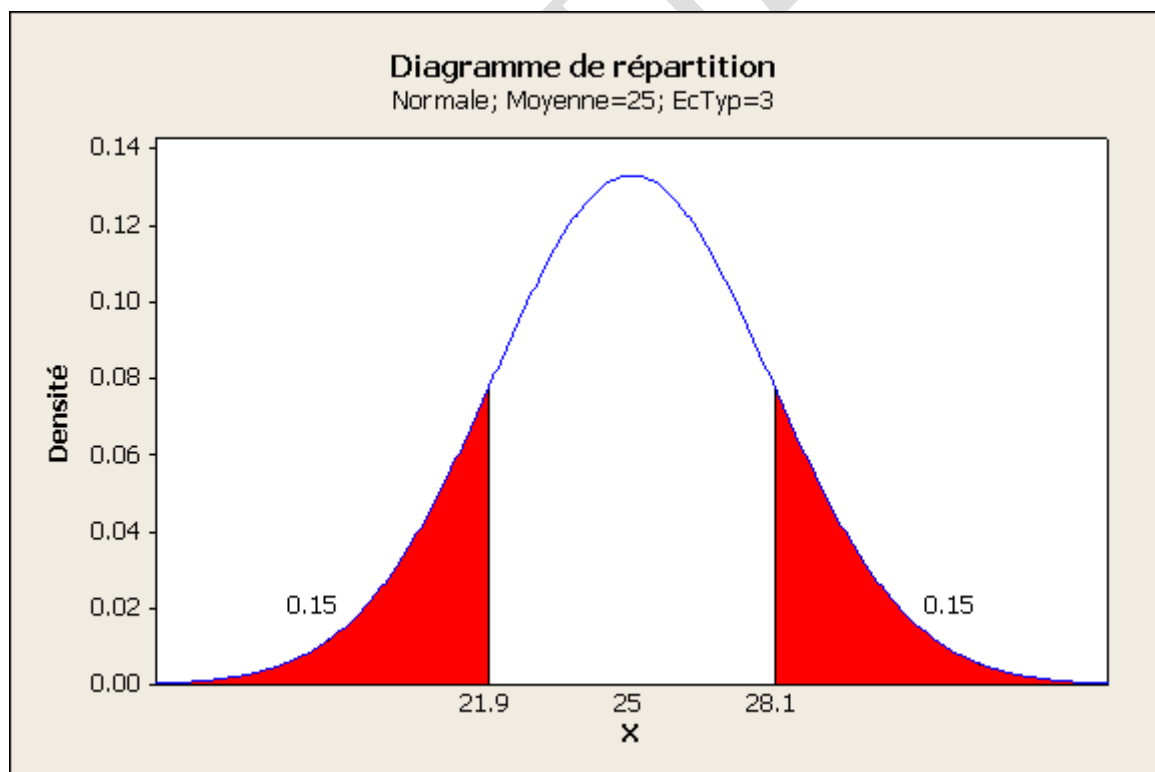


Après avoir choisi **Visualisation de la probabilité**, nous validons par **OK**:

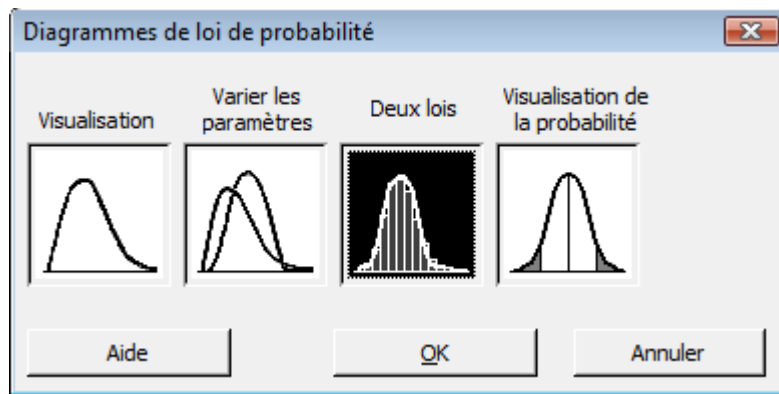




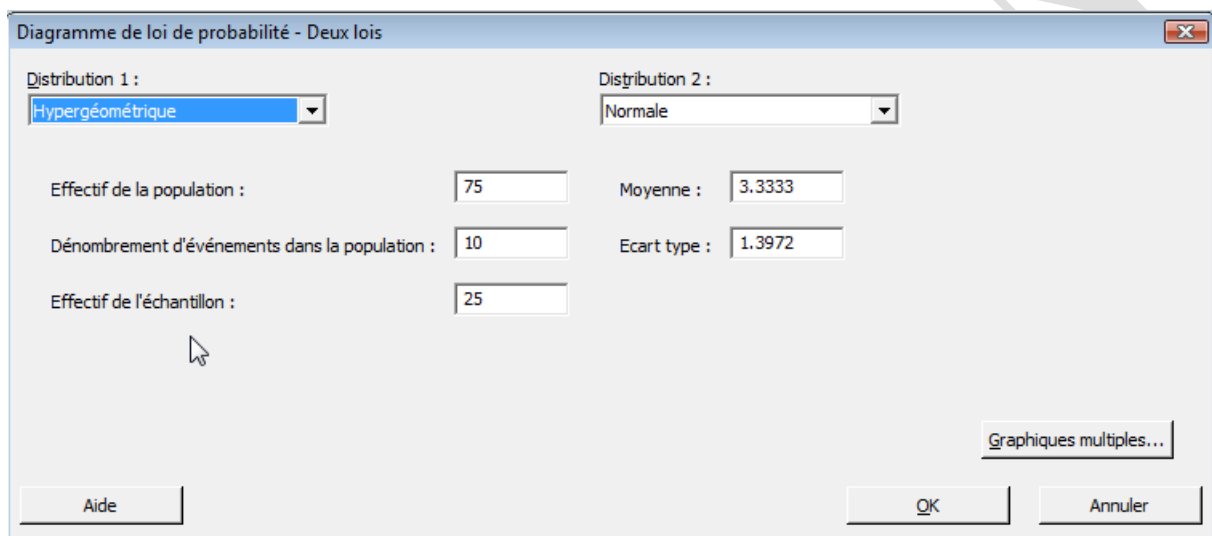
Et nous validons deux fois par **OK** pour avoir au final:



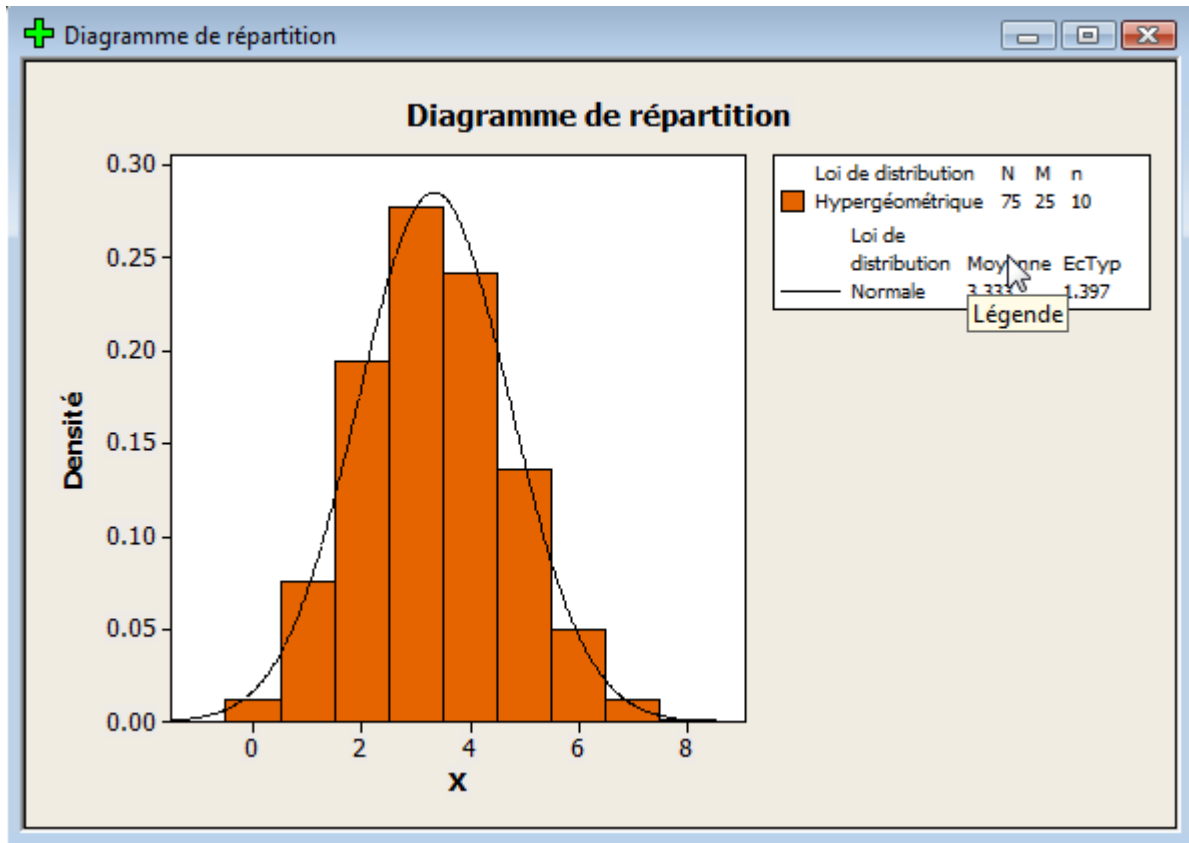
ou encore plus utile, toujours avec le même menu mais cette fois-ci nous choisissons l'option **Deux lois**:



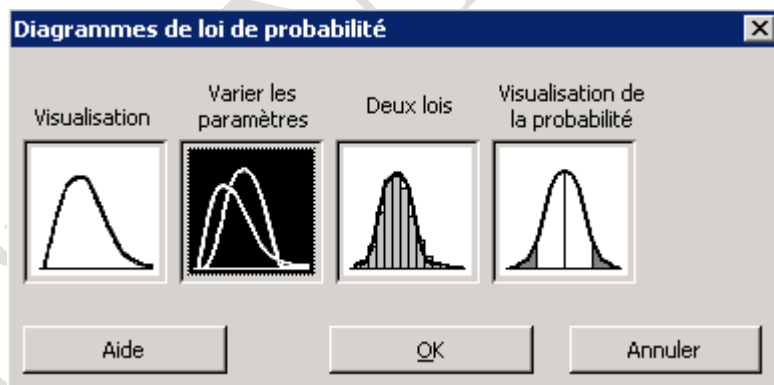
et validons par **OK**:



où nous voulons comparer par exemple une distribution hypergéométrique avec la loi Normale correspondante comme en cours de statistique théorique et nous avons vu qu'il était très malaisé de faire cela avec Microsoft Excel. Il vient alors:

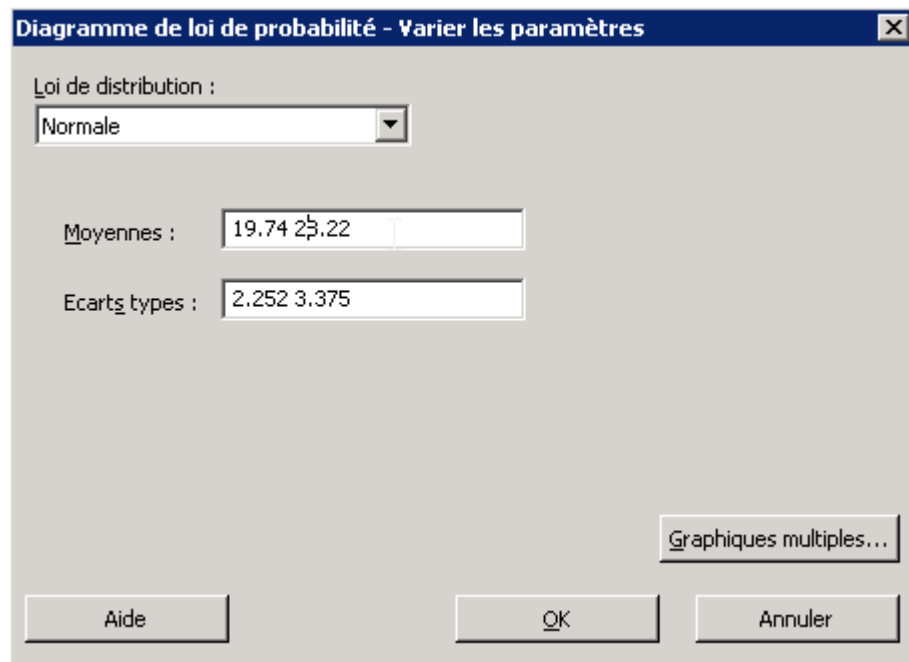


ou encore un autre cas qui a été très utile pour un de mes clients en prenant **Varié les paramètres**:

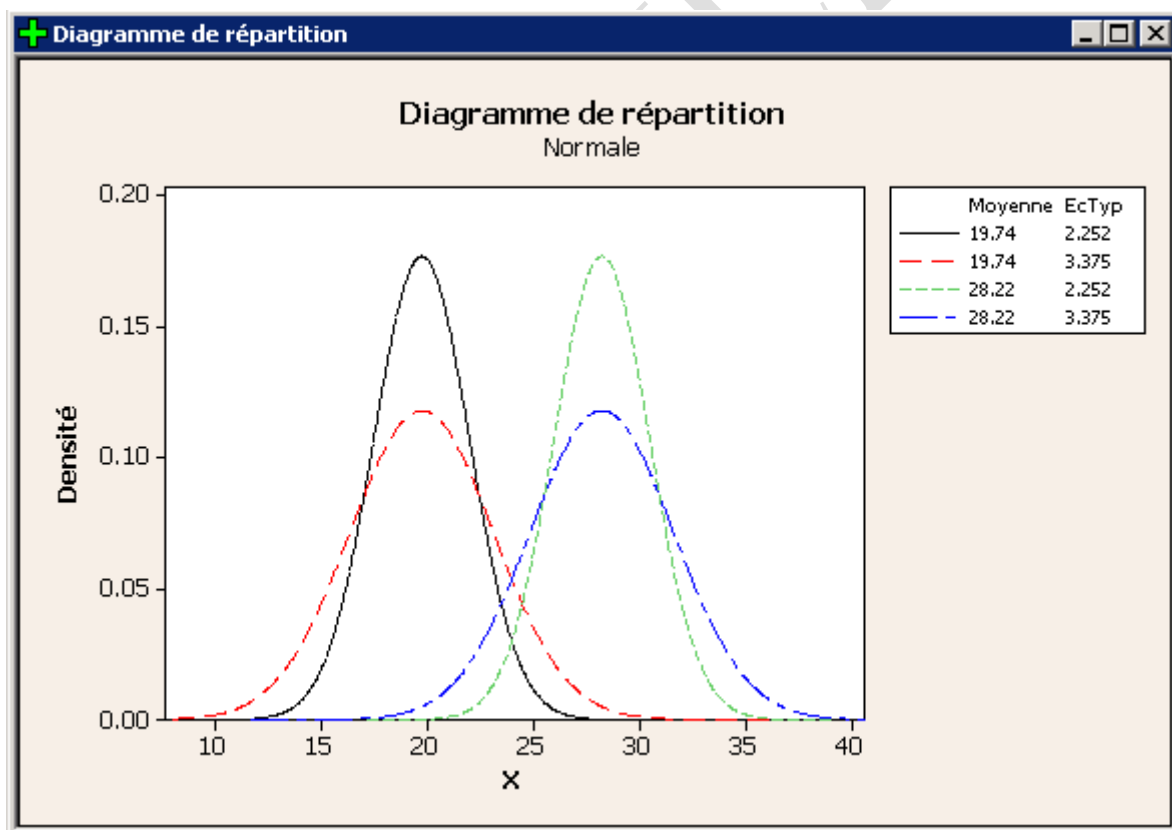


Ensuite on met:





pour obtenir:



et on supprimer les deux tracés qui ne nous intéressent pas (clic dessus avec la souris et **Delete**):

## 8.17. Exercice 25.: Analyse de Pareto

Minitab® Statistical Software 15.1.1.0

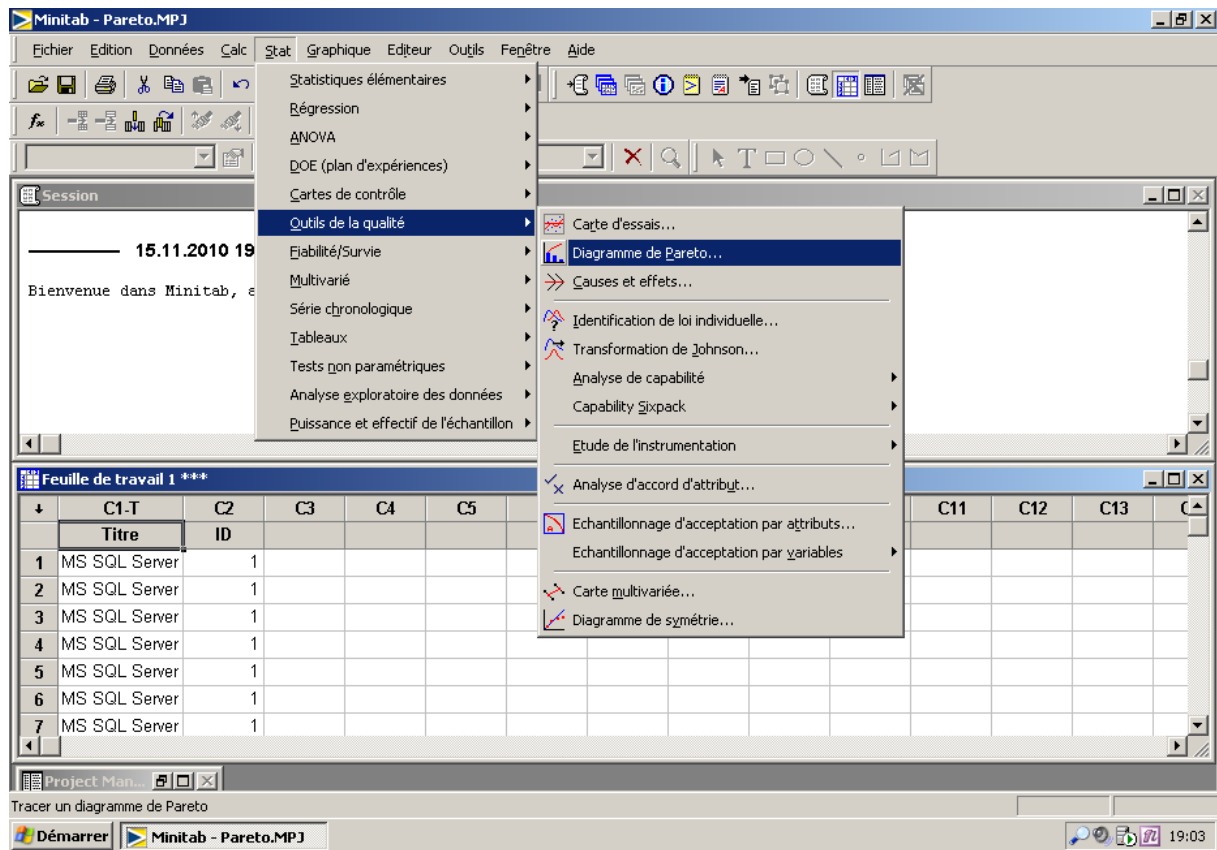
Une entreprise prestataire services vend des formations continues en entreprise. Elle a extrait de sa base de données la liste de tous les cours vendus en ne prenant que le nom du domaine concerné et le numéro d'identifiant correspondant (car Microsoft Excel ne sait pas faire d'analyse de Pareto automatique sur une base autre que des chiffres). Ce qui correspond à 2528 lignes de données avec 43 types de formations (ID) dans un fichier nommé *Pareto.mpj*:

The screenshot shows the Minitab Statistical Software 15.1.1.0 interface. The main window displays a data table with the following structure:

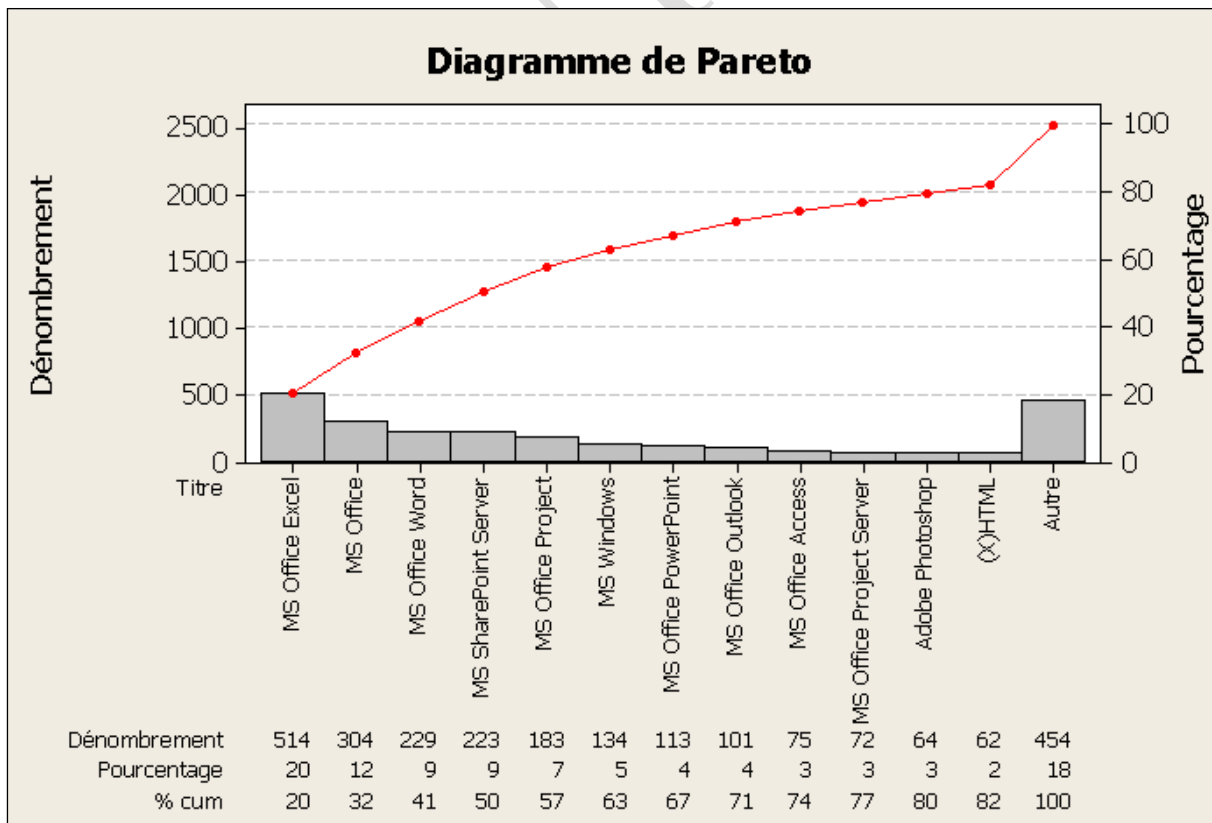
	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Titre	ID												
1	MS SQL Server	1												
2	MS SQL Server	1												
3	MS SQL Server	1												
4	MS SQL Server	1												
5	MS SQL Server	1												
6	MS SQL Server	1												
7	MS SQL Server	1												

The session window shows the date and time: 15.11.2010 19:01:57. The status bar indicates the current worksheet is 'Feuille de travail en cours : Feuille de travail 1'.

Ensuite, nous allons dans le menu **Stat/Outils de qualité/Diagramme de Pareto...**:



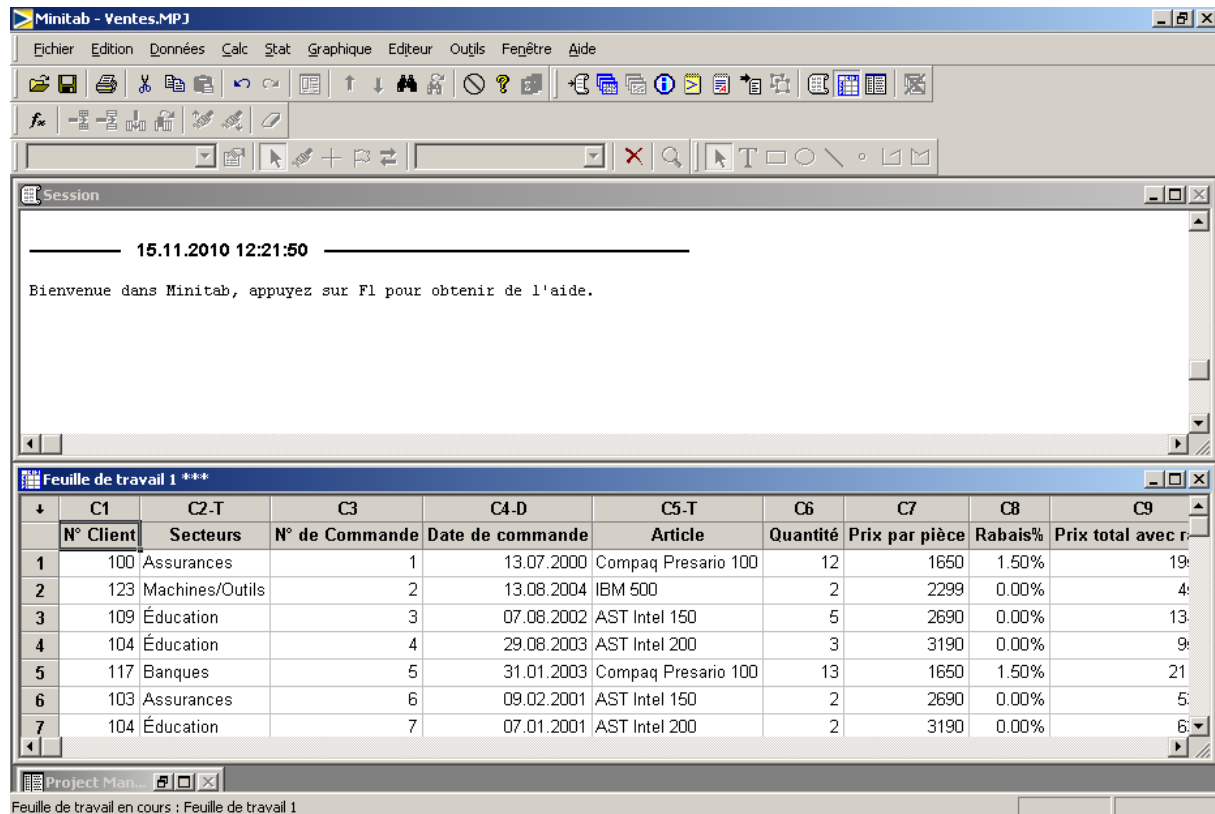
Nous obtenons alors:



## 8.18. Exercice 26.: Schéma branche et feuilles

Minitab® Statistical Software 15.1.1.0

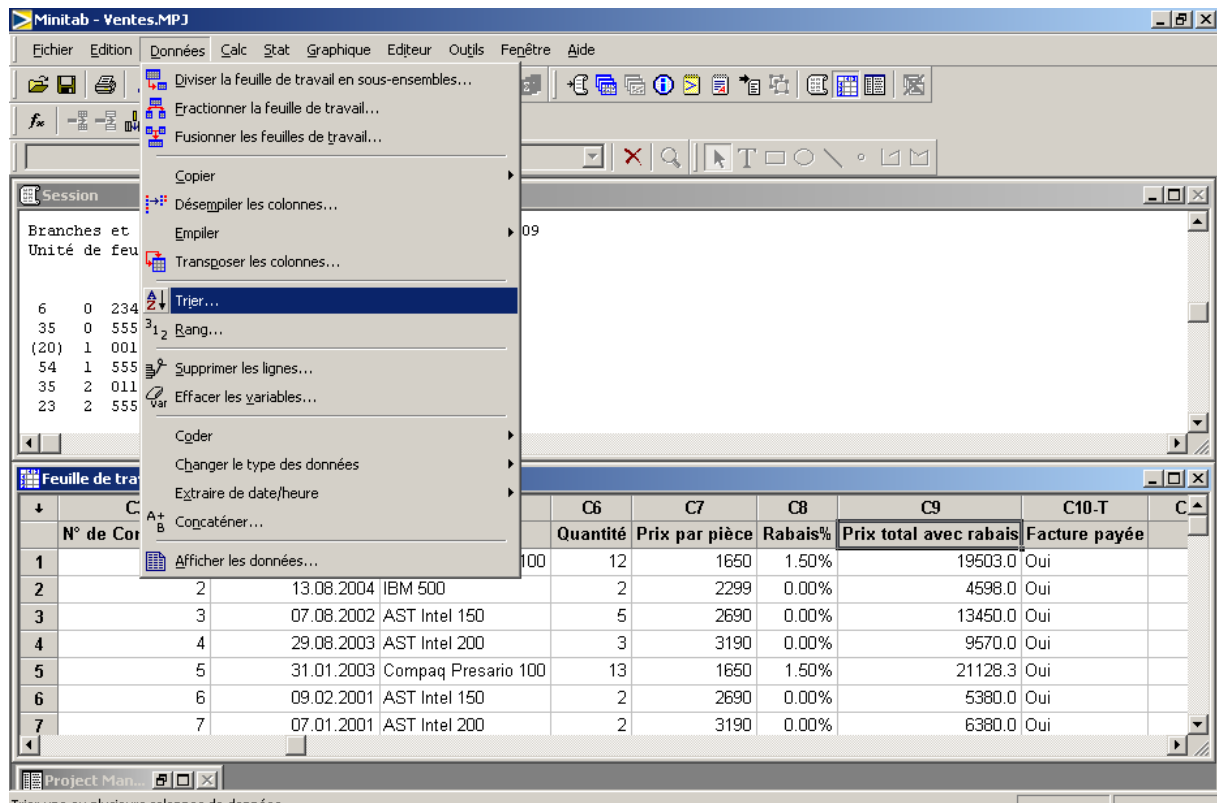
Ouvrez le fichier projet *Ventes.mpj*:



The screenshot shows the Minitab interface. The top window is the 'Session' window, displaying the date and time '15.11.2010 12:21:50' and a welcome message: 'Bienvenue dans Minitab, appuyez sur F1 pour obtenir de l'aide.' Below this is the 'Feuille de travail 1 \*\*\*' window, which contains a data table with 9 columns and 7 rows of data.

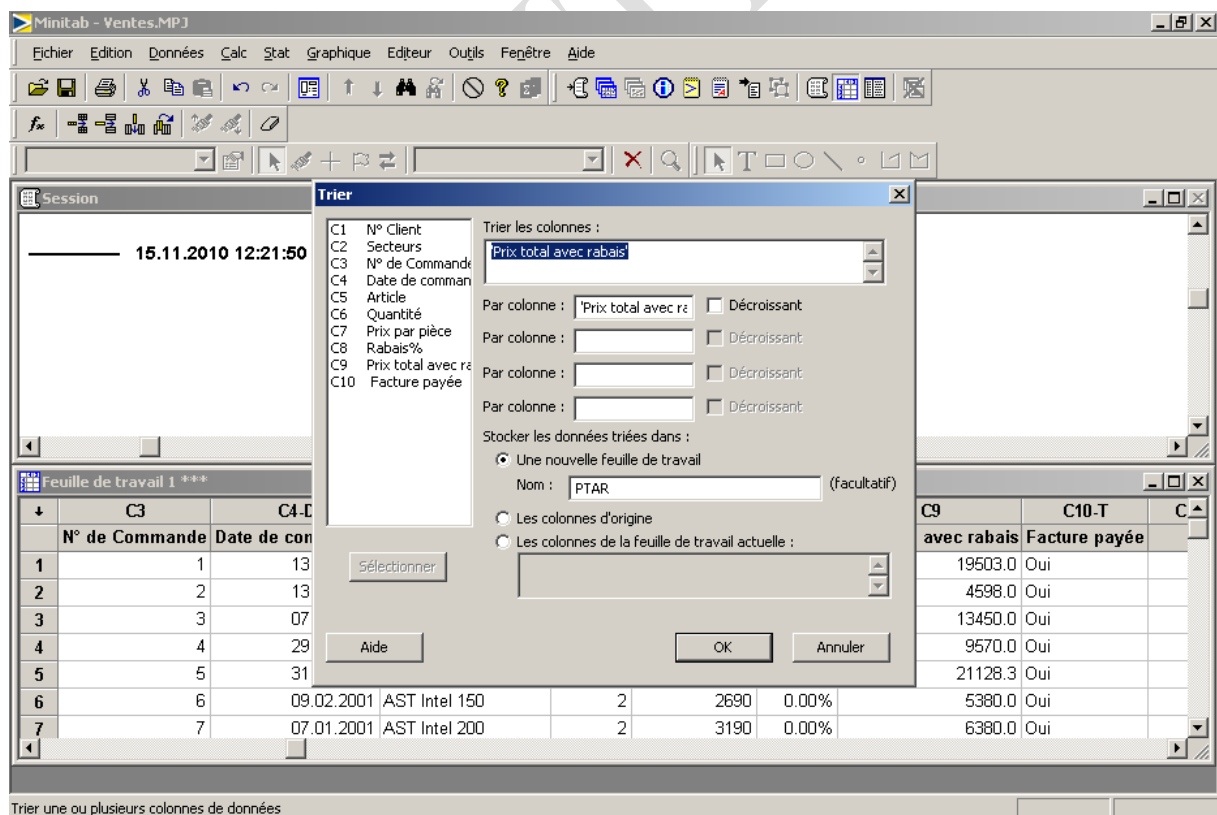
	C1	C2-T	C3	C4-D	C5-T	C6	C7	C8	C9
	N° Client	Secteurs	N° de Commande	Date de commande	Article	Quantité	Prix par pièce	Rabais%	Prix total avec r
1	100	Assurances	1	13.07.2000	Compaq Presario 100	12	1650	1.50%	19:
2	123	Machines/Outils	2	13.08.2004	IBM 500	2	2299	0.00%	4:
3	109	Éducation	3	07.08.2002	AST Intel 150	5	2690	0.00%	13:
4	104	Éducation	4	29.08.2003	AST Intel 200	3	3190	0.00%	9:
5	117	Banques	5	31.01.2003	Compaq Presario 100	13	1650	1.50%	21:
6	103	Assurances	6	09.02.2001	AST Intel 150	2	2690	0.00%	5:
7	104	Éducation	7	07.01.2001	AST Intel 200	2	3190	0.00%	6:

Nous allons d'abord trier la colonne **Prix total avec rabais** dans une nouvelle colonne sur une nouvelle feuille. Nous allons alors dans le menu **Données/Trier**:



Trier une ou plusieurs colonnes de données

Ce qui donne:



Trier une ou plusieurs colonnes de données

Ce qui donne:

Minitab - Ventes.MPJ

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

Session

15.11.2010 12:21:50

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Prix total avec rabais												
1	2690.0												
2	3190.0												
3	4598.0												
4	4598.0												
5	4598.0												
6	4598.0												
7	5380.0												
8	5380.0												

Feuille de travail en cours : PTAR

Démarrer Minitab - Microsoft Word Minitab - Ventes.MPJ 12:57

Ensuite, nous allons dans le menu **Graphique/Branches et feuilles** ou dans le menu **Stat/Analyse exploratoire des données/Branches et feuilles** (puisqu'les deux chemins mènent au même outil!):

Minitab - Ventes.MPJ

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

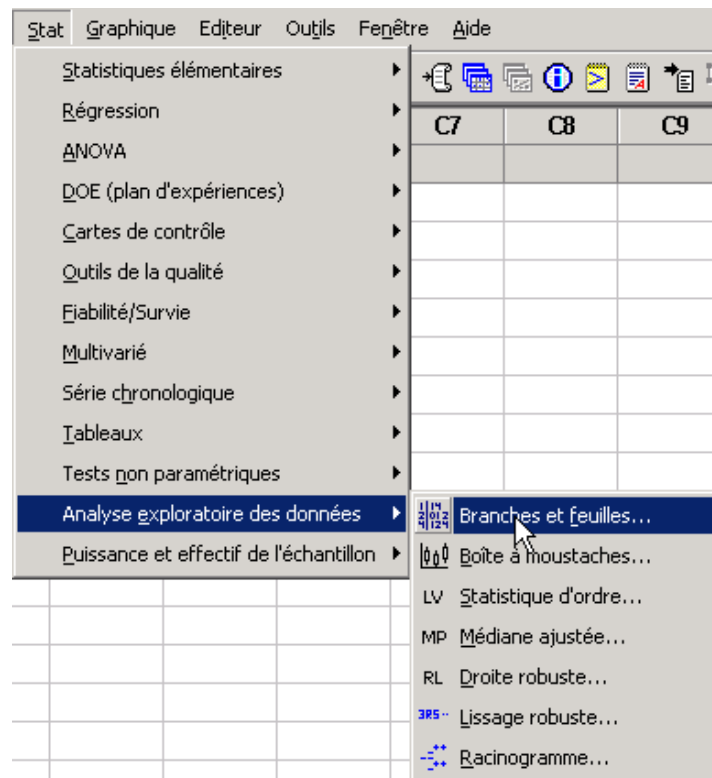
Session

15.11.2010 12:21:50

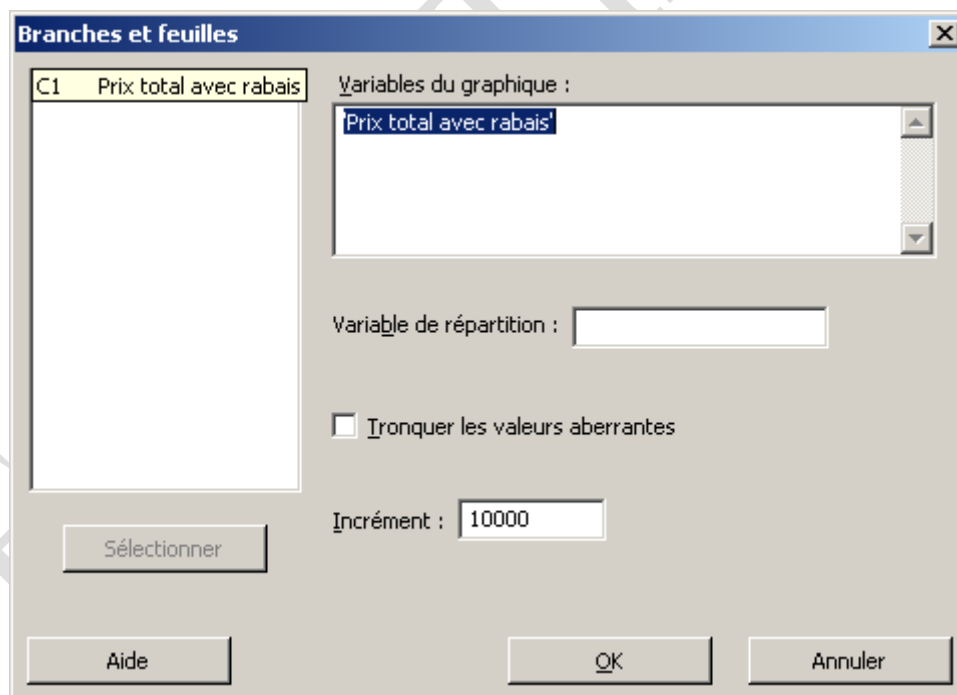
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Prix total avec rabais												
1	2690.0												
2	3190.0												
3	4598.0												
4	4598.0												
5	4598.0												
6	4598.0												
7	5380.0												
8	5380.0												

Tracer des diagrammes à branches et feuilles

Démarrer Minitab - Microsoft Word Minitab - Ventes.MPJ 12:58



Nous avons alors:



Et validons par **OK** pour obtenir dans la fenêtre de session:

The screenshot shows the Minitab interface with a session window displaying a pseudo-graphic and a data table window showing the corresponding data.

**Session Window:**

```
Branches et feuilles de Prix total avec rabais N = 109
Unité de feuille = 1000

35 0 2344445555666666666688888899999999
(39) 1 0011112222333333333355555577788888999
35 2 0111112234445555578
16 3 01113446
8 4 09
6 5 4
5 6 288
2 7 1
1 8 5
```

**PTAR \*\*\* Data Table:**

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Prix total avec rabais												
100	34563.7												
101	36811.5												
102	40848.0												
103	49104.0												
104	54795.3												
105	62623.2												
106	68074.6												
107	68074.6												

Feuille de travail en cours : PTAR

Voilà comment se lit ce pseudo-graphique:

- (35, 0, 2344455...) Il y a donc 35 valeurs comprises entre 0-10'000 et parmi ces valeurs il y en une qui est proche de 2'000 (2), une qui est proche de 3'000 (3), 4 qui sont proches de 4'000 (4444), 4 qui sont proches de 5'000 (5555) etc...

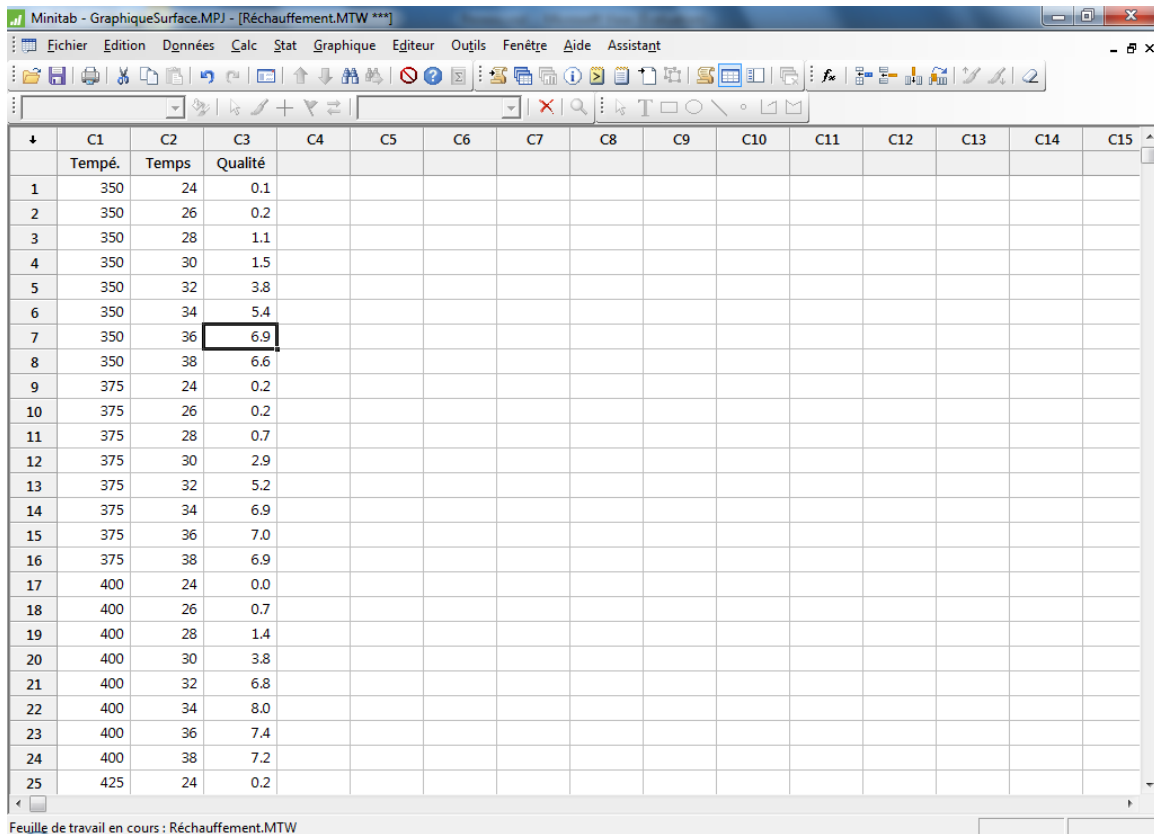
- ((39), 1, 001111222...) La parenthèse (39) signifie que la valeur médiane est entre 39'000 et 40'000. Le 39 signifie qu'il y a 39 valeurs comprises entre 10'000-20'000, etc...



## 8.19. Exercice 27.: Diagramme de surface en 3D

Minitab® Statistical Software 17.1

Le graphique de surface est très utile dans certaines situations (sinon il n'existerait pas...). Voyons donc un exemple basé sur les données du fichier *GraphiqueSurface.mpj*:



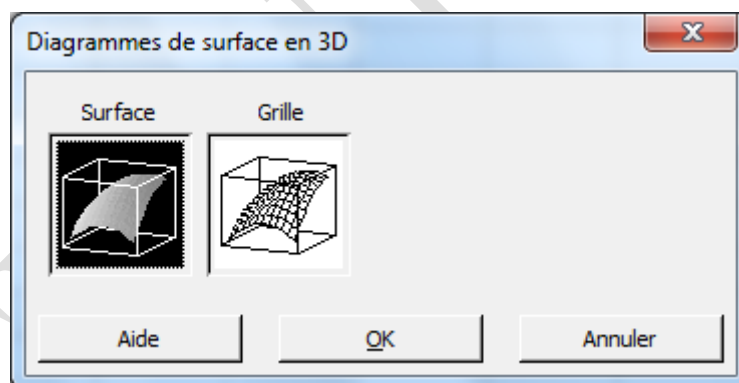
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
	Tempé.	Temps	Qualité												
1	350	24	0.1												
2	350	26	0.2												
3	350	28	1.1												
4	350	30	1.5												
5	350	32	3.8												
6	350	34	5.4												
7	350	36	6.9												
8	350	38	6.6												
9	375	24	0.2												
10	375	26	0.2												
11	375	28	0.7												
12	375	30	2.9												
13	375	32	5.2												
14	375	34	6.9												
15	375	36	7.0												
16	375	38	6.9												
17	400	24	0.0												
18	400	26	0.7												
19	400	28	1.4												
20	400	30	3.8												
21	400	32	6.8												
22	400	34	8.0												
23	400	36	7.4												
24	400	38	7.2												
25	425	24	0.2												

Nous allons ensuite dans le menu **Graphique/Diagramme de surface en 3D...**:

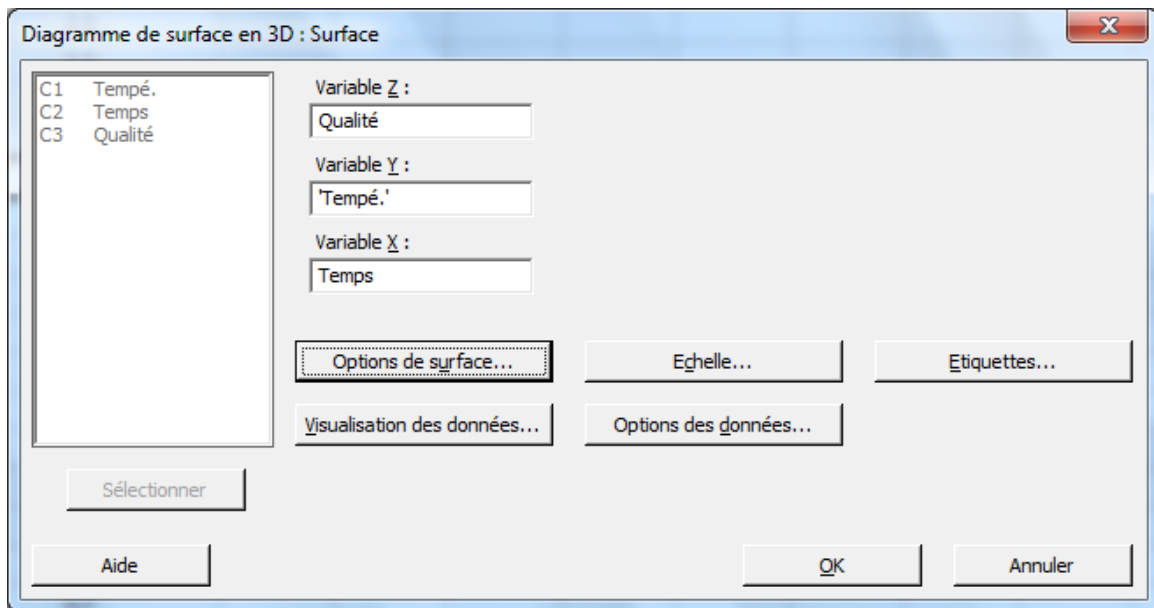
Diagramme de surface en 3D  
Examiner la relation entre une variable de réponse (Z) et deux variables de prédiction (X et Y) en visualisant une surface en 3D de la réponse prévue.

	C1	C2	C3
	Tempé.	Temps	Qualité
1	350	24	0.1
2	350	26	0.2
3	350	28	1.1
4	350	30	1.5
5	350	32	3.8
6	350	34	5.4
7	350	36	6.9
8	350	38	6.6
9	375	24	0.2
10	375	26	0.2
11	375	28	0.7
12	375	30	2.9
13	375	32	5.2
14	375	34	6.9
15	375	36	7.0
16	375	38	6.9
17	400	24	0.0
18	400	26	0.7
19	400	28	1.4
20	400	30	3.8
21	400	32	6.8
22	400	34	8.0
23	400	36	7.4
24	400	38	7.2
25	425	24	0.2

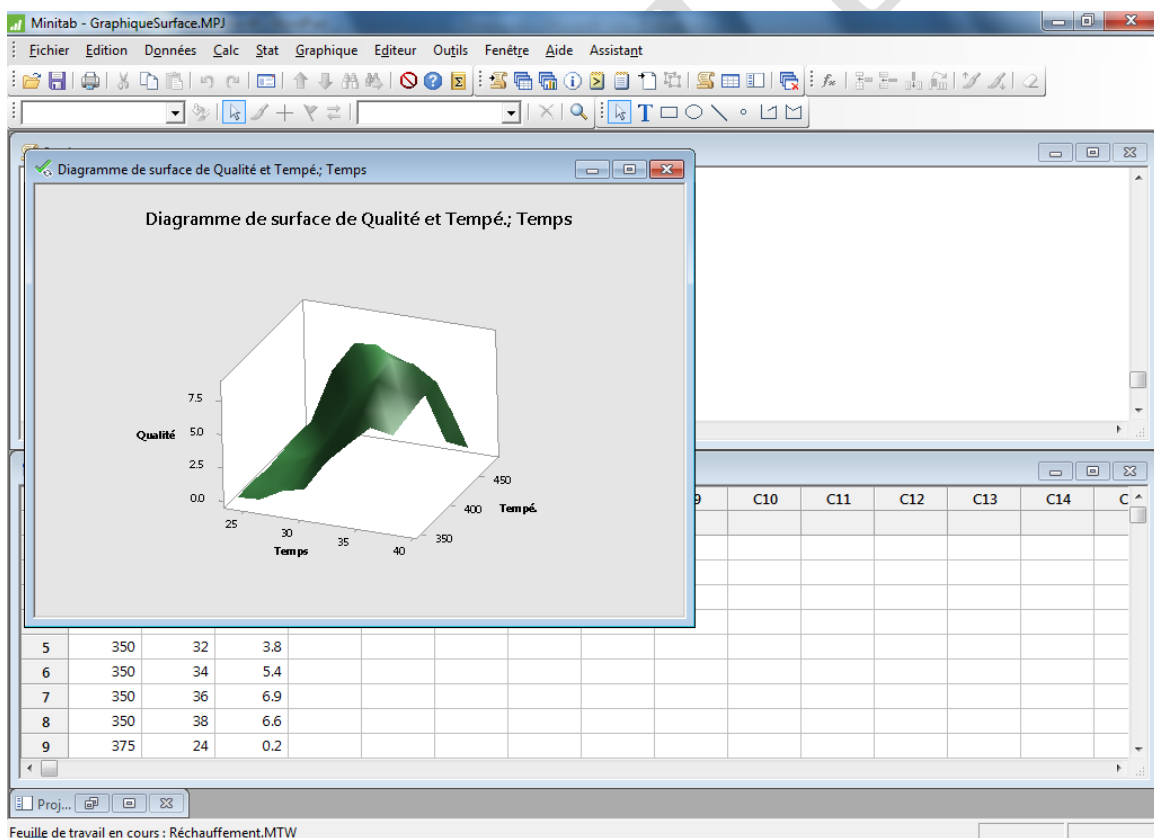
Nous avons alors à l'écran:



Nous prenons l'option **Surface** et validons par **OK**:



Que nous validons par OK pour obtenir un résultat plus que discutable en comparaison de R ou MATLAB:



Bref... juste ne pas oublier la barre d'outils masquée par défaut qui permet de manipuler les graphiques 3D...

## 8.20. Exercice 28.: Graphique de contours...

Minitab® Statistical Software 17.1

Sur les mêmes données voyons un type de graphique que nous retrouverons très souvent dans certains plans d'expériences.

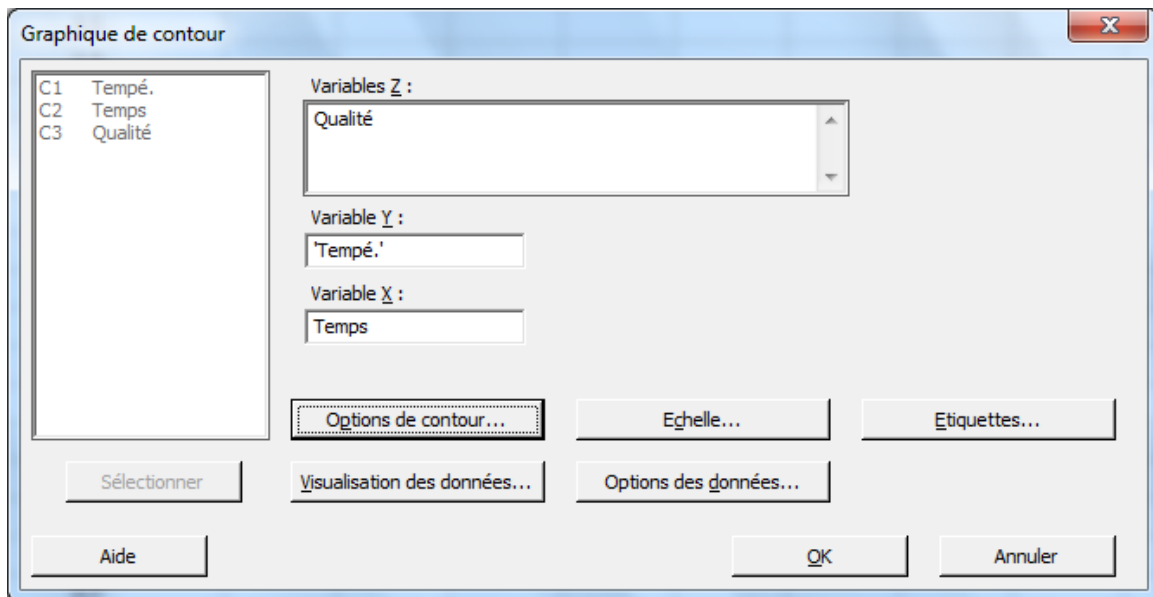
En utilisant les mêmes données qu'avant nous allons dans le menu **Graphique/Graphique de contour...**:

The screenshot shows the Minitab interface with a data table and the 'Graphique' menu open. The 'Graphique de contour...' option is highlighted. A tooltip provides the following description:

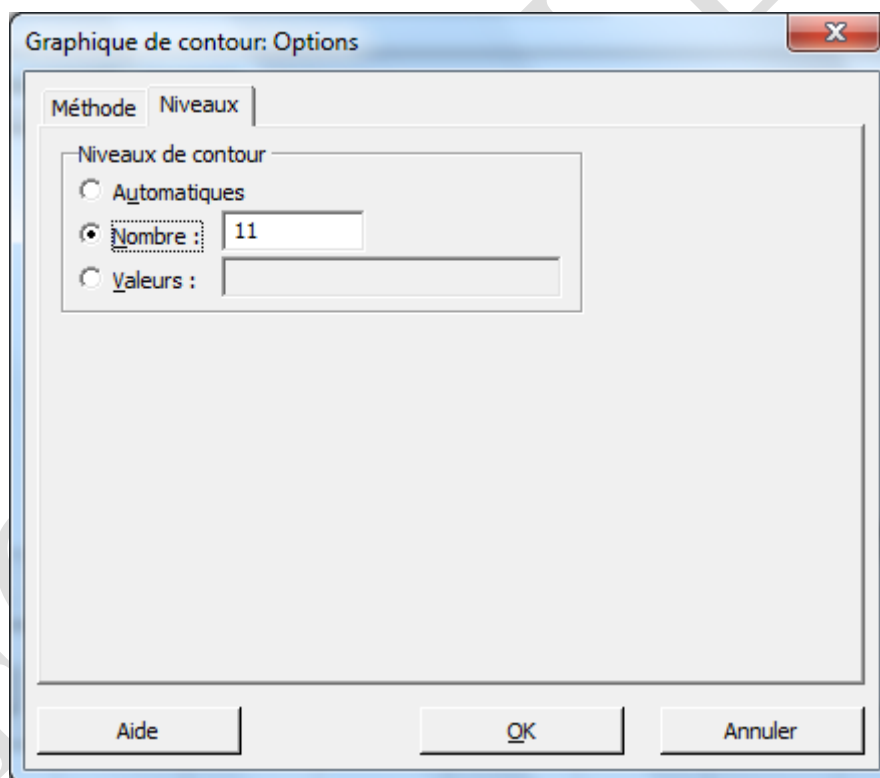
**Graphique de contour**  
Examiner la relation entre une variable de réponse (Z) et deux variables de prédiction (X et Y) en visualisant les contours discrets de la variable de réponse prévue.

	C1	C2	C3
	Tempé.	Temps	Qualité
4	350	30	1.5
5	350	32	3.8
6	350	34	5.4
7	350	36	6.9
8	350	38	6.6
9	375	24	0.2
10	375	26	0.2
11	375	28	0.7
12	375	30	2.9
13	375	32	5.2
14	375	34	6.9
15	375	36	7.0
16	375	38	6.9
17	400	24	0.0
18	400	26	0.7
19	400	28	1.4
20	400	30	3.8
21	400	32	6.8
22	400	34	8.0
23	400	36	7.4
24	400	38	7.2
25	425	24	0.2
26	425	26	0.2
27	425	28	1.8
28	425	30	4.8

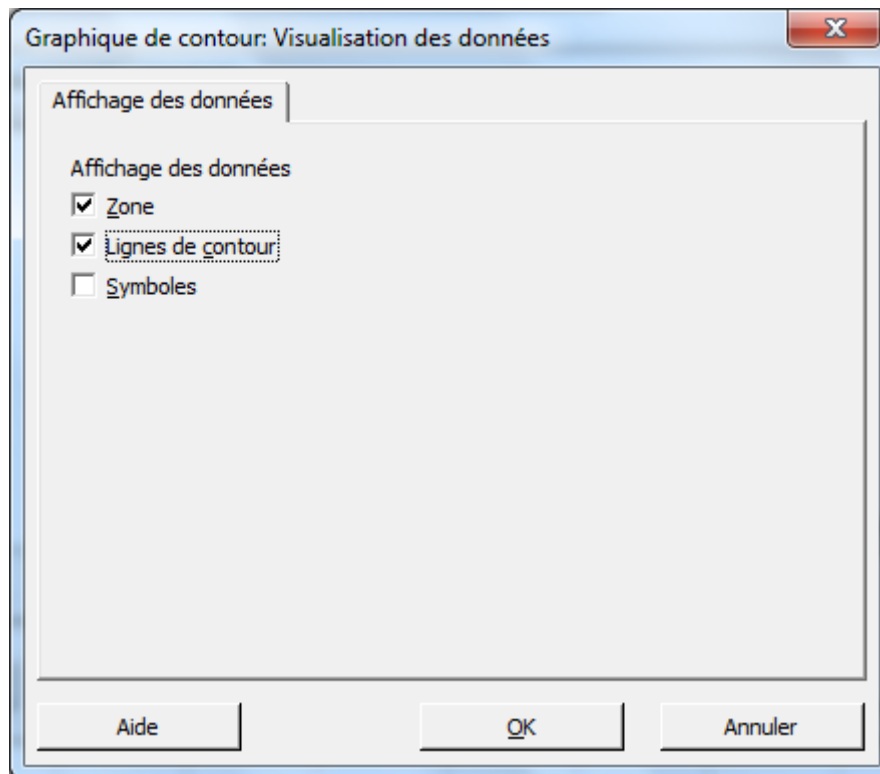
Nous avons alors la boîte de dialogue suivante:



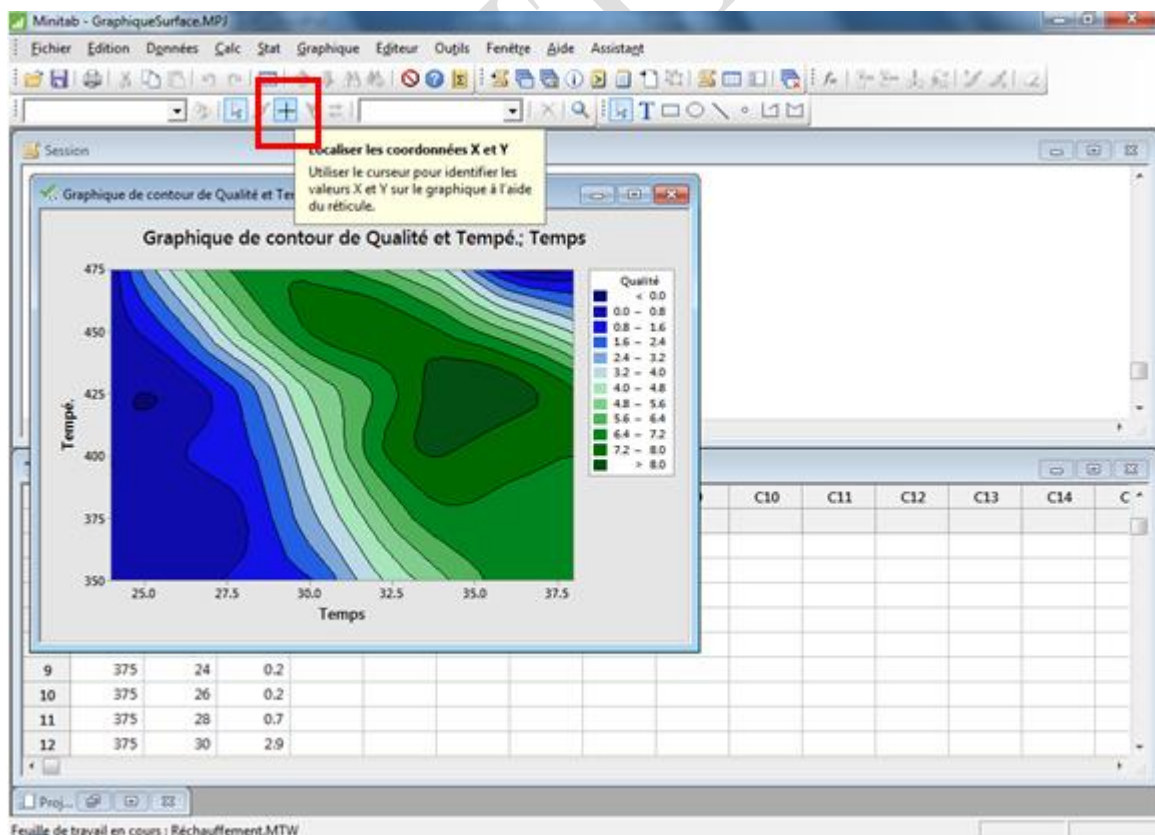
Ensuite nous cliquons sur **Options de contour...** pour améliorer le visuel par défaut en mettant le nombre de niveaux de contour au maximum (c'est-à-dire 11):



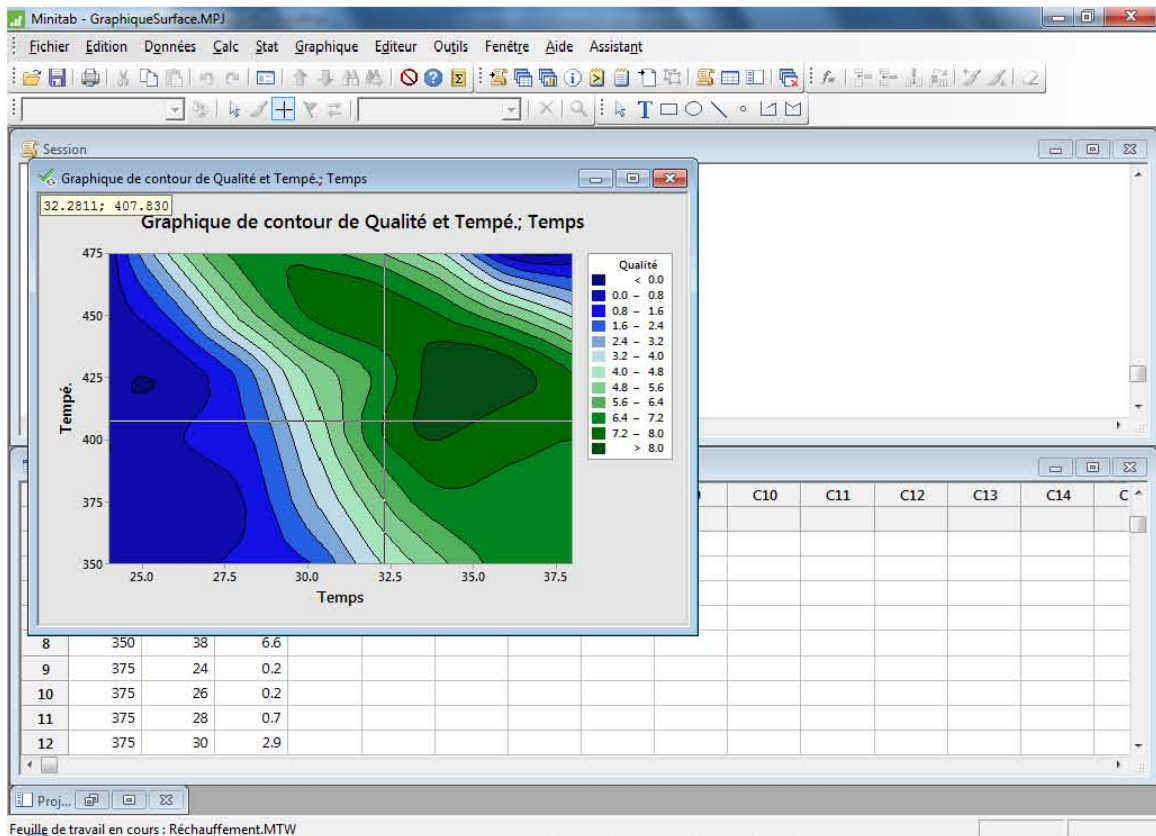
Et dans le bouton **Visualisation des données...**, pour les mêmes raisons esthétiques, nous activons **Lignes de contour**:



Nous validons le tout par **OK** deux fois pour obtenir un graphique dont il est important pour la suite de remarquer qu'un bouton **Localiser les coordonnées X et Y** est disponible!:



Si vous l'activez, vous pouvez alors obtenir dans une petite bulle sur le graphique la position de votre souris dans l'espace du graphique en temps réel:

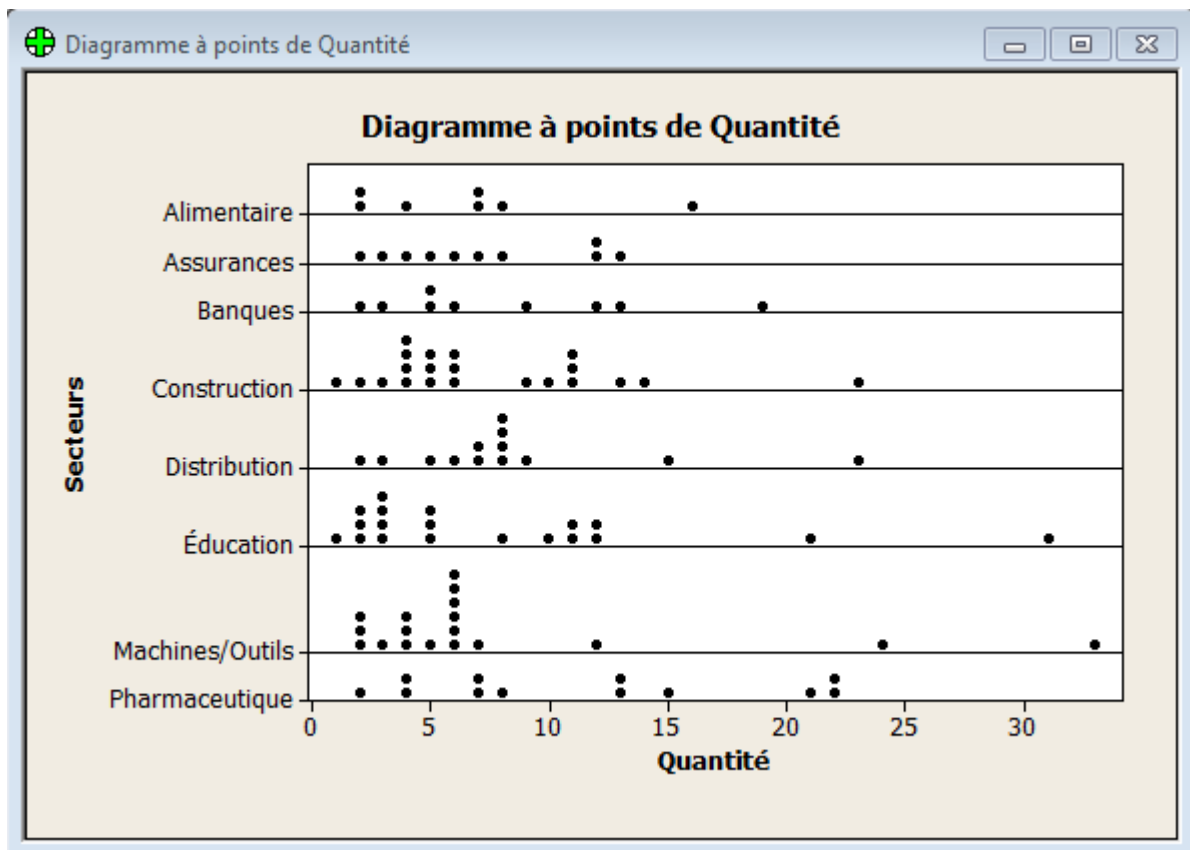


## 8.21. Exercice 29.: Mode de balayage...

Minitab® Statistical Software 15.1.1.0

Voyons maintenant l'utilité du petit outil de balayage de Minitab.

Considérons pour cet exemple, le graphique suivant que nous avons fait plus haut:

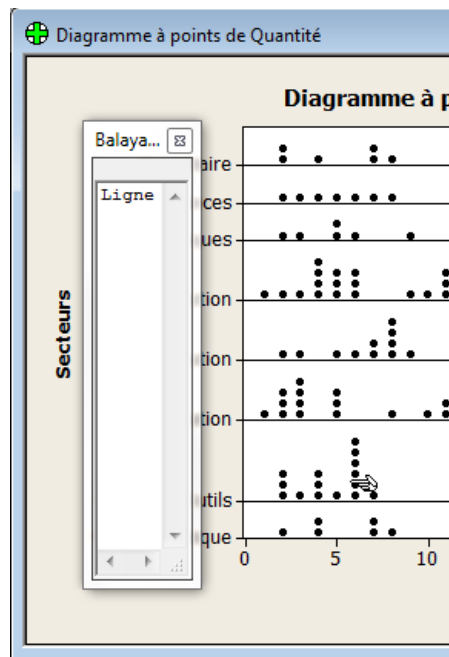


Maintenant, dans la barre d'outils de Modification de graphiques, cliquez sur l'outil **Mode de balayage**:

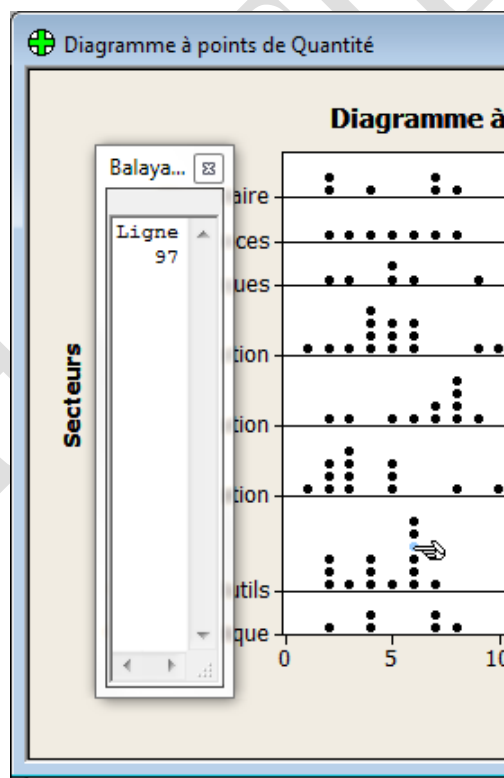


Vous verrez une petite fenêtre apparaître et le curseur de la souris changé si vous passez sur un point du graphique:

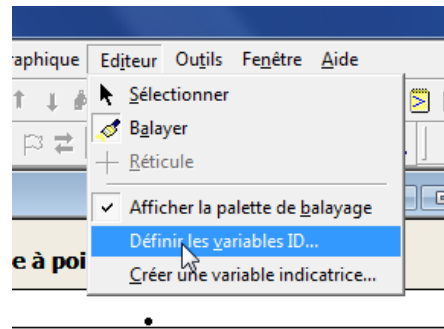




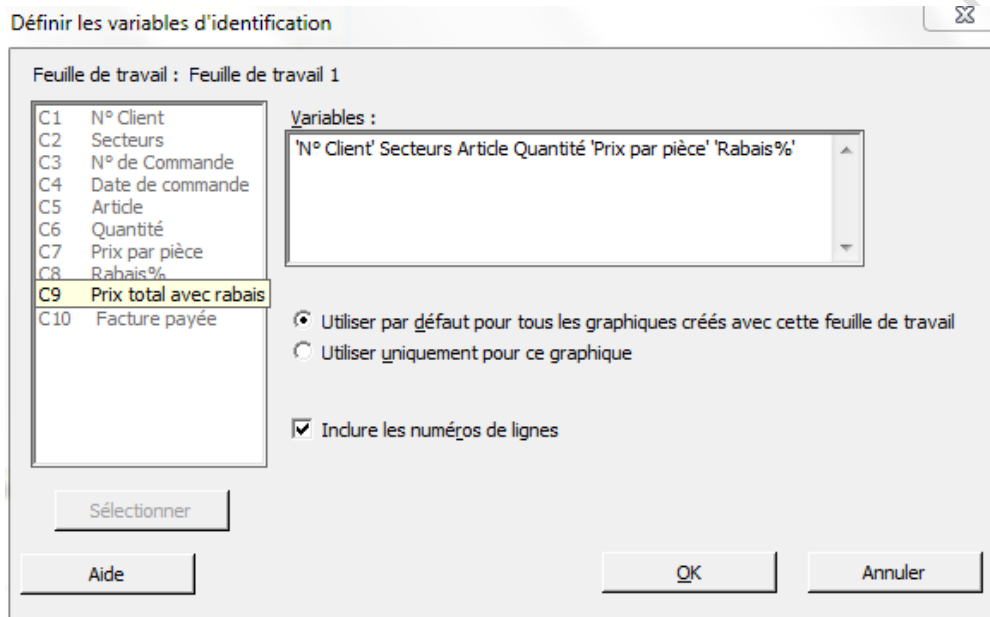
Si on clique sur un point du graphique, on obtient bêtement le numéro de ligne du point (ce qui n'est évidemment pas plus intéressant que la bulle que l'on obtient en survolant le point lui-même!!!!):



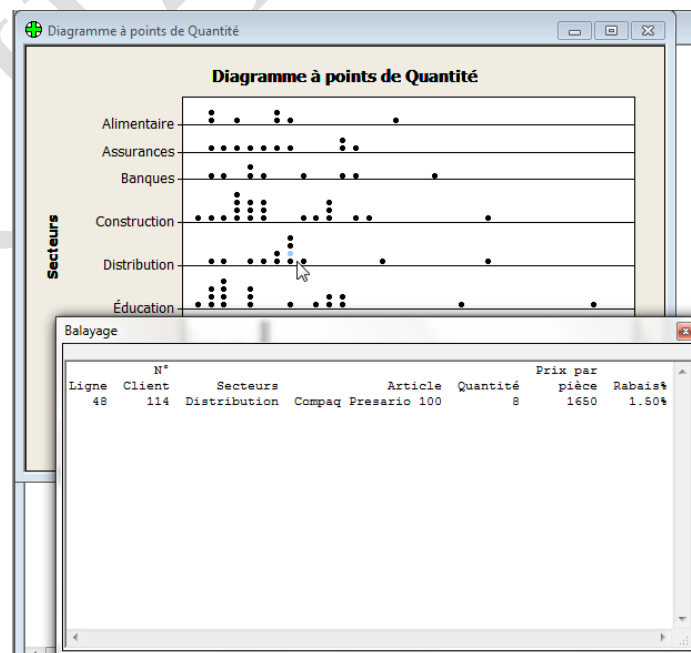
Mais on peut faire mieux en réalité! Effectivement, en allant dans le **menu Définir les variables ID...**:



Et en mettant par exemple:



On obtient après avoir validé par **OK** et en cliquant sur un point du graphique:



Ce qui est nettement plus intéressant!

## 9. Probabilités

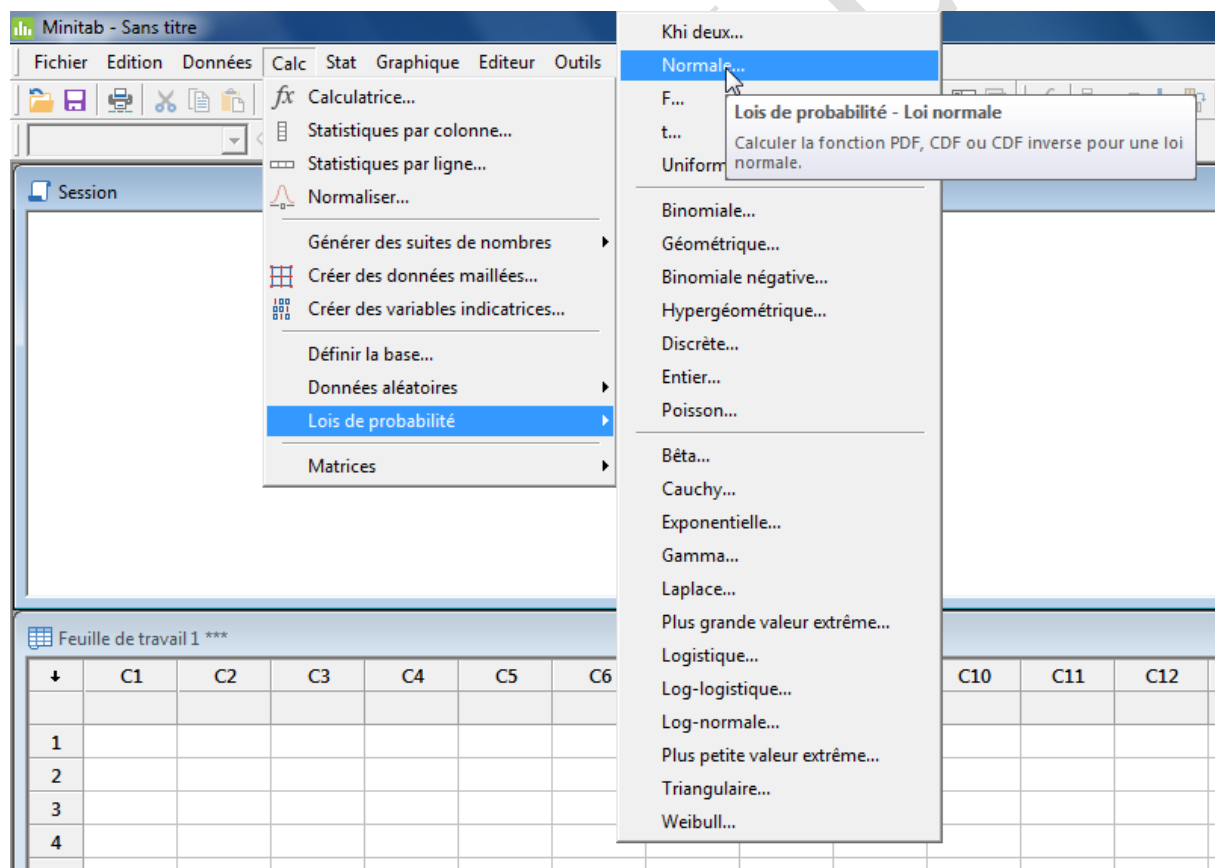
J'ai créé ce chapitre en 2018 suite à des discussions avec certains clients et participants à mes formations. Je n'avais pas pensé que ce serait un jour nécessaire, mais certains doutes seront probablement levés pour ceux qui se posent certaines questions que je préciserai pour chacun des exemples ci-dessous.

### 9.1. Exercice 30.: Calculer une probabilité cumulée

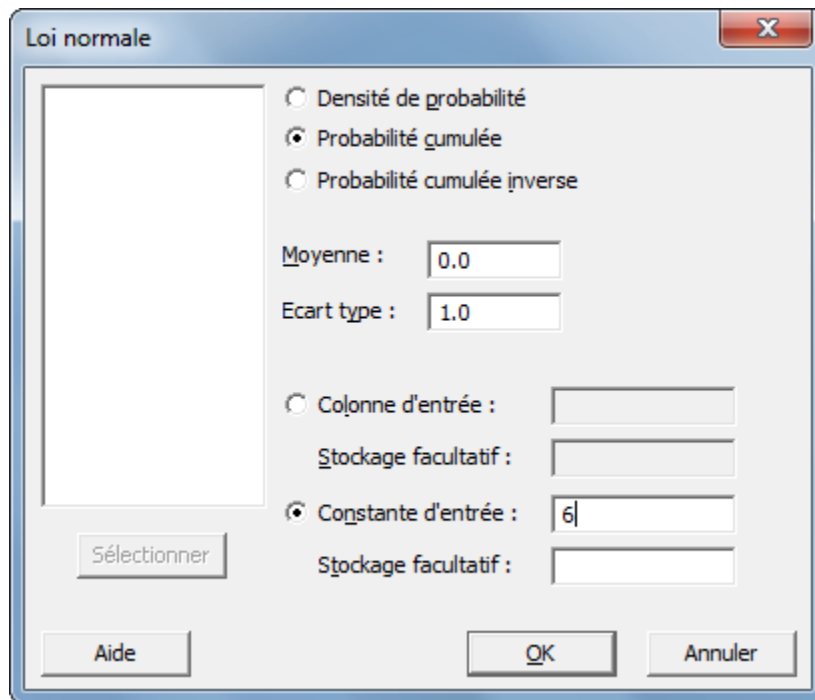
Minitab® Statistical Software 18.1

Commençons par le calcul de probabilité cumulée d'une distribution potentiellement quelconque mais surtout de la loi Normale centrée réduite afin de vérifier si nous pouvons retomber sur les valeurs "canoniques" de la méthodologie Six Sigma.

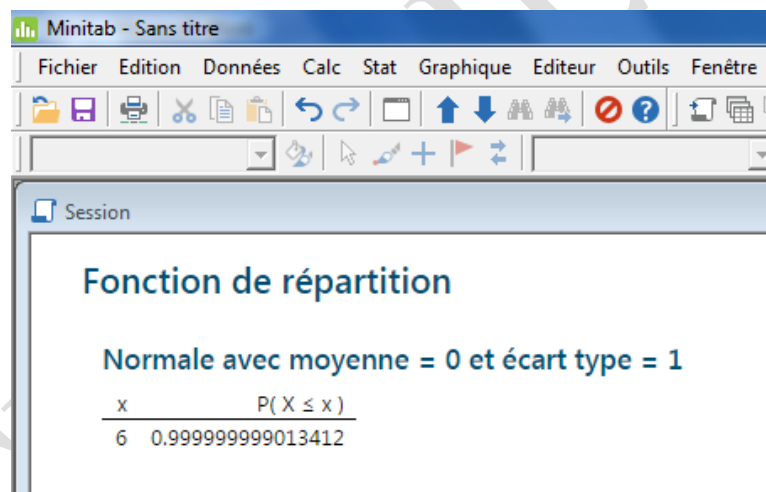
Pour cela, nous allons dans le menu **Calc/Lois de probabilité/Normale....**:



Nous laissons les valeurs par défaut pour la moyenne et l'écart-type (correspondants bien à une loi normale centrée réduite) mais nous mettons 6 pour la constante d'entrée car pour une loi normale centrée réduite, selon correspond donc à 6 Sigma (en unilatéral):



Nous obtenons alors:



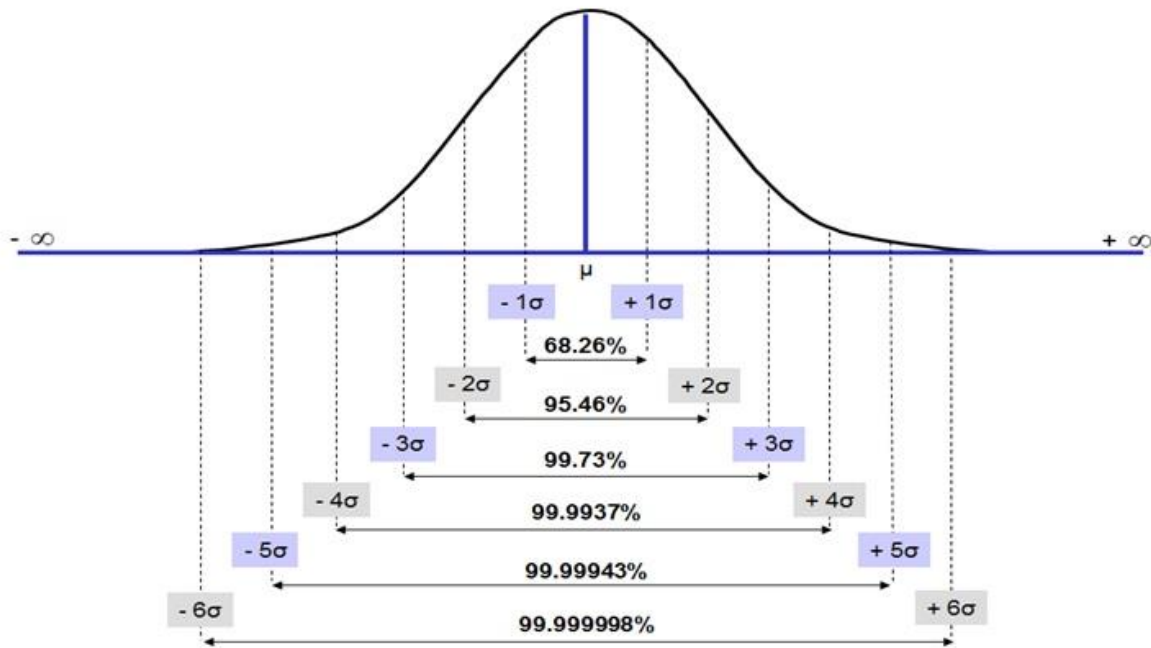
Soit une probabilité cumulée à partir de la moyenne de:

$$0.99999999013412 - 0.5 = 49.9999990134\%$$

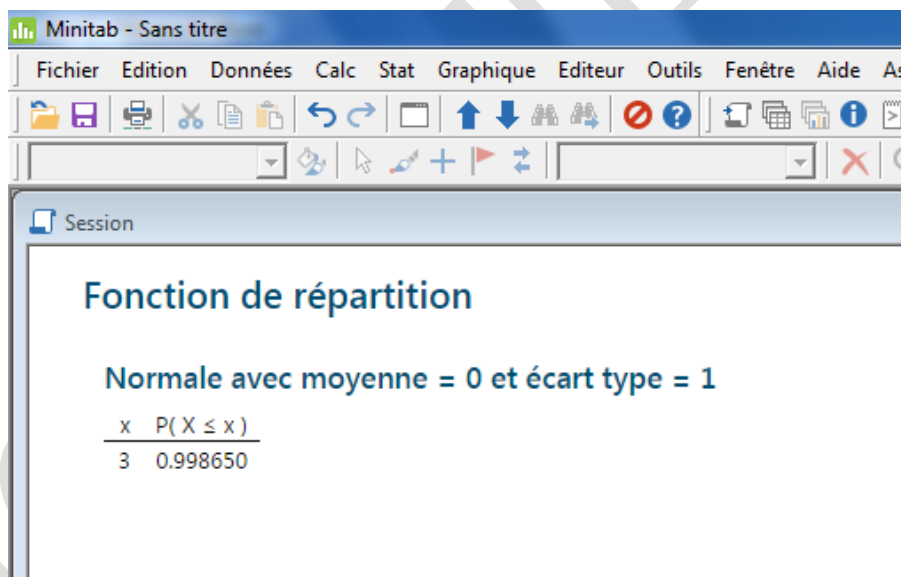
Donc en bilatéral cela correspond à:

$$2 * 49.9999990134\% = 99.9999980268\%$$

Soit les fameux 0.002 parts par million (mais on a vu dans le cours théorique que la méthode 6 Six Sigma tolère 1.5 Sigma de décalage de la moyenne par rapport à la cible de 0 ce qui donne réalité 3.4 ppm) !



Et on peut s'amuser à faire le même calcul avec aussi la fameux 3 sigma:



## 9.2. Exercice 31.: Visualiser la probabilité cumulée (unilatéral ou bilatéral)

Minitab® Statistical Software 17.1

Il arrive souvent que l'on calcule les paramètres lois de probabilités afin de prendre des décisions relativement aux données de la fonction de répartition.

Voyons comment procéder par exemple pour une loi Normale d'espérance 25 et d'écart-type de 3.

Nous allons dans le menu **Graphique/Diagramme de loi de probabilité...**:

The screenshot shows the Minitab Statistical Software 17.1 interface. The 'Graphique' menu is open, and the option 'Diagramme de loi de probabilité...' is selected. The session window displays the following information:

Normale avec moyenne = 25 et  $\sigma = 3$

$$P(X \leq x) = \Phi\left(\frac{x - 25}{3}\right)$$

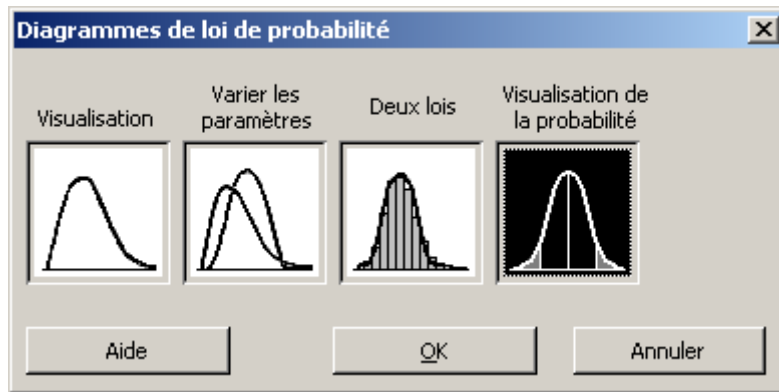
x	0.66	26.2374
---	------	---------

The worksheet window shows the following data:

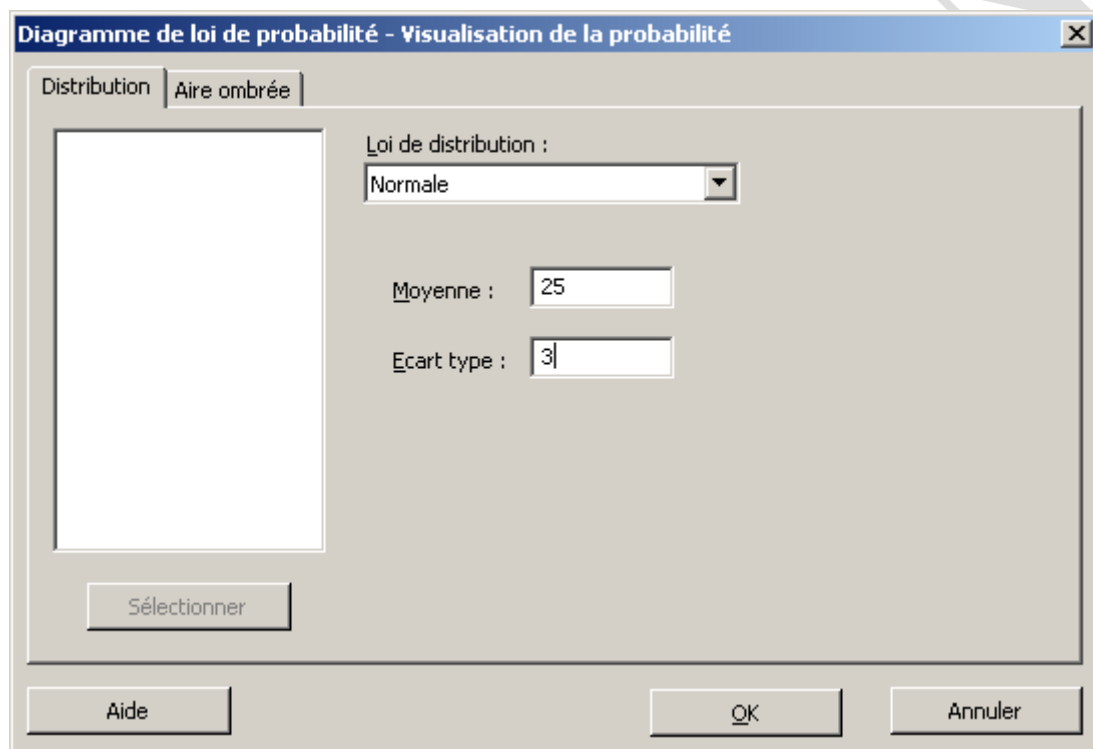
	C1	C2	C3
1	0.66		
2			
3			
4			
5			
6			
7			

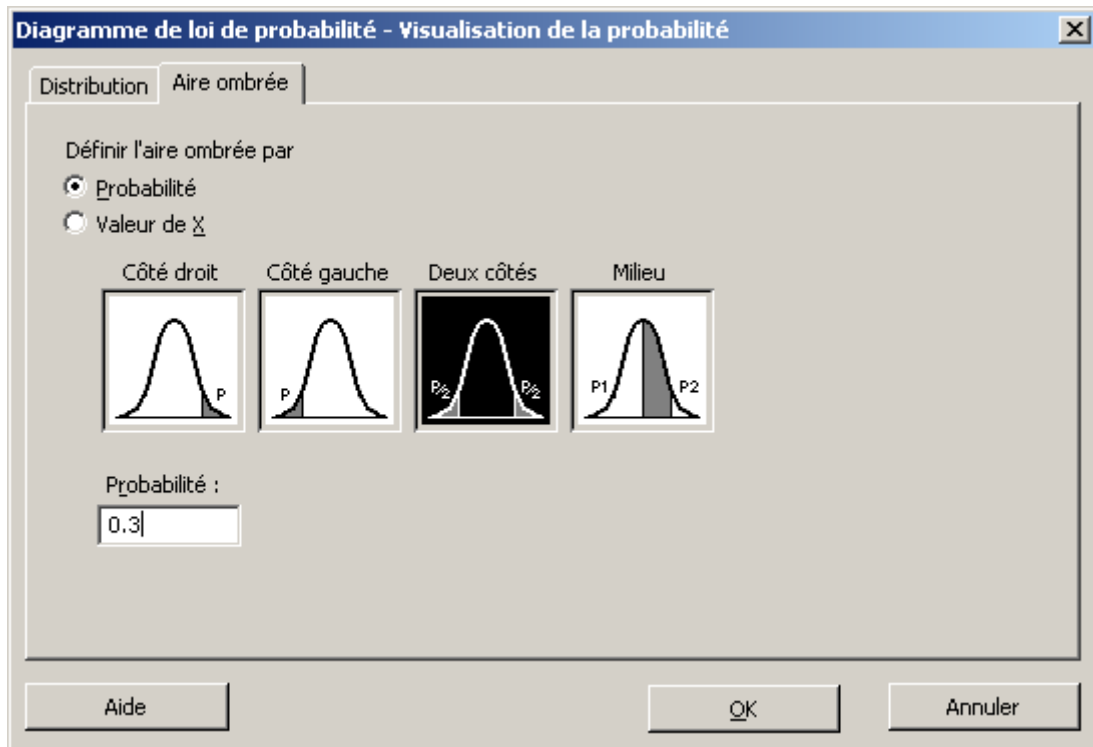
The taskbar at the bottom shows the following applications: Démarrer, Minitab - Microsoft Word, and Minitab - BoxPlot.MPJ. The system clock shows 15:18.

Nous prenons une option qui est longue à faire avec Microsoft Excel (à titre de comparaison c'est plus long dans Excel mais au moins on peut en superposer autant que l'on veut dans Excel ce qui est très utile dans la pratique et qui manque à Minitab):

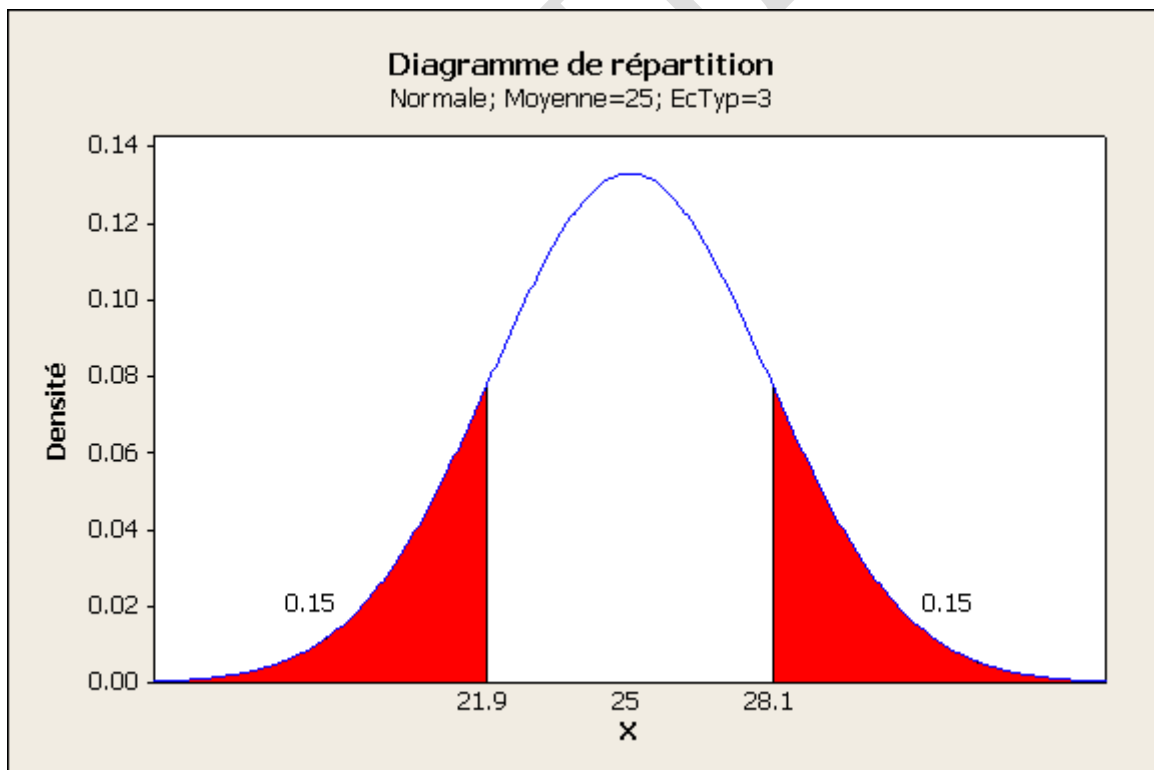


Après avoir choisi **Visualisation de la probabilité**, nous validons par **OK**:





Et nous validons deux fois par **OK** pour avoir au final:





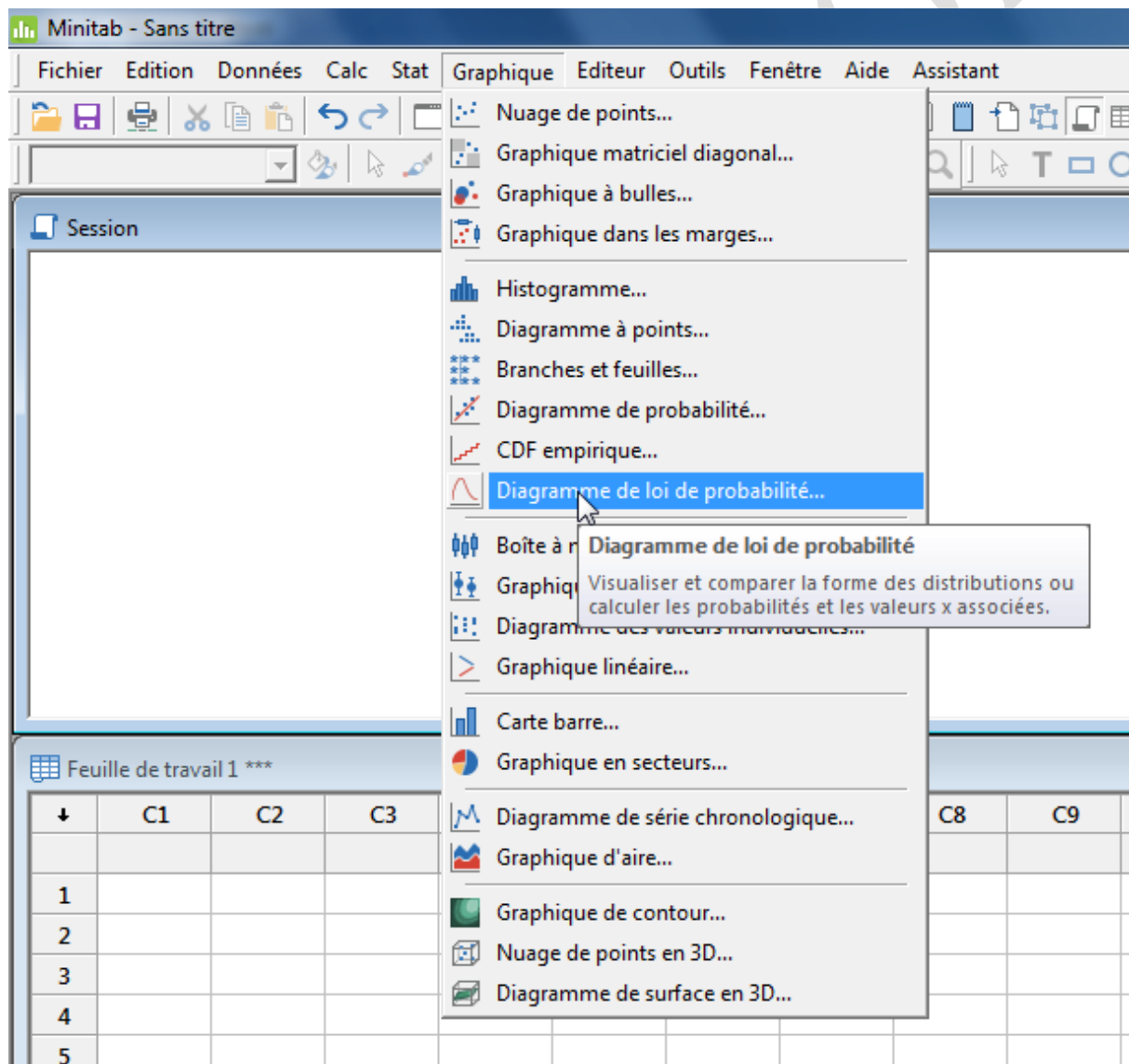
### 9.3. Exercice 32.: Convergence de la loi Binomiale en loi Normale

Minitab® Statistical Software 18.1

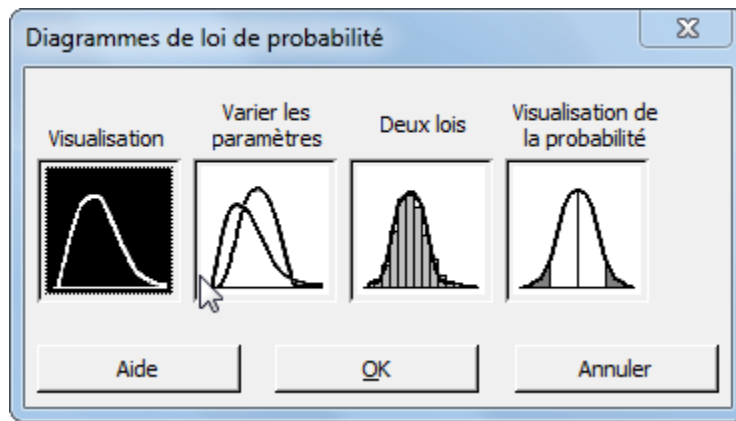
Plusieurs participants à mes formations m'ont demandé à partir de combien de tirages aléatoires la loi binomiale avec des événements d'une certaine probabilité fixée d'avance tend vers une loi Normale.

Nous allons voir ici une approche qualitative (visuelle) pour confirmer la convergence (mais on verra plus tard qu'en réalité il faudrait faire un test de Normalité statistique).

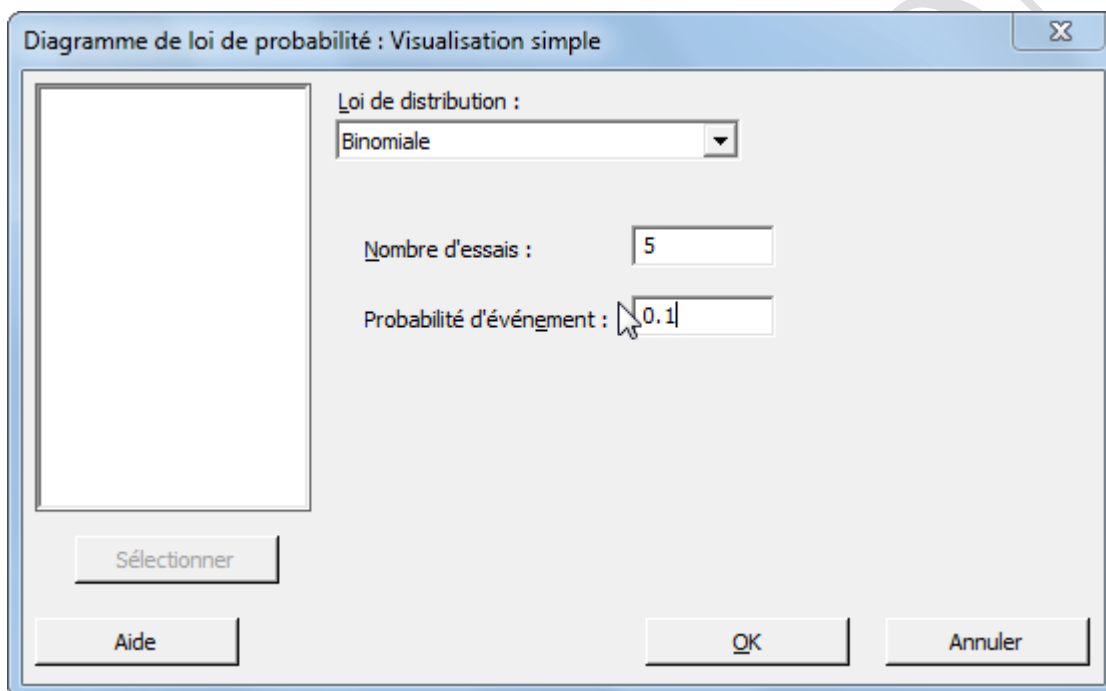
Donc pour générer une variable aléatoire suivant une loi Normale, on va dans le menu **Graphique/Diagramme de loi de probabilité...**:



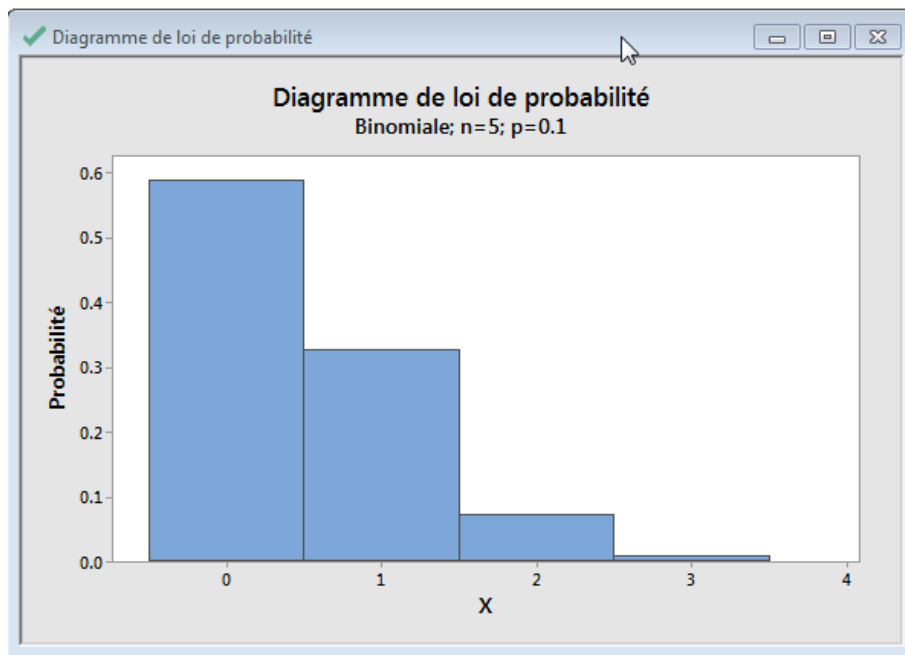
Ensuite on prend la première visualisation:



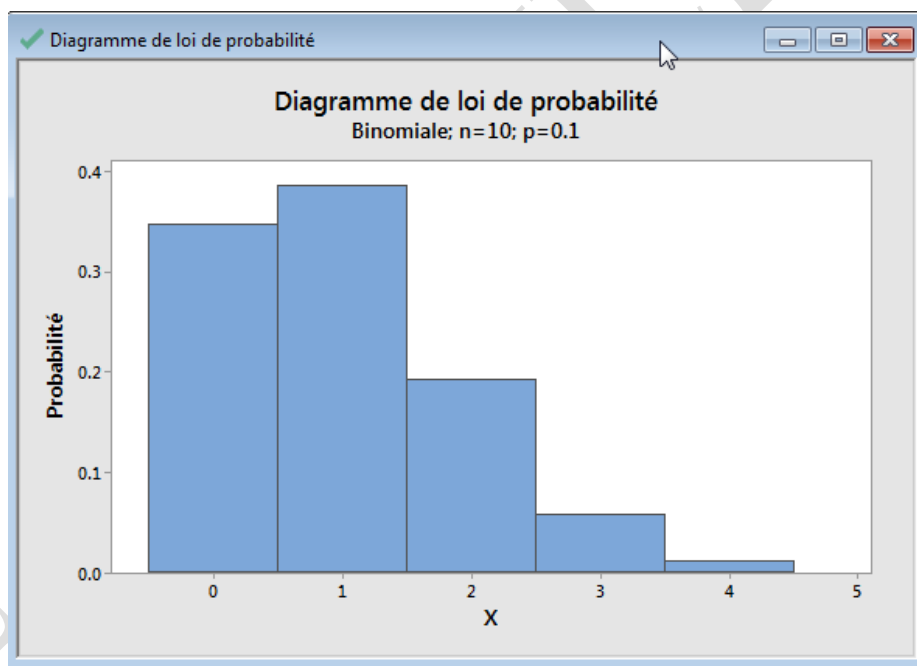
Et on mettra un cas assez typique dans une industrie non exigeante qui sera:



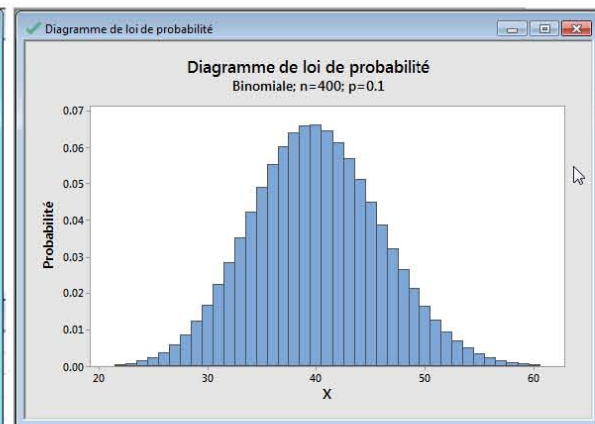
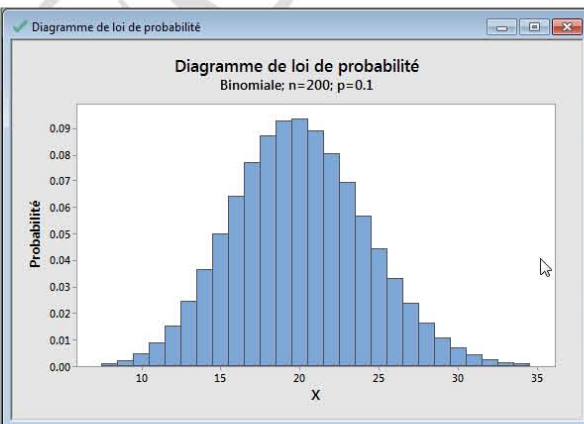
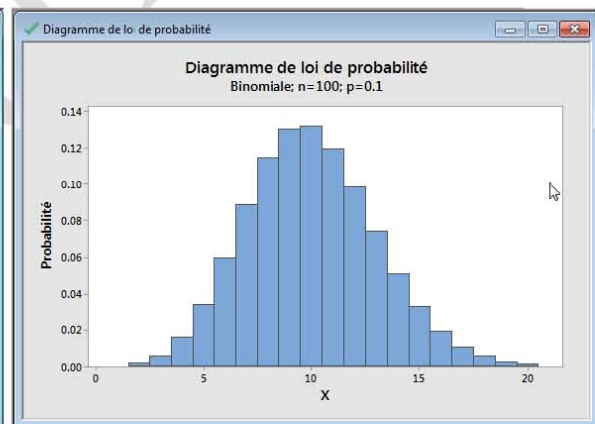
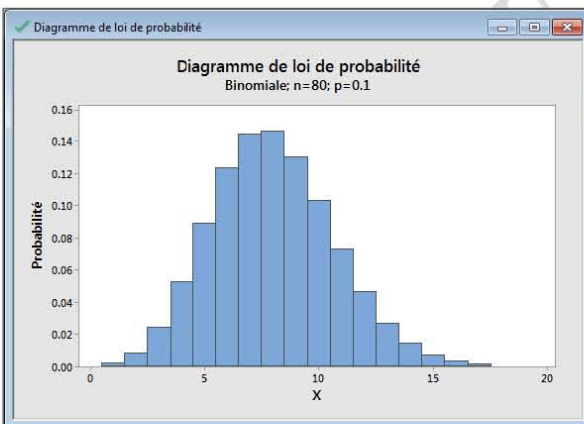
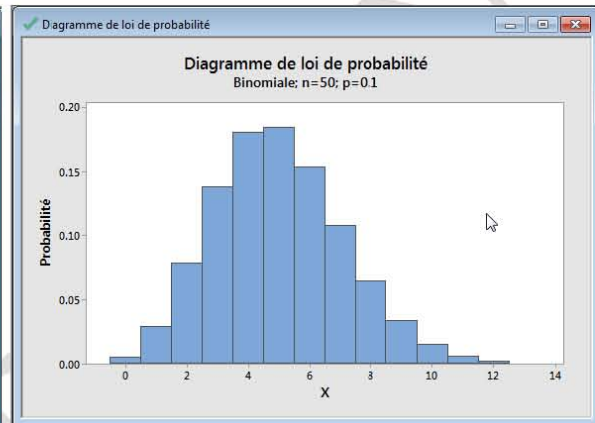
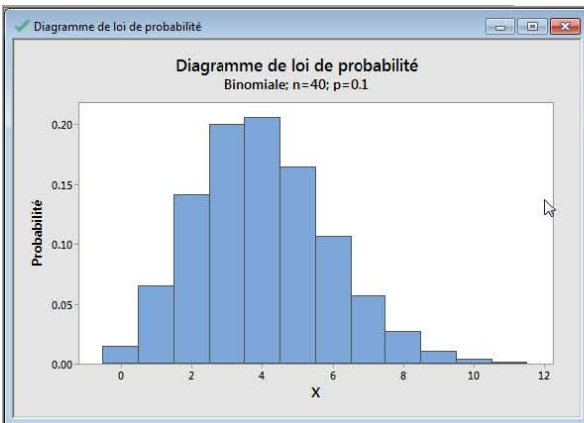
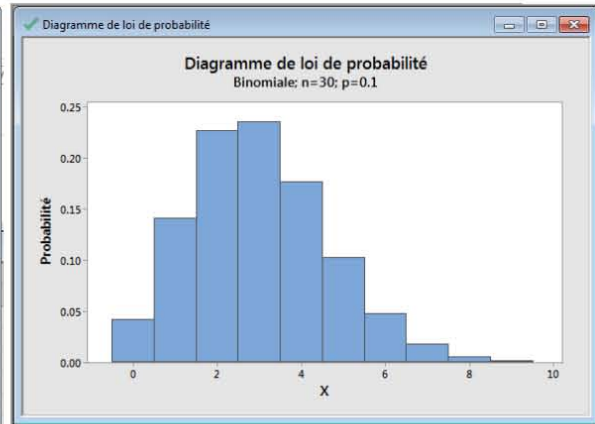
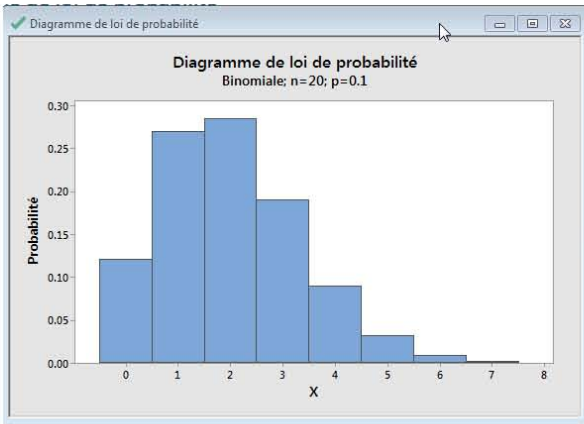
On valide par **OK** pour obtenir:

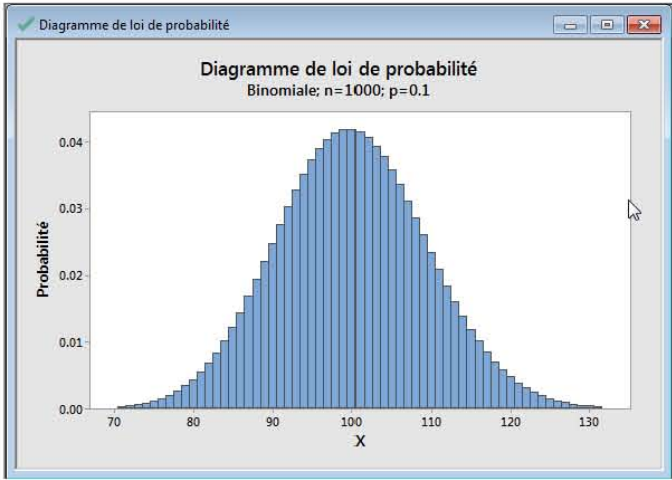


Il est clair que cela ne suite pas une loi Normal a priori. Essayons avec 10 tirages:



C'est mieux mais toujours pas ça... Poursuivons avec des valeurs de plus en plus grandes de tirages:





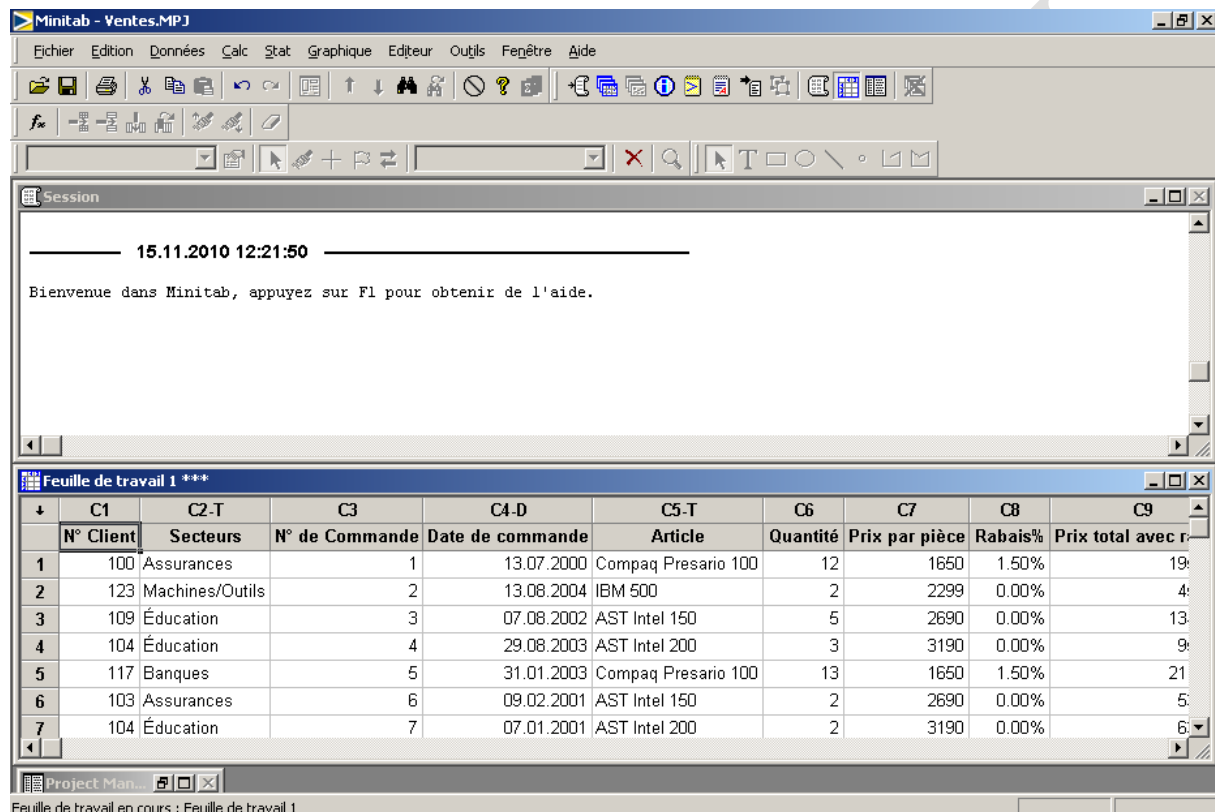
ÉCHANTILLON

# 10. Statistiques paramétriques

## 10.1. Exercice 33.: Statistiques descriptives

Minitab® Statistical Software 15.1.1.0

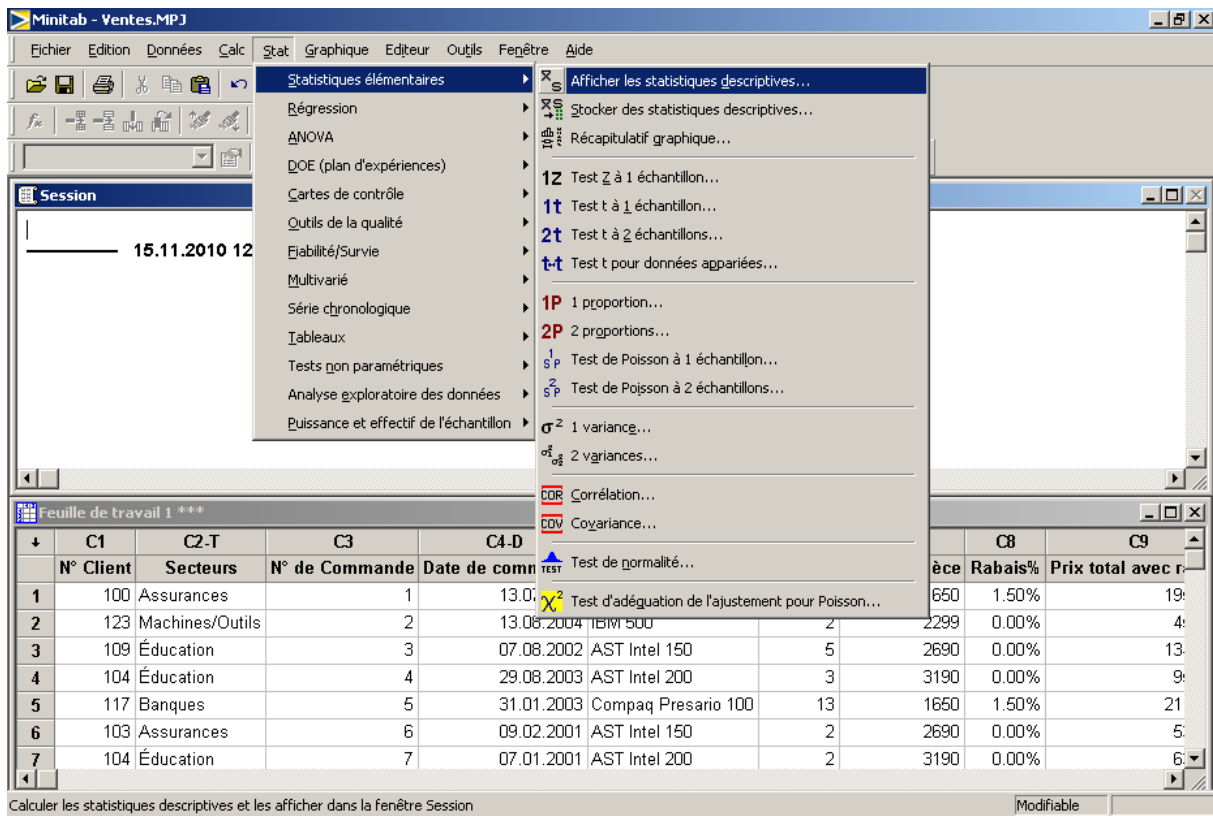
Ouvrez le fichier projet *Ventes.mpj*:



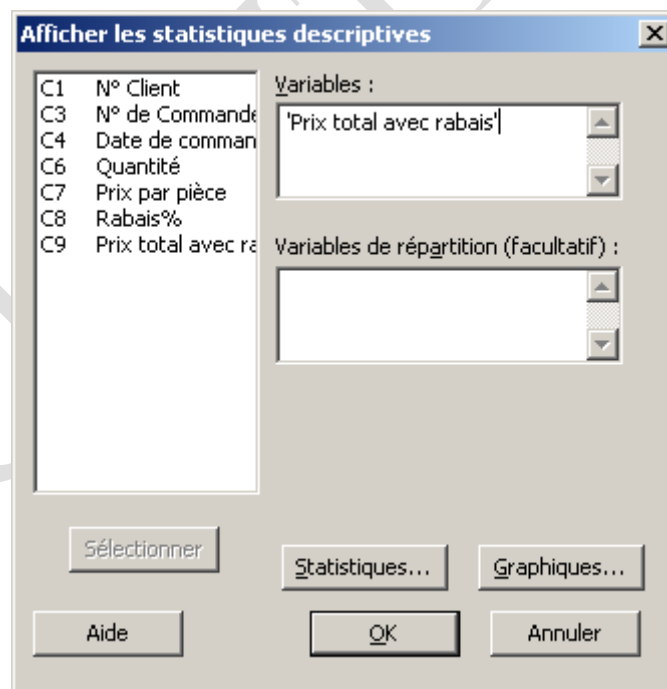
The screenshot shows the Minitab Statistical Software 15.1.1.0 interface. The main window is titled "Minitab - Ventes.MPJ". The menu bar includes "Fichier", "Edition", "Données", "Calc", "Stat", "Graphique", "Editeur", "Outils", "Fenêtre", and "Aide". The toolbar contains various icons for file operations, editing, and analysis. The "Session" window displays the date and time "15.11.2010 12:21:50" and a welcome message: "Bienvenue dans Minitab, appuyez sur F1 pour obtenir de l'aide." The "Feuille de travail 1 \*\*\*" window shows a worksheet with the following data:

	C1	C2-T	C3	C4-D	C5-T	C6	C7	C8	C9
	N° Client	Secteurs	N° de Commande	Date de commande	Article	Quantité	Prix par pièce	Rabais%	Prix total avec r
1	100	Assurances	1	13.07.2000	Compaq Presario 100	12	1650	1.50%	19:
2	123	Machines/Outils	2	13.08.2004	IBM 500	2	2299	0.00%	4:
3	109	Éducation	3	07.08.2002	AST Intel 150	5	2690	0.00%	13:
4	104	Éducation	4	29.08.2003	AST Intel 200	3	3190	0.00%	9:
5	117	Banques	5	31.01.2003	Compaq Presario 100	13	1650	1.50%	21:
6	103	Assurances	6	09.02.2001	AST Intel 150	2	2690	0.00%	5:
7	104	Éducation	7	07.01.2001	AST Intel 200	2	3190	0.00%	6:

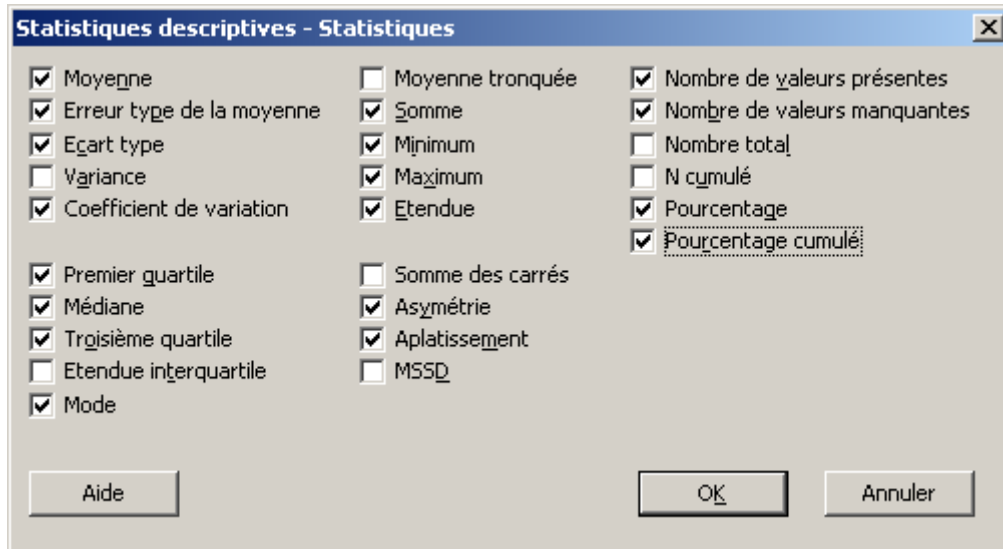
Au même titre que l'utilitaire d'analyse de Microsoft Excel, nous souhaiterions afficher des statistiques descriptives élémentaires relativement aux données de la colonne *Prix total avec rabais*. Pour cela, nous allons dans le menu **Stat/Statistiques élémentaires/Afficher des statistiques descriptives...**:



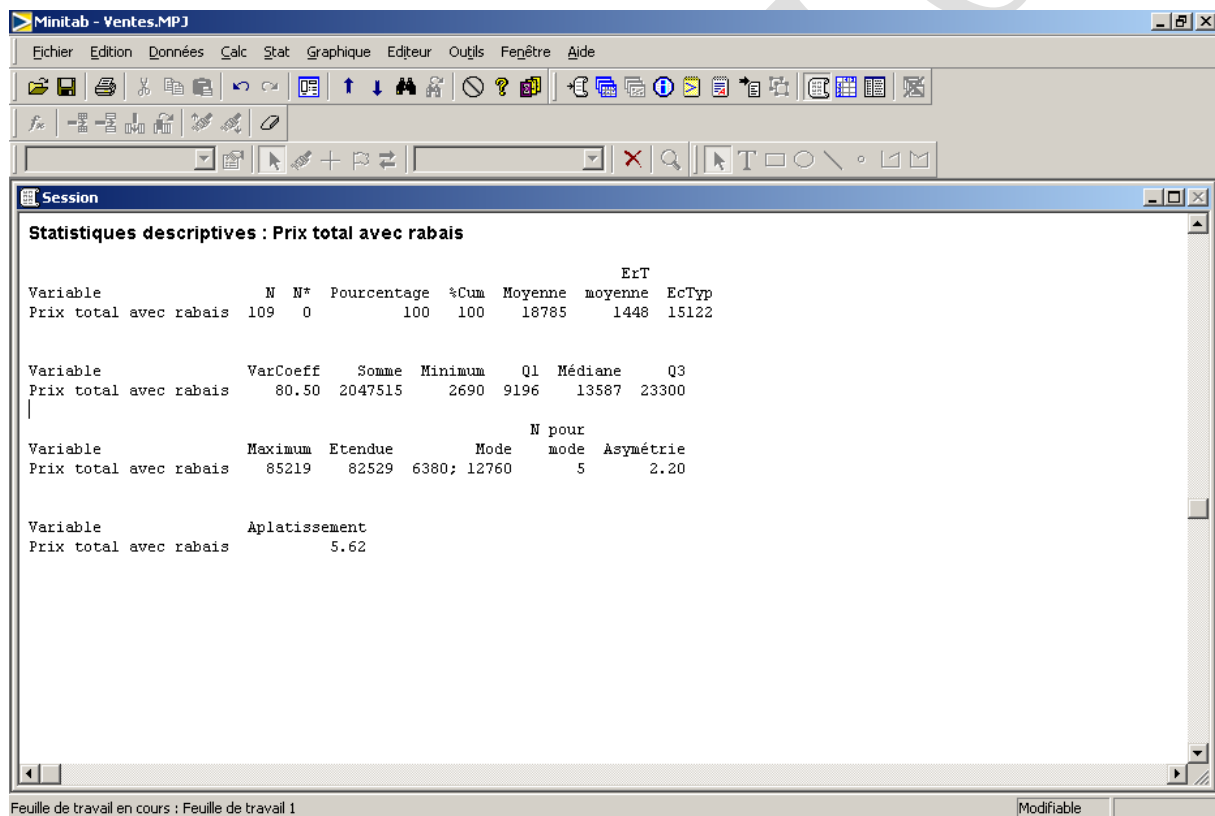
Nous avons alors:



Nous cliquons sur le bouton **Statistiques** pour y prendre:

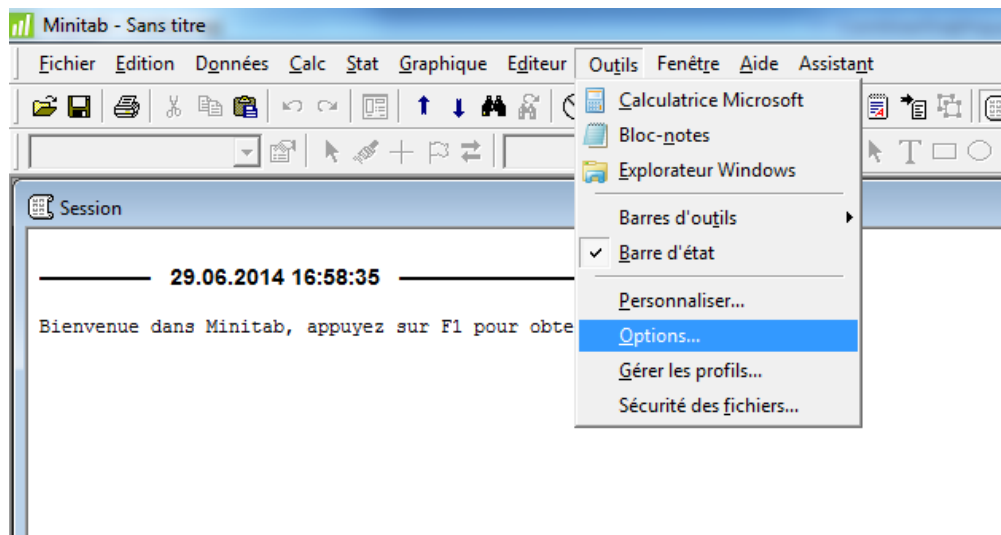


Et nous validons deux fois par **OK** pour obtenir dans la fenêtre de session:

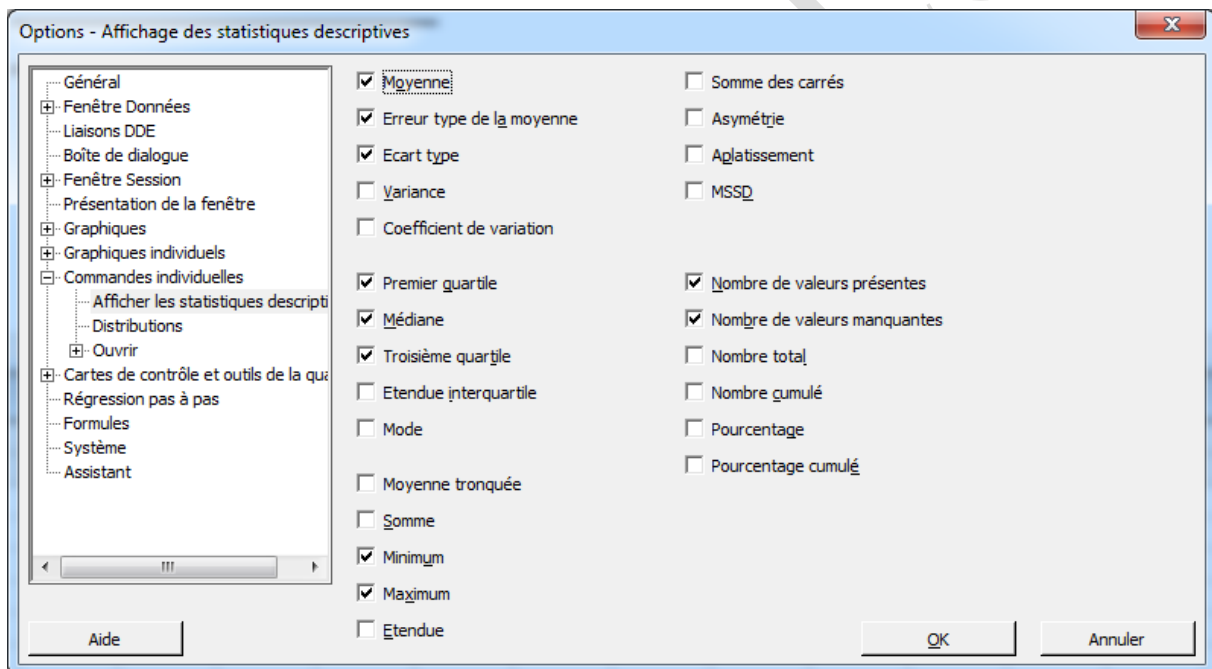


Si nous voulons changer les statistiques descriptives actives par défaut alors cela peut se faire dans les options du logiciel en allant dans **Outils/Options...**:





Et ensuite, nous allons dans **Commandes individuelles/Afficher les statistiques descriptives**:

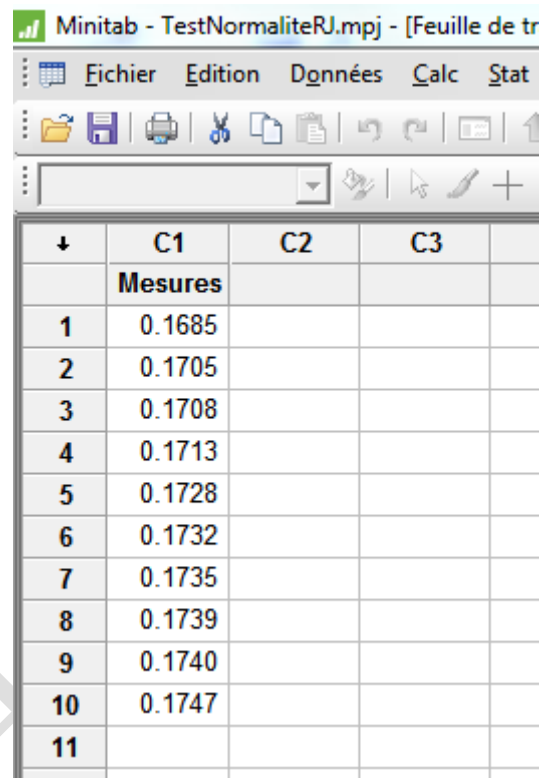


## 10.2. Exercice 34.: Kurtosis et Skewness

Minitab® Statistical Software 17.1.

Comme il y a un grand nombre de méthodes pour calculer le Kurtosis (aplatissement) et le Skewness<sup>1</sup> (asymétrie), je tiens toujours à les traiter à part.

Considérons les données suivantes:

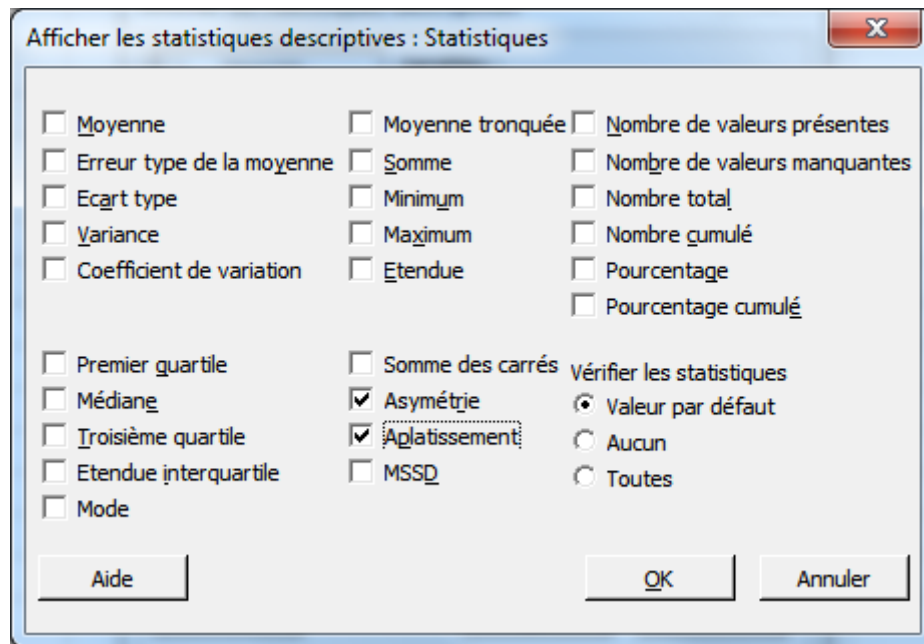


The screenshot shows the Minitab interface with a data table. The table has four columns: C1, C2, and C3. The first column, C1, is labeled 'Mesures' and contains 11 rows of data. The values in C1 are: 0.1685, 0.1705, 0.1708, 0.1713, 0.1728, 0.1732, 0.1735, 0.1739, 0.1740, 0.1747, and an empty cell for row 11. The other columns (C2 and C3) are empty.

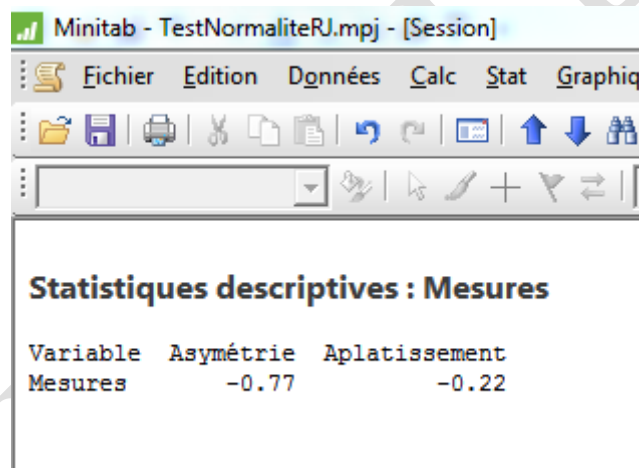
	C1	C2	C3
	Mesures		
1	0.1685		
2	0.1705		
3	0.1708		
4	0.1713		
5	0.1728		
6	0.1732		
7	0.1735		
8	0.1739		
9	0.1740		
10	0.1747		
11			

Nous faisons une analyse de statistiques descriptives comme avant mais en nous concentrant sur le Kurtosis et le Skewness:

<sup>1</sup> Au moins 11 méthodes de calculer le Skewness. Voir l'étude de Tabor, J. (2010), "Investigating the Investigative Task: Testing for Skewness - An Investigation of Different Test Statistics and their Power to Detect Skewness," Journal of Statistics Education, 18, 1-13. [www.amstat.org/publications/jse/v18n2/tabor.pdf](http://www.amstat.org/publications/jse/v18n2/tabor.pdf)



nous obtenons:



Dans Minitab, nous avons les règles suivantes:

- Si le **skewness** est nul, la distribution est symétrique (cela ne signifiant pas que la symétrie a lieu sur un pic de la distribution car dans le cas bimodal l'axe de symétrie peut-être entre les deux valeurs modales symétriques). Si le skewness est positif, la distribution (la valeur modale/médiane) est penchée à droite (ou il y a des valeurs extrêmes à droite). Si le skewness est négatif, la distribution (la valeur modale/médiane) est penchée à gauche (ou il y a des valeurs extrêmes à gauche)
- Si le **kurtosis** est nul (**mesokurtique**) alors l'aplatissement est similaire à celui d'une distribution Normale. Si la valeur est supérieure à 0 alors la distribution d'intérêt (**leptokurtique**) est alors plus haute que celle d'une distribution Normale à moyenne égale (inversement - **platikurtique**) - si le kurtosis est négatif bien évidemment)

Sinon à titre de comparaison (le premier est appelé "coefficient d'asymétrie de Fisher-Pearson" et le second "coefficient d'asymétrie ajusté et standardisé de Fisher-Pearson"):

$$\gamma_1 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left( \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \right)^3}$$

$$\gamma_2 = \gamma_1 \sqrt{\frac{n}{(n-1)(n-2)}}$$

$$\gamma_3 = \gamma_1 \left( \frac{n-1}{n} \right)^{3/2}$$

et:

$$\kappa_1 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left( \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \right)^4}$$

$$\kappa_2 = \frac{(N-1)((N+1)\kappa_1 - 3(N+1))}{(N-2)(N-3)} + 3$$

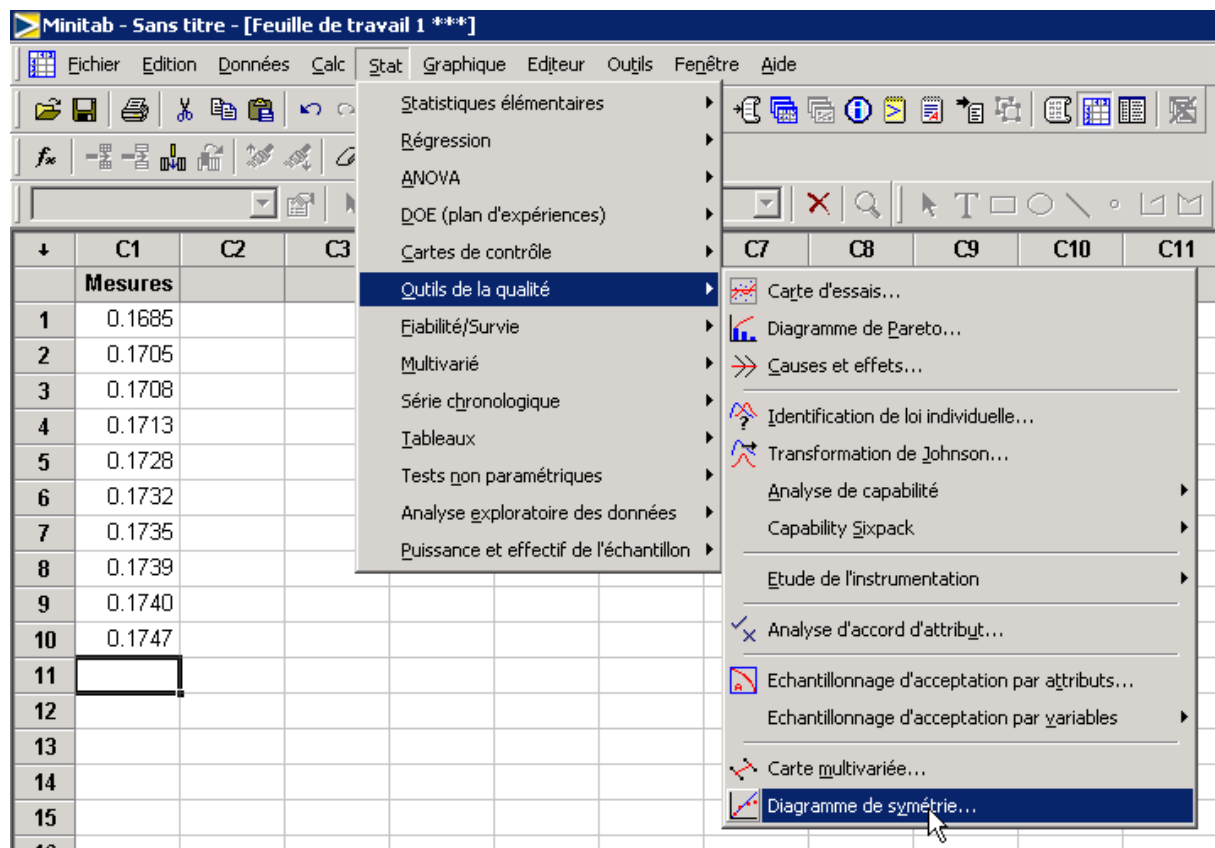
$$\kappa_3 = \frac{N(N+1)}{(N-1)(N-2)(N-3)} \kappa_1 - \frac{3(N-1)^2}{(N-2)(N-3)}$$

$$\kappa_4 = \kappa_1 - 3$$

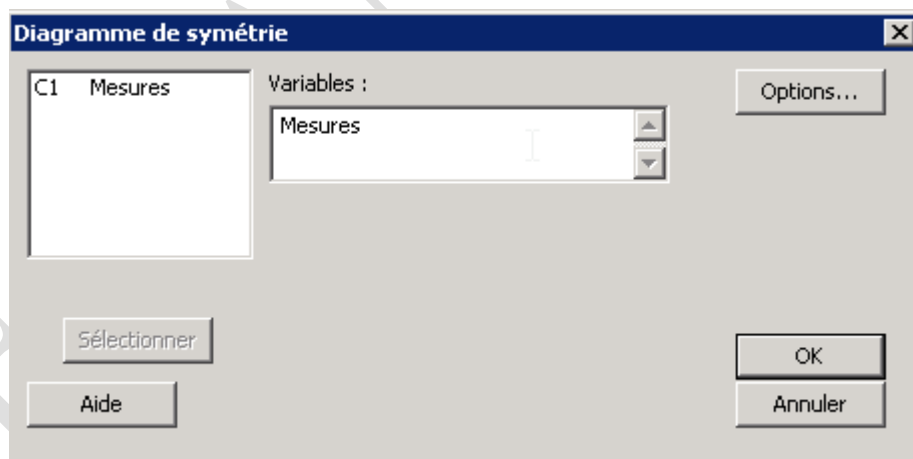
$$\kappa_5 = (\kappa_4 + 3) \left( 1 - \frac{1}{N} \right)^2 - 3$$

Skewness	R	-0.647( $\gamma_1$ ), -0.77( $\gamma_2$ ), -0.552( $\gamma_3$ )
	MATLAB	-0.647( $\gamma_1$ ), -0.77( $\gamma_2$ )
	Mathematica	-0.647( $\gamma_1$ )
	Minitab, SPSS, Excel	-0.77 ( $\gamma_3$ )
Kurtosis	R	-1.114( $\kappa_5$ ), -0.22 ( $\kappa_3$ ), -0.67 ( $\kappa_4$ )
	MATLAB	2.327 ( $\kappa_1$ ), 2.775( $\kappa_2$ )
	Mathematica	2.327( $\kappa_1$ )
	Minitab, SPSS, Excel	-0.22 ( $\kappa_3$ )

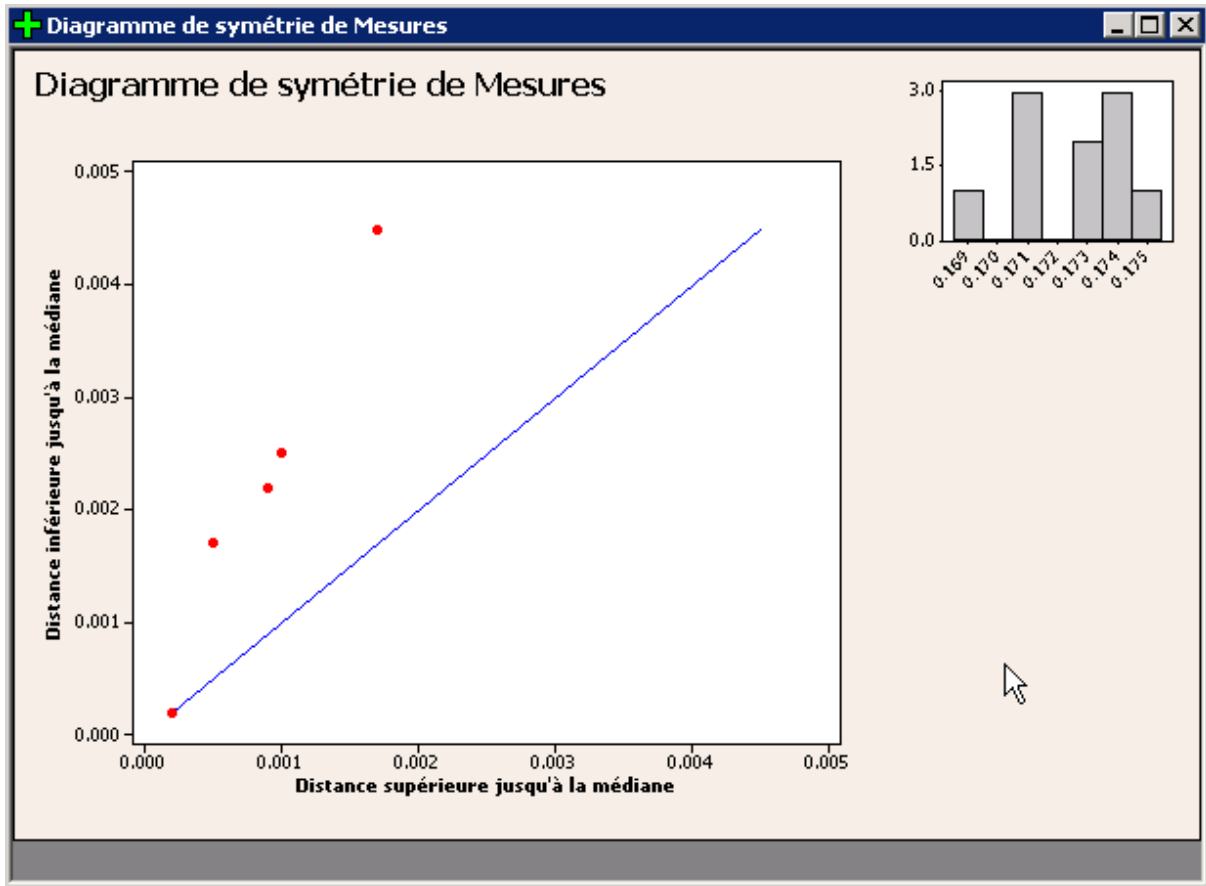
Minitab propose également une autre manière qualitative d'analyser la symétrie d'une distribution en allant dans le menu **Stat/Outils de la qualité/Diagramme de symétrie...**:



Il vient alors (les **Options...** contiennent juste le titre du graphique):



Et si nous validons par **OK**, il vient:



Comment lire ce graphique?

Eh bien dans un premier temps il faut calculer la médiane qui dans le cas présent vaut 0.1730.

Une fois ceci fait, nous trions les données:

0.1685
0.1705
0.1708
0.1713
0.1728
0.1732
0.1735
0.1739
0.174
0.1747

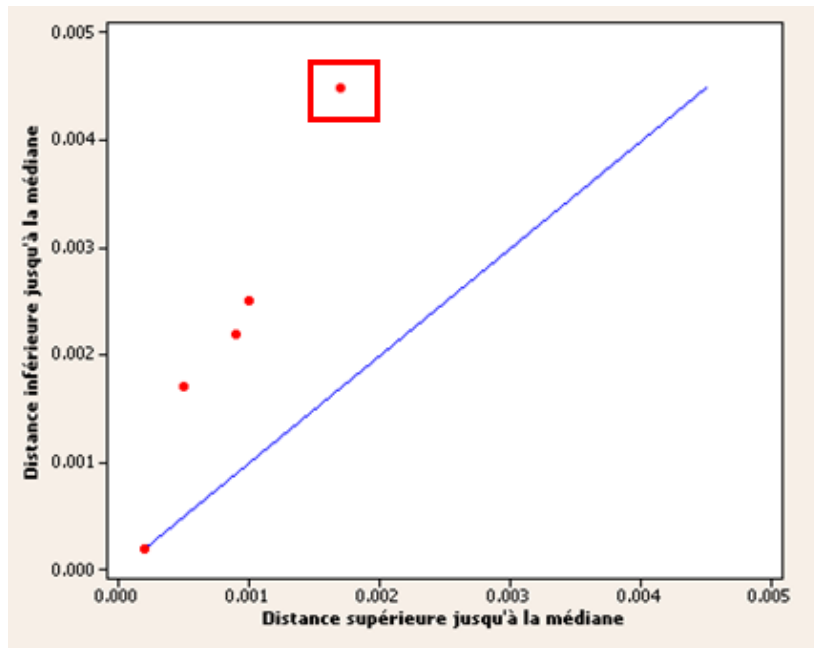
Et voilà le système de fonctionnement du graphique (nous faisons l'exemple avec deux des points visibles sur le graphique): D'abord nous prenons les deux premiers points extrêmes:

$$\{0.1685, 0.1747\}$$

et nous calculons leur distance absolue à la médiane:

$$\{|0.1685-0.1730|=0.0045,|0.1747-0.1730|=0.0017\}$$

Ce qui correspond au point:



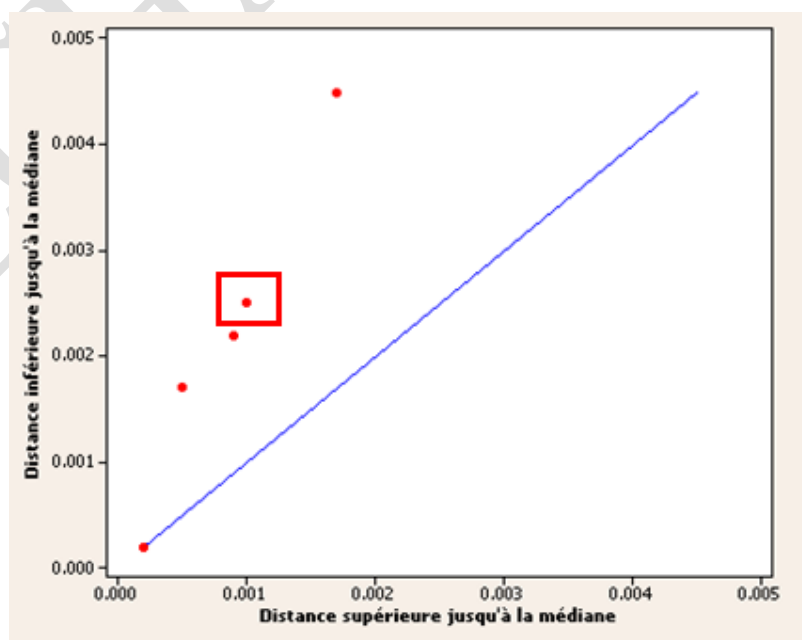
Ensuite pour un deuxième point, nous prenons les deux prochaines mesures extrêmes internes:

$$\{|0.1705-0.1730|=0.0025,|0.174-0.1730|=0.001\}$$

Ce qui correspond au point suivant:

$$\{|0.1705-0.1730|=0.0025,|0.174-0.1730|=0.001\}$$

Ce qui donne:



et ainsi de suite...

ÉCHANTILLON



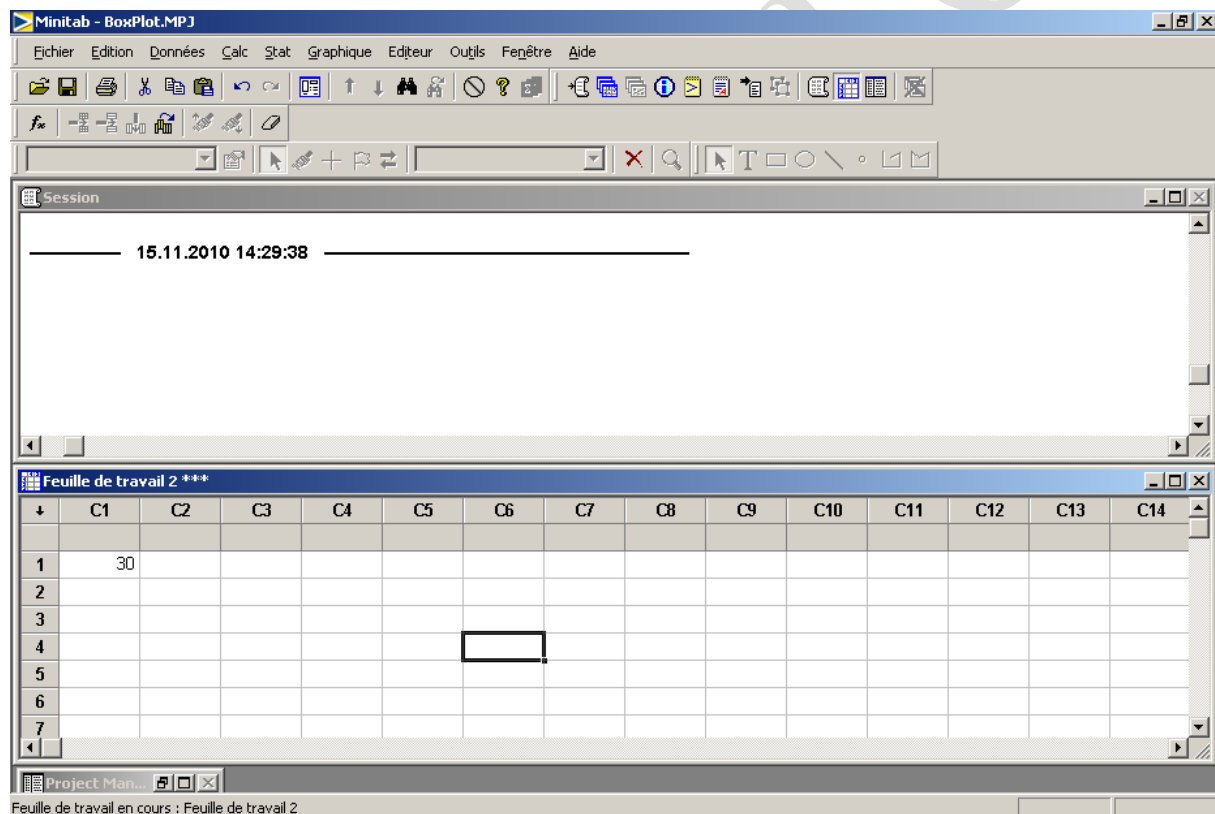
## 10.3. Exercice 35.: Calculer la probabilité cumulée et la probabilité inverse

Minitab® Statistical Software 15.1.1.0

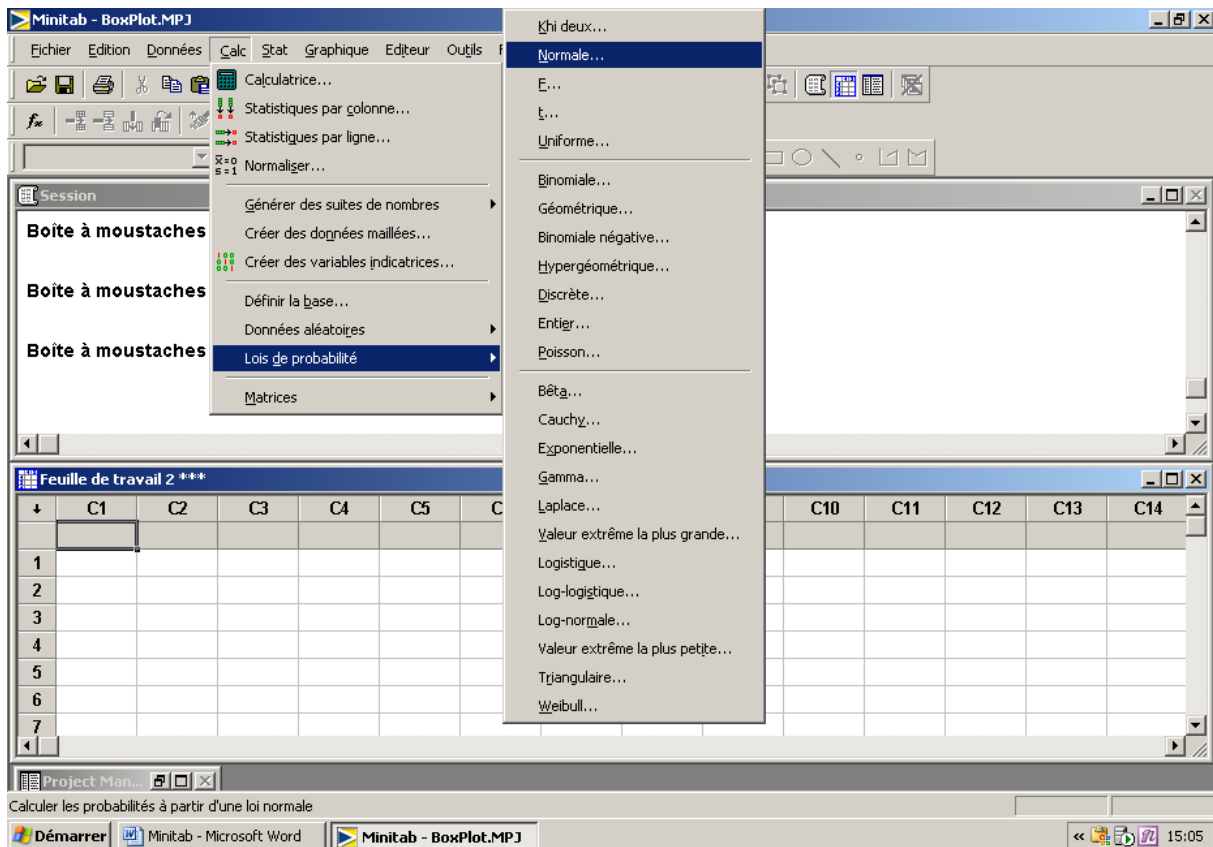
Il arrive souvent que l'on calcule les paramètres lois de probabilités afin de prendre des décisions relativement aux données de la fonction de répartition. Les cas les plus courants étant la loi Normale et la loi Beta.

Voyons comment procéder par exemple pour une loi Normale d'espérance 25 et d'écart-type de 3. Nous souhaiterons savoir qu'elle est la probabilité cumulée de finir un projet le 30<sup>ème</sup> jour.

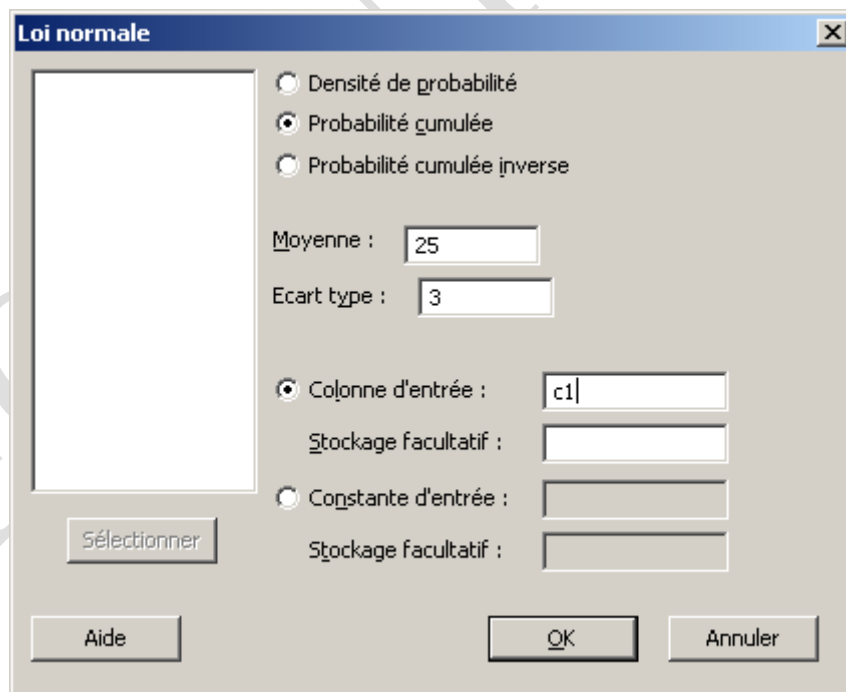
Nous créons un nouveau projet et nous écrivons le nombre 30 dans la première ligne de la première colonne:



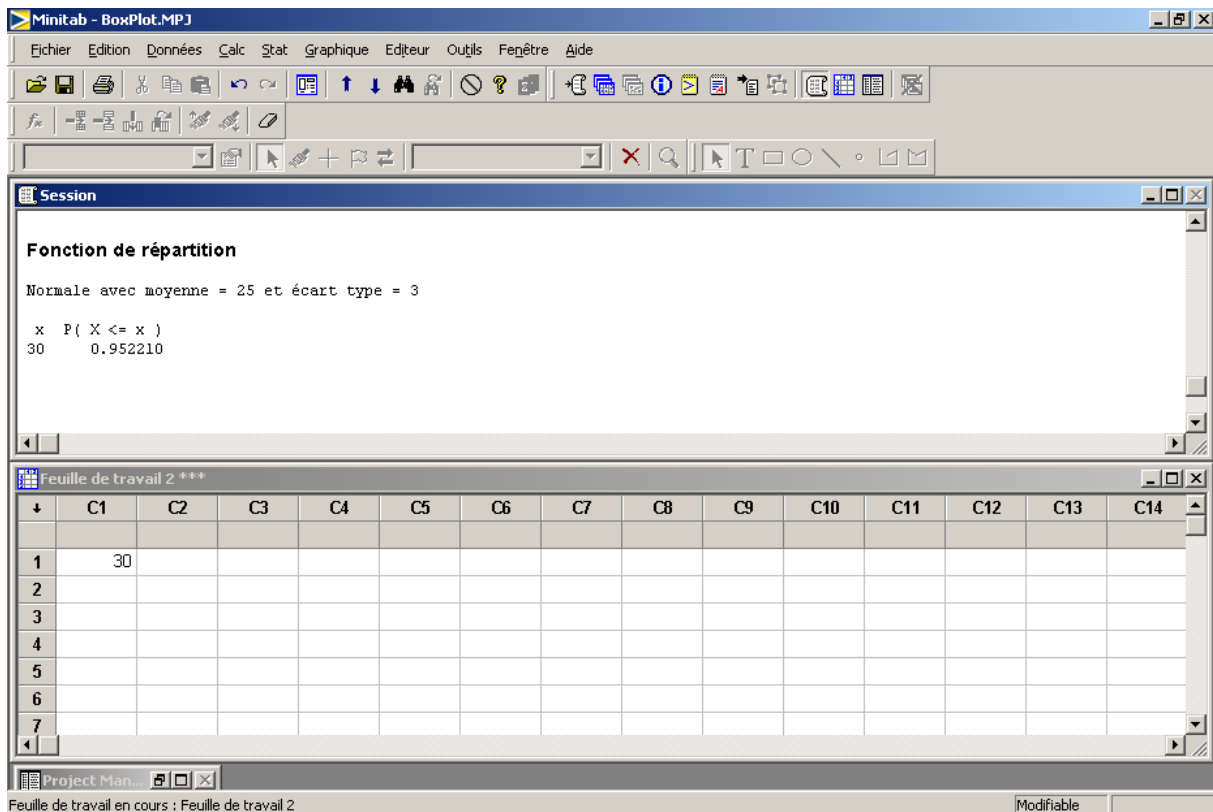
nous allons dans le menu **Stat/Lois de probabilité/Normale...**:



Vient alors:



Nous validons par **OK** pour obtenir dans la fenêtre de session:



The screenshot shows the Minitab software interface. The top window, titled "Session", displays the following text:

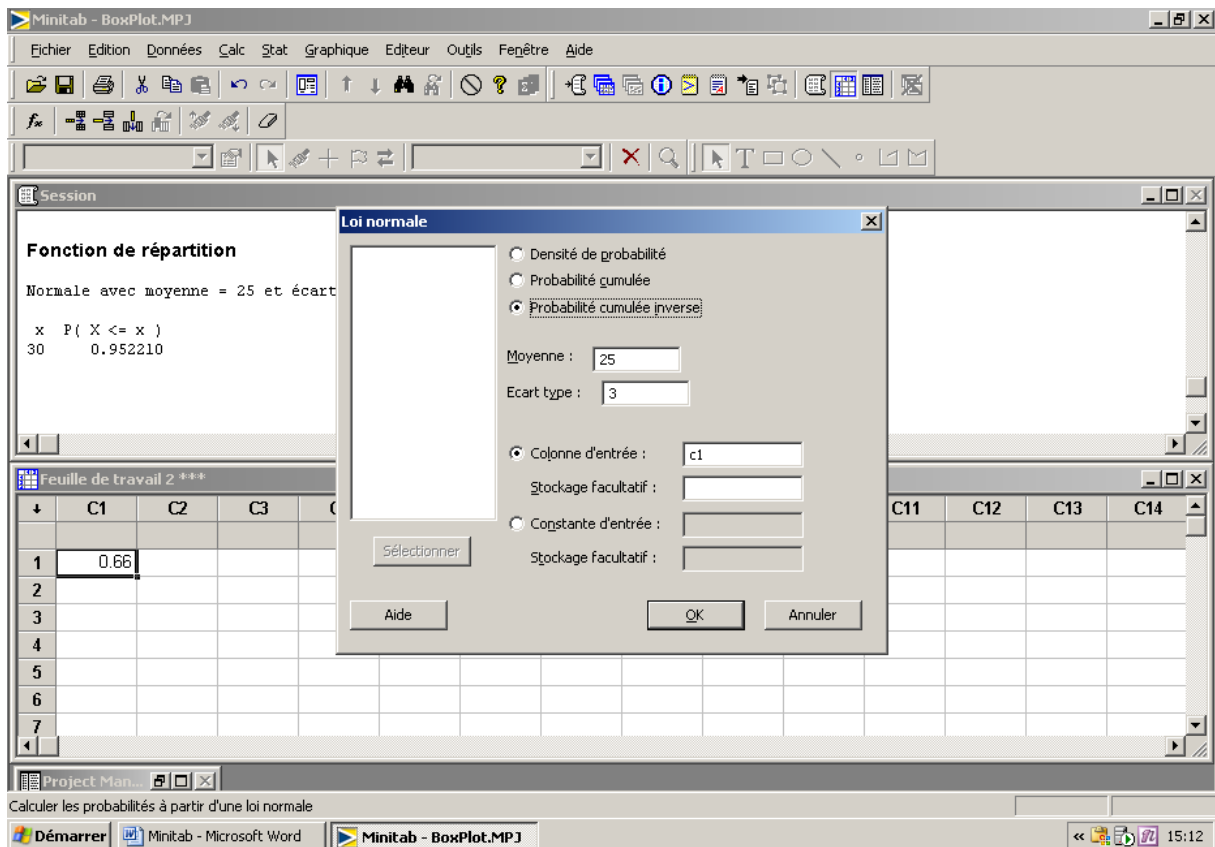
**Fonction de répartition**  
Normale avec moyenne = 25 et écart type = 3

x	P( X <= x )
30	0.952210

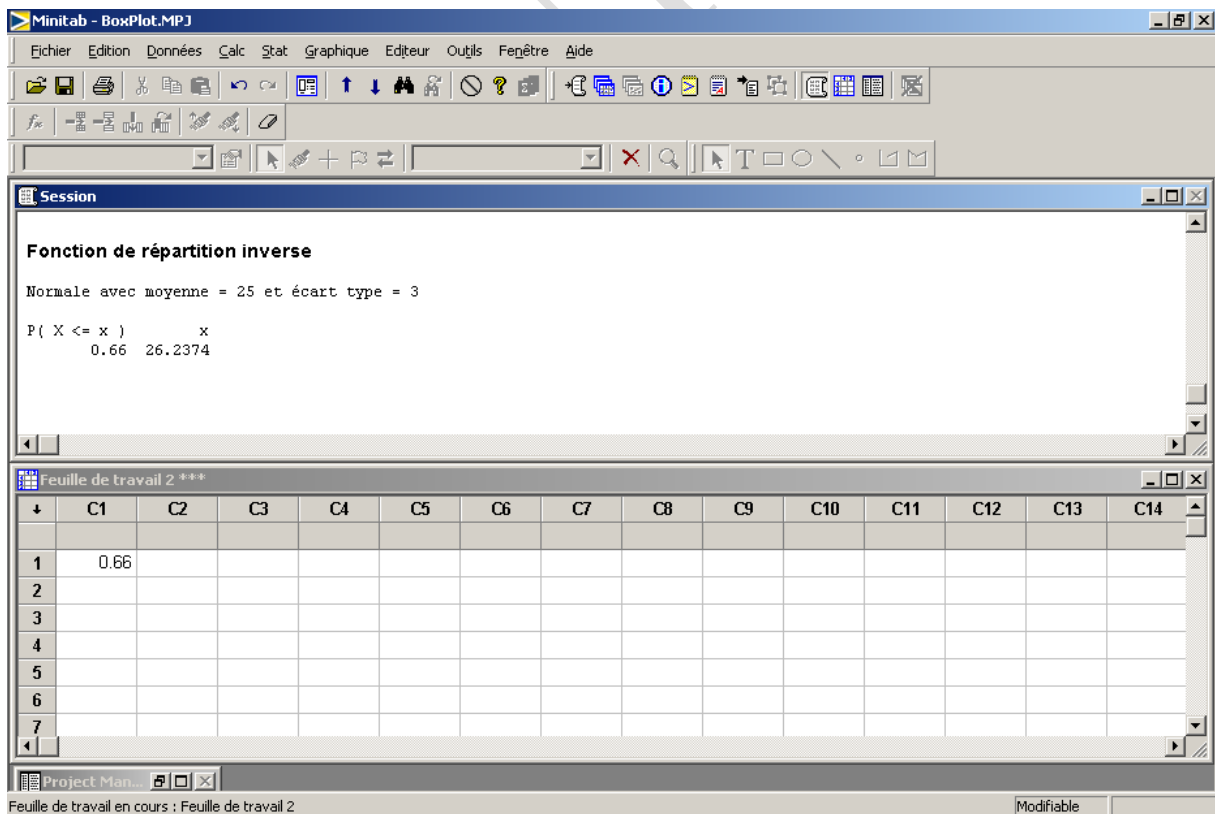
The bottom window, titled "Feuille de travail 2 \*\*\*", shows a spreadsheet with 14 columns (C1 to C14) and 7 rows (1 to 7). The value 30 is entered in cell C1, row 1.

Nous voyons ici un exemple typique où il est plus simple de travailler avec Microsoft Excel.

Évidemment nous pouvons faire l'inverse en nous demandant à qu'elle durée correspondant 66% de probabilité cumulée. Nous écrivons alors 66% dans la première ligne de la première colonne et retournons au même endroit:



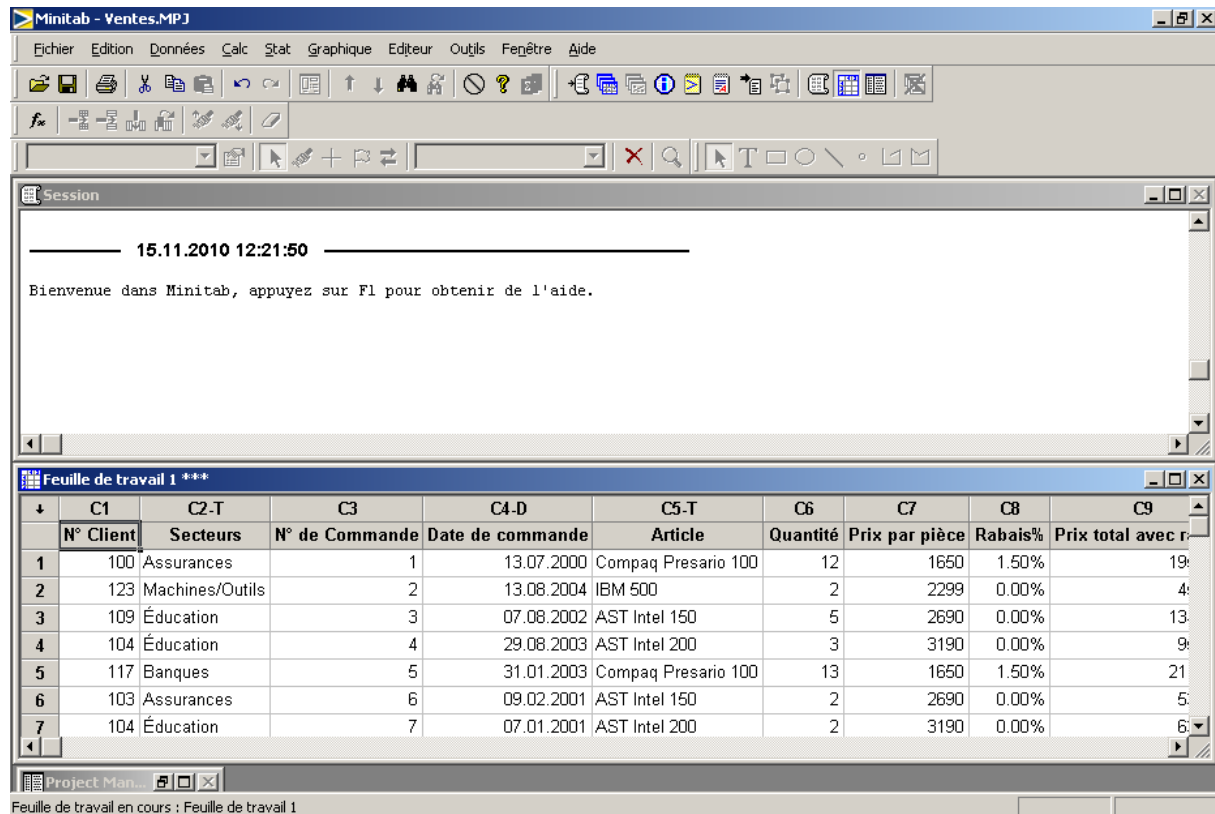
Et en validant par **OK** nous obtenons dans la fenêtre de session:



## 10.4. Exercice 36.: Mettre des données à plat (table de fréquence)

Minitab® Statistical Software 15.1.1.0

Ouvrez le fichier projet *Ventes.mpj*:



Session

15.11.2010 12:21:50

Bienvenue dans Minitab, appuyez sur F1 pour obtenir de l'aide.

Feuille de travail 1 \*\*\*

	C1	C2-T	C3	C4-D	C5-T	C6	C7	C8	C9
	N° Client	Secteurs	N° de Commande	Date de commande	Article	Quantité	Prix par pièce	Rabais%	Prix total avec r
1	100	Assurances	1	13.07.2000	Compaq Presario 100	12	1650	1.50%	19:
2	123	Machines/Outils	2	13.08.2004	IBM 500	2	2299	0.00%	4:
3	109	Éducation	3	07.08.2002	AST Intel 150	5	2690	0.00%	13:
4	104	Éducation	4	29.08.2003	AST Intel 200	3	3190	0.00%	9:
5	117	Banques	5	31.01.2003	Compaq Presario 100	13	1650	1.50%	21:
6	103	Assurances	6	09.02.2001	AST Intel 150	2	2690	0.00%	5:
7	104	Éducation	7	07.01.2001	AST Intel 200	2	3190	0.00%	6:

Project Man... [Icon] [Icon] [Icon]

Feuille de travail en cours : Feuille de travail 1

Nous souhaiterions faire l'équivalent d'un petit tableau de synthèse du type de ce que fait Microsoft Excel avec les tableaux croisés dynamiques.

Allez dans le menu **Stat/Tableaux/Trier à plat les variables individuelles...**:

The screenshot shows the Minitab interface with the 'Tableaux' menu open. The spreadsheet below shows sales data with columns for client number, sector, order number, date, article, quantity, price per piece, discount percentage, and total price with discount.

	C1	C2-T	C3	C4-D	C5-T	C6	C7	C8	C9
	N° Client	Secteurs	N° de Commande	Date de commande	Article	Quantité	Prix par pièce	Rabais%	Prix total avec r
1	100	Assurances	1	13.07.2000	Compaq Presario 100	12	1650	1.50%	19:
2	123	Machines/Outils	2	13.08.2004	IBM 500	2	2299	0.00%	4:
3	109	Éducation	3	07.08.2002	AST Intel 150	5	2690	0.00%	13:
4	104	Éducation	4	29.08.2003	AST Intel 200	3	3190	0.00%	9:
5	117	Banques	5	31.01.2003	Compaq Presario 100	13	1650	1.50%	21:
6	103	Assurances	6	09.02.2001	AST Intel 150	2	2690	0.00%	5:
7	104	Éducation	7	07.01.2001	AST Intel 200	2	3190	0.00%	6:

Nous aurons alors toujours sur la base de l'analyse des secteurs d'activité:

The dialog box 'Tri à plat des variables individuelles' shows a list of variables on the left and a 'Variables:' list on the right containing 'Secteurs'. The 'Affichage' section has checkboxes for 'Dénombrements' and 'Pourcentages' which are checked, and 'Dénombrements cumulés' and 'Pourcentages cumulés' which are unchecked.

En validant par **OK** dans la fenêtre de sortie:

Minitab - Ventes.MPJ

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

Session

Secteurs	Dénombrement	Pourcentage
Alimentaire	7	6.42
Assurances	10	9.17
Banques	9	8.26
Construction	21	19.27
Distribution	13	11.93
Éducation	19	17.43
Machines/Outils	18	16.51
Pharmaceutique	12	11.01
N=	109	

Feuille de travail 1 \*\*\*

	C1	C2-T	C3	C4-D	C5-T	C6	C7	C8	C9
	N° Client	Secteurs	N° de Commande	Date de commande	Article	Quantité	Prix par pièce	Rabais%	Prix total avec r
1	100	Assurances	1	13.07.2000	Compaq Presario 100	12	1650	1.50%	19:
2	123	Machines/Outils	2	13.08.2004	IBM 500	2	2299	0.00%	4:
3	109	Éducation	3	07.08.2002	AST Intel 150	5	2690	0.00%	13:
4	104	Éducation	4	29.08.2003	AST Intel 200	3	3190	0.00%	9:
5	117	Banques	5	31.01.2003	Compaq Presario 100	13	1650	1.50%	21:
6	103	Assurances	6	09.02.2001	AST Intel 150	2	2690	0.00%	5:
7	104	Éducation	7	07.01.2001	AST Intel 200	2	3190	0.00%	6:

Project Man.

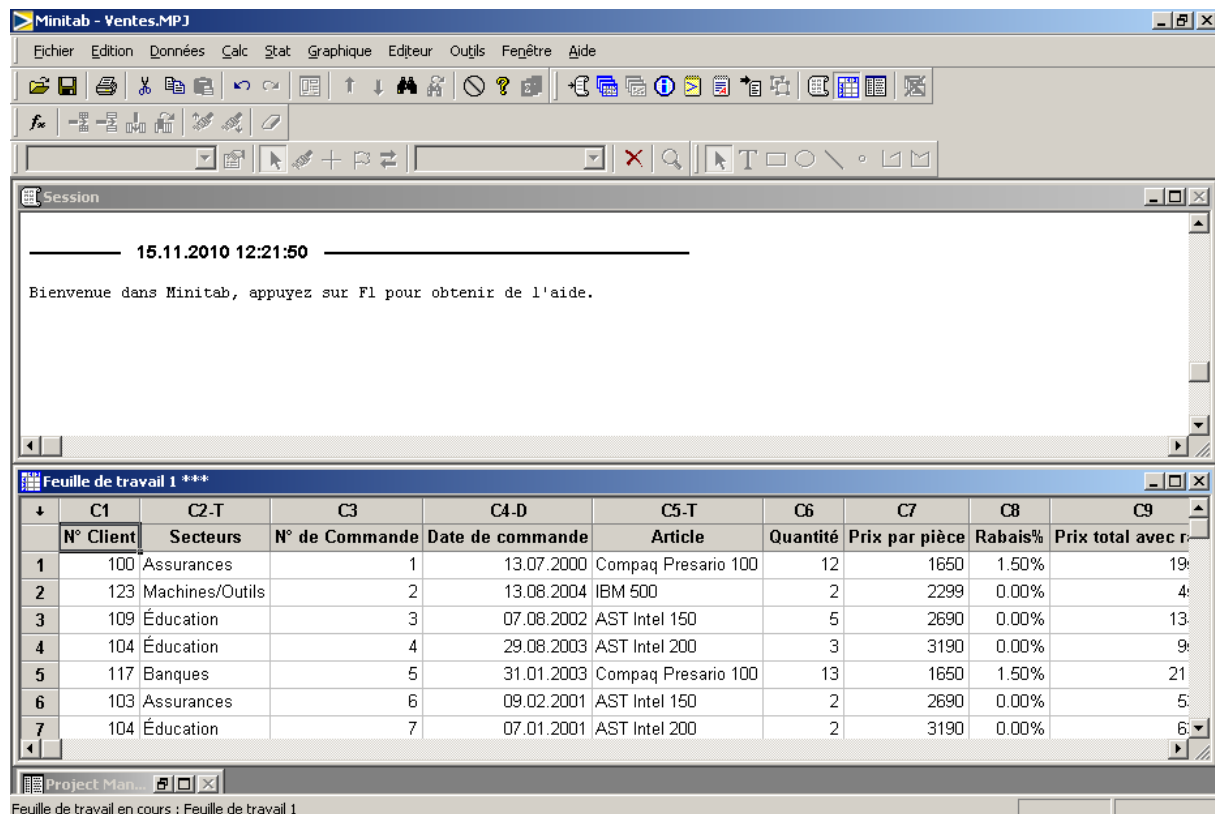
Feuille de travail en cours : Feuille de travail 1

Modifiable

## 10.5. Exercice 37.: Mettre des données à plat (table de contingence)

Minitab® Statistical Software 15.1.1.0

Ouvrez le fichier projet *Ventes.mpj*:



The screenshot shows the Minitab interface. The 'Session' window displays the following text:

```

15.11.2010 12:21:50
Bienvenue dans Minitab, appuyez sur F1 pour obtenir de l'aide.

```

The 'Feuille de travail 1' window displays the following data table:

	C1	C2-T	C3	C4-D	C5-T	C6	C7	C8	C9
	N° Client	Secteurs	N° de Commande	Date de commande	Article	Quantité	Prix par pièce	Rabais%	Prix total avec r
1	100	Assurances	1	13.07.2000	Compaq Presario 100	12	1650	1.50%	19:
2	123	Machines/Outils	2	13.08.2004	IBM 500	2	2299	0.00%	4:
3	109	Éducation	3	07.08.2002	AST Intel 150	5	2690	0.00%	13:
4	104	Éducation	4	29.08.2003	AST Intel 200	3	3190	0.00%	9:
5	117	Banques	5	31.01.2003	Compaq Presario 100	13	1650	1.50%	21:
6	103	Assurances	6	09.02.2001	AST Intel 150	2	2690	0.00%	5:
7	104	Éducation	7	07.01.2001	AST Intel 200	2	3190	0.00%	6:

Nous souhaiterions faire l'équivalent d'un petit tableau de contingence en % des Totaux du type de ce que fait Microsoft Excel avec les tableaux croisés dynamiques.

Pour cela, allons dans le menu **Stat/Tableaux/Tableaux à entrées multiples et Khi-deux...**:



The screenshot shows the Minitab interface with the 'Stat' menu open. The 'Tableaux' option is selected, and a sub-menu is visible with 'Tableau à entrées multiples et Khi deux...' highlighted. The main window displays a data table with the following content:

	C1	C2-T	C3	C4-D	C5-T	C6	C7	C8	C9
	N° Client	Secteurs	N° de Commande	Date de commande	Article	Quantité	Prix par pièce	Rabais%	Prix total avec r
1	100	Assurances	1	13.07.2000	Compaq Presario 100	12	1650	1.50%	19
2	123	Machines/Outils	2	13.08.2004	IBM 500	2	2299	0.00%	4
3	109	Éducation	3	07.08.2002	AST Intel 150	5	2690	0.00%	13
4	104	Éducation	4	29.08.2003	AST Intel 200	3	3190	0.00%	9
5	117	Banques	5	31.01.2003	Compaq Presario 100	13	1650	1.50%	21
6	103	Assurances	6	09.02.2001	AST Intel 150	2	2690	0.00%	5
7	104	Éducation	7	07.01.2001	AST Intel 200	2	3190	0.00%	6

Prenons:

The dialog box 'Tableau à entrées multiples et Khi deux' is shown with the following configuration:

- Variables de catégorie :**
  - Pour les lignes : Secteurs
  - Pour les colonnes : Article
  - Pour les couches : (empty)
- Les effectifs sont dans : (empty) (facultatif)
- Affichage :**
  - Dénombrements
  - Pourcentages des lignes
  - Pourcentages des colonnes
  - Total des pourcentages

Buttons: Sélectionner, Aide, Khi deux..., Autres statistiques..., Options..., OK, Annuler.

Et validons par **OK** pour obtenir dans la fenêtre de session:

**Tableau de statistiques : Secteurs, Article**

Lignes : Secteurs    Colonnes : Article

Secteurs	Article	Alimentaire	Banques	Education	Pharmaceutique	Tous
Compaq	Presario	1.835	1.835	1.835	1.835	1.835
Compaq	Intel	1.835	1.835	1.835	1.835	1.835
Compaq	Intel	1.835	1.835	1.835	1.835	1.835
Compaq	Presario	1.835	1.835	1.835	1.835	1.835
Compaq	Tous	1.835	1.835	1.835	1.835	1.835
IBM	500	1.835	1.835	1.835	1.835	1.835
Tous		1.835	1.835	1.835	1.835	1.835

## 10.6. Exercice 38.: Test d'indépendance du Khi-deux d'une table de contingence

Minitab® Statistical Software 15.1.1.0

Dans le cours de Méthodes Numériques nous avons travaillé avec un tableau de contingence:

	Feuillus	Résineux	Mixtes	Total par dép.
L'Aisne (A)	103'488	7'240	620	111'350
L'Oise (O)	103'813	7'263	622	111'700
La Somme (S)	46'098	3'225	276	49'600
Total	253'400	17'730	1'520	272'650

Nous voulions savoir si le nombre d'arbre dépendait réellement des régions dans lesquelles ils poussent ou si ces valeurs ne sont que dues au hasard de l'échantillon?

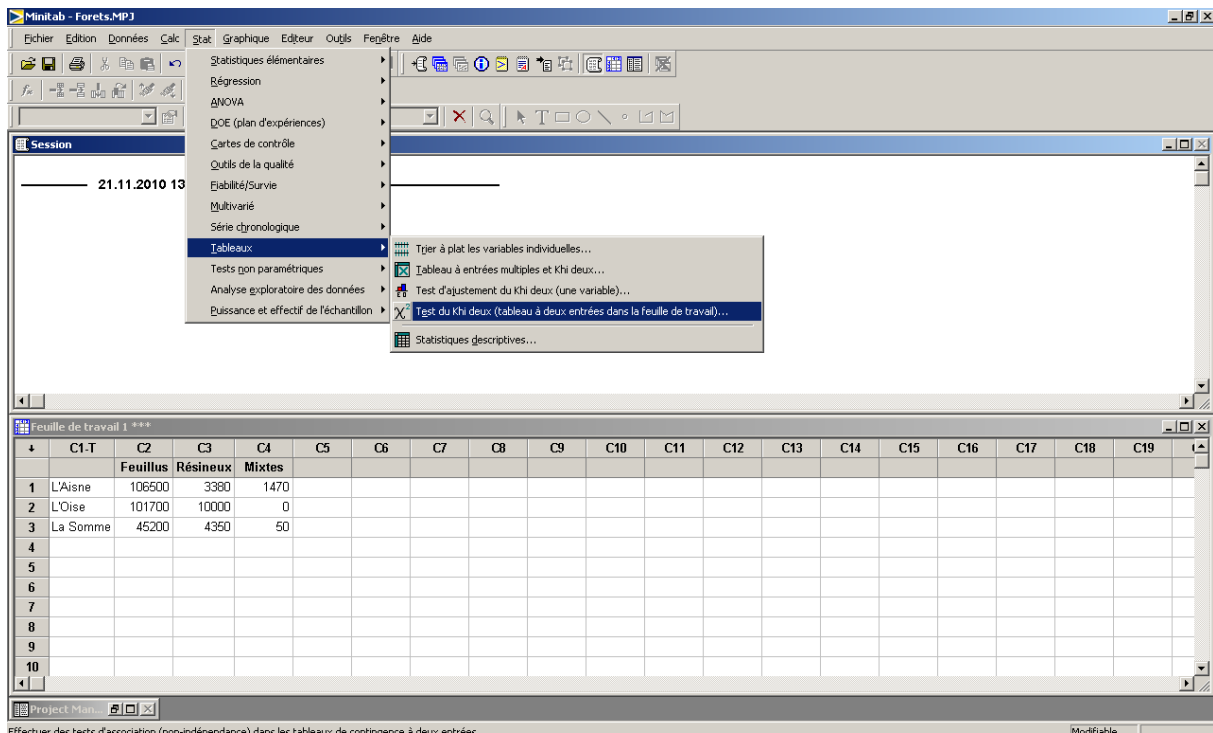
Nous avons vu qu'une bonne technique consistait à utiliser un test du Khi-deux et que le résultat donnait que les différences étaient significatives. Faisons la même analyse avec Minitab® Statistical Software.

Ouvrez le fichier *Forets.mpj*:

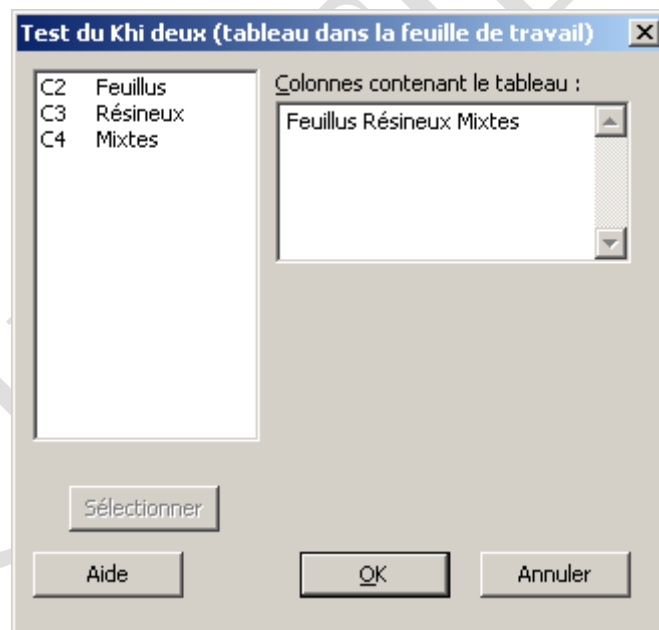
The screenshot shows the Minitab interface. The 'Session' window is open, displaying the date and time: 21.11.2010 13:23:22. Below it, the 'Feuille de travail 1' window is open, showing a table with the following data:

	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19
1	L'Aisne	106500	3380	1470															
2	L'Oise	101700	10000	0															
3	La Somme	45200	4350	50															

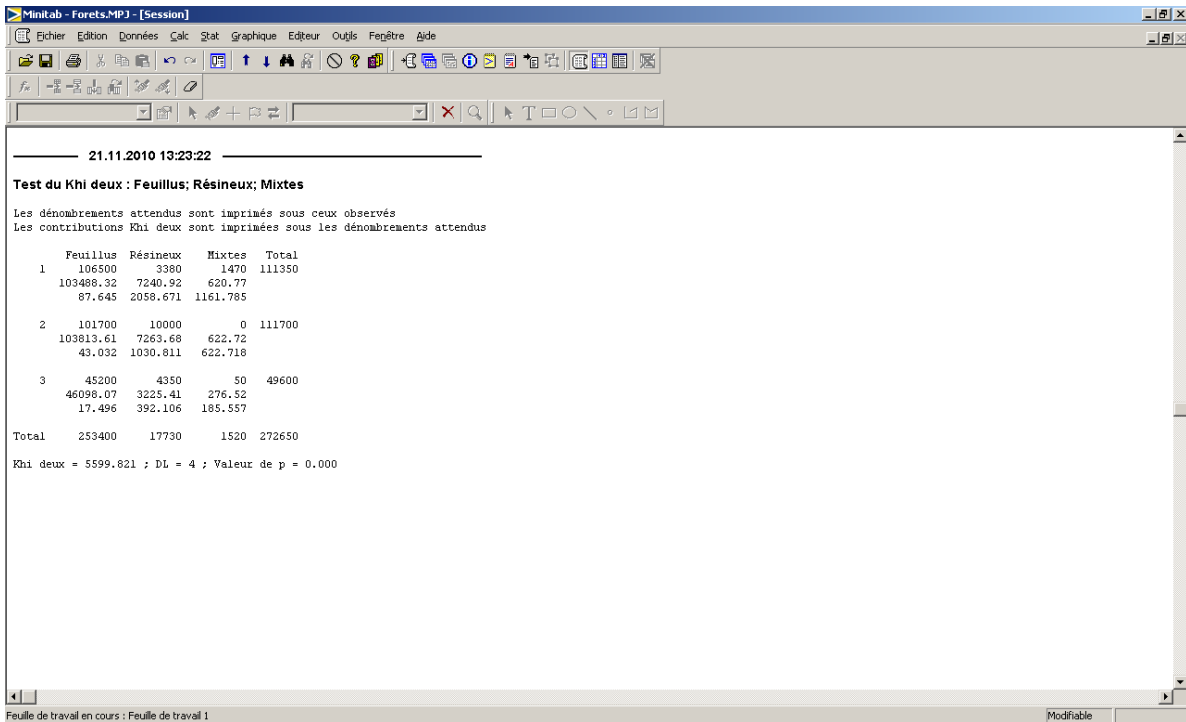
Nous allons dans le menu **Stat/Tableaux/Test du Khi-deux (tableau à deux entrées dans la feuille de travail)**:



Nous avons alors:

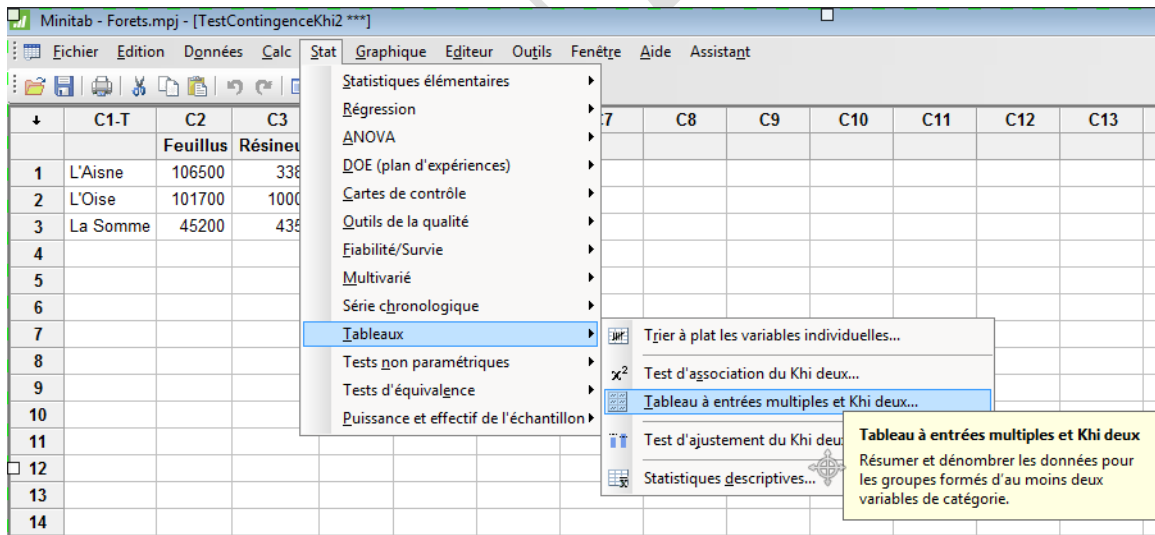


Et si nous validons par **OK** nous obtenons dans la fenêtre de session:

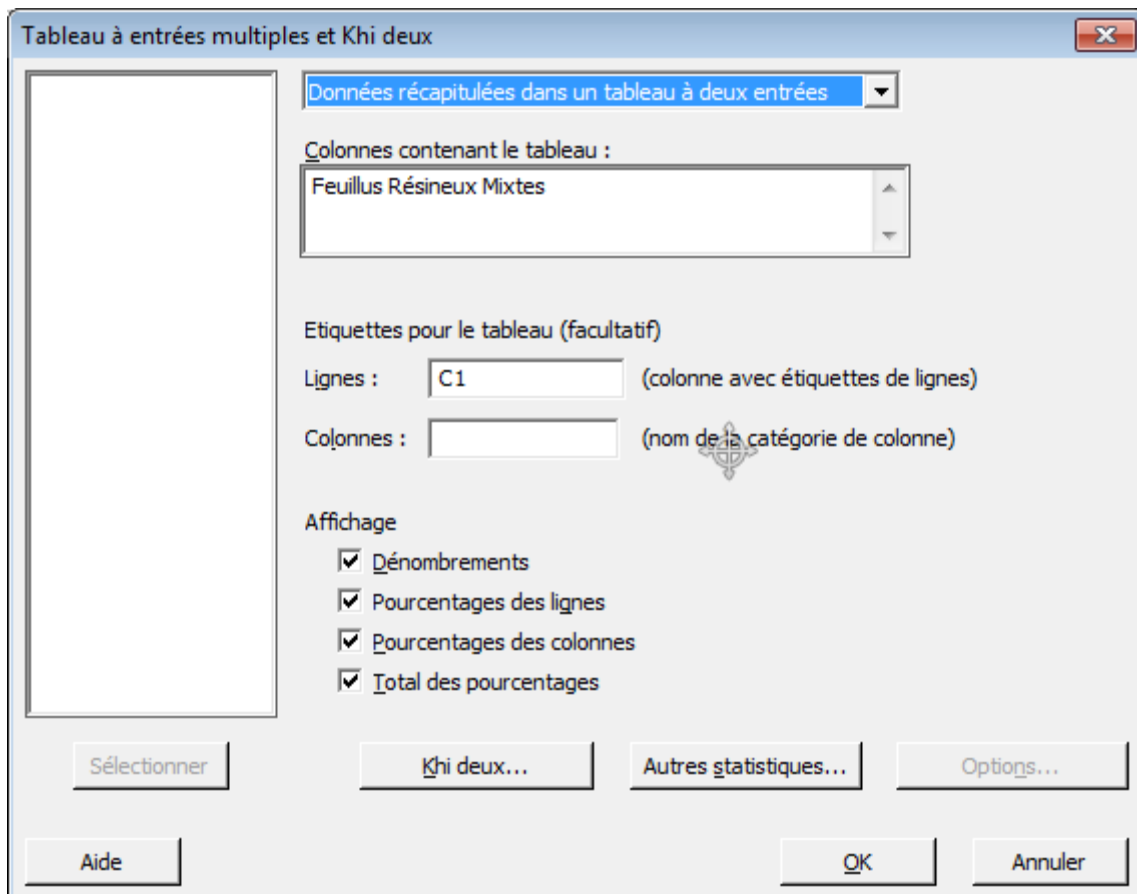


Nous avons donc exactement le même résultat qu'avec les calculs faits avec Microsoft Excel et avec la même conclusion (les populations sont significativement différentes).

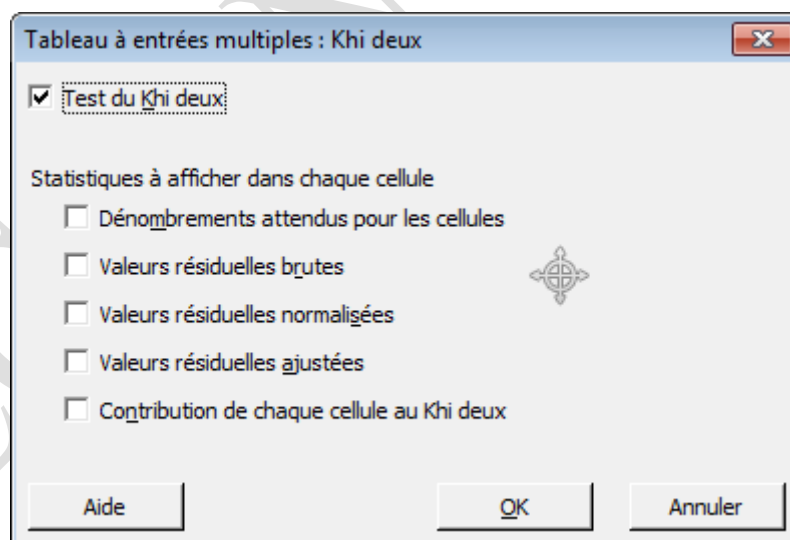
Dans Minitab 17, la procédure a significativement changée. D'abord l'option n'a plus le même nom **Stat/Tableaux/Tableau à entrées multiples et Khi deux...**:



Dans la boîte de dialogue qui apparaît, nous mettons:



et dans le bouton **Khi deux...** nous cochons l'option **Test du Khi-deux**:



Nous validons deux fois par **OK** pour obtenir:

Minitab - Forets.mpj - [Session]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assis

Tableau de statistiques : C1; Colonnes de la feuille de trava

Lignes : C1 Colonnes : Colonnes de la feuille de trava

	Feuillus	Résineux	Mixtes	Total
L'Aisne	106500	3380	1470	111350
	95,64	3,04	1,32	100,00
	42,03	19,06	96,71	40,84
	39,061	1,240	0,539	40,840
L'Oise	101700	10000	0	111700
	91,05	8,95	0,00	100,00
	40,13	56,40	0,00	40,97
	37,301	3,668	0,000	40,968
La Somme	45200	4350	50	49600
	91,13	8,77	0,10	100,00
	17,84	24,53	3,29	18,19
	16,578	1,595	0,018	18,192
Total	253400	17730	1520	272650
	92,94	6,50	0,56	100,00
	100,00	100,00	100,00	100,00
	92,940	6,503	0,557	100,000

Contenu de la cellule :      Dénombrement  
    % de ligne  
    % de colonne  
    % du total

Khi deux de Pearson = 5599,821; DL =4; valeur de p = 0,000  
 Khi deux du rapport de vraisemblance = 6357,222; DL = 4; valeur de p = 0,000

La conclusion reste cependant la même bien heureusement...

Comparons la même analyse via un autre cheminement dans Minitab (la seule différence avec la méthode précédente étant qu'il n'y a pas besoin de cocher la case **Test du Khi-deux** comme nous allons pouvoir le constater...). Nous allons dans **Stat/Tableaux/Test d'association du Khi-deux...**:

The screenshot shows the Minitab interface with the 'Stat' menu open. The 'Tableaux' option is selected, and the 'Test d'association du Khi deux...' option is highlighted. A tooltip for this option reads: 'Déterminer si deux variables de catégorie sont associées.'

	C1-T	C2	C3
		Feuillus	Résineux
1	L'Aisne	106500	3380
2	L'Oise	101700	10000
3	La Somme	45200	4350
4			
5			
6			
7			
8			
9			
10			
11			
12			

Dans la boîte de dialogue qui apparaît nous mettons alors:

The dialog box 'Test d'association du Khi deux' is shown with the following settings:

- Données récapitulées dans un tableau à deux entrées
- Colonnes contenant le tableau : Feuillus Résineux Mixtes
- Etiquettes pour le tableau (facultatif):
  - Lignes : C1 (colonne avec étiquettes de lignes)
  - Colonnes : Lieux (nom de la catégorie de colonne)
- Buttons: Sélectionner, Statistiques..., Options..., Aide, OK, Annuler

Et nous validons par **OK**, ce qui donne:



## Test d'association du Khi deux : C1; Lieux

	Lignes : C1    Colonnes : Lieux			
	Feuillus	Résineux	Mixtes	Total
L'Aisne	106500 103488	3380 7241	1470 621	111350
L'Oise	101700 103814	10000 7264	0 623	111700
La Somme	45200 46098	4350 3225	50 277	49600
Total	253400	17730	1520	272650

Contenu de la cellule  
 Dénombrement  
 Dénombrement attendu

### Test du Khi deux

	Khi deux	DL	Valeur de P
Pearson	5599.821	4	0.000
Rapport de vraisemblance	6357.222	4	0.000

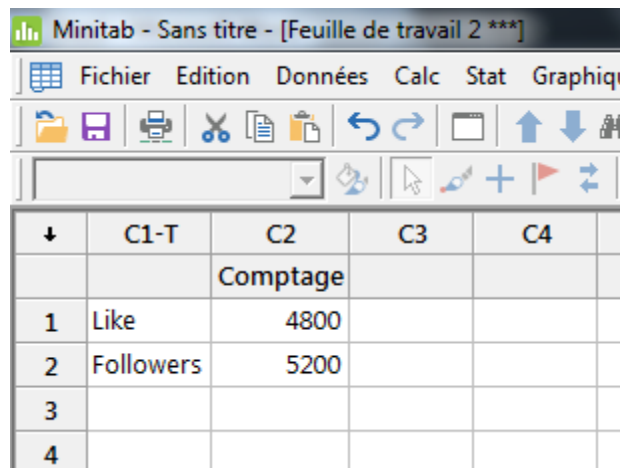
Nous retombons bien sur les mêmes résultats qu'avant mais simplement avec une procédure un tout petit peu simplifiée.

## 10.7. Exercice: Test du Khi-deux de la différence

Minitab® Statistical Software 18.1

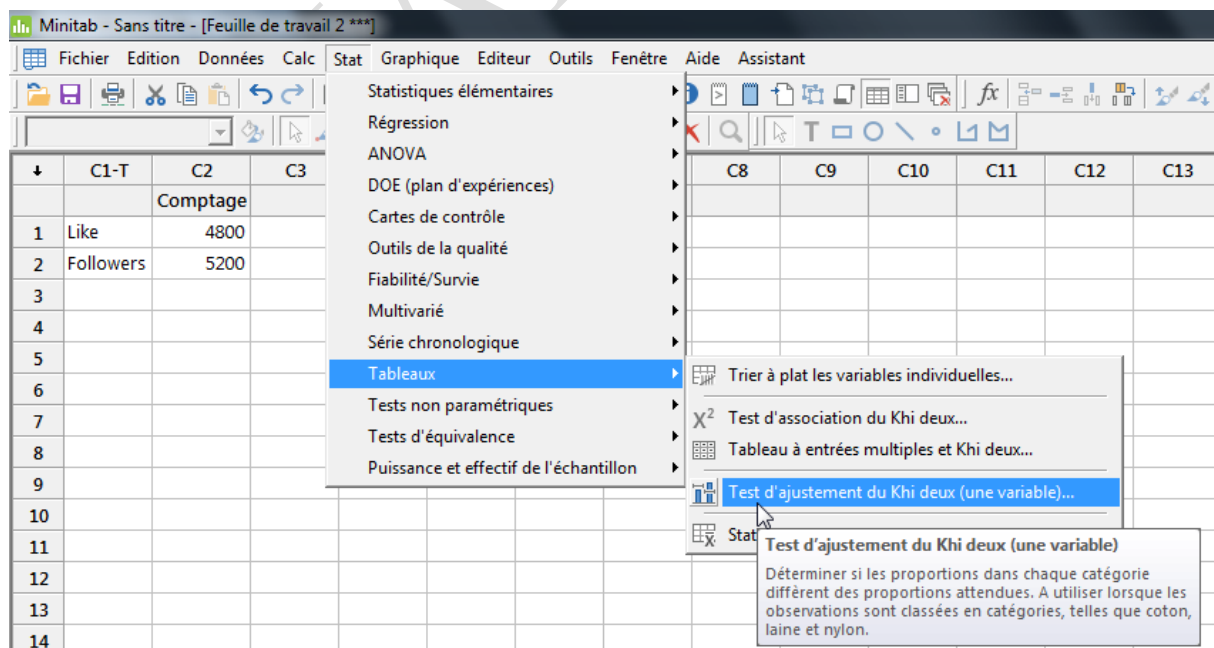
Comme nous l'avons mentionné et vu dans le cours théorique, le test du Khi-deux peut mathématiquement être utilisé pour de nombreux tests différents mais dont l'interprétation diffère. Bien évidemment, comparer deux comptages (sans connaître la proportion correspondante!!!) est un cas particulier du test d'indépendance (ie du test d'homogénéité ou du test d'ajustement selon le point de vue!).

Considérons donc les données suivantes:



	C1-T	C2	C3	C4
		Comptage		
1	Like	4800		
2	Followers	5200		
3				
4				

Nous souhaiterions savoir s'il y a une différence significative. Nous allons donc dans le menu **Stat/Tableaux/Test d'ajustement du Khi-deux (une variable...)...**:

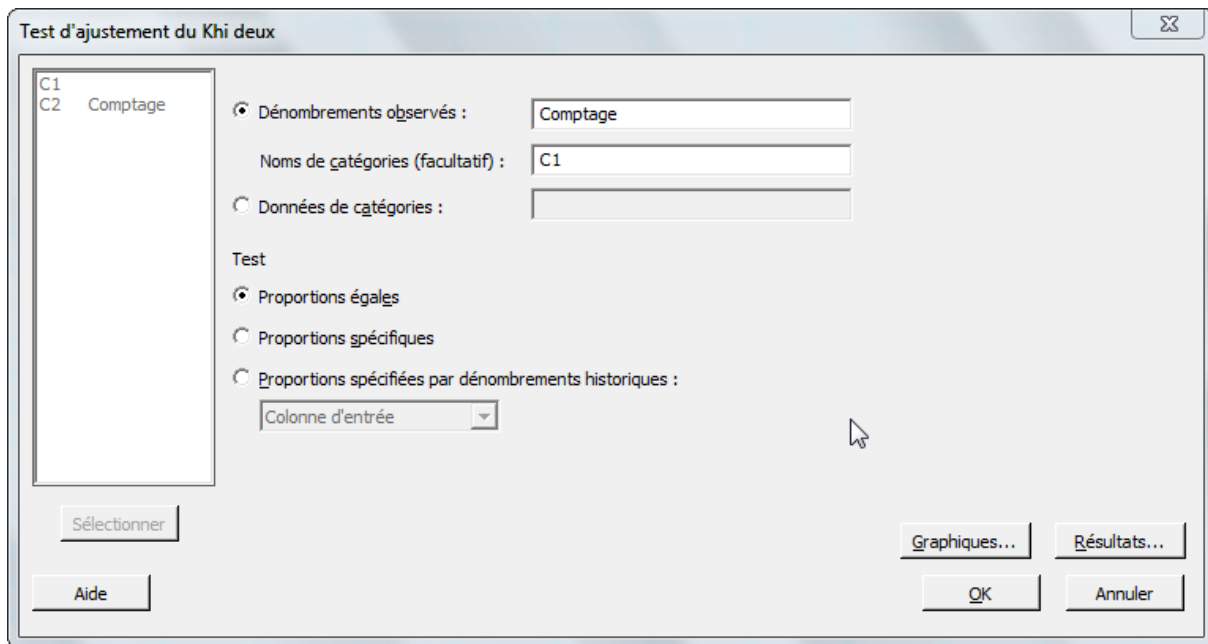


Stat

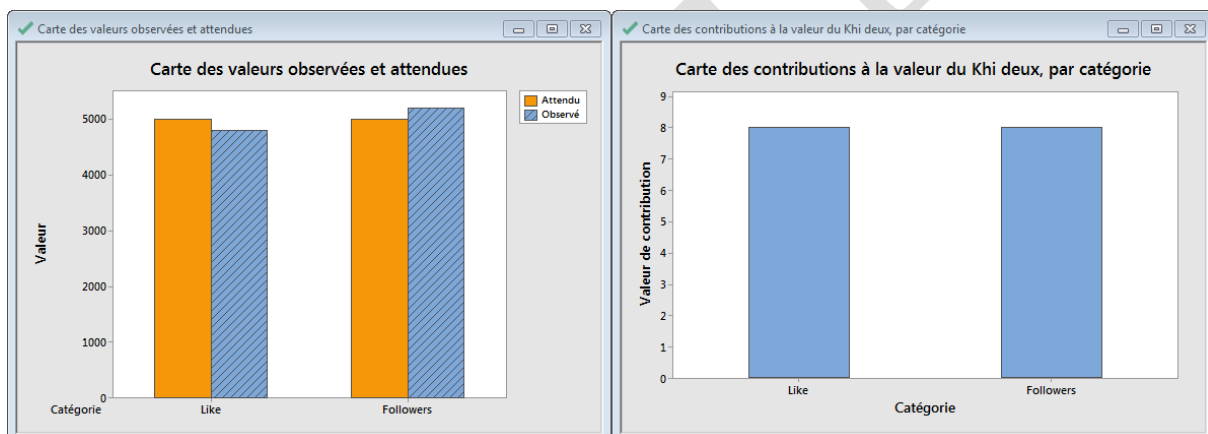
- Statistiques élémentaires
- Régression
- ANOVA
- DOE (plan d'expériences)
- Cartes de contrôle
- Outils de la qualité
- Fiabilité/Survie
- Multivarié
- Série chronologique
- Tableaux**
  - Trier à plat les variables individuelles...
  - $\chi^2$  Test d'association du Khi deux...
  - Tableau à entrées multiples et Khi deux...
  - Test d'ajustement du Khi deux (une variable)...**
- Tests non paramétriques
- Tests d'équivalence
- Puissance et effectif de l'échantillon

**Test d'ajustement du Khi deux (une variable)**  
 Déterminer si les proportions dans chaque catégorie diffèrent des proportions attendues. A utiliser lorsque les observations sont classées en catégories, telles que coton, laine et nylon.

Dans la boîte de dialogue qui apparaît nous mettons:



Ce qui donne graphiquement:



Et au niveau de la fenêtre de session:

### Test d'ajustement du Khi deux pour les dénombrements ... Comptage

Utilisation des noms de catégories dans C1

#### Dénombrements observés et attendus

Catégorie	Observé	Proportion testée	Attendu	Contribution au Khi deux
Like	4800	0.5	5000	8
Followers	5200	0.5	5000	8

#### Test du Khi deux

N	DL	Khi deux	Valeur de P
10000	1	16	0.000

## 10.8. Exercice 39.: Test exact de Fisher

Minitab® Statistical Software 15.1.1.0

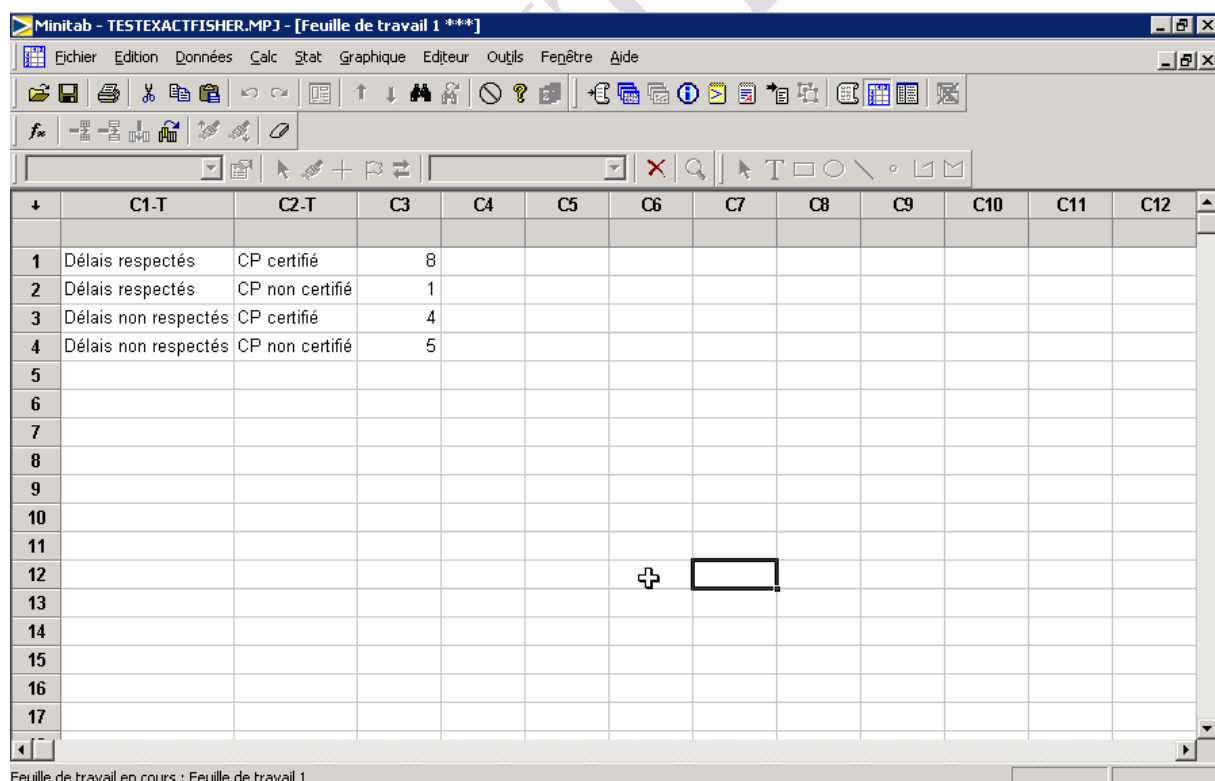
Nous avons étudié en détail dans le cours théorique de méthodes numériques le test exact de Fisher et fait les démonstrations mathématiques avec les tenants et aboutissants permettant ensuite de faire le calcul dans Microsoft Excel. Comme à l'habitude, l'objectif sera ici de vérifier que l'on retombe bien sur le même résultat avec Minitab.

Nous sommes donc partis du tableau suivant:

Projets	Chef de projet certifié	Chef de projet non certifié	Total
Délais respectés	8	1	9
Délais non respectés	4	5	9
Total	12	6	18

Et donc le principe du test est de vérifier si la configuration observée dans le tableau de contingence est une situation extrême par rapport aux situations possibles.

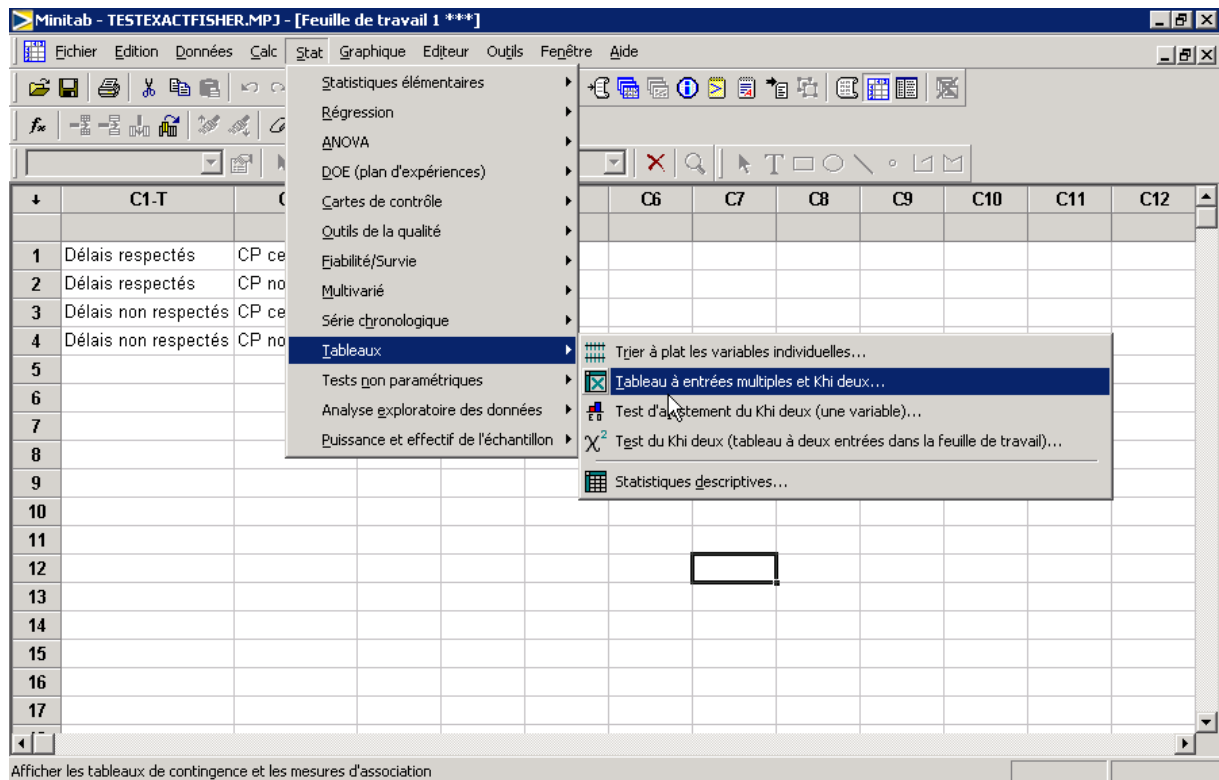
Pour ce faire, dans Minitab, nous créons le fichier suivant:



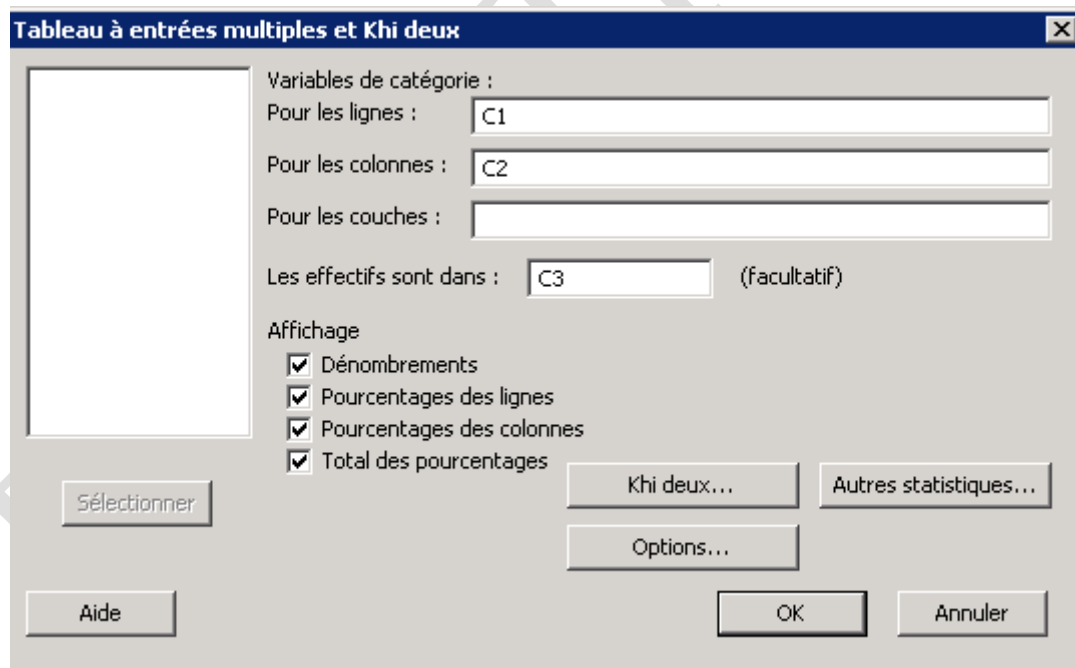
The screenshot shows the Minitab software interface with a data table. The table has 4 rows and 12 columns. The first four rows contain data for project delays and certification status. The first column is labeled '1' through '4'. The second column contains 'Délais respectés' and 'Délais non respectés'. The third column contains 'CP certifié' and 'CP non certifié'. The fourth column contains the counts 8, 1, 4, and 5. The remaining columns are empty.

	C1-T	C2-T	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
1	Délais respectés	CP certifié	8									
2	Délais respectés	CP non certifié	1									
3	Délais non respectés	CP certifié	4									
4	Délais non respectés	CP non certifié	5									
5												
6												
7												
8												
9												
10												
11												
12												
13												
14												
15												
16												
17												

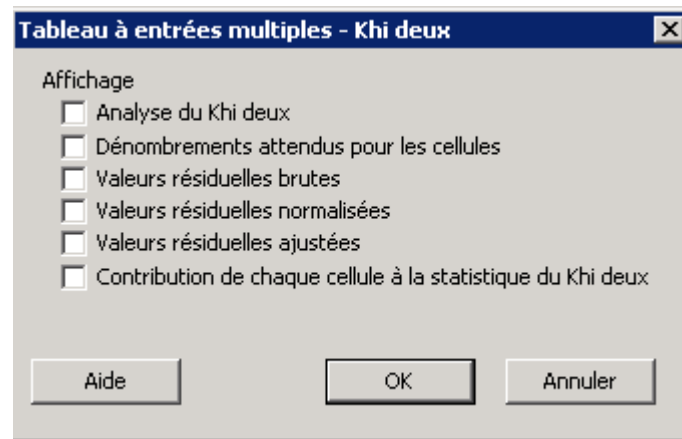
et ensuite nous allons dans le menu **Stat/Tableau à entrées multiples et Khi-deux...**:



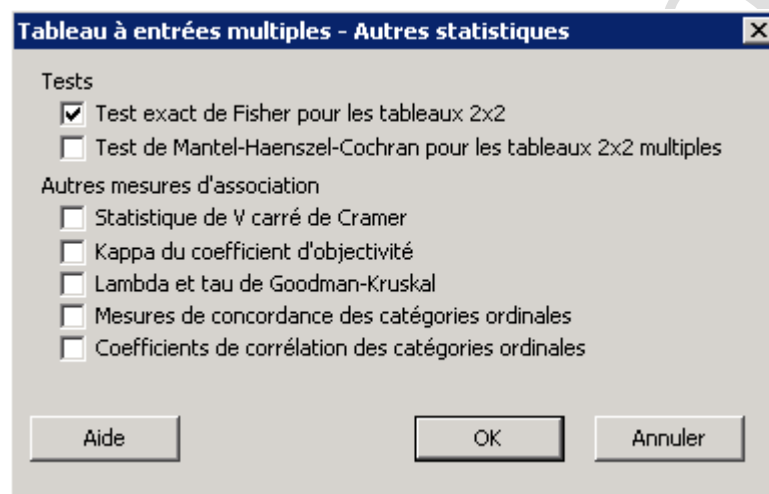
Nous prenons alors:



et dans le bouton **Khi deux...** nous prenons soin de tout décocher:



et dans le bouton **Autres Statistiques...**, nous prenons ce que nous souhaitons contrôler:



et nous validons le tout par **OK** pour obtenir dans la fenêtre de session:

	CP certifié	CP non certifié	Tous
Délais non respectés	4	5	9
	44.44	55.56	100.00
	33.33	83.33	50.00
	22.22	27.78	50.00
Délais respectés	8	1	9
	88.89	11.11	100.00
	66.67	16.67	50.00
	44.44	5.56	50.00
Tous	12	6	18
	66.67	33.33	100.00
	100.00	100.00	100.00
	66.67	33.33	100.00

Contenu de la cellule :  
 Dénombrement  
 % de ligne  
 % de colonne  
 % du total

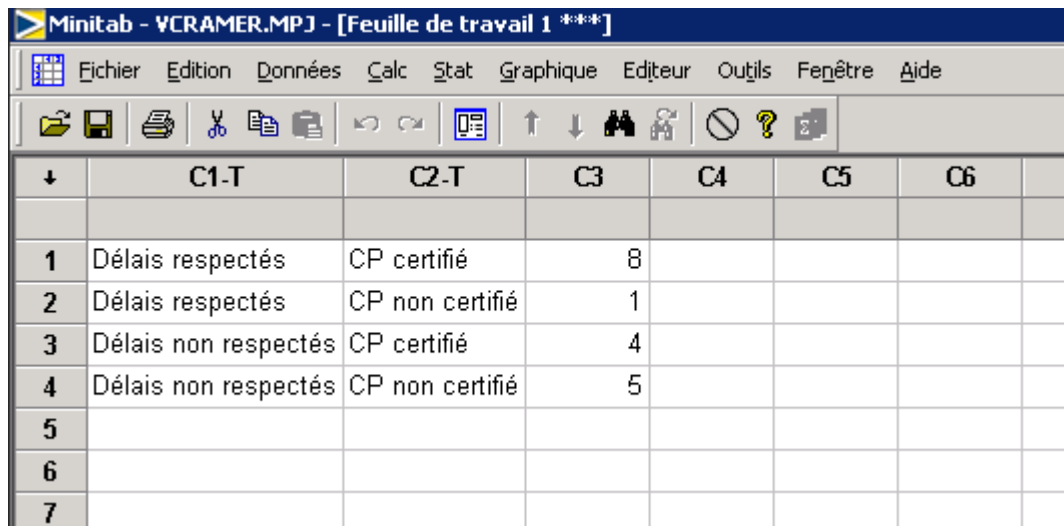
Test exact de Fisher : valeur de P = 0.131222

Nous obtenons donc exactement les mêmes résultats qu'avec Microsoft Excel.

## 10.9. Exercice 40.: V de Cramér (mesure d'association)

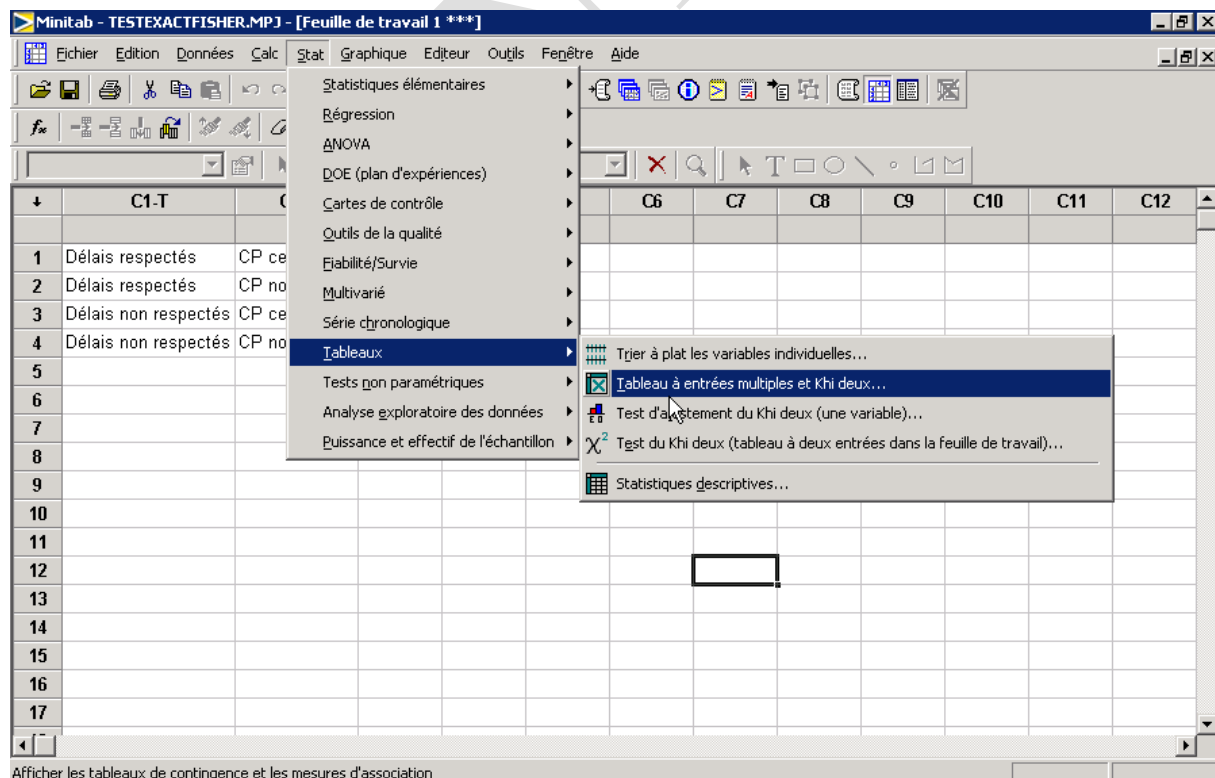
Minitab® Statistical Software 15.1.1.0

Encore une fois, le but est de vérifier les calculs et démonstrations mathématiques faites dans le cours de Méthodes Numériques. Nous partons donc aussi du tableau:



	C1-T	C2-T	C3	C4	C5	C6
1	Délais respectés	CP certifié	8			
2	Délais respectés	CP non certifié	1			
3	Délais non respectés	CP certifié	4			
4	Délais non respectés	CP non certifié	5			
5						
6						
7						

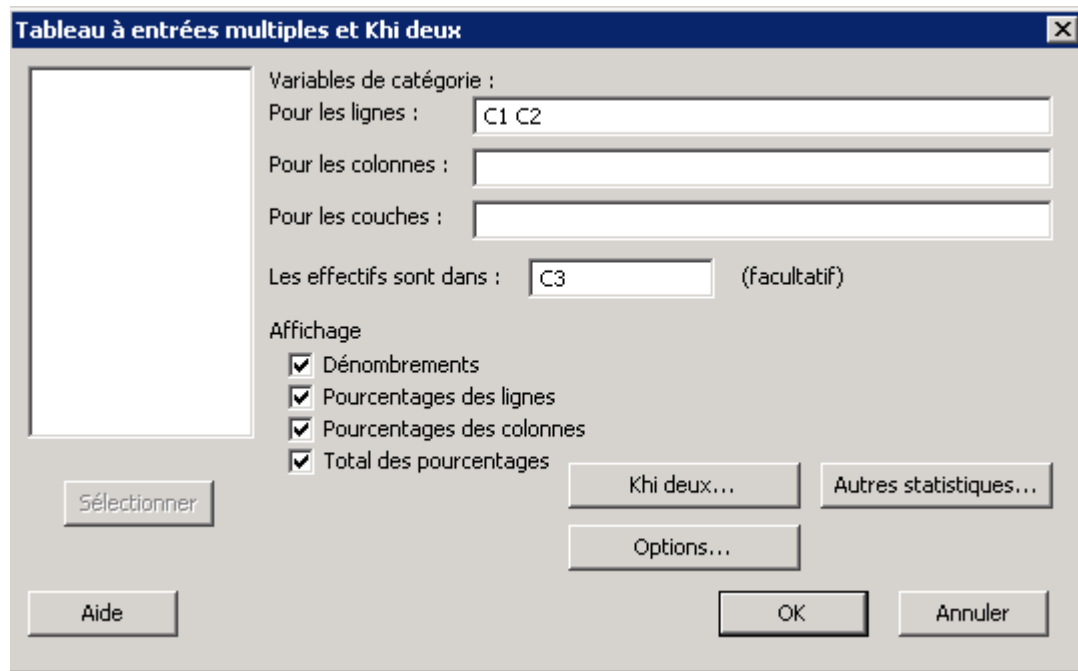
Nous allons dans le menu et ensuite nous allons dans le menu **Stat/Tableau à entrées multiples et Khi-deux...**:



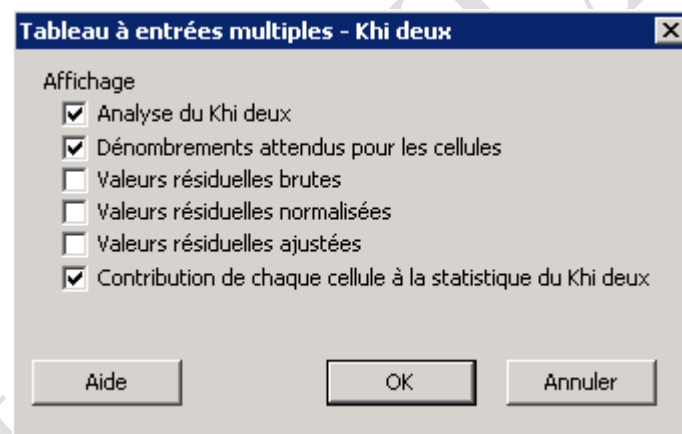
The screenshot shows the Minitab interface with the 'Stat' menu open. The path is: Stat > Tableaux > Tableau à entrées multiples et Khi deux... The 'Tableaux' sub-menu is expanded, showing options like 'Trier à plat les variables individuelles...', 'Tableau à entrées multiples et Khi deux...', 'Test d'ajustement du Khi deux (une variable)...', and 'Test du Khi deux (tableau à deux entrées dans la feuille de travail)...'. The 'Tableau à entrées multiples et Khi deux...' option is highlighted.

et nous prenons un peu tout (déjà nous devons aussi retrouver les mêmes valeurs que le test d'indépendance fait plus haut mais sous une autre forme):

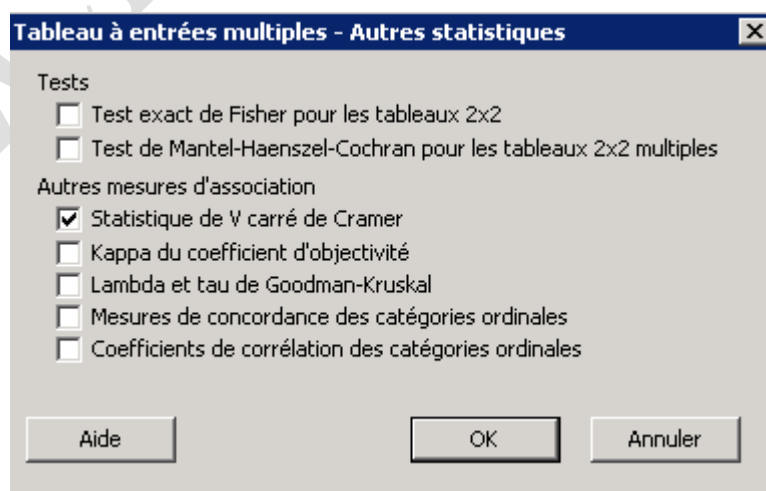




ensuite, nous cliquons sur le bouton **Khi-deux...** pour activer:



et enfin dans le bouton **Autres statistiques...**:



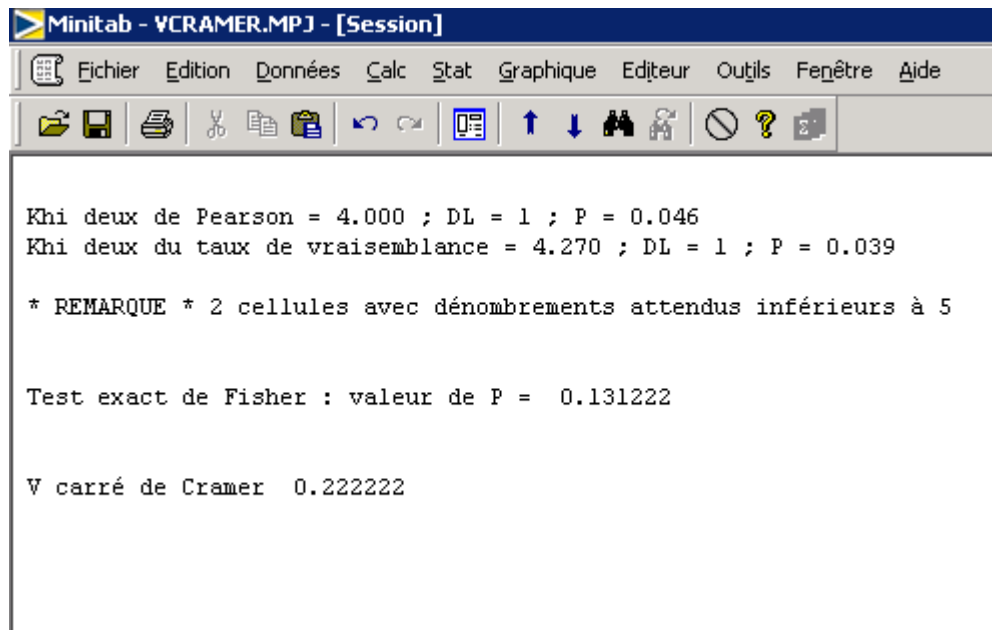
Nous validons tout ce beau monde par **OK** pour obtenir dans la première partie que des valeurs connues et calculées à la main et qui sont conformes à ce que nous avons vu dans le cours théorique mais qui ne contiennent pas le  $V$  de Cramér:

	CP certifié	CP non certifié	Tous
Délais non respectés	4	5	9
	44.44	55.56	100.00
	33.33	83.33	50.00
	22.22	27.78	50.00
	6	3	9
	0.6667	1.3333	*
Délais respectés	8	1	9
	88.89	11.11	100.00
	66.67	16.67	50.00
	44.44	5.56	50.00
	6	3	9
	0.6667	1.3333	*
Tous	12	6	18
	66.67	33.33	100.00
	100.00	100.00	100.00
	66.67	33.33	100.00
	12	6	18
	*	*	*

Contenu de la cellule :  
 Dénombrement  
 % de ligne  
 % de colonne  
 % du total  
 Dénombrement attendu  
 Contribution au Khi deux

Bienvvenue dans Minitab, appuyez sur F1 pour obtenir de l'aide.

et la deuxième partie où nous trouvons ce qui nous intéresse:



```
Minitab - VCRAMER.MPJ - [Session]
Fichier  Edition  Données  Calc  Stat  Graphique  Editeur  Outils  Fenêtre  Aide

Khi deux de Pearson = 4.000 ; DL = 1 ; P = 0.046
Khi deux du taux de vraisemblance = 4.270 ; DL = 1 ; P = 0.039

* REMARQUE * 2 cellules avec dénombrements attendus inférieurs à 5

Test exact de Fisher : valeur de P = 0.131222

V carré de Cramer 0.222222
```

Nous retrouvons donc le Khi-deux de Pearson calculé dans le cours théorique, le DL, la  $p$ -value de 0.046. Ainsi les variables examinées ne sont donc pas indépendantes (la certification a une influence sur les résultats).

Nous n'avons pas vu dans le cours théorique ce qu'est le Khi-deux du taux de vraisemblance donc nous en faisons abstraction pour l'instant (tant que ce ne sera pas au programme).

Nous retrouvons aussi le test exact de Fisher avec sa  $p$ -value déjà obtenue plus haut et aussi calculée dans le cours théorique à la main. Ce test n'est dans le cas présent pas significatif.

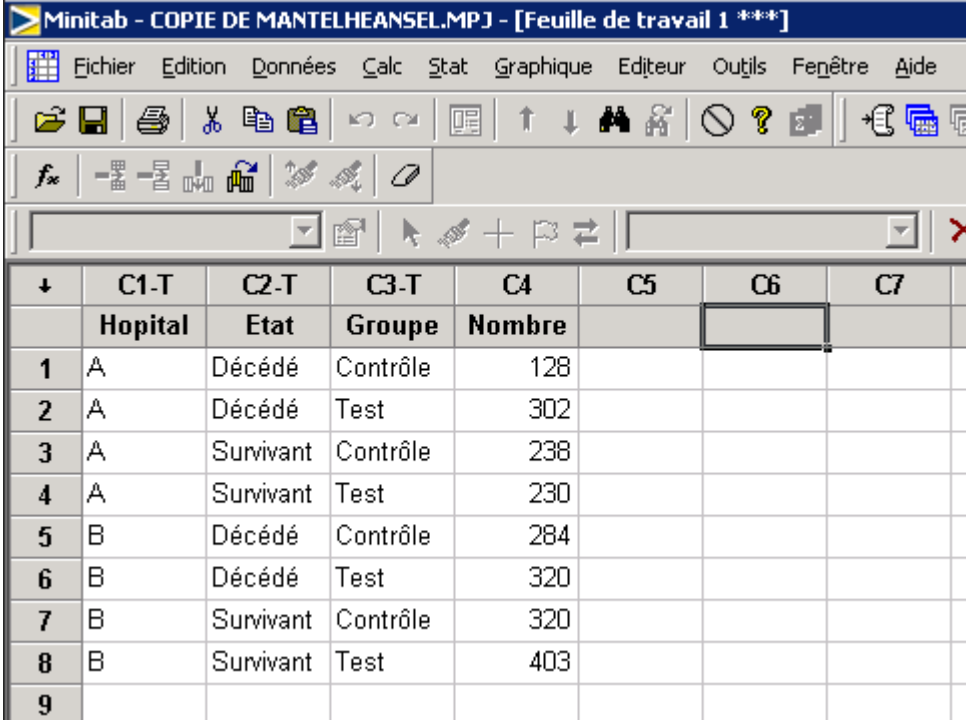
Enfin, nous trouvons le  $V$  de Cramer au carré avec la même valeur que celle calculée dans le cours théorique et qui valide donc la démonstration mathématique. Le  $V$  de Cramer indique ici qu'il y a une relation faible entre les deux facteurs.

## 10.10. Exercice 41.: Test de Mantel-Haenszel-Cochran

Minitab® Statistical Software 15.1.1.0

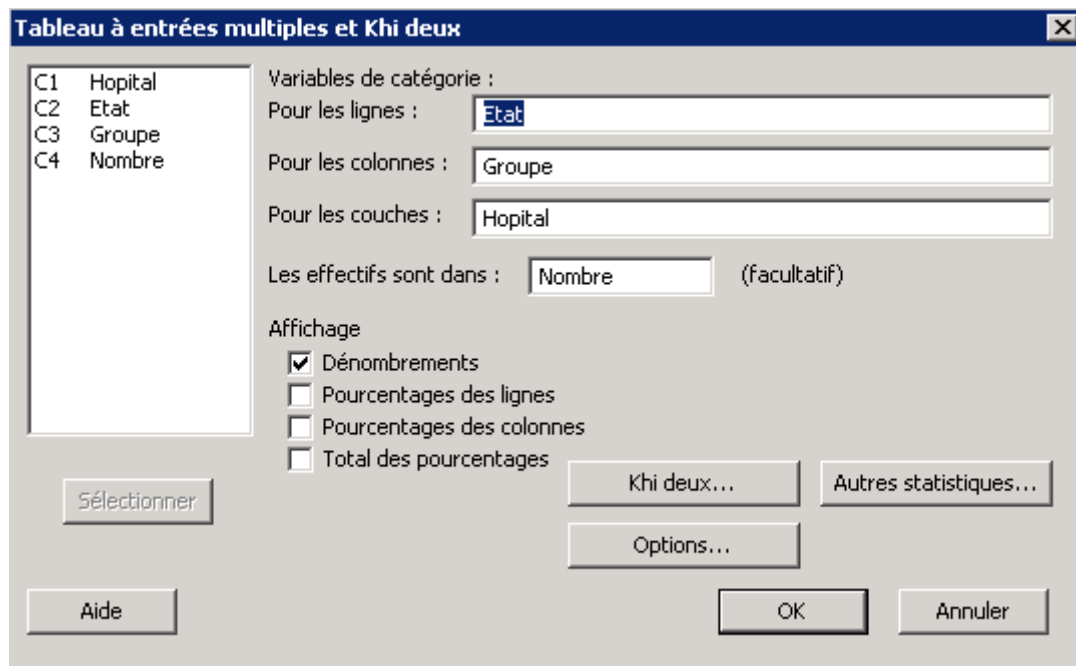
Encore une fois, le but est de vérifier les calculs et démonstrations mathématiques faites dans le cours de Méthodes Numériques.

Nous partons donc des données suivantes (nous avons démontré dans le cours théorique qu'elles ne satisfont pas les conditions pour subir le test MHC mais faisons le quand même...):

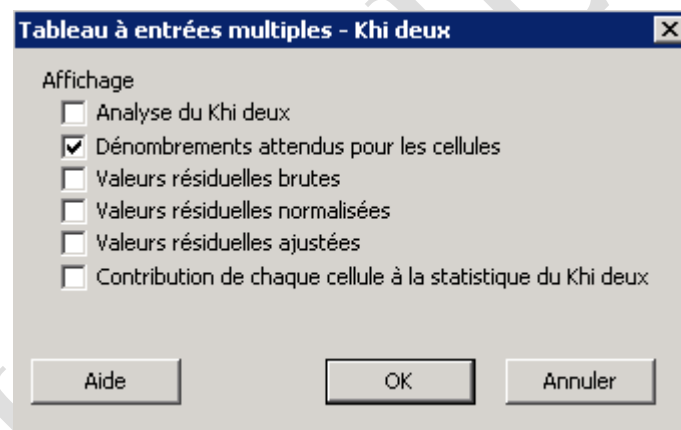


	C1-T	C2-T	C3-T	C4	C5	C6	C7
	Hopital	Etat	Groupe	Nombre			
1	A	Décédé	Contrôle	128			
2	A	Décédé	Test	302			
3	A	Survivant	Contrôle	238			
4	A	Survivant	Test	230			
5	B	Décédé	Contrôle	284			
6	B	Décédé	Test	320			
7	B	Survivant	Contrôle	320			
8	B	Survivant	Test	403			
9							

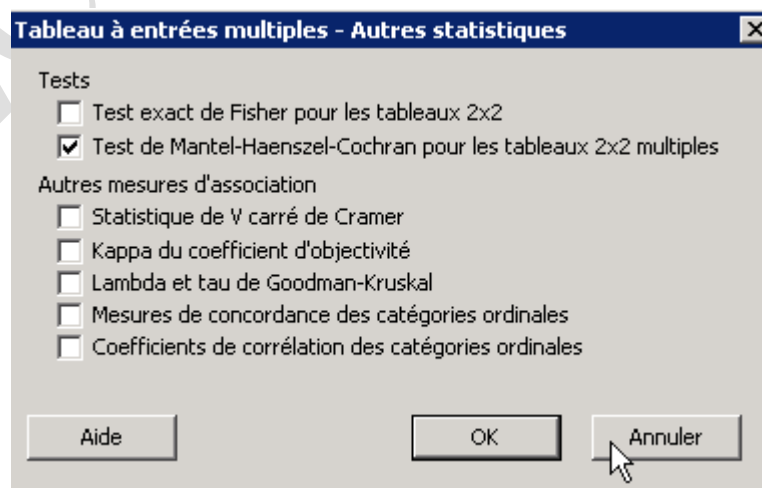
et allons dans le menu **Stat/Tableau à entrées multiples et Khi-deux...**



Dans le bouton **Khi deux...** nous prenons pour avoir l'essentiel (le minimum) tel que calculé manuellement dans le cours théorique:



Dans le bouton **Autres Statistiques...** nous prenons évidemment ce qui nous intéresse ici:



et nous validons le tout par **OK** afin d'obtenir au final dans la fenêtre de session d'abord pour l'Hôpital A:

### Tableau de statistiques : Etat; Groupe; Hopital

En utilisant les effectifs dans Nombre

#### Résultats pour Hopital = A

Lignes : Etat	Colonnes : Groupe		
	Contrôle	Test	Tous
Décédé	128	302	430
	175.3	254.7	430.0
Survivant	238	230	468
	190.7	277.3	468.0
Tous	366	532	898
	366.0	532.0	898.0

Contenu de la cellule :      Dénombrement  
Dénombrement attendu

ce qui correspond bien à ce qui a été calculé dans le cours théorique manuellement. Pour l'Hôpital B:

#### Résultats pour Hopital = B

Lignes : Etat	Colonnes : Groupe		
	Contrôle	Test	Tous
Décédé	284	320	604
	274.9	329.1	604.0
Survivant	320	403	723
	329.1	393.9	723.0
Tous	604	723	1327
	604.0	723.0	1327.0

Contenu de la cellule :      Dénombrement  
Dénombrement attendu

ce qui correspond aussi bien à ce qui a été calculé dans le cours théorique manuellement.

Et finalement:

#### Résultats pour tous les tableaux 2 x 2

Rapport des probabilités de succès commun 0.757173

Statistiques		
de MHC	DL	P
10.4485	1	0.0012275

où comme nous l'avions mentionné dans le cours théorique la correction de continuité est effectivement ajoutée puisque nous retrouvons parfaitement les mêmes résultats. Au niveau de la conclusion, la  $p$ -value étant beaucoup plus petite que les valeurs traditionnelles critiques (10%, 5%, 1%) nous mettons donc en évidence le fait qu'il y a une différence significative entre le groupe de contrôle et de test à travers les différentes strates.

Maintenant nous pouvons aussi faire la même conclusion en observant les deux strates en % :

**Résultats pour Hopital = A**

Lignes : Etat    Colonnes : Groupe

	Contrôle	Test	Total
Décédé	14,25	33,63	47,88
Survivant	26,50	25,61	52,12
Total	40,76	59,24	100,00

Contenu de la cellule :            % du total

**Résultats pour Hopital = B**

Lignes : Etat    Colonnes : Groupe

	Contrôle	Test	Total
Décédé	21,40	24,11	45,52
Survivant	24,11	30,37	54,48
Total	45,52	54,48	100,00

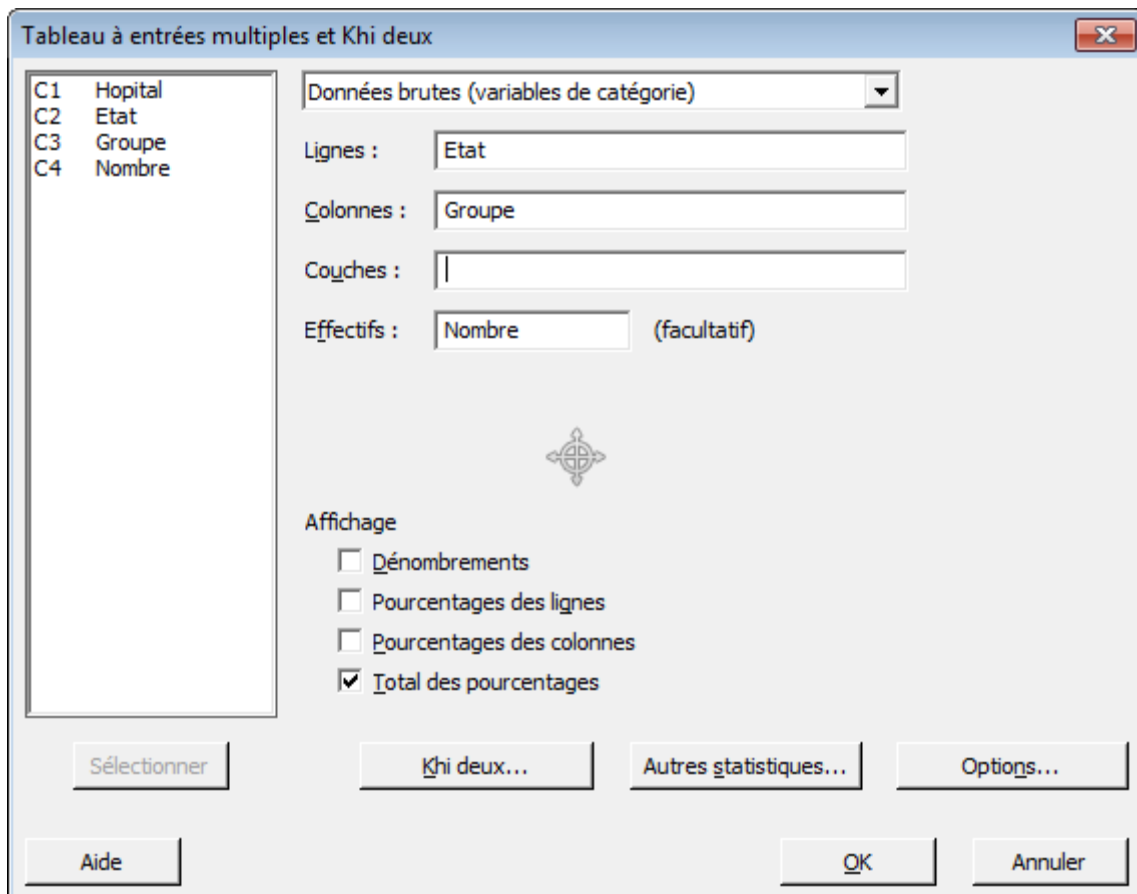
Contenu de la cellule :            % du total

**Résultats pour tous les tableaux 2 x 2**

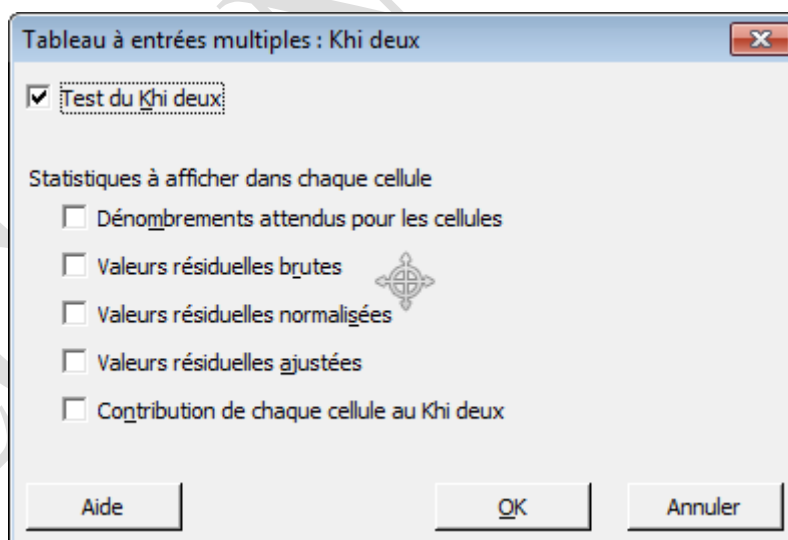
Rapport des probabilités de succès commun 0,757173

Statistiques de MHC	DL	Valeur de P
10,4485	1	0,0012275

Maintenant vient la fameuse question: Et si nous réunissons les deux tableaux?:

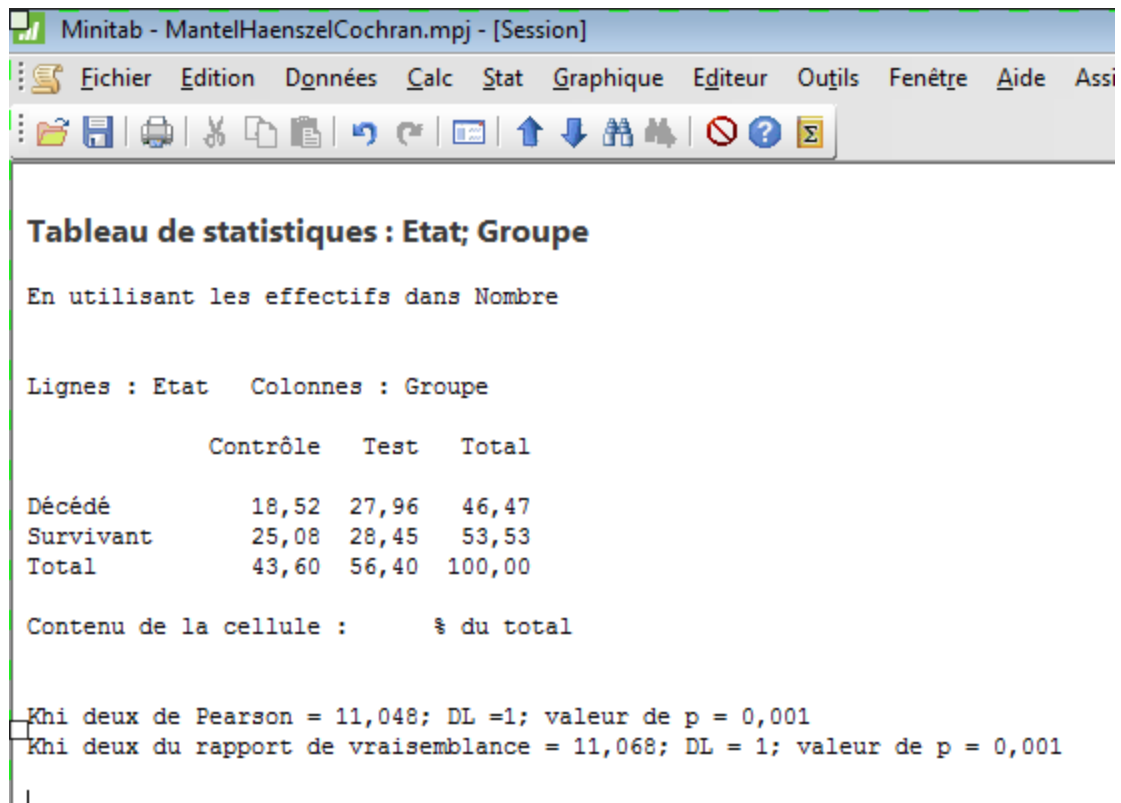


En prenant un test du Khi-deux:



Nous avons alors:





Minitab - MantelHaenszelCochran.mpj - [Session]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assi

Tableau de statistiques : Etat; Groupe

En utilisant les effectifs dans Nombre

Lignes : Etat    Colonnes : Groupe

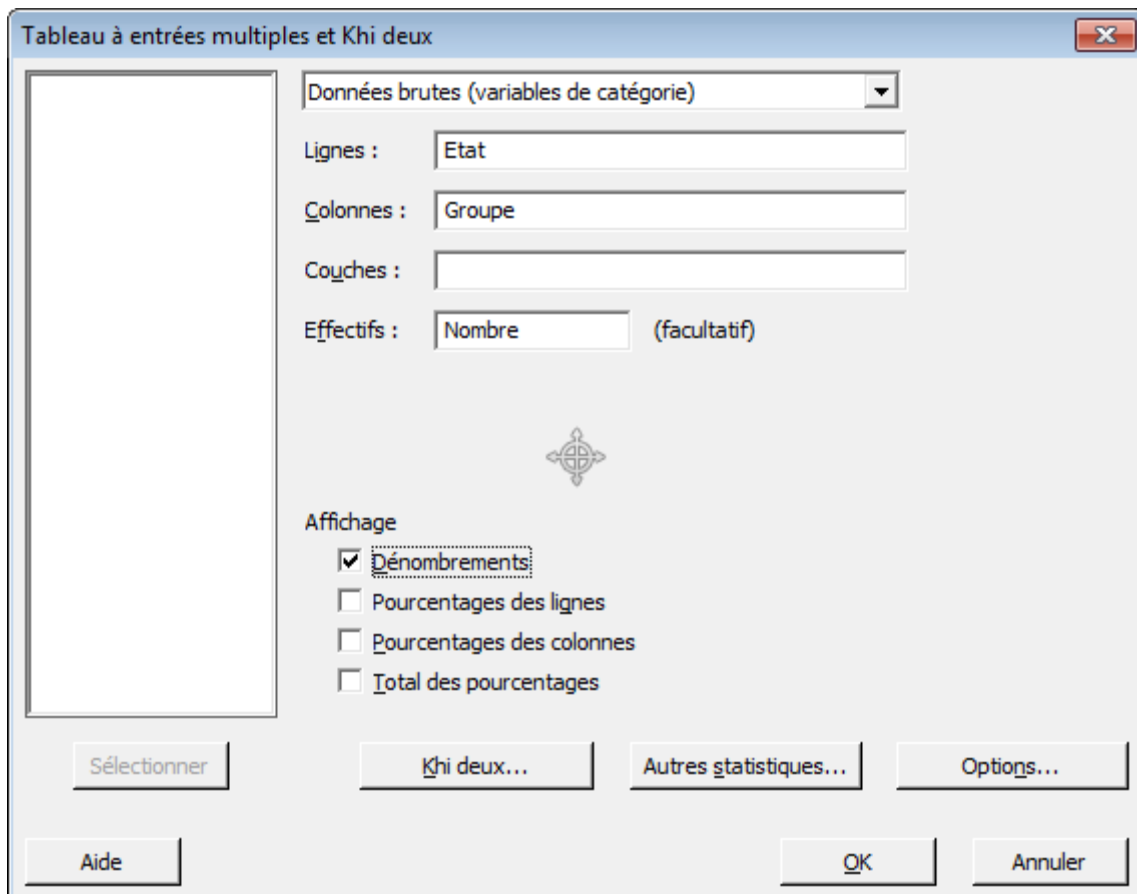
	Contrôle	Test	Total
Décédé	18,52	27,96	46,47
Survivant	25,08	28,45	53,53
Total	43,60	56,40	100,00

Contenu de la cellule :            % du total

Khi deux de Pearson = 11,048; DL =1; valeur de p = 0,001  
Khi deux du rapport de vraisemblance = 11,068; DL = 1; valeur de p = 0,001

Donc le test d'indépendance du Khi-deux nous dit aussi qu'il y a une différence significative même lorsque les tableaux sont réunis!

Ou encore en faisant un tableau de dénombrement unique:



...qui donne:

Minitab - MantelHaenszelCochran.mpj - [Session]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assist

**Tableau de statistiques : Etat; Groupe**

En utilisant les effectifs dans Nombre

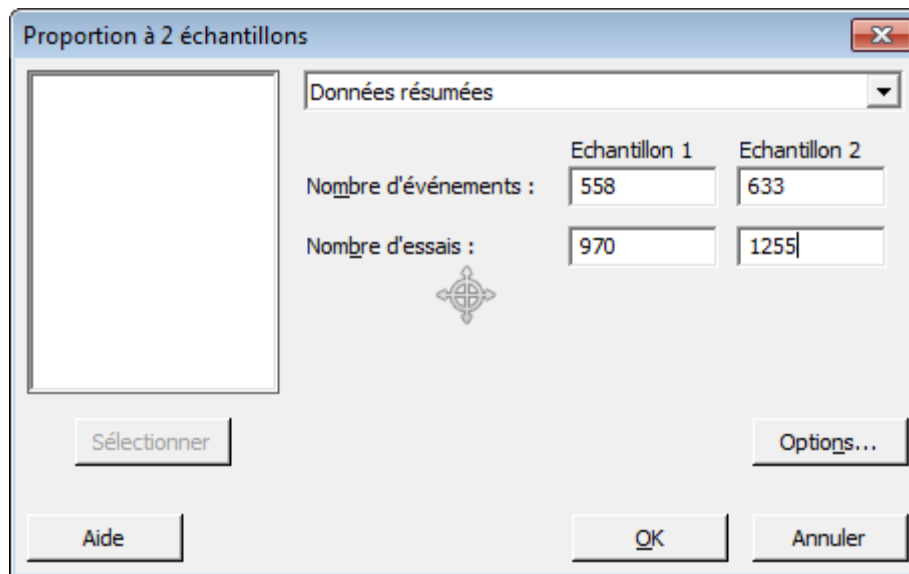
Lignes : Etat Colonnes : Groupe

	Contrôle	Test	Total
Décédé	412	622	1034
Survivant	558	633	1191
Total	970	1255	2225

Contenu de la cellule : Dénombrement

Khi deux de Pearson = 11,048; DL =1; valeur de p = 0,001  
 Khi deux du rapport de vraisemblance = 11,068; DL = 1; valeur de p = 0,001

Et ensuite nous allons pouvoir à l'aide de ces valeurs faire un test de la différence de deux proportions:



Ce qui nous indique bien que les groupes de *Test* et de *Contrôle* sont significativement différents:

**Test et IC pour 2 proportions**

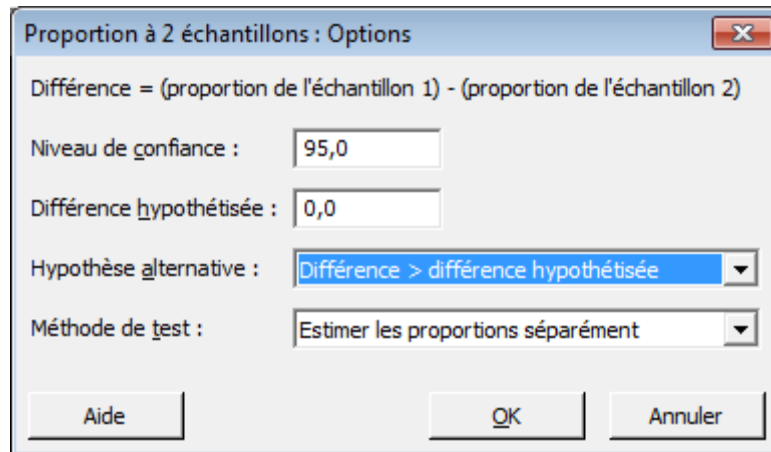
Echantillon	X	N	P échantillon
1	558	970	0,575258
2	633	1255	0,504382

Différence =  $p(1) - p(2)$   
 Estimation de la différence : 0,0708753  
 IC à 95 % pour la différence : (0,0292482; 0,112502)  
 Test de la différence = 0 (et  $\neq$  0) :  $Z = 3,34$   $P = 0,001$

Test exact de Fisher : valeur de  $p = 0,001$

Feuille de travail en cours : TestMHC

Donc il y a bien une différence entre les deux groupes de contrôles mais nous ne savons pas dans quel sens! Or ce qui nous intéresse c'est de savoir si le groupe de contrôle a un meilleur taux de survie que le groupe de test. Dès lors:



Ce qui donne:

Minitab - MantelHaenszelCochran.mpj - [Session]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assistant

### Test et IC pour 2 proportions

Echantillon	X	N	P échantillon
1	558	970	0,575258
2	633	1255	0,504382

Différence =  $p(1) - p(2)$   
Estimation de la différence : 0,0708753  
Borne inférieure (au seuil de 95 %) pour la différence : 0,0359408  
Test de la différence = 0 (fn de > 0) :  $Z = 3,34$   $P = 0,000$

Test exact de Fisher : valeur de  $p = 0,001$

Feuille de travail en cours : TestMHC Modifiable

Et... oui la différence est là aussi significative!

## 10.11. Exercice 42.: Test d'ajustement du Khi-deux

Minitab® Statistical Software 15.1.1.0

Supposons qu'un chercheur tente de déterminer s'il y a ou non une différence statistiquement significative entre le nombre de naissances et le jour de la semaine. Il serait possible d'utiliser ce résultat pour, par exemple, planifier le personnel et l'équipement d'un hôpital.

Supposons que les naissances à un hôpital, pour une certaine période de temps, se répartissent comme suit:

Jour	Lundi	Mardi	Mercredi	Jeudi	Vendredi	Samedi	Dimanche	Total
Fréquence Observées	120	130	125	128	80	70	75	728

Puisqu'il y a au total 728 naissances pour les 7 jours en théorie il devrait y avoir  $728/7=104$  naissances à chaque jour. Nous avons donc maintenant le tableau suivant:

Jour	Lundi	Mardi	Mercredi	Jeudi	Vendredi	Samedi	Dimanche	Total
Fréquence Observée	120	130	125	128	80	70	75	728
Fréquence Théorique	104	104	104	104	104	104	104	728

Nous voulons examiner s'il y a une différence entre les fréquences observées et les fréquences théoriques en nous basant sur un test d'ajustement du Khi-deux (cela revient à comparer une loi de distribution expérimentale à une loi de distribution uniforme). Nous avons déjà fait la démonstration et le calcul à la main dans le cours théorique. Voyons ce que nous obtenons avec Minitab.

Nous partons de:

↓	C1-T	C2
	Jour	Observé
1	Lundi	120
2	Mardi	130
3	Mercredi	125
4	Jeudi	128
5	Vendredi	80
6	Samedi	70
7	Dimanche	75

Nous allons dans le menu **Stat/Tableaux/Test d'ajustement du Khi-deux...**:

The screenshot shows the Minitab interface with the 'Stat' menu open. The 'Tableaux' option is selected, and the submenu is displayed, showing 'Test d'ajustement du Khi deux (une variable)...' as the selected option. The background shows a worksheet with columns C1-T, C2, and C3, and rows 1-10.

	C1-T	C2	C3
	Jour	Observé	
1	Lundi	120	
2	Mardi	130	
3	Mercredi	125	
4	Jeudi	128	
5	Vendredi	80	
6	Samedi	70	
7	Dimanche	75	
8			
9			
10			

et nous mettons:

The screenshot shows the 'Test d'ajustement du Khi deux' dialog box. The 'Dénombrements observés' option is selected, and the 'Observé' field is filled with 'Observé'. The 'Noms de catégories (facultatif)' field is filled with 'Jour'. The 'Test' section has 'Proportions égales' selected. The 'Sélectionner' button is visible, and the 'OK' and 'Annuler' buttons are at the bottom right.

Remarquez que si nous avons créé une colonne avec les données théoriques à comparer, il aurait suffi de cliquer sur **Proportions spécifiques**! Nous validons par **OK** pour obtenir dans la fenêtre de session:

## Test d'ajustement du Khi deux pour les dénombrements observés dans la variable

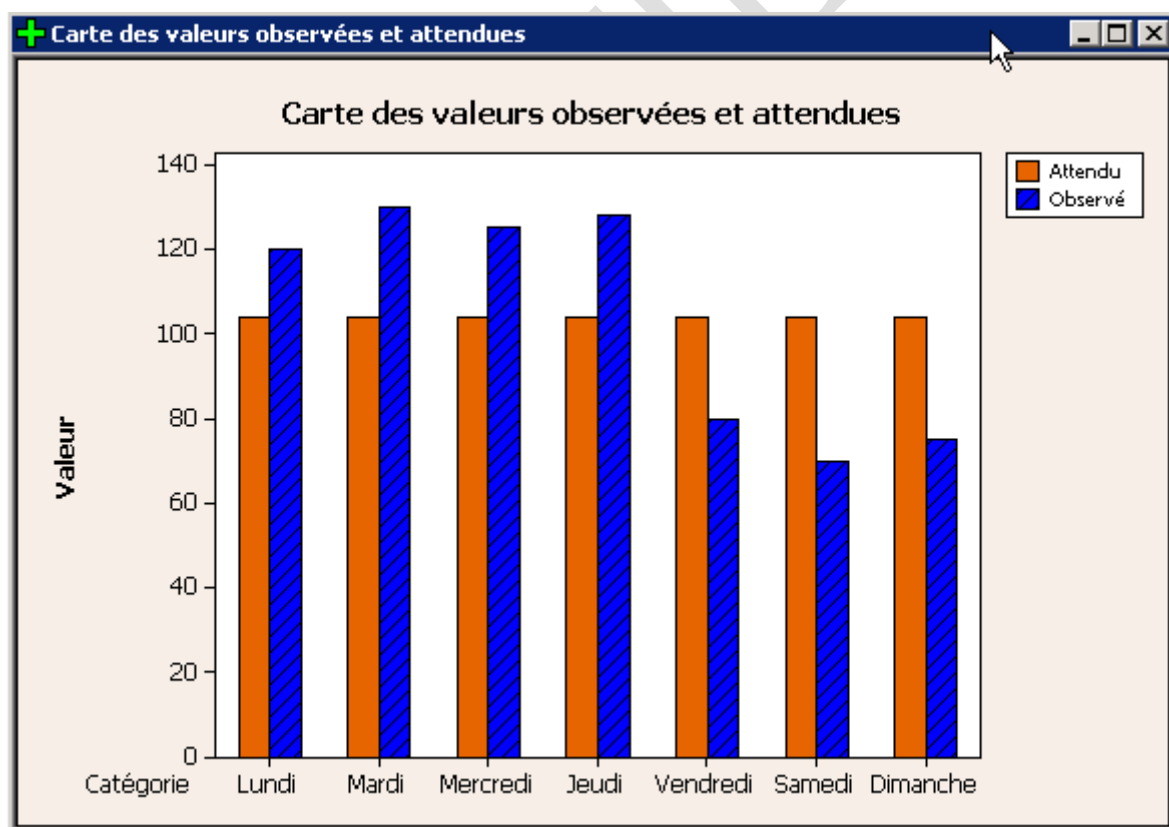
Utilisation des noms de catégories dans Jour

Catégorie	Observé	Proportion testée	Attendu	Contribution au Khi deux
Lundi	120	0.142857	104	2.4615
Mardi	130	0.142857	104	6.5000
Mercredi	125	0.142857	104	4.2404
Jeudi	128	0.142857	104	5.5385
Vendredi	80	0.142857	104	5.5385
Samedi	70	0.142857	104	11.1154
Dimanche	75	0.142857	104	8.0865

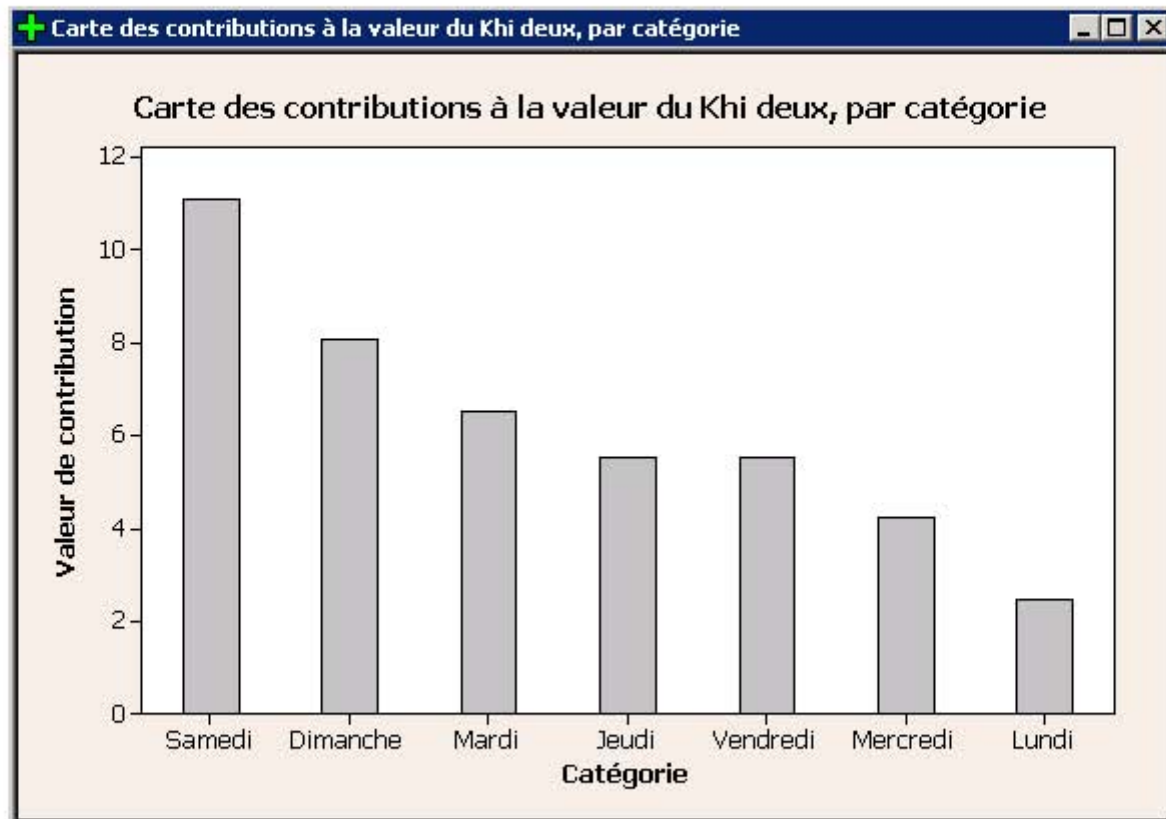
N	DL	Khi deux	Valeur de p
728	6	43.4808	0.000

Il s'agit donc exactement des valeurs obtenues que celles obtenues dans le cours théorique avec la même conclusion étant donnée la  $p$ -value.

Nous obtenons aussi les graphiques suivants:



et:





## 10.12. Exercice 43.: Ajustement d'une loi de Poisson par le Khi-deux

Minitab® Statistical Software 15.1.1.0

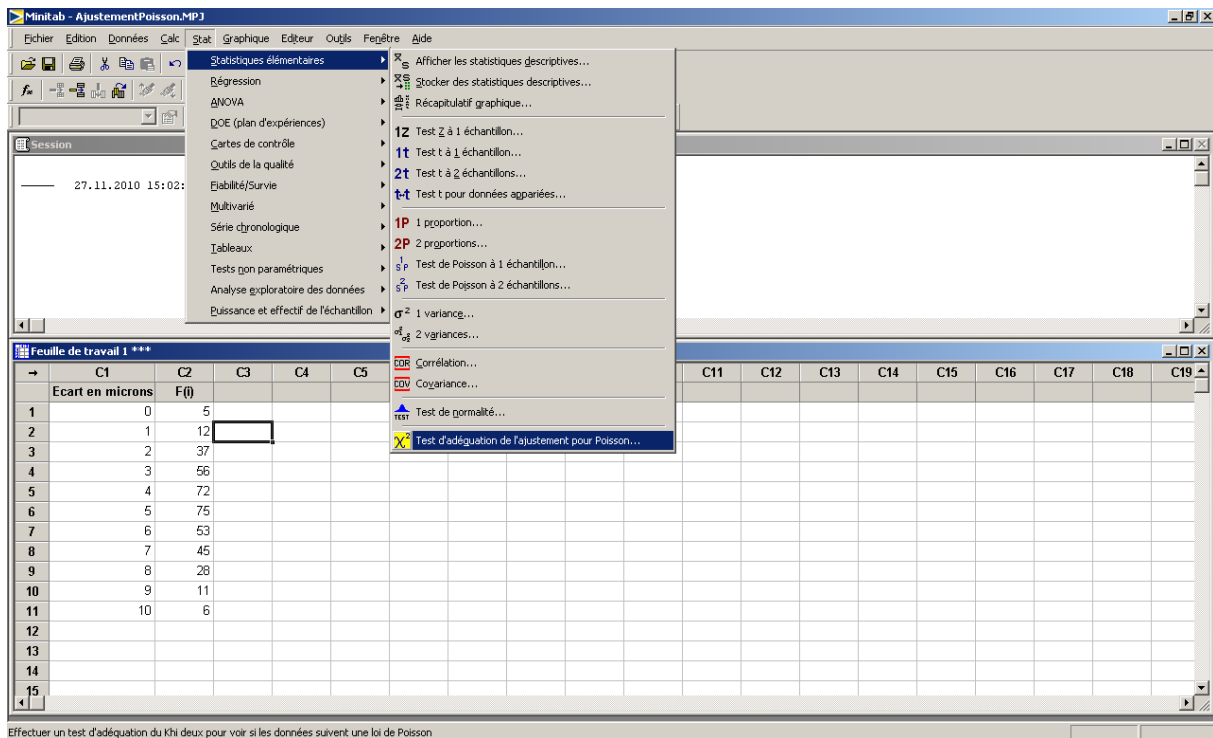
Ouvrez le fichier *Mesures.mpj*:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19
	Ecart en microns	F(i)																	
1	0	5																	
2	1	12																	
3	2	37																	
4	3	56																	
5	4	72																	
6	5	75																	
7	6	53																	
8	7	45																	
9	8	28																	
10	9	11																	
11	10	6																	
12																			
13																			
14																			
15																			

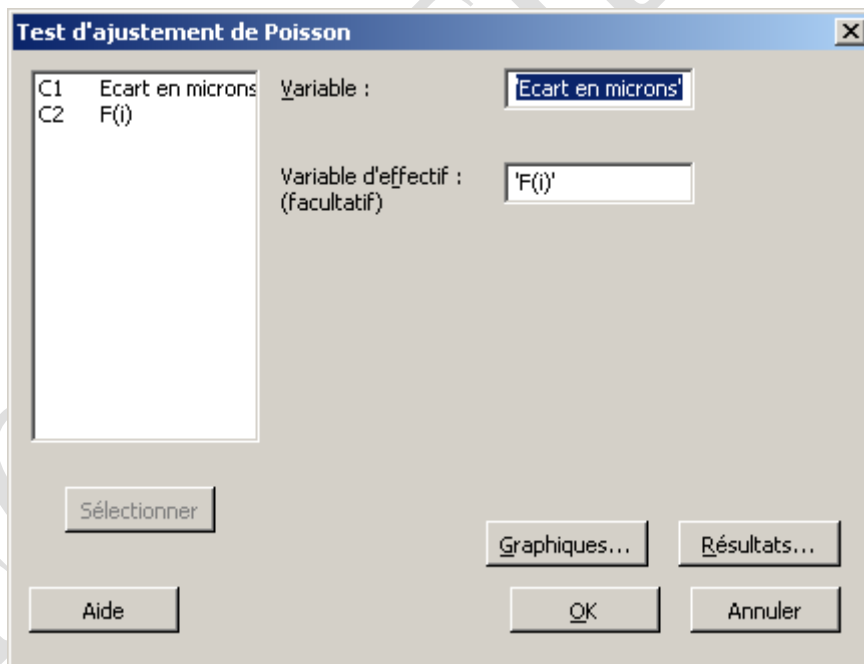
Il s'agit d'un tableau de fréquences de pièces ayant des écarts en micros et il s'agit exactement du même tableau que dans le cours de statistique théorique avec Microsoft Excel.

Nous souhaiterions faire la même analyse mais cette fois-ci avec Minitab® Statistical Software. Il faudra remarquer que contrairement au cours Microsoft Excel où nous avons imposé une espérance de 5 microns pour la moyenne, Minitab® Statistical Software ne nous permet pas de choisir et calcule l'estimateur expérimental de la moyenne. Il y a aura donc une petite différence dans les résultats des calculs qui sera tout à fait normale.

Nous allons dans le menu **Stat/Statistiques élémentaires/Test d'adéquation de l'ajustement du Khi-deux**:



Il vient alors avec les paramètres:



Nous obtenons alors dans la fenêtre de session:

27.11.2010 15:02:07

**Test d'ajustement de la loi de Poisson**

Colonne de données : Ecart en microns  
Colonne des fréquences : F(1)

Moyenne de Poisson pour Ecart en microns = 4.8325

Ecart en microns	Observé	Probabilité de Poisson	Attendu	Contribution au Khi deux
0	5	0.007967	3.1866	1.03191
1	12	0.038498	15.3994	0.75041
2	37	0.093022	37.2088	0.00117
3	56	0.149843	59.9372	0.25863
4	72	0.181029	72.4116	0.00234
5	75	0.174965	69.9858	0.35925
6	53	0.140919	56.3677	0.20121
7	45	0.097285	38.9139	0.95187
8	28	0.058766	23.5064	0.85902
9	11	0.031554	12.6216	0.20835
>=10	6	0.026152	10.4609	1.90232

N	N*	DL	Khi deux	Valeur de p
400	0	9	6.52646	0.686

1 cellule(s) (9.09 %) avec valeur(s) attendue(s) inférieure(s) à 5.

**Carte des valeurs observées et attendues**

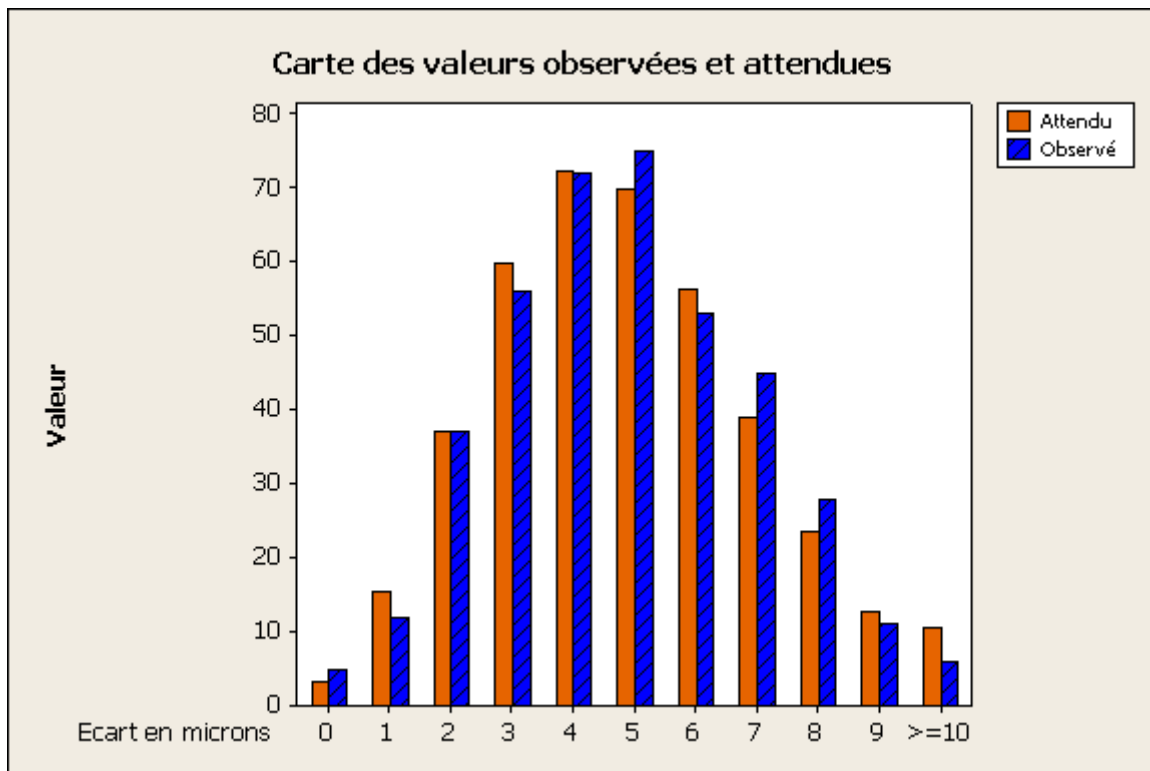
**Carte des contributions à la valeur du Khi deux, par catégorie**

Bienvenue dans Minitab, appuyez sur F1 pour obtenir de l'aide.

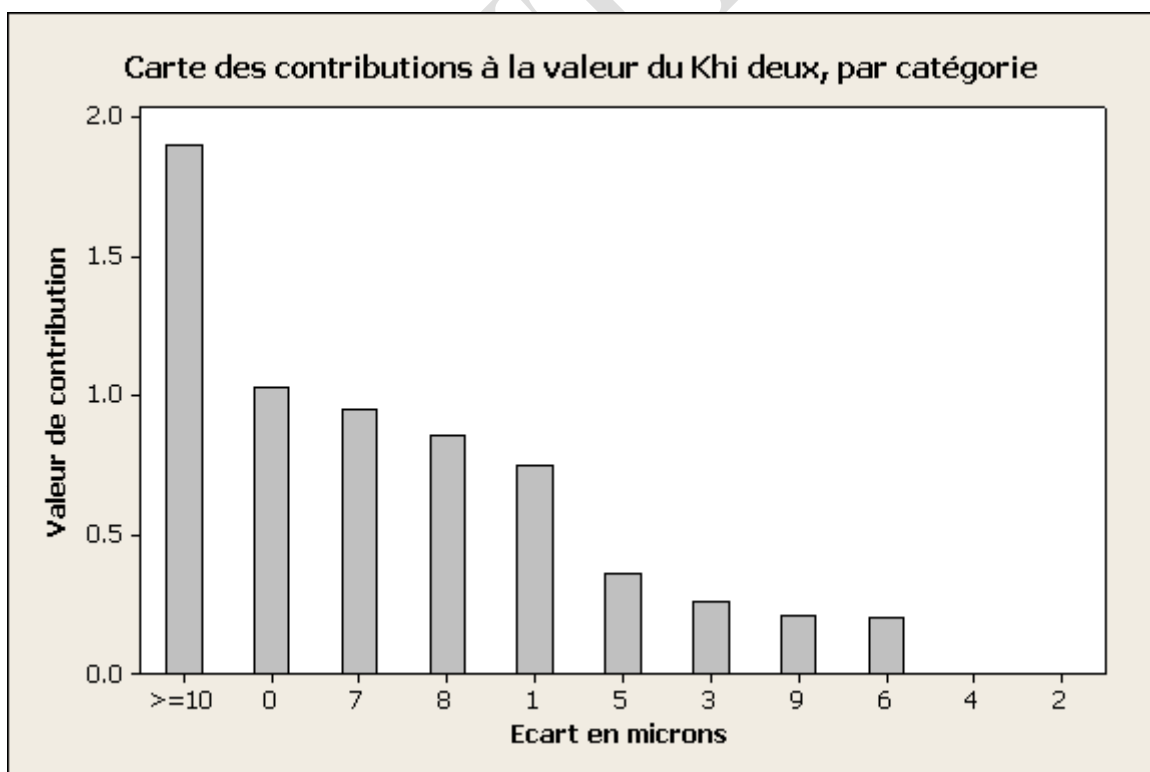
Des valeurs similaires à celles obtenues avec Microsoft Excel avec juste de petites variations car comme nous pouvons le voir, il a utilisé la moyenne expérimentale de 4.8325 au lieu de la moyenne de 5 imposée.

Nous remarquons que Minitab® Statistical Software fait également bien les choses puisqu'il nous avertit qu'une cellule a une valeur attendue inférieure à 5 (la première dans le tableau) ce qui est non recommandé.

Minitab® Statistical Software nous donne également le même graphique que celui que nous avons fait avec Microsoft Excel:



Et un autre graphique que nous n'avons pas fait dans le cours de statistique théorique (nous avons fait par contre le tableau correspondant):

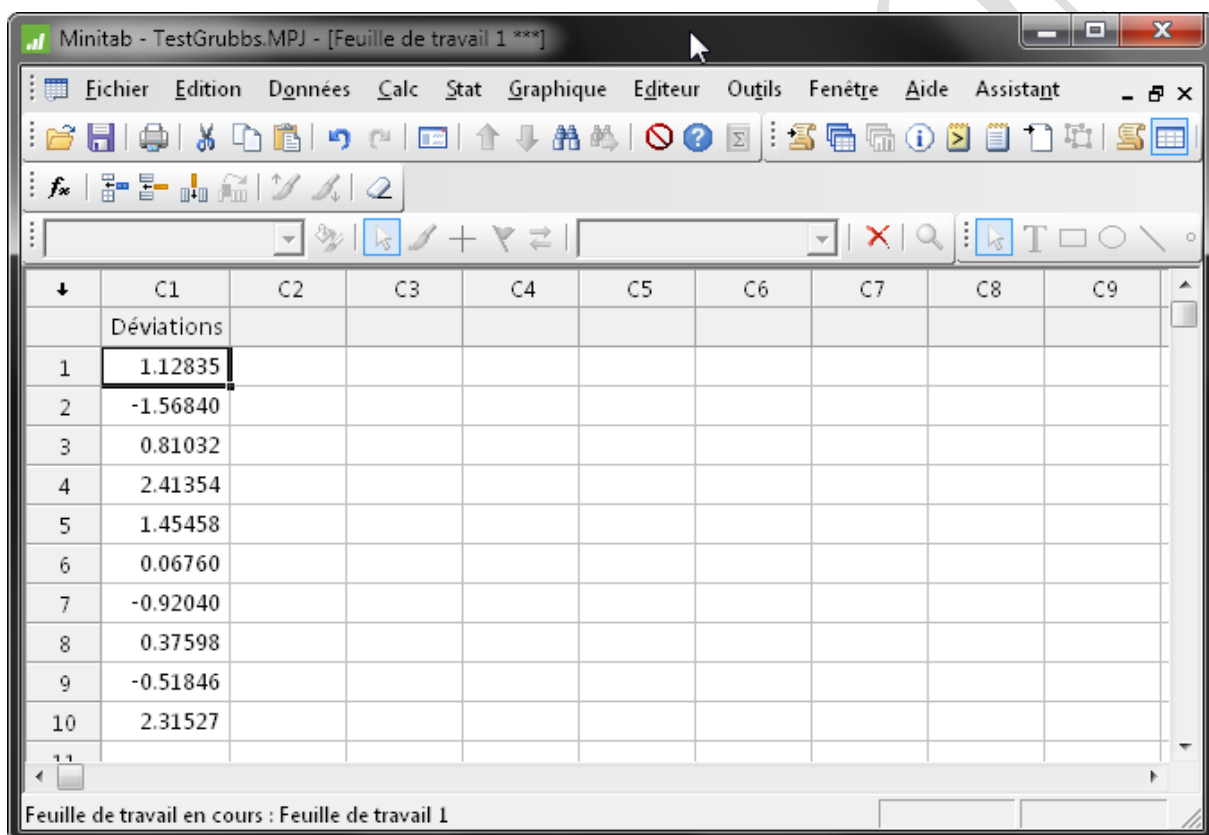


## 10.13. Exercice 44.: Test de Grubbs (test des valeurs aberrantes de Grubbs)

Minitab® Statistical Software 17.1.0

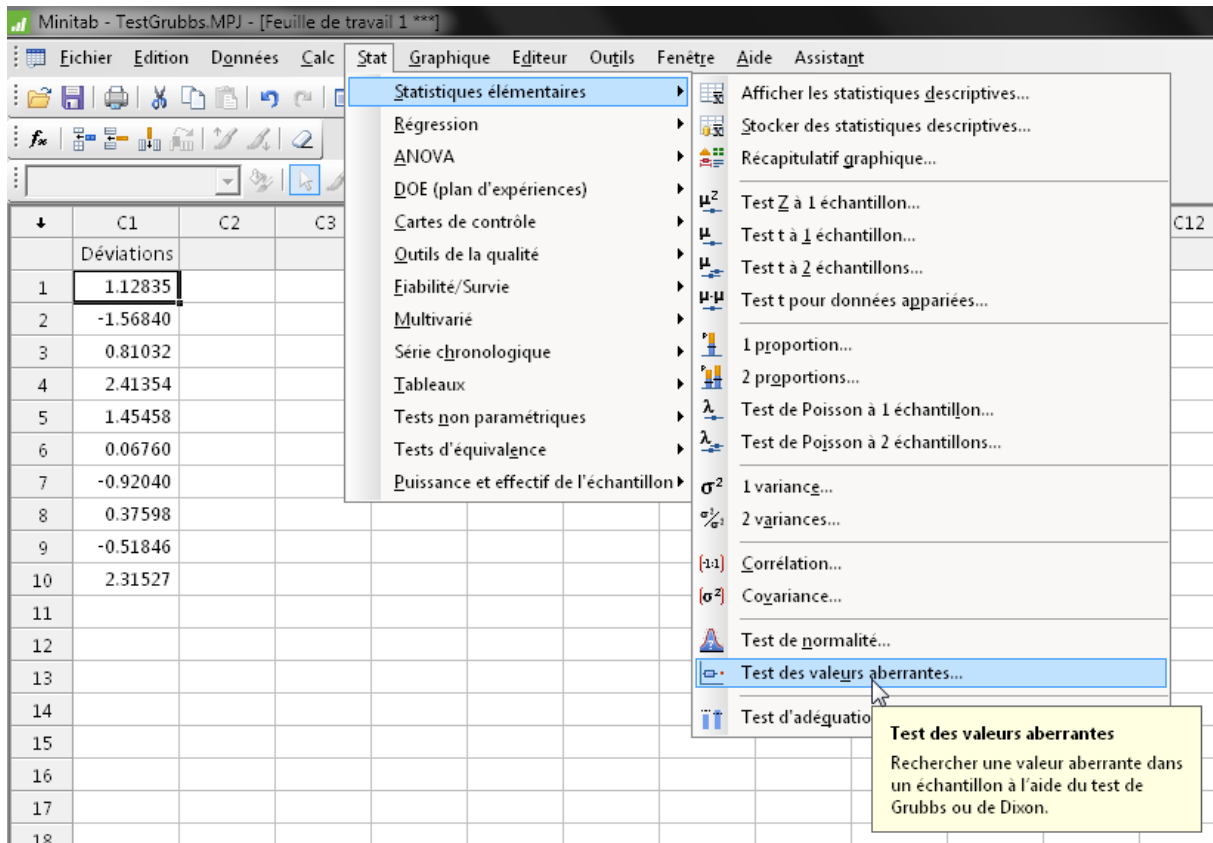
Le but ici va être de vérifier avec la nouveauté de la version 17 de Minitab la validité du test de Grubbs qui y est implémenté que nous avons vu dans le cours théorique et dont nous avons appris à calculer les valeurs critiques à la main pour n'importe quelle distribution.

Nous allons prendre la série de valeurs suivante (et cela n'est pas explicite dans Minitab mais nous allons de suite voir sur quelle distribution théorique implicite il se base pour calculer les valeurs aberrantes):

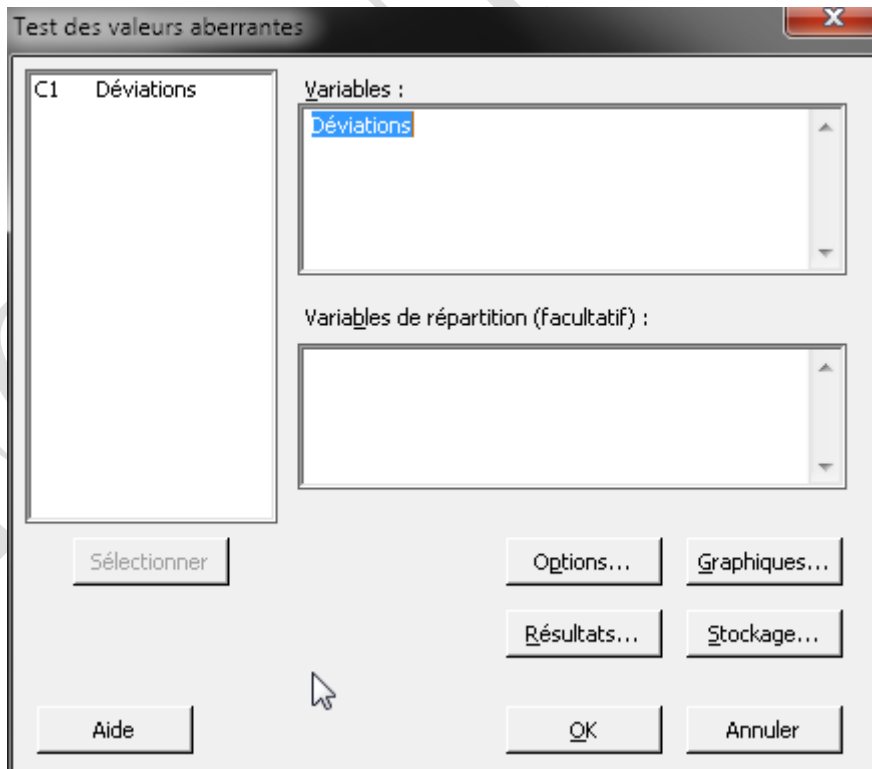


	C1	C2	C3	C4	C5	C6	C7	C8	C9
	Déviations								
1	1.12835								
2	-1.56840								
3	0.81032								
4	2.41354								
5	1.45458								
6	0.06760								
7	-0.92040								
8	0.37598								
9	-0.51846								
10	2.31527								

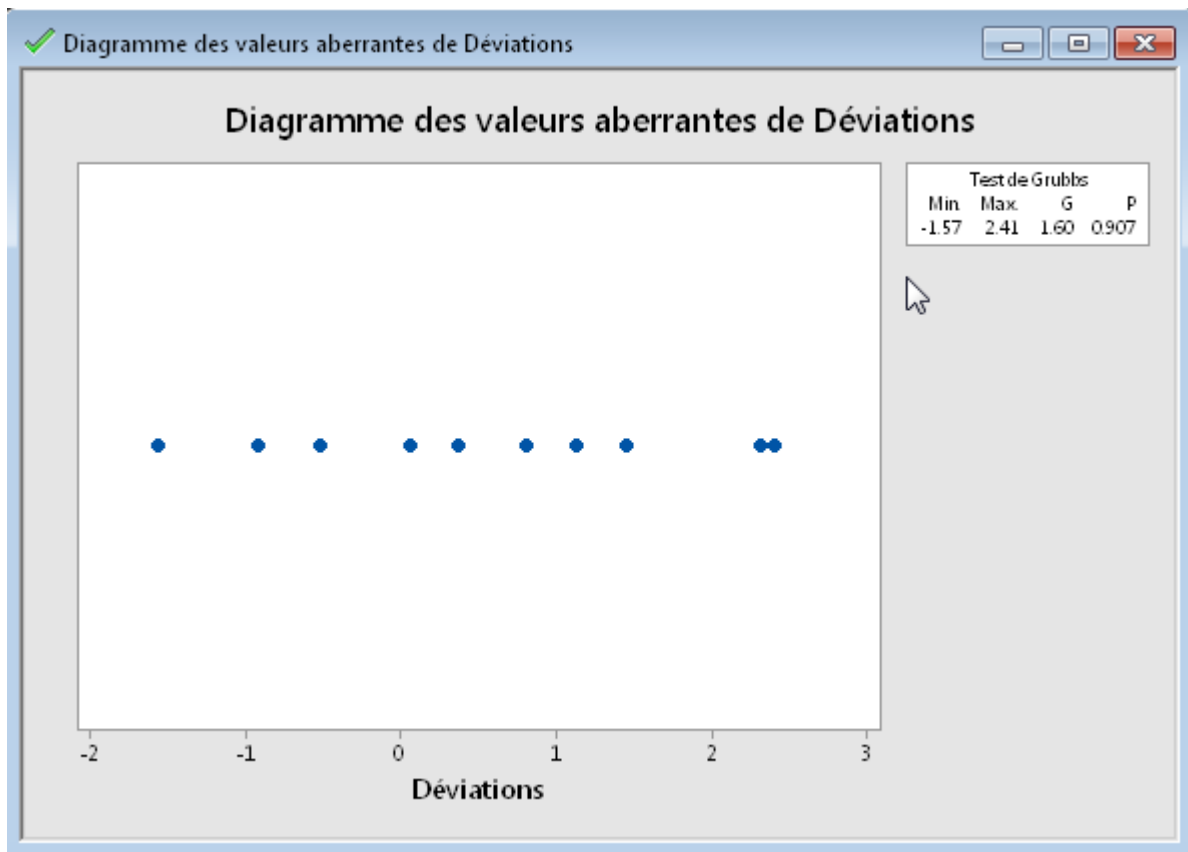
Et nous allons dans le menu **Stat/Statistiques élémentaires/Test des valeurs aberrantes...**:



et nous prenons:



Ce qui donne:



Nous retrouvons donc exactement les mêmes valeurs de G et la même p value que celles calculées à la main dans le cours théorique en se basant sous l'hypothèse que la distribution est Normale! Donc là c'est un peu dangereux que ce ne soit pas indiqué explicitement par Minitab que le calcul de G (et donc la p-value correspondante) se font sous l'hypothèse que la distribution sous-jacente soit Normale. Dommage...!

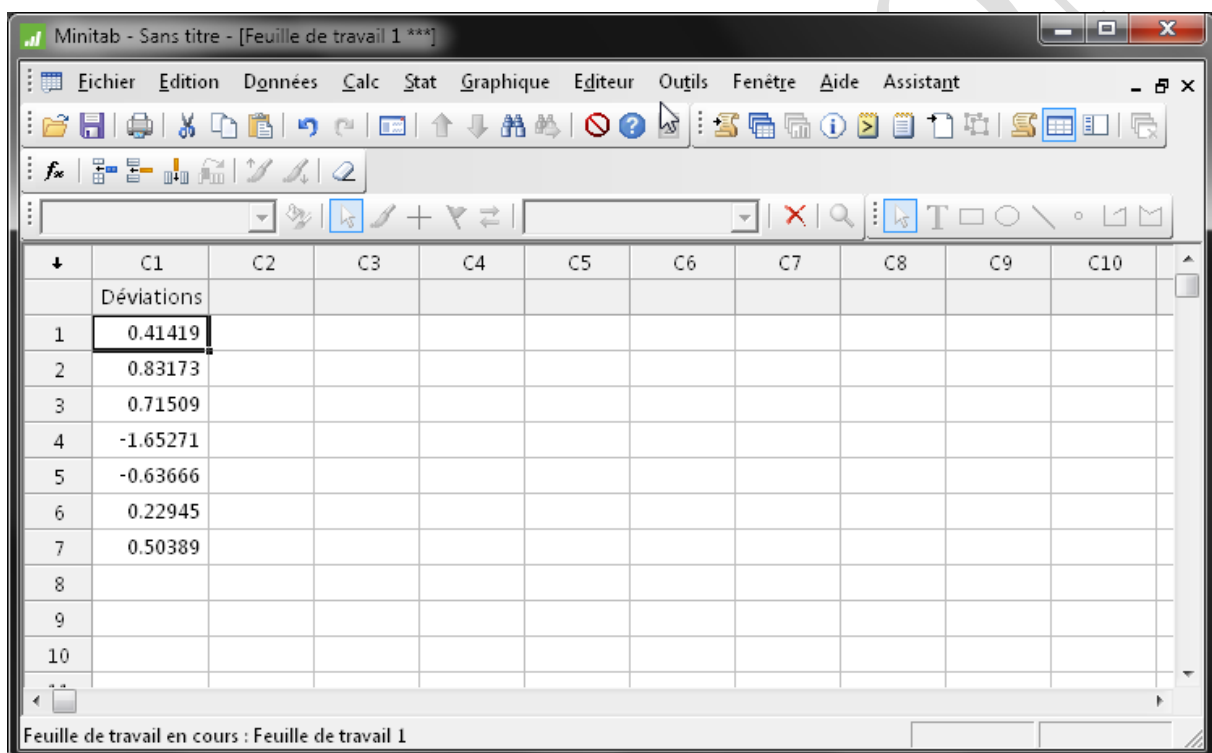
De plus, le graphique ne met pas en évidence par des couleurs les points aberrants ce qui est dommage aussi (bien que l'on puisse obtenir autrement et facilement dans les options du test les points aberrants dans le tableau de la session).

## 10.14. Exercice 45.: Test de Dixon (test des valeurs aberrantes de Dixon)

Minitab® Statistical Software 17.1.0

Le but ici va être de vérifier avec la nouveauté de la version 17 de Minitab la validité du test de Dixon qui y est implémenté que nous avons vu dans le cours théorique et dont nous avons appris à calculer les valeurs critiques à la main pour n'importe quelle distribution.

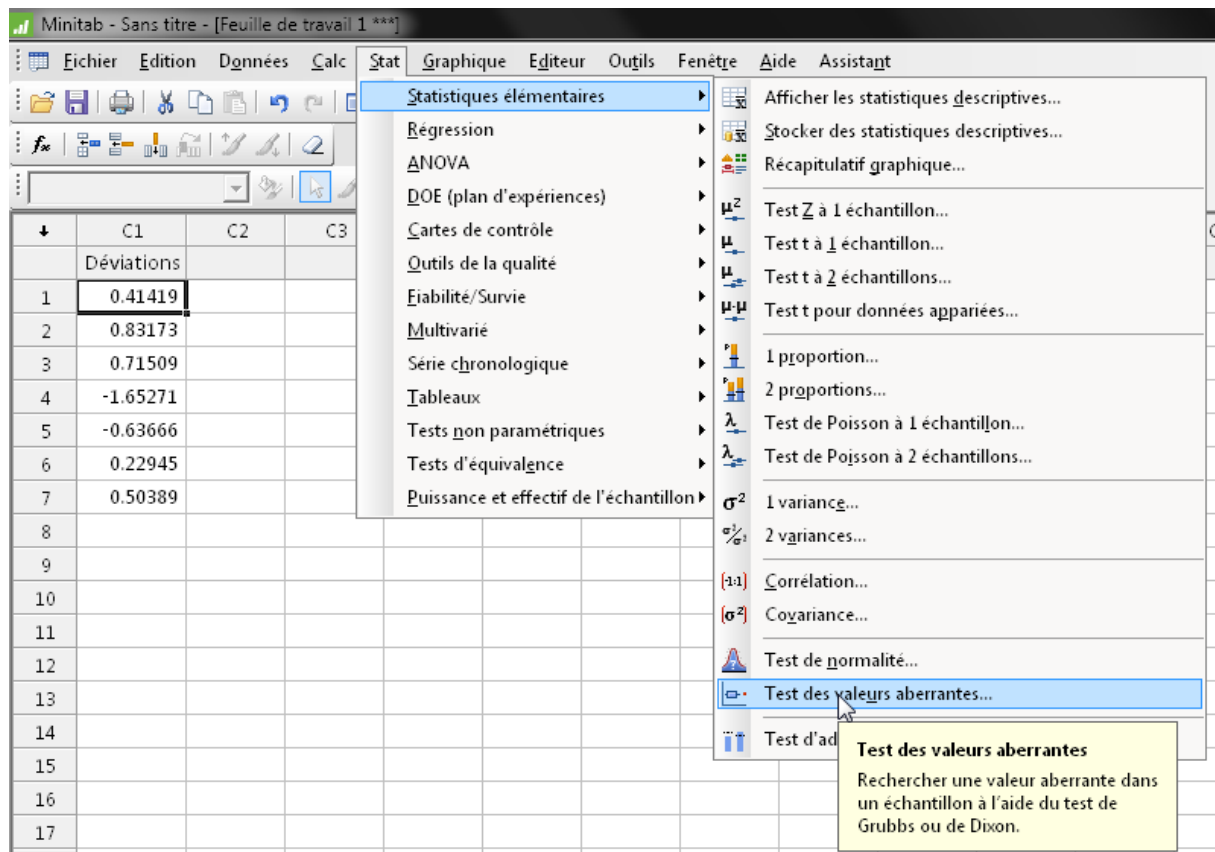
Nous allons prendre la série de valeurs suivante (et cela n'est pas explicite dans Minitab mais nous allons de suite voir sur quelle distribution théorique implicite il se base pour calculer les valeurs aberrantes):



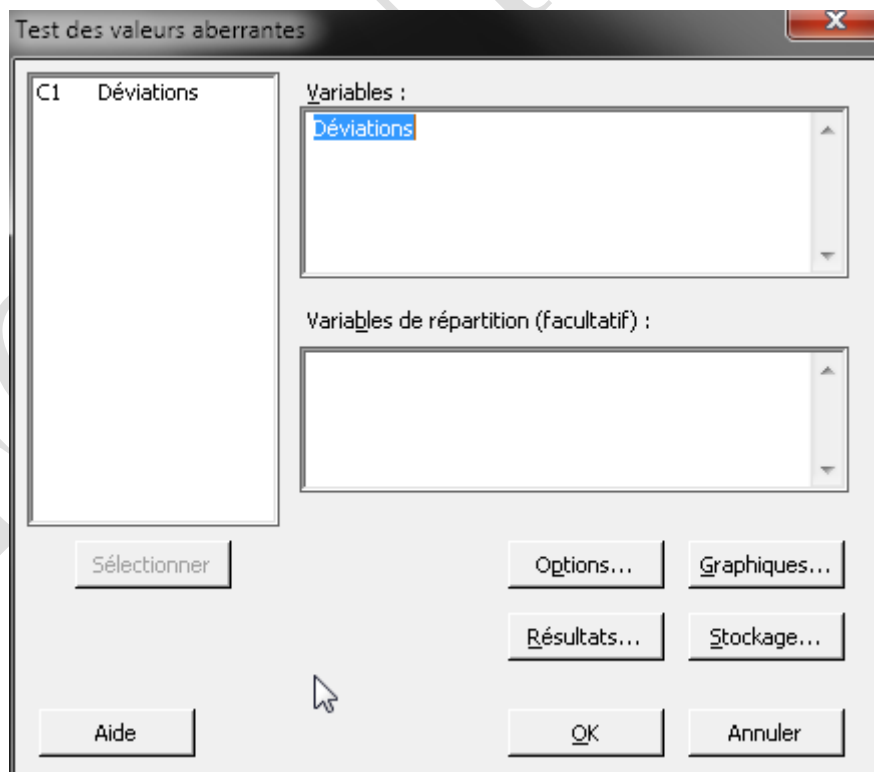
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
	Déviations									
1	0.41419									
2	0.83173									
3	0.71509									
4	-1.65271									
5	-0.63666									
6	0.22945									
7	0.50389									
8										
9										
10										

Et nous allons dans le menu **Stat/Statistiques élémentaires/Test des valeurs aberrantes...**:

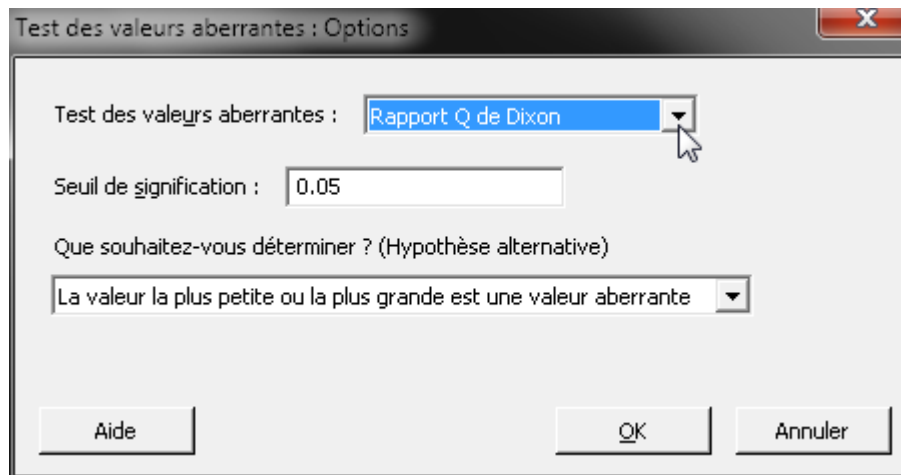




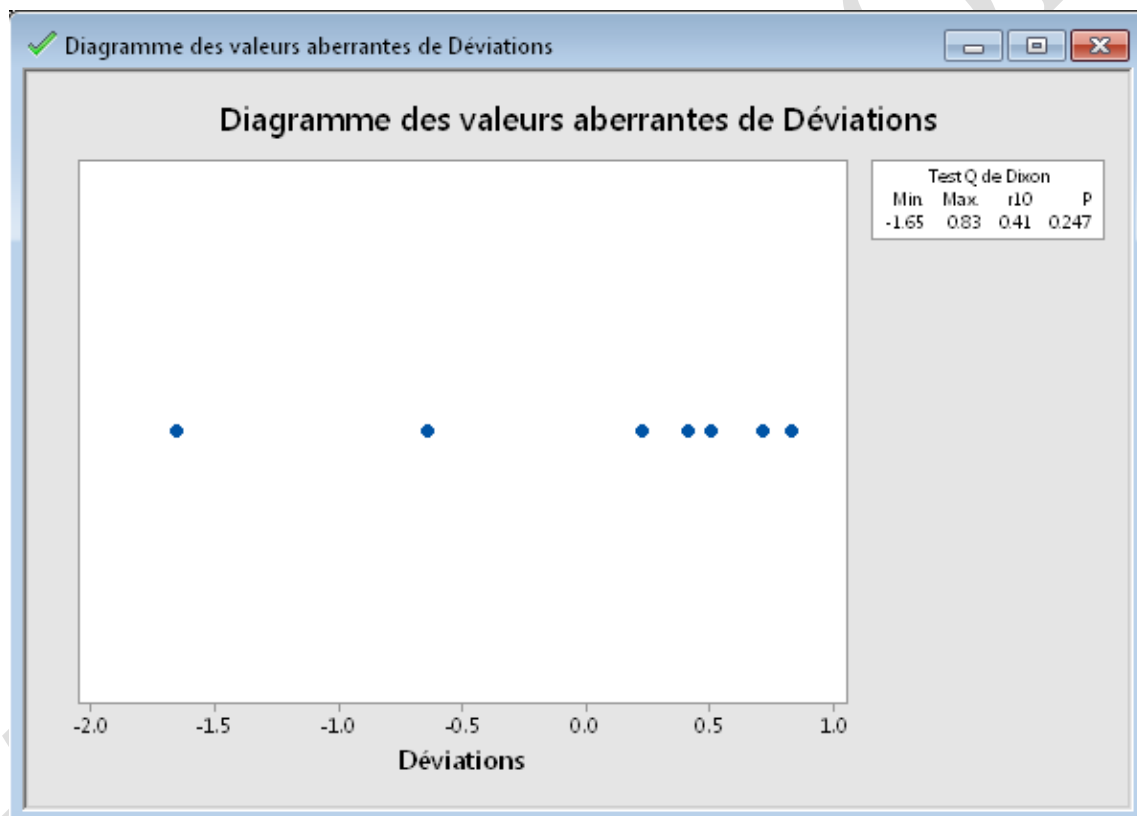
et nous prenons:



Nous cliquons sur **Options...** et nous prenons garde cette fois-ci à sélectionner l'option **Rapport Q de Dixon**:



et nous validons pour obtenir:



Nous retrouvons donc presque les mêmes valeurs de r10 et de la p-value que celles calculées à la main dans le cours théorique en se basant sous l'hypothèse que la distribution est Normale! Donc là c'est un peu dangereux que ce ne soit pas indiqué explicitement par Minitab que le calcul de r10 (et donc la p-value correspondante) se font sous l'hypothèse que la distribution sous-jacente soit Normale. Dommage...!

De plus, le graphique ne met pas en évidence par des couleurs les points aberrants ce qui est dommage aussi (bien que l'on puisse obtenir autrement et facilement dans les options du test les points aberrants dans le tableau de la session).

## 10.15. Exercice 46.: Effectif de l'échantillon pour l'estimation pour loi Normale

Minitab® Statistical Software 16

---

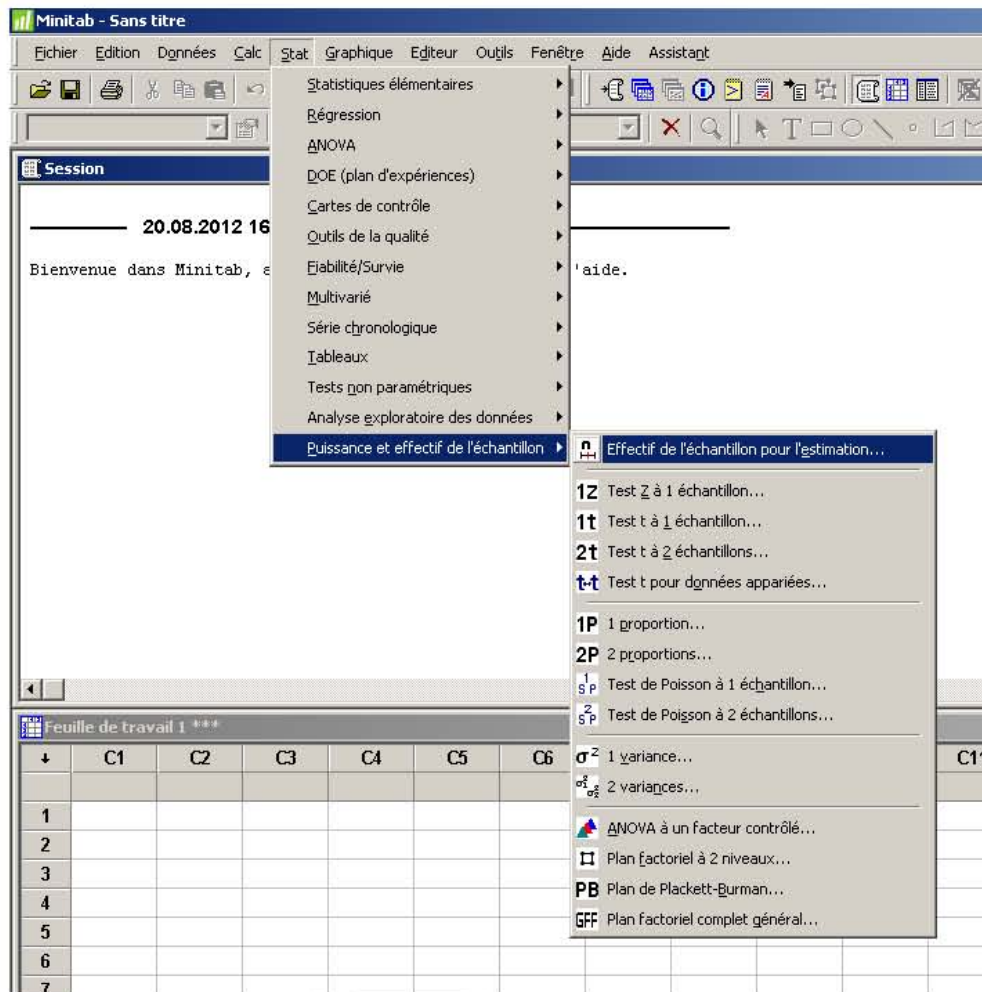
Nous avons démontré dans le cours la relation suivante qui permet de déterminer la taille d'un échantillon pour mettre en évidence une certaine différence entre une moyenne théorique et expérimentale données:

$$n \cong \left( \frac{Z_{\alpha/2} \sigma}{\delta} \right)^2$$

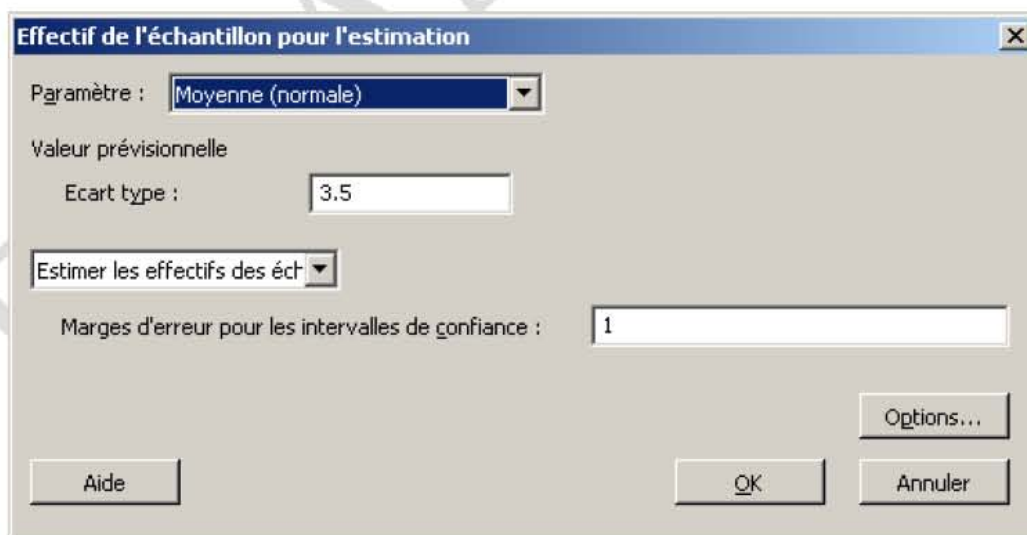
où évidemment la notion de puissance de test ne rentre pas vraiment en compte (puisque posé comme étant égal à 50% comme nous l'avons vu dans le cours théorique!).

Nous souhaiterions dans le cas présent savoir quelle doit être la taille de l'échantillon avec donc implicitement une puissance de 50% pour mettre en évidence une différence (marge d'erreur) de 1 [mm] dans un lot de pièces dont la population totale à un écart-type de 3.5 [mm].

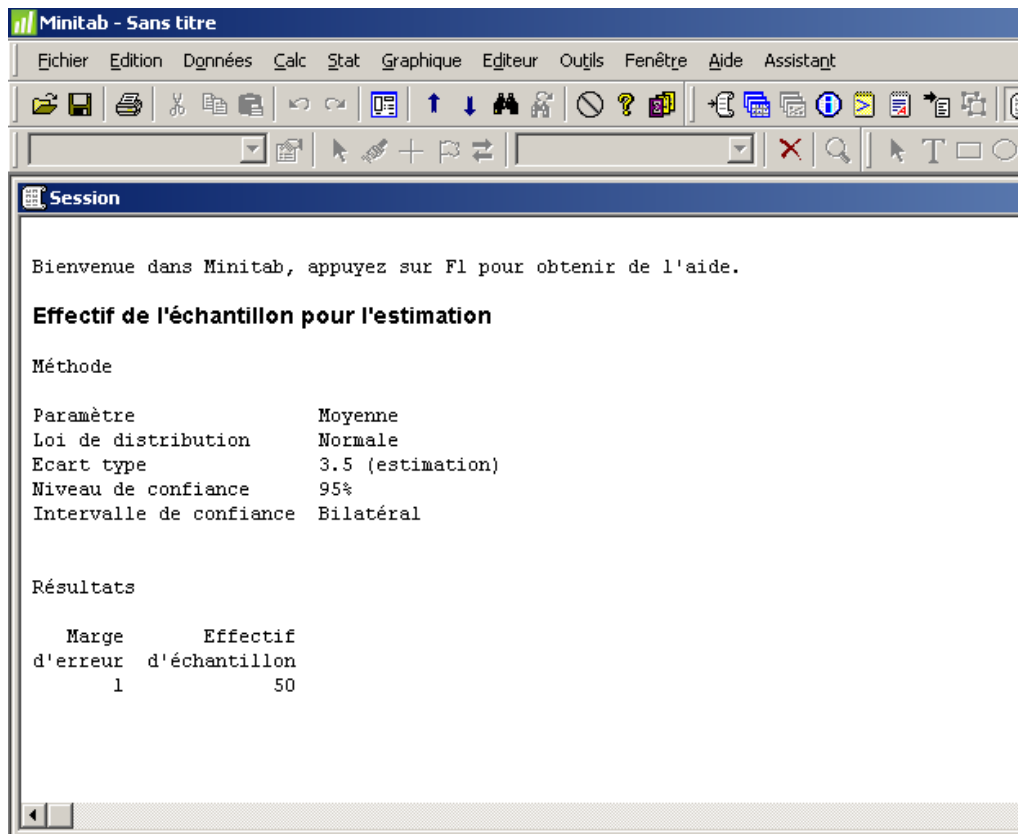
Nous allons pour cela dans le menu **Stat/Puissance et effectif de l'échantillon/Effectif de l'échantillon pour l'estimation...**:



et nous mettons:



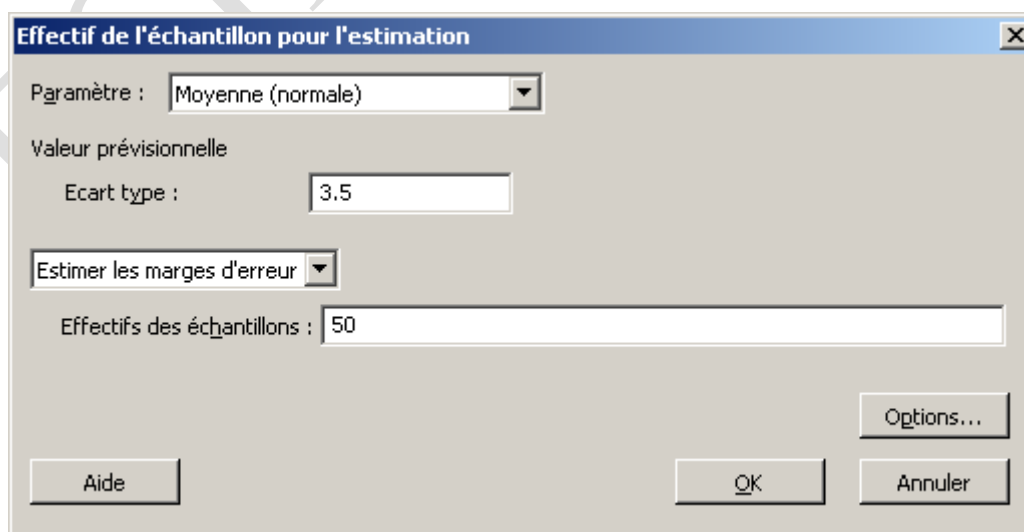
Nous validons par **OK** pour avoir:



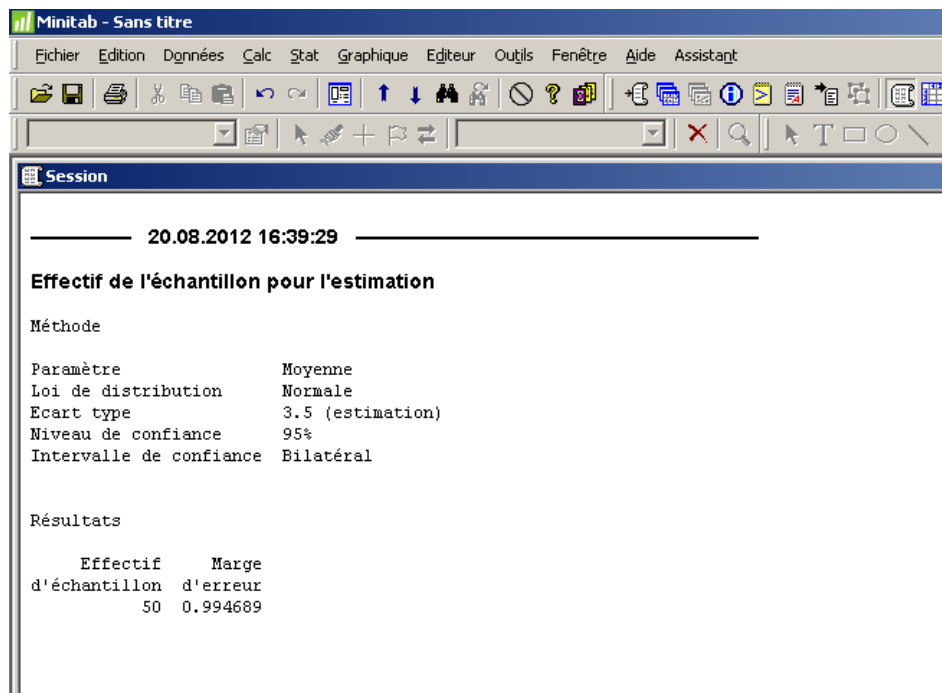
Le lecteur pour vérifier avec les exercices ultérieurs présentant le même calcul avec le choix de la puissance qu'en imposant 50% de puissance qu'il retombe sur presque la même valeur (48 au lieu de 50)! Comparés avec la relation démontrée dans le cours théorique, nous obtenons:

$$n = \left( \frac{Z_{\alpha/2} \sigma}{\delta} \right)^2 = 47.057$$

Nous pouvons nous amuser à inverser le calcul pour voir si nous retombons sur nos pattes:



et nous voyons curieusement que ce n'est pas tout à fait le cas...:



Minitab - Sans titre

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assistant

Session

20.08.2012 16:39:29

**Effectif de l'échantillon pour l'estimation**

Méthode

Paramètre	Moyenne
Loi de distribution	Normale
Ecart type	3.5 (estimation)
Niveau de confiance	95%
Intervalle de confiance	Bilatéral

Résultats

Effectif d'échantillon	Marge d'erreur
50	0.994689

Cet outil utilise une correction de la relation suivante démontrée dans le cours de théorie statistique:

$$n \cong \frac{Z_{\alpha/2}^2 p(1-p)}{\delta^2} = \frac{Z_{\alpha/2}^2 pq}{\delta^2}$$

Si nous faisons les calculs similaires avec la proportion.

## 10.16. Exercice 47.: Puissance d'un test Z à 1 échantillon en bilatéral

Minitab® Statistical Software 15.1.1.0

Un gestionnaire responsable d'une chaîne de production veut savoir combien de temps passe en moyenne un employé à l'usinage sur une pièce donnée. Celui-ci aimerait être dans une estimation se trouvant dans les  $\pm 2$  minutes par rapport à la moyenne et il est connu que l'écart-type d'usinage est de 3 minutes.

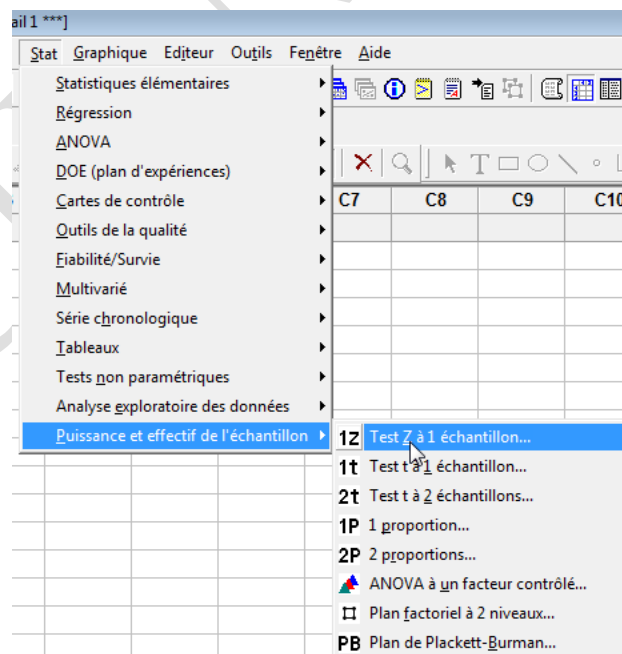
Nous avons déterminé que la taille de l'échantillon de mesures que doit prendre le gestionnaire pour s'assurer un intervalle de confiance à 95% de la moyenne arrondie à l'entier le plus proche est 9 basée sur la relation:

$$n \cong \left( \frac{Z_{\alpha/2} \sigma}{\delta} \right)^2$$

Mais quelle sera la puissance du test?

L'objectif est encore une fois de vérifier que nous obtenons la même chose que le calcul à la main dans le cours de statistique théorique.

Nous allons dans le menu **Stat/Puissance et effectif de l'échantillon/Test Z à 1 échantillon...**:



Nous avons alors (Minitab® Statistical Software calcule toujours le champ qui est laissé vide parmi les 3 premiers):

Puissance et effectif de l'échantillon pour un test Z à 1 échantillon

Spécifier des valeurs pour deux des éléments suivants :

Effectifs des échantillons : 9

Différences : 2

Valeurs de puissance :

Ecart type : 3

Aide Options... Graphique... OK Annuler

et dans **Options...**:

Puissance et effectif de l'échantillon pour un test Z à 1 échantillon - Options

Hypothèse alternative

Inférieure

Différente

Supérieure

Niveau de signification : 0.05

Stocker les effectifs d'échantillons dans :

Stocker les différences dans :

Stocker les valeurs de puissance dans :

Sélectionner Aide OK Annuler

Nous validons par **OK** pour avoir dans la fenêtre de session:

### Puissance et effectif de l'échantillon

Test Z à 1 échantillon

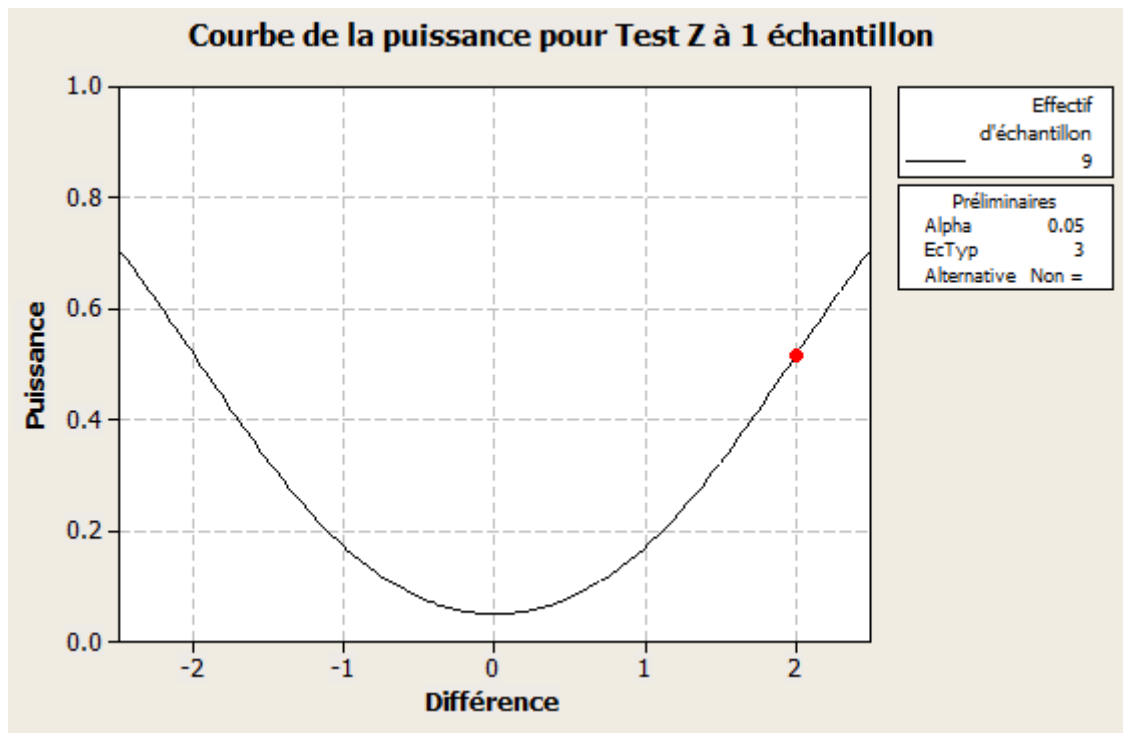
Test de moyenne = nulle (contre moyenne non nulle)

Calcul de puissance pour moyenne = nulle + différence

Alpha = 0.05 Ecart type supposé = 3

	Effectif	
Différence	d'échantillon	Puissance
2	9	0.516005





Nous retrouvons donc pas tout à fait la même valeur de la puissance du test qu'avec Microsoft Excel en utilisant une des relations démontrée dans le cours de statistique théorique:

$$1 - \beta = P\left(Y_{N(0,1)} \leq Z_{\alpha/2} - d\sqrt{n}\right) =$$

$$\text{NORM.DIST}(\text{NORMSINV}(5\%/2) - (-2)/(3/\text{SQRT}(9)); 0; 1; 1) = 0.515967793$$

Il y a donc une différence de:

$$0.0000372 = 0.00372\%$$

avec la valeur donnée par Microsoft Excel. Nous allons donc considérer cela comme négligeable et probablement dû à l'algorithme de calcul pour la loi Normale qui diffère un tout petit peu. Pour résumer nous avons donc  $\beta \cong 50\%$ , soit une chance sur deux de ne pas rejeter l'hypothèse nulle alors qu'elle est fausse.

Indiquons enfin qu'il est possible de faire plusieurs courbes de puissance en écrivant:

Puissance et effectif de l'échantillon pour un test Z à 1 échantillon

Spécifier des valeurs pour deux des éléments suivants :

Effectifs des échantillons : 9 18

Différences : 2 6

Valeurs de puissance :

Ecart type : 3

Options... Graphique...

Aide OK Annuler

Minitab® Statistical Software fait le calcul pour toutes les combinaisons possibles:

### Courbe de la puissance pour Test Z à 1 échantillon

#### Puissance et effectif de l'échantillon

Test Z à 1 échantillon

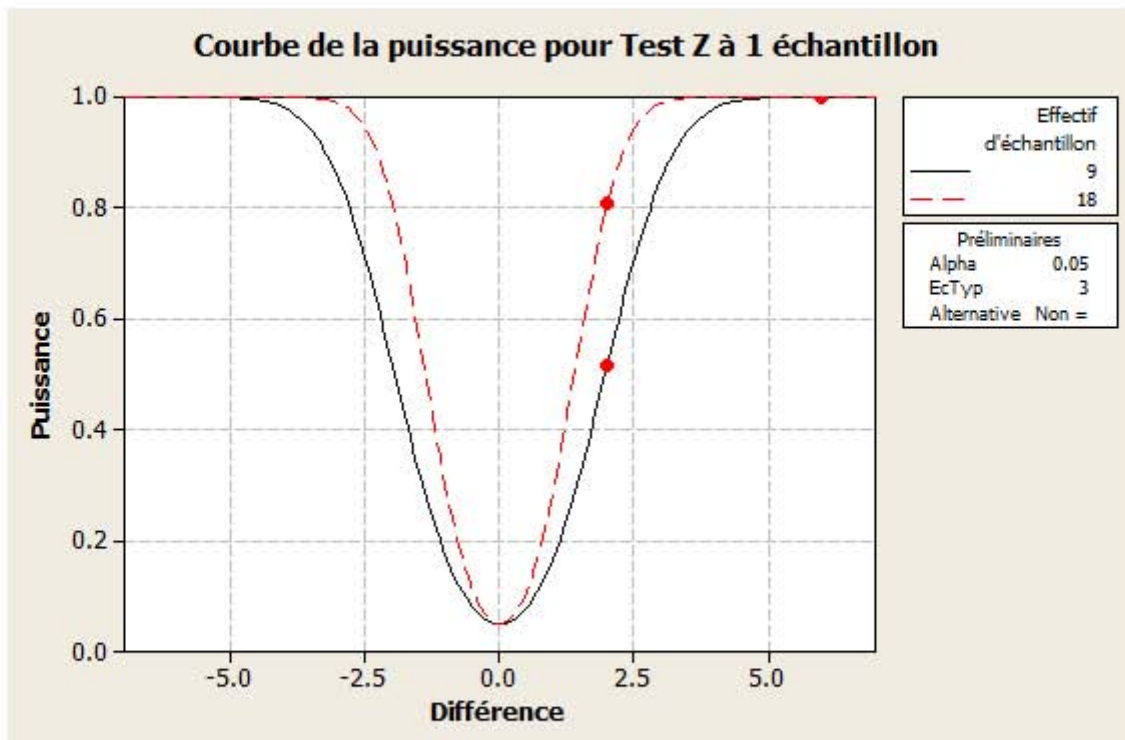
Test de moyenne = nulle (contre moyenne non nulle)

Calcul de puissance pour moyenne = nulle + différence

Alpha = 0.05 Ecart type supposé = 3

Différence	Effectif d'échantillon	Puissance
2	9	0.51601
2	18	0.80743
6	9	0.99997
6	18	1.00000

avec le graphique:



ÉCHANTILLON

## 10.17. Exercice 48.: Taille d'échantillon (effectif) d'un test Z en bilatéral

Minitab® Statistical Software 15.1.1.0

Un gestionnaire responsable d'une chaîne de production veut savoir combien de temps passe en moyenne un employé à l'usinage sur une pièce donnée. Celui-ci aimerait être dans une estimation se trouvant dans les  $\pm 2$  minutes par rapport à la moyenne et il est connu que l'écart-type d'usinage est de 3 minutes. Quel doit être la taille de l'échantillon si nous souhaitons un test avec une puissance de 80% ?

Nous avons déterminé que la taille de l'échantillon de mesures que doit prendre le gestionnaire pour s'assurer un intervalle de confiance à 95% de la moyenne arrondie à l'entier le plus proche est 9 basée sur la relation:

$$n \cong \left( \frac{Z_{\alpha/2} \sigma}{\delta} \right)^2$$

et que cela donnait une puissance d'à peu près 50%. Le but ici est donc de voir à combien va monter  $n$  pour avoir un test avec une puissance de 80%.

La suite est un peu particulière contrairement au calcul à la main car il faut bien se rappeler du chapitre de la "puissance d'un test" que nous avons étudié. Par rapport à notre question, comme la puissance du test n'est pas exigée, cela signifie implicitement que nous tolérons une puissance de 50% (se rappeler de la figure avec les courbes de Gauss qui se superposaient).

Nous avons alors:

Puissance et effectif de l'échantillon pour un test Z à 1 échantillon

Spécifier des valeurs pour deux des éléments suivants :

Effectifs des échantillons :

Différences :

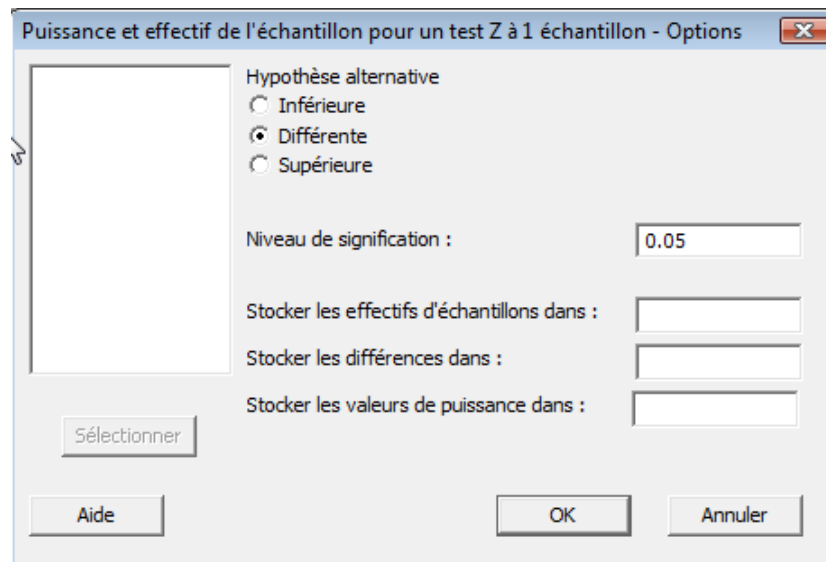
Valeurs de puissance :

Ecart type :

Options... Graphique...

Aide OK Annuler

Nous cliquons sur **Options...** toujours pour faire bien attention d'être à 5% en bilatéral:



Nous validons par **OK** et nous obtenons dans la fenêtre de session:

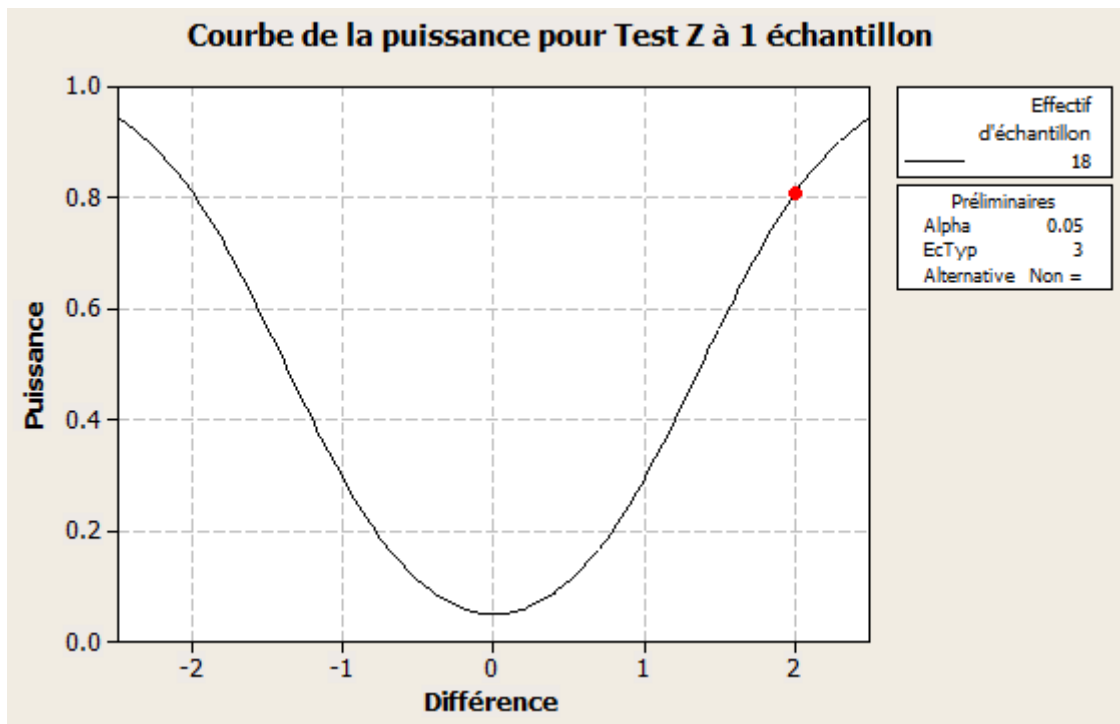
### **Puissance et effectif de l'échantillon**

Test Z à 1 échantillon

Test de moyenne = nulle (contre moyenne non nulle)  
 Calcul de puissance pour moyenne = nulle + différence  
 Alpha = 0.05 Ecart type supposé = 3

	Effectif	Puissance	Puissance
différence	d'échantillon	cible	réelle
2	18	0.8	0.807430

Donc déjà contrairement au calcul à la main, Minitab® Statistical Software arrondi l'effectif d'échantillon à l'entier le plus proche et calcule la puissance réelle du test correspondant à cet arrondi (la puissance est plus grande donc c'est tant mieux!). Minitab® Statistical Software nous sort également la courbe de puissance du test:



Et nous retrouvons bien la même valeur qu'avec Microsoft Excel où nous avons appliqué une des relations démontrée dans le cours théorique (à l'entier près le plus proche):

$$n = \left( \frac{Z_{\alpha/2} - Z_{1-\beta}}{d} \right)^2 = ((\text{NORM.S.INV}(2.5\%) - \text{NORM.S.INV}(80\%)) / (2/3))^2 = 17.6599794$$

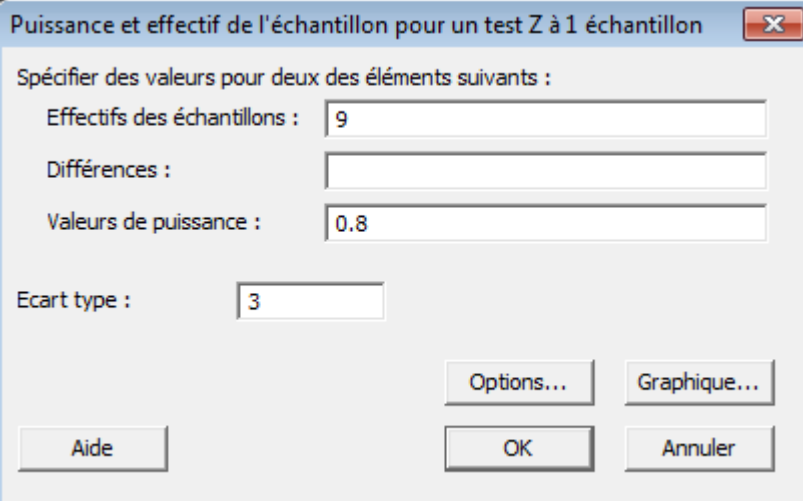
ou avec la deuxième relation démontrée dans le cours théorique:

$$n = \left( \frac{\sigma (Z_{\beta} - Z_{1-\alpha/2})}{\mu - \bar{X}} \right)^2 = ((\text{NORM.S.INV}(20\%) - \text{NORM.S.INV}(1-5\%/2)) / 2)^2 = 17.6599794$$

## 10.18. Exercice 49.: Différence (résolution) d'un test Z en bilatéral

Minitab® Statistical Software 15.1.1.0

Il s'agit toujours de la même démarche mais cette fois, nous voulons déterminer la résolution que nous obtiendrons. Il vient:



Puissance et effectif de l'échantillon pour un test Z à 1 échantillon

Spécifier des valeurs pour deux des éléments suivants :

Effectifs des échantillons : 9

Différences :

Valeurs de puissance : 0.8

Ecart type : 3

Options... Graphique...

Aide OK Annuler

et en validant par **OK** nous avons dans la fenêtre de session:

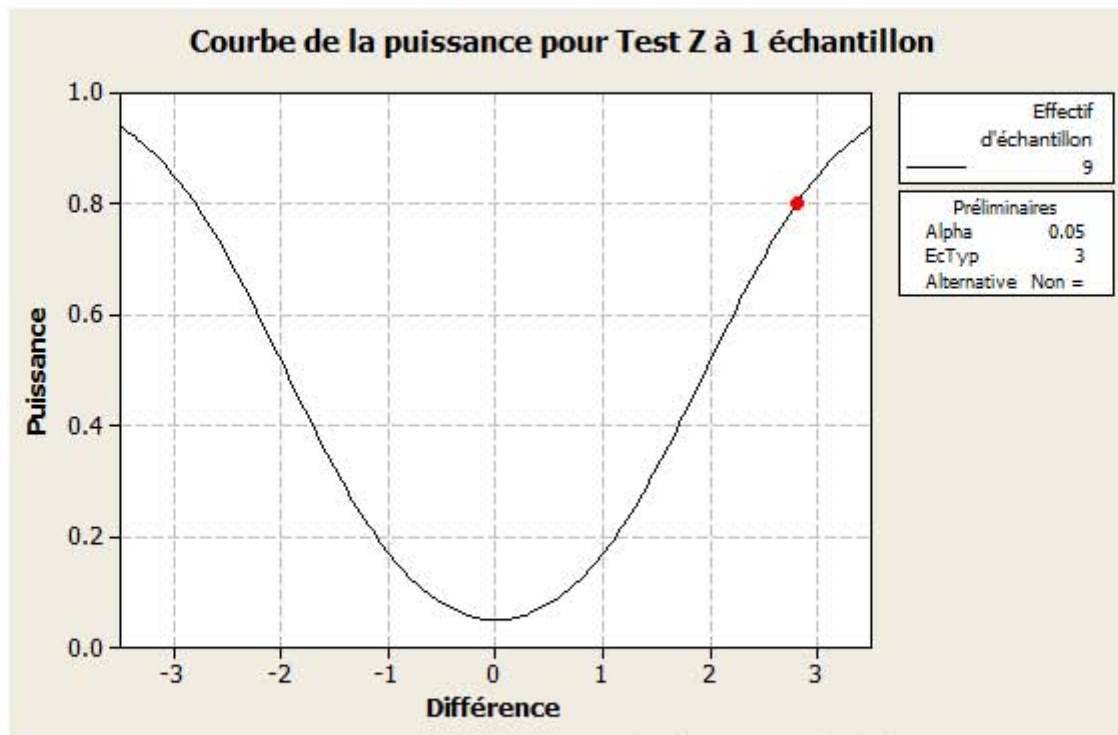
### Puissance et effectif de l'échantillon

Test Z à 1 échantillon

Test de moyenne = nulle (contre moyenne non nulle)  
Calcul de puissance pour moyenne = nulle + différence  
Alpha = 0.05 Ecart type supposé = 3

Effectif d'échantillon	Puissance	Différence
9	0.8	2.80158

et la courbe de puissance correspondante:





## 10.19. Exercice 50.: Intervalle de confiance de la moyenne (test Z à un 1 échantillon)

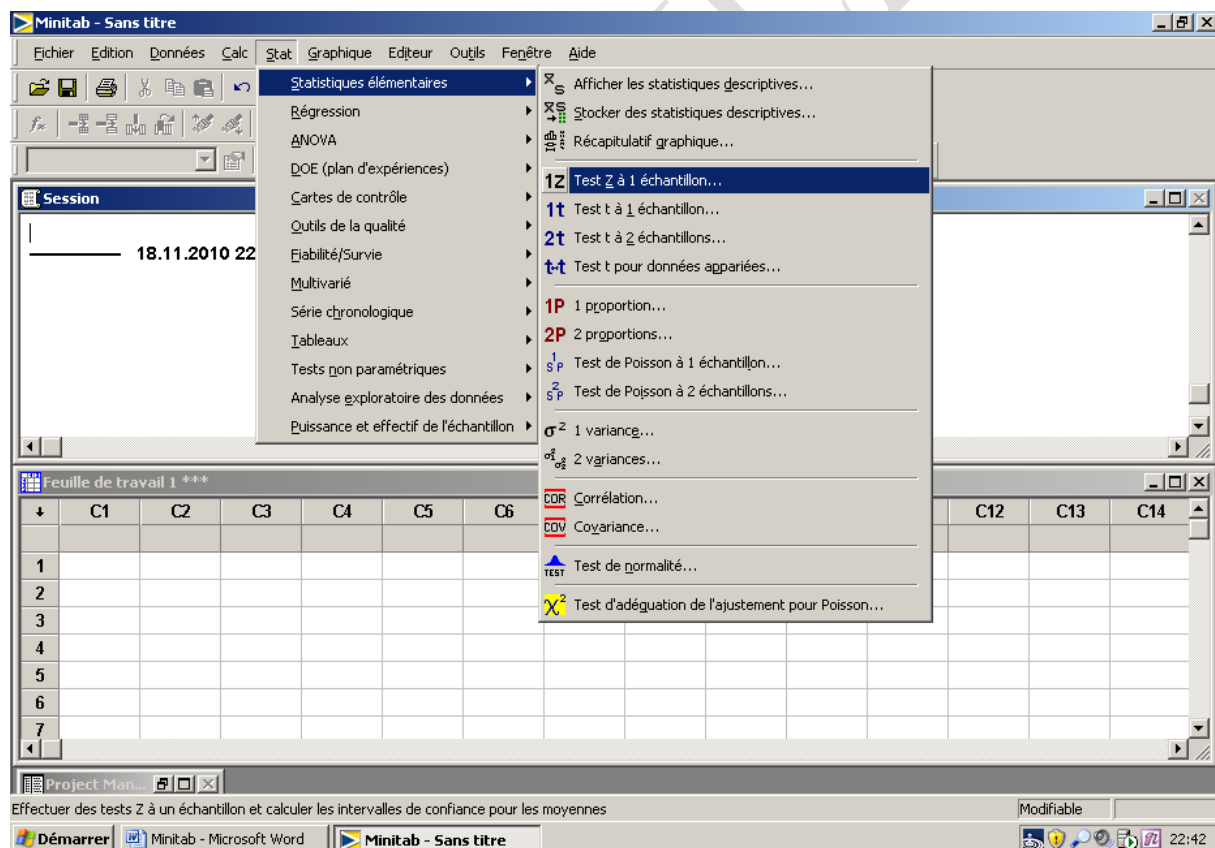
Minitab® Statistical Software 15.1.1.0

Un sondage a été conduit dans des entreprises qui utilisent des panneaux solaires comme source primaire d'électricité. La question qui a été posée était la suivante: Quel est le % de votre électricité provenant du solaire?

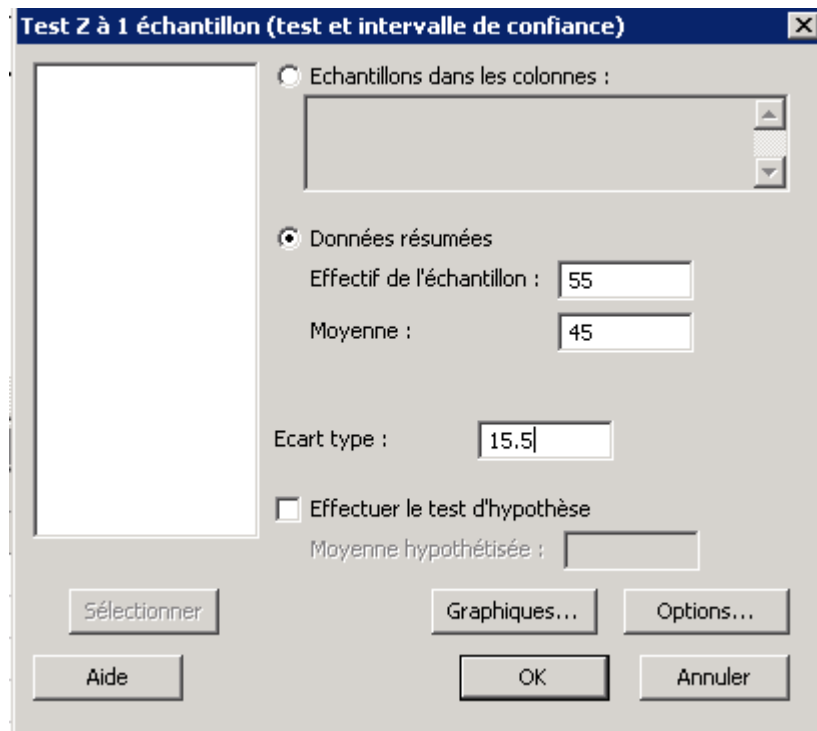
Un échantillon aléatoire de 55 réponses donne une moyenne arithmétique de 45 [kW]. Supposons que l'écart-type standard pour cette question est de 15.5 [kW] (écart-type théorique supposé connu!).

Quel est l'intervalle de confiance représentant 95% de valeurs (nous avons fait ce calcul avec Microsoft Excel).

Nous ouvrons Minitab® Statistical Software et allons dans le menu **Stat/Statistiques élémentaires/Test Z à 1 échantillon...**:



Nous avons alors:



Et nous obtenons alors dans la fenêtre de session:

Minitab - Sans titre

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

Session

**Test Z à 1 échantillon**

Ecart type supposé = 15.5

N	Moyenne	ErT	IC à 95 %
55	45.00	2.09	(40.90; 49.10)

Feuille de travail 1 \*\*\*

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
1												
2												
3												
4												
5												
6												
7												
8												

Feuille de travail en cours : Feuille de travail 1

Ce qui correspond bien à ce que nous avons obtenu avec Microsoft Excel. Évidemment, il n'y a aucune  $p$ -value d'indiquée puisque nous n'avons pas imposé de moyenne hypothétisée dans la boîte de dialogue principale.

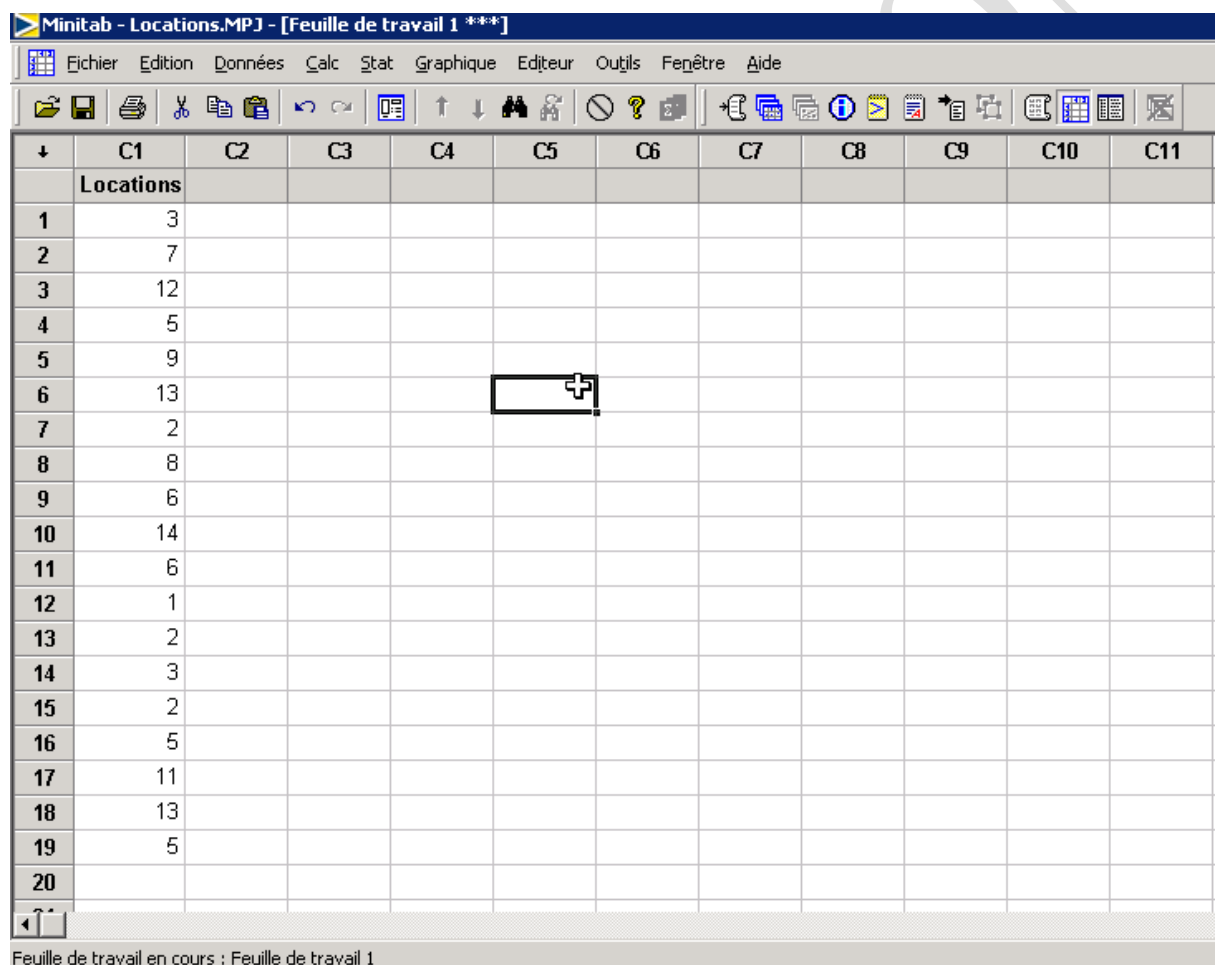
## 10.20. Exercice 51.: Test Z de différence de la moyenne (test Z à un 1 échantillon) en bilatéral

Minitab® Statistical Software 15.1.1.0

Un objectif de location journalier a été imposé à gestionnaire d'une société de location de voitures avec un objectif de 5 et un écart-type de 2. Depuis 19 jours que le garage est ouvert, nous avons les locations suivantes:

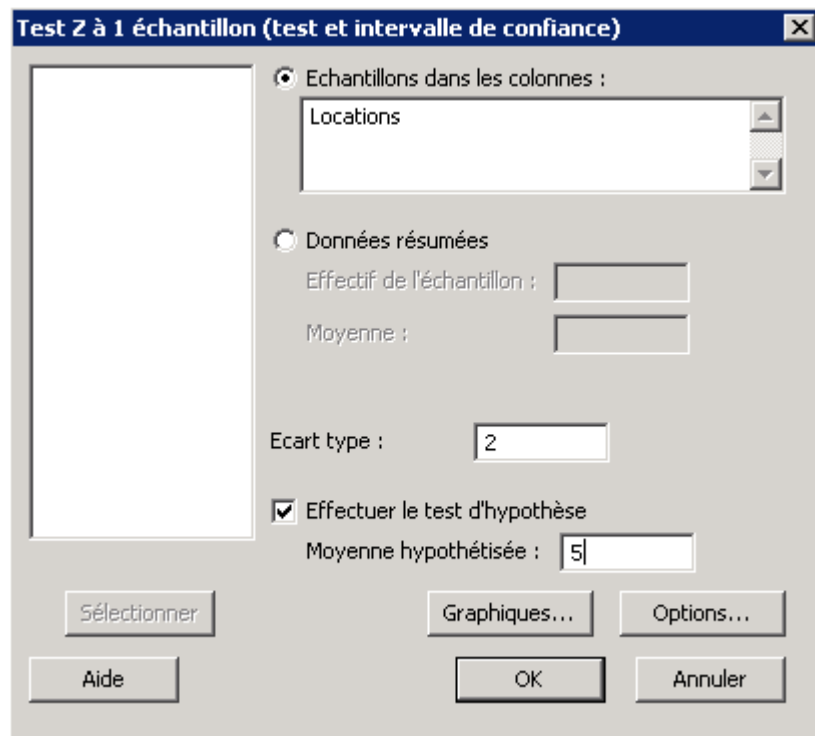
3,7,12,5,9,13,2,8,6,14,6,1,2,3,2,5,11,13,5

ou ouvrez le fichier *Locations.mpj*:

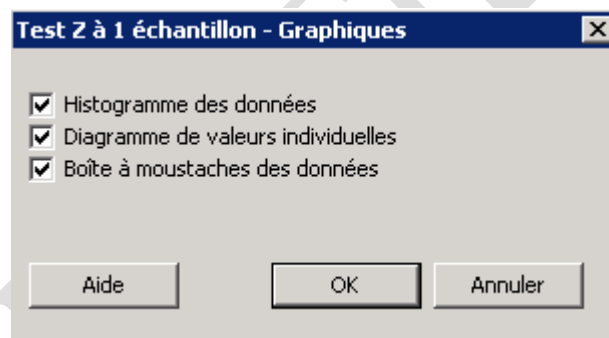


	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
	Locations										
1	3										
2	7										
3	12										
4	5										
5	9										
6	13										
7	2										
8	8										
9	6										
10	14										
11	6										
12	1										
13	2										
14	3										
15	2										
16	5										
17	11										
18	13										
19	5										
20											

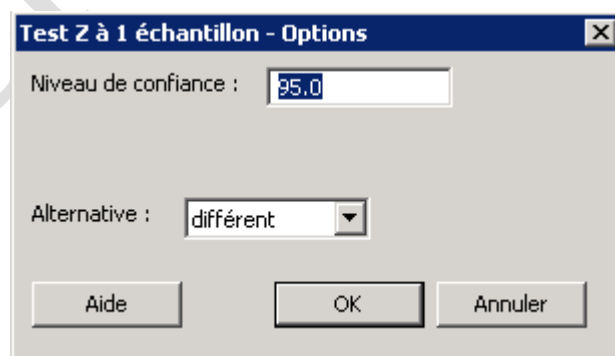
et allez dans le menu **Stat/Statistiques élémentaires/Test Z à 1 échantillon...**:



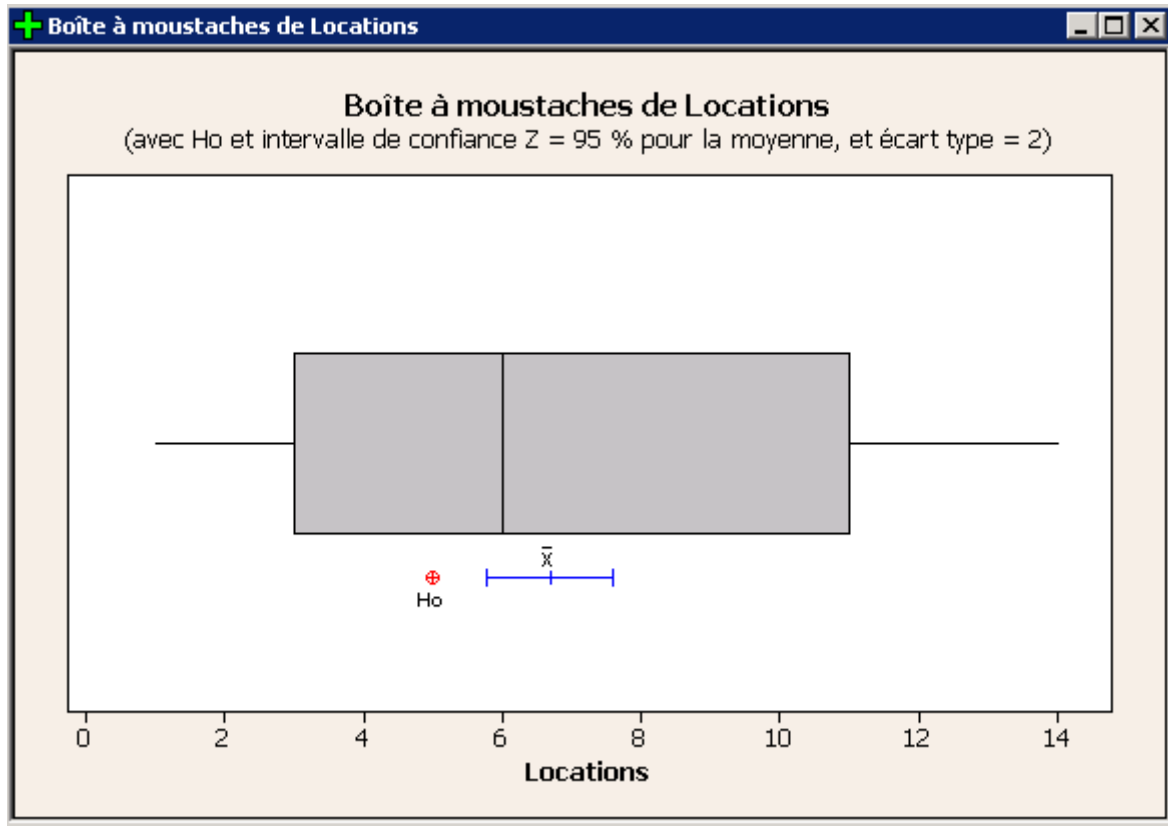
Nous cliquons sur **Graphiques...** pour activer:



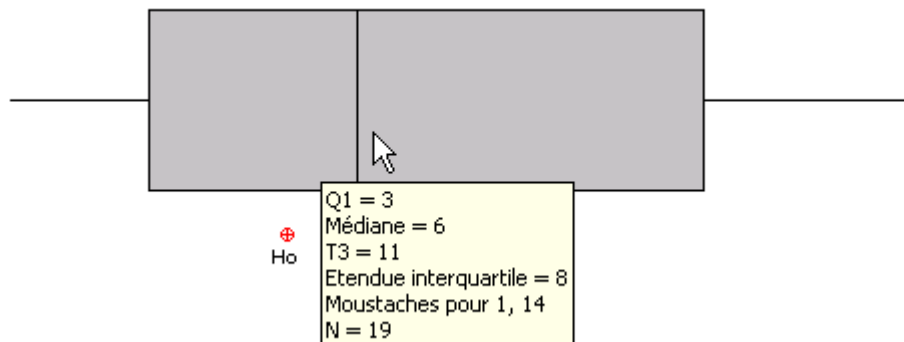
et dans **Options...** nous nous assurons pour commencer à être dans un test bilatéral:



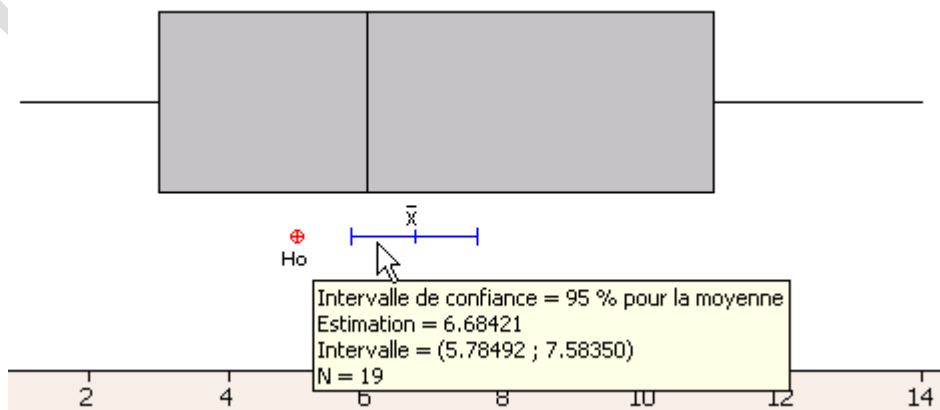
Nous validons le tout par **OK** et d'abord nous obtenons les graphiques:



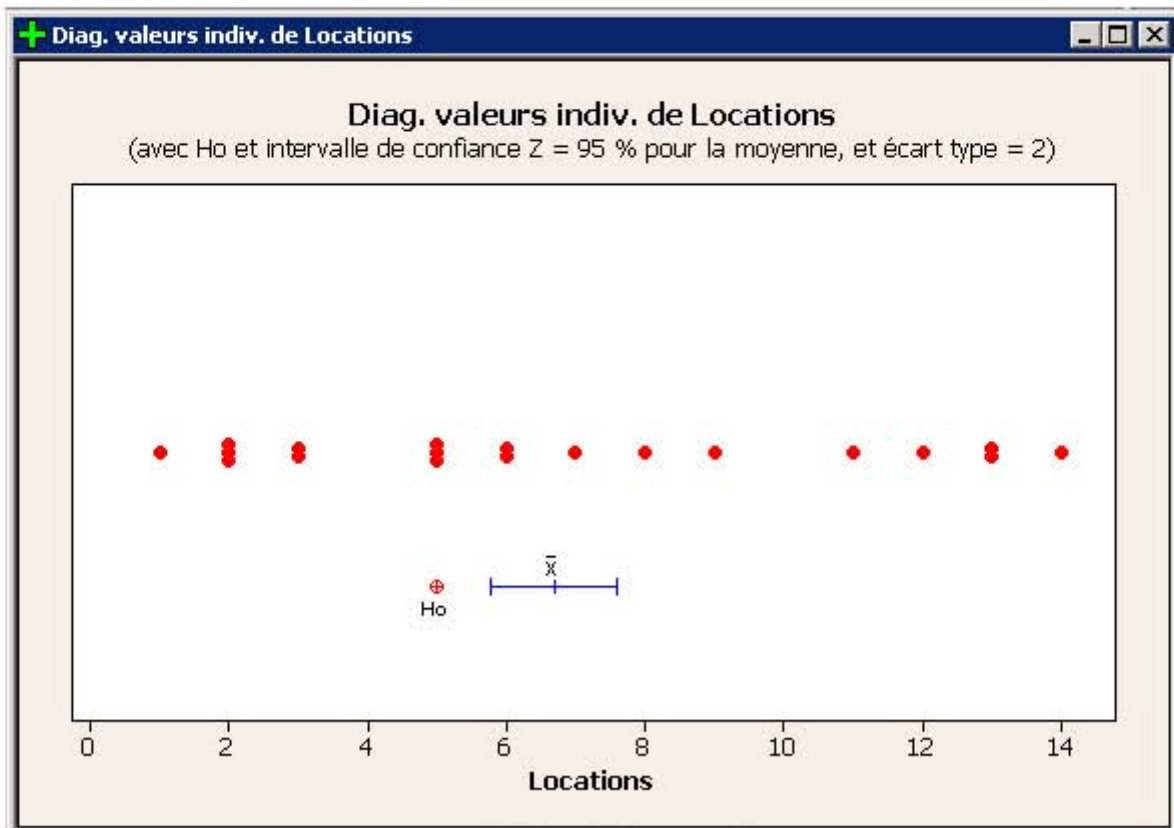
avec l'infobulle sur la boîte à moustache:



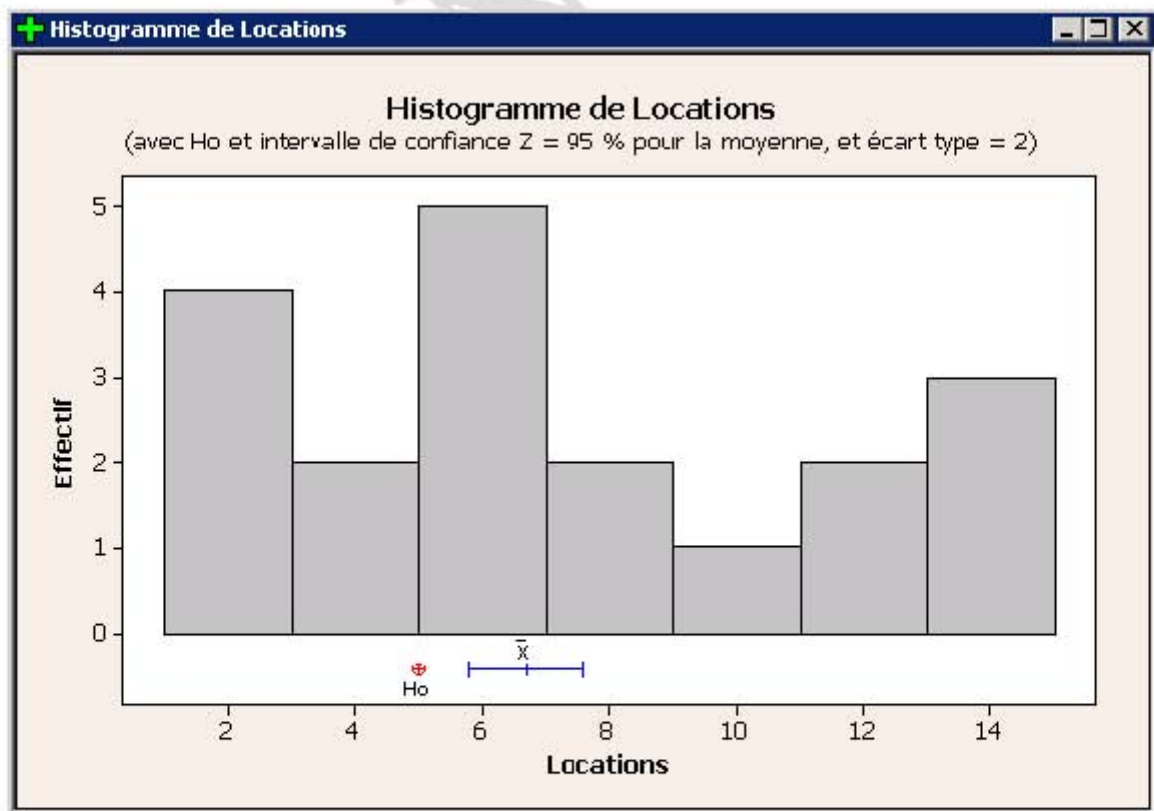
et sur l'intervalle de confiance:



Ensuite, nous avons aussi le diagramme à points:



et l'histogramme:



Mais intéressons-nous plutôt au contenu de la fenêtre de session:

**Test Z à un échantillon : Locations**

Test de mu = 5 en fonction de la différence 5  
Ecart type supposé = 2

Variable	N	Moyenne	EcTyp	ErT moyenne	IC à 95 %	Z	P
Locations	19	6.684	4.230	0.459	(5.785; 7.584)	3.67	0.000

L'intervalle de confiance à 95% ne changera pas quelle que soit la moyenne hypothétisée (ce qui est logique conformément à ce que nous avons vu dans le cours théorique). Au vu de l'intervalle de confiance à 95% il est donc évident que la valeur Z relativement à la moyenne hypothétisée de 5 soit très grande et donc la  $p$ -value très petite.

Par contre, dans le cours théorique nous n'avons pas vu comment calculer le Z dans le cas de la différence avec Microsoft Excel. Donc si jamais voici comment faire (en se rappelant que l'idée sous-jacente est que la différence est nulle):

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{6.684 - 5}{\frac{2}{\sqrt{19}}} = 3.670193 \cong 3.67$$

et donc en bilatéral la  $p$ -value est:

$$=(1-\text{NORM.S.DIST}(3.67019291;\text{TRUE}))= (1-0.999878816)/2=0.0000606$$

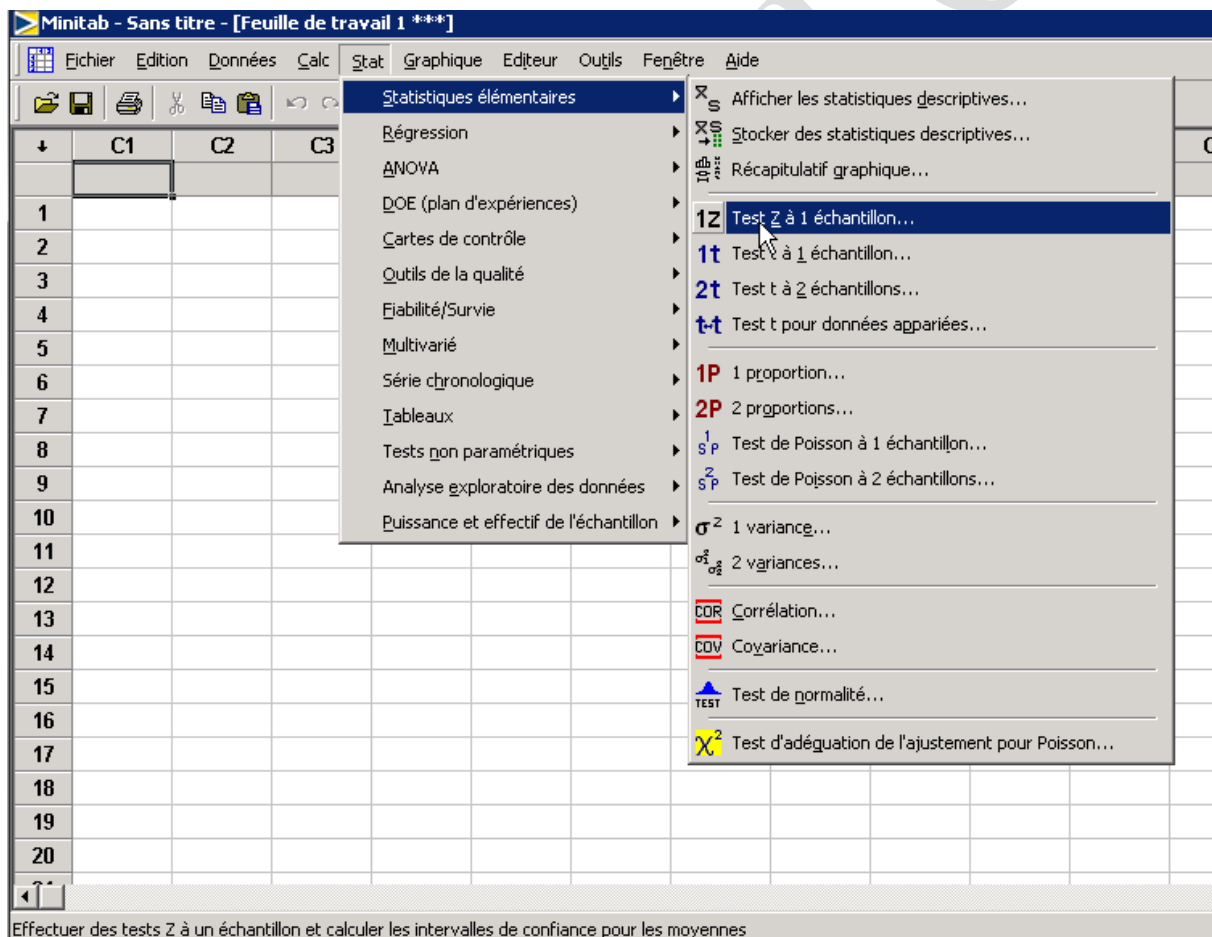
## 10.21. Exercice 52.: Test Z de différence de la moyenne (test Z à un 1 échantillon) en unilatéral gauche

Minitab® Statistical Software 15.1.1.0

Il est connu dans un État que les enfants d'un certain âge ont un poids de 45 kilogrammes et un écart-type de 13 kilogrammes (espérance et écart-type de la population). Une plainte est posée par des parents d'élèves comme quoi les enfants d'une école sont sous-alimentés. Pour cela les parents d'élèves s'appuient sur le fait que 25 enfants du même âge ont un poids moyen de 40.5 kilogrammes.

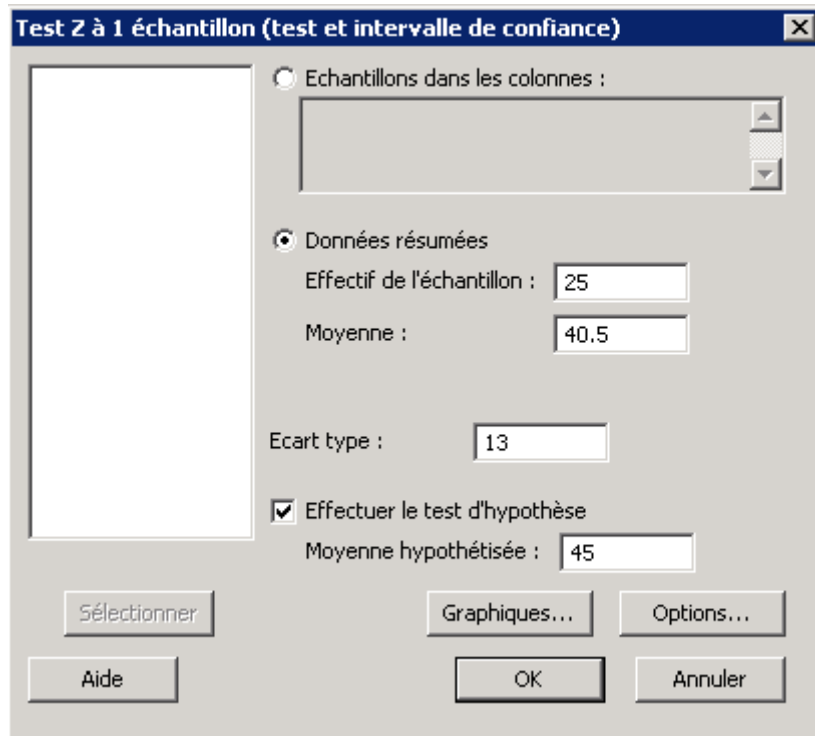
Vérifions si ce sous-poids est significativement inférieur!

Nous allons donc dans le menu ans le menu **Stat/Statistiques élémentaires/Test Z à 1 échantillon...**:

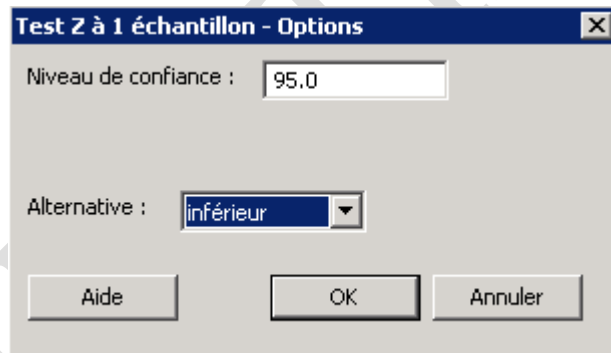


Nous avons alors:





et dans **Options...** nous prenons alors l'hypothèse alternative comme quoi la moyenne mesurée est inférieure à la moyenne hypothétisée:



Nous validons deux fois par **OK** pour obtenir dans la fenêtre de session:

**Test Z à 1 échantillon**

Test de  $\mu = 45$  fn de  $< 45$   
 Ecart type supposé = 13

Devrait être  $\mu \leq 45$  contre  $\mu > 45$  selon moi

N	Moyenne	ErT	Borne supérieure	Z	P
25	40.50	2.60	44.78	-1.73	0.042

Par contre, dans le cours théorique nous n'avons également pas vu comment calculer le Z dans le cas de la différence avec Microsoft Excel. Donc si jamais voici comment faire (en se rappelant que l'idée sous-jacente est que la différence est nulle):

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{40.5 - 45}{\frac{13}{\sqrt{25}}} = -1.73077 \cong -1.73$$

et donc en unilatéral la probabilité cumulée d'être à cette valeur (correspondant directement à la  $p$ -value) est de:

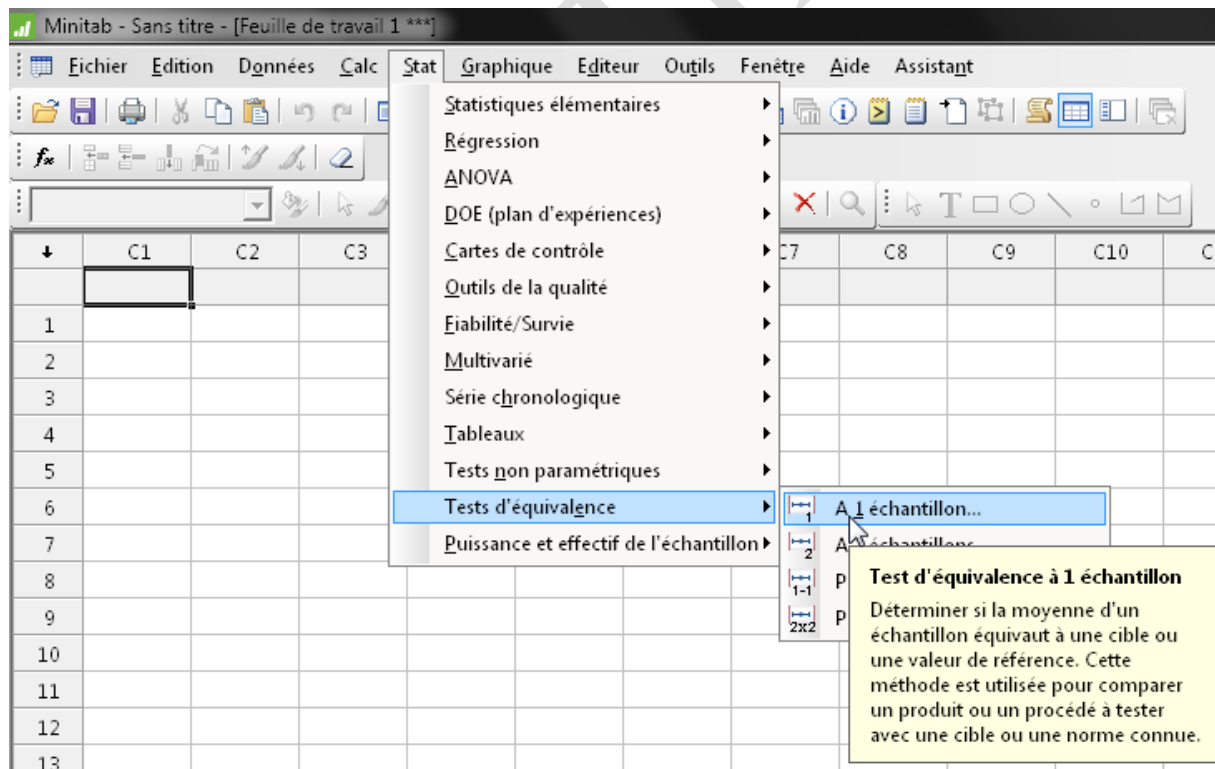
$$= \text{NORM.S.DIST}(-1.730769231; \text{TRUE}) = 0.041746466 = 4.17\%$$

Nous rejetons donc l'hypothèse nulle de l'absence de différence au profit de l'hypothèse alternative. Les enfants sont significativement différemment alimentés à un niveau de 5%.

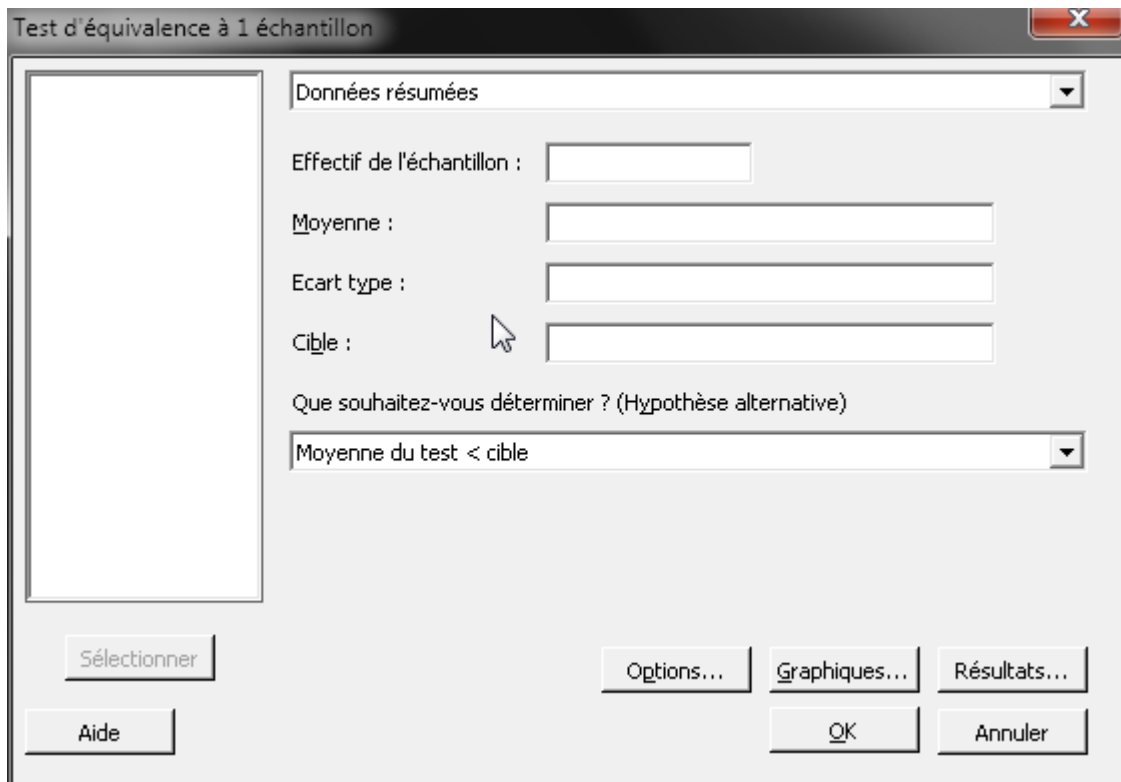
Nous voyons aussi que le résultat est le même si nous calculions la différence dans l'autre sens:

$$Z = \frac{\mu - \bar{X}}{\frac{\sigma}{\sqrt{n}}} = \frac{45 - 40.5}{\frac{13}{\sqrt{25}}} = 1.73077 \cong 1.73$$

Depuis Minitab 17 nous ne pouvons plus procéder ainsi et ce pour des raisons pédagogique, l'équipe de développement a préféré (et c'est judicieux selon moi) mettre cela dans le menu **Stat/Tests d'équivalence**:



Donc si nous voulons reproduire l'exemple ci-dessus nous cliquons sur **A 1 échantillon**:



Test d'équivalence à 1 échantillon

Données résumées

Effectif de l'échantillon :

Moyenne :

Ecart type :

Cible :

Que souhaitez-vous déterminer ? (Hypothèse alternative)

Moyenne du test < cible

Sélectionner

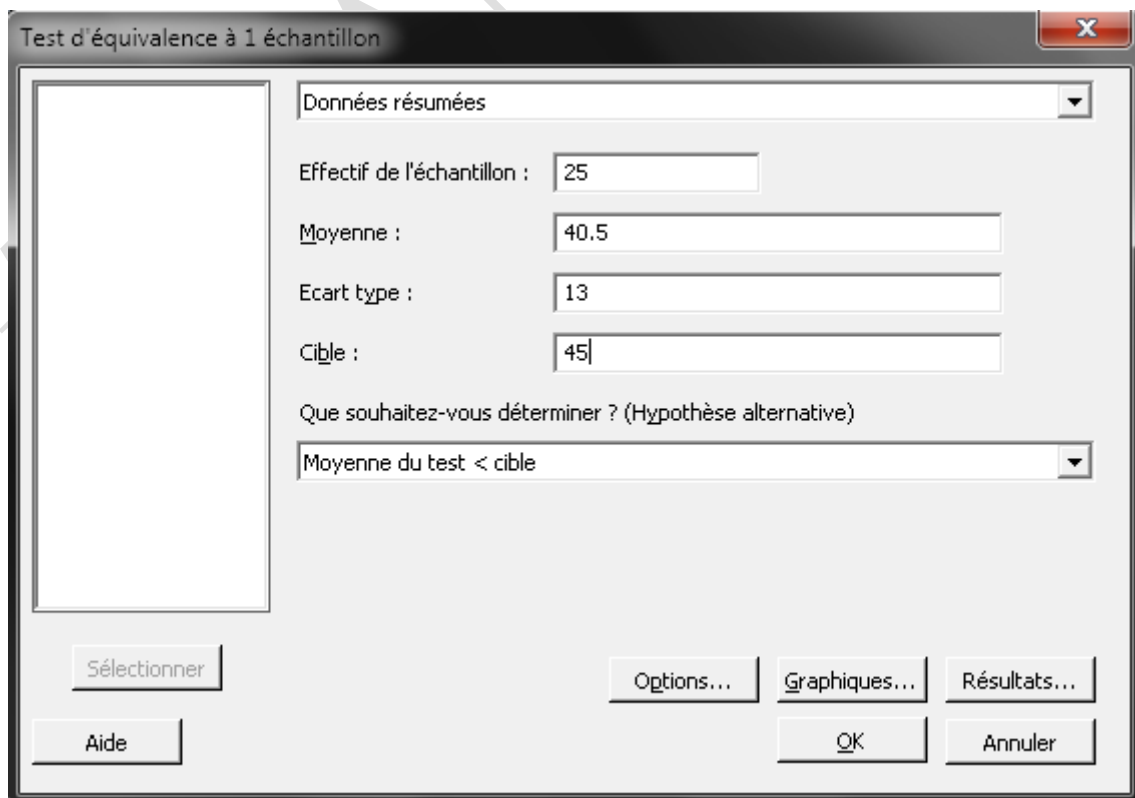
Aide

Options... Graphiques... Résultats...

OK Annuler

Au vu de l'aspect de la boîte de dialogue nous concluons rapidement qu'il s'agira non pas d'un test Z unilatéral mais d'un test de Student en unilatéral (ce qui n'est plus vraiment explicite par contre avec cette nouvelle version).

Nous allons quand même mettre nos données dedans en supposant qu'elles satisfont les hypothèses ad hoc comme toujours:



Test d'équivalence à 1 échantillon

Données résumées

Effectif de l'échantillon : 25

Moyenne : 40.5

Ecart type : 13

Cible : 45

Que souhaitez-vous déterminer ? (Hypothèse alternative)

Moyenne du test < cible

Sélectionner

Aide

Options... Graphiques... Résultats...

OK Annuler

Nous validons par **OK** pour obtenir dans la fenêtre de session:

The screenshot shows the Minitab software interface with the following content in the session window:

**Test d'équivalence à un échantillon**

Méthode

Cible = 45

Statistiques descriptives

N	Moyenne	EcTyp	ErT
25	40.5	13	2.6000

Différence : moyenne - cible

Différence	ErT	Borne supérieure de 95 %	Limite supérieure
-4.5000	2.6000	-0.051707	0

La borne supérieure est inférieure à 0. Déclaration moyenne < cible possible.

Test

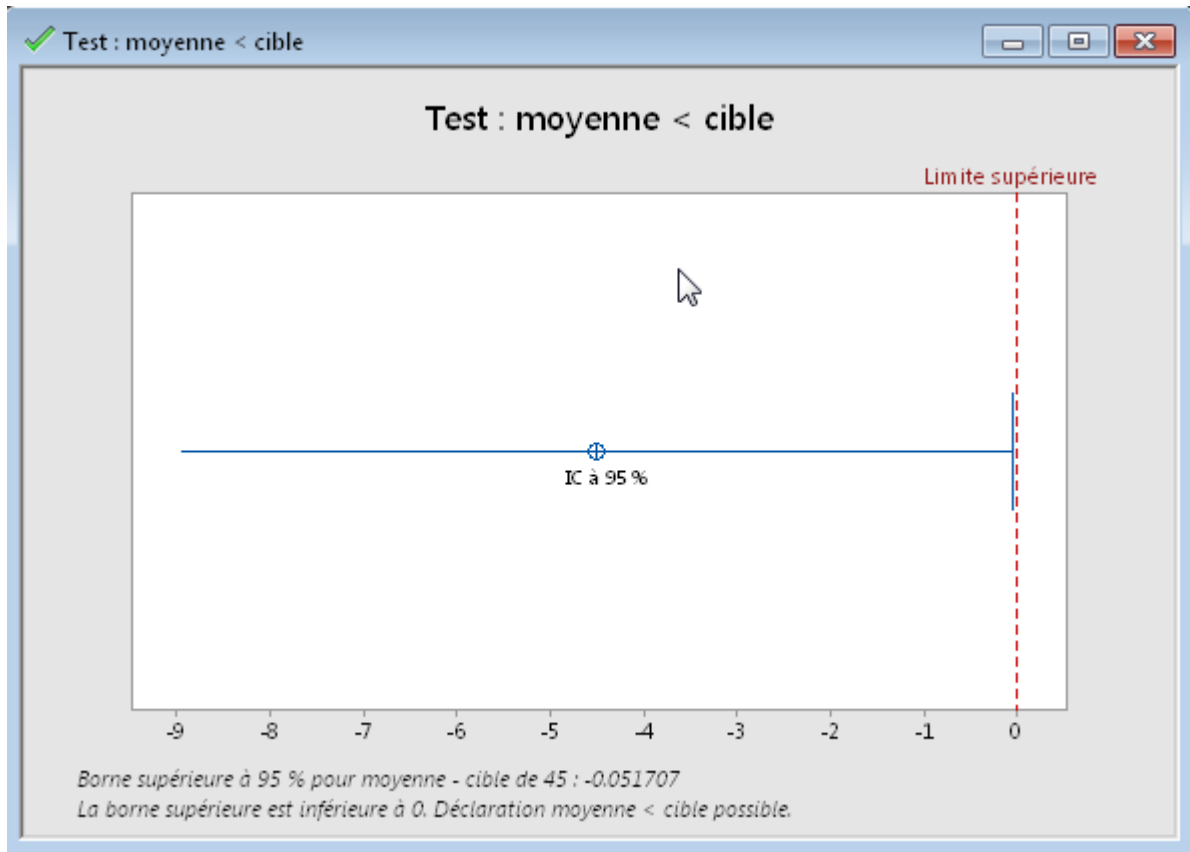
Hypothèse nulle : Moyenne - cible  $\geq$  0  
 Hypothèse alternative : Moyenne - cible < 0  
 Niveau d' $\alpha$  : 0.05

DL	Valeur de T	Valeur de P
24	-1.7308	0.048

Valeur de p  $\leq$  0.05. Déclaration moyenne < cible possible.

Feuille de travail en cours : Feuille de travail 1

et le graphique suivant:



Dans le cours théorique nous n'avons également pas vu comment calculer ce résultat dans le cas de la différence avec Microsoft Excel. Donc si jamais voici comment faire (en se rappelant que l'idée sous-jacente est que la différence est nulle):

$$T = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{40.5 - 45}{\frac{13}{\sqrt{25}}} = -1.73077 \cong -1.73$$

et donc en unilatéral la probabilité cumulée d'être à cette valeur (correspondant directement à la  $p$ -value) est de:

$$=T.DIST(-1.73;25-1;TRUE)=0.048236=4.823\%$$

Nous ne rejetons donc l'hypothèse nulle au profit de l'hypothèse alternative. Les enfants sont donc bien sous-alimentés à un niveau de 5%. C'est la raison pour laquelle Minitab précise: *Déclaration moyenne < cible possible.*

Et ainsi de suite pour tous les tests d'équivalences de Minitab qui ne sont en réalité que des tests classiques de Student en unilatéral, bilatéral avec ou sans cible fixées tels que nous les avons vu dans le cours théorique (excepté pour le plan croisé).

## 10.22. Exercice 53.: Puissance d'un test t à 1 échantillon en bilatéral

Minitab® Statistical Software 15.1.1.0

Un gestionnaire responsable d'une chaîne de production veut savoir combien de temps passe en moyenne un employé à l'usinage sur une pièce donnée. Celui-ci aimerait être dans une estimation se trouvant dans les  $\pm 2$  minutes par rapport à la moyenne et l'écart-type d'usinage est estimé à 3 minutes. Quelle est la puissance du test sachant que la de taille de l'échantillon est de 9 unités?

Pour faire ce calcul, nous utilisons toujours la version studentisée de la relation:

$$1 - \beta = P\left(Y_{N(0,1)} \leq Z_{\alpha/2} - d\sqrt{n}\right) =$$

$$\text{NORM.DIST}(\text{NORMSINV}(5\%/2) - (-2)/(3/\text{SQRT}(9)); 0; 1; 1) = 0.515967793$$

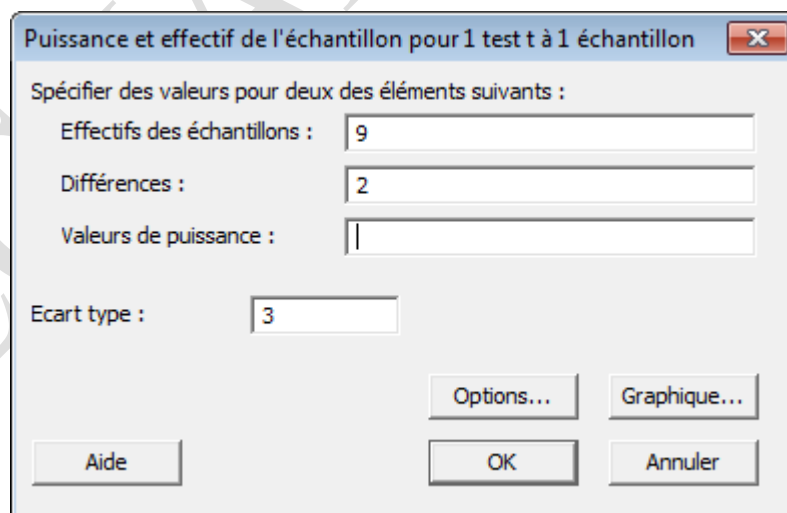
Ce qui devient alors:

$$1 - \beta = P\left(Y_{T(n-1)} \leq T_{\alpha/2, n-1} - d\sqrt{n}\right) =$$

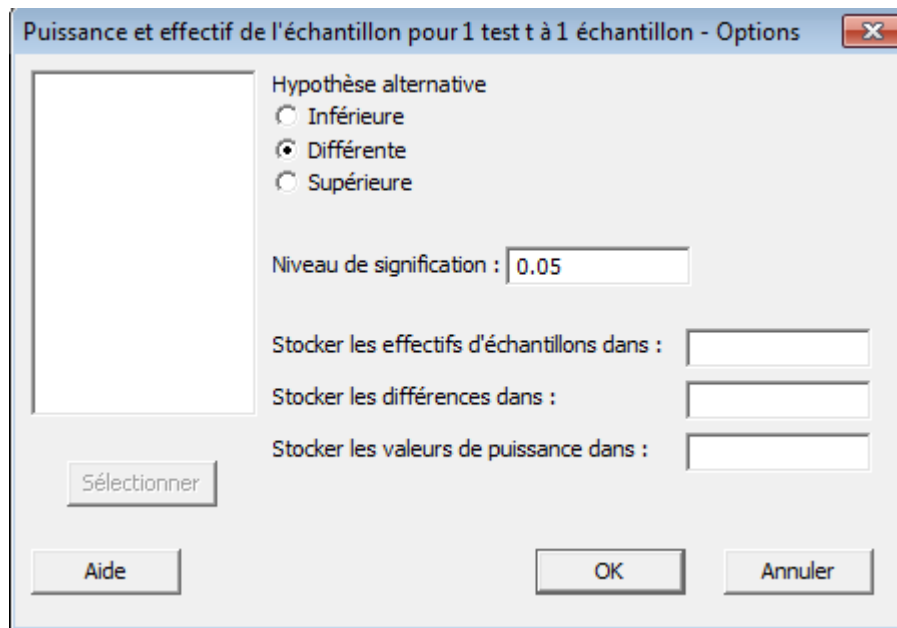
$$\text{T.DIST}(\text{T.INV}(5\%/2; 9-1) - (-2)/(3/\text{SQRT}(9)); 9-1; 1) = 0.383706664$$

Vérifions avec Minitab® Statistical Software:

Nous allons dans le menu **Stat/Puissance et effectif de l'échantillon/Test t à 1 échantillon...**:



Nous vérifions que nous sommes bien en bilatéral dans **Options...**:



Nous validons par **OK** pour avoir dans la fenêtre de session:

### **Puissance et effectif de l'échantillon**

Test t à 1 échantillon

Test de moyenne = nulle (contre moyenne non nulle)

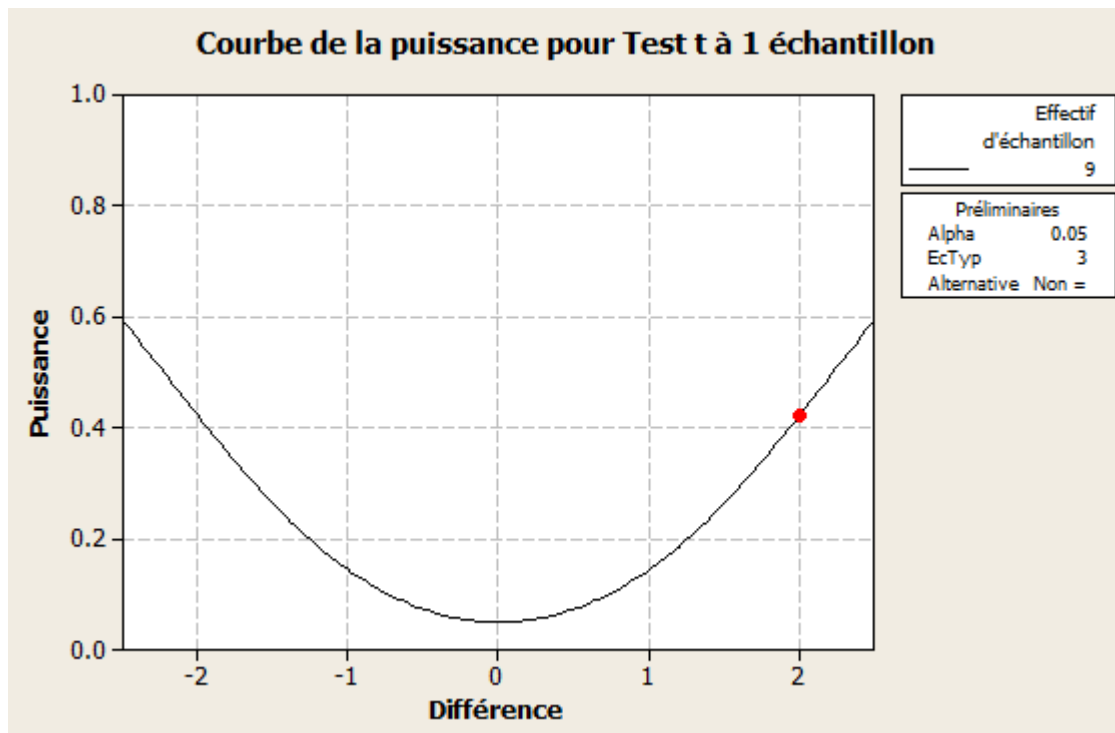
Calcul de puissance pour moyenne = nulle + différence

Alpha = 0.05 Ecart type supposé = 3

	Effectif		
Différence	d'échantillon	Puissance	
2	9	0.421052	

### **Courbe de la puissance pour Test t à 1 échantillon**

et le graphique de puissance:



Nous avons une valeur nettement différente (de l'ordre de 4% dans cet exemple particulier) que celle obtenu avec Microsoft Excel. Il faut donc que je regarde s'il s'agit d'une erreur dans les calculs théoriques ou si c'est une spécificité de l'algorithme utilisé par Minitab® Statistical Software.



## 10.23. Exercice 54.: Taille d'échantillon (effectif) d'un test t à 1 échantillon en bilatéral

Minitab® Statistical Software 15.1.1.0

Un gestionnaire responsable d'une chaîne de production veut savoir combien de temps passe en moyenne un employé à l'usinage sur une pièce donnée. Celui-ci aimerait être dans une estimation se trouvant dans les  $\pm 2$  minutes par rapport à la moyenne et l'écart-type d'usinage est estimé à 3 minutes avec une puissance du test de 80%. Quelle est la taille de l'échantillon nécessaire.

Pour faire ce calcul, nous utilisons la version studentisée de la relation:

$$1 - \beta = P\left(Y_{N(0,1)} \leq Z_{\alpha/2} - d\sqrt{n}\right) =$$

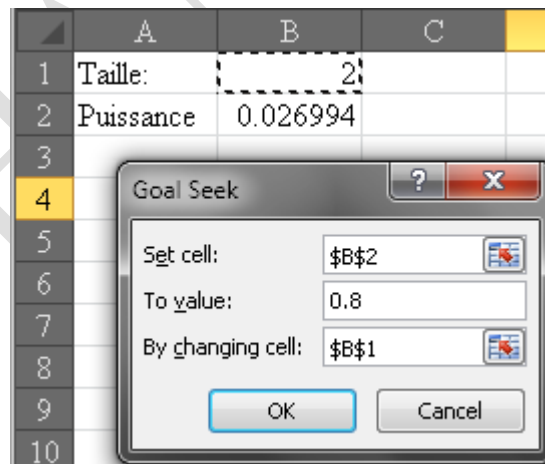
$$\text{NORM.DIST}(\text{NORMSINV}(5\%/2) - (-2)/(3/\text{SQRT}(9)); 0; 1; 1) = 0.515967793$$

Ce qui devient alors:

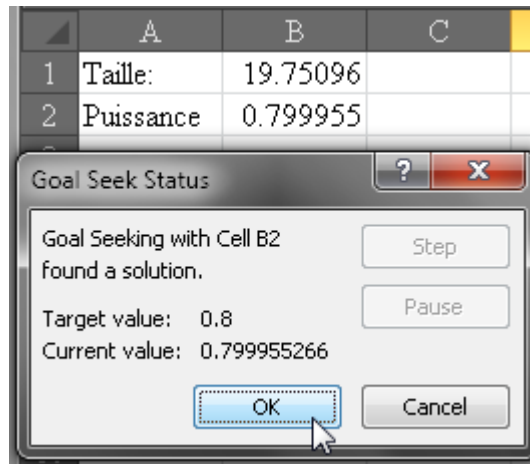
$$1 - \beta = P\left(Y_{T(n-1)} \leq T_{\alpha/2, n-1} - d\sqrt{n}\right) =$$

$$\text{T.DIST}(\text{T.INV}(5\%/2; B1-1) - (-2)/(3/\text{SQRT}(B1)); B1-1; 1)$$

et pour résoudre cette équation en A1 avec Microsoft Excel, nous avons utilisé l'outil Valeur Cible.

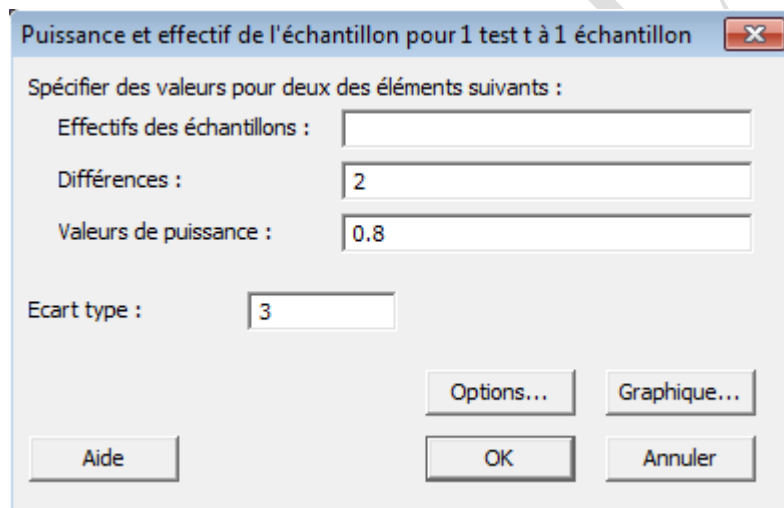


ce qui nous avait donné:

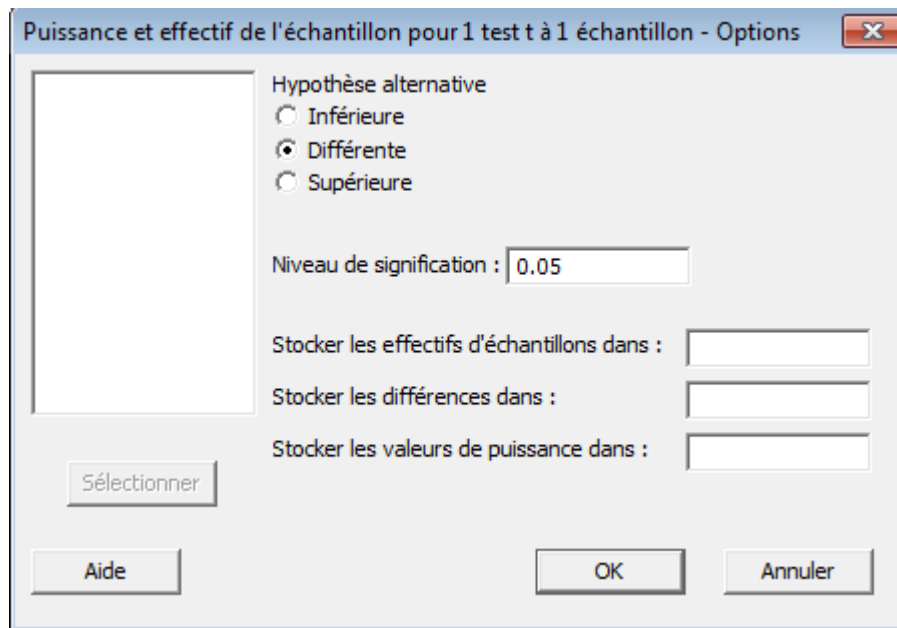


ce qui donne 20 à l'entier le plus proche! Vérifions avec Minitab® Statistical Software:

Nous allons dans le menu **Stat/Puissance et effectif de l'échantillon/Test t à 1 échantillon...**:



Nous vérifions que nous sommes bien en bilatéral dans **Options...**:



Nous validons par **OK** pour avoir dans la fenêtre de session:

### **Puissance et effectif de l'échantillon**

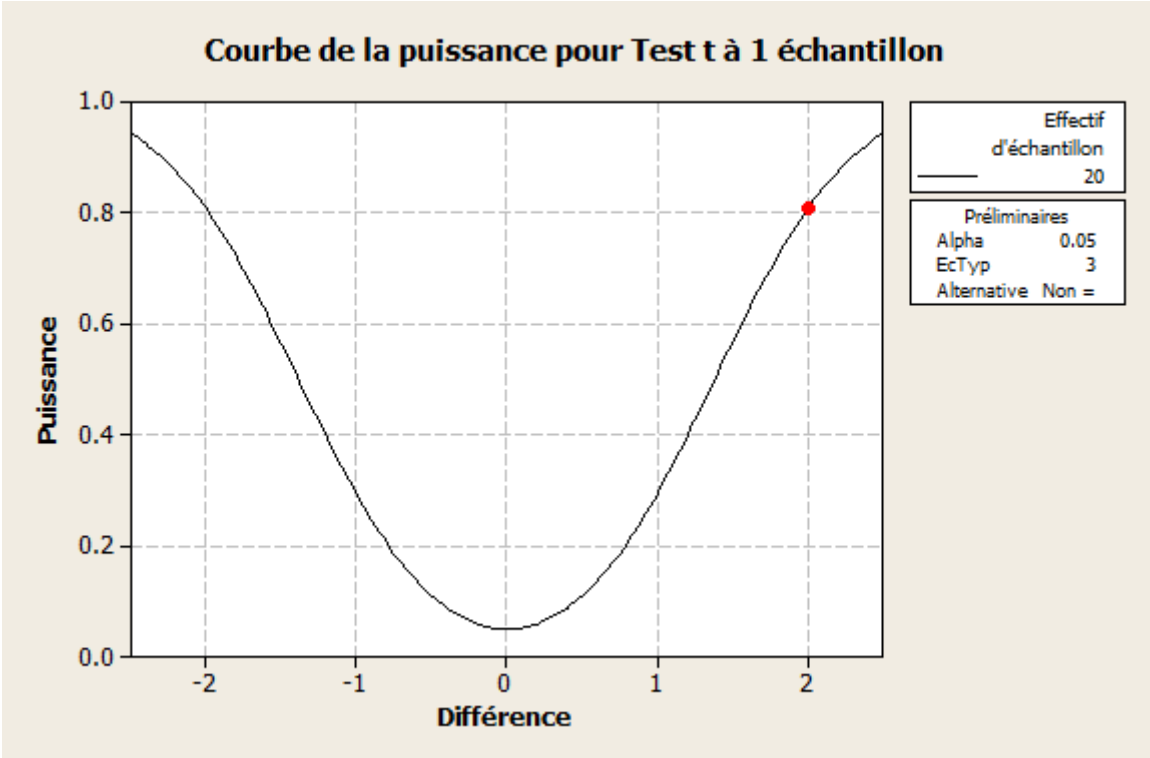
Test t à 1 échantillon

Test de moyenne = nulle (contre moyenne non nulle)  
 Calcul de puissance pour moyenne = nulle + différence  
 Alpha = 0.05 Ecart type supposé = 3

	Effectif	Puissance	Puissance
Différence	d'échantillon	cible	réelle
2	20	0.8	0.807292

### **Courbe de la puissance pour Test t à 1 échantillon**

et le graphique de puissance correspondant:



Nous retrouvons donc la même valeur qu'avec Microsoft Excel.

## 10.24. Exercice 55.: Intervalle de confiance de la moyenne (test t à un 1 échantillon)

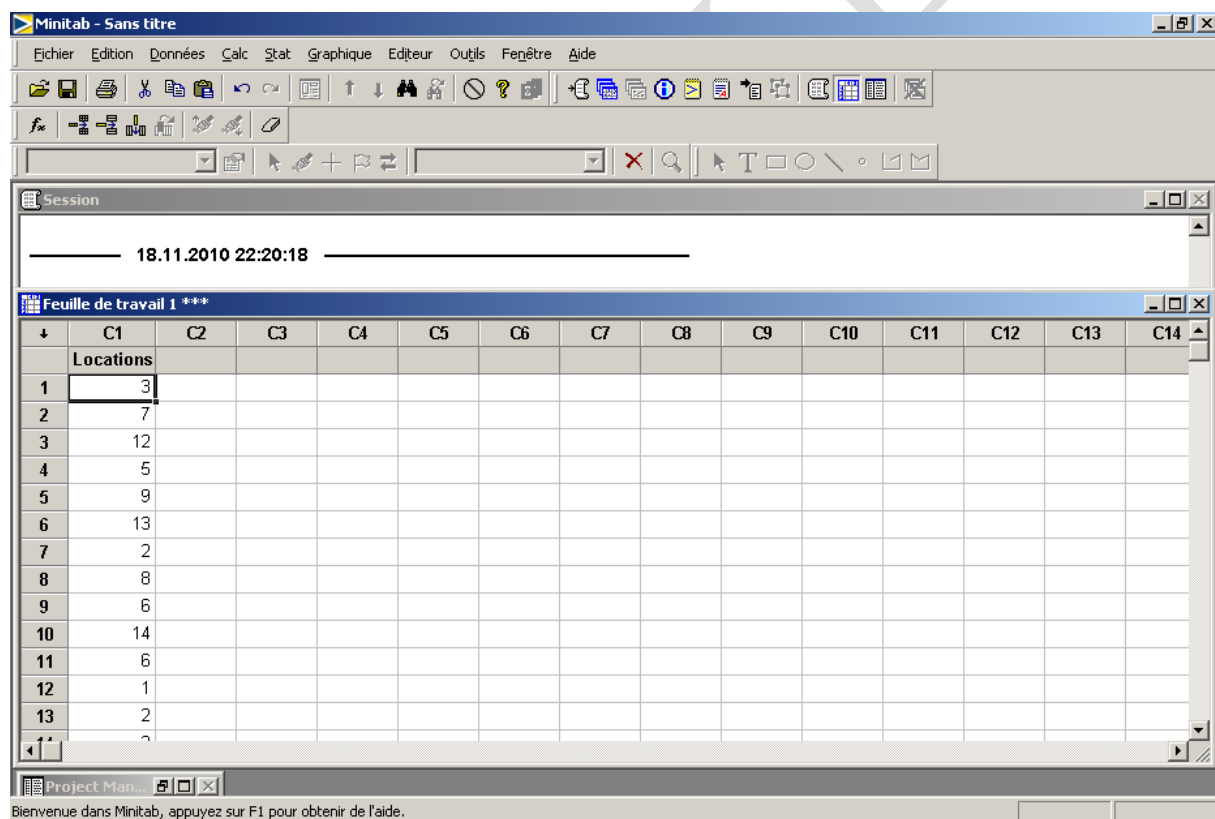
Minitab® Statistical Software 15.1.1.0

Il est donc courant dans le domaine de l'analyse statistique de devoir calculer l'intervalle de confiance de la moyenne. Nous allons reprendre le même exemple que dans le cours Microsoft Excel mais avec Minitab® Statistical Software.

Un gestionnaire d'une société de location de voitures souhaiterait estimer le nombre moyen de fois que des véhicules de luxe sont loués par mois. Il prend pour cela un échantillon aléatoire de 19 voitures de luxes et obtient la suite suivante de nombre de voitures louées:

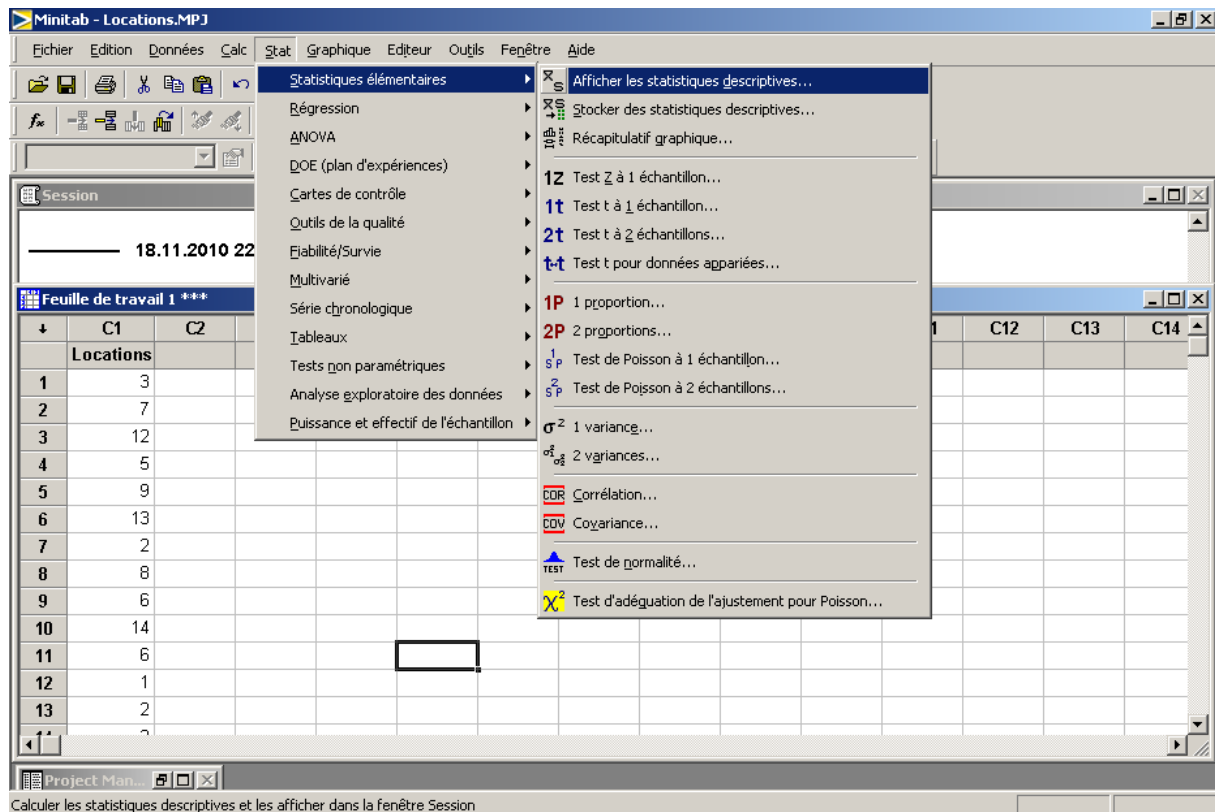
3,7,12,5,9,13,2,8,6,14,6,1,2,3,2,5,11,13,5

Nous devons déterminer l'intervalle à 95% de la moyenne. Donc dans Minitab® Statistical Software nous reportons ces valeurs ou ouvrons le fichier *Locations.mpj*:



	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
1	3													
2	7													
3	12													
4	5													
5	9													
6	13													
7	2													
8	8													
9	6													
10	14													
11	6													
12	1													
13	2													
14	3													
15	2													
16	5													
17	11													
18	13													
19	5													

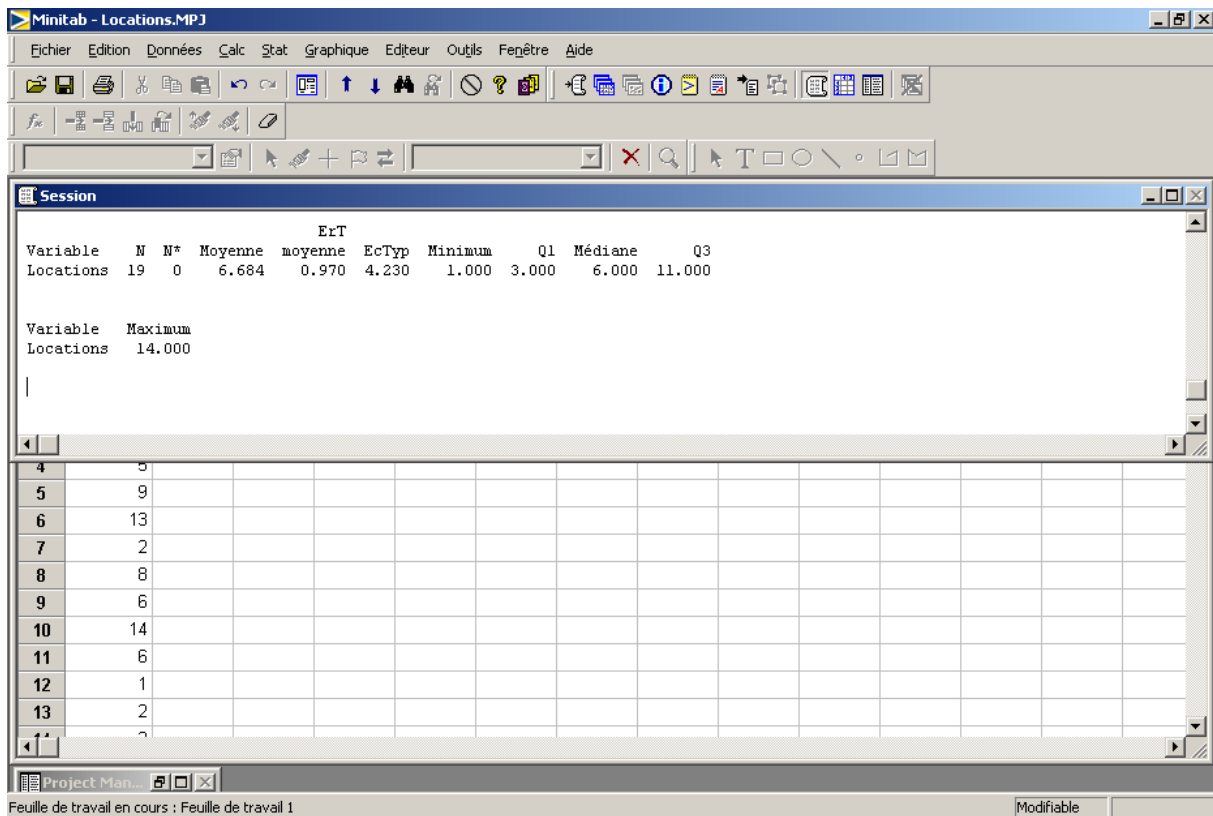
Pour comparer les résultats avec les calculs effectués à la main dans Microsoft Excel avons besoin de l'estimateur de maximum de vraisemblance de l'écart-type (donc c'est un intervalle de confiance avec variance théorique inconnue). Nous allons donc dans le menu **Stat/Statistiques élémentaires/Afficher les statistiques descriptives...**:



Nous avons alors:

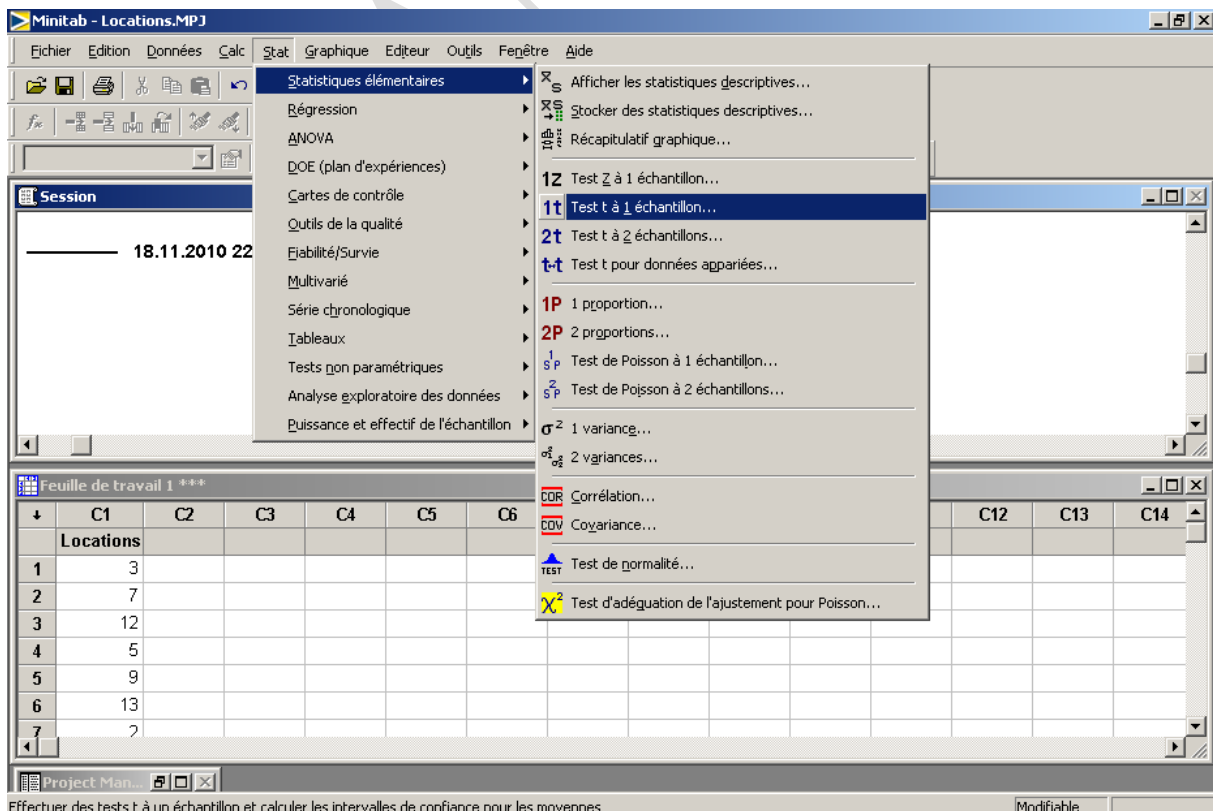


Nous validons par **OK** pour obtenir dans la fenêtre de session:

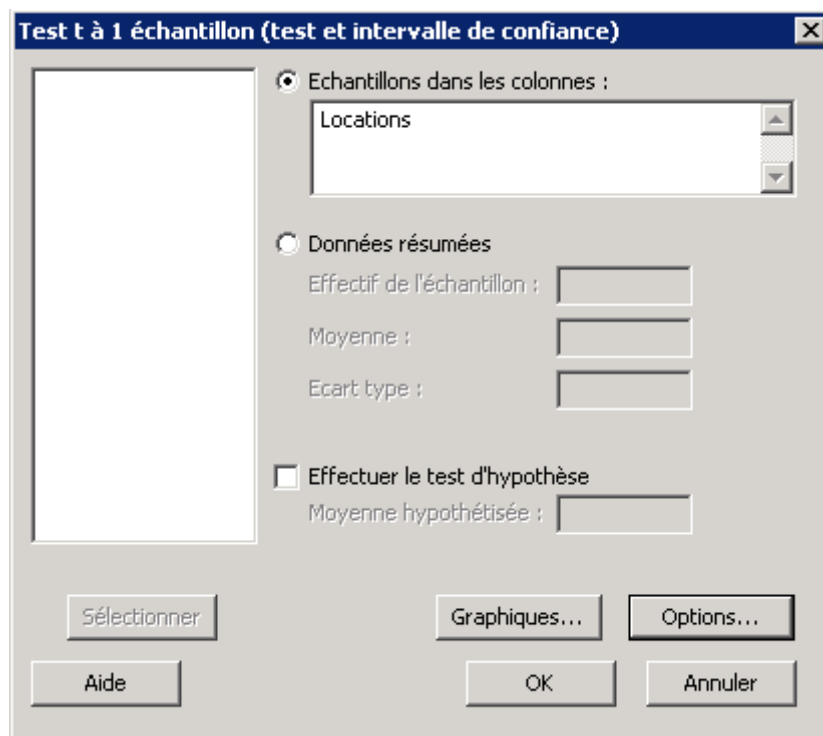


Nous y voyons l'écart-type de 4.23 ce qui correspond parfaitement à ce que nous avons avec les calculs à la main.

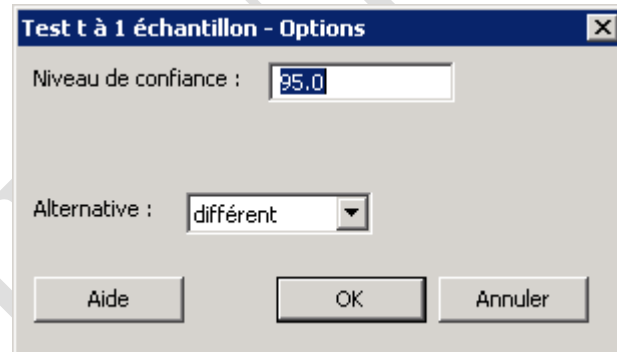
Maintenant, nous allons dans le menu **Stat/Statistiques élémentaires/Test t à un échantillon...**:



Pour avoir:



et pour vérifier que nous sommes bien en bilatéral à un niveau de 95%, nous cliquons sur le bouton **Options...**:



Et nous validons par **OK**:



**Minitab - Sans titre**

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

Session

Bienvenue dans Minitab, appuyez sur F1 pour obtenir de l'aide.

**Test T à un échantillon : Locations**

Variable	N	Moyenne	EcTyp	ErT moyenne	IC à 95 % (4.645; 8.723)
Locations	19	6.684	4.230	0.970	

Feuille de travail 1 \*\*\*

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
	Locations											
1	3											
2	7											
3	12											
4	5											
5	9											
6	13											
7	2											
8	8											

Feuille de travail en cours : Feuille de travail 1

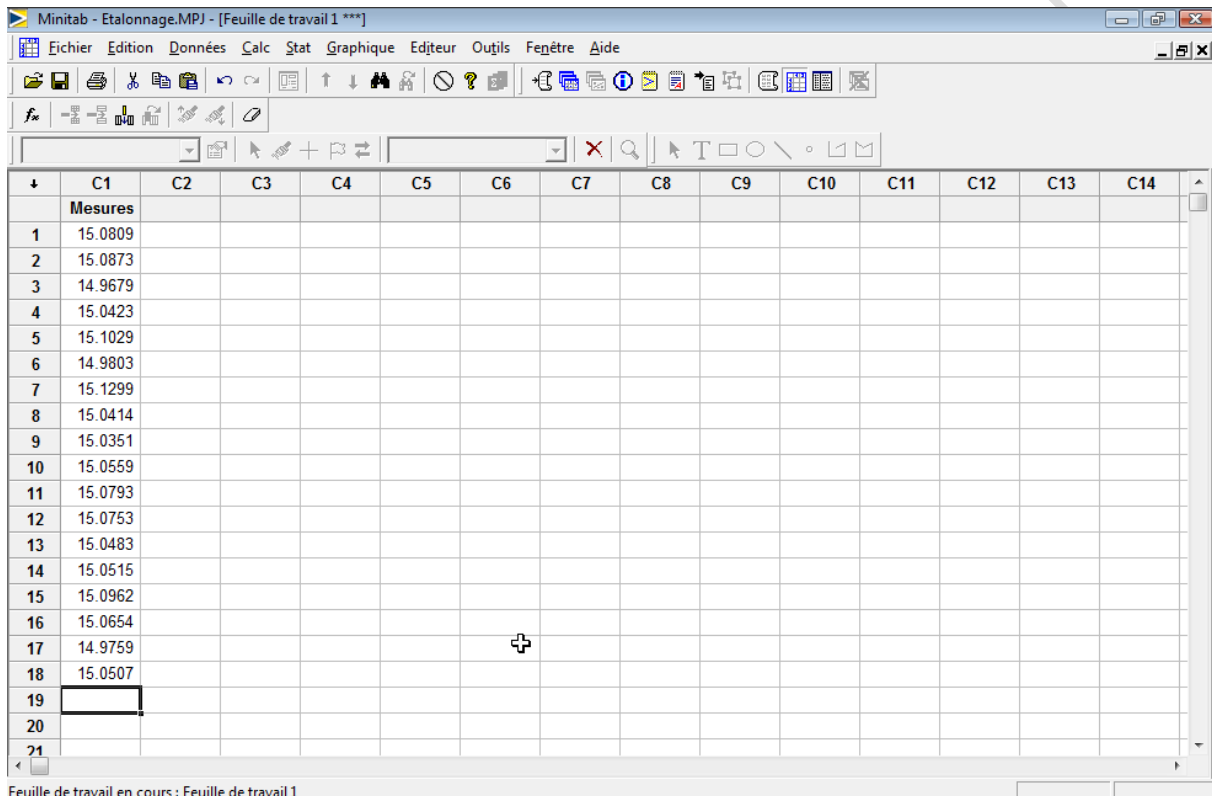
Modifiable

Nous retrouvons le même intervalle de confiance qu'avec Microsoft Excel. Évidemment il n'y a aucune  $p$ -value d'indiqué puisque nous n'avons pas imposé de moyenne hypothétisée dans la boîte de dialogue principale. Et nous pouvons faire à nouveau les mêmes exemples que pour le test  $Z$  mais ce serait un peu trop répétitif donc nous nous en passerons...

## 10.25. Exercice 56.: Test t-Student de différence de la moyenne bilatéral d'un échantillon

Minitab® Statistical Software 15.1.1.0

Nous devons étalonner un appareil à l'aide d'une pièce étalon dont nous savons que le diamètre réel est de 15 microns. Nous avons obtenu les mesures suivantes (fichier *Etalonnage.mpj*):

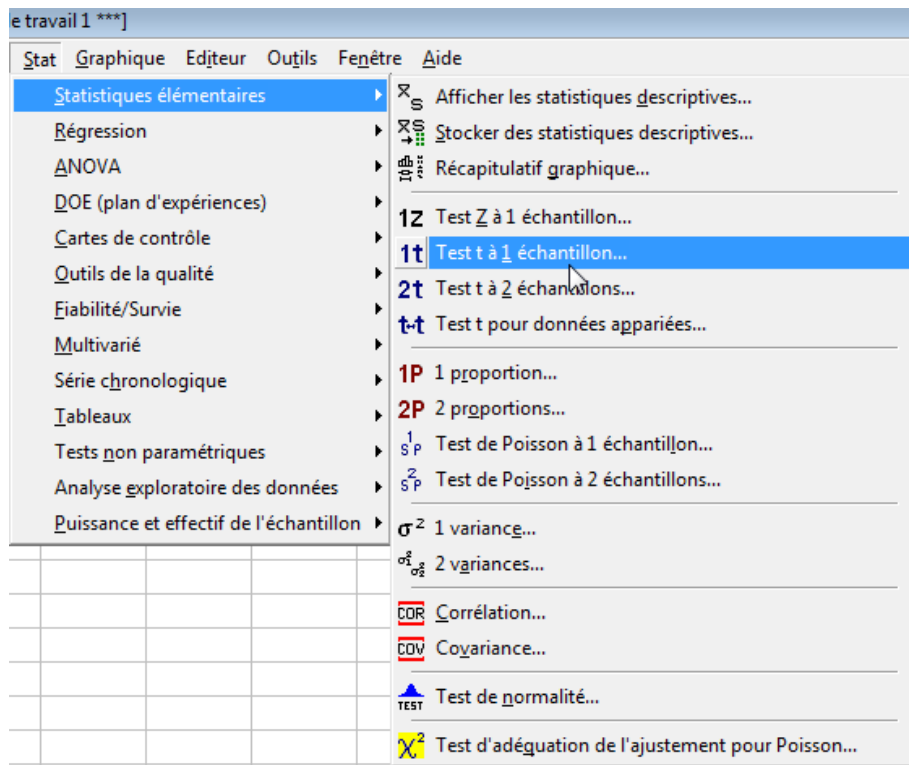


The screenshot shows the Minitab Statistical Software interface. The main window displays a data table with 18 rows of measurements in column C1. The table is titled 'Mesures' and has columns labeled C1 through C14. The data values are as follows:

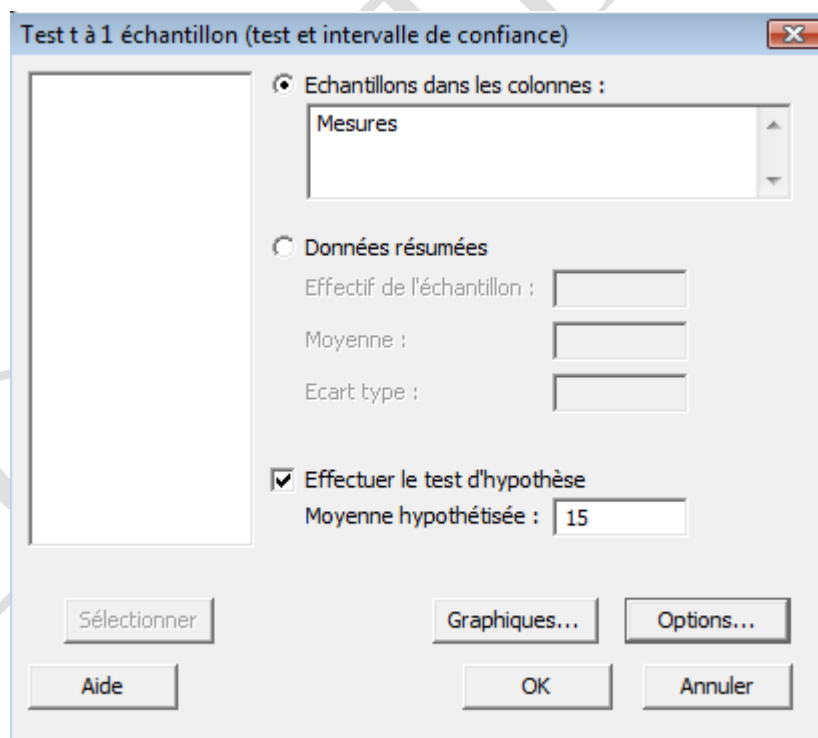
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Mesures													
1	15.0809													
2	15.0873													
3	14.9679													
4	15.0423													
5	15.1029													
6	14.9803													
7	15.1299													
8	15.0414													
9	15.0351													
10	15.0559													
11	15.0793													
12	15.0753													
13	15.0483													
14	15.0515													
15	15.0962													
16	15.0654													
17	14.9759													
18	15.0507													
19														
20														
21														

Nous souhaiterions vérifier un intervalle de confiance de 95% que nous sommes proches la moyenne vraie (pas trop différent).

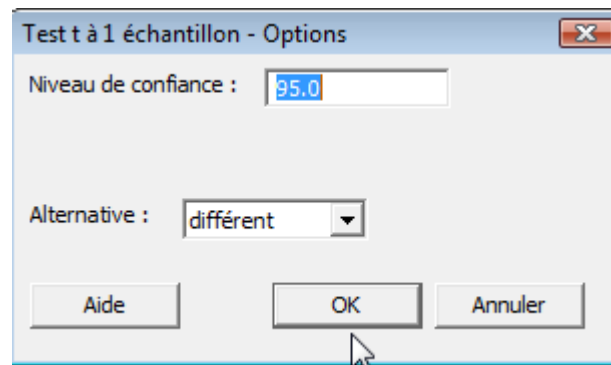
Nous allons dans le menu **Stat/Statistiques élémentaires/Test t à 1 échantillon...**:



Vient alors :



Nous cliquons sur le bouton **Options...** pour vérifier que tout soit conforme:



Nous obtenons:

### Test T à un échantillon : Mesures

Test de  $\mu = 15$  en fonction de la différence 15

Variable	N	Moyenne	EcTyp	ErT moyenne	IC à 95 %	T	P
Mesures	18	15.0537	0.0438	0.0103	(15.0319; 15.0755)	5.20	0.000

L'hypothèse nulle étant que la moyenne arithmétique mesurée est égale 15 microns (l'hypothèse alternative étant qu'elle ne le soit pas...) et donc rejetée. Il y a donc une différence significative.

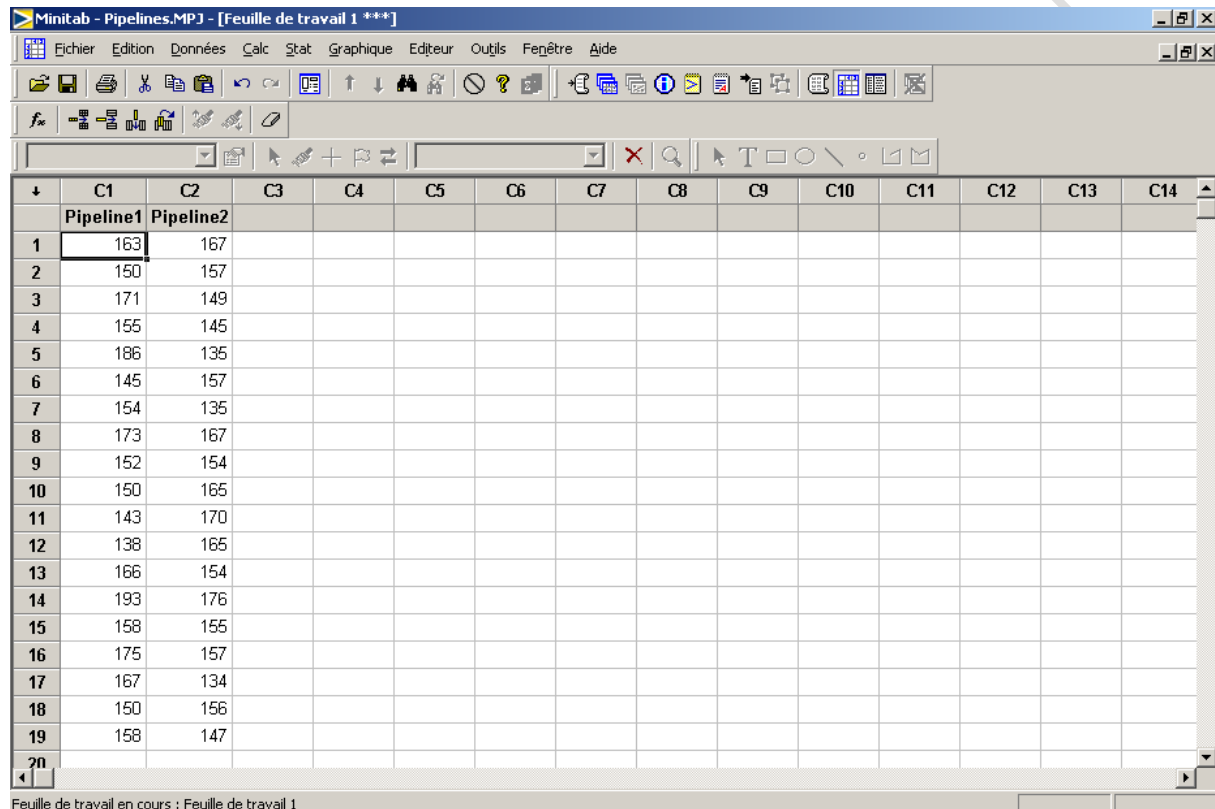
Nous voyons cependant que nous n'obtenons pas les mêmes intervalles de confiance (IC) qu'avec Microsoft Excel ni la même valeur pour  $T$ . Cependant, selon moi il y a une aberration ici avec Minitab® Statistical Software. Si l'IC à 95% était bien de 15.0319 et 15.0775, alors 15.0537 se trouve bien dans l'intervalle et nous ne pourrions pas rejeter l'hypothèse. Or, la  $p$ -value nous montre bien qu'il faut la rejeter. De plus, l'IC devrait être centrée sur la moyenne théorique de 15...

Affaire à suivre... pour savoir si c'est moi qui ai fait une erreur de compréhension (et donc in extenso des erreurs de calcul dans Microsoft Excel) ou si c'est Minitab® Statistical Software.

## 10.26. Exercice 57.: Test t-Student homoscédastique bilatéral d'égalité de la moyenne

Minitab® Statistical Software 15.1.1.0

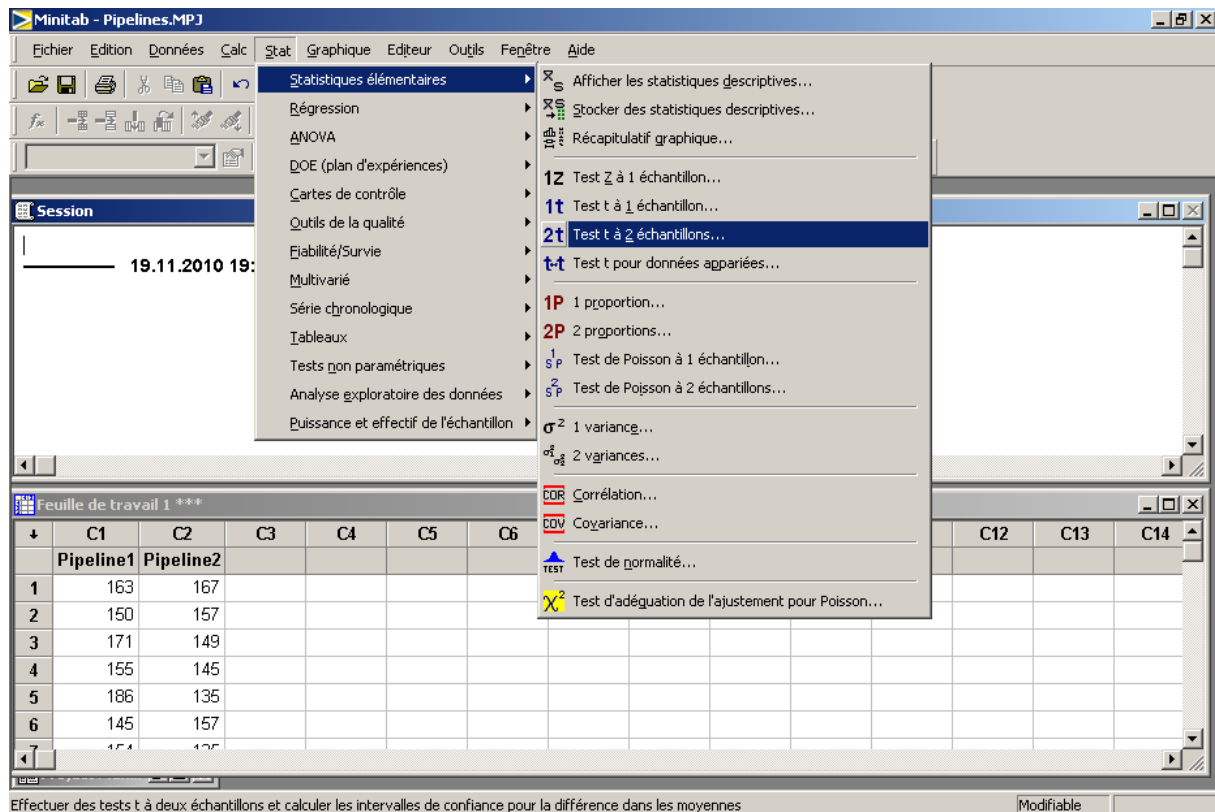
Le débit d'eau en litre par minutes entre plusieurs points équidistants de deux pipe-lines 1 et 2 a été mesuré. Le tableau suivant (il s'agit du même tableau que celui utilisé pour l'exemple du test t pour données appariées) a été obtenu (fichier *Pipelines.mpj*):



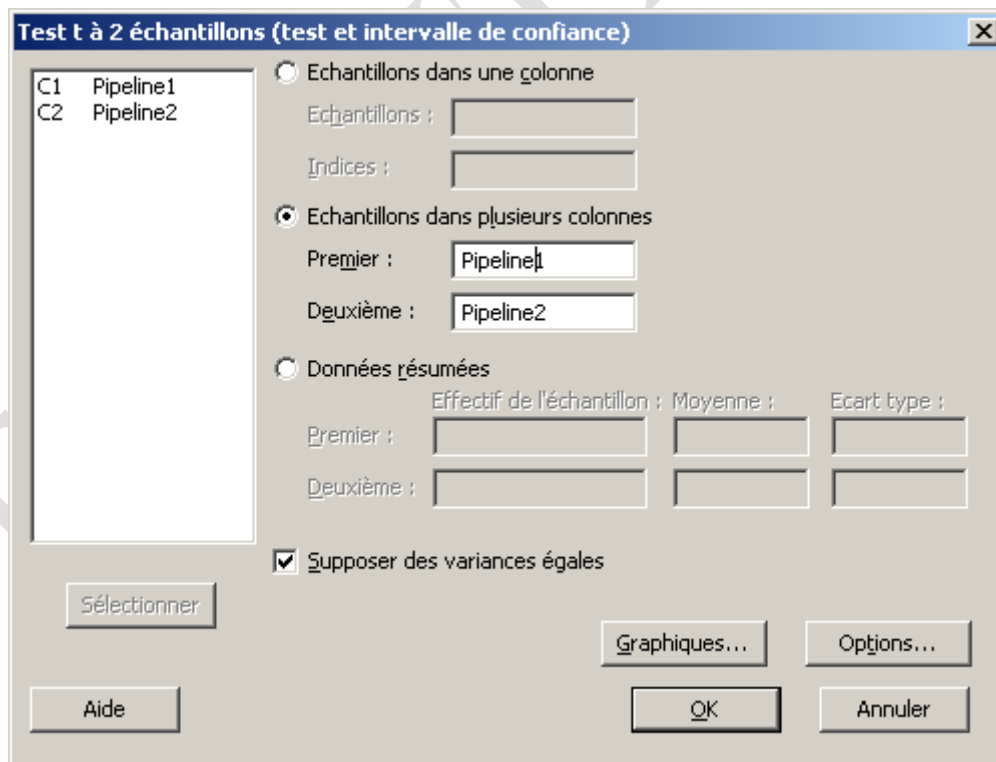
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Pipeline1	Pipeline2												
1	163	167												
2	150	157												
3	171	149												
4	155	145												
5	186	135												
6	145	157												
7	154	135												
8	173	167												
9	152	154												
10	150	165												
11	143	170												
12	138	165												
13	166	154												
14	193	176												
15	158	155												
16	175	157												
17	167	134												
18	150	156												
19	158	147												
20														

Un ingénieur souhaite déterminer s'il y a statistiquement une différence significative entre les deux pipelines à un niveau de confiance de 95% en supposant les variances égales. Nous allons voir si nous retrouvons les mêmes résultats que dans le cours Microsoft Excel et dans le cours de statistique théorique.

Nous allons dans le menu **Stat/Statistiques élémentaires/Test t à 2 échantillons...**:



Nous avons alors:



Et en validant par **OK** nous obtenons effectivement les mêmes résultats que dans Microsoft Excel et dans le cours de statistique théorique:

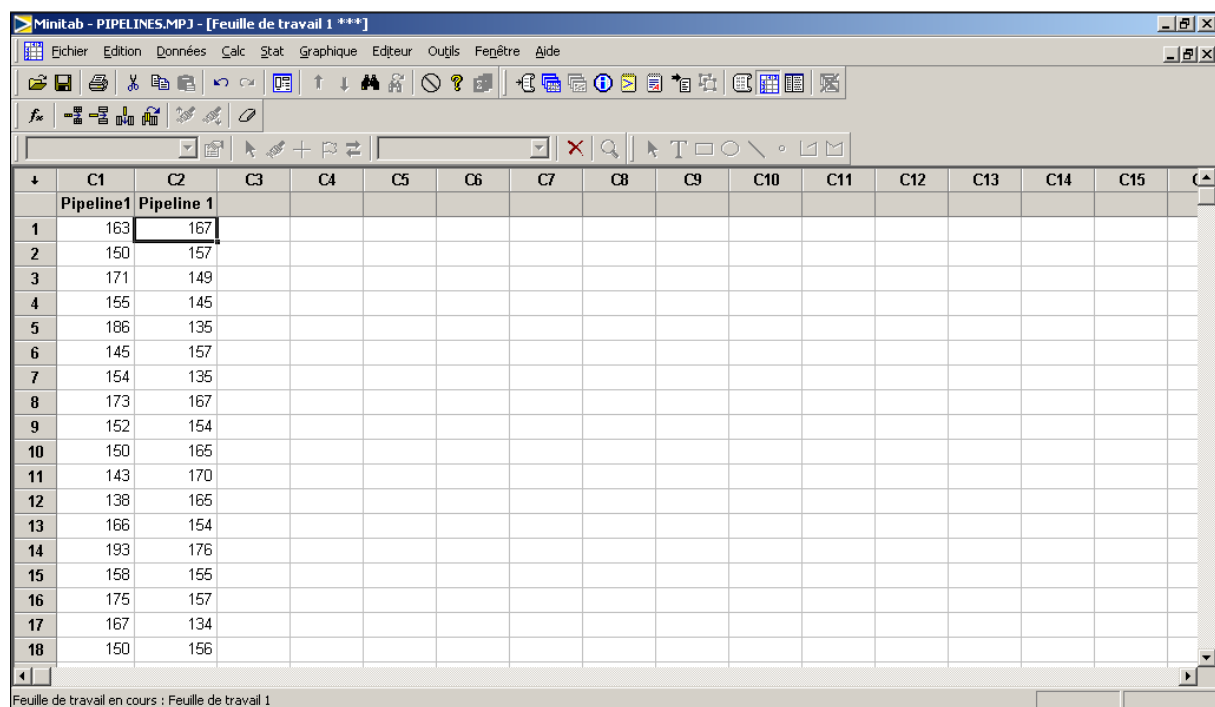
Différence = mu (Pipeline1) - mu (Pipeline2)  
 Estimation de la différence : 5.37  
 Limites de confiance (à 95 %) pour la différence : (-3.41 ; 14.15)  
 Test t de la différence = 0 (en fonction de la différence) : Valeur de T = 1.24  
 Valeur de p = 0.223 DL = 36  
 Les deux utilisent l'écart type regroupé = 13.3463

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Pipeline1	Pipeline2												
1	163	167												
2	150	157												
3	171	149												
4	155	145												
5	186	135												
6	145	157												
7	174	127												

## 10.27. Exercice 58.: Test t-Student pour données appariées

Minitab® Statistical Software 15.1.1.0

Le débit d'eau en litre par minutes entre plusieurs points équidistants d'un même pipeline avant et après révision a été mesuré. Le tableau suivant a été obtenu:

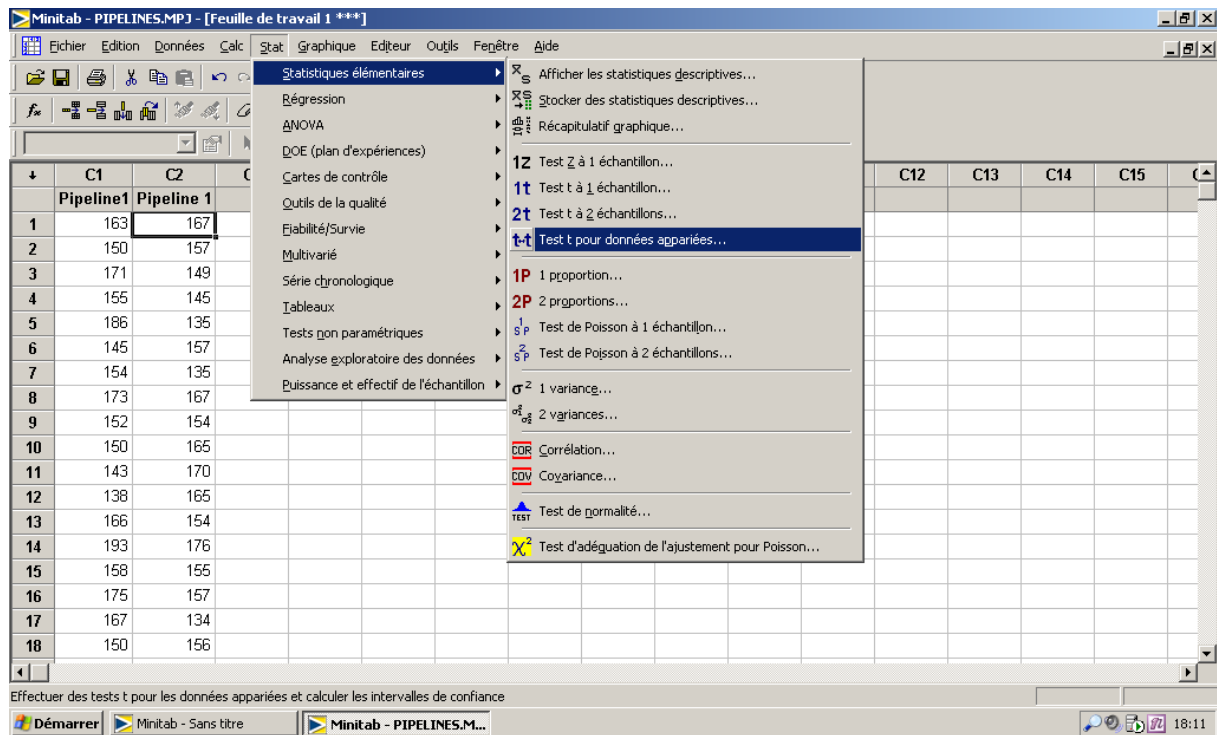


	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
	Pipeline1	Pipeline 1													
1	163	167													
2	150	157													
3	171	149													
4	155	145													
5	186	135													
6	145	157													
7	154	135													
8	173	167													
9	152	154													
10	150	165													
11	143	170													
12	138	165													
13	166	154													
14	193	176													
15	158	155													
16	175	157													
17	167	134													
18	150	156													

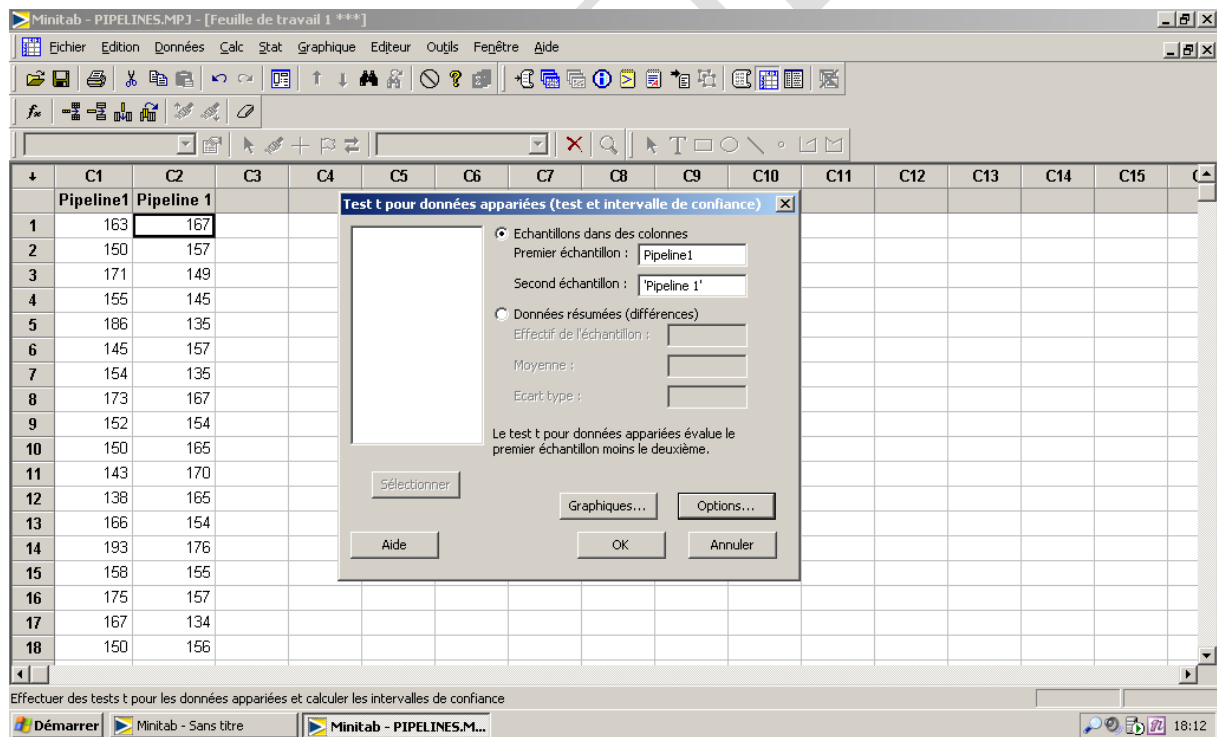
Un ingénieur souhaite déterminer s'il y a statistiquement une différence significative entre les pipelines avant et après révision à un niveau de confiance de 95% en bilatéral (sous l'hypothèse que les données sont paires).

Pour déterminer cela, nous allons dans le menu **Stat/Statistiques élémentaires/Test t pour données appariées...**:

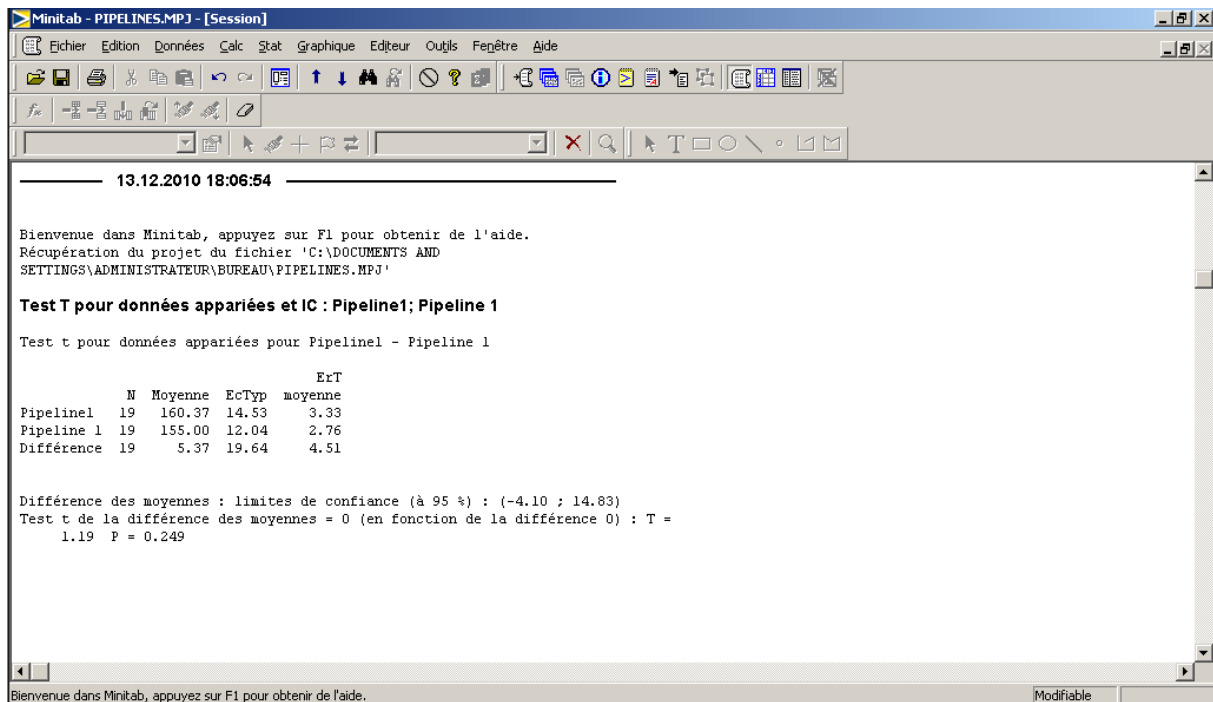




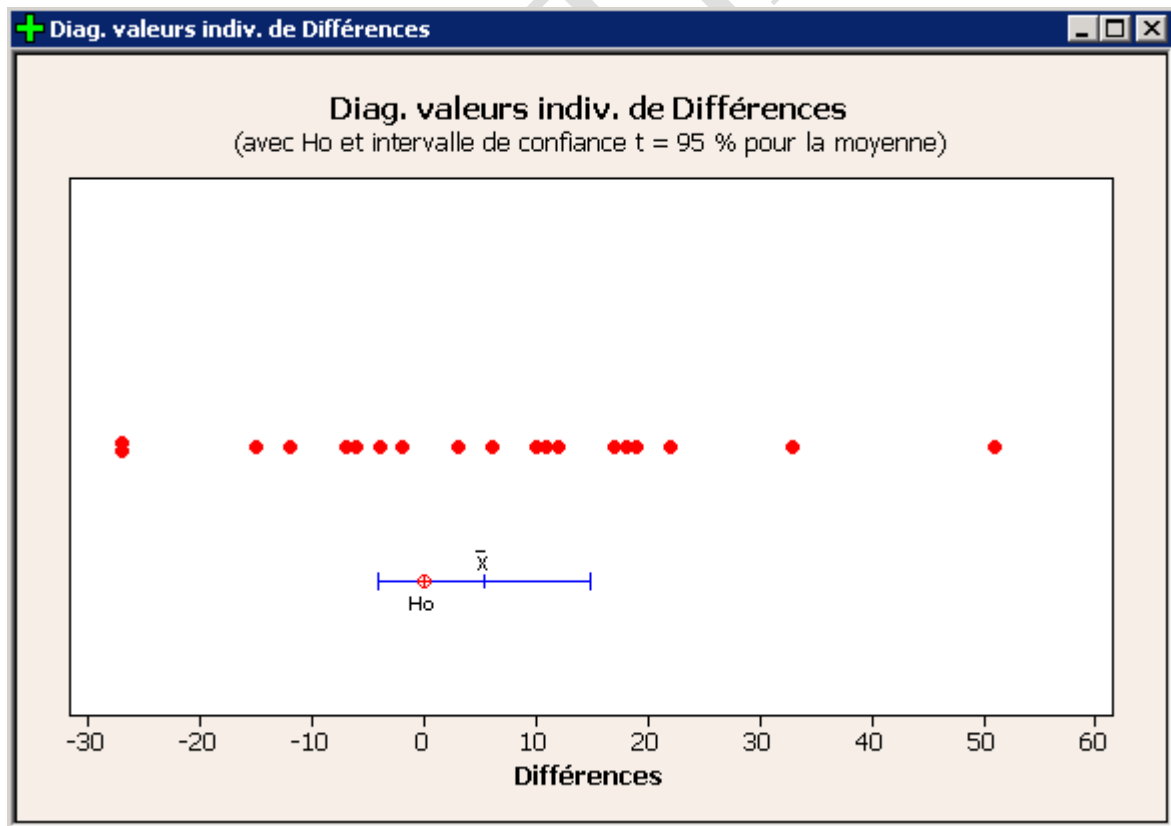
Vient alors:

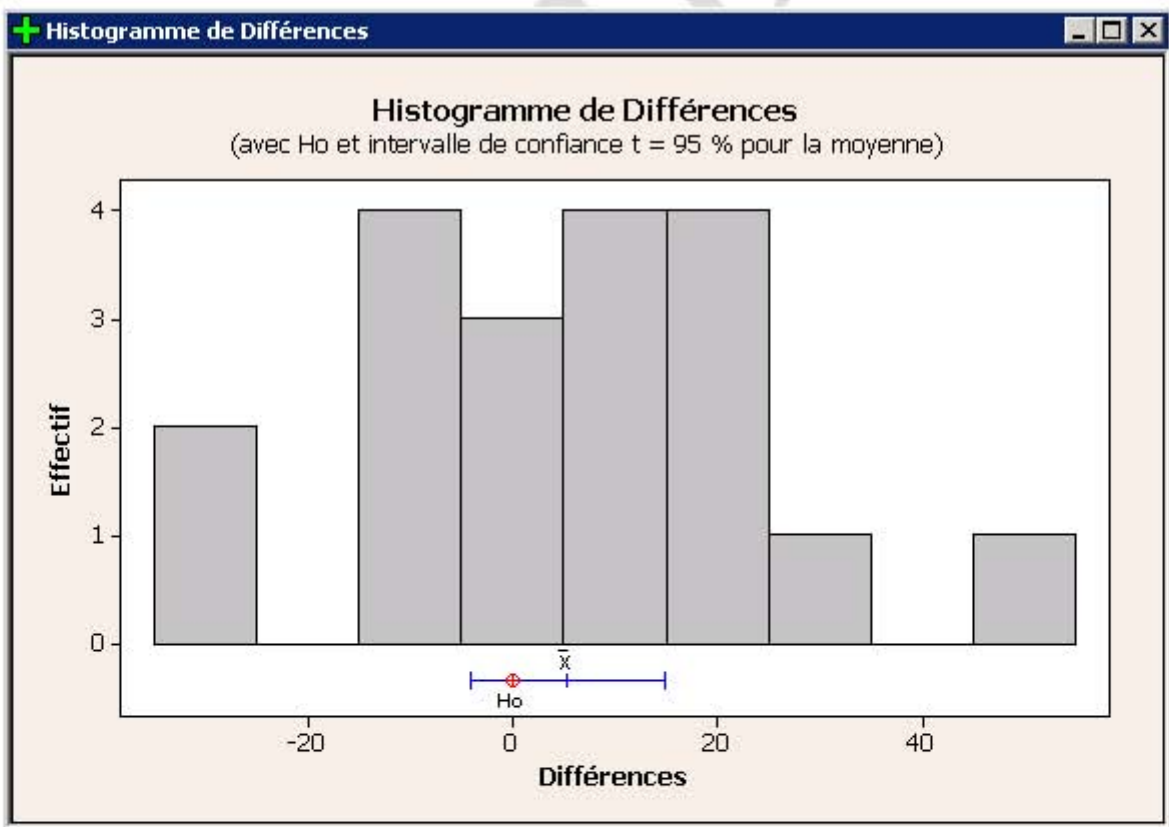
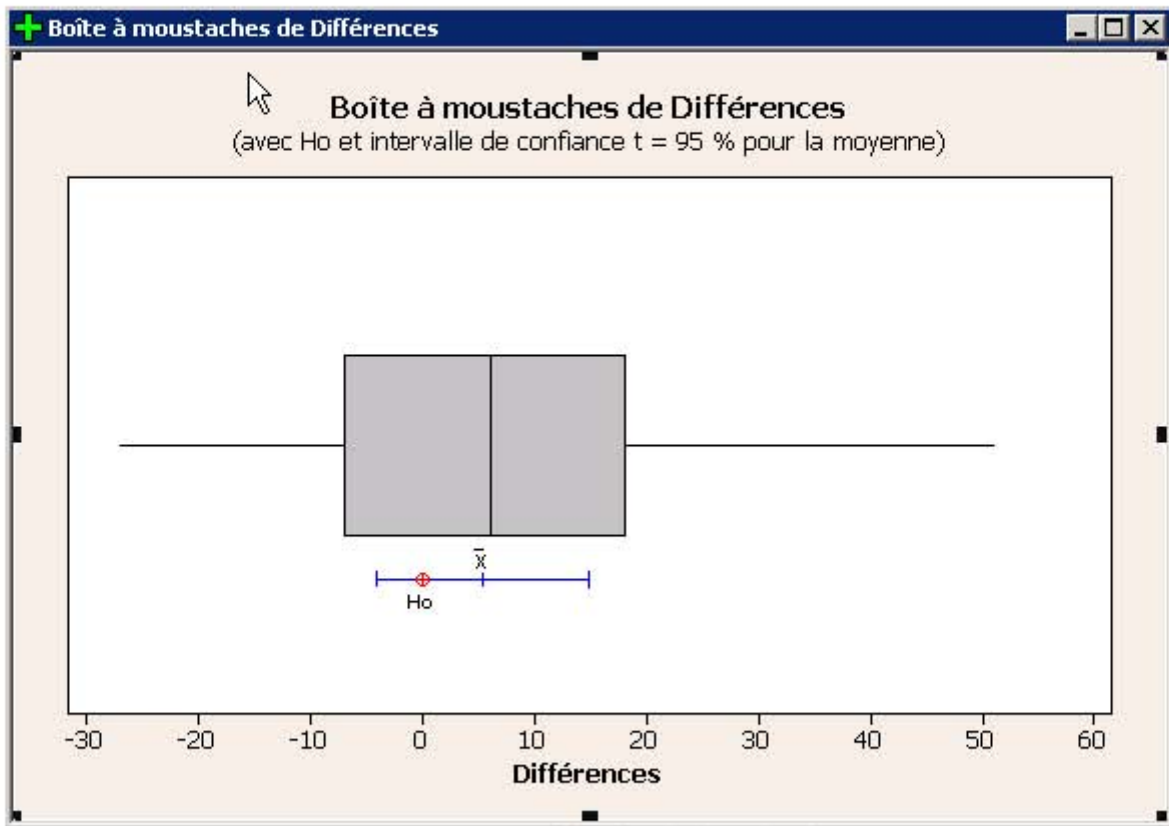


Nous validons par **OK**:



Nous retrouvons donc exactement les valeurs calculées à la main dans le cours de statistique théorique et avec Microsoft Excel! Nous avons aussi les graphiques suivant si nous les activons:

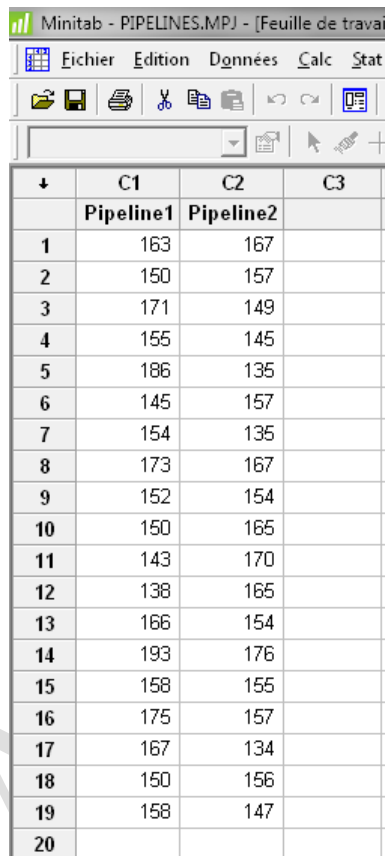




## 10.28. Exercice 59.: Test t-Student hétéroscédastique bilatéral d'égalité de la moyenne (test t de Welch)

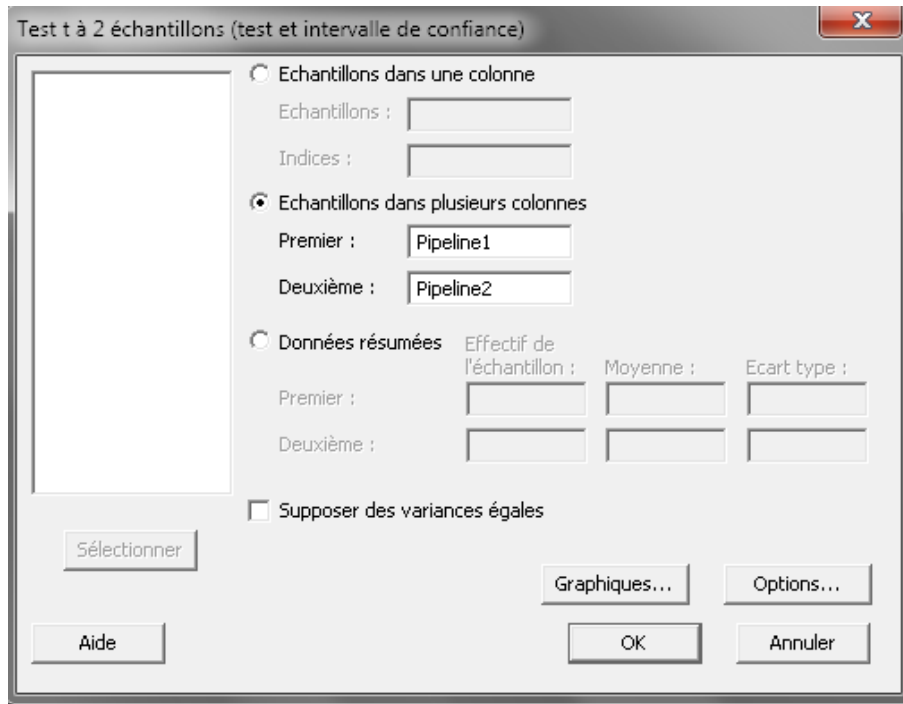
Minitab® Statistical Software 16.1.1.0

Nous repartons des mêmes données que dans le cas de l'exemple homoscdastique:



	C1	C2	C3
	Pipeline1	Pipeline2	
1	163	167	
2	150	157	
3	171	149	
4	155	145	
5	186	135	
6	145	157	
7	154	135	
8	173	167	
9	152	154	
10	150	165	
11	143	170	
12	138	165	
13	166	154	
14	193	176	
15	158	155	
16	175	157	
17	167	134	
18	150	156	
19	158	147	
20			

Et comme avant nous lançons le test à 2 échantillons à la différence que nous laissons **Supposer des variances égales décochées**:



Nous validons par **OK** pour obtenir dans la fenêtre de session:

**Test T à 2 échantillons et IC : Pipeline1; Pipeline2**

Test t à 2 échantillons de Pipelinel et Pipeline2

	N	Moyenne	EcTyp	ErT moyenne
Pipeline1	19	160.4	14.5	3.3
Pipeline2	19	155.0	12.0	2.8

Différence = mu (Pipelinel) - mu (Pipeline2)  
 Estimation de la différence : 5.37  
 Limites de confiance (à 95 %) pour la différence : (-3.43 ; 14.17)  
 Test t de la différence = 0 (contre pas =) : Valeur de T = 1.24 Valeur de p = 0.224 DL = 34

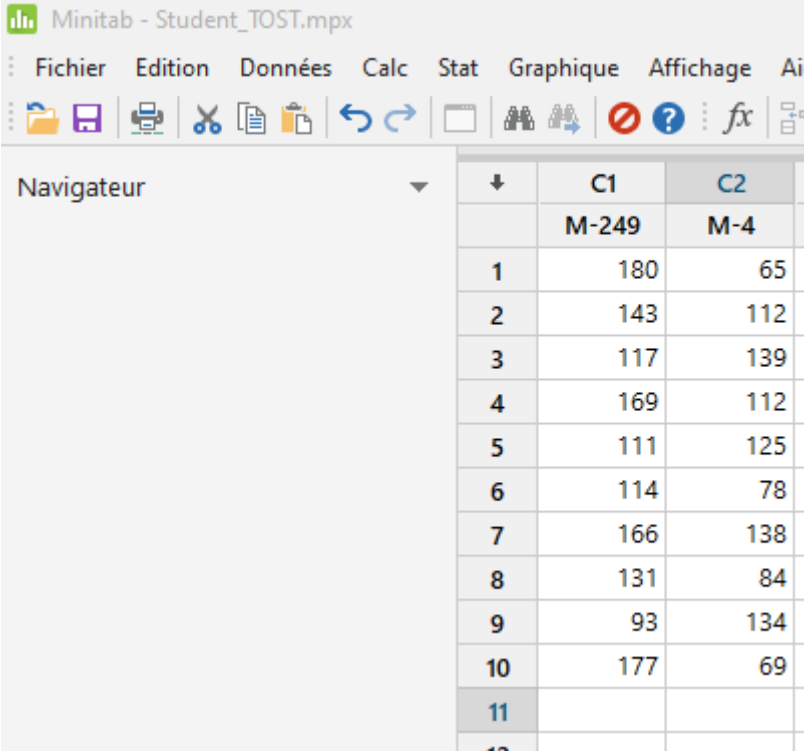
Et les valeurs (et particulièrement les DL), correspondent bien à la relation démontrée dans le cours théorique puisqu'elles prennent la valeur entière de:

$$f \cong \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{n_1 - 1} + \frac{(S_2^2/n_1)^2}{n_2 - 1}}$$

## 10.29. Exercice 60.: TOST (Two One-Sided Test) de Student

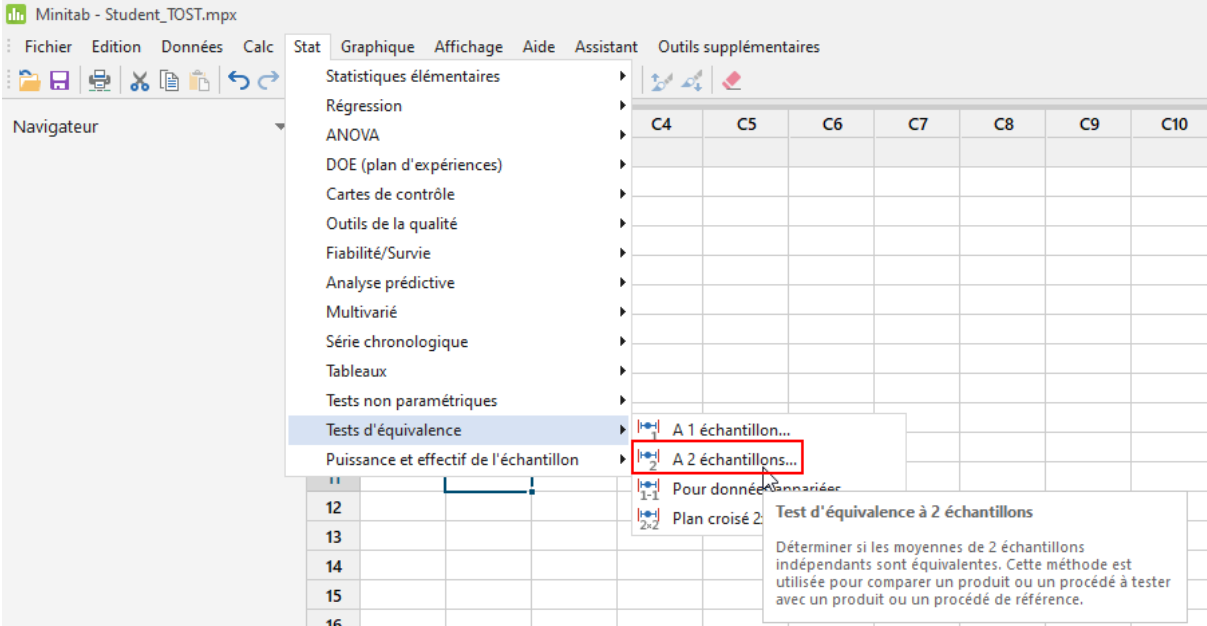
Minitab® Statistical Software 19.2020.1

Nous allons reproduire ici le fameux test TOST de Student et particulièrement l'exemple calculé à la main dans le cours théorique. Nous partons donc des données suivantes (fichier *Student\_TOST.mpx*):



	C1	C2
	M-249	M-4
1	180	65
2	143	112
3	117	139
4	169	112
5	111	125
6	114	78
7	166	138
8	131	84
9	93	134
10	177	69
11		
12		

Ensuite nous allons dans le menu **Stat/Test d'équivalence/A 2 échantillons**:



Stat

- Statistiques élémentaires
- Régression
- ANOVA
- DOE (plan d'expériences)
- Cartes de contrôle
- Outils de la qualité
- Fiabilité/Survie
- Analyse prédictive
- Multivarié
- Série chronologique
- Tableaux
- Tests non paramétriques
- Tests d'équivalence**
  - A 1 échantillon...
  - A 2 échantillons...**
  - Pour données appariées
  - Plan croisé 2
- Puissance et effectif de l'échantillon

**Test d'équivalence à 2 échantillons**  
Déterminer si les moyennes de 2 échantillons indépendants sont équivalentes. Cette méthode est utilisée pour comparer un produit ou un procédé à tester avec un produit ou un procédé de référence.

Nous prenons les options ad-hoc:

Test d'équivalence à 2 échantillons

C1	M-249
C2	M-4

Echantillons dans plusieurs colonnes

Echantillon de test : M-249

Echantillon de référence : M-4

Hypothèse sur : Moyenne du test - moyenne de référence

Que souhaitez-vous **Moyenne du test - moyenne de référence**

Limite inférieure < moyenne du test - moyenne de référence < limite supérieure

Limite inférieure :

Limite supérieure :   Multiplier par la moyenne de référence

Options... Graphiques... Résultats...

Aide OK Annuler

et aussi:

Test d'équivalence à 2 échantillons

Echantillons dans plusieurs colonnes

Echantillon de test : M-249

Echantillon de référence : M-4

Hypothèse sur : Moyenne du test - moyenne de référence

Que souhaitez-vous déterminer ? (Hypothèse alternative)

Limite inférieure < moyenne du test - moyenne de référence < limite supérieure

**Limite inférieure < moyenne du test - moyenne de référence < limite supérieure**

Moyenne du test > moyenne de référence

Moyenne du test < moyenne de référence

Moyenne du test - moyenne de référence > limite inférieure

Moyenne du test - moyenne de référence < limite supérieure

Options... Graphiques... Résultats...

Aide Sélectionner

Pour les limites nous prendrons bien évidemment les mêmes que dans le cours théorique:

The dialog box 'Test d'équivalence à 2 échantillons' contains the following settings:

- Echantillons dans plusieurs colonnes: [dropdown]
- Echantillon de test: 'M-249'
- Echantillon de référence: 'M-4'
- Hypothèse sur: 'Moyenne du test - moyenne de référence'
- Que souhaitez-vous déterminer ? (Hypothèse alternative): 'Limite inférieure < moyenne du test - moyenne de référence < limite supérieure'
- Limite inférieure: '-20'
- Limite supérieure: '20'
- Multiplier par la moyenne de référence

Buttons: Sélectionner, Aide, Options..., Graphiques..., Résultats..., OK, Annuler.

Voyons ce que propose le bouton **Options...**:

The 'Options' dialog box contains the following settings:

- Niveau de risque de déclaration d'équivalence à tort (alpha): '0.05'
- Utiliser l'intervalle de confiance (1 - 2 alpha) x 100 %
- Supposer les variances égales

Buttons: Aide, OK, Annuler.

et le bouton **Graphiques...**:

The 'Graphiques' dialog box contains the following settings:

- Diagramme d'équivalence
- Histogramme
- Diagramme des valeurs individuelles
- Boîte à moustaches

Buttons: Aide, OK, Annuler.

Nous obtenons alors dans la console exactement les mêmes valeurs que celles calculées dans le cours théorique avec les mêmes conclusions:



Test d'équivalence à deux ...

FEUILLE DE TRAVAIL 1

Test d'équivalence à deux échantillons : M-249, M-4

Méthode

Moyenne du test = moyenne de M-249  
 Moyenne de référence = moyenne de M-4  
 Les variances n'ont pas été supposées égales pour l'analyse.

Statistiques descriptives

Variable	N	Moyenne	EcTyp	ErT moyenne
M-249	10	140.10	31.324	9.9056
M-4	10	105.60	29.148	9.2174

Différence : moyenne(M-249) - moyenne(M-4)

Différence	ErT	IC 95% pour équivalence	Intervalle d'équivalence
34.500	13.531	(0, 58.0382)	(-20, 20)

L'IC ne se situe pas dans l'intervalle d'équivalence. Déclaration d'équivalence impossible.

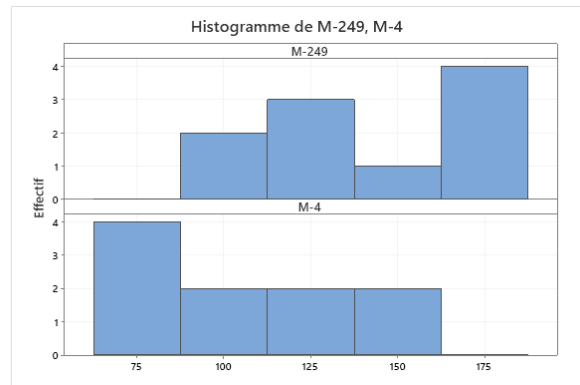
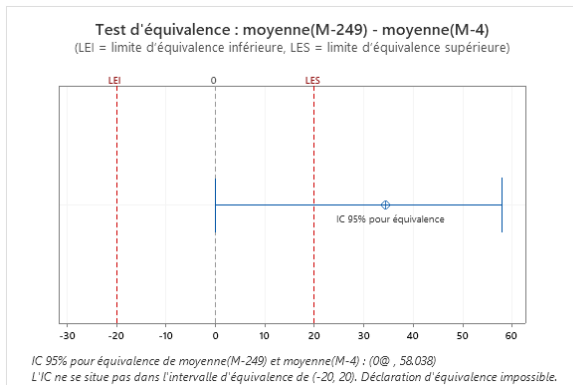
Test

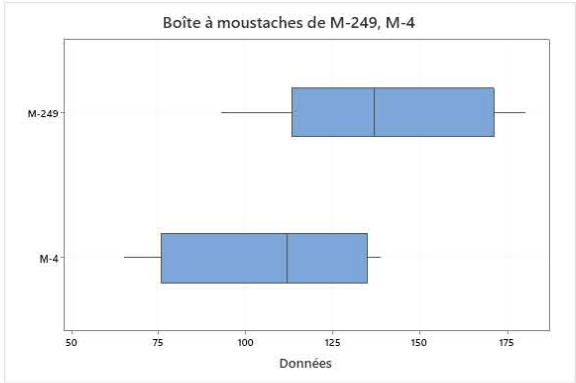
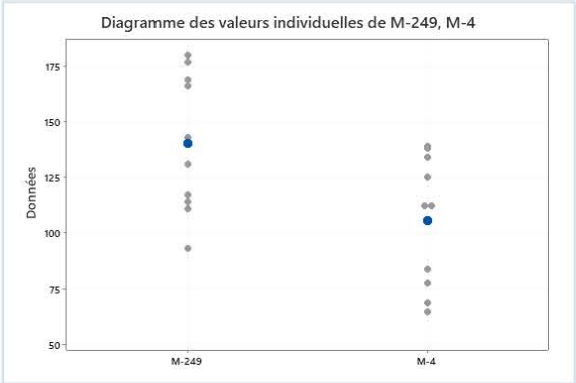
Hypothèse nulle : Différence  $\leq$  -20 ou différence  $\geq$  20  
 Hypothèse alternative : -20 < différence < 20  
 Niveau d' $\alpha$  : 0.05

Hypothèse nulle	DL	Valeur de T	Valeur de P
Différence $\leq$ -20	17	4.0279	0.000
Différence $\geq$ 20	17	1.0716	0.851

La plus grande des deux valeurs de p est 0.851. Déclaration d'équivalence impossible.

Avec les graphiques suivants:





ÉCHANTILLON

## 10.30. Exercice 61.: Taille d'échantillon (effectif) d'un test p (proportion) à 1 échantillon en bilatéral

Minitab® Statistical Software 16.2.1

Encore une fois nous souhaiterions déterminer la taille d'échantillon permettant de mettre en évidence une certaine différence de proportion dans le but de vérifier la relation démontrée dans le cours de statistique théorique et qui est:

$$n = \left( \frac{Z_{\beta} \sqrt{\hat{p}\hat{q}} - Z_{1-\alpha/2} \sqrt{pq}}{p - \hat{p}} \right)^2$$

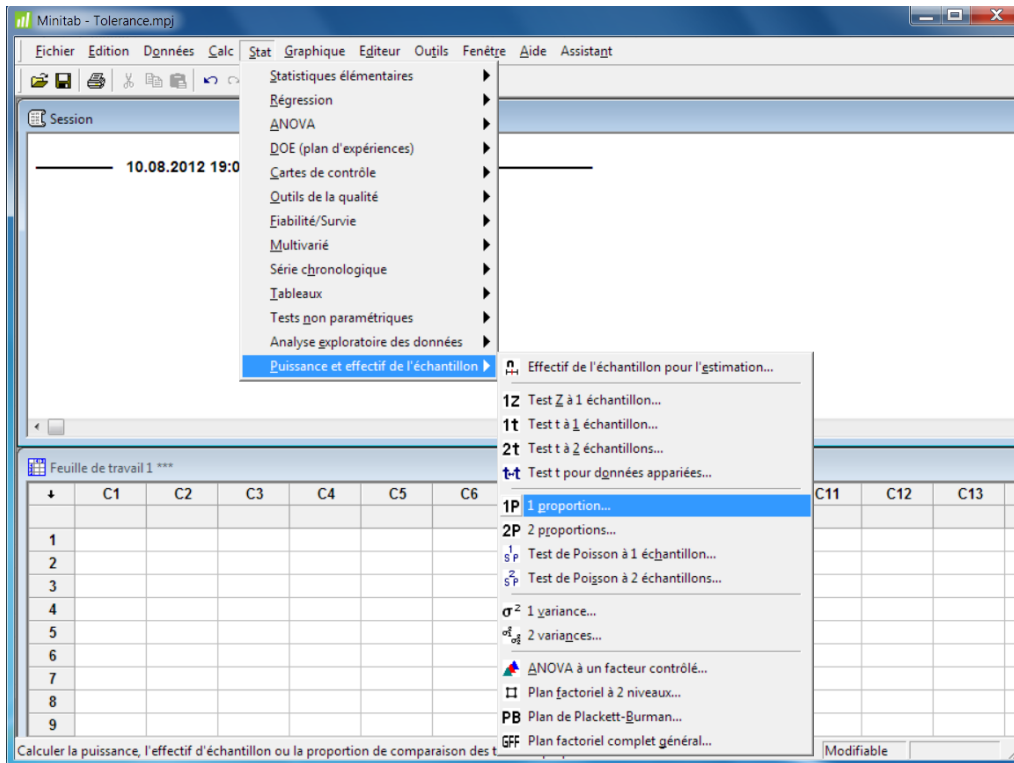
où dans le cas d'une puissance à 50%, nous retrouvons la relation connue:

$$n \cong \frac{Z_{\alpha/2}^2 p(1-p)}{\delta^2} = \frac{Z_{\alpha/2}^2 pq}{\delta^2}$$

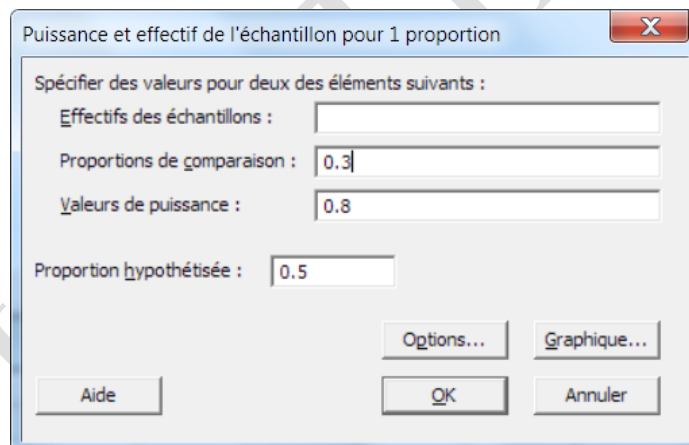
Imaginons que nous souhaiterions mettre en évidence une différence de 20% dans une étude (ou sondage) où nous nous attendons à avoir 50% de proportion expérimentale (donc une proportion de comparaison de 50-20=30%)

À une puissance de 80%, nous souhaiterions donc savoir qu'elle est la taille de l'échantillon qui nous permettra de ne pas maintenir l'hypothèse nulle à tort.

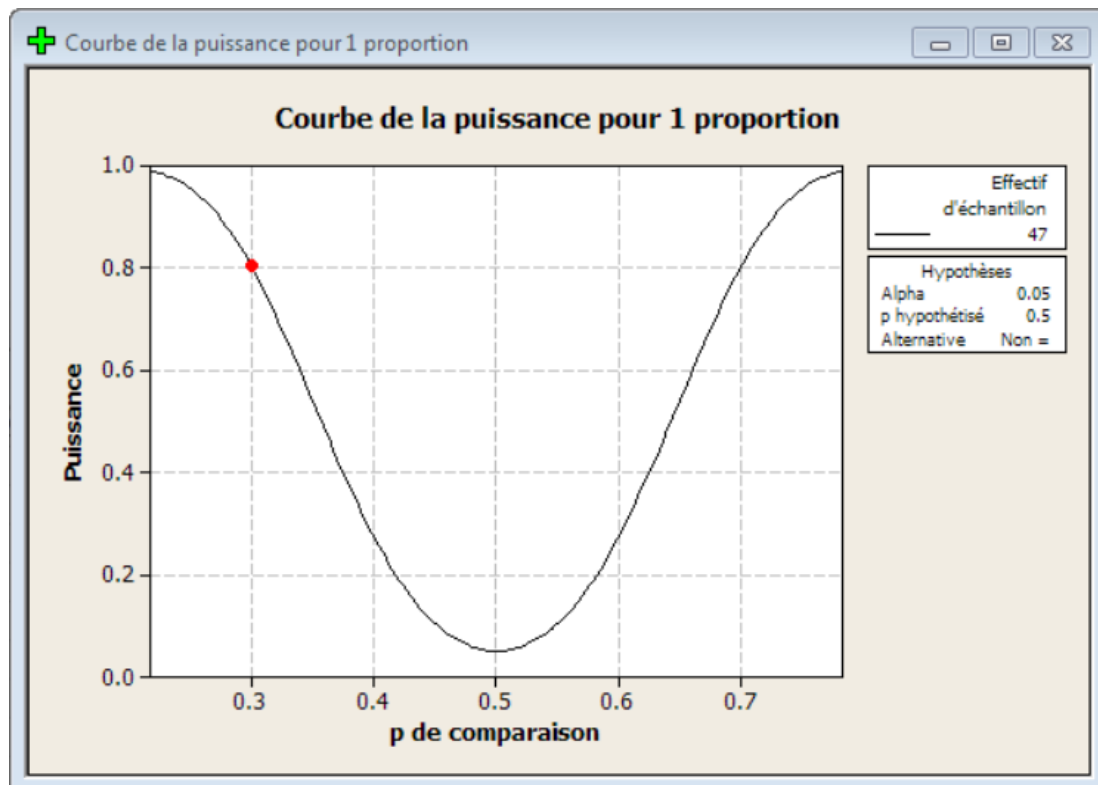
Pour cela, nous allons dans le menu **Stat/Puissance et effectif de l'échantillon/1 Proportion...**:



Nous mettons:



et nous validons par **OK** pour obtenir:



et dans la fenêtre de session:

Test pour 1 proportion

Proportion testée = 0.5 contre (proportion non = 0.5)  
Alpha = 0.05

p de comparaison	Effectif d'échantillon	Puissance cible	Puissance réelle
0.3	47	0.8	0.803325

**Courbe de la puissance pour 1 proportion**

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
1													
2													
3													
4													
5													
6													
7													
8													
9													

Comparons au résultat théorique avec Microsoft Excel:

$$n = \left( \frac{Z_{\beta} \sqrt{\hat{p}\hat{q}} - Z_{1-\alpha/2} \sqrt{pq}}{p - \hat{p}} \right)^2$$

$$= ((\text{NORMSINV}(20\%) * \text{SQRT}(0.5 * (1 - 0.5)) - \text{NORMSINV}(1 - 5\%/2) * \text{SQRT}(0.3 * (1 - 0.3))) / (0.3 - 0.5))^2 = 43.4926$$

La différence avec Minitab provenant d'un facteur de correction de continuité qu'utilise Minitab. Mais il vaut mieux avoir un peu trop d'échantillons que pas assez.

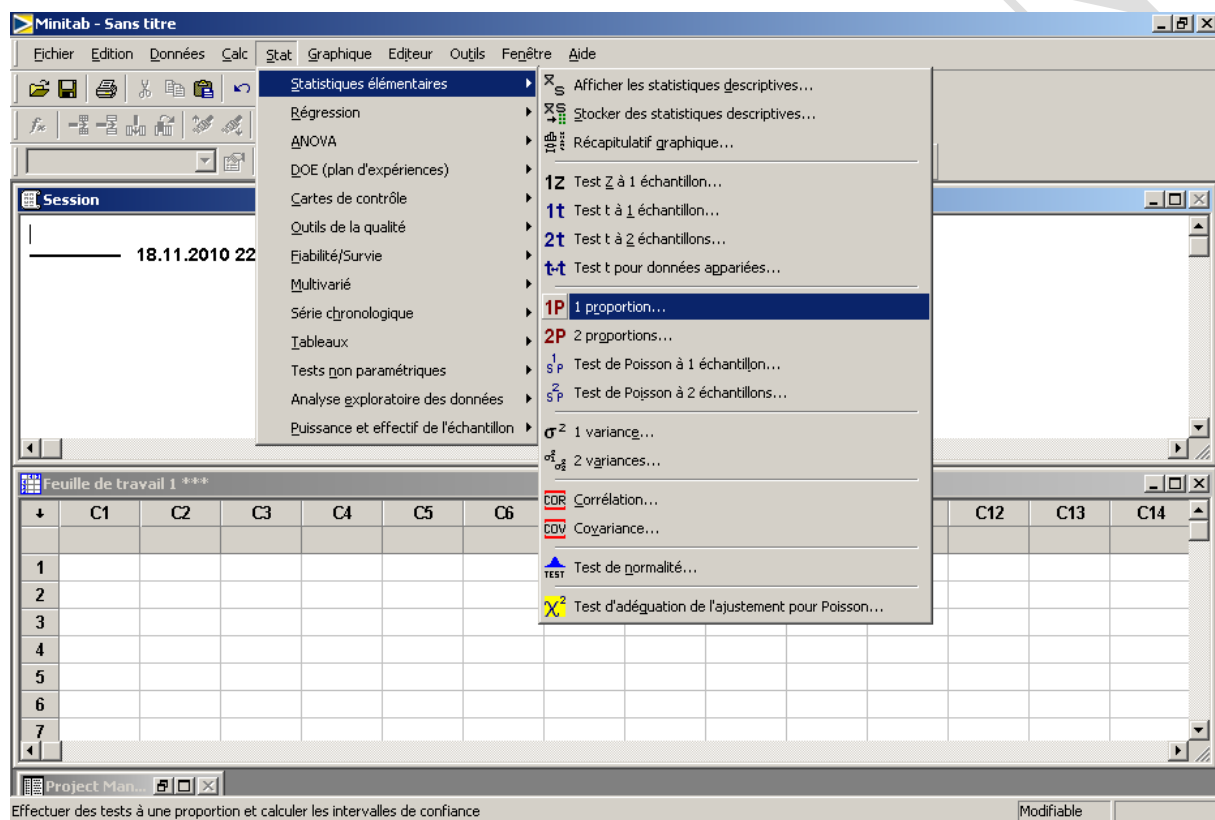
## 10.31. Exercice 62.: Intervalle de confiance de la proportion

Minitab® Statistical Software 15.1.1.0

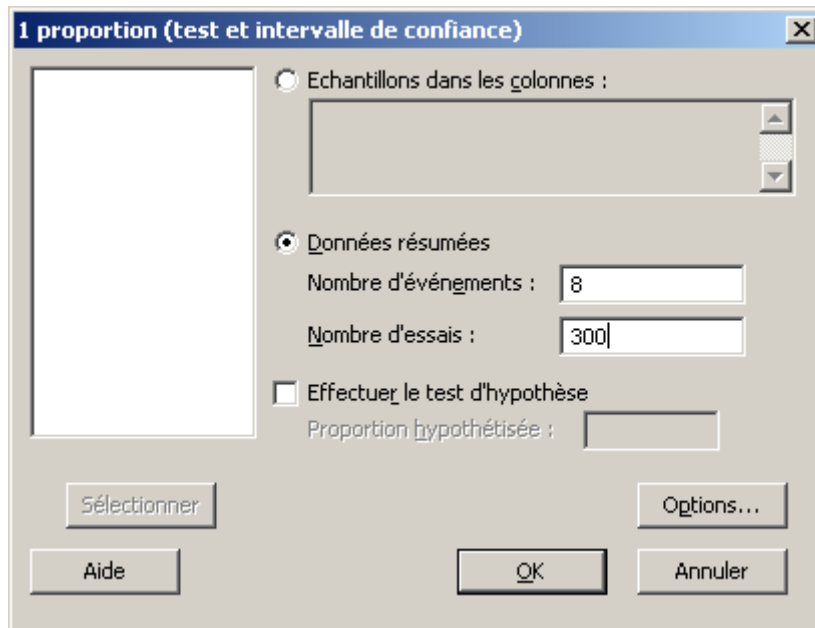
Dans une production de 300 pièces nous en avons trouvées 8 de défectueuses. Quelle est l'intervalle de confiance à 95% de la proportion de pièces défectueuses?

Nous avons fait ce calcul déjà avec Microsoft Excel et nous souhaitons le vérifier.

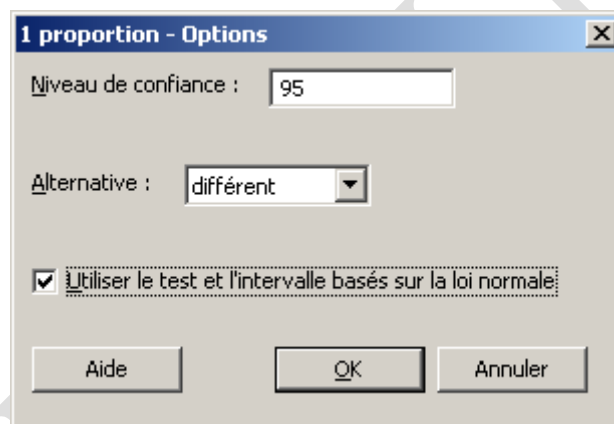
Dans Minitab® Statistical Software (à vide), nous allons dans le menu **Stat/Statistiques élémentaires/1 proportion...**:



Ce qui donne:

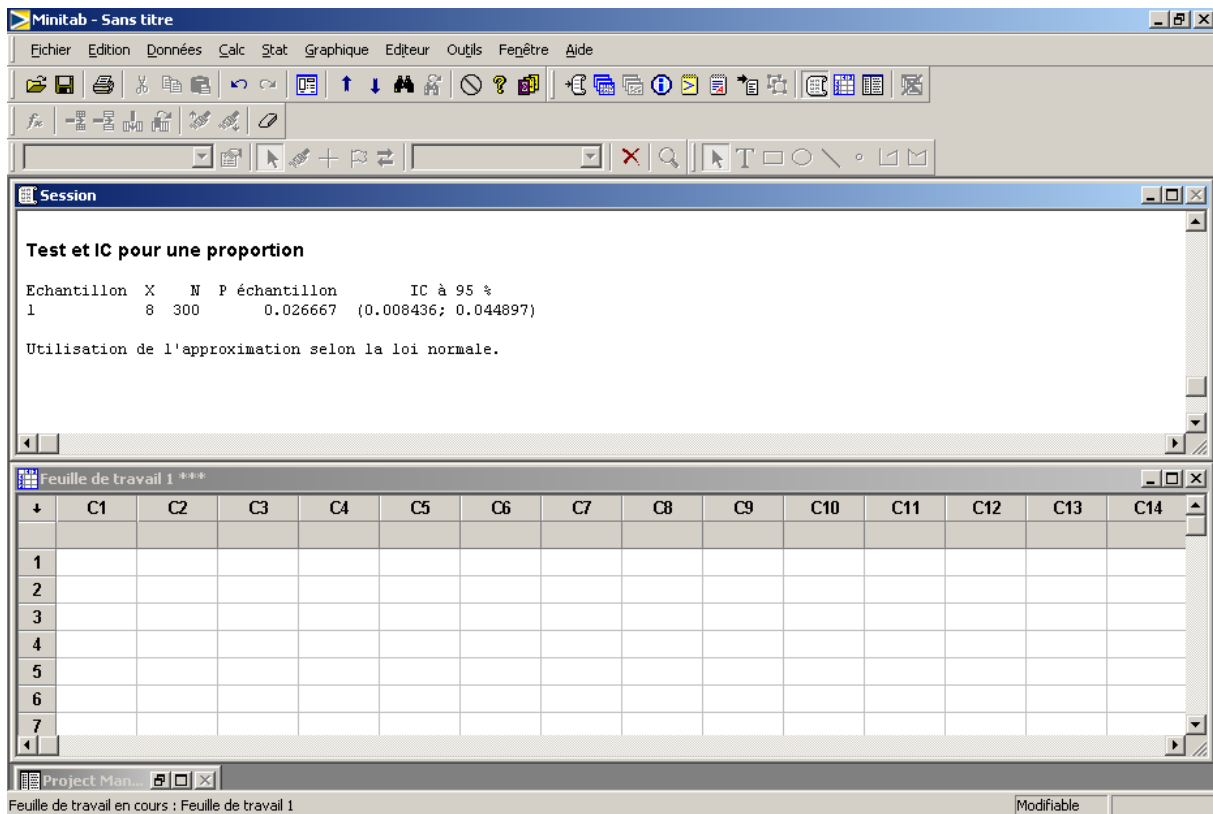


Et nous allons dans le bouton **Options**:



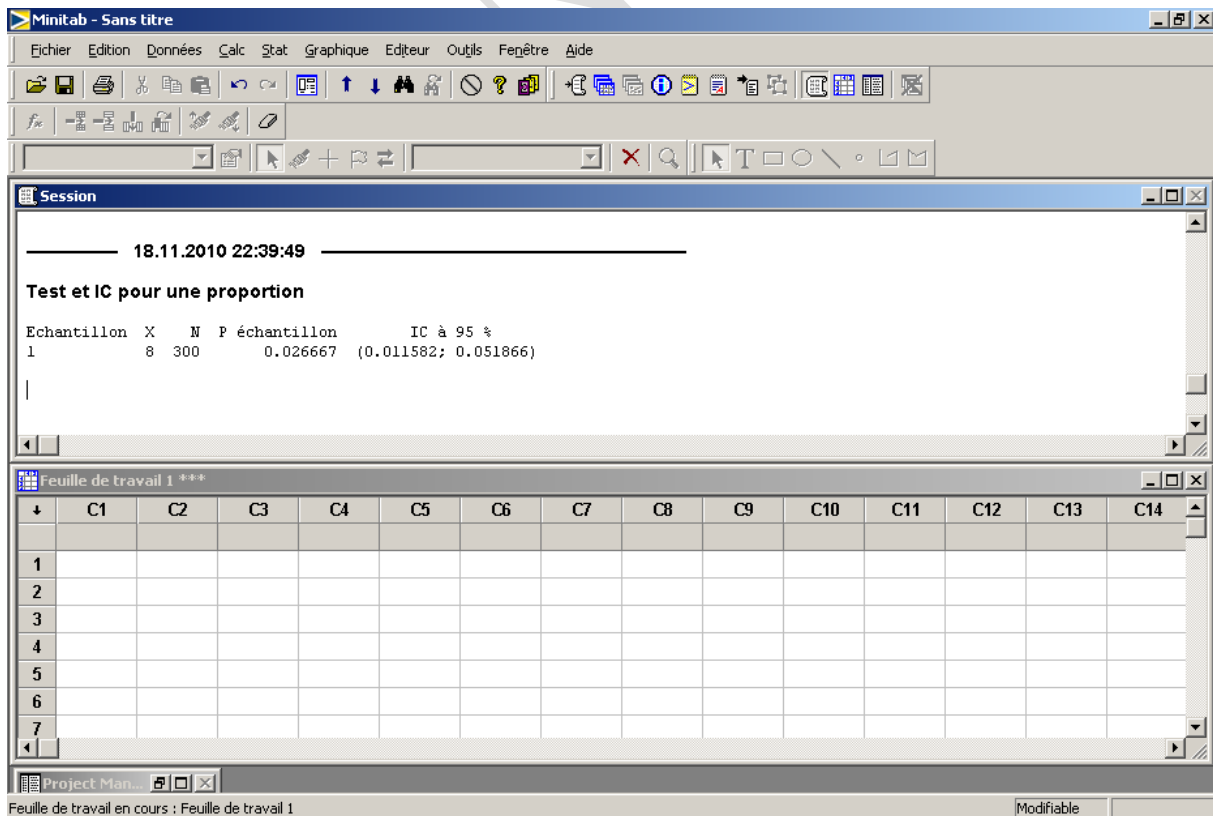
Nous validons alors deux fois par **OK** avons alors dans la fenêtre de session:





Nous avons donc bien le même résultat que dans Microsoft Excel (et donc que dans le cours de statistique théorique).

Si nous n'avions pas activé l'approximation par une loi normale, nous aurions:



## 10.32. Exercice 63.: Taille d'échantillon (effectif) d'un test p (proportion) à 2 échantillons en bilatéral

Minitab® Statistical Software 16.2.1

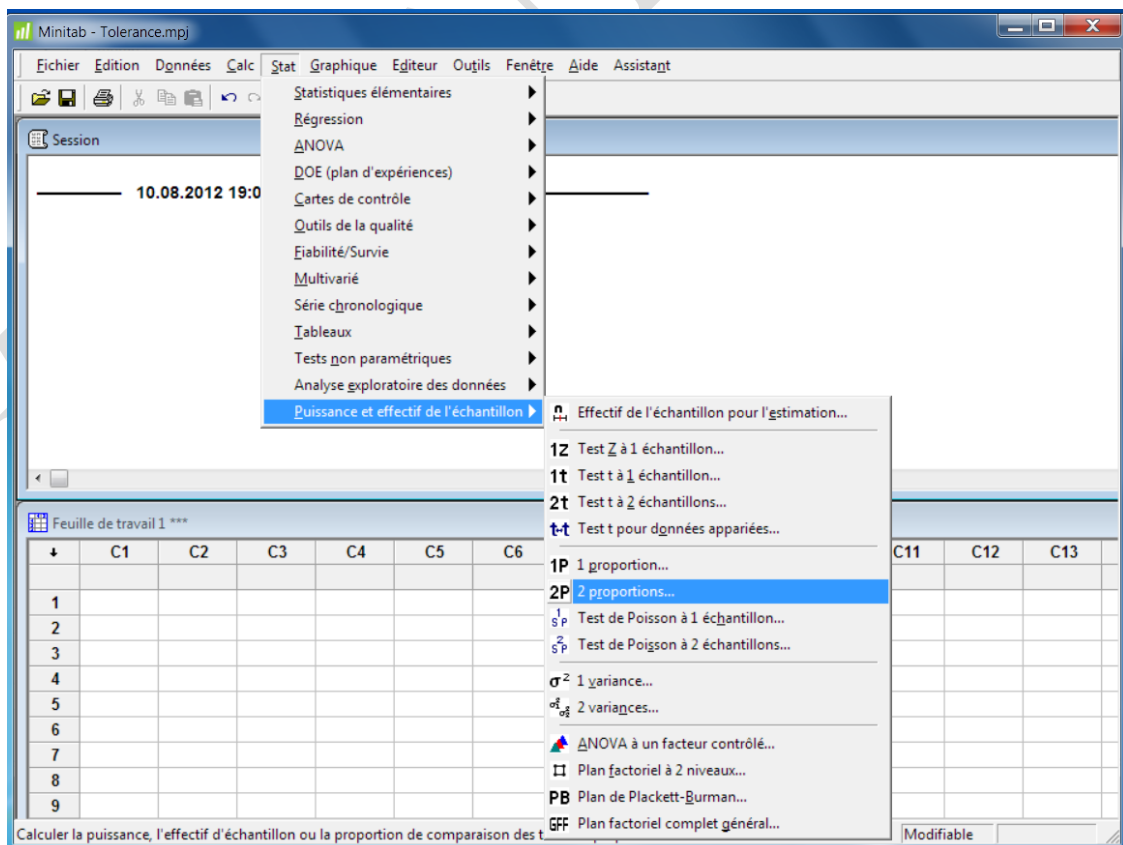
Encore une fois nous souhaiterions déterminer la taille de deux échantillons permettant de mettre en évidence une certaine différence de proportions dans le but de vérifier la relation démontrée dans le cours de statistique théorique et qui est:

$$n = \left( \frac{Z_{\beta} \sqrt{\hat{p}_1 \hat{q}_1 + \hat{p}_2 \hat{q}_2} - Z_{1-\alpha/2} \sqrt{2 \hat{p} \hat{q}}}{\hat{p} - (\hat{p}_2 - \hat{p}_1)} \right)^2$$

Imaginons pour cela que nous souhaiterions mettre en évidence une variation de 20% dans la différence de deux proportions qui sont estimées à priori comme valant 30% et respectivement 50%.

À une puissance de 80%, nous souhaiterions donc savoir qu'elle est la taille des deux échantillons (puisque leur taille est supposée égale dans les développements théoriques) qui nous permettra de ne pas maintenir l'hypothèse nulle à tort.

Pour cela, nous allons dans le menu **Stat/Puissance et effectif de l'échantillon/2 Proportions...**:



Nous mettons:

Puissance et effectif de l'échantillon pour 2 proportions

Spécifier des valeurs pour deux des éléments suivants :

Effectifs des échantillons :

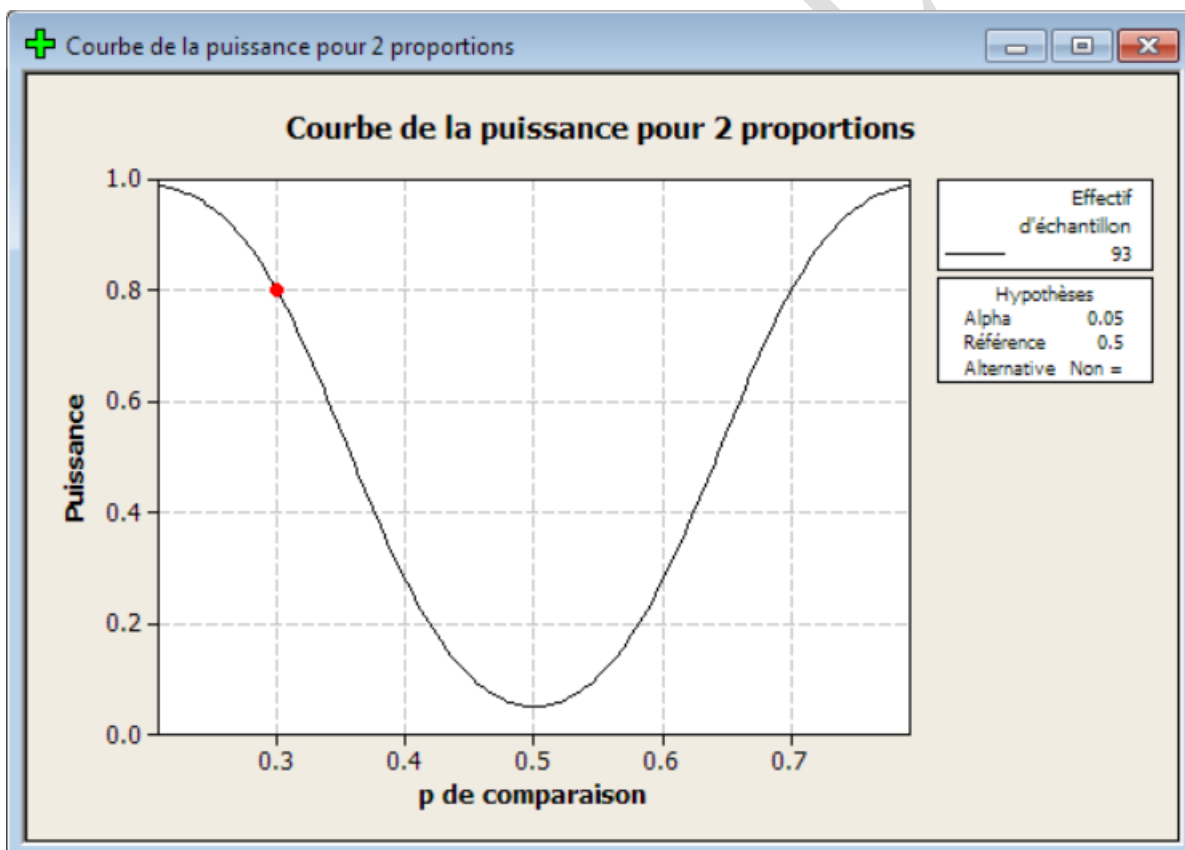
Proportions de comparaison ( $p_1$ ) :

Valeurs de puissance :

Proportion de référence ( $p_2$ ) :

Aide Options... Graphique... OK Annuler

et nous validons par **OK** pour obtenir:



et dans la fenêtre de session:

The screenshot shows the Minitab interface. The 'Session' window displays the following text:

```

Test de comparaison de p = référence (contre p non =)
Calcul de la puissance pour référence = 0.5
Alpha = 0.05

p de      Effectif  Puissance  Puissance
comparaison d'échantillon  cible  réelle
0.3                93      0.8    0.800006

L'effectif de l'échantillon est pour chaque groupe.

Courbe de la puissance pour 2 proportions

```

The 'Feuille de travail 1 \*\*\*' window shows an empty spreadsheet grid with columns C1 to C13 and rows 1 to 9.

Comparons au résultat théorique avec Microsoft Excel:

$$n = \left( \frac{Z_{\beta} \sqrt{\hat{p}\hat{q}} - Z_{1-\alpha/2} \sqrt{pq}}{p - \hat{p}} \right)^2$$

$$= ((\text{NORM.S.INV}(20\%) * \text{SQRT}(0.5 * 0.5 + 0.3 * 0.7) - \text{NORM.S.INV}(1 - 5\% / 2) * \text{SQRT}(2 * 0.4 * 0.6)) / (0.4 - (0.5 - 0.3)))^2 = 92.998$$

La différence avec Minitab est donc nulle puisque celui-ci arrondit à l'entier le plus proche.

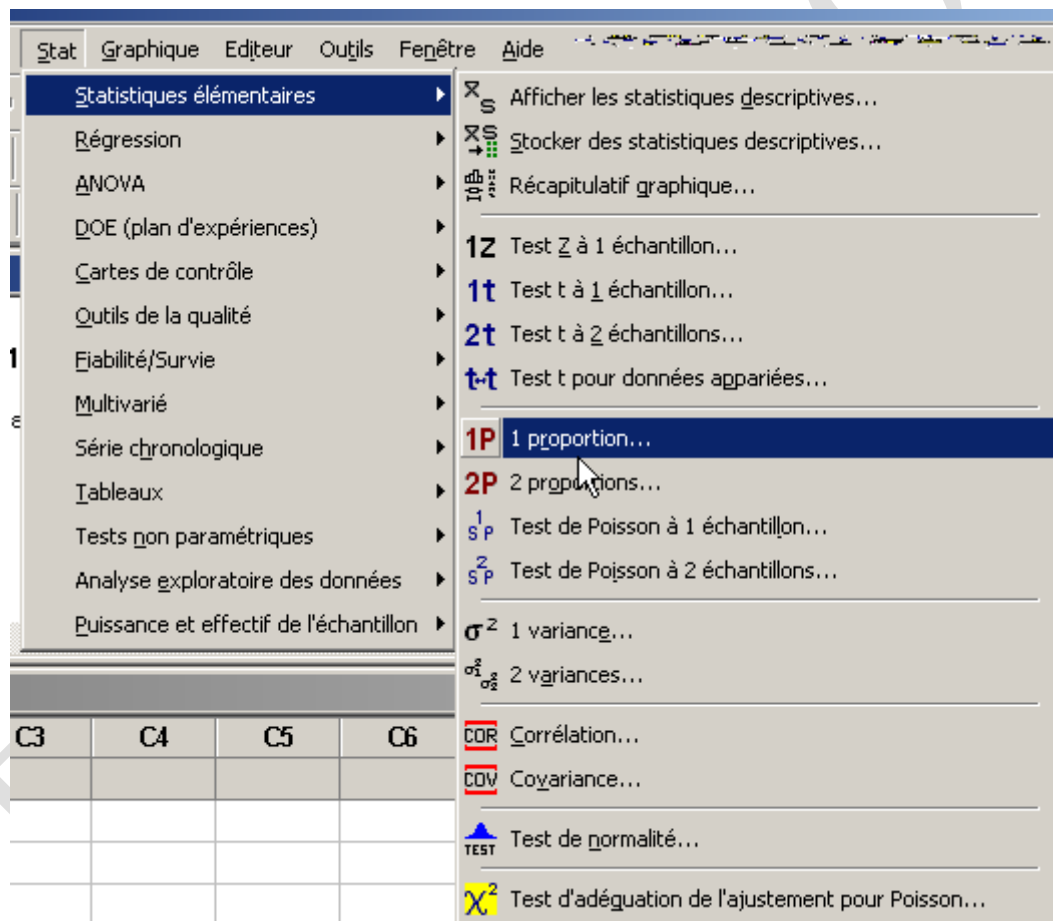
## 10.33. Exercice 64.: Comparaison de proportions sur une même population (test binomial exact)

Minitab® Statistical Software 15.1.1.0

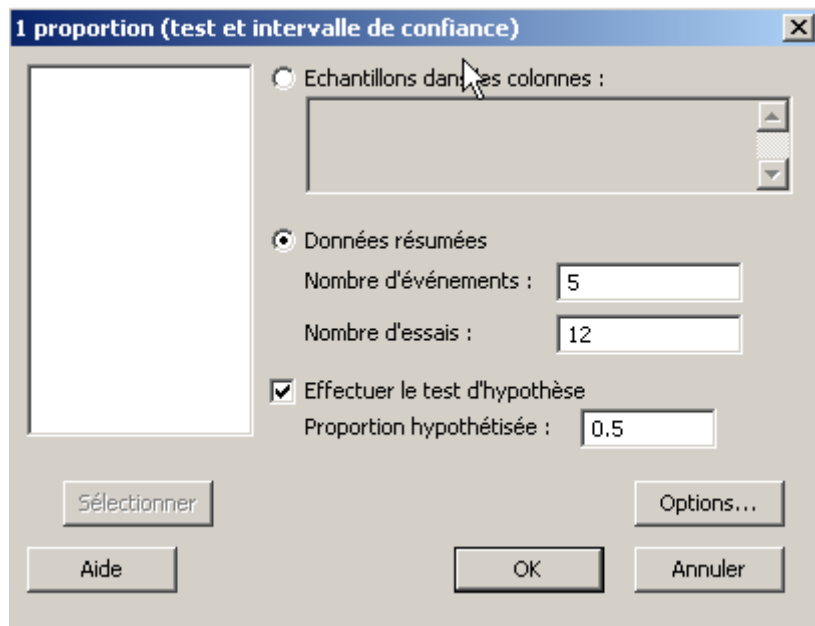
D'une petite population ayant deux caractéristiques  $x$  et  $y$  particulières qui nous intéressaient et pour laquelle nous nous attendions à avoir un parfait équilibre tel que  $x = y$  (donc 50%/50%) nous avons en réalité obtenu  $x = 5$  et  $y = 7$  (donc la population est de taille 12).

Nous souhaiterions donc savoir si la différence est significative avec une certitude de 95% ou simplement due au hasard?

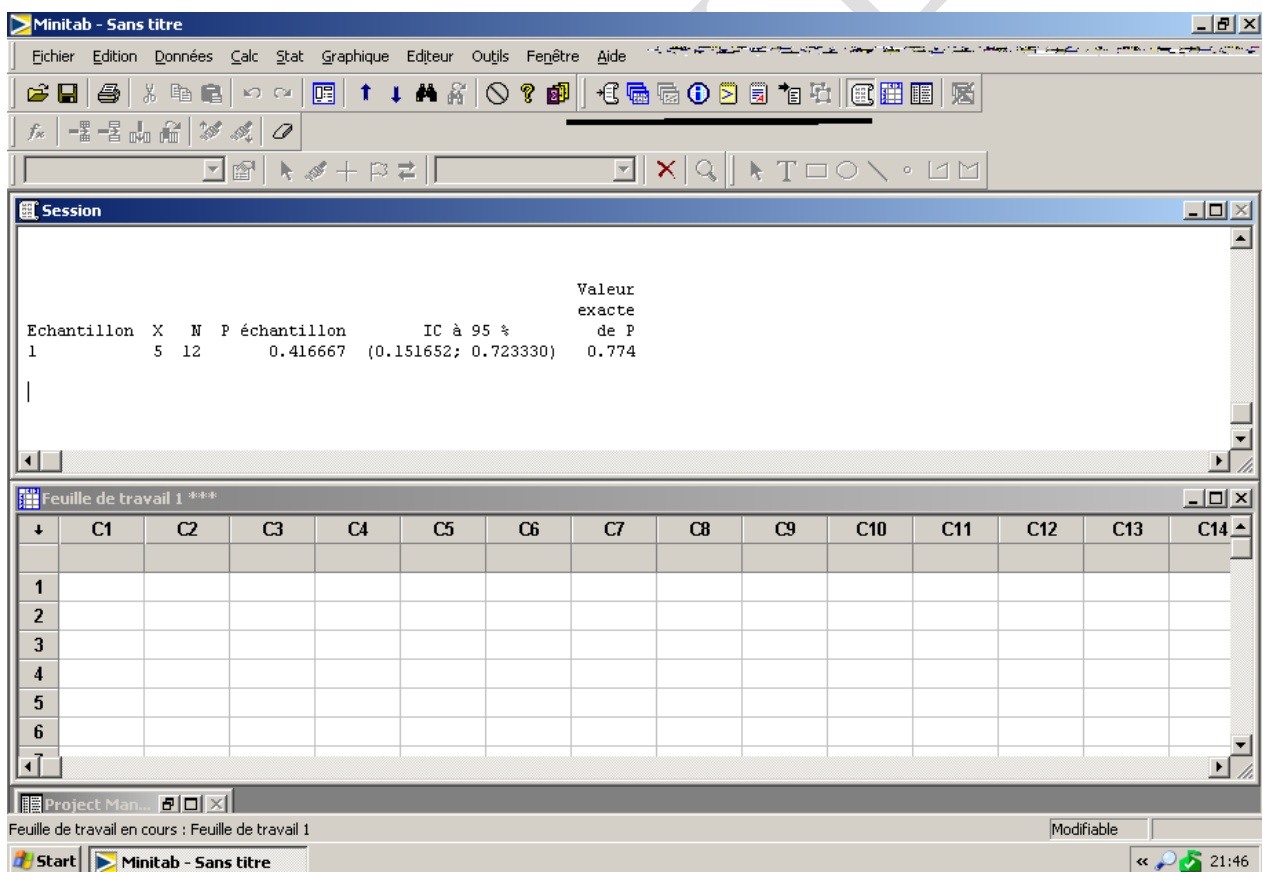
Nous allons dans le menu **Stat/Statistiques élémentaires/1 proportion...**:



Nous avons alors:



ce qui donne après validation:



Nous nous retrouvons alors avec les mêmes résultats que Microsoft Excel.

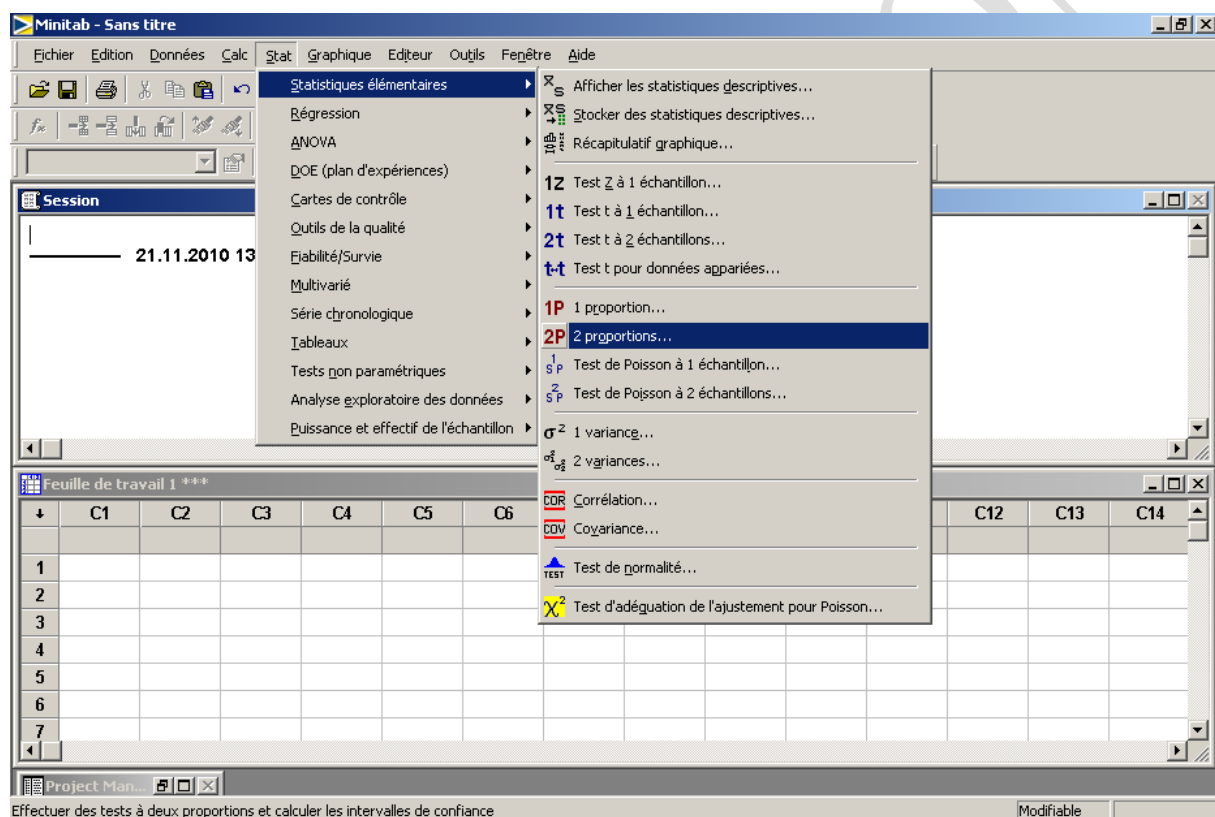
## 10.34. Exercice 65.: Comparaison de proportions sur 2 échantillons indépendants

Minitab® Statistical Software 15.1.1.0

Dans le cadre d'un plan d'échantillonnage nous avons prélevé sur un premier lot de 50 échantillons, 48 en parfait états. Dans un second lot de 30 échantillons, 26 étaient en bon état.

Nous souhaiterions donc savoir si la différence est significative avec une certitude de 95% ou simplement due au hasard?

Nous allons dans le menu **Stat/Statistiques élémentaires/2 proportions...**:



Nous avons alors:

2 proportions (test et intervalle de confiance)

Echantillons dans une colonne  
Echantillons :   
Indices :

Echantillons dans plusieurs colonnes  
Premier :   
Deuxième :

Données résumées  
Evénements : Essais :  
Premier : 48 50  
Deuxième : 26 30

Sélectionner

Options...

Aide OK Annuler

Nous cliquons sur le bouton **Options...** pour configurer relativement à la démonstration vue dans le cours de statistique théorique:

2 proportions - Options

Niveau de confiance : 95.0

Tester la différence : 0.0

Alternative : différent

Utiliser l'estimation regroupée de p pour le test

Aide OK Annuler

Ce qui donne après validation par deux **OK**:



**Test et IC pour deux proportions**

Echantillon	X	N	P échantillon
1	48	50	0.960000
2	26	30	0.866667

Différence =  $p(1) - p(2)$   
Estimation de la différence : 0.0933333  
Limites de confiance (à 95 %) pour la différence : (-0.0398844 ; 0.226551)  
Test de la différence = 0 (en fonction de la différence 0) :  $Z = 1.53$   $P = 0.125$

Test exact de Fisher : valeur p = 0.190

\* REMARQUE \* L'approximation selon la loi normale risque de ne pas être exacte pour les petits échantillons.

Nous nous retrouvons alors avec les mêmes résultats que Microsoft Excel (excepté le test de Fisher exact que nous n'avions pas calculé pendant le cours théorique pour cet exemple particulier).

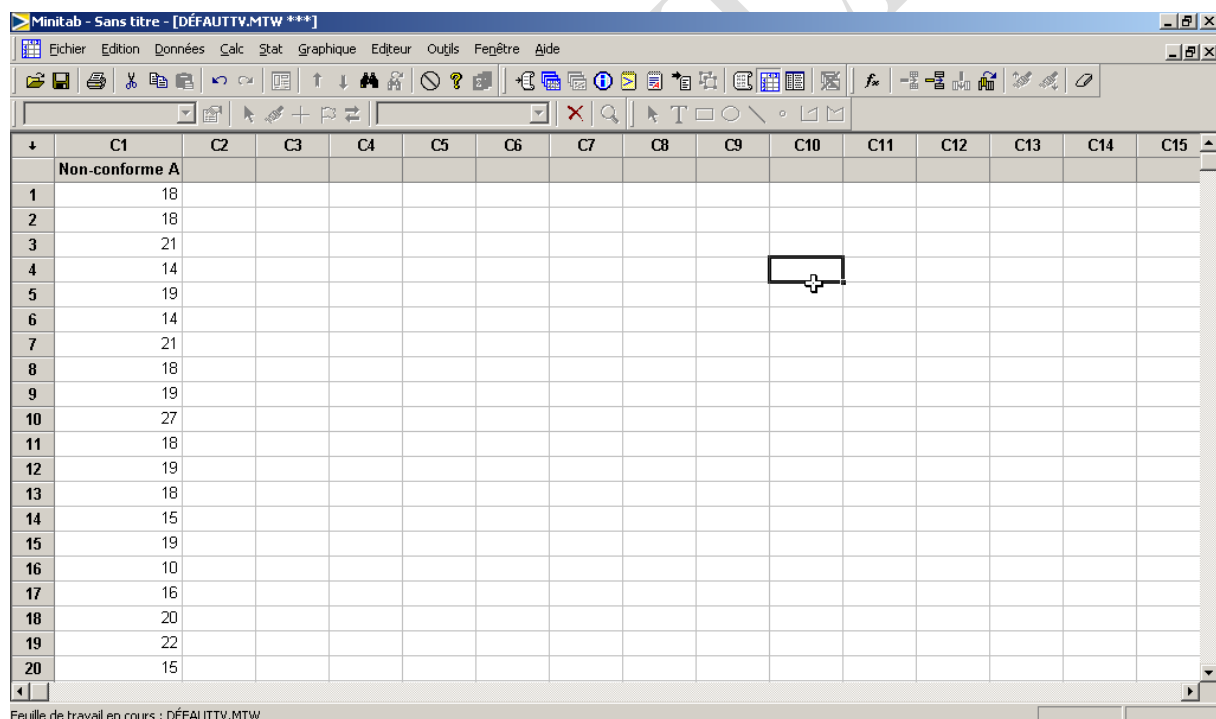
## 10.35. Exercice 66.: Test de Poisson à un échantillon unilatéral/bilatéral

Minitab® Statistical Software 15.1.1.0

Le but est encore une fois de vérifier si nous trouvons les mêmes résultats que dans le cours théorique et que les calculs appliqués avec Microsoft Excel. Considérons le premier exemple vu dans le cours théorique et inspiré de l'aide de Minitab:

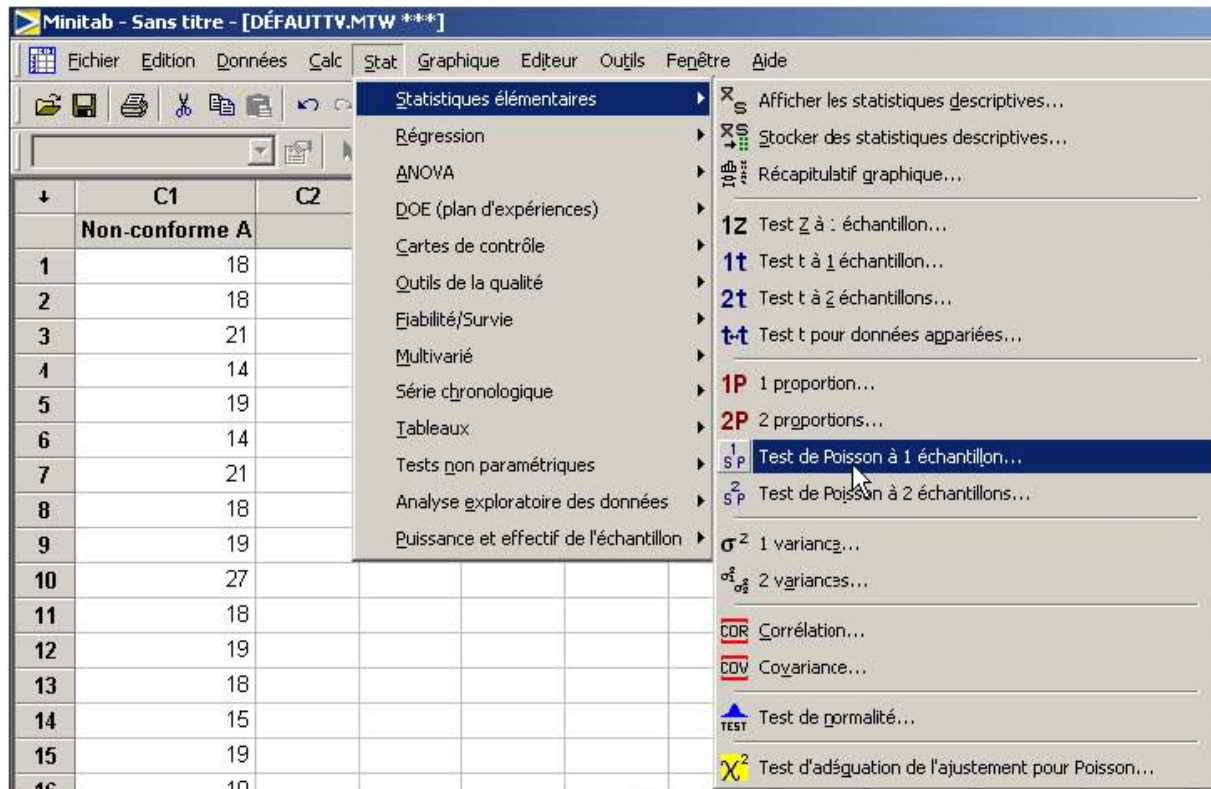
**Unilatéral:** Une société fabrique des télévisions en quantité constante et a mesuré le nombre d'appareils défectueux produits chaque trimestre pendant les dix dernières années (donc 4 fois 10 mesures = 40 trimestres). La direction décide que le nombre maximum acceptable d'unités défectueuses est de 20 par trimestre et souhaite déterminer si l'usine satisfait à ces exigences (sous l'hypothèse que la distribution des défectueux suive une loi de Poisson) à un niveau de confiance de 5%.

Nous ouvrons donc le fichier *Poisson1Echantillon.mpj* comportant 40 mesures (donc une pour chaque trimestre):

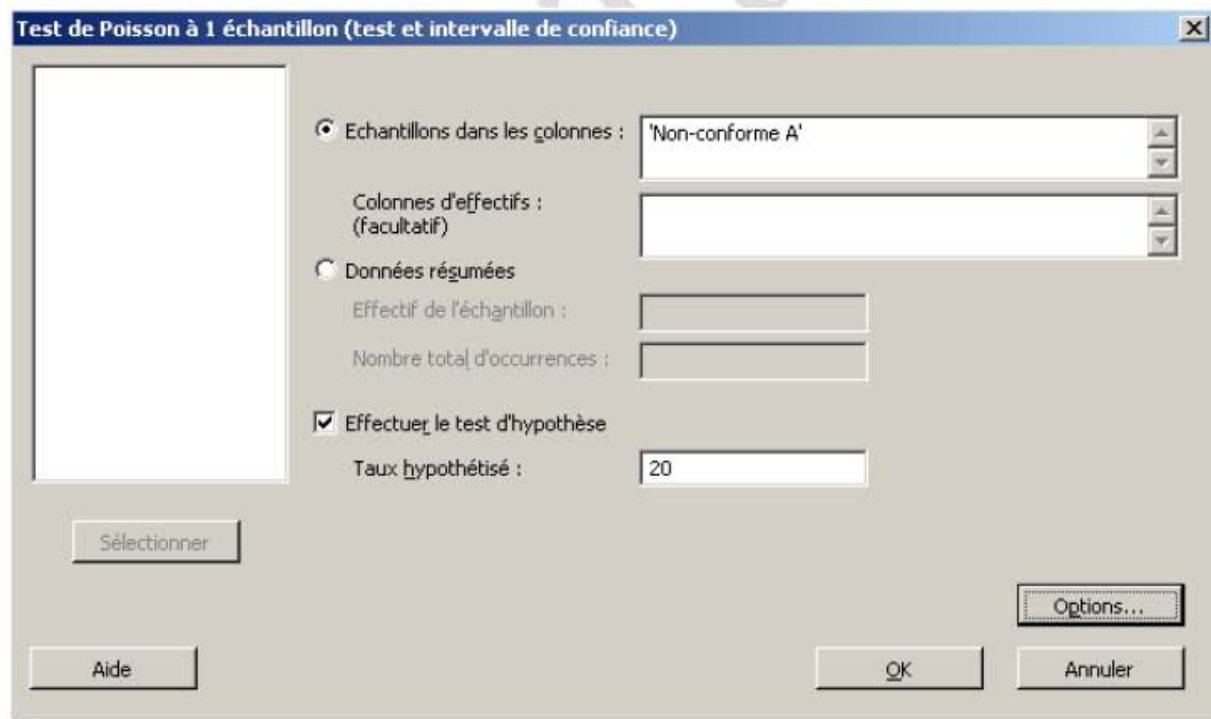


	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
	Non-conforme A														
1	18														
2	18														
3	21														
4	14														
5	19														
6	14														
7	21														
8	18														
9	19														
10	27														
11	18														
12	19														
13	18														
14	15														
15	19														
16	10														
17	16														
18	20														
19	22														
20	15														

Nous allons dans le menu **Stat/Statistiques élémentaires/Test de Poisson à 1 échantillon...**:



et nous mettons:



et dans le bouton **Options...**:

Test de Poisson à 1 échantillon - Options

"Longueur" de l'observation (temps, éléments, aire, volume, etc.) : 1.0

Niveau de confiance : 95.0

Alternative : inférieur

Utiliser le test et l'intervalle de confiance en fonction de la loi de distribution normale

Aide OK Annuler

et nous validons le tout par **OK**:

#### Test et limites de confiance pour un test de Poisson à un échantillon: Non-conforme A

Test du taux = 20 contre taux inférieur à 20

Variable	Nombre total d'occurrences	N	Taux d'occurrence	Borne supérieure	Valeur exacte de P
Non-conforme A	713	40	17.8250	18.9628	0.001

"Longueur" de l'observation = 1.

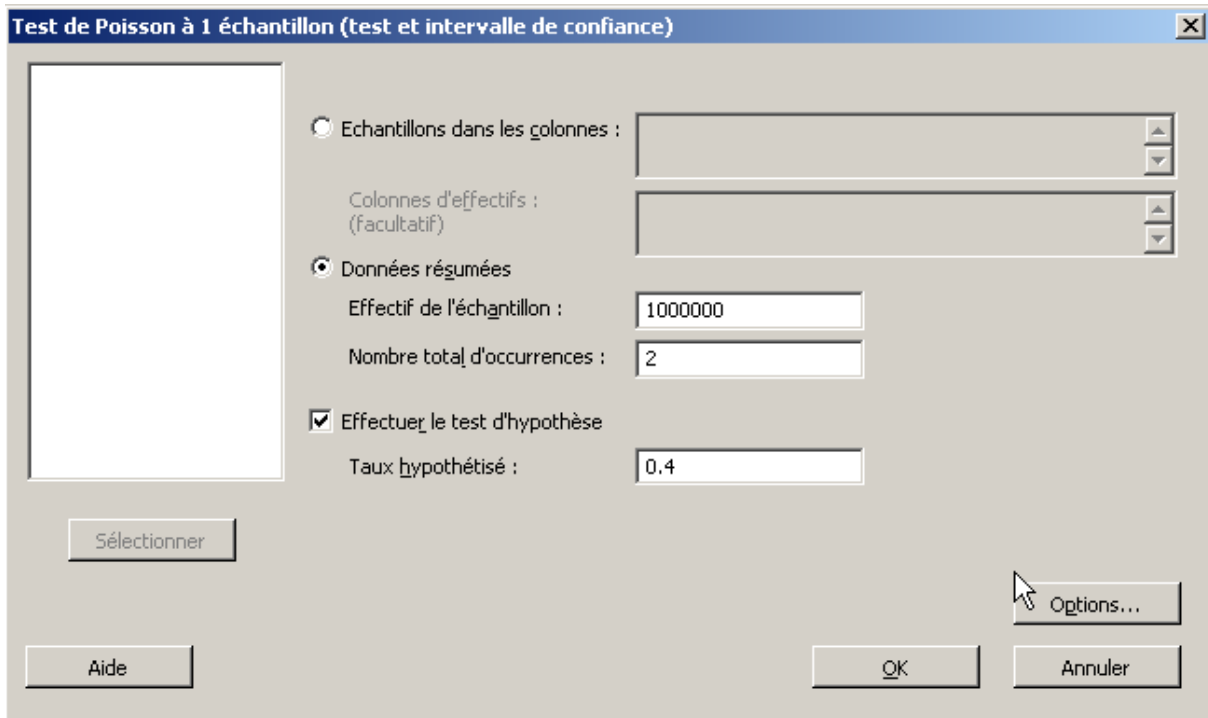
Nous retrouvons presque les mêmes résultats que les calculs effectués avec le tableur (uniquement la borne inférieure donnée par Minitab diffère au deuxième chiffre après la virgule de celle obtenue avec Microsoft Excel).

Nous pouvons remarquer une chose un peu curieuse: Minitab semble utiliser l'approximation Normale (abusive) de la loi de Poisson... affaire à suivre...

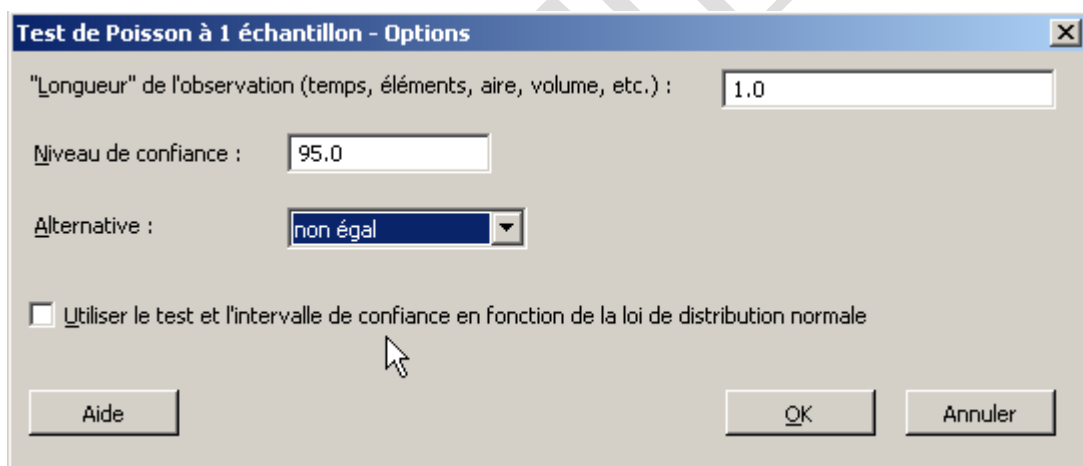
Maintenant, refaisons l'exemple calculé et démontré dans le cours théorique avec l'intervalle en bilatéral.

**Bilatéral:** Une compagnie d'aviation a eu 2 deux crashes en 1'000'000 de vols (événement très rare). Quelle est l'intervalle de confiance en bilatéral à 95% sachant qu'au niveau mondial le nombre d'accidents par millions est de 0.4.

Il n'y a ici aucun fichier à ouvrir! Nous allons simplement dans le menu **Stat/Statistiques élémentaires/Test de Poisson à 1 échantillon...** et nous y mettons:



Nous allons dans le bouton **Options...**:



Il n'y a rien à changer! Nous validons le tout par **OK** deux fois pour obtenir:

Test du taux = 0.4 contre taux différent de 0.4

Echantillon	Nombre total d'occurrences	N	Taux d'occurrence	Limites de confiance = 95 %	Valeur exacte de P
1	2	1000000	0.000002	(0.000000, 0.000007)	0.000

"Longueur" de l'observation = 1.

et ici nous voyons une bêtise monumentale de Minitab qui met tous sous forme de taux! Ce qui fait que l'intervalle devient quasiment inexploitable! Avec Microsoft Excel, nous avons obtenus à partir de la relation démontrée dans le cours théorique:

$$\frac{\chi^2_{5\%/2} (2 \cdot (2+1))}{2} \leq \lambda \leq \frac{\chi^2_{1-5\%/2} (2 \cdot (2+1))}{2}$$

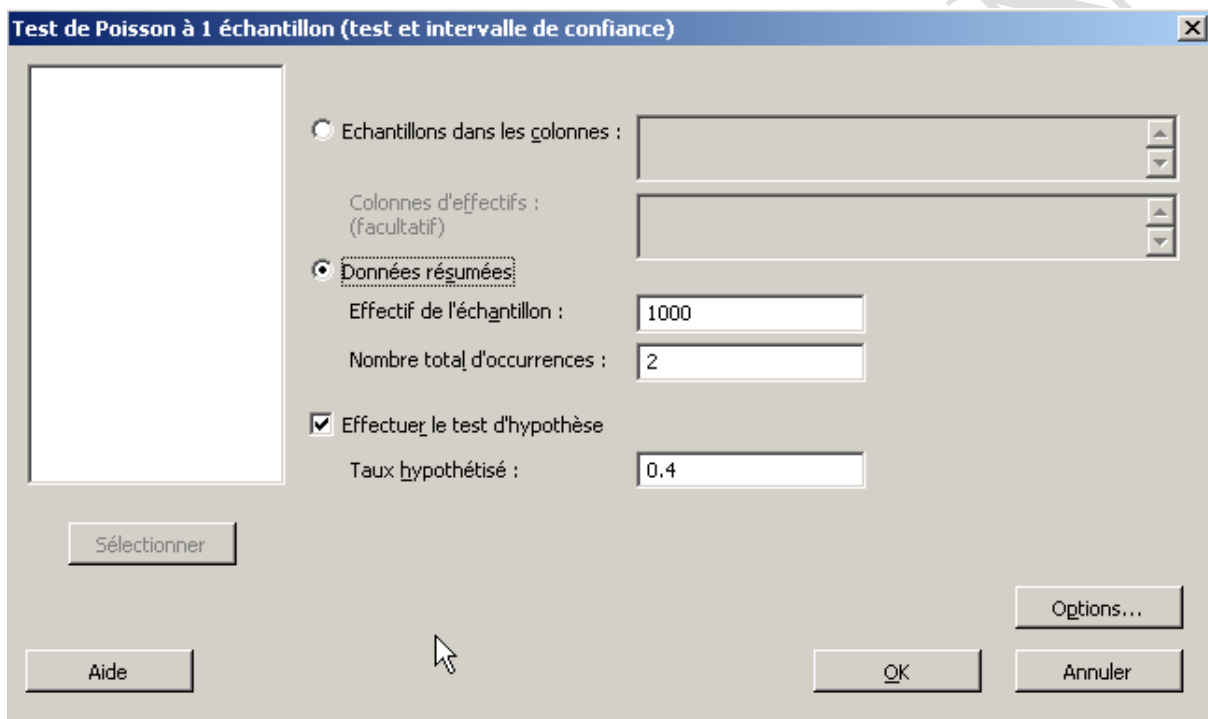
Le résultant suivant:

$$0.618 \leq \lambda \leq 7.224$$

ce qui divisé par 1'000'000 fait:

$$0.000000618 \leq \lambda \leq 0.000007224$$

Bon heureusement nous pouvons tricher en mettant par exemple:



ce qui donne alors:

Test du taux = 0.4 contre taux différent de 0.4

Echantillon	Nombre total d'occurrences	N	Taux d'occurrence	Limites de confiance = 95 %	Valeur exacte de P
1	2	1000	0.002000	(0.000242, 0.007225)	0.000

"Longueur" de l'observation = 1.

et là nous voyons une curiosité: Minitab, si nous comparons avec Microsoft Excel, calcule les bornes de la manière suivante:

$$=LOI.KHIDEUX.INVERSE(1-5\%/2;2*(2+1))/2=7.224$$

Ce qui est conforme pour la borne supérieur, mais pour la borne inférieure il calcule ainsi:

$$=LOI.KHIDEUX.INVERSE(1-5\%/2;2*2))/2=0.242$$

Ce qui est donc bizarre, car normalement les degrés de liberté de la loi du Khi-deux devraient a priori être identiques. En l'occurrence, l'acceptation ou le rejet de l'hypothèse nulle diffère ici totalement en fonction de la méthode de calcul. Affaire à suivre donc!

ÉCHANTILLON

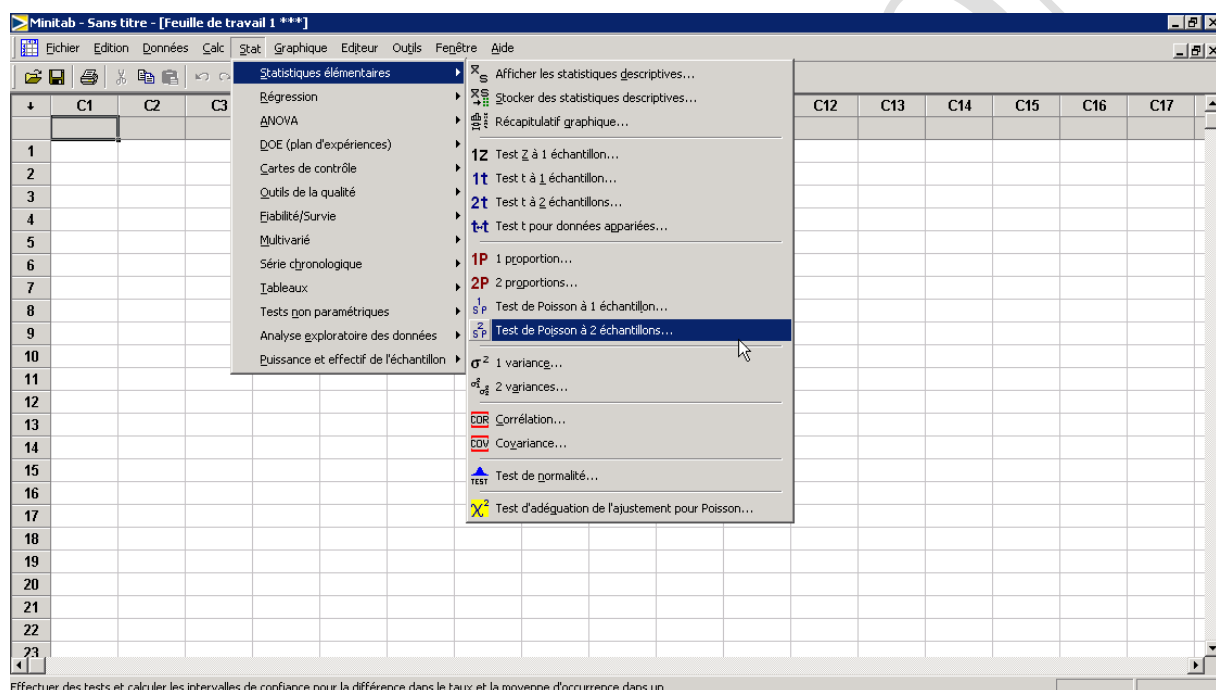
## 10.36. Exercice 67.: Test de Poisson à deux échantillons bilatéral

Minitab® Statistical Software 15.1.1.0

Nous allons ici considérer une variante de l'exemple précédent.

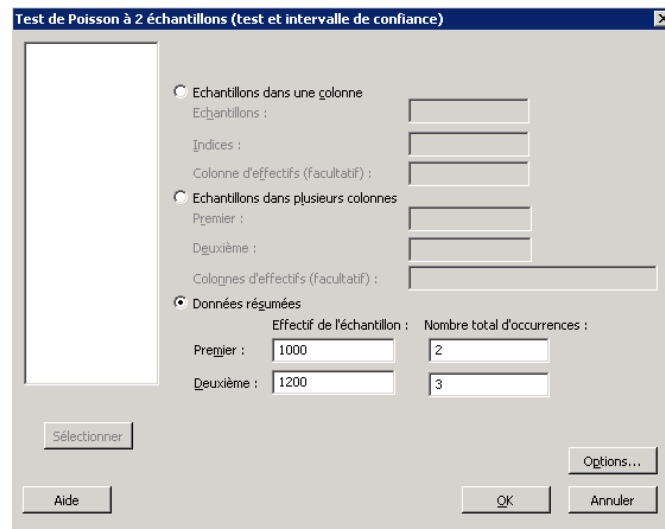
Une compagnie d'aviation a eu 2 deux crashes en 1'000'000 de vols (événement très rare). Une autre compagnie a eu 3 crashes en 1'200'000 vols. Quel est l'intervalle de confiance en bilatéral à 95% en supposant que la différence devrait être nulle.

Nous allons dans le menu **Stat/Statistiques élémentaires/Test de Poisson à 2 échantillons...**:

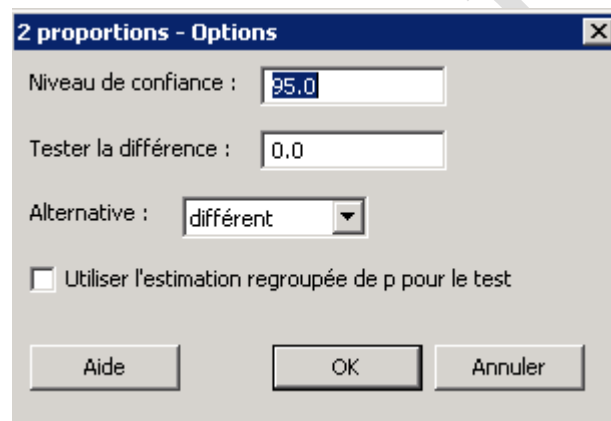


nous y mettons (pour la même raison que dans l'exemple précédent, nous avons réduit la voilure de l'effectif de l'échantillon):





Nous cliquons sur **Options...** pour vérifier que nous sommes bien en bilatéral avec le niveau de confiance voulu:



Nous obtenons alors:

Minitab - Sans titre - [Session]

Fichier Edition Données Calc Stat Graphique Éditeur Outils Fenêtre Aide

Test et limites de confiance pour les tests de Poisson à deux échantillons

Echantillon	Nombre total d'occurrences	N	Taux d'occurrence
1	2	1000	0.0020
2	3	1200	0.0025

Différence = taux (1) - taux (2)  
 Estimation de la différence : -0.0005  
 Limites de confiance (à 95 %) pour la différence : (-0.00446055, 0.00346055)  
 Test de la différence = 0 (en fonction de la différence 0) : Z = -0.25 Valeur p = 0.805

Valeur p du test exact = 1.000

\* REMARQUE \* L'approximation selon la loi normale risque de ne pas être exacte pour les nombres totaux d'occurrences trop faibles.

Donc la différence de proportion de -0.0005 étant dans l'intervalle de confiance, nous ne rejetons pas l'hypothèse nulle quoi la différence entre les deux compagnies est non significative au seuil de 5%.

Par curiosité, si nous faisons le même test mais avec le test de 2 proportions, nous avons alors:

```
Minitab - Sans titre - [Session]
Fichier  Edition  Données  Calc  Stat  Graphique  Editeur  Outils  Fenêtre  Aide

Test et IC pour deux proportions

Echantillon  X      N  P échantillon
1             2    1000  0.002000
2             3    1200  0.002500

Différence = p (1) - p (2)
Estimation de la différence : -0.0005
Limites de confiance (à 95 %) pour la différence : (-0.00445608 ; 0.00345608)
Test de la différence = 0 (en fonction de la différence 0) : Z = -0.25 P =
0.804

Test exact de Fisher : valeur p = 1.000

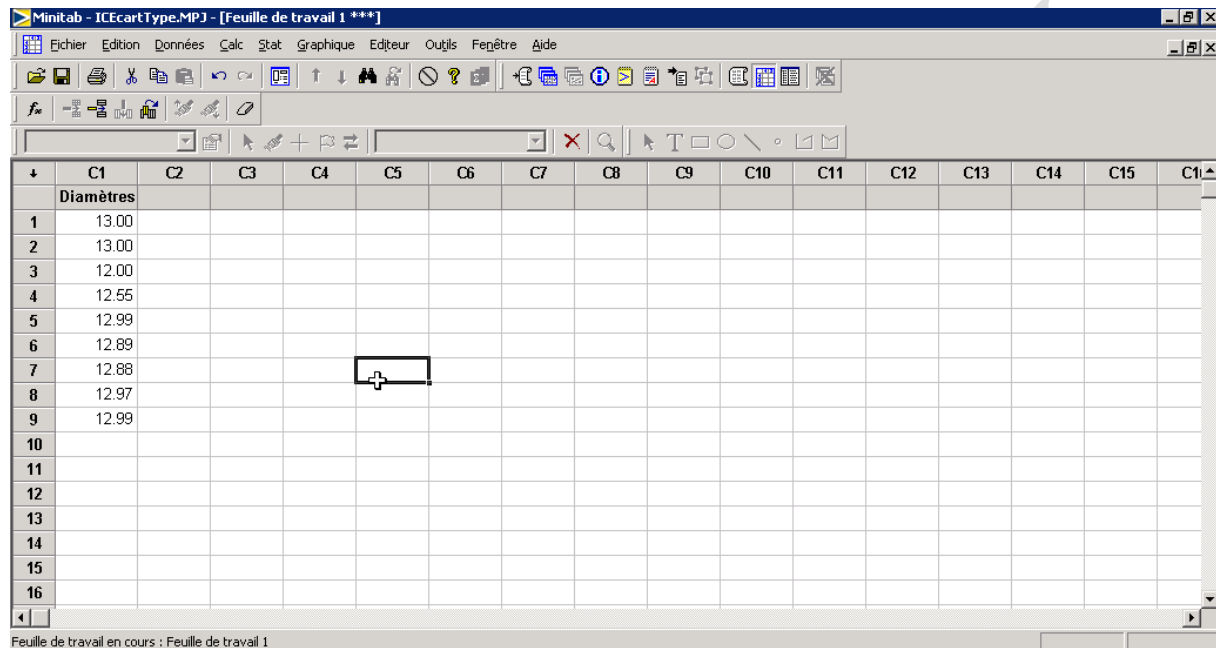
* REMARQUE * L'approximation selon la loi normale risque de ne pas être exacte
pour les petits échantillons.
```

Nous retrouvons donc presque les mêmes valeurs pour l'intervalle de confiance avec le test de 2 proportions. Mais la convergence entre le test de Poisson pour la différence 2 échantillons et celui du test de la différence de deux proportions n'est valable que si la proportion (respectivement le taux) est petite et le nombre d'échantillons  $n$  est grand.

## 10.37. Exercice 68.: Intervalle de confiance de l'écart-type (test du khi-2 de la variance)

Minitab® Statistical Software 15.1.1.0

Un échantillon de 9 vis a été tiré d'une ligne de production et la mesure de leur diamètre en [mm] a été reportée ci-dessous:



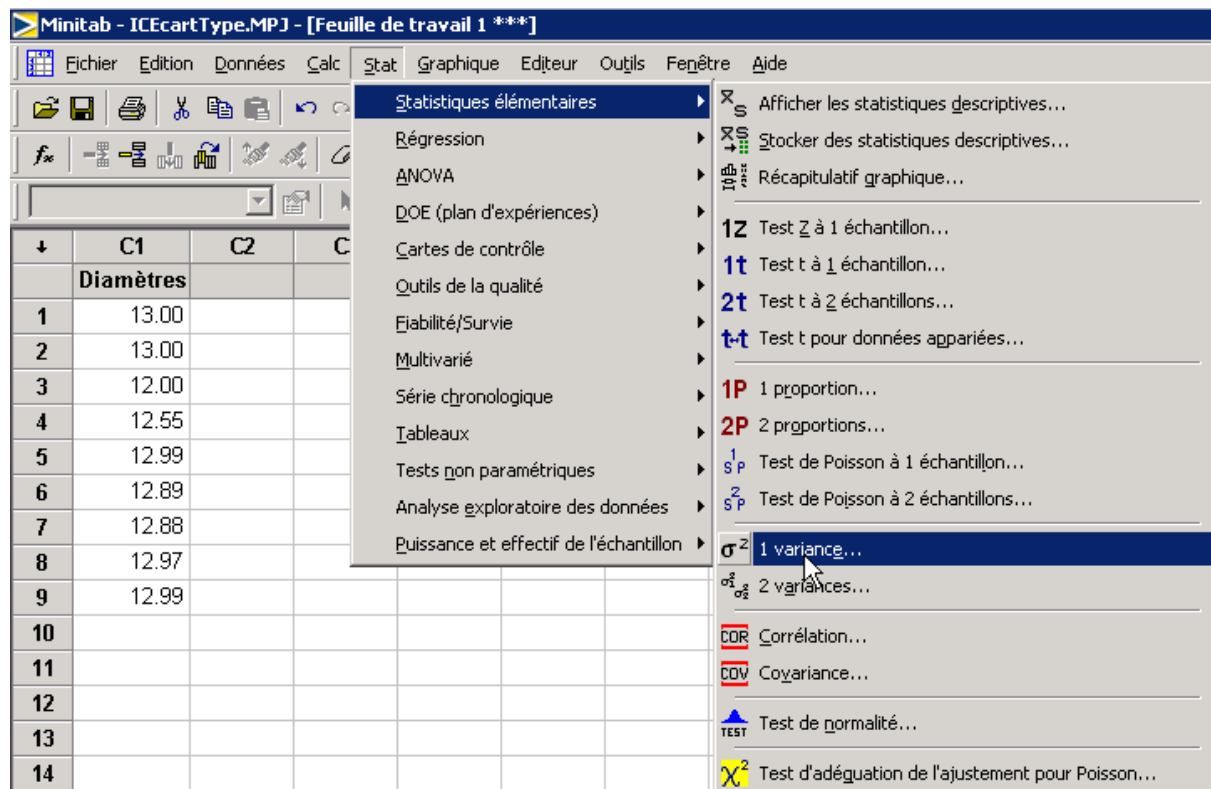
The screenshot shows the Minitab Statistical Software interface. The main window displays a data table with 16 columns (C1 to C16) and 16 rows. The first column (C1) is labeled 'Diamètres'. The data values for rows 1 through 9 are: 13.00, 13.00, 12.00, 12.55, 12.99, 12.89, 12.88, 12.97, and 12.99. A mouse cursor is visible over the cell in row 7, column C5. The status bar at the bottom indicates 'Feuille de travail en cours : Feuille de travail 1'.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16
1	13.00															
2	13.00															
3	12.00															
4	12.55															
5	12.99															
6	12.89															
7	12.88															
8	12.97															
9	12.99															
10																
11																
12																
13																
14																
15																
16																

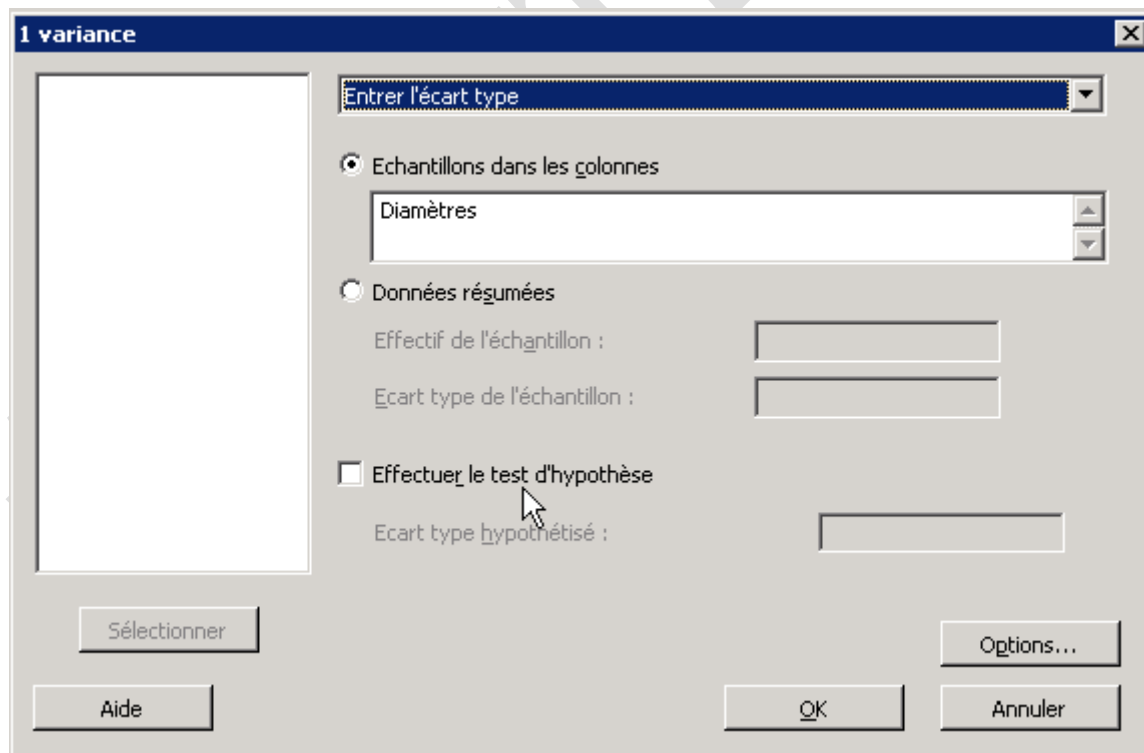
Nous voulons calculer l'écart-type et estimer l'intervalle de confiance de l'écart-type à 95%.

L'objectif va être encore une fois de vérifier que nous obtenons les mêmes résultats que ceux démontrés dans les cours théorique de statistique et qu'avec l'utilitaire d'analyse de Microsoft Excel.

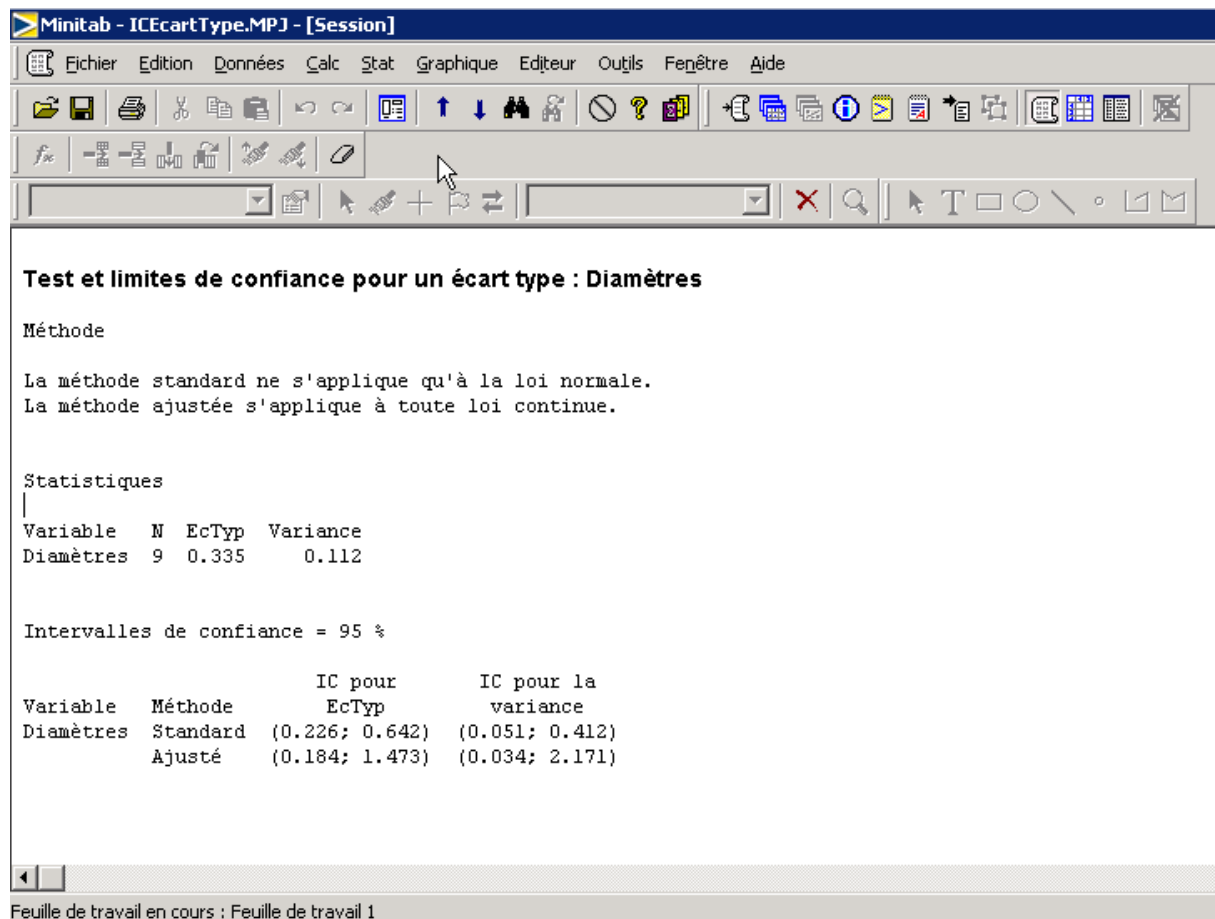
Nous allons dans le menu **Stat/Statistiques élémentaires/1 variance...**:



Nous avons alors:



Nous validons par **OK**:



**Minitab - ICEcartType.MPJ - [Session]**

Fichier Edition Données Calc Stat Graphique Éditeur Outils Fenêtre Aide

Test et limites de confiance pour un écart type : Diamètres

Méthode

La méthode standard ne s'applique qu'à la loi normale.  
La méthode ajustée s'applique à toute loi continue.

Statistiques

Variable	N	EcTyp	Variance
Diamètres	9	0.335	0.112

Intervalles de confiance = 95 %

Variable	Méthode	IC pour EcTyp	IC pour la variance
Diamètres	Standard	(0.226; 0.642)	(0.051; 0.412)
	Ajusté	(0.184; 1.473)	(0.034; 2.171)

Feuille de travail en cours : Feuille de travail 1

et tout est parfait! Nous retrouvons les mêmes résultats que dans le cours théorique et qu'avec Microsoft Excel.

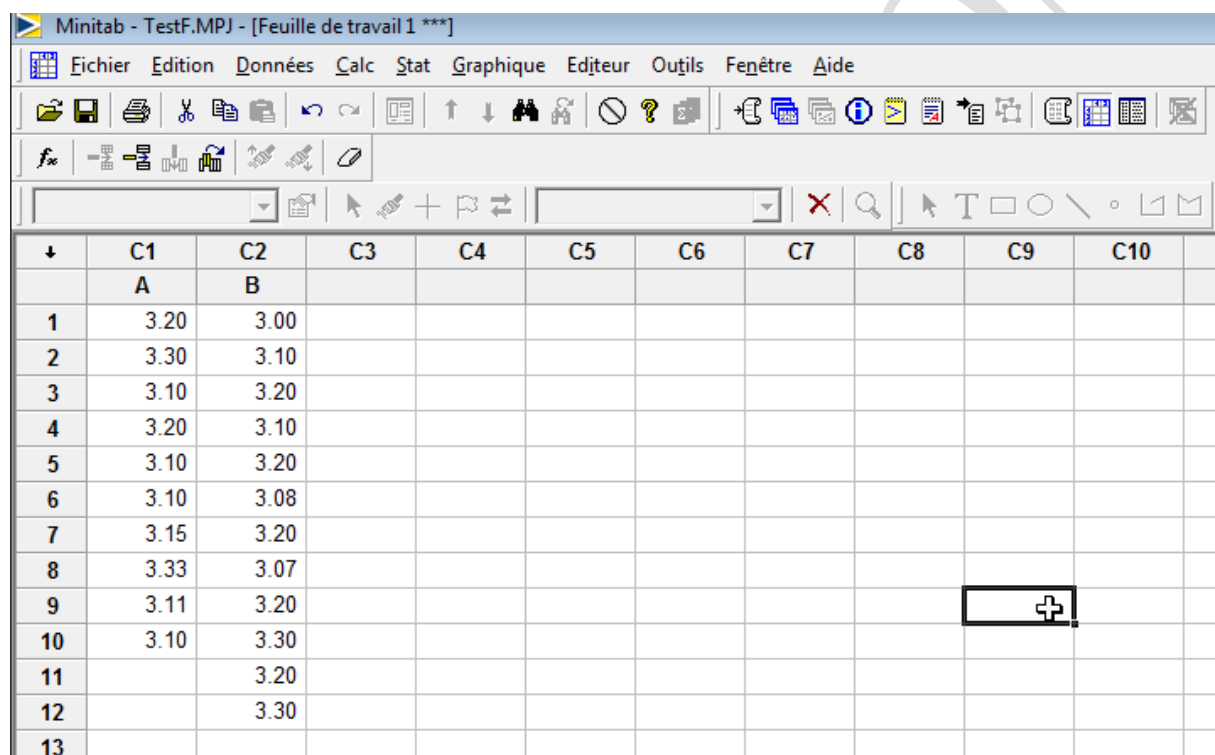
## 10.38. Exercice 69.: Test de Fisher d'égalité des variances

Minitab® Statistical Software 15.1.1.0

Ouvrez le fichier *TestF.mpj*:

Une société reçoit des joints d'étanchéité de deux fournisseurs. Le gestionnaire AQ (assurance qualité) de la société veut comparer la variance de l'épaisseur des joints des deux fournisseurs pour vérifier si les variances sont égales avec un niveau de confiance  $\alpha = 0.05$ .

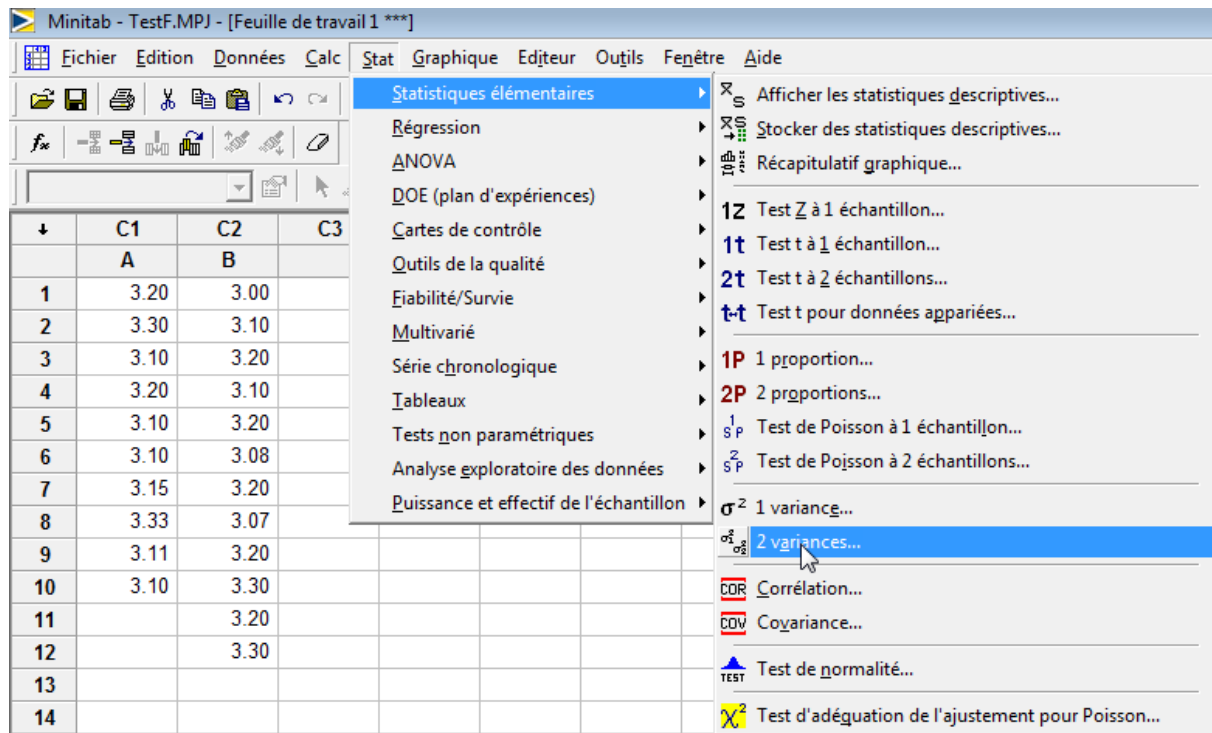
Pour cela, il prend un échantillon de 10 joints du fournisseur A et de 12 joints du fournisseur B. Voici le tableau avec l'épaisseur des joints:



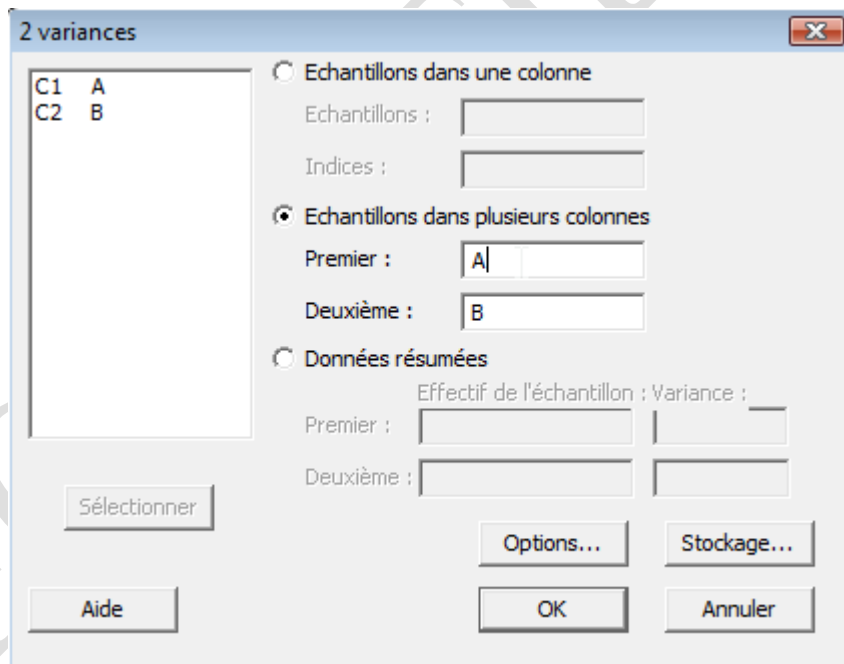
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
	A	B								
1	3.20	3.00								
2	3.30	3.10								
3	3.10	3.20								
4	3.20	3.10								
5	3.10	3.20								
6	3.10	3.08								
7	3.15	3.20								
8	3.33	3.07								
9	3.11	3.20								
10	3.10	3.30								
11		3.20								
12		3.30								
13										

L'objectif va être encore une fois de vérifier que nous obtenons les mêmes résultats que ceux démontrés dans les cours théorique de statistique et qu'avec l'utilitaire d'analyse de Microsoft Excel.

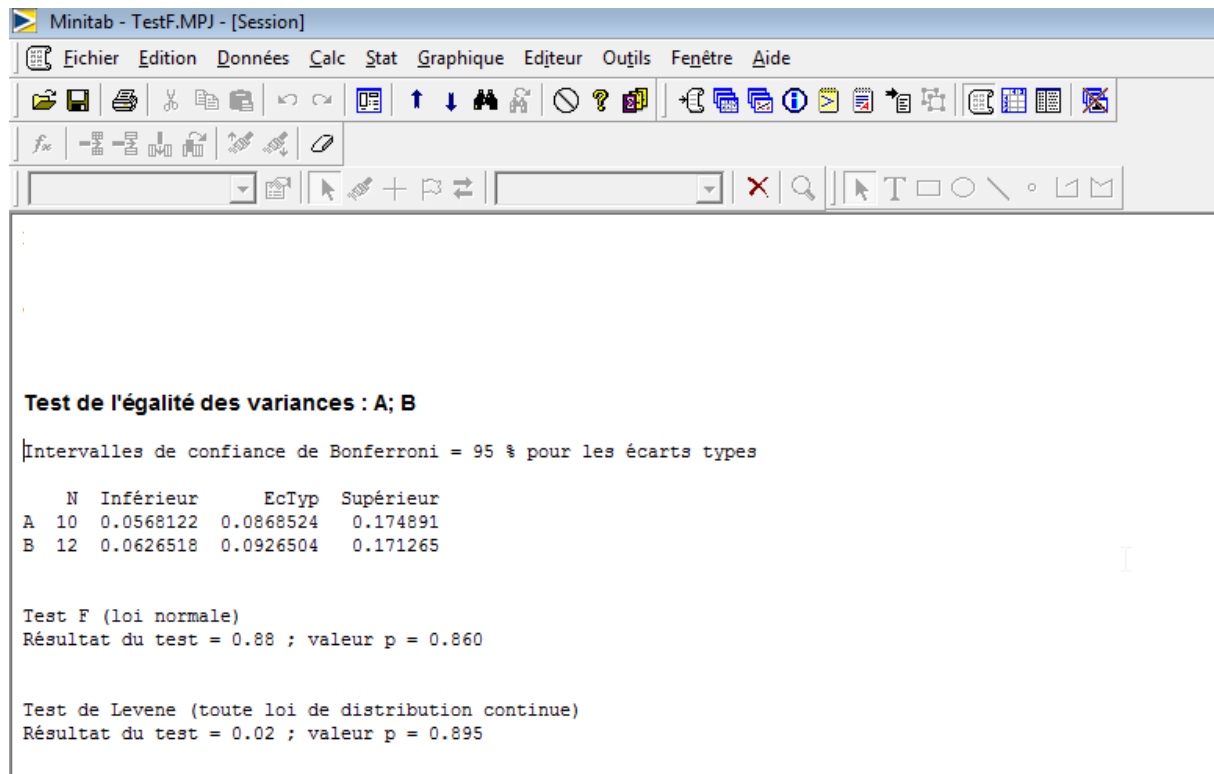
Nous allons dans le menu **Stat/Statistiques élémentaires/2 variances...**:



Nous avons:



Nous validons par **OK** et nous trouvons alors dans la fenêtre de sessions:



Minitab - TestF.MPJ - [Session]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

Test de l'égalité des variances : A; B

Intervalles de confiance de Bonferroni = 95 % pour les écarts types

	N	Inférieur	EcTyp	Supérieur
A	10	0.0568122	0.0868524	0.174891
B	12	0.0626518	0.0926504	0.171265

Test F (loi normale)  
Résultat du test = 0.88 ; valeur p = 0.860

Test de Levene (toute loi de distribution continue)  
Résultat du test = 0.02 ; valeur p = 0.895

Remarque: Nous voyons ci-dessus que Minitab® Statistical Software 15 utilise la correction de Bonferroni que nous avons étudié en détails dans le cours théorique pour l'intervalle de confiance des écart-types des échantillons.

Le résultat est à comparer par curiosité avec ce que donne par la version 17 juste pour rigoler... (observez bien l'intervalle de confiance des écarts-types de A et de B qui sont significativement différents!):



**Test et IC pour 2 variances : A; B**

Méthode

Hypothèse nulle  $\sigma(A) / \sigma(B) = 1$   
Hypothèse alternative  $\sigma(A) / \sigma(B) \neq 1$   
Seuil de signification  $\alpha = 0,05$

La méthode F de Fisher a été utilisée. Elle est exacte uniquement pour des données normales.

Statistiques

Variable	N	EcTyp	Variance	IC à 95 % pour les écarts types
A	10	0,087	0,008	(0,060; 0,159)
B	12	0,093	0,009	(0,066; 0,157)

Rapport des écarts types = 0,937  
Rapport de variances = 0,879

Intervalles de confiance = 95 %

Méthode	IC du rapport des écarts types	IC du rapport des variances
F	(0,495; 1,854)	(0,245; 3,438)

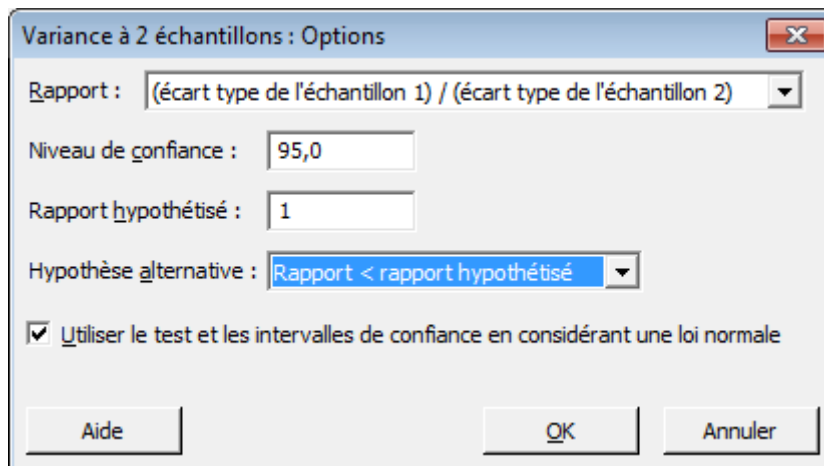
Tests

Méthode	DL1	DL2	Statistique du test	Valeur de P
F	9	11	0,88	0,860

Nous voyons que le résultat du test est comme dans Microsoft Excel en ce qui concerne dans un premier temps le quantile du test de Fisher (pour une distribution sous-jacente normalement distribuée) de 0.88. La  $p$ -value est par contre de 0.860 en bilatéral. Nous avons vu effectivement que Microsoft Excel donnait le test en unilatéral et nous en concluons donc in extenso que Minitab® Statistical Software donne la  $p$ -value en bilatéral ce que nous pouvons facilement constater en allant dans le bouton **Options...** de la boîte de dialogue du test.

Bon dans le cas présent de toute façon, la  $p$ -value en bilatérale étant de 0.860 (donnée par Minitab® Statistical Software) et la unilatérale étant de 0.43 (donnée par Microsoft Excel), dans les deux cas, l'hypothèse de l'égalité des variances ne peut pas être rejetée à un niveau alpha de 5% (donc les variances sont probablement égales!).

Nous pouvons retrouver le résultat de Microsoft Excel de la manière suivante avec Minitab dans la boîte de dialogue du test de Fisher nous cliquons sur **Options...** pour prendre:



En validant deux fois par **OK** nous obtenons:

Minitab - TestF.mpj - [Session]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

Hypothèse nulle  $\sigma(A) / \sigma(B) = 1$   
 Hypothèse alternative  $\sigma(A) / \sigma(B) < 1$   
 Seuil de signification  $\alpha = 0,05$

La méthode F de Fisher a été utilisée. Elle est exacte uniquement pour des données normales.

Statistiques

Variable	N	EcTyp	Variance	Borne supérieure de 95 % pour les écarts types
A	10	0,087	0,008	0,143
B	12	0,093	0,009	0,144

Rapport des écarts types = 0,937  
 Rapport de variances = 0,879

Intervalles de confiance unilatéraux = 95 %

Méthode	Borne supérieure du rapport des écarts types	Borne supérieure du rapport de variances
F	1,651	2,726

Tests

Méthode	DL1	DL2	Statistique du test	Valeur de P
F	9	11	0,88	0,430

Attention! Les versions antérieures et égales à Minitab® Statistical Software 16 n'appliquent pas la correction de Bonferroni ou le test de Levene si des données résumées sont saisies!

ÉCHANTILLON

## 10.39. Exercice 70.: Test de Levene d'égalité de deux variances

Minitab® Statistical Software 15.1.1.0

Le but ici va être de vérifier les résultats obtenus avec Microsoft Excel par simulation d'abord dans le cas de deux variances. Dans l'exercice suivant, nous ferons la même vérification mais avec un tableau d'ANOVA.

Rappelons d'abord que nous avons montré dans le cours de statistique théorique que le test de Levene est en fait une approche empirique du test de Fisher en 4 versions avec des transformations empiriques des observations:

- Levene1 (dit aussi "Levene absolu"):

$$z_{ij} = |x_{ij} - \bar{x}_i|$$

- Levene1 (variante de Brown-Forsythe avec médiane donc version non paramétrique):

$$z_{ij} = |x_{ij} - \bar{x}_{M,i}|$$

- Levene2 (dit aussi "Levene Quadratique"):

$$z_{ij} = (x_{ij} - \bar{x}_i)^2$$

- Levene3:

$$z_{ij} = \log\left((x_{ij} - \bar{x}_i)^2\right)$$

- Levene4:

$$z_{ij} = \sqrt{|x_{ij} - \bar{x}_i|}$$

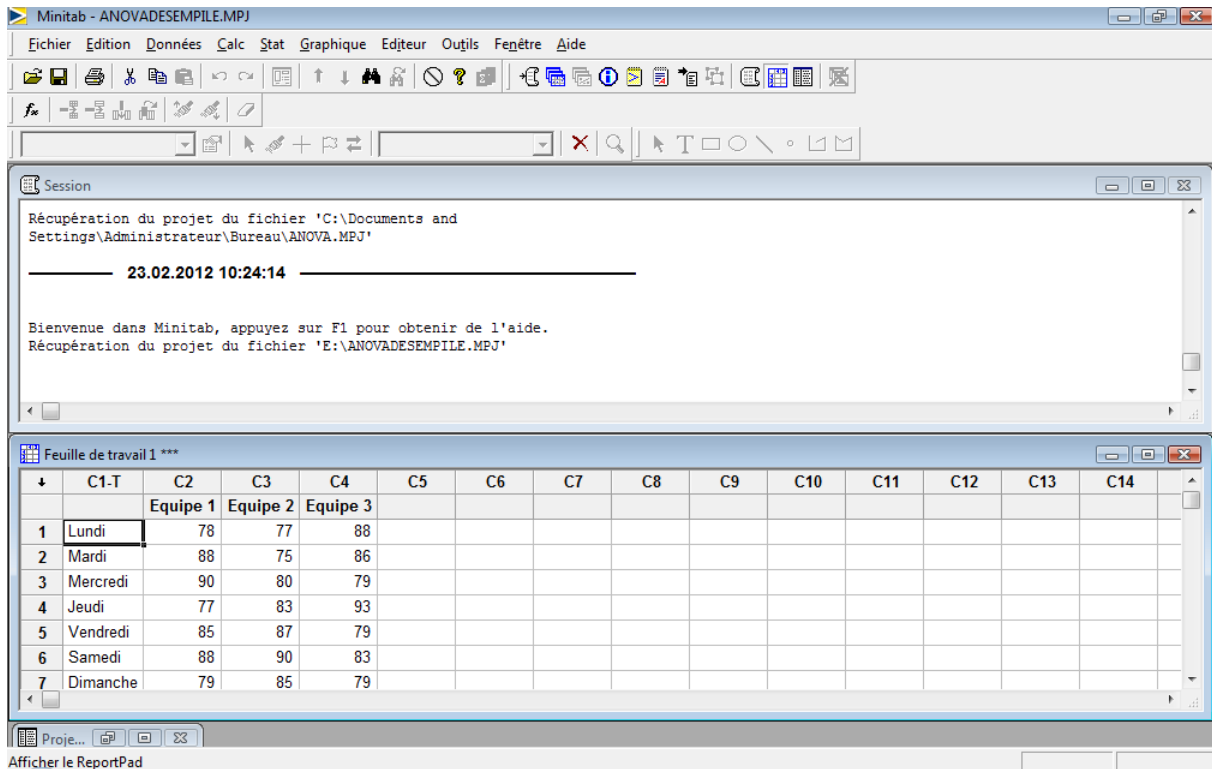
Par simulation de Monte-Carlo nous avons effectivement observé dans le cours théorique que le test de Fisher traditionnel est sensible à la non normalité des données (faux négatifs) alors que la deuxième variante du test de Levene (Brown-Forsythe) l'était beaucoup... beaucoup moins et il serait donc cohérent que Minitab utilise cette version.

*Remarque:* Rappelons que Levene lui-même ne l'a démontré que par des simulations de Monte-Carlo (et seulement pour la première variante!).

Nous souhaitons donc vérifier ici si nous retrouvons les mêmes résultats numériques qu'avec Microsoft Excel et déterminer exactement quelle variante Minitab utilise... (puisque cela n'est a priori pas communiqué dans la documentation).

Nous allons donc reprendre le tableau du fichier *ANOVA-DESEMPILE.mpj* et faire une comparaison des variances seulement de deux équipes et vérifier que nous retombons bien sur les valeurs obtenues dans Microsoft Excel.

Nous partons donc du tableau:



The screenshot displays the Minitab interface. The main window shows a data table with the following content:

	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
		Equipe 1	Equipe 2	Equipe 3										
1	Lundi	78	77	88										
2	Mardi	88	75	86										
3	Mercredi	90	80	79										
4	Jeudi	77	83	93										
5	Vendredi	85	87	79										
6	Samedi	88	90	83										
7	Dimanche	79	85	79										

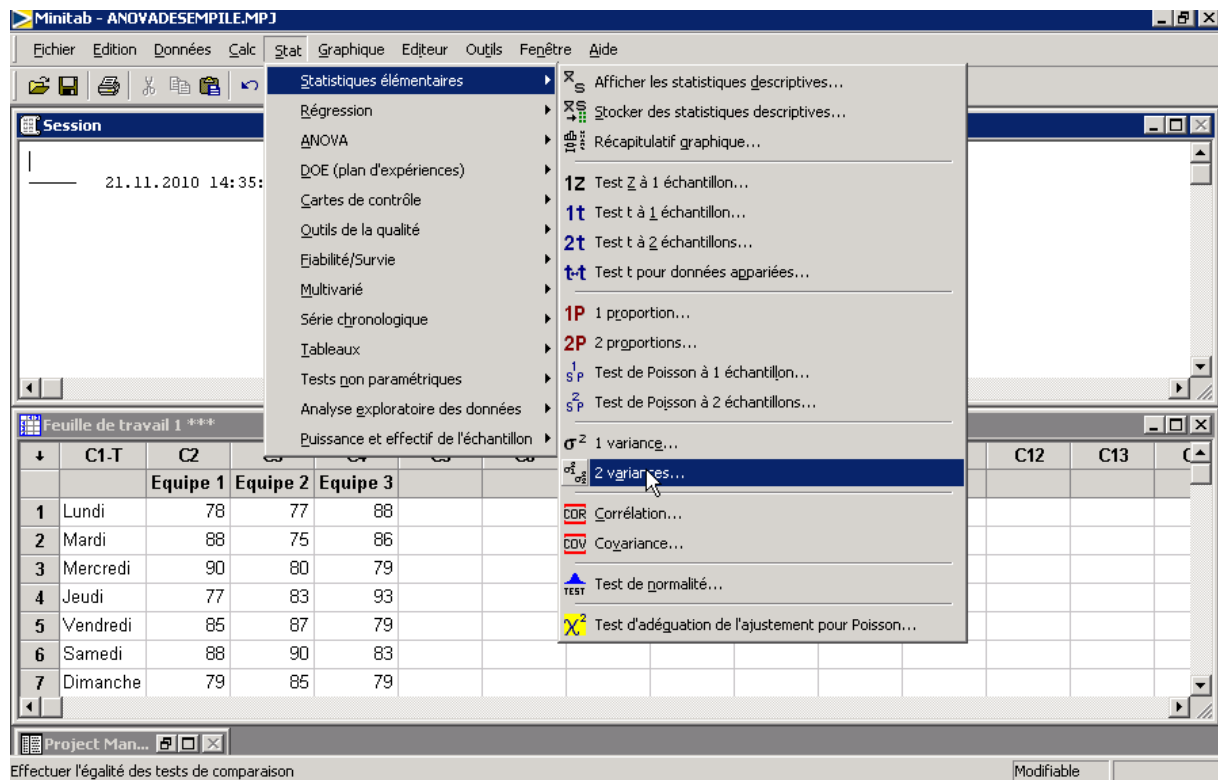
The Session window above the table shows the following text:

```
Session
Récupération du projet du fichier 'C:\Documents and
Settings\Administrateur\Bureau\ANOVA.MPJ'

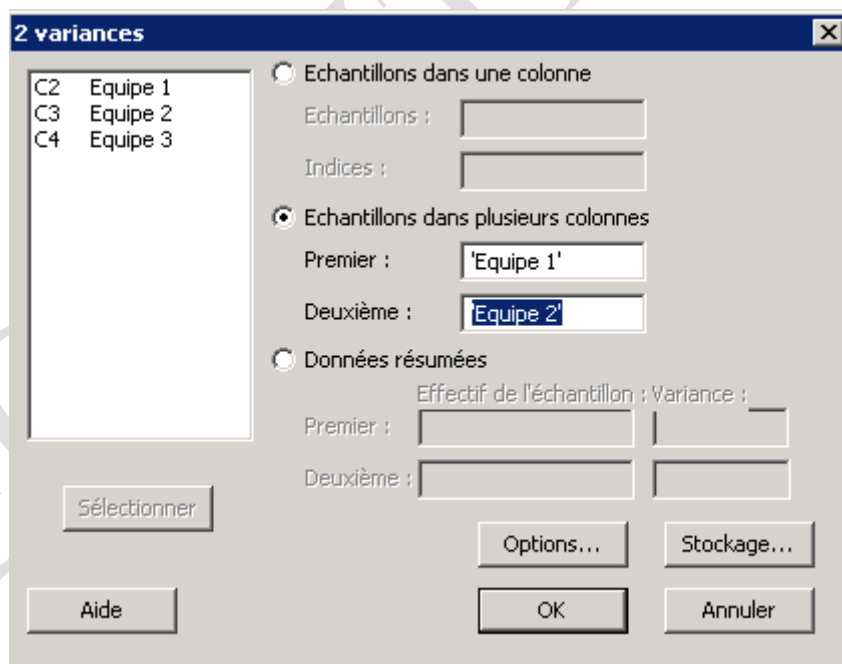
----- 23.02.2012 10:24:14 -----

Bienvenue dans Minitab, appuyez sur F1 pour obtenir de l'aide.
Récupération du projet du fichier 'E:\ANOVADESEMPILE.MPJ'
```

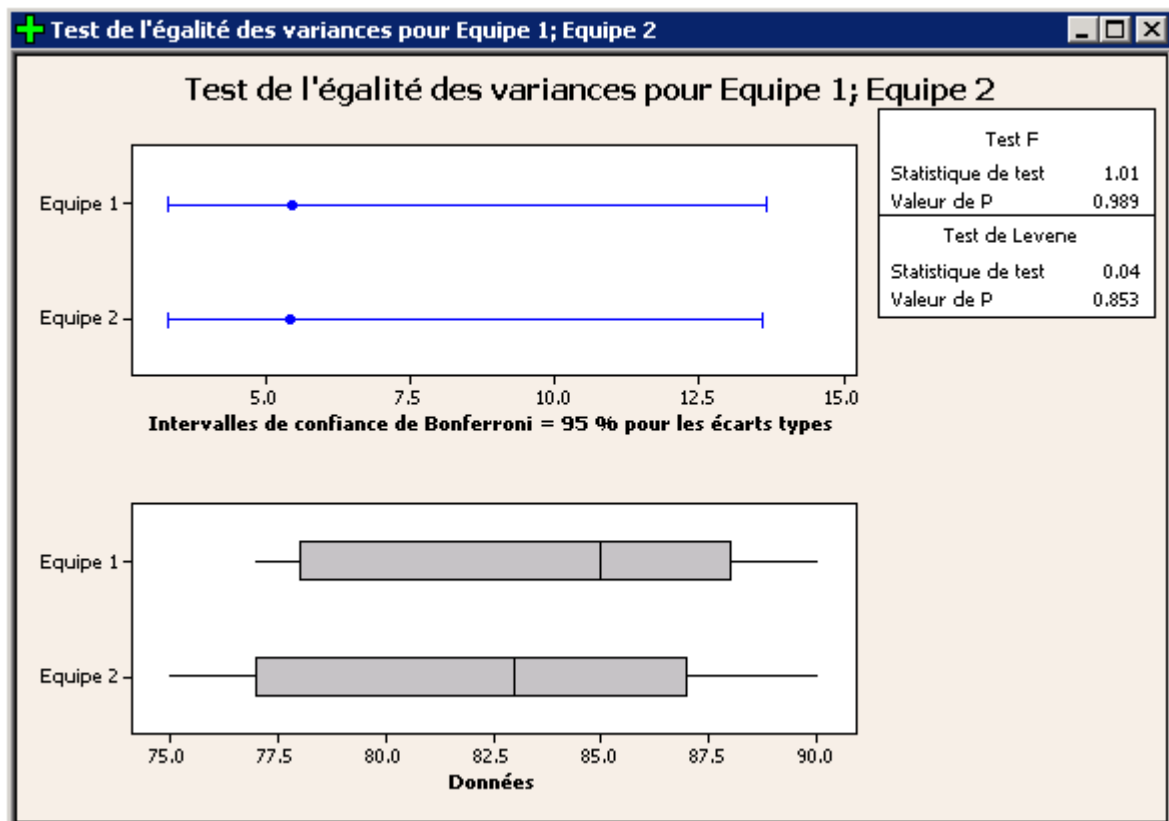
et nous allons dans le menu **Stat/Statistiques élémentaires/2 variances...**:



et nous prenons donc que deux équipes:



et validons par **OK** pour obtenir comme graphique:



et dans la fenêtre de session:

### Test de l'égalité des variances : Equipe 1; Equipe 2

Intervalle de confiance de Bonferroni = 95 % pour les écarts types

	N	Inférieur	EcTyp	Supérieur
Equipe 1	7	3.30756	5.44234	13.6903
Equipe 2	7	3.28890	5.41163	13.6130

Test F (loi normale)  
 Résultat du test = 1.01 ; valeur p = 0.989

Test de Levene (toute loi de distribution continue)

Résultat du test = 0.04 ; valeur p = 0.853

et ce qui nous intéresse est la partie mise en évidence ci-dessus

Nous retrouvons la même valeur que le résultat du test (rapport  $Msk/MSE$ ) que dans Microsoft Excel mais avec moins de précision puisque nous avons obtenu 0.0360. Donc nous pouvons déjà conclure que Minitab utilise la variante de Brown-Forsythe.

Concernant la  $p$ -value nous avons obtenu également la même avec Microsoft Excel mais avec plus de précision.

Donc... objectif atteint! Nous ne ferons cependant pas l'exemple qui consiste à transformer le tableau entier en l'adaptant pour une ANOVA car franchement... autant le faire dans Microsoft Excel!

ÉCHANTILLON

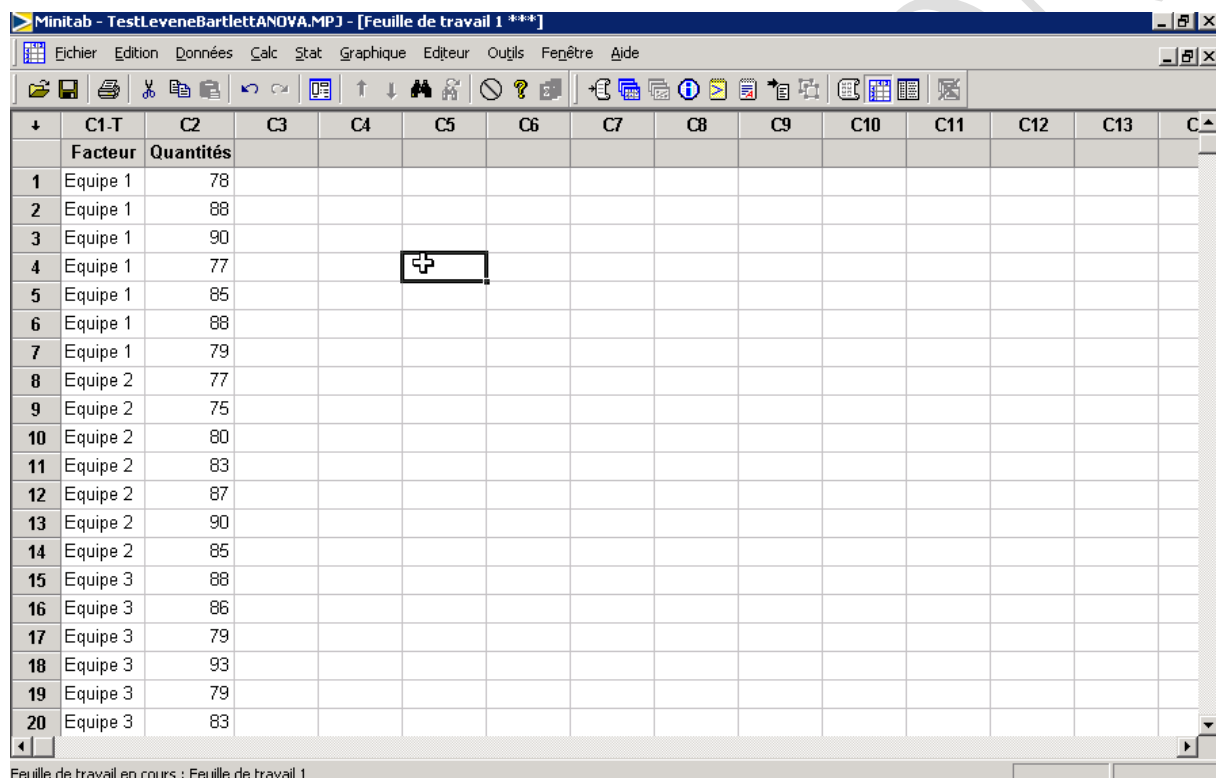


## 10.40. Exercice 71.: Test de Levene et Bartlett d'égalité des variances d'une ANOVA

Minitab® Statistical Software 15.1.1.0

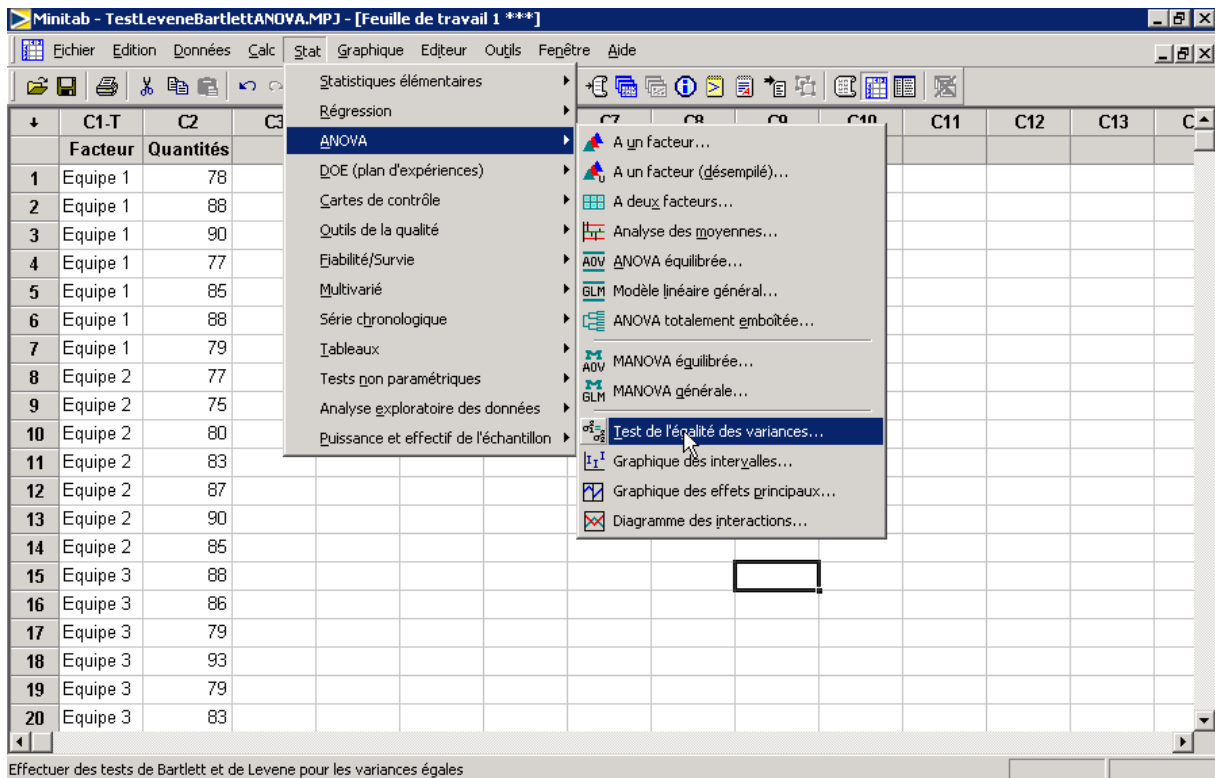
L'objectif ici va être le cas plus intéressant d'application du test de Levene aux données d'une ANOVA empilée et de vérifier encore une fois si les résultats sont conformes aux simulations de Monte-Carlo effectuées dans le cours théoriques et à l'application numérique effectuée dans Microsoft Excel.

Nous partons donc du tableau suivant *TestLeveneBartlettANOVA.mpj*:

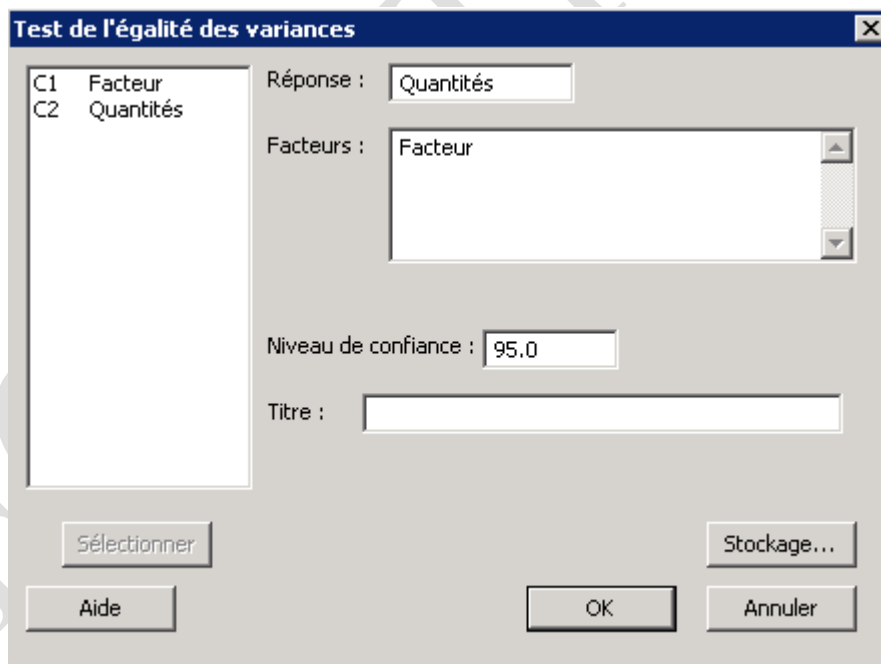


	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Facteur	Quantités												
1	Equipe 1	78												
2	Equipe 1	88												
3	Equipe 1	90												
4	Equipe 1	77												
5	Equipe 1	85												
6	Equipe 1	88												
7	Equipe 1	79												
8	Equipe 2	77												
9	Equipe 2	75												
10	Equipe 2	80												
11	Equipe 2	83												
12	Equipe 2	87												
13	Equipe 2	90												
14	Equipe 2	85												
15	Equipe 3	88												
16	Equipe 3	86												
17	Equipe 3	79												
18	Equipe 3	93												
19	Equipe 3	79												
20	Equipe 3	83												

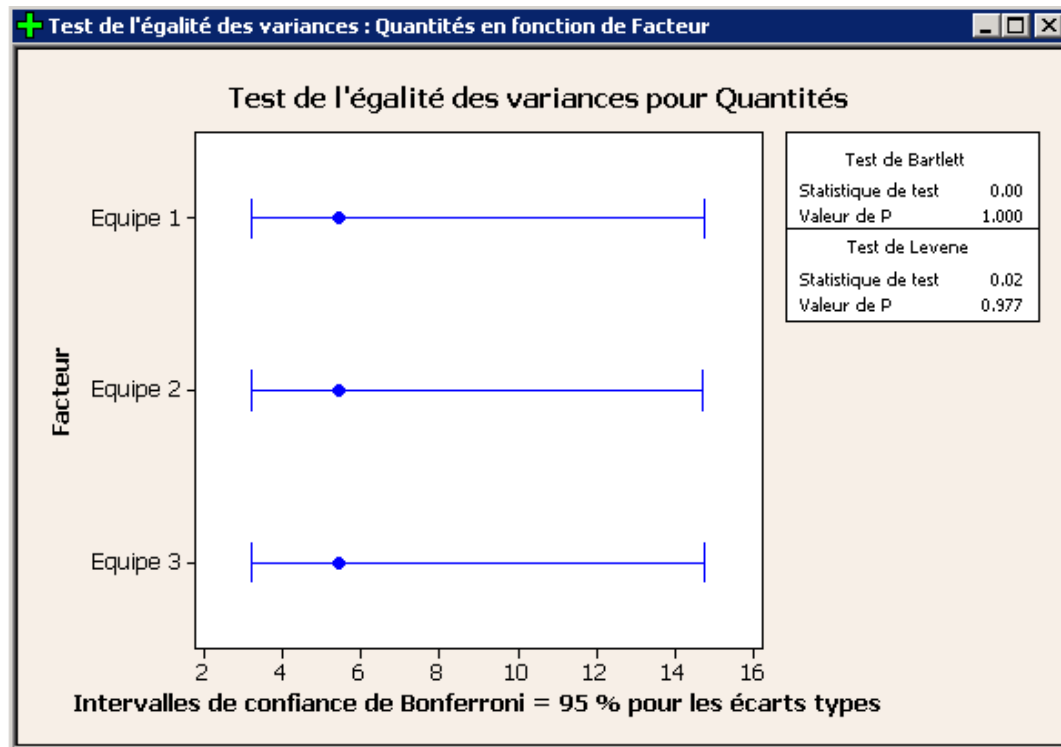
Nous allons dans le menu **Stat/ANOVA/Test de l'égalité des variances...**:



nous y mettons:



Nous validons par **OK** pour obtenir d'abord comme graphique:



et dans la fenêtre de session:

### Test de l'égalité des variances : Quantités en fonction de Facteur

Intervalles de confiance de Bonferroni = 95 % pour les écarts types

Facteur	N	Inférieur	EcTyp	Supérieur
Equipe 1	7	3.20766	5.44234	14.7671
Equipe 2	7	3.18955	5.41163	14.6837
Equipe 3	7	3.19991	5.42920	14.7314

Test de Bartlett (loi normale)  
 Résultat du test = 0.00 ; valeur p = 1.000

Test de Levene (toute loi de distribution continue)  
 Résultat du test = 0.02 ; valeur p = 0.977

D'abord les intervalles de confiance sont conformes à l'intervalle classique vu dans le cours théorique pour l'écart-type.

Ensuite, nous ferons abstraction du test de Bartlett dont nous n'avons pas vu la démonstration mathématique dans le cours théorique puisque les détails de celle-ci est introuvable dans les livres et même dans l'article d'origine de Bartlett lui-même.

Ce qui va nous intéresser est donc la valeur du test de Levene. Minitab indique comme "résultat du test" la valeur de 0.02. Ce qui est conforme à Microsoft Excel où nous avons obtenu 0.02307. Concernant la  $p$ -value Minitab nous donne 0.977 ce qui encore une fois est conforme à Microsoft Excel puisque nous avons obtenu 0.9772. Donc tout est OK pour les vérifications. Sujet clos!

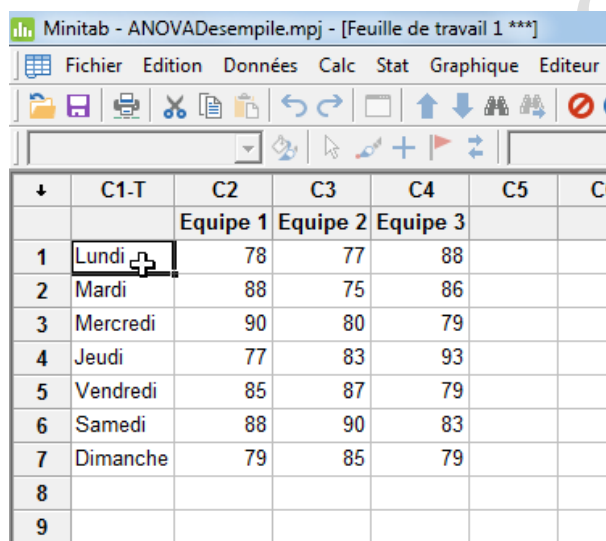
## 10.41. Exercice 72.: Test de Dunnett de comparaison des moyennes à une valeur de contrôle

Minitab® Statistical Software 18.1

Nous allons vérifier ici si la manière dont le test de Dunnett de comparaison des moyennes correspond bien à ce que nous avons vu dans le cours théorique!

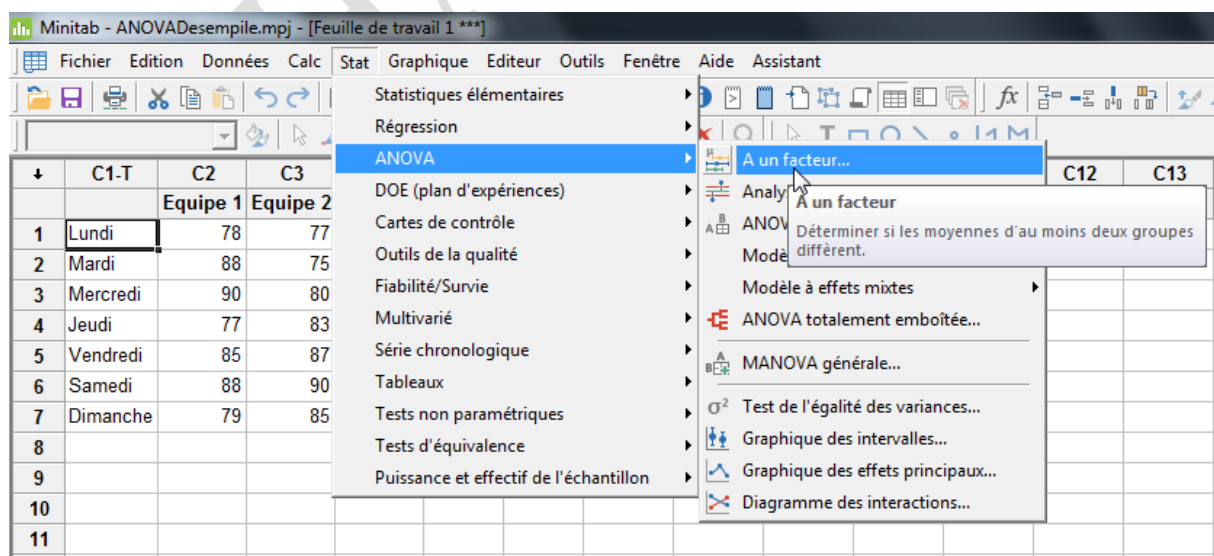
D'abord rappelons que la différence majeure par rapport aux tests précédents est que nous devons choisir un niveau de référence (d'où son intérêt)!

Nous partons du jeu de données du fichier *ANOVADesempilee*:



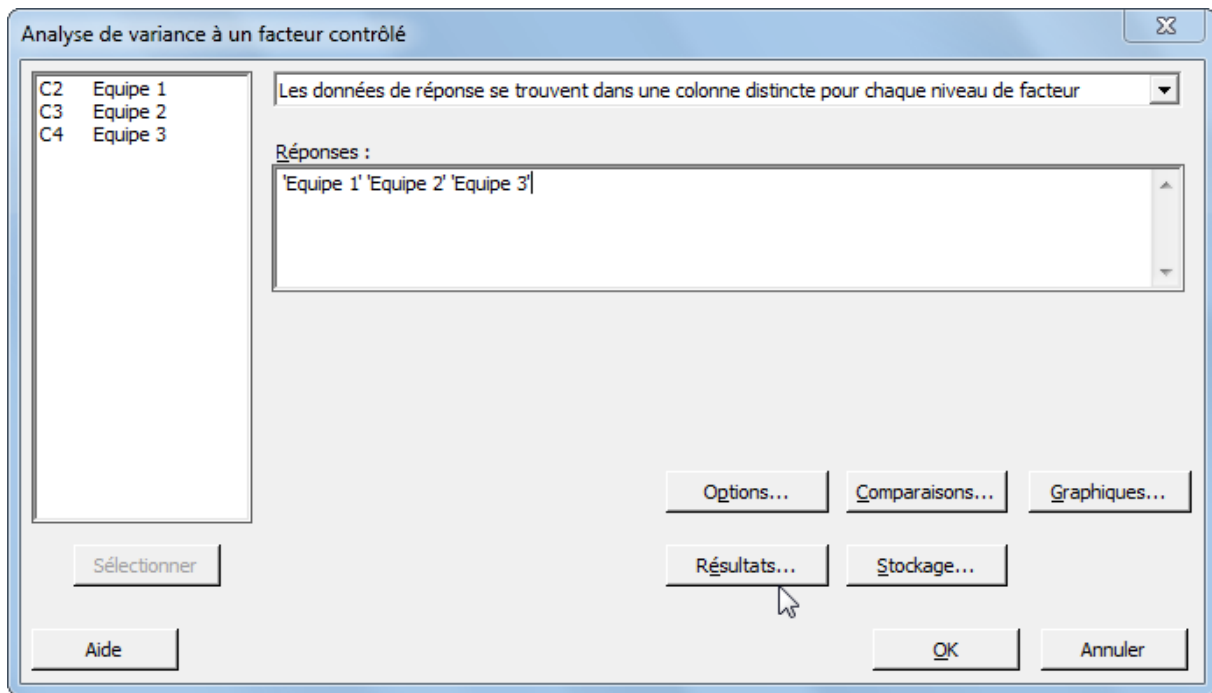
	C1-T	C2	C3	C4	C5	C6
		Equipe 1	Equipe 2	Equipe 3		
1	Lundi	78	77	88		
2	Mardi	88	75	86		
3	Mercredi	90	80	79		
4	Jeudi	77	83	93		
5	Vendredi	85	87	79		
6	Samedi	88	90	83		
7	Dimanche	79	85	79		
8						
9						

Ensuite, nous allons dans le menu **Stat/ANOVA/A un facteur...**:

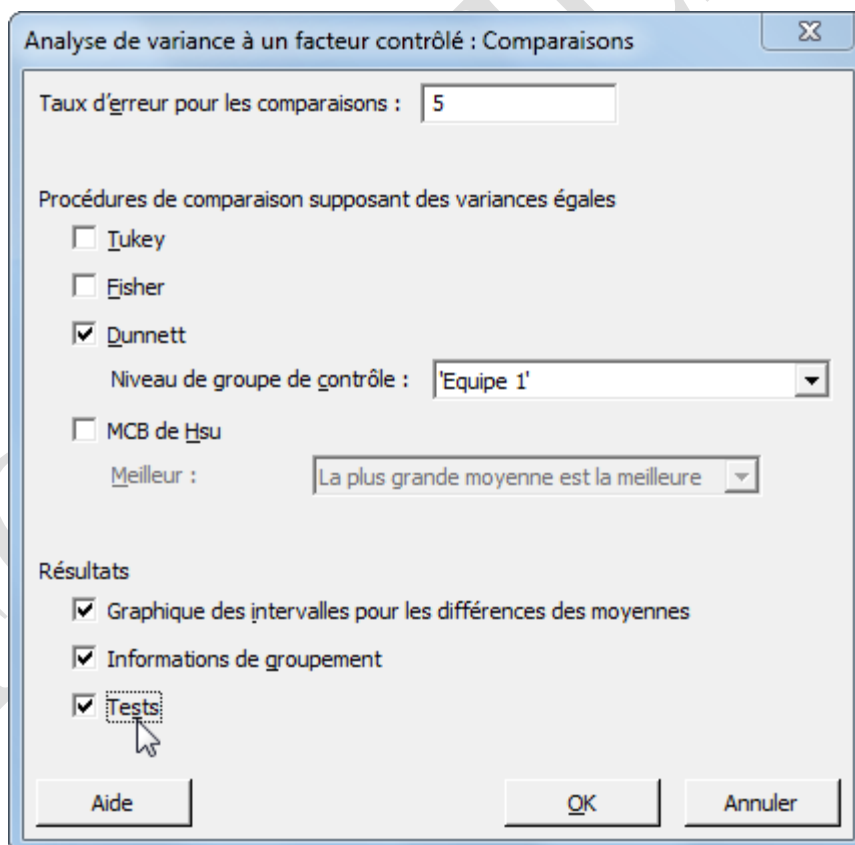


The screenshot shows the Minitab software interface with the 'Stat' menu open and 'ANOVA' selected. The 'ANOVA' submenu is also open, showing 'A un facteur...' as the selected option. A tooltip for 'A un facteur' is visible, stating: 'Déterminer si les moyennes d'au moins deux groupes diffèrent.'

Nous mettons alors les trois niveaux dans la boîte de dialogue qui nous intéresse:



Et cliquons sur le bouton **Options...** pour y choisir et cocher les options suivantes:



Ce qui donne dans la fenêtre la classique ANOVA:

## ANOVA à un facteur contrôlé : Equipe 1; Equipe 2; Equipe 3

### Méthode

Hypothèse nulle Toutes les moyennes sont égales  
 Hypothèse alternative Toutes les moyenne ne sont pas égales  
 Seuil de signification  $\alpha = 0.05$

*Les variances ont été supposées égales pour l'analyse.*

### Informations sur les facteurs

Facteur	Niveaux	Valeurs
Facteur	3	Equipe 1; Equipe 2; Equipe 3

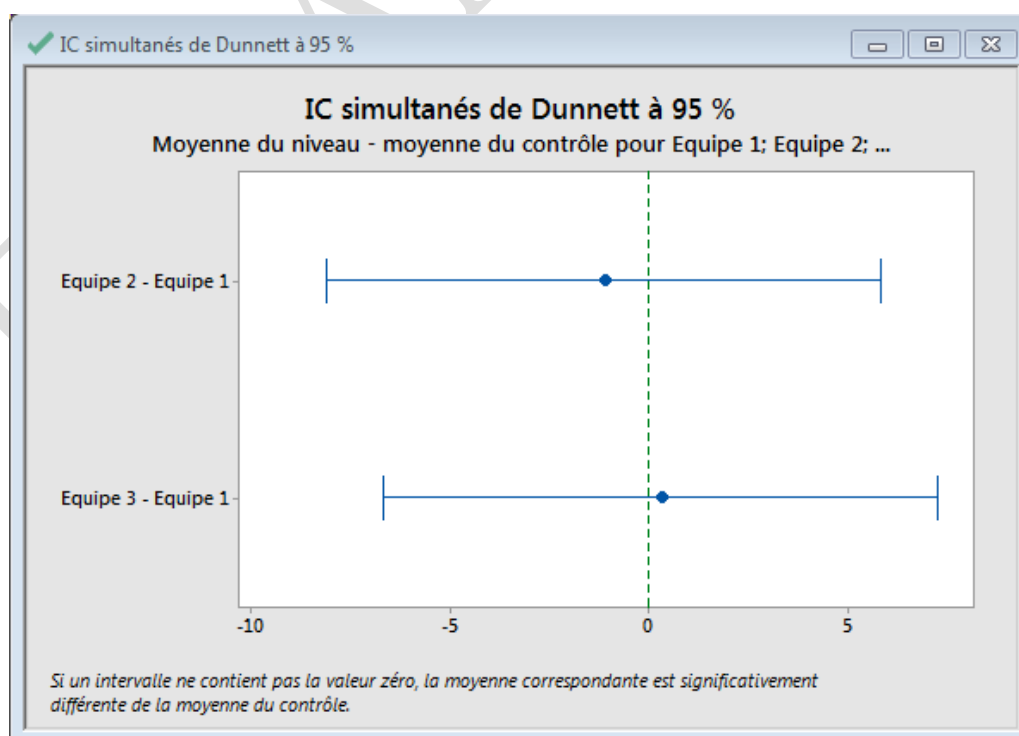
### Analyse de variance

Source	DL	SomCar		Valeur F	Valeur de p
		ajust	CM ajust		
Facteur	2	8.000	4.000	0.14	0.874
Erreur	18	530.286	29.460		
Total	20	538.286			

### Récapitulatif du modèle

S	R carré	R carré	R carré
		(ajust)	(prév)
5.42774	1.49%	0.00%	0.00%

Mais spécifique au test de Dunnett nous avons d'abord le graphique suivant:



Et dans la fenêtre de session relativement au test de Dunnett nous avons les informations spécifiques d'intérêt ci-dessous:

### Moyennes

Facteur	N	Moyenne	EcTyp	IC à 95 %
Equipe 1	7	83.57	5.44	(79.26; 87.88)
Equipe 2	7	82.43	5.41	(78.12; 86.74)
Equipe 3	7	83.86	5.43	(79.55; 88.17)

*Ecart type regroupé = 5.42774*

### Comparaisons multiples de Dunnett avec un contrôle

#### Informations de groupement avec la méthode de Dunnett et un niveau de confiance de 95 %

Facteur	N	Moyenne	Groupement
Equipe 1 (contrôle)	7	83.57	A
Equipe 3	7	83.86	A
Equipe 2	7	82.43	A

*Les moyennes non étiquetées avec la lettre A sont significativement différentes de la moyenne du niveau de contrôle.*

#### Tests de simultanéité de Dunnett pour la moyenne du niveau – moyenne du contrôle

Différence des niveaux	Différence des moyennes	Erreur type de la différence		Valeur de T	Valeur de p ajustée
		de la différence	IC à 95 %		
Equipe 2 - Equipe 1	-1.14	2.90	(-8.10; 5.82)	-0.39	0.894
Equipe 3 - Equipe 1	0.29	2.90	(-6.67; 7.24)	0.10	0.993

*Niveau de confiance individuel = 97.25 %*

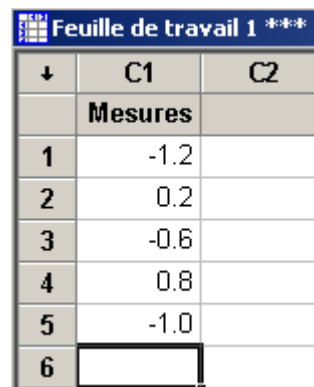
Nous retombons donc parfaitement sur les valeurs calculées à la main dans le cours théorique en utilisant les tables de Dunnett et nous ne sommes pas loin en première approximation non plus de la  $p$ -value sans utiliser les tables de Dunnett!

## 10.42. Exercice 73.: Test de normalité de Kolmogorov-Smirnov

Minitab® Statistical Software 15.1.1.0

Nous souhaitons ici simplement vérifier que nous obtenons les mêmes résultats que ceux obtenus dans Microsoft Excel avec la méthode Monte Carlo (et in extenso vérifier les résultats obtenus lors de l'étude théorique du test de Kolmogorov-Smirnov).

Nous ouvrons donc le fichier *TestNormaliteKS.mpj* qui contient des mesures de variations d'une cote de pièce par rapport à une cible dans un processus de fabrication sous contrôle statistique:



↓	C1	C2
	<b>Mesures</b>	
1	-1.2	
2	0.2	
3	-0.6	
4	0.8	
5	-1.0	
6		

Nous voulons faire un test d'ajustement de Kolmogorov-Smirnov avec une loi Normale centrée réduite à un risque alpha de 5%. Pour cela, nous allons dans le menu **Stat/Statistiques élémentaires/Test de normalité...**:



The screenshot shows the Minitab interface. The 'Stat' menu is open, and 'Test de normalité...' is highlighted. The background shows a worksheet with the following data:

	C1	C2	C3	C4	C5	C6
	Mesures					
1	-1.2					
2	0.2					
3	-0.6					

Nous avons alors à prendre:

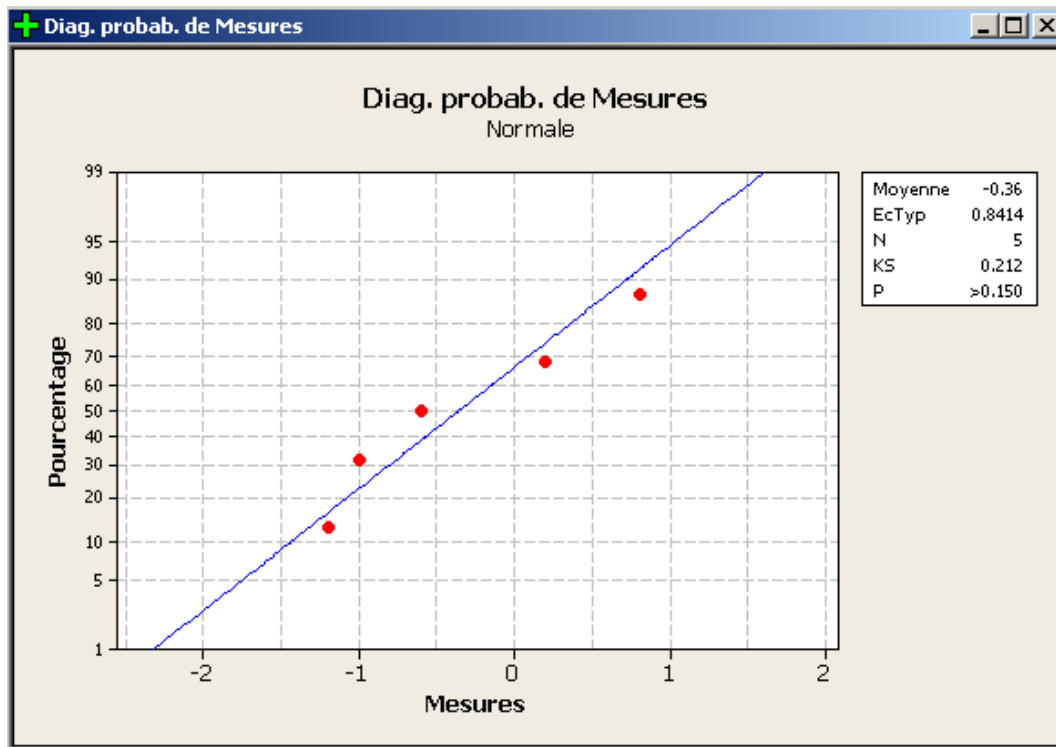
The 'Test de normalité' dialog box is shown with the following settings:

- Variable : Mesures
- Lignes des percentiles:
  - Aucune
  - Aux valeurs Y : [ ]
  - Aux valeurs : [ ]
- Tests de normalité:
  - Anderson-Darling
  - Ryan-Joiner (semblable à Shapiro-Wilk)
  - Kolmogorov-Smirnov
- Titre : [ ]

Buttons: Sélectionner, Aide, OK, Annuler

Où comme le lecteur pourra observer, nous n'avons **pas le choix du niveau alpha ni de la variance et de l'espérance de la loi normale de l'hypothèse nulle...** (c'est bien évidemment à cause du choix de la méthode statistique comme nous l'avons vu dans le cours théorique donc Minitab utilise a priori la méthode de Lilliefors) Nous allons donc voir ce qu'il en est un peu plus loin...

Nous validons par **OK**:



Bon que constatons-nous par rapport à ce que nous avons fait dans Microsoft Excel?:

1. Le logiciel calcule la moyenne arithmétique d'où le  $\mu = -0.36$  pour l'utiliser comme espérance de la loi Normale de l'hypothèse nulle (vous pouvez vérifier avec Microsoft Excel qu'il s'agit bien de l'écart-type).
2. Le logiciel calcule un écart-type sur une méthode qui nous est inconnue. Il utilise peut-être les statistiques des valeurs extrêmes (en tout cas c'est **très énervant que cela ne soit pas communiqué!!!**). Ce qui est sûr, c'est qu'il utilise cette valeur de  $\sigma = 0.8414$  comme écart-type de la loi Normale de l'hypothèse nulle.
3. En utilisant ces deux informations et sachant que Minitab® Statistical Software va se baser sur la loi Normale  $N(\mu, \sigma^2)$  si nous construisons la table vue dans le cours de statistique théorique:

$x$	$\hat{F}(x)$	$\Phi(x, \mu, \sigma^2)$	$ \hat{F}(x) - \Phi(x, \mu, \sigma^2) $
-1.2 <sup>-</sup>	0	0.159	0.159
-1.2	0.2	0.159	0.041
-1.0 <sup>-</sup>	0.2	0.223	0.023
-1.0	0.4	0.223	0.177
-0.6 <sup>-</sup>	0.4	0.388	0.012
-0.6	0.6	0.388	0.212
+0.2 <sup>-</sup>	0.6	0.747	0.147
+0.2	0.8	0.747	0.053

---

+0.8 <sup>-</sup>	0.8	0.916	0.116
+0.8	1	0.916	0.085

Nous y retrouvons donc la valeur de 0.212. Rare sont les logiciels à noter cette valeur *KS*.

Bref, le test est juste mais nous donne des informations peu utiles (à part la *p*-value) et nous propose peu d'options pour le choix de l'hypothèse nulle.

Pour ceux qui veulent une approche ludique de comparaison des trois tests de Normalité:

<http://blog.minitab.com/blog/the-statistical-mentor/anderson-darling-ryan-joiner-or-kolmogorov-smirnov-which-normality-test-is-the-best>

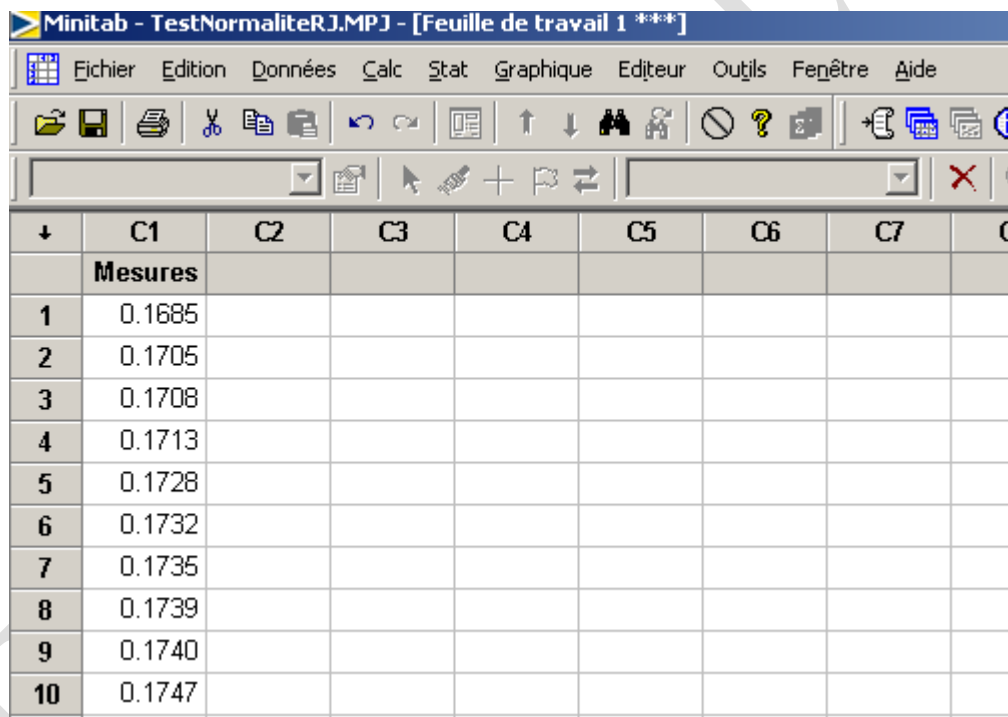
## 10.43. Exercice 74.: Test de normalité de Ryan-Joiner (ie Shapiro-Wilk)

Minitab® Statistical Software 15.1.1.0

Nous souhaitons ici vérifier que nous obtenons les mêmes résultats que ceux obtenus dans Microsoft Excel avec la méthode Monte Carlo (et in extenso vérifier les résultats obtenus lors de l'étude théorique du test de Ryan-Joiner).

Rappelons que l'avantage de ce test non paramétrique de normalité (donc basé sur les rangs) est sa simplicité mais qu'il n'est pas contre pas adapté lorsque que trop de valeurs identiques se répètent.

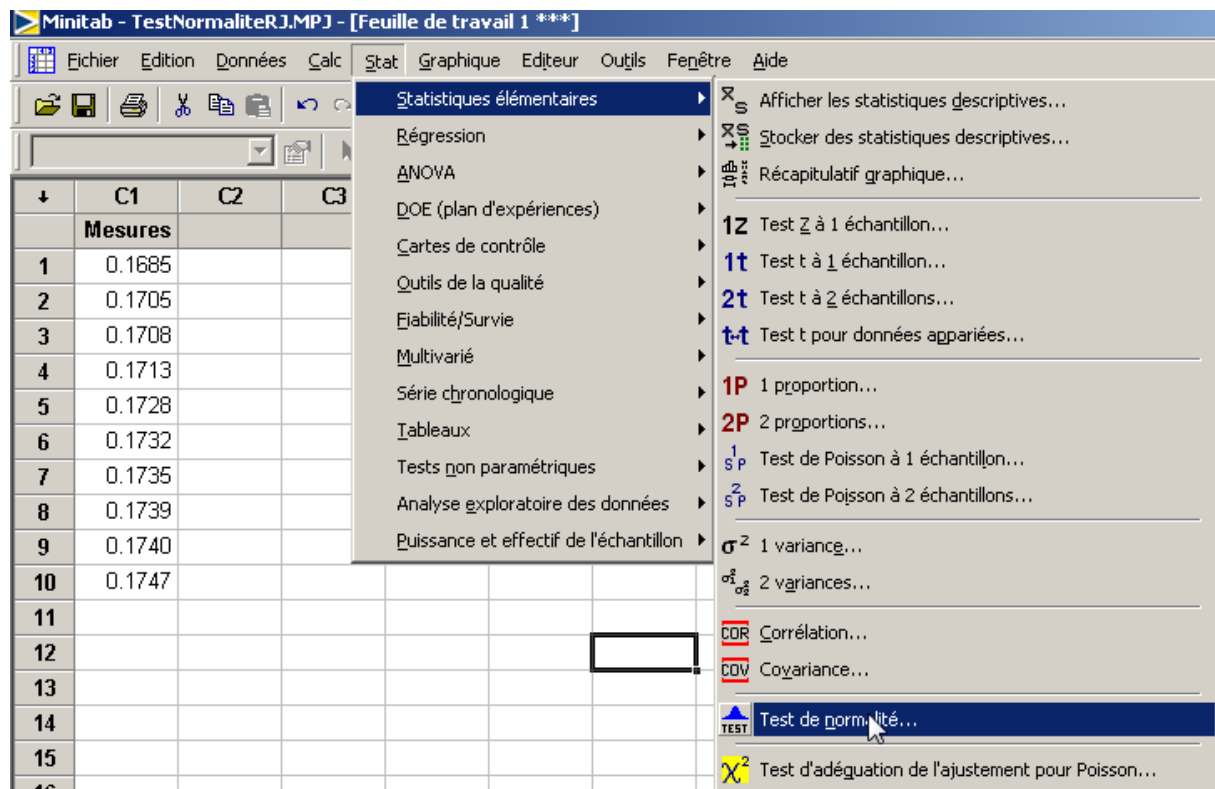
Nous ouvrons donc le fichier *TestNormaliteRJ.mpj* qui contient les 10 mesures ordonnées (triées) d'une cote de pièce (lancement de fabrication) dans un processus de fabrication sous contrôle statistique:



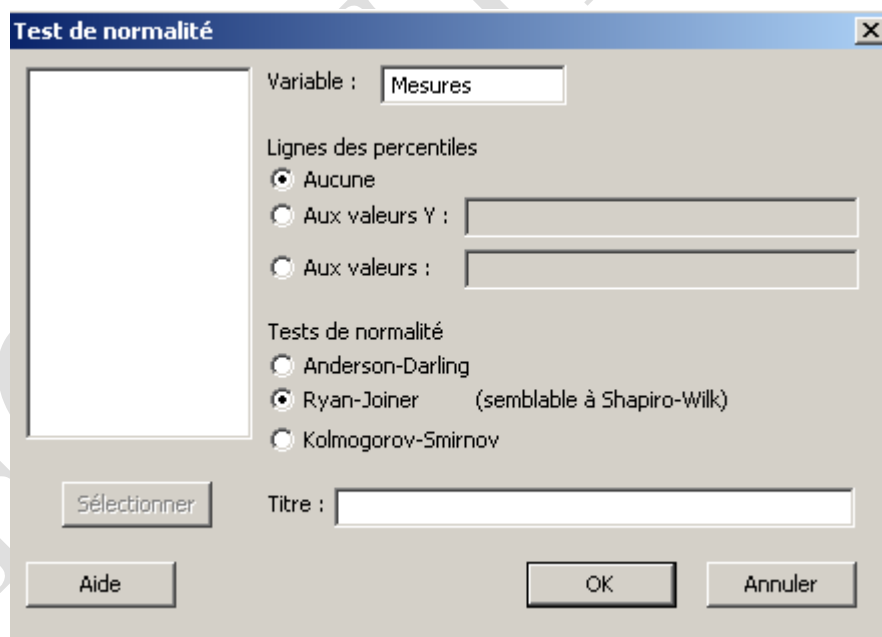
The screenshot shows the Minitab interface with a data table. The table has 10 rows and 8 columns (C1 to C8). The first column (C1) is labeled 'Mesures' and contains the following values: 0.1685, 0.1705, 0.1708, 0.1713, 0.1728, 0.1732, 0.1735, 0.1739, 0.1740, and 0.1747. The other columns (C2 to C8) are empty.

	C1	C2	C3	C4	C5	C6	C7	C8
	<b>Mesures</b>							
1	0.1685							
2	0.1705							
3	0.1708							
4	0.1713							
5	0.1728							
6	0.1732							
7	0.1735							
8	0.1739							
9	0.1740							
10	0.1747							

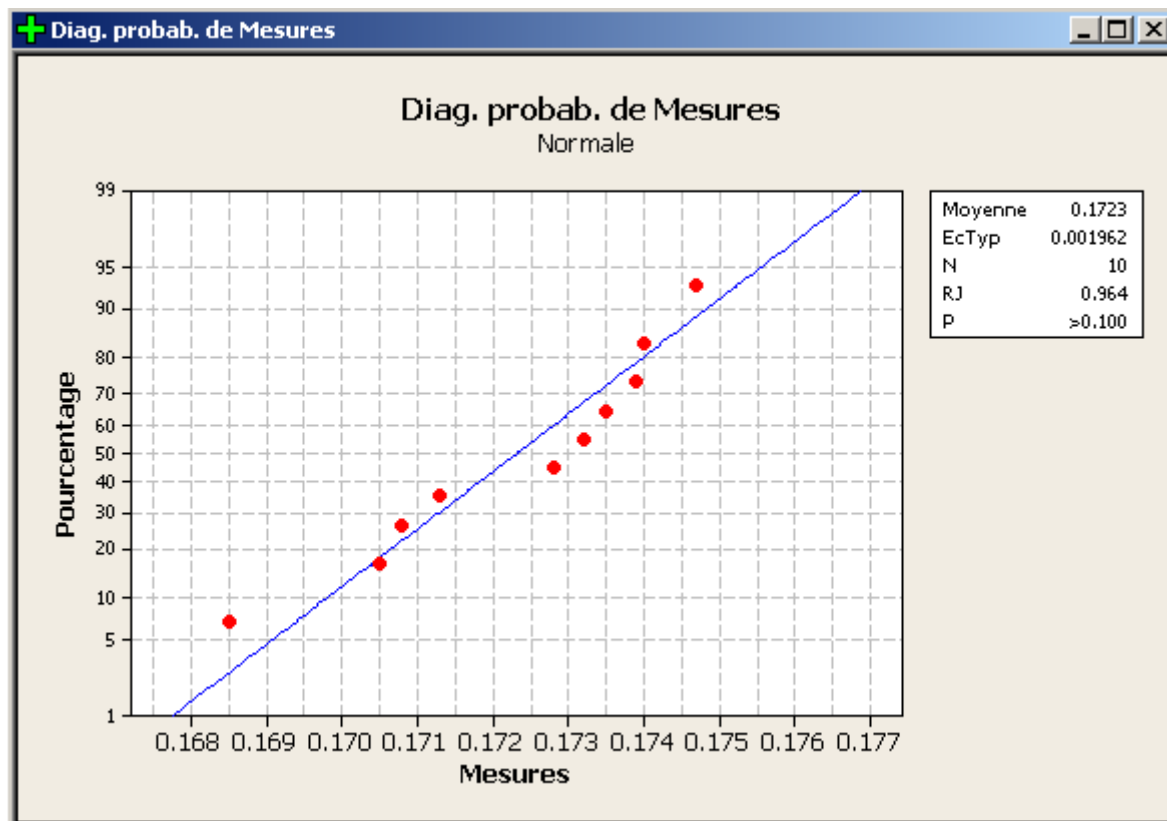
Nous allons dans le menu **Stat/Statistiques élémentaires/Test de Normalité...**:



Nous prenons les paramètres suivants en précisant bien que nous voulons le test de Ryan-Joiner:



et nous validons par **OK** pour obtenir:



Nous obtenons donc exactement le même coefficient  $RJ$  que celui obtenu dans Microsoft Excel à l'aide d'une simulation de Monte-Carlo des coefficients  $b_{i/10}$  du test de Ryan-Joiner (donc Minitab n'utilise par l'approximation du Z-score).

Le carré du coefficient de  $RJ$  doit donner approximativement le coefficient de SW (Shapiro-Wilk). En faisant le carré de la valeur ci-dessus, nous obtenons 0.929296.

Le logiciel R, nous donne:

```
> X2 <- c(0.1685,0.1705,0.1708,0.1713,0.1728,0.1732,0.1735,0.1739,0.1740,0.1747)
> shapiro.test(X2)
```

```
Shapiro-Wilk normality test
```

```
data: X2
W = 0.9246, p-value = 0.3971
```

ce qui est pas trop mal mais la  $p$ -value est cependant très différente! Toutefois la conclusion du test est la même: nous ne rejetons pas l'hypothèse nulle à un niveau de confiance de 95%, la données peuvent être considérée statistiquement comme suivant une loi Normale.

Pour ceux qui veulent une approche ludique de comparaison des trois tests de Normalité:

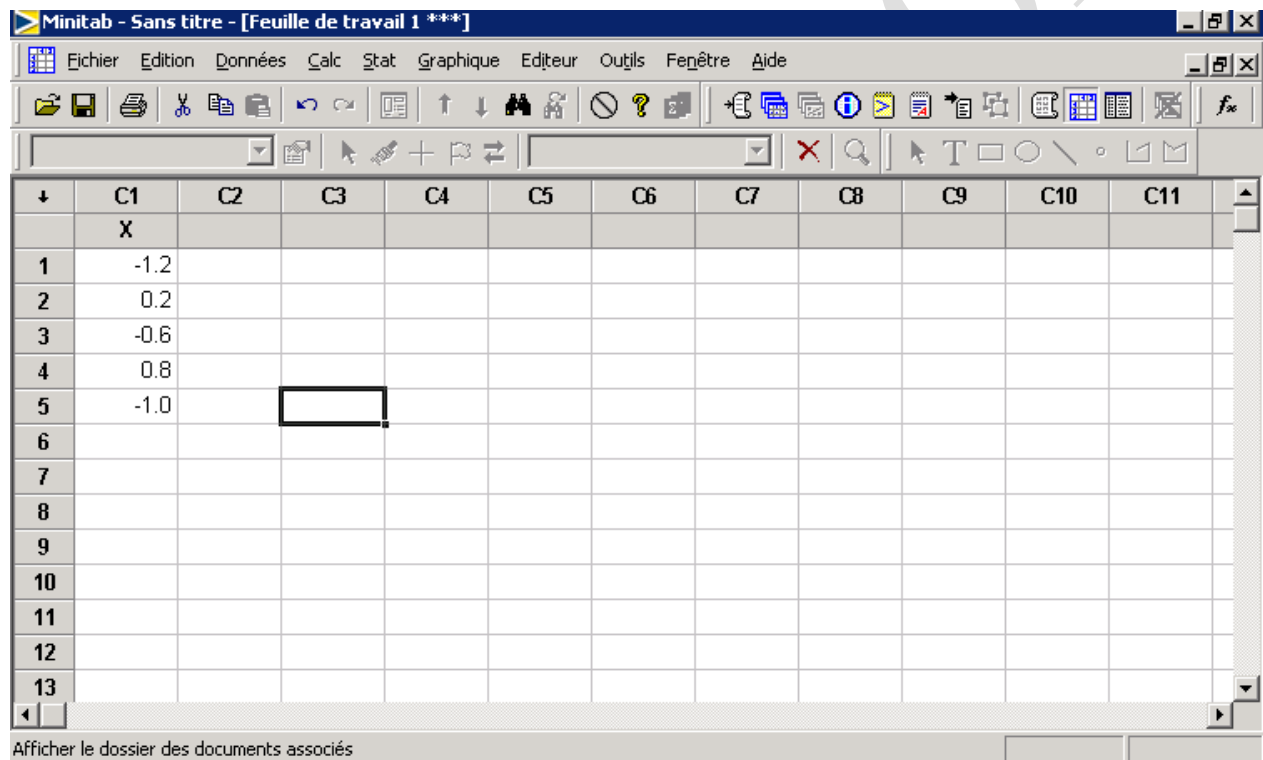
<http://blog.minitab.com/blog/the-statistical-mentor/anderson-darling-ryan-joiner-or-kolmogorov-smirnov-which-normality-test-is-the-best>

## 10.44. Exercice 75.: Test de normalité d'Anderson-Darling (ie Agostino-Stephens)

Minitab® Statistical Software 15.1.1.0

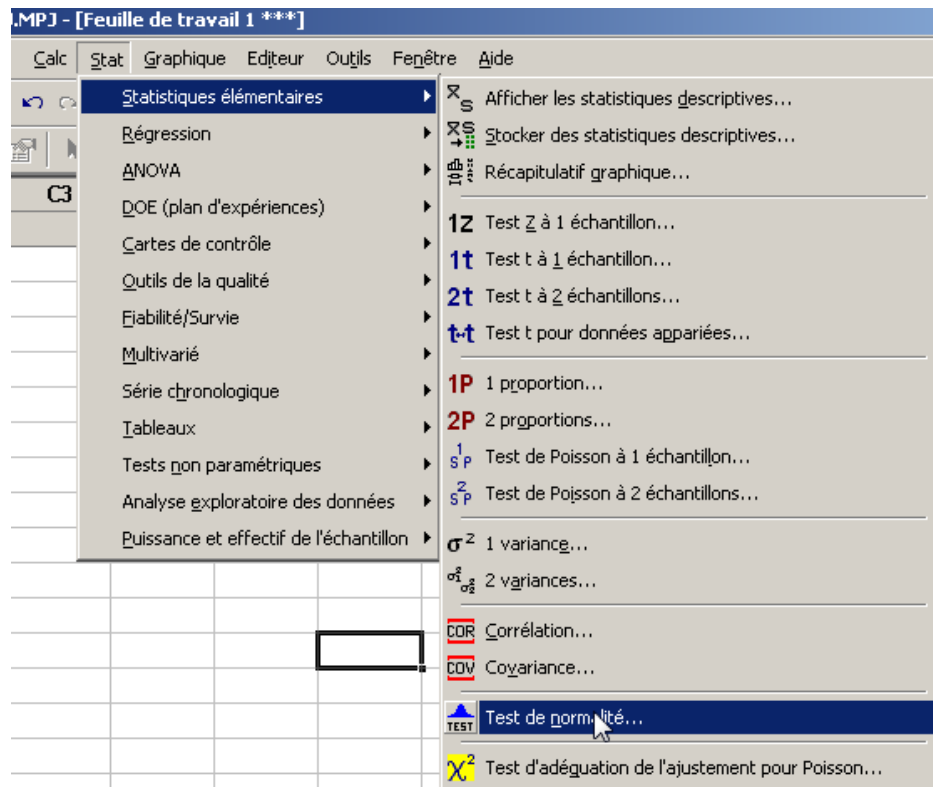
Nous allons ici comme d'habitude vérifier les résultats démontrés dans le cours théorique. Avant de commencer signalons que Minitab permet contrairement à R, Statistica et SPSS de faire le test d'Anderson-Darling pour des échantillons dont le nombre d'individus est compris 10 et 40 (du moins au jour où ces lignes ont été écrites...).

Nous allons donc reprendre les mêmes données que dans le cours théorique (qui ne satisfait donc pas le critère susmentionné mais avait permis de simplifier le nombre de calculs faits à la main):

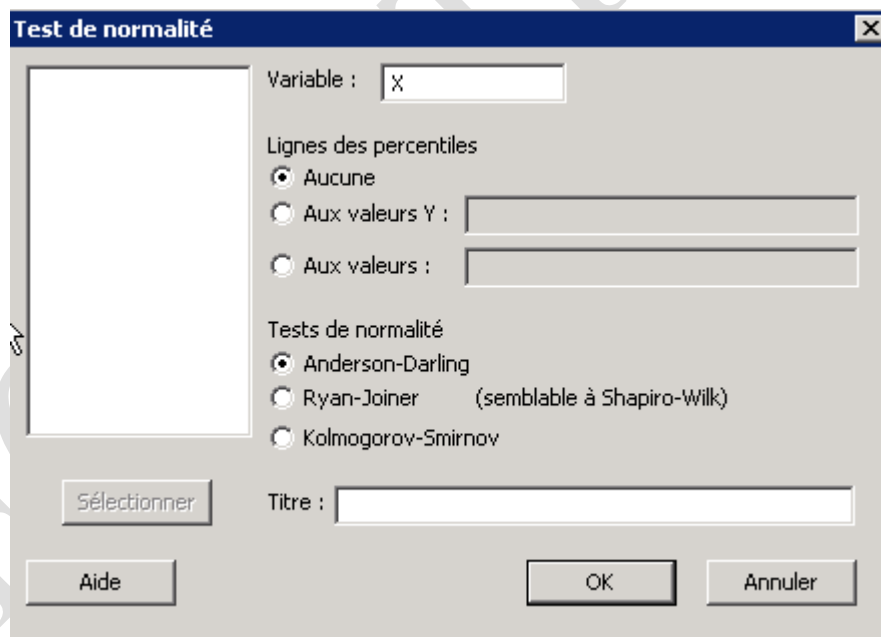


	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
	X										
1	-1.2										
2	0.2										
3	-0.6										
4	0.8										
5	-1.0										
6											
7											
8											
9											
10											
11											
12											
13											

Nous allons dans le menu **Stat/Statistiques élémentaires/Test de Normalité...**:

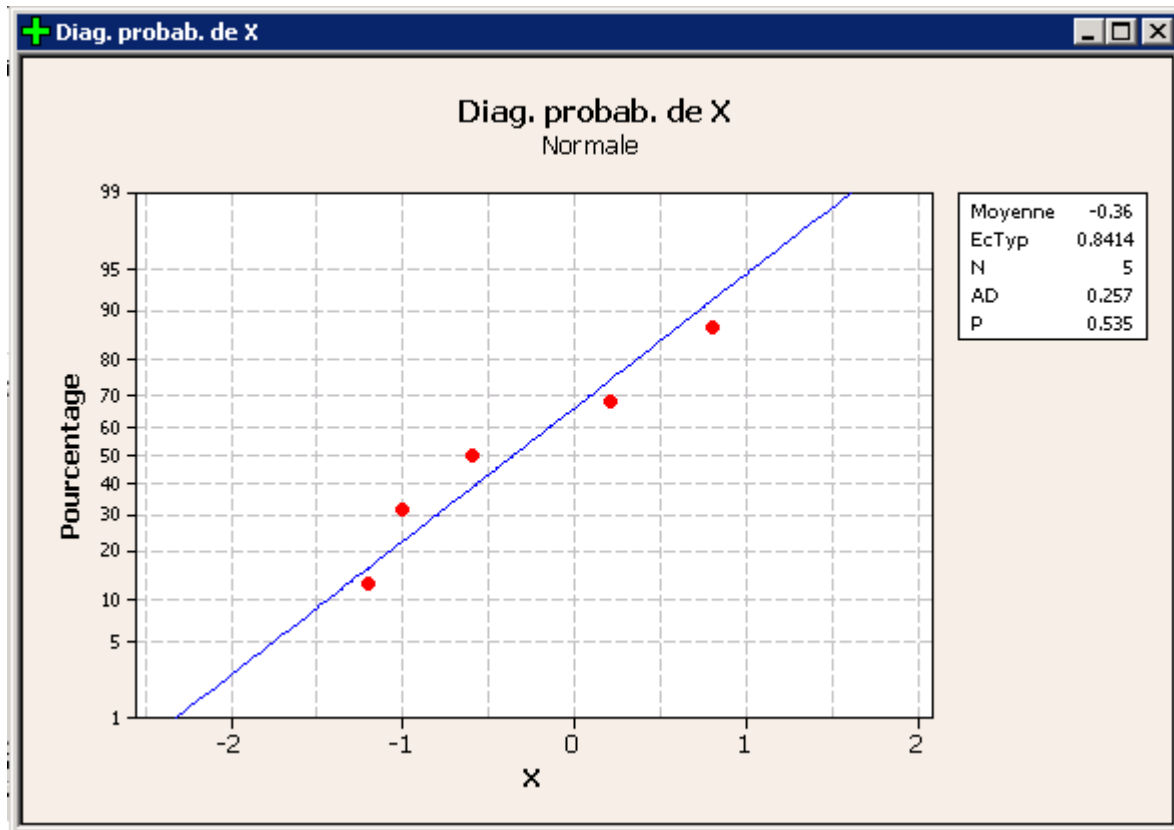


Nous avons alors:



Ce qui donne:





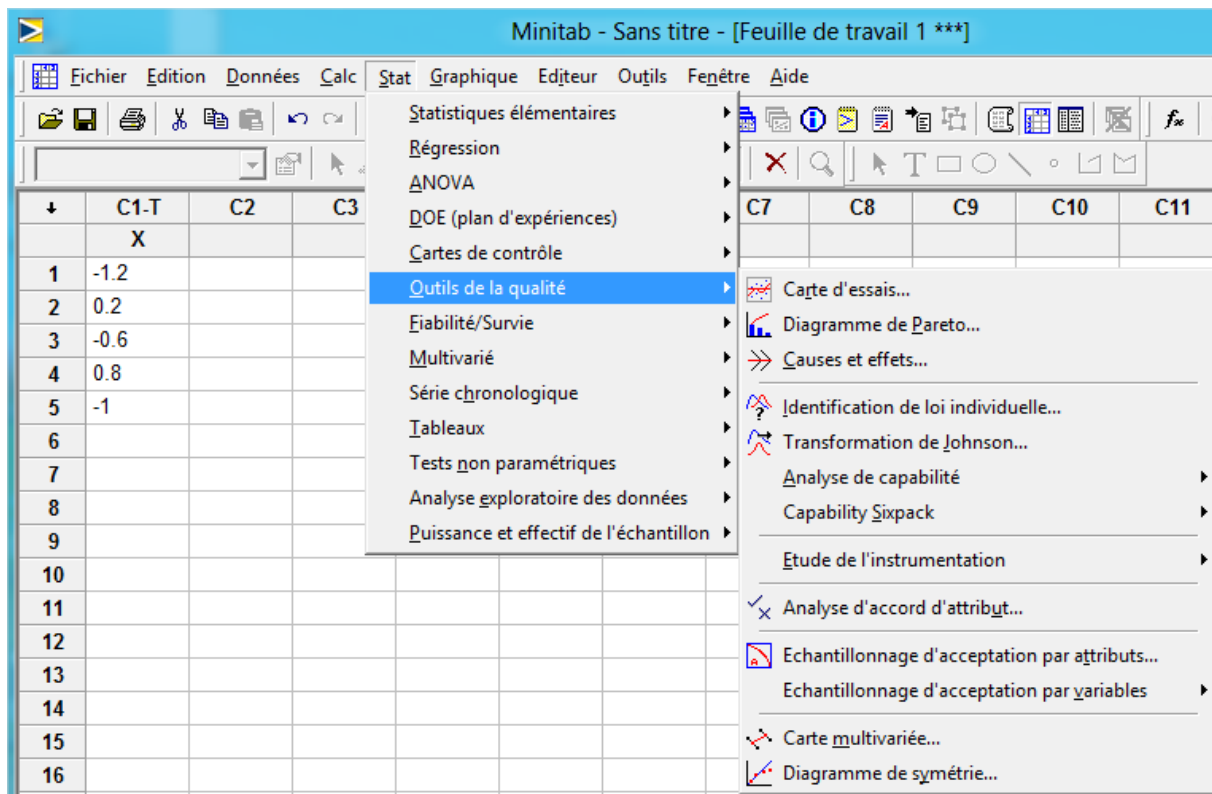
Nous retrouvons alors la même valeur  $AD$  que dans le cours théorique. Nous remarquons que:

1. Minitab fait le test d'équation avec une loi Normale basée sur les estimateurs de la moyenne et de l'écart-type de l'échantillon (dommage qu'on ne puisse choisir comme nous l'avons déjà mentionné pour les autres tests de normalité).
2. Minitab ne donne pas le  $AD^*$  automatiquement mais uniquement quand cela s'avère nécessaire selon les critères de l'équipe de développement.
3. La  $p$ -value est calculée par Minitab sur la base formules empiriques de R.B. D'Agostino et M.A. Stephens. Ce qui fait que la différence est énorme avec celle que nous avons obtenue par Monte-Carlo et qui correspondait à celles calculées par Peter A. W. Lewis.

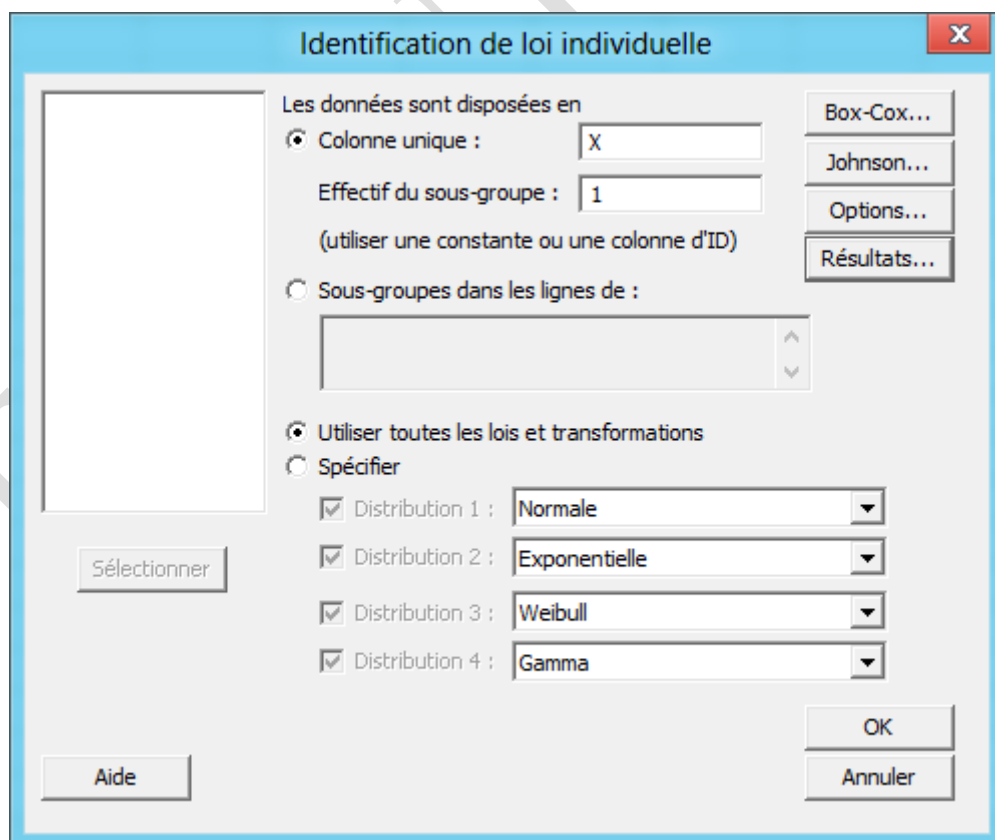
Conclusion:

La  $p$ -value empirique des progiciels de statistiques comme Minitab, SPSS, R et Statistica aura de toute façon moins de faux positifs avec les formules empiriques de R.B. D'Agostino et M.A. Stephens mais elle pourrait avoir bien évidemment en conséquence de plus nombreux faux négatifs donc affaire à suivre...

Nous pouvons cependant aller un peu plus loin. Toujours avec le même set de données allons dans le menu **Stat/Outils de la qualité/Identification de loi individuelle**:



Nous prenons alors le test d'ajustement à d'autres loi (nous laissons celles par défaut même si c'est stupide pour certains loi car leur support n'est pas défini dans les valeurs négatives):



En validant par **OK**, nous obtenons dans la fenêtre de session:

**Identification de la loi de X**

- \* REMARQUE \* Impossible d'ajuster la ou les lois suivantes car les données contiennent des valeurs non positives :  
Exponentielle,  
Log-normale, Weibull, Gamma, Log-logistique
- \* REMARQUE \* Impossible d'effectuer la transformation de Box-Cox car les données contiennent des valeurs non positives.
- \* REMARQUE \* Valeur de P des données initiales > 0,1. Aucune fonction de transformation de Johnson n'est sélectionnée.

## Log-normale à 3 paramètres

- \* ATTENTION \* L'algorithme de Newton-Raphson ne converge pas après 50 itérations.
- \* ATTENTION \* La convergence n'est pas atteinte pour le critère de log de vraisemblance.
- \* ATTENTION \* La convergence n'est pas atteinte pour le critère d'estimation des paramètres.
- \* ATTENTION \* La matrice de variance/covariance des paramètres estimés n'existe pas. Le paramètre de seuil est considéré comme fixe lors du calcul des intervalles de confiance.

## Exponentielle 2 paramètres

- \* ATTENTION \* La matrice de variance/covariance des paramètres estimés n'existe pas. Le paramètre de seuil est considéré comme fixe lors du calcul des intervalles de confiance.

## Weibull 3 paramètres

- \* ATTENTION \* L'algorithme de Newton-Raphson ne converge pas après 50 itérations.
- \* ATTENTION \* La convergence n'est pas atteinte pour le critère d'estimation des paramètres.
- \* ATTENTION \* La matrice de variance/covariance des paramètres estimés n'existe pas. Le paramètre de seuil est considéré comme fixe lors du calcul des intervalles de confiance.

## Gamma 3 paramètres

- \* ATTENTION \* L'algorithme de Newton-Raphson ne converge pas après 50 itérations.
- \* ATTENTION \* La convergence n'est pas atteinte pour le critère de log de vraisemblance.
- \* ATTENTION \* La convergence n'est pas atteinte pour le critère d'estimation des paramètres.
- \* ATTENTION \* La matrice de variance/covariance des paramètres estimés n'existe pas. Le paramètre de seuil est considéré comme fixe lors du

## Log-logistique 3 paramètres

- \* ATTENTION \* L'algorithme de Newton-Raphson ne converge pas après 50 itérations.
- \* ATTENTION \* La convergence n'est pas atteinte pour le critère d'estimation des paramètres.
- \* ATTENTION \* La matrice de variance/covariance des paramètres estimés n'existe pas. Le paramètre de seuil est considéré comme fixe lors du calcul des intervalles de confiance.

## Statistiques descriptives

N	N*	Moyenne	EcTyp	Médiane	Minimum	Maximum	Asymétrie	Aplatissement
5	0	-0,36	0,841427	-0,6	-1,2	0,8	0,606986	-1,57139

## Test d'adéquation de l'ajustement

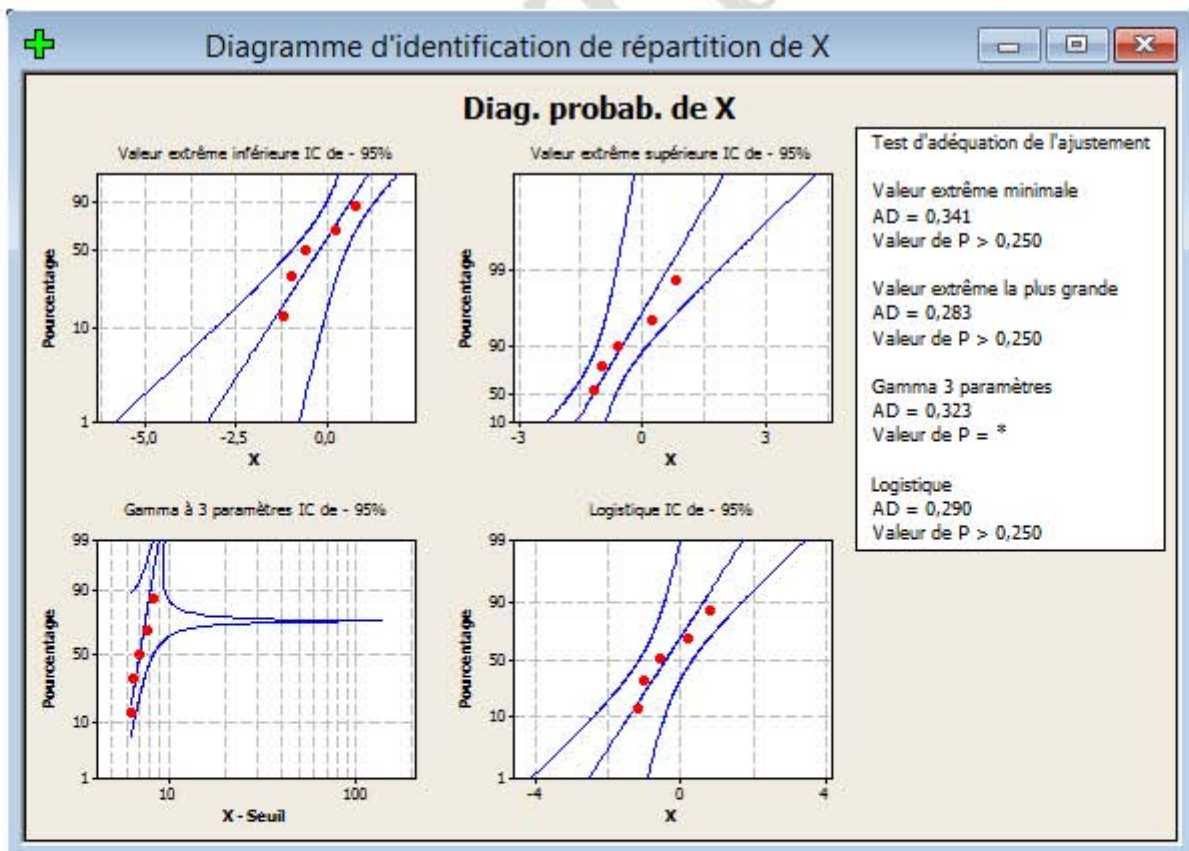
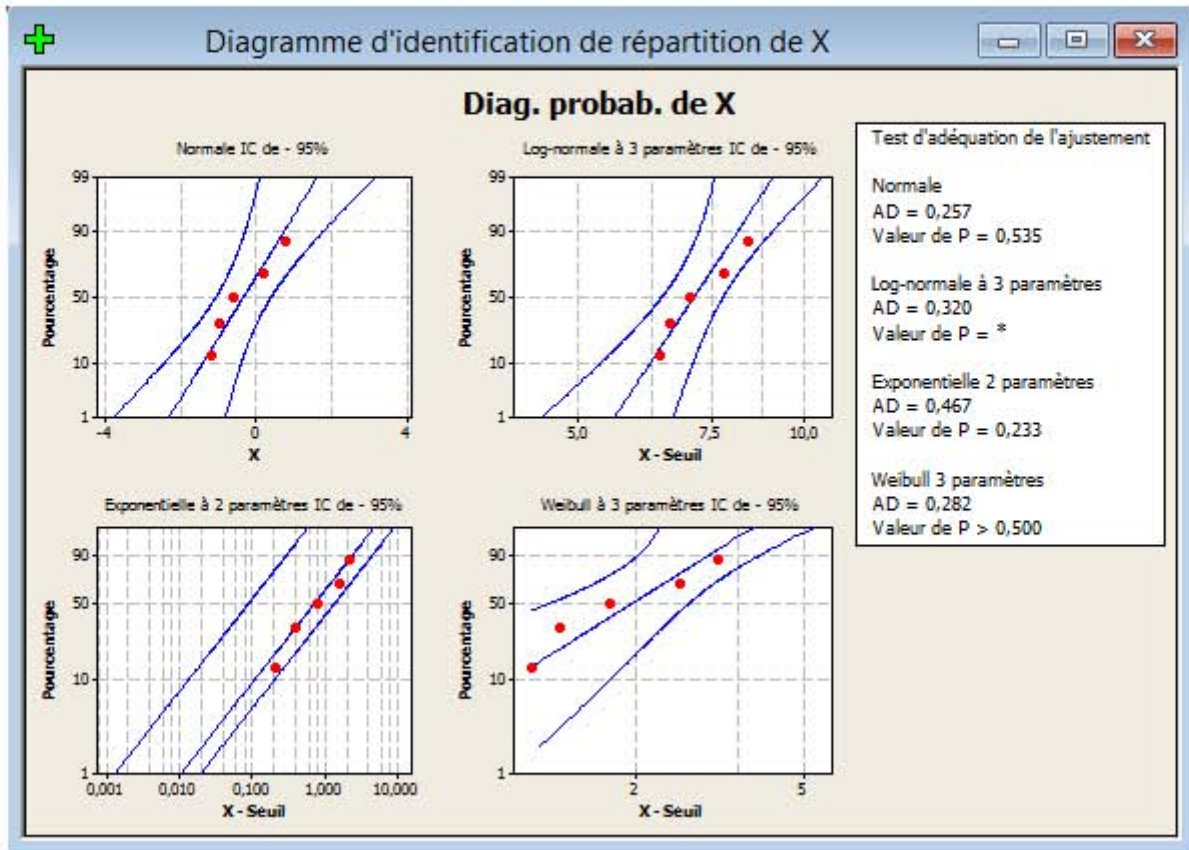
Loi de distribution	AD	P
Normale	0,257	0,535
Log-normale à 3 paramètres	0,320	*
Exponentielle 2 paramètres	0,467	0,233
Weibull 3 paramètres	0,282	>0,500
Valeur extrême minimale	0,341	>0,250
Valeur extrême la plus grande	0,283	>0,250
Gamma 3 paramètres	0,323	*
Logistique	0,290	>0,250
Log-logistique 3 paramètres	0,251	*

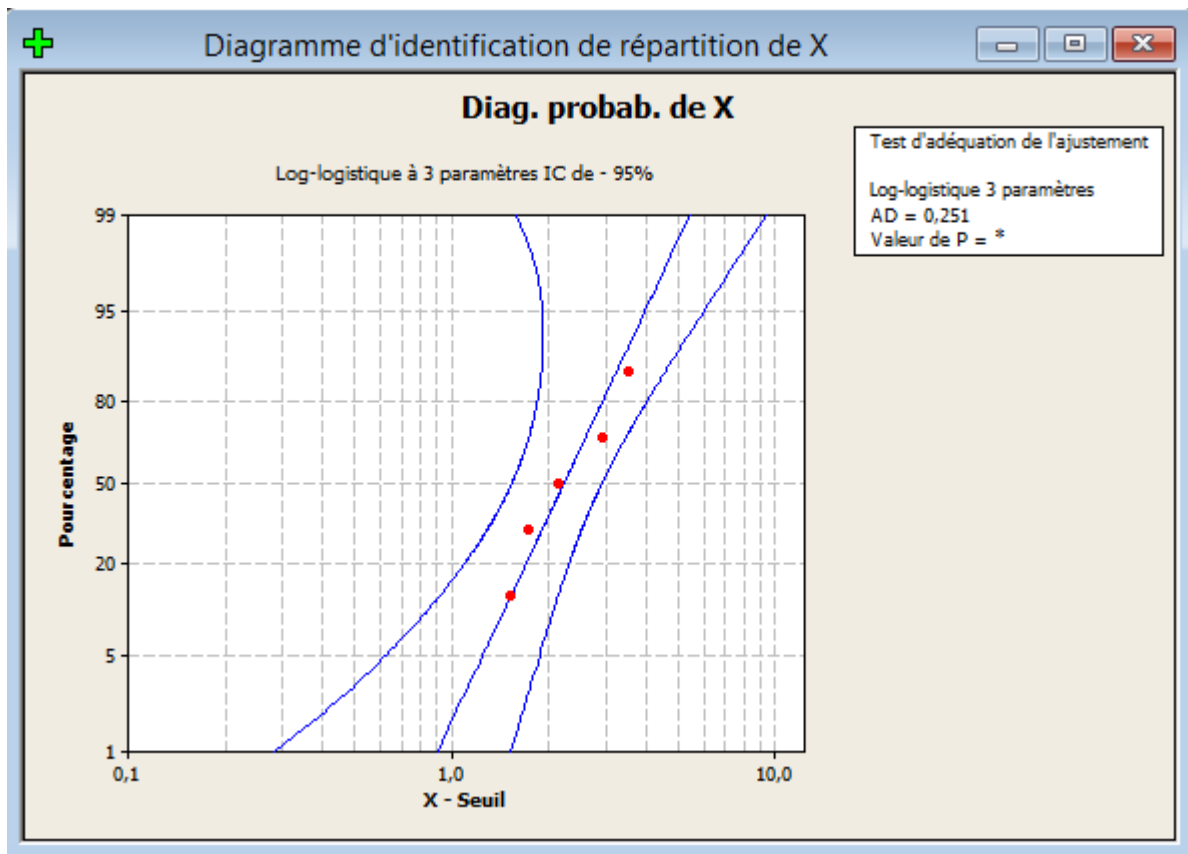
## Estimations de MaxV des paramètres de distribution

Loi de distribution	Emplacement	Forme	Echelle	Seuil
Normal*	-0,36000		0,84143	
Log-normale à 3 paramètres	1,95846		0,10398	-7,61003
Exponentielle 2 paramètres			1,05000	-1,41000
Weibull 3 paramètres		2,88052	2,22423	-2,33402
Valeur extrême minimale	0,02858		0,72830	
Valeur extrême la plus grande	-0,71681		0,58774	
Gamma 3 paramètres		89,13723	0,07882	-7,50580
Logistique	-0,41516		0,46278	
Log-logistique 3 paramètres	0,79931		0,19559	-2,72950

\* Echelle : estimation ajustée de MaxV

Avec les graphiques suivants:





Pour ceux qui veulent une approche ludique de comparaison des trois tests de Normalité:

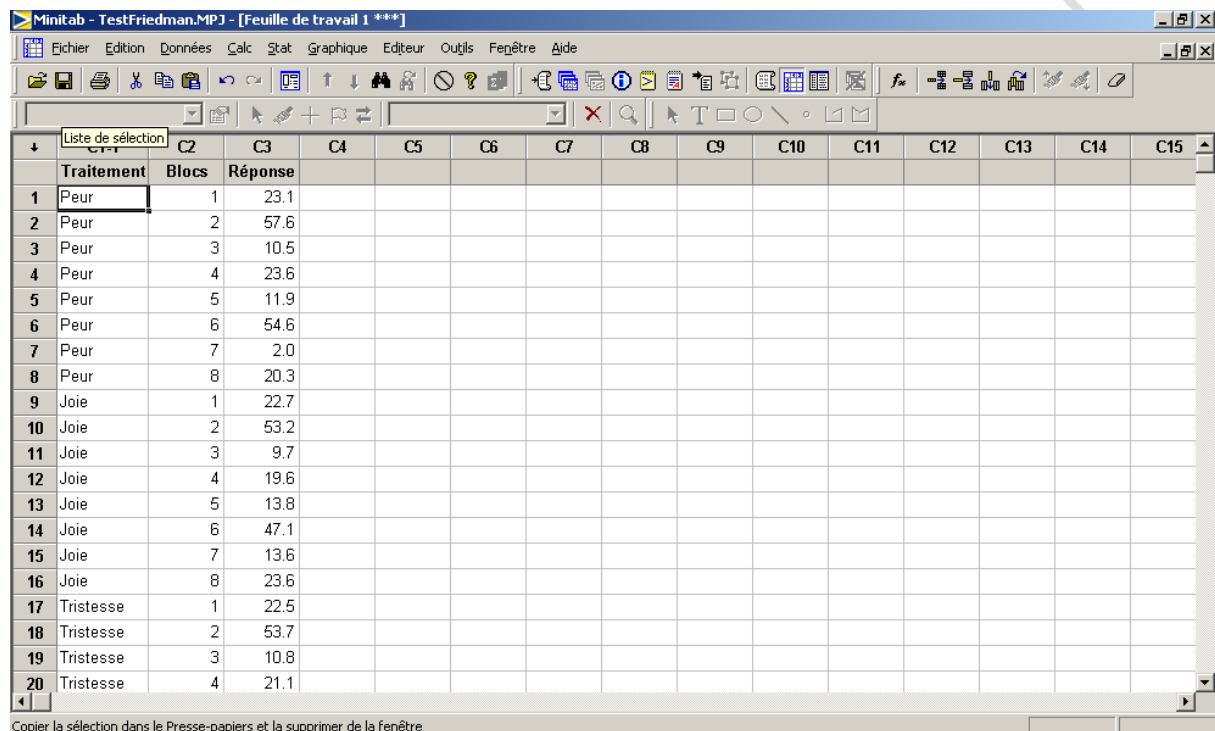
<http://blog.minitab.com/blog/the-statistical-mentor/anderson-darling-ryan-joiner-or-kolmogorov-smirnov-which-normality-test-is-the-best>

# 11. Statistiques non paramétriques

## 11.1. Exercice 76.: Test de Friedman

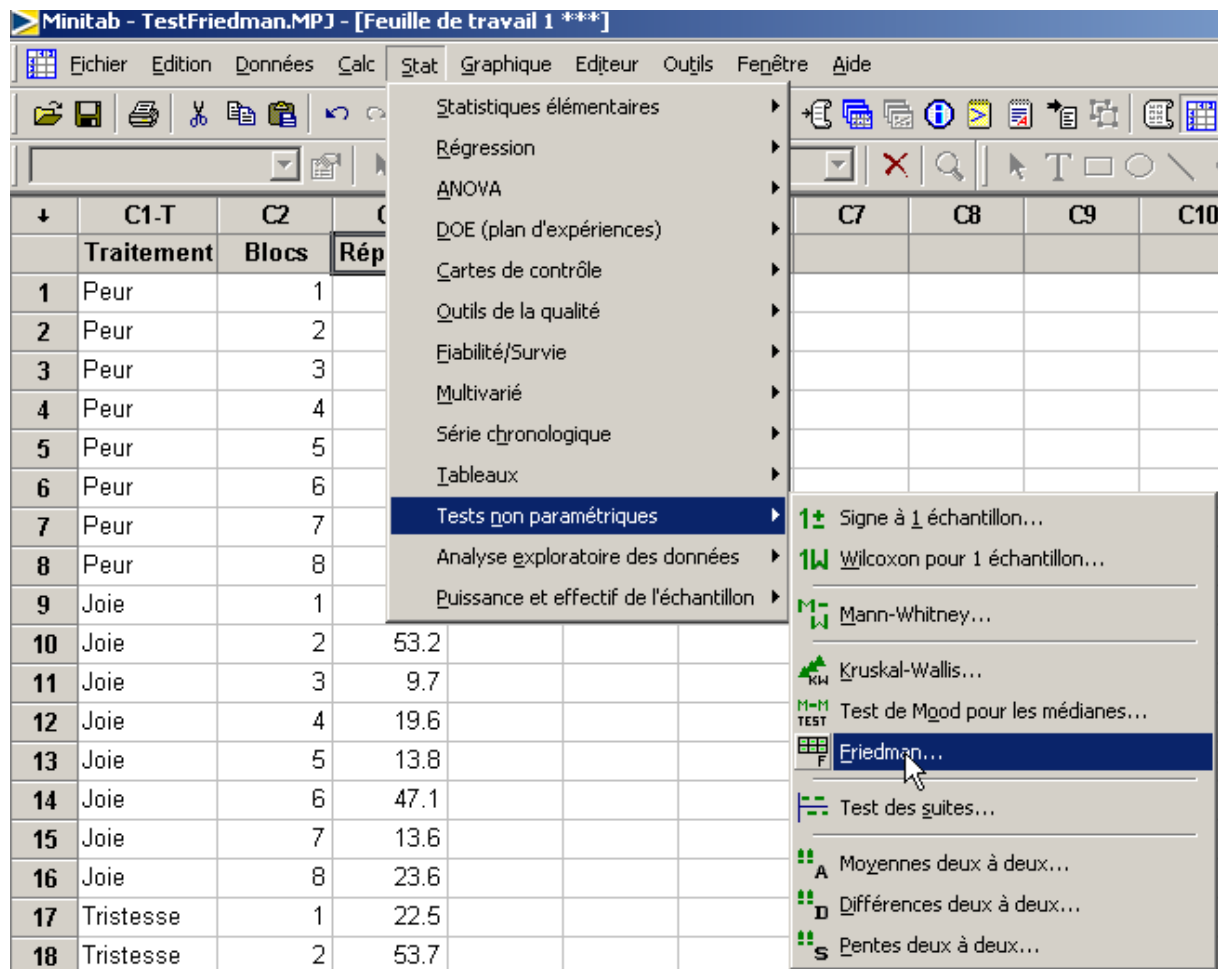
Minitab® Statistical Software 15.1.1.0

Nous allons nous baser sur l'exemple pratique fait aussi dans le cours théorique pour appliquer la pseudo-démonstration (guère convaincante) que nous avons étudiée concernant ce test.



	Liste de sélection	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
	Traitement	Blocs	Réponse												
1	Peur	1	23.1												
2	Peur	2	57.6												
3	Peur	3	10.5												
4	Peur	4	23.6												
5	Peur	5	11.9												
6	Peur	6	54.6												
7	Peur	7	2.0												
8	Peur	8	20.3												
9	Joie	1	22.7												
10	Joie	2	53.2												
11	Joie	3	9.7												
12	Joie	4	19.6												
13	Joie	5	13.8												
14	Joie	6	47.1												
15	Joie	7	13.6												
16	Joie	8	23.6												
17	Tristesse	1	22.5												
18	Tristesse	2	53.7												
19	Tristesse	3	10.8												
20	Tristesse	4	21.1												

Nous allons alors dans le menu **Stat/Tests non paramétriques/Friedman...**:



Nous avons alors:



ce qui donne:



**Test de Friedman : Réponse en fonction de Traitement bloqué par Blocs**

S = 6.45 DL = 3 P = 0.092

Traitement	N	Médiane est.	Somme des rangs
Calme	8	18.506	14.0
Joie	8	19.019	20.0
Peur	8	21.781	27.0
Tristesse	8	18.869	19.0

Grande médiane = 19.544

Nous retrouvons donc exactement les mêmes résultats que dans le cours théorique.

## 11.2. Exercice 77.: Test du signe binomial (dixit: test de la médiane)

Minitab® Statistical Software 15.1.1.0

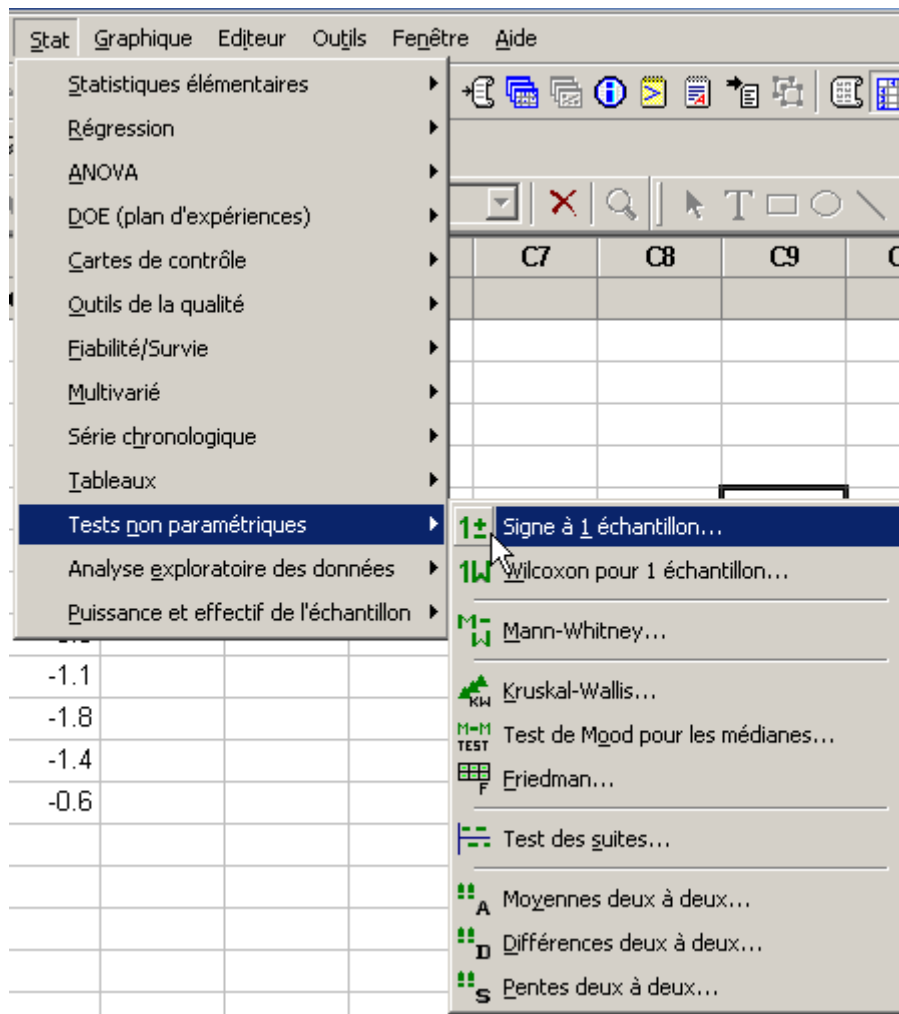
Nous avons effectué deux séries de mesures avec deux méthodes différentes.

Nous souhaiterions savoir la différence est significative ou pas. Le but étant aussi de contrôler si nous avons un résultat différent de celui calculé à la main dans le cours théorique et dans Microsoft Excel.

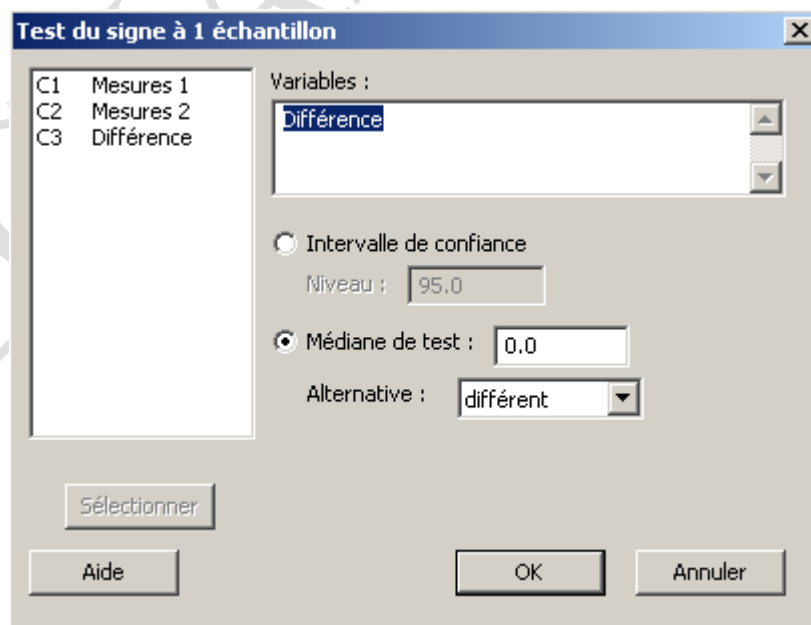
Pour cela, ouvrez le fichier *TestDuSigne.mpj*:

↓	C1	C2	C3
	Mesures 1	Mesures 2	Différence
1	20.4	20.7	-0.3
2	25.4	26.3	-0.9
3	25.6	26.8	-1.2
4	25.6	28.1	-2.5
5	26.6	26.2	0.4
6	28.6	27.3	1.3
7	28.7	29.5	-0.8
8	29.0	32.0	-3.0
9	29.8	30.9	-1.1
10	30.5	32.3	-1.8
11	30.9	32.3	-1.4
12	31.1	31.7	-0.6

et allons dans le menu **Stat/Tests non paramétriques/Signe à 1 échantillon....**:



où nous prenons comme vu dans le cours théorique:



Nous validons pour obtenir dans la fenêtre de session:

The screenshot displays two windows from the Minitab software. The top window, titled 'Session', shows the results of a sign test for the median. The text indicates a test of the median against a difference of 0.00000, with a p-value of 0.0386. Below this, a table summarizes the counts for values above, below, and equal to the median.

	N	Au-dessous	Egal	Au-dessus	P	Médiane
Différence	12	10	0	2	0.0386	-1.000

The bottom window, titled 'feuille de travail 1', shows a data table with 6 rows and 14 columns. The first three columns are labeled 'Mesures 1', 'Mesures 2', and 'Différence'. The data points are as follows:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Mesures 1	Mesures 2	Différence											
1	20.4	20.7	-0.3											
2	25.4	26.3	-0.9											
3	25.6	26.8	-1.2											
4	25.6	28.1	-2.5											
5	26.6	26.2	0.4											
6	28.6	27.3	1.3											

Nous retrouvons toutes les valeurs vues dans le cours de statistique théorique. Nous rejetons donc au vu de la  $p$ -value (inférieure à 5%) que la différence n'est pas significative.

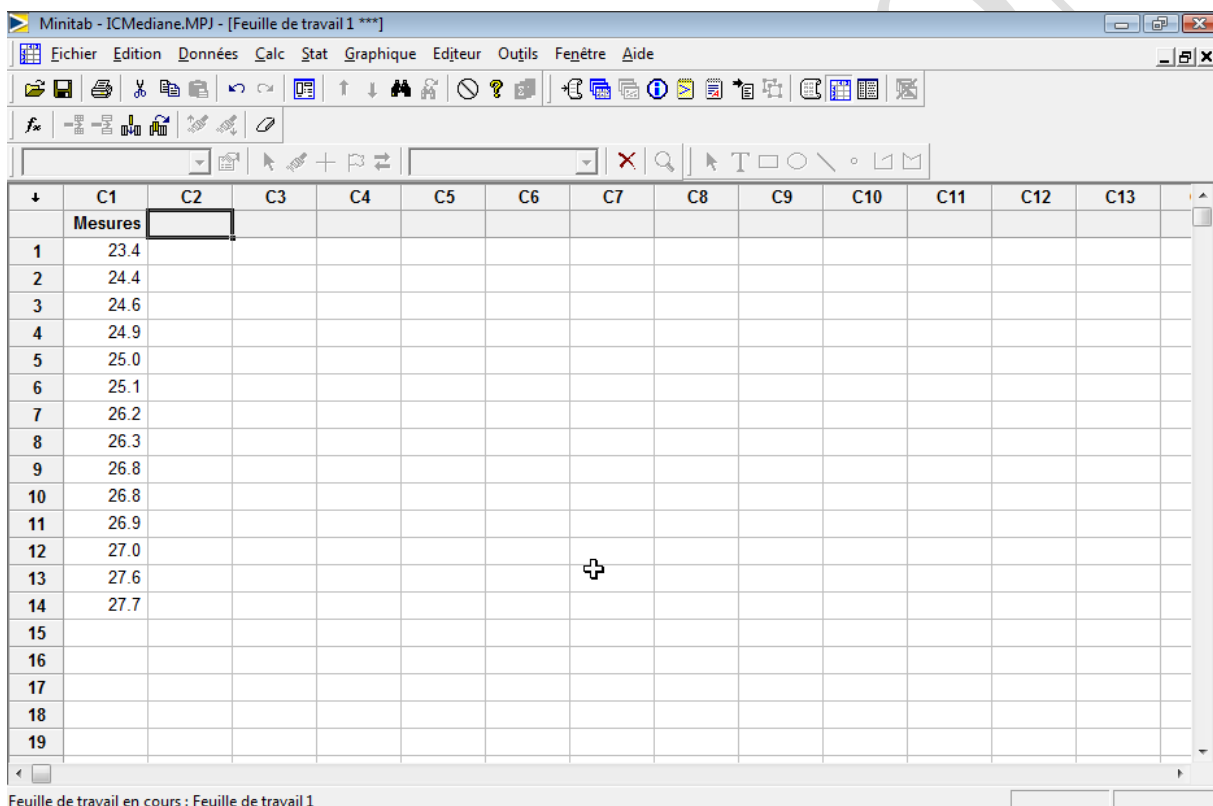
### 11.3. Exercice 78.: Intervalle de confiance de la médiane (via test du signe)

Minitab® Statistical Software 15.1.1.0

Nous avons effectué une série de mesures de valeurs continues.

Nous souhaiterions simplement connaître l'intervalle de confiance de la médiane, le but étant aussi de contrôler que Minitab® Statistical Software utilise bien la méthode que nous avons mentionnée dans le cours de statistique théorique.

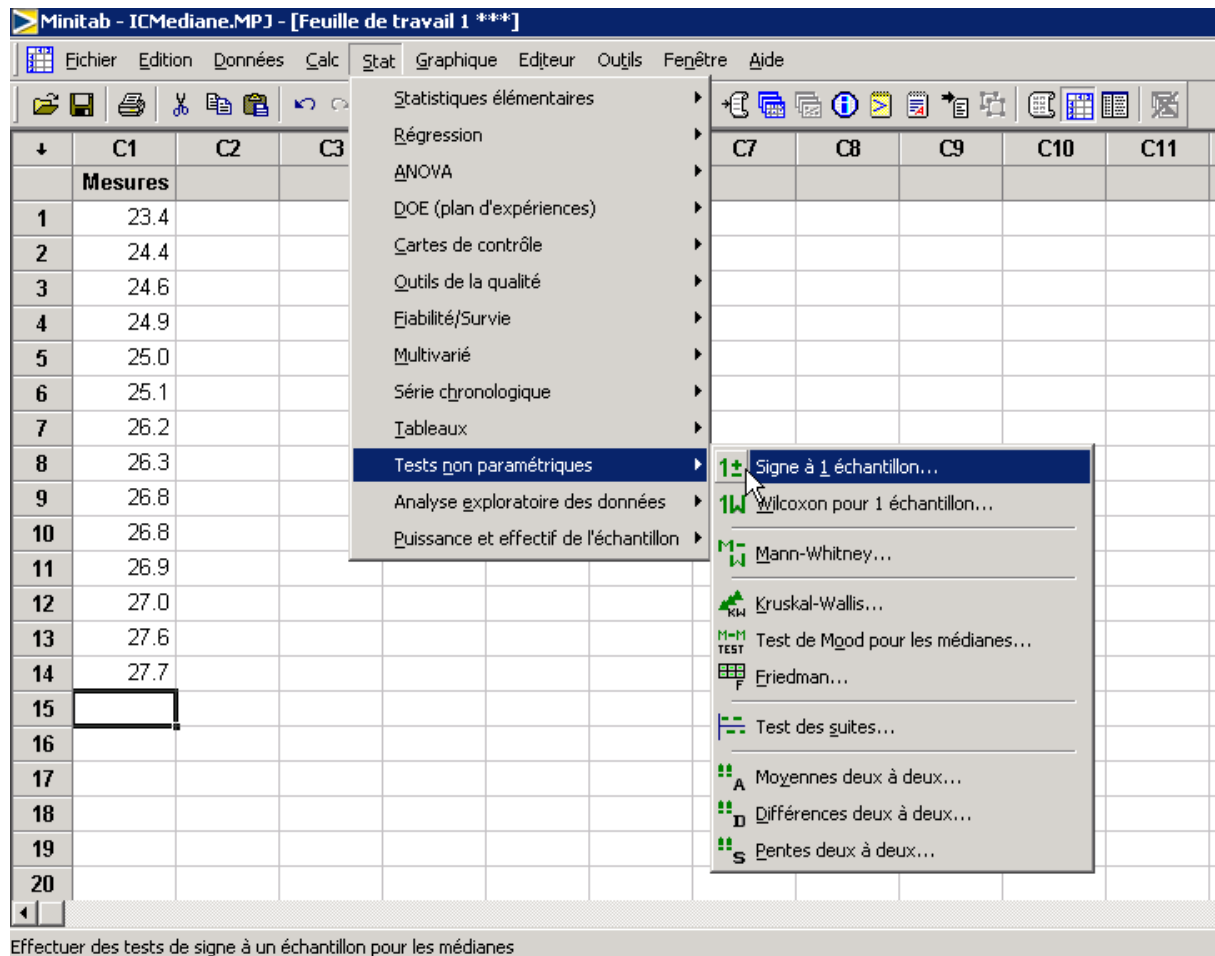
Pour cela, ouvrez le fichier *ICMediane.mpj*:



The screenshot shows the Minitab Statistical Software interface. The title bar reads "Minitab - ICMediane.MPJ - [Feuille de travail 1 \*\*\*]". The menu bar includes "Fichier", "Edition", "Données", "Calc", "Stat", "Graphique", "Editeur", "Outils", "Fenêtre", and "Aide". The toolbar contains various icons for file operations, editing, and statistical analysis. The main window displays a data table with 14 rows of measurements in column C1. The column headers are C1 through C13. The data in C1 is as follows:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Mesures												
1	23.4												
2	24.4												
3	24.6												
4	24.9												
5	25.0												
6	25.1												
7	26.2												
8	26.3												
9	26.8												
10	26.8												
11	26.9												
12	27.0												
13	27.6												
14	27.7												
15													
16													
17													
18													
19													

et allons dans le menu **Stat/Tests non paramétriques/Signe à 1 échantillon....**:

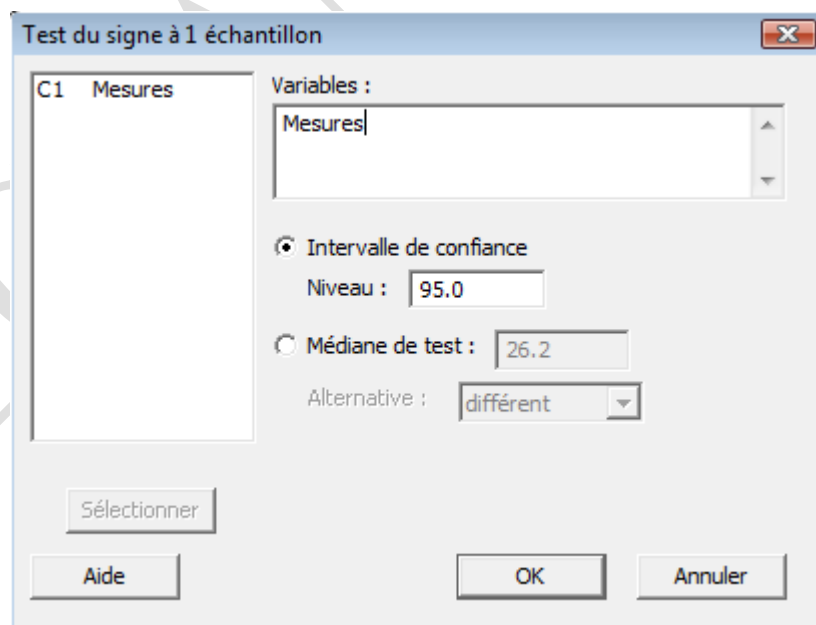


The screenshot shows the Minitab interface with the 'Stat' menu open. The 'Tests non paramétriques' option is selected, and a sub-menu is displayed with 'Signe à 1 échantillon...' as the first option. The background shows a worksheet with a column 'Mesures' containing data from row 1 to 14.

	C1	C2	C3
	Mesures		
1	23.4		
2	24.4		
3	24.6		
4	24.9		
5	25.0		
6	25.1		
7	26.2		
8	26.3		
9	26.8		
10	26.8		
11	26.9		
12	27.0		
13	27.6		
14	27.7		
15			
16			
17			
18			
19			
20			

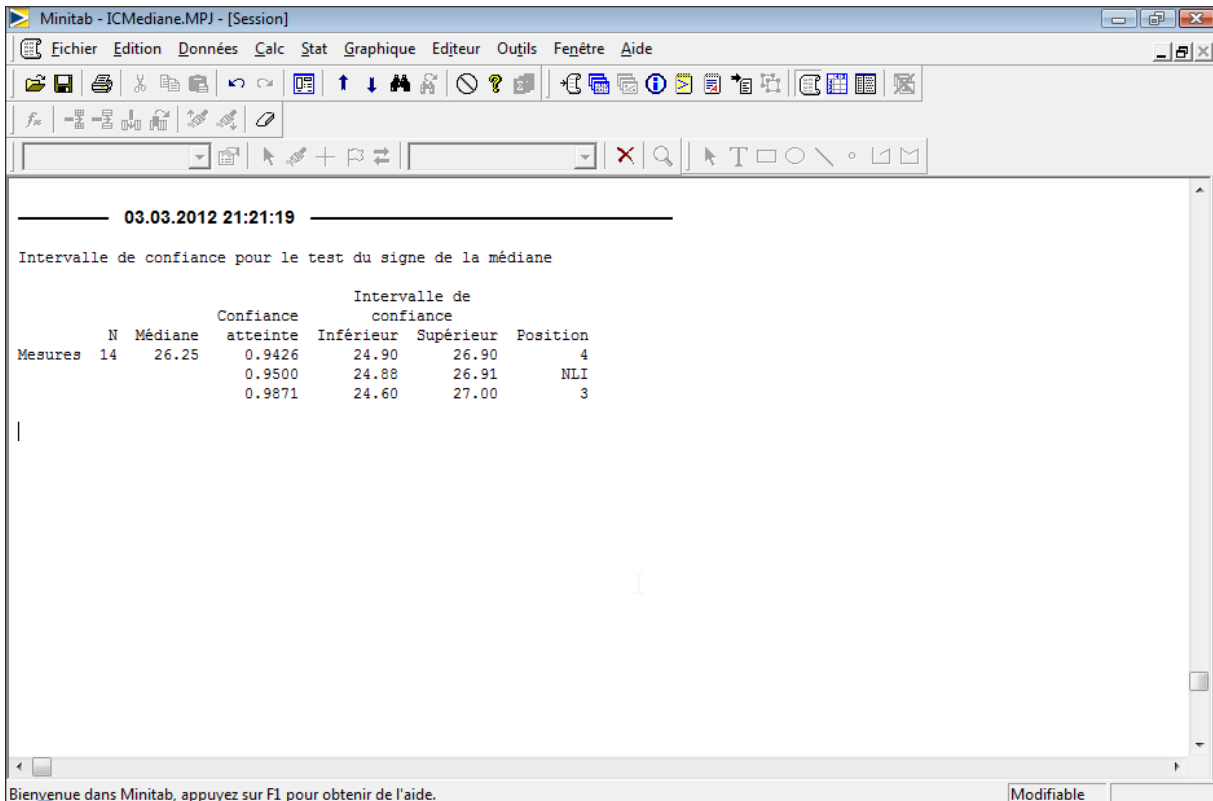
Effectuer des tests de signe à un échantillon pour les médianes

et nous avons alors:



The dialog box 'Test du signe à 1 échantillon' is shown. The variable 'Mesures' is selected in the 'Variables' field. The 'Intervalle de confiance' radio button is selected, with a 'Niveau' of 95.0. The 'Médiane de test' radio button is unselected, with a value of 26.2. The 'Alternative' dropdown is set to 'différent'. Buttons for 'Sélectionner', 'Aide', 'OK', and 'Annuler' are visible.

Nous validons par **OK**:



Nous retrouvons l'intervalle de 94.26% qui était le plus proche que nous pouvions calculer avec Microsoft Excel (voir un peu plus bas).

Nous voyons que Minitab® Statistical Software fait une interpolation non linéaire (NLI) pour déterminer la valeur de l'intervalle souhaité à 95%.

En réalité, Minitab® Statistical Software donne toujours une valeur au-dessous et une au-dessus de la valeur cible souhaitée à cause du comportement discret de la loi binomiale.

A titre de comparaison, nous avons donc avec Microsoft Excel:

	A	B	C	D	E	F	G	H	I
1	Valeurs	Prob. Cumul.	Numéro						
2	23.40	0.09%	1						
3	24.40	0.65%	2		94.26%				
4	24.60	2.87%	3		98.71%				
5	24.90	8.98%	4		26.25				
6	25.00	21.20%	5						
7	25.10	39.53%	6						
8	26.20	60.47%	7						
9	26.30	78.80%	8						
10	26.80	91.02%	9						
11	26.80	97.13%	10						
12	26.90	99.35%	11						
13	27.00	99.91%	12						
14	27.60	99.99%	13						
15	27.70	100.00%	14						

Mesures_1	N	Médiane	Confiance atteinte	Intervalle de confiance		Position
				Inférieur	Supérieur	
14	14	26.25	0.9426	24.90	26.90	4
			0.9500	24.88	26.91	NLI
			0.9871	24.60	27.00	3

soit explicitement:

	A	B	C	D	E	F	G
1	Valeurs	Prob. Cumul.	Numéro				
2	23.4	=LOI BINOMIALE.N(C2;NB(A:A);0.5;1)	1				
3	24.4	=LOI BINOMIALE.N(C3;NB(A:A);0.5;1)	2		=100%-B4-(1-B11)		
4	24.6	=LOI BINOMIALE.N(C4;NB(A:A);0.5;1)	3		=100%-B3-(1-B12)		
5	24.9	=LOI BINOMIALE.N(C5;NB(A:A);0.5;1)	4		=MEDIANE(A2:A15)		
6	25	=LOI BINOMIALE.N(C6;NB(A:A);0.5;1)	5				
7	25.1	=LOI BINOMIALE.N(C7;NB(A:A);0.5;1)	6				
8	26.2	=LOI BINOMIALE.N(C8;NB(A:A);0.5;1)	7				
9	26.3	=LOI BINOMIALE.N(C9;NB(A:A);0.5;1)	8				
10	26.8	=LOI BINOMIALE.N(C10;NB(A:A);0.5;1)	9				
11	26.8	=LOI BINOMIALE.N(C11;NB(A:A);0.5;1)	10				
12	26.9	=LOI BINOMIALE.N(C12;NB(A:A);0.5;1)	11				
13	27	=LOI BINOMIALE.N(C13;NB(A:A);0.5;1)	12				
14	27.6	=LOI BINOMIALE.N(C14;NB(A:A);0.5;1)	13				
15	27.7	=LOI BINOMIALE.N(C15;NB(A:A);0.5;1)	14				

		N		Médiane		Confiance		Intervalle de		Position	
		Mesures_1		14		atteinte		confiance			
				26.25				Inférieur		Supérieur	
						0.9426		24.90		26.90	4
						0.9500		24.88		26.91	NLI
						0.9871		24.60		27.00	3

Nous voyons donc bien que comme mentionné dans le cours théorique, Minitab® Statistical Software utilise bien la loi binomiale. Nous avons aussi mentionné dans le cours théorique qu'il était préférable peut-être d'utiliser la technique du bootstrapping qui donne à 95% un intervalle de 24.90 à 26.90 (valeurs que Minitab® Statistical Software affecte logiquement aux 94.26% de la loi binomiale).

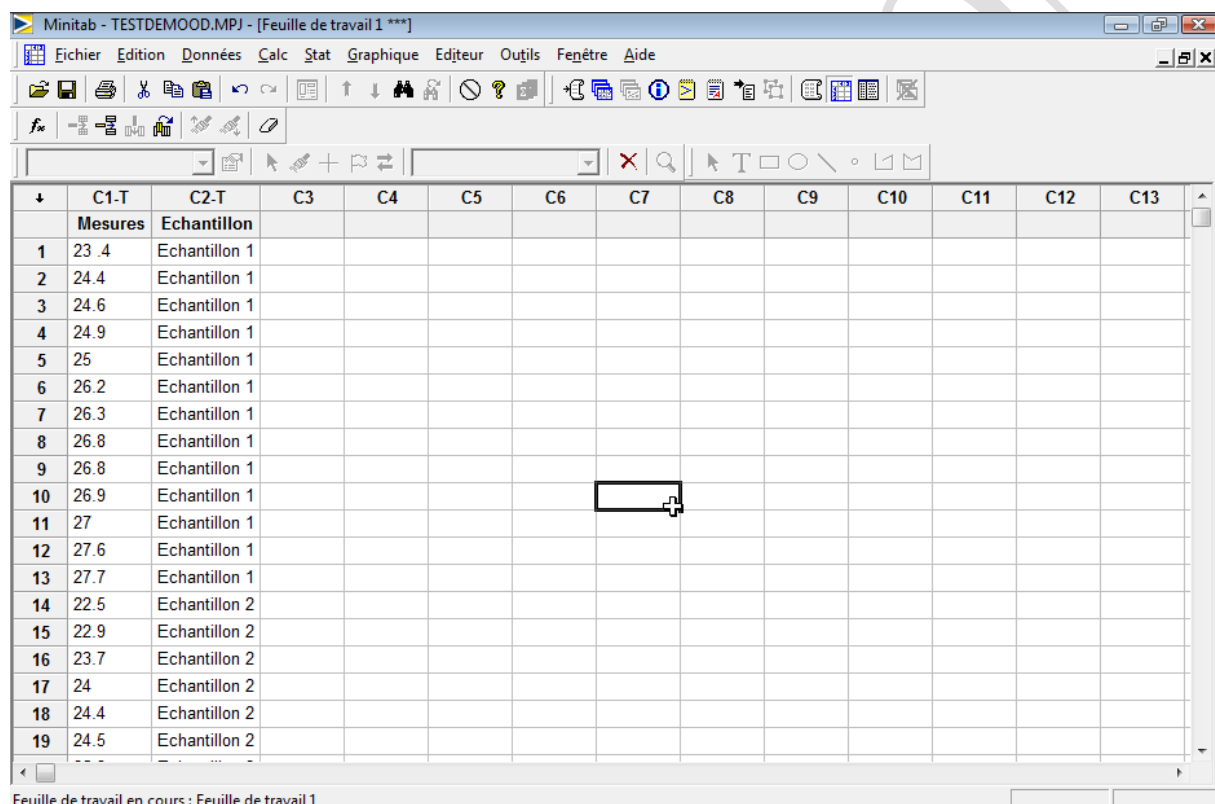


## 11.4. Exercice 79.: Test de Mood (test des médianes)

Minitab® Statistical Software 15.1.1.0

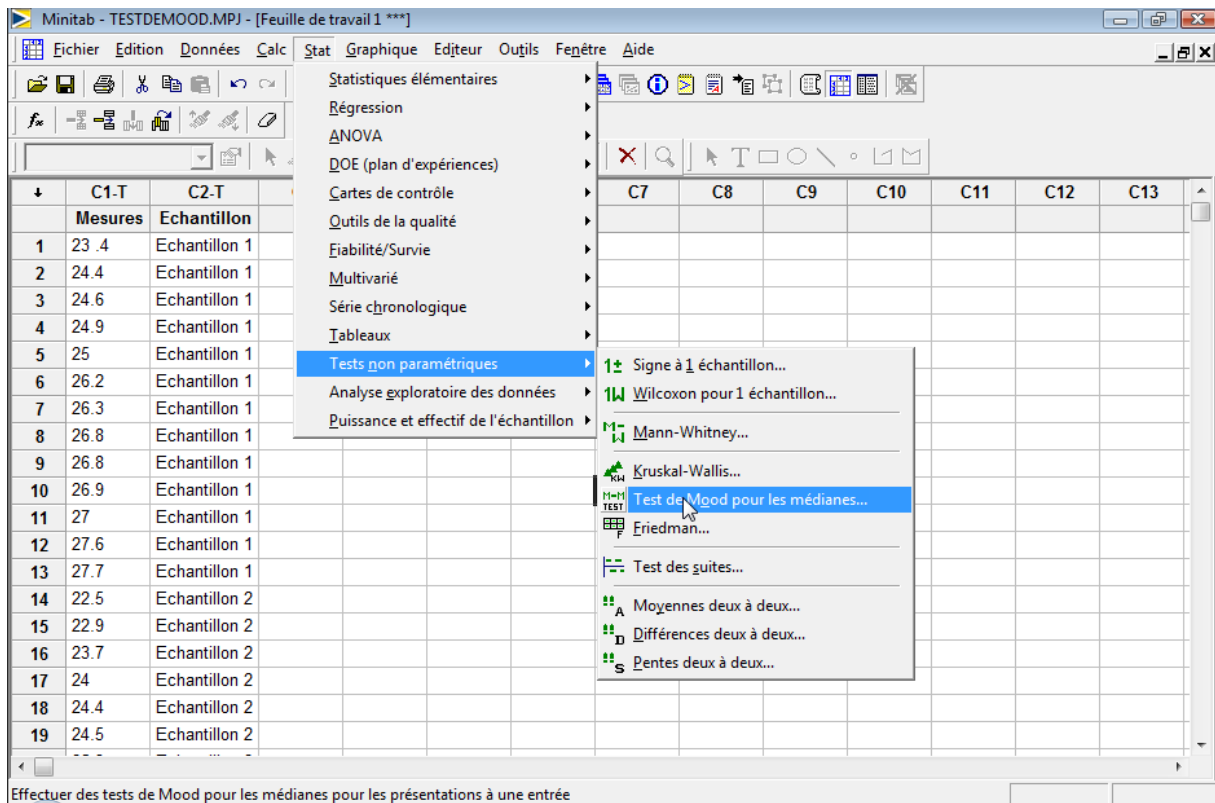
Le but du test de Mood n'est pas de contrôler si des données appariées peuvent être considérées comme égales ou non en se basant sur leur médiane (test des signes), mais simplement de vérifier sur la base d'un tableau de contingence du Khi-deux, si le nombre de valeurs au-dessus ou en-dessous de la médiane de deux échantillons est significativement différent et provient donc de deux populations différentes.

Nous allons voir un exemple et montrer que nous pouvons l'obtenir intégralement à partir du test du signe et du Khi-deux. Considérons le fichier suivant *MoodTest.mpj*:

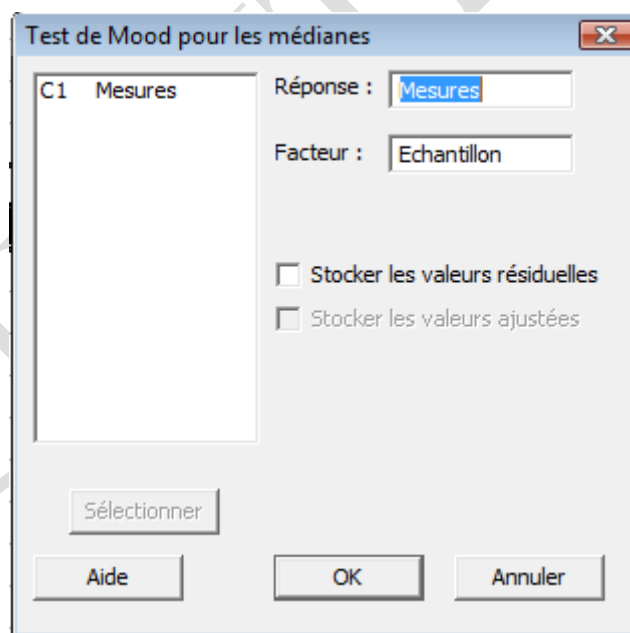


	C1-T	C2-T	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Mesures	Echantillon											
1	23.4	Echantillon 1											
2	24.4	Echantillon 1											
3	24.6	Echantillon 1											
4	24.9	Echantillon 1											
5	25	Echantillon 1											
6	26.2	Echantillon 1											
7	26.3	Echantillon 1											
8	26.8	Echantillon 1											
9	26.8	Echantillon 1											
10	26.9	Echantillon 1											
11	27	Echantillon 1											
12	27.6	Echantillon 1											
13	27.7	Echantillon 1											
14	22.5	Echantillon 2											
15	22.9	Echantillon 2											
16	23.7	Echantillon 2											
17	24	Echantillon 2											
18	24.4	Echantillon 2											
19	24.5	Echantillon 2											

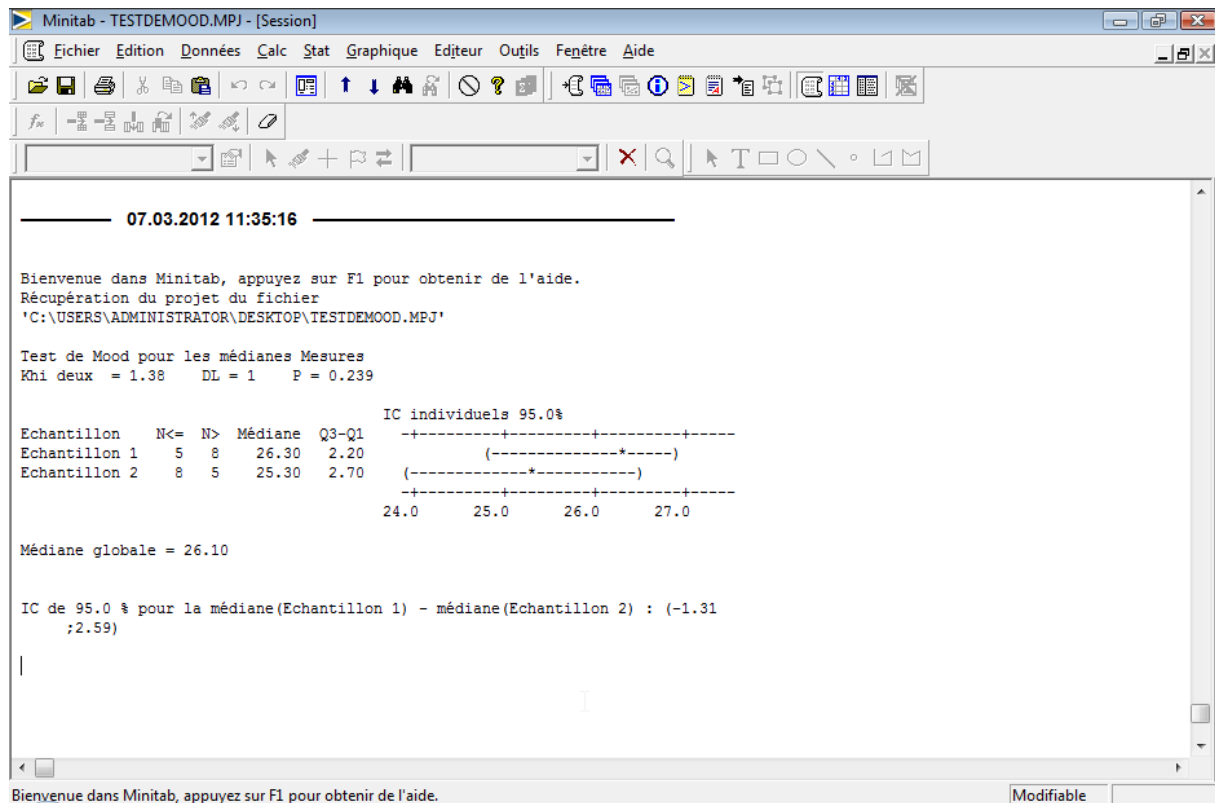
Nous allons dans le menu **Stat/Tests non paramétriques/Test de Mood pour les médianes...**:



Apparaît alors:



Nous validons par **OK**:



Nous allons maintenant vérifier que nous retrouvons toutes les valeurs à partir de deux tests déjà étudiés jusqu'ici excepté pour:

1. L'intervalle entre le premier et le troisième quartile. Nous voyons tout de suite par ailleurs dans le schéma ci-dessus que cet intervalle est faux pour le deuxième échantillon (il l'est aussi pour le premier mais cela se voit mieux avec le deuxième). Effectivement, si l'intervalle de confiance à 95% se situe à peu près à l'oeil entre 24 et 26.5 (soit une différence de 2.5), comment l'espace entre le troisième quartile et le quatrième pourrait-il valoir plus?
2. L'intervalle de confiance de la différence des médianes qui est un mystère pour moi...? Tous les logiciels de statistiques que j'ai testés à ce jour utilisent des méthodes différentes pour cet intervalle et donnent parfois des différences de plus de 100% !!! Un logiciel comme SPSS ne communique simplement pas cette information. Je recommande personnellement à nouveau l'usage du bootstrapping!

Cependant! Nous retrouvons la même médiane globale de 26.10 que dans le cours de statistique théorique (mais bon ça c'est simpliste...) et effectivement 8 valeurs qui sont plus grandes que cette médiane dans le premier échantillon et respectivement 5 pour le deuxième échantillon comme nous l'avons calculé dans le cours théorique avec Microsoft Excel.

Vérifions maintenant que nous arrivons facilement à obtenir la partie suivante du résultat du test:

Echantillon	N<=	N>	Médiane	Q3-Q1	IC individuels 95.0%
Echantillon 1	5	8	26.30	2.20	(-----*-----)

Pour cela, nous prenons les données de l'échantillon 1 seul:

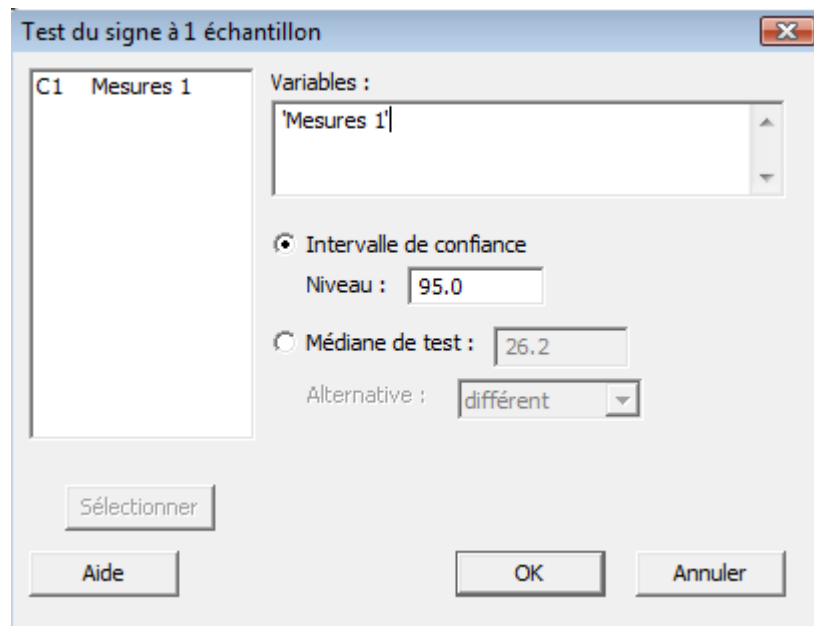
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Mesures 1												
1	23.4												
2	24.4												
3	24.6												
4	24.9												
5	25.0												
6	26.2												
7	26.3												
8	26.8												
9	26.8												
10	26.9												
11	27.0												
12	27.6												
13	27.7												
14													

et nous faisons un test du signe:

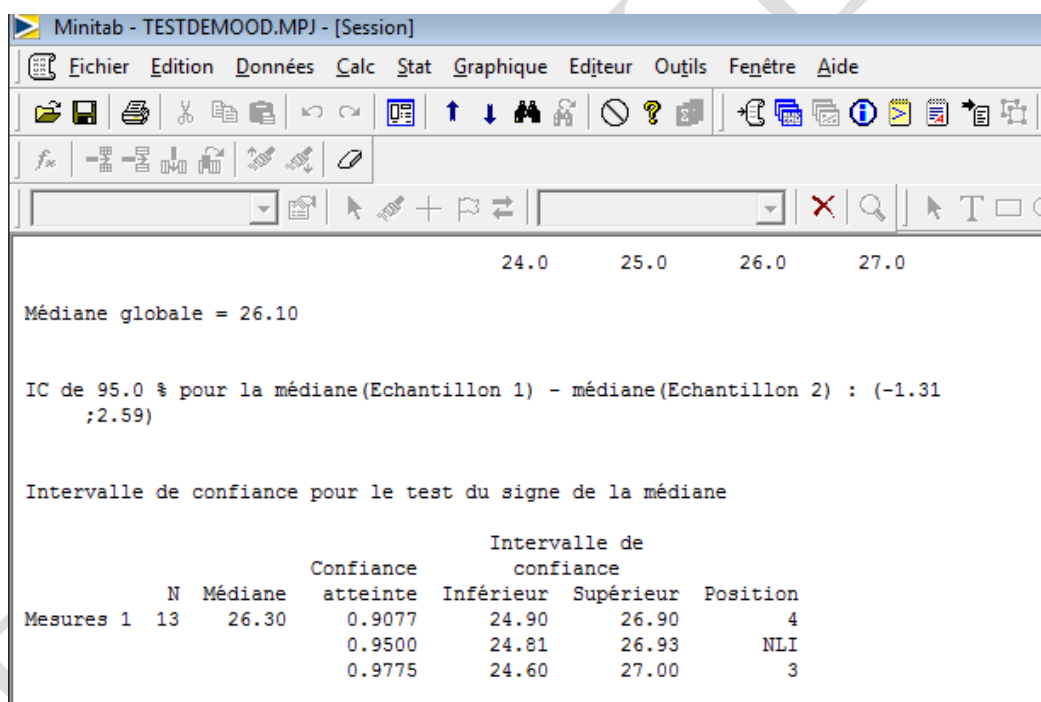
The screenshot shows the Minitab 'Stat' menu with the following options:

- Statistiques élémentaires
- Régression
- ANOVA
- DOE (plan d'expériences)
- Cartes de contrôle
- Outils de la qualité
- Fiabilité/Survie
- Multivarié
- Série chronologique
- Tableaux
- Tests non paramétriques** (highlighted)
  - 1± Signe à 1 échantillon...** (highlighted)
  - 1W Wilcoxon pour 1 échantillon...
  - M-M Mann-Whitney...
  - K-W Kruskal-Wallis...
  - M-M TEST Test de Mgod pour les médianes...
  - F Friedman...
  - Test des suites...
  - A Moyennes deux à deux...
  - D Différences deux à deux...
  - S Pentés deux à deux...
- Analyse exploratoire des données
- Puissance et effectif de l'échantillon

avec un intervalle de confiance de 95%:



Nous obtenons alors:



Les résultats sont plus conformes à ce que nous avons calculé dans le cours théorique avec Microsoft Excel et aussi plus détaillé! Nous remarquons qu'à 95%, la différence de l'intervalle est de 26.93-24.81 soit 2.12. Alors comment la différence entre le 1<sup>er</sup> et 3<sup>ème</sup> quartile pourrait valoir 2.20 comme nous l'indique Minitab® Statistical Software dans le test de Mood? Affaire à suivre...

La démarche est identique et avec les mêmes conclusions pour l'échantillon 2.

Maintenant, vérifions cette partie du test de Mood:

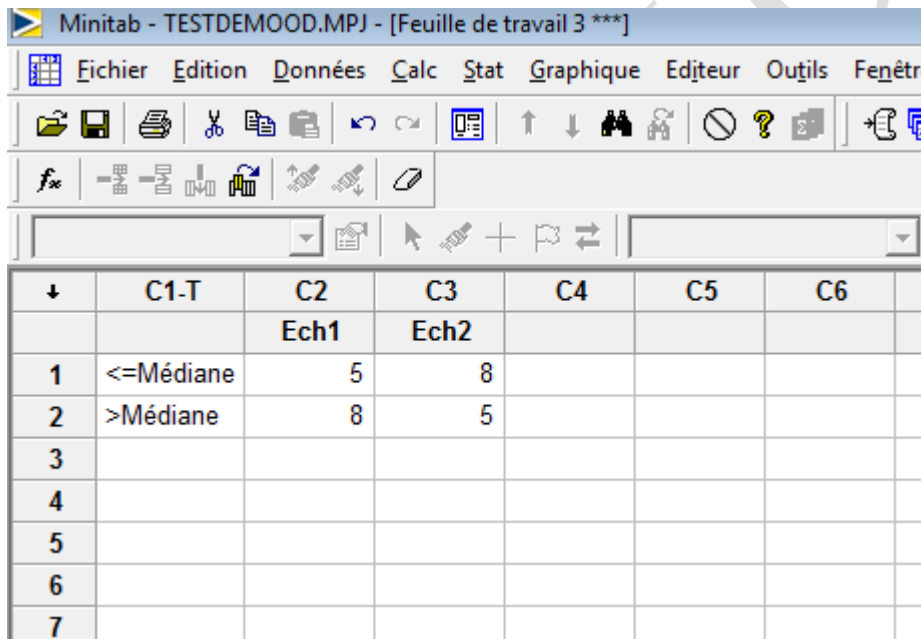
Test de Mood pour les médianes Mesures  
 Khi deux = 1.38 DL = 1 P = 0.239

Il s'agit donc d'un simple test du Khi-deux.

Donc si nous créons notre table de contingence comme nous l'avons appris dans le cours théorique:

	Échantillon 1	Échantillon 2	Total
<=Médiane	5	8	13
>Médiane	8	5	13
Total	13	13	25

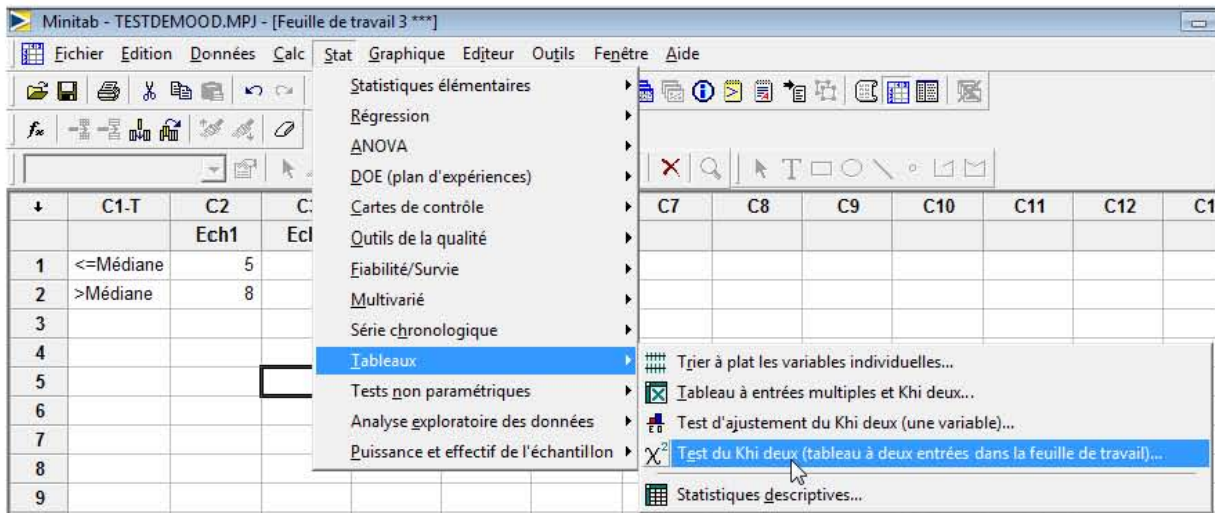
et que nous la reproduisons dans Minitab® Statistical Software:



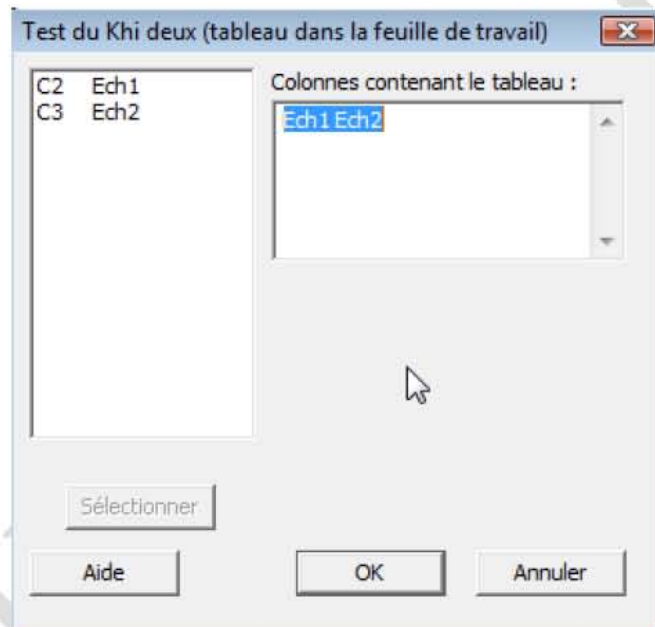
The screenshot shows the Minitab interface with a spreadsheet containing the following data:

↓	C1-T	C2	C3	C4	C5	C6
		Ech1	Ech2			
1	<=Médiane	5	8			
2	>Médiane	8	5			
3						
4						
5						
6						
7						

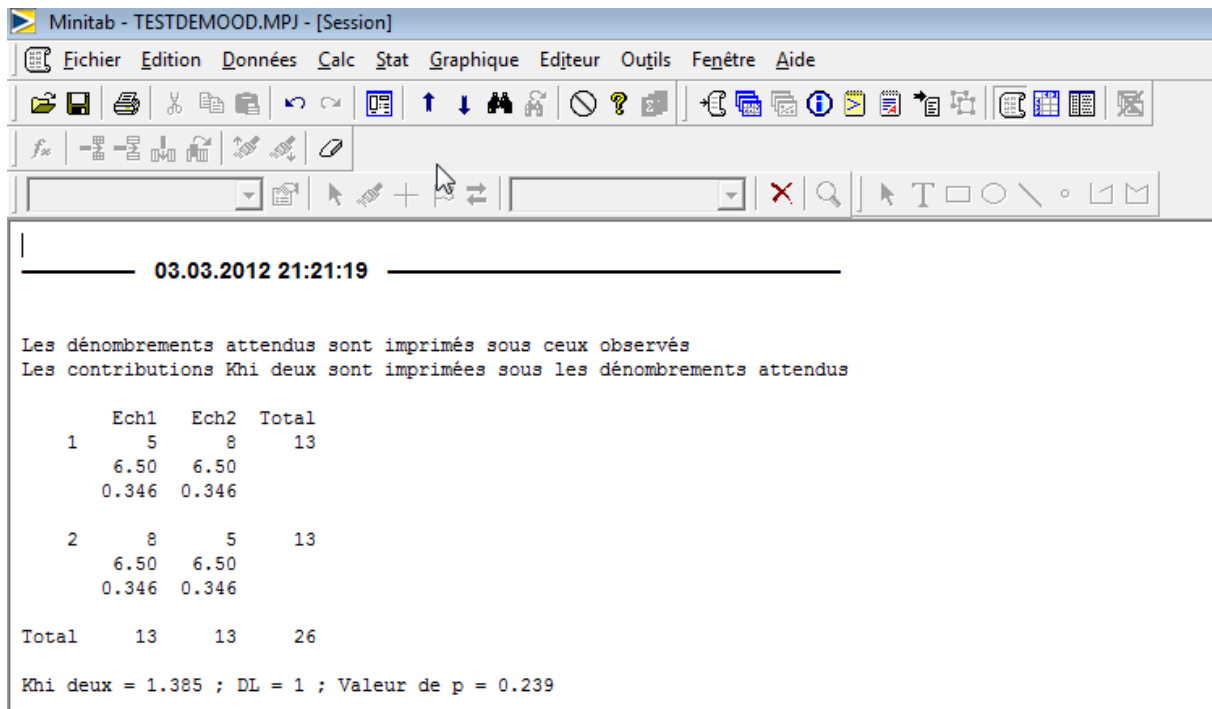
et exécutons un test du Khi-deux comme nous l'avons déjà vu plus haut:



avec les paramètres triviaux:



Nous obtenons:



Minitab - TESTDEMOOD.MPJ - [Session]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

03.03.2012 21:21:19

Les dénombrements attendus sont imprimés sous ceux observés  
Les contributions Khi deux sont imprimées sous les dénombrements attendus

	Ech1	Ech2	Total
1	5	8	13
	6.50	6.50	
	0.346	0.346	
2	8	5	13
	6.50	6.50	
	0.346	0.346	
Total	13	13	26

Khi deux = 1.385 ; DL = 1 ; Valeur de p = 0.239

Ce qui est parfaitement conforme.

Donc dans le cas présent, qu'elle est la conclusion? Rappelons que le test de Mood (dans le cas de Minitab) compare à l'aide d'une table de contingence... le contingent de valeurs au-dessus ou en-dessous de la médiane et fait un test du Khi-deux. Comme dans le cas présent, la  $p$ -value est de 0.239 nous ne considérons pas les deux populations comme différentes.



## 11.5. Exercice 80.: Test de la somme des rangs de (Wilcoxon)Mann-Whitney pour deux échantillons indépendants

Minitab® Statistical Software 15.1.1.0

Comme d'habitude (en chantant...) le but est de voir si nous retrouvons les résultats numériques calculés avec Microsoft Excel en utilisant les relations démontrées dans le cours de statistique théorique.

Par contre nous n'allons pas reprendre les mêmes valeurs qu'en dans le cours théorique car l'exemple y est trop petit (pour des raisons pédagogiques). Nous allons donc d'abord faire l'exemple avec Microsoft Excel et ensuite avec Minitab.

Considérons les deux séries  $X$  et respectivement  $Y$  de valeurs suivantes:

197, 162, 57, 108, 53.5, 55, 77, 39, 66, 48, 121, 79, 309

50.5, 50, 557, 42, 23, 26, 45, 96, 113, 30, 33, 45

Nous combinons les valeurs et indiquerons les rangs respectifs sous forme de tableau pour la suite:

Valeurs combinées	Variable	Rang global
50.5	$X$	11
50	$X$	10
557	$X$	25
42	$X$	6
23	$X$	1
26	$X$	2
45	$X$	7
96	$X$	18
113	$X$	20
30	$X$	3
33	$X$	4
45	$X$	7
197	$Y$	23
162	$Y$	22
57	$Y$	14
108	$Y$	19
53.5	$Y$	12
55	$Y$	13
77	$Y$	16
39	$Y$	5
66	$Y$	15
48	$Y$	9
121	$Y$	21
79	$Y$	17

309	Y	24
-----	---	----

Ce qui donne au final les rangs respectifs:

Variable	Rang global	Variable	Rang global
X	11	Y	23
X	10	Y	22
X	25	Y	14
X	6	Y	19
X	1	Y	12
X	2	Y	13
X	7	Y	16
X	18	Y	5
X	20	Y	15
X	3	Y	9
X	4	Y	21
X	7	Y	17
		Y	24
<b>Somme:</b>	<b>114</b>		<b>210</b>

Nous choisirons comme base la variable X. Nous avons alors pour l'espérance de la somme de ces rangs:

$$E(W_S) = E\left(\sum_{i=1}^{n_x} R_i\right) = \frac{1}{2} n_x (N+1) = \frac{13(25+1)}{2} = 169$$

et nous avons:

$$V(W_S) = V\left(\sum_{i=1}^{n_x} R_i\right) = \frac{n_x n_y (n_x + n_y + 1)}{12} = \frac{n_x n_y (N+1)}{12} = 338$$

Nous avons alors:

$$Z = N(0,1) = \frac{W_s - \frac{1}{2} n_x (N+1)}{\sqrt{\frac{n_x n_y (N+1)}{12}}} = \frac{210 - 169}{\sqrt{422.5}} \cong 2.23$$

Cela correspond donc à une probabilité cumulée:

$$= \text{LOI.NORMALE.STANDARD.N}(2.23; \text{VRAI}) = 0.987$$

Ce qui correspond en bilatéral à une  $p$ -value de (soit la même que Minitab 15.1.2 au deux dixièmes près):

$$p\text{-value} = 2 \cdot (1 - 0.987) = 2 \cdot 0.013 = 0.026 = 2.6\%$$

Donc à un seuil de 5%, nous devons rejeter l'hypothèse nulle comme quoi les deux distributions sont symétriques et identiques.

Dans Minitab, nous avons respectivement:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
	X	Y									
1	197.0	50.5									
2	162.0	50.0									
3	57.0	557.0									
4	108.0	42.0									
5	53.5	23.0									
6	55.0	26.0									
7	77.0	45.0									
8	39.0	96.0									
9	66.0	113.0									
10	48.0	30.0									
11	121.0	33.0									
12	79.0	45.0									
13	309.0										
14											
15											
16											
17											
18											
19											

Et nous allons dans le menu **Stat/Tests non paramétriques/Mann-Withney...** (nom au passage qui n'est pas très bien choisi au vu du nombre de test comportant le terme "Mann-Withney"...):

Minitab - WilcoxonMannWithney.MPJ - [Feuille de travail 1 \*\*\*]

Stat > Tests non paramétriques > Mann-Whitney...

	C1	C2	C3
	X	Y	
1	197.0	50.5	
2	162.0	50.0	
3	57.0	557.0	
4	108.0	42.0	
5	53.5	23.0	
6	55.0	26.0	
7	77.0	45.0	
8	39.0	96.0	
9	66.0	113.0	
10	48.0	30.0	
11	121.0	33.0	
12	79.0	45.0	
13	309.0		
14			
15			
16			
17			
18			
19			
20			

Effectuer des tests de Mann-Whitney à deux échantillons pour les médianes

Nous avons alors:

**Mann-Whitney**

Premier échantillon : X

Second échantillon : Y

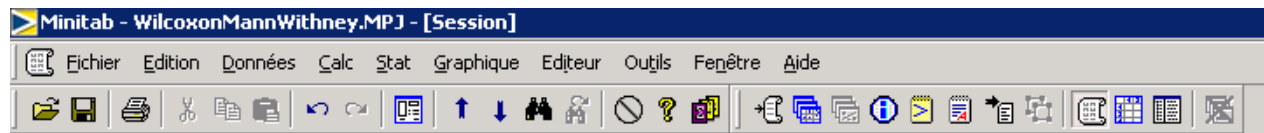
Niveau de confiance : 95.0

Alternative : différent

Sélectionner

Aide OK Annuler

et nous validons:



29.08.2012 17:38:20

### Test de Mann-Whitney et IC : X; Y

	N	Médiane
X	13	77.0
Y	12	45.0

L'estimation ponctuelle de ETA1-ETA2 est 27.0  
Le pourcentage IC (95.3) pour ETA1-ETA2 est (4.5 ; 75.0)  
W = 210.0  
Le test de ETA1 = ETA2 contre ETA1 différent de ETA2 est significatif sur  
0.0276  
Le test est significatif sur 0.0276 (ajusté pour les nombres de même grandeur)

Commentons un peu ce résultat. D'abord, nous avons effectivement vu dans le cours théorique les valeurs des deux médianes calculées simplement avec Microsoft Excel. La différence de l'estimation ponctuelle est un mystère... par ailleurs nous ne la trouvons pas dans la démonstration théorique du test et il en est de même pour le pourcentage de l'intervalle de confiance (qui de toute manière peut aussi être obtenue par Bootstrapping).

La valeur de  $W$  de 210 est conforme à la somme des rangs de l'échantillon  $X$  tel que démontré dans le cours théorique. La  $p$ -value est un peu plus élevée que ce que nous avons obtenu avec la relation démontrée dans le cours théorique qui nous avait permis d'obtenir une valeur 0.026 (mais cette différence n'est pas suffisante pour changer le résultat du test).

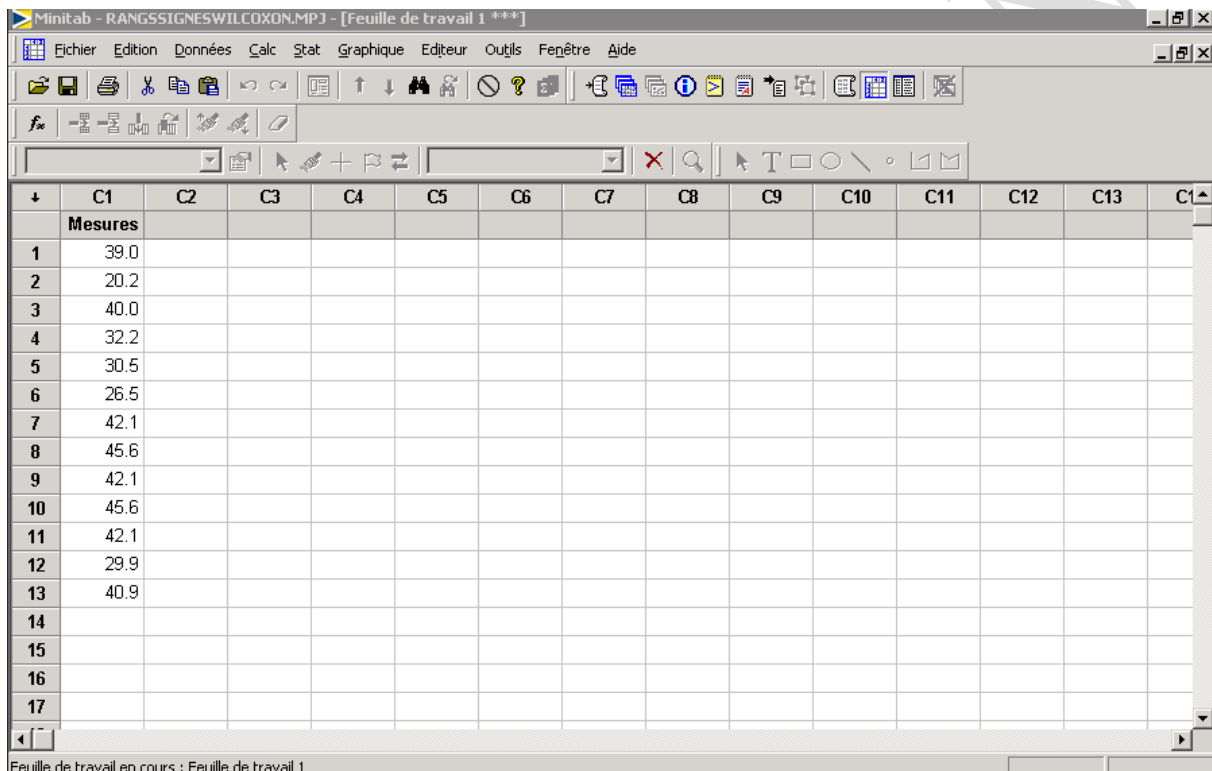
Nous ne parlerons pas des autres valeurs n'ayant pas vu dans le cours théorique les moyens de les démontrer mathématiquement.

## 11.6. Exercice 81.: Test de la somme des rangs signés de Wilcoxon pour 1 échantillon

Minitab® Statistical Software 15.1.1.0

Encore une fois, nous souhaitons comparer les démonstrations faites dans le cours de statistiques théorique avec l'application numérique approximée dans Microsoft Excel avec le résultat obtenu avec Minitab.

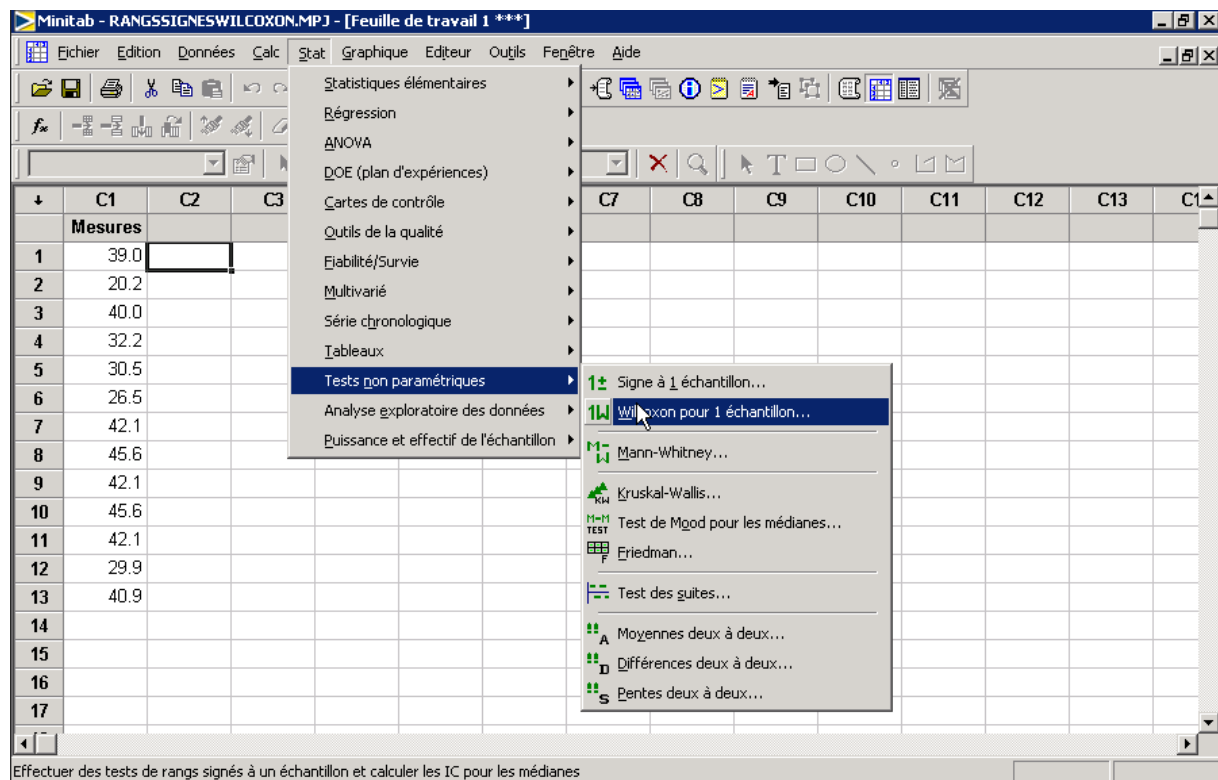
Nous partons donc du même échantillon que dans le cours théorique:



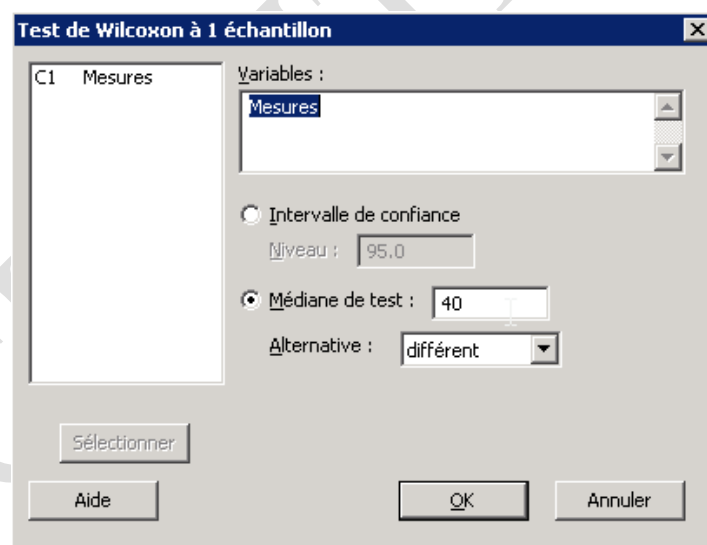
The screenshot shows the Minitab software interface with a data table. The table has 14 columns labeled C1 through C14 and 17 rows. The first row is labeled 'Mesures'. The data in column C1 is as follows:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
1	39.0													
2	20.2													
3	40.0													
4	32.2													
5	30.5													
6	26.5													
7	42.1													
8	45.6													
9	42.1													
10	45.6													
11	42.1													
12	29.9													
13	40.9													
14														
15														
16														
17														

Nous allons dans le menu **Stat/Tests non paramétriques/Wilcoxon pour 1 échantillon...**:



Nous mettons alors comme dans le cours théorique la médiane échantillonnale de 40 en bilatéral:



et nous obtenons:

### Test des rangs signés de Wilcoxon : Mesures

Test de médiane = 40.00 contre médiane différente de 40.00

	N pour	Statistique	Médiane
	N	test	P
		de Wilcoxon	estimée
Mesures	13	12	26.0 0.327 36.55

Nous avons donc la même "statistique de Wilcoxon" de 26 ce qui conforte notre façon de calculer les rangs et nous dit que Minitab utilise les rangs positifs pour indiquer la somme des rangs et nous avons un  $p$ -value de 32.7 à comparer à celle de 30.76% obtenue dans le cours (voir le support de cours MATLAB pour comprendre la différence).

Donc dans le cas présent, nous ne rejetons donc pas l'hypothèse nulle comme quoi la médiane échantillonnale est proche de la médiane vraie et que la distribution est symétrique.

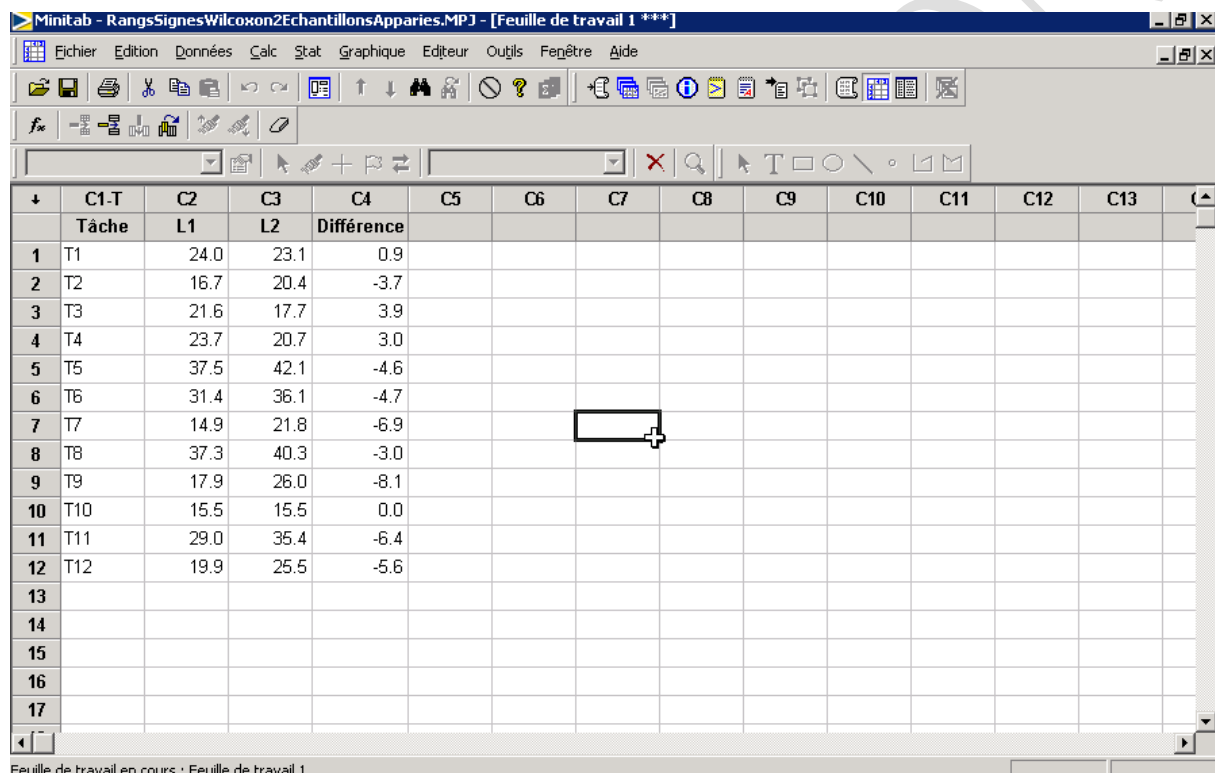


## 11.7. Exercice 82.: Test de la somme des rangs signés de Wilcoxon pour 2 échantillons appariés

Minitab® Statistical Software 15.1.1.0

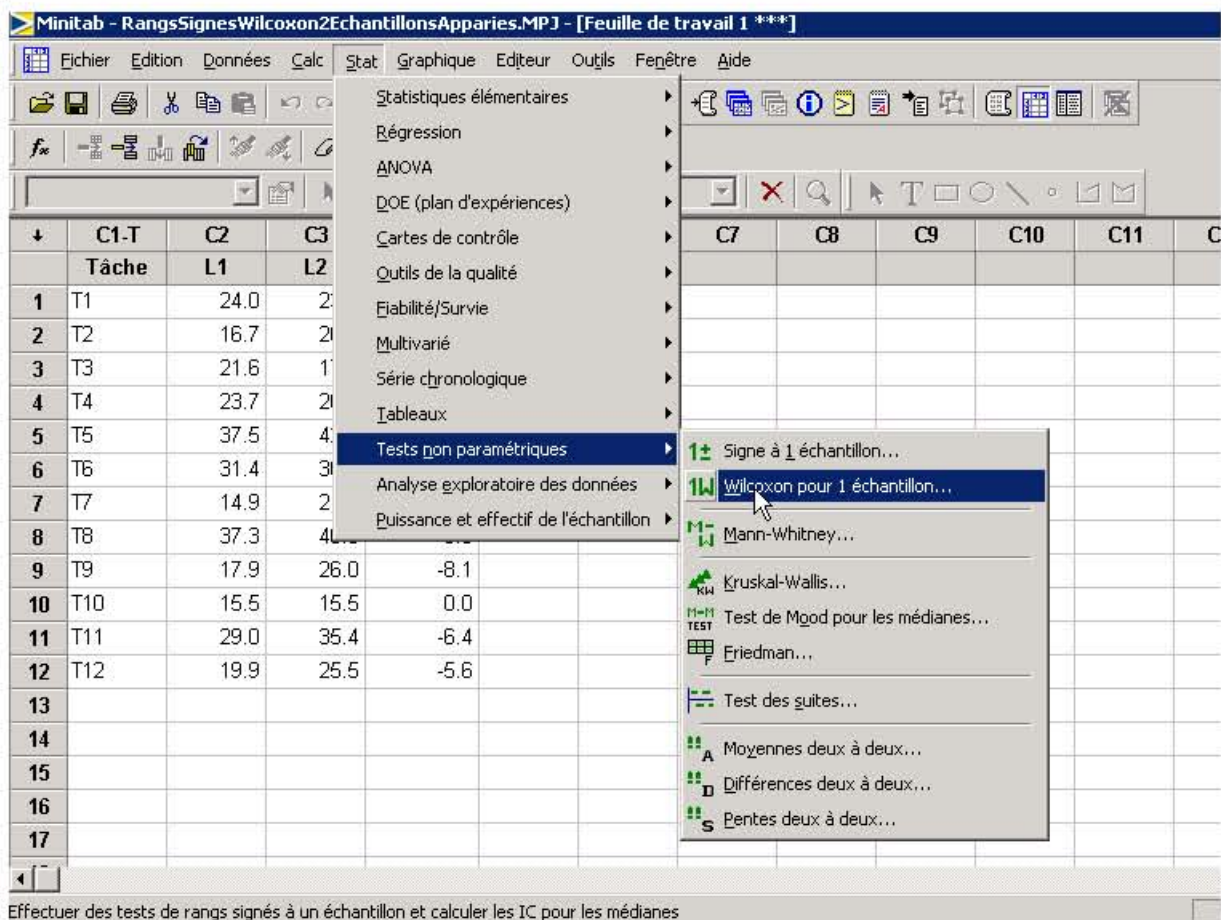
Nous voulons appliquer le test des rangs signés de Wilcoxon pour 2 échantillons appariés comme démontré dans le cours de statistique théorique bien que celui-ci ne soit pas disponible explicitement dans Minitab et ce afin de comparer au résultat calculé avec Microsoft Excel dans le cadre de l'approximation par une loi Normale.

Nous partons donc du même cas que dans le cours de statistique théorique:



	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Tâche	L1	L2	Différence									
1	T1	24.0	23.1	0.9									
2	T2	16.7	20.4	-3.7									
3	T3	21.6	17.7	3.9									
4	T4	23.7	20.7	3.0									
5	T5	37.5	42.1	-4.6									
6	T6	31.4	36.1	-4.7									
7	T7	14.9	21.8	-6.9									
8	T8	37.3	40.3	-3.0									
9	T9	17.9	26.0	-8.1									
10	T10	15.5	15.5	0.0									
11	T11	29.0	35.4	-6.4									
12	T12	19.9	25.5	-5.6									
13													
14													
15													
16													
17													

Nous allons dans le menu **Stat/Tests non paramétriques/Wilcoxon pour 1 échantillon...**:



et nous mettons:



Nous obtenons alors dans la fenêtre de session:

**Test des rangs signés de Wilcoxon : Différence**

Test de médiane = 0.000000 contre médiane différente de 0.000000

	N pour	Statistique		Médiane	
	N	test	de Wilcoxon	P	estimée
Différence	12	11	8.5	0.033	-3.100

Nous retrouvons la statistique de Wilcoxon 8.5 et une  $p$ -value de 3.3% donc une valeur bien différente que celle obtenue avec Microsoft Excel (puisque Minitab ne fait pas l'approximation par une loi Normale) mais toutefois la conclusion est la même!

## 12. ANOVA paramétriques ou non paramétriques (balancées!)

Nous le savons, le nombre d'ANOVA possibles et imaginables existe en grand nombre. Nous allons donc ici uniquement nous concentrer sur celles étudiées et démontrées en détails dans le cours théorique.

### 12.1. Exercice 83.: Puissance du test de l'ANOVA (WP)

Minitab® Statistical Software 15.1.1.0

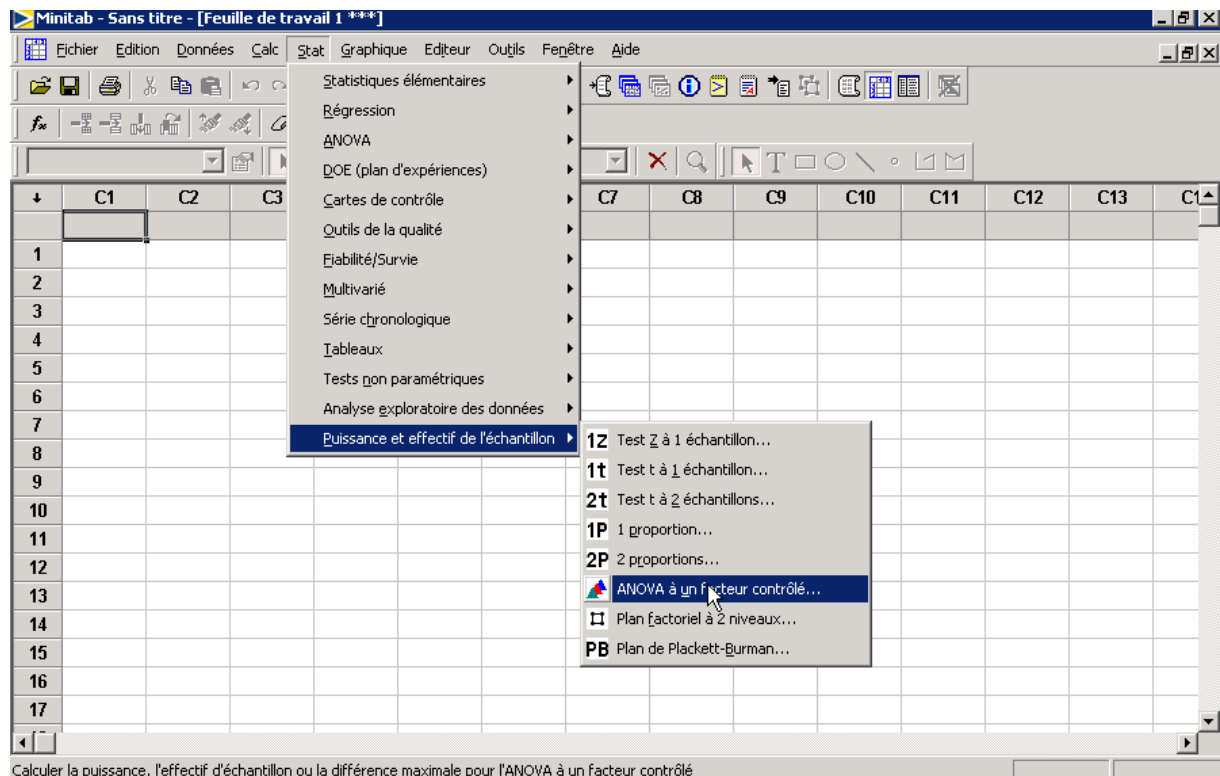
Nous avons vu dans le cours de statistique théorique que certains développements de l'ANOVA nécessitent des hypothèses très fortes. Nous pouvons alors (et devons!) nous poser la question de la taille des échantillons avant exécution de l'analyse en imposant une puissance à ce test.

Nous avons vu en faisant des simulations dans Microsoft Excel que les conditions étaient assez contraignantes dans la pratique pour avoir une puissance de test satisfaisante.

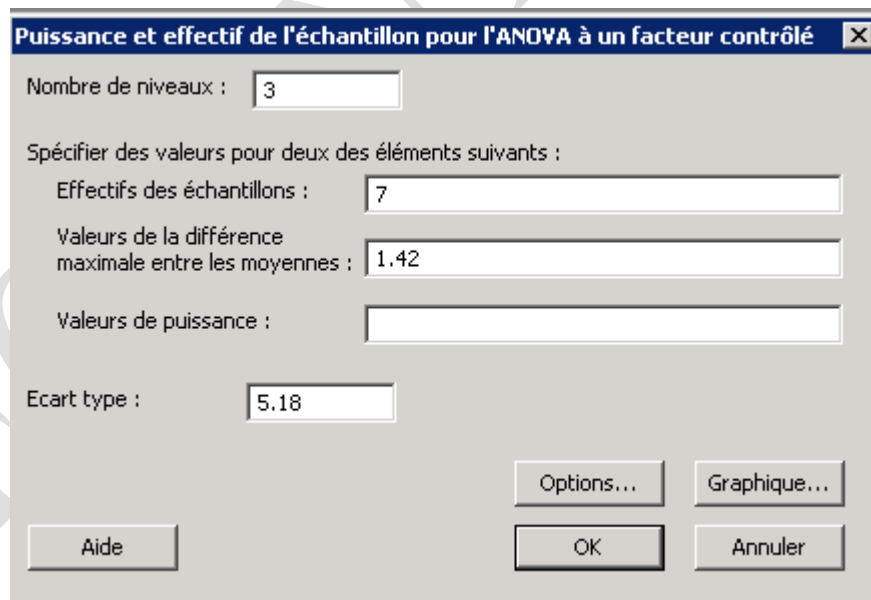
Pour voir cela, reprenons l'exemple suivant (bon faire de l'analyse de la puissance a posteriori n'a pas vraiment de sens comme nous l'avons démontré dans le cours théorique mais voyons quand même pour l'idée...):

	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
		Equipe 1	Equipe 2	Equipe 3									
1	Lundi	78	77	88									
2	Mardi	88	75	86									
3	Mercredi	90	80	79									
4	Jeudi	77	83	93									
5	Vendredi	85	87	79									
6	Samedi	88	90	83									
7	Dimanche	79	85	79									
8													
9													
10													
11													
12													
13													
14													
15													
16													
17													
18													
19													
20													
21													
22													
23													

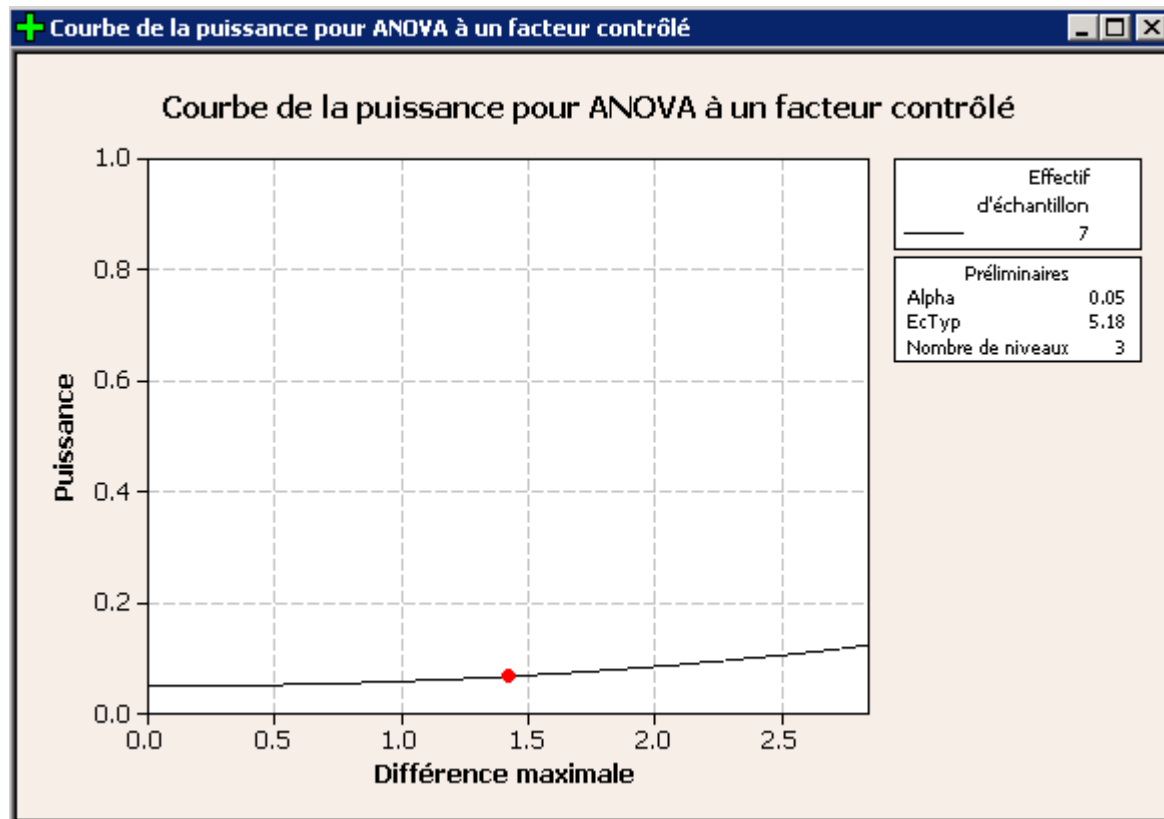
et faisons un analyse de la puissance des mesures ci-dessous. Pour cela, nous allons dans le menu **Stat/Puissance et effectif de l'échantillon/ANOVA à 1 facteur contrôlé...**:



et nous y mettons ce qui correspond à notre tableau:



pour obtenir à un seuil de 5%:



et donc la puissance de notre test est finalement misérable ce qui comme nous l'avons démontré dans le cours théorique est normal puisque la  $p$ -value de ce test est très grande.

Maintenant, faisons les choses comme elles devraient l'être, en sachant que la valeur minimale de la puissance d'un test pour être recevable est a priori de 80%.

Regardons typiquement ce que cela nous donne puisque nous ne pouvons pas jouer avec le nombre de jours dans une semaine, ni avec le nombre d'équipes, ni avec l'écart-type...:

Number of levels: 3

Specify values for two of the following elements:

Sample sizes: 7

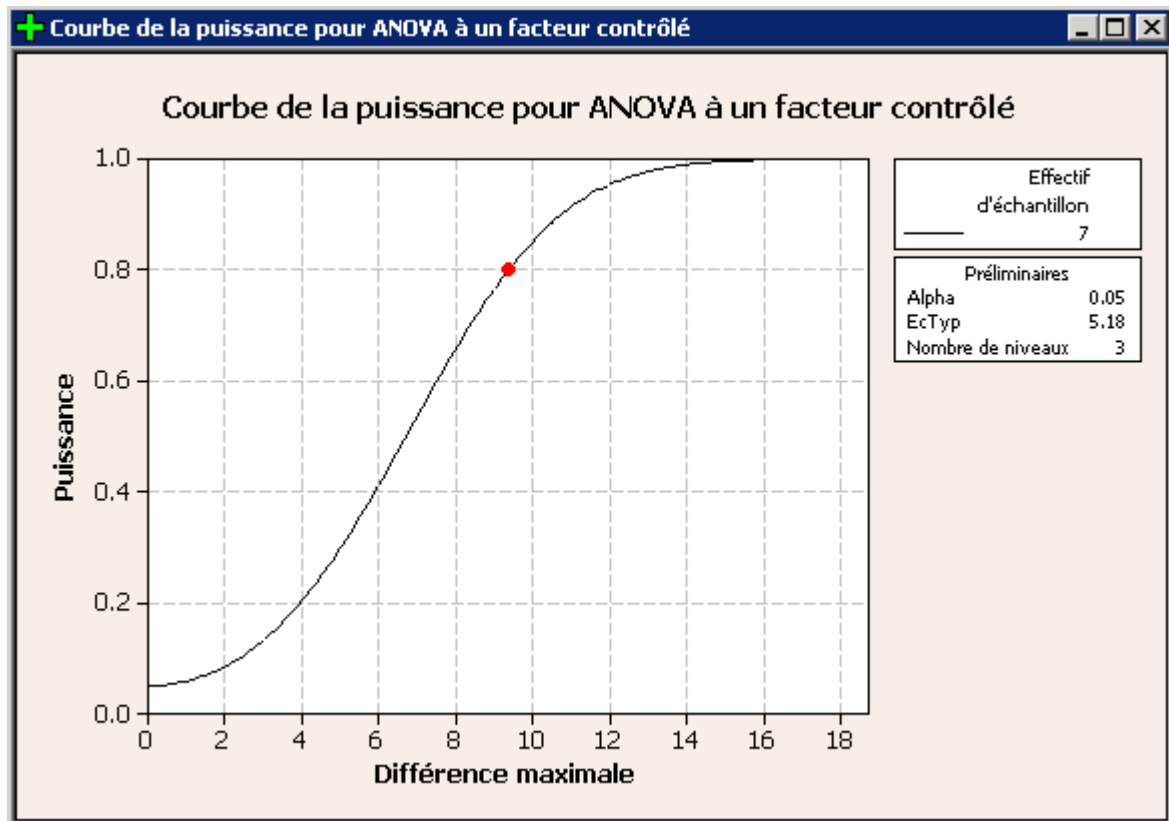
Maximum difference between means: [ ]

Power values: 0.8

Standard deviation: 5.18

Buttons: Aide, Options..., Graphique..., OK, Annuler

Il vient alors:



Ce qui signifie donc qu'en imposant une puissance de 80% de l'ANOVA une différence de 10 (maximum) entre les moyennes des échantillons sera tolérée (ce qui explique aussi pourquoi nous avons obtenu un  $p$ -value très grande avec les valeurs d'origine).

Mais par rapport à ce que nous avons vu dans le cours théorique et complètement indépendamment de l'exemple choisi ici, regardons par curiosité le cas fréquent en laboratoire suivant:

Nombre de niveaux :

Spécifier des valeurs pour deux des éléments suivants :

Effectifs des échantillons :

Valeurs de la différence maximale entre les moyennes :

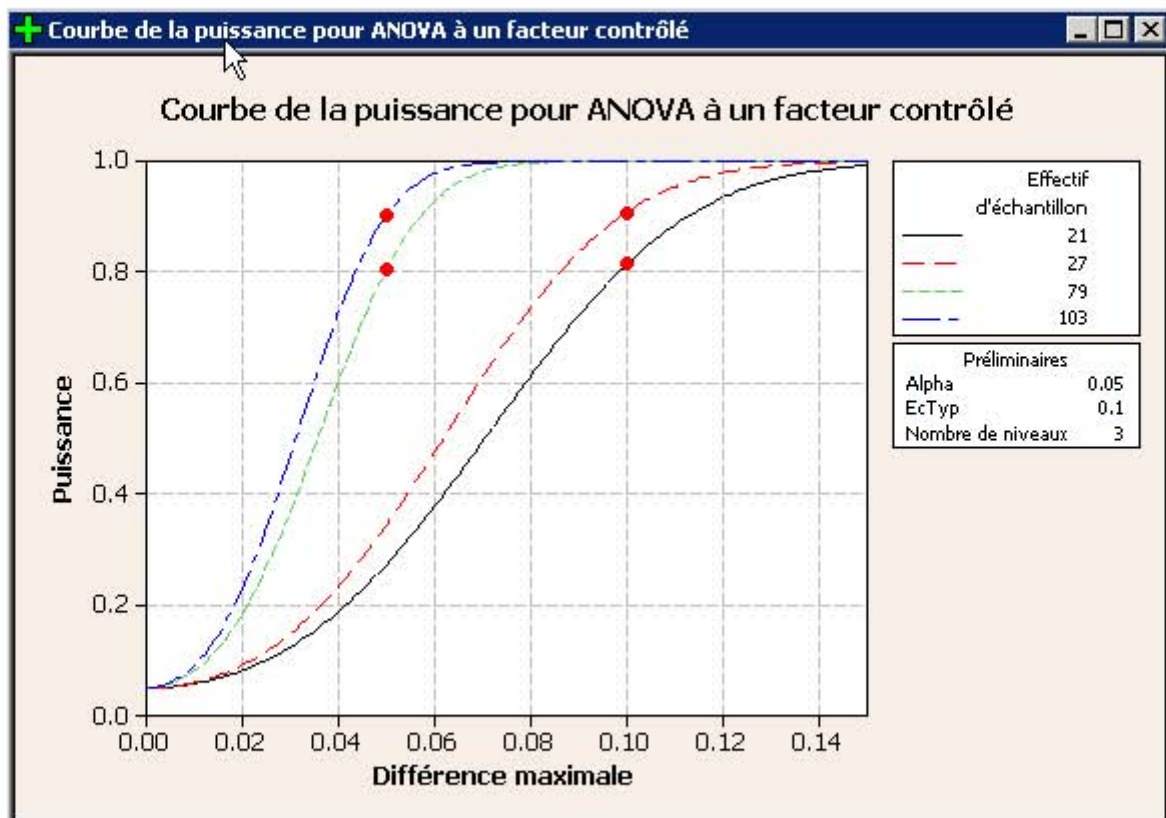
Valeurs de puissance :

Ecart type :

Options... Graphique...

Aide OK Annuler

Ce qui nous donne des courbes riches d'enseignement:



ÉCHANTILLON



## 12.2. Exercice 84.: Analyses graphiques des interactions et effets principaux

Minitab® Statistical Software 15.1.1.0

Le but de cet exemple ne sera pas de vérifier que quoi que ce soit est conforme à Microsoft Excel puisqu'il s'agit seulement de graphiques faisant de simples moyennes arithmétiques. Mais il s'agit cependant d'un sujet important pour bien comprendre ce qu'il y a derrière l'ANOVA à deux facteurs avec répétition et le concept d'interaction.

Exceptionnellement, nous allons détailler les résultats et reprendre les explications du cours de statistique théorique.

Définitions (rappels du cours théorique):

D1. Nous disons qu'il y a absence d'interaction quand la moyenne des réponses d'un facteur en fonction de ses niveaux varie de la même amplitude et avec le même signe que la moyenne des réponses d'un autre facteur en fonction de ses niveaux. Nous disons alors que les courbes de réponses dans le diagramme des interactions sont parallèles.

Remarque: Le parallélisme des réponses est normal en situation d'absence d'interaction, car cela signifie que quel que soit le niveau de l'un ou l'autre des facteurs, la variation (si elle existe), de la réponse sera toujours la même de la même amplitude. Ce qui est caractéristique de l'indépendance (du moins localement).

D2. Nous disons que deux facteurs sont en interaction quand la moyenne des réponses d'un facteur en fonction de ses niveaux ne varie pas de la même amplitude ou/et pas avec le même signe que la moyenne des réponses d'un autre facteur en fonction de ses niveaux. Nous disons alors que les courbes de réponses dans le diagramme des interactions ne sont pas parallèles.

Remarque: L'absence d'interaction est une hypothèse très forte et une observation rare. Souvent, nous avons des interactions ou fortes interactions.

### 12.2.1. Méthodes classiques directement liées aux options ANOVA de Minitab

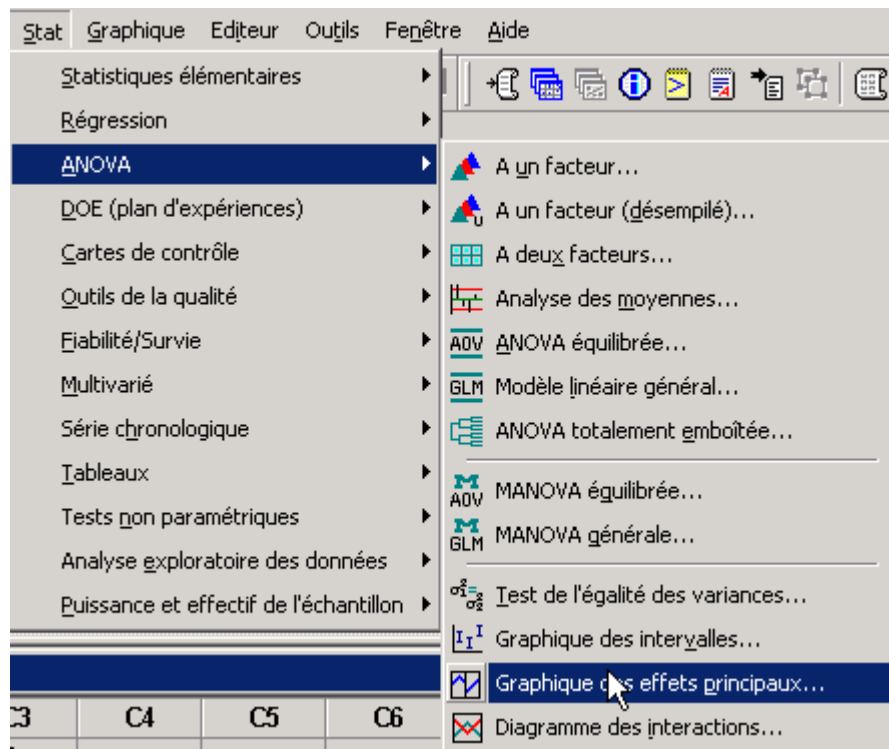
Considérons le petit tableau suivant sans mesurées répétées:

Feuille de travail 1 ***			
↓	C1-T	C2-T	C3
	Facteur1	Facteur2	Valeur
1	Niveau 1	Niveau 1	3
2	Niveau 1	Niveau 2	3
3	Niveau 2	Niveau 1	3
4	Niveau 2	Niveau 2	3

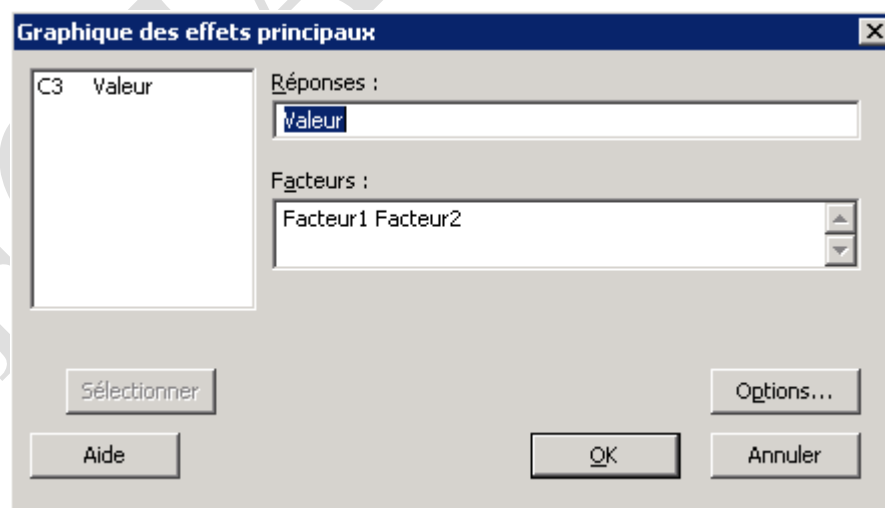
ou autrement présenté:

	Facteur 2	
Facteur 1	Niveau 1	Niveau 2
Niveau 1	3	3
Niveau 2	3	3

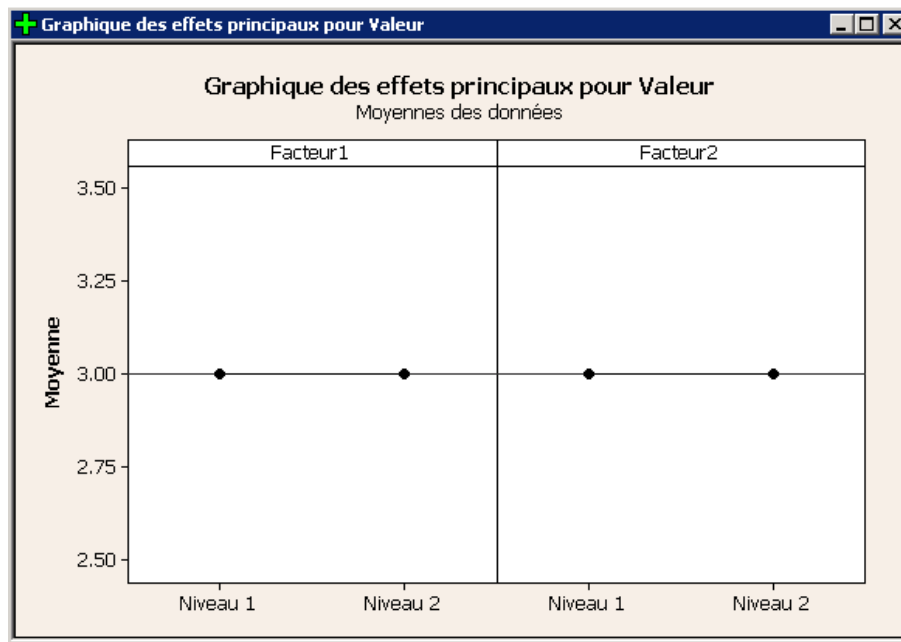
Nous souhaiterions d'abord observer le graphique des effets principaux. Pour cela nous allons dans le menu **Stat/ANOVA/Graphique des effets principaux...**:



Vient alors la boîte de dialogue suivante:

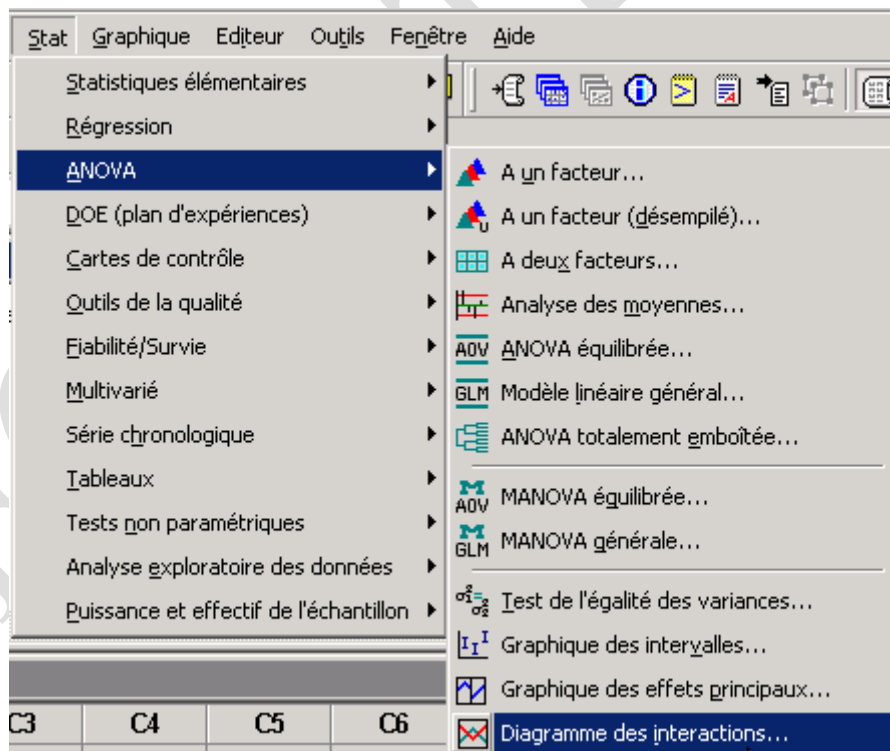


où nous avons mis les réponses et facteurs ad hoc. Nous avons alors le graphique des effets principaux suivant:

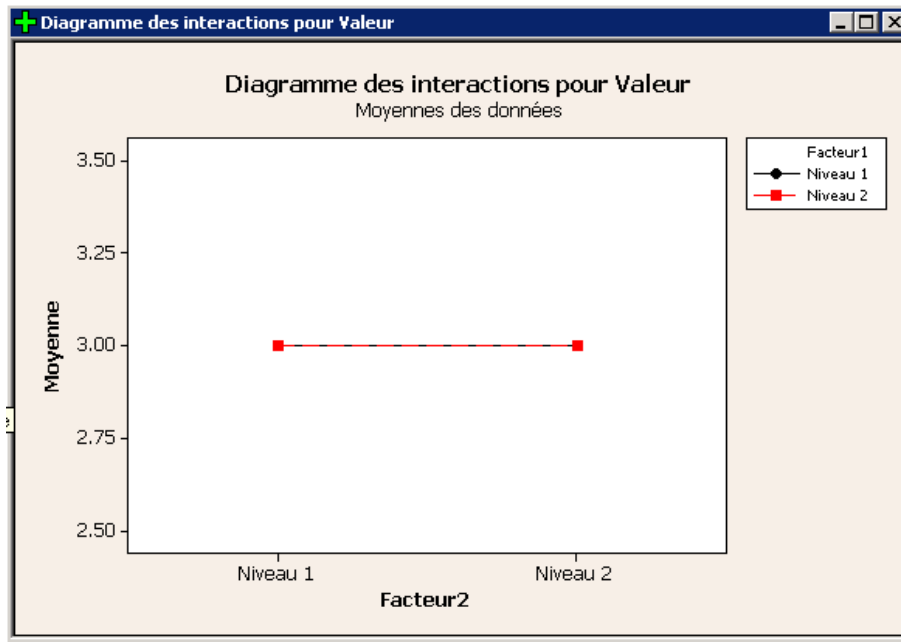


Nous voyons bien qu'aucun facteur n'a un effet principal sur quoi que ce soit. Ce qui est relativement intuitif étant donné le contenu de tableau précédent.

Maintenant, nous souhaiterions observer les interactions. Nous allons alors dans le menu **Stat/ANOVA/Diagramme des interactions...**:



et nous aurons alors sans surprise:



et comme c'est un graphique d'interaction (d'après son nom), nous pouvons constater que les facteurs n'interagissent pas entre eux (ou se neutralisent c'est selon...). Nous disons alors qu'il n'y a **(a priori) aucun effet ni aucune interaction (localement)**. Au fait dans certaines expériences, l'absence d'interaction est une hypothèse très forte et donc souvent rare. Raison pour laquelle il faut faire attention aux mots choisis lors de l'interprétation des graphiques d'interaction (car ne pas passer par les calculs purs est délicat pour cette étape voir non scientifique!).

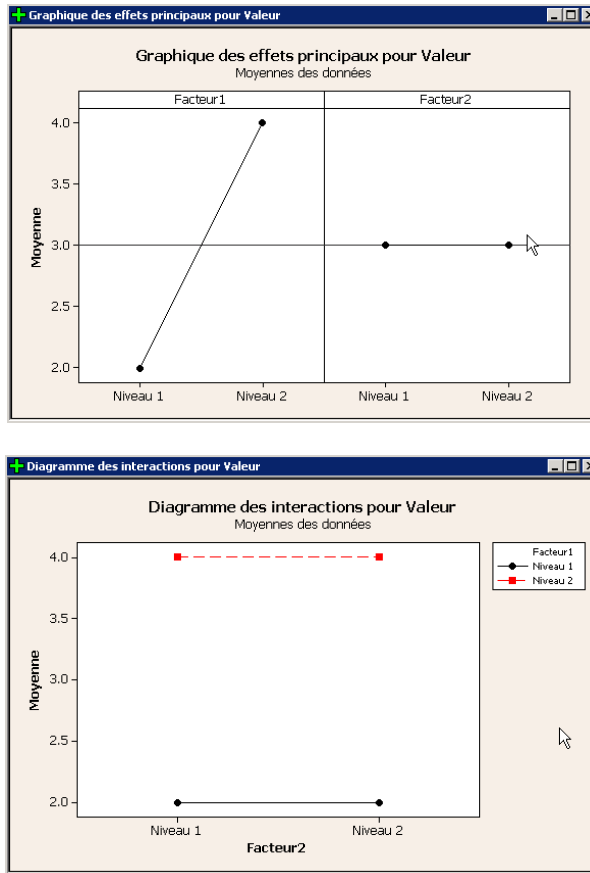
Maintenant voyons le tableau suivant:

Feuille de travail 1 ***			
↓	C1-T	C2-T	C3
	Facteur1	Facteur2	Valeur
1	Niveau 1	Niveau 1	2
2	Niveau 1	Niveau 2	2
3	Niveau 2	Niveau 1	4
4	Niveau 2	Niveau 2	4

ou autrement présenté:

Facteur 1	Facteur 2	
	Niveau 1	Niveau 2
Niveau 1	2	2
Niveau 2	4	4

Il nous paraît clair que le Facteur 1 semble avoir une influence sur la réponse. Mais voyons cela sous forme de graphique. Nous avons alors pour les deux graphiques (la procédure est identique en tout point à ce que nous avons fait précédemment):



Il apparaît clairement dans le graphique du dessus, qu'il y a uniquement le niveau du Facteur 1 qui influence la réponse. Alors que le Facteur 2 n'influence en rien la réponse. Nous disons alors qu'il y a effet principal (localement) du Facteur 1.

Sur le graphique de droite (d'interactions), nous avons la même information, mais sous une forme différente. Nous voyons que quel que soit le niveau du Facteur 2, les réponses sont horizontales et donc celui-ci n'influence en rien les résultats. Nous sommes donc dans une situation où **(a priori) l'effet principal est (localement) le Facteur 1 et en absence interactions entre les facteurs.**

Voyons maintenant le tableau suivant:

Feuille de travail 1 ***			
↓	C1-T	C2-T	C3
	Facteur1	Facteur2	Valeur
1	Niveau 1	Niveau 1	4
2	Niveau 1	Niveau 2	2
3	Niveau 2	Niveau 1	4
4	Niveau 2	Niveau 2	2

ou autrement présenté:

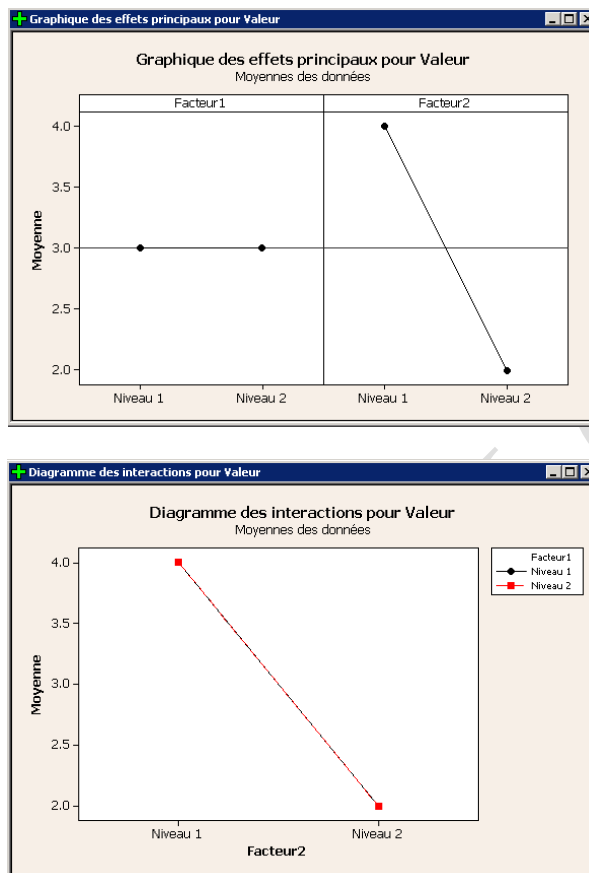
		Facteur 2	
Facteur 1	Niveau 1	Niveau 2	
Niveau 1	4	2	

Niveau 2

4

2

Nous pouvons observer que Facteur 2 a une influence mais pas le Facteur 1. Mais voyons aussi cela sous forme graphique:



Nous observons bien sur le graphique du dessus que le Facteur 1 n'a aucune influence. Sur le graphique de droite c'est moins évident (car il faudrait inverser les catégories et séries du graphique pour avoir le même graphique qu'avant) mais la superposition des deux droites montre que le Facteur 1 n'a pas d'influence. Nous disons alors qu'il y a **(a priori) effet principal (localement) du Facteur2 et en absence d'interactions entre les facteurs.**

Considérons maintenant le tableau suivant:

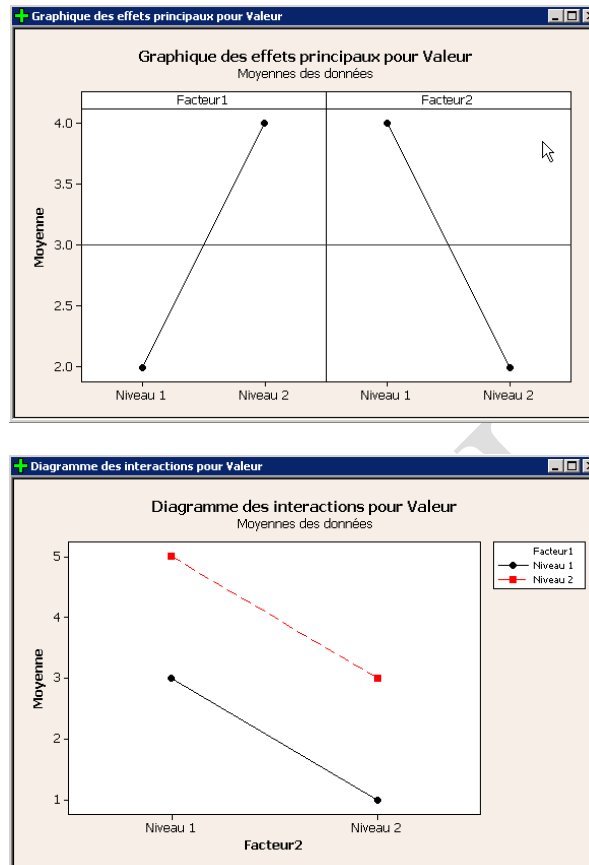
Feuille de travail 1 ***			
↓	C1-T	C2-T	C3
	Facteur1	Facteur2	Valeur
1	Niveau 1	Niveau 1	3
2	Niveau 1	Niveau 2	1
3	Niveau 2	Niveau 1	5
4	Niveau 2	Niveau 2	3

ou autrement présenté:

Facteur 1
Niveau 1
Facteur 2
Niveau 2

Niveau 1	3	1
Niveau 2	5	3

Nous voyons que les deux facteurs ont une influence sur la réponse. Ce que montrent bien les deux graphiques ci-dessous:



Nous observons bien sur le graphique du dessus que le Facteur 1 a une influence sur la réponse et qu'il est de même du Facteur 2 (et en plus de la même amplitude quel que soit le sens!). Sur le graphique de droite, c'est moins évident, mais la même conclusion est valable. Nous disons alors que les **(a priori) deux facteurs sont (localement) significatifs et sans interactions.**

Ici il n'est pas facile de se rendre compte qu'il n'y a pas d'interactions. Il faut avoir repéré la méthode: le passage d'un niveau de facteur à un autre implique toujours une variation identique (de 2 dans le cas présent).

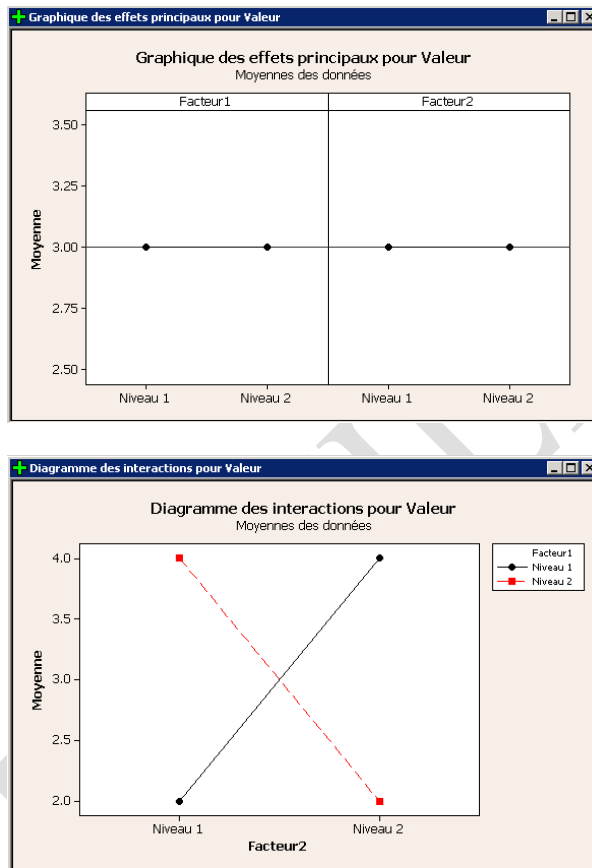
Considérons maintenant le tableau suivant:

Feuille de travail 1 ***			
↓	C1-T	C2-T	C3
	Facteur1	Facteur2	Valeur
1	Niveau 1	Niveau 1	2
2	Niveau 1	Niveau 2	4
3	Niveau 2	Niveau 1	4
4	Niveau 2	Niveau 2	2

ou autrement présenté:

		Facteur 2	
Facteur 1	Niveau 1		Niveau 2
Niveau 1		2	4
Niveau 2		4	2

qui sous cette forme n'est pas triviale à interpréter. Mais avec les graphique on a tout de suite des informations plus pertinentes:



Nous observons bien sur le graphique du dessus, qu'aucun des facteurs n'a d'influence sur la réponse a priori (même graphique qu'au tout début avec la même moyenne). Le graphique de droite nous donne une information complémentaire par contre!!!: Les facteurs ont une influence croisée et comme cette influence croisée est de même amplitude, les effets s'annulent. Nous disons alors que les **deux facteurs sont (localement) en interaction F1\*F2**.

Considérons maintenant le tableau suivant:

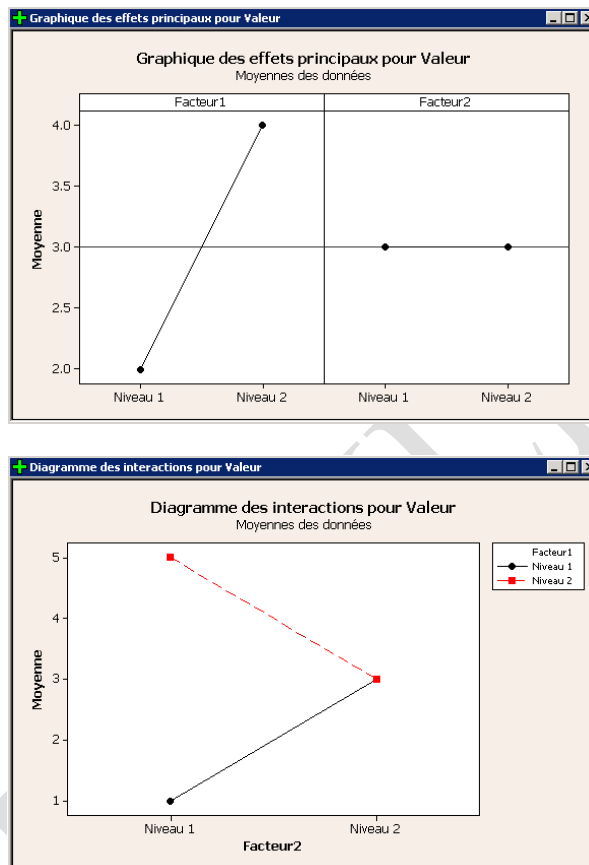
Feuille de travail 1 ***			
↓	C1-T	C2-T	C3
	Facteur1	Facteur2	Valeur
1	Niveau 1	Niveau 1	1
2	Niveau 1	Niveau 2	3
3	Niveau 2	Niveau 1	5
4	Niveau 2	Niveau 2	3



ou autrement présenté:

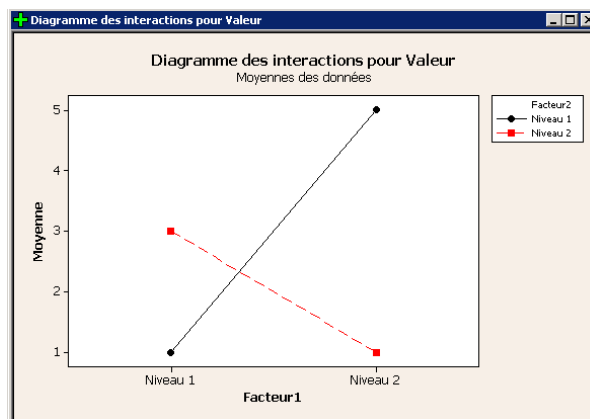
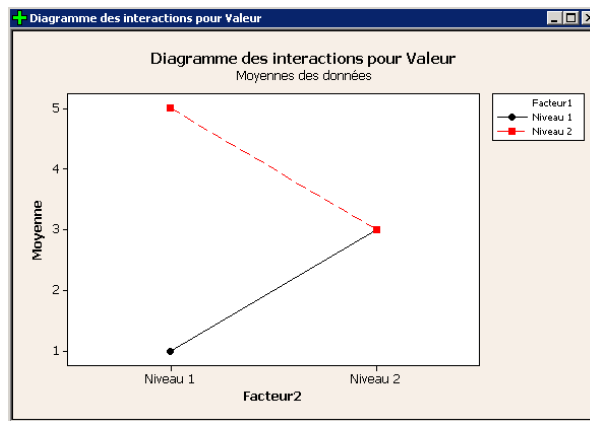
Facteur 1	Facteur 2	
	Niveau 1	Niveau 2
Niveau 1	1	3
Niveau 2	5	3

Ce qui nous donne les deux graphiques suivants:



Nous observons bien sur le graphique du dessus que le Facteur 1 semble avoir une influence et que le Facteur 2 non (en moyenne!). Le graphique de droite nous donne lui aussi, encore une fois, une information complémentaire!!!: C'est que les facteurs sont en interaction. Mais grâce au graphique de dessus, nous pouvons observer que c'est le Facteur 1 qui a une influence significative dans le comportement de la réponse!

Par contre, si l'interaction n'est pas évidente, rien ne vous empêche de faire le graphique de droite ci-dessus, en inversant les séries et catégories. Ce qui donnera dès lors en mettant l'original et le nouveau côte à côte:



Nous disons alors que nous avons **(a priori) deux facteurs (localement) en interaction F1\*F2 où l'influence du Facteur 1 est significative.**

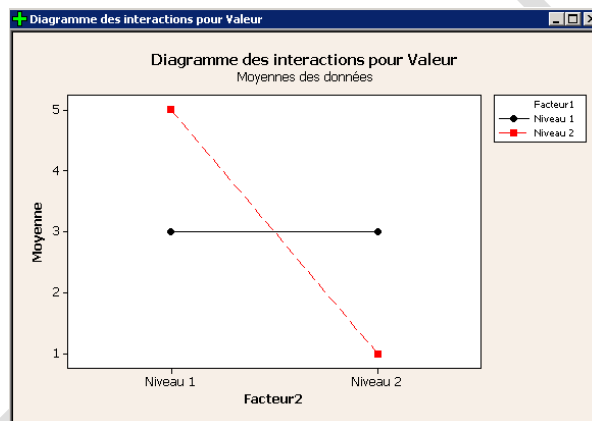
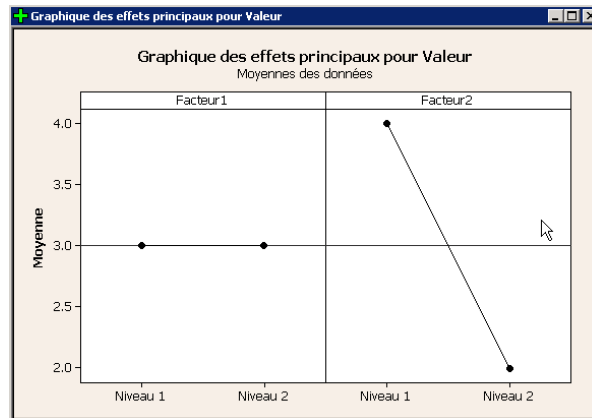
Considérons maintenant le tableau suivant:

Feuille de travail 1 ***			
↓	C1-T	C2-T	C3
	Facteur1	Facteur2	Valeur
1	Niveau 1	Niveau 1	3
2	Niveau 1	Niveau 2	3
3	Niveau 2	Niveau 1	5
4	Niveau 2	Niveau 2	1

ou autrement présenté:

		Facteur 2	
Facteur 1		Niveau 1	Niveau 2
Niveau 1		3	3
Niveau 2		5	1

Ce qui nous donne les deux graphiques suivants:



Nous disons alors que nous avons **(a priori) deux facteurs (localement) en interaction F1\*F2 où l'influence du Facteur 2 est significative.**

Considérons maintenant le tableau suivant:

Feuille de travail 1 ***			
↓	C1-T	C2-T	C3
	Facteur1	Facteur2	Valeur
1	Niveau 1	Niveau 1	1
2	Niveau 1	Niveau 2	1
3	Niveau 2	Niveau 1	5
4	Niveau 2	Niveau 2	1

ou autrement présenté:

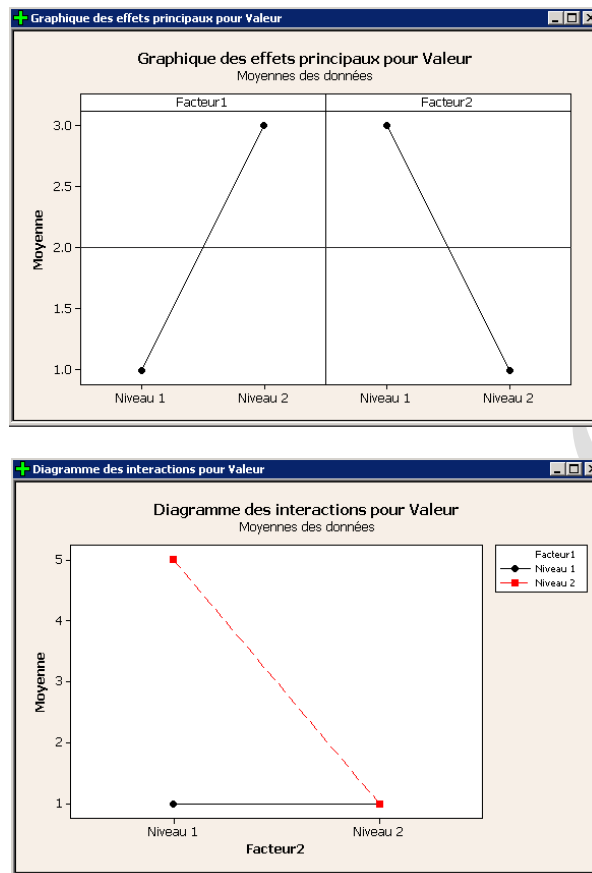
		Facteur 2	
Facteur 1		Niveau 1	Niveau 2
Niveau 1		1	1

Niveau 2

5

1

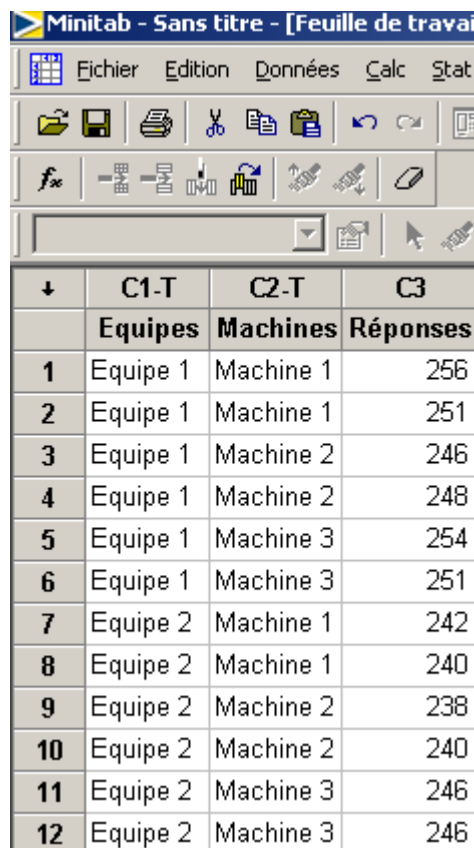
Ce qui nous donne les deux graphiques suivants:



Nous disons alors que nous avons **(a priori) deux facteurs (localement) en interaction F1\*F2 où l'influence des deux facteurs est significative.**

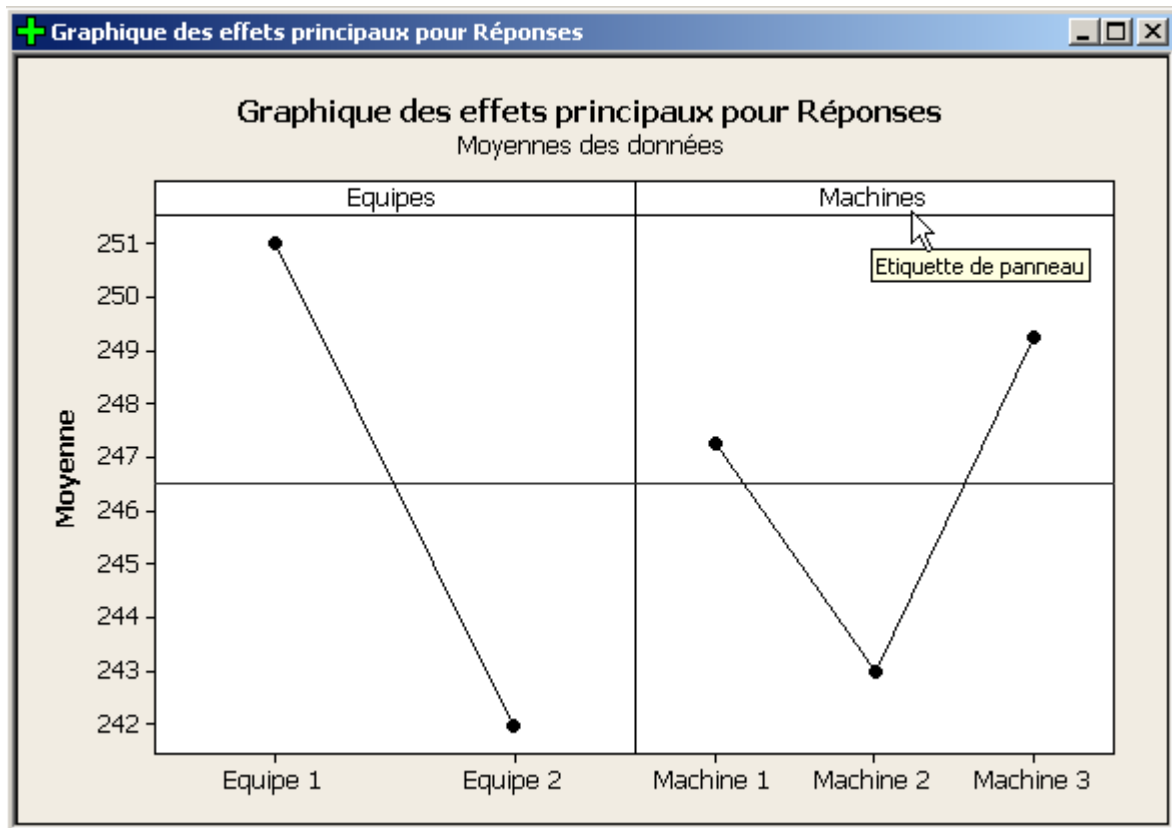
### 12.2.2. Méthode liées aux options Qualité de Minitab

Il existe une autre manière dans Minitab de visualiser les interactions et les effets principaux et jouant avec les données ci-dessous:

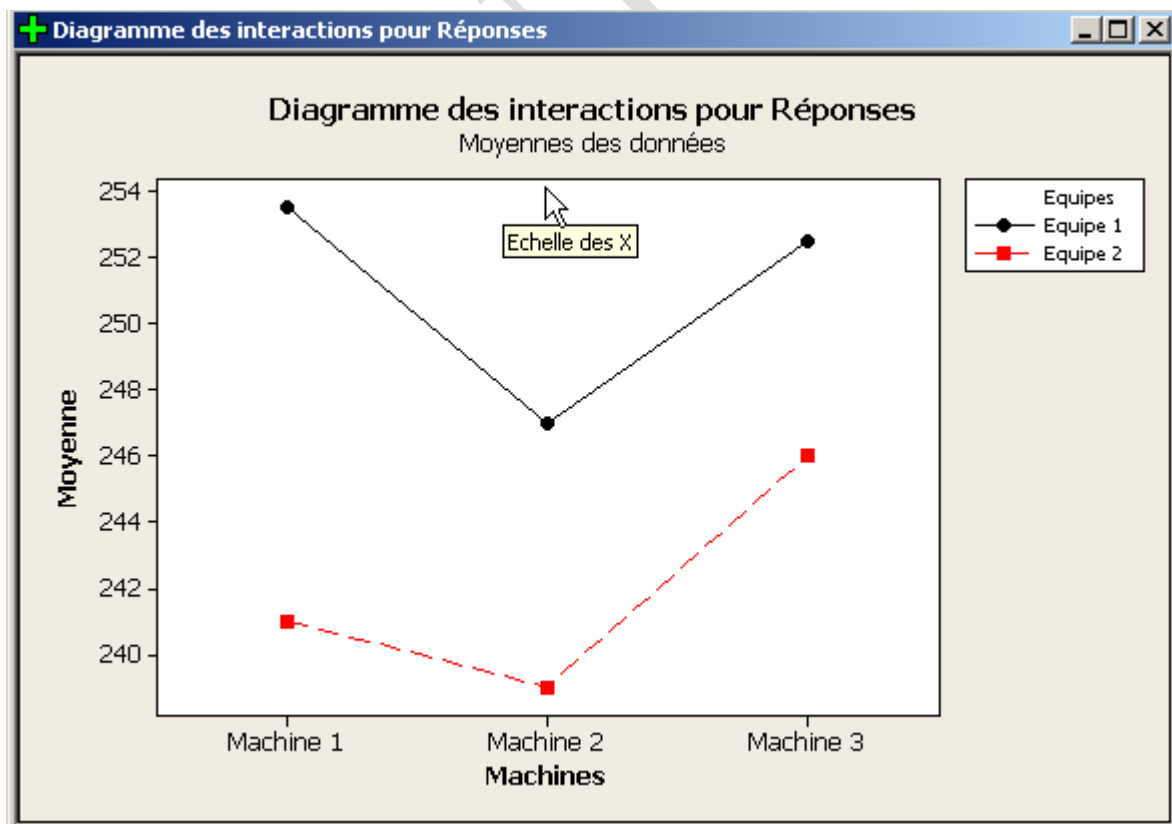


	C1-T	C2-T	C3
	Equipes	Machines	Réponses
1	Equipe 1	Machine 1	256
2	Equipe 1	Machine 1	251
3	Equipe 1	Machine 2	246
4	Equipe 1	Machine 2	248
5	Equipe 1	Machine 3	254
6	Equipe 1	Machine 3	251
7	Equipe 2	Machine 1	242
8	Equipe 2	Machine 1	240
9	Equipe 2	Machine 2	238
10	Equipe 2	Machine 2	240
11	Equipe 2	Machine 3	246
12	Equipe 2	Machine 3	246

Les graphiques vus précédemment donnent pour les facteurs principaux:

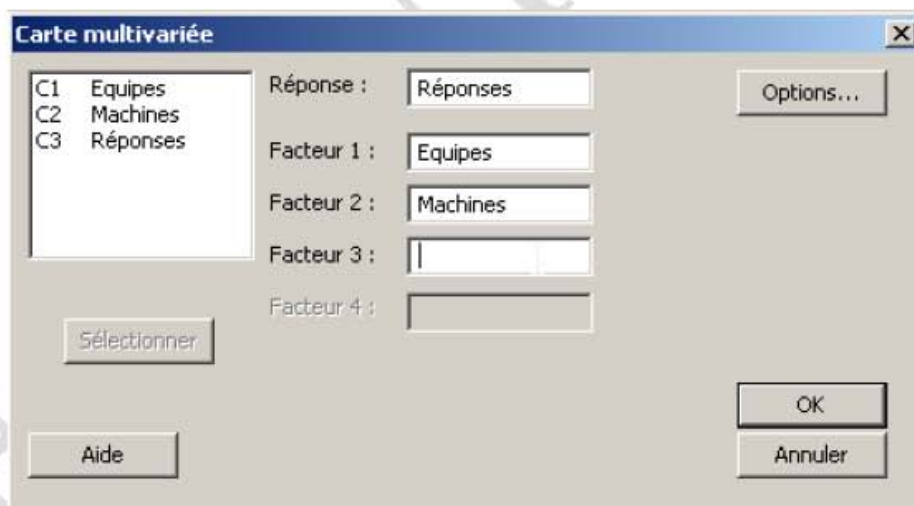
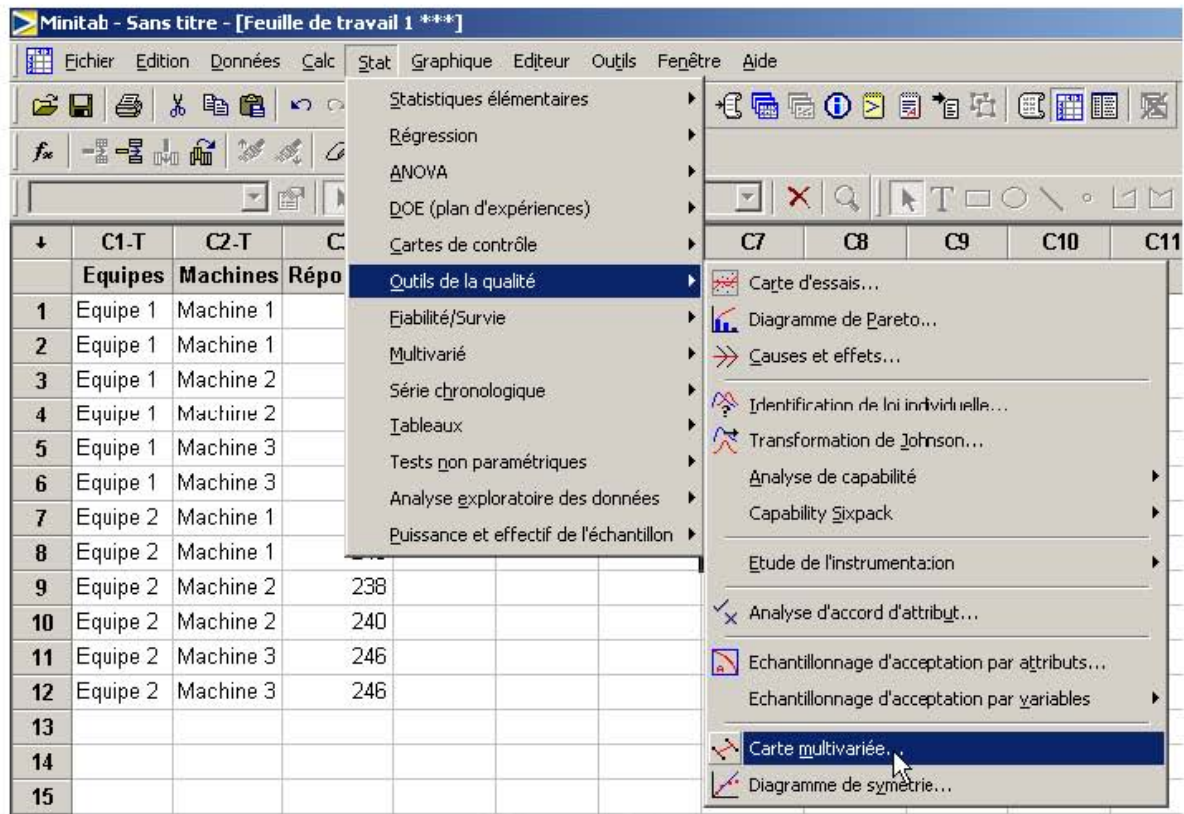


ainsi que pour les interactions:

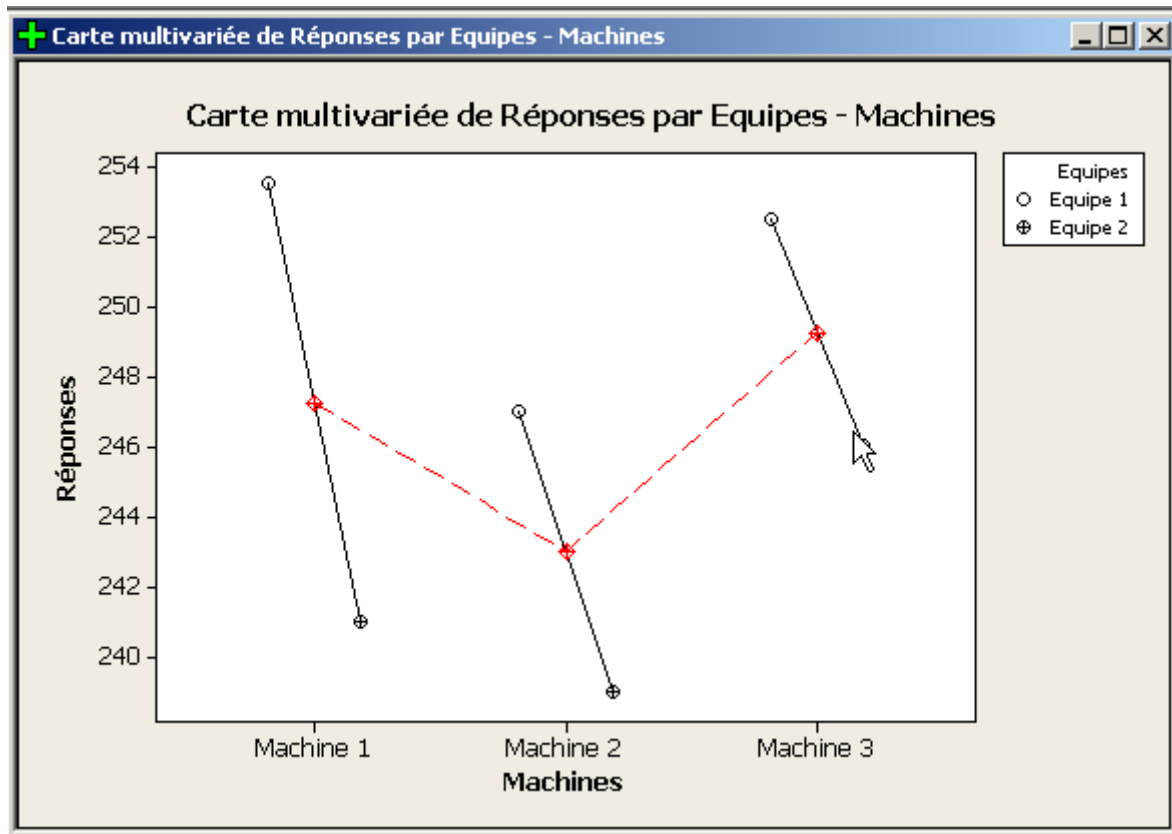


Il y a donc clairement pas d'interaction puisque pas de croisements...

Voyons maintenant l'autre manière de visualiser ceci en allant dans le menu **Stat/Outils de la qualité/Carte multivariée...**:



Ce qui donne:



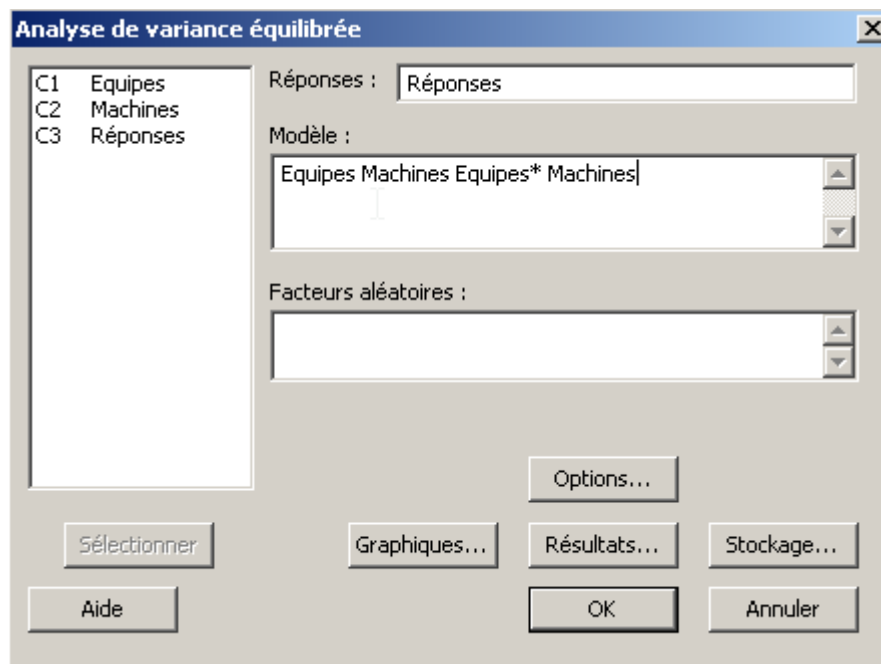
et rien dans la fenêtre de session. Comme nous pouvons le voir le graph ci-dessus est juste un visuel alternatif au graphe précédent.

Nous pouvons vérifier numériquement et statistiquement la présence d'interactions en allant dans le menu **Stat/ANOVA/ANOVA équilibrée...**:

	C1-T	C2-T	C
	Equipes	Machines	Répo
1	Equipe 1	Machine 1	
2	Equipe 1	Machine 1	
3	Equipe 1	Machine 2	
4	Equipe 1	Machine 2	
5	Equipe 1	Machine 3	
6	Equipe 1	Machine 3	
7	Equipe 2	Machine 1	
8	Equipe 2	Machine 1	
9	Equipe 2	Machine 2	238
10	Equipe 2	Machine 2	240
11	Equipe 2	Machine 3	246
12	Equipe 2	Machine 3	246



Et en mettant:



Si nous validons par **OK**, nous avons alors:

Analyse de la variance pour Réponses

Source	DL	Somme des carrés	CM	F	P
Equipes	1	243.000	243.000	63.39	0.000
Machines	2	81.500	40.750	10.63	0.011
Equipes*Machines	2	19.500	9.750	2.54	0.158
Erreur	6	23.000	3.833		
Total	11	367.000			

Donc l'interaction n'est pas significative.

## 12.3. Exercice 85.: Analyse de la variance à un facteur fixe (ANOVA-1) désempilé

Minitab® Statistical Software 15.1.1.0

Imaginons une entreprise faisant les trois huit. Nous avons trois équipes qui travaillent sur une même machine. Nous souhaitons vérifier avec un seuil de confiance de 95% s'il y a une différence de productivité moyenne entre les trois équipes sur une semaine de travail.

Ouvrez le fichier *ANOVA-DESEMPILE.mpj*:

The screenshot shows the Minitab interface. The Session window displays the following text:

```

Récupération du projet du fichier 'C:\Documents and
Settings\Administrateur\Bureau\ANOVA.MPJ'

----- 23.02.2012 10:24:14 -----

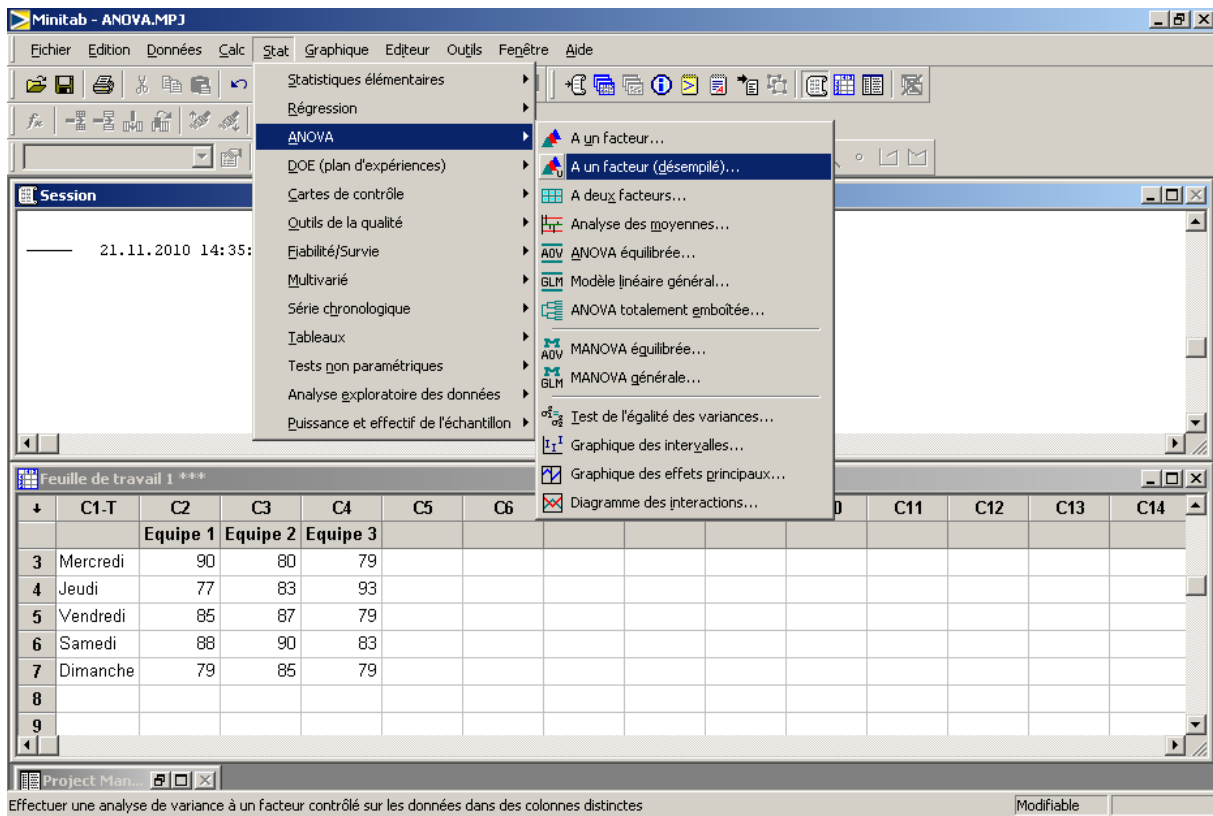
Bienvenue dans Minitab, appuyez sur F1 pour obtenir de l'aide.
Récupération du projet du fichier 'E:\ANOVADESEMPILE.MPJ'

```

The Worksheet window shows a data table with the following structure:

	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
		Equipe 1	Equipe 2	Equipe 3										
1	Lundi	78	77	88										
2	Mardi	88	75	86										
3	Mercredi	90	80	79										
4	Jeudi	77	83	93										
5	Vendredi	85	87	79										
6	Samedi	88	90	83										
7	Dimanche	79	85	79										

Et allez dans le menu **Stat/ANOVA/A un facteur (désempilé)...**:



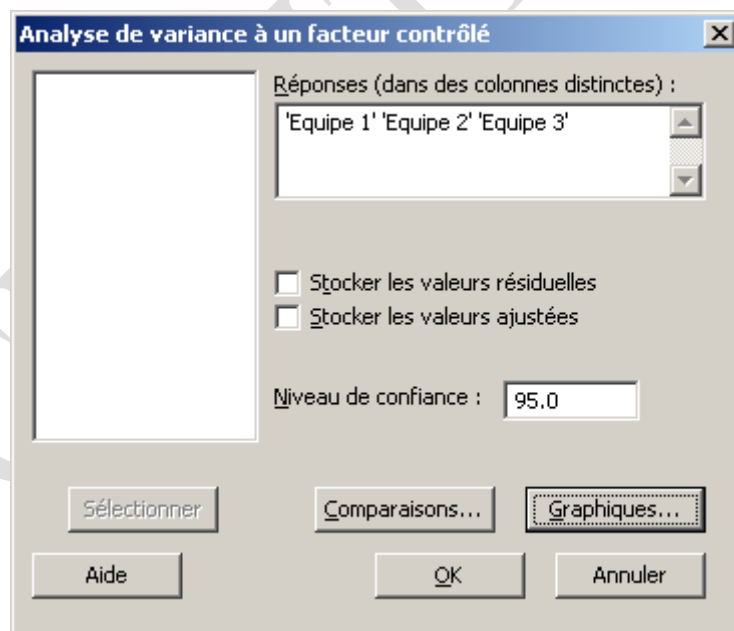
Statistiques élémentaires  
Régression  
ANOVA  
DOE (plan d'expériences)  
Cartes de contrôle  
Outils de la qualité  
Fiabilité/Survie  
Multivarié  
Série chronologique  
Tableaux  
Tests non paramétriques  
Analyse exploratoire des données  
Puissance et effectif de l'échantillon

A un facteur...  
A un facteur (déséquilibré)...  
A deux facteurs...  
Analyse des moyennes...  
ADV ANOVA équilibrée...  
GLM Modèle linéaire général...  
ANOVA totalement emboîtée...  
MANOVA équilibrée...  
GLM MANOVA générale...  
Test de l'égalité des variances...  
Graphique des intervalles...  
Graphique des effets principaux...  
Diagramme des interactions...

	C1-T	C2	C3	C4	C5	C6		C11	C12	C13	C14
		Equipe 1	Equipe 2	Equipe 3							
3	Mercredi	90	80	79							
4	Jeudi	77	83	93							
5	Vendredi	85	87	79							
6	Samedi	88	90	83							
7	Dimanche	79	85	79							
8											
9											

Effectuer une analyse de variance à un facteur contrôlé sur les données dans des colonnes distinctes

Nous avons alors:



Analyse de variance à un facteur contrôlé

Réponses (dans des colonnes distinctes) :

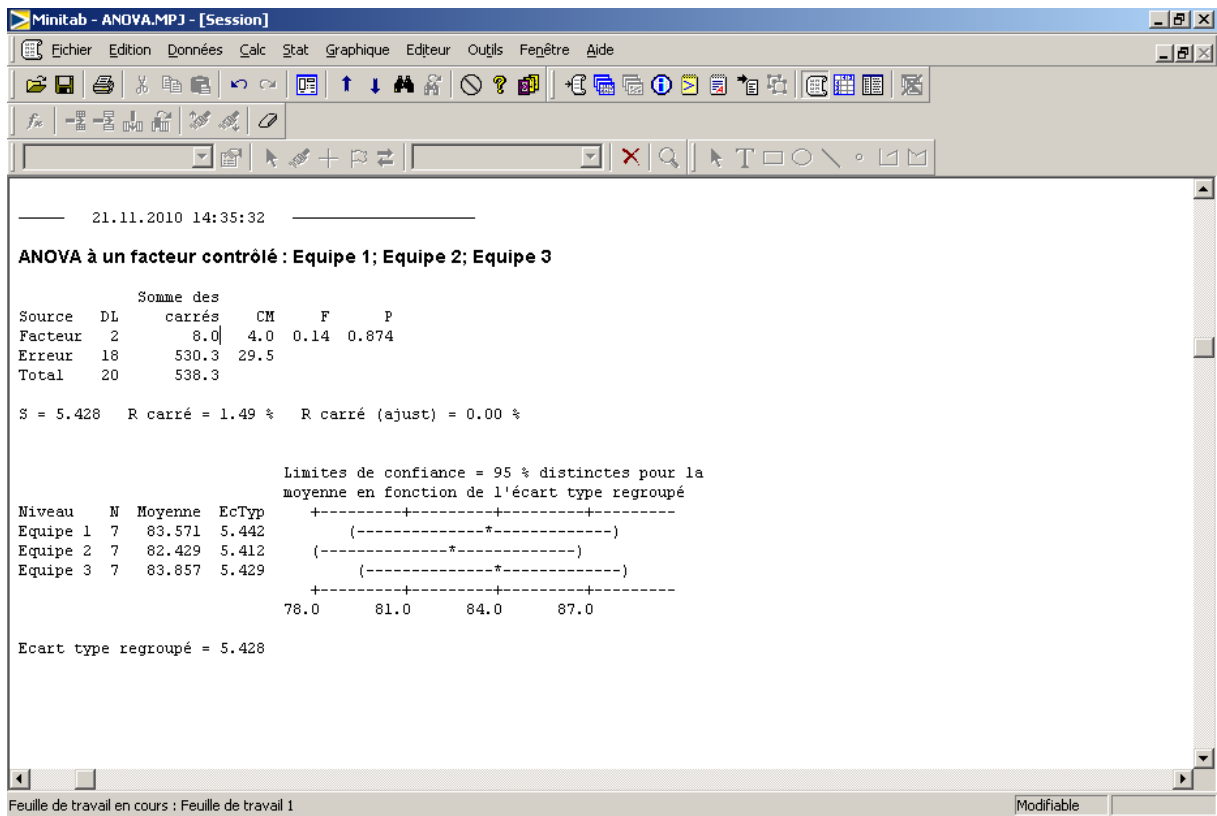
'Equipe 1' 'Equipe 2' 'Equipe 3'

Stocker les valeurs résiduelles  
 Stocker les valeurs ajustées

Niveau de confiance : 95.0

Sélectionner Comparaisons... Graphiques...  
Aide OK Annuler

Nous validons par **OK** ce qui donne dans la fenêtre de session:



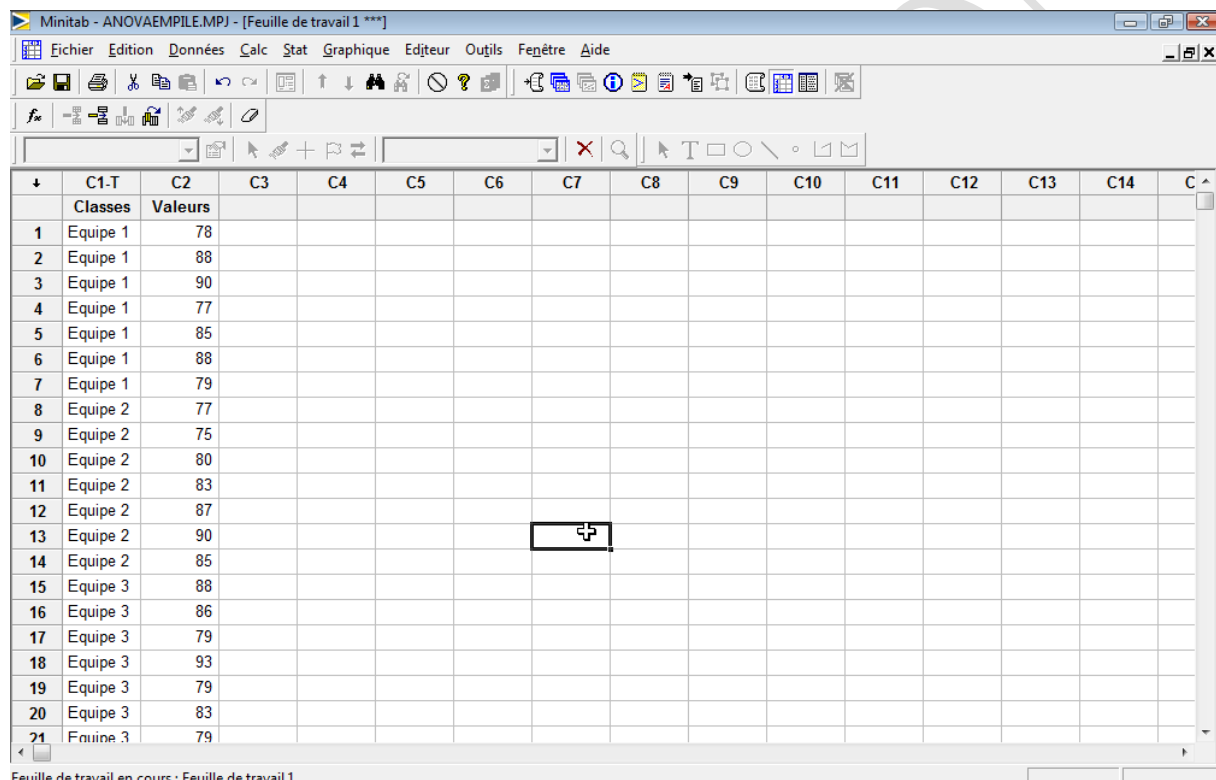
Nous avons donc les mêmes résultats avec les mêmes conclusions que dans Microsoft Excel (mais un peu moins précis).

## 12.4. Exercice 86.: Analyse de la variance à un facteur fixe (ANOVA-1) empilé

Minitab® Statistical Software 15.1.1.0

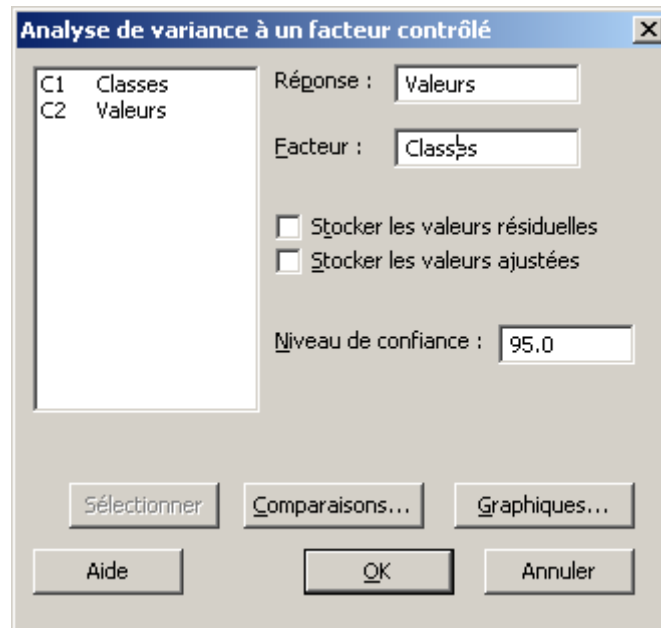
Imaginons une entreprise faisant les trois huit. Nous avons trois équipes qui travaillent sur une même machine. Nous souhaitons vérifier avec un seuil de confiance de 95% s'il y a une différence de productivité moyenne entre les trois équipes sur une semaine de travail.

Ouvrez le fichier *ANOVA-EMPILE.mpj* (il s'agit du même fichier que précédemment mais les données sont empilées):

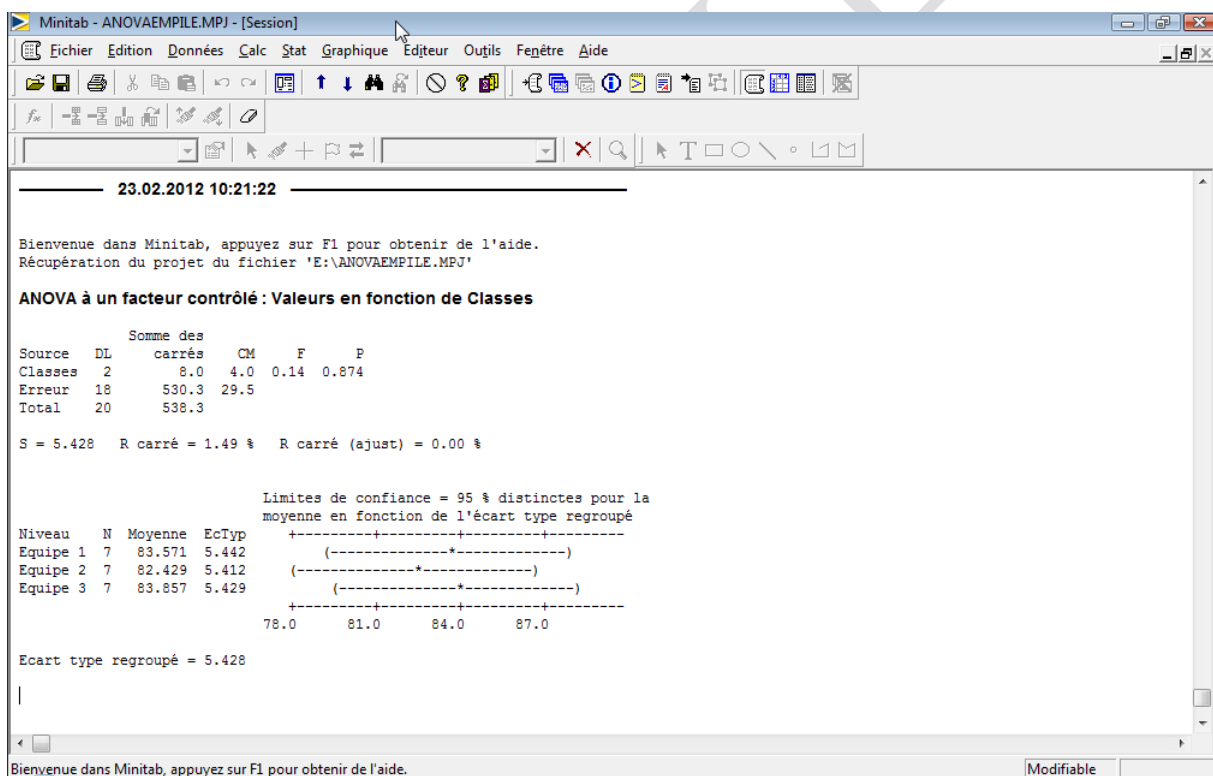


	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C ^
	Classes	Valeurs													
1	Equipe 1	78													
2	Equipe 1	88													
3	Equipe 1	90													
4	Equipe 1	77													
5	Equipe 1	85													
6	Equipe 1	88													
7	Equipe 1	79													
8	Equipe 2	77													
9	Equipe 2	75													
10	Equipe 2	80													
11	Equipe 2	83													
12	Equipe 2	87													
13	Equipe 2	90													
14	Equipe 2	85													
15	Equipe 3	88													
16	Equipe 3	86													
17	Equipe 3	79													
18	Equipe 3	93													
19	Equipe 3	79													
20	Equipe 3	83													
21	Equipe 3	79													

Nous allons dans le menu **Stat/ANOVA/A un facteur...**:



Ce qui donne au final le même résultat que l'ANOVA précédente (ouff!):



## 12.5. Exercice 87.: Test de Tukey

Minitab® Statistical Software 15.1.1.0

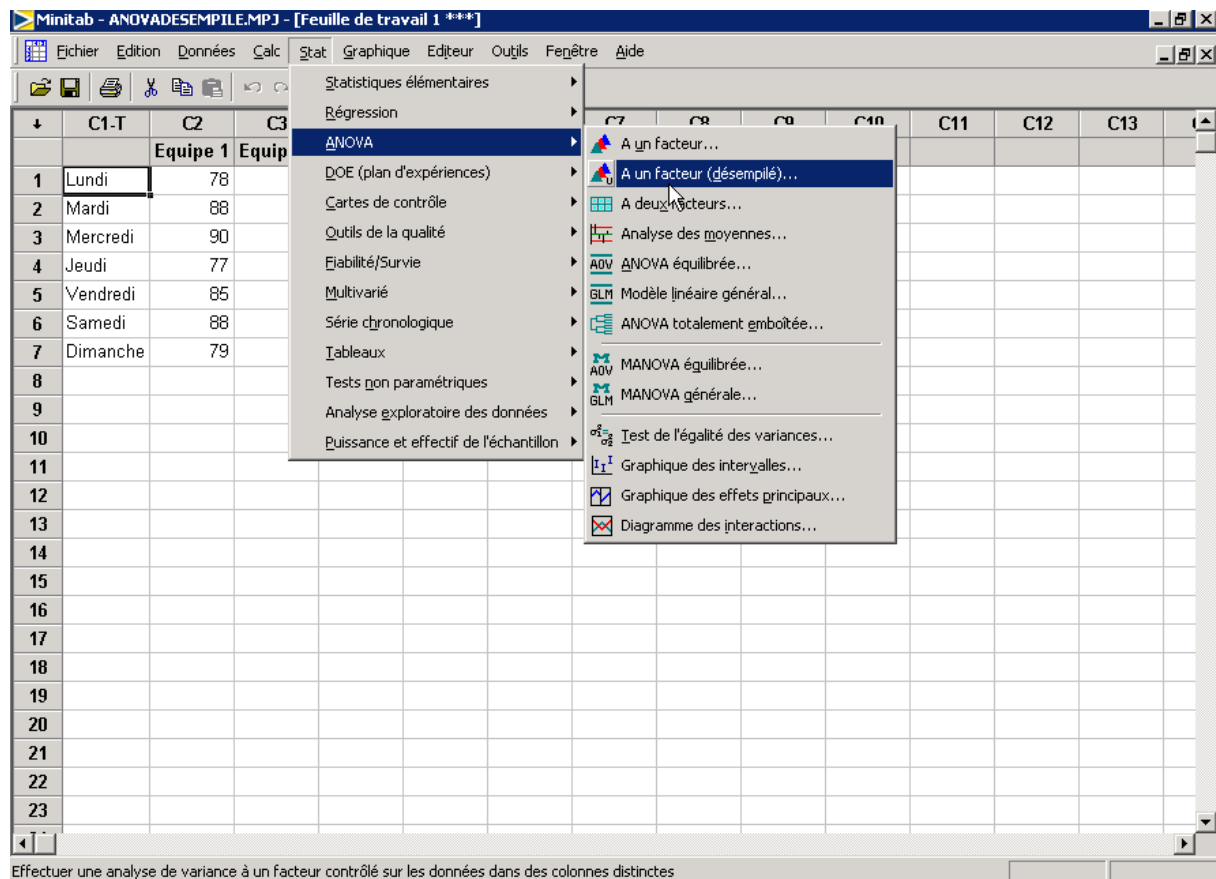
Ici, nous allons vérifier si les concepts démontrés dans le cours de statistique théorique concernant le test de Tukey sont appliqués de façon identique dans Minitab.

Nous n'avons pas en classe simulé les distributions de l'étendue studentisée plus par flemme que par manque de temps (les simulations de Monte-Carlo se basant toujours sur le même principe...), donc il n'y aura pas de comparaison par rapport avec Microsoft Excel. Nous ferons ici juste une vérification des résultats renvoyés par Minitab en utilisant les tables des étendues studentisées.

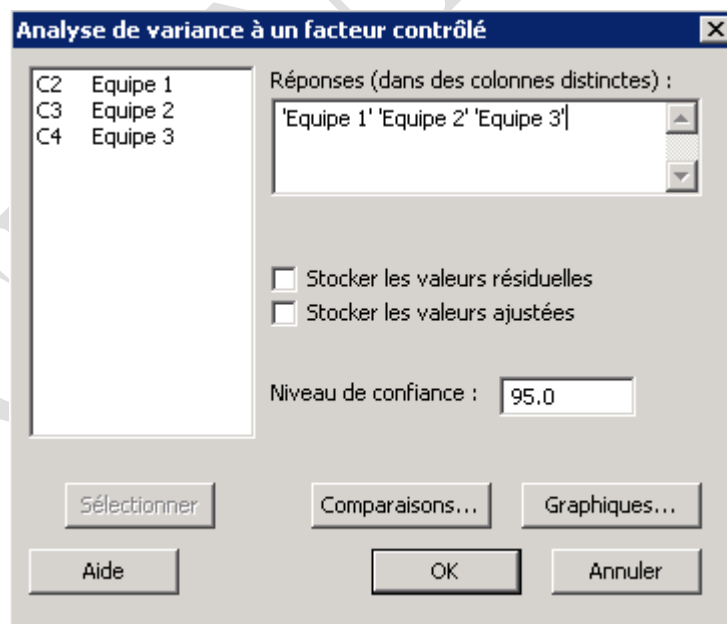
Nous partons toujours du même tableau que lors de notre étude de l'ANOVA à un facteur contrôlé et de la puissance du test de l'ANOVA.

	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
		Equipe 1	Equipe 2	Equipe 3									
1	Lundi	78	77	88									
2	Mardi	88	75	86									
3	Mercredi	90	80	79									
4	Jeudi	77	83	93									
5	Vendredi	85	87	79									
6	Samedi	88	90	83									
7	Dimanche	79	85	79									
8													
9													
10													
11													
12													
13													
14													
15													
16													
17													
18													
19													
20													
21													
22													
23													

Ensuite, nous allons dans le menu **Stat/ANOVA/A un facteur (désempilé)...**:

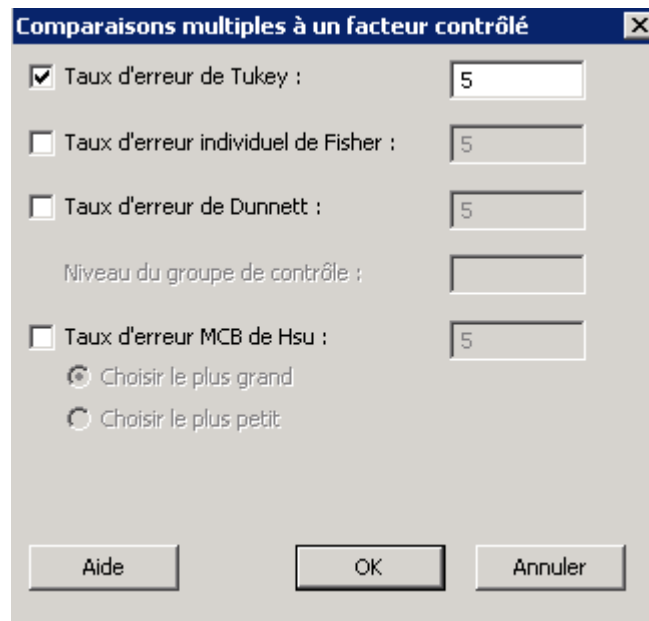


et nous y mettons:



Nous cliquons sur le bouton **Comparaisons...**:



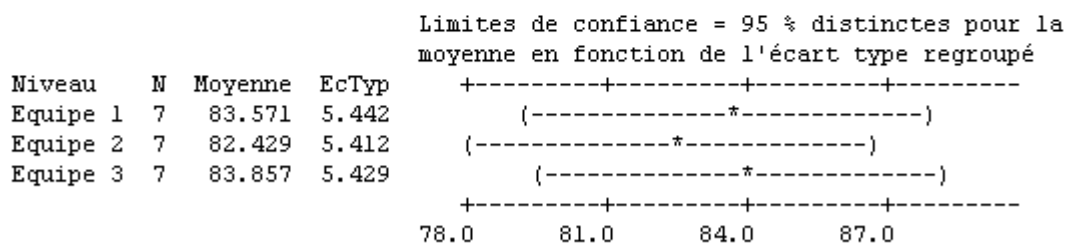


et nous activons le **Taux d'erreur de Tukey** à un seuil de 5% seulement (le **Taux d'erreur individuel de Fisher** étant simplement le test de Fisher vu plus haut appliqué par paire d'échantillons). Nous validons tout ce beau monde par **OK** pour obtenir d'abord la première partie classique:

**ANOVA à un facteur contrôlé : Equipe 1; Equipe 2; Equipe 3**

Source	DL	Somme des carrés	CM	F	P
Facteur	2	8.0	4.0	0.14	0.874
Erreur	18	530.3	29.5		
Total	20	538.3			

S = 5.428 R carré = 1.49 % R carré (ajust) = 0.00 %



Ecart type regroupé = 5.428

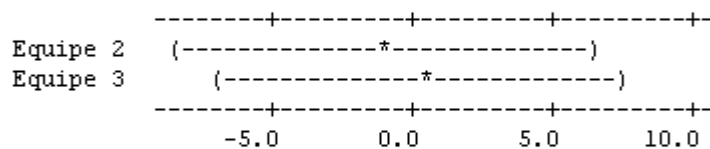
Bon évidemment nous acceptons l'hypothèse nulle mais regardons quand même la suite de la fenêtre de session qui nous donne les informations concernant le test de Tukey (puisque la puissance du test est médiocre).

Intervalles de confiance simultanés de Tukey = 95 %  
Toutes les comparaisons deux à deux

Niveau de confiance individuel = 98.00 %

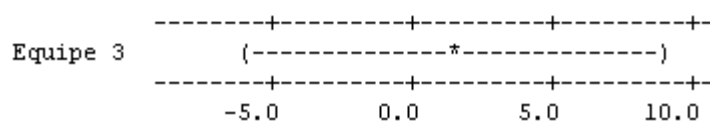
Equipe 1 soustrait de :

	Inférieur	Centré	Supérieur
Equipe 2	-8.549	-1.143	6.263
Equipe 3	-7.120	0.286	7.692



Equipe 2 soustrait de :

	Inférieur	Centré	Supérieur
Equipe 3	-5.977	1.429	8.834



Bon c'est un peu basique comme compte rendu mais maintenant vérifions si nous sommes à peu près en conformité avec ce qui a été vu dans le cours de statistique théorique et en étant un peu plus explicite.

Bon d'abord comme nous avons 3 échantillons, ne pouvons faire que les comparaisons par paires suivantes (1,2);(1,3) ce qui correspond au premier groupe Minitab et de la paire (2,3) qui correspond au deuxième groupe Minitab ci-dessus.

Bon d'abord, nous allons repérer dans les tables ce qui nous intéresse (bon ici c'est à 5% plutôt que 2%...):

$dI$	2	3
1	17,9688	26,9755
2	6,0849	8,3308
3	4,5007	5,9096
4	3,9265	5,0403
5	3,6354	4,6017
6	3,4605	4,3390
7	3,3439	4,1648
8	3,2612	4,0410
9	3,1991	3,9485
10	3,1511	3,8768
11	3,1127	3,8195
12	3,0813	3,7728
13	3,0553	3,7341
14	3,0332	3,7014
15	3,0143	3,6734
16	2,9980	3,6491
17	2,9837	3,6280
18	2,9712	3,6093
19	2,9600	3,5927
20	2,9500	3,5779

[http://davidmlane.com/hyperstat/sr\\_table.html](http://davidmlane.com/hyperstat/sr_table.html)

Nous avons alors conformément à ce qui a été vu dans le cours théorique (donc nous savons d'avance que les intervalles que nous allons obtenir à 5% seront plus réduits que ceux de Minitab qui sont à 2%):

$$R_{\mu,crit} = Q_{k,N-k,1-\alpha} \sqrt{\frac{MSE}{n}} = Q_{3,21-3,1-5\%} \sqrt{\frac{29,5}{7}} = 3,6093 \sqrt{\frac{29,5}{7}} \cong 7,4094$$

Il vient alors que:

$$(\bar{x}_i - \bar{x}_j) - R_{\mu,crit} \leq \mu_i - \mu_j \leq (\bar{x}_i - \bar{x}_j) + R_{\mu,crit}$$

donnera respectivement pour le cas le plus extrême (qui est suffisant):

$$\begin{aligned} (\bar{x}_3 - \bar{x}_2) - R_{\mu,crit} &\leq \mu_3 - \mu_2 \leq (\bar{x}_3 - \bar{x}_2) + R_{\mu,crit} \\ (83,857 - 82,429) - 7,4094 &\leq \mu_3 - \mu_2 \leq (83,857 - 82,429) + 7,4094 \\ 1,429 - 7,4094 &\leq \mu_3 - \mu_2 \leq 1,429 + 7,4094 \\ -5,980 &\leq \mu_3 - \mu_2 \leq 8,838 \end{aligned}$$

Donc 0 étant bien dans l'intervalle, nous ne rejetons pas l'hypothèse d'égalité des moyennes. Nous ne sommes pas trop trop loin du résultat de Minitab.

Mais nous avons surtout:

$$\begin{aligned} \max(\bar{x}_i) - \min(\bar{x}_i) &\leq Q_{k,N-k,1-\alpha} \sqrt{\frac{MSE}{n}} \\ 1,428 &\leq 7,4094 \end{aligned}$$

et la même conclusion en découle.

## 12.6. Exercice 88.: Analyse de la variance à deux facteurs fixes (ANOVA-2) sans répétitions

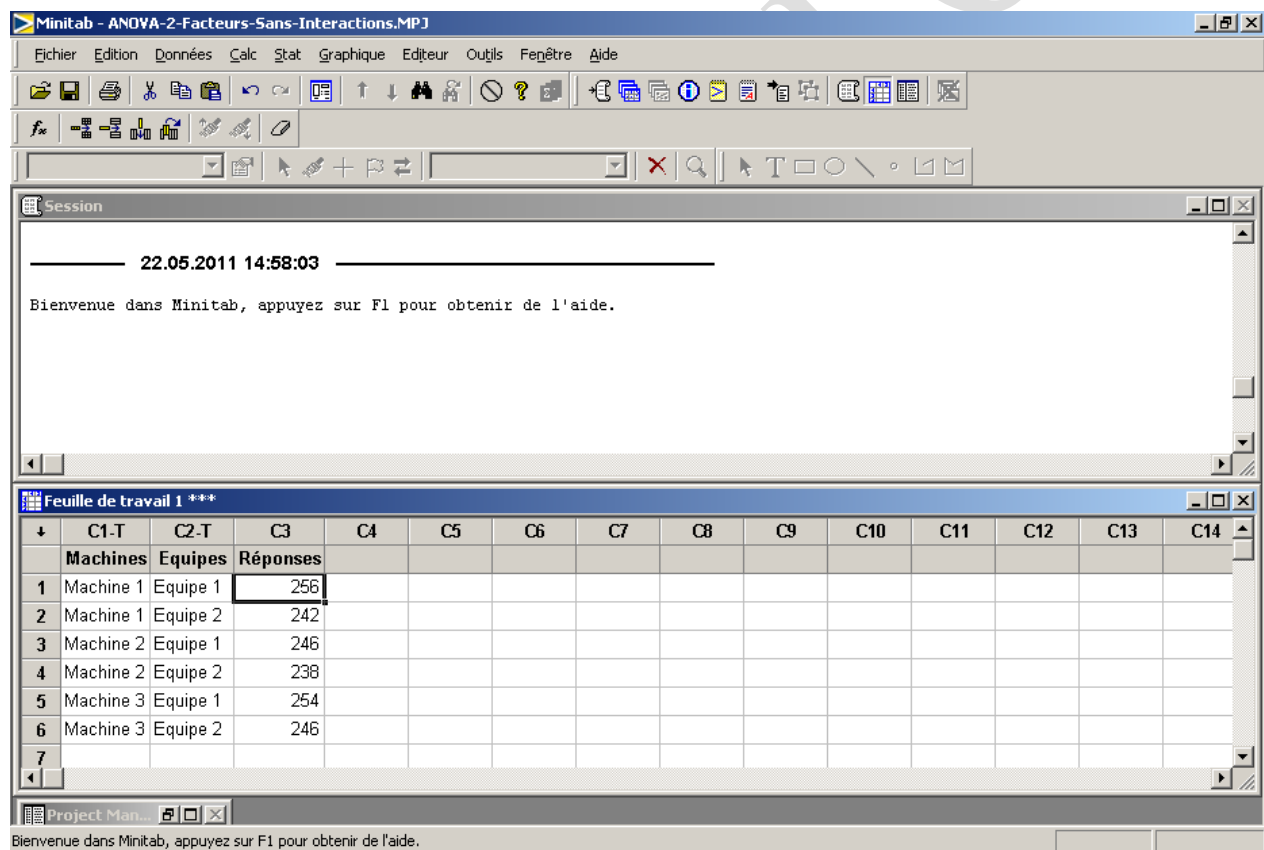
Minitab® Statistical Software 15.1.1.0

### 12.6.1. Minitab 16 et antérieur

Imaginons une entreprise faisant les trois huit. Nous avons trois équipes qui travaillent sur une même machine. Nous souhaitons vérifier avec un seuil de confiance de 95% s'il y a une différence de productivité moyenne entre les trois équipes sur une semaine de travail (hypothèse que les moyennes sont égales).

Remarque: La variable est donc la Machine et sa modalité à trois niveaux!

Ouvrez le fichier *ANOVA-2-Facteurs-Sans-Repetitions.mpj*:



The screenshot shows the Minitab interface with a data table open. The table has columns labeled C1-T, C2-T, C3, C4, C5, C6, C7, C8, C9, C10, C11, C12, C13, and C14. The data is as follows:

	C1-T	C2-T	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Machines	Equipes	Réponses											
1	Machine 1	Equipe 1	256											
2	Machine 1	Equipe 2	242											
3	Machine 2	Equipe 1	246											
4	Machine 2	Equipe 2	238											
5	Machine 3	Equipe 1	254											
6	Machine 3	Equipe 2	246											
7														

Nous allons dans le menu **Stat/ANOVA/A deux facteurs...**:

The screenshot shows the Minitab interface with the 'Stat' menu open and 'ANOVA' selected. The 'ANOVA' submenu is also open, with 'A deux facteurs...' highlighted. Below the menu, a worksheet titled 'Feuille de travail 1 \*\*\*' is visible, containing data for a two-factor ANOVA.

	C1-T	C2-T	C3	C4	C5	C6	10	C11	C12	C13	C14
	Machines	Equipes	Réponses								
1	Machine 1	Equipe 1	256								
2	Machine 1	Equipe 2	242								
3	Machine 2	Equipe 1	246								
4	Machine 2	Equipe 2	238								
5	Machine 3	Equipe 1	254								
6	Machine 3	Equipe 2	246								
7											

Statistiques élémentaires  
Régression  
ANOVA  
DOE (plan d'expériences)  
Cartes de contrôle  
Outils de la qualité  
Fiabilité/Survie  
Multivarié  
Série chronologique  
Tableaux  
Tests non paramétriques  
Analyse exploratoire des données  
Puissance et effectif de l'échantillon

A un facteur...  
A un facteur (déséquilibré)...  
A deux facteurs...  
Analyse des moyennes...  
ANOVA équilibrée...  
Modèle linéaire général...  
ANOVA totalement emboîtée...  
MANOVA équilibrée...  
MANOVA générale...  
Test de l'égalité des variances...  
Graphique des intervalles...  
Graphique des effets principaux...  
Diagramme des interactions...

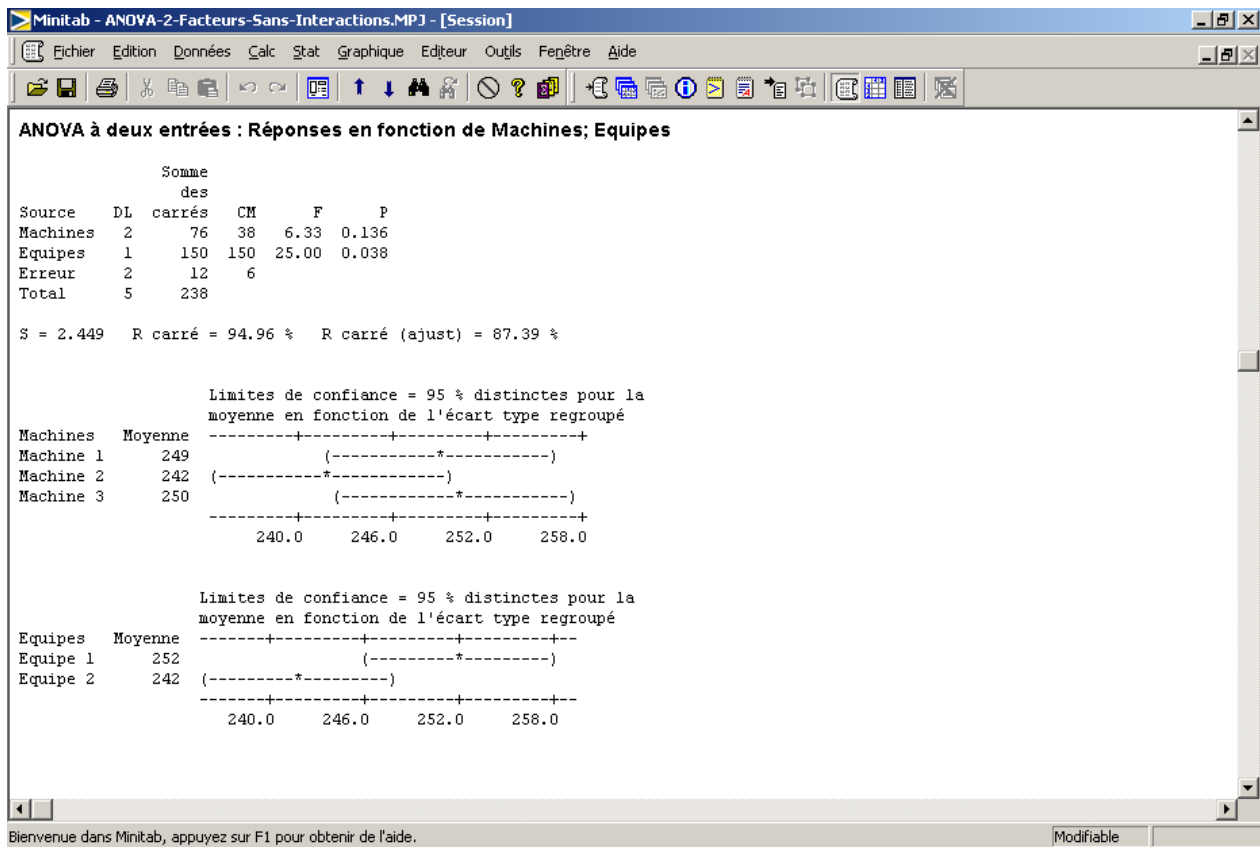
Effectuer une analyse à deux facteurs de la variance pour les données équilibrées

Pour avoir:

The dialog box 'Analyse de variance à deux facteurs contrôlés' is shown with the following settings:

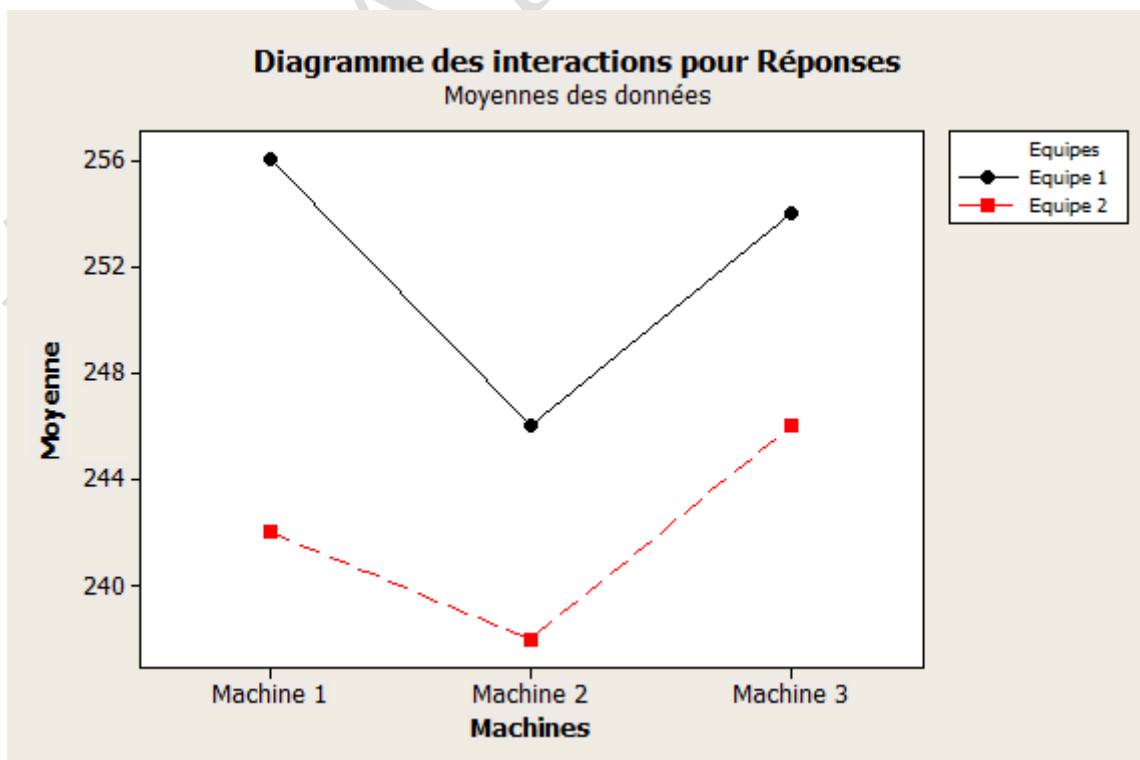
- Réponse : Réponses
- Facteur de ligne : Machines  Afficher les moyennes
- Facteur de colonne : Equipes  Afficher les moyennes
- Stocker les valeurs résiduelles
- Stocker les valeurs ajustées
- Niveau de confiance : 95.0
- Ajuster le modèle additif
- Buttons: Sélectionner, Aide, Graphiques..., OK, Annuler

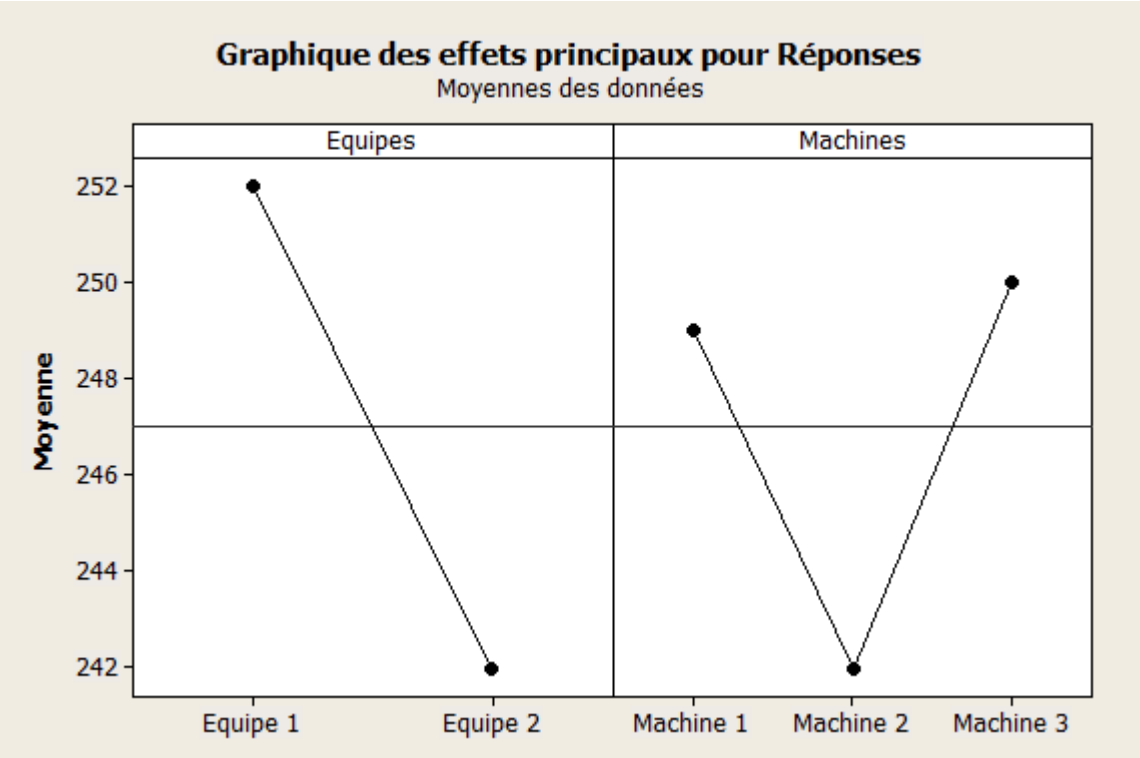
On clique sur **OK**:



Nous obtenons donc les mêmes résultats avec les mêmes conclusions qu'avec Microsoft Excel.

Et pour info, voici les graphiques des interactions et facteurs principaux obtenus en procédant comme montré dans l'exercice relatif à ce sujet:





ÉCHANTILLON

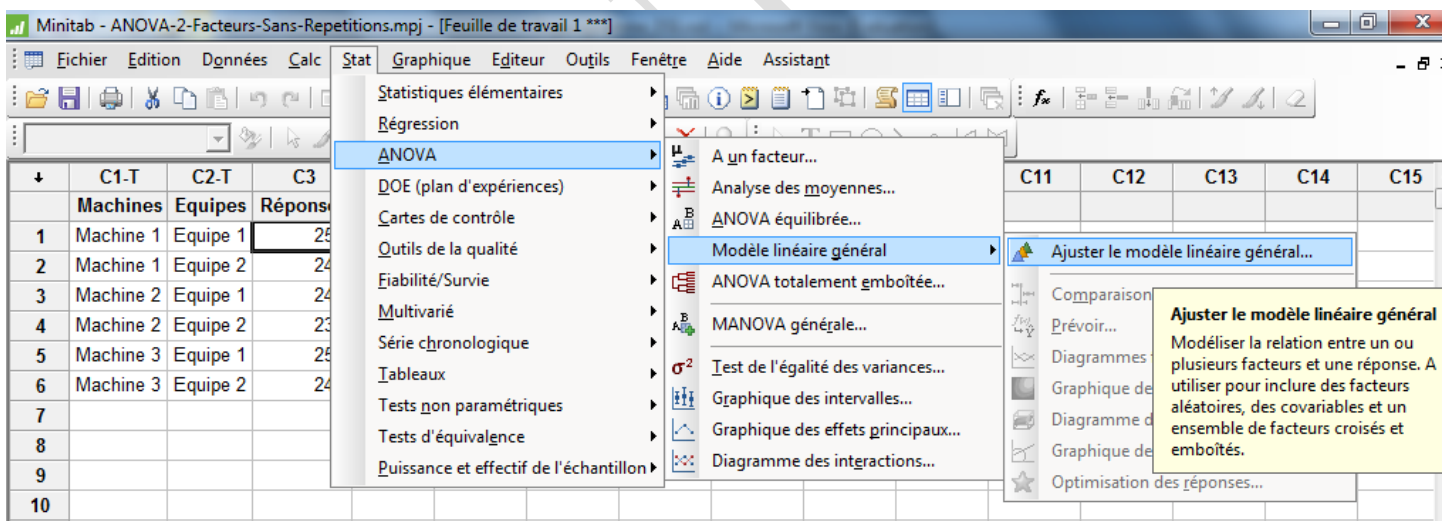
### 12.6.2. Minitab 17

Depuis Minitab 17 la procédure A complément changée...

Voyons cela toujours avec le même fichier:

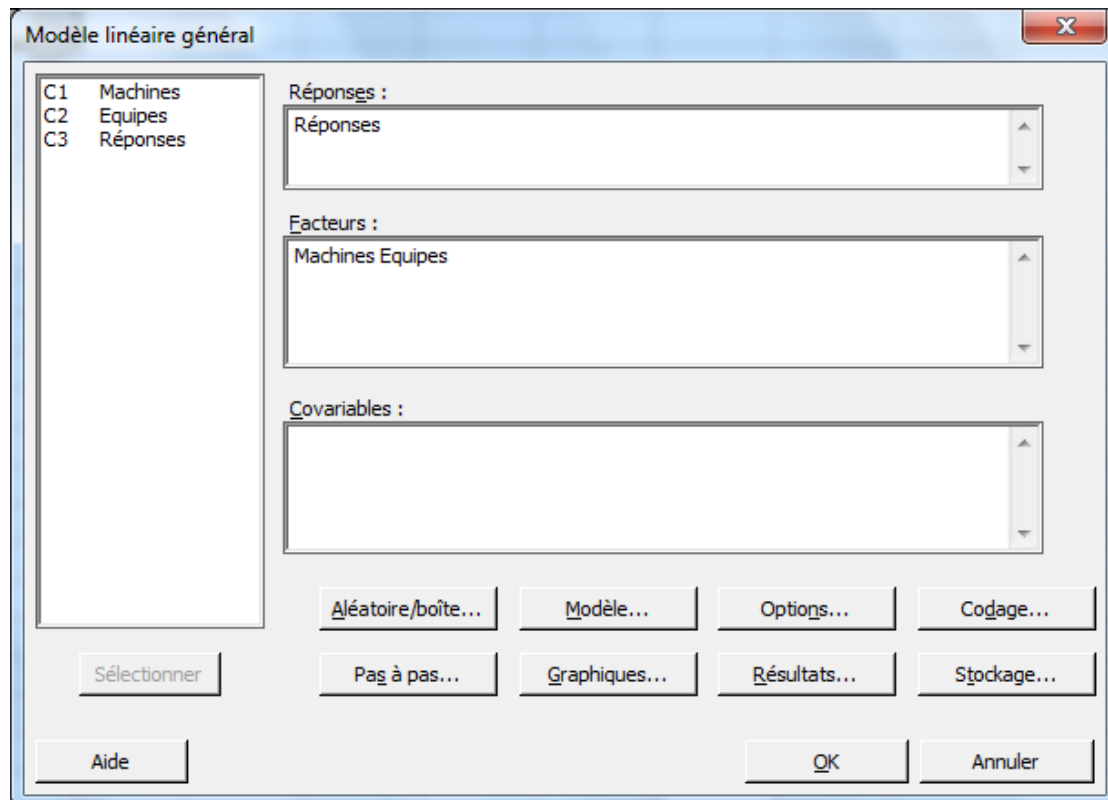
	C1-T	C2-T	C3	C4	C5	C6	C7	C8	C9	C10	C11
	Machines	Equipes	Réponses								
1	Machine 1	Equipe 1	256								
2	Machine 1	Equipe 2	242								
3	Machine 2	Equipe 1	246								
4	Machine 2	Equipe 2	238								
5	Machine 3	Equipe 1	254								
6	Machine 3	Equipe 2	246								
7											
8											
9											
10											
11											
12											

Nous allons ensuite dans le menu **Stat/ANOVA/Modèle linéaire général/Ajuster le modèle linéaire général...**:

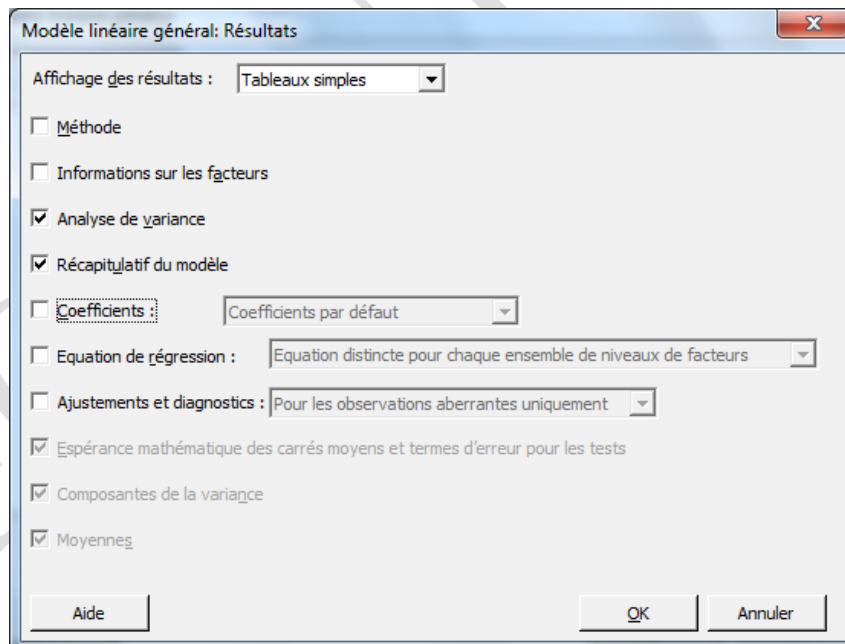


Apparaît alors la boîte de dialogue suivante dans laquelle nous mettons les paramètres habituels:

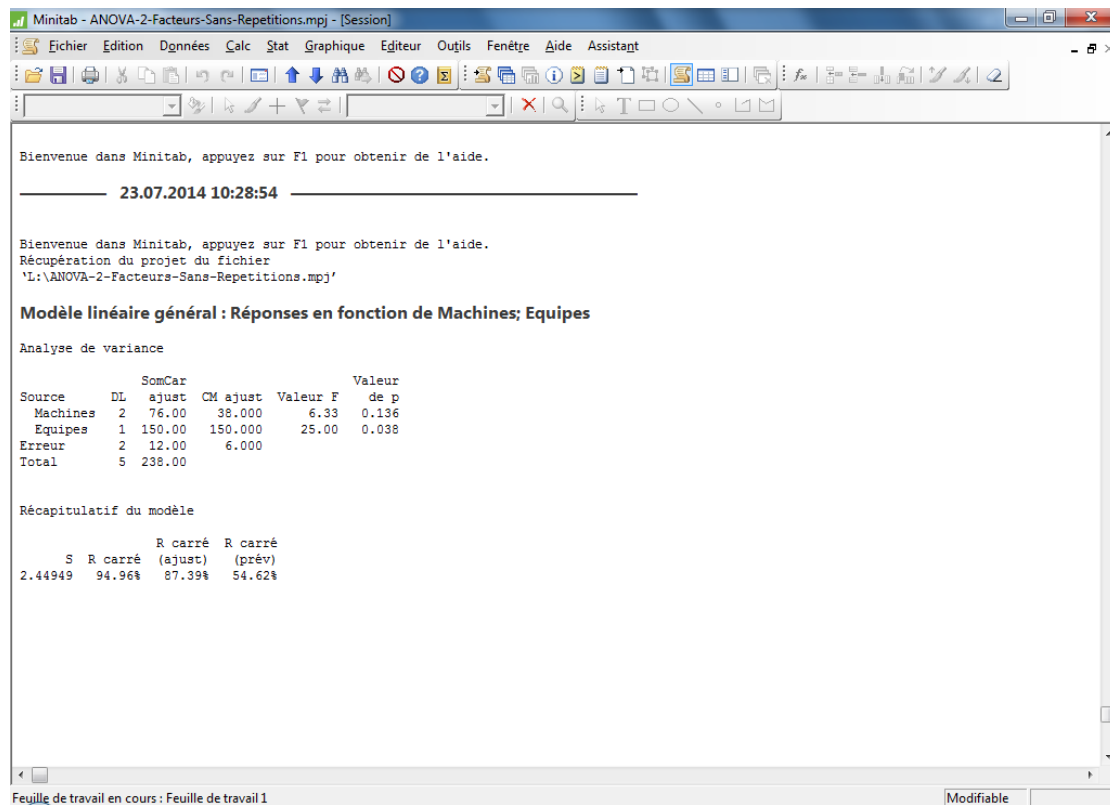




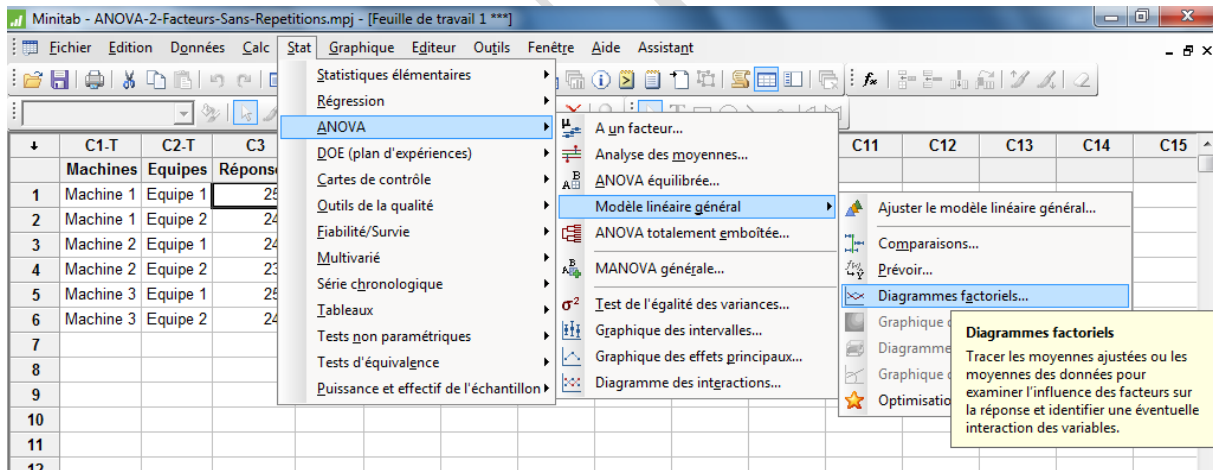
Ensuite, nous cliquons sur **Résultats...** pour se concentrer uniquement à ré-obtenir la même sortie que les versions précédentes l'ANOVA::



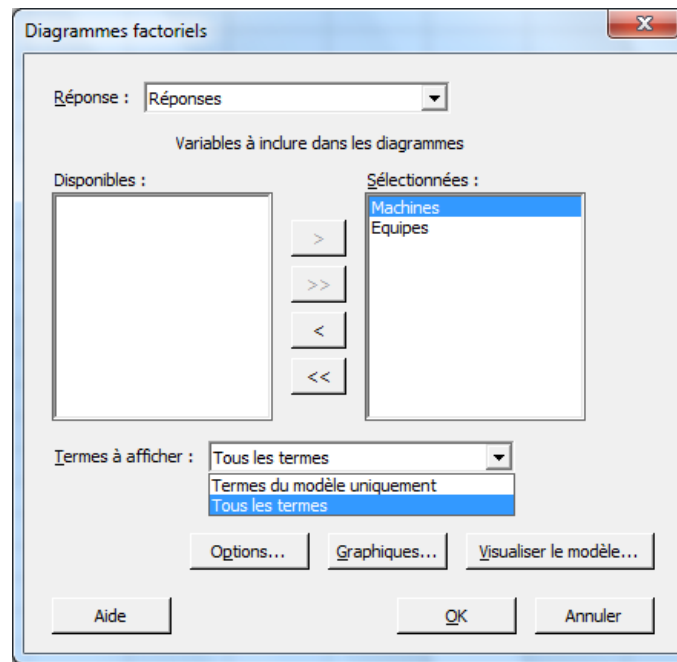
Nous validons deux fois par **OK** pour obtenir:



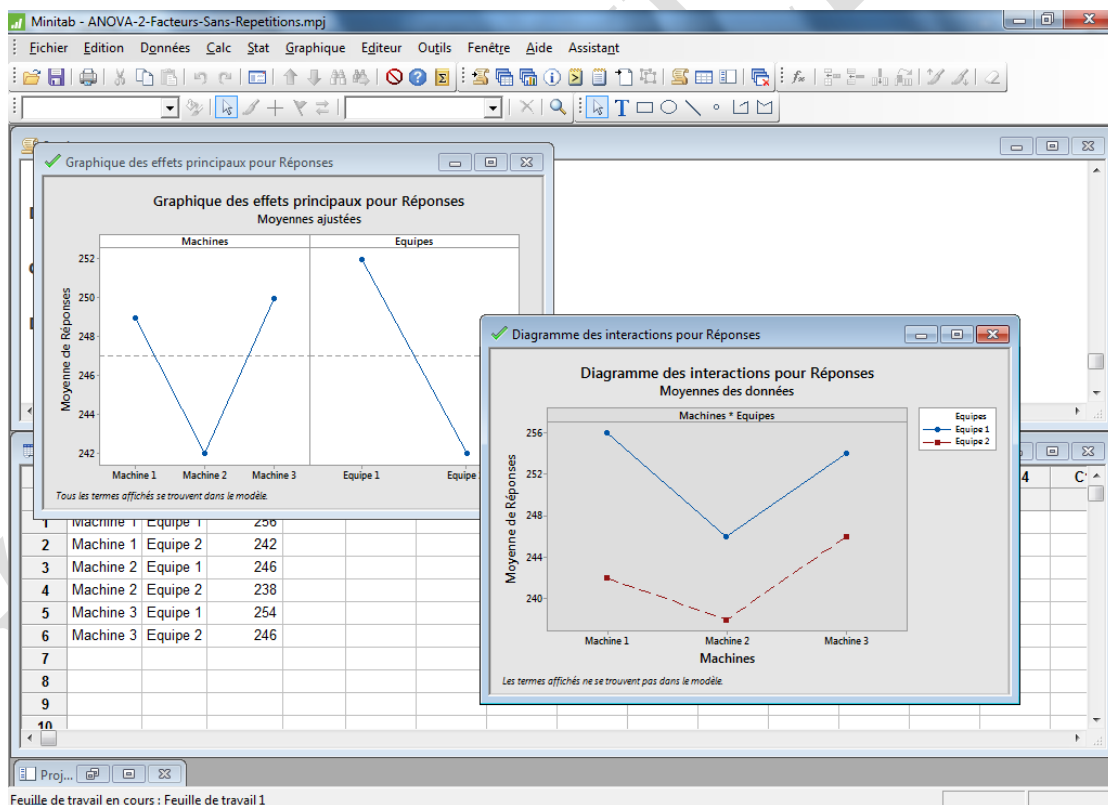
Nous retrouvons bien la même chose mais il nous manque les graphiques. Pour cela, nous allons dans le menu **Stat/ANOVA/Modèle linéaire général/Diagramme factoriels...**



Ce qui nous ouvre:



Si nous validons par **OK** nous réobtenons nos graphiques des versions précédentes:



## 12.7. Exercice 89.: Analyse de la variance carré latin

Minitab® Statistical Software 15.1.1.0

Nous allons comme à l'habitude reproduire l'exemple du cours théorique et vérifier que nous obtenons bien les mêmes résultats. Pour cela, nous partons des données suivantes reproduisant le tableau carré latin du cours théorique:

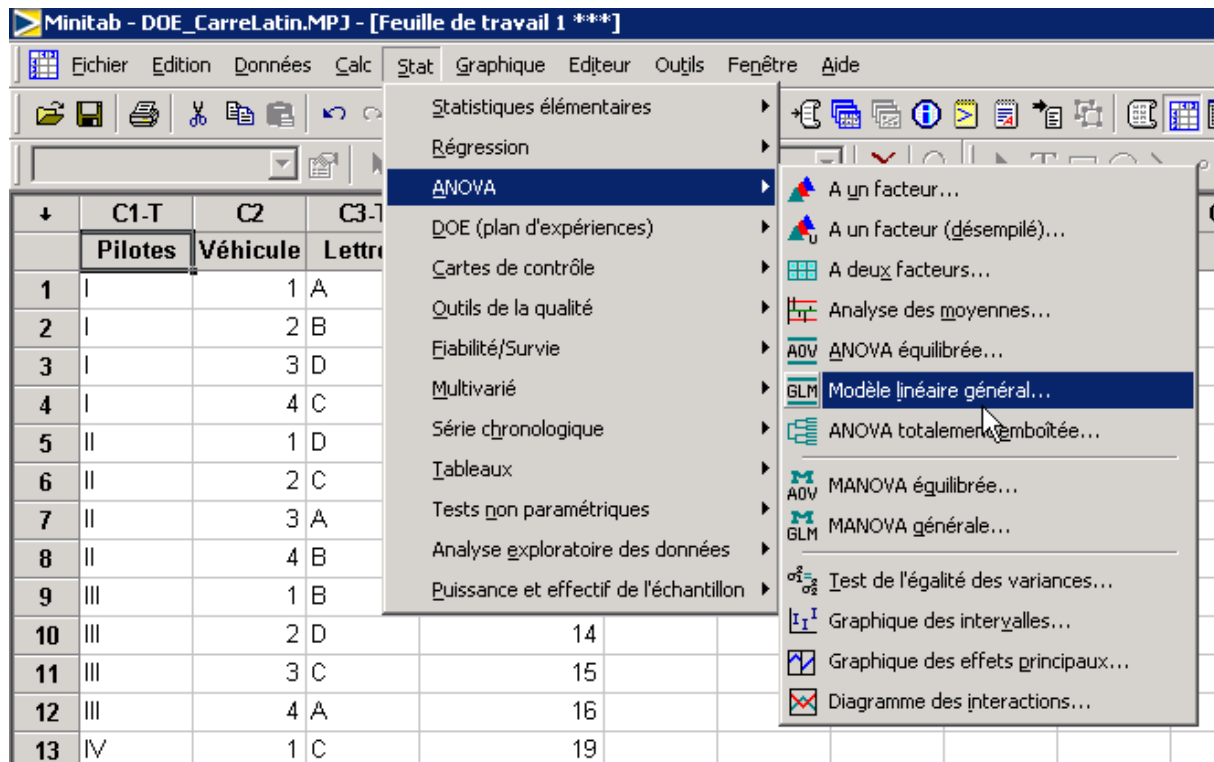
Considérons par exemple le cas d'une analyse de rapidité de 4 véhicules différents {1,2,3,4} par 4 pilotes différents {I,II,III,IV} sur 4 routes différentes {A, B, D, C} et les valeurs mesurées sont les consommations de carburant pour parcourir une distance fixe:

		Véhicules			
		1	2	3	4
Pilotes	I	A	B	D	C
	19	24	23	26	
	II	D	C	A	B
	23	24	19	30	
III	B	D	C	A	
15	14	15	16		
IV	C	A	B	D	
19	18	19	16		

Nous ouvrons pour cela le fichier *DOE\_CarreLatin.mpj*:

	C1-T	C2	C3-T	C4	C5	C6	C7
	Pilotes	Véhicule	Lettres	Consommation			
1	I	1	A	19			
2	I	2	B	24			
3	I	3	D	23			
4	I	4	C	26			
5	II	1	D	23			
6	II	2	C	24			
7	II	3	A	19			
8	II	4	B	30			
9	III	1	B	15			
10	III	2	D	14			
11	III	3	C	15			
12	III	4	A	16			
13	IV	1	C	19			
14	IV	2	A	18			
15	IV	3	B	19			
16	IV	4	D	16			
17							

Nous allons dans le menu **Stat/ANOVA/Modèle linéaire général...**:



Minitab - DOE\_CarreLatin.MPJ - [Feuille de travail 1 \*\*\*]

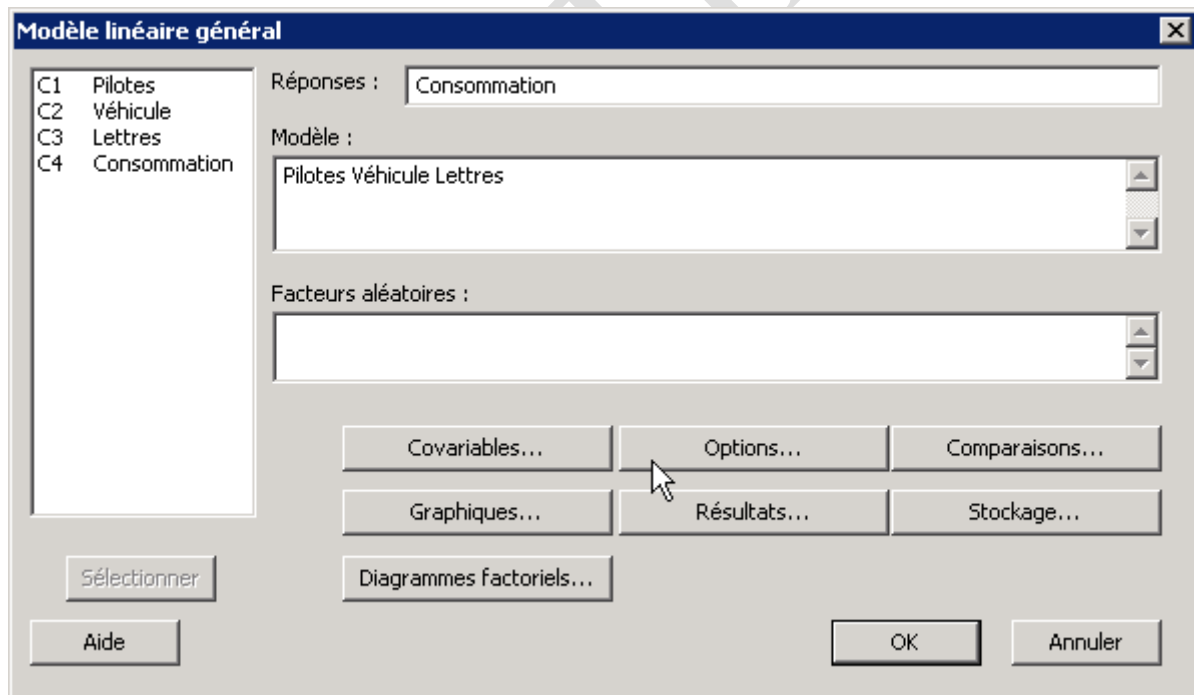
Fichier Edition Données Calc Stat Graphique Éditeur Outils Fenêtre Aide

Statistiques élémentaires  
Régression  
**ANOVA**  
DOE (plan d'expériences)  
Cartes de contrôle  
Outils de la qualité  
Fiabilité/Survie  
Multivarié  
Série chronologique  
Tableaux  
Tests non paramétriques  
Analyse exploratoire des données  
Puissance et effectif de l'échantillon

A un facteur...  
A un facteur (déséquilibré)...  
A deux facteurs...  
Analyse des moyennes...  
ANOVA équilibrée...  
**Modèle linéaire général...**  
ANOVA totalement emboîtée...  
MANOVA équilibrée...  
MANOVA générale...  
Test de l'égalité des variances...  
Graphique des intervalles...  
Graphique des effets principaux...  
Diagramme des interactions...

	C1-T	C2	C3-T
	Pilotes	Véhicule	Lettre
1	I	1	A
2	I	2	B
3	I	3	D
4	I	4	C
5	II	1	D
6	II	2	C
7	II	3	A
8	II	4	B
9	III	1	B
10	III	2	D
11	III	3	C
12	III	4	A
13	IV	1	C

et nous y mettons:



**Modèle linéaire général**

Réponses : Consommation

Modèle : Pilotes Véhicule Lettres

Facteurs aléatoires :

Covariables... Options... Comparaisons...  
Graphiques... Résultats... Stockage...  
Sélectionner Diagrammes factoriels...  
Aide OK Annuler

**Remarque:** La démarche est exactement la même pour une ANOVA Gréco-Latin...!!!

Nous validons par **OK** pour obtenir:

Minitab - DOE\_CarreLatin.MPJ - [Session]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

**Modèle linéaire général : Consommation en fonction de Pilotes; Véhicule; ...**

Facteur	Type	Niveaux	Valeurs
Pilotes	fixe	4	I; II; III; IV
Véhicule	fixe	4	1; 2; 3; 4
Lettres	fixe	4	A; B; C; D

Analyse de la variance pour Consommation, avec utilisation de la somme des carrés ajustée pour les tests

Source	DL	SomCar séq	SomCar ajust	CM ajust	F	P
Pilotes	3	216.000	216.000	72.000	13.50	0.004
Véhicule	3	24.000	24.000	8.000	1.50	0.307
Lettres	3	40.000	40.000	13.333	2.50	0.156
Erreur	6	32.000	32.000	5.333		
Total	15	312.000				

S = 2.30940    R carré = 89.74 %    R carré (ajust) = 74.36 %

Observations aberrantes pour Consommation

Observation	Consommation	Valeur ajustée	Ajust ErT	Valeur résiduelle	Valeur résiduelle normalisée
16	16.0000	19.0000	1.8257	-3.0000	-2.12 R

R indique une observation ayant une valeur résiduelle normalisée importante

Les valeurs du tableau de l'ANOVA sont donc conformes.

## 12.8. Exercice 90.: Analyse de la variance à deux facteurs fixes (ANOVA-2) avec répétitions (et sans interactions)

Minitab® Statistical Software 15.1.1.0

Imaginons une entreprise faisant les trois huit. Nous avons trois équipes qui travaillent sur une même machine. Nous souhaitons vérifier avec un seuil de confiance de 95% s'il y a une différence de productivité moyenne entre les trois équipes sur une semaine de travail (hypothèse que les moyennes sont égales).

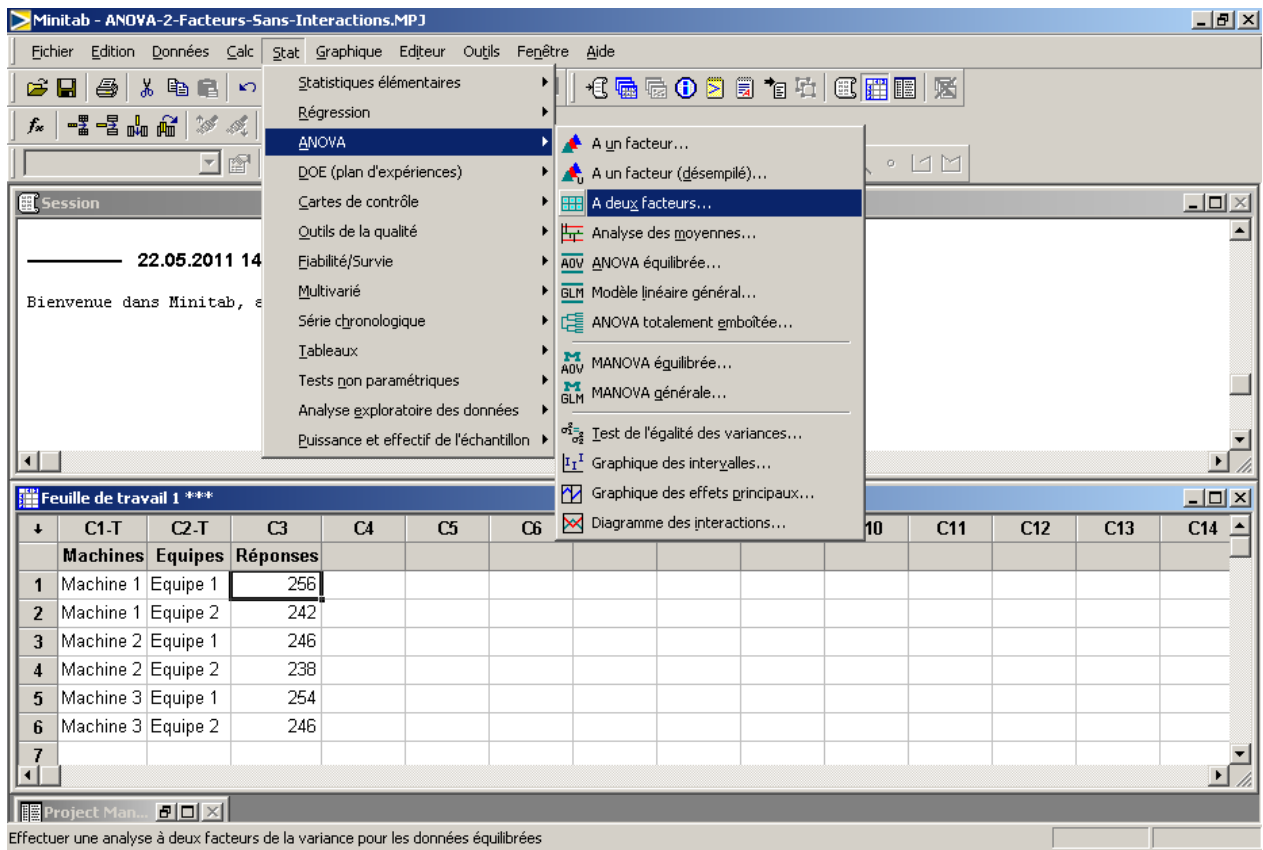
Nous savons également que l'ANOVA à deux facteurs fixes est importante dans les études R&R croisées et nous verrons donc lors de l'étude de cette dernière si nous retombons sur les mêmes résultats!

Remarque: La variable est donc la Machine et sa modalité a trois niveaux!

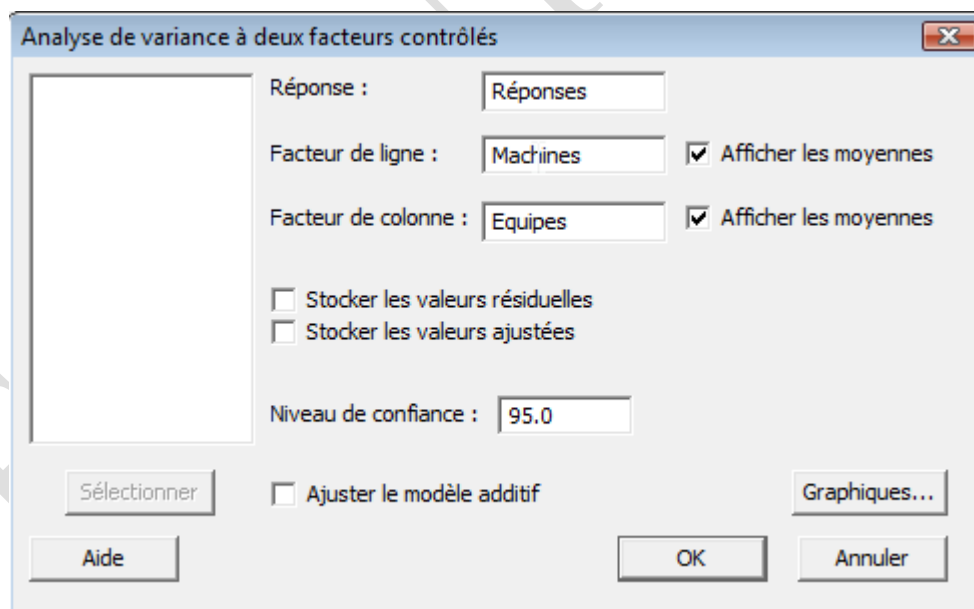
Ouvrez le fichier *ANOVA-2-Facteurs-Avec-Repetitions.mpj*:

	C1-T	C2-T	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Equipes	Machines	Réponses											
1	Equipe 1	Machine 1	256											
2	Equipe 1	Machine 1	251											
3	Equipe 1	Machine 2	246											
4	Equipe 1	Machine 2	248											
5	Equipe 1	Machine 3	254											
6	Equipe 1	Machine 3	251											
7	Equipe 2	Machine 1	242											
8	Equipe 2	Machine 1	240											
9	Equipe 2	Machine 2	238											
10	Equipe 2	Machine 2	240											
11	Equipe 2	Machine 3	246											
12	Equipe 2	Machine 3	246											
13														
14														
15														
16														
17														

Nous allons dans le menu **Stat/ANOVA/A deux facteurs...**:

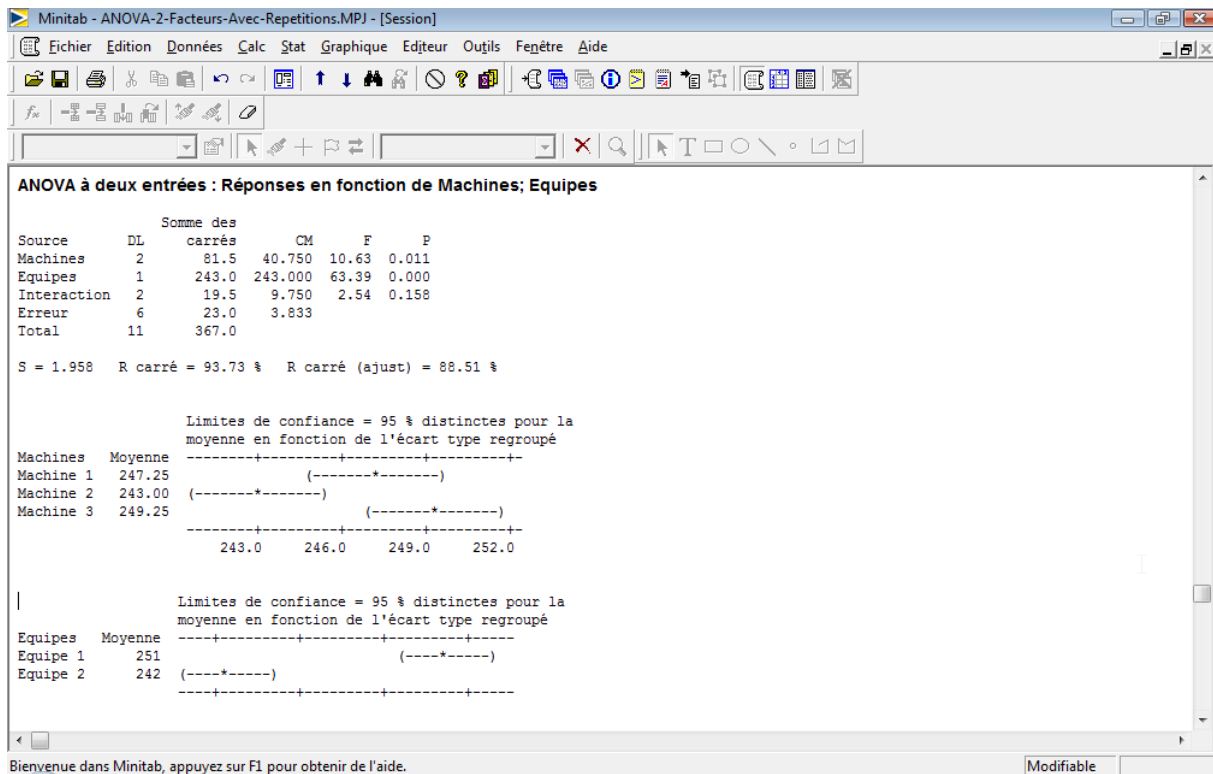


Pour avoir:



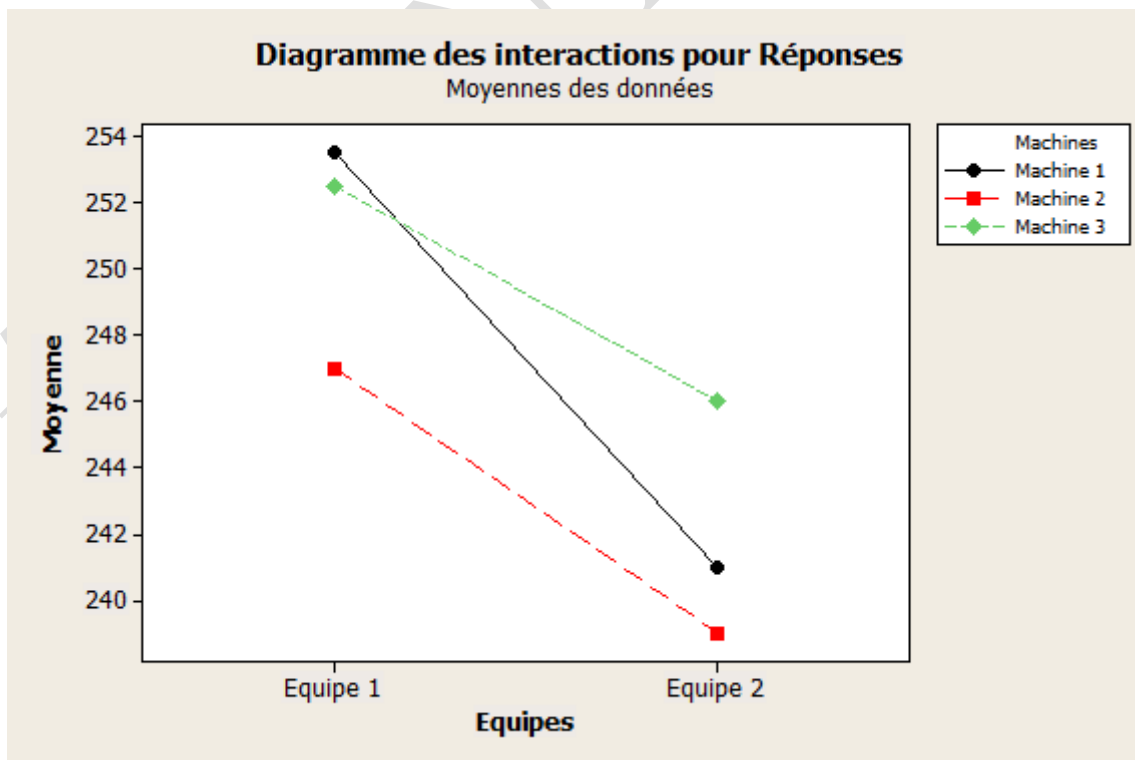
On clique sur **OK**:

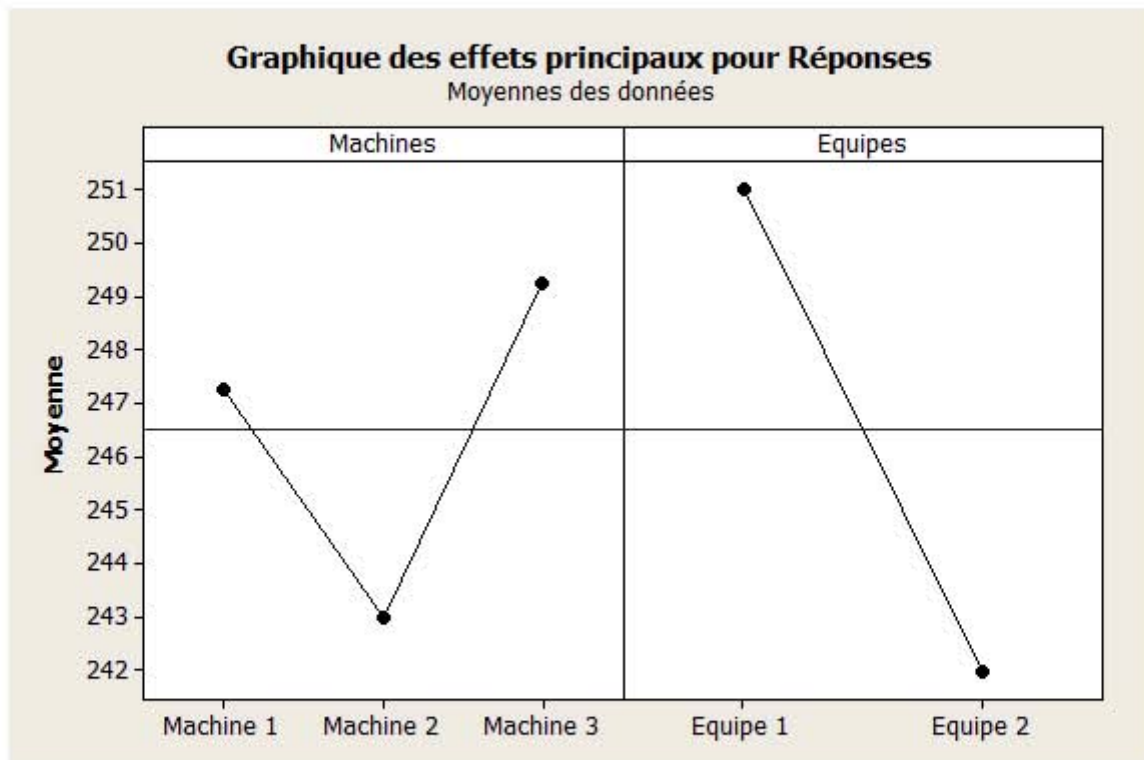




Nous obtenons donc les mêmes résultats avec les mêmes conclusions qu'avec Microsoft Excel.

Et pour info, voici les graphiques des interactions et facteurs principaux obtenus en procédant comme montré dans l'exercice relatif à ce sujet:





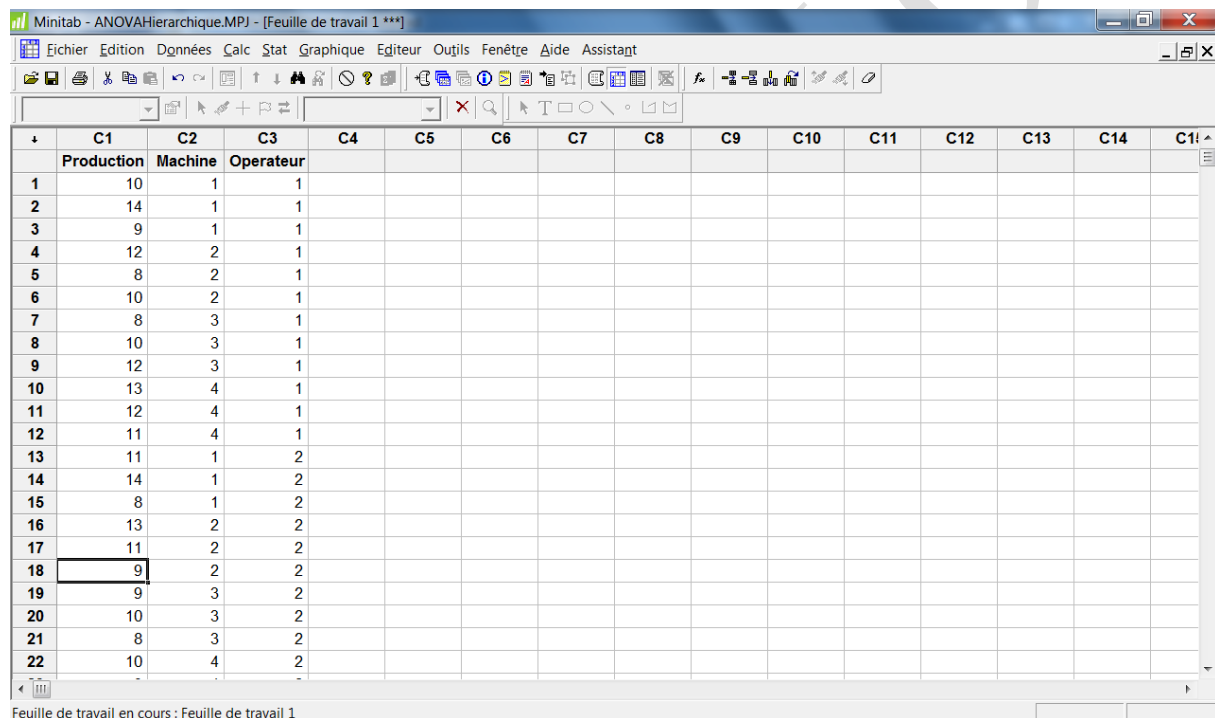
## 12.9. Exercice 91.: Analyse de la variance Imbriquée/Hiérarchisée (nested ANOVA)

Minitab® Statistical Software 16.1.2.0

Le but ici va être de vérifier si avec Minitab nous retrouvons les résultats de l'analyse d'une ANOVA hiérarchisée faite avec Microsoft Excel lors des démonstrations mathématiques dans le cours théorique de l'origine de l'ANOVA hiérarchisée.

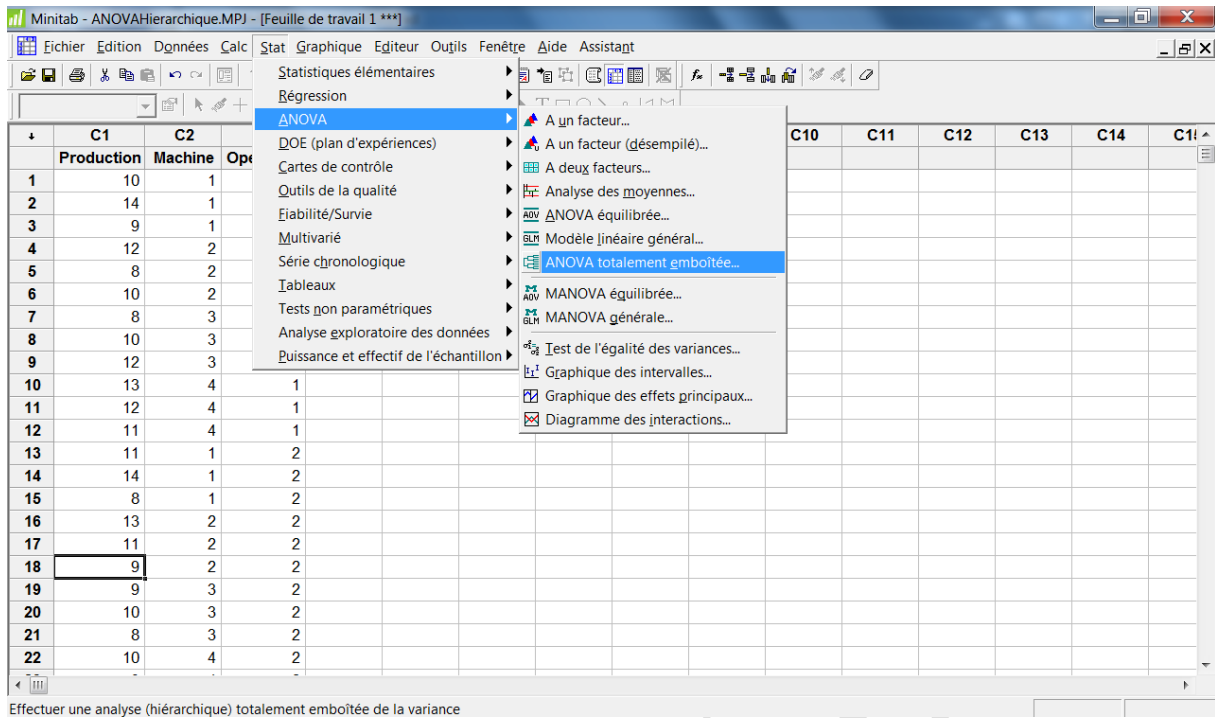
Nous savons également que l'ANOVA hiérarchisée est importante dans les études R&R emboîtées et nous verrons donc lors de l'étude de cette dernière si nous retombons sur les mêmes résultats!

Nous partons donc des données identiques à celles du cours théorique:

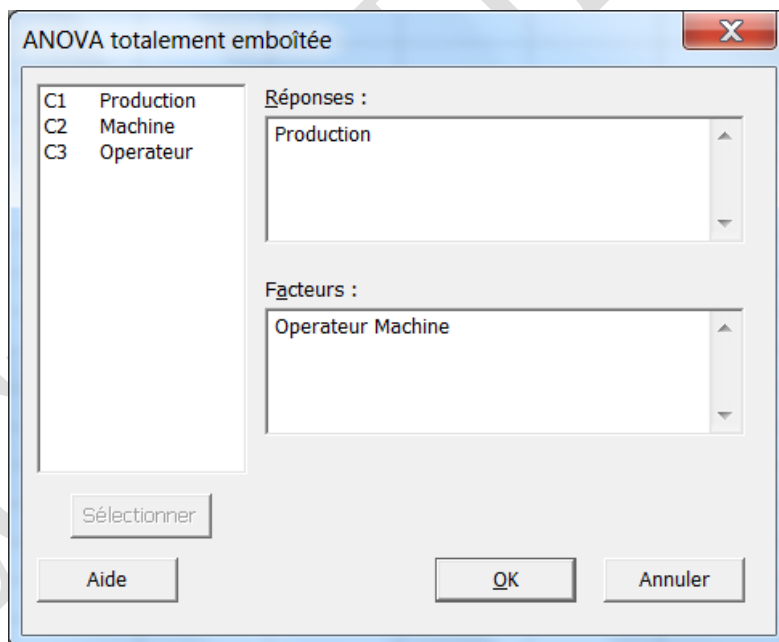


	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
	Production	Machine	Operateur												
1	10	1	1												
2	14	1	1												
3	9	1	1												
4	12	2	1												
5	8	2	1												
6	10	2	1												
7	8	3	1												
8	10	3	1												
9	12	3	1												
10	13	4	1												
11	12	4	1												
12	11	4	1												
13	11	1	2												
14	14	1	2												
15	8	1	2												
16	13	2	2												
17	11	2	2												
18	9	2	2												
19	9	3	2												
20	10	3	2												
21	8	3	2												
22	10	4	2												

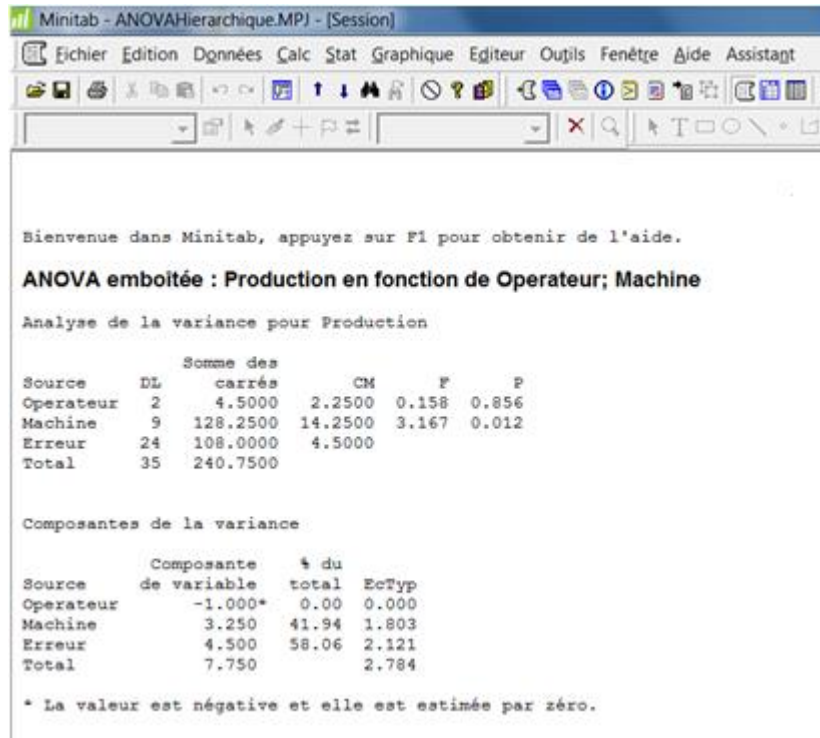
Nous allons dans le menu **Stat/ANOVA/ANOVA totalement emboîtée...**:



Nous prenons comme paramètres de réponses et des facteurs (le premier facteur doit être celui de premier niveau et ainsi de suite!!):



Nous validons par **OK** et obtenons alors:



Puisque la variance de la composante de variable opérateur est négative, nous devinons ici que Minitab fait en réalité une ANOVA emboîte avec les *Opérateurs* automatiquement gérés en tant que variable aléatoire et non en facteurs fixe!!!! **Et comme on ne peut pas avoir un facteur fixe imbriqué dans un facteur aléatoire, alors le facteur *Machine* est lui aussi forcément traité comme facteur aléatoire!!!!**

Nous retrouvons sinon bien les valeurs calculées (du moins pour celles visibles dans la capture d'écran) à la main et dans Microsoft Excel relativement aux démonstrations mathématiques faites dans le cours théorique.

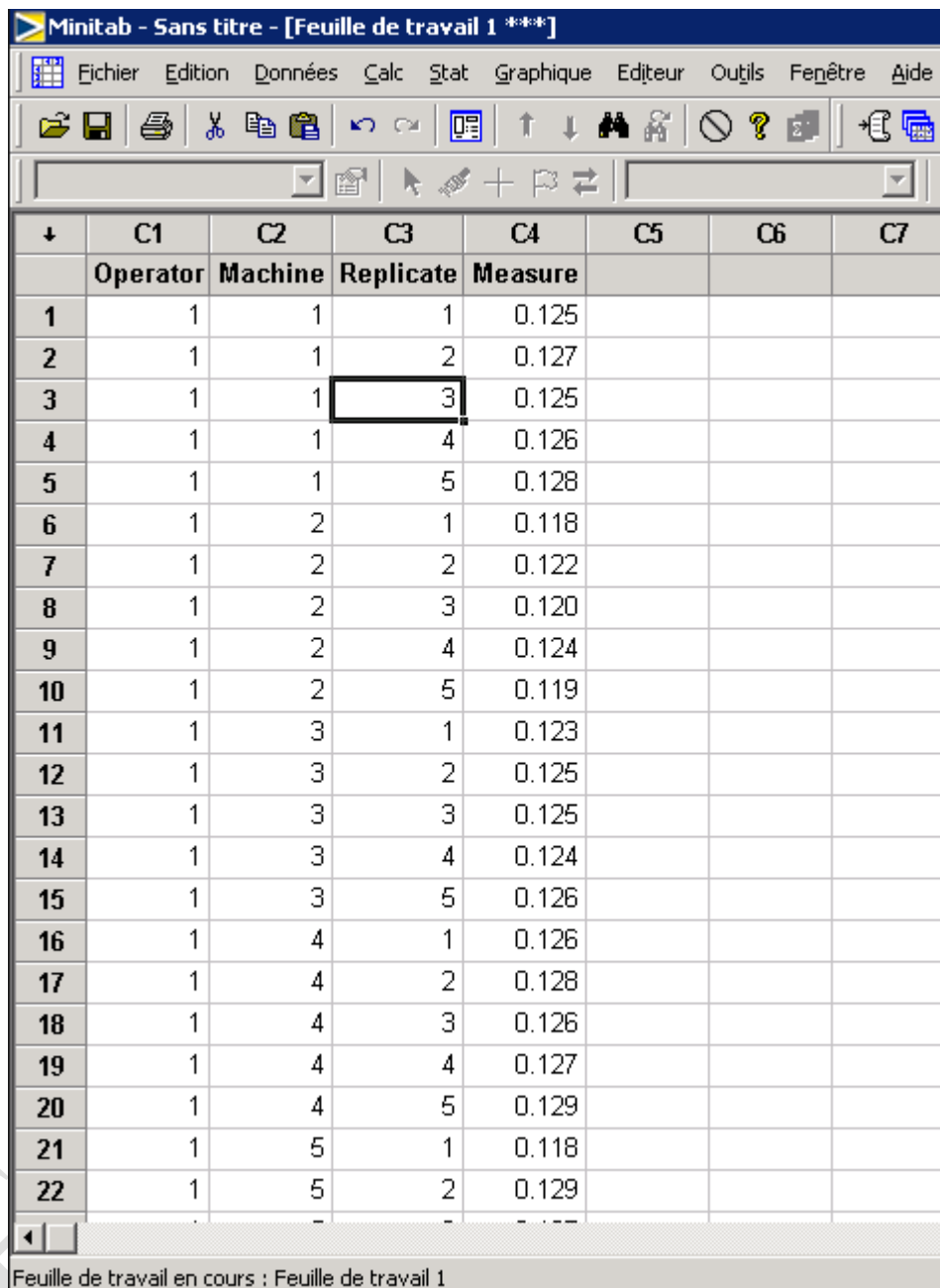
Ou en reprenant l'exemple du NIST pour ceux qui ne seraient pas convaincus:

<http://www.itl.nist.gov/div898/handbook/ppc/section2/ppc233.htm>

	<u>Machines</u>				
	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>
<b>Operator Day</b>	0.125	0.118	0.123	0.126	0.118
	0.127	0.122	0.125	0.128	0.129
	0.125	0.120	0.125	0.126	0.127
	0.126	0.124	0.124	0.127	0.120
	0.128	0.119	0.126	0.129	0.121
<b>Operator Night</b>	0.124	0.116	0.122	0.126	0.125
	0.128	0.125	0.121	0.129	0.123
	0.127	0.119	0.124	0.125	0.114
	0.126	0.125	0.126	0.130	0.124

	0.129	0.120	0.125	0.124	0.117
--	-------	-------	-------	-------	-------

Nous avons alors:



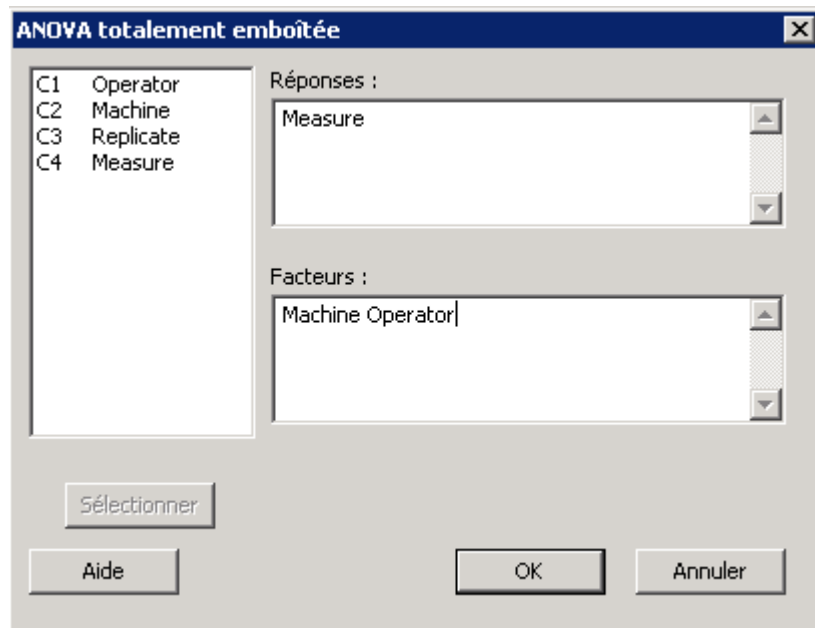
Minitab - Sans titre - [Feuille de travail 1 \*\*\*]

Échier Edition Données Calc Stat Graphique Éditeur Outils Fenêtre Aide

	C1	C2	C3	C4	C5	C6	C7
	Operator	Machine	Replicate	Measure			
1	1	1	1	0.125			
2	1	1	2	0.127			
3	1	1	3	0.125			
4	1	1	4	0.126			
5	1	1	5	0.128			
6	1	2	1	0.118			
7	1	2	2	0.122			
8	1	2	3	0.120			
9	1	2	4	0.124			
10	1	2	5	0.119			
11	1	3	1	0.123			
12	1	3	2	0.125			
13	1	3	3	0.125			
14	1	3	4	0.124			
15	1	3	5	0.126			
16	1	4	1	0.126			
17	1	4	2	0.128			
18	1	4	3	0.126			
19	1	4	4	0.127			
20	1	4	5	0.129			
21	1	5	1	0.118			
22	1	5	2	0.129			

Feuille de travail en cours : Feuille de travail 1

Nous prenons comme paramètres de réponses et des facteurs (le premier facteur doit être celui de premier niveau et ainsi de suite!):



Ce qui après validation nous donne bien:

**ANOVA emboîtée : Mesure en fonction de Machine; Operator**

Analyse de la variance pour Mesure

Source	DL	Somme des carrés	CM	F	P
Machine	4	0.0003	0.0001	20.384	0.003
Operator	5	0.0000	0.0000	0.430	0.825
Erreur	40	0.0003	0.0000		
Total	49	0.0007			

À comparer avec le résultat communiqué sur la page du NIST:

Source	Sum of Squares	Deg. of Freedom	Mean Square	$F_0$
Machine	3.03e-4	4	7.58e-5	20.38
Operator(Machine)	1.86e-5	5	3.72e-6	0.428
Residuals	3.46e-4	40	8.70e-6	
Corrected Total	6.68e-4	49		

## 12.10. Exercice 92.: Split-plot ANOVA (ANOVA en parcelles)

Minitab® Statistical Software 18.1

Il est temps de passer au non moins fameuses ANOVA en parcelles divisées. Alors je tiens à rappeler que:

1. C'est un sujet sensible et qu'il n'est pas facile de choisir les bonnes ANOVA pour des expériences complexes (même les meilleurs statisticiens peuvent se tromper)
2. Les ANOVA split-plot continuent à faire l'objet d'assez nombreuses publications scientifiques et donc il y a plusieurs "formules" pour les calculs. Vous ne risquez pas toujours de tomber sur les mêmes résultats d'un logiciel à un autre...
3. Vous allez voir que Minitab à l'heure actuelle (version 18.1) n'est pas vraiment fait pour les ANOVA split-plot car il ne sait pas les afficher selon la méthode traditionnelle....

Alors allons-y! Nous partons des mêmes données que celle du cours théorique et qui pour rappel viennent du livre *A first course in biometry for agriculture students* de A.A. Rayner publié en 1969:

(A) Treatment	Blocks				Whole-plot (C) treatment totals
	1	2	3	4	
1A	56	45	43	46	
1B	61	58	55	56	
1C	63	53	49	48	
1D	65	61	60	63	
1E	60	61	50	53	
<b>(B) Whole-plot totals</b>	<b>305</b>	<b>278</b>	<b>257</b>	<b>266</b>	<b>1,106</b>
2A	66	57	50	50	
2B	59	55	51	52	
2C	66	58	52	55	
2D	53	53	48	55	
2E	73	77	77	65	
<b>(B) Whole-plot totals</b>	<b>317</b>	<b>300</b>	<b>278</b>	<b>277</b>	<b>1,172</b>
3A	60	50	45	48	
3B	60	59	54	54	
3C	65	56	50	50	
3D	60	58	56	60	
3E	62	68	67	60	
<b>(B) Whole-plot totals</b>	<b>307</b>	<b>291</b>	<b>272</b>	<b>272</b>	<b>1,142</b>
<b>(C) Block totals</b>	<b>929</b>	<b>869</b>	<b>807</b>	<b>815</b>	<b>3,420</b>

(D) Method of cultivation	Varieties					Methods totals
	A	B	C	D	E	
1	190	230	213	249	224	1,106
2	223	217	231	209	292	1,172
3	203	227	221	234	257	1,142
Variety totals	616	674	665	692	773	3,420

Ce qui donne dans Minitab:

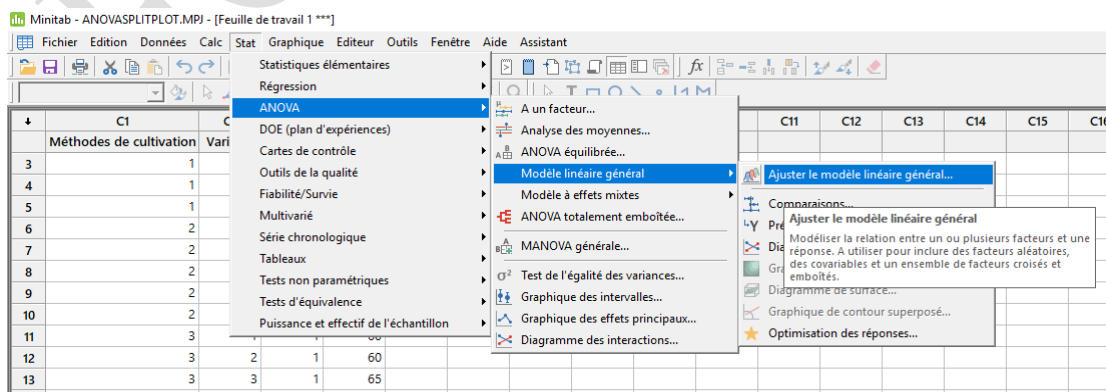


Minitab - ANOVASPLITPLOT.MPJ - [Feuille de travail 1 \*\*\*]

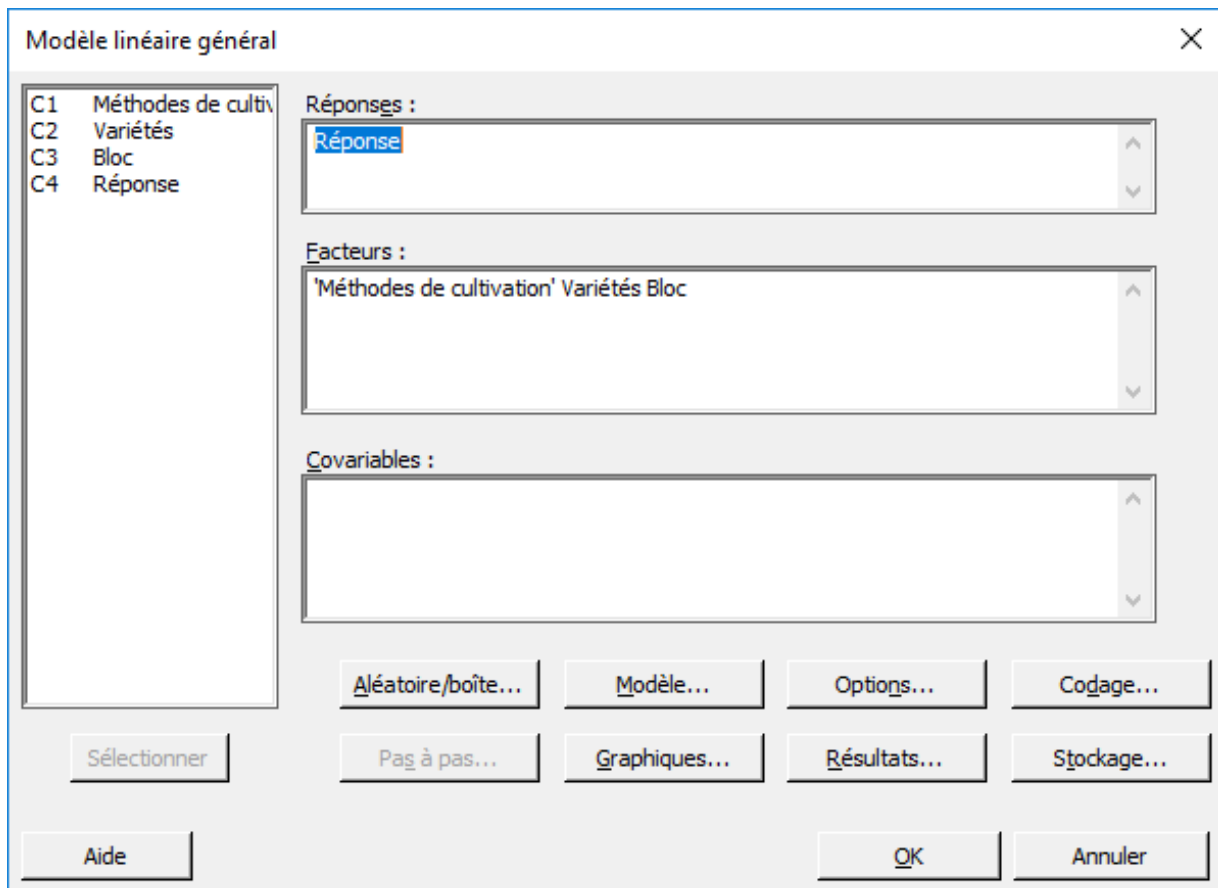
Fichier Edition Données Calc Stat Graphique Editeur Out

	C1	C2	C3	C4
	Méthodes de cultivation	Variétés	Bloc	Réponse
3	1	3	1	63
4	1	4	1	65
5	1	5	1	60
6	2	1	1	66
7	2	2	1	59
8	2	3	1	66
9	2	4	1	53
10	2	5	1	73
11	3	1	1	60
12	3	2	1	60
13	3	3	1	65
14	3	4	1	60
15	3	5	1	62
16	1	1	2	45
17	1	2	2	58
18	1	3	2	53
19	1	4	2	61
20	1	5	2	61
21	2	1	2	57
22	2	2	2	55
23	2	3	2	58
24	2	4	2	53
25	2	5	2	77
26	3	1	2	50
27	3	2	2	59
28	3	3	2	56
29	3	4	2	58
30	3	5	2	68
31	1	1	3	43
32	1	2	3	55
33	1	3	3	49
34	1	4	3	60
35	1	5	3	50
36	2	1	3	50
37	2	2	3	51
38	2	3	3	52
39	2	4	3	48
40	2	5	3	77

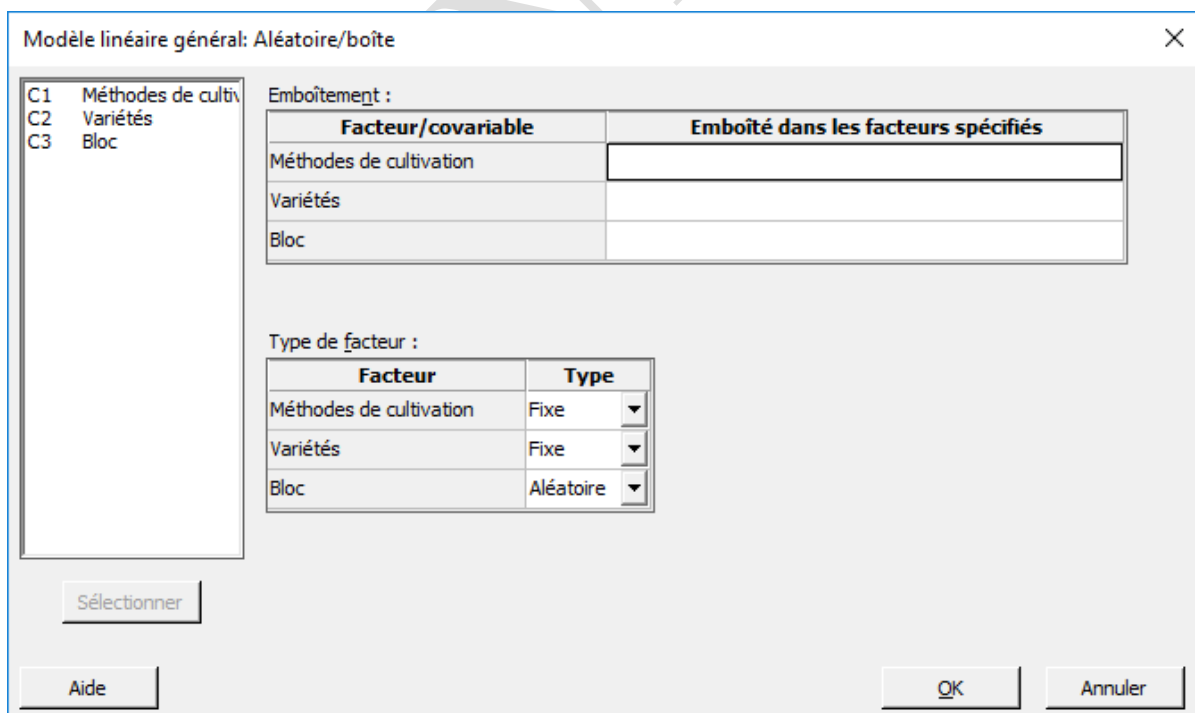
Ensuite, nous allons dans le menu ANOVA/Modèle linéaire général/Ajuster le modèle linéaire général...:



Nous y mettons:



Et là cela commence à être subtile... Cliquer sur le bouton **Aléatoire/boîte...** et y mettre:



On clique sur **OK** et ensuite on va sur le bouton **Modèle...** pour prendre (encore plus subtil car il faut vraiment connaître la théorie mathématique!):

Modèle linéaire général: Modèle

Facteurs et covariables : Ajouter des termes en sélectionnant les facteurs, les prédicteurs et les termes du modèle :

'Méthodes de cultivation'  
Variétés  
Bloc

Interactions jusqu'à l'ordre : 2 Ajouter

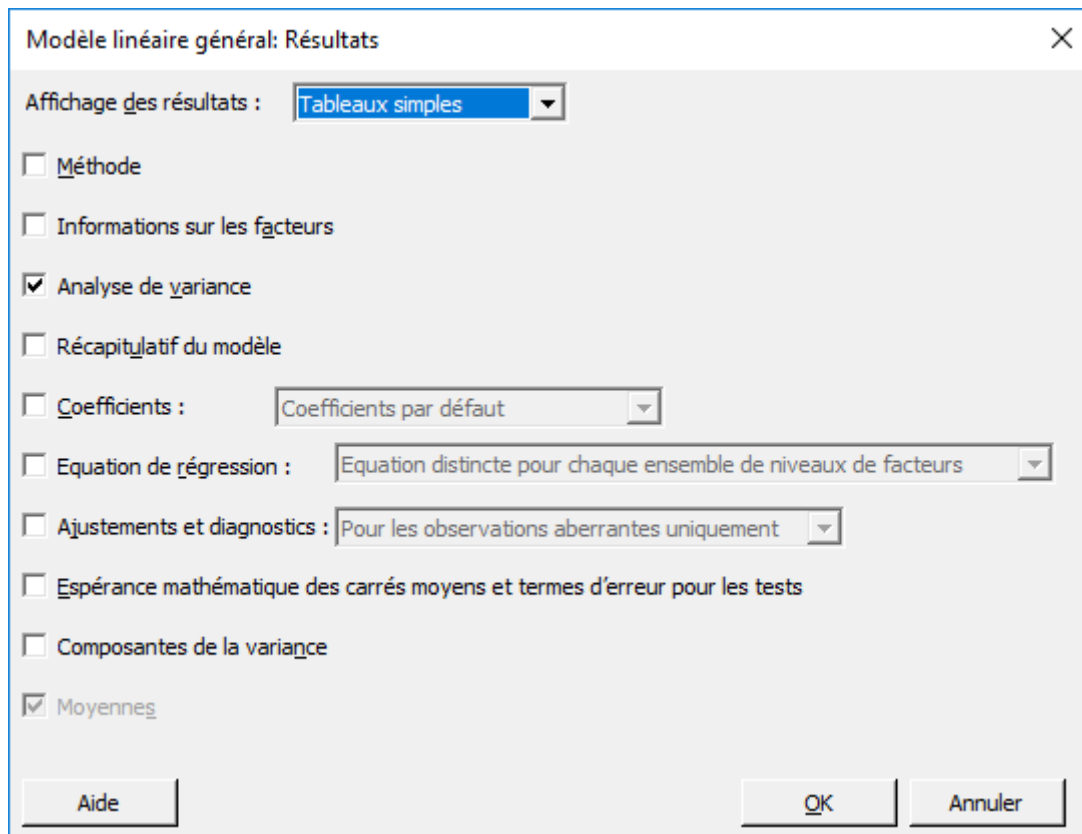
Termes, covariables et facteurs croisés dans le modèle Ajouter

Termes dans le modèle : Valeur par défaut X ↓ ↑

'Méthodes de cultivation'  
Variétés  
Bloc  
'Méthodes de cultivation'\*Variétés  
'Méthodes de cultivation'\*Bloc

Aide OK Annuler

On clique sur **OK** et ensuite sur **Résultats...** car uniquement le tableau de l'ANOVA nous intéresse:



Et on valide deux fois par **OK** pour obtenir:

**Modèle linéaire général : Réponse en fonction de ... ion; Variétés; Bloc**

**Analyse de variance**

Source	DL	SomCar ajust	CM ajust	Valeur F	Valeur de p
Méthodes de cultivation	2	109.20	54.600	21.55	0.002
Variétés	4	1089.17	272.292	18.96	0.000
Bloc	3	638.40	212.800	84.00	0.000
Méthodes de cultivation*Variétés	8	875.13	109.392	7.62	0.000
Méthodes de cultivation*Bloc	6	15.20	2.533	0.18	0.982
Erreur	36	516.90	14.358		
Total	59	3244.00			

Feuille de travail en cours : Feuille de travail 1

Et nous obtenons bien après 30 seconde de clics (...), exactement les mêmes résultats que ceux calculés (très péniblement) à la main dans le cours théorique et que du livre de référence!

### 12.11. Exercice 93.: Test de Kruskal-Wallis (ANOVA non paramétrique)

Minitab® Statistical Software 15.1.1.0

Basons-nous par hommage sur l'exemple original de l'article de Kruskal-Wallis comme nous l'avons fait dans le cours théorique:

Standard		Modification 1		Modification 2	
Output	Rank	Output	Rank	Output	Rank
340	5	339	4	347	10
345	9	333	2	343	7
330	1	344	8	349	11
342	6			355	12
338	3				
					Sum
<i>n</i>	5		3		4
<i>R</i>	24		14		40
<i>R</i> <sup>2</sup> / <i>n</i>	115.2		65.333		400.
					580.533

Checks :  $\sum n = N = 12$

$\sum R = \frac{1}{2}N(N+1) = 78$

$H = \frac{12 \times 580.533}{12 \times 13} - 3 \times 13 = 5.656 \simeq \chi^2(2)$

from (1.2)

$\Pr[\chi^2(2) \geq 5.656] = 0.059$

from [9] or [13]

$\Pr[H(5, 4, 3) \geq 5.656] = 0.049$

from Table 6.1

Donc les productions de trois machines différentes (une standard, une modifiée par un premier expert, une dernière modifiée par une équipe interne). Dans Minitab, cela se mettra sous la forme:

	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
	Machine	Production													
1	Standard	340													
2	Standard	345													
3	Standard	330													
4	Standard	342													
5	Standard	338													
6	Modifiée 1	339													
7	Modifiée 1	333													
8	Modifiée 1	344													
9	Modifiée 2	347													
10	Modifiée 2	343													
11	Modifiée 2	349													
12	Modifiée 2	355													
13															
14															
15															
16															
17															
18															
19															
20															

Ensuite, nous allons dans le menu **Stat/Tests non paramétriques/Kruskal-Wallis...**:

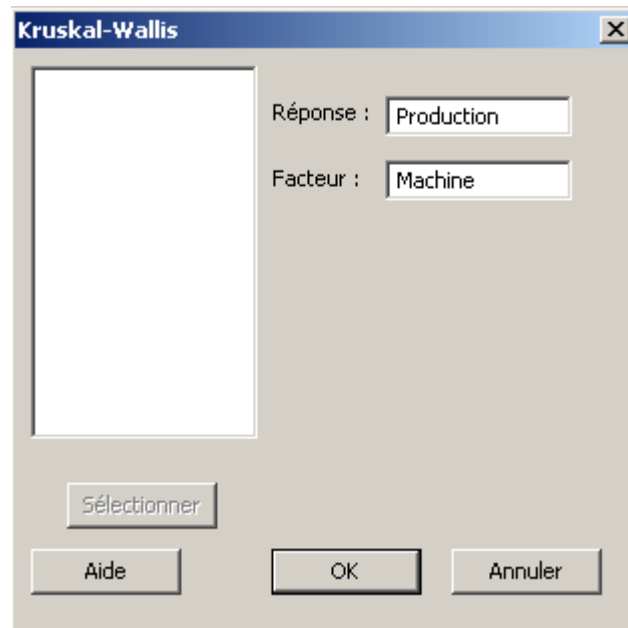
The screenshot shows the Minitab 'Stat' menu with the following options:

- Statistiques élémentaires
- Régression
- ANOVA
- DOE (plan d'expériences)
- Cartes de contrôle
- Outils de la qualité
- Fiabilité/Survie
- Multivarié
- Série chronologique
- Tableaux
- Tests non paramétriques** (highlighted)
- Analyse exploratoire des données
- Puissance et effectif de l'échantillon

The sub-menu for 'Tests non paramétriques' is open, showing the following options:

- Signe à 1 échantillon...
- Wilcoxon pour 1 échantillon...
- Mann-Whitney...
- Kruskal-Wallis...** (highlighted)
- Test de Mood pour les médianes...
- Friedman...
- Test des suites...
- Moyennes deux à deux...
- Différences deux à deux...
- Pentes deux à deux...

Nous entrons (il y a peu d'options... donc difficile de se tromper):



Nous obtenons dans la fenêtre de session:

### Test de Kruskal-Wallis : Production en fonction de Machine

Test de Kruskal-Wallis sur Production

Machine	N	Médiane	Rang moyen	Z
Modifiée 1	3	339.0	4.7	-1.02
Modifiée 2	4	348.0	10.0	2.38
Standard	5	340.0	4.8	-1.38
Global	12		6.5	

B H = 5.66 DL = 2 P = 0.059

\* REMARQUE \* Un ou plusieurs petits échantillons

Nous retrouvons donc la même chose que les calculs effectués à la main par Kruskal-Wallis avec la relation "pseudo-démontrée" (à l'envers...) dans le cours de statistique théorique. Dans le cas présent, à un niveau  $\alpha$  de 5%, nous sommes donc à la limite avec l'approximation par une loi de Khi-deux. Comme l'ont montré Kruskal et Wallis, une simulation par Monte-Carlo donne une  $p$ -value de 0.049.

Bref, dans cette situation il conviendrait plutôt de rejeter l'hypothèse nulle comme quoi les productions sont similaires. Et donc privilégier le fait que celles-ci soient plutôt différentes. Une recommandation et de refaire le test par paire des mesures pour voir ci qui est significativement différent deux par deux.

## 12.12. Exercice 94.: ANCOVA (Analyse de la Covariance)

Minitab® Statistical Software 15.1.1.0

Notre but ici est de vérifier si nous réobtenons bien les résultats des calculs effectués à la main dans le cours théorique et ce avec quel niveau de détails.

D'abord, nous reprenons les données du cours théorique qui sous forme esthétique est donnée pour rappel par:

	Méthode A			Méthode B	
Sujet	$X_a$	$Y_a$	Sujets	$X_b$	$Y_b$
a1	5	20	b1	7	19
a2	10	23	b2	12	26
a3	12	30	b3	27	33
a4	9	25	b4	24	35
a5	23	34	b5	18	30
a6	21	40	b6	22	31
a7	14	27	b7	26	34
a8	18	38	b8	21	28
a9	6	24	b9	14	23
a10	13	31	b10	9	22
Moyenne	13.1	29.2		18.0	28.1

Ce qui nous donne dans Minitab:



Minitab - ANCOVA.MPJ - [Feuille de travail 1 \*\*\*]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

	C1-T	C2	C3	C4-T	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Sujet	Before	After	Method									
3	a3	12	30	A									
4	a4	9	25	A									
5	a5	23	34	A									
6	a6	21	40	A									
7	a7	14	27	A									
8	a8	18	38	A									
9	a9	6	24	A									
10	a10	13	31	A									
11	b1	7	19	B									
12	b2	12	26	B									
13	b3	27	33	B									
14	b4	24	35	B									
15	b5	18	30	B									
16	b6	22	31	B									
17	b7	26	34	B									
18	b8	21	28	B									
19	b9	14	23	B									
20	b10	9	22	B									
21													

Feuille de travail en cours : Feuille de travail 1

Ensuite nous allons dans le menu **Stat/ANOVA/Modèle linéaire général...**:

Minitab - ANCOVA.MPJ - [Feuille de travail 1 \*\*\*]

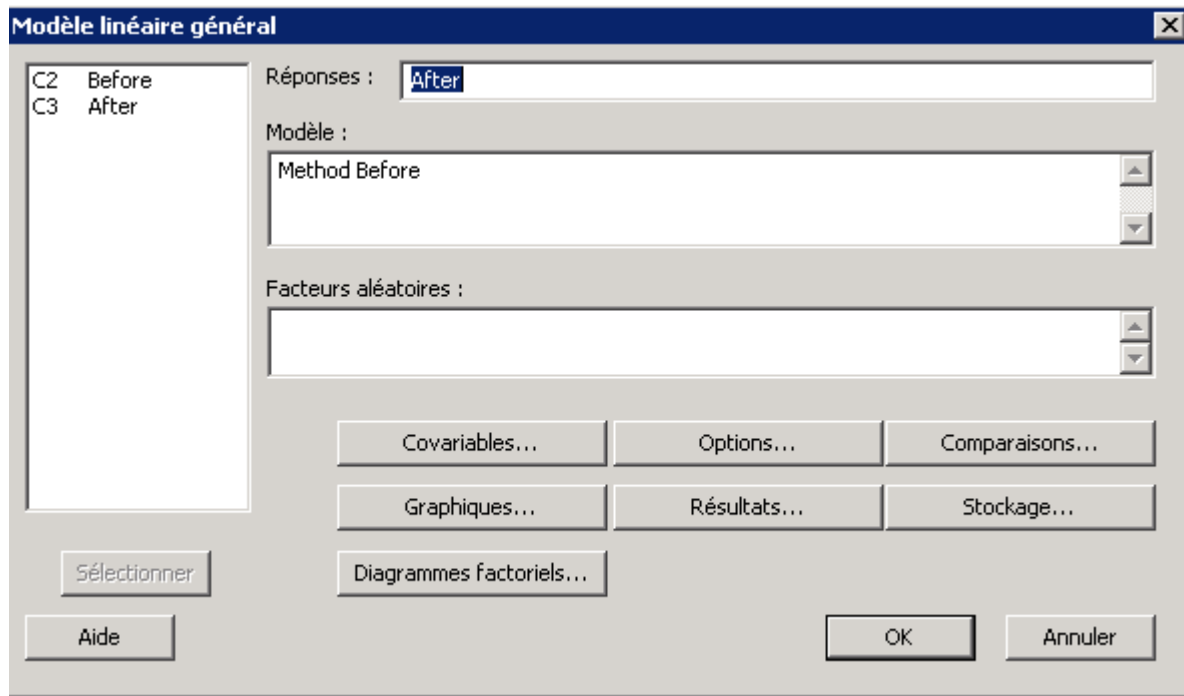
Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

Statistiques élémentaires  
Régression  
**ANOVA**  
DOE (plan d'expériences)  
Cartes de contrôle  
Outils de la qualité  
Fiabilité/Survie  
Multivarié  
Série chronologique  
Tableaux  
Tests non paramétriques  
Analyse exploratoire des données  
Puissance et effectif de l'échantillon

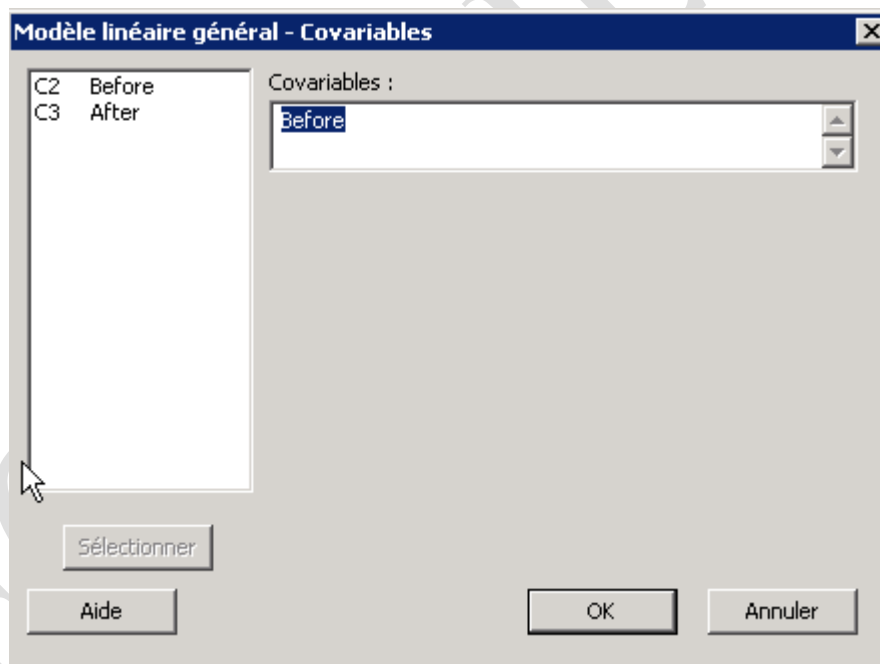
A un facteur...  
A un facteur (désempilé)...  
A deux facteurs...  
Analyse des moyennes...  
ANOVA équilibrée...  
**GLM** **Modèle linéaire général...**  
ANOVA totalement emboîtée...  
MANOVA équilibrée...  
MANOVA générale...  
Test de l'égalité des variances...  
Graphique des intervalles...  
Graphique des effets principaux...  
Diagramme des interactions...

	C1-T	C2	C3	C4-T	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Sujet	Before	After	Method									
3	a3	12	30	A									
4	a4	9	25	A									
5	a5	23	34	A									
6	a6	21	40	A									
7	a7	14	27	A									
8	a8	18	38	A									
9	a9	6	24	A									
10	a10	13	31	A									
11	b1	7	19	B									
12	b2	12	26	B									
13	b3	27	33	B									
14	b4	24	35	B									

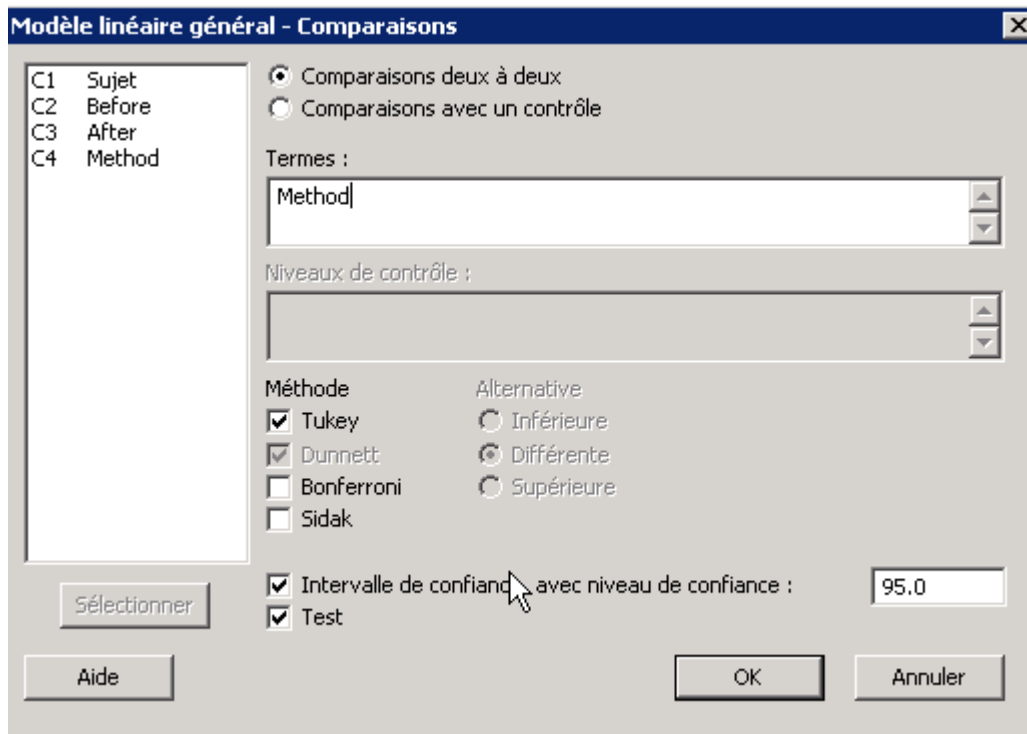
et dans la boîte de dialogue qui apparaît nous mettons:



Et nous cliquons sur le bouton **Covariables...** pour mettre:



Ensuite, nous allons dans le bouton **Comparaisons...** pour prendre:



Nous validons ensuite deux fois par **OK** pour obtenir:

**Modèle linéaire général : After en fonction de Method**

Facteur Type Niveaux Valeurs  
 Method fixe 2 A; B

Analyse de la variance pour After, avec utilisation de la somme des carrés ajustée pour les tests

Source	DL	SomCar séq	SomCar ajust	CM ajust	F	P
Method	1	6.05	115.31	115.31	16.03	0.001
Before	1	540.18	540.18	540.18	75.07	0.000
Erreur	17	122.32	122.32	7.20		
Total	19	668.55				

S = 2.68241 R carré = 81.70 % R carré (ajust) = 79.55 %

Terme	Coeff	Coef ErT	T	P
Constante	15.783	1.602	9.85	0.000
Before	0.82748	0.09550	8.66	0.000

Nous retrouvons en rouge les mêmes résultats que ceux calculés à la main dans le cours théorique. Donc les méthodes d'enseignement sous les hypothèses de construction de l'ANCOVA diffèrent bien.

Sinon nous avons aussi dans la fenêtre de session:

Intervalles de confiance simultanés de Tukey = 95.0 %  
 Variable de réponse After  
 Toutes les comparaisons deux à deux sur les niveaux de Method  
 Method = A soustrait de :

Method	Inférieur	Centré	Supérieur	
B	-7.871	-5.155	-2.438	(-----*-----)

-7.5      -5.0      -2.5      0.0

Tests de simultanéité de Tukey  
 Variable de réponse After  
 Toutes les comparaisons deux à deux sur les niveaux de Method  
 Method = A soustrait de :

Method	Différence des moyennes	Erreur type de la différence	Valeur de T	Valeur de p ajustée
B	-5.155	1.288	-4.003	0.0009

Nous retrouvons donc ici les différences ajustées et la fameuse valeur de -5.16 que nous avons obtenus à la main dans le cours théorique. La méthode *B* ayant une différence négative par rapport à la méthode *A* cette dernière est alors la plus efficace!

## 12.13. Exercice 95.: MANOVA (multivariate ANOVA)

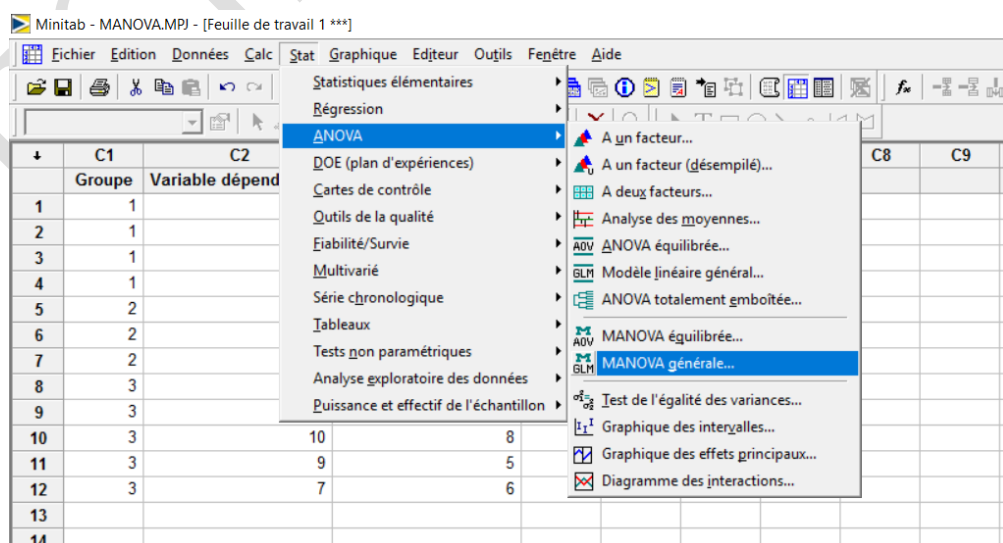
Minitab® Statistical Software 15.1.1.0

Notre but ici est de vérifier si nous réobtenons bien les résultats des calculs effectués à la main dans le cours théorique et ce avec quel niveau de détails.

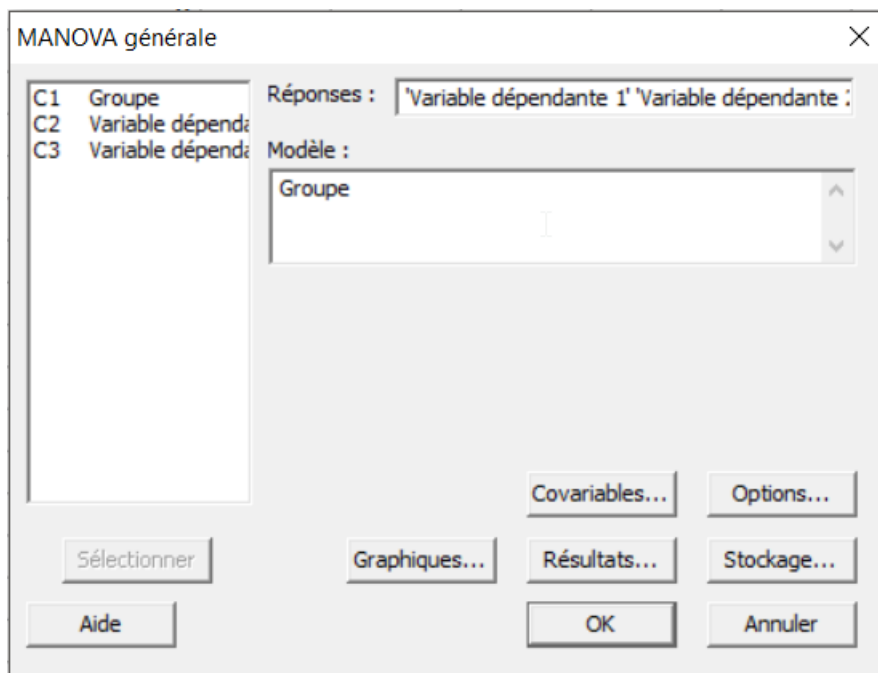
D'abord, nous reprenons les données du cours théorique qui sous forme esthétique est donnée pour rappel par (avec des catégories fictives nominatives pour cette fois car cela facilite la lecture):

	C1	C2	C3	C4	C5	C6
	Groupe	Variable dépendante 1	Variable dépendante 2			
1	1	2	3			
2	1	3	4			
3	1	5	4			
4	1	2	5			
5	2	4	8			
6	2	5	6			
7	2	6	7			
8	3	7	6			
9	3	8	7			
10	3	10	8			
11	3	9	5			
12	3	7	6			
13						

Ensuite, nous allons dans le menu **Stat/ANOVA/MANOVA généralisée...**:



Nous prenons:



Nous validons sur **OK** pour obtenir dans la fenêtre de session:

**Modèle linéaire général : Variable dép, Variable dép en fonction de Groupe**

MANOVA de Groupe  
 $s = 2$      $m = -0.5$      $n = 3.0$

Critère	Statistique de test	F	Nombre	DL		P
				Dénominateur		
De Wilk	0.08967	9.358	4	16	0.000	
Lawley-Hotelling	5.78618	10.126	4	14	0.000	
De Pillai	1.30178	8.390	4	18	0.001	
De Roy	4.89426					

Donc nous obtenons la même valeur de Wilk que dans le cours théorique et que R. Concernant la  $p$ -value elle est la même que dans R.

## 13. Régressions et interpolations

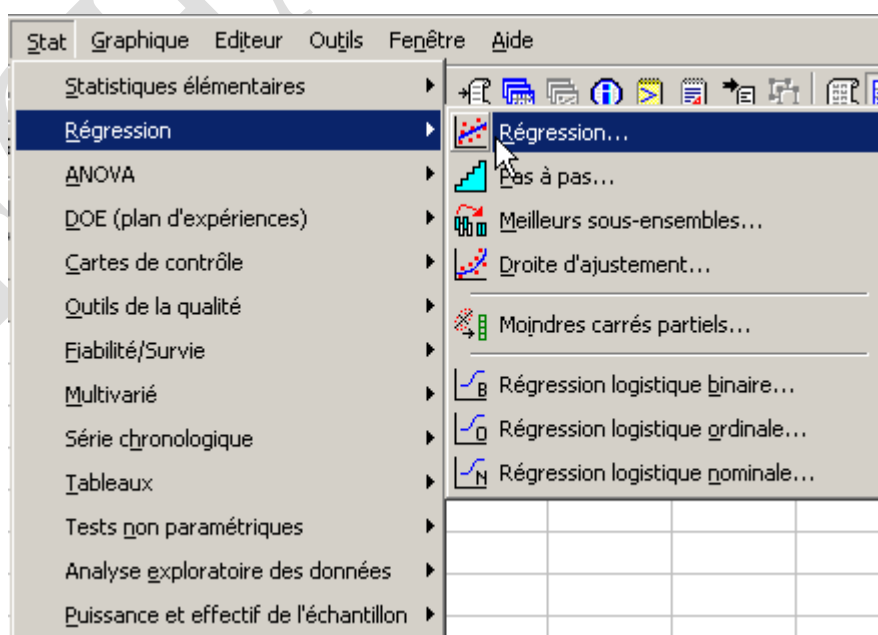
### 13.1. Exercice 96.: Effectuer une régression linéaire simple (modèle linéaire Gaussien)

Minitab® Statistical Software 15.1.1.0

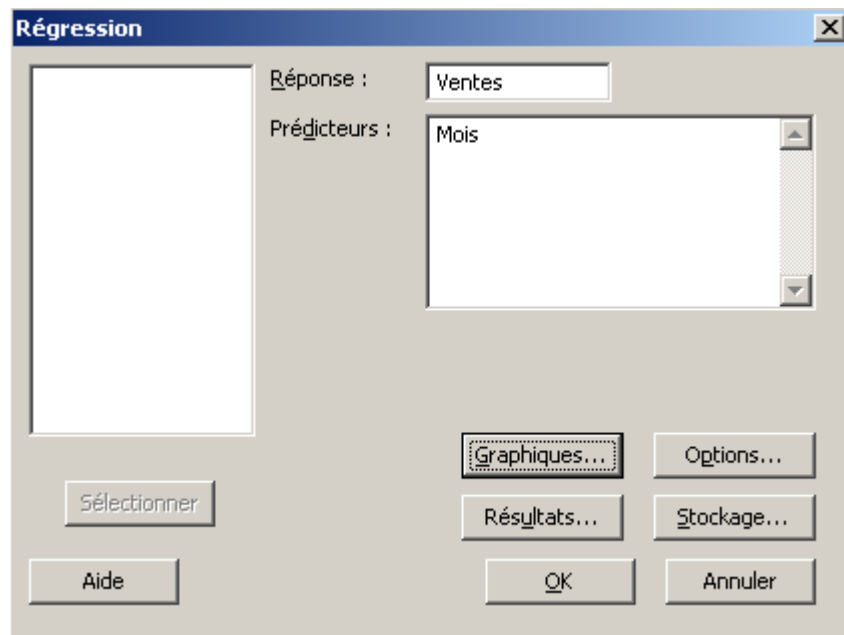
Ouvrez le fichier *RegressionLinaire.mpj* qui donne le nombre d'un certain type de produit vendu en fonction du mois depuis qu'une entreprise a été créée (vous remarquerez que certains mois, le produit en question n'a pas été vendu). Sur l'hypothèse que la relation est linéaire, faisons une analyse par le modèle linéaire Gaussien conformément aux démonstrations effectuées dans le cours de statistique théorique et des exemples faits avec Microsoft Excel (afin de voir les éventuelles différences):

	C1	C2
	Mois	Ventes
1	3	4
2	6	9
3	7	12
4	8	15
5	9	17
6	10	16
7	11	17
8	12	18
9	15	18

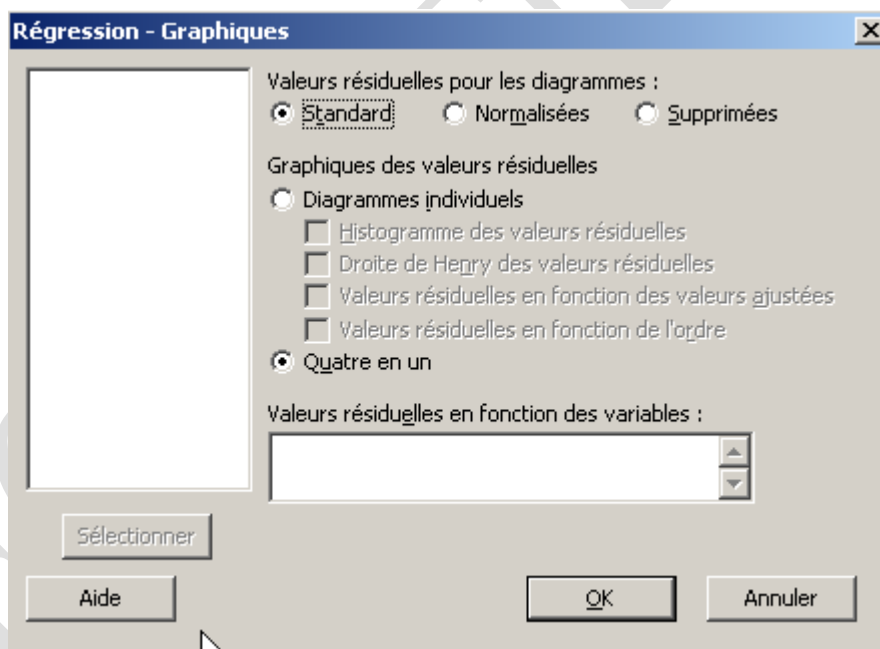
Ensuite, nous allons dans le menu **Stat/Régression/Régression**:



et nous voyons que nous pouvons faire une régression linéaire multiple mais pour l'instant, nous nous concentrerons sur la version simple (à un variable explicative: une covariable):

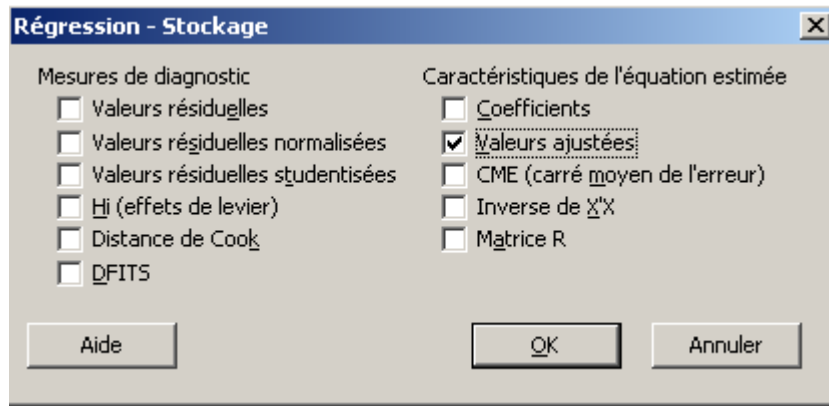


Ensuite, nous cliquons sur le bouton **Graphiques...**:



Pour y cocher **Quatre en un**, ensuite dans la boîte principale, nous cliquons sur le bouton **Stockage...** pour dire que nous souhaiterions voir dans la feuille Minitab® Statistical Software les valeurs calculées à partir du modèle:





Nous validons deux fois par **OK** pour obtenir dans la fenêtre de session les informations suivantes (correspondant à Microsoft Excel mais avec quelques informations supplémentaires)

L'équation de régression est  
 Ventes = 3.02 + 1.22 Mois

Prédicteur	Coeff	Coef ErT	T	P
Constante	3.020	2.151	1.40	0.203
Mois	1.2200	0.2241	5.44	0.001

S = 2.24117    R carré = 80.9 %    R carré (ajust) = 78.2 %

Analyse de variance

Source	DL	Somme des carrés	CM	F	P
Régression	1	148.84	148.84	29.63	0.001
Erreur résiduelle	7	35.16	5.02		
Total	8	184.00			

Observations aberrantes

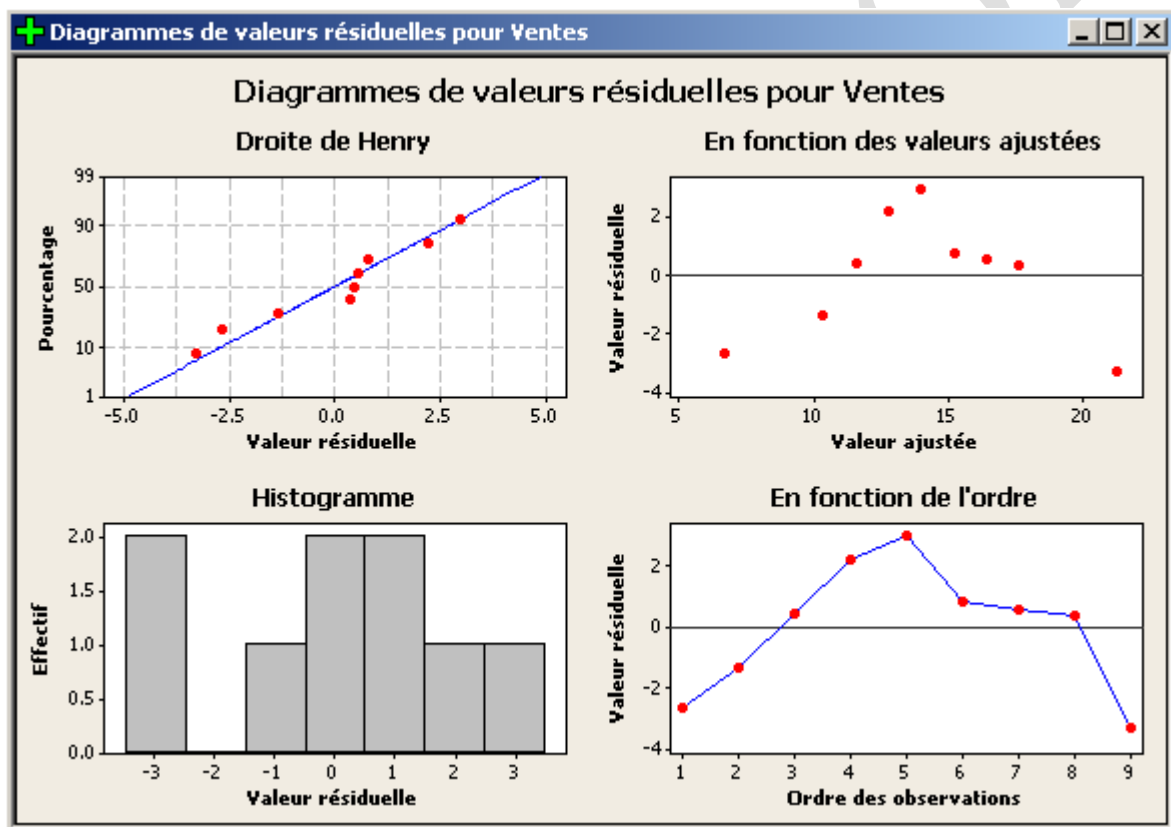
Observation	Mois	Ventes	Valeur ajustée	Ajust ErT	Valeur résiduelle	Valeur résiduelle normalisée
9	15.0	18.000	21.320	1.538	-3.320	-2.04R

R indique une observation ayant une valeur résiduelle normalisée importante :

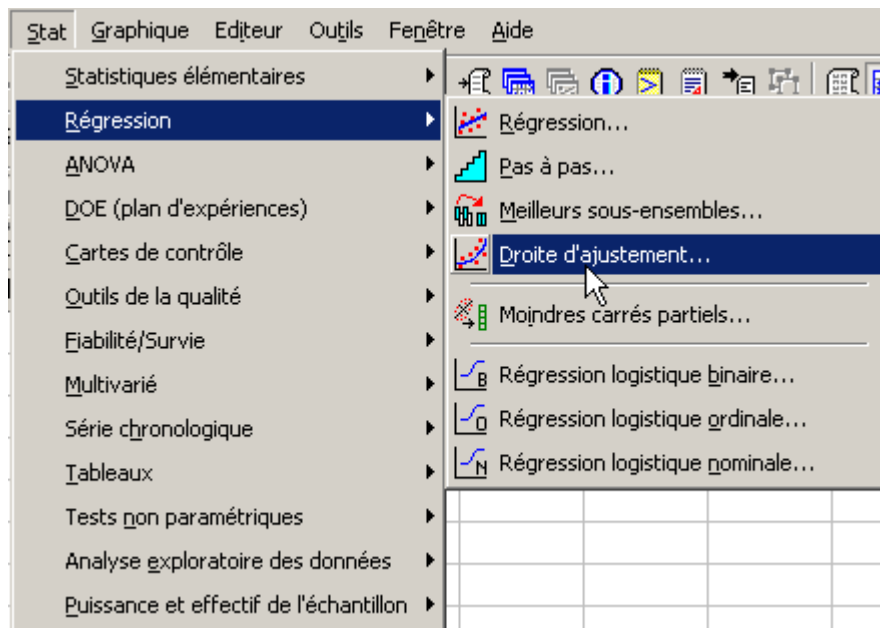
dans la feuille nous avons bien les valeurs ajustées qui ont été stockées (comme demandé):

↓	C1	C2	C3
	Mois	Ventes	AJUSTEES1
1	3	4	6.68
2	6	9	10.34
3	7	12	11.56
4	8	15	12.78
5	9	17	14.00
6	10	16	15.22
7	11	17	16.44
8	12	18	17.66
9	15	18	21.32

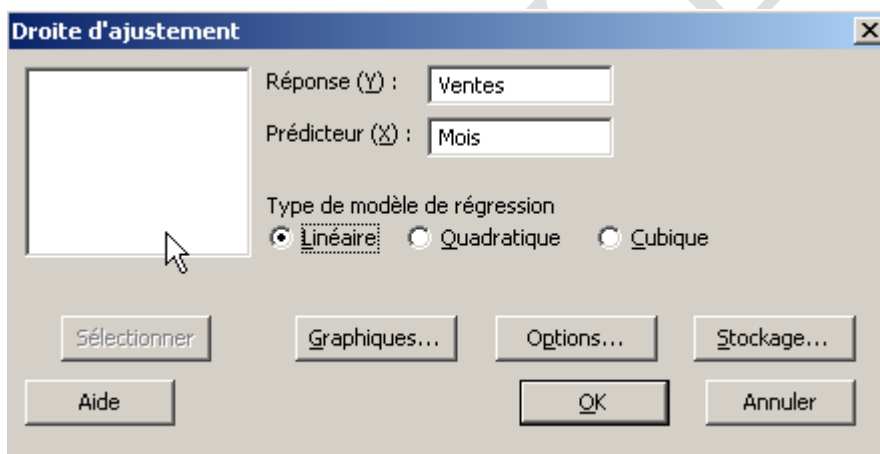
et le graphique suivant sur les résidus:



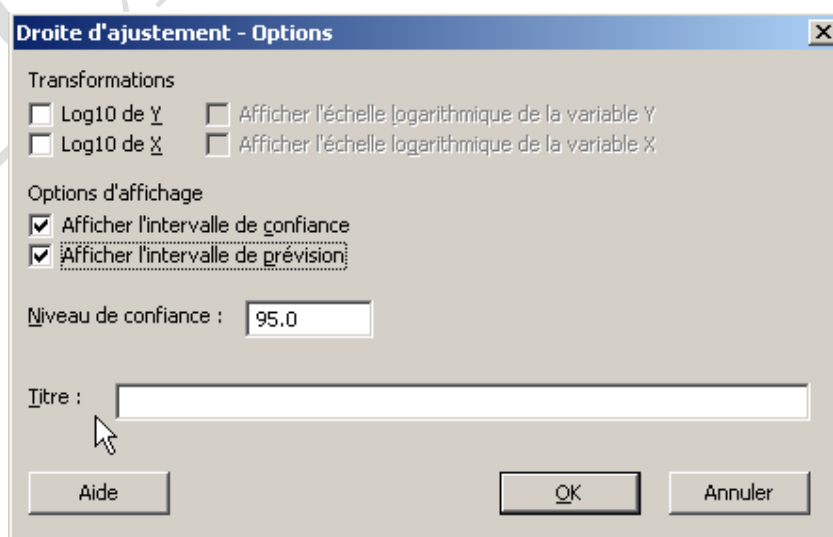
Mais qu'en est-il de la droite de régression??? Cet outil ne le donne pas visuellement. Pour ce faire, il nous faut aller dans le menu **Stats/Régression/Droite d'ajustement...**:



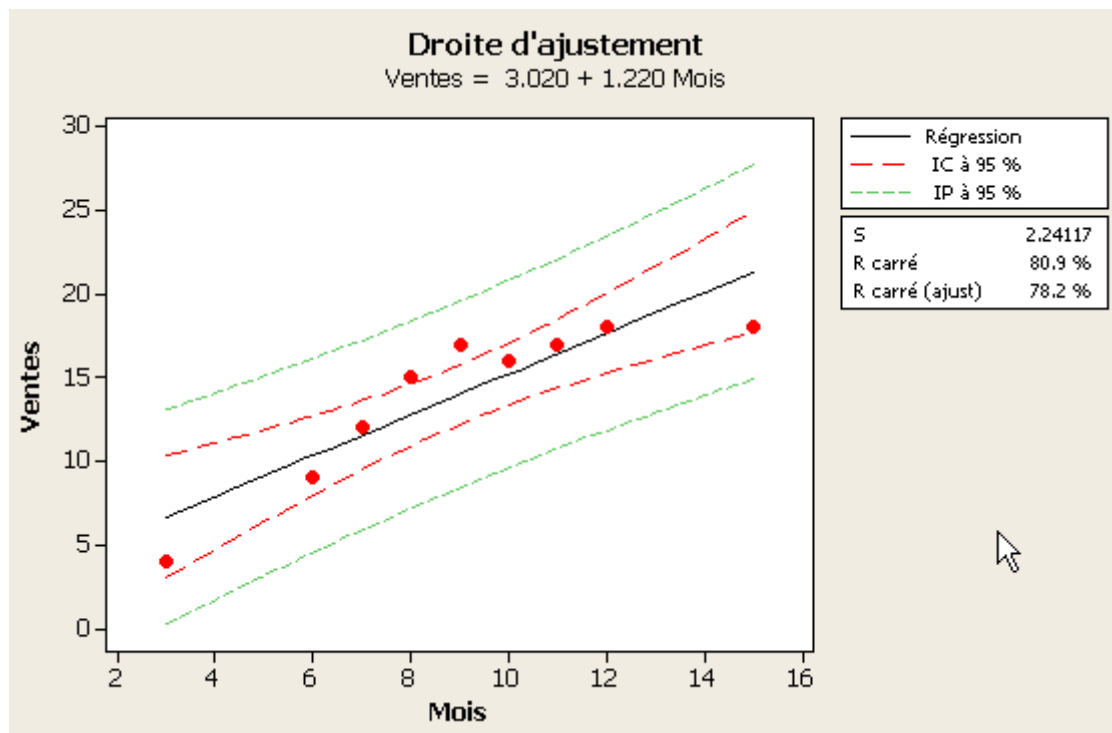
Nous y prenons:



Nous cliquons sur le bout **Options...** pour activer l'intervalle de confiance et l'intervalle de prévision vu dans le cours de statistique théorique:



et nous validons deux fois par **OK** pour obtenir:



Soit la même chose qu'avec Microsoft Excel mais beaucoup plus rapidement (pas besoin de formules).

Pour les intervalles de confiance des coefficients que l'on peut obtenir facilement qu'à partir de la version 16 ou 17 de Minitab, reportez-vous à l'exercice de la page 493.

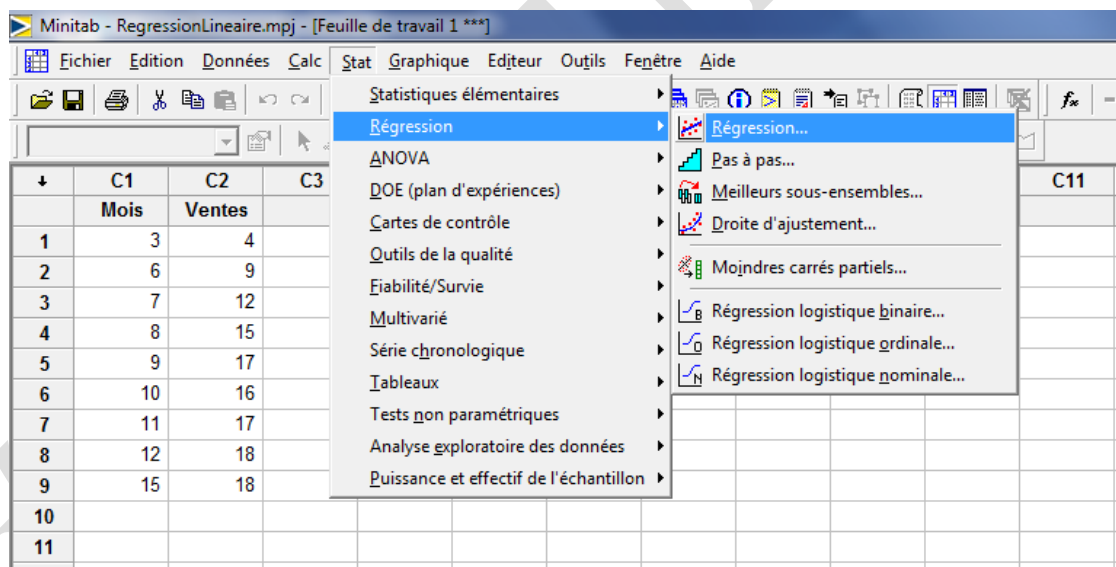
### 13.1.1. Résidus, résidus normalisés, résidus studentisés

Dans le cours théorique, nous avons parlé des 4 types de résidus dans le cours théorique et non sans mal, nous les avons calculés à la main pour une régression linéaire simple. Voyons donc si nous retombons sur les mêmes valeurs.

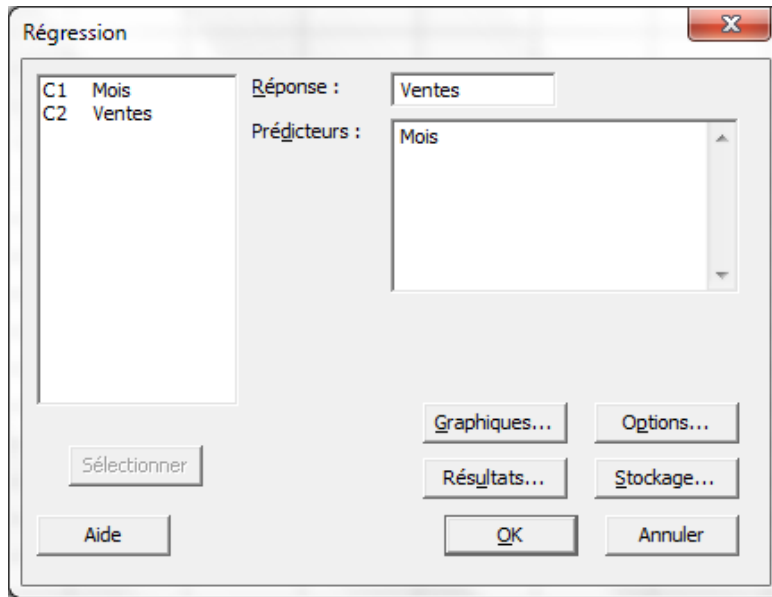
Nous partons toujours des valeurs suivantes:

↓	C1	C2
	Mois	Ventes
1	3	4
2	6	9
3	7	12
4	8	15
5	9	17
6	10	16
7	11	17
8	12	18
9	15	18

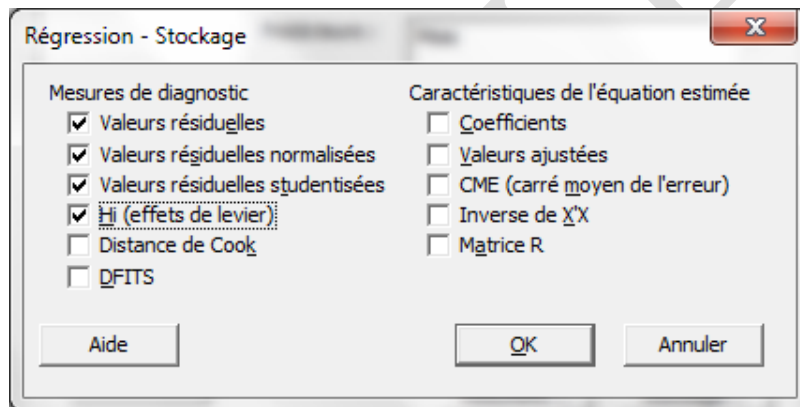
Nous allons dans le menu **Stat/Régression/Régression...**:



Nous mettons:



Nous cliquons sur le bouton **Options...** pour y cocher (remarquez que les valeurs résiduelles supprimées ne sont pas disponibles!):



Nous validons deux fois par **OK** pour obtenir (notez que la 3<sup>ème</sup> colonne est mal nommé... puisqu'il s'agit en réalité des résidus studentisés!):

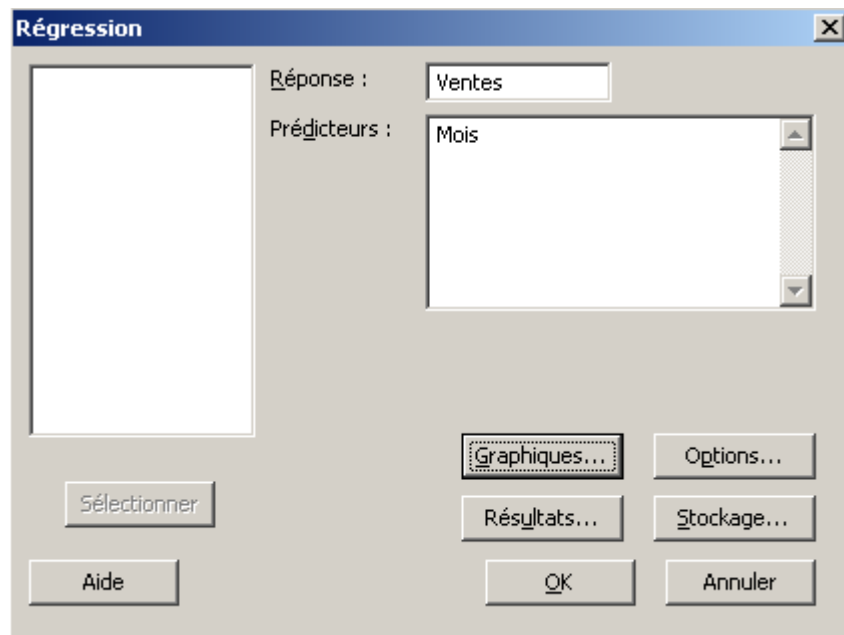
	C1	C2	C3	C4	C5	C6	C7	C8
	Mois	Ventes	RESIDUELLE1	NORMALISEE1	T SUPPR1	HI1		
1	3	4	-2.68	-1.64429	-1.94314	0.471111		
2	6	9	-1.34	-0.66894	-0.64011	0.201111		
3	7	12	0.44	0.21308	0.19792	0.151111		
4	8	15	2.22	1.05660	1.06700	0.121111		
5	9	17	3.00	1.41978	1.55776	0.111111		
6	10	16	0.78	0.37124	0.34713	0.121111		
7	11	17	0.56	0.27120	0.25241	0.151111		
8	12	18	0.34	0.16973	0.15746	0.201111		
9	15	18	-3.32	-2.03695	-2.95509	0.471111		
10								
11								

Nous tombons bien sur les valeurs calculées à la main dans le cours théorique!

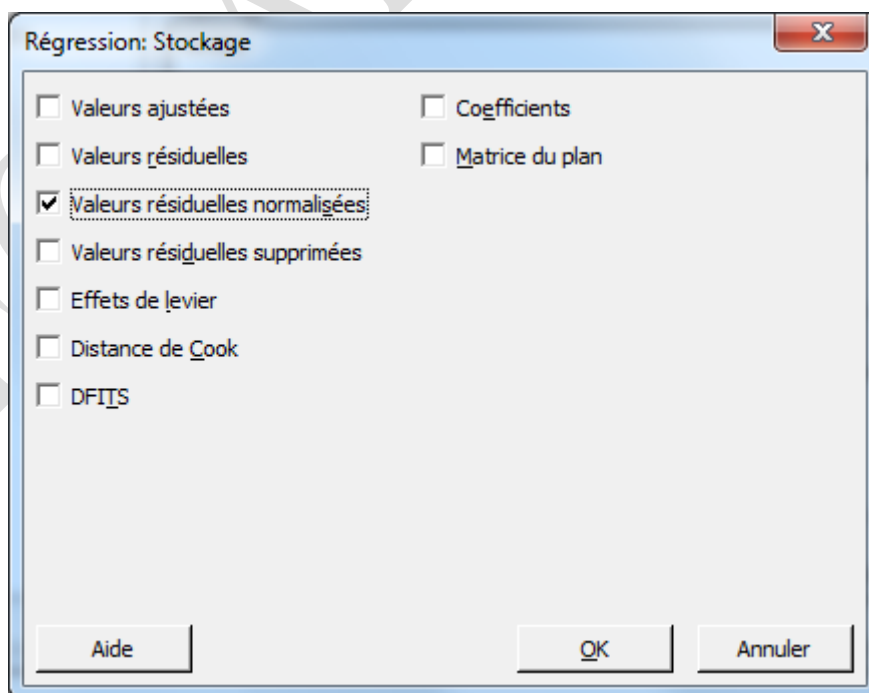
### 13.1.2. Test de l'homoscédasticité des résidus

Si nous avons l'hypothèse d'homoscédasticité et de normalité des résidus qui est satisfaite comme l'exige le modèle linéaire gaussien, alors faire une régression linéaire des résidus devrait nous donner une ordonnée à l'origine nulle avec une pente nulle (ou du moins des valeurs non significatives dans les deux cas).

Comme Minitab ne nous propose directement aucun outil pour cela, nous allons analyser nous-mêmes les résidus en faisant une analyse de la régression de ces derniers:



Nous cliquons sur **Stockage...** pour cocher **Valeurs résiduelles normalisées**:

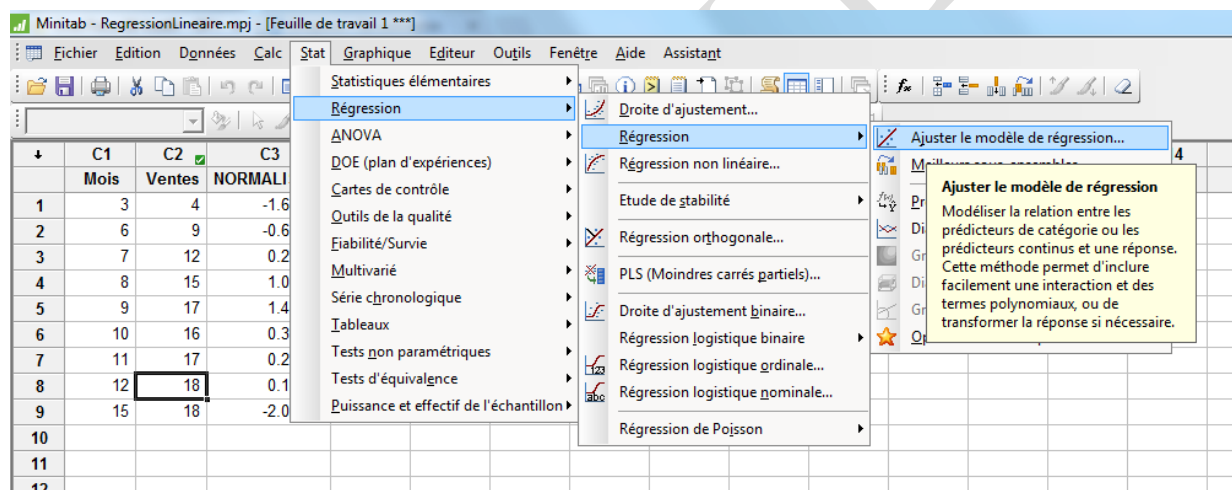


Nous avons alors:

↓	C1	C2 ✓	C3
	Mois	Ventes	NORMALISEE1
1	3	4	-1.64429
2	6	9	-0.66894
3	7	12	0.21308
4	8	15	1.05660
5	9	17	1.41978
6	10	16	0.37124
7	11	17	0.27120
8	12	18	0.16973
9	15	18	-2.03695
10			

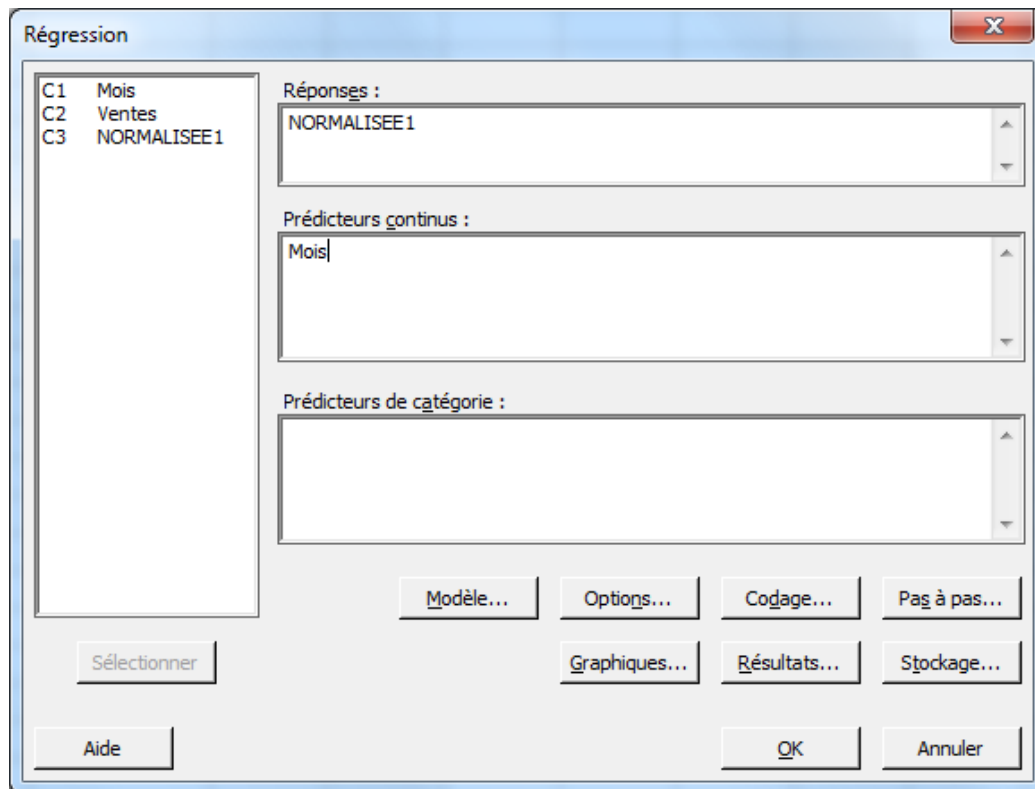
Ensuite, nous allons refaire une régression en passant par le menu (je poursuis avec Minitab 17 car c'est ce que j'avais sous la main lors de la création de ces captures d'écran)

**Stat/Régression/Régression/Ajouter le modèle de régression...** mais cette fois-ci en utilisant les résidus normalisés comme variable à expliquer:

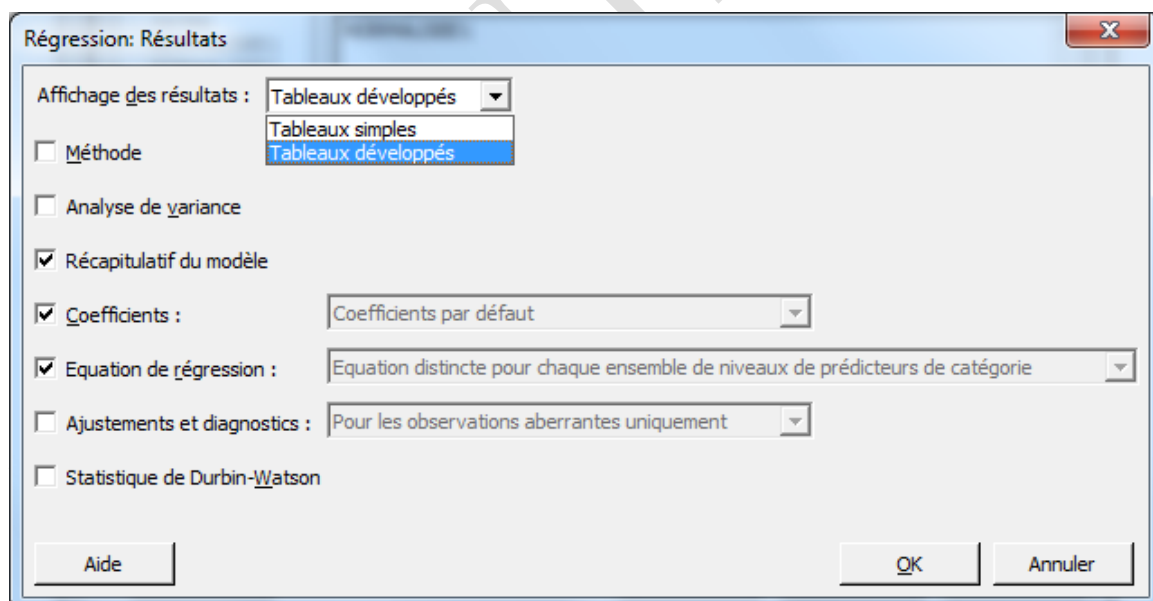


Nous prenons donc:





Nous cliquons sur le bouton **Résultats...** pour prendre **Tableaux développés** comme cela nous aurons aussi les intervalles de confiance des coefficients de la régression:



Nous validons le tout par **OK** pour obtenir:

### Analyse de régression : NORMALISEE1 en fonction de Mois

Récapitulatif du modèle

S	R carré	R carré (ajust)	SomCar-ErrPrév	R carré (prév)
1.23267	0.02%	0.00%	27.8827	0.00%

Coefficients

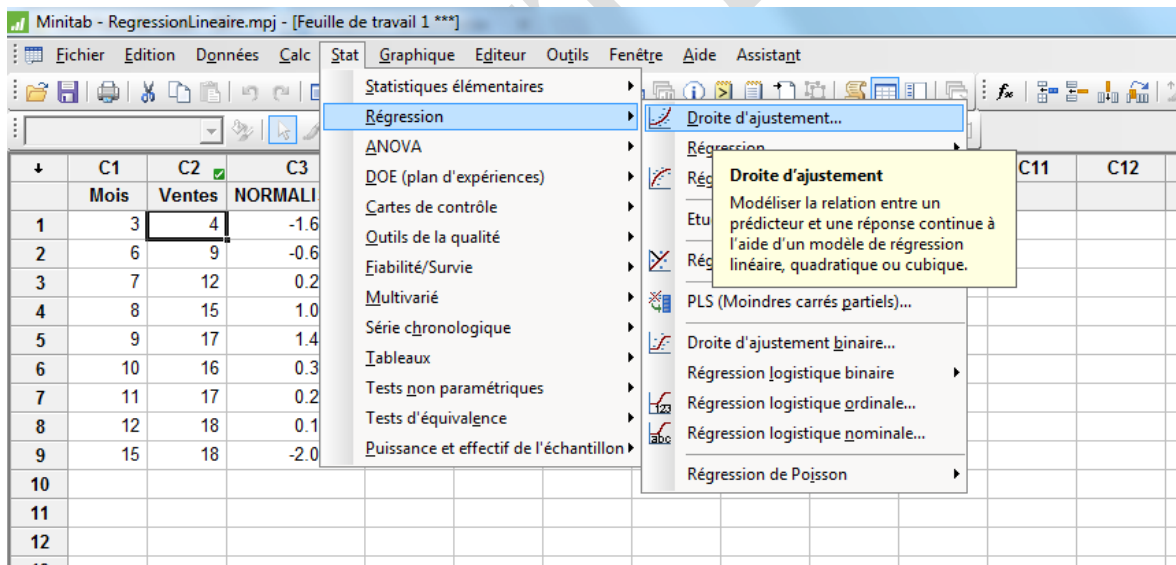
Terme	Coeff	Coef ErT	IC à 95 %	Valeur de T	Valeur de p	FIV
Constante	-0.06	1.18	( -2.85; 2.74)	-0.05	0.963	
Mois	-0.004	0.123	(-0.296; 0.287)	-0.03	0.974	1.00

Equation de régression

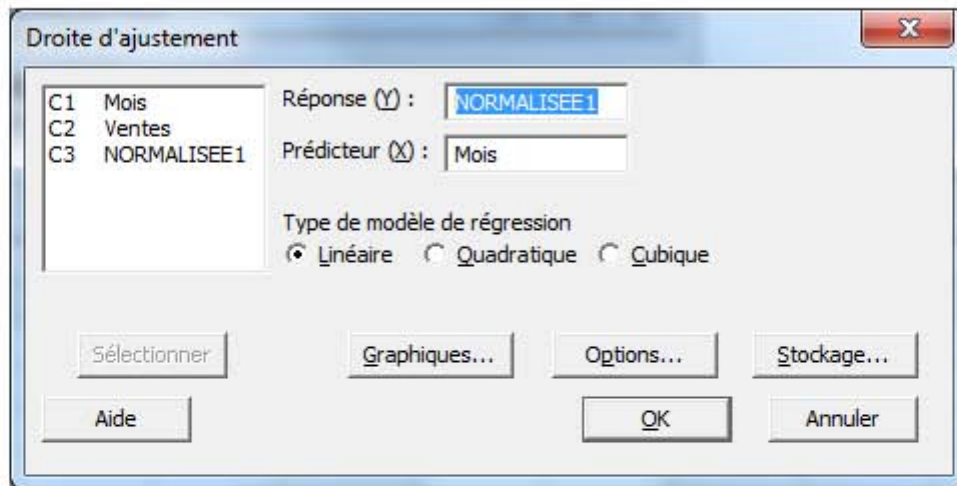
$$\text{NORMALISEE1} = -0.06 - 0.004 \text{ Mois}$$

Donc a posteriori l'homoscédasticité est bien vérifiée!

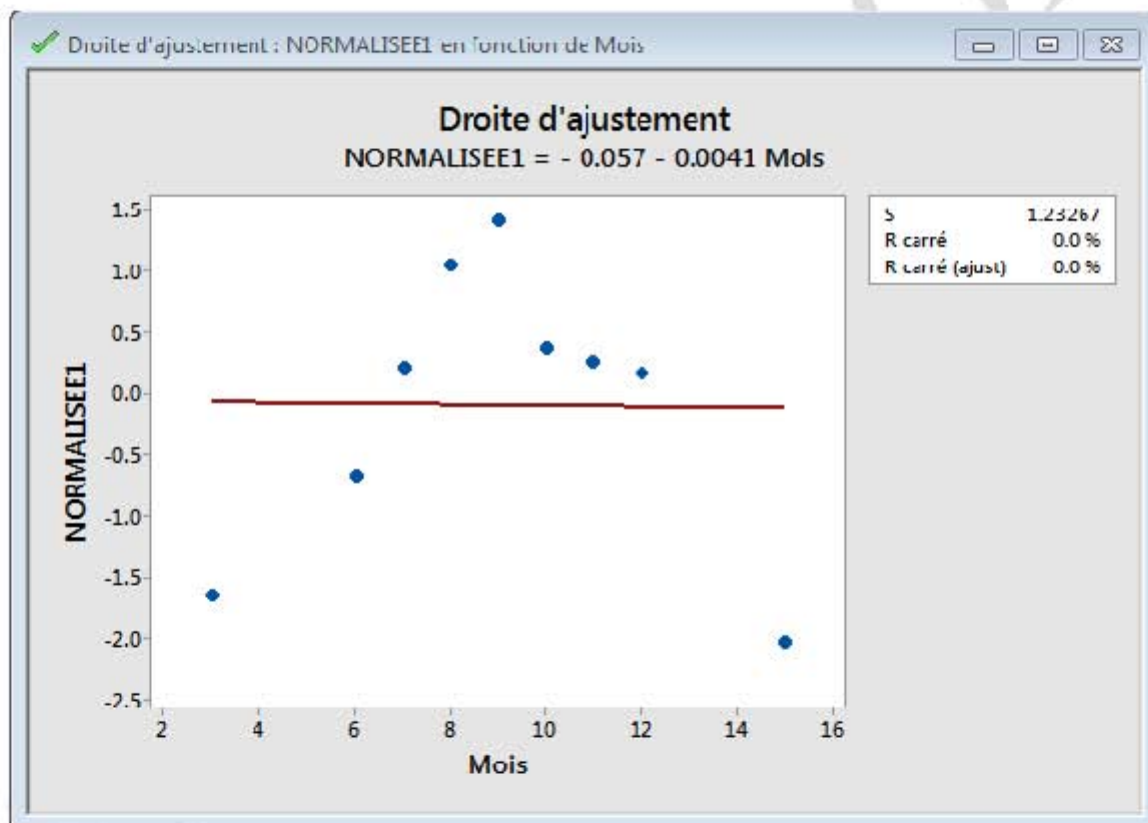
Ensuite si nous voulons tracer la droite de régression des résidus, nous allons dans le menu **Stat/Régression/Droite d'ajustement...**:



et nous prenons:

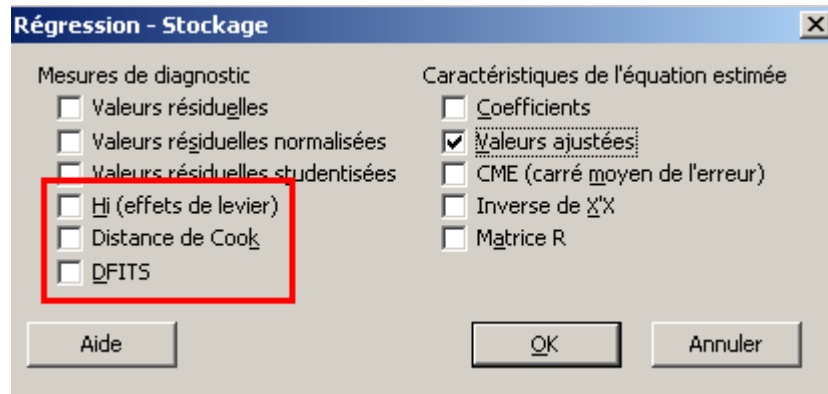


Pour obtenir au final:



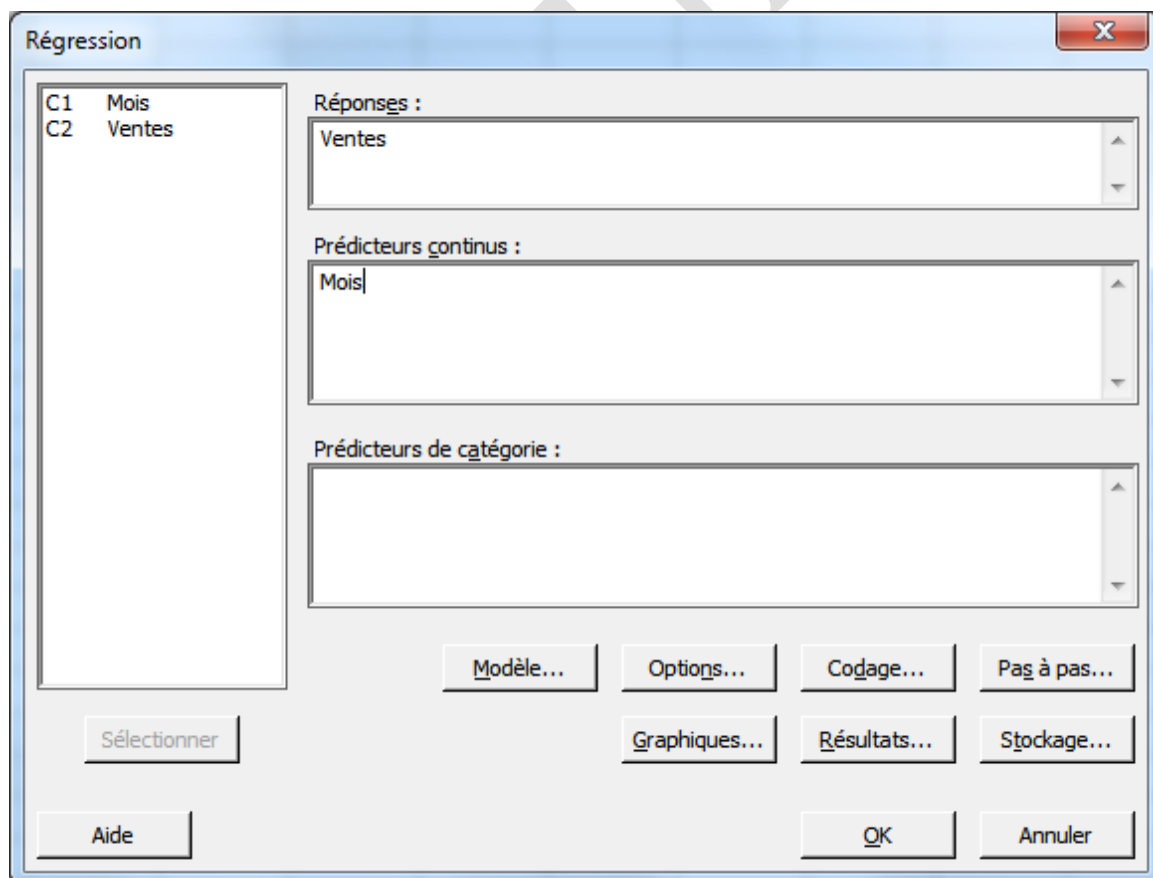
### 13.1.3. Effets de levier des points (valeurs influentes)

Dans le cours théorique nous avons mentionné qu'il existait de nombreuses techniques empiriques pour mesurer l'influence d'un point. Minitab propose les classiques mis en évidence ci-dessous:

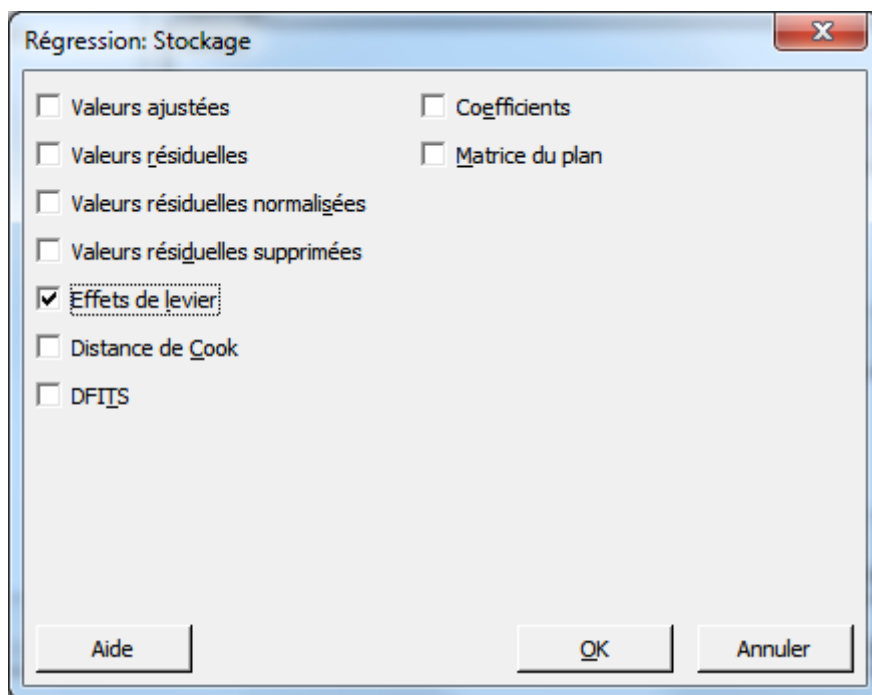


Dans le cadre de ce présent support de cours nous allons nous concentrer sur le seul dont nous avons démontré l'origine mathématique détaillée dans le cours théorique et dont de toute façon la Distance de Cook et le DFITS (DiFFerence In fiTS) découlent: l'effet de levier Hi!

Nous retournons donc dans l'outil de régression:



et dans le bouton **Stockage...**, nous cochons **Effets de levier**:



Quand nous validons le tout par **OK**, nous obtenons:

Minitab - RegressionLineaire.mpj - [Feuille de trava

Fichier Edition Données Calc Stat Gra

↓	C1	C2 ✓	C3	C4
	Mois	Ventes	HI1	
1	3	4	0.471111	
2	6	9	0.201111	
3	7	12	0.151111	
4	8	15	0.121111	
5	9	17	0.111111	
6	10	16	0.121111	
7	11	17	0.151111	
8	12	18	0.201111	
9	15	18	0.471111	
10				

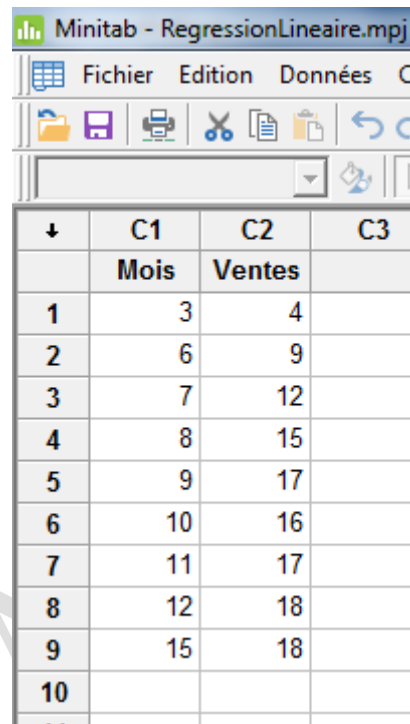
Ce qui est conforme!

## 13.2. Exercice 97.: Régression linéaire simple forcée à l'origine

Minitab® Statistical Software 17.1.0

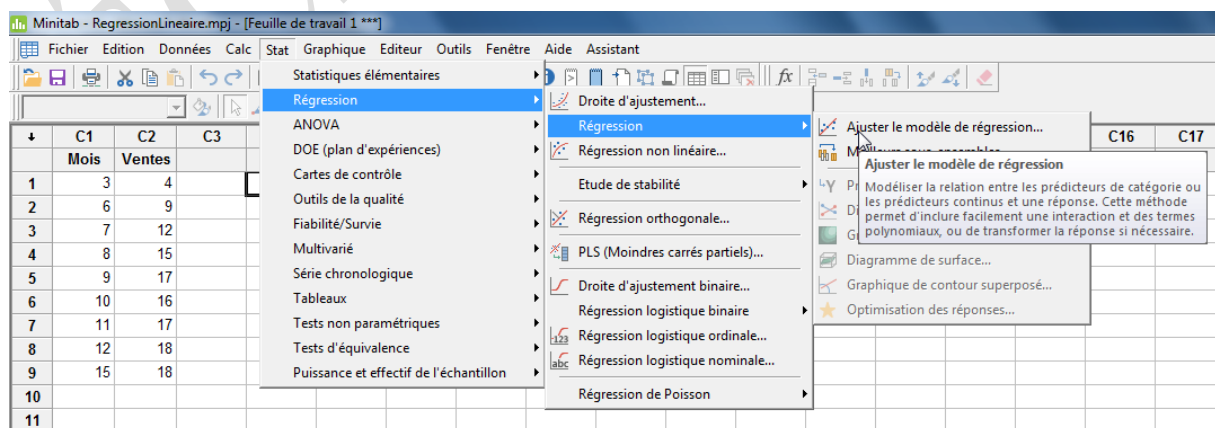
Le fait de forcer la régression à l'origine est très important dans la pratique mais comme nous l'avons étudié dans le cours théorique, c'est également un sujet à controverses. Nous allons vérifier ici si Minitab calcule le coefficient de corrélation conformément à la définition alternative de ce dernier dans le cas où la régression est forcée à l'origine.

Voyons donc comment effectuer une telle régression avec le même jeu de données qu'avant:

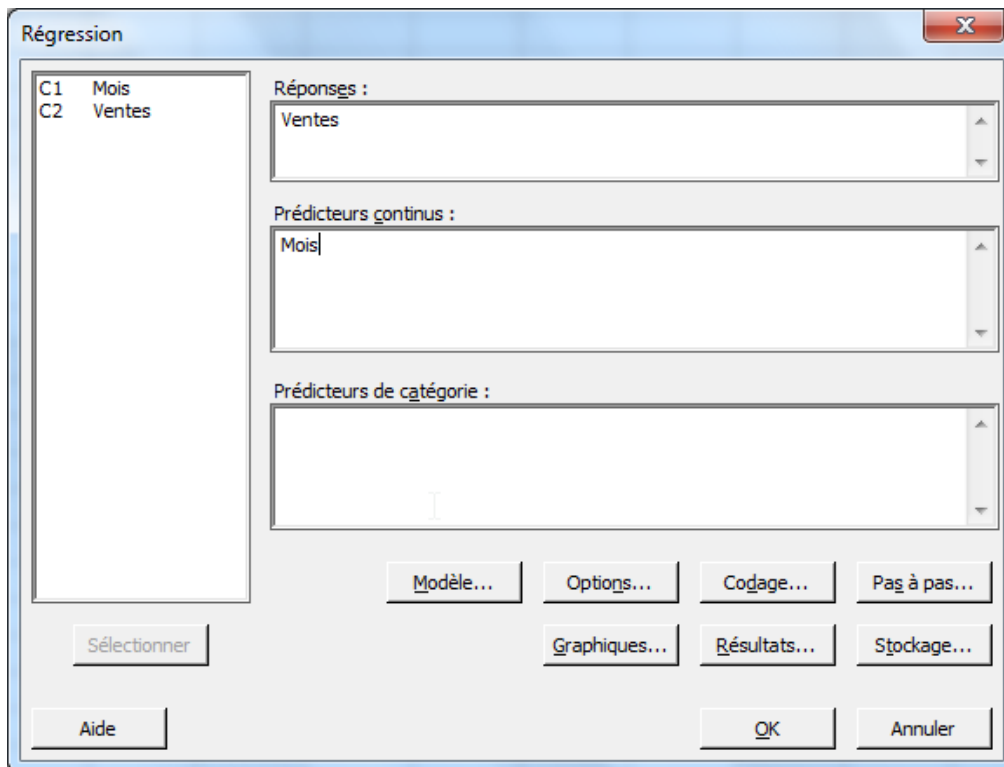


	C1	C2	C3
	Mois	Ventes	
1	3	4	
2	6	9	
3	7	12	
4	8	15	
5	9	17	
6	10	16	
7	11	17	
8	12	18	
9	15	18	
10			
11			

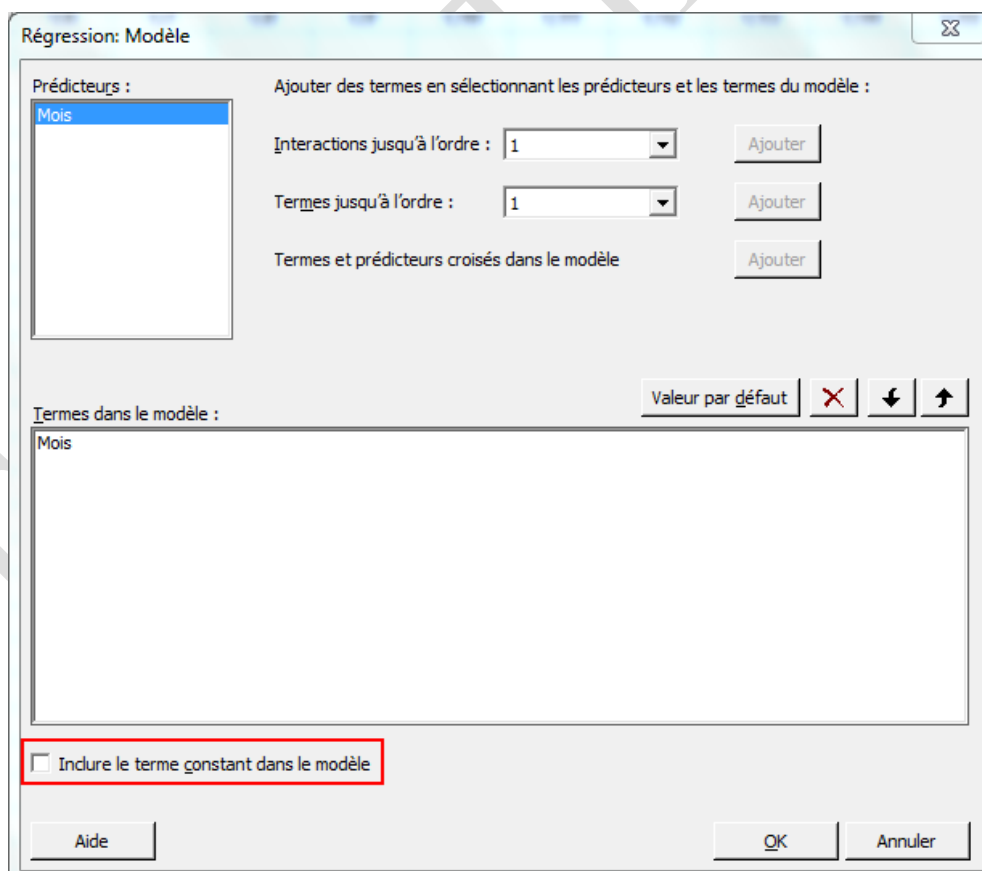
Et nous allons dans le menu **Stat/Régression/Régression/Ajuster le modèle de régression....**:



Dans la boîte de dialogue qui apparaît, nous mettons:



Ensuite, nous cliquons sur le bouton **Modèle...** :



Pour décocher l'option **Inclure le terme constant dans le modèle**. Nous validons deux fois par **OK** pour obtenir:

## Analyse de régression : Ventes en fonction de Mois

### Analyse de variance

Source	DL	SomCar		Valeur	
		ajust	CM ajust	F	de p
Régression	1	1902.94	1902.94	337.84	0.000
Mois	1	1902.94	1902.94	337.84	0.000
Erreur	8	45.06	5.63		
Total	9	1948.00			

### Récapitulatif du modèle

S	R carré	R carré (ajust)	R carré (prév)
2.37333	97.69%	97.40%	96.44%

### Coefficients

Terme	Coeff	Coef ErT	Valeur		FIV
			de T	de p	
Mois	1.5151	0.0824	18.38	0.000	1.00

### Equation de régression

$$\text{Ventes} = 1.5151 \text{ Mois}$$

À comparer avec la sortie donnée par Microsoft Excel:

	A	B	C	D	E	F	G	H	I	J	K
1	RAPPORT DÉTAILLÉ										
2											
3	<i>Statistiques de la régression</i>										
4	Coefficient de détermination multiple	0.98836623									
5	Coefficient de détermination R^2	0.9768678									
6	Coefficient de détermination R^2	0.8518678									
7	Erreur-type	2.37332888									
8	Observations	9									
9											
10	ANALYSE DE VARIANCE										
11		<i>Degré de liberté</i>			<i>mm des carré</i>		<i>me des car</i>		<i>F</i>		<i>eur critique de F</i>
12	Régression	1	1902.93848	1902.93848	337.838312	3.4961E-07					
13	Résidus	8	45.0615199	5.63268999							
14	Total	9	1948								
15											
16		<i>Coefficients</i>									
17	Constante	0	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
18	Mois	1.51507841	0.08242912	18.3803785	7.9029E-08	1.32499652	1.7051603	1.32499652	1.7051603		
19											

À comparer avec R:



```

RGui (64-bit) - [R Console]
File Edit View Misc Packages Windows Help
> mydata<-read.csv("c:/tmp/RegressionLineaireUnivariee.csv",header=T,sep=";")
> mydata
  Mois Ventes
1    3     4
2    6     9
3    7    12
4    8    15
5    9    17
6   10    16
7   11    17
8   12    18
9   15    18
> summary(lm(Ventes~0+Mois,mydata))

Call:
lm(formula = Ventes ~ 0 + Mois, data = mydata)

Residuals:
    Min       1Q   Median       3Q      Max
-4.7262 -0.1809  0.3341  1.3945  3.3643

Coefficients:
      Estimate Std. Error t value Pr(>|t|)
      Mois  1.51508   0.08243   18.38  7.9e-08

Residual standard error: 2.373 on 8 degrees of freedom
Multiple R-squared:  0.9769,    Adjusted R-squared:  0.974
F-statistic: 337.8 on 1 and 8 DF,  p-value: 7.903e-08

> |
    
```

Et dans SPSS 25:

**Model Summary**

Model	R	R Square <sup>b</sup>	Adjusted R Square	Std. Error of the Estimate
1	.988 <sup>a</sup>	.977	.974	2.37333

a. Predictors: Mois  
 b. For regression through the origin (the no-intercept model), R Square measures the proportion of the variability in the dependent variable about the origin explained by regression. This CANNOT be compared to R Square for models which include an intercept.

**ANOVA<sup>a,b</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1902.938	1	1902.938	337.838	.000 <sup>c</sup>
	Residual	45.062	8	5.633		
	Total	1948.000 <sup>d</sup>	9			

a. Dependent Variable: Ventes  
 b. Linear Regression through the Origin  
 c. Predictors: Mois  
 d. This total sum of squares is not corrected for the constant because the constant is zero for regression through the origin.

**Coefficients<sup>a,b</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	Mois	1.515	.082	.988	18.380	.000

a. Dependent Variable: Ventes  
 b. Linear Regression through the Origin

Donc nous voyons que tous les logiciels utilisent la définition alternative du coefficient de corrélation alternatif puisque nous retombons sur la valeur calculée à la main dans le cours théorique de 0.9769!

### 13.3. Exercice 98.: Effectuer une régression linéaire (nouvelles options v16/v17/v18)

Minitab® Statistical Software 17.1.0

---

Il arrive régulièrement que les praticiens de Minitab hésitent à utiliser faire une analyse de plan d'expérience entre les outils de type plans factoriels (peu importe la sous-famille) ou avec des modèles de régression linéaires généralisés.

L'équipe de développement de Minitab voulant faire le choses au plus simple, il proposent alors un outil de régression linéaire depuis la version 17 qui permet d'avoir une modèle de régression un peu analogue à celui d'un plan d'expérience avec presque les mêmes outils d'analyse mais avec cependant un effet moyen seulement et ce par niveau de variable qualitative.

Donc normalement pour des raisons pédagogiques on introduira les analyses dans l'ordre suivant:

1. Analyse de réponses par régression linéaire simple avec ou sans interactions avec ANOVA, surface de réponse, optimisation des réponses, etc. (ces dernières options étant nouvelles dans la v17).
2. Modèle linéaires généralisées (classique mais plus complexe)
3. Plans d'expériences (une des très nombreuses familles disponibles de plans d'expérience qui nécessite une longue formation pour bien comprendre les limites pratiques des modèles mathématiques sous-jacents).

Pour jouer avec les nouveaux outils d'analyse de Minitab, nous prendrons le fichier suivant *ANOVA\_Balancee.mpj*:

Minitab - ANOVA\_Balancede.mpj - [Worksheet 2 \*\*\*]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assistant

	C1-T	C2	C3-T	C4	C5	C6	C7	C8	C9	C10	C
	Lieu	Vitesse	Pression	Response							
1	Ville	40	1.5	32700							
2	Ville	40	2	33430							
3	Ville	40	2.5	31710							
4	Ville	50	1.5	32680							
5	Ville	50	2	34070							
6	Ville	50	2.5	33220							
7	Route	40	1.5	33180							
8	Route	40	2	34430							
9	Route	40	2.5	33570							
10	Route	50	1.5	33270							
11	Route	50	2	33440							
12	Route	50	2.5	32840							
13	Ville	40	1.5	32750							
14	Ville	40	2	33360							
15	Ville	40	2.5	32100							
16	Ville	50	1.5	32270							
17	Ville	50	2	33100							
18	Ville	50	2.5	33700							
19	Route	40	1.5	32160							
20	Route	40	2	34280							
21	Route	40	2.5	33300							
22	Route	50	1.5	33080							
23	Route	50	2	33570							
24	Route	50	2.5	33210							
25	Ville	40	1.5	32960							

Feuille de travail en cours : Worksheet 2

Maintenant, nous allons dans le menu **Stat/Régressions/Ajuster le modèle de régression...**:

Minitab - ANOVA\_Balancede.mpj - [Worksheet 2 \*\*\*]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assistant

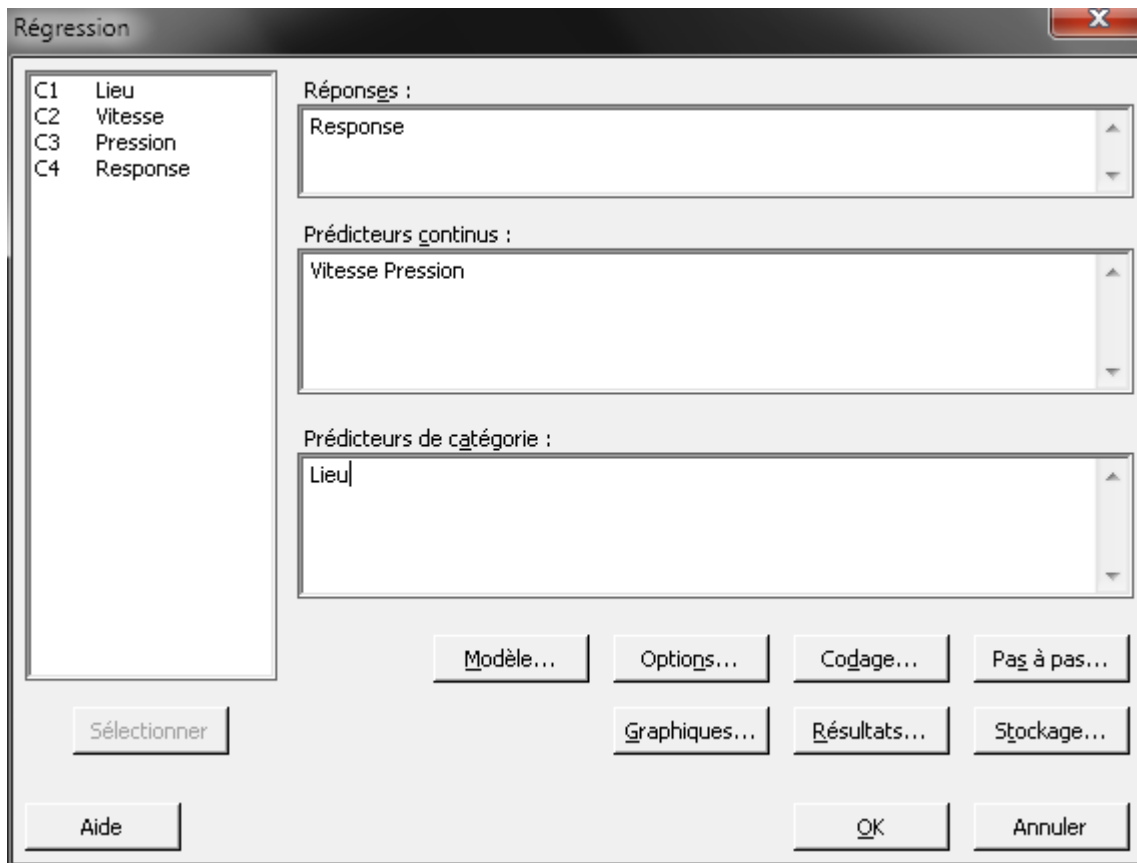
Stat

- Statistiques élémentaires
- Régression
  - Droite d'ajustement...
  - Régression
    - Ajuster le modèle de régression...
    - Meilleurs sous-ensembles
    - Prevoir...
    - Diagrammes fact
    - Graphique de cor
    - Diagramme de su
    - Graphique de cor
    - Optimisation des
  - Régression non linéaire...
  - Etude de stabilité
  - Régression orthogonale...
  - PLS (Moindres carrés partiels)...
  - Droite d'ajustement binaire...
  - Régression logistique binaire
  - Régression logistique ordinale...
  - Régression logistique nominale...
  - Régression de Poisson
- ANOVA
- DOE (plan d'expériences)
- Cartes de contrôle
- Outils de la qualité
- Fiabilité/Survie
- Multivarié
- Série chronologique
- Tableaux
- Tests non paramétriques
- Tests d'équivalence
- Puissance et effectif de l'échantillon

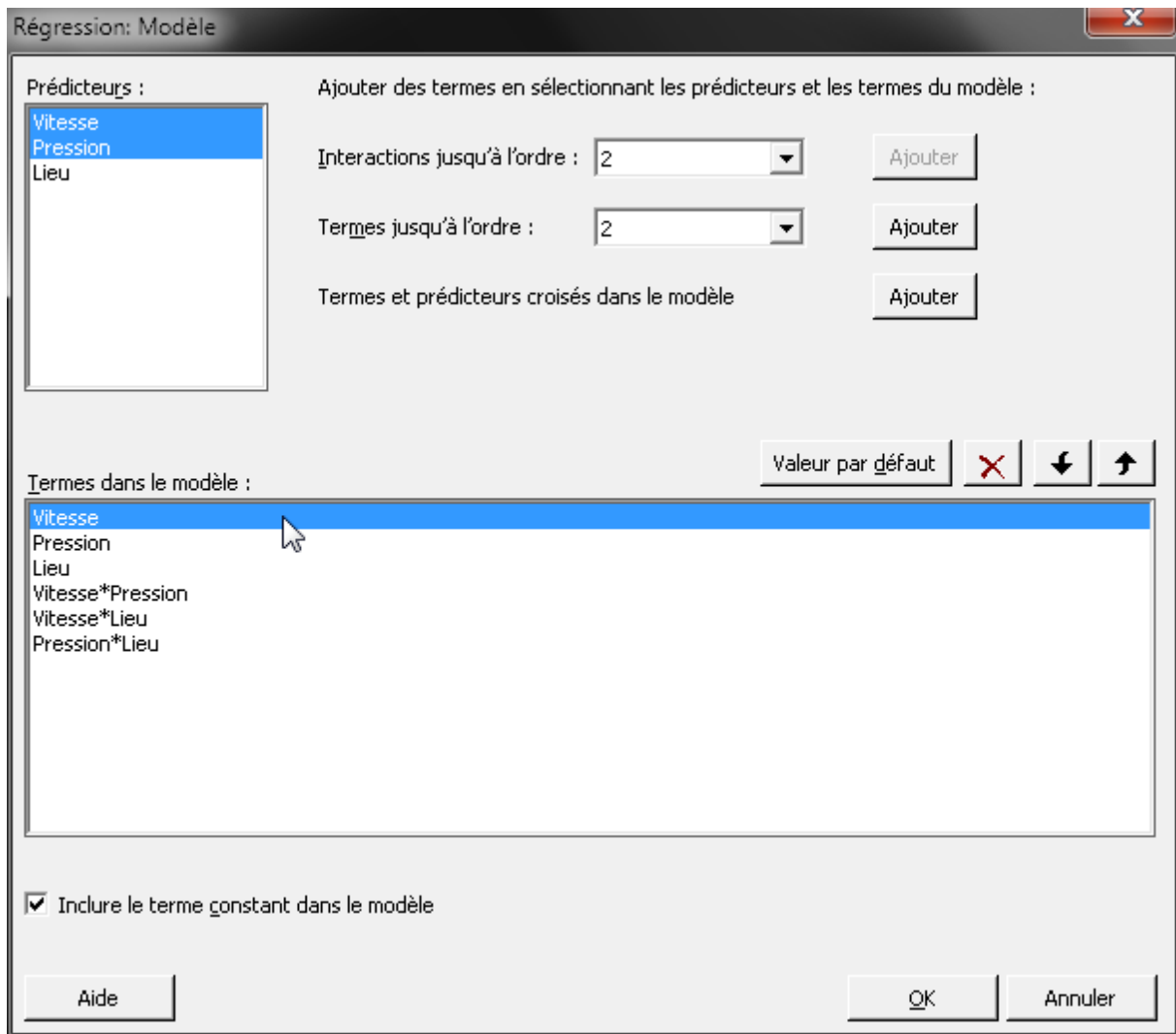
Tableau de données visible en arrière-plan :

	C1-T	C2	C3-T
	Lieu	Vitesse	Pression
1	Ville	40	1.5
2	Ville	40	2
3	Ville	40	2.5
4	Ville	50	1.5
5	Ville	50	2
6	Ville	50	2.5
7	Route	40	1.5
8	Route	40	2
9	Route	40	2.5
10	Route	50	1.5

Dans la boîte de dialogue qui apparaît à l'écran, nous mettons:



Ensuite, nous cliquons sur le bouton **Modèle...** pour y mettre ajouter les mêmes termes d'interactions que dans le cours théorique sur les plans d'expérience (n'oubliez pas que vous devez sélectionner plusieurs facteurs en même temps en haut à gauche):



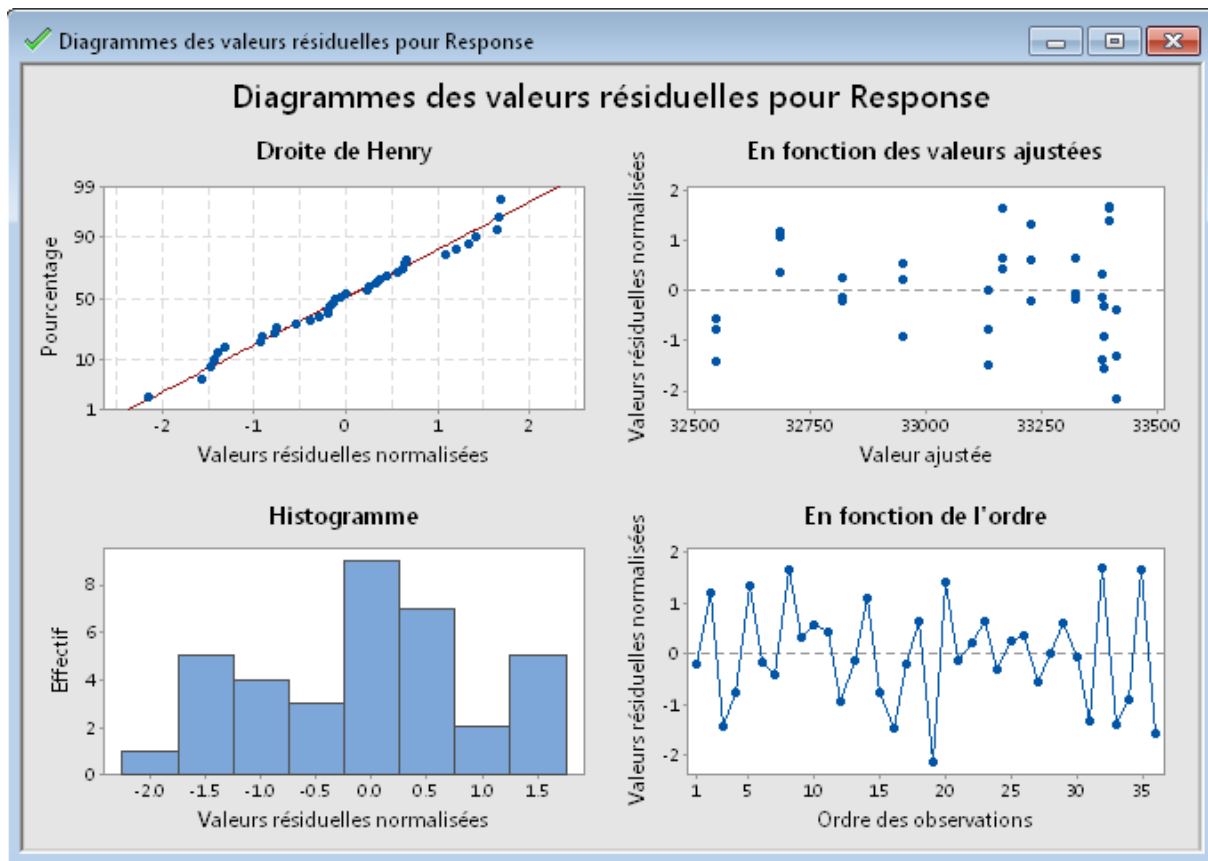
Ensuite de retour dans la boîte de dialogue principale, nous allons dans le bouton **Graphique...**:



Et dans le bouton **Stockage...** de la boîte de dialogue principale, nous prenons:



Et nous validons le tout deux fois par **OK** pour d'abord regarder le diagramme des valeurs résiduelles pour vérifier les conditions d'applications du modèle sous-jacent sont acceptables (c'est le cas):



Ensuite, dans la table, nous trouvons les coefficients des modèles (nous y reviendrons car la sortie n'est pas intuitive):

+	C1-T	C2	C3	C4 <input checked="" type="checkbox"/>	C5
	Lieu	Vitesse	Pression	Response	COEFF1
1	Ville	40	1.5	32700	38123.2
2	Ville	40	2.0	33430	-116.6
3	Ville	40	2.5	31710	-1897.5
4	Ville	50	1.5	32680	0.0
5	Ville	50	2.0	34070	-3344.9
6	Ville	50	2.5	33220	46.7
7	Route	40	1.5	33180	0.0
8	Route	40	2.0	34430	78.0
9	Route	40	2.5	33570	0.0
10	Route	50	1.5	33270	-245.0
11	Route	50	2.0	33440	
12	Route	50	2.5	32840	
13	Ville	40	1.5	32750	
14	Ville	40	2.0	33360	

Dans la fenêtre de session nous nous retrouvons avec une quantité kilométrique d'informations:

**Analyse de régression : Response en fonction de Vitesse; Pression; Lien**

Méthode

Codage des prédicteurs de catégorie (1: 0)

Analyse de variance

Source	Df	Somcar ajust	CM ajust	Valeur F	de p	Valeur
Régression	6	3007251	501208	1.13	0.370	
Vitesse	1	470347	470347	1.06	0.312	
Pression	1	260278	260278	0.59	0.450	
Lien	1	949948	949948	2.14	0.154	
Vitesse*Pression	1	326667	326667	0.74	0.398	
Vitesse*Lien	1	1367340	1367340	3.08	0.090	
Pression*Lien	1	90037	90037	0.20	0.656	
Erreur	29	1286137	443487			
Inadéquation de l'ajustement	5	9459286	1891857	13.35	0.000	
Erreur pure	24	3401851	141744			
Total	35	15868387				

Récapitulatif du modèle

R carré R carré S R carré (ajust) (prév) 2.18% 0.00% 18.95%





## Coefficients

Terme	Coef	Coef ErT	Valeur de T	Valeur de p	FIV
Constante	38123	5154	7.40	0.000	
Vitesse	-117	113	-1.03	0.312	26.00
Pression	-1898	2477	-0.77	0.450	83.00
Lieu					
Ville	-3345	2285	-1.46	0.154	106.00
Vitesse*Pression	46.7	54.4	0.86	0.398	106.00
Vitesse*Lieu					
Ville	78.0	44.4	1.76	0.090	83.00
Pression*Lieu					
Ville	-245	544	-0.45	0.656	26.00

## Equation de régression

Lieu

Route Response = 38123 - 117 Vitesse - 1898 Pression + 46.7 Vitesse\*Pression

Ville Response = 34778 - 39 Vitesse - 2143 Pression + 46.7 Vitesse\*Pression

## Ajustements et diagnostics pour les observations aberrantes

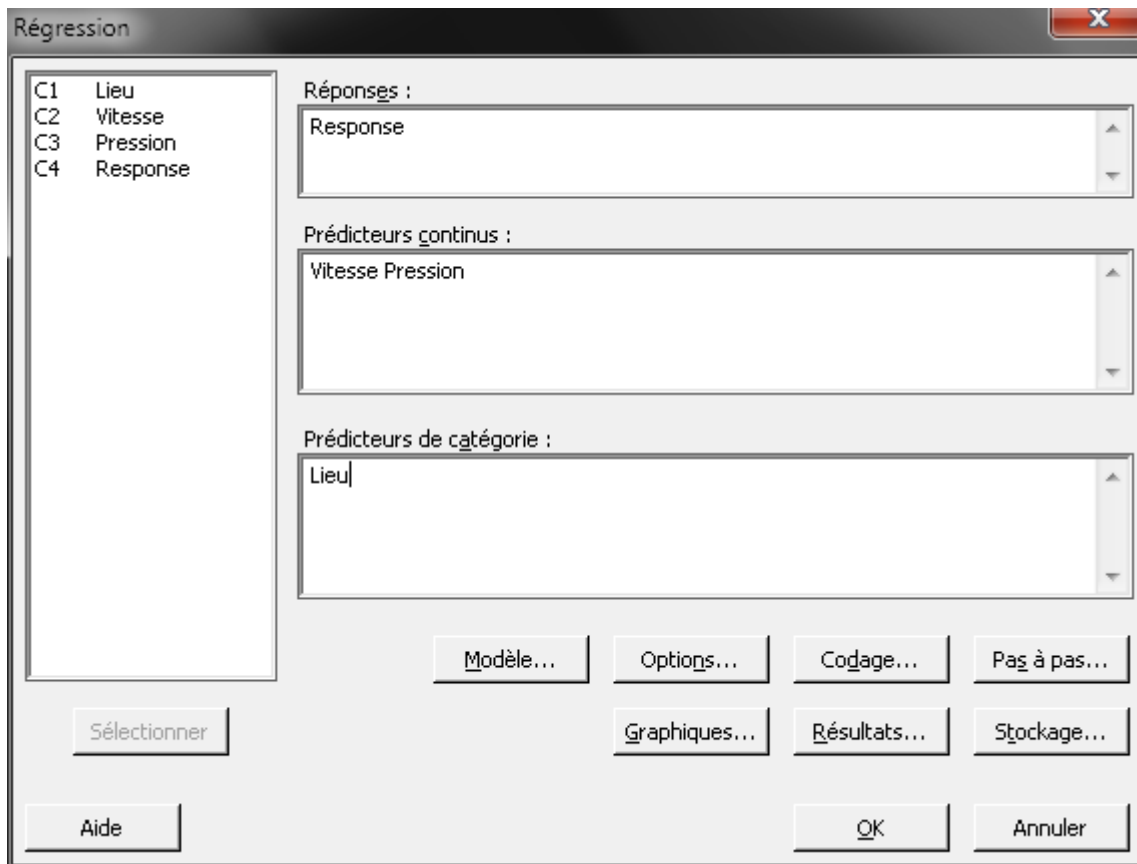
Observation	Response	Valeur ajustée	Résiduelle	Val. résid. norm.	
19	32160	33414	-1254	-2.16	R

R : Valeur résiduelle élevée

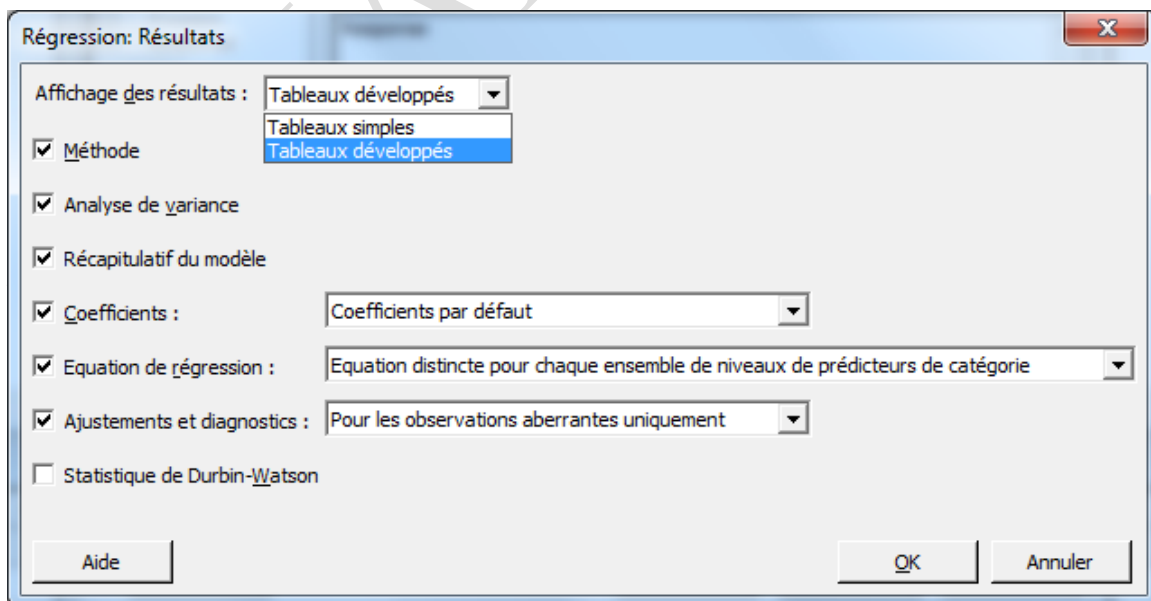
Remarquez l'apparition des équations qui n'apparaissent pas dans Minitab 15!

### 13.3.1. Intervalles de confiance des coefficients

C'est une nouveauté de Minitab 16, nous pouvons enfin les avoir facilement:



En allant dans le bouton Résultats.... et en prenant **Tableaux développés**:



Nous avons alors:

Source	DL	SomCar séq	Contribution	SomCar ajust
Régression	6	3007251	18.95%	3007251
Vitesse	1	223099	1.41%	470347
Pression	1	38400	0.24%	260278
Lieu	1	961707	6.06%	949948
Vitesse*Pression	1	326667	2.06%	326667
Vitesse*Lieu	1	1367340	8.62%	1367340
Pression*Lieu	1	90037	0.57%	90037
Erreur	29	12861137	81.05%	12861137
Inadéquation de l'ajustement	5	9459286	59.61%	9459286
Erreur pure	24	3401851	21.44%	3401851
Total	35	15868387	100.00%	

Source	CM ajust	Valeur F	Valeur de p
Régression	501208	1.13	0.370
Vitesse	470347	1.06	0.312
Pression	260278	0.59	0.450
Lieu	949948	2.14	0.154
Vitesse*Pression	326667	0.74	0.398
Vitesse*Lieu	1367340	3.08	0.090
Pression*Lieu	90037	0.20	0.656
Erreur	443487		
Inadéquation de l'ajustement	1891857	13.35	0.000
Erreur pure	141744		
Total			

Récapitulatif du modèle

S	R carré	R carré (ajust)	SomCar-ErrPrév	R carré (prév)
665.949	18.95%	2.18%	19254178	0.00%

Coefficients

Terme	Coeff	Coef ErT	IC à 95 %	Valeur de T	Valeur de p	FIV
Constante	38123	5154	(27583; 48664)	7.40	0.000	
Vitesse	-117	113	( -348; 115)	-1.03	0.312	26.00
Pression	-1898	2477	(-6963; 3168)	-0.77	0.450	83.00
Lieu						
Ville	-3345	2285	(-8019; 1329)	-1.46	0.154	106.00
Vitesse*Pression	46.7	54.4	(-64.5; 157.9)	0.86	0.398	106.00
Vitesse*Lieu						
Ville	78.0	44.4	(-12.8; 168.8)	1.76	0.090	83.00
Pression*Lieu						
Ville	-245	544	(-1357; 867)	-0.45	0.656	26.00

Equation de régression

Lieu  
 Route Response = 38123 - 117 Vitesse - 1898 Pression + 46.7 Vitesse\*Pression

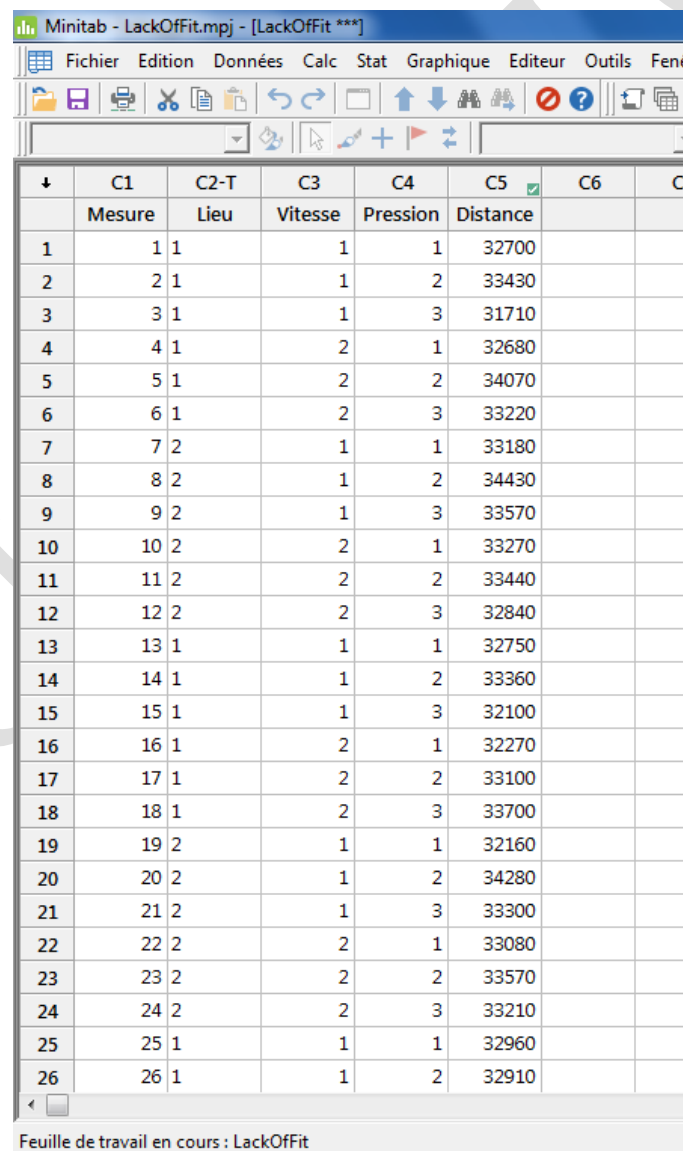
Ville Response = 34778 - 39 Vitesse - 2143 Pression + 46.7 Vitesse\*Pression

### 13.3.2. Inadéquation de l'ajustement (lack-of-fit) et erreur pure

Dans le cours théorique nous avons démontré une astuce mathématique qui permet dans un certain sens de mettre en évidence le manque d'ajustement d'un modèle linéaire (avec ou sans interactions) qui consistait à décomposer l'erreur des résidus en erreur pure et erreur d'ajustement. Même si cette décomposition s'applique aussi à la régression univariée, il nous a semblé plus judicieux de présenter ce sujet ici dans la section concernant la régression multivariée (rappelons que la seule condition est d'avoir de multiples mesures pour une abscisse donnée!).

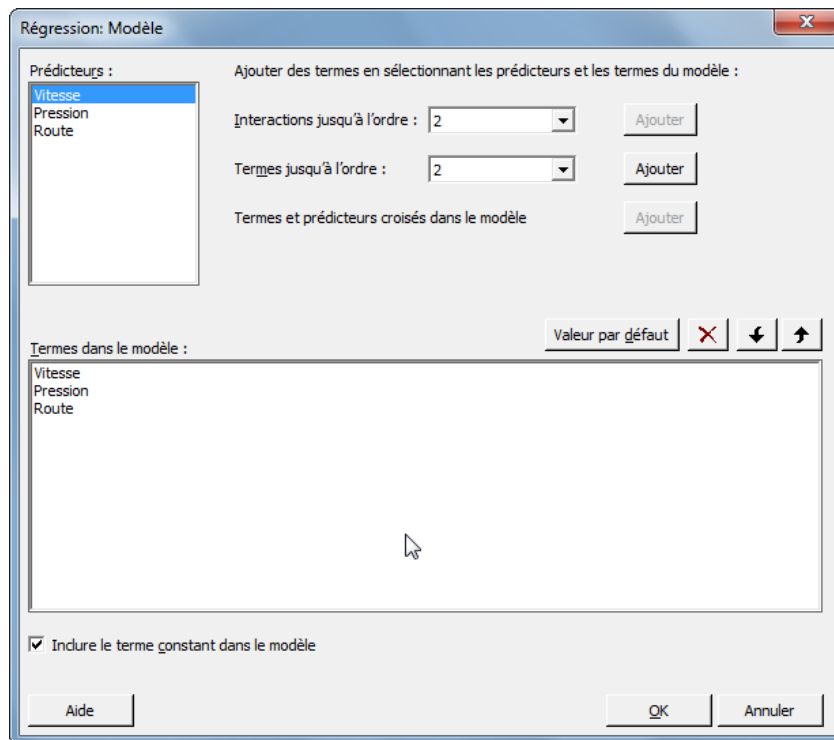
Le but ici va être de voir si nous sommes dans un premier temps capables de reproduire les calculs faits dans Microsoft Excel (à la main) et si nous retombons sur les mêmes résultats que R.

Pour cela, nous partons du jeu de données *LackOfFit.mpj* qui pour rappel est un plan d'expérience d'une série de 12 mesures répétées 3 fois (qui nous est bien connu dans le cours théorique!):



	C1	C2-T	C3	C4	C5	C6	C7
	Mesure	Lieu	Vitesse	Pression	Distance		
1	1	1	1	1	32700		
2	2	1	1	2	33430		
3	3	1	1	3	31710		
4	4	1	2	1	32680		
5	5	1	2	2	34070		
6	6	1	2	3	33220		
7	7	2	1	1	33180		
8	8	2	1	2	34430		
9	9	2	1	3	33570		
10	10	2	2	1	33270		
11	11	2	2	2	33440		
12	12	2	2	3	32840		
13	13	1	1	1	32750		
14	14	1	1	2	33360		
15	15	1	1	3	32100		
16	16	1	2	1	32270		
17	17	1	2	2	33100		
18	18	1	2	3	33700		
19	19	2	1	1	32160		
20	20	2	1	2	34280		
21	21	2	1	3	33300		
22	22	2	2	1	33080		
23	23	2	2	2	33570		
24	24	2	2	3	33210		
25	25	1	1	1	32960		
26	26	1	1	2	32910		

Et nous changeons le modèle comme ci-dessous:



Ce qui nous donne:

### Analyse de variance

Source	DL	SomCar ajust	CM ajust	Valeur F	Valeur de p
Régression	3	1223206	407735	0.89	0.456
Vitesse	1	223099	223099	0.49	0.490
Pression	1	38400	38400	0.08	0.774
Lieu	1	961707	961707	2.10	0.157
Erreur	32	14645181	457662		
Inadéquation de l'ajustement	8	11243331	1405416	9.92	0.000
Erreur pure	24	3401851	141744		
Total	35	15868387			

### Récapitulatif du modèle

S	R carré	R carré (ajust)	R carré (prév)
676.507	7.71%	0.00%	0.00%

### Coefficients

Terme	Coef	Coef ErT	Valeur de T	Valeur de p	FIV
Constante	32640	465	70.21	0.000	
Vitesse	157	226	0.70	0.490	1.00
Pression	40	138	0.29	0.774	1.00
Lieu					
2	327	226	1.45	0.157	1.00

Nous retombons donc bien sur les valeurs attendues et qui de plus sont conformes à celles obtenues à R.

## 13.4. Exercice 99.: Facteur d'Inflation de la Variance (VIF)/colinéarité

Minitab® Statistical Software 15.1.1.0

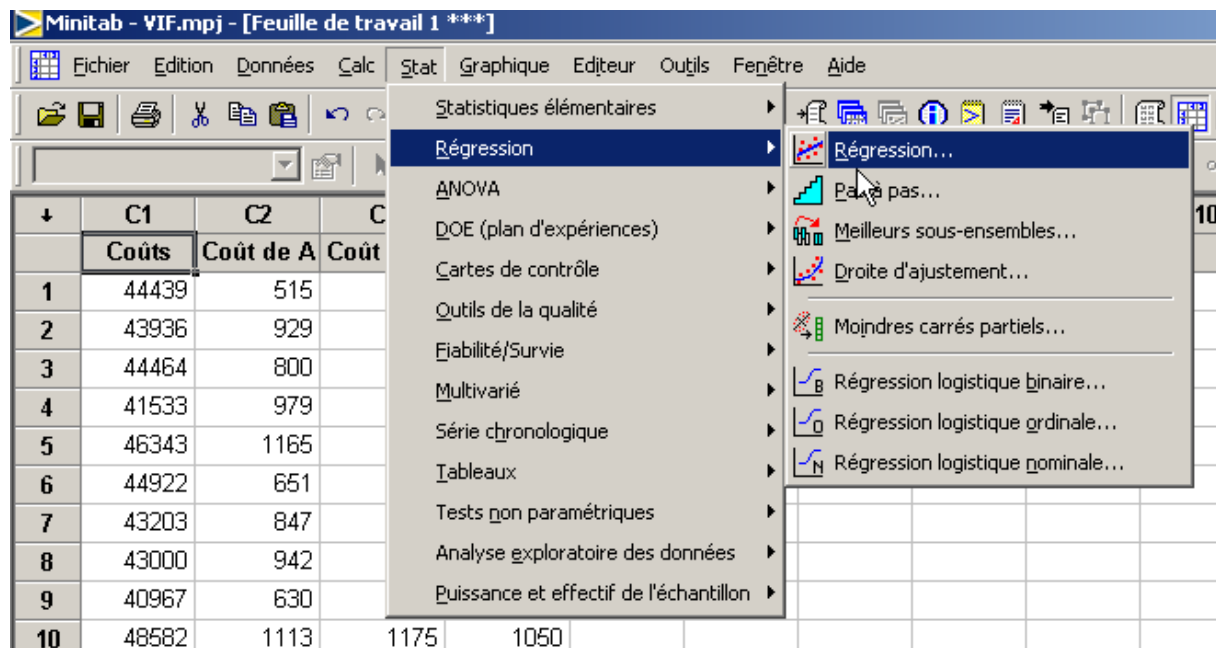
Comme toujours, le but ici va être de vérifier que la relation démontrée dans le cours théorique et les calculs effectués à la main dans Microsoft Excel coïncident avec Minitab pour analyser la multicolinéarité.

Nous partons alors des données suivantes:

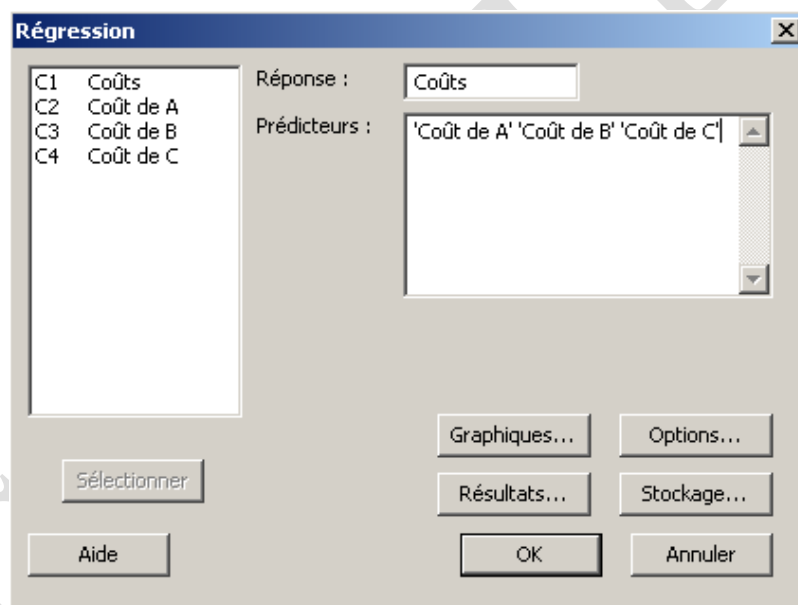
↓	C1	C2	C3	C4	C5	C6
	Coûts	Coût de A	Coût de B	Coût de C		
1	44439	515	541	928		
2	43936	929	692	711		
3	44464	800	710	824		
4	41533	979	675	758		
5	46343	1165	1147	635		
6	44922	651	939	901		
7	43203	847	755	580		
8	43000	942	908	589		
9	40967	630	738	682		
10	48582	1113	1175	1050		
11	45003	1086	1075	984		
12	44303	843	640	828		
13	42070	500	752	708		
14	44353	813	989	804		
15	45968	1190	823	904		
16	47781	1200	1108	1120		
17	43202	731	590	1065		
18	44074	1089	607	1132		
19	44610	786	513	839		
20						
21						

Feuille de travail en cours : Feuille de travail 1

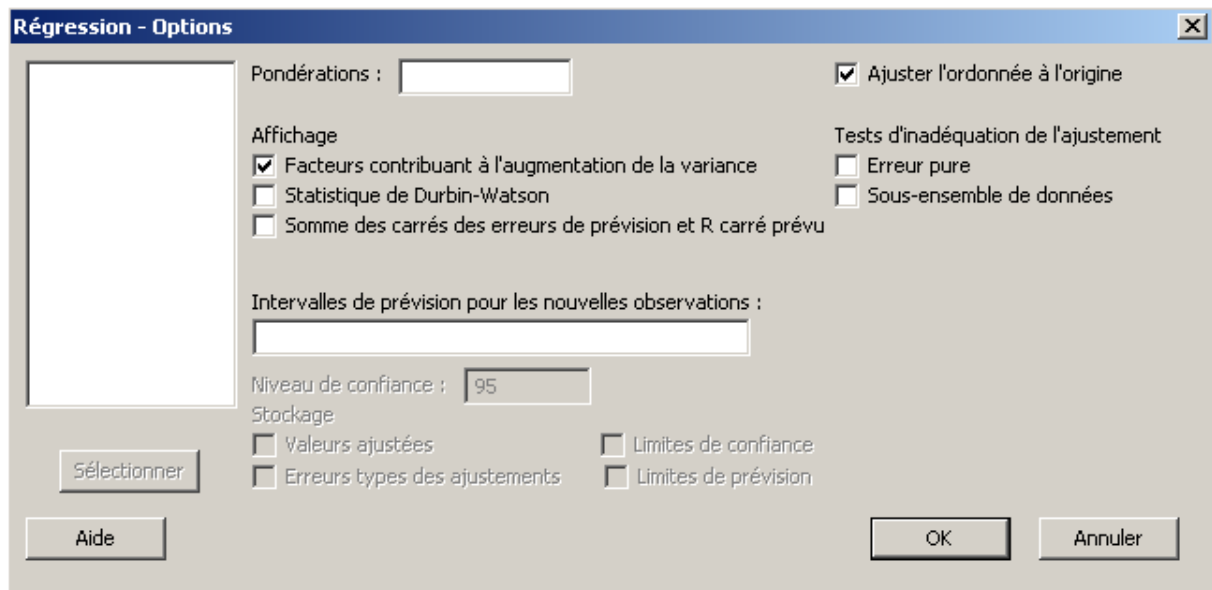
et nous allons dans le menu **Stat/Régression/Régression...**:



et nous prenons:



Cliquez sur Options... et activez **Facteurs contribuant à l'augmentation de la variance**:



Validez deux fois par **OK** pour obtenir:

**Analyse de régression : Coûts en fonction de Coût de A; Coût de B; ...**

L'équation de régression est  
 Coûts = 35103 + 2.07 Coût de A + 4.18 Coût de B + 4.79 Coût de C

Prédicteur	Coeff	Coef ErT	T	P	Facteurs contribuant
					à l'augmentation de
					la variance
Constante	35103	1837	19.11	0.000	
Coût de A	2.066	1.665	1.24	0.234	1.542
Coût de B	4.176	1.681	2.48	0.025	1.436
Coût de C	4.791	1.789	2.68	0.017	1.090

S = 1252.76    R carré = 64.5 %    R carré (ajust) = 57.5 %

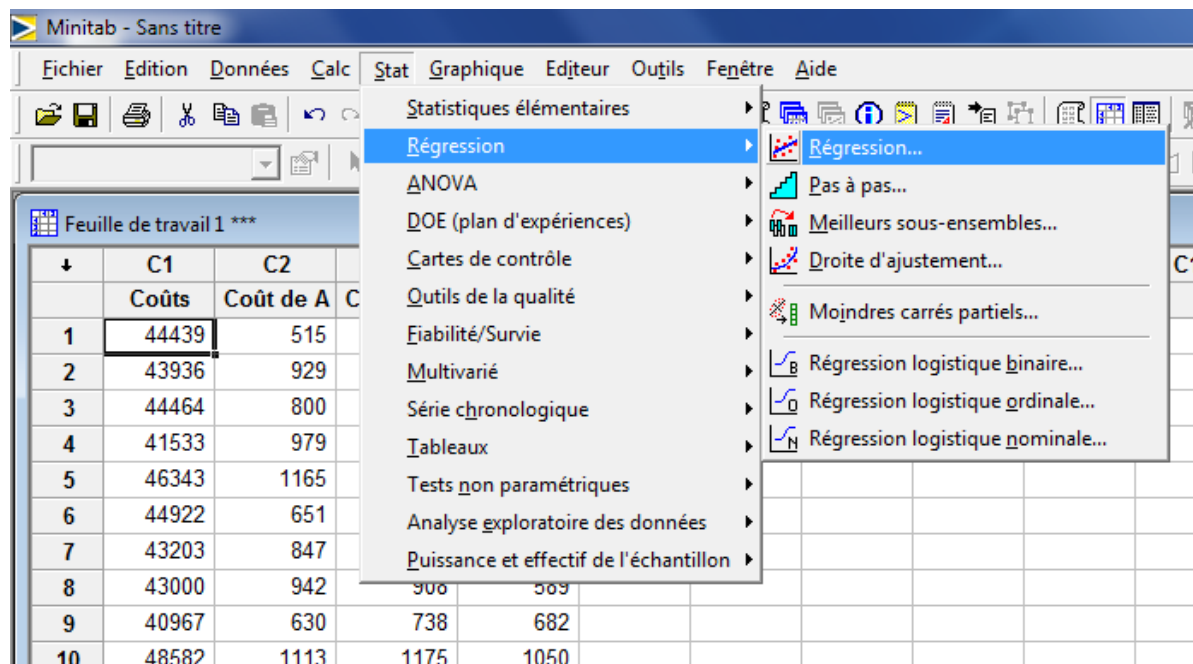
**Analyse de variance**

Source	DL	Somme des carrés	CM	F	P
Régression	3	42856230	14285410	9.10	0.001
Erreur résiduelle	15	23541261	1569417		
Total	18	66397491			

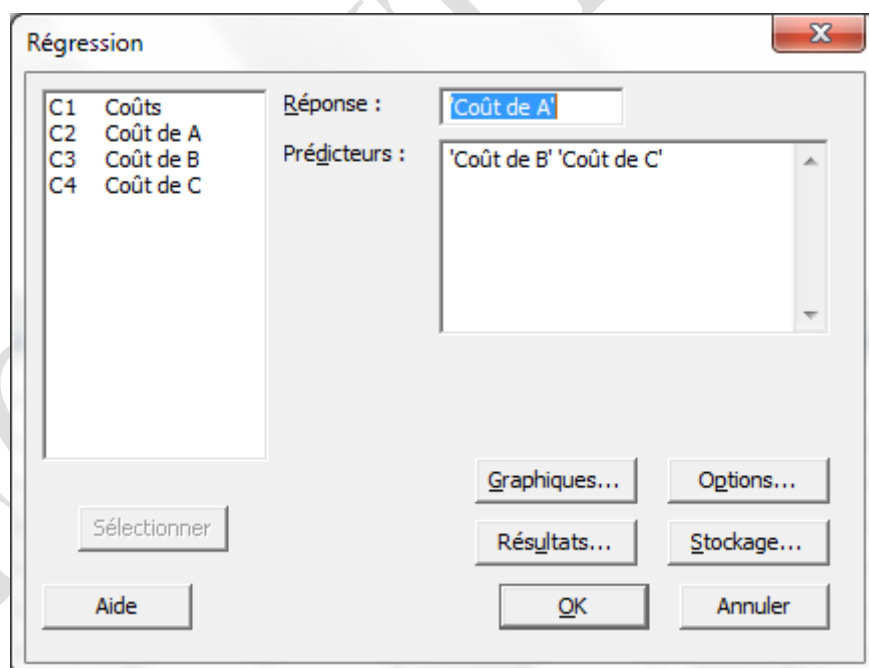
Nous retrouvons effectivement les valeurs calculées manuellement dans le cours théorique et donc les mêmes conclusions.



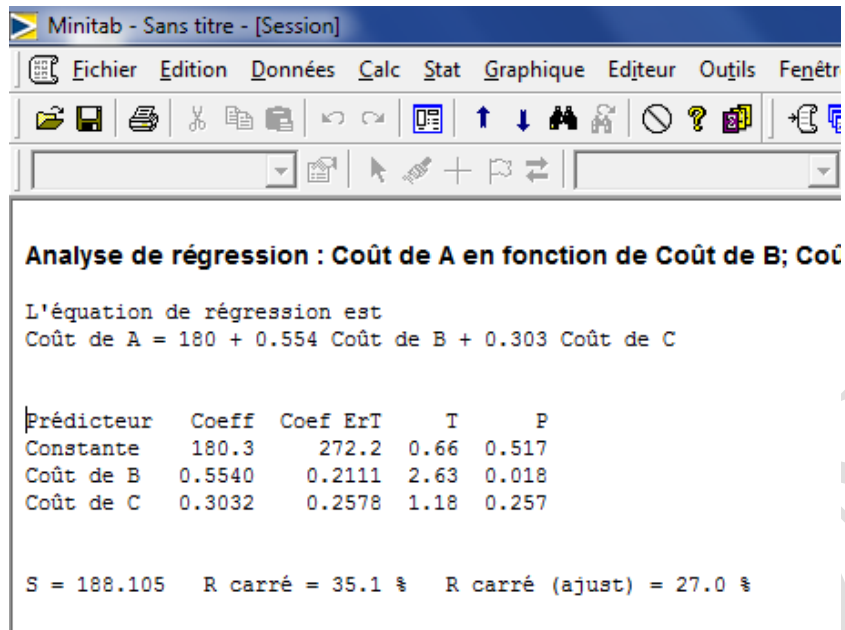
On peut aussi vérifier qu'on retrouve la valeur de 1.542 par la définition même du coefficient de corrélation multiple et de l'expression du VIF. Pour cela, nous allons dans le menu **Stat/Régression/Régression...**:



Et ensuite nous mettons:



Ce qui donne:



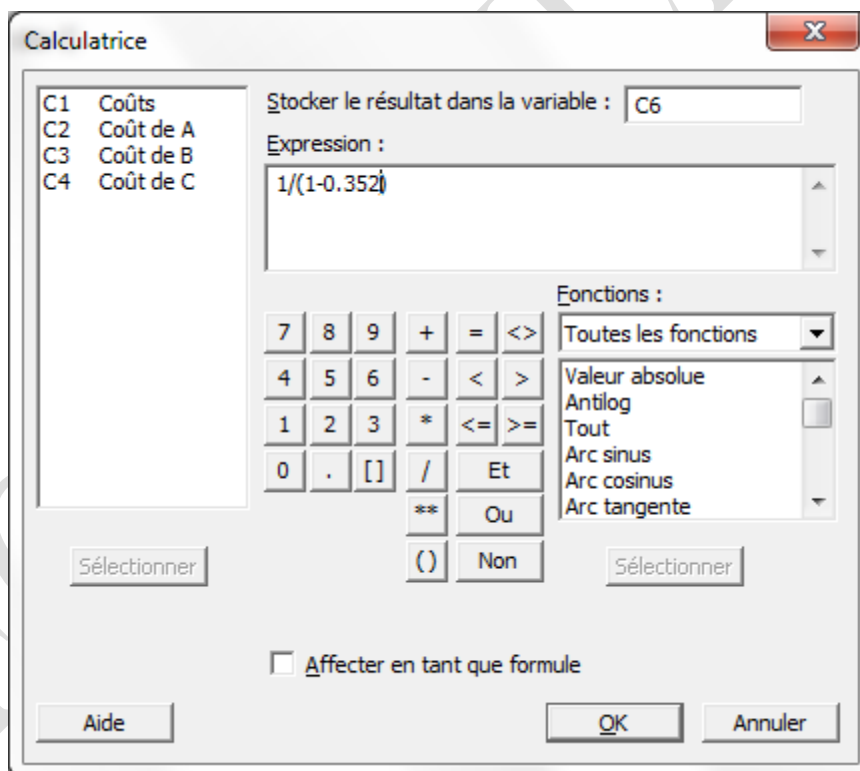
**Analyse de régression : Coût de A en fonction de Coût de B; Coût de C**

L'équation de régression est  
 Coût de A = 180 + 0.554 Coût de B + 0.303 Coût de C

Prédicteur	Coeff	Coef ErT	T	P
Constante	180.3	272.2	0.66	0.517
Coût de B	0.5540	0.2111	2.63	0.018
Coût de C	0.3032	0.2578	1.18	0.257

S = 188.105    R carré = 35.1 %    R carré (ajust) = 27.0 %

Nous voyons donc le R carré multiple qui vaut 35.1%. En injectant ce résultat dans la forme générique du VIF en utilisant la calculatrice Minitab:



Calculatrice

Stocker le résultat dans la variable : C6

Expression :  
 $1/(1-0.352)$

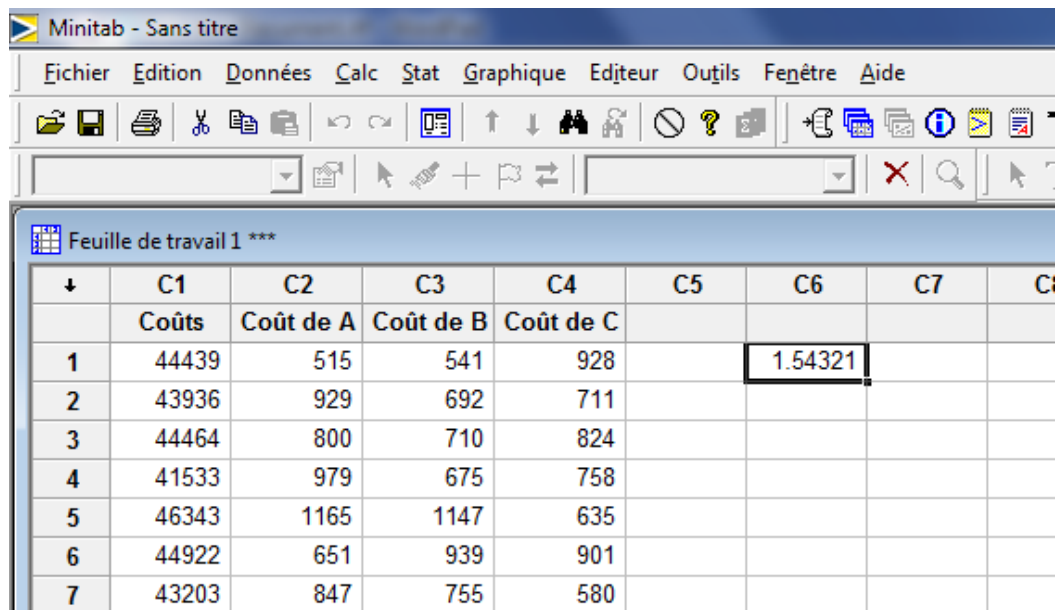
Fonctions :  
 Toutes les fonctions  
 Valeur absolue  
 Antilog  
 Tout  
 Arc sinus  
 Arc cosinus  
 Arc tangente

Sélectionner

Affecter en tant que formule

Aide    OK    Annuler

Cela nous donne:



The screenshot shows the Minitab software interface. The window title is "Minitab - Sans titre". The menu bar includes "Fichier", "Edition", "Données", "Calc", "Stat", "Graphique", "Editeur", "Outils", "Fenêtre", and "Aide". The toolbar contains various icons for file operations, editing, and data analysis. The main window displays a spreadsheet titled "Feuille de travail 1 \*\*\*". The spreadsheet has columns labeled C1 through C7 and rows numbered 1 through 7. The data is as follows:

	C1	C2	C3	C4	C5	C6	C7	C8
	Coûts	Coût de A	Coût de B	Coût de C				
1	44439	515	541	928		1.54321		
2	43936	929	692	711				
3	44464	800	710	824				
4	41533	979	675	758				
5	46343	1165	1147	635				
6	44922	651	939	901				
7	43203	847	755	580				

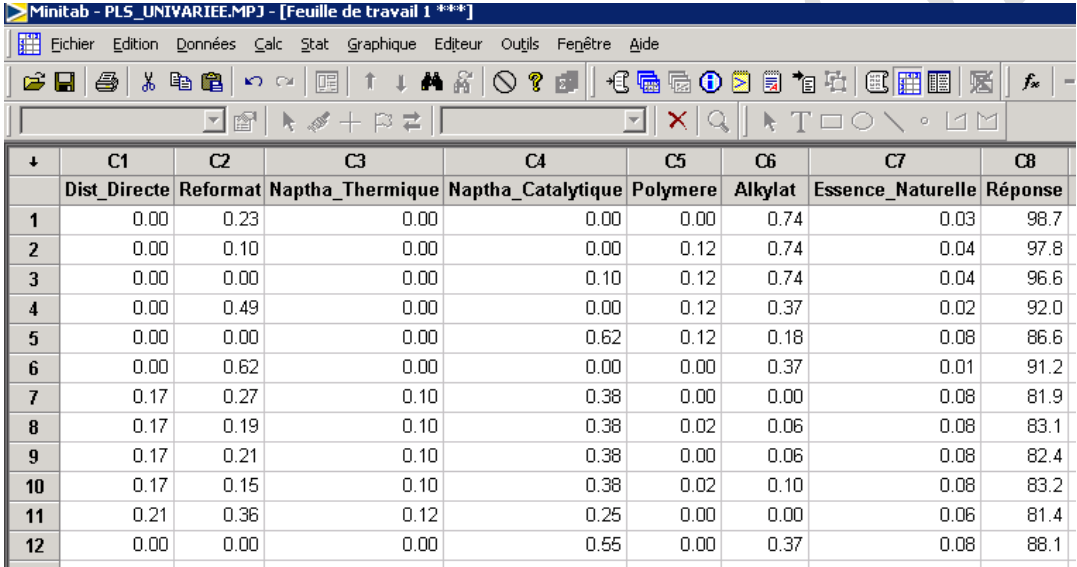
Nous retrouvons bien la valeur de 1.54321 (c'est-à-dire la première valeur du VIF que nous avons obtenu juste précédemment dans la capture d'écran!).

## 13.5. Exercice 100.: Régression (linéaire) par moindres carrés partiels (régression linéaire PLS univariée: PLS1)

Minitab® Statistical Software 15.1.1.0

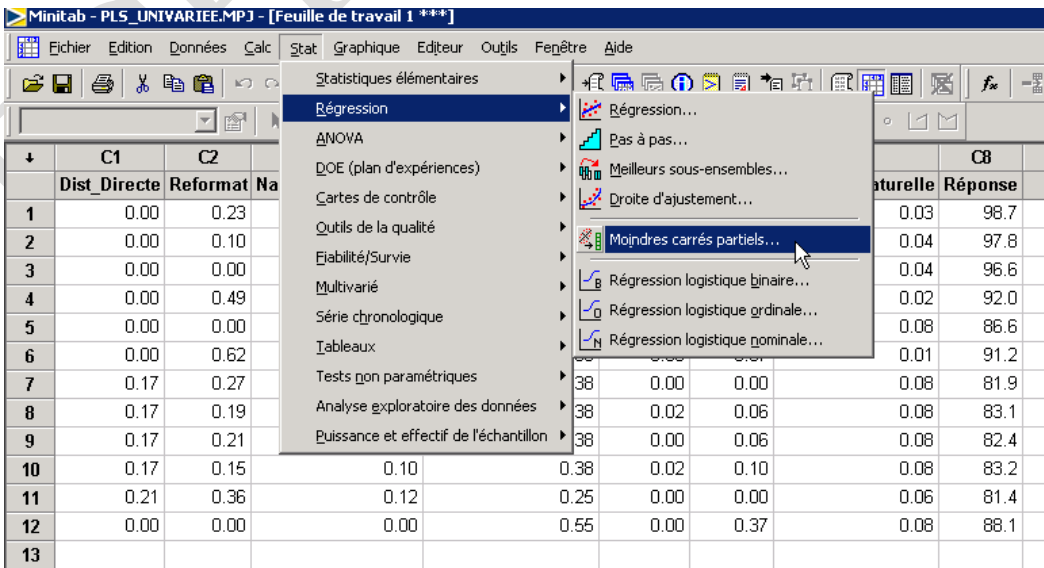
Le but va être ici de vérifier si nous obtenons ou pas les résultats des calculs vu dans le cours théorique lors de la lecture de l'ouvrage de M. Tenenhaus<sup>2</sup> sur la régression PLS univariée (PLS1), c'est-à-dire la régression sur des variables explicatives corrélées avec une unique variable à expliquer.

Nous utiliserons donc les données suivantes:



	C1	C2	C3	C4	C5	C6	C7	C8
	Dist_Directe	Reformat	Naptha_Thermique	Naptha_Catalytique	Polymere	Alkylat	Essence_Naturelle	Réponse
1	0.00	0.23	0.00	0.00	0.00	0.74	0.03	98.7
2	0.00	0.10	0.00	0.00	0.12	0.74	0.04	97.8
3	0.00	0.00	0.00	0.10	0.12	0.74	0.04	96.6
4	0.00	0.49	0.00	0.00	0.12	0.37	0.02	92.0
5	0.00	0.00	0.00	0.62	0.12	0.18	0.08	86.6
6	0.00	0.62	0.00	0.00	0.00	0.37	0.01	91.2
7	0.17	0.27	0.10	0.38	0.00	0.00	0.08	81.9
8	0.17	0.19	0.10	0.38	0.02	0.06	0.08	83.1
9	0.17	0.21	0.10	0.38	0.00	0.06	0.08	82.4
10	0.17	0.15	0.10	0.38	0.02	0.10	0.08	83.2
11	0.21	0.36	0.12	0.25	0.00	0.00	0.06	81.4
12	0.00	0.00	0.00	0.55	0.00	0.37	0.08	88.1

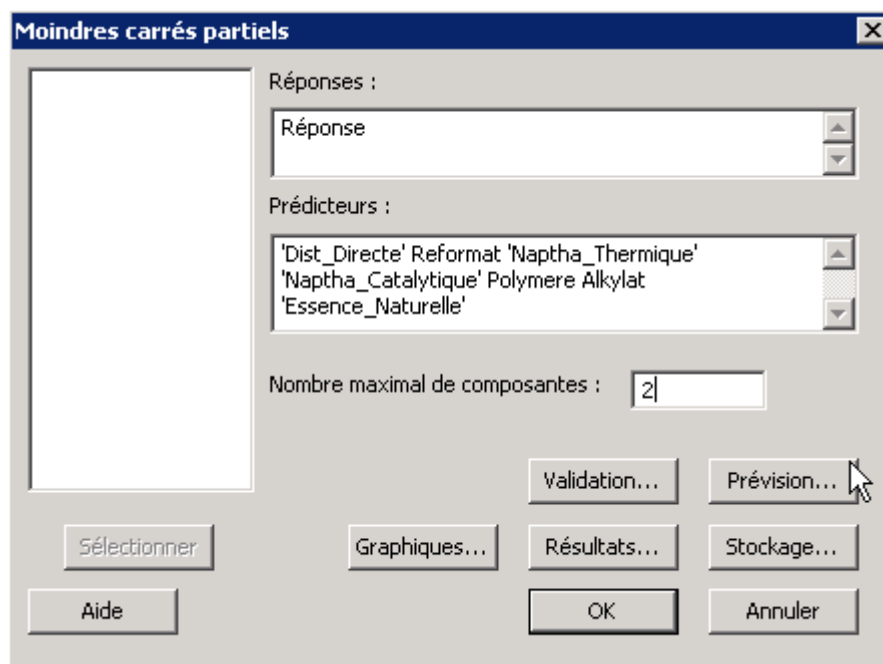
Ensuite, nous allons dans le menu **Stat/Régression/Moindres carrés partiels...**:



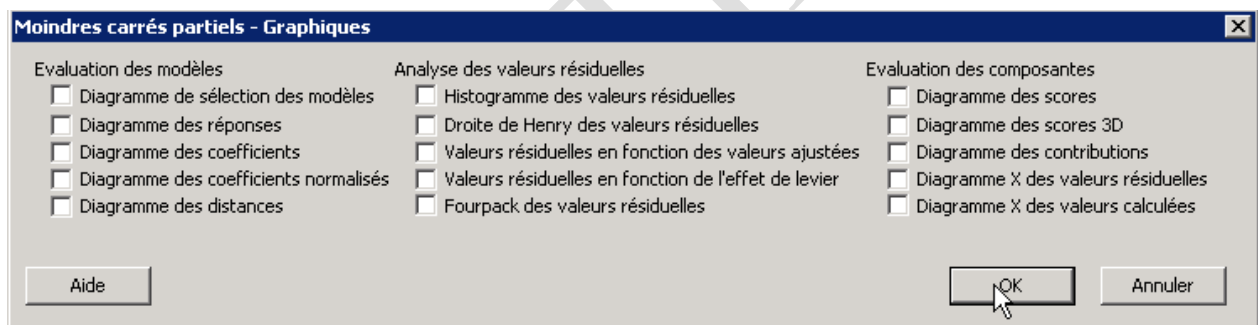
	C1	C2	C3	C4	C5	C6	C7	C8
	Dist_Directe	Reformat	Na				essence Naturelle	Réponse
1	0.00	0.23					0.03	98.7
2	0.00	0.10					0.04	97.8
3	0.00	0.00					0.04	96.6
4	0.00	0.49					0.02	92.0
5	0.00	0.00					0.08	86.6
6	0.00	0.62					0.01	91.2
7	0.17	0.27		0.38	0.00	0.00	0.08	81.9
8	0.17	0.19		0.38	0.02	0.06	0.08	83.1
9	0.17	0.21		0.38	0.00	0.06	0.08	82.4
10	0.17	0.15	0.10	0.38	0.02	0.10	0.08	83.2
11	0.21	0.36	0.12	0.25	0.00	0.00	0.06	81.4
12	0.00	0.00	0.00	0.55	0.00	0.37	0.08	88.1
13								

<sup>2</sup>Michel Tenenhaus, *Régression PLS, Édition Technip, ISBN 2-7108-0735-1, Pages 75-83*

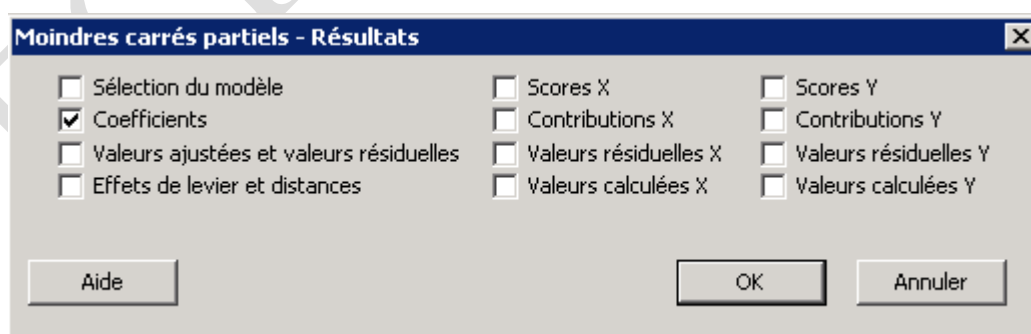
Nous mettons dans la boîte de dialogue qui suit:



Dans le bouton **Graphiques...**, nous décochons tout (bien que le **Diagramme** des réponses soit très très intéressant!):



Dans le bouton **Coefficients**, nous prenons:



Si nous validons le tout, nous obtenons d'abord une ANOVA de la régression (analyse classique):

**Minitab - PLS\_UNIVARIEE.MPJ - [Session]**

Fichier Edition Données Calc Stat Graphique Éditeur Outils Fenêtre Aide

**Régression PLS (moindres carrés partiels) : Réponse en fonction de Dist\_Directe**

Nombre de composantes spécifié : 2

Analyse de la variance pour Réponse

Source	DL	Somme des carrés	CM	F	P
Régression	2	456.730	228.365	185.72	0.000
Erreur résiduelle	9	11.067	1.230		
Total	11	467.797			

Sélection et validation de modèle pour Réponse

Composantes	Variance X	Erreur - Somme des carrés	R carré
1	0.573606	35.7425	0.923594
2	0.726128	11.0666	0.976343

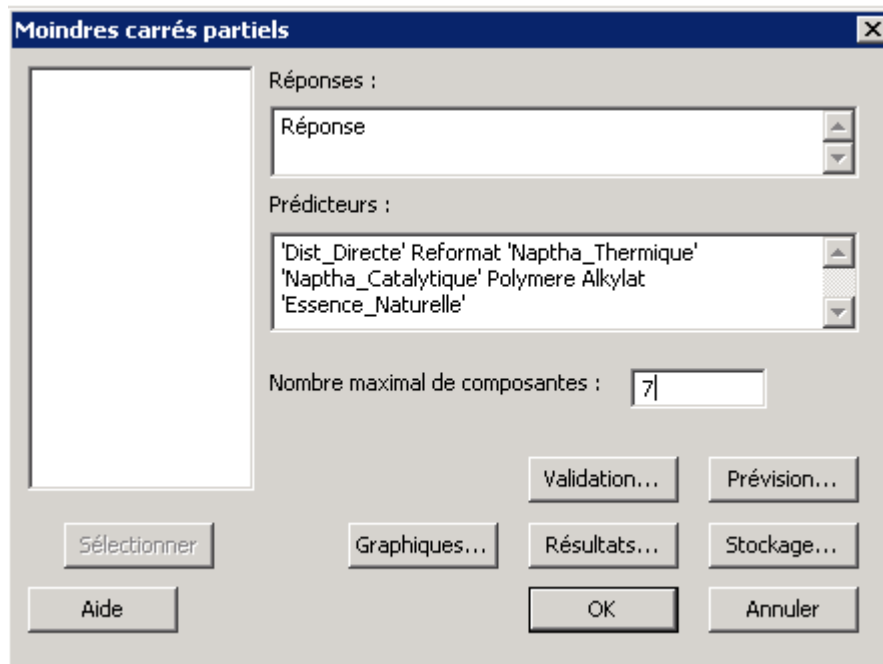
et ensuite la partie qui nous intéresse:

Coefficients de régression

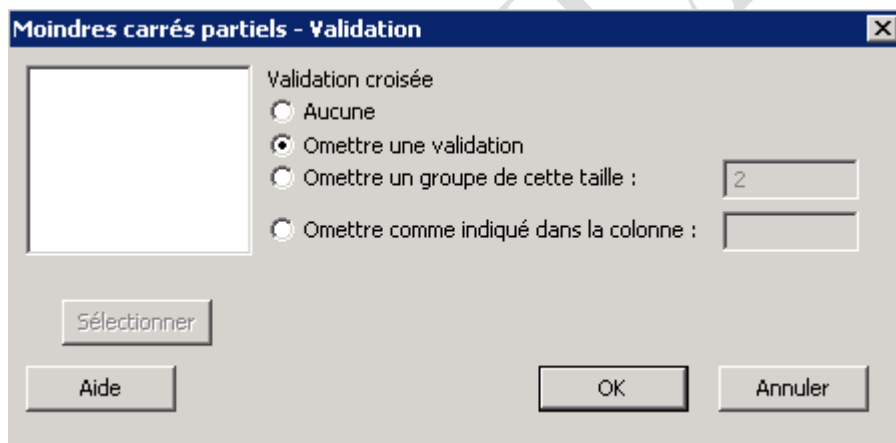
	Réponse	Réponse normalisé
Constante	92.3422	0.000000
Dist_Directe	-12.5636	-0.177801
Reformat	-6.8312	-0.204825
Naptha_Thermique	-21.4140	-0.176741
Naptha_Catalytique	-6.3952	-0.222621
Polymere	3.6776	0.032202
Alkylat	8.9787	0.403275
Essence_Naturelle	-30.6670	-0.127347

qui est à peu près conforme à ce que nous avons lu et étudié dans le cours théorique.

Si maintenant nous souhaitons comme dans le cours de Data Mining demander à Minitab de déterminer automatiquement le meilleur nombre de variables intermédiaires, nous mettons dans la boîte de dialogue le nombre de composantes au maximum possible logique:



et dans le bouton Validation... nous prenons:



ce qui donnera en validant le tout:

Minitab - PLS\_UNIVARIEE.MPJ - [Session]

Fichier Edition Données Calc Stat Graphique Éditeur Outils Fenêtre Aide

Régression PLS (moindres carrés partiels) : Réponse en fonction de Dist\_Directe

Nombre de composantes sélectionnées par validation croisée : 3  
 Nombre d'observations omises par groupe : 1  
 Nombre de composantes en validation croisée : 5  
 Le nombre de composantes spécifié a dû être ajusté.

Analyse de la variance pour Réponse

Source	DL	Somme des carrés	CM	F	P
Régression	3	463.379	154.460	279.69	0.000
Erreur résiduelle	8	4.418	0.552		
Total	11	467.797			

Sélection et validation de modèle pour Réponse

Composantes	Variance X	Erreur - Somme des carrés	R carré	SomCar-ErrPrév	R carré (prév)
1	0.573606	35.7425	0.923594	55.7077	0.880915
2	0.726128	11.0666	0.976343	41.4327	0.911430
3	0.918258	4.4181	0.990556	20.2740	0.956661
4		4.3092	0.990788	21.2424	0.954591
5		3.5219	0.992471	24.5180	0.947588

Nous voyons ci-dessus dans la dernière table que comme pour Tanagra, Minitab prend 3 variables intermédiaires comme étant le résultat optimal. Au final cela donne les coefficients:

#### Coefficients de régression

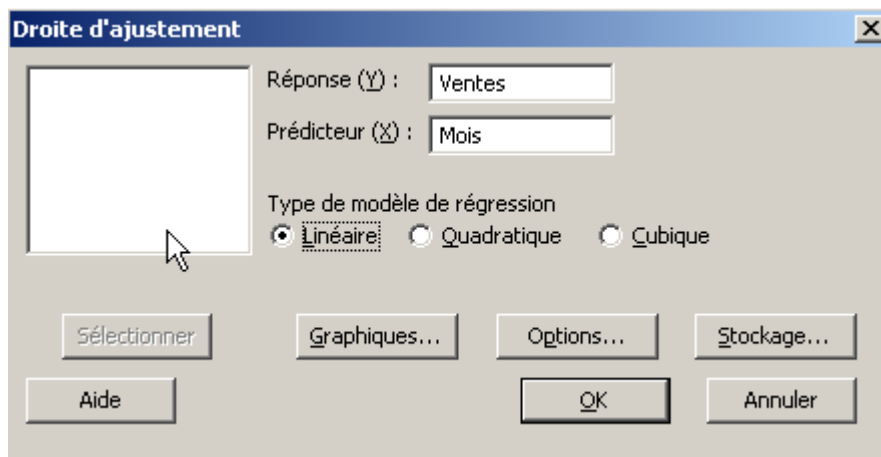
	Réponse	Réponse normalisé
Constante	92.6760	0.000000
Dist_Directe	-9.8283	-0.139092
Reformat	-6.9602	-0.208694
Naptha_Thermique	-16.6662	-0.137555
Naptha_Catalytique	-8.4218	-0.293168
Polymere	-4.3889	-0.038430
Alkylat	10.1613	0.456390
Essence_Naturelle	-34.5290	-0.143384



## 13.6. Exercice 101.: Régression polynomiale cubique ou quadratique

Minitab® Statistical Software 15.1.1.0

Nous avons découvert dans l'exercice précédent l'outil **Droite d'ajustement** qui nous permettait de tracer des courbes (et donc aussi des droites) et qui proposait un modèle cubique et quadratique:



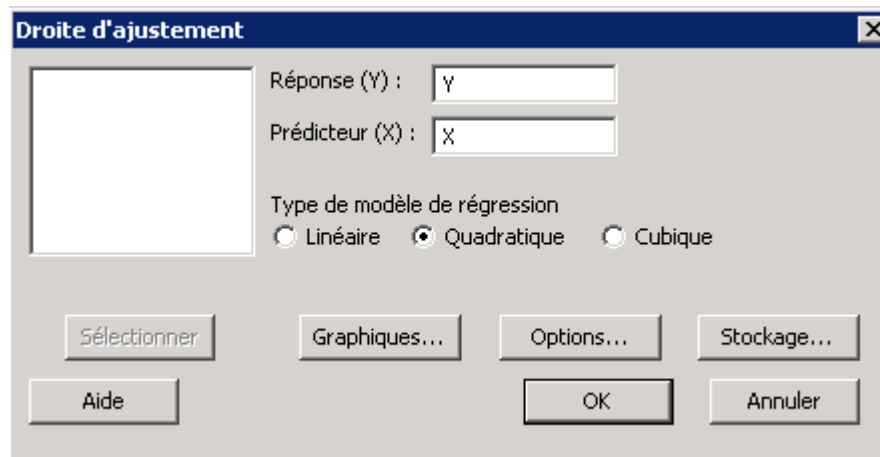
Nous avons vu dans le cours de méthodes numériques qu'il était très souvent possible de ramener nombre de mesures très facilement à un modèle linéaire (raison pour laquelle nous n'avons pas pratiqué dans Microsoft Excel tellement c'est trivial). C'est ce que nous proposons implicitement ici Minitab. Voyons un exemple en considérant les données suivantes que vous saisissez:

	C1	C2	C3	C4	C5	C6	C7
	X	Y					
1	24	236					
2	21	229					
3	34	258					
4	37	257					
5	29	268					
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							

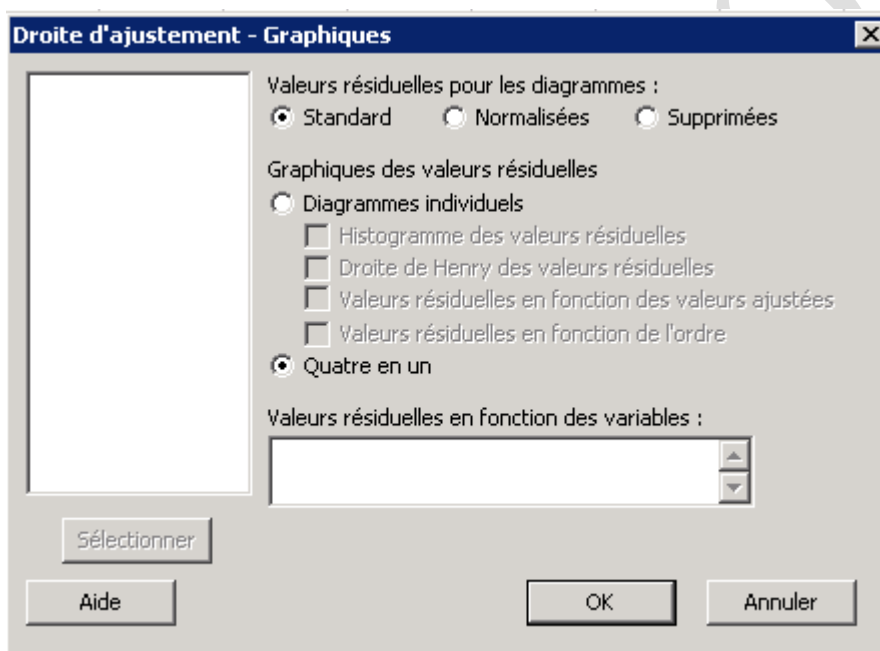
Nous allons ensuite dans le menu **Régression/Droite d'ajustement...**:

	C1	C2	C3
	X	Y	
1	24	236	
2	21	229	
3	34	258	
4	37	257	
5	29	268	
6			
7			
8			
9			
10			
11			

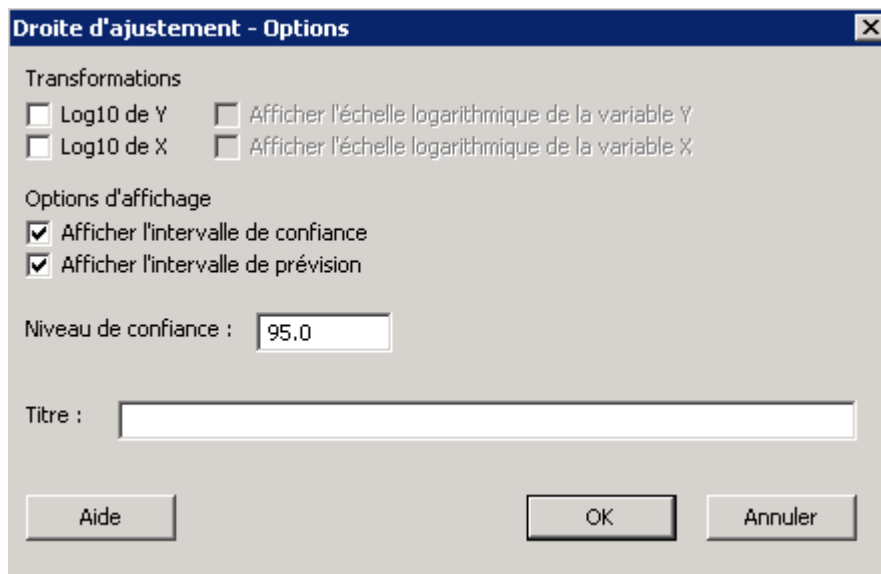
et nous avons alors:



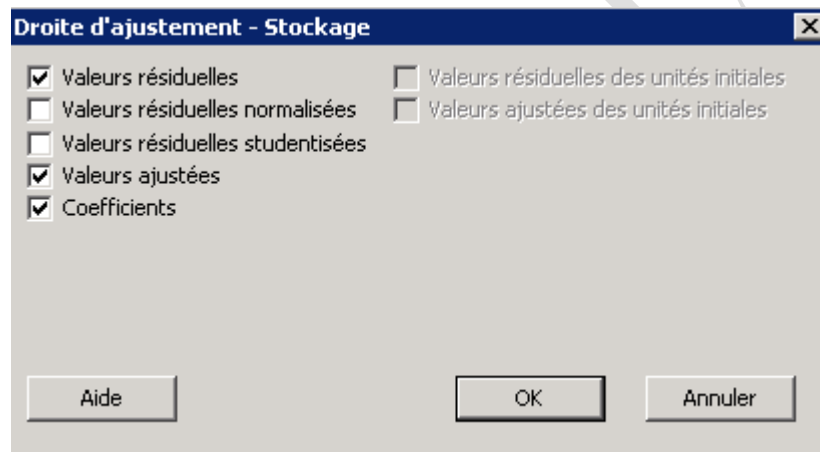
Nous cliquons sur le bouton **Graphiques...** pour mettre les quatre en un:



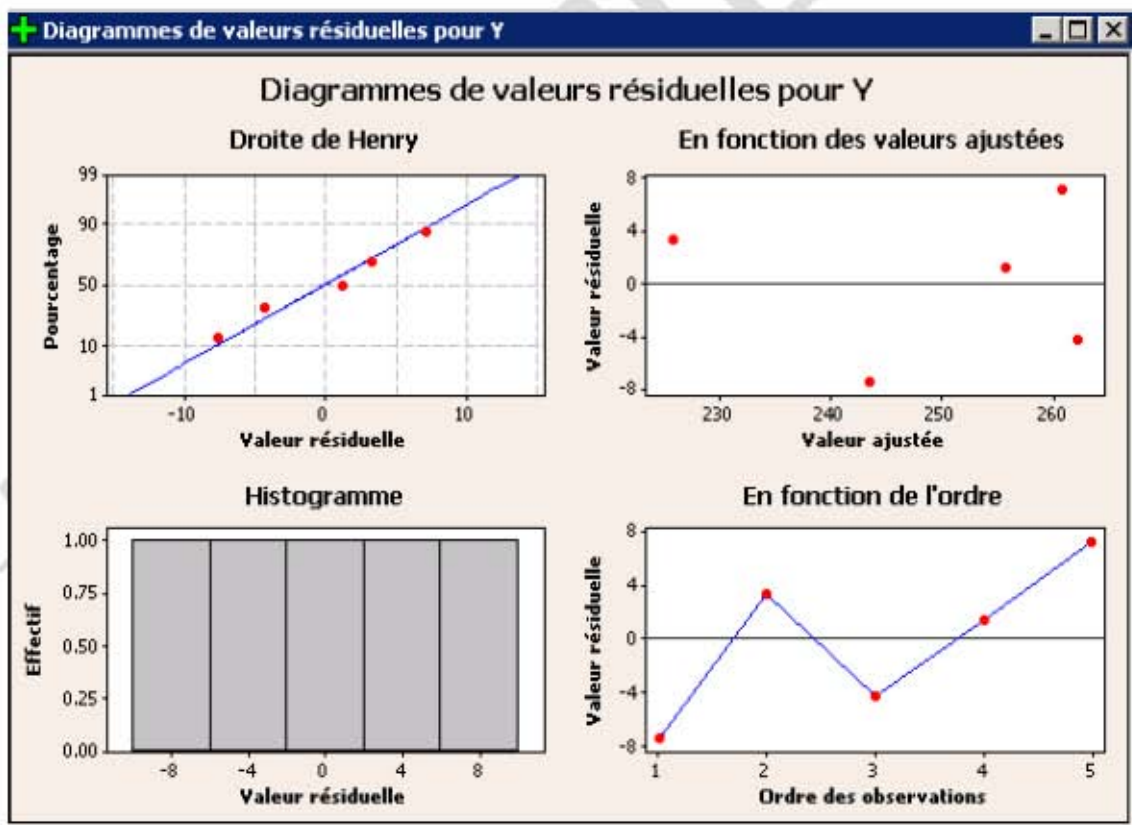
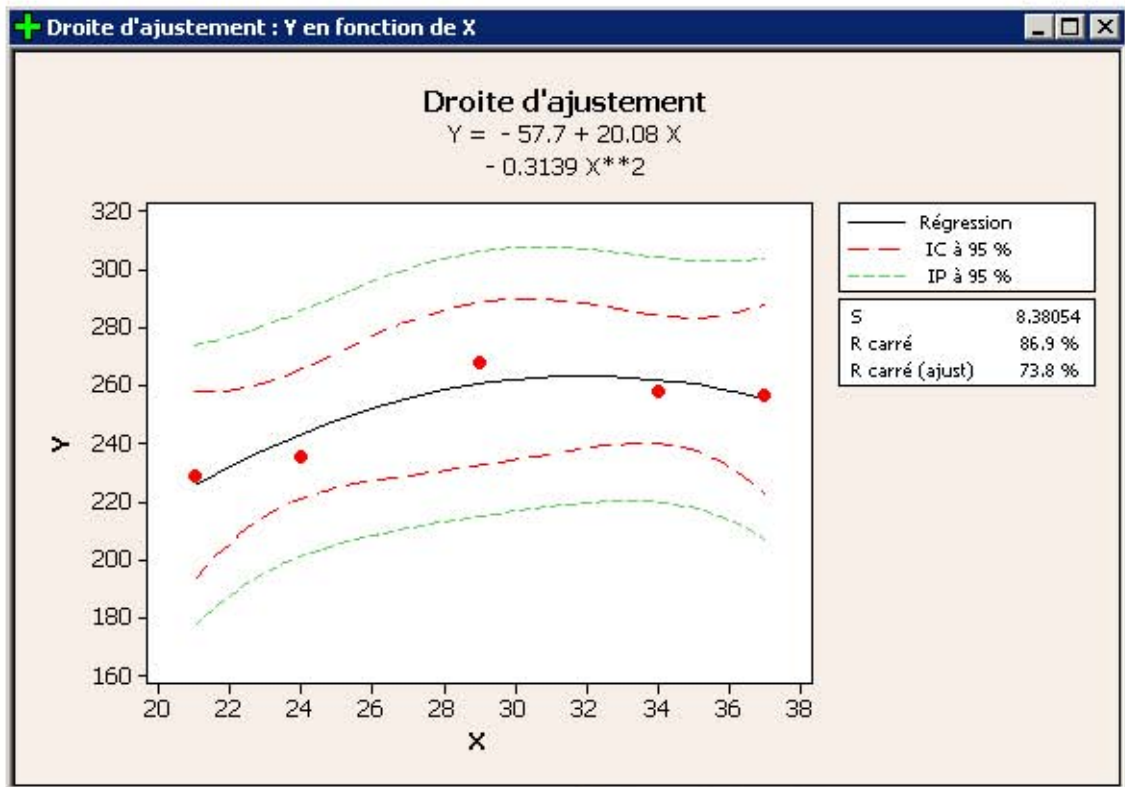
Sur **Options...** pour avoir les différents intervalles:



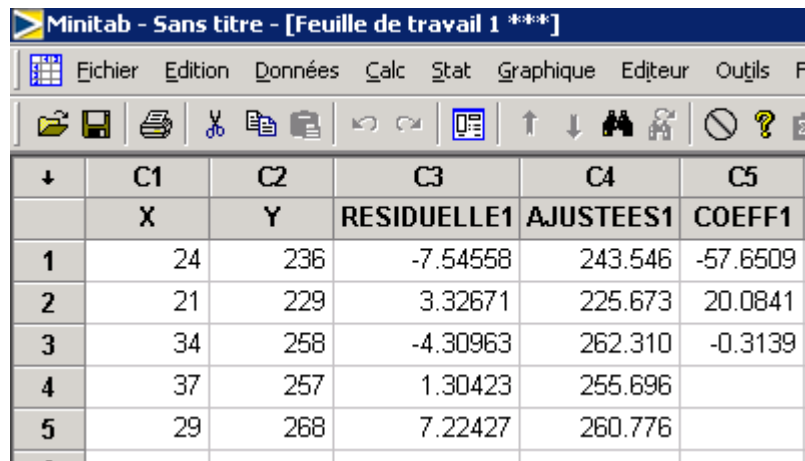
et sur **Stockage...** pour prendre:



et nous validons le tout par **OK**. Nous avons alors d'abord les deux graphiques suivants:



et la feuille suivante:



The screenshot shows the Minitab interface with a table of regression results. The table has five columns: C1 (X), C2 (Y), C3 (RESIDUELLE1), C4 (AJUSTEES1), and C5 (COEFF1). The data is as follows:

	C1	C2	C3	C4	C5
	X	Y	RESIDUELLE1	AJUSTEES1	COEFF1
1	24	236	-7.54558	243.546	-57.6509
2	21	229	3.32671	225.673	20.0841
3	34	258	-4.30963	262.310	-0.3139
4	37	257	1.30423	255.696	
5	29	268	7.22427	260.776	

avec les coefficients qui doivent être lus ainsi:

$$y = -57.6509 + 20.0841x - 0.3139x^2$$

La fenêtre de session est quant à elle vide...

## 13.7. Exercice 102.: Effectuer une régression pas à pas (ascendante et descendante)

Minitab® Statistical Software 15.1.1.0

La régression pas à pas est simplement une technique qui "construit" la fonction de régression en prenant soit:

- En mode ascendant: ajoute les variables explicatives au modèle qui maximisent le coefficient de régression.
- En mode descendant: retire du modèle les variables explicatives qui minimisent le coefficient de régression.

Dans la pratique on privilégiera donc le pas à pas descendant la majorité du temps (méthode "backward").

Ouvrez le fichier *RegressionPasAPas.mpj*:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Mois	Coûts	Coût de A	Coût de B	Coût de C								
1	1	44439	515	541	928								
2	2	43936	929	692	711								
3	3	44464	800	710	824								
4	4	41533	979	675	758								
5	5	46343	1165	1147	635								
6	6	44922	651	939	901								
7	7	43203	847	755	580								
8	8	43000	942	908	589								
9	9	40967	630	738	682								
10	10	48582	1113	1175	1050								
11	11	45003	1086	1075	984								
12	12	44303	843	640	828								
13	13	42070	500	752	708								
14	14	44353	813	989	804								
15	15	45968	1190	823	904								
16	16	47781	1200	1108	1120								
17	17	43202	731	590	1065								
18	18	44074	1089	607	1132								
19	19	44610	786	513	839								

En faisant une régression linéaire simple comme dans l'exemple antérieur:

The screenshot shows the Minitab interface with the 'Stat' menu open and 'Régression' selected. The 'Régression...' option is highlighted. The background data table is as follows:

	C1	C2	C3	C4	C5
	Mois	Coûts	Coût de A	Coût de B	Coût de C
1	1	44439			
2	2	43936			
3	3	44464			
4	4	41533			
5	5	46343	1		
6	6	44922			
7	7	43203			
8	8	43000			
9	9	40967			
10	10	48582	1		
11	11	45003	1086	1075	984
12	12	44303	843	640	828
13	13	42070	500	752	708
14	14	44353	813	989	804
15	15	45968	1190	823	904
16	16	47781	1200	1108	1120
17	17	43202	731	590	1065
18	18	44074	1089	607	1132
19	19	44610	786	513	839
20					

Statistiques élémentaires  
 Régression  
 ANOVA  
 DOE (plan d'expériences)  
 Cartes de contrôle  
 Outils de la qualité  
 Fiabilité/Survie  
 Multivarié  
 Série chronologique  
 Tableaux  
 Tests non paramétriques  
 Analyse exploratoire des données  
 Puissance et effectif de l'échantillon

Régression...  
 Pas à pas...  
 Meilleurs sous-ensembles...  
 Droite d'ajustement...  
 Moindres carrés partiels...  
 Régression logistique binaire...  
 Régression logistique ordinale...  
 Régression logistique nominale...

Effectuer une régression à l'aide de l'estimation des moindres carrés

avec:

The 'Régression' dialog box shows the following configuration:

- Réponse :** Coûts
- Prédicteurs :** 'Coût de A' 'Coût de B' 'Coût de C'

Buttons: Sélectionner, Aide, Graphiques..., Options..., Résultats..., Stockage..., OK, Annuler.



nous obtenons:

### Analyse de régression : Coûts en fonction de Coût de A; Coût de B; ...

L'équation de régression est

$$\text{Coûts} = 35103 + 2.07 \text{ Coût de A} + 4.18 \text{ Coût de B} + 4.79 \text{ Coût de C}$$

Prédicteur	Coeff	Coef ErT	T	P
Constante	35103	1837	19.11	0.000
Coût de A	2.066	1.665	1.24	0.234
Coût de B	4.176	1.681	2.48	0.025
Coût de C	4.791	1.789	2.68	0.017

S = 1252.76    R carré = 64.5 %    R carré (ajust) = 57.5 %

### Analyse de variance

Source	DL	Somme des carrés	CM	F	P
Régression	3	42856230	14285410	9.10	0.001
Erreur résiduelle	15	23541261	1569417		
Total	18	66397491			

Source	DL	SomCar séq
Coût de A	1	23958339

Nous avons fait cela juste pour l'information et pour pouvoir comparer par rapport à ce qui va suivre.

Maintenant, allons dans le menu **Stat/Régression/Pas à pas...**:

Effectuer une régression pas à pas pour identifier les sous-ensembles utiles de prédicteurs

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Mois	Coûts	Coût de A	Coût de B	Coût de C								
1	1	44439	515	541	928								
2	2	43936	929	692	711								
3	3	44464	800	710	824								
4	4	41533	979	675	758								
5	5	46343	1165	1147	635								
6	6	44922	651	939	901								

Nous avons alors:

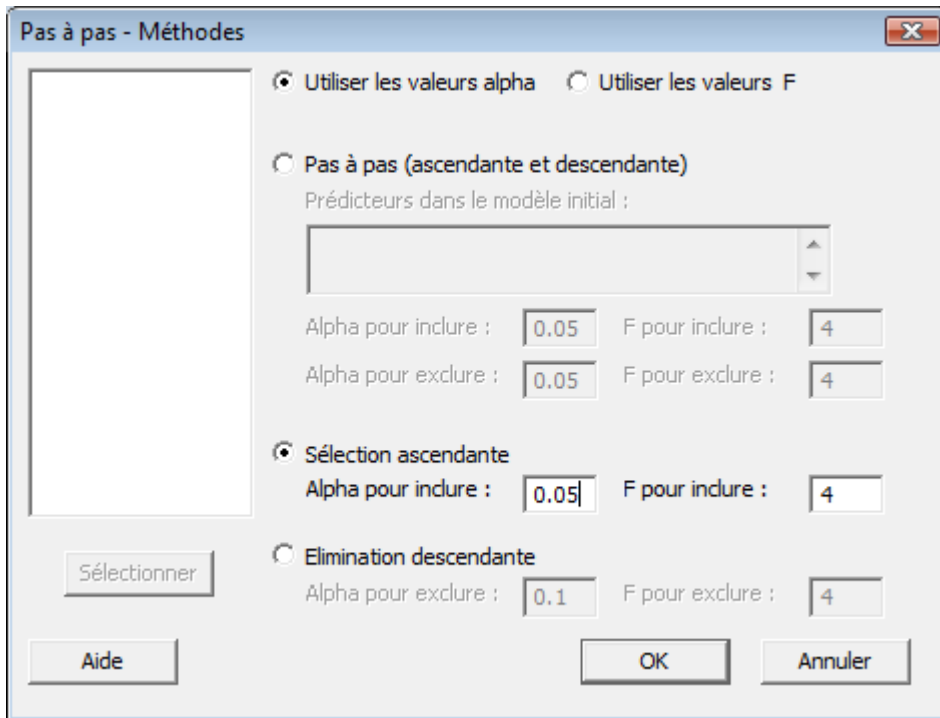
Réponse : Coûts

Prédicteurs : 'Coût de A' 'Coût de B' 'Coût de C'

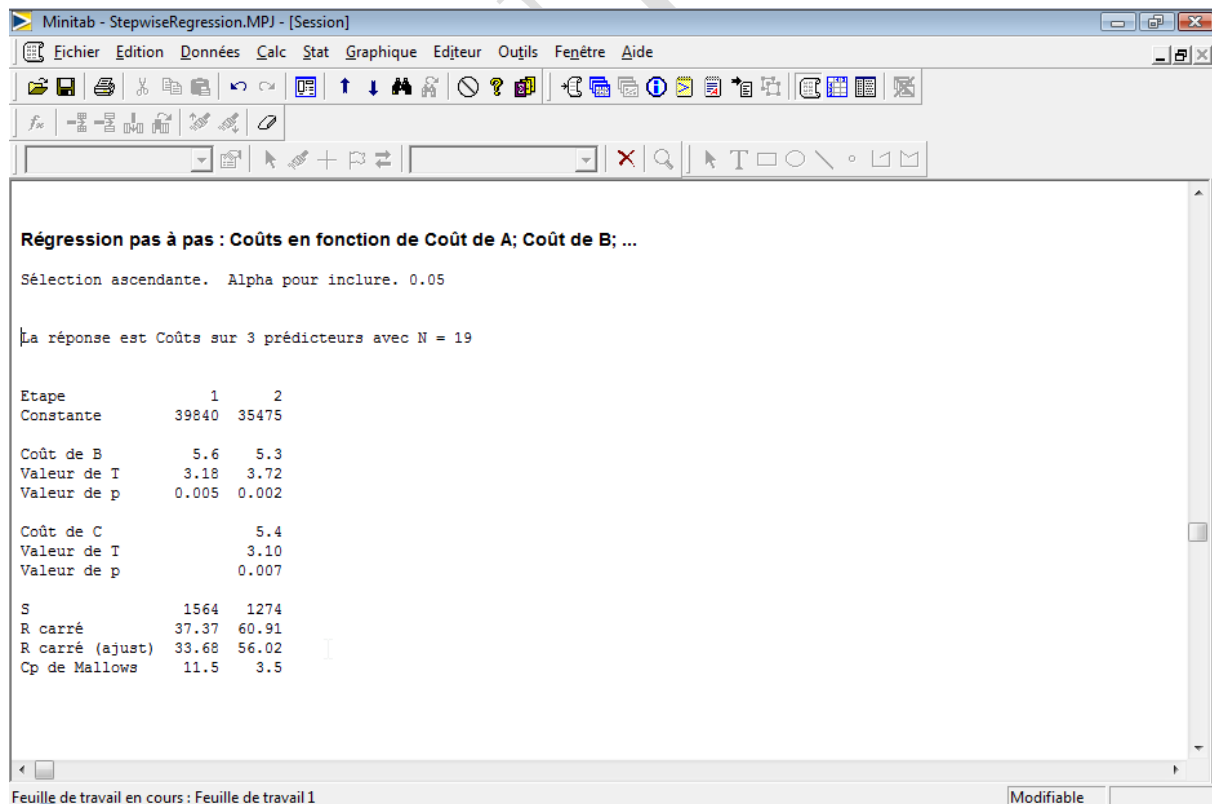
Prédicteurs à inclure dans chaque modèle :

Sélectionner Méthodes... Options... Aide OK Annuler

Cliquez sur **Méthodes...**:



Prenez le modèle **Sélection ascendante** (le plus intéressant pour rappel!!!!) pour commencer (modèle qui construit le modèle variable explicative par variable explicative en maximisant le coefficient de régression) et mettez un **alpha** (seui au-dessus duquel une variable sera incluse que si sa *p*-value en est inférieur) de 0.05 (. Validez deux fois par **OK**:



Le lecteur pourra effectivement vérifier ce résultat (en faisant trois fois une régression simple pour chacune des variables explicatives seule) que si on doit choisir un modèle avec une seule

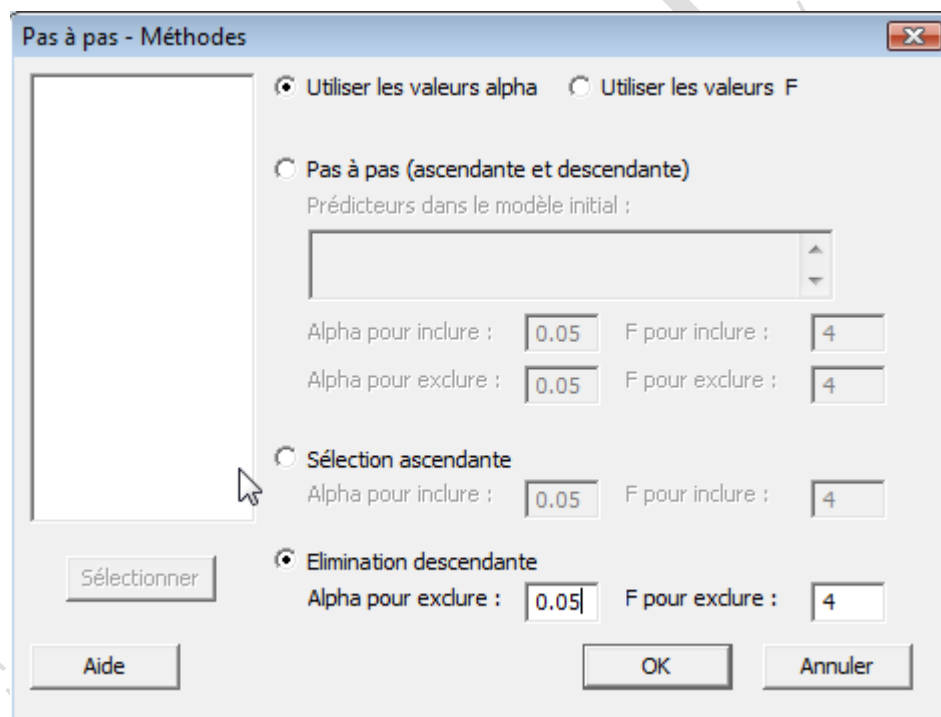
et unique variable explicative *Coût de B* est la meilleure candidate car on obtient alors le plus grand coefficient de  $R$  carré (exprimé curieusement en pourcents dans le cas présent... dans Minitab® Statistical Software???)

Ensuite, le lecteur pourra encore vérifier ce résultat (en faisant deux fois une régression simple pour chacun des couples de variables explicatives *BA* et *BC*) que si on doit choisir un modèle avec deux variables explicative *Coût de C* est la meilleure candidate pour la deuxième variable explicative car on obtient alors le plus grand coefficient de  $R$  carré.

Ensuite, la troisième étape n'a pas lieu car nous sommes en-dessous de la  $p$ -value imposée de 0.05.

Remarque: Le paramètre *Cp de Mallows* étant selon moi très difficile à interpréter (je n'ai jamais vu un article scientifique sérieux expliquant rigoureusement avec quel raisonnement il a été défini), nous allons donc en faire abstraction.

Faisons maintenant la régression pas à pas en déconstruisant le modèle. Nous retournons dans la boîte de dialogue pas à pas:



Nous sélectionnons **Élimination descendante** et nous validons par **OK** pour obtenir dans la fenêtre de session suivante:

**Régression pas à pas : Coûts en fonction de Coût de A; Coût de B; ...**

Elimination descendante. Alpha pour exclusion : 0.05

La réponse est Coûts sur 3 prédicteurs avec N = 19

Etape	1	2
Constante	35103	35475
Coût de A	2.1	
Valeur de T	1.24	
Valeur de p	0.234	
Coût de B	4.2	5.3
Valeur de T	2.48	3.72
Valeur de p	0.025	0.002
Coût de C	4.8	5.4
Valeur de T	2.68	3.10
Valeur de p	0.017	0.007
S	1253	1274
R carré	64.54	60.91
R carré (ajust)	57.45	56.02
Cp de Mallows	4.0	3.5

et encore une fois, il est très simple de vérifier le résultat en faisant de multiples régressions simples.

## 13.8. Exercice 103.: Régression pas à pas combinatoire (meilleur sous-ensemble)

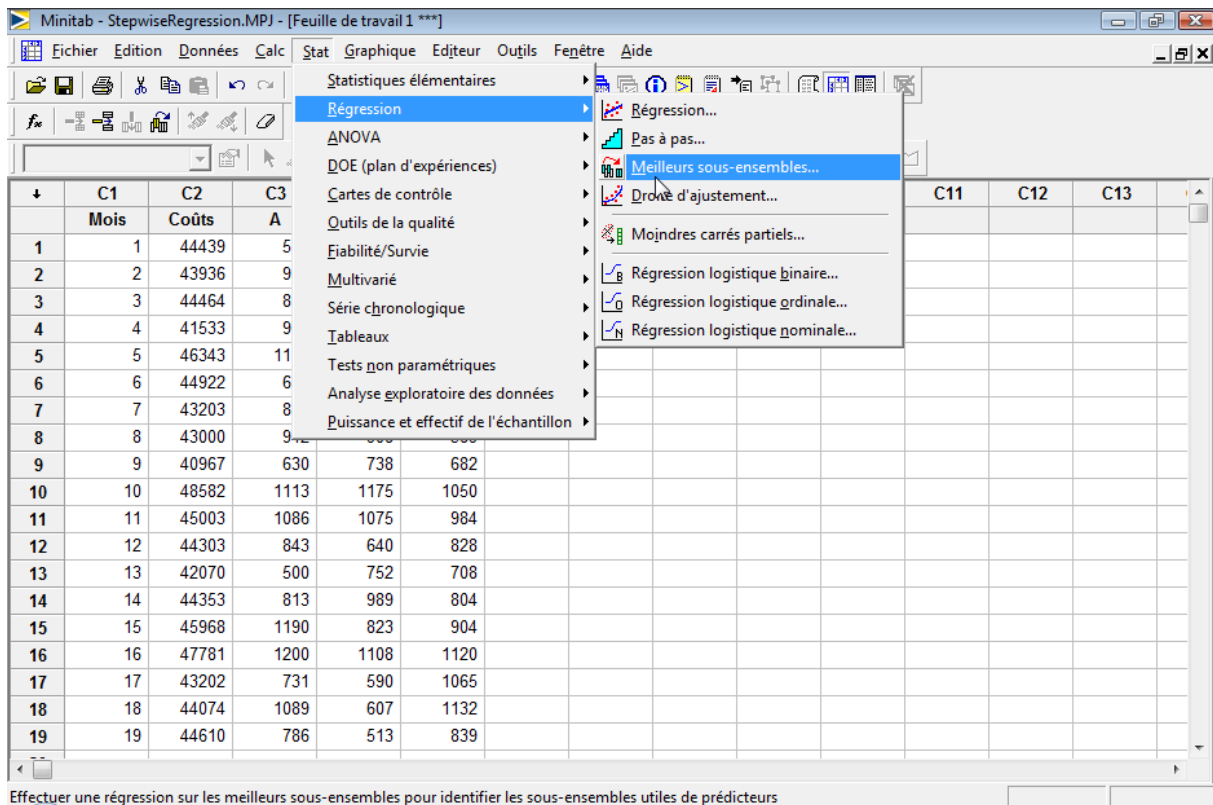
Minitab® Statistical Software 15.1.1.0

Pour la régression pas à pas, aussi appelée "régression linéaire hiérarchique" par certains américains (mais il ne faut pas confondre avec l'ANOVA hiérarchique et sa régression linéaire correspondante!), nous pourrions demander au logiciel, toujours à partir du même fichier ci-dessous:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Mois	Coûts	A	B	C								
1	1	44439	515	541	928								
2	2	43936	929	692	711								
3	3	44464	800	710	824								
4	4	41533	979	675	758								
5	5	46343	1165	1147	635								
6	6	44922	651	939	901								
7	7	43203	847	755	580								
8	8	43000	942	908	589								
9	9	40967	630	738	682								
10	10	48582	1113	1175	1050								
11	11	45003	1086	1075	984								
12	12	44303	843	640	828								
13	13	42070	500	752	708								
14	14	44353	813	989	804								
15	15	45968	1190	823	904								
16	16	47781	1200	1108	1120								
17	17	43202	731	590	1065								
18	18	44074	1089	607	1132								
19	19	44610	786	513	839								

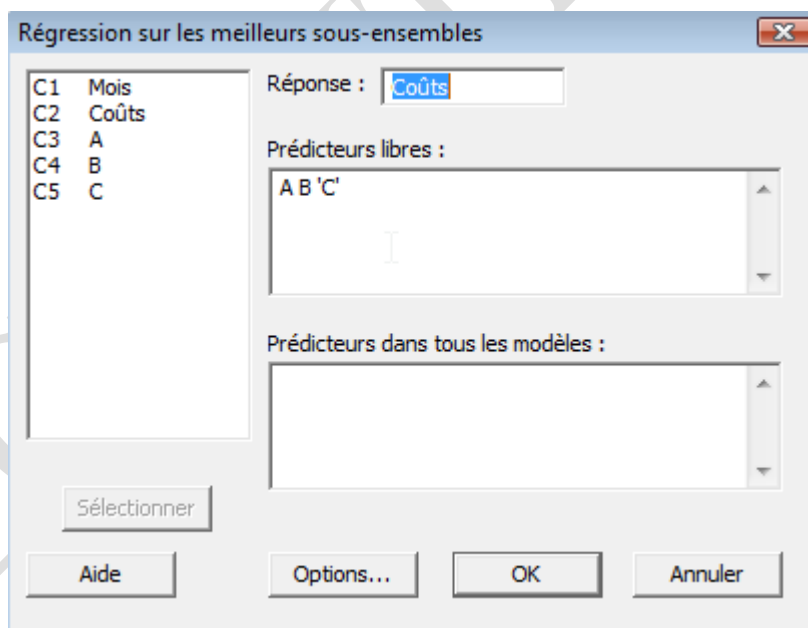
de faire une combinatoire de toutes les variables explicatives en utilisant la méthode de régression pas à pas ascendant.

Pour cela, nous allons dans le menu **Stat/Régression/Meilleurs sous-ensembles...**:



Effectuer une régression sur les meilleurs sous-ensembles pour identifier les sous-ensembles utiles de prédicteurs

Nous avons alors:



Nous validons par **OK** et nous obtenons:

02.03.2012 13:03:10

Régression sur les meilleurs sous-ensembles : Coûts en fonction de A; B; ...

Réponse : Coûts

Vars	R carré	R carré (ajust)	Cp de Mallows	S	A	B	C
1	37.4	33.7	11.5	1564.1	X		
1	36.1	32.3	12.0	1580.0	X		
2	60.9	56.0	3.5	1273.7	X	X	
2	50.0	43.7	8.2	1441.0	X	X	
3	64.5	57.5	4.0	1252.8	X	X	X

Bienvenue dans Minitab, appuyez sur F1 pour obtenir de l'aide. Modifiable

Comme nous pouvons le voir, le logiciel commence par la variable  $B$  (*Coût de B*) qui a donc le meilleur coefficient de régression. Ensuite, il montre la deuxième variable explicative seule qui est  $A$  (*Coût de A*). Il ne montre pas  $C$  car c'est la plus mauvaise des variables explicatives.

Le logiciel fait ensuite la même chose avec les couples. Il fait ainsi  $BC$  car c'est le meilleur couple. Ensuite il fait  $AC$  car c'est le deuxième meilleur. Et encore une fois il ne fait pas le dernier par élimination.

Et pour finir on fait les trois... Nous retrouvons donc les mêmes calculs qu'avec Microsoft Excel et IBM SPSS.

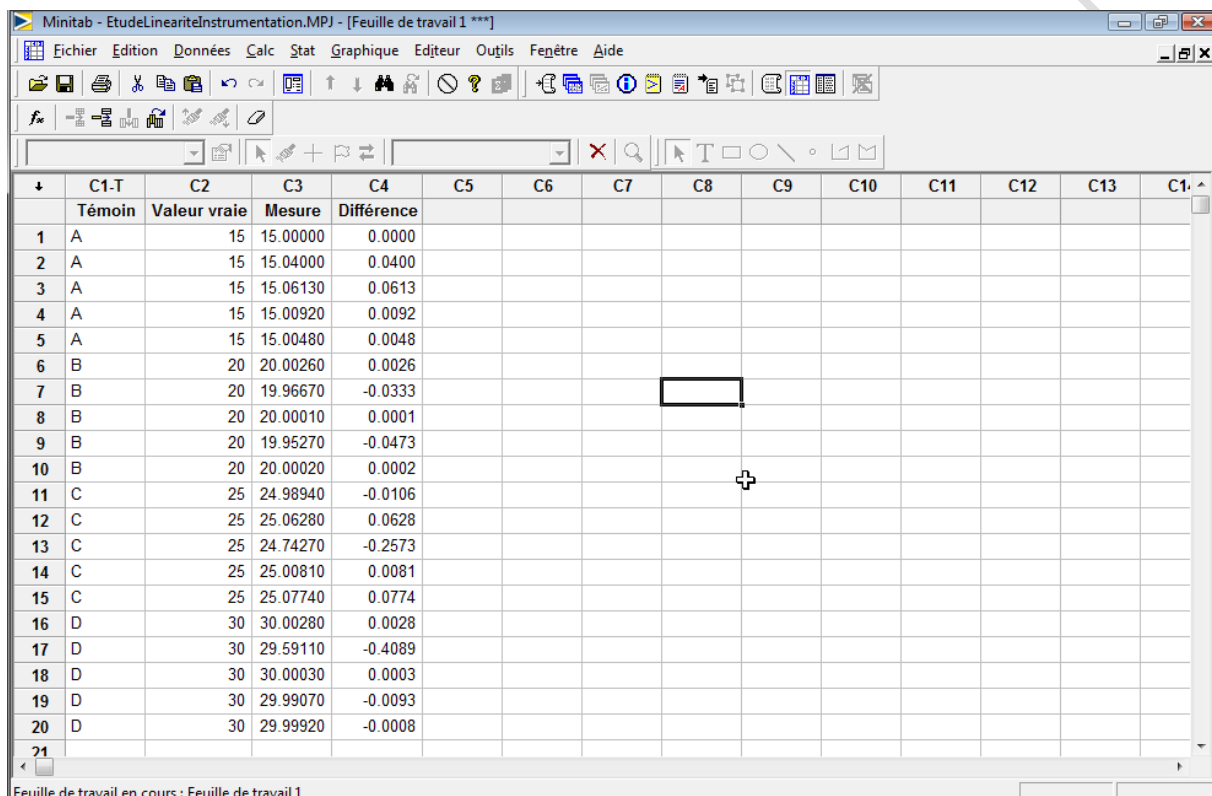


## 13.9. Exercice 104.: Étude de linéarité de l'instrumentation et biais de mesure

Minitab® Statistical Software 15.1.1.0

Il s'agit ici de voir simplement une application particulière de la régression linéaire simple.

Si dans un laboratoire, nous avons fait de mesures sur des échantillons témoins étalonnés et que nous avons obtenus le tableau suivant:



	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Témoin	Valeur vraie	Mesure	Différence										
1	A	15	15.00000	0.0000										
2	A	15	15.04000	0.0400										
3	A	15	15.06130	0.0613										
4	A	15	15.00920	0.0092										
5	A	15	15.00480	0.0048										
6	B	20	20.00260	0.0026										
7	B	20	19.96670	-0.0333										
8	B	20	20.00010	0.0001										
9	B	20	19.95270	-0.0473										
10	B	20	20.00020	0.0002										
11	C	25	24.98940	-0.0106										
12	C	25	25.06280	0.0628										
13	C	25	24.74270	-0.2573										
14	C	25	25.00810	0.0081										
15	C	25	25.07740	0.0774										
16	D	30	30.00280	0.0028										
17	D	30	29.59110	-0.4089										
18	D	30	30.00030	0.0003										
19	D	30	29.99070	-0.0093										
20	D	30	29.99920	-0.0008										
21														

En supposant que les différences (écarts) entre les mesures  $M$  et les valeurs réelles des témoins étalons  $T$  soient linéairement dépendantes (donc plus la valeur vraie est grande plus la différence de mesure le sera aussi), nous allons donc avoir un modèle théorique du type (pour chaque témoin étalon):

$$M = aT + b + \varepsilon$$

Dans le cas idéal, nous devrions évidemment avoir:

$$M = T + \varepsilon$$

mais dans la pratique il arrive plutôt que l'instrument ait un biais et donc nous nous attendons à avoir:

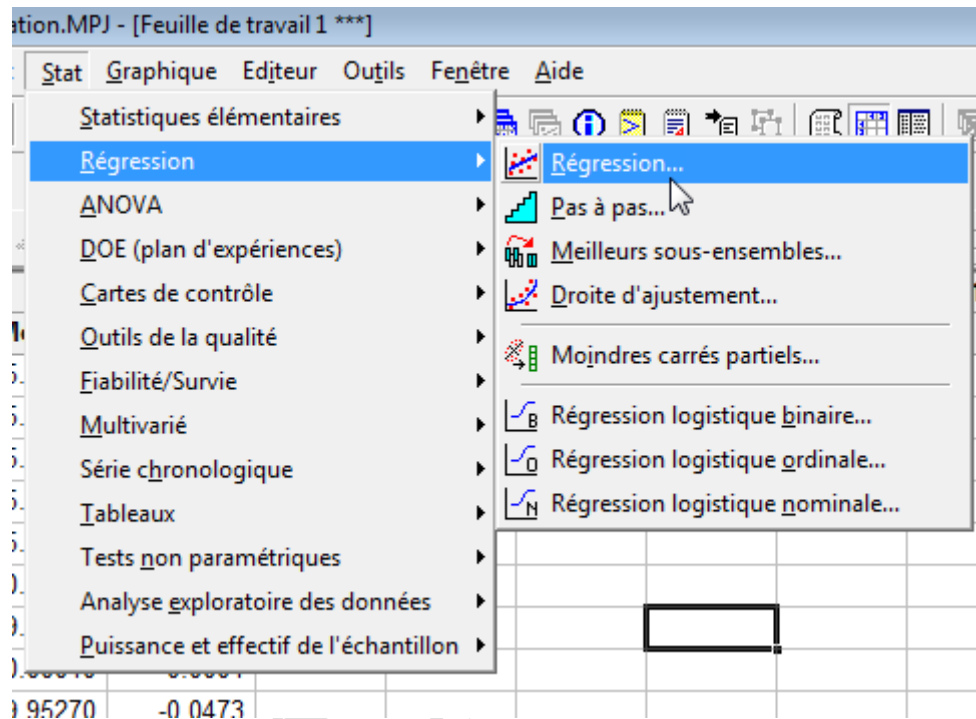
$$M = aT + \varepsilon$$

Il en va de même pour l'écart, où nous nous attendons à avoir:

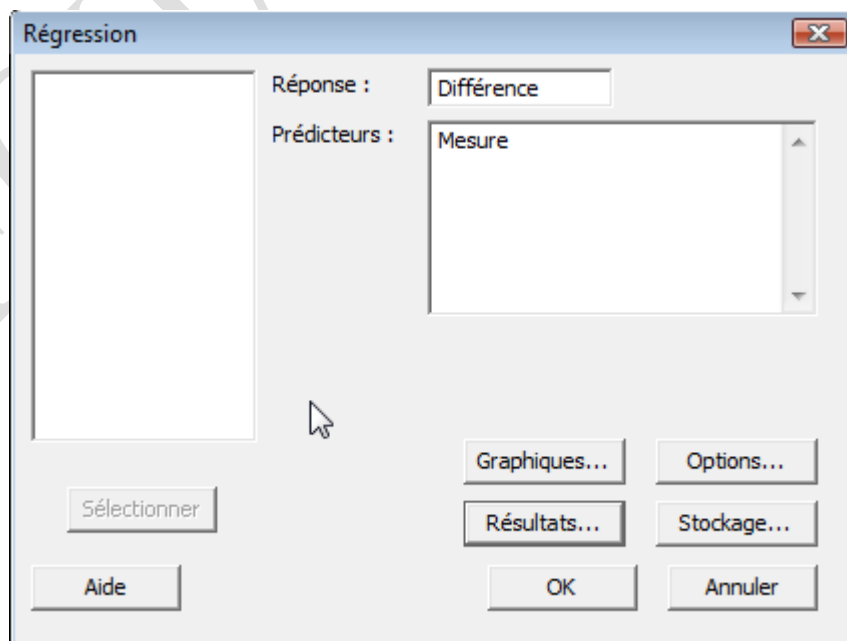
$$M - T = D = \alpha M + \beta + \varepsilon$$

Il est donc intéressant pour un expérimentateur de connaître le biais  $\beta$  (ordonnée à l'origine) au repos de son instrumentation (toujours sous l'hypothèse de linéarité!).

Ainsi, l'étude de linéarité de biais de l'instrumentation consiste à faire une simple régression linéaire des différences en fonction de la mesure étalon. Donc avec les données précédentes, dans Minitab® Statistical Software, allons d'abord dans le menu **Stat/Régression/Régression...**:



avec:



Nous validons par **OK** pour obtenir dans la fenêtre de session:

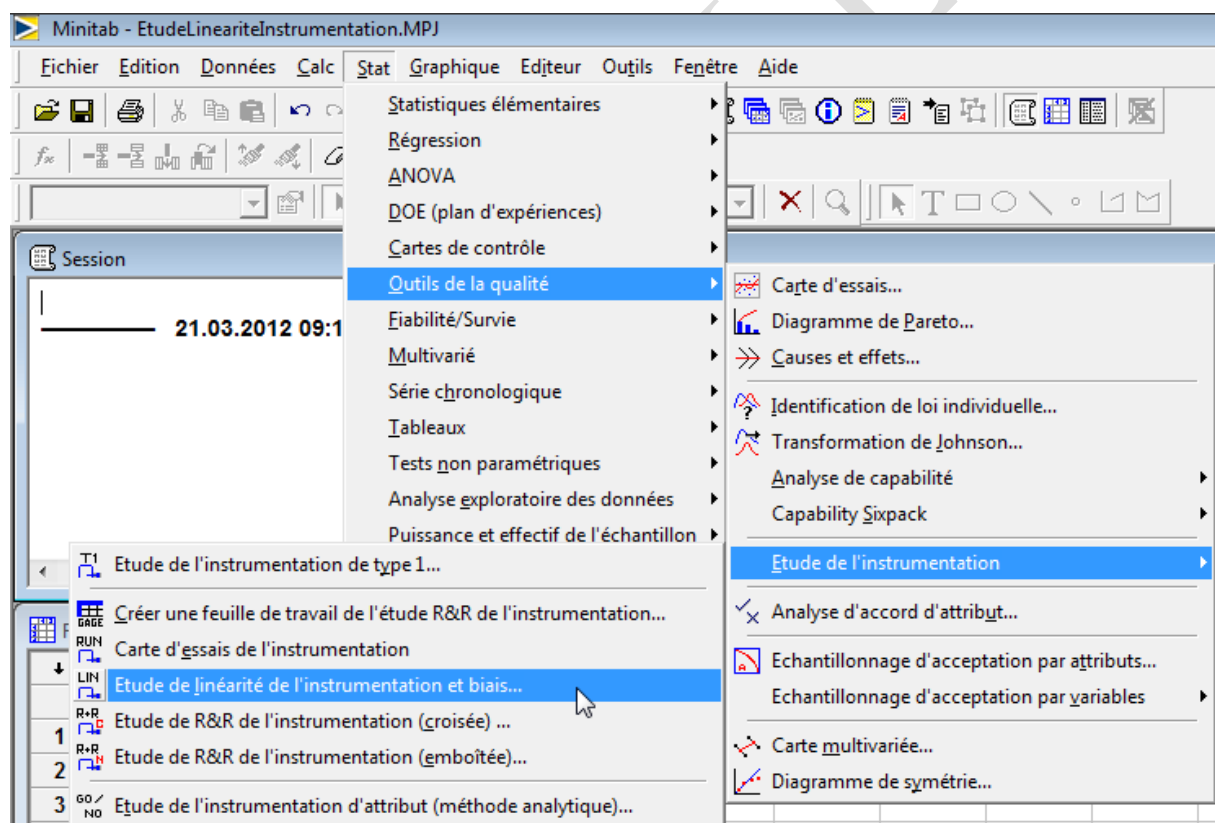
### Analyse de régression : Différence en fonction de Mesure

L'équation de régression est  
 Différence = 0.115 - 0.00624 Mesure

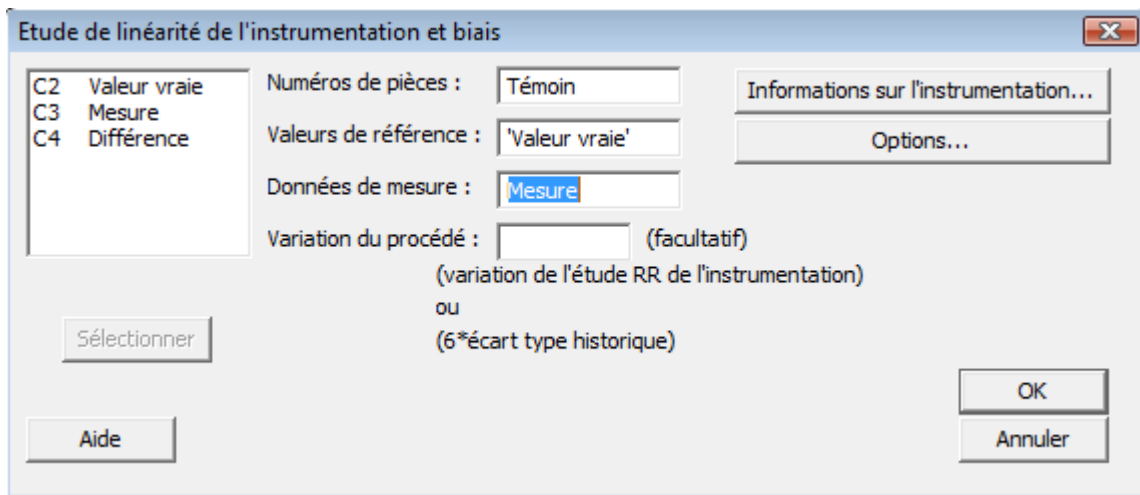
Prédicteur	Coeff	Coef ErT	T	P
Constante	0.1153	0.1021	1.13	0.274
Mesure	-0.006237	0.004410	-1.41	0.174

S = 0.109552    R carré = 10.0 %    R carré (ajust) = 5.0 %

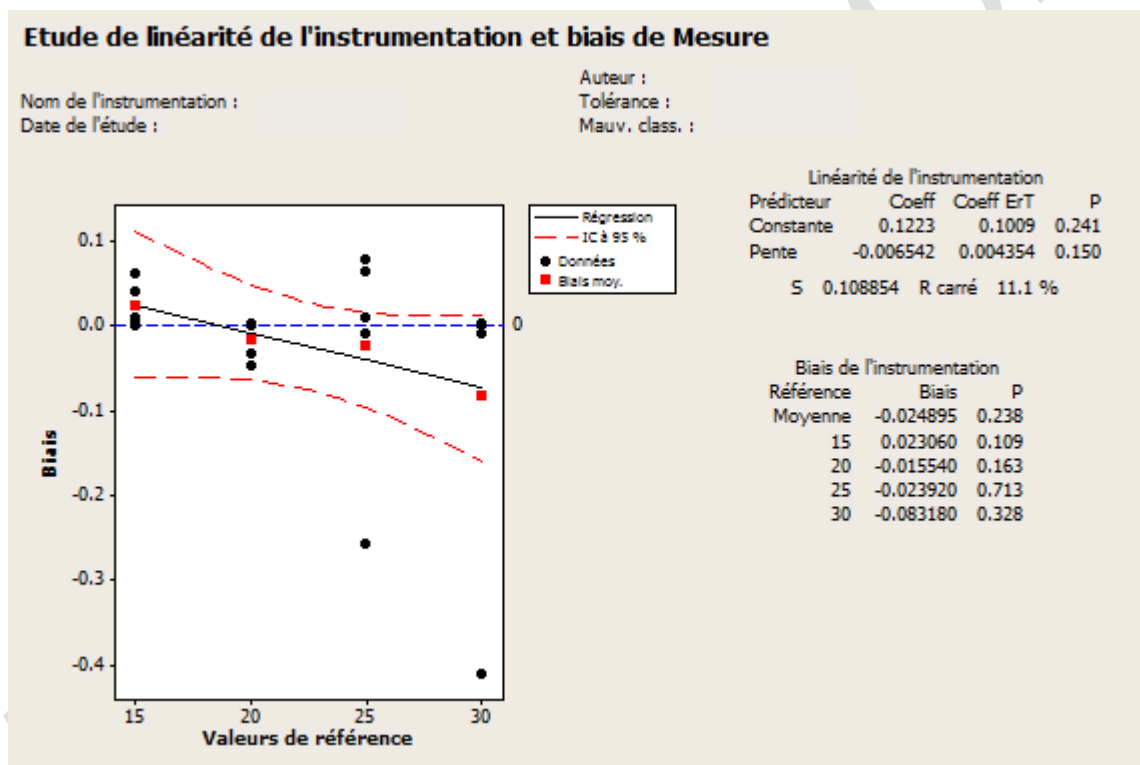
Donc l'instrumentation à un biais non significatif de 0.1153 mais par contre le modèle de semble pas linéaire. Mais continuons tout de même en voyant maintenant la même analyse mais avec un outil qui prémache une synthèse plus pertinente. Nous allons dans le menu **Stat/Outil de la qualité/Étude de l'instrumentation/Étude de linéarité de l'instrumentation et biais...**:



Avec:



Nous validons par **OK** pour avoir:



On note qu'il est dommage qu'il n'y ait pas de test de Fisher de la courbure contrairement au logiciel de Reliasoft...

Décortiquons cela (outre le graphique car nous avons déjà vu plus haut comment le générer et nous avons fait les démonstrations mathématiques dans le cours théorique) en commençant par:

Linéarité de l'instrumentation				
Prédicteur	Coeff	Coeff ErT	P	
Constante	0.1223	0.1009	0.241	
Pente	-0.006542	0.004354	0.150	
S 0.108854 R carré 11.1 %				

Nous retrouvons donc presque les mêmes valeurs qu'avec l'outil de régression linéaire standard qui pour rappel étaient:

**Analyse de régression : Différence en fonction de Mesure**

L'équation de régression est  
Différence = 0.115 - 0.00624 Mesure

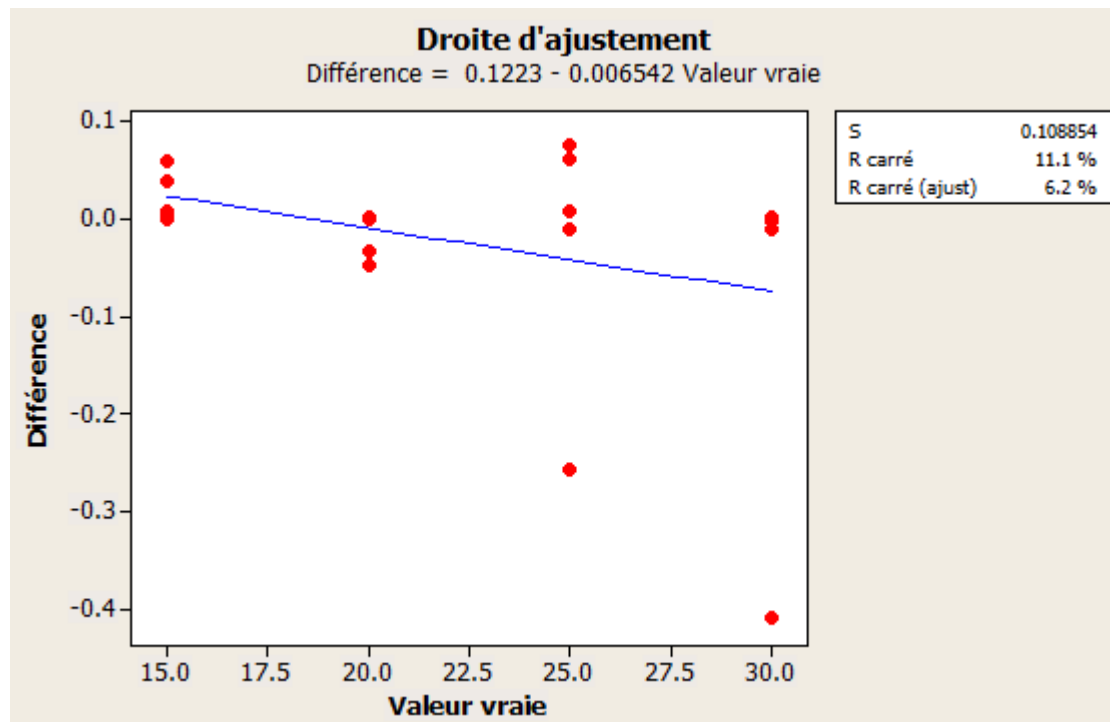
Prédicteur	Coeff	Coef ErT	T	P
Constante	0.1153	0.1021	1.13	0.274
Mesure	-0.006237	0.004410	-1.41	0.174

S = 0.109552 R carré = 10.0 % R carré (ajust) = 5.0 %

La question est alors: quelles sont les bonnes valeurs (même si les paramètres sont dans les deux cas significatifs)? Eh bien si nous nous basons sur les calculs faits dans le cours théorique et avec Microsoft Excel, c'est le premier. Par ailleurs avec Microsoft Excel nous obtenons bien le premier:

SUMMARY OUTPUT								
<i>Regression Statistics</i>								
Multiple R		0.333821903						
R Square		0.111437063						
Adjusted R Square		0.062072455						
Standard Error		0.108853793						
Observations		20						
<i>ANOVA</i>								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	1	0.026748602	0.026748602	2.257428308	0.150315682			
Residual	18	0.213284667	0.011849148					
Total	19	0.240033269						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	0.1223	0.100946854	1.211528588	0.241358481	-0.089781471	0.334381471	-0.089781471	0.334381471
Valeur vraie	-0.006542	0.004354152	-1.502474062	0.150315682	-0.015689733	0.002605733	-0.015689733	0.002605733

et avec Minitab® Statistical Software en utilisant non pas l'outil régression mais l'outil Droite d'ajustement, nous retrouvons bien:



Je ne m'explique pas actuellement cette différence alors que si nous regardons les exercices de régression faits avant, nous avons les mêmes résultats (ce qui devrait être le cas!) entre l'outil **Régression** et l'outil **Droite d'ajustement**! Affaire à suivre...

Parlons maintenant de la partie:

Biais de l'instrumentation		
Référence	Biais	P
Moyenne	-0.024895	0.238
15	0.023060	0.109
20	-0.015540	0.163
25	-0.023920	0.713
30	-0.083180	0.328

Le biais de -0.024895 est simplement la moyenne arithmétique des biais (ordonnées à l'origine des régressions linéaires de chaque sous-ensemble de valeurs):

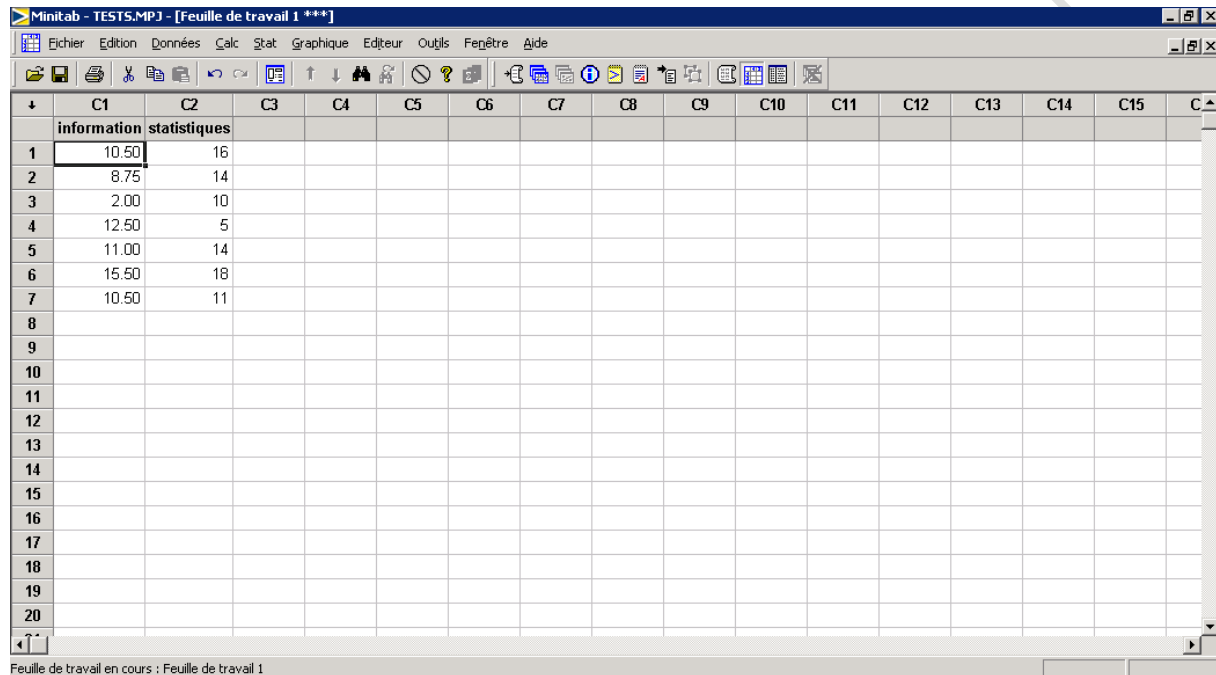
0.023060, -0.015540, -0.023920, -0.083180

que l'on peut facilement retrouver. Par contre, il ne m'a pas été possible de retrouver les  $p$ -values que Minitab® Statistical Software retourne. Affaire à suivre aussi...

## 13.10. Exercice 105.: Effectuer un calcul de la corrélation $R$ (de Pearson) simple

Minitab® Statistical Software 16

Ouvrez le fichier *Tests.mpj* qui contient les points obtenus par des employés qui ont passé des tests de compétences dans une entreprise d'ingénierie (chaque ligne correspondant à un employé):

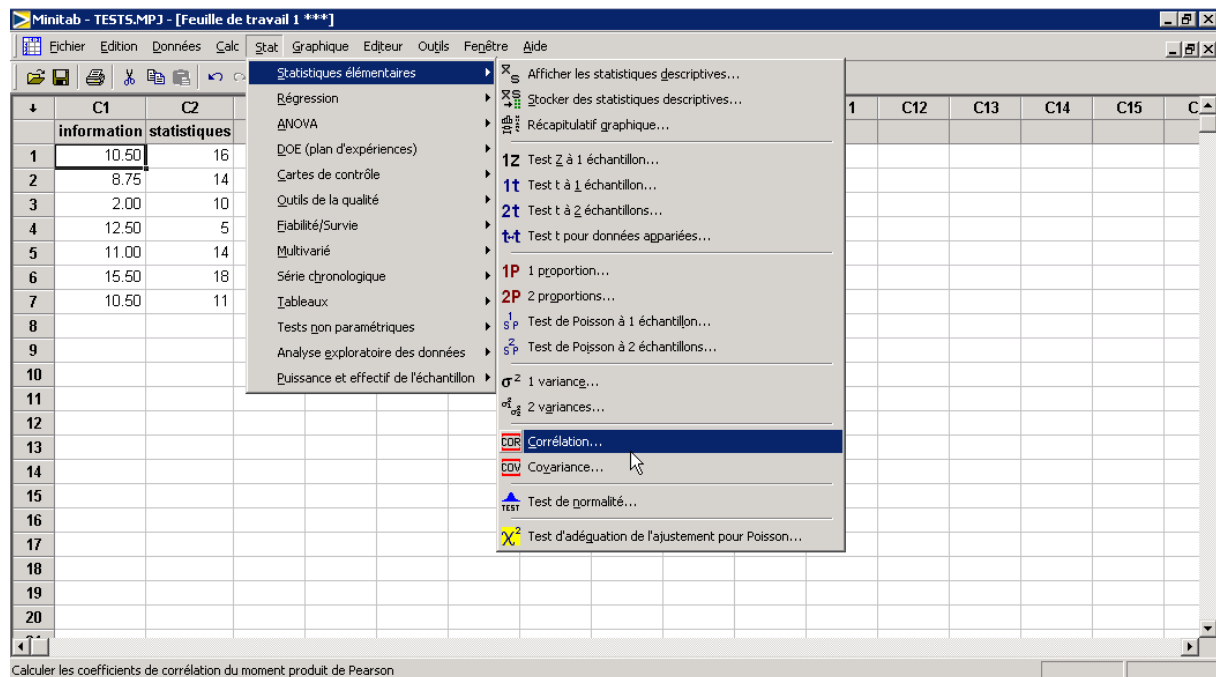


The screenshot shows the Minitab Statistical Software 16 interface. The title bar reads "Minitab - TESTS.MPJ - [Feuille de travail 1 \*\*\*]". The menu bar includes "Fichier", "Edition", "Données", "Calc", "Stat", "Graphique", "Editeur", "Outils", "Fenêtre", and "Aide". The toolbar contains various icons for file operations, editing, and statistical analysis. The worksheet grid has columns labeled C1 through C15 and rows numbered 1 through 20. The first two columns, C1 and C2, contain data for two variables. The first row of data is labeled "information" and "statistiques". The data points are as follows:

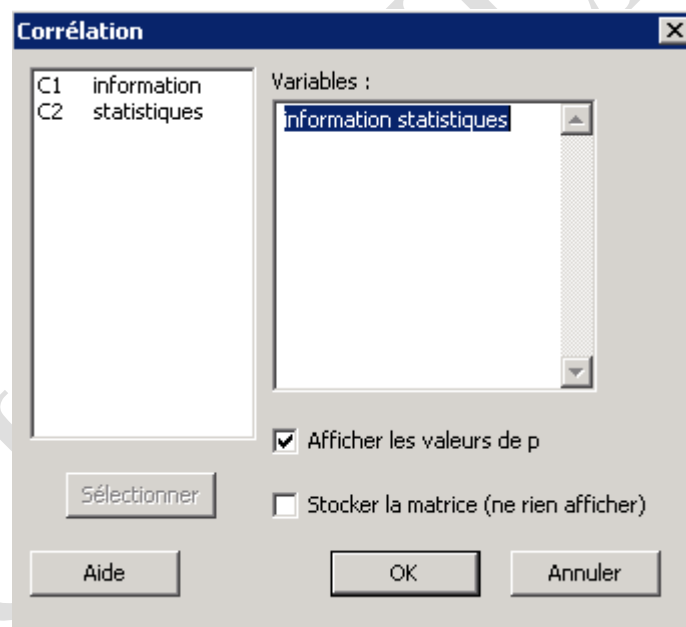
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
1	10.50	16													
2	8.75	14													
3	2.00	10													
4	12.50	5													
5	11.00	14													
6	15.50	18													
7	10.50	11													
8															
9															
10															
11															
12															
13															
14															
15															
16															
17															
18															
19															
20															

Nous souhaiterions savoir si les compétences sont corrélées (facteur de corrélation  $R$ ) et si nous obtenons le même résultat qu'avec Microsoft Excel:

Nous allons dans le menu **Stat/Statistiques élémentaires/Corrélation...**:



Nous avons alors:



Ce qui nous donne dans la fenêtre de session:

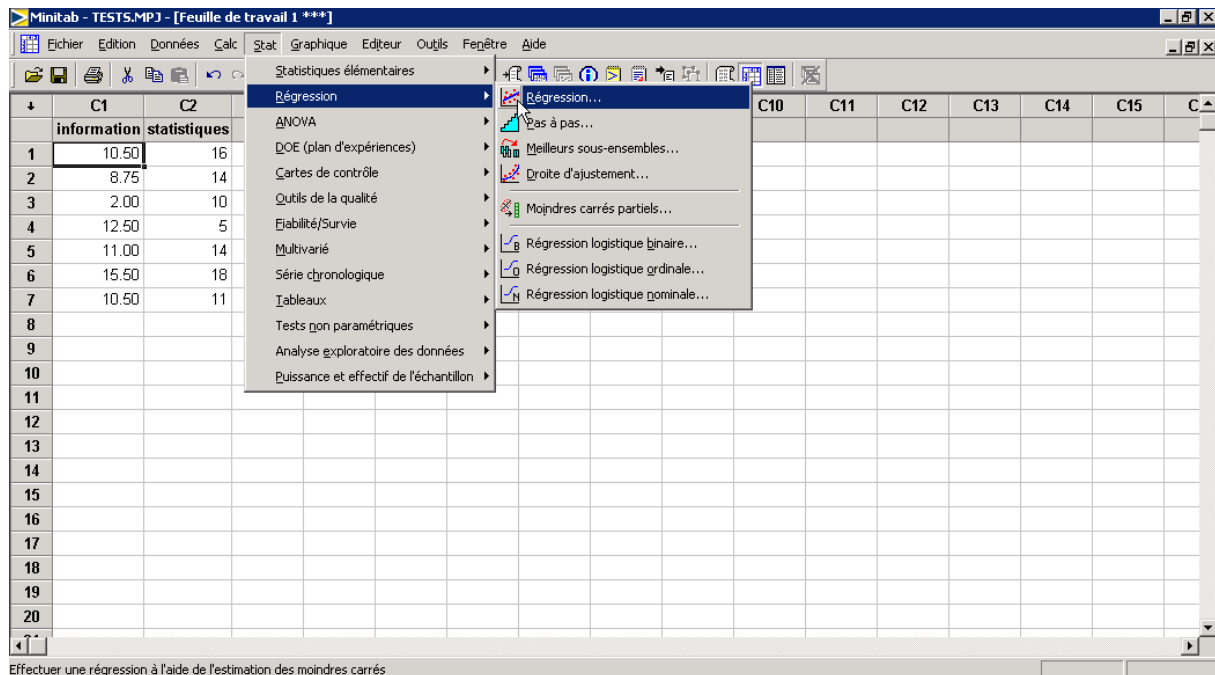
### Corrélations : information; statistiques

```
Corrélation de Pearson de information et statistiques = 0.298
Valeur de p = 0.516
```

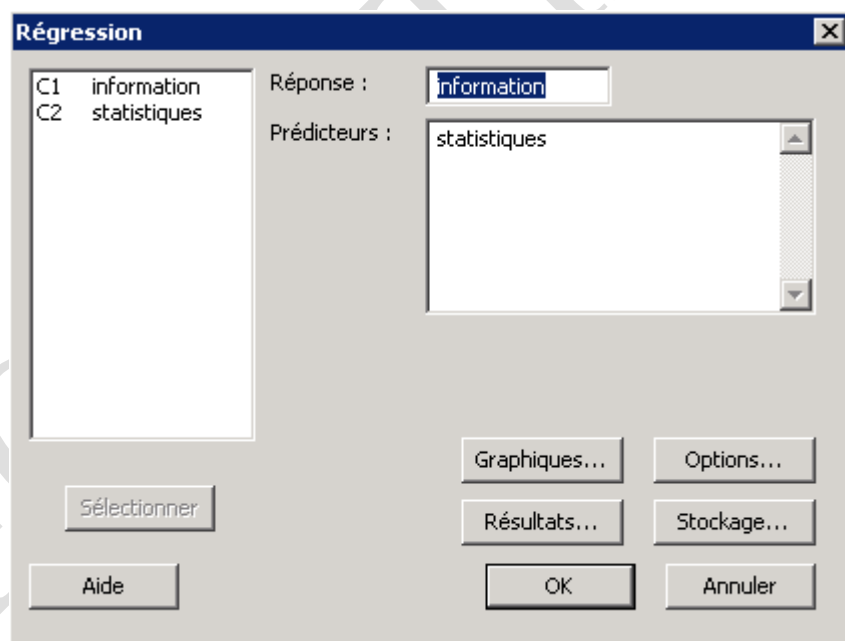
Ce qui correspond bien au  $R$  que donne Microsoft Excel (puisque ce dernier donne le  $R^2$  il faut prendre la racine carré du coefficient de corrélation de Microsoft Excel pour voir que cela est bien conforme).



Par ailleurs, nous retrouvons aussi la racine carrée de la valeur de Minitab si nous faisons une régression linéaire simple:



avec:



en validant par **OK** nous avons dans la fenêtre de session le morceau suivant:

**Analyse de régression : information en fonction de statistiques**

L'équation de régression est  
 information = 6.50 + 0.287 statistiques

Prédicteur	Coeff	Coef ErT	T	P
Constante	6.498	5.419	1.20	0.284
statistiques	0.2871	0.4108	0.70	0.516

S = 4.34223 R carré = 8.9 % R carré (ajust) = 0.0 %

**Analyse de variance**

Source	DL	Somme des carrés	CM	F	P
Régression	1	9.21	9.21	0.49	0.516
Erreur résiduelle	5	94.27	18.85		
Total	6	103.48			

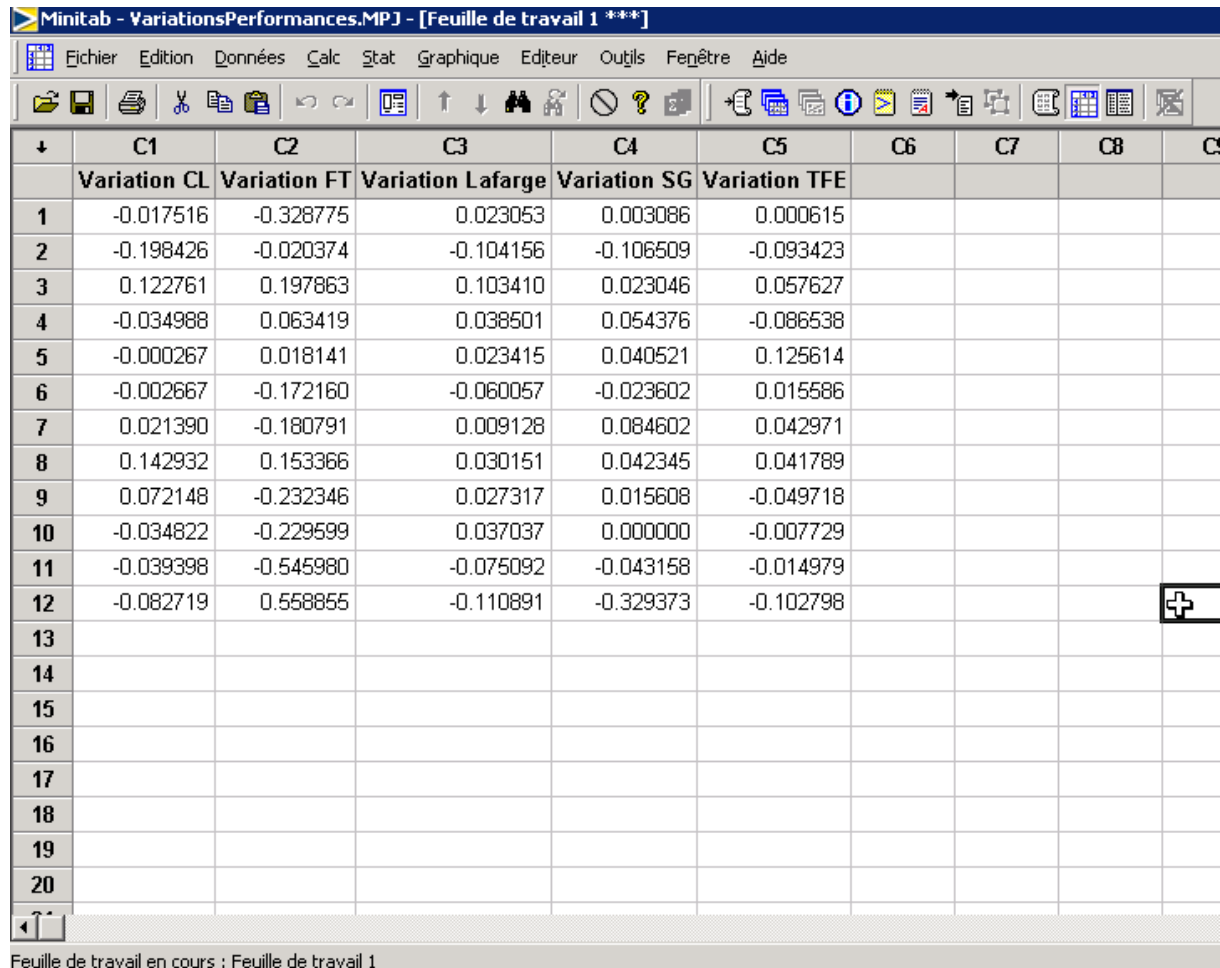
et racine carrée de 8.9% donne bien le résultat obtenu plus haut de 0.298

Donc ce n'est pas parce qu'un employé est bon avec la statistique qu'il est bon avec l'informatique... (cela se confirme dans la pratique...).

## 13.11. Exercice 106.: Matrice de corrélation $R$ de Pearson

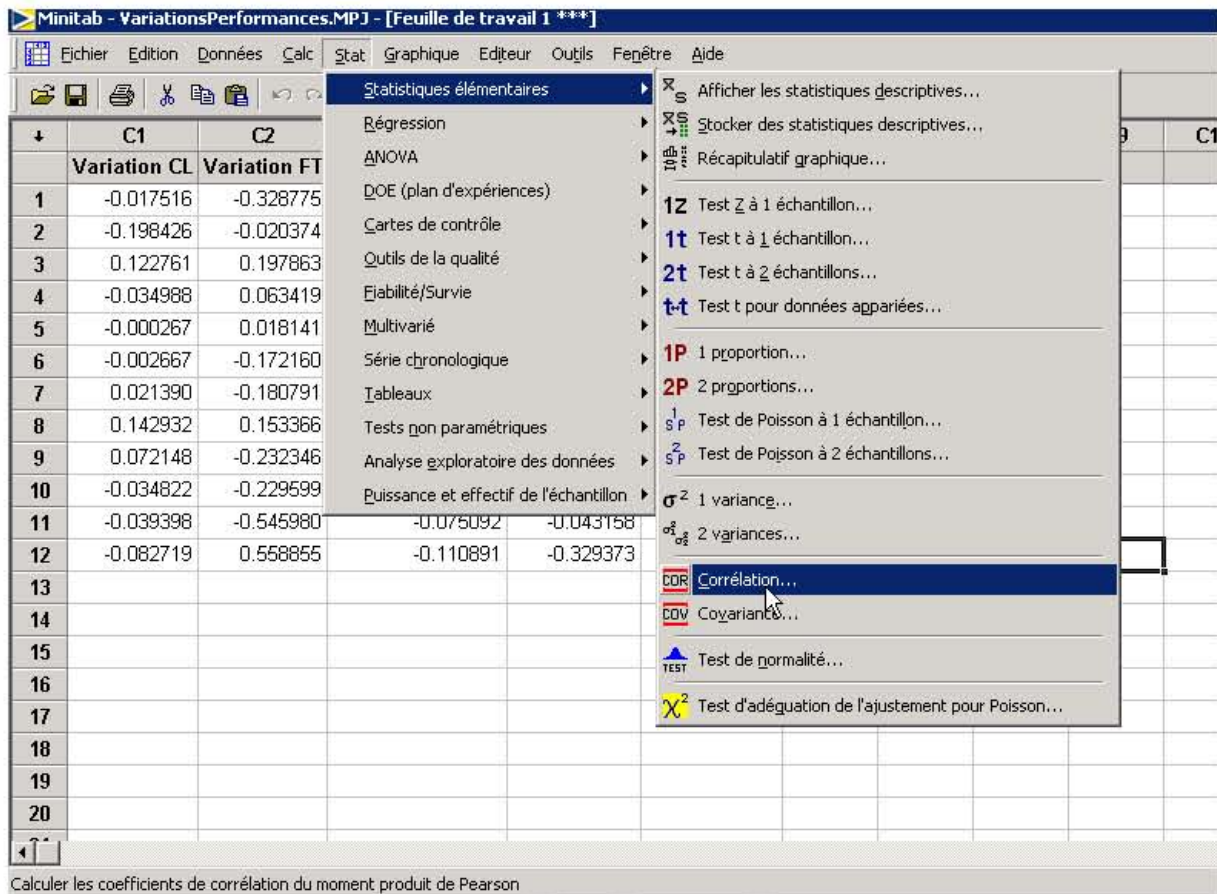
Minitab® Statistical Software 15.1.1.0

Ouvrez le fichier *VariationsPerformances.mpj*:

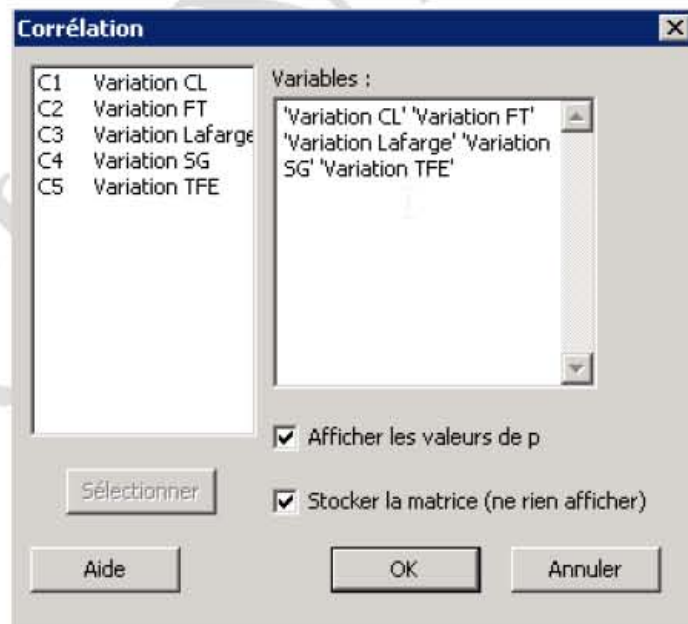


	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
	Variation CL	Variation FT	Variation Lafarge	Variation SG	Variation TFE					
1	-0.017516	-0.328775	0.023053	0.003086	0.000615					
2	-0.198426	-0.020374	-0.104156	-0.106509	-0.093423					
3	0.122761	0.197863	0.103410	0.023046	0.057627					
4	-0.034988	0.063419	0.038501	0.054376	-0.086538					
5	-0.000267	0.018141	0.023415	0.040521	0.125614					
6	-0.002667	-0.172160	-0.060057	-0.023602	0.015586					
7	0.021390	-0.180791	0.009128	0.084602	0.042971					
8	0.142932	0.153366	0.030151	0.042345	0.041789					
9	0.072148	-0.232346	0.027317	0.015608	-0.049718					
10	-0.034822	-0.229599	0.037037	0.000000	-0.007729					
11	-0.039398	-0.545980	-0.075092	-0.043158	-0.014979					
12	-0.082719	0.558855	-0.110891	-0.329373	-0.102798					
13										
14										
15										
16										
17										
18										
19										
20										

Nous souhaitons calculer la matrice de corrélation pour vérifier que nous obtenons la même que dans le cours Microsoft Excel. Nous allons donc à nouveau dans le menu Nous allons dans le menu **Stat/Statistiques élémentaires/Corrélation...**:



Avec:



où nous avons coché l'option **Stocker la matrice** car elle est affichée plus élégamment et de façon plus complète que dans la fenêtre de session. Nous validons par OK et alors rien n'apparaît, ni dans le feuille, ni dans le fenêtre de session. Il faut aller dans le menu **Données/Afficher les données...**:

	C1	C5	C6	C7	C8	C9
	Variation CL	Variation TFE				
1	-0.017516	0.000615				
2	-0.198426	-0.093423				
3	0.122767	0.057627				
4	-0.034986	-0.086538				
5	-0.000267	0.125614				
6	-0.002667	0.015586				
7	0.021390	0.042971				
8	0.142932	0.041789				
9	0.072148	-0.049718				
10	-0.034822	-0.007729				
11	-0.039396	-0.014979				
12	-0.082719	-0.102798				
13						
14						
15						
16						
17						

et prendre:

en validant par **OK** nous obtenons:

**Affichage des données**

Matrice CORRELATIO1

1.00000	0.04781	0.73339	0.53702	0.59299
0.04781	1.00000	-0.03864	-0.49978	-0.15762
0.73339	-0.03864	1.00000	0.73919	0.52883
0.53702	-0.49978	0.73919	1.00000	0.59172
0.59299	-0.15762	0.52883	0.59172	1.00000

Soit une matrice de corrélation plus élégante que l'option par défaut mais sans les légendes...  
 Pour rappel, avec Microsoft Excel nous avons obtenu:

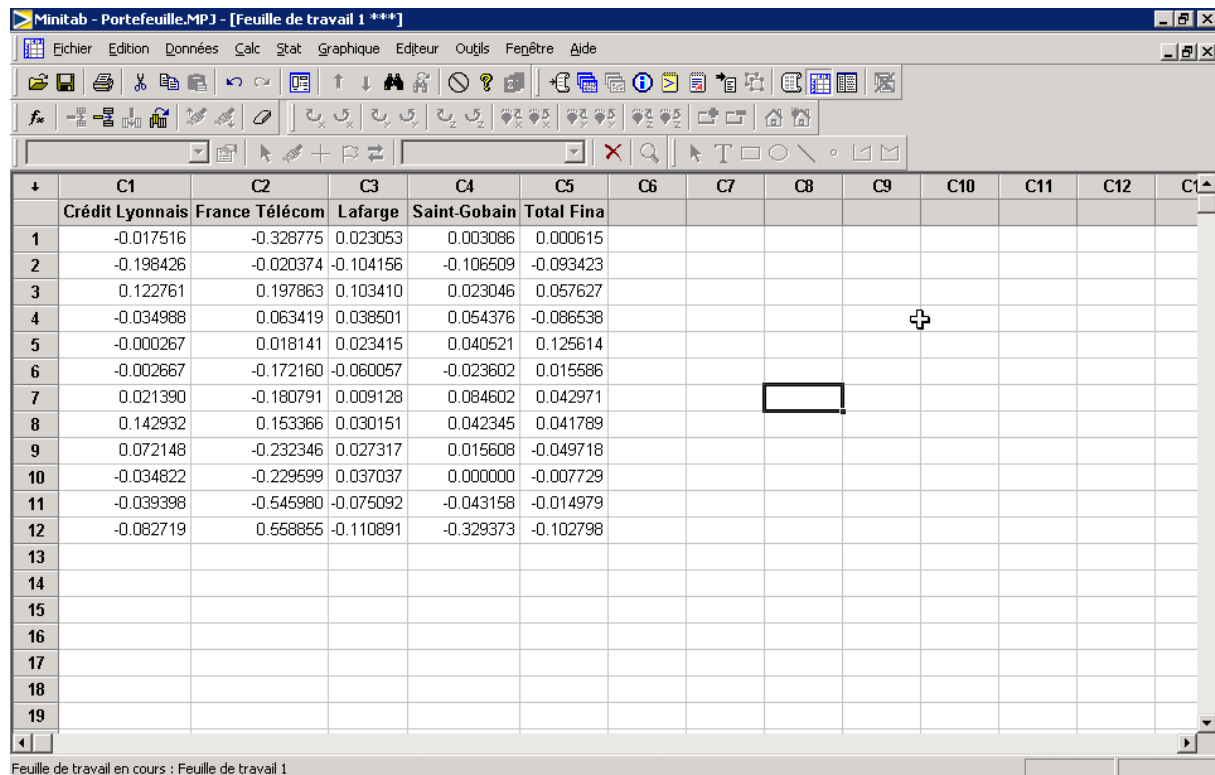
	A	B	C	D	E	F
1		<i>Variation CL</i>	<i>Variation FT</i>	<i>Variation Lafarge</i>	<i>Variation SG</i>	<i>Variation TFE</i>
2	Variation CL	1				
3	Variation FT	0.047808998	1			
4	Variation Lafarge	0.733388875	-0.038639388	1		
5	Variation SG	0.537019415	-0.499780883	0.739191421	1	
6	Variation TFE	0.592994327	-0.157621731	0.528833682	0.591723389	1

Donc c'est parfaitement conforme.

## 13.12. Exercice 107.: Effectuer un calcul de la covariance

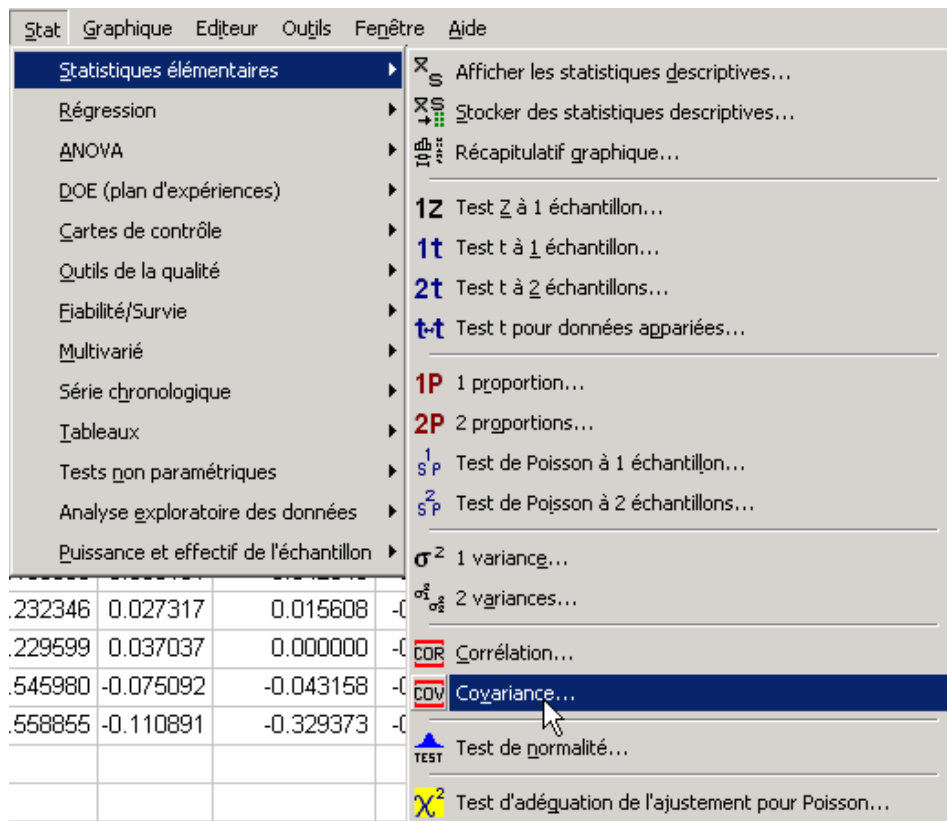
Minitab® Statistical Software 15.1.1.0

Ouvrez le fichier *Portefeuille.mpj* avec les données de rendement que nous avons utilisées dans le cours d'analyse financière et dans l'exercice précédent:

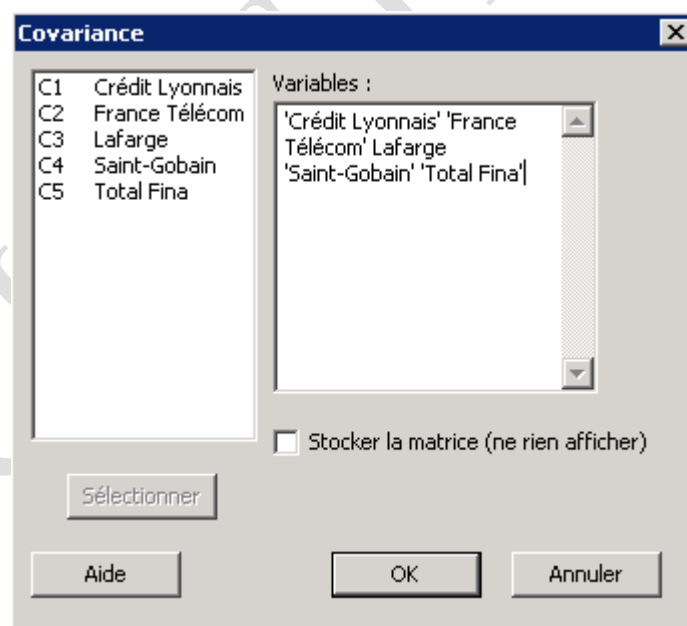


	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Crédit Lyonnais	France Télécom	Lafarge	Saint-Gobain	Total Fina								
1	-0.017516	-0.328775	0.023053	0.003086	0.000615								
2	-0.198426	-0.020374	-0.104156	-0.106509	-0.093423								
3	0.122761	0.197863	0.103410	0.023046	0.057627								
4	-0.034988	0.063419	0.038501	0.054376	-0.086538								
5	-0.000267	0.018141	0.023415	0.040521	0.125614								
6	-0.002667	-0.172160	-0.060057	-0.023602	0.015586								
7	0.021390	-0.180791	0.009128	0.084602	0.042971								
8	0.142932	0.153366	0.030151	0.042345	0.041789								
9	0.072148	-0.232346	0.027317	0.015608	-0.049718								
10	-0.034822	-0.229599	0.037037	0.000000	-0.007729								
11	-0.039398	-0.545980	-0.075092	-0.043158	-0.014979								
12	-0.082719	0.558855	-0.110891	-0.329373	-0.102798								
13													
14													
15													
16													
17													
18													
19													

Allez dans le menu **Stat/Statistiques élémentaires/Covariance...**:



Nous avons alors:



En validant il vient dans la fenêtre de sessions la matrice des variances-covariances sous une forme peu habituelle (raison pour laquelle il est préférable de stocker la matrice et de l'afficher plus tard comme nous l'avons fait dans l'exercice précédent):



**Covariances : Crédit Lyonnais; France Télécom; Lafarge; Saint-Gobain; Total Fina**

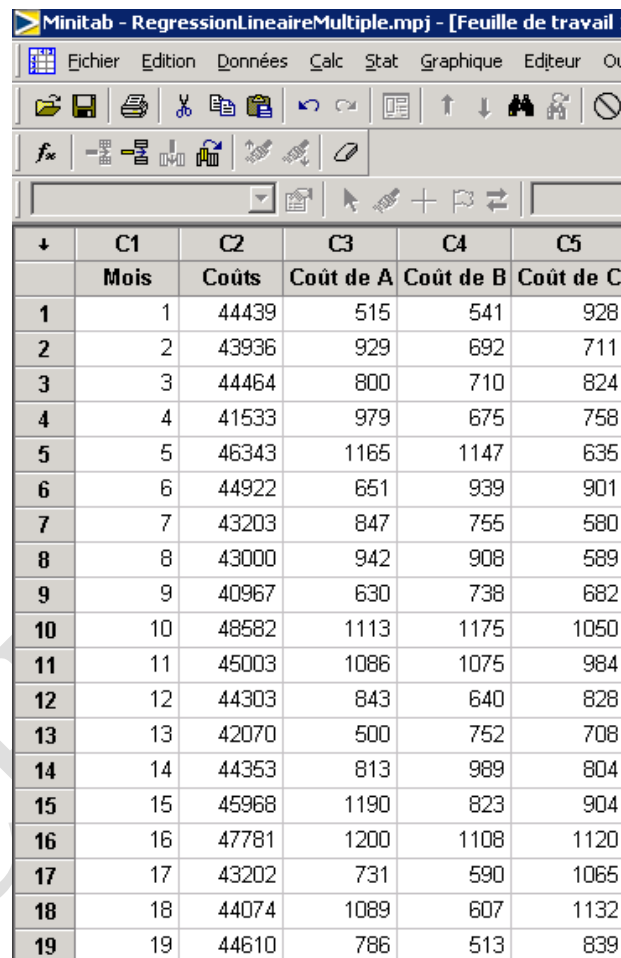
	Crédit Lyonnais	France Télécom	Lafarge
Crédit Lyonnais	0.00831448		
France Télécom	0.00125522	0.08290560	
Lafarge	0.00443929	-0.00073856	0.00440680
Saint-Gobain	0.00536049	-0.01575319	0.00537174
Total Fina	0.00371970	-0.00312211	0.00241502
	Saint-Gobain	Total Fina	
Saint-Gobain	0.01198378		
Total Fina	0.00445611	0.00473239	

Donc nous remarquons dans un premier temps que Microsoft Excel est beaucoup plus pratique par rapport à l'obtention de cette matrice. Quelque chose de très bien par contre! Minitab® Statistical Software utilise l'estimateur non biaisé de la variance alors que Microsoft Excel fait usage de la variance vraie (estimateur biaisé). Il est cependant dommage que Minitab® Statistical Software ne propose pas quelle variance nous voulons utiliser (ce qui serait quand même un minimum pour un logiciel de statistiques).

### 13.13. Exercice 108.: Effectuer une régression linéaire multiple (modèle linéaire Gaussien)

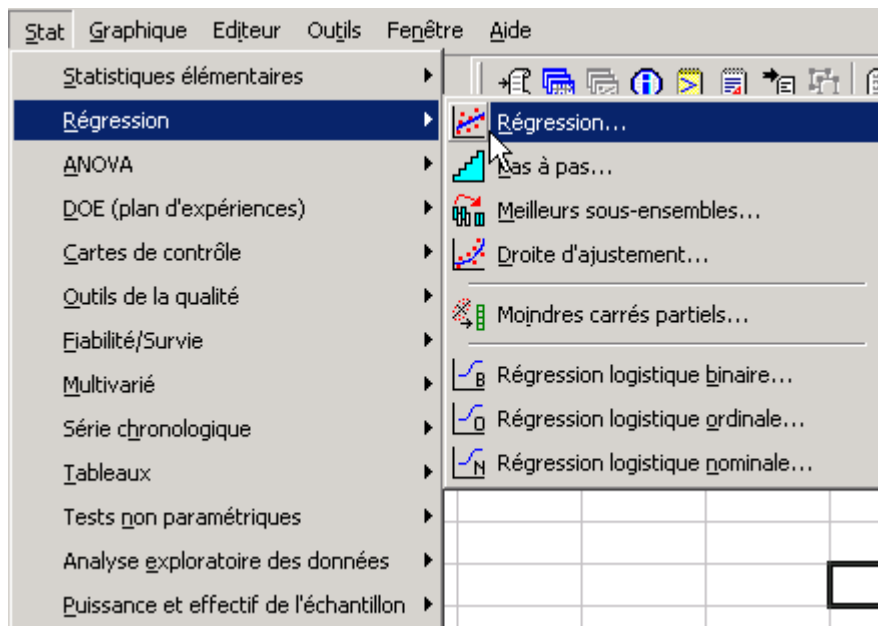
Minitab® Statistical Software 15.1.1.0

Ouvrez le fichier *RegressionLineaireMultiple.mpj* qui donne les ventes d'un produit en fonction des mois et de l'investissement dans certaines stratégies marketing. Sur l'hypothèse que la relation est linéaire, faisons une analyse par le modèle linéaire Gaussien conformément aux démonstrations effectuées dans le cours de statistique théorique et des exemples faits avec Microsoft Excel (afin de voir les éventuelles différences):

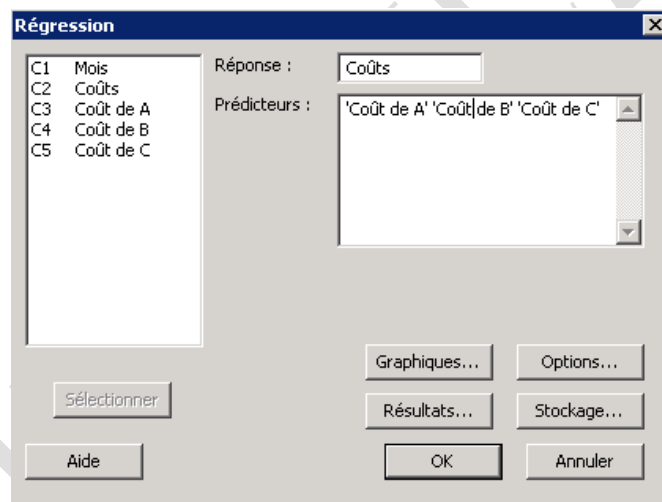


↓	C1	C2	C3	C4	C5
	Mois	Coûts	Coût de A	Coût de B	Coût de C
1	1	44439	515	541	928
2	2	43936	929	692	711
3	3	44464	800	710	824
4	4	41533	979	675	758
5	5	46343	1165	1147	635
6	6	44922	651	939	901
7	7	43203	847	755	580
8	8	43000	942	908	589
9	9	40967	630	738	682
10	10	48582	1113	1175	1050
11	11	45003	1086	1075	984
12	12	44303	843	640	828
13	13	42070	500	752	708
14	14	44353	813	989	804
15	15	45968	1190	823	904
16	16	47781	1200	1108	1120
17	17	43202	731	590	1065
18	18	44074	1089	607	1132
19	19	44610	786	513	839

Nous allons dans le menu **Stat/Régression/Régression...**:



où nous faisons le minimum vital (pour ne pas répéter ce qui a déjà été vu avec la régression linéaire simple):



Nous validons par **OK** pour obtenir:

**Analyse de régression : Coûts en fonction de Coût de A; Coût de B; ...**

L'équation de régression est

$$\text{Coûts} = 35103 + 2.07 \text{ Coût de A} + 4.18 \text{ Coût de B} + 4.79 \text{ Coût de C}$$

Prédicteur	Coeff	Coef ErT	T	P
Constante	35103	1837	19.11	0.000
Coût de A	2.066	1.665	1.24	0.234
Coût de B	4.176	1.681	2.48	0.025
Coût de C	4.791	1.789	2.68	0.017

S = 1252.76    R carré = 64.5 %    R carré (ajust) = 57.5 %

**Analyse de variance**

Source	DL	Somme des carrés	CM	F	P
Régression	3	42856230	14285410	9.10	0.001
Erreur résiduelle	15	23541261	1569417		
Total	18	66397491			

Source	DL	SomCar séq
Coût de A	1	23958339
Coût de B	1	7647915
Coût de C	1	11249975

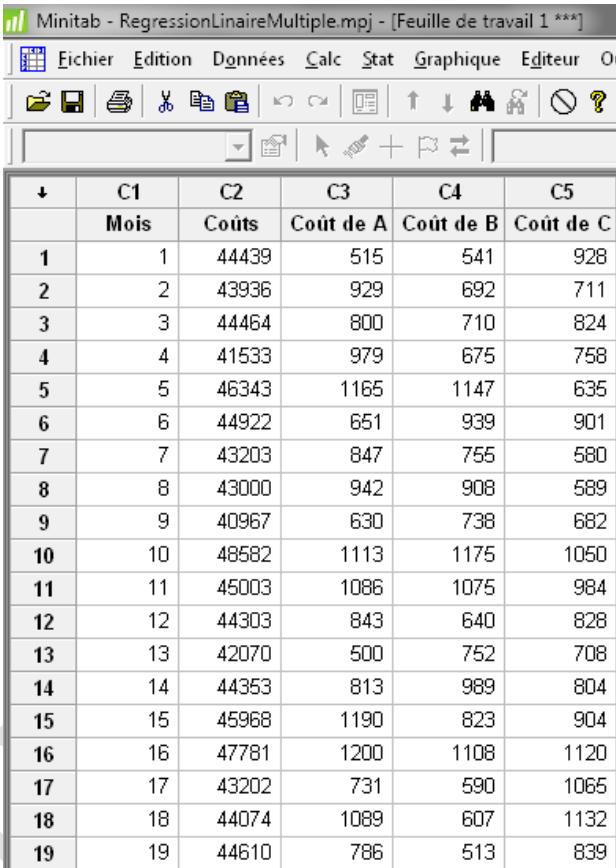
Nous retrouvons donc les mêmes valeurs que dans Microsoft Excel et elles sont donc conformes à ce que nous avons calculé dans l'étude théorique de la régression linéaire multiple.

## 13.14. Exercice 109.: Déterminer les corrélations partielles

Minitab® Statistical Software 16

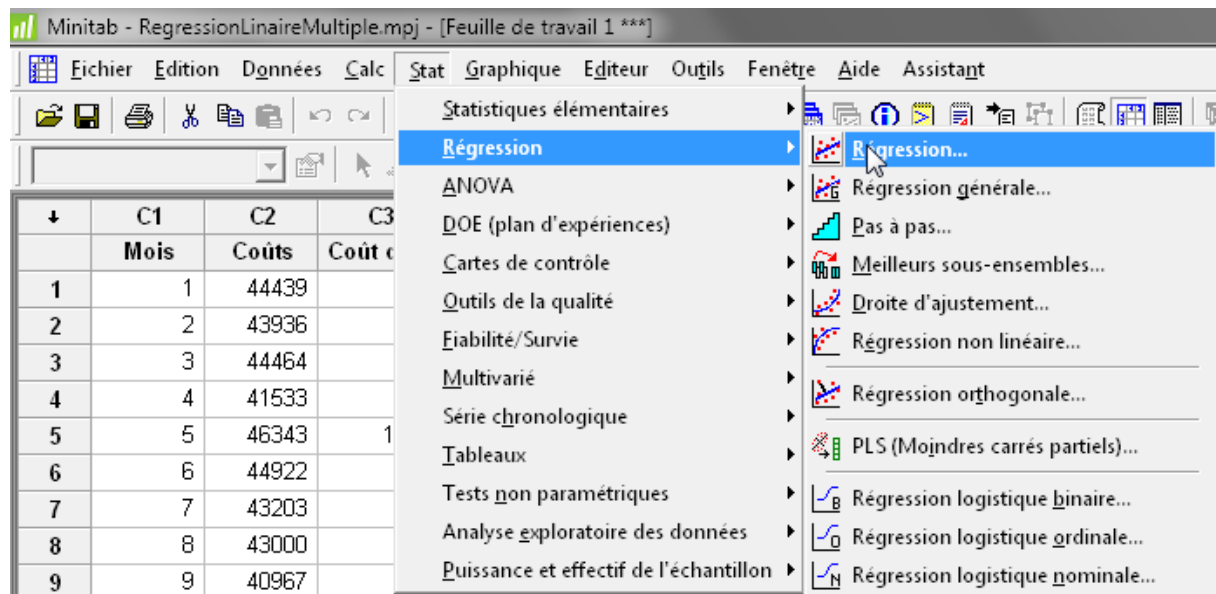
Minitab n'a malheureusement pas d'outil pour déterminer rapidement les corrélations partielles. Il va donc falloir jouer un peu comme dans Microsoft Excel.

Nous reprenons les données suivantes:

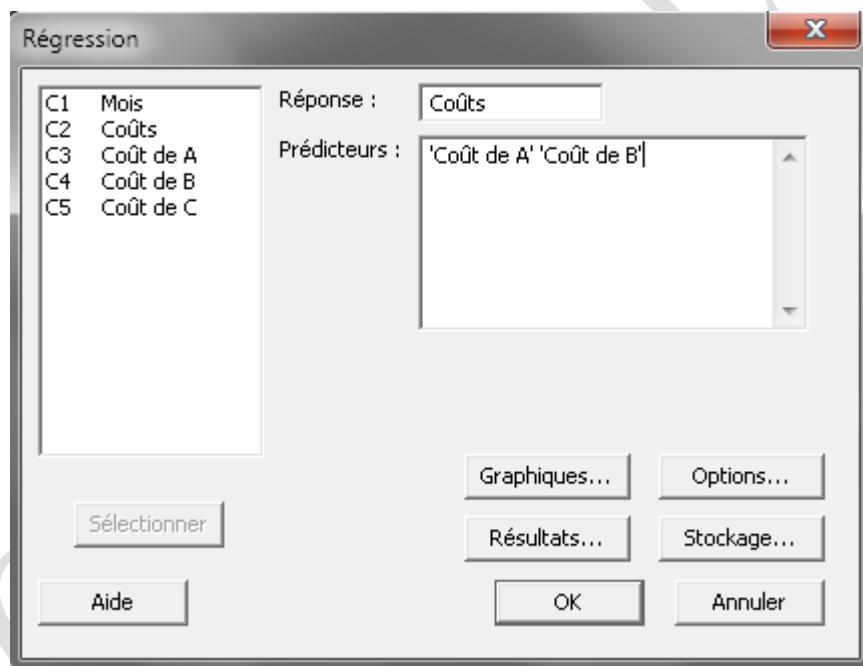


	C1	C2	C3	C4	C5
	Mois	Coûts	Coût de A	Coût de B	Coût de C
1	1	44439	515	541	928
2	2	43936	929	692	711
3	3	44464	800	710	824
4	4	41533	979	675	758
5	5	46343	1165	1147	635
6	6	44922	651	939	901
7	7	43203	847	755	580
8	8	43000	942	908	589
9	9	40967	630	738	682
10	10	48582	1113	1175	1050
11	11	45003	1086	1075	984
12	12	44303	843	640	828
13	13	42070	500	752	708
14	14	44353	813	989	804
15	15	45968	1190	823	904
16	16	47781	1200	1108	1120
17	17	43202	731	590	1065
18	18	44074	1089	607	1132
19	19	44610	786	513	839

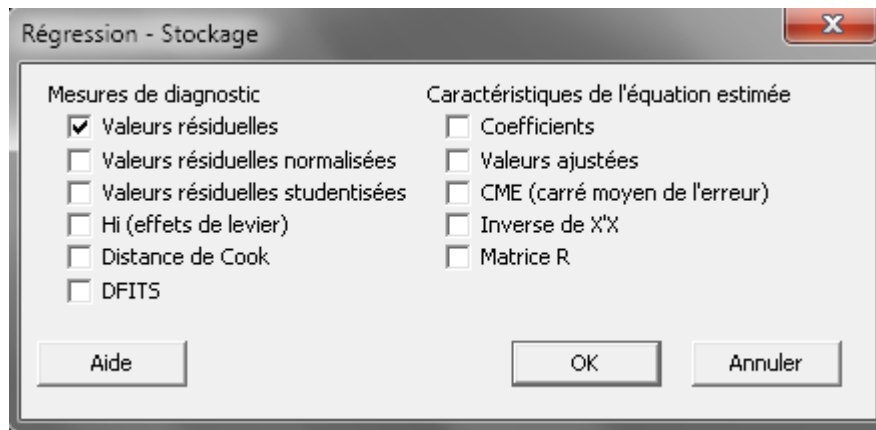
Nous allons dans le menu **Stat/Régression/Régression...**:



Pour avoir la corrélation partielle par rapport aux coûts de C nous prenons:



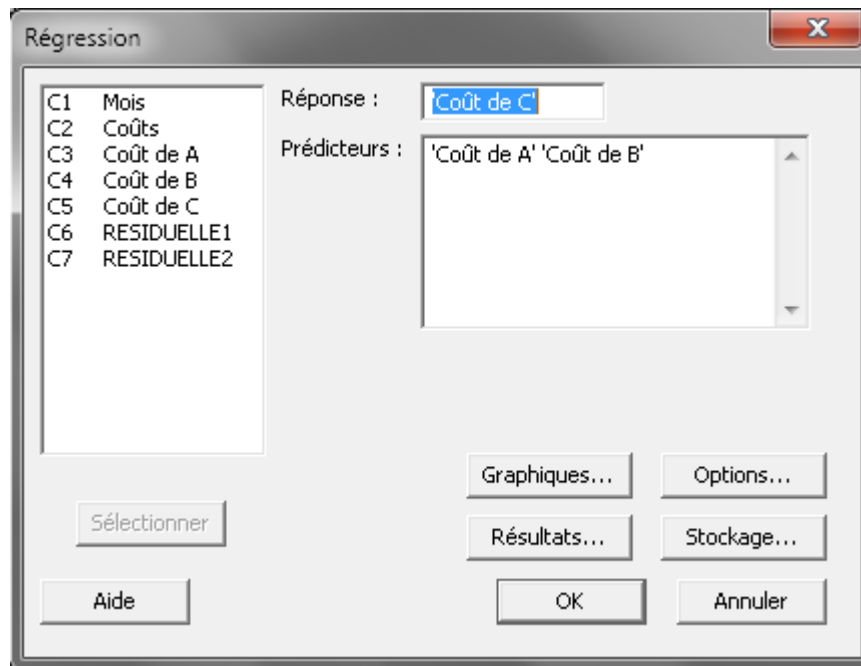
Et il nous faut stocker les valeurs résiduelles en cliquant sur le bouton **Stockage...**:



Nous validons deux fois par **OK**:

↓	C1	C2	C3	C4	C5	C6
	Mois	Coûts	Coût de A	Coût de B	Coût de C	RESIDUELLE1
1	1	44439	515	541	928	2302.33
2	2	43936	929	692	711	-133.87
3	3	44464	800	710	824	756.49
4	4	41533	979	675	758	-2640.34
5	5	46343	1165	1147	635	-190.07
6	6	44922	651	939	901	864.82
7	7	43203	847	755	580	-826.76
8	8	43000	942	908	589	-1910.04
9	9	40967	630	738	682	-2278.77
10	10	48582	1113	1175	1050	2118.46
11	11	45003	1086	1075	984	-1001.83
12	12	44303	843	640	828	710.84
13	13	42070	500	752	708	-795.32
14	14	44353	813	989	804	-427.12
15	15	45968	1190	823	904	547.27
16	16	47781	1200	1108	1120	1275.49
17	17	43202	731	590	1065	166.59
18	18	44074	1089	607	1132	-214.06
19	19	44610	786	513	839	1675.88

Ensuite nous relançons la régression mais en prenant:



Nous aurons alors:

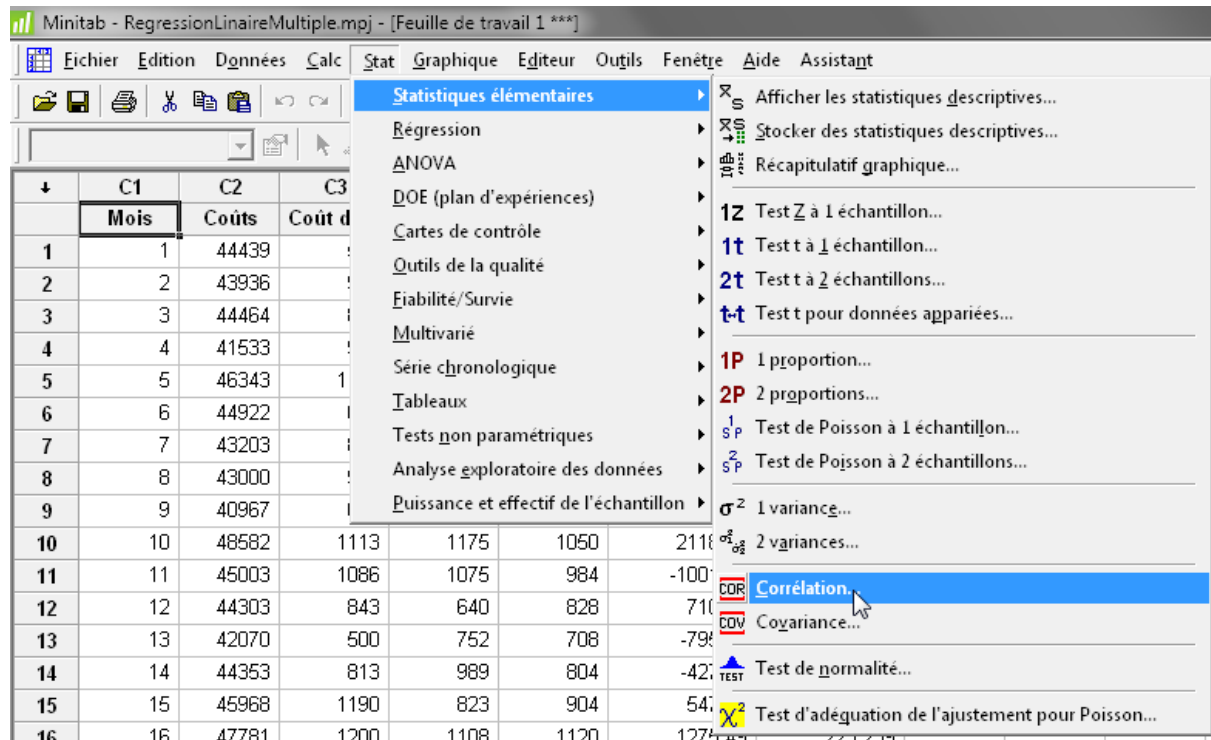
Minitab - RegressionLineaireMultiple.mpj - [Feuille de travail 1 \*\*\*]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assistant

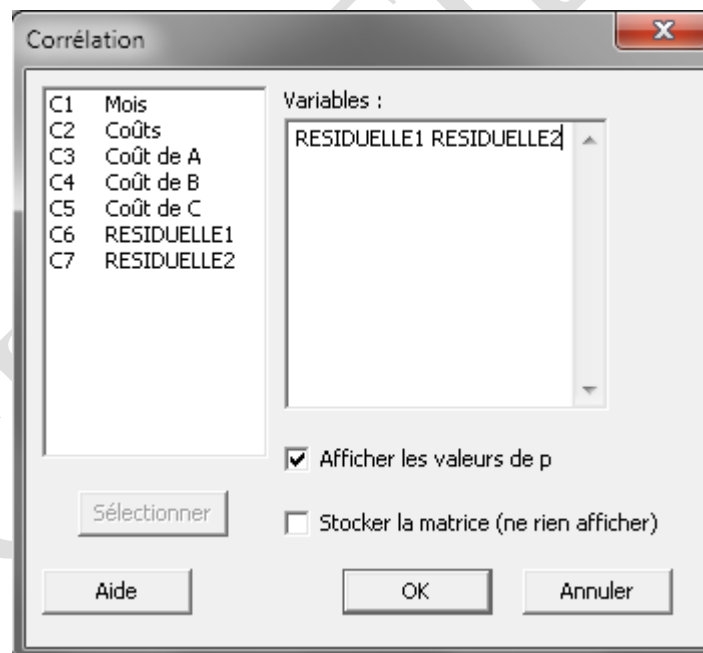
↓	C1	C2	C3	C4	C5	C6	C7
	Mois	Coûts	Coût de A	Coût de B	Coût de C	RESIDUELLE1	RESIDUELLE2
1	1	44439	515	541	928	2302.33	153.490
2	2	43936	929	692	711	-133.87	-156.871
3	3	44464	800	710	824	756.49	-8.171
4	4	41533	979	675	758	-2640.34	-124.727
5	5	46343	1165	1147	635	-190.07	-248.609
6	6	44922	651	939	901	864.82	131.219
7	7	43203	847	755	580	-826.76	-259.940
8	8	43000	942	908	589	-1910.04	-260.339
9	9	40967	630	738	682	-2278.77	-102.689
10	10	48582	1113	1175	1050	2118.46	182.890
11	11	45003	1086	1075	984	-1001.83	113.819
12	12	44303	843	640	828	710.84	-22.574
13	13	42070	500	752	708	-795.32	-41.133
14	14	44353	813	989	804	-427.12	-3.237
15	15	45968	1190	823	904	547.27	-19.092
16	16	47781	1200	1108	1120	1275.49	223.239
17	17	43202	731	590	1065	166.59	238.754
18	18	44074	1089	607	1132	-214.06	213.482
19	19	44610	786	513	839	1675.88	-9.511

Ensuite, nous allons dans le menu **Statistiques Élémentaires/Corrélation...**:





Et dans la boîte de dialogue qui apparaît, nous prenons les résidus par paires:



**Corrélations : RESIDUELLE1; RESIDUELLE2**

Corrélation de Pearson de RESIDUELLE1 et RESIDUELLE2 = 0.569  
 Valeur de p = 0.011

et ainsi de suite pour les autres combinaisons.

## 13.15. Exercice 110.: Effectuer une régression non linéaire

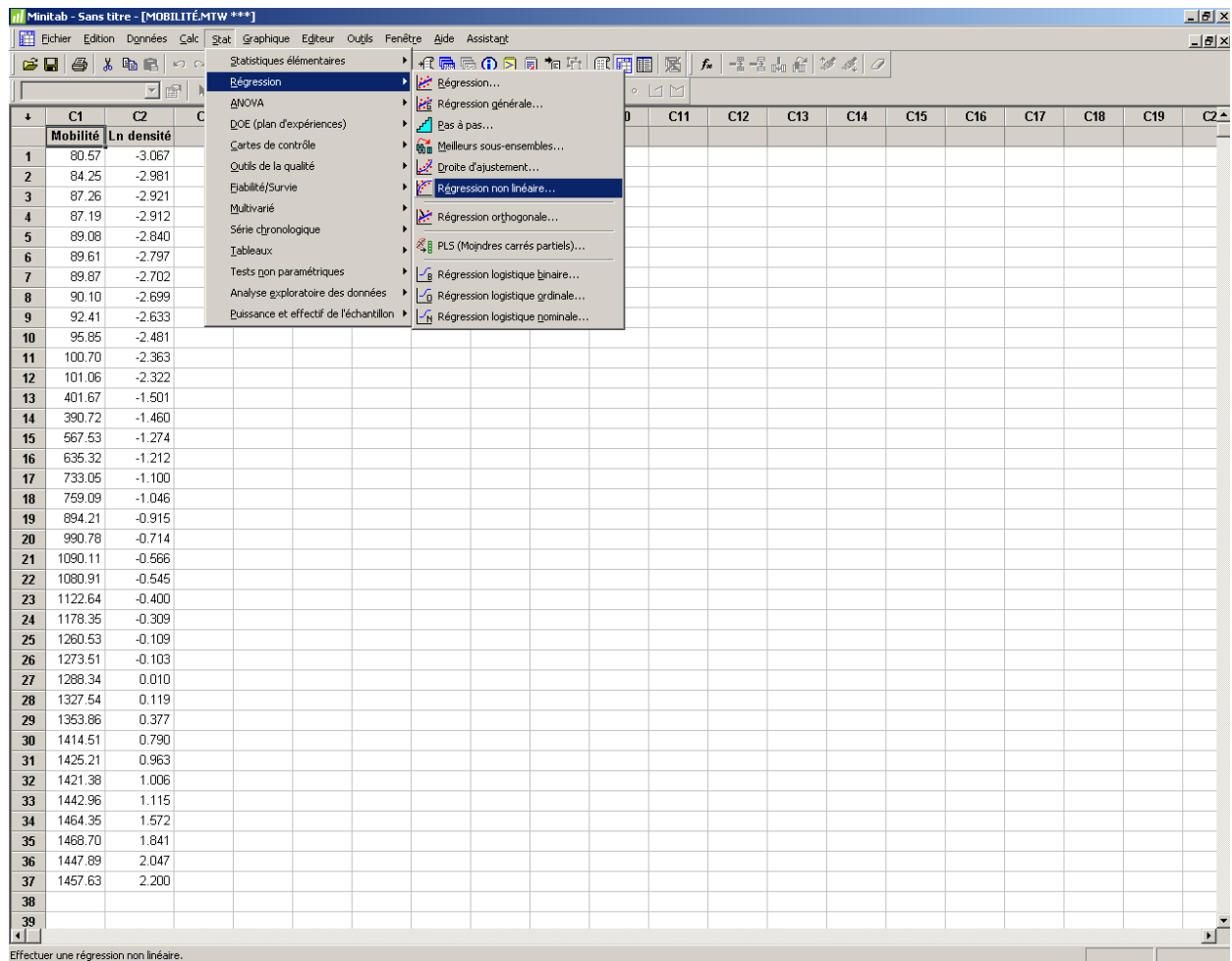
Minitab® Statistical Software 16

Nous voulons comprendre la relation entre la mobilité des électrons en semi-conducteurs et le logarithme népérien de la densité. Des recherches antérieures suggèrent qu'un modèle rationnel non linéaire (rapport de 2 valeurs polynomiales) à 7 paramètres fournit une adéquation de l'ajustement. Les méthodes abordées ici produisent les valeurs de début du modèle.

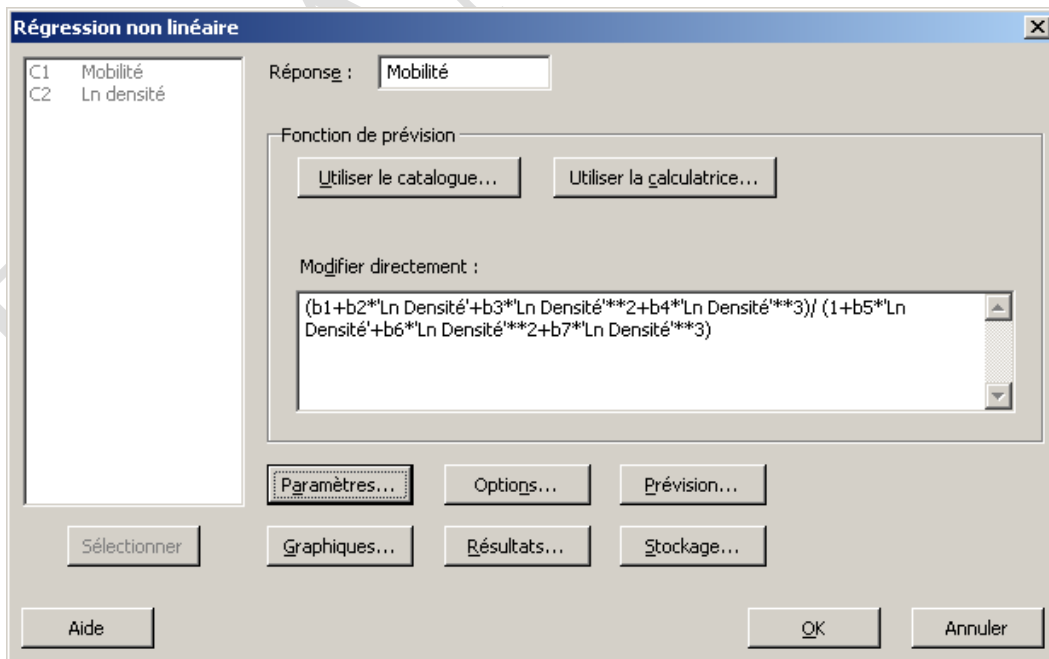
Ouvrez la feuille de travail *GaussNewton.mtw*.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20
	Mobilité	Ln densité																		
1	80.57	-3.067																		
2	84.25	-2.981																		
3	87.26	-2.921																		
4	87.19	-2.912																		
5	89.08	-2.840																		
6	89.61	-2.797																		
7	89.87	-2.702																		
8	90.10	-2.699																		
9	92.41	-2.633																		
10	95.85	-2.481																		
11	100.70	-2.363																		
12	101.06	-2.322																		
13	401.67	-1.501																		
14	390.72	-1.460																		
15	567.53	-1.274																		
16	635.32	-1.212																		
17	733.05	-1.100																		
18	759.09	-1.046																		
19	894.21	-0.915																		
20	990.78	-0.714																		
21	1090.11	-0.566																		
22	1060.91	-0.545																		
23	1122.64	-0.400																		
24	1178.35	-0.309																		
25	1260.53	-0.109																		
26	1273.51	-0.103																		
27	1288.34	0.010																		
28	1327.54	0.119																		
29	1353.86	0.377																		
30	1414.51	0.790																		
31	1425.21	0.963																		
32	1421.38	1.006																		
33	1442.96	1.115																		
34	1464.35	1.572																		
35	1468.70	1.841																		
36	1447.89	2.047																		
37	1457.63	2.200																		
38																				
39																				

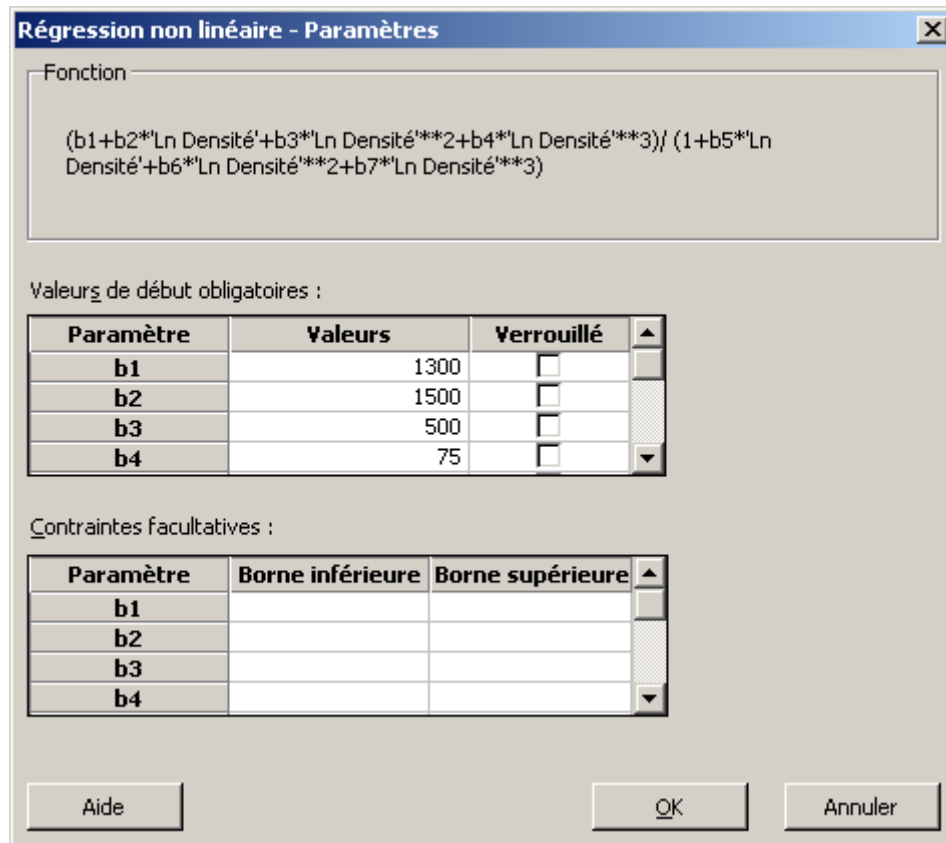
Nous allons dans le menu **Stat/Régression/Régression non linéaire...**:



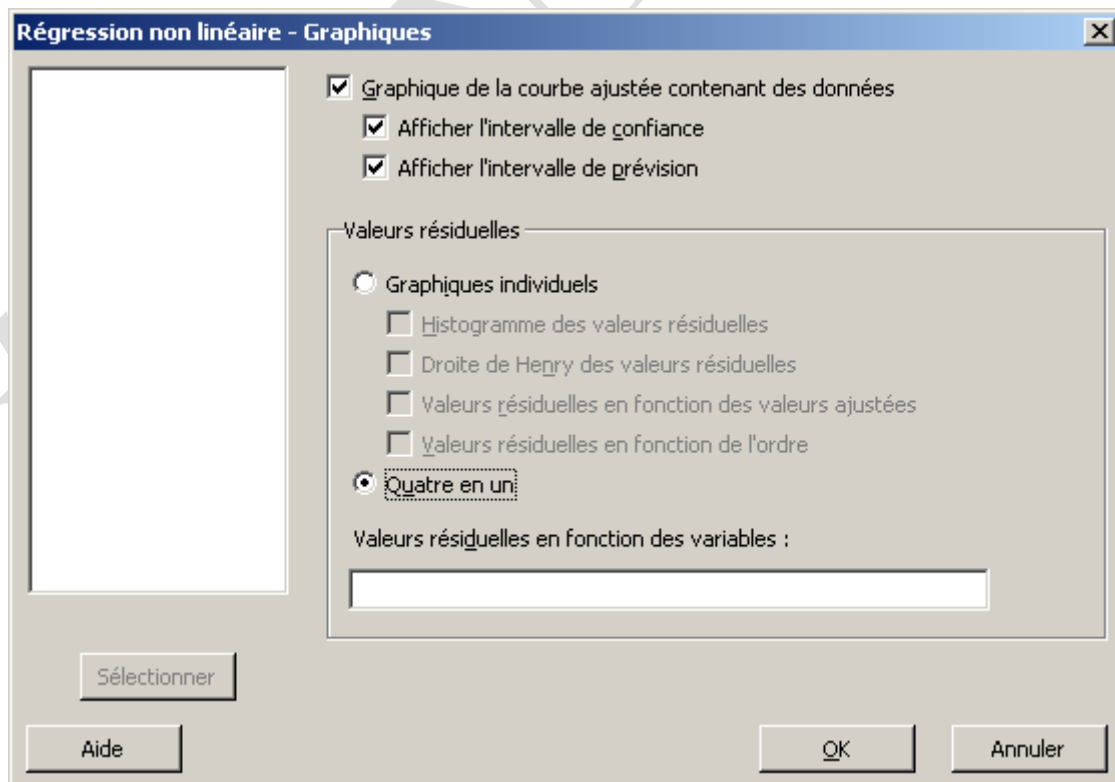
et dans la boîte de dialogue qui apparaît, nous mettons:



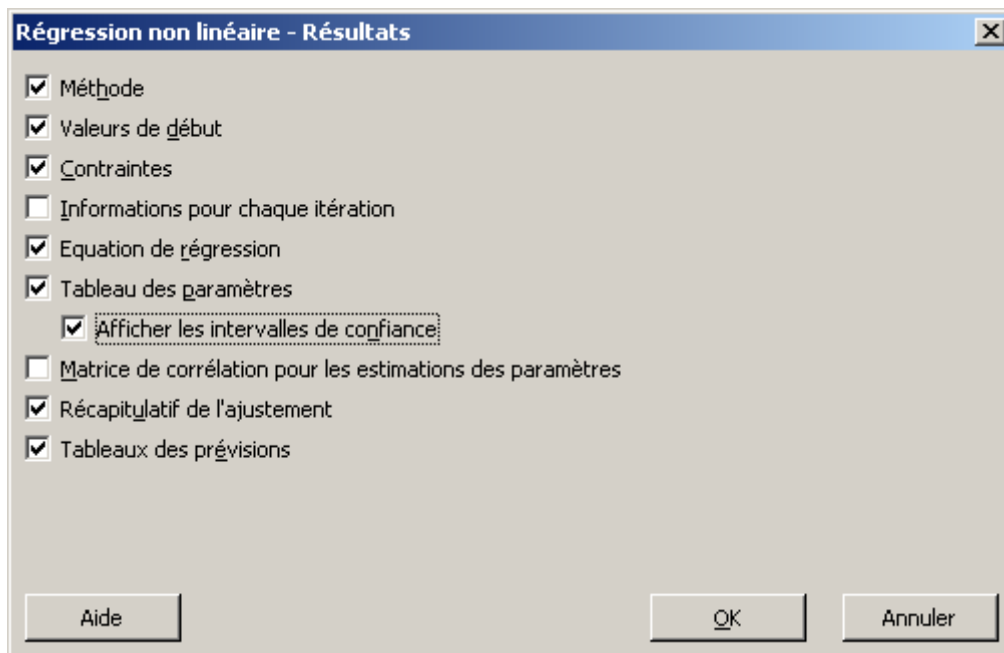
Nous cliquons ensuite sur le bouton **Paramètres** et sous **Valeurs de début obligatoires**, dans l'ordre de b1 à b7, nous entrons une valeur par cellule: 1300, 1500, 500, 75, 1, 0,4, 0,05.



Nous cliquons sur **OK** et ensuite sur le bouton **Graphique** de la boîte de dialogue principale et nous mettons les paramètres visibles ci-dessous:



Nous cliquons sur **OK** et ensuite sur le bouton **Résultats** (toujours de la boîte de dialogue principale) pour y activer les paramètres suivants:



Cliquez sur **OK** autant de fois que nécessaire pour valider le tout et nous obtenons dans la fenêtre de session:

Bienvenue dans Minitab, appuyez sur F1 pour obtenir de l'aide.

**Résultats pour : MOBILITE.MTW**

Regression non lineaire : Mobilité = (b1 + b2 \* Ln densité) + ...

Méthode  
 Algorithme Gauss-Newton  
 Itérations maximales 200  
 Tolérance 0.00001

Valeurs initiales des paramètres

Paramètre	Valeur
b1	1300
b2	1500
b3	500
b4	75
b5	1
b6	0.4
b7	0.05

Equation

$$\text{Mobilité} = (1288.14 + 1491.08 * \ln(\text{densité}) + 583.238 * \ln(\text{densité})^2 + 75.4167 * \ln(\text{densité})^3) / (1 + 0.966295 * \ln(\text{densité}) + 0.397973 * \ln(\text{densité})^2 + 0.0497273 * \ln(\text{densité})^3)$$

Estimations des paramètres

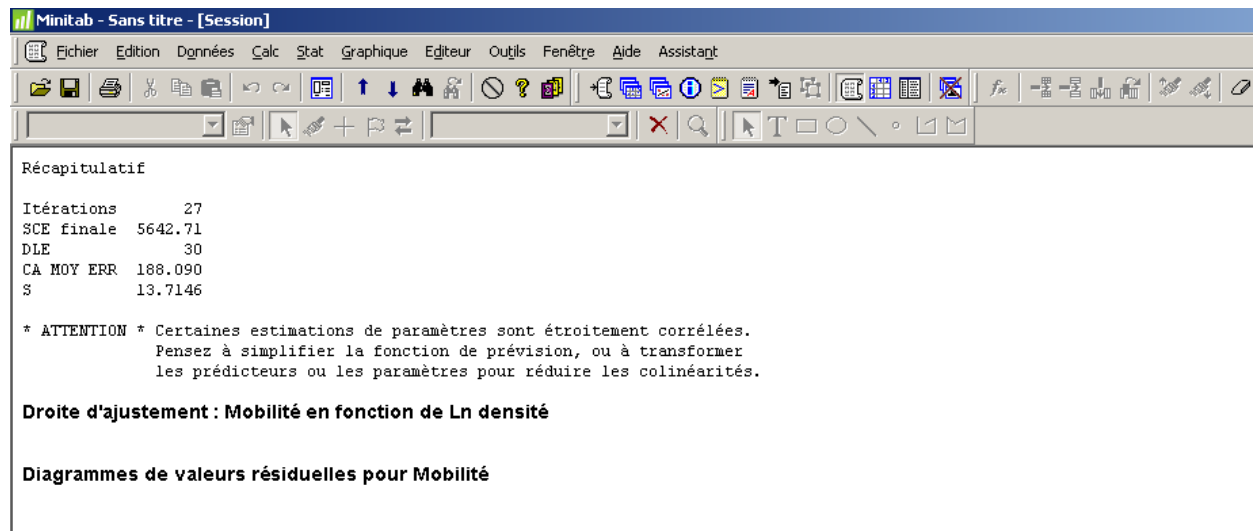
Paramètre	Estimation	SE	IC à 95 %
b1	1288.14	4.6648	(1278.59; 1297.71)
b2	1491.08	39.5711	(1381.50; 1548.27)
b3	583.24	28.6986	(502.36; 625.87)
b4	75.42	5.5675	(59.58; 83.57)
b5	0.97	0.0313	(0.88; *)
b6	0.40	0.0150	(0.36; *)
b7	0.05	0.0066	(0.03; 0.06)

Mobilité = (b1 + b2 \* Ln densité + b3 \* Ln densité^2 + b4 \* Ln densité^3) / (1 + b5 \* Ln densité + b6 \* Ln densité^2 + b7 \* Ln densité^3)

Inadéquation de l'ajustement

Aucune répétition.

Minitab ne peut pas effectuer le test d'inadéquation de l'ajustement en fonction de l'erreur pure.



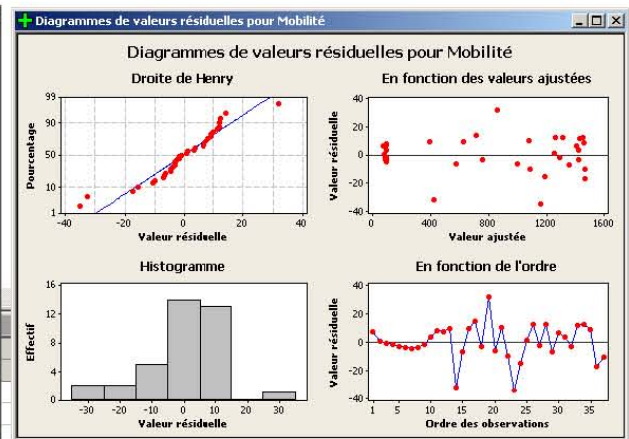
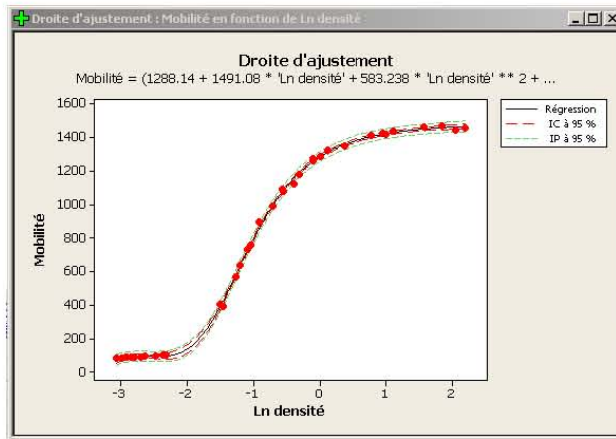
Minitab a donc obtenu les critères de convergence en 27 itérations en utilisant l'algorithme de Gauss-Newton avec la fonction de prévision et les valeurs de début de paramètres spécifiées. Toutefois, la convergence proprement dite ne garantit pas l'ajustement optimal du modèle, ni la minimisation de la somme des carrés de l'erreur (SCE). Il est possible d'obtenir une convergence avec des valeurs de paramètres incorrectes (SCE minimale locale, mauvaises valeurs de début ou fonction de prévision erronée, par exemple). Il est donc impératif d'examiner les valeurs des paramètres, la droite d'ajustement et les graphiques des valeurs résiduelles pour s'assurer que le modèle convient aux données et que l'algorithme converge sur le minimum global de la SCE.

Pour ce fichier de données, Minitab n'effectue pas de test d'inadéquation de l'ajustement car il n'y a aucune répétition.

Pour la régression non linéaire, Minitab ne calcule pas la valeur du coefficient de corrélation ou de  $p$  pour le modèle global car ces valeurs n'ont généralement aucune signification hors du contexte du modèle linéaire. Par conséquent, lorsque les chercheurs évaluent l'ajustement et comparent les modèles non linéaires concurrents, ils les choisissent souvent en fonction de leur connaissance du domaine concerné, de celui avec la plus petite SCE finale ou de la valeur  $S$  (Standard Error of Estimate) et du résultat graphique.

La valeur  $S$  est généralement plus facile à interpréter, à la fois en soi et par rapport aux autres valeurs, car  $S$  est exprimé dans les mêmes unités que la variable de réponse (mobilité des électrons)

Pour les graphiques, nous obtenons:





## 13.16. Exercice 111.: Régression logistique (logit) à variable catégorielle (qualitative) binaire (dixit: scoring)

Minitab® Statistical Software 15.1.1.0

Comme nous l'avons vu dans le cours de statistique théorique, le problème d'une régression linéaire simple (bivariée) ou multiple à variable expliquée de type binaire c'est que le modèle théorique peut prendre toute valeur dans  $\mathbb{R}$  ce qui est évidemment une aberration puisque dans le cas d'une variable expliquée de type 0/1, le modèle théorique ne doit jamais donner une valeur inférieure à 0 et supérieure 1. Or c'est malheureusement ce que fera une régression linéaire...

Pour cela nous avons démontré qu'il était possible d'utiliser une autre technique mathématique qui nous assurait que la condition susmentionnée soit satisfaite (régression binomiale!). Nous allons reprendre le même exemple mais au lieu de faire les calculs à la main, nous les faisons donc avec Minitab® Statistical Software.

Nous avons alors dans le fichier *RegressionLogistique.mpj* le tableau suivant des crédits procurés pas une banque avec le statut de bon débiteur (1) ou mauvais débiteur (0):

The screenshot shows the Minitab Statistical Software interface. The main window displays a data table with the following content:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Montant crédit	Status											
1	27200	0											
2	27200	0											
3	27200	0											
4	27200	0											
5	27200	0											
6	27200	0											
7	27200	0											

The session window shows the following text:

```

05.01.2011 06:57:27

Bienvenue dans Minitab, appuyez sur F1 pour obtenir de l'aide.
Récupération du projet du fichier 'C:\DOCUMENTS AND
SETTINGS\ADMINISTRATEUR\BUREAU\REGRESSIONLOGISTIQUE.MPJ'

```

Ensuite, nous allons dans le menu **Stat/Régression/Régression logistique binaire...**:

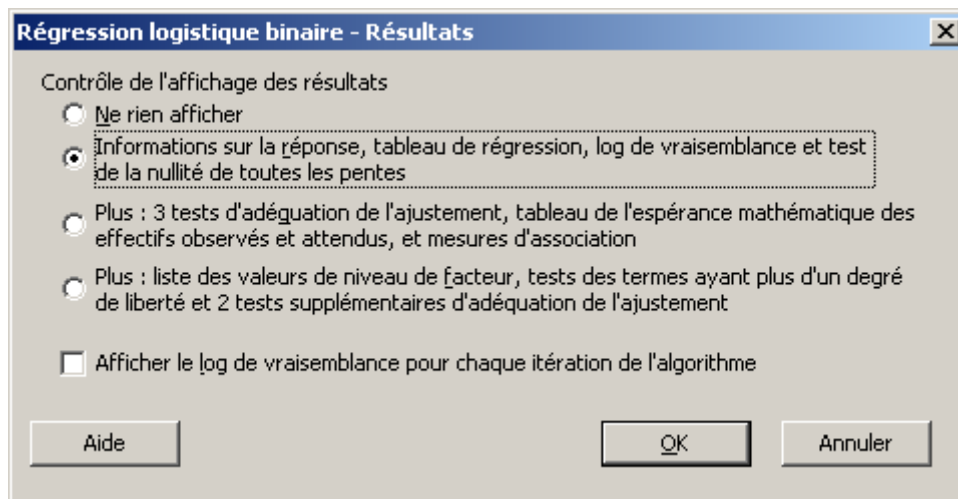
The screenshot shows the Minitab interface with the 'Stat' menu open and 'Régression' selected. The 'Régression logistique binaire...' option is highlighted. Below the menu, a spreadsheet window titled 'Feuille de travail 1 \*\*\*' is visible, containing data for 'Montant crédit' and 'Status'.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Montant crédit	Status											
1	27200	0											
2	27200	0											
3	27200	0											
4	27200	0											
5	27200	0											
6	27200	0											
7	27200	0											

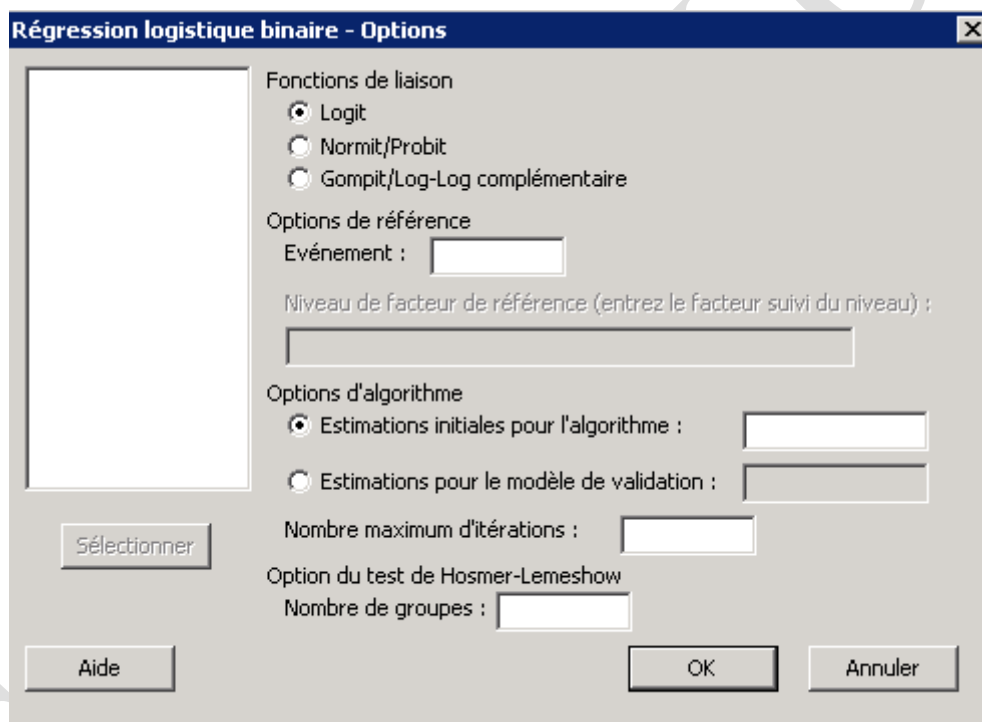
Nous y mettons:

The dialog box 'Régression logistique binaire' is shown. The 'Réponse en format réponse/effectif' radio button is selected. The 'Réponse' field contains 'Status'. The 'Modèle' field contains 'Montant crédit'. The 'Facteurs (facultatif)' field is empty. Buttons for 'Graphiques...', 'Options...', 'Résultats...', 'Stockage...', 'Aide', 'OK', and 'Annuler' are visible.

Dans le bouton **Résultats...** afin de faire comme dans le cours théorique, vous prenez:



et dans le bouton **Options** nous vérifions qu'il prenne bien le modèle **Logit** qui correspond à celui calculé à la main dans le cours théorique:



Malheureusement Minitab® Statistical Software ne propose pas d'afficher le graphique de la régression logistique donc nous validons par **OK** deux fois pour obtenir dans la fenêtre de session:

**Régression logistique binaire : Status en fonction de Montant crédit**

Fonction de liaison : Logit

Informations de réponse

Variable	Valeur	Dénombrement
Status	1	91 (Événement)
	0	45
Total		136

Tableau de régression logistique

Prédicteur	Coeff	Coef ErT	Z	P	Rapport des probabilités de succès	
					IC à 95 % Inférieur	IC à 95 % Supérieur
Constante	-61.3183	12.0224	-5.10	0.000		
Montant crédit	0.0022110	0.0004309	5.13	0.000	1.00	1.00

Prédicteur Supérieur

Constante	
Montant crédit	1.00

Feuille de travail en cours : Feuille de travail 1

Les  $p$ -value étant inférieures à 0, nous pouvons admettre que les coefficients et la constantes sont bien non nuls.

Cela signifie que la probabilité cumulée de valoir 1 est alors donnée par (avec  $x$  en millier de francs):

$$P(X \leq x) = \frac{1}{1 + e^{-(-61.3183 + 0.002211x)}}$$

résultat à comparer avec les calculs faits à la main dans le cours de statistique théorique:

$$P(X \leq x) = \frac{1}{1 + e^{-(-51.1116 + 0.0018371x)}}$$

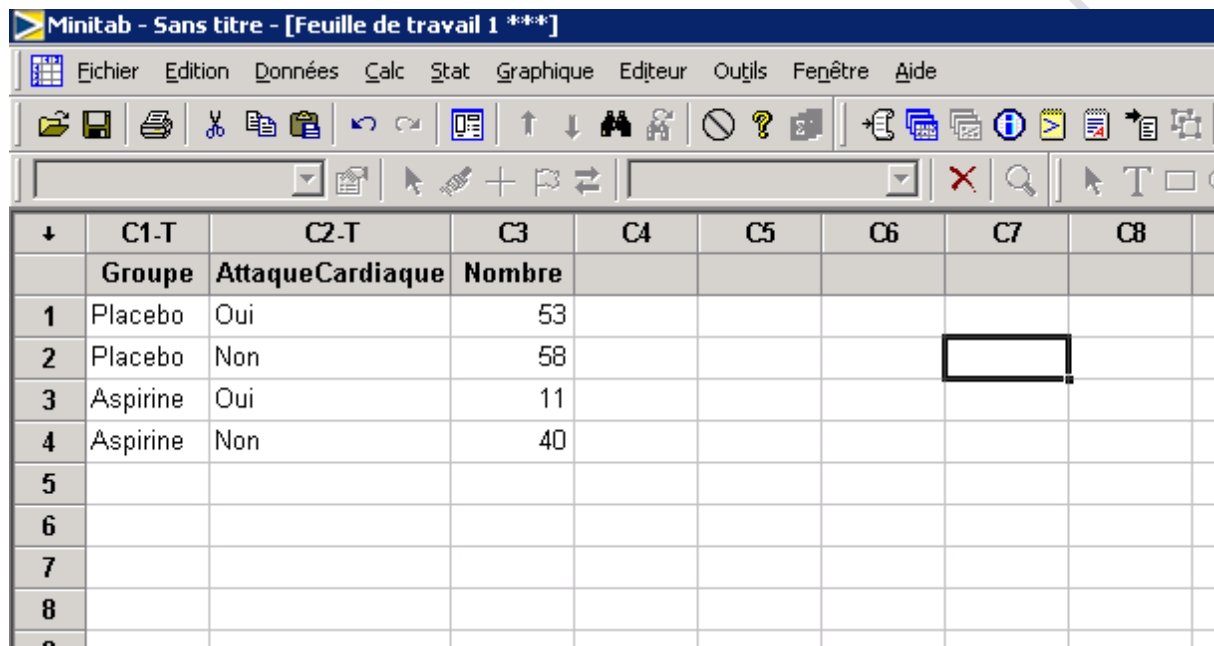
Donc il y a eu une relativement grosse différence au niveau des paramètres ce qui est normal puisqu'en dans le cours théorique nous avons utilisé la méthode des moindres carrés alors que les logiciels statistiques utilisent la "log-vraisemblance" empirique (log-likelihood).

## 13.17. Exercice 112.: Odds Ratio et intervalle de confiance

Minitab® Statistical Software 15.1.1.0

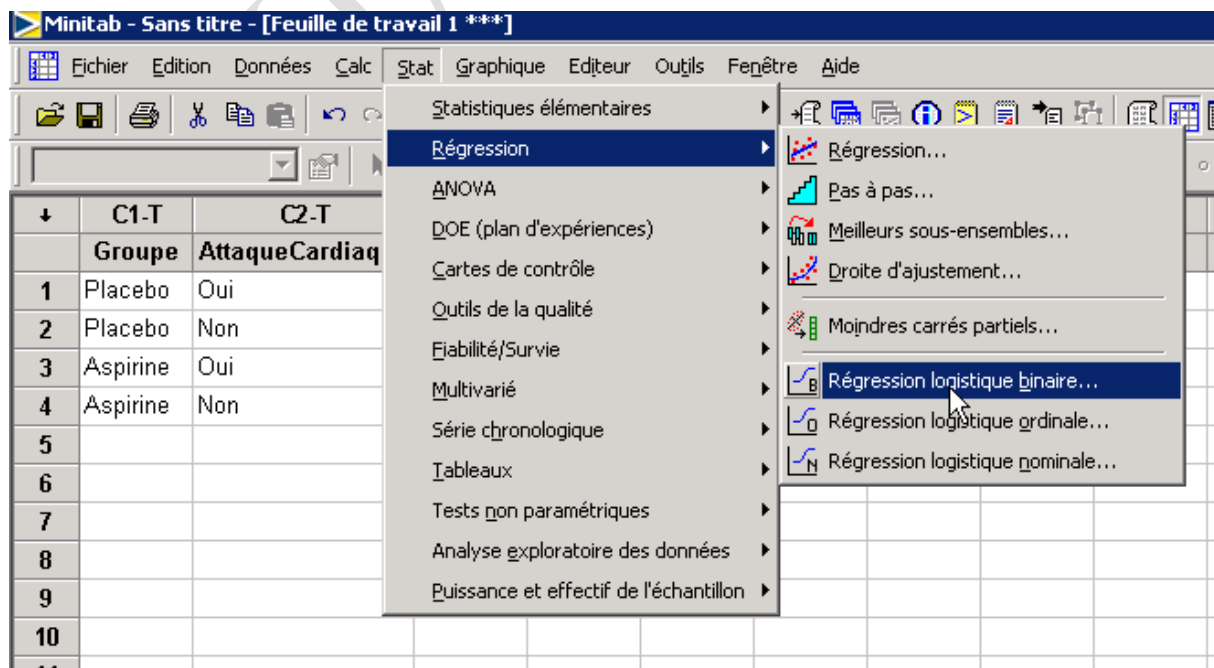
Le but va être ici de vérifier que nous retrouvons le même intervalle de confiance que l'odds ratio calculé à la main durant la cours théorique statistique suite à la démonstration mathématique de la construction de l'intervalle de confiance.

Nous partons donc des données suivantes:



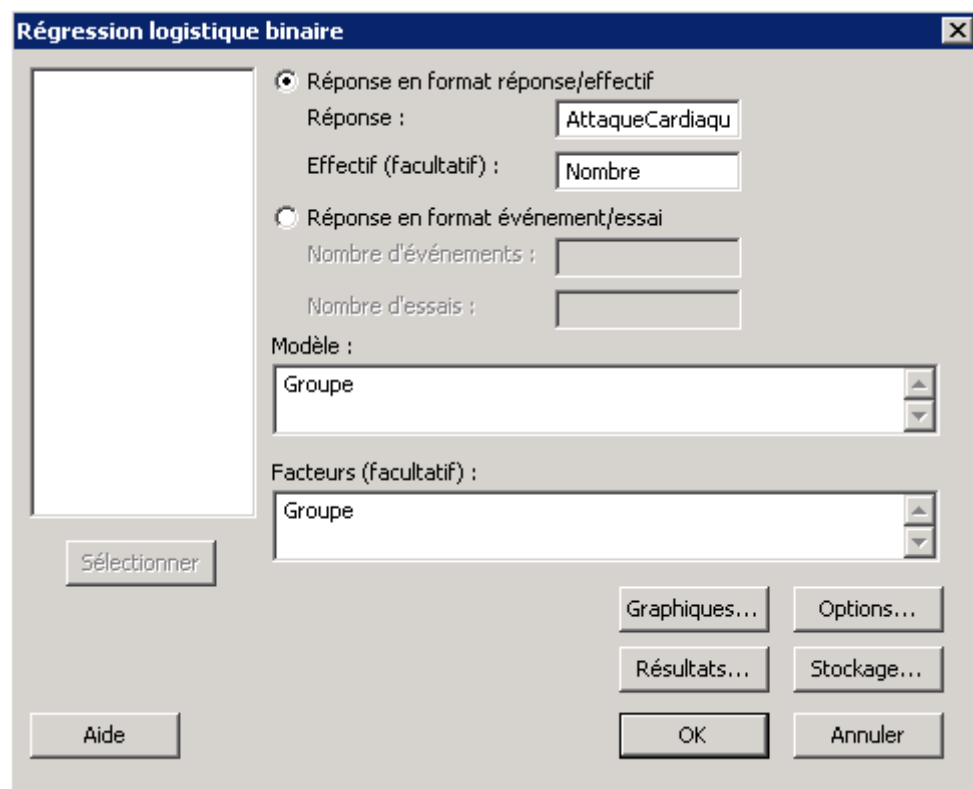
	C1-T	C2-T	C3	C4	C5	C6	C7	C8
	Groupe	AttaqueCardiaque	Nombre					
1	Placebo	Oui	53					
2	Placebo	Non	58					
3	Aspirine	Oui	11					
4	Aspirine	Non	40					
5								
6								
7								
8								
9								

et nous allons dans le menu **Stat/Régression/Régression logistique binaire...**:

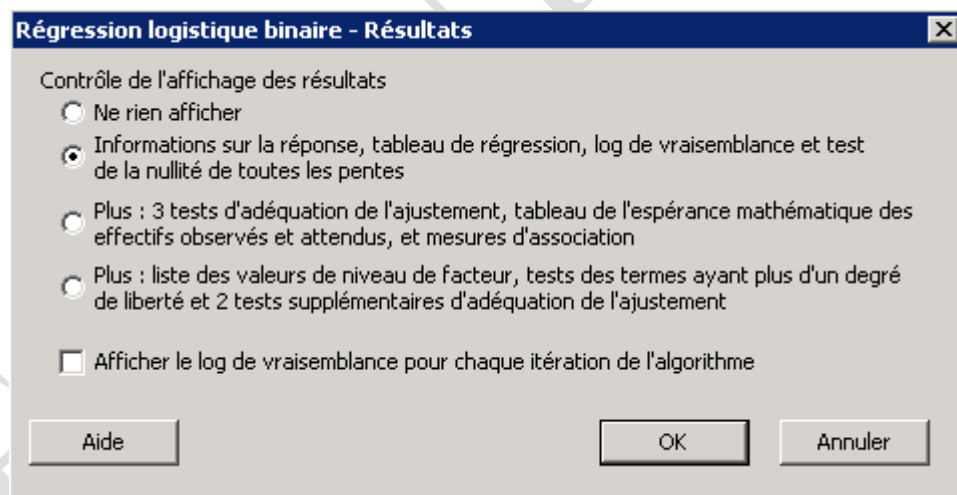


	C1-T	C2-T
	Groupe	AttaqueCardiaq
1	Placebo	Oui
2	Placebo	Non
3	Aspirine	Oui
4	Aspirine	Non
5		
6		
7		
8		
9		
10		
11		

Nous prenons:



Dans le bouton **Résultats...** nous prenons:



Nous validons deux fois par **OK** pour obtenir:

**Minitab - Sans titre - [Session]**

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

Régression logistique binaire : AttaqueCardiaque en fonction de Groupe

Fonction de liaison : Logit

Informations de réponse

Variable	Valeur	Dénombrement
AttaqueCardiaque	Oui	64 (Evénement)
	Non	98
	Total	162

Effectif : Nombre

Tableau de régression logistique

Prédicteur	Coeff	Coef ErT	Z	P	Rapport des	IC à 95 %	
					probabilités de succès	Inférieur	Supérieur
Constante	-1.29098	0.340454	-3.79	0.000			
Groupe							
Placebo	1.20083	0.389895	3.08	0.002	3.32	1.55	7.14

Nous retrouvons donc bien l'Odds Ratio de 3.32 tel que calculé dans le cours théorique et l'intervalle de confiance de 1.55 et 7.14 au deuxième chiffre après la virgule près!!!

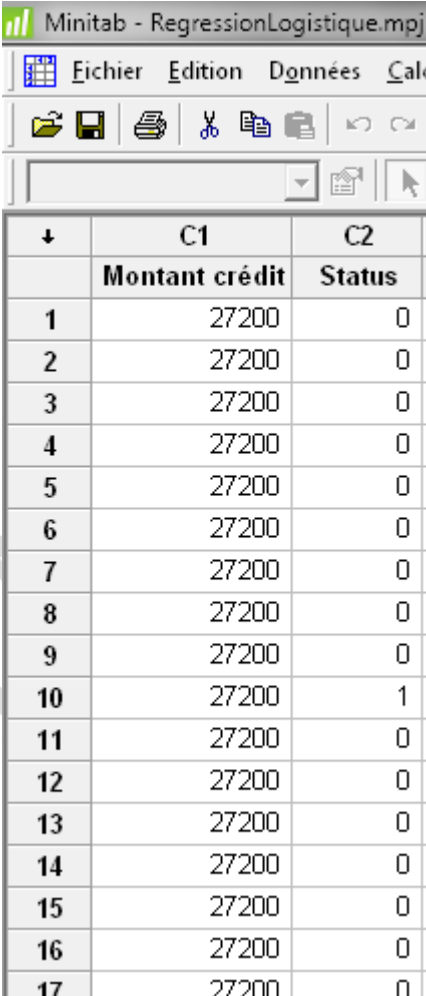
On regrettera cependant la présentation des informations qui n'est pas des meilleures...

### 13.18. Exercice 113.: Courbe ROC pour régression logistique (logit) à variable catégorielle (qualitative) binaire

Minitab® Statistical Software 15.1.1.0

Le but ici va être de vérifier si nous obtenons bien la même courbe ROC (Receiver Operating Characteristic) que celle calculée dans le cours théorique à la main lors de l'étude théorique du concept.

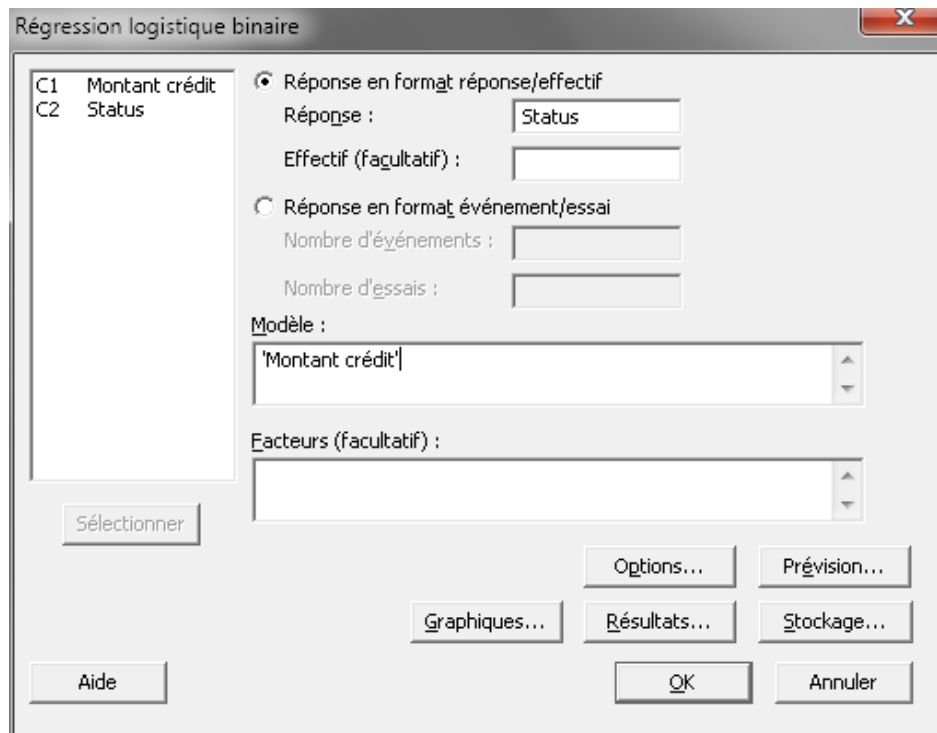
Nous allons reprendre les données des 136 lignes de l'exemple précédent mais avec les valeurs binaires numérisées telles que:



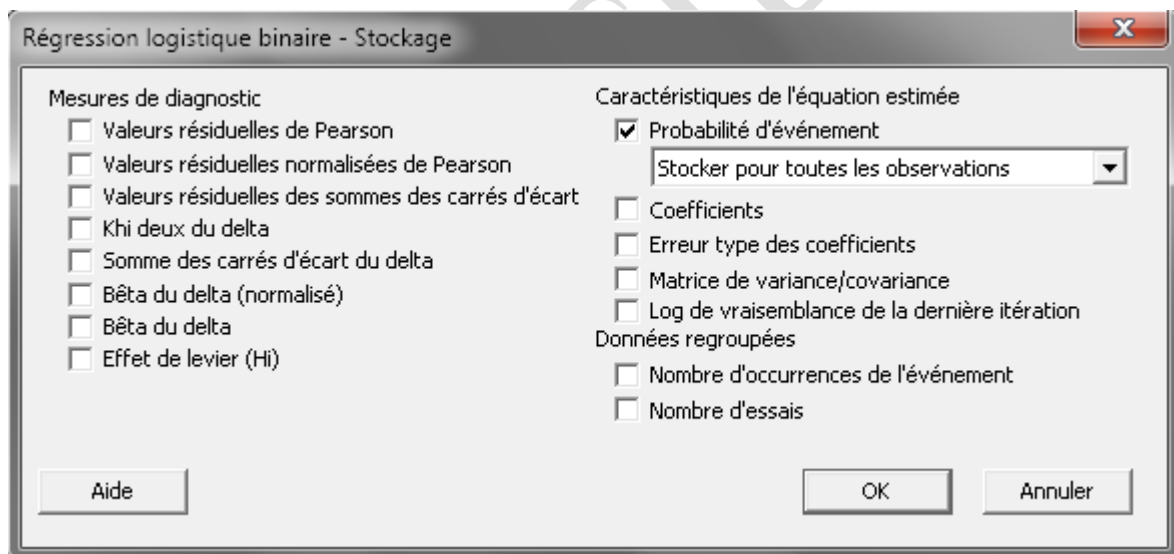
↓	C1	C2
	Montant crédit	Status
1	27200	0
2	27200	0
3	27200	0
4	27200	0
5	27200	0
6	27200	0
7	27200	0
8	27200	0
9	27200	0
10	27200	1
11	27200	0
12	27200	0
13	27200	0
14	27200	0
15	27200	0
16	27200	0
17	27200	0

et nous lançons la régression logistique exactement en passant au même endroit que dans l'exercice précédent et avec les mêmes paramètres:

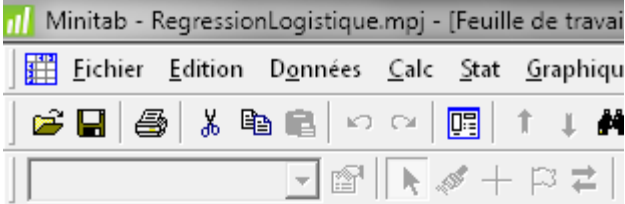




à la différence que nous allons cliquer sur Stockage... pour cocher **Probabilité d'événement**:



Nous validons par **OK** pour obtenir (bon pour les autres lignes vous référer à votre écran):



Minitab - RegressionLogistique.mpj - [Feuille de travail]

Échier Édition Données Calc Stat Graphique

↓	C1	C2	C3
	Montant crédit	Status	PROB EVEN1
1	27200	0	0.235364
2	27200	0	0.235364
3	27200	0	0.235364
4	27200	0	0.235364
5	27200	0	0.235364
6	27200	0	0.235364
7	27200	0	0.235364
8	27200	0	0.235364
9	27200	0	0.235364
10	27200	1	0.235364
11	27200	0	0.235364
12	27200	0	0.235364
13	27200	0	0.235364
14	27200	0	0.235364
15	27200	0	0.235364
16	27200	0	0.235364
17	27200	0	0.235364
18	27200	0	0.235364

Nous obtenons donc les mêmes probabilités d'événements que celles calculées trivialement dans le cours théorique.

Et ensuite?? Eh bien avec les versions 16 et antérieures de Minitab il va falloir aller sur la page Internet suivante:

<http://www.minitab.com/en-US/support/macros/default.aspx?action=code&id=141>

**Minitab** Software for Quality Improvement. United States - English

Products Training Support Academic Store Company Theater Search...

Home > Support > Macros > Macros > ROC (Receiver Operating Characteristic) Curve Code

**Support**

- Knowledgebase / FAQ
- Downloads
- Documentation
- Licensing
- Macros**
  - Download All Macros
  - ANOVA & Regression
  - Analysis of Count Data
  - Data Manipulation
  - Education
  - General Statistics
  - Graphs
  - Math
  - Miscellaneous
  - Nonparametrics
  - Quality Control and Designed Experiments
  - Time Series
  - Other Macro Resources
- Online Support

## ROC (Receiver Operating Characteristic) Curve Code

► How to save this page as a macro

```
macro
#####
#
# Name:      ROCBLR.MAC
# Version:   Release 15 and 16
# Written by: Mindy Tomlinson
#
# This macro performs three functions as a follow up to a binary logistic
# (BLR) regression analysis:
#
# 1. Generates a classification table
# 2. Generates an ROC (Receiver Operating Characteristic) curve and
#    calculates the area under the curve
# 3. Given the event probabilities, stores an event probability for each
#    row, not just the first instance of a unique set of predictor values
#
# Before running this macro,
# 1. Ensure that the response data is in response/frequency format, which
#    is the default format.
# 2. Perform the binary logistic regression analysis AND store the event
#    probabilities in the worksheet.
#
#
```

Vous copiez le code dans un fichier nommé ROCBLR.MAC et suivez les étapes de de la page 1216. Ensuite, vous allez donc dans le menu **Édition/Editeur de ligne de commande**:

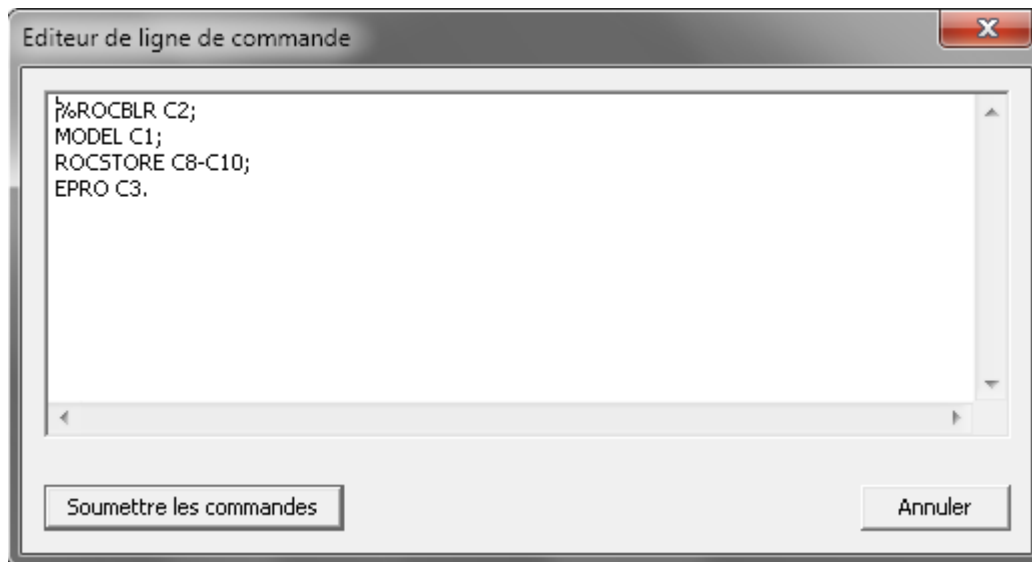
Minitab - RegressionLogistique.mpj - [Feuille de travail 1 \*\*\*]

Fichier Édition Données Calc Stat Graphique Editeur

- Annuler (Ctrl+Z)
- Répéter (Ctrl+Y)
- Effacer les cellules (Suppr)
- Supprimer les cellules
- Copier les cellules (Ctrl+C)
- Couper les cellules (Ctrl+X)
- Coller les cellules (Ctrl+V)
- Coller la liaison
- Liaisons de feuille de travail
- Sélectionner toutes les cellules (Ctrl+A)
- Modifier la dernière boîte de dialogue (Ctrl+E)
- Editeur de ligne de commande (Ctrl+L)**

1				
2				
3				
4				
5				
6				
7				
8				
9				
10	27200	1	0.235364	
11	27200	0	0.235364	

et nous mettons conformément à la documentation de la macro:



Nous obtenons alors dans la fenêtre de session:

Exécution depuis le fichier : C:\Program Files (x86)\Minitab\Minitab 16\français\Macros\ROCBLR.MAC

-----  
 Status is classified as '1' when the event probability  
 for an observation is greater than or equal to 0.5.  
 -----

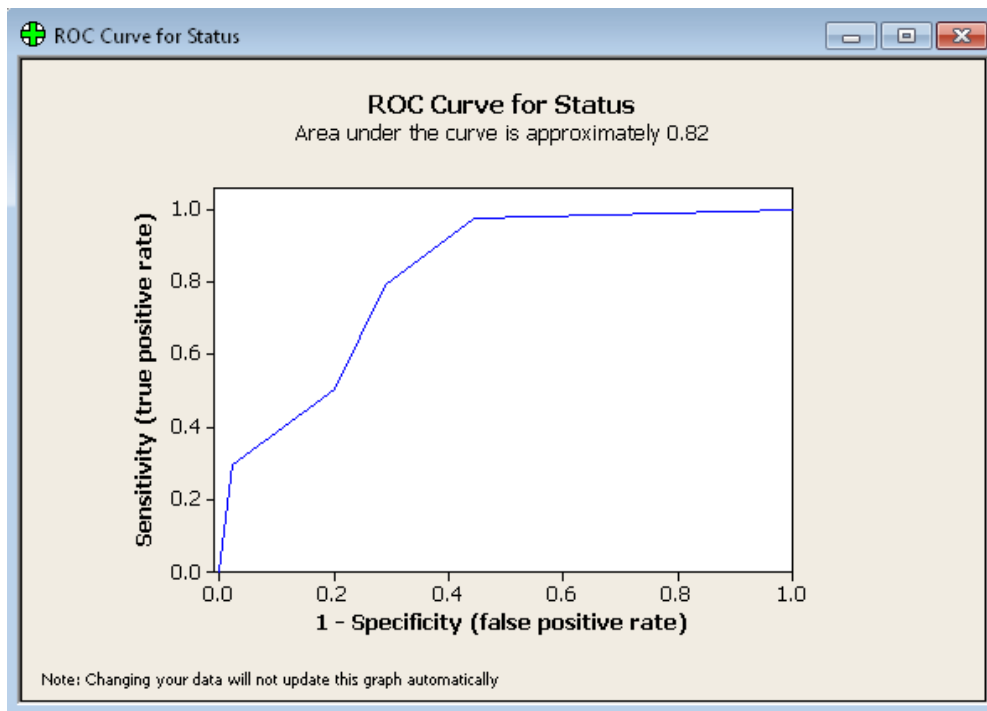
Rows: Observed Status  
 Columns: Number of observations that are correctly or incorrectly classified  
 -----

	Classification		
	Correct	Incorrect	Total
1	72	19	91
0	32	13	45
Total	104	32	136

-----  
 ROC Curve:

Sensitivity = P( Classified Status = 1 | Observed Status = 1 )  
 Specificity = P( Classified Status = 0 | Observed Status = 0 )

et comme courbe ROC:



et dans la feuille de travail (du moins pour les quelques premières lignes visibles) où l'on voit que Minitab traite le calcul dans un ordre très particulier dans sa Macro...:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
	Montant crédit	Status	PROB EVEN1					p	Sensitivity	Specificity
1	27200	0	0.235364					1.00	0.00000	1.00000
2	27200	0	0.235364					0.82	0.29670	0.97778
3	27200	0	0.235364					0.83	0.29670	0.97778
4	27200	0	0.235364					0.84	0.29670	0.97778
5	27200	0	0.235364					0.85	0.29670	0.97778
6	27200	0	0.235364					0.86	0.29670	0.97778
7	27200	0	0.235364					0.87	0.29670	0.97778
8	27200	0	0.235364					0.88	0.29670	0.97778
9	27200	0	0.235364					0.89	0.29670	0.97778
10	27200	1	0.235364					0.90	0.29670	0.97778
11	27200	0	0.235364					0.91	0.29670	0.97778
12	27200	0	0.235364					0.92	0.29670	0.97778
13	27200	0	0.235364					0.93	0.29670	0.97778
14	27200	0	0.235364					0.94	0.29670	0.97778
15	27200	0	0.235364					0.95	0.29670	0.97778
16	27200	0	0.235364					0.96	0.29670	0.97778
17	27200	0	0.235364					0.97	0.29670	0.97778
18	27200	0	0.235364					0.98	0.29670	0.97778
19	27200	0	0.235364					0.99	0.29670	0.97778
20	27200	0	0.235364					0.78	0.50549	0.80000
21	27200	0	0.235364					0.79	0.50549	0.80000
22	27200	0	0.235364					0.80	0.50549	0.80000
23	27200	0	0.235364					0.81	0.50549	0.80000
24	27200	0	0.235364					0.49	0.79121	0.71111
25	27200	0	0.235364					0.50	0.79121	0.71111

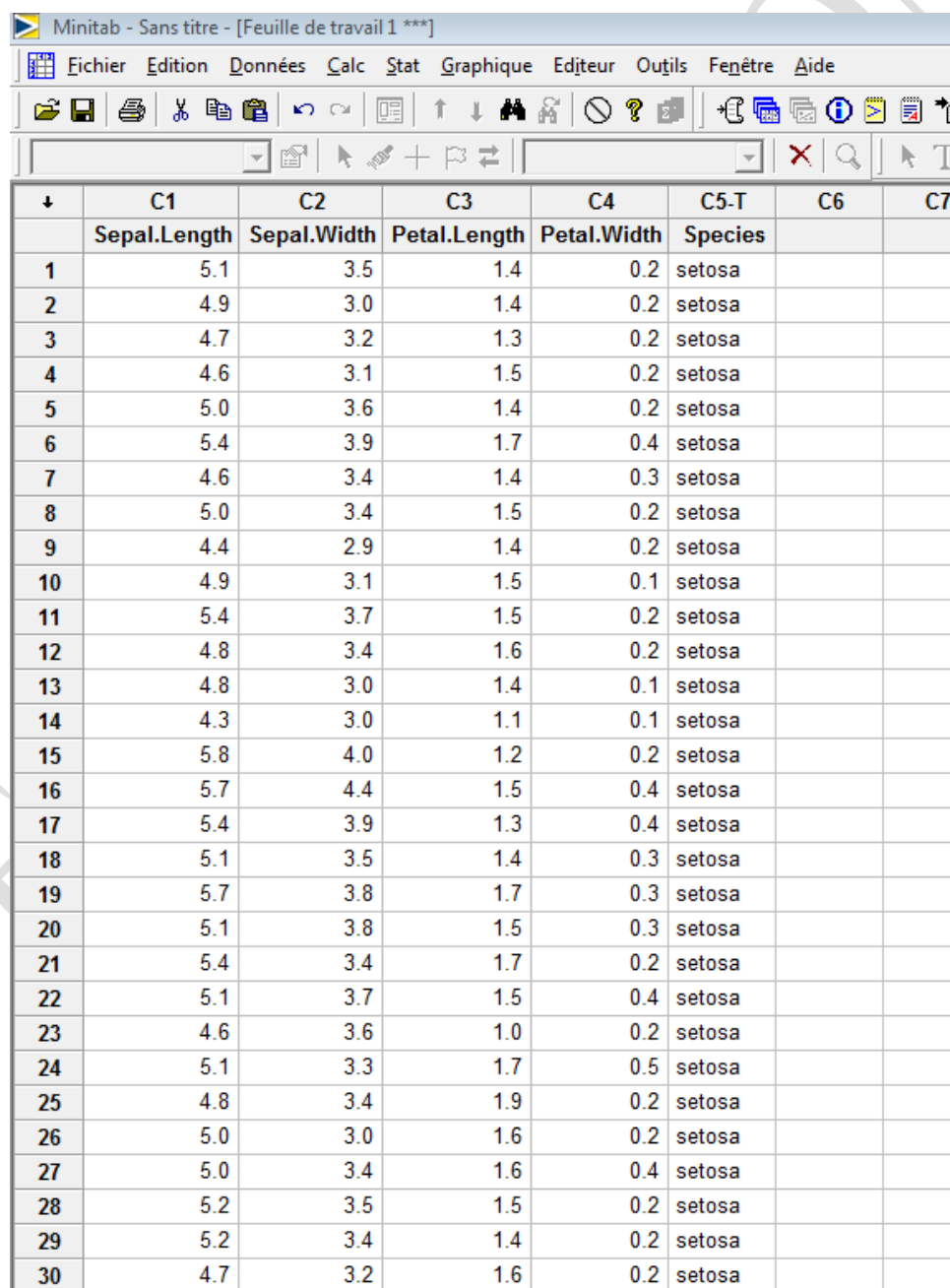
## 13.19. Exercice 114.: Régression logistique multinomiale

Minitab® Statistical Software 15.1.1.0

Nous avons démontré (non sans mal...) dans le cours théorique que la régression logistique multinomiale était une généralisation de la régression logistique binomiale.

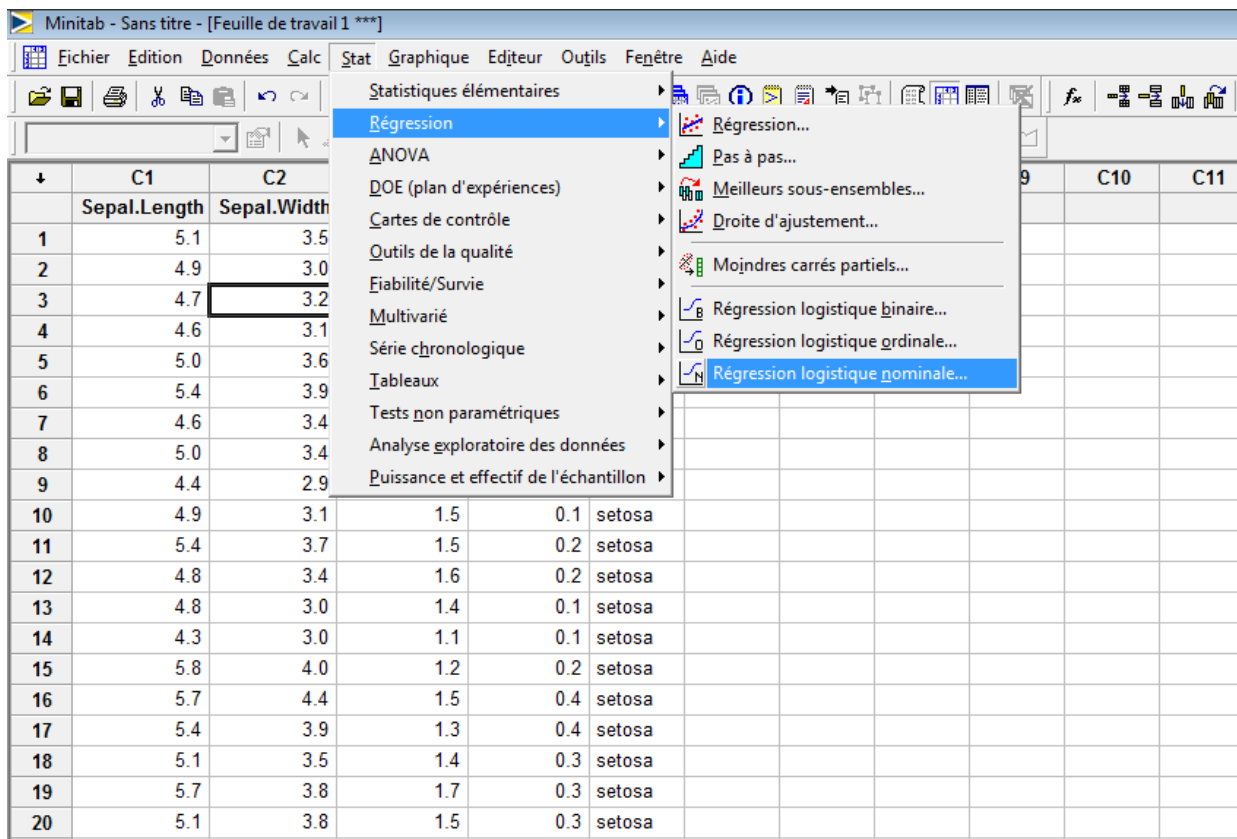
Cependant au vu de la lourdeur des calculs, nous nous sommes abstenus de faire tout calcul à la main ou même dans Microsoft Excel. Le but ici va donc être de vérifier si nous obtenons au moins les mêmes résultats qu'avec Tanagra ou R.

Pour cela nous partons du même jeu de données Iris de Fisher du fichier *Iris.mpj*:

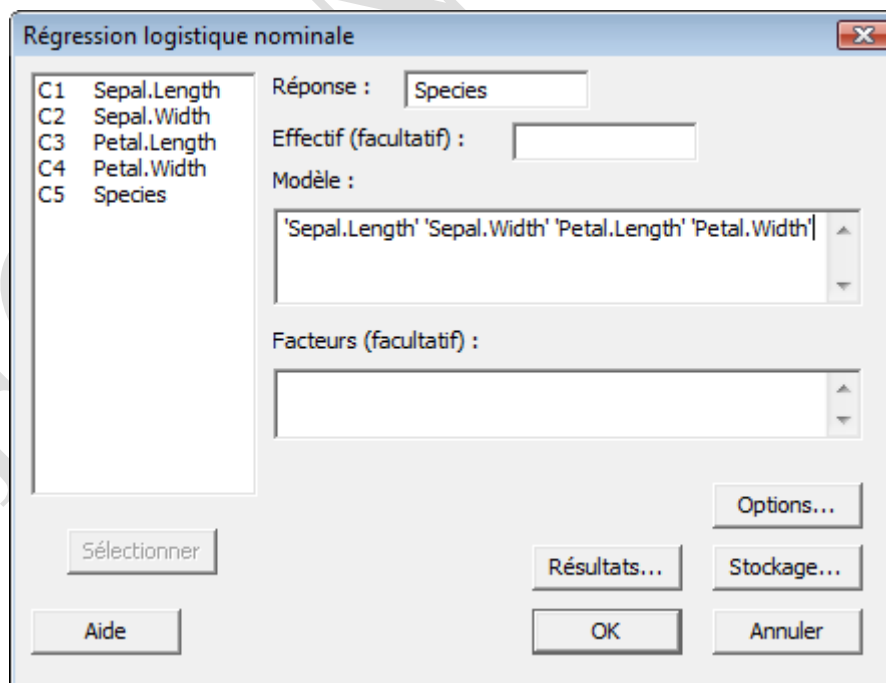


↓	C1	C2	C3	C4	C5-T	C6	C7
	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species		
1	5.1	3.5	1.4	0.2	setosa		
2	4.9	3.0	1.4	0.2	setosa		
3	4.7	3.2	1.3	0.2	setosa		
4	4.6	3.1	1.5	0.2	setosa		
5	5.0	3.6	1.4	0.2	setosa		
6	5.4	3.9	1.7	0.4	setosa		
7	4.6	3.4	1.4	0.3	setosa		
8	5.0	3.4	1.5	0.2	setosa		
9	4.4	2.9	1.4	0.2	setosa		
10	4.9	3.1	1.5	0.1	setosa		
11	5.4	3.7	1.5	0.2	setosa		
12	4.8	3.4	1.6	0.2	setosa		
13	4.8	3.0	1.4	0.1	setosa		
14	4.3	3.0	1.1	0.1	setosa		
15	5.8	4.0	1.2	0.2	setosa		
16	5.7	4.4	1.5	0.4	setosa		
17	5.4	3.9	1.3	0.4	setosa		
18	5.1	3.5	1.4	0.3	setosa		
19	5.7	3.8	1.7	0.3	setosa		
20	5.1	3.8	1.5	0.3	setosa		
21	5.4	3.4	1.7	0.2	setosa		
22	5.1	3.7	1.5	0.4	setosa		
23	4.6	3.6	1.0	0.2	setosa		
24	5.1	3.3	1.7	0.5	setosa		
25	4.8	3.4	1.9	0.2	setosa		
26	5.0	3.0	1.6	0.2	setosa		
27	5.0	3.4	1.6	0.4	setosa		
28	5.2	3.5	1.5	0.2	setosa		
29	5.2	3.4	1.4	0.2	setosa		
30	4.7	3.2	1.6	0.2	setosa		

Ensuite, nous allons dans le menu **Stat/Régression/Régression logistique nominale...**:



Dans la boîte de dialogue qui apparaît nous mettons:



Et nous validons par **OK** pour obtenir:

**Régression logistique nominale : Species en fonction de Sepal.Length; Sepal.Wid**

\* ATTENTION \* L'algorithme n'a pas convergé après 20 itérations.  
 \* ATTENTION \* La convergence n'est pas atteinte pour le critère d'estimation des paramètres.  
 \* ATTENTION \* Les résultats risquent de ne pas être fiables.  
 \* ATTENTION \* Essayez d'augmenter le nombre maximum d'itérations.

## Informations de réponse

Variable	Valeur	Dénombrement
Species	virginica	50 (Evénement de référence)
	versicolor	50
	setosa	50
	Total	150

## Tableau de régression logistique

Prédicteur	Coeff	Coef ErT	Z	P
Logit 1 : (versicolor/virginica)				
Constante	42.6378	25.7077	1.66	0.097
Sepal.Length	2.46522	2.39430	1.03	0.303
Sepal.Width	6.68089	4.47956	1.49	0.136
Petal.Length	-9.42939	4.73721	-1.99	0.047
Petal.Width	-18.2861	9.74261	-1.88	0.061
Logit 2 : (setosa/virginica)				
Constante	34.8385	25715.3	0.00	0.999
Sepal.Length	10.1828	7634.52	0.00	0.999
Sepal.Width	12.5823	3558.58	0.00	0.997
Petal.Length	-24.1203	5436.32	-0.00	0.996
Petal.Width	-35.0769	8578.96	-0.00	0.997

Prédicteur	Rapport des probabilités de succès	IC à 95 %	
		Inférieur	Supérieur
Logit 1 : (versicolor/virginica)			
Constante			
Sepal.Length	11.77	0.11	1284.40
Sepal.Width	797.03	0.12	5182683.15
Petal.Length	0.00	0.00	0.87
Petal.Width	0.00	0.00	2.25
Logit 2 : (setosa/virginica)			
Constante			
Sepal.Length	26445.10	0.00	*
Sepal.Width	291354.28	0.00	*
Petal.Length	0.00	0.00	*
Petal.Width	0.00	0.00	*

Log de vraisemblance = -5.949

Vérifier que toutes les pentes sont à zéro : G = 317.685, DL = 8, Valeur de p = 0.000



## Tests d'adéquation de l'ajustement

Méthode	Khi deux	DL	P
Pearson	13.1862	288	1.000
Somme des carrés d'écart	11.8985	288	1.000

Nous pouvons conclure que nous sommes très très très loin des résultats de Tanagra. Par contre, nous sommes très proches des résultats des deux packages de R avec le même genre de conclusions.

## 13.20. Exercice 115.: Régression orthogonale (régression de Deming)

Minitab® Statistical Software 16.1.2

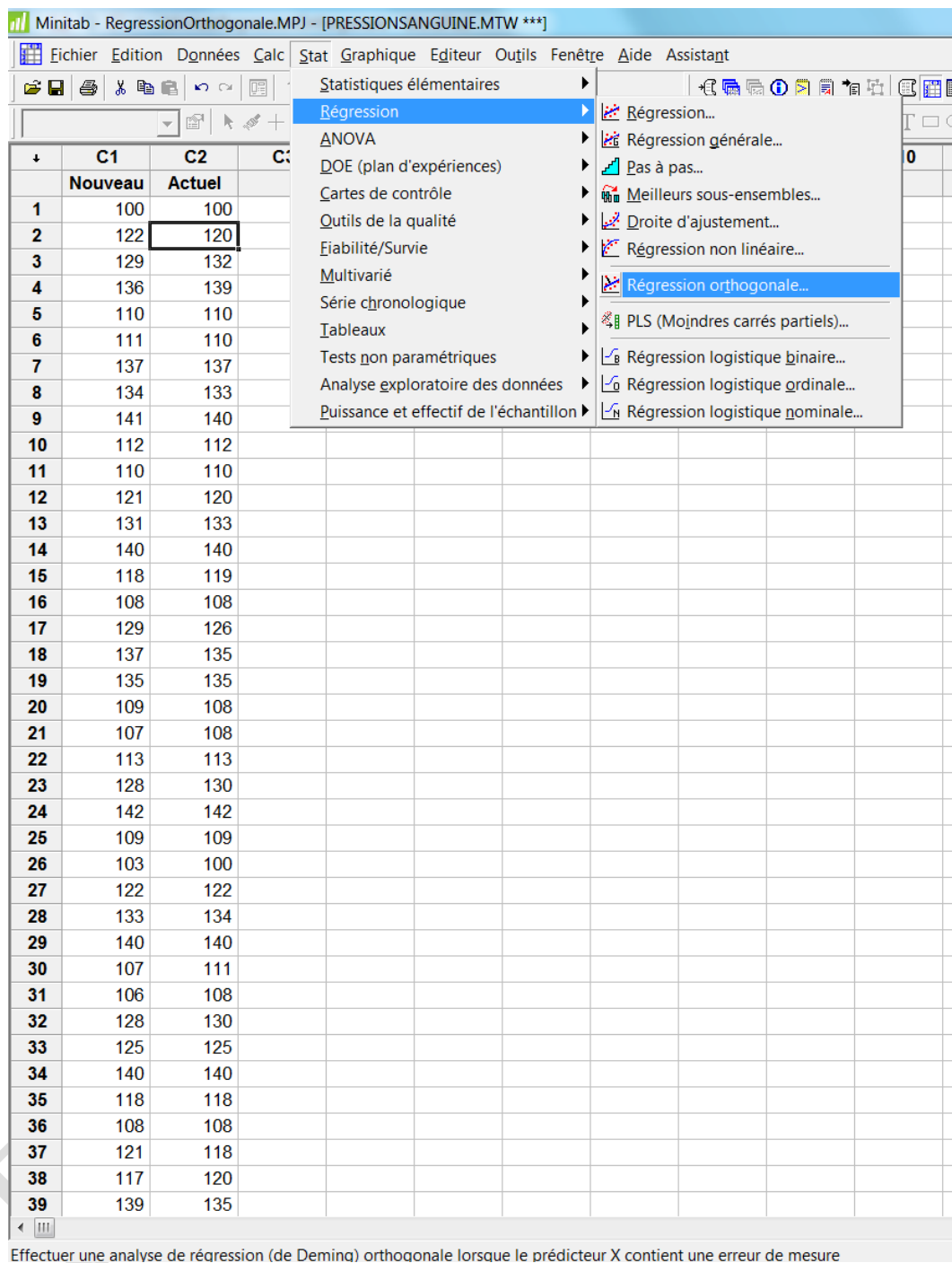
Nous allons encore une fois vérifier ici si les calculs effectués à la main et dans Microsoft Excel suite aux démonstrations mathématiques d'une partie des éléments de la régression de Deming effectuées dans le cours théorique.

Pour cela, nous utiliserons les 39 lignes suivantes (prises en partie de l'aide Minitab) qui correspondent à la mesure de la pression sanguine avec deux instruments de 39 individus (chaque individu ayant mesuré sa pression sanguine une fois avec l'ancien et une fois avec le nouvel instrument):

	C1	C2	C3	C4	C5	C6
	Nouveau	Actuel				
1	100	100				
2	122	120				
3	129	132				
4	136	139				
5	110	110				
6	111	110				
7	137	137				
8	134	133				
9	141	140				
10	112	112				
11	110	110				
12	121	120				
13	131	133				
14	140	140				
15	118	119				
16	108	108				
17	129	126				
18	137	135				
19	135	135				
20	109	108				
21	107	108				
22	113	113				
23	128	130				
24	142	142				
25	109	109				
26	103	100				
27	122	122				
28	133	134				
29	140	140				
30	107	111				
31	106	108				
32	128	130				
33	125	125				
34	140	140				
35	118	118				
36	108	108				
37	121	118				
38	117	120				
39	139	135				

Feuille de travail en cours : PRESSIONSANGUINE.MTW

Nous allons dans le menu **Stat/Régression/Régression orthogonale...**:

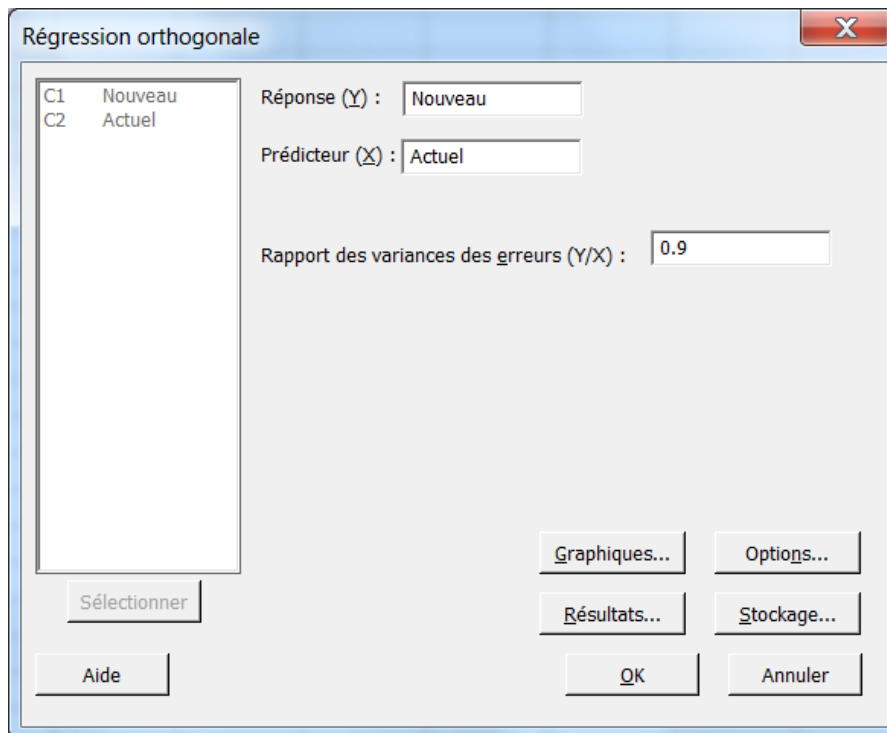


The screenshot shows the Minitab interface with the 'Stat' menu open. The 'Régression' submenu is selected, and 'Régression orthogonale...' is highlighted. The background spreadsheet contains the following data:

	C1	C2	C3
	Nouveau	Actuel	
1	100	100	
2	122	120	
3	129	132	
4	136	139	
5	110	110	
6	111	110	
7	137	137	
8	134	133	
9	141	140	
10	112	112	
11	110	110	
12	121	120	
13	131	133	
14	140	140	
15	118	119	
16	108	108	
17	129	126	
18	137	135	
19	135	135	
20	109	108	
21	107	108	
22	113	113	
23	128	130	
24	142	142	
25	109	109	
26	103	100	
27	122	122	
28	133	134	
29	140	140	
30	107	111	
31	106	108	
32	128	130	
33	125	125	
34	140	140	
35	118	118	
36	108	108	
37	121	118	
38	117	120	
39	139	135	

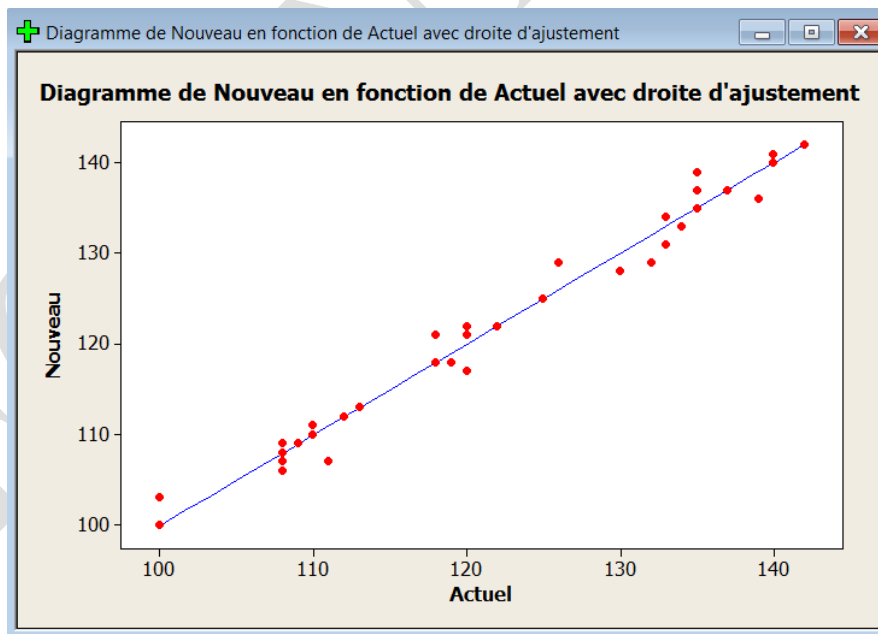
Effectuer une analyse de régression (de Deming) orthogonale lorsque le prédicteur X contient une erreur de mesure

Nous avons alors:

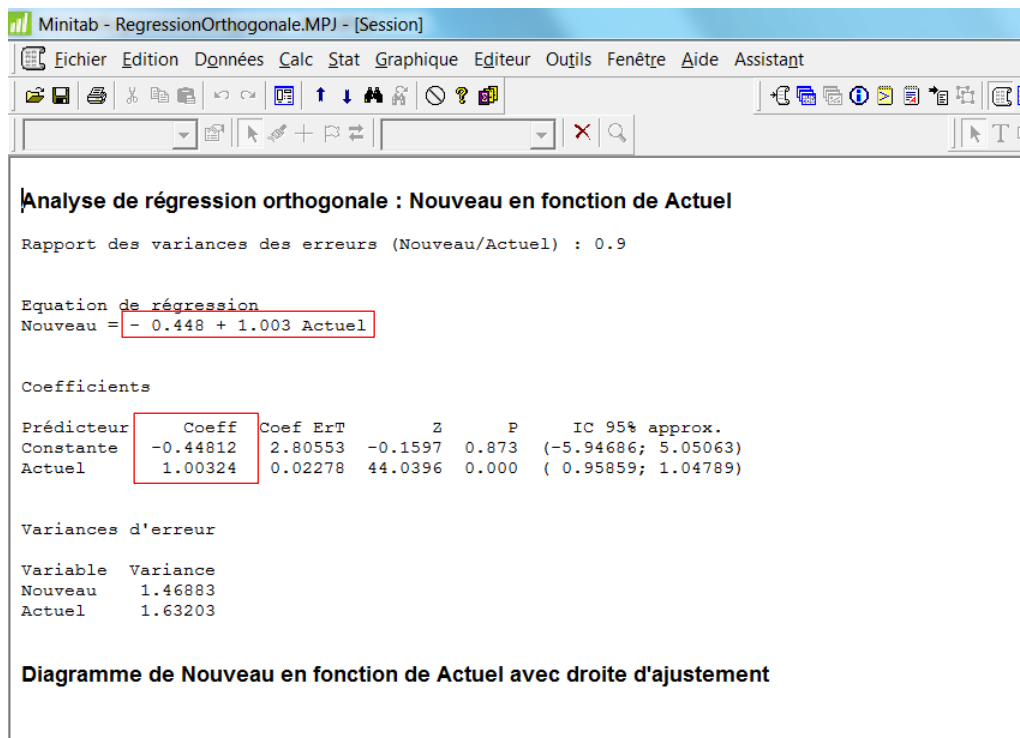


Nous y retrouvons le rapport des variances des erreurs qui est le paramètre "problématique" mais toutefois indispensable pour mener à bien cette comparaison (pouvant être déterminé dans la pratique par des mesures historiques précédant la comparaison en question).

Nous obtenons alors le graphique suivant (qui a peu d'intérêt):



et nous obtenons surtout dans la fenêtre de session:



où j'ai entouré de rouge les valeurs dont nous avons démontré théoriquement l'origine (et fait les calculs à la main). Cependant seuls ils n'ont guère d'intérêt car il nous faut absolument les intervalles de confiance.

Nous voyons alors dans le résultat ci-dessus que "1" est compris dans l'intervalle de confiance de la pente ce qui est un premier bon signe pour dire que les deux appareils sont statistiques non significativement différents.

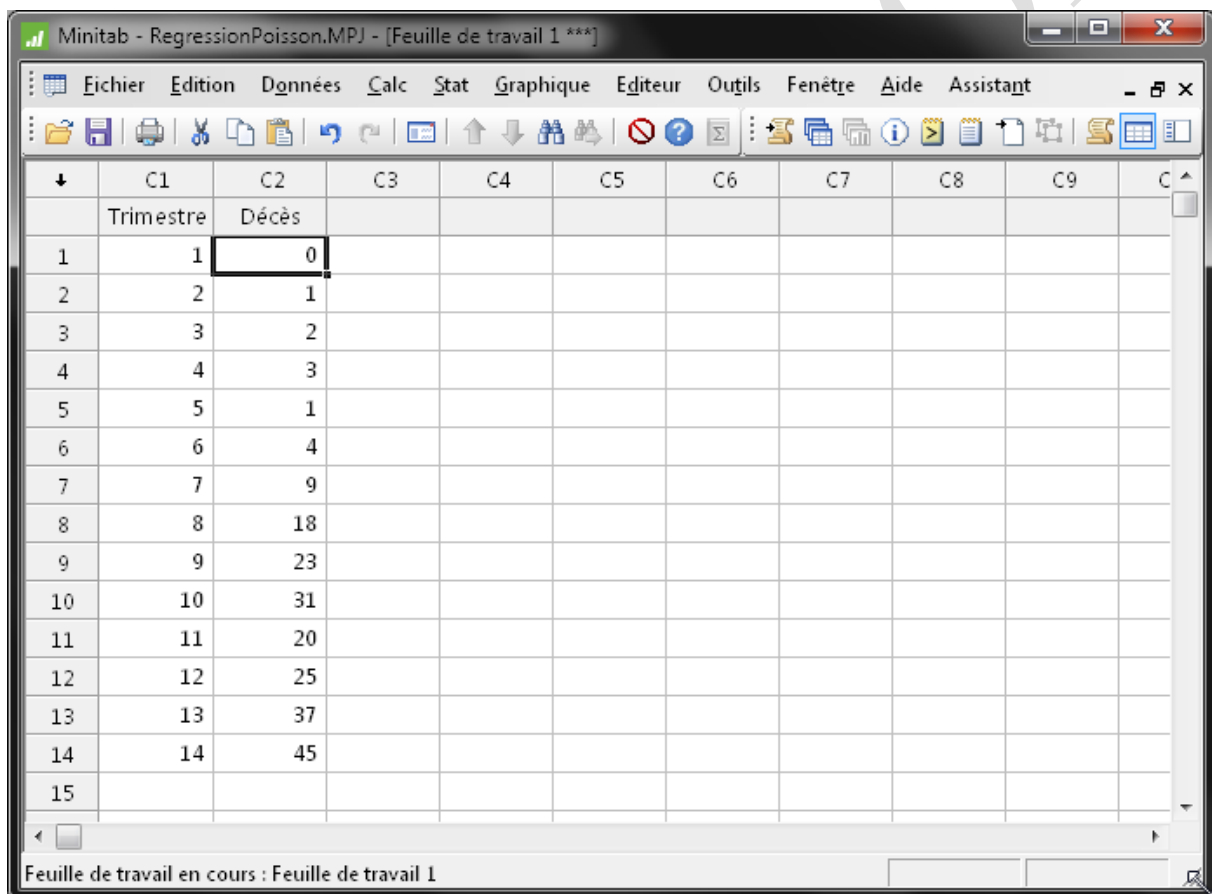
En ce qui concerne la Constante (ordonnée à l'origine) Minitab la valeur de l'intervalle de confiance est énorme!!! Mais le 0 est dans l'intervalle donc on peut considérer encore une fois que les deux appareils sont statistiques non significativement différents.

## 13.21. Exercice 116.: Régression de Poisson log-linéaire (GLM Poisson)

Minitab® Statistical Software 17.1

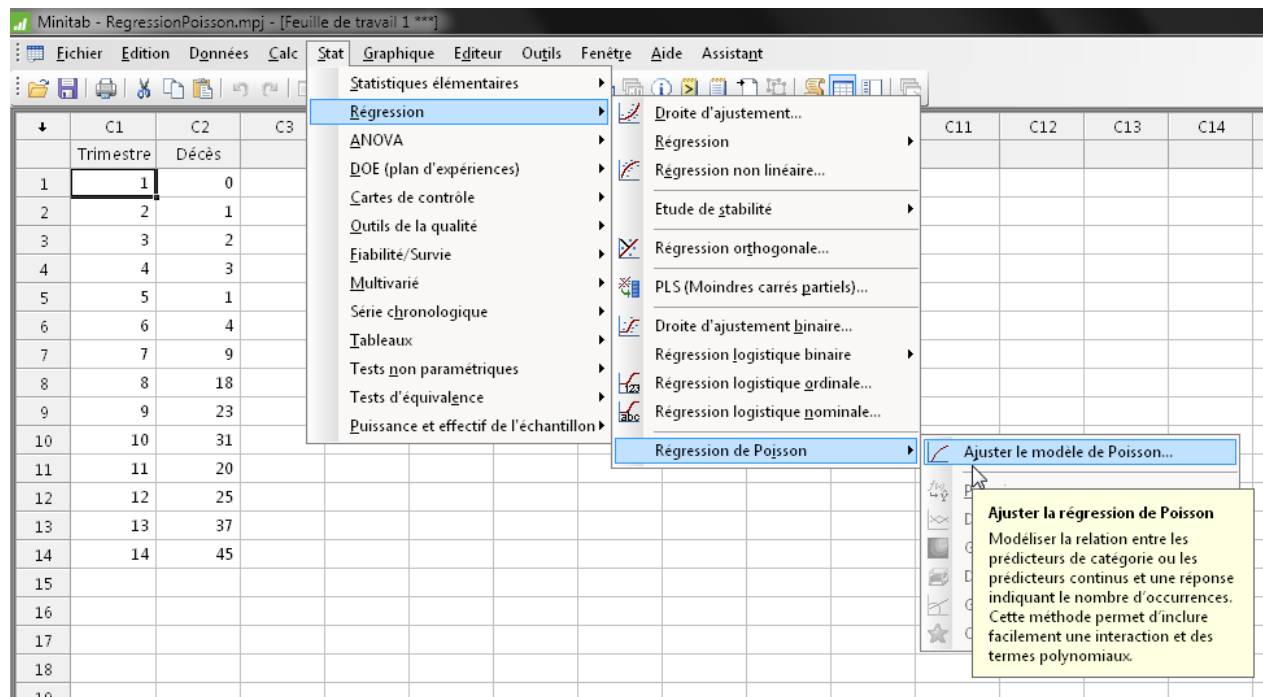
Nous allons encore une fois vérifier ici si les calculs effectués à la main et dans Microsoft Excel suite aux démonstrations mathématiques d'une partie des éléments de la régression linéaire de Poisson (événements rares) avec lien log-linéaire.

Nous partons donc des données suivantes (sans revenir sur les remarques et critiques du modèle de régression généralisée de Poisson faite déjà longuement dans le cours théorique) où nous pouvons considérer dans le cas industriel que les décès sont des arrêts définitifs de machines (elles ont rendues l'âme comme on dit):

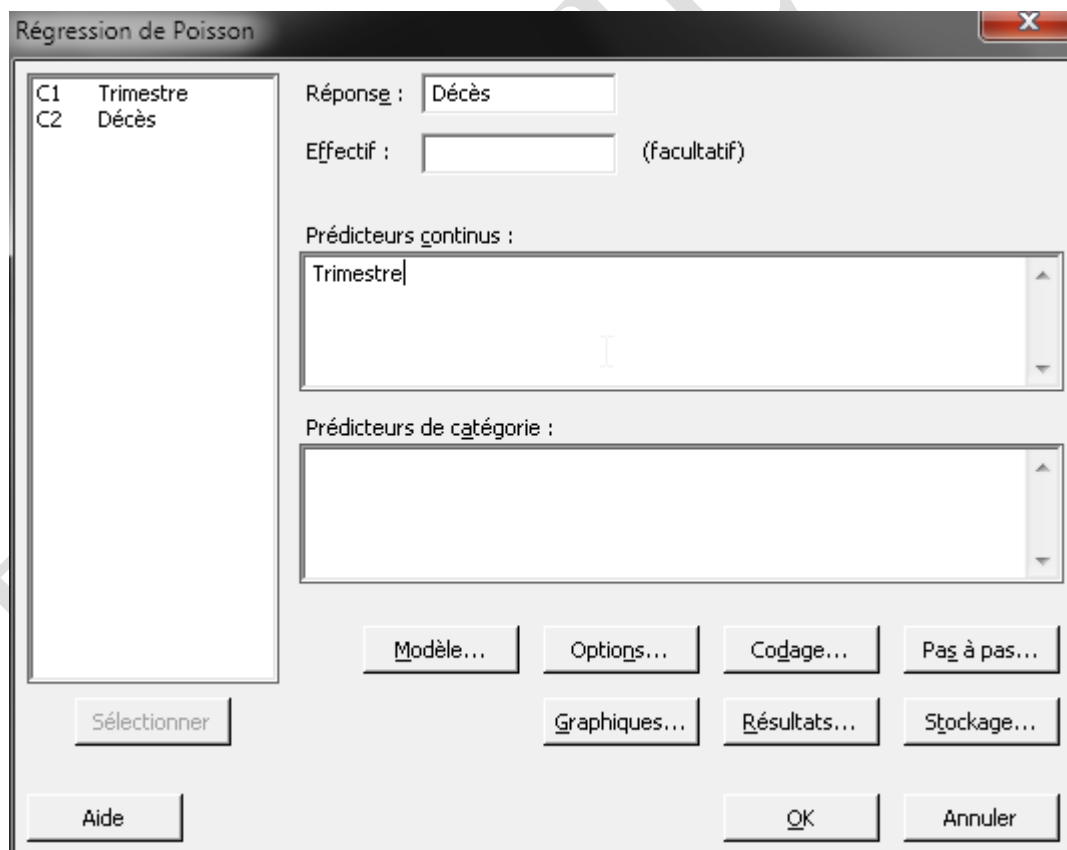


	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
	Trimestre	Décès								
1	1	0								
2	2	1								
3	3	2								
4	4	3								
5	5	1								
6	6	4								
7	7	9								
8	8	18								
9	9	23								
10	10	31								
11	11	20								
12	12	25								
13	13	37								
14	14	45								
15										

Nous allons dans le menu **Stat/Régression/Régression de Poisson/Ajuster le modèle de Poisson...**:



Ensuite, nous mettons les paramètres triviaux dans le cadre de cet exemple:



La boîte de dialogue que l'on obtient en cliquant sur le bouton **Modèle...** ne contient que du classique:

Régression de Poisson: Modèle

Prédicteurs :

Trimestre

Ajouter des termes en sélectionnant les prédicteurs et les termes du modèle :

Interactions jusqu'à l'ordre : 1 Ajouter

Termes jusqu'à l'ordre : 1 Ajouter

Termes et prédicteurs croisés dans le modèle Ajouter

Termes dans le modèle :

Valeur par défaut X ↕ ↗

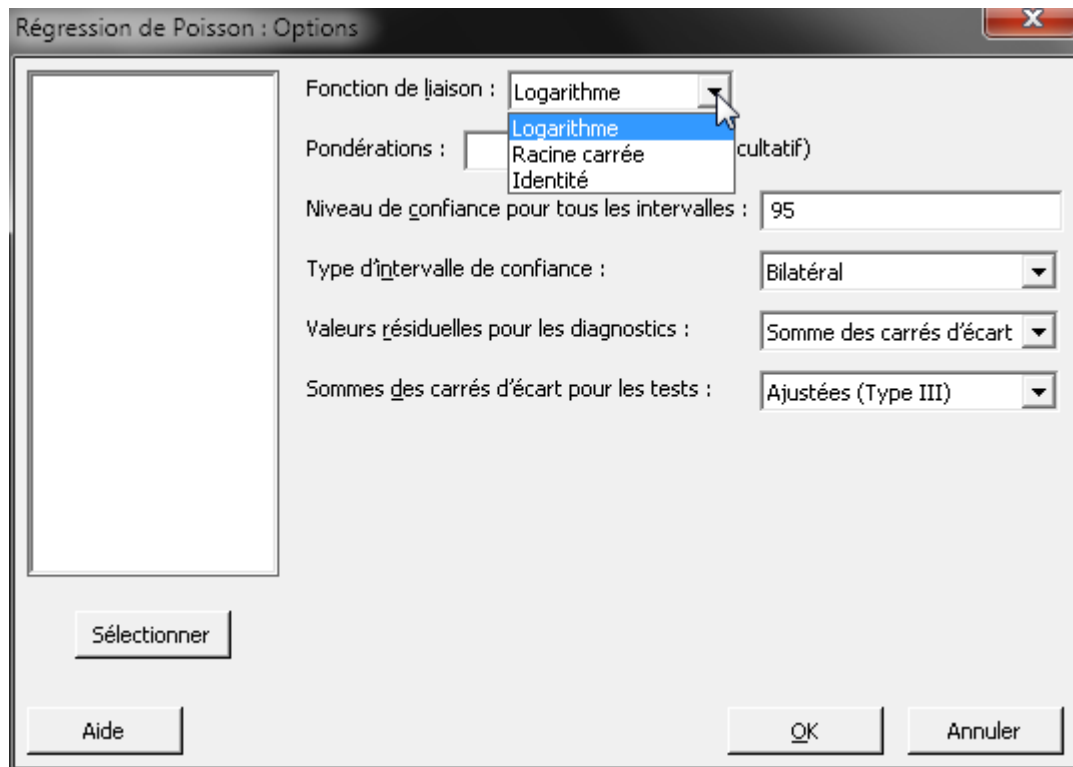
Trimestre

Inclure le terme constant dans le modèle

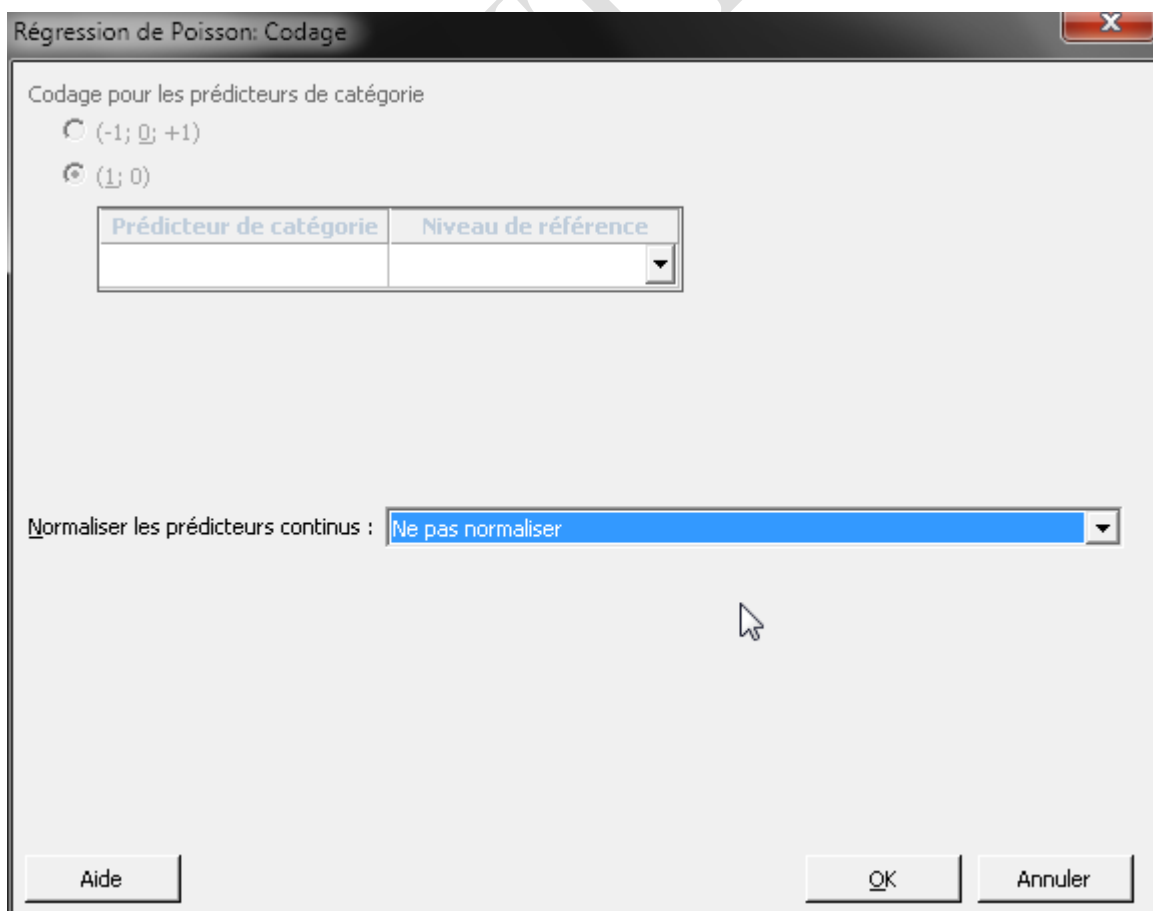
Aide OK Annuler

Ensuite en cliquant sur le bouton **Options...** de la boîte de dialogue principale, nous spécifions que nous voulons un lien de type **logarithme**:

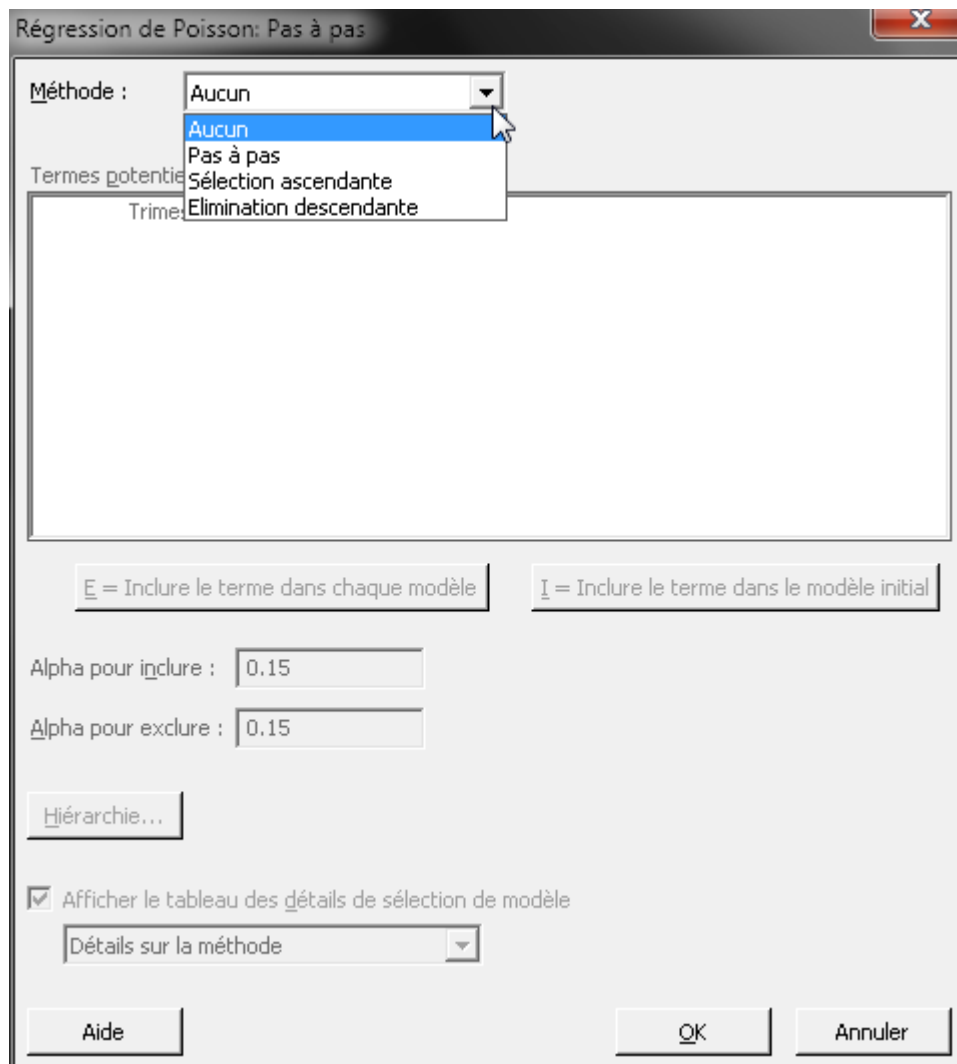




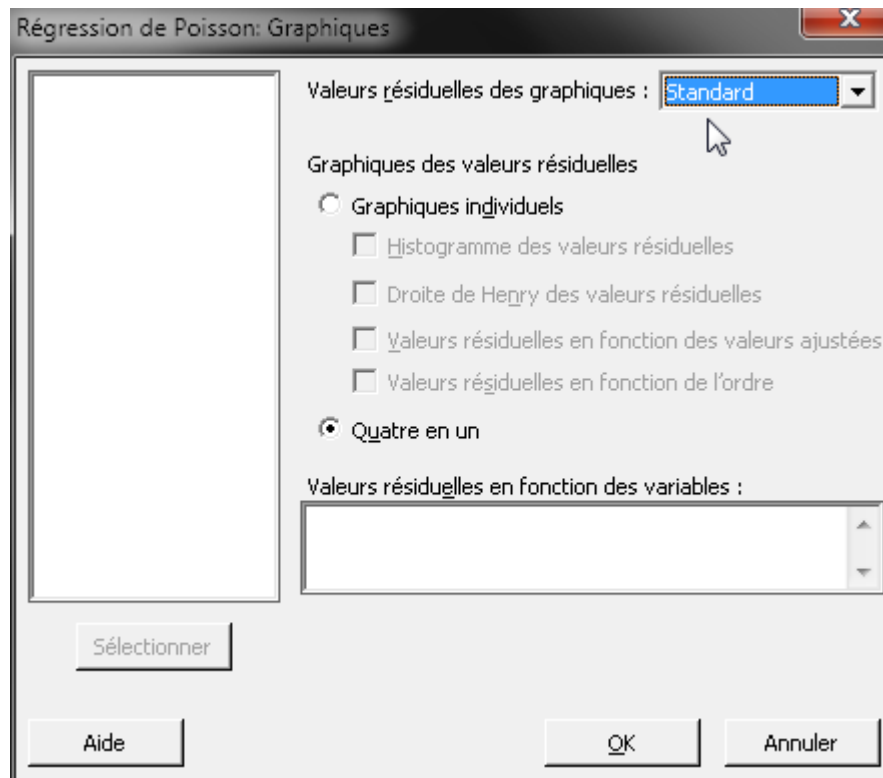
Dans le bouton **Codage...** de la boîte de dialogue principale, nous n'aurons rien besoin de mettre puisque nous n'avons pas de variables catégorielles:



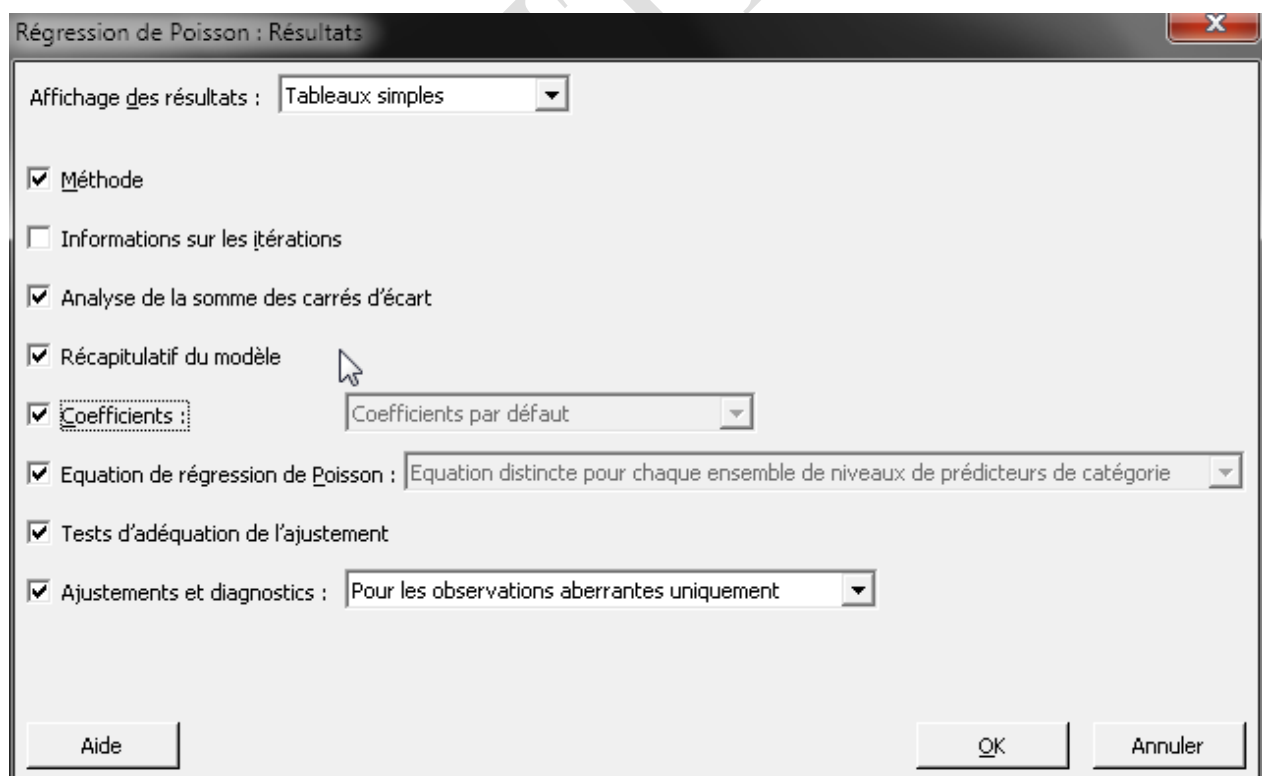
Ensuite, dans le bouton **Pas à pas...** de la boîte de dialogue principale, nous ne choisissons aucune méthode spécifique par les techniques classiques de régression (sélection de variables explicatives):



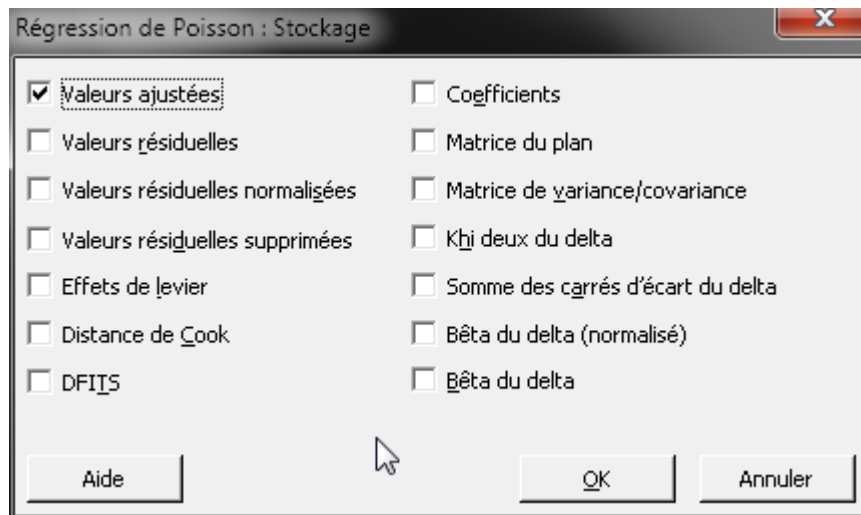
Nous prenons ensuite dans le bouton **Graphiques...** de la boîte de dialogue principale les 4 graphiques en un:



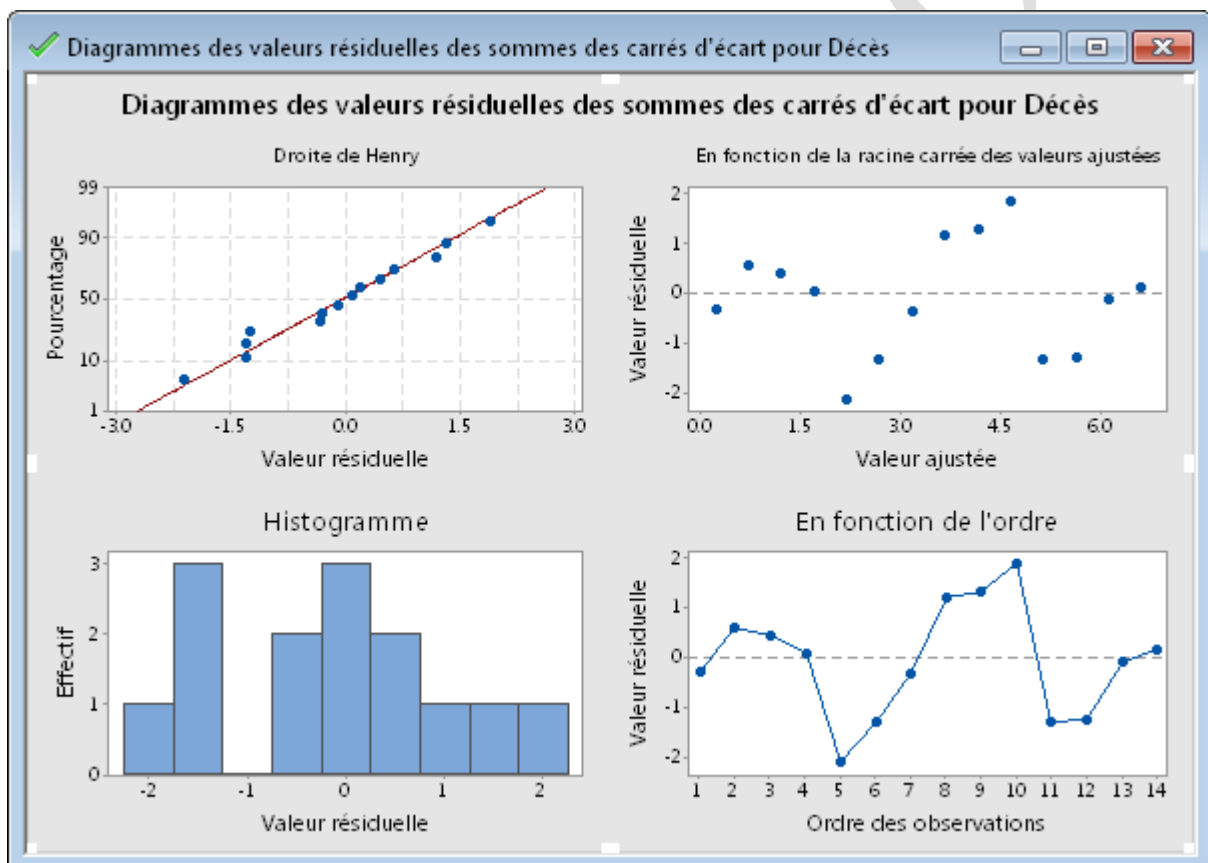
Ensuite, nous avons quelques choix de reports dans la feuille de travail en cliquant sur le bouton **Résultats...** de la boîte de dialogue principale:



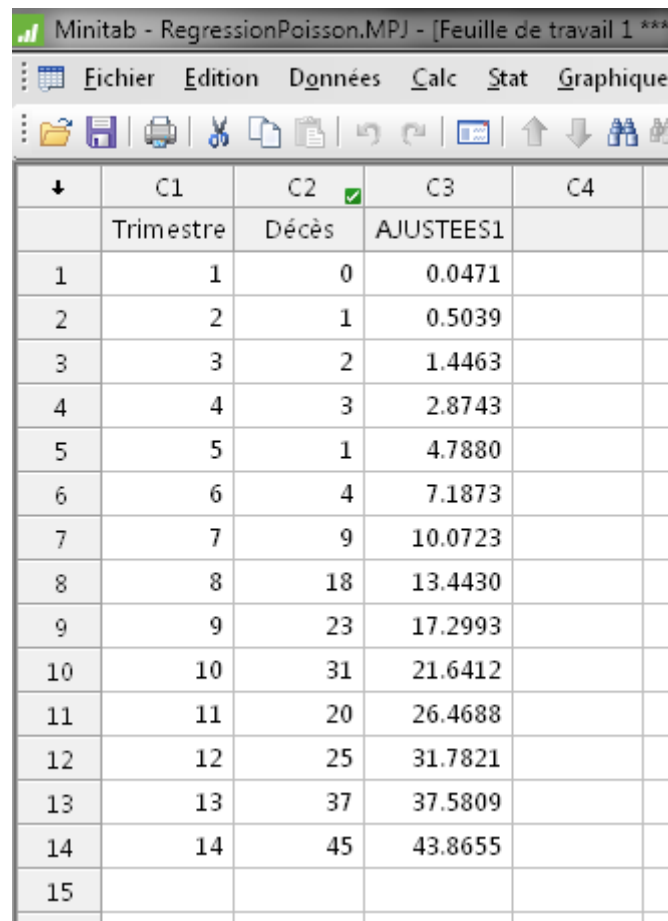
Et dans le bouton **Stockage...** de la boîte de dialogue principale:



Ce qui donne au niveau graphique:



et dans le fenêtre dans la feuille de travail:



The screenshot shows the Minitab interface with a worksheet titled "RegressionPoisson.MPJ". The menu bar includes "Fichier", "Edition", "Données", "Calc", "Stat", and "Graphique". The toolbar contains icons for file operations, editing, and navigation. The data table is as follows:

	C1	C2 <input checked="" type="checkbox"/>	C3	C4
	Trimestre	Décès	AJUSTEES1	
1	1	0	0.0471	
2	2	1	0.5039	
3	3	2	1.4463	
4	4	3	2.8743	
5	5	1	4.7880	
6	6	4	7.1873	
7	7	9	10.0723	
8	8	18	13.4430	
9	9	23	17.2993	
10	10	31	21.6412	
11	11	20	26.4688	
12	12	25	31.7821	
13	13	37	37.5809	
14	14	45	43.8655	
15				

et dans la fenêtre de session:

## Analyse de régression de Poisson : Décès en fonction de Trimestre

### Méthode

Fonction de liaison Logarithme népérien  
Lignes utilisées 14

### Tableau des sommes des carrés d'écart

Source	DL	Somme des carrés d'écart ajustée	Moyenne ajustée	Khi deux	Valeur de p
Régression	1	177.62	177.619	177.62	0.000
Trimestre	1	177.62	177.619	177.62	0.000
Erreur	12	29.65	2.471		
Total	13	207.27			

### Récapitulatif du modèle

R carré de la somme des carrés d'écart	R carré (ajust) de la somme des carrés d'écart	AIC
85.69%	85.21%	86.58

### Coefficients

Terme	Coeff	Coef ErT	FIV
Constante	0.340	0.251	
Trimestre	0.2565	0.0220	1.00

Equation de régression

Décès = exp(Y')

$$Y' = 0.340 + 0.2565 \text{ Trimestre}$$

Tests d'adéquation de l'ajustement

Test	DL	Estimation	Moyenne	Khi deux	Valeur de p
Somme des carrés d'écart	12	29.65352	2.47113	29.65	0.003
Pearson	12	28.84730	2.40394	28.85	0.004

Ajustements et diagnostics pour les observations aberrantes

Observation	Décès	Valeur ajustée	Résiduelle	Val. résid. norm.	
1	0.00	1.82	-1.91	-2.00	R
5	1.00	5.06	-2.21	-2.34	R
8	18.00	10.93	1.95	2.06	R
9	23.00	14.13	2.16	2.27	R
10	31.00	18.26	2.71	2.84	R
14	45.00	50.96	-0.85	-1.16	X

R : Valeur résiduelle élevée

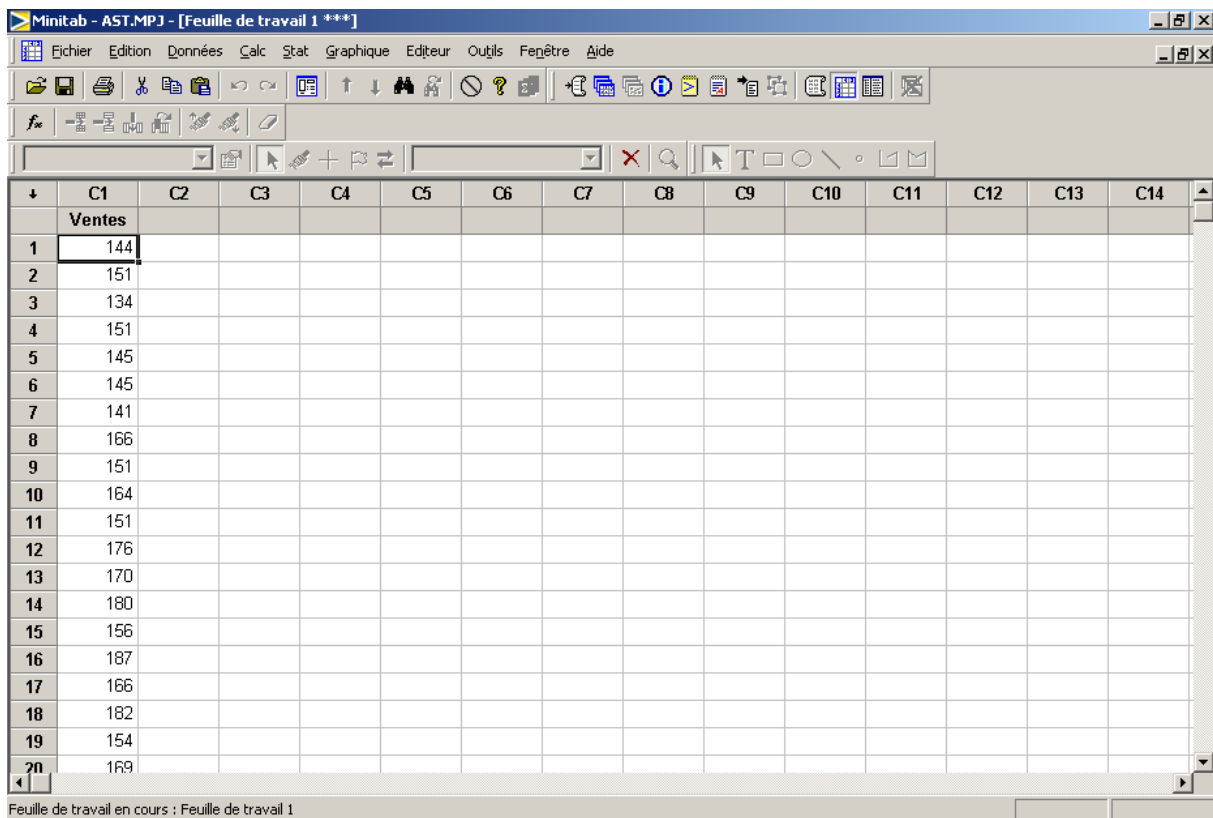
X : Valeur de X aberrante

où nous avons mis en rouge la partie que nous avons calculé à la main dans le cours théorique. Comme nous pouvons le constater il y a donc concordance parfaite!

## 13.22. Exercice 117.: Effectuer une autocorrélation et une autocorrélation partielle

Minitab® Statistical Software 15.1.1.0

Ouvrez le fichier *AST.mpj*:

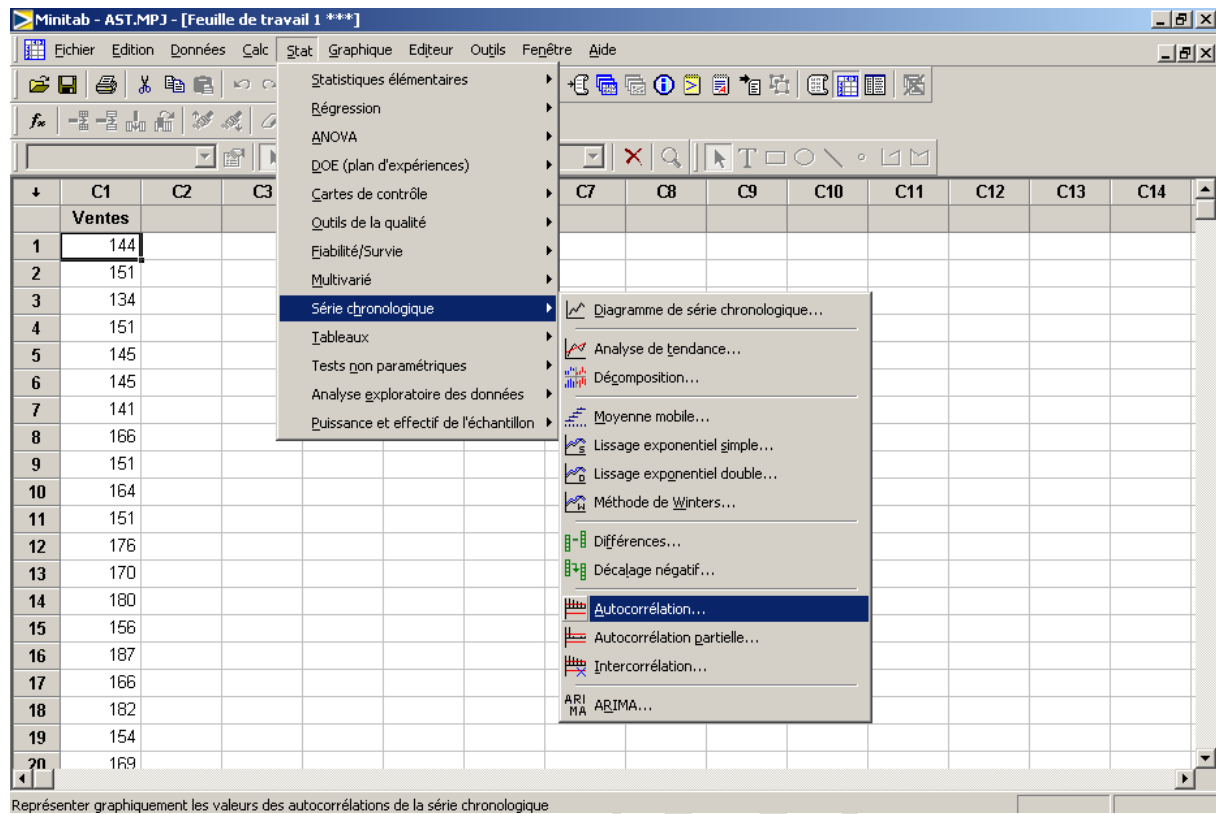


The screenshot shows the Minitab Statistical Software interface. The title bar reads "Minitab - AST.MPJ - [Feuille de travail 1 \*\*\*]". The menu bar includes "Fichier", "Edition", "Données", "Calc", "Stat", "Graphique", "Editeur", "Outils", "Fenêtre", and "Aide". The toolbar contains various icons for file operations, editing, and analysis. The worksheet grid has columns labeled C1 through C14. Column C1 is titled "Ventures" and contains the following data values from row 1 to row 20: 144, 151, 134, 151, 145, 145, 141, 166, 151, 164, 151, 176, 170, 180, 156, 187, 166, 182, 154, 169. The status bar at the bottom indicates "Feuille de travail en cours : Feuille de travail 1".

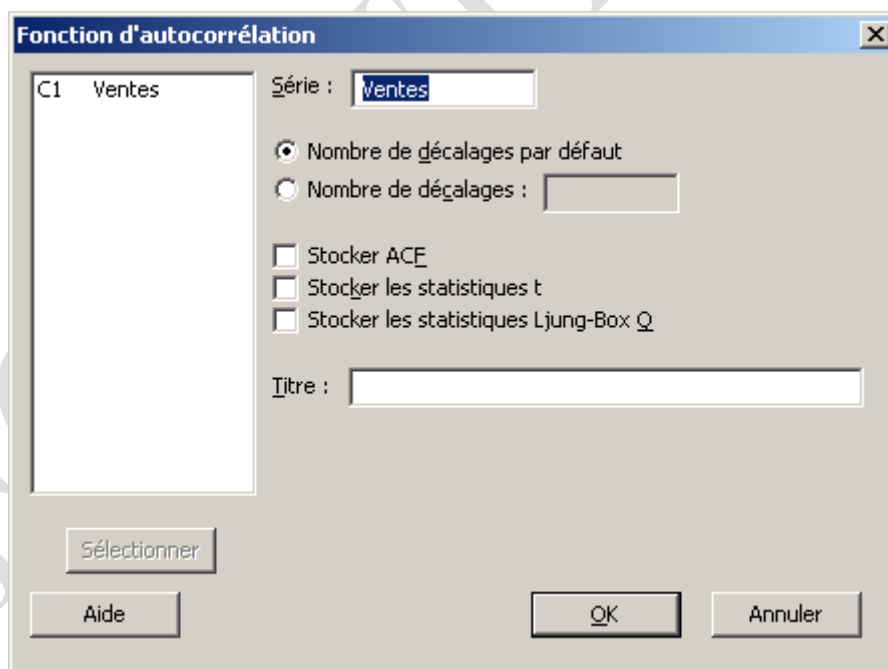
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Ventes													
1	144													
2	151													
3	134													
4	151													
5	145													
6	145													
7	141													
8	166													
9	151													
10	164													
11	151													
12	176													
13	170													
14	180													
15	156													
16	187													
17	166													
18	182													
19	154													
20	169													

Et allez dans le menu **Stat/Série chronologique/Autocorrélation...**:

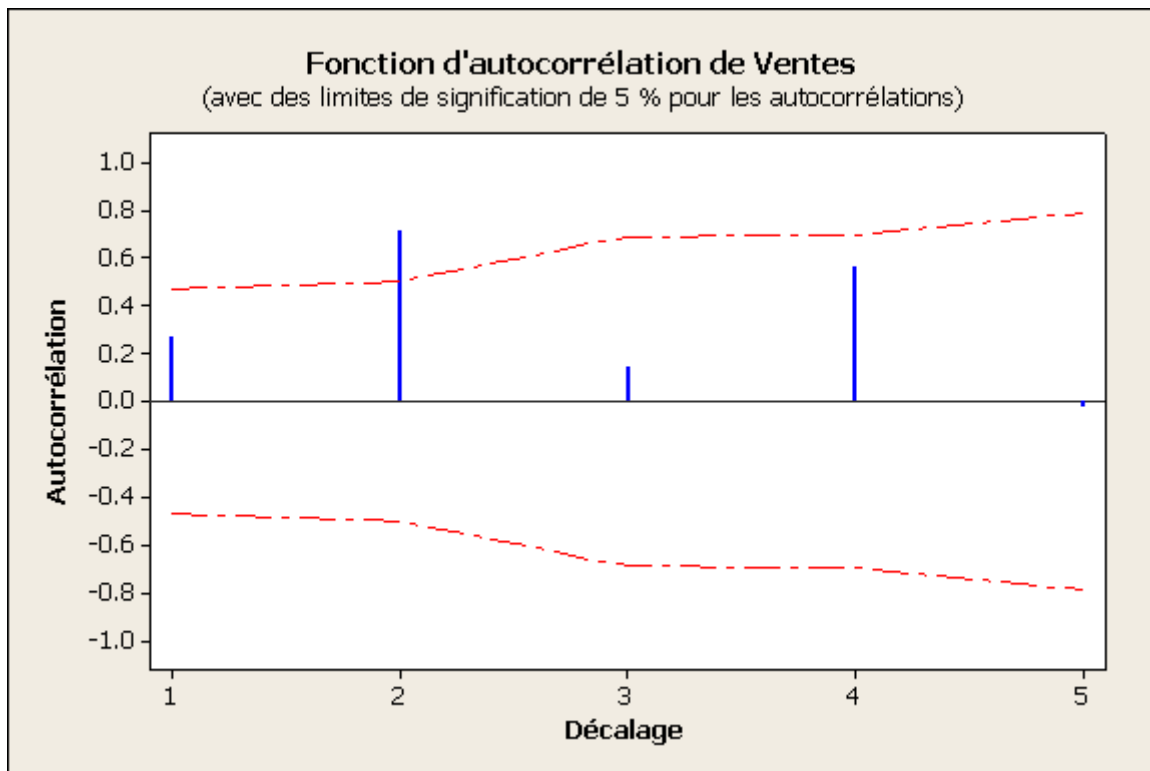




Vient alors:



Validez par **OK**:



Avec dans la fenêtre de session:

**Session**

Décalage	ACF	T	LB0
1	0.269597	1.21	1.68
2	0.712261	2.98	14.08
3	0.142952	0.43	14.61
4	0.563178	1.70	23.34
5	-0.027251	-0.07	23.36

**Autocorrélation de Ventes**

**Feuille de travail 1**

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Ventes													
1	144													
2	151													
3	134													
4	151													
5	145													
6	145													
7	141													

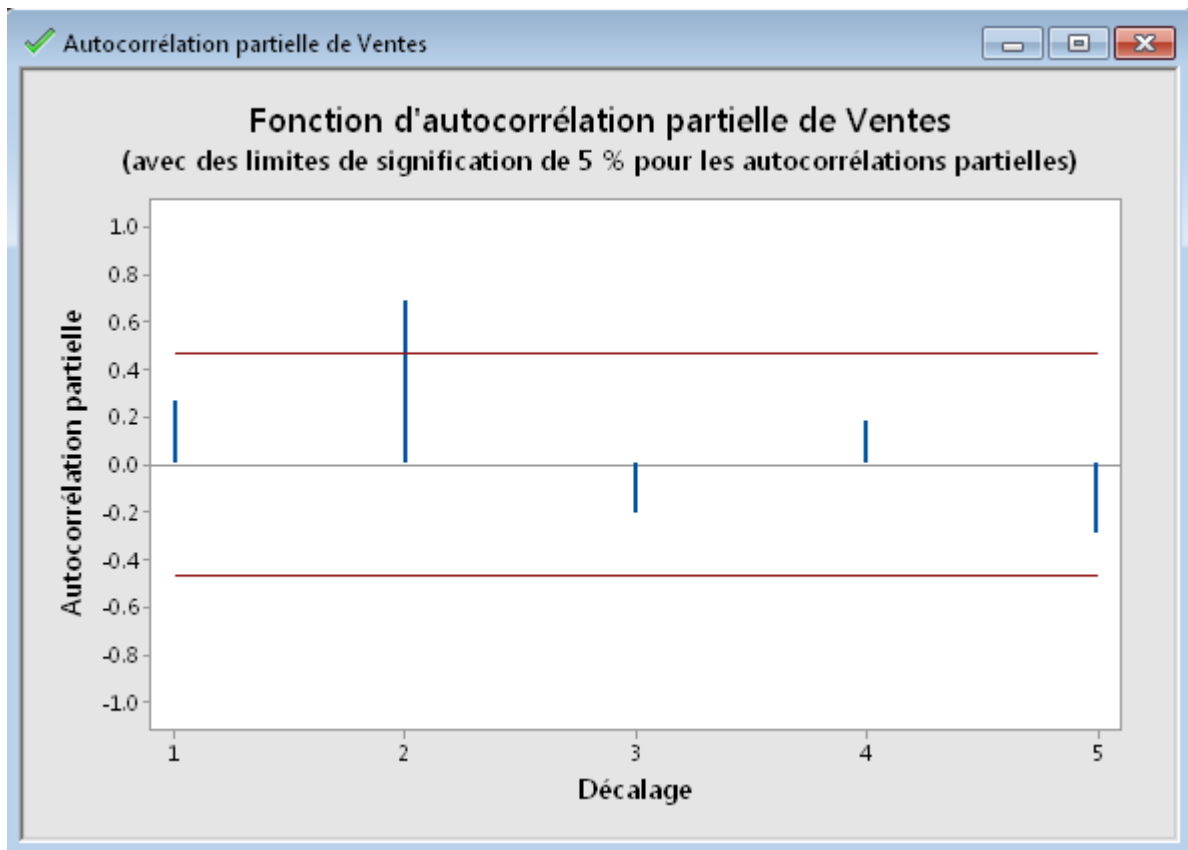
Nous retrouvons les mêmes valeurs que dans le cours théorique et que nous avons calculées à l'aide de Microsoft Excel.

Maintenant, faisons une autocorrélation partielle de ces mêmes données:

The screenshot shows the Minitab interface with the 'Stat' menu open. The 'Série chronologique' option is selected, and its sub-menu is displayed. The 'Autocorrelation partielle...' option is highlighted. A tooltip for 'Autocorrelation partielle' is visible, explaining its purpose: 'Mesurer le degré de corrélation des observations entre elles à différents points temporels après avoir supprimé l'autocorrelation pour les observations entre les différents points temporels. A utiliser avec la fonction d'autocorrelation afin d'identifier les composantes d'un modele ARIMA.'

	C1	C2	C3	C8	C9	C10	C11	C12
	Ventes							
1	144							
2	151							
3	134							
4	151							
5	145							
6	145							
7	141							
8	166							
9	151							
10	164							
11	151							
12	176							
13	170							
14	180							
15	156							
16	187							
17	166							
18	182							
19	154							
20	169							
21								
22								
23								
24								

Ce qui nous donne comme graphique:



et dans la fenêtre de session:

#### Fonction d'autocorrélation partielle : Ventes

Décalage	PACF	T
1	0.269597	1.21
2	0.689709	3.08
3	-0.210973	-0.94
4	0.179683	0.80
5	-0.287960	-1.29

Ce qui est donc conforme à ce que nous avons calculé dans le cours théorique en utilisant Microsoft Excel.

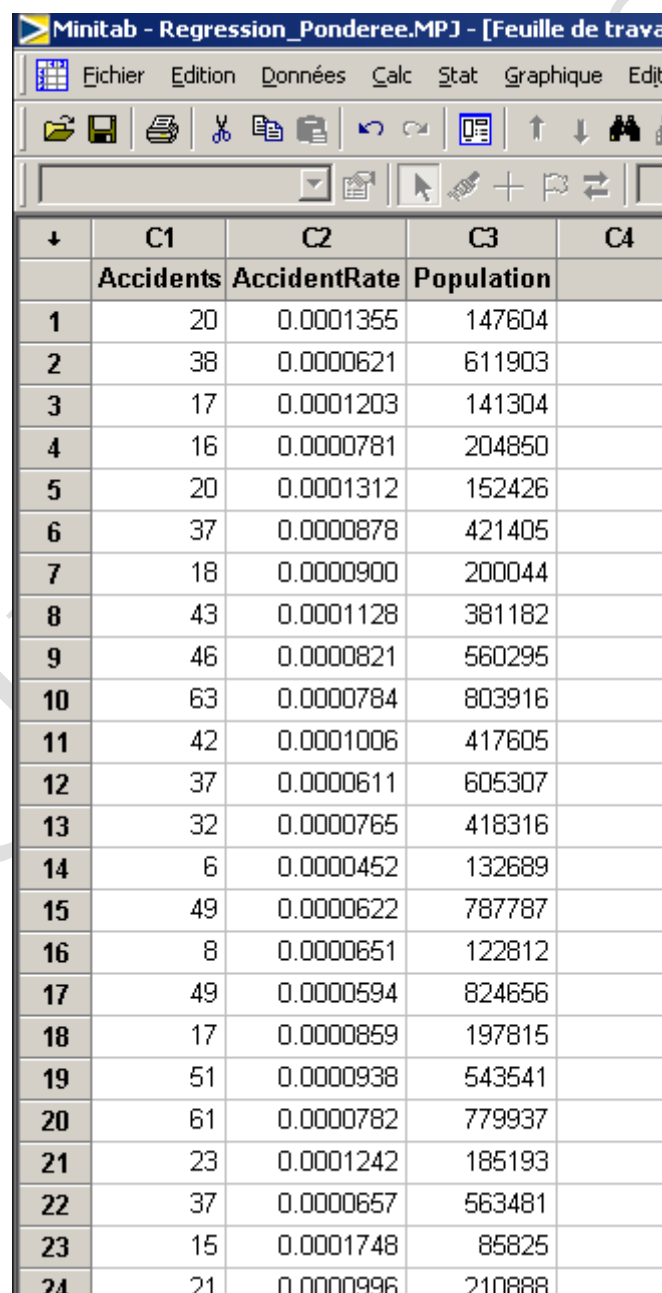
### 13.23. Exercice 118.: Régression pondérée

Minitab® Statistical Software 15.1.1.0

J'ai donc gardé la régression pondérée pour la fin de ce chapitre concernant les régressions. Effectivement, la difficulté récurrente de déterminer les poids en font finalement un outil plus rarement utilisé que tous les précédents outils (du moins selon mon expérience personnelle).

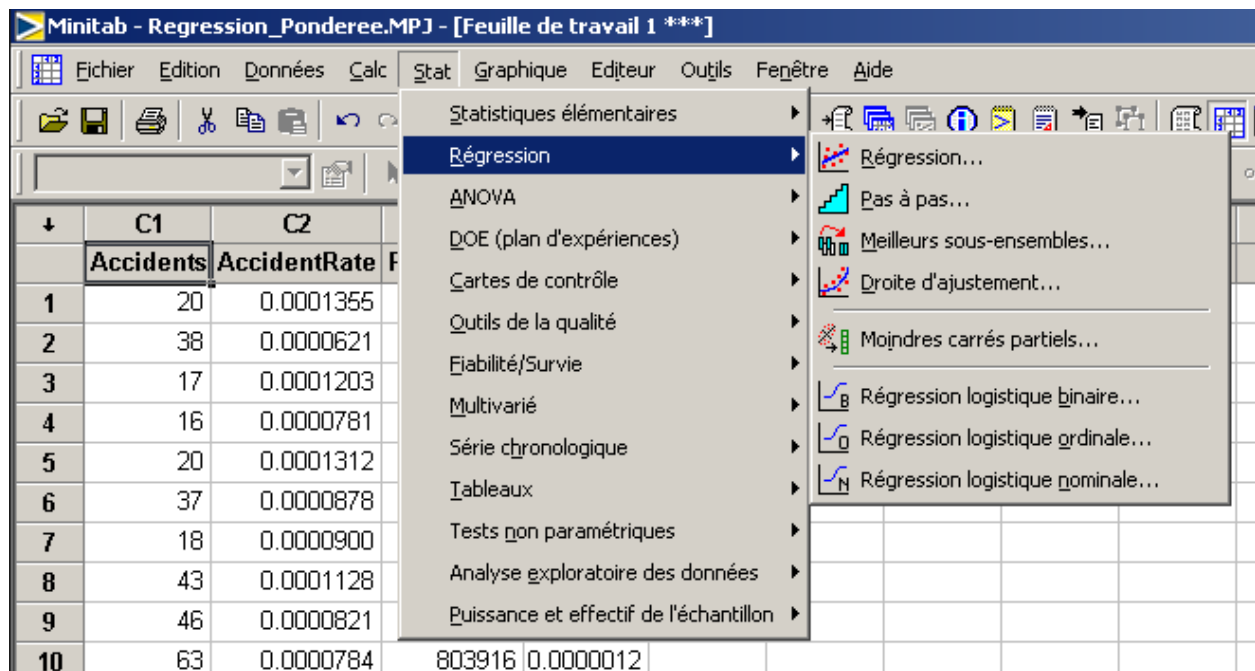
Voyons donc un exemple dont nous n'avons pas fait les calculs à la main dans le cours théorique mais dont nous avons étudié la mathématique sous-jacente suffisamment en détails!

Pour cela nous ouvrons le fichier *RegressionPonderee.mpj*:

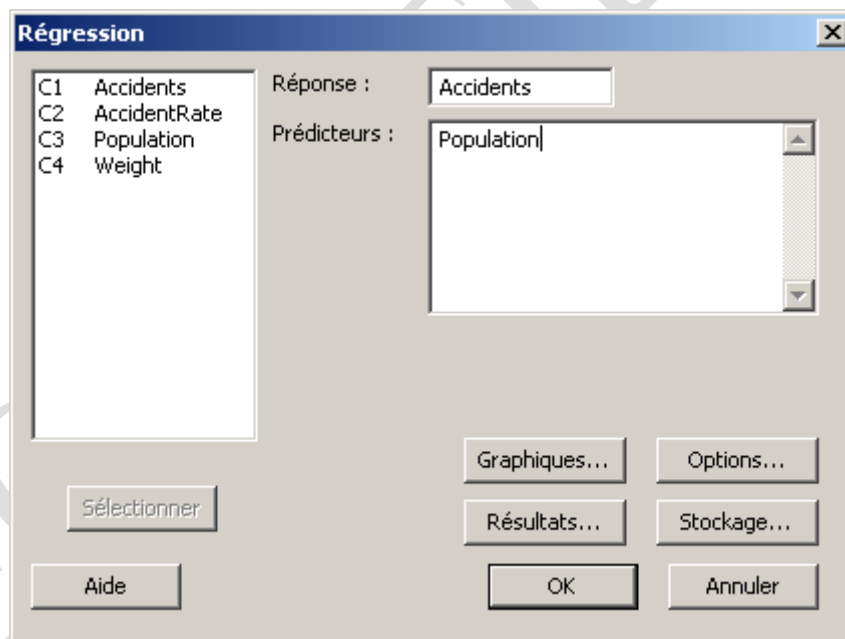


	C1	C2	C3	C4
	Accidents	AccidentRate	Population	
1	20	0.0001355	147604	
2	38	0.0000621	611903	
3	17	0.0001203	141304	
4	16	0.0000781	204850	
5	20	0.0001312	152426	
6	37	0.0000878	421405	
7	18	0.0000900	200044	
8	43	0.0001128	381182	
9	46	0.0000821	560295	
10	63	0.0000784	803916	
11	42	0.0001006	417605	
12	37	0.0000611	605307	
13	32	0.0000765	418316	
14	6	0.0000452	132689	
15	49	0.0000622	787787	
16	8	0.0000651	122812	
17	49	0.0000594	824656	
18	17	0.0000859	197815	
19	51	0.0000938	543541	
20	61	0.0000782	779937	
21	23	0.0001242	185193	
22	37	0.0000657	563481	
23	15	0.0001748	85825	
24	21	0.0000996	210888	

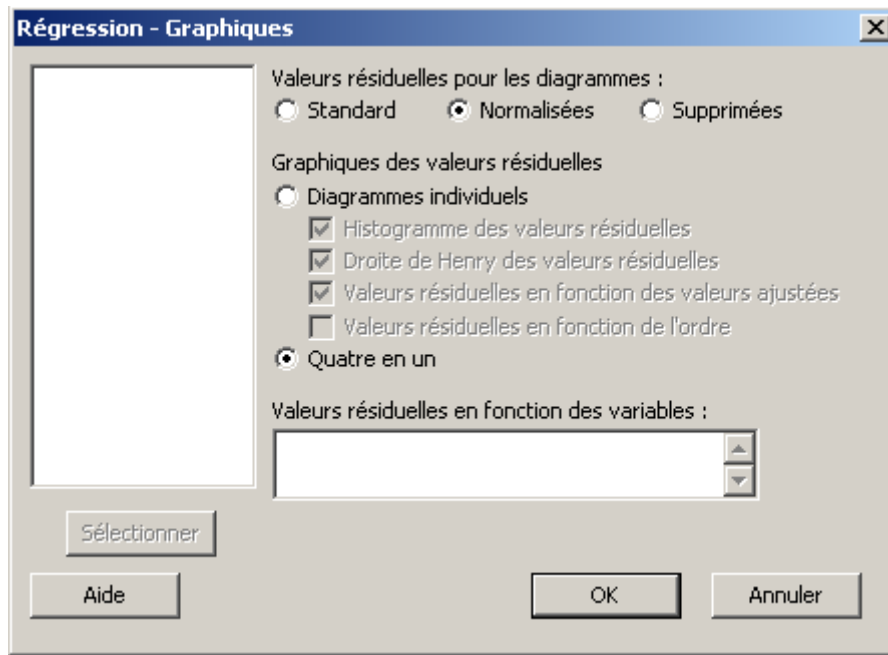
Nous allons dans le menu **Stat/Statistiques élémentaires/Régression...**:



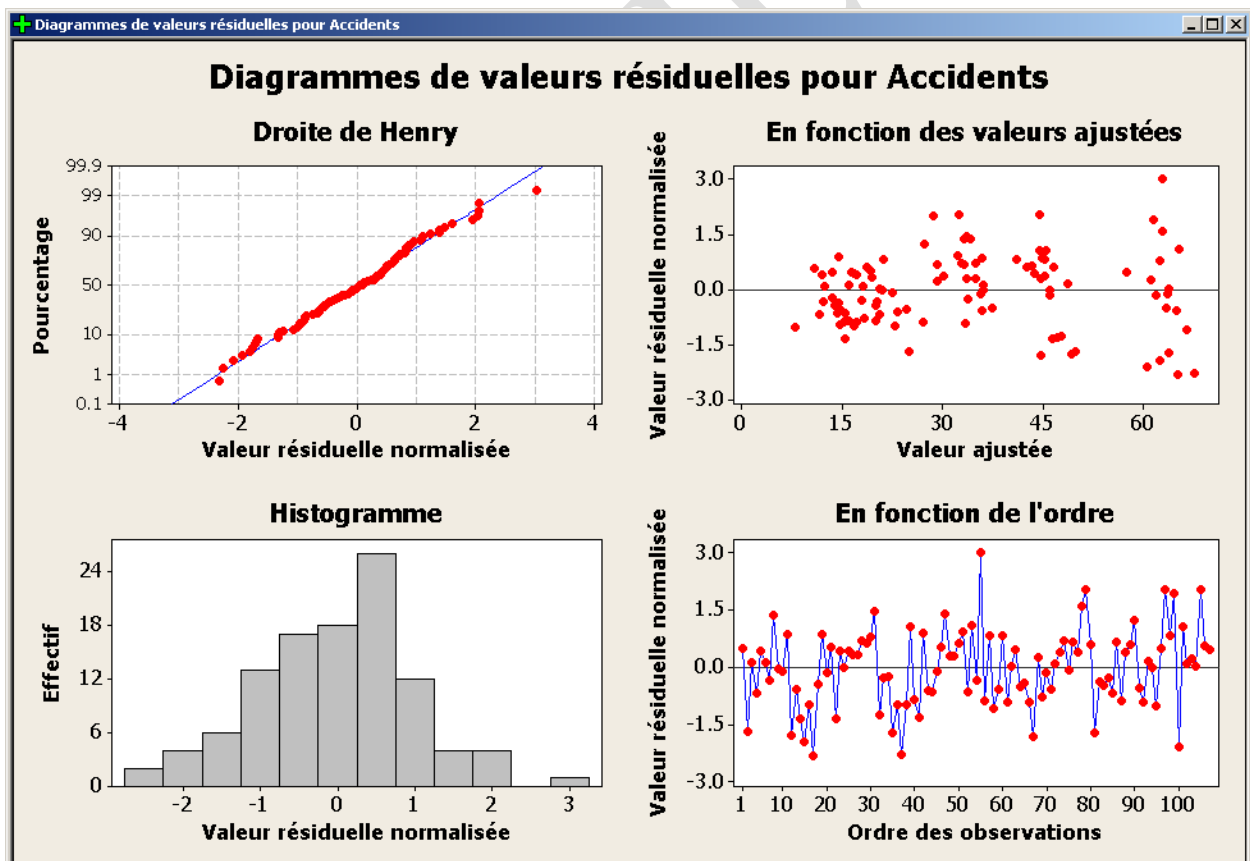
D'abord faisons une régression classique:



Avec quelques graphs usuels:



Ce qui donne graphiquement (notez le graph en haut à droite qui montre bien que le résidus ne sont pas homoscédastiques):



Et quantitativement cela donne:

**Analyse de régression : Accidents en fonction de Population**

L'équation de régression est

Accidents = 5.82 + 0.000072 Population

Prédicteur	Coeff	Coef ErT	T	P
Constante	5.822	1.313	4.43	0.000
Population	0.00007186	0.00000286	25.15	0.000

S = 7.08465 R carré = 85.8 % R carré (ajust) = 85.6 %

## Analyse de variance

Source	DL	Somme des carrés	CM	F	P
Régression	1	31741	31741	632.39	0.000
Erreur résiduelle	105	5270	50		
Total	106	37011			

## Observations aberrantes

Observation	Population	Accidents	Valeur ajustée	Ajust ErT	Valeur résiduelle
17	824656	49.000	65.084	1.413	-16.084
37	860256	52.000	67.642	1.503	-15.642
55	794016	84.000	62.882	1.338	21.118
79	317227	43.000	28.619	0.717	14.381
97	371482	47.000	32.517	0.687	14.483
100	760860	46.000	60.499	1.257	-14.499
105	538845	59.000	44.545	0.803	14.455

Observation	Valeur résiduelle normalisée
17	-2.32R
37	-2.26R
55	3.04R
79	2.04R
97	2.05R
100	-2.08R
105	2.05R

R indique une observation ayant une valeur résiduelle normalisée importante

Nous voyons que les résidus augmentent avec et à mesure que la population augmente. L'idée serait donc de pondérer les résidus par l'inverse de la population. Pour cela nous allons d'abord créer une colonne *Poids*:



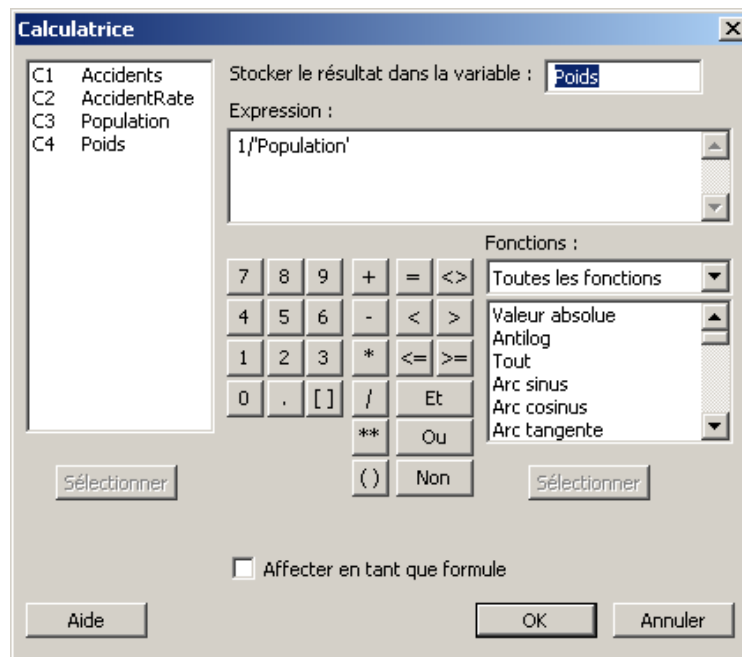
	C1	C2	C3	C4
	Accidents	AccidentRate	Population	Poids
1	20	0.0001355	147604	
2	38	0.0000621	611903	
3	17	0.0001203	141304	
4	16	0.0000781	204850	
5	20	0.0001312	152426	
6	37	0.0000878	421405	
7	18	0.0000900	200044	
8	43	0.0001128	381182	
9	46	0.0000821	560295	
10	63	0.0000784	803916	
11	42	0.0001006	417605	
12	37	0.0000611	605307	
13	32	0.0000765	418316	

Ensuite nous allons dans le menu **Calc/Calculatrice...**:

The screenshot shows the Minitab 'Calc' menu with the following options:

- Calculatrice...
- Statistiques par colonne...
- Statistiques par ligne...
- Normaliser... (with  $\bar{x}=0$  and  $s=1$  options)
- Générer des suites de nombres
- Créer des données maillées...
- Créer des variables indicatrices...
- Définir la base...
- Données aléatoires
- Lois de probabilité
- Matrices

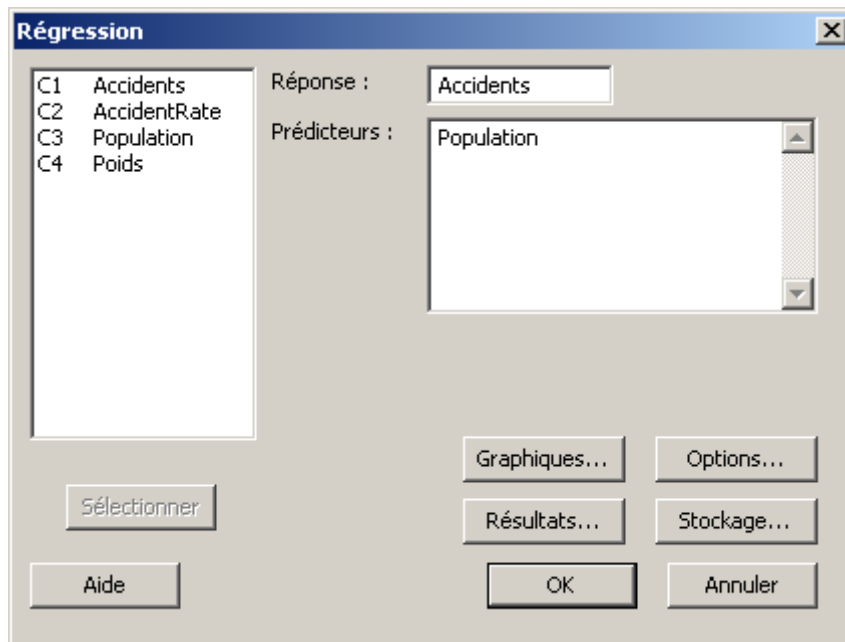
Et prenons:



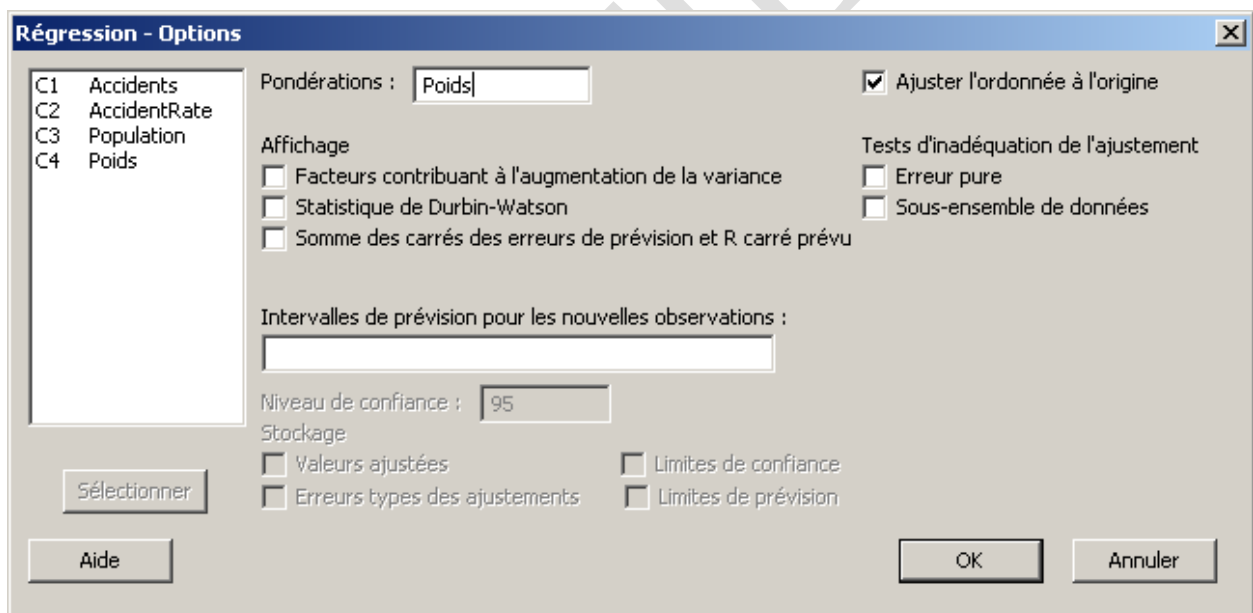
Ce qui donner après validation sur **OK**:

	C1	C2	C3	C4
	Accidents	AccidentRate	Population	Poids
1	20	0.0001355	147604	0.0000068
2	38	0.0000621	611903	0.0000016
3	17	0.0001203	141304	0.0000071
4	16	0.0000781	204850	0.0000049
5	20	0.0001312	152426	0.0000066
6	37	0.0000878	421405	0.0000024
7	18	0.0000900	200044	0.0000050
8	43	0.0001128	381182	0.0000026
9	46	0.0000821	560295	0.0000018
10	63	0.0000784	803916	0.0000012
11	42	0.0001006	417605	0.0000024
12	37	0.0000611	605307	0.0000017
13	32	0.0000765	418316	0.0000024
14	6	0.0000452	132689	0.0000075
15	49	0.0000622	787787	0.0000013
16	8	0.0000651	122812	0.0000081
17	49	0.0000594	824656	0.0000012
18	17	0.0000859	197815	0.0000051
19	51	0.0000938	543541	0.0000018
20	61	0.0000782	779937	0.0000013

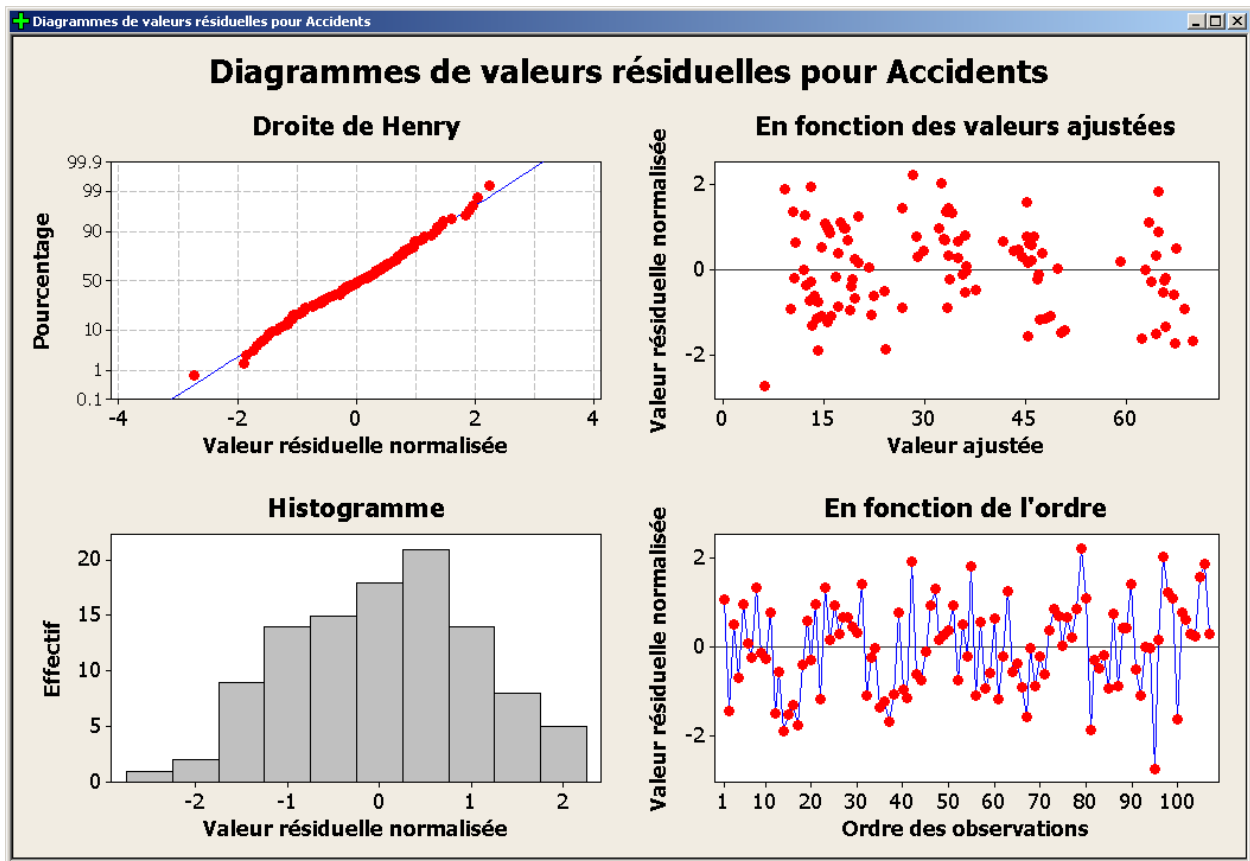
Nous relançons la régression comme avant:



Mais cette fois nous cliquons sur **Options...** pour définir le champ **Pondérations**:



Nous validons le tout deux fois par **OK** pour obtenir d'abord graphiquement:



Remarquez que les résidus ont une bien meilleure forme maintenant (qualitativement parlant bien sûr!) et pour l'analyse quantitative, nous avons:

**Analyse de régression : Accidents en fonction de Population**

Analyse pondérée avec les pondérations de Poids

L'équation de régression est  
 Accidents = 3.82 + 0.000077 Population

Prédicteur	Coeff	Coef ErT	T	P
Constante	3.8152	0.8672	4.40	0.000
Population	0.00007698	0.00000286	26.88	0.000

S = 0.0117756    R carré = 87.3 %    R carré (ajust) = 87.2 %

Analyse de variance

Source	DL	Somme des carrés	CM	F	P
Régression	1	0.10021	0.10021	722.70	0.000
Erreur résiduelle	105	0.01456	0.00014		
Total	106	0.11477			

Observations aberrantes

Observation	Population	Accidents	Valeur ajustée	Ajust ErT	Valeur résiduelle
79	317227	43.0000	28.2359	0.6000	14.7641
95	30000	1.0000	6.1247	0.8026	-5.1247
97	371482	47.0000	32.4125	0.6770	14.5875

Observation	Valeur résiduelle normalisée
79	2.24R
95	-2.73RX
97	2.04R

R indique une observation ayant une valeur résiduelle normalisée importante  
 X indique une observation ayant un effet de levier important.

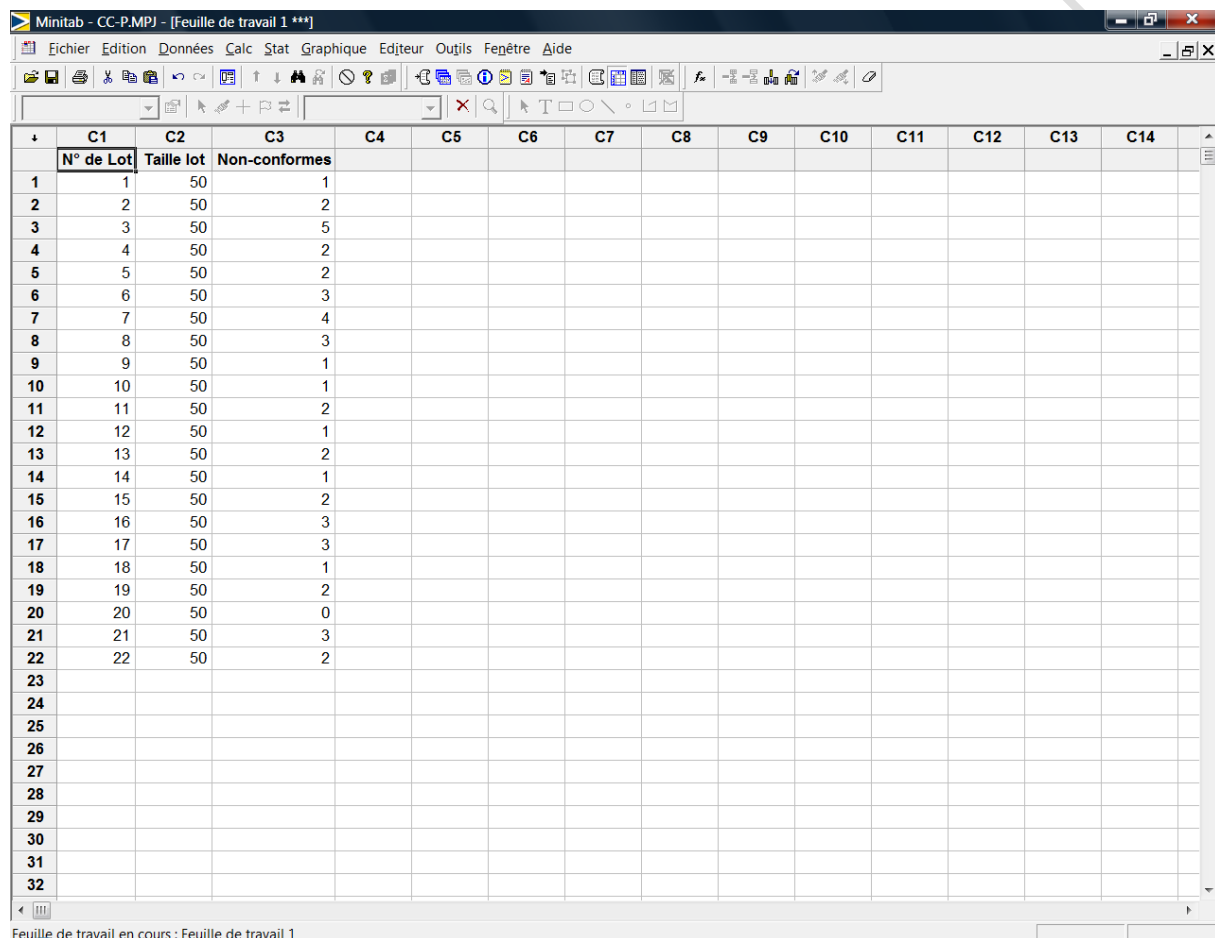
Nous avons donc aussi pour le coup moins de valeurs aberrantes comme nous pouvions nous y attendre!

# 14. Cartes de contrôle (CC)

## 14.1. Exercice 119.: CC par attributs p (proportion)

Minitab® Statistical Software 15.1.1.0

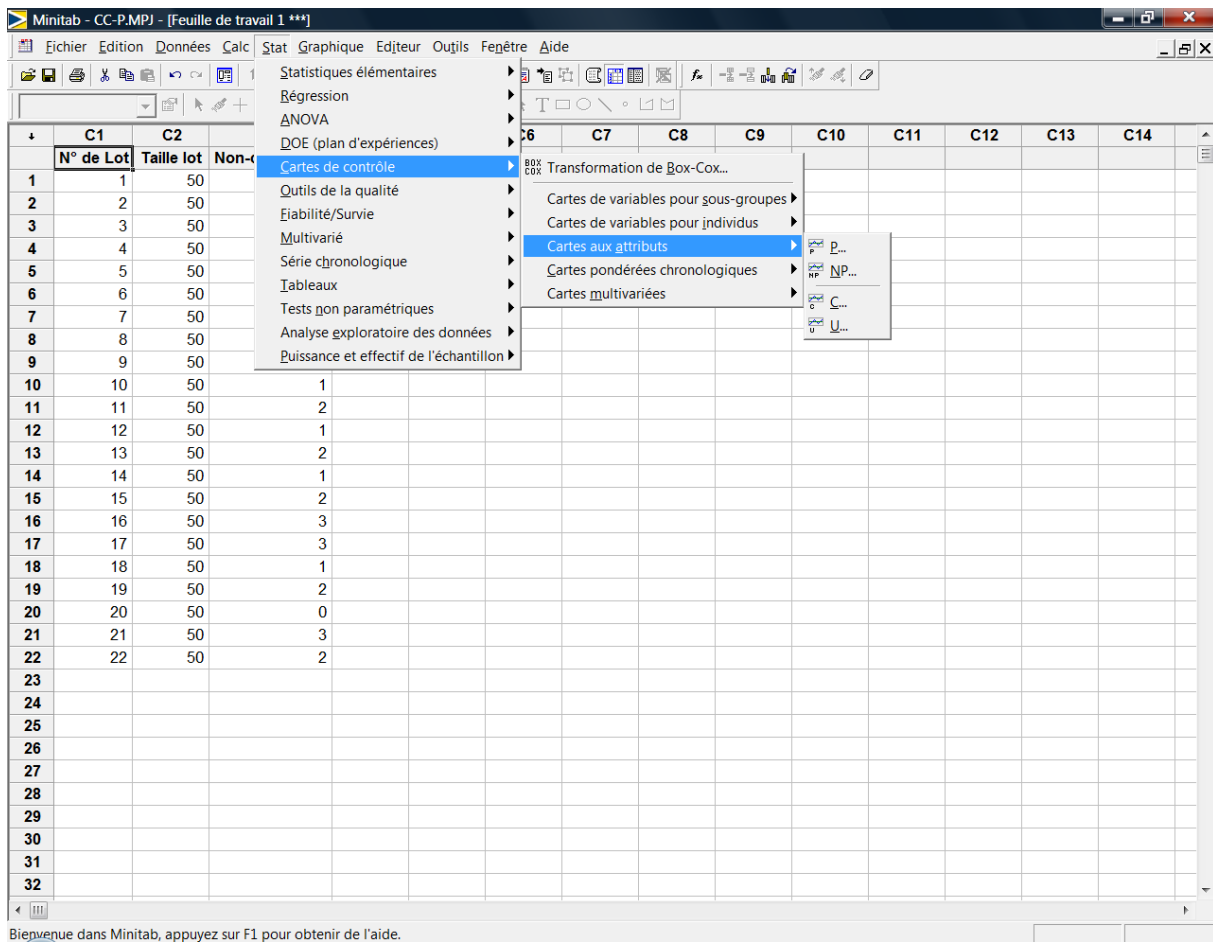
Dans le cadre de contrôle de lots nous avons obtenus les mesures suivantes (fichier *CC-P.mpj*):



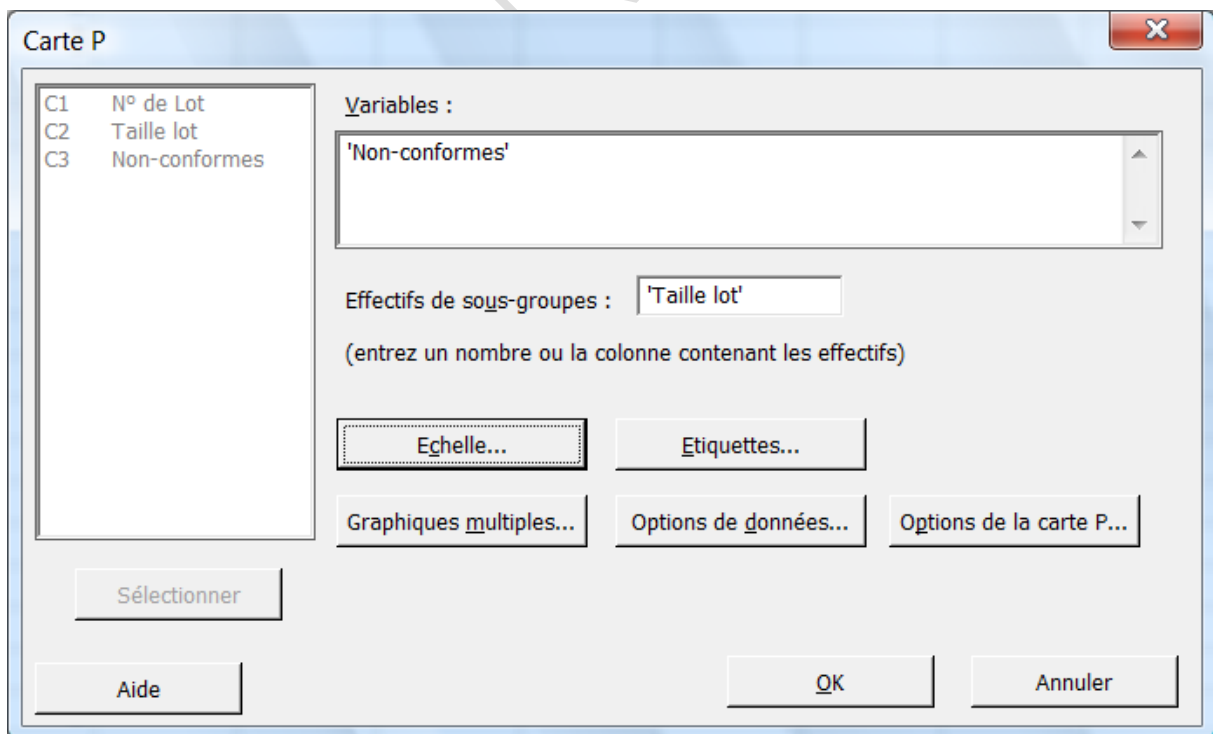
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	N° de Lot	Taille lot	Non-conformes											
1	1	50	1											
2	2	50	2											
3	3	50	5											
4	4	50	2											
5	5	50	2											
6	6	50	3											
7	7	50	4											
8	8	50	3											
9	9	50	1											
10	10	50	1											
11	11	50	2											
12	12	50	1											
13	13	50	2											
14	14	50	1											
15	15	50	2											
16	16	50	3											
17	17	50	3											
18	18	50	1											
19	19	50	2											
20	20	50	0											
21	21	50	3											
22	22	50	2											
23														
24														
25														
26														
27														
28														
29														
30														
31														
32														

Nous souhaiterions tracer une carte de contrôle par attributs de type  $p$  et comparer avec les calculs théoriques effectués dans le cours de statistique théorique et dans le cours Microsoft Excel.

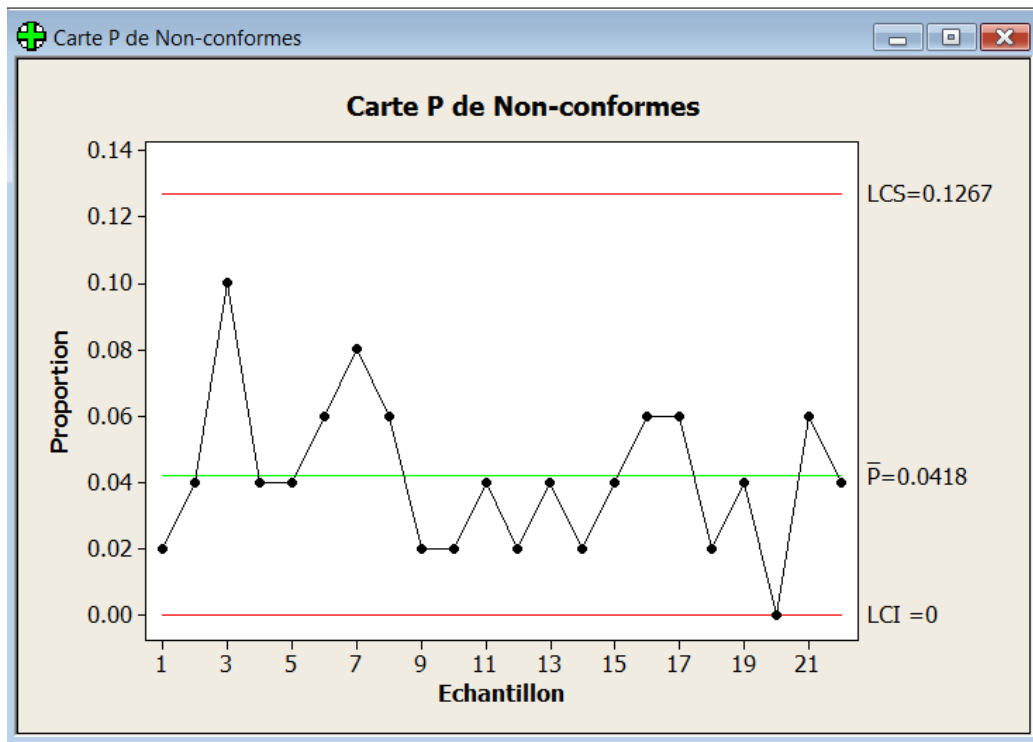
Dans Minitab® Statistical Software nous allons dans le menu **Stat/Cartes de contrôle/Cartes aux attributs/P...**:



Nous avons alors la boîte de dialogue suivante:



Nous validons par **OK**:



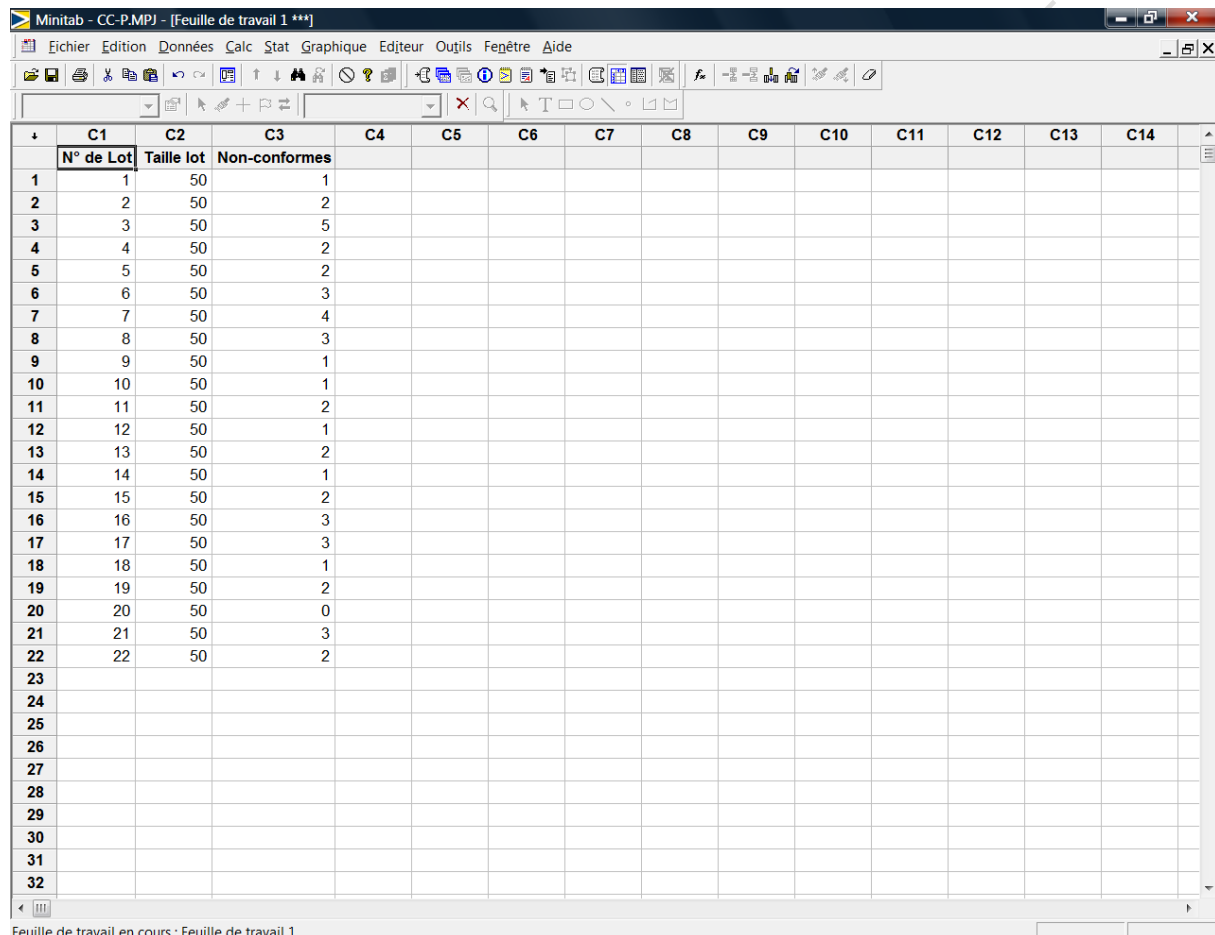
Nous obtenons donc les mêmes valeurs et la même carte que dans le cours de statistique théorique et Microsoft Excel.



## 14.2. Exercice 120.: CC par attributs de type np (comptage par proportion)

Minitab® Statistical Software 15.1.1.0

Dans le cadre de contrôle de lots nous avons obtenu les mesures suivantes (fichier *CC-NP.mpj*):

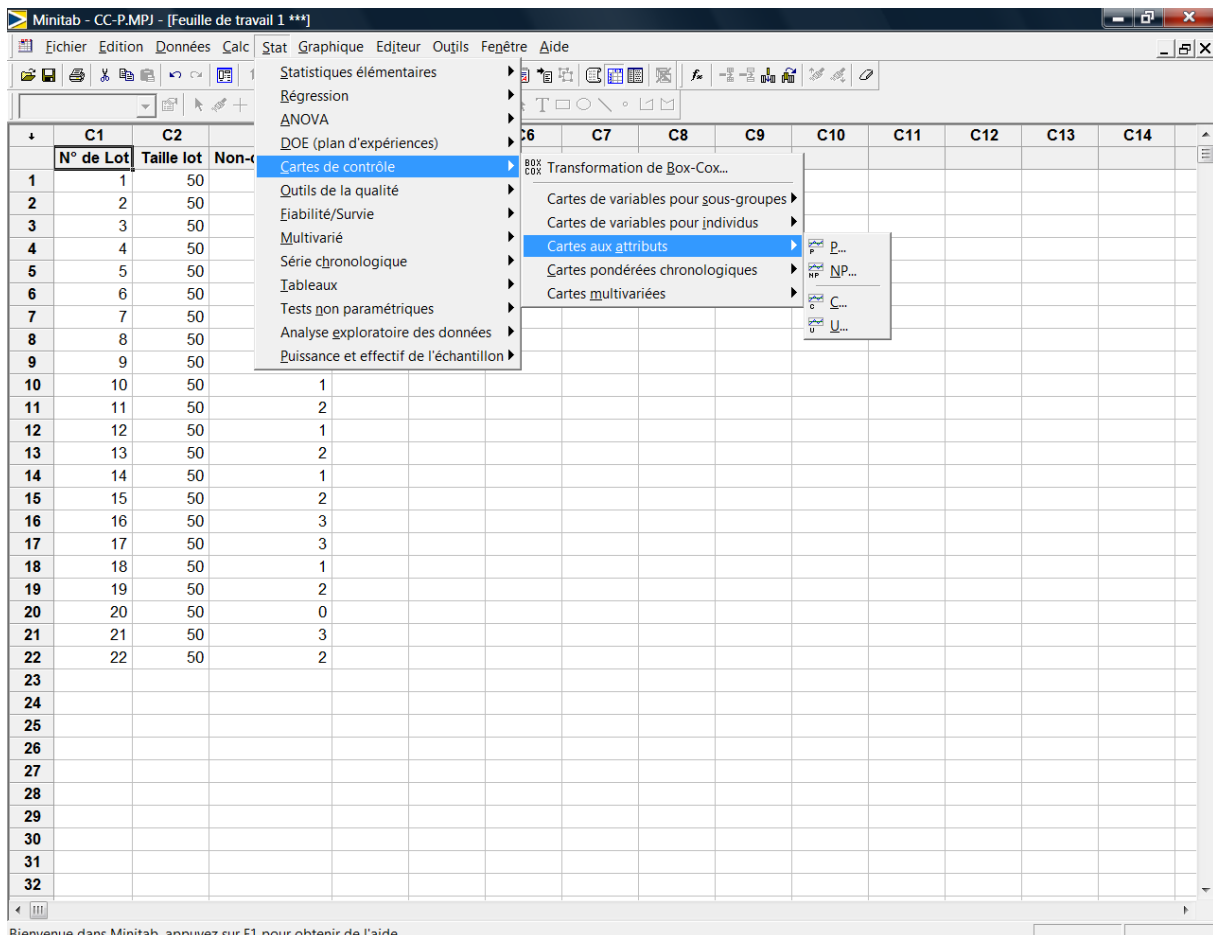


The screenshot shows the Minitab Statistical Software interface. The main window displays a data table with the following columns: C1 (N° de Lot), C2 (Taille lot), C3 (Non-conformes), and columns C4 through C14. The data is as follows:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	N° de Lot	Taille lot	Non-conformes											
1	1	50	1											
2	2	50	2											
3	3	50	5											
4	4	50	2											
5	5	50	2											
6	6	50	3											
7	7	50	4											
8	8	50	3											
9	9	50	1											
10	10	50	1											
11	11	50	2											
12	12	50	1											
13	13	50	2											
14	14	50	1											
15	15	50	2											
16	16	50	3											
17	17	50	3											
18	18	50	1											
19	19	50	2											
20	20	50	0											
21	21	50	3											
22	22	50	2											
23														
24														
25														
26														
27														
28														
29														
30														
31														
32														

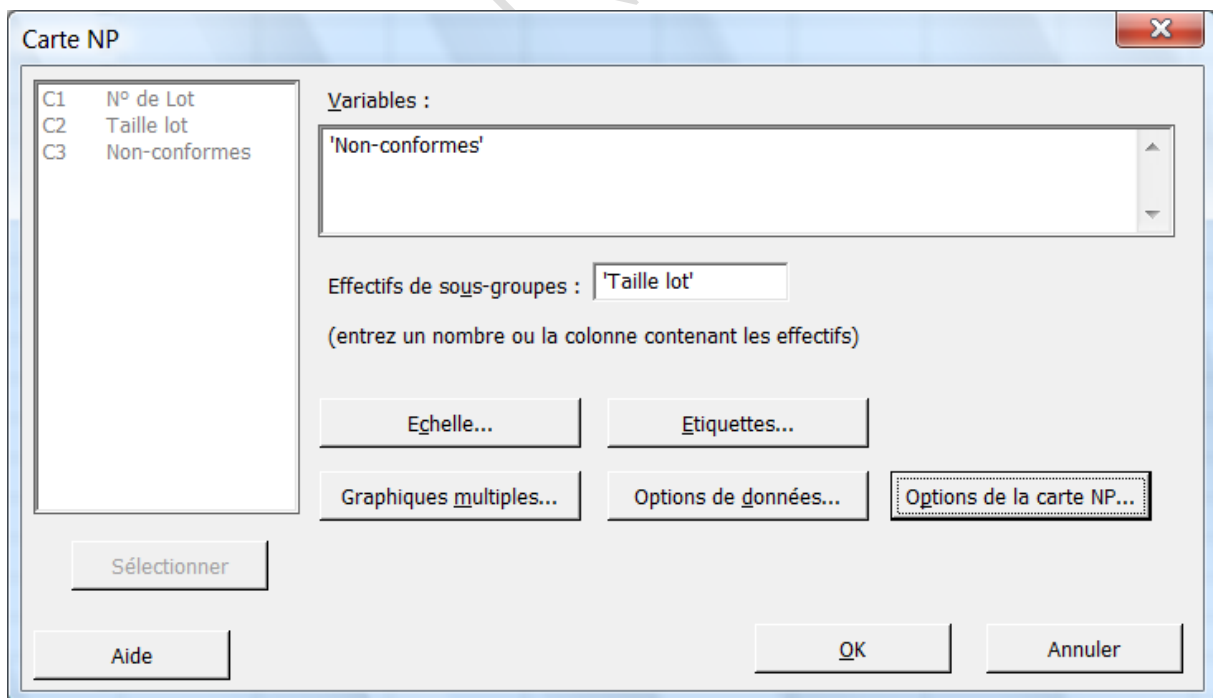
Nous souhaiterions tracer une carte de contrôle par attributs de type  $np$  et comparer avec les calculs théoriques effectués dans le cours de statistique théorique et dans le cours Microsoft Excel.

Dans Minitab® Statistical Software nous allons dans le menu **Stat/Cartes de contrôle/Cartes aux attributs/NP...**:

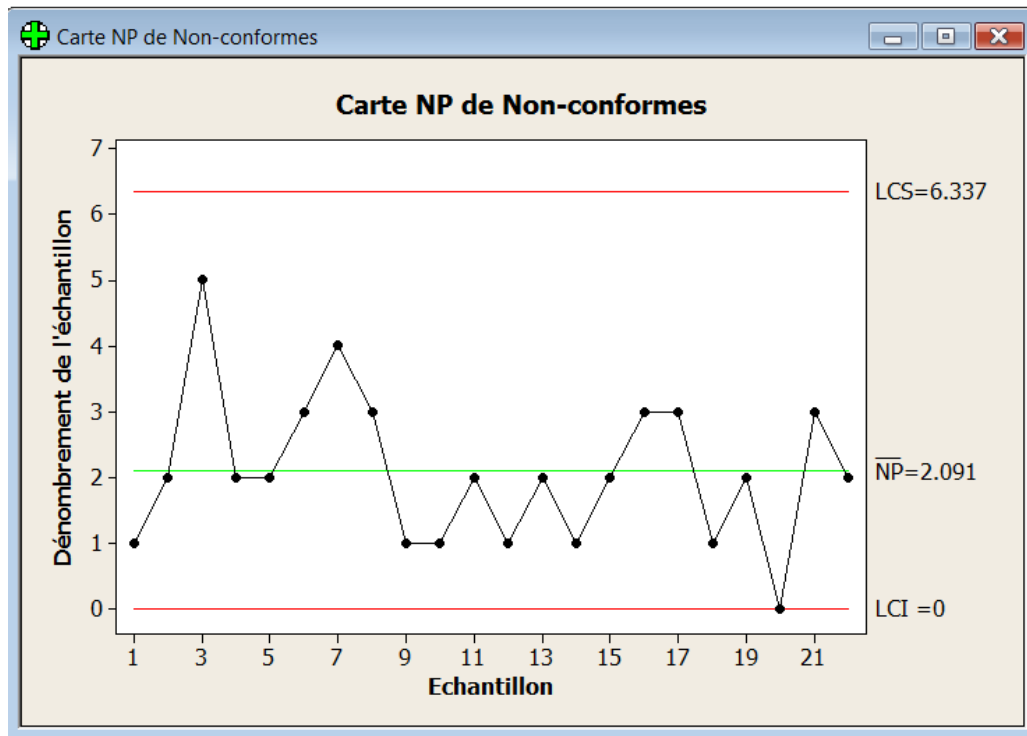


Bienvenue dans Minitab, appuyez sur F1 pour obtenir de l'aide.

Nous avons alors la boîte de dialogue suivante:



Nous validons par **OK**:



Nous obtenons donc les mêmes valeurs et la même carte que dans le cours de statistique théorique et Microsoft Excel.

## 14.3. Exercice 121.: CC par attributs de type C

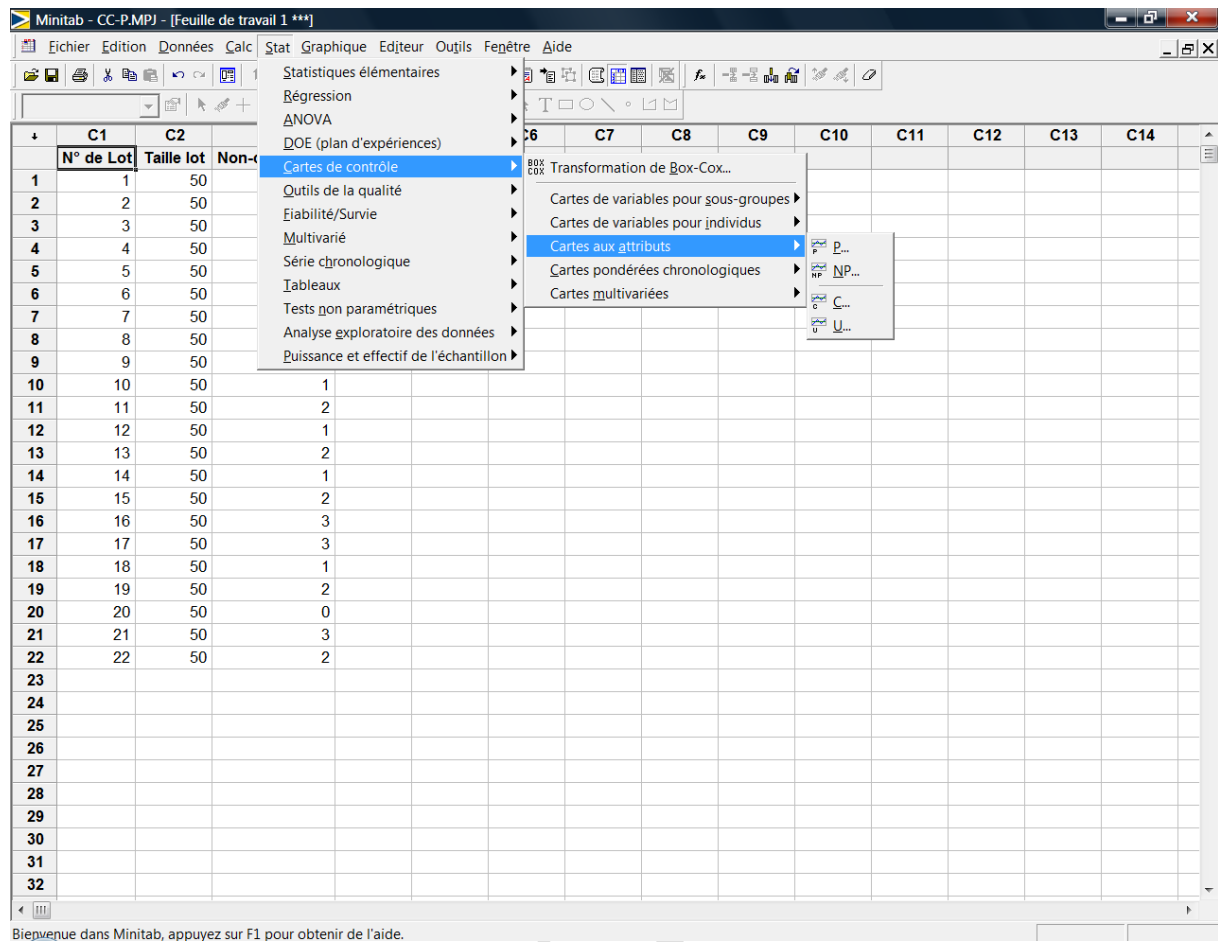
Minitab® Statistical Software 15.1.1.0

Dans le cadre de contrôle de lots nous avons obtenu les mesures suivantes (fichier *CC-C.mpj*):

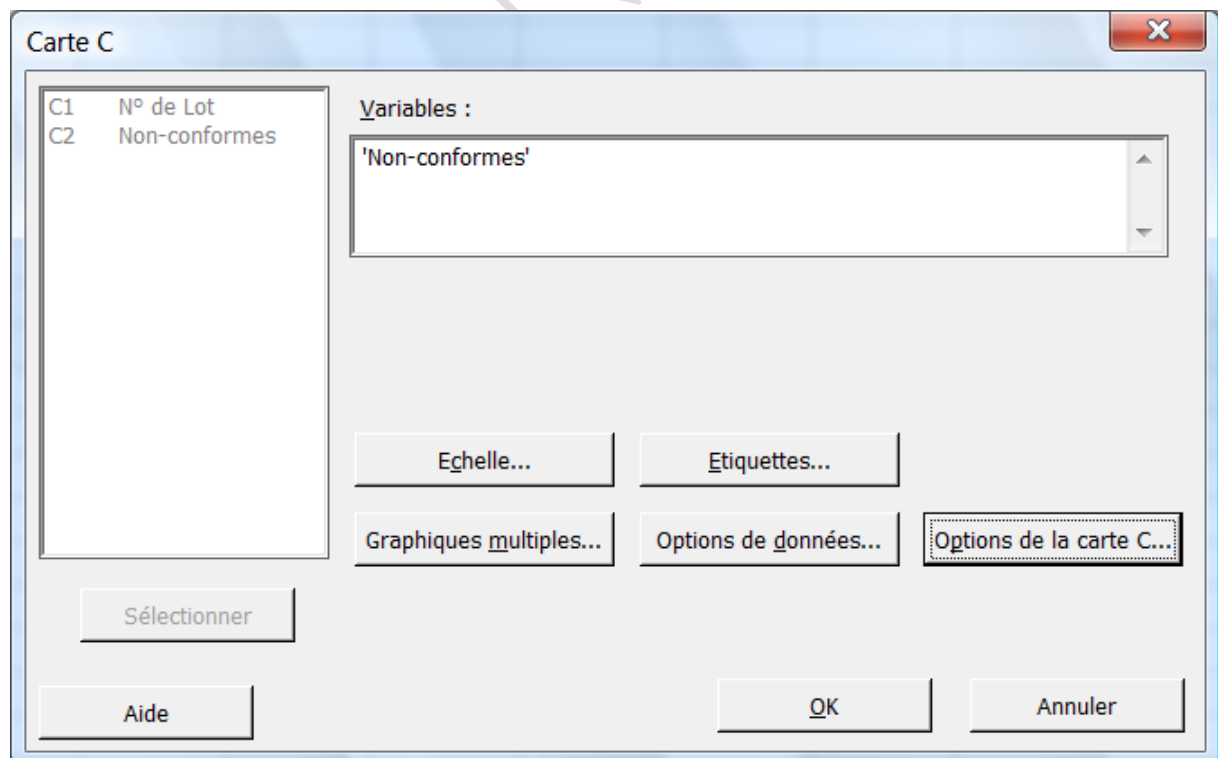
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	N° de Lot	Non-conformes												
1	1	1												
2	2	2												
3	3	5												
4	4	2												
5	5	2												
6	6	3												
7	7	4												
8	8	3												
9	9	1												
10	10	1												
11	11	2												
12	12	1												
13	13	2												
14	14	1												
15	15	2												
16	16	3												
17	17	3												
18	18	1												
19	19	2												
20	20	0												
21	21	3												
22	22	2												
23														
24														
25														
26														
27														
28														
29														
30														
31														
32														

Nous souhaiterions tracer une carte de contrôle par attributs de type *c* et comparer avec les calculs théoriques effectués dans le cours de statistique théorique et dans le cours Microsoft Excel.

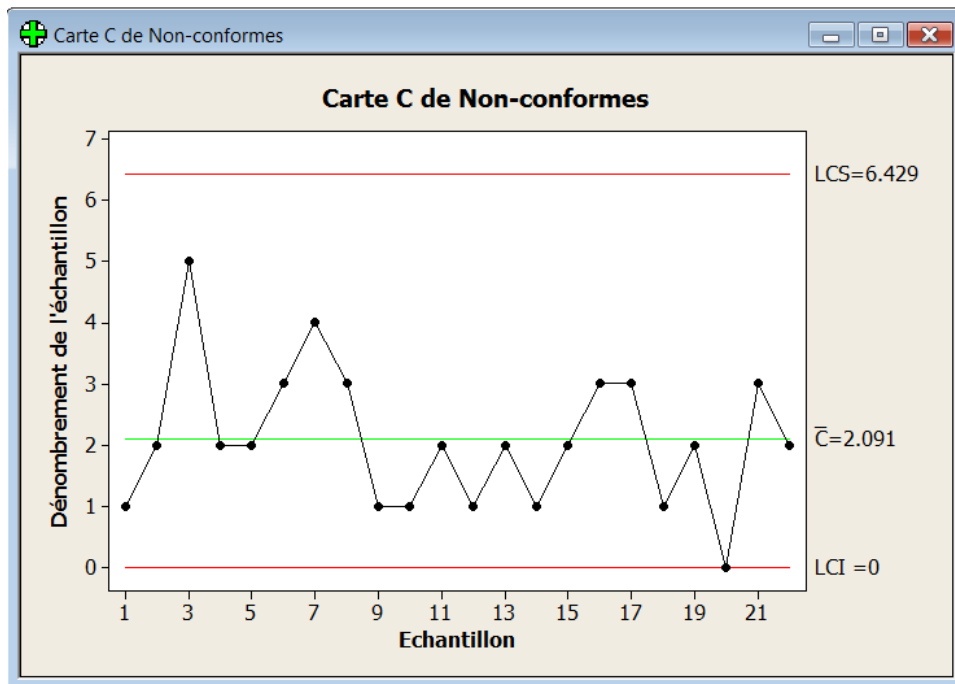
Dans Minitab® Statistical Software nous allons dans le menu **Stat/Cartes de contrôle/Cartes aux attributs/C...**:



Nous avons alors la boîte de dialogue suivante:



Nous validons par **OK**:

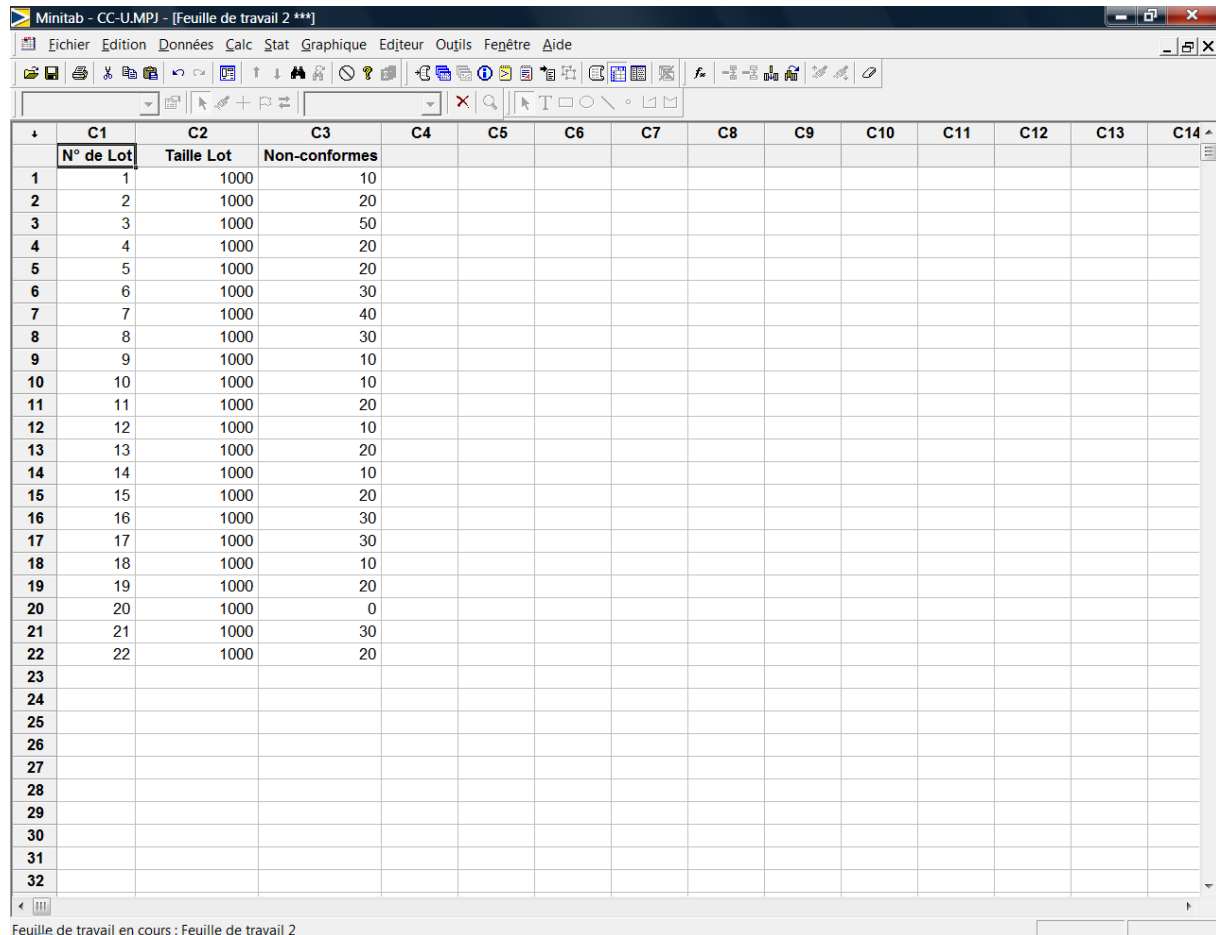


Nous obtenons donc les mêmes valeurs et la même carte que dans le cours de statistique théorique et Microsoft Excel.

## 14.4. Exercice 122.: CC par attributs de type U

Minitab® Statistical Software 15.1.1.0

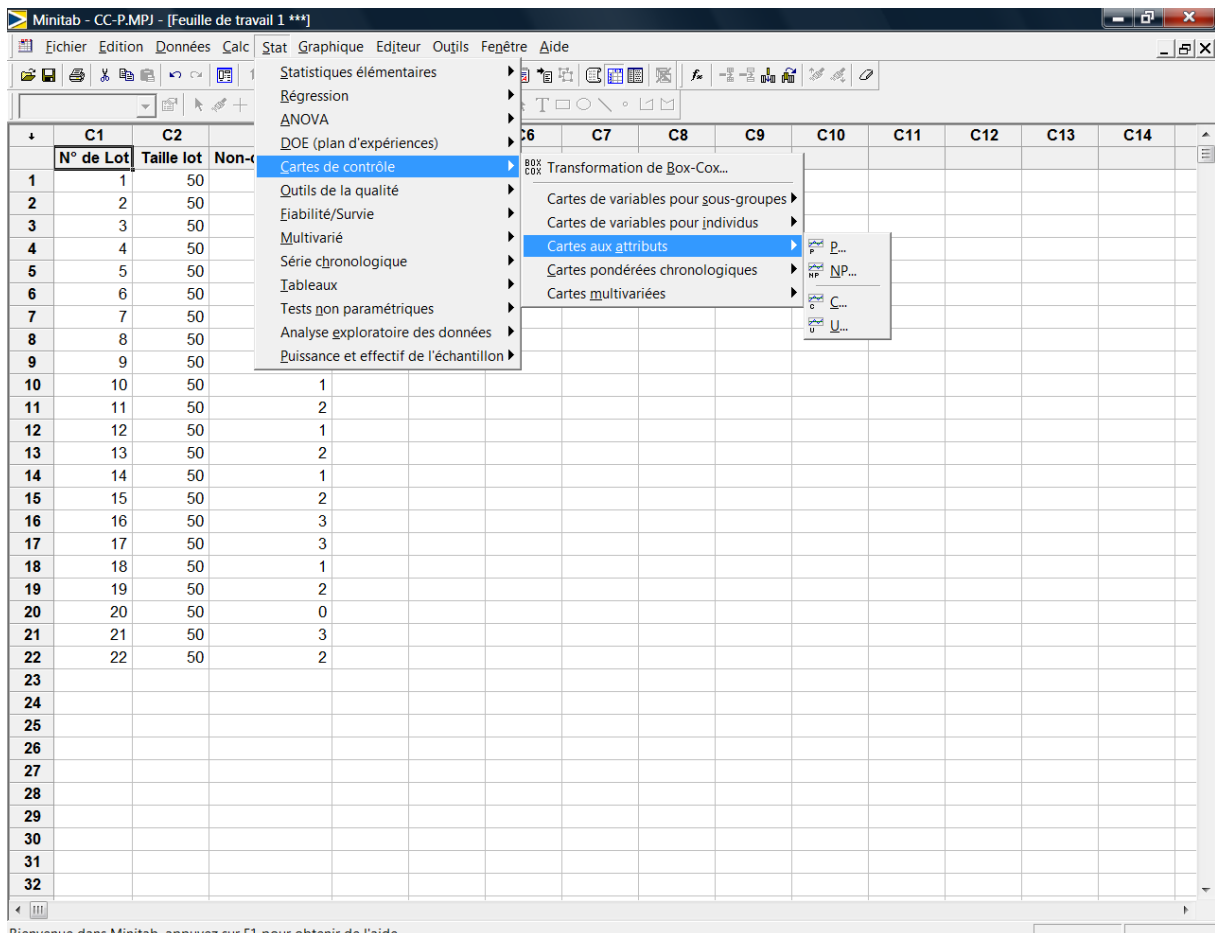
Dans le cadre de contrôle de lots nous avons obtenu les mesures suivantes (fichier *CC-U.mpj*):



	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	N° de Lot	Taille Lot	Non-conformes											
1	1	1000	10											
2	2	1000	20											
3	3	1000	50											
4	4	1000	20											
5	5	1000	20											
6	6	1000	30											
7	7	1000	40											
8	8	1000	30											
9	9	1000	10											
10	10	1000	10											
11	11	1000	20											
12	12	1000	10											
13	13	1000	20											
14	14	1000	10											
15	15	1000	20											
16	16	1000	30											
17	17	1000	30											
18	18	1000	10											
19	19	1000	20											
20	20	1000	0											
21	21	1000	30											
22	22	1000	20											
23														
24														
25														
26														
27														
28														
29														
30														
31														
32														

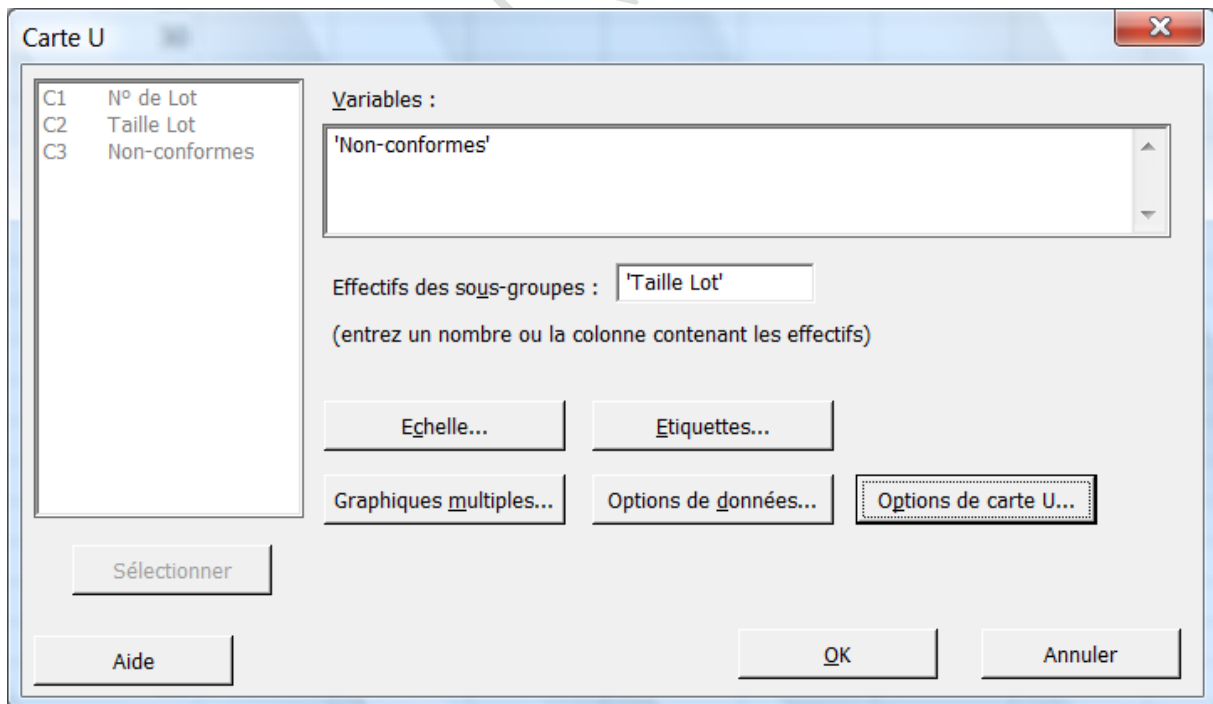
Nous souhaiterions tracer une carte de contrôle par attributs de type  $u$  et comparer avec les calculs théoriques effectués dans le cours de statistique théorique et dans le cours Microsoft Excel.

Dans Minitab® Statistical Software nous allons dans le menu **Stat/Cartes de contrôle/Cartes aux attributs/U...**:



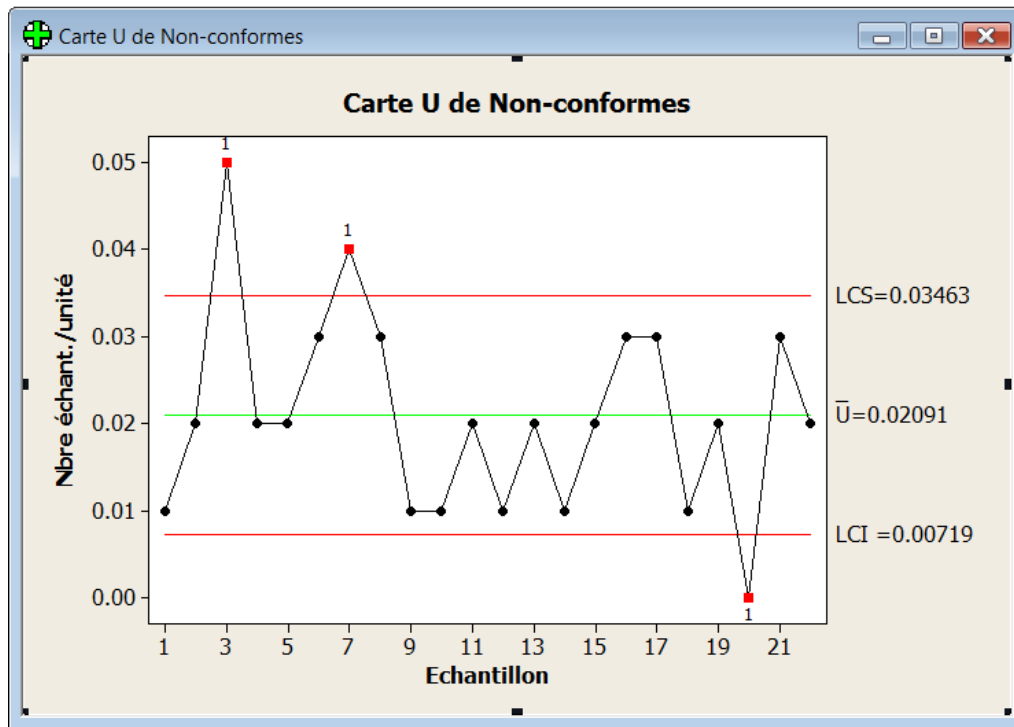
Bienvenue dans Minitab. appuyez sur F1 pour obtenir de l'aide.

Nous avons alors la boîte de dialogue suivante:



Nous validons par **OK**:

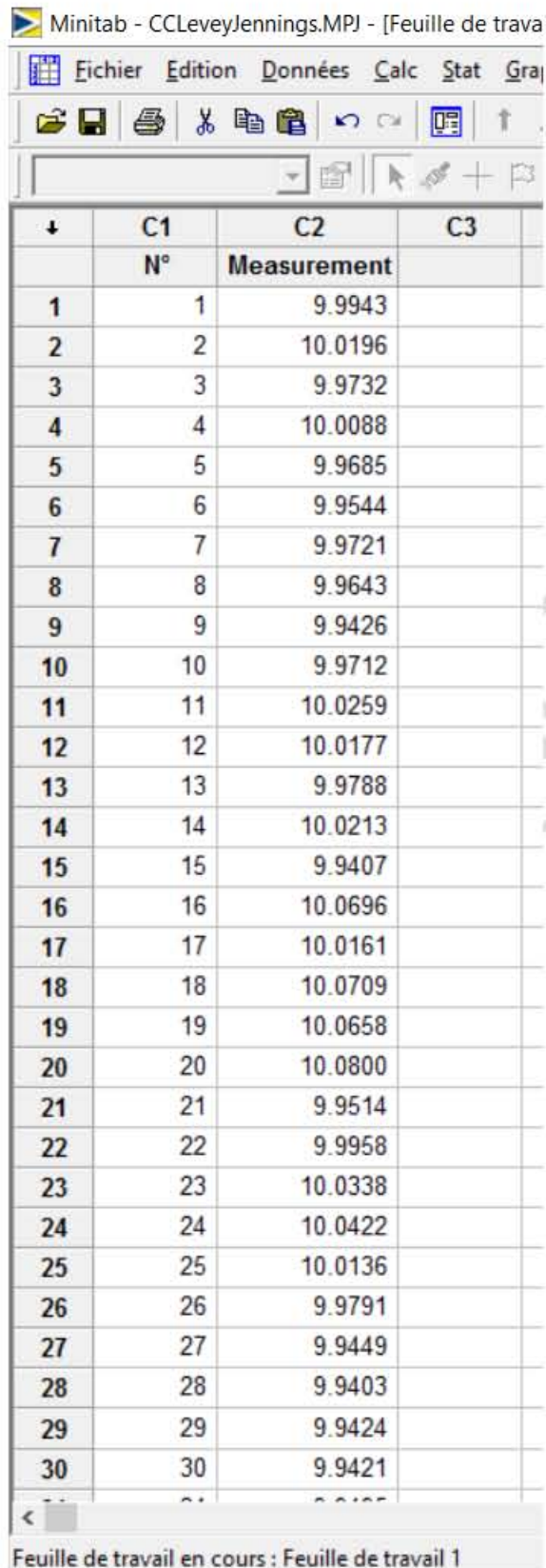




Nous obtenons donc les mêmes valeurs et la même carte que dans le cours de statistique théorique et Microsoft Excel.

## 14.5. Exercice 123.: CC par mesures de Levey-Jennings

Minitab® Statistical Software 15.1.1.0



Minitab - CCLeveyJennings.MPJ - [Feuille de trava

Fichier Edition Données Calc Stat Gra

	C1	C2	C3
	N°	Measurement	
1	1	9.9943	
2	2	10.0196	
3	3	9.9732	
4	4	10.0088	
5	5	9.9685	
6	6	9.9544	
7	7	9.9721	
8	8	9.9643	
9	9	9.9426	
10	10	9.9712	
11	11	10.0259	
12	12	10.0177	
13	13	9.9788	
14	14	10.0213	
15	15	9.9407	
16	16	10.0696	
17	17	10.0161	
18	18	10.0709	
19	19	10.0658	
20	20	10.0800	
21	21	9.9514	
22	22	9.9958	
23	23	10.0338	
24	24	10.0422	
25	25	10.0136	
26	26	9.9791	
27	27	9.9449	
28	28	9.9403	
29	29	9.9424	
30	30	9.9421	

Feuille de travail en cours : Feuille de travail 1

Minitab - CCLevyJennings.MPJ - [Feuille de travail 1 \*\*\*]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

Statistiques élémentaires

- Statistiques descriptives
  - Afficher les statistiques descriptives...
  - Stocker des statistiques descriptives...
  - Récapitulatif graphique...
- Régression
- ANOVA
- DOE (plan d'expériences)
- Cartes de contrôle
- Outils de la qualité
- Fiabilité/Survie
- Multivarié
- Série chronologique
- Tableaux
- Tests non paramétriques
- Analyse exploratoire des données
- Puissance et effectif de l'échantillon

	C1	C2
	N°	Measurement
1	1	9.9943
2	2	10.0196
3	3	9.9732
4	4	10.0088
5	5	9.9685
6	6	9.9544
7	7	9.9721
8	8	9.9643
9	9	9.9426
10	10	9.9712
11	11	10.0259
12	12	10.0177
13	13	9.9788
14	14	10.0213
15	15	9.9407
16	16	10.0696

- 1Z Test Z à 1 échantillon...
- 1t Test t à 1 échantillon...
- 2t Test t à 2 échantillons...
- t-t Test t pour données appariées...
- 1P 1 proportion...
- 2P 2 proportions...
- 1SP Test de Poisson à 1 échantillon...
- 2SP Test de Poisson à 2 échantillons...
- σ<sup>2</sup> 1 variance...
- σ<sub>1</sub><sup>2</sup> σ<sub>2</sub><sup>2</sup> 2 variances...
- CDR Corrélation...
- COV Covariance...
- TEST Test de normalité...
- χ<sup>2</sup> Test d'adéquation de l'ajustement pour Poisson...

Stocker des statistiques descriptives

C1	N°
C2	Measurement

Variables :

Measurement

Variables de répartition (facultatif) :

Sélectionner

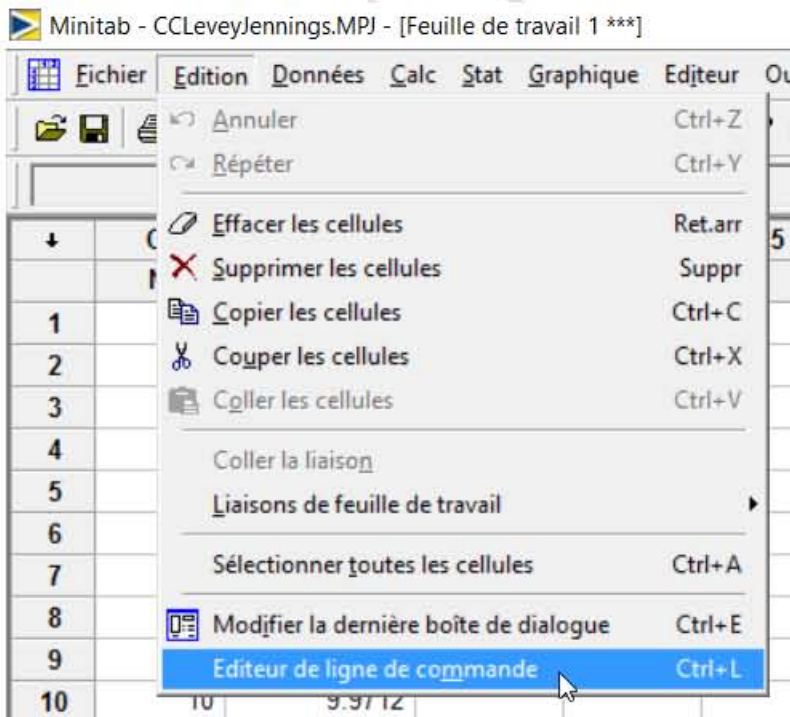
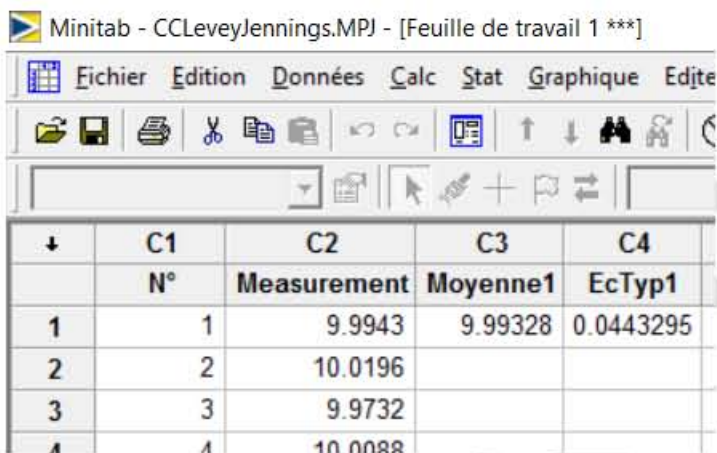
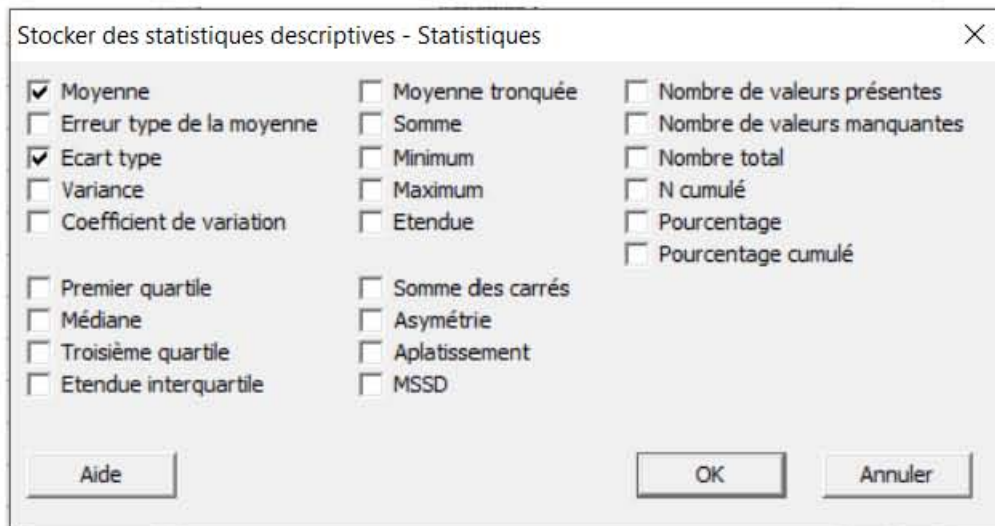
Statistiques...

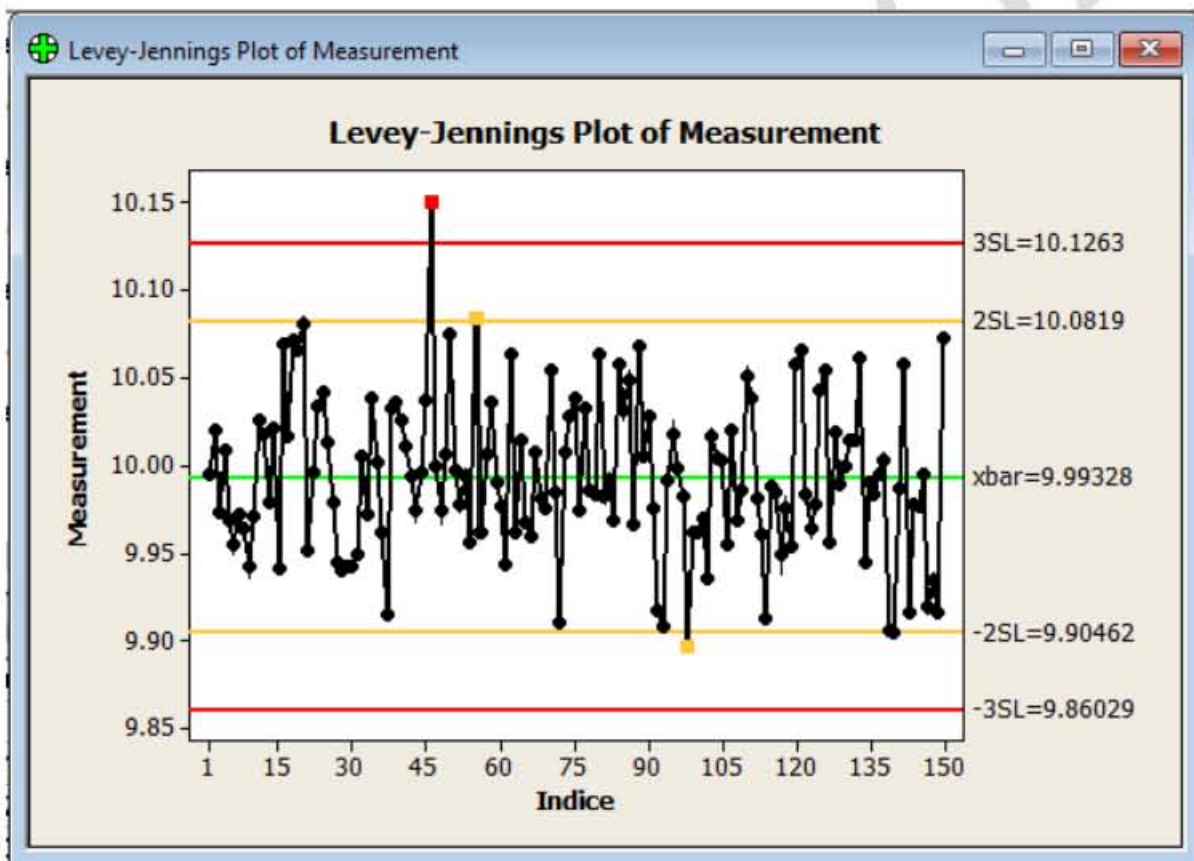
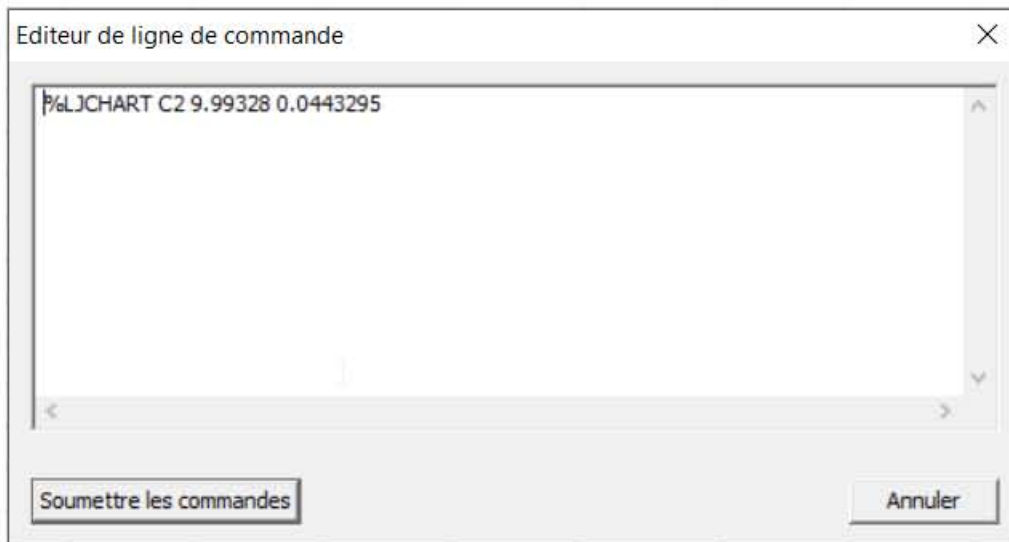
Options...

Aide

OK

Annuler





## 14.6. Exercice 124.: CC par mesures S (S barre-S)

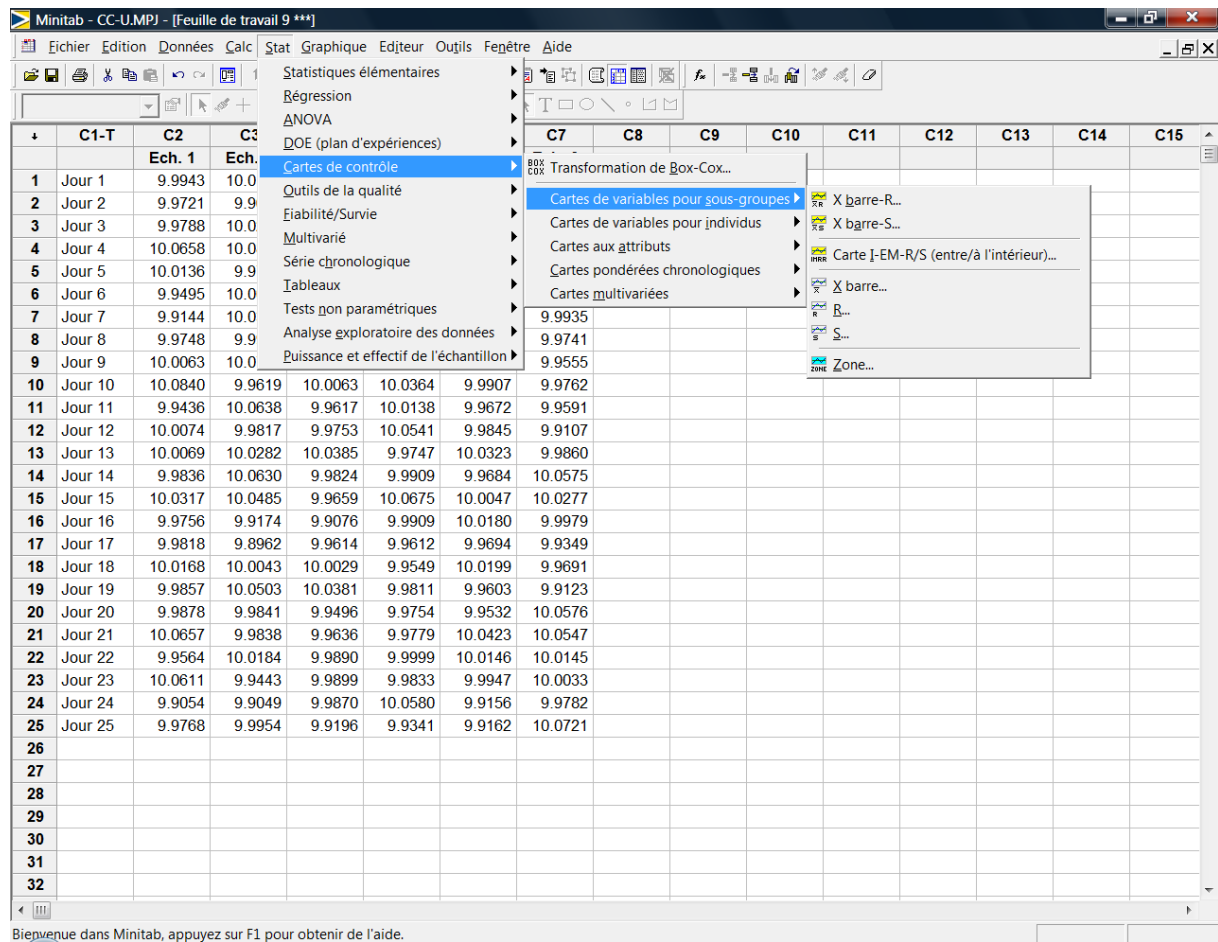
Minitab® Statistical Software 15.1.1.0

Dans le cadre de contrôle d'une production sous contrôle statistique nous avons obtenu les mesures suivantes (fichier *CC-Sbarre.mpj*):

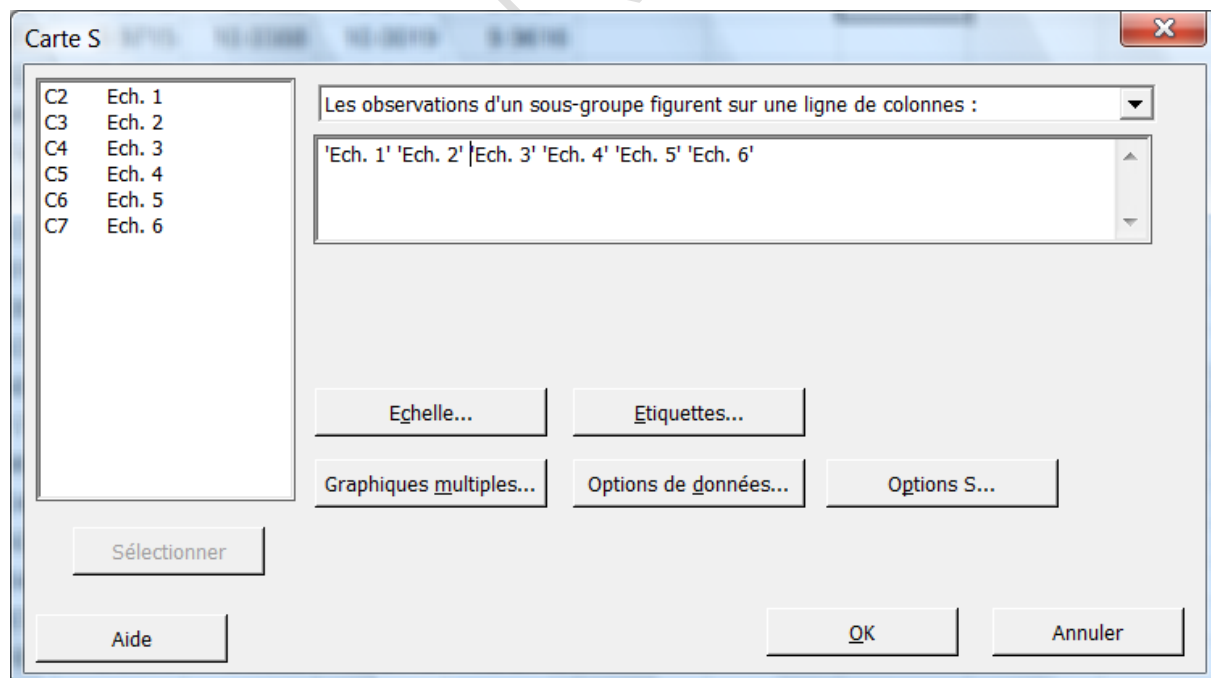
	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
		Ech. 1	Ech. 2	Ech. 3	Ech. 4	Ech. 5	Ech. 6								
1	Jour 1	9.9943	10.0196	9.9732	10.0088	9.9685	9.9544								
2	Jour 2	9.9721	9.9643	9.9426	9.9712	10.0259	10.0177								
3	Jour 3	9.9788	10.0213	9.9407	10.0696	10.0161	10.0709								
4	Jour 4	10.0658	10.0800	9.9514	9.9958	10.0338	10.0422								
5	Jour 5	10.0136	9.9791	9.9449	9.9403	9.9424	9.9421								
6	Jour 6	9.9495	10.0057	9.9715	10.0388	10.0019	9.9616								
7	Jour 7	9.9144	10.0321	10.0361	10.0253	10.0109	9.9935								
8	Jour 8	9.9748	9.9962	10.0369	10.1500	9.9996	9.9741								
9	Jour 9	10.0063	10.0749	9.9971	9.9780	9.9934	9.9555								
10	Jour 10	10.0840	9.9619	10.0063	10.0364	9.9907	9.9762								
11	Jour 11	9.9436	10.0638	9.9617	10.0138	9.9672	9.9591								
12	Jour 12	10.0074	9.9817	9.9753	10.0541	9.9845	9.9107								
13	Jour 13	10.0069	10.0282	10.0385	9.9747	10.0323	9.9860								
14	Jour 14	9.9836	10.0630	9.9824	9.9909	9.9684	10.0575								
15	Jour 15	10.0317	10.0485	9.9659	10.0675	10.0047	10.0277								
16	Jour 16	9.9756	9.9174	9.9076	9.9909	10.0180	9.9979								
17	Jour 17	9.9818	9.8962	9.9614	9.9612	9.9694	9.9349								
18	Jour 18	10.0168	10.0043	10.0029	9.9549	10.0199	9.9691								
19	Jour 19	9.9857	10.0503	10.0381	9.9811	9.9603	9.9123								
20	Jour 20	9.9878	9.9841	9.9496	9.9754	9.9532	10.0576								
21	Jour 21	10.0657	9.9838	9.9636	9.9779	10.0423	10.0547								
22	Jour 22	9.9564	10.0184	9.9890	9.9999	10.0146	10.0145								
23	Jour 23	10.0611	9.9443	9.9899	9.9833	9.9947	10.0033								
24	Jour 24	9.9054	9.9049	9.9870	10.0580	9.9156	9.9782								
25	Jour 25	9.9768	9.9954	9.9196	9.9341	9.9162	10.0721								
26															
27															
28															
29															
30															
31															
32															

Nous souhaiterions tracer une carte de contrôle par mesures groupées de type S et comparer avec les calculs théoriques effectués dans le cours de statistique théorique et dans le cours Microsoft Excel.

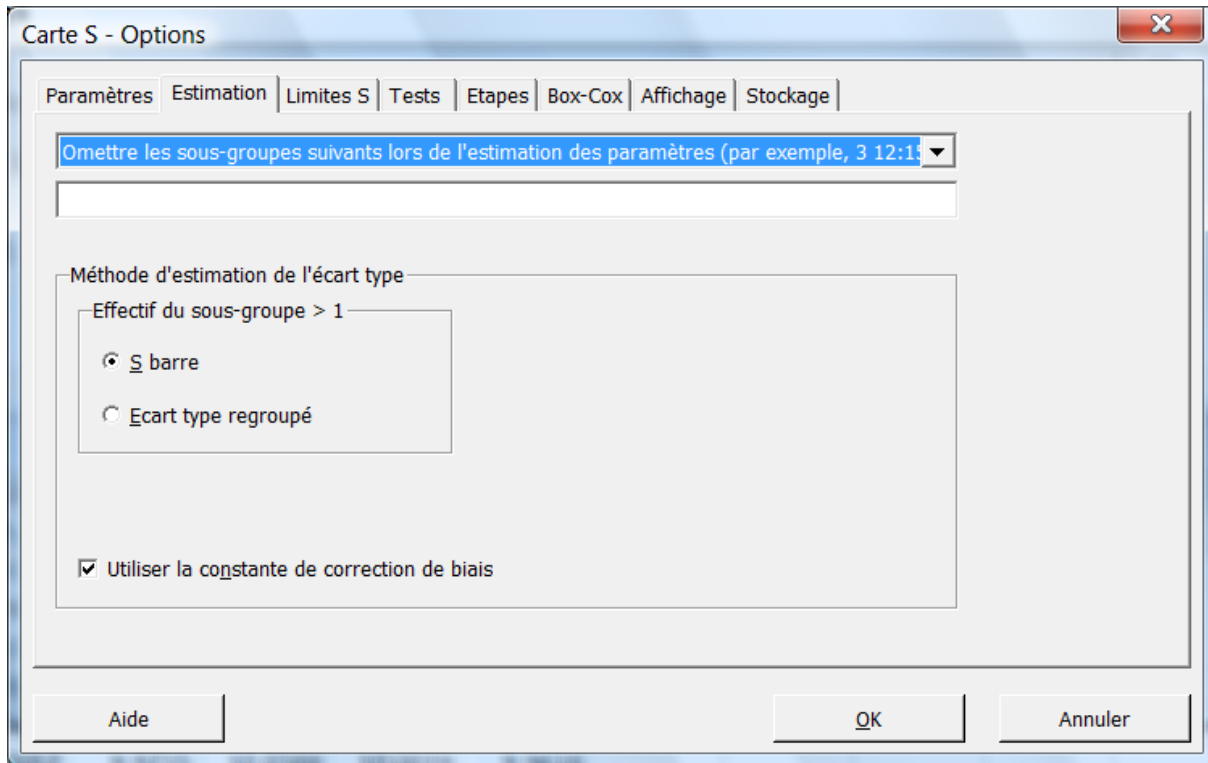
Dans Minitab® Statistical Software nous allons dans le menu **Stat/Cartes de contrôle/Cartes de variables pour sous-groupes/S...**:



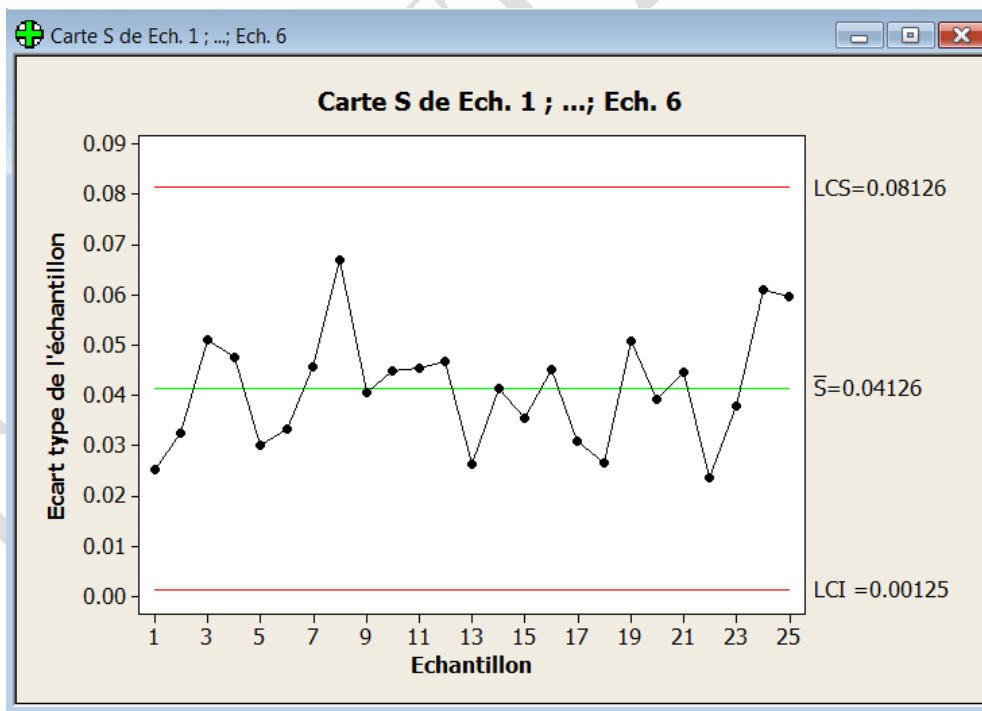
Nous avons alors la boîte de dialogue suivante:



Nous cliquons sur **Options S...**:



pour cocher l'option **S barre** car c'est un peu le but quand même... (curieux que ce ne soit pas coché par défaut dans Minitab® Statistical Software...). Nous validons deux fois par **OK**:



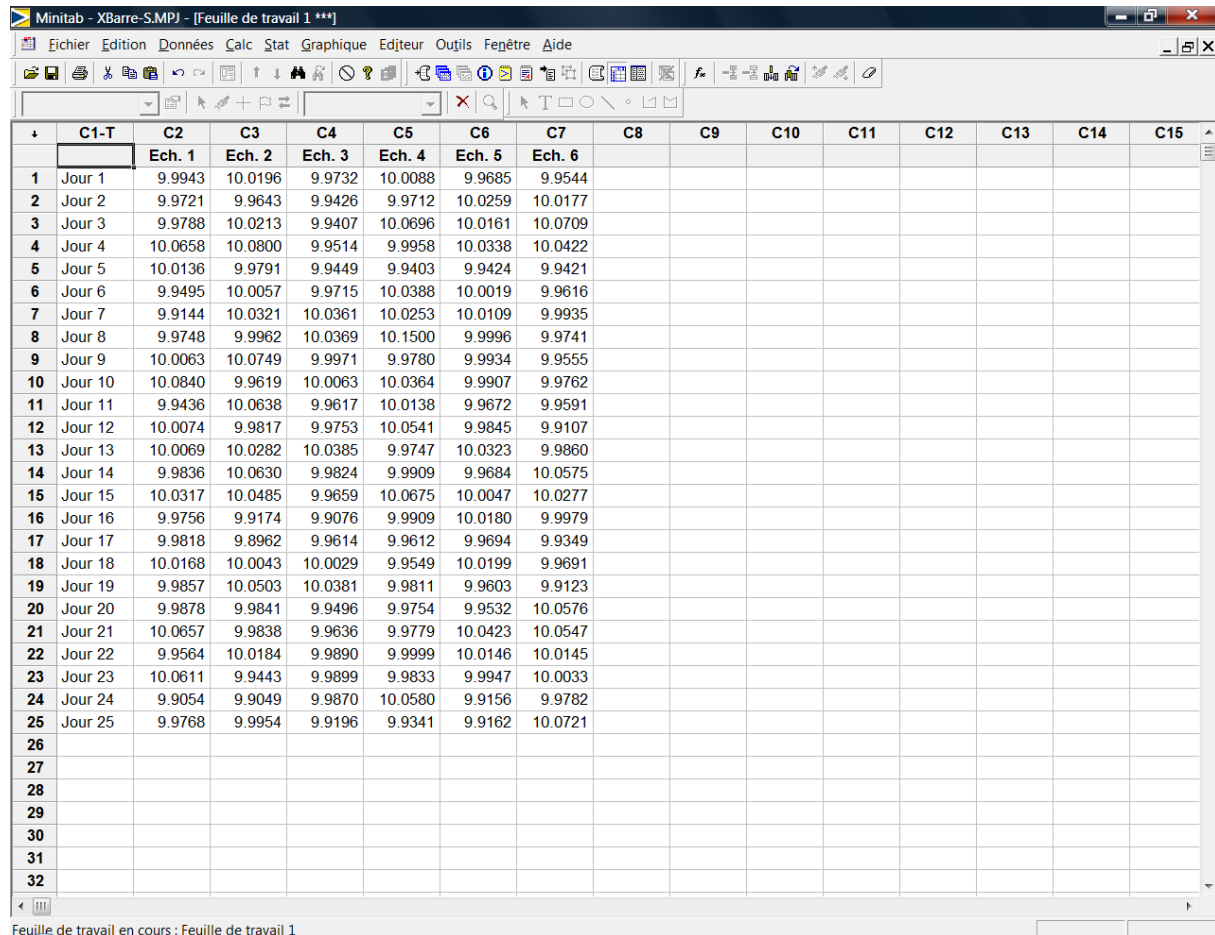
Nous obtenons donc les mêmes valeurs et la même carte que dans le cours de statistique théorique et Microsoft Excel.



## 14.7. Exercice 125.: CC par mesures X barre-S simple

Minitab® Statistical Software 15.1.1.0

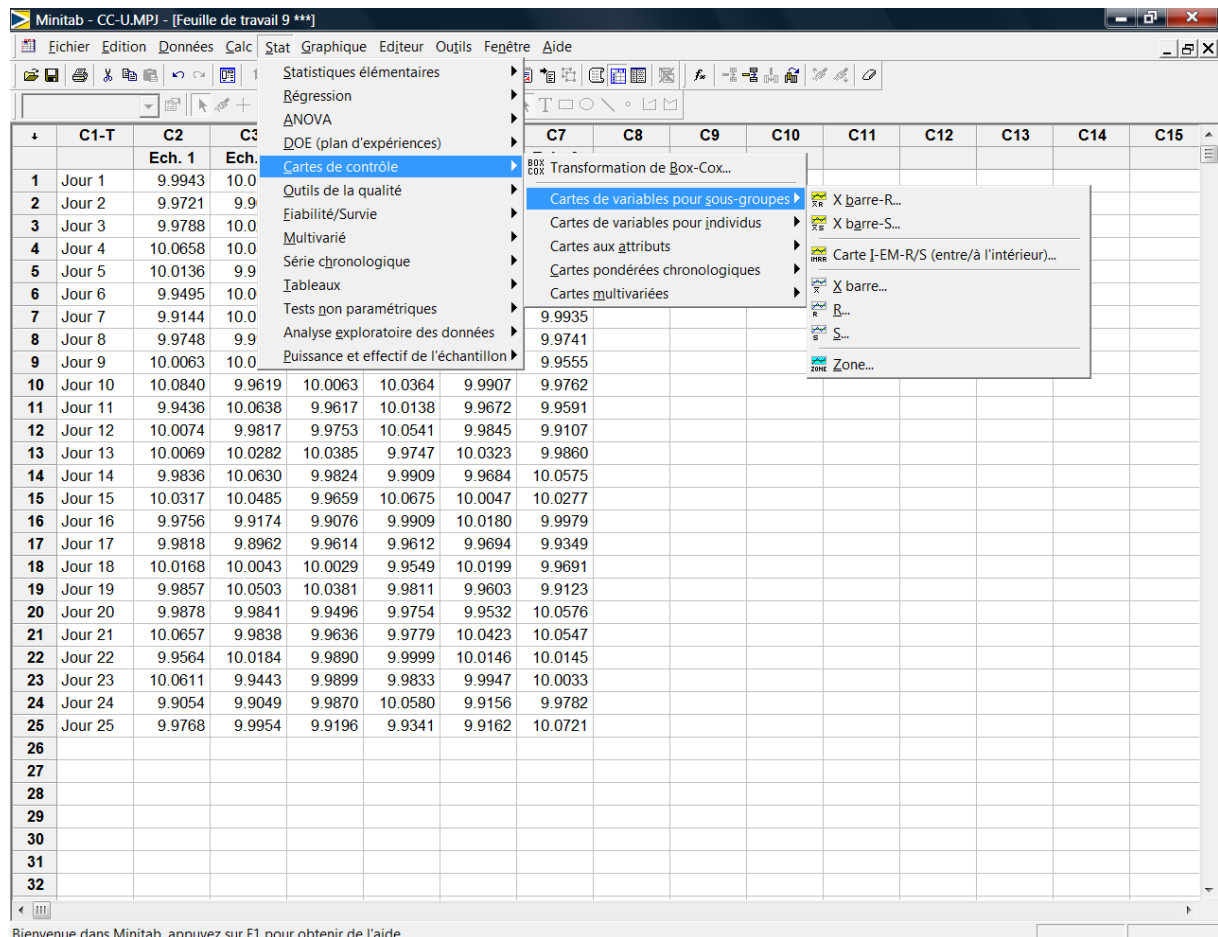
Dans le cadre de contrôle d'une production sous contrôle statistique nous avons obtenu les mesures suivantes (fichier *Xbarre-S.mpj*):



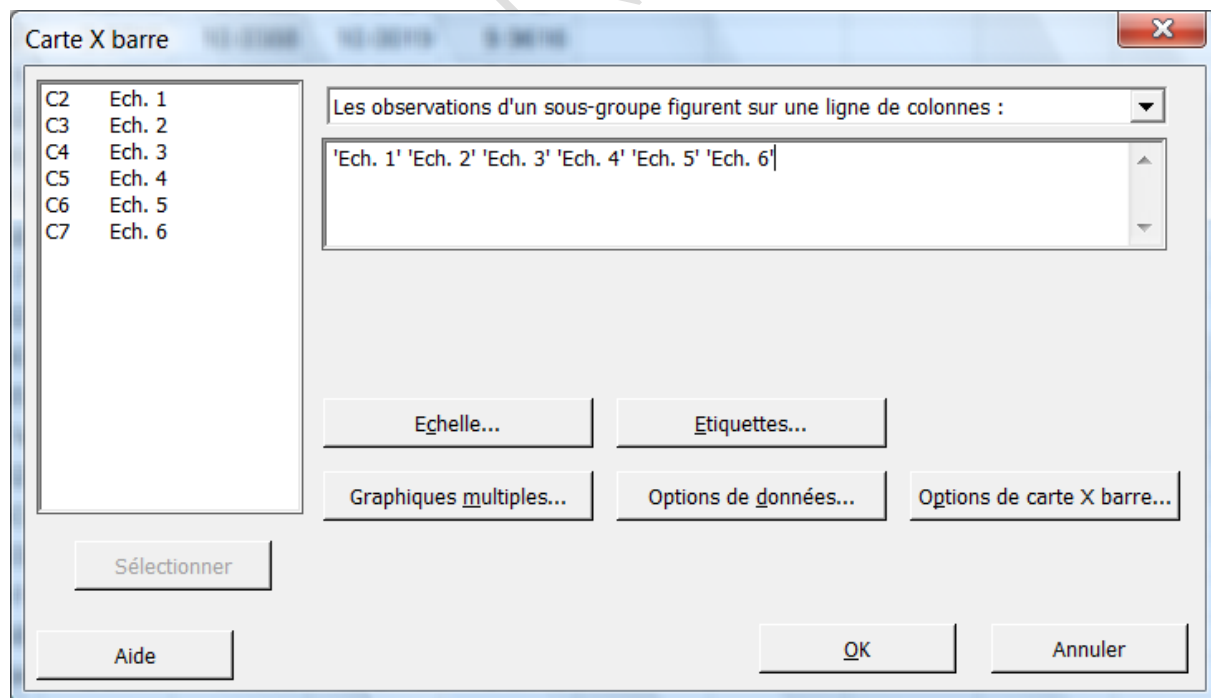
	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
		Ech. 1	Ech. 2	Ech. 3	Ech. 4	Ech. 5	Ech. 6								
1	Jour 1	9.9943	10.0196	9.9732	10.0088	9.9685	9.9544								
2	Jour 2	9.9721	9.9643	9.9426	9.9712	10.0259	10.0177								
3	Jour 3	9.9788	10.0213	9.9407	10.0696	10.0161	10.0709								
4	Jour 4	10.0658	10.0800	9.9514	9.9958	10.0338	10.0422								
5	Jour 5	10.0136	9.9791	9.9449	9.9403	9.9424	9.9421								
6	Jour 6	9.9495	10.0057	9.9715	10.0388	10.0019	9.9616								
7	Jour 7	9.9144	10.0321	10.0361	10.0253	10.0109	9.9935								
8	Jour 8	9.9748	9.9962	10.0369	10.1500	9.9996	9.9741								
9	Jour 9	10.0063	10.0749	9.9971	9.9780	9.9934	9.9555								
10	Jour 10	10.0840	9.9619	10.0063	10.0364	9.9907	9.9762								
11	Jour 11	9.9436	10.0638	9.9617	10.0138	9.9672	9.9591								
12	Jour 12	10.0074	9.9817	9.9753	10.0541	9.9845	9.9107								
13	Jour 13	10.0069	10.0282	10.0385	9.9747	10.0323	9.9860								
14	Jour 14	9.9836	10.0630	9.9824	9.9909	9.9684	10.0575								
15	Jour 15	10.0317	10.0485	9.9659	10.0675	10.0047	10.0277								
16	Jour 16	9.9756	9.9174	9.9076	9.9909	10.0180	9.9979								
17	Jour 17	9.9818	9.8962	9.9614	9.9612	9.9694	9.9349								
18	Jour 18	10.0168	10.0043	10.0029	9.9549	10.0199	9.9691								
19	Jour 19	9.9857	10.0503	10.0381	9.9811	9.9603	9.9123								
20	Jour 20	9.9878	9.9841	9.9496	9.9754	9.9532	10.0576								
21	Jour 21	10.0657	9.9838	9.9636	9.9779	10.0423	10.0547								
22	Jour 22	9.9564	10.0184	9.9890	9.9999	10.0146	10.0145								
23	Jour 23	10.0611	9.9443	9.9899	9.9833	9.9947	10.0033								
24	Jour 24	9.9054	9.9049	9.9870	10.0580	9.9156	9.9782								
25	Jour 25	9.9768	9.9954	9.9196	9.9341	9.9162	10.0721								
26															
27															
28															
29															
30															
31															
32															

Nous souhaiterions tracer une carte de contrôle par mesures groupées de type *X barre-S* et comparer avec les calculs théoriques effectués dans le cours de statistique théorique et dans le cours Microsoft Excel.

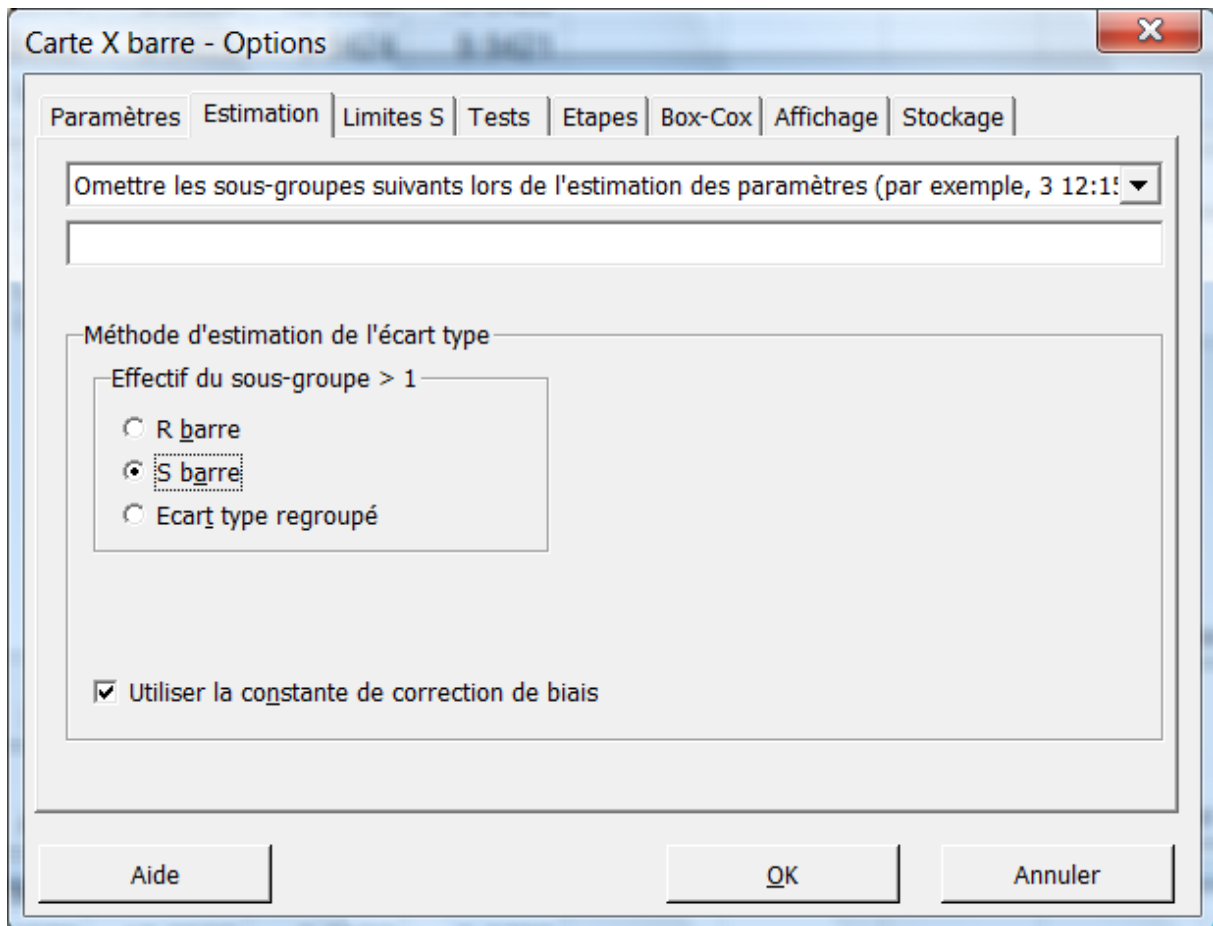
Dans Minitab® Statistical Software nous allons dans le menu **Stat/Cartes de contrôle/Cartes de variables pour sous-groupes/X barre...**:



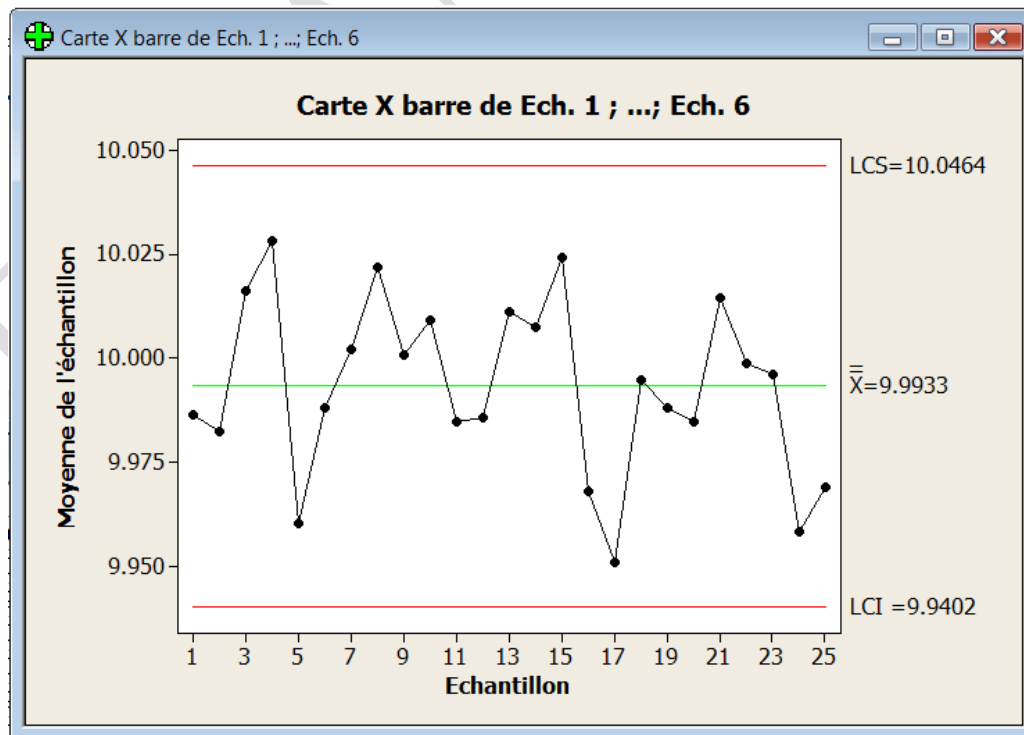
Nous avons alors la boîte de dialogue suivante:



Nous cliquons sur **Options de carte X barre...**:



pour cocher l'option **S barre** car c'est un peu le but quand même. Nous validons deux fois par **OK**:



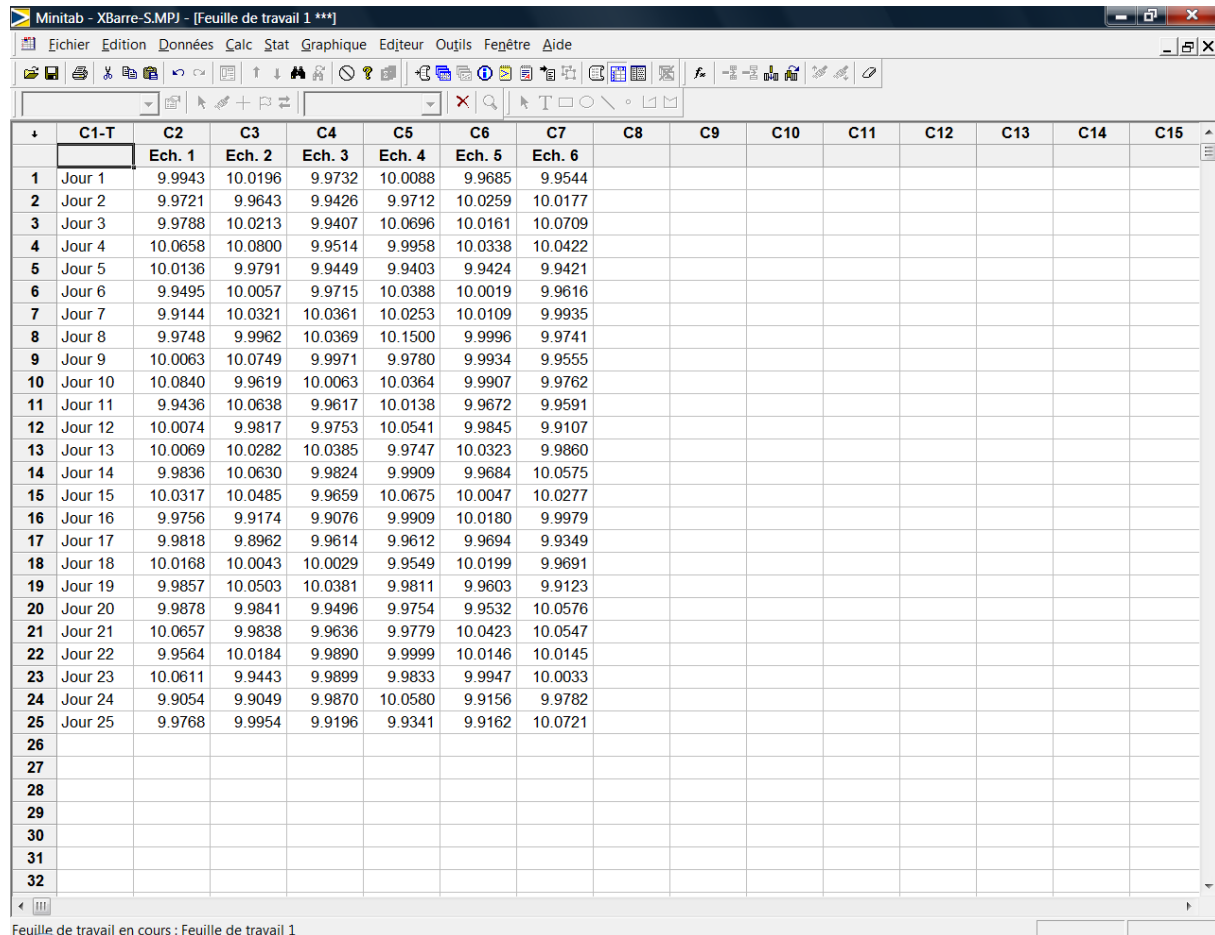
Nous obtenons donc les mêmes valeurs et la même carte que dans le cours de statistique théorique et Microsoft Excel.

ÉCHANTILLON

## 14.8. Exercice 126.: CC par mesures X barre-S barre combinées

Minitab® Statistical Software 15.1.1.0

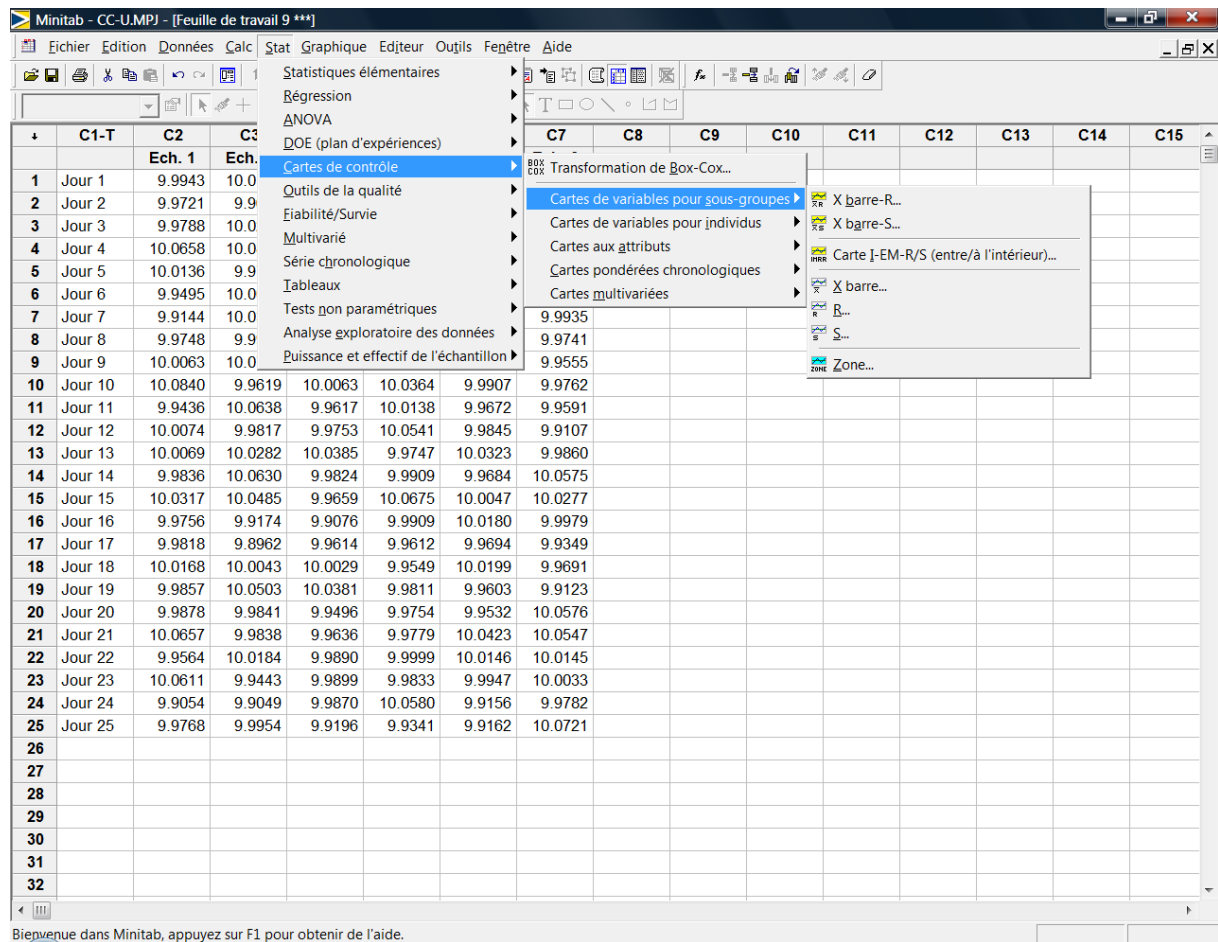
Dans le cadre de contrôle d'une production sous contrôle statistique nous avons obtenu les mesures suivantes (fichier CC Xbarre-S.mpj):



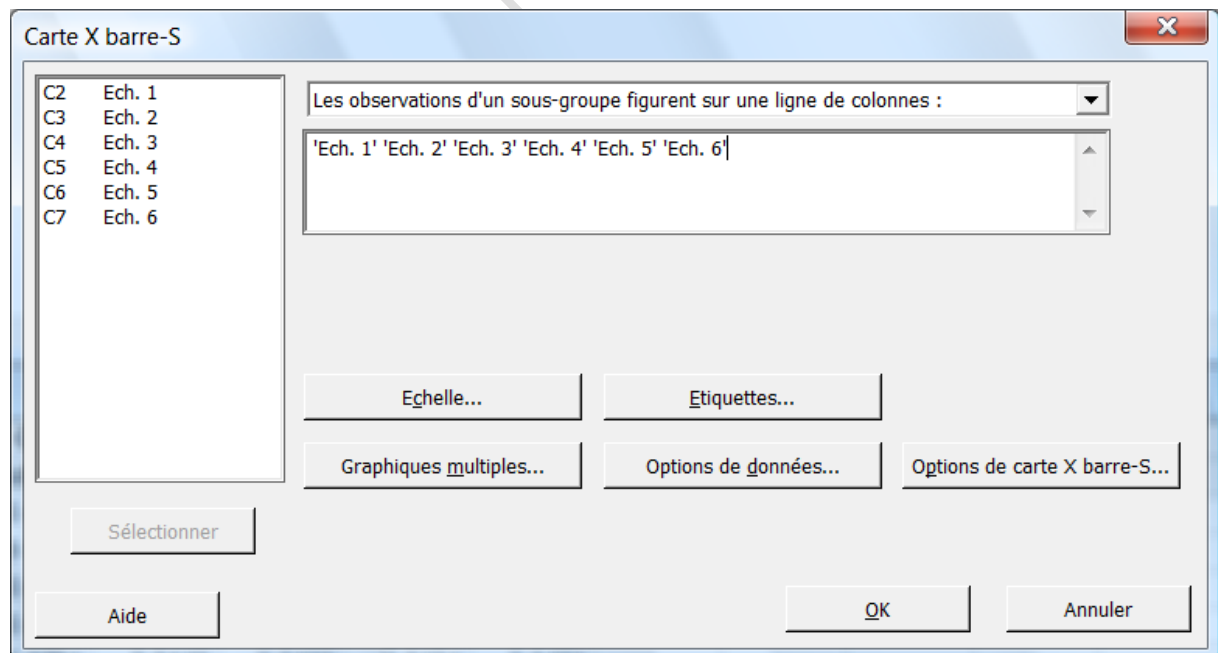
	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
		Ech. 1	Ech. 2	Ech. 3	Ech. 4	Ech. 5	Ech. 6								
1	Jour 1	9.9943	10.0196	9.9732	10.0088	9.9685	9.9544								
2	Jour 2	9.9721	9.9643	9.9426	9.9712	10.0259	10.0177								
3	Jour 3	9.9788	10.0213	9.9407	10.0696	10.0161	10.0709								
4	Jour 4	10.0658	10.0800	9.9514	9.9958	10.0338	10.0422								
5	Jour 5	10.0136	9.9791	9.9449	9.9403	9.9424	9.9421								
6	Jour 6	9.9495	10.0057	9.9715	10.0388	10.0019	9.9616								
7	Jour 7	9.9144	10.0321	10.0361	10.0253	10.0109	9.9935								
8	Jour 8	9.9748	9.9962	10.0369	10.1500	9.9996	9.9741								
9	Jour 9	10.0063	10.0749	9.9971	9.9780	9.9934	9.9555								
10	Jour 10	10.0840	9.9619	10.0063	10.0364	9.9907	9.9762								
11	Jour 11	9.9436	10.0638	9.9617	10.0138	9.9672	9.9591								
12	Jour 12	10.0074	9.9817	9.9753	10.0541	9.9845	9.9107								
13	Jour 13	10.0069	10.0282	10.0385	9.9747	10.0323	9.9860								
14	Jour 14	9.9836	10.0630	9.9824	9.9909	9.9684	10.0575								
15	Jour 15	10.0317	10.0485	9.9659	10.0675	10.0047	10.0277								
16	Jour 16	9.9756	9.9174	9.9076	9.9909	10.0180	9.9979								
17	Jour 17	9.9818	9.8962	9.9614	9.9612	9.9694	9.9349								
18	Jour 18	10.0168	10.0043	10.0029	9.9549	10.0199	9.9691								
19	Jour 19	9.9857	10.0503	10.0381	9.9811	9.9603	9.9123								
20	Jour 20	9.9878	9.9841	9.9496	9.9754	9.9532	10.0576								
21	Jour 21	10.0657	9.9838	9.9636	9.9779	10.0423	10.0547								
22	Jour 22	9.9564	10.0184	9.9890	9.9999	10.0146	10.0145								
23	Jour 23	10.0611	9.9443	9.9899	9.9833	9.9947	10.0033								
24	Jour 24	9.9054	9.9049	9.9870	10.0580	9.9156	9.9782								
25	Jour 25	9.9768	9.9954	9.9196	9.9341	9.9162	10.0721								
26															
27															
28															
29															
30															
31															
32															

Nous souhaiterions tracer une carte de contrôle par mesures groupées de type *X barre* combinée à une *S barre* (soit les deux exercices précédents en un seul coup) et comparer avec les calculs théoriques effectués dans le cours de statistique théorique et dans le cours Microsoft Excel.

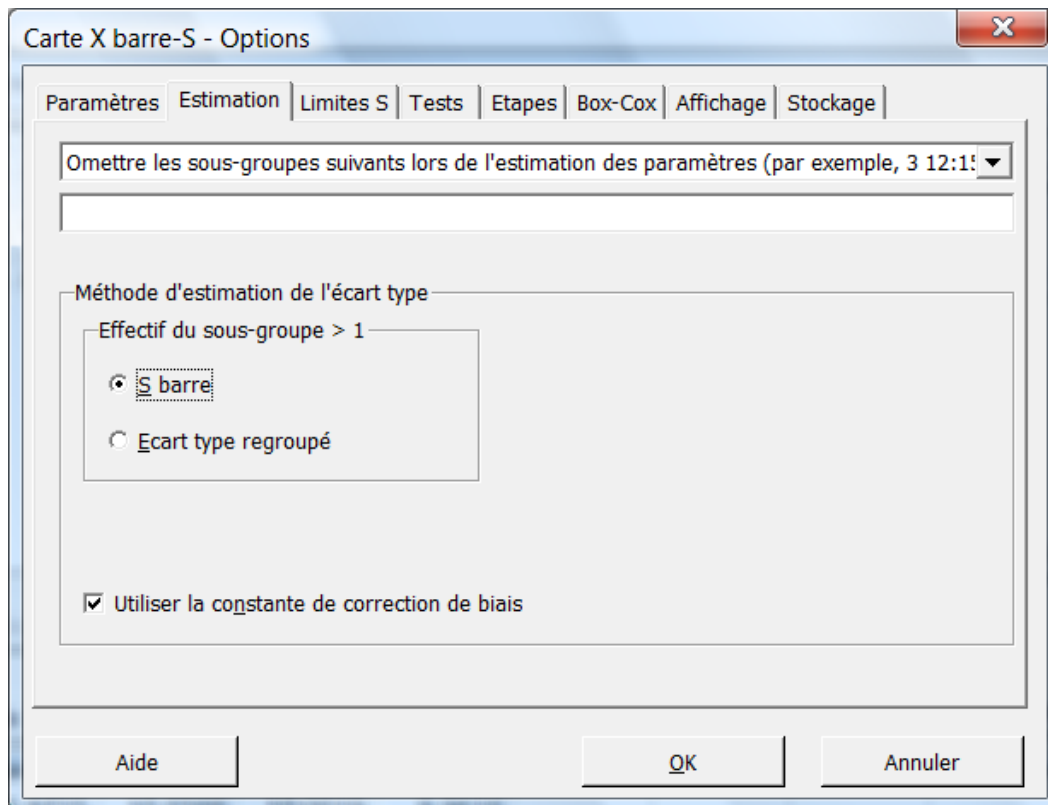
Dans Minitab® Statistical Software nous allons dans le menu **Stat/Cartes de contrôle/Cartes de variables pour sous-groupes/X barre-S...**:



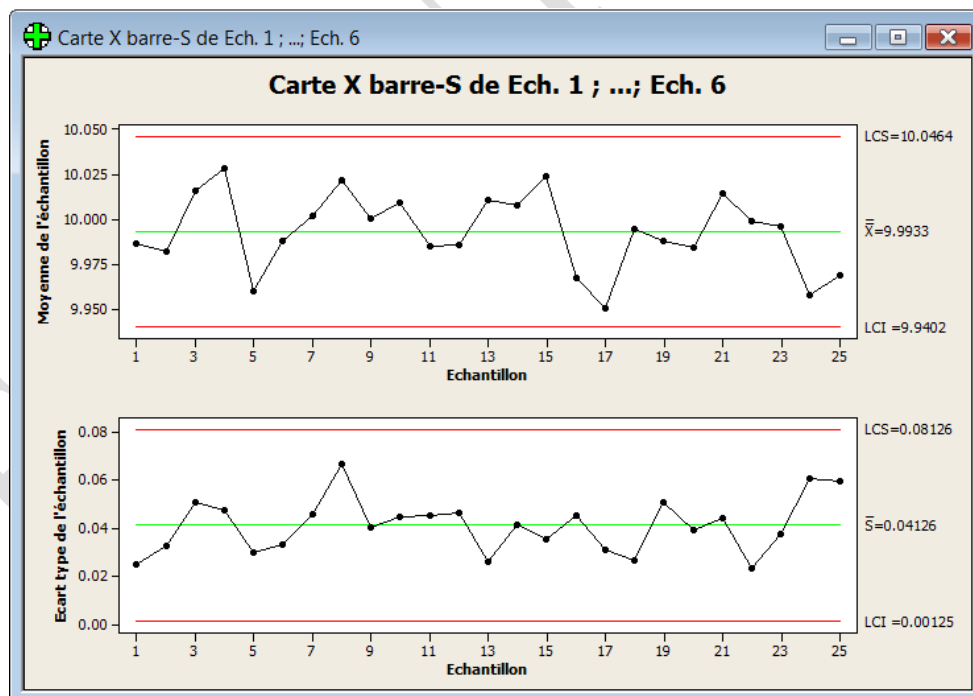
Nous avons alors la boîte de dialogue suivante:



Nous cliquons sur **Options de carte X barre-S...**:



pour cocher l'option **S barre** car c'est un peu le but quand même. Nous validons deux fois par **OK**:

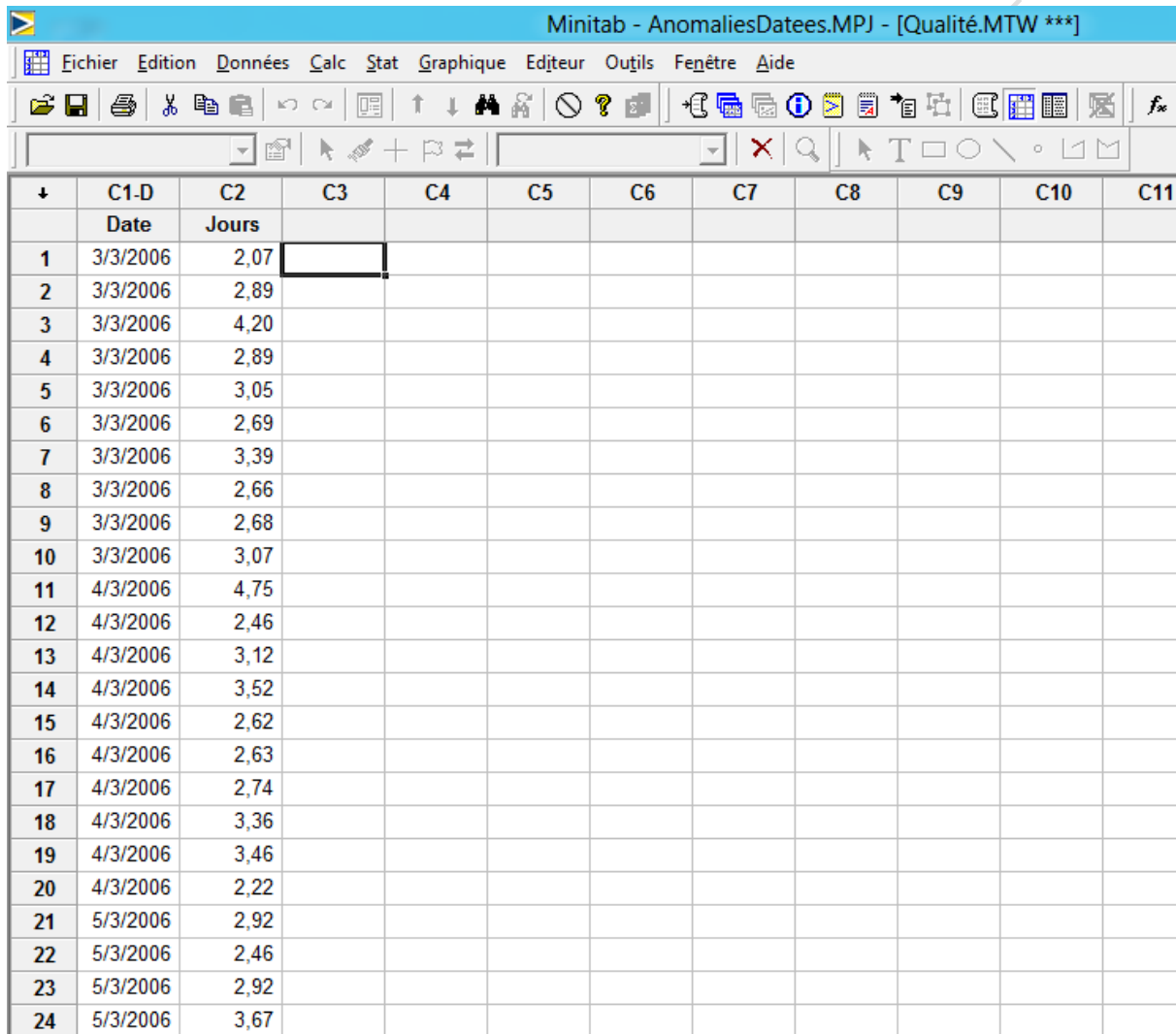


et donc nous retrouvons bien les graphiques des deux exercices précédents qui correspondent donc toujours à ce qui a été calculé dans le cours de statistique théorique et Microsoft Excel.

## 14.9. Exercice 127.: CC par mesures X barre-S barre combinées groupées par dates

Minitab® Statistical Software 15.1.1.0

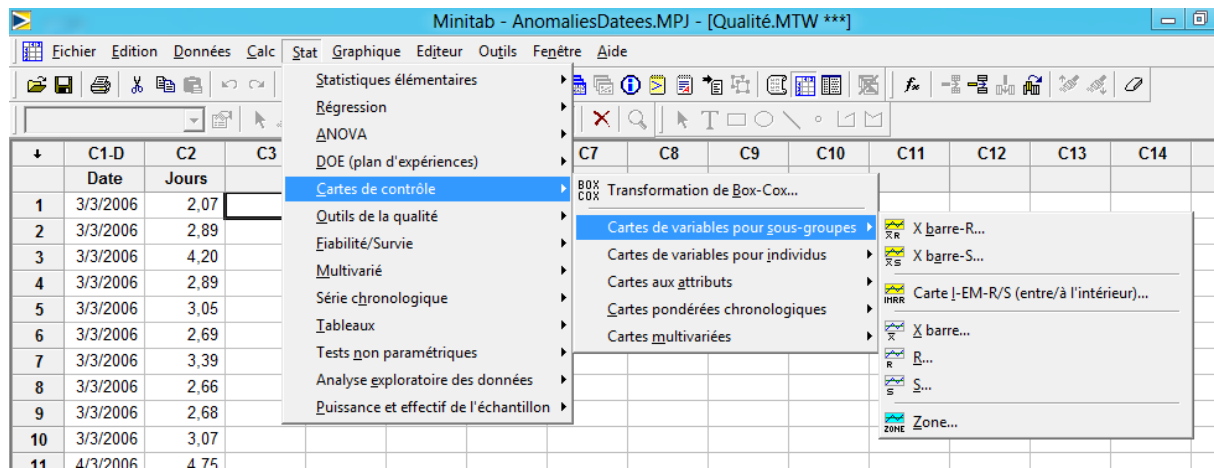
Dans le cadre de contrôle d'une production sous contrôle statistique nous avons obtenu les données suivantes (fichier AnomaliesDatees.mpj):



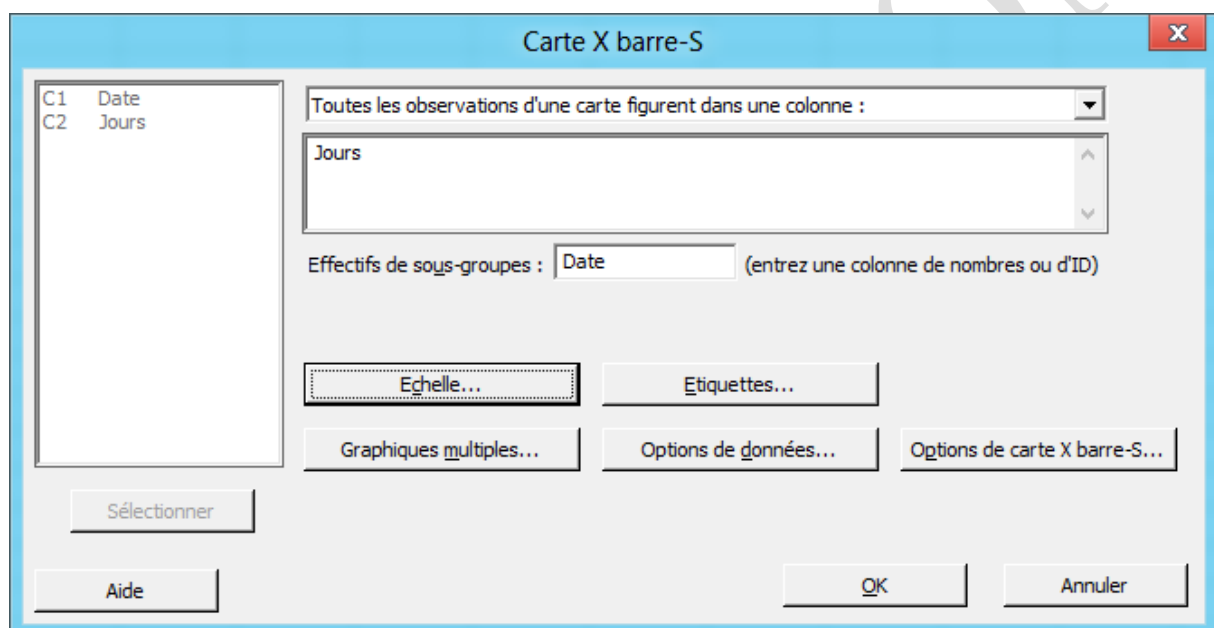
	C1-D	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
	Date	Jours									
1	3/3/2006	2,07									
2	3/3/2006	2,89									
3	3/3/2006	4,20									
4	3/3/2006	2,89									
5	3/3/2006	3,05									
6	3/3/2006	2,69									
7	3/3/2006	3,39									
8	3/3/2006	2,66									
9	3/3/2006	2,68									
10	3/3/2006	3,07									
11	4/3/2006	4,75									
12	4/3/2006	2,46									
13	4/3/2006	3,12									
14	4/3/2006	3,52									
15	4/3/2006	2,62									
16	4/3/2006	2,63									
17	4/3/2006	2,74									
18	4/3/2006	3,36									
19	4/3/2006	3,46									
20	4/3/2006	2,22									
21	5/3/2006	2,92									
22	5/3/2006	2,46									
23	5/3/2006	2,92									
24	5/3/2006	3,67									

Nous voulons effectuer une analyse X barre-S barre combinée. Pour cela, nous allons dans Minitab® Statistical Software nous allons dans le menu **Stat/Cartes de contrôle/Cartes de variables pour sous-groupes/X barre-S...**:

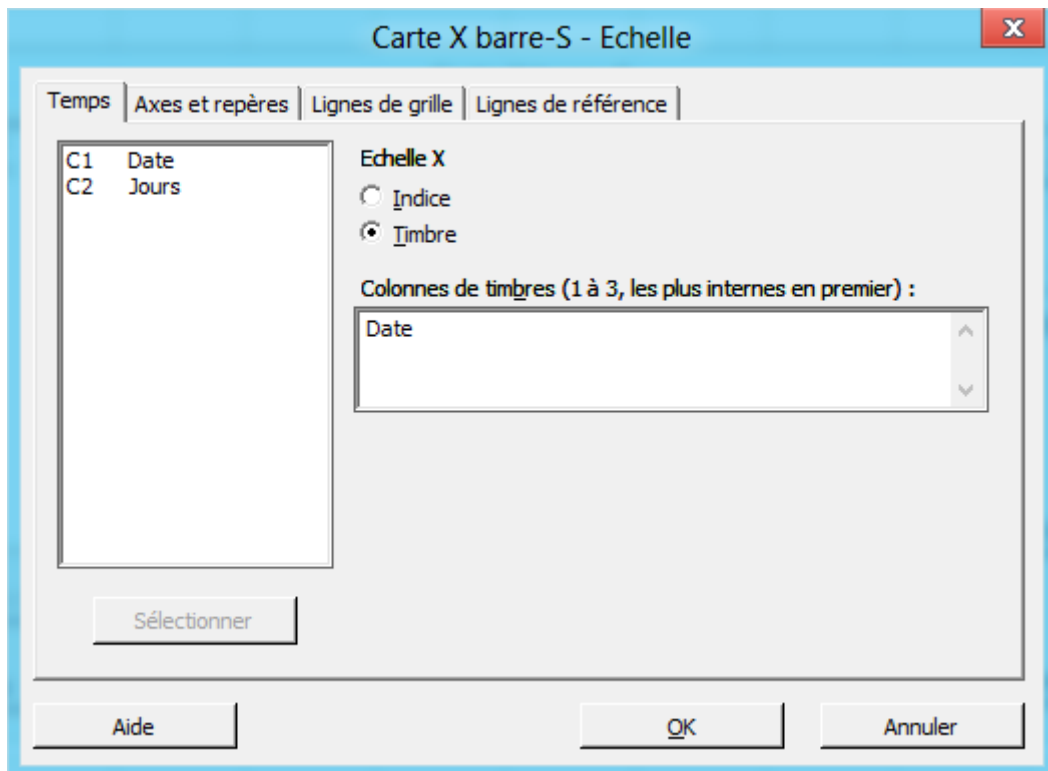




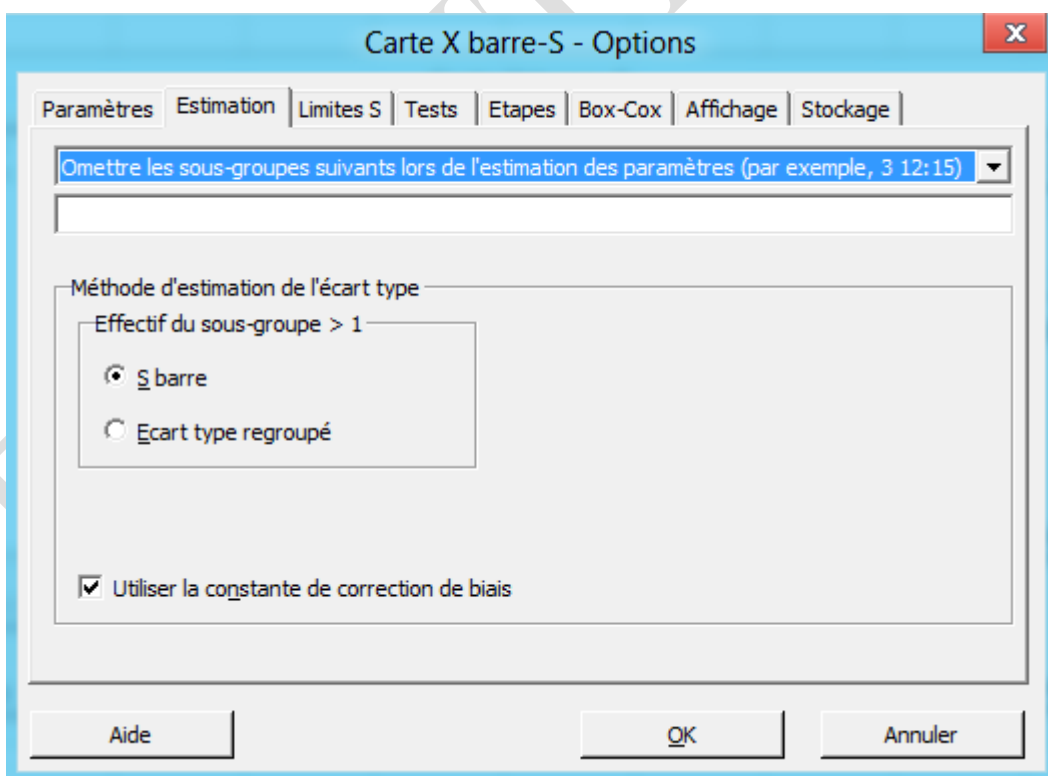
et nous prenons:



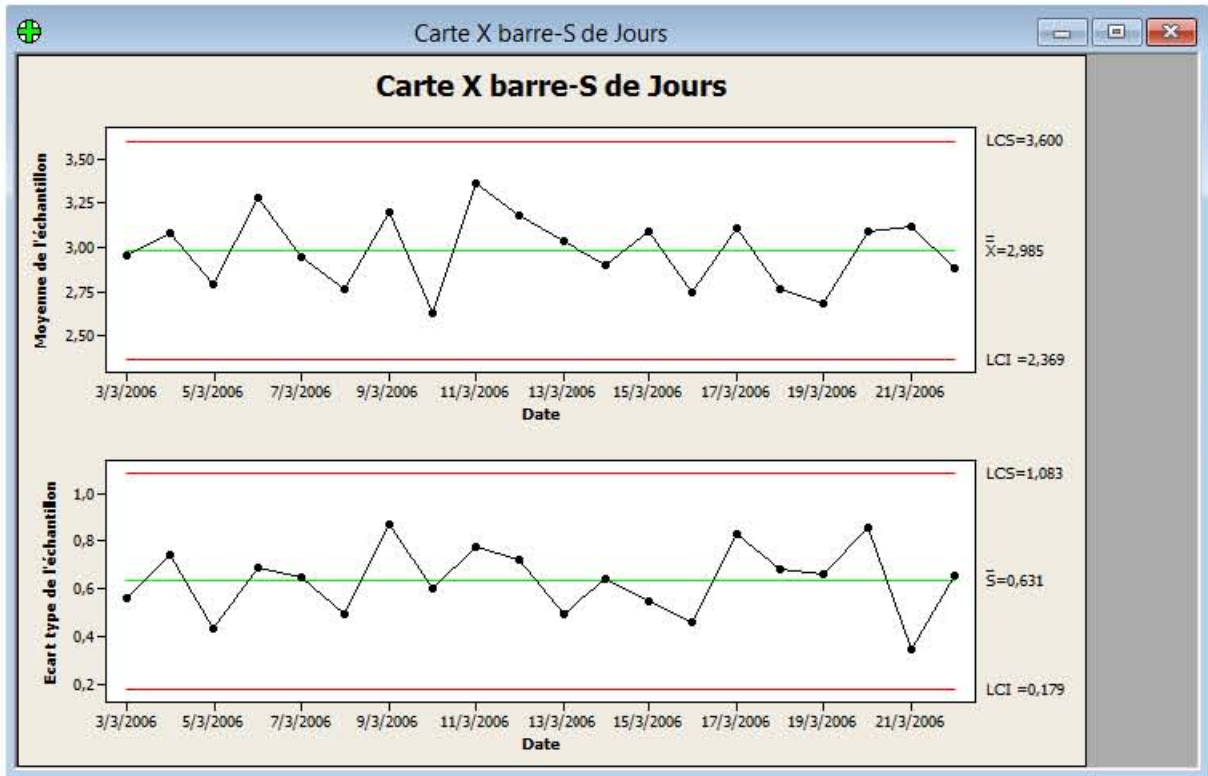
Ensuite, pour avoir le temps en abscisse, on clique sur le bouton **Echelle...** pour y mettre:



Nous validons par OK et dans **Options de carte X barre-S...**, nous prenons (car c'est quand même un peu le but...):



et nous validons tout cela par **OK** autant de fois que nécessaire pour obtenir au final nos diagrammes avec les dates:



## 14.10. Exercice 128.: CC par mesures $X$ barre- $S_p$ avec variance groupées et facteur de correction

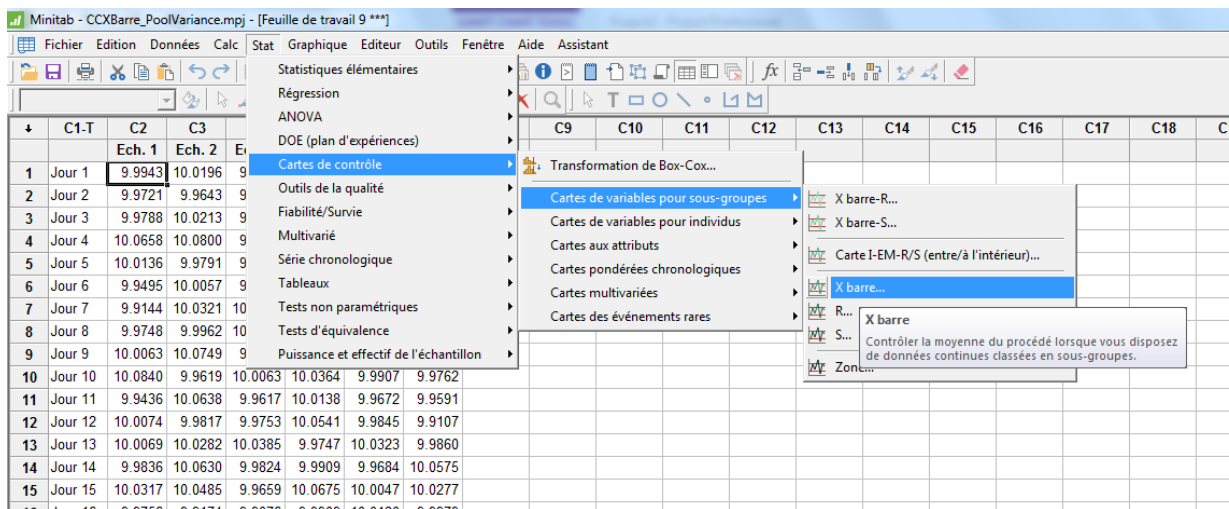
Minitab® Statistical Software 17.1.3

Dans le cadre de contrôle d'une production sous contrôle statistique nous avons obtenu les mesures suivantes (fichier *CCXBarre\_PoolVariance.mpj*):

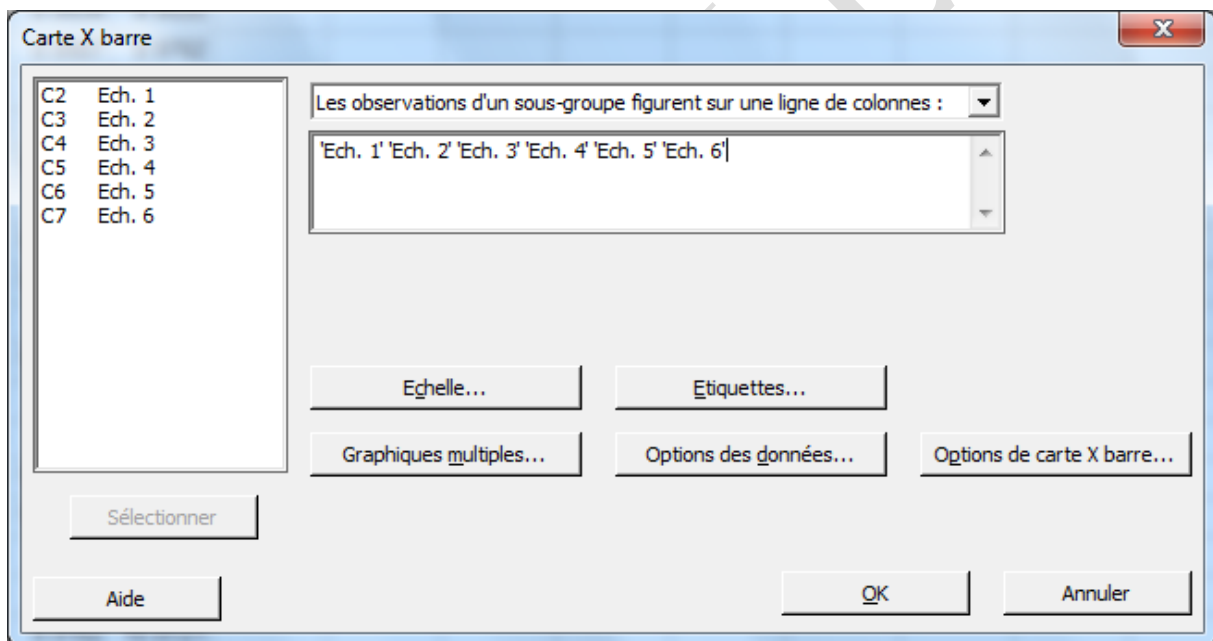
	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
		Ech. 1	Ech. 2	Ech. 3	Ech. 4	Ech. 5	Ech. 6				
1	Jour 1	9.9943	10.0196	9.9732	10.0088	9.9685	9.9544				
2	Jour 2	9.9721	9.9643	9.9426	9.9712	10.0259	10.0177				
3	Jour 3	9.9788	10.0213	9.9407	10.0696	10.0161	10.0709				
4	Jour 4	10.0658	10.0800	9.9514	9.9958	10.0338	10.0422				
5	Jour 5	10.0136	9.9791	9.9449	9.9403	9.9424	9.9421				
6	Jour 6	9.9495	10.0057	9.9715	10.0388	10.0019	9.9616				
7	Jour 7	9.9144	10.0321	10.0361	10.0253	10.0109	9.9935				
8	Jour 8	9.9748	9.9962	10.0369	10.1500	9.9996	9.9741				
9	Jour 9	10.0063	10.0749	9.9971	9.9780	9.9934	9.9555				
10	Jour 10	10.0840	9.9619	10.0063	10.0364	9.9907	9.9762				
11	Jour 11	9.9436	10.0638	9.9617	10.0138	9.9672	9.9591				
12	Jour 12	10.0074	9.9817	9.9753	10.0541	9.9845	9.9107				
13	Jour 13	10.0069	10.0282	10.0385	9.9747	10.0323	9.9860				
14	Jour 14	9.9836	10.0630	9.9824	9.9909	9.9684	10.0575				
15	Jour 15	10.0317	10.0485	9.9659	10.0675	10.0047	10.0277				
16	Jour 16	9.9756	9.9174	9.9076	9.9909	10.0180	9.9979				
17	Jour 17	9.9818	9.8962	9.9614	9.9612	9.9694	9.9349				
18	Jour 18	10.0168	10.0043	10.0029	9.9549	10.0199	9.9691				
19	Jour 19	9.9857	10.0503	10.0381	9.9811	9.9603	9.9123				
20	Jour 20	9.9878	9.9841	9.9496	9.9754	9.9532	10.0576				
21	Jour 21	10.0657	9.9838	9.9636	9.9779	10.0423	10.0547				
22	Jour 22	9.9564	10.0184	9.9890	9.9999	10.0146	10.0145				
23	Jour 23	10.0611	9.9443	9.9899	9.9833	9.9947	10.0033				
24	Jour 24	9.9054	9.9049	9.9870	10.0580	9.9156	9.9782				
25	Jour 25	9.9768	9.9954	9.9196	9.9341	9.9162	10.0721				
26											
27											

Nous souhaiterions tracer une carte de contrôle par mesures groupées de type  $X$  barre  $S_p$  et comparer avec les calculs théoriques effectués dans le cours de statistique théorique et dans le cours Microsoft Excel.

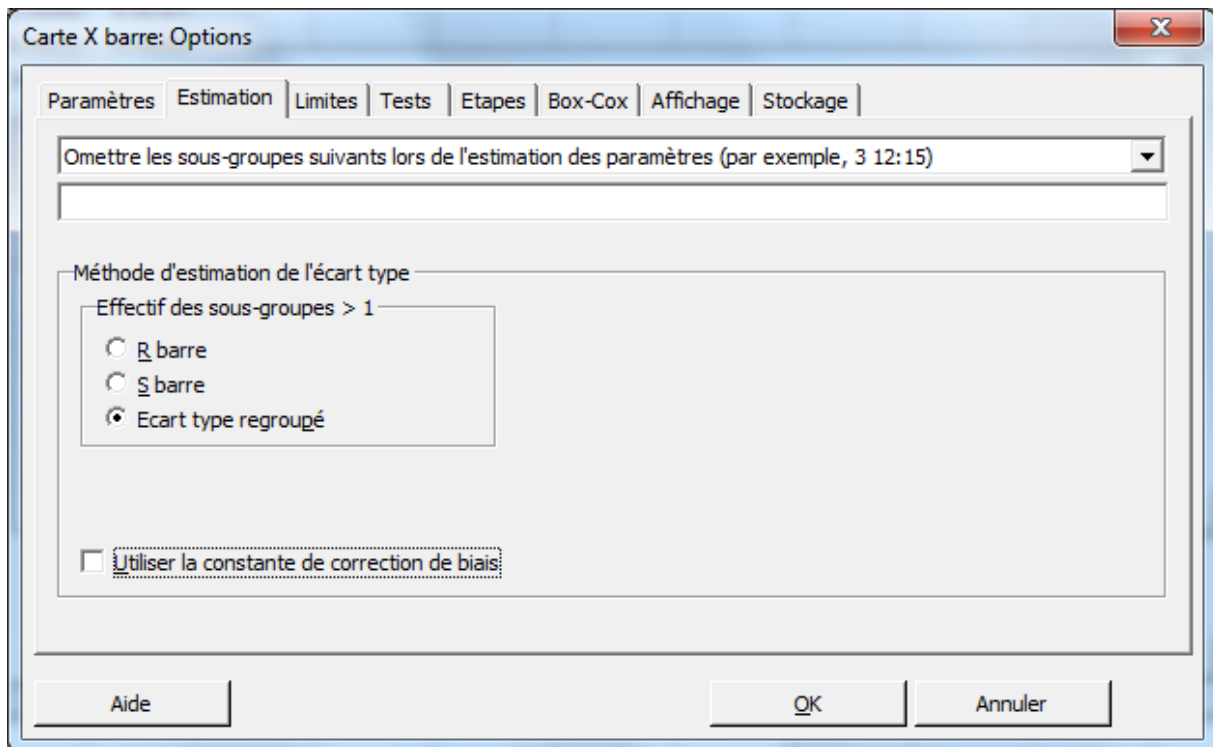
Dans Minitab® Statistical Software nous allons dans le menu **Stat/Cartes de contrôle/Cartes de variables pour sous-groupes/X barre...**:



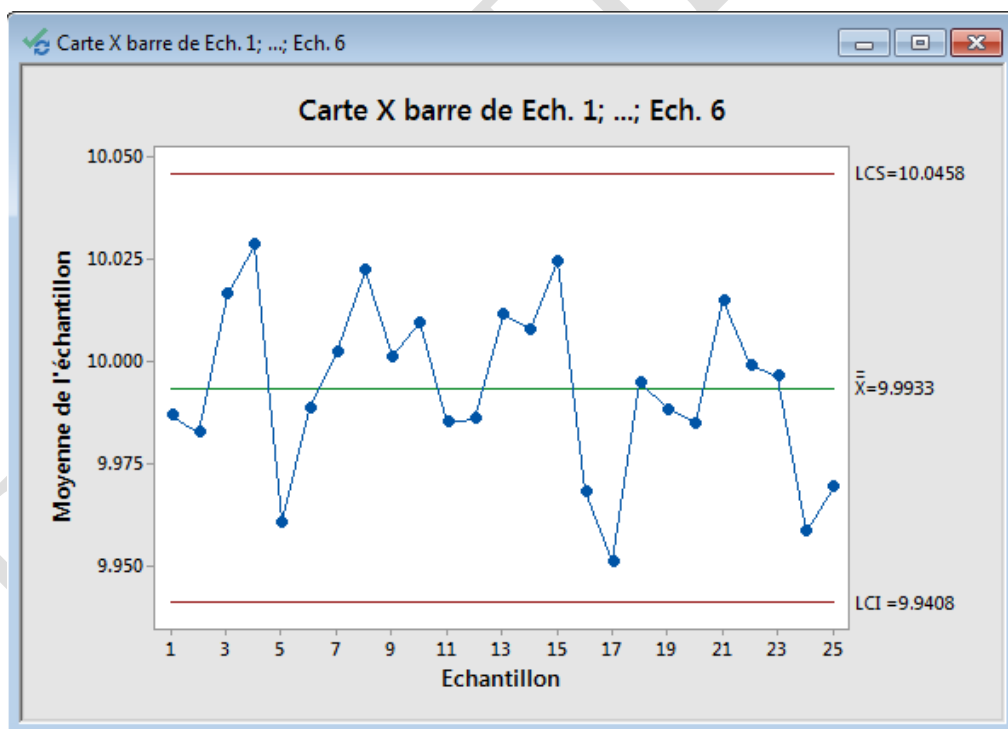
Nous prenons alors bien évidemment les colonnes correspondantes:



et dans l'onglet **Estimation** nous sélection Ecart type regroupé et cochons la case **Utiliser la constant de correction de biais**:



Nous obtenons alors:

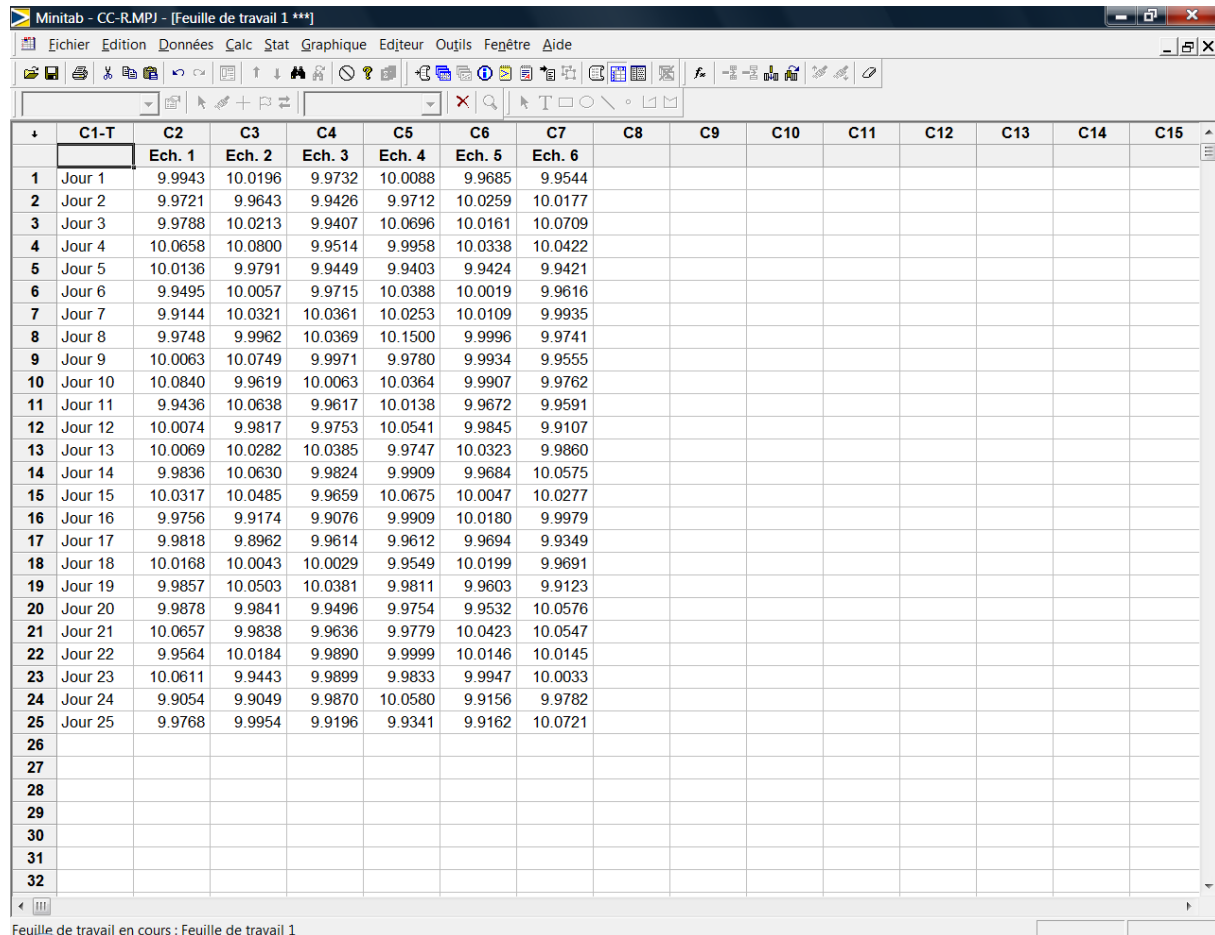


et donc le graphique et les valeurs qui correspondent donc toujours à ce qui a été calculé dans le cours de statistique théorique et Microsoft Excel.

## 14.11. Exercice 129.: CC par mesures R barre-R barre simple

Minitab® Statistical Software 15.1.1.0

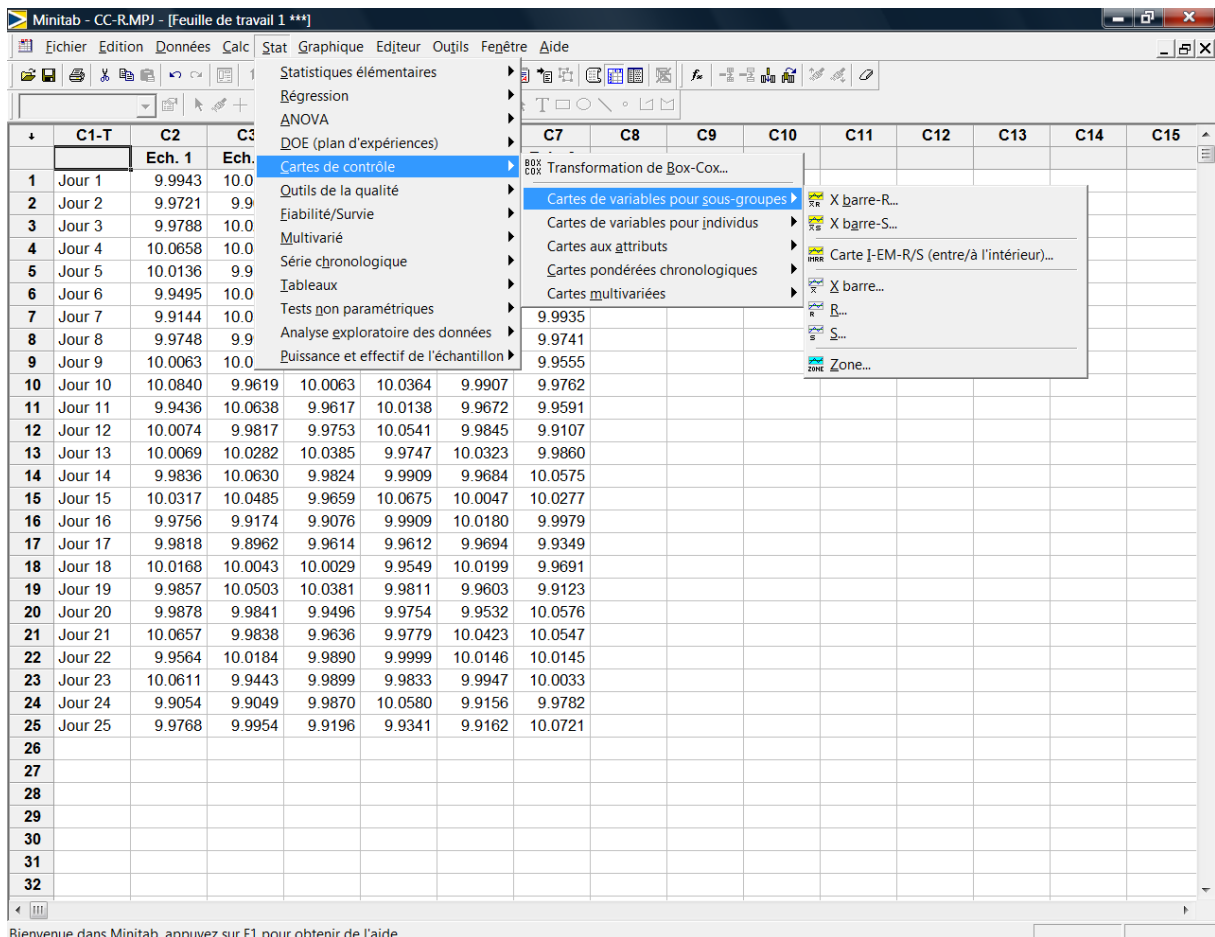
Dans le cadre de contrôle d'une production sous contrôle statistique nous avons obtenu les mesures suivantes (fichier *CC Rbarre-R.mpj*):



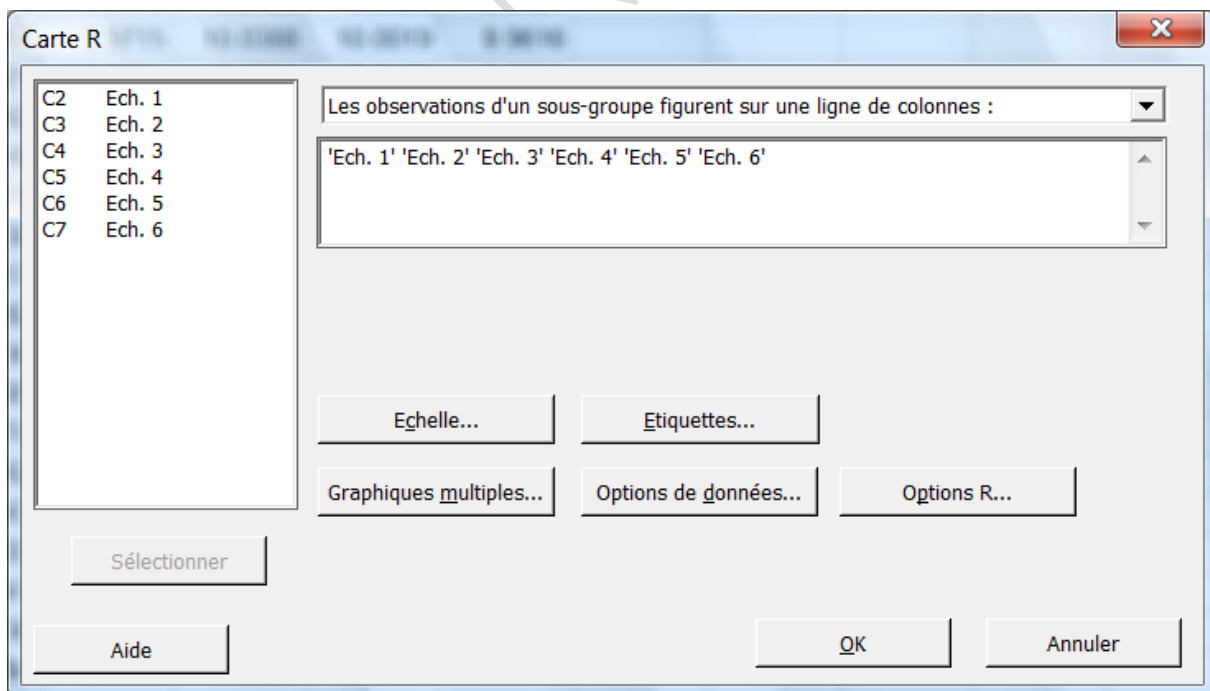
	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
		Ech. 1	Ech. 2	Ech. 3	Ech. 4	Ech. 5	Ech. 6								
1	Jour 1	9.9943	10.0196	9.9732	10.0088	9.9685	9.9544								
2	Jour 2	9.9721	9.9643	9.9426	9.9712	10.0259	10.0177								
3	Jour 3	9.9788	10.0213	9.9407	10.0696	10.0161	10.0709								
4	Jour 4	10.0658	10.0800	9.9514	9.9958	10.0338	10.0422								
5	Jour 5	10.0136	9.9791	9.9449	9.9403	9.9424	9.9421								
6	Jour 6	9.9495	10.0057	9.9715	10.0388	10.0019	9.9616								
7	Jour 7	9.9144	10.0321	10.0361	10.0253	10.0109	9.9935								
8	Jour 8	9.9748	9.9962	10.0369	10.1500	9.9996	9.9741								
9	Jour 9	10.0063	10.0749	9.9971	9.9780	9.9934	9.9555								
10	Jour 10	10.0840	9.9619	10.0063	10.0364	9.9907	9.9762								
11	Jour 11	9.9436	10.0638	9.9617	10.0138	9.9672	9.9591								
12	Jour 12	10.0074	9.9817	9.9753	10.0541	9.9845	9.9107								
13	Jour 13	10.0069	10.0282	10.0385	9.9747	10.0323	9.9860								
14	Jour 14	9.9836	10.0630	9.9824	9.9909	9.9684	10.0575								
15	Jour 15	10.0317	10.0485	9.9659	10.0675	10.0047	10.0277								
16	Jour 16	9.9756	9.9174	9.9076	9.9909	10.0180	9.9979								
17	Jour 17	9.9818	9.8962	9.9614	9.9612	9.9694	9.9349								
18	Jour 18	10.0168	10.0043	10.0029	9.9549	10.0199	9.9691								
19	Jour 19	9.9857	10.0503	10.0381	9.9811	9.9603	9.9123								
20	Jour 20	9.9878	9.9841	9.9496	9.9754	9.9532	10.0576								
21	Jour 21	10.0657	9.9838	9.9636	9.9779	10.0423	10.0547								
22	Jour 22	9.9564	10.0184	9.9890	9.9999	10.0146	10.0145								
23	Jour 23	10.0611	9.9443	9.9899	9.9833	9.9947	10.0033								
24	Jour 24	9.9054	9.9049	9.9870	10.0580	9.9156	9.9782								
25	Jour 25	9.9768	9.9954	9.9196	9.9341	9.9162	10.0721								
26															
27															
28															
29															
30															
31															
32															

Nous souhaiterions tracer une carte de contrôle par mesures groupées de type *R barre R* et comparer avec les calculs théoriques effectués dans le cours de statistique théorique et dans le cours Microsoft Excel.

Dans Minitab® Statistical Software nous allons dans le menu **Stat/Cartes de contrôle/Cartes de variables pour sous-groupes/R...**:

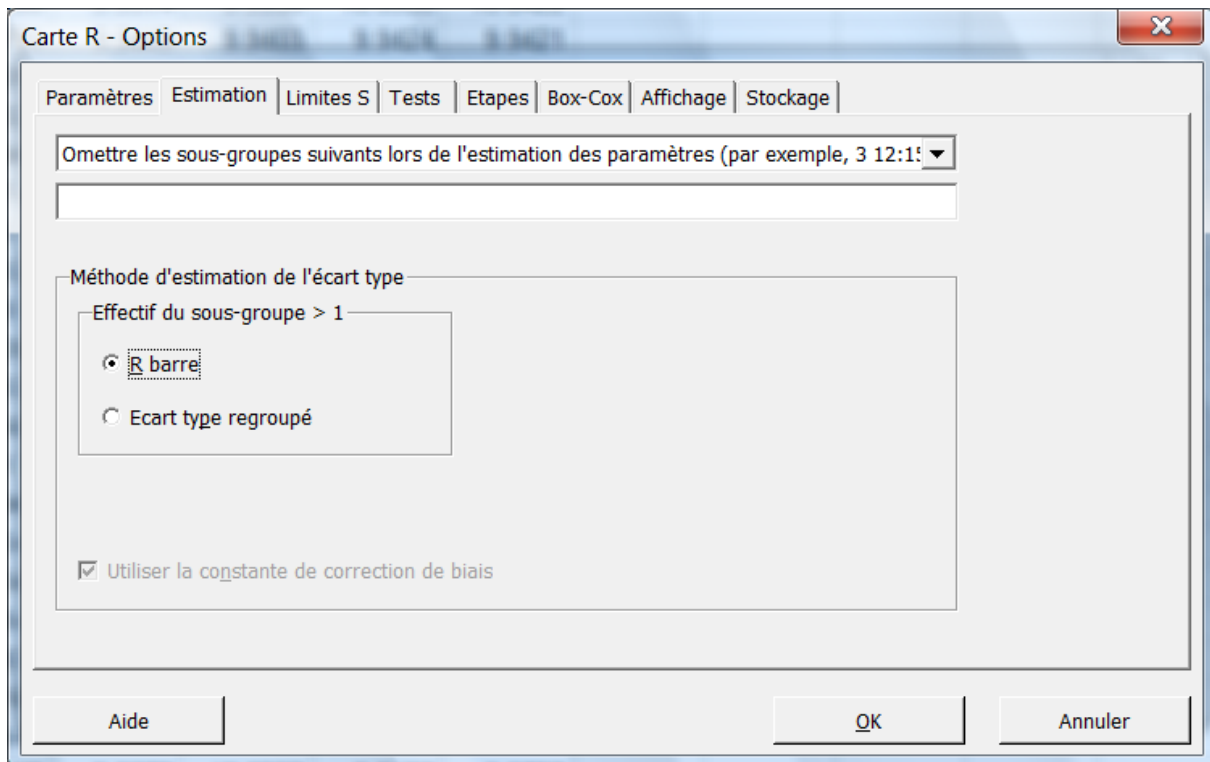


Nous avons alors la boîte de dialogue suivante:

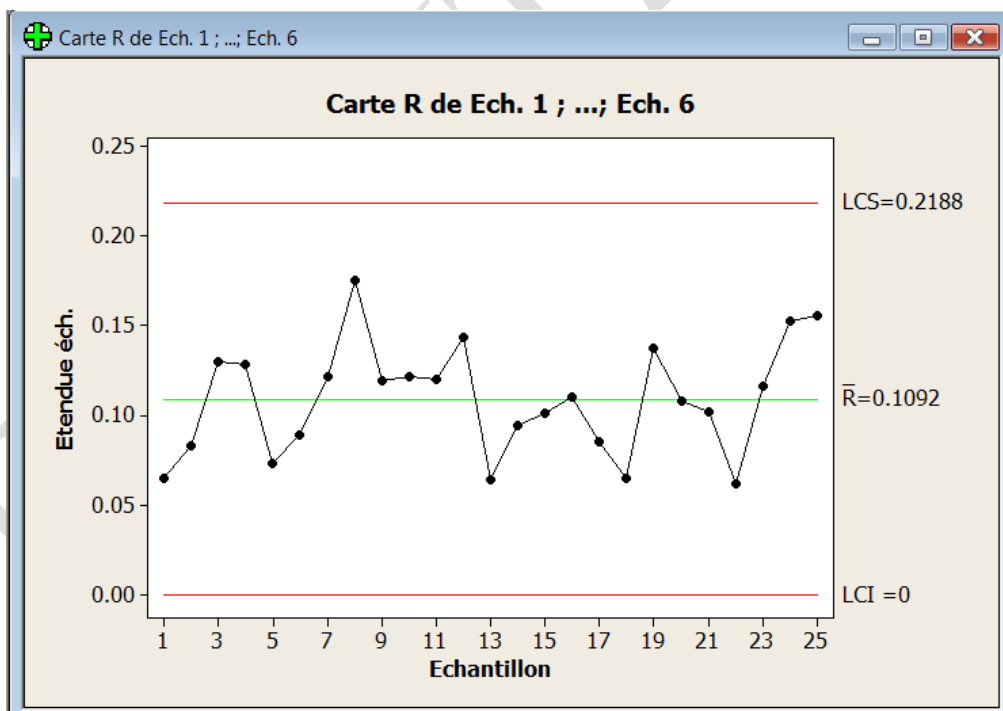


Nous cliquons sur **Options R...**:





pour cocher l'option **R barre** car c'est un peu le but quand même. Nous validons deux fois par **OK**:

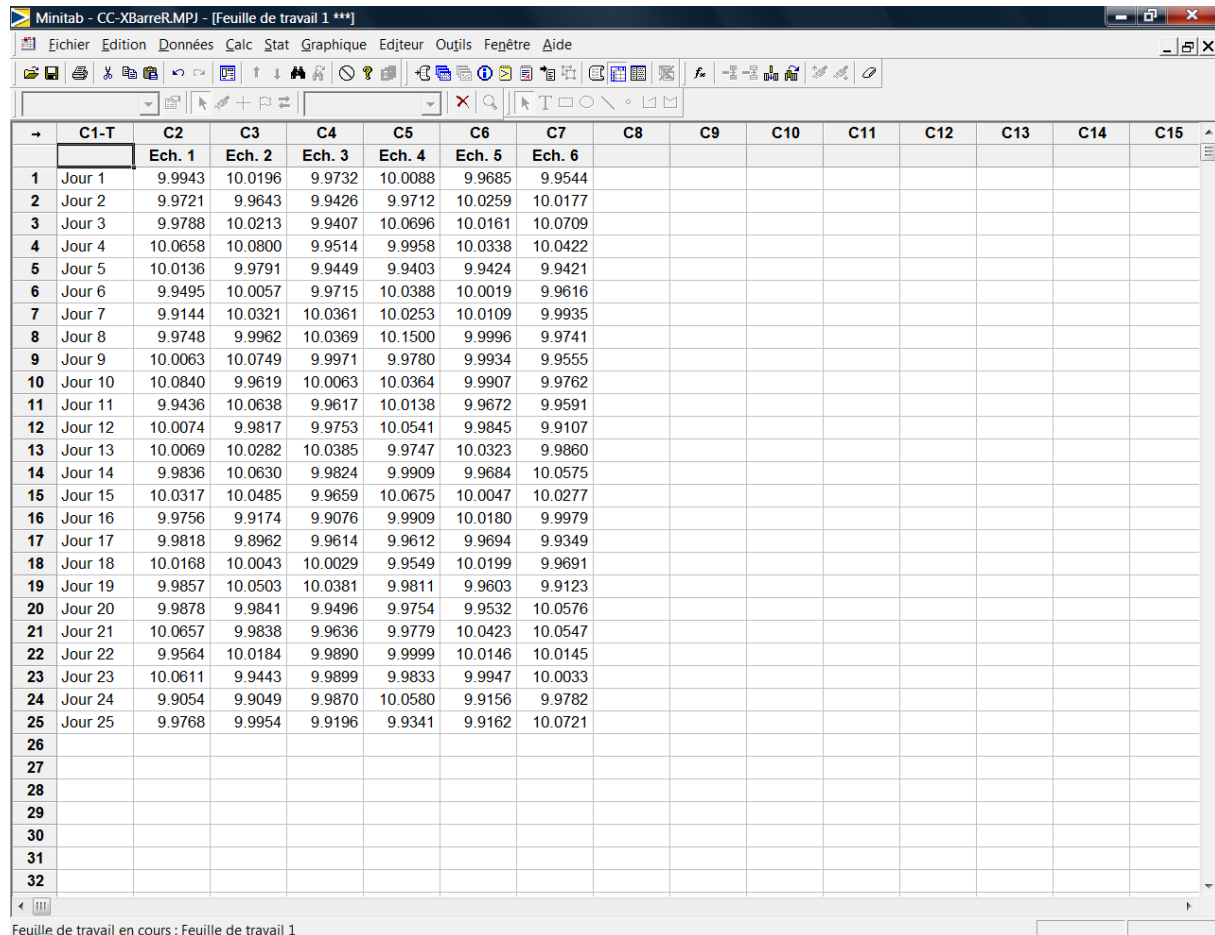


Nous obtenons donc les mêmes valeurs et la même carte que dans le cours de statistique théorique et Microsoft Excel.

## 14.12. Exercice 130.: CC par mesures X barre-R barre simple

Minitab® Statistical Software 15.1.1.0

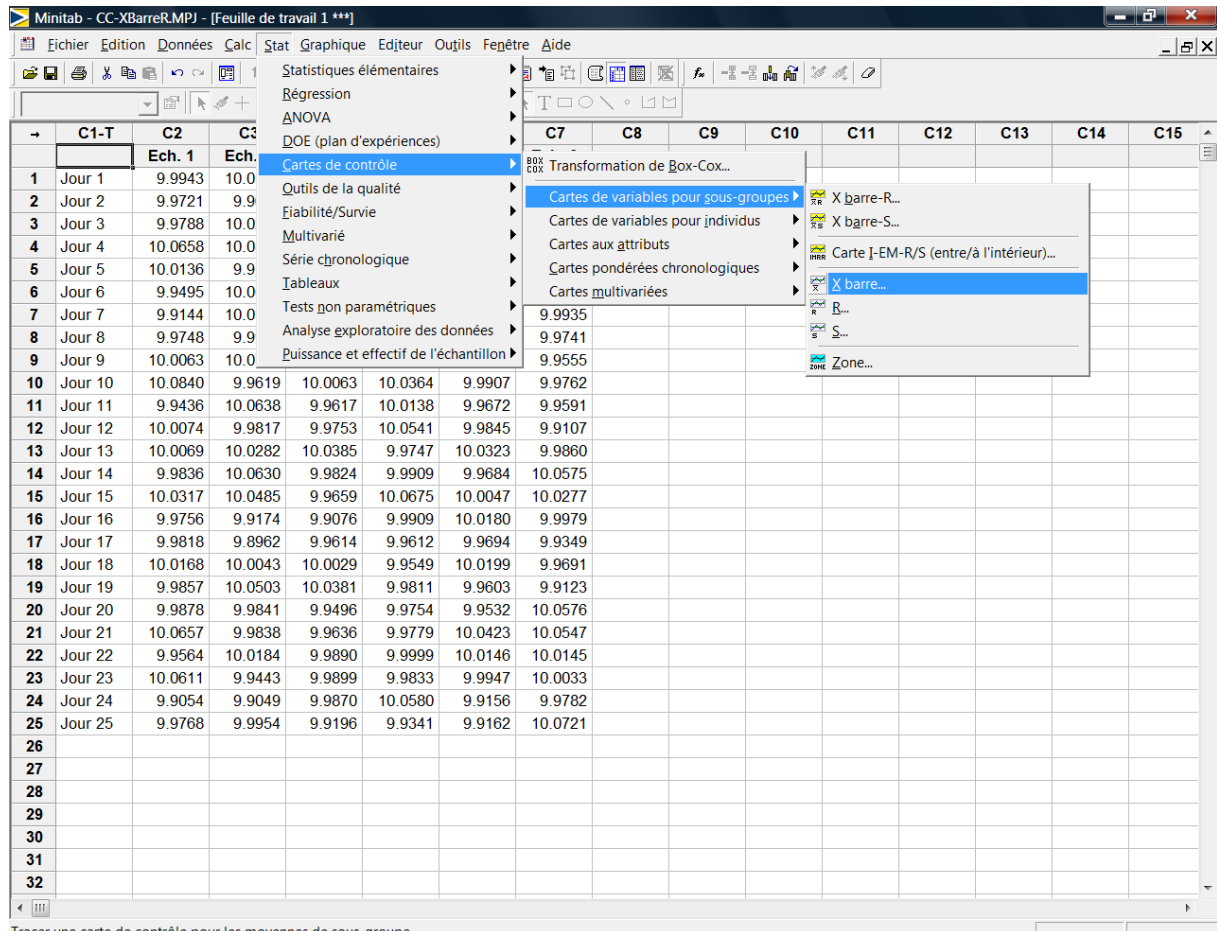
Dans le cadre de contrôle d'une production sous contrôle statistique nous avons obtenu les mesures suivantes (fichier *CC XBarre-R.mpj*):



	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
		Ech. 1	Ech. 2	Ech. 3	Ech. 4	Ech. 5	Ech. 6								
1	Jour 1	9.9943	10.0196	9.9732	10.0088	9.9685	9.9544								
2	Jour 2	9.9721	9.9643	9.9426	9.9712	10.0259	10.0177								
3	Jour 3	9.9788	10.0213	9.9407	10.0696	10.0161	10.0709								
4	Jour 4	10.0658	10.0800	9.9514	9.9958	10.0338	10.0422								
5	Jour 5	10.0136	9.9791	9.9449	9.9403	9.9424	9.9421								
6	Jour 6	9.9495	10.0057	9.9715	10.0388	10.0019	9.9616								
7	Jour 7	9.9144	10.0321	10.0361	10.0253	10.0109	9.9935								
8	Jour 8	9.9748	9.9962	10.0369	10.1500	9.9996	9.9741								
9	Jour 9	10.0063	10.0749	9.9971	9.9780	9.9934	9.9555								
10	Jour 10	10.0840	9.9619	10.0063	10.0364	9.9907	9.9762								
11	Jour 11	9.9436	10.0638	9.9617	10.0138	9.9672	9.9591								
12	Jour 12	10.0074	9.9817	9.9753	10.0541	9.9845	9.9107								
13	Jour 13	10.0069	10.0282	10.0385	9.9747	10.0323	9.9860								
14	Jour 14	9.9836	10.0630	9.9824	9.9909	9.9684	10.0575								
15	Jour 15	10.0317	10.0485	9.9659	10.0675	10.0047	10.0277								
16	Jour 16	9.9756	9.9174	9.9076	9.9909	10.0180	9.9979								
17	Jour 17	9.9818	9.8962	9.9614	9.9612	9.9694	9.9349								
18	Jour 18	10.0168	10.0043	10.0029	9.9549	10.0199	9.9691								
19	Jour 19	9.9857	10.0503	10.0381	9.9811	9.9603	9.9123								
20	Jour 20	9.9878	9.9841	9.9496	9.9754	9.9532	10.0576								
21	Jour 21	10.0657	9.9838	9.9636	9.9779	10.0423	10.0547								
22	Jour 22	9.9564	10.0184	9.9890	9.9999	10.0146	10.0145								
23	Jour 23	10.0611	9.9443	9.9899	9.9833	9.9947	10.0033								
24	Jour 24	9.9054	9.9049	9.9870	10.0580	9.9156	9.9782								
25	Jour 25	9.9768	9.9954	9.9196	9.9341	9.9162	10.0721								
26															
27															
28															
29															
30															
31															
32															

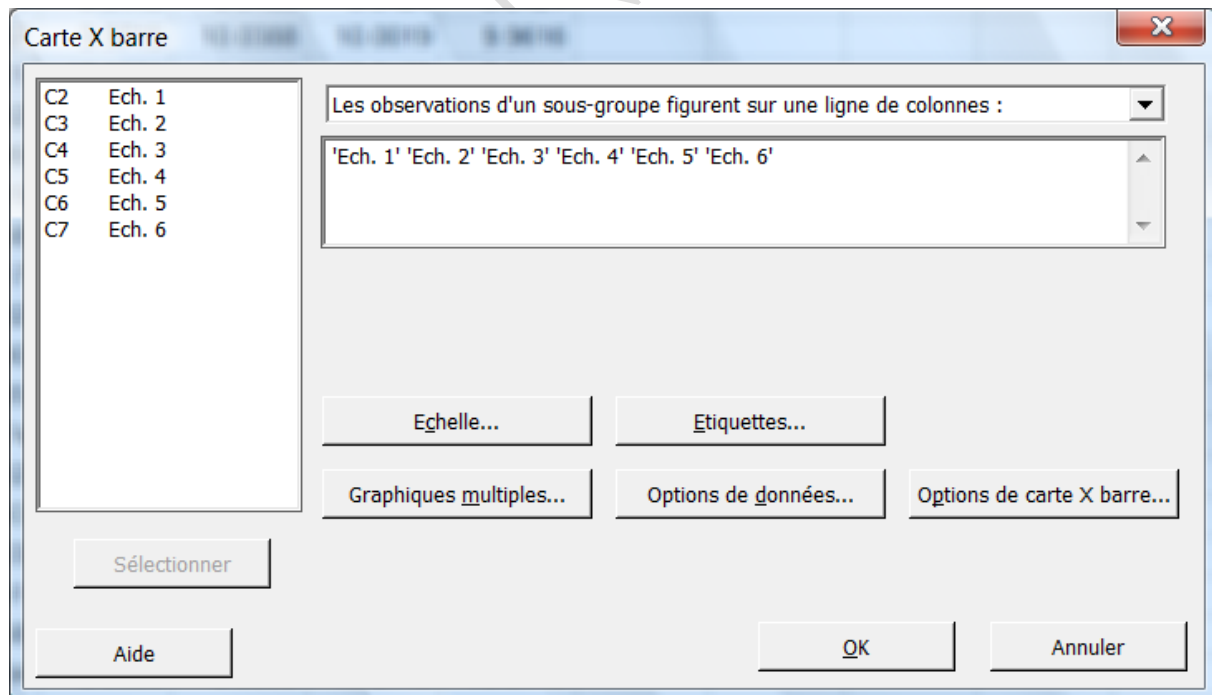
Nous souhaiterions tracer une carte de contrôle par mesures groupées de type *X barre R* et comparer avec les calculs théoriques effectués dans le cours de statistique théorique et dans le cours Microsoft Excel.

Dans Minitab® Statistical Software nous allons dans le menu **Stat/Cartes de contrôle/Cartes de variables pour sous-groupes/X barre...**:

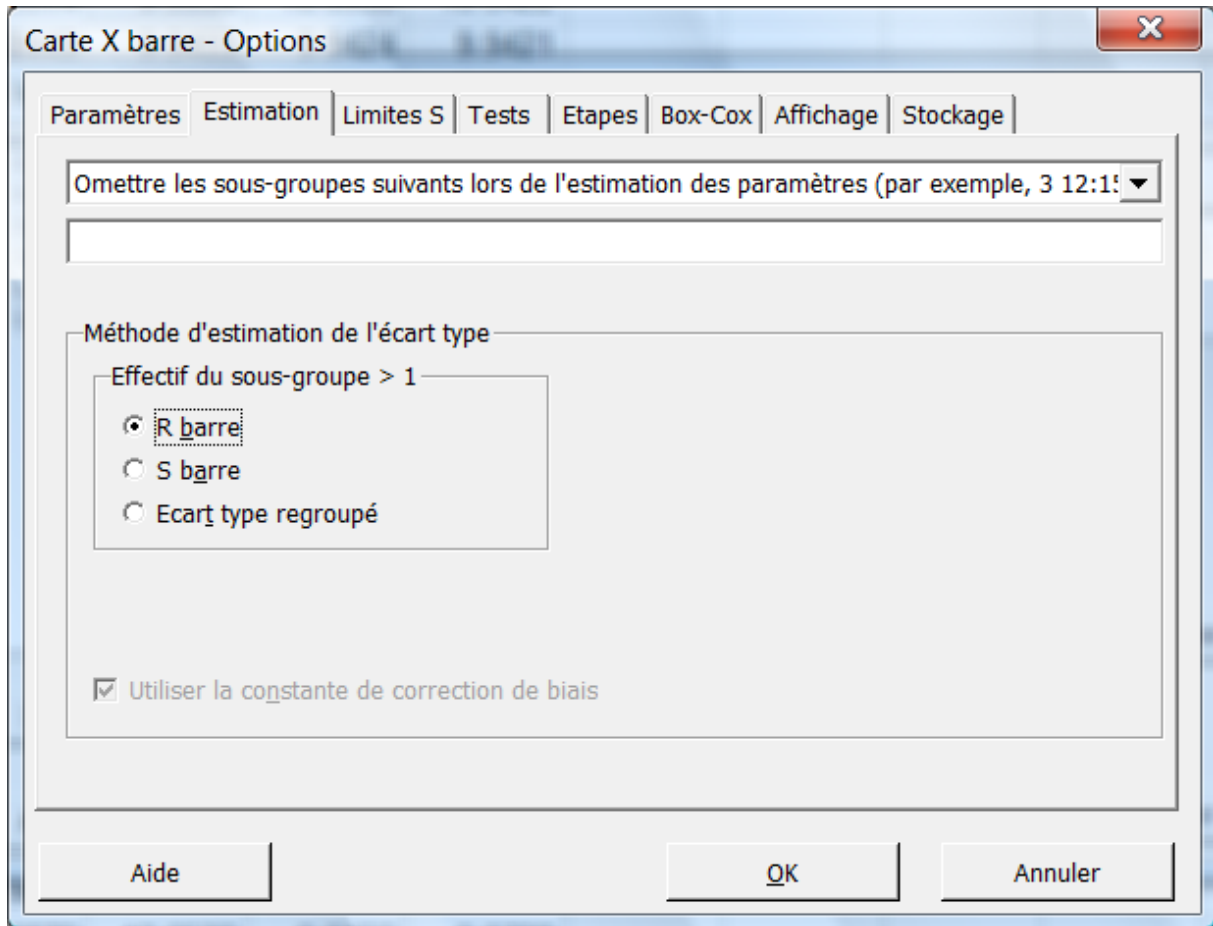


Tracer une carte de contrôle pour les moyennes de sous-aroupe

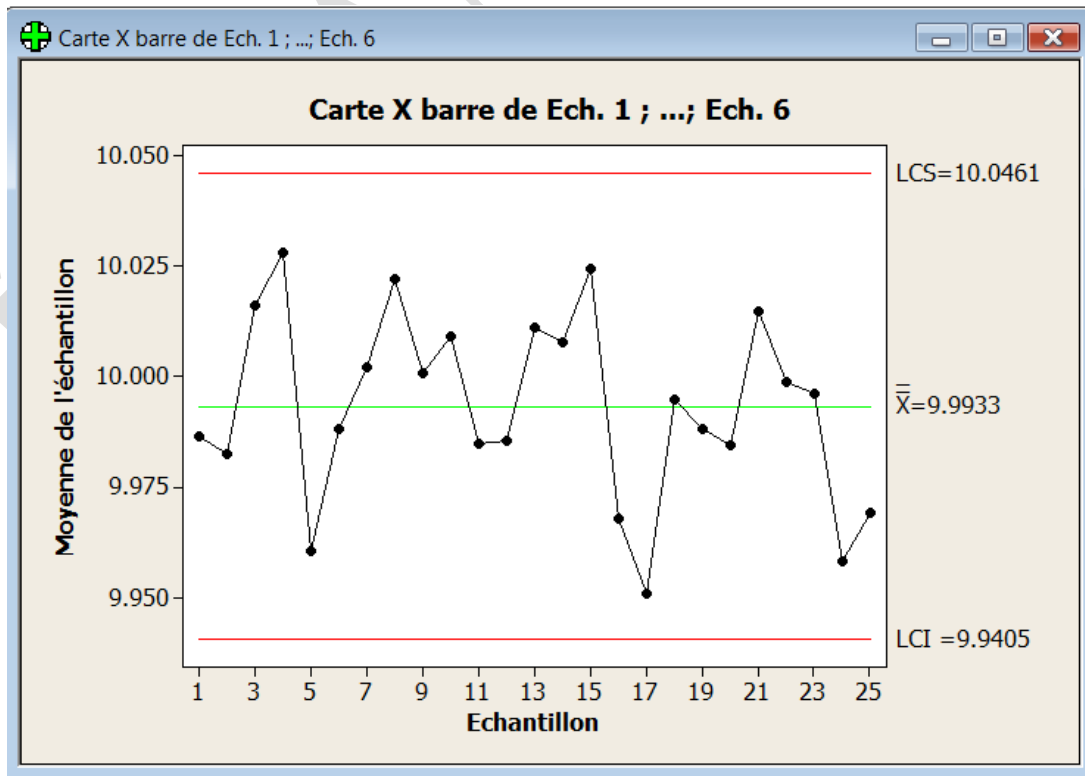
Nous avons alors la boîte de dialogue suivante:



Nous cliquons sur **Options de carte X barre...**:



pour cocher l'option **R barre** car c'est un peu le but quand même. Nous validons deux fois par **OK**:



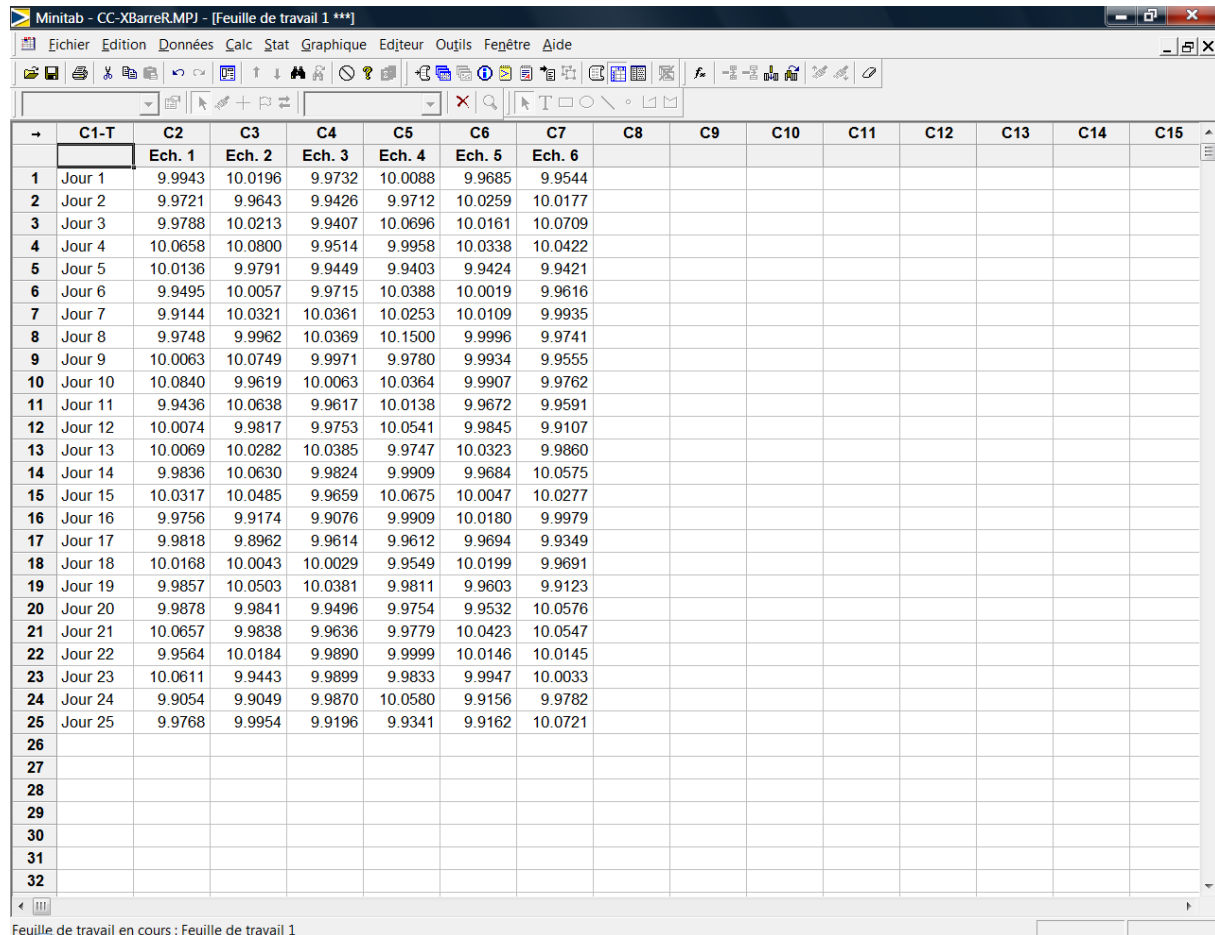
Nous obtenons donc les mêmes valeurs et la même carte que dans le cours de statistique théorique et Microsoft Excel.

ÉCHANTILLON

## 14.13. Exercice 131.: CC par mesures X barre-R barre combinées

Minitab® Statistical Software 15.1.1.0

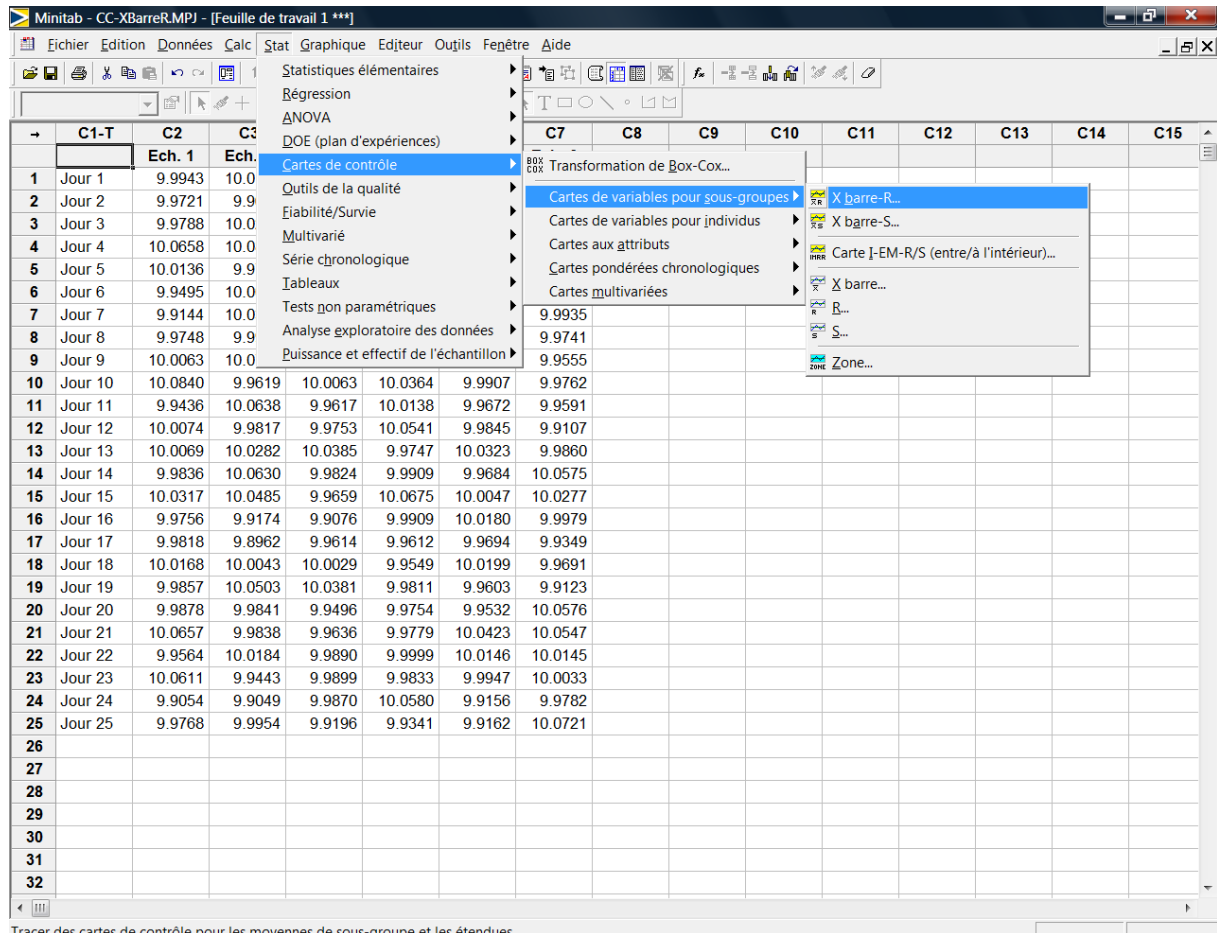
Dans le cadre de contrôle d'une production sous contrôle statistique nous avons obtenu les mesures suivantes (fichier *CC Xbarre-R.mpj*):



	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
		Ech. 1	Ech. 2	Ech. 3	Ech. 4	Ech. 5	Ech. 6								
1	Jour 1	9.9943	10.0196	9.9732	10.0088	9.9685	9.9544								
2	Jour 2	9.9721	9.9643	9.9426	9.9712	10.0259	10.0177								
3	Jour 3	9.9788	10.0213	9.9407	10.0696	10.0161	10.0709								
4	Jour 4	10.0658	10.0800	9.9514	9.9958	10.0338	10.0422								
5	Jour 5	10.0136	9.9791	9.9449	9.9403	9.9424	9.9421								
6	Jour 6	9.9495	10.0057	9.9715	10.0388	10.0019	9.9616								
7	Jour 7	9.9144	10.0321	10.0361	10.0253	10.0109	9.9935								
8	Jour 8	9.9748	9.9962	10.0369	10.1500	9.9996	9.9741								
9	Jour 9	10.0063	10.0749	9.9971	9.9780	9.9934	9.9555								
10	Jour 10	10.0840	9.9619	10.0063	10.0364	9.9907	9.9762								
11	Jour 11	9.9436	10.0638	9.9617	10.0138	9.9672	9.9591								
12	Jour 12	10.0074	9.9817	9.9753	10.0541	9.9845	9.9107								
13	Jour 13	10.0069	10.0282	10.0385	9.9747	10.0323	9.9860								
14	Jour 14	9.9836	10.0630	9.9824	9.9909	9.9684	10.0575								
15	Jour 15	10.0317	10.0485	9.9659	10.0675	10.0047	10.0277								
16	Jour 16	9.9756	9.9174	9.9076	9.9909	10.0180	9.9979								
17	Jour 17	9.9818	9.8962	9.9614	9.9612	9.9694	9.9349								
18	Jour 18	10.0168	10.0043	10.0029	9.9549	10.0199	9.9691								
19	Jour 19	9.9857	10.0503	10.0381	9.9811	9.9603	9.9123								
20	Jour 20	9.9878	9.9841	9.9496	9.9754	9.9532	10.0576								
21	Jour 21	10.0657	9.9838	9.9636	9.9779	10.0423	10.0547								
22	Jour 22	9.9564	10.0184	9.9890	9.9999	10.0146	10.0145								
23	Jour 23	10.0611	9.9443	9.9899	9.9833	9.9947	10.0033								
24	Jour 24	9.9054	9.9049	9.9870	10.0580	9.9156	9.9782								
25	Jour 25	9.9768	9.9954	9.9196	9.9341	9.9162	10.0721								
26															
27															
28															
29															
30															
31															
32															

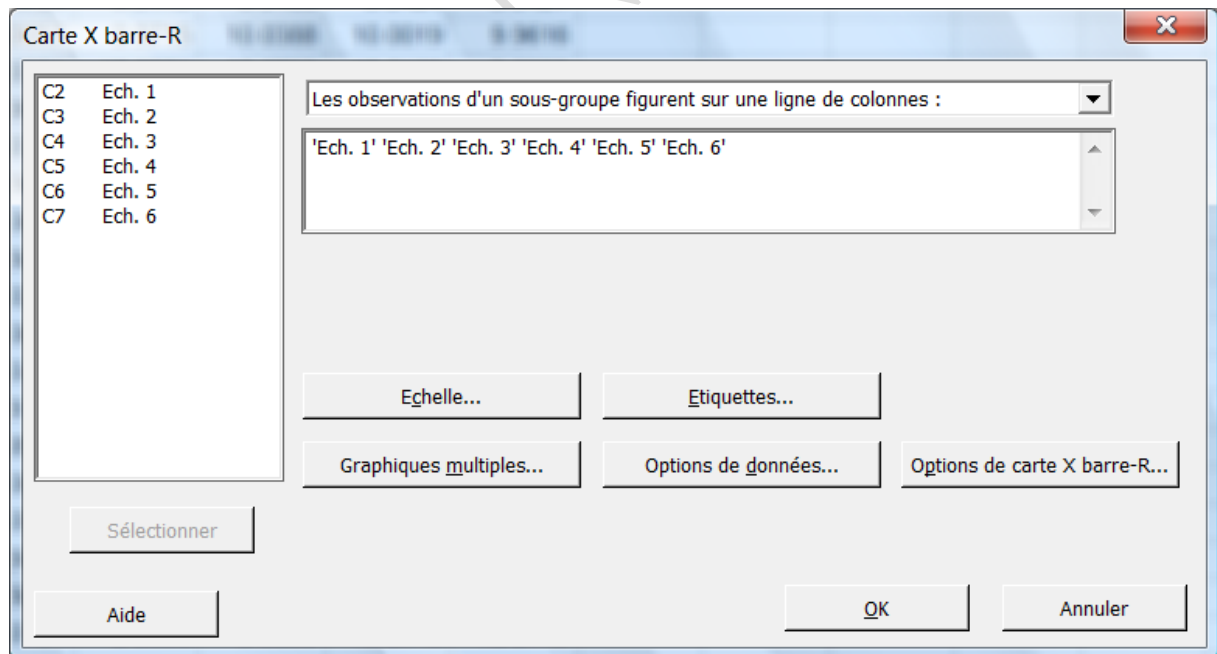
Nous souhaiterions tracer une carte de contrôle par mesures groupées de type *X barre* combinée à une *R barre* (soit les deux exercices précédents en un seul coup) et comparer avec les calculs théoriques effectués dans le cours de statistique théorique et dans le cours Microsoft Excel.

Dans Minitab® Statistical Software nous allons dans le menu **Stat/Cartes de contrôle/Cartes de variables pour sous-groupes/X barre-R...**:

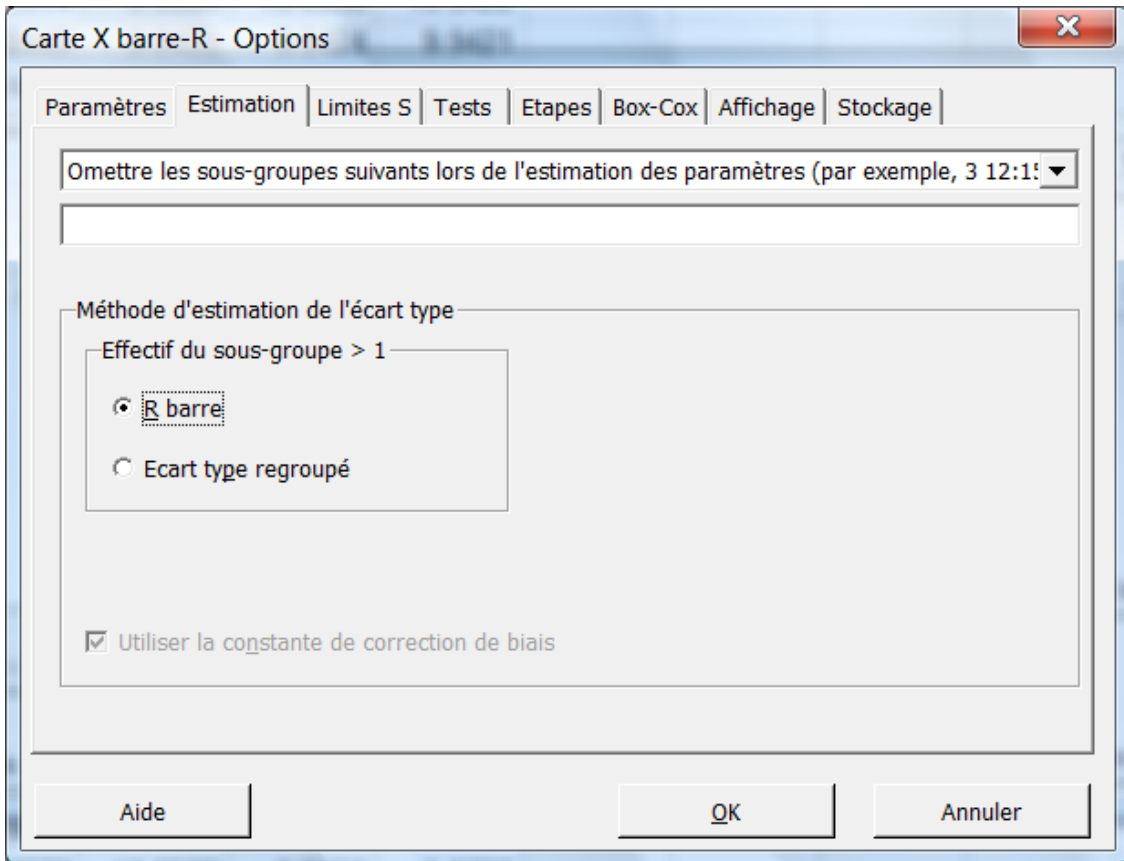


Tracer des cartes de contrôle pour les moyennes de sous-groupe et les étendues

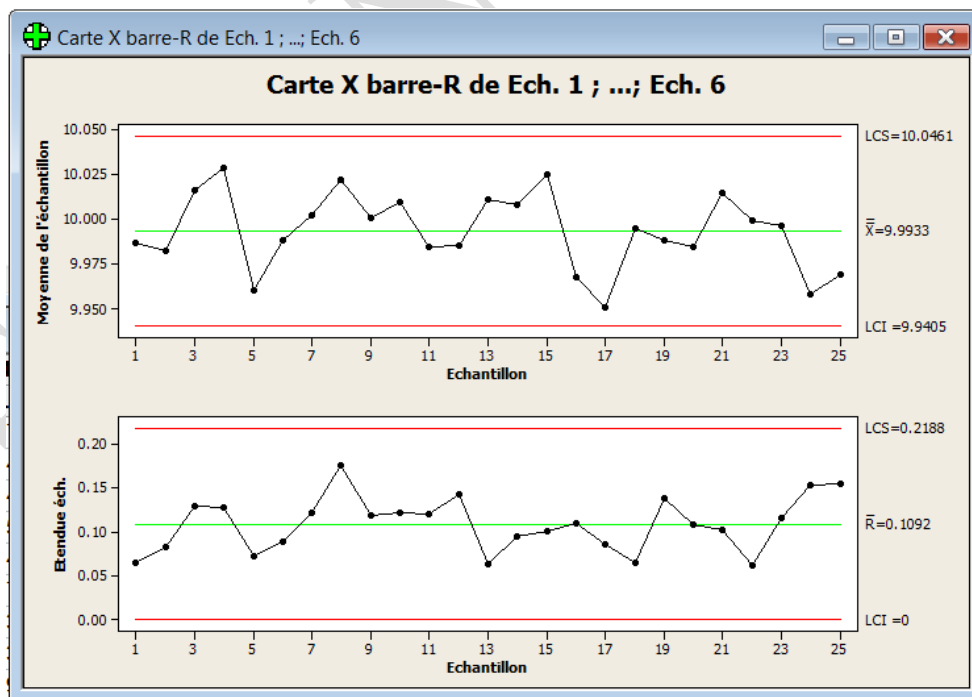
Nous avons alors la boîte de dialogue suivante:



Nous cliquons sur le bouton **Options de carte X barre-R**:



pour cocher l'option **R barre** car c'est un peu le but quand même. Nous validons deux fois par **OK**:



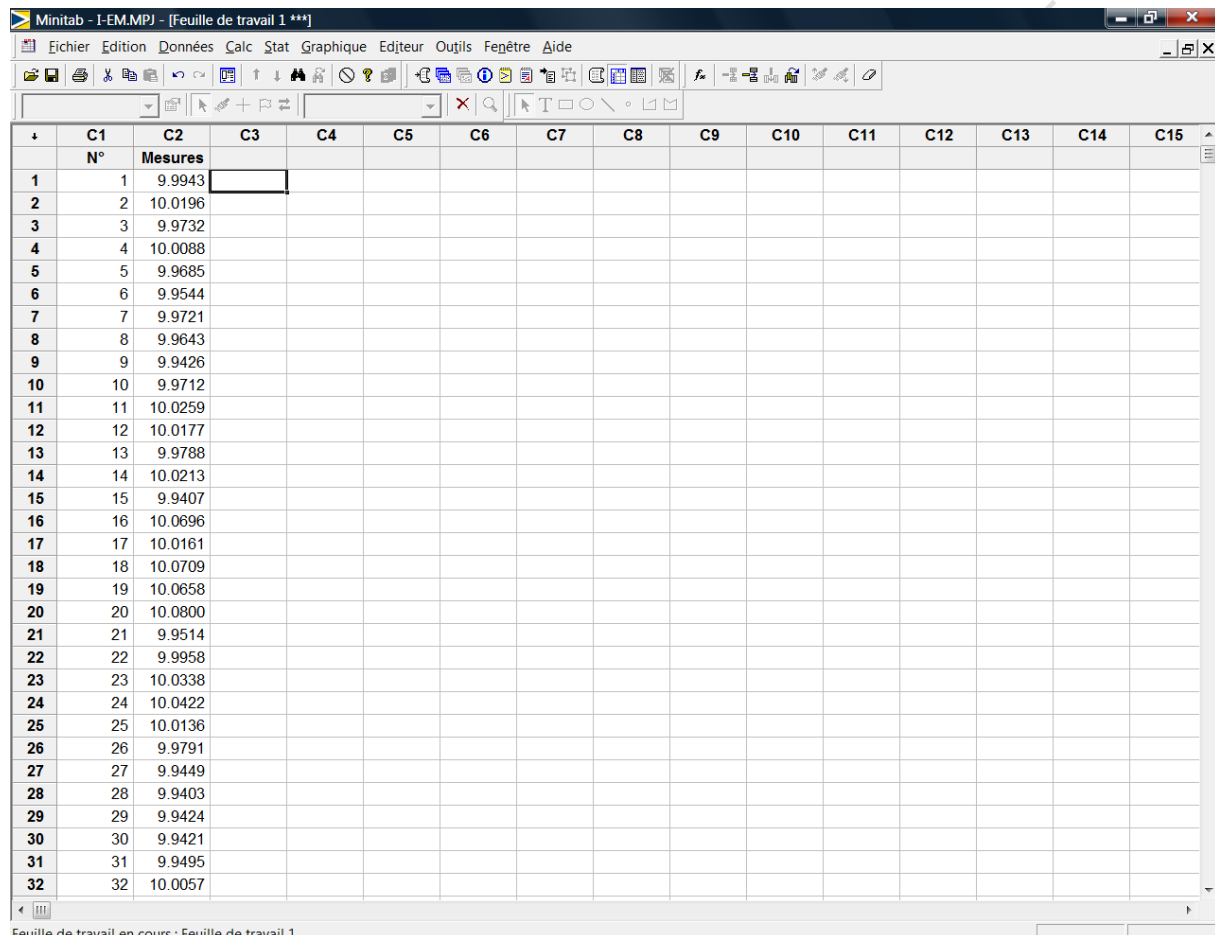
et donc nous retrouvons bien les graphiques des deux exercices précédents qui correspondent donc toujours à ce qui a été calculé dans le cours de statistique théorique et Microsoft Excel.



## 14.14. Exercice 132.: CC par mesures individuelles par l'étendue mobile

Minitab® Statistical Software 15.1.1.0

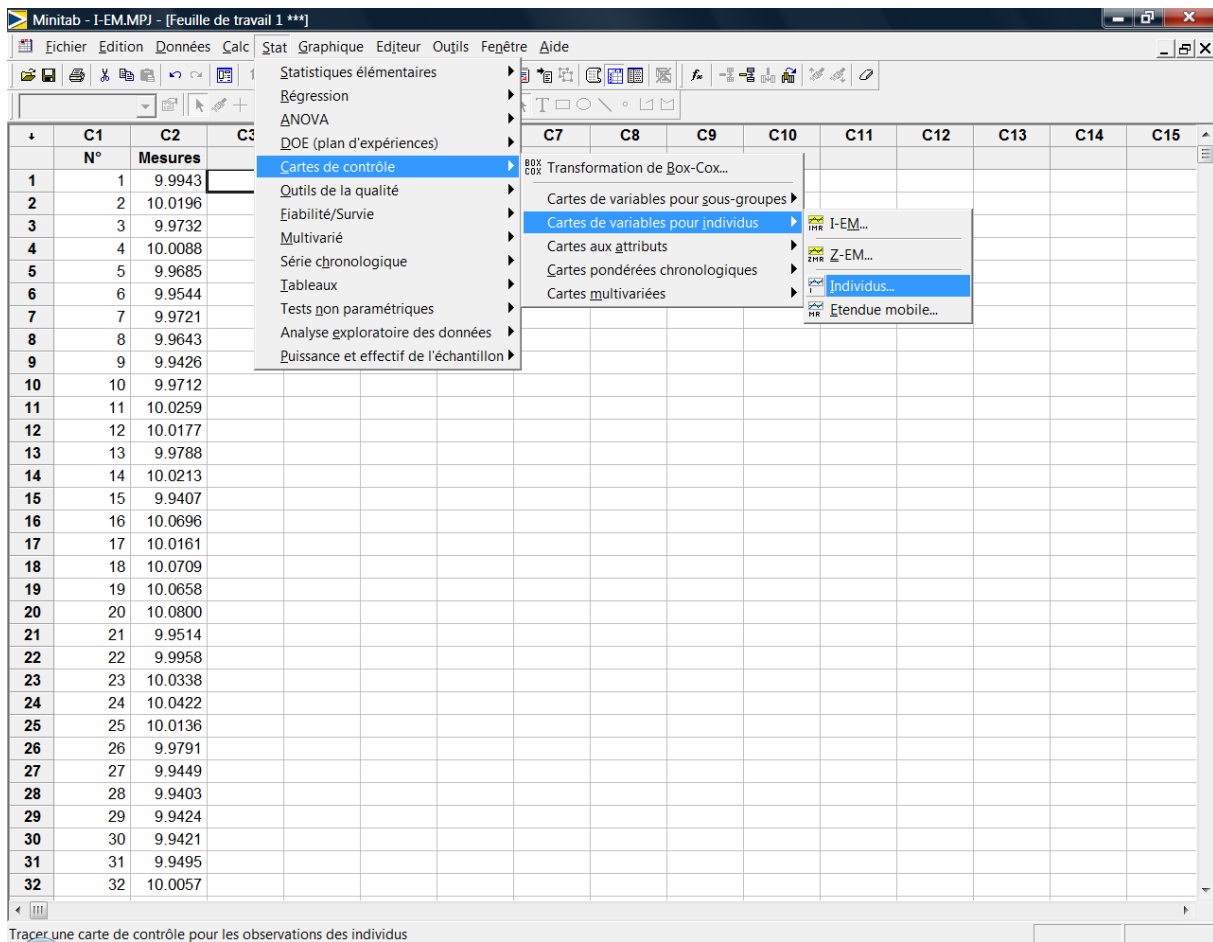
Dans le cadre de contrôle d'une production sous contrôle statistique nous avons obtenu les mesures suivantes (fichier *CC I-EM.mpj*):



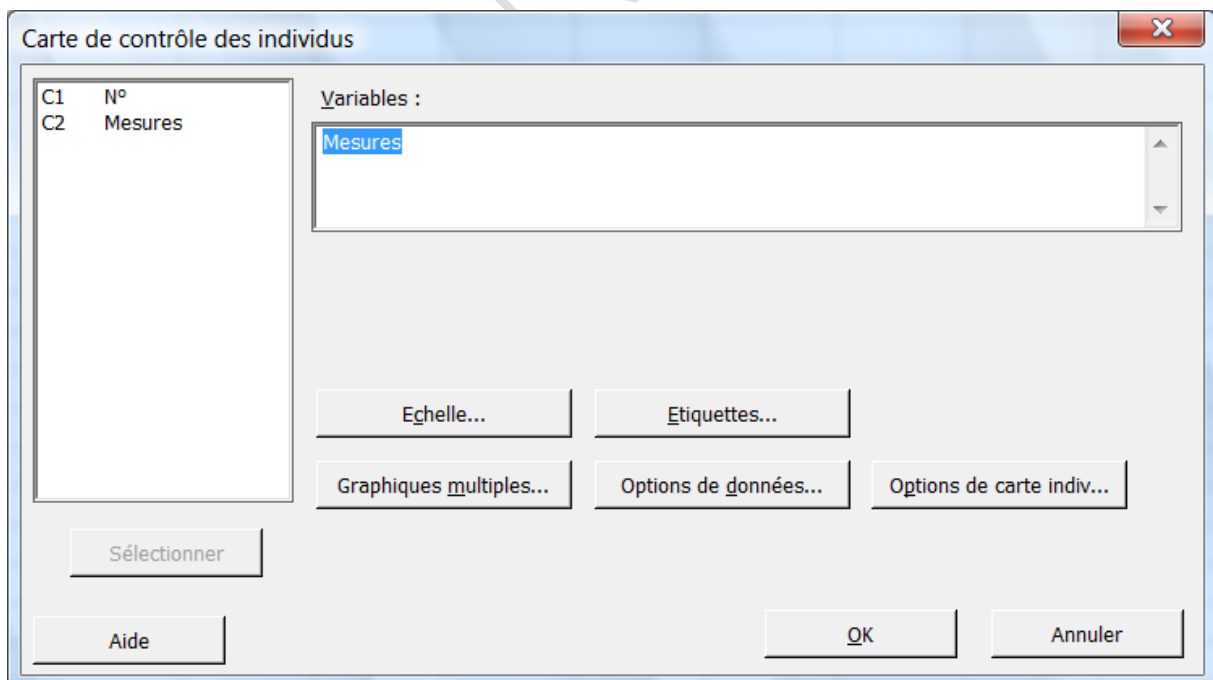
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
	N°	Mesures													
1	1	9.9943													
2	2	10.0196													
3	3	9.9732													
4	4	10.0088													
5	5	9.9685													
6	6	9.9544													
7	7	9.9721													
8	8	9.9643													
9	9	9.9426													
10	10	9.9712													
11	11	10.0259													
12	12	10.0177													
13	13	9.9788													
14	14	10.0213													
15	15	9.9407													
16	16	10.0696													
17	17	10.0161													
18	18	10.0709													
19	19	10.0658													
20	20	10.0800													
21	21	9.9514													
22	22	9.9958													
23	23	10.0338													
24	24	10.0422													
25	25	10.0136													
26	26	9.9791													
27	27	9.9449													
28	28	9.9403													
29	29	9.9424													
30	30	9.9421													
31	31	9.9495													
32	32	10.0057													

Nous souhaiterions tracer une carte de contrôle par mesures groupées de type *I-EM barre EM* et comparer avec les calculs théoriques effectués dans le cours de statistique théorique et dans le cours Microsoft Excel.

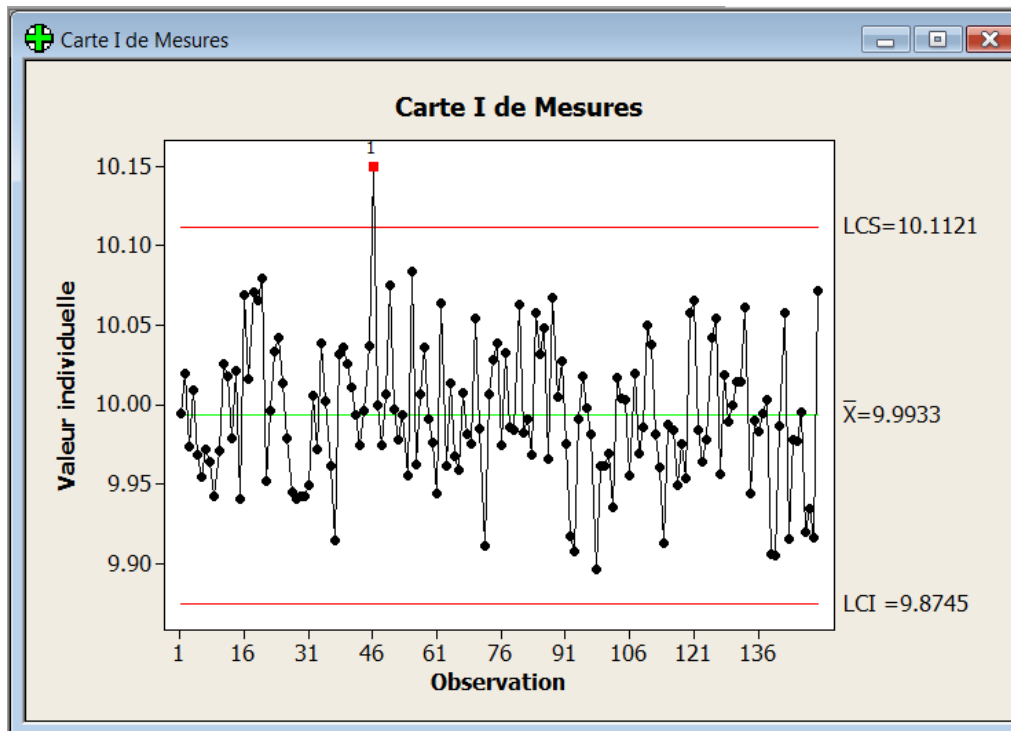
Dans Minitab® Statistical Software nous allons dans le menu **Stat/Cartes de contrôle/Cartes de variables pour individus/Individus...**:



Nous avons alors la boîte de dialogue suivante:



Nous validons par **OK**:

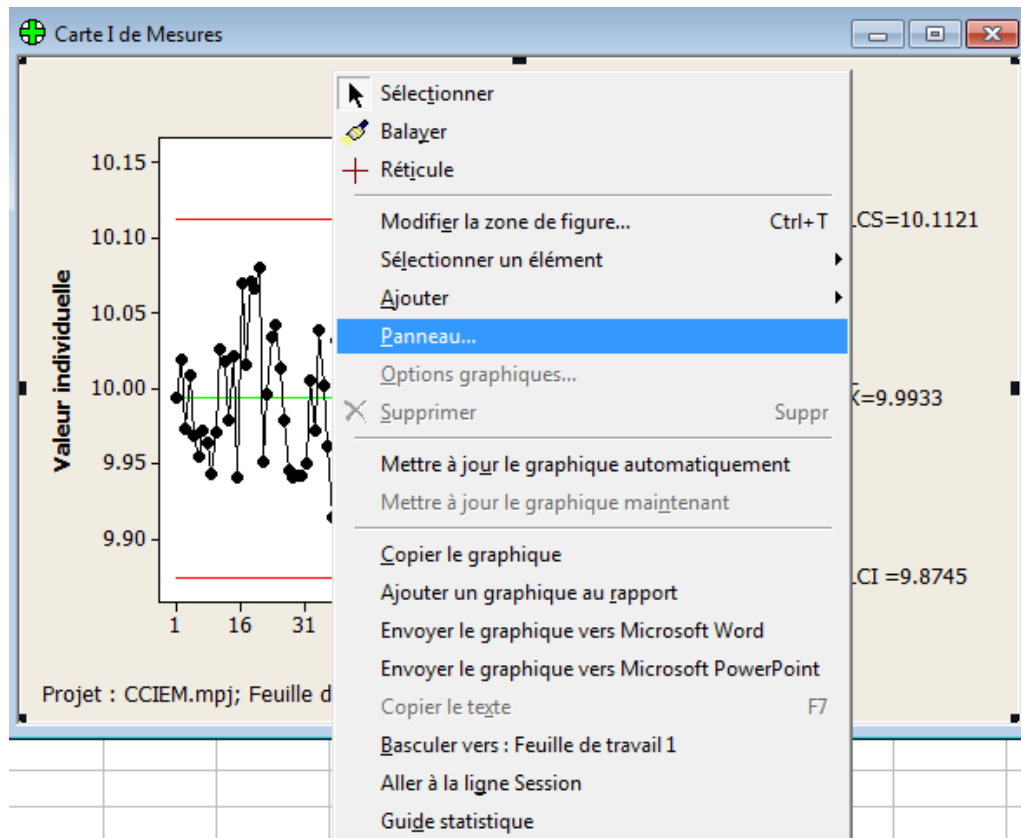


Nous obtenons donc les mêmes valeurs et la même carte que dans le cours de statistique théorique et Microsoft Excel.

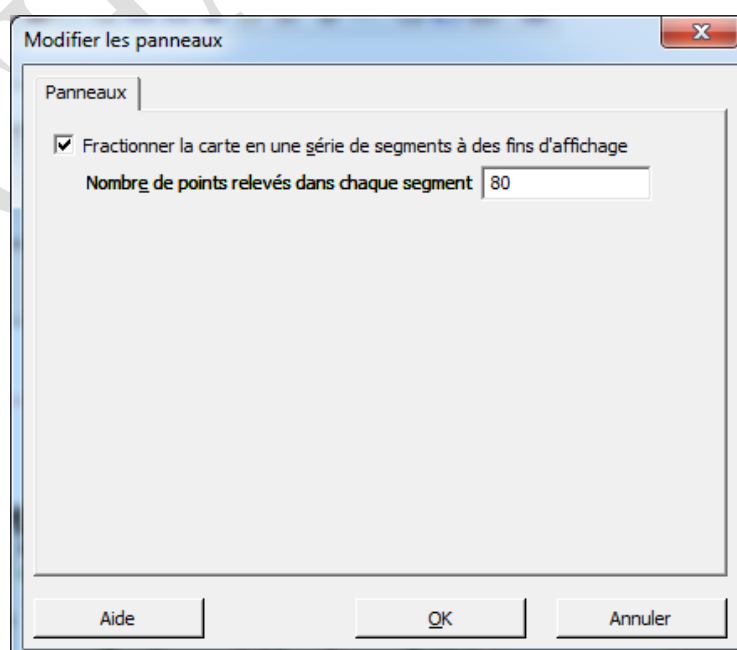
### 14.14.1. Couper une carte de contrôle quelconque

Au passage nous pouvons remarquer que cette CC est relativement dense, il serait judicieux la couper en deux morceaux. Voyons comment procéder:

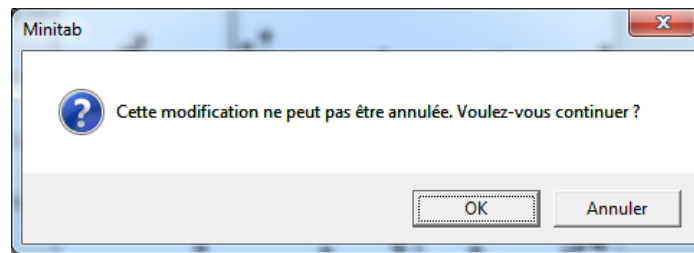
Pour ce faire, nous faisons un clic droit sur le graphique et sélectionnons l'option **Panneau...**:



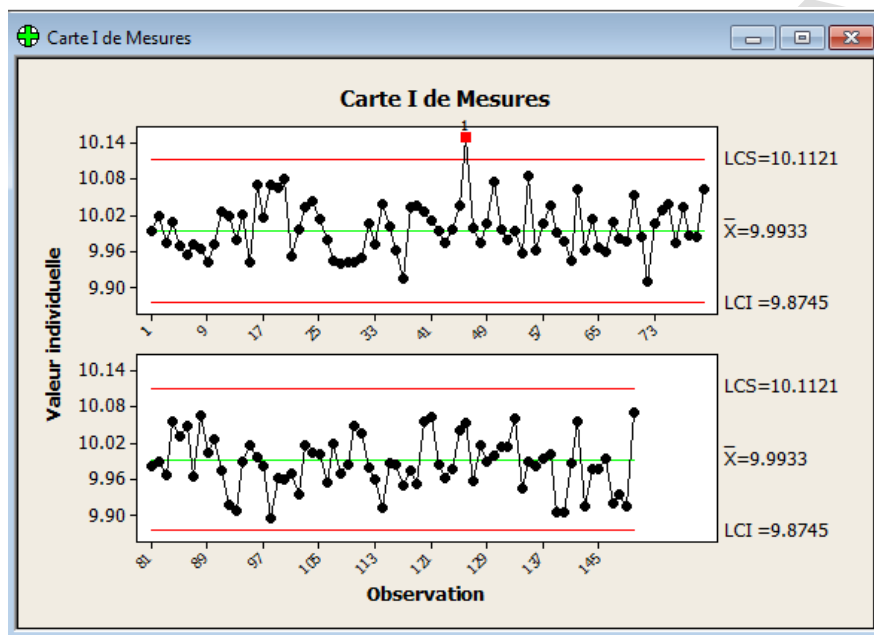
Apparaît alors la boîte de dialogue suivante:



où nous pouvons choisir le nombre de points pour le fractionnement. Si nous validons par **OK**, nous obtenons:



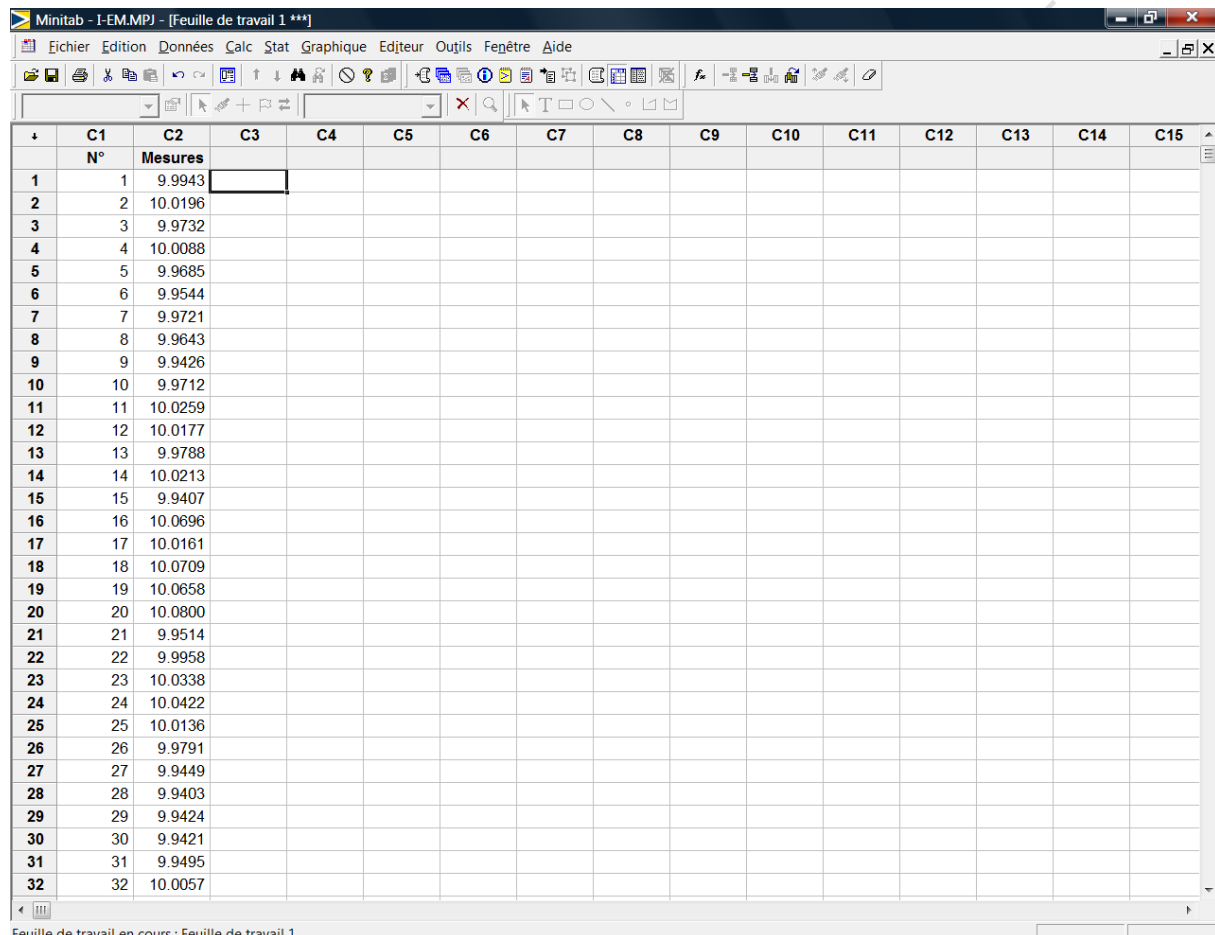
si nous confirmons, nous avons au final:



## 14.15. Exercice 133.: CC par mesures individuelles X barre par l'étendue mobile

Minitab® Statistical Software 15.1.1.0

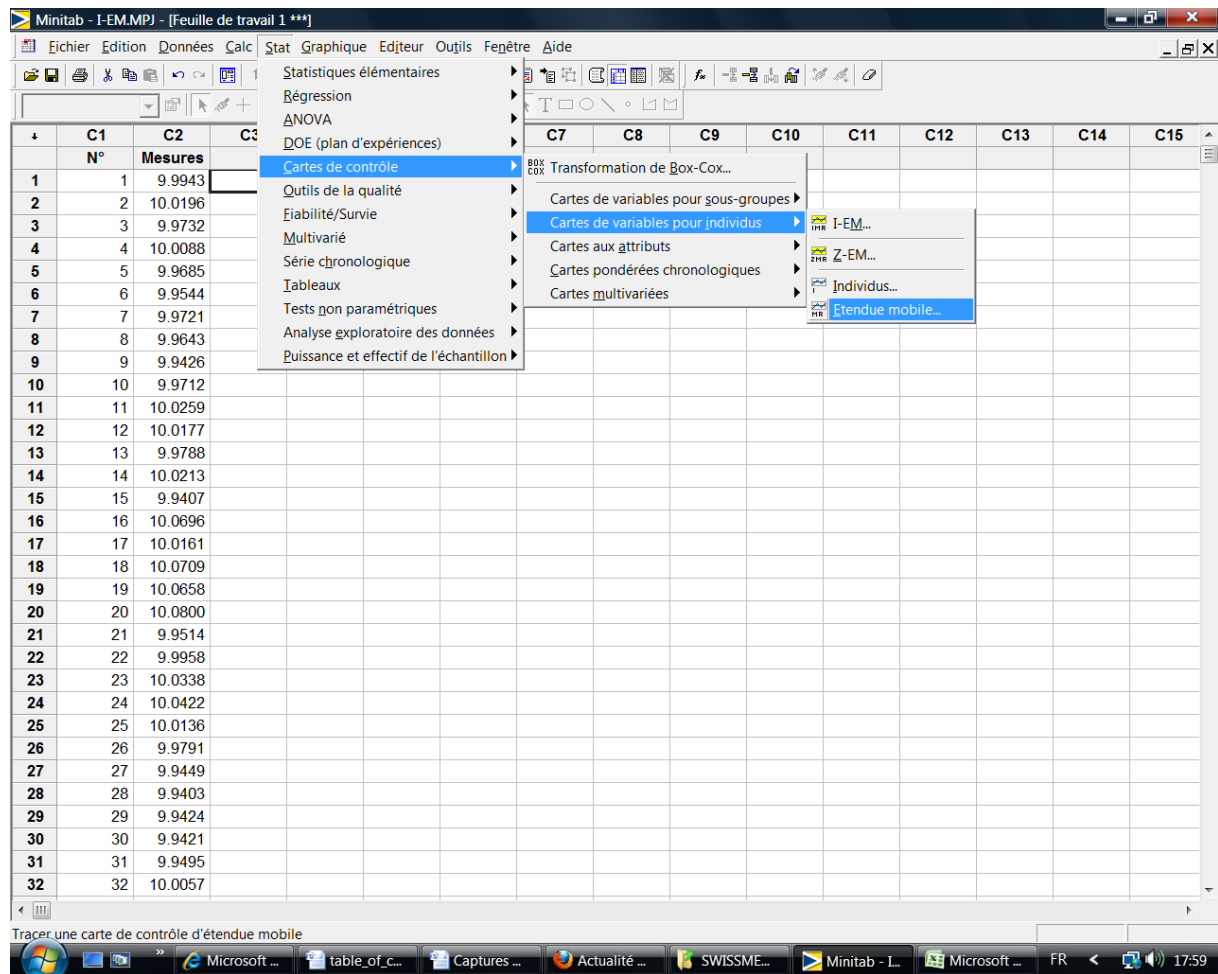
Dans le cadre de contrôle d'une production sous contrôle statistique nous avons obtenu les mesures suivantes (fichier *CC I-EM.mpj*):



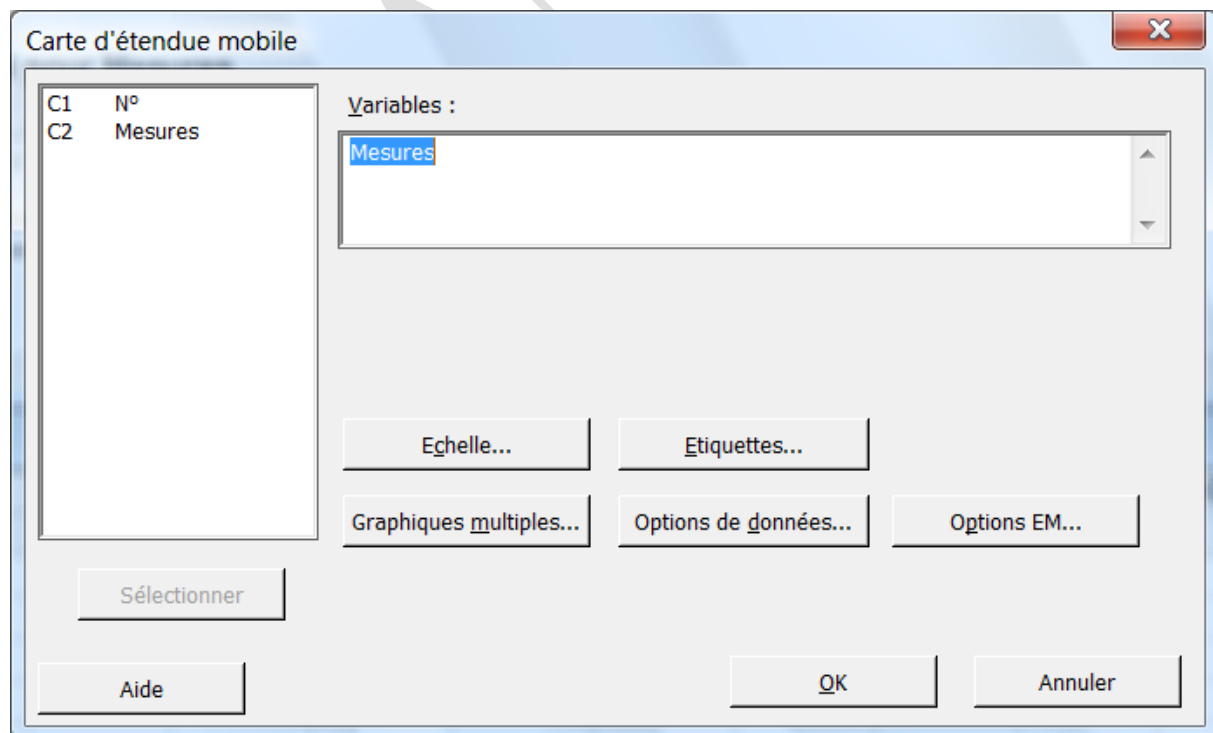
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
	N°	Mesures													
1	1	9.9943													
2	2	10.0196													
3	3	9.9732													
4	4	10.0088													
5	5	9.9685													
6	6	9.9544													
7	7	9.9721													
8	8	9.9643													
9	9	9.9426													
10	10	9.9712													
11	11	10.0259													
12	12	10.0177													
13	13	9.9788													
14	14	10.0213													
15	15	9.9407													
16	16	10.0696													
17	17	10.0161													
18	18	10.0709													
19	19	10.0658													
20	20	10.0800													
21	21	9.9514													
22	22	9.9958													
23	23	10.0338													
24	24	10.0422													
25	25	10.0136													
26	26	9.9791													
27	27	9.9449													
28	28	9.9403													
29	29	9.9424													
30	30	9.9421													
31	31	9.9495													
32	32	10.0057													

Nous souhaiterions tracer une carte de contrôle par mesures groupées de type *I-EM barre EM* et comparer avec les calculs théoriques effectués dans le cours de statistique théorique et dans le cours Microsoft Excel.

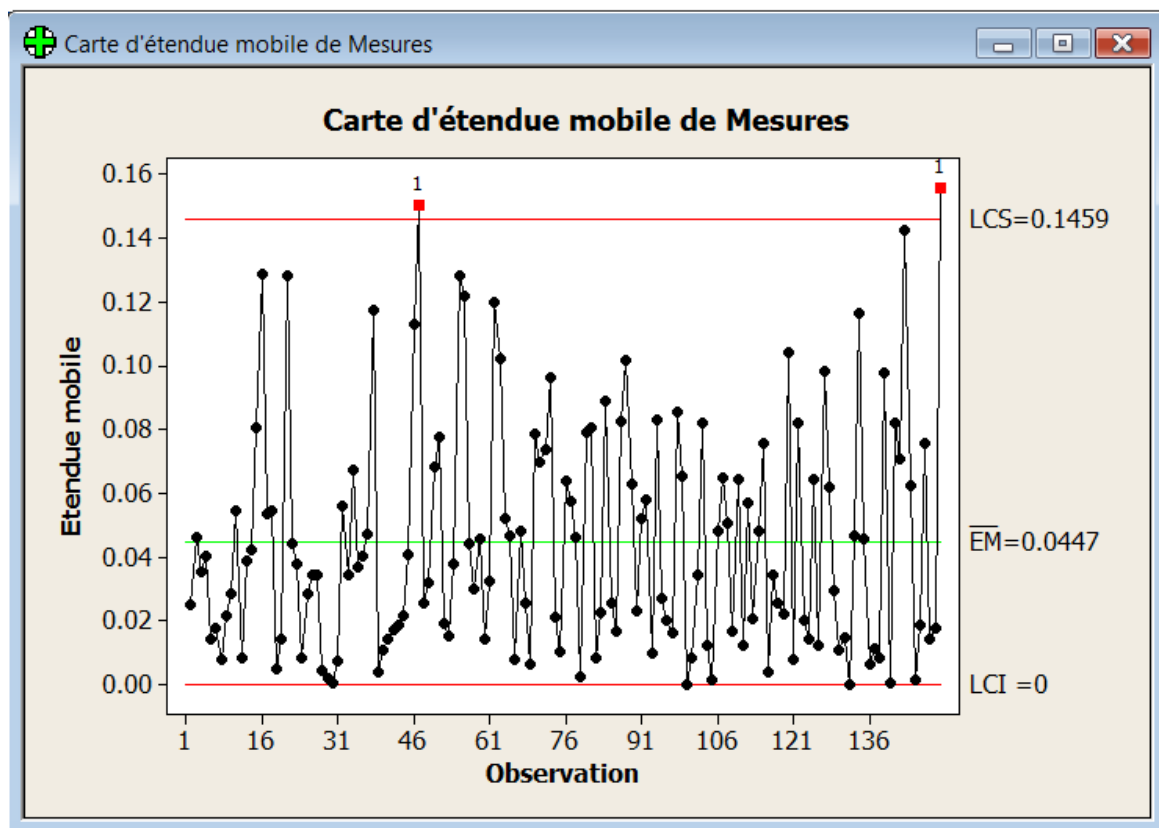
Dans Minitab® Statistical Software nous allons dans le menu **Stat/Cartes de contrôle/Cartes de variables pour individus/Etendue Mobile...**:



Nous avons alors la boîte de dialogue suivante:



Nous validons par **OK**:



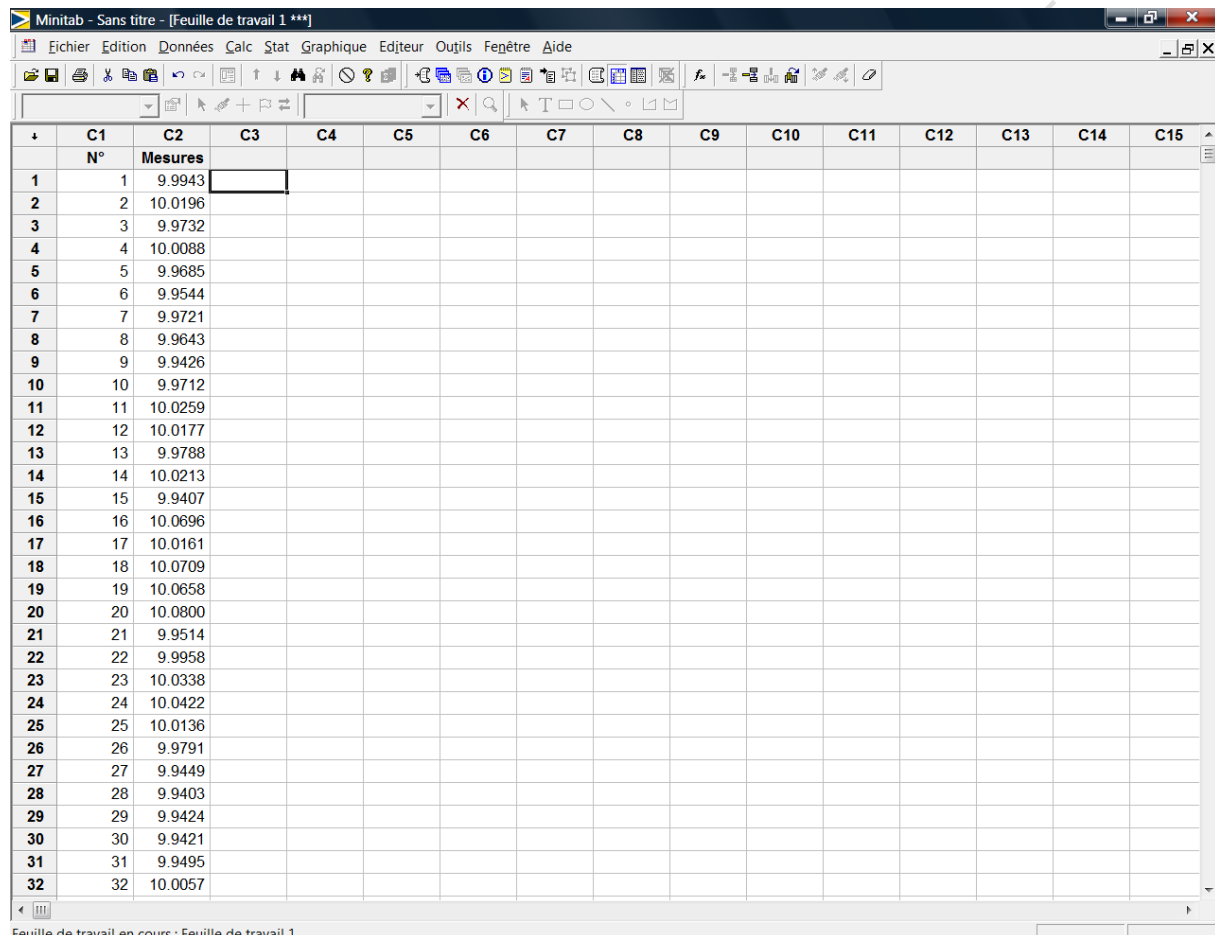
Nous obtenons donc les mêmes valeurs et la même carte que dans le cours de statistique théorique et Microsoft Excel.



## 14.16. Exercice 134.: CC par mesures individuelles X barre-EM barre combinées

Minitab® Statistical Software 15.1.1.0

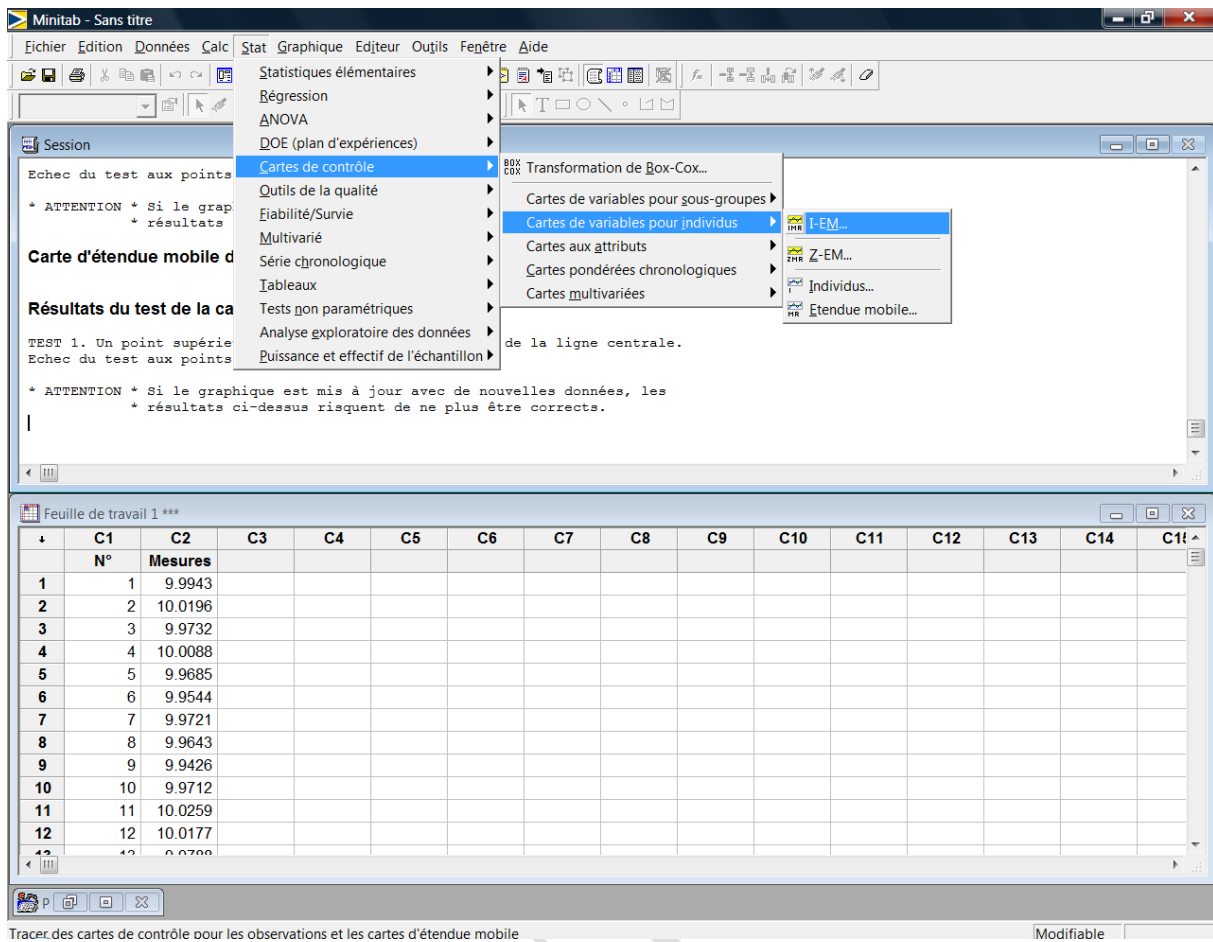
Dans le cadre de contrôle d'une production sous contrôle statistique nous avons obtenu les mesures suivantes (fichier *CC CC I-EM.mpj*):



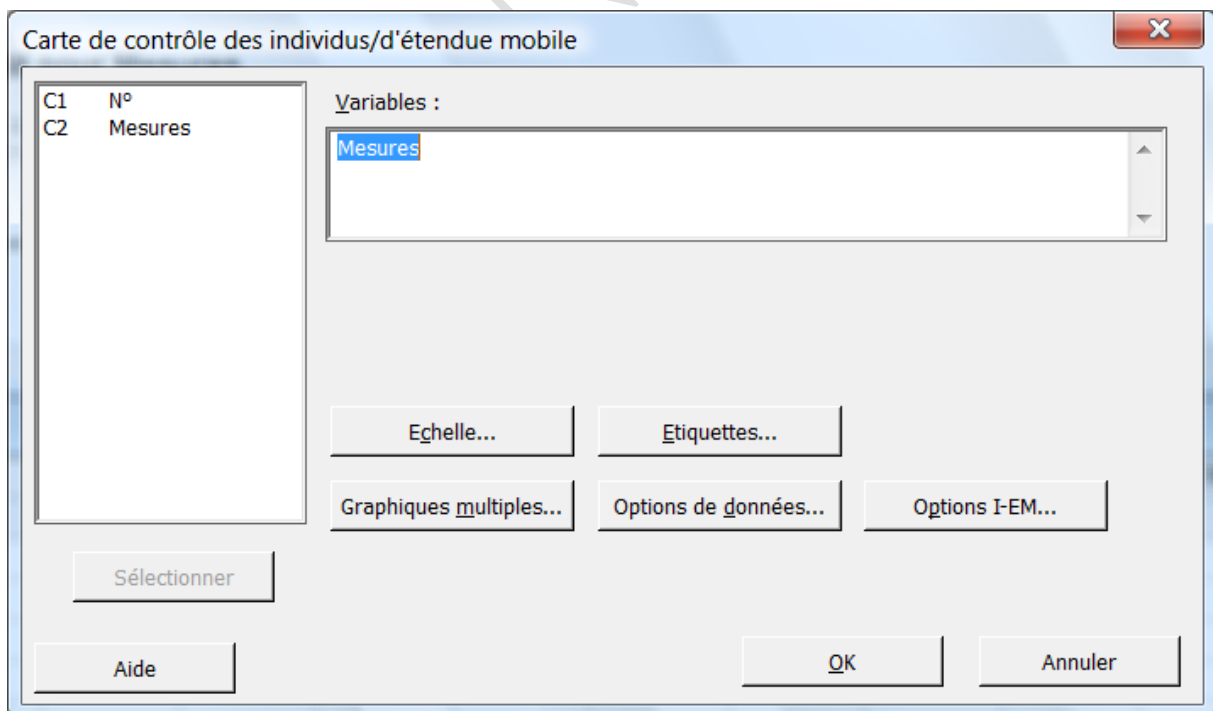
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
	N°	Mesures													
1	1	9.9943													
2	2	10.0196													
3	3	9.9732													
4	4	10.0088													
5	5	9.9685													
6	6	9.9544													
7	7	9.9721													
8	8	9.9643													
9	9	9.9426													
10	10	9.9712													
11	11	10.0259													
12	12	10.0177													
13	13	9.9788													
14	14	10.0213													
15	15	9.9407													
16	16	10.0696													
17	17	10.0161													
18	18	10.0709													
19	19	10.0658													
20	20	10.0800													
21	21	9.9514													
22	22	9.9958													
23	23	10.0338													
24	24	10.0422													
25	25	10.0136													
26	26	9.9791													
27	27	9.9449													
28	28	9.9403													
29	29	9.9424													
30	30	9.9421													
31	31	9.9495													
32	32	10.0057													

Nous souhaiterions tracer une carte de contrôle par mesures groupées de type *X barre* combinée à une *EM barre* (soit les deux exercices précédents en un seul coup) et comparer avec les calculs théoriques effectués dans le cours de statistique théorique et dans le cours Microsoft Excel.

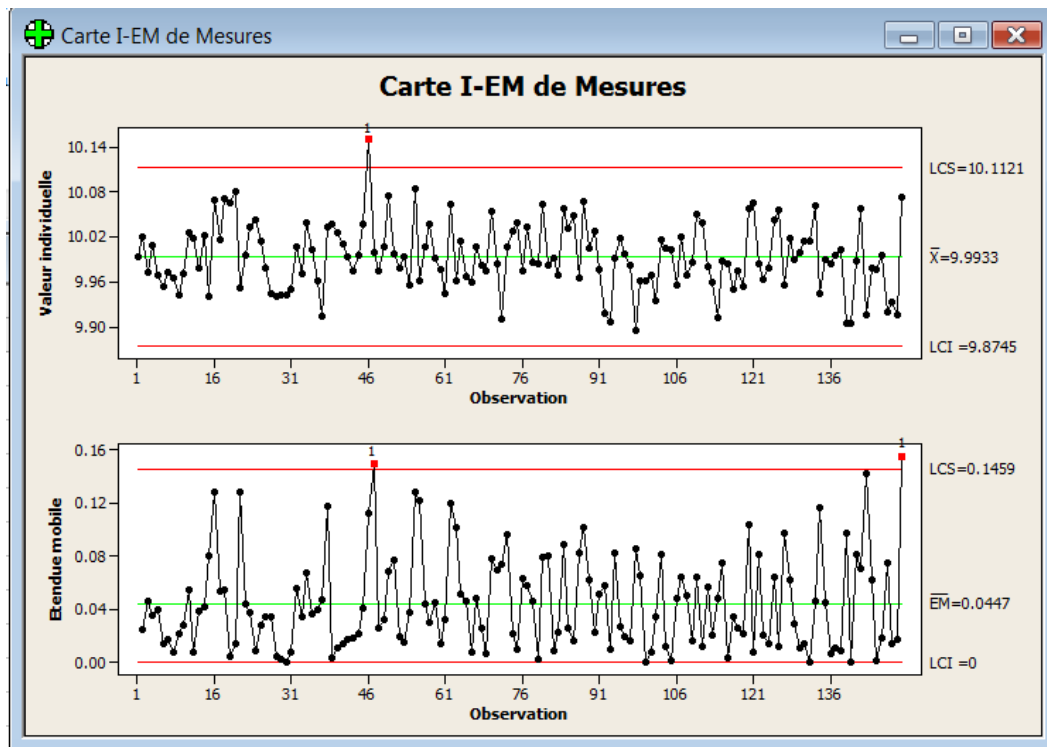
Dans Minitab® Statistical Software nous allons dans le menu **Stat/Cartes de contrôle/Cartes de variables pour sous-groupes/I-EM...**:



Nous avons alors la boîte de dialogue suivante:



Nous validons par **OK**:

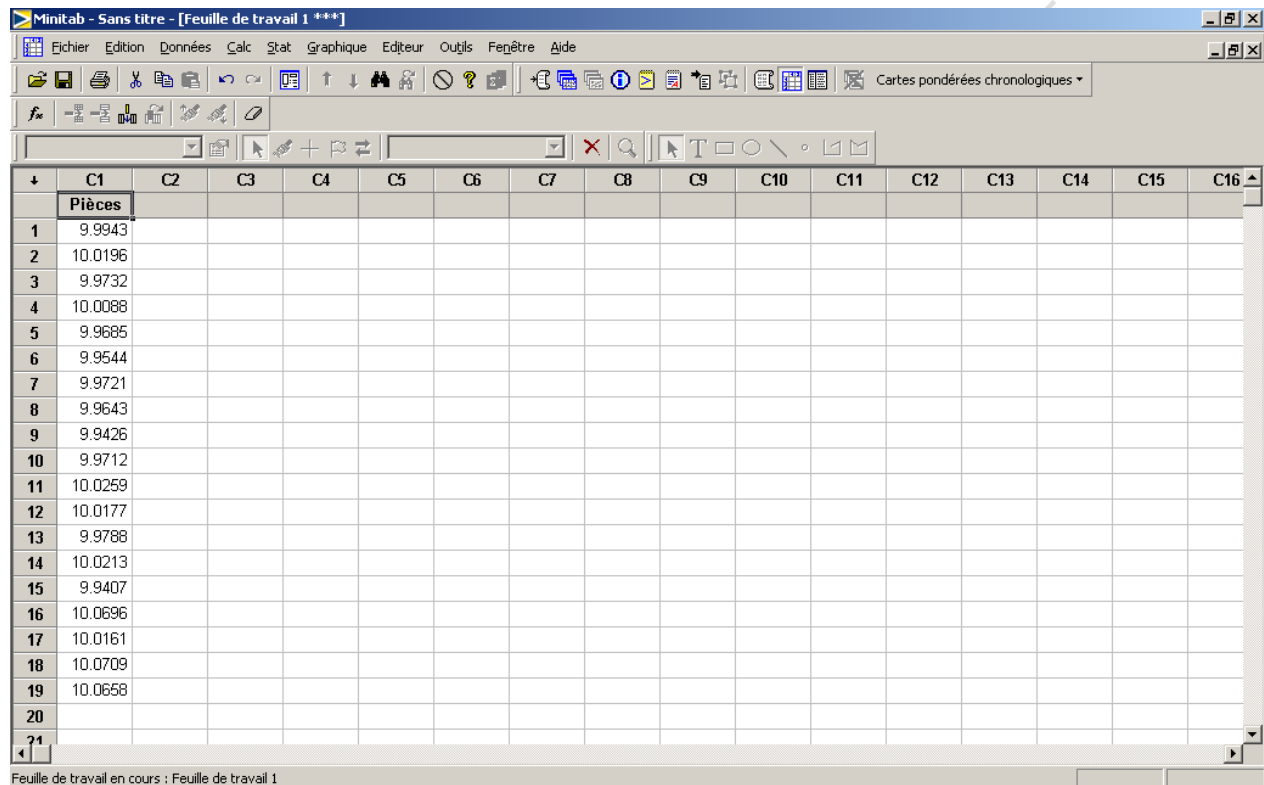


et donc nous retrouvons bien les graphiques des deux exercices précédents qui correspondent donc toujours à ce qui a été calculé dans le cours de statistique théorique et Microsoft Excel.

## 14.17. Exercice 135.: CC individuelle Moyenne Mobile (MA) avec limites basées sur l'étendue mobile

Minitab® Statistical Software 15.1.1.0

Dans le cadre de contrôle d'une production sous contrôle statistique nous avons obtenu les mesures suivantes (fichier *CC MA.mpj*):

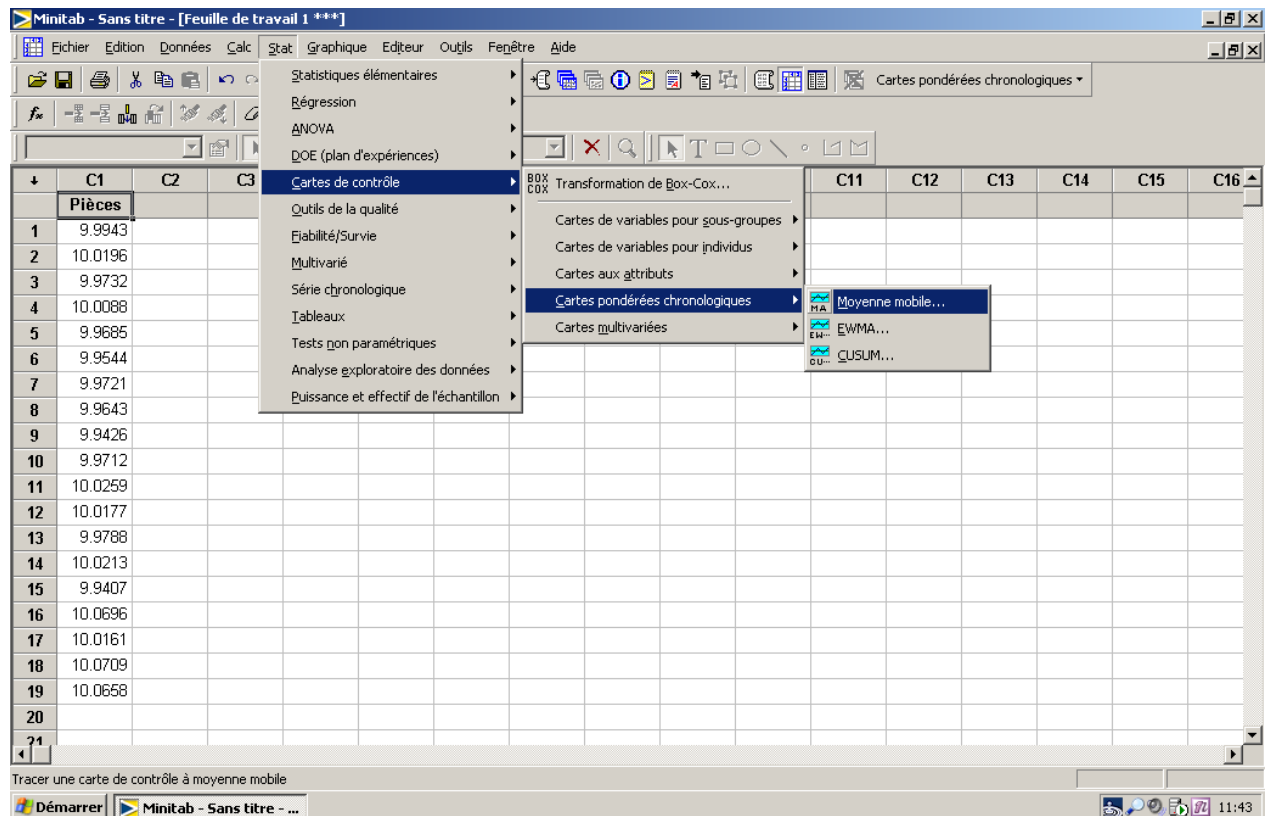


The screenshot shows the Minitab Statistical Software interface. The main window displays a data table with 19 rows of measurements in column C1. The table is titled 'Pièces' in the first row. The measurements are as follows:

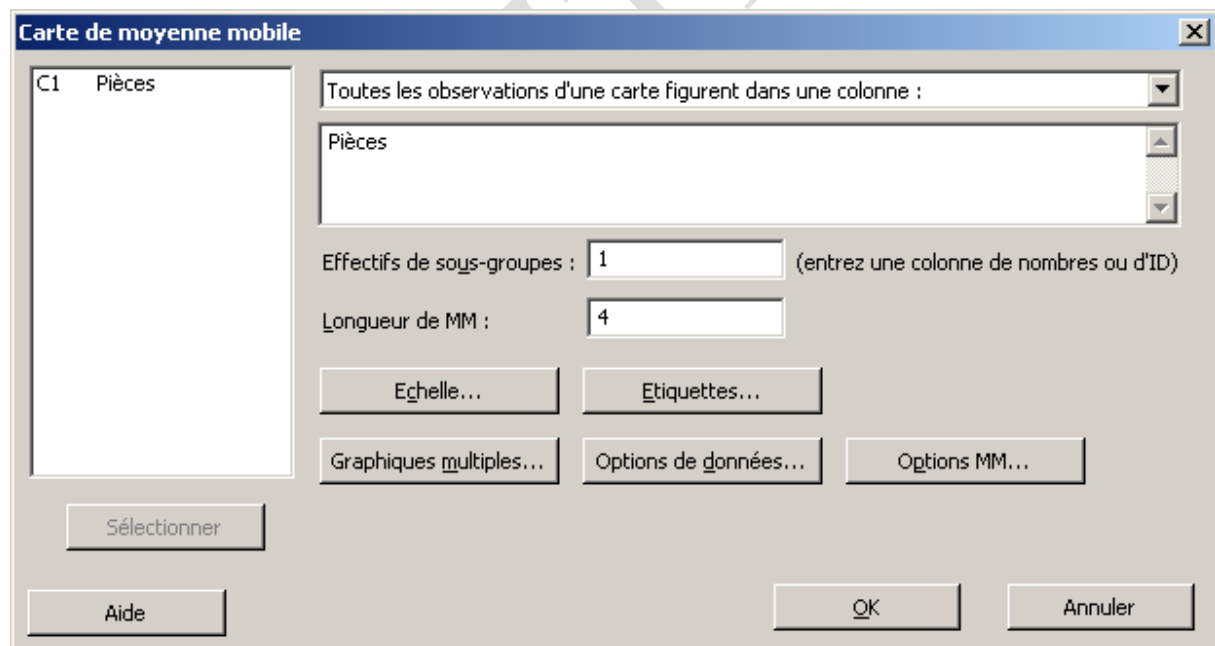
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16
	Pièces															
1	9.9943															
2	10.0196															
3	9.9732															
4	10.0088															
5	9.9685															
6	9.9544															
7	9.9721															
8	9.9643															
9	9.9426															
10	9.9712															
11	10.0259															
12	10.0177															
13	9.9788															
14	10.0213															
15	9.9407															
16	10.0696															
17	10.0161															
18	10.0709															
19	10.0658															
20																
21																

Nous souhaiterions tracer une carte de contrôle par mesures individuelle autocorrélée de type *Moyenne mobile* avec limites basées sur l'étendue mobile et comparer avec les calculs théoriques effectués dans le cours de statistique théorique et dans le cours Microsoft Excel.

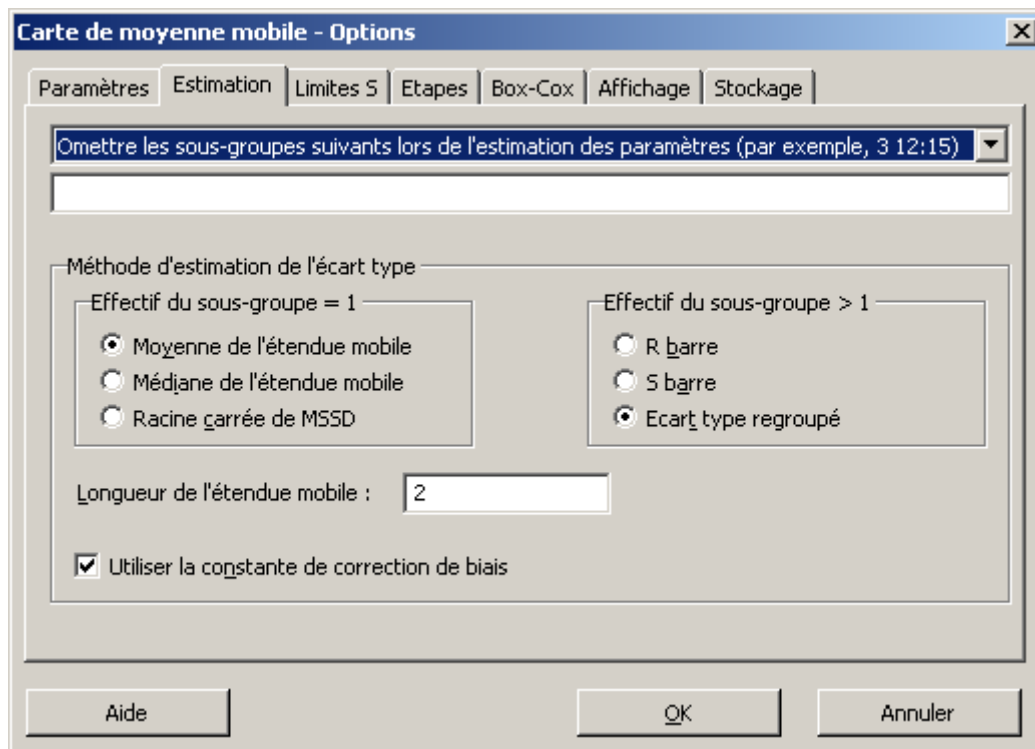
Dans Minitab® Statistical Software nous allons dans le menu **Stat/Cartes de contrôle/Cartes pondérées chronologiques/Moyenne Mobile...**:



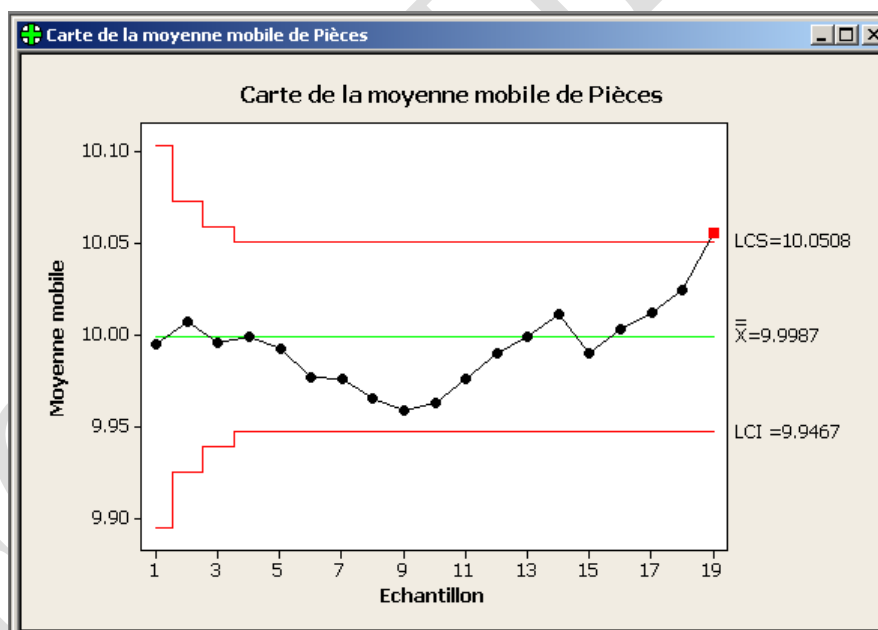
Nous avons alors la boîte de dialogue suivante:



où nous avons mis les mêmes paramètres que dans le cours Microsoft Excel et de statistiques théoriques ( $h = 4$  et effectif du sous-groupe valant 1). Nous cliquons ensuite sur **Options MM...** pour vérifier que les limites seront bien calculées sur la base de l'étendue mobile:



Nous validons deux fois par **OK** pour obtenir:

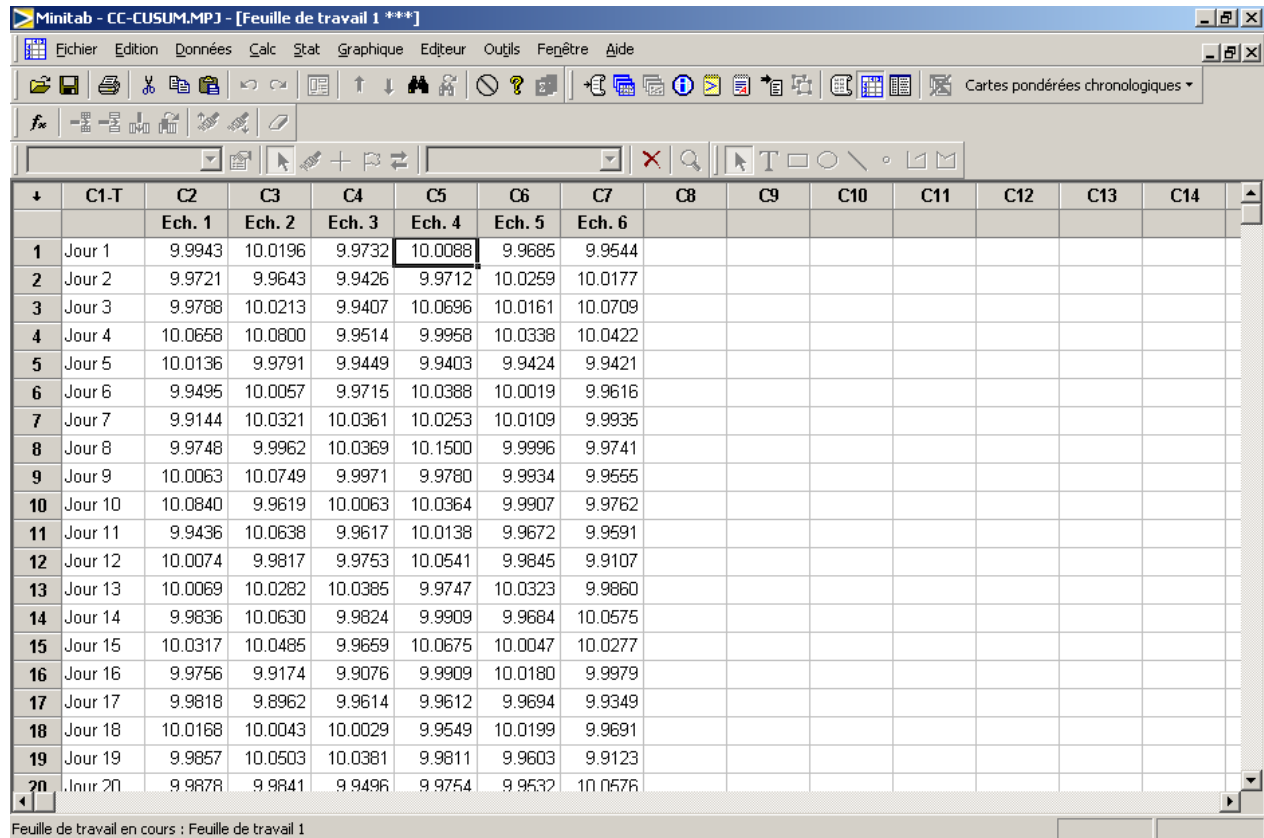


Nous obtenons donc les mêmes valeurs  $M_t$  et la même carte que dans le cours de statistique théorique et Microsoft Excel à la différence des 3 premiers points (qui selon moi ne devraient pas être présentés sur la carte).

## 14.18. Exercice 136.: CC CUSUM V-Masque avec échantillons

Minitab® Statistical Software 15.1.1.0

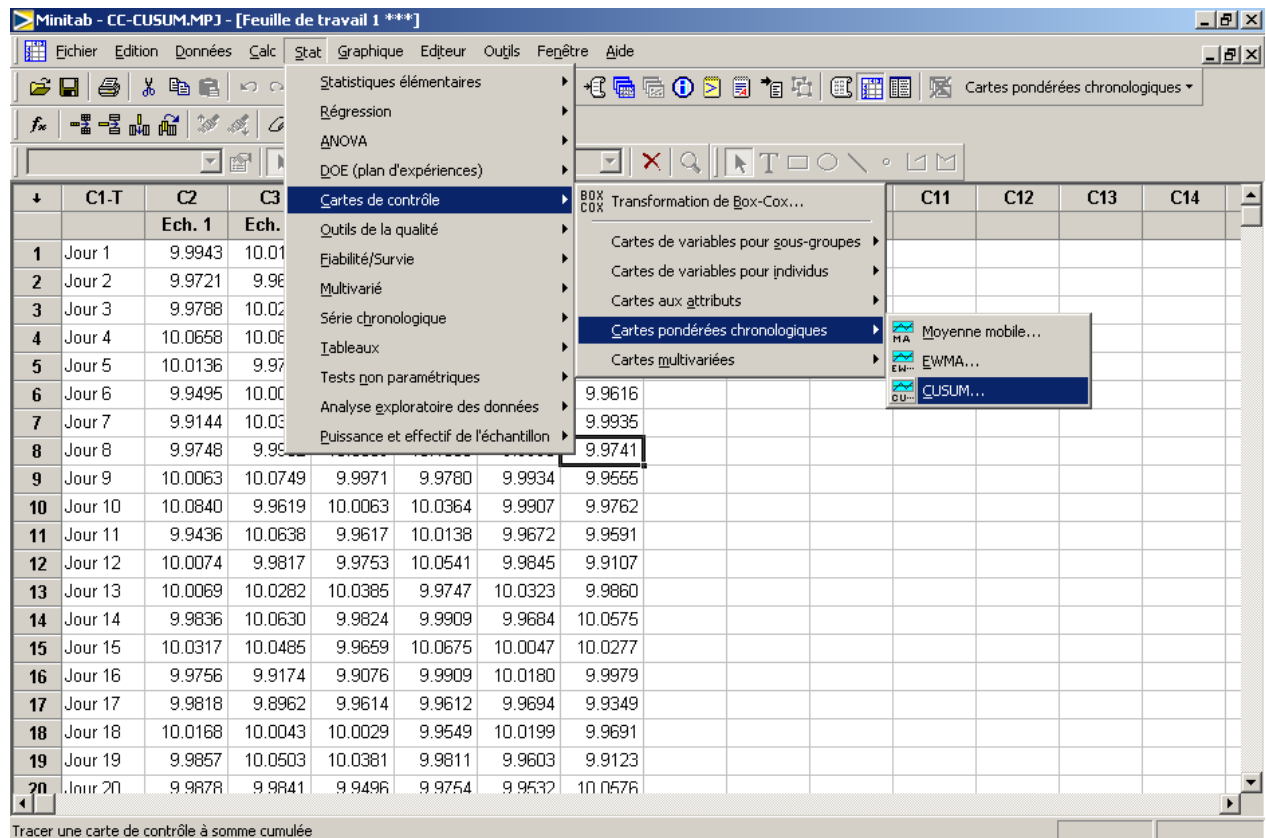
Dans le cadre de contrôle d'une production sous contrôle statistique nous avons obtenu les mesures suivantes (fichier *CC CUSUM.mpj*):



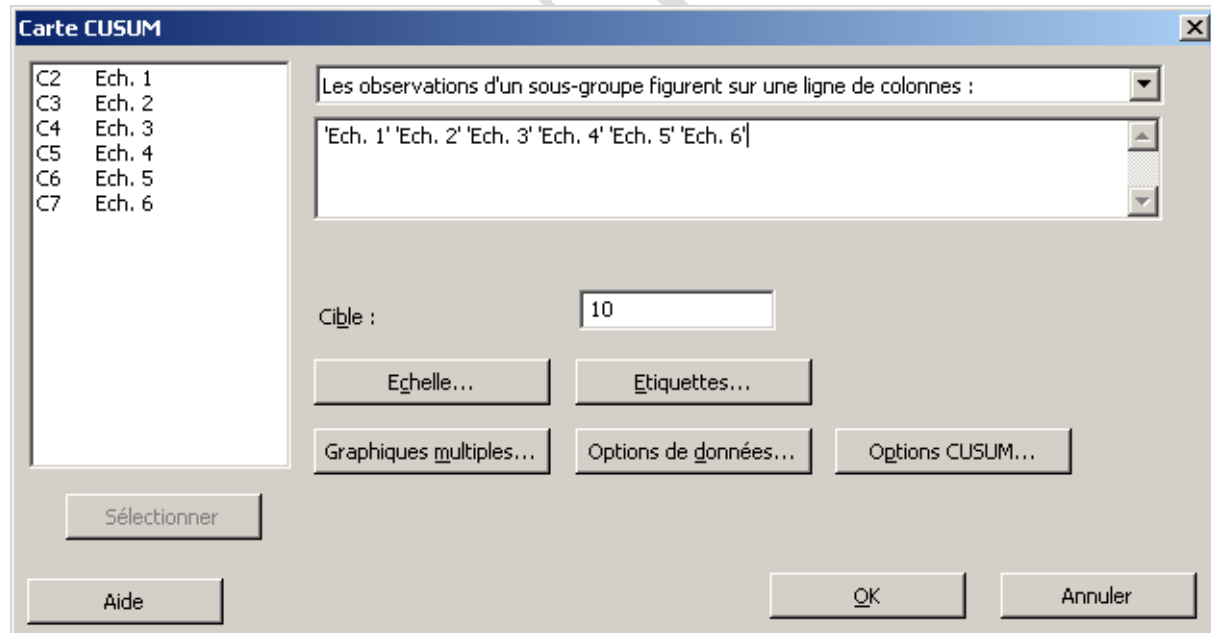
	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
		Ech. 1	Ech. 2	Ech. 3	Ech. 4	Ech. 5	Ech. 6							
1	Jour 1	9.9943	10.0196	9.9732	10.0088	9.9685	9.9544							
2	Jour 2	9.9721	9.9643	9.9426	9.9712	10.0259	10.0177							
3	Jour 3	9.9788	10.0213	9.9407	10.0696	10.0161	10.0709							
4	Jour 4	10.0658	10.0800	9.9514	9.9958	10.0338	10.0422							
5	Jour 5	10.0136	9.9791	9.9449	9.9403	9.9424	9.9421							
6	Jour 6	9.9495	10.0057	9.9715	10.0388	10.0019	9.9616							
7	Jour 7	9.9144	10.0321	10.0361	10.0253	10.0109	9.9935							
8	Jour 8	9.9748	9.9962	10.0369	10.1500	9.9996	9.9741							
9	Jour 9	10.0063	10.0749	9.9971	9.9780	9.9934	9.9555							
10	Jour 10	10.0840	9.9619	10.0063	10.0364	9.9907	9.9762							
11	Jour 11	9.9436	10.0638	9.9617	10.0138	9.9672	9.9591							
12	Jour 12	10.0074	9.9817	9.9753	10.0541	9.9845	9.9107							
13	Jour 13	10.0069	10.0282	10.0385	9.9747	10.0323	9.9860							
14	Jour 14	9.9836	10.0630	9.9824	9.9909	9.9684	10.0575							
15	Jour 15	10.0317	10.0485	9.9659	10.0675	10.0047	10.0277							
16	Jour 16	9.9756	9.9174	9.9076	9.9909	10.0180	9.9979							
17	Jour 17	9.9818	9.8962	9.9614	9.9612	9.9694	9.9349							
18	Jour 18	10.0168	10.0043	10.0029	9.9549	10.0199	9.9691							
19	Jour 19	9.9857	10.0503	10.0381	9.9811	9.9603	9.9123							
20	Jour 20	9.9878	9.9841	9.9496	9.9754	9.9532	10.0576							

Nous souhaiterions tracer une carte de contrôle autocorrélée par mesures groupées de type *CUSUM V-Masque* et comparer avec les calculs théoriques effectués dans le cours de statistique théorique et dans le cours Microsoft Excel.

Dans Minitab® Statistical Software nous allons dans le menu **Stat/Cartes de contrôle/Cartes pondérées chronologiques/CUSUM...**:

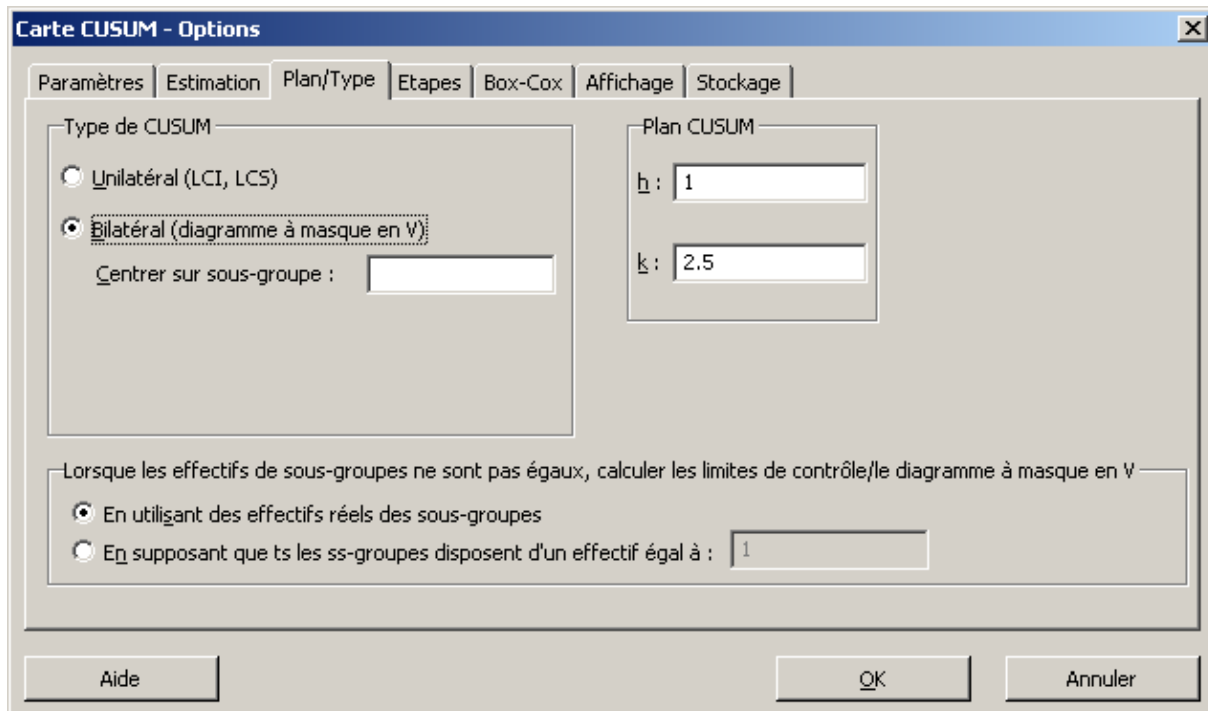


Nous avons alors la boîte de dialogue suivante:



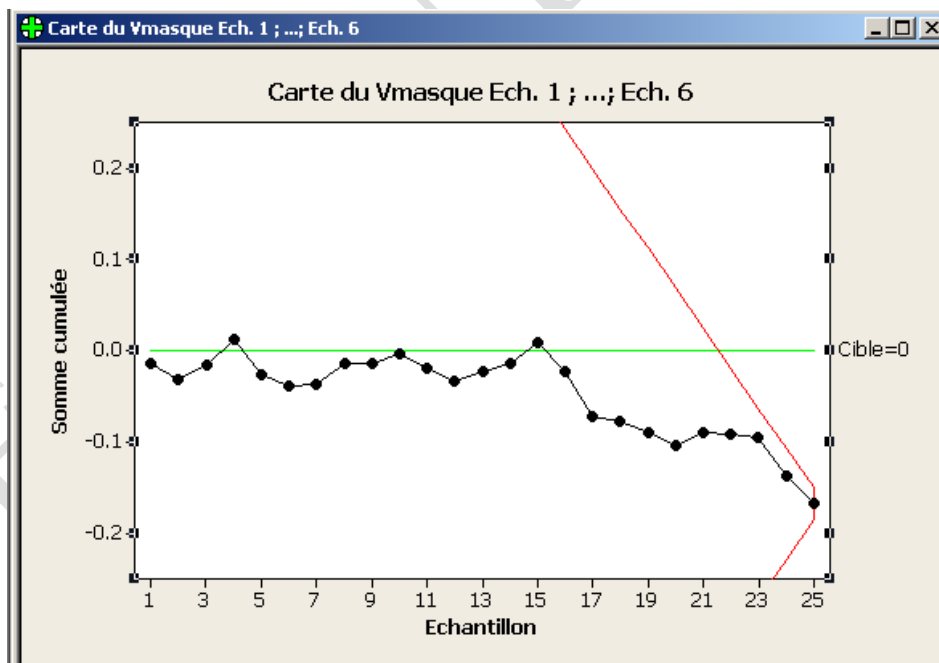
Ensuite, pour avoir une carte de contrôle de type V-Masque il faut aller cliquer sur le bouton **Options CUSUM** et aller dans l'onglet **Plan/Type**:





où nous avons par rapport au cours théorique la même valeur que  $h$ , mais  $\delta$  n'est pas demandé puisque le logiciel passe directement par le paramètre  $K$  vu dans le cours théorique et il le demande en % (les calculs à la main ont pour valeur de  $K$  0.025)

Nous avons alors en validant deux fois par **OK**:



Nous obtenons donc les mêmes valeurs  $C_i$  (un logiciel comme JMP se base lui sur les  $S_i$ ) et la même carte que dans le cours de statistique théorique et Microsoft Excel.

Depuis Minitab 17 la cible n'est plus dessinée en mode V-Masque curieusement et il faut adapter l'échelle! Comme cela devient difficilement lisible il faut alors ajouter la grille sur le graph en passant par le menu **Edition/Ajouter/Grille...**

ÉCHANTILLON

## 14.19. Exercice 137.: CC EWMA avec échantillons

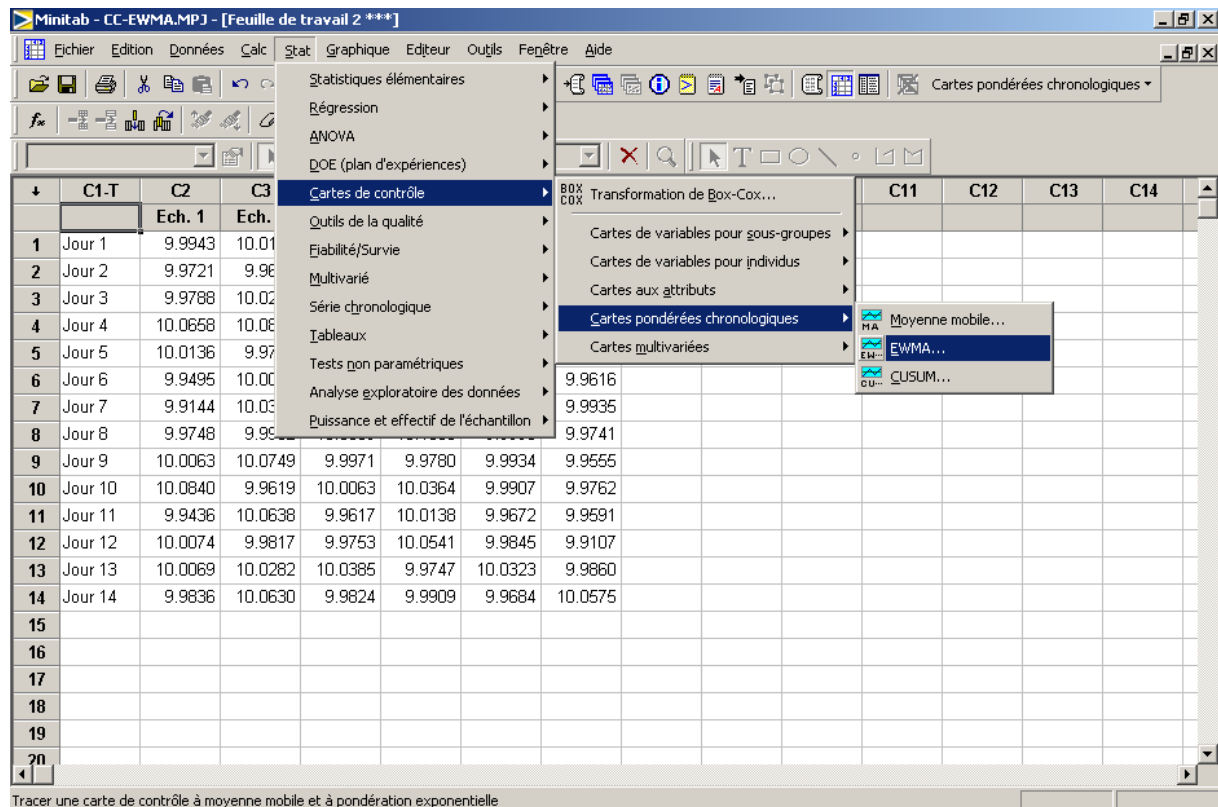
Minitab® Statistical Software 15.1.1.0

Dans le cadre de contrôle d'une production sous contrôle statistique nous avons obtenu les mesures suivantes (fichier *CC EWMA Ech.mpj*):

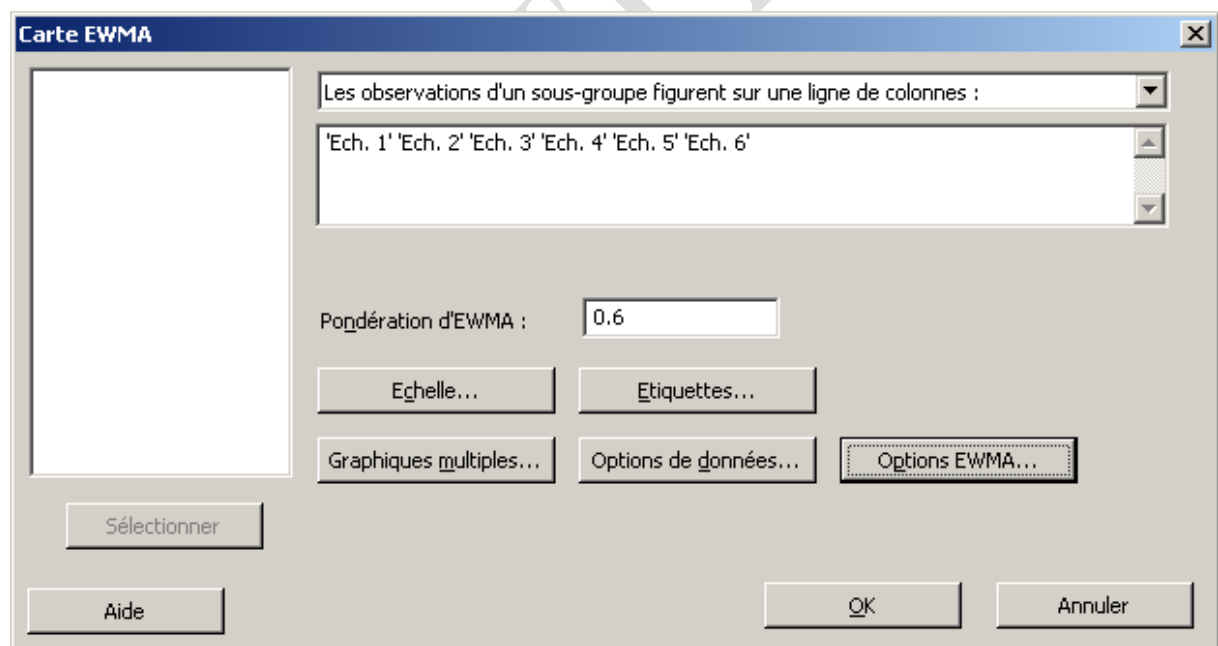
	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
		Ech. 1	Ech. 2	Ech. 3	Ech. 4	Ech. 5	Ech. 6							
1	Jour 1	9.9943	10.0196	9.9732	10.0088	9.9685	9.9544							
2	Jour 2	9.9721	9.9643	9.9426	9.9712	10.0259	10.0177							
3	Jour 3	9.9788	10.0213	9.9407	10.0696	10.0161	10.0709							
4	Jour 4	10.0658	10.0800	9.9514	9.9958	10.0338	10.0422							
5	Jour 5	10.0136	9.9791	9.9449	9.9403	9.9424	9.9421							
6	Jour 6	9.9495	10.0057	9.9715	10.0388	10.0019	9.9616							
7	Jour 7	9.9144	10.0321	10.0361	10.0253	10.0109	9.9935							
8	Jour 8	9.9748	9.9962	10.0369	10.1500	9.9996	9.9741							
9	Jour 9	10.0063	10.0749	9.9971	9.9780	9.9934	9.9555							
10	Jour 10	10.0840	9.9619	10.0063	10.0364	9.9907	9.9762							
11	Jour 11	9.9436	10.0638	9.9617	10.0138	9.9672	9.9591							
12	Jour 12	10.0074	9.9817	9.9753	10.0541	9.9845	9.9107							
13	Jour 13	10.0069	10.0282	10.0385	9.9747	10.0323	9.9860							
14	Jour 14	9.9836	10.0630	9.9824	9.9909	9.9684	10.0575							
15														
16														
17														
18														
19														
20														

Nous souhaiterions tracer une carte de contrôle autocorrélée par mesures groupées de type *EWMA* avec groupes d'échantillons et comparer avec les calculs théoriques effectués dans le cours de statistique théorique et dans le cours Microsoft Excel.

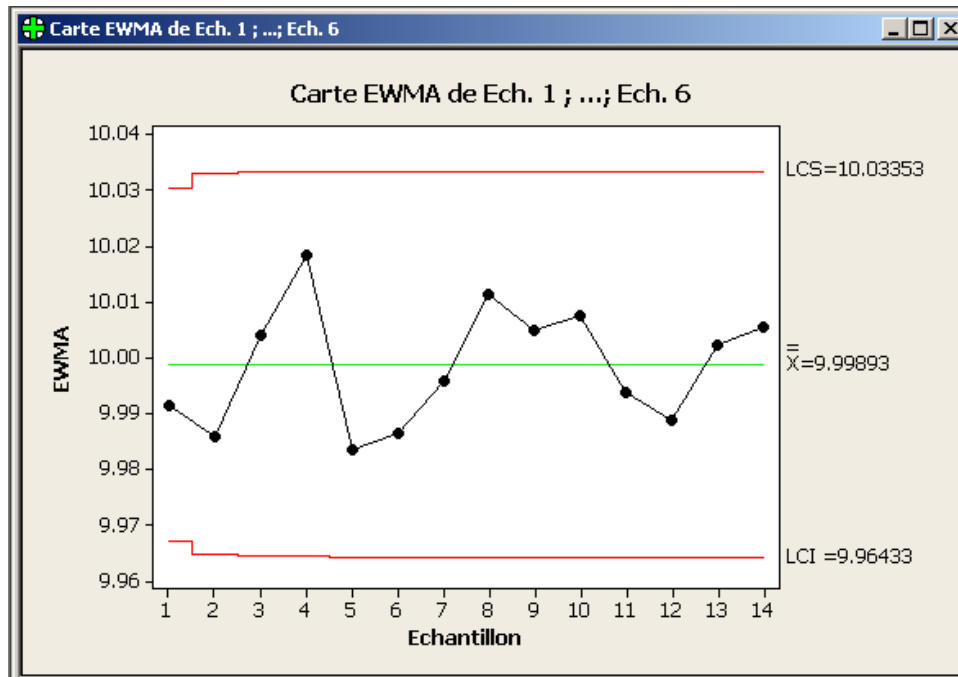
Dans Minitab® Statistical Software nous allons dans le menu **Stat/Cartes de contrôle/Cartes pondérées chronologiques/EWMA...**:



Nous avons alors la boîte de dialogue suivante:



Nous validons par **OK**:

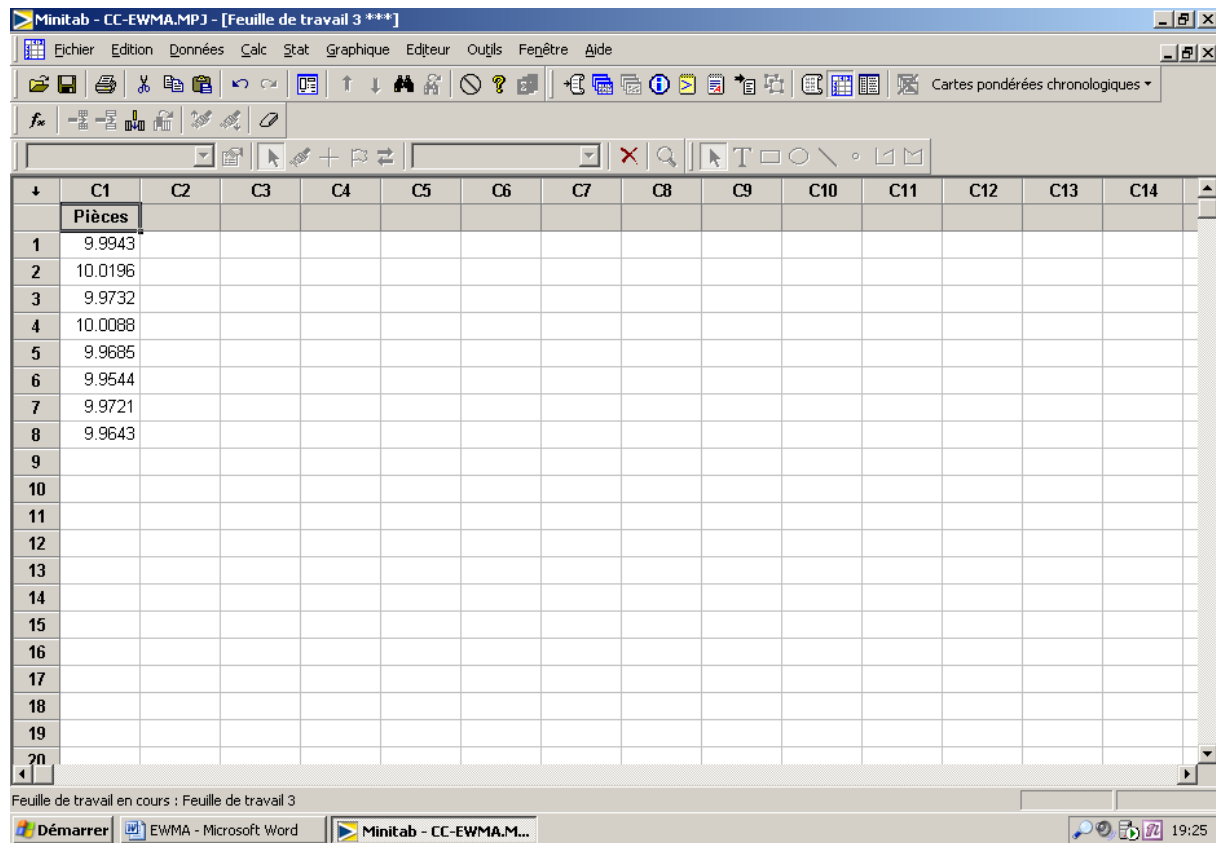


Nous retrouvons donc quasiment les mêmes valeurs que dans le cours théorique de statistiques et que le cours Microsoft Excel. Effectivement, nous y avons fait une approximation en ce qui concerne les limites de contrôle (indépendantes du temps  $t$ ) et la valeur du point de départ d'où la petite différence.

## 14.20. Exercice 138.: CC EWMA individuelle

Minitab® Statistical Software 15.1.1.0

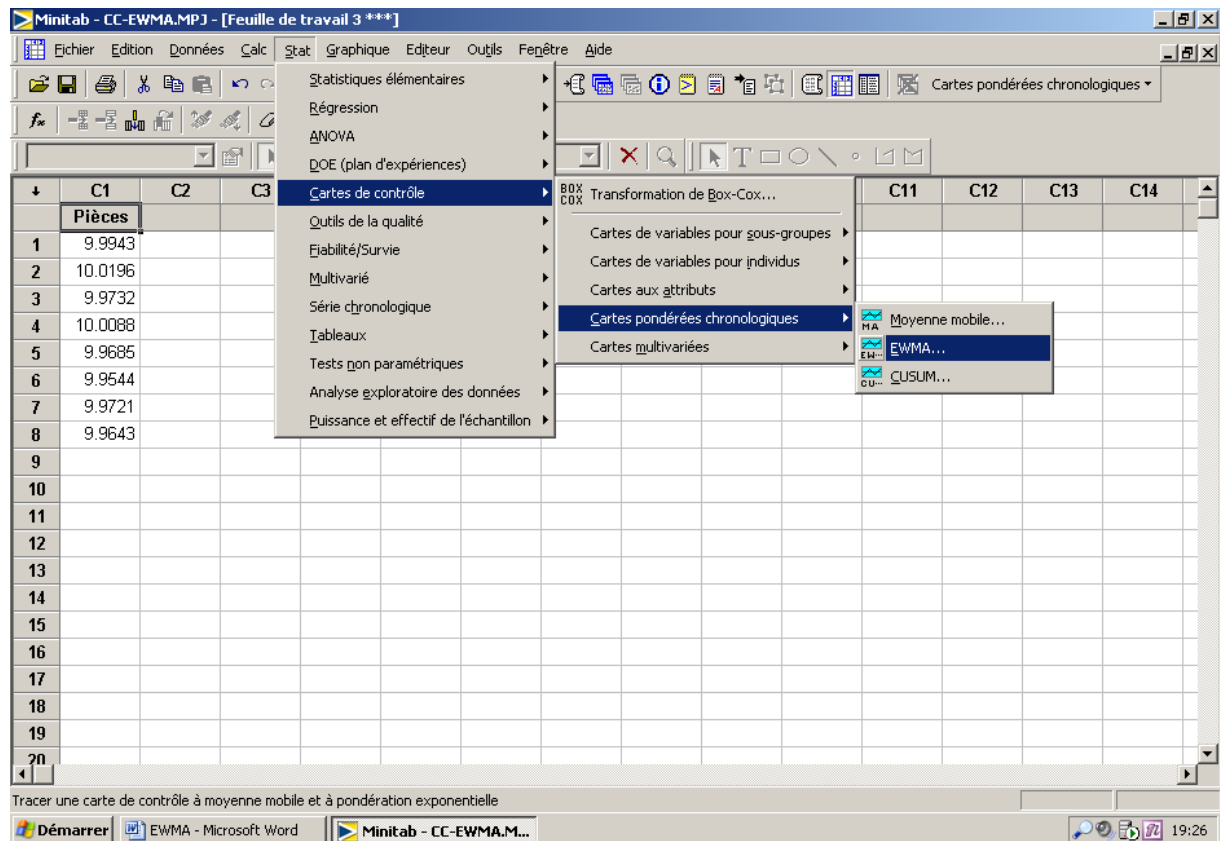
Dans le cadre de contrôle d'une production sous contrôle statistique nous avons obtenu les mesures suivantes (fichier *CC EWMA Ind.mpj*):



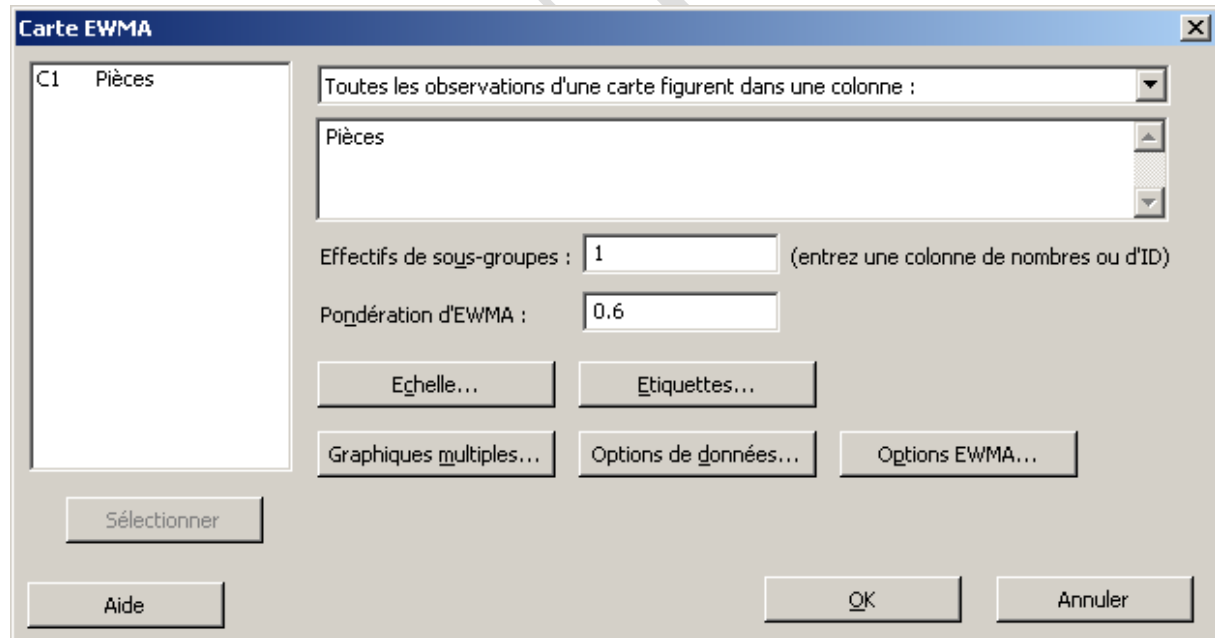
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Pièces													
1	9.9943													
2	10.0196													
3	9.9732													
4	10.0088													
5	9.9685													
6	9.9544													
7	9.9721													
8	9.9643													
9														
10														
11														
12														
13														
14														
15														
16														
17														
18														
19														
20														

Nous souhaiterions tracer une carte de contrôle autocorrélée par mesures individuelles de type *EWMA* et comparer avec les calculs théoriques effectués dans le cours de statistique théorique et dans le cours Microsoft Excel.

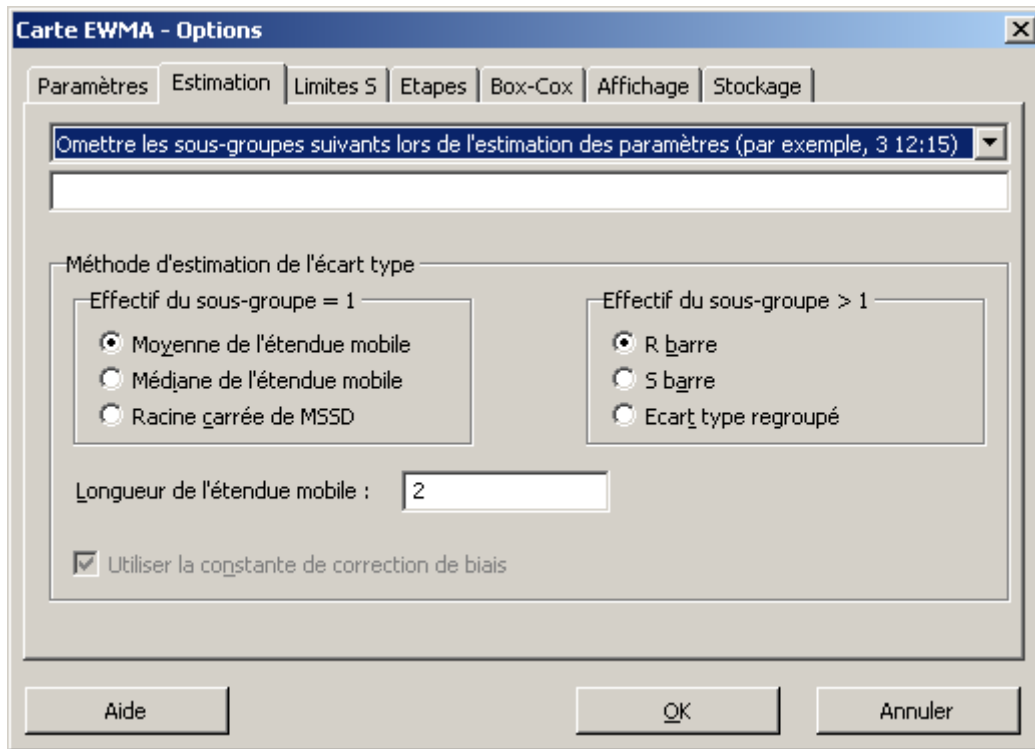
Dans Minitab® Statistical Software nous allons dans le menu **Stat/Cartes de contrôle/Cartes pondérées chronologiques/EWMA...**:



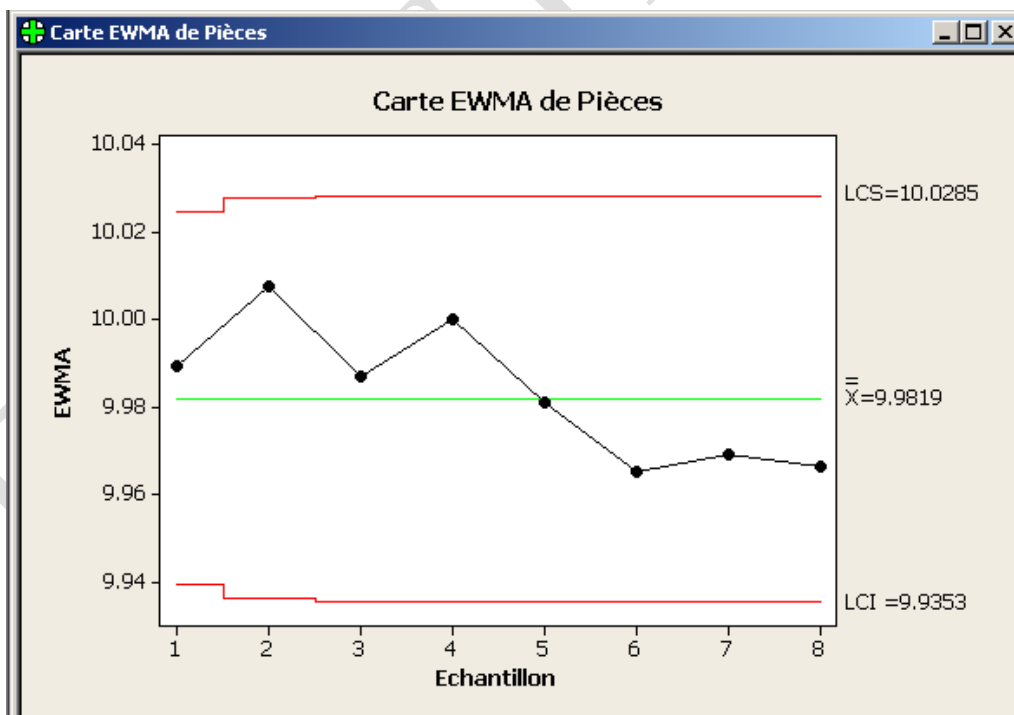
Nous avons alors la boîte de dialogue suivante:



Nous cliquons sur **Options EWMA...**:



et nous nous assurons que **Moyenne de l'étendue mobile** soit activée afin d'être conforme à ce que nous avons fait dans le cours théorique de Statistique et Microsoft Excel. Nous validons enfin deux fois par **OK**:



Nous retrouvons donc quasiment les mêmes valeurs que dans le cours théorique de statistiques et que le cours Microsoft Excel. Effectivement, nous y avons fait une approximation en ce qui concerne les limites de contrôle (indépendantes du temps  $t$ ) et la valeur du point de départ d'où la petite différence.



## 14.21. Exercice 139.: CC combinée X Barre-R et R Barre-R avec deux phases (avant/après)

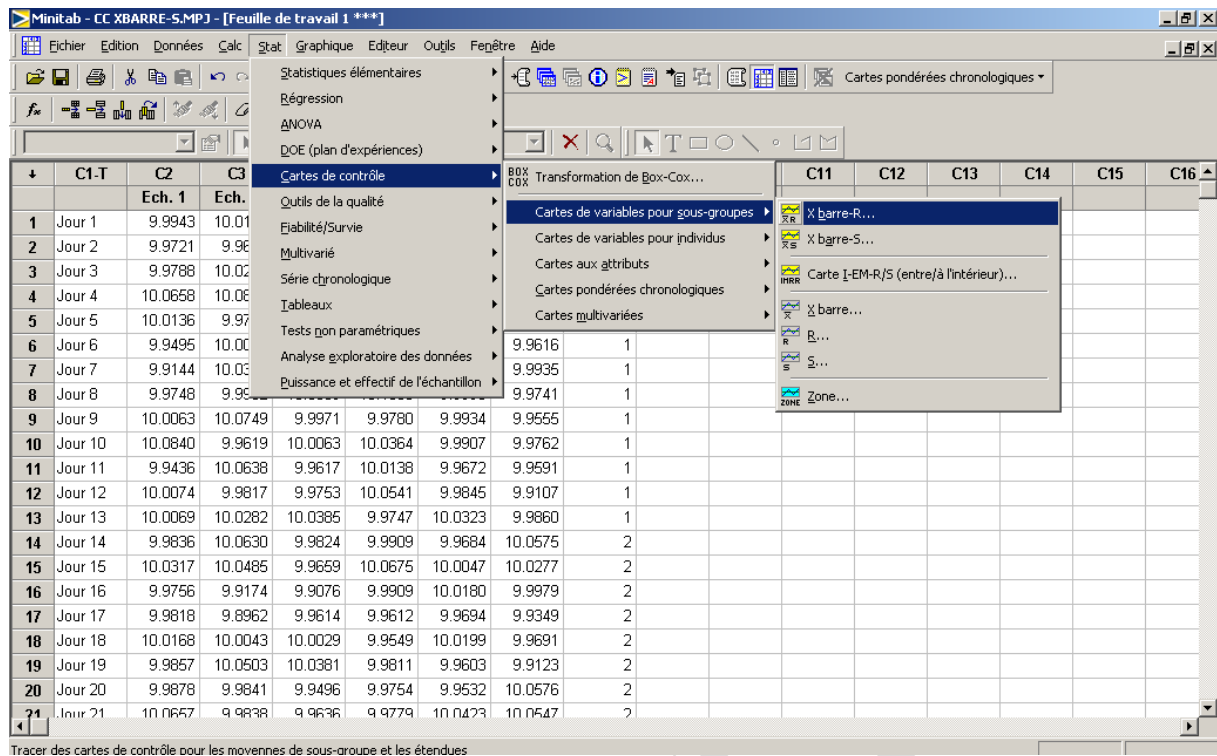
Minitab® Statistical Software 15.1.1.0

Dans le cadre de contrôle d'une production sous contrôle statistique nous avons obtenu les mesures suivantes (fichier *CC XBARRE-R Avant-After.mpj*):

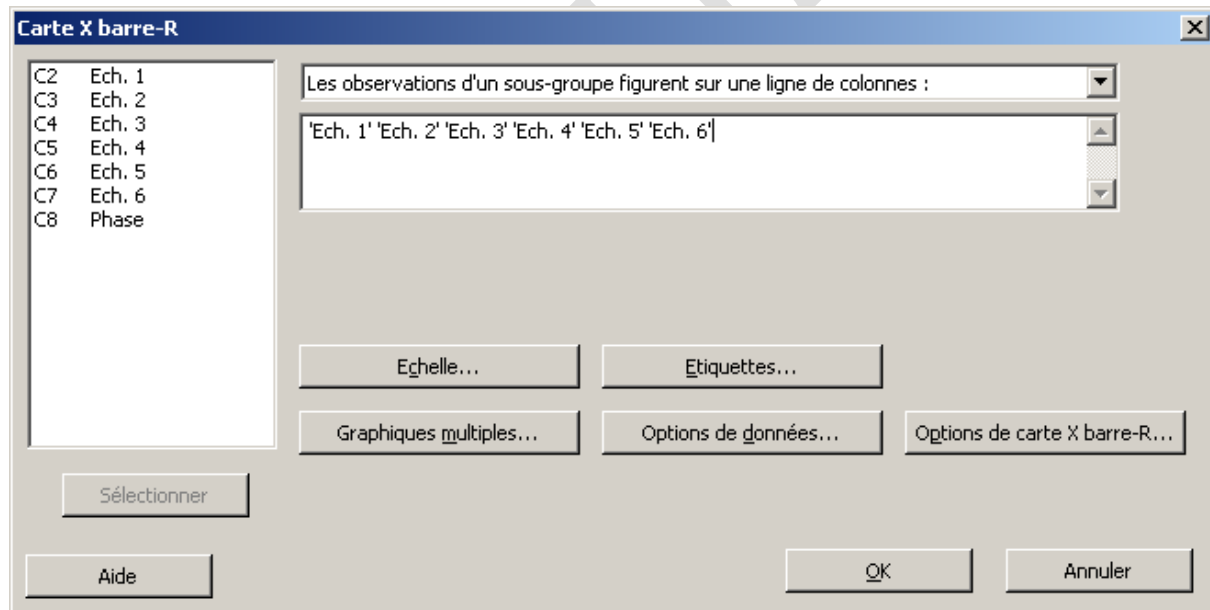
	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16
		Ech. 1	Ech. 2	Ech. 3	Ech. 4	Ech. 5	Ech. 6	Phase								
1	Jour 1	9.9943	10.0196	9.9732	10.0088	9.9685	9.9544	1								
2	Jour 2	9.9721	9.9643	9.9426	9.9712	10.0259	10.0177	1								
3	Jour 3	9.9788	10.0213	9.9407	10.0696	10.0161	10.0709	1								
4	Jour 4	10.0658	10.0800	9.9514	9.9958	10.0338	10.0422	1								
5	Jour 5	10.0136	9.9791	9.9449	9.9403	9.9424	9.9421	1								
6	Jour 6	9.9495	10.0057	9.9715	10.0388	10.0019	9.9616	1								
7	Jour 7	9.9144	10.0321	10.0361	10.0253	10.0109	9.9935	1								
8	Jour 8	9.9748	9.9962	10.0369	10.1500	9.9996	9.9741	1								
9	Jour 9	10.0063	10.0749	9.9971	9.9780	9.9934	9.9555	1								
10	Jour 10	10.0840	9.9619	10.0063	10.0364	9.9907	9.9762	1								
11	Jour 11	9.9436	10.0638	9.9617	10.0138	9.9672	9.9591	1								
12	Jour 12	10.0074	9.9817	9.9753	10.0541	9.9845	9.9107	1								
13	Jour 13	10.0069	10.0282	10.0385	9.9747	10.0323	9.9860	1								
14	Jour 14	9.9836	10.0630	9.9824	9.9909	9.9684	10.0575	2								
15	Jour 15	10.0317	10.0485	9.9659	10.0675	10.0047	10.0277	2								
16	Jour 16	9.9756	9.9174	9.9076	9.9909	10.0180	9.9979	2								
17	Jour 17	9.9818	9.8962	9.9614	9.9612	9.9694	9.9349	2								
18	Jour 18	10.0168	10.0043	10.0029	9.9549	10.0199	9.9691	2								
19	Jour 19	9.9857	10.0503	10.0381	9.9811	9.9603	9.9123	2								
20	Jour 20	9.9878	9.9841	9.9496	9.9754	9.9532	10.0576	2								
21	Jour 21	10.0657	9.9838	9.9636	9.9779	10.0473	10.0547	2								

Nous souhaiterions tracer une carte de contrôle par mesures groupées de type combinée X Barre-R et R Barre-R afin de voir qu'il est très facile avec Minitab® Statistical Software (au même titre qu'avec Microsoft Excel) d'obtenir une carte de contrôle avec deux phases (ou plus...).

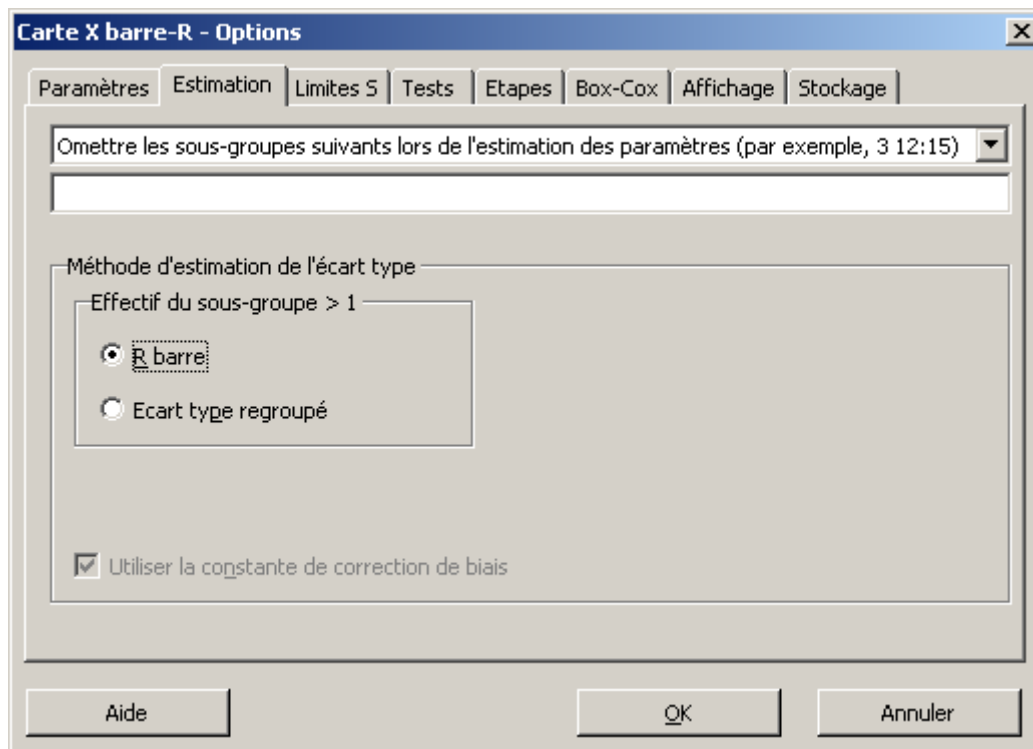
Dans Minitab® Statistical Software nous allons dans le menu **Stat/Cartes de contrôle/X-Barre R...**:



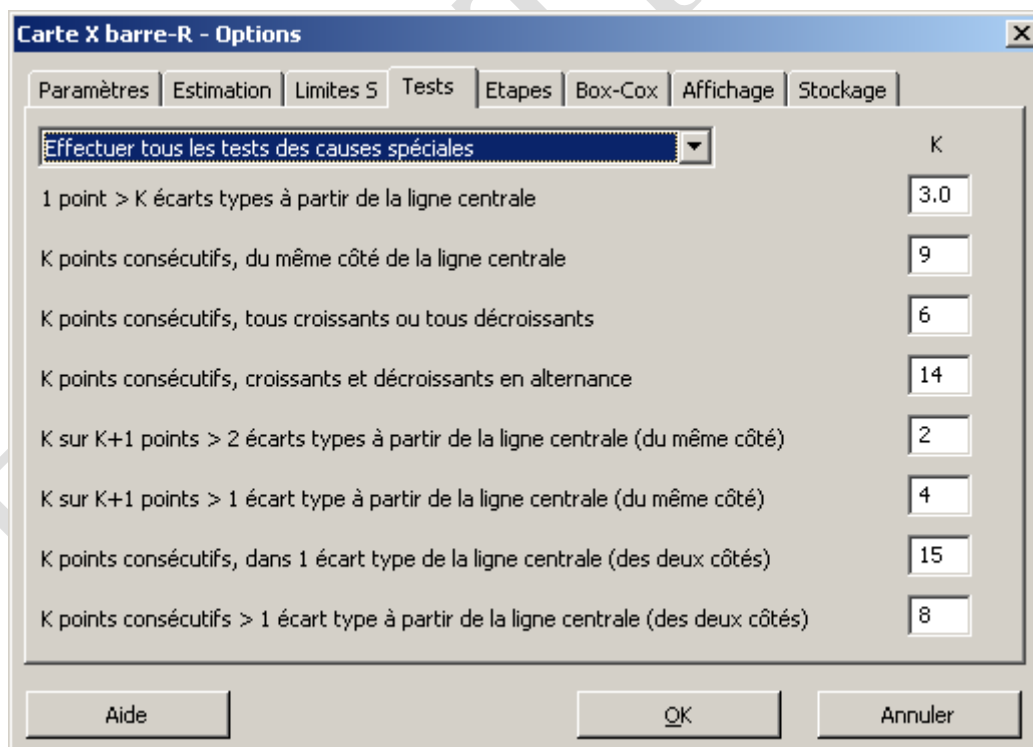
Nous avons alors la boîte de dialogue suivante:



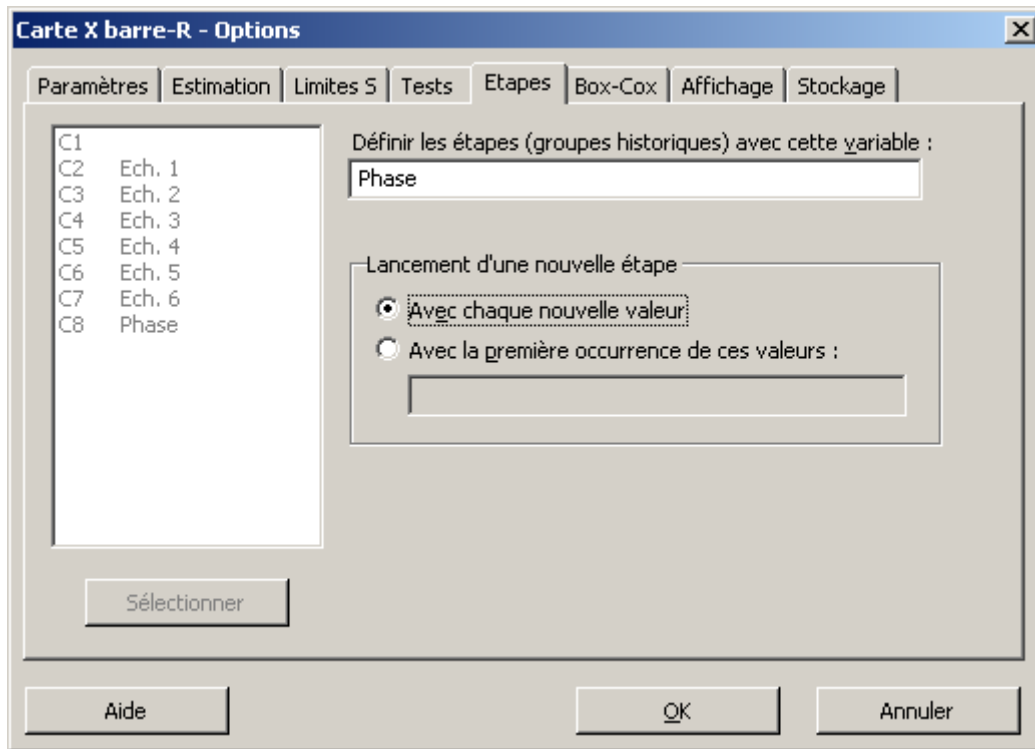
Nous cliquons sur **Options de carte X barre-R...** pour s'assurer d'abord que l'**Estimation** utilise la même méthode opératoire que celle vue dans le cours théorique:



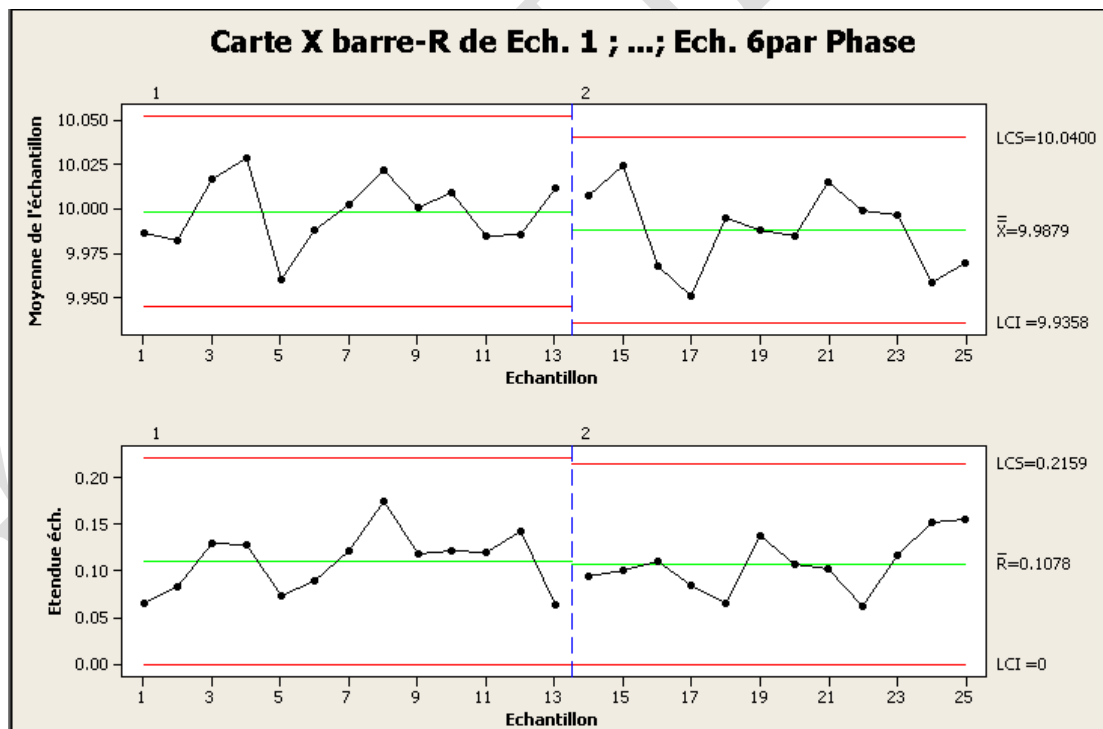
Nous en profitons pour activer toutes les règles de type WECO que Minitab® Statistical Software propose en allant dans l'onglet **Tests**:



et nous allons enfin dans l'onglet **Etapas** pour signaler au logiciel la colonne qui lui permettra de distinguer le commencement d'une nouvelle phase:



Nous validons deux fois par OK pour obtenir au final:



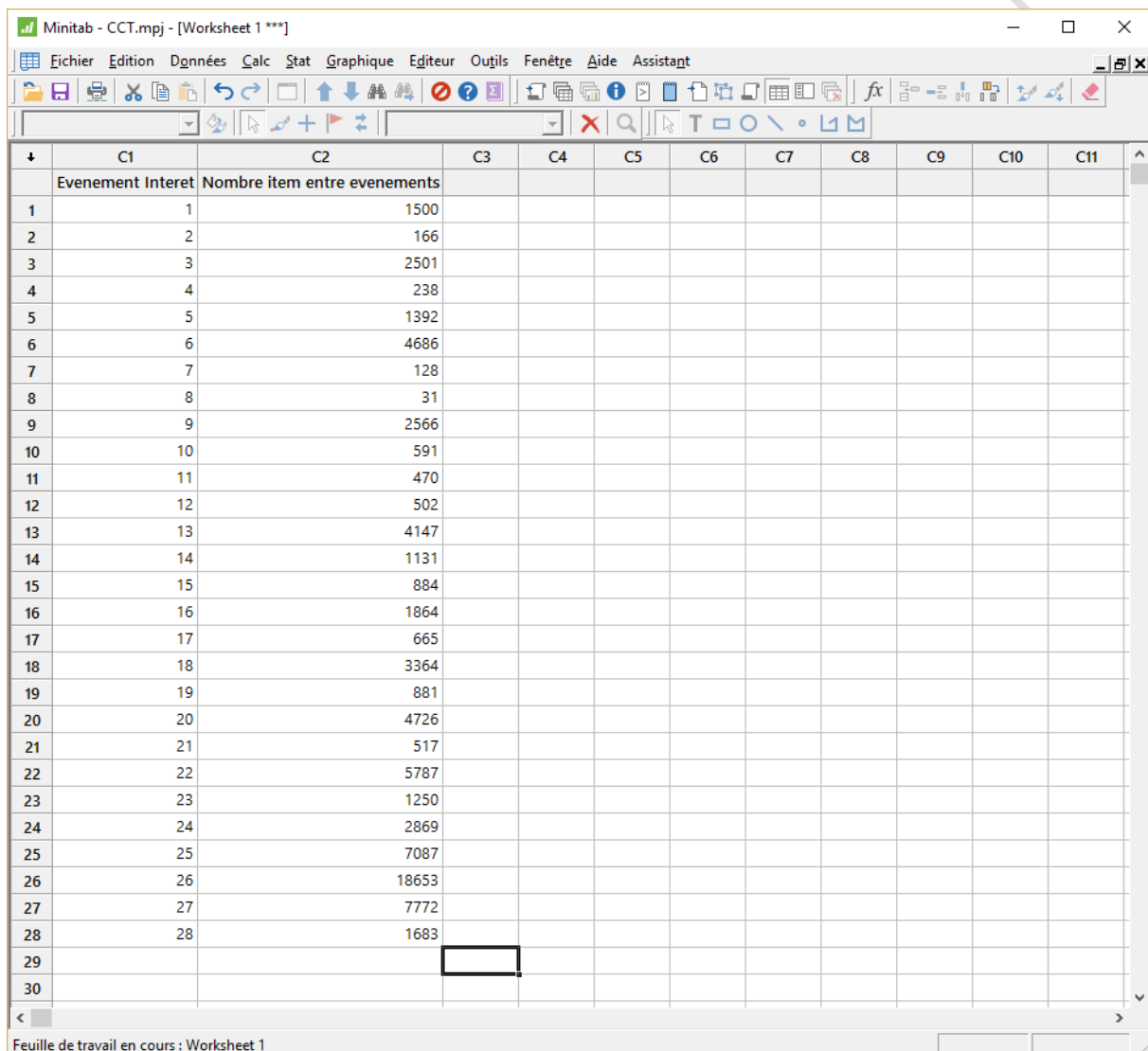
Pour voir qu'il n'y a finalement pas d'amélioration notable entre la phase 1 et 2.

## 14.22. Exercice 140.: CC T (aux fréquences)

Minitab® Statistical Software R17

Depuis la version 16 de Minitab nous avons enfin la carte de contrôle aux fréquences basée sur la loi exponentielle qui est disponible. Voyons cette dernière et comparons là aux développements effectués dans le cours théorique.

D'abord nous reprenons les mêmes données:



The screenshot shows the Minitab interface with a worksheet titled 'CCT.mpj - [Worksheet 1 \*\*\*]'. The worksheet contains two columns: C1, labeled 'Evenement Interet', and C2, labeled 'Nombre item entre evenements'. The data is as follows:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
	Evenement Interet	Nombre item entre evenements									
1	1	1500									
2	2	166									
3	3	2501									
4	4	238									
5	5	1392									
6	6	4686									
7	7	128									
8	8	31									
9	9	2566									
10	10	591									
11	11	470									
12	12	502									
13	13	4147									
14	14	1131									
15	15	884									
16	16	1864									
17	17	665									
18	18	3364									
19	19	881									
20	20	4726									
21	21	517									
22	22	5787									
23	23	1250									
24	24	2869									
25	25	7087									
26	26	18653									
27	27	7772									
28	28	1683									
29											
30											

Ensuite, nous allons dans **Stat/Cartes de contrôle/Carte des événements rares/T...**:

The screenshot shows the Minitab interface with the 'Stat' menu open. The 'Cartes de contrôle' option is selected, and the 'T' dialog box is open. The spreadsheet data is as follows:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Evenement Interet	Nombre item											
1	1												
2	2												
3	3												
4	4												
5	5												
6	6												
7	7												
8	8												
9	9												
10	10									591			
11	11									470			
12	12									502			
13	13									4147			
14	14									1131			
15	15									884			
16	16									1864			
17	17									665			
18	18									3364			
19	19									881			
20	20									4726			
21	21									517			
22	22									5787			
23	23									1250			
24	24									2869			
25	25									7087			
26	26									18653			
27	27									7772			
28	28									1683			
29													
30													

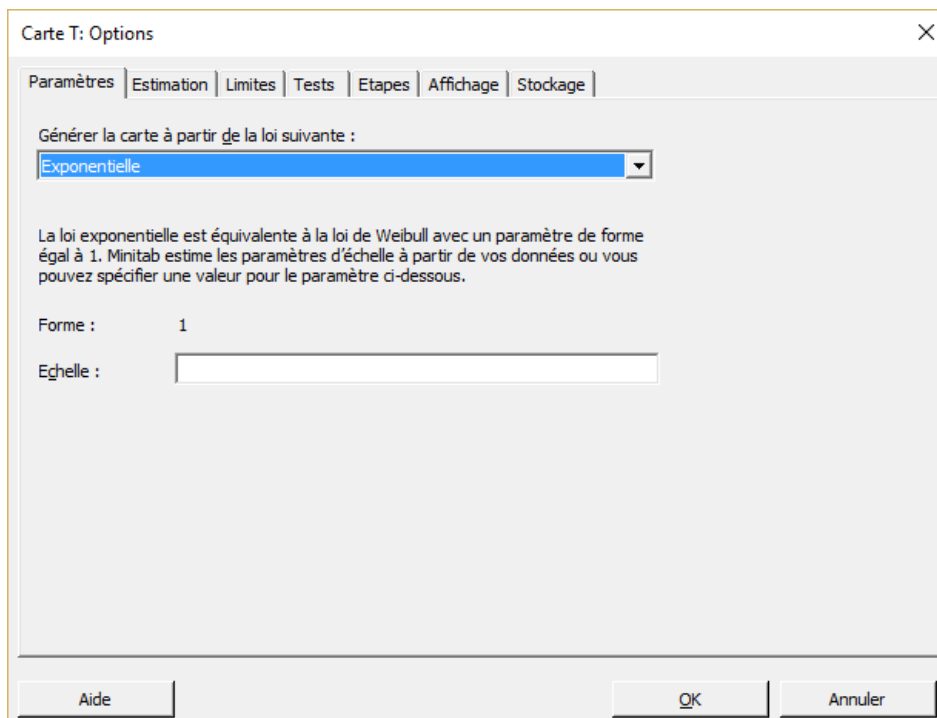
Nous remplissons comme ci-dessous:

The 'Carte T' dialog box is shown with the following settings:

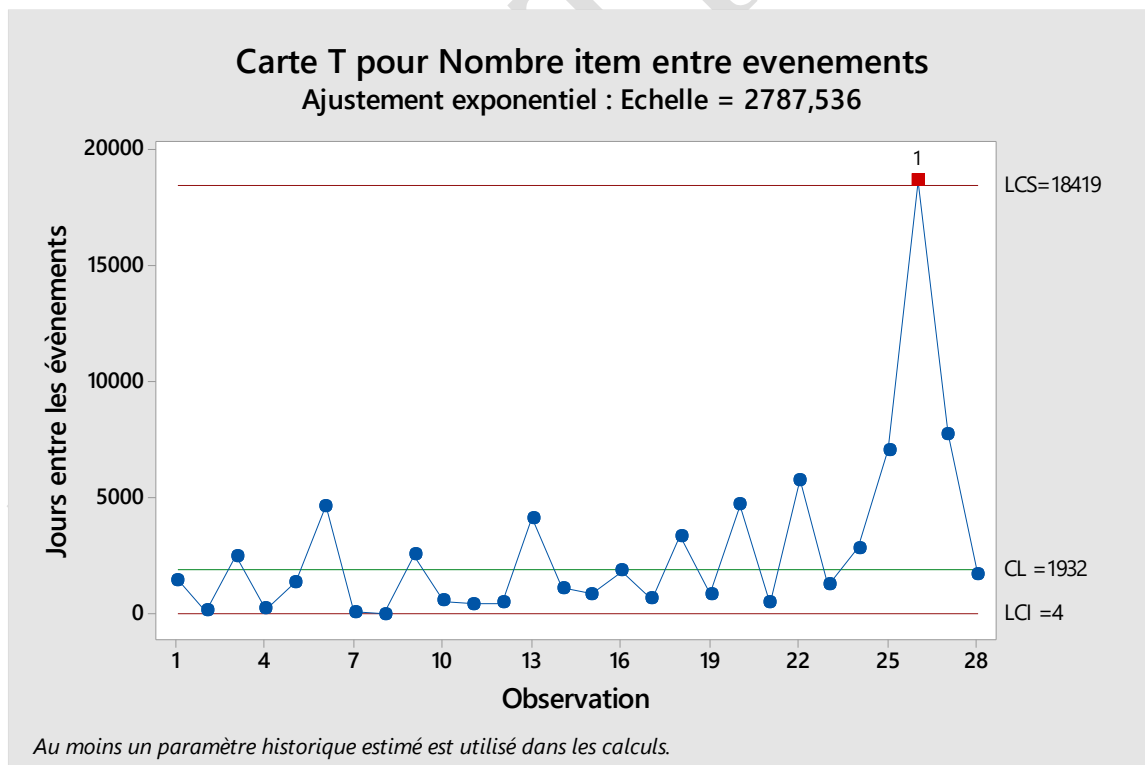
- Format des données :** Nombre d'intervalles entre les événements
- Unités pour les intervalles :** Jours
- Variables :** 'Nombre item entre evenements'

Buttons at the bottom include: Sélectionner, Echelle..., Etiquettes..., Graphiques multiples..., Options des données..., Options de carte T..., Aide, OK, and Annuler.

et nous cliquons sur **Options de carte T...** pour prendre **Exponentielle**:



Ce qui donne:

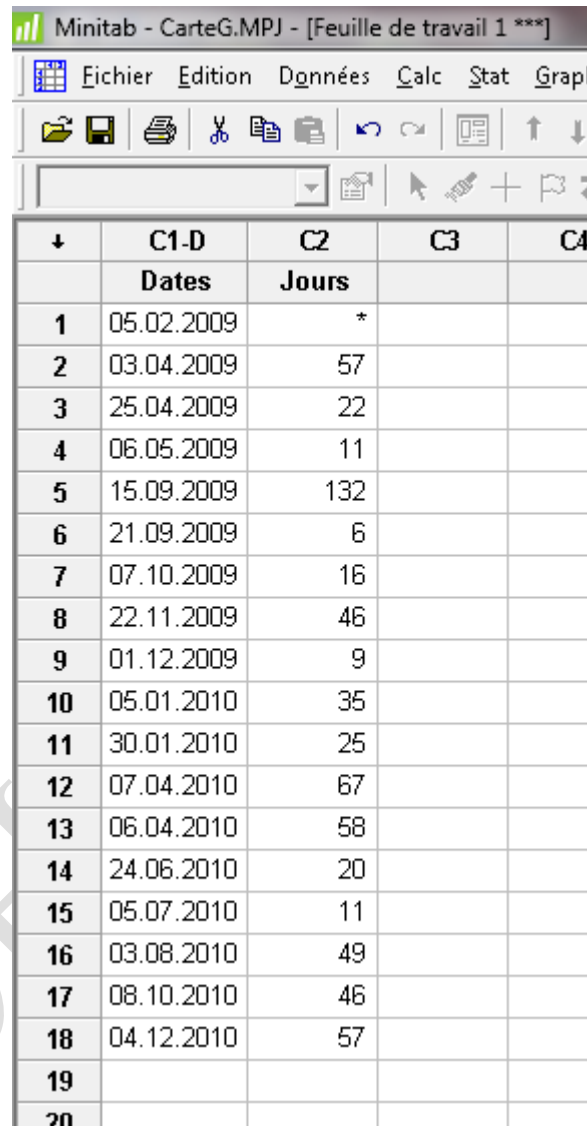


Nous voyons que les résultats sont significativement différents par rapport au cours théorique. La carte de contrôle de Minitab est moins sensible que celle que nous avons établie dans le cours théorique et même si Minitab communique les formules dans l'aide du logiciel ils ne communiquent pas explicitement la démarche intellectuelle et mathématique et donc il est difficile de juger qui est plus exact que l'autre.

## 14.23. Exercice 141.: CC G

Minitab® Statistical Software R16.2.1

Dans le cadre de contrôle d'événements rares, nous avons obtenu les mesures suivantes (fichier *CarteG.mpj*):



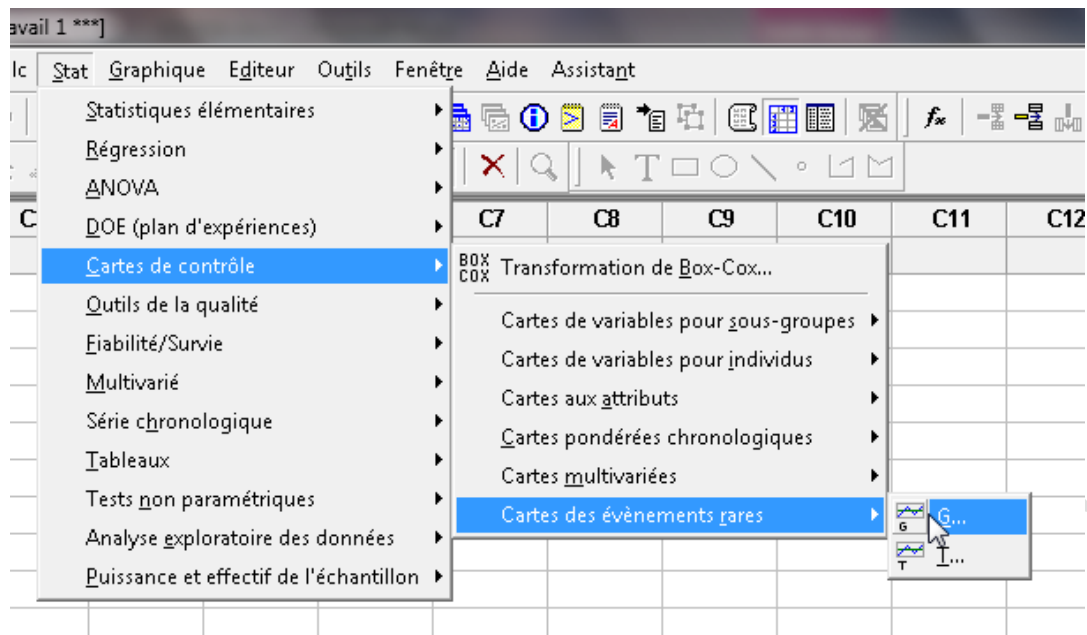
	C1-D	C2	C3	C4
	Dates	Jours		
1	05.02.2009	*		
2	03.04.2009	57		
3	25.04.2009	22		
4	06.05.2009	11		
5	15.09.2009	132		
6	21.09.2009	6		
7	07.10.2009	16		
8	22.11.2009	46		
9	01.12.2009	9		
10	05.01.2010	35		
11	30.01.2010	25		
12	07.04.2010	67		
13	06.04.2010	58		
14	24.06.2010	20		
15	05.07.2010	11		
16	03.08.2010	49		
17	08.10.2010	46		
18	04.12.2010	57		
19				
20				

où nous pouvons-voir que nous avons pris le cas particulier où... il y a exactement un événement mesuré par jour et donc le nombre d'événements entre chaque anomalie correspond au nombre de jours entre chaque anomalie aussi...

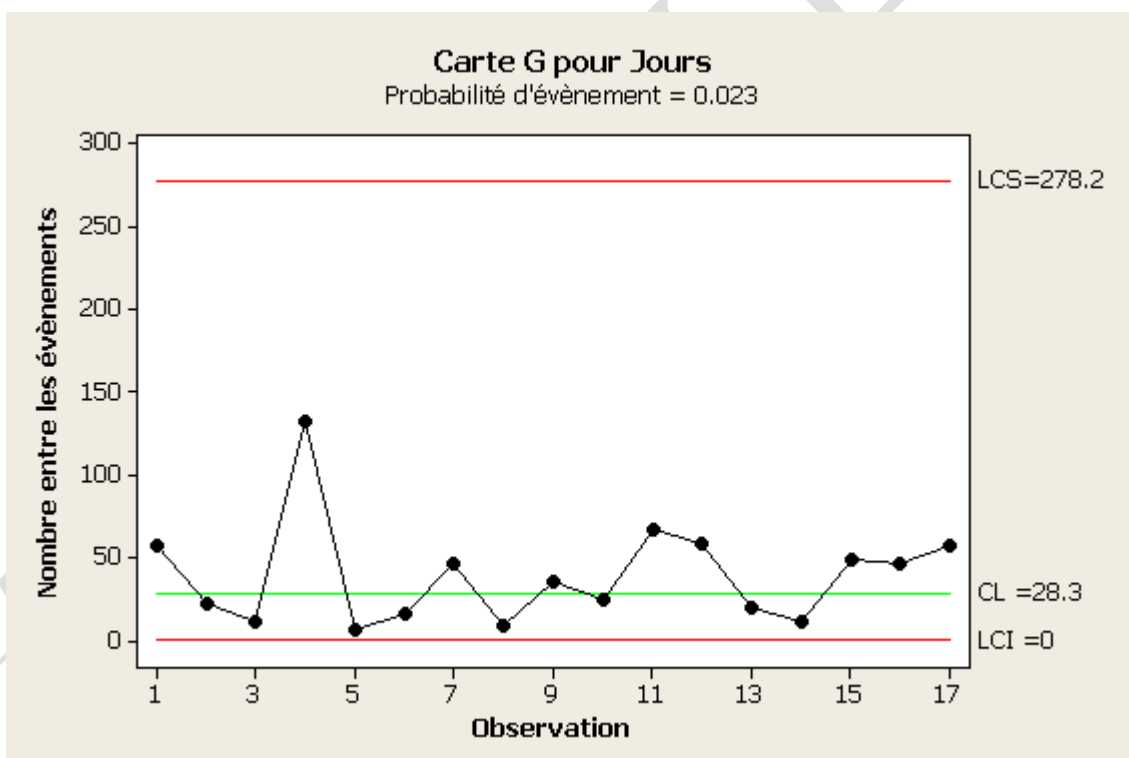
Nous souhaiterions tracer une carte de contrôle par des événements G afin de voir qu'il est très facile avec Minitab® Statistical Software (au même titre qu'avec Microsoft Excel) d'obtenir une carte de contrôle.

Dans Minitab® Statistical Software (16.2.1) nous allons dans le menu **Stat/Cartes de contrôle/Cartes des événements rares/G...**:





Nous obtenons alors:



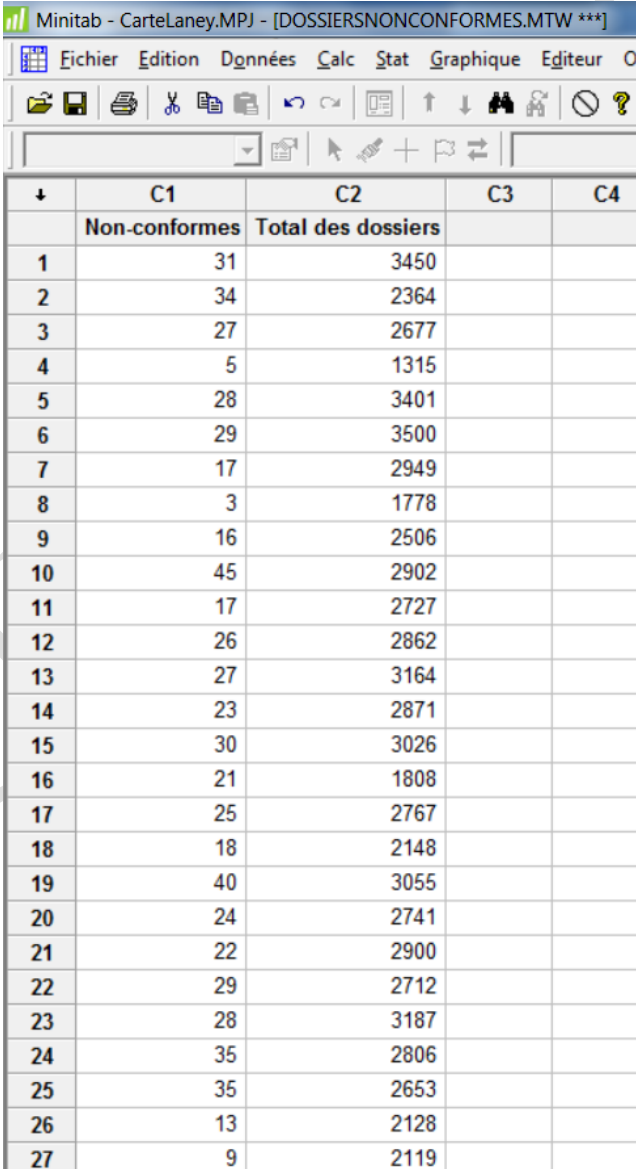
où, comme nous l'avons spécifié lors de l'étude théorique, pour un logiciel comme Minitab® Statistical Software, la limite de contrôle inférieure est définie au percentile 0.0013499 de la distribution géométrique (en prenant la valeur décimale la plus proche par interpolation linéaire). La limite de contrôle supérieure est définie au percentile 0.99865 (en prenant aussi la valeur décimale la plus proche par interpolation linéaire). La ligne centrale est définie au percentile 0.5 (en prenant la valeur décimale la plus proche par interpolation linéaire).

## 14.24. Exercice 142.: CC P' de Laney

Minitab® Statistical Software R16.2.1

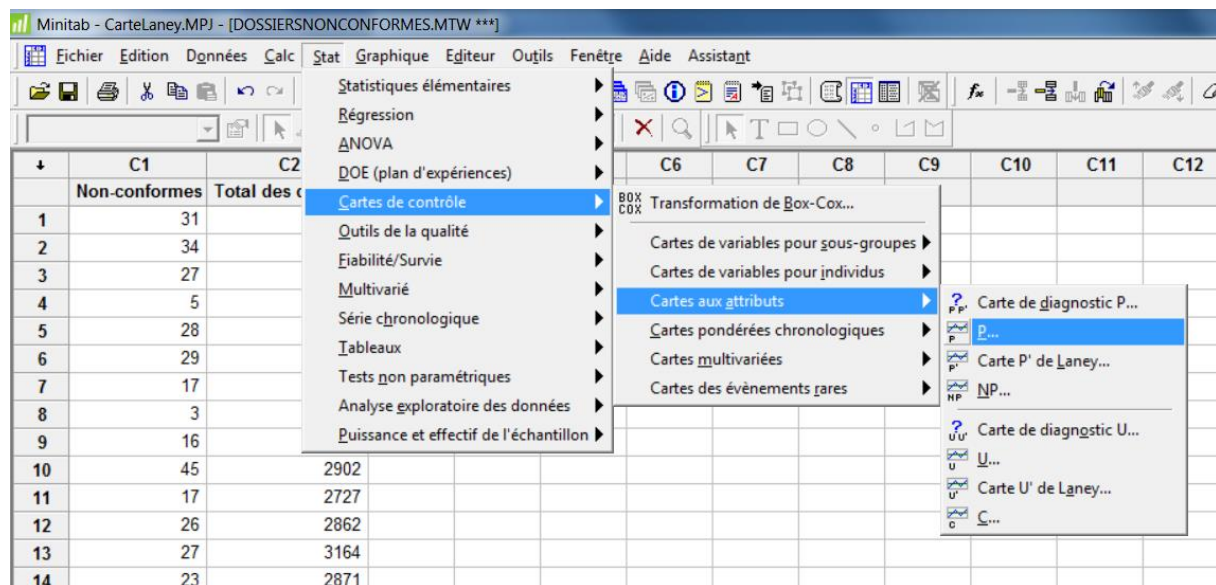
Le but va être ici de vérifier si nous obtenons les mêmes résultats que le cas d'application numérique calculé avec Microsoft Excel suite à la démonstration mathématique "au pouce" du concept de construction de carte de Laney. Nous savons qu'il y a peu de chance d'obtenir exactement les mêmes valeurs que le logiciel Minitab étant donné que nous avons précisé dans le cours théorique qu'il existait des corrections au modèle théorique (raison pour laquelle pour des études sérieuses il faut toujours avoir au moins trois logiciels de statistiques sous la main!).

Ouvrez le fichier CarteLaney\_P.mpj qui est une partie de l'exemple fourni avec Minitab:

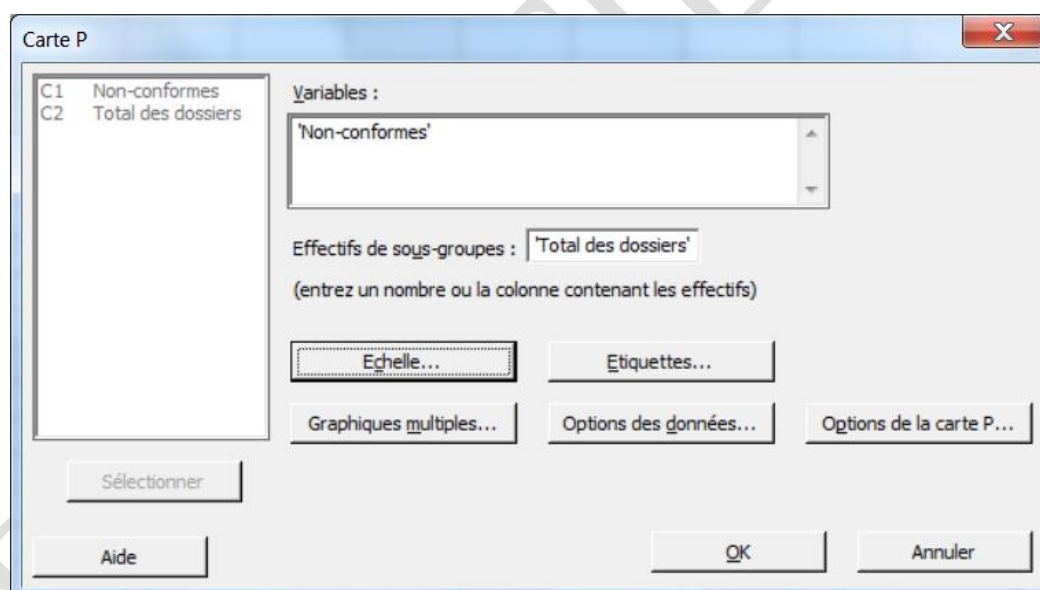


	C1	C2	C3	C4
	<b>Non-conformes</b>	<b>Total des dossiers</b>		
1	31	3450		
2	34	2364		
3	27	2677		
4	5	1315		
5	28	3401		
6	29	3500		
7	17	2949		
8	3	1778		
9	16	2506		
10	45	2902		
11	17	2727		
12	26	2862		
13	27	3164		
14	23	2871		
15	30	3026		
16	21	1808		
17	25	2767		
18	18	2148		
19	40	3055		
20	24	2741		
21	22	2900		
22	29	2712		
23	28	3187		
24	35	2806		
25	35	2653		
26	13	2128		
27	9	2119		

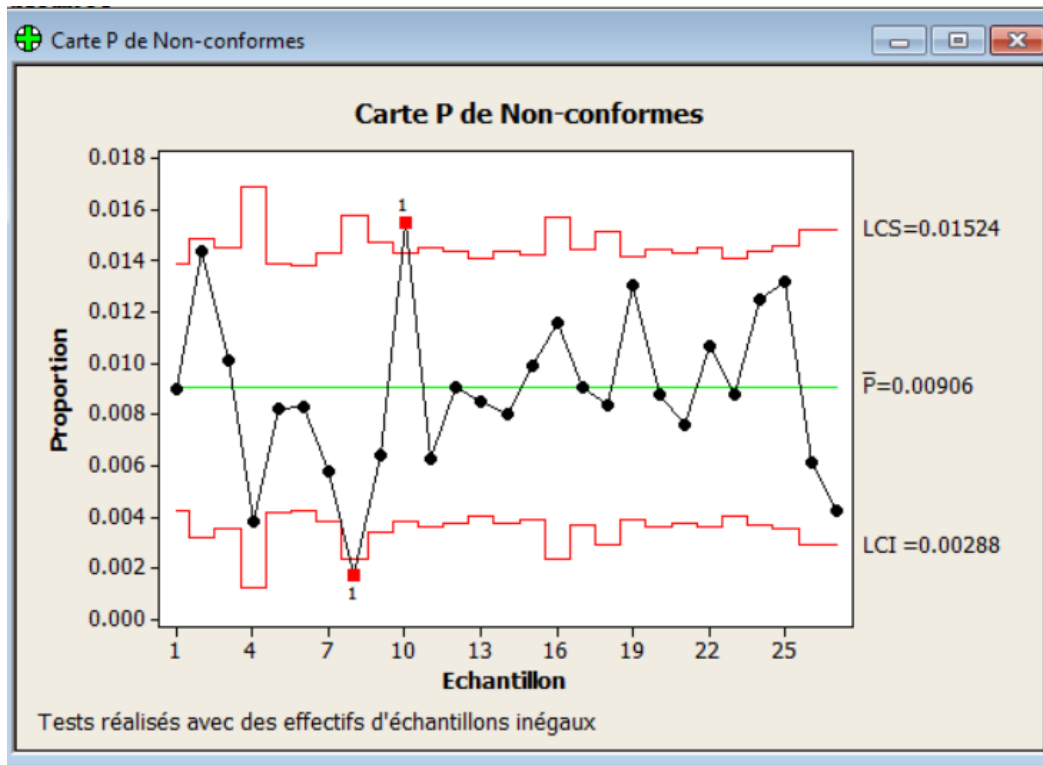
Ensuite, nous allons dans le menu **Stat/Cartes de contrôle/Cartes aux attributs/P...** pour voir d'abord à quoi ressemble ses données dans une carte de contrôle  $p$  standard et la comparer avec une carte  $p$  de Laney:



Et nous avons alors:



Nous validons par **OK** pour obtenir:



Bon il n'y a pas grand-chose à en dire car nous n'y avons pas d'alarmes systématiques (donc faire appel à la carte de Laney est discutable!). Maintenant faisons ce qui nous intéresse! Nous allons dans le menu **Stat/Cartes de contrôle/Cartes aux attributs/Carte P' de Laney...**:

Minitab - CarteLaney.MPJ - [DOSSIERSNONCONFORMES.MTW \*\*\*]

Stat Graphique Editeur Outils Fenêtre Aide Assistant

Statistiques élémentaires

Régression

ANOVA

DOE (plan d'expériences)

Cartes de contrôle

Outils de la qualité

Fiabilité/Survie

Multivarié

Série chronologique

Tableaux

Tests non paramétriques

Analyse exploratoire des données

Puissance et effectif de l'échantillon

	C1	C2
	Non-conformes	Total des c
1	31	
2	34	
3	27	
4	5	
5	28	
6	29	
7	17	
8	3	
9	16	
10	45	2902
11	17	2727
12	26	2862
13	27	3164
14	23	2871

Transformation de Box-Cox...

Cartes de variables pour sous-groupes

Cartes de variables pour individus

Cartes aux attributs

Cartes pondérées chronologiques

Cartes multivariées

Cartes des événements rares

Carte de diagnostic P...

P...

Carte P' de Laney...

NP...

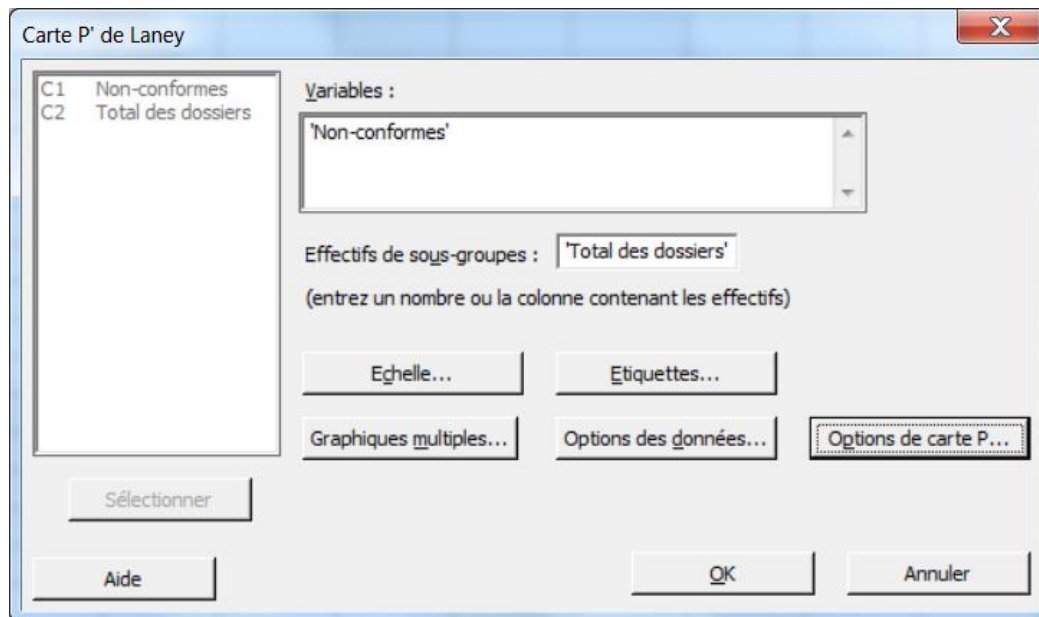
Carte de diagnostic U...

U...

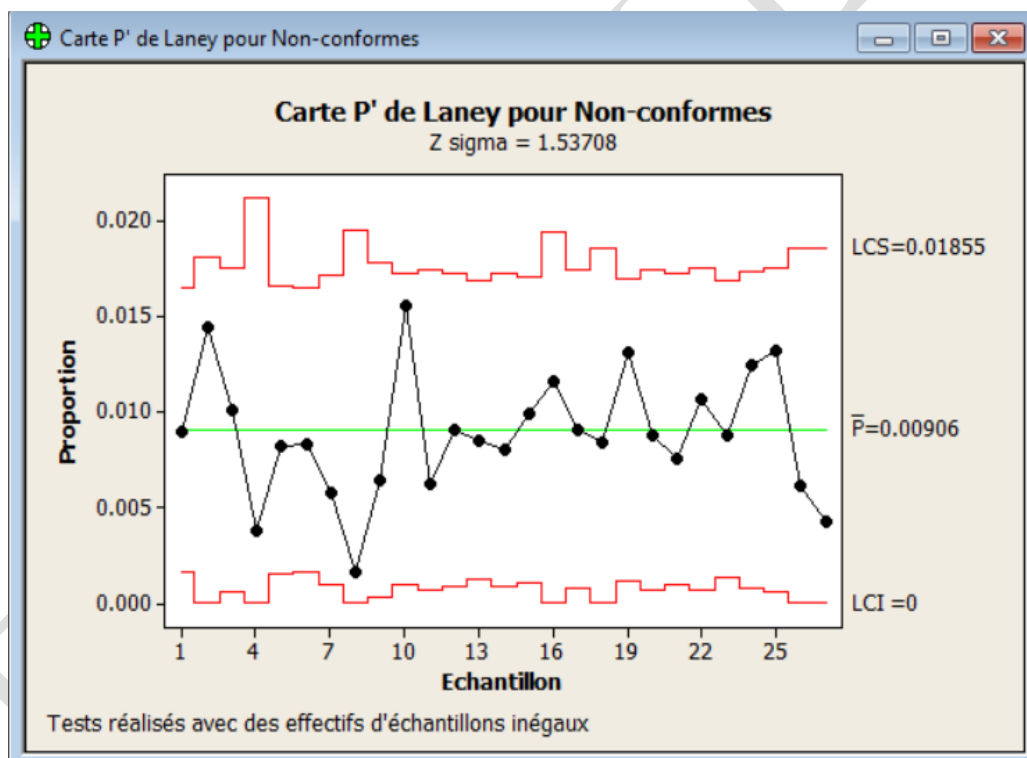
Carte U' de Laney...

C...

Nous avons alors:

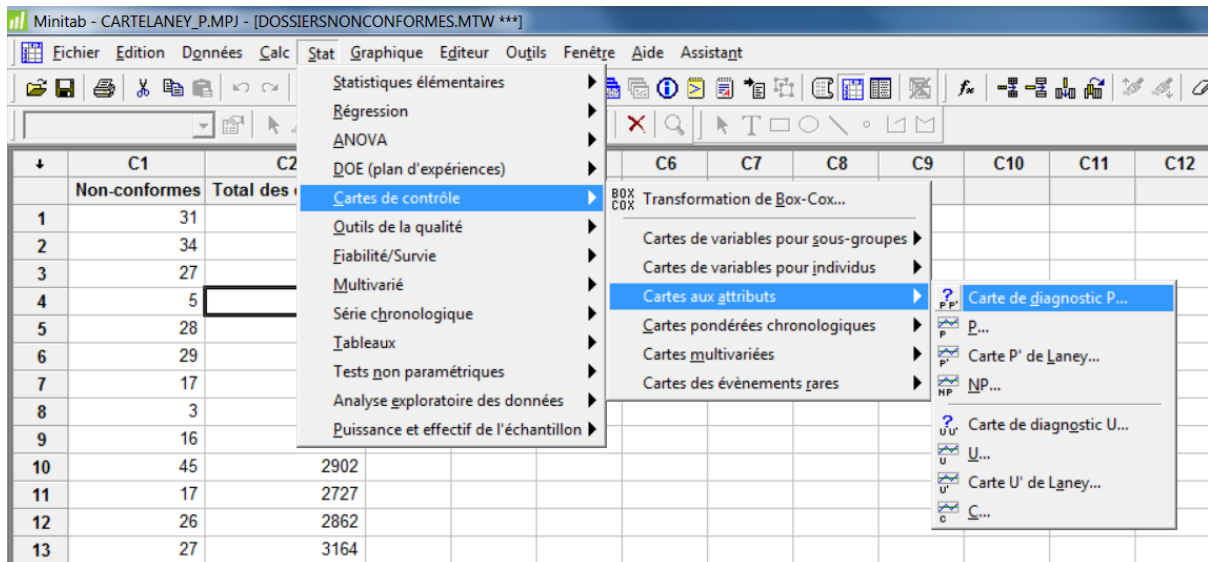


Nous obtenons alors:

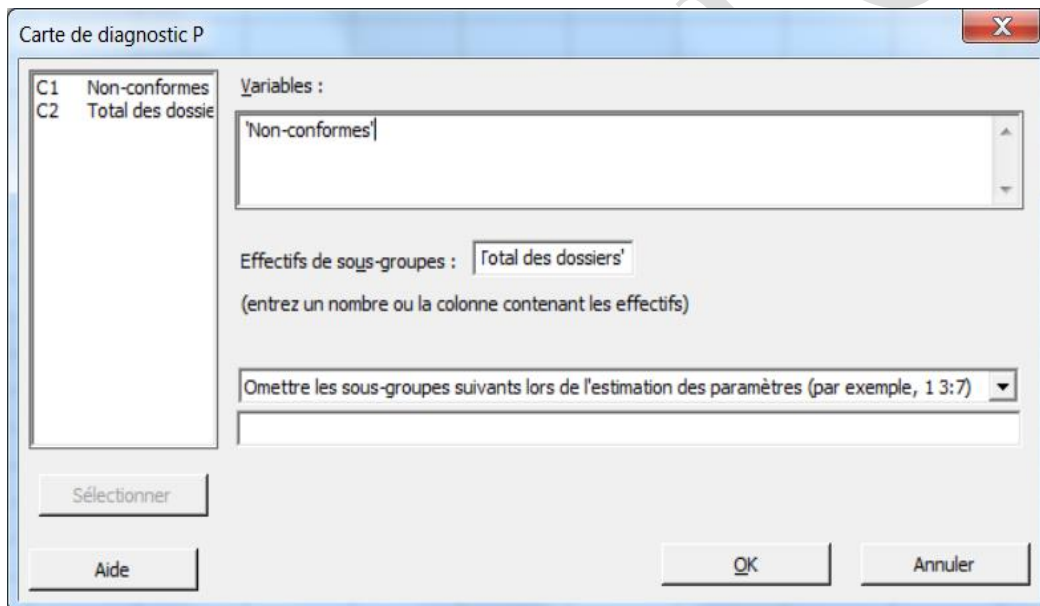


Donc nous voyons que la carte de contrôle P' est un peu moins sensible. Nous ne retrouvons aussi effectivement pas exactement les mêmes valeurs que lors des calculs à la main. Mais cela était (a priori) prévisible!

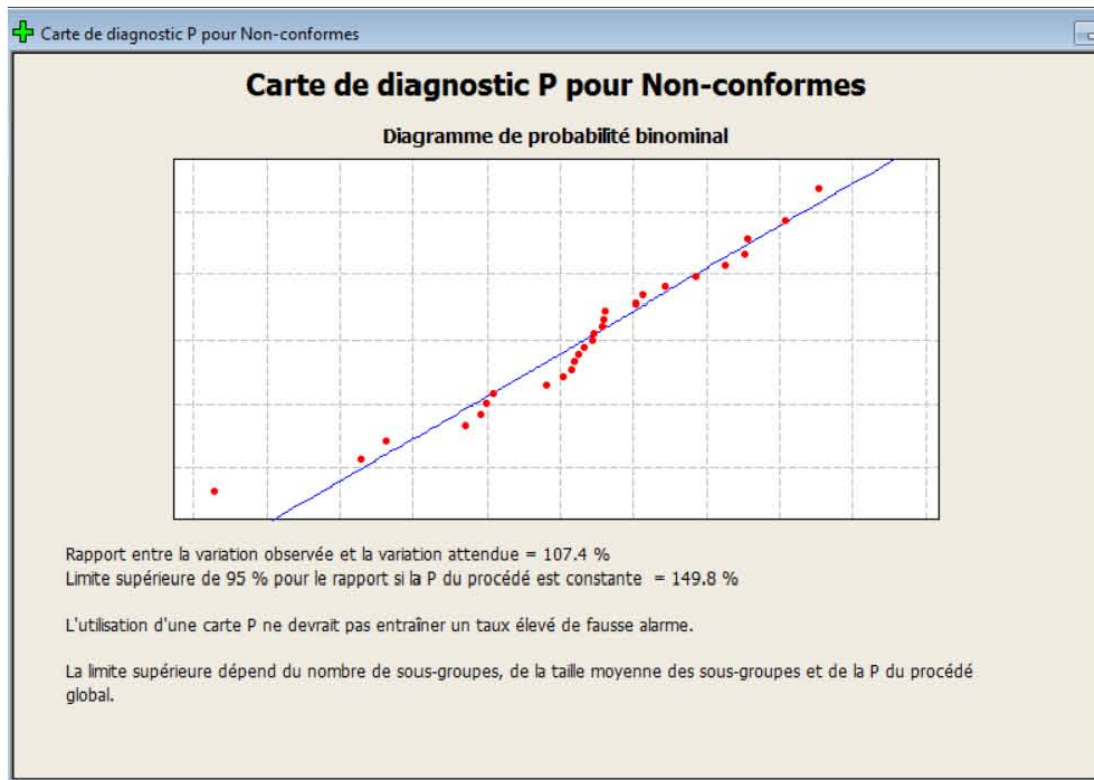
Pour aider (mais cela n'empêche pas les discussions d'expert!). Minitab propose un outil d'aide au choix de la meilleure carte  $p$ . Ainsi, en allant dans le menu **Stat/Cartes de contrôle/Cartes aux attributs/Carte de diagnostic P...**:



Nous avons alors:



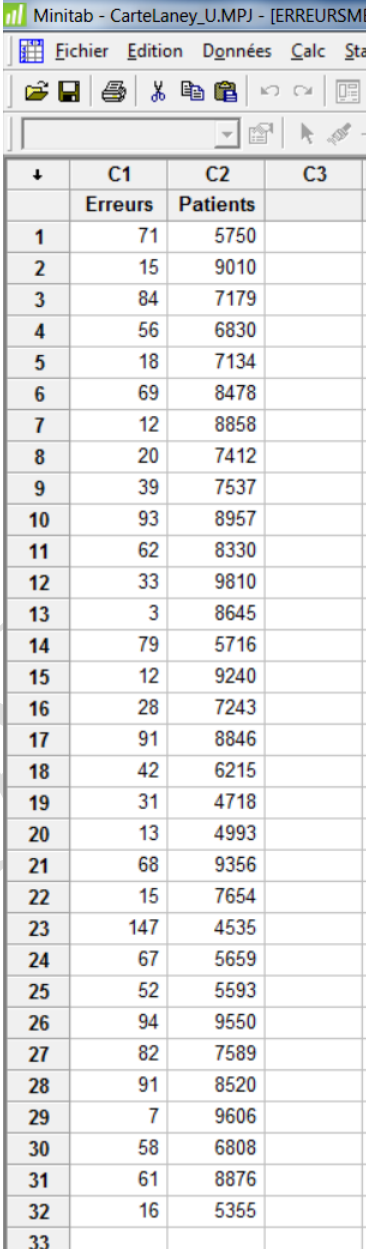
Il vient alors (la conclusion est conforme visuellement parlant mais scientifiquement parlant, nous n'avons pas traité ni démontré quoi que ce soit par rapport à cet outil d'aide à la décision dans le cours théorique):



## 14.25. Exercice 143.: CC U' de Laney

Minitab® Statistical Software R16.2.1

Nous pouvons constater la même chose avec les mêmes conclusions avec la carte  $u$  de Laney mais cette fois avec des données différentes. Nous partons aussi de l'exemple fourni avec Minitab:



	C1	C2	C3
	Erreurs	Patients	
1	71	5750	
2	15	9010	
3	84	7179	
4	56	6830	
5	18	7134	
6	69	8478	
7	12	8858	
8	20	7412	
9	39	7537	
10	93	8957	
11	62	8330	
12	33	9810	
13	3	8645	
14	79	5716	
15	12	9240	
16	28	7243	
17	91	8846	
18	42	6215	
19	31	4718	
20	13	4993	
21	68	9356	
22	15	7654	
23	147	4535	
24	67	5659	
25	52	5593	
26	94	9550	
27	82	7589	
28	91	8520	
29	7	9606	
30	58	6808	
31	61	8876	
32	16	5355	
33			

Et d'abord nous faisons une carte classique pour comparaison ultérieure en allant dans le menu **Stat/Cartes de contrôle/Cartes aux attributs/U...**:



	C1	C2	C3
	Erreurs	Patients	
1	71	5750	
2	15	9010	
3	84	7179	
4	56	6830	
5	18	7134	
6	69	8478	
7	12	8858	
8	20	7412	
9	39	7537	
10	93	8957	
11	62	8330	
12	33	9810	
13	3	8645	

Nous avons alors:

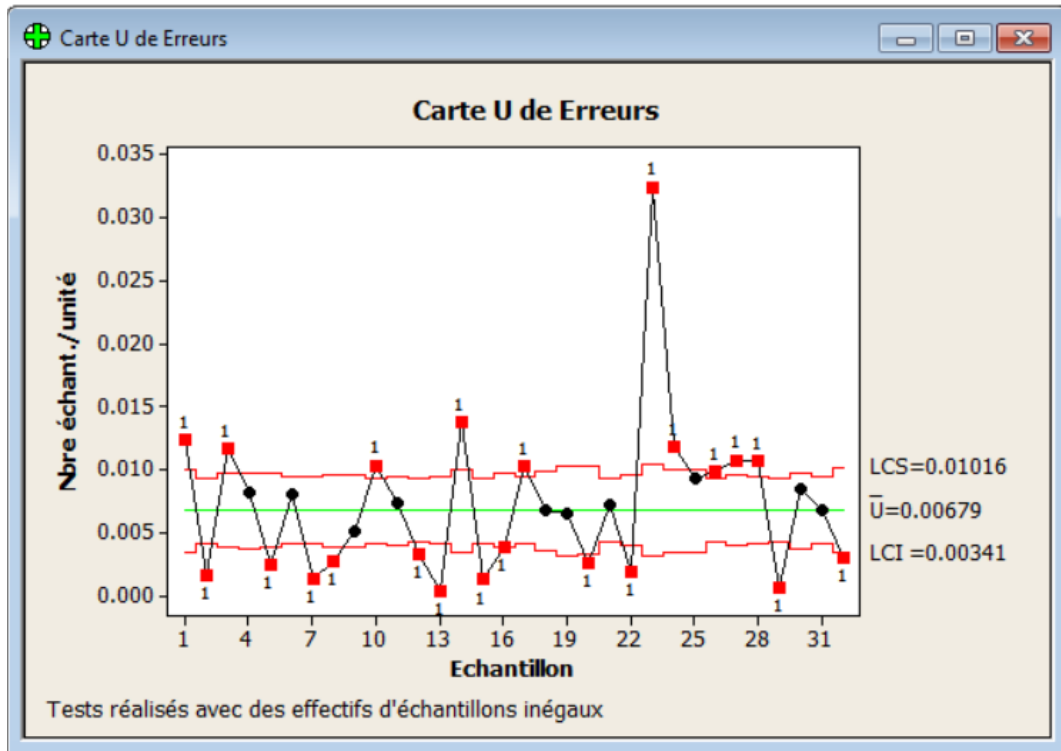
Carte U

Variables :  
Erreurs

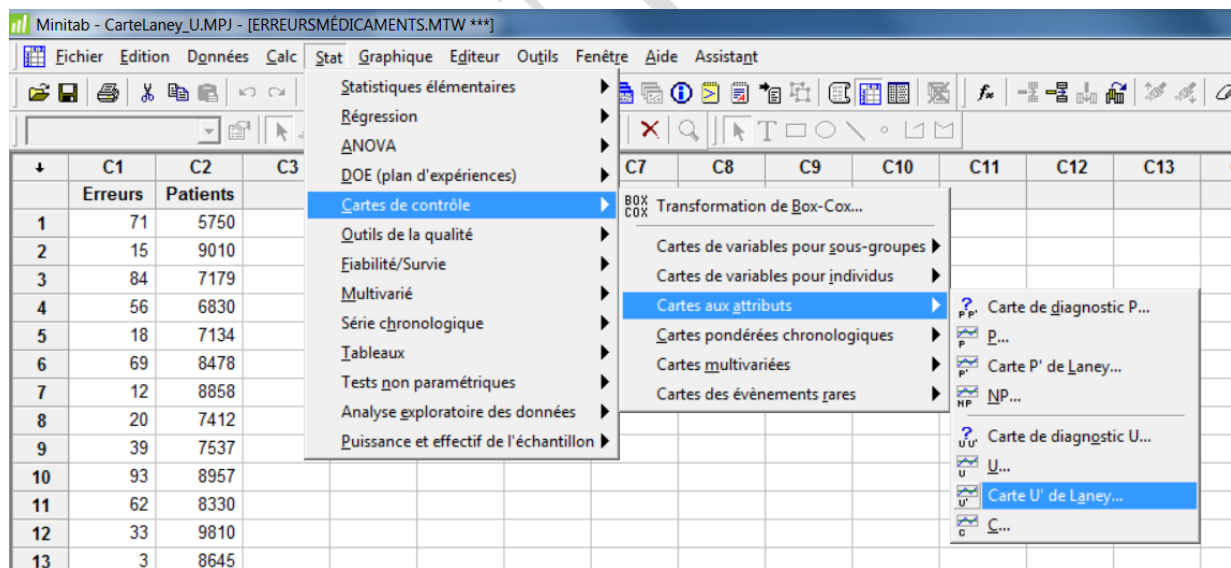
Effectifs des sous-groupes : Patients  
(entrez un nombre ou la colonne contenant les effectifs)

Echelle... Etiquettes...  
Graphiques multiples... Options des données... Options de carte U...  
Sélectionner Aide OK Annuler

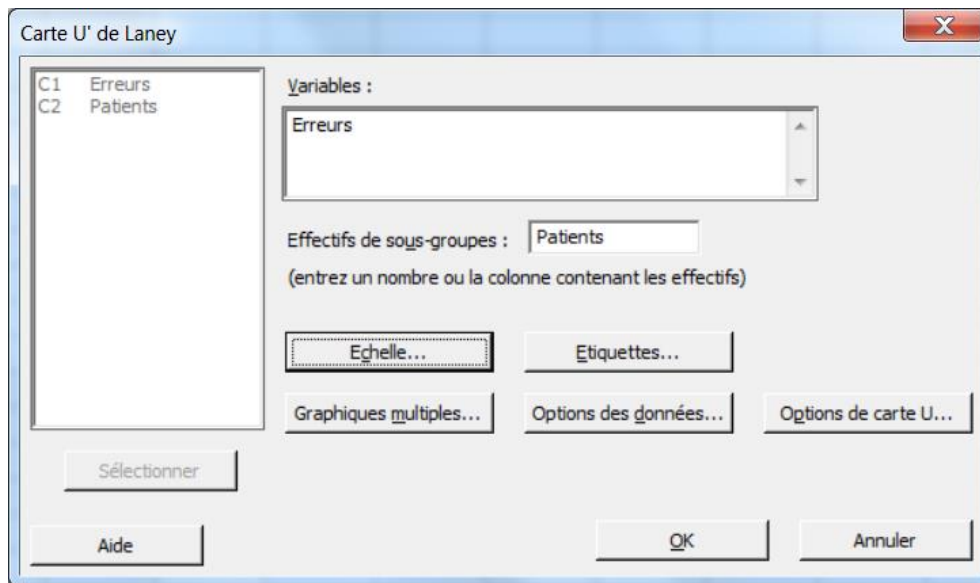
et en cliquant sur **OK**:



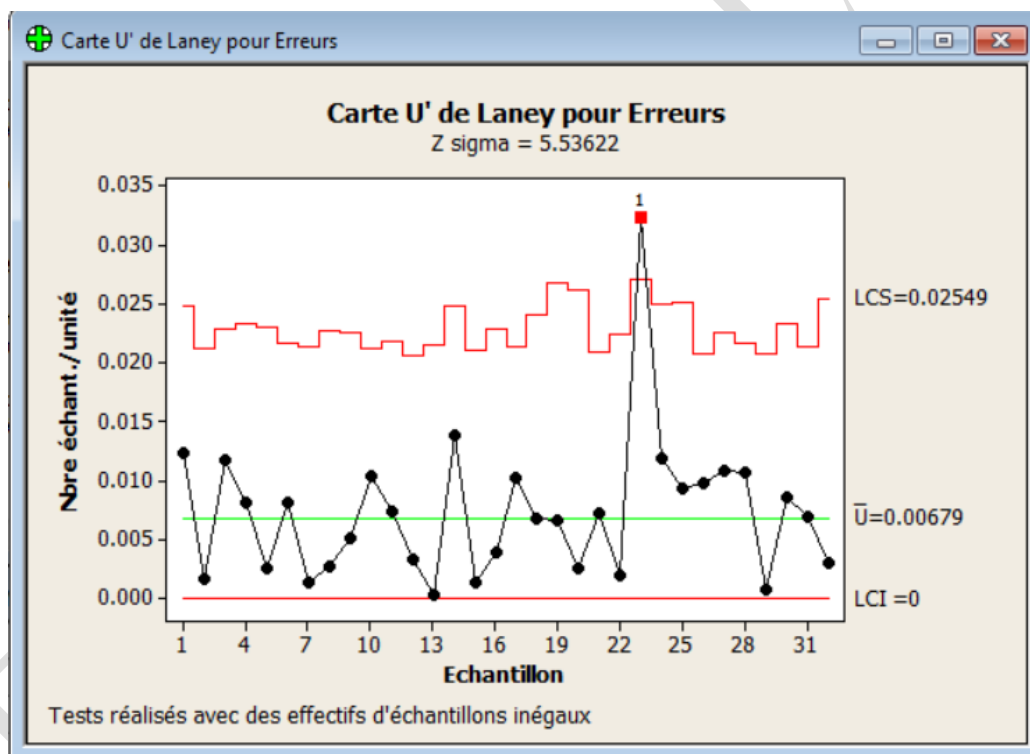
Donc dans le cas présent... soit en s'étrangle.... Soit en pose des questions. Vient alors la carte de  $u$  de Laney en allant dans le menu **Stat/Cartes de contrôle/Cartes aux attributs/Carte U' de Laney...**:



Ce qui nous donne:

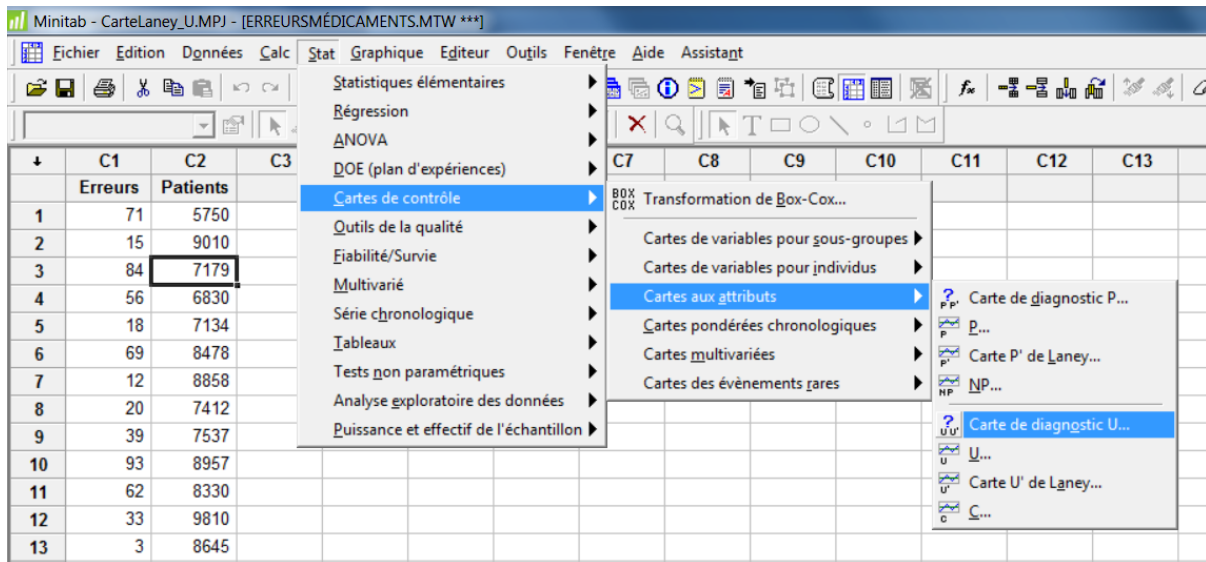


et si nous cliquons sur **OK**:

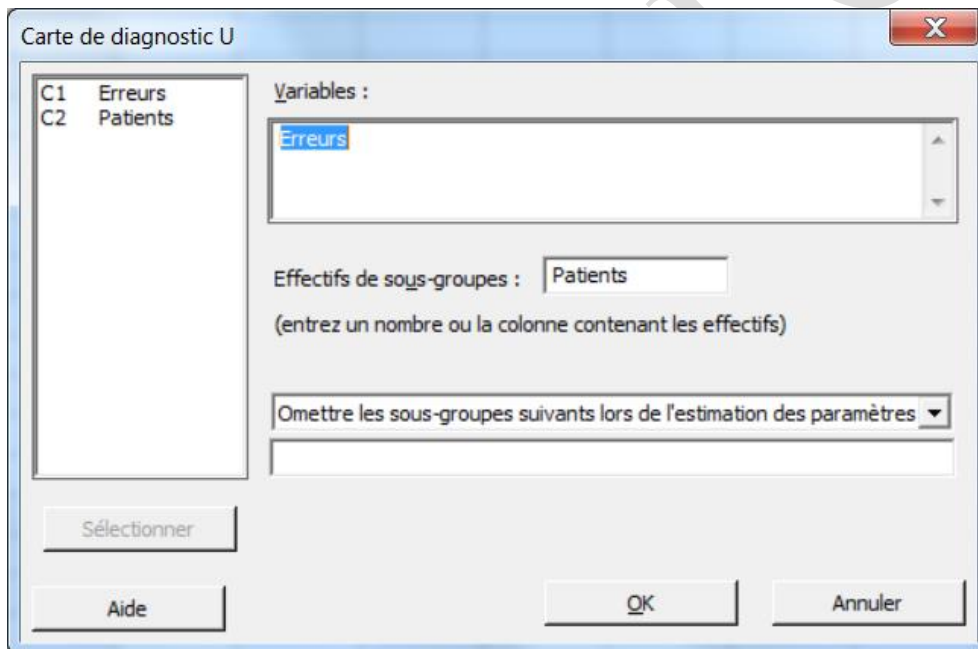


Ce qui est nettement plus réaliste mais.... sujet toutefois à discussion.

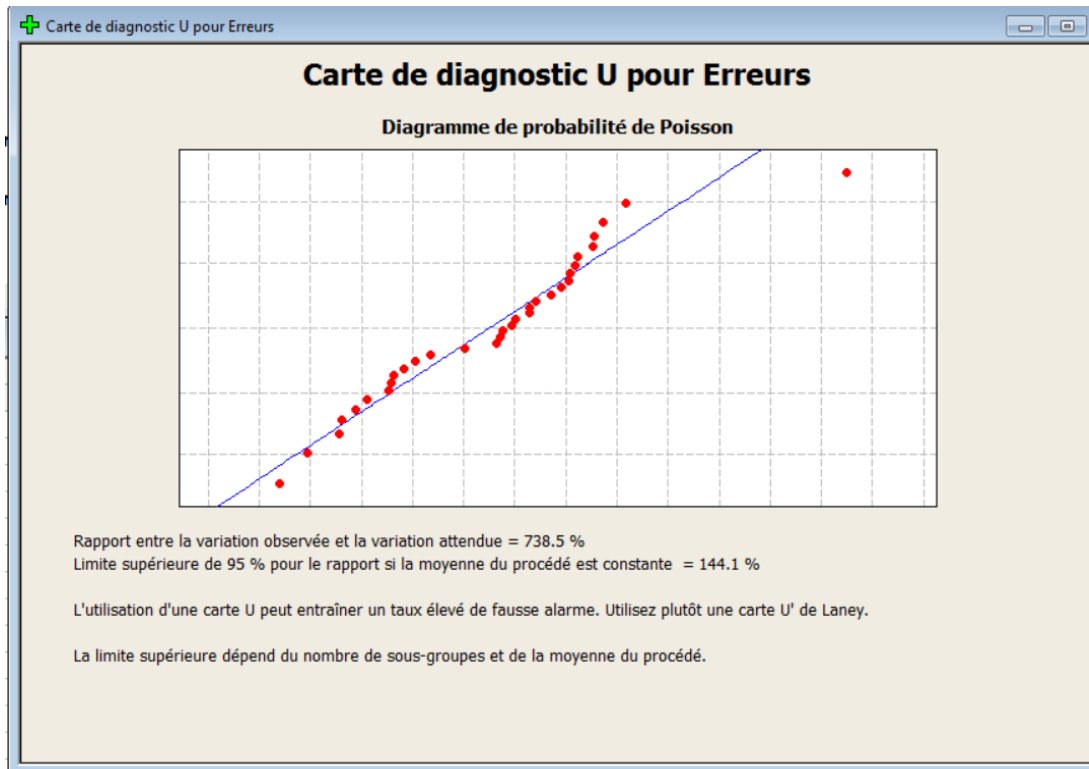
Au même titre que pour la carte  $p$  de Laney, il existe un outil d'aide à la décision quant au choix de la carte. Pour cela, nous allons dans le menu **Stat/Cartes de contrôle/Cartes aux attributs/Carte de diagnostic U...**:



Ce qui nous donne:



Il vient alors (la conclusion est conforme visuellement parlant mais scientifiquement parlant, nous n'avons pas traité ni démontré quoi que ce soit par rapport à cet outil d'aide à la décision dans le cours théorique):



Là encore, rappelons l'importance pour des études sérieuses qu'il faut toujours avoir au moins trois progiciels de statistiques sous la main!

## 14.26. Exercice 144.: Carte de contrôle de Hotelling (T2)

Minitab® Statistical Software 15.1.1.0

Comme nous l'avons mentionné dans le cours théorique il existe dans la pratique plusieurs types de cartes multivariées de Hotelling  $T^2$  (trois à ma connaissance):

- Carte de Hotelling  $T^2$  avec USL par loi du khi-2
- Carte de Hotelling  $T^2$  avec USL par loi Beta
- Carte de Hotelling  $T^2$  avec USL par loi de Fisher

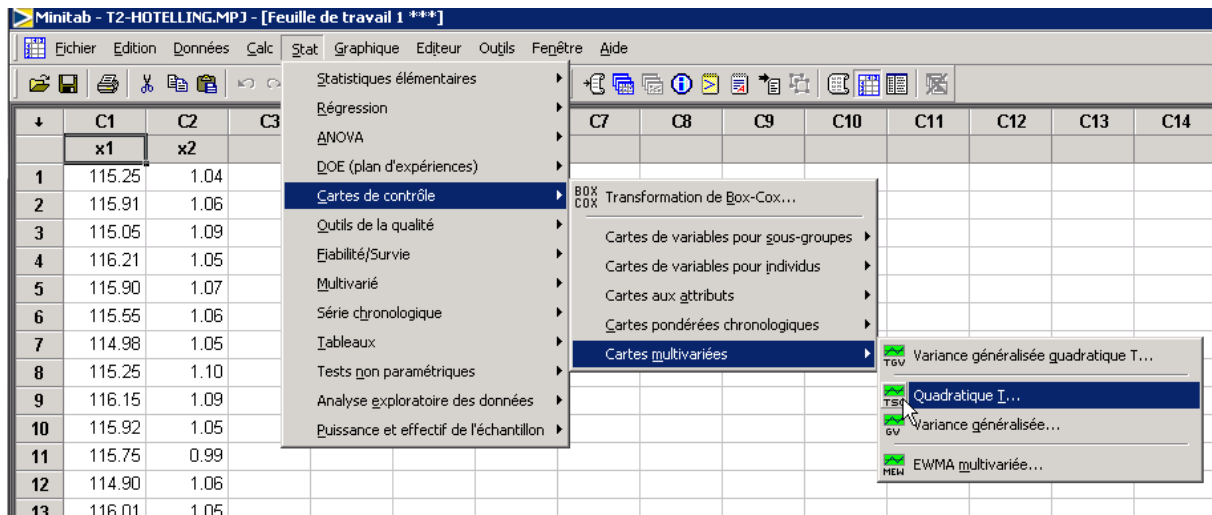
C'est la raison pour laquelle en fonction du logiciel de statistiques utilisé vous aurez une USL qui variera du simple au double voire au triple pour les mêmes données de base.

Dans le cadre présent, les concepteurs de Minitab ont choisi de faire le Carte de Hotelling de Fisher en se basant sur les différences séquentielle des mesures ce qui est un choix un peu neutre dans le cadre où le processus n'est pas parfaitement sous contrôle statistique (Normalité des données).

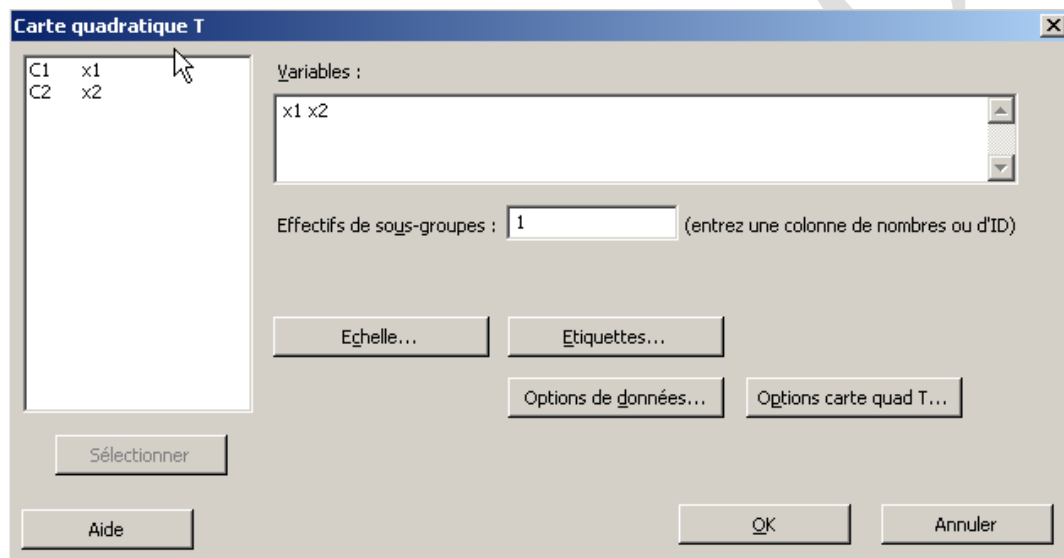
Nous partons donc des mêmes données que dans le cours théorique:

	C1	C2	C3	C4	C5	C6	C7
	x1	x2					
1	115.25	1.04					
2	115.91	1.06					
3	115.05	1.09					
4	116.21	1.05					
5	115.90	1.07					
6	155.55	1.06					
7	114.98	1.05					
8	115.25	1.10					
9	116.15	1.09					
10	115.92	1.05					
11	115.75	0.99					
12	114.90	1.06					
13	116.01	1.05					
14	115.83	1.07					
15	115.29	1.11					
16	115.63	1.04					
17	115.40	1.03					
18	115.58	1.05					
19	115.72	1.06					
20	115.40	1.05					

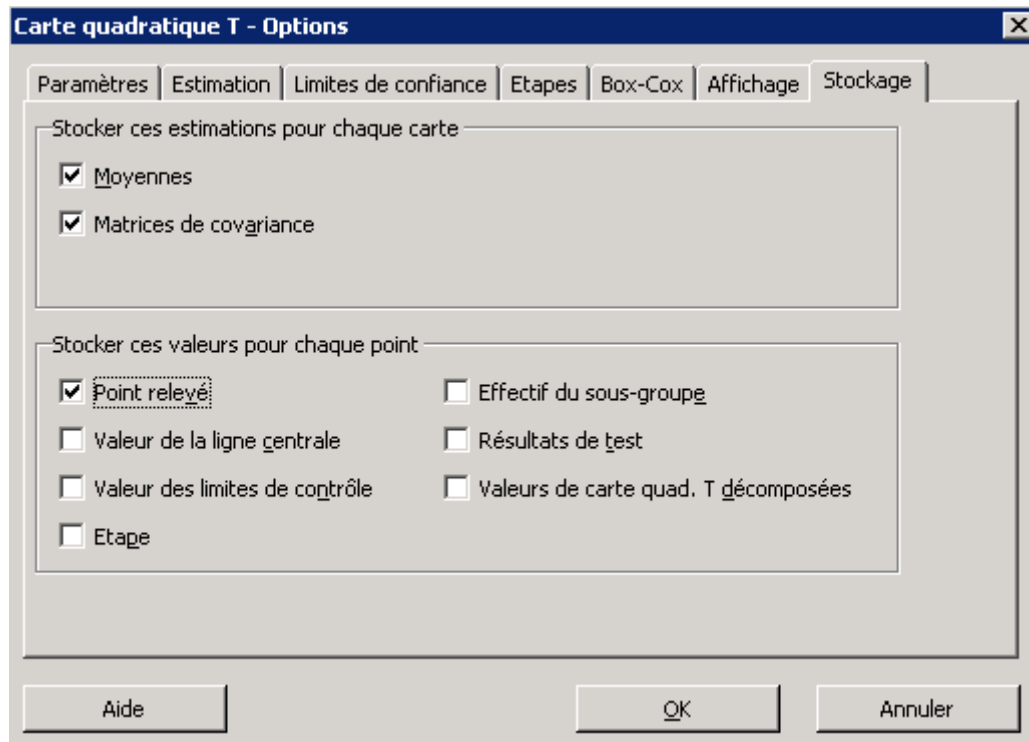
Ensuite, nous allons dans le menu **Stat/Cartes de contrôle/Cartes multivariées/Quadratique T...**:



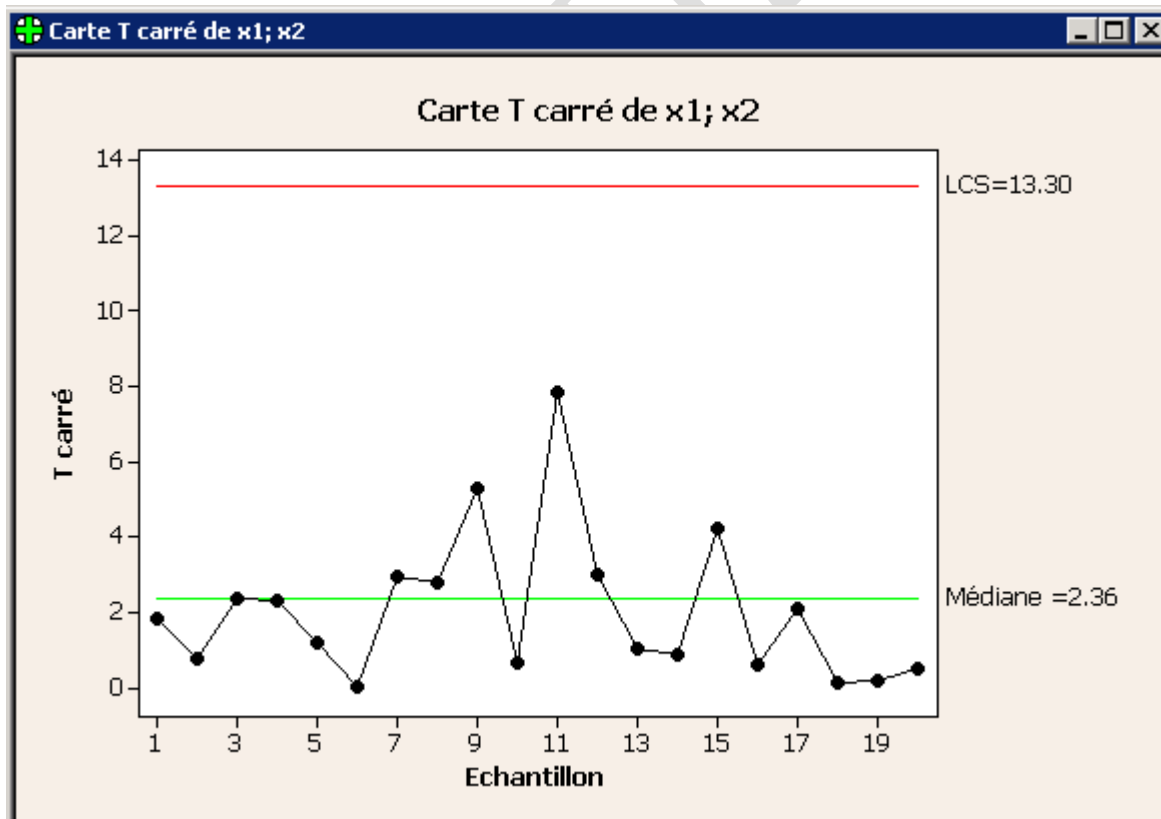
Dans la première boîte de dialogue qui apparaît, nous prenons:



En cliquant sur le bouton **Options carte quad T...** nous allons dans l'onglet **Stockage** pour prendre (afin de comparer avec les calculs faits à la main dans le cours théorique):



en validant le tout par OK nous obtenons pour graphique:

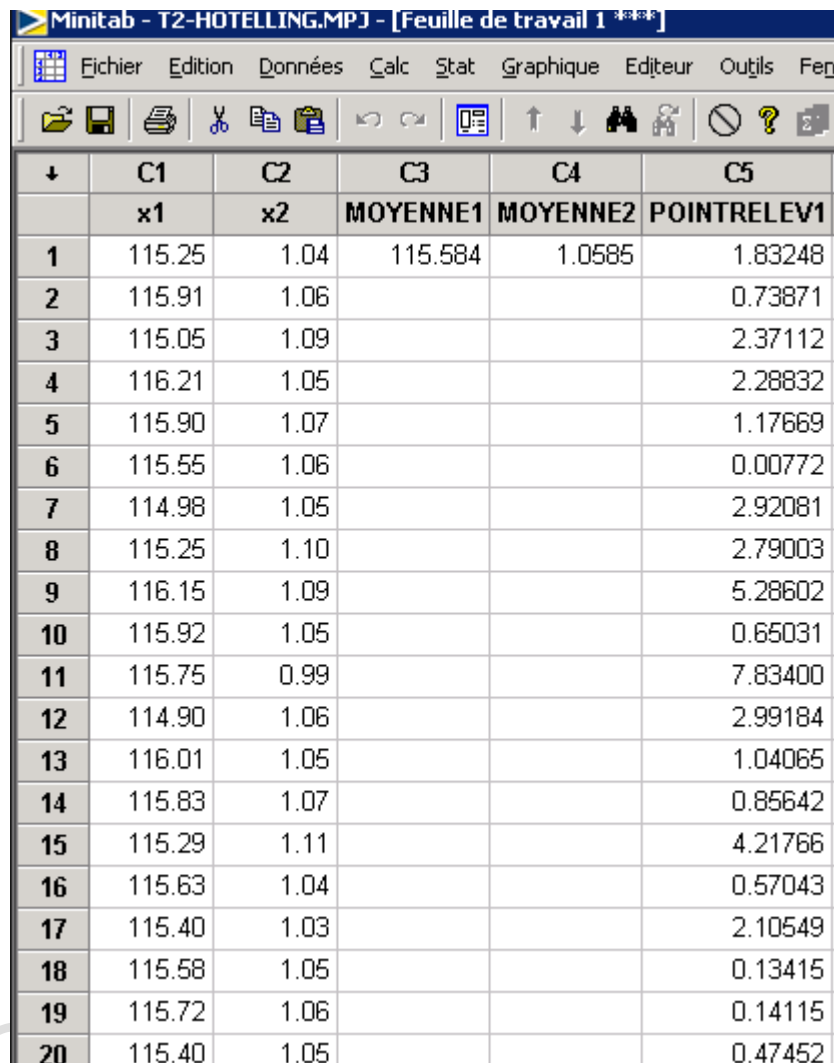


Nous voyons que la LCS est très proche de celle calculée avec la loi du Khi-2 dans le cours théorique avec un seuil de 99.9% (mais en réalité Minitab n'utilise pas le modèle du Khi-2 et n'est pas à un seuil de 99.9% mais autre chose à un seuil de 99.88%...????). L'origine de la médiane de 2.36 est inconnue.



Suivant les logiciels et le modèle théorique utilisé, la limite supérieure est comprise entre 9 et 23... (je conseille alors de prendre la plus petite si le procédé est prouvé comme étant stable pour des raisons évidentes de qualité)

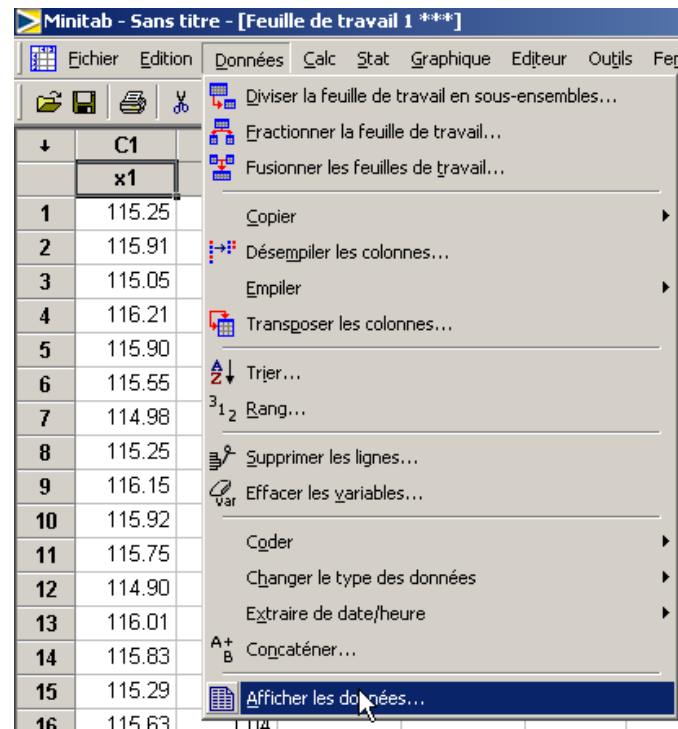
Au niveau des données brutes, Minitab nous donne:



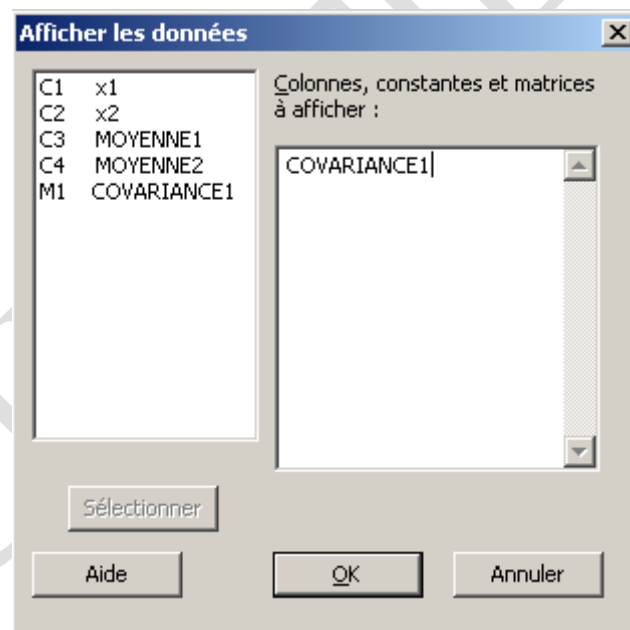
	C1	C2	C3	C4	C5
	x1	x2	MOYENNE1	MOYENNE2	POINTRELEV1
1	115.25	1.04	115.584	1.0585	1.83248
2	115.91	1.06			0.73871
3	115.05	1.09			2.37112
4	116.21	1.05			2.28832
5	115.90	1.07			1.17669
6	115.55	1.06			0.00772
7	114.98	1.05			2.92081
8	115.25	1.10			2.79003
9	116.15	1.09			5.28602
10	115.92	1.05			0.65031
11	115.75	0.99			7.83400
12	114.90	1.06			2.99184
13	116.01	1.05			1.04065
14	115.83	1.07			0.85642
15	115.29	1.11			4.21766
16	115.63	1.04			0.57043
17	115.40	1.03			2.10549
18	115.58	1.05			0.13415
19	115.72	1.06			0.14115
20	115.40	1.05			0.47452

Les moyennes sont correctes, nous remarquons cependant que Minitab pour les points relevés utilise en réalité la matrice de variance-covariance des différences séquentielle des mesures.

Voyons alors ce qu'il a stocké comme matrice des variances-covariances en allant dans le menu **Données/Afficher les données**:



et nous prenons la matrice M1:



Nous validons par **OK** pour obtenir dans la fenêtre de session:



et c'est là que nous constatons un choix bizarre des concepteurs de Minitab. Ils ont choisi de montrer la matrice des variances-covariances des mesures telles que nous l'avons calculée bien évidemment dans le cours théorique mais pourtant ils utilisent la matrice des variances-covariances des différences séquentielle pour faire le calcul des points relevés. Alors certes dans la pratique c'est celle indiquée est qui la plus utile mais pour ceux qui veulent retrouver les calculs faits en arrière-plan c'est trompeur. Il aurait été judicieux de mettre les deux matrices de variances-covariances dans la mémoire de Minitab.

Mais il se peut que nous devions imposer la matrice des variances-covariances dans certains cas. Voyons comment faire en pensant assez intuitivement qu'en reprenant et en imposant la même matrice des variances-covariances, nous allons avoir le même graphique.

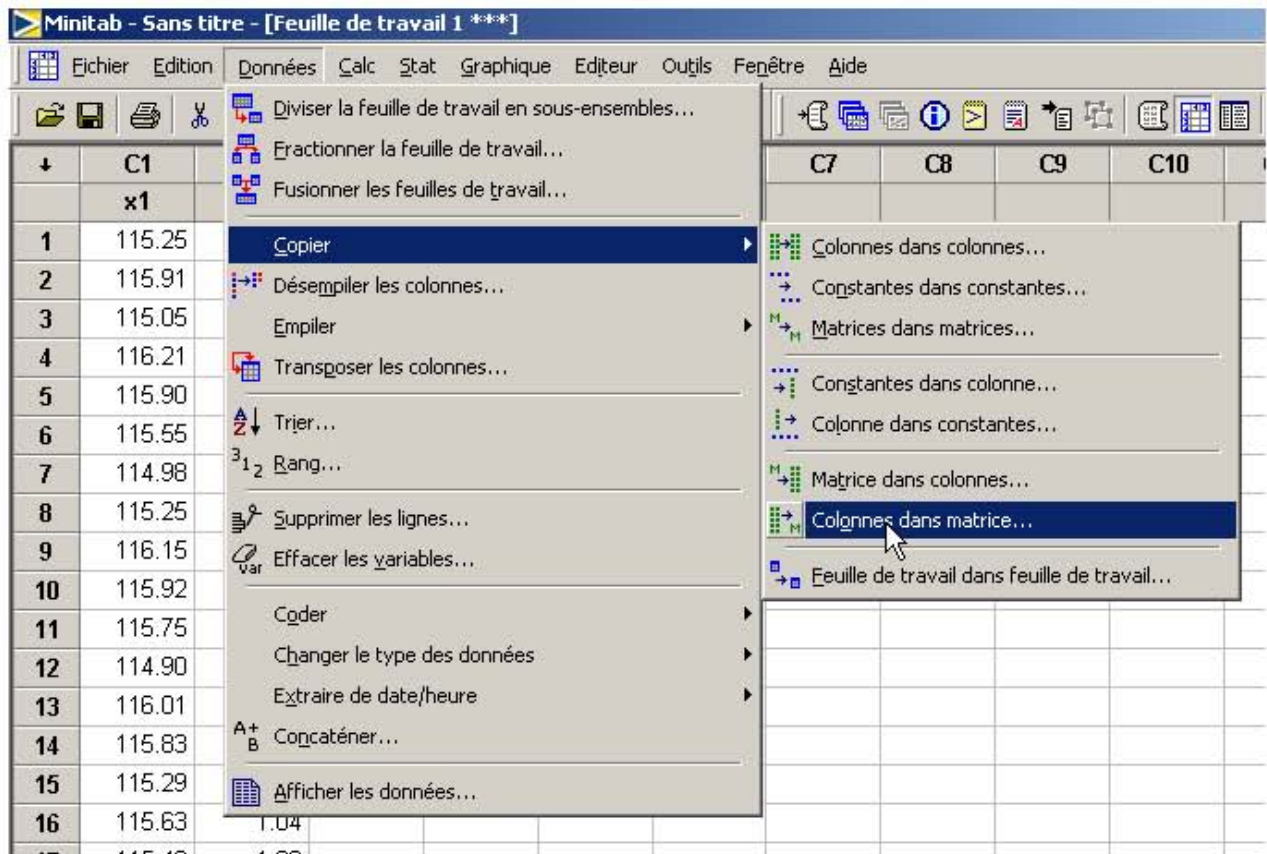
Alors nous saisissons les données (naïvement en pensant bien faire...):

Minitab - T2-HOTELLING.MPJ - [Feuille de travail 1 \*\*\*]

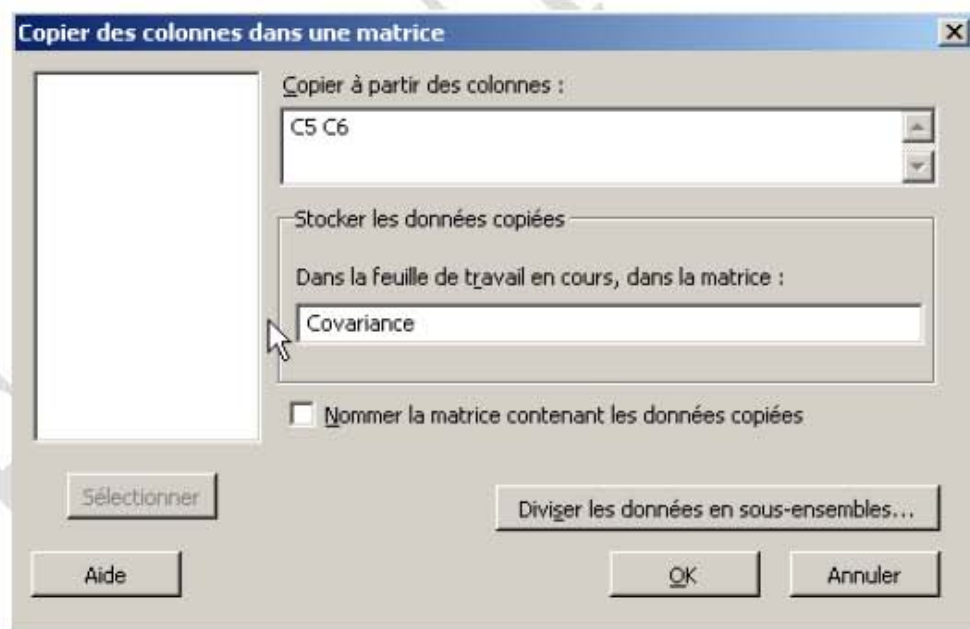
Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre

	C1	C2	C3	C4	C5	C6
	x1	x2				
1	115.25	1.04			0.148299	-0.0014621
2	115.91	1.06			-0.001462	0.0006976
3	115.05	1.09				
4	116.21	1.05				
5	115.90	1.07				
6	115.55	1.06				
7	114.98	1.05				
8	115.25	1.10				
9	116.15	1.09				
10	115.92	1.05				
11	115.75	0.99				
12	114.90	1.06				
13	116.01	1.05				
14	115.83	1.07				
15	115.29	1.11				
16	115.63	1.04				

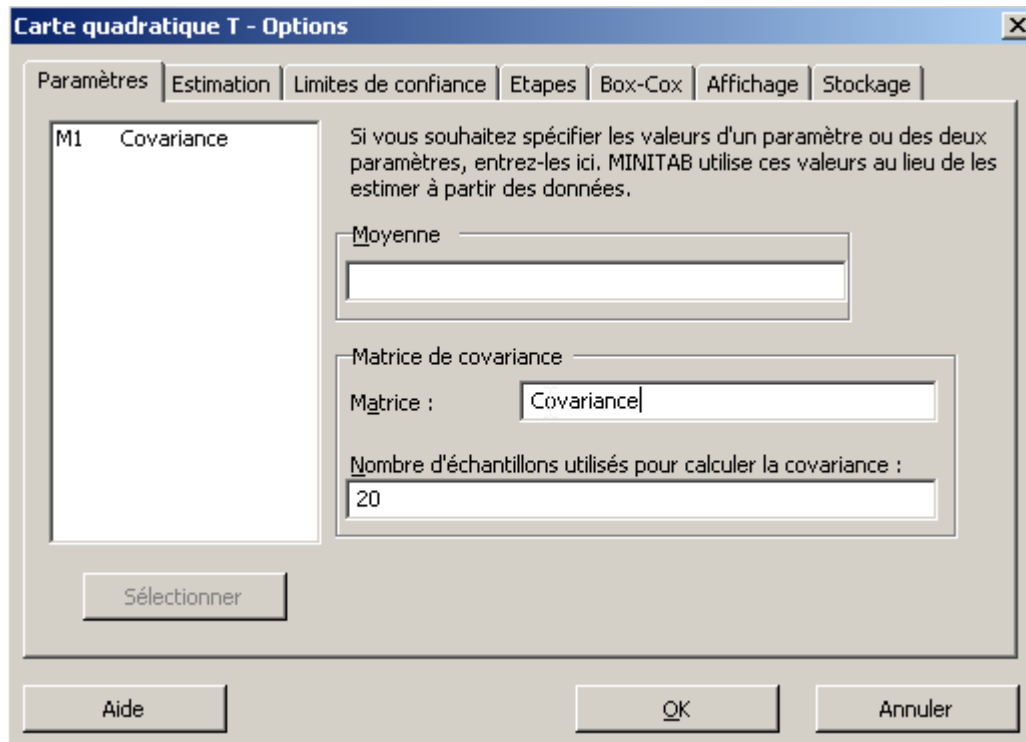
Ensuite, nous allons dans le menu **Données/Copier/Colonnes dans matrice...**:



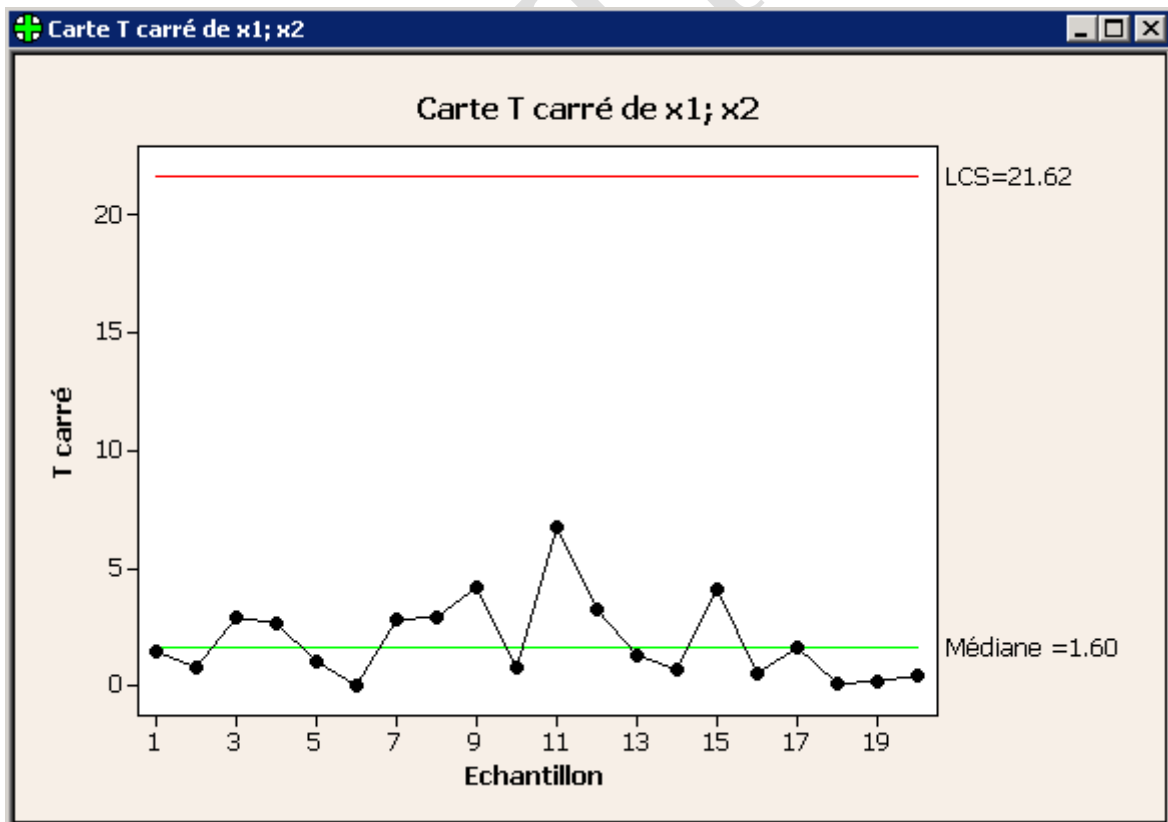
et nous mettons:



Nous validons bien évidemment par **OK** et retournons créer une carte quadratique T de Hotelling mais cette fois-ci en spécifiant la matrice de covariance:



En validant nous obtenons la même carte de contrôle mais toutefois pas avec la même limite ni la même médiane:



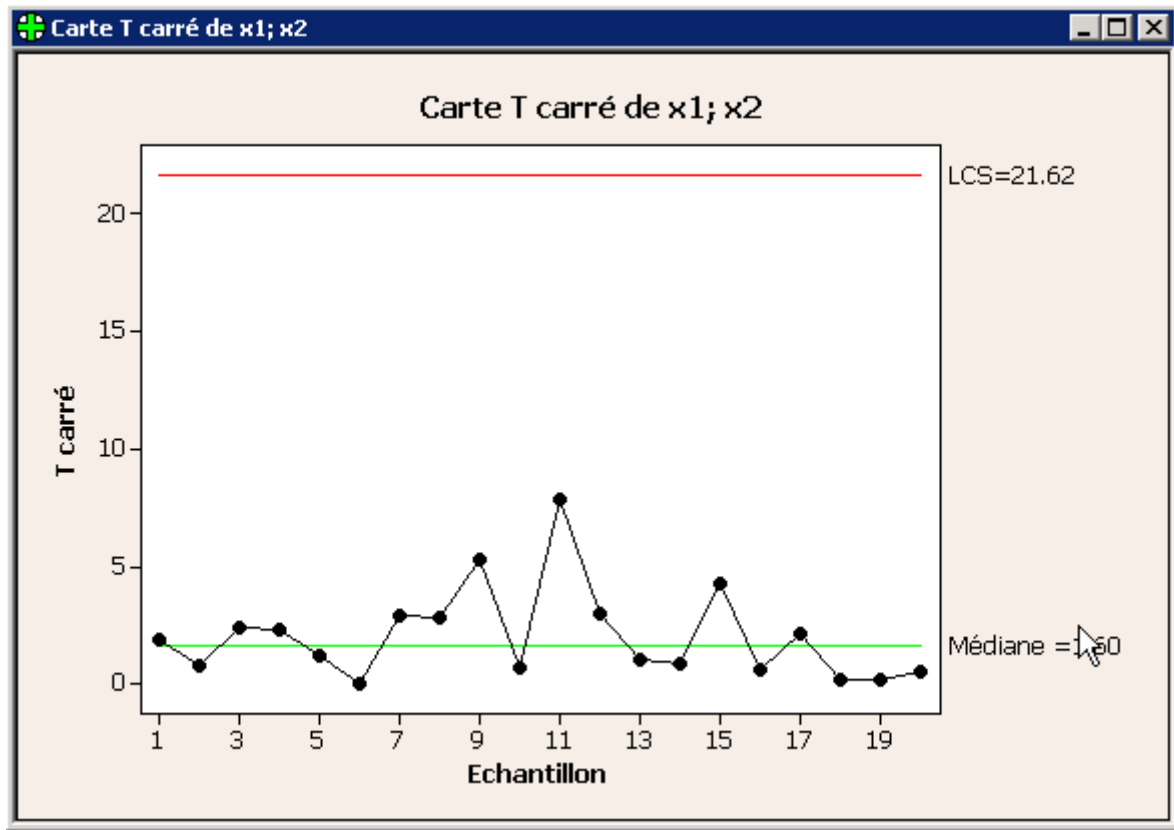
Les points relevés correspondent par contre cette fois-ci bien à ce qui a été calculé dans le cours théorique (je ne sais pas si vous suivez parce que c'est assez confus....) mais par contre plus la limite supérieure indiquée sur le graphique...:

	C1	C2	C3	C4	C5	C6	C7
	x1	x2					POINTRELEV1
1	115.25	1.04			0.148299	-0.0014621	1.44741
2	115.91	1.06			-0.001462	0.0006976	0.74916
3	115.05	1.09					2.93031
4	116.21	1.05					2.65041
5	115.90	1.07					0.98601
6	115.55	1.06					0.00978
7	114.98	1.05					2.76584
8	115.25	1.10					2.88896
9	116.15	1.09					4.17276
10	115.92	1.05					0.80066
11	115.75	0.99					6.72974
12	114.90	1.06					3.19506
13	116.01	1.05					1.25078
14	115.83	1.07					0.69191
15	115.29	1.11					4.04033
16	115.63	1.04					0.49097
17	115.40	1.03					1.57338
18	115.58	1.05					0.10685
19	115.72	1.06					0.13653
20	115.40	1.05					0.38401

Alors comment retomber sur nos pattes? Eh bien en utilisant la matrice des variances-covariances des différences des séquences de mesure que nous avons calculée dans le cours théorique:

	C1	C2	C3	C4	C5	C6
	x1	x2				
1	115.25	1.04			0.174487	-0.0037158
2	115.91	1.06			-0.003716	0.0006289
3	115.05	1.09				
4	116.21	1.05				
5	115.90	1.07				
6	115.55	1.06				
7	114.98	1.05				

et nous relançons la carte de Hotelling et là surprise.... (!), le graphique ne donne toujours pas la même limite supérieur qu'au début (????):



mais avec toutefois (??? O\_o) les points relevés:

Minitab - T2-HOTELLING.MPJ - [Feuille de travail 1 ***]							
Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide							
↓	C1	C2	C3	C4	C5	C6	C7
	x1	x2					POINTRELEV1
1	115.25	1.04			0.174487	-0.0037158	1.83248
2	115.91	1.06			-0.003716	0.0006289	0.73871
3	115.05	1.09					2.37112
4	116.21	1.05					2.28832
5	115.90	1.07					1.17669
6	115.55	1.06					0.00772
7	114.98	1.05					2.92081
8	115.25	1.10					2.79003
9	116.15	1.09					5.28602
10	115.92	1.05					0.65031
11	115.75	0.99					7.83400
12	114.90	1.06					2.99184
13	116.01	1.05					1.04065
14	115.83	1.07					0.85642
15	115.29	1.11					4.21766
16	115.63	1.04					0.57043
17	115.40	1.03					2.10549
18	115.58	1.05					0.13415
19	115.72	1.06					0.14115
20	115.40	1.05					0.47452

qui correspondent bien à la première carte de Contrôle... Bref comme à l'habitude et comme je le dis dans tous mes cours, normalement on prend trois logiciels de statistiques et on compare et on se rend compte qu'aucun ne donne la même chose. Donc on prend soit la voie du milieu soit le cas au pire...

ÉCHANTILLON



## 15. Outils de la qualité

### 15.1. Exercice 145.: Test de normalité et Analyse de capacité (Normale)

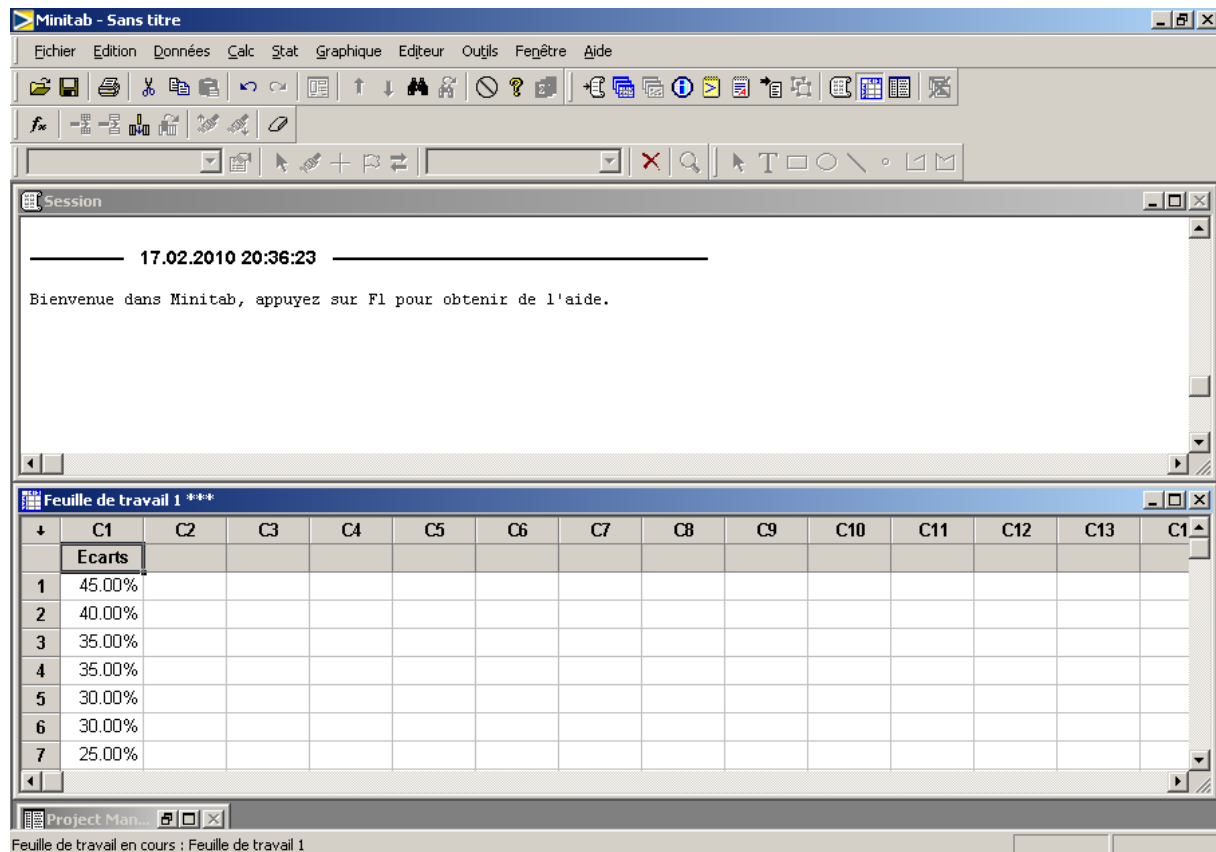
Minitab® Statistical Software 15.1.1.0

Les déviations temporelles en % de 25 tâches donnent les valeurs suivantes après clôture d'un projet:

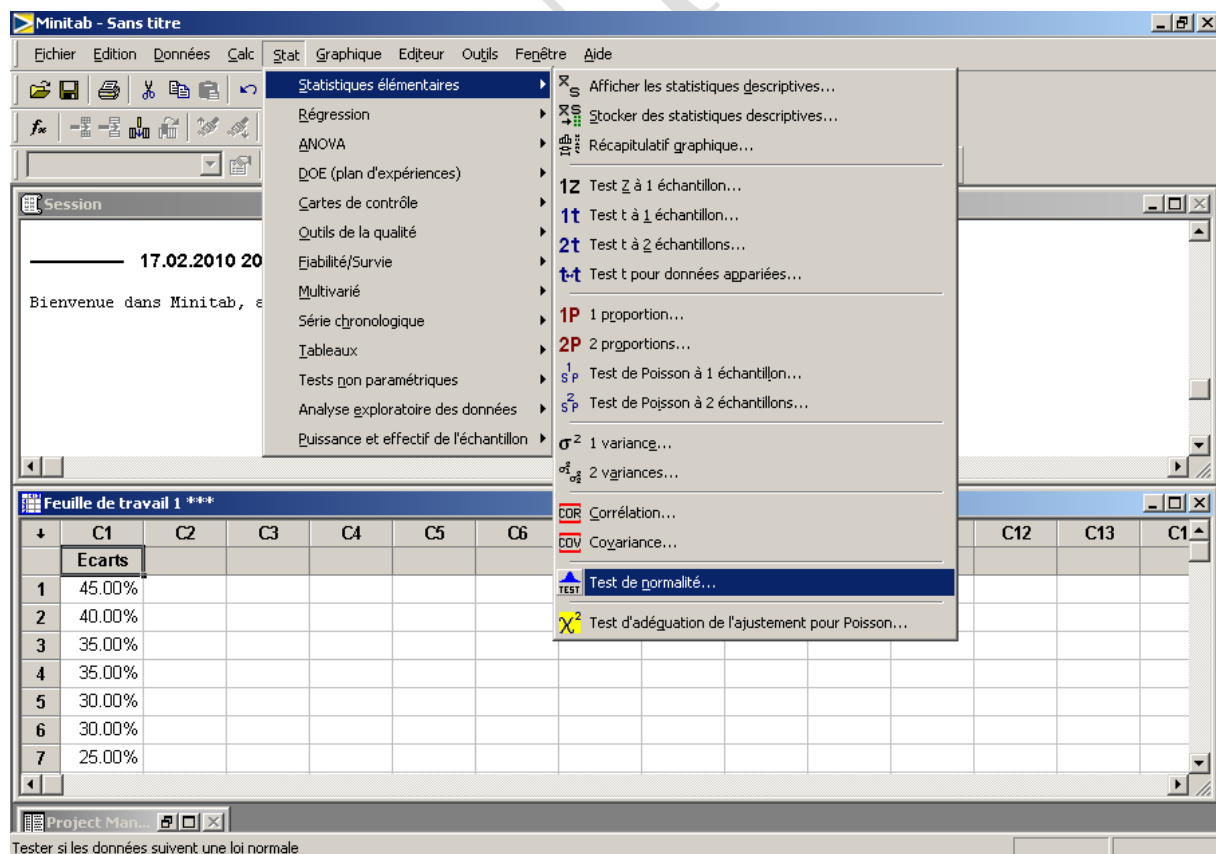
A	B
Ecart	Ecart
45%	15%
40%	15%
35%	15%
35%	10%
30%	10%
30%	10%
25%	5%
25%	5%
25%	0%
20%	-5%
20%	-5%
20%	-10%
20%	

Nous souhaiterions savoir si les estimations du chef de projet sont sous contrôle (test de normalité de Kolmogorov-Smirnov sous Minitab® Statistical Software) et faire une analyse de la capacité des déviations avec Minitab® Statistical Software en prenant l'intervalle  $LSL/USL$  de  $[-0.1, 0.2]$  avec une cible nulle.

Nous ouvrons Minitab® Statistical Software et reportons les données dans une unique colonne (ou nous ouvrons le fichier *CapabiliteProjet.mpj*):



Nous allons dans le menu **Stat/Statistiques élémentaires/Test de normalité...**:

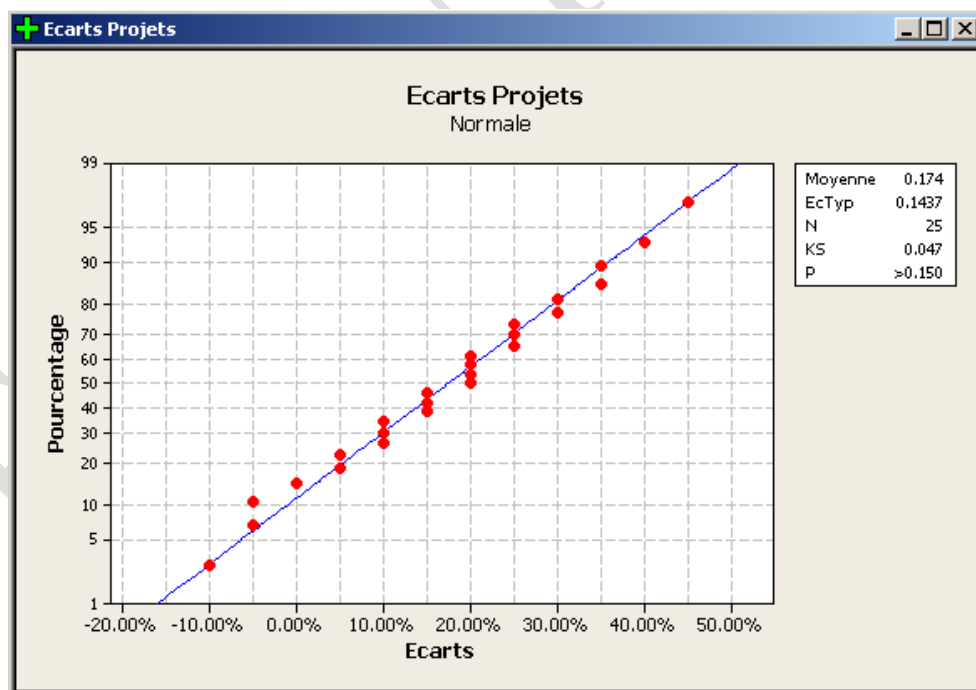


Nous remplissons la boîte de dialogue:

The screenshot shows the Minitab interface. A dialog box titled 'Test de normalité' is open, with 'Ecart' as the variable. Under 'Lignes des percentiles', 'Aucune' is selected. Under 'Tests de normalité', 'Kolmogorov-Smirnov' is selected. The title of the test is 'Ecart Projets'. The background worksheet shows the following data for 'Ecart':

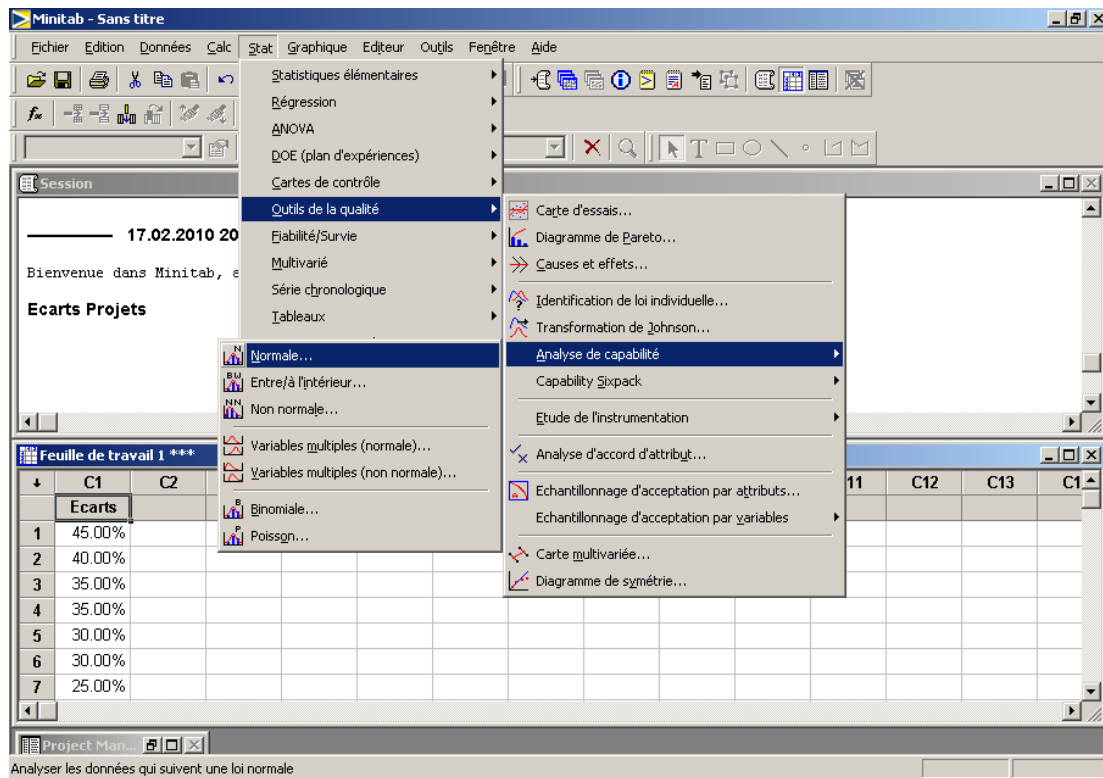
	C1	C2	C3
	Ecart		
1	45.00%		
2	40.00%		
3	35.00%		
4	35.00%		
5	30.00%		
6	30.00%		
7	25.00%		

et validons par **OK** ce qui donne:

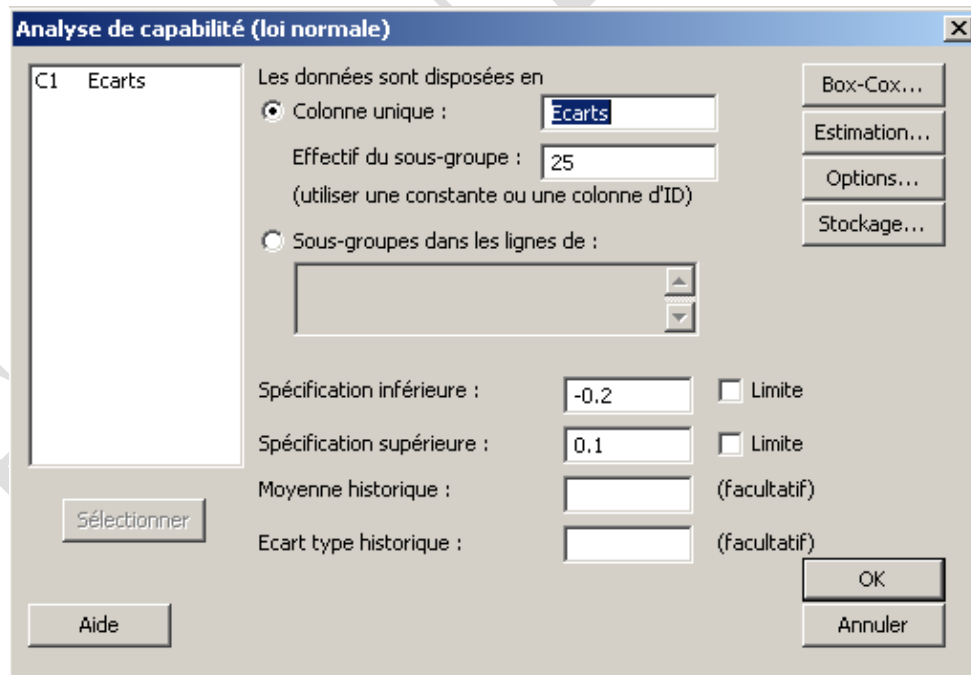


La  $p$ -value étant supérieur à 0.05 (puisque Minitab® Statistical Software indique qu'elle est plus grande que 0.150) fait que nous ne pouvons rejeter l'hypothèse de normalité.

Pour faire l'analyse de capabilité nous allons maintenant dans **Outils de la qualité/Analyse de capabilité/Normale...**:

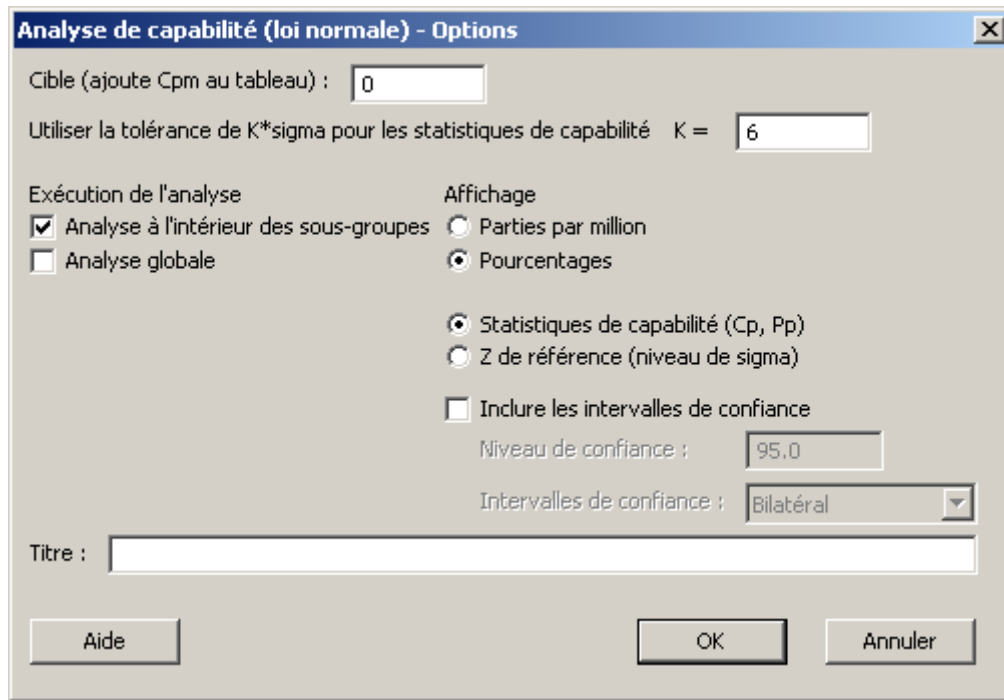


Nous paramétrons l'analyse de capacité<sup>3</sup>:

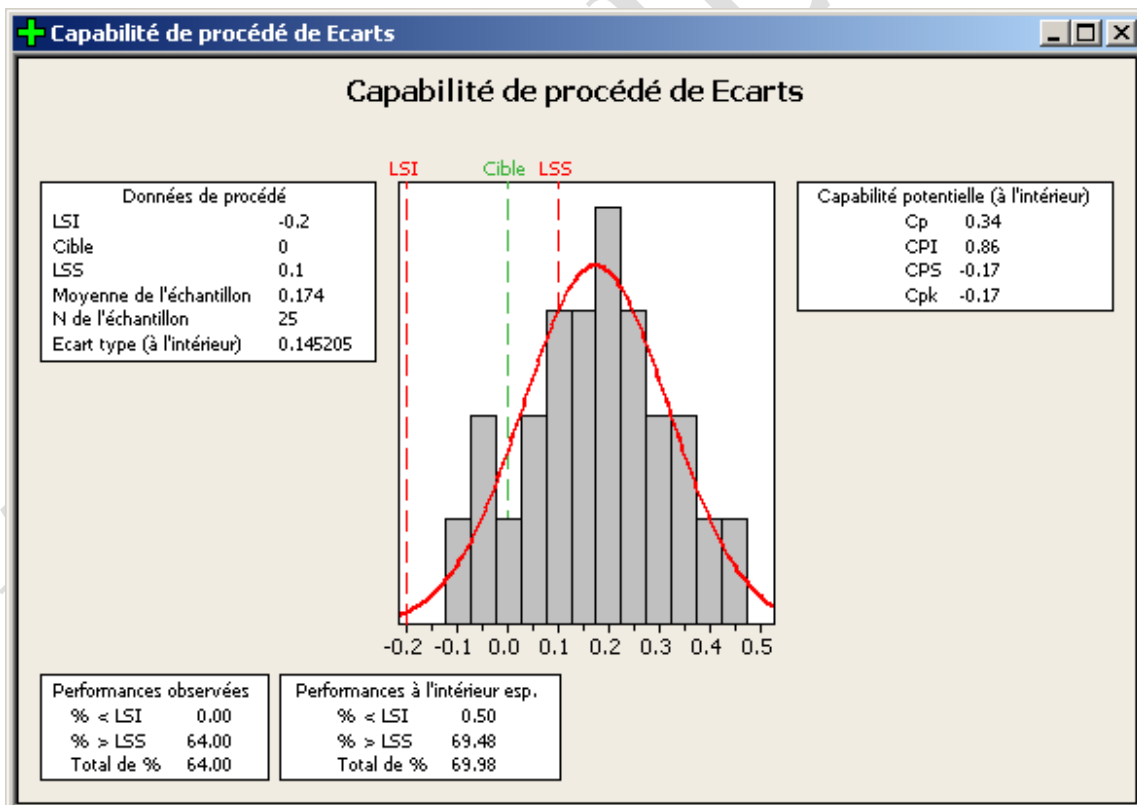


nous cliquons sur le bouton **Options...** et complétons la boîte de dialogue:

<sup>3</sup> **Attention!** Cocher "limite" pour une de vos spécifications dans Minitab donne lieu à une analyse de capacité unilatérale. Si vous définissez les deux limites de spécification comme bornes, toutes les statistiques de capacité apparaîtront sous forme d'astérisques (\*) dans les résultats.



Graphiquement nous avons finalement en validant deux fois par **OK**:



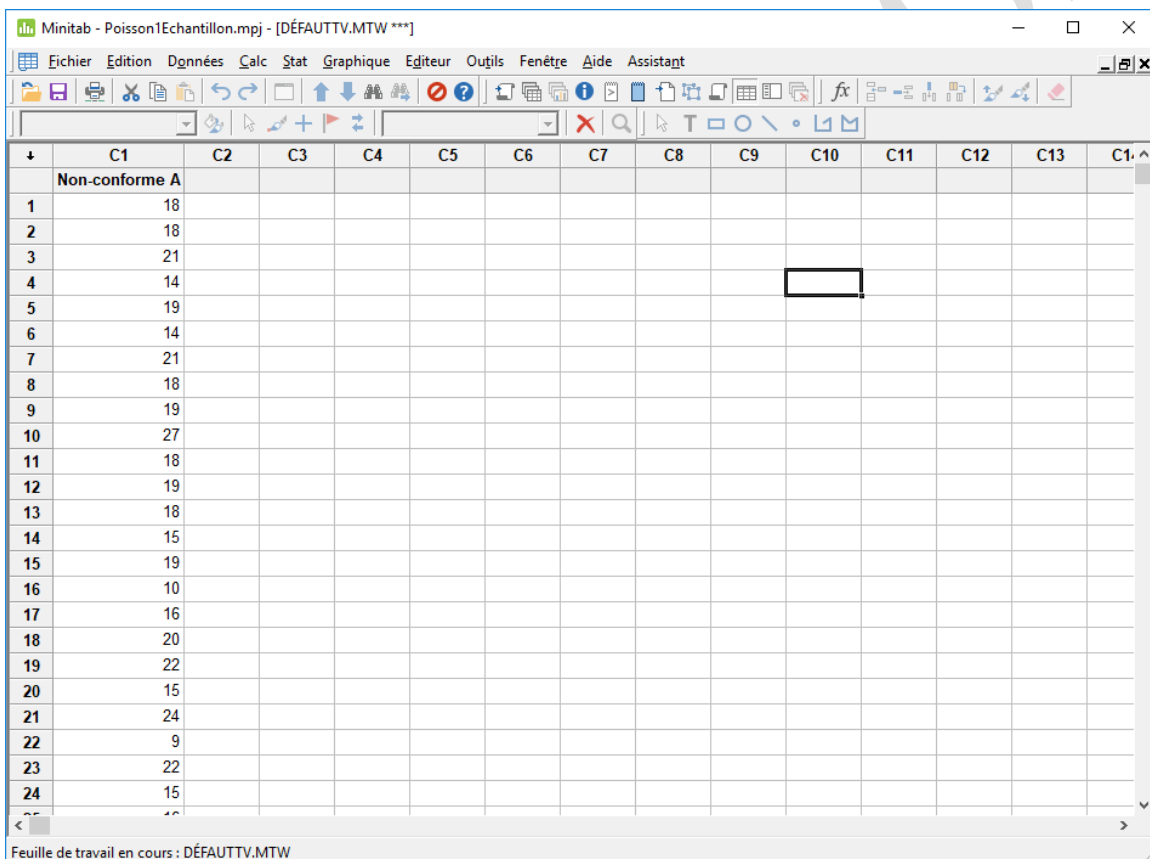
Nous retrouvons donc dans ce graphique toutes les notions vues et démontrées mathématiquement en détail dans le cours théorique.

## 15.2. Exercice 146.: Analyse de capabilité de Poisson (événements rares)

Minitab® Statistical Software 18.1

Nous reprenons l'exemple (fictif) vu plus haut d'une société fabrique des télévisions en quantité constante et a mesuré le nombre d'appareils défectueux produits chaque trimestre pendant les dix dernières années (donc 4 fois 10 mesures = 40 trimestres). La direction décide que le nombre maximum acceptable d'unités défectueuses est de 20 par trimestre et souhaite déterminer si l'usine satisfait à ces exigences (sous l'hypothèse que la distribution des défectueux suive une loi de Poisson) à un niveau de confiance de 5%.

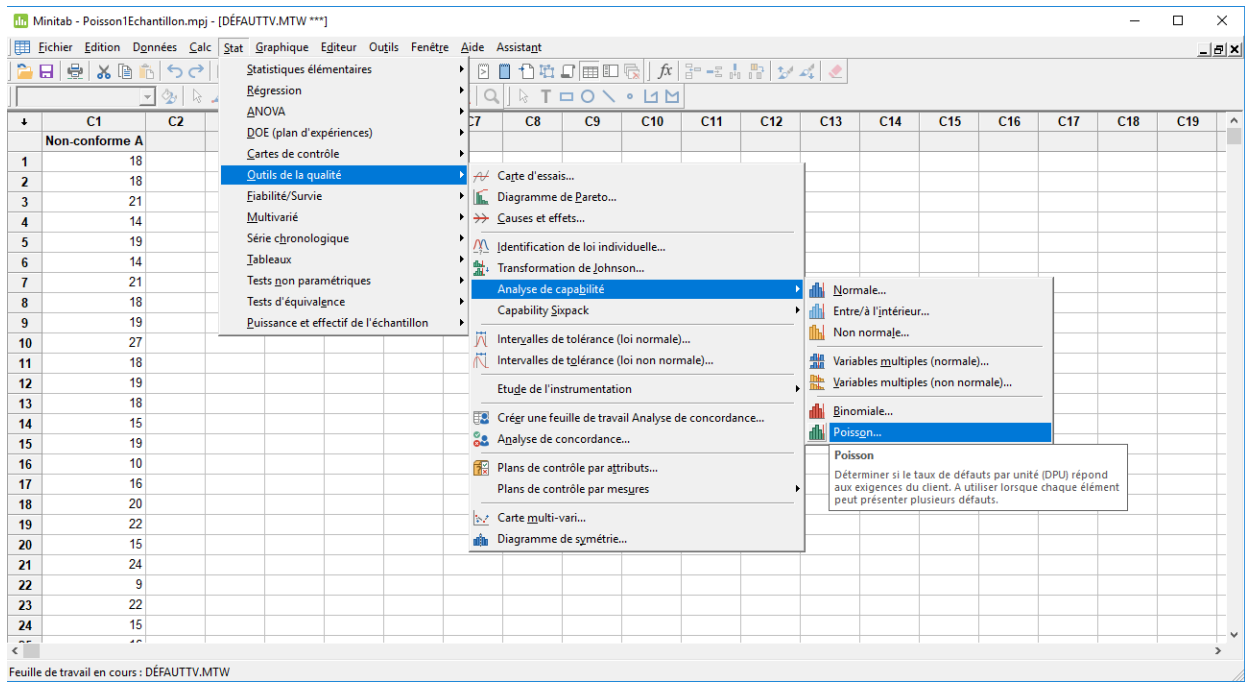
Nous ouvrons donc le fichier *Poisson1Echantillon.mpj* comportant 40 mesures (donc une pour chaque trimestre):



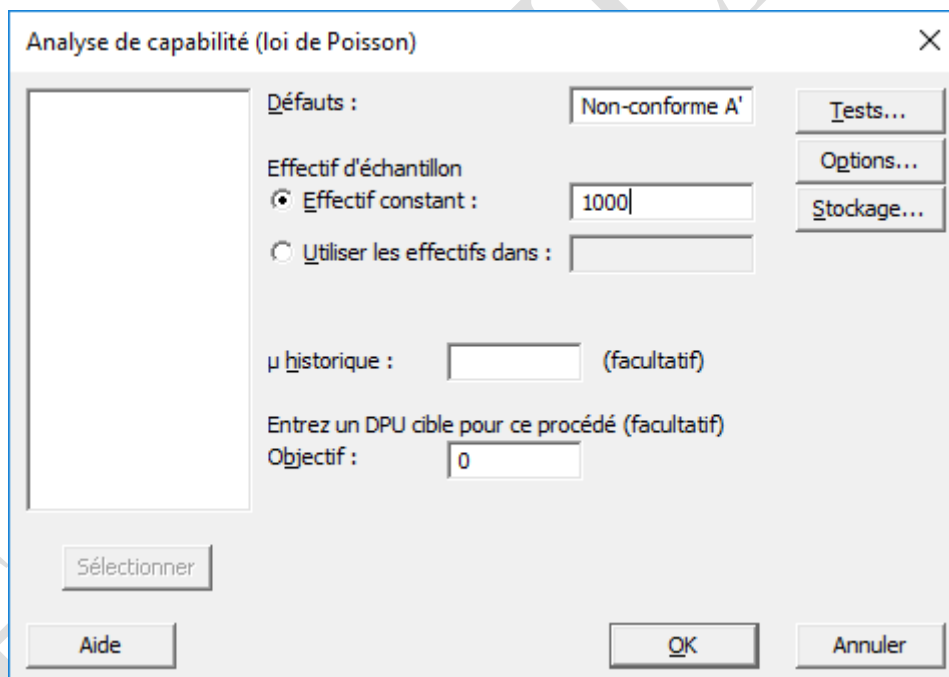
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Non-conforme A													
1	18													
2	18													
3	21													
4	14													
5	19													
6	14													
7	21													
8	18													
9	19													
10	27													
11	18													
12	19													
13	18													
14	15													
15	19													
16	10													
17	16													
18	20													
19	22													
20	15													
21	24													
22	9													
23	22													
24	15													

Nous assumons que la société en question aura toujours produit un nombre constant de 1000 téléviseurs par jour.

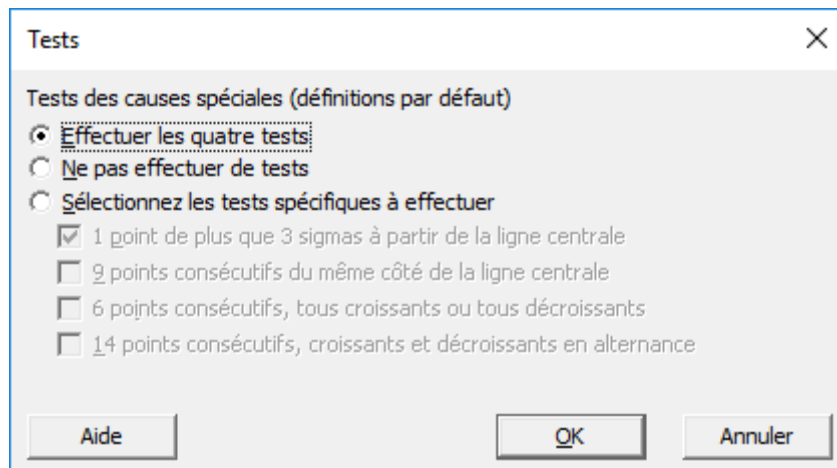
Nous allons ensuite dans **Outils de la qualité/Analyse de capabilité/Poisson...**:



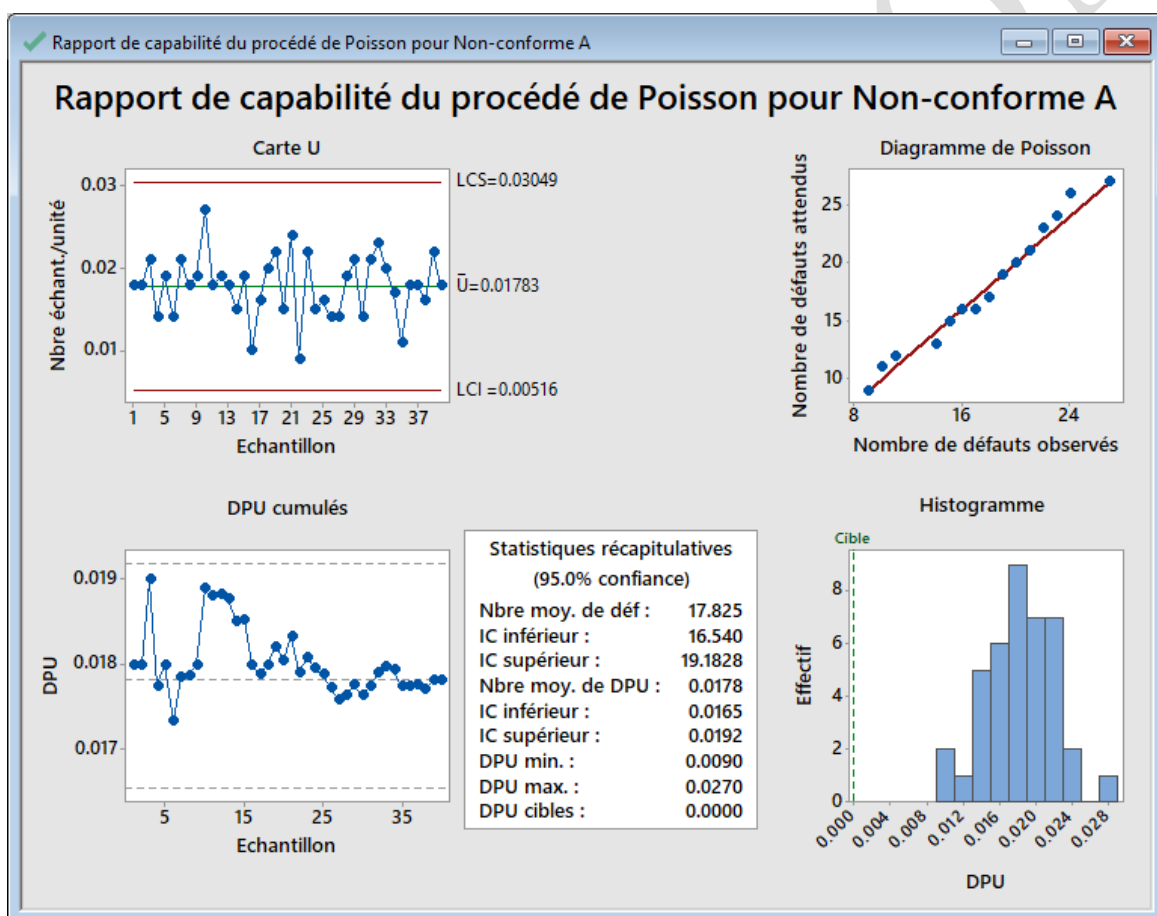
Ensuite, nous mettons les options suivantes:



Dans Tests..., nous prenons l'option **Effectuer les quatre tests**:



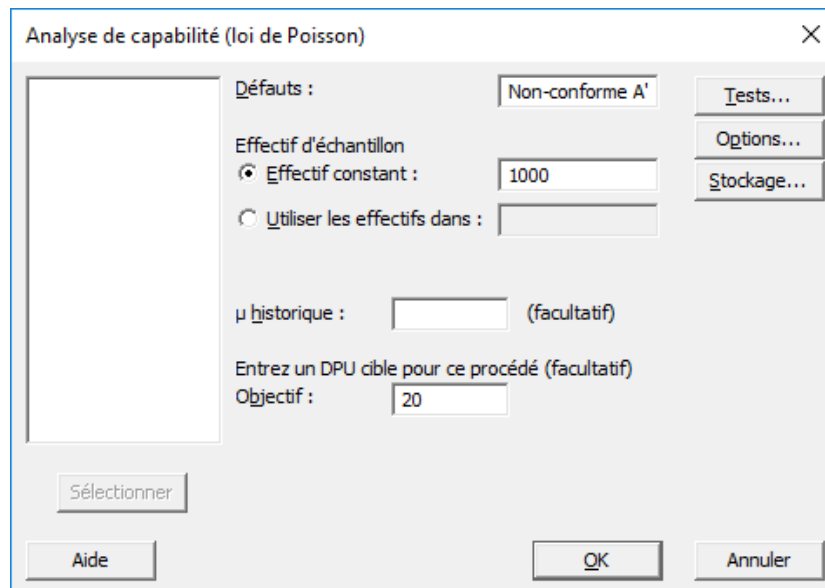
Ce qui donne:



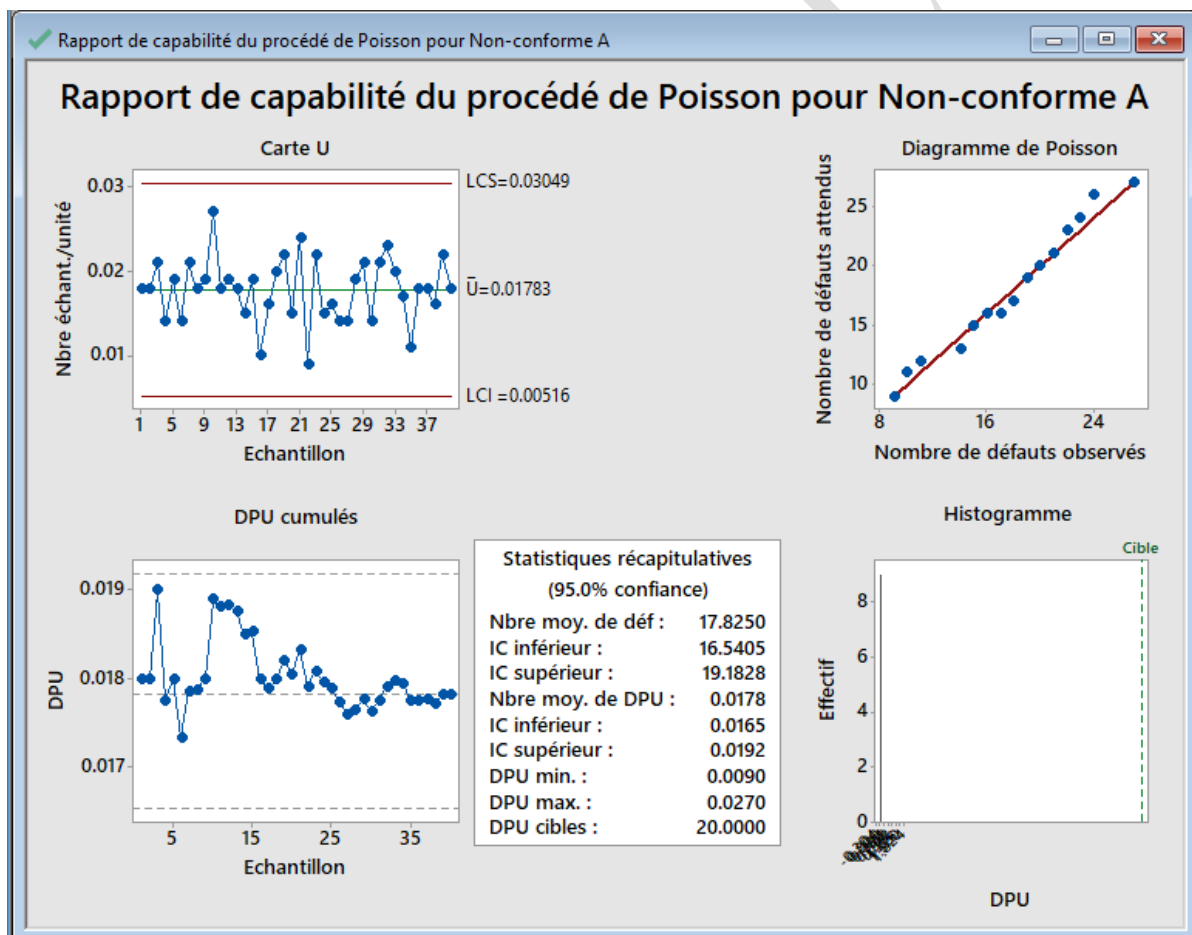
Nous voyons que tous les calculs sont conformes et utilisent bien les relations démontrées dans le cours théorique!

Maintenant refaisons la même manipulation mais en mettant un objectif comme illustré ci-dessous:





Nous comprenons au vu du résultat pourquoi ce paramètre est facultatif....:

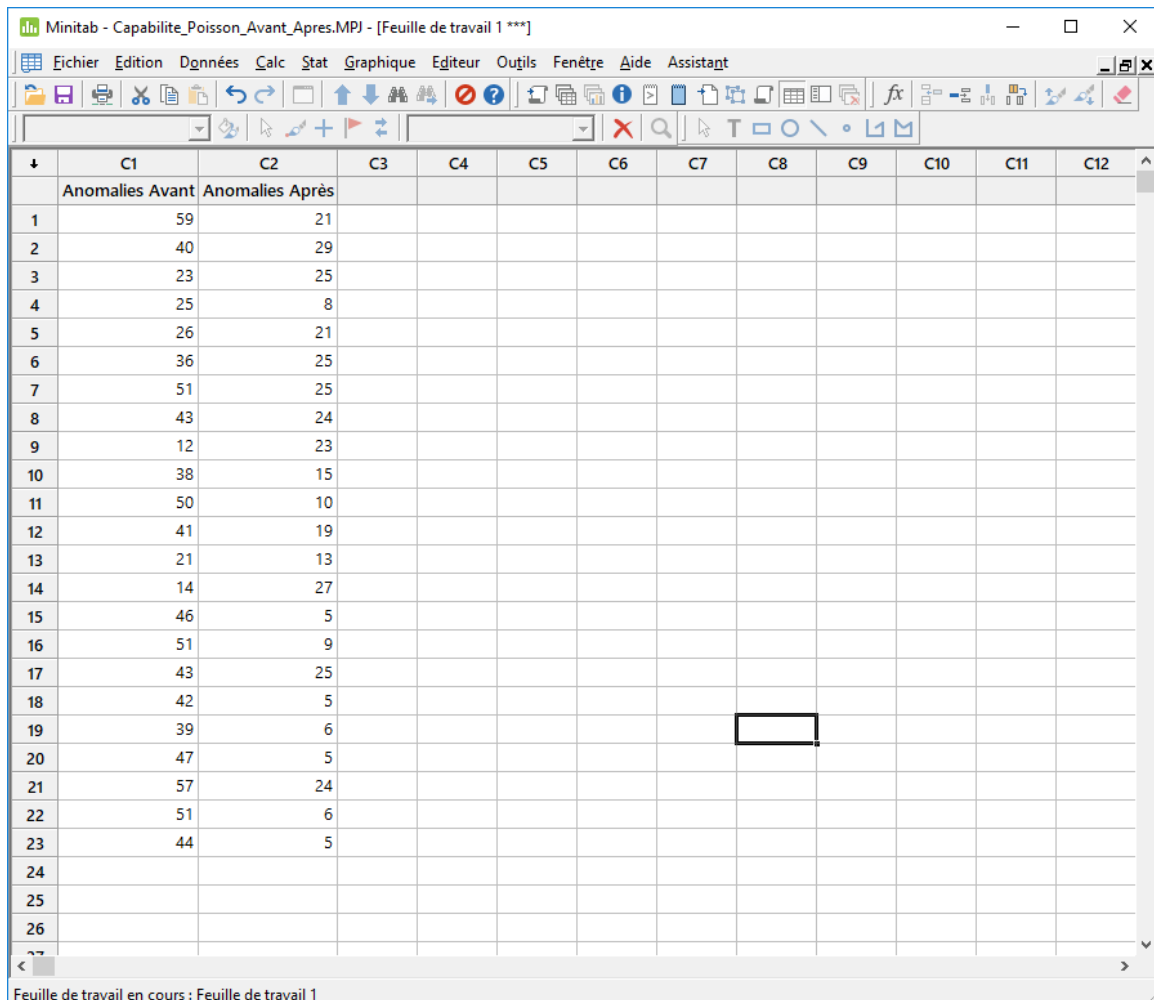


### 15.3. Exercice 147.: Analyse de capabilité de Poisson avant/après

Minitab® Statistical Software 18.1

Voyons ici une application qui combine plusieurs notions vues précédemment et donc aussi démontrées dans le cours théorique.

Considérons les données suivantes du fichier *Capabilite\_Poisson\_Avant\_Apres.mpj*:



The screenshot shows the Minitab software interface with a data table. The table has 12 columns labeled C1 through C12. The first two columns are labeled 'Anomalies Avant' and 'Anomalies Après'. The data is as follows:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
	Anomalies Avant	Anomalies Après										
1	59	21										
2	40	29										
3	23	25										
4	25	8										
5	26	21										
6	36	25										
7	51	25										
8	43	24										
9	12	23										
10	38	15										
11	50	10										
12	41	19										
13	21	13										
14	14	27										
15	46	5										
16	51	9										
17	43	25										
18	42	5										
19	39	6										
20	47	5										
21	57	24										
22	51	6										
23	44	5										
24												
25												
26												
27												

Après nous allons dans le menu **Assistant/Analyse de capabilité Avant/Après...**:

Minitab - Capabilite\_Poisson\_Avant\_Apres.MPJ - [Feuille de travail 1 \*\*\*]

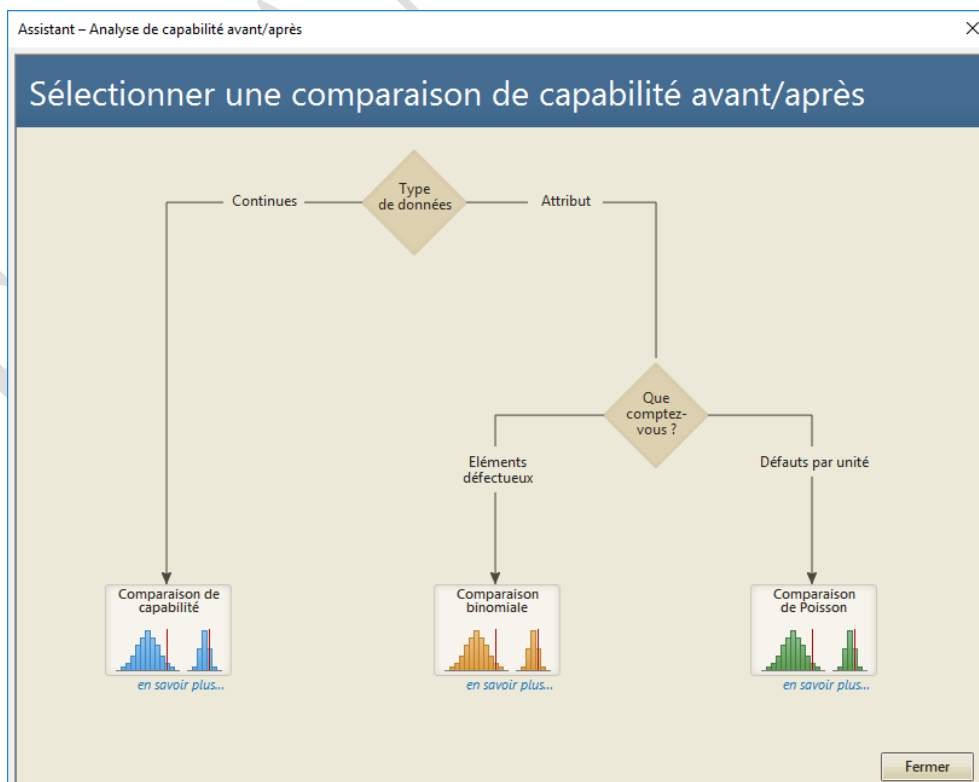
	C1	C2	C3	C4	C5	C6
	Anomalies Avant	Anomalies Après				
1	59	21				
2	40	29				
3	23	25				
4	25	8				
5	26	21				
6	36	25				
7	51	25				
8	43	24				
9	12	23				
10	38	15				
11	50	10				
12	41	19				
13	21	13				
14	14	27				
15	46	5				
16	51	9				
17	43	25				
18	42	5				
19	39	6				
20	47	5				
21	57	24				
22	51	6				
23	44	5				
24						
25						
26						
27						

Assistant

- Analyse des systèmes de mesure (MSA)...
- Analyse de capacité...
- Analyse graphique...
- Tests d'hypothèse...
- Régression...
- DOE (plan d'expériences)
- Analyse de capacité Avant/Après...**
- Cartes de contrôle
- Cartes de contrôle

Analyse de capacité (avant/après)  
Utilisez cette option pour évaluer si le procédé est plus en mesure de répondre aux exigences des clients suite à la modification apportée.

Apparaît alors:



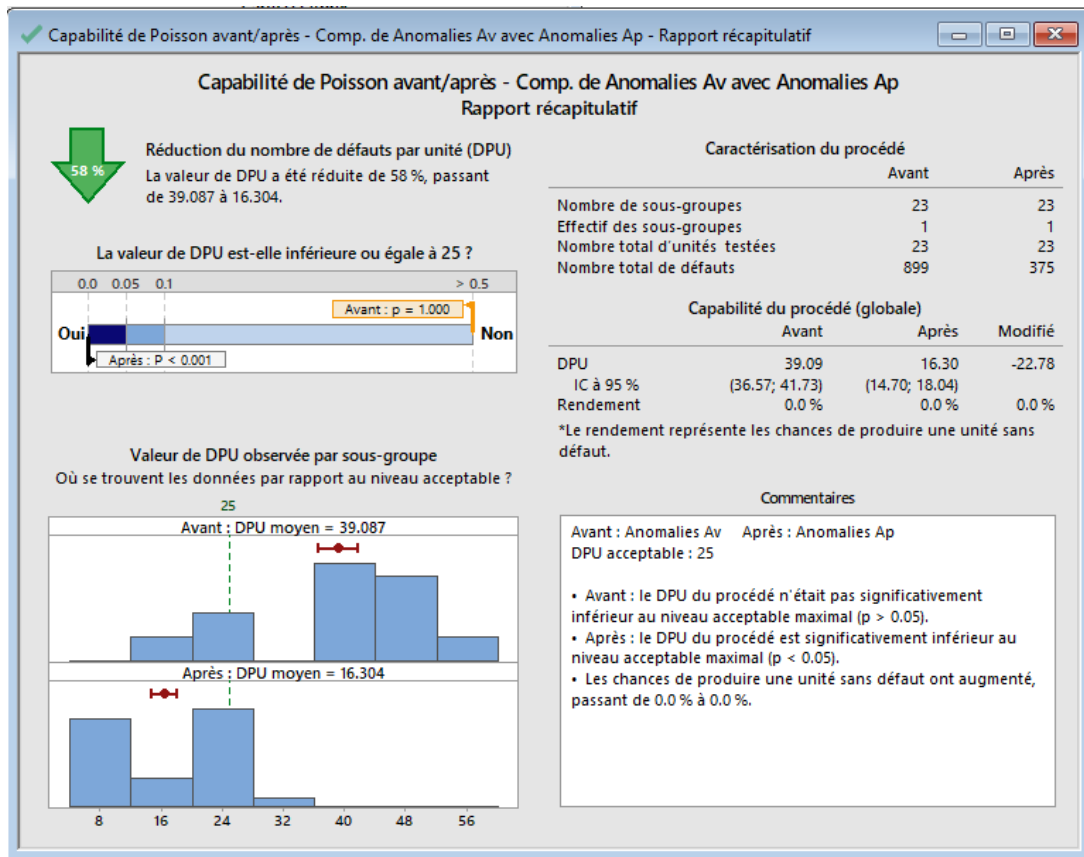
Nous sélectionnons **Comparaison de Poisson** et mettons:

The screenshot shows the Minitab dialog box for 'Analyse de capabilité (avant/après) du pourcentage de défectueux par unité (Poisson)'. It is divided into three sections: 'Données du procédé de référence', 'Données du procédé amélioré', and 'Configuration du test'. In the 'Données du procédé de référence' section, 'Colonne du nombre de défectueux' is set to 'Anomalies Avant', and 'Effectif constant pour tous les sous-groupes' is set to 1. The 'Données du procédé amélioré' section has 'Colonne du nombre de défectueux' set to 'Anomalies Après' and 'Effectif constant pour tous les sous-groupes' set to 1. In the 'Configuration du test' section, 'Quel nombre de défectueux par unité voulez-vous accepter?' has 'Maximum' set to 25. There is a 'Sélectionner' button on the left and 'OK' and 'Annuler' buttons at the bottom right.

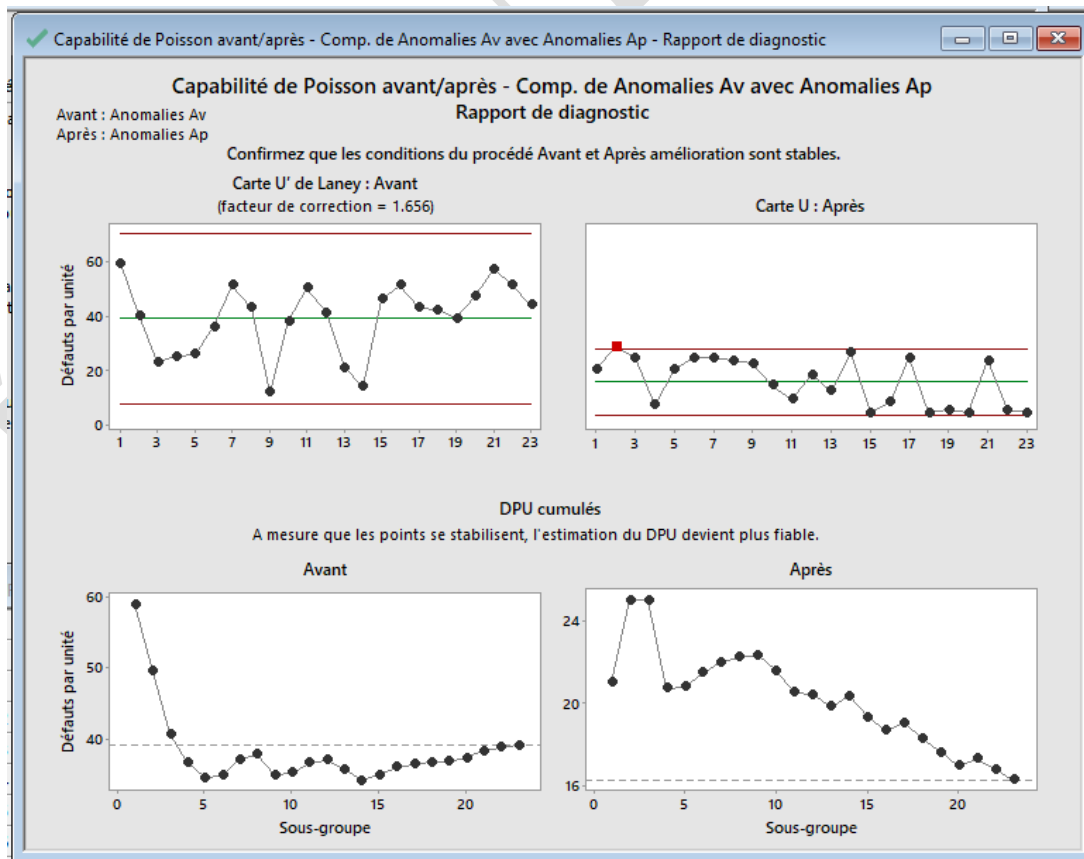
Nous avons alors le message suivant une fois que nous cliquons sur **OK**:

The screenshot shows a Minitab warning dialog box with a yellow warning icon. The text reads: 'Vos données Avant présentent une variation supérieure à celle attendue. En raison de cette situation, appelée surdispersion, les limites de contrôle d'une carte U sont trop étroites pour les données, ce qui entraîne la génération d'un plus grand nombre de fausses alarmes. Utilisez la carte U' de Laney au lieu de la carte U. La carte U' corrige les limites de contrôle afin de tenir compte de la surdispersion. Voulez-vous utiliser la carte U' de Laney?'. At the bottom, there are 'Oui' and 'Non' buttons.

Et nous obtenons comme premier graphique:



et comme deuxième graphique:



et au final un résumé:

Vérifier	Statut	Description
Stabilité		Pour les données Avant et Après amélioration, le nombre de défauts par unité est stable. Aucun point n'est hors contrôle.
Nombre de sous-groupes		Les données Avant et Après amélioration ont moins de 25 sous-groupes. Pour une analyse de capabilité, il est en général recommandé de collecter au moins 25 sous-groupes pendant une période assez longue pour identifier les différentes sources de variation du procédé.
Variation attendue		Bien que la variation dans vos données ne corresponde pas à la variation attendue (sur <a href="#">Tableau</a> ou sous-dispersion), vous avez opté pour la carte U' de Laney qui corrige les limites de contrôle afin de tenir compte de cette condition. La carte U' figurant dans le rapport de diagnostic doit indiquer les alarmes de manière appropriée.
Quantité de données		Les intervalles de confiance à 95 % pour le nombre de défauts par unité sont les suivants : Avant : 36.57; 41.73 Après : 14.70; 18.04 Si ces intervalles sont trop larges pour votre application, vous pouvez collecter davantage de données pour augmenter la précision.

## 15.4. Exercice 148.: Transformations de Box-Cox

Minitab® Statistical Software 15.1.1.0

Bon le but ici ne sera pas de débattre de la justesse de ces transformations (ou de leur origine scientifique) qui ont donc pour rappel de "normaliser"<sup>4</sup> des données. Ceci a déjà été fait dans le cours théorique et de plus, il suffit de Googler un peu pour se faire sa propre opinion sur cette technique "d'ingénierie statistique".

Rappelons cependant les deux relations empiriques définies par Box et Cox dans leur article d'origine de 1964 (il s'agit d'une capture d'écran pour ceux qui douteraient):

### 3. TRANSFORMATION OF THE DEPENDENT VARIABLE

We work with a parametric family of transformations from  $y$  to  $y^{(\lambda)}$ , the parameter  $\lambda$ , possibly a vector, defining a particular transformation. Two important examples considered here are

$$y^{(\lambda)} = \begin{cases} \frac{y^\lambda - 1}{\lambda} & (\lambda \neq 0), \\ \log y & (\lambda = 0), \end{cases} \quad (1)$$

and

$$y^{(\lambda)} = \begin{cases} \frac{(y + \lambda_2)^{\lambda_1} - 1}{\lambda_1} & (\lambda_1 \neq 0), \\ \log(y + \lambda_2) & (\lambda_1 = 0). \end{cases} \quad (2)$$

basé sur la minimisation de l'écart-type (selon ce qui est écrit dans l'article d'origine) et une variante

For the simple power transformation

$$z^{(\lambda)} = \frac{y^\lambda - 1}{\lambda \bar{y}^{\lambda-1}}$$

where  $\bar{y}$  is the geometric mean of the observations.

Minitab® Statistical Software intègre cette dernière relation. Au vu de la forme de la dernière expression, il paraît clair que les transformations de Box-Cox ne marchent que si et seulement si:

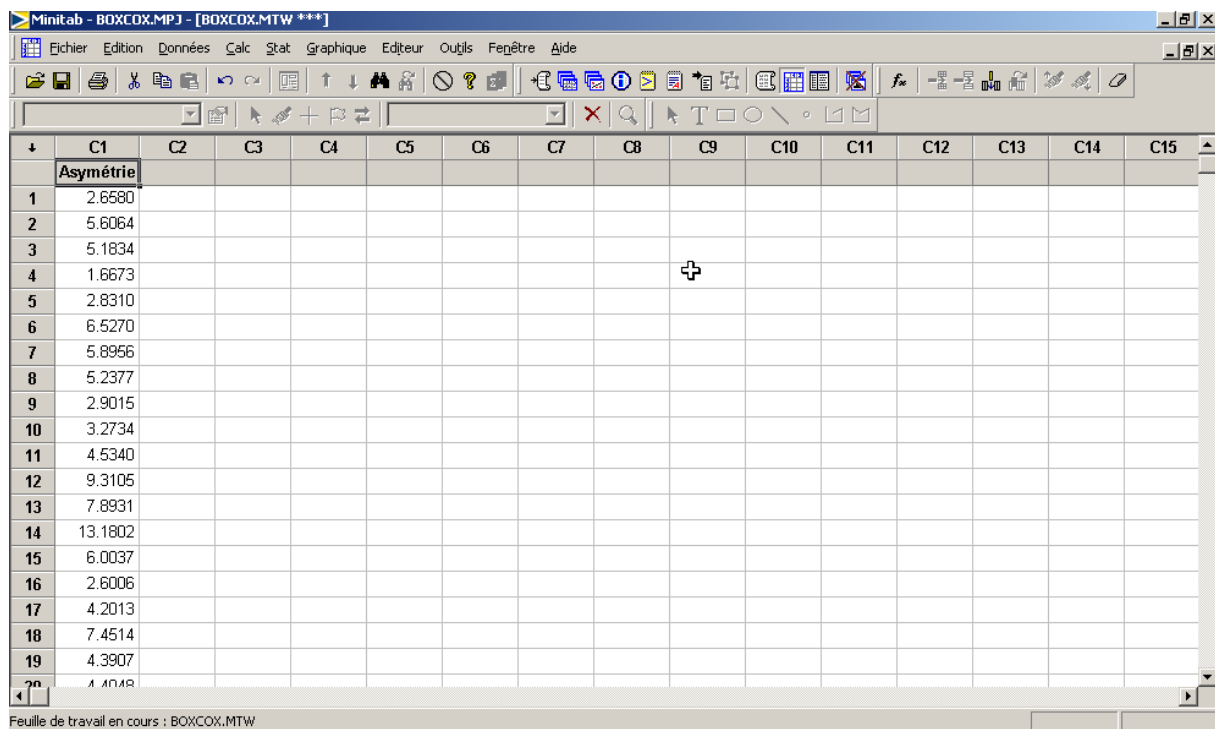
$$y > 0$$

ce qui limite de façon importante son utilité.

Nous n'allons pas nous casser la tête dans le présent exemple. Nous allons juste faire usage du fichier exemple donné par Minitab® Statistical Software et juste comparer le résultat à celui que nous obtenons avec le solveur de Microsoft Excel.

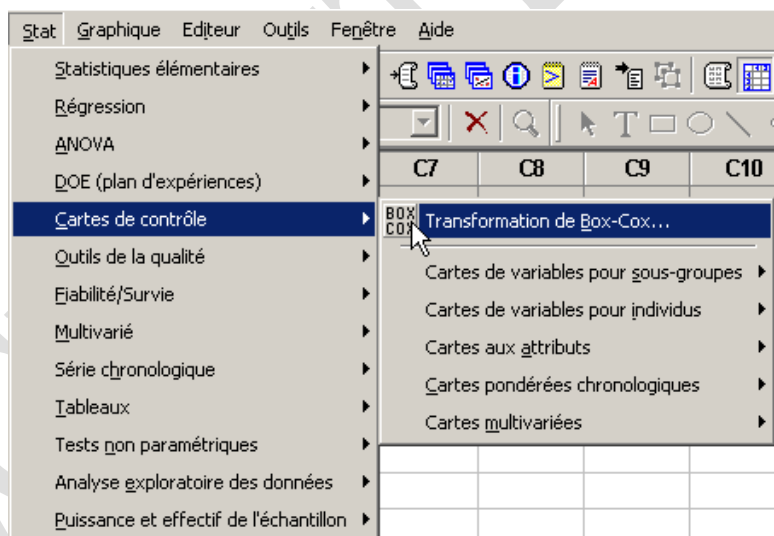
<sup>4</sup> Dans le sens: faire que la distribution ressemble le plus possible à une loi Normale

Ouvrez donc le fichier *BoxCox.mpj* avec ses 125 données:



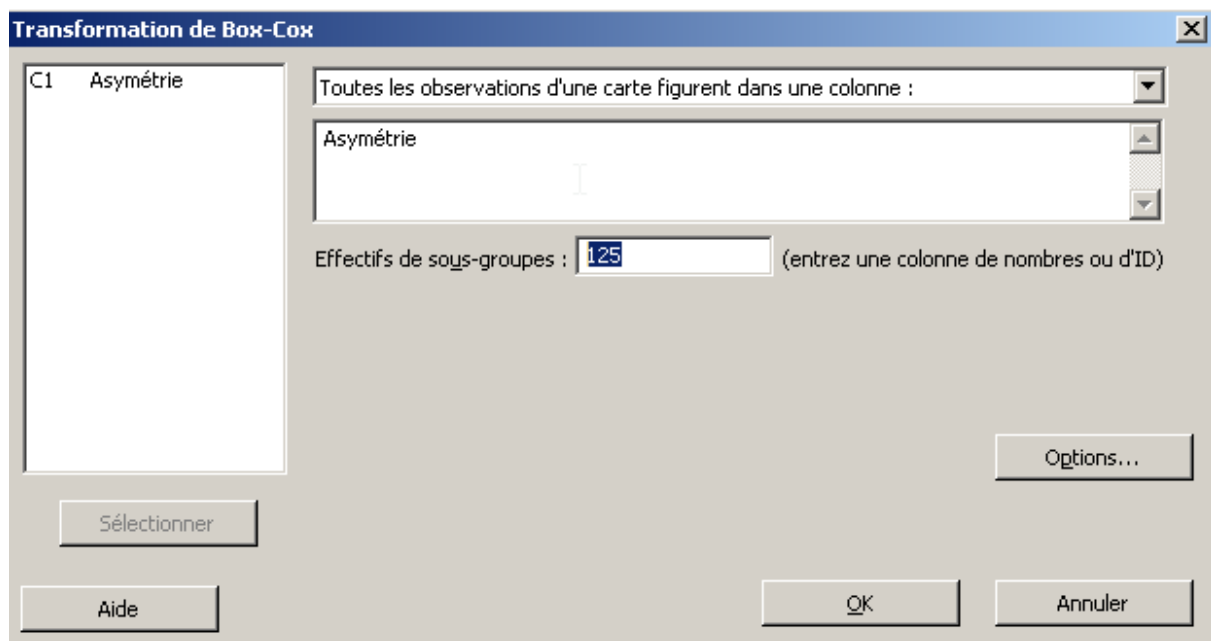
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
	Asymétrie														
1	2.6580														
2	5.6064														
3	5.1834														
4	1.6673														
5	2.8310														
6	6.5270														
7	5.8956														
8	5.2377														
9	2.9015														
10	3.2734														
11	4.5340														
12	9.3105														
13	7.8931														
14	13.1802														
15	6.0037														
16	2.6006														
17	4.2013														
18	7.4514														
19	4.3907														
20	4.4048														

Ensuite, allez dans le menu **Stat/Cartes de contrôle/Transformation de Box-Cox...**:



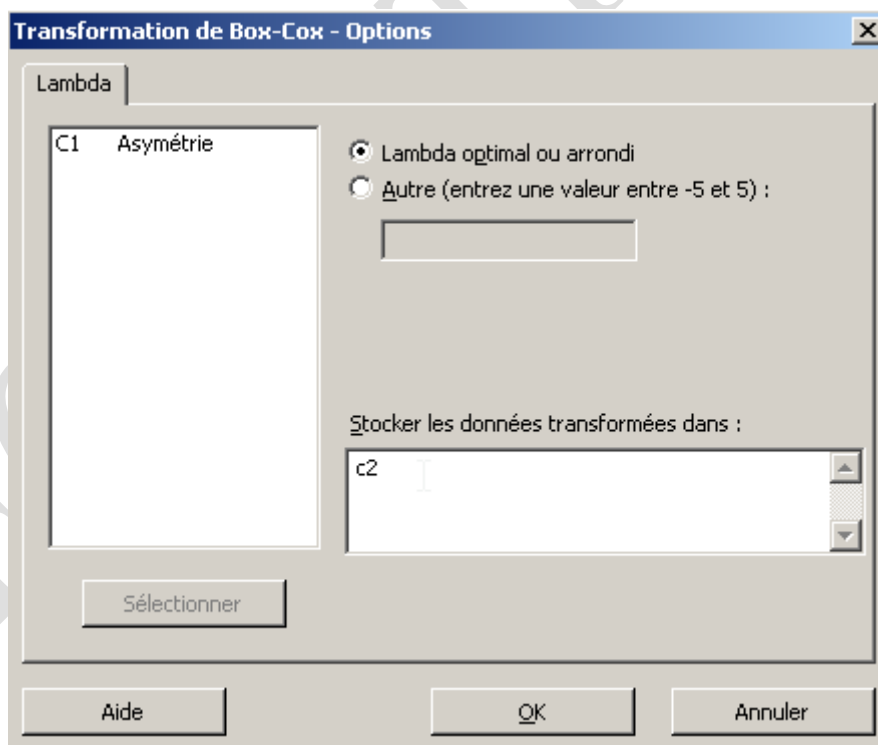
Nous mettons alors:



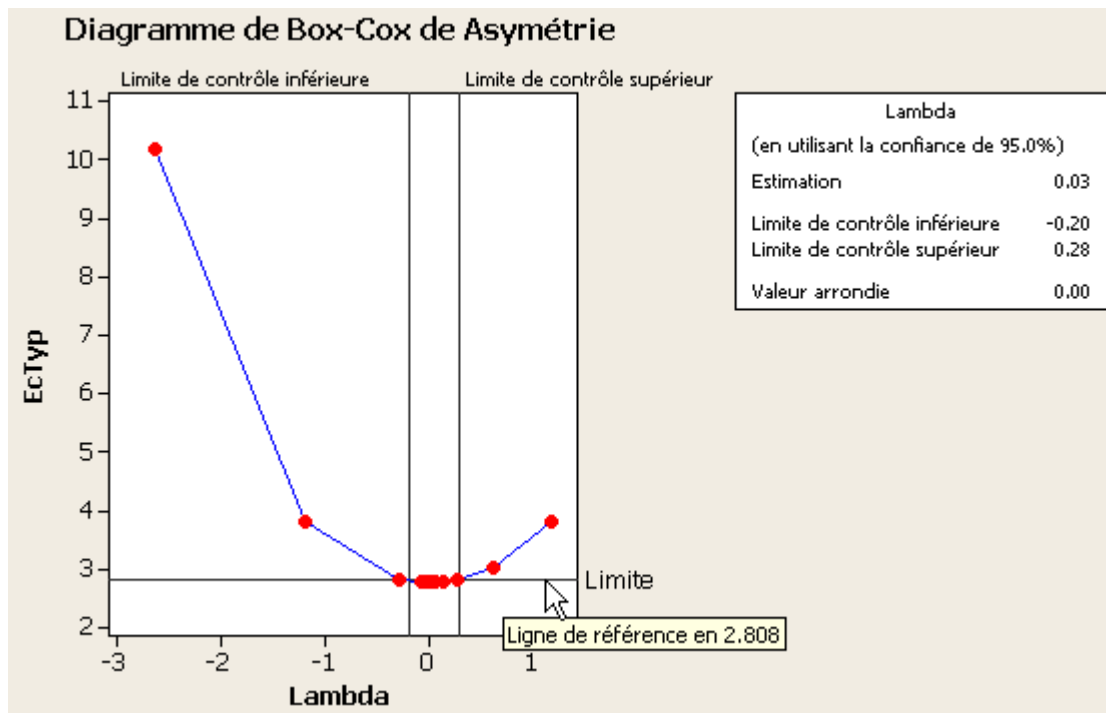


Si nous tapons 125 pour l'effectif du sous-groupe, c'est simplement pour pouvoir faciliter la comparaison avec Microsoft Excel et aussi parce que nous ne souhaitons pas faire usage de la transformation pour une carte de contrôle.

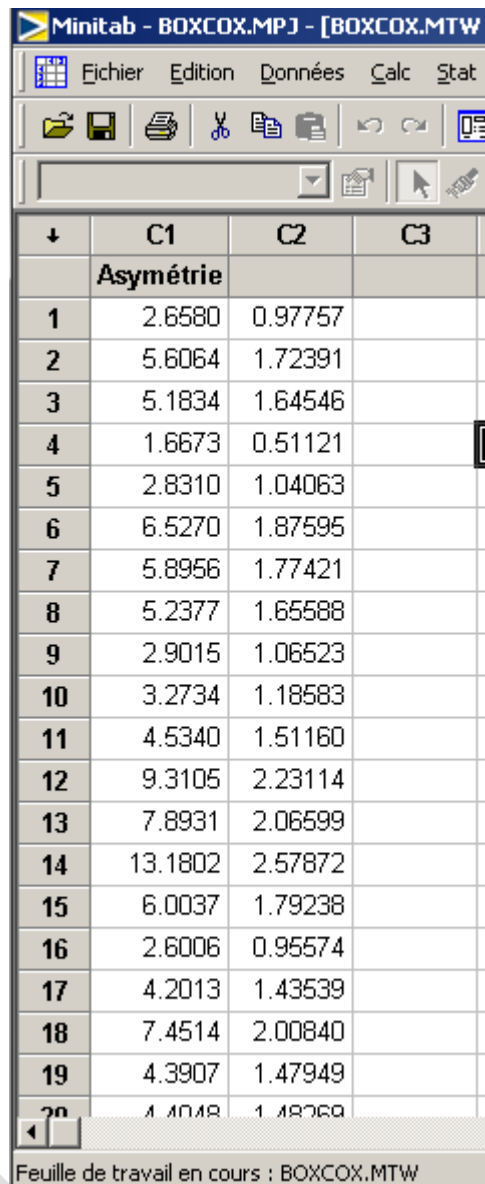
Nous cliquons ensuite le bouton **Options...**:



et en validant plusieurs fois par **OK**, nous avons d'abord un graphique:



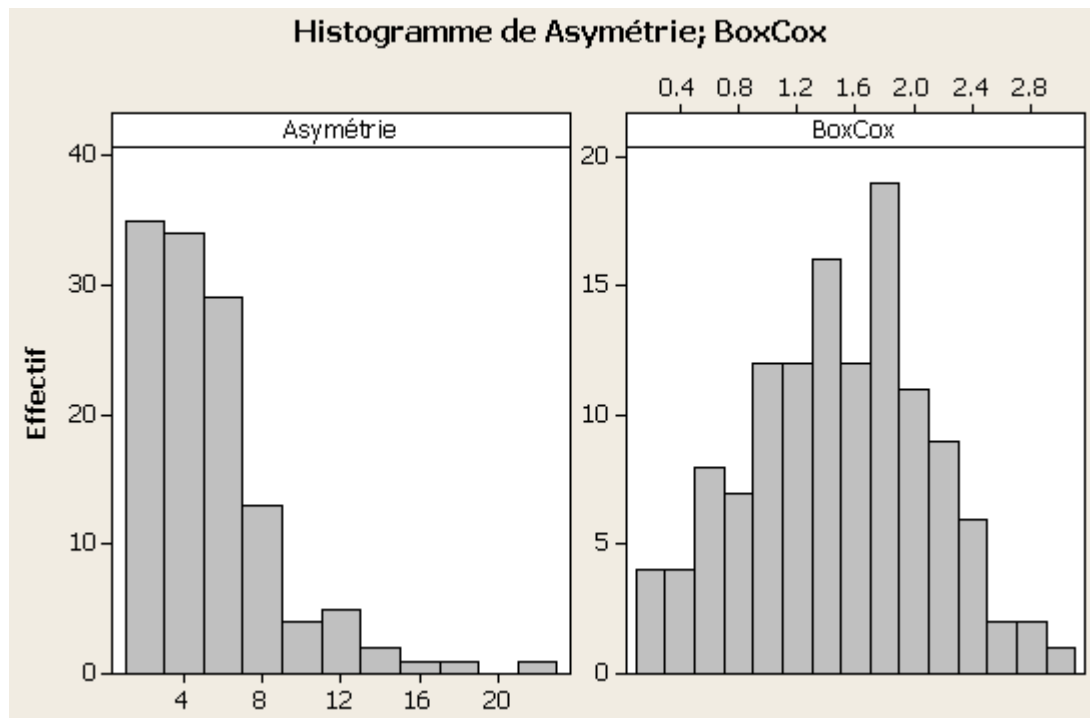
qui montre comment l'écart-type est minimisé en fonction de la valeur de lambda et également l'intervalle de confiance de ce dernier à 95%. Sur la feuille, nous avons:



	C1	C2	C3
	<b>Asymétrie</b>		
1	2.6580	0.97757	
2	5.6064	1.72391	
3	5.1834	1.64546	
4	1.6673	0.51121	
5	2.8310	1.04063	
6	6.5270	1.87595	
7	5.8956	1.77421	
8	5.2377	1.65588	
9	2.9015	1.06523	
10	3.2734	1.18583	
11	4.5340	1.51160	
12	9.3105	2.23114	
13	7.8931	2.06599	
14	13.1802	2.57872	
15	6.0037	1.79238	
16	2.6006	0.95574	
17	4.2013	1.43539	
18	7.4514	2.00840	
19	4.3907	1.47949	
20	1.4048	1.48269	

Feuille de travail en cours : BOXCOX.MTW

Si nous comparons les deux distributions (avant et après) avec les outils graphiques habituels (exercices effectués au début de ce support):



Maintenant, observons quelle valeur de lambda nous obtenons en utilisant le solveur de Microsoft Excel. Avec la version anglaise de Microsoft Excel, la préparation du problème donne:

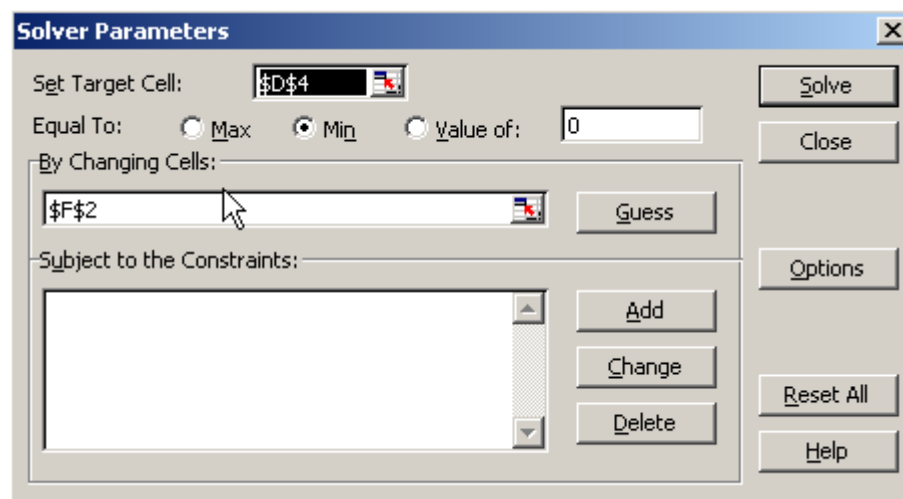
Microsoft Excel - BoxCox

File Edit View Insert Format Tools Data Window Help

Arial 10 B I U %

	A	B	C	D	E	F
1	Asymétrie	Transfo		Geo Mean		Lambda
2	2.658	$= (A2^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$		$= \text{GEOMEAN}(A2:A126)$		10
3	5.6064	$= (A3^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$		StDev		
4	5.1834	$= (A4^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$		$= \text{STDEV}(B:B)$		
5	1.6673	$= (A5^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				
6	2.831	$= (A6^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				
7	6.527	$= (A7^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				
8	5.8956	$= (A8^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				
9	5.2377	$= (A9^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				
10	2.9015	$= (A10^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				
11	3.2734	$= (A11^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				
12	4.534	$= (A12^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				
13	9.3105	$= (A13^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				
14	7.8931	$= (A14^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				
15	13.1802	$= (A15^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				
16	6.0037	$= (A16^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				
17	2.6006	$= (A17^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				
18	4.2013	$= (A18^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				
19	7.4514	$= (A19^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				
20	4.3907	$= (A20^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				
21	4.4048	$= (A21^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				
22	11.1304	$= (A22^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				
23	2.0317	$= (A23^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				
24	2.4576	$= (A24^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				
25	4.192	$= (A25^{\$F\$2-1}) / (\$F\$2 * \$D\$2^{\$F\$2-1})$				

Ensuite, nous paramétrons le solveur comme ci-dessous:



et nous validons par **Solve**, ce qui donne:

	A	B	C	D	E	F
1	Asymétrie	Transfo		Geo Mean		Lambda
2	2.658	4.181331696		4.4147226		0.031796
3	5.6064	7.462240184		StDev		
4	5.1834	7.113709236		2.7595179		
5	1.6673	2.17034024				
6	2.831	4.455526569				
7	6.527	8.140205299				
8	5.8956	7.686162334				
9	5.2377	7.159959278				
10	2.9015	4.56263772				
11	3.2734	5.089009634				
12	4.534	6.521007656				
13	9.3105	9.736910768				
14	7.8931	8.99225338				
15	13.1802	11.31690529				
16	6.0037	7.767140793				
17	2.6006	4.086525783				
18	4.2013	6.184681552				
19	7.4514	8.733519				
20	4.3907	6.379177126				
21	4.4048	6.393329821				
22	11.1304	10.54632328				
23	2.0317	3.019046666				
24	2.4576	3.841229469				
25	4.192	6.174914033				

Nous retrouvons donc presque les mêmes valeurs que dans Minitab® Statistical Software.

Par ailleurs, voici comment Minitab® Statistical Software calcule l'écart-type (capture d'écran de leur KB655):

If the subgroup size is greater than 1, Minitab uses the pooled standard deviation method.

$$SD = \frac{s_p}{c_4(d)} \text{ where } d = \sum (n_i - 1) + 1 \text{ degrees of freedom}$$

$$s_p = \sqrt{\frac{\sum_i \sum_j (w_{ij} - \bar{w}_i)^2}{\sum_i (n_i - 1)}}$$

If the subgroup size is equal to 1, Minitab uses the average moving range method.

$$SD = \frac{\overline{MR}}{d_2(r)} \text{ where } r = \text{the number of observation used in the moving range (} r=2 \text{ by default).}$$

$$MR_i = \text{the } i\text{th moving range} = \max[w_i, \dots, w_{i+r}] - \min[w_i, \dots, w_{i+r}], \text{ for } i=r, \dots, n$$

$$\overline{MR} = \frac{(MR_r + \dots + MR_n)}{(n - r + 1)}$$

Knowledgebase ID 655: <http://www.minitab.com/support/answers/answer.aspx?ID=655>

## 15.5. Exercice 149.: Transformations de Johnson (Yeo-Johnson)

Minitab® Statistical Software 15.1.1.0

Bon le but ici ne sera pas de débattre de la justesse de ces transformations (ou de leur origine scientifique) qui ont donc pour rappel de "normaliser"<sup>5</sup> des données. Ceci a aussi déjà été fait dans le cours théorique et de plus, il suffit de Googler un peu pour se faire sa propre opinion sur cette technique "d'ingénierie statistique".

Voici les transformations définies empiriquement par Johnson:

$$t: x \rightarrow t(x) = a + b \cdot g\left(\frac{x-c}{d}\right)$$

Johnson établit alors une classification des principales lois en trois systèmes associées chacun à une transformation  $g$  particulière et donc à une famille de transformations  $t$ :

$$g(z) = \log(z)$$

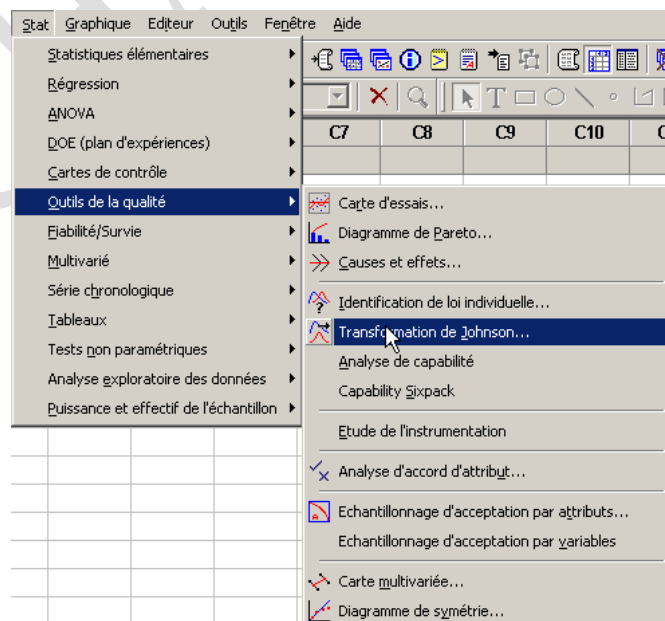
$$g(z) = \log\left(\frac{z}{1-z}\right)$$

$$g(z) = \operatorname{arcsinh}(z)$$

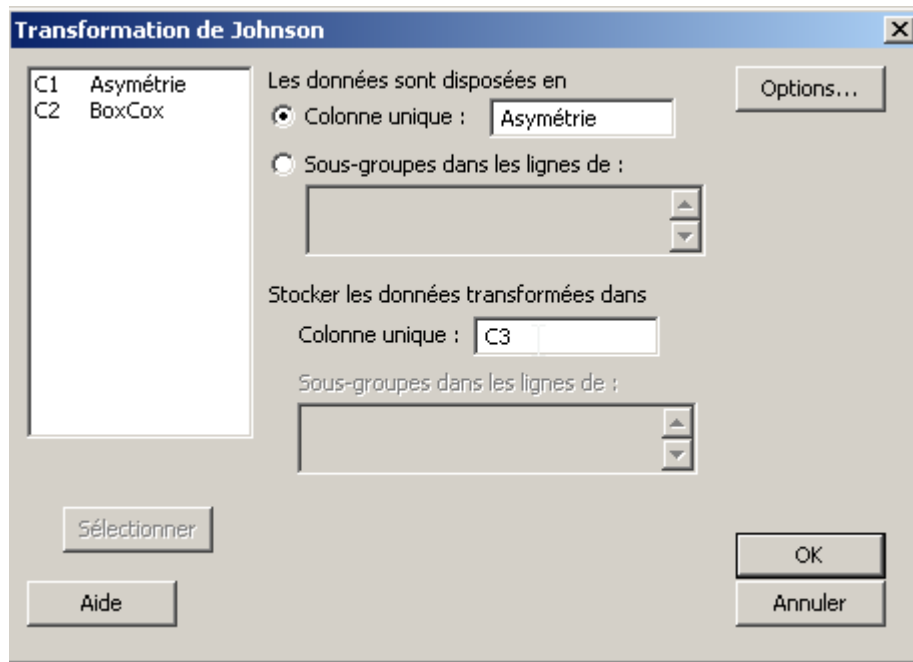
respectivement dénommées Sl, Sb et Su.

Nous allons pour cet exemple reprendre exactement le même fichier avec les mêmes données que pour l'exemple des transformations de Box-Cox.

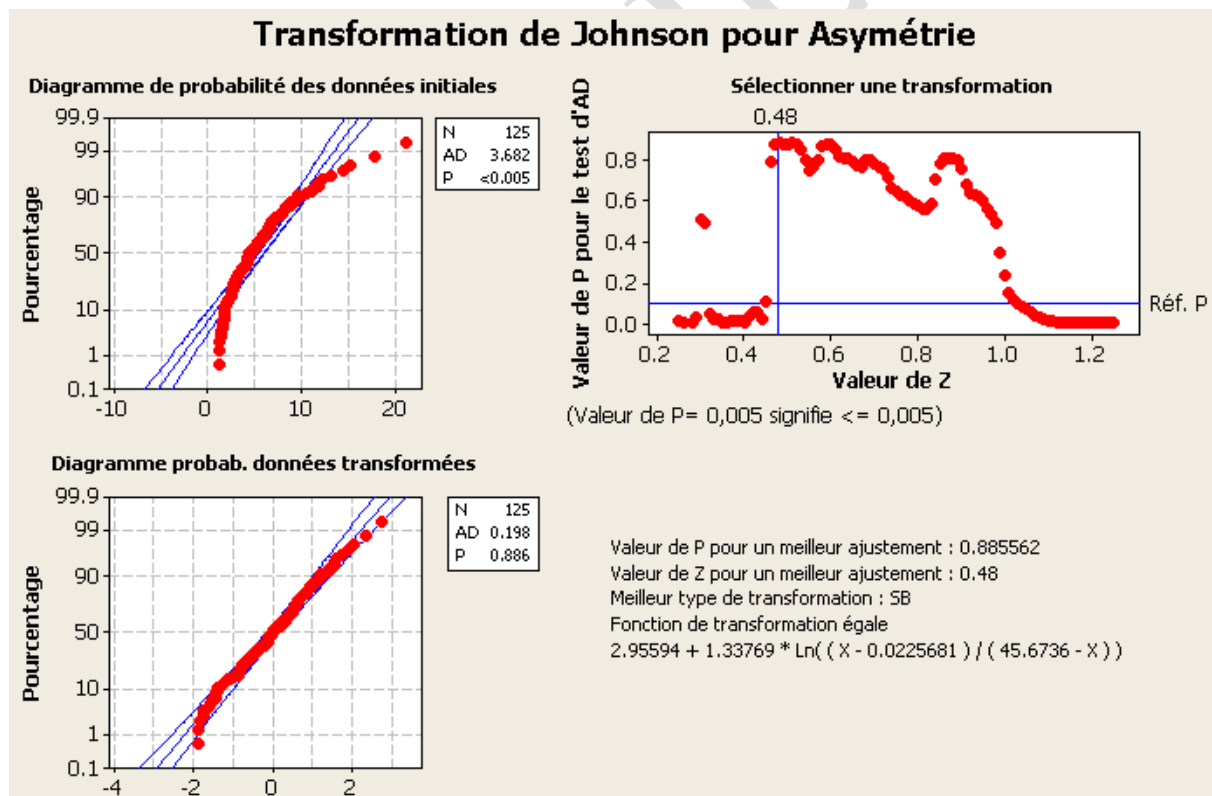
Nous allons dans le menu **Stat/Outils de la qualité/Transformation de Johnson...**:



<sup>5</sup> Dans le sens: faire que la distribution ressemble le plus possible à une loi Normale

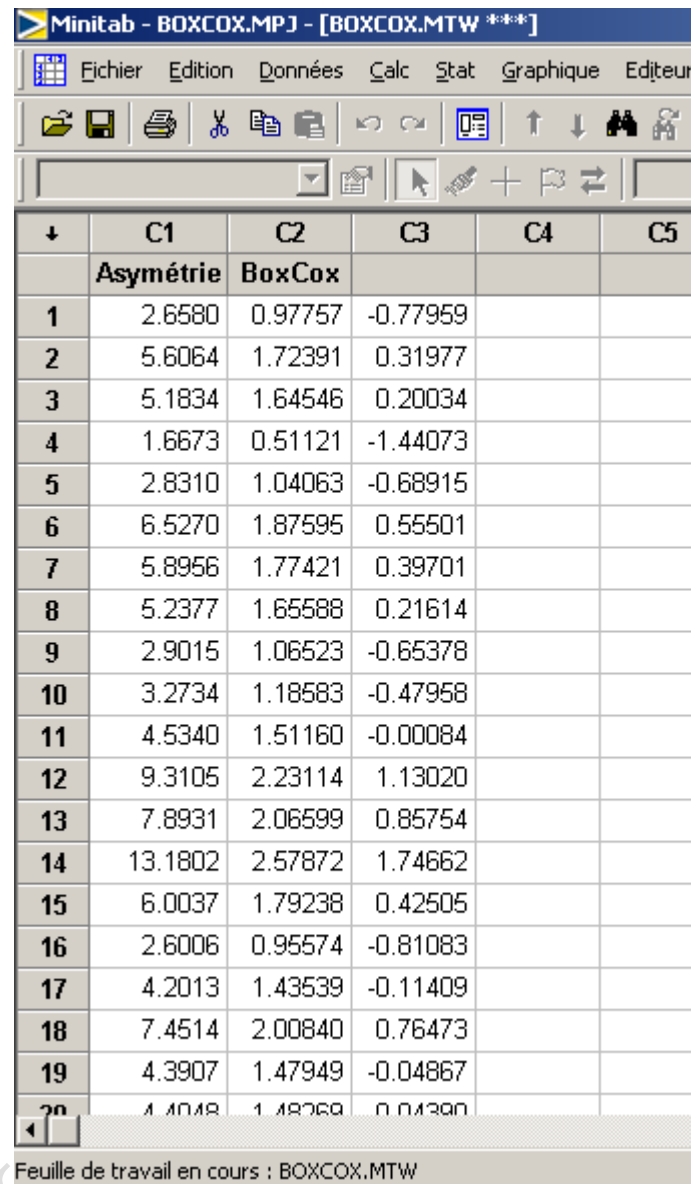


Nous validons par **OK** plusieurs fois et obtenons d'abord le graphique suivant (dommage que la transformation de Box-Cox ne fasse pas le même...):



ce qui donne comme tableau:





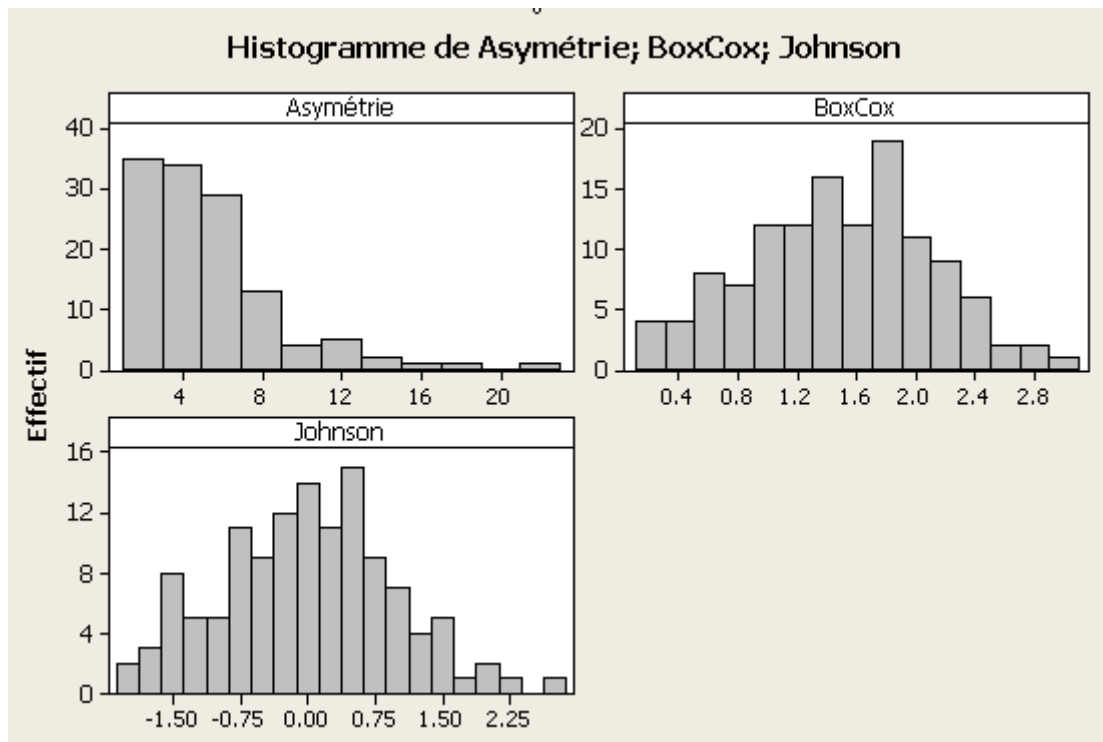
Minitab - BOXCOX.MPJ - [BOXCOX.MTW \*\*\*]

Fichier Edition Données Calc Stat Graphique Editeur

	C1	C2	C3	C4	C5
	<b>Asymétrie</b>	<b>BoxCox</b>			
1	2.6580	0.97757	-0.77959		
2	5.6064	1.72391	0.31977		
3	5.1834	1.64546	0.20034		
4	1.6673	0.51121	-1.44073		
5	2.8310	1.04063	-0.68915		
6	6.5270	1.87595	0.55501		
7	5.8956	1.77421	0.39701		
8	5.2377	1.65588	0.21614		
9	2.9015	1.06523	-0.65378		
10	3.2734	1.18583	-0.47958		
11	4.5340	1.51160	-0.00084		
12	9.3105	2.23114	1.13020		
13	7.8931	2.06599	0.85754		
14	13.1802	2.57872	1.74662		
15	6.0037	1.79238	0.42505		
16	2.6006	0.95574	-0.81083		
17	4.2013	1.43539	-0.11409		
18	7.4514	2.00840	0.76473		
19	4.3907	1.47949	-0.04867		
20	4.4048	1.48769	0.04390		

Feuille de travail en cours : BOXCOX.MTW

Si nous comparons maintenant les trois colonnes:



La transformation de Johnson est donc peut-être plus pertinente que la Box-Cox puisque les données peuvent aussi être négatives.

## 15.6. Exercice 150.: Analyse de capacité court terme et long temps et niveau de qualité

Minitab® Statistical Software 15.1.1.0

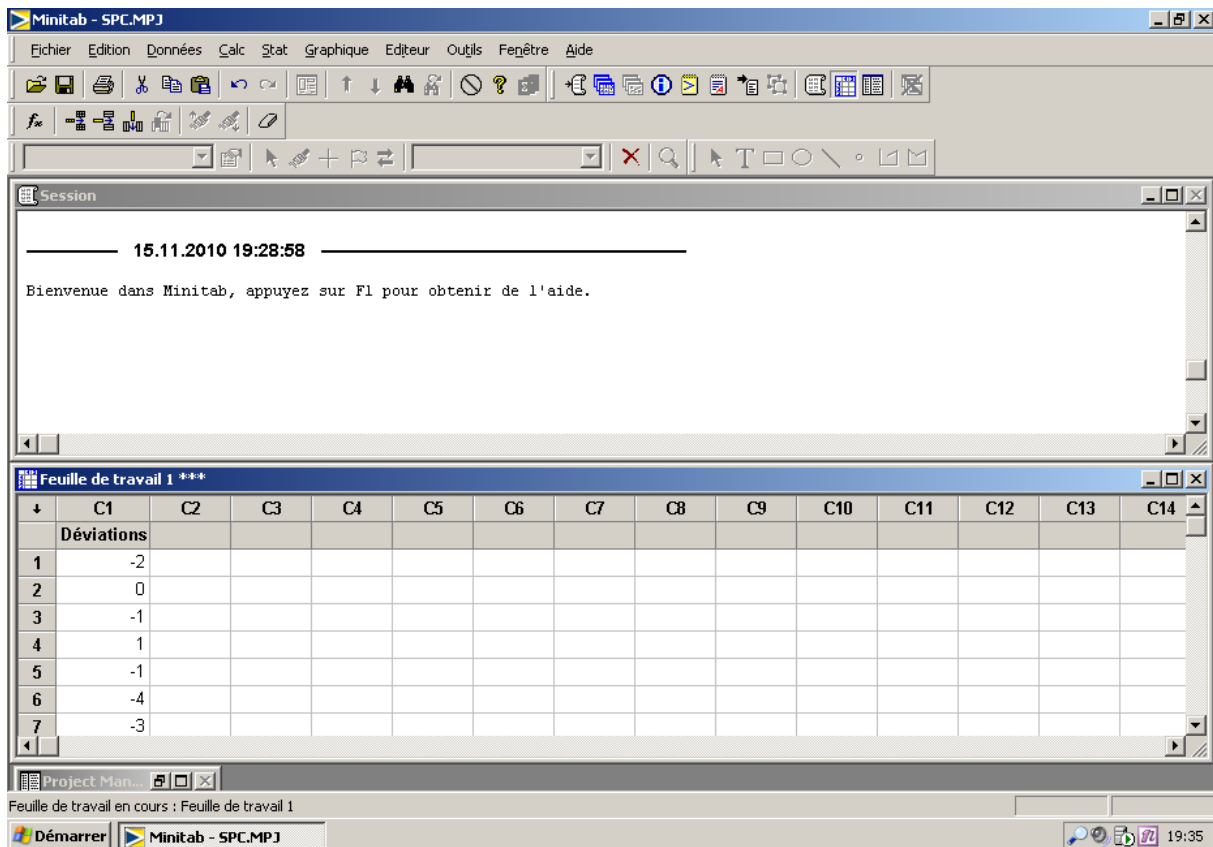
Considérons une nouvelle petite production de  $n = 50$  pièces par lot de 10 (afin d'ajuster en cours de production). La mesure de côtes de 5 pièces chaque heure pendant 10 heures avec une tolérance de  $10 \pm 0.07$  soit en termes de centièmes un étendue de:

$$E = 2 \cdot 7 = 14$$

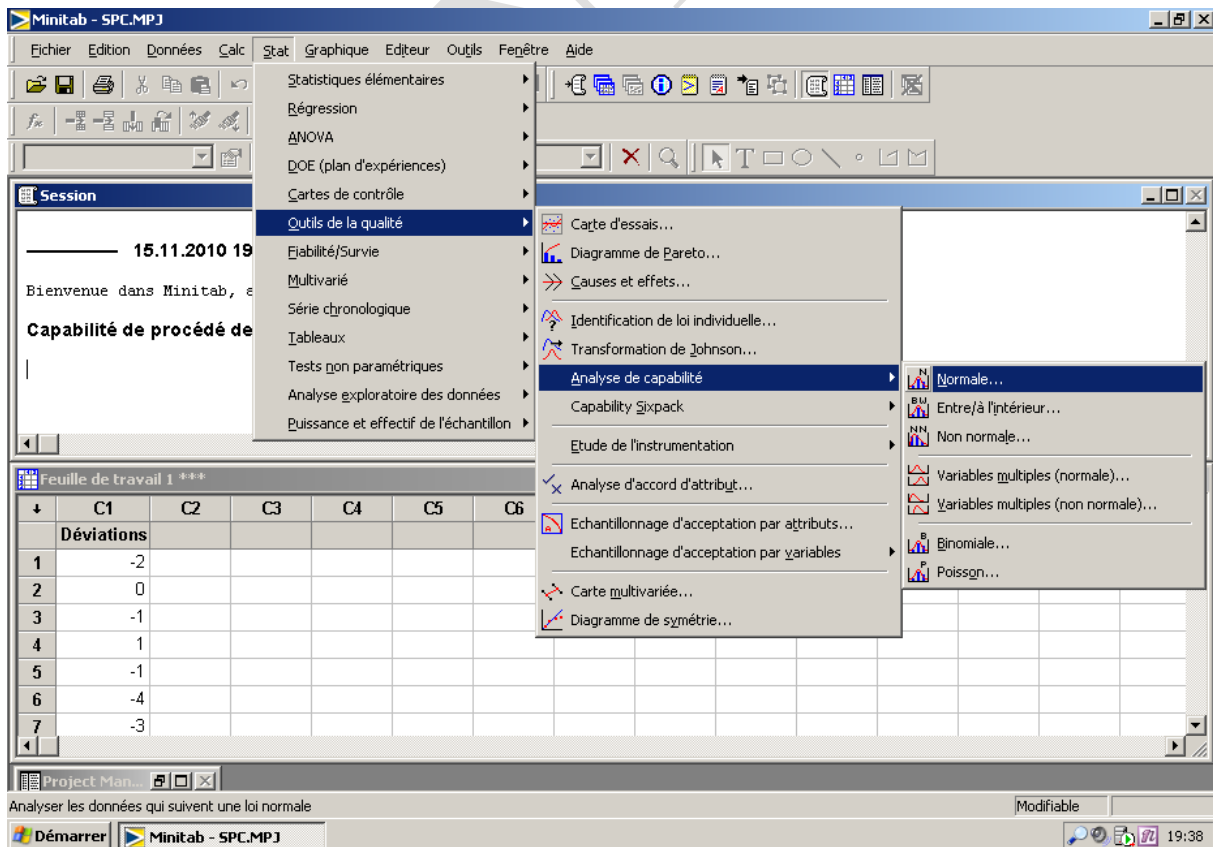
et une cible de  $T = 0$  (en termes d'écart). Nous avons les données suivantes:

	1	2	3	4	5	6	7	8	9	10
1	-2	-4	-1	0	4	0	3	0	1	-1
2	0	-3	0	-2	1	-2	0	1	-1	2
3	-1	0	-3	-1	0	0	-1	-1	3	1
4	1	1	-2	2	2	0	1	0	4	0
5	-1	-1	-3	0	0	3	3	2	1	0
$\mu$	-0.6	-1.4	-1.8	-0.2	1.4	0.2	1.2	0.4	1.6	0.5
$\sigma$	1.14	2.07	1.30	1.48	1.67	1.79	1.79	1.14	1.95	1.14

Dans le cours de maîtrise statistique des procédés, nous avons calculé à la main tous les indicateurs de capacité long terme et court terme. Nous souhaiterions maintenant automatiser la procédure et retrouver les mêmes valeurs avec Minitab® Statistical Software. Pour cela nous ouvrons un fichier *SPC.mpj* qui contient le tableau ci-dessus en une seule colonne:



Pour faire l'analyse de capabilité nous allons maintenant dans **Outils de la qualité/Analyse de capabilité/Normale...**:



Et nous paramétrons conformément aux prélèvements:

**Analyse de capabilité (loi normale)**

Les données sont disposées en

**Colonne unique :**

**Effectif du sous-groupe :**   
(utiliser une constante ou une colonne d'ID)

**Sous-groupes dans les lignes de :**

**Spécification inférieure :**   Limite

**Spécification supérieure :**   Limite

**Moyenne historique :**  (facultatif)

**Ecart type historique :**  (facultatif)

Box-Cox...  
Estimation...  
Options...  
Stockage...

Sélectionner

Aide

OK  
Annuler

nous cliquons sur le bouton **Options...**:

**Analyse de capabilité (loi normale) - Options**

Cible (ajoute Cpm au tableau) :

Utiliser la tolérance de  $K \cdot \sigma$  pour les statistiques de capabilité  $K =$

**Exécution de l'analyse**

Analyse à l'intérieur des sous-groupes

Analyse globale

**Affichage**

Parties par million

Pourcentages

Statistiques de capabilité (Cp, Pp)

Z de référence (niveau de sigma)

Inclure les intervalles de confiance

Niveau de confiance :

Intervalles de confiance :

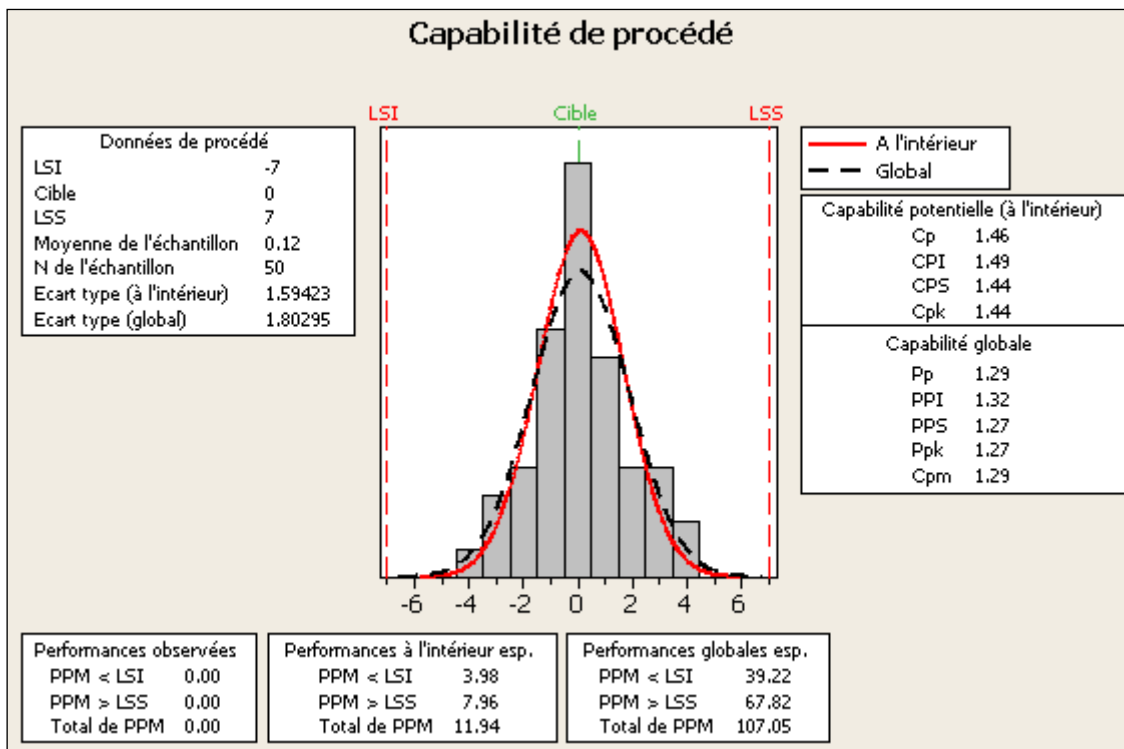
Titre :

Aide

OK

Annuler

et validons deux fois par **OK** et nous voyons que nous obtenons les chiffres obtenus dans le cours SPC:



Nous retrouvons donc dans ce graphique tous les notions vues et démontrées mathématiquement en détail dans le cours théorique.

## 15.7. Exercice 151.: Intervalle de confiance de la capacité

Minitab® Statistical Software 15.1.1.0

Le but ici va être de vérifier l'intervalle de confiance de la capacité sur les données de l'exemple antérieur en vérifiant que l'on retombe bien sur celui calculé dans le cours théorique suite à la démonstration mathématique:

A	B
<b>Ecart</b>	<b>Ecart</b>
45%	15%
40%	15%
35%	15%
35%	10%
30%	10%
30%	10%
25%	5%
25%	5%
25%	0%
20%	-5%
20%	-5%
20%	-10%
20%	

Pour faire l'analyse de capacité nous allons maintenant dans **Outils de la qualité/Analyse de capacité/Normale...**:

The screenshot shows the Minitab software interface. The 'Outils de la qualité' (Quality Tools) menu is open, and the path 'Analyse de capacité' (Capability Analysis) > 'Normale...' (Normal...) is selected. The 'Feuille de travail 1' (Worksheet 1) is visible in the background, containing the data from the table above. The status bar at the bottom indicates 'Analyser les données qui suivent une loi normale' (Analyze data that follow a normal distribution).

Nous paramétrons l'analyse de capabilité:

**Analyse de capabilité (loi normale)**

Les données sont disposées en

Colonne unique :

Effectif du sous-groupe :   
(utiliser une constante ou une colonne d'ID)

Sous-groupes dans les lignes de :

Spécification inférieure :   Limite

Spécification supérieure :   Limite

Moyenne historique :  (facultatif)

Ecart type historique :  (facultatif)

Box-Cox...  
Estimation...  
Options...  
Stockage...

Sélectionner

Aide

OK  
Annuler

nous cliquons sur le bouton **Options...** et complétons la boîte de dialogue surtout la partie en rouge:

**Analyse de capabilité (loi normale) - Options**

Cible (ajoute Cpm au tableau) :

Utiliser la tolérance de  $K \cdot \sigma$  pour les statistiques de capabilité  $K =$

Exécution de l'analyse

Analyse à l'intérieur des sous-groupes

Analyse globale

Affichage

Parties par million

Pourcentages

Statistiques de capabilité (Cp, Pp)

Z de référence (niveau de sigma)

Inclure les intervalles de confiance

Niveau de confiance :

Intervalles de confiance :

Titre :

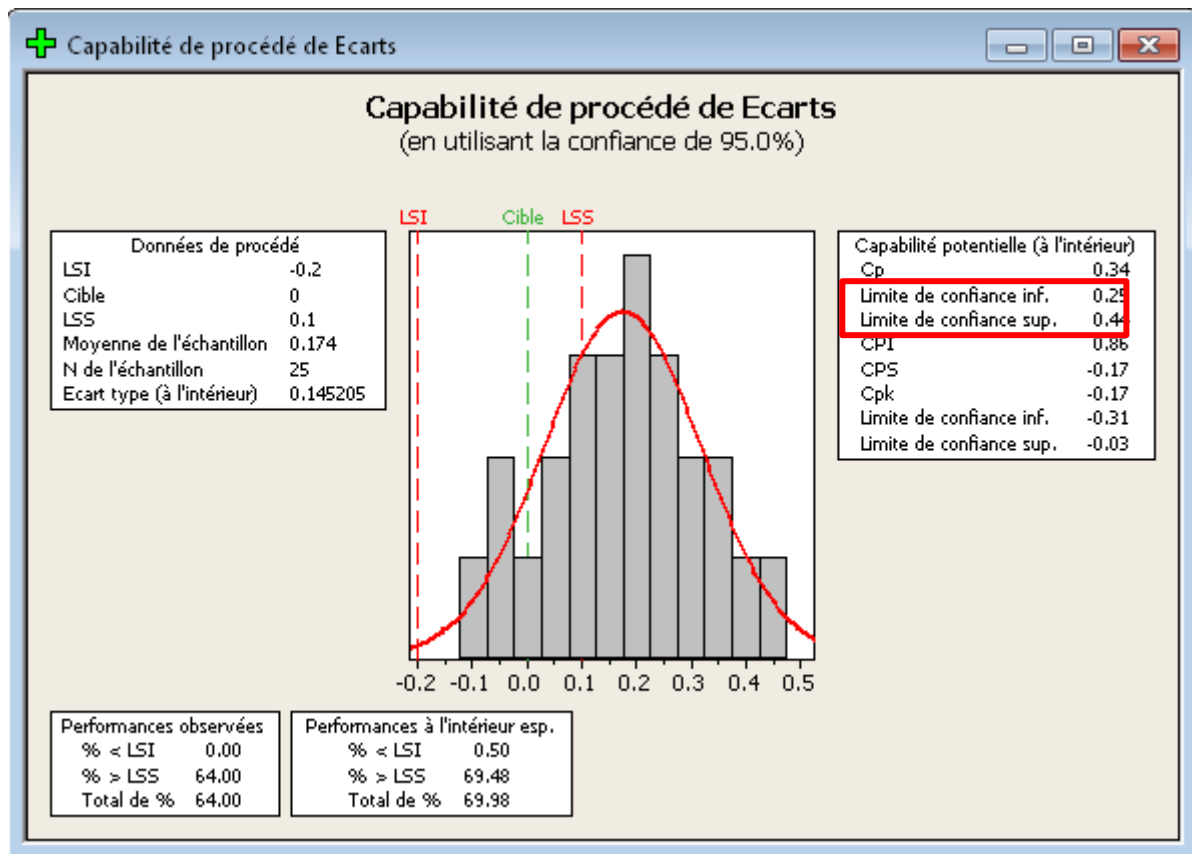
Aide

OK

Annuler

Nous avons alors en validant par **OK**:





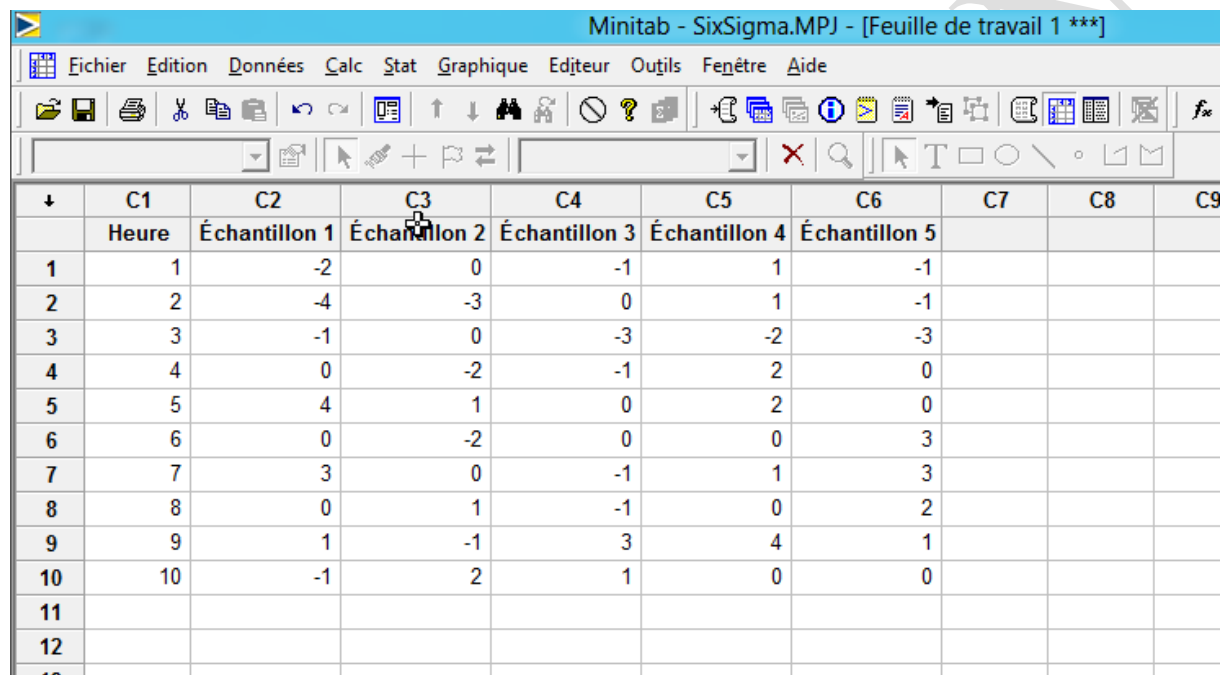
où nous avons mis en rouge ce qui nous intéresse! Nous retrouvons bien les valeurs calculés dans le cours théorique!

## 15.8. Exercice 152.: Analyse SixPack

Minitab® Statistical Software 15.1.1.0

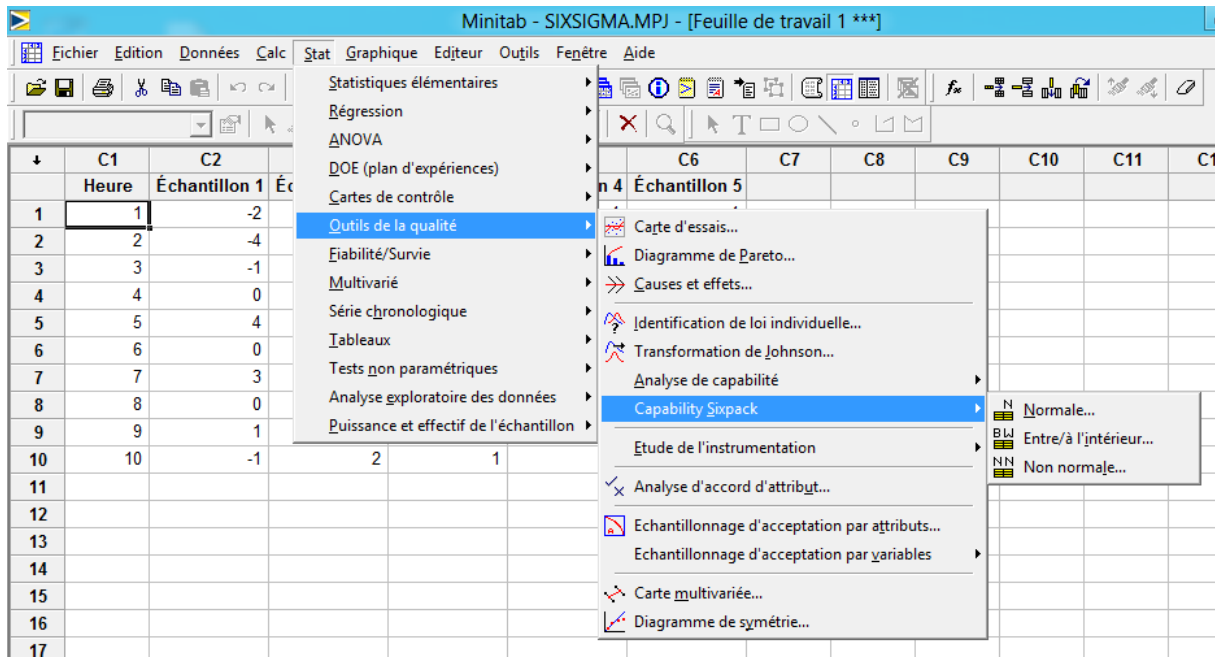
Maintenant que nous savons faire une analyse de capabilité et des cartes de contrôles avec Minitab et que nous savons aussi en faire la démonstration mathématique et le calcul manuel comme vu dans le cours théorique, nous pouvons voir dans Minitab un outil fort pratique qui combine les deux concepts et là aussi vérifier que nous retrouvons tout ce qui a été calculé dans le cours théorique.

Nous partons du même jeu de données que dans le cours théorique:

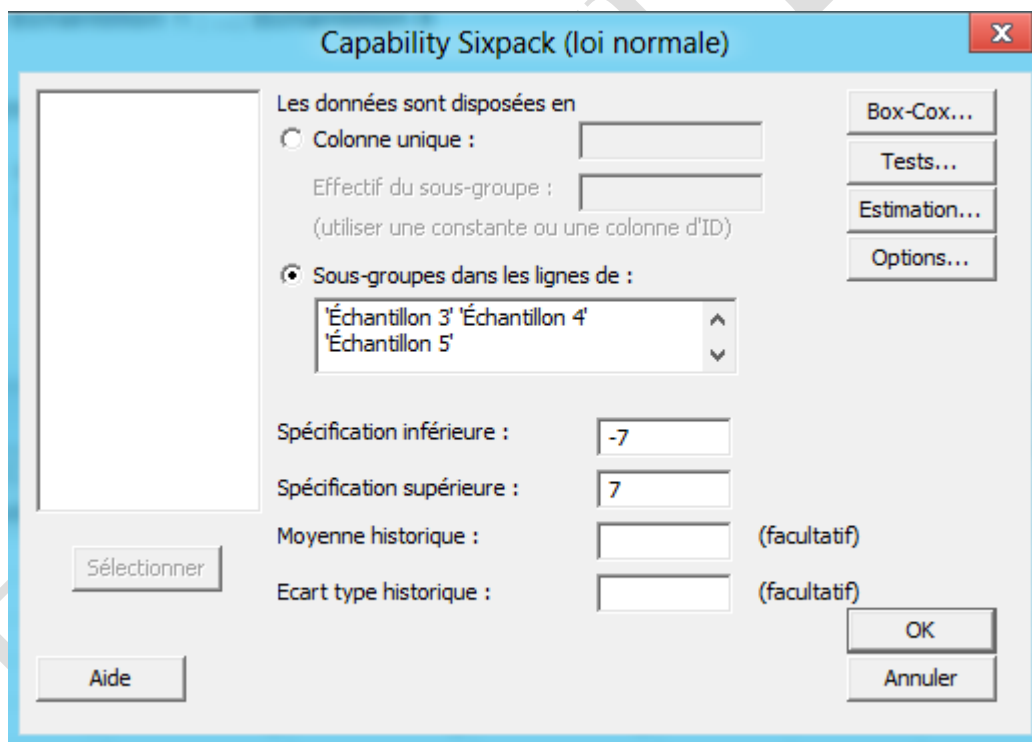


	C1	C2	C3	C4	C5	C6	C7	C8	C9
	Heure	Échantillon 1	Échantillon 2	Échantillon 3	Échantillon 4	Échantillon 5			
1	1	-2	0	-1	1	-1			
2	2	-4	-3	0	1	-1			
3	3	-1	0	-3	-2	-3			
4	4	0	-2	-1	2	0			
5	5	4	1	0	2	0			
6	6	0	-2	0	0	3			
7	7	3	0	-1	1	3			
8	8	0	1	-1	0	2			
9	9	1	-1	3	4	1			
10	10	-1	2	1	0	0			
11									
12									

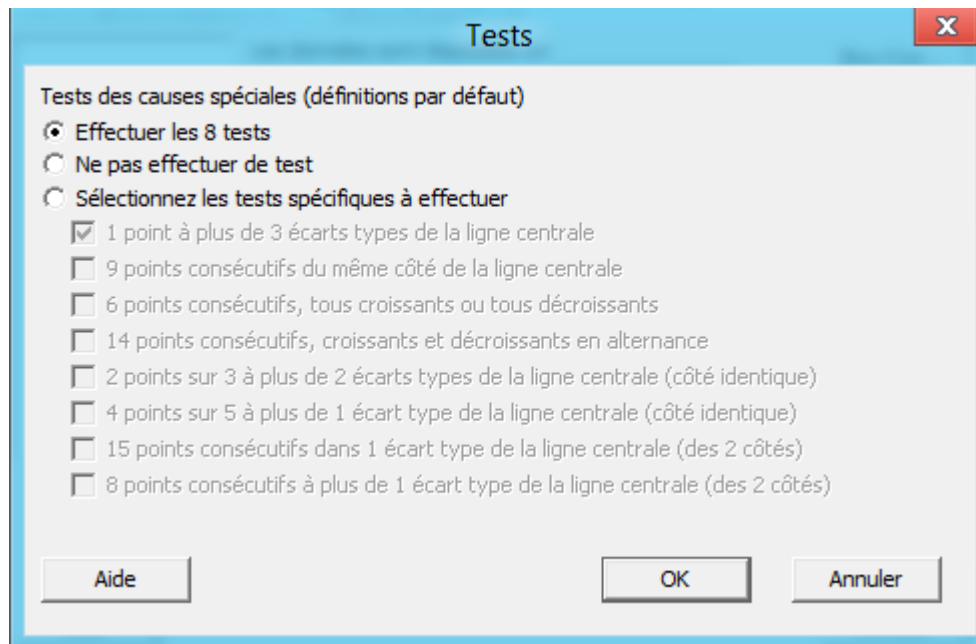
Ensuite, nous allons dans la menu **Stat/Outils de la qualité/Capability Sixpack/Normale...**:



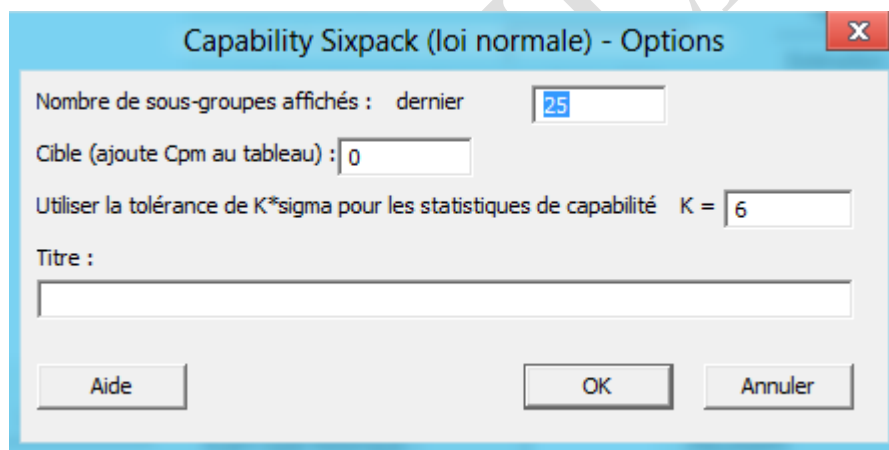
Et nous mettons alors:



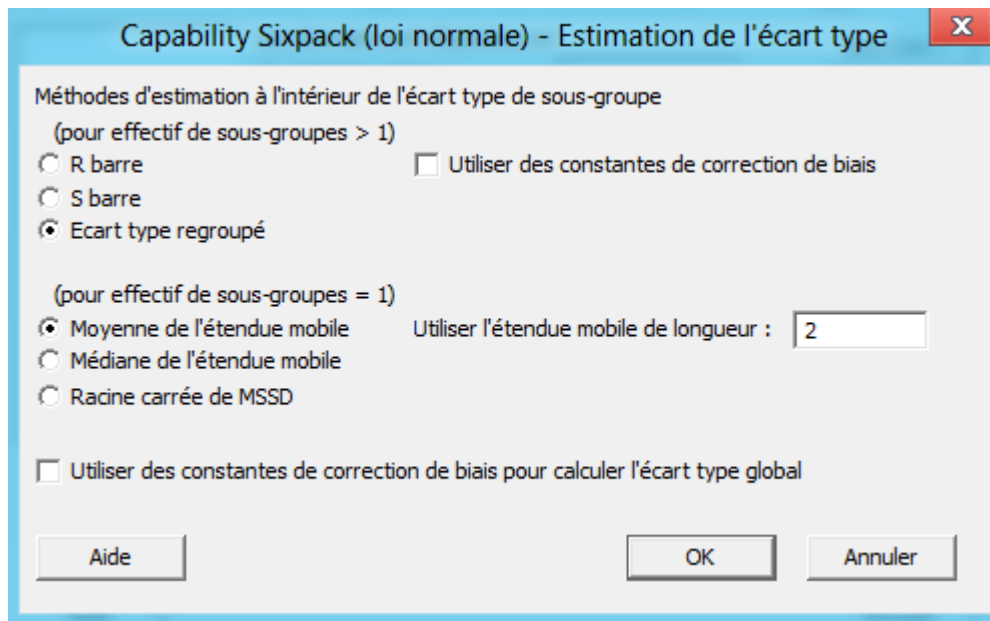
Dans le bouton **Tests...** nous prenons tous les tests de la WECO:



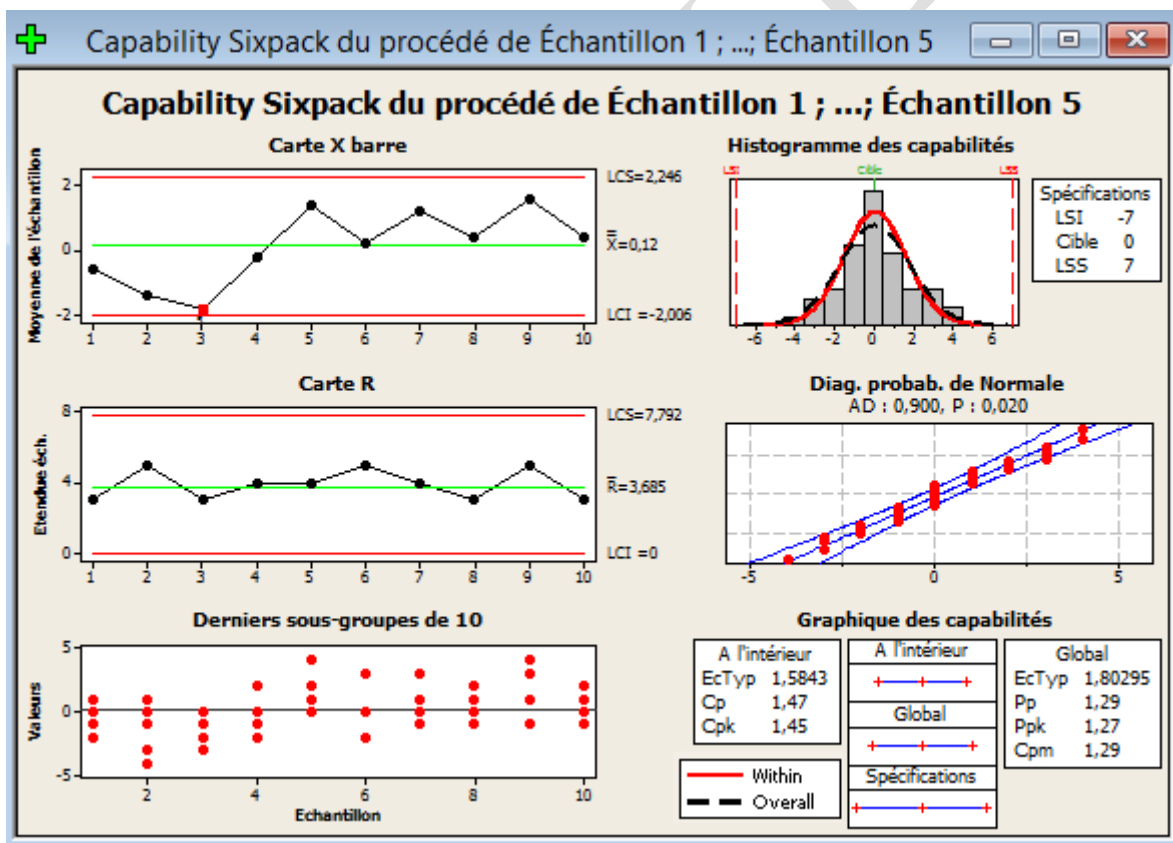
Dans le bouton **Options...** nous prenons (toujours pour être conforme aux calculs faits à la main dans le cours théorique):



Comme les échantillons sont petit (inférieur à 10), nous devrions prendre la carte R barre mais si nous voulons retrouver les calculs faits dans le cours théorique, il nous faudra laisser l'option **Ecart Type regroupé** active dans le bouton **Estimation...** (car dans Minitab les indices de Capacité et non pas seulement la carte de contrôle dépendent de la manière dont l'écart-type est calculé):



Nous validons tout ce beau monde par OK pour obtenir:



avec dans la fenêtre de session:

```
TEST 5. 2 points sur 3 supérieurs à 2 écarts types à partir de la ligne
centrale (sur un côté de LC).
Echec du test aux points : 3
```

Nous remarquons que nous retrouvons tous les calculs faits à la main au deuxième chiffre après la virgule près. Nous voyons aussi que le test d'Anderson-Darling nous amène en toute rigueur à rejeter l'hypothèse de Normalité mais comme c'est pour l'exemple...

ÉCHANTILLON

## 15.9. Exercice 153.: Étude de l'instrumentation de type I

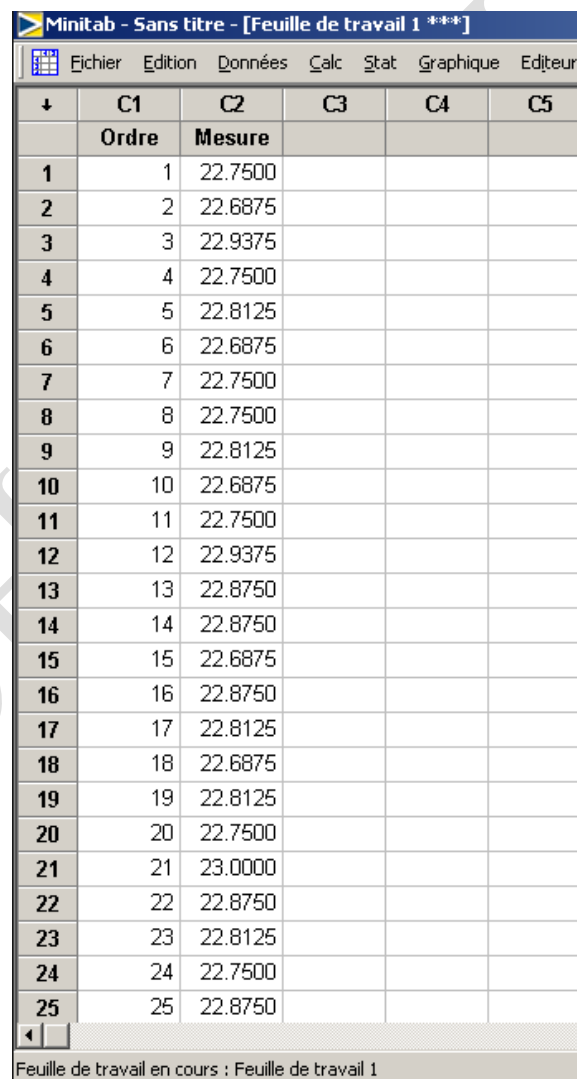
Minitab® Statistical Software 15.1.1.0

Bon nous n'allons pas nous étendre sur ce sujet qui serait a priori d'origine plus empirique qu'autre chose et donc certains paramètres sont plus que discutables... donc il n'y a rien à démontrer mathématiquement mais seulement à détailler la manière dont les calculs sont faits.

### Attention!

N'oubliez juste pas que ce type d'étude devrait être fait avant les études R&R (on vérifie la précision de l'instrument de mesure avant de faire des tests R&R logiquement...!).

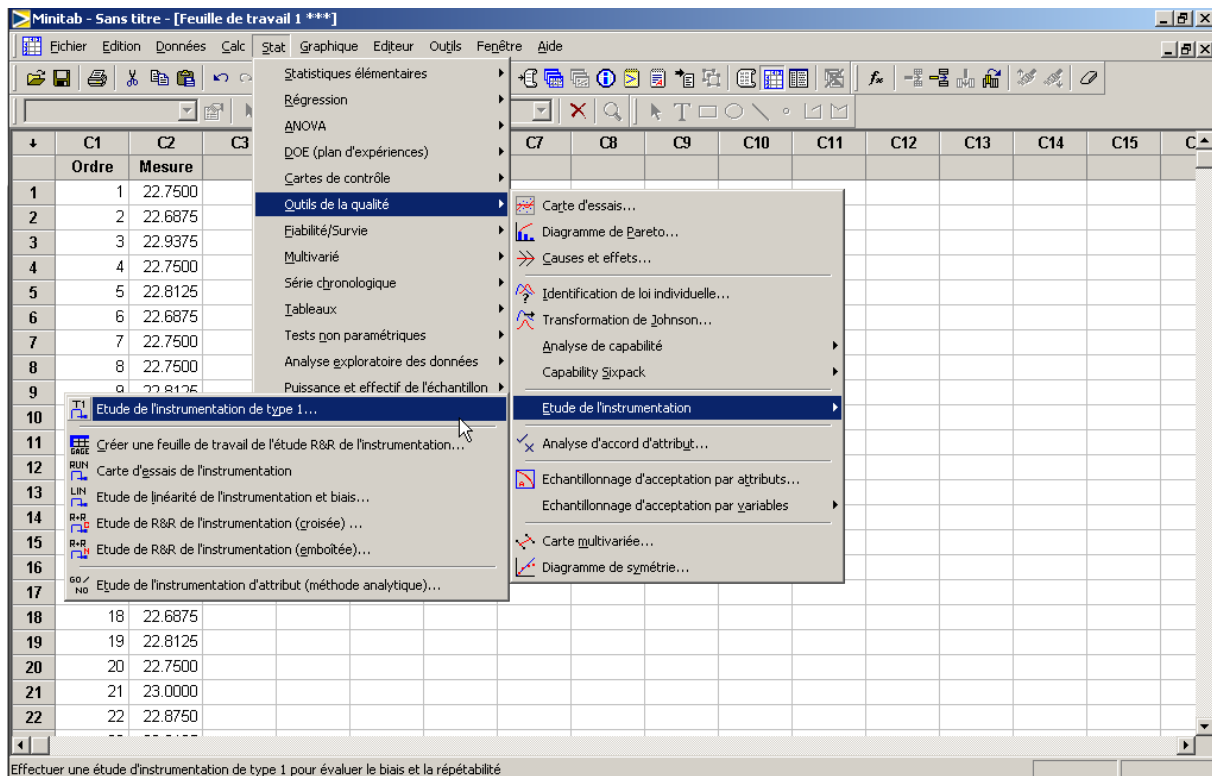
Considérons donc le jeu de données suivant consistant dans les mesures répétées d'un élément unique:



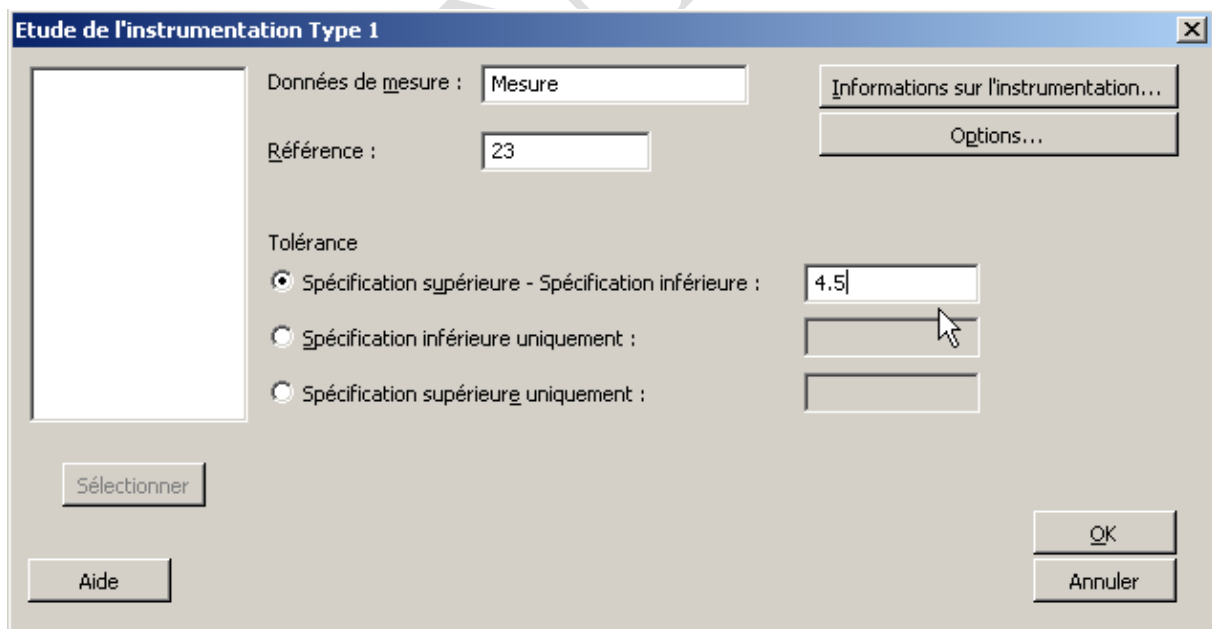
	C1	C2	C3	C4	C5
	Ordre	Mesure			
1	1	22.7500			
2	2	22.6875			
3	3	22.9375			
4	4	22.7500			
5	5	22.8125			
6	6	22.6875			
7	7	22.7500			
8	8	22.7500			
9	9	22.8125			
10	10	22.6875			
11	11	22.7500			
12	12	22.9375			
13	13	22.8750			
14	14	22.8750			
15	15	22.6875			
16	16	22.8750			
17	17	22.8125			
18	18	22.6875			
19	19	22.8125			
20	20	22.7500			
21	21	23.0000			
22	22	22.8750			
23	23	22.8125			
24	24	22.7500			
25	25	22.8750			

Ces mesures proviennent d'un appareil d'acquisition dont nous souhaiterions déterminer la capacité.

Pour cela, nous allons dans le menu **Stat/Outils de la qualité/Étude de l'instrumentation/Étude de l'instrumentation de type I...**:

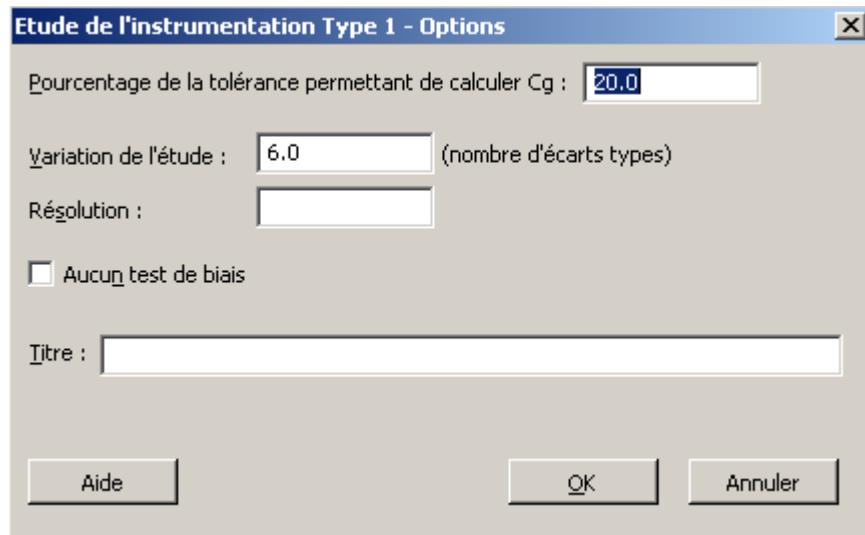


Dans la boîte de dialogue qui apparaît, nous prenons:

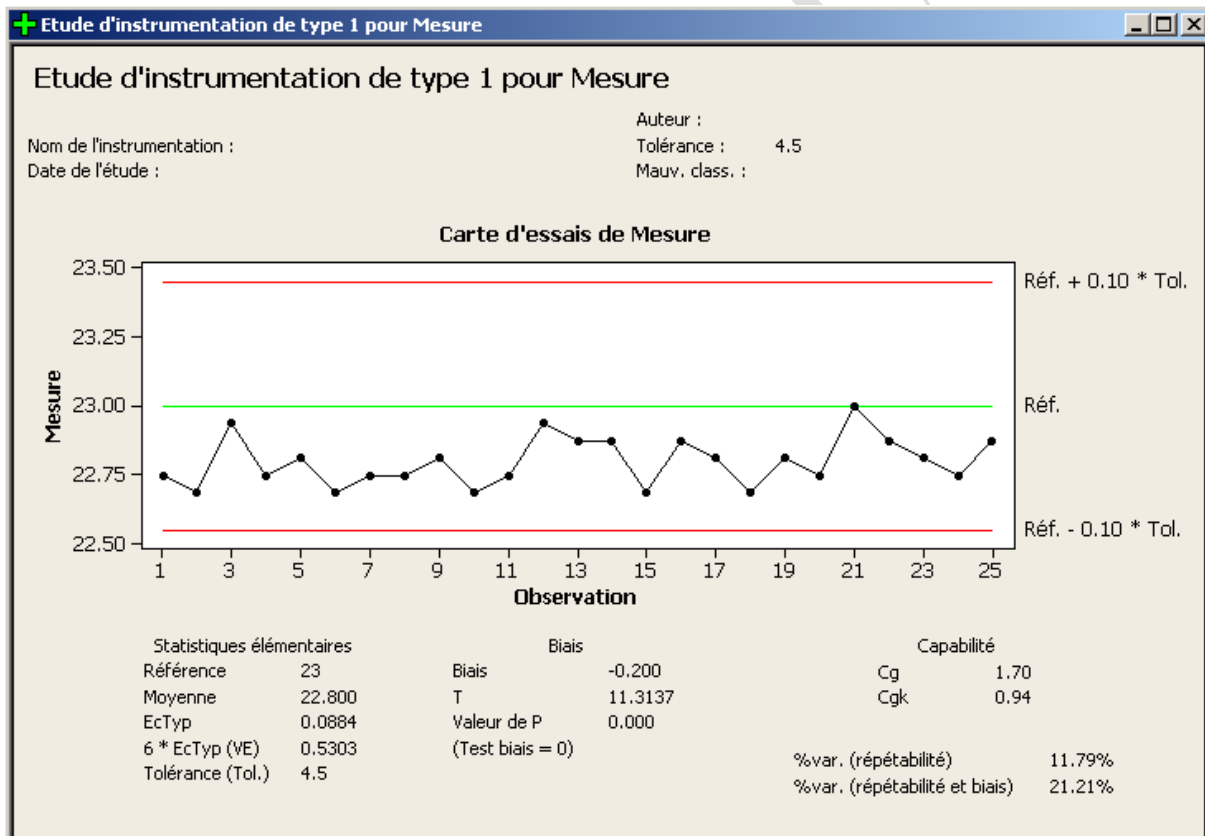


Dans le bouton **Options...** nous laissons les valeurs par défaut:





Nous validons le tout par **OK** pour obtenir le graphique suivante (il n'y a rien dans la fenêtre de sessions):



Vérifions tout cela avec le fichier Microsoft Excel ci-dessous:

	A	B
1	Ordre	Mesure
2	1	22.75
3	2	22.6875
4	3	22.9375
5	4	22.75
6	5	22.8125
7	6	22.6875
8	7	22.75
9	8	22.75
10	9	22.8125
11	10	22.6875
12	11	22.75
13	12	22.9375
14	13	22.875
15	14	22.875
16	15	22.6875
17	16	22.875
18	17	22.8125
19	18	22.6875
20	19	22.8125
21	20	22.75
22	21	23
23	22	22.875
24	23	22.8125
25	24	22.75
26	25	22.875

Le premier groupe d'informations est trivial à recalculer et à comprendre dans le présent contexte:

Statistiques élémentaires	
Référence	23
Moyenne	22.800
EcTyp	0.0884
6 * EcTyp (VE)	0.5303
Tolérance (Tol.)	4.5

Cela donne:

	D	E
1	Cible:	23
2	Moyenne:	=MOYENNE(B2:B26)
3	Écart-type:	=ECARTYPE(B2:B26)
4	6*StdDev:	=6*E3
5	Tolérance choisie:	4.5

Soit:

	D	E
1	Cible:	23
2	Moyenne:	22.800
3	Écart-type:	0.0884
4	6*StdDev:	0.5303
5	Tolérance choisie:	4.5

Maintenant poursuivons avec la partie un peu plus technique (mais niveau première année de licence à tout casser et encore...):

Biais	
Biais	-0.200
T	11.3137
Valeur de P	0.000
(Test biais = 0)	

Cela donne:

	D	E
1	Cible:	23
2	Moyenne:	=MOYENNE(B2:B26)
3	Écart-type:	=ECARTYPE(B2:B26)
4	6*StdDev:	=6*E3
5	Tolérance choisie:	4.5
6		
7	Biais	=E1-E2
8	n:	=NB(B2:B26)
9	T:	=E7/(E3/RACINE(E8))
10	Pvalue:	=LOI.STUDENT(ABS(E9);E8-1;2)

Soit (bon pour le biais le signe n'est pas le même mais il suffit d'inverser la soustraction):

	D	E
1	Cible:	23
2	Moyenne:	22.800
3	Écart-type:	0.0884
4	6*StdDev:	0.5303
5	Tolérance choisie:	4.5
6		
7	Biais	0.200
8	n:	25
9	T:	11.3137
10	Pvalue:	0.000

Donc là aussi rien de spécial. On fait simplement un test unilatéral de Student de l'erreur-standard estimée comme nous l'avons déjà traité en longueur dans le cours théorique. La question étant plutôt... Est-ce que cela a un sens de faire un tel calcul dans le cas présent? Et puis en réalité le biais est négatif donc nous sommes toujours en-dessous de la valeur du gabarit.

Réponse a priori: Oui pourquoi pas mais encore faudrait-il vérifier la normalité des données comme à l'habitude. Après dans notre cas, l'interprétation est que la déviation à la moyenne est très significative (donc la procédure de mesure n'est pas bonne).

Maintenant passons à l'autre partie:

Capabilité	
Cg	1.70
Cgk	0.94

Ce qui donne:

	D	E
12	K:	20
13	Cg:	$=((E12/100*E5)/E4)$
14	Cgk:	$=(0.1*E5-ABS(E7))/(3*E3)$

Soit:

	D	E
1	Cible:	23
2	Moyenne:	22.800
3	Écart-type:	0.0884
4	6*StdDev:	0.5303
5	Tolérance choisie:	4.5
6		
7	Biais	-0.200
8	n:	25
9	T:	11.3137
10	Pvalue:	0.000
11		
12	K:	20
13	Cg:	1.70
14	Cgk:	0.94

Bon que dire sur  $C_g$  tel que défini dans Minitab (et dans les industries qui l'utilise). Ici la capacité de gabarit est acceptable (car au-delà du 1.33 défini empiriquement par l'industrie concernée). Objectivement, en faisant abstraction de cet indicateur empirique, nous pouvons effectivement grossièrement dire que la capacité de gabarit est acceptable vu la petitesse de l'écart-type.

Pour le  $C_{gk}$  dont la relation est aussi a priori complètement empirique on se passera de commentaires...

Pour la dernière partie:

%var. (répétabilité)	11.79%
%var. (répétabilité et biais)	21.21%

Les valeurs proviennent de:

	D	E
16	%Var(Repeatability):	$=20/E13$
17	%Var(R&B):	$=E12/E14$

Ce qui donne:

	D	E
1	Cible:	23
2	Moyenne:	22.800
3	Écart-type:	0.0884
4	6*StdDev:	0.5303
5	Tolérance choisie:	4.5
6		
7	Biais	-0.200
8	n:	25
9	T:	11.3137
10	Pvalue:	0.000
11		
12	K:	20
13	Cg:	1.70
14	Cgk:	0.94
15		
16	%Var(Repeatability):	11.79
17	%Var(R&B):	21.21

Bref là aussi ce type d'indicateurs... sont plus que discutables.

Bon il faut savoir que d'un logiciel à l'autre (j'en ai testé trois différents: Minitab, R et QIMacros) aucun ne donne les mêmes valeurs (du moins à ce jour) et ce pour presque toutes les valeurs retournées (presque toutes mais pas toutes...).

## 15.10. Exercice 154.: Étude R&R (reproductibilité/répétabilité) croisée pour données continues

Minitab® Statistical Software 15.1.1.0

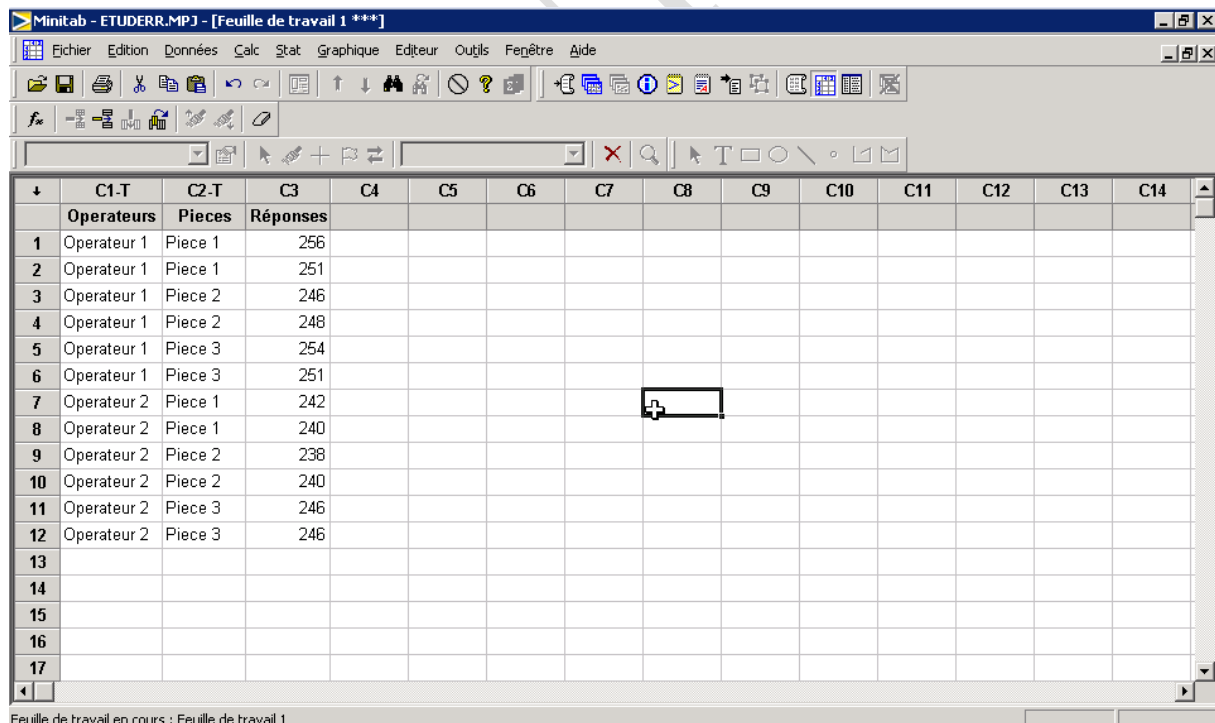
Une étude R&R mesure l'erreur de précision en faisant mesurer par exemple une pièce par plusieurs fois et par des personnes différentes et dont les pièces sont interchangeables d'une personne à l'autre (d'où le terme "croisé" et souvent le fait que cela soit considéré à juste titre comme des études R&R pour des méthodes non-désconstructives).

Étant donné que cette pièce ne change pas de taille, une variation dans les résultats doit représenter la répétabilité de la mesure et sa reproductibilité par différentes personnes.

Cette étude est très simple et basée sur l'ANOVA à deux facteurs avec répétitions et **facteurs aléatoires** dont on peut déduire les données de l'étude R&R (reproductibilité/répétabilité) à partir du tableau de l'ANOVA (TAV). Pour plus d'exemples, le lecteur peut se référer à la norme ISO/TR 12888:2011.

Considérons donc les données suivantes qui correspondent à une ANOVA à deux facteurs avec répétition (il s'agit des mêmes valeurs numériques que l'exercice de l'ANOVA à deux facteurs sans répétition vu plus haut mais détourné).

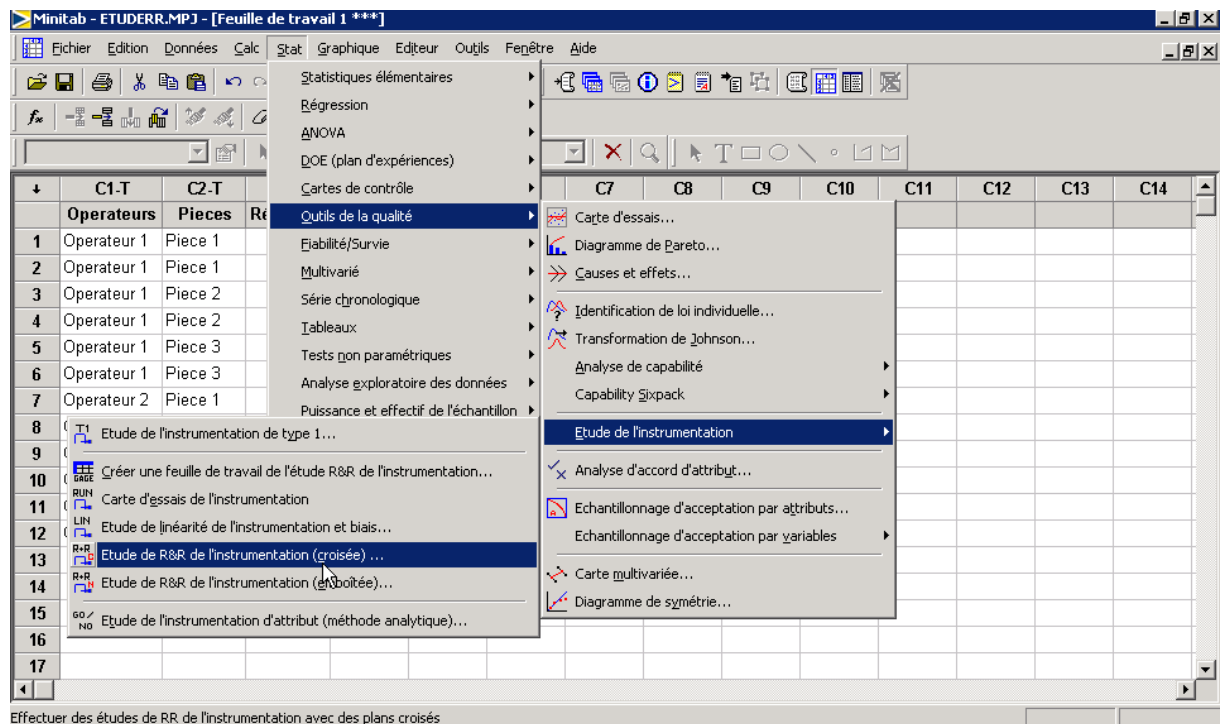
Ouvrez le fichier *EtudeRRContinu.mpj*:



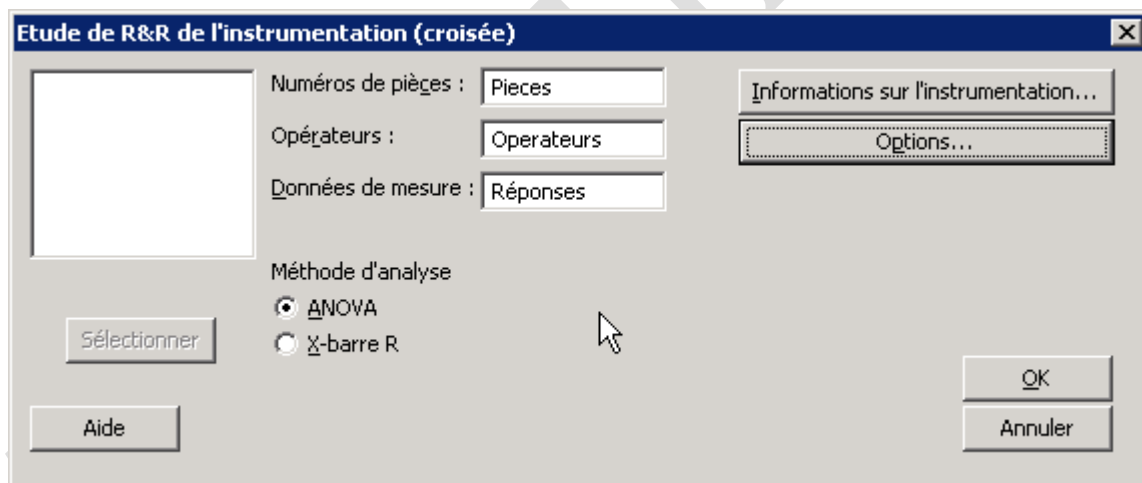
	C1-T	C2-T	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Operateurs	Pieces	Réponses											
1	Operateur 1	Piece 1	256											
2	Operateur 1	Piece 1	251											
3	Operateur 1	Piece 2	246											
4	Operateur 1	Piece 2	248											
5	Operateur 1	Piece 3	254											
6	Operateur 1	Piece 3	251											
7	Operateur 2	Piece 1	242											
8	Operateur 2	Piece 1	240											
9	Operateur 2	Piece 2	238											
10	Operateur 2	Piece 2	240											
11	Operateur 2	Piece 3	246											
12	Operateur 2	Piece 3	246											
13														
14														
15														
16														
17														

Nous avons donc 3 opérateurs qui mesurent le diamètre en microns de deux pièces.

Nous allons dans le menu **Stat/Outils de la qualité/Étude de l'instrumentation/Étude de R&R de l'instrumentation (croisée) ...**:



et nous nous utiliserons seulement sur une étude R&R (reproductibilité/répétabilité) basée sur l'ANOVA (je trouve cela plus pertinent que le carte de contrôle mentionnée):

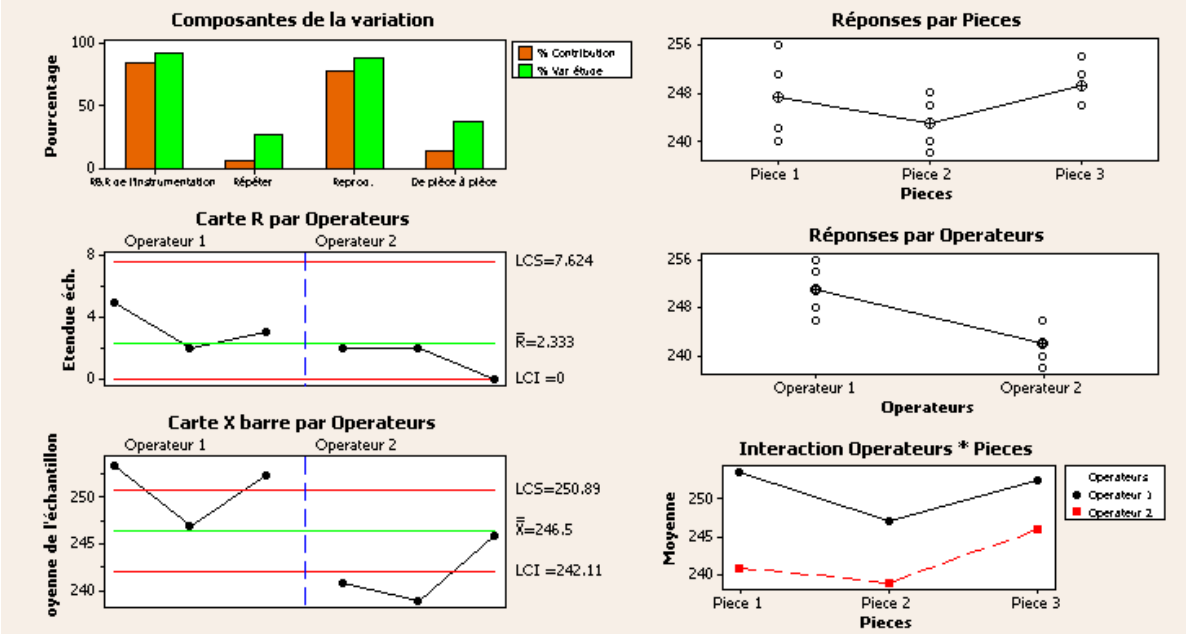


Nous validons par **OK** pour avoir:

### R&R (ANOVA) d'instrumentation pour Réponses

Nom de l'instrumentation :  
Date de l'étude :

Auteur :  
Tolérance :  
Mauv. class. :



Gardez en tête que dans les cartes de contrôle, les LCI et LCS sont calculées statistiquement selon la méthode que nous avons vu dans lors de notre étude des cartes de contrôles. Et souvenez-vous que (LCS/LCI) – qui sont calculés - ce n'est pas les (LSS/LSI) imposés !!!!!

Dans la fenêtre de session où nous retrouvons implicitement l'idée que:

$$\sigma_{total}^2 = \sigma_{pièces}^2 + \sigma_{opérateurs}^2 + \sigma_{pièces \times opérateurs}^2 + \sigma_{erreur}^2$$



**Etude R&R d'instrumentation - Méthode ANOVA****Tableau ANOVA à double entrée avec interaction**

Source	DL	Somme des carrés	CM	F	P
Pieces	2	81.5	40.750	4.1795	0.193
Operateurs	1	243.0	243.000	24.9231	0.038
Pieces * Operateurs	2	19.5	9.750	2.5435	0.158
Répétabilité	6	23.0	3.833		
Total	11	367.0			

Alpha pour supprimer le terme d'interaction = 0.25

**R&R de l'instrumentation**

Source	CompVar	% contribution (de CompVar)
R&R d'instrumentation total	45.6667	85.49
Répétabilité	3.8333	7.18
Reproductibilité	41.8333	78.32
Operateurs	38.8750	72.78
Operateurs*Pieces	2.9583	5.54
De pièce à pièce	7.7500	14.51
Variation totale	53.4167	100.00

Source	Ecart type	Variations d'étude (6 * écart type)	% var étude (%VE)
R&R d'instrumentation total	6.75771	40.5463	92.46
Répétabilité	1.95789	11.7473	26.79
Reproductibilité	6.46787	38.8072	88.50
Operateurs	6.23498	37.4099	85.31
Operateurs*Pieces	1.71998	10.3199	23.53
De pièce à pièce	2.78388	16.7033	38.09
Variation totale	7.30867	43.8520	100.00

Nombre de catégories distinctes = 1

**Bon décortiquons maintenant tout cela en commençant par les données de la fenêtre de sessions!!**

Nous avons donc en premier une table de l'ANOVA à deux facteurs avec répétitions et facteurs aléatoires (vous pouvez vérifier que c'est une ANOVA à facteurs aléatoires en exécutant l'outil habituel correspondant dans Minitab):

**Etude R&R d'instrumentation - Méthode ANOVA****Tableau ANOVA à double entrée avec interaction**

Source	DL	Somme des carrés	CM	F	P
Pieces	2	81.5	40.750	4.1795	0.193
Operateurs	1	243.0	243.000	24.9231	0.038
Pieces * Operateurs	2	19.5	9.750	2.5435	0.158
Répétabilité	6	23.0	3.833		
Total	11	367.0			

aucune surprise car c'est la même que dans l'exercice que nous avons fait plus haut et que nous avons démontré théoriquement et calculé à la main dans le cours de statistique théorique. La seule subtilité et que "**l'erreur résiduelle**" se nomme maintenant "**répétabilité**".

Les conclusions quant aux  $p$ -values de cette table se font comme à l'habitude pour une ANOVA normale.

Ensuite, vient le tableau suivant (la deuxième colonne est évidente puisqu'il ne s'agit que de la contribution de chaque valeur de la première colonne en % du total):

**R&R de l'instrumentation**

Source	CompVar	% contribution (de CompVar)
R&R d'instrumentation total	45.6667	85.49
Répétabilité	3.8333	7.18
Reproductibilité	41.8333	78.32
Operateurs	38.8750	72.78
Operateurs*Pieces	2.9583	5.54
De pièce à pièce	7.7500	14.51
Variation totale	53.4167	100.00

Pour le calculer, rappelons que nous avons 2 opérateurs ( $r = 2$ ), 3 pièces ( $k = 3$ ) et deux répétitions pour chaque mesure ( $n = 2$ ).

La première ligne du tableau *R&R d'instrumentation total* étant juste un total, il nous faut calculer la deuxième ligne que voici:

Répétabilité	3.8333
--------------	--------

et nous voyons de suite dans le tableau de l'ANOVA que:

**Etude R&R d'instrumentation - Méthode ANOVA**

**Tableau ANOVA à double entrée avec interaction**

Source	DL	Somme des carrés	CM	F	P
Pieces	2	81.5	40.750	4.1795	0.193
Operateurs	1	243.0	243.000	24.9231	0.038
Pieces * Operateurs	2	19.5	9.750	2.5435	0.158
Répétabilité	6	23.0	3.833		
Total	11	367.0			

Alpha pour supprimer le terme d'interaction = 0.25

**R&R de l'instrumentation**

Source	CompVar	% contribution (de CompVar)
R&R d'instrumentation total	45.6667	85.49
Répétabilité	3.8333	7.18
Reproductibilité	41.8333	78.32
Operateurs	38.8750	72.78
Operateurs*Pieces	2.9583	5.54
De pièce à pièce	7.7500	14.51
Variation totale	53.4167	100.00

Ce qui de par la définition mathématique de l'erreur résiduelle (variance intra-groupe) il est compréhensible qu'elle soit assimilée à la répétabilité.

La ligne *Reproductibilité* est aussi juste un total des deux lignes du dessous. Il nous suffit de calculer ces dernières pour donc avoir le total. Voyons donc comment est calculée la valeur à partir de la table de l'ANOVA:

Operateurs 38.8750

Nous avons:

$$\text{Opérateurs} = \frac{(\text{CM:Operateurs} - \text{CM:Pieces*Operateurs})}{\text{nombre de pièces} \times \text{réplications}} = \frac{243 - 9.75}{3 \cdot 2} = 38.875$$

Bingo! La reproductibilité des opérateurs s'interprète donc comme la variance pure uniquement due aux opérateurs (d'où la soustraction avec la variance d'interaction), divisé par le nombre de mesures par chaque opérateur. Il s'agit donc en quelque sorte de la variance purement intra-opérateur.

Voyons maintenant d'où vient la valeur suivante:

Operateurs\*Pieces 2.9583

Nous avons (c'est assez logique):

$$\text{Opérateurs*Pièces} = \frac{(\text{CM:Pièces*Opérateurs} - \text{CM:Répétabilité})}{\text{nombre répétitions}} = \frac{9.75 - 3.833}{2} = 2.9585$$

presque la même valeur... (il faudrait voir comment Minitab® Statistical Software fait les arrondis...).

Nous pouvons donc interpréter cette valeur comme étant la variance purement due à l'interaction (d'où la soustraction) divisé par le nombre de mesure. Il s'agit donc en quelque sorte de la variance purement intra opérateur et intra pièce.

Il ne reste plus qu'à calculer:

De pièce à pièce 7.7500

qui logiquement est:

$$\text{De pièce à pièce} = \frac{(\text{CM:Pièces} - \text{CM:Pièces*Opérateurs})}{\text{nombre d'opérateurs} \cdot \text{nombre répétitions}} = \frac{40.75 - 9.75}{2 \cdot 2} = 7.75$$

Bon maintenant que nous avons déterminé toutes les valeurs à la main pour vérification, que pouvons-nous dire globalement de ce tableau?:

#### R&R de l'instrumentation

Source	CompVar	% contribution (de CompVar)
R&R d'instrumentation total	45.6667	85.49
Répétabilité	3.8333	7.18
Reproductibilité	41.8333	78.32
Opérateurs	38.8750	72.78
Opérateurs*Pièces	2.9583	5.54
De pièce à pièce	7.7500	14.51
Variation totale	53.4167	100.00

Déjà il était clair que le tableau de l'ANOVA nous signalait clairement au vu des  $p$ -values que le choix des opérateurs semble avoir un effet sur les valeurs mais pas celui sur les pièces. Ayant connaissance de cela, le tableau R&R (reproductibilité/répétabilité) de l'instrumentation nous permet de voir que la principale variation se trouve lors d'un changement d'un opérateur moins que lors d'un changement de pièce pour un opérateur donné. Il y a donc un problème entre les opérateurs plutôt que dans la façon avec laquelle chaque opérateur mesure avec sa propre technique le diamètre de la pièce.

Pour un responsable qualité, il serait bien d'entraîner les exécutants afin d'avoir la ligne *R&R d'instrumentation total* qui converge asymptotiquement vers la valeur *De pièce à pièce*.

Maintenant, voyons le troisième tableau:

Source	Ecart type	Variations	
		d'étude (6 * écart type)	% var étude (%VE)
R&R d'instrumentation total	6.75771	40.5463	92.46
Répétabilité	1.95789	11.7473	26.79
Reproductibilité	6.46787	38.8072	88.50
Opérateurs	6.23498	37.4099	85.31
Opérateurs*Pièces	1.71998	10.3199	23.53
De pièce à pièce	2.78388	16.7033	38.09
Variation totale	7.30867	43.8520	100.00

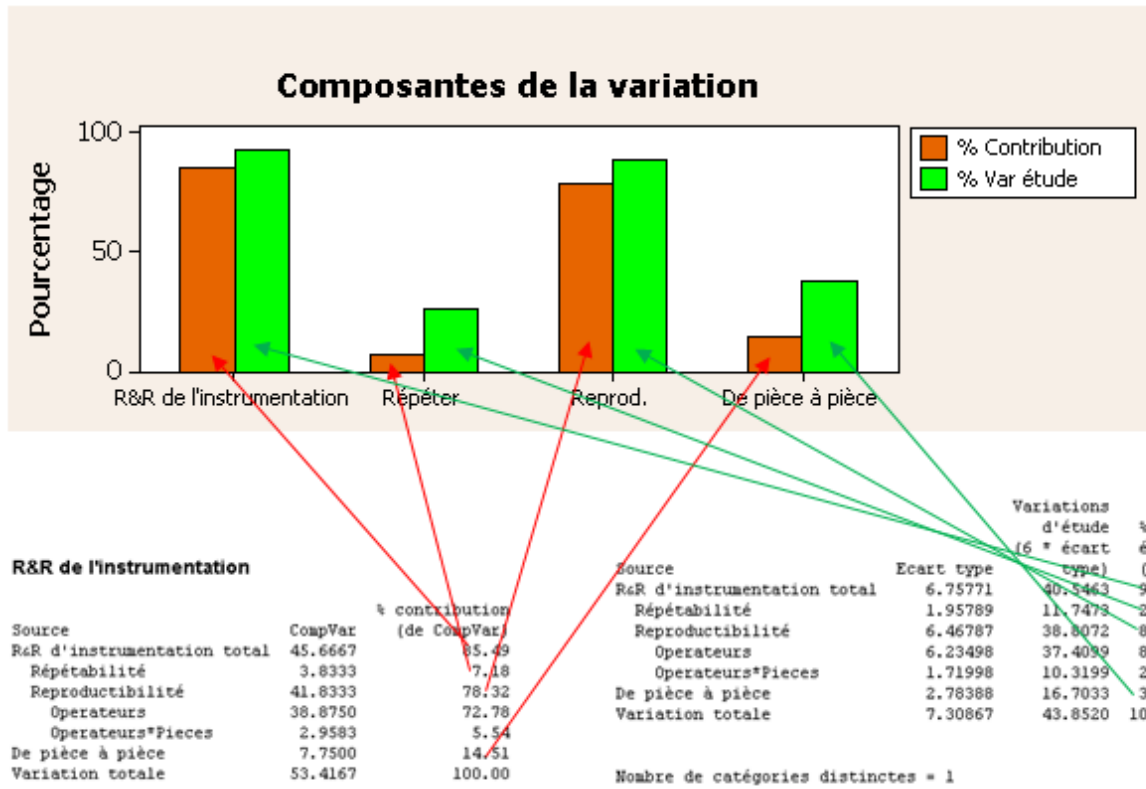
Nombre de catégories distinctes = 1

La première colonne étant nommée *Ecart type*, nous nous rendons compte qu'il s'agit simplement de la racine carrée de la variance du tableau précédent puisque:

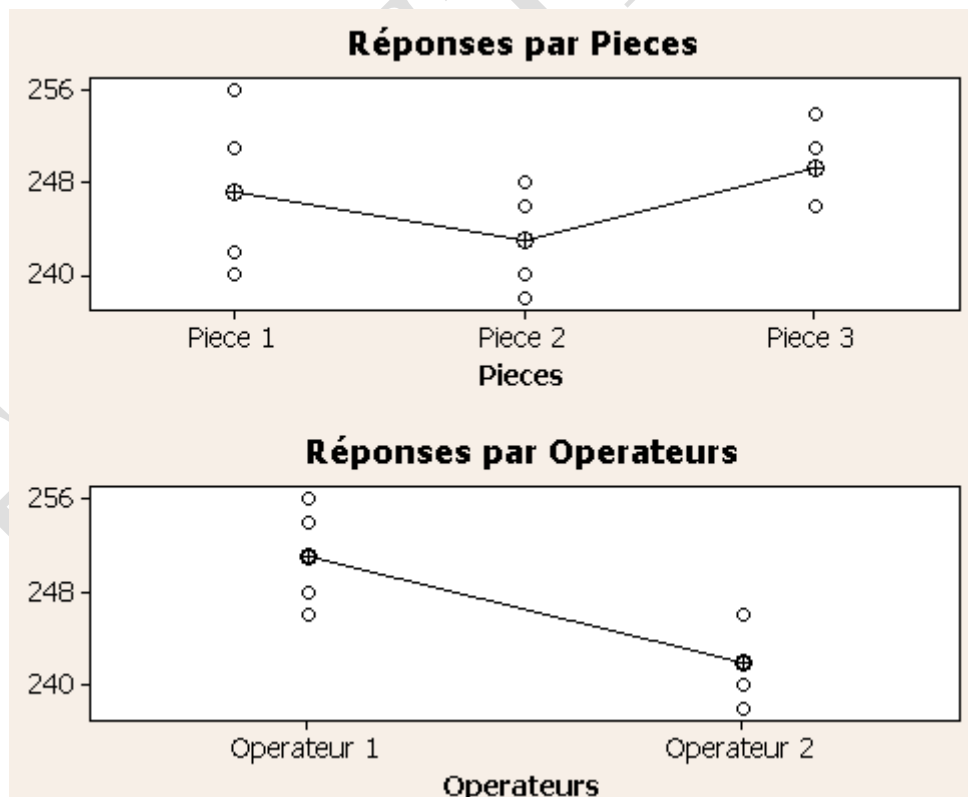
$$\sqrt{V(X)} = \sigma$$

Il s'agit donc simplement d'une autre manière de lire en utilisant les intervalles de confiance basés sur une loi Normale. Ainsi, à  $6\sigma$ , nous voyons que les opérateurs ont 92.46% de probabilité cumulée d'avoir une dispersion de  $40.5362/2 \sim 20$  microns par rapport à la moyenne. Ce qui est énorme!!! Donc dans le cas présent il faut définir les objectifs avec la direction et les clients pour savoir à combien il faut diminuer cette valeur.

Maintenant passons aux graphiques en commençant par le premier pour montrer d'où viennent ses valeurs:



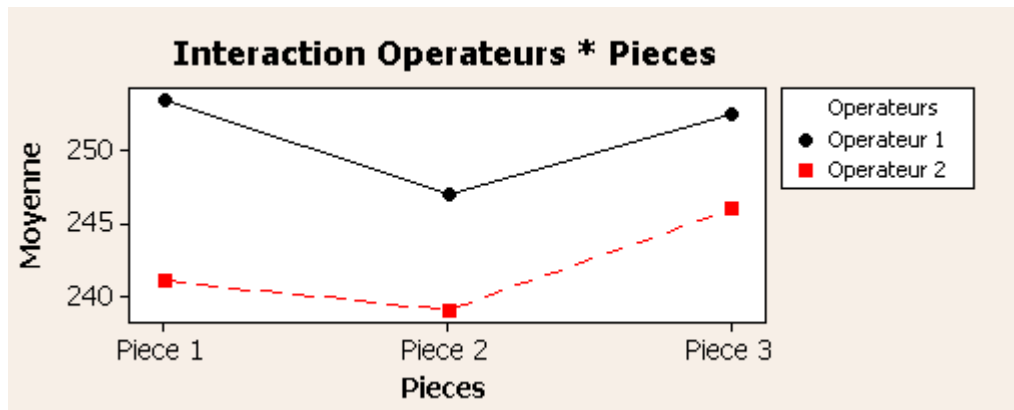
c'est très simple... il n'y pas grand-chose à en dire. Ensuite, pour les deux graphiques:



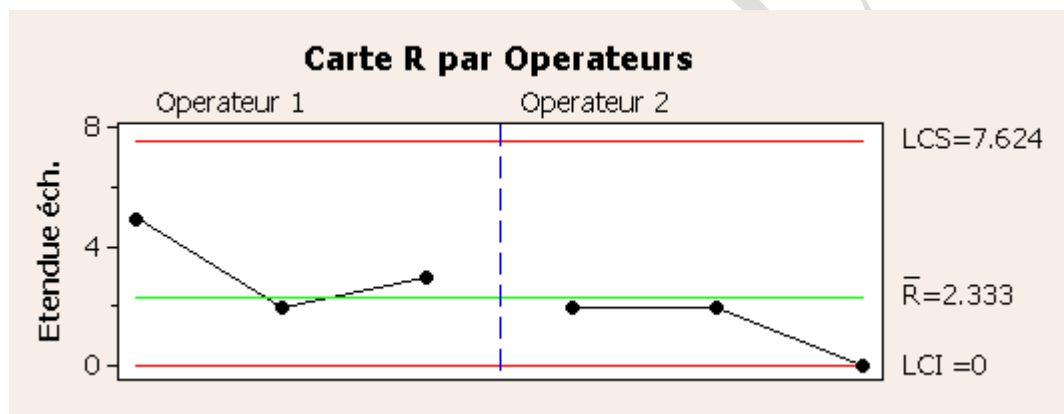
On a bien sur chacun des graphiques à chaque fois les quatre mesures + la médiane

Il n'y a pas grand-chose à dire non plus outre qu'encore une fois, on voit facilement qu'il y a un problème plus prononcé entre opérateurs qu'entre pièces.

Le graphique suivant est aussi un classique de l'ANOVA:



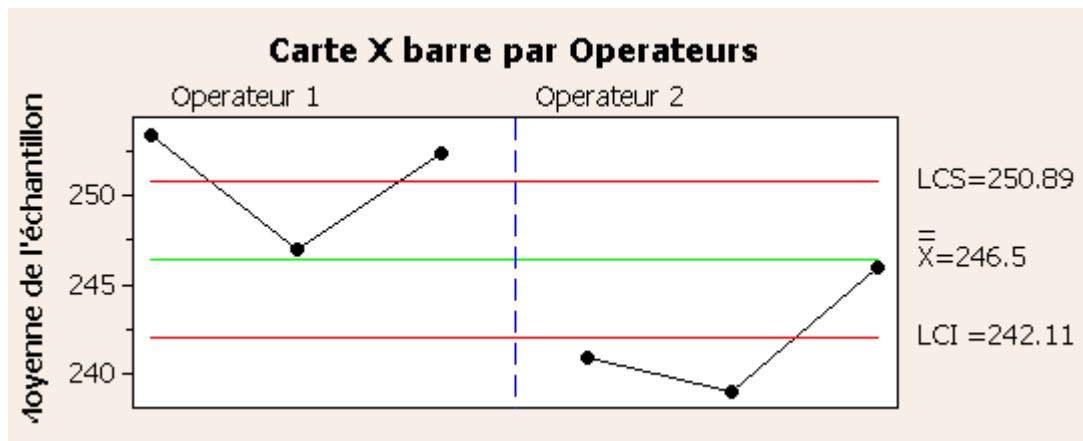
Un peu plus surprenant, c'est de retrouver des cartes de contrôle. Nous avons en premier:



il s'agit donc d'une carte de contrôle  $R$ -bar  $R$  que nous allons retrouver plus loin et dont nous avons démontré le mécanisme dans le cours de statistique théorique. C'est carte n'a aucun raison d'exister dans le cas présent pour les raisons suivantes (opinion personnelle mais à débattre en classe):

1. Il n'y pas de flèche du temps dans une étude R&R (reproductibilité/répétabilité)
2. Normalement chaque point de la carte représente une étendue et il est clair que dans le cas présent une étendue de deux mesures n'a aucun sens
3. Bien que la carte  $R$ -bar  $R$  soit fait pour moins de dix mesures par échantillonnage, il ne faut pas exagérer en considérant qu'elle fonctionne aussi pour deux mesures par échantillonnage...
4. Normalement cette carte représente les mesures pour une même famille de pièces et ceci n'est pas spécifié dans l'exécution du test par le logiciel.

Enfin le dernier graphique qui est encore une carte de contrôle:



a encore moins de raison d'être présente! Elle n'est statistiquement absolument pas significative dans le cas présent. Mais nous allons en débattre en classe.

Bref, ces deux derniers graphiques n'ont donc de sens que si l'on fait mesurer de nombreuses fois (plus que deux fois en tout cas....) une même pièce à un opérateur.



## 15.11. Exercice 155.: Créer une feuille de travail pour l'analyse R&R

Minitab® Statistical Software 15.1.1.0

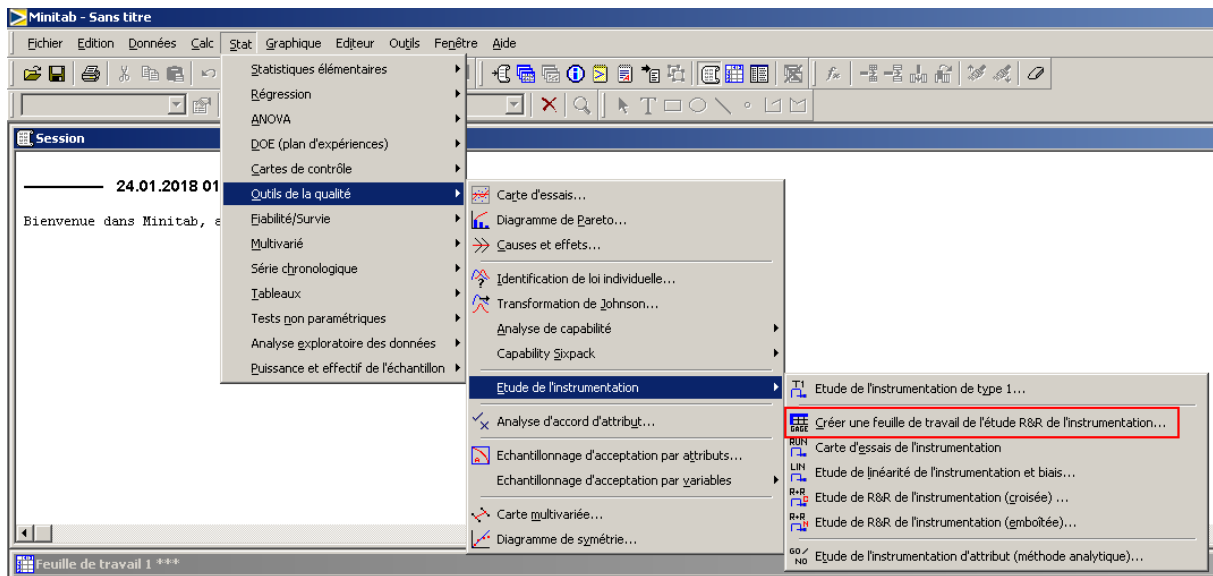
Dans l'exemple précédent nous avons fait usage du tableau suivant:

	C1-T	C2-T	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Operateurs	Pieces	Réponses											
1	Operateur 1	Piece 1	256											
2	Operateur 1	Piece 1	251											
3	Operateur 1	Piece 2	246											
4	Operateur 1	Piece 2	248											
5	Operateur 1	Piece 3	254											
6	Operateur 1	Piece 3	251											
7	Operateur 2	Piece 1	242											
8	Operateur 2	Piece 1	240											
9	Operateur 2	Piece 2	238											
10	Operateur 2	Piece 2	240											
11	Operateur 2	Piece 3	246											
12	Operateur 2	Piece 3	246											
13														
14														
15														
16														
17														

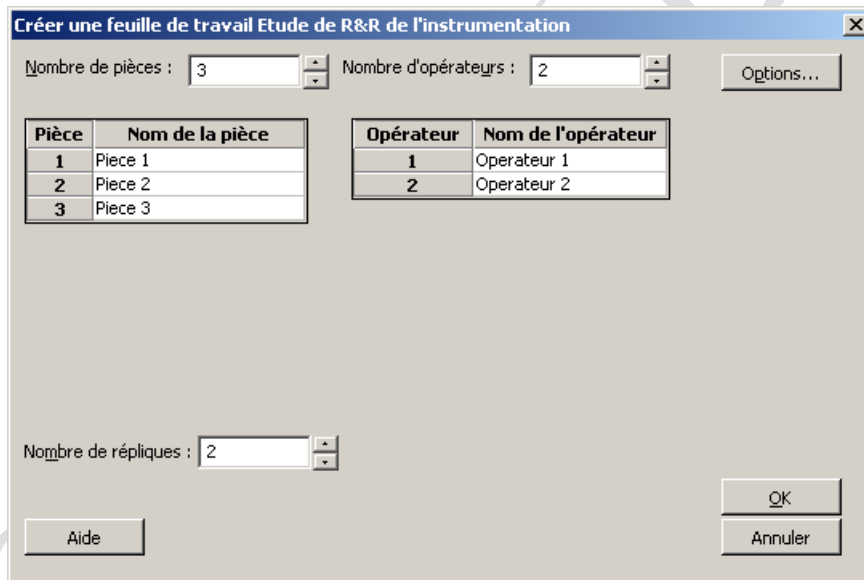
On peut se demander s'il n'y pas une méthode/outil rapide dans Minitab pour construire un tel tableau?

La réponse est bien évidemment: *Oui!* Voyons cela!

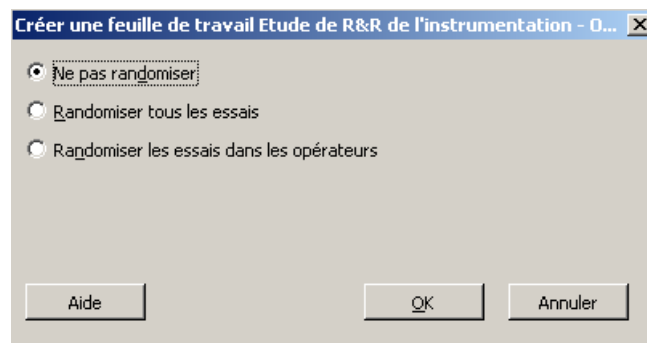
Nous allons dans le menu **Outils de la qualité/Etude de l'instrumentation/Créer une feuille de travail...** :



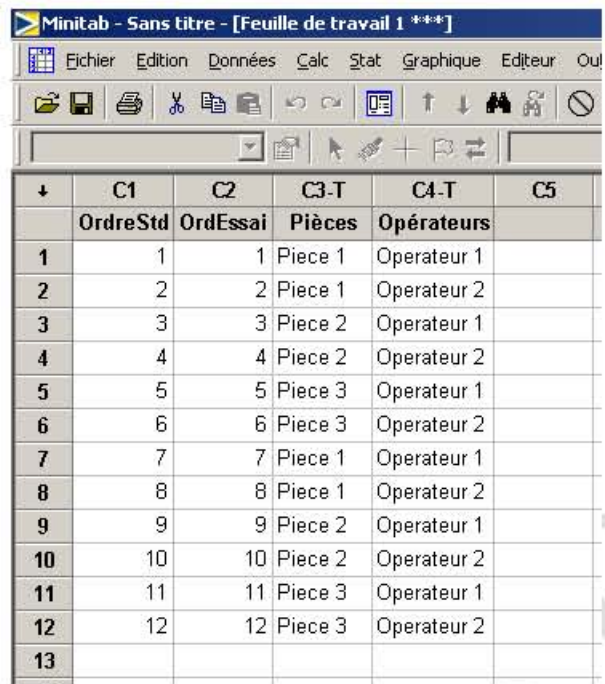
Et pour reproduire au mieux l'exemple précédent, dans la boîte de dialogue qui apparaît, nous mettons:



Dans le bouton **Options...** nous prenons *Ne pas randomiser* si nous voulons effectivement pouvoir reproduire l'exemple précédent au mieux:



Et nous cliquons deux fois sur **OK** pour obtenir:



The image shows a screenshot of a Minitab spreadsheet window titled "Minitab - Sans titre - [Feuille de travail 1 \*\*\*]". The window contains a data table with 13 rows and 5 columns. The columns are labeled C1, C2, C3-T, C4-T, and C5. The data in the table is as follows:

	C1	C2	C3-T	C4-T	C5
	OrdreStd	OrdEssai	Pièces	Opérateurs	
1	1	1	Piece 1	Operateur 1	
2	2	2	Piece 1	Operateur 2	
3	3	3	Piece 2	Operateur 1	
4	4	4	Piece 2	Operateur 2	
5	5	5	Piece 3	Operateur 1	
6	6	6	Piece 3	Operateur 2	
7	7	7	Piece 1	Operateur 1	
8	8	8	Piece 1	Operateur 2	
9	9	9	Piece 2	Operateur 1	
10	10	10	Piece 2	Operateur 2	
11	11	11	Piece 3	Operateur 1	
12	12	12	Piece 3	Operateur 2	
13					

Et voilà!

## 15.12. Exercice 156.: Étude R&R (reproductibilité/répétabilité) emboîtée pour données continues

Minitab® Statistical Software 18.1

Une étude R&R mesure l'erreur de précision en faisant mesurer par exemple une pièce par plusieurs fois et par des personnes différentes et dont les pièces ne sont pas interchangeables d'une personne à l'autre (d'où le terme "emboîté" et souvent le fait que cela soit considéré à juste titre comme des études R&R pour des méthodes destructives).

Étant donné que cette pièce ne change pas de taille, une variation dans les résultats doit représenter la répétabilité de la mesure et sa reproductibilité par différentes personnes.

Cette étude est très simple et basée sur l'ANOVA à deux facteurs emboîtée avec répétition dont on peut déduire les données de l'étude R&R (reproductibilité/répétabilité) à partir du tableau de l'ANOVA (TAV). Pour plus d'exemples, le lecteur peut à nouveau se référer à la norme ISO/TR 12888:2011.

Considérons donc les données suivantes qui correspondent à une ANOVA à deux facteurs hiérarchique avec répétition (il s'agit des mêmes valeurs numériques que l'exercice de l'ANOVA à deux facteurs sans répétition vu plus haut mais détourné).

Ouvrez le fichier *ANOVAHierarchique.mpj*:

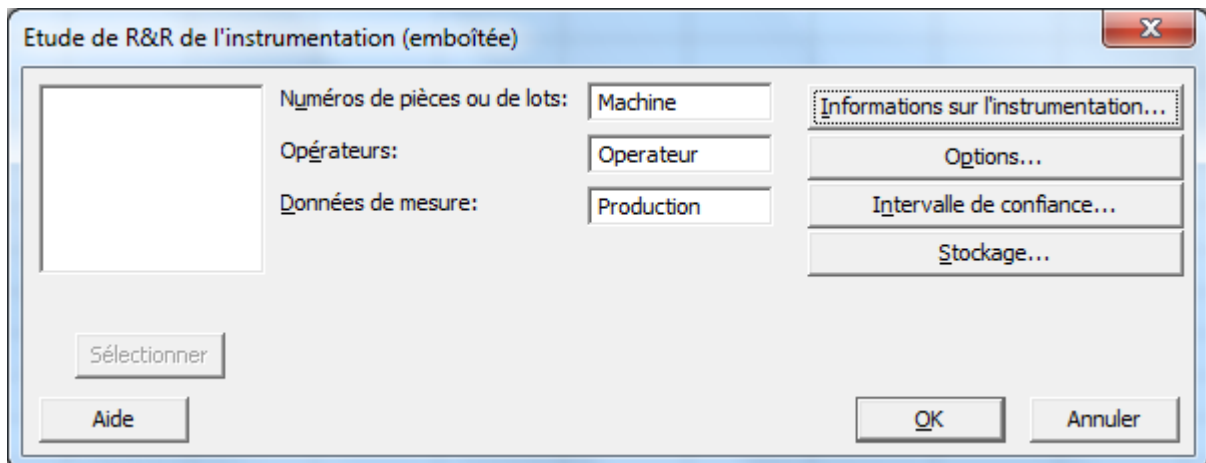
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19
	Production	Machine	Operateur																
1	10	1	1																
2	14	1	1																
3	9	1	1																
4	12	2	1																
5	8	2	1																
6	10	2	1																
7	8	3	1																
8	10	3	1																
9	12	3	1																
10	13	4	1																
11	12	4	1																
12	11	4	1																
13	11	1	2																
14	14	1	2																
15	8	1	2																
16	13	2	2																
17	11	2	2																
18	9	2	2																
19	9	3	2																
20	10	3	2																
21	8	3	2																
22	10	4	2																
23	9	4	2																
24	8	4	2																
25	13	1	3																
26	10	1	3																
27	16	1	3																

Dont voici la suite des données:

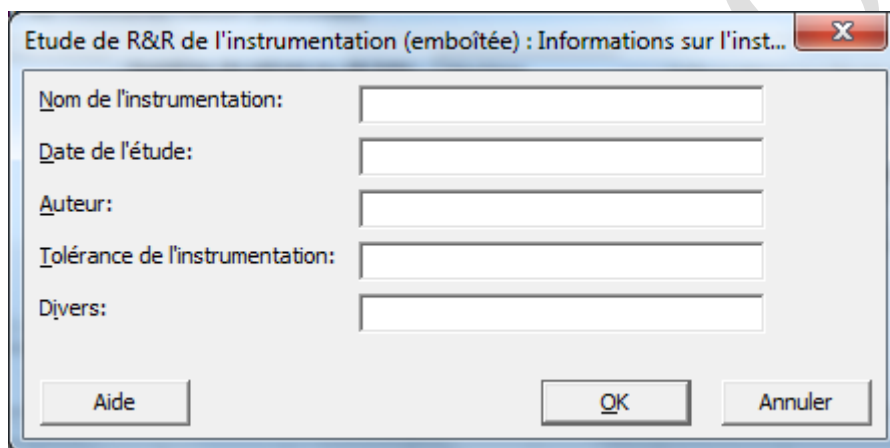
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19
26	10	1	3																
27	16	1	3																
28	14	2	3																
29	13	2	3																
30	12	2	3																
31	7	3	3																
32	9	3	3																
33	5	3	3																
34	10	4	3																
35	7	4	3																
36	4	4	3																

Nous allons ensuite dans le menu **Outils de la qualité / Etude de l'instrumentation / Etude de R&R de l'instrumentation (emboîtée) ...** :

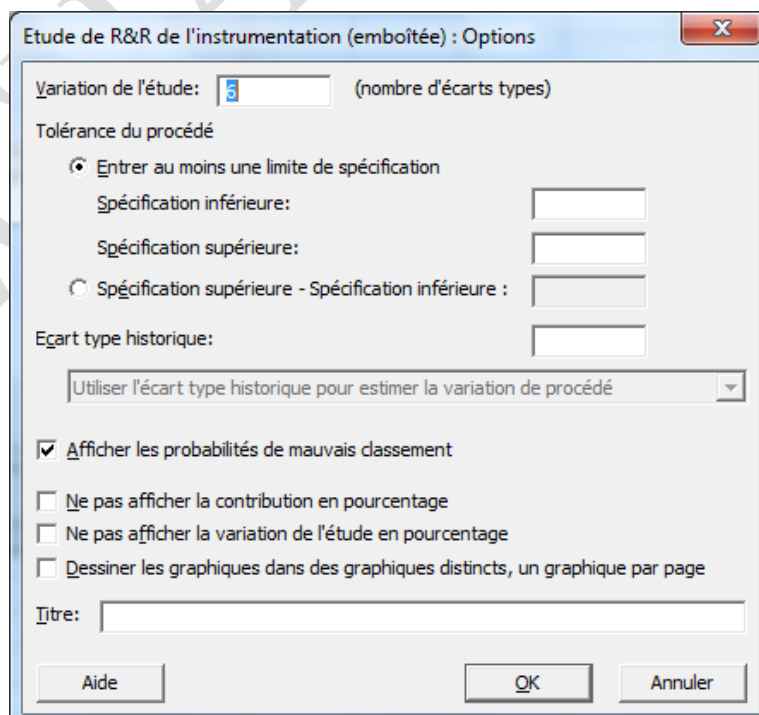
Nous avons alors la boîte de dialogue suivante:



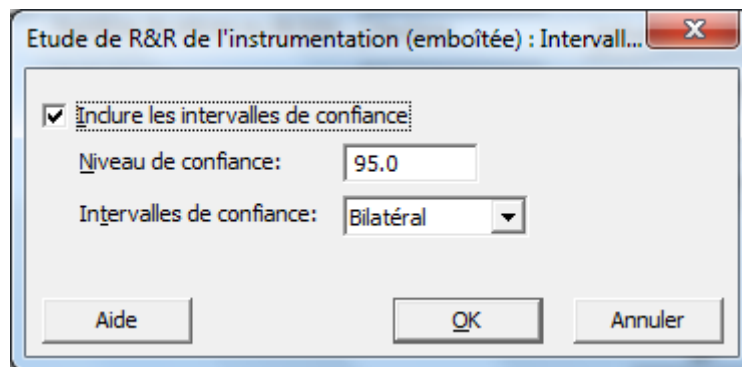
Si on clique sur le premier bouton **Informations sur l'instrumentation...** on obtient juste des métadonnées pour les graphiques:



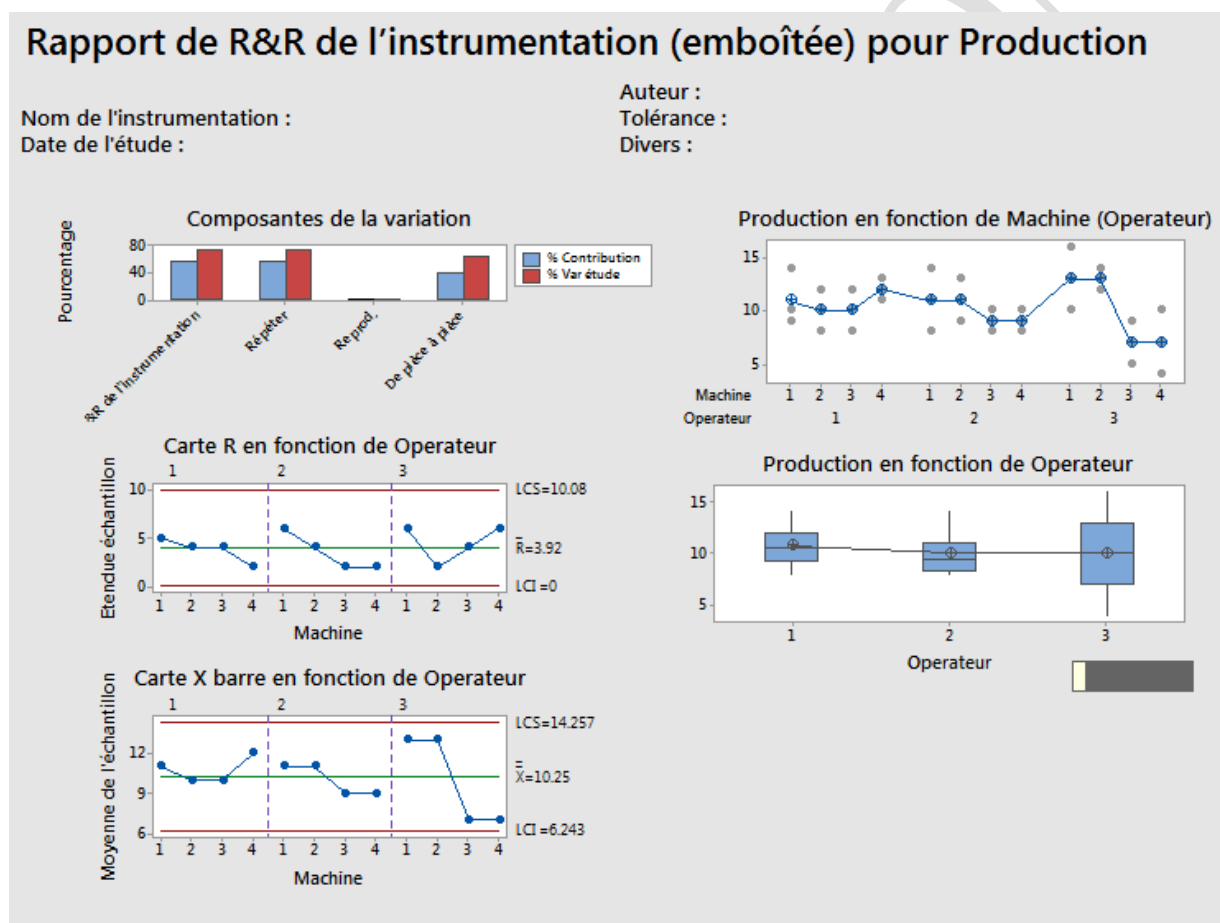
Dans **Options...** c'est déjà plus intéressant:



Et dans **Intervalle de confiance...** on prendra:



Ce qui donne le graphique suivant:



Gardez en tête que dans la carte de contrôle, les LCI et LCS sont calculées statistiquement selon la méthode que nous avons vu dans lors de notre étude des cartes de contrôles. Et souvenez-vous que (LCS/LCI) – qui sont calculés - ce n'est pas le (LSS/LSI) imposé !!!!!

Et dans la fenêtre de session nous obtenons:

## Etude R&R de l'instrumentation - ANOVA emboîtée

### R&R de l'instrumentation (emboîtée) pour Production

Source	DL	Somme des carrés	CM	F	P
Operateur	2	4.50	2.25	0.15789	0.856
Machine (Operateur)	9	128.25	14.25	3.16667	0.012
Répétabilité	24	108.00	4.50		
Total	35	240.75			

### Composantes de la variance

Source	CompVar	IC à 95 %	% contribution (de CompVar)	IC à 95 %
R&R de l'instrumentation totale	4.50	(1.337; 12.729)	58.06	(2.56; 93.15)
Répétabilité	4.50	(2.744; 8.709)	58.06	(24.80; 100.00)
Reproductibilité	0.00	(0.000; 7.072)	0.00	(0.00; 43.29)
De pièce à pièce	3.25	(0.382; 14.293)	41.94	(6.85; 97.44)
Variation totale	7.75	(5.533; 19.110)	100.00	

Nous verrons après que la reproductibilité aura toujours une estimation ponctuelle égale à 0.00000 car c'est inhérent au principe de considérer les opérateurs (machines) comme variables aléatoires et donc qui ne devraient pas avoir mathématiquement de variation et ce par définition (puisque'ils peuvent être mathématiquement infiniment proches).

### Evaluation de l'instrumentation

Source	Ecart type	IC à 95 %	Variations d'étude (6 × écart type)	IC à 95 %
R&R de l'instrumentation totale	2.12132	(1.156; 3.568)	12.7279	(6.937; 21.407)
Répétabilité	2.12132	(1.656; 2.951)	12.7279	(9.938; 17.706)
Reproductibilité	0.00000	(0.000; 2.659)	0.0000	(0.000; 15.956)
De pièce à pièce	1.80278	(0.618; 3.781)	10.8167	(3.711; 22.684)
Variation totale	2.78388	(2.352; 4.371)	16.7033	(14.114; 26.229)

Source	% var étude (%VE)	IC à 95 %
R&R de l'instrumentation totale	76.20	(16.01; 96.51)
Répétabilité	76.20	(49.80; 100.00)
Reproductibilité	0.00	(0.00; 65.79)
De pièce à pièce	64.76	(26.17; 98.71)
Variation totale	100.00	

Nombre de catégories distinctes = 1

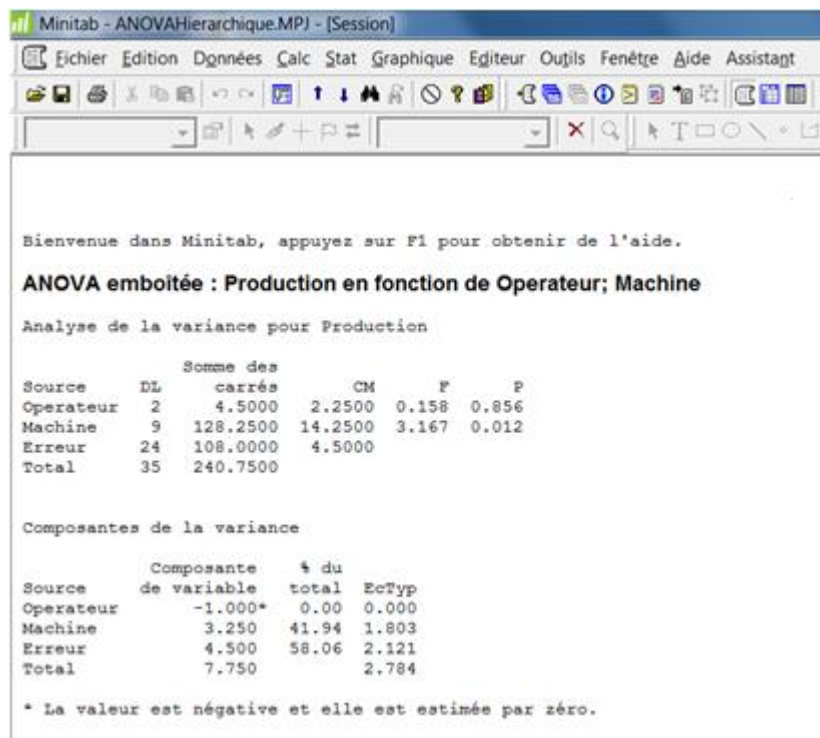
IC à 95 % = (0.383531; 8.71706)

Analysons tout cela en commençant par l'ANOVA emboîtée:

Source	DL	Somme des carrés	CM	F	P
Operateur	2	4.50	2.25	0.15789	0.856
Machine (Operateur)	9	128.25	14.25	3.16667	0.012
Répétabilité	24	108.00	4.50		
Total	35	240.75			



Déjà nous remarquons ici que nous retrouvons les mêmes valeurs que lors de notre étude de l'ANOVA hiérarchique (en se rappelant que l'ANOVA hiérarchique prend automatiquement les *Opérateurs* et *Machines* en tant que variable aléatoire!!!). Effectivement, rappelons que nous avons obtenu:



Le fait que tous les facteurs soient considérés comme aléatoires par défaut peut poser problème dans la pratique! Nous verrons comment résoudre cela lors de notre découverte de l'étude R&R étendue.

Pour le reste, voici suite à la demande d'un participant! Commençons par la répétabilité (moyen de mesure):

### Etude R&R de l'instrumentation - ANOVA emboîtée

#### R&R de l'instrumentation (emboîtée) pour Production

Source	DL	Somme des carrés	CM	F	P
Operateur	2	4.50	2.25	0.15789	0.856
Machine (Operateur)	9	128.25	14.25	3.16667	0.012
Répétabilité	24	108.00	4.50		
Total	35	240.75			

#### Composantes de la variance

Source	CompVar	IC à 95 %	% contribution (de CompVar)	IC à 95 %
R&R de l'instrumentation totale	4.50	(1.337; 12.729)	58.06	(2.56; 93.15)
Répétabilité	4.50	(2.744; 8.709)	58.06	(24.80; 100.00)
Reproductibilité	0.00	(0.000; 7.072)	0.00	(0.00; 43.29)
De pièce à pièce	3.25	(0.382; 14.293)	41.94	(6.85; 97.44)
Variation totale	7.75	(5.533; 19.110)	100.00	

La répétabilité est donc juste la valeur CM des résidus de l'ANOVA. Jusque-là rien de nouveau, c'est comme dans la R&R croisée!

Passons maintenant à la reproductibilité (propre aux opérateurs pour rappel!):

## Etude R&R de l'instrumentation - ANOVA emboîtée

### R&R de l'instrumentation (emboîtée) pour Production

Source	DL	Somme des carrés	CM	F	P
Operateur	2	4.50	2.25	0.15789	0.856
Machine (Operateur)	9	128.25	14.25	3.16667	0.012
Répétabilité	24	108.00	4.50		
Total	35	240.75			

### Composantes de la variance

Source	CompVar	IC à 95 %	% contribution (de CompVar)	IC à 95 %
R&R de l'instrumentation totale	4.50	(1.337; 12.729)	58.06	(2.56; 93.15)
Répétabilité	4.50	(2.744; 8.709)	58.06	(24.80; 100.00)
Reproductibilité	0.00	(0.000; 7.072)	0.00	(0.00; 43.29)
De pièce à pièce	3.25	(0.382; 14.293)	41.94	(6.85; 97.44)
Variation totale	7.75	(5.533; 19.110)	100.00	

La reproductibilité sera toujours nulle par construction dans une ANOVA dont les facteurs sont aléatoires comme nous l'avons déjà mentionné (sinon quoi ils ne seraient pas aléatoires!).

Passons maintenant au pièce à pièce:

## Etude R&R de l'instrumentation - ANOVA emboîtée

### R&R de l'instrumentation (emboîtée) pour Production

Source	DL	Somme des carrés	CM	F	P
Operateur	2	4.50	2.25	0.15789	0.856
Machine (Operateur)	9	128.25	14.25	3.16667	0.012
Répétabilité	24	108.00	4.50		
Total	35	240.75			

### Composantes de la variance

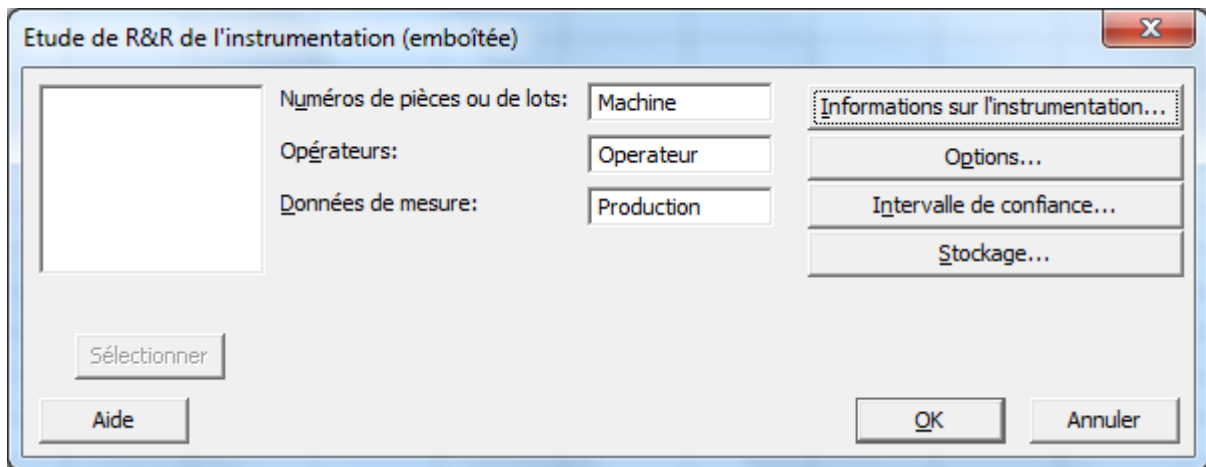
Source	CompVar	IC à 95 %	% contribution (de CompVar)	IC à 95 %
R&R de l'instrumentation totale	4.50	(1.337; 12.729)	58.06	(2.56; 93.15)
Répétabilité	4.50	(2.744; 8.709)	58.06	(24.80; 100.00)
Reproductibilité	0.00	(0.000; 7.072)	0.00	(0.00; 43.29)
De pièce à pièce	3.25	(0.382; 14.293)	41.94	(6.85; 97.44)
Variation totale	7.75	(5.533; 19.110)	100.00	

Qui est donné aussi pour le coup par:

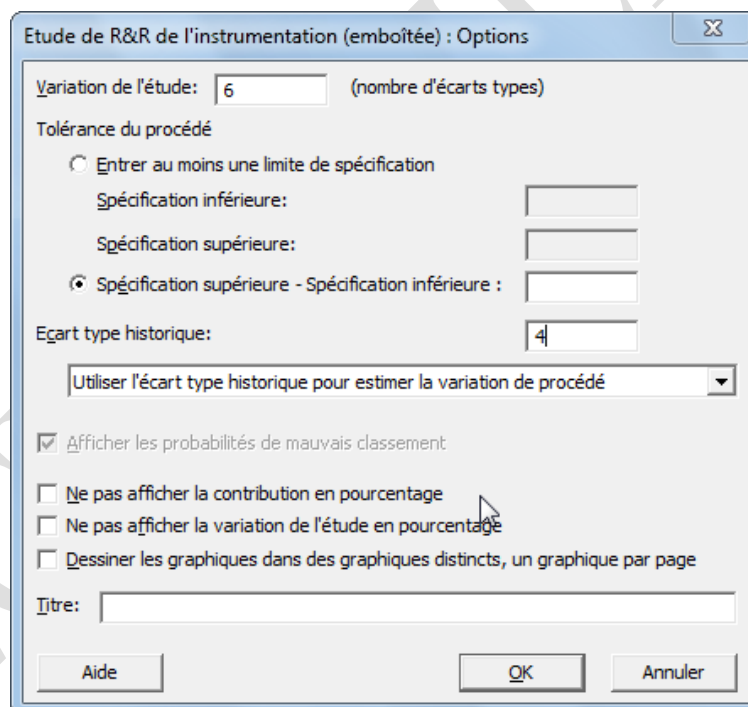
$$\text{De pièce à pièce} = \frac{(\text{CM:Machine(Opérateur)} - \text{CM:Répétabilité})}{\text{nombre de réplifications}} = \frac{14.25 - 4.50}{3} = 3.25$$

## Ajout d'une limite LSL/USL

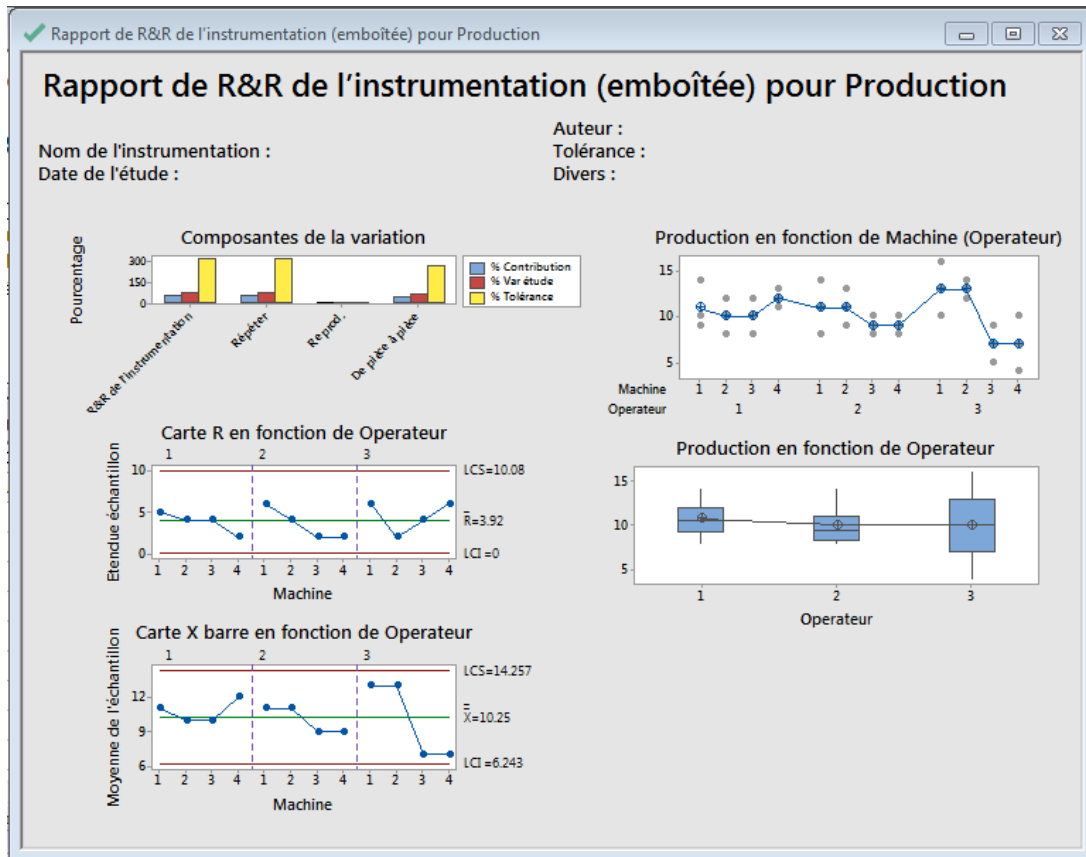
Lorsque l'on lance l'analyse, via le bouton **Options**:



On peut définir un certain nombre de sigma (écarts types) via le champ **Variation de l'étude** ainsi que des spécifications USL/LSL (**Tolérance du procédé**). Si nous faisons les choix suivants:



Ce qui dans le graphique donne une nouvelle barre en jaune nommée **% Tolérance**:



Et pour le coup l'évaluation de l'instrumentation contient de nouvelles colonnes mises en évidence ci-dessous:

### Evaluation de l'instrumentation

Source	Ecart type	IC à 95 %	Variations d'étude	
			(6 × écart type)	IC à 95 %
R&R de l'instrumentation totale	2.12132	(1.156; 3.568)	12.7279	(6.937; 21.407)
Répétabilité	2.12132	(1.656; 2.951)	12.7279	(9.938; 17.706)
Reproductibilité	0.00000	(0.000; 2.659)	0.0000	(0.000; 15.956)
De pièce à pièce	1.80278	(0.618; 3.781)	10.8167	(3.711; 22.684)
Variation totale	2.78388	(2.352; 4.371)	16.7033	(14.114; 26.229)

Source	% var étude (%VE)	IC à 95 %	% tolérance	
			(VE/Tolér.)	IC à 95 %
R&R de l'instrumentation totale	76.20	(16.01; 96.51)	318.20	(173.43; 535.17)
Répétabilité	76.20	(49.80; 100.00)	318.20	(248.46; 442.66)
Reproductibilité	0.00	(0.00; 65.79)	0.00	(0.00; 398.90)
De pièce à pièce	64.76	(26.17; 98.71)	270.42	(92.76; 567.10)
Variation totale	100.00		417.58	(352.85; 655.72)

Alors d'abord, pour rappel... la colonne *Ecart Type* est simplement la racine carrée de la colonne *CompVar*:

### Composantes de la variance

Source	CompVar	IC à 95 %	% contribution (de CompVar)	IC à 95 %
R&R de l'instrumentation totale	4.50	(1.337; 12.729)	58.06	(2.56; 93.15)
Répétabilité	4.50	(2.744; 8.709)	58.06	(24.80; 100.00)
Reproductibilité	0.00	(0.000; 7.072)	0.00	(0.00; 43.29)
De pièce à pièce	3.25	(0.382; 14.293)	41.94	(6.85; 97.44)
Variation totale	7.75	(5.533; 19.110)	100.00	

Les autres colonnes sont triviales jusqu'à:

### Evaluation de l'instrumentation

Source	Ecart type	Variations d'étude		
		IC à 95 %	(6 × écart type)	IC à 95 %
R&R de l'instrumentation totale	2.12132	(1.156; 3.568)	12.7279	(6.937; 21.407)
Répétabilité	2.12132	(1.656; 2.951)	12.7279	(9.938; 17.706)
Reproductibilité	0.00000	(0.000; 2.659)	0.0000	(0.000; 15.956)
De pièce à pièce	1.80278	(0.618; 3.781)	10.8167	(3.711; 22.684)
Variation totale	2.78388	(2.352; 4.371)	16.7033	(14.114; 26.229)

Source	% var étude (%VE)	IC à 95 %	% tolérance (VE/Tolér.)	IC à 95 %
R&R de l'instrumentation totale	76.20	(16.01; 96.51)	318.20	(173.43; 535.17)
Répétabilité	76.20	(49.80; 100.00)	318.20	(248.46; 442.66)
Reproductibilité	0.00	(0.00; 65.79)	0.00	(0.00; 398.90)
De pièce à pièce	64.76	(26.17; 98.71)	270.42	(92.76; 567.10)
Variation totale	100.00		417.58	(352.85; 655.72)

Donc d'où vient le 318.20 par exemple? Eh bien :

$$\% \text{Tolerance} = \frac{\text{var Etude}}{USL - LSL} = \frac{6\hat{\sigma}_i}{USL - LSL} = \frac{12.7279}{4} \% = 318.20$$

## 15.13. Exercice 157.: Étude R&R (reproductibilité/répétabilité) étendue pour données continues

Minitab® Statistical Software 18.1

Alors voici une analyse qui va nous permettre de sélectionner plus que 2 niveaux et aussi de choisir quels facteurs de l'ANOVA seront fixes et lesquels seront aléatoires.

### R&R emboîtée avec facteurs aléatoires

Pour bien comprendre l'utilisation de cette nouvelle option disponible depuis Minitab 16, le mieux est d'apprendre à reproduire l'Étude R&R emboîtée que nous avons faite précédemment et qui pour rappel nous avait donnée:

#### Etude R&R de l'instrumentation - ANOVA emboîtée

##### R&R de l'instrumentation (emboîtée) pour Production

Source	DL	Somme des carrés	CM	F	P
Operateur	2	4.50	2.25	0.15789	0.856
Machine (Operateur)	9	128.25	14.25	3.16667	0.012
Répétabilité	24	108.00	4.50		
Total	35	240.75			

##### Composantes de la variance

Source	CompVar	IC à 95 %	% contribution (de CompVar)	IC à 95 %
R&R de l'instrumentation totale	4.50	(1.337; 12.729)	58.06	(2.56; 93.15)
Répétabilité	4.50	(2.744; 8.709)	58.06	(24.80; 100.00)
Reproductibilité	0.00	(0.000; 7.072)	0.00	(0.00; 43.29)
De pièce à pièce	3.25	(0.382; 14.293)	41.94	(6.85; 97.44)
Variation totale	7.75	(5.533; 19.110)	100.00	

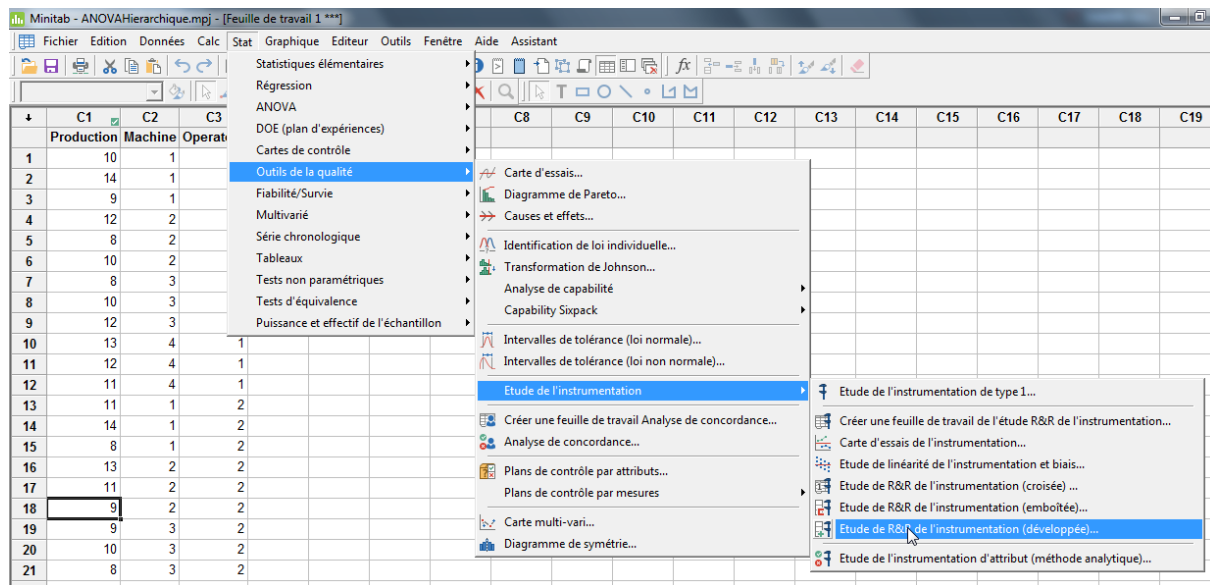
Donc repartons des mêmes données:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19
	Production	Machine	Operateur																
1	10	1	1																
2	14	1	1																
3	9	1	1																
4	12	2	1																
5	8	2	1																
6	10	2	1																
7	8	3	1																
8	10	3	1																
9	12	3	1																
10	13	4	1																
11	12	4	1																
12	11	4	1																
13	11	1	2																
14	14	1	2																
15	8	1	2																
16	13	2	2																
17	11	2	2																
18	9	2	2																
19	9	3	2																
20	10	3	2																
21	8	3	2																
22	10	4	2																
23	9	4	2																
24	8	4	2																
25	13	1	3																
26	10	1	3																
27	16	1	3																

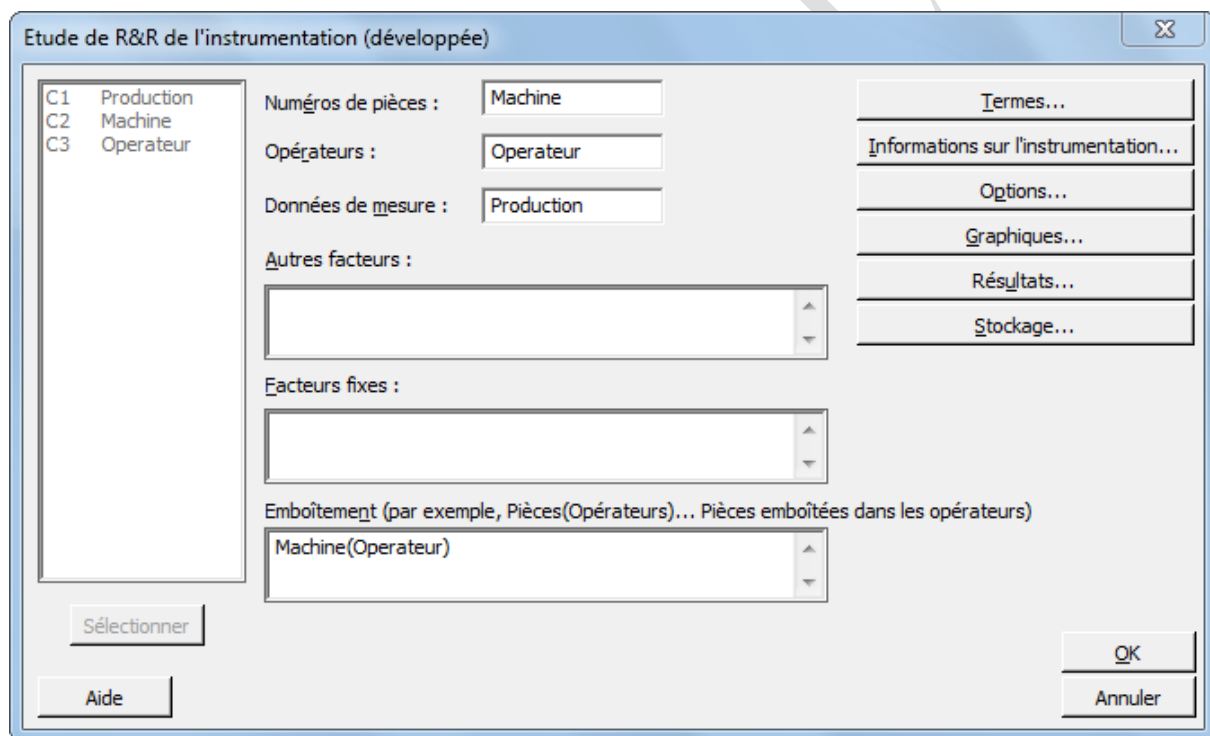
Dont voici la suite des données:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19
	Production	Machine	Operateur																
26	10	1	3																
27	16	1	3																
28	14	2	3																
29	13	2	3																
30	12	2	3																
31	7	3	3																
32	9	3	3																
33	5	3	3																
34	10	4	3																
35	7	4	3																
36	4	4	3																
37																			
38																			
39																			
40																			
41																			
42																			
43																			
44																			
45																			
46																			
47																			
48																			
49																			
50																			
51																			
52																			

Nous allons dans le menu **Stat/Outils de la qualité/Etude de l'instrumentation/Etude R&R de l'instrumentation étendue...**:



Nous avons alors la boîte de dialogue que nous avons rempli comme ci-dessous:



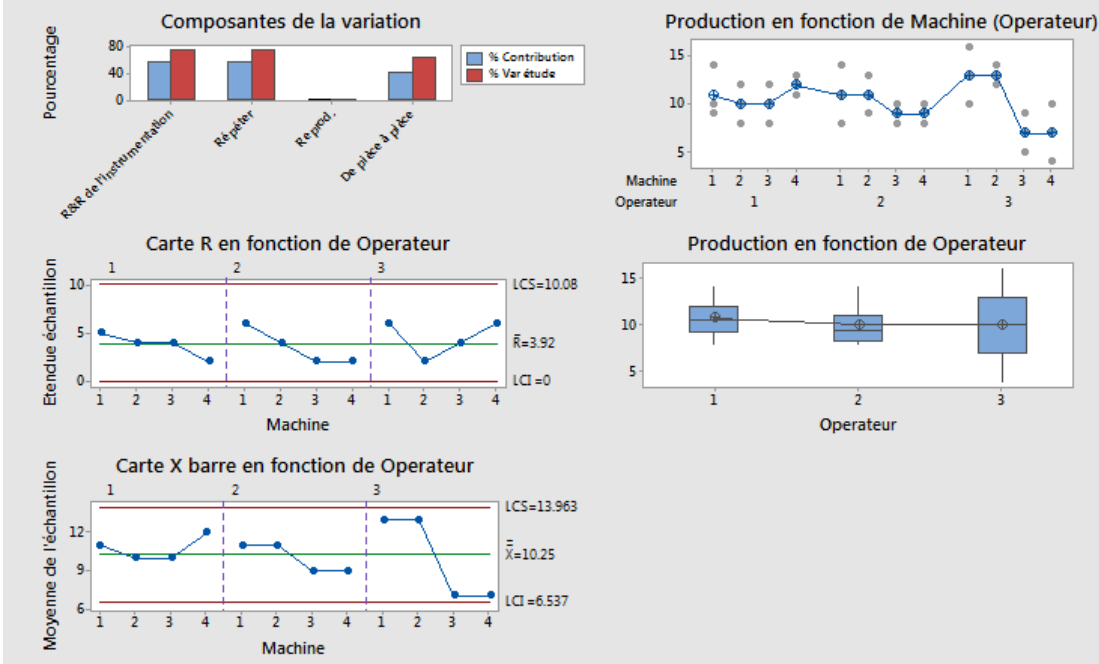
Remarquez que nous n'avons rien mis dans **Facteurs fixes** et que nous avons explicité le type d'**Emboîtement**! Ce qui suppose implicitement que nous imposons tous les facteurs comme étant aléatoires comme vont nous les prouver les captures d'écran suivantes:



## Rapport de R&R de l'instrumentation (développée) pour Production

Nom de l'instrumentation :  
Date de l'étude :

Auteur :  
Tolérance :  
Divers :



Le graphique est effectivement exactement le même que pour l'étude R&R emboîtée!!!!

Concernant la fenêtre de session, nous avons:

### Etude R&R de l'instrumentation : Production en fonction ... Operateur

#### Informations sur les facteurs

Facteur	Type	Niveaux	Valeurs
Machine(Operateur)	Aléatoire	12	1(1); 2(1); 3(1); 4(1); 1(2); 2(2); 3(2); 4(2); 1(3); 2(3); 3(3); 4(3)
Operateur	Aléatoire	3	1; 2; 3

#### Tableau ANOVA comprenant tous les termes

Source	DL	SomCar séq	SomCar ajust	CM ajust	Valeur F	Valeur de p
Operateur	2	4.500	4.500	2.250	0.16	0.856
Machine(Operateur)	9	128.250	128.250	14.250	3.17	0.012
Répétabilité	24	108.000	108.000	4.500		
Total	35	240.750				

$\alpha$  pour supprimer le terme d'interaction = 0.05

#### Composantes de la variance

Source	CompVar	% contribution (de CompVar)
R&R de l'instrumentation totale	4.50	58.06
Répétabilité	4.50	58.06
Reproductibilité	0.00	0.00
Operateur	0.00	0.00
De pièce à pièce	3.25	41.94
Machine(Operateur)	3.25	41.94
Variation totale	7.75	100.00

Donc outre le premier tableau qui nous indique bien que nous avons pris tous les facteurs comme étant aléatoires:

### Informations sur les facteurs

Facteur	Type	Niveaux	Valeurs
Machine(Operateur)	Aléatoire	12	1(1); 2(1); 3(1); 4(1); 1(2); 2(2); 3(2); 4(2); 1(3); 2(3); 3(3); 4(3)
Operateur	Aléatoire	3	1; 2; 3

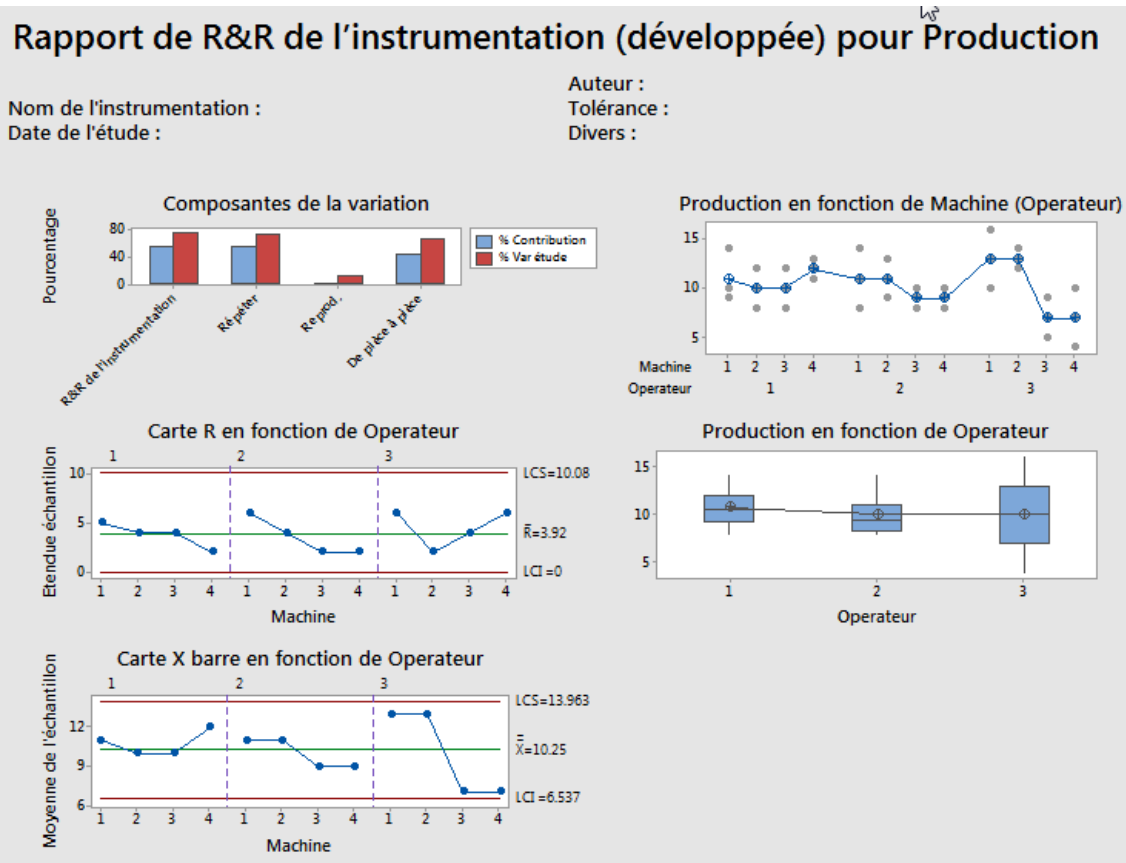
tout le reste de la fenêtre de session est parfaitement identique à l'étude R&R emboîtée que nous avons fait précédemment.

### R&R emboîtée avec facteurs fixes

Mais maintenant quid d'une étude R&R emboîtée avec les mêmes facteurs mais cette fois-ci... FIXES????!!!!

Eh bien il nous suffit de renseigner l'assistant comme suit:

Si nous validons, nous obtenons d'abord le graphique suivant:



Déjà nous pouvons constater avec le graphique ci-dessus et dans cet exemple particulier, que le seul graphique qui pourrait avoir une différence (*Composantes de la variation* en haut à gauche) est avec un niveau de zoom tel que nous ne pouvons pas distinguer la moindre variation (mais en réalité il y a une petite variation comme nous allons pouvoir le constater dans la fenêtre de session).

Maintenant, voyons la fenêtre de session:

Etude R&R de l'instrumentation : Production en fonction ... Operateur

Informations sur les facteurs

Facteur	Type	Niveaux	Valeurs
Machine(Operateur)	Fixe	12	1(1); 2(1); 3(1); 4(1); 1(2); 2(2); 3(2); 4(2); 1(3); 2(3); 3(3); 4(3)
Operateur	Fixe	3	1; 2; 3

Tableau ANOVA comprenant tous les termes

Source	DL	SomCar séq	SomCar ajust	CM ajust	Valeur F	Valeur de p
Machine(Operateur)	9	128.250	128.250	14.250	3.17	0.012
Operateur	2	4.500	4.500	2.250	0.50	0.613
Répétabilité	24	108.000	108.000	4.500		
Total	35	240.750				

$\alpha$  pour supprimer le terme d'interaction = 0.05

Composantes de la variance

Source	CompVar	% contribution (de CompVar)
R&R de l'instrumentation totale	4.6250	56.49
Répétabilité	4.5000	54.96
Reproductibilité	0.1250	1.53
Operateur	0.1250	1.53
De pièce à pièce	3.5625	43.51
Machine(Operateur)	3.5625	43.51
Variation totale	8.1875	100.00

**Evaluation de l'instrumentation**

Source	Ecart type	Variations d'étude (6 × écart type)	% var étude (%VE)
R&R de l'instrumentation totale	2.15058	12.9035	75.16
Répétabilité	2.12132	12.7279	74.14
Reproductibilité	0.35355	2.1213	12.36
Opérateur	0.35355	2.1213	12.36
De pièce à pièce	1.88746	11.3248	65.96
Machine(Opérateur)	1.88746	11.3248	65.96
Variation totale	2.86138	17.1683	100.00

Nombre de catégories distinctes = 1

Donc il est important ici de noter que nous sommes bien passés à un modèle emboîté à facteurs fixes:

**Informations sur les facteurs**

Facteur	Type	Niveaux	Valeurs
Machine(Opérateur)	Fixe	12	1(1); 2(1); 3(1); 4(1); 1(2); 2(2); 3(2); 4(2); 1(3); 2(3); 3(3); 4(3)
Opérateur	Fixe	3	1; 2; 3

Comparons maintenant les deux ANOVA! D'abord un rappel du tableau de l'ANOVA avec **facteurs aléatoires**:

**Tableau ANOVA comprenant tous les termes**

Source	DL	SomCar séq	SomCar ajust	CM ajust	Valeur F	Valeur de p
Opérateur	2	4.500	4.500	2.250	0.16	0.856
Machine(Opérateur)	9	128.250	128.250	14.250	3.17	0.012
Répétabilité	24	108.000	108.000	4.500		
Total	35	240.750				

$\alpha$  pour supprimer le terme d'interaction = 0.05

Et donc celui à **facteurs fixes**:

**Tableau ANOVA comprenant tous les termes**

Source	DL	SomCar séq	SomCar ajust	CM ajust	Valeur F	Valeur de p
Machine(Opérateur)	9	128.250	128.250	14.250	3.17	0.012
Opérateur	2	4.500	4.500	2.250	0.50	0.613
Répétabilité	24	108.000	108.000	4.500		
Total	35	240.750				

$\alpha$  pour supprimer le terme d'interaction = 0.05

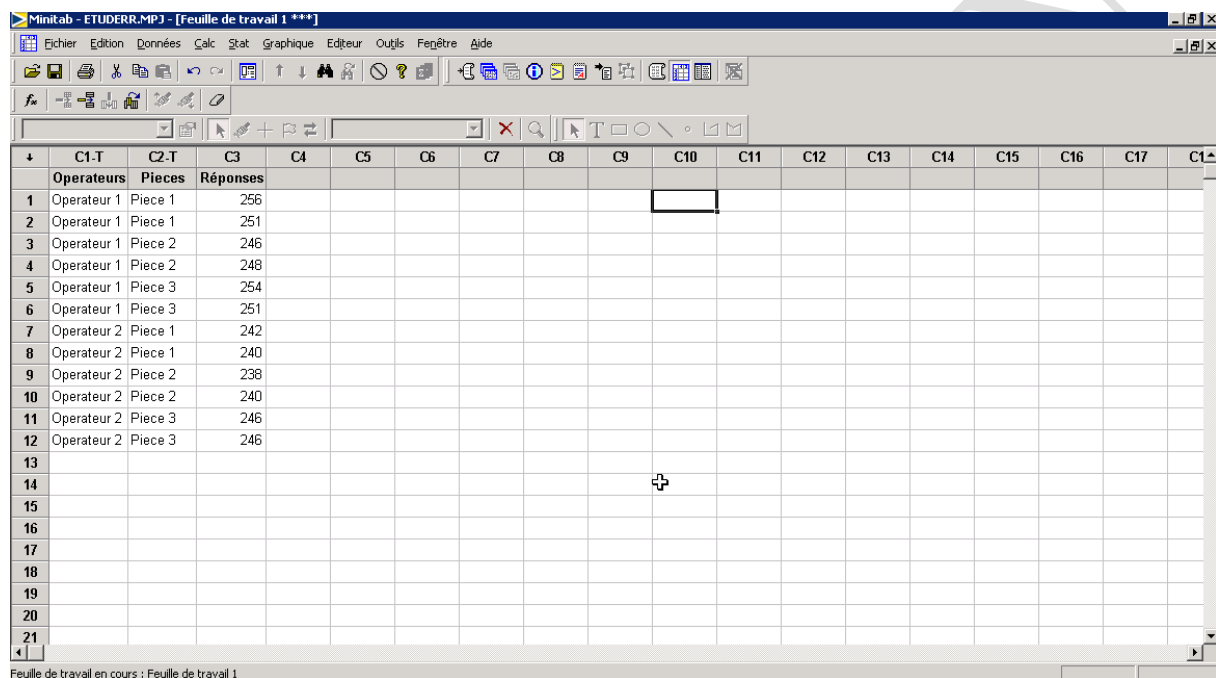
Nous remarquons donc qu'une des *p*-valeurs diffère un peu mais que les conclusions restent toutefois les mêmes.

## 15.14. Exercice 158.: Cartes d'essais de l'instrumentation

Minitab® Statistical Software 15.1.1.0

Toujours dans le cadre de l'étude R&R (reproductibilité/répétabilité), il existe un petit graphique très pratique visuellement qui s'appelle une *Carte d'essai de l'instrumentation* qui n'utilise aucune notation mathématique.

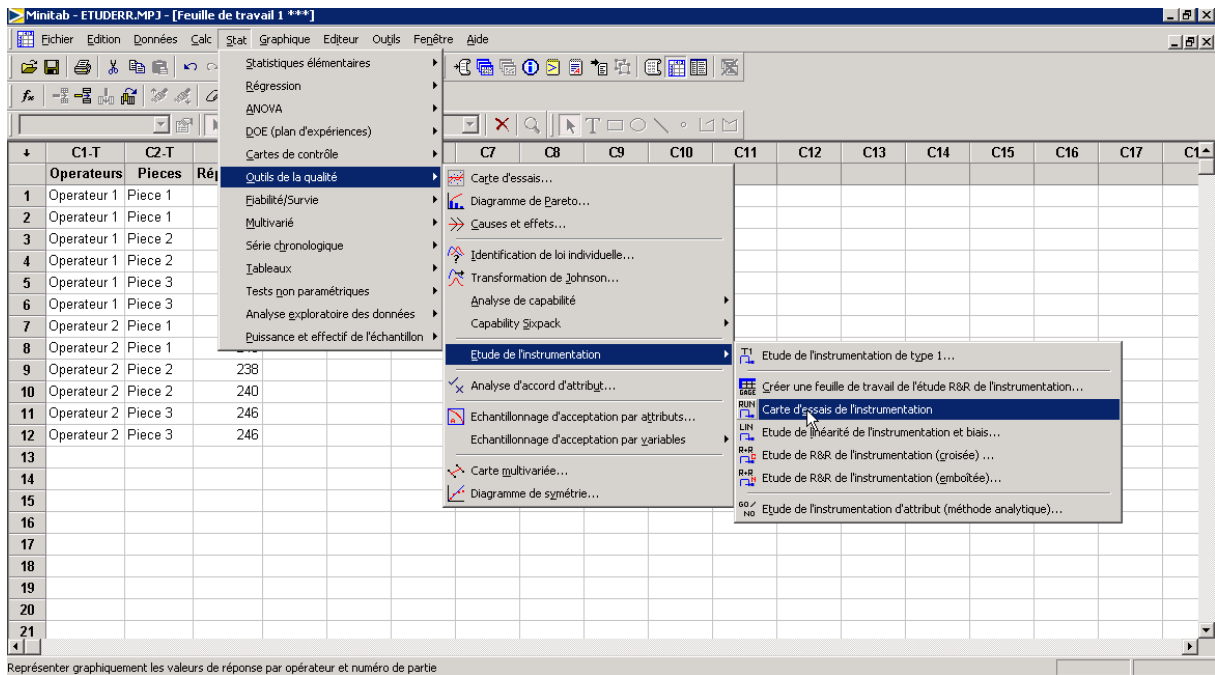
Pour voir cela, nous restons avec le fichier de l'exercice précédent (*EtudeRR.mpj* ou *EtudeRRContinue.mpj*):



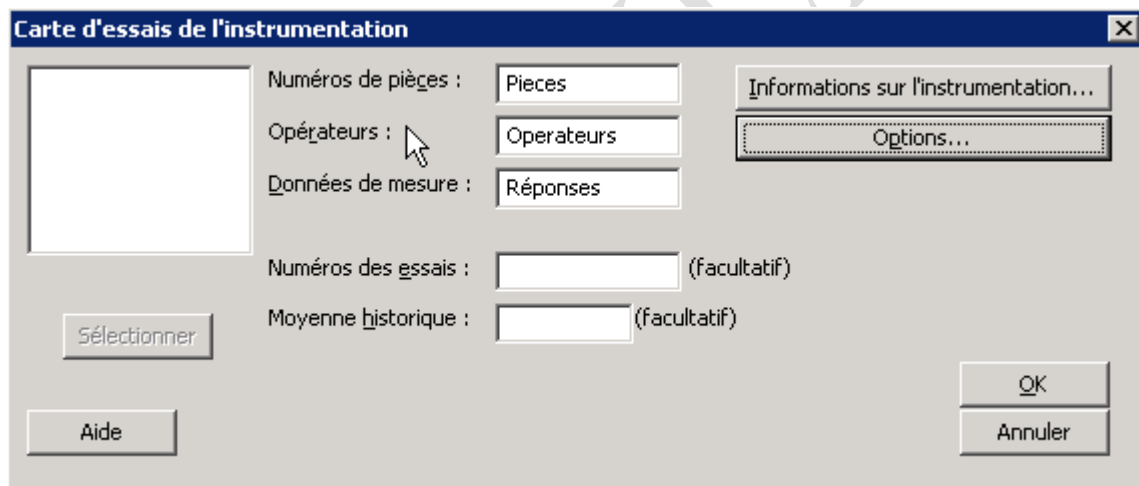
The screenshot shows the Minitab software interface with a data table. The table has columns labeled C1-T, C2-T, C3, C4, C5, C6, C7, C8, C9, C10, C11, C12, C13, C14, C15, C16, C17, and C18. The first three columns are 'Operateurs', 'Pièces', and 'Réponses'. The data is as follows:

	C1-T	C2-T	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18
	Operateurs	Pièces	Réponses															
1	Operateur 1	Piece 1	256															
2	Operateur 1	Piece 1	251															
3	Operateur 1	Piece 2	246															
4	Operateur 1	Piece 2	248															
5	Operateur 1	Piece 3	254															
6	Operateur 1	Piece 3	251															
7	Operateur 2	Piece 1	242															
8	Operateur 2	Piece 1	240															
9	Operateur 2	Piece 2	238															
10	Operateur 2	Piece 2	240															
11	Operateur 2	Piece 3	246															
12	Operateur 2	Piece 3	246															
13																		
14																		
15																		
16																		
17																		
18																		
19																		
20																		
21																		

Nous allons dans le menu **Stat/Outils de la qualité/Étude de l'instrumentation/Carte d'essai de l'instrumentation...**:



et nous avons alors:

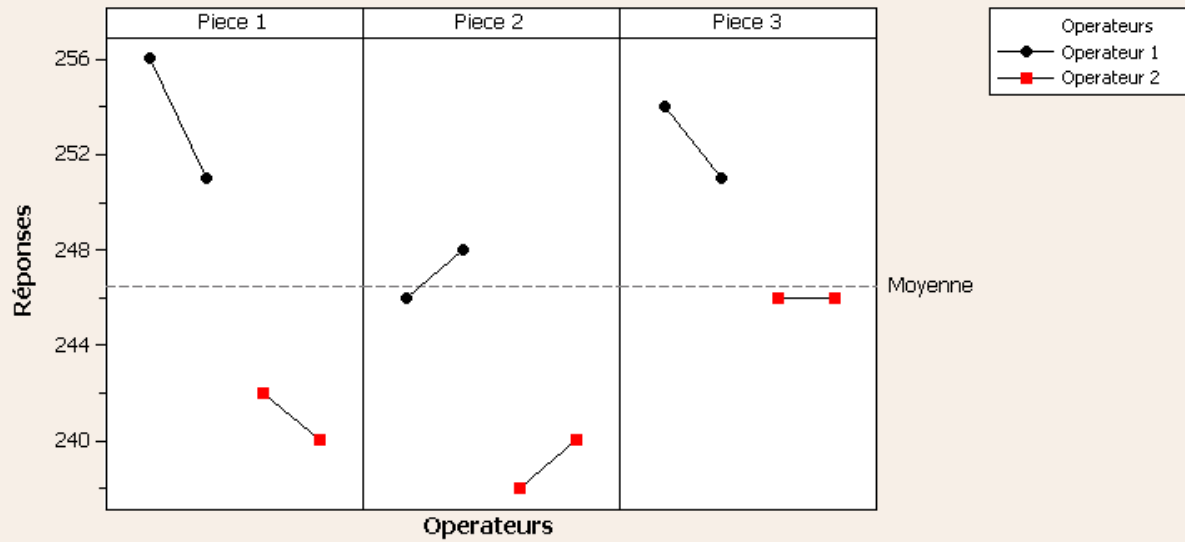


Nous cliquons sur **OK**:

**Carte d'essais de l'instrumentation de Réponses par Pièces, Operateurs**

Nom de l'instrumentation :  
 Date de l'étude :

Auteur :  
 Tolérance :  
 Mauv. class. :



Variable du panneau : Pieces

Comme vous pouvez le constater c'est basique. Minitab ne génère aucune statistique.

## 15.15. Exercice 159.: Étude R&R (reproductibilité/répétabilité) pour données par attributs

Minitab® Statistical Software 15.1.1.0

---

Les études R&R pour données par attribut comparent plusieurs décideurs (opérateurs) appelés à décider de l'acceptabilité de plusieurs services (produits ou investissements), à différents moments. Les résultats sont utilisés pour évaluer autant la reproductibilité (si les décideurs sont d'accord entre eux) que la répétabilité (si chaque décideur est systématiquement en accord avec ses propres résultats).

Pour plus d'exemples, le lecteur pourra se référer à la norme ISO/TR 14468:2010.

Bien évidemment, des biais peuvent surgir dans le système de décision. Donc même si tous sont unanimes, il se peut qu'ils aient tous fait le mauvais choix.

Vu que les évaluations par attribut sont souvent de nature subjective, il n'est toujours facile de déterminer quelle est la bonne décision.

Un point important dans ce type d'analyses, c'est d'avoir un nombre impair d'estimateurs (donneurs d'opinions) ainsi que pour chaque analyse un nombre impair d'opinions.

Raison pour laquelle dans le tableau ci-dessous, nous avons 3 estimateurs et 3 opinions pour chaque ligne.

Ouvrez donc le fichier *EtudeRRAttribut.mpj*:

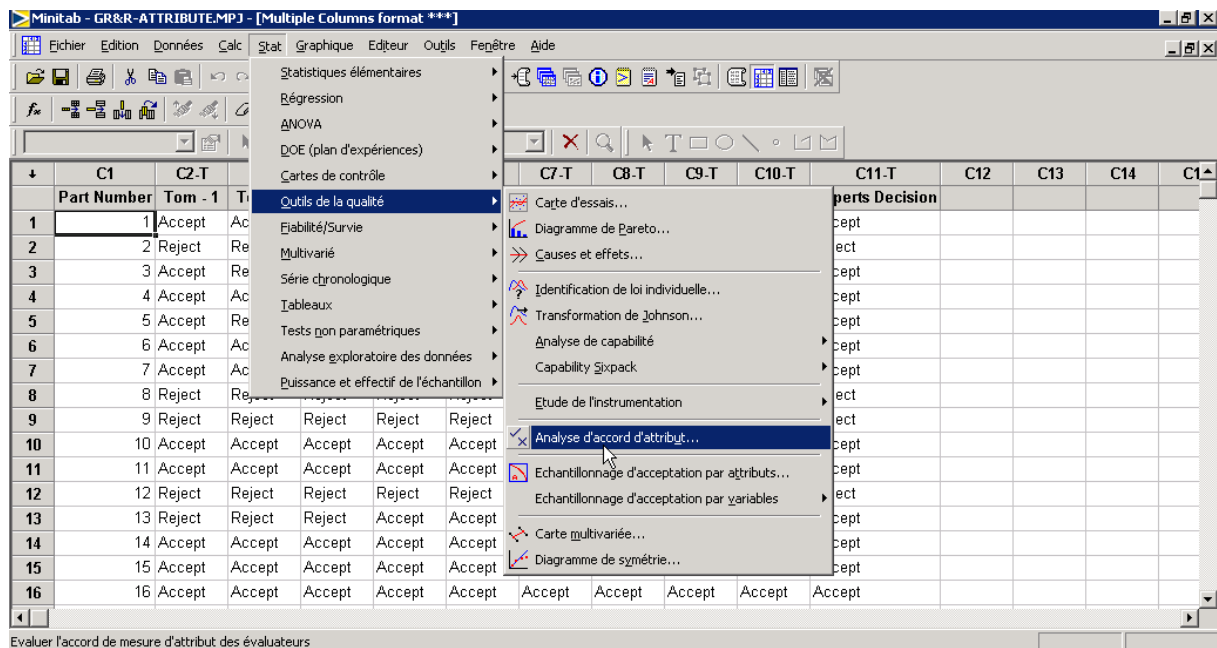


Minitab - GR&R-ATTRIBUTE.MPJ - [Multiple Columns format \*\*\*]

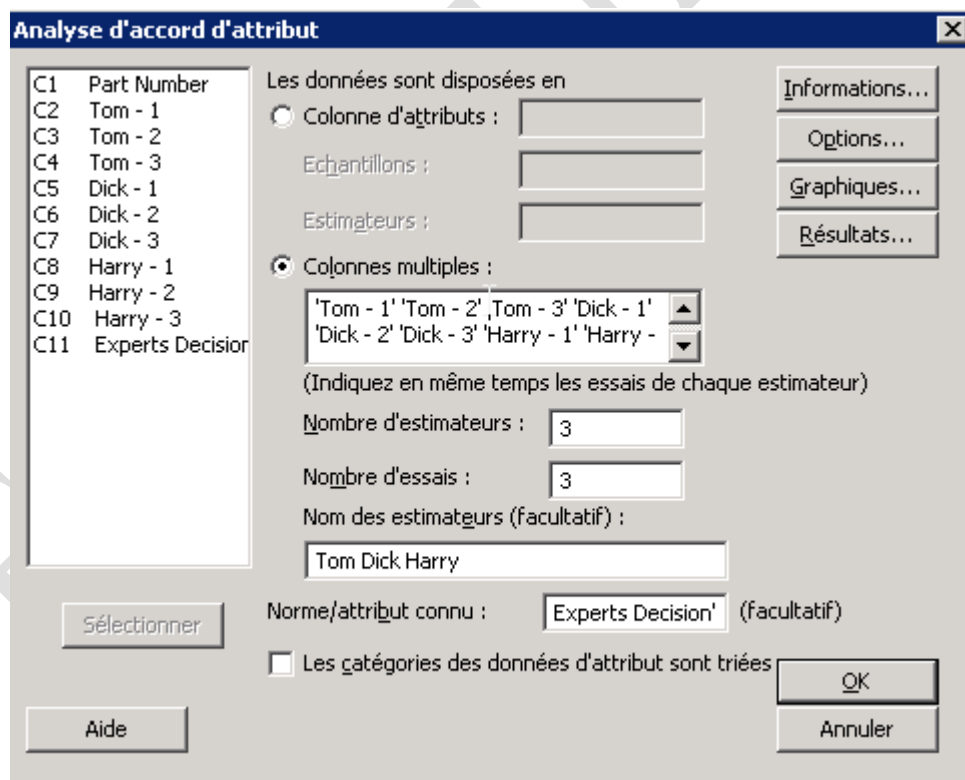
Fichier Edition Données Calc Stat Graphique Éditeur Outils Fenêtre Aide

	C1	C2-T	C3-T	C4-T	C5-T	C6-T	C7-T	C8-T	C9-T	C10-T	C11-T
	Part Number	Tom - 1	Tom - 2	Tom - 3	Dick - 1	Dick - 2	Dick - 3	Harry - 1	Harry - 2	Harry - 3	Experts Decision
1	1	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept
2	2	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject
3	3	Accept	Reject	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept
4	4	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept
5	5	Accept	Reject	Reject	Accept	Accept	Accept	Accept	Accept	Accept	Accept
6	6	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept
7	7	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept
8	8	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject
9	9	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject
10	10	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept
11	11	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept
12	12	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject
13	13	Reject	Reject	Reject	Accept	Accept	Accept	Accept	Accept	Accept	Accept
14	14	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept
15	15	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept
17	17	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject
18	18	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept
19	19	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept
20	20	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept
21	21	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject
22	22	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject
23	23	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject
24	24	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept
25	25	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Accept
26	26	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject	Reject
27	27	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept
28	28	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept
29	29	Reject	Reject	Reject	Accept	Accept	Accept	Accept	Reject	Reject	Accept
30	30	Reject	Accept	Reject	Accept	Accept	Accept	Accept	Accept	Accept	Accept

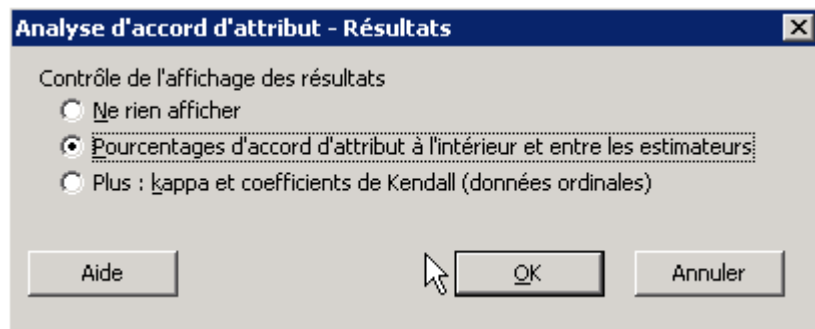
et nous allons dans le menu **Stat/Outils de la qualité/Analyse d'accord d'attribut...** (qui depuis Minitab 16 se nomme **Analyse de concordance...**):



Comme les d'accords sont représentés sur plusieurs colonnes, nous prenons *Colonnes multiples* (je préfère effectivement cette dernière configuration – plus visuelle - même si lorsqu'on travaille avec des SGBD, la structure en *Colonnes* d'attributs est plus facilement gérable):



Nous allons dans **Résultats...** pour désactiver certaines statistiques empiriques:



Nous validons par **OK**:

**Analyse d'accord d'attribut de Tom - 1; Tom - 2; Tom - 3; Dick - 1; ...**

**Entre estimateurs**

Accord d'évaluation

Estimateur	# Inspection	# Correspondance	Pourcentage	IC à 95 %
Tom	30	27	90.00	(73.47; 97.89)
Dick	30	30	100.00	(90.50; 100.00)
Harry	30	29	96.67	(82.78; 99.92)

# Correspondance : l'estimateur est conforme avec lui-même après les tests.

**Chaque estimateur en fonction du standard**

Accord d'évaluation

Estimateur	# Inspection	# Correspondance	Pourcentage	IC à 95 %
Tom	30	24	80.00	(61.43; 92.29)
Dick	30	29	96.67	(82.78; 99.92)
Harry	30	28	93.33	(77.93; 99.18)

# Correspondance : les évaluations de tous les estimateurs, après essais, concordent avec le standard connu.

Non-accord d'attribut

Estimateur	# Reject /		# Accept /		# Mixte
	Accept	Pourcentage	Reject	Pourcentage	
Tom	3	14.29	0	0.00	3
Dick	1	4.76	0	0.00	0
Harry	1	4.76	0	0.00	1

Estimateur	Pourcentage
Tom	10.00
Dick	0.00
Harry	3.33

# Reject/Accept : évaluations réalisées sur les essais = Reject/ standard = Accept.

# Accept/Reject : évaluations réalisées sur les essais = Accept/ standard = Reject.

# Mixte : les évaluations réalisées sur les essais ne sont pas identiques.

### Entre estimateurs

Accord d'évaluation

# Inspection	# Correspondance	Pourcentage	IC à 95 %
30	25	83.33	(65.28; 94.36)

# Correspondance : les évaluations de tous les estimateurs concordent entre elles.

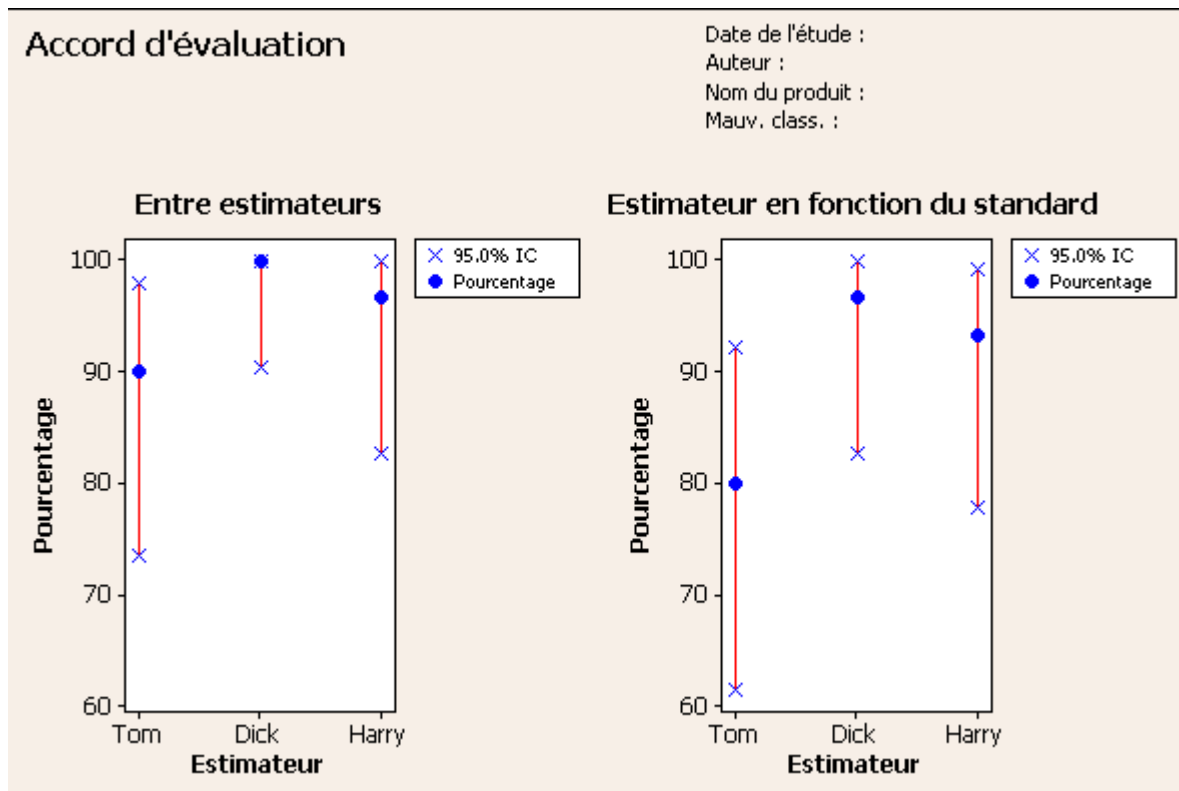
### Tous les estimateurs en fonction du standard

Accord d'évaluation

# Inspection	# Correspondance	Pourcentage	IC à 95 %
30	24	80.00	(61.43; 92.29)

# Correspondance : les évaluations de tous les estimateurs concordent avec le standard connu.

avec le graphique:



Nous souhaitons pouvoir déterminer tous les calculs à la main pour en comprendre le mécanisme. Voyons donc cela... petit à petit. Commençons par le premier tableau qui analyse les accords **Entre estimateurs**:

**Entre estimateurs**

Accord d'évaluation

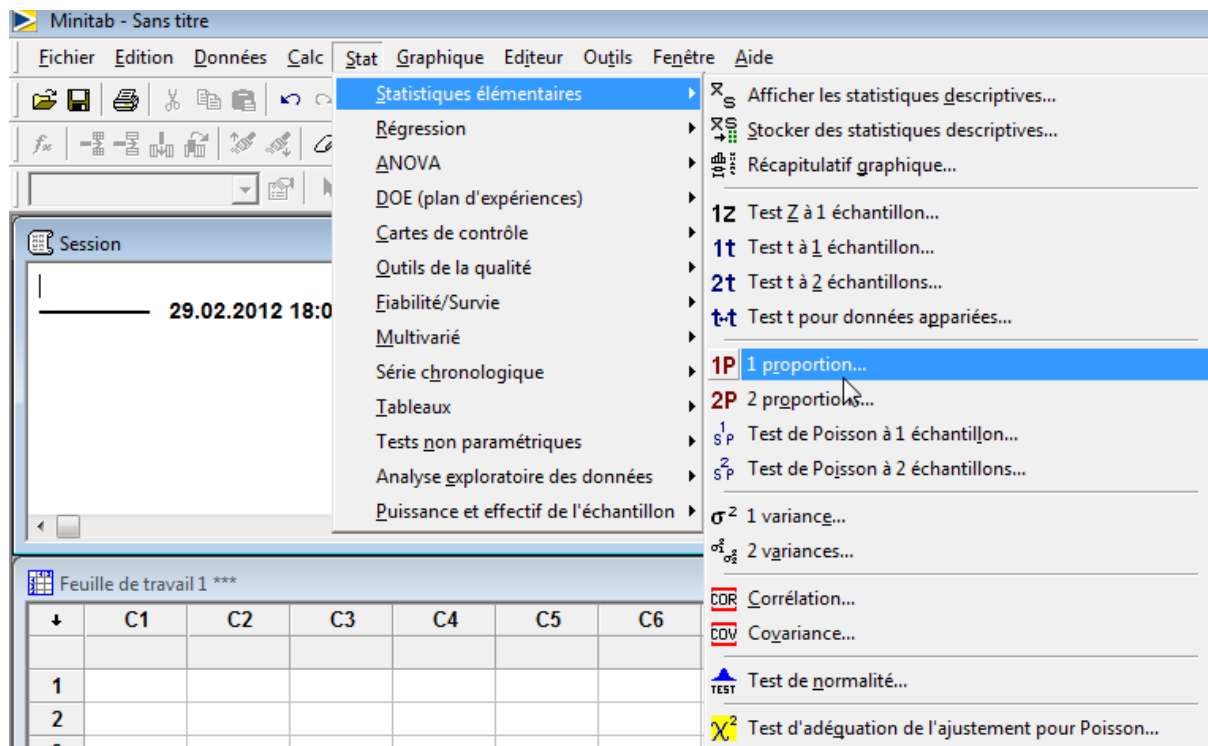
Estimateur	# Inspection	# Correspondance	Pourcentage	IC à 95 %
Tom	30	27	90.00	(73.47; 97.89)
Dick	30	30	100.00	(90.50; 100.00)
Harry	30	29	96.67	(82.78; 99.92)

# Correspondance : l'estimateur est conforme avec lui-même après les tests.

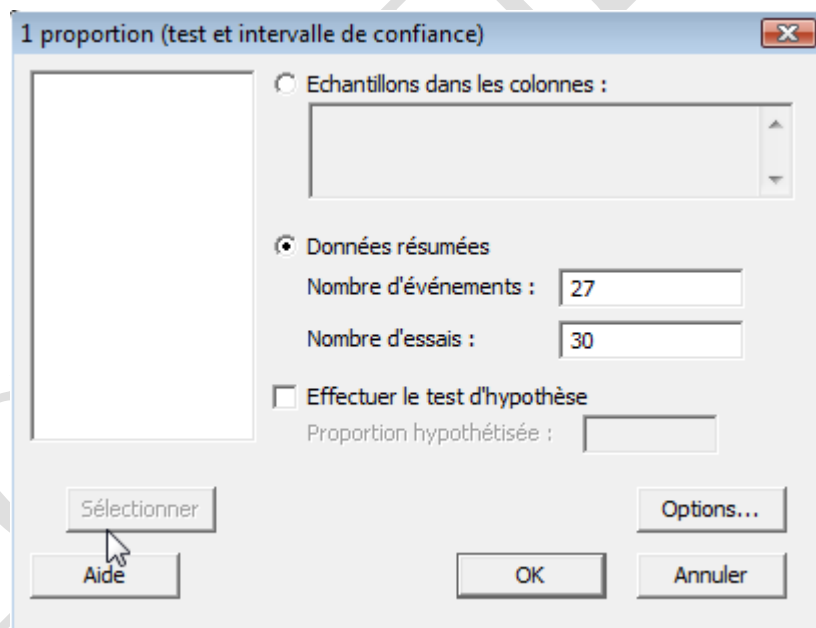
La colonne *# Inspection* est évidente... il s'agit du nombre d'échantillons analysés (pièces, projets, décisions ou autres...).

La colonne *# Correspondance* est évidente aussi. Il s'agit pour chaque donneur d'opinion (estimateur) le nombre d'échantillons analysés pour lesquels l'avis général de chaque estimateur était identique à l'avis général des autres estimateurs (on ne prend pas du tout en compte l'avis de l'expert!). Pas besoin de calculs, il suffit de prendre 60 secondes pour vérifier ligne par ligne dans le tableau d'origine et compter. Je laisse le soin au lecteur de faire cela.

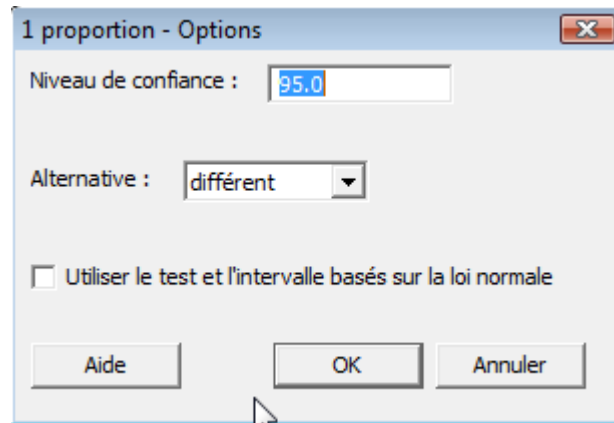
La colonne *Pourcentage* est simple le rapport entre *# Correspondance* et *# Inspection*. L'intervalle de confiance est lui simplement calculé à partir du test binomial exact. Mais comme c'est long à faire, montrons que nous pouvons le réobtenir en faisant à nouveau (nous avons déjà présenté cet outil plus haut mais avec l'approximation en loi Normale):



ce qui nous amène à:



mais dans les bouton Options... nous ne cochons pas cette fois-ci **Utiliser le test et l'intervalle** basés sur la loi normale.



et si nous validons nous obtenons:

### Test et IC pour une proportion

Echantillon	X	N	P échantillon	IC à 95 %
1	27	30	0.900000	(0.734712; 0.978883)

donc ce qui correspond bien à 73.47% et 97.89%.

Passons au deuxième tableau qui analyse la non concordance d'attributs **entre les estimateurs et l'expert**:

```
# Correspondance : les évaluations de tous les estimateurs, après essais,
concordent avec le standard connu.
```

```
Non-accord d'attribut
```

Estimateur	# Reject / Accept	Pourcentage	# Accept / Reject	Pourcentage	# Mixte
Tom	3 / 0	14.29	0 / 0	0.00	3
Dick	1 / 0	4.76	0 / 0	0.00	0
Harry	1 / 0	4.76	0 / 0	0.00	1

Estimateur	Pourcentage
Tom	10.00
Dick	0.00
Harry	3.33

```
# Reject/Accept : évaluations réalisées sur les essais = Reject/ standard =
Accept.
```

```
# Accept/Reject : évaluations réalisées sur les essais = Accept/ standard =
Reject.
```

```
# Mixte : les évaluations réalisées sur les essais ne sont pas identiques.
```

La colonne *#Reject/Accept* avec les valeurs (3,1,1) donne simplement le nombre de fois que chaque estimateur rejette alors que l'expert accepte. Pour vérifier, il suffit de compter dans la table d'origine.

La première colonne *Pourcentage* avec les valeurs (14.29%, 4.76%,4.76%) est simplement le rapport entre la colonne *#Reject/Accept* et le nombre de lignes pour lesquelles l'expert a

accepté (dans le cas présent, l'expert a accepté 21 fois sur 30). Donc il s'agit de (3/21,1/21,1/21)

La colonne *#Accept/Reject* avec les valeurs (0,0,0) donne simplement le nombre de fois que chaque estimateur accepte alors que l'expert rejette. Pour vérifier, il suffit de compter dans la table d'origine.

La deuxième colonne *Pourcentage* avec les valeurs (0 %, 0%,0 %) est simplement le rapport entre la colonne *# Accept/Reject* et le nombre de lignes pour lesquelles l'expert a accepté (dans le cas présent, l'expert a rejeté 9 fois sur 30). Donc il s'agit de (0/9,0/9,0/9)

La colonne *#Mixte* avec les valeurs (3,0,1) donne simplement le nombre de fois que chaque estimateur a un mélange d'opinions (les opinions unilatéralement contraires à l'experte ne sont donc pas comptabilisées) et ce quel que soit la conclusion de l'expert.

La dernière colonne *Pourcentage* avec les valeurs (10%,0%,3.33%) donne simplement le rapport en pourcentage entre la colonne *#Mixte* et le nombre de lignes de la table (soit une division par 30).

Le prochain tableau est donc:

```

Entre estimateurs
Accord d'évaluation
# Inspection # Correspondance Pourcentage IC à 95 %
          30          25          83.33 (65.28; 94.36)
# Correspondance : les évaluations de tous les estimateurs concordent entre
elles.

```

Il s'agit simplement d'une analyse globale non nominative entre estimateurs et toujours avec un test binomial exact de la proportion.

Effectivement, la colonne *#Inspection* donne le nombre de lignes du tableau, la colonne *#Correspondance* donne le nombre de lignes où les estimateurs étaient parfaitement d'accord entre eux (donc les 9 opinions identiques). La colonne pourcentage donne simplement le rapport 25/30.

Le dernier tableau:

```

Tous les estimateurs en fonction du standard
Accord d'évaluation
# Inspection # Correspondance Pourcentage IC à 95 %
          30          24          80.00 (61.43; 92.29)
# Correspondance : les évaluations de tous les estimateurs concordent avec le
standard connu.

```

est identique au précédent à la différence que c'est la comparaisons par rapport à l'expert.



## 15.16. Exercice 160.: Kappa de Cohen (analyse d'accord d'attribut)

Minitab® Statistical Software 15.1.1.0

Encore une fois, le but est d'appliquer les calculs et démonstrations faits à la main pendant le cours théorique de méthodes numériques et de vérifier que nous retrouvons les mêmes valeurs avec Minitab.

Nous partons donc fichier *KappaFleiss\_Cohen.mpj*:

	C1-T	C2	C3	C4	C5	C6	C7	C8	C9
	Estimateur	Echantillon	Attribut						
1	Alice	1	0						
2	Alice	2	1						
3	Alice	3	0						
4	Alice	4	0						
5	Alice	5	1						
6	Alice	6	0						
7	Alice	7	1						
8	Alice	8	1						
9	Alice	9	0						
10	Alice	10	0						
11	Alice	11	1						
12	Bob	1	0						
13	Bob	2	1						
14	Bob	3	1						
15	Bob	4	0						
16	Bob	5	0						
17	Bob	6	0						
18	Bob	7	1						
19	Bob	8	0						
20	Bob	9	0						
21	Bob	10	0						
22	Bob	11	1						
23									

Nous allons dans le menu **Stat/Outils de la qualité/Analyse d'accord d'attribut...**:

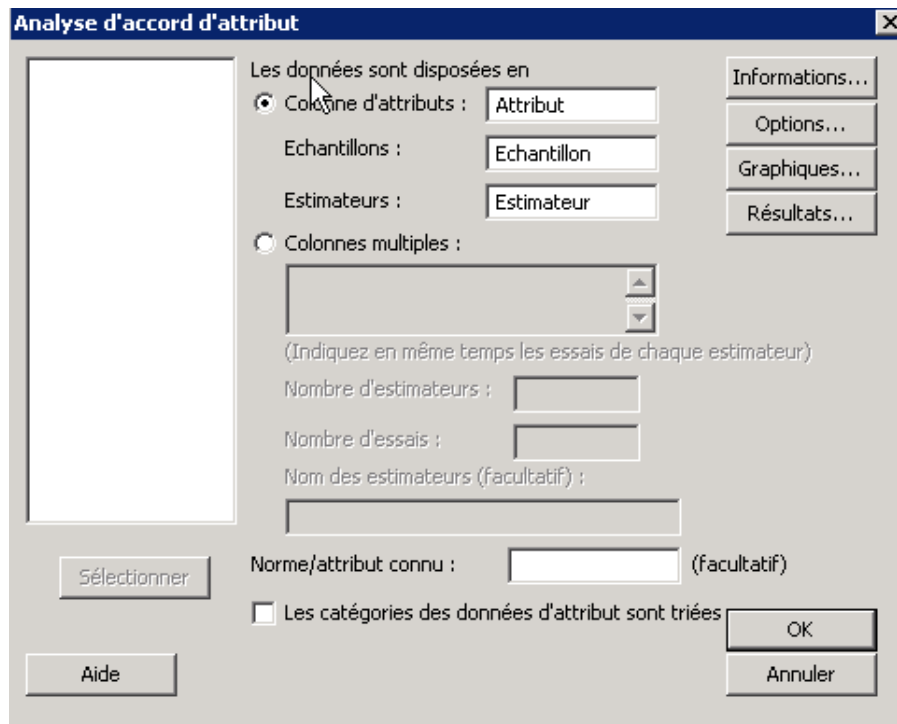
Minitab - KappaFleiss\_Cohen.MPJ - [Feuille de travail 1 \*\*\*]

Stat > Outils de la qualité > Analyse d'accord d'attribut...

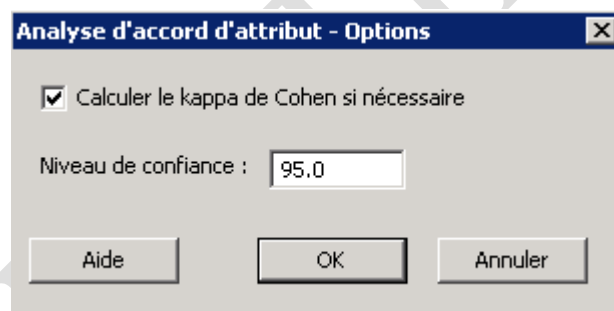
	C1-T	C2	A
	Estimateur	Echantillon	
1	Alice	1	
2	Alice	2	
3	Alice	3	
4	Alice	4	
5	Alice	5	
6	Alice	6	
7	Alice	7	
8	Alice	8	
9	Alice	9	
10	Alice	10	
11	Alice	11	1
12	Bob	1	0
13	Bob	2	1
14	Bob	3	1
15	Bob	4	0
16	Bob	5	0
17	Bob	6	0
18	Bob	7	1
19	Bob	8	0
20	Bob	9	0
21	Bob	10	0
22	Bob	11	1
23			

Evaluer l'accord de mesure d'attribut des évaluateurs

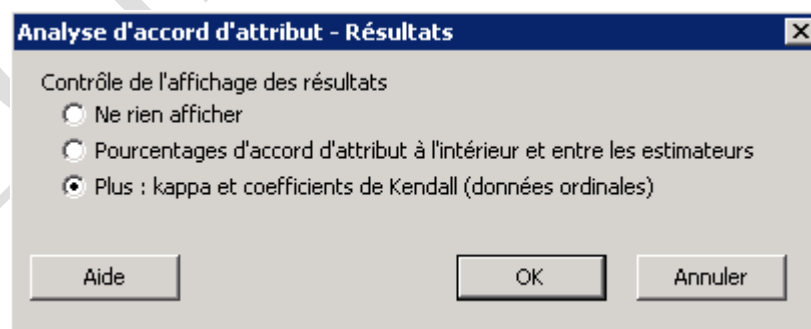
Et nous prenons:



Nous cliquons sur le bouton **Options...** pour demander à Minitab de nous calculer aussi le Kappa de Cohen. Nous avons alors:



Nous cliquons ensuite sur le bouton **Résultats...** pour s'assurer d'avoir bien:



et nous validons par **OK** autant de fois que nécessaire. Nous obtenons alors:

Minitab - KappaFleiss\_Cohen.MPJ - [Session]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

**Entre estimateurs**

Accord d'évaluation

# Inspection	# Correspondance	Pourcentage	IC à 95 %
11	8	72.73	(39.03; 93.98)

# Correspondance : les évaluations de tous les estimateurs concordent entre elles.

Statistiques Kappa de Fleiss

Réponse	Kappa	Kappa ErT	Z	P (contre valeur supérieure à 0)
0	0.435897	0.301511	1.44571	0.0741
1	0.435897	0.301511	1.44571	0.0741

Statistiques kappa de Cohen

Réponse	Kappa	Kappa ErT	Z	P (contre valeur supérieure à 0)
0	0.440678	0.296225	1.48765	0.0684
1	0.440678	0.296225	1.48765	0.0684

\* REMARQUE \* Essai unique pour chaque estimateur. Aucun pourcentage d'accord

Bienvenue dans Minitab, appuyez sur F1 pour obtenir de l'aide. Modifiable

Soit les mêmes résultats que lors des calculs à la main effectués dans le cours de méthodes numériques lors de la démonstration mathématique en ce qui concerne le kappa de Cohen mis à évidence en rouge ci-dessus (comme le kappa de Fleiss n'est pas dans le cours théorique, nous ferons comme s'il n'existait pas...).

## 15.17. Exercice 161.: Plan d'échantillonnage par mesures

Minitab® Statistical Software 15.1.1.0

---

Dans le cadre de notre exemple fait dans le cours de statistique - le producteur de boissons gazeuses - qui reçoit des lots de 10'000 bouteilles dont la résistance à la pression (LSL) doit être supérieure à 22.5 avec un écart-type sous contrôle statistique et valant 1.5, le fournisseur et le client ont chacun fixé les risques qu'ils sont prêts à courir.

Ainsi, nous avons que le fournisseur souhaite une probabilité cumulée (risque)  $\alpha$  de 5% que le client rejette à tort le lot avec moins de 5% ( $D$ ) de non-conformes. De son côté le client souhaite une probabilité cumulée (risque)  $\beta$  de 10% d'accepter à tort un lot avec plus de 5% ( $D$ ) de non-conformes.

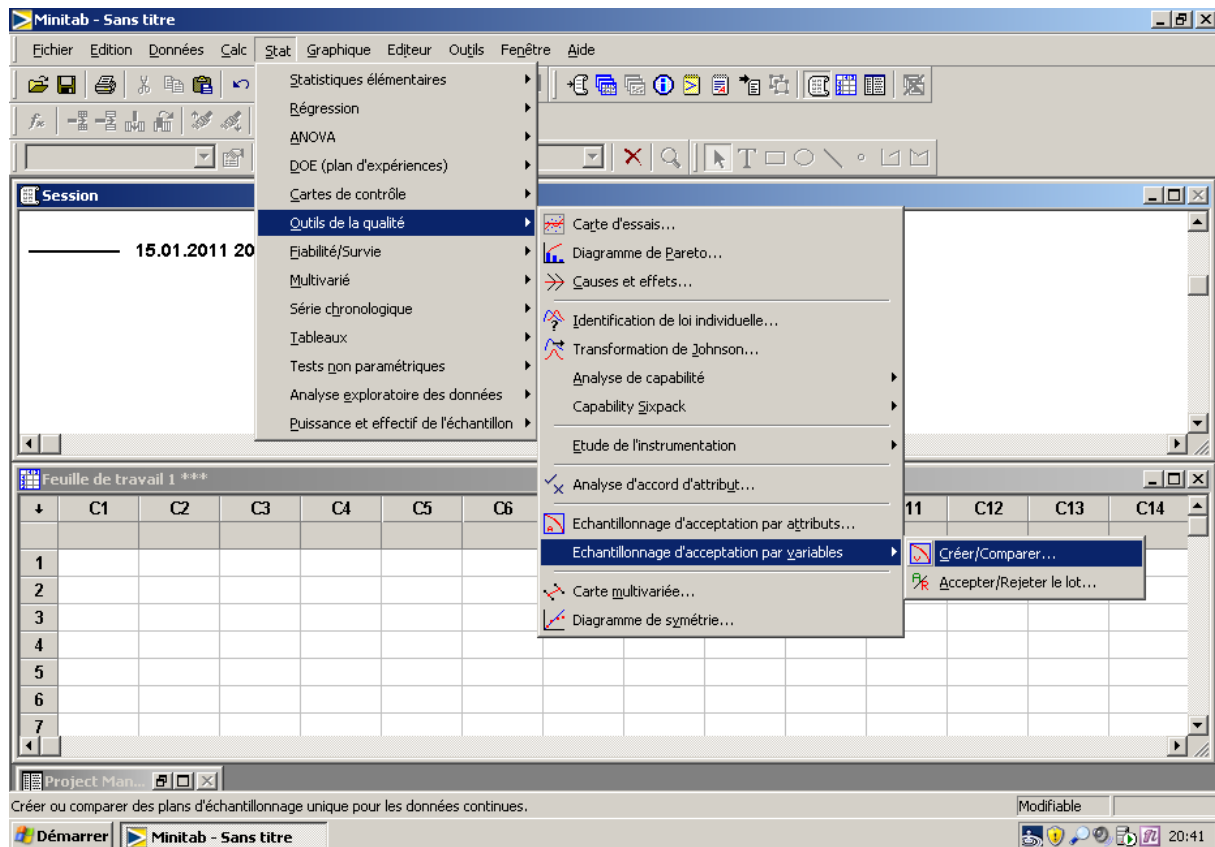
$$P(\text{Refuser un lot} | D \leq D_f) \leq \alpha \Rightarrow P(\text{Refuser un lot} | D \leq 1\%) \leq 5\%$$

$$P(\text{Accepter un lot} | D \geq D_c) \leq \beta \Rightarrow P(\text{Accepter un lot} | D \geq 5\%) \leq 10\%$$

Nous devons alors en utilisant cette norme définir le N.Q.A. Comme mentionné dans la définition, en général le N.Q.A. est proche de  $D_f$  (mais pas égal!). Donc le N.Q.A. sera posé comme valant 1%...

Nous souhaitons comparer les valeurs données par Minitab® Statistical Software et celles obtenues dans le cours statistique théorique basées sur la norme AF-X06-023.

Dans Minitab® Statistical Software nous allons dans le menu **Stat/Outils de la qualité/Echantillonnage d'acceptation par variables/Créer/comparer...**:



et vient la fenêtre suivante à remplir comme indiqué:

**Echantillonnage d'acceptation par variables (Créer/Comparer)**

Créer un plan d'échantillonnage

Options...  
Graphiques...

Unités des niveaux de qualité : Pourcentage de défaillances

Niveau de qualité acceptable (AQL) : 1

Niveau de qualité rejetaable (RQL ou LTPD) : 5

Risque pour le fournisseur (Alpha) : 0.05

Risque pour l'acheteur (Bêta) : 0.10

Spécification inférieure : 22.5

Spécification supérieure :

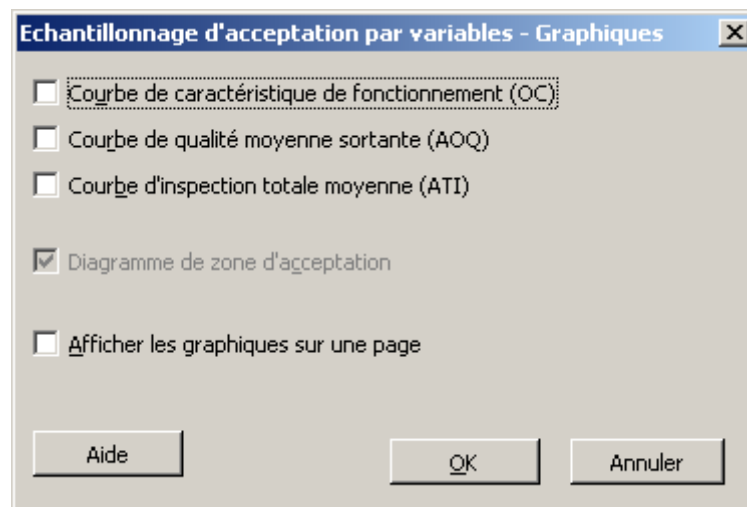
Ecart type historique : 1.5 (facultatif)

Taille du lot : 10000

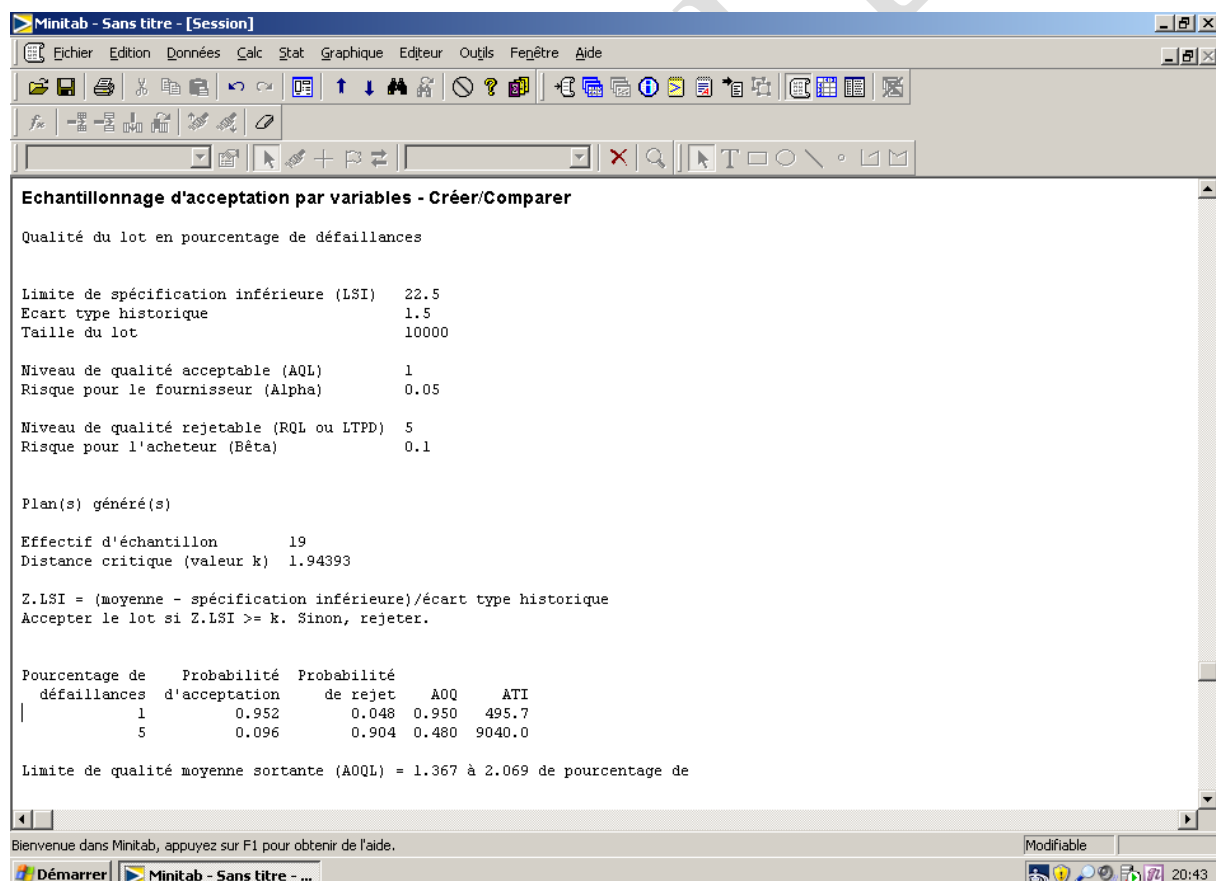
Aide

OK  
Annuler

Nous allons dans le bouton **Graphiques...**:



nous décochons tout pour la simple raison que nous n'en avons pas fait les démonstrations dans le cours de statistique théorique. Nous validons alors deux fois par **OK** pour obtenir dans la fenêtre de session les informations suivantes:



Ce qui correspond assez bien à ce que nous avons obtenu dans le cours de statistique théorique où l'effectif de l'échantillon était de 18 pour les calculs à la main et 25 selon la norme (Minitab® Statistical Software donne donc 19) et une distance critique de 1.943 pour les calculs à la main et 1.97 selon la norme (Minitab® Statistical Software donne donc 1.94393).

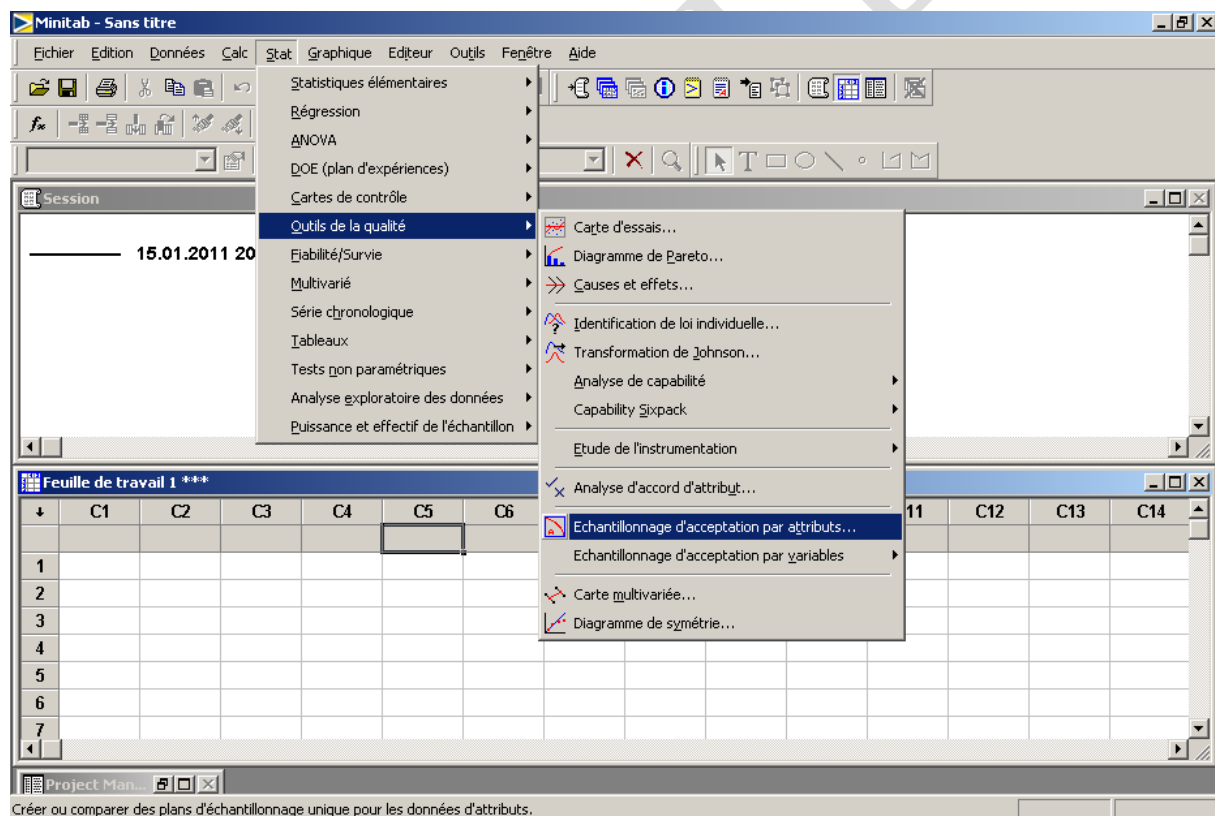
## 15.18. Exercice 162.: Plan d'échantillonnage par attribut

Minitab® Statistical Software 15.1.1.0

Un lot de bouteilles est livré sous forme de lots de 10'000 unités correspondant à  $N$ . Nous cherchons à mettre en place un plan de contrôle de la réception par attributs avec la probabilité cumulée (risque)  $\alpha$  de 1% que le client rejette à tort le lot avec moins de 2.5% ( $P_\alpha$ ) de non-conformes. De son côté le client souhaite une probabilité cumulée (risque)  $\beta$  de 10% d'accepter à tort un lot avec plus de 5% ( $P_\beta$ ) de non-conformes.

Le but est de déterminer donc la valeur de  $A$  (nombre de non-conformes limite acceptable) et de  $n$  (taille de l'échantillon à prélever) et de la comparer aux valeurs obtenues à l'aide nomographe dans le cours de statistique théorique.

Dans Minitab® Statistical Software nous allons dans le menu **Stat/Outils de la qualité/Echantillonnage d'acceptation par attributs....**



Nous avons alors la fenêtre suivante:



Echantillonnage d'acceptation par attributs

Créer un plan d'échantillonnage

Type de mesure : Conforme/Non-conforme

Unités des niveaux de qualité : Pourcentage de défaillances

Niveau de qualité acceptable (AQL) : 2.5

Niveau de qualité rejetable (RQL ou LTPD) : 5

Risque pour le fournisseur (Alpha) : 0.01

Risque pour l'acheteur (Bêta) : 0.10

Taille du lot : 10000

Aide Options... Graphiques... OK Annuler

Si la taille du lot était petite, nous pourrions aller dans le bouton **Options** et y cocher **Utiliser la loi hypergéométrique pour le lot isolé**:

Echantillonnage d'acceptation par attributs - Options

Utiliser la loi hypergéométrique pour le lot isolé

Entrer des niveaux de qualité supplémentaires pour calculer les probabilités d'acceptation :

(Unités : Pourcentage de défaillances)

Vous pouvez augmenter légèrement les valeurs alpha et bêta pour les plans alternatifs avec effectifs d'échantillons inférieurs

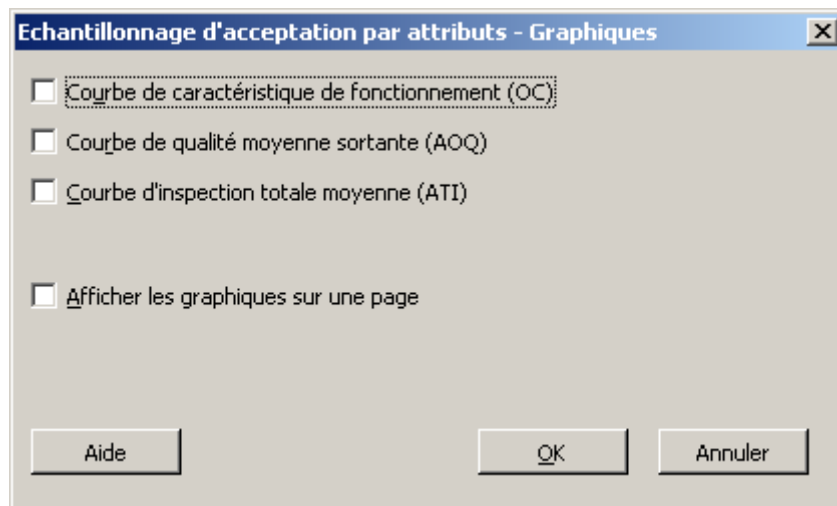
Valeur alpha maximale autorisée :

Valeur bêta maximale autorisée :

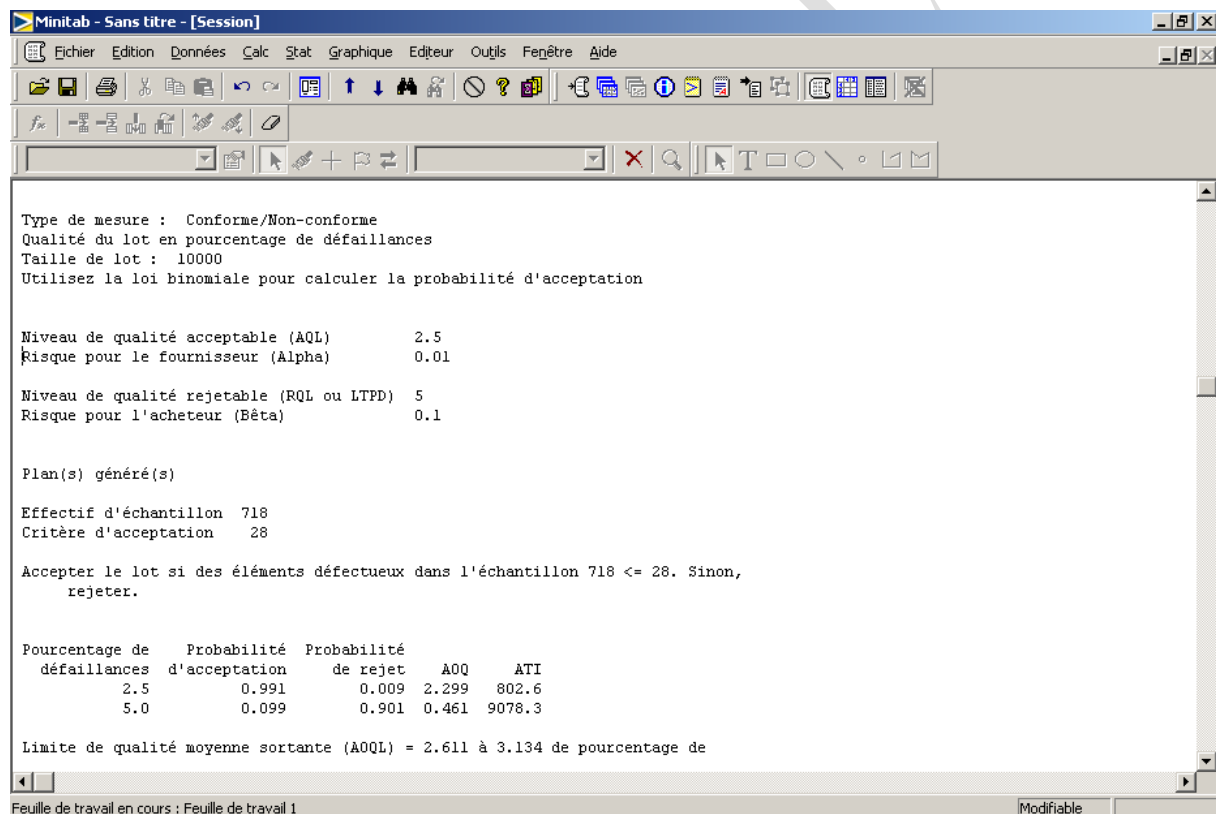
Aide OK Annuler

mais ici le lot est très grand nous pouvons donc nous en passer.

Dans le bouton **Graphiques...**:



nous décochons tout pour la simple raison que nous n'en avons pas fait les démonstrations dans le cours de statistique théorique. Nous validons alors deux fois par **OK** pour obtenir dans la fenêtre de session les informations suivantes:



Ce qui correspond assez bien à ce que nous avons obtenu dans le cours de statistique théorique où l'effectif de l'échantillon était de 700 (Minitab® Statistical Software donne donc 718) et un critère d'acceptation de 30 (Minitab® Statistical Software donne donc 28).

## Plan de contrôle de tolérance 0 (C=0/"Accept on None")

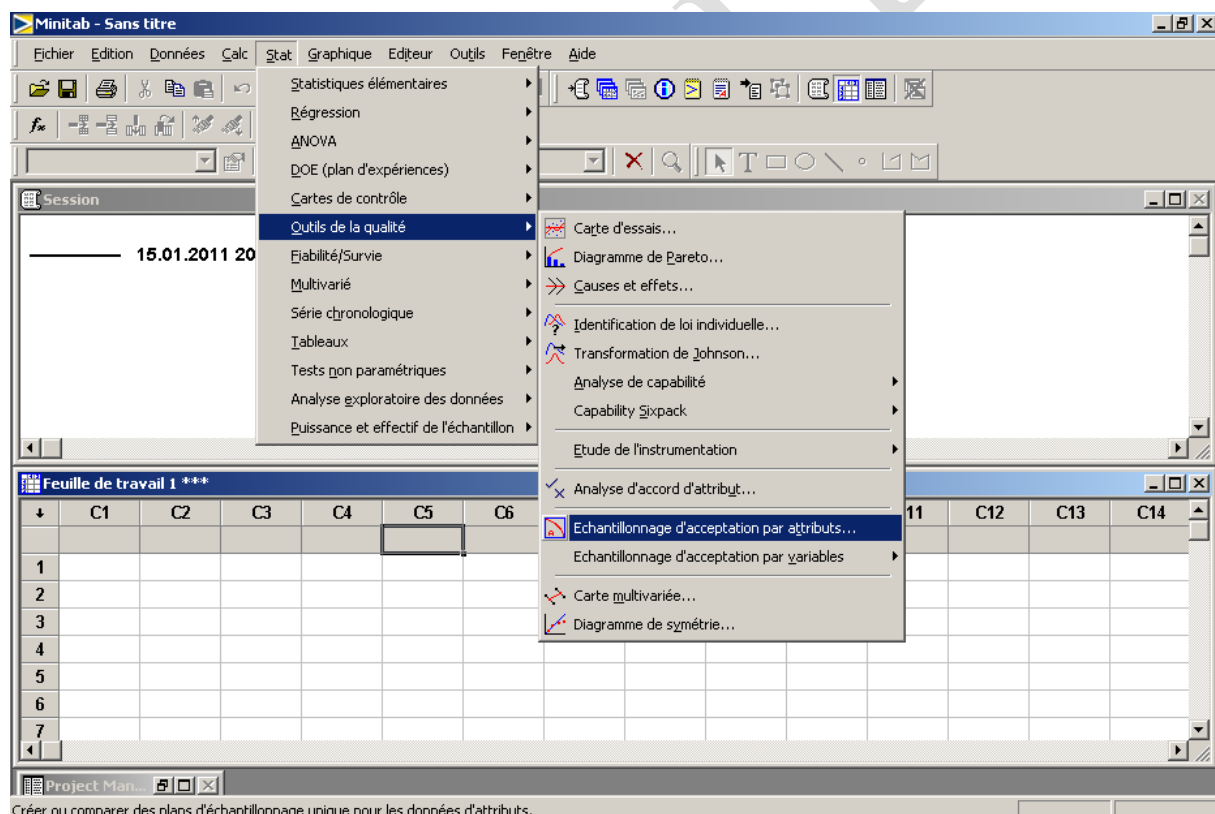
Certaines organisations sont tenues de rejeter un lot entier lorsqu'un seul défaut est détecté, par exemple dans le cadre d'un programme à zéro défectueux. Dans ces cas de figure, l'entreprise utilise des plans  $C = 0$ , aussi appelés "plans de tolérance 0", qui sont conçus pour avoir un critère d'acceptation (C) de 0.

Vous pouvez utiliser un plan  $C=0$  lorsque les conditions suivantes sont réunies :

- Vous avez défini un NQR (niveau de qualité acceptable) et une valeur bêta, ou ceux-ci sont implicites.
- Vous n'utilisez pas de NQA (niveau de qualité acceptable).

Si la qualité entrante est excellente, un plan  $C = 0$  peut vous aider à réduire les quantités d'inspection. Cela dit, si la qualité entrante n'est pas pratiquement parfaite, vous rejetterez plus de lots et contrôlerez probablement plus d'articles.

Dans Minitab® Statistical Software, pour construire un tel plan, nous allons dans le menu **Stat/Outils de la qualité/Echantillonnage d'acceptation par attributs...**:



Nous avons alors la fenêtre suivante où il faut mettre les valeurs comme indiquées (presque impossible à deviner si on n'est pas les concepteurs du logiciel Minitab!):

Echantillonnage d'acceptation par attributs

Créer un plan d'échantillonnage

Type de mesure : Conforme/Non-conforme

Unités des niveaux de qualité : Pourcentage de défaillances

Niveau de qualité acceptable (AQL) : 0,05

Niveau de qualité rejetable (RQL ou LTPD) : 0,5

Risque pour le fournisseur (Alpha) : 0,89

Risque pour l'acheteur (Bêta) : 0,1

Taille du lot : 200

Aide

Options...

Graphiques...

OK

Annuler

Dans le bouton **Options** on cochera **Utiliser la loi hypergéométrique pour le lot isolé**:

Echantillonnage d'acceptation par attributs - Options

Utiliser la loi hypergéométrique pour le lot isolé

Entrer des niveaux de qualité supplémentaires pour calculer les probabilités d'acceptation :

(Unités : Pourcentage de défaillances)

Vous pouvez augmenter légèrement les valeurs alpha et bêta pour les plans alternatifs avec effectifs d'échantillons inférieurs

Valeur alpha maximale autorisée :

Valeur bêta maximale autorisée :

Aide

OK

Annuler

Ce qui donne dans la fenêtre de session:

**Echantillonnage d'acceptation par attributs**

Type de mesure : Conforme/Non-conforme  
 Qualité du lot en pourcentage de défaillances  
 Taille de lot : 200  
 Utilisez la loi binomiale pour calculer la probabilité d'acceptation

Niveau de qualité acceptable (AQL) 0,05  
 Risque pour le fournisseur (Alpha) 0,89

Niveau de qualité rejetable (RQL ou LTPD) 0,5  
 Risque pour l'acheteur (Bêta) 0,1

Plan(s) généré(s)

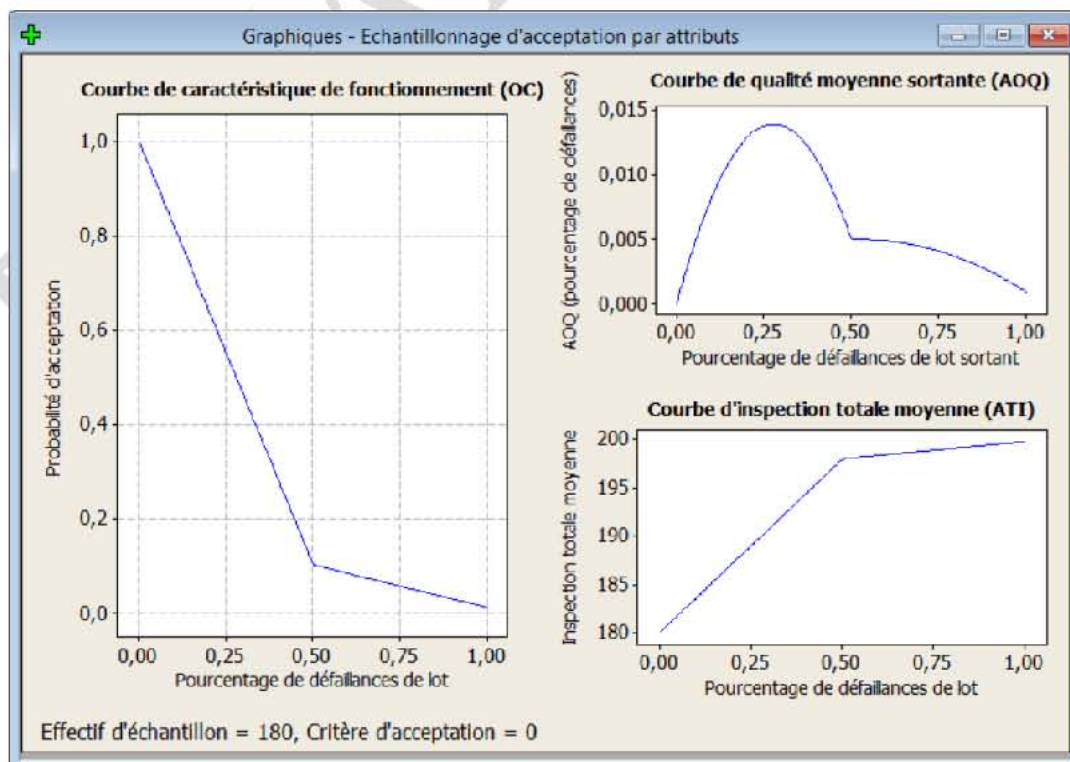
Effectif d'échantillon 180  
 Critère d'acceptation 0

Accepter le lot si des éléments défectueux dans l'échantillon 180  $\leq$  0. Sinon, rejeter.

Pourcentage de défaillances	Probabilité d'acceptation	Probabilité de rejet	AOQ	ATI
0,05	0,910	0,090	0,005	181,8
0,50	0,100	0,900	0,005	198,0

Limite de qualité moyenne sortante (AOQL) = 0,005 à 0,500 de pourcentage de défaillances.

et graphiquement:

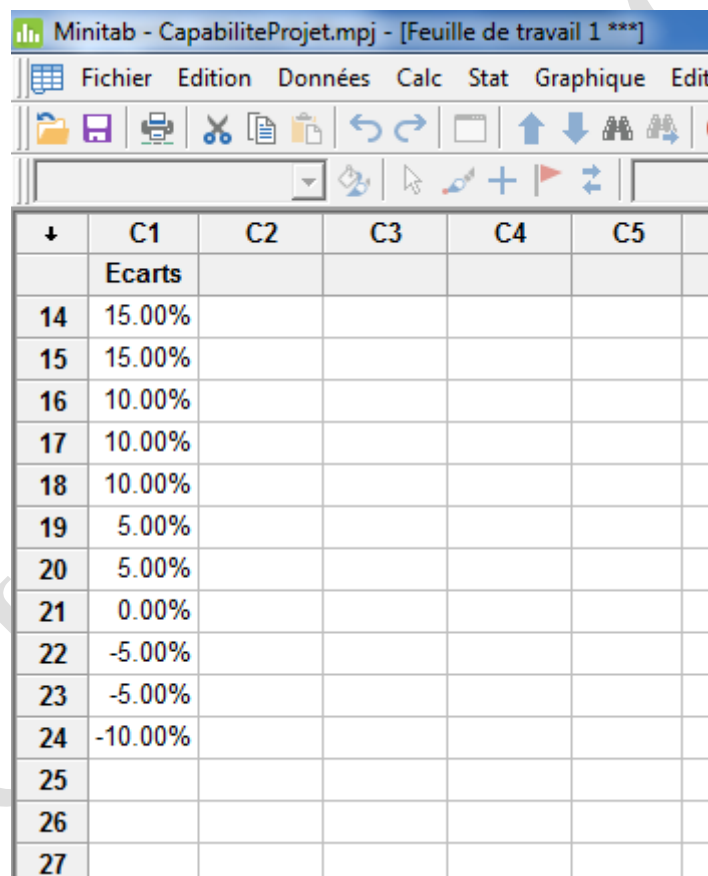


## 15.19. Exercice 163.: Intervalle de tolérance paramétrique pour loi Normale

Minitab® Statistical Software 18.10

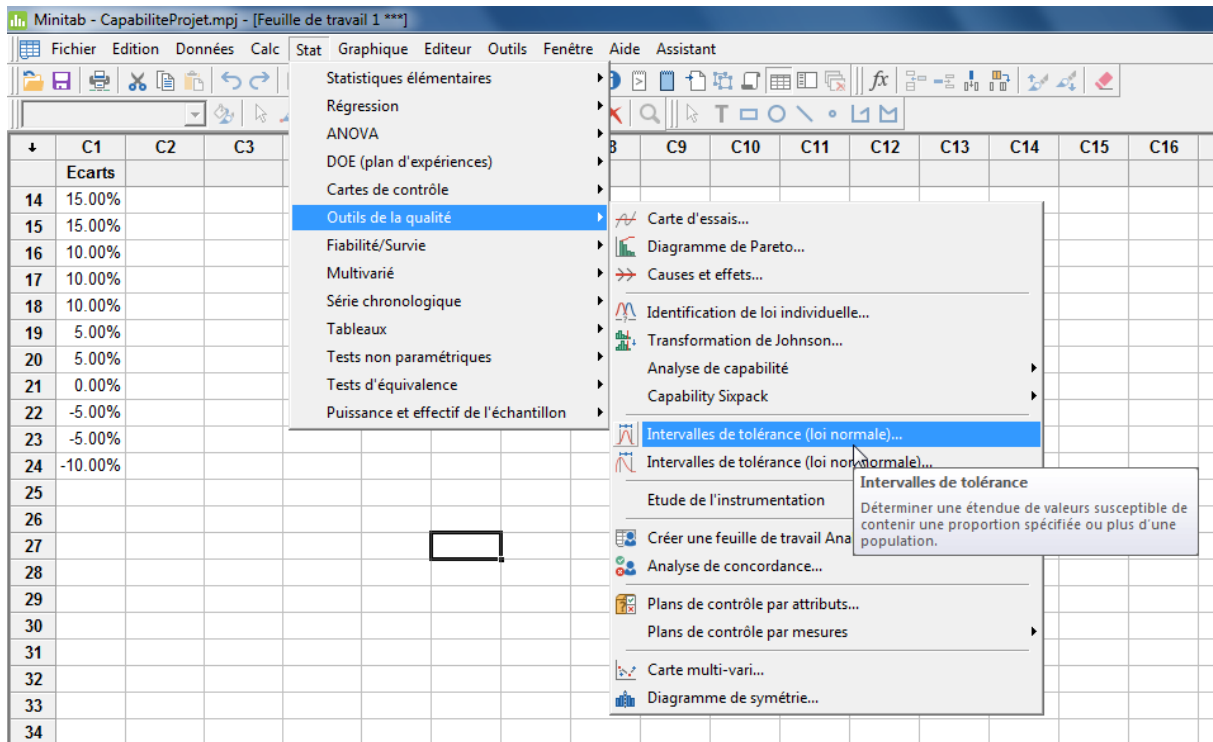
Nous avons étudié en détails dans le cours théorique ce qu'était l'intervalle de tolérance (paramétrique) pour une loi Normale et à quoi cela pouvait être extrêmement utile dans l'industrie et les services. Voyons donc de quoi il s'agit en reprenant le jeu de données *CapabiliteProjet.mpj* (et sans oublier que les données de l'intégrale qui ne peut être résolue à ce jour que numériquement est tabulée dans la norme ISO 16269-6).

Nous partons donc du même jeu de données que pour le test de Normalité et l'analyse de capabilité que nous avons fait plus tôt (14 dernières lignes visible ici seulement):

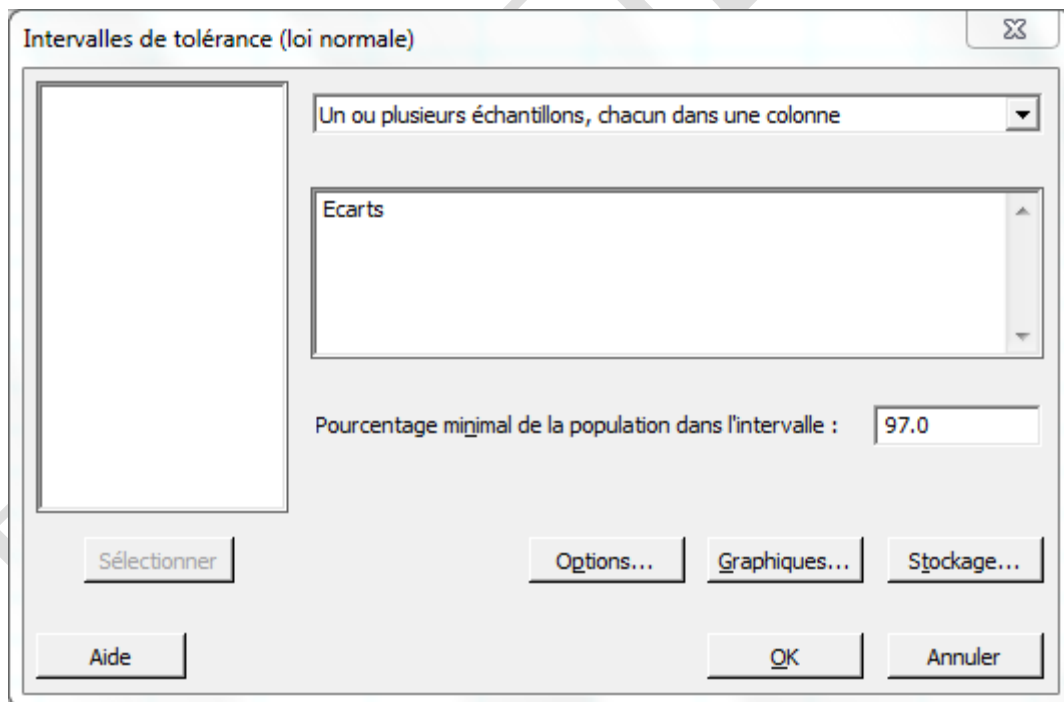


↓	C1	C2	C3	C4	C5
	<b>Ecart</b>				
14	15.00%				
15	15.00%				
16	10.00%				
17	10.00%				
18	10.00%				
19	5.00%				
20	5.00%				
21	0.00%				
22	-5.00%				
23	-5.00%				
24	-10.00%				
25					
26					
27					

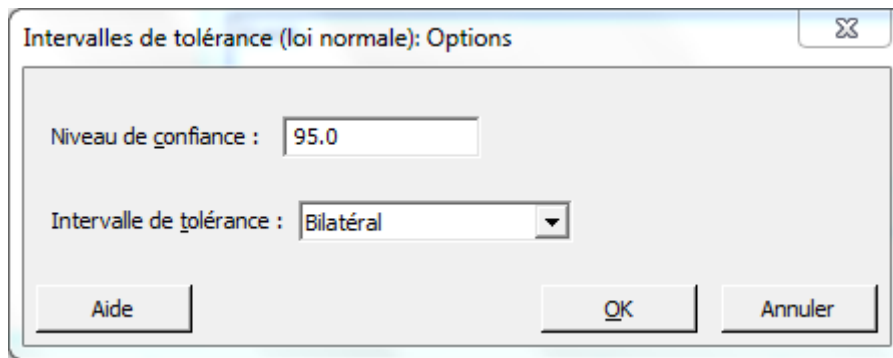
Nous allons dans le menu **Stat/Outils de la qualité/Intervalle de tolérance (loi normale)...**:



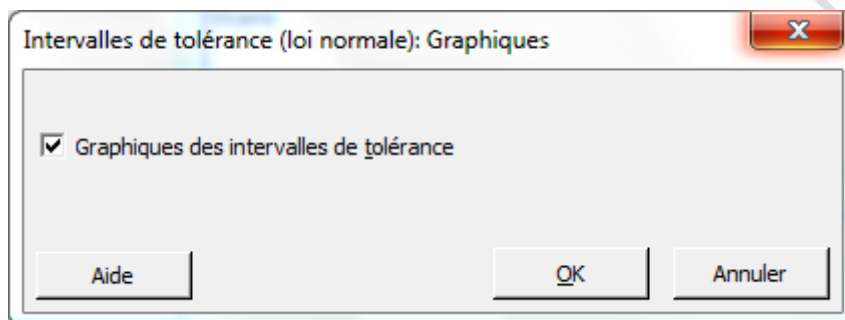
Et nous y mettons:



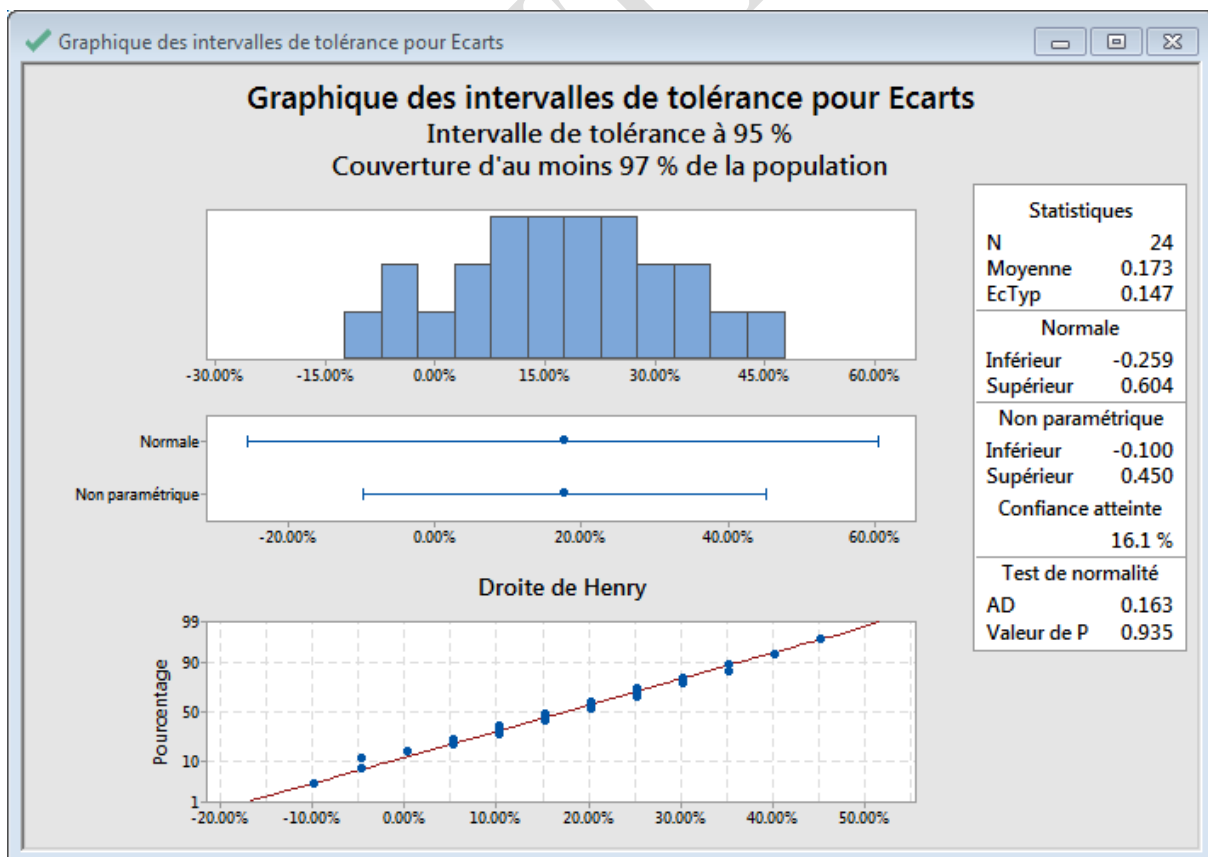
Nous prenons un intervalle de tolérance en bilatéral à 95% en cliquant sur le bouton **Options...**:



Le bouton Graphiques... ne propose malheureusement pas grand-chose (devrait permettre d'afficher ou non la version non-paramétrique):



Ce qui donne au final:





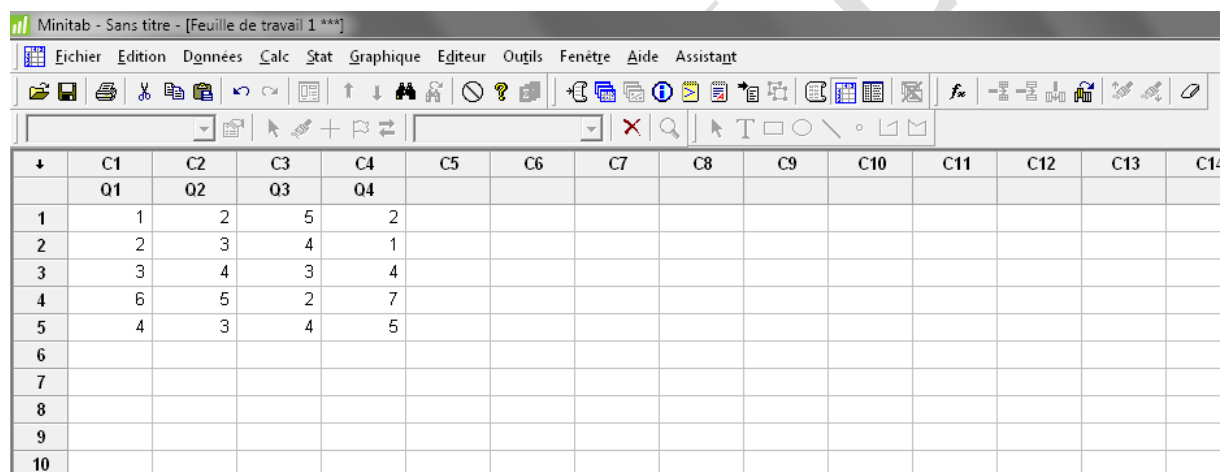
## 16. Statistiques multivariées (data mining/machine learning/fouille de données)

### 16.1. Exercice 164.: Étude de la cohérence avec l'alpha de Cronbach

Minitab® Statistical Software 16.1.1.0

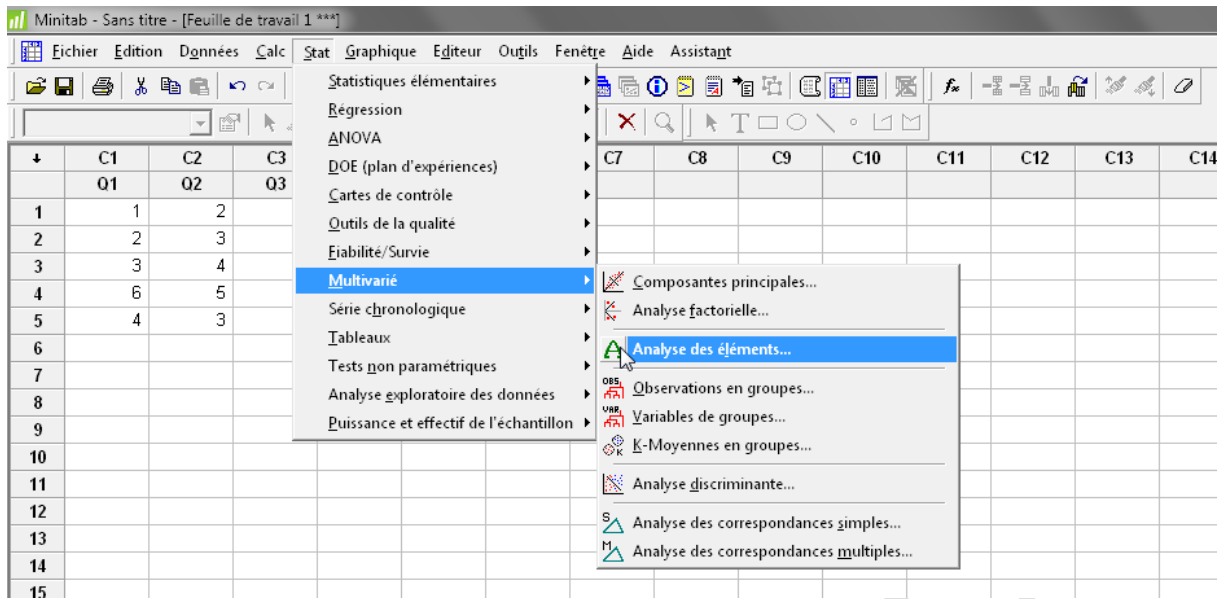
Comme à l'habitude, le but va être de vérifier ici que nous retombons sur les calculs faits à la main dans le cours théorique suite à la démonstration mathématique de l'origine de l'alpha de Cronbach.

Nous partons donc des données suivantes:

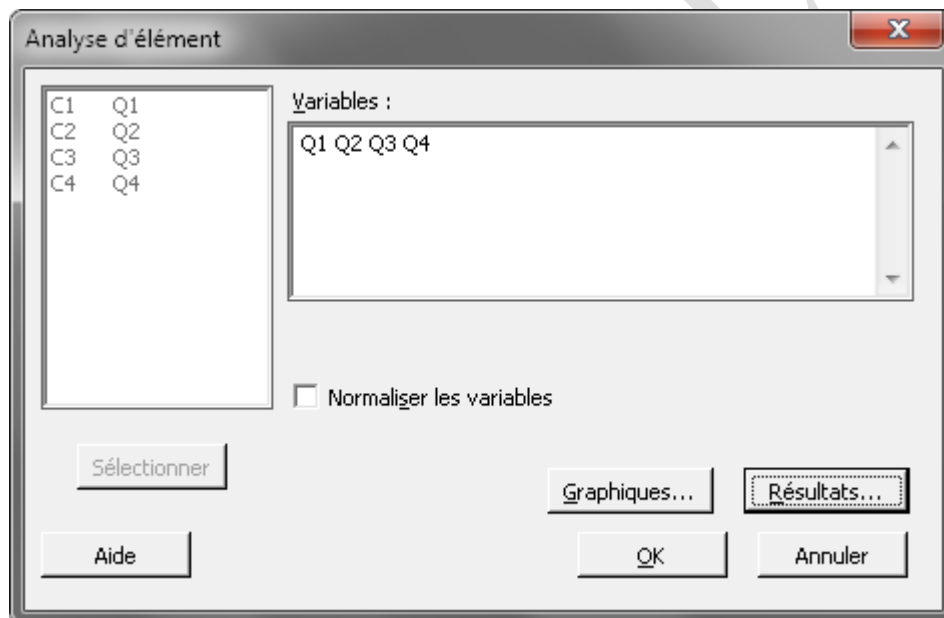


	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	O1	O2	O3	O4										
1	1	2	5	2										
2	2	3	4	1										
3	3	4	3	4										
4	6	5	2	7										
5	4	3	4	5										
6														
7														
8														
9														
10														

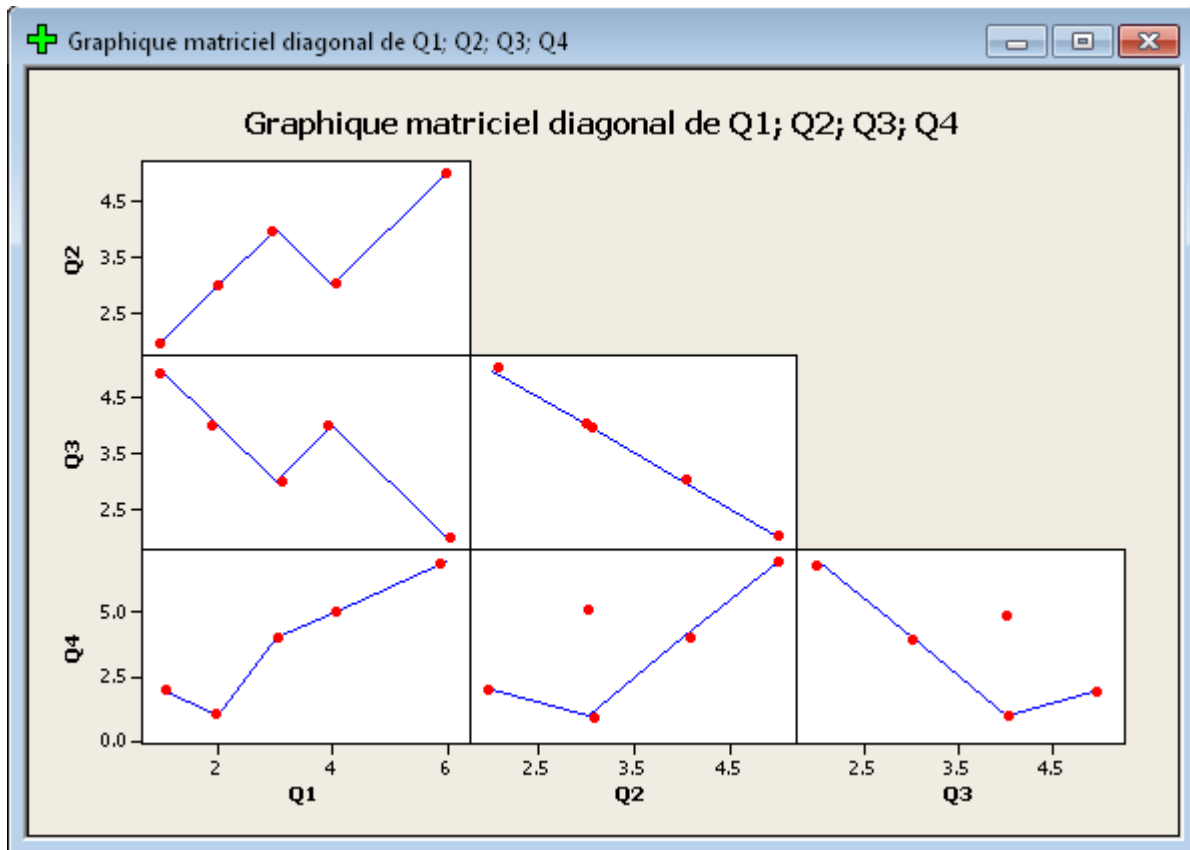
Nous allons dans le menu **Stat/Multivarié/Analyse des éléments...**:



Nous remplissons tel que visible ci-dessous et laissons les options standards:



Nous validons par OK pour obtenir le graphique matriciel simple (dont l'utilité est discutable selon moi):



et surtout vient dans la fenêtre de session ce qui nous intéresse:

Minitab - Sans titre - [Session]

Echier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assistant

Analyse des éléments de Q1; Q2; Q3; Q4

Matrice de corrélation

	Q1	Q2	Q3
Q2	0.866		
Q3	-0.866	-1.000	
Q4	0.936	0.771	-0.771

Contenu de la cellule : Corrélation de Pearson

Élément et statistiques totales

Variable	Dénombrement		
	total	Moyenne	EcTyp
Q1	5	3.200	1.924
Q2	5	3.400	1.140
Q3	5	3.600	1.140
Q4	5	3.800	2.387
Total	5	14.000	4.243

Alpha de Cronbach = 0.4444

Statistiques d'élément omises

Variable omise	Moyenne totale	EcTyp totale	Corr totale	Corr multiple	Alpha de Cronbach
	ajust.	ajust.	élé-m-ajust	quadratique	
Q1	10.800	2.387	0.9363	0.9279	-0.6842
Q2	10.600	3.362	0.7045	1.0000	0.0796
Q3	10.400	5.225	-0.8897	1.0000	0.9121
Q4	10.200	1.924	0.9363	0.8830	-1.0541

et nous retrouvons bien donc tout ce qui a été calculé à la main dans le cours avec le tableur Microsoft Excel. Il est juste dommage de constater que nous n'avons pas de seuil de confiance à un niveau alpha donné comme le fait le logiciel MedCalc™ par exemple.

ÉCHANTILLON

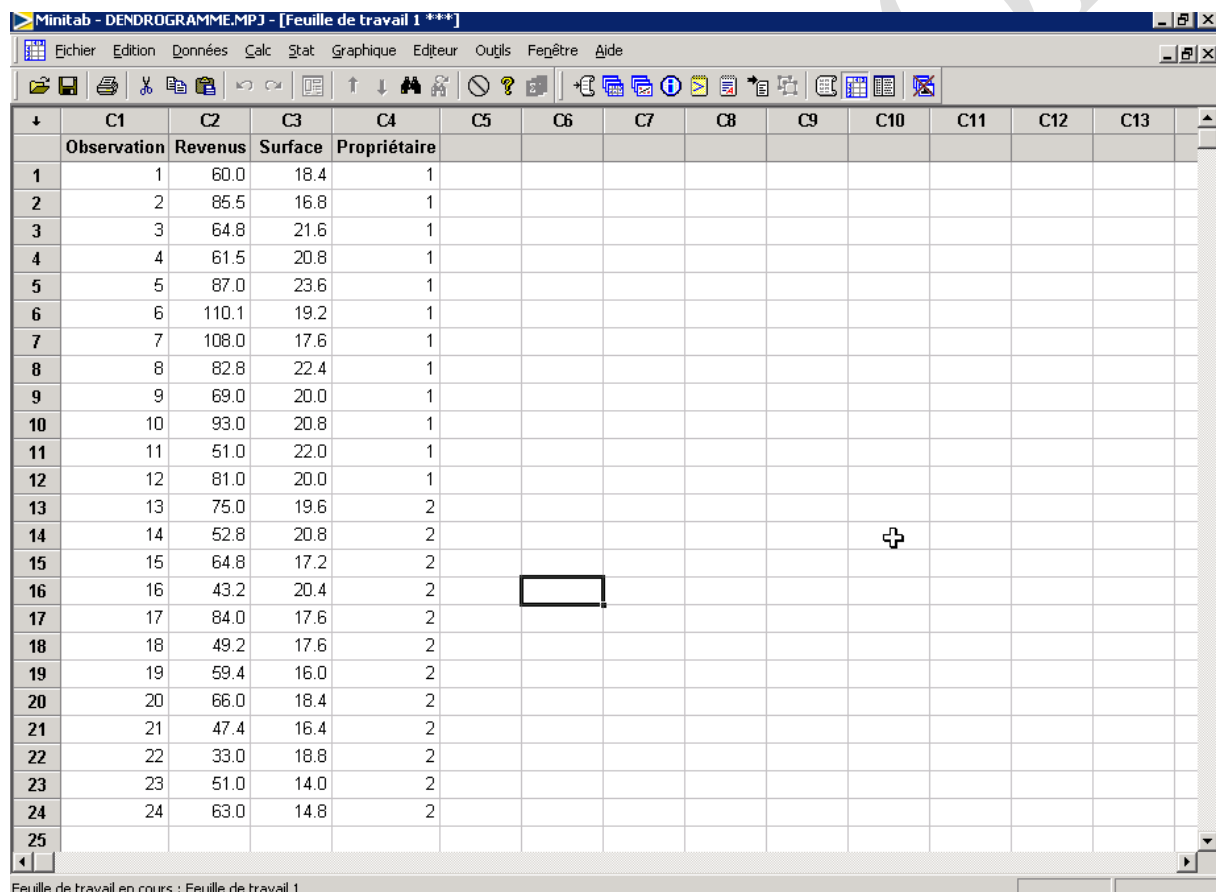
## 16.2. Exercice 165.: Clustering dendrogramme

Minitab® Statistical Software 15.1.1.0

Dans le cours de méthodes numériques nous avons parlé, démontré et implémenté dans Microsoft Excel les algorithmes pour les CART, les dendrogrammes ainsi que la méthode  $K$ -means pour pouvoir faire du regroupement.

Le but va être ici de vérifier si nous obtenons bien avec Minitab le même dendrogramme ou groupement  $K$ -means que dans le cours théorique (puisque les CART ne sont pas disponibles à ce jour dans Minitab).

Pour vérifier cela, nous partons exactement du même échantillon de données que dans le cours théorique (ou vous ouvrez le fichier *Dendrogramme.mpj*):



	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Observation	Revenus	Surface	Propriétaire									
1	1	60.0	18.4	1									
2	2	85.5	16.8	1									
3	3	64.8	21.6	1									
4	4	61.5	20.8	1									
5	5	87.0	23.6	1									
6	6	110.1	19.2	1									
7	7	108.0	17.6	1									
8	8	82.8	22.4	1									
9	9	69.0	20.0	1									
10	10	93.0	20.8	1									
11	11	51.0	22.0	1									
12	12	81.0	20.0	1									
13	13	75.0	19.6	2									
14	14	52.8	20.8	2									
15	15	64.8	17.2	2									
16	16	43.2	20.4	2									
17	17	84.0	17.6	2									
18	18	49.2	17.6	2									
19	19	59.4	16.0	2									
20	20	66.0	18.4	2									
21	21	47.4	16.4	2									
22	22	33.0	18.8	2									
23	23	51.0	14.0	2									
24	24	63.0	14.8	2									
25													

Ensuite, nous allons dans le menu **Stat/Multivarié/Observations en groupes...**:

The screenshot shows the Minitab interface with the 'Stat' menu open. The 'Observations en groupes...' option is highlighted. The background shows a data table with columns C1 to C13 and rows 1 to 25.

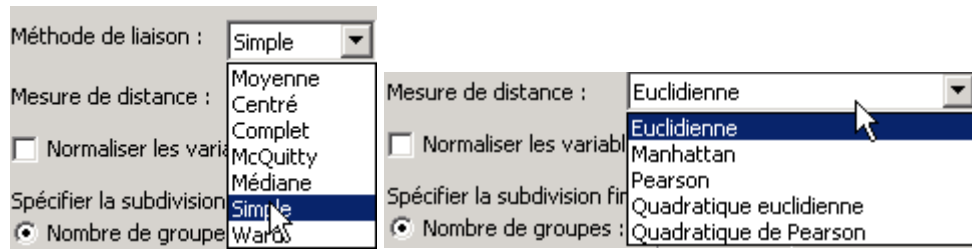
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Observation	Revenus	Surface	Propriétaire									
1	1	60.0											
2	2	85.5											
3	3	64.8											
4	4	61.5											
5	5	87.0											
6	6	110.1											
7	7	108.0											
8	8	82.8											
9	9	69.0											
10	10	93.0											
11	11	51.0	22.0	1									
12	12	81.0	20.0	1									
13	13	75.0	19.6	2									
14	14	52.8	20.8	2									
15	15	64.8	17.2	2									
16	16	43.2	20.4	2									
17	17	84.0	17.6	2									
18	18	49.2	17.6	2									
19	19	59.4	16.0	2									
20	20	66.0	18.4	2									
21	21	47.4	16.4	2									
22	22	33.0	18.8	2									
23	23	51.0	14.0	2									
24	24	63.0	14.8	2									
25													

et nous y mettons:

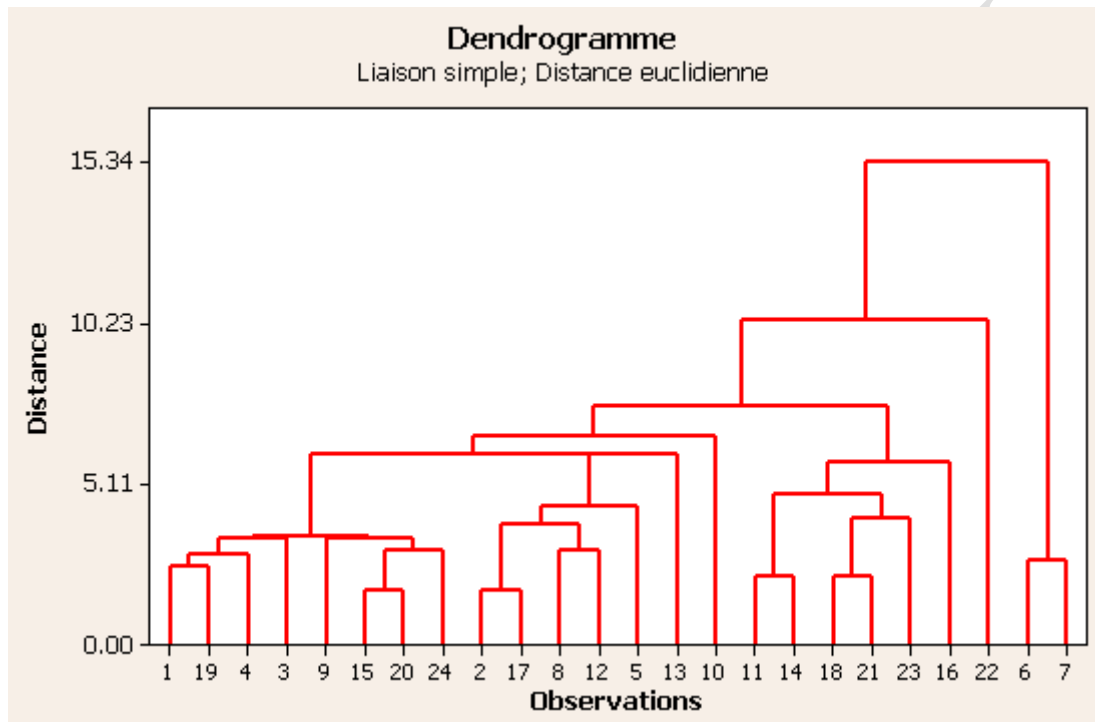
The dialog box 'Observations en groupes' is shown. It contains the following settings:

- Matrice de distance ou variables : Revenus Surface
- Méthode de liaison : Simple
- Mesure de distance : Euclidienne
- Normaliser les variables
- Spécifier la subdivision finale par:
  - Nombre de groupes : 1
  - Niveau de similarité : [ ]
- Montrer le dendrogramme
- Buttons: Sélectionner, Aide, Stockage..., Personnaliser..., OK, Annuler

Remarquez que nous y retrouvons les méthodes de liaisons ainsi que les mesures de distances dont nous avons parlées dans le cours théorique (même si nous allons garder la **Simple** et **Euclidienne**):



Nous gardons donc tout tel que et validons par **OK** pour obtenir d'abord le dendrogramme suivant (il est dommage que l'on ne voie pas bien l'ordre de traitement des groupements):



et dans la fenêtre de sessions le listing suivant:

**Analyse hiérarchique des observations : Revenus; Surface**

Distance euclidienne, Liaison simple

Etapas de fusion

Etape	Nombre de groupes	Niveau de similarité	Niveau de distance	Groupes liés		Nouveau groupe	Nombre d'obs. dans le nouveau groupe
1	23	97.7989	1.6971	15	20	15	2
2	22	97.7951	1.7000	2	17	2	2
3	21	97.1942	2.1633	11	14	11	2
4	20	97.1942	2.1633	18	21	18	2
5	19	96.7914	2.4739	1	19	1	2
6	18	96.5758	2.6401	6	7	6	2
7	17	96.3292	2.8302	1	4	1	3
8	16	96.1090	3.0000	15	24	15	3
9	15	96.1090	3.0000	8	12	8	2
10	14	95.5959	3.3956	1	3	1	4
11	13	95.5902	3.4000	9	15	9	4
12	12	95.5674	3.4176	1	9	1	8
13	11	95.0171	3.8419	2	8	2	4
14	10	94.7797	4.0249	18	23	18	3
15	9	94.3346	4.3681	2	5	2	5
16	8	93.8341	4.7539	11	18	11	5
17	7	92.4774	5.8000	11	16	11	6
18	6	92.2007	6.0133	2	13	2	6
19	5	92.2007	6.0133	1	2	1	14
20	4	91.4123	6.6212	1	10	1	15
21	3	90.1565	7.5895	1	11	1	21
22	2	86.6088	10.3247	1	22	1	22
23	1	80.1072	15.3375	1	6	1	24

Subdivision finale  
 Nombre de groupes : 1

Groupes	Nombre d'observations	Dans la somme de groupe des carrés	Distance moyenne du centré	Distance maximale du centré
Groupes1	24	9146.30	16.1994	41.6633

où nous voyons que tout est conforme à ce que nous avons calculé dans le cours théorique avec Microsoft Excel.

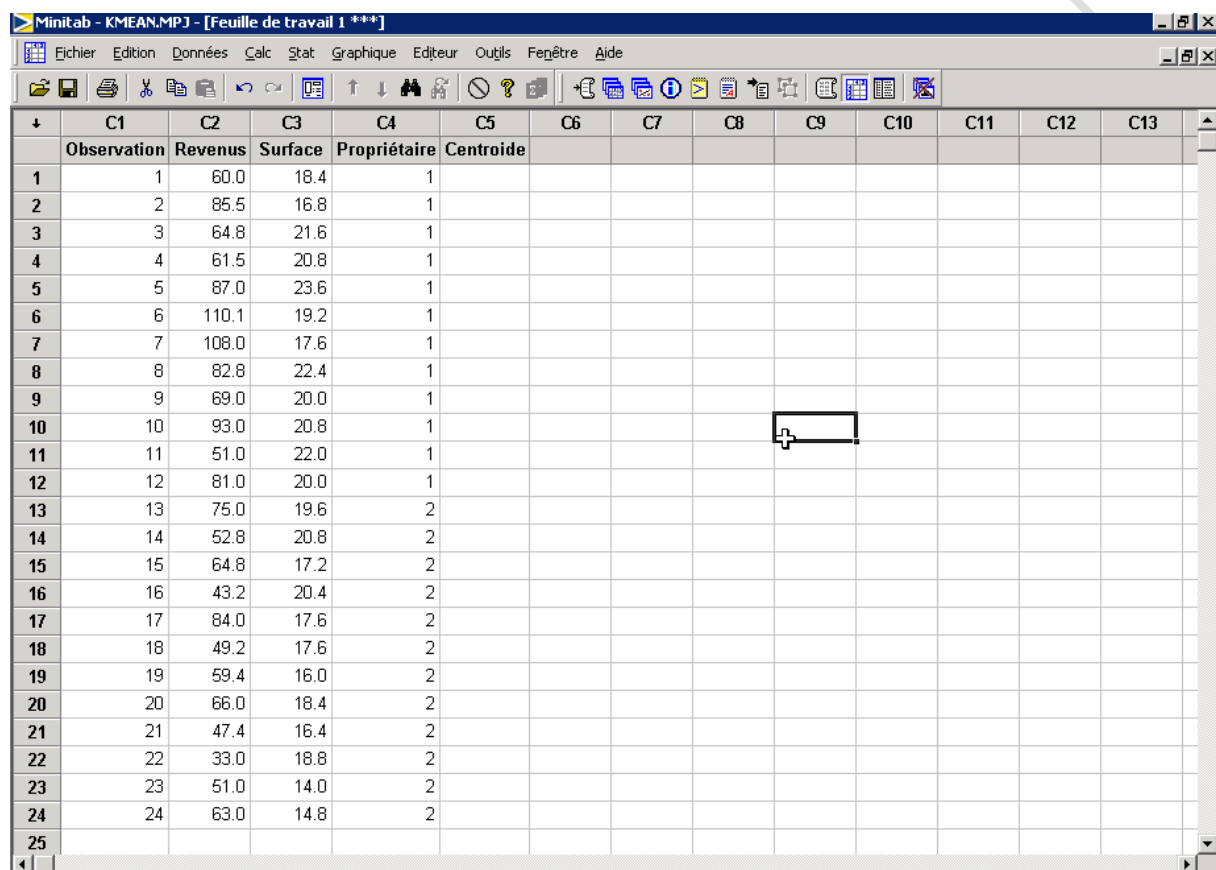


## 16.3. Exercice 166.: K-means (k-moyennes en groupes)

Minitab® Statistical Software 15.1.1.0

Comme d'habitude, le but est de vérifier ce que nous avons vu dans le cours théorique de Méthodes Numériques concernant les k-Means (toujours du clustering) et que nous avons essayé d'approcher avec Microsoft Excel et son solveur (sachant qu'il n'était donc pas possible d'avoir deux extremums en même temps avec cet outil).

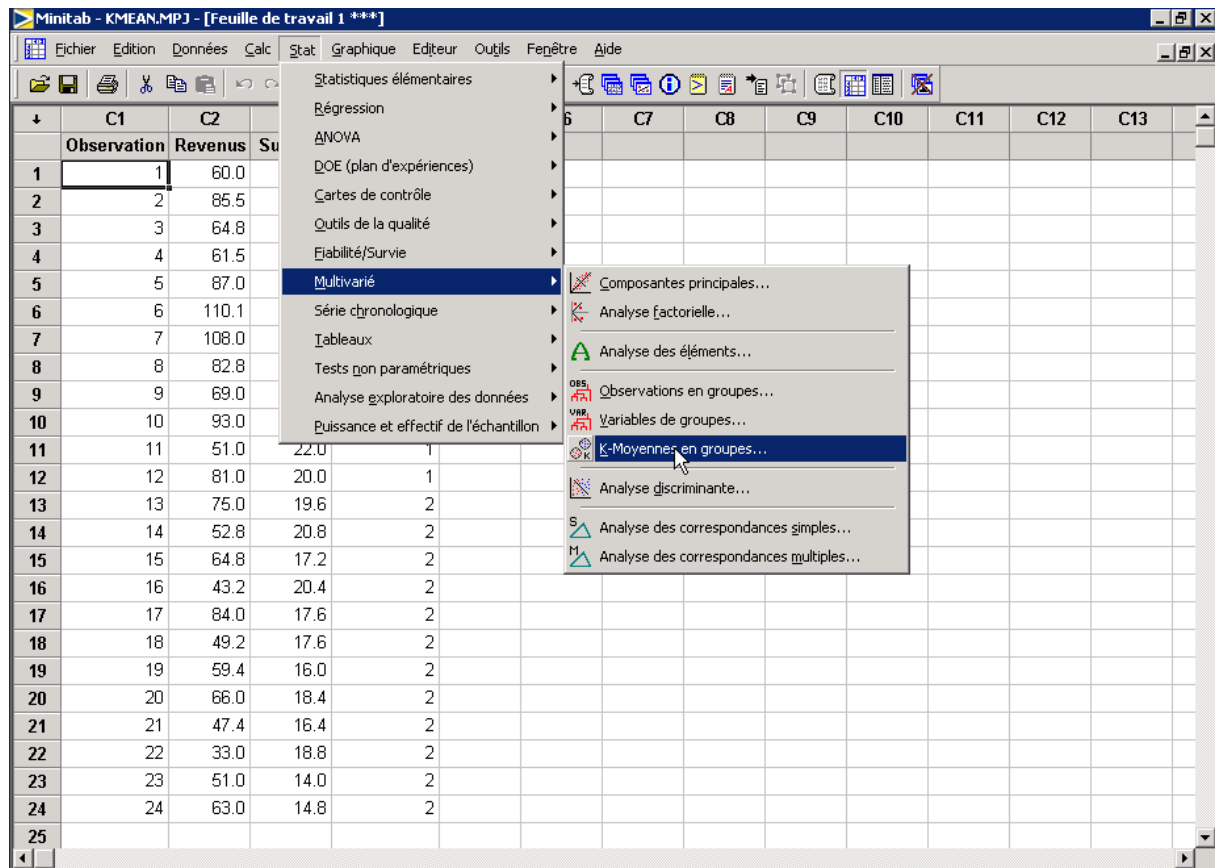
Nous partons donc toujours des mêmes données (ouvrez le fichier *kmeans.mpj*):



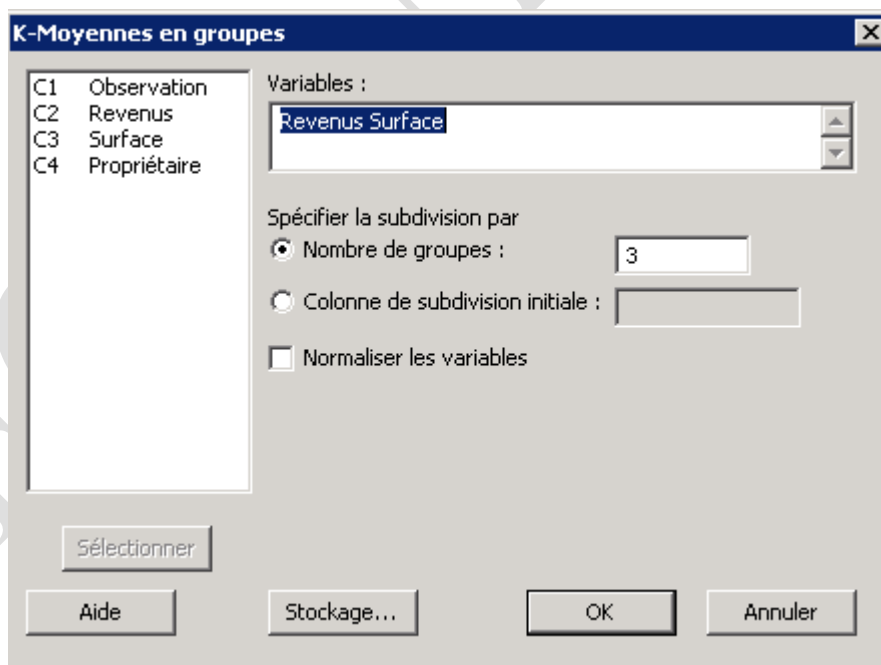
The screenshot shows the Minitab interface with a data table. The table has 25 rows and 13 columns. The first five columns are labeled: Observation, Revenus, Surface, Propriétaire, and Centroide. The data is as follows:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Observation	Revenus	Surface	Propriétaire	Centroide								
1	1	60.0	18.4	1									
2	2	85.5	16.8	1									
3	3	64.8	21.6	1									
4	4	61.5	20.8	1									
5	5	87.0	23.6	1									
6	6	110.1	19.2	1									
7	7	108.0	17.6	1									
8	8	82.8	22.4	1									
9	9	69.0	20.0	1									
10	10	93.0	20.8	1									
11	11	51.0	22.0	1									
12	12	81.0	20.0	1									
13	13	75.0	19.6	2									
14	14	52.8	20.8	2									
15	15	64.8	17.2	2									
16	16	43.2	20.4	2									
17	17	84.0	17.6	2									
18	18	49.2	17.6	2									
19	19	59.4	16.0	2									
20	20	66.0	18.4	2									
21	21	47.4	16.4	2									
22	22	33.0	18.8	2									
23	23	51.0	14.0	2									
24	24	63.0	14.8	2									
25													

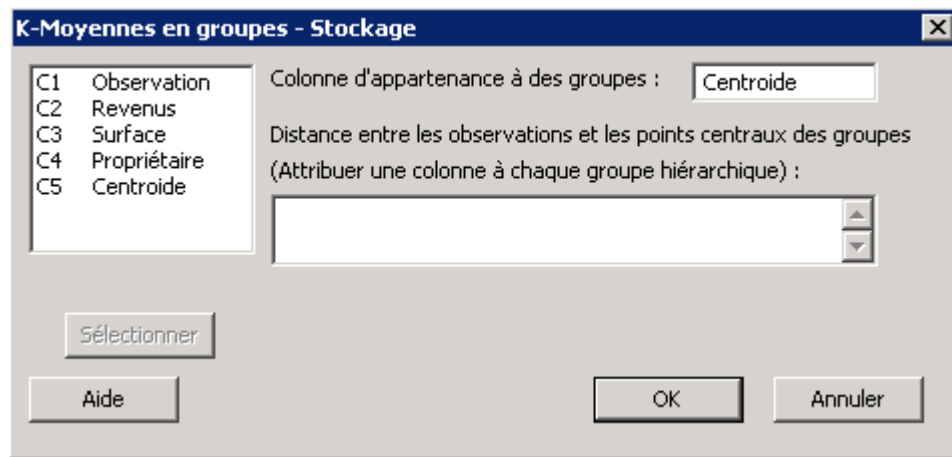
Nous allons ensuite dans le menu **Stat/Multivarié/K-Moyennes en groupes...**:



et nous prenons:

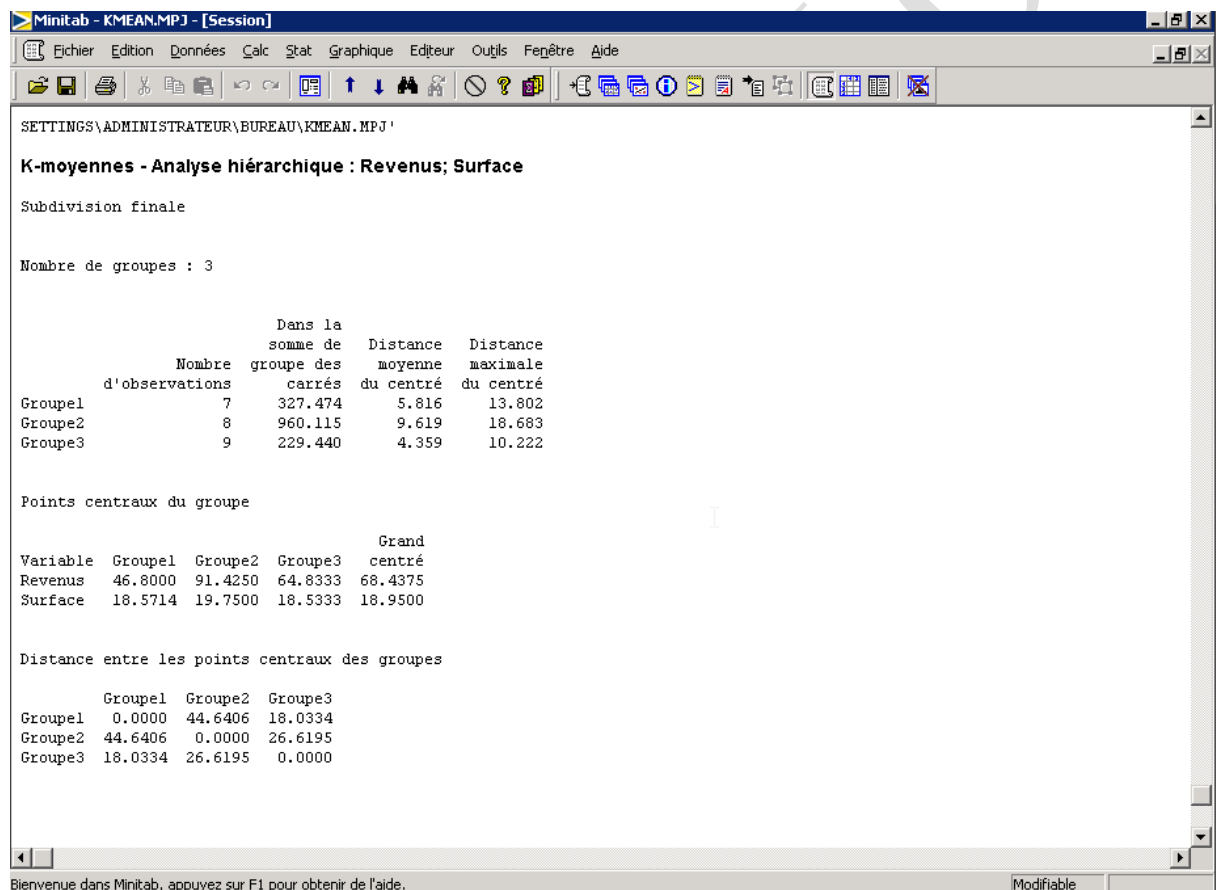


et nous cliquons aussi (pour comparer avec Microsoft Excel plus facilement) sur le bouton **Stockage...**:



pour avoir donc les appartenances des individus signalées automatiquement.

Nous validons le tout par **OK** pour obtenir (sans graphique ce qui est assez lamentable...):



et:

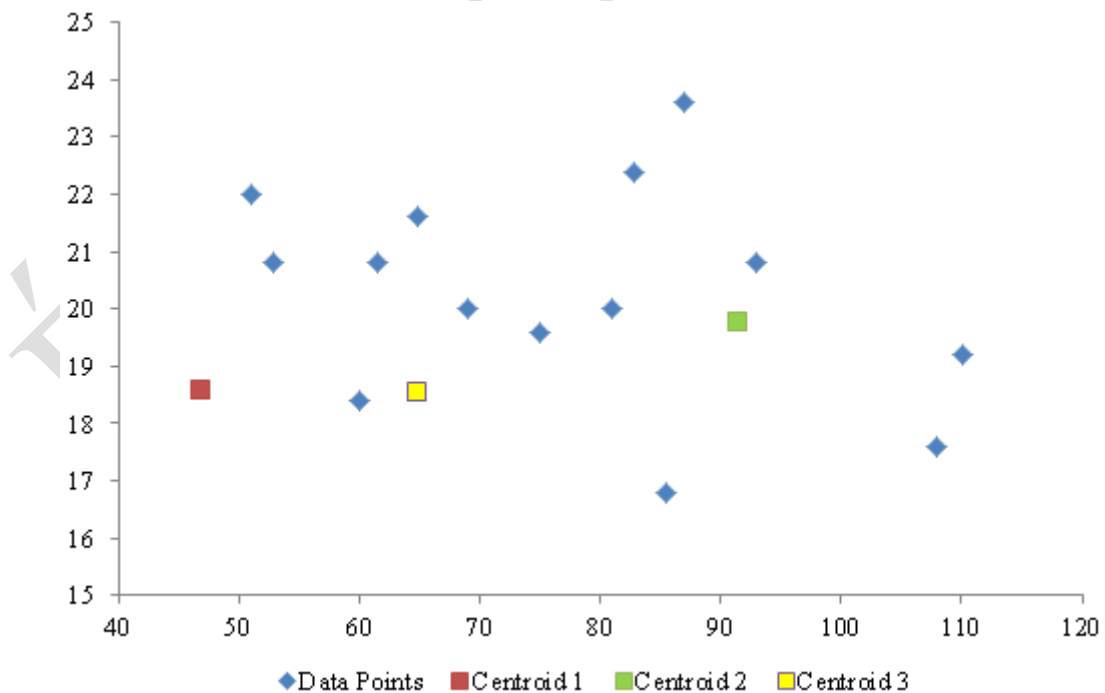
Minitab - KMEAN.MPJ - [Feuille de travail 1 \*\*\*]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Observation	Revenus	Surface	Propriétaire	Centroïde								
1	1	60.0	18.4	1	3								
2	2	85.5	16.8	1	2								
3	3	64.8	21.6	1	3								
4	4	61.5	20.8	1	3								
5	5	87.0	23.6	1	2								
6	6	110.1	19.2	1	2								
7	7	108.0	17.6	1	2								
8	8	82.8	22.4	1	2								
9	9	69.0	20.0	1	3								
10	10	93.0	20.8	1	2								
11	11	51.0	22.0	1	1								
12	12	81.0	20.0	1	2								
13	13	75.0	19.6	2	3								
14	14	52.8	20.8	2	1								
15	15	64.8	17.2	2	3								
16	16	43.2	20.4	2	1								
17	17	84.0	17.6	2	2								
18	18	49.2	17.6	2	1								
19	19	59.4	16.0	2	3								
20	20	66.0	18.4	2	3								
21	21	47.4	16.4	2	1								
22	22	33.0	18.8	2	1								
23	23	51.0	14.0	2	1								
24	24	63.0	14.8	2	3								
25													

Feuille de travail en cours : Feuille de travail 1

Le résultat de Minitab représenté avec le fichier Microsoft Excel que nous avons utilisé pendant le cours théorique:

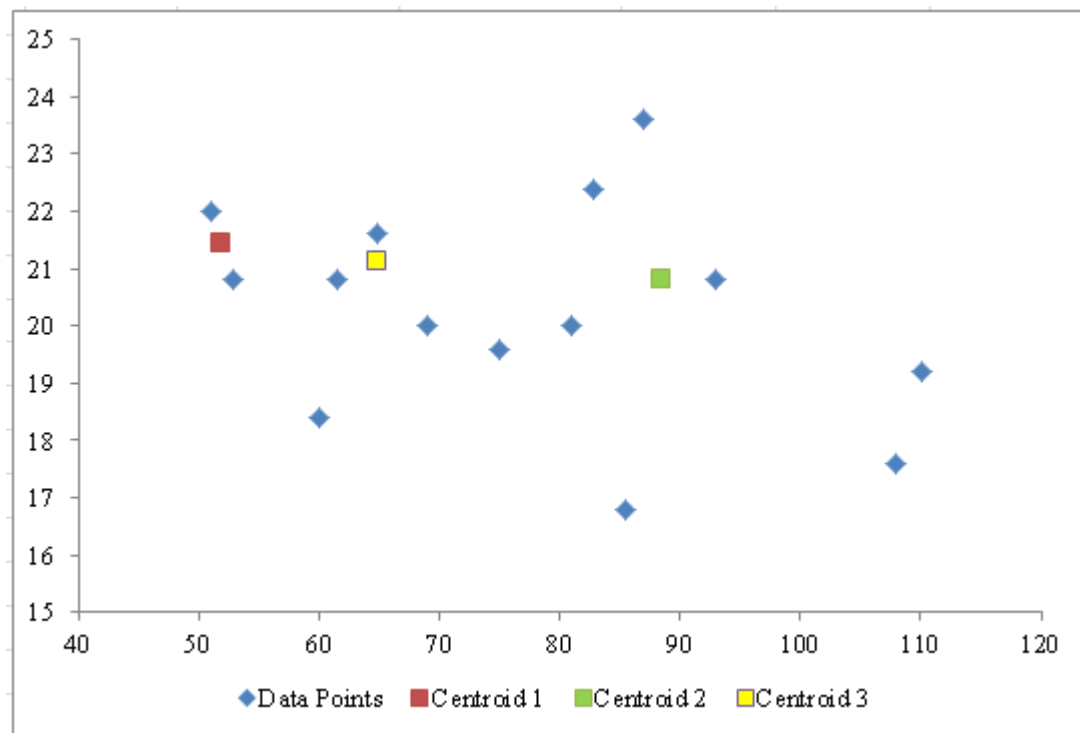


N°	Distance to CENTROID 1	Distance to CENTROID 2	Distance to CENTROID 3	CLASS	Minimum Distance
1	13.20111275	31.45398425	4.835137824	Cluster3	4.835137824
2	38.74051959	6.618770656	20.73925789	Cluster2	6.618770656
3	18.2530112	26.68919491	3.06688079	Cluster3	3.06688079
4	14.86797424	29.94341539	4.030982235	Cluster3	4.030982235
5	40.51329187	5.865417717	22.73838248	Cluster2	5.865417717
6	63.30312108	18.68309731	45.2716094	Cluster2	18.68309731
7	61.20770881	16.71386027	43.17678818	Cluster2	16.71386027
8	36.20301338	9.022922198	18.37807601	Cluster2	9.022922198
9	22.24591868	22.42639349	4.417306621	Cluster3	4.417306621
10	46.25372048	1.89291442	28.25775854	Cluster2	1.89291442
11	5.421743074	40.48756754	14.26107281	Cluster1	5.421743074
12	34.22982468	10.42799717	16.23309575	Cluster2	10.42799717
13	28.21875295	16.42568492	10.22250643	Cluster3	10.22250643
14	6.400520132	38.63926921	12.24492702	Cluster1	6.400520132
15	18.05216713	26.74683392	1.333715779	Cluster3	1.333715779
16	4.0377937	48.22938031	21.7136878	Cluster1	4.0377937
17	37.21268088	7.730014554	19.18940952	Cluster2	7.730014554
18	2.589134597	42.2797011	15.66113399	Cluster1	2.589134597
19	12.85970832	32.24380755	5.994860948	Cluster3	5.994860948
20	19.20076504	25.46081548	1.17429033	Cluster3	1.17429033
21	2.252771173	44.15227203	17.56334017	Cluster1	2.252771173
22	13.80189327	58.43272307	31.83441719	Cluster1	13.80189327
23	6.207873868	40.83188858	14.55716311	Cluster1	6.207873868
24	16.63320348	28.85278366	4.159148685	Cluster3	4.159148685
				<b>Sum</b>	<b>107.6200568</b>

À la question: *Retrouvons-nous avec Minitab les mêmes résultats qu'avec Microsoft Excel?*  
 La réponse est bien évidemment négative puisque le solveur de Microsoft Excel ne peut rendre extremum qu'un paramètre à la fois. Par contre, les groupes obtenus sont exactement les mêmes (voir ci-dessous les captures d'écran obtenues pendant le cours théorique).

Remarque: Nous obtenons aussi les mêmes résultats que ceux vus dans le cours de Data Mining avec Tanagra.

Pour rappel, le solveur GRG ou Évolutionnaire de Microsoft Excel donnait un résultat différent (voir ci-dessous) mais... la distance entre les groupes n'est pas aussi grande que le résultat trouvé par Minitab (car dans Microsoft Excel on peut minimiser la distance des points au centre d'un groupe mais pas en même temps maximiser la distance entre les centres des groupes):



N°	Distance to CENTROID 1	Distance to CENTROID 2	Distance to CENTROID 3	CLASS	Minimum Distance
1	8.686145481	28.48473914	5.555442106	Cluster3	5.555442106
2	33.95821155	4.950389833	21.10594524	Cluster2	4.950389833
3	12.9425742	23.59400135	0.477777452	Cluster3	0.477777452
4	9.661830186	26.88129768	3.357419365	Cluster3	3.357419365
5	35.20850988	3.099333368	22.29615695	Cluster2	3.099333368
6	58.28401019	21.77945743	45.29913525	Cluster2	21.77945743
7	56.27175266	19.88209564	43.30189304	Cluster2	19.88209564
8	30.95670218	5.799123839	18.00352711	Cluster2	5.799123839
9	17.20077021	19.39885723	4.307502314	Cluster3	4.307502314
10	41.14621233	4.618784727	28.16012	Cluster2	4.618784727
11	1.031892127	37.39973251	13.86943333	Cluster1	1.031892127
12	29.17637255	7.427301486	16.19730555	Cluster2	7.427301486
13	23.21348054	13.43728488	10.27194633	Cluster3	10.27194633
14	1.131438639	35.58129473	12.04610393	Cluster1	1.131438639
15	13.61444197	23.85835688	3.924271737	Cluster3	3.924271737
16	8.719337751	45.18328904	21.65385312	Cluster1	8.719337751
17	32.3685296	5.44053634	19.47967347	Cluster2	5.44053634
18	4.660308126	39.31382598	16.03381111	Cluster1	4.660308126
19	9.291493451	29.38026971	7.474521323	Cluster3	7.474521323
20	14.46188899	22.51232855	2.960067986	Cluster3	2.960067986
21	6.719788539	41.21954285	18.07017151	Cluster1	6.719788539
22	19.04074726	55.41831282	31.92644519	Cluster1	19.04074726
23	7.477049022	37.99931313	15.56746531	Cluster1	7.477049022
24	12.9636544	26.08670472	6.586776722	Cluster3	6.586776722
				<b>Sum of</b>	<b>93.68990466</b>

## 16.4. Exercice 167.: Analyse Linéaire Discriminante

Minitab® Statistical Software 16.1

Nous partons des mêmes données que celles utilisées dans le cours théorique aussi:

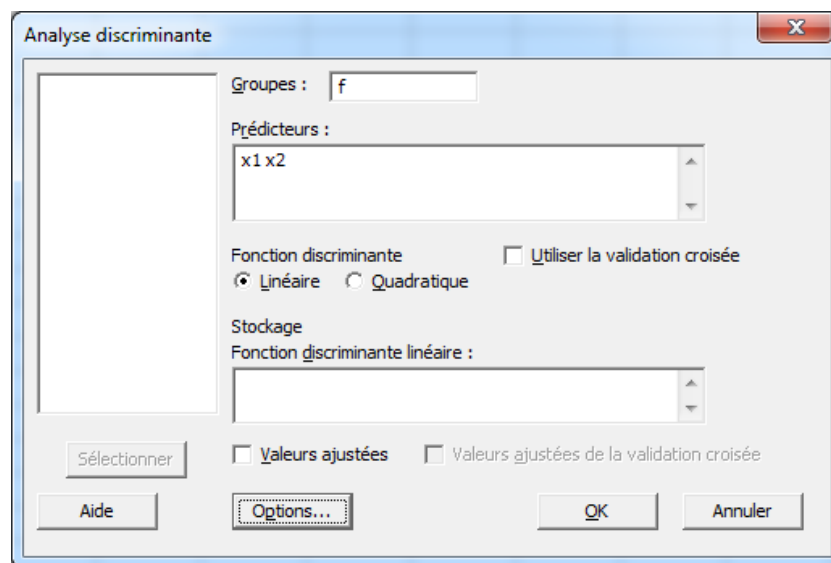
	C1	C2	C3-T
	x1	x2	f
1	4	2	a
2	2	4	a
3	2	3	a
4	3	6	a
5	4	4	a
6	9	10	b
7	6	8	b
8	9	5	b
9	8	7	b
10	10	8	b
11			

Ensuite, nous allons dans le menu **Stat/Multivarié/Analyse discriminante...**:

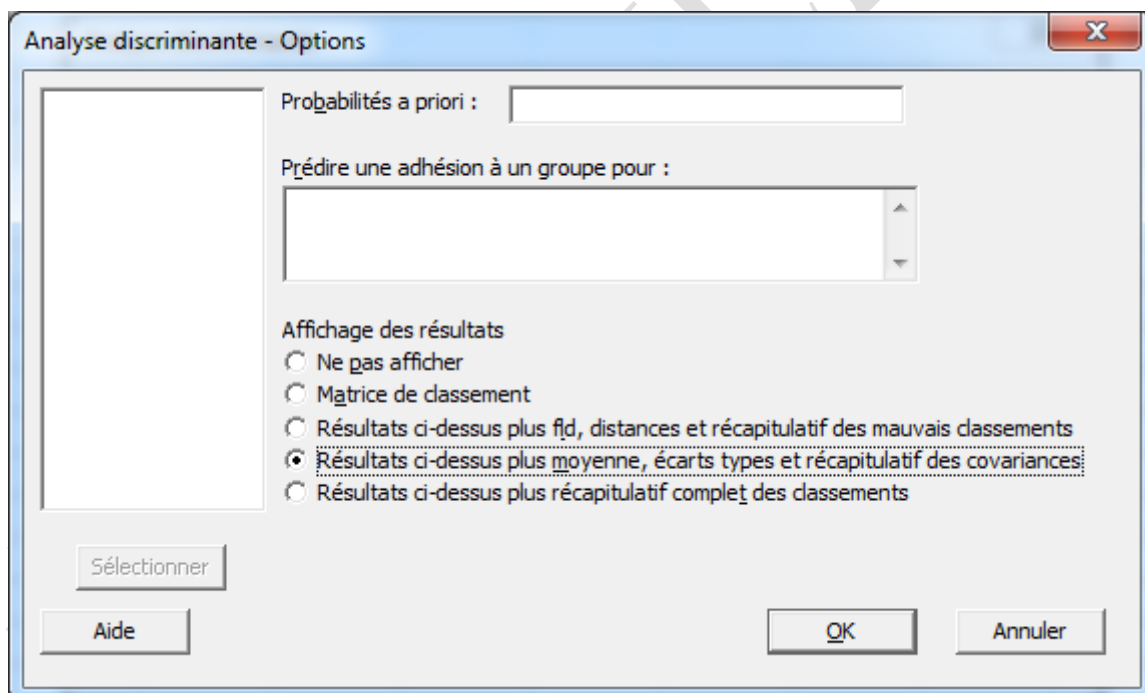
Stat

- Statistiques élémentaires
- Régression
- ANOVA
- DOE (plan d'expériences)
- Cartes de contrôle
- Outils de la qualité
- Fiabilité/Survie
- Multivarié**
  - Composantes principales...
  - Analyse factorielle...
  - Analyse des éléments...
  - Observations en groupes...
  - Variation de groupes...
  - K-Moyennes en groupes...
  - Analyse discriminante...**
  - Analyse des correspondances simples...
  - Analyse des correspondances multiples...
- Série chronologique
- Tableaux
- Tests non paramétriques
- Analyse exploratoire des données
- Puissance et effectif de l'échantillon

Dans la boîte de dialogue qui apparaît, nous mettons:



Ensuite, nous cliquons sur **Options...** et nous prenons la quatrième option afin de pouvoir vérifier les détails par rapport au cours théorique:



Nous validons deux fois par **OK** et nous obtenons:



Minitab - AnalyseDiscriminante.MPJ - [Session]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide As

Analyse discriminante : f en fonction de x1; x2

Méthode linéaire pour la réponse : f

Prédicteurs : x1; x2

Groupe	a	b
Dénombrement	5	5

Récapitulatif du classement

Mettre dans groupe	Groupe vrai	
	a	b
a	5	0
b	0	5
Nombre total	5	5
N correct	5	5
Proportion	1.000	1.000

N = 10                      N correct = 10                      Proportion correcte = 1.000

Distance quadratique entre les groupes

	a	b
a	0.0000	24.4013
b	24.4013	0.0000

Jusque-là rien de spécial à part qu'afficher la distance entre les groupes presque au début n'est peut-être pas un choix pertinent.

Nous avons ensuite:

Fonction discriminante linéaire pour les groupes

	a	b
Constante	-5.758	-34.163
x1	1.953	5.369
x2	1.488	3.056

Variable	Moyenne de regroupement	Moyennes du groupe	
		a	b
x1	5.7000	3.0000	8.4000
x2	5.7000	3.8000	7.6000

Variable	Ecart type de regroupement	EcTyp du groupe	
		a	b
x1	1.285	1.000	1.517
x2	1.658	1.483	1.817

Matrice de covariance groupée

	x1	x2
x1	1.6500	
x2	-0.1500	2.7500

Matrice de covariance du groupe a

	x1	x2
x1	1.00000	
x2	-0.25000	2.20000

Matrice de covariance du groupe b

	x1	x2
x1	2.30000	
x2	-0.05000	3.30000

Là aussi nous retrouvons tous les calculs faits dans le cours théorique. La sortie pourrait cependant être améliorée par un graphique (quand cela est possible en 2D et 3D) avec la frontière de séparation, les plots de densité des lois Normales et également les vecteurs propres de la droite de projection.

## 16.5. Exercice 168.: Analyse en composantes principales (ACP) paramétrique

Minitab® Statistical Software 15.1.1.0

Soit les données suivantes concernant des fleurs (mêmes données "Iris de Fisher" que dans le cours théorique):

Fleur n°	$x_1$	$x_2$	$x_3$
1	5.1	3.5	1.4
2	4.9	3.0	1.4
3	4.7	3.2	1.3
4	4.6	3.1	1.5
5	5.0	3.6	1.4
6	7.0	3.2	4.7
7	6.4	3.2	4.5
8	6.9	3.1	4.9
9	5.5	2.3	4.0
10	6.5	2.8	4.6
11	6.3	3.3	6.0
12	5.8	2.7	5.1
13	7.1	3.0	5.9
14	6.3	2.9	5.6
15	6.5	3.0	5.8

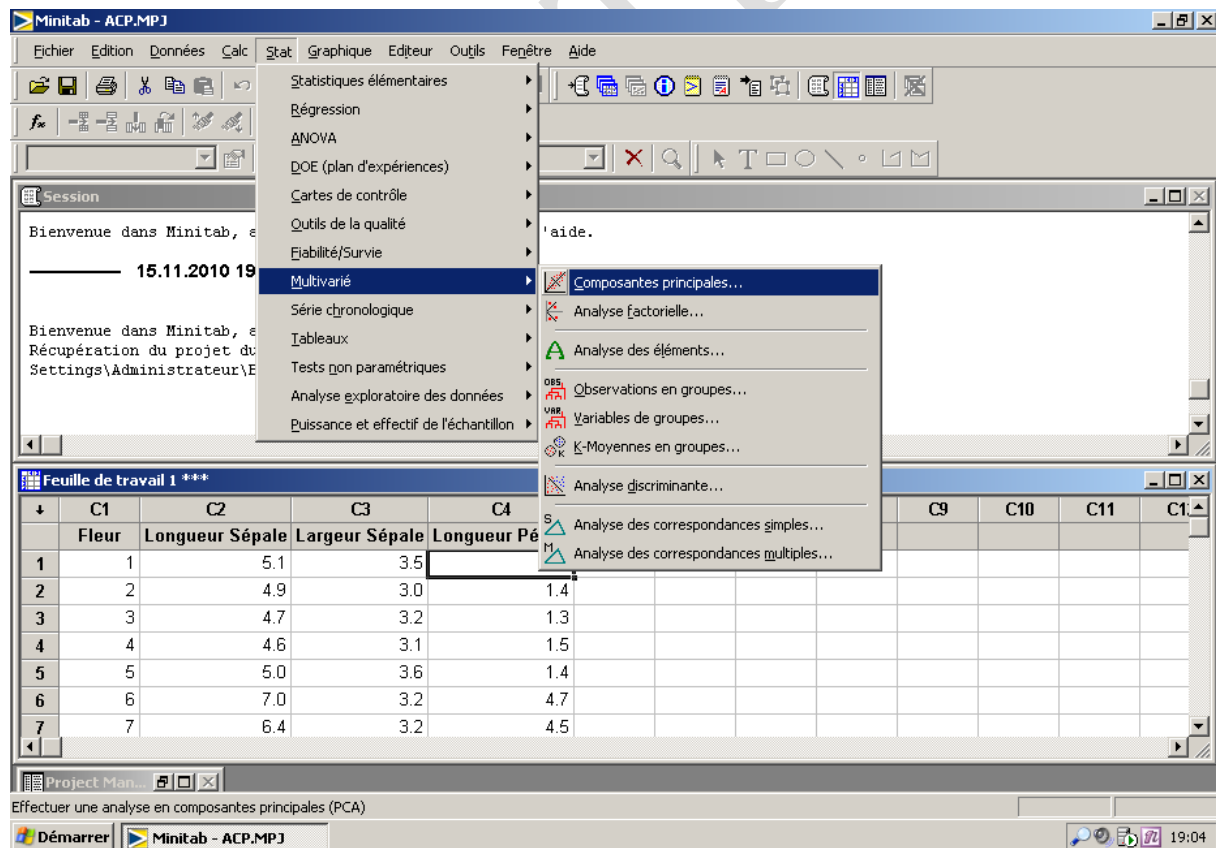
Nous souhaiterions effectuer une A.C.P. (paramétrique non régularisée) avec Minitab® Statistical Software (car Microsoft Excel n'a aucun outil intégré par défaut) afin d'identifier des regroupements possibles et déterminez avec une approche de corrélation les valeurs propres aussi des différentes composantes

Nous créons un nouveau projet et y mettons les données:

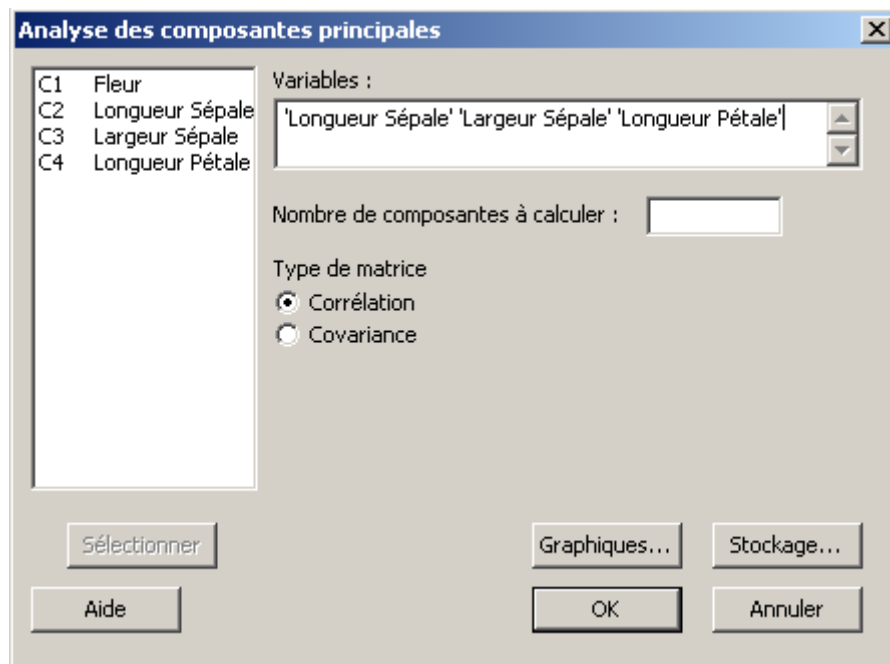
↓	C1	C2	C3	C4
	Fleur	Longueur Sépale	Largeur Sépale	Longueur Pétale
1	1	5.1	3.5	1.4
2	2	4.9	3.0	1.4
3	3	4.7	3.2	1.3
4	4	4.6	3.1	1.5
5	5	5.0	3.6	1.4
6	6	7.0	3.2	4.7
7	7	6.4	3.2	4.5
8	8	6.9	3.1	4.9
9	9	5.5	2.3	4.0
10	10	6.5	2.8	4.6
11	11	6.3	3.3	6.0
12	12	5.8	2.7	5.1
13	13	7.1	3.0	5.9
14	14	6.3	2.9	5.6
15	15	6.5	3.0	5.8

ou nous ouvrons le fichier *ACP.mpj*.

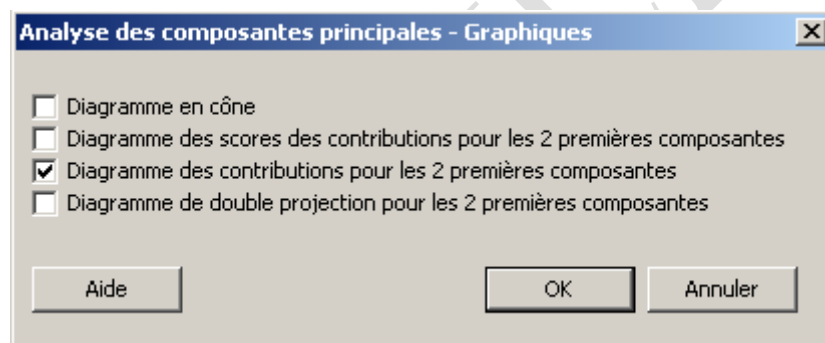
Ensuite, nous allons dans le menu **Stat/Multivarié/Composantes principales...**:



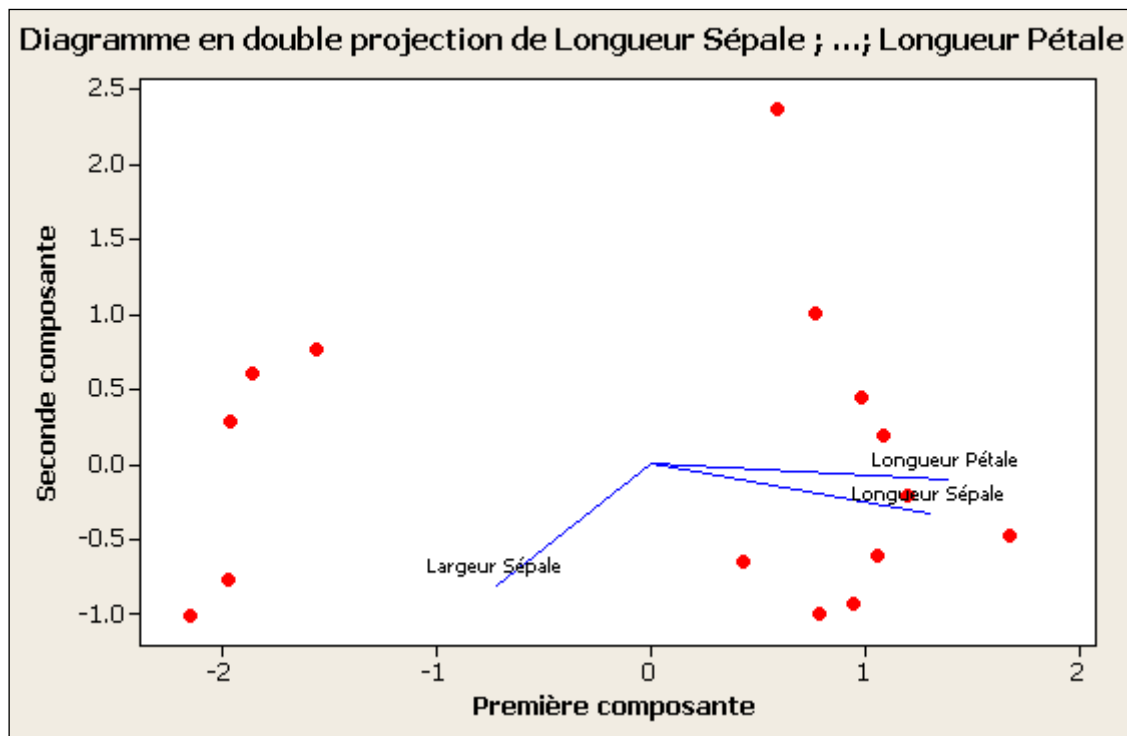
Ensuite nous prenons les données comme indiqué ci-dessous:



et en cliquant sur le bouton **Graphiques...**:



Nous validons le tout en cliquant deux fois sur **OK**:



Dans le graphique ci-dessus nous ne retrouvons ni les valeurs calculées à la main dans le cours théorique ni celles retournées par R ou MATLAB.

En ce qui concerne les valeurs propres, Minitab® Statistical Software nous donne dans la fenêtre d'exécution:

#### Analyse en composantes principales : Longueur Sép; Largeur Sépa; Longueur Pét

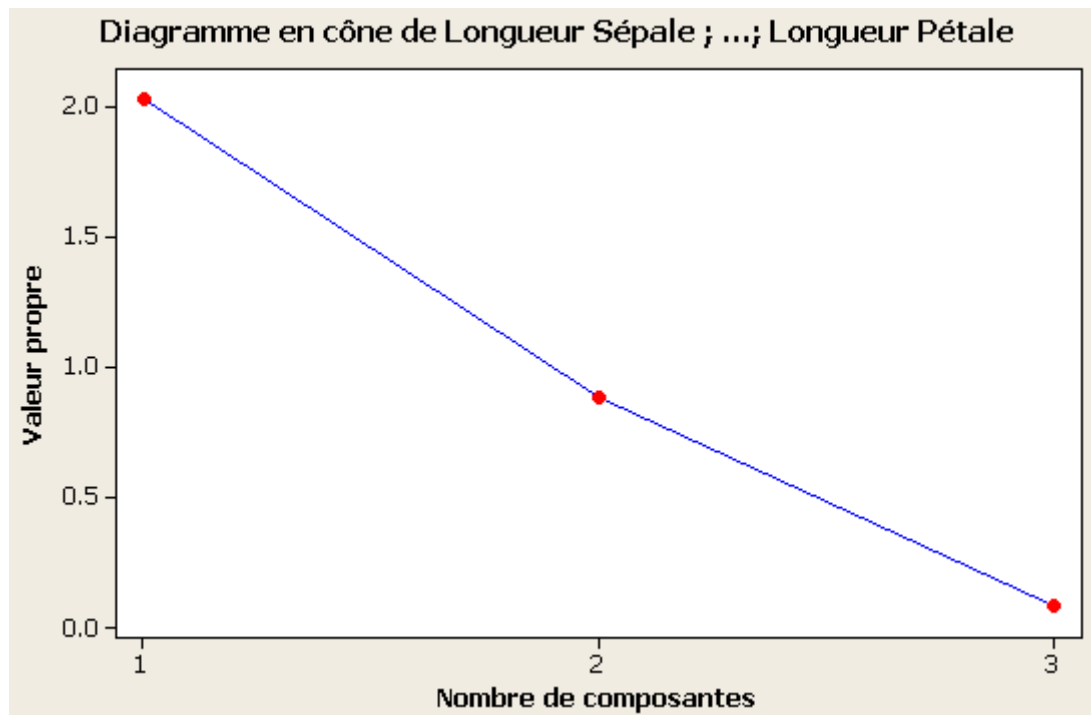
Analyse des valeurs et vecteurs propres de la matrice de corrélation

Valeur propre	2.0302	0.8846	0.0853
Proportion	0.677	0.295	0.028
Cumulatif	0.677	0.972	1.000

Variable	PC1	PC2	PC3
Longueur Sépale	0.641	-0.380	0.667
Largeur Sépale	-0.353	-0.917	-0.184
Longueur Pétale	0.682	-0.117	-0.722

avec le graphique<sup>6</sup> (dont l'utilité est tout discutable) pour représenter l'amplitude des valeurs propres en fonction du numéro de la valeur propre:

<sup>6</sup> Pour rappel il s'agit d'un "scree plot" (en anglais "scree" ce sont les éboulis qui descendent des montagnes).

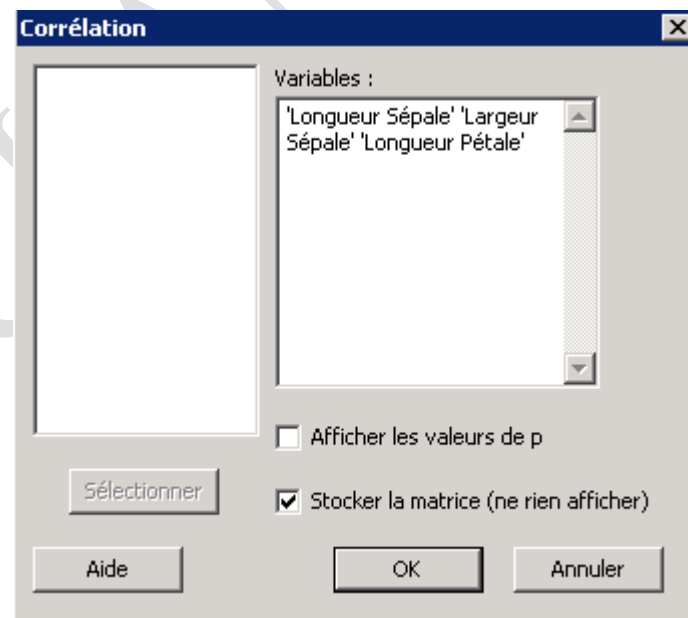


Nous pouvons vérifier que nous obtenons la même matrice de corrélation que dans le cours théorique avec le calcul fait dans Microsoft Excel ainsi que les valeurs données par Minitab ci-dessus à la main. Pour cela, nous allons dans le menu **Stat/Statistiques élémentaires/Corrélation...**:

Calculer les coefficients de corrélation du moment produit de Pearson

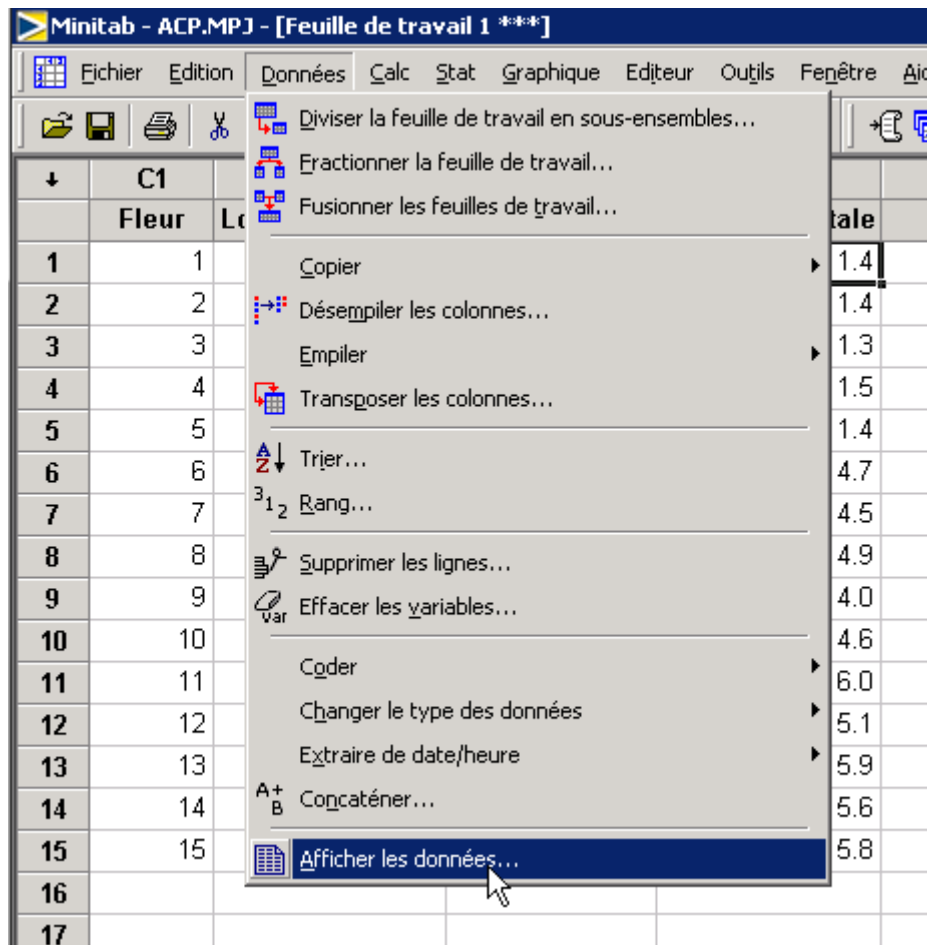
	C1	C2
	Fleur	Longueur Sépale
1	1	5.1
2	2	4.7
3	3	4.9
4	4	4.4
5	5	5.0
6	6	7.0
7	7	6.4
8	8	6.3
9	9	5.8
10	10	6.6
11	11	6.3
12	12	5.8
13	13	7.1
14	14	6.3
15	15	6.5
16		
17		
18		
19		
20		
21		

Nous y prenons:

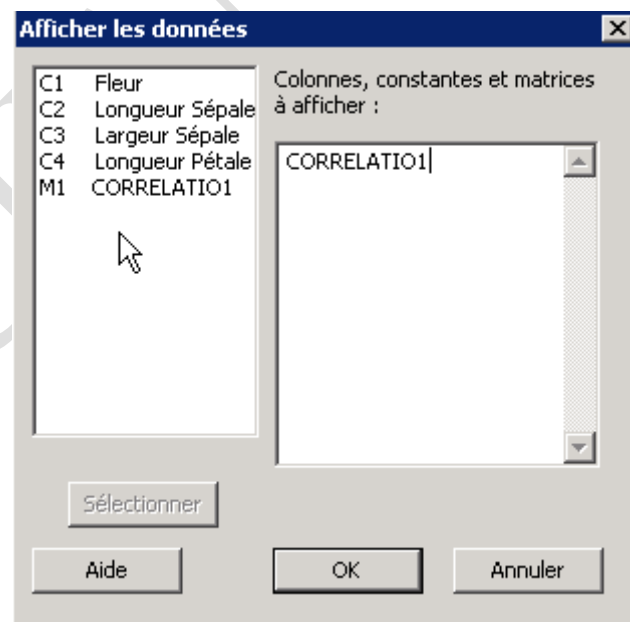


Nous validons par **OK**. Ensuite, nous allons dans le menu **Données/Afficher les données...**:





et nous prenons:



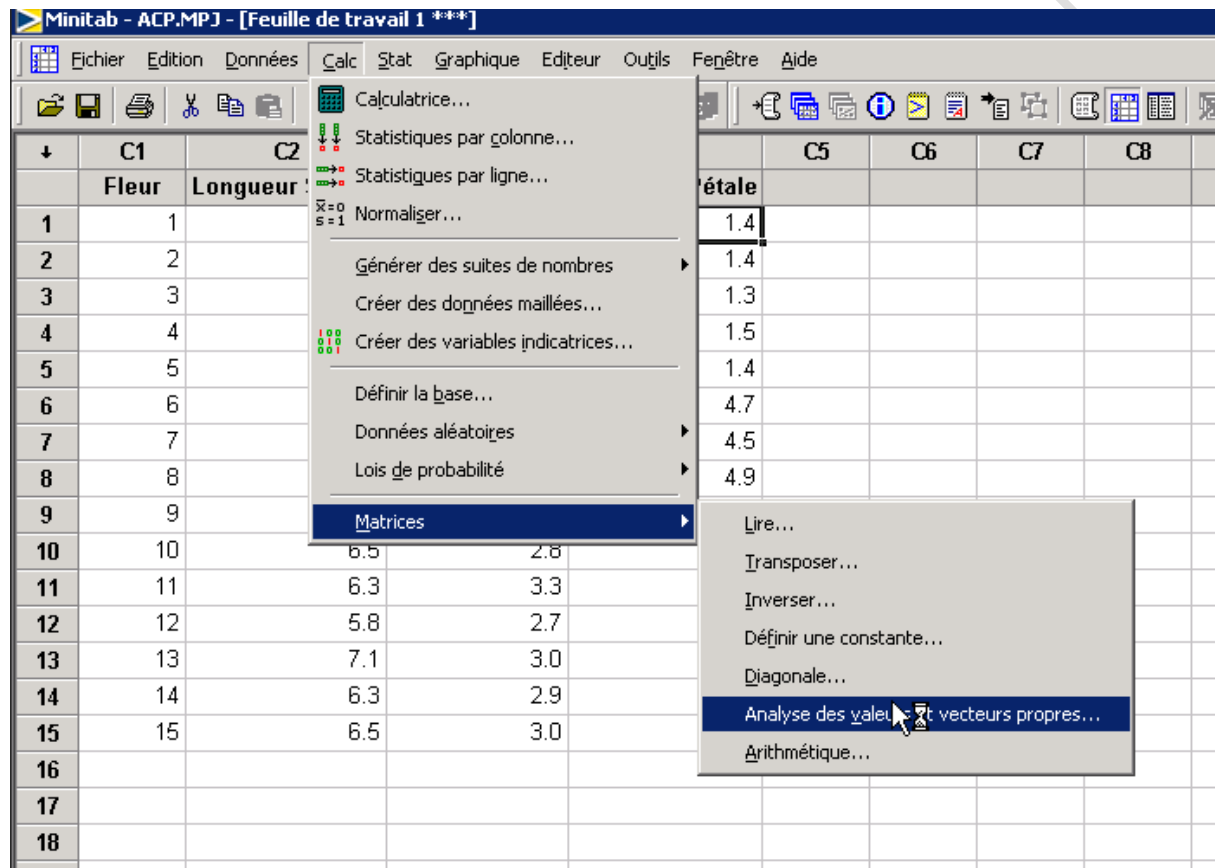
Nous retrouvons après validation par **OK** la matrice de corrélation calculée dans le cours théorique:

### Affichage des données

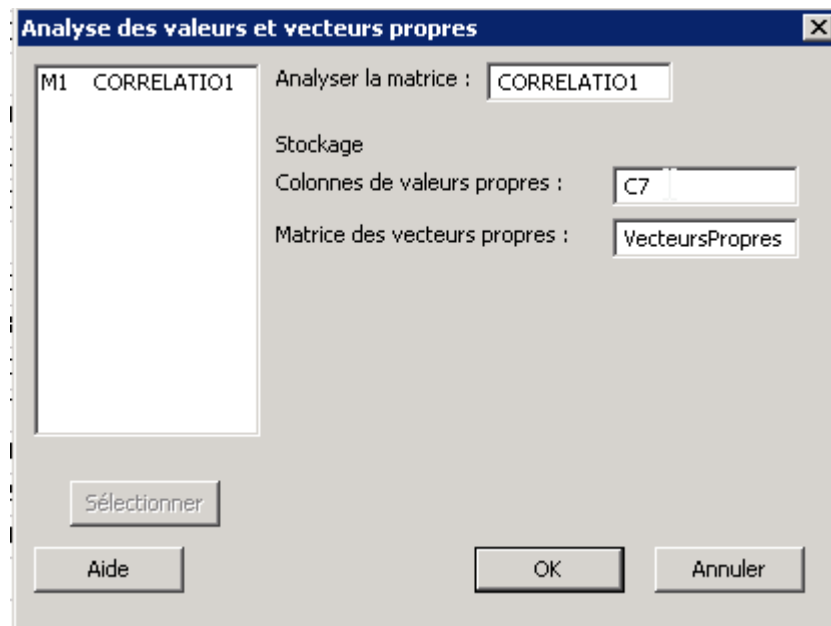
Matrice CORRELATI01

1.00000	-0.16090	0.88545
-0.16090	1.00000	-0.38179
0.88545	-0.38179	1.00000

Demandons à Minitab de nous sortir les valeurs propres et vecteurs propres de la matrice de corrélation. Pour cela, nous allons dans le menu **Calc/Matrices/Analyse des valeurs et vecteurs propres...**:



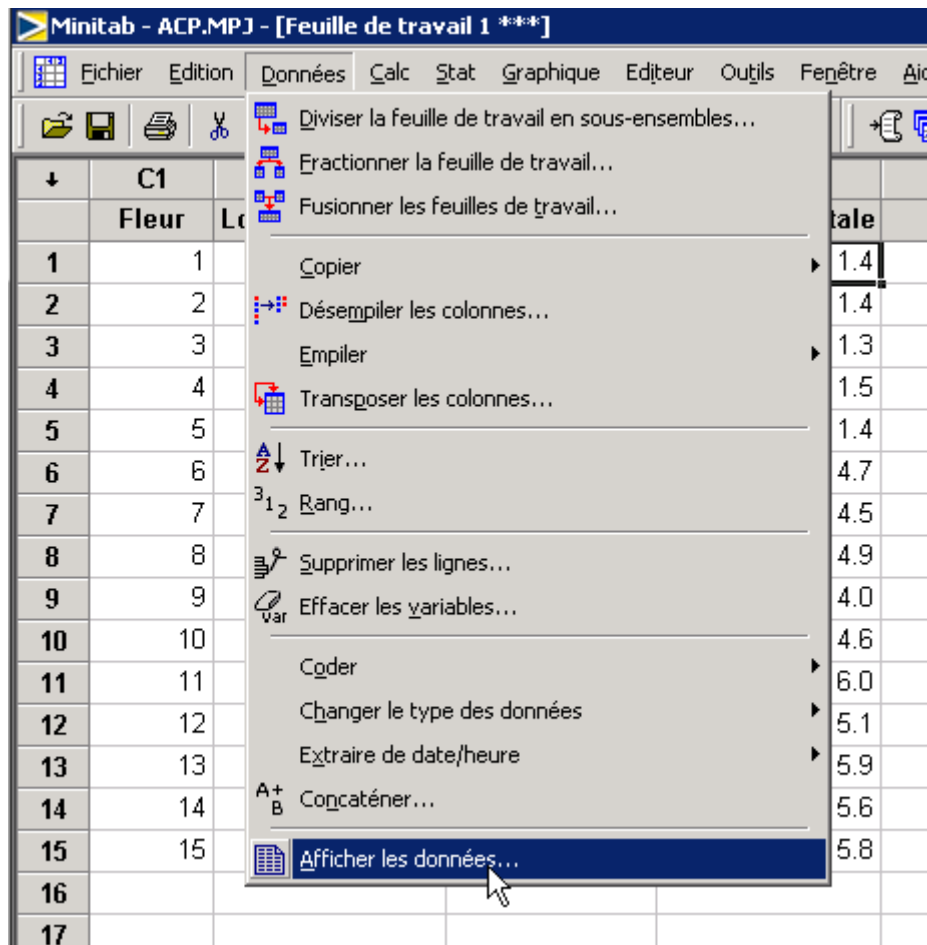
et nous prenons:



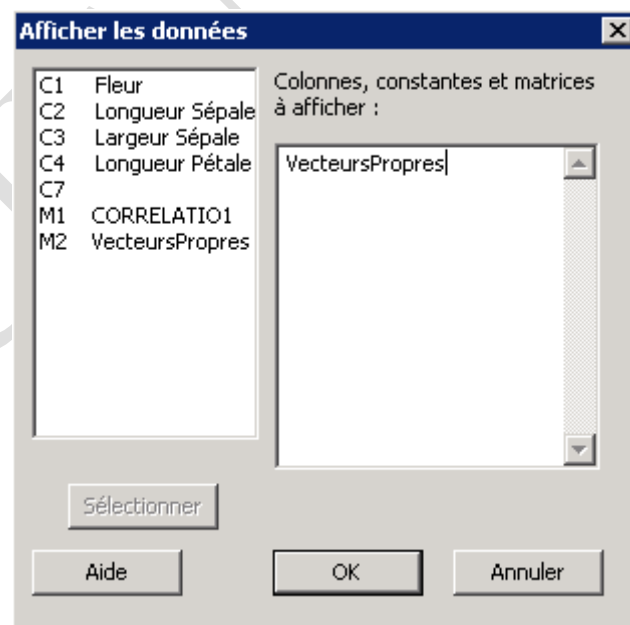
Nous validons par **OK** pour obtenir dans un premier temps dans la feuille:

	C1	C2	C3	C4	C5	C6	C7
	Fleur	Longueur Sépale	Largeur Sépale	Longueur Pétale			
1	1	5.1	3.5	1.4			2.03018
2	2	4.9	3.0	1.4			0.88455
3	3	4.7	3.2	1.3			0.08526
4	4	4.6	3.1	1.5			
5	5	5.0	3.6	1.4			
6	6	7.0	3.2	4.7			
7	7	6.4	3.2	4.5			
8	8	6.9	3.1	4.9			

Ce qui est conforme à ce que nous avons calculé aussi à la main dans le cours théorique et que Minitab nous a déjà donné plus haut dans le rapport complet de l'analyse en composantes principales. Ensuite, voyons les valeurs des vecteurs propres (système de trois équations à trois inconnues que nous n'avons pas résolu dans le cours théorique car trop trivial). Pour les afficher nous retournons dans le menu **Données/Afficher les données...**:



et nous prenons:



en validant par **OK**, nous avons dans la fenêtre de sessions les trois vecteurs propres mis les uns à côté des autres:

**Affichage des données**

Matrice VecteursPropres

-0.641002	0.380302	-0.666698
0.352751	0.917415	0.184163
-0.681677	0.117129	0.722217

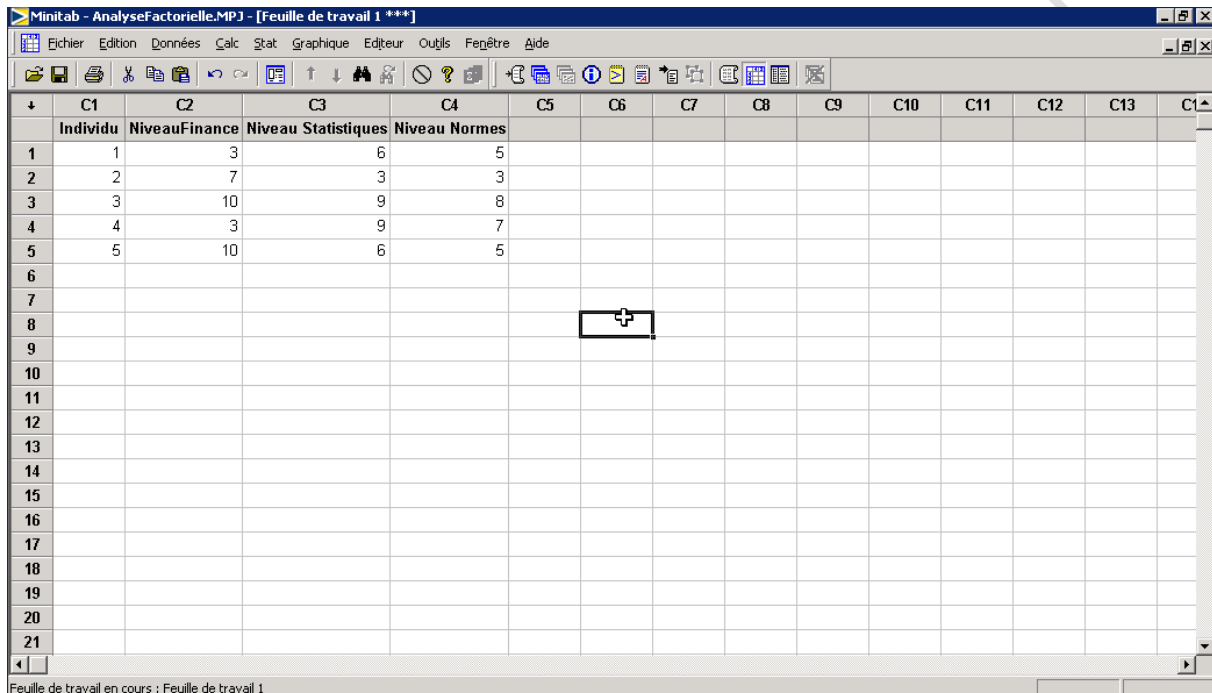
et nous retrouvons dans les trois vecteurs de base (propres) du plan principal.

ÉCHANTILLON

## 16.6. Exercice 169.: Analyse factorielle exploratoire (méthode ACP sans rotation!)

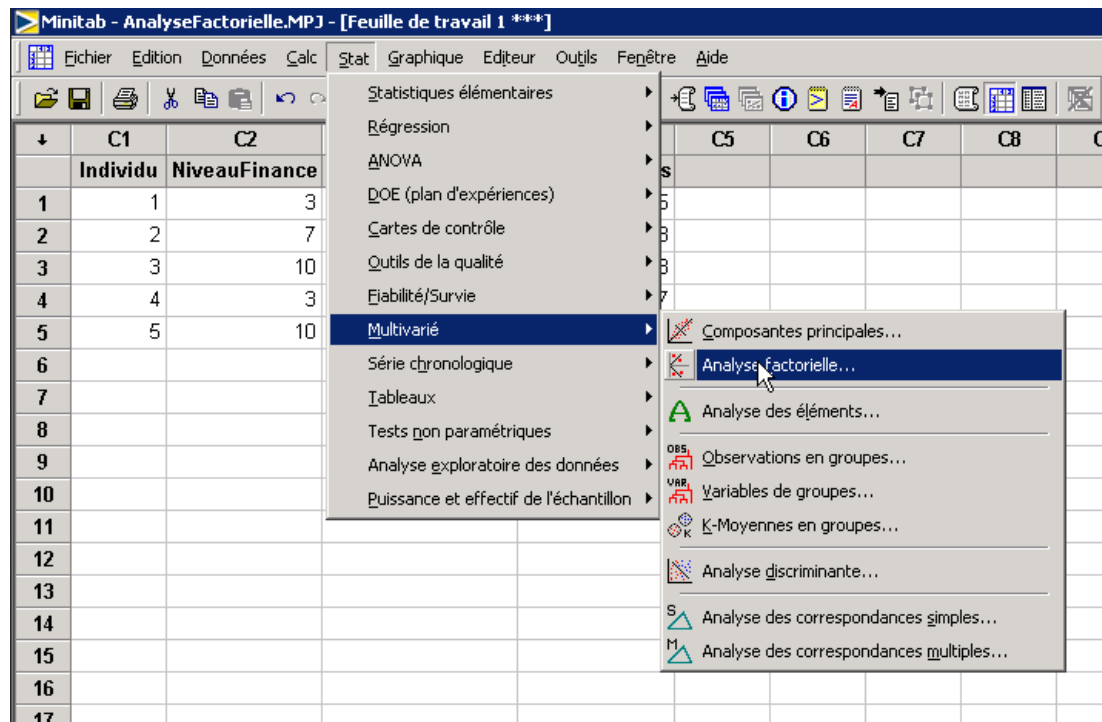
Minitab® Statistical Software 15.1.1.0

Le but va être ici de vérifier que les résultats calculés à la main (et donc les conclusions y relatives) sont conformes aux démonstrations mathématiques faites dans le cours avec l'exemple particulier que nous avons choisi:

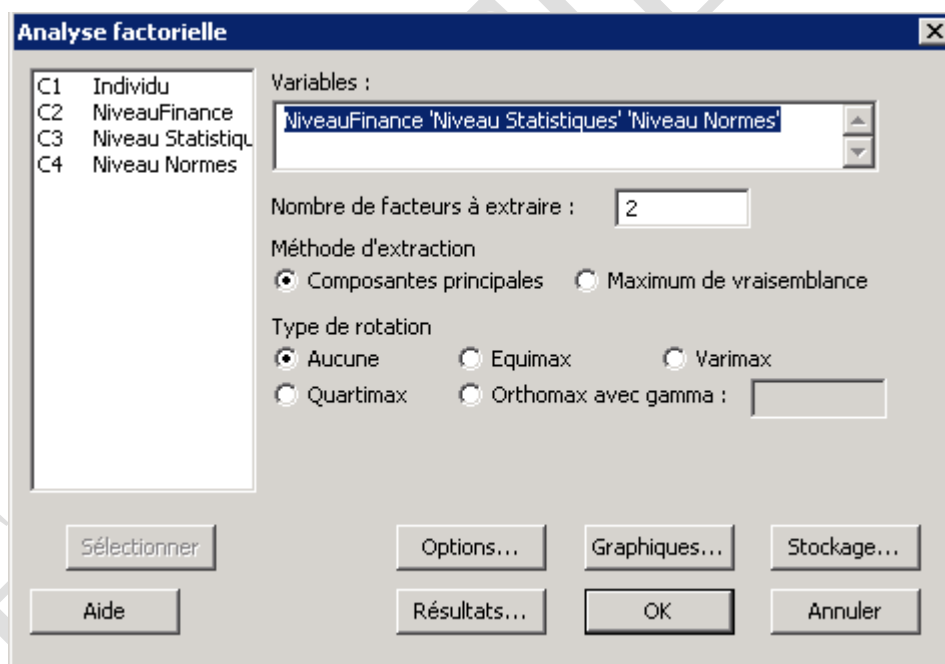


	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Individu	NiveauFinance	Niveau Statistiques	Niveau Normes										
1	1	3	6	5										
2	2	7	3	3										
3	3	10	9	8										
4	4	3	9	7										
5	5	10	6	5										
6														
7														
8														
9														
10														
11														
12														
13														
14														
15														
16														
17														
18														
19														
20														
21														

Ensuite, nous allons dans le menu **Stat/Multivarié/Analyse factorielle...**:

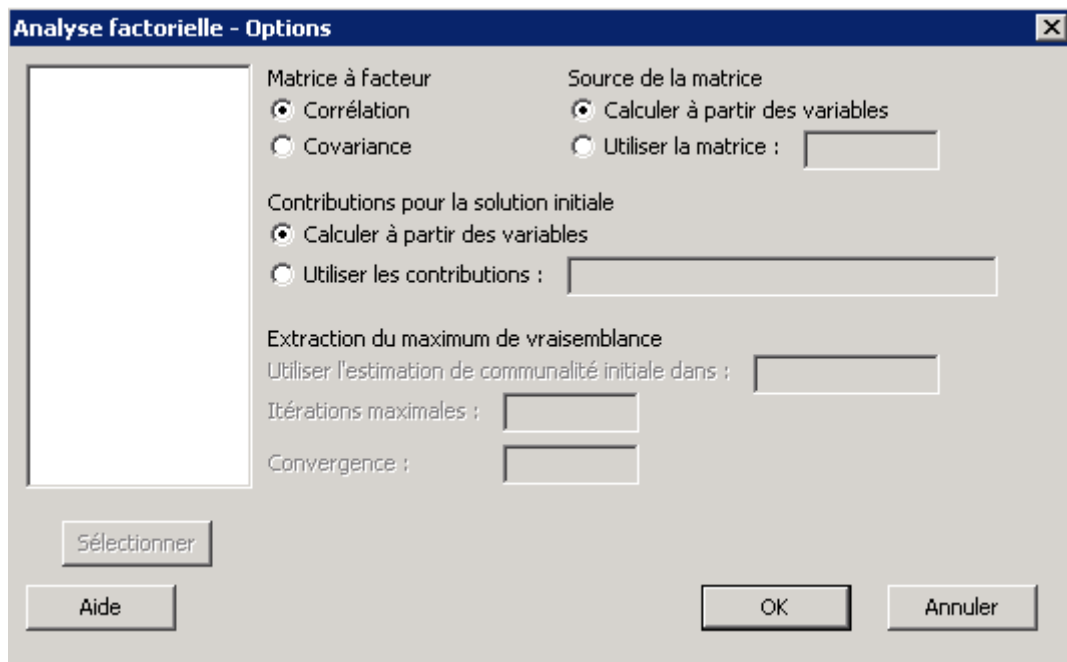


Dans la boîte de dialogue qui apparaît, nous mettons:

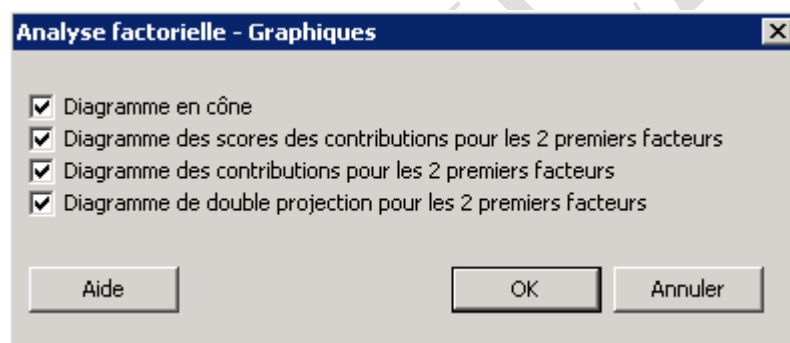


Comme vous pouvez le voir, nous allons utiliser la technique des **Composantes principales** et **Aucune** rotation.

Nous cliquons ensuite sur **Options...** pour prendre:



et dans **Graphiques...**, nous prenons:



et nous validons le tout par **OK**. Nous obtenons alors d'abord dans la fenêtre de session:



**Analyse factorielle : NiveauFinance; Niveau Statistiques; Niveau Normes**

Analyse factorielle du composant principal de la matrice de corrélation

Saturations de facteurs et communalités sans rotations

Variable	Facteur1	Facteur2	Communalité
NiveauFinance	0.030	1.000	1.000
Niveau Statistiques	0.994	-0.082	0.995
Niveau Normes	0.996	0.051	0.995
Variance	1.9815	1.0083	2.9898
% var	0.660	0.336	0.997

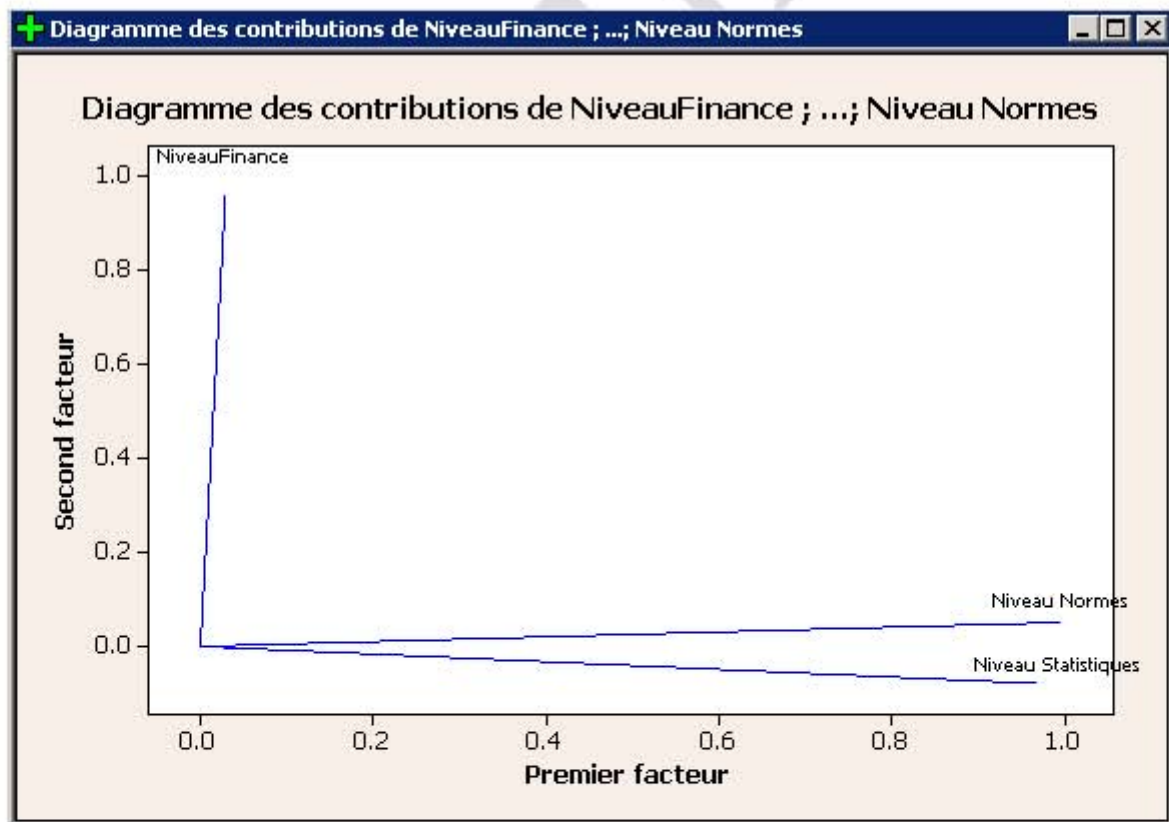
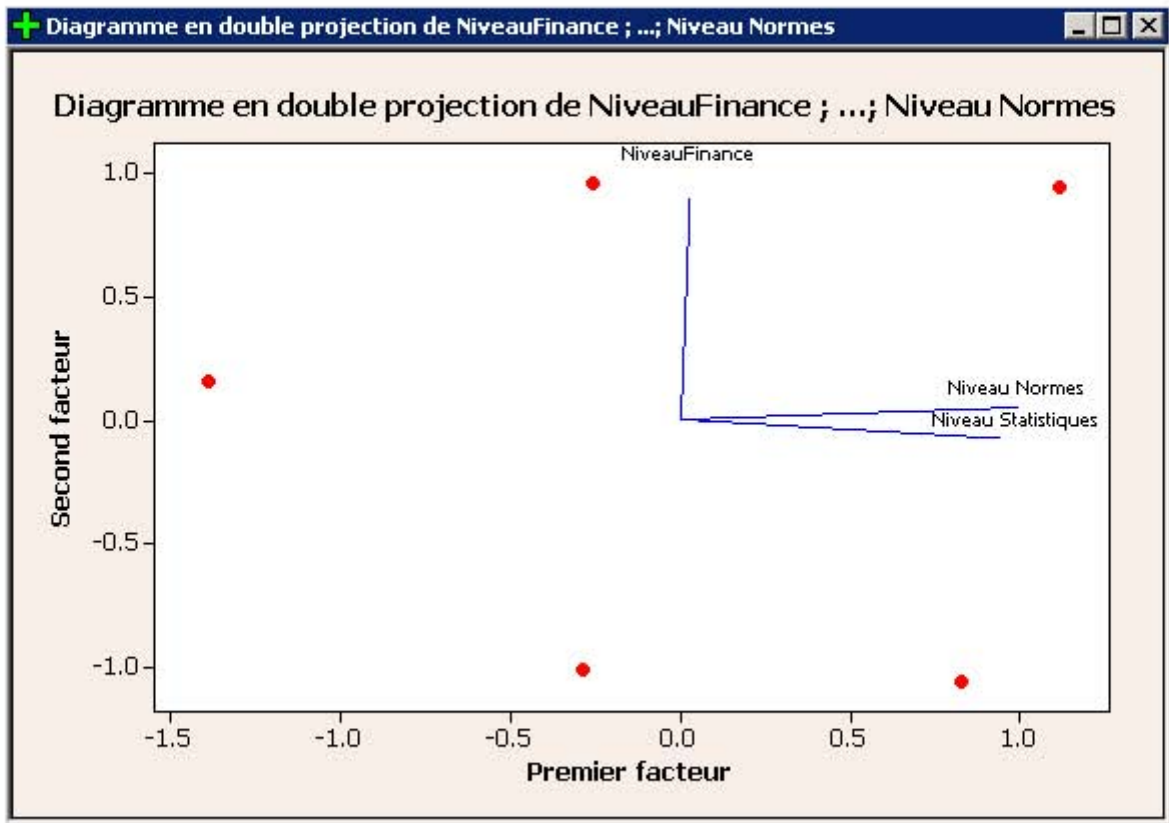
Coefficients des scores de facteur

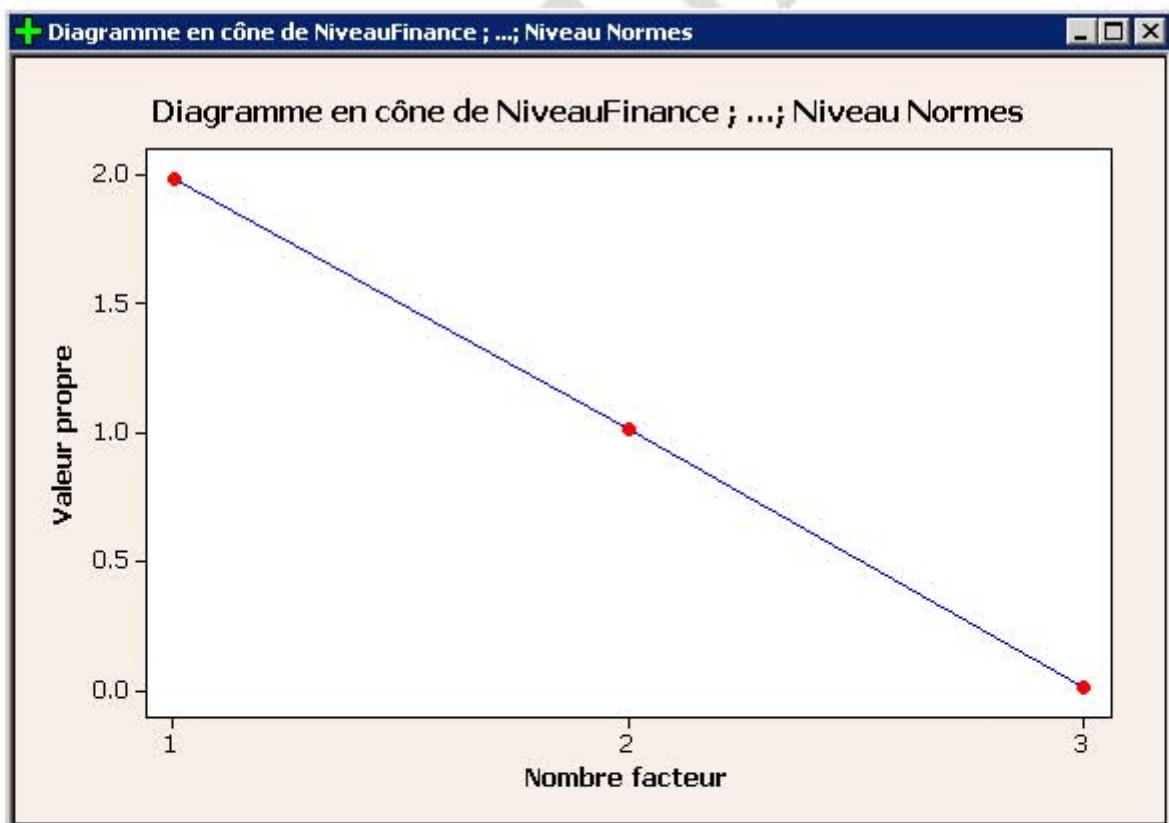
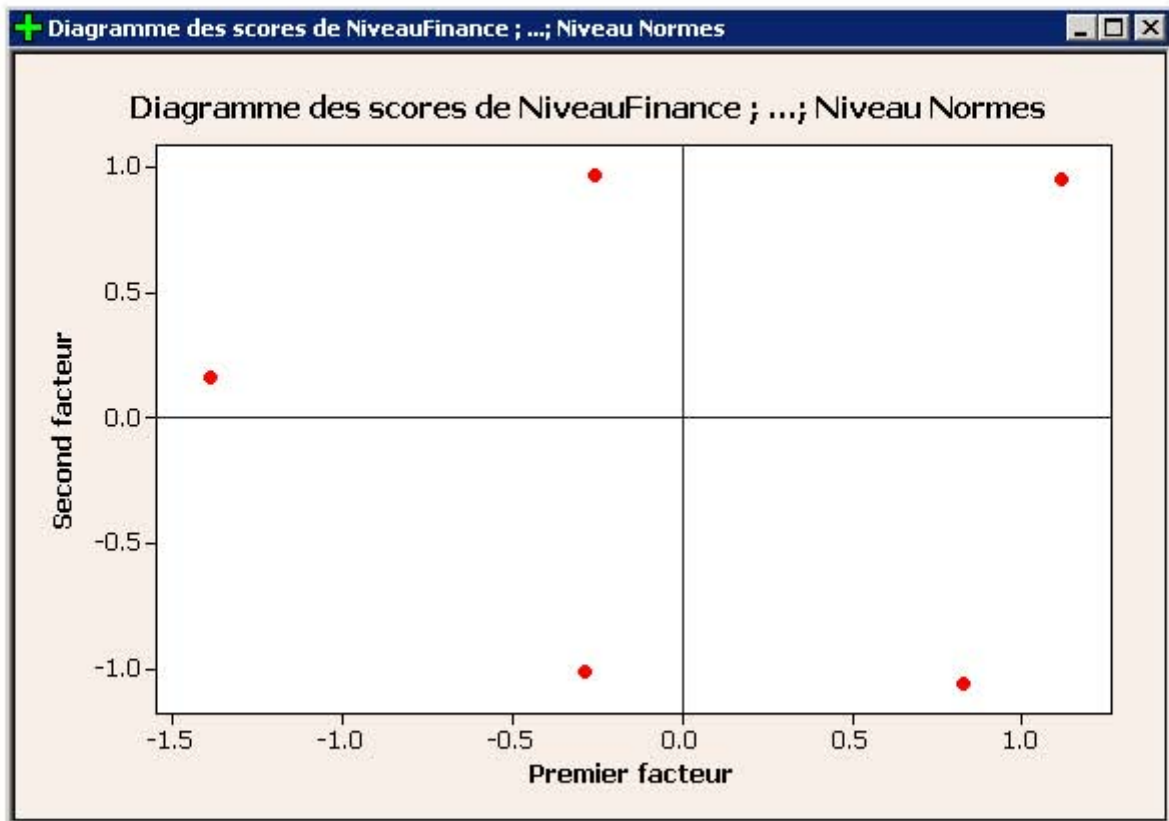
Variable	Facteur1	Facteur2
NiveauFinance	0.015	0.991
Niveau Statistiques	0.502	-0.081
Niveau Normes	0.503	0.051

Nous retrouvons bien les coefficients (loadings/saturations) des 2 facteurs du modèle linéaire implicites ainsi que la communalité calculé à la main ainsi que la variance totale qu'ils représentent évidemment directement et le % du total de la variance expliquée.

Dessous, nous retrouvons les coefficients des scores de facteur dont nous avons mentionné dans le cours théorique l'utilité toute relative...

et au niveau des graphiques, nous obtenons exactement les mêmes que si nous faisons un analyse par composantes principale des données (graphiques dont l'utilité est aussi tout à fait discutable)





L'analyse en composantes principale des mêmes données nous sort le tableau suivant dans la fenêtre de session:

### Analyse en composantes principales : NiveauFinanc; Niveau Stati; Niveau Norme

Analyse des valeurs et vecteurs propres de la matrice de corrélation

Valeur propre	1.9815	1.0083	0.0102
Proportion	0.660	0.336	0.003
Cumulatif	0.660	0.997	1.000

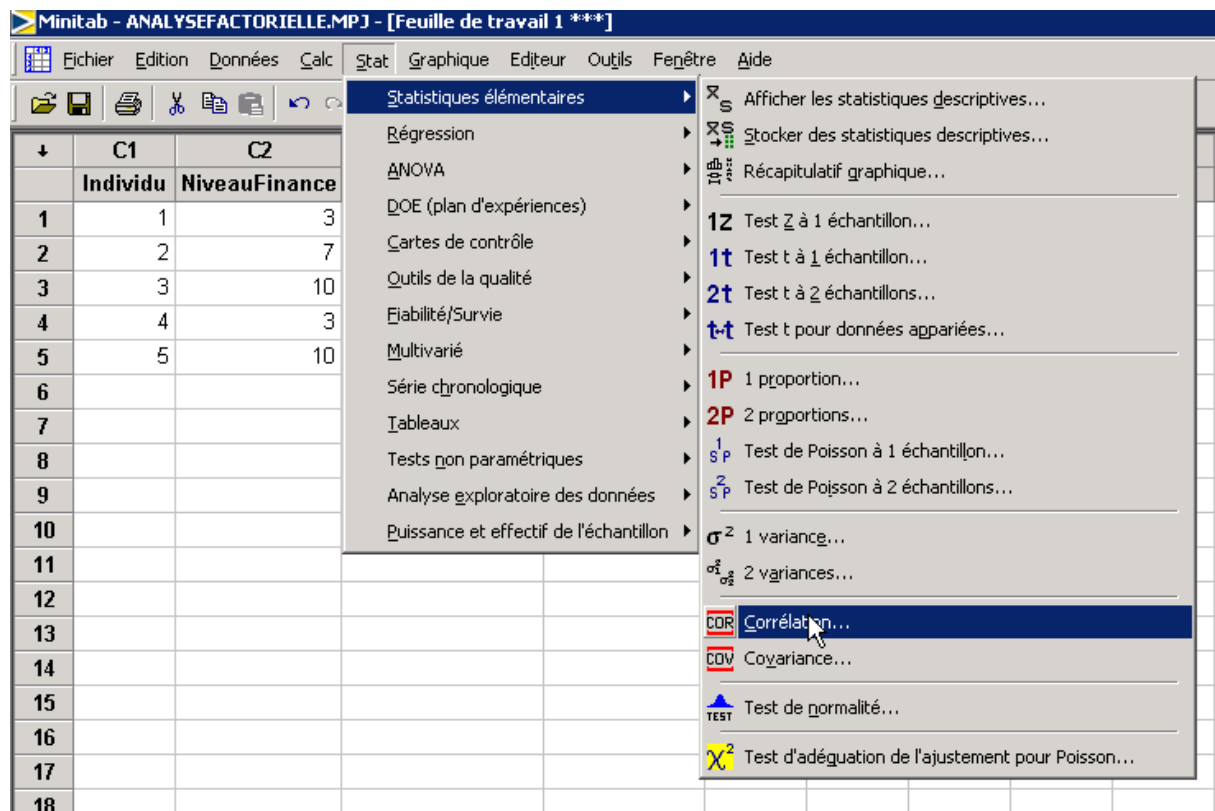
Variable	PC1	PC2
NiveauFinance	0.021	0.995
Niveau Statistiques	0.706	-0.081
Niveau Normes	0.708	0.051

Nous retrouvons bien les valeurs propres utilisées dans le calcul à la main dans le cours théorique et que l'analyse factorielle utilise pour déterminer les coefficients (loadings/saturations).

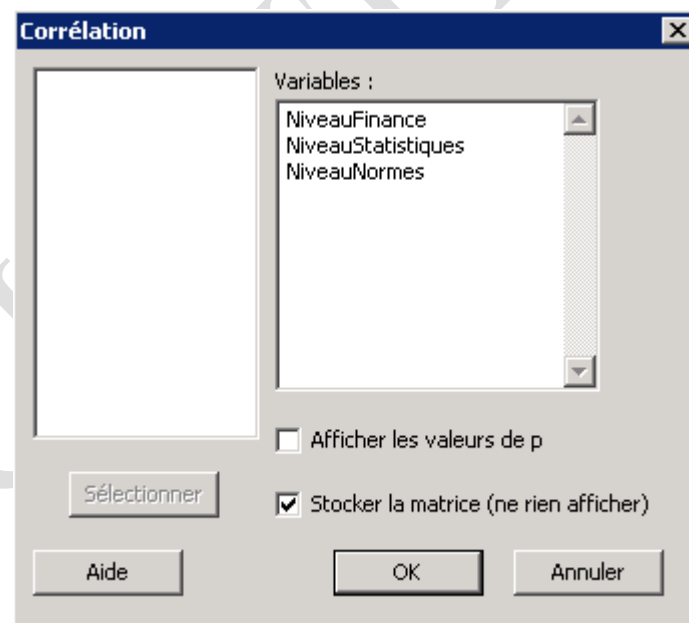
Nous pouvons reproduire le résultat aussi en plusieurs étapes dans Minitab. Nous partons toujours de:

	C1	C2	C3	C4	C5
	Individu	NiveauFinance	NiveauStatistiques	NiveauNormes	
1	1	3	6	5	
2	2	7	3	3	
3	3	10	9	8	
4	4	3	9	7	
5	5	10	6	5	
6					
7					
8					
9					
10					

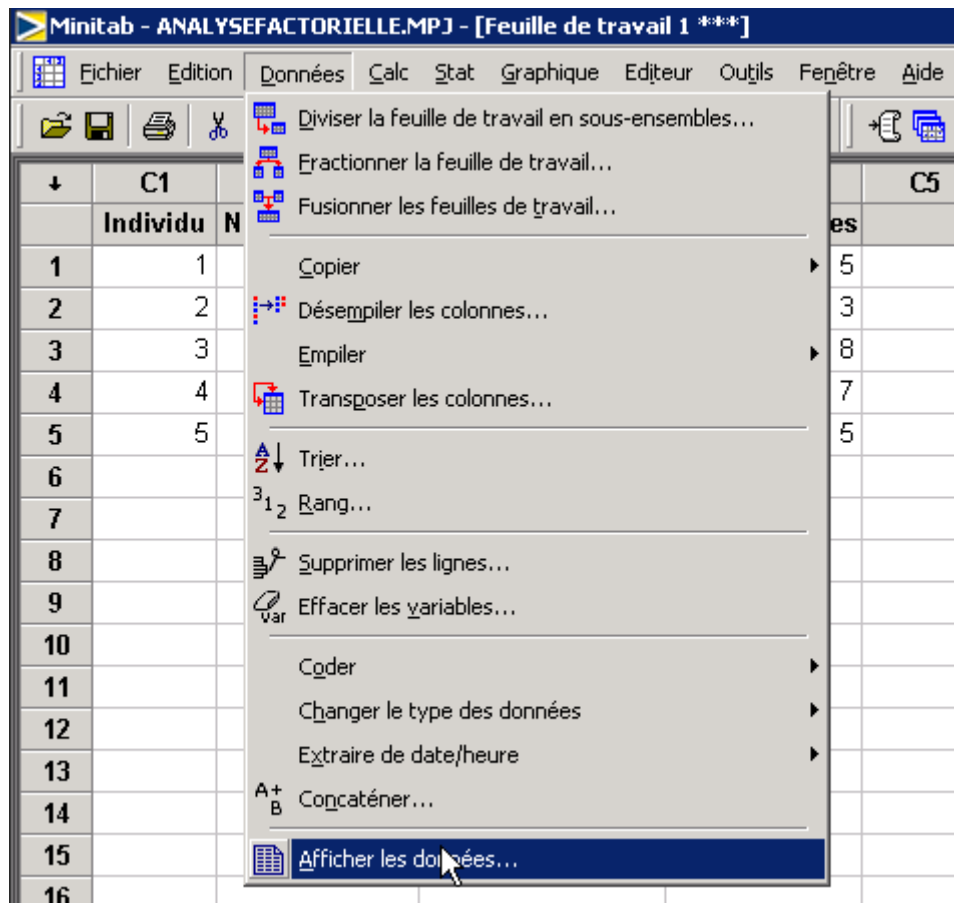
et nous calculons la matrice de corrélation en allant dans le menu **Stat/Statistiques élémentaires/Corrélation...**:



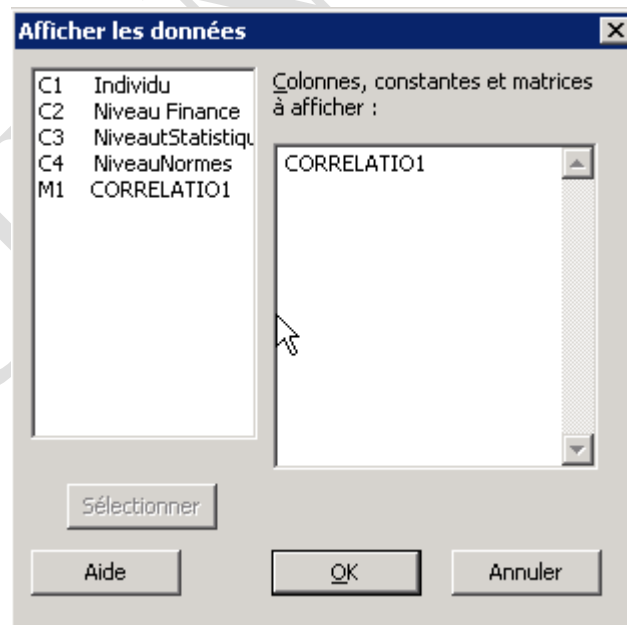
Nous prenons alors:



Et affichons la matrice de corrélation dans la fenêtre de session en allant dans le menu **Données/Afficher les données...**:



ce qui nous amène:



ce qui donne (matrice que nous avons obtenue aussi dans Microsoft Excel lors des calculs à la main suite à la démonstration mathématique du modèle d'analyse factorielle):

### Affichage des données

Matrice CORRELATIO1

1.00000	-0.05112	0.08045
-0.05112	1.00000	0.98102
0.08045	0.98102	1.00000

Ensuite, il nous faut déterminer les valeurs et vecteurs propres de la matrice de corrélation. Ce va être beaucoup plus simple que de résoudre à la main par radicaux de l'équation du troisième degré du déterminant de la matrice dont la forme a déjà été démontrée dans le cours d'Algèbre Linéaire.

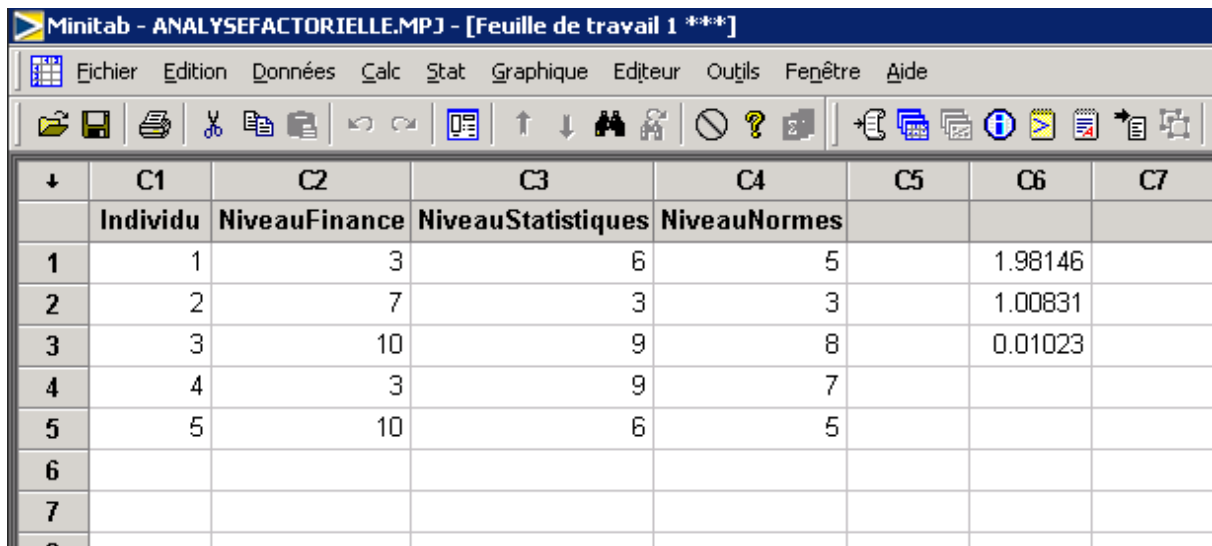
Pour cela, nous allons dans le menu **Matrice/Analyse des valeurs et vecteurs propres...**:

The screenshot shows the Minitab interface with the 'Matrices' menu open. The menu options are:

- Lire...
- Transposer...
- Inverser...
- Définir une constante...
- Diagonale...
- Analyse des valeurs et vecteurs propres...** (highlighted)
- Arithmétique...

The background shows a worksheet with columns C1 (Individu) and C2 (NiveauFin) containing data for 17 rows.

ce qui donne les trois valeurs propres recherchées:



The screenshot shows the Minitab software interface with a menu bar (Fichier, Edition, Données, Calc, Stat, Graphique, Éditeur, Outils, Fenêtre, Aide) and a toolbar. Below the toolbar is a data table with the following content:

↓	C1	C2	C3	C4	C5	C6	C7
	Individu	NiveauFinance	NiveauStatistiques	NiveauNormes			
1	1	3	6	5		1.98146	
2	2	7	3	3		1.00831	
3	3	10	9	8		0.01023	
4	4	3	9	7			
5	5	10	6	5			
6							
7							
8							

Nous pouvons contrôler les valeurs propres données par Minitab avec Maple (ou avec Tanagra et R):

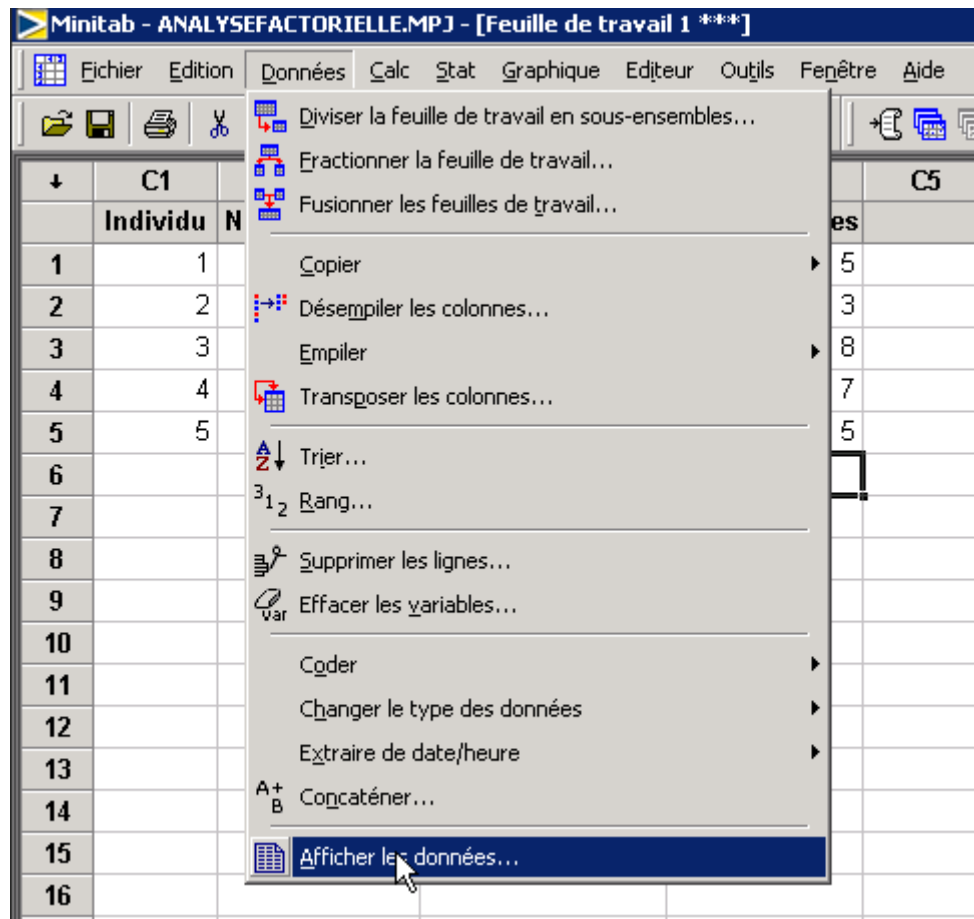
```
>A:=array([[1,-0.05112,0.08045],[-0.05112,1,0.98102],[0.08045,0.98102,1]]);
>evalf(Eigenvals(A));
```

Ce qui donne bien:

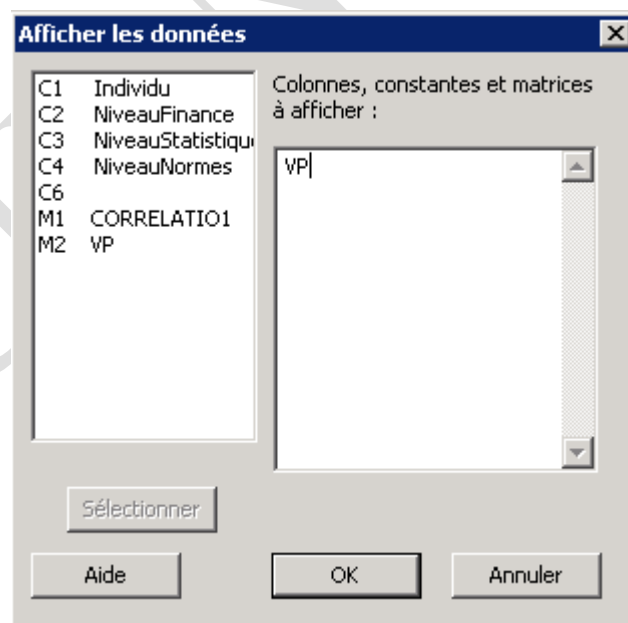
[\[.010233253, 1.008306522, 1.981460230\]](#)

Pour afficher dans Minitab la matrices des vecteurs propres  $S$ , nous allons ensuite à nouveau dans le menu **Données/Afficher les données...**:





et nous prenons:



Nous validons par **OK** pour obtenir la matrice des vecteurs propres données par Minitab:

**Affichage des données**

Matrice VP

```

0.021221  0.995383  0.093602
0.706236 -0.081193  0.703306
0.707659  0.051181 -0.704698

```

Nous pouvons vérifier cela avec Maple (ou avec Tanagra et R):

```

> A:=array([[1,-0.05112,0.08045],[-0.05112,1,0.98102],[0.08045,0.98102,1]]);
> lambda := evalf(Eigenvals(A,vecs));
> print(vecs);

```

Ce qui donne:

$$\begin{bmatrix} -.09360383077 & -.9953833219 & .02122180483 \\ -.7033053770 & .08119452457 & .7062357952 \\ .7046984254 & -.05118096633 & .7076585603 \end{bmatrix}$$

Nous voyons donc qu'il y a un souci dans les signes avec Minitab (mais bon ce sont les mêmes que dans R)! Mais cela n'a pas d'impact pour la suite.

Pour la suite, nous allons continuer avec Microsoft Excel car c'est pénible avec Minitab. Nous avons donc pour l'instant la matrice diagonale des valeurs propres rangées dans l'ordre décroissant et la matrice des vecteurs propres:

	A	B	C
1	Matrice diagonale valeurs propres		
2	1.98146	0	0
3	0	1.00831	0
4	0	0	0.01023
5			
6	Matrice des vecteurs propres		
7	0.021221	0.995383	0.093602
8	0.706236	-0.081193	0.703306
9	0.707659	0.05118	-0.704698

Nous prenons la racine des composantes de la diagonale de la première matrice. Ce qui donne:

	A	B	C
1	Matrice diagonale valeurs propres		
2	1.40764342	0	0
3	0	1.0041464	0
4	0	0	0.10114346
5			
6	Matrice des vecteurs propres		
7	0.021221	0.995383	0.093602
8	0.706236	-0.081193	0.703306
9	0.707659	0.05118	-0.704698

et nous multiplions ces deux matrices (ce qui équivaut en réalité à multiplier chaque vecteur propre par la racine carrée de la valeur propre qui l'a engendré!) pour obtenir:

	A	B	C
1	Matrice diagonale valeurs propres		
2	1.40764342	0	0
3	0	1.0041464	0
4	0	0	0.10114346
5			
6	Matrice des vecteurs propres		
7	0.021221	0.995383	0.093602
8	0.706236	-0.081193	0.703306
9	0.707659	0.05118	-0.704698
10			
11	Produit:		
12	0.0298716	0.99951026	0.00946723
13	0.99412846	-0.0815297	0.0711348
14	0.99613154	0.05139221	-0.0712756

et nous retrouvons donc bien les coefficients des facteurs du modèle linéaire! Bingo!

Soit avec les formules explicites:

	A	B	C
1	Matrice diagonale valeurs propres		
2	=RACINE(1.98146)	0	0
3	0	=RACINE(1.00831)	0
4	0	0	=RACINE(0.01023)
5			
6	Matrice des vecteurs propres		
7	0.021221	0.995383	0.093602
8	0.706236	-0.081193	0.703306
9	0.707659	0.05118	-0.704698
10			
11	Produit:		
12	=PRODUITMAT(A7:C9;A2:C4)	=PRODUITMAT(A7:C9;A2:C4)	=PRODUITMAT(A7:C9;A2:C4)
13	=PRODUITMAT(A7:C9;A2:C4)	=PRODUITMAT(A7:C9;A2:C4)	=PRODUITMAT(A7:C9;A2:C4)
14	=PRODUITMAT(A7:C9;A2:C4)	=PRODUITMAT(A7:C9;A2:C4)	=PRODUITMAT(A7:C9;A2:C4)

## 16.7. Exercice 170.: Analyse factorielle exploratoire (méthode ACP avec rotation VariMax!)

Minitab® Statistical Software 15.1.1.0

Le but va être ici d'observer si la rotation empirique VariMax de maximisation des variances des lignes par rotations itératives 2D de l'exemple précédent va nous aider à mieux comprendre l'influence des différents coefficients que nous avons obtenus:

### Analyse factorielle : NiveauFinance; Niveau Statistiques; Niveau Normes

Analyse factorielle du composant principal de la matrice de corrélation

Saturations de facteurs et communalités sans rotations

Variable	Facteur1	Facteur2	Communalité
NiveauFinance	0.030	1.000	1.000
Niveau Statistiques	0.994	-0.082	0.995
Niveau Normes	0.996	0.051	0.995
Variance	1.9815	1.0083	2.9898
% var	0.660	0.336	0.997

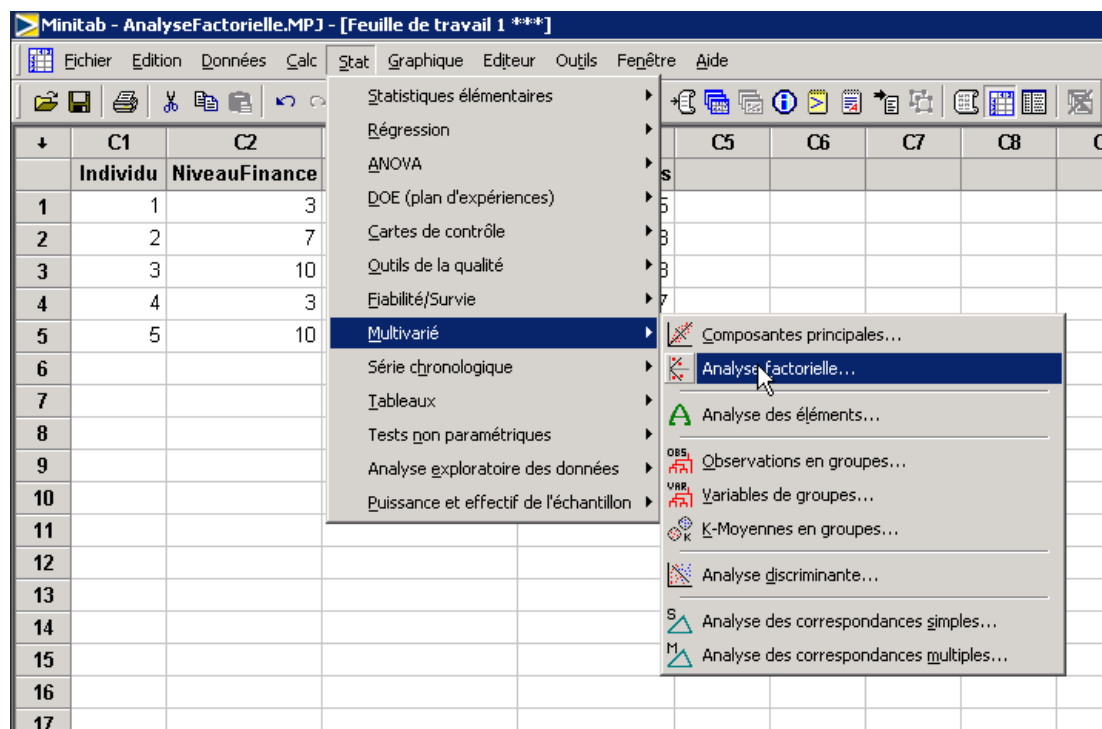
Coefficients des scores de facteur

Variable	Facteur1	Facteur2
NiveauFinance	0.015	0.991
Niveau Statistiques	0.502	-0.081
Niveau Normes	0.503	0.051

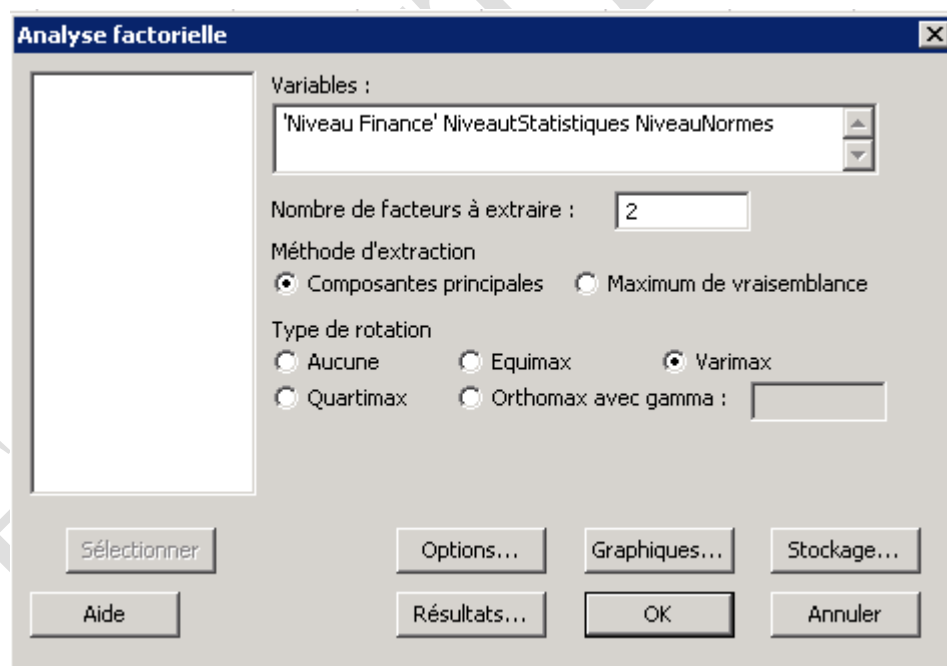
Donc nous repartons des données suivantes:

	C1	C2	C3	C4	C5
	Individu	NiveauFinance	NiveauStatistiques	NiveauNormes	
1	1	3	6	5	
2	2	7	3	3	
3	3	10	9	8	
4	4	3	9	7	
5	5	10	6	5	
6					
7					
8					
9					
10					

Ensuite, nous allons dans le menu **Stat/Multivarié/Analyse factorielle...**:

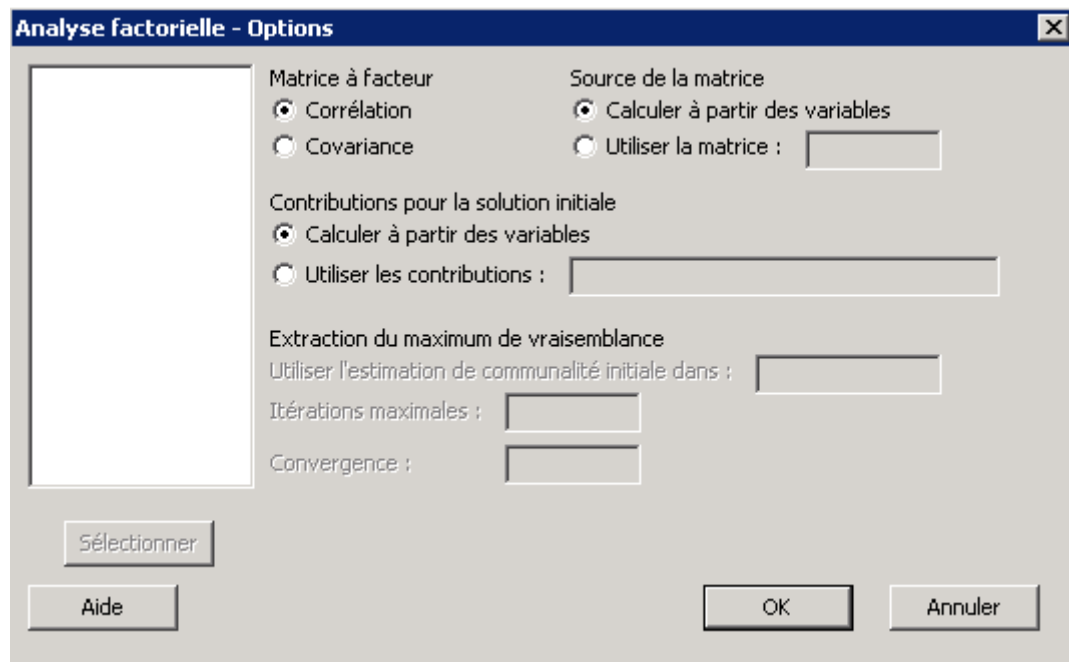


Dans la boîte de dialogue qui apparaît, nous mettons:



Comme vous pouvez le voir, nous allons utiliser la technique des **Composantes principales** et **Aucune** rotation.

Nous cliquons ensuite sur **Options...** pour prendre:



et nous obtenons comme sortie:

Analyse factorielle du composant principal de la matrice de corrélation

Saturations de facteurs et communalités sans rotations

Variable	Facteur1	Facteur2	Communalité
Niveau Finance	0.030	1.000	1.000
NiveautStatistiques	0.994	-0.082	0.995
NiveauNormes	0.996	0.051	0.995
Variance	1.9815	1.0083	2.9898
% var	0.660	0.336	0.997

Saturations de facteurs et communalités avec rotations  
Rotation varimax

Variable	Facteur1	Facteur2	Communalité
Niveau Finance	0.007	1.000	1.000
NiveautStatistiques	0.996	-0.059	0.995
NiveauNormes	0.995	0.074	0.995
Variance	1.9810	1.0088	2.9898
% var	0.660	0.336	0.997

Coefficients des scores de facteur

Variable	Facteur1	Facteur2
Niveau Finance	-0.007	0.991
NiveautStatistiques	0.503	-0.069

Nous voyons d'abord que Minitab nous redonne les coefficients sans rotations ce qui est très utile. Nous voyons dans le cas présent que la rotation VariMax (comme nous l'avions calculé dans le cours théorique) n'apporte bien évidemment aucune plus-value.

ÉCHANTILLON

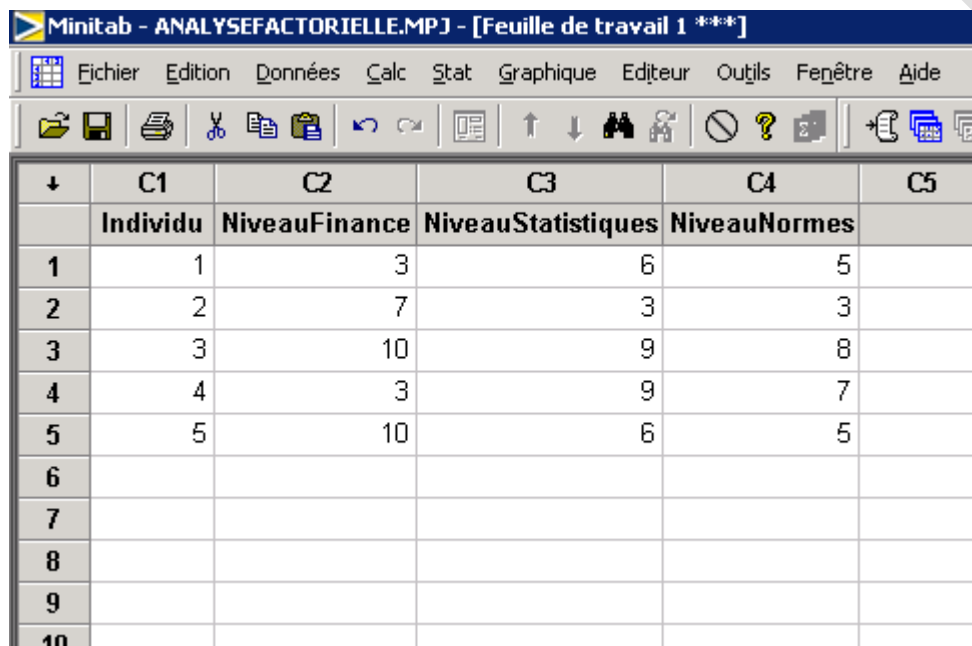


## 16.8. Exercice 171.: Analyse factorielle exploratoire (méthode ACP avec rotation OrthoMax!)

Minitab® Statistical Software 15.1.1.0

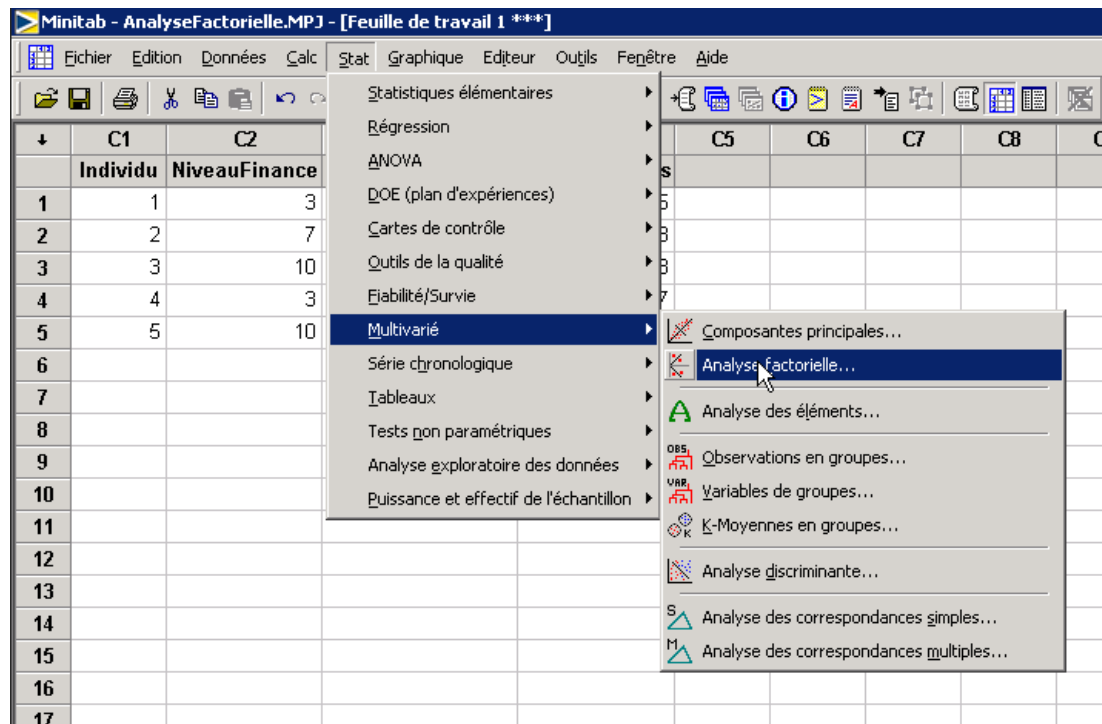
Le but va être ici d'observer en prenant le cas particulier de la généralisation OrthoMax de la rotation empirique VariMax que nous retombons bien sur les valeurs VariMax.

Donc nous repartons des données suivantes:

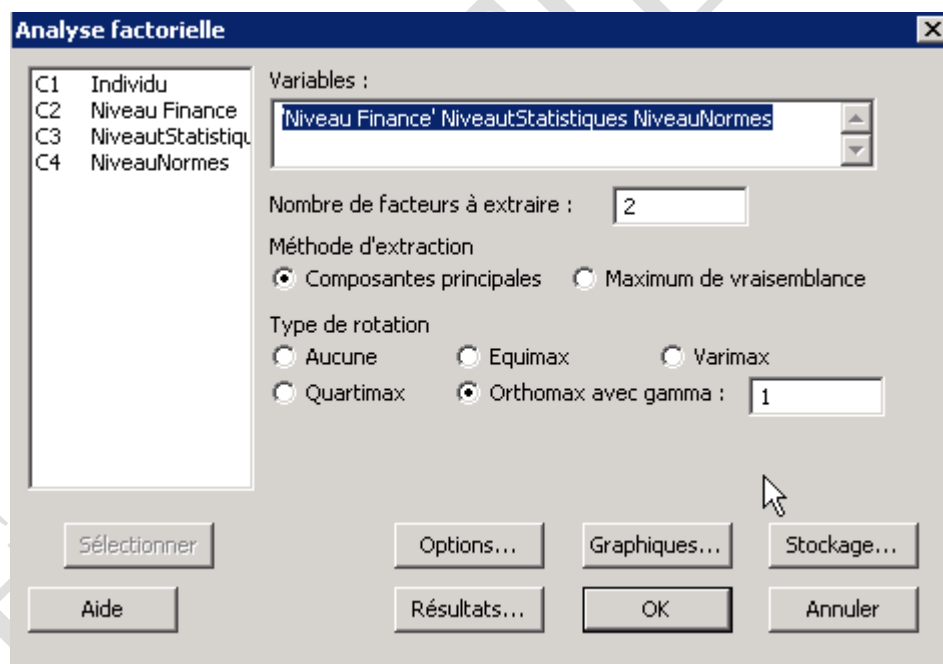


↓	C1	C2	C3	C4	C5
	Individu	NiveauFinance	NiveauStatistiques	NiveauNormes	
1	1	3	6	5	
2	2	7	3	3	
3	3	10	9	8	
4	4	3	9	7	
5	5	10	6	5	
6					
7					
8					
9					
10					

Ensuite, nous allons dans le menu **Stat/Multivarié/Analyse factorielle...**:

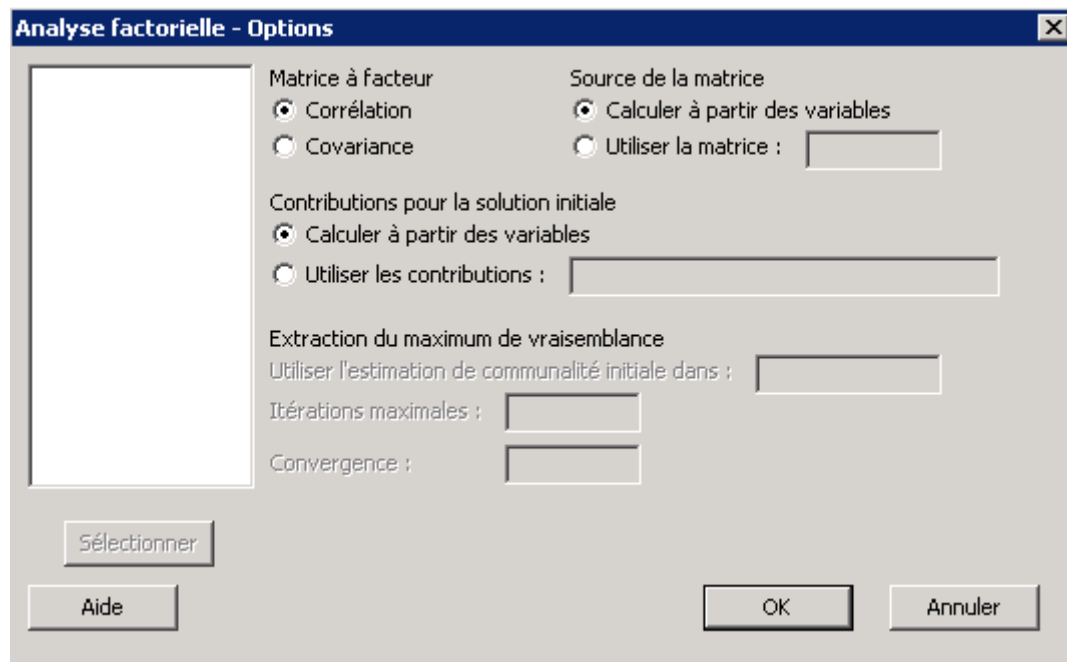


Dans la boîte de dialogue qui apparaît, nous mettons:



Comme vous pouvez le voir, nous allons utiliser la technique des **Composantes principales** et **Aucune** rotation.

Nous cliquons ensuite sur **Options...** pour prendre:



et nous obtenons comme sortie effectivement la même chose que la VariMax:

Analyse factorielle du composant principal de la matrice de corrélation

Saturations de facteurs et communalités sans rotations

Variable	Facteur1	Facteur2	Communalité
Niveau Finance	0.030	1.000	1.000
NiveautStatistiques	0.994	-0.082	0.995
NiveauNormes	0.996	0.051	0.995
Variance	1.9815	1.0083	2.9898
% var	0.660	0.336	0.997

Saturations de facteurs et communalités avec rotations  
Rotation varimax

Variable	Facteur1	Facteur2	Communalité
Niveau Finance	0.007	1.000	1.000
NiveautStatistiques	0.996	-0.059	0.995
NiveauNormes	0.995	0.074	0.995
Variance	1.9810	1.0088	2.9898
% var	0.660	0.336	0.997

Coefficients des scores de facteur

Variable	Facteur1	Facteur2
Niveau Finance	-0.007	0.991
NiveautStatistiques	0.503	-0.069

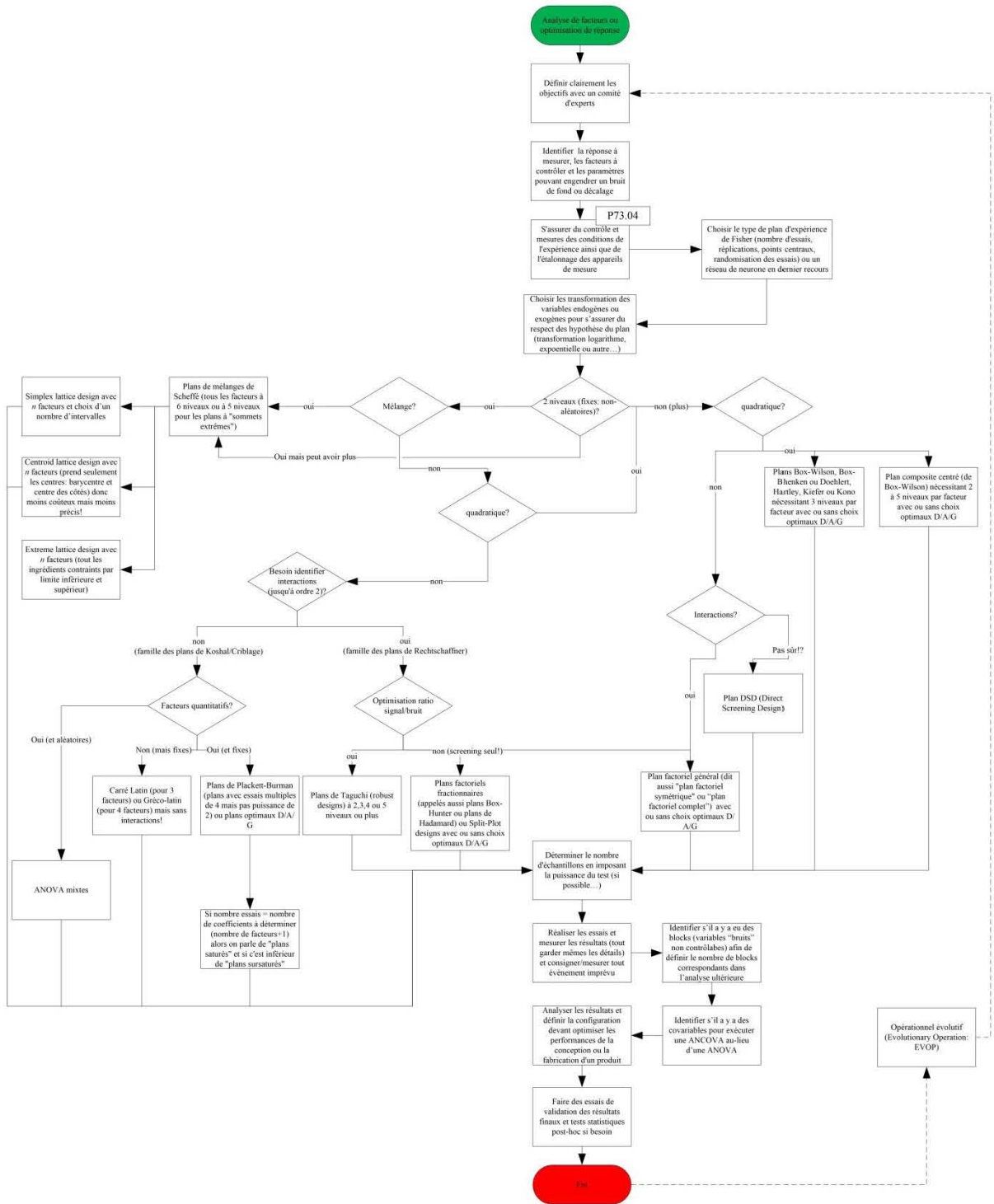
# 17. Plans d'expériences

Rappels sur les types de plans:

- **Plans de Fisher** qui inclut tous les sous-types de plans par hommage à cet illustre personnage des statistiques
- **Plans de Koshal/Plan de criblage** (aussi appelées "plans un à la fois") permettant d'analyser juste les facteurs principaux mais sans aucune interaction
- **Plans saturés** qui sont des plans de Koshal/Criblage simplement sans réplication des mesures
- **Plans factoriels généraux** ou **plans factoriels complets** à 1, 2, 3,  $n$  niveaux linéaire avec sous interactions (mais pas quadratiques ou supérieur!)
- **Plans factoriels à 2 niveaux** ou **Plans de Box & Hunter** ou encore **Plans basés sur la matrice de Hadamard** qui sont des plans factoriels (donc non quadratiques) avec ou sans interactions
- **Plans sursaturés** ou **Plans factoriels fractionnaires à 2 niveaux** (aliasés ou dit aussi "avec générateurs") avec ou sans interactions (mais pas quadratiques ou supérieur!) qui contiennent moins d'expériences ("runs" en anglais) qu'il n'y a de coefficients à déterminer
- **Plans factoriels de Plackett-Burman** dont l'ordre est un multiple de quatre mais pas une puissance de deux (donc 12, 20, 24, 36, 40, 44 etc.) et on ne peut pas analyser les interactions
- **Plans Definitive Screening Design** qui sont orthogonaux, ayant un nombre d'essais égal au double de nombre de facteurs + 1, contrairement aux plans de résolution III ont les effets principaux indépendants des interactions de facteurs doubles, contrairement aux plans de résolution IV les interactions à deux facteurs ne sont pas complètement liés avec d'autres couples de facteurs, il est possible de savoir s'il y a une courbure et d'en déterminer directement les facteurs responsables (contrairement aux autres plans où on ne sait pas les facteurs responsables), etc.
- **Plans de Taguchi** qui sont des plans factoriels complets particuliers dans le sens où les facteurs peuvent avoir de nombreux niveaux mais ils doivent en avoir tous le même nombre. Certaines tables de Taguchi peuvent gérer les interactions, d'autres de façon limitée et d'autres pas du tout!
- **Plan optimaux D, G, A, V** qui sont des plans factoriels complets généraux à  $n$  facteurs dont le nombre d'essais est minimisé algorithmiquement par la minimisation d'une expression incluant de la variance de la matrice d'information.
- **Plans de surface à deux niveaux composites centrés** qui sont des plans d'expérience quadratiques

- **Plans de surface de Box-Behnken** qui sont des plans d'expérience quadratiques nécessitant 3 niveaux pour tous les facteurs qui ont l'avantage d'être rotatifs
- **Plans de surface de Box-Wilson** qui sont des plans d'expérience quadratiques nécessitant 5 niveaux pour tous les facteurs qui ont l'avantage d'être rotatifs
- **Plans de Rechtschaffner** sont les plans où toutes les interactions d'ordre trois et supérieurs sont négligées
- **Plans split-plot** ou **Plans en parcelles** divisées à 2 niveaux (facteurs difficiles à changer) très utilisés en ingénierie agroalimentaire ou dans les cultures biologiques dont un des facteurs (parmi une quantité indéterminée) est difficile voire impossible à randomiser
- **Plans de Scheffé** basés sur l'hypothèse que l'on peut construire le modèle de mélange sur le principe des plans de surface avec un astuce de changement de variable via la contrainte de mélange et qui contient donc implicitement: les **simplex lattice design**, les **centroid lattice design**, les **extreme lattice design**.
- **Plans Carrés Latins** qui sont des ANOVA à trois facteurs imbriqués mais en version réduite car TOUS les niveaux du troisième facteur ne sont pas appliqués pour toutes les combinaisons de chacun des deux premiers facteurs mais seulement une fois avec la configuration particulière que chaque niveau du troisième facteur doit se trouver qu'une seule fois dans chaque ligne et qu'une seule fois dans chacun des colonnes des deux premiers (donc cela élimine l'analyse des interactions)!
- **Plans Greco-Latin** qui sont au fait la superposition des deux plans carrés latins. Donc impossible aussi d'analyser les interactions avec ce type de plans!
- **Plans par réseaux de neurone** qui ne nécessitent aucune hypothèse de départ mais qui ne permettent pas l'inférence statistique sans bootstrap (donc pas d'ANOVA, ni de screening de facteurs!).

Et pour ceux qui auraient de la peine à décider et à s'y retrouver voici un processus de décision (mis à jour une fois tous les deux ans suite à des retours clients et à de nouveaux développements en mathématiques):



## 17.1. Exercice 172.: Construction de plans d'expériences

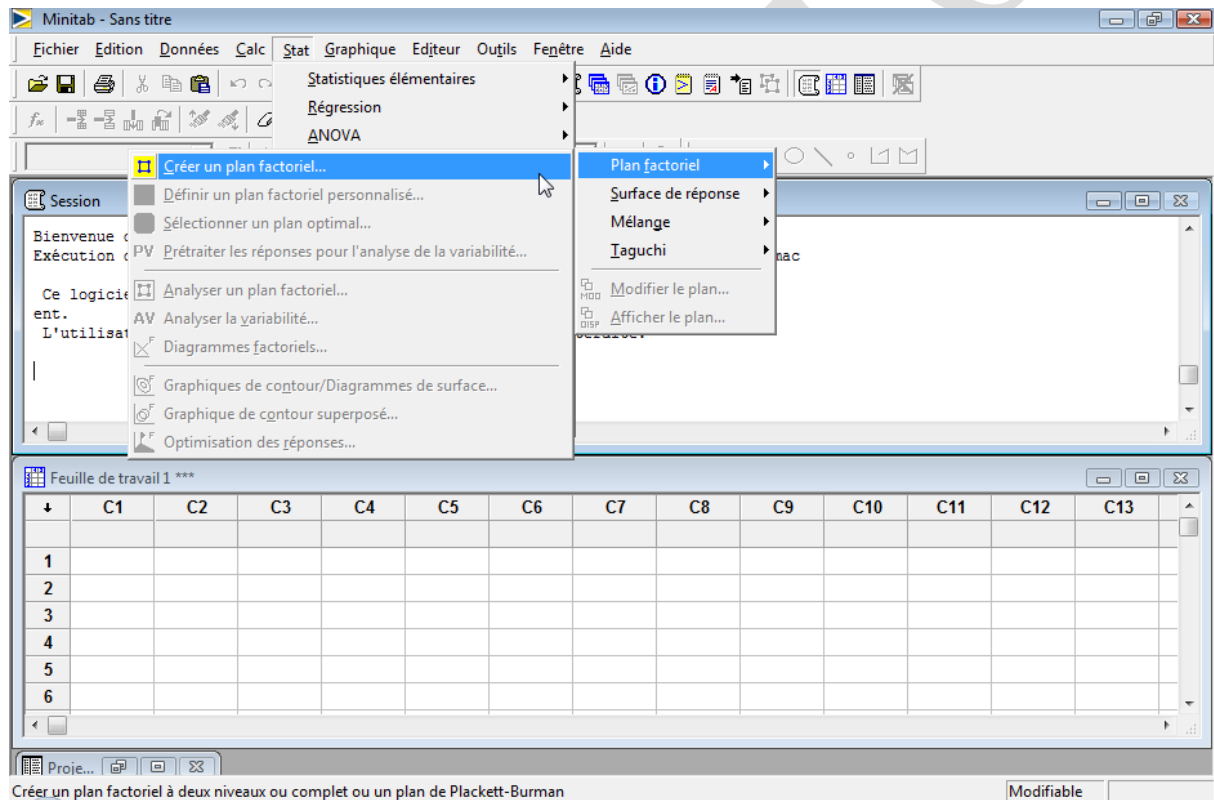
Minitab® Statistical Software 15.1.1.0

Ici, le lecteur n'aura pas besoin d'ouvrir de fichiers car le seul objectif ici est de vérifier que Minitab® Statistical Software propose les plans d'expérience de base que nous avons étudié dans le cours théorique sur le même sujet avant de passer à des exemples concrets.

**Remarque:** Attention!!! Après avoir créé un plan d'expérience, à ce jour, dans Minitab, de nombreux paramètres (et non pas tous!) du plan d'expérience ne peuvent plus être modifiés. Il faut alors recommencer à zéro (ce qui est assez pénible!).

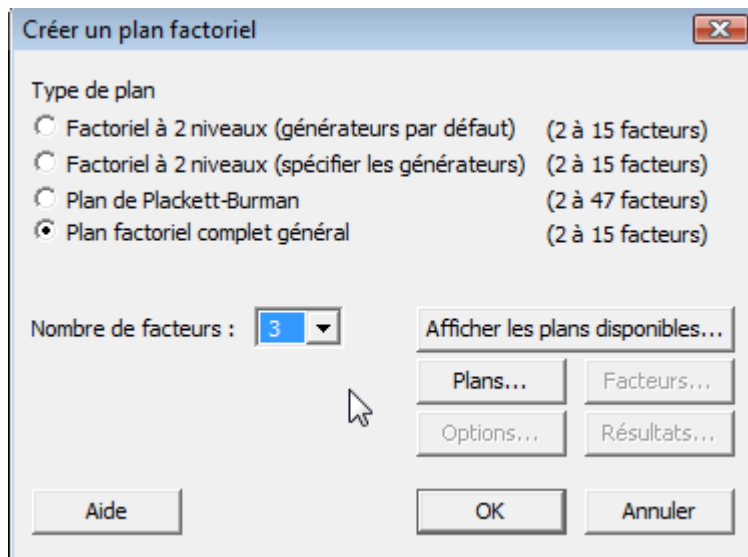
### 17.1.1. Construction d'un plan factoriel complet

Pour commencer, nous allons dans le menu **Stat/DOE (plan d'expériences)/Plan factoriel/Créer un plan factoriel...**:



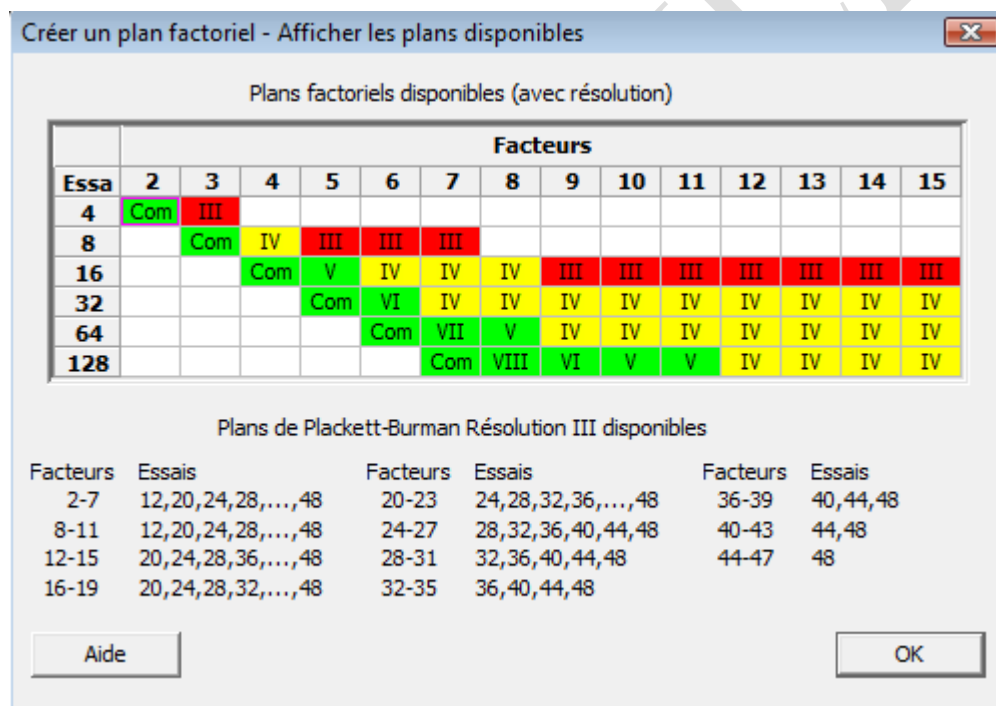
Bien que le nom soit "plan factoriel", nous pouvons créer en réalité d'autres plans que des plans factoriels comme nous allons le voir. Donc le nom n'est pas forcément très bien choisi.

Vient alors la boîte de dialogue suivante:



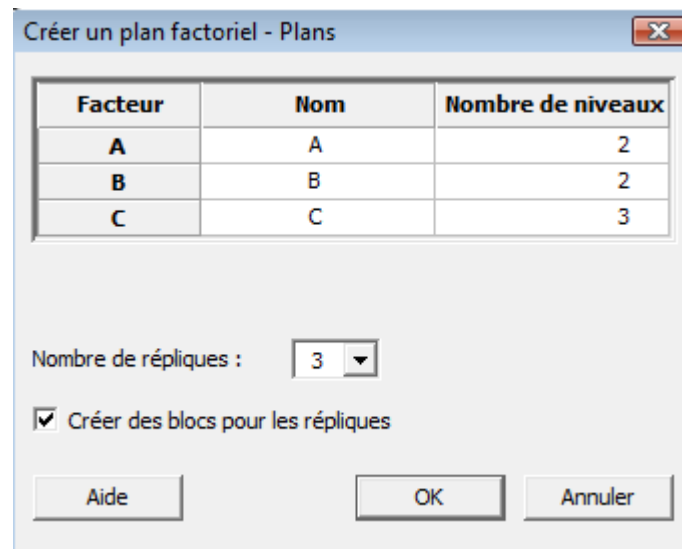
Nous prenons d'abord le classique **Plan factoriel complet général** correspondant à l'exercice d'optimisation fait lors du cours de statistique théorique.

Notez qu'en cliquant sur **Afficher les plans disponibles...** nous avons:



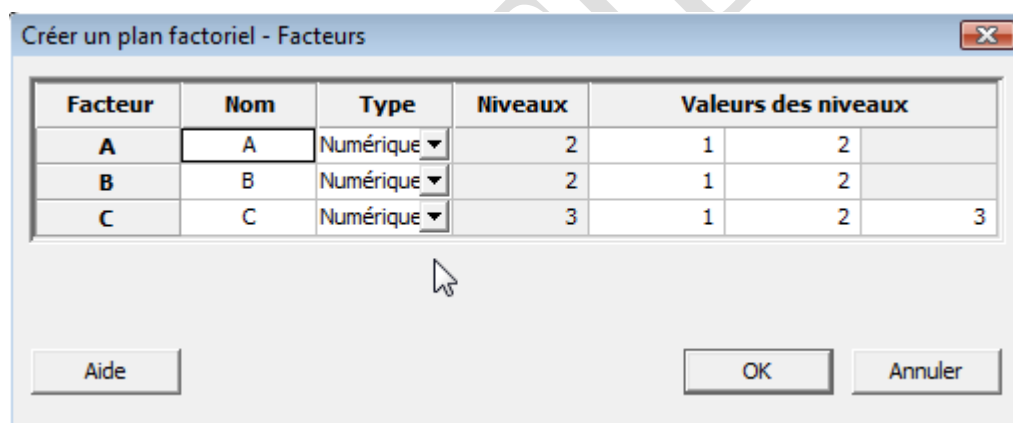
Ensuite, nous cliquons sur le bouton **Plans** pour y mettre conformément toujours au même exercice:



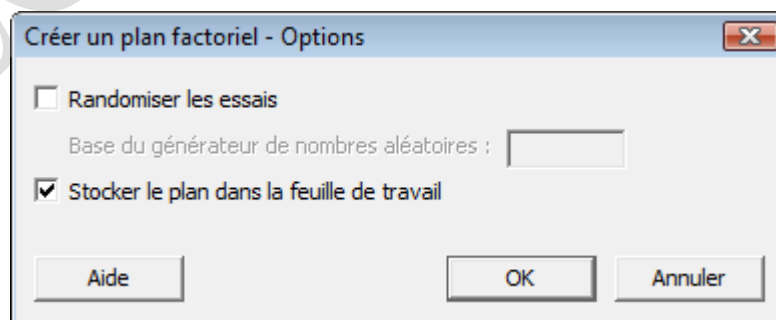


où nous avons comme pour l'exercice fait lors du cours de statistique théorique, signalé que nous avons **3 Nombre de répliques** et nous demandons à Minitab® Statistical Software de **Créer des blocs pour les répliques**.

Ensuite, nous cliquons sur le bouton **Facteurs** pour contrôler que tout est conforme relativement à ce que nous avons fait dans le cours de statistique théorique:



Nous cliquons ensuite sur le bouton **Options** pour décocher **Randomiser les essais** et pour activer **Stocker le plan** dans la feuille de travail:



Nous validons le tout par **OK**. Nous obtenons alors dans la fenêtre de session:

### Plan factoriel à plusieurs niveaux

Facteurs : 3 Répliques : 3  
 Essais de base : 12 Nombre total d'essais : 36  
 Blocs de base : 1 Nombre total de blocs : 3

Nombre de niveaux : 2; 2; 3

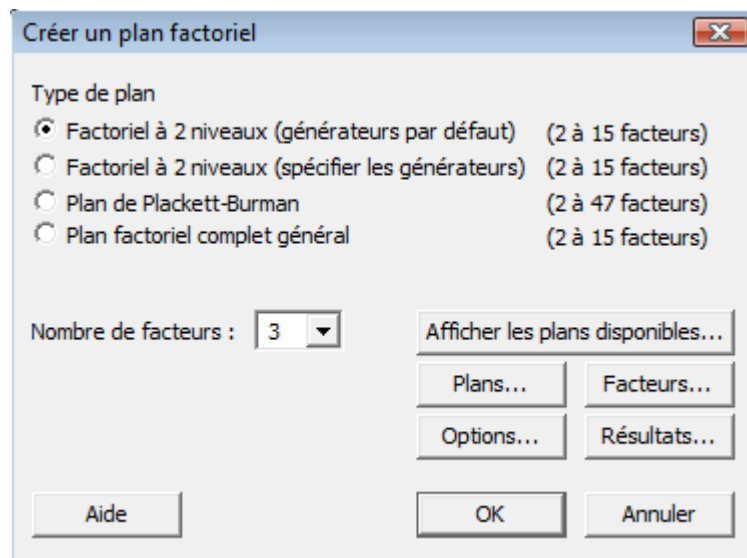
où tout est conforme à ce que nous avons vu dans le cours de statistique théorique, ainsi que le plan d'expérience ci-dessous qui est lui aussi totalement conforme:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	OrdreStd	OrdEssai	TypePt	Blocs	A	B	C						
1	1	1	1	1	1	1	1						
2	2	2	1	1	1	1	2						
3	3	3	1	1	1	1	3						
4	4	4	1	1	1	2	1						
5	5	5	1	1	1	2	2						
6	6	6	1	1	1	2	3						
7	7	7	1	1	2	1	1						
8	8	8	1	1	2	1	2						
9	9	9	1	1	2	1	3						
10	10	10	1	1	2	2	1						
11	11	11	1	1	2	2	2						
12	12	12	1	1	2	2	3						
13	13	13	1	2	1	1	1						
14	14	14	1	2	1	1	2						
15	15	15	1	2	1	1	3						
16	16	16	1	2	1	2	1						
17	17	17	1	2	1	2	2						
18	18	18	1	2	1	2	3						
19	19	19	1	2	2	1	1						

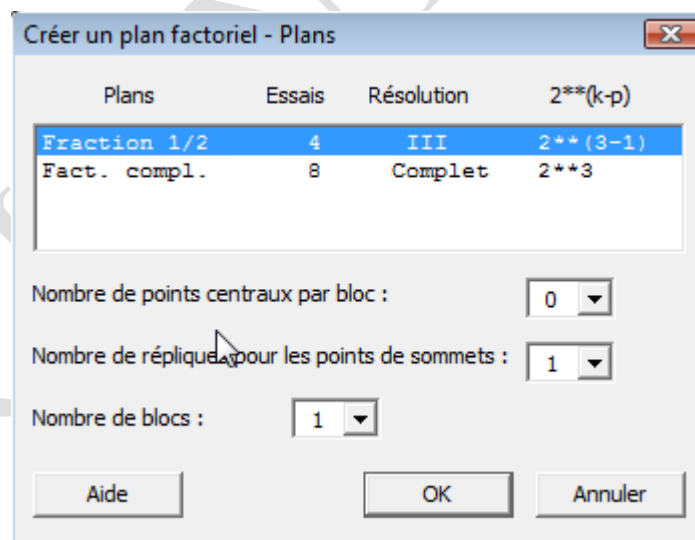
### 17.1.2. Construction d'un plan factoriel fractionnaire à 2 niveaux

Maintenant voyons comment générer un plan fractionnaire (dans le cas présent ce sera un plan factoriel fractionnaire comme celui vu dans le cours théorique) où nous laissons le soin à Minitab® Statistical Software de choisir les meilleurs générateurs (alias).

Nous commençons donc par:



où comme dans le cours de statistique théorique, nous prenons **3 Nombre de facteurs**. Ensuite, nous cliquons sur le bouton **Plans...**:



Et nous prenons le plan **Fraction 1/2** et nous cliquons sur **OK** pour tout valider:

**Plan factoriel fractionnaire**

Facteurs : 3 Plan de base : 3; 4 Résolution : III  
 Essais : 4 Répliques : 1 Fraction : 1/2  
 Blocs : 1 Points centraux (total) : 0

\* REMARQUE \* Certains effets principaux sont confondus avec des interactions à deux facteurs.

Alias initiales (indépendantes) : C = AB

Structure des alias

- I + ABC
- A + BC
- B + AC
- C + AB

Nous voyons déjà dans la fenêtre de sessions que nous avons le même générateur et les mêmes alias que ceux vus dans le cours théorique. Le plan correspondant est alors donné par Minitab® Statistical Software:

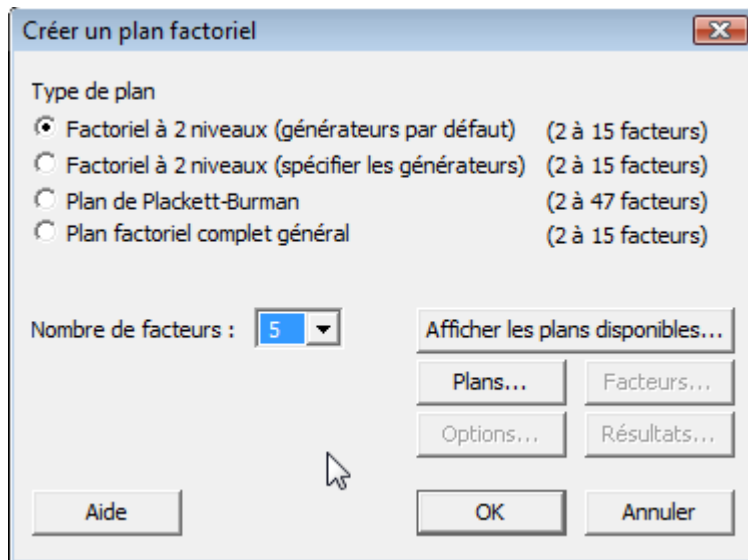
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	C						
1	2	1	1	1	1	-1	-1						
2	1	2	1	1	-1	-1	1						
3	3	3	1	1	-1	1	-1						
4	4	4	1	1	1	1	1						
5													
6													
7													
8													
9													
10													
11													
12													
13													
14													
15													
16													
17													
18													
19													

et nous retrouvons donc bien la même matrice (à la différence de l'ordre des lignes car je n'ai pas décoché l'option de randomisation des essais).

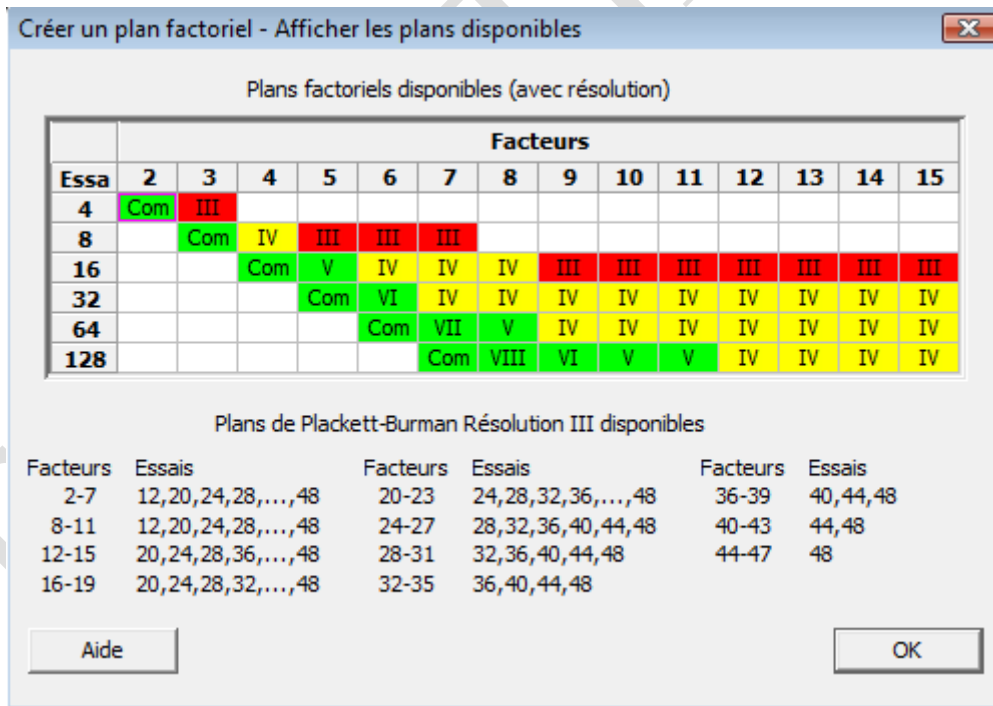
Maintenant voyons le vrai intérêt d'un logiciel pour la base des plans d'expérience. Comme nous l'avons mentionné dans le cours de statistique théorique, au-delà de 4 facteurs il devient

pénible de déterminer la matrice d'Hadamard pour les plans fractionnaires et les alias optimaux. Alors voyons comment Minitab® Statistical Software nous facilite la tâche.

Nous retournons dans la boîte de dialogue suivante:

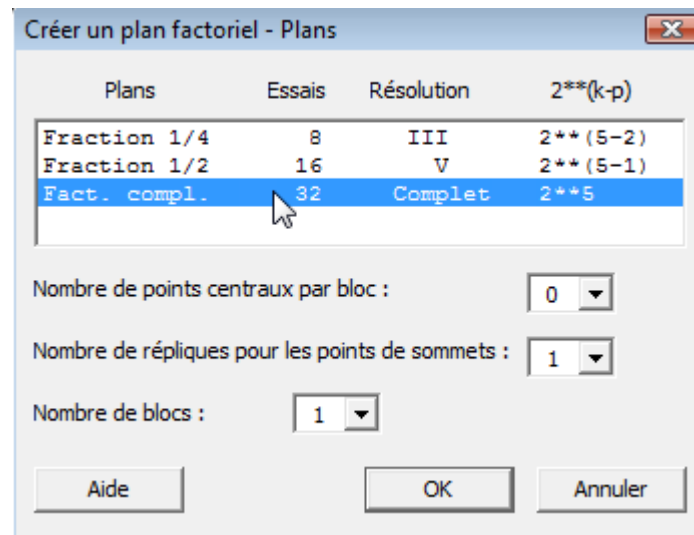


et nous choisissons **5 Nombre de facteurs**. Nous cliquons sur **Afficher les plans disponibles...** et nous avons alors:

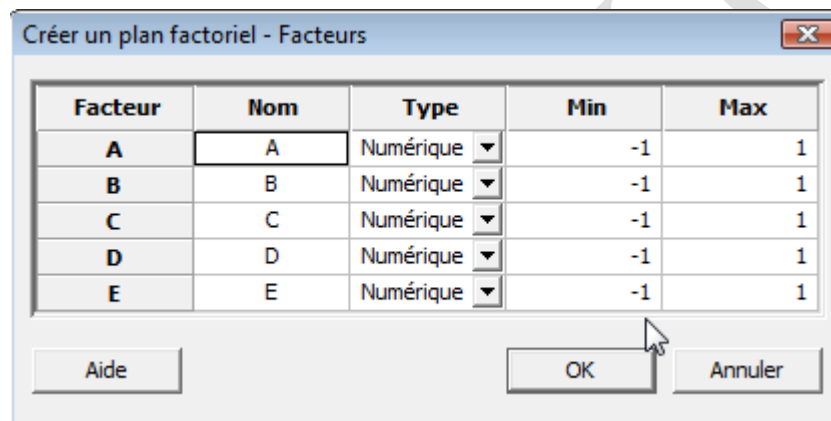


Nous voyons ainsi que pour 5 facteurs, Minitab® Statistical Software nous propose un plan complet à 32 essais, un plan factoriel fractionnaire (1/2) de résolution V de 16 essais et un plan factoriel fractionnaire (1/4) de résolution III dit alors "**plan de screening**" (dangereux à utiliser quand même!!!).

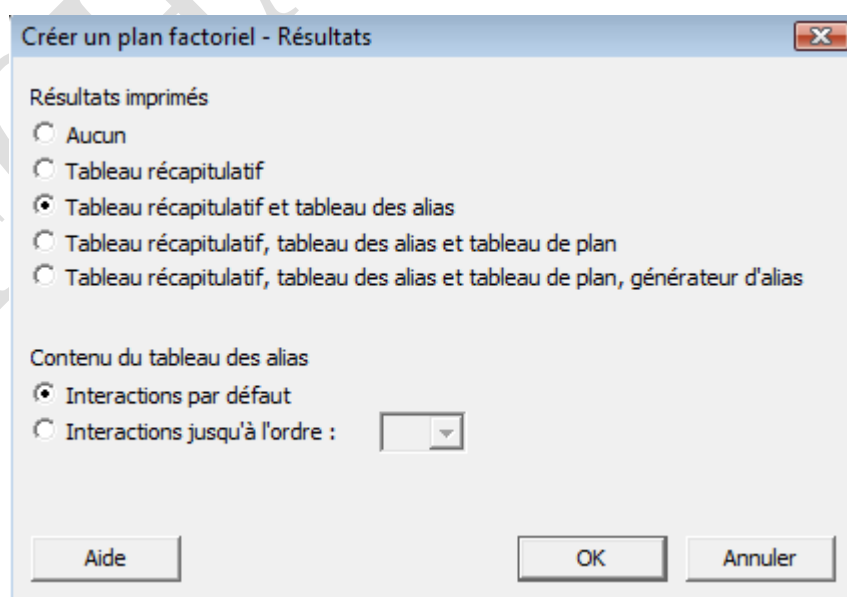
Voyons donc cela en prenant d'abord le classique plan complet:



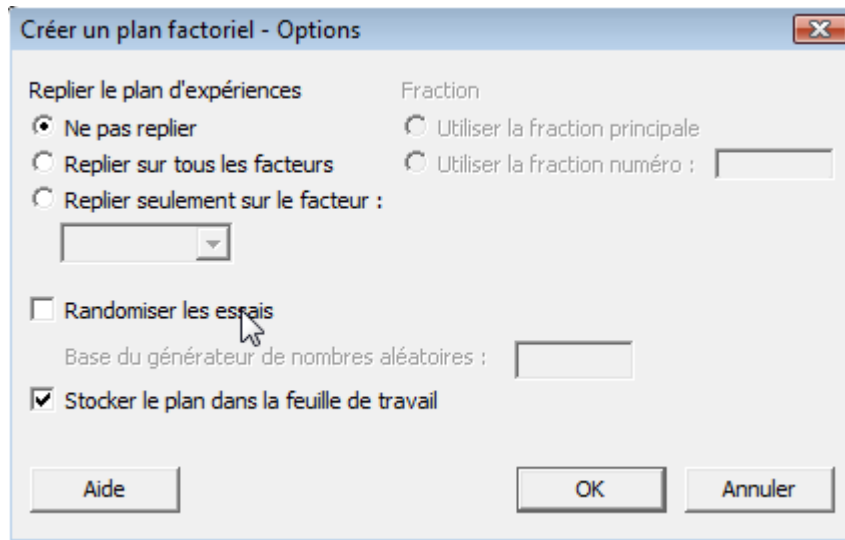
et en prenant des facteurs factoriels:



Demandons à Minitab® Statistical Software de nous afficher le **Tableau récapitulatif et tableau des alias** dans la fenêtre de session:



Ensuite, demandons-lui de ne pas **Randomiser les essais**:



Validons le tout par **OK**. Nous avons alors:

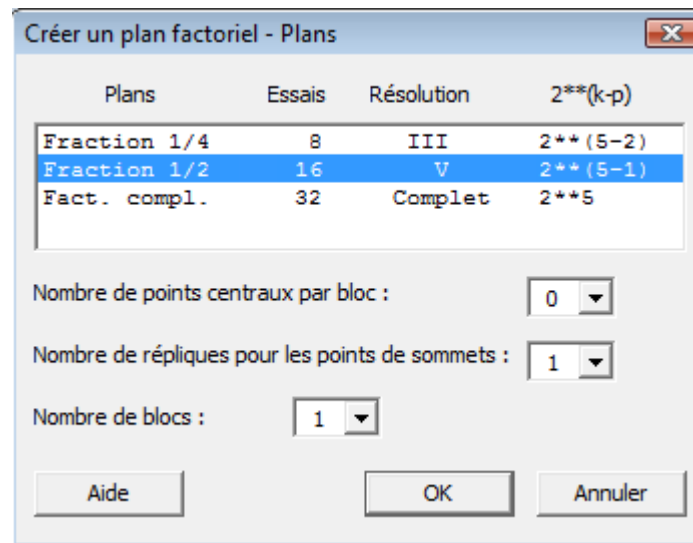
### Plan factoriel complet

```
Facteurs : 5 Plan de base : 5; 32
Essais : 32 Répliques : 1
Blocs : 1 Points centraux (total) : 0
```

Tous les termes sont indépendants des alias.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	C	D	E				
1	1	1	1	1	-1	-1	-1	-1	-1				
2	2	2	1	1	1	-1	-1	-1	-1				
3	3	3	1	1	-1	1	-1	-1	-1				
4	4	4	1	1	1	1	-1	-1	-1				
5	5	5	1	1	-1	-1	1	-1	-1				
6	6	6	1	1	1	-1	1	-1	-1				
7	7	7	1	1	-1	1	1	-1	-1				
8	8	8	1	1	1	1	1	-1	-1				
9	9	9	1	1	-1	-1	-1	1	-1				
10	10	10	1	1	1	-1	-1	1	-1				
11	11	11	1	1	-1	1	-1	1	-1				
12	12	12	1	1	1	1	-1	1	-1				
13	13	13	1	1	-1	-1	1	1	-1				
14	14	14	1	1	1	-1	1	1	-1				
15	15	15	1	1	-1	1	1	1	-1				
16	16	16	1	1	1	1	1	1	-1				
17	17	17	1	1	-1	-1	-1	-1	1				
18	18	18	1	1	1	-1	-1	-1	1				
19	19	19	1	1	-1	1	-1	-1	1				

Donc tout est bon (à part que comme d'habitude avec Minitab® Statistical Software, il est dommage qu'il ne propose pas d'afficher les interactions). Ensuite, voyons le plan fractionnaire 1/2 de résolution V. Il va être intéressant de voir ici quels sont les alias choisis par défaut par le logiciel. Nous avons donc:



Nous validons par **OK** et nous obtenons:

```
Facteurs : 5 Plan de base : 5; 16 Résolution : V
Essais : 16 Répliques : 1 Fraction : 1/2
Blocs : 1 Points centraux (total) : 0
```

Alias initiales (indépendantes) : E = ABCD

|  
Structure des alias

I + ABCDE

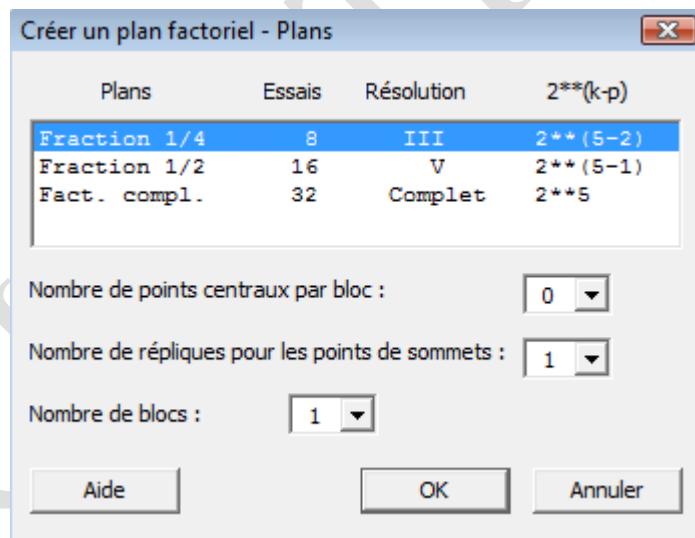
A + BCDE  
 B + ACDE  
 C + ABDE  
 D + ABCE  
 E + ABCD  
 AB + CDE  
 AC + BDE  
 AD + BCE  
 AE + BCD  
 BC + ADE  
 BD + ACE  
 BE + ACD  
 CD + ABE  
 CE + ABD  
 DE + ABC

Donc nous voyons la liste de tous les alias, ce qui est très intéressant! Par ailleurs, nous pouvons voir que le choix de ceux-ci est même intuitif sans faire de grande théorie. Ce qui donne pour plan:



	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	C	D	E				
1	1	1	1	1	-1	-1	-1	-1	1				
2	2	2	1	1	1	-1	-1	-1	-1				
3	3	3	1	1	-1	1	-1	-1	-1				
4	4	4	1	1	1	1	-1	-1	1				
5	5	5	1	1	-1	-1	1	-1	-1				
6	6	6	1	1	1	1	-1	1	-1				
7	7	7	1	1	-1	1	1	-1	1				
8	8	8	1	1	1	1	1	-1	-1				
9	9	9	1	1	-1	-1	-1	1	-1				
10	10	10	1	1	1	-1	-1	1	1				
11	11	11	1	1	-1	1	-1	1	1				
12	12	12	1	1	1	1	-1	1	-1				
13	13	13	1	1	-1	-1	1	1	1				
14	14	14	1	1	1	-1	1	1	-1				
15	15	15	1	1	-1	1	1	1	-1				
16	16	16	1	1	1	1	1	1	1				
17													
18													
19													

Maintenant voyons le plan factoriel fractionnaire 1/4 de résolution III. Il est très intéressant de voir comment les alias vont être choisis:



En validant par **OK**, nous obtenons:

### Plan factoriel fractionnaire

Facteurs : 5 Plan de base : 5; 8 Résolution : III  
 Essais : 8 Répliques : 1 Fraction : 1/4  
 Blocs : 1 Points centraux (total) : 0

\* REMARQUE \* Certains effets principaux sont confondus avec des interactions à deux facteurs.

Alias initiales (indépendantes) : D = AB; E = AC

Structure des alias

I + ABD + ACE + BCDE

A + BD + CE + ABCDE

B + AD + CDE + ABCE

C + AE + BDE + ABCD

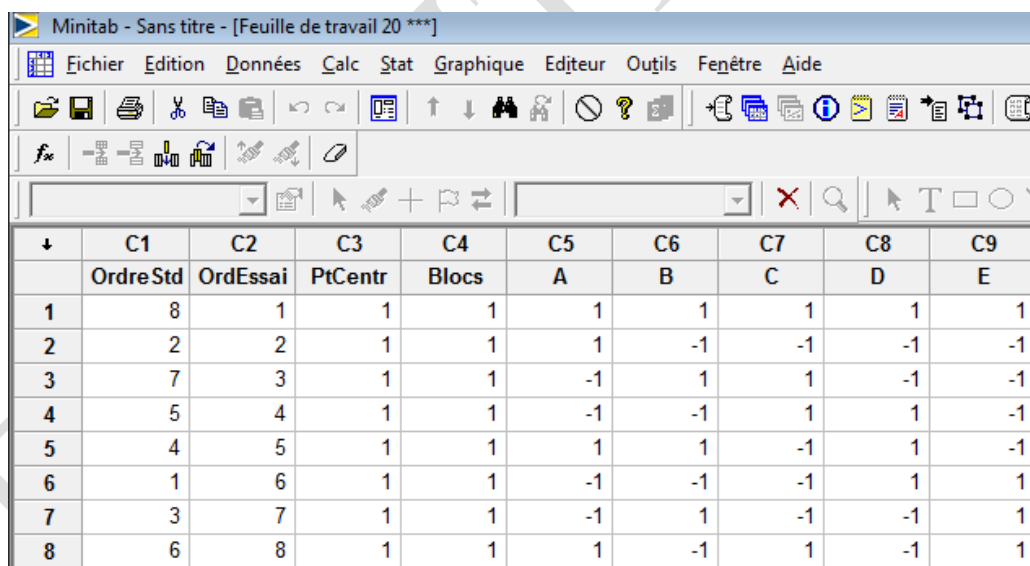
D + AB + BCE + ACDE

E + AC + BCD + ABDE

BC + DE + ABE + ACD

BE + CD + ABC + ADE

Il est donc très intéressant de pouvoir observer le détail du choix des alias. Le plan quant à lui donne:



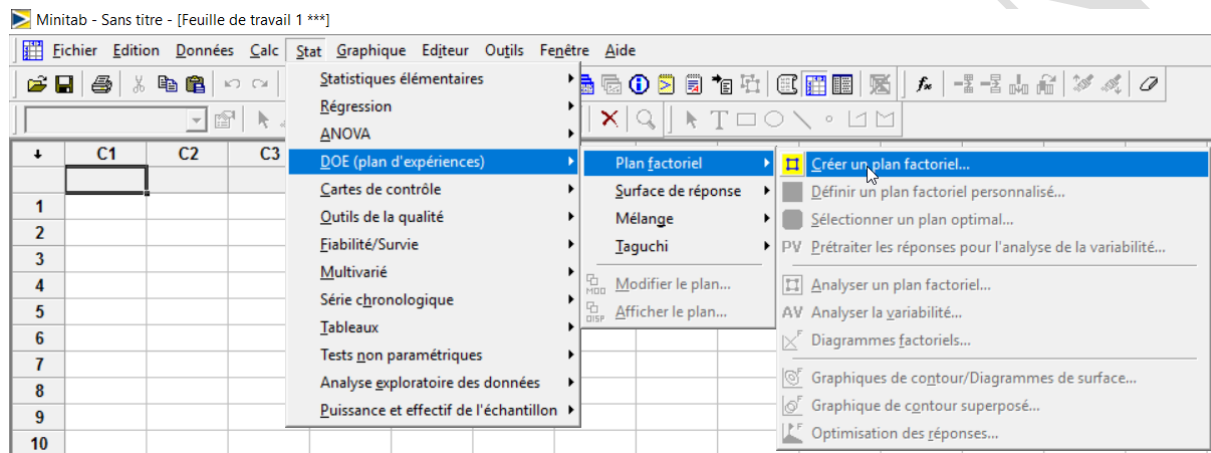
↓	C1	C2	C3	C4	C5	C6	C7	C8	C9
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	C	D	E
1	8	1	1	1	1	1	1	1	1
2	2	2	1	1	1	-1	-1	-1	-1
3	7	3	1	1	-1	1	1	-1	-1
4	5	4	1	1	-1	-1	1	1	-1
5	4	5	1	1	1	1	-1	1	-1
6	1	6	1	1	-1	-1	-1	1	1
7	3	7	1	1	-1	1	-1	-1	1
8	6	8	1	1	1	-1	1	-1	1

### Choix de la fraction du plan fractionnaire

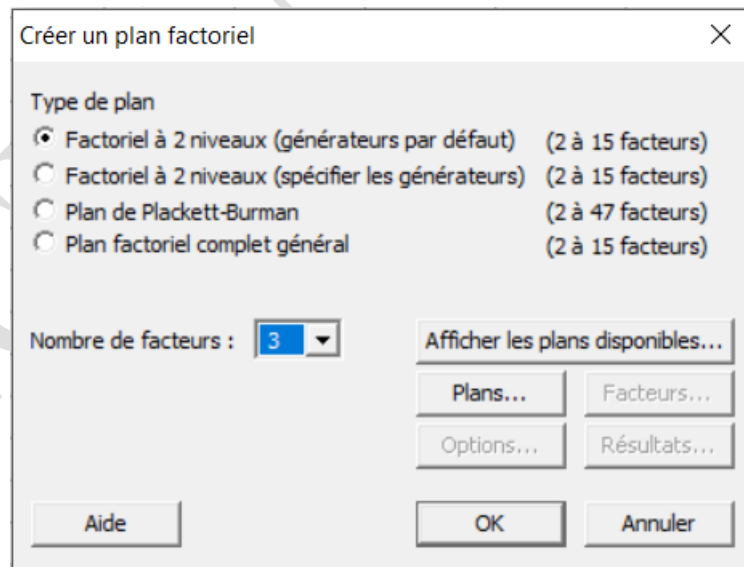
Nous avons vu dans le cours théorique qu'un plan fractionnaire était justement le plan initial fractionné en plusieurs morceaux de taille égale. Or nous avons bien vu alors qu'il fallait choisir une des fractions pour travailler (même si au fait travailler sur chacune des fractions donne par symétrie un résultat identique).

Nous souhaiterions ici vérifier que nous pouvons bien retrouver ces fractions avec Minitab et surtout dans le cas de l'exemple fait à la main dans le cours théorique: le plan  $2^3$  fractionnaire.

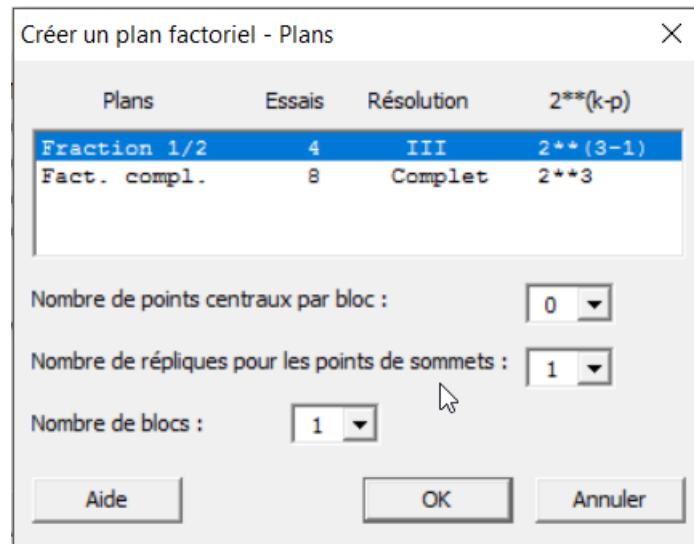
Donc d'abord pour avoir la fraction principale selon Minitab nous allons dans **DOE (plan d'expérience)/Plan factoriel/Créer un plan factoriel...**:



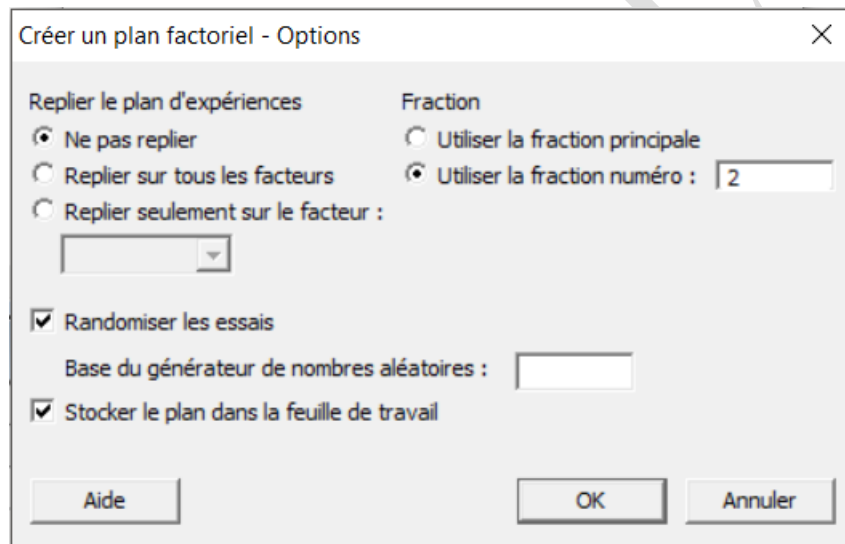
Nous choisissons Factoriel à 2 niveaux (générateurs par défaut):



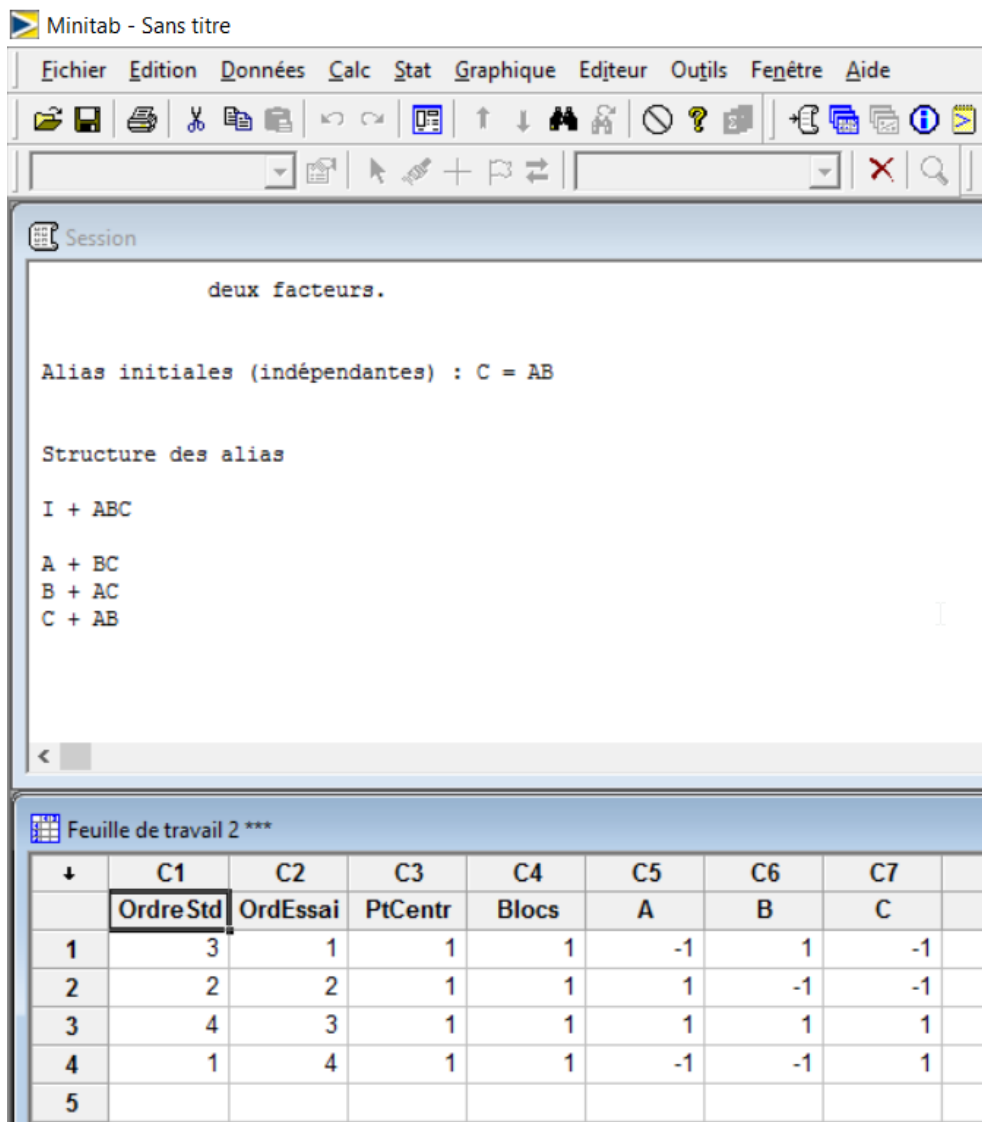
Nous cliquons sur **Plans...** :



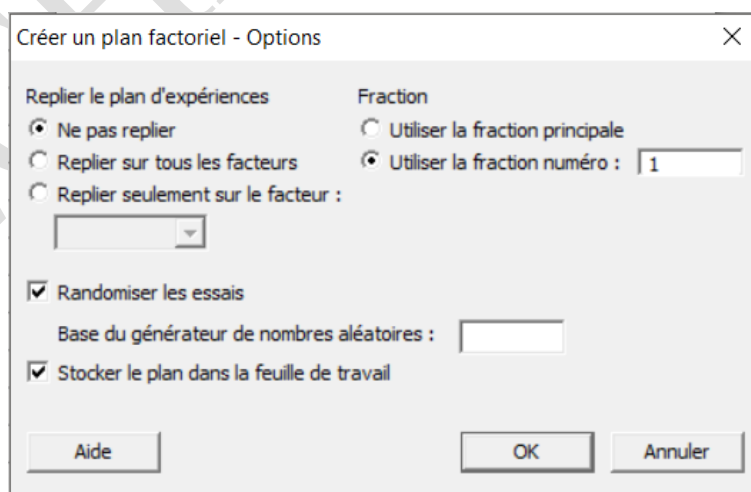
et nous prenons donc le plan fractionnaire! Ensuite nous validons par **OK** et cliquons sur **Options...** :



et nous spécifions bien **Utiliser la fraction numéro : 2**. Cela donne après avoir validé deux fois par **OK**:



et en ré-itérant avec la deuxième fraction:



Nous obtenons alors:

Minitab - Sans titre

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

Session

\* REMARQUE \* Certains effets principaux sont confondus avec des interactions à deux facteurs.

Alias initiales (indépendantes) :  $C = -AB$

Structure des alias

I - ABC

A - BC

B - AC

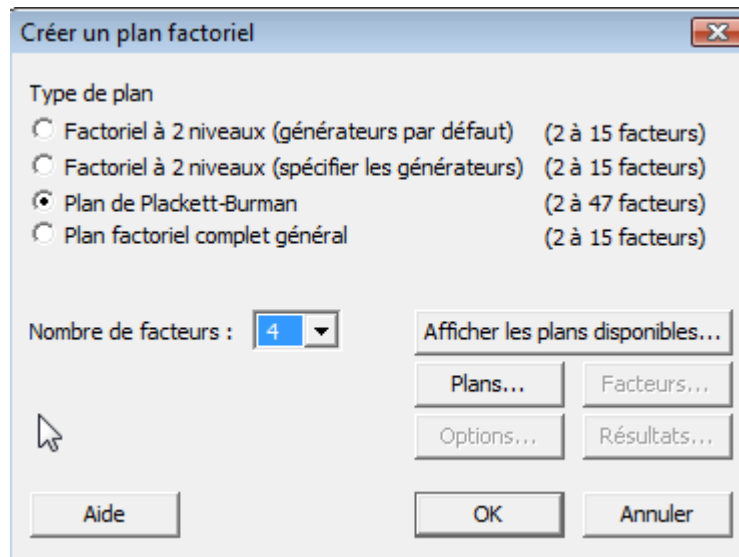
C - AB

Feuille de travail 1 \*\*\*

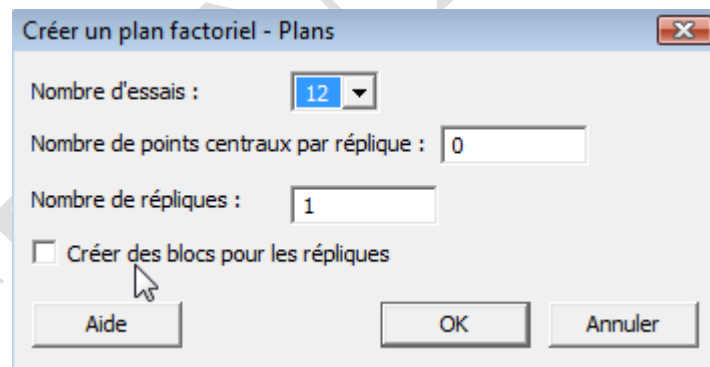
↓	C1	C2	C3	C4	C5	C6	C7	C8
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	C	
1	3	1	1	1	-1	1	1	
2	2	2	1	1	1	-1	1	
3	4	3	1	1	1	1	-1	
4	1	4	1	1	-1	-1	-1	
5								

*Plans de Plackett-Burman*

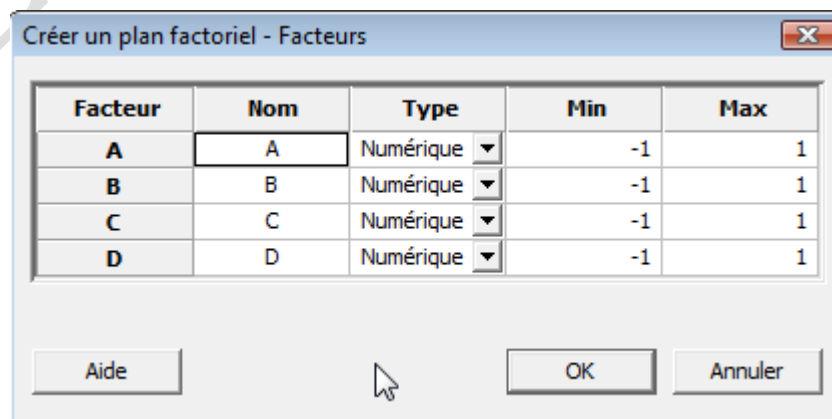
Maintenant, voyons les plans de Plackett-Burman que nous avons aussi étudié dans le cours de statistique théorique. Faisons l'exemple avec le plus petit nombre de facteurs où ces plans commencent à être vraiment intéressants comme nous l'avons vu dans le cours théorique. Nous avons alors:



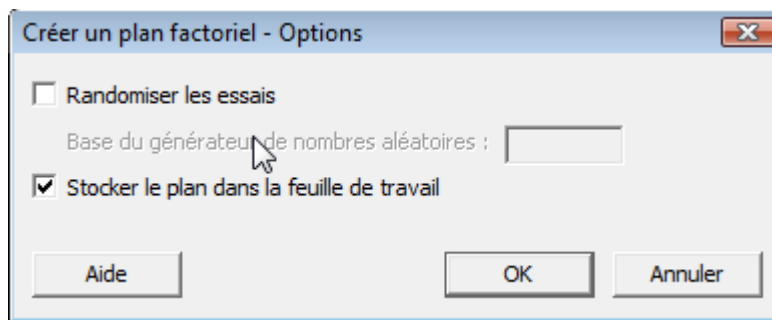
Nous prenons 12 essais comme nous l'avons fait dans le cours théorique (le lecteur remarquera que le nombre d'essais proposés correspond bien à ce que nous avons vu dans le cours théorique lors de la démonstration des matrices de Hadamard):



Nous vérifions que nous aurons bien un plan de Plackett-Burman de type factoriel:



Nous désactivons la randomisation des essais:



Et nous validons tout cela par **OK** et nous obtenons:

### Plan de Plackett-Burman

```
Facteurs :      4      Répliques :      1
Essais de base : 12     Nombre total d'essais : 12
Blocs de base :  1     Nombre total de blocs :  1
```

Nous observons donc que la fenêtre de session n'affiche pas les alias (ce qui est dommage puisque les plans de Plackett-Burman sont des plans à résolution III comme nous l'avons démontré dans le cours théorique). Le plan obtenu est lui plutôt surprenant. Nous n'avons que 4 colonnes pour les facteurs de base. Certes il n'est pas utile d'en avoir plus mais cela aurait été sympathique de pouvoir choisir...

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	OrdreStd	OrdEssai	TypePt	Blocs	A	B	C	D					
1	1	1	1	1	1	-1	1	-1					
2	2	2	1	1	1	1	-1	1					
3	3	3	1	1	-1	1	1	-1					
4	4	4	1	1	1	-1	1	1					
5	5	5	1	1	1	1	-1	1					
6	6	6	1	1	1	1	1	-1					
7	7	7	1	1	-1	1	1	1					
8	8	8	1	1	-1	-1	1	1					
9	9	9	1	1	-1	-1	-1	1					
10	10	10	1	1	1	-1	-1	-1					
11	11	11	1	1	-1	1	-1	-1					
12	12	12	1	1	-1	-1	-1	-1					
13													
14													
15													
16													
17													
18													
19													



### 17.1.3. Construction de plans factoriels optimaux D/A/G

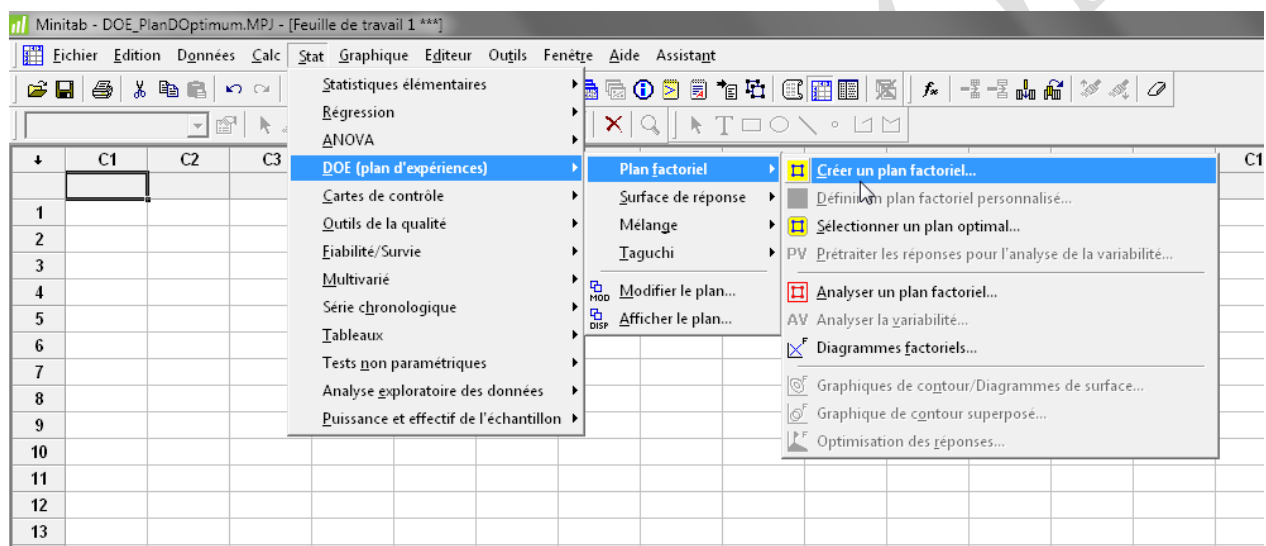
Le but va être ici, comme dans le cours théorique, de reproduire une version simplifiée de l'exemple disponible ici:

<http://www.itl.nist.gov/div898/handbook/pri/section5/pri521.htm>

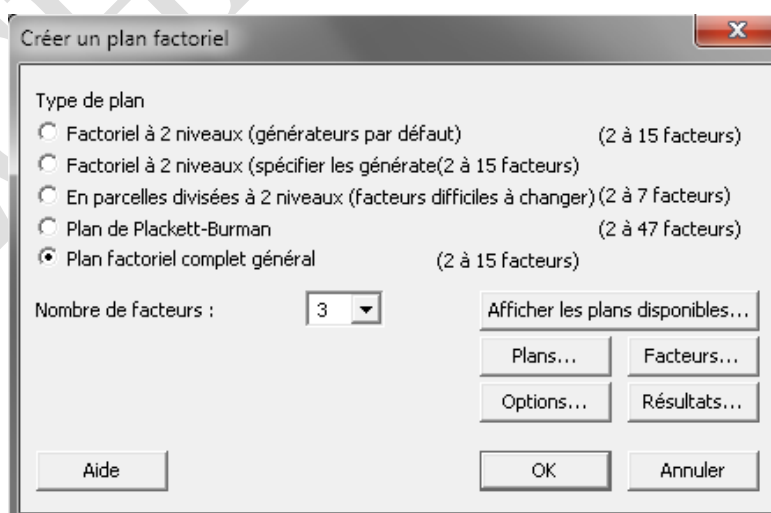
en considérant uniquement la partie linéaire (car a priori Minitab ne permet pas d'avoir des puissances des facteurs avec les version 16 et antérieures), c'est-à-dire:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

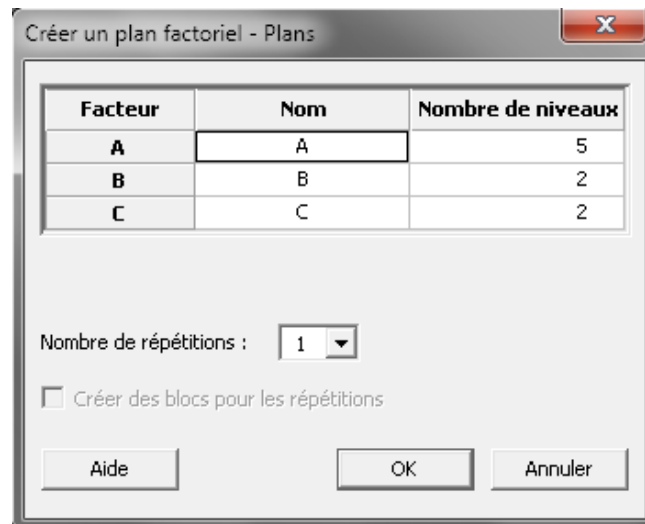
On commence en allant dans le menu **Stat/DOE (plan d'expérience)/Plan factoriel/Créer un plan factoriel...**:



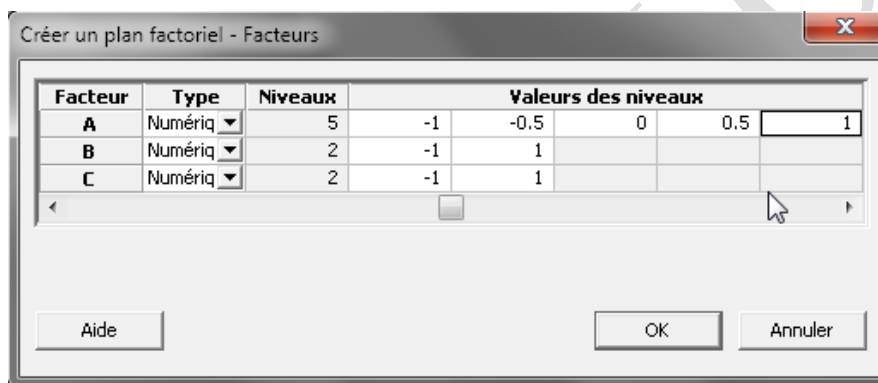
On prend comme sur le site du NIST un **Plan factoriel complet général** de 3 facteurs:



avec respectivement 5;2;2 niveaux comme sur le site du NIST:



On met les mêmes valeurs de niveaux que sur le site du NIST:



On ne randomise pas les essais afin de bien observer qu'on obtient la même matrice d'expérience que le NIST:

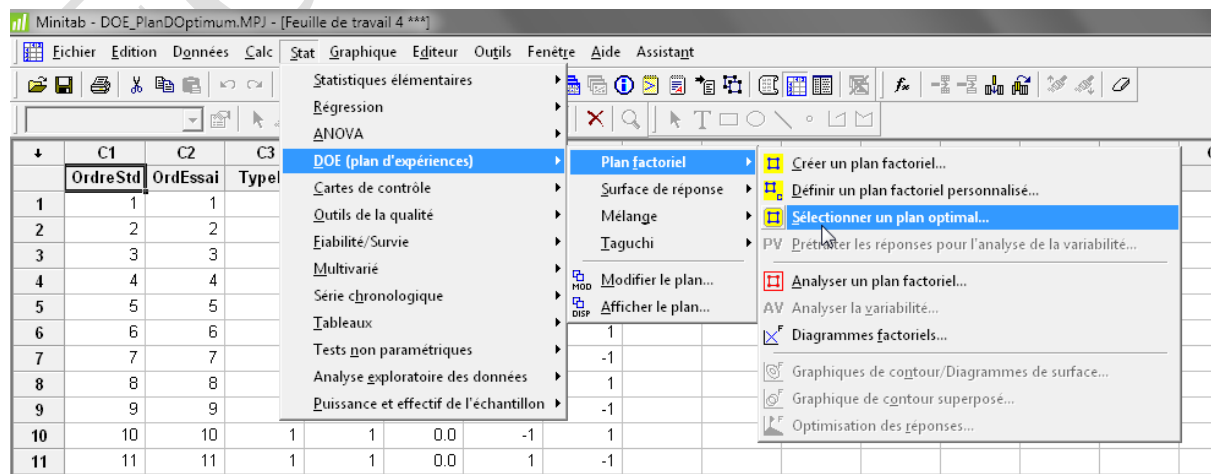


Ce qui est bien le cas! Jusque-là tout va bien:

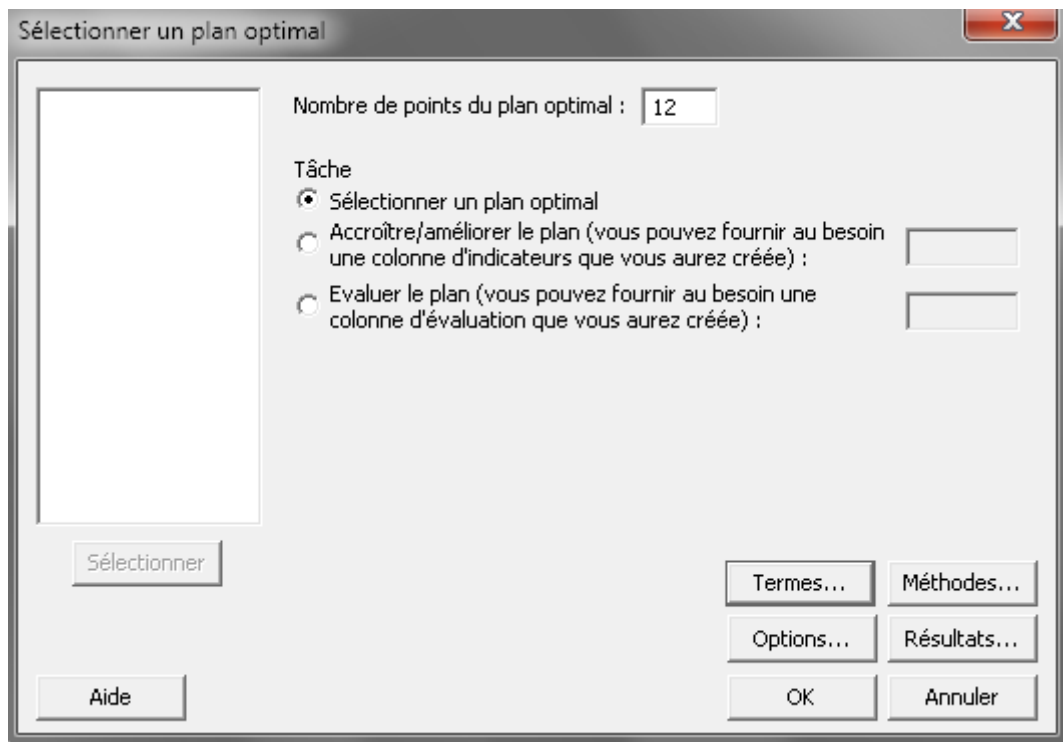
	C1	C2	C3	C4	C5	C6	C7
	OrdreStd	OrdEssai	TypePt	Blocs	A	B	C
1	1	1	1	1	-1.0	-1	-1
2	2	2	1	1	-1.0	-1	1
3	3	3	1	1	-1.0	1	-1
4	4	4	1	1	-1.0	1	1
5	5	5	1	1	-0.5	-1	-1
6	6	6	1	1	-0.5	-1	1
7	7	7	1	1	-0.5	1	-1
8	8	8	1	1	-0.5	1	1
9	9	9	1	1	0.0	-1	-1
10	10	10	1	1	0.0	-1	1
11	11	11	1	1	0.0	1	-1
12	12	12	1	1	0.0	1	1
13	13	13	1	1	0.5	-1	-1
14	14	14	1	1	0.5	-1	1
15	15	15	1	1	0.5	1	-1
16	16	16	1	1	0.5	1	1
17	17	17	1	1	1.0	-1	-1
18	18	18	1	1	1.0	-1	1
19	19	19	1	1	1.0	1	-1
20	20	20	1	1	1.0	1	1

soit exactement le même tableau avec le même ordre que dans les cours théorique.

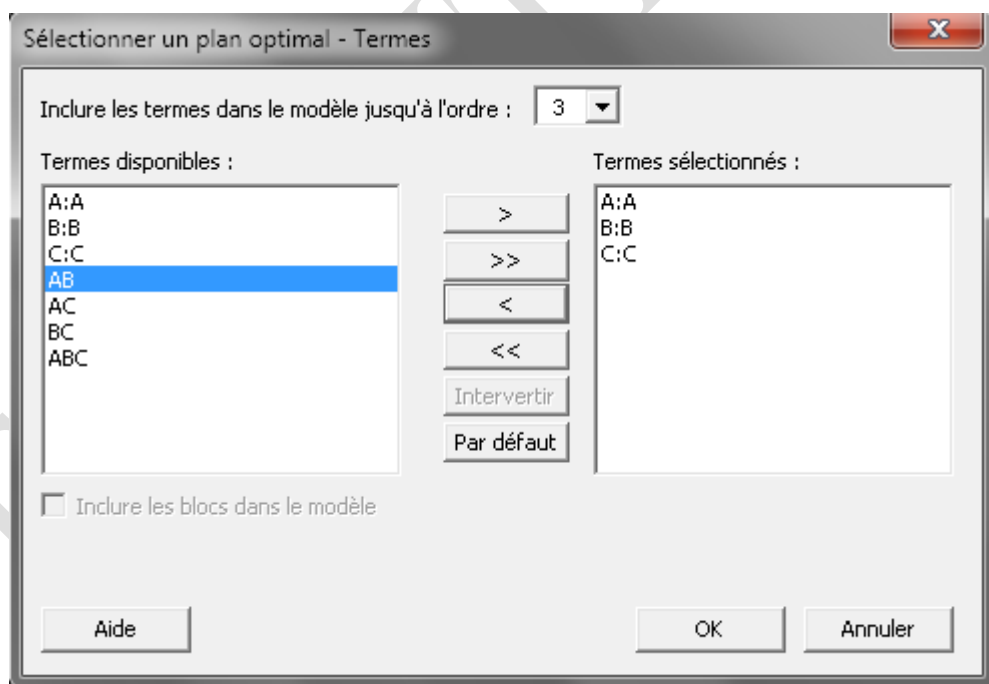
Ensuite on veut en sortir un plan optimal comme sur le NIST alors nous allons dans le menu DOE (plan d'expérience)/Plan factoriel/Sélectionner un plan optimal...:



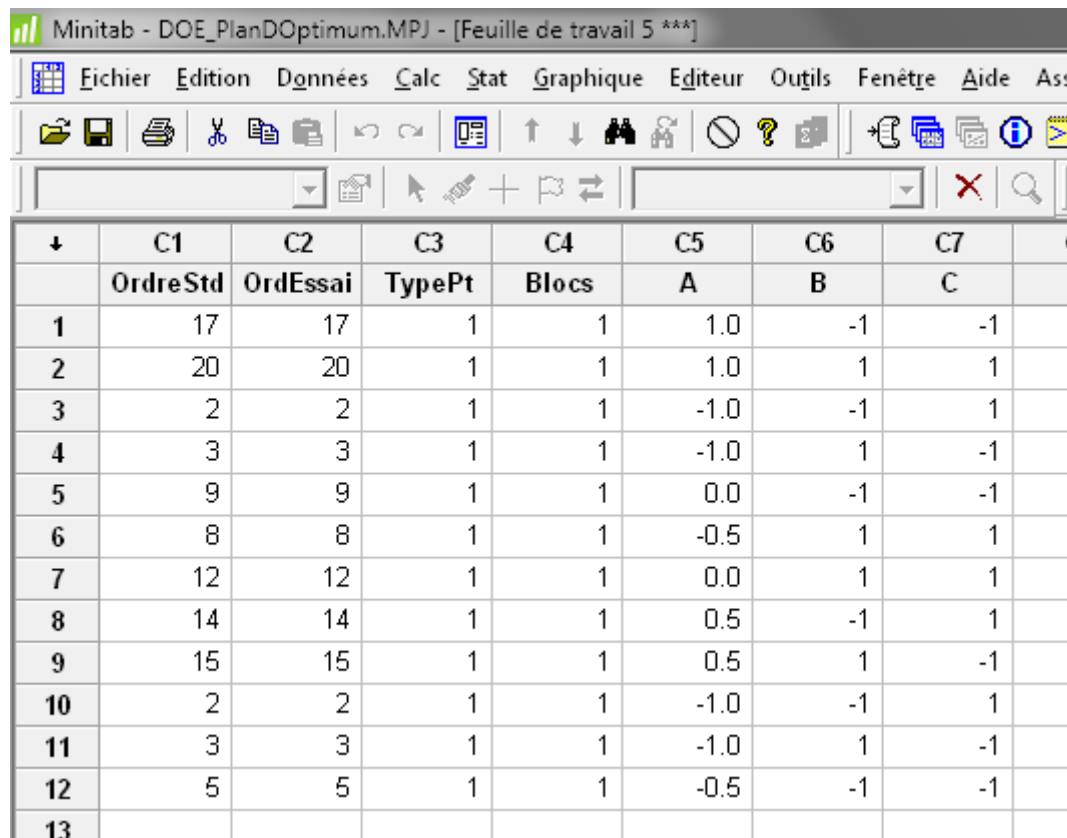
On prend:



avec les termes:



On obtient la matrice d'expérience (ce qui est acceptable mais a priori pas optimal mais ça ce n'est pas grave car cela dépend de l'algorithme utilisé):



↓	C1	C2	C3	C4	C5	C6	C7
	OrdreStd	OrdEssai	TypePt	Blocs	A	B	C
1	17	17	1	1	1.0	-1	-1
2	20	20	1	1	1.0	1	1
3	2	2	1	1	-1.0	-1	1
4	3	3	1	1	-1.0	1	-1
5	9	9	1	1	0.0	-1	-1
6	8	8	1	1	-0.5	1	1
7	12	12	1	1	0.0	1	1
8	14	14	1	1	0.5	-1	1
9	15	15	1	1	0.5	1	-1
10	2	2	1	1	-1.0	-1	1
11	3	3	1	1	-1.0	1	-1
12	5	5	1	1	-0.5	-1	-1
13							

Ensuite, Minitab fait à notre connaissance une bizarrerie, il recode implicitement les colonnes sous la forme suivante:

Matrix XMAT1

```

1  0  1  0  0 -1 -1
1  0  0  1  0 -1 -1
1  0  0  0  1  1 -1
1  0  0  0  1 -1  1
1  0  1  0  0  1 -1
1  0  0  1  0  1  1
1  0  1  0  0 -1  1
1  1  0  0  0  1  1
1  0  0  1  0  1  1
1 -1 -1 -1 -1  1 -1
1  1  0  0  0 -1 -1
1 -1 -1 -1 -1 -1  1

```

où les 4 premières colonnes de la matrice d'expérience ci-dessus (en excluant bien sûr la 1ère colonne remplie de 1) correspondent aux niveaux -1, -0.5, 0 et 0.5 du facteur A, (1 étant le niveau de référence).

La première colonne de la matrice d'expérience contient donc 1 à chaque fois que -1 apparaît dans la matrice de données pour le facteur A, -1 à chaque fois que le niveau de référence 1 apparaît dans la matrice de données pour le facteur A, et 0 pour toutes les autres valeurs.

Dès lors sous ces conditions, nous retrouvons effectivement les valeurs mises en évidence en rouge retournées dans la fenêtre de session par Minitab si nous faisons les calculs dans un logiciel spécialisé ou un tableur:

**Résultats pour : Feuille de travail 5****Plan optimal : A; B; C**

Plan factoriel sélectionné à l'aide de l'optimalité D

Nombre de points de plan candidats : 20

Nombre de points de plan dans le plan optimal : 12

Termes du modèle : A; B; C

Plan initial généré par la méthode séquentielle

Plan initial amélioré par la méthode des échanges

Nombre de points de plan échangés : 1

## Plan optimal

Numéro des lignes des points du plan sélectionné : 17; 20; 2; 3; 9; 8; 12; 14;  
15; 2; 3; 5

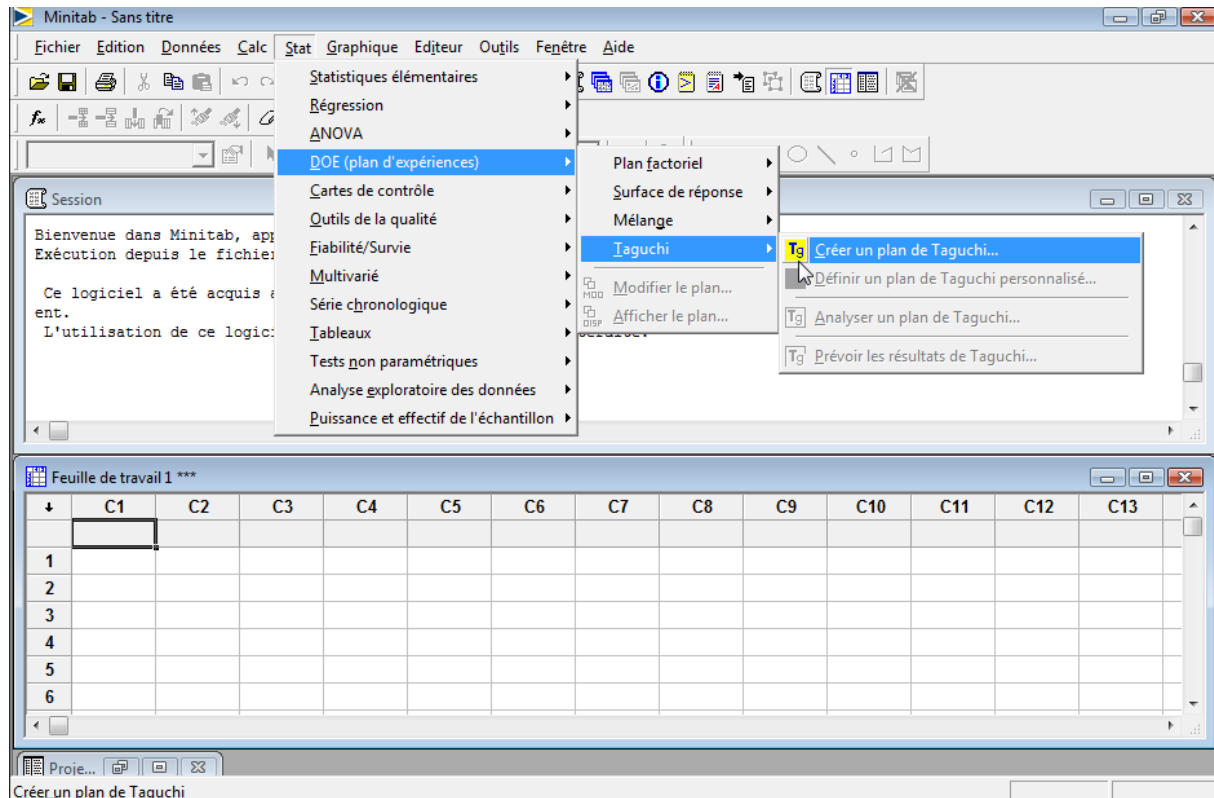
Conditionnement :	4.76569
Optimalité D (déterminant de $XTX$ ) :	230400
Optimalité A (trace de $\text{inv}(XTX)$ ) :	1.66667
Optimalité G (effet de levier moyen/effet de levier maximum) :	0.875
Optimalité V (effet de levier moyen) :	0.583333
Effet de levier maximum :	0.666667

Pour ce qui concerne les indices d'optimalité D/A/G étudiés dans le cours théorique, nous retrouvons le même que celui calculé à la main en ce qui concerne le D (même si la matrice d'expérience de Minitab est douteuse) et pour les indices A/G nous retrouvons des valeurs similaires mais pas tout à fait égales.

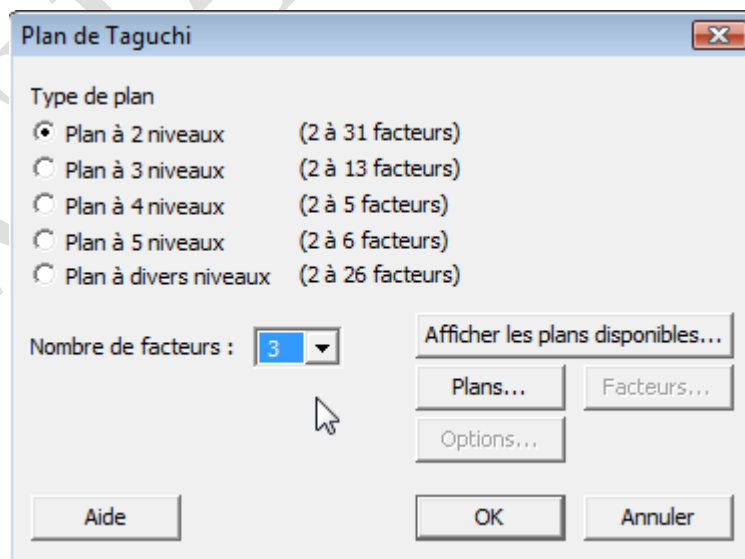
Donc prudence d'ici à ce que ceci soit éclairci par l'équipe Minitab (je suis en attente d'une réponse détaillée de leur part après leur avoir posé la question en septembre 2013)!

### 17.1.4. Construction d'un plan de Taguchi

Maintenant, pour en finir avec la base des plans d'expérience, voyons encore les plans de Taguchi. Nous allons dans le menu **DOE (plan d'expériences)/Taguchi/Créer un plan de Taguchi...**:



Nous allons prendre un **Plan à 2 niveaux** (donc un plan factoriel à 2 niveaux) avec la valeur **3** pour **nombre de facteurs** comme dans le cours théorique:



Nous cliquons par curiosité sur **Afficher les plans disponibles...**:

Plan de Taguchi - Afficher les plans disponibles

Plans de Taguchi disponibles (avec nombre de facteurs)

Plans	A un niveau plans			
	2 niveaux	3 niveaux	4 niveaux	5 niveaux
L4	2-3			
L8	2-7			
L9		2-4		
L12	2-11			
L16	2-15			
L16			2-5	
L25				2-6
L27		2-13		
L32	2-31			

A un niveau
  Mixte à 2-3 niveaux
  Mixte à 2-4 niveaux
  Mixte

Aide OK

Nous voyons qu'il y a donc la majorité des plans de Taguchi qui sont disponibles dans Minitab® Statistical Software. Il n'y pas les graphes linéaires par contre ce qui pourrait être jugé par certains comme étant un manque. Voilà ce que nous trouvons par exemple dans l'onglet **Mixte à 2-3 niveaux**:

Plan de Taguchi - Afficher les plans disponibles

Plans de Taguchi disponibles (avec nombre de facteurs)

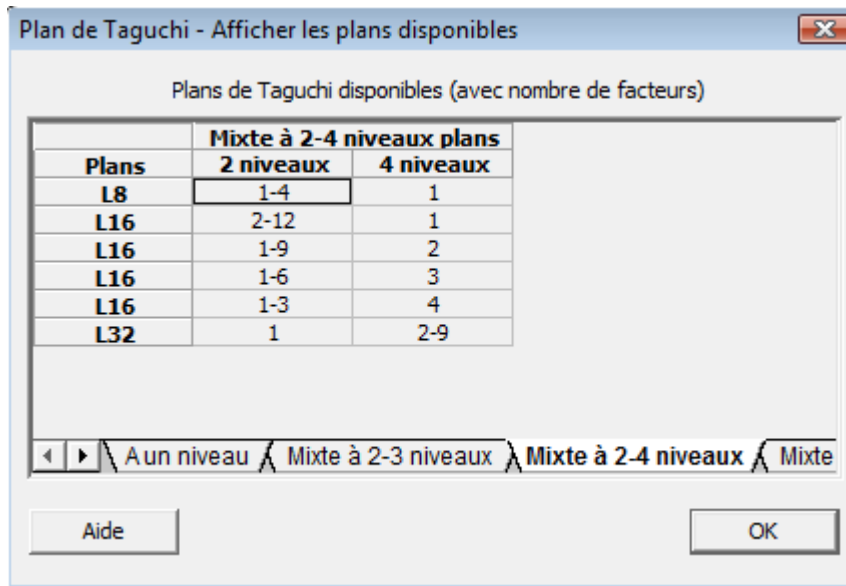
Plans	Mixte à 2-3 niveaux plans	
	2 niveaux	3 niveaux
L18	1	1-7
L36	1-11	2-12
L36	1-3	13
L54	1	3-25

A un niveau
  Mixte à 2-3 niveaux
  Mixte à 2-4 niveaux
  Mixte

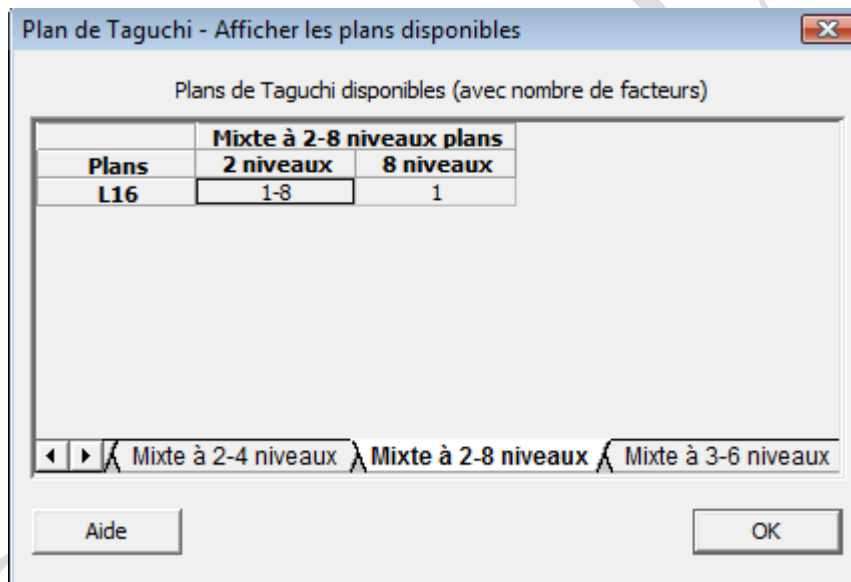
Aide OK

et dans l'onglet **Mixte à 2-4 niveaux**:

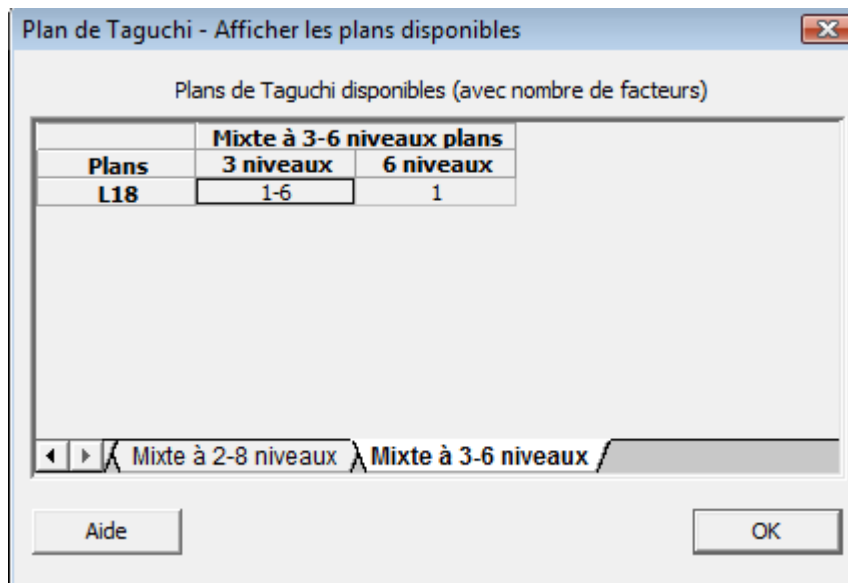




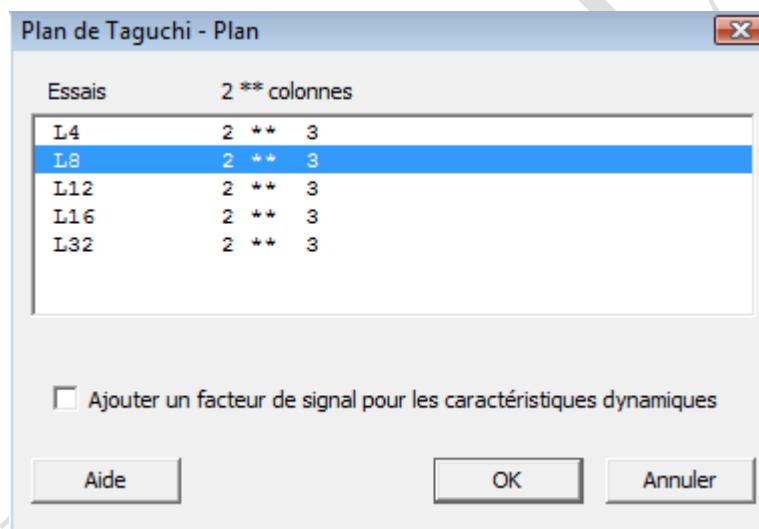
et dans l'onglet **Mixte à 2-8 niveaux**:



et enfin dans l'onglet **Mixte à 3-6 niveaux**:



Bon ceci étant vu, cliquons sur le bouton **Plans...** pour avoir:



Nous prenons le plan L8 comme celui étudié dans le cours théorique. Ensuite, nous cliquons sur le bouton **Facteurs...**:

Plan de Taguchi - Facteurs

Affecter les facteurs

Aux colonnes du répertoire, comme spécifié ci-dessous

Pour permettre l'estimation des interactions sélectionnées

Facteu	Nom	Valeurs des niveaux	Colonne	Niveau
A	A	1 2	1	2
B	B	1 2	2	2
C	C	1 2	4	2

Aide

Nous vérifions la façon dont les niveaux seront notés (nous laissons la notation traditionnelle de Taguchi). Nous voyons aussi que Minitab® Statistical Software choisi comme nous l'avons vu dans le cours théorique de mettre les trois facteurs A, B et C dans les colonnes respectives 1, 2 et 4.

Nous validons le tout par **OK** et obtenons d'abord dans la fenêtre de session:

### Plan de Taguchi

Plan de Taguchi orthogonal

L8 (2\*\*3)

Facteurs : 3  
Essais : 8

Colonnes de L8 (2\*\*7) Répertoire

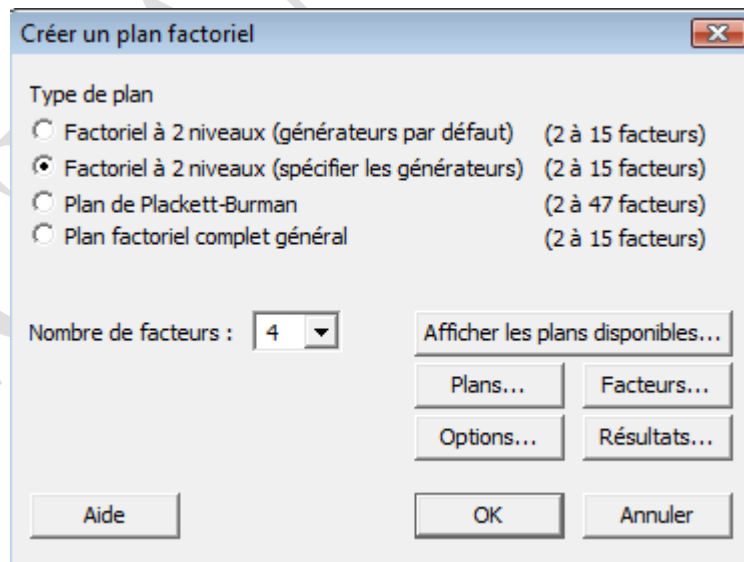
1 2 4

c'est un petit résumé. Il faut dire que la société responsable de Minitab® Statistical Software avait la marge de faire beaucoup mieux. Mais bon on ne va pas se plaindre, c'est déjà pas mal. Le plan lui est donné:

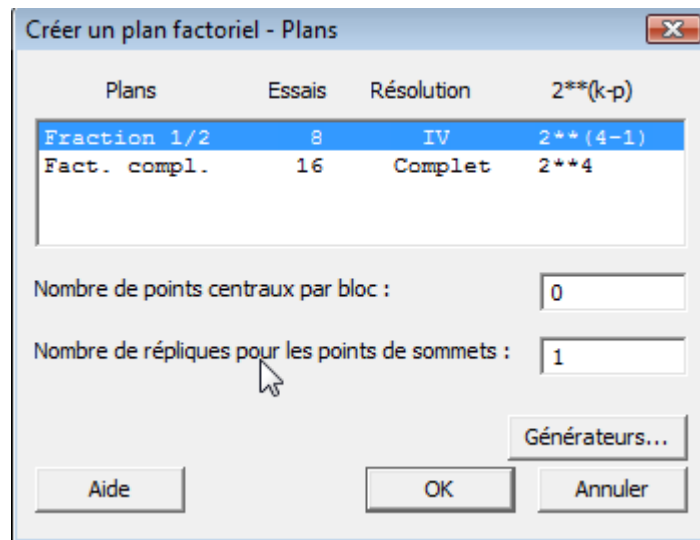
	C1	C2	C3	C4	C5
	A	B	C		
1	1	1	1		
2	1	1	2		
3	1	2	1		
4	1	2	2		
5	2	1	1		
6	2	1	2		
7	2	2	1		
8	2	2	2		
9					
10					

et là encore une fois, nous avons le minimum, minimorum, ce qui est dommage bien que ce soit juste. Il aurait été bien de pouvoir choisir si nous voulions la table de Taguchi complète ou non.

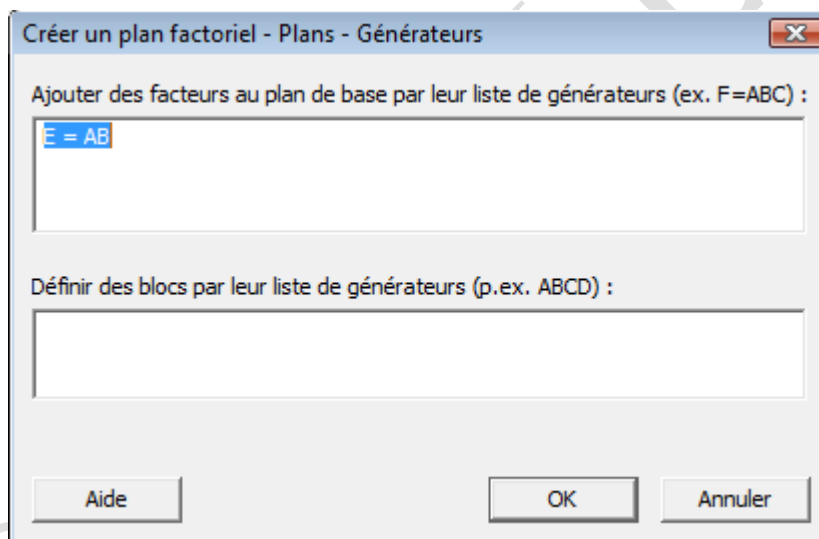
Maintenant, voyons si nous retrouvons ce plan de Taguchi avec le plan factoriel "classique". Nous allons donc le menu déjà vu plus haut:



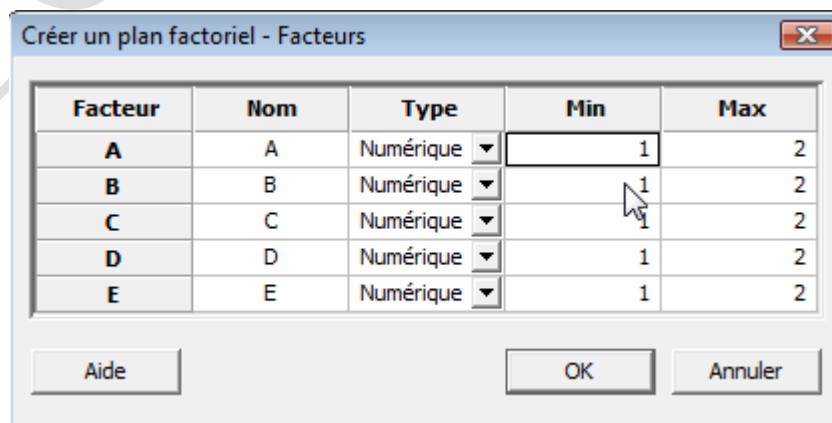
Nous allons dans le bouton **Plans...**:



et prenons le plan fractionnaire 1/2 de résolution IV. Nous savons que pour obtenir le plan de Taguchi à partir d'un plan classique, nous devons au moins spécifier un générateur. Nous prendrons donc la 5<sup>ème</sup> colonne (E) comme étant la multiplication de A et B (AB) soit:



Nous définissons aussi la notation des facteurs en cliquant sur le bouton **Facteurs...** afin d'avoir la même notation que celle de Taguchi:



et nous validons le tout par **OK** pour obtenir dans la fenêtre de session:

**Plan factoriel fractionnaire**

Facteurs : 5 Plan de base : 4; 8 Résolution : III  
 Essais : 8 Répliques : 1 Fraction : 1/4  
 Blocs : 1 Points centraux (total) : 0

\* REMARQUE \* Certains effets principaux sont confondus avec des interactions à deux facteurs.

Alias initiales (indépendantes) : D = ABC; E = AB

Structure des alias (jusqu'à l'ordre 4)

I + ABE + CDE + ABCD

- A + BE + BCD + ACDE
- B + AE + ACD + BCDE
- C + DE + ABD + ABCE
- D + CE + ABC + ABDE
- E + AB + CD + ABCDE
- AC + BD + ADE + BCE
- AD + BC + ACE + BDE

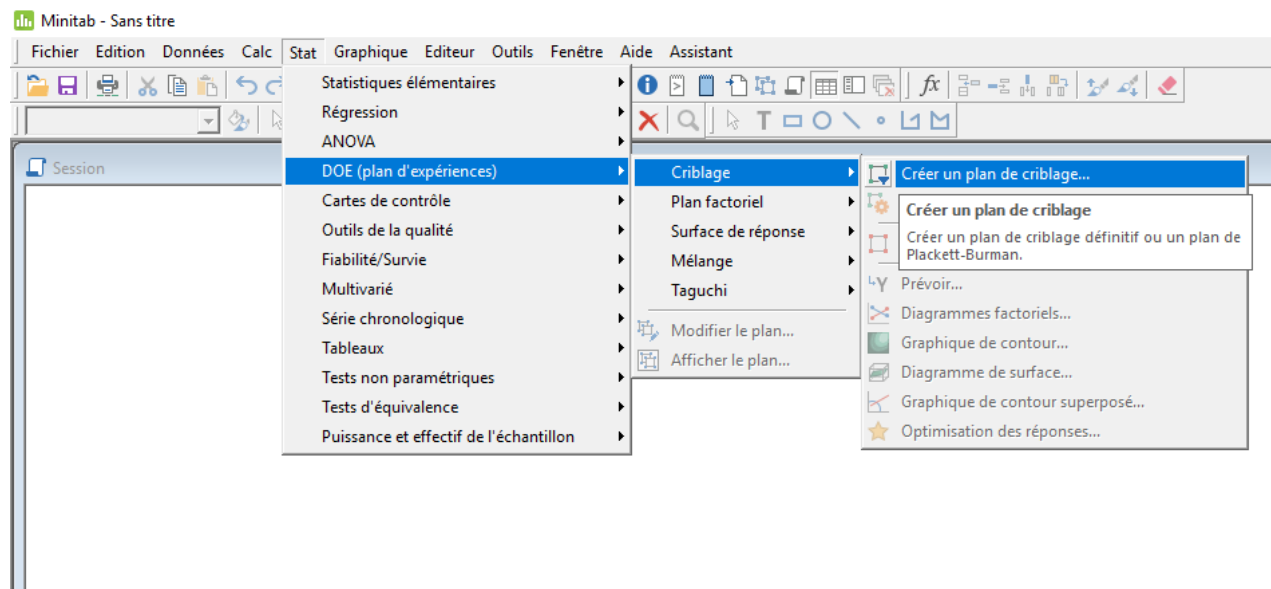
Nous voyons ainsi les alias qu'il nous faudrait prendre pour avoir la même chose qu'une table de Taguchi (nous avons représenté en rouge l'équivalent de la table de Taguchi). Évidemment c'est du bricolage mais cela montre juste qu'on peut retomber dessus.

↓	C1	C2	C3	C4	C5	C6	C7	C8	C9
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	C	D	E
1	1	1	1	1	1	1	1	1	2
2	2	2	1	1	2	1	1	2	1
3	3	3	1	1	1	2	1	2	1
4	4	4	1	1	2	2	1	1	2
5	5	5	1	1	1	1	2	2	2
6	6	6	1	1	2	1	2	1	1
7	7	7	1	1	1	2	2	1	1
8	8	8	1	1	2	2	2	2	2
9									

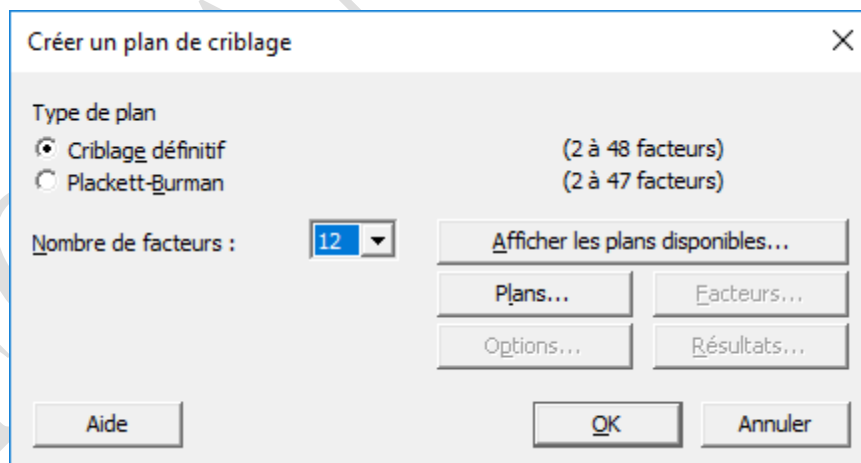
### 17.1.5. Création d'un plan de criblage définitif (Minitab >18)

Comme pour les autres plans, nous allons d'abord ici juste vérifier que nous arrivons à retomber sur plan d'expérience vu dans le cours théorique. Ici le but va être de vérifier si on peut effectivement reproduire le plan  $C(m=12, n=2)$  vu dans le cours théorique (on va pas tous les faire non plus...).

Pour cela, à partir de Minitab 18, nous allons dans le menu **DOE (plan d'expériences)/Criblage/Créer un plan de criblage...**:

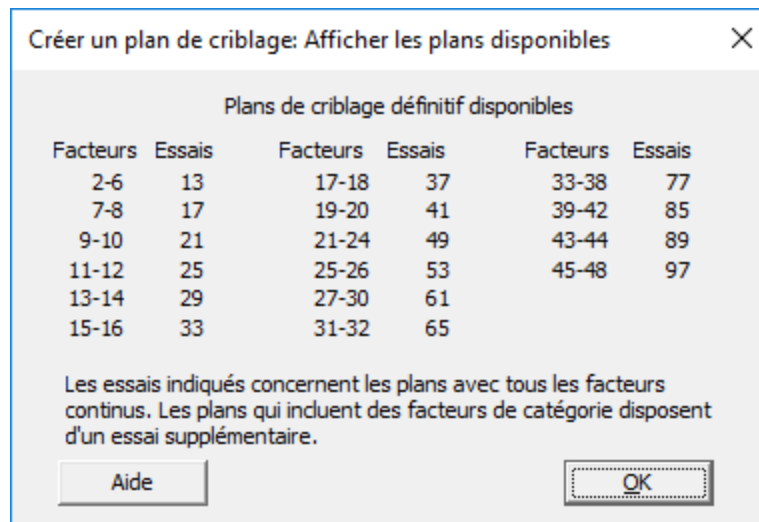


Ce qui nous donne:

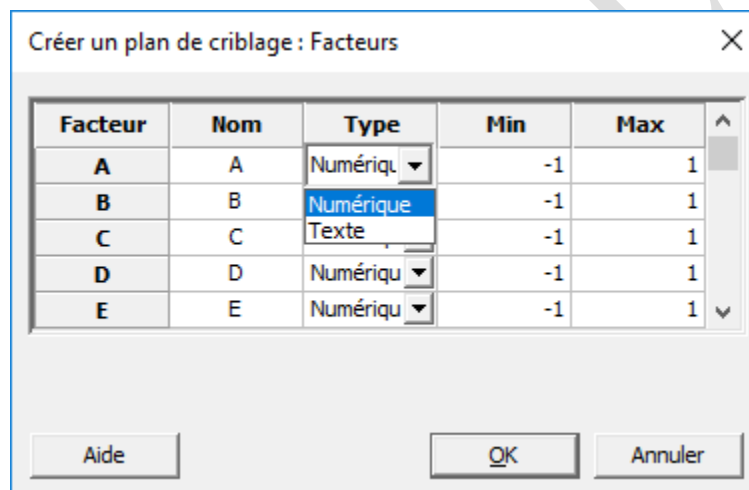


où nous choisissons le plan de **Criblage définitif** (pour rappel en anglais: Definitive Screening Design) avec 12 facteurs.

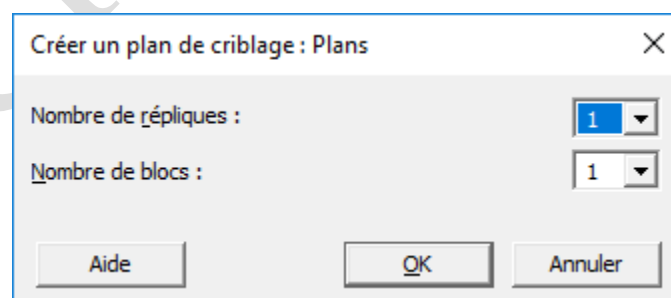
Nous cliquons sur le bouton **Afficher les plans disponibles...** pour constater que nous retrouvons ce que nous avons vu dans le cours théorique:



En cliquant sur le bouton Facteurs... nous avons le classique (notez que le troisième niveau 0 n'est pas indiqué et rappelez-vous que les facteurs catégoriels diminuent la robustesse des analyses de plans d'expérience en général):

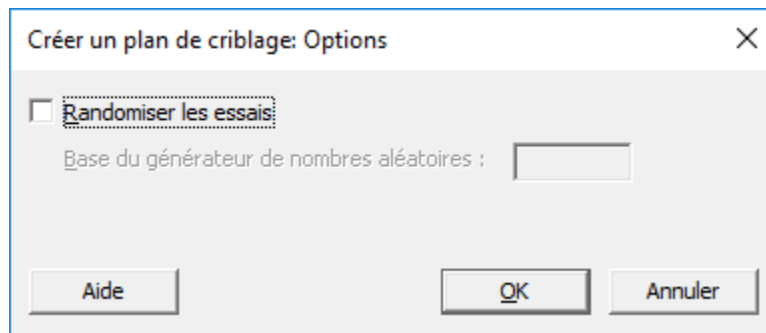


Ensuite, nous cliquons sur le bouton **Plans...**:



Et pour faciliter le contrôle, dans Options... nous allons désactiver la randomisation:





Et nous validons deux fois par **OK** pour obtenir:

Minitab - Sans titre - [Feuille de travail 2 \*\*\*]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assistant

↓	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C
	OrdreStd	OrdEssai	TypePt	Blocs	A	B	C	D	E	F	G	H	J	K	L	M	
1	1	1	2	1	0	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	1	0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
3	3	3	2	1	1	0	-1	-1	-1	-1	1	-1	1	1	1	1	1
4	4	4	2	1	-1	0	1	1	1	1	-1	1	-1	-1	-1	-1	-1
5	5	5	2	1	1	1	0	1	1	-1	1	-1	-1	1	-1	-1	-1
6	6	6	2	1	-1	-1	0	-1	-1	1	-1	1	1	-1	1	1	1
7	7	7	2	1	1	1	-1	0	1	1	-1	-1	-1	-1	1	1	1
8	8	8	2	1	-1	-1	1	0	-1	-1	1	1	1	1	-1	-1	-1
9	9	9	2	1	1	1	-1	-1	0	1	-1	1	1	1	-1	-1	-1
10	10	10	2	1	-1	-1	1	1	0	-1	1	-1	-1	-1	1	1	1
11	11	11	2	1	1	1	1	-1	-1	0	1	1	-1	-1	1	-1	-1
12	12	12	2	1	-1	-1	-1	1	1	0	-1	-1	1	1	-1	1	1
13	13	13	2	1	1	-1	-1	1	1	-1	0	1	1	-1	1	-1	-1
14	14	14	2	1	-1	1	1	-1	-1	1	0	-1	-1	1	-1	1	1
15	15	15	2	1	1	1	1	1	-1	-1	-1	0	1	-1	-1	1	1
16	16	16	2	1	-1	-1	-1	-1	1	1	1	0	-1	1	1	-1	-1
17	17	17	2	1	1	-1	1	1	-1	1	-1	-1	0	1	1	-1	-1
18	18	18	2	1	-1	1	-1	-1	1	-1	1	1	0	-1	-1	1	1
19	19	19	2	1	1	-1	-1	1	-1	1	1	1	-1	0	-1	1	1
20	20	20	2	1	-1	1	1	-1	1	-1	-1	-1	1	0	1	-1	-1
21	21	21	2	1	1	-1	1	-1	1	-1	-1	1	-1	1	0	1	1
22	22	22	2	1	-1	1	-1	1	-1	1	1	-1	1	-1	0	-1	-1
23	23	23	2	1	1	-1	1	-1	1	1	1	-1	1	-1	-1	0	0
24	24	24	2	1	-1	1	-1	1	-1	-1	-1	1	-1	1	1	0	0
25	25	25	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
26																	

Ce qui est bien conforme à ce que nous avons vu dans le cours théorique (à part que le format dans le cours théorique était plus... "lisible")!

## 17.2. Exercice 173.: Comparaison d'un plan d'expérience identique en factoriel complet général et factoriel à 2 niveaux

Minitab® Statistical Software 15.1.1.0

Avant de passer au gros exercice fait à la main dans le cours théorique il va être utile de commencer par un cas hyper simpliste qui avait consisté à résoudre à la main le système suivant:

$$a_0 \cdot 1 + a_1 (40) + a_2 (1.5) + a_{12} (40)(1.5) \cong 32'700$$

$$a_0 \cdot 1 + a_1 (50) + a_2 (1.5) + a_{12} (50)(1.5) \cong 32'680$$

$$a_0 \cdot 1 + a_1 (40) + a_2 (3) + a_{12} (40)(3) \cong 31'710$$

$$a_0 \cdot 1 + a_1 (50) + a_2 (3) + a_{12} (50)(3) \cong 33'220$$

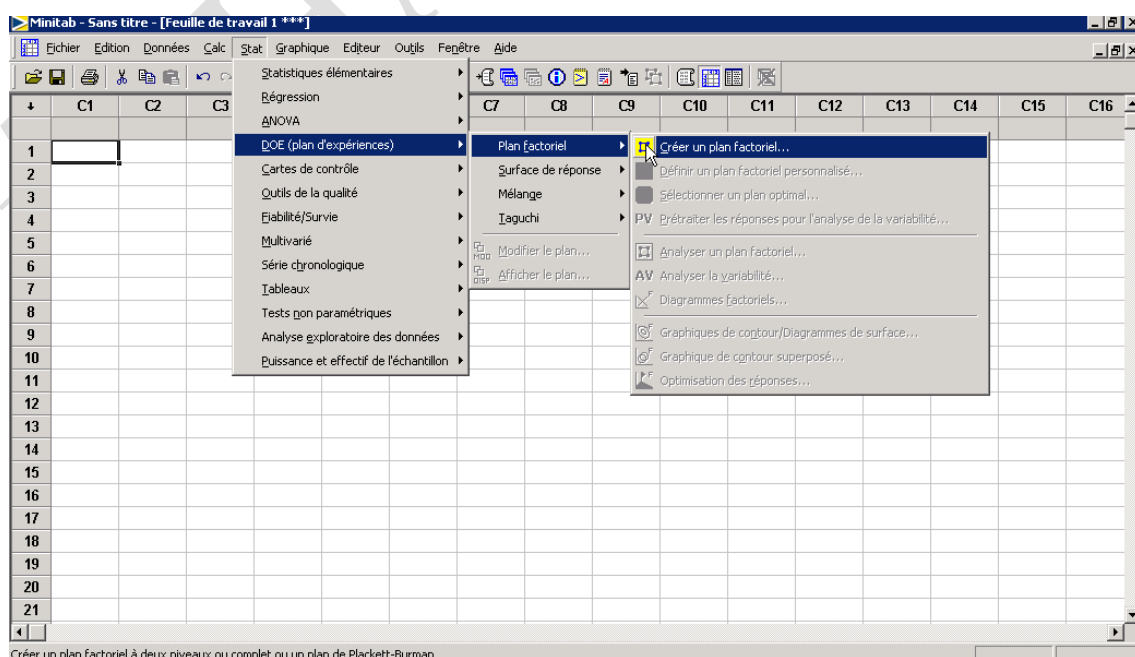
Ce qui nous avait donné sous forme matricielle:

$$\begin{bmatrix} 1 & 40 & 1.5 & 60 \\ 1 & 50 & 1.5 & 75 \\ 1 & 40 & 3 & 120 \\ 1 & 50 & 3 & 150 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_{12} \end{bmatrix} = \begin{bmatrix} 32'700 \\ 32'680 \\ 31'710 \\ 33'220 \end{bmatrix}$$

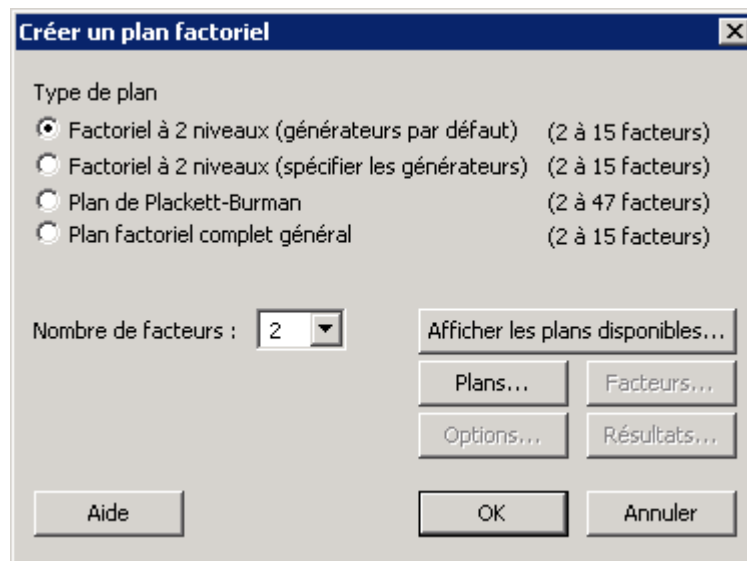
La solution déterminée avait été:

$$y \cong 39'890 - 155x_1 - 4740x_2 + 102x_{12}$$

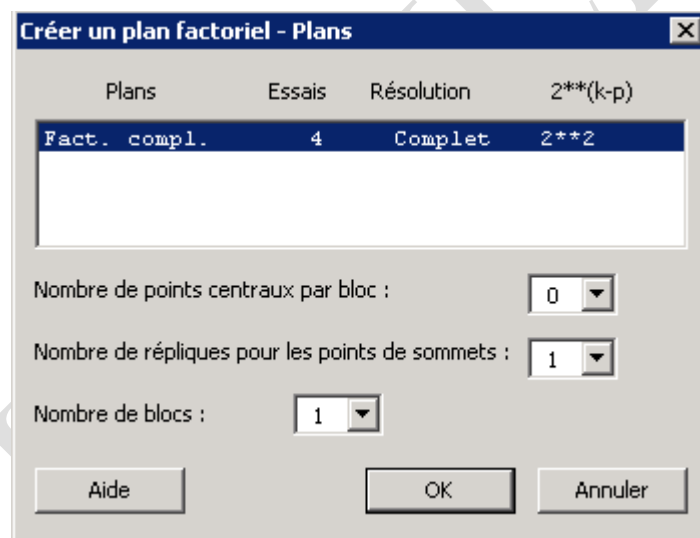
Vérifions cela avec Minitab. Nous allons dans **Stat/DOE (plan d'expériences)/Plan factoriel/Créer un plan factoriel...**:



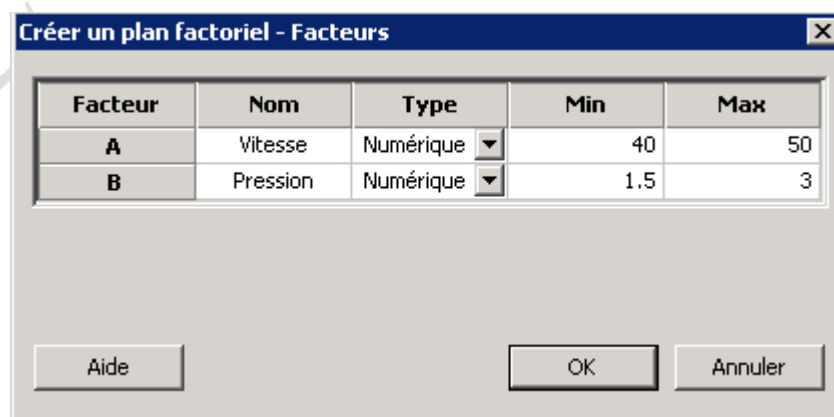
et nous allons prendre un plan **Factoriel à 2 niveaux (générateurs par défaut)** avec **2** facteurs:



Nous cliquons sur le bouton **Plans...** pour avoir le choix logique:



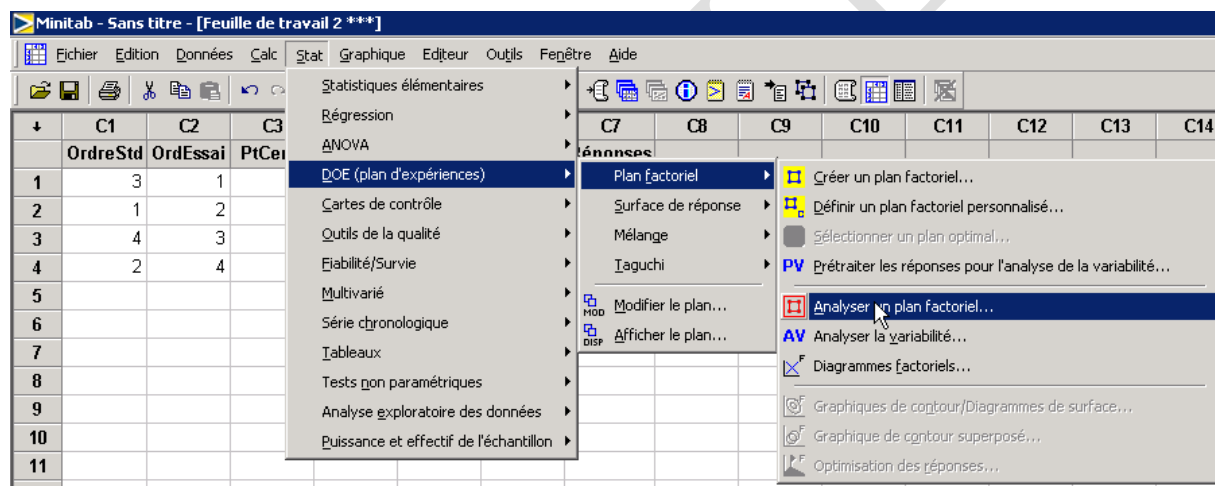
Ensuite, nous cliquons sur le bouton **Facteurs...** pour y mettre:



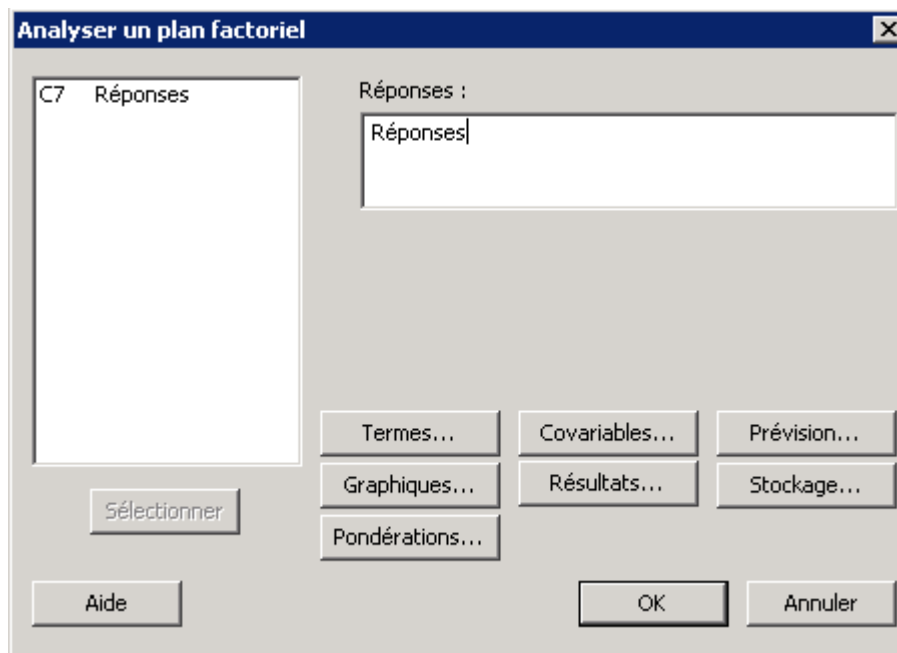
Et nous validons le tout par **OK** et nous saisissons ensuite les valeurs mesurées (Réponses) dans la feuille:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
	OrdreStd	OrdEssai	PtCentr	Blocs	Vitesse	Pression	Réponses				
1	3	1	1	1	40	3.0	31710				
2	1	2	1	1	40	1.5	32700				
3	4	3	1	1	50	3.0	33220				
4	2	4	1	1	50	1.5	32680				
5											
6											
7											
8											

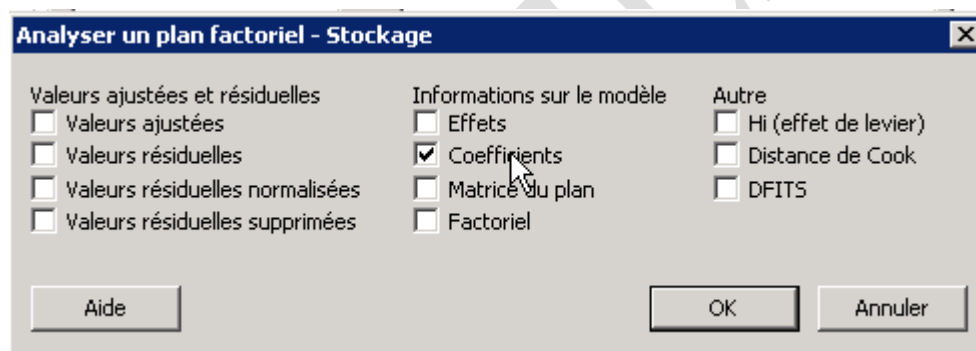
et nous lançons l'analyse en allant dans le menu **Stat/DOE (plan d'expériences)/Plan factoriel/Analyser un plan factoriel...**:



et nous avons alors:



Une seule chose nous intéresse dans cet exercice! Les coefficients!! Dès lors, nous allons dans le bouton **Stockage...** et sélectionnons la case **Coefficients**:



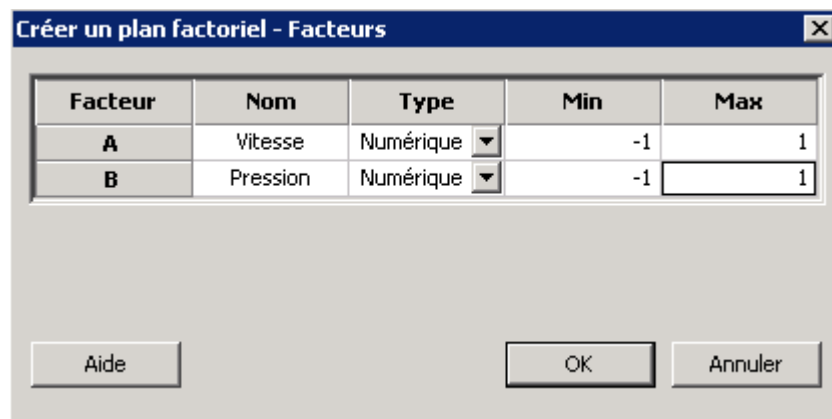
En validant le tout par **OK**, nous avons alors dans la feuille:

	C1	C2	C3	C4	C5	C6	C7	C8	C9
	OrdreStd	OrdEssai	PtCentr	Blocs	Vitesse	Pression	Réponses	COEFF1	
1	3	1	1	1	40	3.0	31710	39890	
2	1	2	1	1	40	1.5	32700	-155	
3	4	3	1	1	50	3.0	33220	-4740	
4	2	4	1	1	50	1.5	32680	102	
5									

Soit exactement les mêmes coefficients que ceux calculés à la main:

$$y \cong 39'890 - 155x_1 - 4740x_2 + 102x_1x_2$$

Bien évidemment si le tableau d'expérience avait été en données codées avec par exemple:



Cela n'aurait bien évidemment pas donné le même comme le montre la capture d'écran du résultat d'analyse ci-dessous:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
	OrdreStd	OrdEssai	PtCentr	Blocs	Vitesse	Pression	Réponses	COEFF1			
1	1	1	1	1	-1	-1	32700	32577.5			
2	3	2	1	1	-1	1	31710	372.5			
3	4	3	1	1	1	1	33220	-112.5			
4	2	4	1	1	1	-1	32680	382.5			
5											

Coefficients trivialement calculés avec la relation démontrée dans le cours théorique:

$$a_0 = \frac{1}{4}(y_1 + y_2 + y_3 + y_4)$$

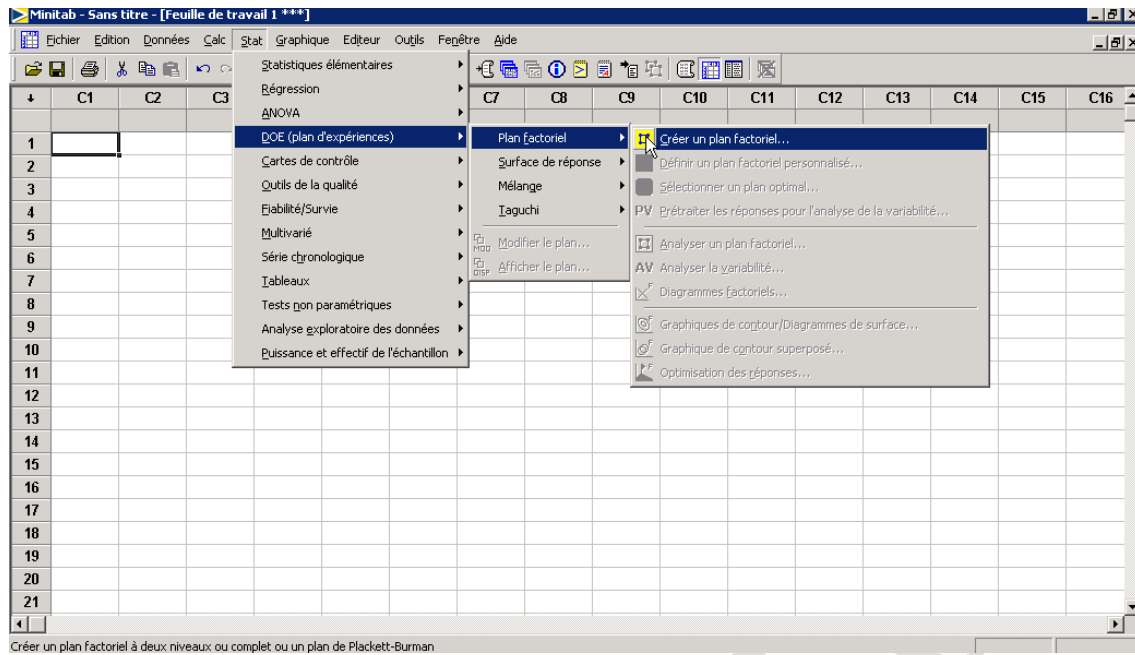
$$a_1 = \frac{1}{4}(-y_1 + y_2 - y_3 + y_4)$$

$$a_2 = \frac{1}{4}(-y_1 - y_2 + y_3 + y_4)$$

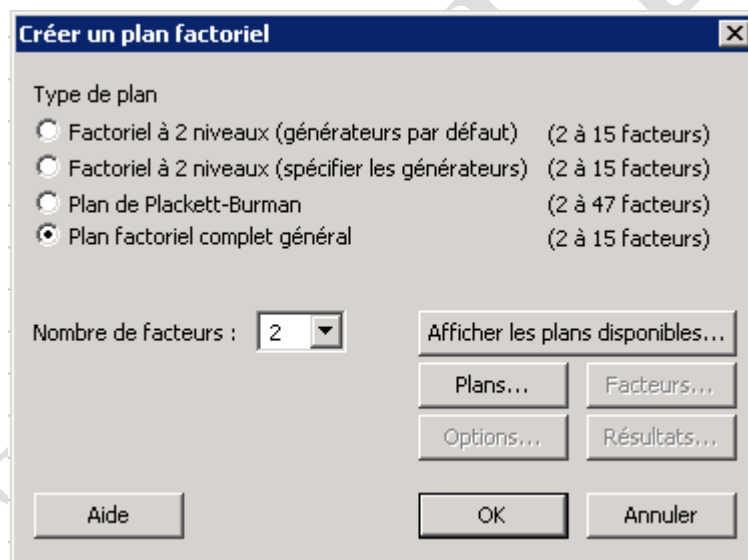
$$a_{12} = \frac{1}{4}(y_1 - y_2 - y_3 + y_4)$$

Bon tout est normal et rassurant!

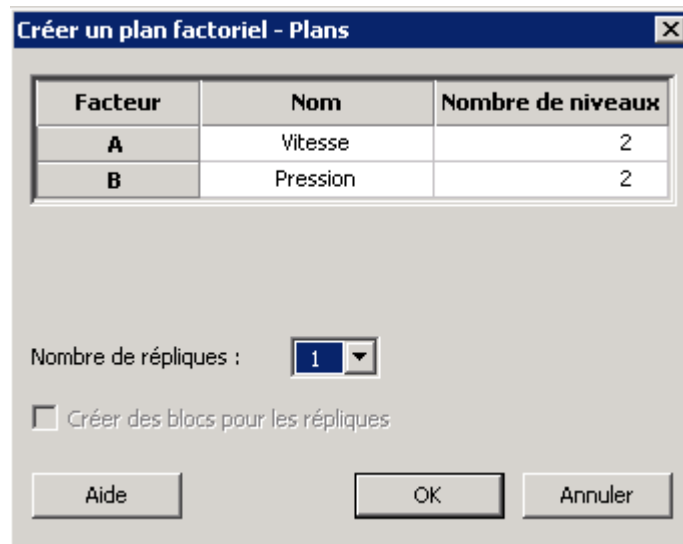
Maintenant procédons à une autre variante dont le résultat est moins intuitif. Nous recommençons en allant dans le menu **Stat/DOE (plan d'expériences)/Plan factoriel/Créer un plan factoriel...**:



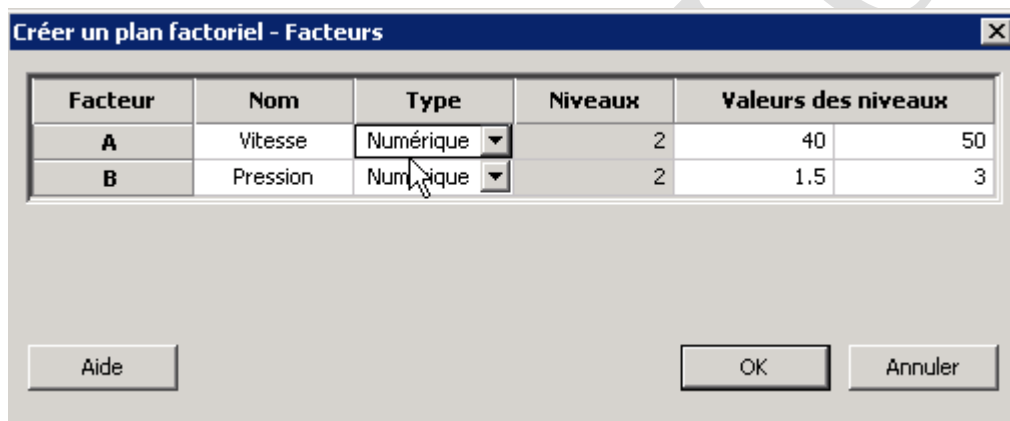
et nous prenons cette fois-ci un **Plan factoriel complet général** avec 2 facteurs:



et nous cliquons sur **Plans...** pour avoir:



Ensuite, nous cliquons sur **Facteurs...** pour y mettre:



Nous validons le tout par **OK** pour ensuite mettre les valeurs mesurées (remarquez que nous avons le même tableau qu'au début de l'exercice donc a priori nous devrions avoir le même résultat):

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
	OrdreStd	OrdEssai	TypePt	Blocs	Vitesse	Pression	Réponses				
1	3	1	1	1	50	1.5	32680				
2	4	2	1	1	50	3.0	33220				
3	2	3	1	1	40	3.0	31710				
4	1	4	1	1	40	1.5	32700				
5											
6											

et en lançant l'analyse exactement comme pour les exemples précédents, nous obtenons:



	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
	OrdreStd	OrdEssai	TypePt	Blocs	Vitesse	Pression	Réponses	COEFF1			
1	3	1	1	1	50	1.5	32680	32577.5			
2	4	2	1	1	50	3.0	33220	-372.5			
3	2	3	1	1	40	3.0	31710	112.5			
4	1	4	1	1	40	1.5	32700	382.5			
5											
6											
7											

Soit un résultat complètement différent à :

$$y \cong 39'890 - 155x_1 - 4740x_2 + 102x_1x_2$$

Au fait si le lecteur a bien fait attention Minitab renvoi les coefficients pour les facteurs codés obtenus par :

$$a_0 = \frac{1}{4}(y_1 + y_2 + y_3 + y_4)$$

$$a_1 = \frac{1}{4}(-y_1 + y_2 - y_3 + y_4)$$

$$a_2 = \frac{1}{4}(-y_1 - y_2 + y_3 + y_4)$$

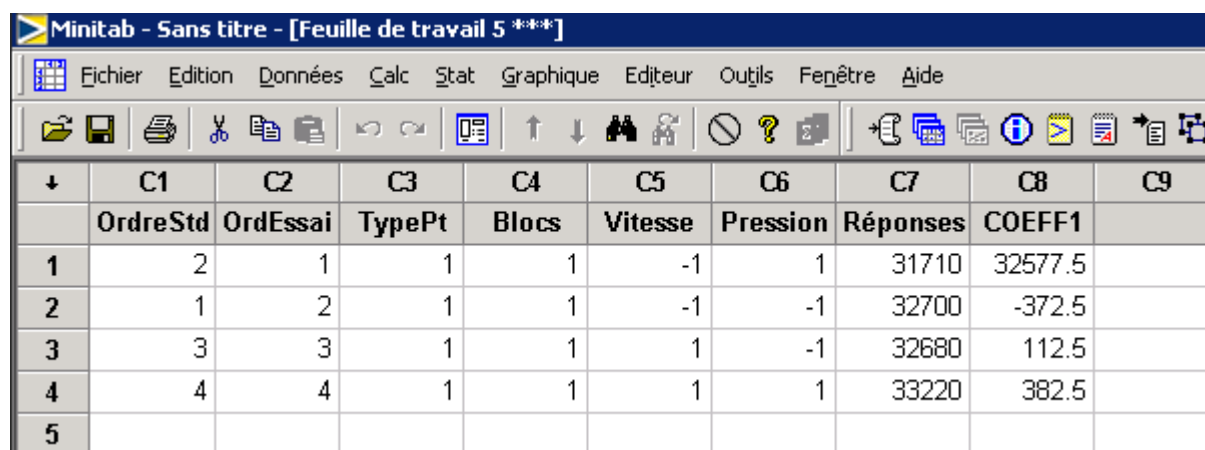
$$a_{12} = \frac{1}{4}(y_1 - y_2 - y_3 + y_4)$$

Pour la simple raison qu'il n'est pas possible de faire autrement dans le cas général!

Pour vérifier, nous pouvons relance le même plan d'expérience et en codant les facteurs:

Facteur	Nom	Type	Niveaux	Valeurs des niveaux	
A	Vitesse	Numérique	2	-1	1
B	Pression	Numérique	2	-1	1

Nous arrivons au même résultat (donc avec les mêmes coefficients!!):



Minitab - Sans titre - [Feuille de travail 5 \*\*\*]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

	C1	C2	C3	C4	C5	C6	C7	C8	C9
	OrdreStd	OrdEssai	TypePt	Blocs	Vitesse	Pression	Réponses	COEFF1	
1	2	1	1	1	-1	1	31710	32577.5	
2	1	2	1	1	-1	-1	32700	-372.5	
3	3	3	1	1	1	-1	32680	112.5	
4	4	4	1	1	1	1	33220	382.5	
5									

C.Q.F.D....

### 17.3. Exercice 174.: Analyse de la variance multifactorielle

Minitab® Statistical Software 15.1.1.0

Avant de commencer à analyser une expérience avec les outils de plans d'expérience, il nous semble pertinent d'abord d'étudier les données avec une ANOVA multifactorielle pour comparer ensuite les résultats avec les outils dédiés aux plans d'expérience.

Rappelons que nous avons étudié dans le cours de statistique théorique une grande partie des plans d'expérience sur la base des données du tableau visible ci-dessous. Nous allons procéder cette fois-ci de même mais avec le logiciel Minitab® Statistical Software. Donc nous avons:

N° Essai	$x_1$	$x_2$	$x_3$	Valeurs	de	y
1	1	1	1	32'700	32'750	32'960
2	1	1	2	33'430	33'360	32'910
3	1	1	3	31'710	32'100	32'220
4	1	2	1	32'680	32'270	33'130
5	1	2	2	34'070	33'100	33'610
6	1	2	3	33'220	33'700	33'285
7	2	1	1	33'180	32'160	32'640
8	2	1	2	34'430	34'280	34'460
9	2	1	3	33'570	33'300	32'570
10	2	2	1	33'270	33'080	32'415
11	2	2	2	33'440	33'570	34'204
12	2	2	3	32'840	33'210	32'470

Dans Minitab® Statistical Software, nous voulons comme dans le cours théorique, faire d'abord l'analyse de la variance à un facteur pour le **type de route uniquement** (facteur 1: où 1 est la ville et 2 est la route).

Dans Minitab® Statistical Software, nous avons alors avec le fichier *ANOVA\_Balancee\_Desempillee.mpj*:

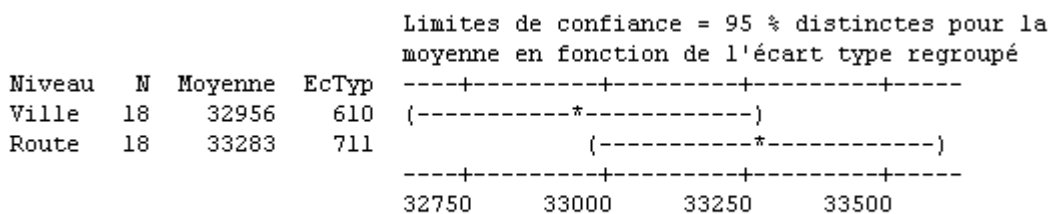
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Ville	Route	40 km/h	50 km/h	1.5 kg	2 kg	2.5 kg							
1	32700	33180	32700	32680	32700	33430	31710							
2	33430	34430	33430	34070	32680	34070	33220							
3	31710	33570	31710	33220	33180	34430	33570							
4	32680	33270	33180	33270	33270	33440	32840							
5	34070	33440	34430	33440	32750	33360	32100							
6	33220	32840	33570	32840	32270	33100	33700							
7	32750	32160	32750	32270	32160	34280	33300							
8	33360	34280	33360	33100	33080	33570	33210							
9	32100	33300	32100	33700	32960	32910	32220							
10	32270	33080	32160	33080	33130	33610	33285							
11	33100	33570	34280	33570	32640	34460	32570							
12	33700	33210	33300	33210	32415	34204	32470							
13	32960	32640	32960	33130										
14	32910	34460	32910	33610										
15	32220	32570	32220	33285										
16	33130	32415	32640	32415										
17	33610	34204	34460	34204										
18	33285	32470	32570	32470										
19														
20														
21														

et nous faisons une ANOVA à un facteur désémpilé uniquement sur les colonnes *Ville* et *Route* en suivant la procédure déjà vue plus haut en détails. Nous avons alors (sans vérifier que toutes les conditions d'application soient respectées):

**ANOVA à un facteur contrôlé : Ville; Route**

Source	DL	Somme des carrés	CM	F	P
Facteur	1	961707	961707	2.19	0.148
Erreur	34	14906680	438432		
Total	35	15868387			

S = 662.1 R carré = 6.06 % R carré (ajust) = 3.30 %

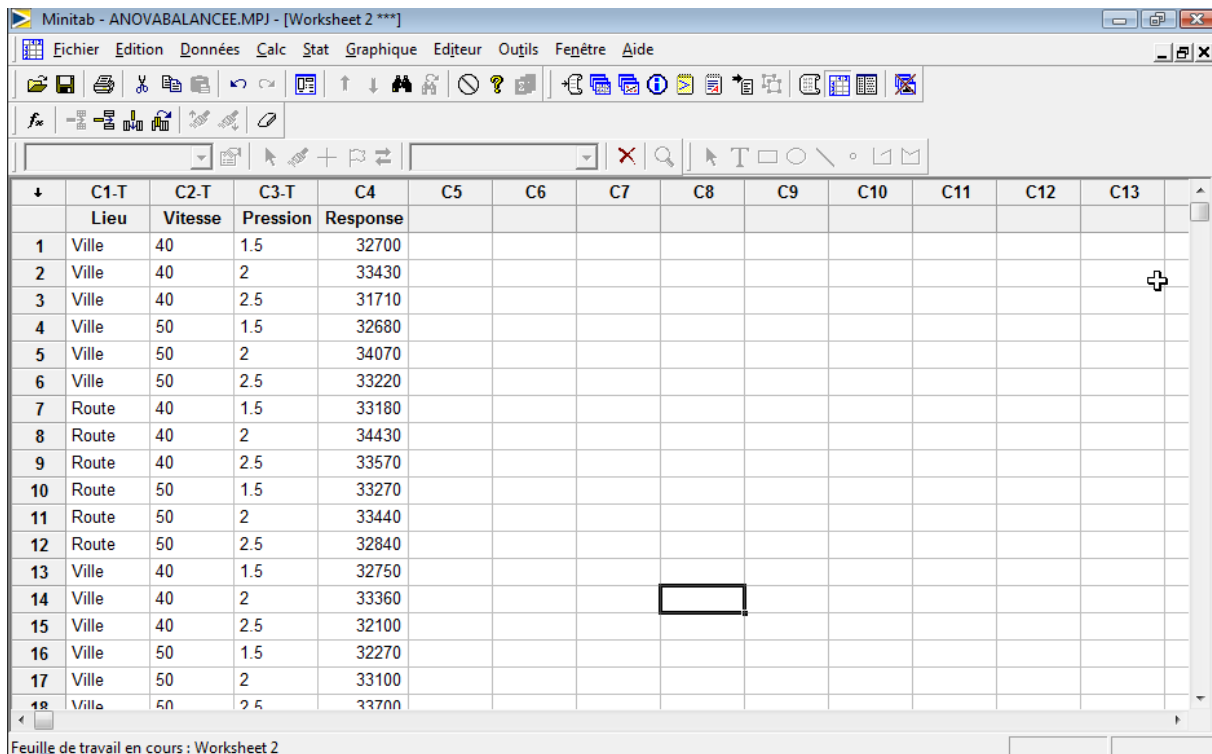


Ecart type regroupé = 662

Nous retrouvons la même chose qu'avec le calcul à la main et qu'avec Microsoft Excel. Donc il apparaît que le lieu n'influence pas de façon significative la distance parcourue (puisque l'hypothèse nulle d'égalité des moyennes réussit le test).

Nous pourrions faire de même pour chaque facteur comme dans le cours théorique, mais ce qui nous intéresse ici c'est l'analyse de la variance multifactorielle et la comparer au résultat obtenu plus haut.

Donc d'abord, dans Minitab® Statistical Software, nous préparons le tableau complet (mais seulement une partie apparaît dans la capture d'écran ci-dessous) ou ouvrons le fichier ANOVABalancee.mpj:



	C1-T	C2-T	C3-T	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Lieu	Vitesse	Pression	Response									
1	Ville	40	1.5	32700									
2	Ville	40	2	33430									
3	Ville	40	2.5	31710									
4	Ville	50	1.5	32680									
5	Ville	50	2	34070									
6	Ville	50	2.5	33220									
7	Route	40	1.5	33180									
8	Route	40	2	34430									
9	Route	40	2.5	33570									
10	Route	50	1.5	33270									
11	Route	50	2	33440									
12	Route	50	2.5	32840									
13	Ville	40	1.5	32750									
14	Ville	40	2	33360									
15	Ville	40	2.5	32100									
16	Ville	50	1.5	32270									
17	Ville	50	2	33100									
18	Ville	50	2.5	33700									

Nous allons donc dans le menu **Stat/ANOVA/ANOVA équilibrée...**:

The screenshot shows the Minitab interface with the 'Stat' menu open and 'ANOVA' selected. The 'ANOVA équilibrée...' option is highlighted. The background data table is as follows:

	C1-T	C2-T	C3-T				
	Lieu	Vitesse	Pressi				
1	Ville	40	1.5				
2	Ville	40	2				
3	Ville	40	2.5				
4	Ville	50	1.5				
5	Ville	50	2				
6	Ville	50	2.5				
7	Route	40	1.5				
8	Route	40	2				
9	Route	40	2.5			33570	
10	Route	50	1.5			33270	
11	Route	50	2			33440	
12	Route	50	2.5			32840	
13	Ville	40	1.5			32750	
14	Ville	40	2			33360	
15	Ville	40	2.5			32100	
16	Ville	50	1.5			32270	
17	Ville	50	2			33100	
18	Ville	50	2.5			33700	

et voulons d'abord voir ce qu'il se passe si nous ne prenons que le *Lieu* comme variable catégorielle:

The dialog box 'Analyse de variance équilibrée' is shown with the following settings:

- Réponses : Response
- Modèle : Lieu
- Facteurs aléatoires : (empty)

Buttons at the bottom: Sélectionner, Graphiques..., Résultats..., Stockage..., Aide, OK, Annuler.

Nous validons par **OK** et nous obtenons dans la fenêtre de session:

### ANOVA : Response en fonction de Lieu

Facteur Type Niveaux Valeurs  
 Lieu fixe 2 Route; Ville

Analyse de la variance pour Response

Source	DL	Somme des carrés	CM	F	P
Lieu	1	961707	961707	2.19	0.148
Erreur	34	14906680	438432		
Total	35	15868387			

S = 662.142 R carré = 6.06 % R carré (ajust) = 3.30 %

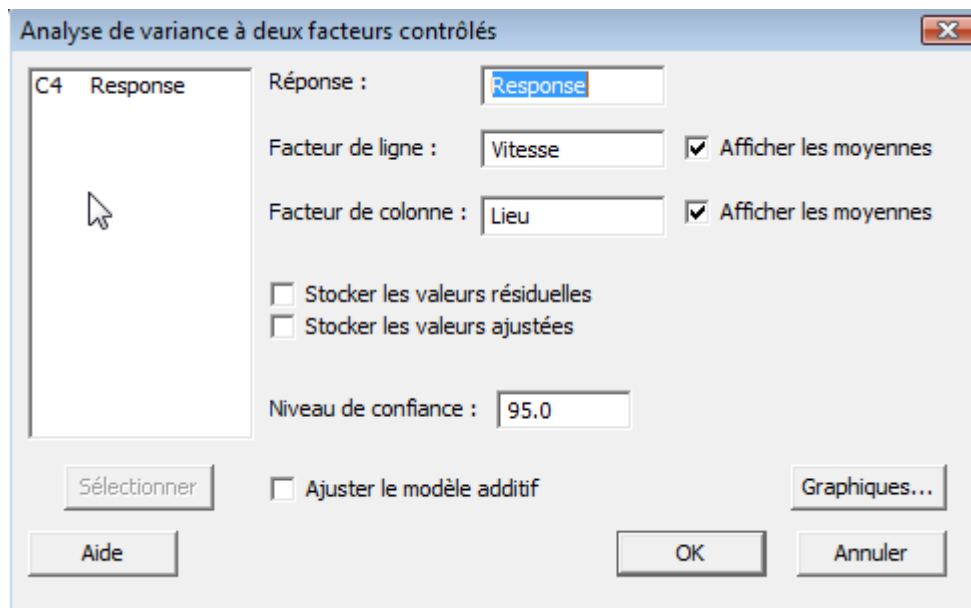
Soit exactement les mêmes valeurs que l'ANOVA à un facteur. Résultat logique puisque l'ANOVA multifactorielle est un cas général de l'ANOVA à un facteur et à deux facteurs.

Faisons maintenant le même contrôle mais avec l'ANOVA à deux facteurs. Nous souhaitons vérifier si le facteur 1 (route ou ville) représentant le *Lieu* ainsi que le facteur 2 (40 [km/h] = 1, 50 [km/h] = 2) représentant la *Vitesse* ont une influence sur la distance parcourue.

Donc toujours en partant du même tableau:

	C1-T	C2-T	C3-T	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	Lieu	Vitesse	Pression	Response									
1	Ville	40	1.5	32700									
2	Ville	40	2	33430									
3	Ville	40	2.5	31710									
4	Ville	50	1.5	32680									
5	Ville	50	2	34070									
6	Ville	50	2.5	33220									
7	Route	40	1.5	33180									
8	Route	40	2	34430									
9	Route	40	2.5	33570									
10	Route	50	1.5	33270									
11	Route	50	2	33440									
12	Route	50	2.5	32840									
13	Ville	40	1.5	32750									
14	Ville	40	2	33360									
15	Ville	40	2.5	32100									
16	Ville	50	1.5	32270									
17	Ville	50	2	33100									
18	Ville	50	2.5	33700									

Nous allons dans le menu **Stat/ANOVA/A deux facteurs...**:

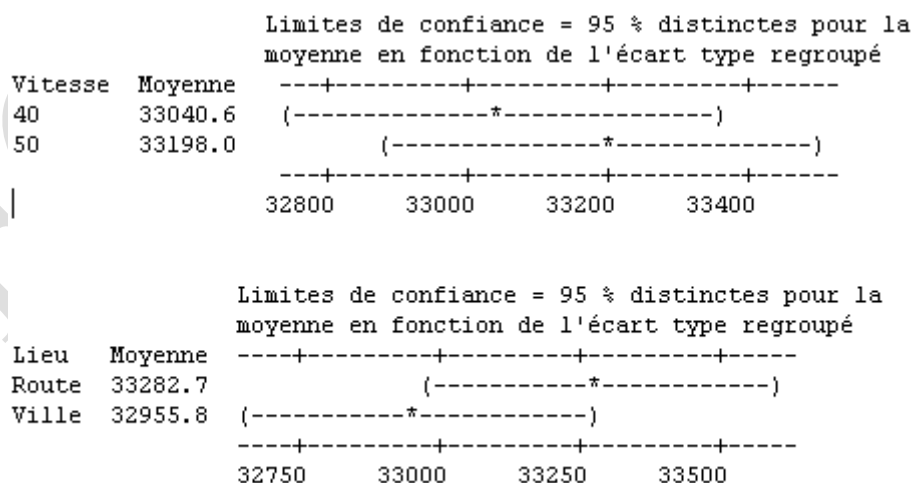


Nous validons par **OK** pour obtenir dans la fenêtre de session:

**ANOVA à deux entrées : Response en fonction de Vitesse; Lieu**

Source	DL	Somme des carrés	CM	F	P
Vitesse	1	223099	223099	0.54	0.469
Lieu	1	961707	961707	2.31	0.138
Interaction	1	1367340	1367340	3.29	0.079
Erreur	32	13316241	416133		
Total	35	15868387			

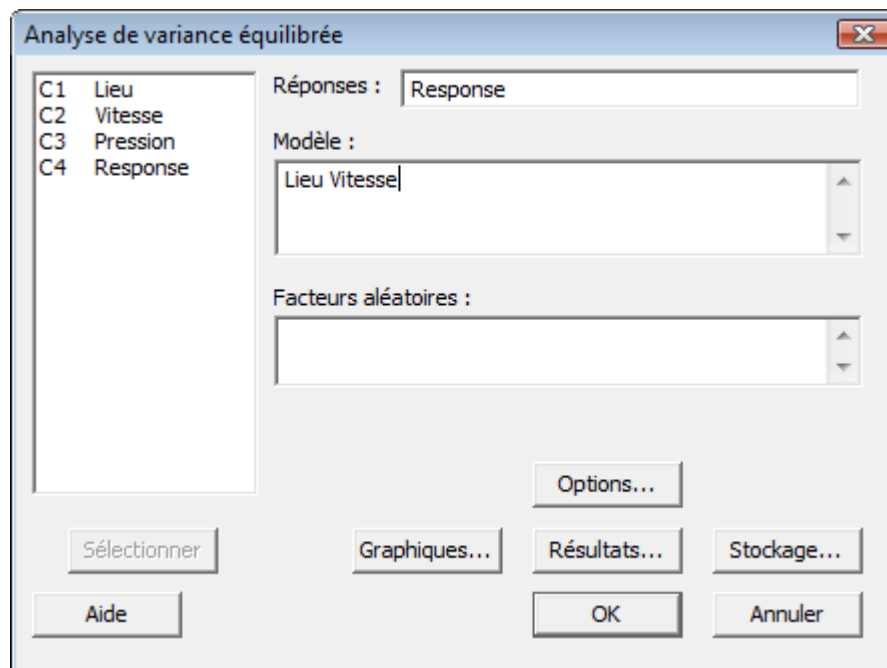
S = 645.1 R carré = 16.08 % R carré (ajust) = 8.22 %



Nous voyons donc que seule l'interaction a un effet significatif.



Vérifions maintenant que nous avons le même résultat avec l'outil d'analyse multifactorielle (bien évidemment cela devrait être le cas). Nous allons donc dans le menu **Stat/ANOVA/ANOVA équilibrée...**:



Nous validons par **OK**:

#### **ANOVA : Response en fonction de Lieu; Vitesse**

Facteur	Type	Niveaux	Valeurs
Lieu	fixe	2	Route; Ville
Vitesse	fixe	2	40; 50

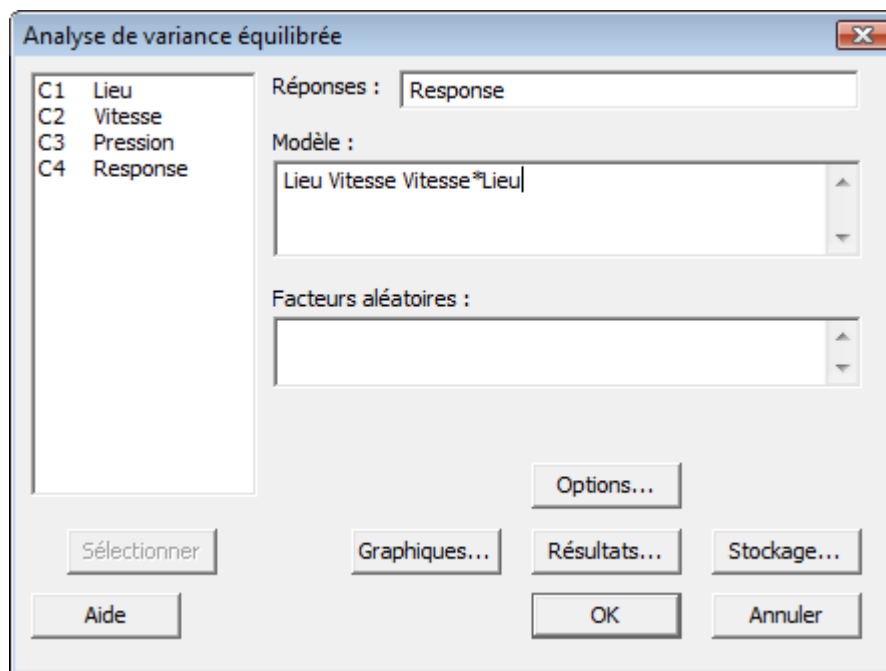
Analyse de la variance pour Response

Source	DL	Somme des carrés	CM	F	P
Lieu	1	961707	961707	2.16	0.151
Vitesse	1	223099	223099	0.50	0.484
Erreur	33	14683581	444957		
Total	35	15868387			

S = 667.051    R carré = 7.47 %    R carré (ajust) = 1.86 %

Le lecteur pourrait être surpris que nous ne trouvons pas les mêmes valeurs pour le test de Fisher et la  $p$ -value. Mais au fait c'est normal pour l'instant car si nous faisons bien attention, il manque le terme d'interaction et donc l'erreur totale n'est pas la même que si nous faisons l'ANOVA à deux facteurs (par contre la somme des carrés et le CM sont identiques à l'ANOVA à deux facteurs faite plus haut). Donc ce que nous avons fait ici c'est une ANOVA à deux facteurs en imposant qu'il n'y avait aucune interaction (ce qui est parfois très intéressant comme possibilité). Ce qui est bien évidemment un peu stupide car la vitesse dépend de lieu. Alors comment faire l'ANOVA multifactorielle pour obtenir l'ANOVA à deux

facteurs faite plus haut? Eh bien simplement en mettant les paramètres suivants dans la fenêtre de l'ANOVA équilibrée:



en validant par **OK**, nous obtenons dans la fenêtre de session:

#### ANOVA : Response en fonction de Lieu; Vitesse

Facteur	Type	Niveaux	Valeurs
Lieu	fixe	2	Route; Ville
Vitesse	fixe	2	40; 50

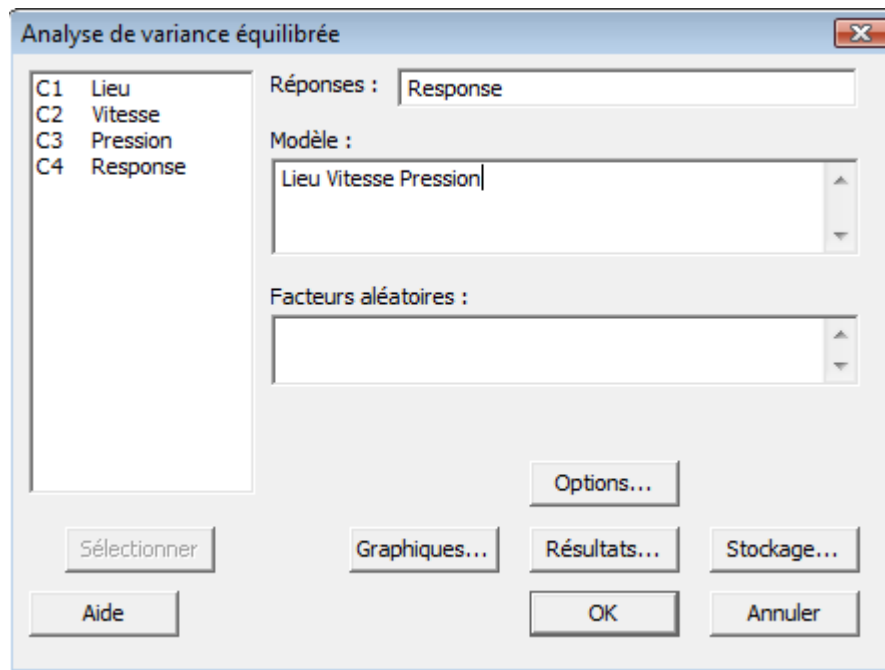
Analyse de la variance pour Response

Source	DL	Somme des carrés	CM	F	P
Lieu	1	961707	961707	2.31	0.138
Vitesse	1	223099	223099	0.54	0.469
Lieu*Vitesse	1	1367340	1367340	3.29	0.079
Erreur	32	13316241	416133		
Total	35	15868387			

S = 645.083    R carré = 16.08 %    R carré (ajust) = 8.22 %

Outre la précision de certains termes et l'ordre des informations, nous retrouvons donc bien exactement les mêmes valeurs que l'ANOVA équilibrée (multifactorielle).

Maintenant faisons ce dont quoi cet outil est vraiment fait pour. Commençons par une ANOVA à trois facteurs (fixes) sans interactions:



Nous validons par **OK** et obtenons dans la fenêtre de session:

#### **ANOVA : Response en fonction de Lieu; Vitesse; Pression**

Facteur	Type	Niveaux	Valeurs
Lieu	fixe	2	Route; Ville
Vitesse	fixe	2	40; 50
Pression	fixe	3	1.5; 2; 2.5

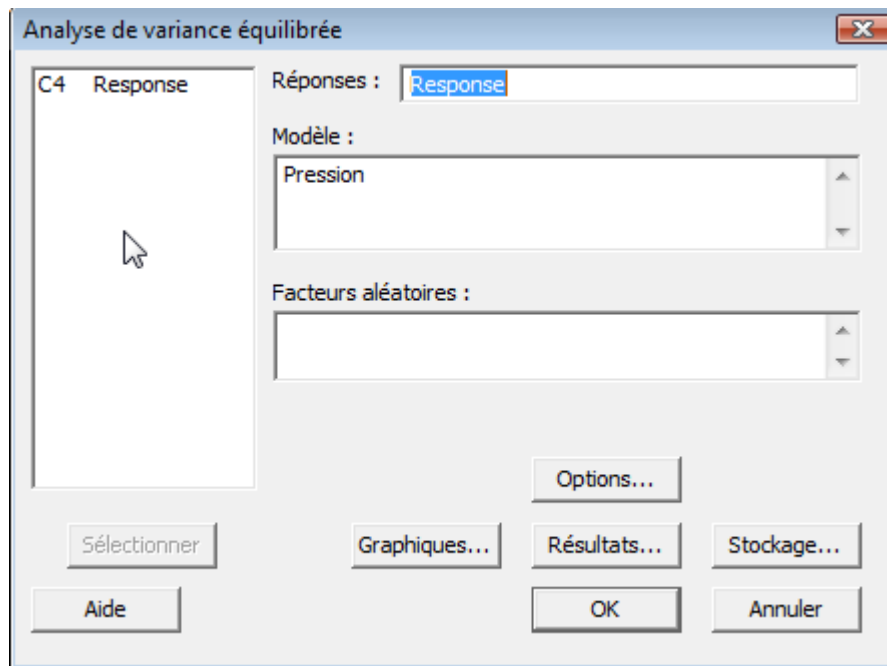
Analyse de la variance pour Response

Source	DL	Somme des carrés	CM	F	P
Lieu	1	961707	961707	3.85	0.059
Vitesse	1	223099	223099	0.89	0.352
Pression	2	6943967	3471983	13.91	0.000
Erreur	31	7739615	249665		
Total	35	15868387			

S = 499.665    R carré = 51.23 %    R carré (ajust) = 44.93 %

Donc en ne considérant aucune interaction, les facteurs Lieu et Vitesse ont toujours une influence significative mais pas la pression.

Si nous faisons la pression seule:



Nous obtenons alors dans la fenêtre de sessions:

### **ANOVA : Response en fonction de Pression**

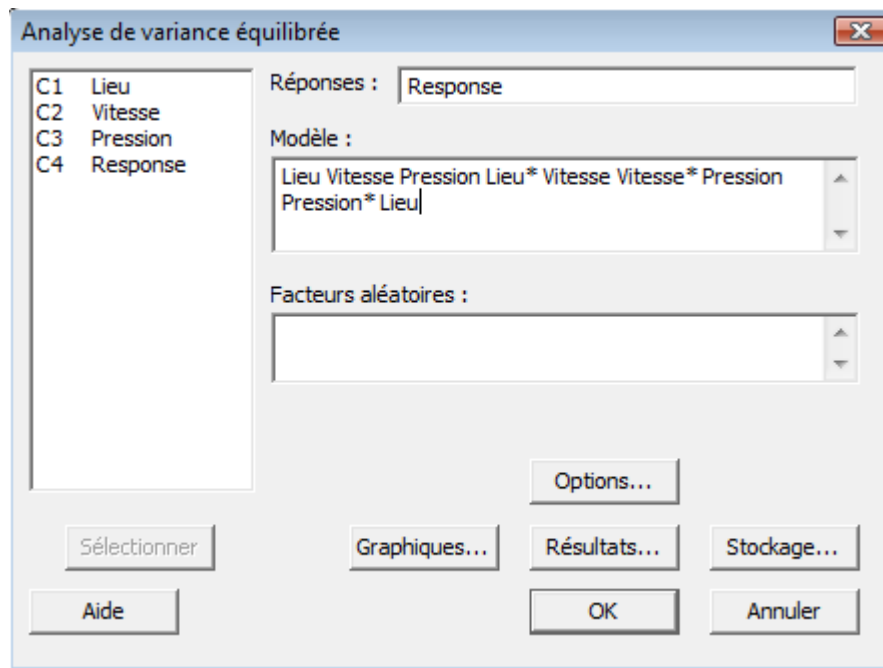
Facteur	Type	Niveaux	Valeurs
Pression	fixe	3	1.5; 2; 2.5

Analyse de la variance pour Response

Source	DL	Somme des carrés	CM	F	P
Pression	2	6943967	3471983	12.84	0.000
Erreur	33	8924421	270437		
Total	35	15868387			

S = 520.036    R carré = 43.76 %    R carré (ajust) = 40.35 %

Et à nouveau avec les trois facteurs et les toutes les interactions du deuxième ordre:



Nous obtenons alors dans la fenêtre de sessions:

### ANOVA : Response en fonction de Lieu; Vitesse; Pression

Facteur	Type	Niveaux	Valeurs
Lieu	fixe	2	Route; Ville
Vitesse	fixe	2	40; 50
Pression	fixe	3	1.5; 2; 2.5

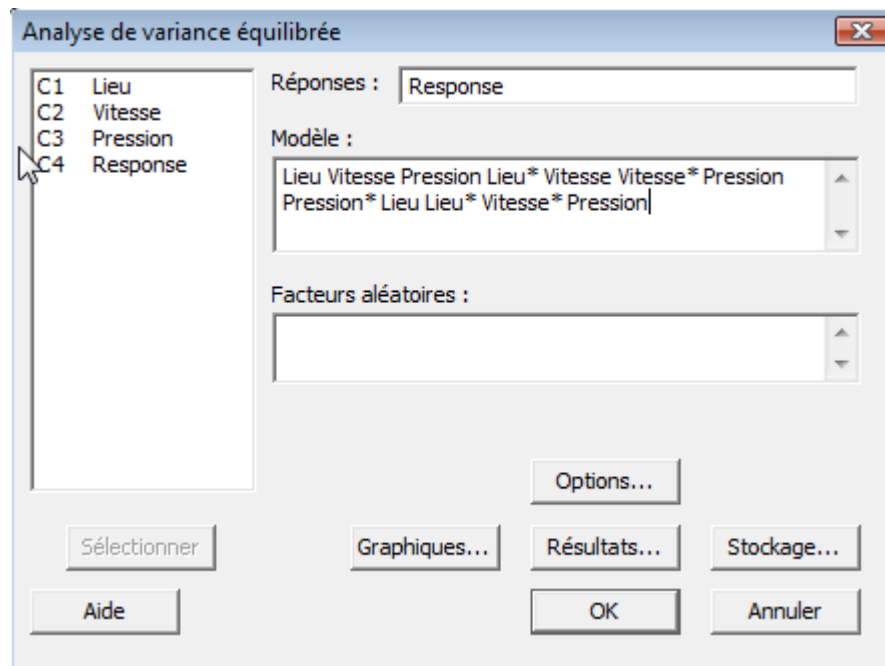
Analyse de la variance pour Response

Source	DL	Somme des carrés	CM	F	P
Lieu	1	961707	961707	4.93	0.035
Vitesse	1	223099	223099	1.14	0.295
Pression	2	6943967	3471983	17.81	0.000
Lieu*Vitesse	1	1367340	1367340	7.01	0.014
Vitesse*Pression	2	741020	370510	1.90	0.170
Lieu*Pression	2	561782	280891	1.44	0.255
Erreur	26	5069472	194980		
Total	35	15868387			

S = 441.565    R carré = 68.05 %    R carré (ajust) = 56.99 %

La pression reste donc toujours non significative. Par contre, Nous voyons que toutes les interactions ont une influence très significative.

Maintenant, faisons avec la triple interaction:



Nous obtenons alors dans la fenêtre de sessions:

### ANOVA : Response en fonction de Lieu; Vitesse; Pression

Facteur	Type	Niveaux	Valeurs
Lieu	fixe	2	Route; Ville
Vitesse	fixe	2	40; 50
Pression	fixe	3	1.5; 2; 2.5

Analyse de la variance pour Response

Source	DL	Somme des carrés	CM	F	P
Lieu	1	961707	961707	6.78	0.016
Vitesse	1	223099	223099	1.57	0.222
Pression	2	6943967	3471983	24.49	0.000
Lieu*Vitesse	1	1367340	1367340	9.65	0.005
Vitesse*Pression	2	741020	370510	2.61	0.094
Lieu*Pression	2	561782	280891	1.98	0.160
Lieu*Vitesse*Pression	2	1667622	833811	5.88	0.008
Erreur	24	3401851	141744		
Total	35	15868387			

S = 376.489    R carré = 78.56 %    R carré (ajust) = 68.74 %

et donc l'interaction des trois facteurs est aussi très significative.

## 17.4. Exercice 175.: Test de Fisher de la courbure

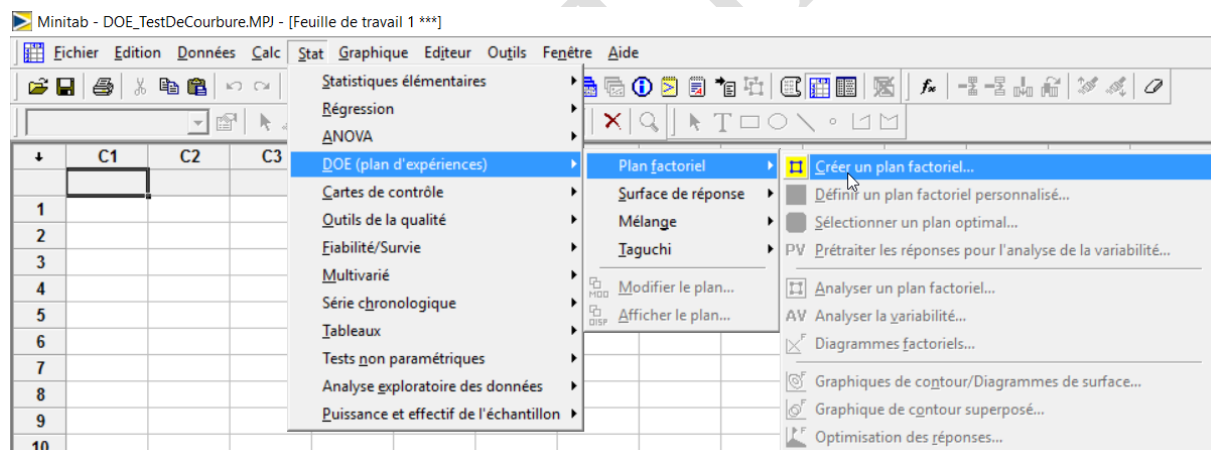
Minitab® Statistical Software 15.1.1.0

Comme nous l'avons vu dans le cours théorique, le test de courbure ne peut être effectué que si:

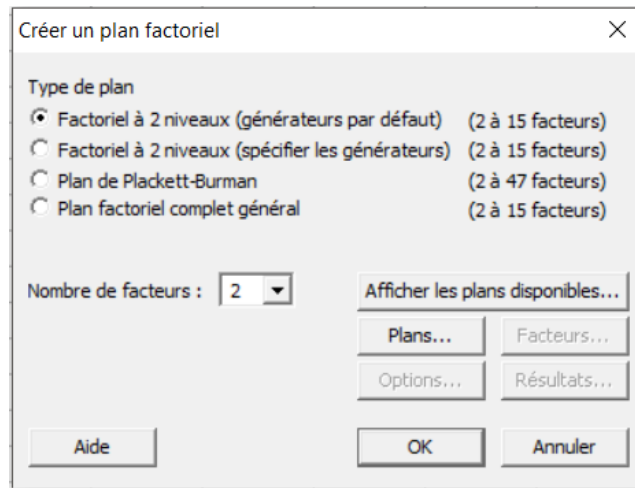
1. Les facteurs ont que 2 niveaux (donc le test ne marche pas pour les plans factoriels généralisés!!!)
2. Un des facteurs est une variable qualitative (mais les méthodes de calcul avec l'ordinateur permet de faire sauter cette limite simplement!)
3. Il faut au moins 3 répétitions

Pour pouvoir faire un exemple, nous allons reprendre l'expérience faite avant, mais avec un seul bloc (ce qui permet de vérifier les résultats à la main au besoin plus facilement pour le lecteur ou le formateur lors des cours qu'il dispense) et en éliminant la variable qualitative des lieux. Pour cela, nous procédons donc ainsi:

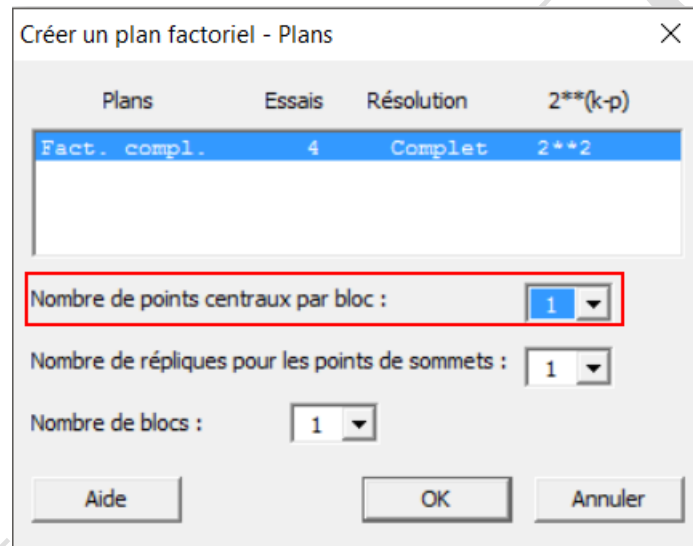
D'abord nous créons la plan d'expérience en allant dans **Stat/DOE (plan d'expérience)/Plan factoriel/Créer un plan factoriel...:**



Nous prenons ensuite l'option **Factoriel à 2 niveaux (générateurs par défaut)**:



et nous cliquons sur **Plans...** pour choisir **1 Nombre de points centraux par bloc** (toujours histoire de faciliter les calculs à la main):



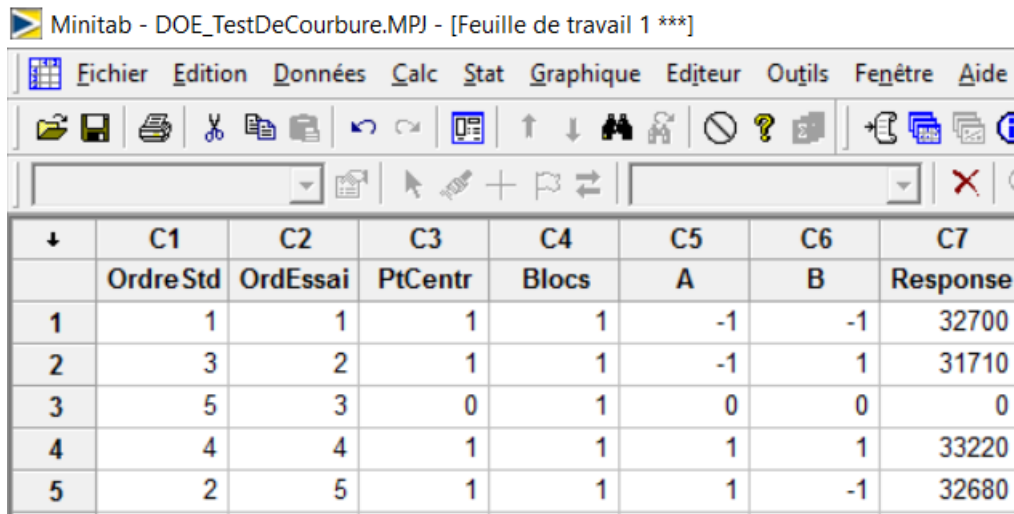
On valide par **OK** deux fois pour obtenir:

Minitab - DOE\_TestDeCourbure.MPJ - [Feuille de travail 1 \*\*\*]

	C1	C2	C3	C4	C5	C6	C7
	Ordre Std	OrdEssai	PtCentr	Blocs	A	B	
1	1	1	1	1	-1	-1	
2	3	2	1	1	-1	1	
3	5	3	0	1	0	0	
4	4	4	1	1	1	1	
5	2	5	1	1	1	-1	
6							



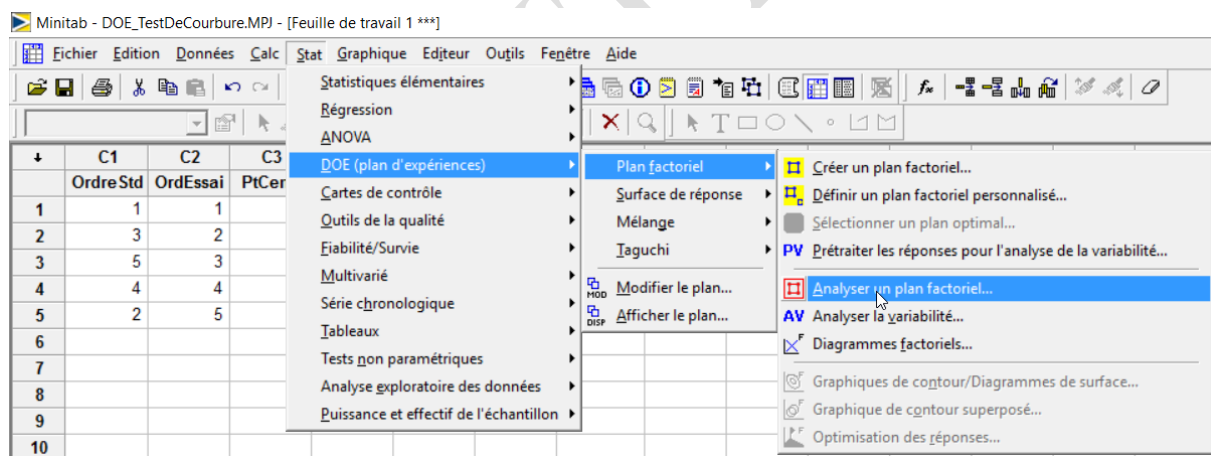
On saisit les valeurs d'un des blocs en prenant la *Ville* comme référence et pour les deux niveaux de la *Pression* uniquement les valeurs extrêmes des trois niveaux:



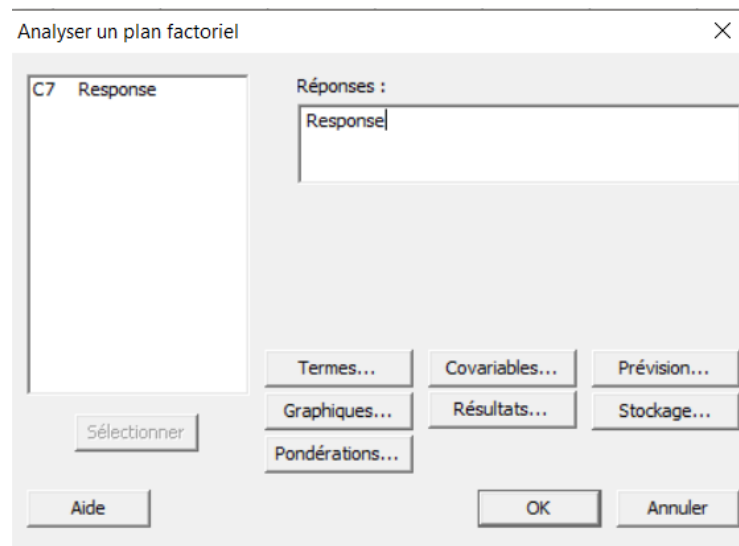
	C1	C2	C3	C4	C5	C6	C7
	Ordre Std	OrdEssai	PtCentr	Blocs	A	B	Response
1	1	1	1	1	-1	-1	32700
2	3	2	1	1	-1	1	31710
3	5	3	0	1	0	0	0
4	4	4	1	1	1	1	33220
5	2	5	1	1	1	-1	32680

Bien évidemment le point central dans notre situation ne nécessite aucune mesure car avec une vitesse nulle et une pression nulle des pneus, la distance est nulle... (merci de mettre en marche ici votre "bullshit detector" !!!).

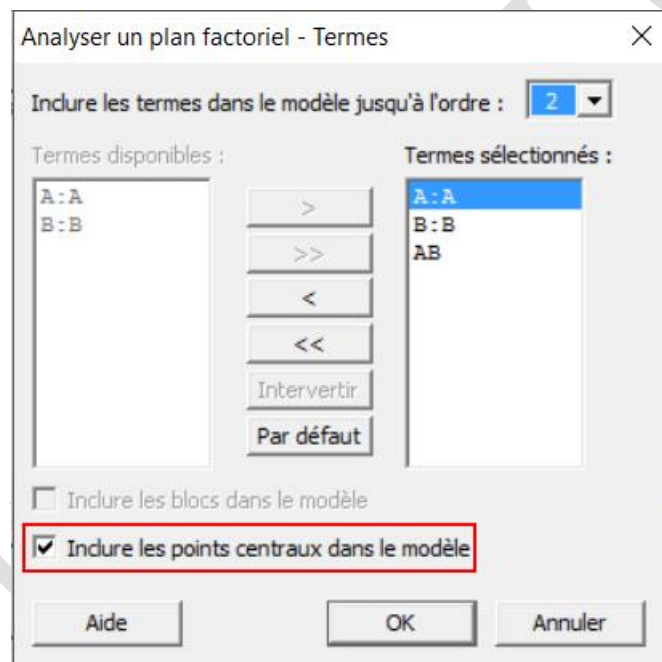
Ensuite on lance l'analyse en allant dans **Stat/DOE (plan d'expérience)/Plan factoriel/Analyser un plan factoriel...**:



Dans la boîte de dialogue qui apparaît alors, nous prenons les réponses adéquates:



Maintenant nous cliquons sur **Termes**:



pour s'assurer que **Inclure les points centraux dans le modèle** est actif. Nous cliquons deux fois sur **OK** pour obtenir:

Effets et coefficients estimés pour Response (unités codées)

Terme	Effet	Coeff
Constante		32578
A	745	373
B	-225	-113
A*B	765	382
Pt ctr		-32578

S = \* SomCar-ErrPrév = \*

Analyse de la variance pour Response (unités codées)

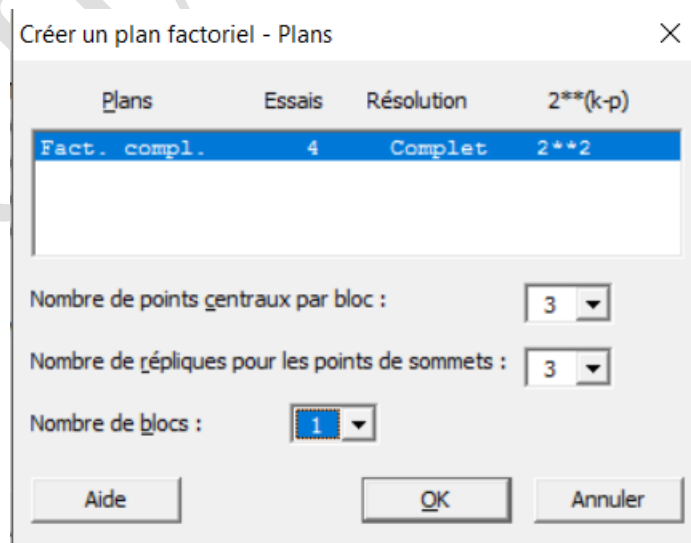
Source	DL	SomCar séq	SomCar ajust	CM ajust	F
Effets principaux	2	605650	605650	302825	*
2 Interactions - Nombre de facteurs	1	585225	585225	585225	*
Courbure	1	849034805	849034805	849034805	*
Erreur résiduelle	0	*	*	*	*
Total	4	850225680			

Source	P
Effets principaux	*
2 Interactions - Nombre de facteurs	*
Courbure	*
Erreur résiduelle	
Total	

Structure des alias

I  
A  
B  
A\*B

Donc là ce qui est intéressant c'est que Minitab ne communique pas les valeurs. C'est normal car nous avons démontré dans le cours théorique qu'il faut au moins 3 réplifications puisque un des degrés de liberté de la distribution de Fisher est  $n-2$ . Dès lors refaisons la même chose avec 3 réplicas pour tous les points:



Nous validons par **OK**. Et pour les facteurs nous mettons (histoire de faciliter la saisie):

Facteur	Nom	Type	Min	Max
A	Vitesse	Numérique	40	50
B	Pression	Numérique	1.5	2.5

et pour faciliter la saisie nous allons aussi désactiver la randomisation des essais:

Créer un plan factoriel - Options

Replier le plan d'expériences

- Ne pas replier
- Replier sur tous les facteurs
- Replier seulement sur le facteur :

Fraction

- Utiliser la fraction principale
- Utiliser la fraction numéro :

**Randomiser les essais**

Base du générateur de nombres aléatoires :

Stocker le plan dans la feuille de travail

Ce qui donne:

Minitab - DOE\_TestDeCourbure\_AvecReplication.MPJ - [Feuille de travail 1 \*\*\*]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

↓	C1	C2	C3	C4	C5	C6	C7
	Ordre Std	OrdEssai	PtCentr	Blocs	Vitesse	Pression	
1	1	1	1	1	40	1.5	
2	2	2	1	1	50	1.5	
3	3	3	1	1	40	2.5	
4	4	4	1	1	50	2.5	
5	5	5	1	1	40	1.5	
6	6	6	1	1	50	1.5	
7	7	7	1	1	40	2.5	
8	8	8	1	1	50	2.5	
9	9	9	1	1	40	1.5	
10	10	10	1	1	50	1.5	
11	11	11	1	1	40	2.5	
12	12	12	1	1	50	2.5	
13	13	13	0	1	45	2.0	
14	14	14	0	1	45	2.0	
15	15	15	0	1	45	2.0	
16							
17							

Nous rajoutons les réponses:

↓	C1	C2	C3	C4	C5	C6	C7	C8
	Ordre Std	OrdEssai	PtCentr	Blocs	Vitesse	Pression	Response	
1	1	1	1	1	40	1.5	32700	
2	2	2	1	1	50	1.5	32680	
3	3	3	1	1	40	2.5	31710	
4	4	4	1	1	50	2.5	33220	
5	5	5	1	1	40	1.5	33180	
6	6	6	1	1	50	1.5	33270	
7	7	7	1	1	40	2.5	33570	
8	8	8	1	1	50	2.5	32840	
9	9	9	1	1	40	1.5	32750	
10	10	10	1	1	50	1.5	32270	
11	11	11	1	1	40	2.5	32100	
12	12	12	1	1	50	2.5	33700	
13	13	13	0	1	45	2.0	0	
14	14	14	0	1	45	2.0	0	
15	15	15	0	1	45	2.0	0	
16								

Nous lançons l'analyse comme avant.

Effets et coefficients estimés pour Response (unités codées)

Terme	Effet	Coeff	Coef	ErT	T	P
Constante		32833		156.5	209.75	0.000
Vitesse	328	164		156.5	1.05	0.319
Pression	48	24		156.5	0.15	0.880
Vitesse*Pression	465	232		156.5	1.49	0.168
Pt ctr		-32833		350.0	-93.80	0.000

S = 542.248      SomCar-ErrPrév = 3855246827  
 R carré = 99.89 %      R carré (prév) = 0.00 %      R carré (ajust) = 99.84 %

Analyse de la variance pour Response (unités codées)

Source	DL	SomCar séq	SomCar ajust	CM ajust
Effets principaux	2	330417	330417	165208
2 Interactions - Nombre de facteurs	1	648675	648675	648675
Courbure	1	2587135335	2587135335	2587135335
Erreur résiduelle	10	2940333	2940333	294033
Erreur pure	10	2940333	2940333	294033
Total	14	2591054760		

Source	F	P
Effets principaux	0.56	0.587
2 Interactions - Nombre de facteurs	2.21	0.168
Courbure	8798.78	0.000
Erreur résiduelle		
Erreur pure		
Total		

Observations aberrantes pour Response

Observation	OrdreStd	Response	Valeur ajustée	Ajust ErT	Valeur résiduelle	Valeur résiduelle normalisée
7	7	33570.0	32460.0	313.1	1110.0	2.51R

R indique une observation ayant une valeur résiduelle normalisée importante



Coefficients estimés pour Response à l'aide des données des unités non codées

Terme	Coeff
Constante	39628.3
Vitesse	-153.167
Pression	-4136.67
Vitesse*Pression	93.0000
Pt ctr	-32832.5

Structure des alias

I

Vitesse

Pression

Vitesse\*Pression

Et donc il y a clairement une courbure et donc un plan de surface de réponse devrait être exécuté après le screening. Par contre comme nous l'avons démontré dans le cours théorique... cela ne nous dit pas quels sont les facteurs responsables de la courbure.

**Attention!!!! Il n'est pas possible de faire donc de diagramme de contour et de diagramme de surface dans Minitab s'il y a des points centraux!!!!!!**

## 17.5. Exercice 176.: Analyse du modèle d'un plan factoriel complet général

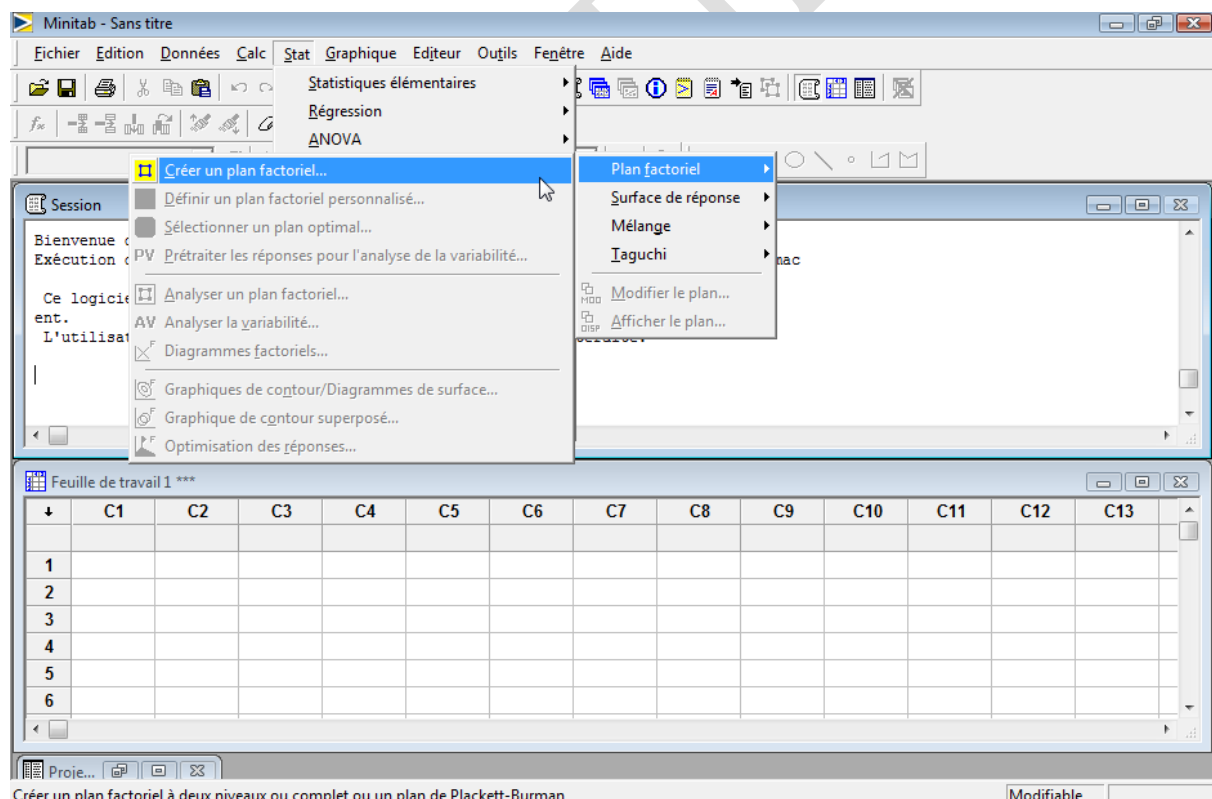
Minitab® Statistical Software 15.1.1.0

Dans le cours de statistique théorique, nous avons étudié en détails un plan d'expérience permettant de modéliser et donc d'optimiser la distance parcourue par un véhicule. Nous avons obtenus sans interactions, le modèle mathématique suivant:

$$Y = 33'119.3 + x_1 \underbrace{\begin{bmatrix} -163.4 \\ 163.4 \end{bmatrix}}_{\Sigma=0} + x_2 \underbrace{\begin{bmatrix} -78.7 \\ 78.7 \end{bmatrix}}_{\Sigma=0} + x_3 \underbrace{\begin{bmatrix} -349.7 \\ 619.4 \\ -269.7 \end{bmatrix}}_{\Sigma=0}$$

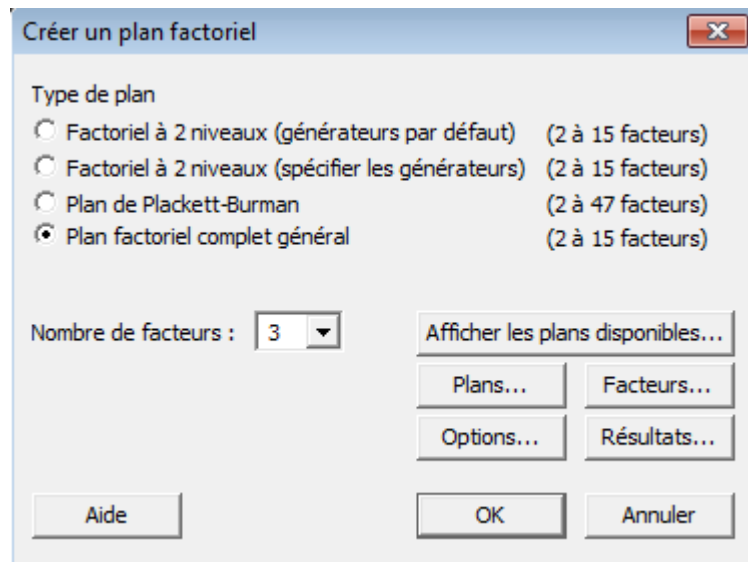
L'objectif va être de vérifier si nous retrouvons le même modèle avec Minitab® Statistical Software. Pour cela, nous allons d'abord créer le plan factoriel comme dans l'exercice précédent.

Nous allons donc dans le menu **Stat/DOE (plan d'expériences)/Plan factoriel/Créer un plan factoriel...**:

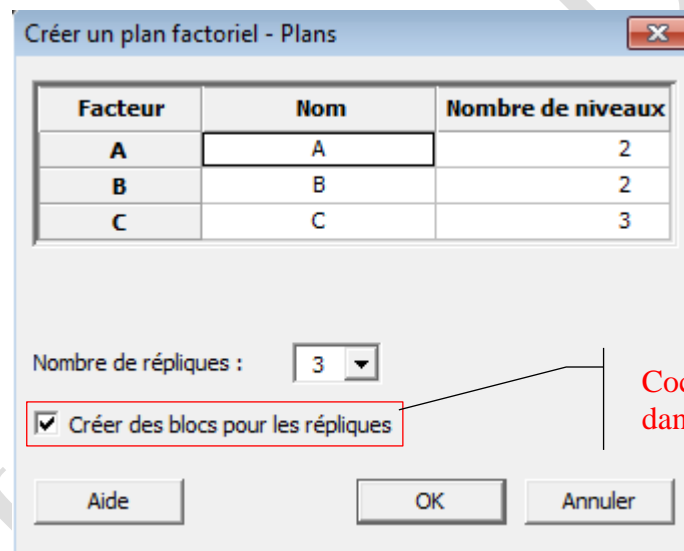


Nous prenons un **Plan factoriel complet général** à 3 niveaux:





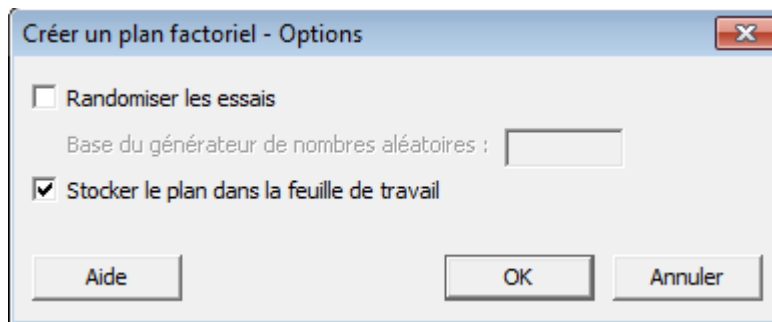
Nous cliquons sur le bouton **Plans...**:



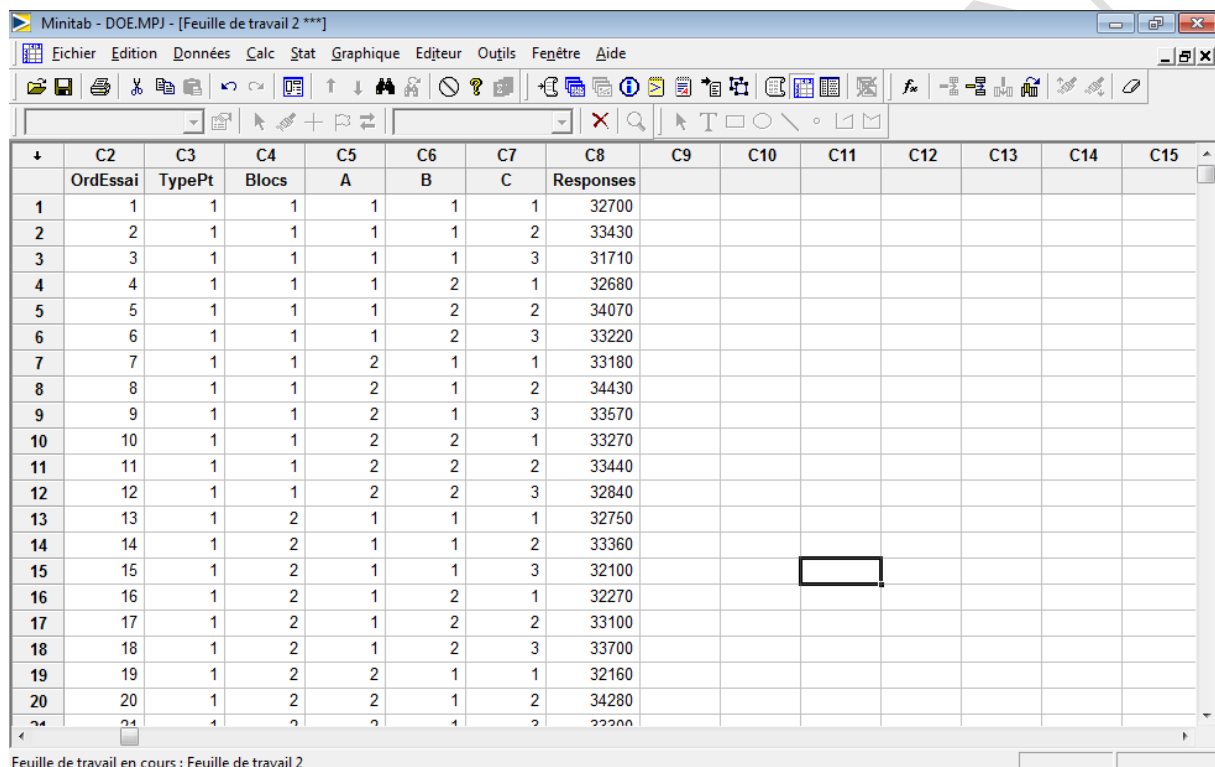
sans oublier de spécifier nous avons 3 répliques comme nous l'avons fait dans le cours de statistique théorique.

**Remarque: Il n'y pas le choix de faire des points centraux pour ce type de plan, même si tous les facteurs sont à 2 niveaux. Pour avoir des points centraux, il faudra passer par l'option Factoriel à 2 niveaux et prendre ensuite un plan complet... O\_o**

Ensuite, nous allons dans le bouton **Options...**:



Nous validons par **OK** et nous reportons toutes les mesures utilisées dans le cours théorique (ou ouvrir le fichier *DOE\_FactorielComplet\_3Facteurs*):

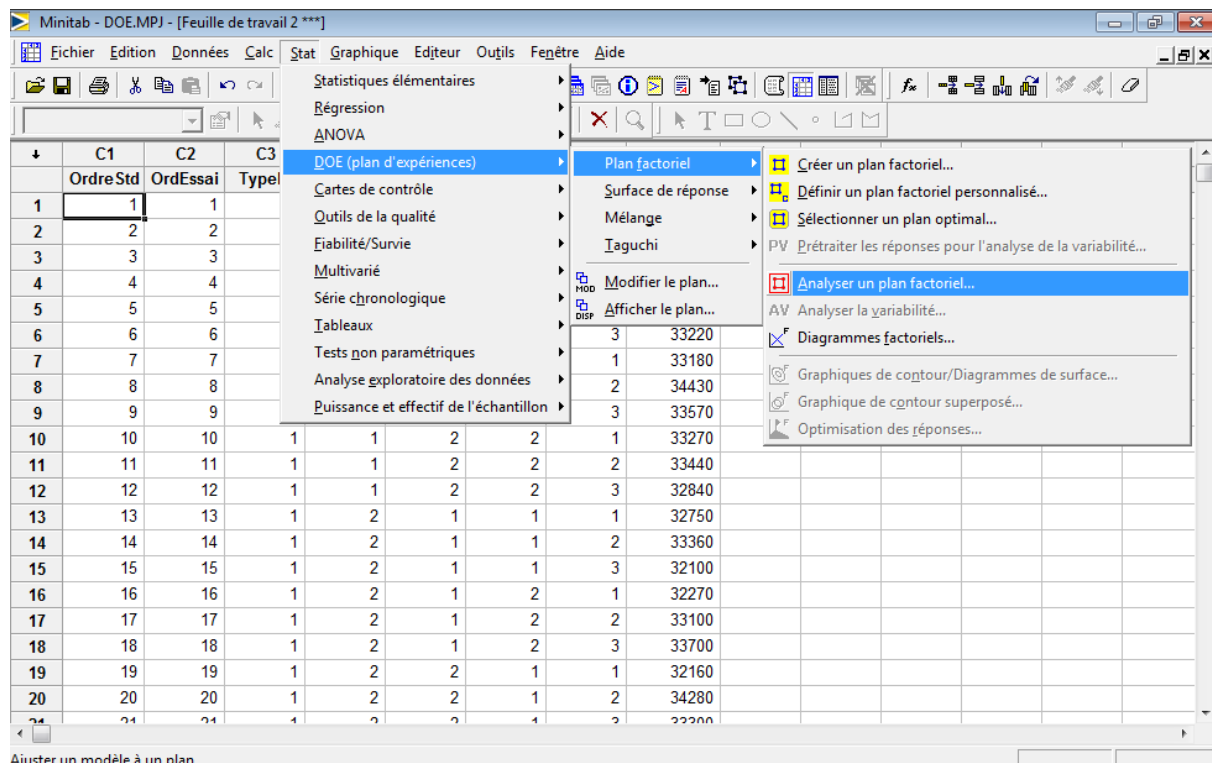


ce qui correspond bien au tableau du cours théorique qui était pour rappel:

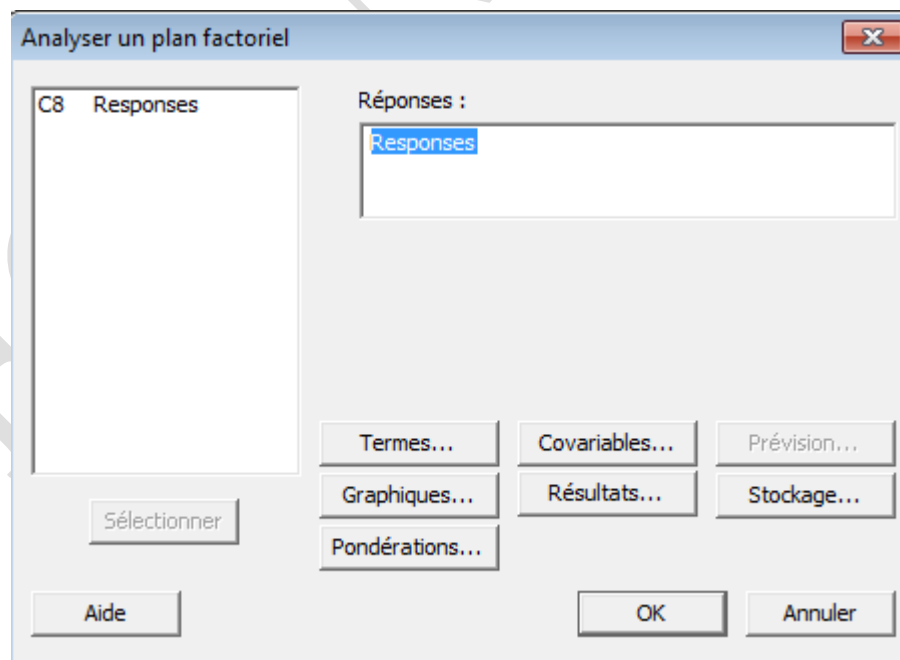
N° Essai	$x_1$	$x_2$	$x_3$	Valeurs	de	y
1	1	1	1	32'700	32'750	32'960
2	1	1	2	33'430	33'360	32'910
3	1	1	3	31'710	32'100	32'220
4	1	2	1	32'680	32'270	33'130
5	1	2	2	34'070	33'100	33'610
6	1	2	3	33'220	33'700	33'285
7	2	1	1	33'180	32'160	32'640
8	2	1	2	34'430	34'280	34'460
9	2	1	3	33'570	33'300	32'570
10	2	2	1	33'270	33'080	32'415
11	2	2	2	33'440	33'570	34'204
12	2	2	3	32'840	33'210	32'470

### 17.5.1. Analyse sans blocs

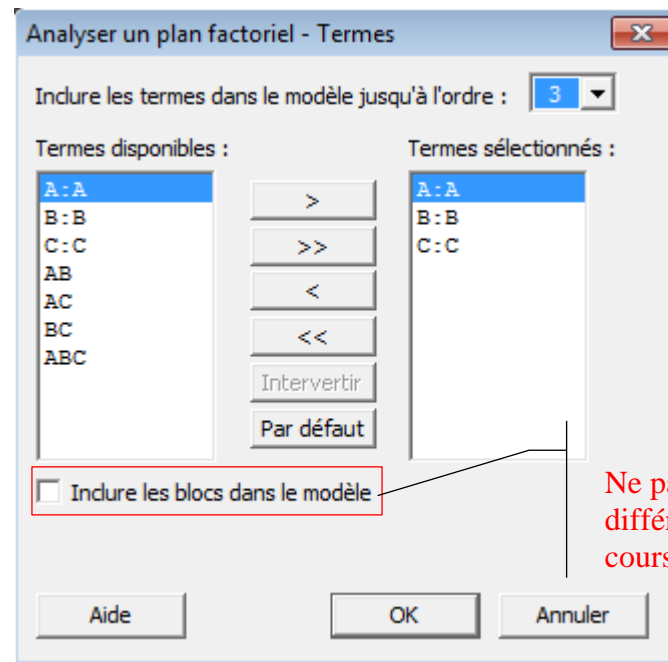
Ensuite, nous lançons l'analyse du plan en allant dans le menu **Stat/DOE (plan d'expériences)/Plan factoriel/Analyser un plan factoriel**:



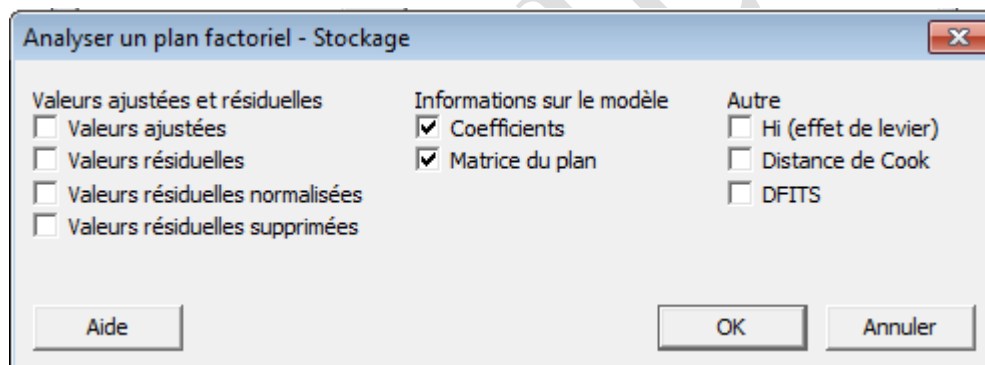
Nous mettons les réponses:



Nous cliquons sur le bouton **Termes...** pour y prendre les facteurs simples sans les interactions (dans un premier temps):

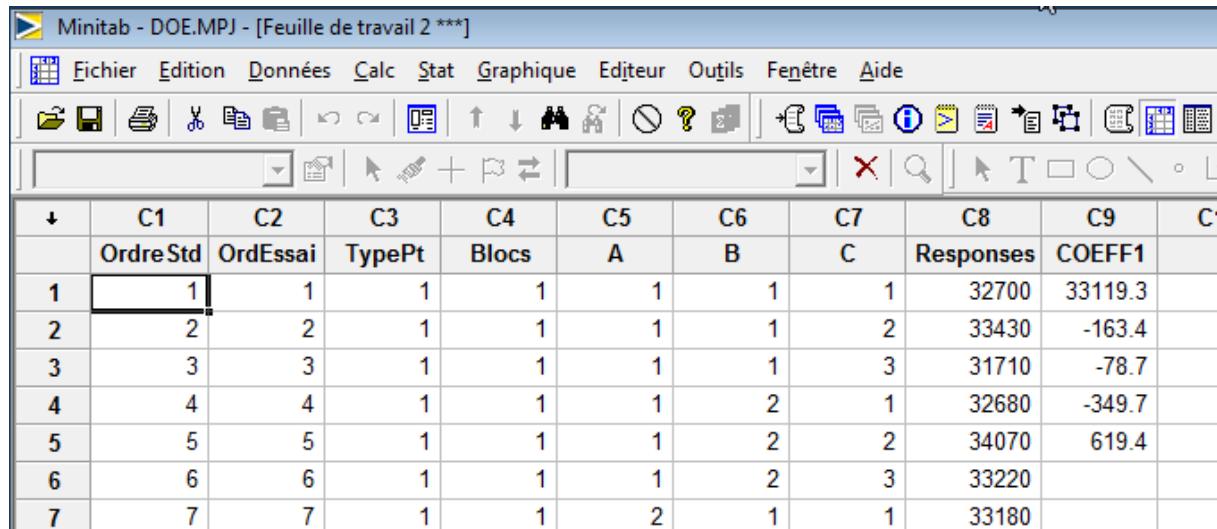


Ensuite, nous cliquons sur le bouton **Stockage** pour dire que nous voulons les **Coefficients** et la **Matrice du plan**:



**Attention!!!** Avec un plan factoriel complet général nous n'avons pas l'option d'obtenir les **Effets** (voir plus loin l'analyse de Pareto d'un plan factoriel fractionnaire).

Nous validons par **OK** et nous nous retrouvons avec une nouvelle colonne **COEFF1**:



	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
	OrdreStd	OrdEssai	TypePt	Blocs	A	B	C	Responses	COEFF1	
1	1	1	1	1	1	1	1	32700	33119.3	
2	2	2	1	1	1	1	2	33430	-163.4	
3	3	3	1	1	1	1	3	31710	-78.7	
4	4	4	1	1	1	2	1	32680	-349.7	
5	5	5	1	1	1	2	2	34070	619.4	
6	6	6	1	1	1	2	3	33220		
7	7	7	1	1	2	1	1	33180		

qui contient bien le modèle théorique sans interactions:

$$Y = 33'119.3 + x_1 \begin{bmatrix} -163.4 \\ 163.4 \end{bmatrix} + x_2 \begin{bmatrix} -78.7 \\ 78.7 \end{bmatrix} + x_3 \begin{bmatrix} -349.7 \\ 619.4 \\ -269.7 \end{bmatrix}$$

$\underbrace{\hspace{1.5cm}}_{\Sigma=0}$ 
 $\underbrace{\hspace{1.5cm}}_{\Sigma=0}$ 
 $\underbrace{\hspace{1.5cm}}_{\Sigma=0}$

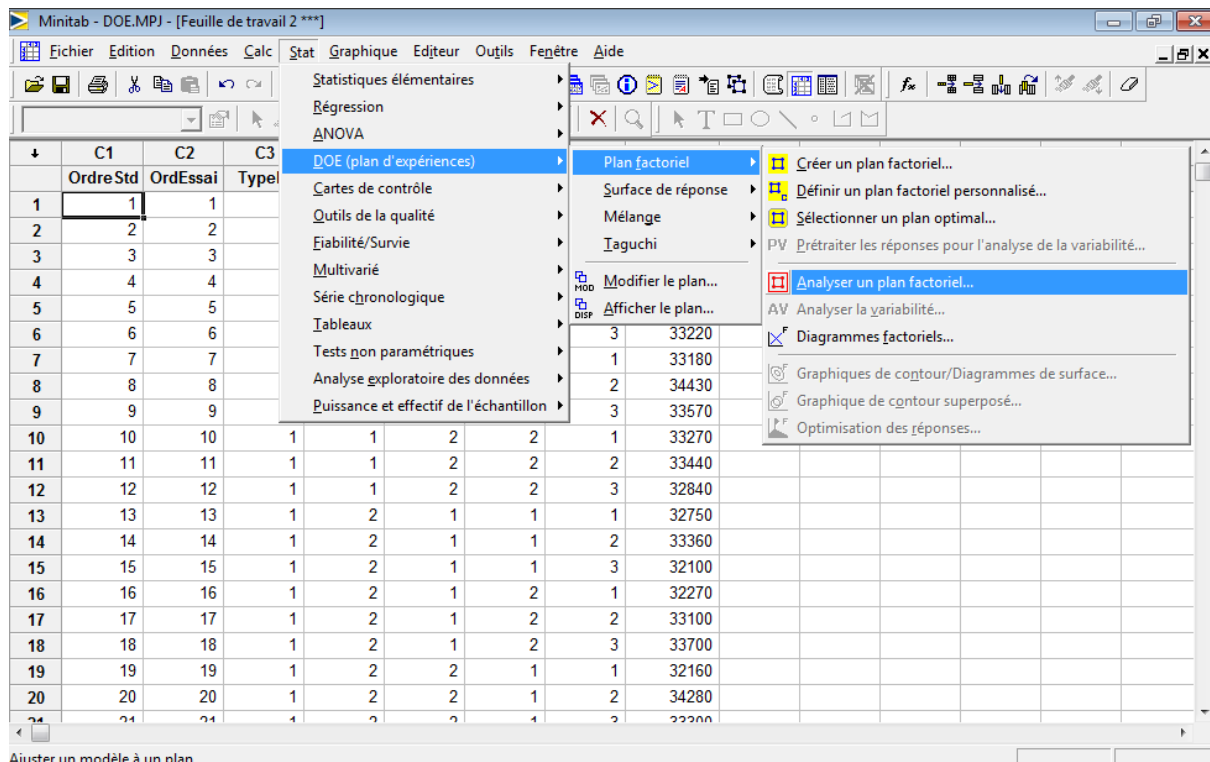
mais sous forme condensée, comme nous l'avions mentionné dans le cours théorique.

Ensuite, dans le cours théorique, nous avons calculé à la main le modèle avec des interactions du deuxième ordre. Nous avons obtenus:

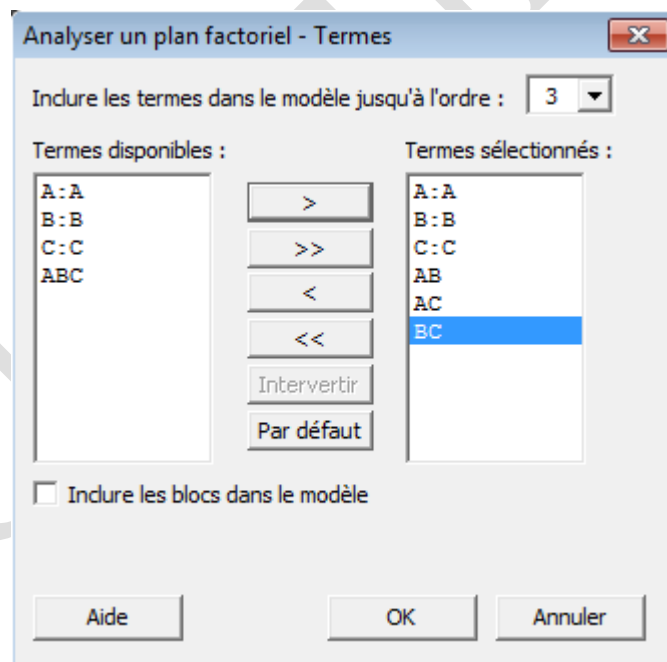
$$y = 33119.3 + A \begin{bmatrix} -163.4 \\ 163.4 \end{bmatrix} + B \begin{bmatrix} -78.7 \\ 78.7 \end{bmatrix} + C \begin{bmatrix} -349.7 \\ 619.4 \\ -269.7 \end{bmatrix} + AB \begin{bmatrix} -194.9 \\ 194.9 \\ 194.9 \\ -194.9 \end{bmatrix} + AC \begin{bmatrix} 142.2 \\ -161.9 \\ 19.7 \\ -142.2 \\ 161.9 \\ -19.7 \end{bmatrix} + BC \begin{bmatrix} 40.8 \\ 151.7 \\ -192.5 \\ -40.8 \\ -151.7 \\ 192.5 \end{bmatrix}$$

$\underbrace{\hspace{1.5cm}}_{\Sigma=0}$ 
 $\underbrace{\hspace{1.5cm}}_{\Sigma=0}$ 
 $\underbrace{\hspace{1.5cm}}_{\Sigma=0}$ 
 $\underbrace{\hspace{1.5cm}}_{\Sigma=0}$ 
 $\underbrace{\hspace{1.5cm}}_{\Sigma=0}$ 
 $\underbrace{\hspace{1.5cm}}_{\Sigma=0}$

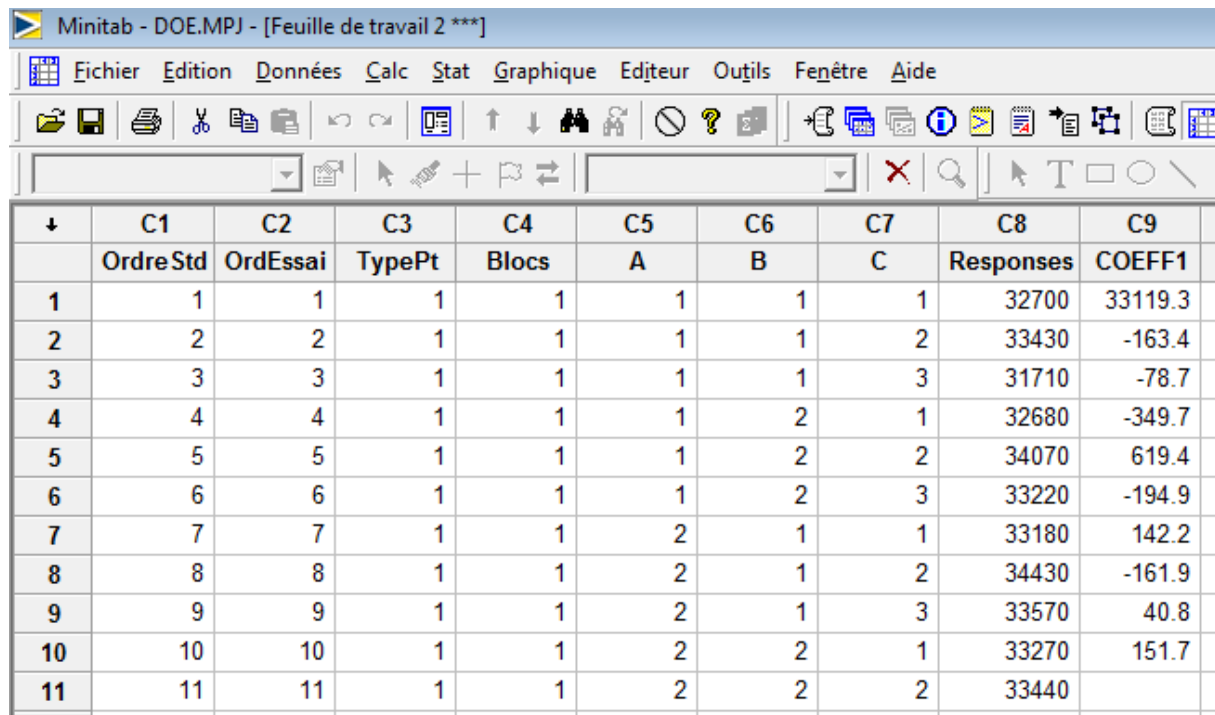
Voyons si Minitab® Statistical Software nous donnera cela sous la même forme (bon nous savons déjà que non mais vérifions quand même). Donc encore une fois, nous allons dans le même menu:



Mais cette fois-ci, nous prenons les interactions jusqu'au deuxième ordre:



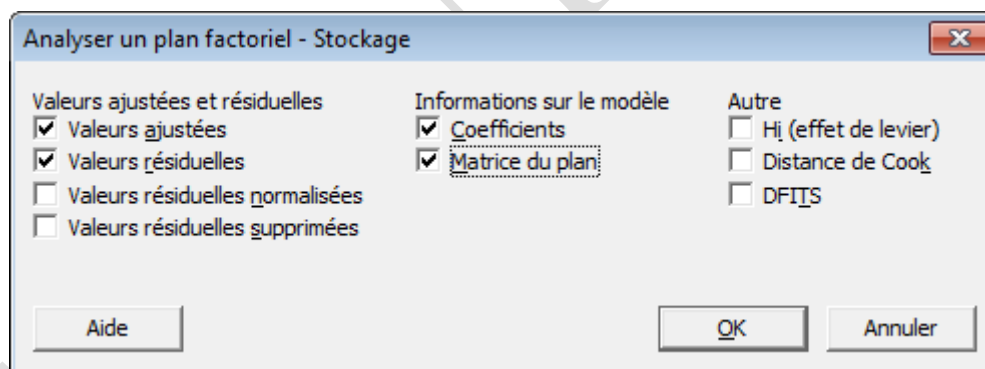
Nous validons par **OK** pour obtenir (nous avons effacé la colonne des résultats précédents):



	C1	C2	C3	C4	C5	C6	C7	C8	C9
	Ordre Std	OrdEssai	TypePt	Blocs	A	B	C	Responses	COEFF1
1	1	1	1	1	1	1	1	32700	33119.3
2	2	2	1	1	1	1	2	33430	-163.4
3	3	3	1	1	1	1	3	31710	-78.7
4	4	4	1	1	1	2	1	32680	-349.7
5	5	5	1	1	1	2	2	34070	619.4
6	6	6	1	1	1	2	3	33220	-194.9
7	7	7	1	1	2	1	1	33180	142.2
8	8	8	1	1	2	1	2	34430	-161.9
9	9	9	1	1	2	1	3	33570	40.8
10	10	10	1	1	2	2	1	33270	151.7
11	11	11	1	1	2	2	2	33440	

Nous retrouvons bien les coefficients sous une forme très condensée comme nous l'avions mentionné dans le cours théorique.

Si maintenant, nous recommençons la procédure mais en activant cette fois-ci les **Valeurs ajustées** et les **Valeurs résiduelles**:



Nous nous retrouvons avec:

↓	C1-T	C2	C3	C4	C5	C6	C7	C8	C9	C10
	Lieu	Vitesse	Pression	Response	Ordre Std	OrdEssai	Blocs	TypePt	AJUSTEES1	COEFF1
1	Ville	40	1.5	32700	1	1	1	1	32803.3	33119.3
2	Ville	40	2.0	33430	2	2	1	1	33233.3	163.4
3	Ville	40	2.5	31710	3	3	1	1	32010.0	-163.4
4	Ville	50	1.5	32680	4	4	1	1	32693.3	-78.7
5	Ville	50	2.0	34070	5	5	1	1	33593.3	78.7
6	Ville	50	2.5	33220	6	6	1	1	33401.7	-349.7
7	Route	40	1.5	33180	7	7	1	1	32660.0	619.4
8	Route	40	2.0	34430	8	8	1	1	34390.0	-269.7
9	Route	40	2.5	33570	9	9	1	1	33146.7	194.9
10	Route	50	1.5	33270	10	10	1	1	32921.7	-194.9
11	Route	50	2.0	33440	11	11	1	1	33738.0	-194.9
12	Route	50	2.5	32840	12	12	1	1	32840.0	194.9
13	Ville	40	1.5	32750	13	13	1	1	32803.3	-142.2
14	Ville	40	2.0	33360	14	14	1	1	33233.3	161.9
15	Ville	40	2.5	32100	15	15	1	1	32010.0	-19.7
16	Ville	50	1.5	32270	16	16	1	1	32693.3	142.2
17	Ville	50	2.0	33100	17	17	1	1	33593.3	-161.9
18	Ville	50	2.5	33700	18	18	1	1	33401.7	19.7
19	Route	40	1.5	32160	19	19	1	1	32660.0	40.8
20	Route	40	2.0	34280	20	20	1	1	34390.0	151.7
21	Route	40	2.5	33300	21	21	1	1	33146.7	-192.5
22	Route	50	1.5	33080	22	22	1	1	32921.7	-40.8
23	Route	50	2.0	33570	23	23	1	1	33738.0	-151.7
24	Route	50	2.5	33210	24	24	1	1	32840.0	192.5
25	Ville	40	1.5	32960	25	25	1	1	32803.3	-287.8
26	Ville	40	2.0	32910	26	26	1	1	33233.3	58.1
27	Ville	40	2.5	32220	27	27	1	1	32010.0	229.7
28	Ville	50	1.5	33130	28	28	1	1	32693.3	287.8
29	Ville	50	2.0	33610	29	29	1	1	33593.3	-58.1
30	Ville	50	2.5	33285	30	30	1	1	33401.7	-229.7
31	Route	40	1.5	32640	31	31	1	1	32660.0	287.8
32	Route	40	2.0	34460	32	32	1	1	34390.0	-58.1
33	Route	40	2.5	32570	33	33	1	1	33146.7	-229.7
34	Route	50	1.5	32415	34	34	1	1	32921.7	-287.8
35	Route	50	2.0	34204	35	35	1	1	33738.0	58.1
36	Route	50	2.5	32470	36	36	1	1	32840.0	229.7

Nous nous retrouvons donc avec les mêmes valeurs ajustées (valeurs du modèle théorique) et les mêmes valeurs résiduelles (qui sont simplement la différence entre la colonne *Réponse* et la colonne *AJUSTEES1*).

Et nous avons aussi:



**Régression factorielle générale : Response en fonction de Lieu; Vitesse; ...**

## Informations sur les facteurs

Facteur	Niveaux	Valeurs
Lieu	2	Route; Ville
Vitesse	2	40; 50
Pression	3	1.5; 2.0; 2.5

## Analyse de variance

Source	DL	SomCar ajust	CM ajust	Valeur F	Valeur de p
Modèle	11	12466537	1133322	8.00	0.000
Linéaire	4	8128773	2032193	14.34	0.000
Lieu	1	961707	961707	6.78	0.016
Vitesse	1	223099	223099	1.57	0.222
Pression	2	6943967	3471983	24.49	0.000
Interactions à 2 facteur(s)	5	2670142	534028	3.77	0.012
Lieu*Vitesse	1	1367340	1367340	9.65	0.005
Lieu*Pression	2	561782	280891	1.98	0.160
Vitesse*Pression	2	741020	370510	2.61	0.094
Interactions à 3 facteur(s)	2	1667622	833811	5.88	0.008
Lieu*Vitesse*Pression	2	1667622	833811	5.88	0.008
Erreur	24	3401851	141744		
Total	35	15868387			

## Récapitulatif du modèle

S	R carré	R carré (ajust)	R carré (prév)
376.489	78.56%	68.74%	51.76%

## Coefficients

Terme	Coeff	Coef ErT	Valeur de T	Valeur de p	FIV
Constante	33119.3	62.7	527.81	0.000	
Lieu					
Route	163.4	62.7	2.60	0.016	1.00
Vitesse					
40	-78.7	62.7	-1.25	0.222	1.00
Pression					
1.5	-349.7	88.7	-3.94	0.001	1.33
2.0	619.4	88.7	6.98	0.000	1.33
Lieu*Vitesse					
Route 40	194.9	62.7	3.11	0.005	1.00
Lieu*Pression					
Route 1.5	-142.2	88.7	-1.60	0.122	1.33
Route 2.0	161.9	88.7	1.82	0.081	1.33
Vitesse*Pression					
40 1.5	40.8	88.7	0.46	0.650	1.33
40 2.0	151.7	88.7	1.71	0.100	1.33
Lieu*Vitesse*Pression					

Route 40 1.5	-287.8	88.7	-3.24	0.003	1.33
Route 40 2.0	58.1	88.7	0.65	0.519	1.33

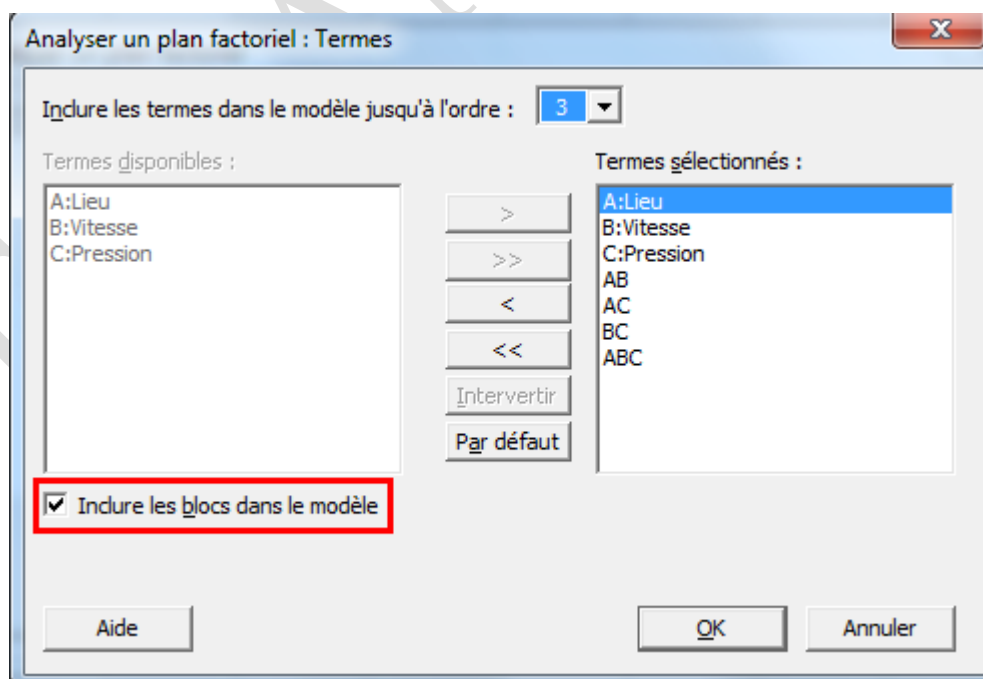
Equation de régression

```
Response = 33119.3 + 163.4 Lieu_Route - 163.4 Lieu_Ville - 78.7 Vitesse_40
+ 78.7 Vitesse_50 - 349.7 Pression_1.5 + 619.4 Pression_2.0
- 269.7 Pression_2.5 + 194.9 Lieu*Vitesse_Route 40
- 194.9 Lieu*Vitesse_Route 50 - 194.9 Lieu*Vitesse_Ville 40
+ 194.9 Lieu*Vitesse_Ville 50 - 142.2 Lieu*Pression_Route 1.5
+ 161.9 Lieu*Pression_Route 2.0 - 19.7 Lieu*Pression_Route 2.5
+ 142.2 Lieu*Pression_Ville 1.5 - 161.9 Lieu*Pression_Ville 2.0
+ 19.7 Lieu*Pression_Ville 2.5 + 40.8 Vitesse*Pression_40 1.5
+ 151.7 Vitesse*Pression_40 2.0 - 192.5 Vitesse*Pression_40 2.5
- 40.8 Vitesse*Pression_50 1.5 - 151.7 Vitesse*Pression_50 2.0
+ 192.5 Vitesse*Pression_50 2.5 - 287.8 Lieu*Vitesse*Pression_Route
40 1.5 + 58.1 Lieu*Vitesse*Pression_Route 40 2.0
+ 229.7 Lieu*Vitesse*Pression_Route 40 2.5
+ 287.8 Lieu*Vitesse*Pression_Route 50 1.5
- 58.1 Lieu*Vitesse*Pression_Route 50 2.0
- 229.7 Lieu*Vitesse*Pression_Route 50 2.5
+ 287.8 Lieu*Vitesse*Pression_Ville 40 1.5
- 58.1 Lieu*Vitesse*Pression_Ville 40 2.0
- 229.7 Lieu*Vitesse*Pression_Ville 40 2.5
- 287.8 Lieu*Vitesse*Pression_Ville 50 1.5
+ 58.1 Lieu*Vitesse*Pression_Ville 50 2.0
+ 229.7 Lieu*Vitesse*Pression_Ville 50 2.5
```

Notez que l'ANOVA est une ANOVA à facteurs fixes!!!!

### 17.5.2. Analyse avec blocs

Si nous reproduisons les étapes ci-dessus mais en incluant cette fois-ci les blocs:



Nous obtenons:

## Régression factorielle générale : Response en fonction de Blocs; Lieu; ...

### Informations sur les facteurs

Facteur	Niveaux	Valeurs
Lieu	2	Ville; Route
Vitesse	2	40; 50
Pression	3	1.5; 2.0; 2.5

### Analyse de variance

Source	DL	SomCar ajust	CM ajust	Valeur F	Valeur de p
Modèle	13	12620181	970783	6.58	0.000
<b>Blocs</b>	<b>2</b>	<b>153644</b>	<b>76822</b>	<b>0.52</b>	<b>0.601</b>
Linéaire	4	8128773	2032193	13.76	0.000
Lieu	1	961707	961707	6.51	0.018
Vitesse	1	223099	223099	1.51	0.232
Pression	2	6943967	3471983	23.52	0.000
Interactions à 2 facteur(s)	5	2670142	534028	3.62	0.015
Lieu*Vitesse	1	1367340	1367340	9.26	0.006
Lieu*Pression	2	561782	280891	1.90	0.173
Vitesse*Pression	2	741020	370510	2.51	0.104
Interactions à 3 facteur(s)	2	1667622	833811	5.65	0.010
Lieu*Vitesse*Pression	2	1667622	833811	5.65	0.010
Erreur	22	3248206	147646		
Total	35	15868387			

### Récapitulatif du modèle

S	R carré	R carré (ajust)	R carré (prév)
384.247	79.53%	67.43%	45.19%

Les blocs sont alors "simplement" considérés comme des facteurs dans le sens où nous n'avons pas pu nous remettre dans les mêmes conditions initiales lors des mesures dans des blocs différents (ce qui peut arriver!).

Cependant même si ils ont des coefficients in extenso, ils sont éliminés du modèle par Minitab ce qui est là aussi relativement logique:

## Coefficients

Terme	Coeff	Coef ErT	Valeur de T	Valeur de p	FIV
Constante	33119.3	64.0	517.16	0.000	
Blocs					
1	92.4	90.6	1.02	0.319	1.33
2	-45.9	90.6	-0.51	0.617	1.33
Lieu					
Ville	-163.4	64.0	-2.55	0.018	1.00
Vitesse					
40	-78.7	64.0	-1.23	0.232	1.00
Pression					
1.5	-349.7	90.6	-3.86	0.001	1.33
2.0	619.4	90.6	6.84	0.000	1.33
Lieu*Vitesse					
Ville 40	-194.9	64.0	-3.04	0.006	1.00
Lieu*Pression					
Ville 1.5	142.2	90.6	1.57	0.131	1.33
Ville 2.0	-161.9	90.6	-1.79	0.088	1.33
Vitesse*Pression					
40 1.5	40.8	90.6	0.45	0.657	1.33
40 2.0	151.7	90.6	1.68	0.108	1.33
Lieu*Vitesse*Pression					
Ville 40 1.5	287.8	90.6	3.18	0.004	1.33
Ville 40 2.0	-58.1	90.6	-0.64	0.528	1.33

## Equation de régression

```

Response = 33119.3 - 163.4 Lieu_Ville + 163.4 Lieu_Route - 78.7 Vitesse_40
+ 78.7 Vitesse_50 - 349.7 Pression_1.5 + 619.4 Pression_2.0
- 269.7 Pression_2.5 - 194.9 Lieu*Vitesse_Ville 40
+ 194.9 Lieu*Vitesse_Ville 50 + 194.9 Lieu*Vitesse_Route 40
- 194.9 Lieu*Vitesse_Route 50 + 142.2 Lieu*Pression_Ville 1.5
- 161.9 Lieu*Pression_Ville 2.0 + 19.7 Lieu*Pression_Ville 2.5
- 142.2 Lieu*Pression_Route 1.5 + 161.9 Lieu*Pression_Route 2.0
- 19.7 Lieu*Pression_Route 2.5 + 40.8 Vitesse*Pression_40 1.5
+ 151.7 Vitesse*Pression_40 2.0 - 192.5 Vitesse*Pression_40 2.5
- 40.8 Vitesse*Pression_50 1.5 - 151.7 Vitesse*Pression_50 2.0
+ 192.5 Vitesse*Pression_50 2.5 + 287.8 Lieu*Vitesse*Pression_Ville
40 1.5 - 58.1 Lieu*Vitesse*Pression_Ville 40 2.0
- 229.7 Lieu*Vitesse*Pression_Ville 40 2.5
- 287.8 Lieu*Vitesse*Pression_Ville 50 1.5
+ 58.1 Lieu*Vitesse*Pression_Ville 50 2.0
+ 229.7 Lieu*Vitesse*Pression_Ville 50 2.5
- 287.8 Lieu*Vitesse*Pression_Route 40 1.5
+ 58.1 Lieu*Vitesse*Pression_Route 40 2.0
+ 229.7 Lieu*Vitesse*Pression_Route 40 2.5
+ 287.8 Lieu*Vitesse*Pression_Route 50 1.5
- 58.1 Lieu*Vitesse*Pression_Route 50 2.0
- 229.7 Lieu*Vitesse*Pression_Route 50 2.5

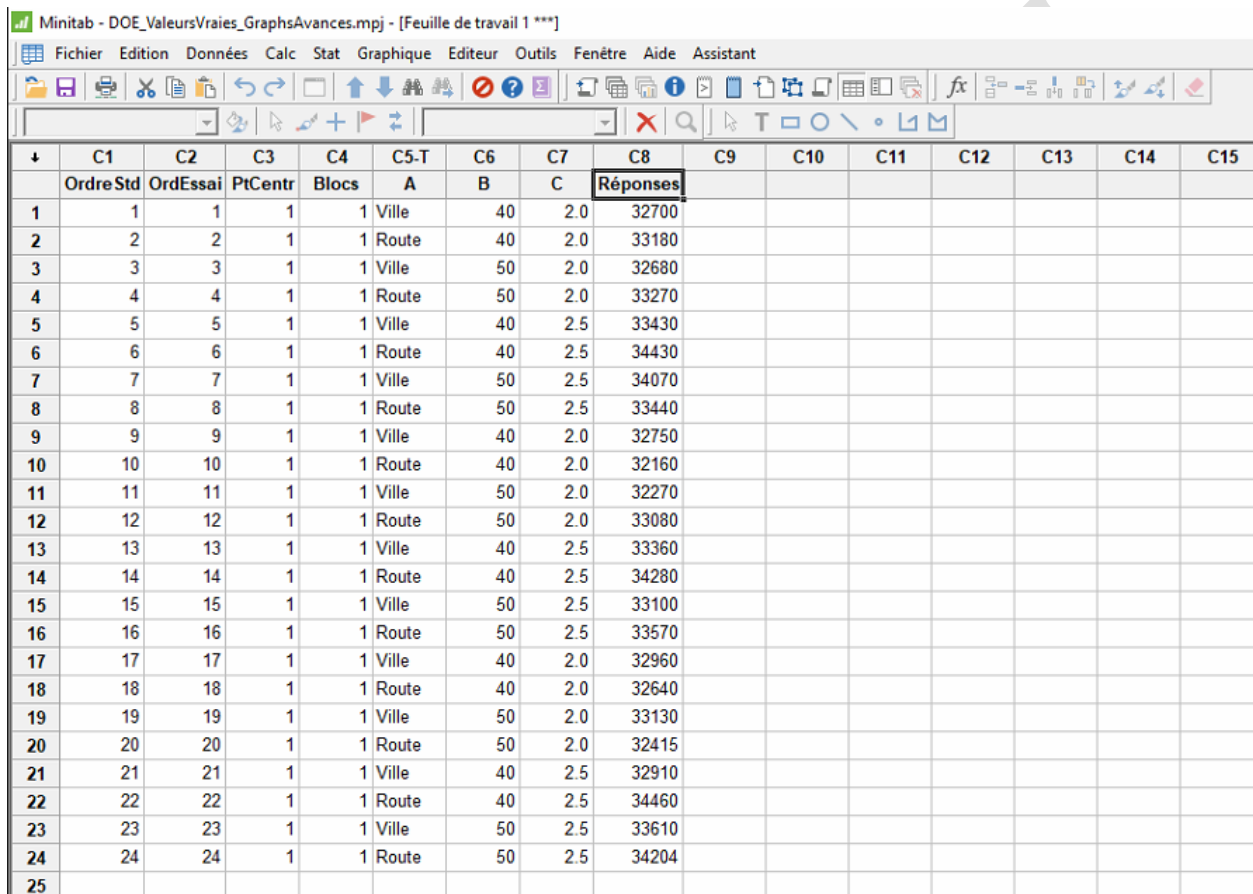
```

### 17.5.3. Analyse avec données normalisées

Lorsque les données mesurées à chaque sommet du plan d'expérience ne suivent clairement pas une loi Normale (en admettant qu'une analyse dans ce sens a été préalablement menée), il est possible de demander à Minitab d'effectuer une transformation de normalité.

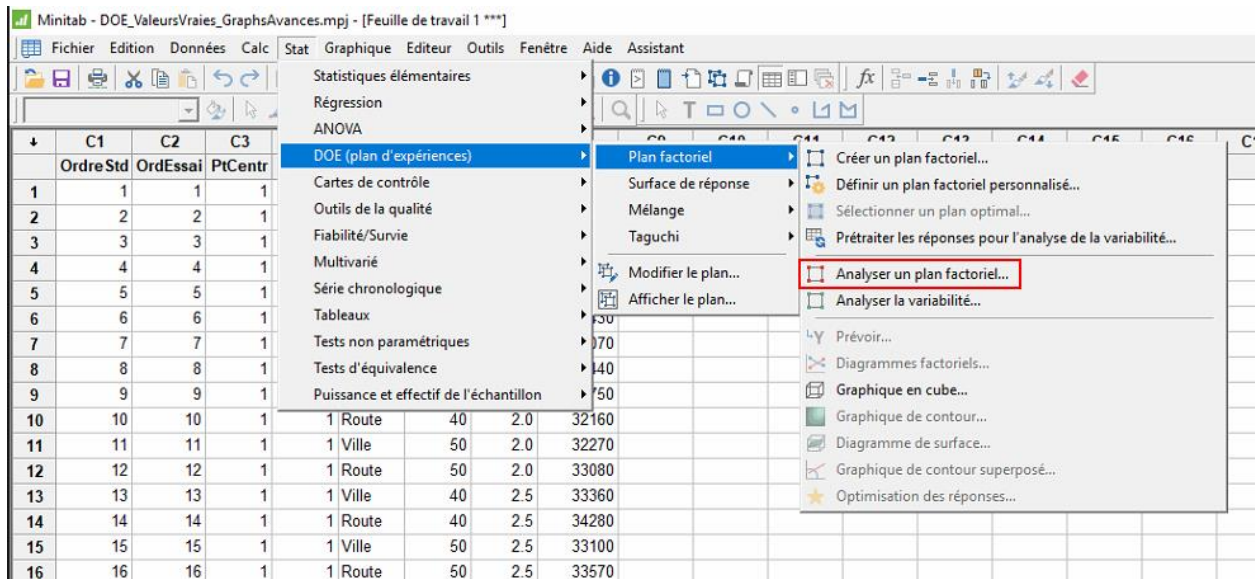
Gardez en tête que cela n'a évidemment un sens que pour les plans d'expériences avec réplication!

Pour l'exemple considérons le fichier DOE\_ValeursVraies\_GraphsAvances.mpj:

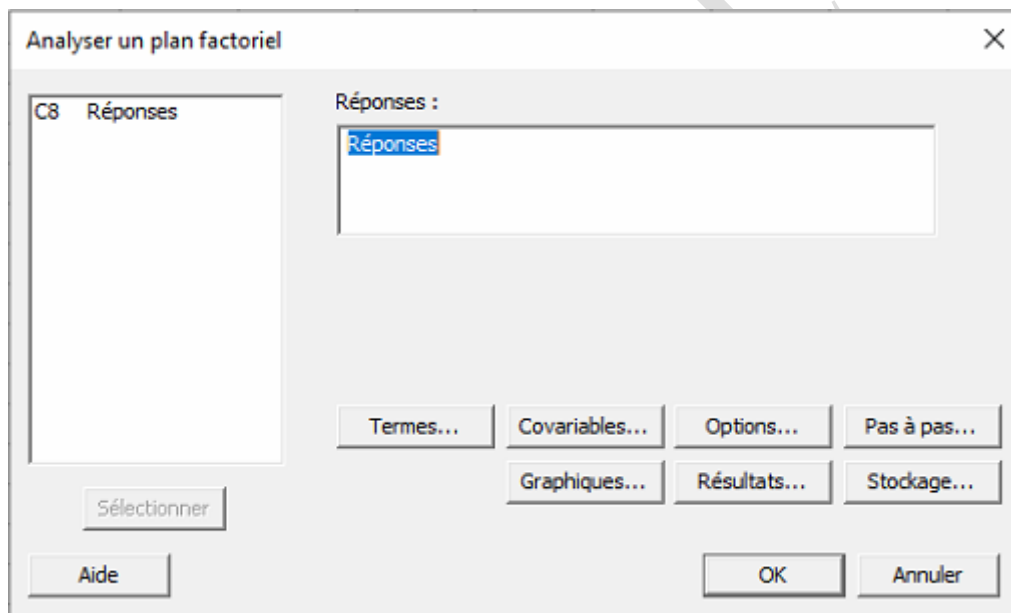


	C1	C2	C3	C4	C5-T	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	C	Réponses							
1	1	1	1	1	Ville	40	2.0	32700							
2	2	2	1	1	Route	40	2.0	33180							
3	3	3	1	1	Ville	50	2.0	32680							
4	4	4	1	1	Route	50	2.0	33270							
5	5	5	1	1	Ville	40	2.5	33430							
6	6	6	1	1	Route	40	2.5	34430							
7	7	7	1	1	Ville	50	2.5	34070							
8	8	8	1	1	Route	50	2.5	33440							
9	9	9	1	1	Ville	40	2.0	32750							
10	10	10	1	1	Route	40	2.0	32160							
11	11	11	1	1	Ville	50	2.0	32270							
12	12	12	1	1	Route	50	2.0	33080							
13	13	13	1	1	Ville	40	2.5	33360							
14	14	14	1	1	Route	40	2.5	34280							
15	15	15	1	1	Ville	50	2.5	33100							
16	16	16	1	1	Route	50	2.5	33570							
17	17	17	1	1	Ville	40	2.0	32960							
18	18	18	1	1	Route	40	2.0	32640							
19	19	19	1	1	Ville	50	2.0	33130							
20	20	20	1	1	Route	50	2.0	32415							
21	21	21	1	1	Ville	40	2.5	32910							
22	22	22	1	1	Route	40	2.5	34460							
23	23	23	1	1	Ville	50	2.5	33610							
24	24	24	1	1	Route	50	2.5	34204							
25															

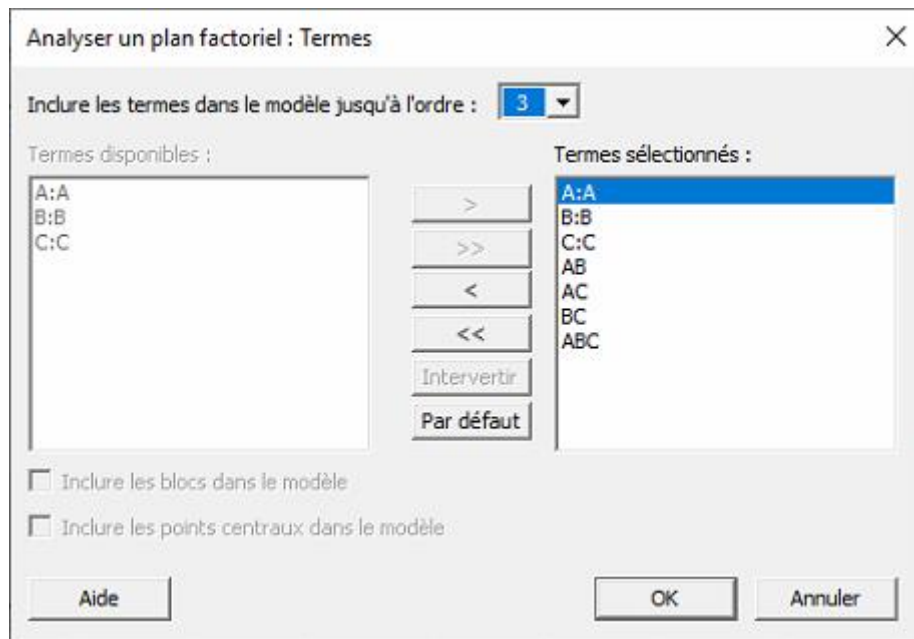
Lançons une analyse:



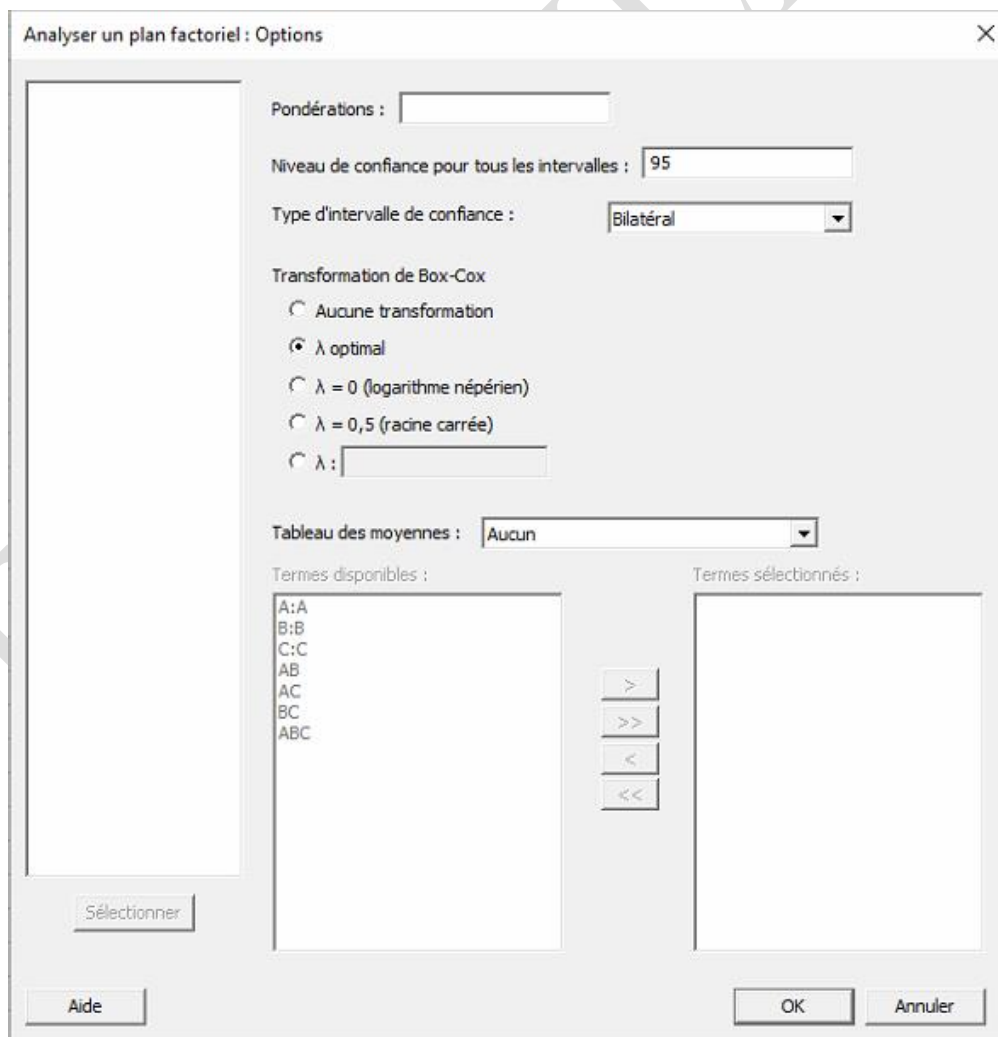
Dans la réponse nous mettons:



Et dans les **Termes...**:



Mais allons dans le bouton **Options...** (à partir de la version 16 de Minitab!) pour effectuer une transformation de Box-Cox (malheureusement il n'y pas de transformation de Johnson):



Malheureusement il ne semble pas être possible à ce jour de stocker dans une feuille Minitab les valeurs résultantes transformées.

Si nous validons le tout, nous obtenons (observez en particulier les rectangles rouges que j'ai rajouté):

### Régression factorielle : Réponses en fonction de A; B; C

Méthode

Transformation de Box-Cox

Valeur  $\lambda$  arrondie 10  
 Valeur  $\lambda$  estimée 9.97365  
 IC à 95 % pour  $\lambda$  (\*; 28.0251)

Analyse de la variance pour la réponse transformée

Source	DL	SomCar ajust	CM ajust	Valeur F	Valeur de p
Modèle	7	8939717	1277102	9.18	0.000
Linéaire	3	6959305	2319768	16.68	0.000
A	1	959218	959218	6.90	0.018
B	1	25151	25151	0.18	0.676
C	1	5974936	5974936	42.97	0.000
Interactions à 2 facteur(s)	3	1138422	379474	2.73	0.078
A*B	1	289265	289265	2.08	0.168
A*C	1	736761	736761	5.30	0.035
B*C	1	112396	112396	0.81	0.382
Interactions à 3 facteur(s)	1	841989	841989	6.06	0.026
A*B*C	1	841989	841989	6.06	0.026
Erreur	16	2224712	139045		
Total	23	11164429			

Récapitulatif du modèle pour la réponse transformée

S	R carré	R carré (ajust)	R carré (prév)
372.887	80.07%	71.36%	55.16%

Coefficients codés pour la réponse transformée

Terme	Effet	Coeff	Coef ErT	Valeur de T	Valeur de p	FIV
Constante		3389.9	76.1	44.54	0.000	
A	399.8	199.9	76.1	2.63	0.018	1.00
B	-64.7	-32.4	76.1	-0.43	0.676	1.00
C	997.9	499.0	76.1	6.56	0.000	1.00
A*B	-219.6	-109.8	76.1	-1.44	0.168	1.00
A*C	350.4	175.2	76.1	2.30	0.035	1.00
B*C	-136.9	-68.4	76.1	-0.90	0.382	1.00
A*B*C	-374.6	-187.3	76.1	-2.46	0.026	1.00

Equation de régression en unités non codées

$$\frac{(\text{Réponses}^{\lambda-1})}{(\lambda * g^{\lambda-1})} = -6352 - 15561 A + 117 B + 4459 C + 315 A*B + 7444 A*C - 54.7 B*C - 149.8 A*B*C$$

$$\lambda = 10; g = 33247.8 \text{ (correspond à la moyenne géométrique de Réponses)}$$

Et si nous lançons une optimisation des réponses, nous obtenons:



**Optimisation des réponses : Réponses**

Paramètres

Réponse	Objectif	Inférieur	Cible	Supérieur	Pondération	Importance
Réponses	Maximum	32160	34460		1	1

Solution

Solution	A	B	C	Réponses Valeur ajustée	Désirabilité composite
1	Route	40	2.5	34390.8	0.969917

Prévisions de réponses multiples

Variable	Configuration
A	Route
B	40
C	2.5

Réponse	Valeur ajustée	IC à 95 %	IP à 95 %
Réponses	34390.8	(34038.3; 34713.5)	(33649.5; 35011.3)

Sans la transformation normale, nous obtenons pour la régression factorielle:

**Régression factorielle : Réponses en fonction de A; B; C**

Analyse de variance

Source	DL	SomCar ajust	CM ajust	Valeur F	Valeur de p
Modèle	7	7863165	1123309	7.76	0.000
Linéaire	3	6362840	2120947	14.66	0.000
A	1	720720	720720	4.98	0.040
B	1	7385	7385	0.05	0.824
C	1	5634735	5634735	38.94	0.000
Interactions à 2 facteur(s)	3	782375	260792	1.80	0.187
A*B	1	153760	153760	1.06	0.318
A*C	1	554800	554800	3.83	0.068
B*C	1	73815	73815	0.51	0.485
Interactions à 3 facteur(s)	1	717950	717950	4.96	0.041
A*B*C	1	717950	717950	4.96	0.041
Erreur	16	2314967	144685		
Total	23	10178133			

Récapitulatif du modèle

S	R carré	R carré (ajust)	R carré (prév)
380.375	77.26%	67.30%	48.82%

Coefficients codés

Terme	Effet	Coeff	Coef ErT	Valeur de T	Valeur de p	FIV
Constante		33254.1	77.6	428.29	0.000	
A	346.6	173.3	77.6	2.23	0.040	1.00
B	-35.1	-17.5	77.6	-0.23	0.824	1.00
C	969.1	484.5	77.6	6.24	0.000	1.00
A*B	-160.1	-80.0	77.6	-1.03	0.318	1.00
A*C	304.1	152.0	77.6	1.96	0.068	1.00
B*C	-110.9	-55.5	77.6	-0.71	0.485	1.00
A*B*C	-345.9	-173.0	77.6	-2.23	0.041	1.00

Equation de régression en unités non codées

$$\text{Réponses} = 24559 - 14484 A + 96 B + 3935 C + 295 A*B + 6835 A*C - 44.4 B*C - 138.4 A*B*C$$

Remarque que les coefficient de régression sont significativement différents! En ce qui concerne l'optimisation, les résultats ne sont dans ce particulier pas trop différents:

### Optimisation des réponses : Réponses

#### Paramètres

Réponse	Objectif	Inférieur	Cible	Supérieur	Pondération	Importance
Réponses	Maximum	32160	34460		1	1

#### Solution

Solution	A	B	C	Réponses	
				Valeur ajustée	Désirabilité composite
1	Route	40	2.5	34390	0.969565

#### Prévisions de réponses multiples

Variable	Configuration
A	Route
B	40
C	2.5

Réponse	Valeur ajustée	ErT ajust	IC à 95 %	IP à 95 %
Réponses	34390	220	(33924; 34856)	(33459; 35321)

## 17.6. Exercice 177.: Analyse graphique des effets principaux et interactions du modèle d'un plan factoriel complet avec répétitions

Minitab® Statistical Software 15.1.1.0

Nous partons toujours du même modèle que celui utilisé précédemment:

	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
	OrdEssai	TypePt	Blocs	A	B	C	Responses							
1	1	1	1	1	1	1	32700							
2	2	1	1	1	1	2	33430							
3	3	1	1	1	1	3	31710							
4	4	1	1	1	2	1	32680							
5	5	1	1	1	2	2	34070							
6	6	1	1	1	2	3	33220							
7	7	1	1	2	1	1	33180							
8	8	1	1	2	1	2	34430							
9	9	1	1	2	1	3	33570							
10	10	1	1	2	2	1	33270							
11	11	1	1	2	2	2	33440							
12	12	1	1	2	2	3	32840							
13	13	1	2	1	1	1	32750							
14	14	1	2	1	1	2	33360							
15	15	1	2	1	1	3	32100							
16	16	1	2	1	2	1	32270							
17	17	1	2	1	2	2	33100							
18	18	1	2	1	2	3	33700							
19	19	1	2	2	1	1	32160							
20	20	1	2	2	1	2	34280							

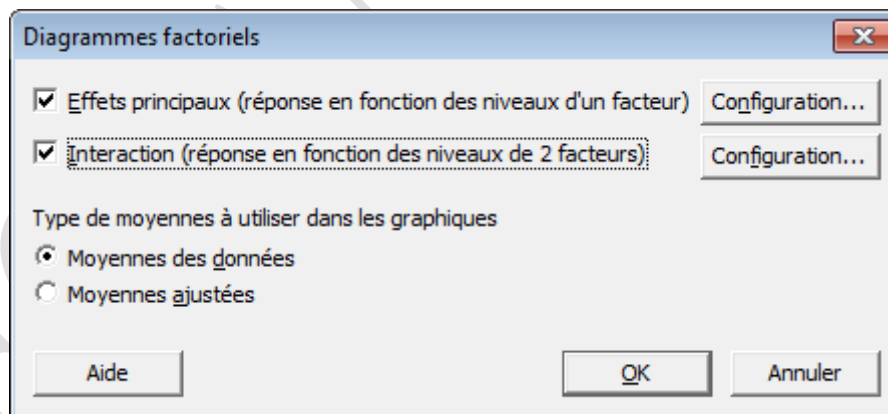
Feuille de travail en cours : Feuille de travail 2

## Diagrammes factoriels (interactions et facteurs principaux)

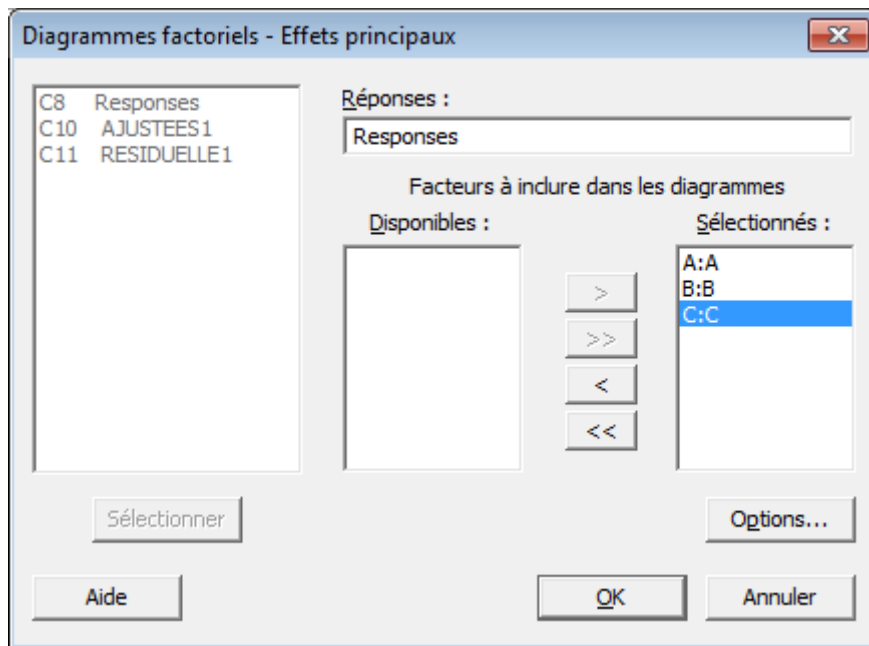
Nous allons dans le menu **Stat/DOE (plan d'expériences)/Plan factoriel/Diagrammes factoriels...**:

	C1	C2	C3	Type1
	OrdreStd	OrdEssai		
1	1	1		
2	2	2		
3	3	3		
4	4	4		
5	5	5		
6	6	6		
7	7	7		
8	8	8		
9	9	9		
10	10	10	1	1
11	11	11	1	1
12	12	12	1	2
13	13	13	1	2
14	14	14	1	2
15	15	15	1	2
16	16	16	1	2
17	17	17	1	2
18	18	18	1	2
19	19	19	1	2
20	20	20	1	2
21	21	21	1	2

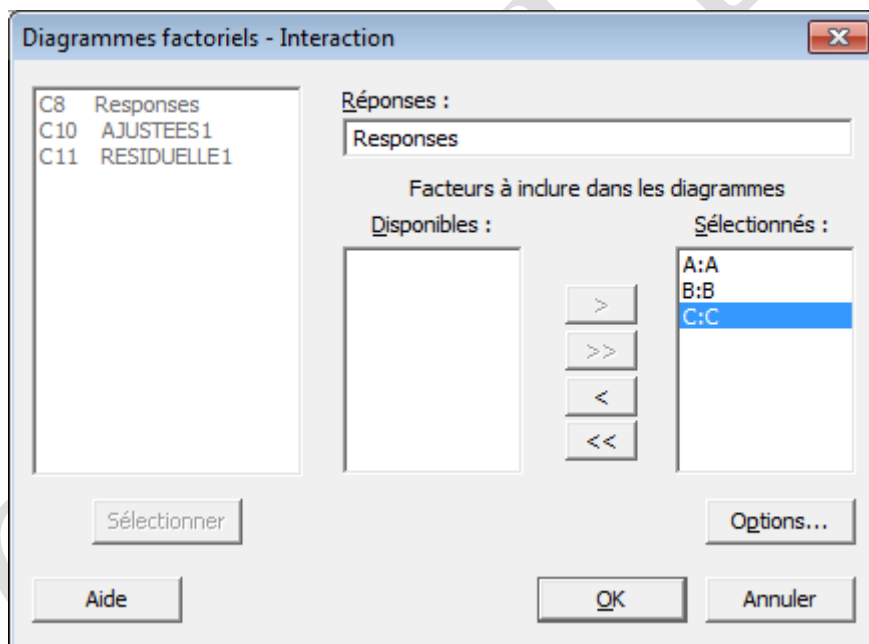
et nous cochoons:



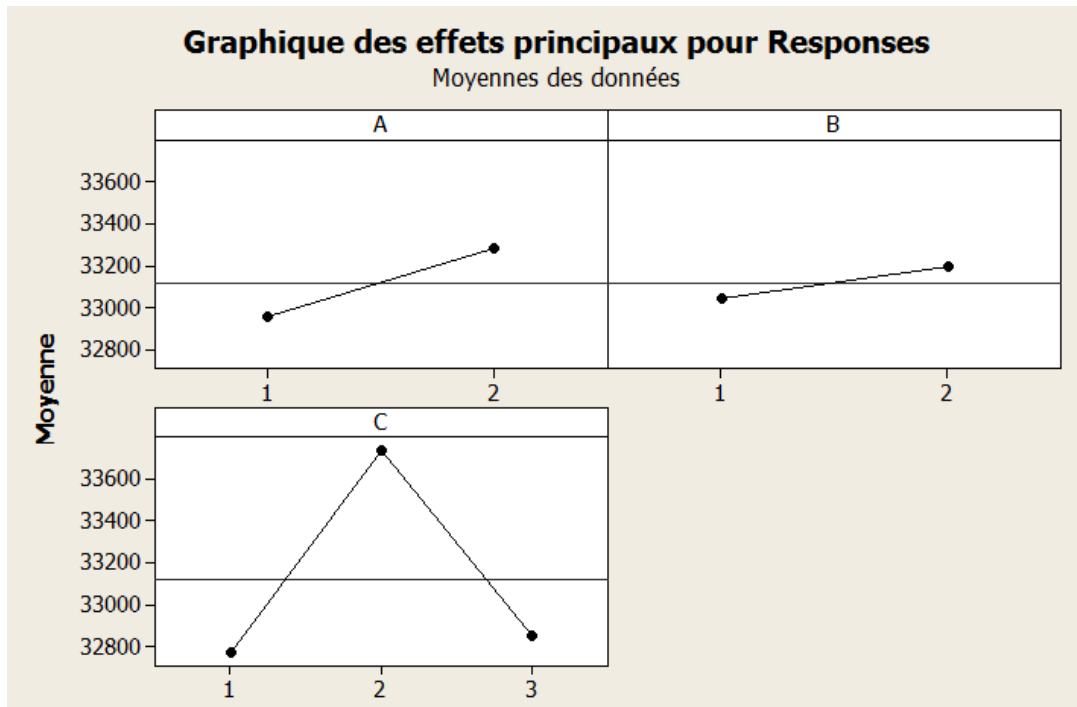
Nous cliquons sur le bouton **Configuration...** pour les effets principaux pour prendre:



et de même pour les interactions:



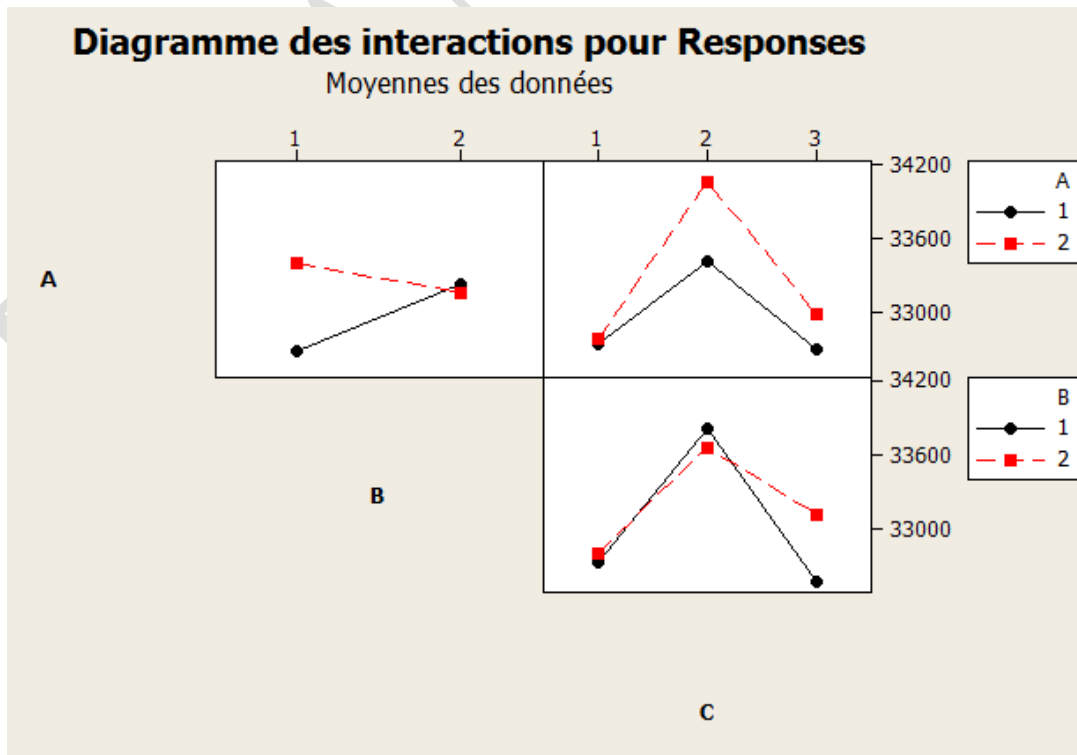
Ce qui nous donne:



Nous retrouvons ci-dessus dans un premier temps pour chaque sous-graphique la moyenne expérimentale générale qui est 33'119.3.

Ensuite, nous retrouvons dans chaque graphique, la moyenne expérimentale de la variable d'observation (distance parcourue) pour chacun des niveaux. Moyennes identiques que celles calculées dans le cours théorique.

Ensuite pour les interactions, nous obtenons le "profileur" suivant:



Nous voyons donc qu'il y a une interaction importante entre A et B (les calculs faits à la main dans le cours théorique nous l'avaient montré). Pour les autres, nous retrouvons aussi les mêmes résultats que dans le cours théorique.

## Effets et Diagrammes de Pareto des Effets (PEF)

Ah voilà un diagramme qui amène souvent la même question: *Comment sont calculées les valeurs sur ce diagramme de Pareto ainsi que les effets????*

Voyons d'abord comment obtenir ce diagramme dans un cas simple avec seulement les trois facteurs principaux.

**Attention!!!!** Le diagramme de Pareto et les valeurs des effets ne sont disponibles que pour les plans factoriels à 2 niveaux!!!

Nous partons de:

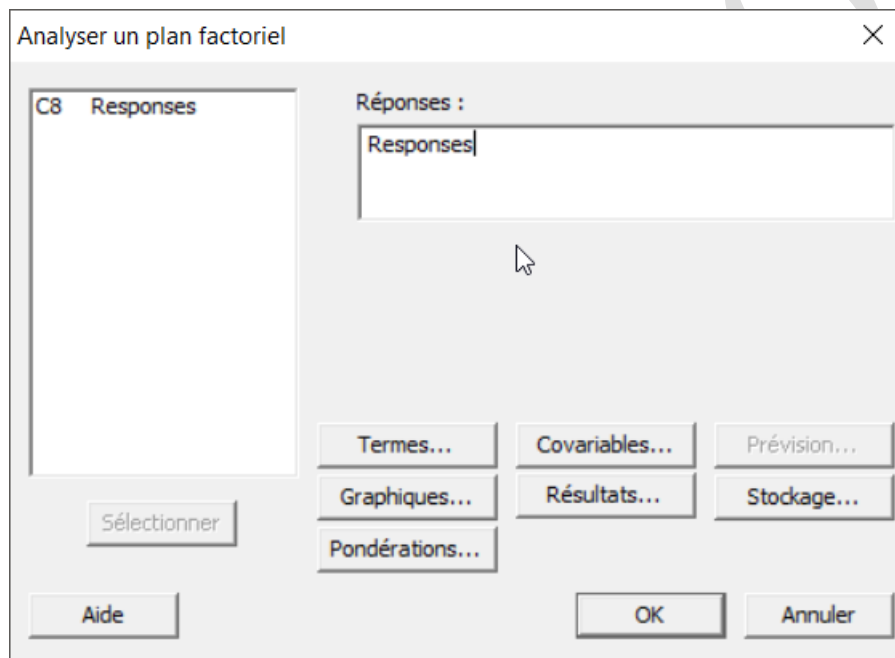
↓	C1	C2	C3	C4	C5	C6	C7	C8
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	C	Réponses
2	2	2	1	1	1	-1	-1	33180
3	3	3	1	1	-1	1	-1	32680
4	4	4	1	1	1	1	-1	33270
5	5	5	1	1	-1	-1	1	33430
6	6	6	1	1	1	-1	1	34430
7	7	7	1	1	-1	1	1	34070
8	8	8	1	1	1	1	1	33440
9	9	9	1	1	-1	-1	-1	32750
10	10	10	1	1	1	-1	-1	32160
11	11	11	1	1	-1	1	-1	32270
12	12	12	1	1	1	1	-1	33080
13	13	13	1	1	-1	-1	1	33360
14	14	14	1	1	1	-1	1	34280
15	15	15	1	1	-1	1	1	33100
16	16	16	1	1	1	1	1	33570
17	17	17	1	1	-1	-1	-1	32960
18	18	18	1	1	1	-1	-1	32640
19	19	19	1	1	-1	1	-1	33130
20	20	20	1	1	1	1	-1	32415
21	21	21	1	1	-1	-1	1	32910
22	22	22	1	1	1	-1	1	34460
23	23	23	1	1	-1	1	1	33610
24	24	24	1	1	1	1	1	34204

Nous allons dans **DOE (plan d'expériences)/Plan factoriel/Analyser un plan factoriel...**:

Minitab - DOE\_FactorielComplet\_3Facteurs.mpj - [Feuille de travail 2 \*\*\*]

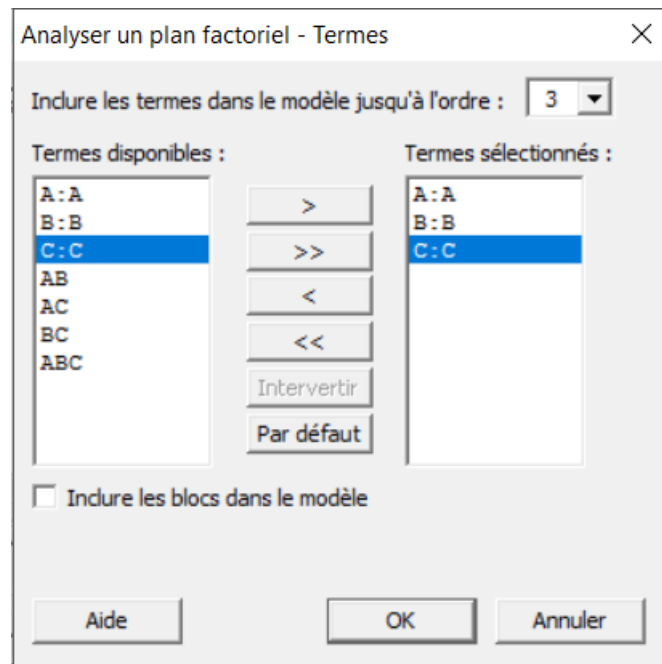
	C1	C2	C3							
	OrdreStd	OrdEssai	Type							
1	1	1								
2	2	2								
3	3	3								
4	4	4								
5	5	5								
6	6	6								
7	7	7								
8	8	8								
9	9	9								
10	10	10	1	1	2	2	1	33270		
11	11	11	1	1	2	2	2	33440		

Nous mettons:

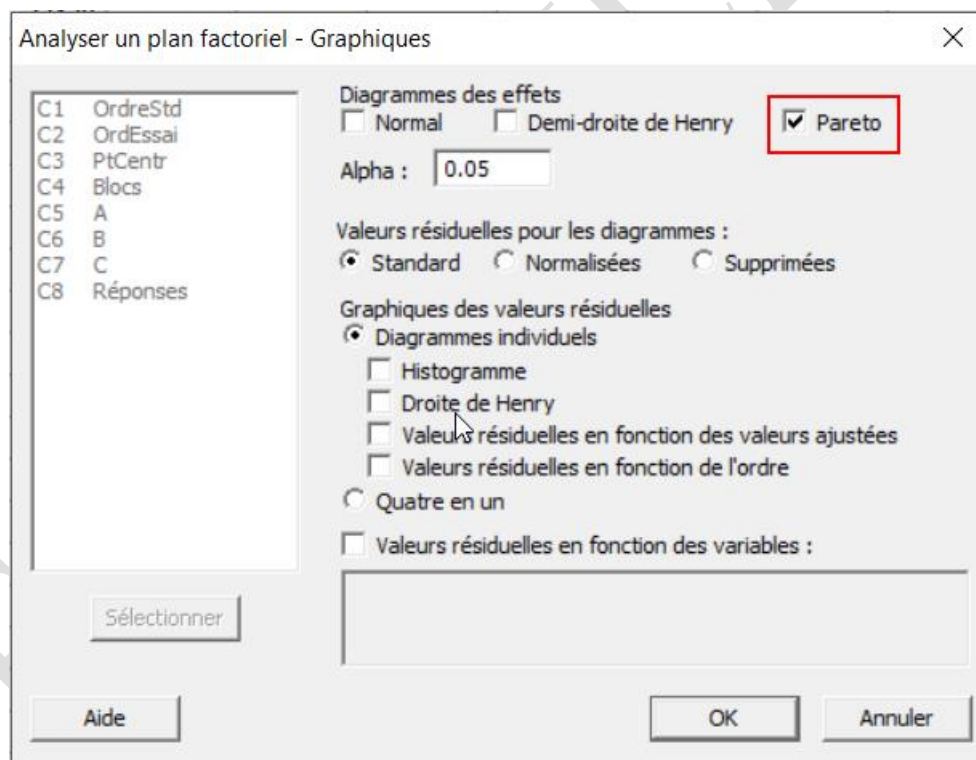


Nous cliquons sur le bouton **Termes...** pour prendre seulement les facteurs principaux (afin de simplifier les calculs à la main):





Nous validons par **OK** et cliquons sur **Graphs...** pour y prendre:



Et nous validons deux fois par **OK** pour obtenir:

**Plan factoriel : Réponses en fonction de A, B, C**

Effets et coefficients estimés pour Réponses (unités codées)

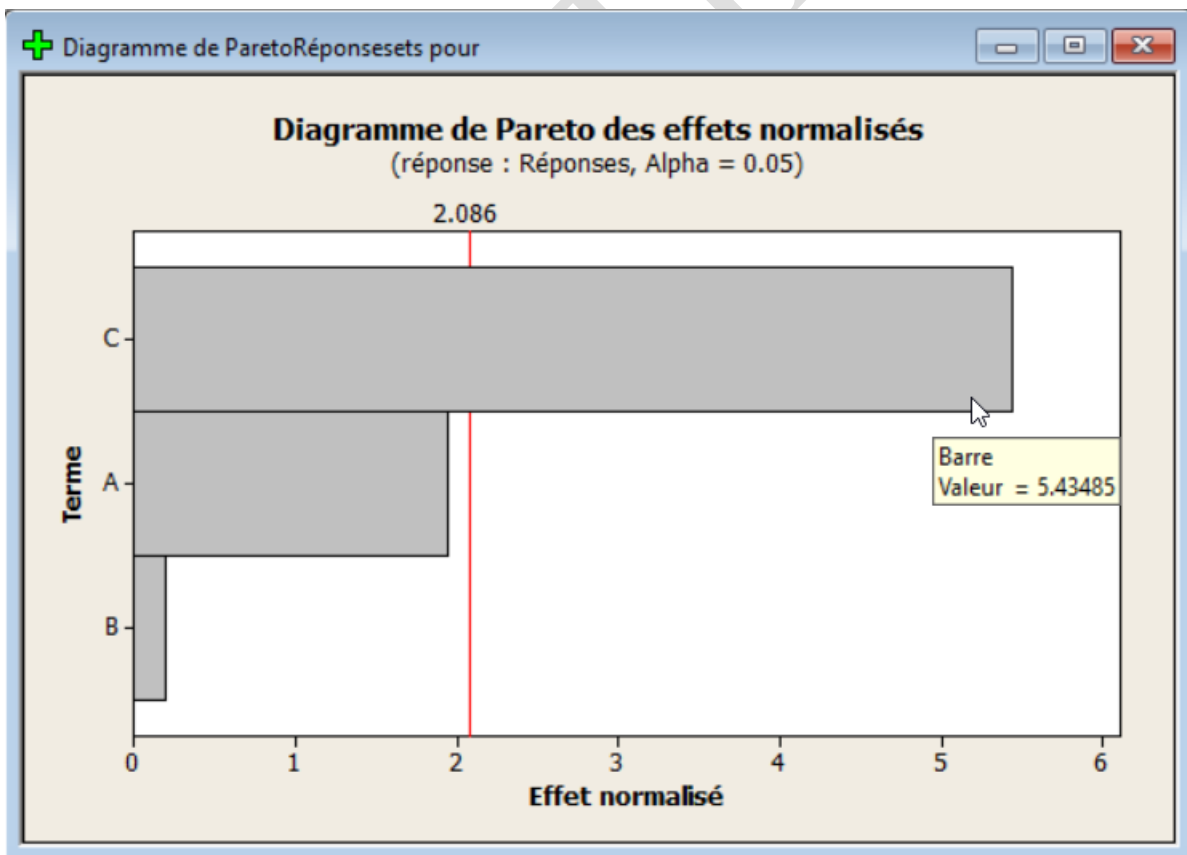
Terme	Effet	Coef	Coef ErT	T	P
Constante		33254.1	89.15	372.99	0.000
A	346.6	173.3	89.15	1.94	0.066
B	-35.1	-17.5	89.15	-0.20	0.846
C	969.1	484.5	89.15	5.43	0.000

S = 436.766      SomCar-ErrPrév = 5494021  
 R carré = 62.51 %      R carré (prév) = 46.02 %      R carré (ajust) = 56.89 %

Analyse de la variance pour Réponses (unités codées)

Source	DL	SomCar séq	SomCar ajust	CM ajust	F	P
Effets principaux	3	6362840	6362840	2120947	11.12	0.000
Erreur résiduelle	20	3815292	3815292	190765		
Inadéquation	4	1500325	1500325	375081	2.59	0.076
Erreur pure	16	2314967	2314967	144685		
Total	23	10178133				

et le graph qui nous intéresse:



Donc d'abord nous pouvons observer que les "Effet normalisé" sont simplement les valeurs absolues prises de:

### Plan factoriel : Réponses en fonction de A, B, C

Effets et coefficients estimés pour Réponses (unités codées)

Terme	Effet	Coeff	Coef ErT	T	P
Constante		33254.1	89.15	372.99	0.000
A	346.6	173.3	89.15	1.94	0.066
B	-35.1	-17.5	89.15	-0.20	0.846
C	969.1	484.5	89.15	5.43	0.000

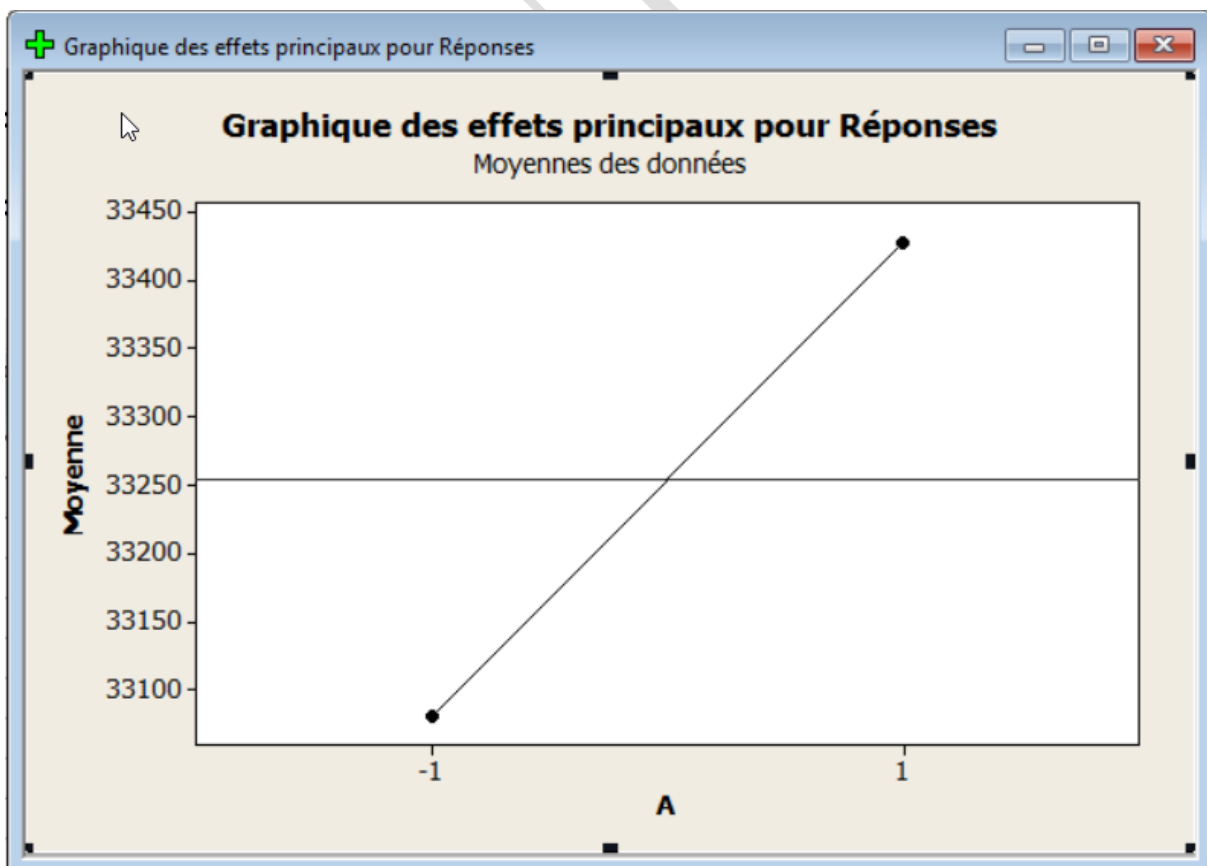
Or cette statistique de Student nous avons vu la démonstration de l'origine de son calcul dans le cours théorique!

Maintenant concentrons-nous sur:

Effets et coefficients estimés pour Réponses (unités codées)

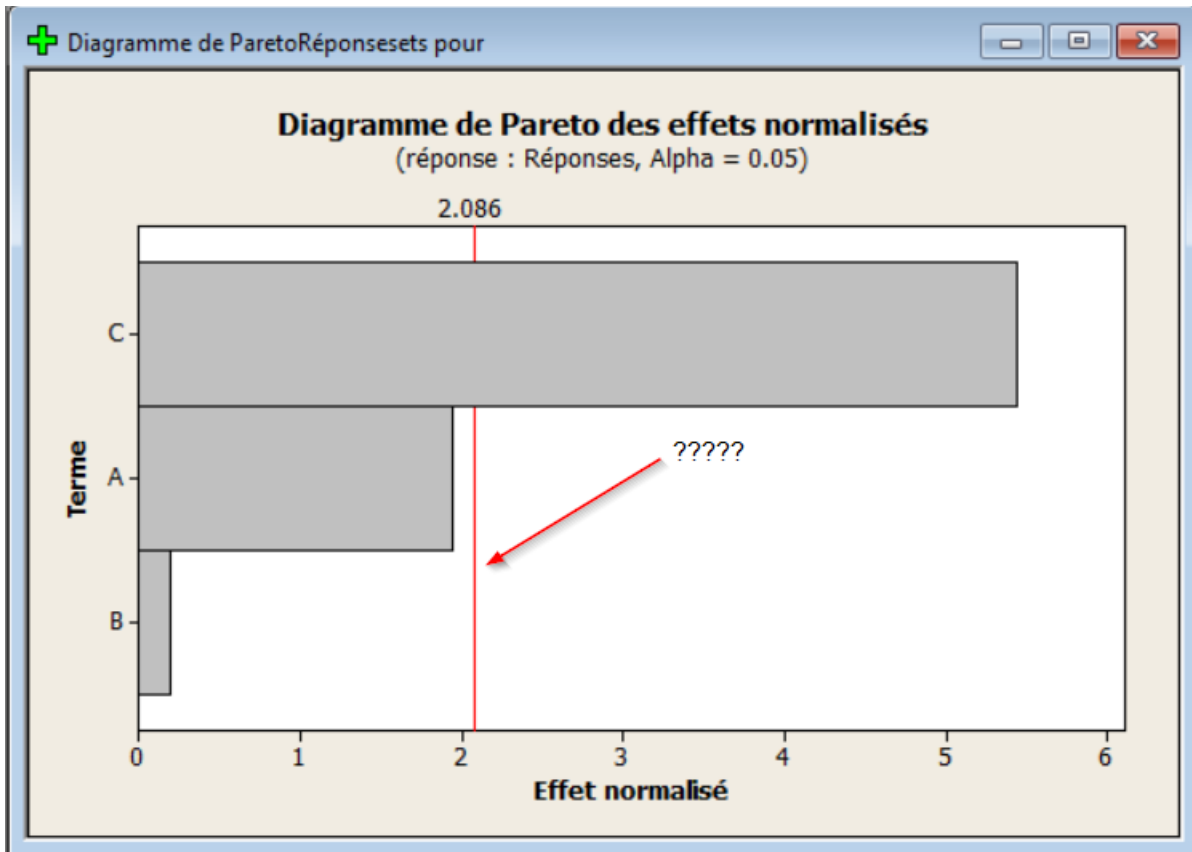
Terme	Effet	Coeff	Coef ErT	T	P
Constante		33254.1	89.15	372.99	0.000
A	346.6	173.3	89.15	1.94	0.066
B	-35.1	-17.5	89.15	-0.20	0.846
C	969.1	484.5	89.15	5.43	0.000

L'origine de ces valeurs nous est en réalité connue, effectivement si nous faisons le graphique des effets principaux pour le facteur A, nous obtenons:



La différence d'amplitude entre le point le plus bas et le point le plus haut et le point le plus bas est égale à l'effet de 346.6! Comme nous l'avons démontré dans le cours théorique, dans le cas d'un facteur à deux niveaux c'est égal au double du coefficient du modèle!

Maintenant la question qui va nous intéresser est d'où vient la ligne rouge appelée "Marge d'erreur"????



Alors Minitab calcule cela ainsi (reproductible avec Microsoft Excel 2007 et supérieur):

$$=LOI.STUDENT.INVERSE.N(1-\alpha/2,DDL \text{ Erreurs résiduelles})$$

Donc dans notre cas comme  $\alpha=5\%$  et:

Analyse de la variance pour Réponses (unités codées)

Source	DL	SomCar séq	SomCar ajust	CM ajust	F	P
Effets principaux	3	6362840	6362840	2120947	11.12	0.000
Erreur résiduelle	20	3815292	3815292	190765		
Inadéquation	4	1500325	1500325	375081	2.59	0.076
Erreur pure	16	2314967	2314967	144685		
Total	23	10178133				

Donc ce qui donne avec Microsoft Excel:

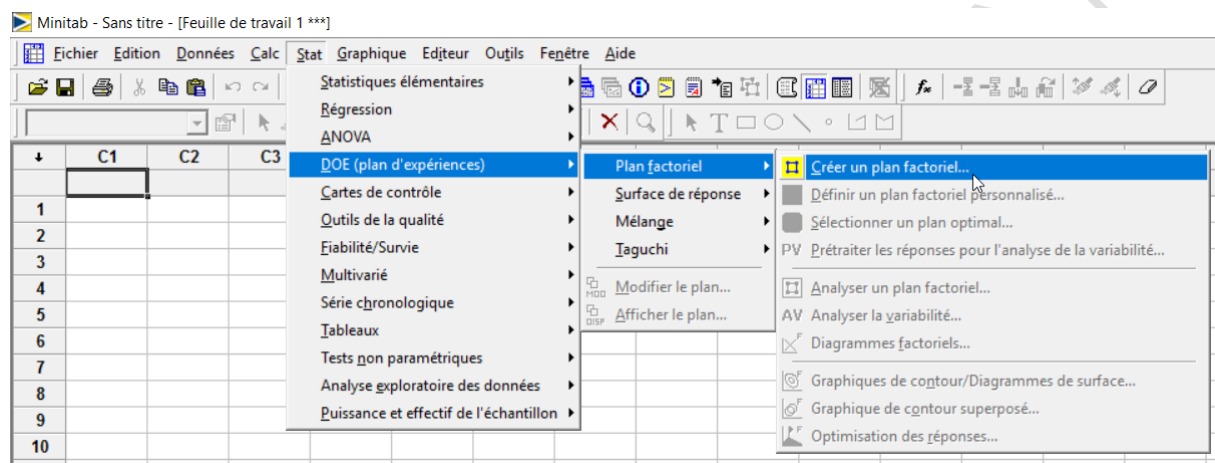
$$=LOI.STUDENT.INVERSE.N(1-5\%/2,20)= 2.085963447$$

## 17.7. Exercice 178.: Lenth's PSE pour un plan factoriel à 2 niveaux sans répétitions

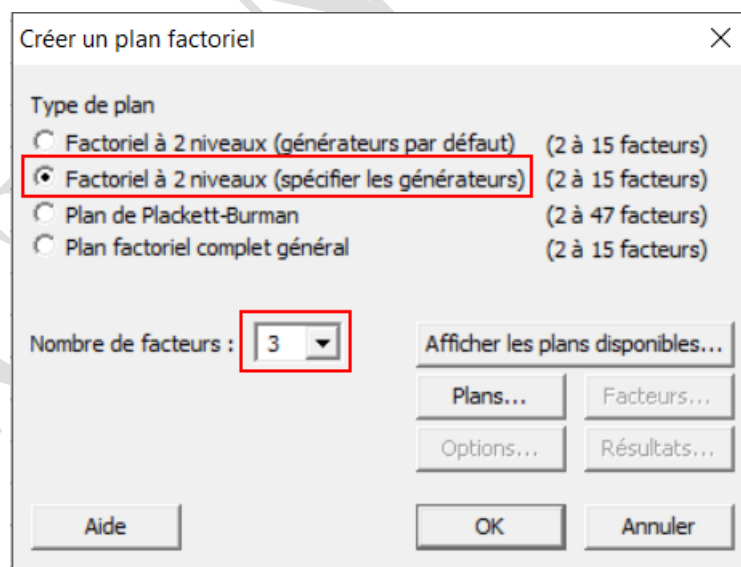
Minitab® Statistical Software 15.1.1.0

Nous allons vérifier ici dans le cas des plans non répliqués si nous obtenons les mêmes résultats que les calculs à la main concernant la pseudo erreur de Lenth.

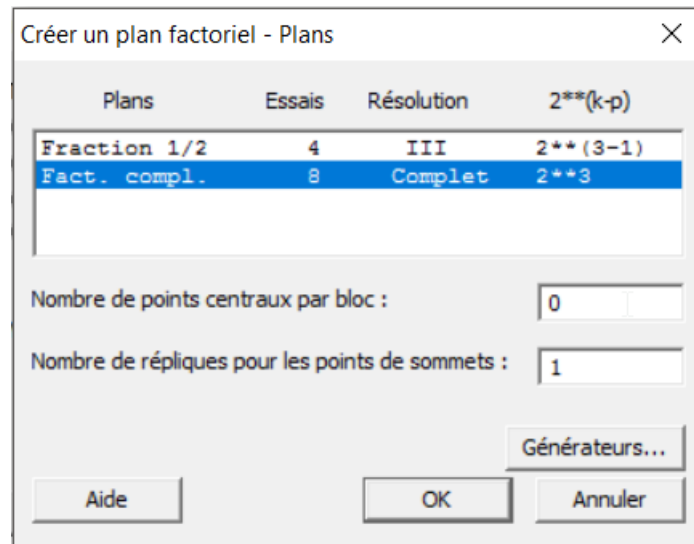
Donc d'abord nous construisons le plan! Nous allons donc dans **Stat/DOE (plan d'expériences)/Créer un plan factoriel...**:



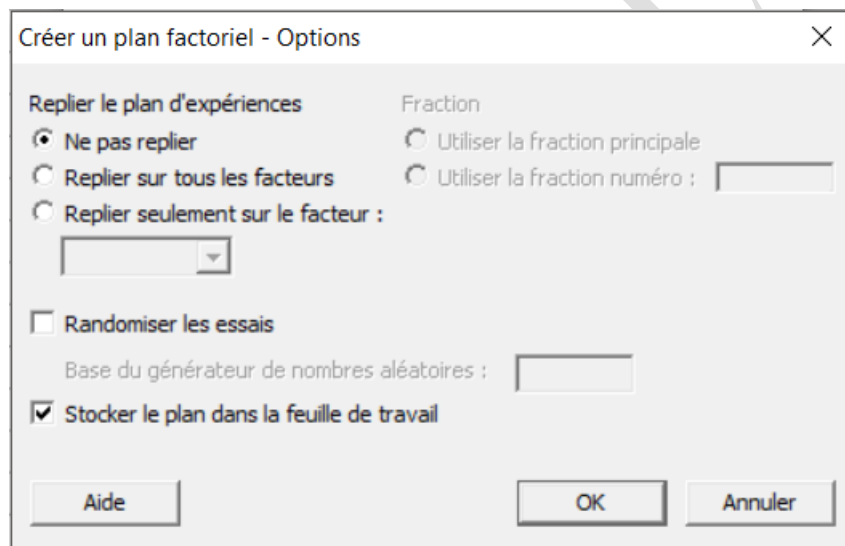
Nous prenons:



Nous cliquons sur **Plans...**:



Nous validons par **OK** puis nous cliquons sur **Options...** pour désactiver l'option **Randomiser les essais**:



Nous validons deux fois par **OK** et saisissons les mesures:

Minitab - DOE\_LenthPSE.MPJ - [Feuille de travail 2 \*\*\*]

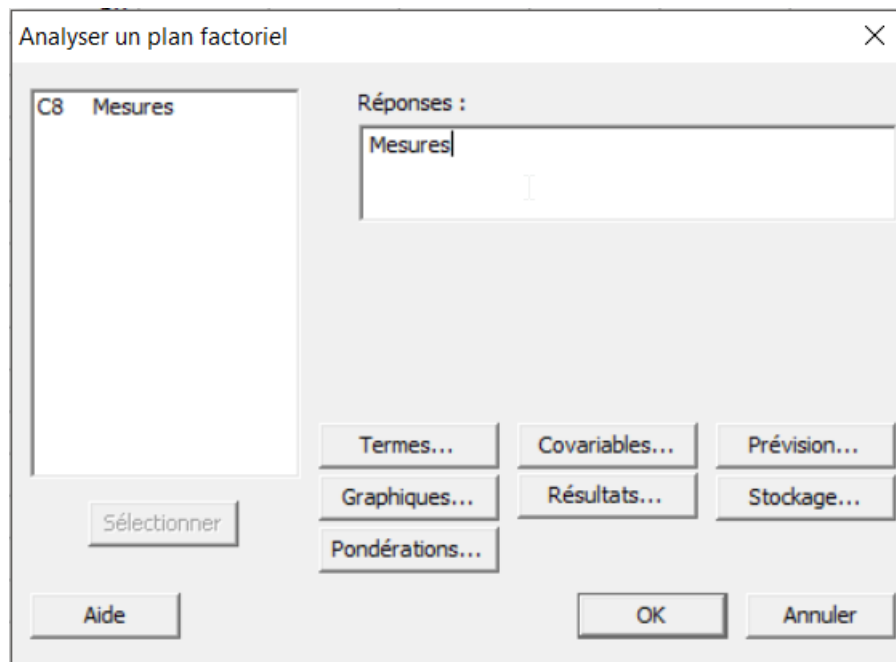
	C1	C2	C3	C4	C5	C6	C7	C8
	Ordre Std	OrdEssai	PtCentr	Blocs	A	B	C	Mesures
1	1	1	1	1	-1	-1	-1	59
2	2	2	1	1	1	-1	-1	74
3	3	3	1	1	-1	1	-1	50
4	4	4	1	1	1	1	-1	69
5	5	5	1	1	-1	-1	1	50
6	6	6	1	1	1	-1	1	81
7	7	7	1	1	-1	1	1	46
8	8	8	1	1	1	1	1	79

Ensuite, nous lançons l'analyse en allant dans **Stat/DOE (plan d'expériences)/Analyser un plan factoriel...**:

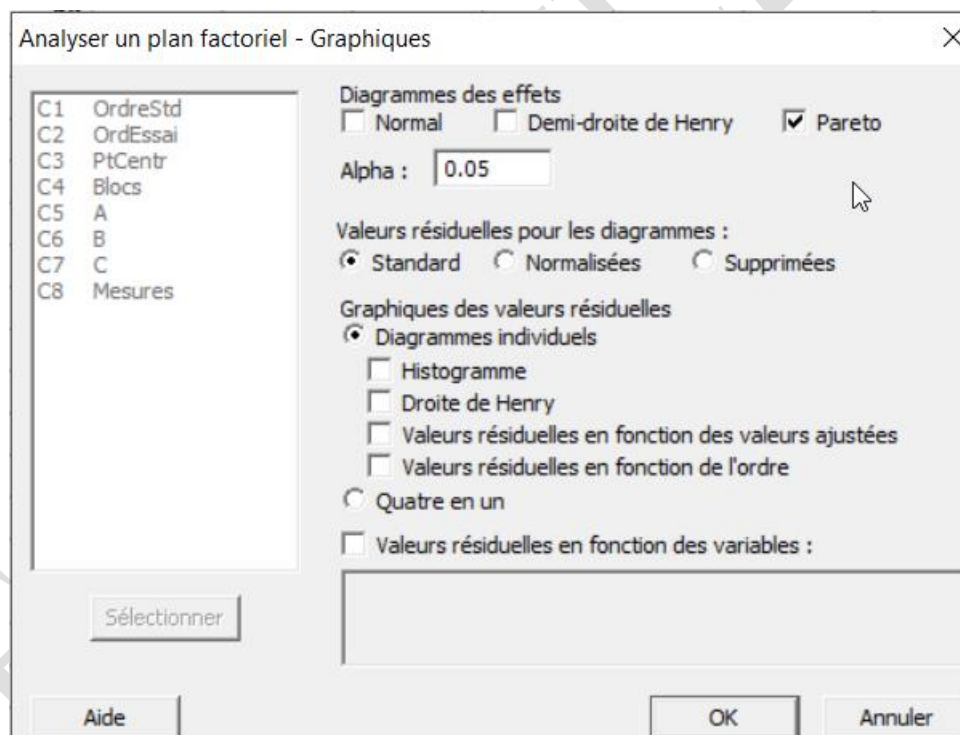
Minitab - DOE\_LenthPSE.MPJ - [Feuille de travail 2 \*\*\*]

The screenshot shows the Minitab interface with the 'Stat' menu open. The path 'Stat > DOE (plan d'expériences) > Plan factoriel > Analyser un plan factoriel...' is highlighted. The background shows a portion of the data table with columns C1, C2, and C3.

ce qui nous donne:



Nous cliquons sur **Graphiques...** pour cocher **Pareto**:



et nous validons deux fois par **OK**. Ce qui donne pour la fenêtre de sessions:

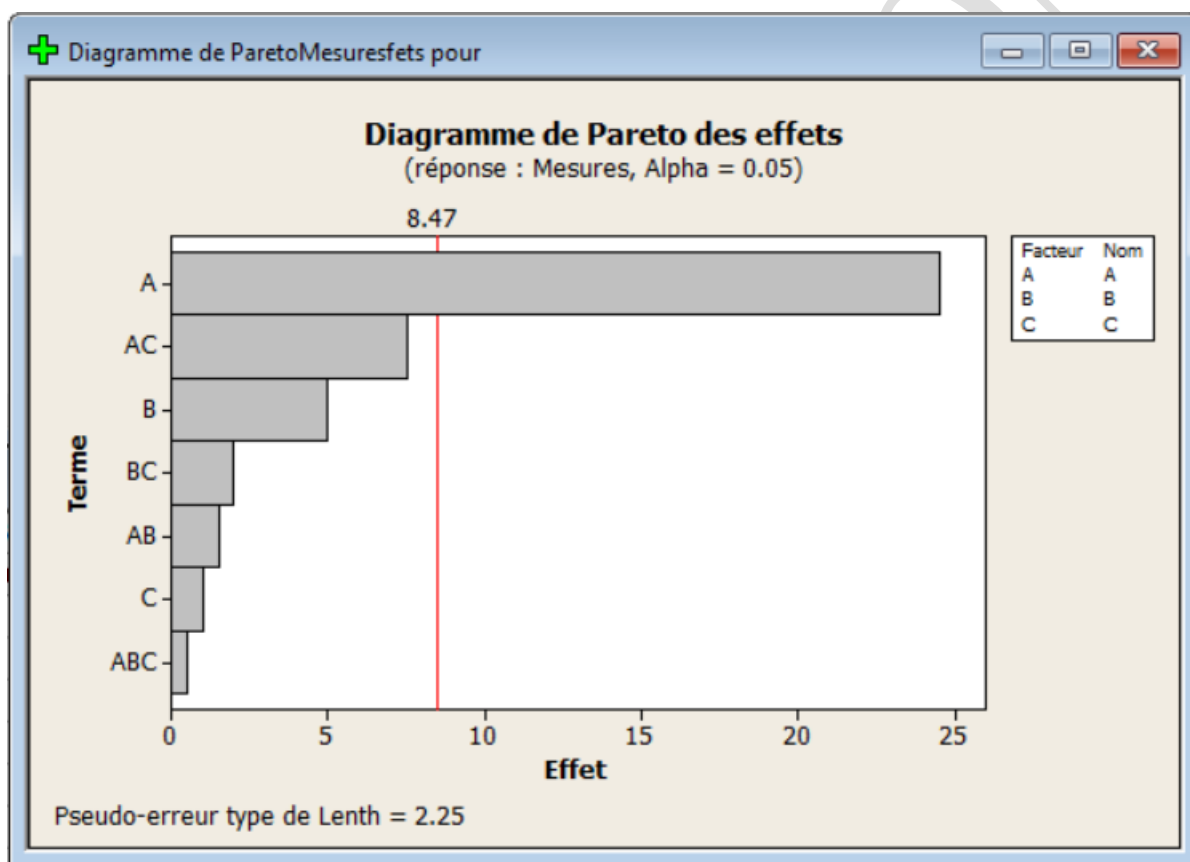


**Plan factoriel : Mesures en fonction de A, B, C**

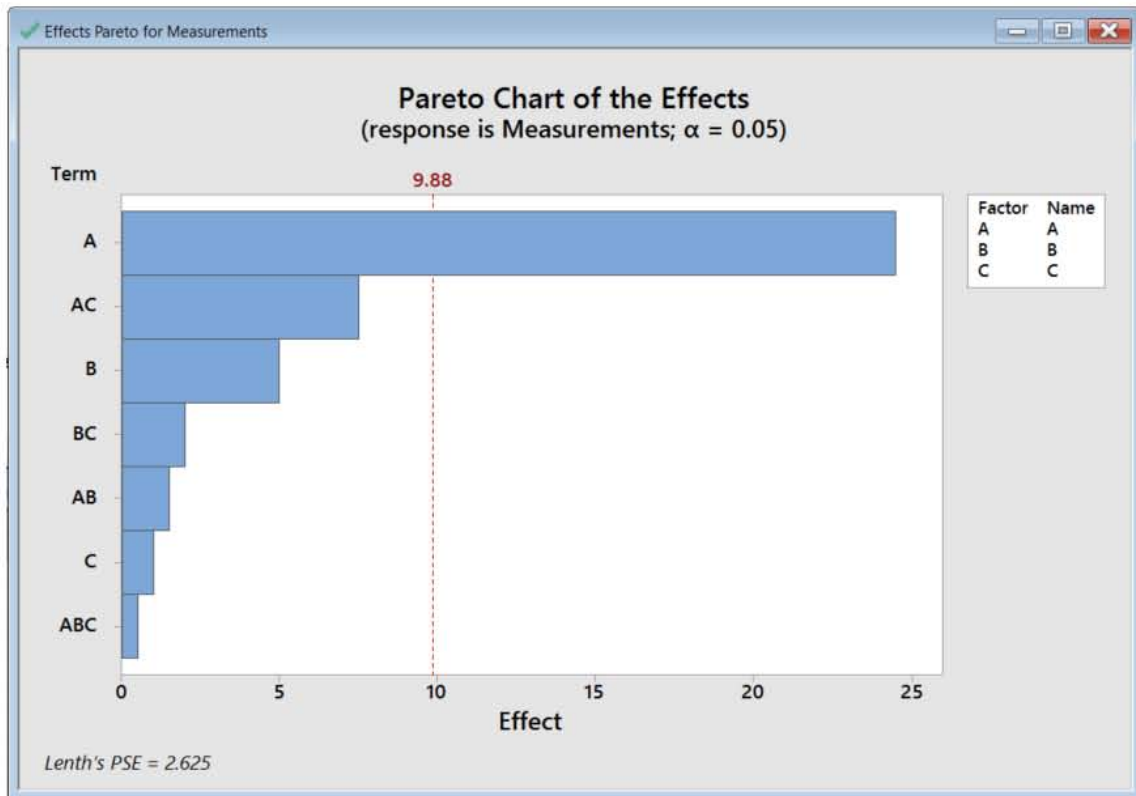
Effets et coefficients estimés pour Mesures (unités codées)

Terme	Effet	Coeff
Constante		63.500
A	24.500	12.250
B	-5.000	-2.500
C	1.000	0.500
A*B	1.500	0.750
A*C	7.500	3.750
B*C	2.000	1.000
A*B*C	-0.500	-0.250

et pour le graph de Pareto:



Comme nous pouvons le voir, les calculs du diagramme de Pareto ne correspondent pas au cas du cours théorique. Mais en procédant identiquement avec Minitab 16 et ultérieur nous obtenons les bonnes valeurs:

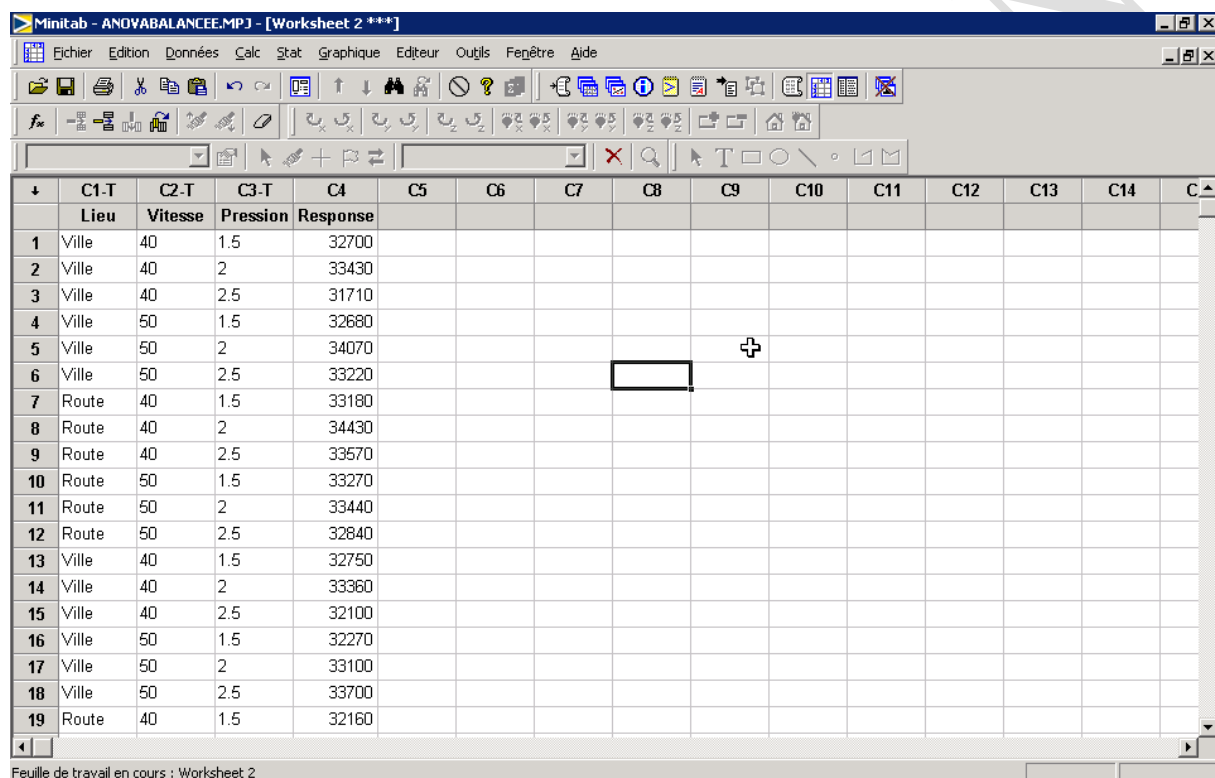


## 17.8. Exercice 179.: Modèle linéaire généralisé d'un plan d'expérience

Minitab® Statistical Software 15.1.1.0

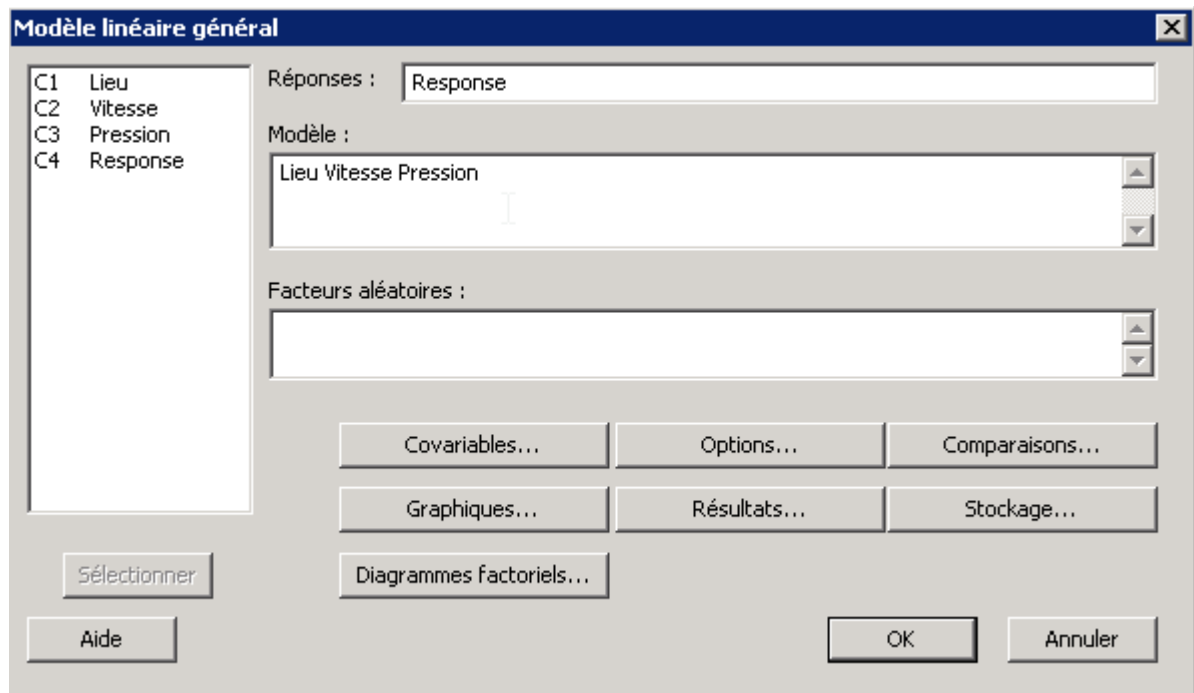
Nous allons voir maintenant le modèle linéaire généralisé en utilisant les données du plan d'expérience précédent mais en imaginant que nous ne savons pas que cela vienne d'un plan d'expérience. Nous allons voir évidemment que nous allons retomber sur mêmes résultats.

Donc nous partons du fichier *ANOVABalancee.mpj* qui contient donc nos 36 mesures avec répliquations:

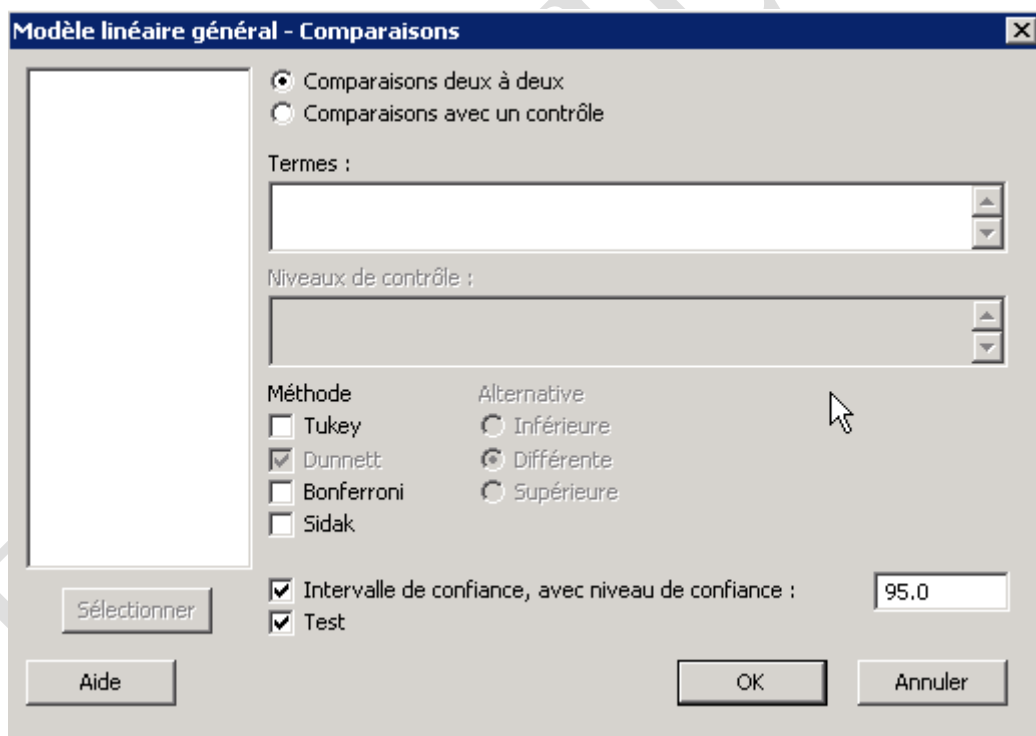


	C1-T	C2-T	C3-T	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
	Lieu	Vitesse	Pression	Response											
1	Ville	40	1.5	32700											
2	Ville	40	2	33430											
3	Ville	40	2.5	31710											
4	Ville	50	1.5	32680											
5	Ville	50	2	34070											
6	Ville	50	2.5	33220											
7	Route	40	1.5	33180											
8	Route	40	2	34430											
9	Route	40	2.5	33570											
10	Route	50	1.5	33270											
11	Route	50	2	33440											
12	Route	50	2.5	32840											
13	Ville	40	1.5	32750											
14	Ville	40	2	33360											
15	Ville	40	2.5	32100											
16	Ville	50	1.5	32270											
17	Ville	50	2	33100											
18	Ville	50	2.5	33700											
19	Route	40	1.5	32160											

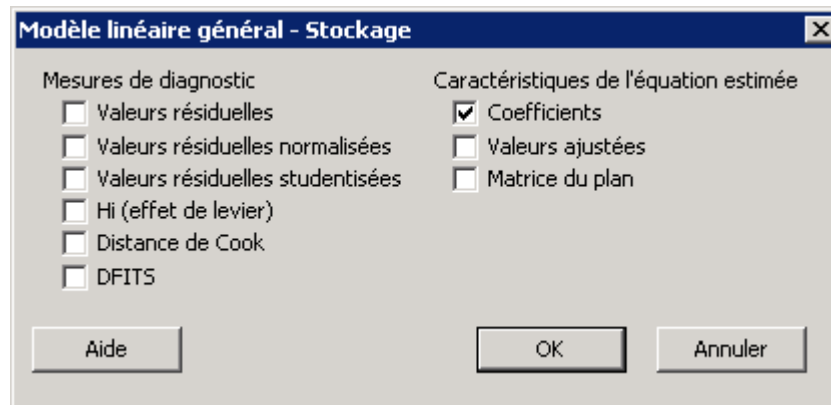
Nous allons dans le menu **Stat/ANOVA/Modèle linéaire général...** et nous commençons par un modèle à trois variables explicatives sans interactions (évidemment nous allons pouvoir vérifier que cet outil gère la variable Catégorielle qu'est le *Lieu*):



et nous cliquons sur le bouton **Comparaisons...**:



pour désactiver le test de Tukey que nous n'avons pas encore vu. Ensuite, nous cliquons sur le menu **Stockage**:



et nous cochons l'option **Coefficients**. Ensuite dans le bouton

et nous validons le tout par **OK** et nous avons alors dans la fenêtre de sessions:

**Modèle linéaire général : Response en fonction de Lieu; Vitesse; ...**

Facteur	Type	Niveaux	Valeurs
Lieu	fixe	2	Route; Ville
Vitesse	fixe	2	40; 50
Pression	fixe	3	1.5; 2; 2.5

Analyse de la variance pour Response, avec utilisation de la somme des carrés ajustée pour les tests

Source	DL	SomCar séq	SomCar ajust	CM ajust	F	P
Lieu	1	961707	961707	961707	3.85	0.059
Vitesse	1	223099	223099	223099	0.89	0.352
Pression	2	6943967	6943967	3471983	13.91	0.000
Erreur	31	7739615	7739615	249665		
Total	35	15868387				

S = 499.665 R carré = 51.23 % R carré (ajust) = 44.93 %

Observations aberrantes pour Response

Observation	Response	Valeur ajustée	Ajust ErT	Valeur résiduelle	Valeur résiduelle normalisée
18	33700.0	32764.9	186.2	935.1	2.02 R

R indique une observation ayant une valeur résiduelle normalisée importante

et dans la feuille, nous retrouvons:

↓	C1-T	C2-T	C3-T	C4	C5	C6
	Lieu	Vitesse	Pression	Response	COEFF1	
1	Ville	40	1.5	32700	33119.3	
2	Ville	40	2	33430	163.4	
3	Ville	40	2.5	31710	-78.7	
4	Ville	50	1.5	32680	-349.7	
5	Ville	50	2	34070	619.4	
6	Ville	50	2.5	33220		
7	Route	40	1.5	33180		
8	Route	40	2	34430		
9	Route	40	2.5	33570		
10	Route	50	1.5	33270		
11	Route	50	2	33440		
12	Route	50	2.5	32840		
13	Ville	40	1.5	32750		
14	Ville	40	2	33360		
15	Ville	40	2.5	32100		
16	Ville	50	1.5	32270		
17	Ville	50	2	33100		
18	Ville	50	2.5	33700		
19	Route	40	1.5	32160		

Feuille de travail en cours : Worksheet 2

Analysons un peu tout cela. D'abord concernant l'ANOVA, nous retrouvons exactement la même que celle effectuée dans les exercices précédents. Nous retrouvons également exactement les mêmes coefficients que dans les exercices vus plus haut ainsi qu'avec le calcul à la main fait dans le cours théorique:

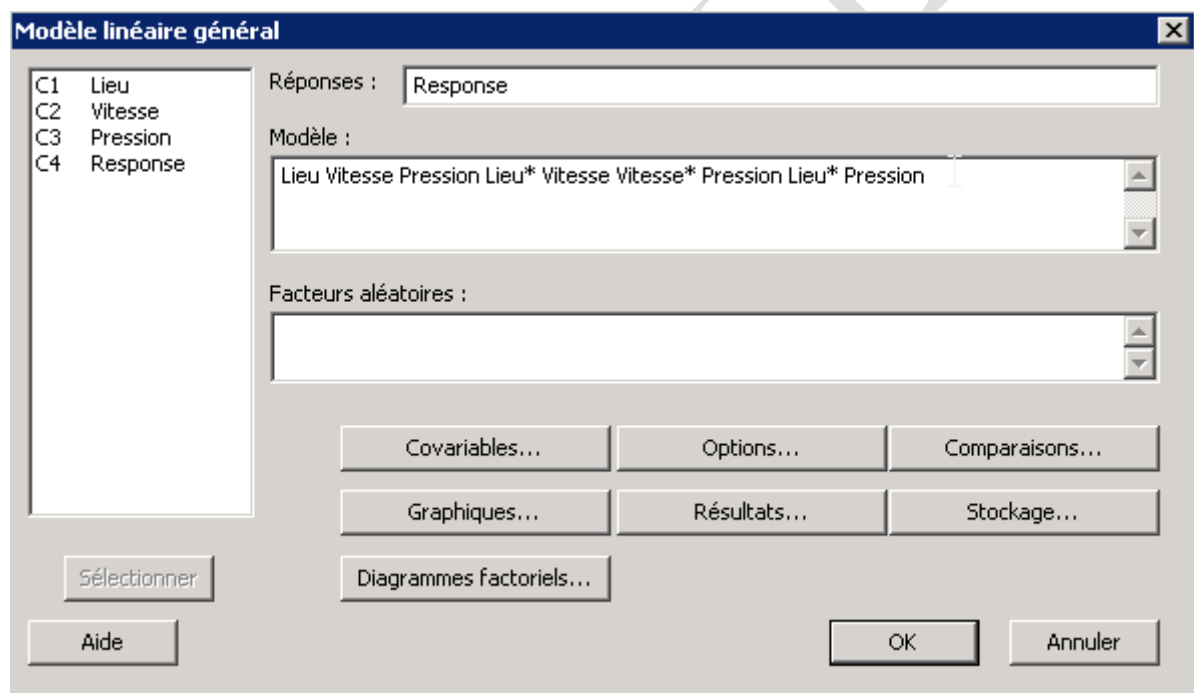
$$Y = 33'119.3 + x_1 \begin{bmatrix} -163.4 \\ 163.4 \end{bmatrix} + x_2 \begin{bmatrix} -78.7 \\ 78.7 \end{bmatrix} + x_3 \begin{bmatrix} -349.7 \\ 619.4 \\ -269.7 \end{bmatrix}$$

$\sum=0$                        $\sum=0$                        $\sum=0$

Dans Minitab 17 la sortie est encore plus conforme qu'à l'époque par rapport à la théorie comme le montre la capture ci-dessous:

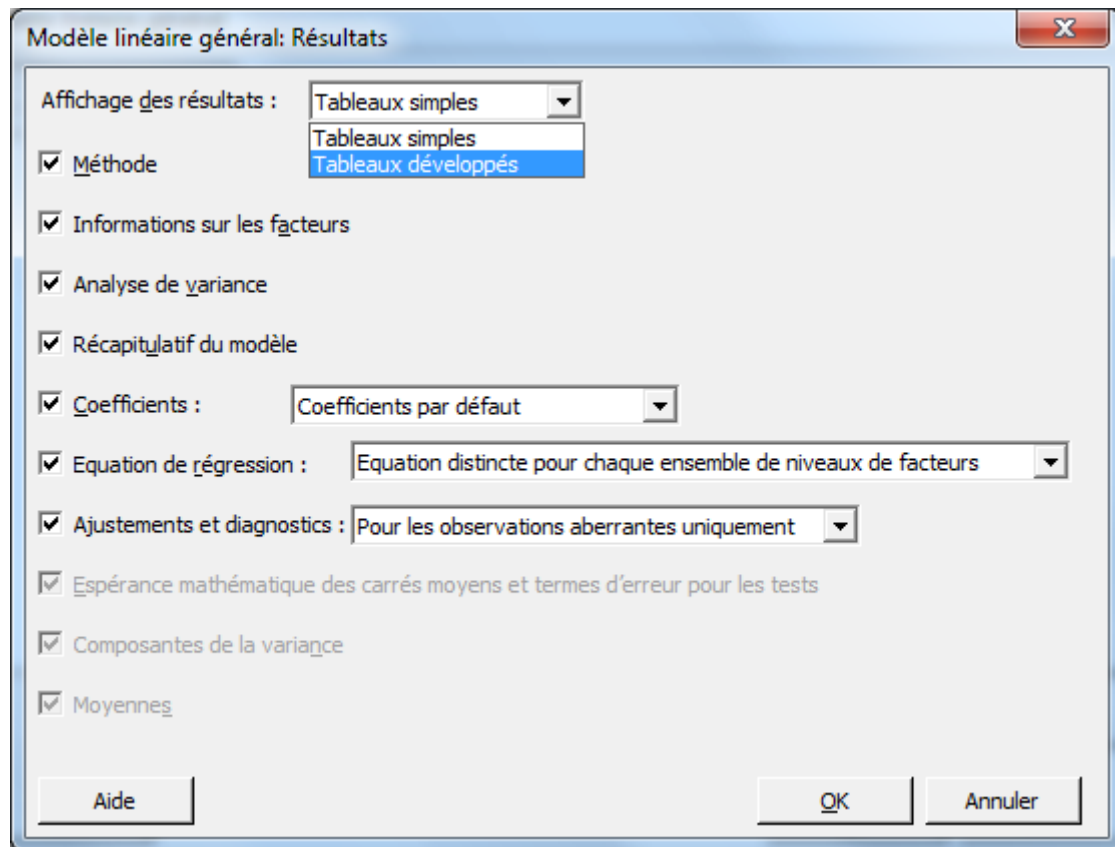
↓	C1-T	C2	C3	C4 <input checked="" type="checkbox"/>	C5
	Lieu	Vitesse	Pression	Response	COEFF1
1	Ville	40	1.5	32700	33119.3
2	Ville	40	2.0	33430	163.4
3	Ville	40	2.5	31710	-163.4
4	Ville	50	1.5	32680	-78.7
5	Ville	50	2.0	34070	78.7
6	Ville	50	2.5	33220	-349.7
7	Route	40	1.5	33180	619.4
8	Route	40	2.0	34430	-269.7
9	Route	40	2.5	33570	
10	Route	50	1.5	33270	
11	Route	50	2.0	33440	
12	Route	50	2.5	32840	
13	Ville	40	1.5	32750	

Il en va de même si nous introduisons toutes les interactions:



le lecteur pourra vérifier qu'il retrouve encore une fois tous les résultats vus plus haut. Ce qui est réconfortant!

Une chose intéressante avec Minitab 16 et 17 est l'option directe de l'équation dans la fenêtre de session et des intervalles de confiance. Pour ce dernier il faut aller dans **Résultats...** et choisir **Tableaux développés**:



Nous obtenons alors:



Minitab - ANOVA\_Balancee.mpj - [Session]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assistant

Analyse de variance

Source	DL	SomCar séq	Contribution	SomCar ajust	CM ajust
Lieu	1	961707	6.06%	961707	961707
Vitesse	1	223099	1.41%	223099	223099
Pression	2	6943967	43.76%	6943967	3471983
Erreur	31	7739615	48.77%	7739615	249665
Inadéquation de l'ajustement	7	4337764	27.34%	4337764	619681
Erreur pure	24	3401851	21.44%	3401851	141744
Total	35	15868387	100.00%		

Source	Valeur F	Valeur de p
Lieu	3.85	0.059
Vitesse	0.89	0.352
Pression	13.91	0.000
Erreur		
Inadéquation de l'ajustement	4.37	0.003
Erreur pure		
Total		

Récapitulatif du modèle

S	R carré	R carré (ajust)	SomCar-ErrPrév	R carré (prév)
499.665	51.23%	44.93%	10437607	34.22%

Coefficients

Terme	Coeff	Coef ErT	IC à 95 %	Valeur de T	Valeur de p	FIV
Constante	33119.3	83.3	(32949.4; 33289.1)	397.70	0.000	
Lieu						
Route	163.4	83.3	( -6.4; 333.3)	1.96	0.059	1.00
Vitesse						
40	-78.7	83.3	( -248.6; 91.1)	-0.95	0.352	1.00
Pression						
1.5	-350	118	( -590; -109)	-2.97	0.006	1.33
2.0	619	118	( 379; 860)	5.26	0.000	1.33

Equation de régression

Response = 33119.3 + 163.4 Lieu\_Route - 163.4 Lieu\_Ville - 78.7 Vitesse\_40 + 78.7 Vitesse\_50 - 350 Pression\_1.5 + 619 Pression\_2.0 - 270 Pression\_2.5

Ajustements et diagnostics pour les observations aberrantes

Observation	Response	Valeur ajustée	ErT ajust	IC à 95 %	Résiduelle	Val. résid. norm.
18	33700	32765	186	(32385; 33145)	935	2.02

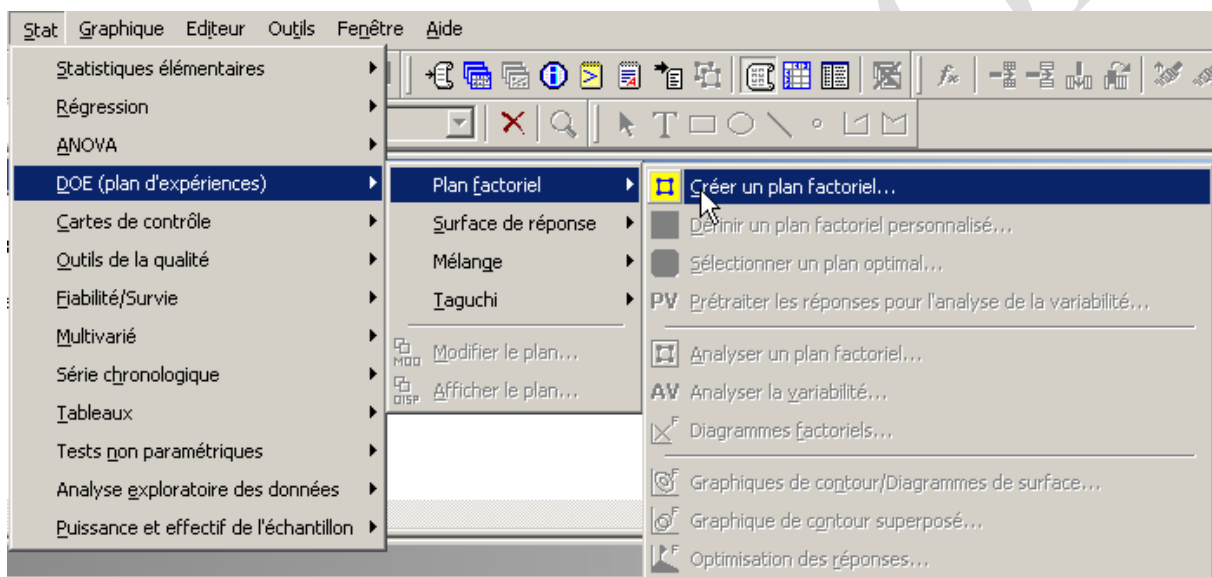
## 17.9. Exercice 180.: Analyse d'un plan factoriel fractionnaire avec choix des alias

Minitab® Statistical Software 15.1.1.0

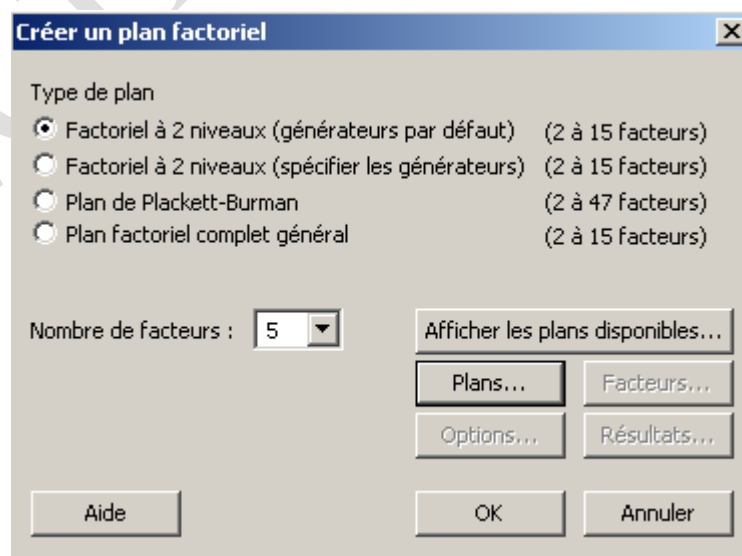
Nous allons voir ici comment faire un plan d'expérience factoriel fractionnaire à 5 facteurs du type  $2^{5-2}$  avec le choix des générateurs. Mais d'abord nous allons comparer les générateurs que choisit automatiquement Minitab à ceux que nous allons ensuite imposer.

### 17.9.1. Choix automatique des alias (confusions)

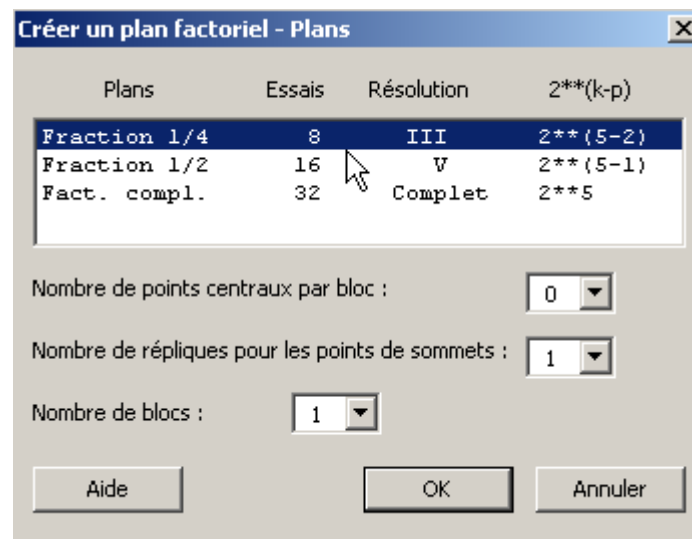
Nous allons donc dans le menu **Stat/DOE (plan d'expériences)/Plan factoriel/Créer un plan factoriel...**:



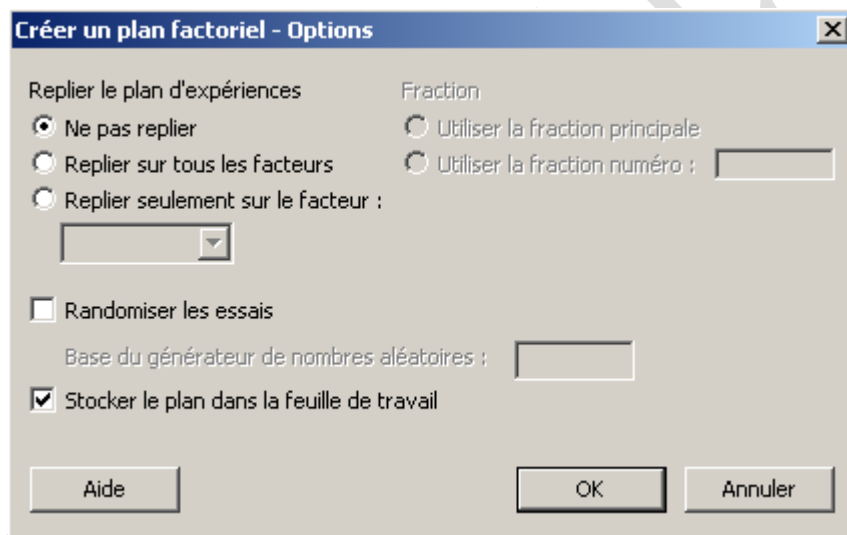
Pour prendre:



Nous cliquons sur **Plans...** pour prendre:



Nous désactivons la randomisation des essais dans le bouton **Options...**:



En validant par **OK**, nous avons alors dans la fenêtre de session (les générateurs sont ainsi que les alias conformes à ce qui a été démontré dans le cours théorique):

**Plan factoriel fractionnaire**

Facteurs : 5 Plan de base : 5; 8 Résolution : III  
 Essais : 8 Répliques : 1 Fraction : 1/4  
 Blocs : 1 Points centraux (total) : 0

\* REMARQUE \* Certains effets principaux sont confondus avec des interactions à deux facteurs.

Alias initiales (indépendantes) : D = AB; E = AC

Structure des alias

I + ABD + ACE + BCDE

A + BD + CE + ABCDE

B + AD + CDE + ABCE

C + AE + BDE + ABCD

D + AB + BCE + ACDE

E + AC + BCD + ABDE

BC + DE + ABE + ACD

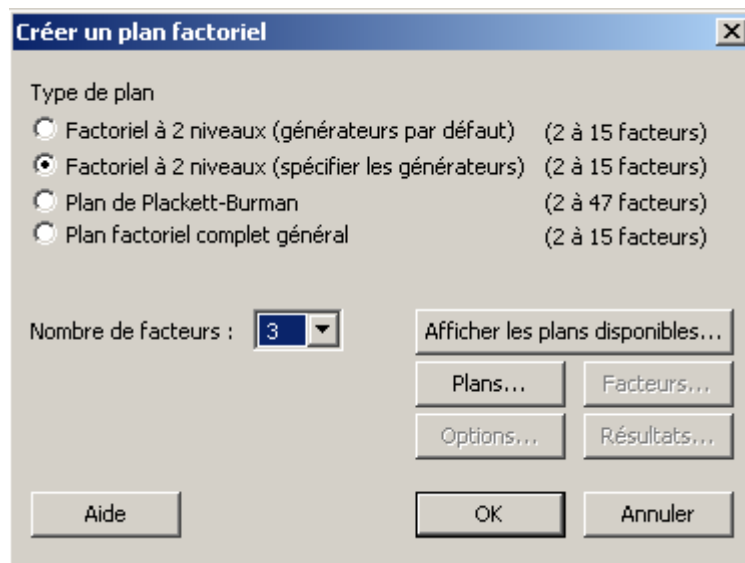
BE + CD + ABC + ADE

avec le plan:

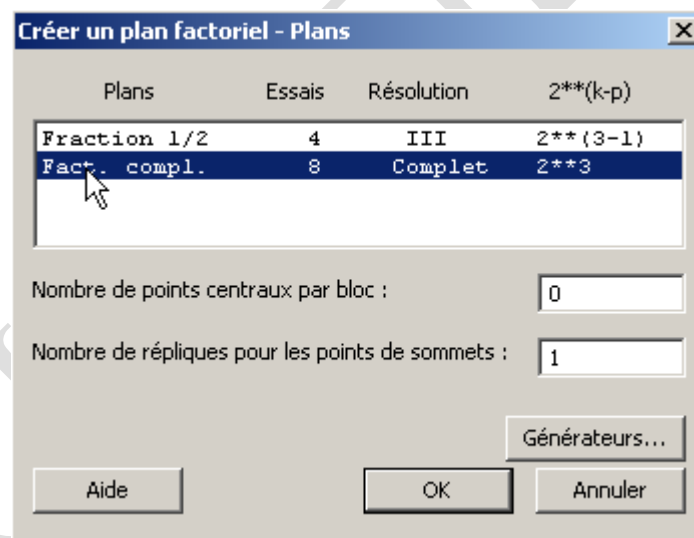
Feuille de travail 3 ***									
↓	C1	C2	C3	C4	C5	C6	C7	C8	C9
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	C	D	E
1	1	1	1	1	-1	-1	-1	1	1
2	2	2	1	1	1	-1	-1	-1	-1
3	3	3	1	1	-1	1	-1	-1	1
4	4	4	1	1	1	1	-1	1	-1
5	5	5	1	1	-1	-1	1	1	-1
6	6	6	1	1	1	-1	1	-1	1
7	7	7	1	1	-1	1	1	-1	-1
8	8	8	1	1	1	1	1	1	1

### 17.9.2. Choix manuel des alias (confusions)

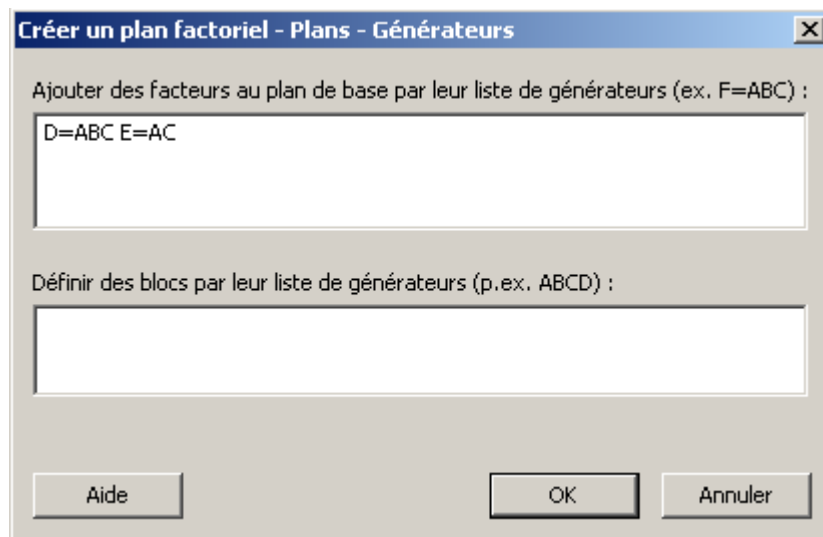
Maintenant, refaisons la même chose mais en imposant les facteurs:



Si nous cliquons sur le bouton **Plans...** pour choisir a priori un plan factoriel complet avec 8 essais:



et nous cliquons sur le bouton **Générateurs...**:



Remarque: Dans le cas de notre exemple avec la voiture {A: route, B: pression, C:vitesse} nous pouvons imaginer que nous avons en fait deux variables en plus: {D: Type de conduite, E: Dureté suspension).

Donc nous prenons 3 facteurs et générons deux autres facteurs avec cette méthode basée sur notre retour sur expérience (REX). Nous désactivons aussi la randomisation, pour obtenir dans la fenêtre de session:

**Plan factoriel fractionnaire**

Facteurs : 5    Plan de base :                    3; 8    Résolution : III  
 Essais : 8    Répliques :                    1    Fraction : 1/4  
 Blocs : 1    Points centraux (total) : 0

\* REMARQUE \* Certains effets principaux sont confondus avec des interactions à deux facteurs.

Alias initiales (indépendantes) : D = ABC; E = AC

Structure des alias (jusqu'à l'ordre 3)

I + ACE + BDE

A + CE + BCD + ABDE  
 B + DE + ACD + ABCE  
 C + AE + ABD + BCDE  
 D + BE + ABC + ACDE  
 E + AC + BD + ABCDE  
 AB + CD + ADE + BCE  
 AD + BC + ABE + CDE  
 ABCD

et le plan correspondant (où nous pouvons vérifier que les générateurs sont bien respectés):

↓	C1	C2	C3	C4	C5	C6	C7	C8	C9
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	C	D	E
1	1	1	1	1	-1	-1	-1	-1	1
2	2	2	1	1	1	-1	-1	1	-1
3	3	3	1	1	-1	1	-1	1	1
4	4	4	1	1	1	1	-1	-1	-1
5	5	5	1	1	-1	-1	1	1	-1
6	6	6	1	1	1	-1	1	-1	1
7	7	7	1	1	-1	1	1	-1	-1
8	8	8	1	1	1	1	1	1	1
9									

Nous voyons bien que des facteurs principaux sont aliasés avec des interactions d'ordre deux, donc si les hypothèses de départ (voir cours théorique) ne sont pas satisfaites, il y aura ambiguïté!

La comparaison côte à côte de la structure d'alias automatique et celle que nous venons d'imposer donne:

Alias initiales (indépendantes) : D = AB; E = AC    Alias initiales (indépendantes) : D = ABC; E = AC

Structure des alias

I + ABD + ACE + BCDE

A + BD + CE + ABCDE  
 B + AD + CDE + ABCE  
 C + AE + BDE + ABCD  
 D + AB + BCE + ACDE  
 E + AC + BCD + ABDE  
 BC + DE + ABE + ACD  
 BE + CD + ABC + ADE

Structure des alias (jusqu'à l'ordre 3)

I + ACE + BDE

A + CE + BCD + ABDE  
 B + DE + ACD + ABCE  
 C + AE + ABD + BCDE  
 D + BE + ABC + ACDE  
 E + AC + BD + ABCDE  
 AB + CD + ADE + BCE  
 AD + BC + ABE + CDE  
 ABCD

Dans le tableau de gauche tout est conforme... (il y a 8 lignes ce qui est rassurant). Dans le tableau de droite, alors que nous n'avons fait que de redéfinir le générateur D=ABC en fin de compte, nous avons une première curiosité... Il y a 9 lignes... (je n'en ai pas encore trouvé la raison).

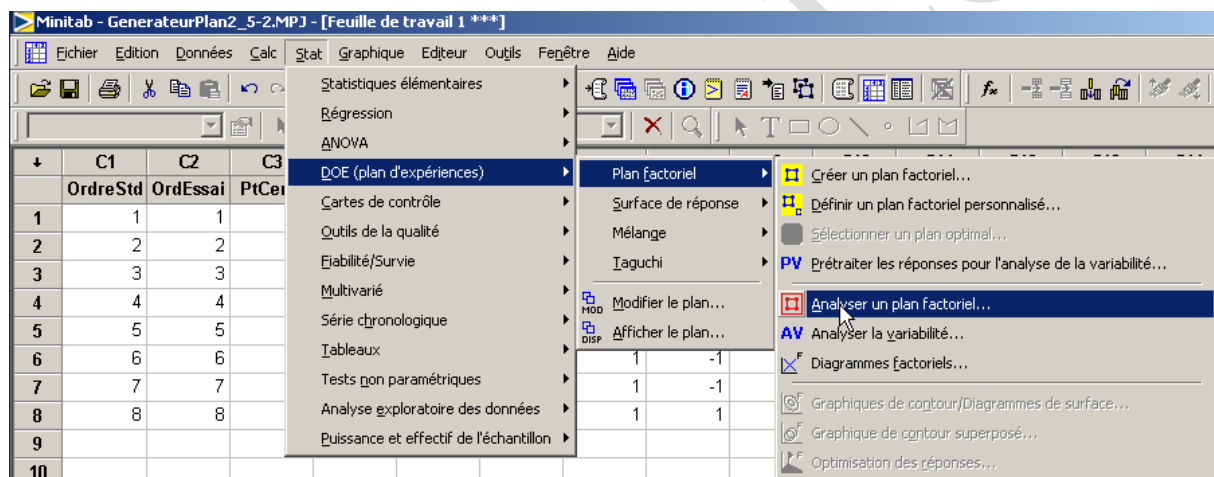


### 17.9.3. Analyse du plan d'expérience aliasé (avec confusions)

Maintenant, nous saisissons les mesures effectuées (pas de répliques ce qui est dommage...):

↓	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	C	D	E	Mesures
1	1	1	1	1	-1	-1	-1	-1	1	27.4
2	2	2	1	1	1	-1	-1	1	-1	31.1
3	3	3	1	1	-1	1	-1	1	1	26.6
4	4	4	1	1	1	1	-1	-1	-1	32.4
5	5	5	1	1	-1	-1	1	1	-1	31.4
6	6	6	1	1	1	-1	1	-1	1	29.3
7	7	7	1	1	-1	1	1	-1	-1	27.5
8	8	8	1	1	1	1	1	1	1	15.5

Nous lançons l'analyse en allant dans le menu **Stat/DOE (plan d'expérience)/Plan factoriel/Analyser un plan factoriel...**:



Nous obtenons alors dans la fenêtre de sessions:

Effets et coefficients estimés pour Mesures (unités codées)

Terme	Effet	Coeff
Constante		27.650
A	-1.150	-0.575
B	-4.300	-2.150
C	-3.450	-1.725
D	-3.000	-1.500
E	-5.900	-2.950
A*B	-1.950	-0.975
A*D	-4.550	-2.275

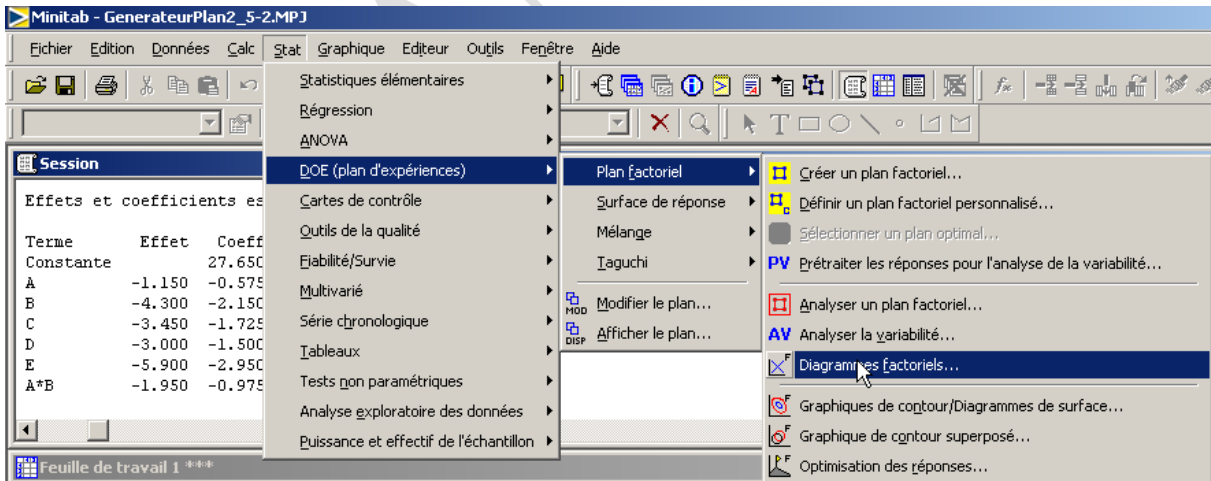
S = \* SomCar-ErrPrév = \*

Analyse de la variance pour Mesures (unités codées)

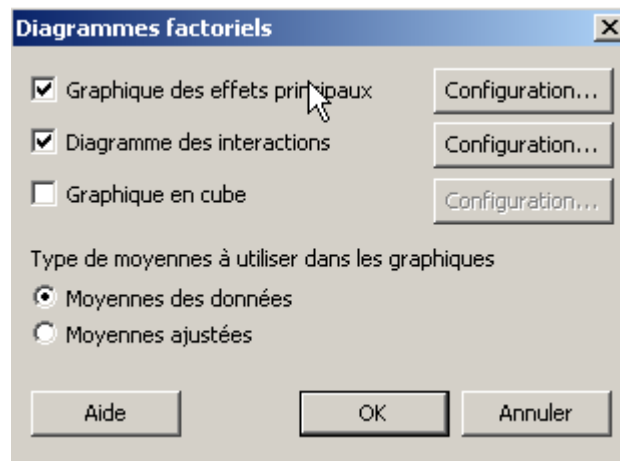
Source	DL	SomCar séq	SomCar ajust	CM ajust	F	P
Effets principaux	5	151.05	151.05	30.21	*	*
2 Interactions - Nombre de facteurs	2	49.01	49.01	24.50	*	*
Erreur résiduelle	0	*	*	*		
Total	7	200.06				

Donc le choix des alias quel qu'il soit est mal parti car au vu de l'amplitude des coefficients et de leurs effets on va droit dans le mur. L'ANOVA elle aussi annonce de mauvaises nouvelles... les 5 effets principaux sont significatifs ainsi que les interactions (d'où l'étoile).

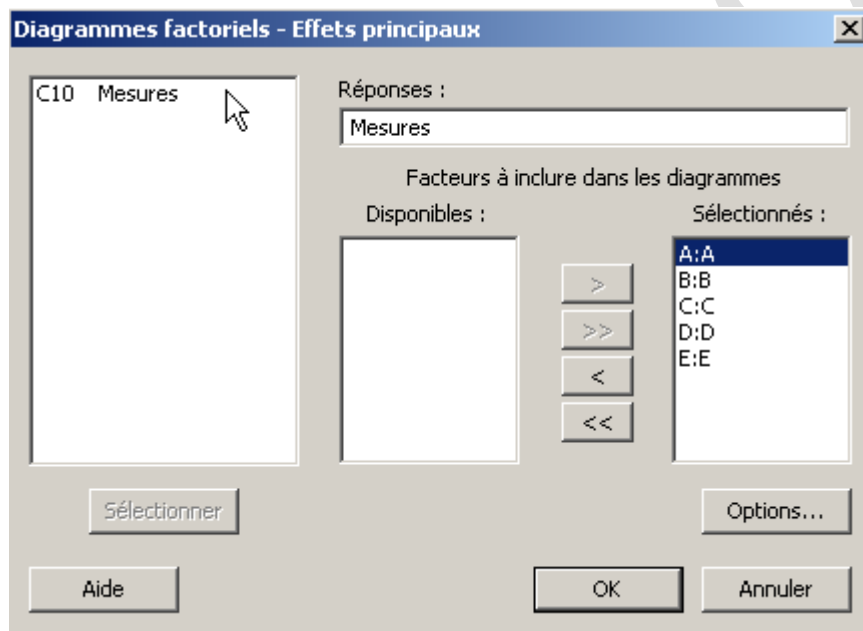
Maintenant, voyons les diagrammes factoriels pour vérifier si effectivement les facteurs interagissent en allant dans **Stat/DOE (plan d'expériences)/Plan factoriel/Diagrammes factoriels...**:



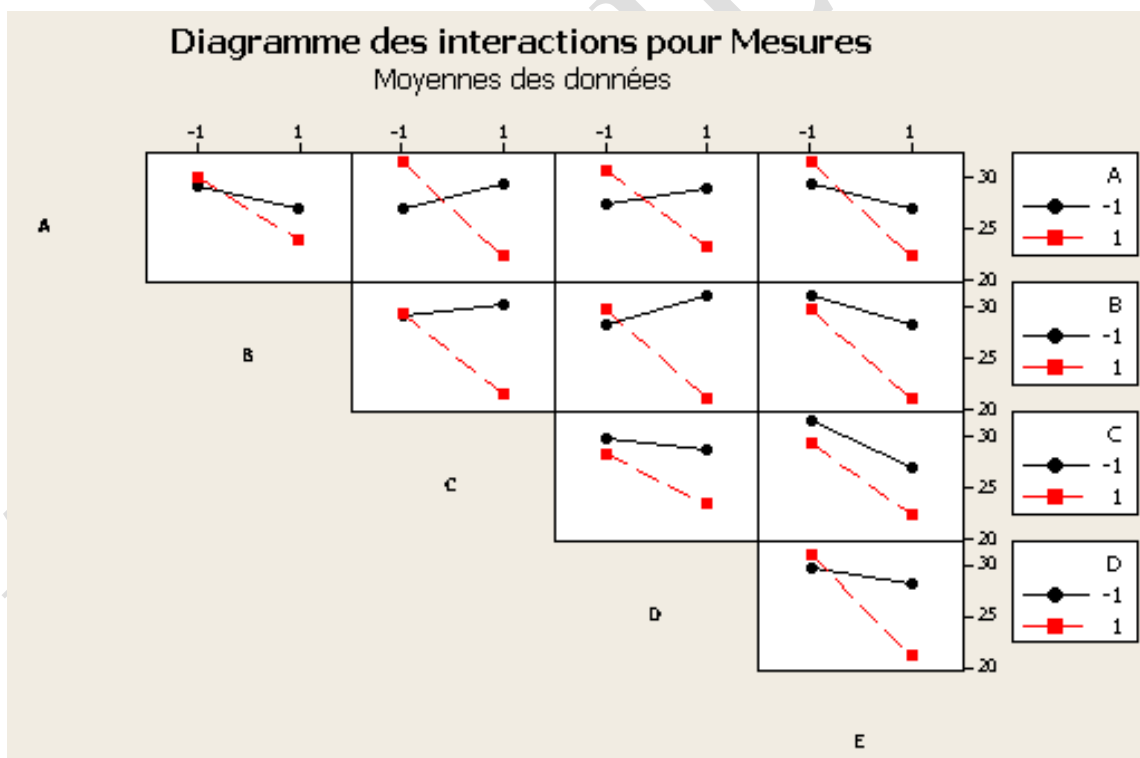
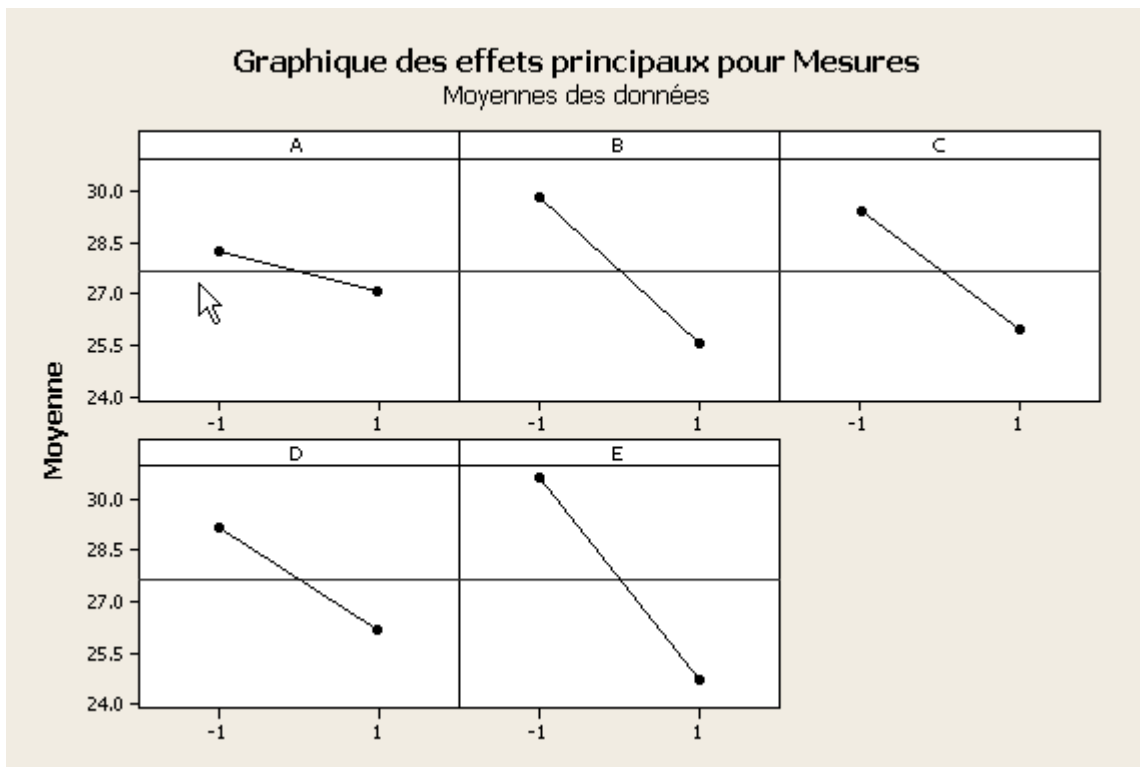
Nous avons alors:



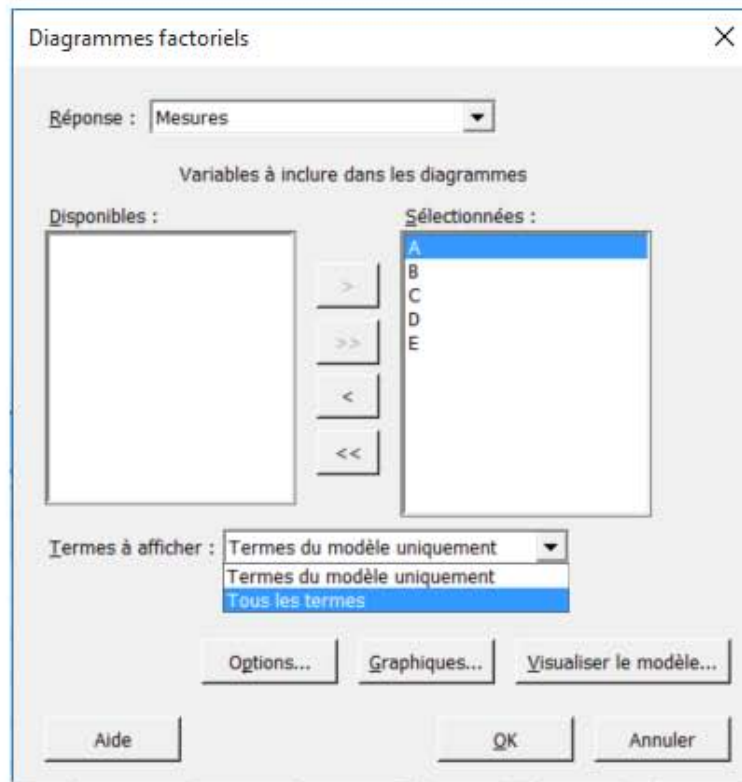
Nous cliquons sur le premier et deuxième bouton **Configuration...** et nous prenons tous les facteurs:



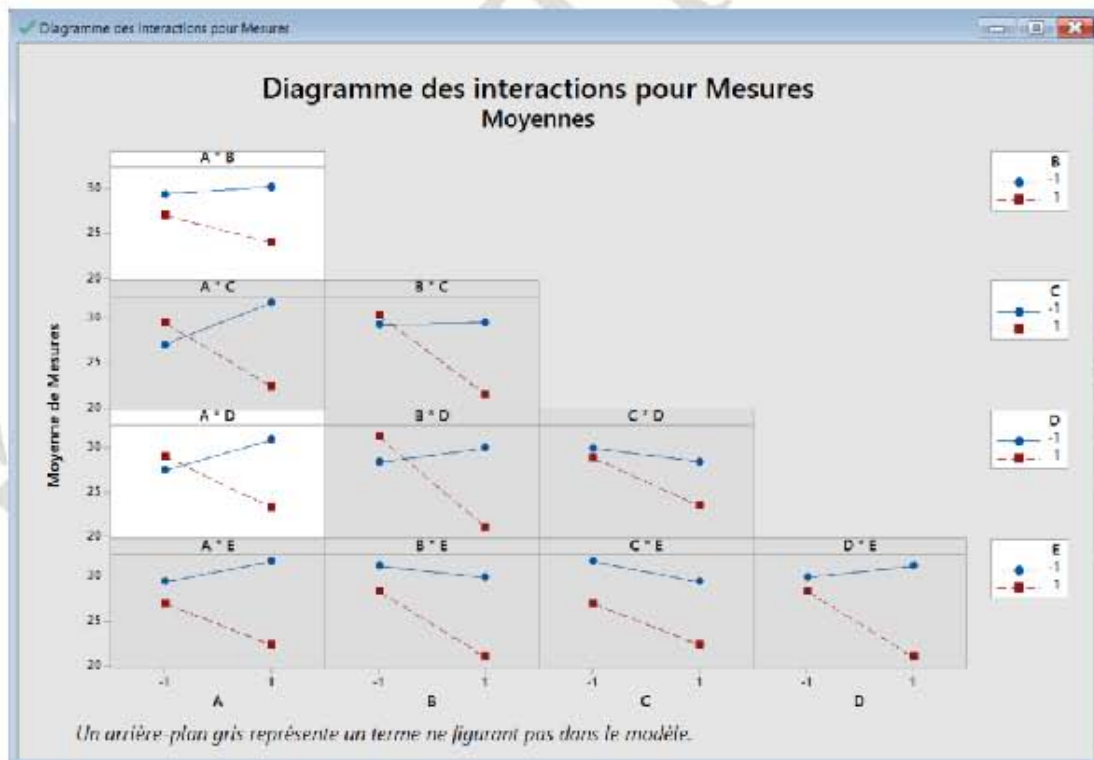
Nous avons alors:



Attention depuis **Minitab 16** pour avoir un graph similaire, toujours en passant par **Stat/DOE (plan d'expériences)/Plan factoriel/Diagrammes factoriels...** il faut faire attention à prendre l'option **Tous les termes**:



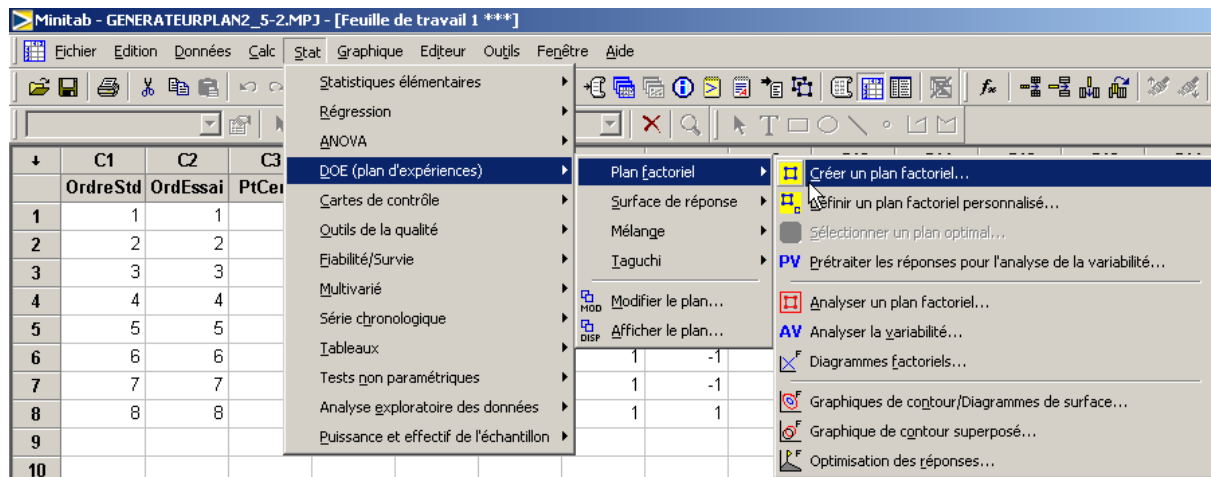
Ce qui donnera:



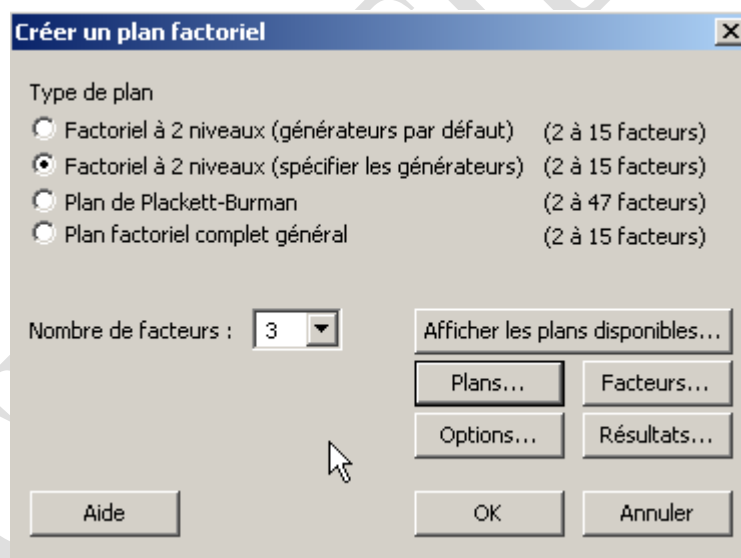
Donc au niveau des interactions ce n'est pas la joie (pour rappel, des courbes confondues traduisent une interaction nulle; des courbes très différentes traduisent une interaction forte). Il y en a un peu partout et elles ne sont donc pas négligeables comme nous l'a montré l'ANOVA.

### 17.9.4. Analyse du plan d'expérience désalié

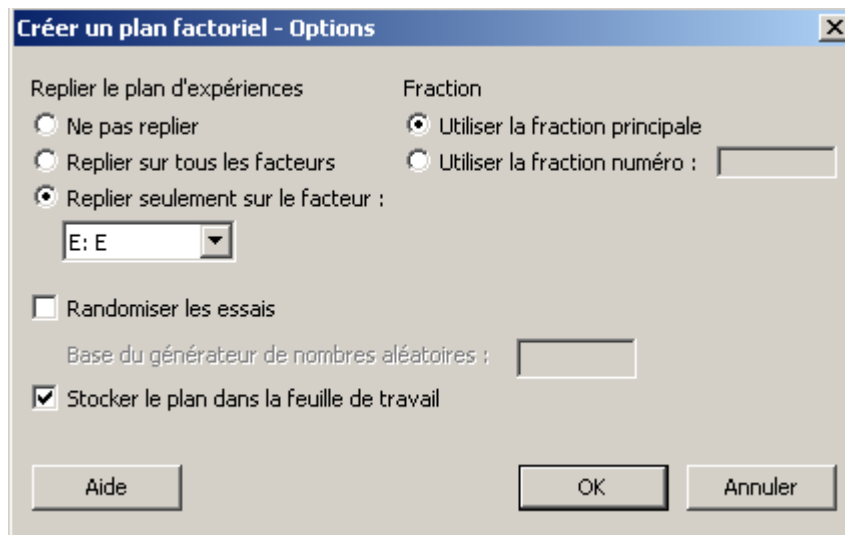
Il est donc nécessaire de faire des essais complémentaires. Pour cela nous allons reprendre le plan et nous allons désalié le facteur E. Pour ce faire, nous allons reprendre le plan en désalié le facteur E dans le menu **Stat/DOE (plan d'expériences)/Plan factoriel/Créer un plan factoriel...**:



Et nous repartons des mêmes paramètres qu'avant avec les mêmes générateurs:



Mais cette fois-ci, nous cliquons sur le bouton **Options...** et nous prenons:



et nous validons deux fois par **OK** et nous obtenons dans la fenêtre de session:

```
Alias initiales indépendantes (avant repli) : D = ABC; E = AC
```

```
Repliés sur les facteurs : E
```

```
Structure des alias (jusqu'à l'ordre 3)
```

```
|
I
```

```
A + BCD
B + ACD
C + ABD
D + ABC
E + ABCDE
AB + CD
AC + BD
AD + BC
AE + BCDE
BE + ACDE
CE + ABDE
DE + ABCE
ABE + CDE
ACE + BDE
ADE + BCE
ABCD
```

Donc forcément c'est un peu mieux au niveau des 4 facteurs principaux qui ne sont plus aliasés avec E excepté E lui-même. Et nous obtenons comme plan factoriel:

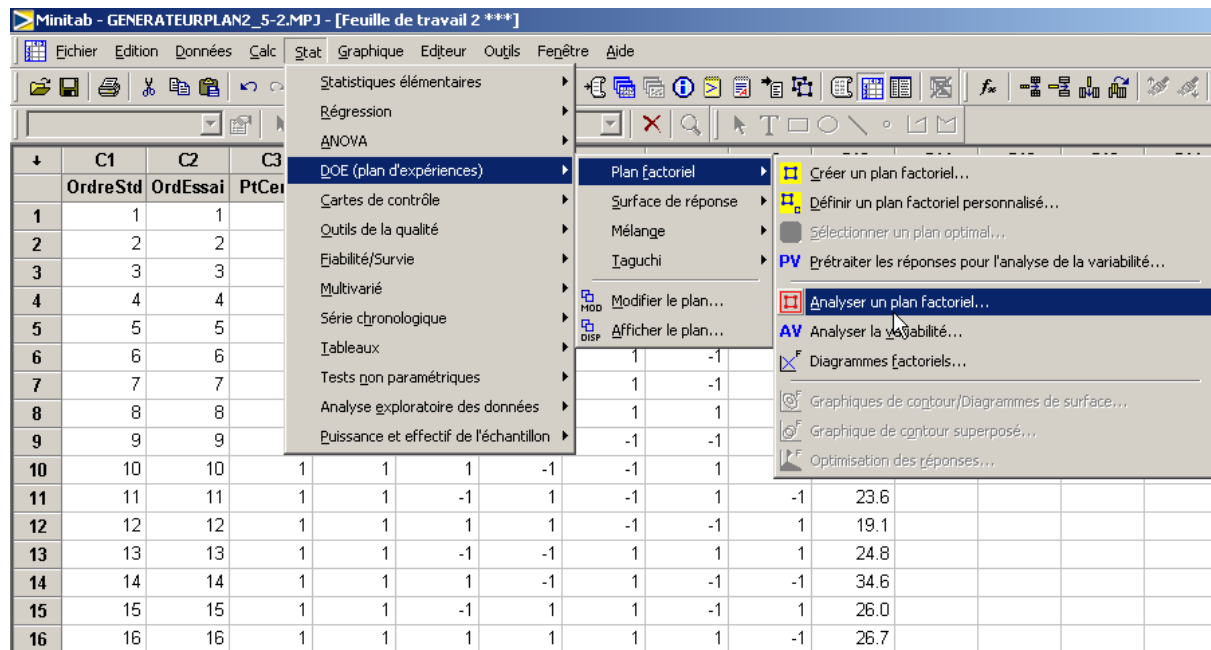
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	C	D	E	
1	1	1	1	1	-1	-1	-1	-1	1	
2	2	2	1	1	1	-1	-1	1	-1	
3	3	3	1	1	-1	1	-1	1	1	
4	4	4	1	1	1	1	-1	-1	-1	
5	5	5	1	1	-1	-1	1	1	-1	
6	6	6	1	1	1	-1	1	-1	1	
7	7	7	1	1	-1	1	1	-1	-1	
8	8	8	1	1	1	1	1	1	1	
9	9	9	1	1	-1	-1	-1	-1	-1	
10	10	10	1	1	1	-1	-1	1	1	
11	11	11	1	1	-1	1	-1	1	-1	
12	12	12	1	1	1	1	-1	-1	1	
13	13	13	1	1	-1	-1	1	1	1	
14	14	14	1	1	1	-1	1	-1	-1	
15	15	15	1	1	-1	1	1	-1	1	
16	16	16	1	1	1	1	1	1	-1	

Donc le repli impose bien évidemment plus d'essais (il faut alors se poser la question si c'est le meilleur choix...). Et donc nous ajoutons aux premiers essais, 8 autres. Ce qui donnera:

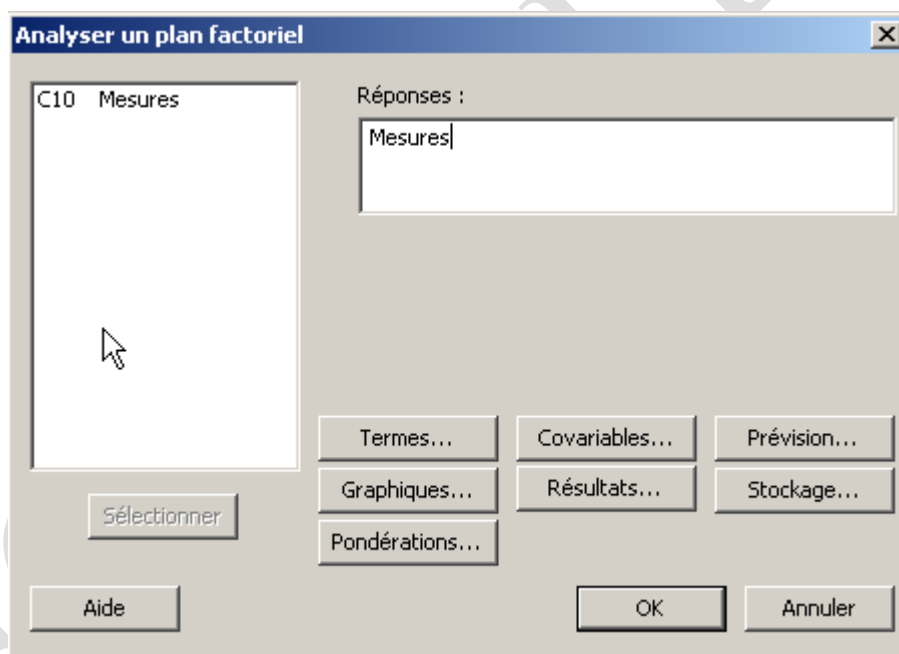
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	C	D	E	Mesures
1	1	1	1	1	-1	-1	-1	-1	1	27.4
2	2	2	1	1	1	-1	-1	1	-1	31.1
3	3	3	1	1	-1	1	-1	1	1	26.6
4	4	4	1	1	1	1	-1	-1	-1	32.4
5	5	5	1	1	-1	-1	1	1	-1	31.4
6	6	6	1	1	1	-1	1	-1	1	29.3
7	7	7	1	1	-1	1	1	-1	-1	27.5
8	8	8	1	1	1	1	1	1	1	15.5
9	9	9	1	1	-1	-1	-1	-1	-1	27.0
10	10	10	1	1	1	-1	-1	1	1	17.0
11	11	11	1	1	-1	1	-1	1	-1	23.6
12	12	12	1	1	1	1	-1	-1	1	19.1
13	13	13	1	1	-1	-1	1	1	1	24.8
14	14	14	1	1	1	-1	1	-1	-1	34.6
15	15	15	1	1	-1	1	1	-1	1	26.0
16	16	16	1	1	1	1	1	1	-1	26.7

Ensuite, nous analysons le plan factoriel en allant dans le menu **Stat/DOE (plan d'expérience)/Plan factoriel/Analyser un plan factoriel...**:





Pour prendre:



et nous validons par **OK** pour obtenir dans la fenêtre de session:

Informations sur les alias pour les termes du modèle.  
Les termes totalement confondus ont été supprimés de l'analyse.

I + A\*B\*C\*D

### Plan factoriel : Mesures en fonction de A; B; C; D; E

Effets et coefficients estimés pour Mesures (unités codées)

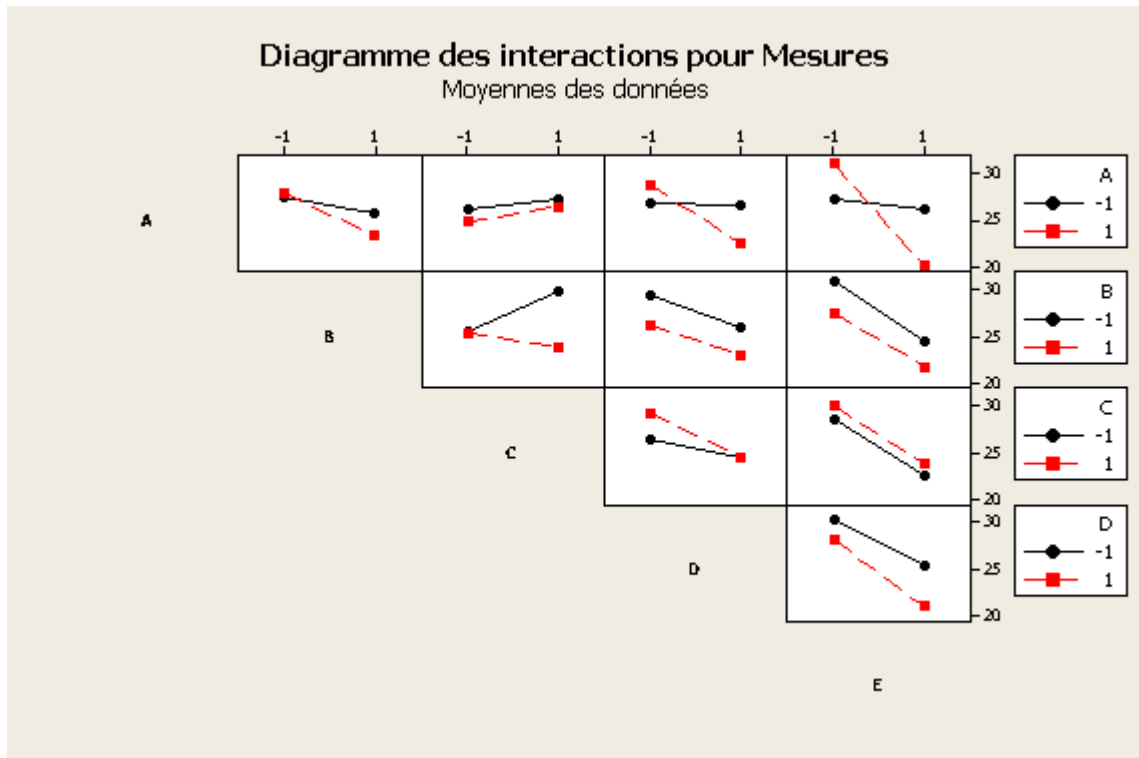
Terme	Effet	Coeff
Constante		26.250
A	-1.075	-0.537
B	-3.150	-1.575
C	1.450	0.725
D	-3.325	-1.662
E	-6.075	-3.037
A*B	-1.425	-0.713
A*C	0.175	0.087
A*D	-2.950	-1.475
A*E	-4.900	-2.450
B*E	0.325	0.162
C*E	-0.075	-0.037
D*E	-1.150	-0.575
A*B*E	-1.600	-0.800
A*C*E	2.800	1.400
A*D*E	-0.525	-0.263

S = \* SomCar-ErrPrév = \*

Analyse de la variance pour Mesures (unités codées)

Source	DL	SomCar séq	SomCar ajust	CM ajust	F	P
Effets principaux	5	244.57	244.57	48.91	*	*
2 Interactions - Nombre de facteurs	7	144.83	144.83	20.69	*	*
3 Interactions - Nombre de facteurs	3	42.70	42.70	14.23	*	*
Erreur résiduelle	0	*	*	*		
Total	15	432.10				

et le diagramme des interactions donne:



Et donc nous sommes dans une situation un peu plus confortable pour tirer les conclusions nécessaires à l'expérience.

## 17.10. Exercice 181.: Plan factoriel avec points centraux

Minitab® Statistical Software 15.1.1.0

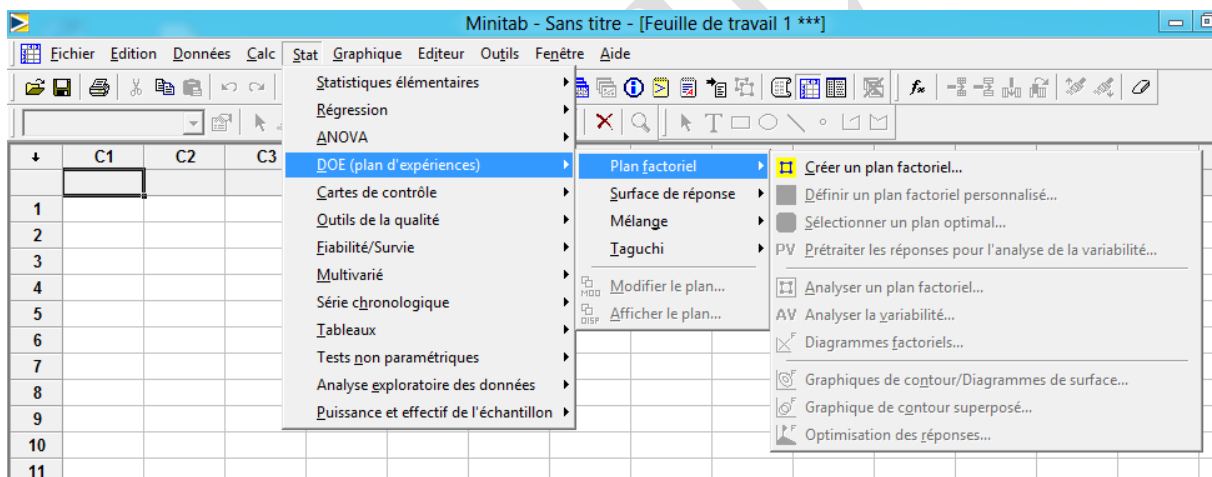
Une manière assez robuste d'accepter ou de rejeter le modèle linéaire supposé jusqu'à maintenant est d'utiliser des points centraux (sinon quoi on passera aux plans de surface de réponse). Voyons comment.

Considérons un plan basé sur deux facteurs qui sont respectivement:

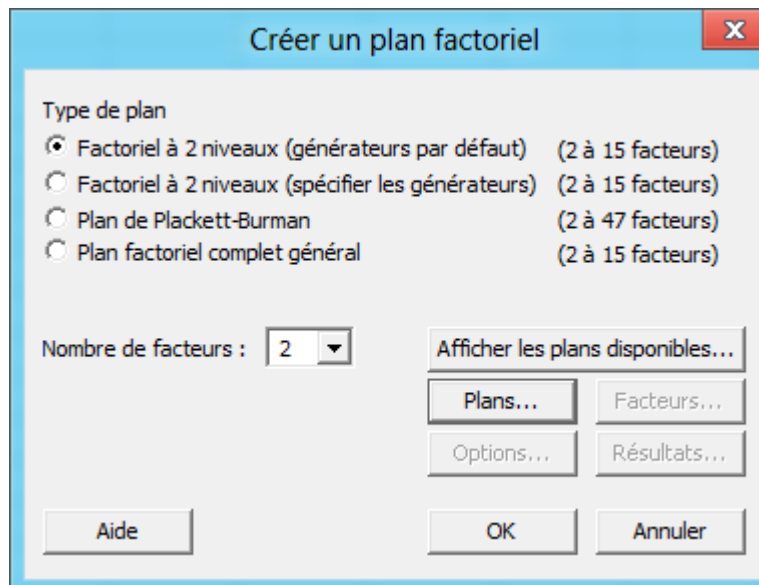
- La vitesse d'avancement  $A$  de l'outil d'usinage en mètres par minute à 2 niveaux
- La vitesse de coupe  $B$  de l'outil en mètres par seconde à 2 niveaux.

On mesure la rugosité et le nombre de pics par unité de longueur. Le but étant d'avoir un modèle (si possible...) qui permet de déterminer la vitesse d'avancement et la vitesse de coupe afin de minimiser ces deux paramètres sachant que les deux facteurs interagissent et que nous avons mesuré deux points centraux pour le plan d'expérience.

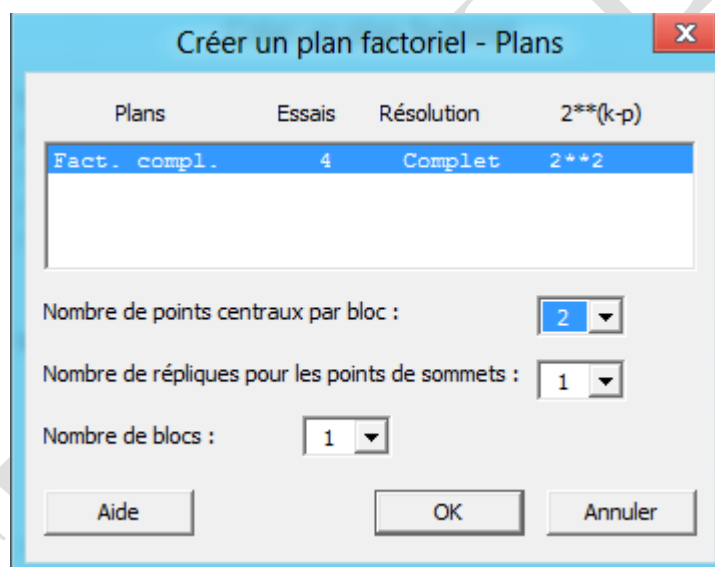
Nous allons donc dans le menu **DOE (plan d'expérience)/Plan factoriel/Créer un plan factoriel...**:



Nous prenons un plan factoriel à 2 niveaux avec les générateurs par défaut:



Nous cliquons sur le bouton **Plans...** pour bien spécifier le **Nombre de points centraux par bloc**:



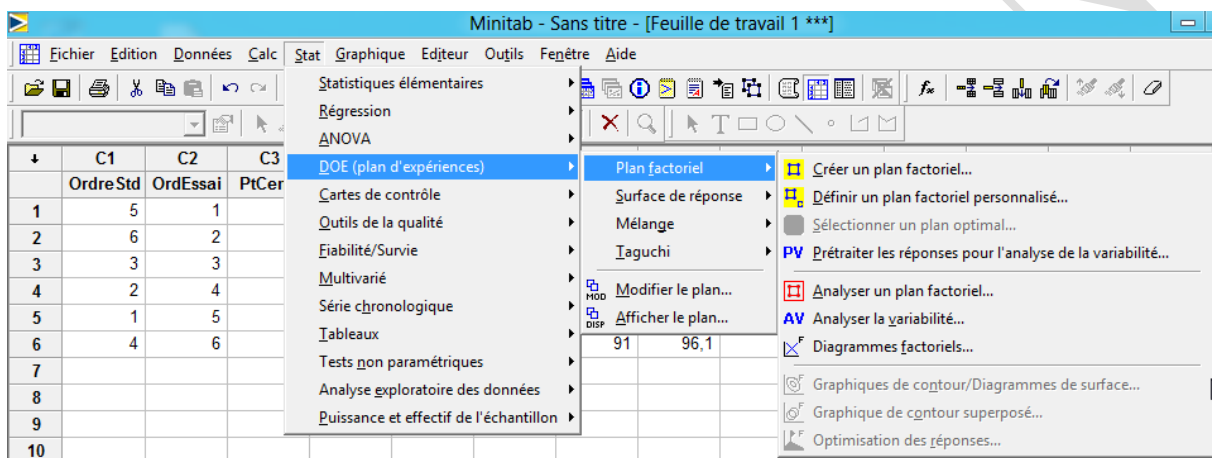
Nous validons le tout par **OK** pour obtenir:

↓	C1	C2	C3	C4	C5	C6	C7
	Ordre Std	OrdEssai	PtCentr	Blocs	A	B	
1	5	1	0	1	0	0	
2	6	2	0	1	0	0	
3	3	3	1	1	-1	1	
4	2	4	1	1	1	-1	
5	1	5	1	1	-1	-1	
6	4	6	1	1	1	1	
7							

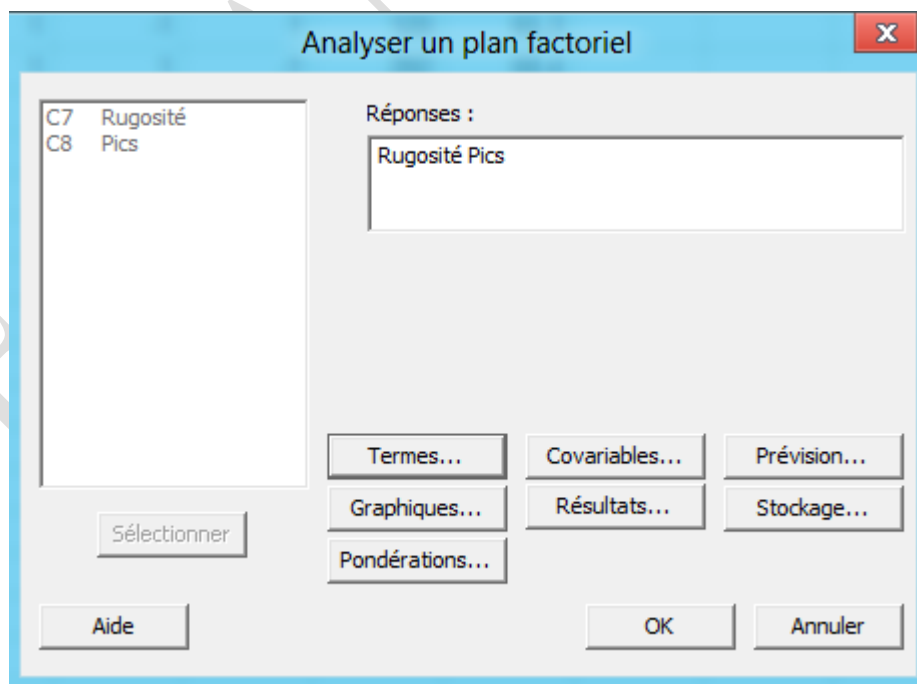
Nous y mettons les mesures:

↓	C1	C2	C3	C4	C5	C6	C7	C8	C9
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	Rugosité	Pics	
1	5	1	0	1	0	0	233	63,8	
2	6	2	0	1	0	0	235	61,9	
3	3	3	1	1	-1	1	120	65,3	
4	2	4	1	1	1	-1	282	68,4	
5	1	5	1	1	-1	-1	194	77,8	
6	4	6	1	1	1	1	91	96,1	
7									

Nous lançons l'analyse en allant dans le menu **DOE (plan d'expérience)/Plan factoriel/Analyse un plan factoriel...**:



Nous prenons:



Et validons le tout par **OK** pour nous intéresser uniquement aux deux sorties suivantes de la fenêtre de session:

**Plan factoriel : Rugosité en fonction de A; B**

Effets et coefficients estimés pour Rugosité (unités codées)

Terme	Effet	Coeff	Coef ErT	T	P
Constante		171,75	0,7071	242,89	0,003
A	29,50	14,75	0,7071	20,86	0,030
B	-132,50	-66,25	0,7071	-93,69	0,007
A*B	-58,50	-29,25	0,7071	-41,37	0,015
Pt ctr		62,25	1,2247	50,83	0,013

**Plan factoriel : Pics en fonction de A; B**

Effets et coefficients estimés pour Pics (unités codées)

Terme	Effet	Coeff	Coef ErT	T	P
Constante		76,90	0,6718	114,48	0,006
A	10,70	5,35	0,6718	7,96	0,080
B	7,60	3,80	0,6718	5,66	0,111
A*B	20,10	10,05	0,6718	14,96	0,042
Pt ctr		-14,05	1,1635	-12,08	0,053

Avant d'aller plus loin dans l'analyse comparons les points centraux par exemple pour la rugosité. Nous avons donc expérimentalement:

$$y_0(0,0) = \frac{233 + 235}{2} = 234$$

Alors que la théorie nous donne un coefficient:

$$a_0 = 171.75$$

La différence entre les deux donne:

$$y_0(0,0) - a_0 = 234 - 171.75 = 62.25$$

Valeur que l'on retrouve dans la table renvoyée par Minitab et mise en évidence en rouge ci-dessus.

Au vu de la différence qui est très significative, il va de soi que le modèle linéaire est à rejeter.

## 17.11. Exercice 182.: Analyse du modèle d'un plan factoriel personnalisé avec covariable

Minitab® Statistical Software 18.1

Rappelons le résultat que nous avons obtenu lors de l'analyse de deux groupes test à une méthode d'enseignement (voir page 471):

	Méthode A			Méthode B	
Sujet	X <sub>a</sub>	Y <sub>a</sub>	Sujets	X <sub>b</sub>	Y <sub>b</sub>
a1	5	20	b1	7	19
a2	10	23	b2	12	26
a3	12	30	b3	27	33
a4	9	25	b4	24	35
a5	23	34	b5	18	30
a6	21	40	b6	22	31
a7	14	27	b7	26	34
a8	18	38	b8	21	28
a9	6	24	b9	14	23
a10	13	31	b10	9	22
Moyenne	13.1	29.2		18.0	28.1

et nous avons obtenu:

### Modèle linéaire général : After en fonction de Method

```
Facteur Type Niveaux Valeurs
Method fixe 2 A; B
```

Analyse de la variance pour After, avec utilisation de la somme des carrés ajustée pour les tests

Source	DL	SomCar séq	SomCar ajust	CM ajust	F	P
Method	1	6.05	115.31	115.31	16.03	0.001
Before	1	540.18	540.18	540.18	75.07	0.000
Erreur	17	122.32	122.32	7.20		
Total	19	668.55				

S = 2.68241 R carré = 81.70 % R carré (ajust) = 79.55 %

Terme	Coeff	Coef ErT	T	P
Constante	15.783	1.602	9.85	0.000
Before	0.82748	0.09550	8.66	0.000

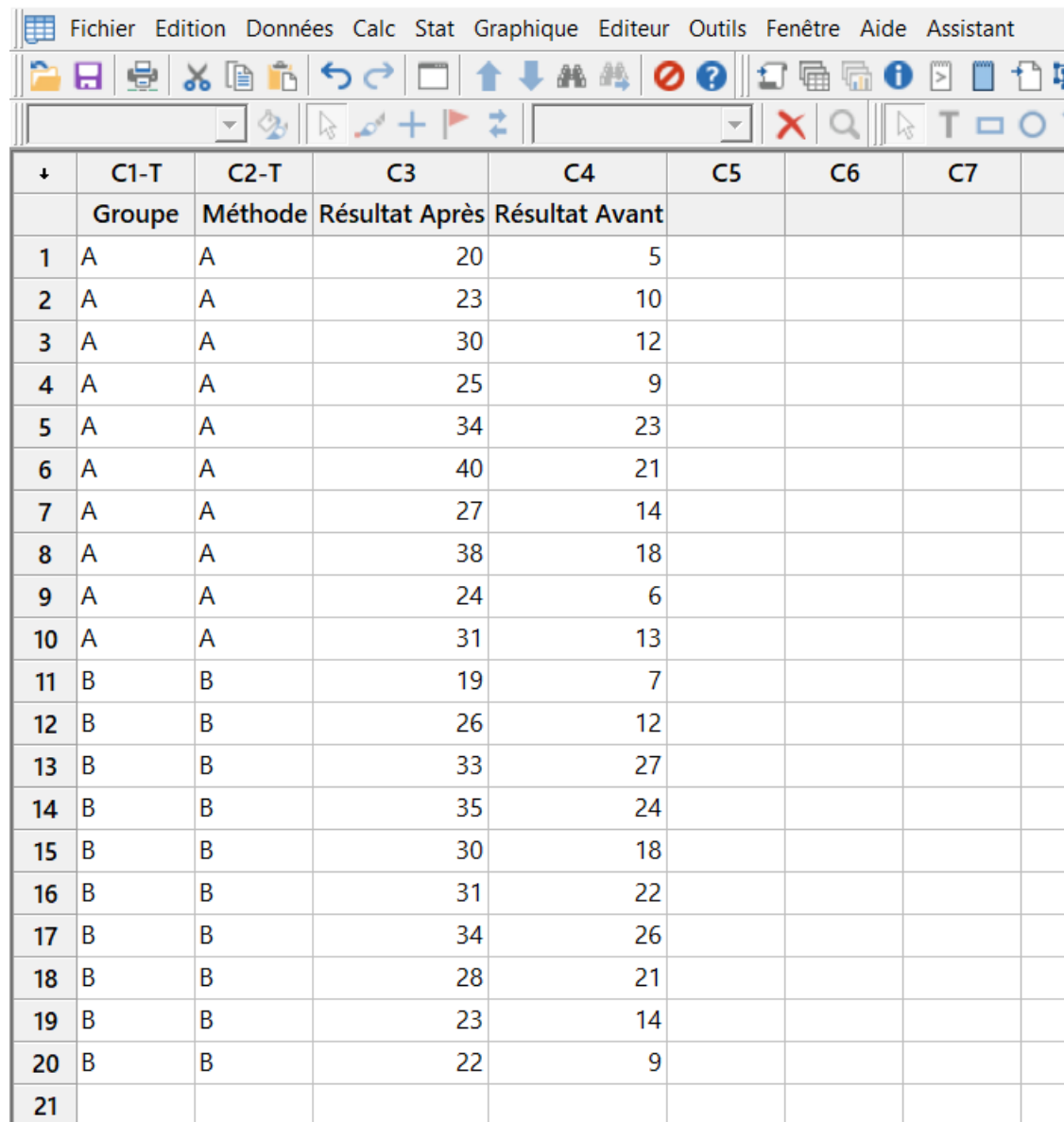


Maintenant, nous souhaiterions déterminer quel est le meilleur groupe et la meilleure méthode en utilisant un plan d'expérience.

Pour cela nous préparons la feuille suivante:

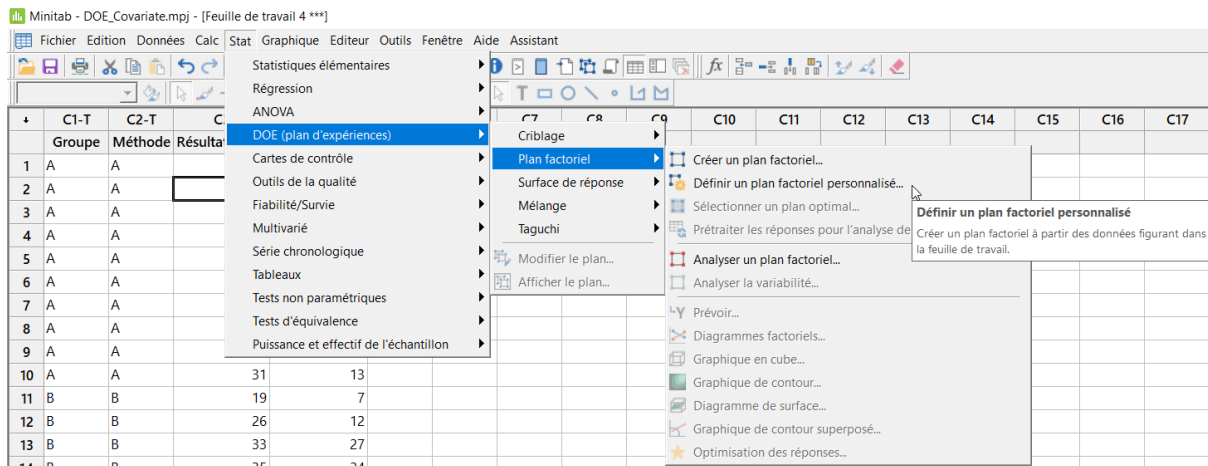
Minitab - DOE\_Covariate.mpj - [Feuille de travail 4 \*\*\*]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assistant

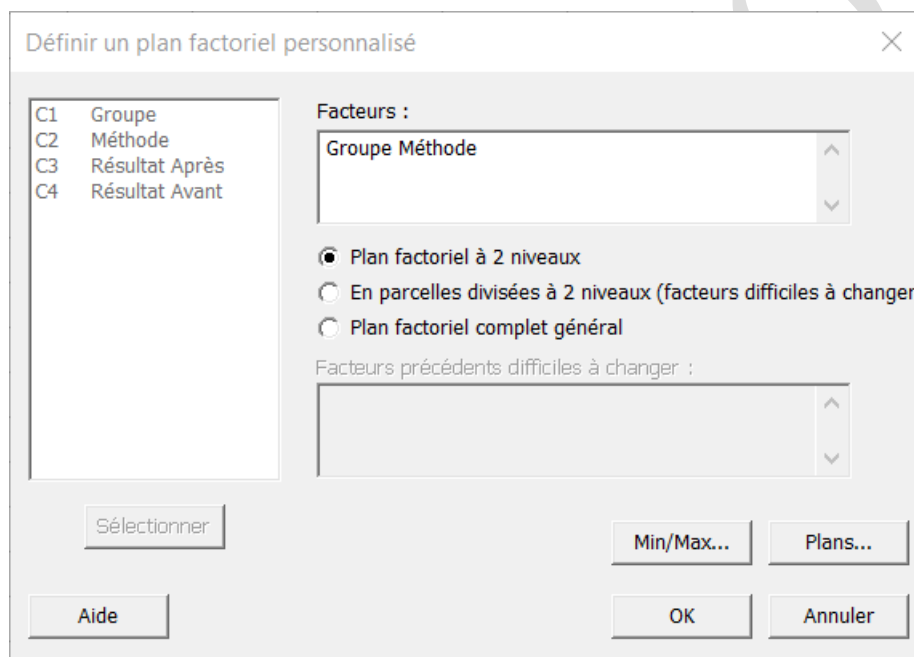


↓	C1-T	C2-T	C3	C4	C5	C6	C7
	Groupe	Méthode	Résultat Après	Résultat Avant			
1	A	A	20	5			
2	A	A	23	10			
3	A	A	30	12			
4	A	A	25	9			
5	A	A	34	23			
6	A	A	40	21			
7	A	A	27	14			
8	A	A	38	18			
9	A	A	24	6			
10	A	A	31	13			
11	B	B	19	7			
12	B	B	26	12			
13	B	B	33	27			
14	B	B	35	24			
15	B	B	30	18			
16	B	B	31	22			
17	B	B	34	26			
18	B	B	28	21			
19	B	B	23	14			
20	B	B	22	9			
21							

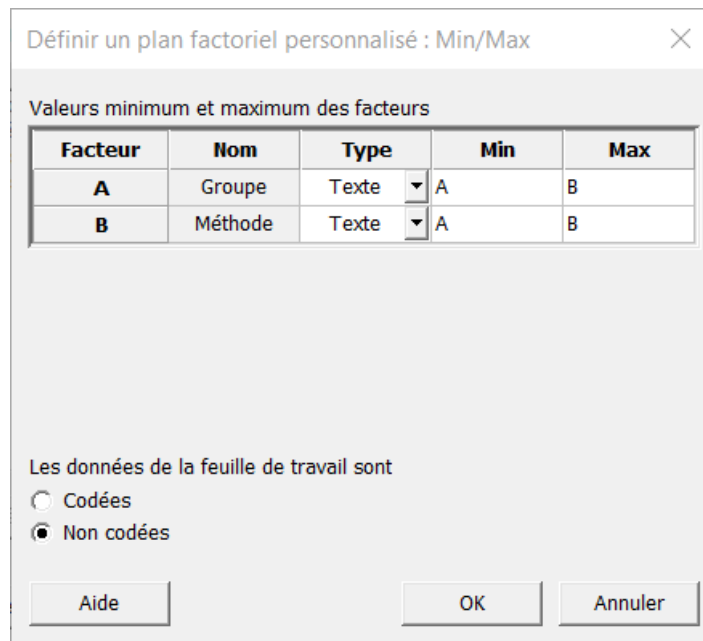
Ensuite, nous allons dans le menu **Stat/DOE (plan d'expérience)/Plan factoriel/Définir un plan factoriel personnalisé ...** car aucun plan standard ne satisfait bien évidemment à ce type de configuration:



Ensuite nous mettons:



et nous cliquons sur **Min/Max** pour mettre:



et nous validons par **OK** deux fois pour obtenir (nous prenons **Non codées** car des variables catégorielles ne peuvent par construction pas être codées au sens des plans d'expérience):

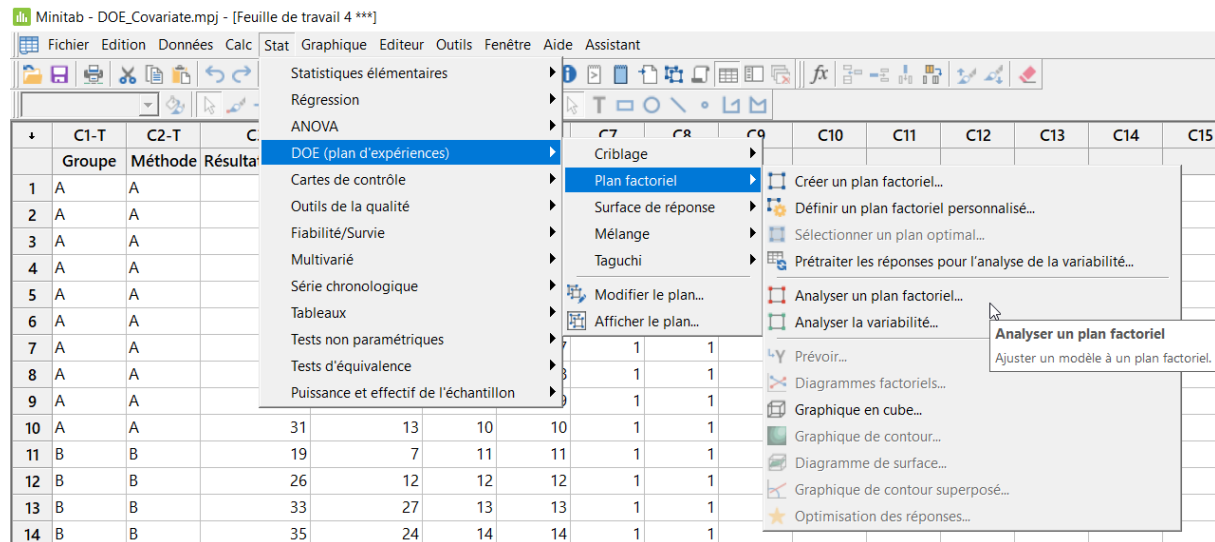
Minitab - DOE\_Covariate.mpj - [Feuille de travail 4 \*\*\*]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assistant

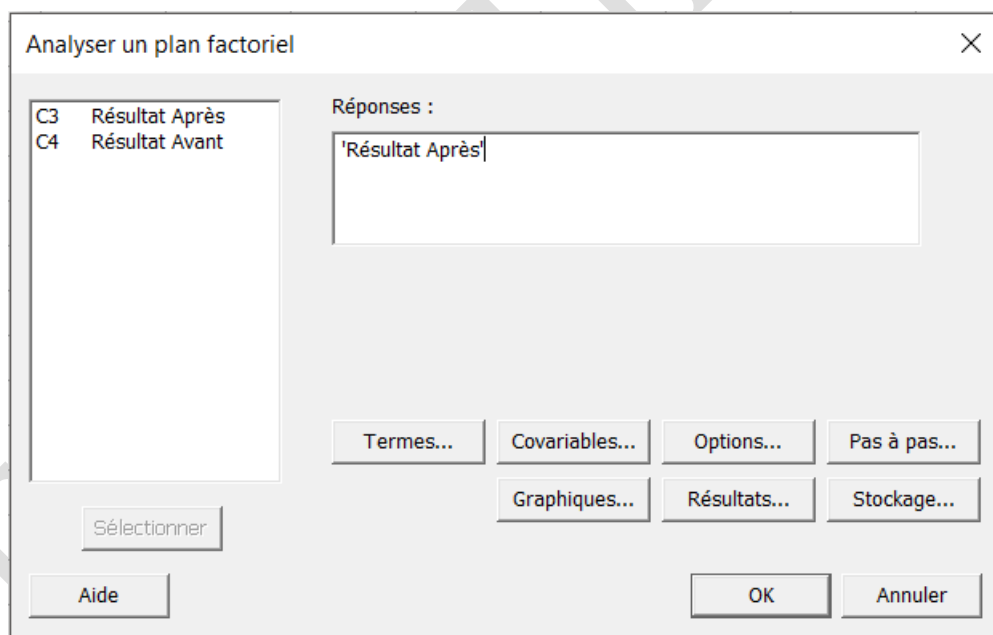
	C1-T	C2-T	C3	C4	C5	C6	C7	C8
	Groupe	Méthode	Résultat Après	Résultat Avant	OrdreStd	OrdEssai	Blocs	PtCentr
1	A	A	20	5	1	1	1	1
2	A	A	23	10	2	2	1	1
3	A	A	30	12	3	3	1	1
4	A	A	25	9	4	4	1	1
5	A	A	34	23	5	5	1	1
6	A	A	40	21	6	6	1	1
7	A	A	27	14	7	7	1	1
8	A	A	38	18	8	8	1	1
9	A	A	24	6	9	9	1	1
10	A	A	31	13	10	10	1	1
11	B	B	19	7	11	11	1	1
12	B	B	26	12	12	12	1	1
13	B	B	33	27	13	13	1	1
14	B	B	35	24	14	14	1	1
15	B	B	30	18	15	15	1	1
16	B	B	31	22	16	16	1	1
17	B	B	34	26	17	17	1	1
18	B	B	28	21	18	18	1	1
19	B	B	23	14	19	19	1	1
20	B	B	22	9	20	20	1	1

Il est intéressant de noter que comme il s'agit de variables non codées au sens des plans d'expérience, Minitab les mets simplement en tant que point centraux!

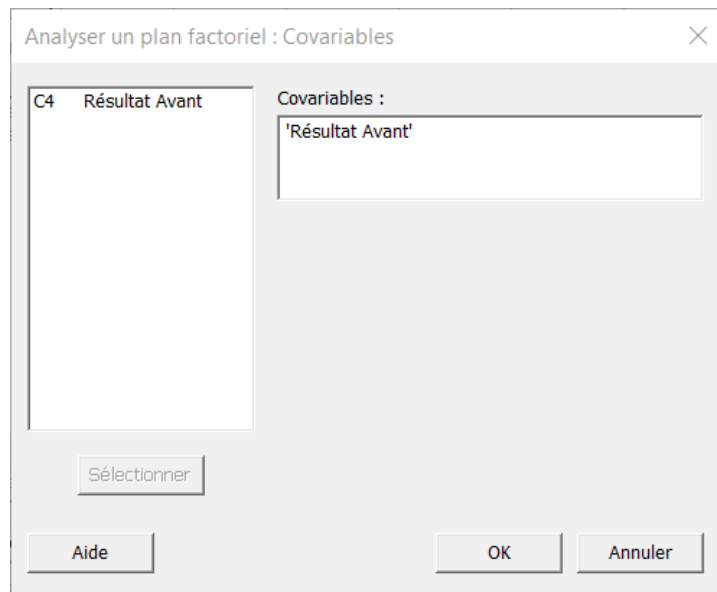
Maintenant lançons l'analyse en allant dans **Stat/DOE (plan d'expérience)/Plan factoriel/Analyser un plan factoriel....**:



Nous prenons:



Nous cliquons sur **Covariables** et y mettons (notez que nous n'aurions pas pu mettre ce facteur en tant que covariable si nous l'avions mis dans **Réponses** ci-dessus):



Nous avons alors (comparez les résultats avec l'ANCOVA standard):

Minitab - DOE\_Covariate.mpj - [Session]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assistant

Régression factorielle : Résultat Après en fonction de ... upe, Méthode

\* REMARQUE \* Ce plan n'est pas orthogonal.

Les termes suivants sont totalement confondus avec d'autres termes et ont été supprimés :  
Méthode, Groupe\*Méthode

**Analyse de variance**

Source	DL	SomCar		Valeur F	Valeur de p
		ajust	CM ajust		
Modèle	2	546.2	273.115	37.96	0.000
Covariables	1	540.2	540.180	75.07	0.000
Résultat Avant	1	540.2	540.180	75.07	0.000
Linéaires	1	115.3	115.306	16.03	0.001
Groupe	1	115.3	115.306	16.03	0.001
Erreur	17	122.3	7.195		
Total	19	668.5			

**Récapitulatif du modèle**

S	R carré	R carré (ajust)	R carré (prév)
2.68241	81.70%	79.55%	74.62%

**Coefficients codés**

Terme	Effet	Coeff	Coef ErT	Valeur de T	Valeur de p	FIV
Constante		15.78	1.60	9.85	0.000	
Résultat Avant		0.8275	0.0955	8.66	0.000	1.15
Groupe	-5.155	-2.577	0.644	-4.00	0.001	1.15

**Equation de régression en unités non codées**

Résultat Après = 15.78 + 0.8275 Résultat Avant - 2.577 Groupe

**Structure des alias**

Facteur	Nom
A	Groupe
B	Méthode

Alias

I + AB  
A + B

Donc nous ici que nous sommes coincés car tout logiquement, la variable *Méthode* a été supprimée (il ne pouvait en garder qu'une)!

Nous voyons cependant que les *Groupe* ET le *Résultats avant* ont une influence significative.

Comme le *Groupe* est totalement corrélé à la *Méthode*, si nous cherchons à optimiser une des variables, nous optimiserons alors les deux à la fois!

Nous allons aborder plus tard la méthode qui permet dans Minitab de trouver les valeurs des variables optimales et le lecteur pourra alors vérifier que c'est la méthode A qui est la plus efficace.

## 17.12. Exercice 183.: Analyse du modèle d'un plan de Plackett-Burman

Minitab® Statistical Software 15.1.1.0

Le but ici va être de vérifier ce que nous avons vu dans le cours théorique. D'abord, il sera considéré comme évident que les plans PB sont intéressants qu'à partir de 4 facteurs et ce pour 12, 20, 24, 28, etc. selon la règle (pour rappel...):

Multiple de 4 mais pas puissance de 2

essais (le reste étant simplement des plans factoriels fractionnaires particuliers). Nous rappelons que les plans PB permettent de réduire un petit peu (4 facteurs: 12 essais) voir énormément (11 facteurs: 12 essais) le nombre de mesures.

Nous allons voir aussi que Minitab ne communique pas les alias des plans PB pour les raisons suivantes indiquées sur le site de Minitab:

<http://www.minitab.com/fr-FR/support/answers/answer.aspx?id=530>

### Alias structure for Plackett-Burman designs - ID 530

Revised: 7/18/2008

Applies to

- Minitab 16

#### Description

Does Minitab display the alias structure for Plackett-Burman designs?

#### Solution

Minitab does not display the alias structure for Plackett-Burman designs because it is usually very messy. In many Plackett-Burman designs, as shown in "Design and Analysis of Experiments" by Douglas C. Montgomery, each main effect is partially confounded with all 2-factor interactions not involving itself. These designs are not meant to model interactions. They are meant to be used as screening designs, where you have a large number of factors, and you want to identify which have important main effects.

When you analyze your design in Minitab, you will see a message that says: "\* NOTE \* There is partial confounding, no alias table was printed."

Reference:

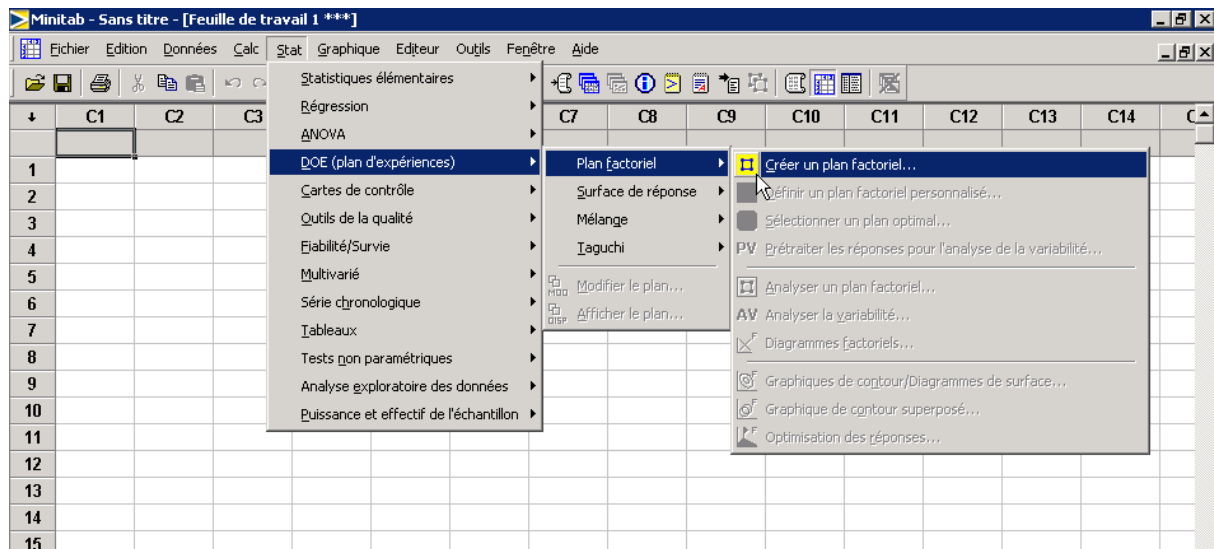
D. C. Montgomery (2004). *Design and Analysis of Experiments*, Sixth Edition. John Wiley & Sons, Inc.

Il est donc à noter que pour ce type de plans, les alias étant très conséquents, cela complique grandement l'analyse. Il est donc conseillé d'y réfléchir à deux fois avant d'utiliser un plan PB.

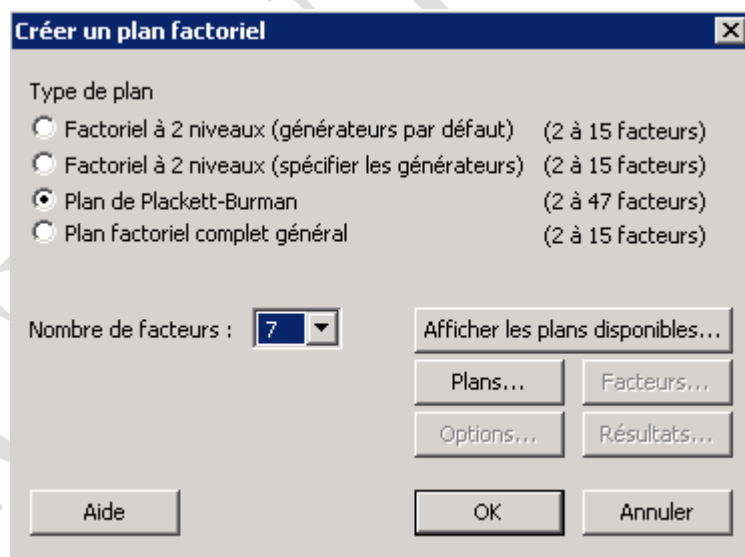
Pour l'exemple qui va suivre, nous allons exceptionnellement reprendre l'excellent exemple que Monsieur Jaques Goupy propose dans son magnifique ouvrage pédagogique sur les plans d'expérience consacré à JMP mais que nous allons mettre à la sauce Minitab...

Nous avons donc une machine dont les résultats souhaités (il y a deux résultats qui nous intéressent!) sont dépendants de 7 facteurs. Nous prendrons le plus petit plan de 12 essais afin de voir si nous pouvons en tirer quelque chose et au besoin nous ferons des expériences supplémentaires.

Nous allons donc dans le menu **Stat/DOE (plan d'expérience)/Plan factoriel/Créer un plan factoriel...**:



Nous prenons un plan de Plackett-Burman avec 7 facteurs:



Nous cliquons sur le bouton **Plans...**:



**Créer un plan factoriel - Plans**

Nombre d'essais : 12

Nombre de points centraux par réplique : 0

Nombre de répliques : 1

Créer des blocs pour les répliques

Aide OK Annuler

et nous laissons tout tel quel (12 essais étant le minimum de toute manière!).

Ensuite, nous désactivons la randomisation des essais en allant dans le bouton **Options...**:

**Créer un plan factoriel - Options**

Randomiser les essais

Base du générateur de nombres aléatoires :

Stocker le plan dans la feuille de travail

Aide OK Annuler

et nous validons le tout par **OK** et après avoir saisi les mesures des deux réponses d'intérêt, nous avons:

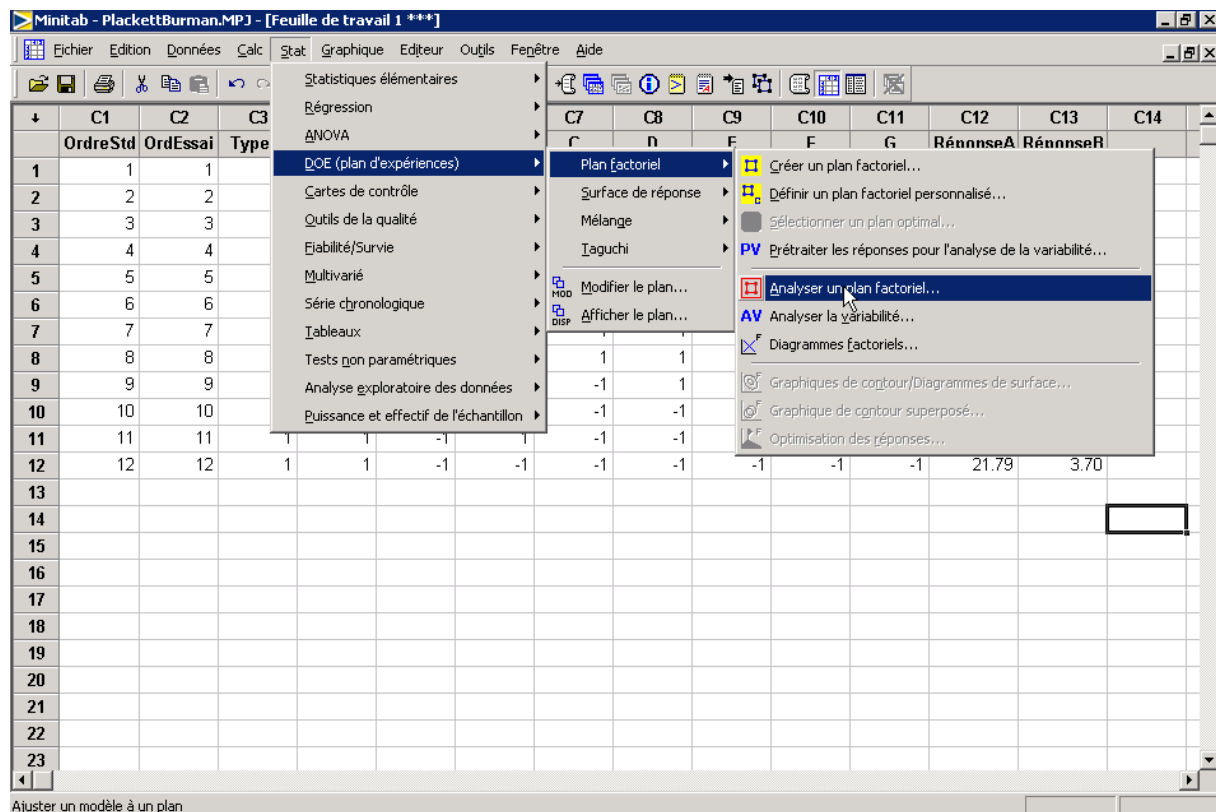
Minitab - PlackettBurman.MPJ - [Feuille de travail 1 \*\*\*\*]

Echier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

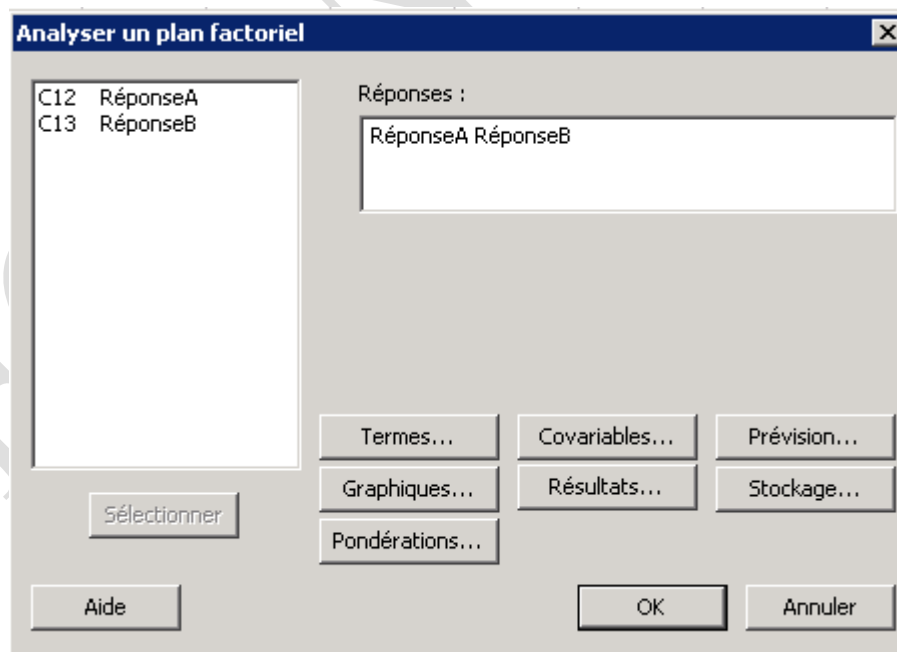
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	OrdreStd	OrdEssai	TypePt	Blocs	A	B	C	D	E	F	G	RéponseA	RéponseB	
1	1	1	1	1	1	-1	1	-1	-1	-1	1	24.24	4.40	
2	2	2	1	1	1	1	-1	1	-1	-1	-1	23.61	3.94	
3	3	3	1	1	-1	1	1	-1	1	-1	-1	24.02	3.58	
4	4	4	1	1	1	-1	1	1	-1	1	-1	23.36	3.90	
5	5	5	1	1	1	1	-1	1	1	-1	1	24.84	4.20	
6	6	6	1	1	1	1	1	-1	1	1	-1	22.87	4.00	
7	7	7	1	1	-1	1	1	1	-1	1	1	24.59	30.78	
8	8	8	1	1	-1	-1	1	1	1	-1	1	24.84	4.40	
9	9	9	1	1	-1	-1	-1	1	1	1	-1	23.13	4.06	
10	10	10	1	1	1	-1	-1	-1	1	1	1	24.64	3.57	
11	11	11	1	1	-1	1	-1	-1	-1	1	1	21.66	4.11	
12	12	12	1	1	-1	-1	-1	-1	-1	-1	-1	21.79	3.70	
13														
14														
15														
16														
17														
18														
19														
20														
21														
22														
23														

Feuille de travail en cours : Feuille de travail 1

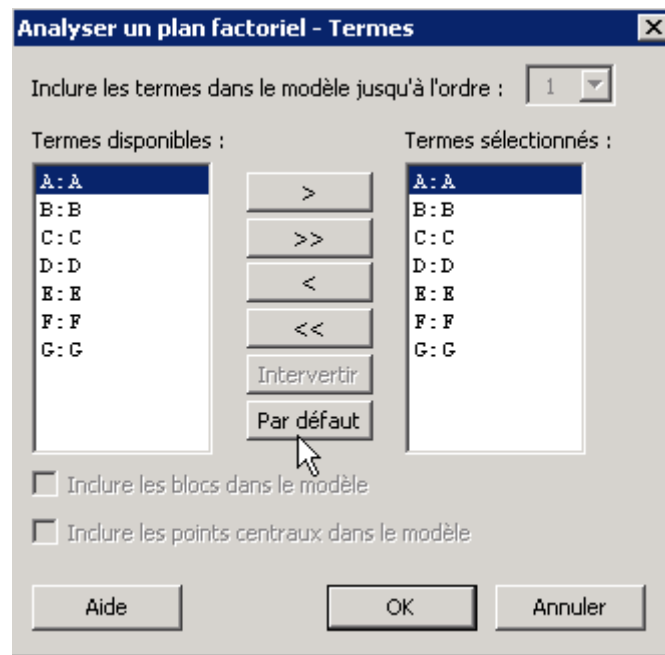
Nous allons ensuite dans le menu **Stat/DOE (plan d'expériences)/Plan factoriel/Analyser un plan factoriel...**:



Nous avons alors:



Si nous cliquons sur le bouton **Termes...**:



nous n'avons bien évidemment que le choix des facteurs principaux!

En validant par **OK** il vient alors (accrochez-vous...!) d'abord dans la fenêtre de session:

**Plan de Plackett-Burman**

Facteurs : 7 Répliques : 1  
 Essais de base : 12 Nombre total d'essais : 12  
 Blocs de base : 1 Nombre total de blocs : 1

**Plan factoriel : RéponseA; RéponseB****Plan factoriel : RéponseA en fonction de A; B; C; D; E; F; G**

Effets et coefficients estimés pour RéponseA (unités codées)

Terme	Effet	Coeff	Coef ErT	T	P
Constante		23.6325	0.2419	97.69	0.000
A	0.5883	0.2942	0.2419	1.22	0.291
B	-0.0683	-0.0342	0.2419	-0.14	0.895
C	0.7083	0.3542	0.2419	1.46	0.217
D	0.8583	0.4292	0.2419	1.77	0.151
E	0.8483	0.4242	0.2419	1.75	0.154
F	-0.5150	-0.2575	0.2419	-1.06	0.347

S = 0.837979

SomCar-ErrPrév = 25.2795

R carré = 79.29 %

R carré (prév) = 0.00 %

R carré (ajust) = 43.04 %

Analyse de la variance pour RéponseA (unités codées)

Source	DL	SomCar séq	SomCar		F	P
			ajust	CM ajust		
Effets principaux	7	10.753	10.753	1.5361	2.19	0.234
Erreur résiduelle	4	2.809	2.809	0.7022		
Total	11	13.561				

**Plan factoriel : RéponseB en fonction de A; B; C; D; E; F; G**

Effets et coefficients estimés pour RéponseB (unités codées)

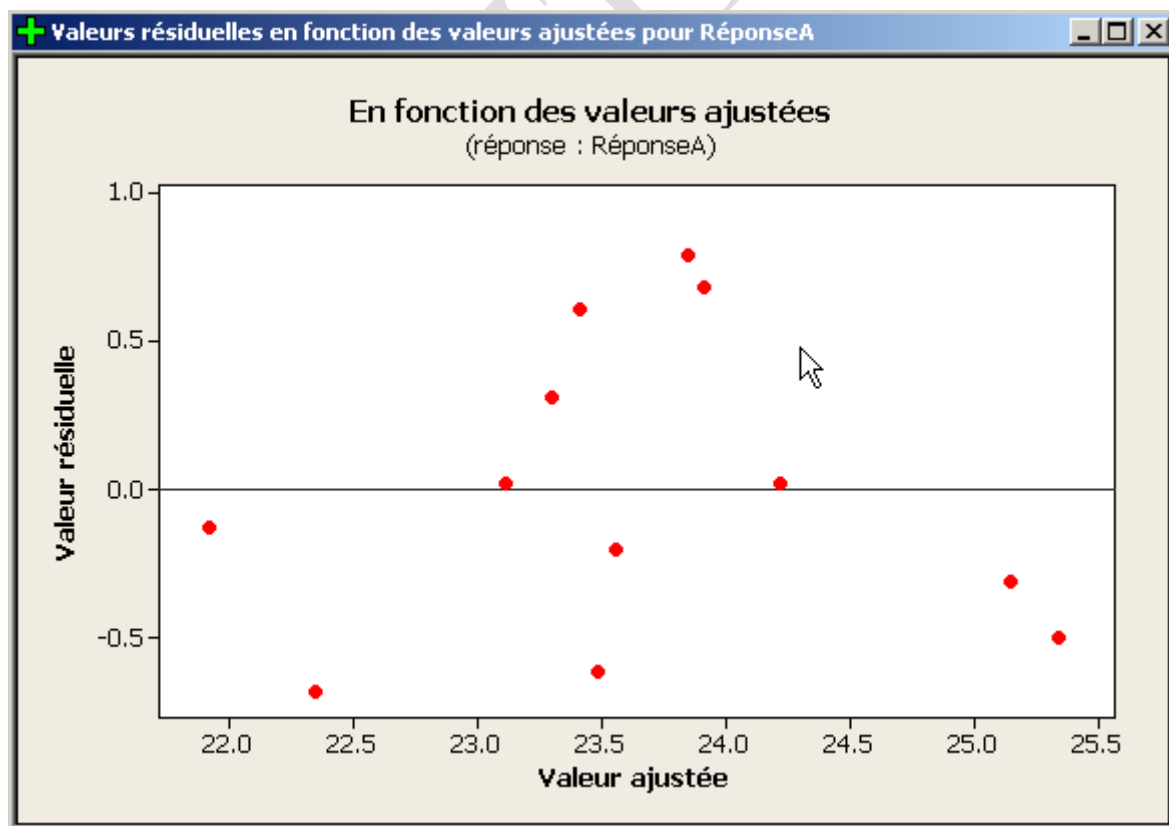
Terme	Effet	Coef	Coef ErT	T	P
Constante		6.220	2.181	2.85	0.046
A	-4.437	-2.218	2.181	-1.02	0.367
B	4.430	2.215	2.181	1.02	0.367
C	4.580	2.290	2.181	1.05	0.353
D	4.653	2.327	2.181	1.07	0.346
E	-4.503	-2.252	2.181	-1.03	0.360
F	4.367	2.183	2.181	1.00	0.373
G	4.713	2.357	2.181	1.08	0.341

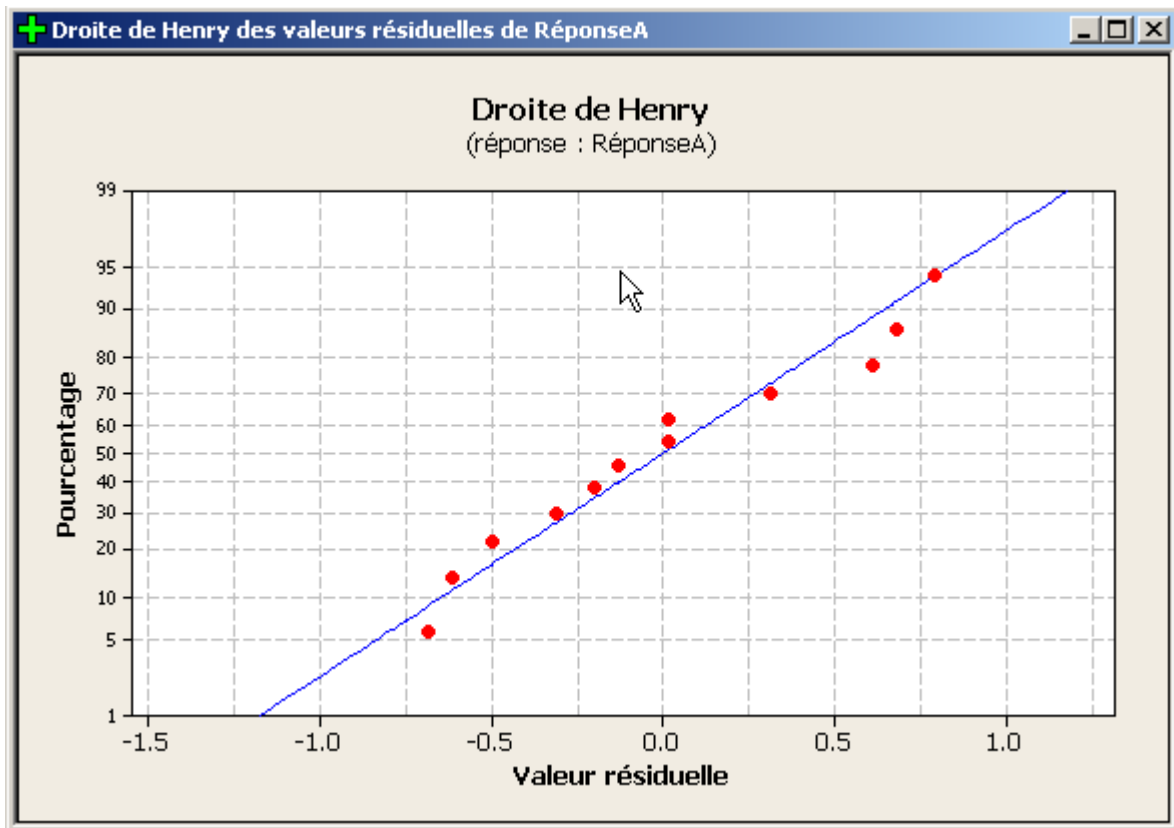
S = 7.55582      SomCar-ErrPrév = 2055.26  
R carré = 65.34 %      R carré (prév) = 0.00 %      R carré (ajust) = 4.69 %

Analyse de la variance pour RéponseB (unités codées)

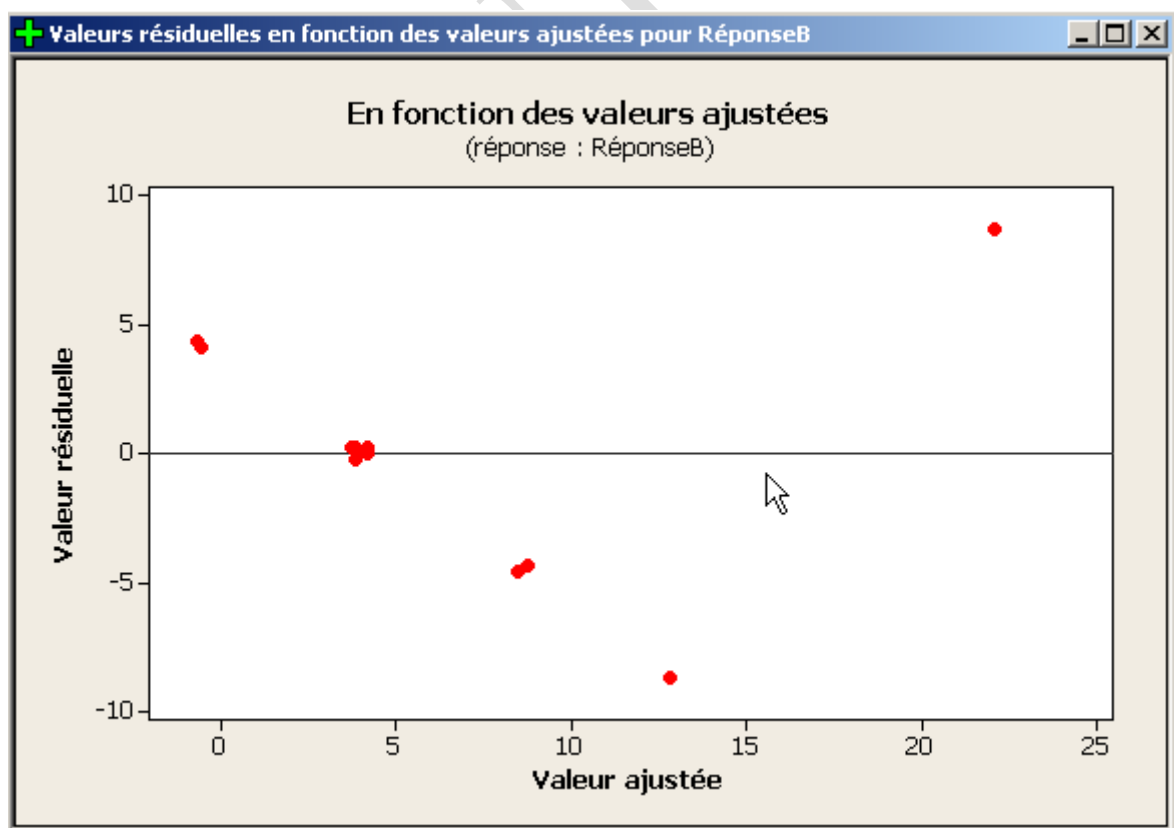
Source	DL	SomCar séq	SomCar ajust	CM ajust	F	P
Effets principaux	7	430.5	430.5	61.50	1.08	0.501
Erreur résiduelle	4	228.4	228.4	57.09		
Total	11	658.9				

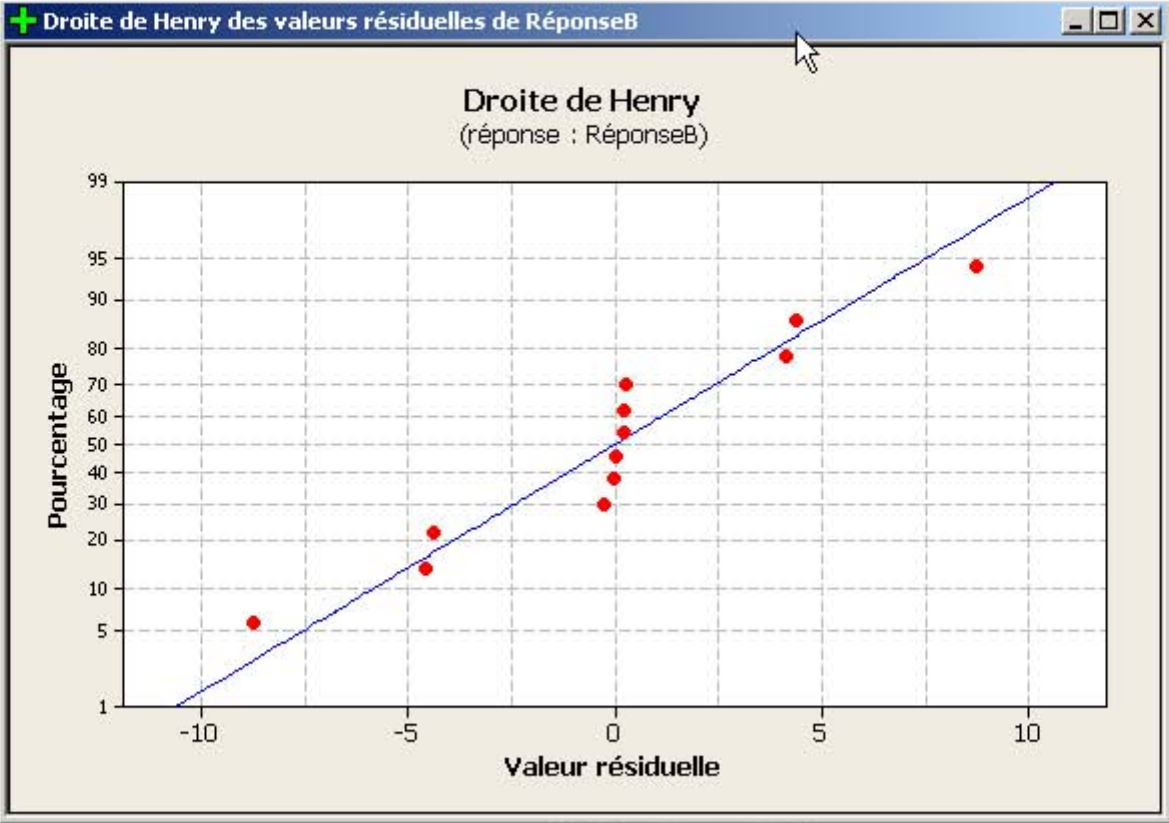
Pour les graphes, voici ceux de la réponse A:





et ceux de la réponse B:





ÉCHANTILLON

## 17.13. Exercice 184.: Analyse graphique avancée de base de plans factoriels

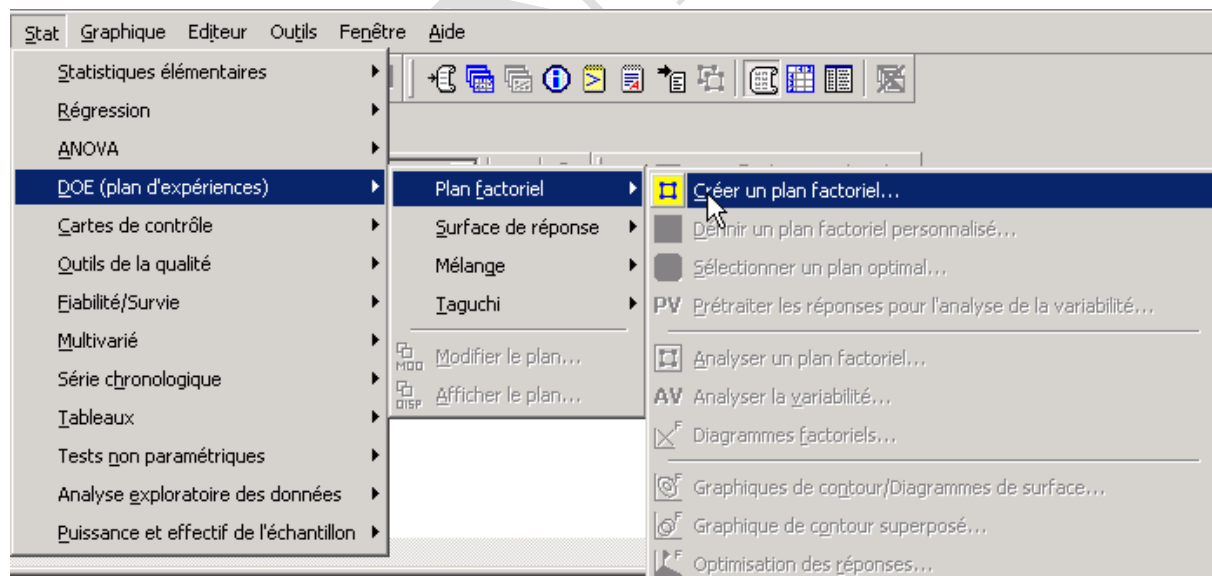
Minitab® Statistical Software 15.1.1.0

Ce que nous allons voir maintenant ne marche a priori à ce jour (version 15 de Minitab® Statistical Software) que si et seulement si nous avons des plans d'expériences factoriels complets ou fractionnaires à deux niveaux. Cela limite bien évidemment de manière très importante l'intérêt de l'outil.

Nous allons reprendre l'exemple vu dans le cours théorique mais simplifié pour avoir uniquement un plan factoriel complet à 2 niveaux:

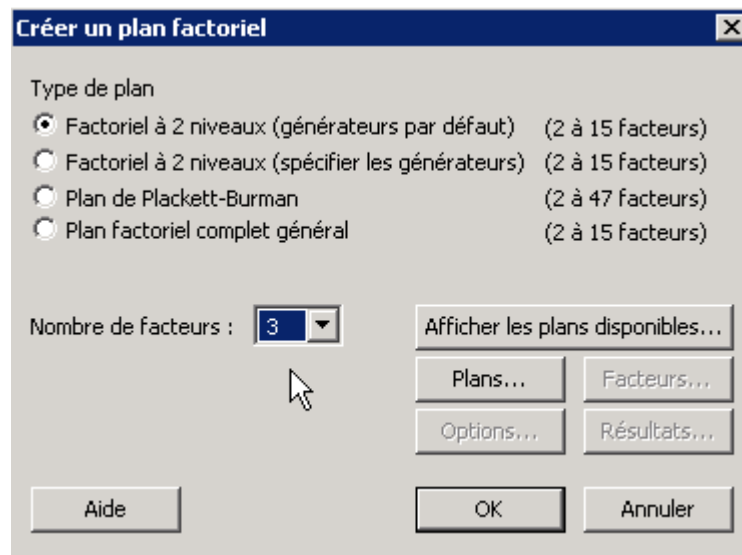
N° Essai	$x_1$	$x_2$	$x_3$	Valeurs	de	y
1	1	1	1	32'700	32'750	32'960
2	1	1	2	33'430	33'360	32'910
3	1	2	1	32'680	32'270	33'130
4	1	2	2	34'070	33'100	33'610
5	2	1	1	33'180	32'160	32'640
6	2	1	2	34'430	34'280	34'460
7	2	2	1	33'270	33'080	32'415
8	2	2	2	33'440	33'570	34'204

Nous créons alors le plan factoriel complet à 2 niveaux correspondant:

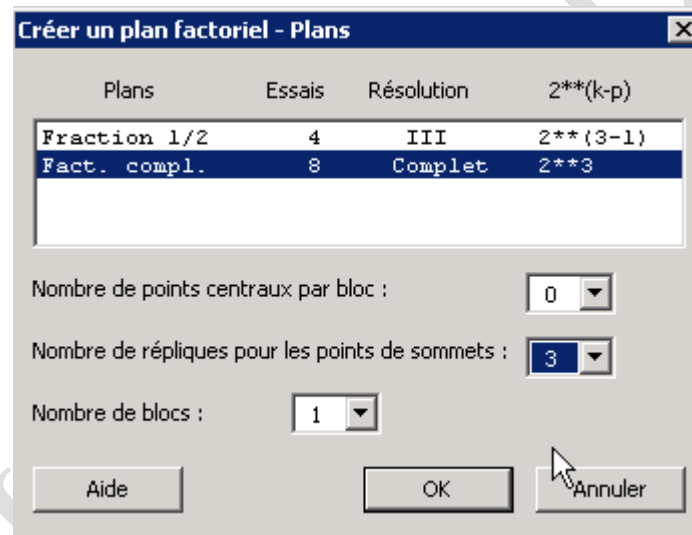


Nous sommes obligés de prendre l'option **Factoriel à 2 niveaux (générateurs par défaut)**:

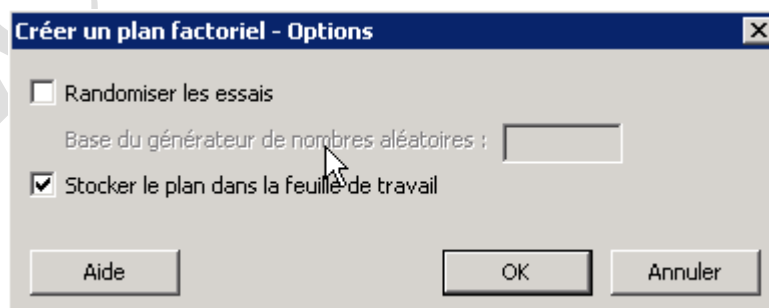




Nous cliquons sur le bouton **Plans...** et prenons un plan **Fact. compl.** avec **3** pour l'option **Nombre de répliques pour les points de sommets**:



Nous désactivons la randomisation:



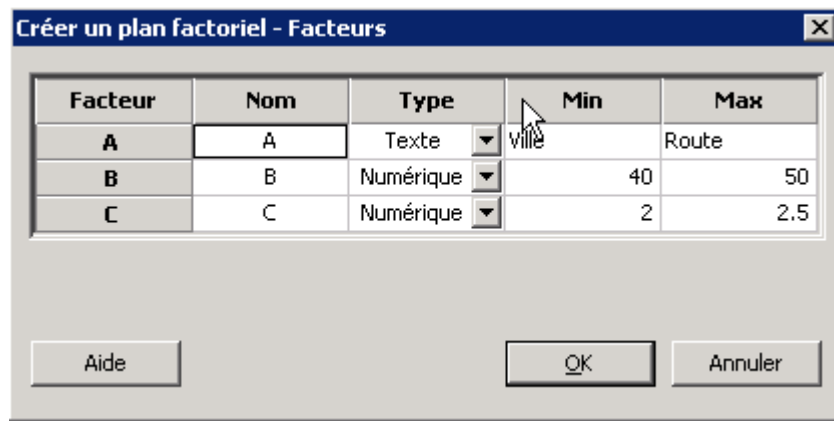
Nous avons alors:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	C								
1	1	1	1	1	-1	-1	-1								
2	2	2	1	1	1	1	-1								
3	3	3	1	1	-1	1	-1								
4	4	4	1	1	1	1	-1								
5	5	5	1	1	-1	-1	1								
6	6	6	1	1	1	1	1								
7	7	7	1	1	-1	1	1								
8	8	8	1	1	1	1	1								
9	9	9	1	1	-1	-1	-1								
10	10	10	1	1	1	-1	-1								
11	11	11	1	1	-1	1	-1								
12	12	12	1	1	1	1	-1								
13	13	13	1	1	-1	-1	1								
14	14	14	1	1	1	-1	1								
15	15	15	1	1	-1	1	1								
16	16	16	1	1	1	1	1								
17	17	17	1	1	-1	-1	-1								
18	18	18	1	1	1	-1	-1								
19	19	19	1	1	-1	1	-1								

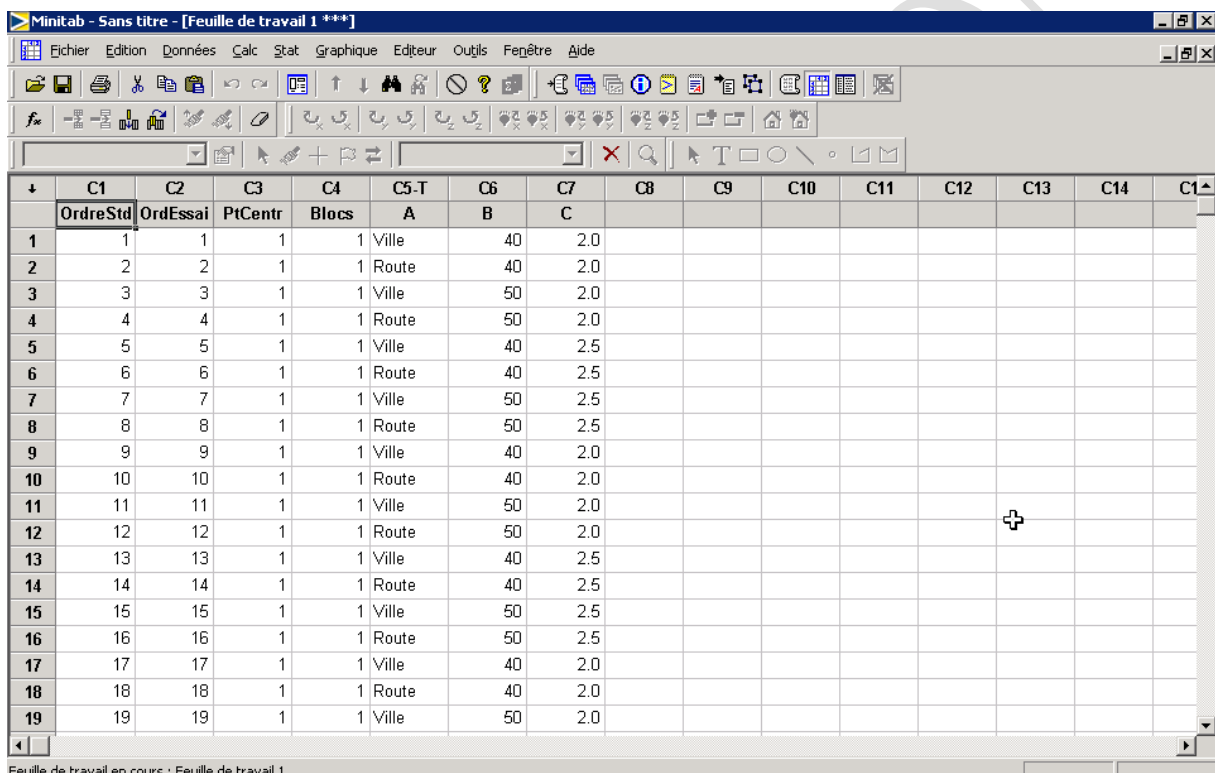
Nous remplissons toutes les réponses (vous pouvez aussi ouvrir le fichier *DOE\_ValeursCodees\_GraphesAvances*):

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	C	Réponses							
1	1	1	1	1	-1	-1	-1	32700							
2	2	2	1	1	1	-1	-1	33180							
3	3	3	1	1	-1	1	-1	32680							
4	4	4	1	1	1	1	-1	33270							
5	5	5	1	1	-1	-1	1	33430							
6	6	6	1	1	1	-1	1	34430							
7	7	7	1	1	-1	1	1	34070							
8	8	8	1	1	1	1	1	33440							
9	9	9	1	1	-1	-1	-1	32750							
10	10	10	1	1	1	-1	-1	32160							
11	11	11	1	1	-1	1	-1	32270							
12	12	12	1	1	1	1	-1	33080							
13	13	13	1	1	-1	-1	1	33360							
14	14	14	1	1	1	-1	1	34280							
15	15	15	1	1	-1	1	1	33100							
16	16	16	1	1	1	1	1	33570							
17	17	17	1	1	-1	-1	-1	32960							
18	18	18	1	1	1	-1	-1	32640							
19	19	19	1	1	-1	1	-1	33130							

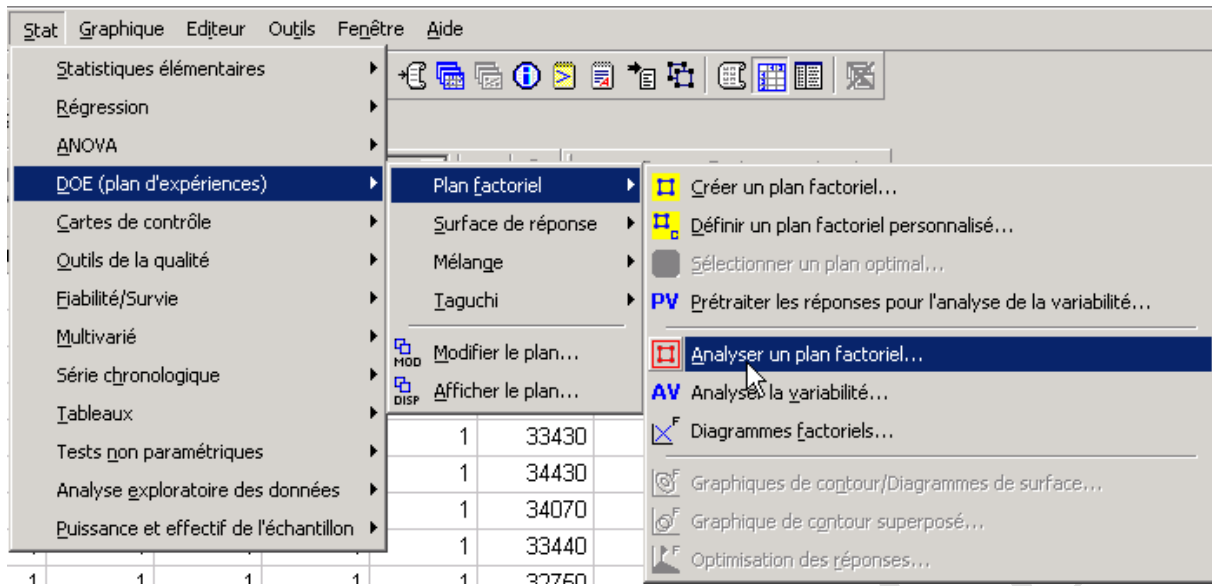
Nous aurions tout aussi bien pu faire:



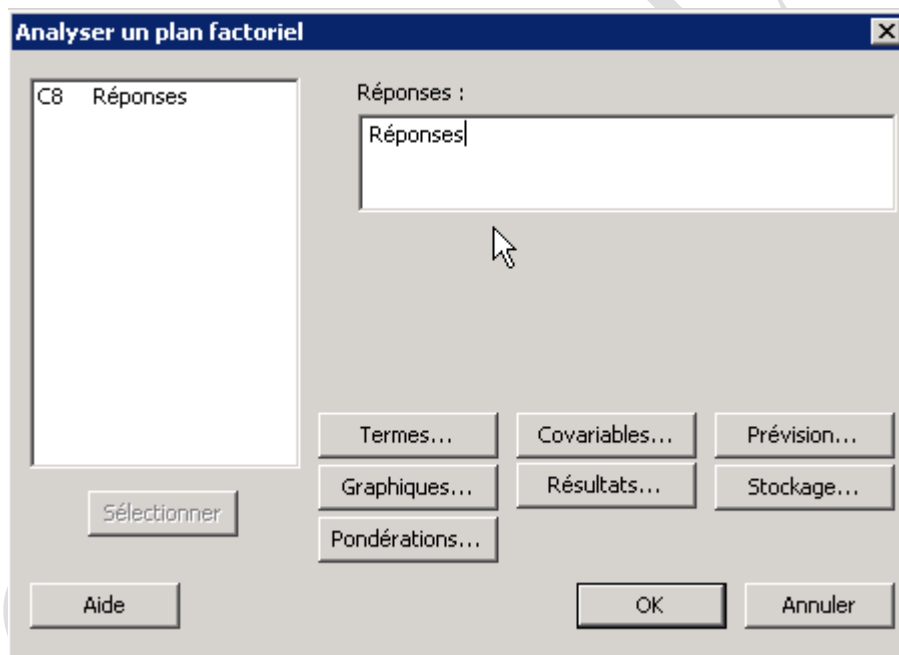
Ce qui nous aurait donné:



Nous allons maintenant dans le menu **Stat/DOE (plan d'expérience)/Plan factoriel/Analyser un plan factoriel...**:



et nous prenons:



Nous cochons et validons le tout par **OK** pour obtenir d'abord dans la fenêtre de session (**attention!!! la sortie à complètement changé dans la version 17 de Minitab!!!**):

**Plan factoriel : Réponses en fonction de A; B; C**

Effets et coefficients estimés pour Réponses (unités codées)

Terme	Effet	Coeff	Coef ErT	T	P
Constante		33254.1	77.64	428.29	0.000
A	346.6	173.3	77.64	2.23	0.040
B	-35.1	-17.5	77.64	-0.23	0.824
C	969.1	484.5	77.64	6.24	0.000
A*B	-160.1	-80.0	77.64	-1.03	0.318
A*C	304.1	152.0	77.64	1.96	0.068
B*C	-110.9	-55.5	77.64	-0.71	0.485
A*B*C	-345.9	-173.0	77.64	-2.23	0.041

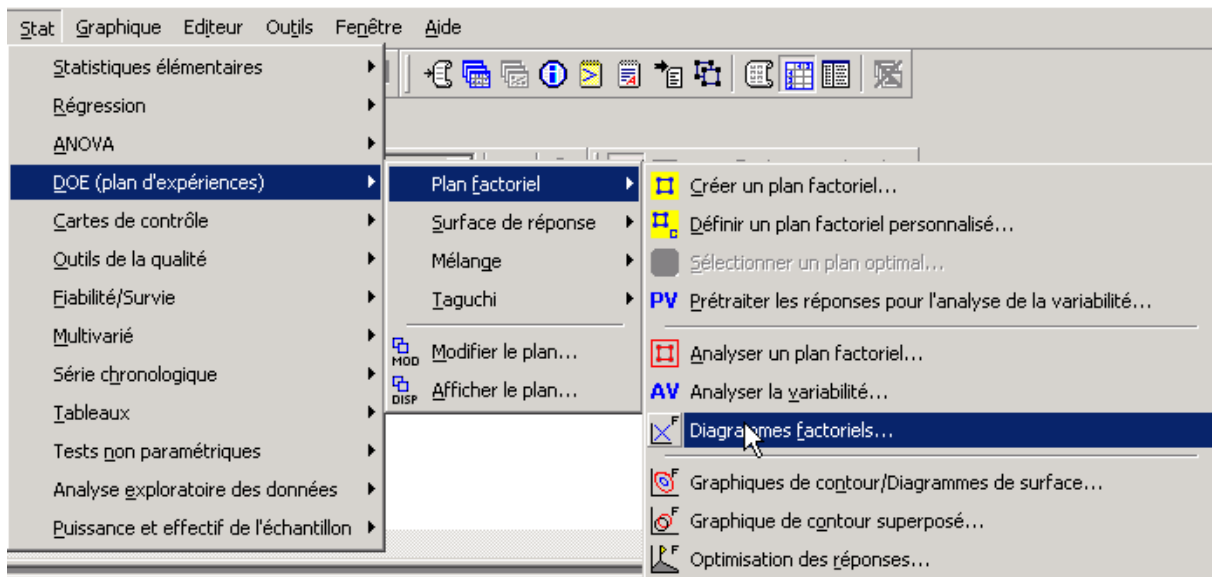
S = 380.375      SomCar-ErrPrév = 5208676  
R carré = 77.26 %      R carré (prév) = 48.82 %      R carré (ajust) = 67.30 %

Analyse de la variance pour Réponses (unités codées)

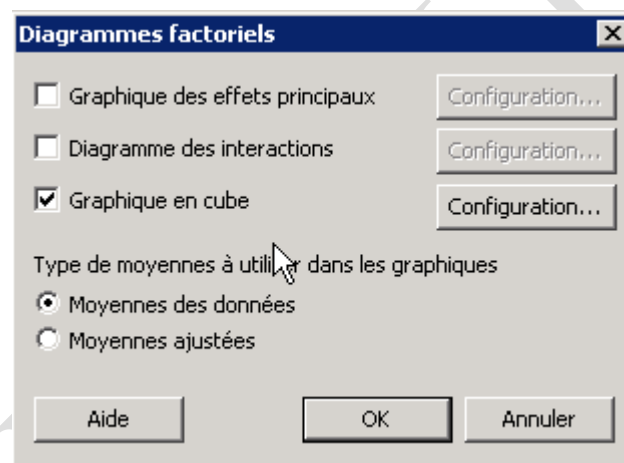
Source	DL	SomCar séq	SomCar ajust	CM ajust	F
Effets principaux	3	6362840	6362840	2120947	14.66
2 Interactions - Nombre de facteurs	3	782375	782375	260792	1.80
3 Interactions - Nombre de facteurs	1	717950	717950	717950	4.96
Erreur résiduelle	16	2314967	2314967	144685	
Erreur pure	16	2314967	2314967	144685	
Total	23	10178133			

Source	P
Effets principaux	0.000
2 Interactions - Nombre de facteurs	0.187
3 Interactions - Nombre de facteurs	0.041
Erreur résiduelle	
Erreur pure	
Total	

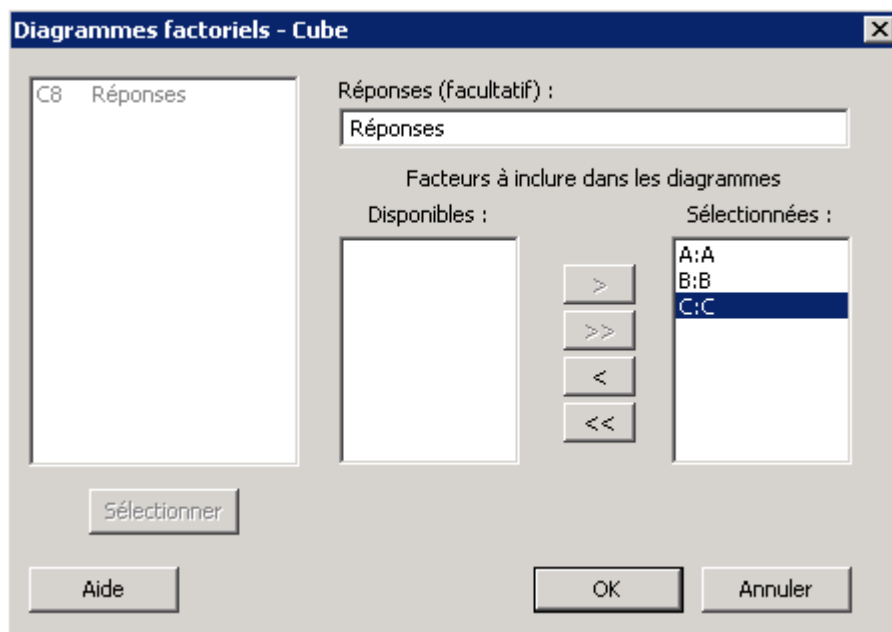
Une fois ceci fait, nous allons dans le menu **Stat/DOE (plan d'expérience)/Plan factoriel/Diagrammes factoriels...**:



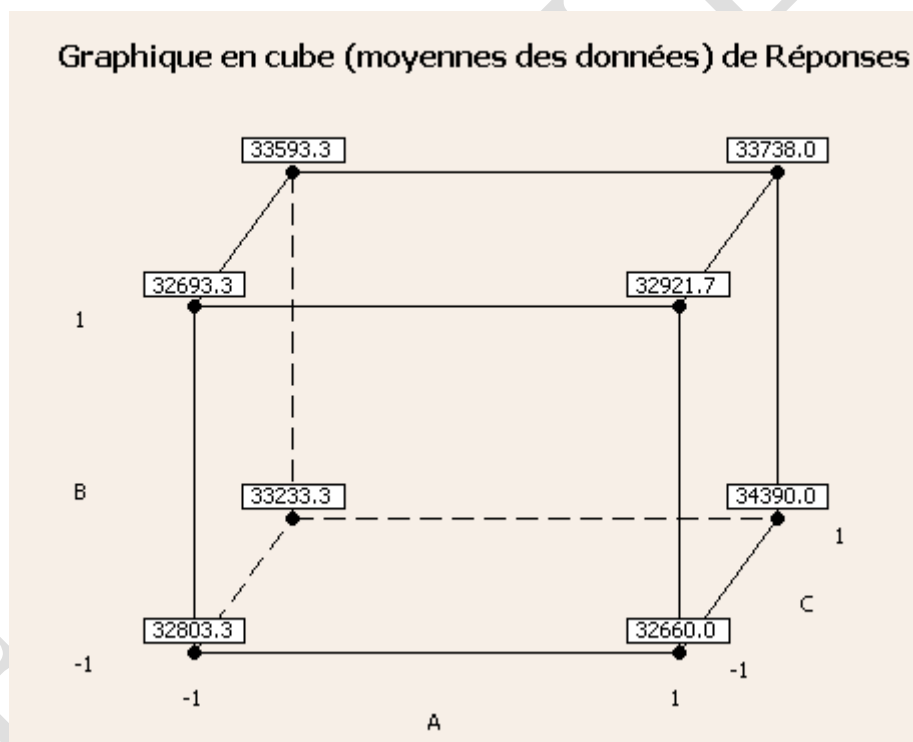
Nous y voyons une "nouauté" disponible qu'avec ce type de plan et qui est le **Graphique en cube**:



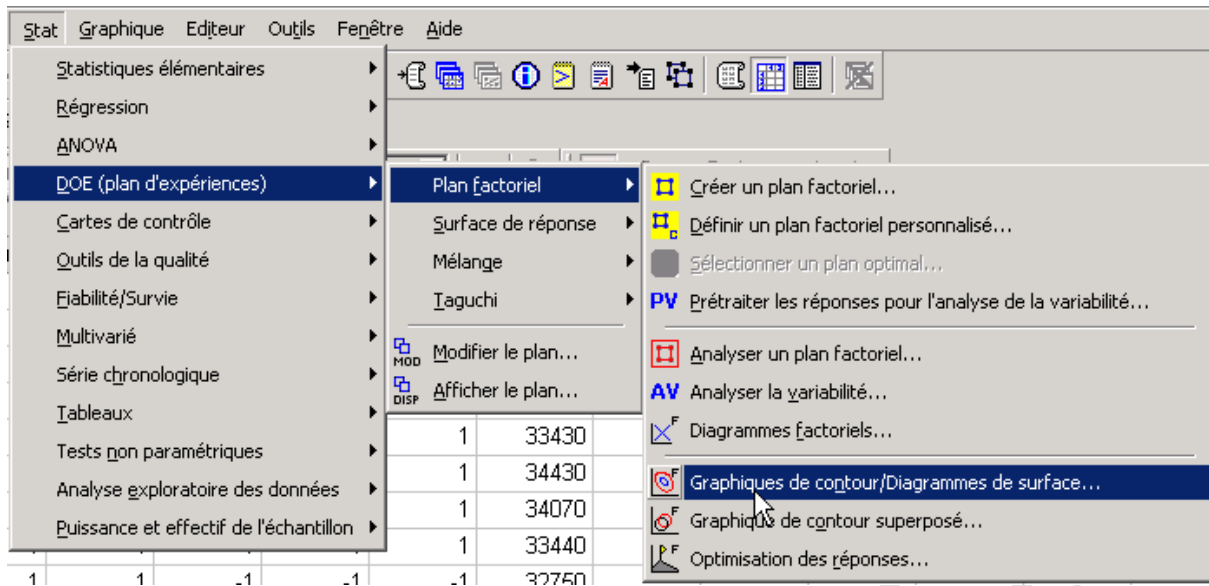
Nous cliquons sur **Configuration...**:



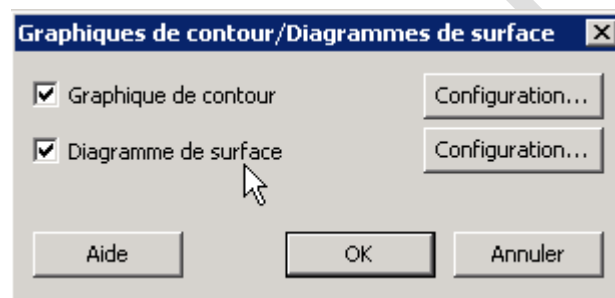
Ce qui donne:



**Les options que nous allons voir maintenant (Graphiques de contour et Optimisation des réponses), s'activent dans le menu que si et seulement si lorsque nous avons lancé l'analyse d'un diagramme factoriel (non général!!!) au préalable.**

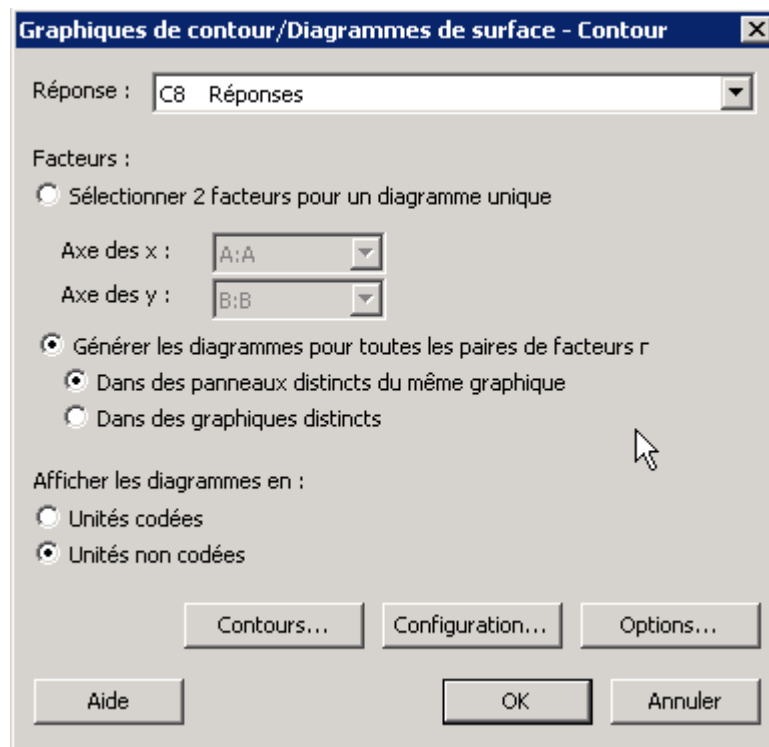


Nous cochons les deux types de graphiques:

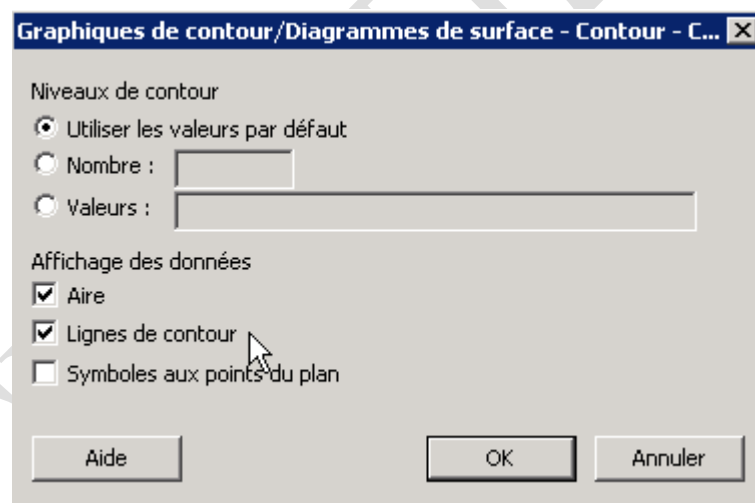


Évidemment, comme nous avons plus de 2 facteurs, il serait difficile de représenter la réponse dans un espace à 4 dimensions... nous allons donc générer dans un même panneau de graphique, toutes les combinaisons (ce qui devient donc difficile à interpréter, d'où l'utiliser assez discutable de l'utilisation de graphiques à ce niveau-là...):

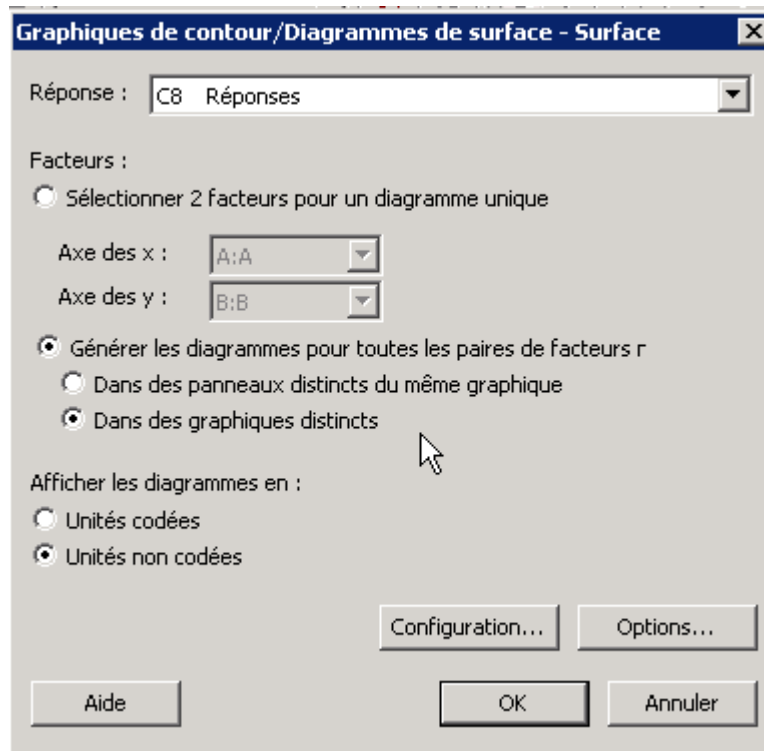




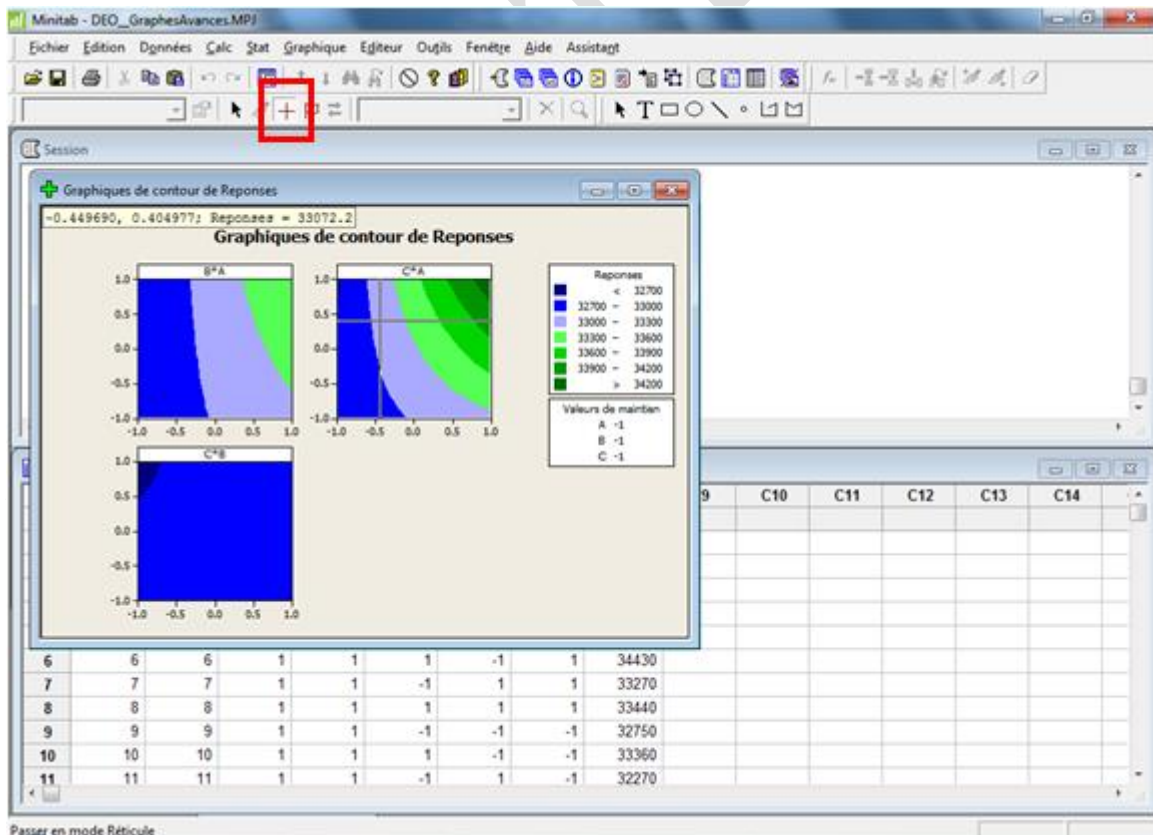
Nous cliquons sur le bouton **Contours...**:



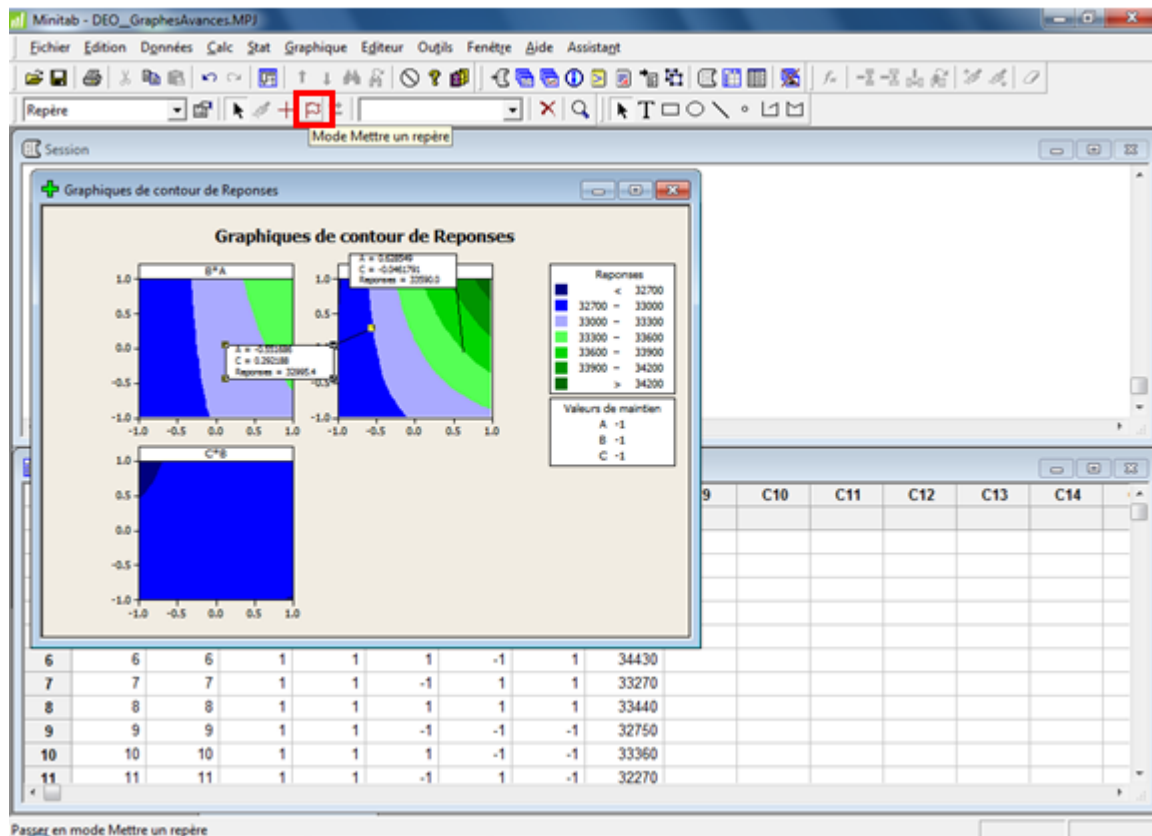
Nous faisons de même pour l'autre diagramme:



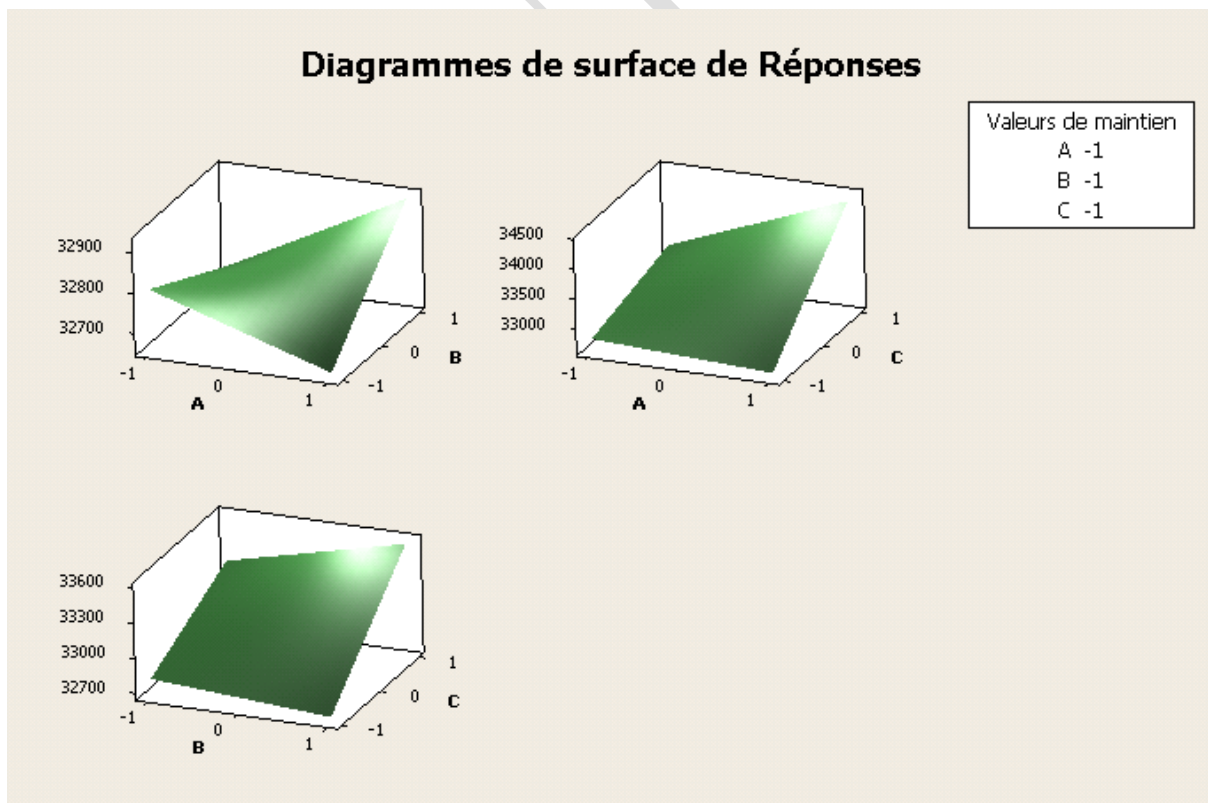
Et nous obtenons les surfaces de réponses projetées avec les isoclines (n'oubliez pas le curseur qui permet d'obtenir les coordonnées d'un point!):



Curieusement voici une autre fonctionnalité utile appelée **Mode Mettre un repère** mais qui ne fonctionne qu'avec les graphiques de contours générés par les plans d'expérience:



Nous pouvons aussi sinon simplement une surface de réponse en 3D simple:



## 17.14. Exercice 185.: Analyse de la variabilité (**WP**)

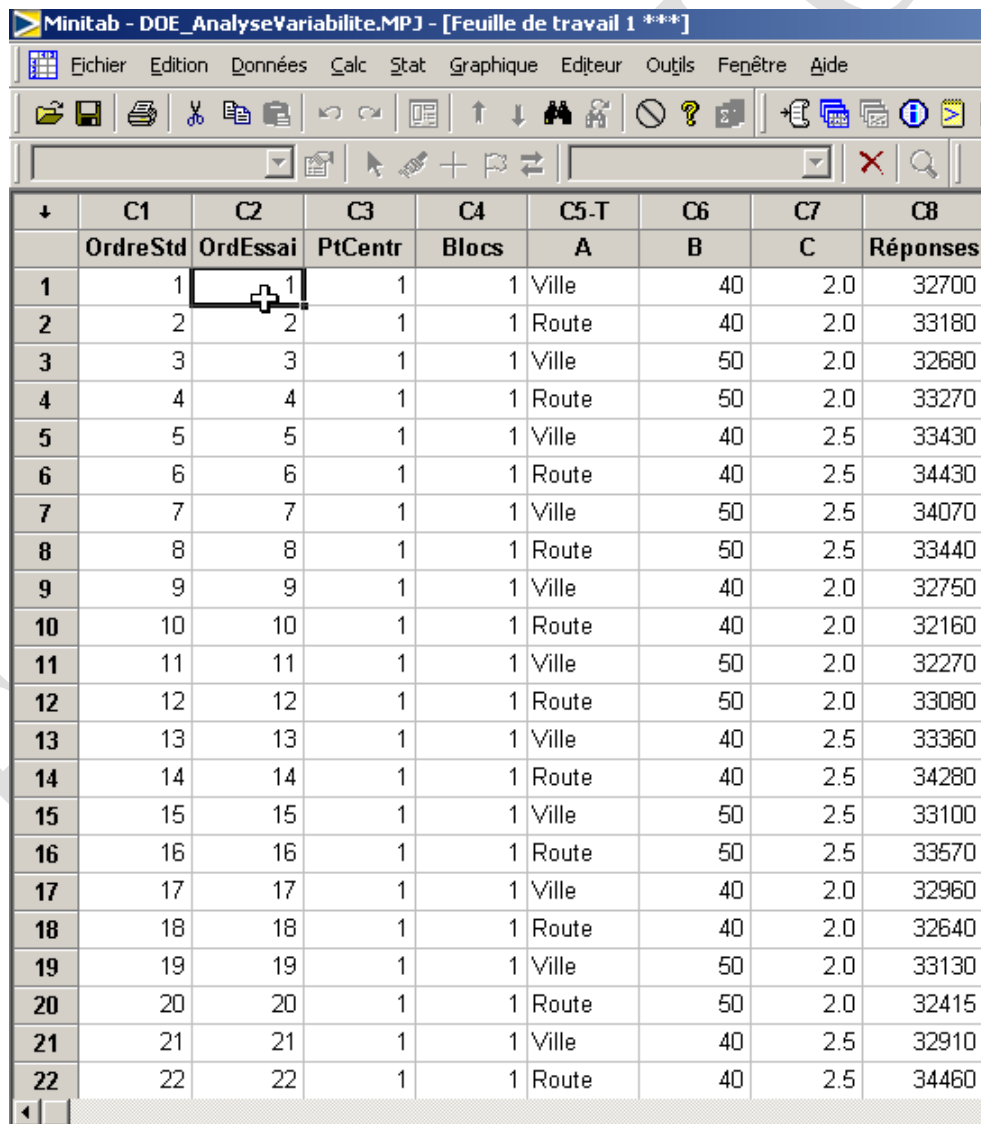
Minitab® Statistical Software 15.1.1.0

Dans un plan factoriel le raisonnement est fait sur les moyennes des points relevés. L'idée de l'analyse de la variabilité est d'atteindre une réponse cible/max ou min comme à l'habitude mais en même temps de "jouer" avec la variance pour la maximiser/minimiser. Il s'agit donc d'un plan d'expérience à double contrainte!

Les conditions d'utilisation pour cet outil sont dans Minitab:

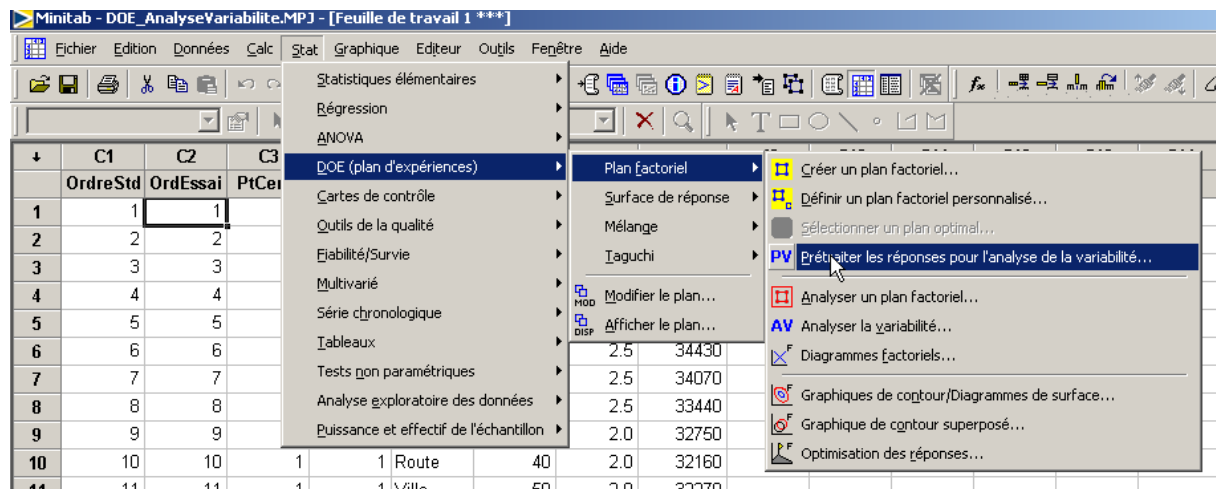
- Analyse de la variabilité ne fonctionne que pour les plans factoriels à deux niveaux!
- Les mesures doivent être répétées pour pouvoir calculer une variance!

Pour découvrir cela, nous allons utiliser le fichier suivant (3 facteurs à 2 niveaux):

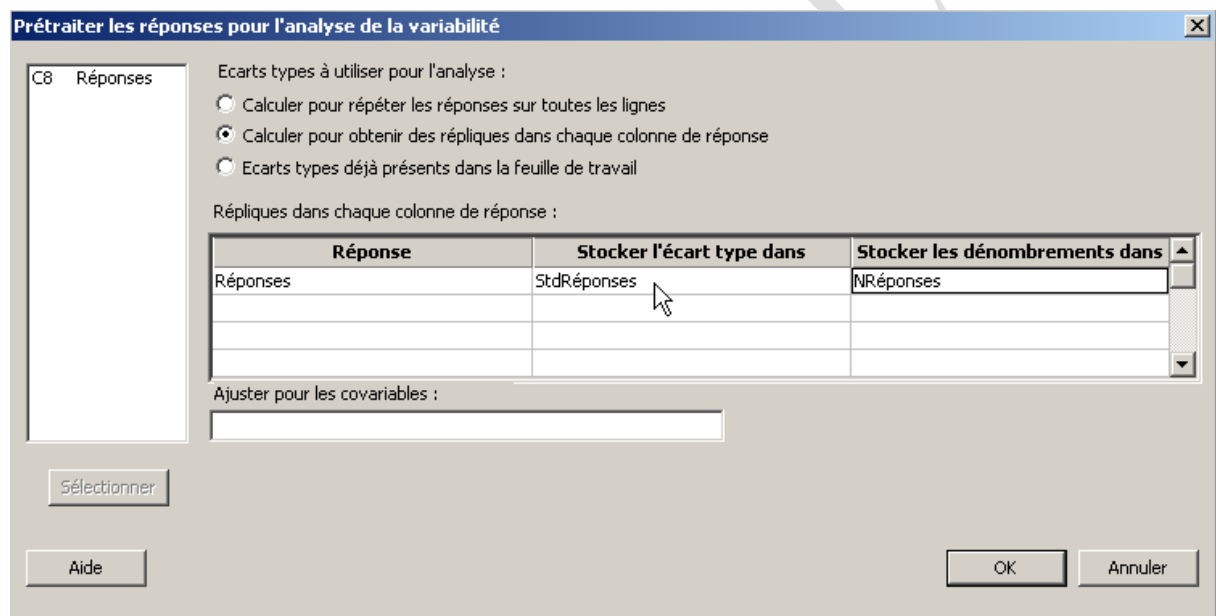


	C1	C2	C3	C4	C5-T	C6	C7	C8
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	C	Réponses
1	1	1	1	1	Ville	40	2.0	32700
2	2	2	1	1	Route	40	2.0	33180
3	3	3	1	1	Ville	50	2.0	32680
4	4	4	1	1	Route	50	2.0	33270
5	5	5	1	1	Ville	40	2.5	33430
6	6	6	1	1	Route	40	2.5	34430
7	7	7	1	1	Ville	50	2.5	34070
8	8	8	1	1	Route	50	2.5	33440
9	9	9	1	1	Ville	40	2.0	32750
10	10	10	1	1	Route	40	2.0	32160
11	11	11	1	1	Ville	50	2.0	32270
12	12	12	1	1	Route	50	2.0	33080
13	13	13	1	1	Ville	40	2.5	33360
14	14	14	1	1	Route	40	2.5	34280
15	15	15	1	1	Ville	50	2.5	33100
16	16	16	1	1	Route	50	2.5	33570
17	17	17	1	1	Ville	40	2.0	32960
18	18	18	1	1	Route	40	2.0	32640
19	19	19	1	1	Ville	50	2.0	33130
20	20	20	1	1	Route	50	2.0	32415
21	21	21	1	1	Ville	40	2.5	32910
22	22	22	1	1	Route	40	2.5	34460

Nous allons ensuite dans le menu **DOE (plan d'expérience)/Plan factoriel/Prétraiter les réponses pour l'analyse de la variabilité...**:



Et nous prenons:



Et nous validons par **OK**.

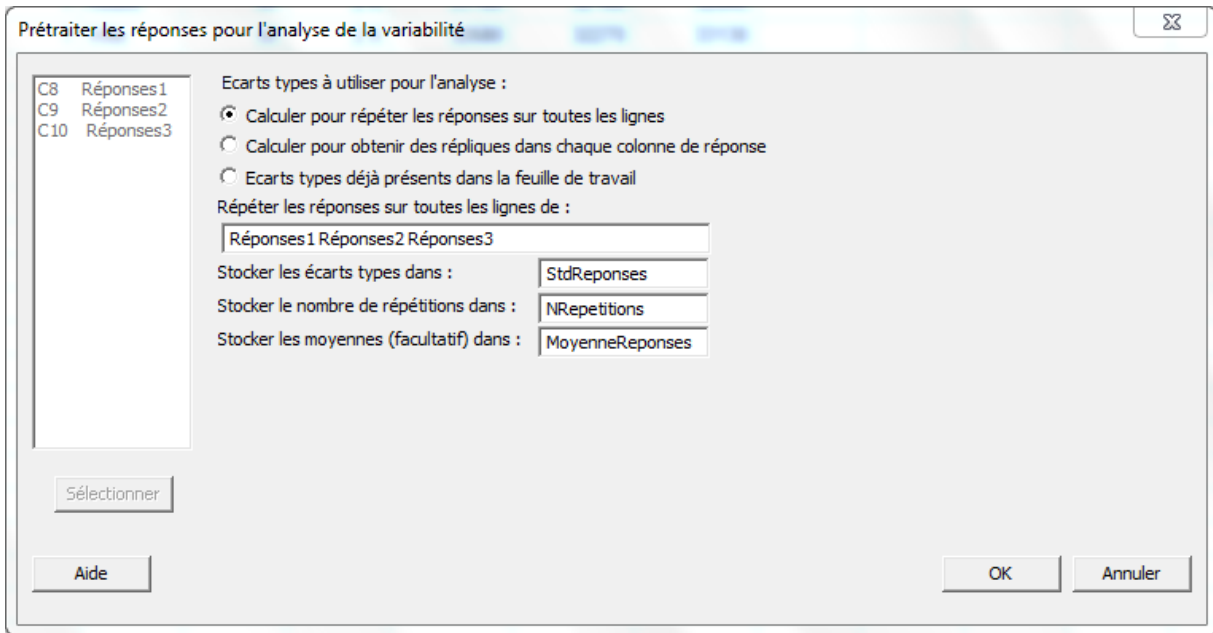
Nous obtenons alors:

	C1	C2	C3	C4	C5-T	C6	C7	C8	C9	C10
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	C	Réponses	StdRéponses	NRéponses
1	1	1	1	1	Ville	40	2.0	32700	137.961	3
2	2	2	1	1	Route	40	2.0	33180	510.294	3
3	3	3	1	1	Ville	50	2.0	32680	430.155	3
4	4	4	1	1	Route	50	2.0	33270	448.952	3
5	5	5	1	1	Ville	40	2.5	33430	282.194	3
6	6	6	1	1	Route	40	2.5	34430	96.437	3
7	7	7	1	1	Ville	50	2.5	34070	485.215	3
8	8	8	1	1	Route	50	2.5	33440	408.769	3
9	9	9	1	1	Ville	40	2.0	32750	*	*
10	10	10	1	1	Route	40	2.0	32160	*	*
11	11	11	1	1	Ville	50	2.0	32270	*	*
12	12	12	1	1	Route	50	2.0	33080	*	*
13	13	13	1	1	Ville	40	2.5	33360	*	*

Remarque: Si les données de réplifications avaient été reportées comme visible ci-dessous:

	C1	C2	C3	C4	C5-T	C6	C7	C8	C9	C10	C11	C12	C13
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	C	Réponses1	Réponses2	Réponses3			
1	1	1	1	1	Ville	40	2.0	32700	32750	32960			
2	2	2	1	1	Route	40	2.0	33180	32160	32640			
3	3	3	1	1	Ville	50	2.0	32680	32270	33130			
4	4	4	1	1	Route	50	2.0	33270	33080	32415			
5	5	5	1	1	Ville	40	2.5	33430	33360	32910			
6	6	6	1	1	Route	40	2.5	34430	34280	34460			
7	7	7	1	1	Ville	50	2.5	34070	33100	33610			
8	8	8	1	1	Route	50	2.5	33440	33570	34204			
9													
10													

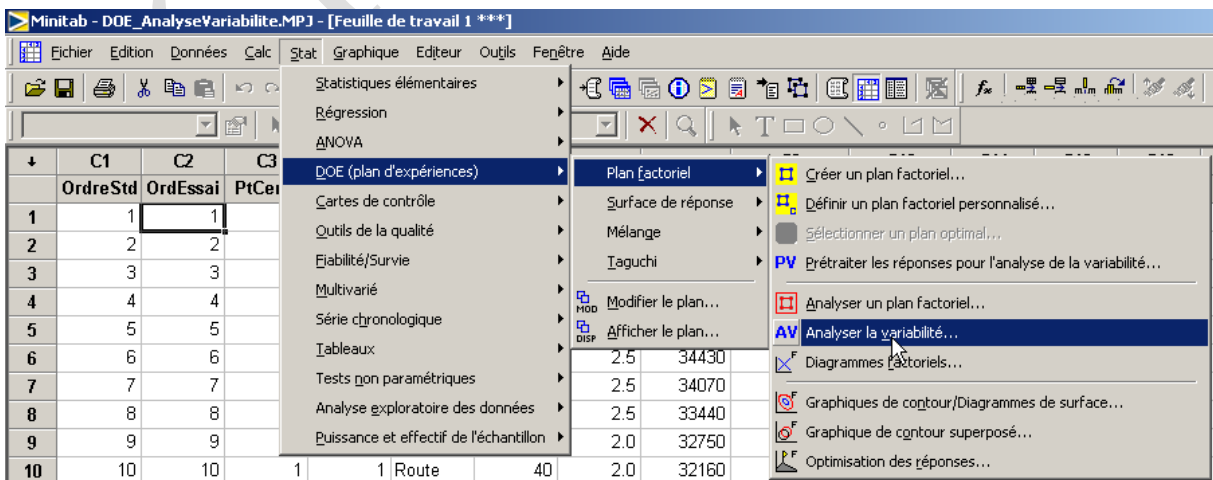
alors il aurait fallu sélectionner l'option **Calculer pour répéter les réponses sur toutes les lignes**:



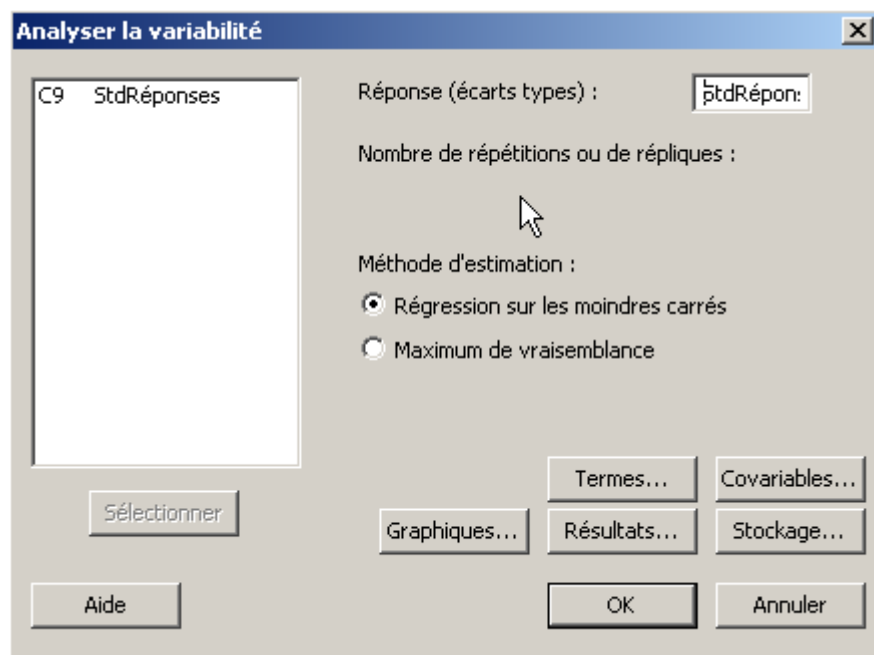
Ce qui aurait donné:

	C1	C2	C3	C4	C5-T	C6	C7	C8	C9	C10	C11	C12	C13
	OrdreStd	OrdEssai	PtCentr	Blocs	A	B	C	Réponses1	Réponses2	Réponses3	StdReponses	NRépetitions	MoyenneReponses
1	1	1	1	1	Ville	40	2.0	32700	32750	32960	137.961	3	32803.3
2	2	2	1	1	Route	40	2.0	33180	32160	32640	510.294	3	32660.0
3	3	3	1	1	Ville	50	2.0	32680	32270	33130	430.155	3	32693.3
4	4	4	1	1	Route	50	2.0	33270	33080	32415	448.952	3	32921.7
5	5	5	1	1	Ville	40	2.5	33430	33360	32910	282.194	3	33233.3
6	6	6	1	1	Route	40	2.5	34430	34280	34460	96.437	3	34390.0
7	7	7	1	1	Ville	50	2.5	34070	33100	33610	485.215	3	33593.3
8	8	8	1	1	Route	50	2.5	33440	33570	34204	408.769	3	33738.0
9													
10													

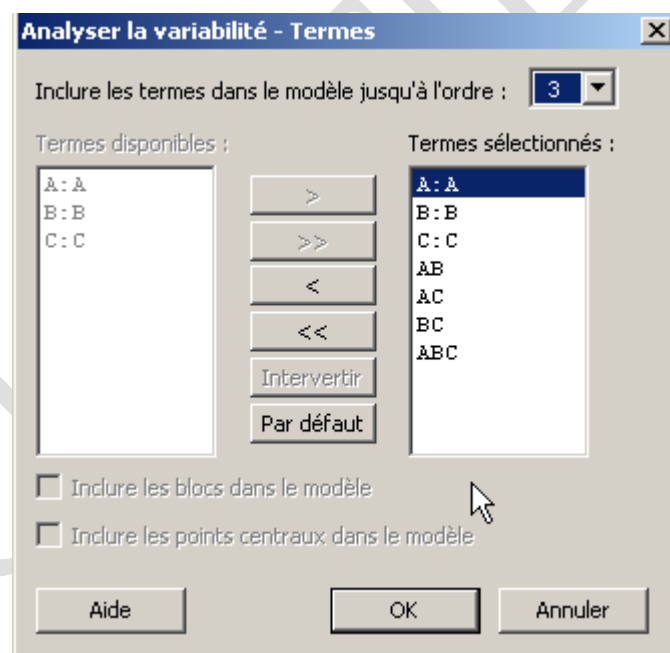
Une fois ceci fait, nous pouvons nous rendre le menu **DOE (plan d'expérience)/Plan factoriel/Analyser la variabilité...**:



Il vient alors:

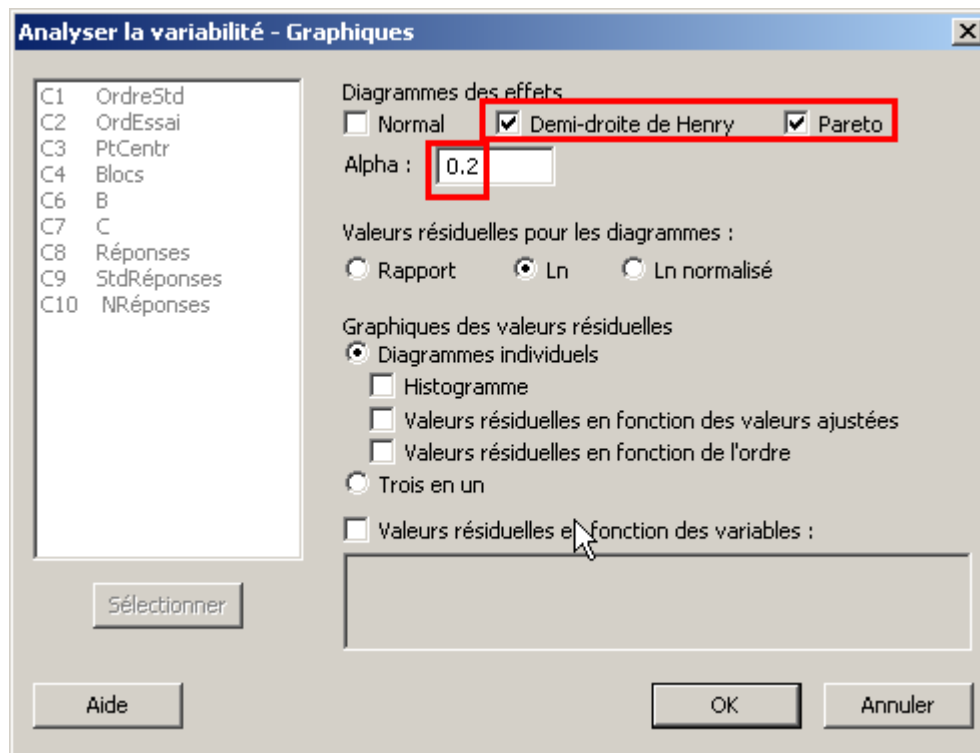


Dans le bouton **Termes...** nous prenons tout:



Et dans le bouton **Graphiques...** (nous utilisons le  $L_n$  pour "Normaliser" un peu les données... selon la méthode de Box-Cox en croisant les doigts que ça ira car nous le verrons avec la droite de Henry plus bas!):





La valeur alpha choisie ici est le choix du seuil de signification qui fait que nous allons éliminer automatiquement les facteurs ou interactions de facteurs dont la  $p$ -value est au-dessus!

En validant le tout par **OK** nous obtenons d'abord dans la fenêtre de sessions:

Estimation des effets et des coefficients de régression du ln de StdRéponses  
(unités codées)

Terme	Effet	Effet de rapport	Coeff
Constante			5.71778
A	0.026409	1.02676	0.01320
B	0.748841	2.11455	0.37442
C	-0.230951	0.79378	-0.11548
A*B	-0.090744	0.91325	-0.04537
A*C	-0.648984	0.52258	-0.32449
B*C	0.244290	1.27671	0.12215
A*B*C	0.541877	1.71923	0.27094

Analyse de la variance de ln de StdRéponses

Source	DL	SomCar séq	SomCar ajust	CM ajust	F	P
Effets principaux	3	2.99002	2.99002	0.9967	*	*
Interactions à 2 facteur(s)	3	2.37866	2.37866	0.7929	*	*
Interactions à 3 facteur(s)	1	1.42805	1.42805	1.4281	*	*
Erreur résiduelle	0	*	*	*		
Total	7	6.79673				

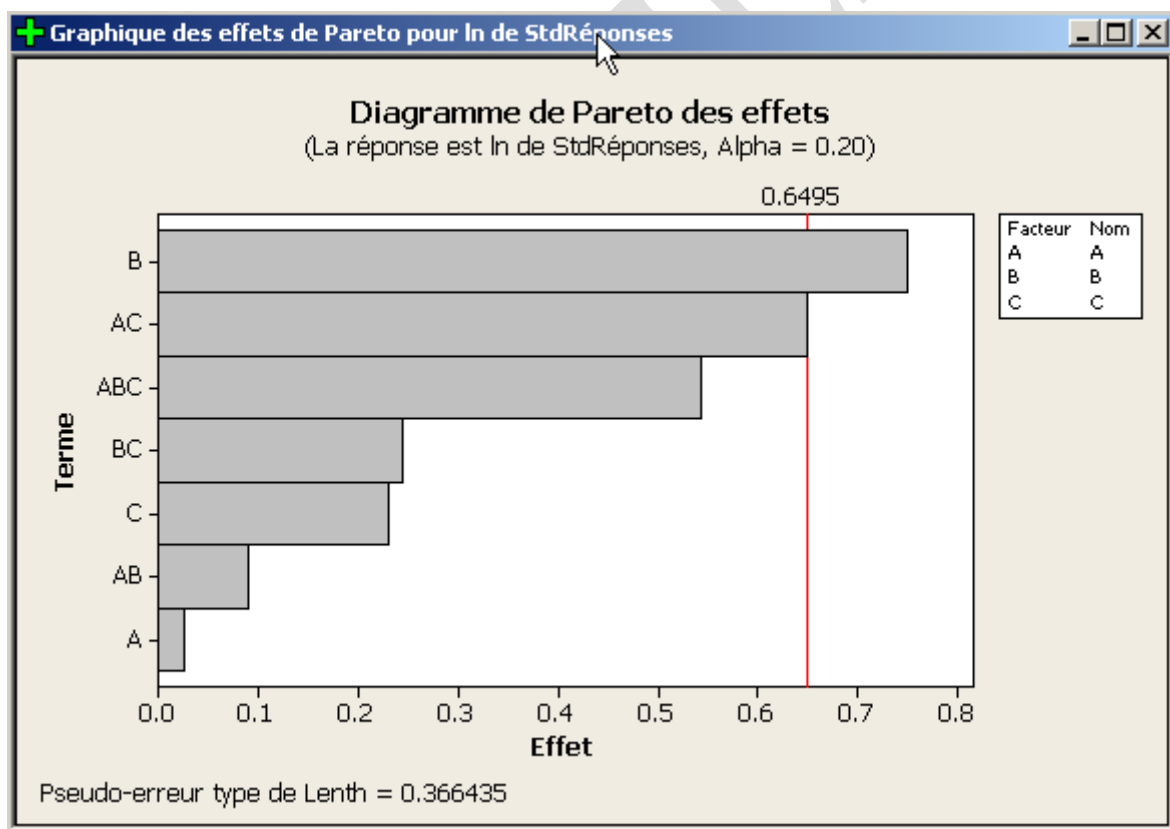
Estimation des effets et des coefficients de régression du ln de StdRéponses  
(unités non codées)

Terme	Coeff
Constante	13.2810
A	25.2880
B	-0.144977
C	-4.85912
A*B	-0.496764
A*C	-11.0518
B*C	0.0977161
A*B*C	0.216751

Structure des alias

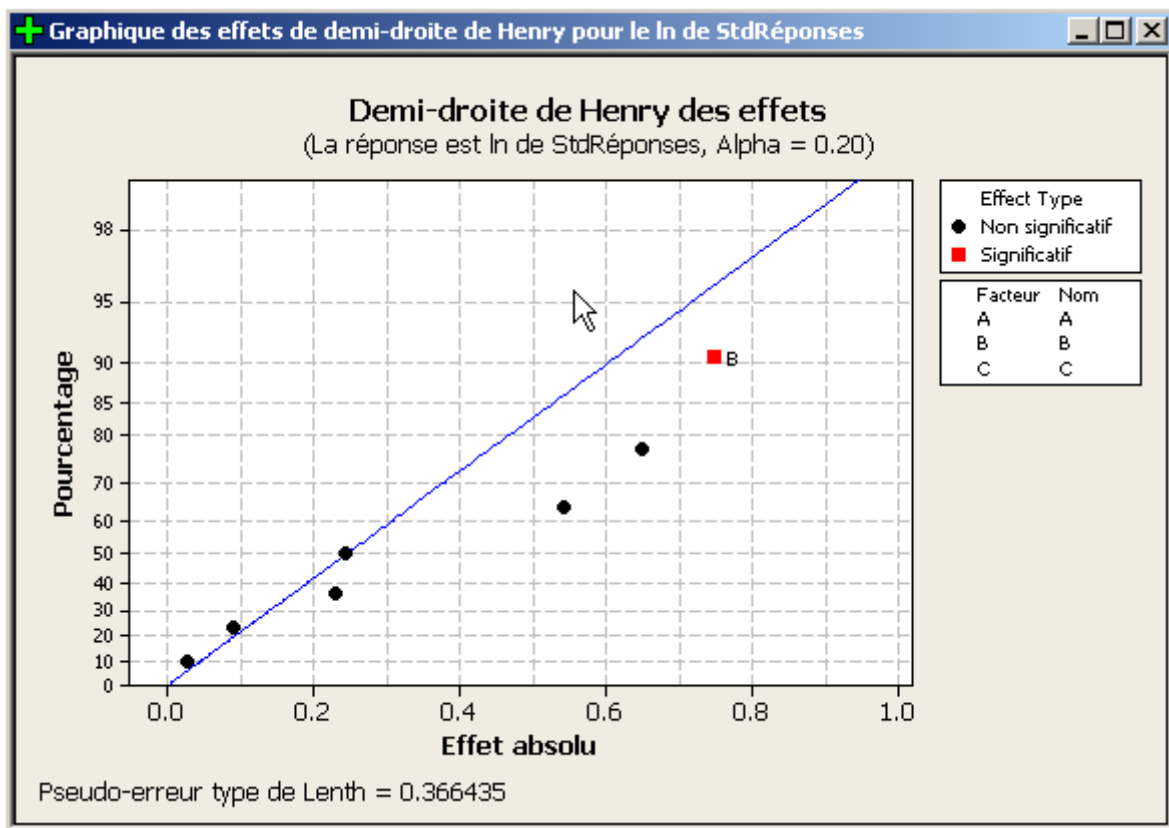
I  
A  
B  
C  
A\*B  
A\*C  
B\*C  
A\*B\*C

Et le premier graphique suivant:

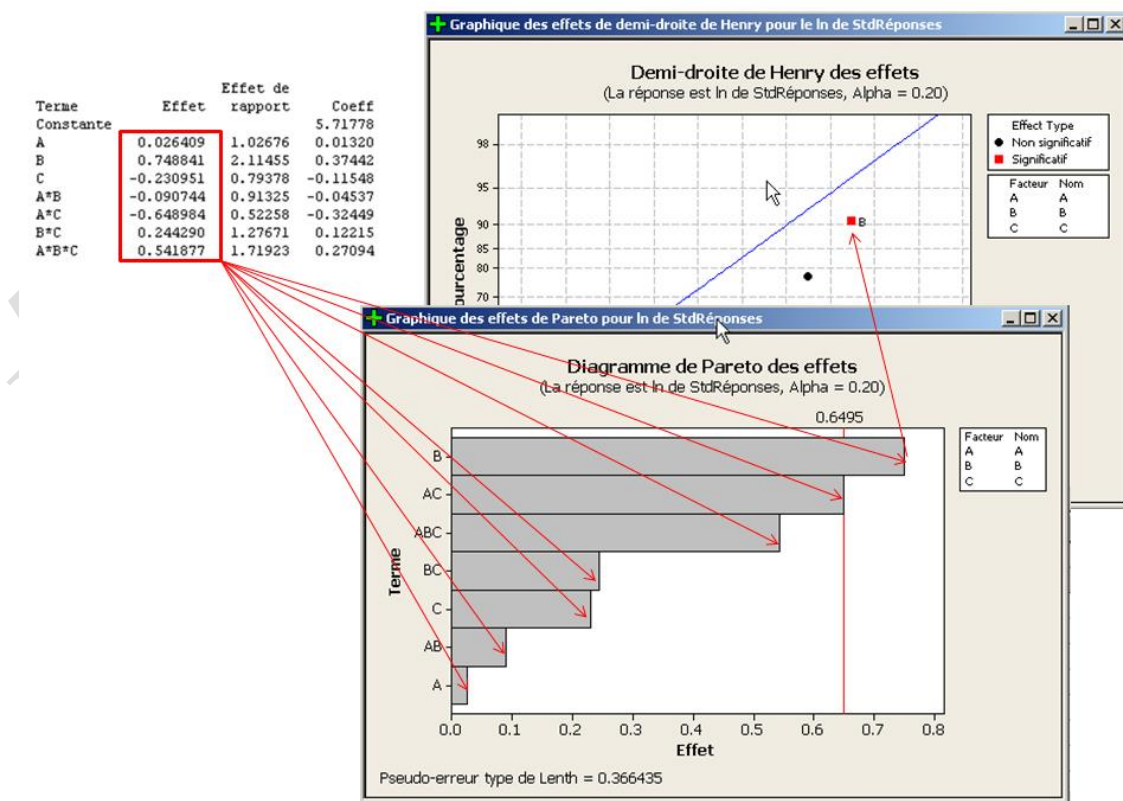


Donc dans l'optique de faire du DOE de la variance (au lieu de la moyenne), nous pourrions éliminer le facteur A du modèle pour commencer puisqu'il n'y a aucune influence sur la variance mais la question est de savoir s'il a influence cependant sur la réponse... Bref c'est un jeu subtil!

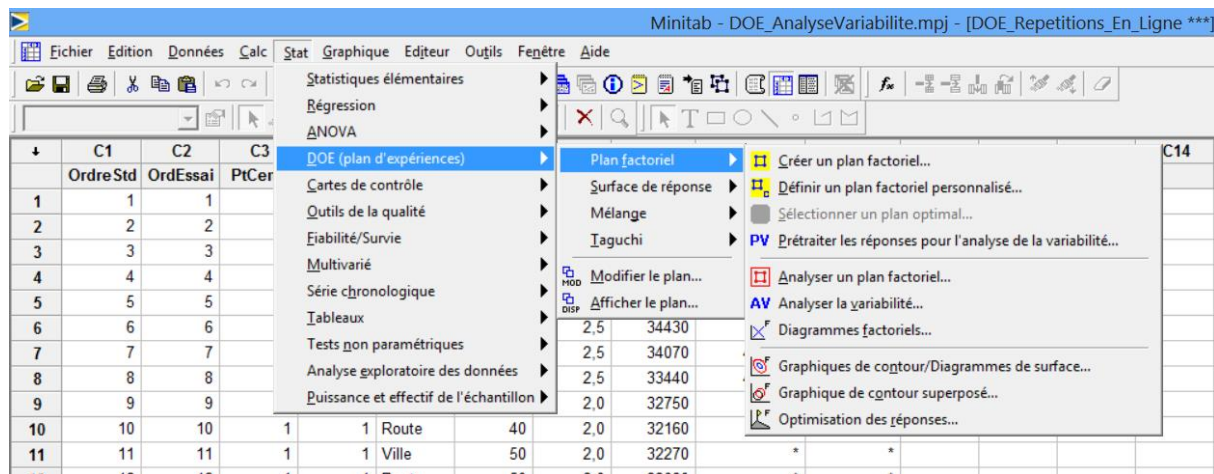
Au niveau du deuxième graphique:



Nous voyons que toutes les données sont bien reliées:



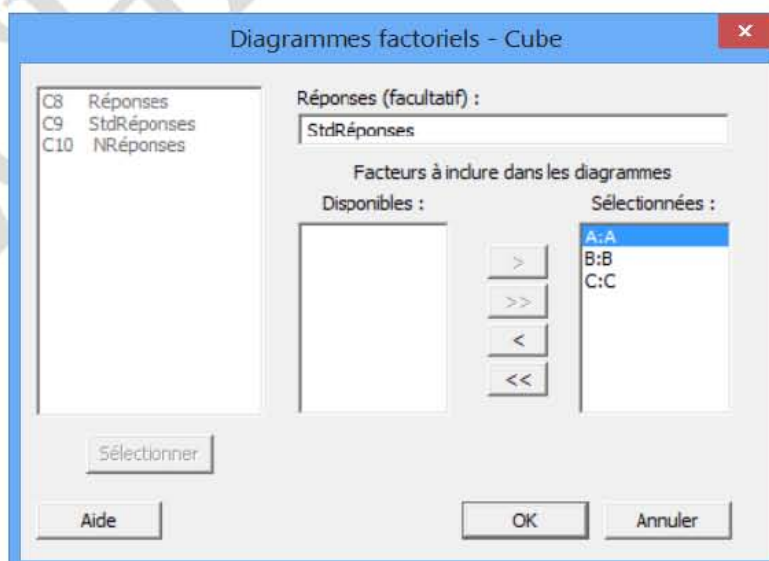
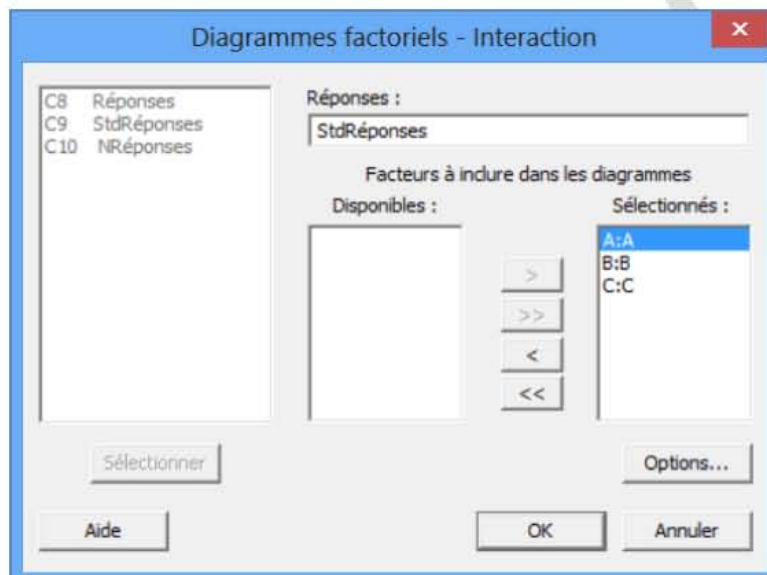
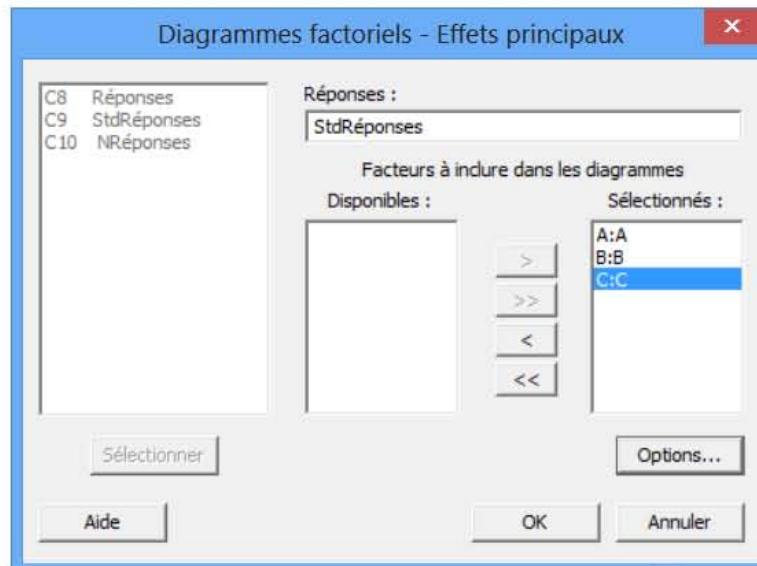
Voyons les diagrammes factoriels correspondants qui permettent de comprendre le paramètre *Effet de rapport* (en gros il s'agit du ratio entre la variable d'intérêt – ici sa variance - au niveau haut et au niveau bas). Pour cela nous allons dans le menu **Stat/DOE (plan d'expériences)/Plan factoriel/Diagrammes factoriels...**:



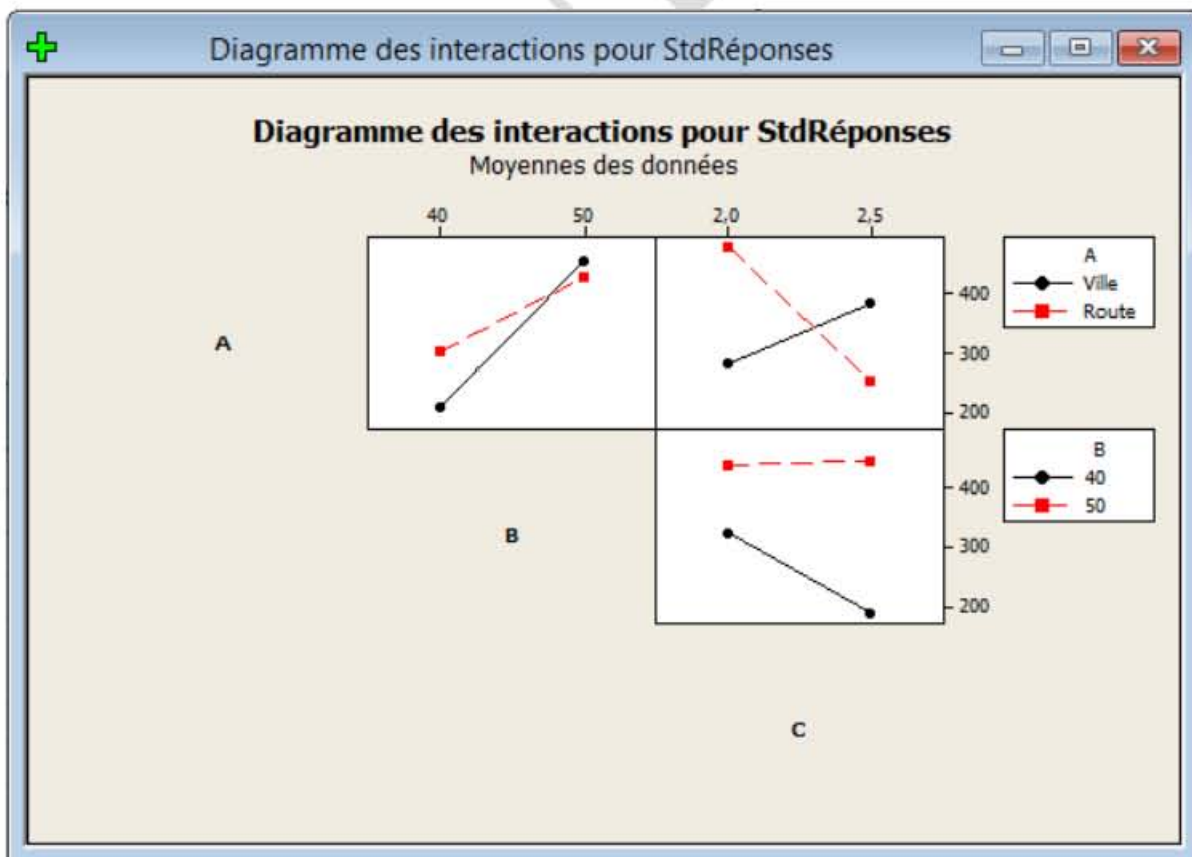
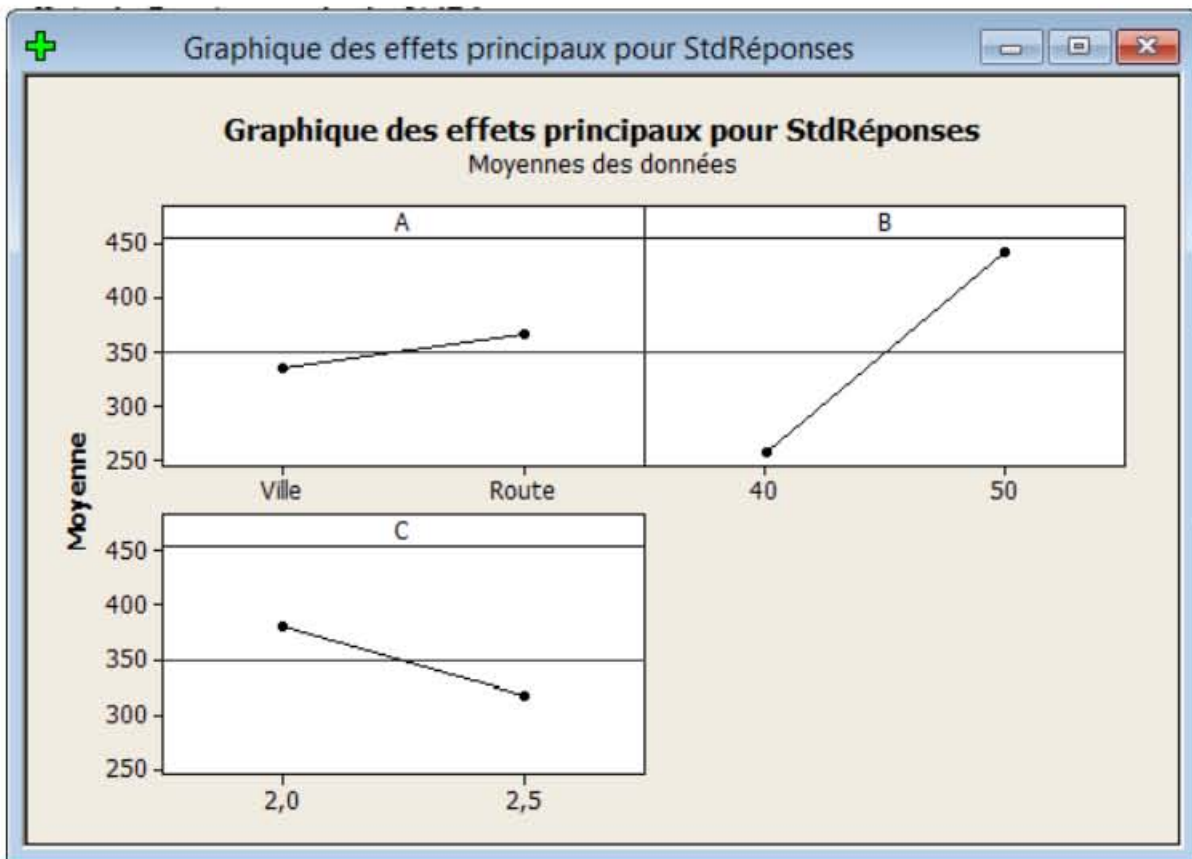
Dans la boîte de dialogue nous allons tout activer (même si ce n'est que la premier et deuxième qui nous intéressent!):



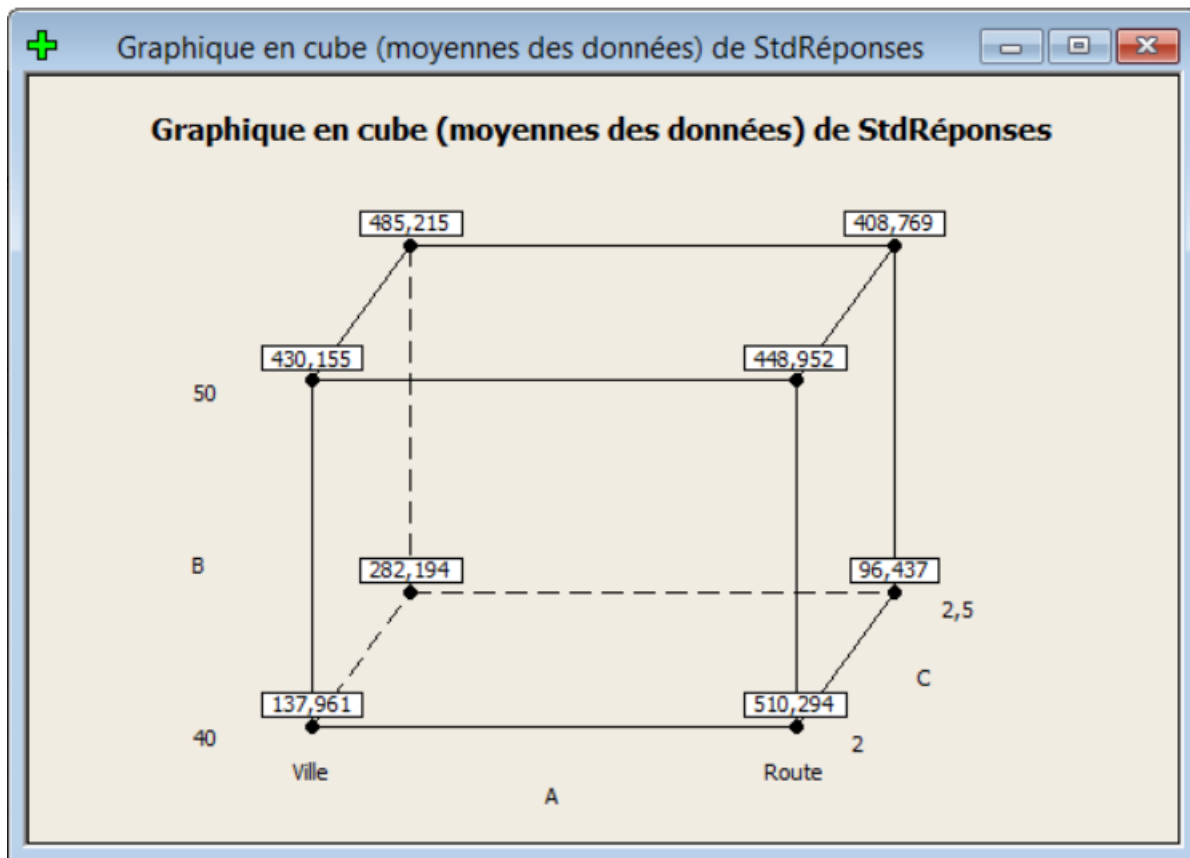
Et dans le bouton de **Configuration...** de chacun nous allons mettre respectivement exactement les mêmes paramètres:



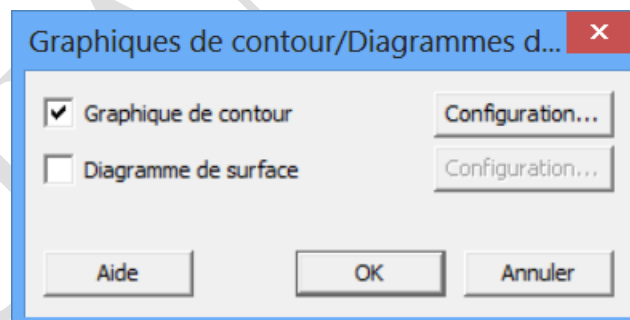
Nous obtenons alors pour la moyenne des variances :



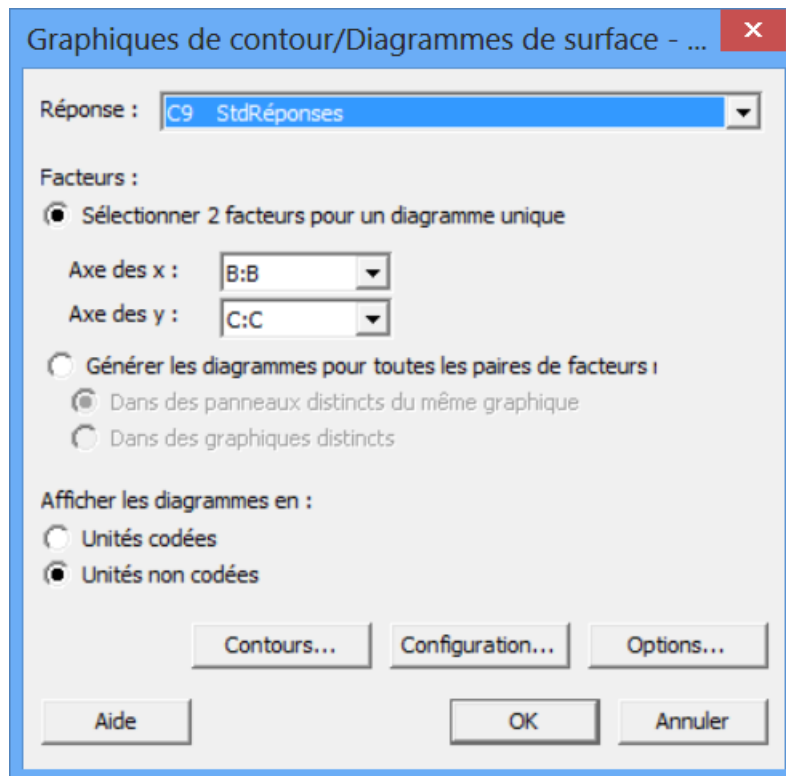
Et le graphique en cube qui peut toujours servir surtout dans le cas à 3 facteurs (et aussi avec les chiffres visibles avec précision dans le graphique):



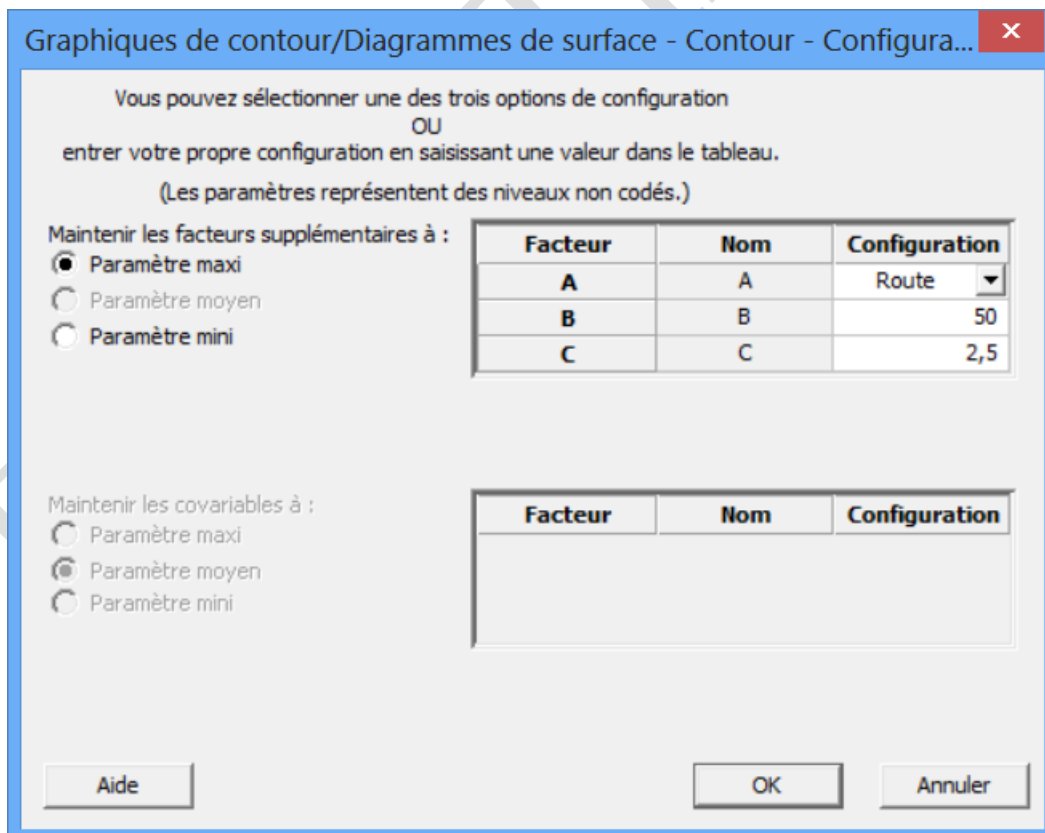
On peut aussi ajouter une analyse par **Diagramme de contour...**:



Et dans configuration on va prendre:

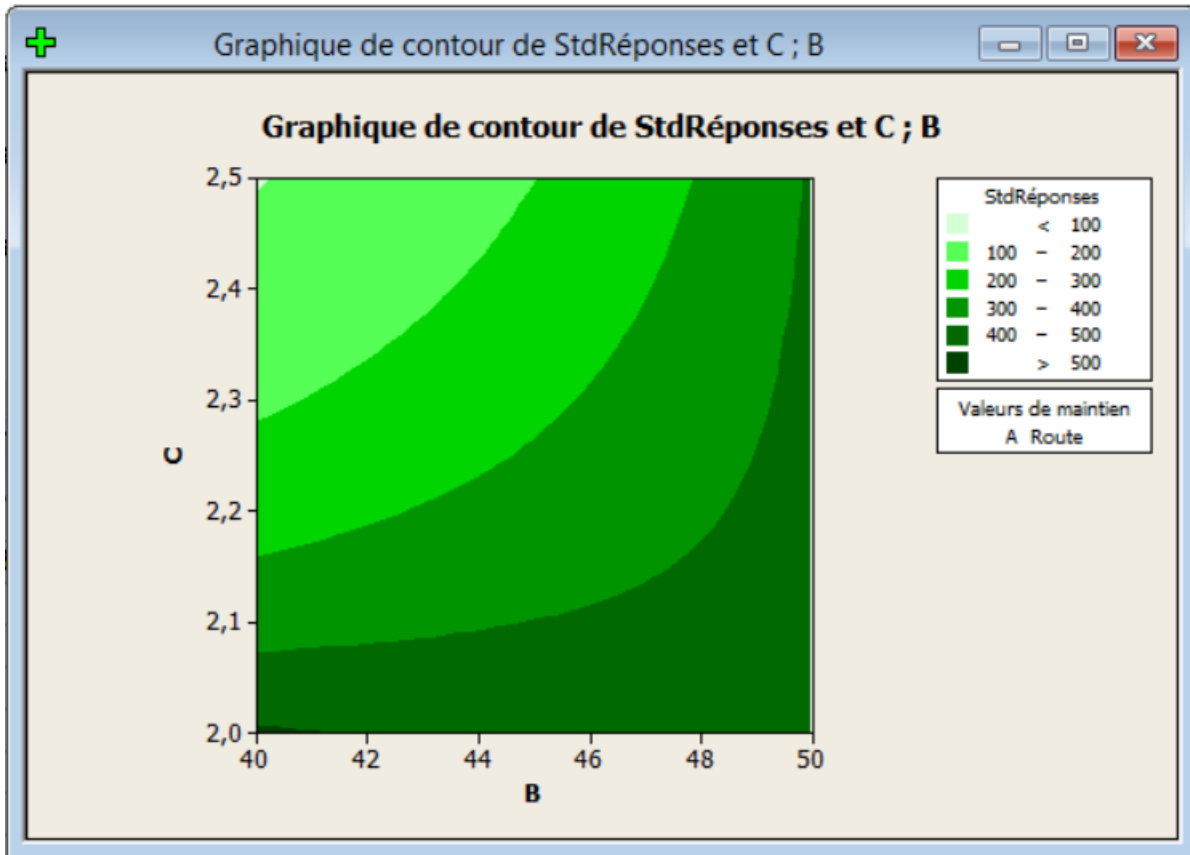


Et dans **Contours...** on va prendre:

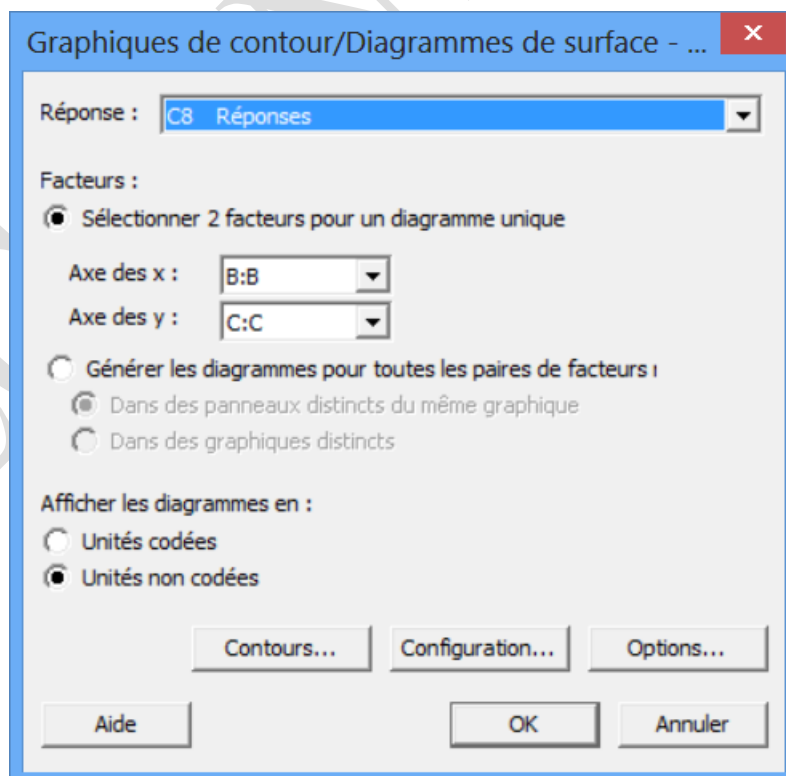


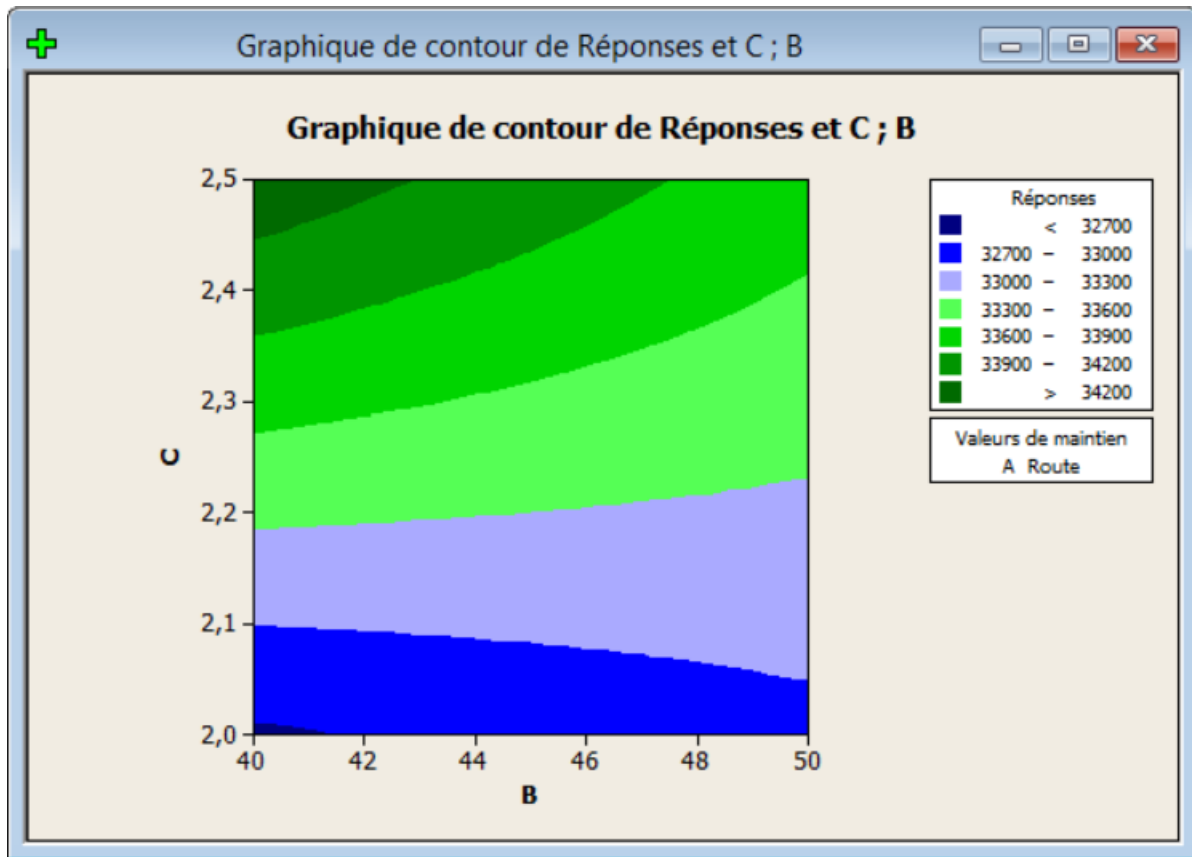
Ce qui donne:



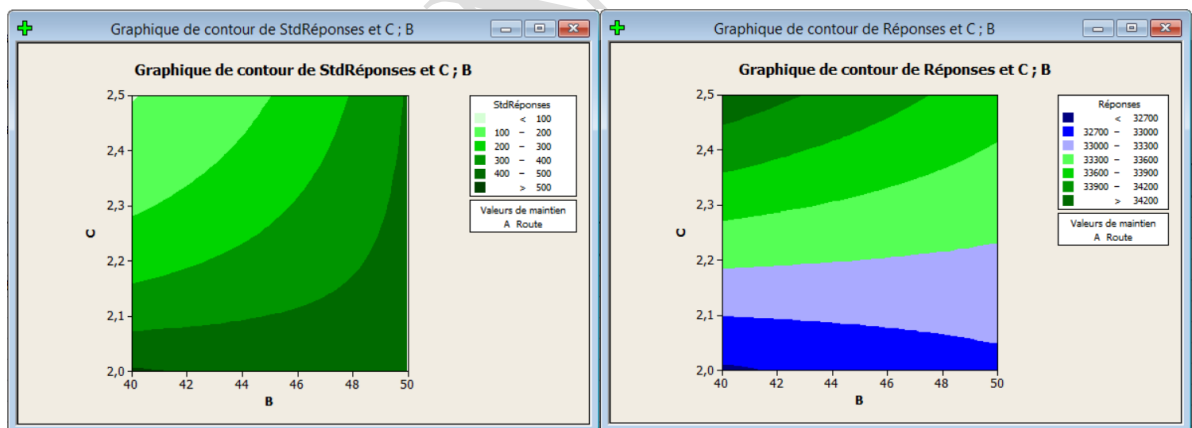


Si on a aussi lancé au préalable une analyse du plan factoriel (et pas que de la variabilité!) et on peut choisir après-coup de faire le même graph pour la variable d'intérêt:

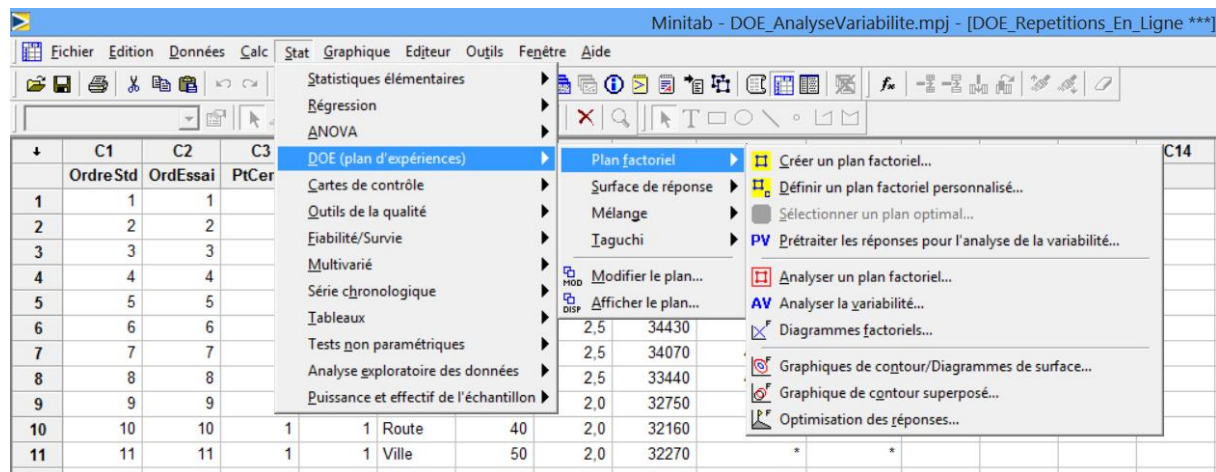




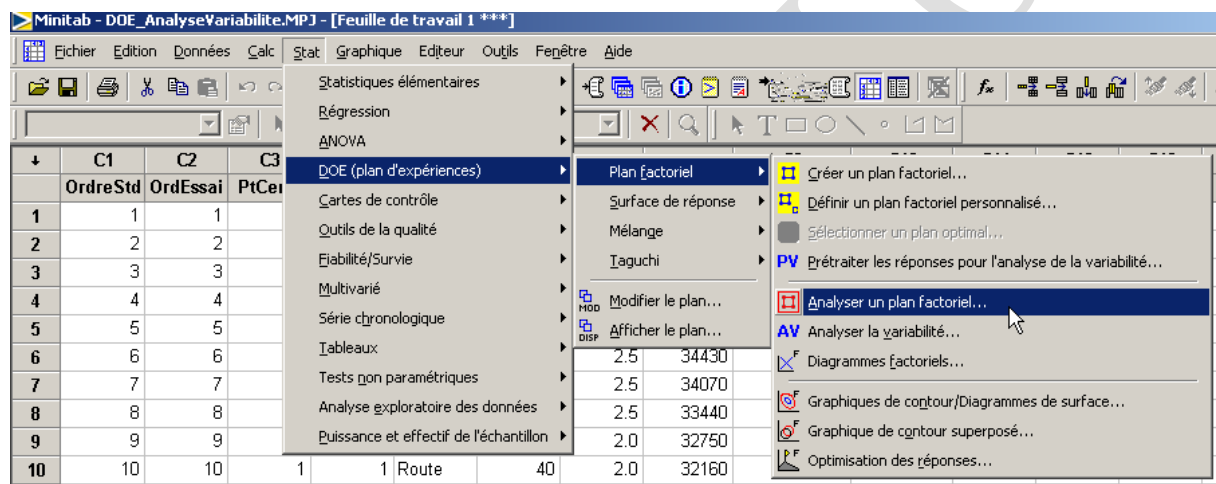
Soit mis côte à côté (souvent on va chercher à minimiser la variance et à maximiser la réponse):



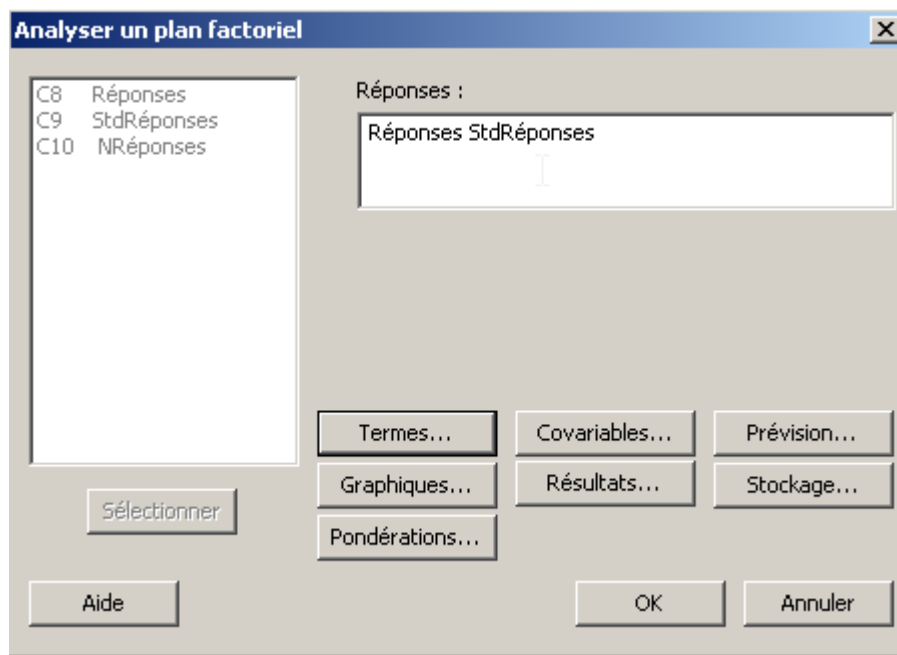
Maintenant faisons une analyse des deux variables en même temps:



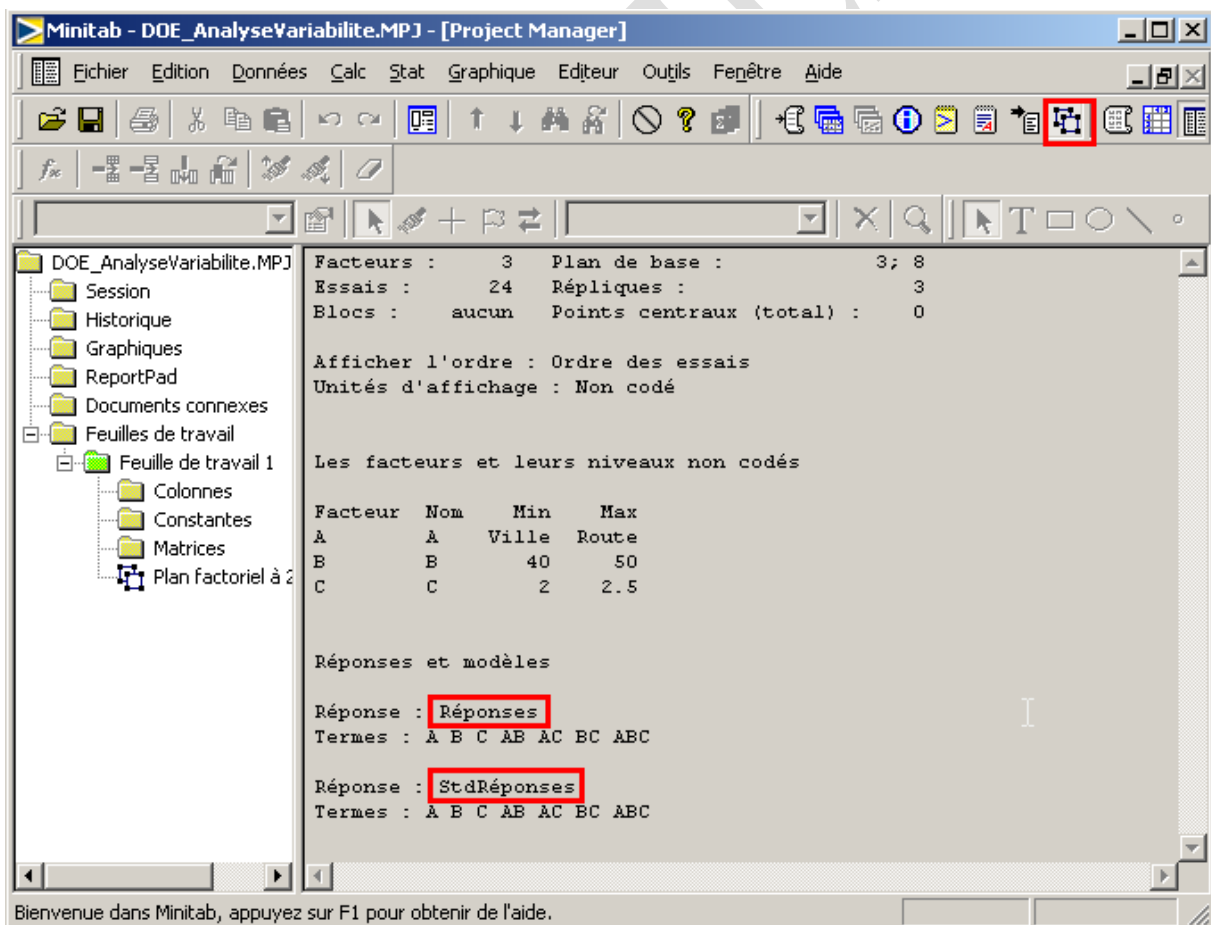
Maintenant voyons comment minimiser la variance tout en maximisant la moyenne. Nous allons pour cela dans le menu **DOE (plan d'expérience)/Plan factoriel/Analyser un plan factoriel...**:



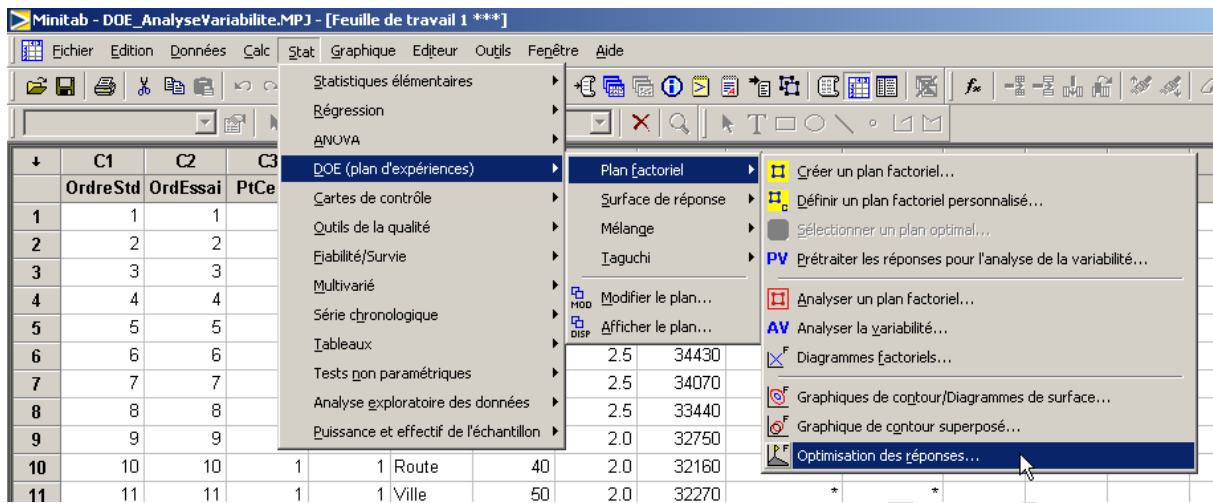
Et nous faisons cette analyse **sur les deux éléments d'intérêt: réponses et écart-type!!!**



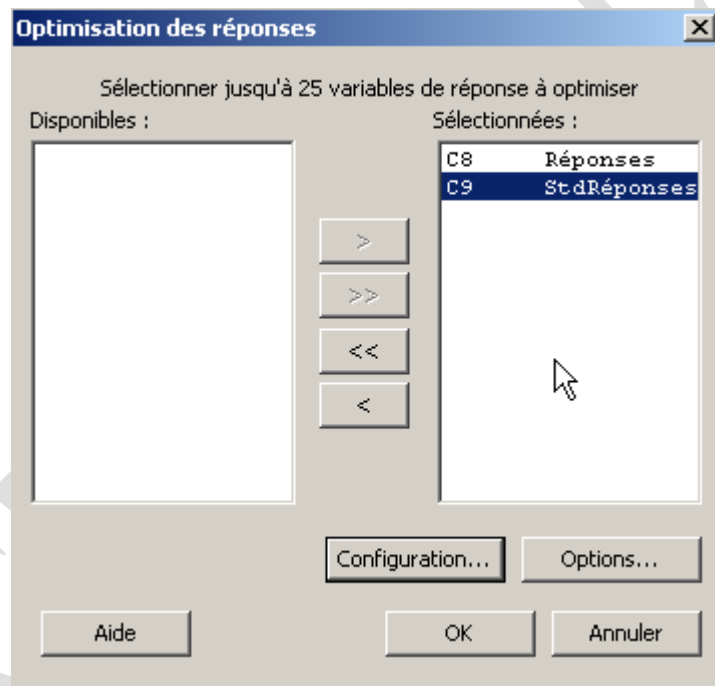
Nous validons par **OK** et allons vérifier dans le **Project Manager** que les deux analyses sont bien présentes:



Ensuite, nous pouvons enfin faire l'optimisation en allant dans le menu **DOE (plan d'expérience)/Plan factoriel/Optimiser les réponses...**:



Et nous prenons nos deux variables d'intérêt:



En validant par **OK**, nous arrivons sur la boîte de dialogue suivante et après c'est de l'art au niveau des pondérations...:

**Optimisation des réponses - Configuration**

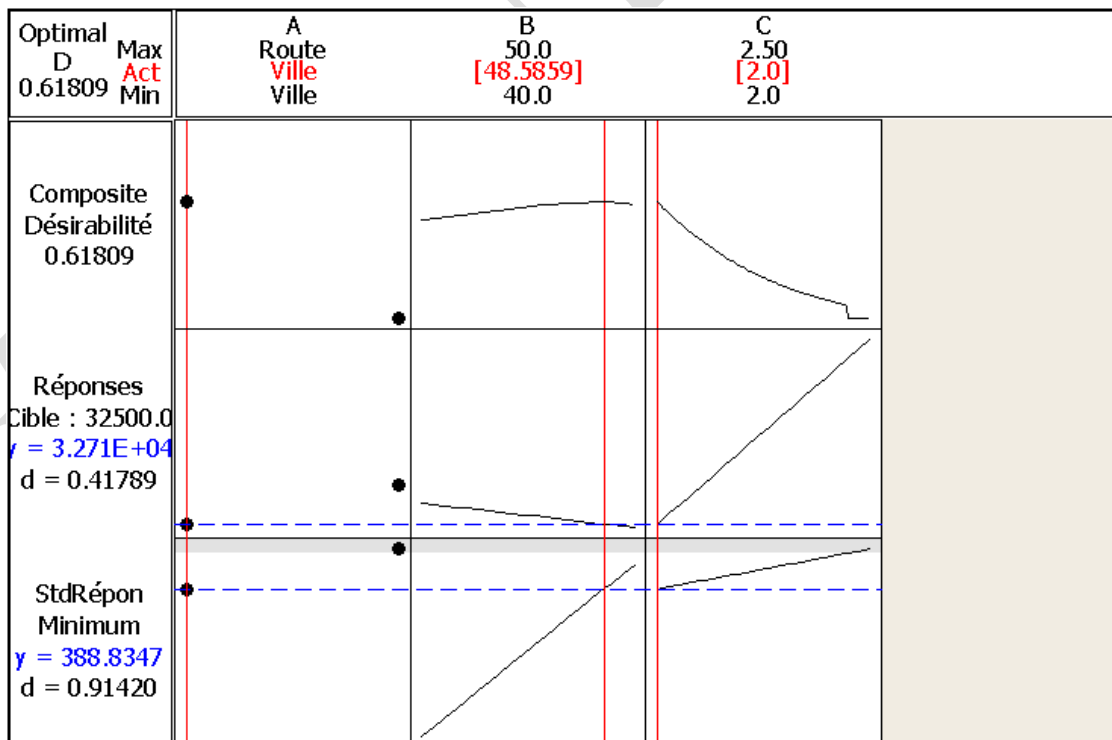
Réponse	Objectif	Inférieur	Cible	Supérieur	Pondération	Importance
C8 Réponses	Cible	30000	32500	35000	10	1
C9 StdRéponses	Réduire		300	450	0.1	1

Fonctions de désirabilité pour différents objectifs - Comment les pondérations agissent sur leurs formes

Minimiser la réponse      Atteindre une valeur cible      Maximiser la réponse

Aide      OK      Annuler

Vient alors le graph résumé:



Voilà... et y'a plus qu'à il faut qu'on...

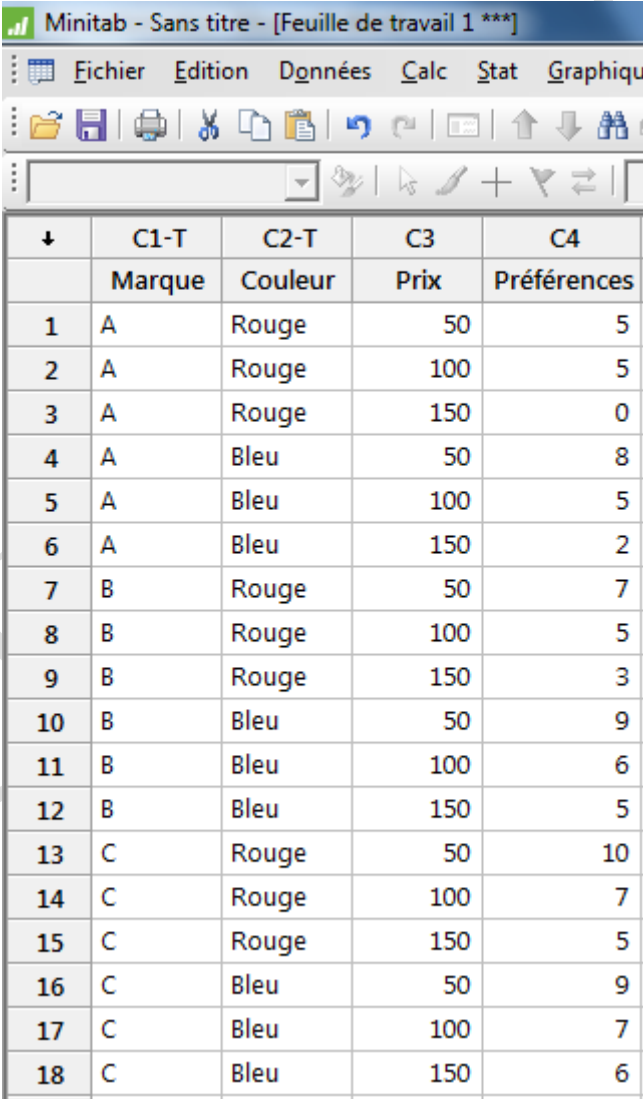
## 17.15. Exercice 186.: Analyse conjointe

Minitab® Statistical Software 17.1.2

Dans le cours théorique nous avons vu que dans le cas d'un modèle linéaire (donc où les variables ne sont pas supposées corrélées entre elles ni être quadratiques) les plans d'expérience pouvaient s'appliquer au domaine Marketing pour déterminer la combinaison de paramètres optimale d'un produit qui maximise la préférence des consommateurs.

Nous allons partir du même jeu de données que dans le cours théorique et allons voir si les résultats correspondent.

Nous avons donc des produits de plusieurs marques avec plusieurs couleurs et à plusieurs prix et nous avons noté le nombre de consommateurs qui préféraient un produit plutôt que l'autre:

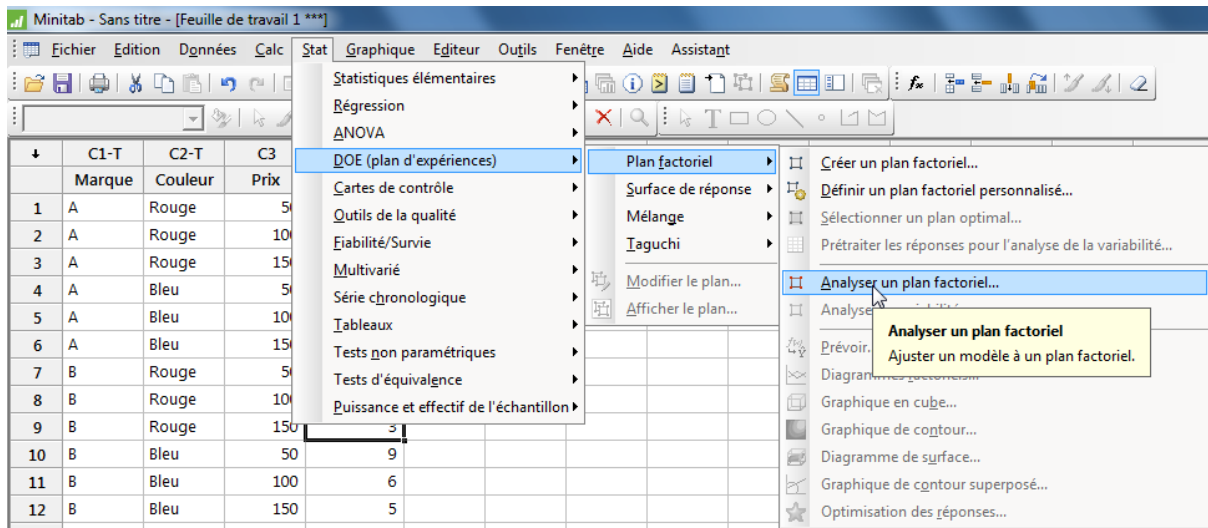


The screenshot shows the Minitab software interface with a data table. The table has 5 columns: a row indicator (1-18), 'C1-T' (Marque), 'C2-T' (Couleur), 'C3' (Prix), and 'C4' (Préférences). The data is as follows:

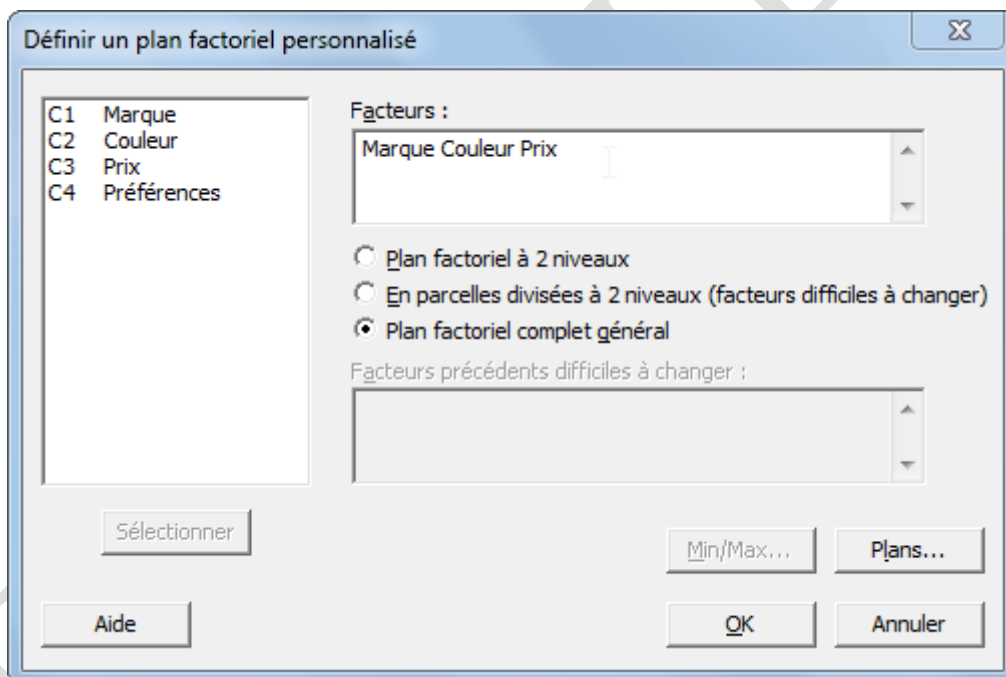
	C1-T	C2-T	C3	C4
	Marque	Couleur	Prix	Préférences
1	A	Rouge	50	5
2	A	Rouge	100	5
3	A	Rouge	150	0
4	A	Bleu	50	8
5	A	Bleu	100	5
6	A	Bleu	150	2
7	B	Rouge	50	7
8	B	Rouge	100	5
9	B	Rouge	150	3
10	B	Bleu	50	9
11	B	Bleu	100	6
12	B	Bleu	150	5
13	C	Rouge	50	10
14	C	Rouge	100	7
15	C	Rouge	150	5
16	C	Bleu	50	9
17	C	Bleu	100	7
18	C	Bleu	150	6

Il s'agit donc d'un plan factoriel complet et nous allons voir comment l'analyser (la procédure étant la même que pour les plans d'expériences habituels).

Nous allons donc d'abord dans la menu **DOE (plan d'expérience)/Plan factoriel/Définir un plan factoriel...**:

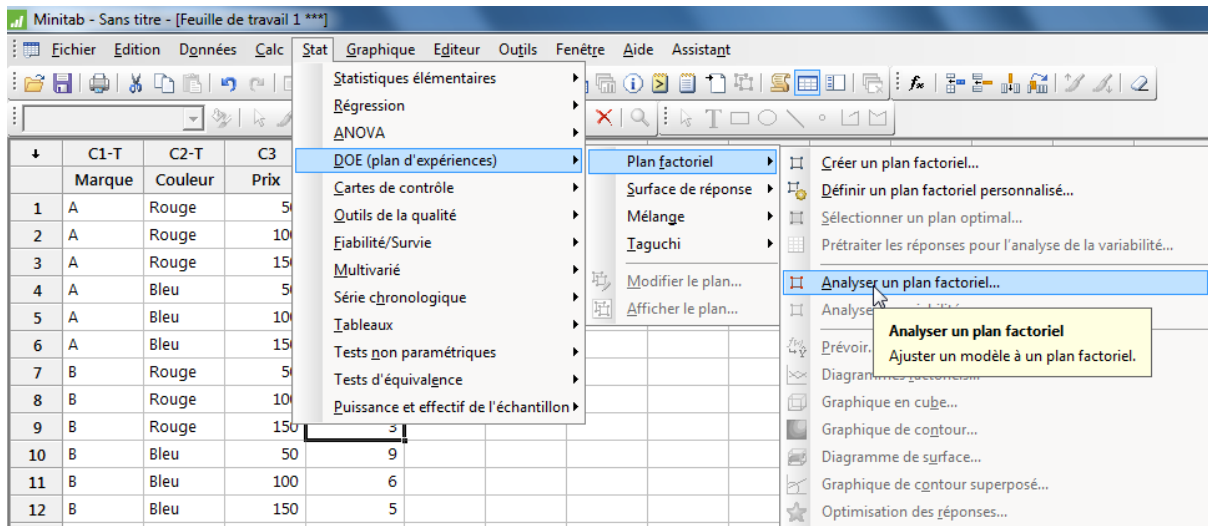


Dans la boîte de dialogue qui apparaît nous mettons:

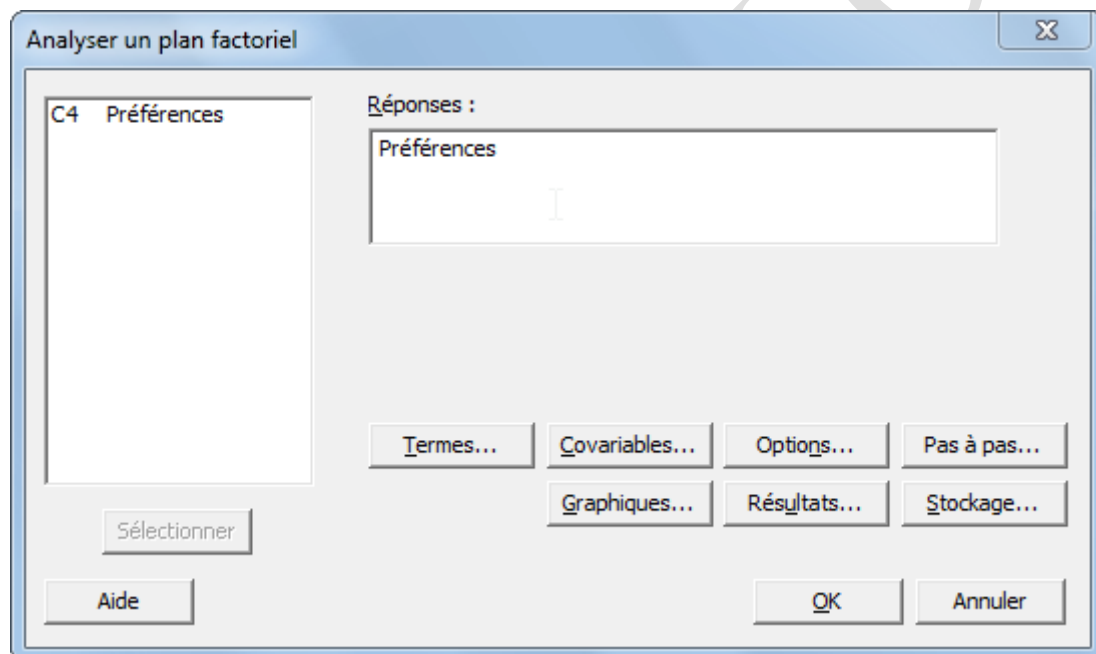


Ensuite, nous allons dans **DOE (plan d'expérience)/Plan factoriel/Analyser un plan factoriel...**:

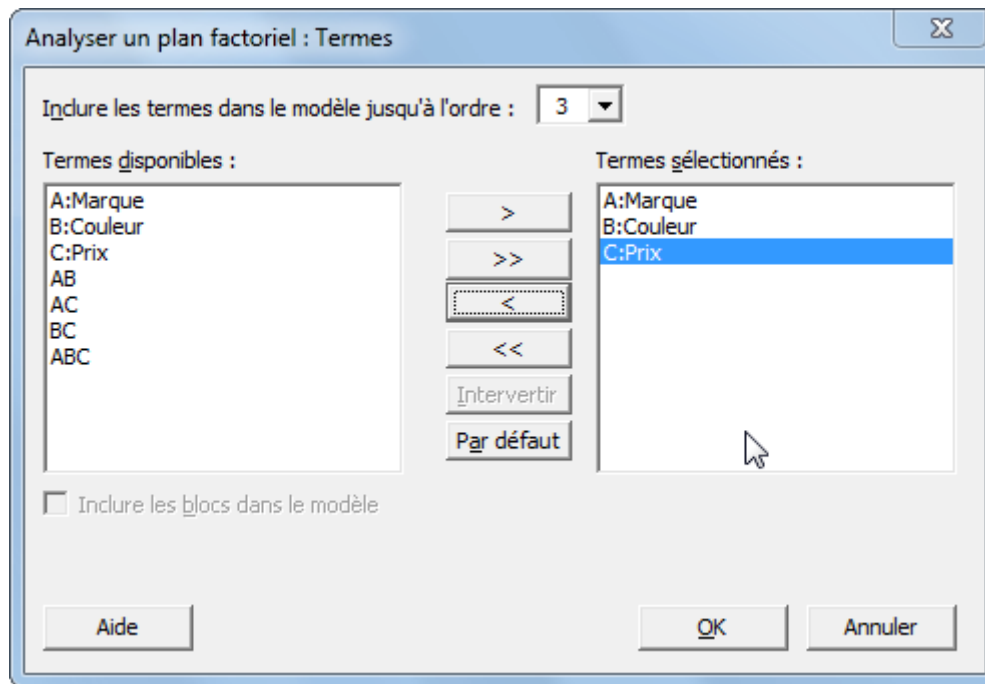




Et dans la boîte de dialogue qui apparaît, nous mettons:



Nous cliquons sur le bouton **Termes...** pour y enlever tous les termes avec interactions:



Nous validons deux fois par **OK** pour obtenir:

### Régression factorielle générale : Préférences en fonction de Marque; Couleur; Prix

#### Informations sur les facteurs

Facteur	Niveaux	Valeurs
Marque	3	A; B; C
Couleur	2	Bleu; Rouge
Prix	3	50; 100; 150

#### Analyse de variance

Source	DL	SomCar ajust	CM ajust	Valeur F	Valeur de p
Modèle	5	96.444	19.2889	21.70	0.000
Linéaire	5	96.444	19.2889	21.70	0.000
Marque	2	30.111	15.0556	16.94	0.000
Couleur	1	5.556	5.5556	6.25	0.028
Prix	2	60.778	30.3889	34.19	0.000
Erreur	12	10.667	0.8889		
Total	17	107.111			

#### Récapitulatif du modèle

S	R carré	R carré (ajust)	R carré (prév)
0.942809	90.04%	85.89%	77.59%

## Coefficients

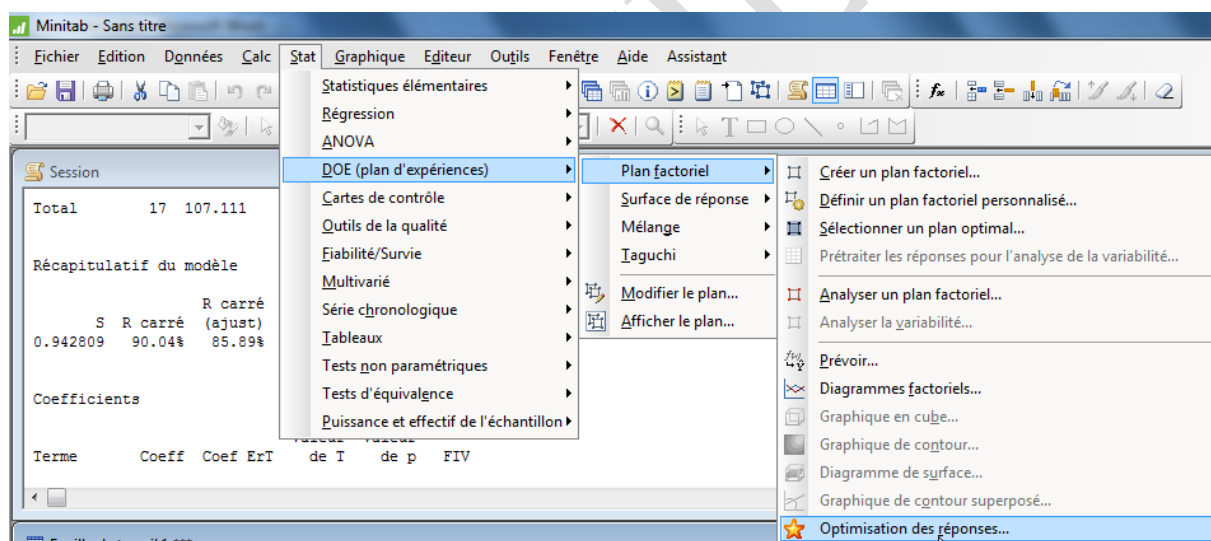
Terme	Coeff	Coef ErT	Valeur de T	Valeur de p	FIV
Constante	5.778	0.222	26.00	0.000	
Marque					
A	-1.611	0.314	-5.13	0.000	1.33
B	0.056	0.314	0.18	0.863	1.33
Couleur					
Bleu	0.556	0.222	2.50	0.028	1.00
Prix					
50	2.222	0.314	7.07	0.000	1.33
100	0.056	0.314	0.18	0.863	1.33

## Equation de régression

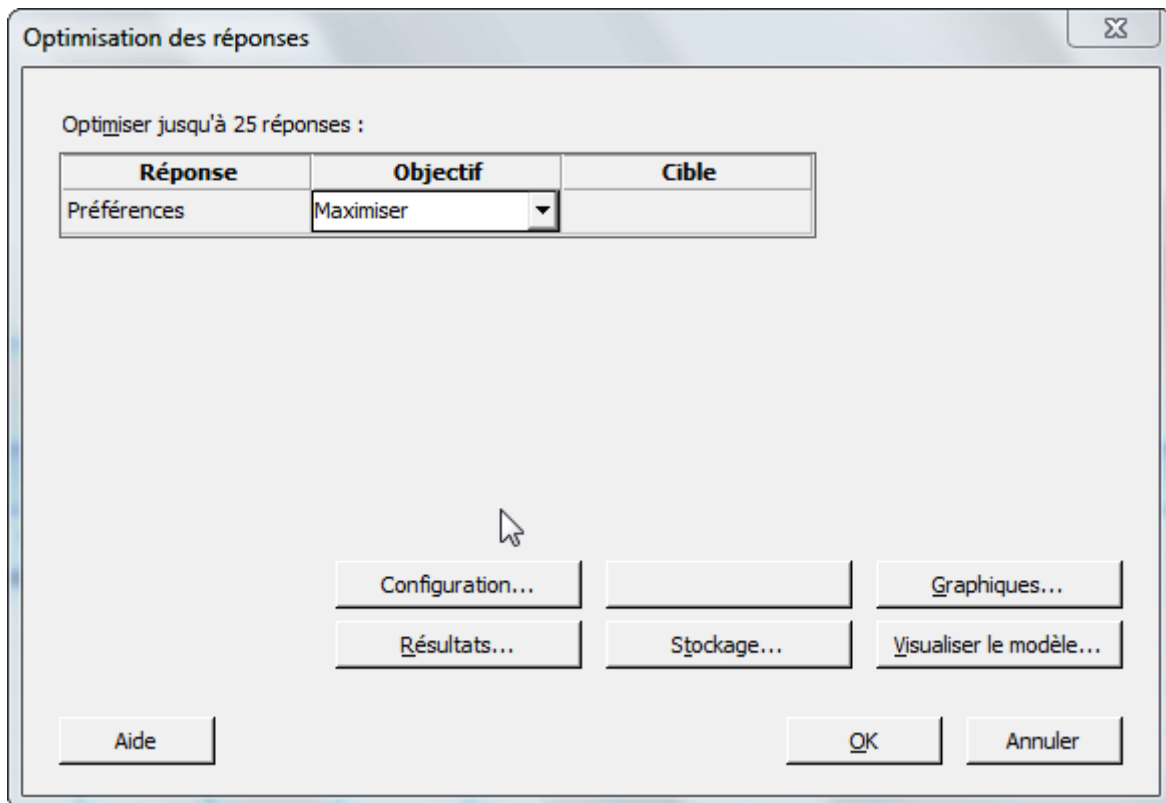
$$\text{Préférences} = 5.778 - 1.611 \text{ Marque\_A} + 0.056 \text{ Marque\_B} + 1.556 \text{ Marque\_C} + 0.556 \text{ Couleur\_Bleu} - 0.556 \text{ Couleur\_Rouge} + 2.222 \text{ Prix\_50} + 0.056 \text{ Prix\_100} - 2.278 \text{ Prix\_150}$$

Donc contrairement à R et Microsoft Excel nous obtenons beaucoup plus rapidement des détails plus précis concernant le modèle (sinon les valeurs numériques restent conformes).

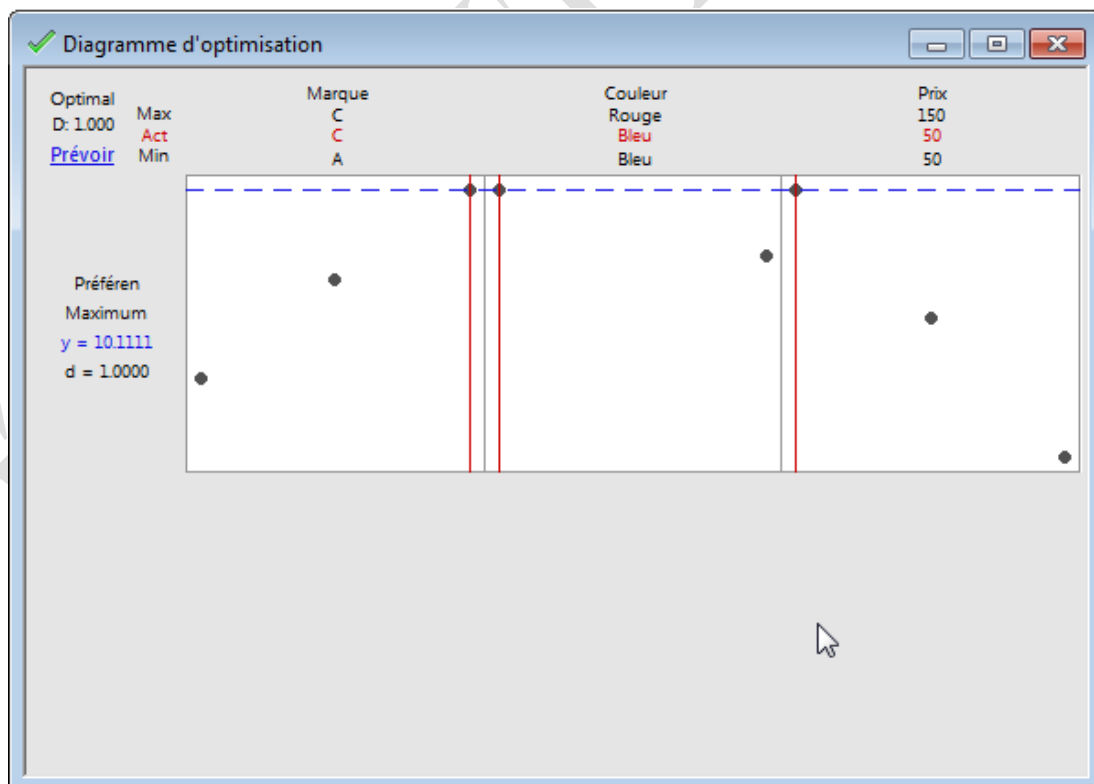
Ensuite nous pouvons utiliser l'optimiseur de Minitab en allant dans **DOE (plan d'expérience)/Plan factoriel/Optimisation des réponses**:



Nous allons bien évidemment chercher à maximiser la réponse:



Et quand nous validons par **OK** nous obtenons:



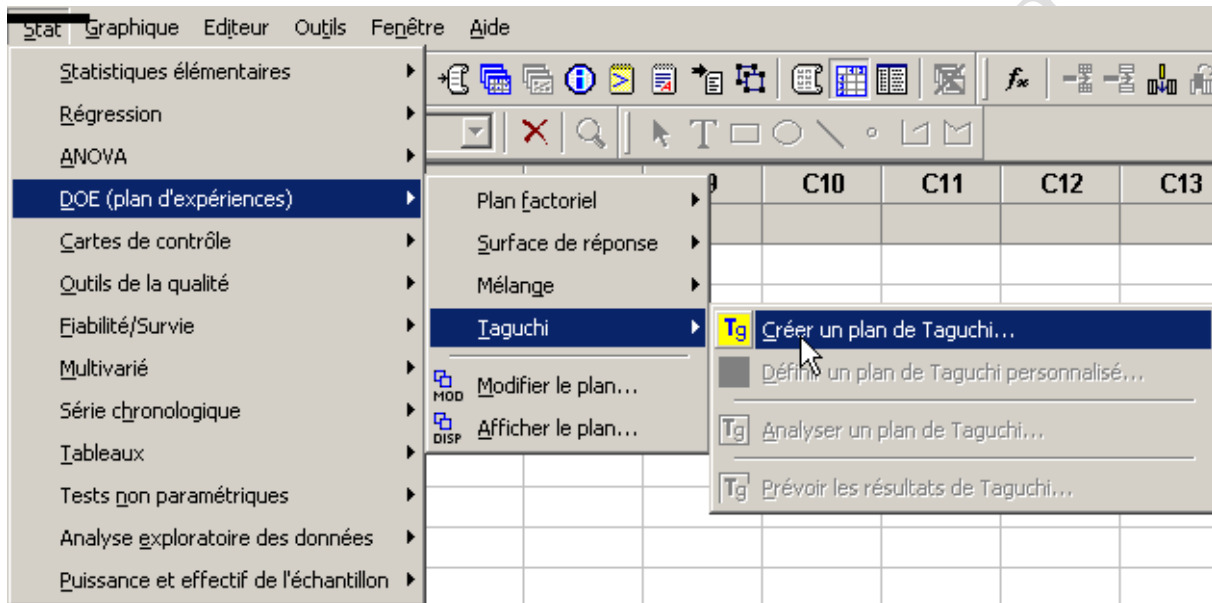
Ce qui est conformes aux résultats attendus!

## 17.16. Exercice 187.: Analyse du bruit (plan de Taguchi)

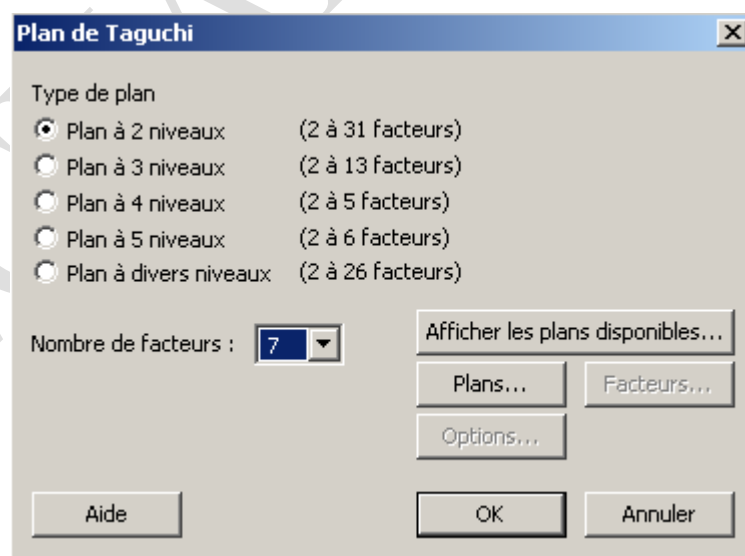
Minitab® Statistical Software 15.1.1.0

Considérons une expérience dépendante de 7 facteurs à 2 niveaux. Nous souhaitons faire une analyse du bruit de Taguchi.

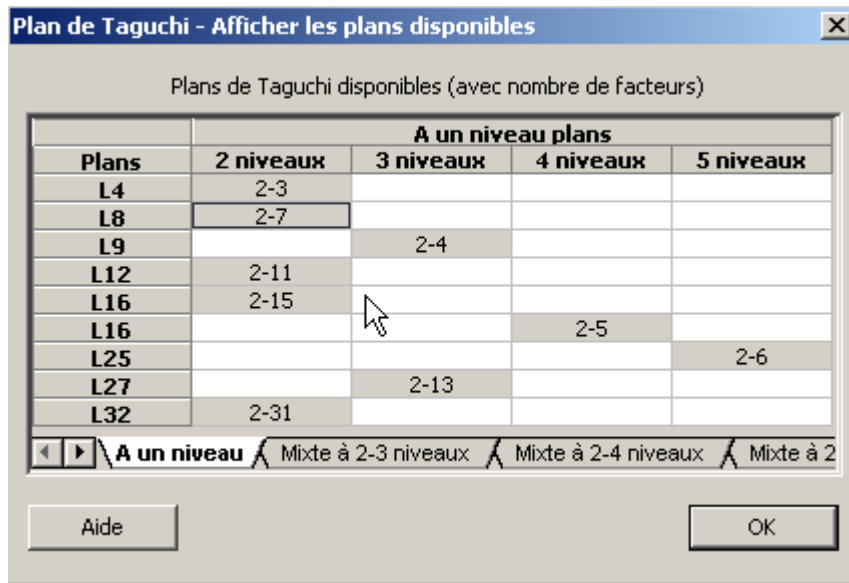
Pour cela, nous allons dans le menu **Stat/DOE (plan d'expériences)/Taguchi/Créer un plan de Taguchi...**:



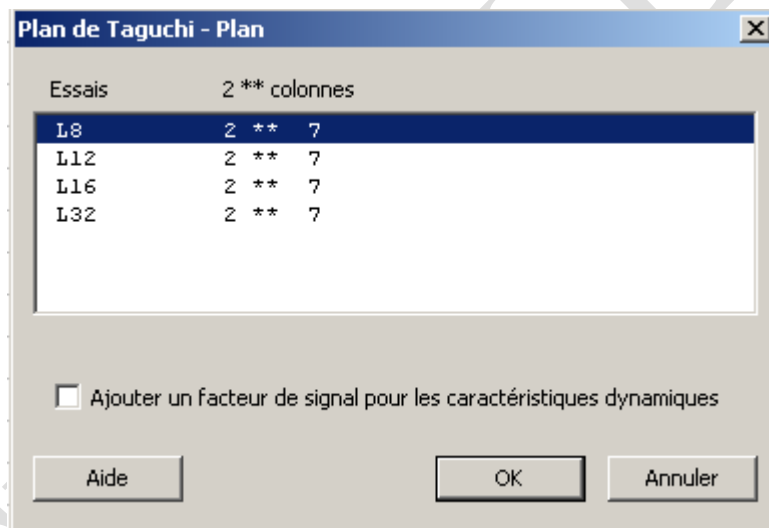
Nous prenons un 7 facteurs:



Ensuite, nous cliquons sur **Afficher les plans disponibles...**:

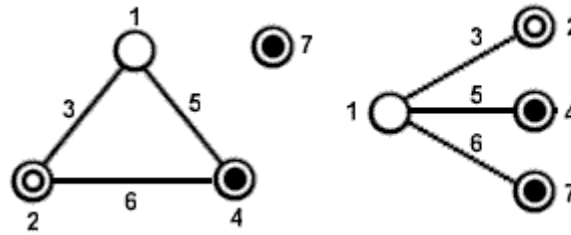


et nous prendrons donc un 2-7. Pour cela, de retour dans la boîte de dialogue précédente, nous cliquons sur le bouton **Plans...**:

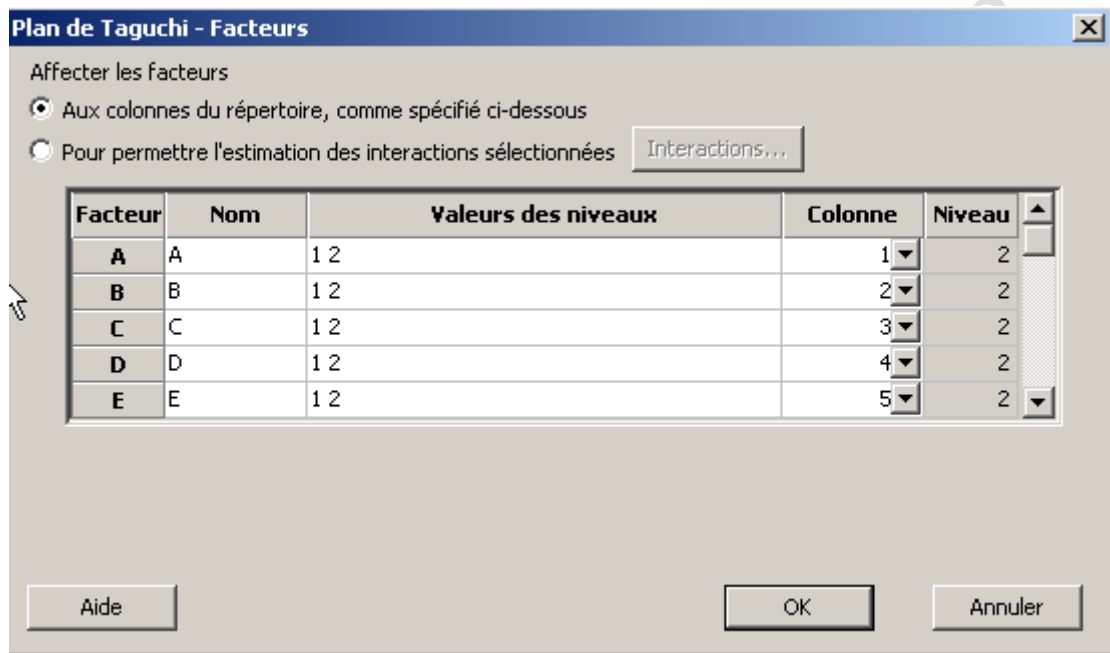


Nous prenons le L8 qui selon Taguchi est donné donc par (micro-rappel exceptionnel du cours théorique):

Essais	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2



Ensuite, nous cliquons sur le bouton **Facteurs** de la boîte de dialogue principale pour vérifier que tout est bon:



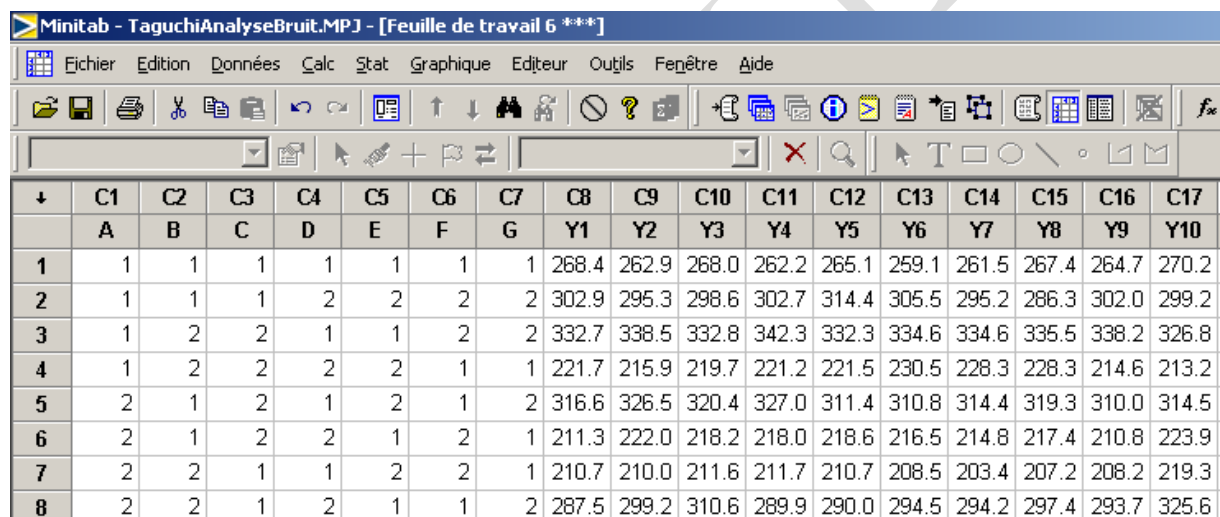
et nous validons par **OK** jusqu'à revenir à la feuille Minitab. Nous obtenons alors le *tableau d'entrée* de Taguchi:

	C1	C2	C3	C4	C5	C6	C7	C8	C9
	A	B	C	D	E	F	G		
1	1	1	1	1	1	1	1		
2	1	1	1	2	2	2	2		
3	1	2	2	1	1	2	2		
4	1	2	2	2	2	1	1		
5	2	1	2	1	2	1	2		
6	2	1	2	2	1	2	1		
7	2	2	1	1	2	2	1		
8	2	2	1	2	1	1	2		
9									
10									
11									
12									

Le *tableau de sortie* de Taguchi va contenir les réponses avec le bruit sur la base du scénario suivant (il y a donc 10 mesures dans des situations différentes pour chaque essai fonction de 7 facteurs):

Variable de réponse	Température et humidité
Y1	50° @ 15%
Y2	50° @ 50%
Y3	60° @ 15%
Y4	60° @ 50%
Y5	70° @ 15%
Y6	70° @ 50%
Y7	80° @ 15%
Y8	80° @ 50%
Y9	90° @ 15%
Y10	90° @ 50%

Nous saisissons alors les valeurs obtenues dans le cadre de l'expérience (ouvrir le fichier *TaguchiAnalyseBruit.mpj*):

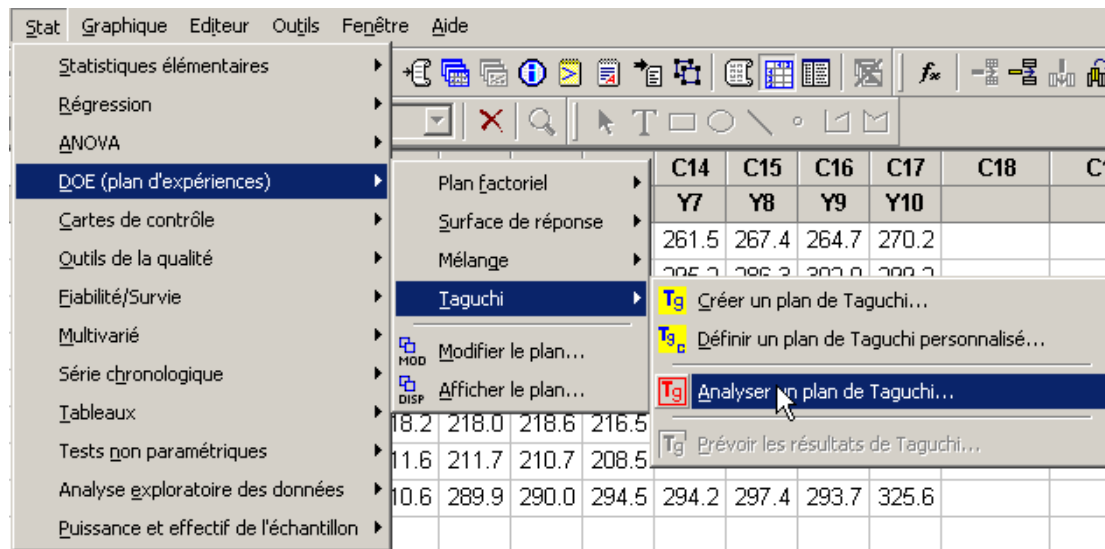


The screenshot shows the Minitab interface with a data table for a Taguchi design experiment. The table has 17 columns (C1 to C17) and 8 rows (1 to 8). The first 7 columns (C1-C7) represent factors A through G, and the last 10 columns (C8-C17) represent response variables Y1 through Y10. The data values are as follows:

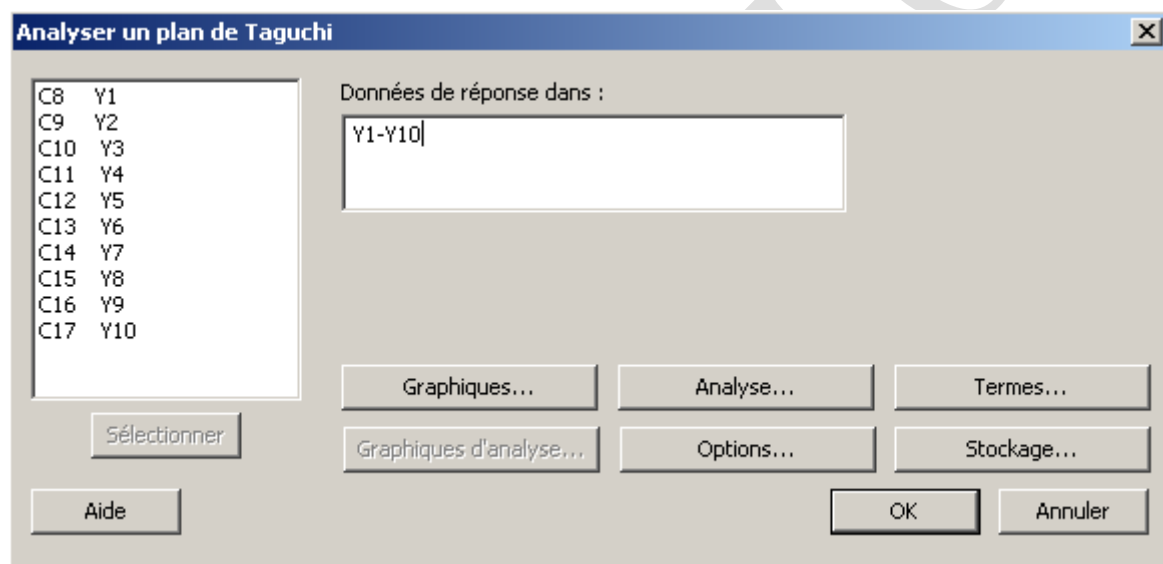
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17
	A	B	C	D	E	F	G	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10
1	1	1	1	1	1	1	1	268.4	262.9	268.0	262.2	265.1	259.1	261.5	267.4	264.7	270.2
2	1	1	1	2	2	2	2	302.9	295.3	298.6	302.7	314.4	305.5	295.2	286.3	302.0	299.2
3	1	2	2	1	1	2	2	332.7	338.5	332.8	342.3	332.3	334.6	334.6	335.5	338.2	326.8
4	1	2	2	2	2	1	1	221.7	215.9	219.7	221.2	221.5	230.5	228.3	228.3	214.6	213.2
5	2	1	2	1	2	1	2	316.6	326.5	320.4	327.0	311.4	310.8	314.4	319.3	310.0	314.5
6	2	1	2	2	1	2	1	211.3	222.0	218.2	218.0	218.6	216.5	214.8	217.4	210.8	223.9
7	2	2	1	1	2	2	1	210.7	210.0	211.6	211.7	210.7	208.5	203.4	207.2	208.2	219.3
8	2	2	1	2	1	1	2	287.5	299.2	310.6	289.9	290.0	294.5	294.2	297.4	293.7	325.6

Nous allons ensuite dans **Stat/DOE (plan d'expériences)/Taguchi/Analyser un plan de Taguchi...**:

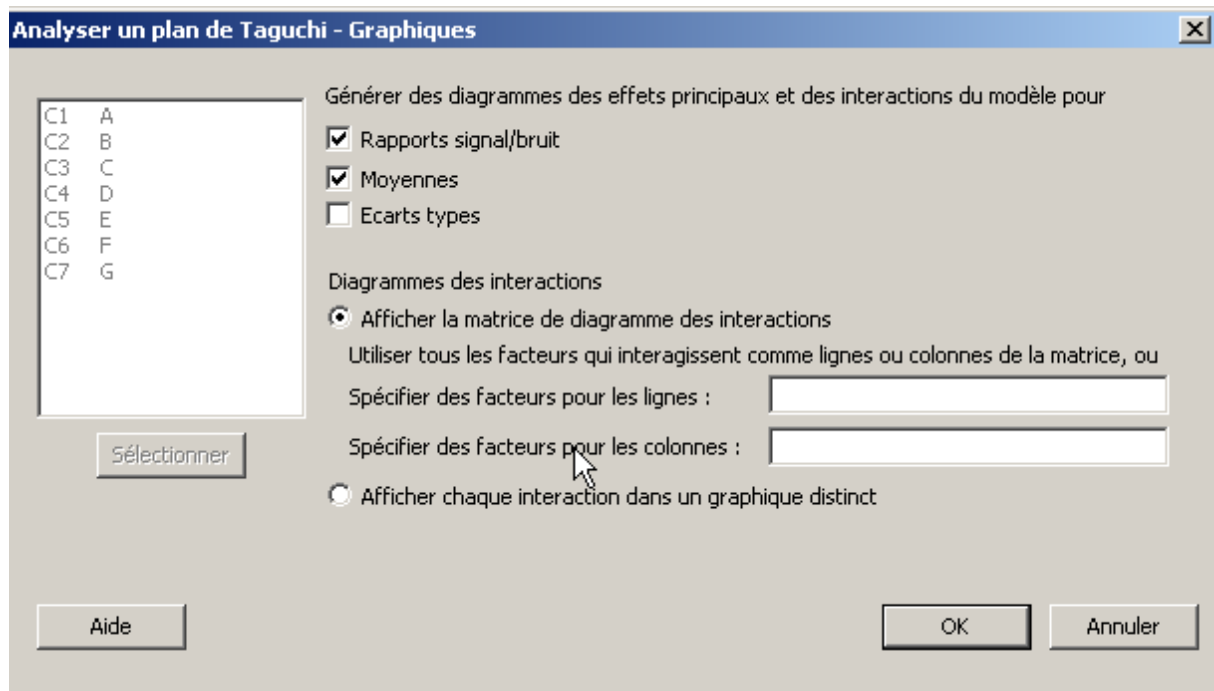




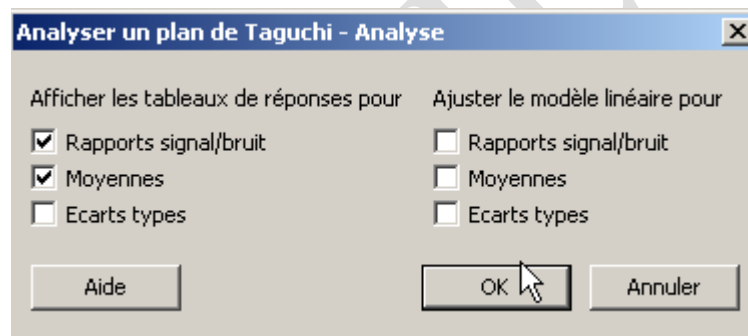
Nous avons alors:



Dans le bouton **Graphique...**, nous laissons les options par défaut:



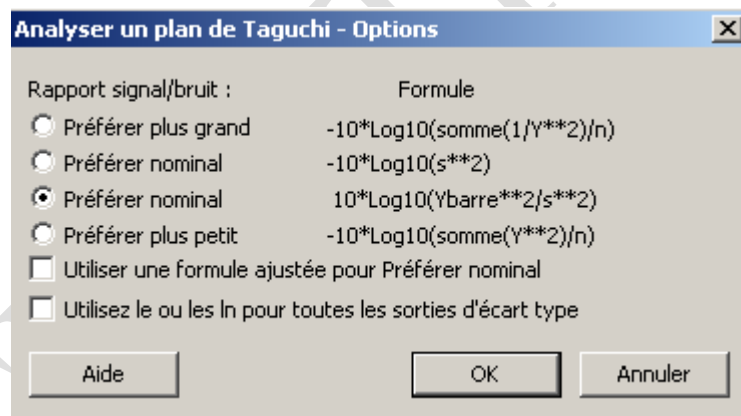
Dans le bouton **Analyse...** aussi:



Dans le bouton **Termes...** aussi (pas d'alias désirés):



Et enfin le bouton le plus important:



Et nous retrouvons ce que nous avons démontré dans le cours théorique de génie industriel pour l'option *Préférer nominal* ainsi que toutes les autres options! Nous validons le tout par **OK** autant de fois que nécessaire et nous obtenons alors dans la fenêtre de session:

**Analyse de Taguchi : Y1; Y2; Y3; Y4; Y5; ... en fonction de A; B; C; D; E; ...**

Tableau des réponses pour les rapports signal/bruit  
Préférer nominal ( $10 \cdot \text{Log}_{10}(\bar{Y}^2/s^2)$ )

Niveau	A	B	C	D	E	F	G
1	34.74	34.58	33.03	35.99	34.52	32.83	34.37
2	32.78	32.94	34.50	31.53	33.00	34.70	33.16
Delta	1.96	1.64	1.47	4.46	1.52	1.88	1.21
Rang	2	4	6	1	5	3	7

Tableau des réponses pour les moyennes

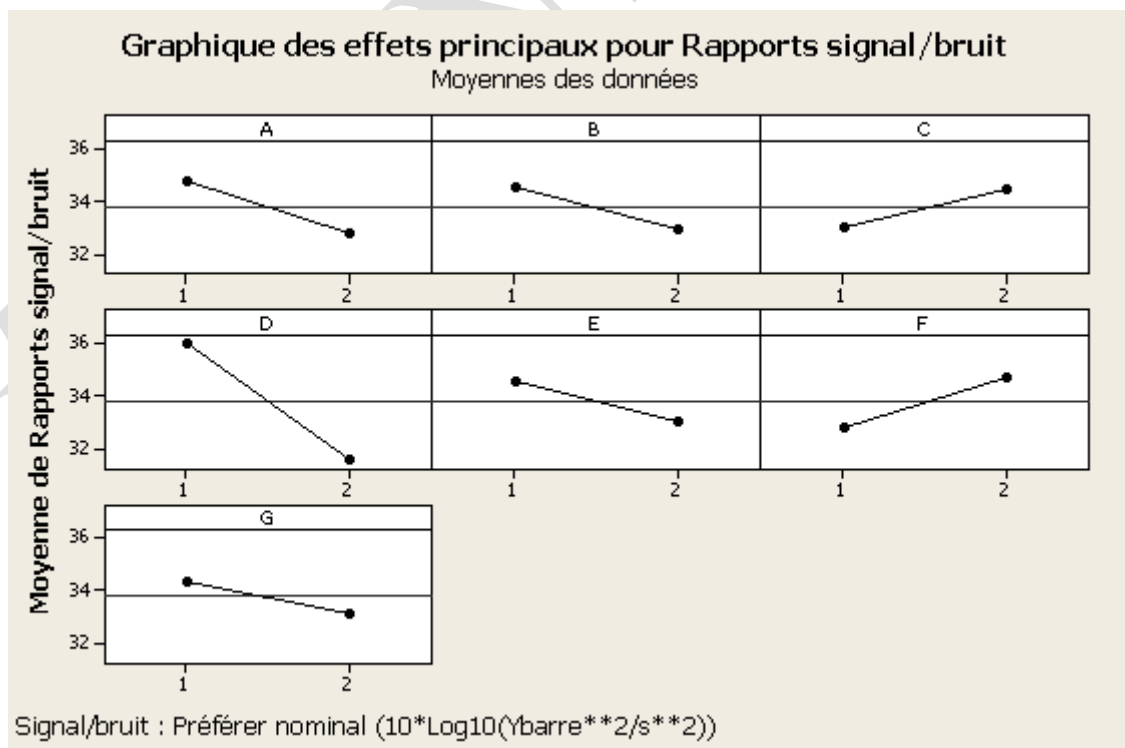
Niveau	A	B	C	D	E	F	G
1	280.4	274.8	268.4	281.8	278.8	275.4	228.4
2	260.7	266.2	272.6	259.3	262.2	265.6	312.6
Delta	19.7	8.7	4.3	22.5	16.6	9.9	84.2
Rang	3	6	7	2	4	5	1

Les questions sont alors:

1. Quel facteur contrôlé à le plus grand impact sur le ratio S/N?
2. Quel facteur contrôlé à le plus petit impact sur le ratio S/N?

**PS: Notez que Minitab met automatiquement des rangs dans chacune des tables ci-dessus...**

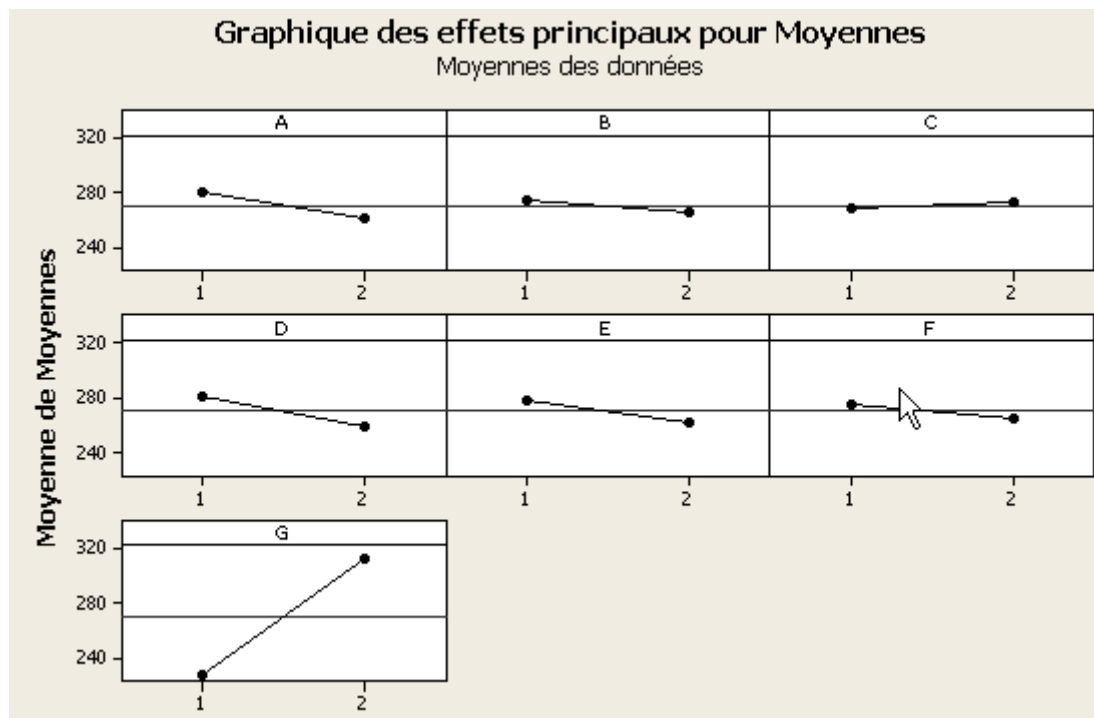
Nous pouvons arriver à la même conclusion que Minitab en analysant le graphique du rapport S/N:



Nous voyons dans le graphique ci dessus que  $A_1B_1C_2D_1E_1F_2G_1$  est la combinaison perdante (provoque le plus de bruit)... Il paraît assez évidant visuellement que c'est D qui influence le

plus le bruit et  $G$  le moins... ce qui est conforme aux rangs donné par le premier tableau dans la fenêtre de sessions.

Avec l'autre graphique, nous voyons que c'est le facteur  $G$  qui influence le plus la réponse moyenne (ce qui correspond au rang donné par Minitab dans le deuxième tableau de la session):

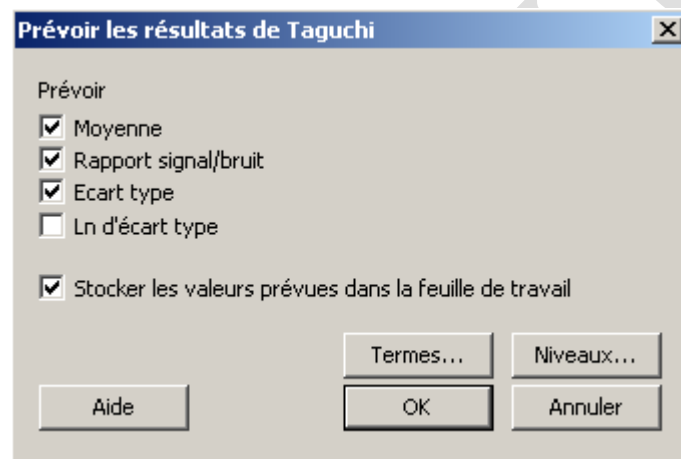
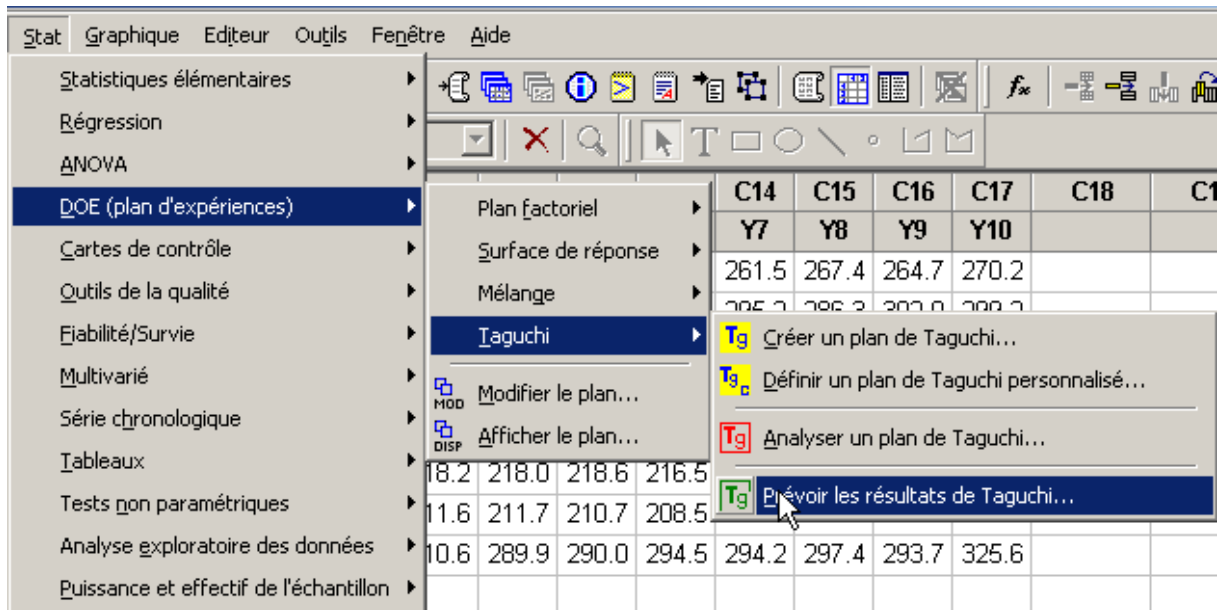


Nous y voyons aussi que la combinaison gagnante est:

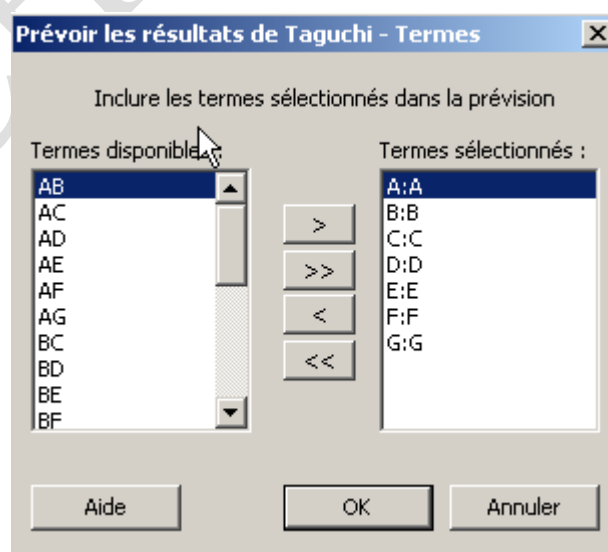
$$A_1 B_1 C_2 D_1 E_1 F_2 G_2$$

Donc finalement c'est  $G$  qui provoque le moins de bruit mais a le plus d'effet sur la réponse. Donc c'est avec ce facteur qu'il va falloir jouer le plus dans le cadre de notre expérience.

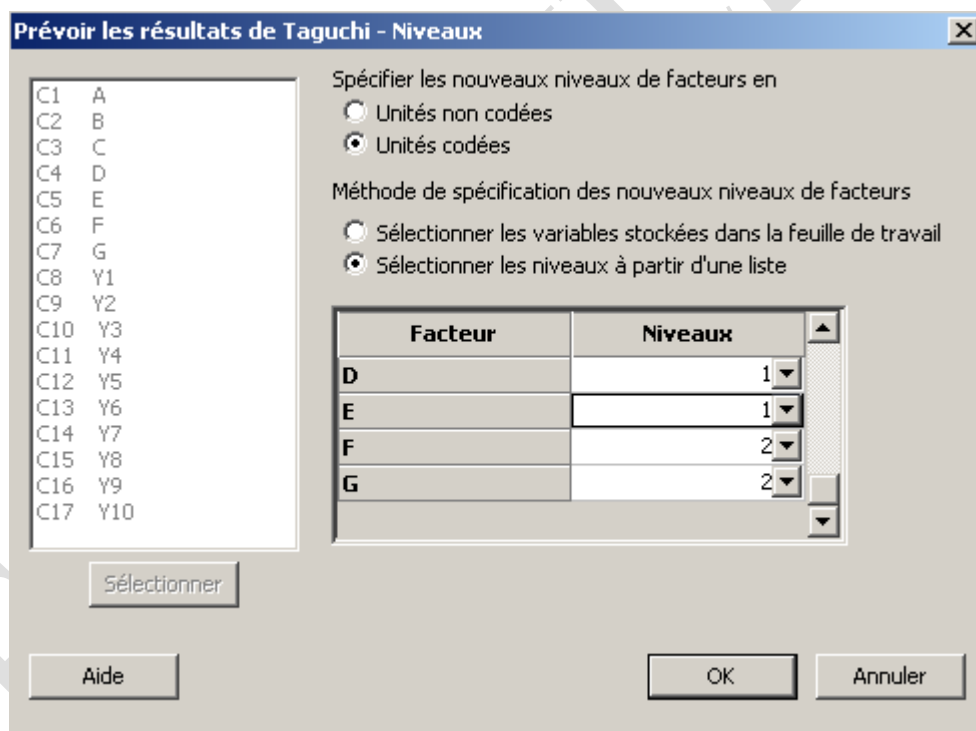
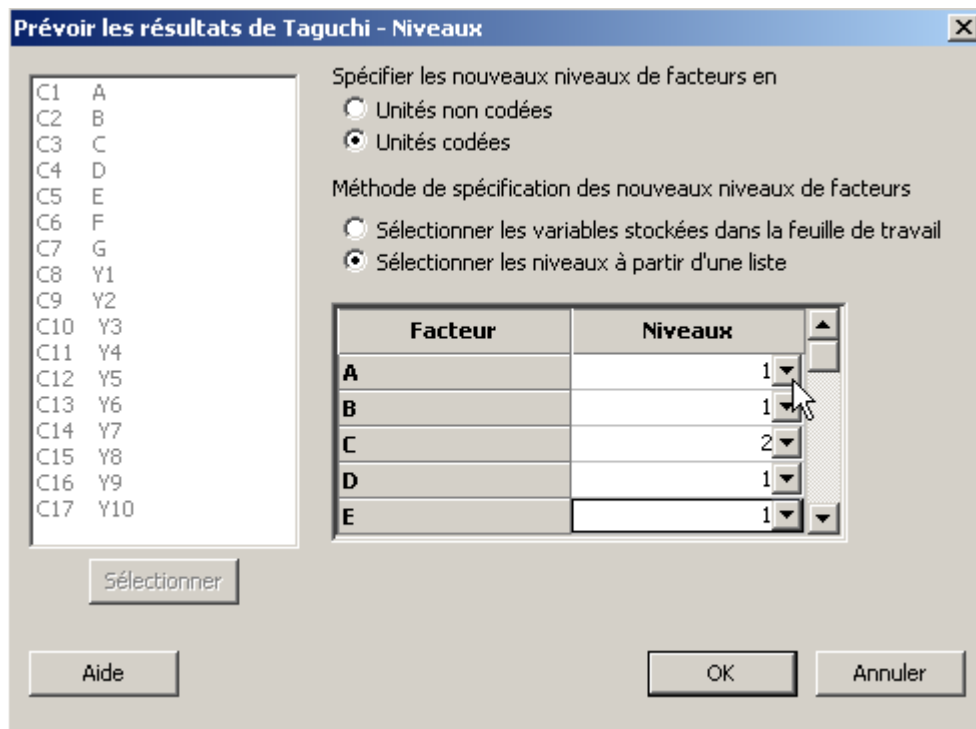
Nous allons maintenant simuler notre choix  $A_1 B_1 C_2 D_1 E_1 F_2 G_2$ . Pour cela, nous allons dans **Stat/DOE (plan d'expériences)/Taguchi/Prévoir les résultats de Taguchi..**:



Pour le bouton **Termes...** nous laissons tel quel le contenu (analyse sur tous les termes souhaitée mais sans les interactions):



Sur le bouton **Niveaux...**:



où nous avons donc fait en sorte de mettre la combinaison optimale:

$$A_1 B_1 C_2 D_1 E_1 F_2 G_2$$

Nous validons tout ce beau monde par **OK** autant de fois que nécessaire pour avoir:

**Valeurs prévues**

Rapport S/B	Moyenne	ECART TYPE
39.6181	343.503	3.05437

Niveaux de facteurs pour les prévisions

A	B	C	D	E	F	G
1	1	2	1	1	2	2



### 17.17. Exercice 188.: Plan d'expérience de Taguchi

Minitab® Statistical Software 15.1.1.0

Le but ici va être de comparer le résultat obtenu avec plan factoriel complet que nous avons fait plus haut avec la voiture mais cette fois avec le plan de Taguchi correspondant. Nous avons donc sans interactions:

$$Y = 33'119.3 + x_1 \underbrace{\begin{bmatrix} -163.4 \\ 163.4 \end{bmatrix}}_{\Sigma=0} + x_2 \underbrace{\begin{bmatrix} -78.7 \\ 78.7 \end{bmatrix}}_{\Sigma=0} + x_3 \underbrace{\begin{bmatrix} -349.7 \\ 619.4 \\ -269.7 \end{bmatrix}}_{\Sigma=0}$$

ou avec interactions (mais sans l'interaction triple):

$$y = 33119.3 + A \underbrace{\begin{bmatrix} -163.4 \\ 163.4 \end{bmatrix}}_{\Sigma=0} + B \underbrace{\begin{bmatrix} -78.7 \\ 78.7 \end{bmatrix}}_{\Sigma=0} + C \underbrace{\begin{bmatrix} -349.7 \\ 619.4 \\ -269.7 \end{bmatrix}}_{\Sigma=0} + AB \underbrace{\begin{bmatrix} -194.9 \\ 194.9 \\ 194.9 \\ -194.9 \end{bmatrix}}_{\Sigma=0} + AC \underbrace{\begin{bmatrix} 142.2 \\ -161.9 \\ 19.7 \\ -142.2 \\ 161.9 \\ -19.7 \end{bmatrix}}_{\Sigma=0} + BC \underbrace{\begin{bmatrix} 40.8 \\ 151.7 \\ -192.5 \\ -40.8 \\ -151.7 \\ 192.5 \end{bmatrix}}_{\Sigma=0}$$

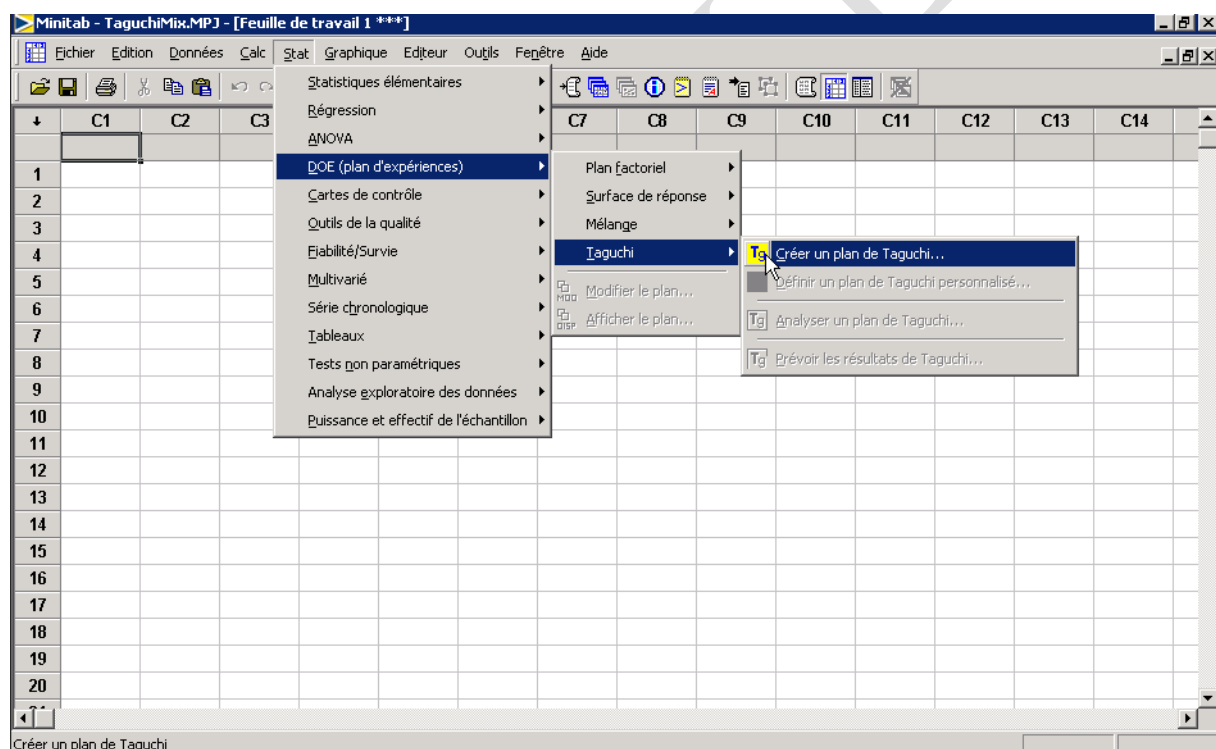
soit avec les niveaux explicites:

$$y = 33119.3 + x_1 \begin{bmatrix} x_{1,1} \\ x_{1,2} \end{bmatrix} + x_2 \begin{bmatrix} x_{2,1} \\ x_{2,2} \end{bmatrix} + x_3 \begin{bmatrix} x_{3,1} \\ x_{3,2} \\ x_{3,3} \end{bmatrix} + x_1 x_2 \begin{bmatrix} x_{1,1} x_{2,1} \\ x_{1,1} x_{2,2} \\ x_{1,2} x_{2,1} \\ x_{1,1} x_{2,2} \end{bmatrix} + x_1 x_3 \begin{bmatrix} x_{1,1} x_{3,1} \\ x_{1,1} x_{3,2} \\ x_{1,1} x_{3,3} \\ x_{1,2} x_{3,1} \\ x_{1,2} x_{3,2} \\ x_{1,2} x_{3,3} \end{bmatrix} + x_2 x_3 \begin{bmatrix} x_{2,1} x_{3,1} \\ x_{2,1} x_{3,2} \\ x_{2,1} x_{3,3} \\ x_{2,2} x_{3,1} \\ x_{2,2} x_{3,2} \\ x_{2,2} x_{3,3} \end{bmatrix}$$

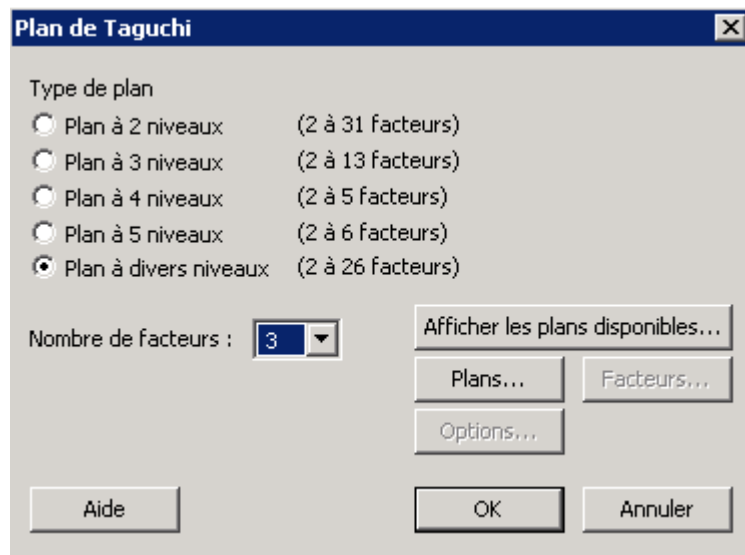
avec le tableau suivant:

N° Essai	$x_1$	$x_2$	$x_3$	Valeurs	de	y
1	1	1	1	32'700	32'750	32'960
2	1	1	2	33'430	33'360	32'910
3	1	1	3	31'710	32'100	32'220
4	1	2	1	32'680	32'270	33'130
5	1	2	2	34'070	33'100	33'610
6	1	2	3	33'220	33'700	33'285
7	2	1	1	33'180	32'160	32'640
8	2	1	2	34'430	34'280	34'460
9	2	1	3	33'570	33'300	32'570
10	2	2	1	33'270	33'080	32'415
11	2	2	2	33'440	33'570	34'204
12	2	2	3	32'840	33'210	32'470

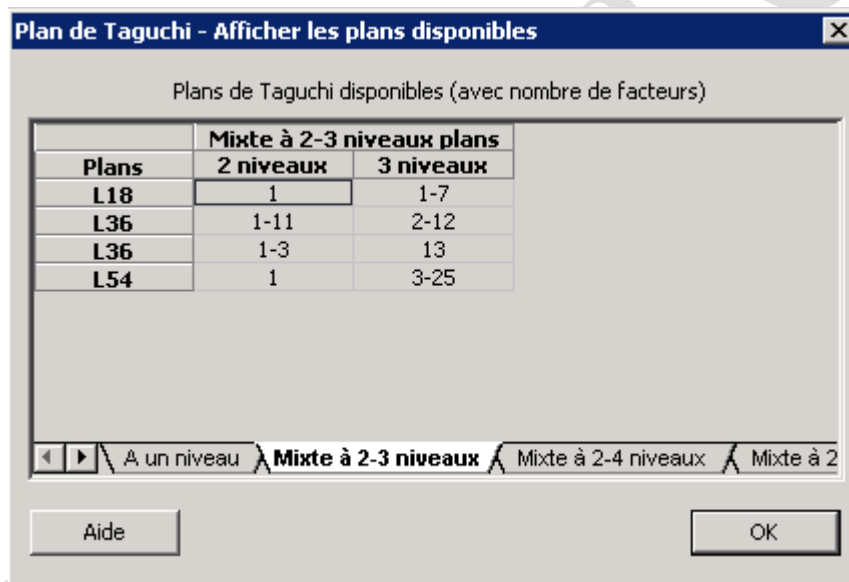
Nous allons donc dans le menu **Stat/DOE (plan d'expériences)/Taguchi/Créer un plan de Taguchi...**:



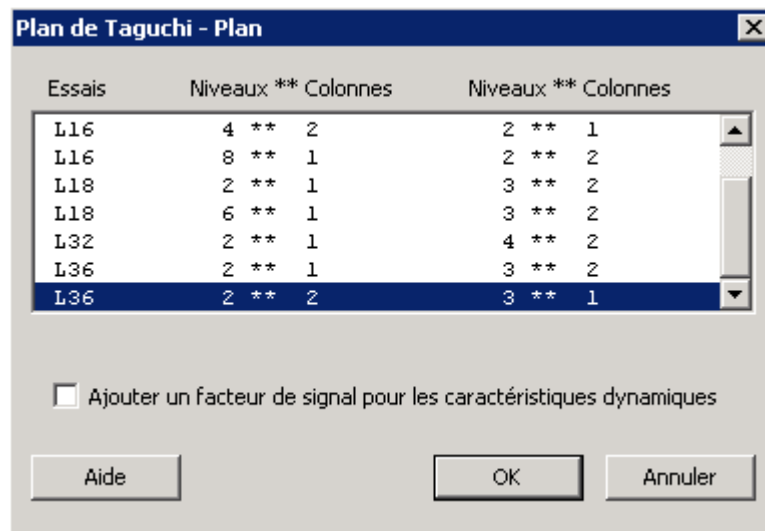
et nous prenons pour l'expérience concernée:



Si nous cliquons sur **Afficher les plans disponibles...**:

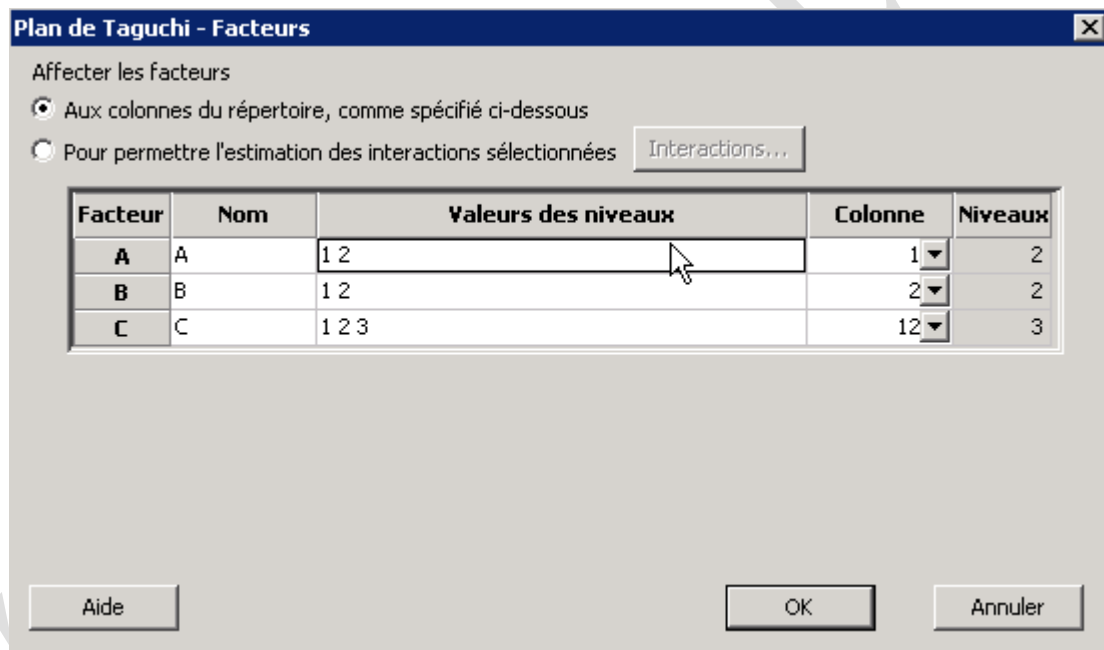


Celui qui va nous intéresser est un L36 (36 expériences, soit 3 fois 12). Donc en cliquant sur le bouton **Plans...** nous prenons 2 colonnes avec 2 niveaux et 1 colonne avec 3 niveaux puisque c'est le seul cas qui s'adapte à notre expérience:

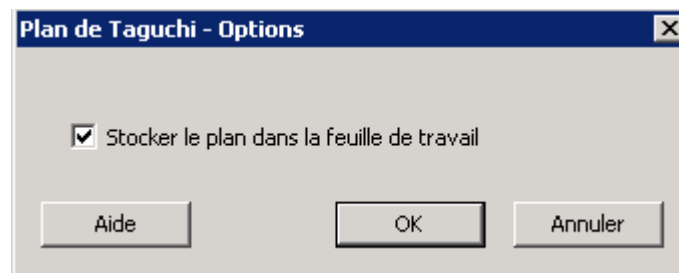


C'est-à-dire 2 colonnes à 2 niveaux et une colonne à 3 niveaux!

Ensuite, nous cliquons sur le bouton **Facteurs...**:



Ensuite, nous vérifions que Minitab mette le tableau dans la feuille:



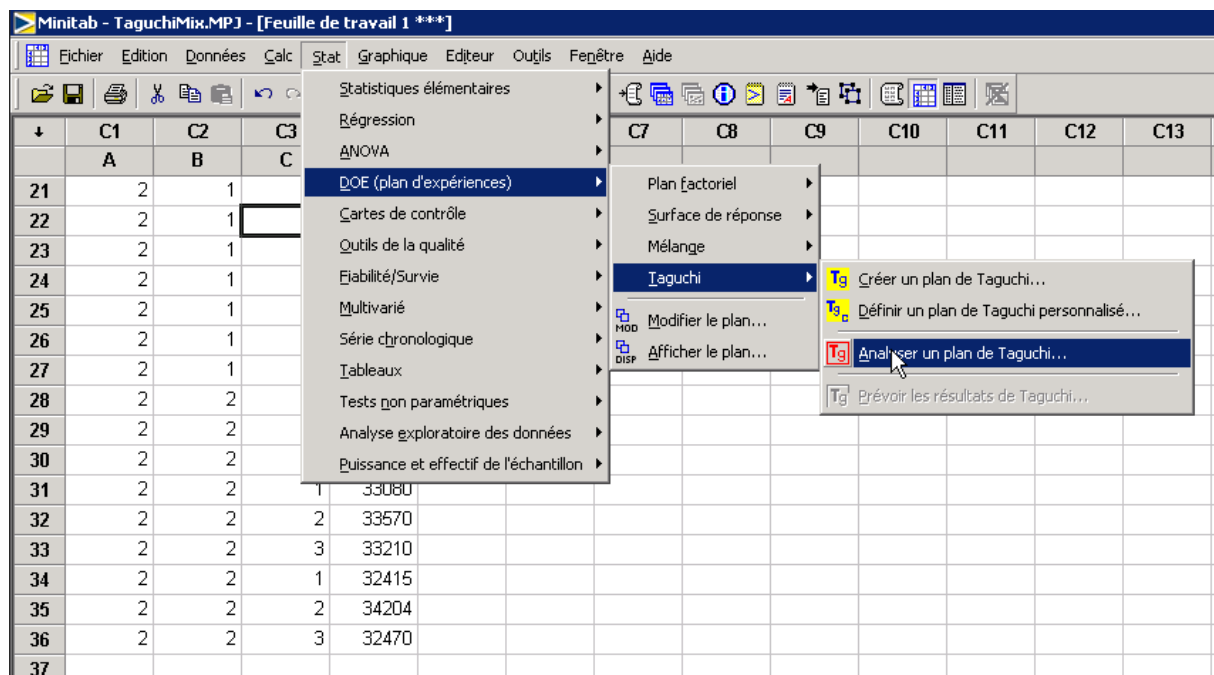
Si nous validons par **OK** deux fois, nous obtenons:

	C1 A	C2 B	C3 C	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
1	1	1	1											
2	1	1	2											
3	1	1	3											
4	1	1	1											
5	1	1	2											
6	1	1	3											
7	1	1	1											
8	1	1	2											
9	1	1	3											
10	1	2	1											
11	1	2	2											
12	1	2	3											
13	1	2	1											
14	1	2	2											
15	1	2	3											
16	1	2	1											
17	1	2	2											
18	1	2	3											
19	2	1	1											
20	2	1	2											

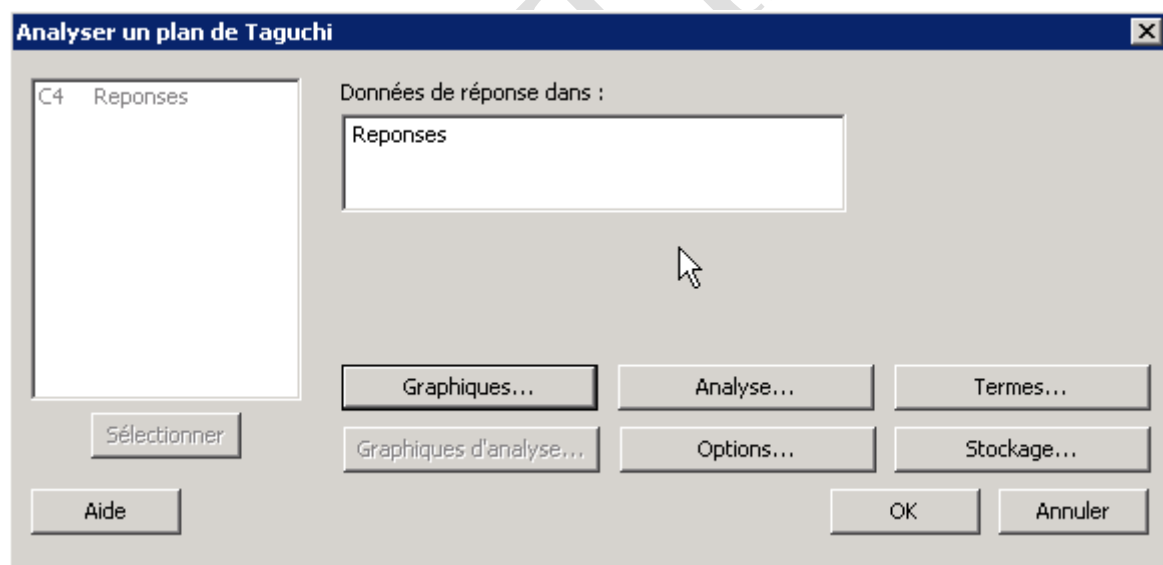
et nous y mettons les mesures de notre tableau initial:

	A	B	C	Reponses
+	C1	C2	C3	C4
1	1	1	1	32700
2	1	1	2	33430
3	1	1	3	31710
4	1	1	1	32750
5	1	1	2	33360
6	1	1	3	32100
7	1	1	1	32960
8	1	1	2	32910
9	1	1	3	32220
10	1	2	1	32680
11	1	2	2	34070
12	1	2	3	33220
13	1	2	1	32270
14	1	2	2	33100
15	1	2	3	33700
16	1	2	1	33130
17	1	2	2	33610
18	1	2	3	33285
19	2	1	1	33180
20	2	1	2	34430
21	2	1	3	33570
22	2	1	1	32160
23	2	1	2	34280
24	2	1	3	33300
25	2	1	1	32640
26	2	1	2	34460
27	2	1	3	32570
28	2	2	1	33270
29	2	2	2	33440
30	2	2	3	32840
31	2	2	1	33080
32	2	2	2	33570
33	2	2	3	33210
34	2	2	1	32415
35	2	2	2	34204
36	2	2	3	32470

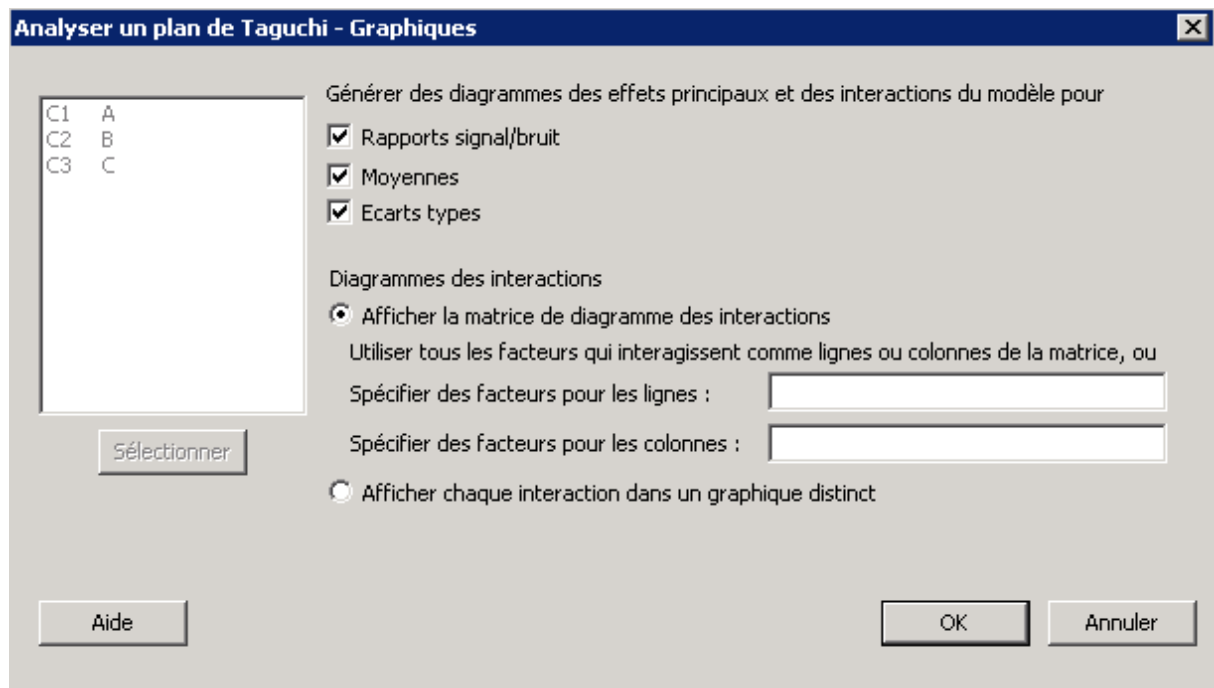
Ensuite, nous lançons l'analyse en allant dans le menu **Stat/DOE (plan d'expériences)/Taguchi/Analyser un plan de Taguchi...**:



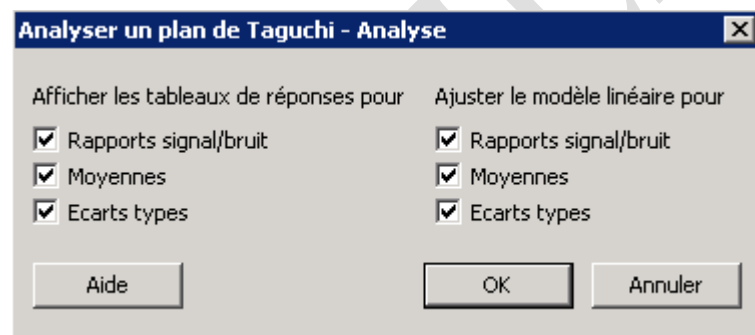
et nous y mettons:



Dans le bouton **Graphiques...** le lecteur peut tout cocher s'il le veut (comme nous l'avons fait ci-dessous) mais l'analyse graphique du rapport signal/bruit ainsi que des écarts-types n'est pas toujours aisée. Par ailleurs nous ne montrerons ici que le graphique des effets principaux et interactions des moyennes.



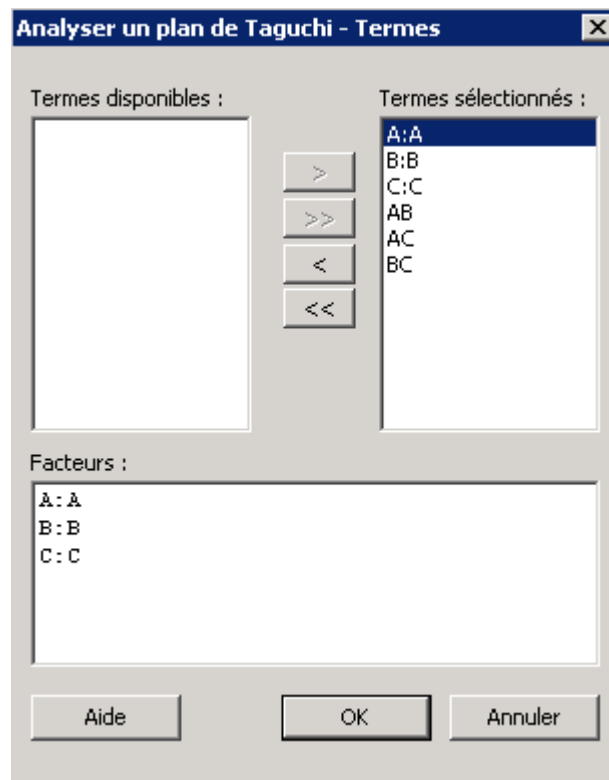
En cliquant sur le bouton **Analyse...** de la boîte de dialogue principale, nous avons:



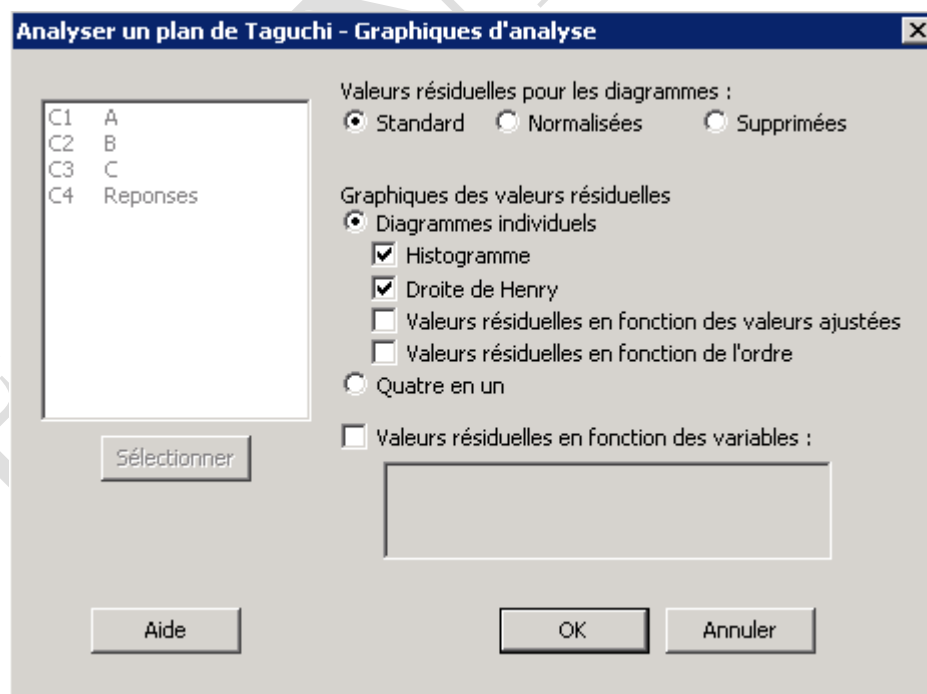
où nous avons tout coché. Mais en réalité, nous n'allons-nous concentrer encore une fois que sur les Moyennes!

En cliquant ensuite sur le bouton **Termes...** de la boîte de dialogues principale, nous allons prendre directement le modèle le plus complet avec interactions:

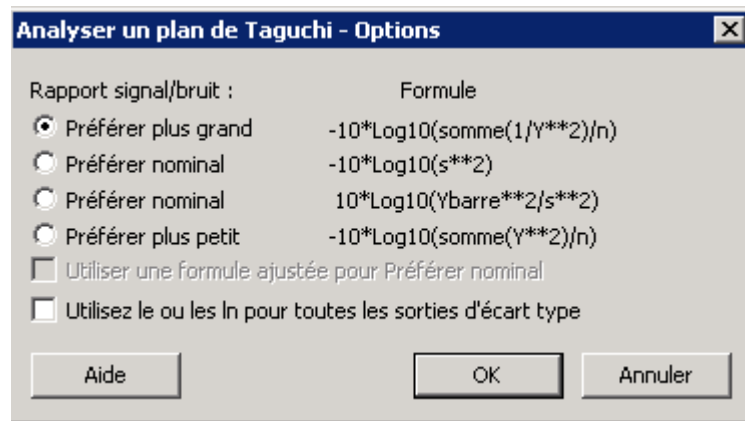




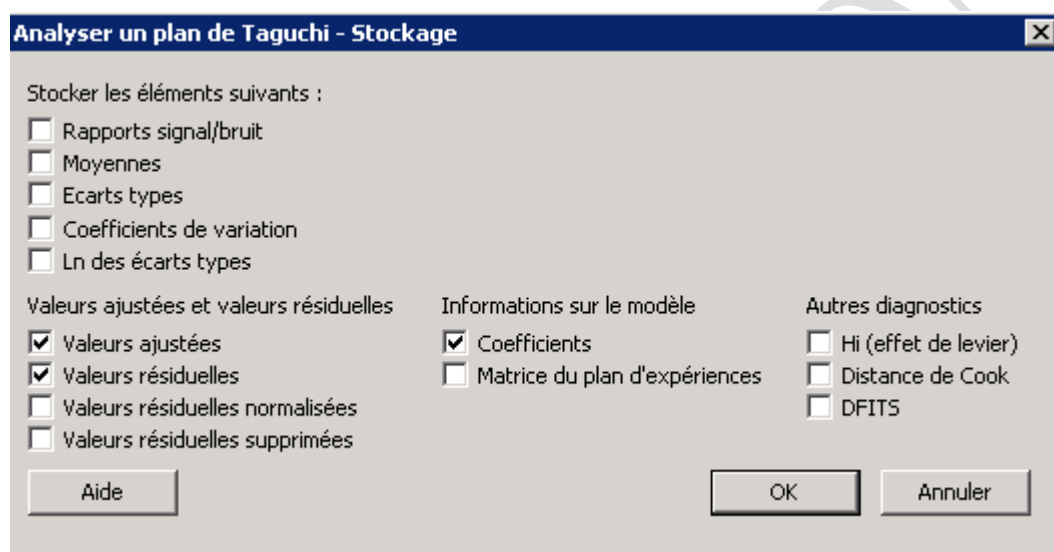
Une fois les termes choisis, nous pouvons cliquer sur le bouton **Graphiques d'Analyse...** de la boîte de dialogue principale où nous avons coché un grand nombre de graphes mais dans le présent support nous n'en utiliserons aucun (le reste étant réservé aux détails du cours théorique):



Ensuite, nous cliquons sur le bouton **Options...** de la boîte de dialogue principale pour dire que ce qui nous intéresse le plus c'est de parcourir la plus grande distance. Donc **Préférer le plus grand**:



Ensuite, dans le bouton **Stockage...** de la boîte de dialogue principale, nous prenons:



où nous ne cochoons pas **Moyennes** car nous souhaitons montrer ultérieurement une autre façon des les ajouter après coup. En validant tout cela par **OK**, nous obtenons dans la fenêtre de session:

Source	DL	Somcar seq	Somcar ajust	CM ajust	F	P
A	1	0.021641	0.021641	0.021641	1.11	0.403
B	1	0.005387	0.005387	0.005387	0.28	0.652
C	2	0.157956	0.157956	0.078978	4.05	0.198
A*B	1	0.030903	0.030903	0.030903	1.59	0.335
A*C	2	0.012523	0.012523	0.006262	0.32	0.757
B*C	2	0.017436	0.017436	0.008718	0.45	0.691
Erreur résiduelle	2	0.038992	0.038992	0.019496		
Total	11	0.284838				

Analyse de la variance pour Rapports signal/bruit

S = 0.1396 R carré = 86.3 % R carré (ajust) = 24.7 %

Terme	Coef	Coef ERT	T	P
Constante	90.3991	0.04031	2242.753	0.000
A 1	-0.0425	0.04031	-1.054	0.403
B 1	-0.0212	0.04031	-0.526	0.652
C 1	-0.0911	0.05700	-1.598	0.251
C 2	0.1618	0.05700	2.839	0.105
A*B 1 1	-0.0507	0.04031	-1.259	0.335
A*C 1 1	0.0374	0.05700	0.656	0.579
A*C 1 2	-0.0414	0.05700	-0.727	0.543
B*C 1 1	0.0114	0.05700	0.199	0.860
B*C 1 2	0.0400	0.05700	0.701	0.556

Coefficients du modèle estimés pour Rapports signal/bruit

Analyse du modèle linéaire : Rapports signal/bruit en fonction de A; B; C

Analyse de Taguchi : Réponses en fonction de A; B; C

1 2 12

Colonnes de L36(2\*\*11 3\*\*12) Répertoire

Essais : 36

Facteurs : 3

L36(2\*\*2 3\*\*1)

Plan de Taguchi orthogonal

Plan de Taguchi

Bienvenue dans Minitab, appuyez sur F1 pour obtenir de l'aide.

**Analyse du modèle linéaire : Moyennes en fonction de A; B; C**

Coefficients du modèle estimés pour Moyennes

Terme	Coeff	Coef ErT	T	P
Constante	33119.3	152.2	217.620	0.000
A 1	-163.4	152.2	-1.074	0.395
B 1	-78.7	152.2	-0.517	0.656
C 1	-349.7	215.2	-1.625	0.246
C 2	619.4	215.2	2.878	0.103
A*B 1 1	-194.9	152.2	-1.281	0.329
A*C 1 1	142.2	215.2	0.661	0.577
A*C 1 2	-161.9	215.2	-0.752	0.530
B*C 1 1	40.8	215.2	0.190	0.867
B*C 1 2	151.7	215.2	0.705	0.554

S = 527.2 R carré = 86.6 % R carré (ajust) = 26.4 %

Analyse de la variance pour Moyennes

Source	DL	SomCar séq	SomCar ajust	CM ajust	F	P
A	1	320569	320569	320569	1.15	0.395
B	1	74366	74366	74366	0.27	0.656
C	2	2314656	2314656	1157328	4.16	0.194
A*B	1	455780	455780	455780	1.64	0.329
A*C	2	187261	187261	93630	0.34	0.748
B*C	2	247007	247007	123503	0.44	0.692
Erreur résiduelle	2	555874	555874	277937		
Total	11	4155512				

Il est intéressant déjà de comparer les coefficients mis en évidence ci-dessus avec le modèle théorique obtenu avec le plan factoriel complet avec interactions:

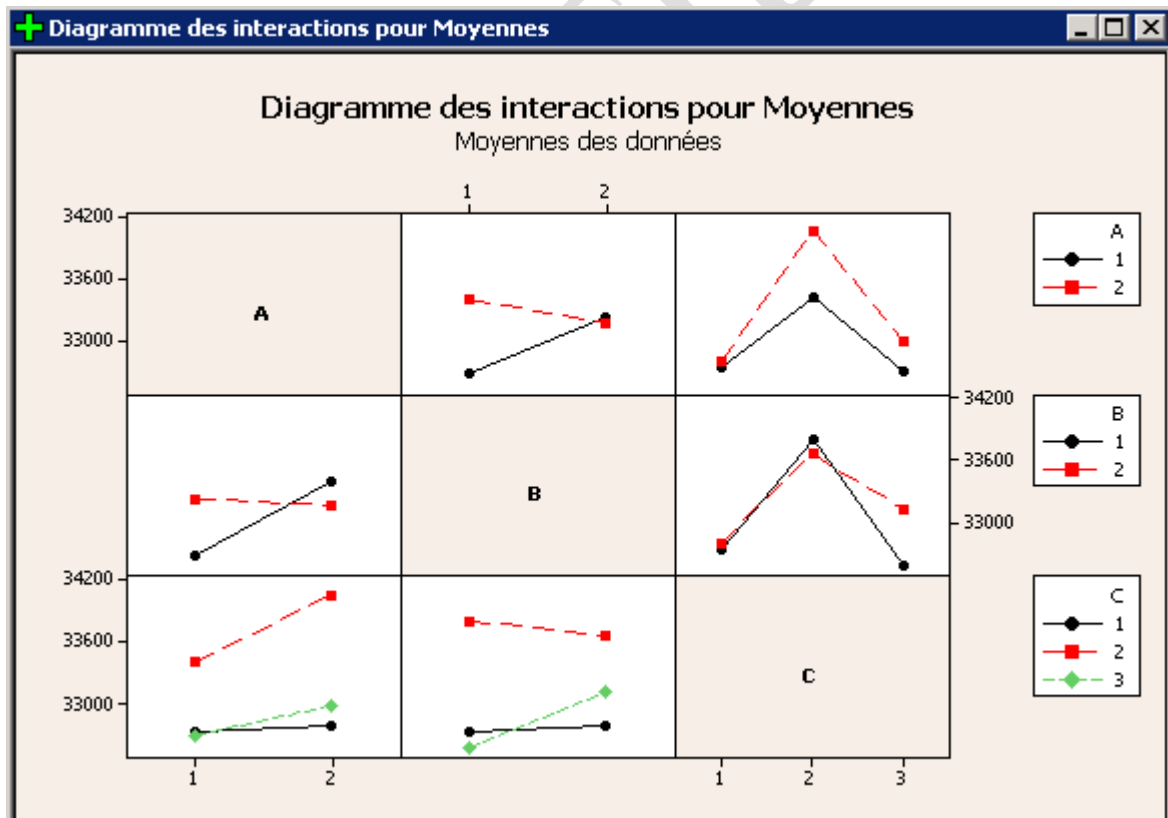
$$y = 33119.3 + A \begin{bmatrix} -163.4 \\ 163.4 \end{bmatrix} + B \begin{bmatrix} -78.7 \\ 78.7 \end{bmatrix} + C \begin{bmatrix} -349.7 \\ 619.4 \\ -269.7 \end{bmatrix} + AB \begin{bmatrix} -194.9 \\ 194.9 \\ 194.9 \\ -194.9 \end{bmatrix} + AC \begin{bmatrix} 142.2 \\ -161.9 \\ 19.7 \\ -142.2 \\ 161.9 \\ -19.7 \end{bmatrix} + BC \begin{bmatrix} 40.8 \\ 151.7 \\ -192.5 \\ -40.8 \\ -151.7 \\ 192.5 \end{bmatrix}$$

$\underbrace{\hspace{1.5cm}}_{\Sigma=0}$     
 $\underbrace{\hspace{1.5cm}}_{\Sigma=0}$     
 $\underbrace{\hspace{1.5cm}}_{\Sigma=0}$     
 $\underbrace{\hspace{1.5cm}}_{\Sigma=0}$     
 $\underbrace{\hspace{1.5cm}}_{\Sigma=0}$     
 $\underbrace{\hspace{1.5cm}}_{\Sigma=0}$

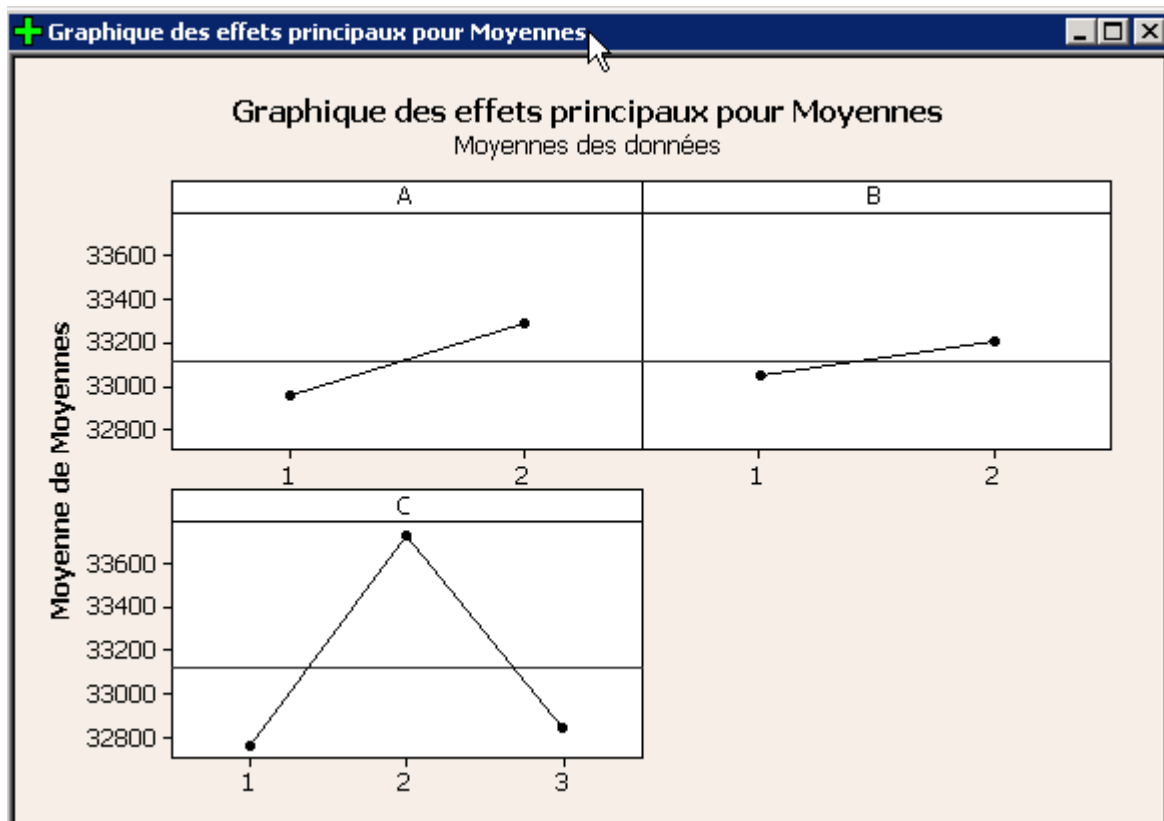
Nous voyons donc que nous obtenons bien la majorité des coefficients du modèle fait avec le plan factoriel complet et conforme à ce que nous avons fait à la main dans le cours théorique. Nous retrouvons ces coefficients dans la colonne C11:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	A	B	C	Reponses	AJUS MOY1	AJUS ECTY1	AJUS SB1	RES MOY1	RES ECTYP1	RES SB1	COEFF MOY1	COEFF ETYP1	COEFF SB1
1	1	1	1	32700	32515.5	194.236	90.2424	287.806	-56.2745	0.0758283	33119.3	351.196	90.3991
2	1	1	2	33430	33291.4	222.757	90.4451	-58.111	59.4371	-0.0142166	-163.4	-40.768	-0.0425
3	1	1	3	31710	32239.7	269.808	90.1667	-229.694	-3.1626	-0.0616116	-78.7	-49.386	-0.0212
4	1	1	1	32750	*	*	*	*	*	*	-349.7	30.644	-0.0911
5	1	1	2	33360	*	*	*	*	*	*	619.4	-33.043	0.1618
6	1	1	3	32100	*	*	*	*	*	*	-194.9	-32.109	-0.0507
7	1	1	1	32960	*	*	*	*	*	*	142.2	-57.015	0.0374
8	1	1	2	32910	*	*	*	*	*	*	-161.9	106.319	-0.0414
9	1	1	3	32220	*	*	*	*	*	*	40.8	-8.327	0.0114
10	1	2	1	32680	32981.1	373.880	90.3635	-287.806	56.2745	-0.0758283	151.7	-79.453	0.0400
11	1	2	2	34070	33535.2	544.652	90.5090	58.111	-59.4371	0.0142166			
12	1	2	3	33220	33172.0	257.238	90.4132	229.694	3.1626	0.0616116			
13	1	2	1	32270	*	*	*	*	*	*			
14	1	2	2	33100	*	*	*	*	*	*			
15	1	2	3	33700	*	*	*	*	*	*			
16	1	2	1	33130	*	*	*	*	*	*			
17	1	2	2	33610	*	*	*	*	*	*			
18	1	2	3	33285	*	*	*	*	*	*			
19	2	1	1	33180	32947.8	454.019	90.3540	-287.806	56.2745	-0.0758283			
20	2	1	2	34430	34331.9	155.874	90.7144	58.111	-59.4371	0.0142166			

Nous avons le graphique suivant qui met particulièrement en évidence une interaction entre les facteurs A et B:



et le graphique des effets principaux:



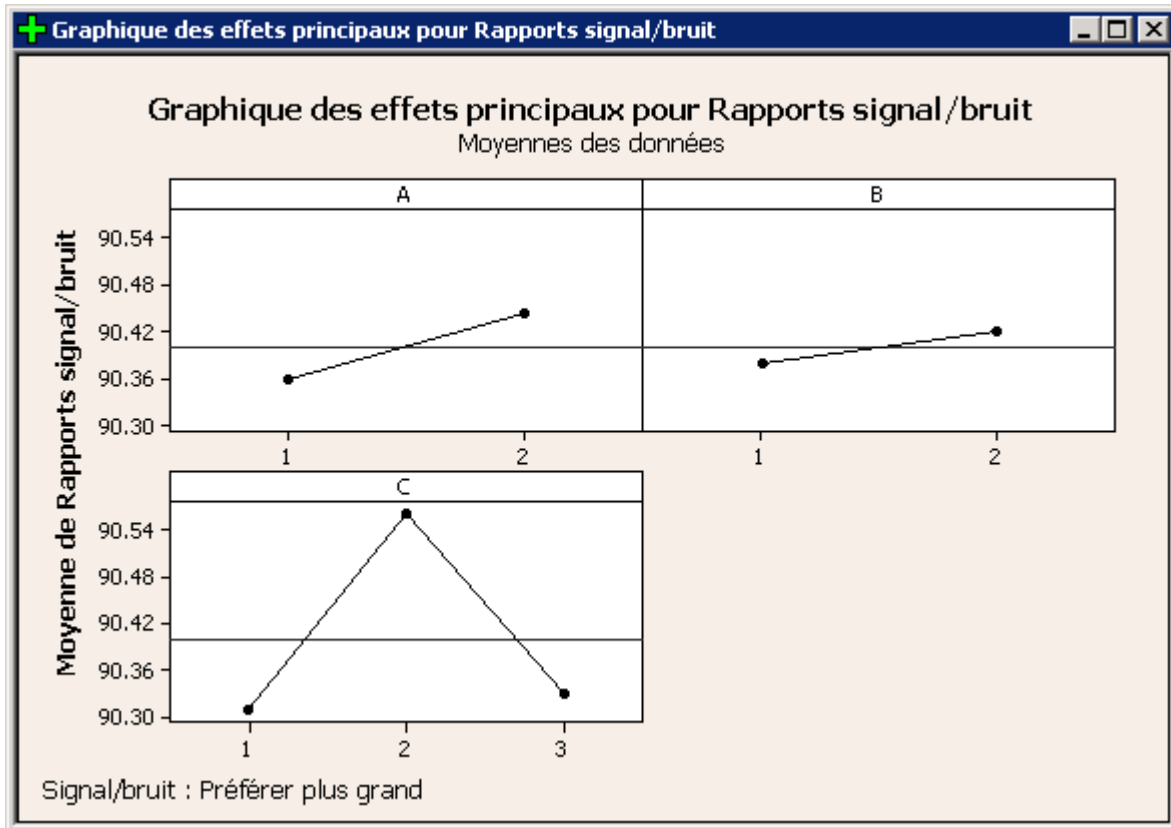
Nous y voyons aussi que la combinaison gagnante est:

$$A_2 B_2 C_2$$

Les questions restantes sont alors:

- Quel facteur contrôlé à le plus grand impact sur le ratio S/N?
- Quel facteur contrôlé à le plus petit impact sur le ratio S/N?

PS: Notez que Minitab met automatiquement des rangs dans chacune des tables ci-dessus...



Nous voyons dans le graphique ci-dessus que  $A_2B_2C_2$  est la combinaison perdante (provoque le plus de bruit)... Il paraît assez évidemment visuellement que c'est C qui influence le plus le bruit et B le moins... ce qui est conforme aux rangs donné par le tableau dans la fenêtre de sessions si nous avons activé l'affichage des données de bruit:

Tableau des réponses pour les rapports signal/bruit  
Préférer plus grand

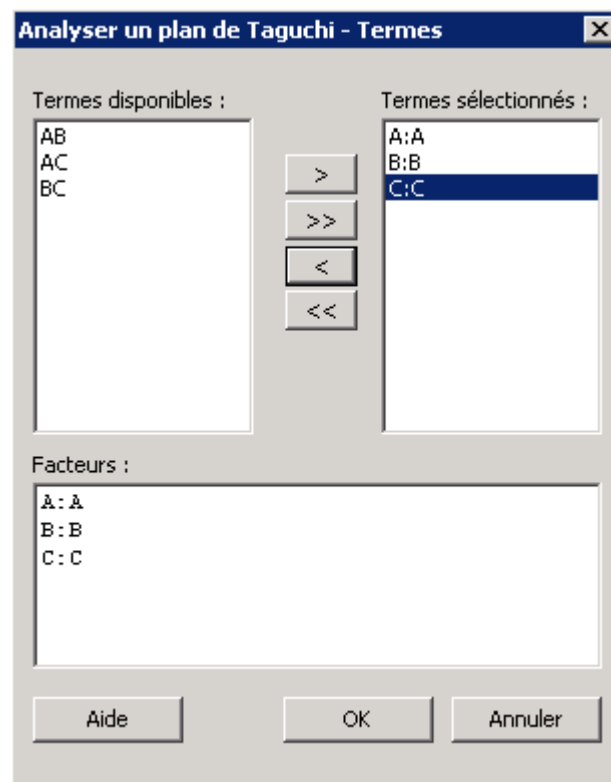
Niveau	A	B	C
1	90.36	90.38	90.31
2	90.44	90.42	90.56
3			90.33
Delta	0.08	0.04	0.25
Rang	2	3	1

Tableau des réponses pour les moyennes

Niveau	A	B	C
1	32956	33041	32770
2	33283	33198	33739
3			32850
Delta	327	157	969
Rang	2	3	1

Pour simuler le modèle linéaire avec ce scénario particulier, il suffit de procéder exactement comme dans l'exercice précédent! (il nous a semblé inutile de refaire la démarche puisque ce qui nous intéresse ici est de comparer à l'analyse faite par plan factoriel complet).

Pour clore, si nous faisons maintenant l'analyse sans interactions:



Nous obtenons dans la fenêtre de session:

### Analyse du modèle linéaire : Rapports signal/bruit en fonction de A; B; C

Coefficients du modèle estimés pour Rapports signal/bruit

Terme	Coeff	Coef ErT	T	P
Constante	90.3991	0.03448	2621.932	0.000
A 1	-0.0425	0.03448	-1.232	0.258
B 1	-0.0212	0.03448	-0.615	0.558
C 1	-0.0911	0.04876	-1.868	0.104
C 2	0.1618	0.04876	3.319	0.013

S = 0.1194 R carré = 64.9 % R carré (ajust) = 44.9 %

Analyse de la variance pour Rapports signal/bruit

Source	DL	SomCar séq	SomCar ajust	CM ajust	F	P
A	1	0.021641	0.021641	0.021641	1.52	0.258
B	1	0.005387	0.005387	0.005387	0.38	0.558
C	2	0.157956	0.157956	0.078978	5.54	0.036
Erreur résiduelle	7	0.099854	0.099854	0.014265		
Total	11	0.284838				



**Analyse du modèle linéaire : Moyennes en fonction de A; B; C**

Coefficients du modèle estimés pour Moyennes

Terme	Coeff	Coef ErT	T	P
Constante	33119.3	131.2	252.434	0.000
A 1	-163.4	131.2	-1.246	0.253
B 1	-78.7	131.2	-0.600	0.567
C 1	-349.7	185.5	-1.885	0.101
C 2	619.4	185.5	3.338	0.012

S = 454.5 R carré = 65.2 % R carré (ajust) = 45.3 %

Analyse de la variance pour Moyennes

Source	DL	SomCar séq	SomCar ajust	CM ajust	F	P
A	1	320569	320569	320569	1.55	0.253
B	1	74366	74366	74366	0.36	0.567
C	2	2314656	2314656	1157328	5.60	0.035
Erreur résiduelle	7	1445921	1445921	206560		
Total	11	4155512				

et dans la table:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	A	B	C	Reponses	AJUS MOY2	AJUS ECTY2	AJUS SB2	RES MOY2	RES ECTYP2	RES SB2	COEFF MOY2	COEFF ETYP2	COEFF SB2
1	1	1	1	32700	32527.4	291.687	90.2444	275.917	-153.726	0.073836	33119.3	351.196	90.3991
2	1	1	2	33430	33496.5	228.000	90.4973	-263.167	54.194	-0.066446	-163.4	-40.768	-0.0425
3	1	1	3	31710	32607.4	263.441	90.2647	-597.417	3.204	-0.159630	-78.7	-49.386	-0.0212
4	1	1	1	32750	*	*	*	*	*	*	-349.7	30.644	-0.0911
5	1	1	2	33360	*	*	*	*	*	*	619.4	-33.043	0.1618
6	1	1	3	32100	*	*	*	*	*	*			
7	1	1	1	32960	*	*	*	*	*	*			
8	1	1	2	32910	*	*	*	*	*	*			
9	1	1	3	32220	*	*	*	*	*	*			
10	1	2	1	32680	32684.9	390.458	90.2867	8.472	39.697	0.000936			
11	1	2	2	34070	33653.9	326.771	90.5397	-60.611	158.443	-0.016421			
12	1	2	3	33220	32764.9	362.213	90.3071	636.806	-101.812	0.167725			
13	1	2	1	32270	*	*	*	*	*	*			
14	1	2	2	33100	*	*	*	*	*	*			
15	1	2	3	33700	*	*	*	*	*	*			
16	1	2	1	33130	*	*	*	*	*	*			
17	1	2	2	33610	*	*	*	*	*	*			
18	1	2	3	33285	*	*	*	*	*	*			
19	2	1	1	33180	32854.3	373.223	90.3293	-194.306	137.071	-0.051099			
20	2	1	2	34430	33823.4	309.536	90.5822	566.611	-213.099	0.146347			

A comparer avec:

$$Y = 33'119.3 + x_1 \begin{bmatrix} -163.4 \\ 163.4 \end{bmatrix} + x_2 \begin{bmatrix} -78.7 \\ 78.7 \end{bmatrix} + x_3 \begin{bmatrix} -349.7 \\ 619.4 \\ -269.7 \end{bmatrix}$$

$\underbrace{\hspace{1.5cm}}_{\Sigma=0}$ 
 $\underbrace{\hspace{1.5cm}}_{\Sigma=0}$ 
 $\underbrace{\hspace{1.5cm}}_{\Sigma=0}$

Nous retrouvons bien tous les mêmes coefficients.

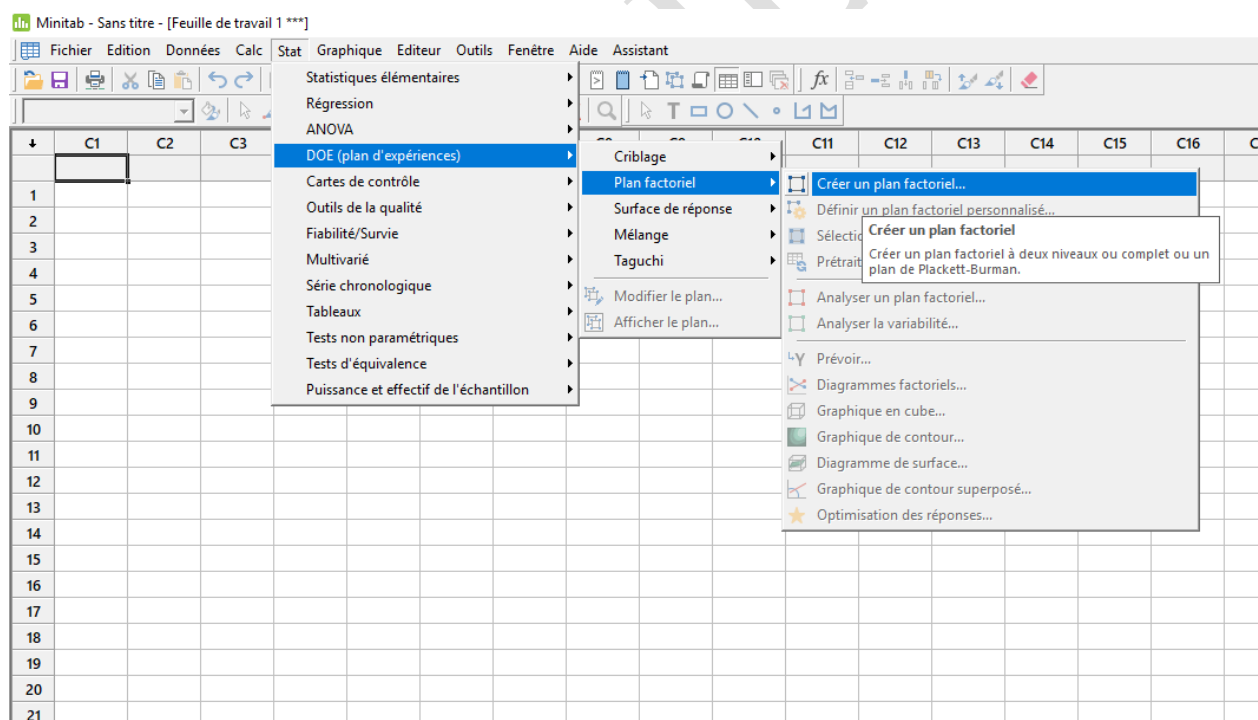
## 17.18. Exercice 189.: Plan en parcelles divisées (split-plot) avec facteurs difficiles à changer

Minitab® Statistical Software 18.1

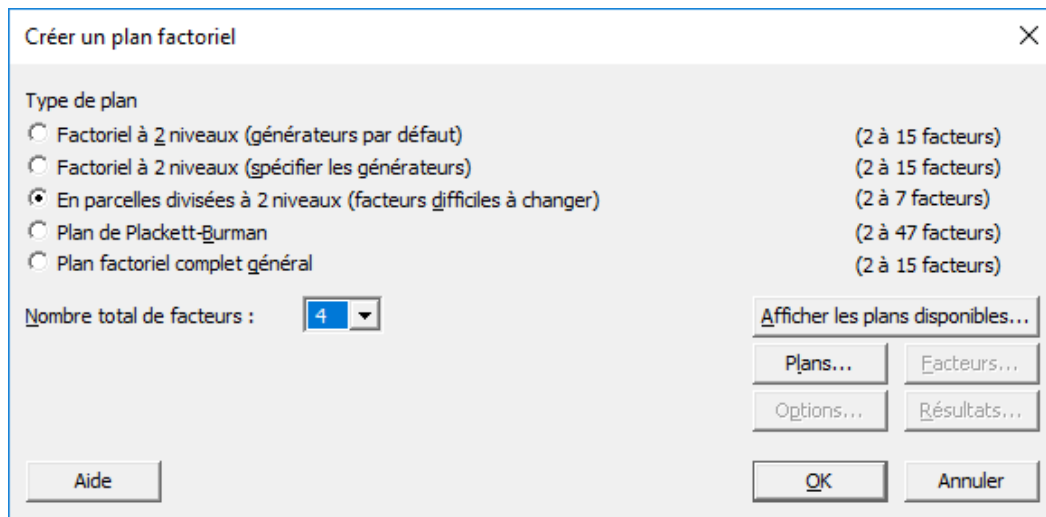
Il est temps de passer aux non moins fameux plans d'expérience factoriels en parcelles divisées. Alors je tiens à rappeler (encore une fois!) que:

1. C'est un sujet sensible et qu'il n'est pas facile de choisir les bons plans d'expérience pour des expériences complexes (même les meilleurs statisticiens peuvent se tromper)
2. Les plans d'expériences split-plot continuent à faire l'objet d'assez nombreuses publications scientifiques et donc il y a plusieurs "formules" pour les calculs. Vous ne risquez pas toujours de tomber sur les mêmes résultats d'un logiciel à un autre...
3. Rappelez-vous dans ce qui va suivre que DAC signifie "Difficile à changer"!

Nous allons partir de zéro. Pour cela nous allons d'abord dans le menu **Stat/DOE (plan d'expériences)/Plan factoriel/Créer un plan factoriel...**:



Nous prenons ensuite l'option **En parcelles divisées à 2 niveaux (facteurs difficiles à changer)** avec **4** facteurs:



Si nous cliquons sur **Afficher les plans disponibles...** (pour rappel DAC signifie "Difficile à Changer", FAC signifie "Facile à changer" et SB signifie "sous-blocs". etc... **Tout est indiqué dans la légende en bas!!!**) nous avons dans le cas à 1 facteur difficile à changer:

Créer un plan factoriel : Afficher les plans disponibles

Plans en parcelles divisées disponibles

FAC	1 facteur DAC								
	2 SB			4 SB			8 SB		16 SB
	1/4	1/2	Complet	1/8	1/4	1/2	1/8	1/4	1/8
1			2 SP Complet						
2			4 SP Complet						
3		4 SP IV	8 SP Complet			4 SP Complet SB + 3FI			
4	4 SP III	8 SP V	16 SP Complet		4 SP IV SB + 2FI	8 SP Complet SB + 4FI		4 SP Complet SB + 2FI	
5	8 SP IV	16 SP VI	32 SP Complet		8 SP V SB + 3FI	16 SP Complet SB + 5FI		8 SP Complet SB + 3FI	
6	16 SP IV	32 SP VII	64 SP Complet	8 SP III SB + 3FI	16 SP VI SB + 3FI	32 SP Complet SB + 6FI	8 SP V SB + 3FI	16 SP Complet SB + 3FI	8 SP Complet SB + 3FI

1 facteur DAC / 2 facteurs DAC / 3 facteurs DAC

Sélectionnez l'onglet concernant le nombre de facteurs difficiles à changer (DAC) afin d'afficher les plans disponibles.

Les plans sont organisés en lignes selon le nombre de facteurs faciles à changer (easy to change, FAC).

Les plans sont organisés en colonnes selon le nombre de sous-blocs (whole plot, SB) et la fraction du plan complet de facteurs faciles à changer (easy to change) dans chaque sous-bloc (whole plot).

Minitab affiche trois caractéristiques pour chaque plan :

- Le nombre d'essais de sous-parcelles (subplot) dans chaque sous-bloc (whole plot).
- La résolution globale du plan.
- Le cas échéant, l'interaction la plus faible que le plan confond avec les sous-blocs (whole plot). Par exemple, 3FI indique une interaction à 3 facteurs.

Aide OK

Dans le cas de 2 facteurs difficiles à changer:

Créer un plan factoriel : Afficher les plans disponibles

Plans en parcelles divisées disponibles

2 facteurs DAC						
	4 SB		8 SB		16 SB	
FAC	1/4	1/2	Complet	1/4	1/2	1/4
1			2 SP Complet			
2			4 SP Complet			
3		4 SP V	8 SP Complet		4 SP Complet SB + 3FI	
4	4 SP IV	8 SP VI	16 SP Complet	4 SP IV SB + 2FI	8 SP Complet SB + 4FI	4 SP Complet SB + 2FI
5	8 SP IV	16 SP VII	32 SP Complet	8 SP V SB + 3FI	16 SP Complet SB + 5FI	8 SP Complet SB + 3FI

1 facteur DAC / 2 facteurs DAC / 3 facteurs DAC

Sélectionnez l'onglet concernant le nombre de facteurs difficiles à changer (DAC) afin d'afficher les plans disponibles.

Les plans sont organisés en lignes selon le nombre de facteurs faciles à changer (easy to change, FAC).

Les plans sont organisés en colonnes selon le nombre de sous-blocs (whole plot, SB) et la fraction du plan complet de facteurs faciles à changer (easy to change) dans chaque sous-bloc (whole plot).

Minitab affiche trois caractéristiques pour chaque plan :

- Le nombre d'essais de sous-parcelles (subplot) dans chaque sous-bloc (whole plot).
- La résolution globale du plan.
- Le cas échéant, l'interaction la plus faible que le plan confond avec les sous-blocs (whole plot). Par exemple, 3FI indique une interaction à 3 facteurs.

Aide OK

et dans le cas de 3 facteurs difficiles à changer:

Créer un plan factoriel : Afficher les plans disponibles

Plans en parcelles divisées disponibles

FAC	3 facteurs DAC		
	8 SB		16 SB
	1/2	Complet	1/2
1		2 SP Complet	
2		4 SP Complet	
3	4 SP VI	8 SP Complet	4 SP Complet SB + 3FI
4	8 SP VII	16 SP Complet	8 SP Complet SB + 4FI

1 facteur DAC / 2 facteurs DAC / **3 facteurs DAC**

Sélectionnez l'onglet concernant le nombre de facteurs difficiles à changer (DAC) afin d'afficher les plans disponibles.

Les plans sont organisés en lignes selon le nombre de facteurs faciles à changer (easy to change, FAC).

Les plans sont organisés en colonnes selon le nombre de sous-blocs (whole plot, SB) et la fraction du plan complet de facteurs faciles à changer (easy to change) dans chaque sous-bloc (whole plot).

Minitab affiche trois caractéristiques pour chaque plan :

- Le nombre d'essais de sous-parcelles (subplot) dans chaque sous-bloc (whole plot).
- La résolution globale du plan.
- Le cas échéant, l'interaction la plus faible que le plan confond avec les sous-blocs (whole plot). Par exemple, 3FI indique une interaction à 3 facteurs.

Aide OK

En cliquant sur le bouton **Plans...** nous avons une information complémentaire utile et important qui est le nombre d'essais. Ainsi pour le cas de 1 facteur difficile à changer et 3 autres facteurs faciles à changer, nous avons alors:

Créer un plan factoriel : Plans

Nombre de facteurs difficiles à changer : 1

Plans	Sous-blocs	Sous-parcelles	Essais	Résolution
Fraction 1/2	2	4	8	IV
Fact. compl.	2	8	16	Complet
Fact. compl.	4	4	16	Complet (SB + 3FI)

Nombre de répliques de sous-bloc : 2

Créer des blocs pour les répliques

Nombre de répliques de sous-parcelles : 1

Aide OK Annuler

Et pour le cas de 2 facteurs difficile à changer et 2 autres facteurs faciles à changer, nous avons alors:

Créer un plan factoriel : Plans

Nombre de facteurs difficiles à changer : 2

Plans	Sous-blocs	Sous-parcelles	Essais	Résolution
Fact. compl.	4	4	16	Complet

Nombre de répliques de sous-bloc : 2

Créer des blocs pour les répliques

Nombre de répliques de sous-parcelles : 1

Aide OK Annuler

et au final dans le cas de 3 facteurs difficiles à changer et 1 autre facteur facile à changer, nous avons alors:

Créer un plan factoriel : Plans

Nombre de facteurs difficiles à changer : 3

Plans	Sous-blocs	Sous-parcelles	Essais	Résolution
Fact. compl.	8	2	16	Complet

Nombre de répliques de sous-bloc : 2

Créer des blocs pour les répliques

Nombre de répliques de sous-parcelles : 1

Aide OK Annuler

Dans notre cas, nous allons donc prendre un plan avec 1 facteur difficile à changer, de type complet avec 2 sous-blocs et 8 sous parcelles avec 2 répliques de sous-blocs (ce qui fera au total  $2 \times 16 = 32$  essais):

Créer un plan factoriel : Plans

Nombre de facteurs difficiles à changer : 1

Plans	Sous-blocs	Sous-parcelles	Essais	Résolution
Fraction 1/2	2	4	8	IV
Fact. compl.	2	8	16	Complet
Fact. compl.	4	4	16	Complet (SB + 3FI)

Nombre de répliques de sous-bloc : 2

Créer des blocs pour les répliques

Nombre de répliques de sous-parcelles : 1

Aide OK Annuler

Nous validons par **OK** et cliquons ensuite sur le bouton **Facteurs**:

Créer un plan factoriel : Facteurs

Facteur	Nom	Type	Min	Max
A[ <b>DAC</b> ]	A[ <b>DAC</b> ]	Numérique	-1	1
B	B	Numérique	-1	1
C	C	Numérique	-1	1
D	D	Numérique	-1	1

DAC indique un facteur difficile à changer.

Aide OK Annuler

Et nous mettons (pour rappel "DAC" signifie "Difficile à Changer"):

Créer un plan factoriel : Facteurs

Facteur	Nom	Type	Min	Max
A[ <b>DAC</b> ]	Température	Numérique	200	275
B	Poids revête	Numérique	20	30
C	Quantité de l	Numérique	10	20
D	Température	Numérique	150	200

DAC indique un facteur difficile à changer.

Aide OK Annuler

soit explicitement:

- **A[**DAC**]**: Température Extrudeuse [**DAC**]
- **B**: Poids revêtement
- **C**: Quantité lubrifiant
- **D**: Température de cuisson

Ensuite, nous validons par **OK** et cliquons sur le bouton Options pour y désactiver l'option **Randomiser les essais**:

Créer un plan factoriel : Options

Replier le plan d'expériences

Ne pas replier
  Replier sur tous les facteurs
  Replier seulement sur le facteur :

Fraction

Utiliser la fraction principale
  Utiliser la fraction numéro :

**Randomiser les essais**

Base du générateur de nombres aléatoires :

Stocker le plan dans la feuille de travail

Aide OK Annuler



Nous validons deux fois par **OK** pour obtenir:

**Plan factoriel complet en parcelles divisées**

**Récapitulatif du plan**

Facteurs :	4	Sous-blocs :	4
Difficile à changer :	1	Nombre d'essais par sous-bloc :	8
Essais :	32	Répliques de sous-bloc :	2
Blocs :	1	Répliques de sous-parcelles :	1

Facteurs difficiles à changer : A  
 Générateurs de sous-blocs : A  
 Tous les termes sont exempts d'alias.

↓	C1	C2	C3	C4	C5	C6	C7	C8	C9
	OrdreStd	OrdEssai	TypePt	Blocs	DC	Température Extrudeuse [DAC]	Poids revêtement	Quantité de lubrifiant	Température de cuisson
1	1	1	1	1	1	200	20	10	150
2	2	2	1	1	1	200	30	10	150
3	3	3	1	1	1	200	20	20	150
4	4	4	1	1	1	200	30	20	150
5	5	5	1	1	1	200	20	10	200
6	6	6	1	1	1	200	30	10	200
7	7	7	1	1	1	200	20	20	200
8	8	8	1	1	1	200	30	20	200
9	9	9	1	1	2	275	20	10	150
10	10	10	1	1	2	275	30	10	150
11	11	11	1	1	2	275	20	20	150
12	12	12	1	1	2	275	30	20	150
13	13	13	1	1	2	275	20	10	200
14	14	14	1	1	2	275	30	10	200
15	15	15	1	1	2	275	20	20	200
16	16	16	1	1	2	275	30	20	200

Feuille de travail en cours : Feuille de travail 1

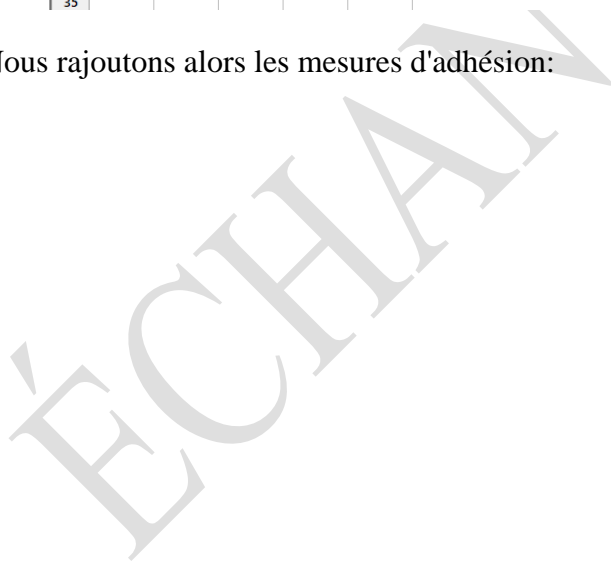
Soit pour les 32 essais (vue globale):

Minitab - Sans titre - [Feuille de travail 1 \*\*\*]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assistant

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
	OrdreStd	OrdEssai	TypePt	Blocs	DC	Température Extrudeuse [DAC]	Poids revêtement	Quantité de lubrifiant	Température de cuisson	
1	1	1	1	1	1	200	20	10	150	
2	2	2	1	1	1	200	30	10	150	
3	3	3	1	1	1	200	20	20	150	
4	4	4	1	1	1	200	30	20	150	
5	5	5	1	1	1	200	20	10	200	
6	6	6	1	1	1	200	30	10	200	
7	7	7	1	1	1	200	20	20	200	
8	8	8	1	1	1	200	30	20	200	
9	9	9	1	1	2	275	20	10	150	
10	10	10	1	1	2	275	30	10	150	
11	11	11	1	1	2	275	20	20	150	
12	12	12	1	1	2	275	30	20	150	
13	13	13	1	1	2	275	20	10	200	
14	14	14	1	1	2	275	30	10	200	
15	15	15	1	1	2	275	20	20	200	
16	16	16	1	1	2	275	30	20	200	
17	17	17	1	1	3	200	20	10	150	
18	18	18	1	1	3	200	30	10	150	
19	19	19	1	1	3	200	20	20	150	
20	20	20	1	1	3	200	30	20	150	
21	21	21	1	1	3	200	20	10	200	
22	22	22	1	1	3	200	30	10	200	
23	23	23	1	1	3	200	20	20	200	
24	24	24	1	1	3	200	30	20	200	
25	25	25	1	1	4	275	20	10	150	
26	26	26	1	1	4	275	30	10	150	
27	27	27	1	1	4	275	20	20	150	
28	28	28	1	1	4	275	30	20	150	
29	29	29	1	1	4	275	20	10	200	
30	30	30	1	1	4	275	30	10	200	
31	31	31	1	1	4	275	20	20	200	
32	32	32	1	1	4	275	30	20	200	
33										
34										
35										

Nous rajoutons alors les mesures d'adhésion:



Minitab - DOE\_SplitPlot.MPJ - [Feuille de travail 1 \*\*\*]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assistant

↓	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
	OrdreStd	OrdEssai	TypePt	Blocs	DC	Température Extrudeuse [DAC]	Poids revêtement	Quantité de lubrifiant	Température de cuisson	Adhésion
1	1	1	1	1	1	200	20	10	150	29.3
2	2	2	1	1	1	200	30	10	150	39.9
3	3	3	1	1	1	200	20	20	150	25.4
4	4	4	1	1	1	200	30	20	150	30.9
5	5	5	1	1	1	200	20	10	200	30.9
6	6	6	1	1	1	200	30	10	200	37.9
7	7	7	1	1	1	200	20	20	200	27.3
8	8	8	1	1	1	200	30	20	200	33.0
9	9	9	1	1	2	275	20	10	150	34.0
10	10	10	1	1	2	275	30	10	150	44.9
11	11	11	1	1	2	275	20	20	150	13.7
12	12	12	1	1	2	275	30	20	150	22.9
13	13	13	1	1	2	275	20	10	200	32.3
14	14	14	1	1	2	275	30	10	200	41.6
15	15	15	1	1	2	275	20	20	200	18.4
16	16	16	1	1	2	275	30	20	200	23.4
17	17	17	1	1	3	200	20	10	150	28.4
18	18	18	1	1	3	200	30	10	150	37.0
19	19	19	1	1	3	200	20	20	150	22.8
20	20	20	1	1	3	200	30	20	150	26.7
21	21	21	1	1	3	200	20	10	200	29.4
22	22	22	1	1	3	200	30	10	200	37.7
23	23	23	1	1	3	200	20	20	200	22.2
24	24	24	1	1	3	200	30	20	200	30.9
25	25	25	1	1	4	275	20	10	150	28.4
26	26	26	1	1	4	275	30	10	150	41.4
27	27	27	1	1	4	275	20	20	150	14.4
28	28	28	1	1	4	275	30	20	150	24.4
29	29	29	1	1	4	275	20	10	200	31.1
30	30	30	1	1	4	275	30	10	200	42.8
31	31	31	1	1	4	275	20	20	200	10.8
32	32	32	1	1	4	275	30	20	200	21.7
33										
34										
35										
36										
37										

Ensuite, nous faisons une analyse en cliquant sur **Stat/DOE (plans d'expériences)/Plan factoriel/Analyser un plan factoriel...**:

Minitab - DOE\_SplitPlot.MPJ - [Feuille de travail 1 \*\*\*]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assistant

Statistiques élémentaires  
Régression  
ANOVA  
DOE (plan d'expériences)  
Cartes de contrôle  
Outils de la qualité  
Fiabilité/Survie  
Multivarié  
Série chronologique  
Tableaux  
Tests non paramétriques  
Tests d'équivalence  
Puissance et effectif de l'échantillon

Criblage  
Plan factoriel  
Surface de réponse  
Mélange  
Taguchi  
Modifier le plan...  
Afficher le plan...

Quantité de lubrifiant Température de cuisson Adhésion

Créer un plan factoriel...  
Définir un plan factoriel personnalisé...  
Sélectionner un plan optimal...  
Prétraiter les réponses pour l'analyse de la variabilité...  
Analyser un plan factoriel...  
Analyser la variabilité...  
Analyser un plan factoriel  
Ajuster un modèle à un plan factoriel.  
Prévoir...  
Diagrammes factoriels...  
Graphique en cube...  
Graphique de contour...  
Diagramme de surface...  
Graphique de contour superposé...  
Optimisation des réponses...

	C1	C2	C3				C8	C9	C10	C11	
	OrdreStd	OrdEssai	TypePt								
1	1	1	1								
2	2	2	1								
3	3	3	1								
4	4	4	1								
5	5	5	1								
6	6	6	1								
7	7	7	1								
8	8	8	1								
9	9	9	1								
10	10	10	1	1	2		200	20			
11	11	11	1	1	2		275	30			
12	12	12	1	1	2		275	20			
13	13	13	1	1	2		275	30			
14	14	14	1	1	2		275	20			
15	15	15	1	1	2		275	30			
16	16	16	1	1	2		275	20			
17	17	17	1	1	3		200	20	10	28.4	
18	18	18	1	1	3		200	30	10	37.0	
19	19	19	1	1	3		200	20	20	22.8	
20	20	20	1	1	3		200	30	20	26.7	
21	21	21	1	1	3		200	20	10	29.4	
22	22	22	1	1	3		200	30	10	37.7	
23	23	23	1	1	3		200	20	200	22.2	
24	24	24	1	1	3		200	30	200	30.9	
25	25	25	1	1	4		275	20	10	28.4	
26	26	26	1	1	4		275	30	10	41.4	
27	27	27	1	1	4		275	20	20	14.4	
28	28	28	1	1	4		275	30	150	24.4	
29	29	29	1	1	4		275	20	10	31.1	
30	30	30	1	1	4		275	30	10	200	42.8
31	31	31	1	1	4		275	20	20	200	10.8
32	32	32	1	1	4		275	30	20	200	21.7
33											
34											
35											
36											

Nous mettons:

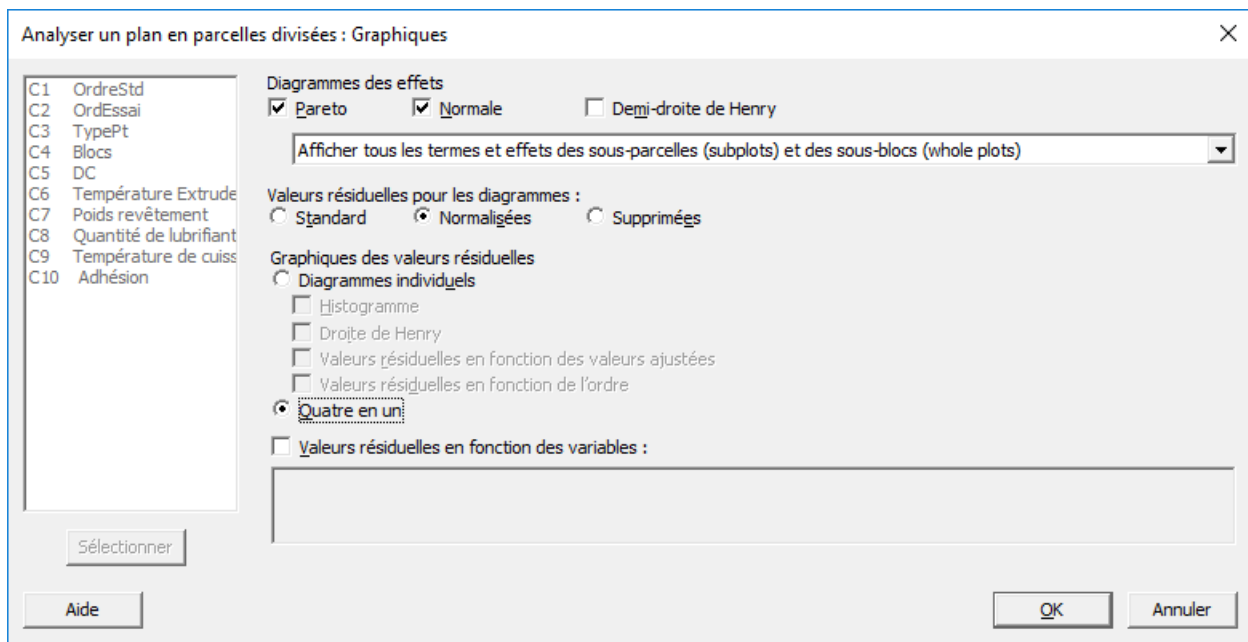
Analyser un plan en parcelles divisées

C10 Adhésion

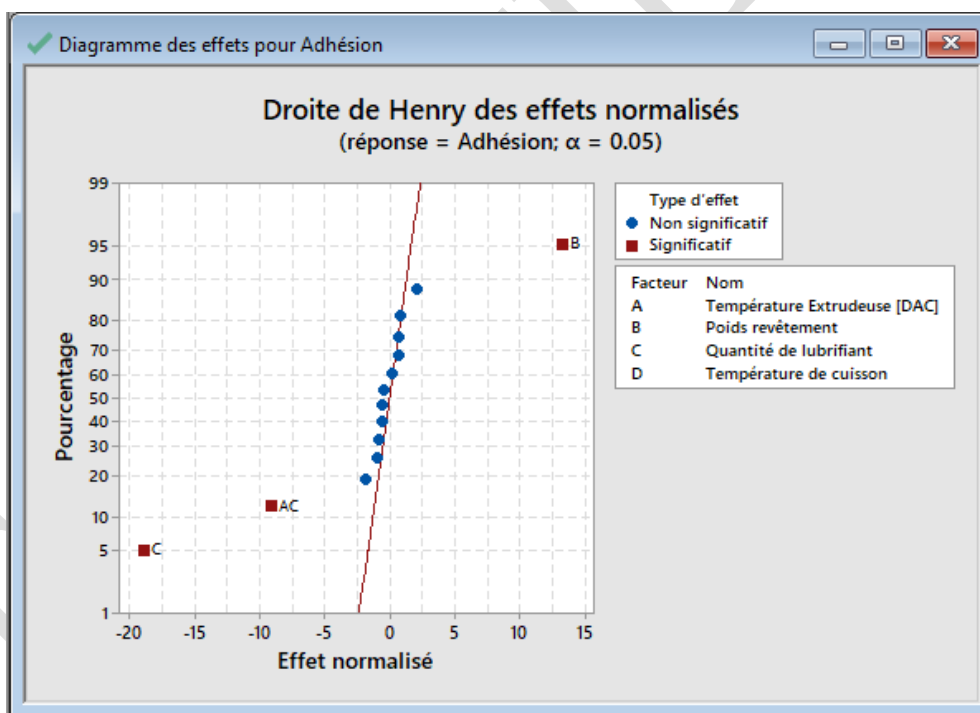
Réponses :  
Adhésion

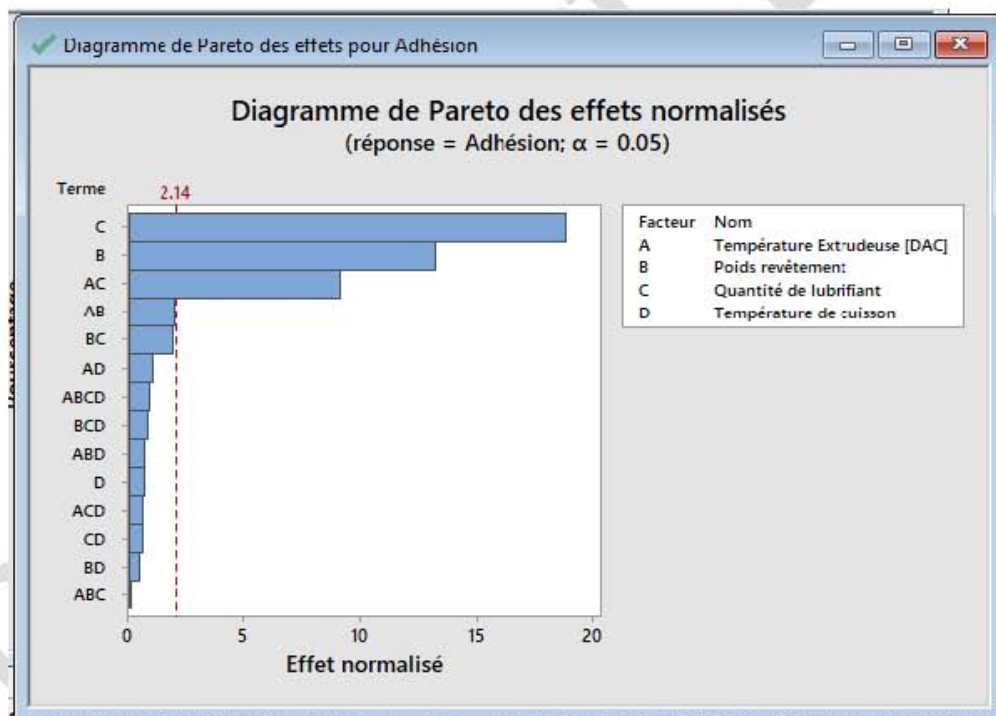
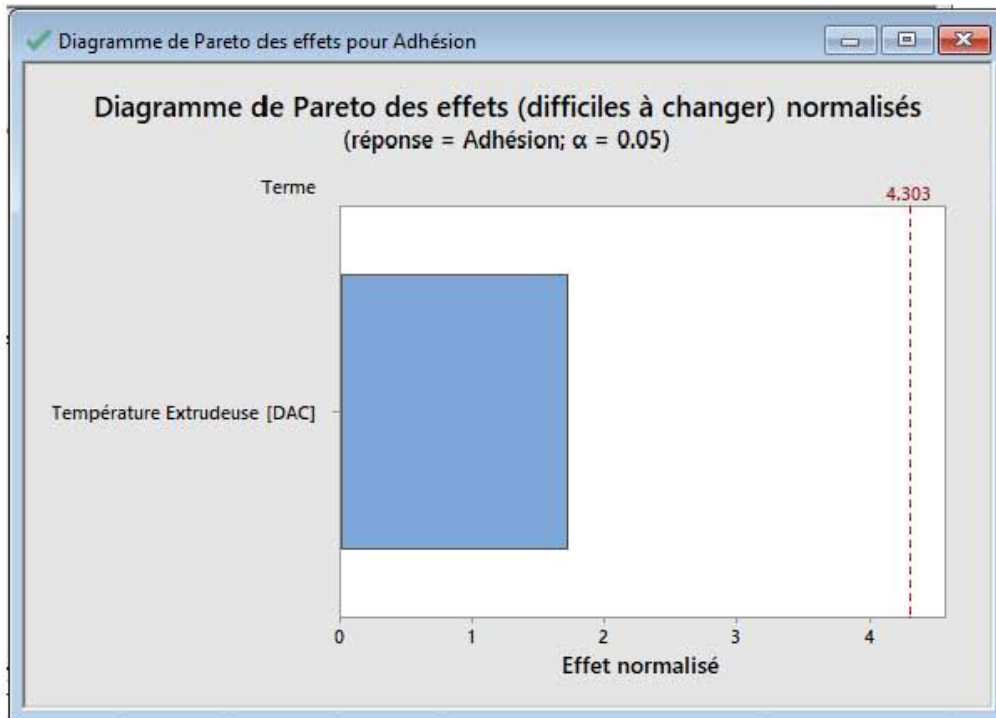
Termes... Covariables... Options...  
Graphiques... Résultats... Stockage...  
Sélectionner Aide OK Annuler

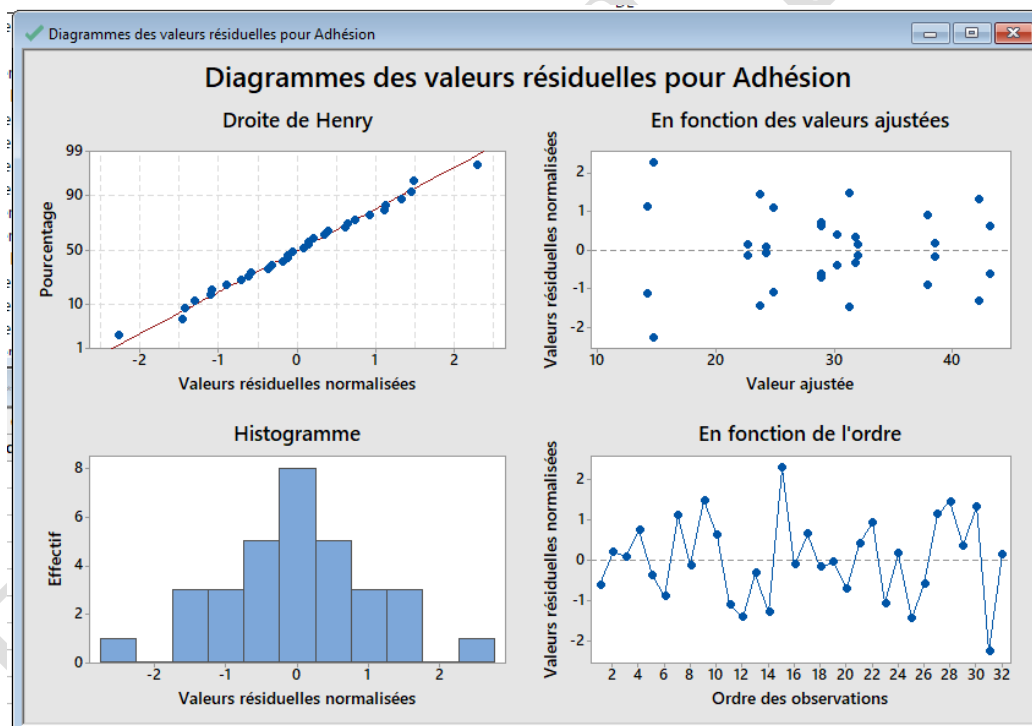
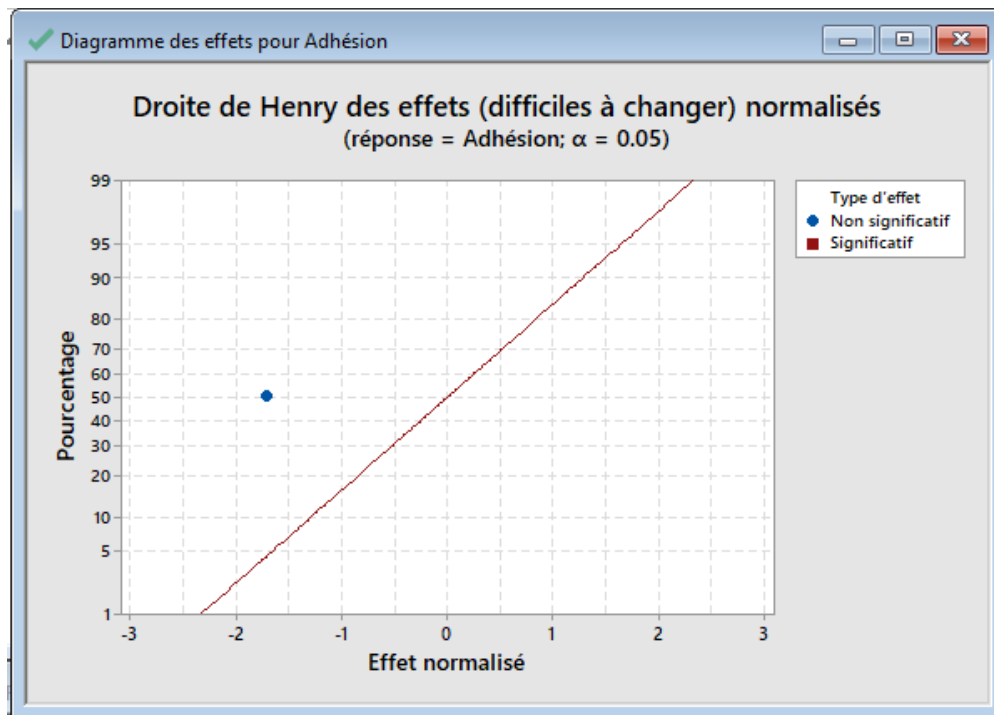
Nous cliquons sur le bouton **Graphiques...** pour prendre:



et nous validons deux fois par **OK** pour obtenir dans l'ordre d'apparition:







Et pour la partie tableaux (assez indigeste vu la longueur des noms des variables):

## Régression factorielle en parcelles divisées : Adhésion en ... de cuisson

### Analyse de variance

Source	DL
Température Extrudeuse [DAC]	1
Erreur SB	2
Poids revêtement	1
Quantité de lubrifiant	1
Température de cuisson	1
Température Extrudeuse [DAC]*Poids revêtement	1
Température Extrudeuse [DAC]*Quantité de lubrifiant	1
Température Extrudeuse [DAC]*Température de cuisson	1
Poids revêtement*Quantité de lubrifiant	1
Poids revêtement*Température de cuisson	1
Quantité de lubrifiant*Température de cuisson	1
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant	1
Température Extrudeuse [DAC]*Poids revêtement*Température de cuisson	1
Température Extrudeuse [DAC]*Quantité de lubrifiant*Température de cuisson	1
Poids revêtement*Quantité de lubrifiant*Température de cuisson	1
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant*Température de cuisson	1
Erreur SP	14
Total	31

Source	SomCar ajust
Température Extrudeuse [DAC]	59.13
Erreur SB	40.17
Poids revêtement	597.72
Quantité de lubrifiant	1226.36
Température de cuisson	1.49
Température Extrudeuse [DAC]*Poids revêtement	14.72
Température Extrudeuse [DAC]*Quantité de lubrifiant	285.01
Température Extrudeuse [DAC]*Température de cuisson	3.71
Poids revêtement*Quantité de lubrifiant	13.13
Poids revêtement*Température de cuisson	0.81
Quantité de lubrifiant*Température de cuisson	1.16
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant	0.03
Température Extrudeuse [DAC]*Poids revêtement*Température de cuisson	1.67
Température Extrudeuse [DAC]*Quantité de lubrifiant*Température de cuisson	1.32
Poids revêtement*Quantité de lubrifiant*Température de cuisson	2.26
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant*Température de cuisson	2.70
Erreur SP	47.94
Total	



Source	CM ajust
Température Extrudeuse [DAC]	59.13
Erreur SB	20.08
Poids revêtement	597.72
Quantité de lubrifiant	1226.36
Température de cuisson	1.49
Température Extrudeuse [DAC]*Poids revêtement	14.72
Température Extrudeuse [DAC]*Quantité de lubrifiant	285.01
Température Extrudeuse [DAC]*Température de cuisson	3.71
Poids revêtement*Quantité de lubrifiant	13.13
Poids revêtement*Température de cuisson	0.81
Quantité de lubrifiant*Température de cuisson	1.16
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant	0.03
Température Extrudeuse [DAC]*Poids revêtement*Température de cuisson	1.67
Température Extrudeuse [DAC]*Quantité de lubrifiant*Température de cuisson	1.32
Poids revêtement*Quantité de lubrifiant*Température de cuisson	2.26
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant*Température de cuisson	2.70
Erreur SP	3.42
Total	

Source	Valeur F
Température Extrudeuse [DAC]	2.94
Erreur SB	5.87
Poids revêtement	174.56
Quantité de lubrifiant	358.16
Température de cuisson	0.43
Température Extrudeuse [DAC]*Poids revêtement	4.30
Température Extrudeuse [DAC]*Quantité de lubrifiant	83.24
Température Extrudeuse [DAC]*Température de cuisson	1.08
Poids revêtement*Quantité de lubrifiant	3.84
Poids revêtement*Température de cuisson	0.24
Quantité de lubrifiant*Température de cuisson	0.34
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant	0.01
Température Extrudeuse [DAC]*Poids revêtement*Température de cuisson	0.49
Température Extrudeuse [DAC]*Quantité de lubrifiant*Température de cuisson	0.39
Poids revêtement*Quantité de lubrifiant*Température de cuisson	0.66
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant*Température de cuisson	0.79
Erreur SP	
Total	



Source	Valeur de p
Température Extrudeuse [DAC]	0.228
Erreur SB	0.014
Poids revêtement	0.000
Quantité de lubrifiant	0.000
Température de cuisson	0.520
Température Extrudeuse [DAC]*Poids revêtement	0.057
Température Extrudeuse [DAC]*Quantité de lubrifiant	0.000
Température Extrudeuse [DAC]*Température de cuisson	0.315
Poids revêtement*Quantité de lubrifiant	0.070
Poids revêtement*Température de cuisson	0.634
Quantité de lubrifiant*Température de cuisson	0.569
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant	0.933
Température Extrudeuse [DAC]*Poids revêtement*Température de cuisson	0.497
Température Extrudeuse [DAC]*Quantité de lubrifiant*Température de cuisson	0.545
Poids revêtement*Quantité de lubrifiant*Température de cuisson	0.430
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant*Température de cuisson	0.389
Erreur SP	
Total	

### Récapitulatif du modèle

S	R carré (SP)	S(SB)	R carré (SB)
1.85042	97.82%	1.44309	59.55%

## Coefficients codés

Terme	Effet
Constante	
Température Extrudeuse [DAC]	-2.719
Poids revêtement	8.644
Quantité de lubrifiant	-12.381
Température de cuisson	0.431
Température Extrudeuse [DAC]*Poids revêtement	1.356
Température Extrudeuse [DAC]*Quantité de lubrifiant	-5.969
Température Extrudeuse [DAC]*Température de cuisson	-0.681
Poids revêtement*Quantité de lubrifiant	-1.281
Poids revêtement*Température de cuisson	-0.319
Quantité de lubrifiant*Température de cuisson	0.381
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant	0.056
Température Extrudeuse [DAC]*Poids revêtement*Température de cuisson	-0.456
Température Extrudeuse [DAC]*Quantité de lubrifiant*Température de cuisson	-0.406
Poids revêtement*Quantité de lubrifiant*Température de cuisson	0.531
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant*Température de cuisson	-0.581

Terme	Coeff
Constante	29.247
Température Extrudeuse [DAC]	-1.359
Poids revêtement	4.322
Quantité de lubrifiant	-6.191
Température de cuisson	0.216
Température Extrudeuse [DAC]*Poids revêtement	0.678
Température Extrudeuse [DAC]*Quantité de lubrifiant	-2.984
Température Extrudeuse [DAC]*Température de cuisson	-0.341
Poids revêtement*Quantité de lubrifiant	-0.641
Poids revêtement*Température de cuisson	-0.159
Quantité de lubrifiant*Température de cuisson	0.191
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant	0.028
Température Extrudeuse [DAC]*Poids revêtement*Température de cuisson	-0.228
Température Extrudeuse [DAC]*Quantité de lubrifiant*Température de cuisson	-0.203
Poids revêtement*Quantité de lubrifiant*Température de cuisson	0.266
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant*Température de cuisson	-0.291

Terme	Coef ErT
Constante	0.792
Température Extrudeuse [DAC]	0.792
Poids revêtement	0.327
Quantité de lubrifiant	0.327
Température de cuisson	0.327
Température Extrudeuse [DAC]*Poids revêtement	0.327
Température Extrudeuse [DAC]*Quantité de lubrifiant	0.327
Température Extrudeuse [DAC]*Température de cuisson	0.327
Poids revêtement*Quantité de lubrifiant	0.327
Poids revêtement*Température de cuisson	0.327
Quantité de lubrifiant*Température de cuisson	0.327
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant	0.327
Température Extrudeuse [DAC]*Poids revêtement*Température de cuisson	0.327
Température Extrudeuse [DAC]*Quantité de lubrifiant*Température de cuisson	0.327
Poids revêtement*Quantité de lubrifiant*Température de cuisson	0.327
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant*Température de cuisson	0.327

Terme	Valeur de T
Constante	36.92
Température Extrudeuse [DAC]	-1.72
Poids revêtement	13.21
Quantité de lubrifiant	-18.93
Température de cuisson	0.66
Température Extrudeuse [DAC]*Poids revêtement	2.07
Température Extrudeuse [DAC]*Quantité de lubrifiant	-9.12
Température Extrudeuse [DAC]*Température de cuisson	-1.04
Poids revêtement*Quantité de lubrifiant	-1.96
Poids revêtement*Température de cuisson	-0.49
Quantité de lubrifiant*Température de cuisson	0.58
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant	0.09
Température Extrudeuse [DAC]*Poids revêtement*Température de cuisson	-0.70
Température Extrudeuse [DAC]*Quantité de lubrifiant*Température de cuisson	-0.62
Poids revêtement*Quantité de lubrifiant*Température de cuisson	0.81
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant*Température de cuisson	-0.89

Terme	Valeur de p
Constante	0.001
Température Extrudeuse [DAC]	0.228
Poids revêtement	0.000
Quantité de lubrifiant	0.000
Température de cuisson	0.520
Température Extrudeuse [DAC]*Poids revêtement	0.057
Température Extrudeuse [DAC]*Quantité de lubrifiant	0.000
Température Extrudeuse [DAC]*Température de cuisson	0.315
Poids revêtement*Quantité de lubrifiant	0.070
Poids revêtement*Température de cuisson	0.634
Quantité de lubrifiant*Température de cuisson	0.569
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant	0.933
Température Extrudeuse [DAC]*Poids revêtement*Température de cuisson	0.497
Température Extrudeuse [DAC]*Quantité de lubrifiant*Température de cuisson	0.545
Poids revêtement*Quantité de lubrifiant*Température de cuisson	0.430
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant*Température de cuisson	0.389

Terme	FIV
Constante	
Température Extrudeuse [DAC]	*
Poids revêtement	1.00
Quantité de lubrifiant	1.00
Température de cuisson	1.00
Température Extrudeuse [DAC]*Poids revêtement	1.00
Température Extrudeuse [DAC]*Quantité de lubrifiant	1.00
Température Extrudeuse [DAC]*Température de cuisson	1.00
Poids revêtement*Quantité de lubrifiant	1.00
Poids revêtement*Température de cuisson	1.00
Quantité de lubrifiant*Température de cuisson	1.00
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant	1.00
Température Extrudeuse [DAC]*Poids revêtement*Température de cuisson	1.00
Température Extrudeuse [DAC]*Quantité de lubrifiant*Température de cuisson	1.00
Poids revêtement*Quantité de lubrifiant*Température de cuisson	1.00
Température Extrudeuse [DAC]*Poids revêtement*Quantité de lubrifiant*Température de cuisson	1.00



## Equation de régression en unités non codées

Adhésion	=	$  \begin{aligned}  & -174 + 0.674 \text{ Température Extrudeuse [DAC]} + 7.54 \text{ Poids revêtement} \\  & + 16.0 \text{ Quantité de lubrifiant} + 0.92 \text{ Température de cuisson} \\  & - 0.0209 \text{ Température Extrudeuse [DAC]} * \text{Poids revêtement} \\  & - 0.0633 \text{ Température Extrudeuse [DAC]} * \text{Quantité de lubrifiant} \\  & - 0.00315 \text{ Température Extrudeuse [DAC]} * \text{Température de cuisson} \\  & - 0.623 \text{ Poids revêtement} * \text{Quantité de lubrifiant} \\  & - 0.0403 \text{ Poids revêtement} * \text{Température de cuisson} \\  & - 0.0724 \text{ Quantité de lubrifiant} * \text{Température de cuisson} \\  & + 0.00220 \text{ Température Extrudeuse [DAC]} * \text{Poids revêtement} * \text{Quantité de lubrifiant} \\  & + 0.000137 \text{ Température Extrudeuse [DAC]} * \text{Poids revêtement} * \text{Température de cuisson} \\  & + 0.000267 \text{ Température Extrudeuse [DAC]} * \text{Quantité de lubrifiant} * \text{Température de cuisson} \\  & + 0.00337 \text{ Poids revêtement} * \text{Quantité de lubrifiant} * \text{Température de cuisson} \\  & - 0.000012 \text{ Température Extrudeuse [DAC]} * \text{Poids revêtement} * \text{Quantité de lubrifiant} * \text{Température de cuisson}  \end{aligned}  $
----------	---	---

*Moyenne de l'équation sur les sous-blocs (whole plots).*

## Structure des alias

Facteur	Nom
A	Température Extrudeuse [DAC]
B	Poids revêtement
C	Quantité de lubrifiant
D	Température de cuisson

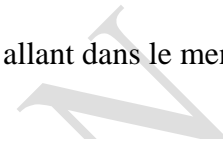
Alias
I
A
B
C
D
AB
AC
AD
BC
BD
CD
ABC
ABD
ACD
BCD
ABCD

## Ajustements et diagnostics pour les observations aberrantes

Observation	Adhésion	Valeur ajustée	Résiduelle	Val. résid. norm.	Valeurs résiduelles du sous-bloc
15	18.40	14.60	2.79	2.28	1.0125 R
31	10.80	14.60	-2.79	-2.28	-1.0125 R

R : Valeur résiduelle élevée

Pour simplifier la lecture, nous allons modifier les noms des facteurs en allant dans le menu **Stat/DOE (plan d'expériences)/Modifier le plan...**:



Minitab - DOE\_SplitPlot.MPJ - [Feuille de travail 1 \*\*\*\*]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assistant

Statistiques élémentaires  
Régression  
ANOVA  
DOE (plan d'expériences)  
Cartes de contrôle  
Outils de la qualité  
Fiabilité/Survie  
Multivarié  
Série chronologique  
Tableaux  
Tests non paramétriques  
Tests d'équivalence  
Puissance et effectif de l'échantillon

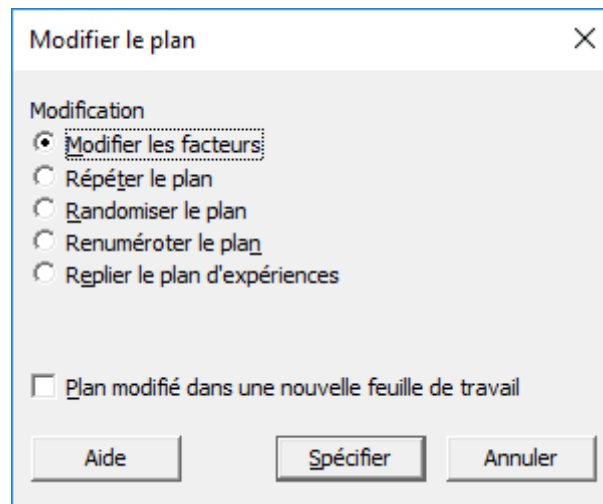
Criblage  
Plan factoriel  
Surface de réponse  
Mélange  
Taguchi  
Modifier le plan...  
Afficher le plan...

Quantité de lubrifiant  
Température de cuisson  
Adhésion

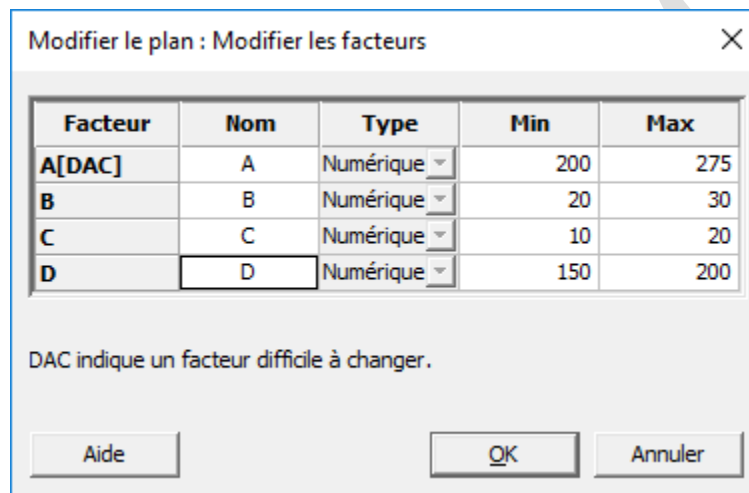
	C8	C9	C10	C11	
	Quantité de lubrifiant	Température de cuisson	Adhésion		
1	10	150	29.3		
2	10	150	39.9		
3	20	150	25.4		
4	20	150	30.9		
5	10	200	30.9		
6	10	200	37.9		
7	200	200	27.3		
8	200	200	33.0		
9	275	20	150	34.0	
10	275	30	10	150	44.9
11	275	20	20	150	13.7
12	275	30	20	150	22.9
13	275	20	10	200	32.3
14	275	30	10	200	41.6
15	275	20	20	200	18.4
16	275	30	20	200	23.4
17	200	20	10	150	28.4
18	200	30	10	150	37.0
19	200	20	20	150	22.8
20	200	30	20	150	26.7
21	200	20	10	200	29.4
22	200	30	10	200	37.7
23	200	20	20	200	22.2
24	200	30	20	200	30.9
25	275	20	10	150	28.4
26	275	30	10	150	41.4
27	275	20	20	150	14.4
28	275	30	20	150	24.4
29	275	20	10	200	31.1
30	275	30	10	200	42.8
31	275	20	20	200	10.8
32	275	30	20	200	21.7
33					
34					
35					

Modifier le plan  
Modifier le plan dans la feuille de travail, y compris les noms des facteurs et la randomisation.

Nous prenons l'option **Modifier les facteurs** et validons par **OK**:



Nous cliquons sur **Spécifier** et mettons:



Nous validons par **OK** pour obtenir:



Minitab - DOE\_SplitPlot.MPJ - [Feuille de travail 1 \*\*\*]

Echier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assistant

↓	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
	OrdreStd	OrdEssai	TypePt	Blocs	DC	A	B	C	D	Adhésion	
1	1	1	1	1	1	200	20	10	150	29.3	
2	2	2	1	1	1	200	30	10	150	39.9	
3	3	3	1	1	1	200	20	20	150	25.4	
4	4	4	1	1	1	200	30	20	150	30.9	
5	5	5	1	1	1	200	20	10	200	30.9	
6	6	6	1	1	1	200	30	10	200	37.9	
7	7	7	1	1	1	200	20	20	200	27.3	
8	8	8	1	1	1	200	30	20	200	33.0	
9	9	9	1	1	2	275	20	10	150	34.0	
10	10	10	1	1	2	275	30	10	150	44.9	
11	11	11	1	1	2	275	20	20	150	13.7	
12	12	12	1	1	2	275	30	20	150	22.9	
13	13	13	1	1	2	275	20	10	200	32.3	
14	14	14	1	1	2	275	30	10	200	41.6	
15	15	15	1	1	2	275	20	20	200	18.4	
16	16	16	1	1	2	275	30	20	200	23.4	
17	17	17	1	1	3	200	20	10	150	28.4	
18	18	18	1	1	3	200	30	10	150	37.0	
19	19	19	1	1	3	200	20	20	150	22.8	
20	20	20	1	1	3	200	30	20	150	26.7	
21	21	21	1	1	3	200	20	10	200	29.4	
22	22	22	1	1	3	200	30	10	200	37.7	
23	23	23	1	1	3	200	20	20	200	22.2	
24	24	24	1	1	3	200	30	20	200	30.9	
25	25	25	1	1	4	275	20	10	150	28.4	
26	26	26	1	1	4	275	30	10	150	41.4	
27	27	27	1	1	4	275	20	20	150	14.4	
28	28	28	1	1	4	275	30	20	150	24.4	
29	29	29	1	1	4	275	20	10	200	31.1	
30	30	30	1	1	4	275	30	10	200	42.8	
31	31	31	1	1	4	275	20	20	200	10.8	
32	32	32	1	1	4	275	30	20	200	21.7	
33											

et si nous relançons l'analyse identiquement à précédemment, nous obtenons un tableau bien plus digeste à lire:

## Régression factorielle en parcelles divisées : Adhésion en ... A; B; C; D

### Analyse de variance

Source	DL	SomCar ajust	CM ajust	Valeur F	Valeur de p
A	1	59.13	59.13	2.94	0.228
Erreur SB	2	40.17	20.08	5.87	0.014
B	1	597.72	597.72	174.56	0.000
C	1	1226.36	1226.36	358.16	0.000
D	1	1.49	1.49	0.43	0.520
A*B	1	14.72	14.72	4.30	0.057
A*C	1	285.01	285.01	83.24	0.000
A*D	1	3.71	3.71	1.08	0.315
B*C	1	13.13	13.13	3.84	0.070
B*D	1	0.81	0.81	0.24	0.634
C*D	1	1.16	1.16	0.34	0.569
A*B*C	1	0.03	0.03	0.01	0.933
A*B*D	1	1.67	1.67	0.49	0.497
A*C*D	1	1.32	1.32	0.39	0.545
B*C*D	1	2.26	2.26	0.66	0.430
A*B*C*D	1	2.70	2.70	0.79	0.389
Erreur SP	14	47.94	3.42		
Total	31				

### Récapitulatif du modèle

S	R carré (SP)	S(SB)	R carré (SB)
1.85042	97.82%	1.44309	59.55%

## Coefficients codés

Terme	Effet	Coef	Coef ErT	Valeur de T	Valeur de p	FIV
Constante		29.247	0.792	36.92	0.001	
A	-2.719	-1.359	0.792	-1.72	0.228	*
B	8.644	4.322	0.327	13.21	0.000	1.00
C	-12.381	-6.191	0.327	-18.93	0.000	1.00
D	0.431	0.216	0.327	0.66	0.520	1.00
A*B	1.356	0.678	0.327	2.07	0.057	1.00
A*C	-5.969	-2.984	0.327	-9.12	0.000	1.00
A*D	-0.681	-0.341	0.327	-1.04	0.315	1.00
B*C	-1.281	-0.641	0.327	-1.96	0.070	1.00
B*D	-0.319	-0.159	0.327	-0.49	0.634	1.00
C*D	0.381	0.191	0.327	0.58	0.569	1.00
A*B*C	0.056	0.028	0.327	0.09	0.933	1.00
A*B*D	-0.456	-0.228	0.327	-0.70	0.497	1.00
A*C*D	-0.406	-0.203	0.327	-0.62	0.545	1.00
B*C*D	0.531	0.266	0.327	0.81	0.430	1.00
A*B*C*D	-0.581	-0.291	0.327	-0.89	0.389	1.00

## Equation de régression en unités non codées

$$\begin{aligned} \text{Adhésion} = & -174 + 0.674 A + 7.54 B + 16.0 C + 0.92 D - 0.0209 A*B - 0.0633 A*C - 0.00315 A*D \\ & - 0.623 B*C - 0.0403 B*D - 0.0724 C*D + 0.00220 A*B*C + 0.000137 A*B*D \\ & + 0.000267 A*C*D + 0.00337 B*C*D - 0.000012 A*B*C*D \end{aligned}$$

Moyenne de l'équation sur les sous-blocs (whole plots).

## Structure des alias

Facteur	Nom
A	A
B	B
C	C
D	D

### Alias

I  
A  
B  
C  
D  
AB  
AC  
AD  
BC  
BD  
CD  
ABC  
ABD  
ACD  
BCD  
ABCD

## Ajustements et diagnostics pour les observations aberrantes

Observation	Adhésion	Valeur ajustée	Résiduelle	Val. résid. norm.	Valeurs résiduelles du sous-bloc	
15	18.40	14.60	2.79	2.28	1.0125	R
31	10.80	14.60	-2.79	-2.28	-1.0125	R

*R : Valeur résiduelle élevée*

## 17.19. Exercice 190.: Plan de criblage définitif (definitive screening design – DSD) (*WP*)

Minitab® Statistical Software 18.1

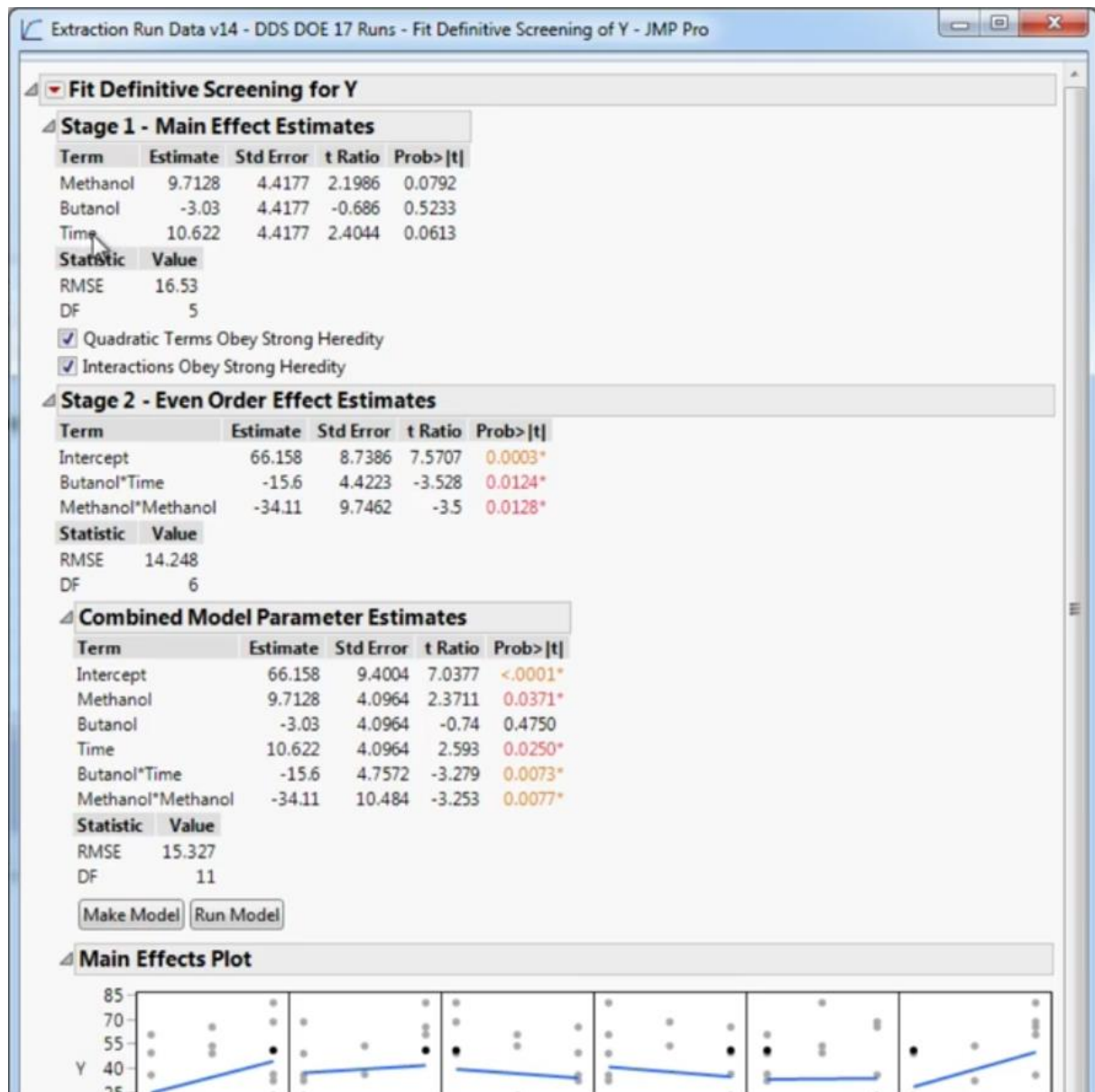
Quelques petits rappels au préalable:

1. Les DSD sont relativement récents. On est parfois surpris que des chercheurs découvrent encore tout récemment des propriétés mathématiques qui permettent de les améliorer.
2. On manque encore de recul sur la robustesse de ses plans. Donc à utiliser avec des pincettes même si pour l'instant il semblerait qu'on ne leur ait trouvé aucune faille.
3. Nous n'avons pas étudié les détails mathématiques de ces plans qui sont basés sur les matrices de conférences et les plans optimaux. Votre professeur est en pleine rédaction d'un chapitre sur le sujet depuis 2 ans mais peine à simplifier les détails des calculs. Il faudra donc accepter que ce que nous allons voir se fera sans preuves.
4. Nous allons prendre ici comme exemple celui qui est fait avec le logiciel JMP qui avait les DSD bien avant Minitab et qui globalement est bien plus performant que ce dernier (du moins à ce jour).

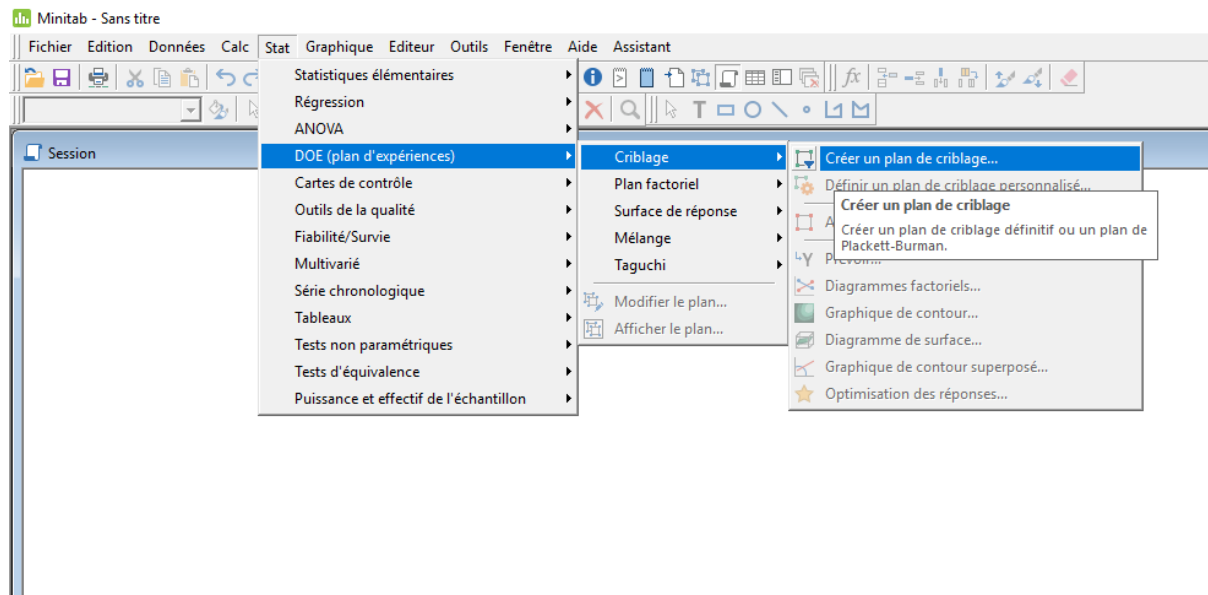
Voici le plan d'expérience de criblage définitif que nous souhaiterions analyser avec Minitab 18.1:

	Methanol	Ethanol	Propanol	Butanol	pH	Time	Y
1	0	10	10	5	6	1	0.13
2	10	10	10	0	9	1	18.57
3	10	0	10	0	6	1.5	32.06
4	5	10	10	10	9	2	65.07
5	0	0	10	10	7.5	1	49.20
6	5	5	5	5	7.5	1.5	53.40
7	0	10	0	10	9	1.5	15.97
8	5	0	0	0	6	1	48.80
9	0	0	10	0	9	2	35.58
10	10	5	10	10	6	2	36.01
11	0	0	0	10	6	2	2.10
12	10	10	0	0	7.5	2	79.94
13	10	0	0	5	9	2	68.19
14	10	0	5	10	9	1	22.39
15	0	10	5	0	6	2	60.32
16	0	5	0	0	9	1	8.65
17	10	10	0	10	6	1	50.75

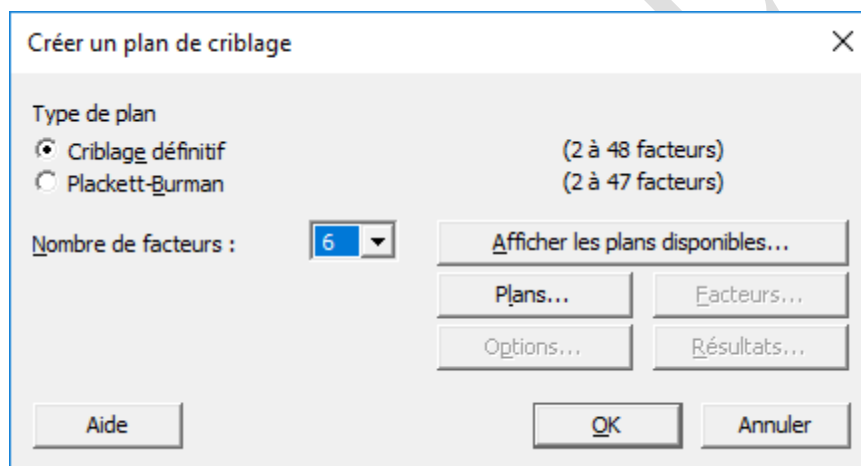
Et le type de résultat auquel nous nous attendons (rappelez-vous dans le cours théorique que j'ai insisté sur l'importance de la reproductibilité avec des exemples fameux dans le problème de la R&D!!!):



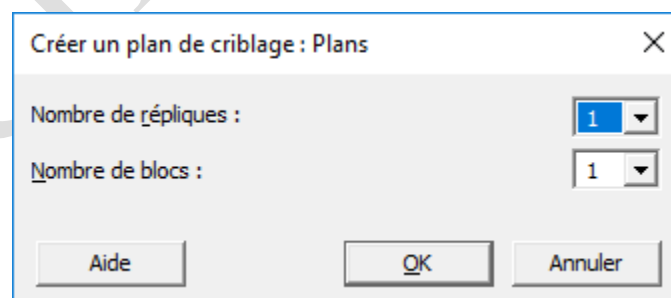
Donc d'abord voyons si nous pouvons reproduire ce plan à 6 facteurs continus et 4 mesures supplémentaires ("extra runs") avec Minitab. Pour cela nous allons dans le menu **Stat/DOE (plan d'expériences)/Criblage/Créer un plan de criblage...**



Ensuite nous mettons 6 facteurs comme ci-dessous:



Dans le bouton **Plans...** nous prenons:



Et ensuite après avoir validé par **OK** et cliqué sur **Facteurs...** nous mettons:

Créer un plan de criblage : Facteurs

Facteur	Nom	Type	Min	Max
A	Méthanol	Numériqu	0	10
B	Ethanol	Numériqu	0	10
C	Propanol	Numériqu	0	10
D	Butanol	Numériqu	0	10
E	pH	Numériqu	6	9

Aide OK Annuler

Créer un plan de criblage : Facteurs

Facteur	Nom	Type	Min	Max
C	Propanol	Numériqu	0	10
D	Butanol	Numériqu	0	10
E	pH	Numériqu	6	9
F	Temps	Numériqu	1	2

Aide OK Annuler

Nous cliquons sur **Options** après avoir validé afin de désactiver la randomisation:

Créer un plan de criblage: Options

Randomiser les essais:

Base du générateur de nombres aléatoires :

Aide OK Annuler

Nous voyons donc qu'il n'y a ni a priori aucune option pour ajouter des mesures complémentaires optimales, ni pour rajouter des points centraux! Nous sommes donc plus limités que le logiciel JMP de SAS ou que l'add-in de Jones & Nachtsheim qui permet dans Minitab de faire des DSD!

Le seul résultat que nous pouvons obtenir est effectivement:



Minitab - Sans titre - [Feuille de travail 1 \*\*\*]

Echier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assistant

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
	OrdreStd	OrdEssai	TypePt	Blocs	Méthanol	Ethanol	Propanol	Butanol	pH	Temps	
1	1	1	2	1	5	10	10	10	9.0	2.0	
2	2	2	2	1	5	0	0	0	6.0	1.0	
3	3	3	2	1	10	5	10	0	6.0	2.0	
4	4	4	2	1	0	5	0	10	9.0	1.0	
5	5	5	2	1	10	10	5	10	6.0	1.0	
6	6	6	2	1	0	0	5	0	9.0	2.0	
7	7	7	2	1	10	0	10	5	9.0	1.0	
8	8	8	2	1	0	10	0	5	6.0	2.0	
9	9	9	2	1	10	0	0	10	7.5	2.0	
10	10	10	2	1	0	10	10	0	7.5	1.0	
11	11	11	2	1	10	10	0	0	9.0	1.5	
12	12	12	2	1	0	0	10	10	6.0	1.5	
13	13	13	0	1	5	5	5	5	7.5	1.5	
14											
15											

Donc nous allons construire un plan de criblage personnalisé en saisissant dans une nouvelle feuille de travail (ou en ouvrant le fichier *DOE\_DSD\_JMP.mpj*):

Minitab - Sans titre - [Feuille de travail 1 \*\*\*]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assi

	C1	C2	C3	C4	C5	C6	C7	C8
	Méthanol	Ethanol	Propanol	Butanol	pH	Temps	Rendement	
1	0	10	10	5	6.0	1.0	0.13	
2	10	10	10	0	9.0	1.0	18.57	
3	10	0	10	0	6.0	1.5	32.06	
4	5	10	10	10	9.0	2.0	65.07	
5	0	0	10	10	7.5	1.0	49.20	
6	5	5	5	5	7.5	1.5	53.40	
7	0	10	0	10	9.0	1.5	15.97	
8	5	0	0	0	6.0	1.0	48.80	
9	0	0	10	0	9.0	2.0	35.58	
10	10	5	10	10	6.0	2.0	36.01	
11	0	0	0	10	6.0	2.0	2.10	
12	10	10	0	0	7.5	2.0	79.94	
13	10	0	0	5	9.0	2.0	68.19	
14	10	0	5	10	9.0	1.0	22.39	
15	0	10	5	0	6.0	2.0	60.32	
16	0	5	0	0	9.0	1.0	8.65	
17	10	10	0	10	6.0	1.0	50.75	
18								
19								

Nous allons ensuite dans le menu **Stat/DOE (plan d'expériences)/Criblage/Définir un plan de criblage personnalisée...**:

Minitab - Sans titre - [Feuille de travail 1 \*\*\*]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assistant

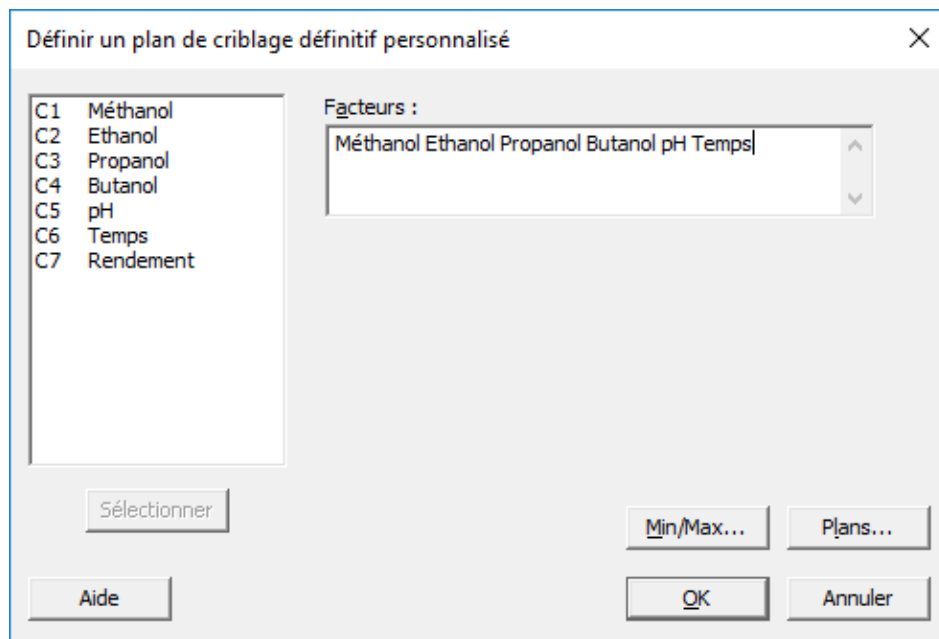
Statistiques élémentaires  
Régression  
ANOVA  
**DOE (plan d'expériences)**  
Cartes de contrôle  
Outils de la qualité  
Fiabilité/Survie  
Multivarié  
Série chronologique  
Tableaux  
Tests non paramétriques  
Tests d'équivalence  
Puissance et effectif de l'échantillon

Criblage  
Plan factoriel  
Surface de réponse  
Mélange  
Taguchi  
Modifier le plan...  
Afficher le plan...

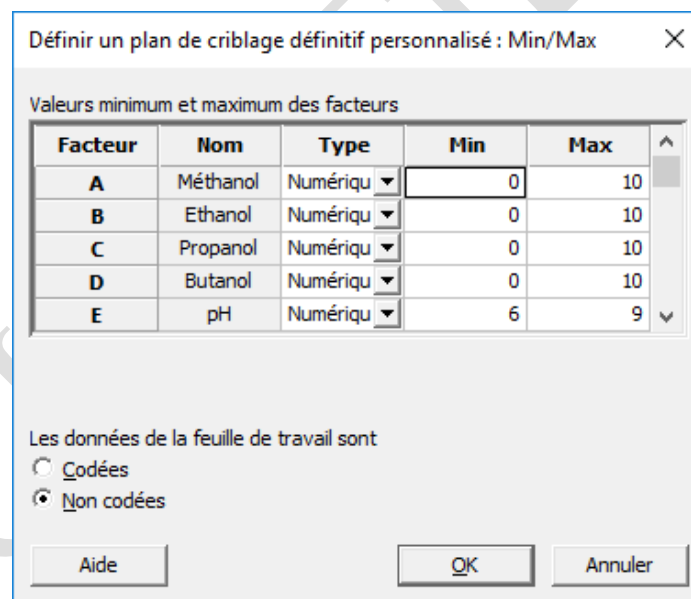
Créer un plan de criblage...  
**Définir un plan de criblage personnalisée...**  
Analyse  
Prévoir  
Diagrammes factoriels...  
Graphique de contour...  
Diagramme de surface...  
Graphique de contour superposé...  
Optimisation des réponses...

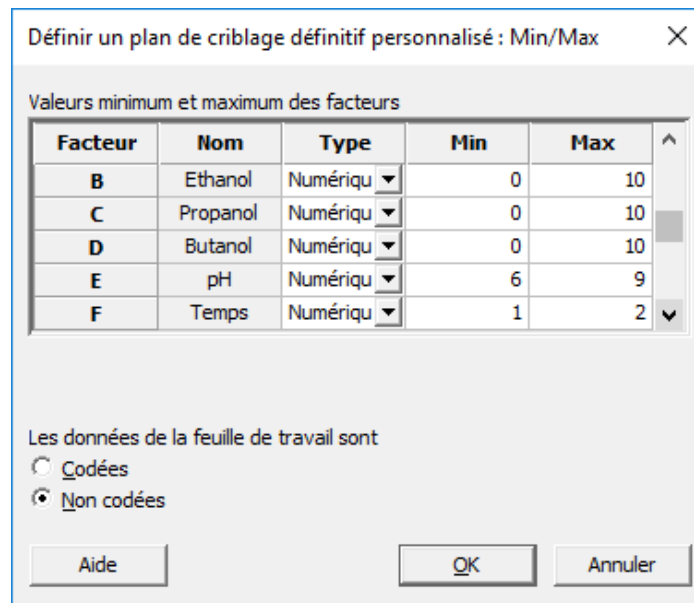
	C1	C2	C3	C4	C5	C6	C7	C8	C15	C16
	Méthanol	Ethanol	Propanol							
1	0	10	10							
2	10	10	10							
3	10	0	10							
4	5	10	10							
5	0	0	10							
6	5	5	5							
7	0	10	0							
8	5	0	0							
9	0	0	10							
10	10	5	10	10	6.0	2.0	36.01			
11	0	0	0	10	6.0	2.0	2.10			
12	10	10	0	0	7.5	2.0	79.94			
13	10	0	0	5	9.0	2.0	68.19			
14	10	0	5	10	9.0	1.0	22.39			
15	0	10	5	0	6.0	2.0	60.32			
16	0	5	0	0	9.0	1.0	8.65			
17	10	10	0	10	6.0	1.0	50.75			
18										
19										
20										
21										

Dans la boîte de dialogue qui apparaît nous mettons:



Nous cliquons sur le bouton **Min/Max...** (c'est de toute façon obligé!) pour vérifier que les niveaux soient les bons:





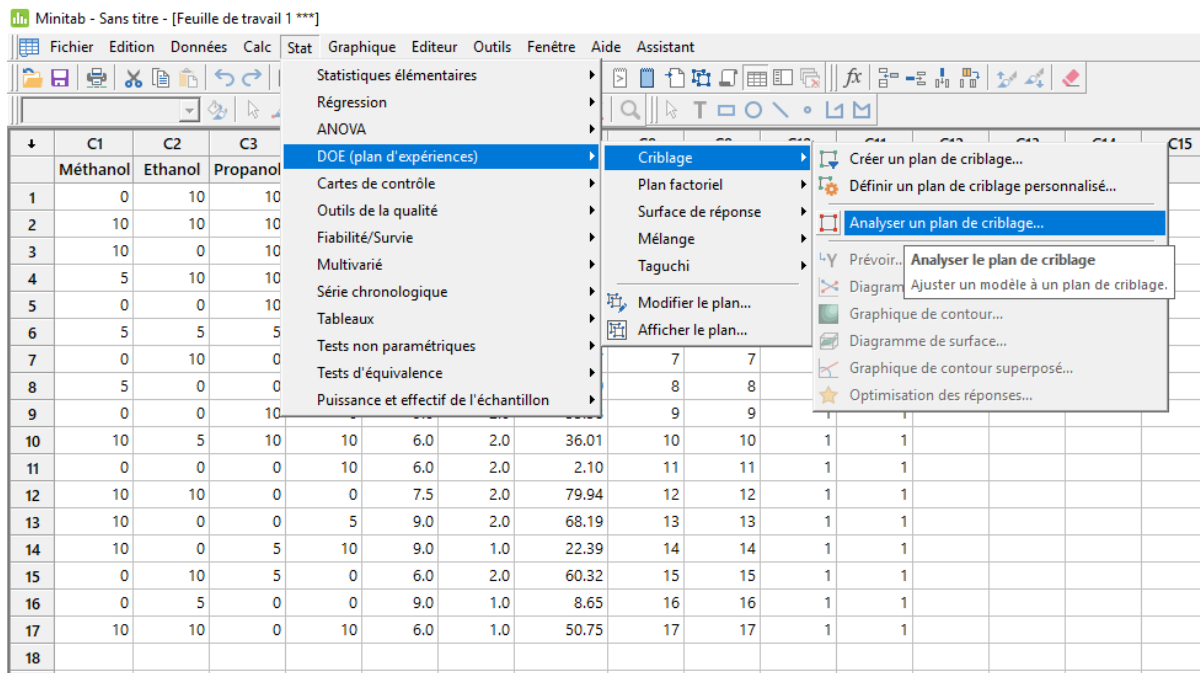
Et nous validons par **OK** pour obtenir:

Minitab - Sans titre - [Feuille de travail 1 \*\*\*]

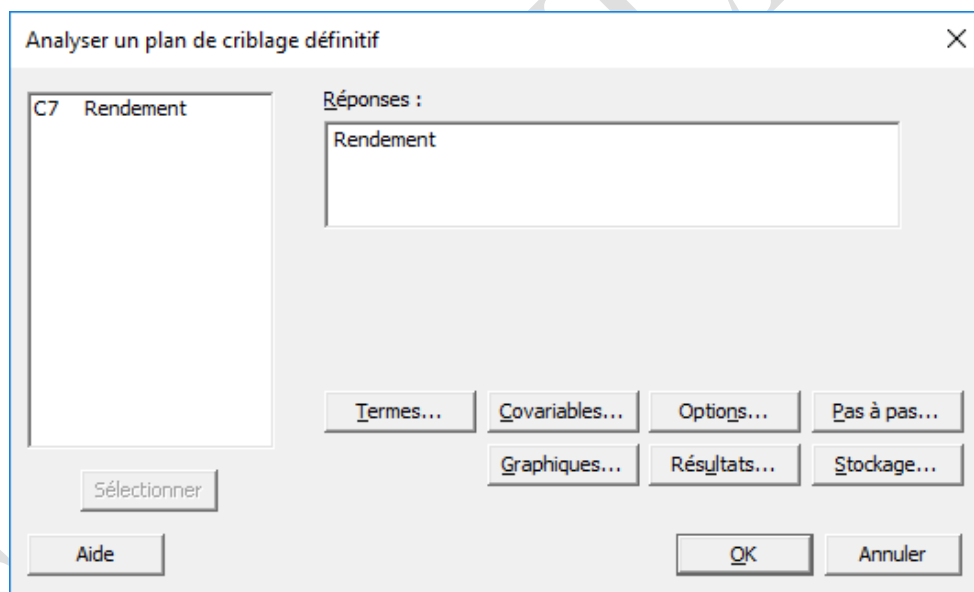
Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide Assistant

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
	Méthanol	Ethanol	Propanol	Butanol	pH	Temps	Rendement	OrdreStd	OrdEssai	Blocs	TypePt	
1	0	10	10	5	6.0	1.0	0.13	1	1	1	1	
2	10	10	10	0	9.0	1.0	18.57	2	2	1	1	
3	10	0	10	0	6.0	1.5	32.06	3	3	1	1	
4	5	10	10	10	9.0	2.0	65.07	4	4	1	1	
5	0	0	10	10	7.5	1.0	49.20	5	5	1	1	
6	5	5	5	5	7.5	1.5	53.40	6	6	1	1	
7	0	10	0	10	9.0	1.5	15.97	7	7	1	1	
8	5	0	0	0	6.0	1.0	48.80	8	8	1	1	
9	0	0	10	0	9.0	2.0	35.58	9	9	1	1	
10	10	5	10	10	6.0	2.0	36.01	10	10	1	1	
11	0	0	0	10	6.0	2.0	2.10	11	11	1	1	
12	10	10	0	0	7.5	2.0	79.94	12	12	1	1	
13	10	0	0	5	9.0	2.0	68.19	13	13	1	1	
14	10	0	5	10	9.0	1.0	22.39	14	14	1	1	
15	0	10	5	0	6.0	2.0	60.32	15	15	1	1	
16	0	5	0	0	9.0	1.0	8.65	16	16	1	1	
17	10	10	0	10	6.0	1.0	50.75	17	17	1	1	
18												
19												

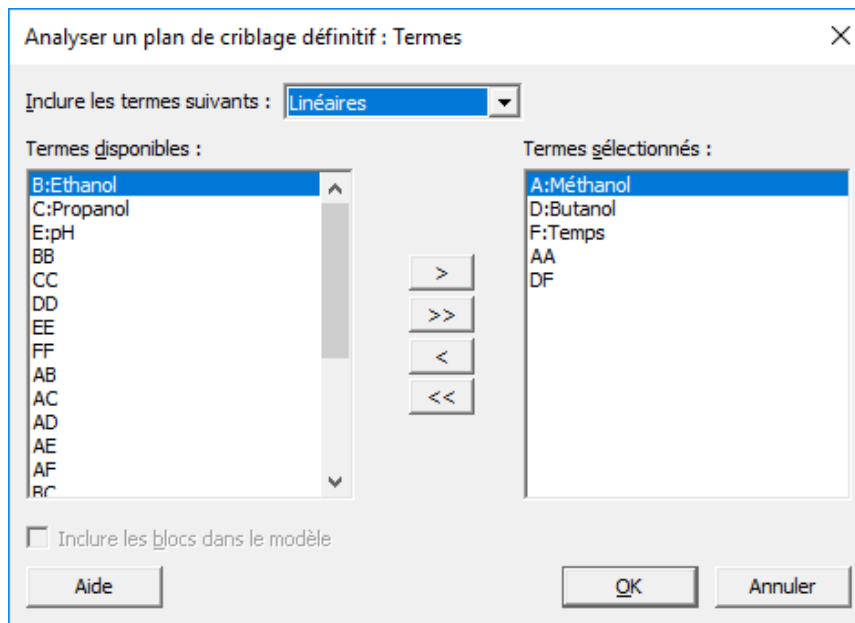
Ensuite, nous allons dans le menu **Stat/DOE (plan d'expériences)/Criblage/Analyser un plan de criblage...**:



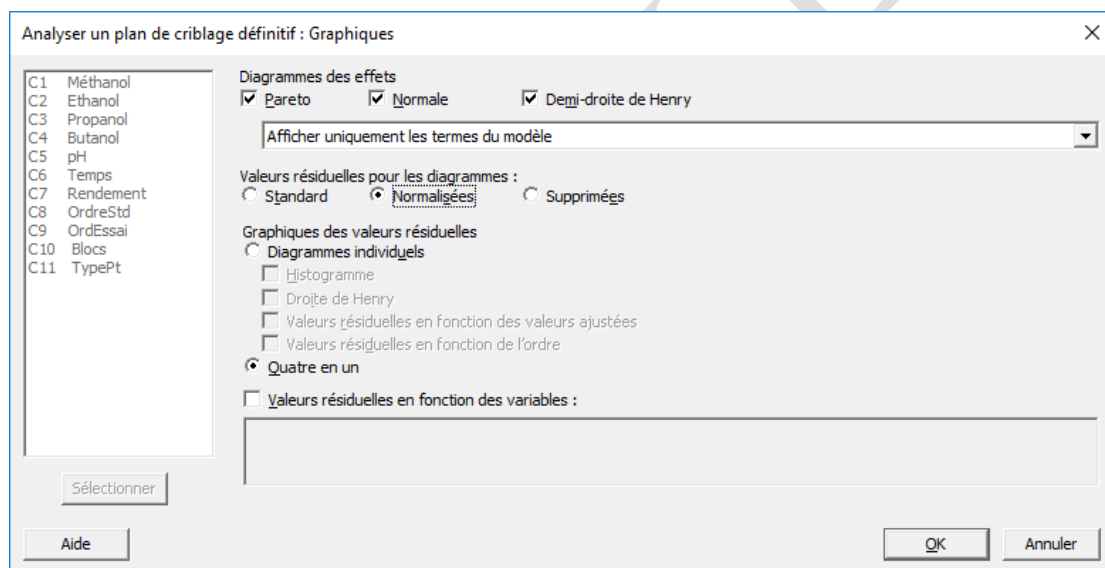
Nous mettons dans la boîte de dialogue qui apparaît:



Nous cliquons sur le bouton **Termes...** pour mettre la même chose qu'avec le logiciel JMP:



Dans le bouton **Graphiques...** nous prenons:



Ce qui nous donne d'abord dans la fenêtre de session:

## Modèle du plan de criblage : Rendement en fonction de ... nol; Temps

## Analyse de variance

Source	DL	SomCar ajust	CM ajust	Valeur F	Valeur de p
Modèle	5	6695.1	1339.0	5.70	0.008
Linéaires	3	3028.8	1009.6	4.30	0.031
Méthanol	1	1320.4	1320.4	5.62	0.037
Butanol	1	128.6	128.6	0.55	0.475
Temps	1	1579.8	1579.8	6.73	0.025
Carré	1	2486.6	2486.6	10.59	0.008
Méthanol*Méthanol	1	2486.6	2486.6	10.59	0.008
Interactions à 2 facteur(s)	1	2526.3	2526.3	10.75	0.007
Butanol*Temps	1	2526.3	2526.3	10.75	0.007
Erreur	11	2584.0	234.9		
Total	16	9279.1			

## Récapitulatif du modèle

S	R carré	R carré (ajust)	R carré (prév)
15.3268	72.15%	59.49%	34.74%

## Coefficients codés

Terme	Coef	Coef ErT	Valeur de T	Valeur de p	FIV
Constante	66.16	9.40	7.04	0.000	
Méthanol	9.71	4.10	2.37	0.037	1.00
Butanol	-3.03	4.10	-0.74	0.475	1.00
Temps	10.62	4.10	2.59	0.025	1.00
Méthanol*Méthanol	-34.1	10.5	-3.25	0.008	1.16
Butanol*Temps	-15.60	4.76	-3.28	0.007	1.16

## Equation de régression en unités non codées

$$\text{Rendement} = -53.3 + 15.59 \text{ Méthanol} + 8.75 \text{ Butanol} + 52.4 \text{ Temps} - 1.364 \text{ Méthanol*Méthanol} - 6.24 \text{ Butanol*Temps}$$

## Structure des alias (jusqu'à l'ordre 2)

Facteur	Nom
A	Méthanol
B	Ethanol
C	Propanol
D	Butanol
E	pH
F	Temps

## Alias

$$I + 0.77 \text{ BB} + 0.51 \text{ CC} + 0.64 \text{ DD} + 0.51 \text{ EE} + 0.64 \text{ FF} - 0.51 \text{ AC} + 0.13 \text{ AD} - 0.13 \text{ AF} + 0.26 \text{ BC} + 0.64 \text{ BD} + 0.26 \text{ BE} + 0.90 \text{ BF} + 0.64 \text{ CD} + 0.77 \text{ CE} + 0.64 \text{ CF} + 0.90 \text{ DE} + 0.90 \text{ EF}$$

A

D

F

$$AA + 0.06 \text{ BB} + 0.38 \text{ CC} + 0.22 \text{ DD} + 0.38 \text{ EE} + 0.22 \text{ FF} + 0.62 \text{ AC} - 0.16 \text{ AD} + 0.16 \text{ AF} - 0.31 \text{ BC} - 0.78 \text{ BD} - 0.31 \text{ BE} - 1.09 \text{ BF} - 0.78 \text{ CD} - 0.94 \text{ CE} - 0.78 \text{ CF} - 1.09 \text{ DE} - 1.09 \text{ EF}$$

$$DF - 0.16 \text{ BB} + 0.23 \text{ CC} + 0.04 \text{ DD} + 0.23 \text{ EE} + 0.04 \text{ FF} + 0.77 \text{ AC} - 0.19 \text{ AD} + 0.19 \text{ AF} + 0.61 \text{ BC} + 0.04 \text{ BD} + 0.61 \text{ BE} - 0.35 \text{ BF} + 0.04 \text{ CD} - 0.16 \text{ CE} + 0.04 \text{ CF} - 0.35 \text{ DE} - 0.35 \text{ EF}$$

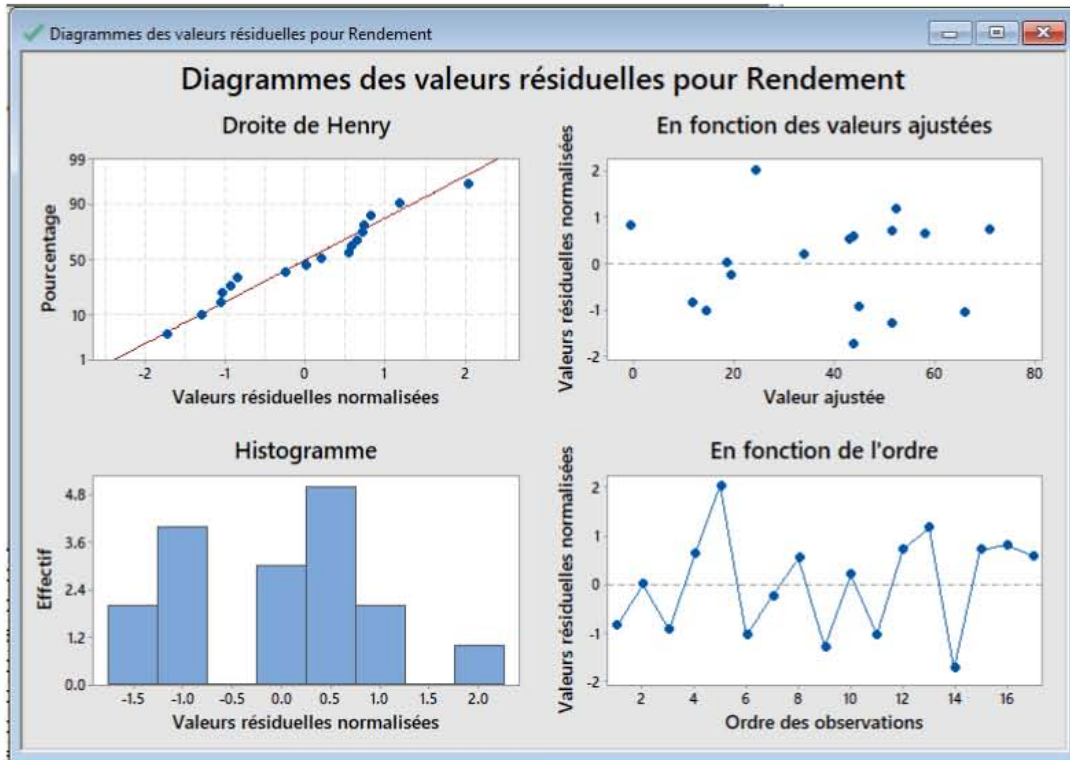
## Ajustements et diagnostics pour les observations aberrantes

Observation	Rendement	Valeur ajustée	Résiduelle	Val. résid. norm.
5	49.20	24.28	24.92	2.03 R

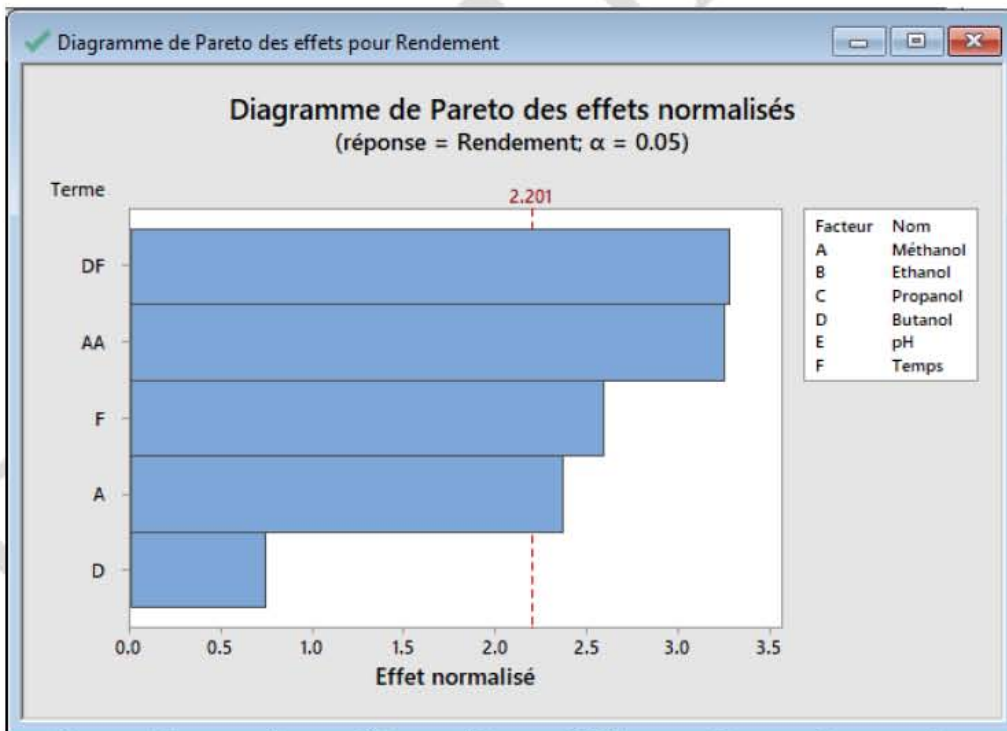
R : Valeur résiduelle élevée

Ce qui est conforme avec JMP à l'exception de la mise en page et de certaines informations extraites.

En ce qui concerne les graphiques nous obtenons:

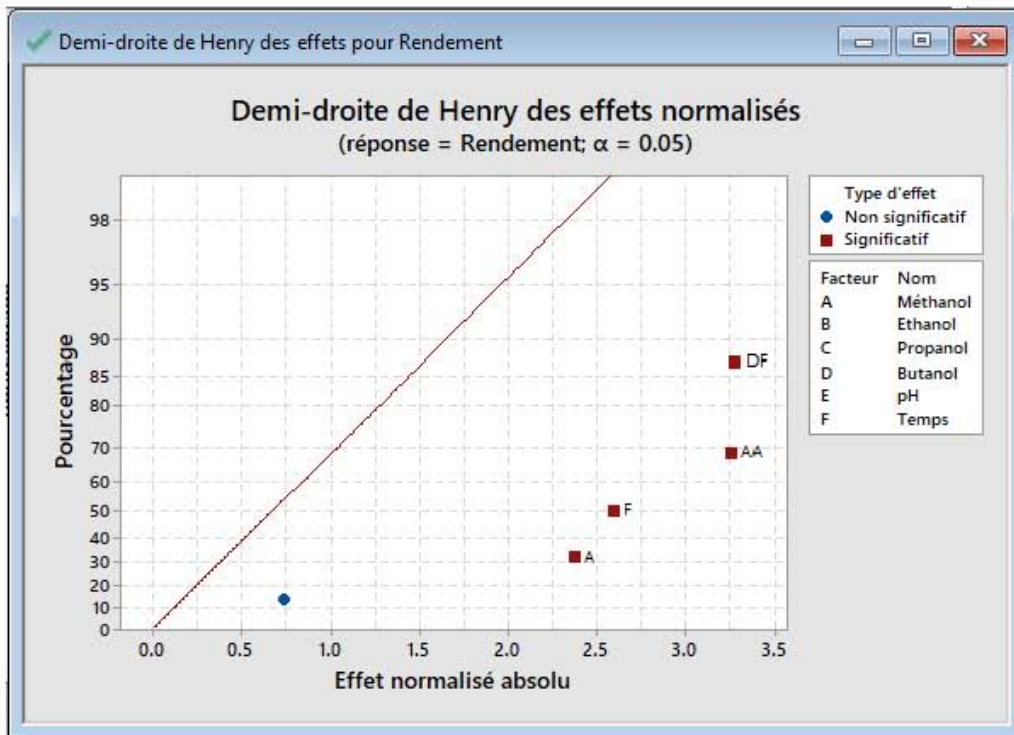


avec:

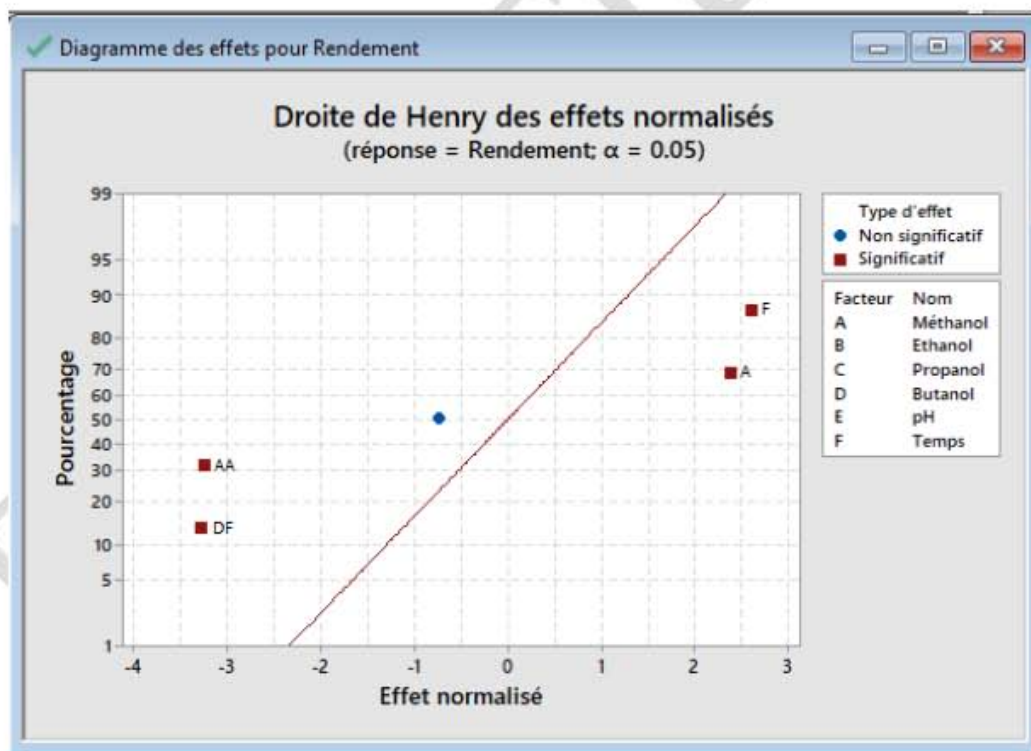


et:

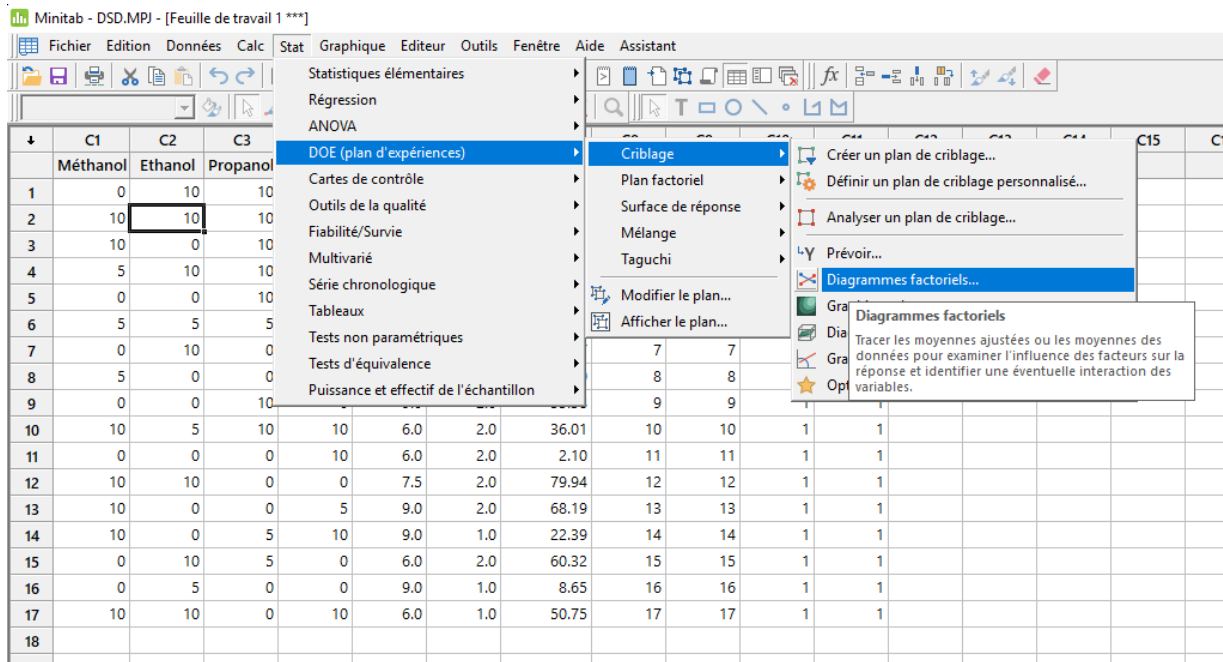




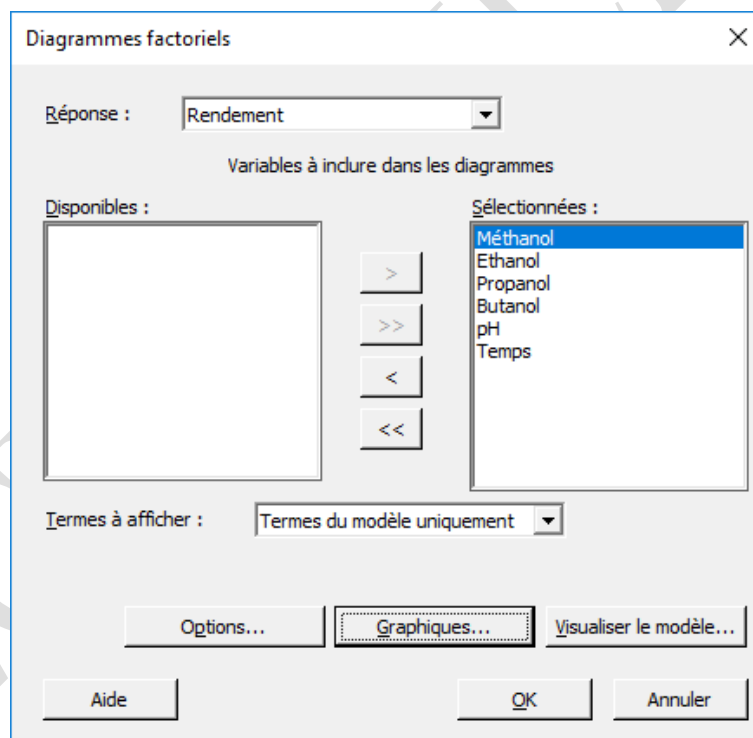
et:



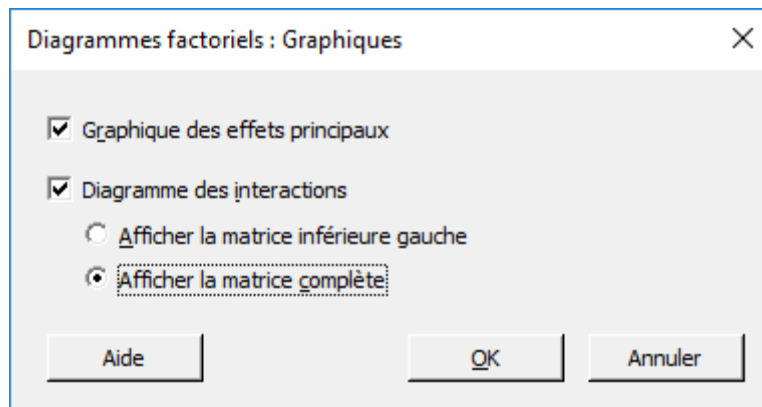
Voyons également par curiosité les diagrammes factoriels:



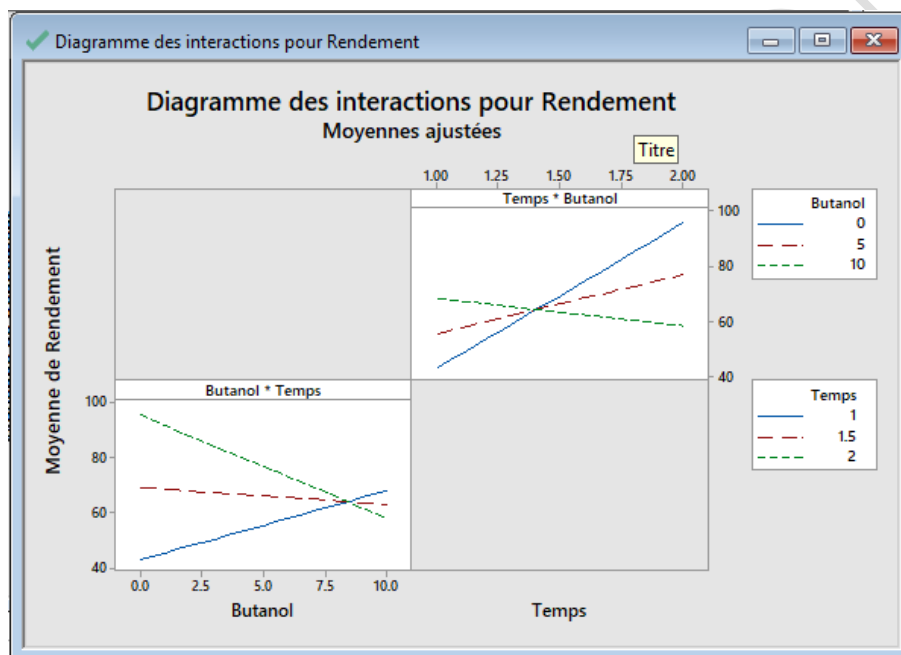
En prenant:



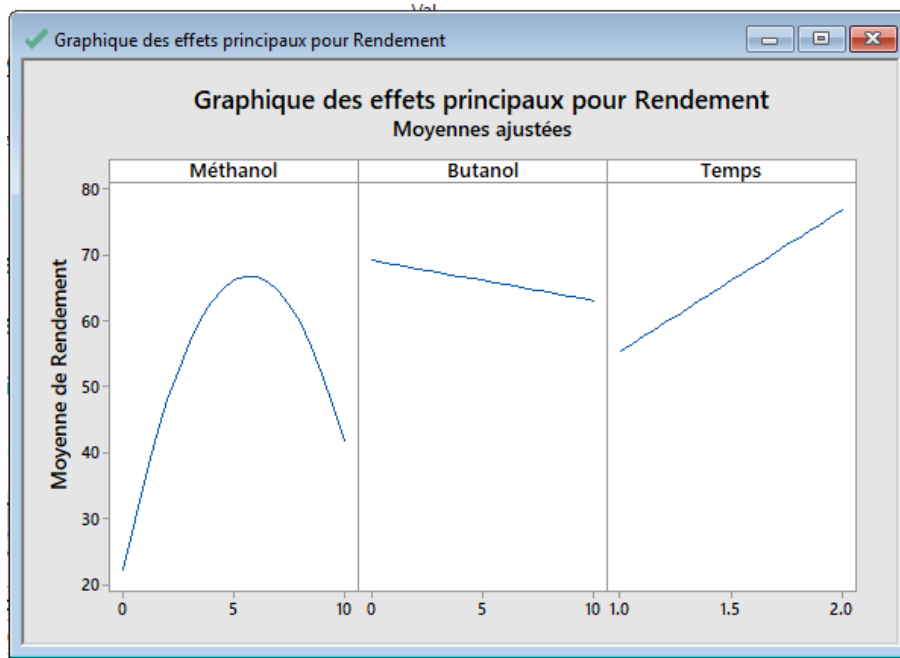
En cliquant sur le bouton **Graphiques...** nous prenons:



Nous validons deux fois par **OK** pour obtenir:



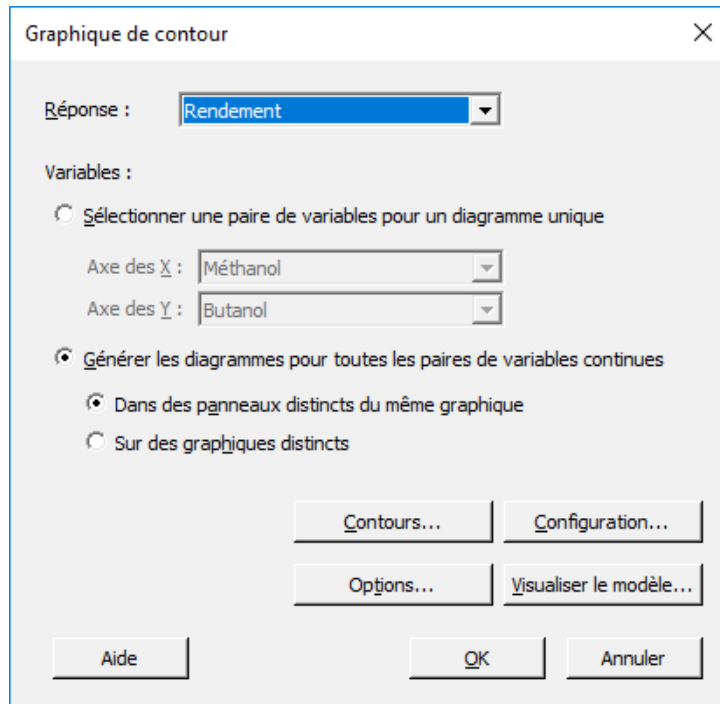
Ainsi que:



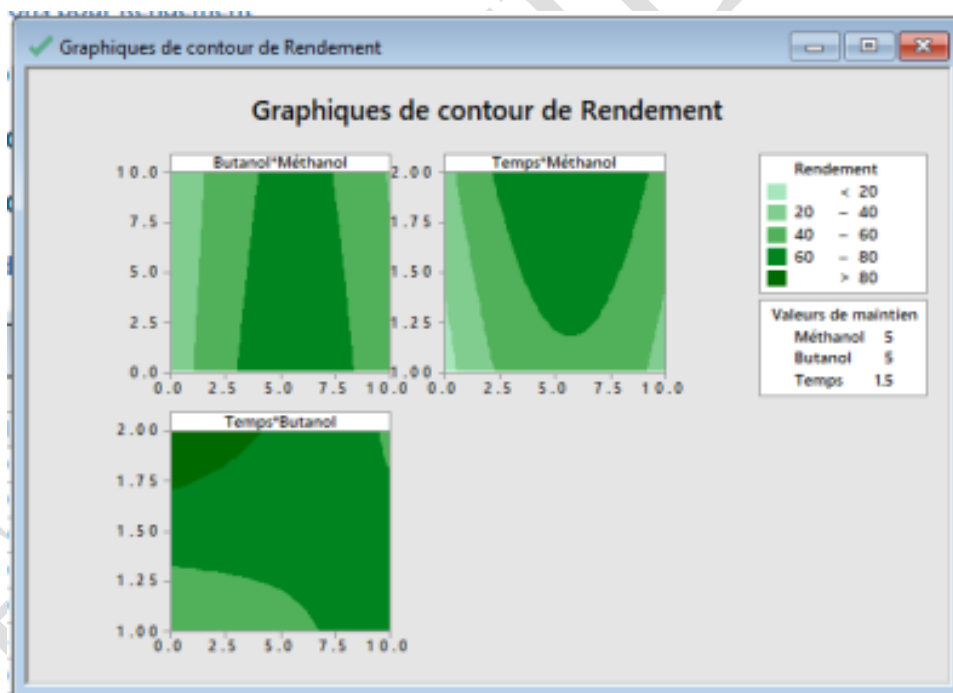
Nous pouvons aussi afficher les graphiques de contour:

	C1	C2	C3														
	Méthanol	Ethanol	Propanol														
1	0	10	10														
2	10	10	10														
3	10	0	10														
4	5	10	10														
5	0	0	10														
6	5	5	5														
7	0	10	0														
8	5	0	0														
9	0	0	10														
10	10	5	10	10	6.0	2.0	36.01	10	10	1	1						
11	0	0	0	10	6.0	2.0	2.10	11	11	1	1						
12	10	10	0	0	7.5	2.0	79.94	12	12	1	1						
13	10	0	0	5	9.0	2.0	68.19	13	13	1	1						
14	10	0	5	10	9.0	1.0	22.39	14	14	1	1						
15	0	10	5	0	6.0	2.0	60.32	15	15	1	1						
16	0	5	0	0	9.0	1.0	8.65	16	16	1	1						
17	10	10	0	10	6.0	1.0	50.75	17	17	1	1						
18																	

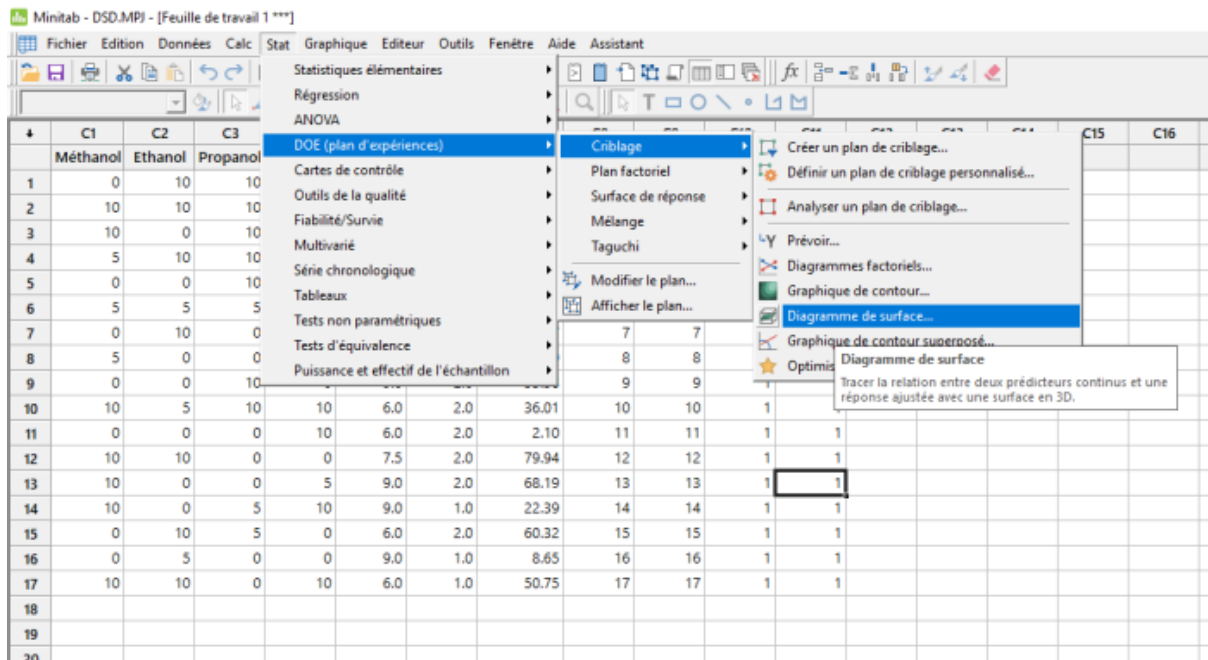
Avec les paramètres suivants:



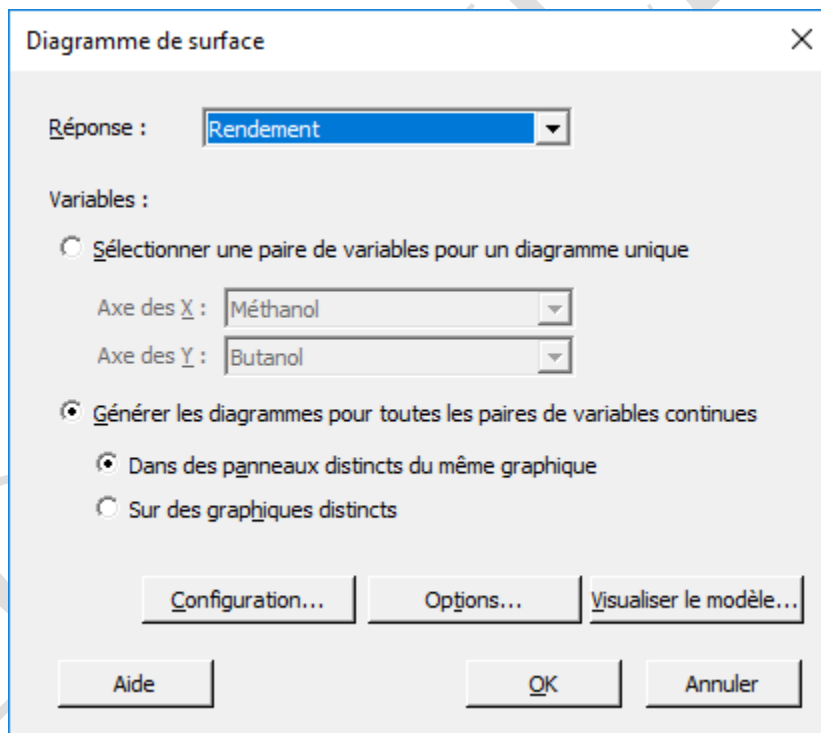
Ce qui donne:



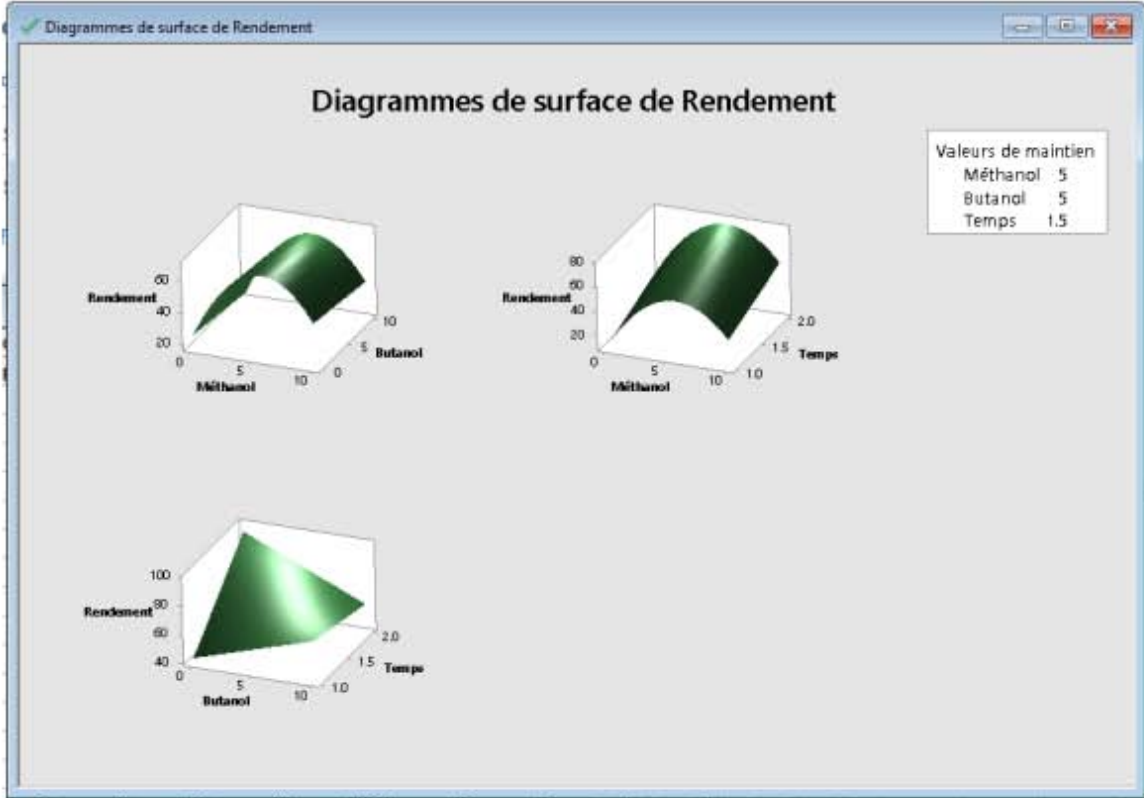
Et nous pouvons aussi afficher les diagrammes de surface:



Avec les paramètres suivants:



Ce qui donne:



etc. etc.

## 17.20. Exercice 191.: Plan de réponse de surface composite face centrée (modèle du second degré ou "modèle quadratique")

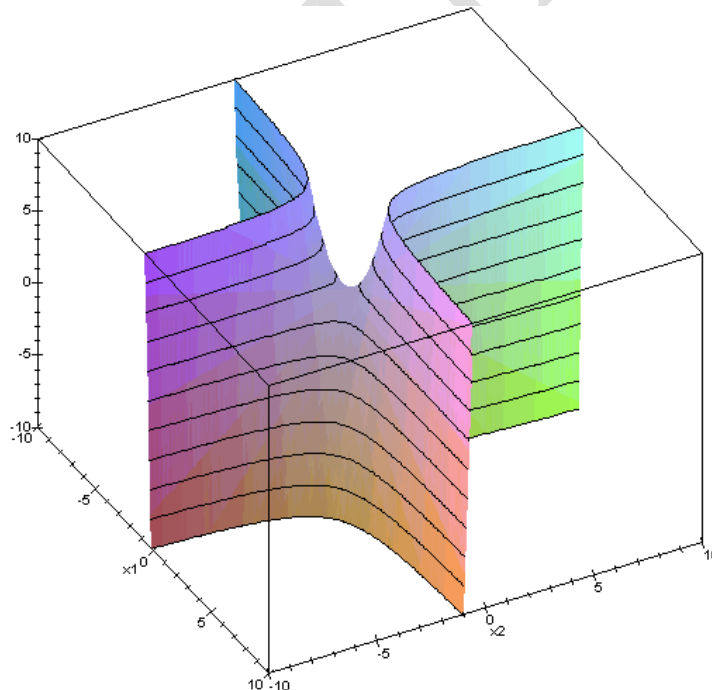
Minitab® Statistical Software 15.1.1.0

Nous n'avons pas étudié dans le cours de génie industriel les détails de la mathématique derrière tous les types de plans de réponse de surface (nous avons cependant démontré mathématiquement et non sans peine l'origine des plans de Box-Behnken) pour la simple raison qu'il s'agit de techniques empiriques. Nous allons donc tenter d'introduire le concept ici même de manière extrêmement condensée (pour plus de détails se référer comme à l'habitude au cours théorique).

Rappelons d'abord que les plans factoriels avec interactions sont en réalité des développements de Taylor-Maclaurin au deuxième ordre et au premier degré (appelé parfois "modèle synergique"):

$$y \cong a_0 + \sum a_i x_i + \sum_{i \neq j} a_{ij} x_i x_j + e$$

et que dans le cas particulier de deux variables nous parlons de modèle parabololoïde hyperbolique représenté typiquement par:



Pour passer à un modèle au second degré (quadratique complet):

$$y = a_0 + \sum a_i x_i + \sum a_{ij} x_i x_j = a_0 + \sum a_i x_i + \sum_{i \neq j} a_{ij} x_i x_j + \sum_i a_{ii} x_i x_i$$

Soit dans le cas de deux variables nous parlons de "surface de réponse" (au-delà il s'agit de volume ou d'hypervolume):



$$y = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2 + \underbrace{a_{11}x_1^2 + a_{22}x_2^2}_{\text{courbure}}$$

Les plans considérant le modèle incluant les trois termes:

$$a_{12}x_1x_2 + \underbrace{a_{11}x_1^2 + a_{22}x_2^2}_{\text{courbure}}$$

sont appelés "**plans composites de Doehlert**". Nous aurons donc à estimer 4 termes:

$a_0$  : terme constant

$a_i$  : terme de premier degré

$a_{ij}$  : terme rectangle

$a_{ii}$  : terme carré

Nous voyons que dans le cas d'un plan à 2 facteurs, il nous faudra estimer 6 coefficients:

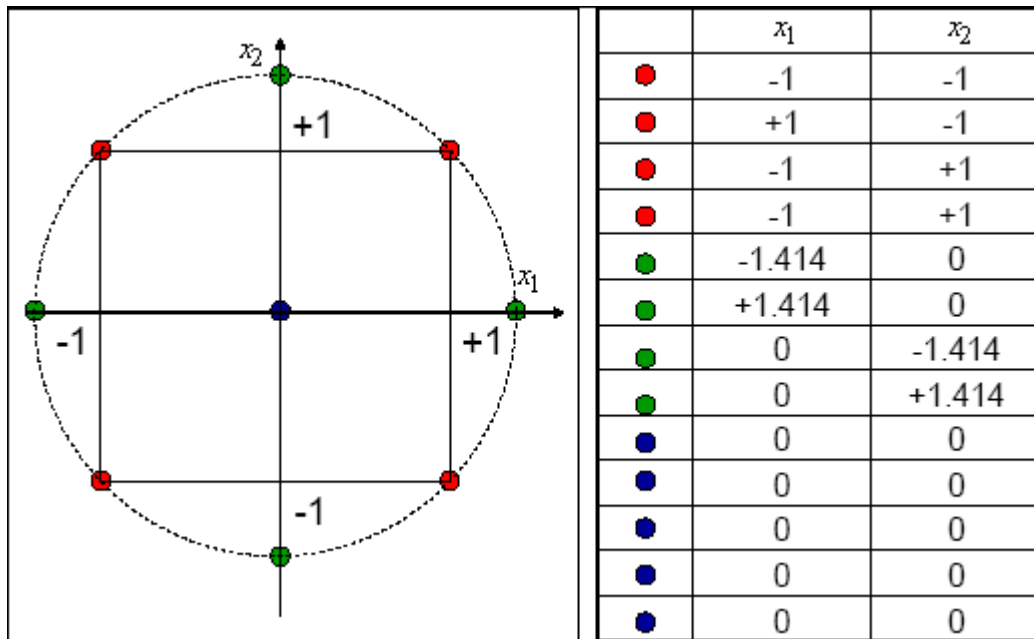
$$y = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2 + a_{11}x_1^2 + a_{22}x_2^2$$

et pour 3 facteurs il nous faudra estimer 10 coefficients:

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3 + a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2$$

L'idée souvent retenue pour passer à un modèle du second degré après une étude factorielle tout en conservant les essais déjà réalisés consiste simplement à compléter par les essais pour l'estimation du modèle de degré plus élevé.

Dans le cas de deux variables, une première méthode empirique consiste à utiliser une "**matrice composite équiradiale**" (cas particulier d'un plan composite général) ce qui impose une forte symétrie à la surface de réponse et implique une isovariance et nous donne aussi 13 essais selon le principe intuitif du cercle circonscrit au carré d'essais:

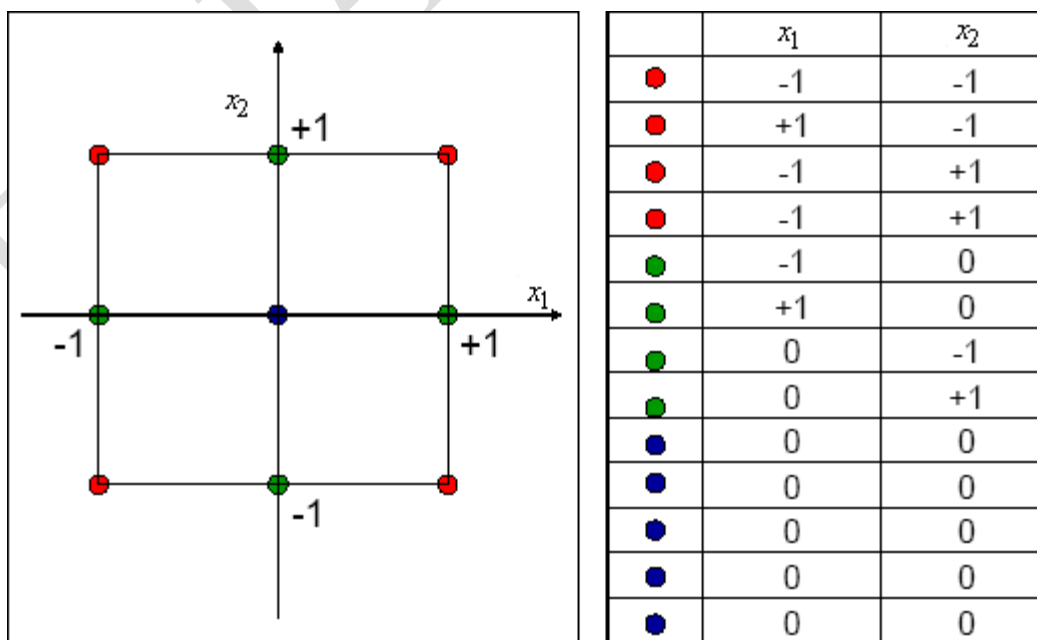


Effectivement si le carré à un demi-sommet de 1, alors sa demi-diagonale étant égale au rayon du cercle circonscrit donnera:

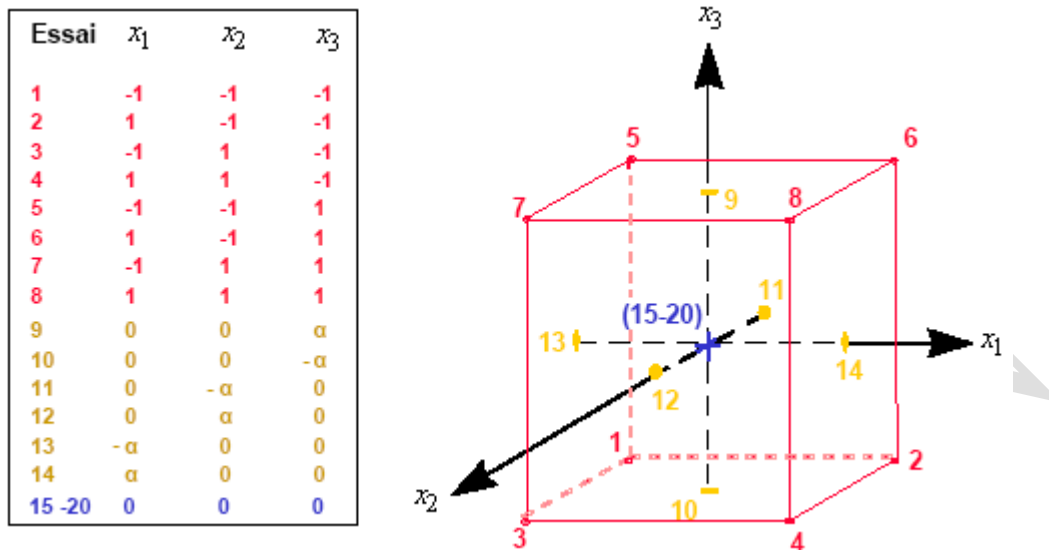
$$R = \sqrt{1^2 + 1^2} = \sqrt{2} \cong 1.414$$

Donc avec cette méthode les points de contrôles ne sont pas au centre des arêtes du rectangle (ou face du cube ou hypercube c'est selon....). Au vu de la construction du plan (matrice) d'expérience donnée ci-dessus, le lecteur aura remarqué que finalement nous avons finalement 5 niveaux par facteurs pour un plan composite équiradial à 3 facteurs.

Une deuxième méthode empirique consiste à utiliser un "matrice composite face-centrée" (cas particulier d'un plan composite) de 13 essais aussi:

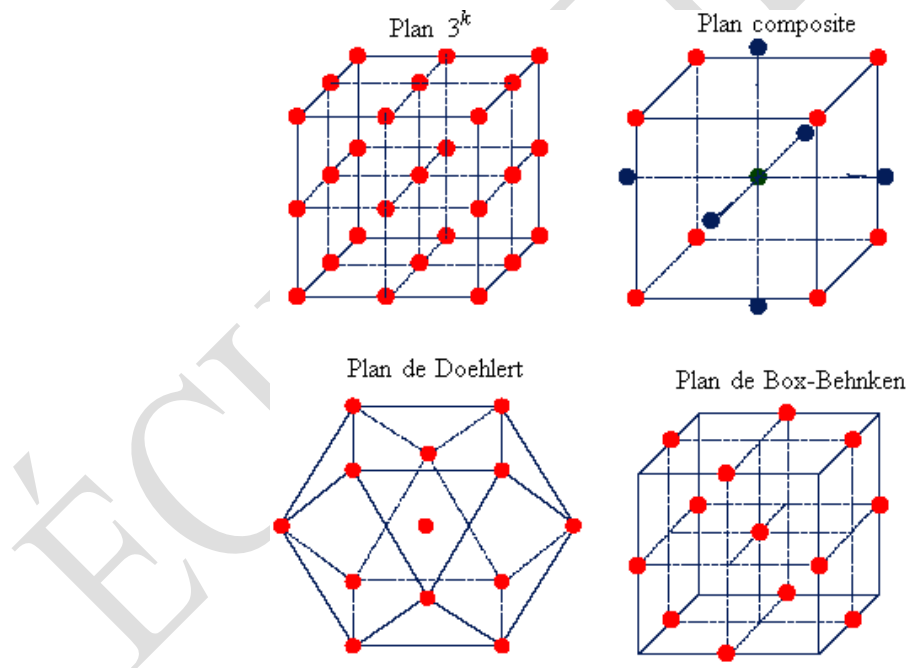


Dans un cas à trois variables (facteurs) nous avons empiriquement 20 essais toujours en nous basant sur le principe d'une sphère circonscrite:



avec  $\alpha = 1.682$  pour la matrice composite équiradiale et  $\alpha = 1$  pour la matrice face-centrée.

Sinon il y a d'autres techniques que les plans face-centrée et équiradiaux et qui sont:



où le "plan composite (centré)" est un plan à  $k$  facteurs composé de  $N_f$  essais d'un plan factoriel complet ou fractionnaire  $2^{k-r}$ , de  $2k$  essais en étoile sur les axes à une distance  $\alpha$  du centre du domaine et où nous avons  $N_0$  essais au centre du domaine. Le coût d'un plan composite centré est donc de:

$$N = 2^{k-r} + 2k + N_0$$

Les règles suivantes sont d'usage pour les plans composites:

R1. Choix de  $\alpha$  :

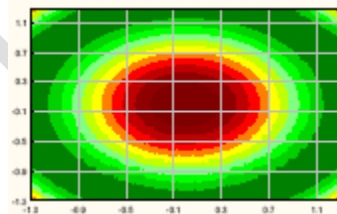
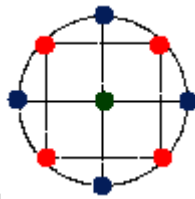
- $\alpha = 1$ : Essais sur les faces du cube et nous nous retrouvons donc avec une matrice face-centrée. C'est un choix commode.
- $\alpha = \sqrt{k}$  : Essais sur une sphère de rayon  $\sqrt{k}$  et nous nous retrouvons donc avec une matrice équiradiale.
- $\alpha = \sqrt[4]{N_f}$  : En prédiction, ce choix assure l'isovariance par rotation (rotatabilité)
- $\alpha = \sqrt{\frac{\sqrt{N_f N} - N_f}{2}}$  : En estimation, assure la (presque)-orthogonalité des estimateurs

R2. Choix de points centraux  $N_0$  :

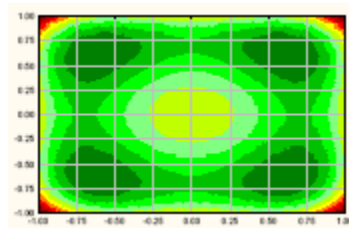
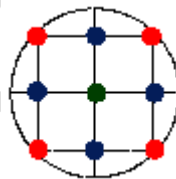
- $N_0 > 1$  : Choisi par l'expérimentateur en fonction de son budget...
- $N_0 = 4\sqrt{N_f} + 4 - 2k$  pour assurer orthogonalité et rotatabilité

Voici un petit résumé:

Rotatif

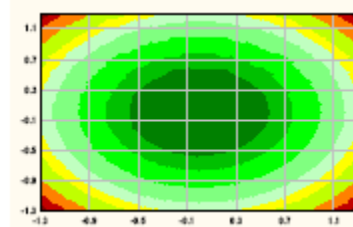
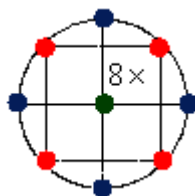


Orthogonal



Rotatif & Orthogonal

8 ×



Et un tableau récapitulatif pour les plans composites pour le choix de  $\alpha, N_0$  et  $r$ :

$k$ (nombre de facteur)	2	3	4	5	5	6	6	7	8
Nombre de termes	6	10	15	21	21	28	28	36	45
$r$	0	0	0	0	1	0	1	1	2
$N_f$	4	8	16	31	16	64	32	64	64
Générateur					I=12356		I=123456	I=1234567	I=12345 I=12678
Nombre d'essais	9	15	25	43	27	77	45	79	81
$\alpha$ Rotatif	1.41	1.68	2	2.38	2	2.83	2.38	2.83	2.83
$\alpha$ Orthogonal (avec $N_0 = 1$ )	1	1.22	1.41	1.6	1.55	1.76	1.72	1.88	2
$\alpha$ Orthogonal (avec $N_0 = 2$ )	1.08	1.29	1.48	1.68	1.61	1.82	1.78	1.94	2.05
$\alpha$ Orthogonal (avec $N_0 = 3$ )	1.15	1.35	1.55	1.72	1.66	1.88	1.84	2	2.11
$N_0$ Rota et orthogonal	8	9	12	17	10	24	15	22	20

Concernant les plans de Doehlert, il s'agit de plans formés de  $N_0$  essais au centre du domaine et de  $k^2 + 2k$  essais répartis "le plus uniformément possible" sur une sphère de rayon 1. Il est souvent "dilaté" pour mieux couvrir le domaine expérimental. Il s'agit d'un plan considéré comme peu coûteux ayant de bonnes propriétés de séquentialité mais avec des propriétés statistiques moins bonnes que pour le plan composite et factoriel.

Le coût de ce type de plan est en toute généralité de:

$$N = k^2 + k + 1$$

Les "plans de Box-Behnken" sont des plans non linéaires pour des facteurs à 3 niveaux. Comme le montre la figure donnée plus haut avec les quatre types de plans courants, les plans de Box-Behnken pour 3 facteurs ont les essais qui sont disposés au milieu des arêtes du cube. Ces plans demandent souvent moins d'essais que les composites généraux.

Pour résumer dans le cas à 3 facteurs à 2 niveaux, nous avons pour les plans les plus courants:

	Composite Centré Équiradial				Composite Face Centrée				Box- Behnken		
Essai	X1	X2	X3	Essai	X1	X2	X3	Essai	X1	X2	X3
1	-1	-1	-1	1	-1	-1	-1	1	-1	-1	0
1	+1	-1	-1	1	+1	-1	-1	1	+1	-1	0
1	-1	+1	-1	1	-1	+1	-1	1	-1	+1	0
1	+1	+1	-1	1	+1	+1	-1	1	+1	+1	0
1	-1	-1	+1	1	-1	-1	+1	1	-1	0	-1
1	+1	-1	+1	1	+1	-1	+1	1	+1	0	-1
1	-1	+1	+1	1	-1	+1	+1	1	-1	0	+1
1	+1	+1	+1	1	+1	+1	+1	1	+1	0	+1
1	-1.682	0	0	1	-1	0	0	1	0	-1	-1
1	+1.682	0	0	1	+1	0	0	1	0	+1	-1
1	0	-1.682	0	1	0	-1	0	1	0	-1	+1
1	0	+1.682	0	1	0	+1	0	1	0	+1	+1
1	0	0	-1.682	1	0	0	-1	3	0	0	0
1	0	0	+1.682	1	0	0	+1				
6	0	0	0	6	0	0	0				
<b>Essais = 20</b>				<b>Essais = 20</b>				<b>Essais = 15</b>			

Cette petite introduction théorique étant faite, passons à l'exemple pratique.

Une nouvelle molécule est développée pour un processus de dépôt de surface. Un comité d'expert a identifié trois facteurs contrôlables d'intérêt:

Facteur	Étendue
---------	---------

Additif	0 à 70 grammes
Temps de réaction	20 à 60 minutes
Température de réaction	100 à 180 degrés celsius

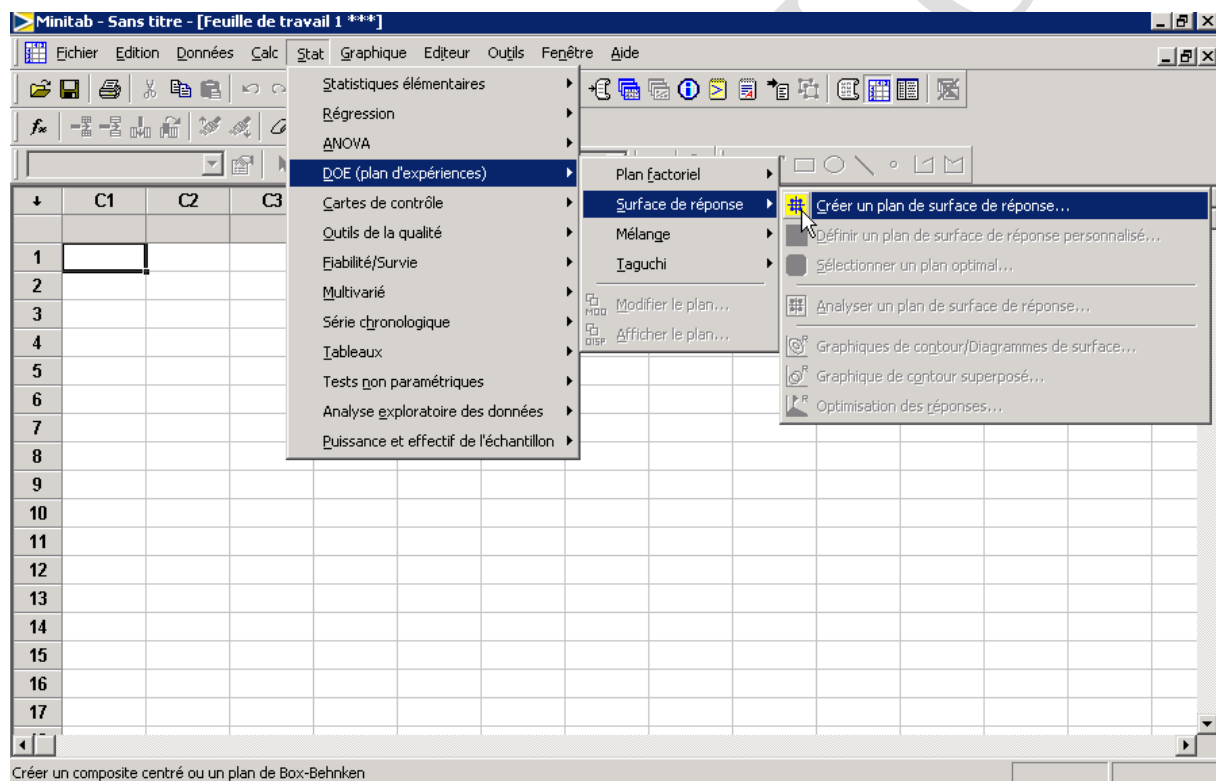
Le taux de la nouvelle molécule est mesuré à la fin de l'expérience et le niveau d'adhésion de la surface est ensuite mesurée. Les exigences sont d'avoir:

$$\% \text{ taux de recouvrement} \geq 91\%$$

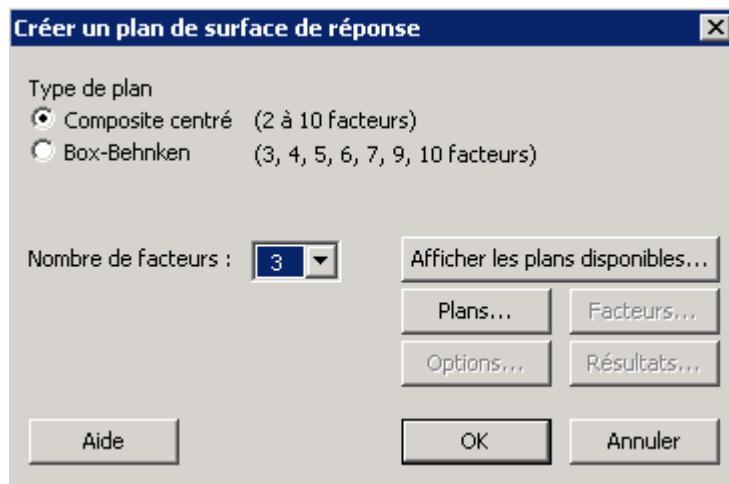
$$\text{adhésion} \geq 45 \left[ \text{g} \cdot \text{cm}^{-2} \right]$$

Nous voulons déterminer les paramètres permettant de satisfaire ces conditions sachant que nous avons les moyens que pour 24 essais et que nous savons que le phénomène est fortement non linéaire.

Nous allons dans Minitab® Statistical Software R15 dans le menu **Stat/DOE (plan d'expériences/Surface de réponse/Créer un plan de surface de réponse...**



Nous prenons un plan composite centré (**car nous avec des facteurs à 2 niveaux et les plans de Box-Behnken sont spécialisés dans les plans à facteurs ayant 3 niveaux**) avec 3 facteurs:



Nous voyons en cliquant sur **Afficher les plans disponibles...**:

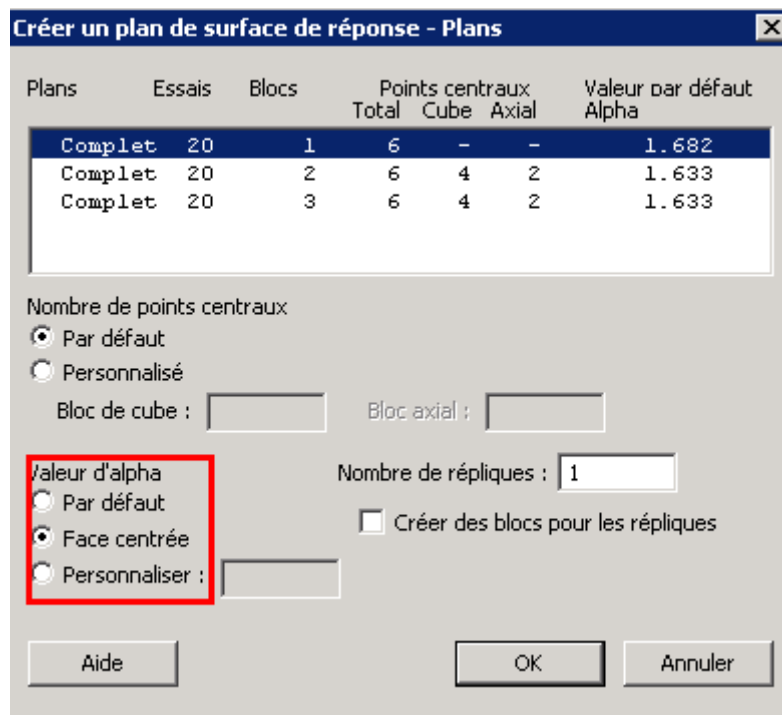
Créer un plan de surface de réponse - Afficher les plans disponibles

Plans de surface de réponse disponibles (avec nombre d'essais)

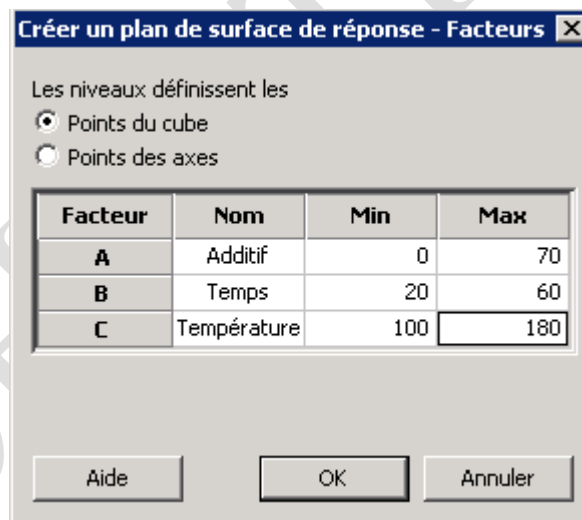
Plan		Facteurs								
		2	3	4	5	6	7	8	9	10
Composite centré (complet)	non divisé en blocs	13	20	31	52	90	152			
	divisé en blocs	14	20	30	54	90	160			
Composite centré (1/2)	non divisé en blocs				32	53	88	154		
	divisé en blocs				33	54	90	160		
Composite centré (1/4)	non divisé en blocs							90	156	
	divisé en blocs							90	160	
Composite centré (1/8)	non divisé en blocs									158
	divisé en blocs									160
Box-Behnken	non divisé en blocs		15	27	46	54	62		130	170
	divisé en blocs			27	46	54	62		130	170

qu'avec 24 essais nous sommes bien partis pour faire un plan de surface. Nous cliquons ensuite sur le bouton **Plans...** de la boîte de dialogue principale:



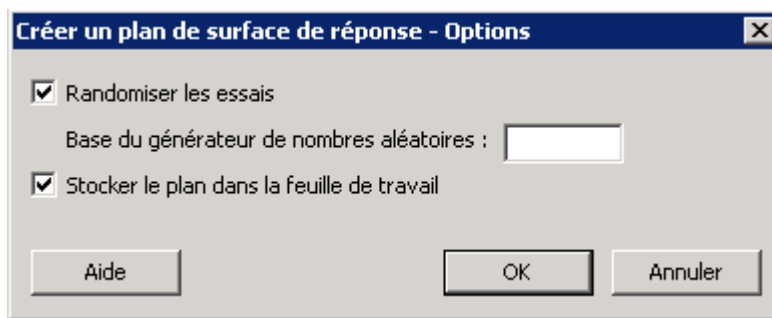


et nous prenons un **Face centrée** avec un nombre de points centraux par défaut (donc 6 dans le cas présent). Ensuite, nous cliquons sur le bouton **Facteurs...** de la boîte de dialogue principale:

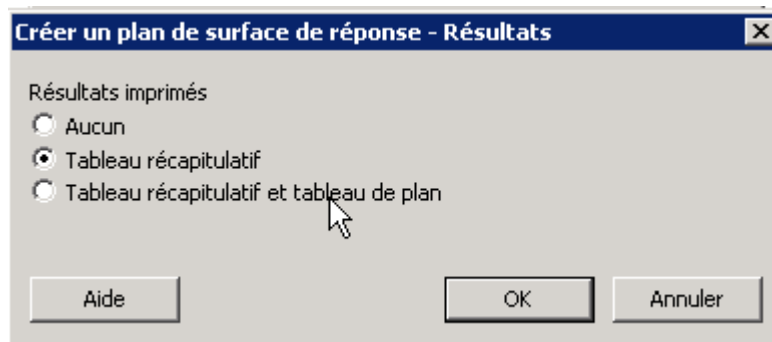


**Avec les plans de surface nous pouvons travailler en unités non codées car lors de l'analyse, les options nous permettent de choisir si nous voulons transformer les données pour que l'ANOVA fonctionne tout de même... (voir plus loin) et ce contrairement aux autres plans (du moins avec la version 15 du logiciel).**

Dans les **Options...** nous allons laisser randomisation des essais:



Dans le bouton **Résultats...** nous ne changeons rien:



et nous validons le tout par **OK**. Nous avons alors dans la fenêtre de session:

### Plan composite centré

Facteurs :	3	Répliques :	1
Essais de base :	20	Nombre total d'essais :	20
Blocs de base :	1	Nombre total de blocs :	1

Factoriel à deux niveaux : factoriel complet

Points du cube :	8
Points centraux en cube :	6
Points des axes :	6
Points centraux dans l'axe :	0

Alpha : 1

et dans la feuille de données (**nous voyons que le point central (35,40,140) s'y répète bien 6 fois**):

Minitab - PlanCompositeCentre.MPJ - [Feuille de travail 2 \*\*\*]

Fichier Edition Données Calc Stat Graphique Éditeur Outils Fenêtre Aide

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	OrdreStd	OrdEssai	TypePt	Blocs	Additif	Temps	Température						
1	17	1	0	1	35	40	140						
2	10	2	-1	1	70	40	140						
3	13	3	-1	1	35	40	100						
4	15	4	0	1	35	40	140						
5	8	5	1	1	70	60	180						
6	7	6	1	1	0	60	180						
7	9	7	-1	1	0	40	140						
8	16	8	0	1	35	40	140						
9	6	9	1	1	70	20	180						
10	5	10	1	1	0	20	180						
11	11	11	-1	1	35	20	140						
12	19	12	0	1	35	40	140						
13	14	13	-1	1	35	40	180						
14	4	14	1	1	70	60	100						
15	12	15	-1	1	35	60	140						
16	3	16	1	1	0	60	100						
17	18	17	0	1	35	40	140						
18	1	18	1	1	0	20	100						
19	2	19	1	1	70	20	100						
20	20	20	0	1	35	40	140						

Feuille de travail en cours : Feuille de travail 2

et nous saisissons les données de l'expérience en faisant bien attention à ne pas nous tromper de ligne:

Minitab - PlanCompositeCentre.MPJ - [Feuille de travail 2 \*\*\*]

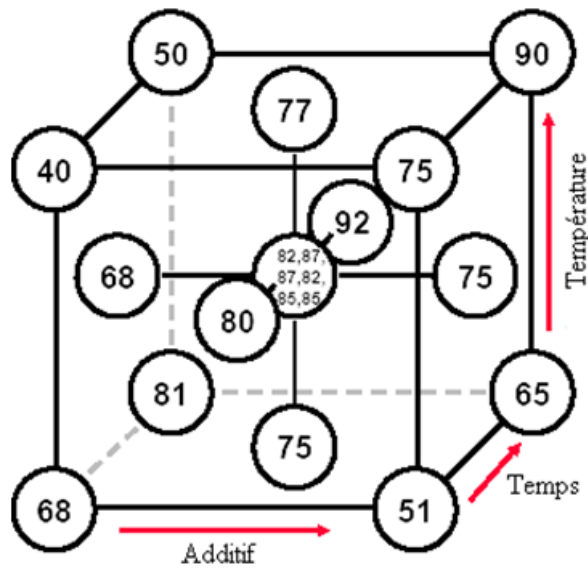
Fichier Edition Données Calc Stat Graphique Éditeur Outils Fenêtre Aide

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
	OrdreStd	OrdEssai	TypePt	Blocs	Additif	Temps	Température		Taux	Adhesion			
1	+	1	0	1	35	40	140		82	37			
2	10	2	-1	1	70	40	140		75	44			
3	13	3	-1	1	35	40	100		75	31			
4	15	4	0	1	35	40	140		87	41			
5	8	5	1	1	70	60	180		90	39			
6	7	6	1	1	0	60	180		50	40			
7	9	7	-1	1	0	40	140		68	24			
8	16	8	0	1	35	40	140		87	41			
9	6	9	1	1	70	20	180		75	31			
10	5	10	1	1	0	20	180		40	37			
11	11	11	-1	1	35	20	140		80	38			
12	19	12	0	1	35	40	140		82	37			
13	14	13	-1	1	35	40	180		77	44			
14	4	14	1	1	70	60	100		65	48			
15	12	15	-1	1	35	60	140		92	41			
16	3	16	1	1	0	60	100		81	10			
17	18	17	0	1	35	40	140		82	37			
18	1	18	1	1	0	20	100		68	3			
19	2	19	1	1	70	20	100		51	40			
20	20	20	0	1	35	40	140		82	37			
21													

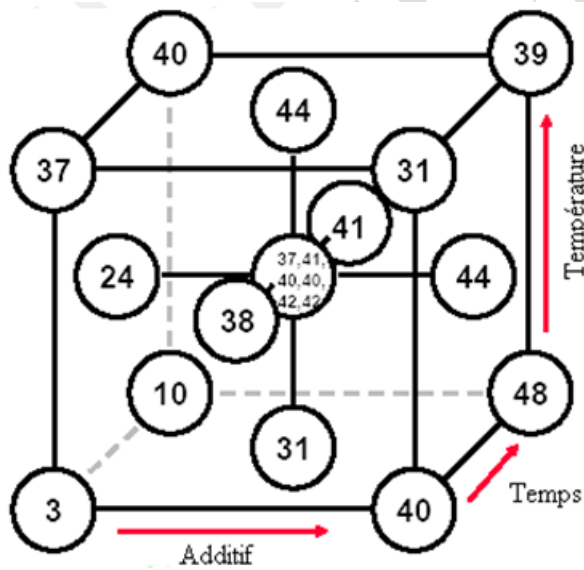
Feuille de travail en cours : Feuille de travail 2

**Nous voyons déjà qu'au niveau des objectifs à atteindre ça va ne pas être simple... puisque aucun point expérimental ne satisfait aux objectifs de l'énoncé jusqu'à maintenant.**

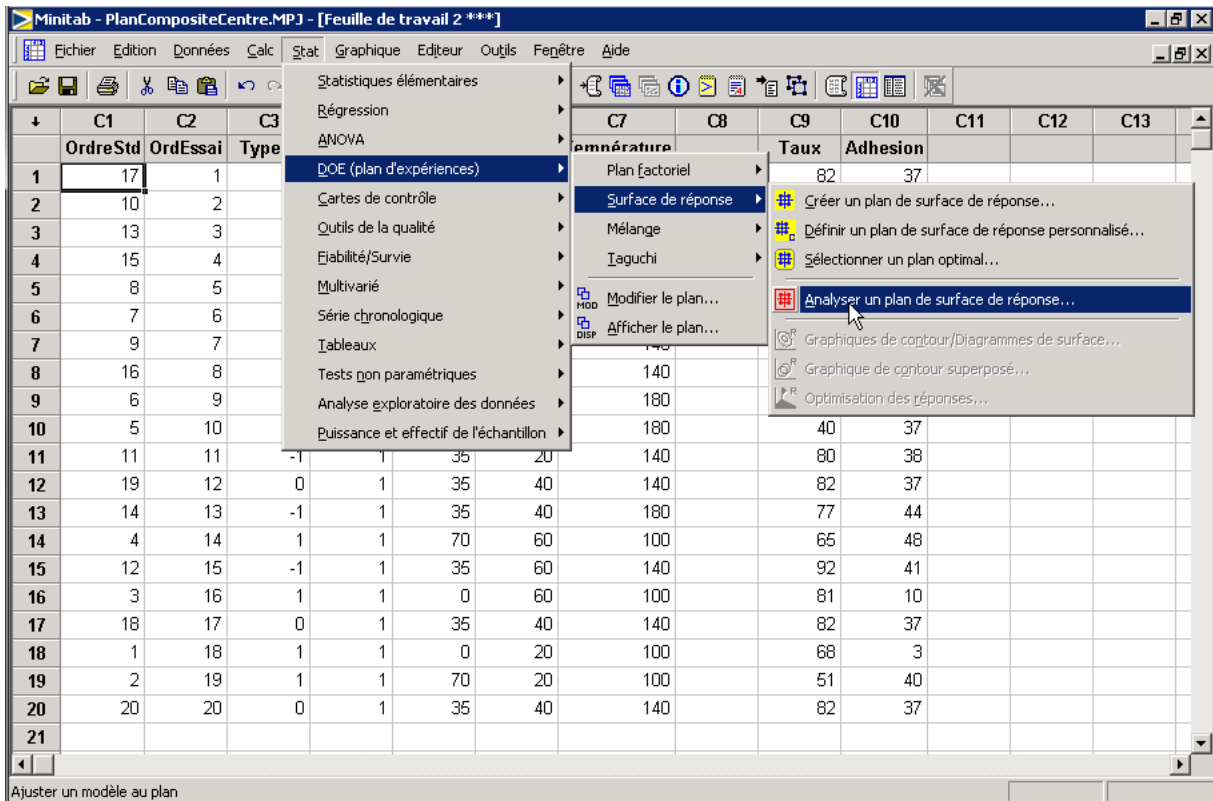
Donc ce que nous avons saisi ci-dessus correspond pour le facteur *Taux* à:



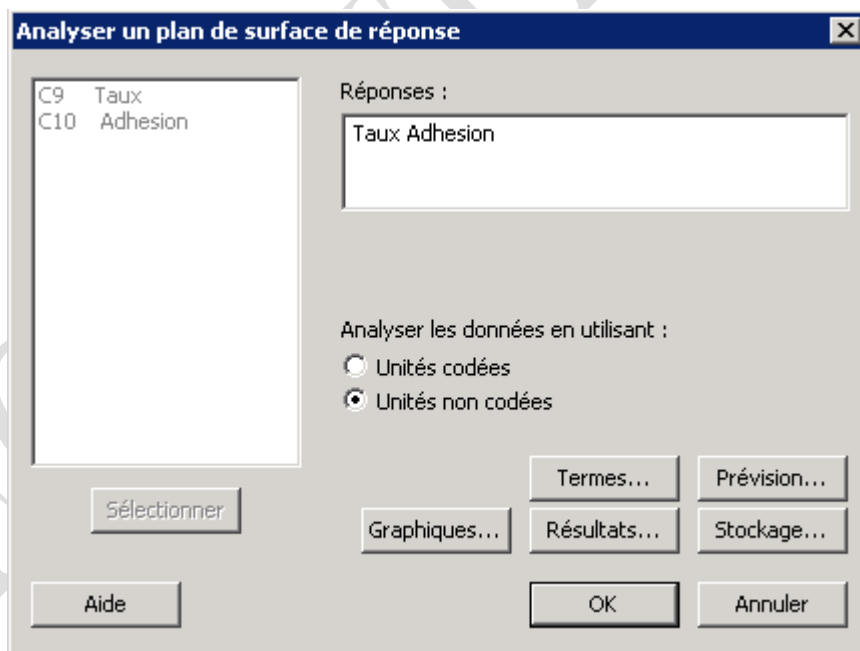
et pour le facteur *Adhésion* à:



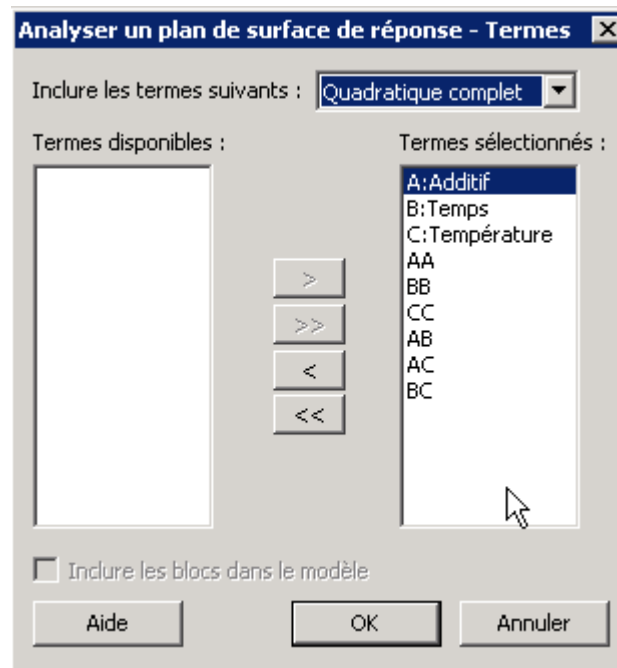
Ensuite, nous allons analyser le plan en allant dans le menu **Stat/DOE (plan d'expériences/Surface de réponse/Analyser un plan de surface de réponse...**



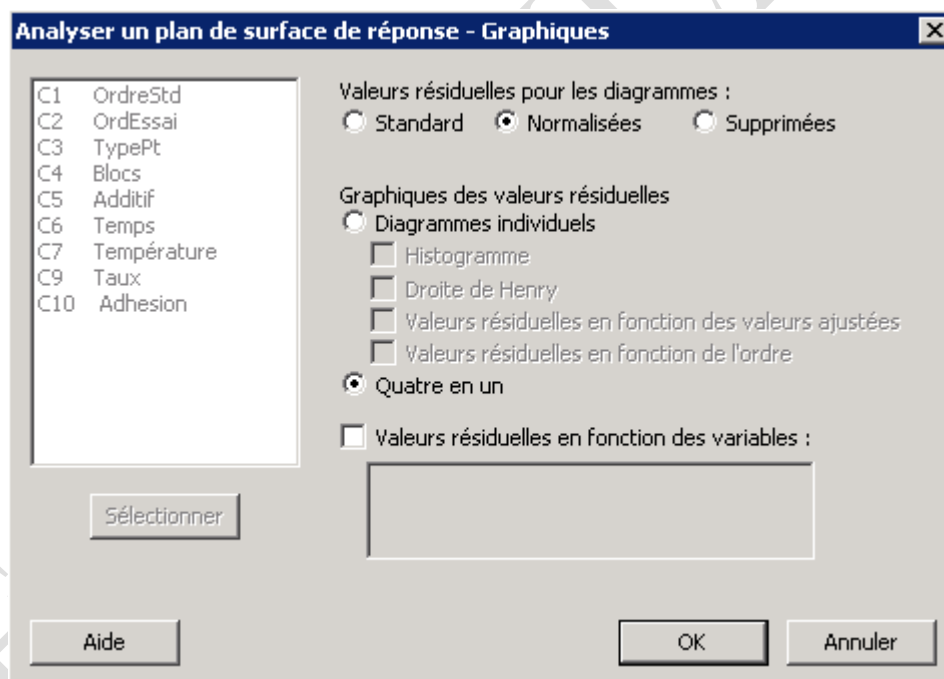
où nous prenons:



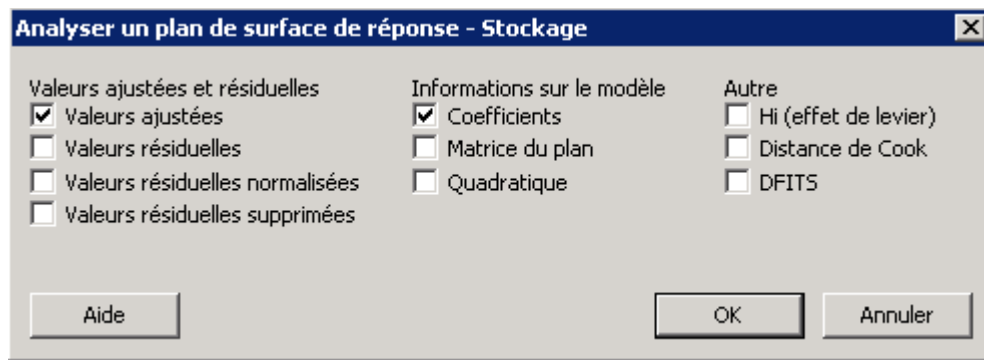
Pour le bouton **Termes...**, nous prenons les coefficients rectangles et les quadratiques (donc tout!):



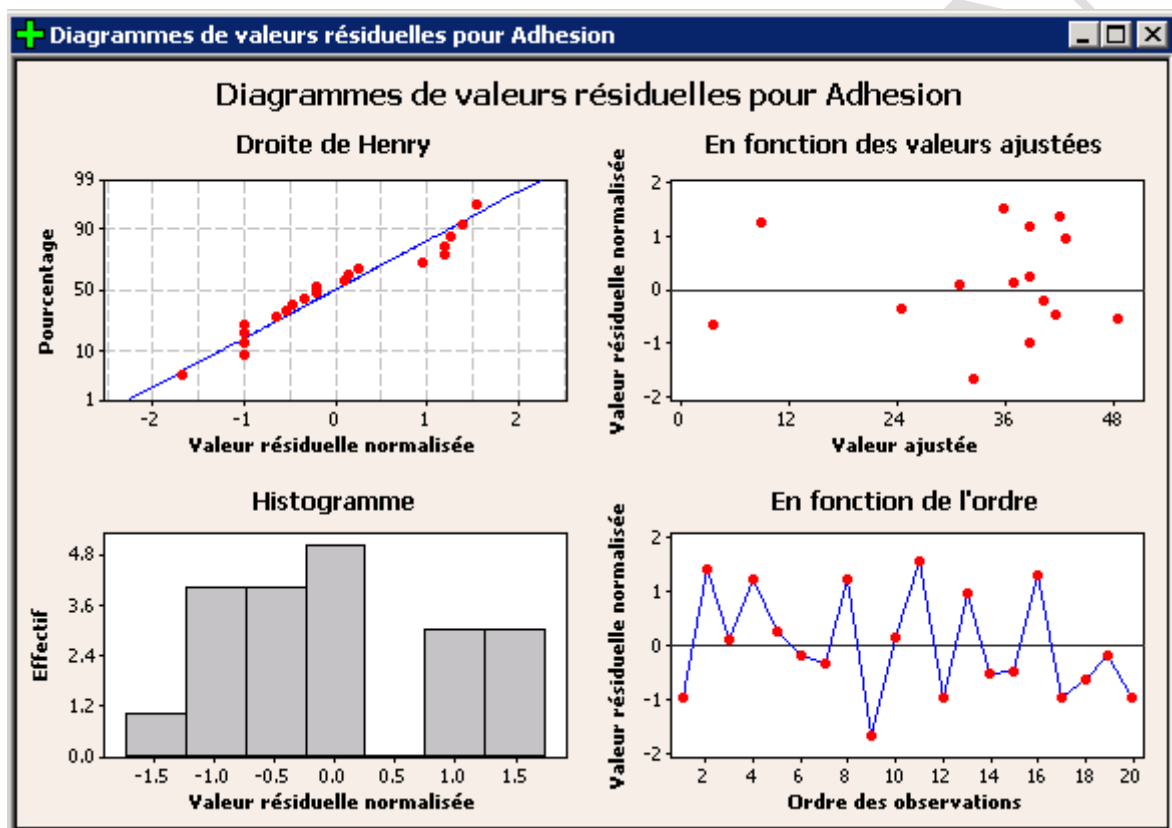
Dans le bouton **Graphiques...** nous prenons:

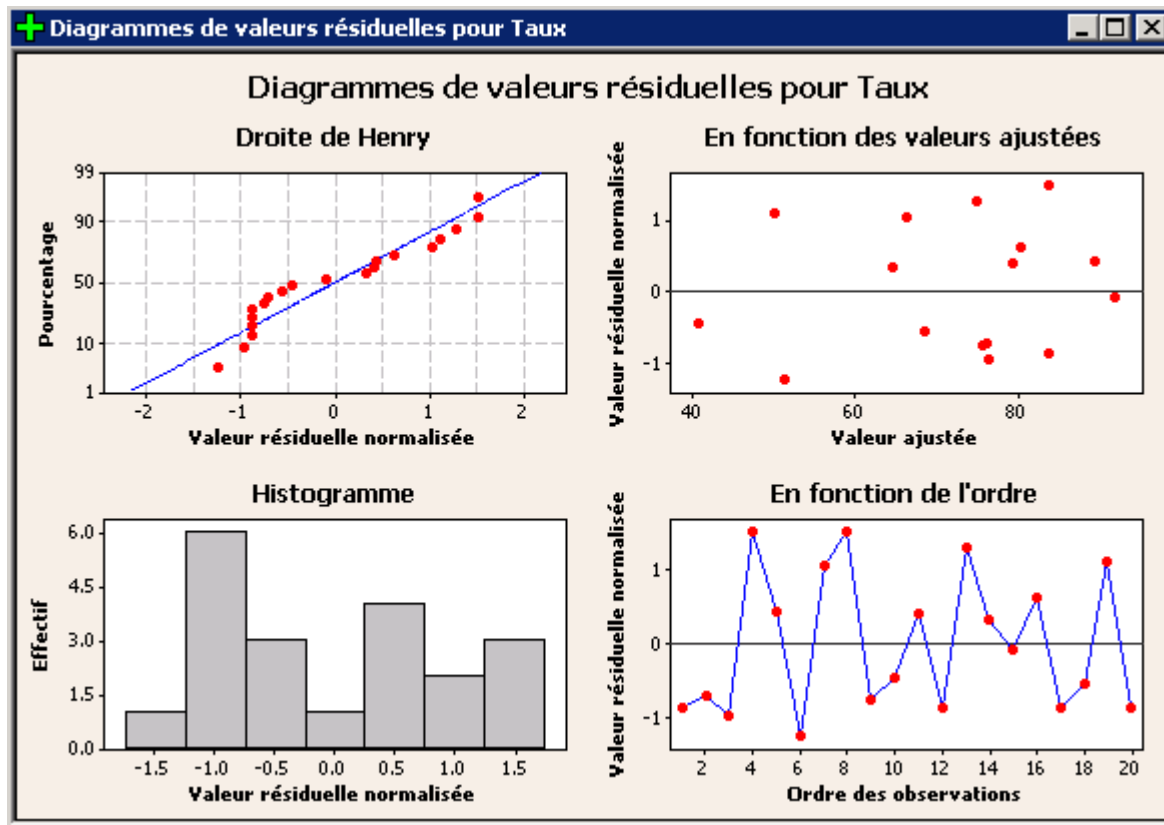


Dans le bouton **Stockage...** nous prenons comme à l'habitude les informations que je considère personnellement comme les plus intéressantes pédagogiquement parlant:



Nous validons le tout par **OK**. Nous obtenons les graphiques suivants:





et dans la fenêtre de session, nous avons:

**Régression de la surface de réponse : Taux en fonction de Additif; Temps; ...**

L'analyse a été effectuée à l'aide de données non codées.

Coefficients de régression estimés pour Taux

Terme	Coeff	Coef ErT	T	P
Constante	11.8682	15.9208	0.745	0.473
Additif	-0.5334	0.1202	-4.439	0.001
Temps	-0.0556	0.3056	-0.182	0.859
Température	1.0709	0.2391	4.479	0.001
Additif*Additif	-0.0103	0.0011	-9.401	0.000
Temps*Temps	0.0048	0.0033	1.425	0.184
Température*Température	-0.0051	0.0008	-6.041	0.000
Additif*Temps	0.0011	0.0011	0.955	0.362
Additif*Température	0.0096	0.0006	17.192	0.000
Temps*Température	-0.0003	0.0010	-0.318	0.757

S = 2.22097      SomCar-ErrPrév = 196.121  
 R carré = 98.66 %    R carré (prév) = 94.67 %    R carré (ajust) = 97.45 %



Source	DL	SomCar	ség	ajust	CM	ajust	F	p
Régression	9	2298.42	2298.42	1087.12	362.373	96.42	0.000	0.000
Linéaire	3	1206.60	244.82	244.82	81.606	21.71	0.000	0.000
Carre	3	244.82	244.82	244.82	81.606	21.71	0.000	0.000
Interaction	3	847.00	847.00	847.00	282.333	75.12	0.000	0.000
Erreur résiduelle	10	37.58	37.58	37.58	3.758			
Inadéquation de l'ajustement	5	16.25	16.25	16.25	3.250	0.76	0.614	
Erreur pure	5	21.33	21.33	21.33	4.267			
Total	19	2336.00						

Analyse de la variance pour Adhesion

S = 1.93860 SomCar-ErrPrév = 167.401 R carré = 98.39 % R carré (prév) = 92.83 % R carré (ajust) = 96.94 %

Term	Coef	Coef ERT	T	P
Constante	-65.0909	13.8967	-4.684	0.001
Additif	1.5505	0.1049	14.782	0.000
Temps	0.2041	0.2668	0.765	0.462
Température	0.7867	0.2087	3.770	0.004
Additif*Additif	-0.0045	0.0010	-4.744	0.001
Temps*Temps	-0.0001	0.0029	-0.039	0.970
Température*Température	-0.0013	0.0007	-1.750	0.111
Additif*Temps	0.0011	0.0010	1.094	0.300
Additif*Température	-0.0073	0.0005	-14.955	0.000
Temps*Température	-0.0006	0.0009	-0.730	0.482

Coefficients de régression estimés pour Adhesion

L'analyse a été effectuée à l'aide de données non codées.

**Régression de la surface de réponse : Adhesion en fonction de Additif; Temps;**

Source	DL	SomCar	ség	ajust	CM	ajust	F	p
Régression	9	3631.62	3631.62	403.514	81.80			
Linéaire	3	656.10	1512.52	1512.52	504.174	102.21		
Carre	3	1512.52	1512.52	1512.52	504.174	102.21		
Interaction	3	1463.00	1463.00	1463.00	487.667	98.86		
Erreur résiduelle	10	49.33	49.33	49.33	4.933			
Inadéquation de l'ajustement	5	15.99	15.99	15.99	3.199	0.48		
Erreur pure	5	33.33	33.33	33.33	6.667			
Total	19	3680.95						

Analyse de la variance pour Taux

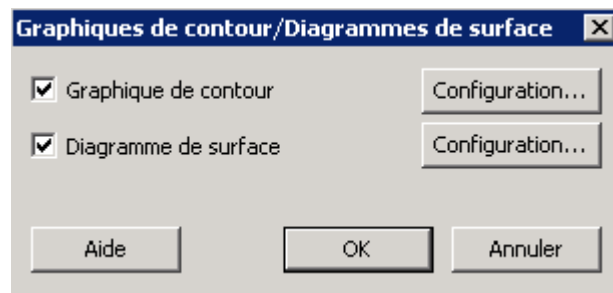
et dans la feuille:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
	OrdreStd	OrdEssai	TypePt	Blocs	Additif	Temps	Température		Taux	Adhesion	AJUSTEES1	AJUSTEES2	COEFF1	COEFF2	
1	17	1	0	1	35	40	140		82	37	83.8364	38.8182	11.8682	-65.0909	
2	10	2	-1	1	70	40	140		75	44	76.1455	42.0727	-0.5334	1.5505	
3	13	3	-1	1	35	40	100		75	31	76.5455	30.8727	-0.0556	0.2041	
4	15	4	0	1	35	40	140		87	41	83.8364	38.8182	1.0709	0.7867	
5	8	5	1	1	70	60	180		90	39	89.5636	38.7818	-0.0103	-0.0045	
6	7	6	1	1	0	60	180		50	40	51.2636	40.1818	0.0048	-0.0001	
7	9	7	-1	1	0	40	140		68	24	66.3455	24.4727	-0.0051	-0.0013	
8	16	8	0	1	35	40	140		87	41	83.8364	38.8182	0.0011	0.0011	
9	6	9	1	1	70	20	180		75	31	75.7636	32.4818	0.0096	-0.0073	
10	5	10	1	1	0	20	180		40	37	40.4636	36.8818	-0.0003	-0.0006	
11	11	11	-1	1	35	20	140		80	38	79.3455	35.8727			
12	19	12	0	1	35	40	140		82	37	83.8364	38.8182			
13	14	13	-1	1	35	40	180		77	44	74.9455	42.6727			
14	4	14	1	1	70	60	100		65	48	64.6636	48.4818			
15	12	15	-1	1	35	60	140		92	41	92.1455	41.6727			
16	3	16	1	1	0	60	100		81	10	80.3636	8.8818			
17	18	17	0	1	35	40	140		82	37	83.8364	38.8182			
18	1	18	1	1	0	20	100		68	3	68.5636	3.5818			
19	2	19	1	1	70	20	100		51	40	49.8636	40.1818			
20	20	20	0	1	35	40	140		82	37	83.8364	38.8182			
21															

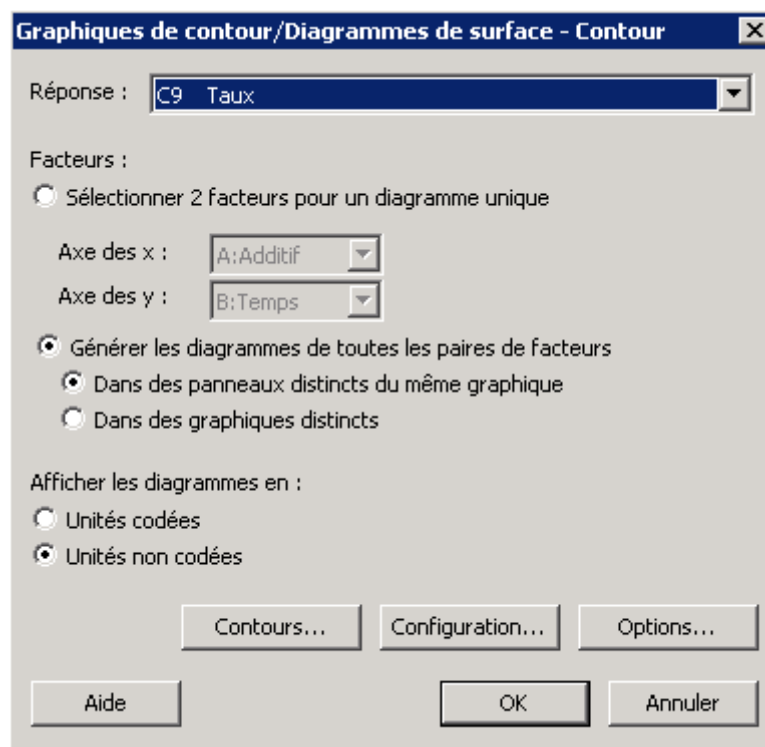
Bon nous avons donc notre modèle théorique avec la connaissance si les coefficients sont significatifs ou non. Mais nous cherchons cependant à savoir si nous pouvons arriver à nos objectifs.

Pour cela, nous allons dans le menu **Stat/DOE (plan d'expériences/Surface de réponse/Graphique de contour/Diagrammes de surface...:**

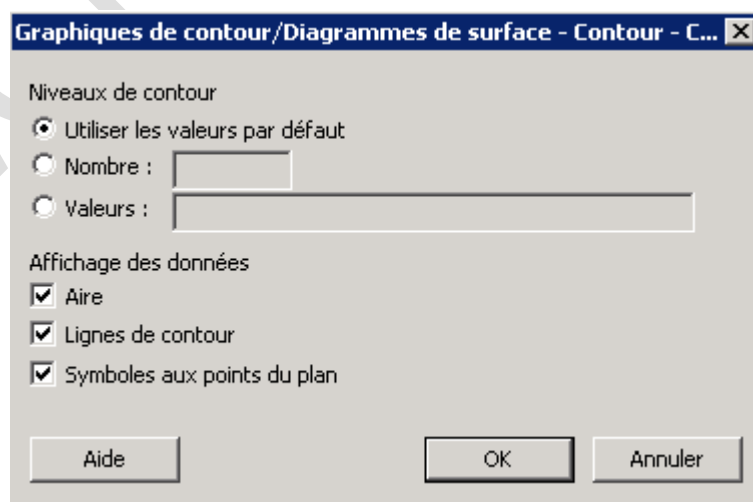
Nous prenons les deux graphiques:



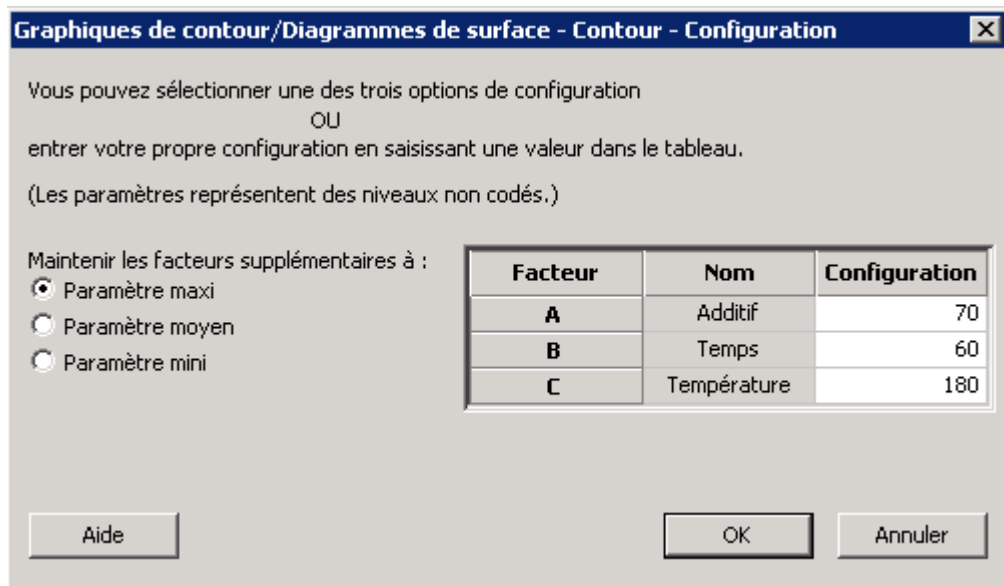
avec pour configuration du premier (c'est la partie la plus longue car il faut le fait pour chacune des réponses):



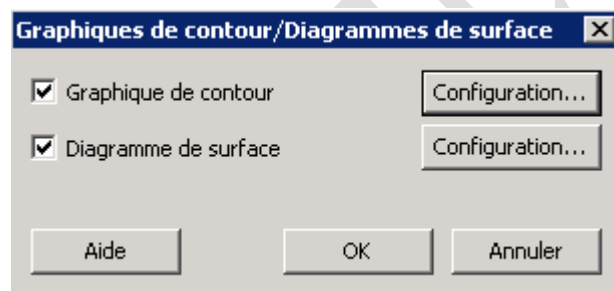
Nous cliquons sur **Contours...**:



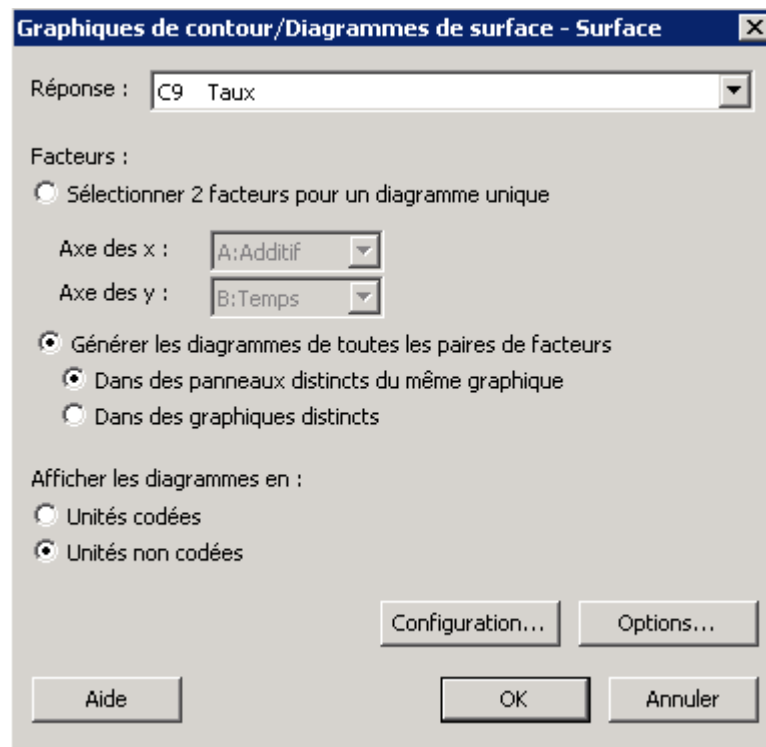
Nous cliquons sur **Configuration...** et au vu des résultats obtenus il paraît clair que nous allons plutôt prendre les valeurs maximales des facteurs:



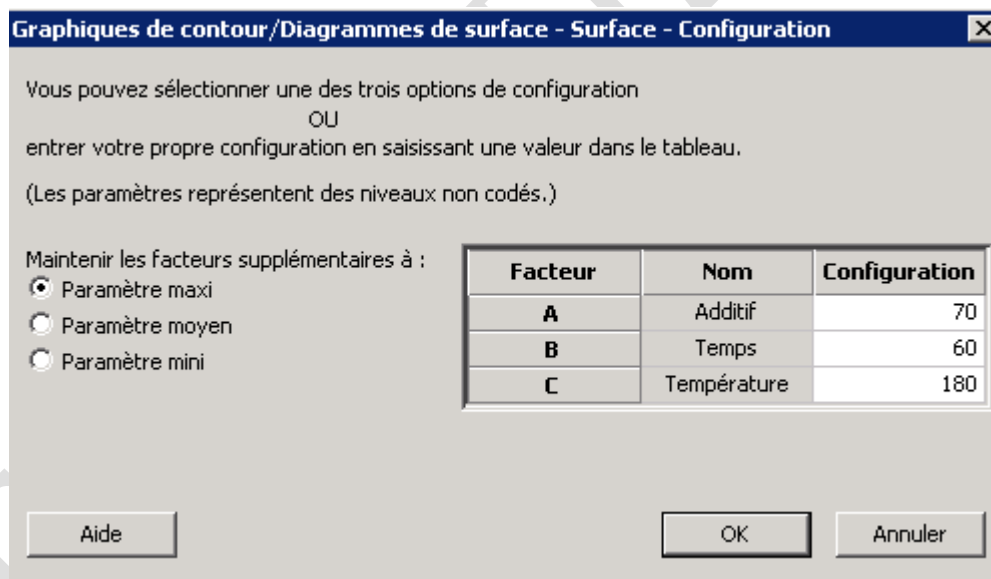
et nous validons par **OK**. Nous configurons ensuite le diagramme de surface:



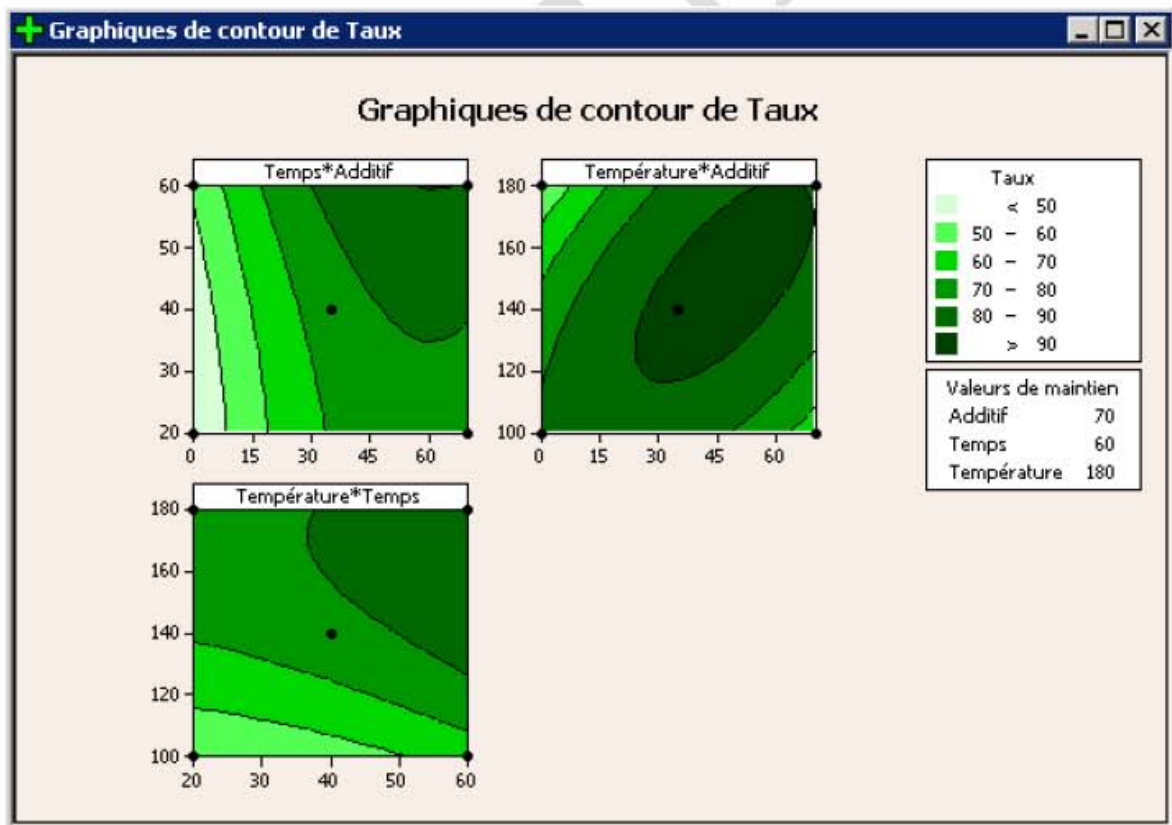
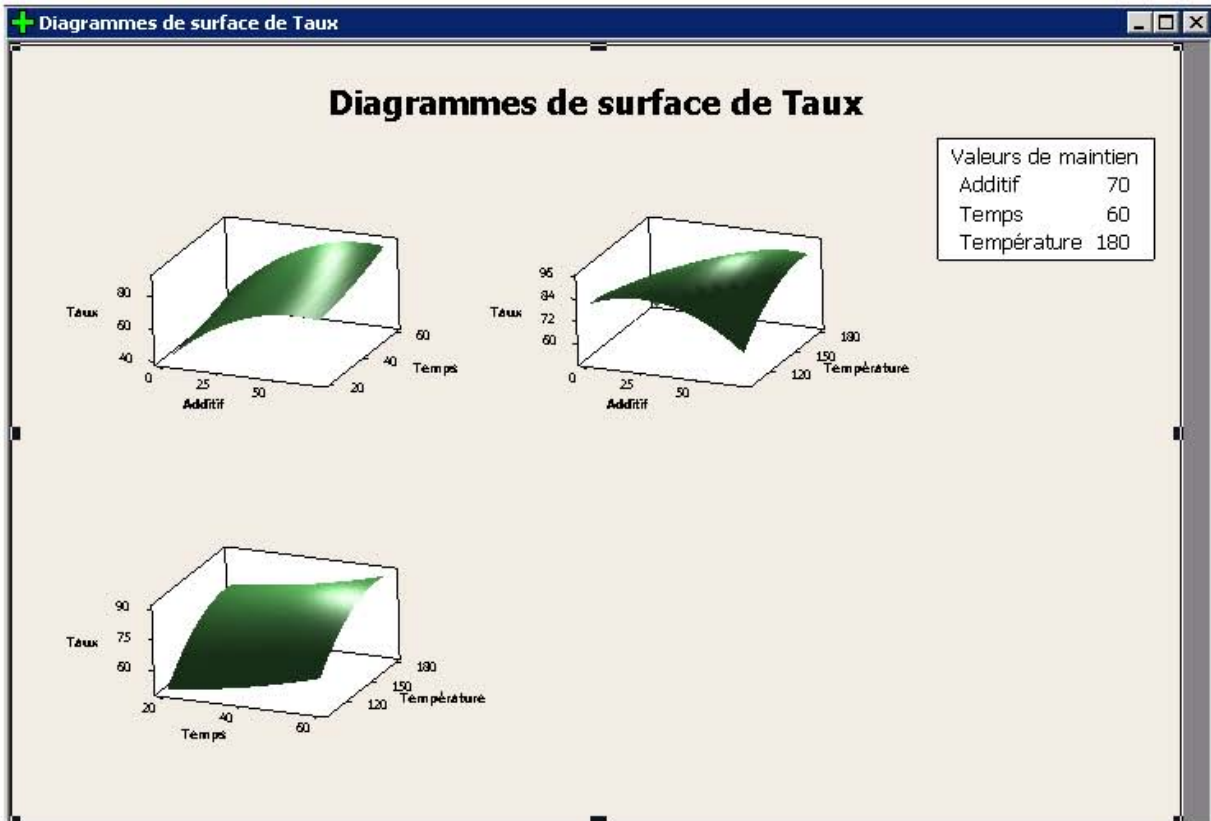
Ce qui nous amène dans:



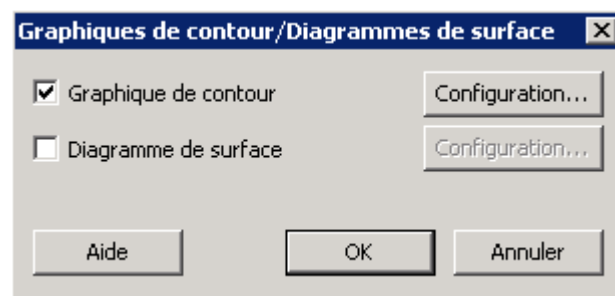
et dans configuration, nous prenons aussi:



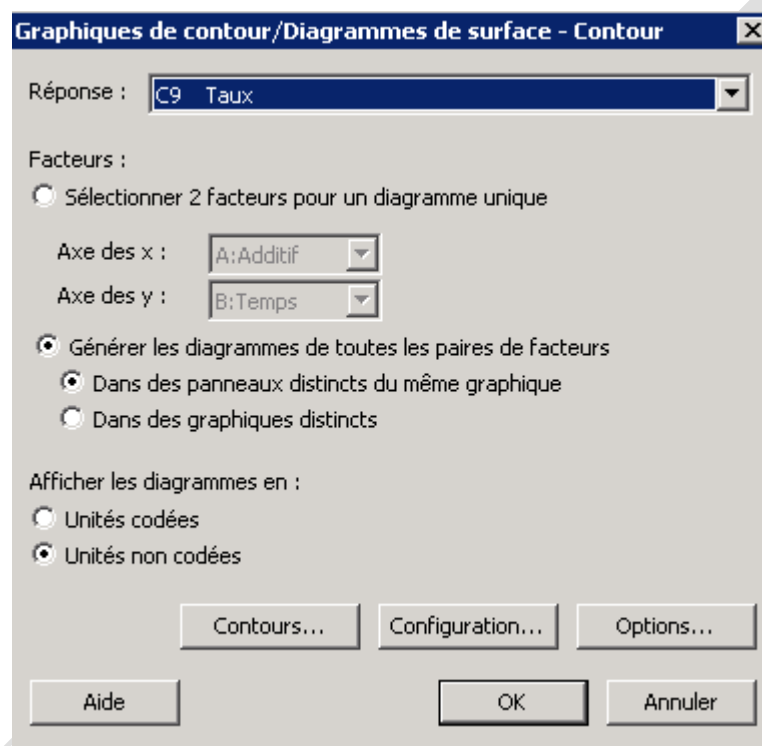
Nous validons le tout par **OK** pour avoir:



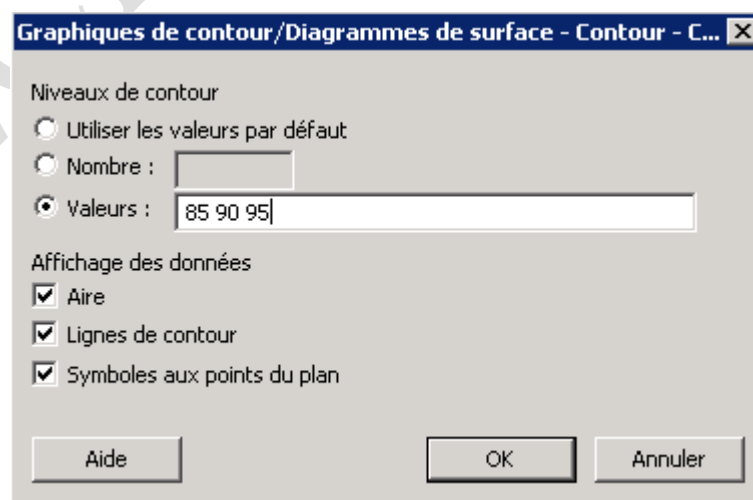
Revenons dans l'analyse graphique en prenant cette fois-ci que les surfaces de contour:



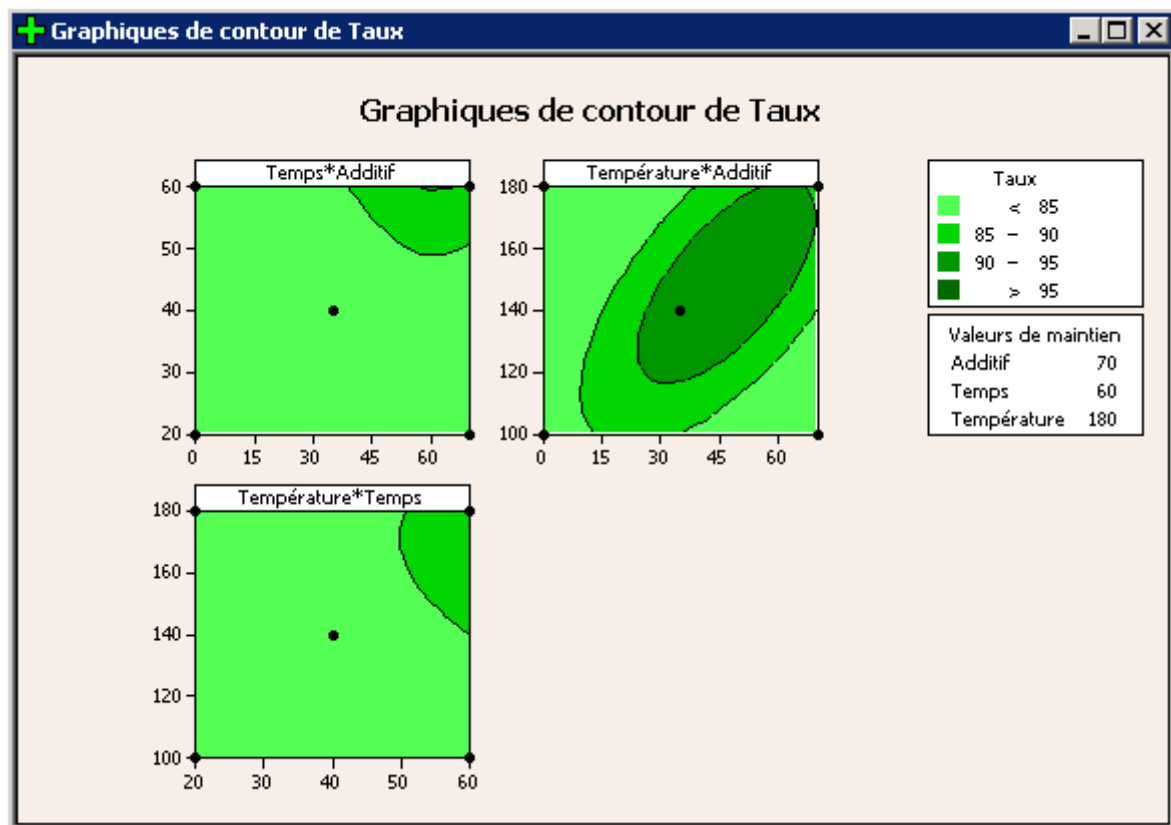
et allons dans la **Configuration...**:



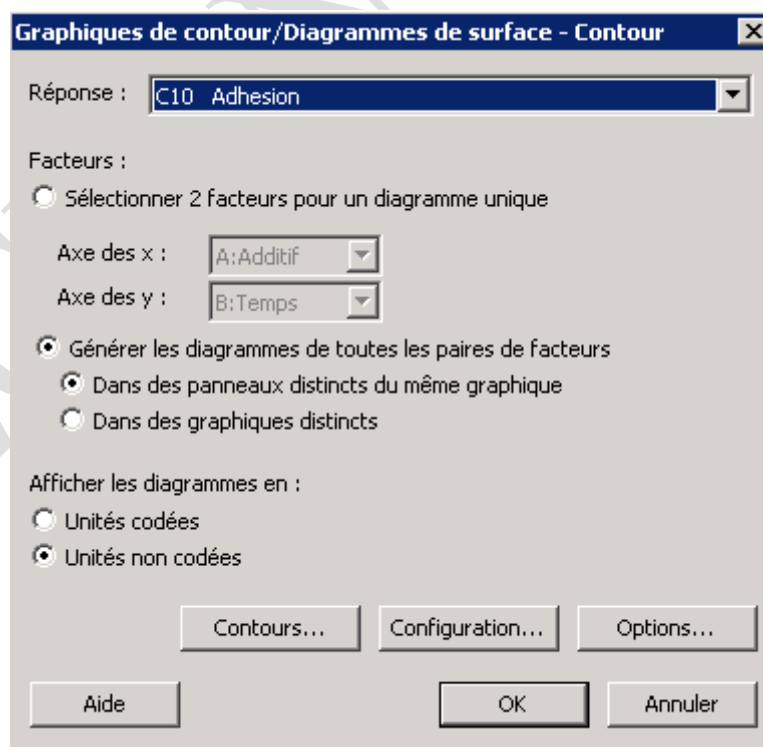
et allons dans **Contours...** pour indiquer trois isoclines qui nous intéressent:



si nous validons par **OK**, nous avons:

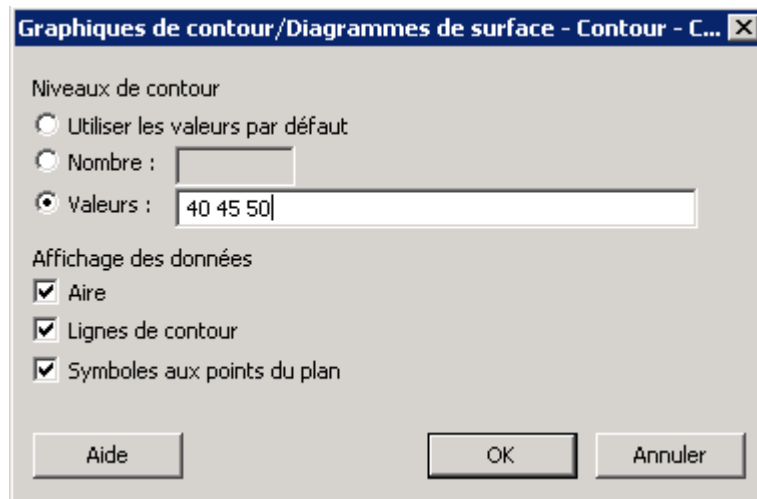


Si nous faisons la même chose pour l'adhésion:

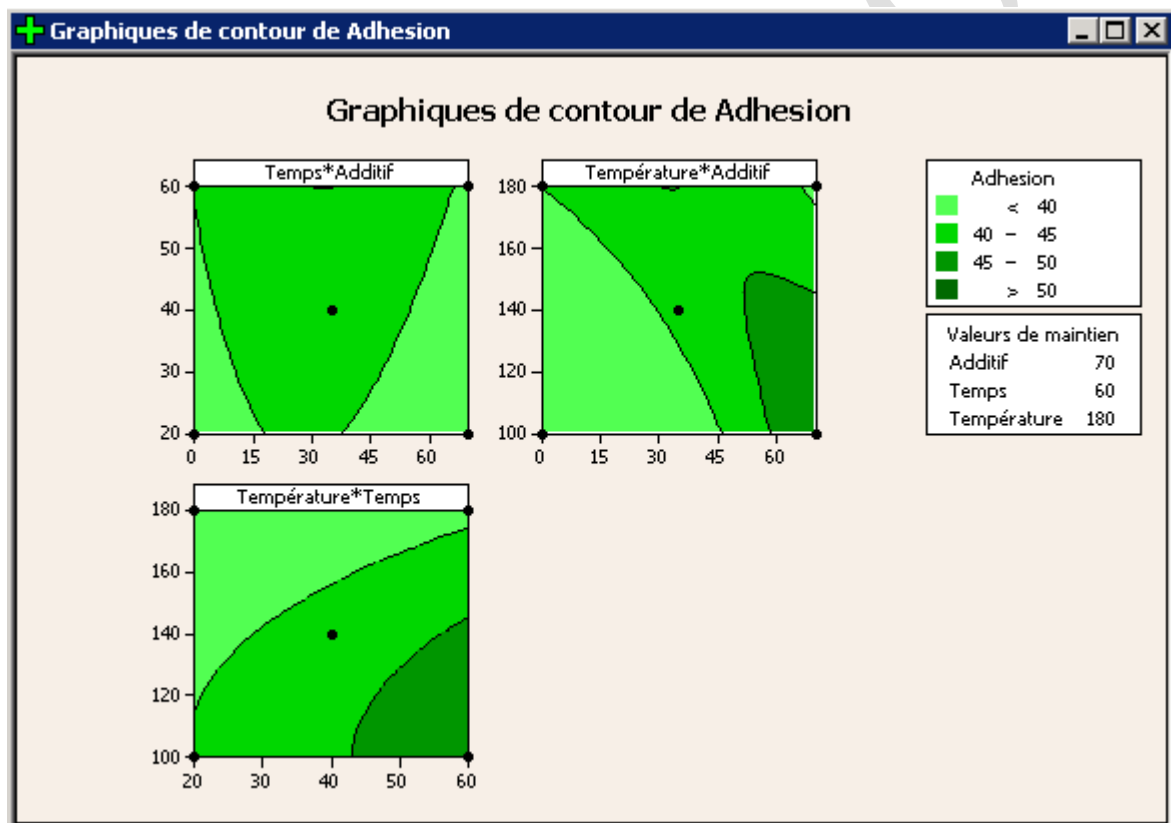


avec les isoclines:

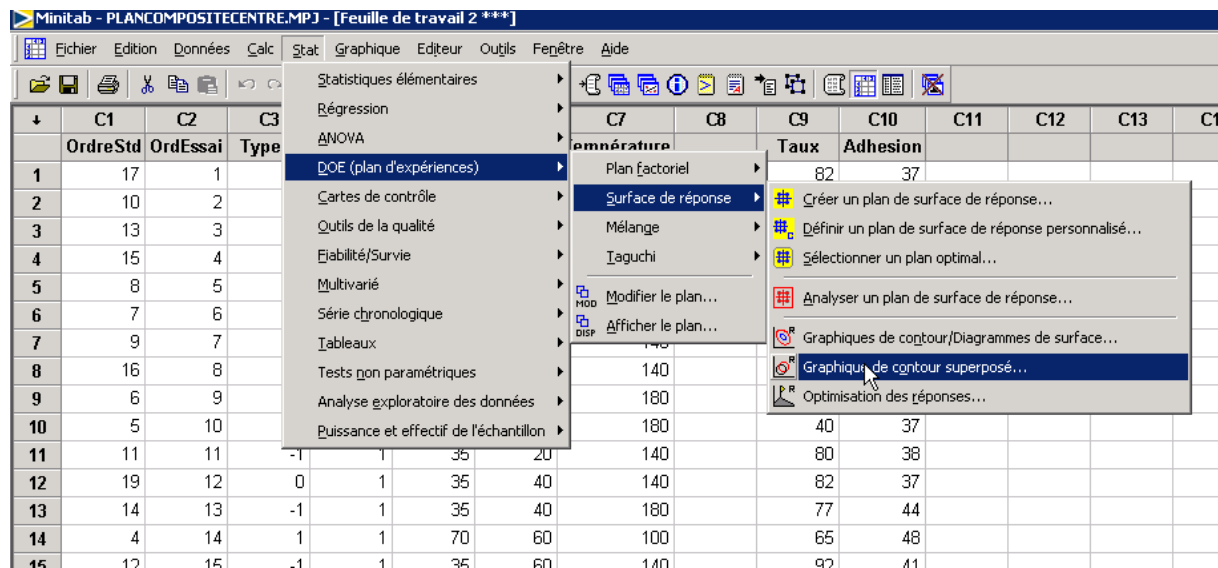




Ce qui donne:



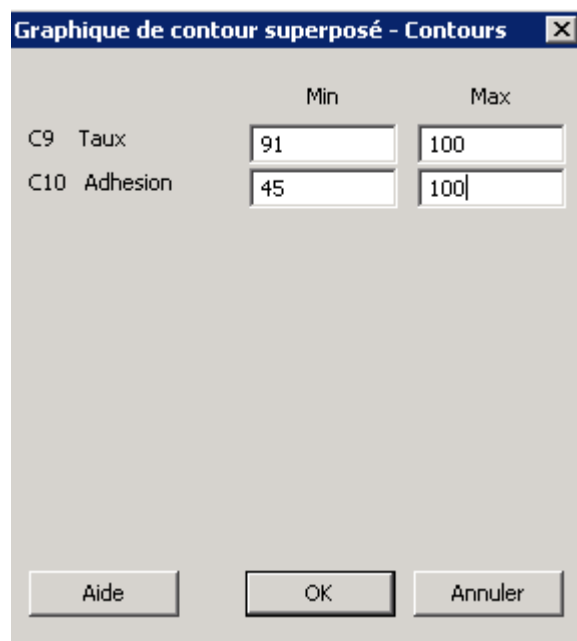
Il est cependant un peu difficile de trouver la zone d'optimum en basculant d'un graph à l'autre, dès lors, nous pouvons aller dans le menu **Stat/DOE (plan d'expériences/Surface de réponse/Graphique de contour superposé...**



Nous prenons:

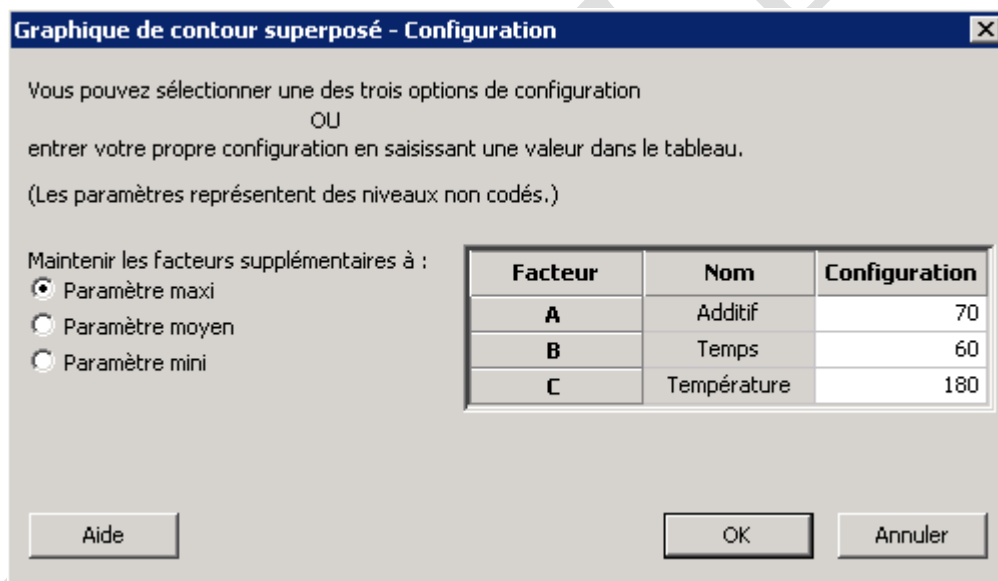


et dans le bouton **Contours...** nous prenons:



	Min	Max
C9 Taux	91	100
C10 Adhesion	45	100

et dans **Configuration...** au vu des graphiques précédents:



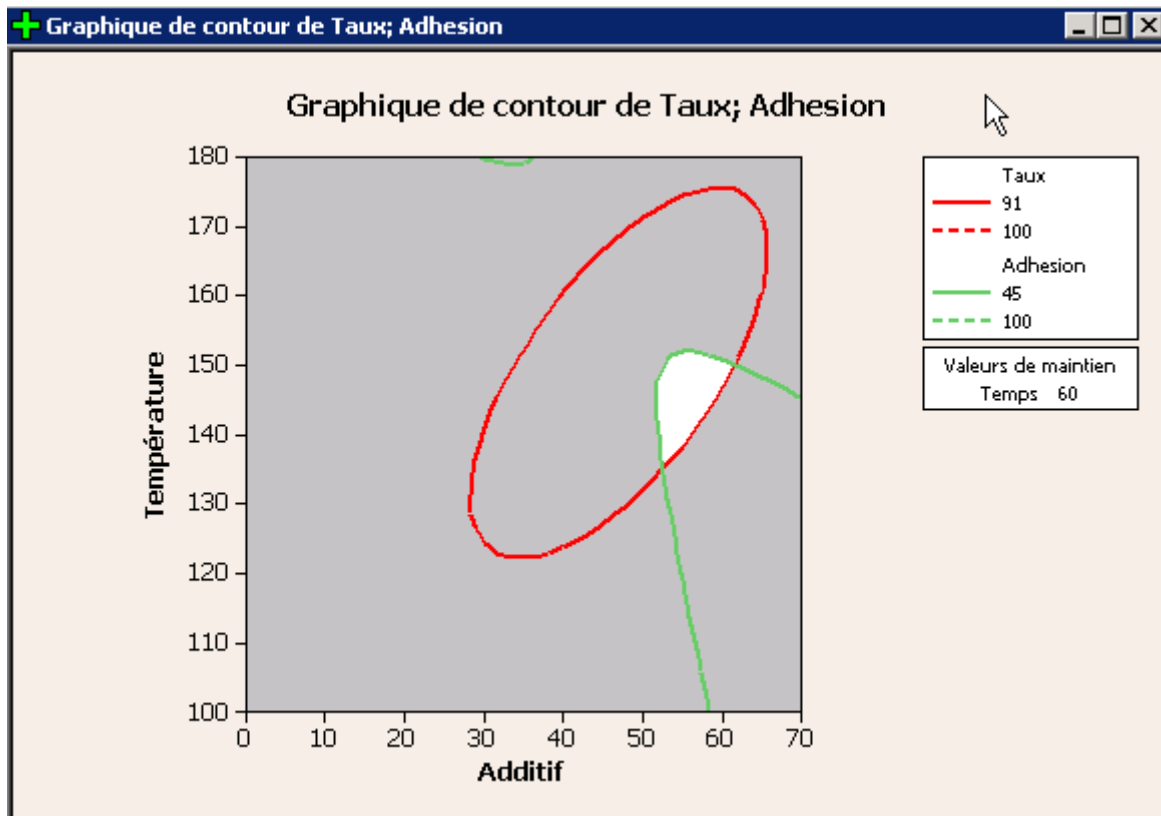
Vous pouvez sélectionner une des trois options de configuration  
OU  
entrer votre propre configuration en saisissant une valeur dans le tableau.  
(Les paramètres représentent des niveaux non codés.)

Maintenir les facteurs supplémentaires à :

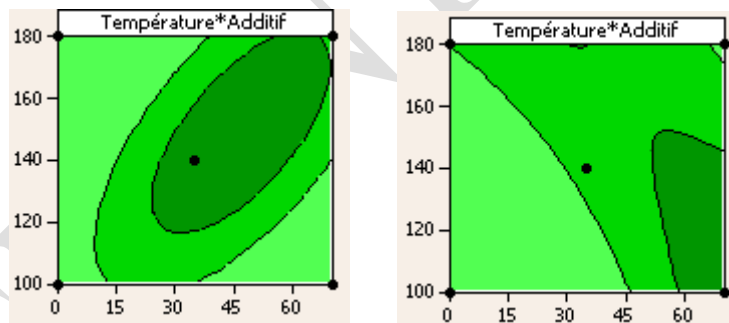
- Paramètre maxi
- Paramètre moyen
- Paramètre mini

Facteur	Nom	Configuration
A	Additif	70
B	Temps	60
C	Température	180

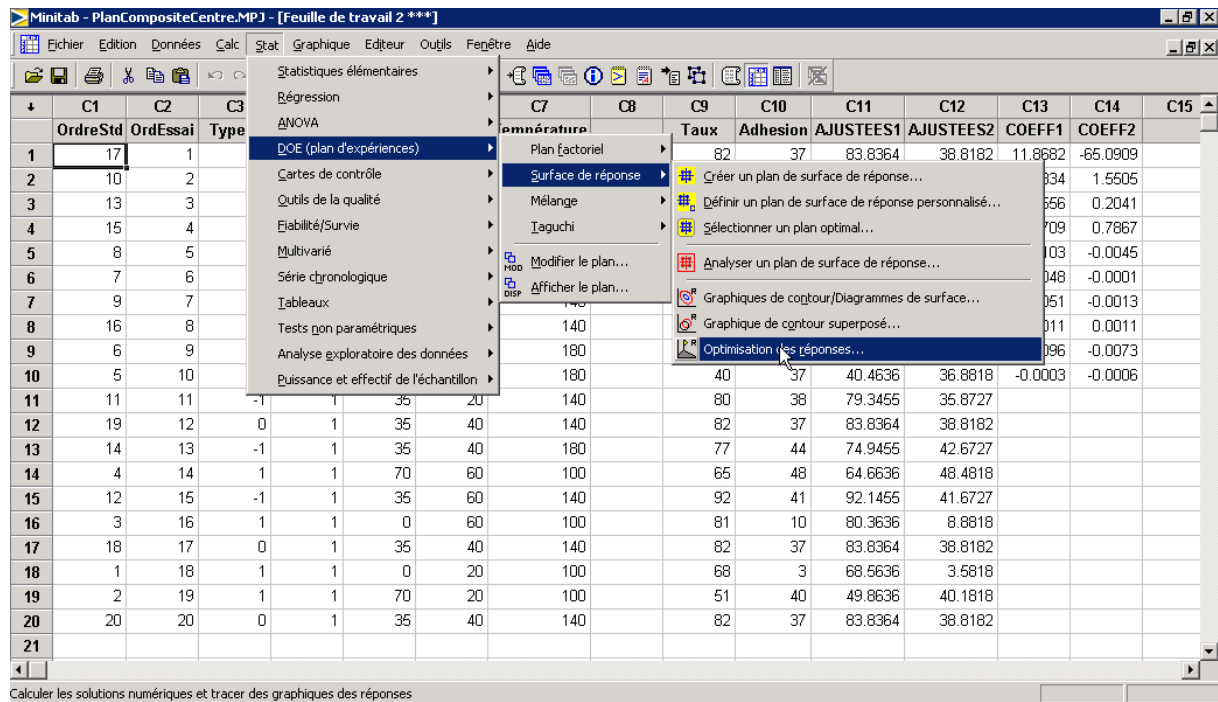
Nous validons le tout par **OK** pour obtenir:



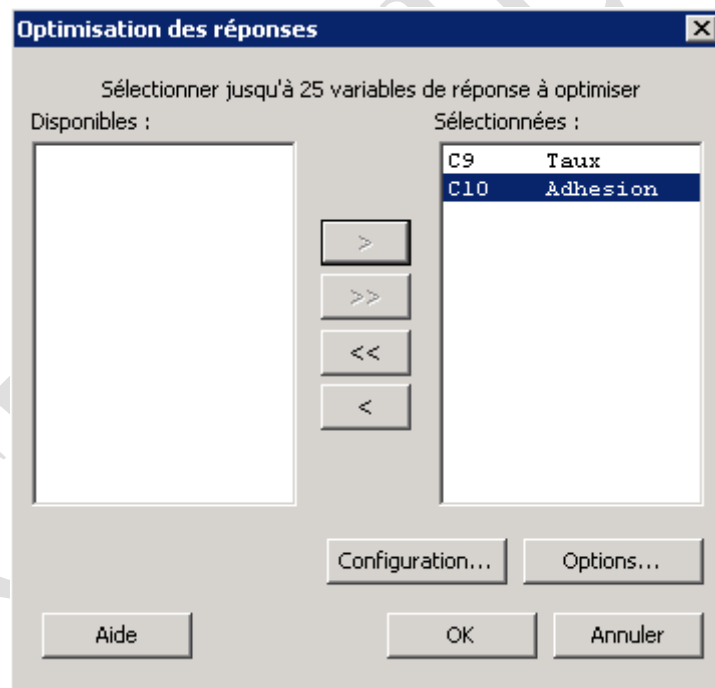
où nous retrouvons bien la superposition des deux petits graphiques suivants vus plus haut respectivement pour le taux et l'adhésion:



Bon l'interprétation qualitative graphique c'est bien sympa mais quand même... nous avons autre chose à faire de notre temps et imaginez dans un cas à 8 facteurs... cela devient impossible! Donc il doit bien y avoir un outil fait pour. Et c'est bien le cas! Nous allons dans le menu **Stat/DOE (plan d'expériences/Surface de réponse/Optimisation des réponses...**



Nous prenons alors les deux réponses:



Nous cliquons sur **Configuration...** pour prendre:

**Optimisation des réponses - Configuration**

Réponse	Objectif	Inférieur	Cible	Supérieur	Pondération	Importance
C9 Taux	Agrandir	91	95		1	1
C10 Adhesion	Agrandir	45	48		1	1

Fonctions de désirabilité pour différents objectifs - Comment les pondérations agissent sur leurs formes

Minimiser la réponse      Atteindre une valeur cible      Maximiser la réponse

Aide      OK      Annuler

Nous avons alors:

Toutes les lignes rouges verticales sont déplaçables!!!

Minitab - PlanCompositeCentre.MPJ - [Diagramme d'optimisation]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

Optimal	Additif	Temps	Températ
Max	70.0	60.0	180.0
D	[55.1515]	[60.0]	[145.2525]
0.13673	0.0	20.0	100.0
Act			
Min			

Composite Désirabilité 0.13673

Taux Maximum  $y = 91.9679$   $d = 0.24198$

Adhesion Maximum  $y = 45.2318$   $d = 0.07726$

OptiPlot

avec dans la fenêtre de session:

### Optimisation des réponses

#### Paramètres

	Objectif	Inférieur	Cible	Supérieur	Pondération	Importer
Taux	Maximum	91	95	95	1	1
Adhesion	Maximum	45	48	48	1	1

#### Solution globale

Additif = 55.1515  
Temps = 60  
Température = 145.253

#### Réponses prévues

Taux = 91.9679 , désirabilité = 0.241977  
Adhesion = 45.2318 , désirabilité = 0.077262

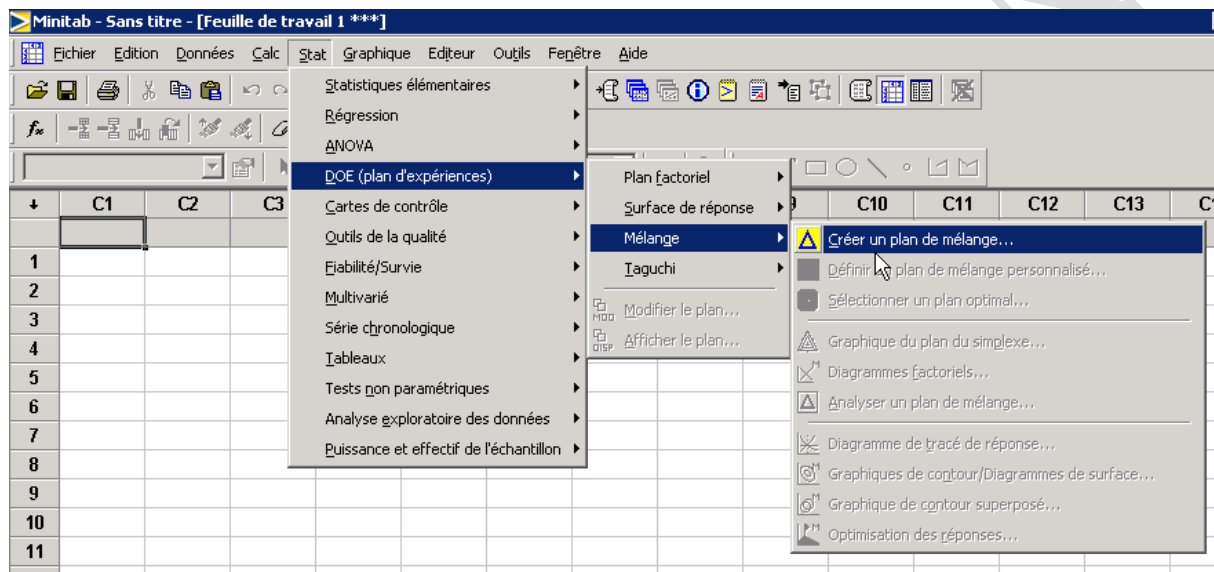
Désirabilité composite = 0.136732

## 17.21. Exercice 192.: Plan de mélange centré du simplexe (sans variable de processus)

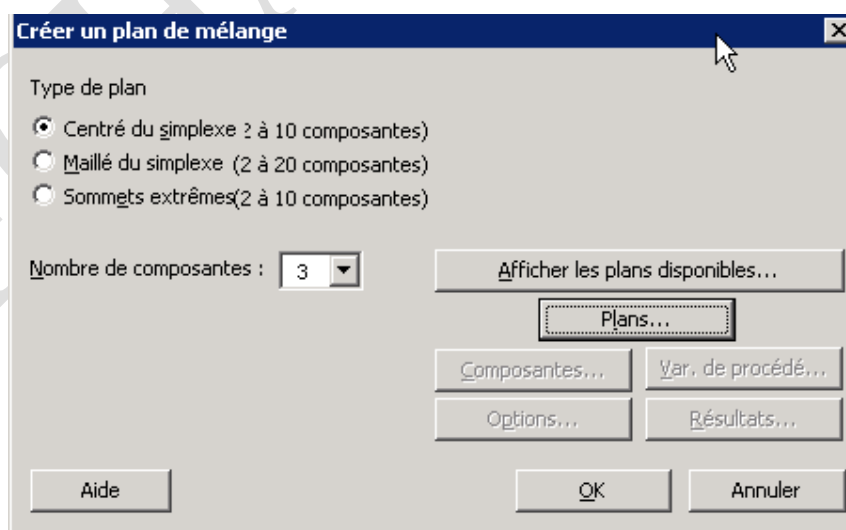
Minitab® Statistical Software 16.2.1

Le but va être de vérifier ici (comme à l'habitude) les calculs appliqués relatifs aux démonstrations mathématiques faites dans le cours théorique relativement aux plans de mélanges avec points aux sommets.

Pour cela, nous allons dans le menu **Stat/DOE (plan d'expériences)/Mélange/Créer un plan de mélange...**:

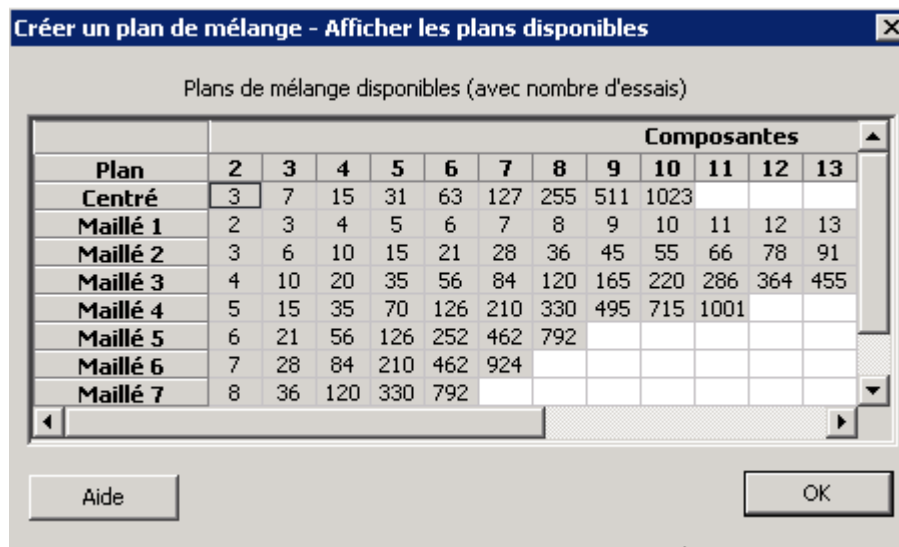


Nous prenons alors un type de plan **Centré du simplexe (2 à 10 composants)** avec comme nombre de composants **3** (comme l'exemple pratique calculé dans le cours théorique):

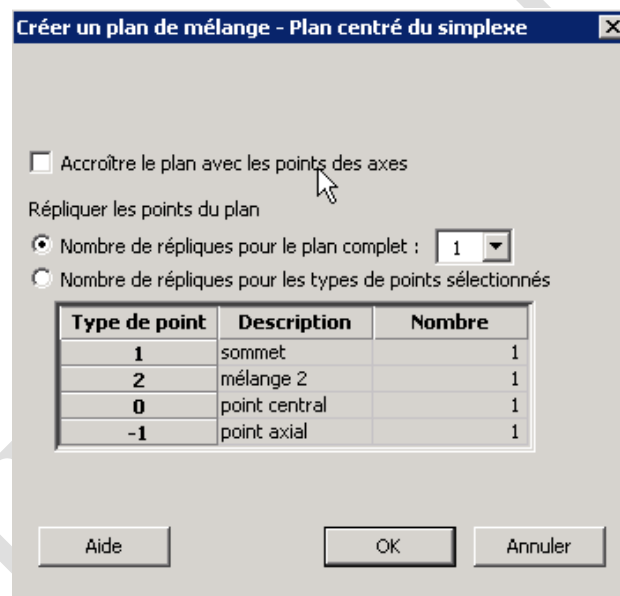


En cliquant sur **Afficher les plans disponibles...**, nous retrouvons les nombres d'essais pour la ligne **Centré** dont nous avons démontré le calcul général dans le cours théorique:

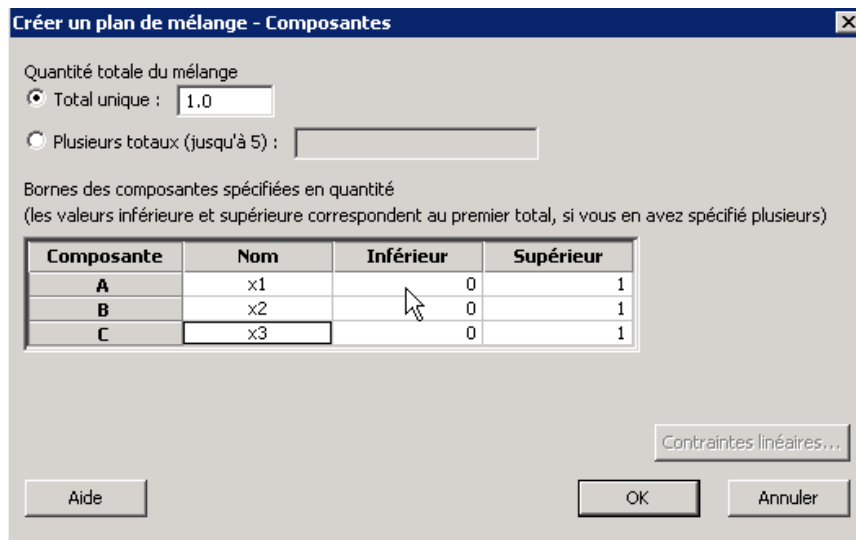




Nous cliquons ensuite sur le bouton **Plans...** pour choisir de décocher **Accroître le plan avec les points des axes** (notions que nous n'avons pas abordée dans le cours théorique):



Nous cliquons ensuite sur le bouton **Composantes...** de la boîte de dialogue principale pour avoir le choix du fameux **Total unique** qui par défaut est déjà à 1 (logique...) et nous renommons les noms des composantes (polyéthylène, polystyrène, polypropylène) comme nous les avons notées dans le cours théorique:



En validant le tout par **OK**, nous obtenons dans la fenêtre de session (nous n'avons pris les captures d'écran que de ce que nous avons vu dans le cours théorique) :

```
Composantes :          3 Points du plan :  7
Variables de procédé :  0 Degré du plan :  3

Total du mélange : 1.00000
```

Bornes des composantes du mélange

Comp	Quantité		Proportion		Pseudo-composante	
	Inférieur	Supérieur	Inférieur	Supérieur	Inférieur	Supérieur
A	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
B	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
C	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000

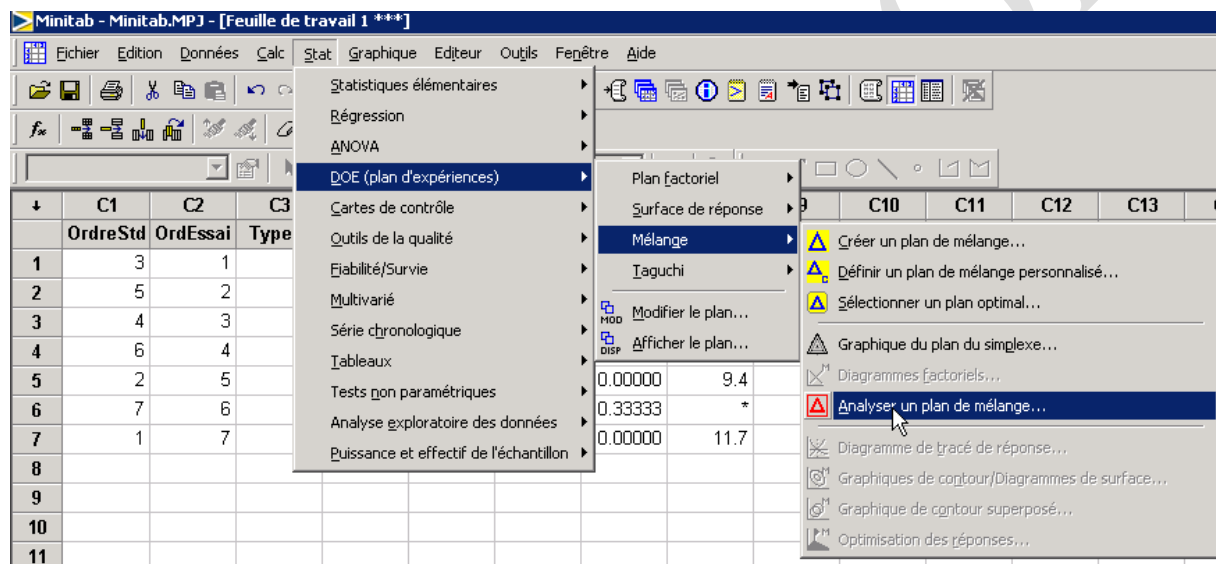
et nous obtenons dans la feuille:

	C1	C2	C3	C4	C5	C6	C7	C8
	OrdreStd	OrdEssai	TypePt	Blocs	x1	x2	x3	
1	3	1	1	1	0.00000	0.00000	1.00000	
2	5	2	2	1	0.50000	0.00000	0.50000	
3	4	3	2	1	0.50000	0.50000	0.00000	
4	6	4	2	1	0.00000	0.50000	0.50000	
5	2	5	1	1	0.00000	1.00000	0.00000	
6	7	6	0	1	0.33333	0.33333	0.33333	
7	1	7	1	1	1.00000	0.00000	0.00000	
8								

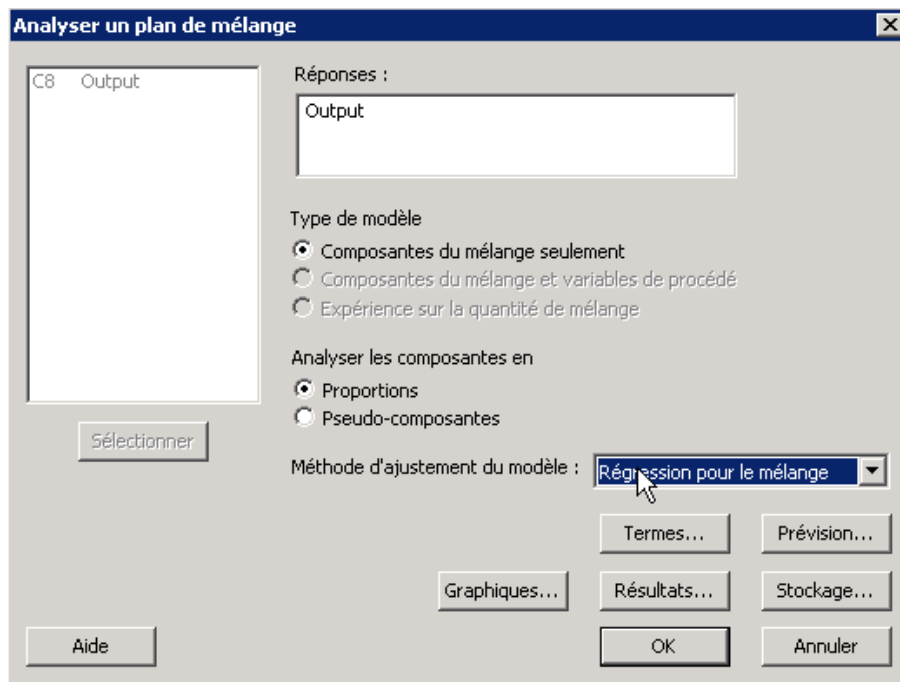
Nous allons y mettre la moyenne des mesures d'élongation (en sautant le point central du simplexe puisque nous n'avons pas traité le sujet dans le cours théorique):

↓	C1	C2	C3	C4	C5	C6	C7	C8
	OrdreStd	OrdEssai	TypePt	Blocs	x1	x2	x3	Output
1	3	1	1	1	0.00000	0.00000	1.00000	16.4
2	5	2	2	1	0.50000	0.00000	0.50000	16.9
3	4	3	2	1	0.50000	0.50000	0.00000	15.3
4	6	4	2	1	0.00000	0.50000	0.50000	10.5
5	2	5	1	1	0.00000	1.00000	0.00000	9.4
6	7	6	0	1	0.33333	0.33333	0.33333	*
7	1	7	1	1	1.00000	0.00000	0.00000	11.7

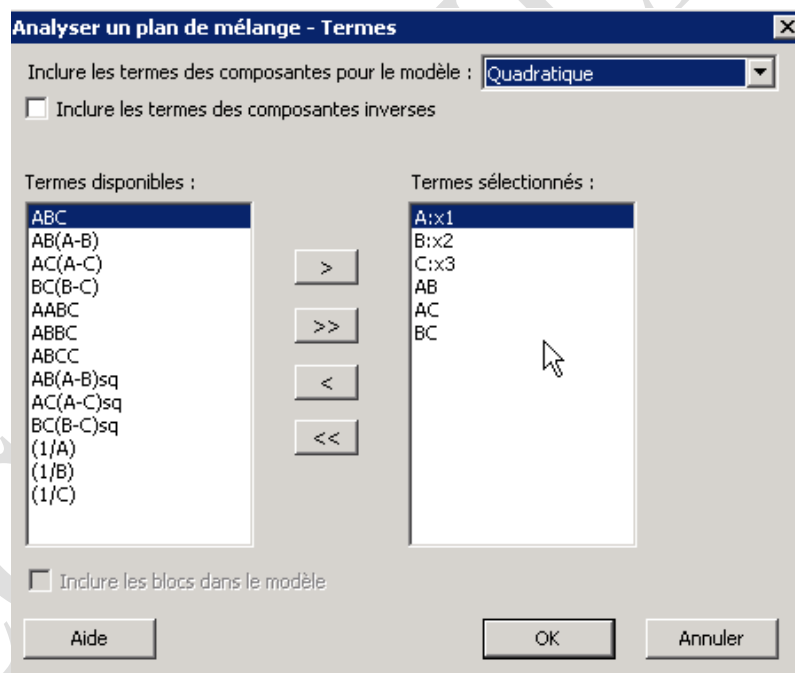
Ensuite, nous allons dans le menu **Stat/DOE (plan d'expériences)/Mélange/Analyser un plan de mélange...**:



pour avoir:



Nous cliquons sur **Termes...** pour s'assurer d'avoir le modèle choisi dans le cours théorique:



Nous validons le tout par **OK** pour obtenir dans la fenêtre de session un premier morceau:

Coefficients de régression estimés pour Output (proportions de composante)

Terme	Coeff	Coef			Facteurs contribuant à l'augmentation de la variance
		ErT	T	P	
x1	11.700	*	*	*	1.500
x2	9.400	*	*	*	1.500
x3	16.400	*	*	*	1.500
x1*x2	19.000	*	*	*	1.500
x1*x3	11.400	*	*	*	1.500
x2*x3	-9.600	*	*	*	1.500

S = \* SomCar-ErrPrév = \*

Alors nous retrouvons bien les valeurs des coefficients calculés à la main mais chose curieuse... Minitab n'arrive pas à calculer l'erreur ni les intervalles de confiance et la  $p$ -value alors que nous avons démontré dans le cours théorique qu'un tel calcul était possible (et nous l'avons fait!). Raison pour laquelle... (encore une fois!) il est important de rappeler qu'il faut toujours plusieurs progiciels de statistiques pour faire des études sérieuses et un peu plus sûres.

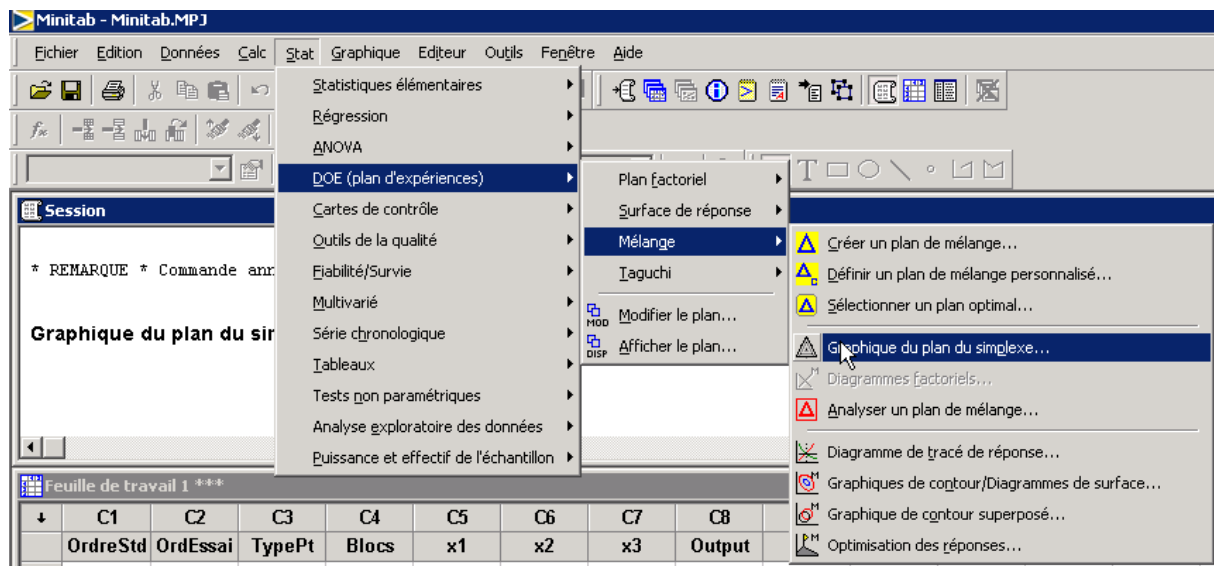
Maintenant voyons le deuxième morceau donné dans la fenêtre de session qui est donc l'ANOVA:

Analyse de la variance pour Output (proportions de composante)

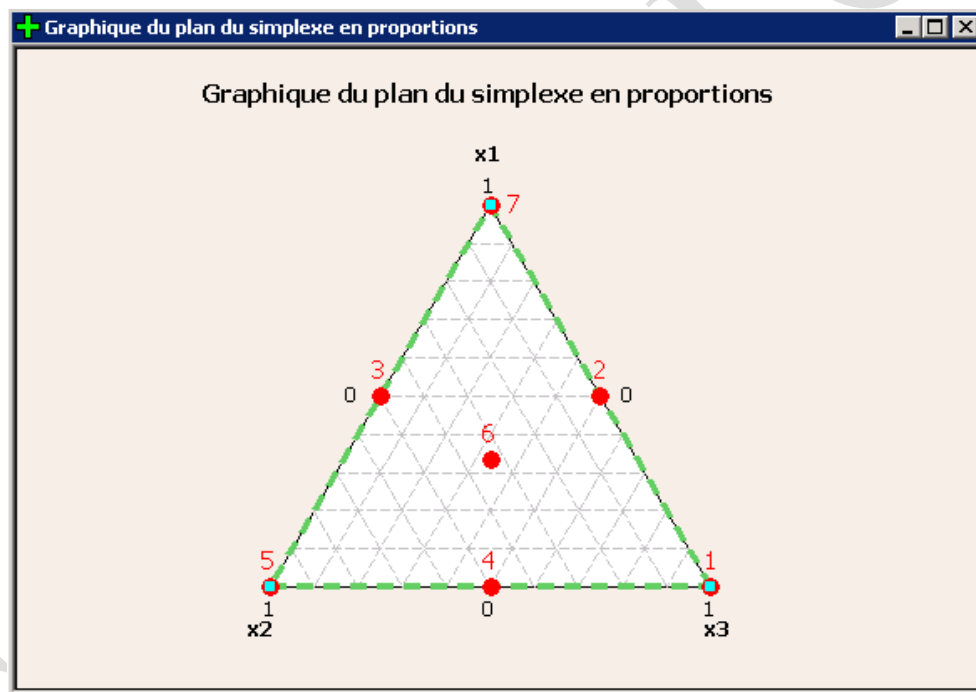
Source	DL	SomCar				F	P
		SomCar séq	ajust	CM ajust			
Régression	5	52.1533	52.1533	10.4307			
Linéaire	2	25.7013	25.4600	12.7300	*	*	
Quadratique	3	26.4520	26.4520	8.8173	*	*	
Erreur résiduelle	0	*	*	*			
Total	5	52.1533					

et là curieusement (c'est probablement à cause de l'absence du point central) rien ne correspond à ce qui a été calculé dans le cours théorique et ni le test de Fisher, ni la  $p$ -value ne sont indiquées alors que nous avons été capable de le faire dans le cours théorique.

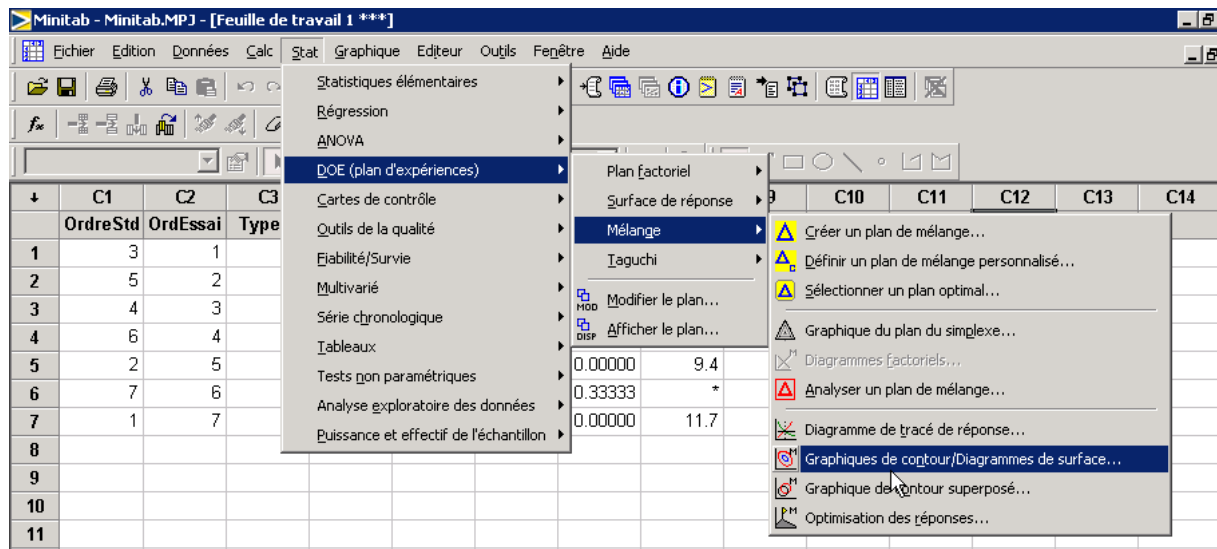
Avant d'aller plus loin et de comparer avec le cours théorique remarquons que Minitab permet une chose amusante qui est l'affichage du simplexe utilisé. Pour cela, nous allons dans le menu **Stat/DOE (plan d'expériences)/Mélange/Graphique du plan du simple...**:



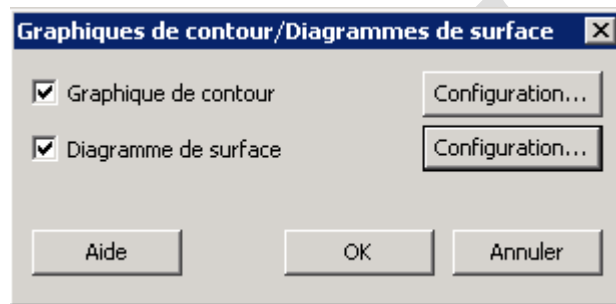
pour obtenir (rien de bien nouveau mais utile à présenter dans les rapport de labo de R&D):



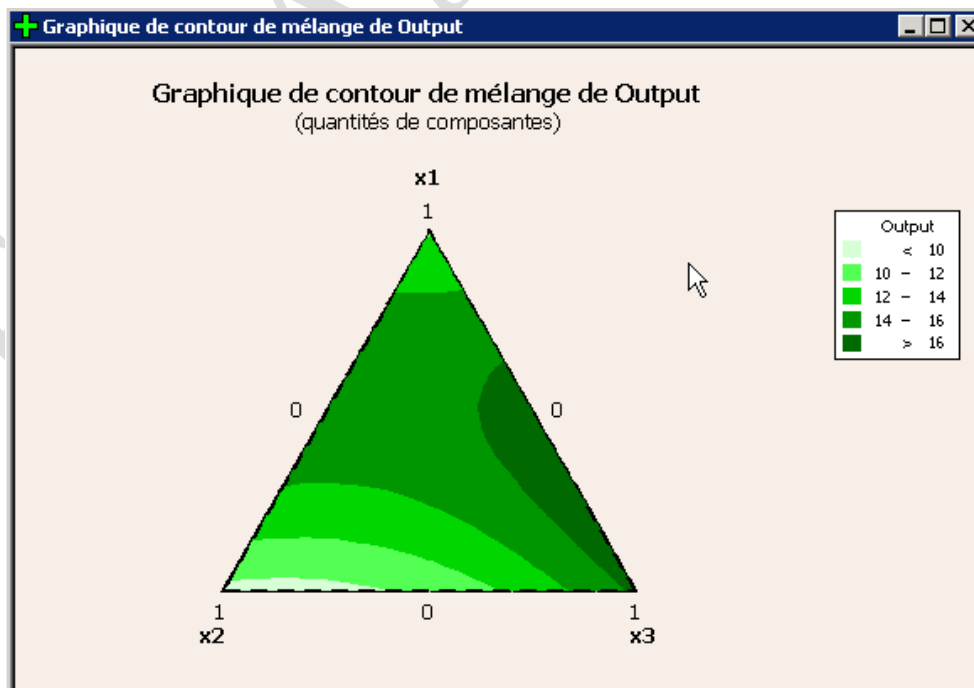
Ensuite, beaucoup plus intéressant, nous pouvons obtenir le graphique 3D de la fonction avec les isoclines en allant dans le menu **Stat/DOE (plan d'expériences)/Mélange/Graphiques de contour/Diagramme de surface...**:



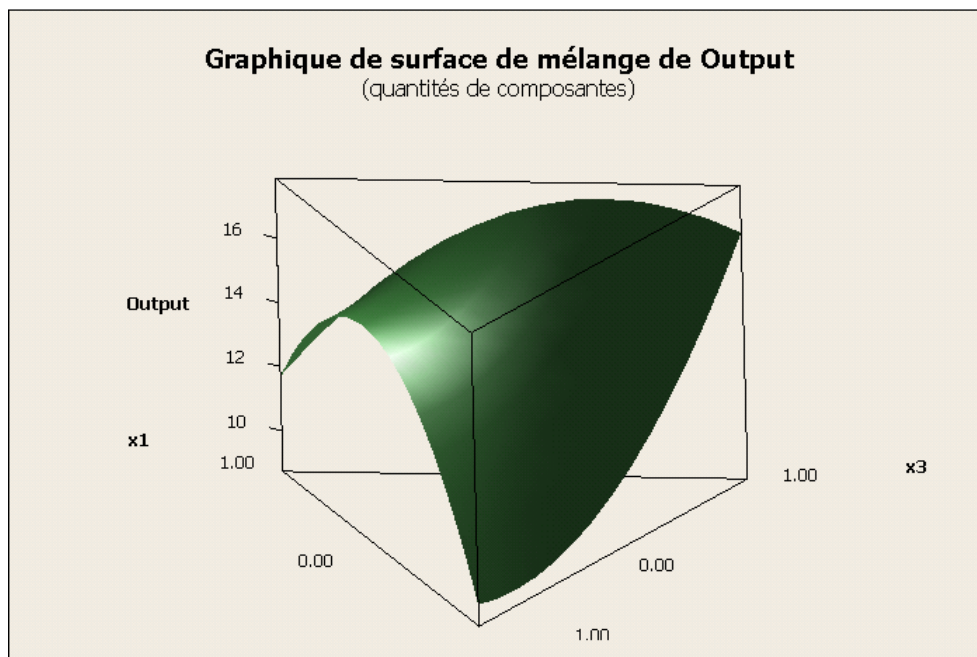
et nous avons:



et en validant par **OK** avec les paramètres de configuration par défaut nous obtenons le premier graphique suivant:

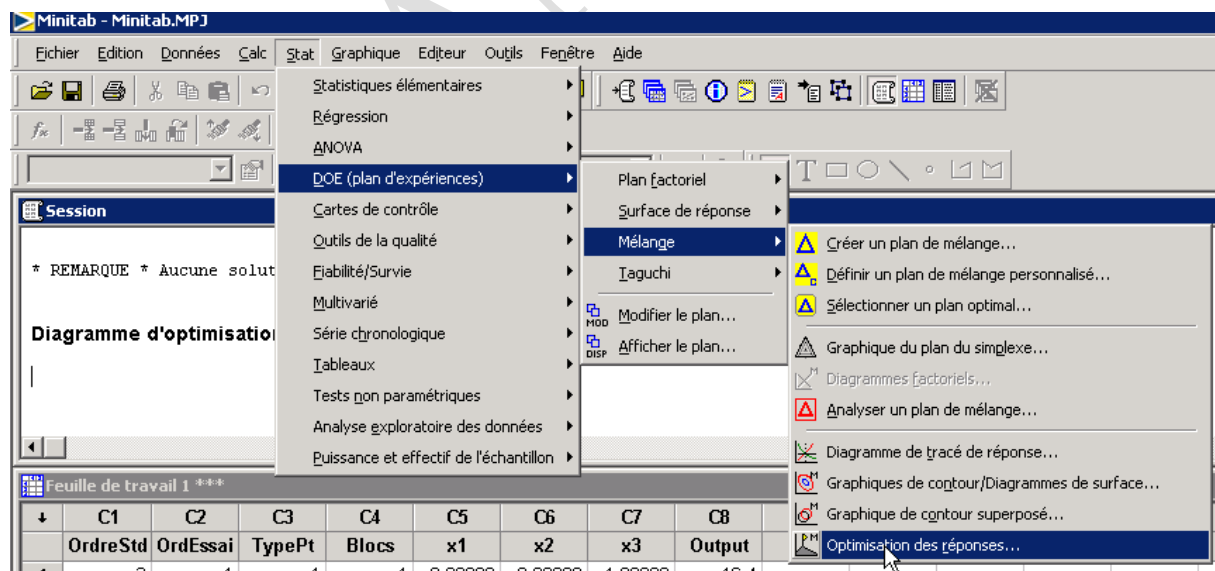


qui est évidemment conforme avec ce que nous avons obtenu dans le cours théorique, ainsi que le graphique 3D suivant (ce n'est pas vraiment toujours très utile mais bon...):



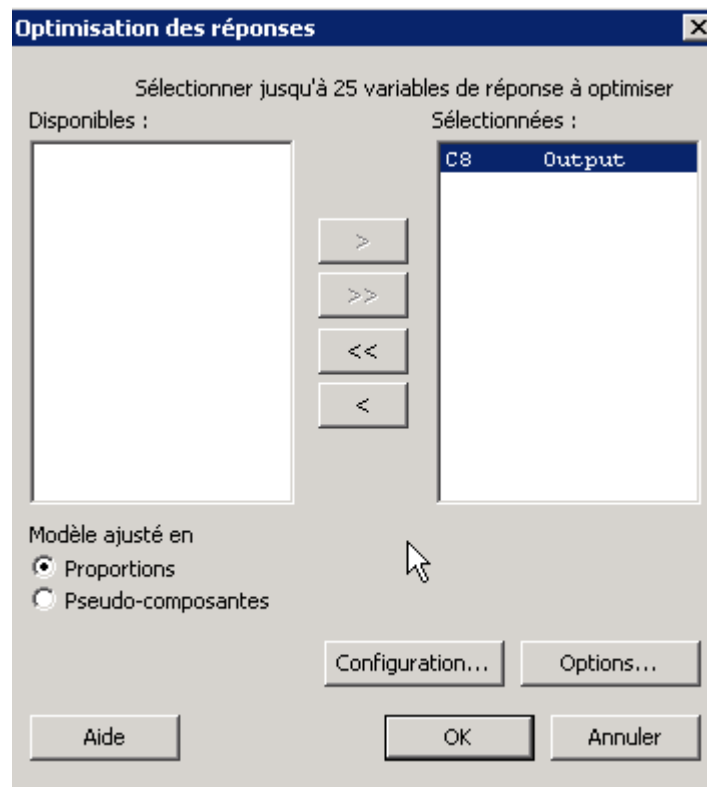
qui est lui aussi conforme à ce que nous avons bien évidemment obtenu dans le cours théorique.

Ensuite, regardons si nous obtenons aussi la même valeur maximale (en admettant que c'est ce que nous recherchons bien évidemment comme dans le cours théorique...). Pour cela, dans Minitab, nous allons dans le menu **Stat/DOE (plan d'expériences)/Mélange/Optimisation des réponses...**:

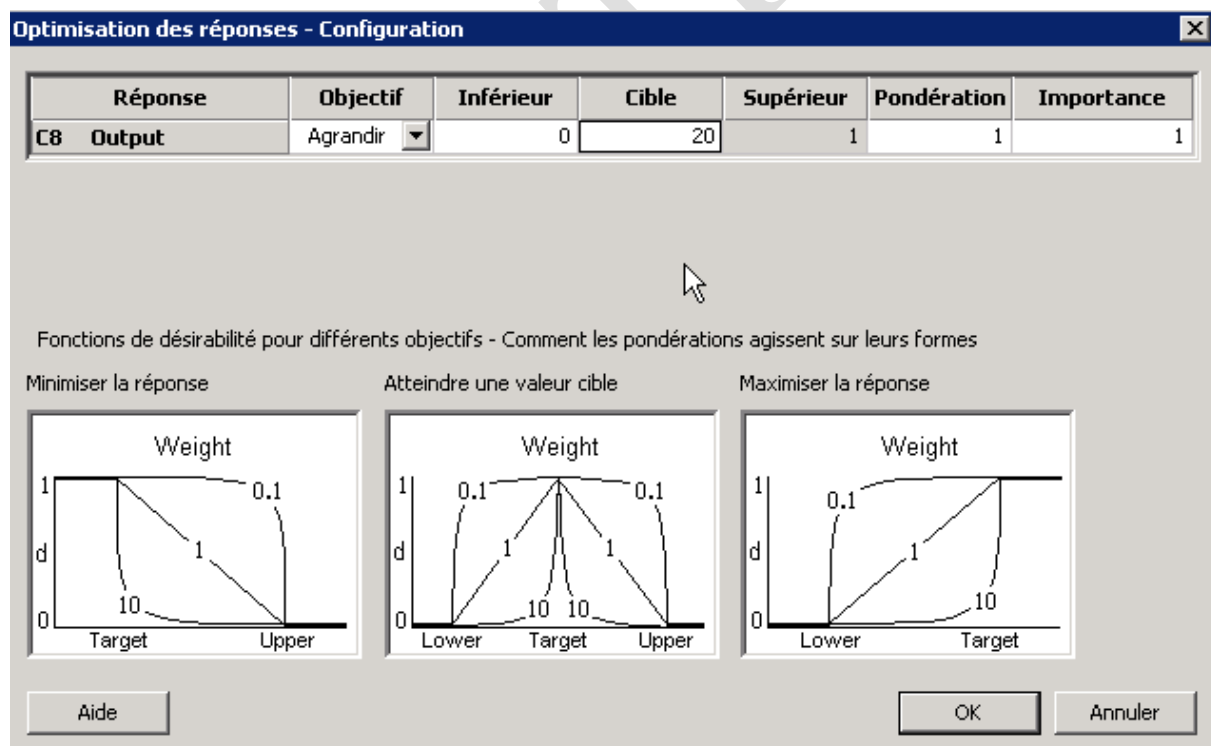


Apparaît alors la boîte de dialogue suivante:

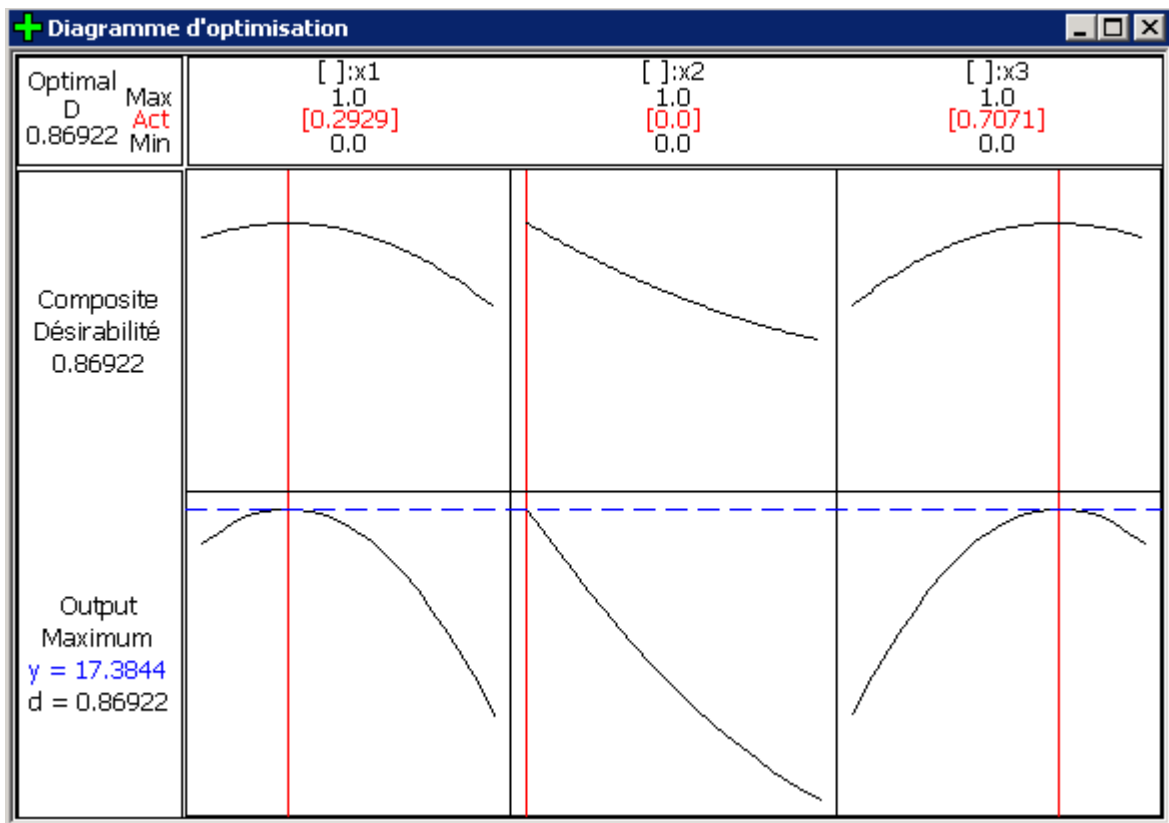




que nous laissons tel quel. Nous cliquons sur le bouton **Configuration...** pour avoir:



où nous avons mis évidemment comme objectif **Agrandir** (merci la traduction... comme à l'habitude...) et arbitrairement nous mettons une cible de 20 (de toute façon ce qui nous intéresse est le maximum et non cette cible). Si nous validons le tout par **OK**, nous obtenons:



Et la réponse optimale est alors de 17.3844 ce qui implique l'absence totale du composant 2 mais le composant 1 en proportion de 0.2929 et le troisième en proportion de 0.7071 (la cible ne peut être atteinte, raison pour laquelle nous avons posé 20... par anticipation).

## 17.22. Exercice 193.: Plan de mélange centré du simplexe avec variables de processus

Minitab® Statistical Software 15.1.1

Le but ici va être de reproduire l'exemple donné ici:

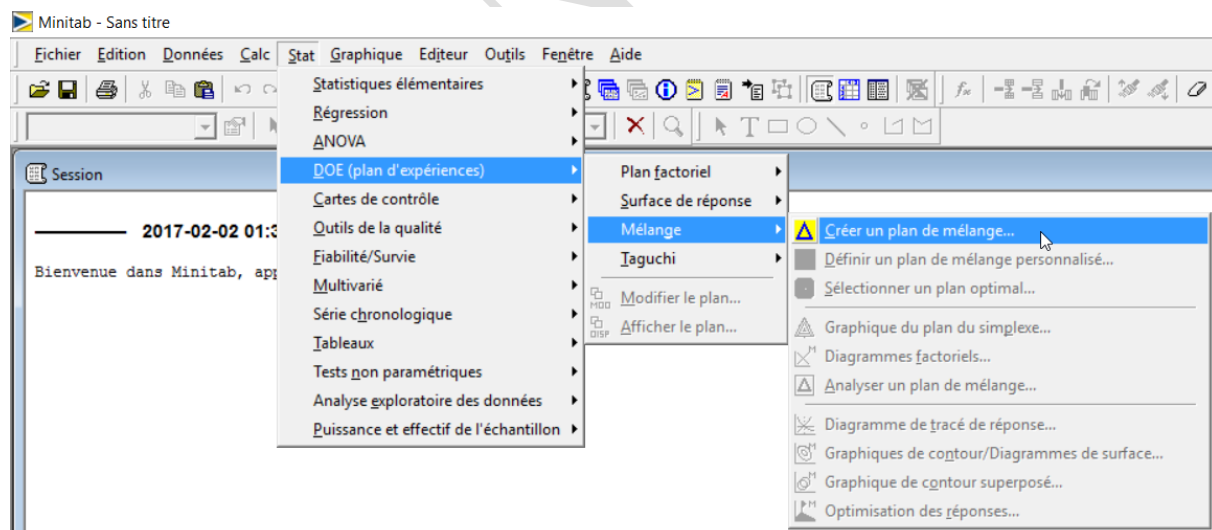
[http://reliawiki.org/index.php/Mixture\\_Design](http://reliawiki.org/index.php/Mixture_Design)

Ce qui est excellent pour comparer aussi les résultats des deux logiciels.

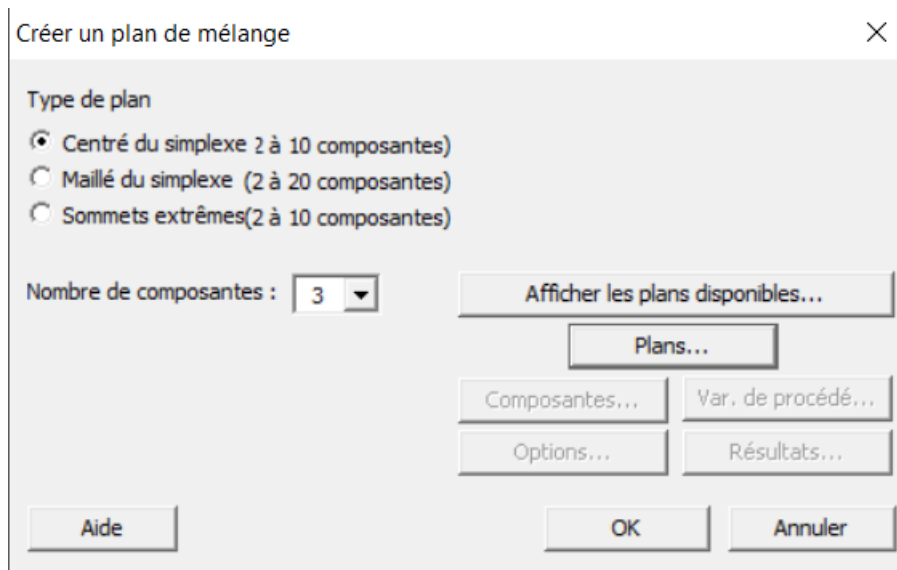
L'énoncé est le suivant au cas où le lien ci-dessus viendrait à ne plus marcher:

Trois sortes de viandes (boeuf, porc et agneau) sont mélangées pour former des galettes de hamburger. La viande comprend 90% du mélange total, les 10% restants étant réservés aux ingrédients aromatisants. Une plan d'expérience du simplexe  $\{p,m\}=\{3,2\}$  (voir cours théorique) avec point central et donc 7 combinaisons est utilisé pour l'expérience. Deux variables de procédé sur la fabrication des galettes sont également étudiées: la température de cuisson et le temps de cuisson. Les valeurs basse et haute température sont 375 °F et 425 °F, et les valeurs de temps faible et élevé sont 25 et 40 minutes. Un plan factoriel complet à deux niveaux est utilisé et affiché ci-dessous avec des valeurs codées. Une des propriétés des hamburgers est la galette. La texture est mesurée par un test de compression qui mesure les grammes de force nécessaires pour percer la surface de la galette.

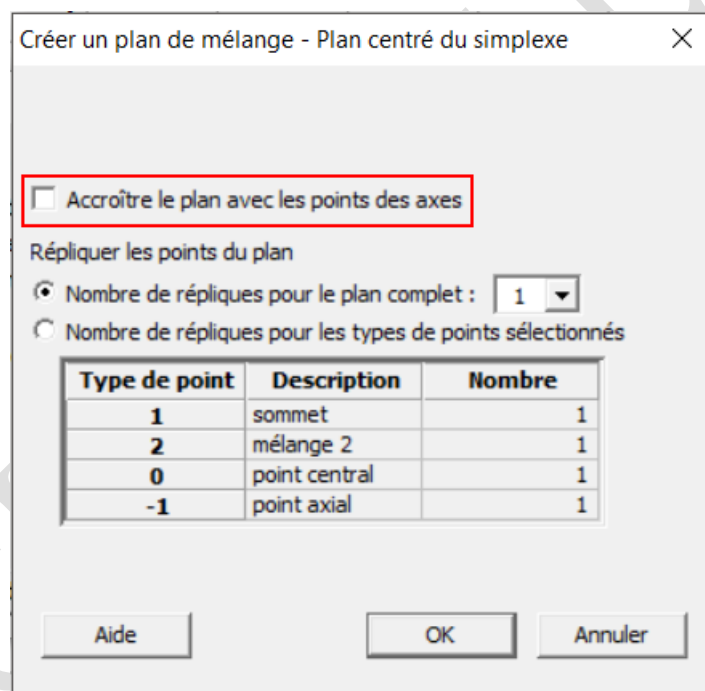
Donc nous allons dans le menu **Stat/DOE (plan d'expérience)/Mélange/Créer un plan de mélange...**:



Nous prenons:



et dans le bouton **Plans...** nous prenons soin de désactiver **Accroître le plan avec les points des axes**:



Et nous cliquons sur le bouton **Var. de procédé** pour prendre:

Créer un plan de mélange - Variables de procédé

Variables de procédé

Aucune

Nombre :  Type de plan :  Fraction numéro :

Variable de procédé	Nom	Type	Min	Max
X1	X1	Numérique	-1	1
X2	X2	Numérique	-1	1
X3	X3	Numérique	-1	1
X4	X4	Numérique	-1	1
X5	X5	Numérique	-1	1
X6	X6	Numérique	-1	1
X7	X7	Numérique	-1	1

Aide OK Annuler

Et dans **Composantes...** nous prenons:

Créer un plan de mélange - Composantes

Quantité totale du mélange

Total unique :

Plusieurs totaux (jusqu'à 5) :

Bornes des composantes spécifiées en quantité  
(les valeurs inférieure et supérieure correspondent au premier total, si vous en avez spécifié plusieurs)

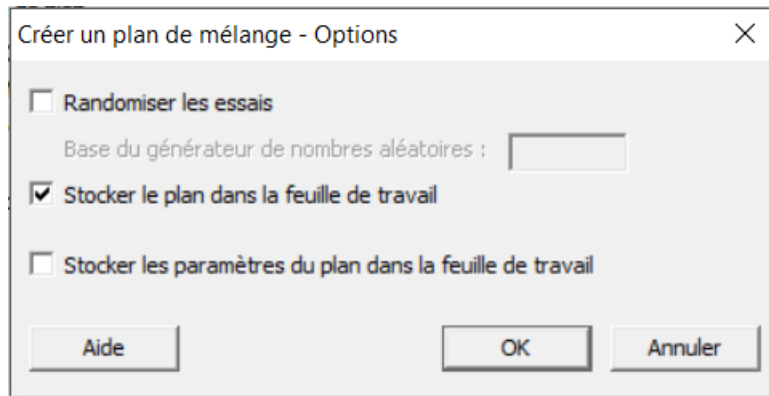
Composante	Nom	Inférieur	Supérieur
A	x1	0	1
B	x2	0	1
C	x3	0	1

Contraintes linéaires...

Aide OK Annuler

**Attention!!!!** Remarquez que nous n'avons pas mis 0.9 ci-dessus (pour le 90% mentionné au début de l'énoncé). Effectivement, la valeur qu'attend Minitab est la valeur normalisée sommée des proportions entre elles et non pas par rapport au grand total.

Nous allons éviter la randomisation pour faciliter la saisie en décochant dans **Options** la case **Randomisation des essais**:



Ce qui donne après avoir validé deux fois par **OK**:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	
	OrdreStd	OrdEssai	TypePt	Blocs	A	B	C	X1	X2	
1	1	1	1	1	1.00000	0.00000	0.00000	-1	-1	
2	2	2	1	1	0.00000	1.00000	0.00000	-1	-1	
3	3	3	1	1	0.00000	0.00000	1.00000	-1	-1	
4	4	4	2	1	0.50000	0.50000	0.00000	-1	-1	
5	5	5	2	1	0.50000	0.00000	0.50000	-1	-1	
6	6	6	2	1	0.00000	0.50000	0.50000	-1	-1	
7	7	7	0	1	0.33333	0.33333	0.33333	-1	-1	
8	8	8	1	1	1.00000	0.00000	0.00000	1	-1	
9	9	9	1	1	0.00000	1.00000	0.00000	1	-1	
10	10	10	1	1	0.00000	0.00000	1.00000	1	-1	
11	11	11	2	1	0.50000	0.50000	0.00000	1	-1	
12	12	12	2	1	0.50000	0.00000	0.50000	1	-1	
13	13	13	2	1	0.00000	0.50000	0.50000	1	-1	
14	14	14	0	1	0.33333	0.33333	0.33333	1	-1	
15	15	15	1	1	1.00000	0.00000	0.00000	-1	1	
16	16	16	1	1	0.00000	1.00000	0.00000	-1	1	
17	17	17	1	1	0.00000	0.00000	1.00000	-1	1	
18	18	18	2	1	0.50000	0.50000	0.00000	-1	1	
19	19	19	2	1	0.50000	0.00000	0.50000	-1	1	
20	20	20	2	1	0.00000	0.50000	0.50000	-1	1	
21	21	21	0	1	0.33333	0.33333	0.33333	-1	1	
22	22	22	1	1	1.00000	0.00000	0.00000	1	1	
23	23	23	1	1	0.00000	1.00000	0.00000	1	1	
24	24	24	1	1	0.00000	0.00000	1.00000	1	1	
25	25	25	2	1	0.50000	0.50000	0.00000	1	1	
26	26	26	2	1	0.50000	0.00000	0.50000	1	1	
27	27	27	2	1	0.00000	0.50000	0.50000	1	1	
28	28	28	0	1	0.33333	0.33333	0.33333	1	1	

Maintenant, nous allons saisir les valeurs données dans le lien au début et dont voici une capture d'écran:

Standard Order	A: Beef	B: Pork	C: Lamb	Z1: Temperature	Z2: Time	Texture (10 <sup>3</sup> gram)
1	1	0	0	-1	-1	1.84
2	0.5	0.5	0	-1	-1	0.67
3	0.5	0	0.5	-1	-1	1.51
4	0	1	0	-1	-1	1.29
5	0	0.5	0.5	-1	-1	1.42
6	0	0	1	-1	-1	1.16
7	0.333	0.333	0.333	-1	-1	1.59
8	1	0	0	1	-1	2.86
9	0.5	0.5	0	1	-1	1.1
10	0.5	0	0.5	1	-1	1.6
11	0	1	0	1	-1	1.53
12	0	0.5	0.5	1	-1	1.81
13	0	0	1	1	-1	1.5
14	0.333	0.333	0.333	1	-1	1.68
15	1	0	0	-1	1	3.01
16	0.5	0.5	0	-1	1	1.21
17	0.5	0	0.5	-1	1	2.32
18	0	1	0	-1	1	1.93
19	0	0.5	0.5	-1	1	2.57
20	0	0	1	-1	1	1.83
21	0.333	0.333	0.333	-1	1	1.94
22	1	0	0	1	1	4.13
23	0.5	0.5	0	1	1	1.67
24	0.5	0	0.5	1	1	2.57
25	0	1	0	1	1	2.26
26	0	0.5	0.5	1	1	3.15
27	0	0	1	1	1	2.22
28	0.333	0.333	0.333	1	1	2.6

Ce qui donne dans Minitab:

Minitab - DOE\_PlanMelanceAvecVariablesProcessus.mpj - [Feuille de travail 2 \*\*\*]

Eichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
	OrdreStd	OrdEssai	TypePt	Blocs	Boeuf	Porc	Agneau	Température	Temps	Texture
1	1	1	1	1	1.00000	0.00000	0.00000	-1	-1	1.84
2	2	2	1	1	0.00000	1.00000	0.00000	-1	-1	1.29
3	3	3	1	1	0.00000	0.00000	1.00000	-1	-1	1.16
4	4	4	2	1	0.50000	0.50000	0.00000	-1	-1	0.67
5	5	5	2	1	0.50000	0.00000	0.50000	-1	-1	1.51
6	6	6	2	1	0.00000	0.50000	0.50000	-1	-1	1.42
7	7	7	0	1	0.33333	0.33333	0.33333	-1	-1	1.59
8	8	8	1	1	1.00000	0.00000	0.00000	1	-1	2.86
9	9	9	1	1	0.00000	1.00000	0.00000	1	-1	1.53
10	10	10	1	1	0.00000	0.00000	1.00000	1	-1	1.50
11	11	11	2	1	0.50000	0.50000	0.00000	1	-1	1.10
12	12	12	2	1	0.50000	0.00000	0.50000	1	-1	1.60
13	13	13	2	1	0.00000	0.50000	0.50000	1	-1	1.81
14	14	14	0	1	0.33333	0.33333	0.33333	1	-1	1.68
15	15	15	1	1	1.00000	0.00000	0.00000	-1	1	3.01
16	16	16	1	1	0.00000	1.00000	0.00000	-1	1	1.93
17	17	17	1	1	0.00000	0.00000	1.00000	-1	1	1.83
18	18	18	2	1	0.50000	0.50000	0.00000	-1	1	1.21
19	19	19	2	1	0.50000	0.00000	0.50000	-1	1	2.32
20	20	20	2	1	0.00000	0.50000	0.50000	-1	1	2.57
21	21	21	0	1	0.33333	0.33333	0.33333	-1	1	1.94
22	22	22	1	1	1.00000	0.00000	0.00000	1	1	4.13
23	23	23	1	1	0.00000	1.00000	0.00000	1	1	2.26
24	24	24	1	1	0.00000	0.00000	1.00000	1	1	2.22
25	25	25	2	1	0.50000	0.50000	0.00000	1	1	1.67
26	26	26	2	1	0.50000	0.00000	0.50000	1	1	2.57
27	27	27	2	1	0.00000	0.50000	0.50000	1	1	3.15
28	28	28	0	1	0.33333	0.33333	0.33333	1	1	2.60

Nous lançons l'analyse en allant dans le menu DOE (plan d'expérience)/Mélange/Analyser un plan de mélange...:

Minitab - DOE\_PlanMelanceAvecVariablesProcessus.mpj - [Feuille de travail 2 \*\*\*]

Eichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

Statistiques élémentaires  
Régression  
ANOVA  
DOE (plan d'expériences)  
Cartes de contrôle  
Outils de la qualité  
Fiabilité/Survie  
Multivarié  
Série chronologique  
Tableaux  
Tests non paramétriques  
Analyse exploratoire des données  
Puissance et effectif de l'échantillon

Plan factoriel  
Surface de réponse  
Mélange  
Taguchi  
Modifier le plan...  
Afficher le plan...

Plan factoriel  
Surface de réponse  
Mélange  
Taguchi  
Modifier le plan...  
Afficher le plan...

Créer un plan de mélange...  
Définir un plan de mélange personnalisé...  
Sélectionner un plan optimal...  
Graphique du plan du simplexe...  
Diagrammes factoriels...  
Analyser un plan de mélange...  
Diagramme de tracé de réponse...  
Graphiques de contour/Diagrammes de surface...  
Graphique de contour superposé...  
Optimisation des réponses...

	C9	C10	C11	C12	C13
	emps	Texture			
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					



Nous avons:

Analyser un plan de mélange

Réponses :  
Texture

Type de modèle  
 Composantes du mélange seulement  
 Composantes du mélange et variables de procédé  
 Expérience sur la quantité de mélange

Analyser les composantes en  
 Proportions  
 Pseudo-composantes

Méthode d'ajustement du modèle : Régression pour le mélange

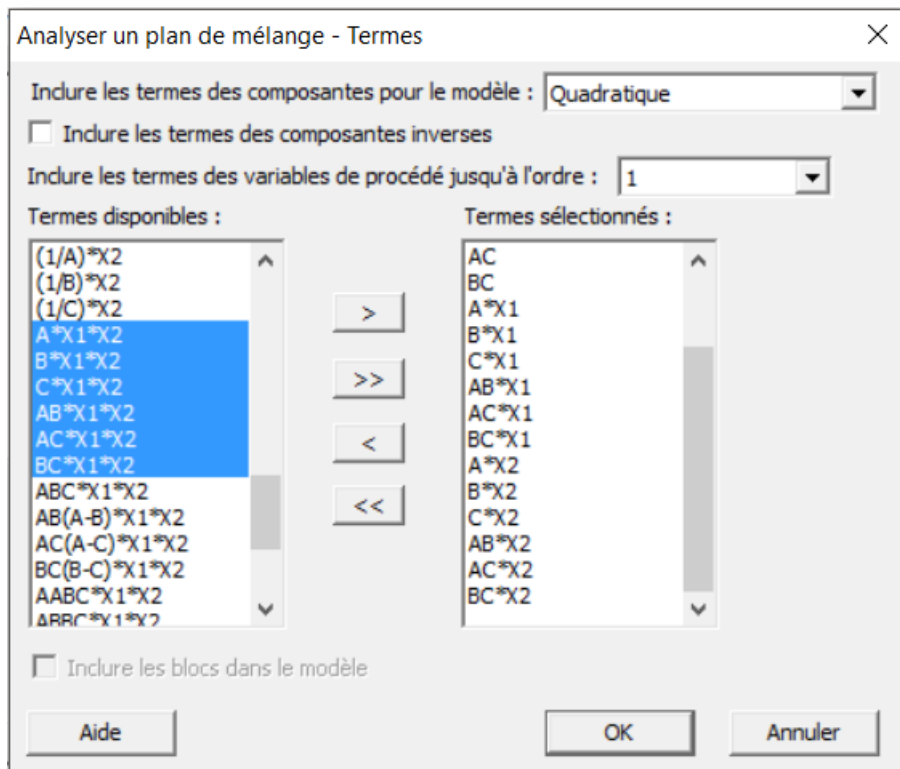
Sélectionner

Termes... Prévission...  
Graphiques... Résultats... Stockage...  
Aide OK Annuler

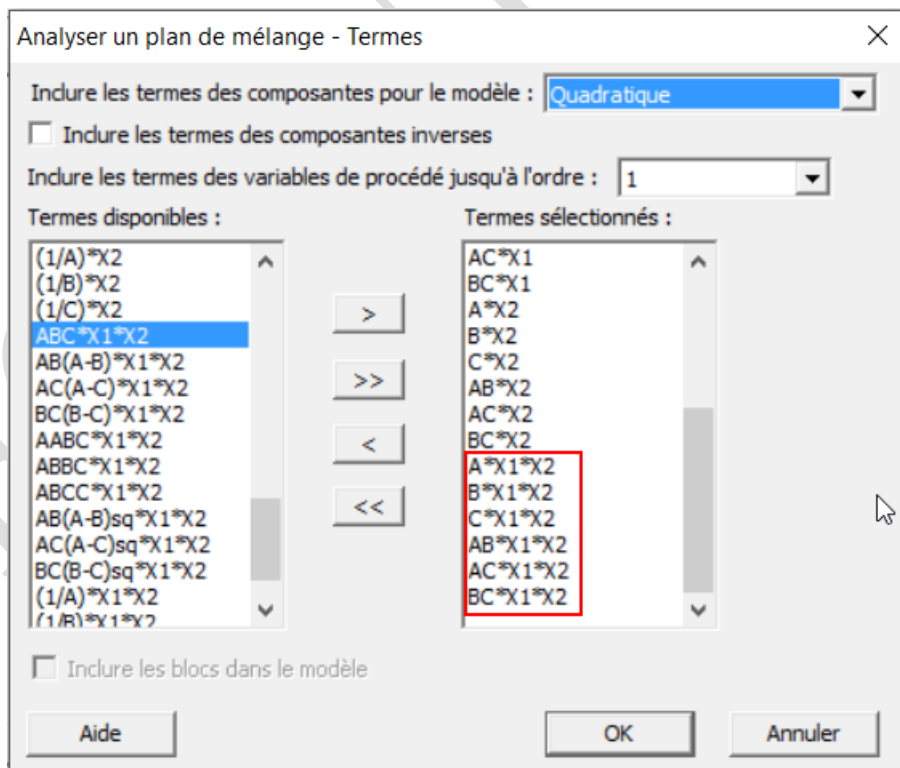
Et pour faire le même modèle que sur l'exemple donné en lien:

Term	Coefficient	Standard Error	T Value	P Value	Variance Inflation Factor
A:Beef	2.9421	0.1236	*	*	1.5989
B:Pork	1.7346	0.1236	*	*	1.5989
C:Lamb	1.6596	0.1236	*	*	1.5989
A • B	-4.4170	0.5680	-7.7766	0.0015	1.5695
A • C	-0.9170	0.5680	-1.6146	0.1817	1.5695
B • C	2.4480	0.5680	4.3099	0.0125	1.5695
Z1 • A	0.5324	0.1236	4.3084	0.0126	1.5989
Z1 • B	0.1399	0.1236	1.1319	0.3209	1.5989
Z1 • C	0.1799	0.1236	1.4557	0.2192	1.5989
Z1 • A • B	-0.4123	0.5680	-0.7260	0.5081	1.5695
Z1 • A • C	-1.0423	0.5680	-1.8352	0.1404	1.5695
Z1 • B • C	0.3727	0.5680	0.6561	0.5476	1.5695
Z2 • A	0.6193	0.1236	5.0117	0.0074	1.5989
Z2 • B	0.3518	0.1236	2.8468	0.0465	1.5989
Z2 • C	0.3568	0.1236	2.8873	0.0447	1.5989
Z2 • A • B	-0.9802	0.5680	-1.7258	0.1595	1.5695
Z2 • A • C	-0.3202	0.5680	-0.5638	0.6030	1.5695
Z2 • B • C	0.9248	0.5680	1.6282	0.1788	1.5695
Z1 • Z2 • A	0.0177	0.1236	0.1433	0.8930	1.5989
Z1 • Z2 • B	0.0152	0.1236	0.1231	0.9080	1.5989
Z1 • Z2 • C	0.0052	0.1236	0.0422	0.9684	1.5989
Z1 • Z2 • A • B	0.0808	0.5680	0.1423	0.8937	1.5695
Z1 • Z2 • A • C	0.2308	0.5680	0.4064	0.7052	1.5695
Z1 • Z2 • B • C	0.2658	0.5680	0.4680	0.6641	1.5695

Nous allons cliquer sur le bouton **Termes** et rajouter les éléments en surbrillance:



Nous avons alors bien 24 termes:



Nous validons deux fois par **OK** en nous obtenons:

**Plan centré du simplexe**

Composantes : 3 Points du plan : 28  
 Variables de procédé : 2 Degré du plan : 3

Total du mélange : 1.00000

Nombre de limites pour chaque dimension

Type de point	1	2	0
Dimension	0	1	2
Nombre	3	3	1

Nombre de points du plan pour chaque type

Type de point	1	2	3	0	-1
Distincts	12	12	0	4	0
Répliques	1	1	0	1	0
Nombre total	12	12	0	4	0

Bornes des composantes du mélange

Comp	Quantité		Proportion		Pseudo-composante	
	Inférieur	Supérieur	Inférieur	Supérieur	Inférieur	Supérieur
A	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
B	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000
C	0.0000	1.0000	0.0000	1.0000	0.0000	1.0000

**Régression pour mélanges : Texture en fonction de Boeuf, Porc, ...**

Coefficients de régression estimés pour Texture (proportions de composante)

Terme	Coeff	Coef	ErT	T	P
Boeuf	2.942	0.1236		*	*
Porc	1.735	0.1236		*	*
Agneau	1.660	0.1236		*	*
Boeuf*Porc	-4.417	0.5680		-7.78	0.001
Boeuf*Agneau	-0.917	0.5680		-1.61	0.182
Porc*Agneau	2.448	0.5680		4.31	0.013
Boeuf*Température	0.532	0.1236		4.31	0.013
Porc*Température	0.140	0.1236		1.13	0.321
Agneau*Température	0.180	0.1236		1.46	0.219
Boeuf*Porc*Température	-0.412	0.5680		-0.73	0.508
Boeuf*Agneau*Température	-1.042	0.5680		-1.84	0.140
Porc*Agneau*Température	0.373	0.5680		0.66	0.548
Boeuf*Temps	0.619	0.1236		5.01	0.007
Porc*Temps	0.352	0.1236		2.85	0.047
Agneau*Temps	0.357	0.1236		2.89	0.045
Boeuf*Porc*Temps	-0.980	0.5680		-1.73	0.159
Boeuf*Agneau*Temps	-0.320	0.5680		-0.56	0.603
Porc*Agneau*Temps	0.925	0.5680		1.63	0.179
Boeuf*Température*Temps	0.018	0.1236		0.14	0.893
Porc*Température*Temps	0.015	0.1236		0.12	0.908
Agneau*Température*Temps	0.005	0.1236		0.04	0.968
Boeuf*Porc*Température*Temps	0.081	0.5680		0.14	0.894
Boeuf*Agneau*Température*Temps	0.231	0.5680		0.41	0.705
Porc*Agneau*Température*Temps	0.266	0.5680		0.47	0.664

Terme	Facteurs contribuant à l'augmentation de la variance
Boeuf	1.599
Porc	1.599
Agneau	1.599
Boeuf*Porc	1.569
Boeuf*Agneau	1.569
Porc*Agneau	1.569
Boeuf*Température	1.599
Porc*Température	1.599
Agneau*Température	1.599
Boeuf*Porc*Température	1.569
Boeuf*Agneau*Température	1.569
Porc*Agneau*Température	1.569
Boeuf*Temps	1.599
Porc*Temps	1.599
Agneau*Temps	1.599
Boeuf*Porc*Temps	1.569
Boeuf*Agneau*Temps	1.569
Porc*Agneau*Temps	1.569
Boeuf*Température*Temps	1.599
Porc*Température*Temps	1.599
Agneau*Température*Temps	1.599
Boeuf*Porc*Température*Temps	1.569
Boeuf*Agneau*Température*Temps	1.569
Porc*Agneau*Température*Temps	1.569

\* REMARQUE \* Les coefficients sont calculés pour les variables de procédé codées.

S = 0.248069      SomCar-ErrPrév = 103.970  
R carré = 98.35 %      R carré (prév) = 0.00 %      R carré (ajust) = 88.85 %

## Analyse de la variance pour Texture (proportions de composante)

Source	DL	SomCar séq	SomCar ajust	CM ajust
Régression	23	14.6619	14.66186	0.63747
Composante seulement				
Linéaire	2	1.9825	4.14465	2.07233
Quadratique	3	5.1982	5.19823	1.73274
Composante * Température				
Linéaire	3	1.6893	1.36384	0.45461
Quadratique	3	0.2692	0.26917	0.08972
Composante * Temps				
Linéaire	3	5.0662	2.59266	0.86422
Quadratique	3	0.3825	0.38252	0.12751
Composante * Température * Temps				
Linéaire	3	0.0507	0.00233	0.00078
Quadratique	3	0.0233	0.02327	0.00776
Erreur résiduelle	4	0.2462	0.24615	0.06154
Total	27	14.9080		

Source	F	P
Régression	10.36	0.017
Composante seulement		
Linéaire	33.68	0.003
Quadratique	28.16	0.004
Composante * Température		
Linéaire	7.39	0.042
Quadratique	1.46	0.352
Composante * Temps		
Linéaire	14.04	0.014
Quadratique	2.07	0.247
Composante * Température * Temps		
Linéaire	0.01	0.998
Quadratique	0.13	0.940
Erreur résiduelle		
Total		

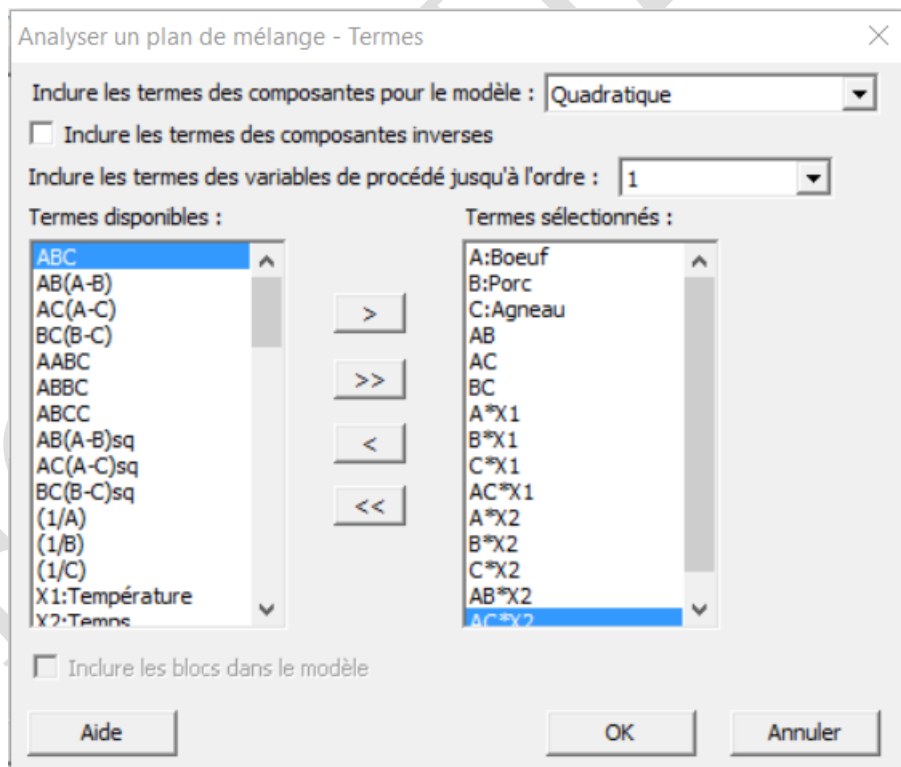
Observations aberrantes pour Texture

Observation	OrdreStd	Texture	Valeur ajustée	Ajust ErT	Valeur résiduelle	Valeur résiduelle normalisée
1	1	1.840	1.808	0.247	0.032	1.47 X
2	2	1.290	1.258	0.247	0.032	1.47 X
3	3	1.160	1.128	0.247	0.032	1.47 X
8	8	2.860	2.837	0.247	0.023	1.04 X
9	9	1.530	1.507	0.247	0.023	1.04 X
10	10	1.500	1.477	0.247	0.023	1.04 X
15	15	3.010	3.011	0.247	-0.001	-0.06 X
16	16	1.930	1.931	0.247	-0.001	-0.06 X
17	17	1.830	1.831	0.247	-0.001	-0.06 X
22	22	4.130	4.111	0.247	0.019	0.86 X
23	23	2.260	2.241	0.247	0.019	0.86 X
24	24	2.220	2.201	0.247	0.019	0.86 X

X indique une observation ayant un effet de levier important.

Nous avons donc bien les mêmes résultats que ceux obtenus avec le logiciel DOE++ de ReliaSoft (ouf!).

Le tableau ci-dessus montre que certains termes ont des *p*-values très importantes, donc, nous pouvons supprimer ces termes du modèle. Nous pouvons également supprimer d'autres termes dont la *p*-value est supérieure à 0.5. Après avoir recalculé avec les 15 termes souhaités:



les résultats finaux sont alors:

### Régression pour mélanges : Texture en fonction de Boeuf, Porc, ...

Coefficients de régression estimés pour Texture (proportions de composante)

Terme	Coeff	Coef	ErT	T	P	Facteurs contribuant
						à l'augmentation de la variance
Boeuf	2.942	0.10199		*	*	1.599
Porc	1.735	0.10199		*	*	1.599
Agneau	1.660	0.10199		*	*	1.599
Boeuf*Porc	-4.417	0.46881	-9.42	0.000		1.569
Boeuf*Agneau	-0.917	0.46881	-1.96	0.072		1.569
Porc*Agneau	2.448	0.46881	5.22	0.000		1.569
Boeuf*Température	0.492	0.09309	5.28	0.000		1.332
Porc*Température	0.137	0.08448	1.62	0.130		1.097
Agneau*Température	0.218	0.09309	2.34	0.036		1.332
Boeuf*Agneau*Température	-1.041	0.46785	-2.22	0.044		1.563
Boeuf*Temps	0.613	0.10193	6.01	0.000		1.597
Porc*Temps	0.433	0.09324	4.65	0.000		1.336
Agneau*Temps	0.438	0.09324	4.70	0.000		1.336
Boeuf*Porc*Temps	-1.023	0.46831	-2.18	0.048		1.566
Boeuf*Agneau*Temps	-0.363	0.46831	-0.78	0.452		1.566

\* REMARQUE \* Les coefficients sont calculés pour les variables de procédé codées.

S = 0.204752      SomCar-ErrPrév = 1.98148  
R carré = 96.34 %      R carré (prév) = 86.71 %      R carré (ajust) = 92.41 %

Analyse de la variance pour Texture (proportions de composante)

Source	DL	SomCar séq	SomCar ajust	CM ajust	F
Régression	14	14.3630	14.36301	1.02593	24.47
Composante seulement					
Linéaire	2	1.9825	4.14465	2.07233	49.43
Quadratique	3	5.1982	5.19823	1.73274	41.33
Composante * Température					
Linéaire	3	1.6893	1.80205	0.60068	14.33
Quadratique	1	0.2074	0.20740	0.20740	4.95
Composante * Temps					
Linéaire	3	5.0662	3.78229	1.26076	30.07
Quadratique	2	0.2194	0.21939	0.10970	2.62
Erreur résiduelle	13	0.5450	0.54500	0.04192	
Total	27	14.9080			

Source	P
Régression	0.000
Composante seulement	
Linéaire	0.000
Quadratique	0.000
Composante * Température	
Linéaire	0.000
Quadratique	0.044
Composante * Temps	
Linéaire	0.000
Quadratique	0.111
Erreur résiduelle	
Total	

On voit ici que la sortie de Minitab est assez "malheureuse". L'idéal serait d'avoir la même chose que DOE++ de Reliasoft, c'est-à-dire:

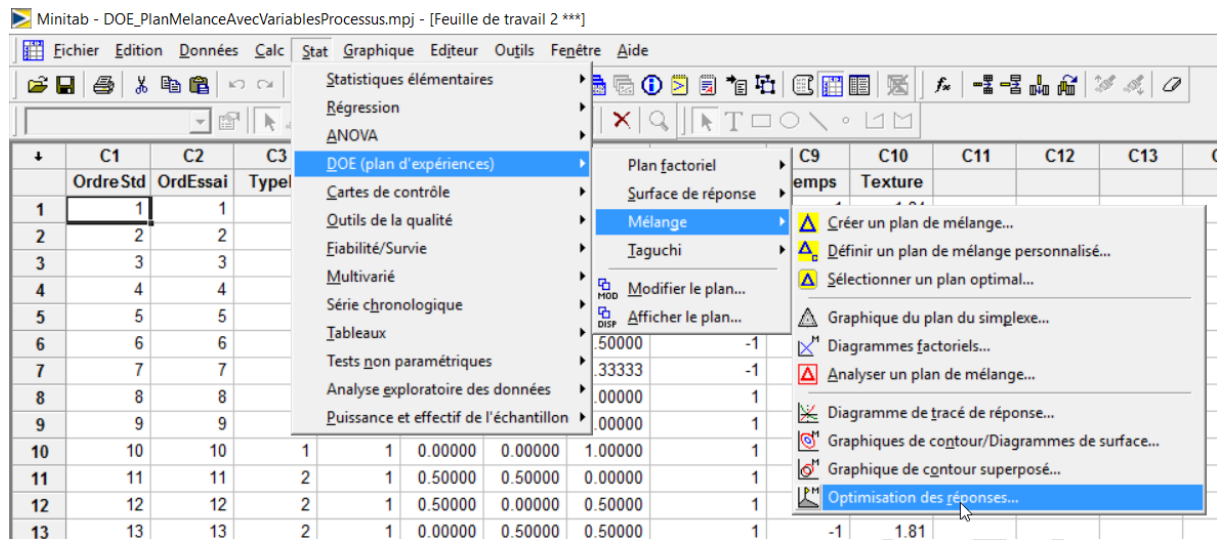
ANOVA Table					
Source of Variation	Degrees of Freedom	Sum of Squares [Partial]	Mean Squares [Partial]	F Ratio	P Value
Model	14	14.5066	1.0362	33.5558	6.8938E-08
Component Only					
Linear	2	4.1446	2.0723	67.1102	1.4088E-07
A • B	1	3.7216	3.7216	120.5208	6.0305E-08
A • C	1	0.1604	0.1604	5.1949	0.0402
B • C	1	1.1431	1.1431	37.0173	3.8782E-05
Component • Z1					
Z1 • A	1	1.1691	1.1691	37.8604	3.4705E-05
Z1 • B	1	0.1095	0.1095	3.5456	0.0823
Z1 • C	1	0.2290	0.2290	7.4172	0.0174
Z1 • A • C	1	0.2074	0.2074	6.7165	0.0224
Component • Z2					
Z2 • A	1	1.6845	1.6845	54.5517	5.3010E-06
Z2 • B	1	0.5059	0.5059	16.3819	0.0014
Z2 • C	1	0.5205	0.5205	16.8556	0.0012
Z2 • A • B	1	0.1782	0.1782	5.7698	0.0320
Z2 • B • C	1	0.1688	0.1688	5.4651	0.0360
Residual	13	0.4014	0.0309		
Lack of Fit	13	0.4014	0.0309		
Total	27	14.9080			

Car c'est beaucoup plus explicite!

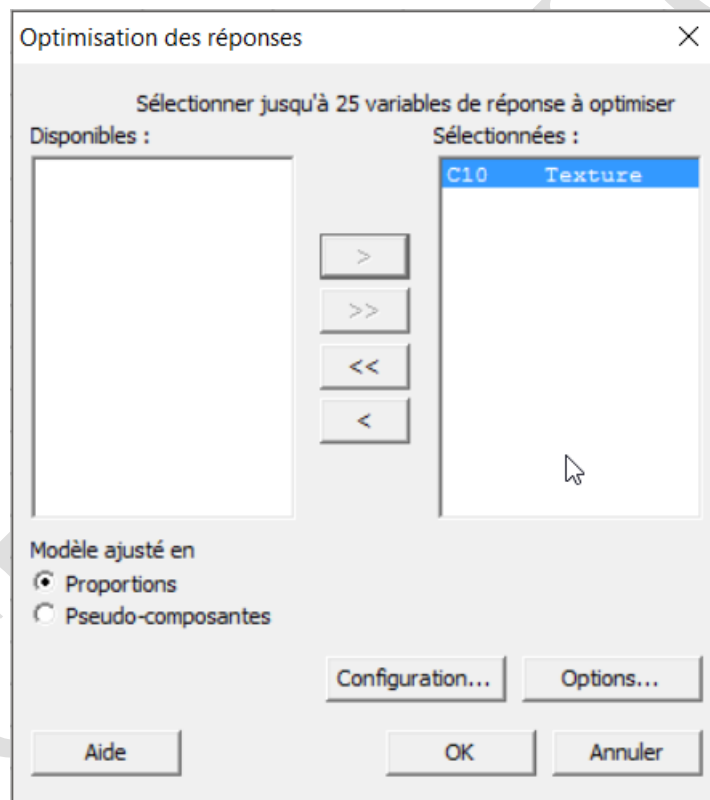
Nous remarquons aussi que les résultats commencent à sérieusement diverger en Minitab et DOE++.

Le tableau ci-dessus montre que les deux facteurs de processus ont des effets significatifs sur la texture des galettes. Comme le modèle est assez compliqué, les meilleurs paramètres pour les variables de processus et pour les composants ne peuvent pas être facilement identifiés.

Nous allons alors dans **DOE (plan d'expérience)/Mélange/Optimisation des réponses...**:



Pour prendre:



Nous cliquons sur **Configuration...** et comme nous voulons une cible de 3 pour la texture avec une étendue acceptable de [2.5,3.5], nous mettons:



Optimisation des réponses - Configuration

Réponse	Objectif	Inférieur	Cible	Supérieur	Pondération	Importance
C10 Texture	Cible	2.5	3	3.5	1	1

Fonctions de désirabilité pour différents objectifs - Comment les pondérations agissent sur leurs formes

Minimiser la réponse      Atteindre une valeur cible      Maximiser la réponse

Aide      OK      Annuler

Nous validons deux fois par **OK** pour obtenir:

Minitab - DOE\_PlanMelanceAvecVariablesProcessus.mpj - [Diagramme d'optimisation]

Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

Optimal	[ ]:Boeuf	[ ]:Porc	[ ]:Agneau	Températ	Temps
D	1.0	1.0	1.0	1.0	1.0
Max	[0.8724]	[0.0]	[0.1276]	[0.2008]	[0.4642]
1.0000	0.0	0.0	0.0	-1.0	-1.0
Act					
Min					

Composite Désirabilité 1.0000

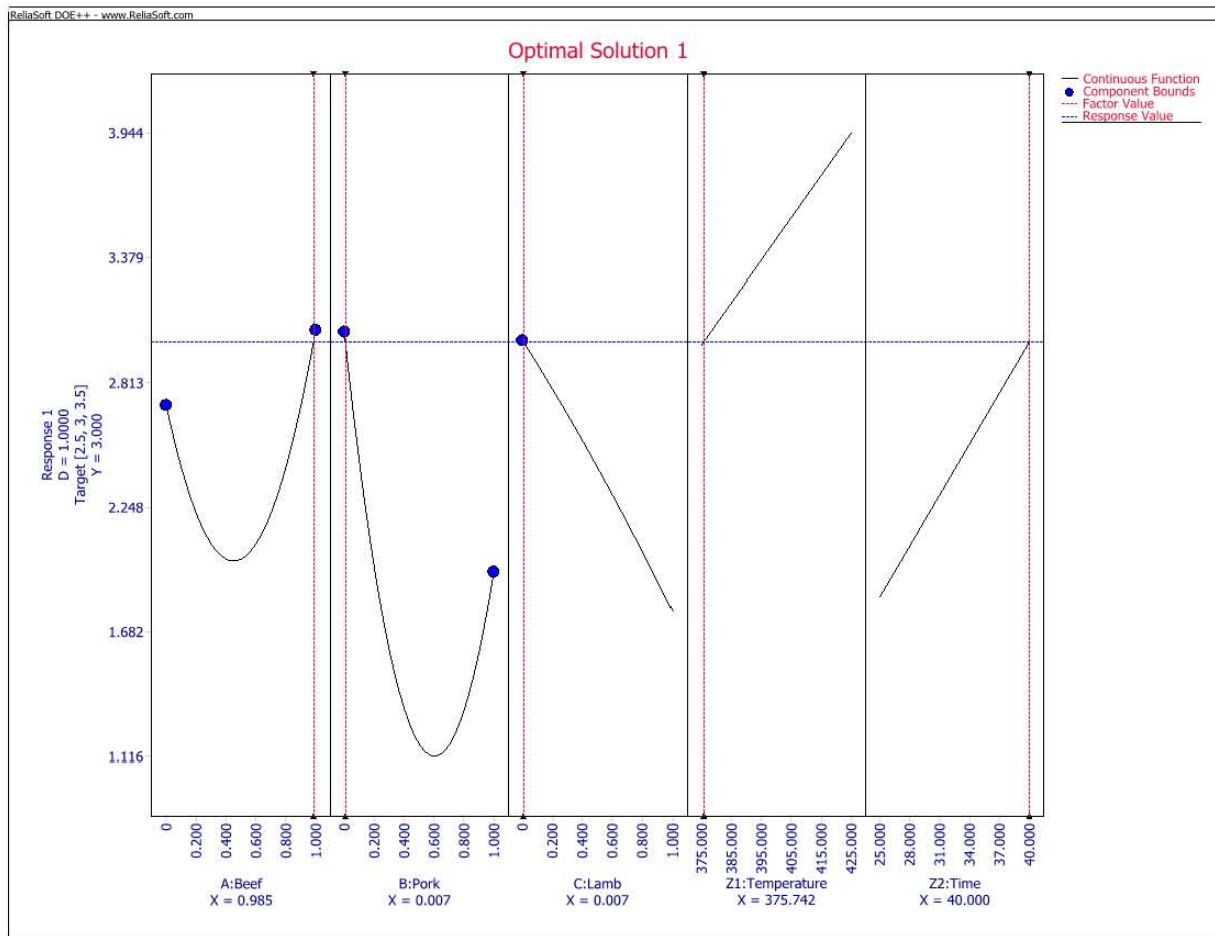
Texture Cible : 3.0  
y = 3.0  
d = 1.0000

C'est-à-dire:

- Bœuf: 87.24%
- Porc: 0%
- Agneau: 12.76%
- Température: 0.2008 c'est-à-dire 405.2 [°F]

- Temps: 0.4642 c'est-à-dire 35.918 [min]

À comparer avec DOE++....:



## 18. Fiabilité/Survie

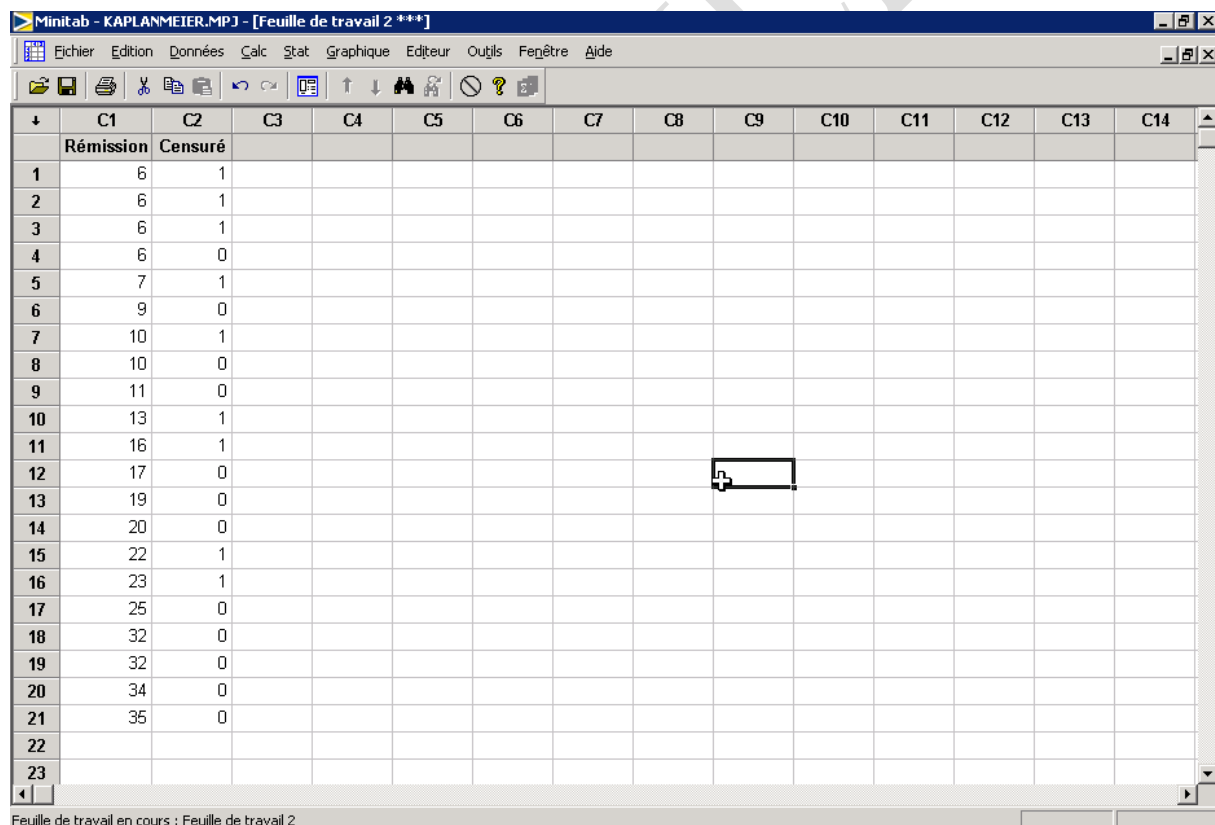
Le but dans ce chapitre va être bien évidemment de comparer les résultats avec ceux obtenus à la main mais surtout avec le logiciel Weibull++ qui reste un standard industriel incontournable dans le domaine!

### 18.1. Exercice 194.: Analyse de données censurées à droite selon modèle non paramétrique de Kaplan-Meier

Minitab® Statistical Software 15.1.1.0

Le but ici va être à nouveau de vérifier si nous obtenons bien les mêmes résultats que ceux calculés à la main dans le cours théorique lorsque nous avons étudié la démonstration de l'estimateur de survie de Kaplan-Meier.

Nous partons du tableau vu dans le cours théorique mais adapté aux exigences de Minitab (il n'est vraiment pas aisé de deviner que c'est sous cette forme que les choses doivent être représentées):



	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Rémission	Censuré												
1	6	1												
2	6	1												
3	6	1												
4	6	0												
5	7	1												
6	9	0												
7	10	1												
8	10	0												
9	11	0												
10	13	1												
11	16	1												
12	17	0												
13	19	0												
14	20	0												
15	22	1												
16	23	1												
17	25	0												
18	32	0												
19	32	0												
20	34	0												
21	35	0												
22														
23														

Ensuite, nous allons dans le menu **Stat/Fiabilité/Survie/Analyse de répartition troncage à droite/Analyse de répartition non paramétrique...**:

Effectuer une analyse de répartition non paramétrique avec troncage à droite/sans troncage

	C1	C2	C3	C7	C8	C9	C10	C11	C12	C13	C14
	Rémission	Censuré									
1	6	1									
2	6	1									
3	6	1									
4	6	0									
5	7										
6	9										
7	10										
8	10										
9	11										
10	13	1									
11	16	1									
12	17	0									
13	19	0									
14	20	0									
15	22	1									
16	23	1									
17	25	0									
18	32	0									
19	32	0									
20	34	0									
21	35	0									
22											
23											

Nous avons alors:

Analyse de répartition non paramétrique - Troncage à droite

Variables :

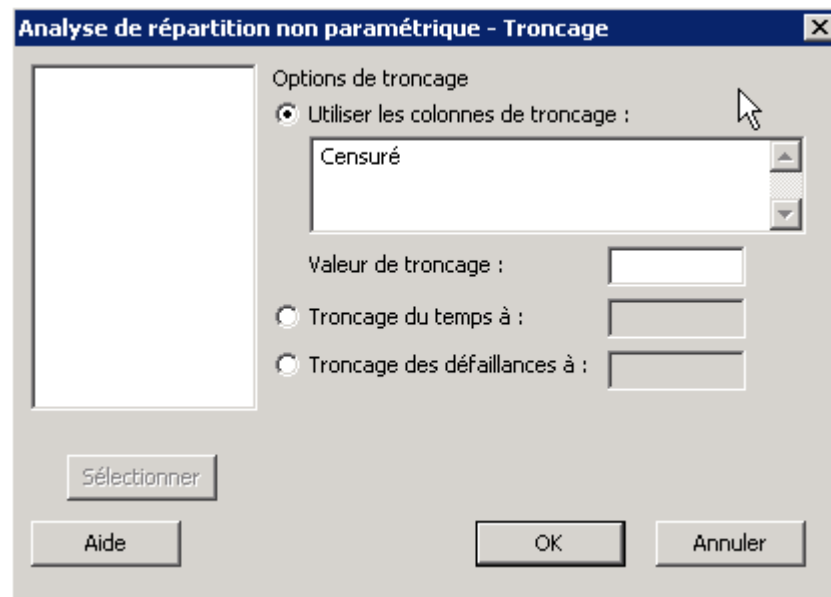
Rémission

Colonnes d'effectifs (facultatif) :

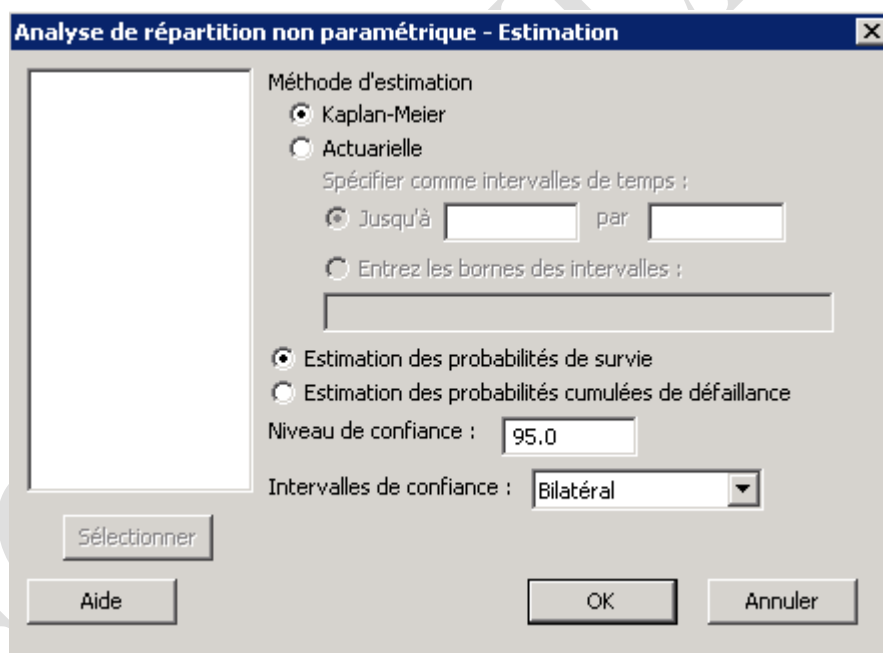
Variable de répartition :

Sélectionner Aide Tronquer... FMode... Estimation... Graphiques... Résultats... Stockage... OK Annuler

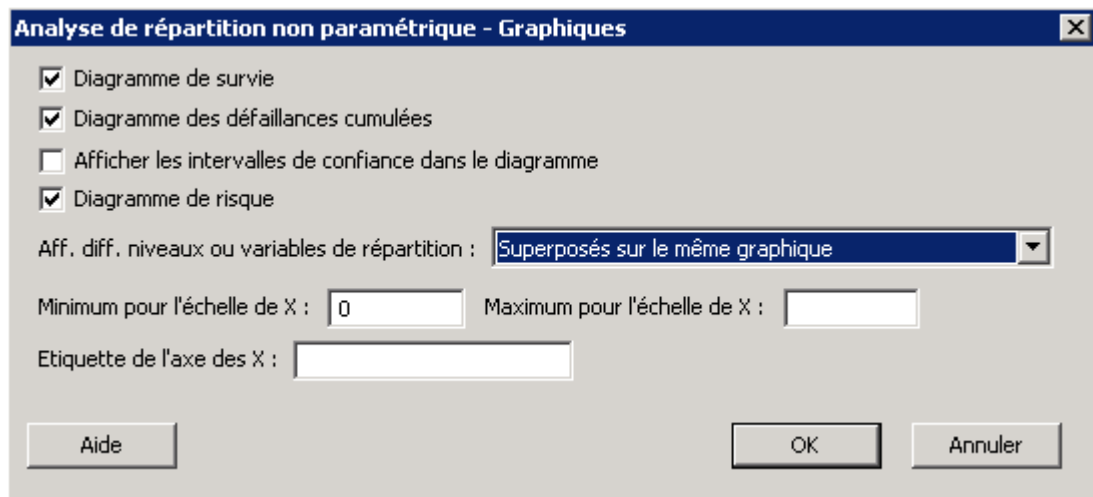
Nous cliquons sur le bouton **Tronquer...** pour y mettre:



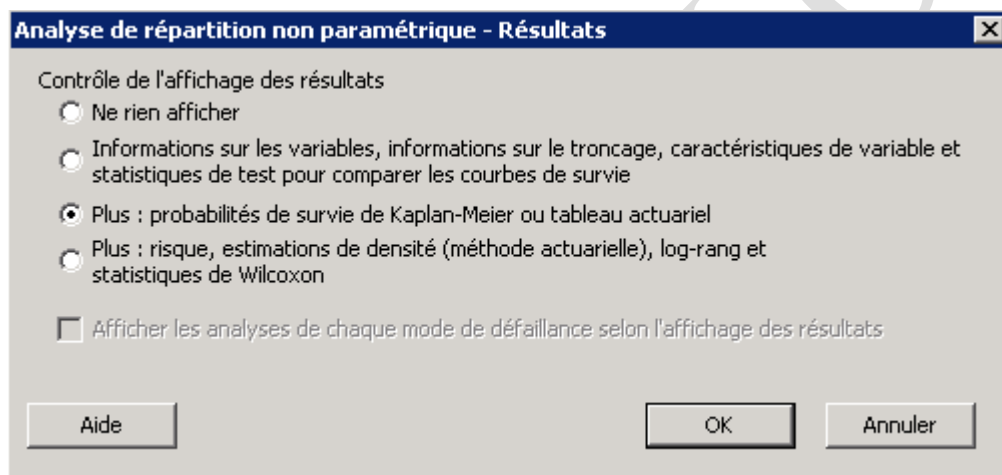
Dans le bouton **Estimation...** de la boîte de dialogue principale, nous prenons bien garde à être en analyse de Kaplan-Meier puisque c'est la seule que nous avons dans le cours théorique:



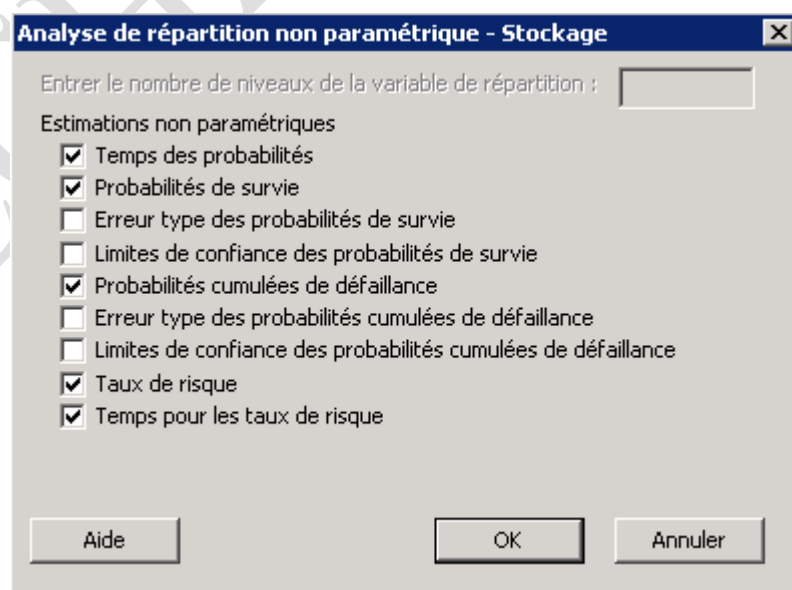
Dans le bouton **Graphique** nous prenons uniquement des options relativement à ce que nous avons démontré mathématiquement dans le cours théorique aussi:



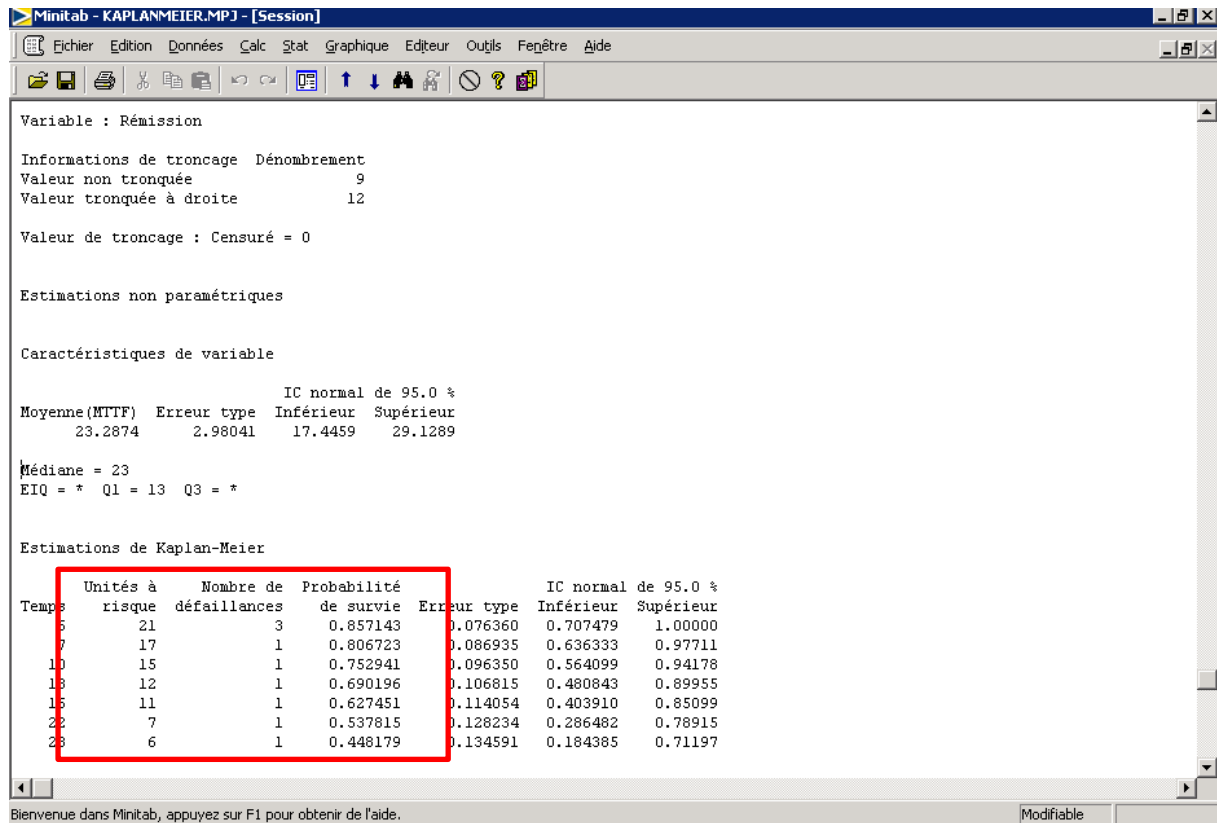
Dans le bouton **Résultats...** de la boîte de dialogue principale nous prenons aussi que des options relatives à ce qui a été démontré mathématiquement dans le cours théorique :



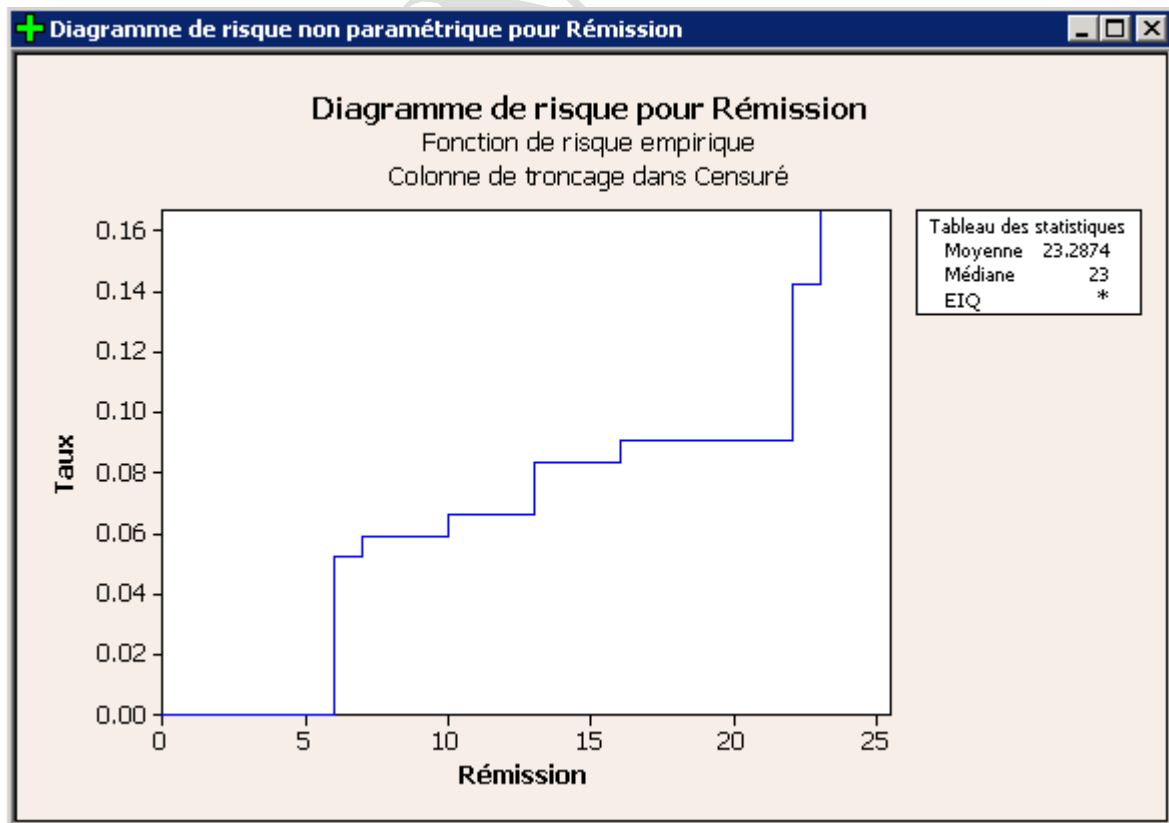
et dans le bouton **Stockage...** de la boîte de dialogue principale, nous prenons:

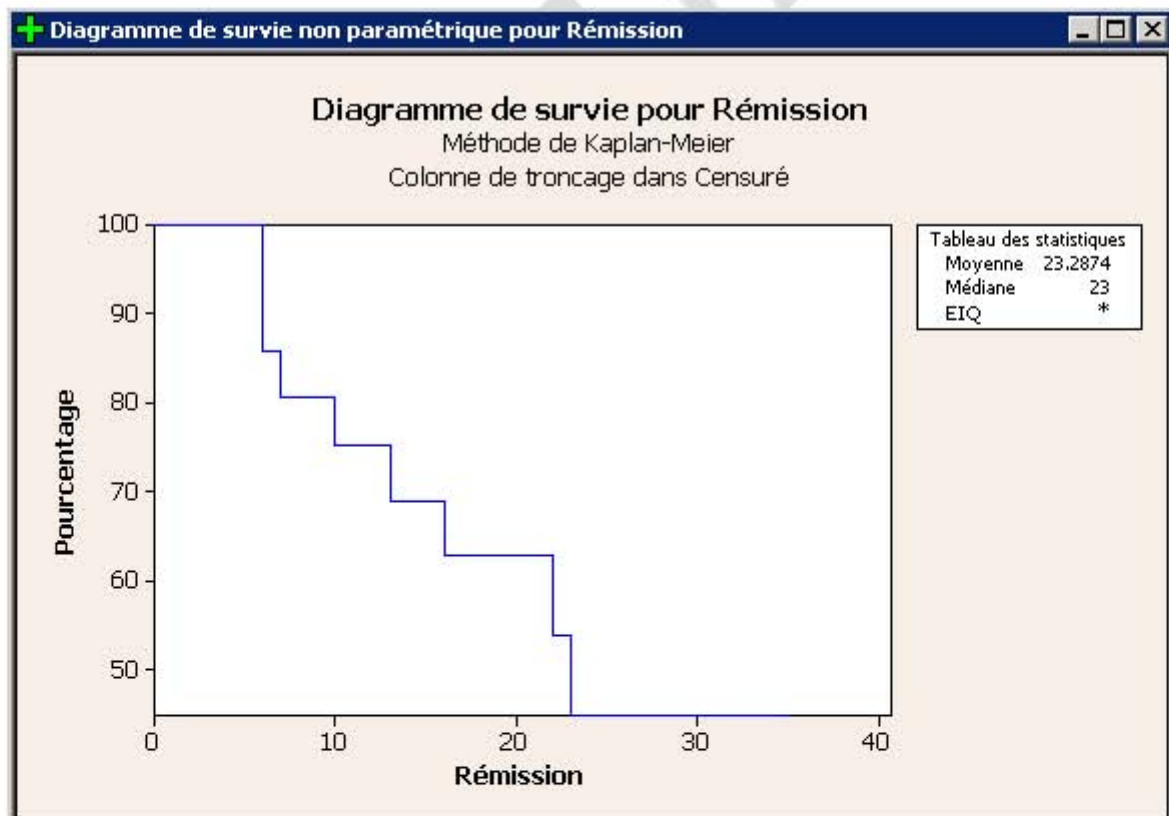
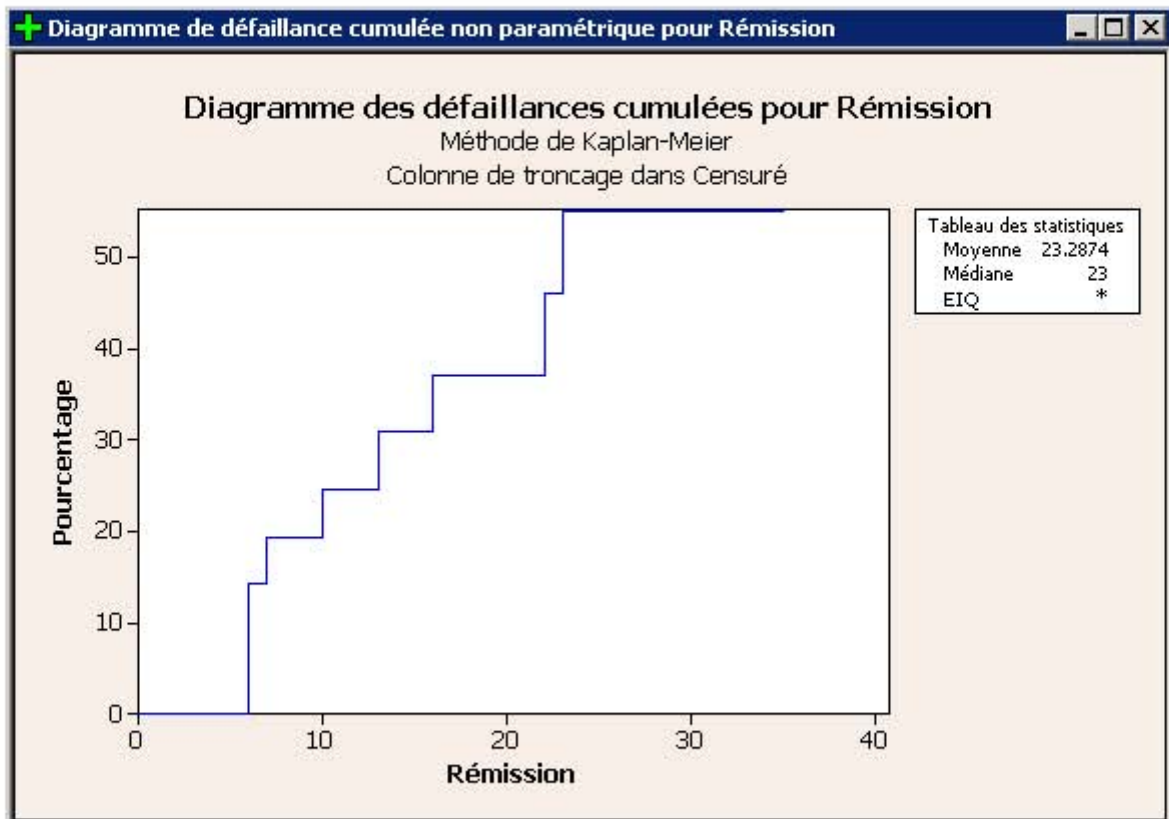


Nous validons tout cela par **OK** pour obtenir d'abord dans la fenêtre de session:



Nous retrouvons donc la même colonne de probabilité de survie que dans le cours théorique (nous n'avons pas calculé dans le cours théorique ce qui est en dehors du rectangle rouge). Et nous avons les trois graphiques suivants:





et donc nous retrouvons les trois mêmes graphiques que dans le cours théorique. Et dans le tableau, nous obtenons:



Minitab - KAPLANMEIER.MPJ - [Feuille de travail 2 \*\*\*]

Echier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
	Rémission	Censuré	TPS PROB1	PROB SURV1	PROB CUM1	TPS TXRISQ1	TX RISQUE1					
1	6	1	6	0.857143	0.142857	6	0.052632					
2	6	1	7	0.806723	0.193277	7	0.058824					
3	6	1	10	0.752941	0.247059	10	0.066667					
4	6	0	13	0.690196	0.309804	13	0.083333					
5	7	1	16	0.627451	0.372549	16	0.090909					
6	9	0	22	0.537815	0.462185	22	0.142857					
7	10	1	23	0.448179	0.551821	23	0.166667					
8	10	0										
9	11	0										
10	13	1										
11	16	1										
12	17	0										
13	19	0										
14	20	0										
15	22	1										
16	23	1										
17	25	0										
18	32	0										
19	32	0										
20	34	0										
21	35	0										
22												
23												

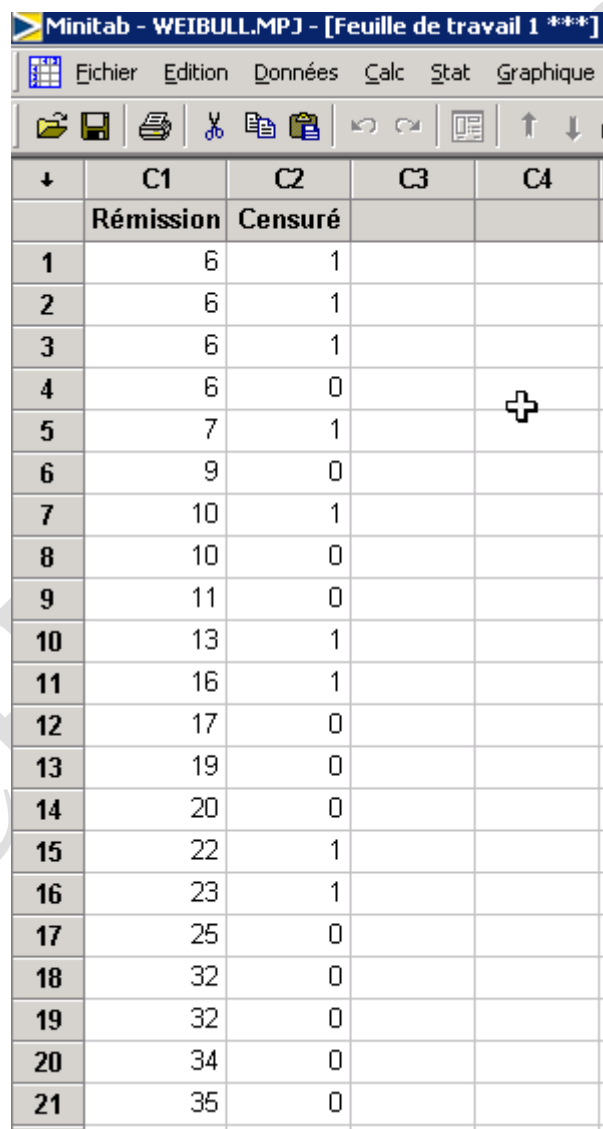
Feuille de travail en cours : Feuille de travail 2

## 18.2. Exercice 195.: Ajustement de données censurées à droite à une distribution de Weibull à deux paramètres

Minitab® Statistical Software 15.1.1.0

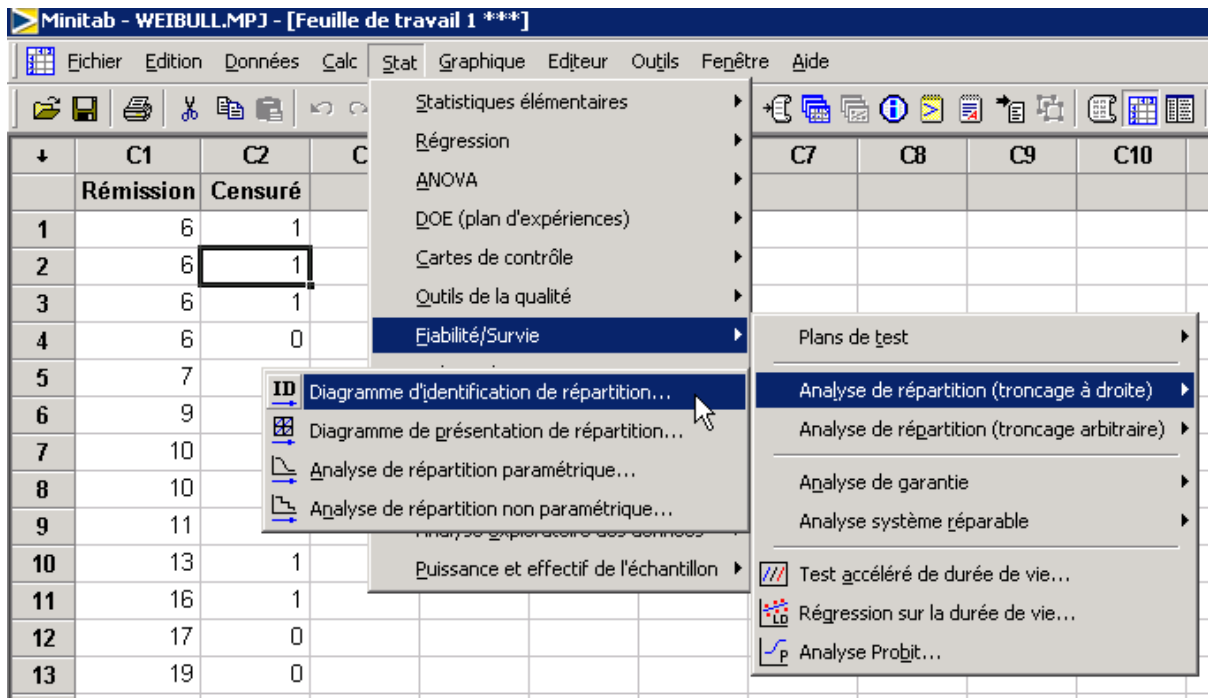
Nous allons ici vérifier si nous retrouvons les résultats des calculs faits dans Microsoft Excel suite à la démonstration faite dans le cours de Statistique théorique.

Nous partons donc des mêmes données que celles pour l'analyse de Kaplan-Meier puisque ce sont celles que nous avons aussi utilisées pour le calcul dans Microsoft Excel après avoir fait le calcul à la main de l'estimateur de maximum de vraisemblance de la loi de Weibull à deux paramètres avec données censurées à droite:

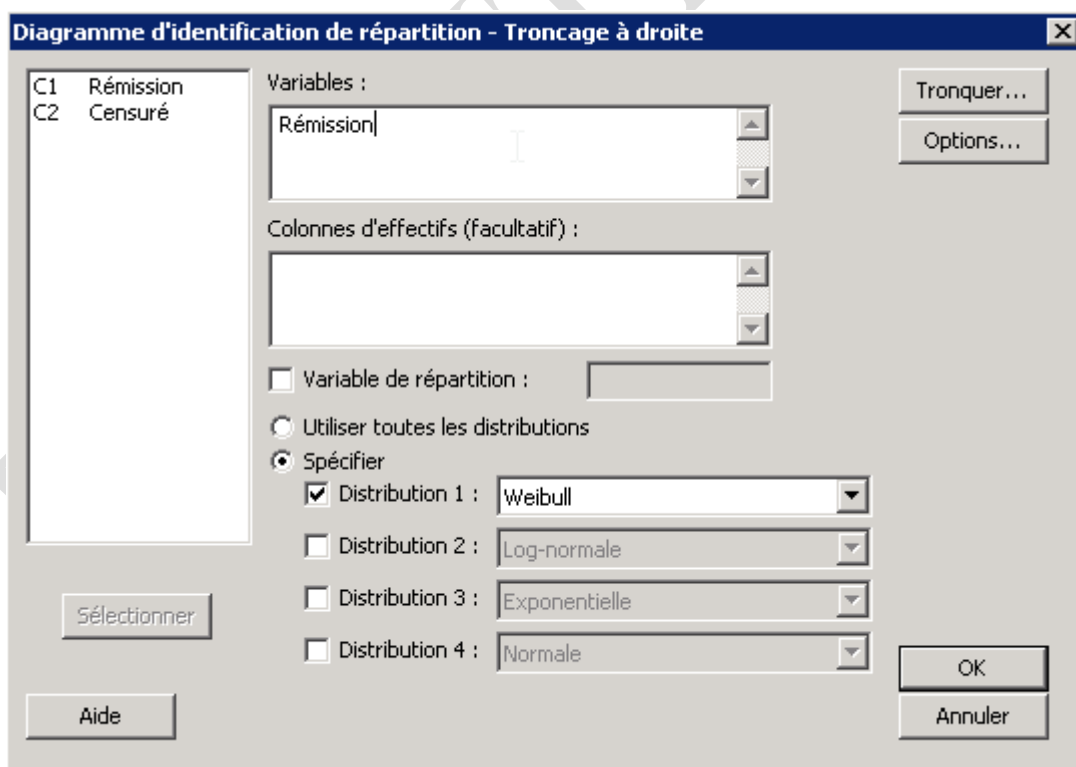


	C1	C2	C3	C4
	Rémission	Censuré		
1	6	1		
2	6	1		
3	6	1		
4	6	0		
5	7	1		
6	9	0		
7	10	1		
8	10	0		
9	11	0		
10	13	1		
11	16	1		
12	17	0		
13	19	0		
14	20	0		
15	22	1		
16	23	1		
17	25	0		
18	32	0		
19	32	0		
20	34	0		
21	35	0		

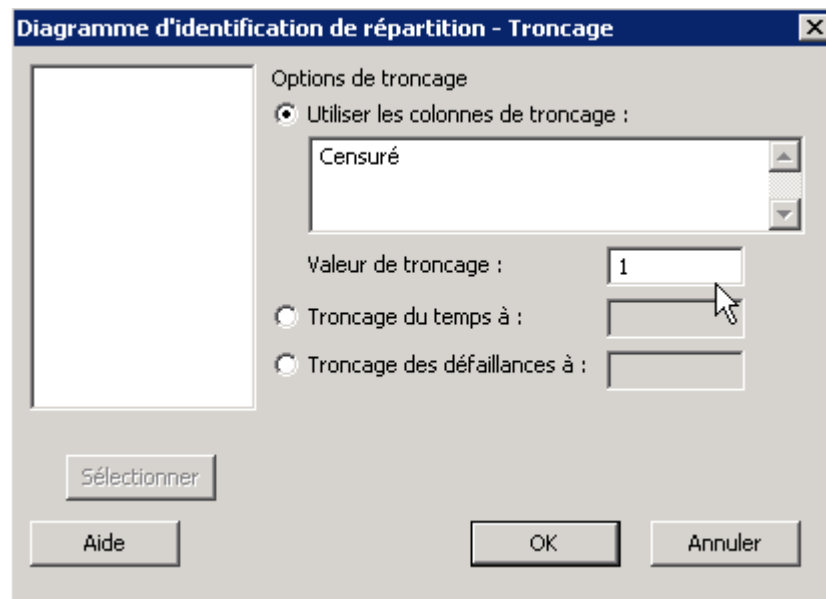
Ensuite, nous allons dans le menu **Stat/Fiabilité/Survie/Analyse de répartition (troncage à droite)/Diagramme d'identification de répartition**:



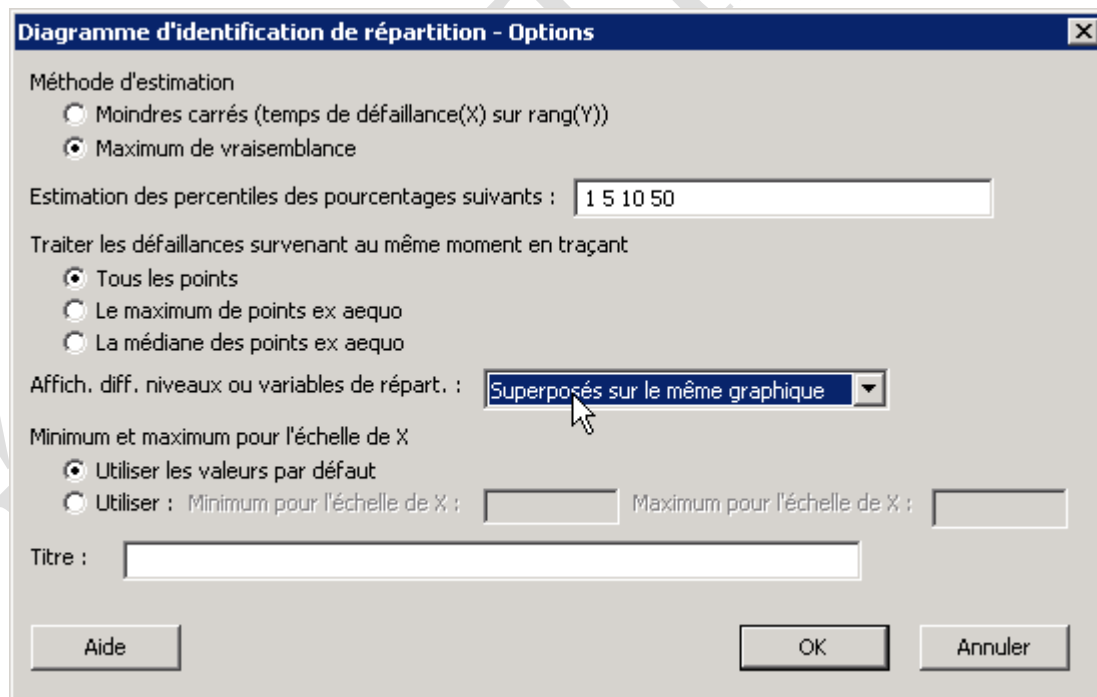
Dans la boîte de dialogue qui apparaît nous mettons:



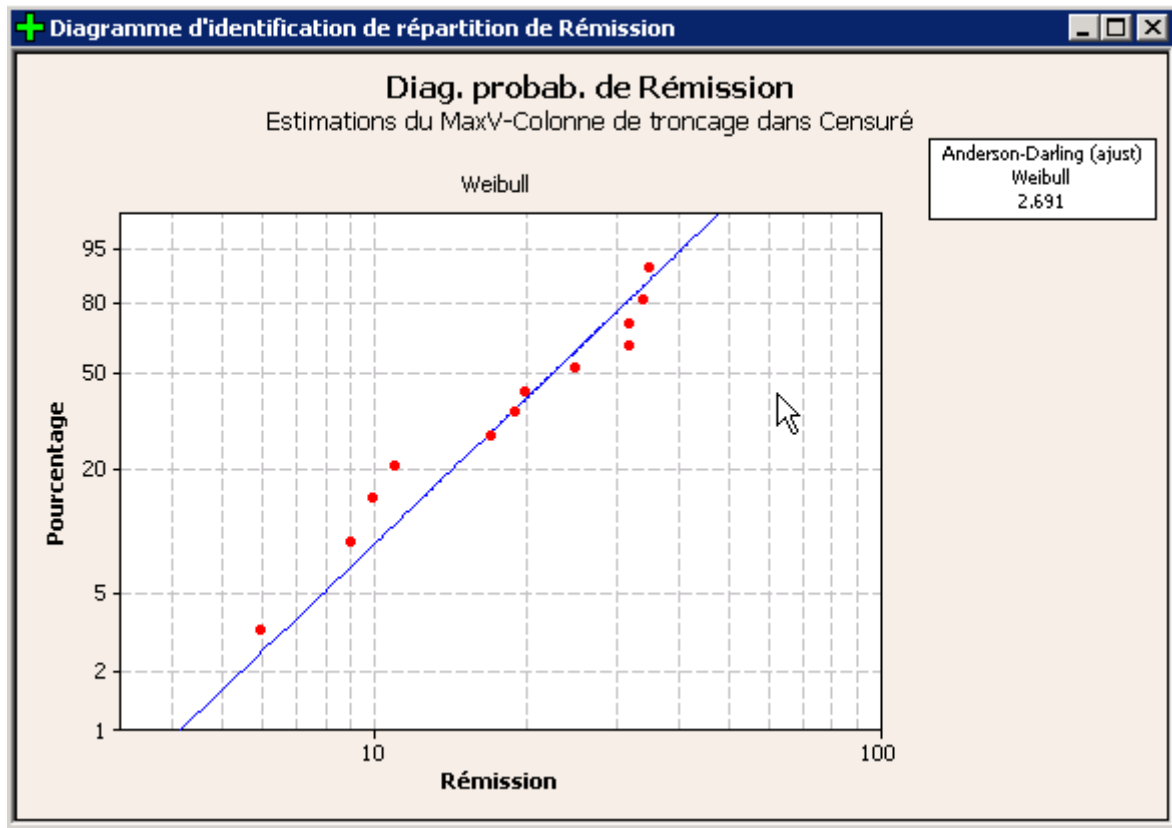
Nous cliquons sur le bouton **Tronquer...** pour y mettre dans la boîte de dialogue qui apparaît les paramètres suivants:



Nous validons par **OK** et cliquons sur le bouton **Options...** pour faire bien attention à prendre l'option **Maximum de vraisemblance** qui correspond à ce que nous avons démontré dans le cours théorique:



Nous validons le tout deux fois par **OK** pour obtenir d'abord un graphique d'ajustement qui est un peu léger dans la version 15 de Minitab (mais qui a été considérablement amélioré dans la version 17):



et dans la fenêtre de sessions, nous avons une version tabulée de la distribution ajustée:

**Diagramme d'identification de répartition : Rémission**

Adéquation de l'ajustement

Loi de distribution Anderson-Darling (ajust)  
Weibull 2.691

Tableau des percentiles

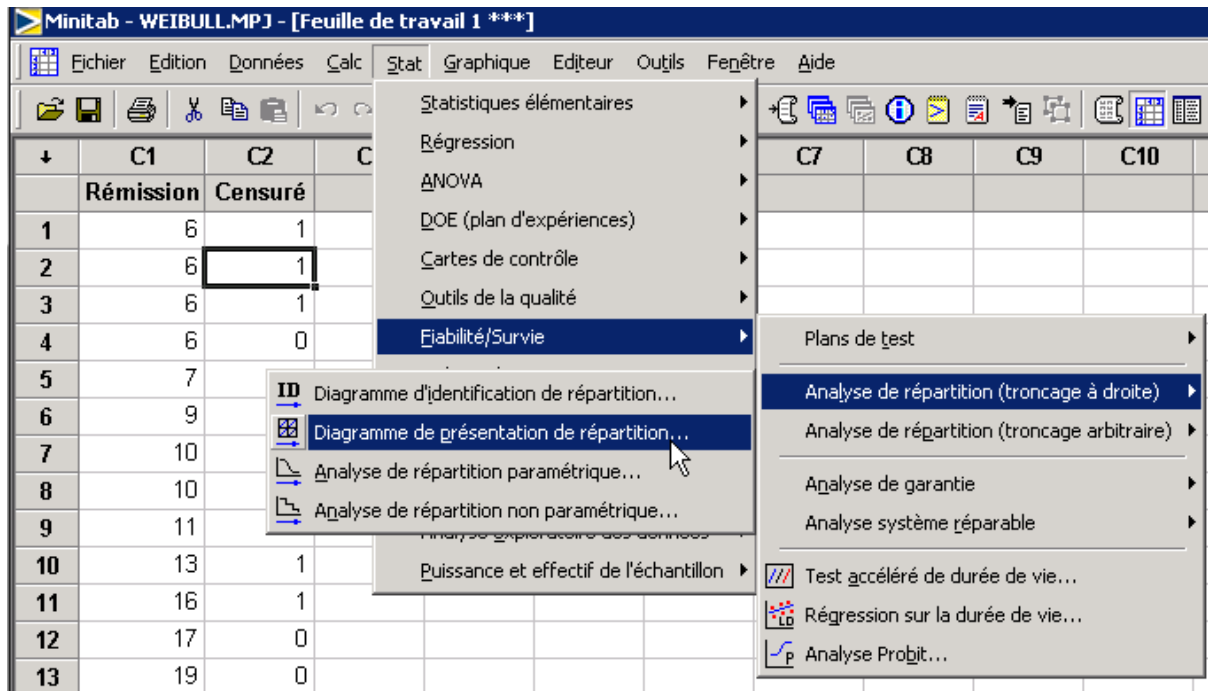
Loi de distribution	Pourcentage	Percentile	Erreur type	IC normal de 95 %	
				Inférieur	Supérieur
Weibull	1	4.16011	1.76713	1.80939	9.56483
Weibull	5	7.96617	2.31172	4.51064	14.0689
Weibull	10	10.6133	2.48093	6.71241	16.7813
Weibull	50	22.4878	2.73051	17.7253	28.5300

Tableau des durées moyennes avant défaillance (MTTF)

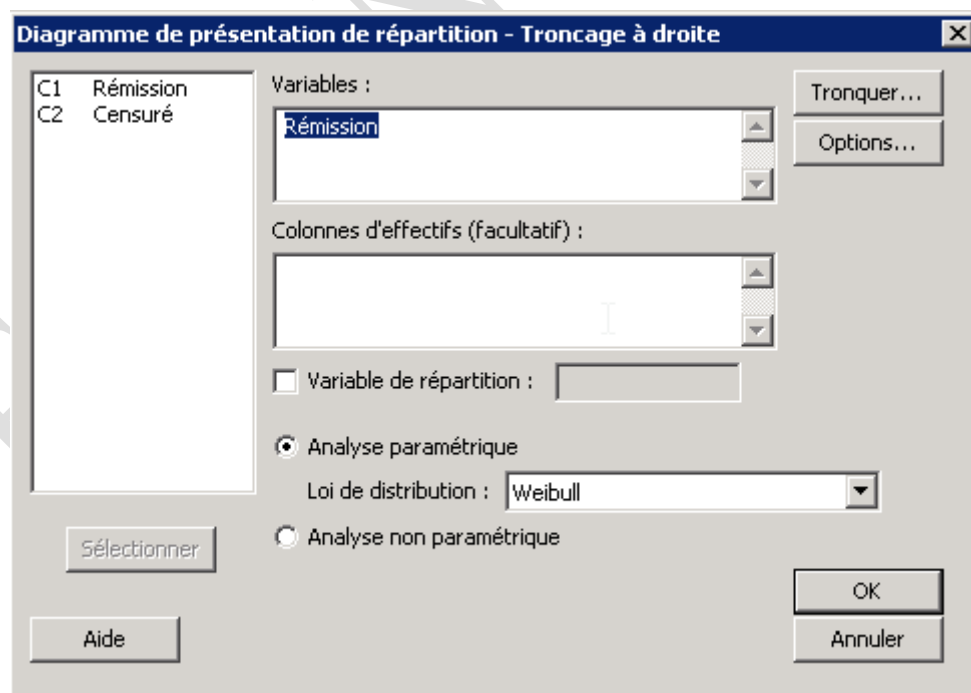
Loi de distribution	Moyenne	Erreur type	IC normal de 95 %	
			Inférieur	Supérieur
Weibull	23.0931	2.67208	18.4073	28.9717

Nous pouvons obtenir mieux toujours avec les mêmes données et en utilisant les mêmes résultats théoriques.

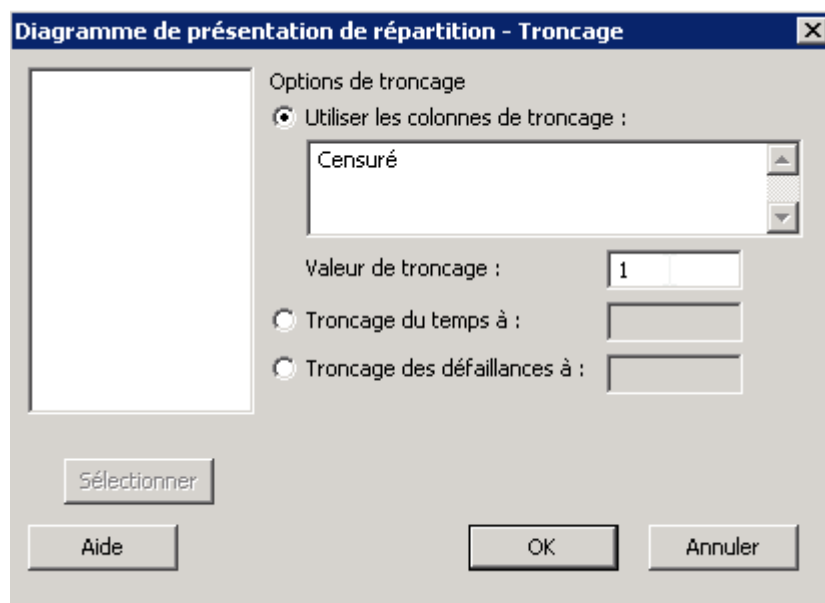
Nous allons dans le menu **Stat/Fiabilité/Survie/Analyse de répartition (troncage à droite)/Diagramme de présentation de répartition**:



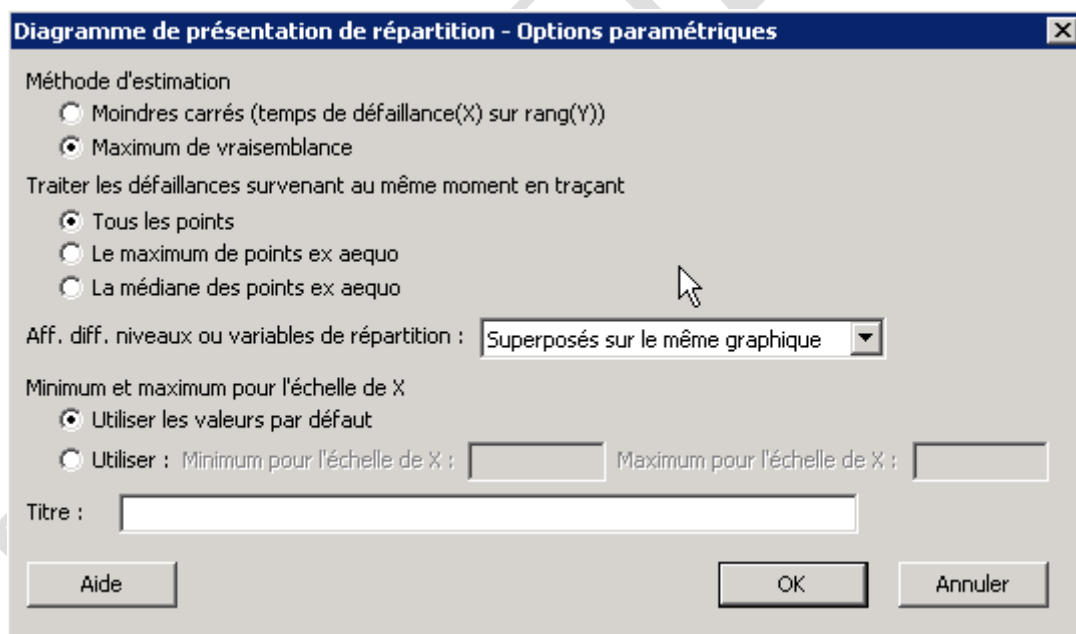
Apparaît alors une boîte de dialogue quasi similaire à l'exemple précédent où nous mettons:



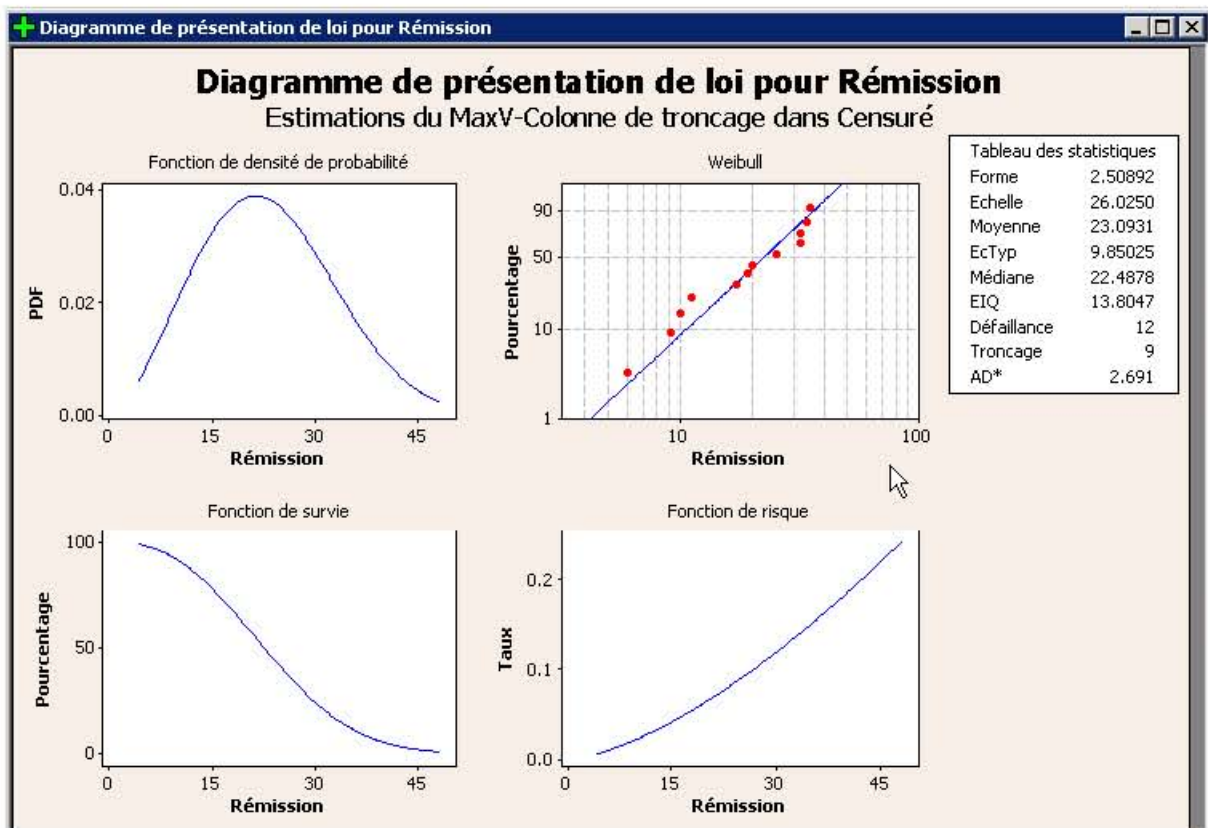
Nous cliquons sur **Tronquer...** pour y mettre:



Et dans le bouton **Options...** nous faisons bien attention à activer l'option Maximum de vraisemblance pour encore une fois être en correspondance avec ce que nous avons vu dans le cours théorique:



Ce qui nous donne:





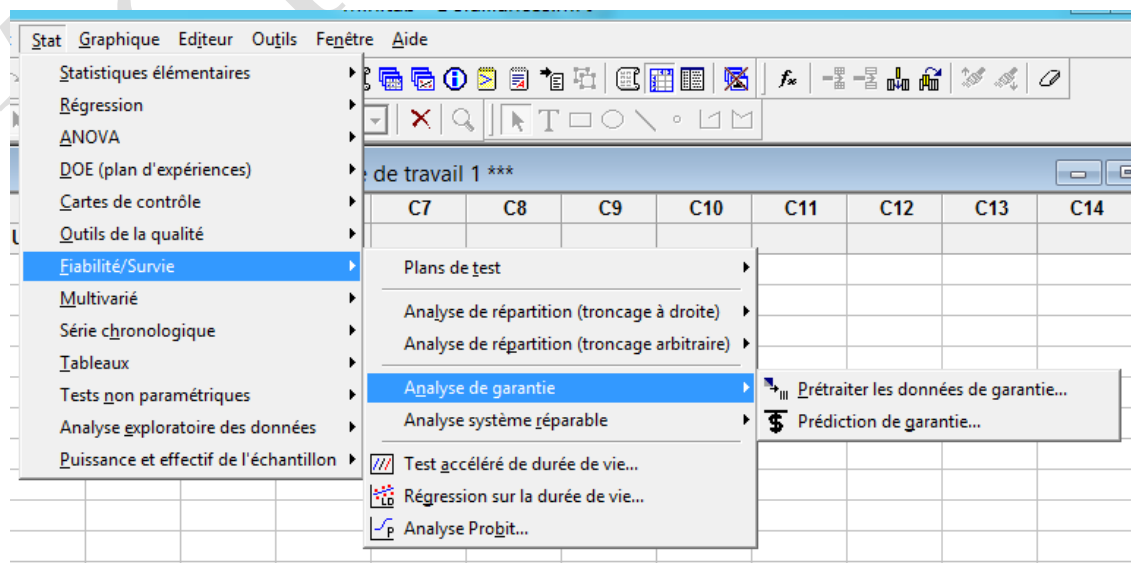
### 18.3. Exercice 196.: Prédiction de garantie basé sur le maximum de vraisemblance d'une loi exponentielle

Minitab® Statistical Software 15.1.1.0

Le but ici est de vérifier les calculs de maximum de vraisemblance de fiabilités démontrés dans le cours théorique et calculés à la main. Pour voir si Minitab donne la même chose, nous entrons les valeurs suivantes dans une feuille:

	C1	C2	C3	C4	C5	C6	C7	C8	C9
	Début	Fin	Unités						
1	0	700	1						
2	700	800	1						
3	800	900	1						
4	900	950	1						
5	1000	*	6						
6									
7									
8									
9									
10									
11									

Ensuite, nous allons dans le menu **Stat/Fiabilité/Survie/Analyse de garantie/Prédiction de garantie...**:



et nous y mettons (dans le cours de Génie Industriel nous avons fait le choix d'une loi exponentielle donc attention à bien prendre celle-ci!):

Prévision de garantie

C1 Début  
C2 Fin  
C3 Unités

Données de garantie transformées

Heure de début : Début

Heure de fin : Fin

Effectif (facultatif) : Unités

Loi de distribution supposée : Exponentielle

Hypothèses de garantie

Durée de garantie : (facultatif)

Coût moyen par défaillance : (facultatif)

Sélectionner

Aide

Prévision...  
Graphiques...  
Options...  
OK  
Annuler

Et nous devons dire à Minitab d'utiliser la technique de maximum de vraisemblance utilisée dans le cours théorique :

Prévision de garantie - Options

Estimations de distribution

Calculer les estimations optimales

Entrer les estimations définies par l'utilisateur :

Méthode d'estimation

Moindres carrés (moment de défaillance (X) sur rang (Y))

Maximum de vraisemblance

Sélectionner

Aide

OK  
Annuler

Nous validons par **OK** autant de fois que nécessaire pour obtenir:

Loi : exponentielle avec moyenne = 2213,93  
Méthode d'estimation : maximum de vraisemblance

Récapitulatif des plaintes actuelles relatives à la garantie

Nombre total d'unités	10
Nombre de défaillances observé	4
Nombre de défaillances attendu	2,52958
IC de Poisson 95 %	(0,426867, 8,05192)

Nombre d'unités à risque pour les périodes futures 6

À comparer avec la moyenne obtenue dans le cours théorique:

$$\mu = \frac{1}{\lambda} = 2337.5$$

Donc il y a une différence significative entre le logiciel et notre calcul à la main ce qui est bizarre et que je ne m'explique pas puisque faire ce type de calcul est trivial. Il est possible que la différence provienne de la manière d'entrer les données dans la feuille Minitab qui n'est pas très bien documentée.

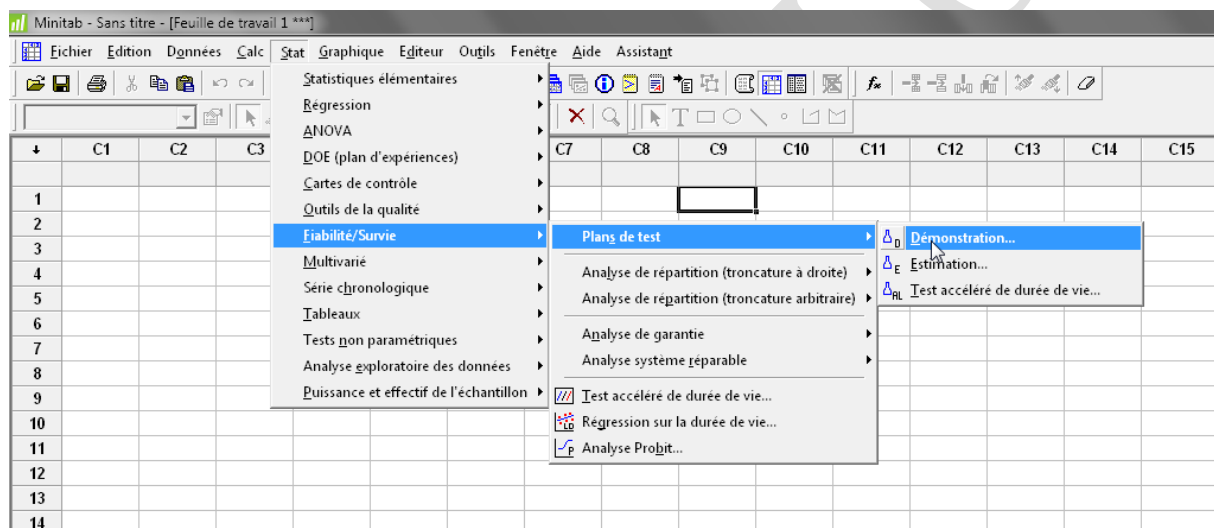
## 18.4. Exercice 197.: Plan de démonstration modèle cumulatif binomial pour taille d'échantillon nécessaire avec temps de test contraint

Minitab® Statistical Software 15.1.1.0

Nous allons ici vérifier si nous retrouvons les résultats des calculs faits dans Microsoft Excel suite à la démonstration faite dans le cours de Génie Industriel.

Nous voulons démontrer une fiabilité de 90% pendant un temps de 100 heures avec un seuil de confiance de 95% sans aucune défaillance. Nous assumons que la fiabilité suit une loi de Weibull de paramètre de forme  $\beta = 1.5$ . Le problème étant que nous n'avons que 48 heures à disposition, quelle quantité de produits devons-nous tester?

Dans Minitab, nous allons dans le menu **Fiabilité/Survie/Plans de test/Démonstration...**:



Dans la boîte de dialogue où "Nombre de tests pour chaque unité" est très mal traduit de l'anglais... nous mettons:

**Plans de test de démonstration**

Valeur minimale à démontrer

Echelle (Weibull ou expo) ou emplacement (autres lois de distribution) :

Percentile : Pourcentage :

Fiabilité : 0,9 Durée : 100

Durée moyenne avant défaillance (MTTF) :

Nombre maximal de défaillances autorisées : 0

Effectifs des échantillons :

Nombre de tests pour chaque unité : 48

Hypothèses pour la loi

Loi de distribution : Weibull

Forme (Weibull) ou échelle (autres lois de distribution) : 1.5

Aide OK Annuler

Dans le bouton Options... nous mettons le niveau de confiance:

**Plans de test de démonstration - Options**

Niveau de confiance : 95

Aide OK Annuler

Ce qui nous donne dans la fenêtre de session:

### Plans de test de démonstration

Plan de test de fiabilité  
 Loi : Weibull, Forme = 1.5  
 Objectif de fiabilité = 0.9, Niveau de confiance cible = 95 %

Test de défaillance	Temps de test	Effectif d'échantillon	Niveau de confiance réel
0	48	86	95.0869

Nous retrouvons donc bien le résultat des calculs effectués à la main ainsi qu'avec Weibull++ comme l'atteste la capture d'écran suivante:

Weibull DRT Reliability

### Design a reliability demonstration test

**What metric would you like to demonstrate?**

Metric	Reliability value at a specific time
Demonstrate this reliability (%)	90
With this confidence level (%)	95
At this time (Hr)	100

**Assume the failure rate behavior is governed by this distribution**

Distribution	With this Beta
2P-Weibull	1.5

**Solve for this value**

Value	Required sample size
With this test time (Hr)	48
With a maximum of this many failures	0

### Results

Sample size	<b>85.4994</b>
-------------	----------------

**Notes**  
This is based on both the assumed failure rate behavior given the specified distribution and the specified acceleration factor.

RDT 1

---

### Reliability Demonstration Test

#### TEST DESIGN

**Test Design Method**  
Parametric Binomial

**Input**  
Units: Hour (Hr)  
Acceleration Factor: 1

**Display Options**  
 Show sample size as integer

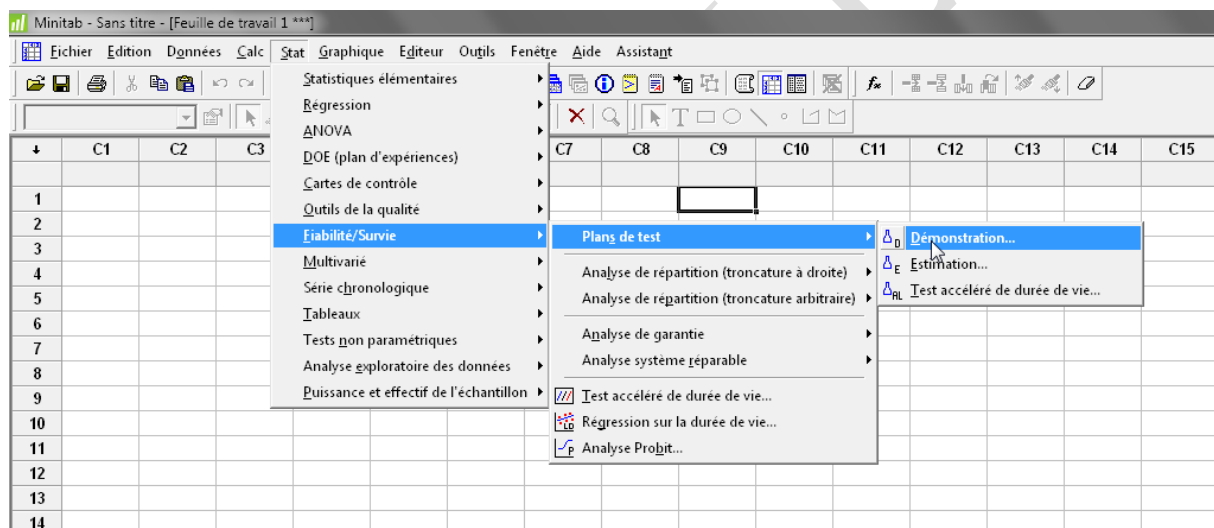
## 18.5. Exercice 198.: Plan de démonstration modèle cumulatif binomial pour temps cumulé de test nécessaire avec taille d'échantillon connue

Minitab® Statistical Software 15.1.1.0

L'idée ici est bien évidemment comme toujours de vérifier les calculs faits à la main suite aux démonstrations théoriques.

Nous allons reprendre l'exemple précédent avec les mêmes paramètres à la différence que cette fois-ci le temps disponible (nécessaire) ne nous est pas connu mais que nous connaissons la taille de l'échantillon qui nous a été mis à disposition. Donc nous cherchons alors le temps de test cumulé nécessaire pour la plan de test de démonstration avec un échantillon de taille 20 (évidemment vu que nous reprenons les données du cas précédent, le résultat devra être très supérieur à 48h.!).

Dans Minitab, nous allons dans le menu **Fiabilité/Survie/Plans de test/Démonstration...**:



Dans la boîte de dialogue qui apparaît, nous mettons:

**Plans de test de démonstration**

Valeur minimale à démontrer

Echelle (Weibull ou expo) ou emplacement (autres lois de distribution) :

Percentile :  Pourcentage :

Fiabilité :  Durée :

Durée moyenne avant défaillance (MTTF) :

Nombre maximal de défaillances autorisées :

Effectifs des échantillons :

Nombre de tests pour chaque unité :

Hypothèses pour la loi

Loi de distribution :

Forme (Weibull) ou échelle (autres lois de distribution) :

Aide

Graphiques...

Options...

OK

Annuler

En oubliant pas dans **Options...** d'avoir:

**Plans de test de démonstration - Options**

Niveau de confiance :

Aide

OK

Annuler

Ce qui nous donne dans la fenêtre de session:

```

Plans de test de démonstration
|
Plan de test de fiabilité
Loi : Weibull, Forme = 1.5
Objectif de fiabilité = 0.9, Niveau de confiance réel = 95 %

      Test de      Effectif      Temps
défaillance d'échantillon de test
          0             20  126.434

```

Nous retrouvons donc bien le résultat des calculs effectués à la main ainsi qu'avec Weibull++ comme l'atteste la capture d'écran suivante:



Weibull DRT Reliability
Weibull DRT MTF

### Design a reliability demonstration test

**What metric would you like to demonstrate?**

Metric	Reliability value at a specific time
Demonstrate this reliability (%)	90
With this confidence level (%)	95
At this time (Hr)	100

**Assume the failure rate behavior is governed by this distribution**

Distribution	2P-Weibull
With this Beta	1.5

**Solve for this value**

Value	Required test time
With this sample size	20
With a maximum of this many failures	0

**Results**

Test time per unit (Hr)	<b>126.4339</b>
-------------------------	-----------------

**Notes**

This is based on both the assumed failure rate behavior given the specified distribution and the specified acceleration factor.

Reliability Demonstration Test

#### TEST DESIGN

**Test Design Method**

Parametric Binomial

**Input**

Units: Hour (Hr)

Acceleration Factor: 1

**Display Options**

Show sample size as integer

RDT 1

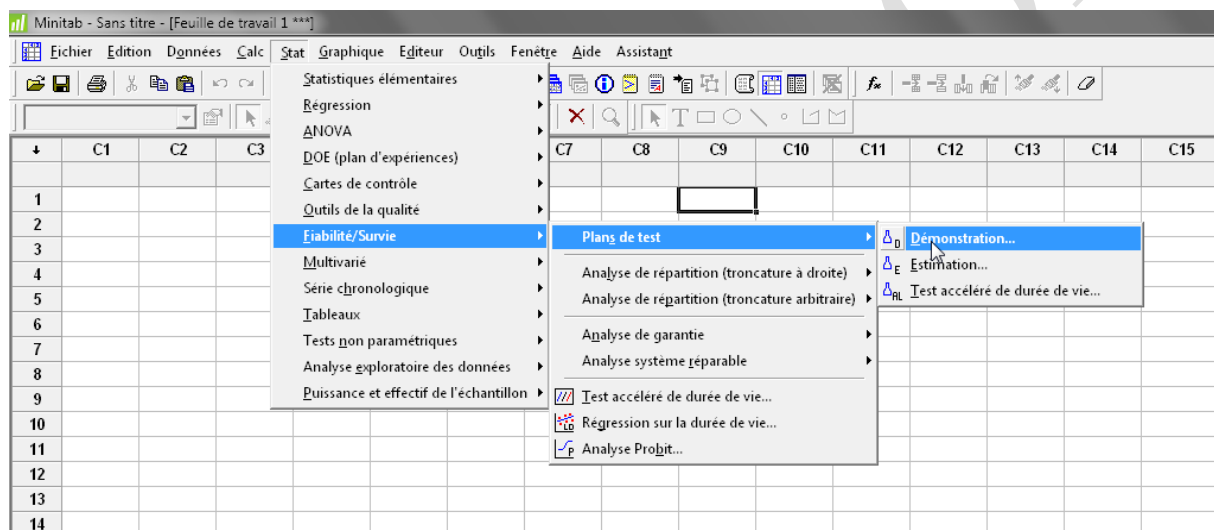
## 18.6. Exercice 199.: Plan de démonstration modèle exponentiel-chi temps cumulé de test nécessaire

Minitab® Statistical Software 15.1.1.0

Le but ici va être comme à l'habitude... de vérifier si nous retrouvons la même valeur avec Minitab que celle calculée à la main après la démonstration mathématique du temps maximum de temps de test pour un seuil de confiance donné.

Donc comme dans le cours théorique, nous allons considérer le cas où nous voulons savoir combien de temps nous devons tester 1 élément pour prouver à un niveau de confiance de 90% que sa MTTF est de 100 (minutes, heures ou mois peu importe l'unité).

Dans Minitab, nous allons dans le menu **Fiabilité/Survie/Plans de test/Démonstration...**:



où nous prenons:

Plans de test de démonstration

Valeur minimale à démontrer

Echelle (Weibull ou expo) ou emplacement (autres lois de distribution) :

Percentile :  Pourcentage :

Fiabilité :  Durée :

Durée moyenne avant défaillance (MTTF) :

Nombre maximal de défaillances autorisées :

Effectifs des échantillons :

Durée des tests pour chaque unité :

Hypothèses pour la loi

Loi de distribution :

Forme (Weibull) ou échelle (autres lois de distribution) :

Aide

Graphiques...

Options...

OK

Annuler

Remarque qu'il est bizarre que Minitab demande l'effectif des échantillons sachant que le logiciel Weibull++ et le modèle théorique que nous avons démontré dans le cours théorique ne nécessitent pas cette information mais nous allons voir cela plus loin.

Dans **Options**, nous mettons:

Plans de test de démonstration - Options

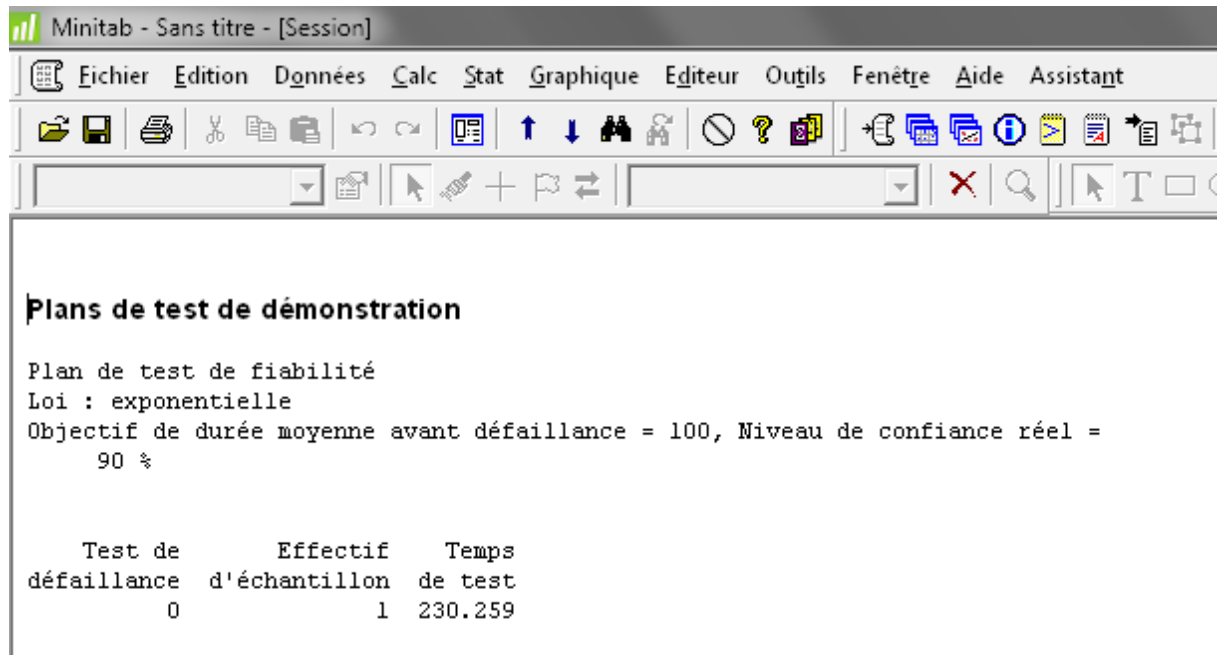
Niveau de confiance :

Aide

OK

Annuler

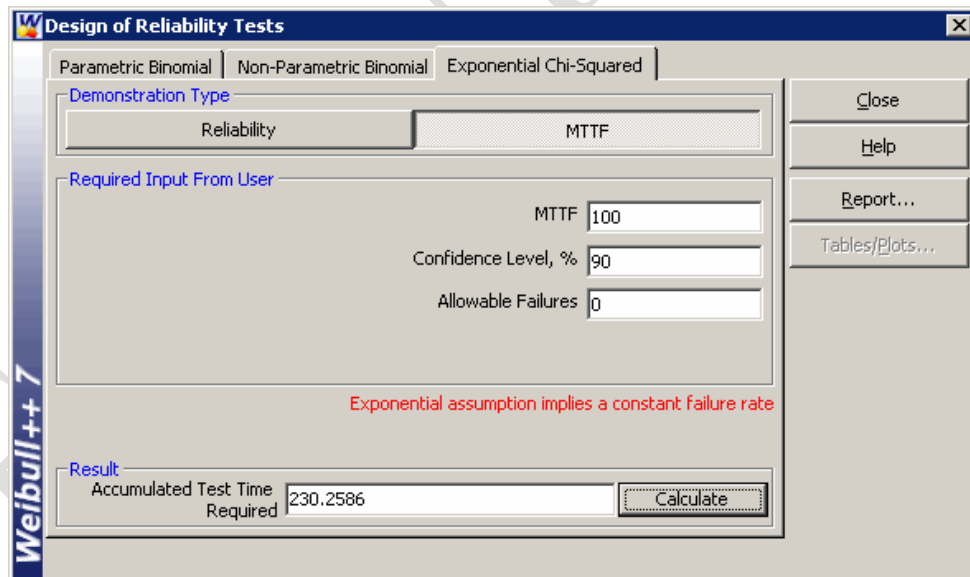
Nous validons le tout par **OK** pour obtenir dans la fenêtre de session:



Ce qui correspond bien à au calcul à la main effectué dans le cours théorique:

$$T = \frac{MTTF \cdot \chi^2_{CL,2(r+1)}}{2} = \frac{MTTF \cdot \chi^2_{0.9,2(0+1)}}{2} = 100 \cdot \text{CHISQ.INV}(0.9;2)/2 = 230.2585$$

et qui correspond bien aussi à la sortie du logiciel Weibull++:



Maintenant, si nous mettons (nous multiplions par 10 la taille de l'échantillon test):

Plans de test de démonstration

Valeur minimale à démontrer

Echelle (Weibull ou expo) ou emplacement (autres lois de distribution) :

Percentile :  Pourcentage :

Fiabilité :  Durée :

Durée moyenne avant défaillance (MTTF) :

Nombre maximal de défaillances autorisées :

Effectifs des échantillons :

Durée des tests pour chaque unité :

Hypothèses pour la loi

Loi de distribution :

Forme (Weibull) ou échelle (autres lois de distribution) :

Aide

Graphiques...

Options...

OK

Annuler

et que nous validons par **OK**, nous voyons que cela divise par 10 le temps de test:

Plans de test de démonstration

Valeur minimale à démontrer

Echelle (Weibull ou expo) ou emplacement (autres lois de distribution) :

Percentile :  Pourcentage :

Fiabilité :  Durée :

Durée moyenne avant défaillance (MTTF) :

Nombre maximal de défaillances autorisées :

Effectifs des échantillons :

Durée des tests pour chaque unité :

Hypothèses pour la loi

Loi de distribution :

Forme (Weibull) ou échelle (autres lois de distribution) :

Aide

Graphiques...

Options...

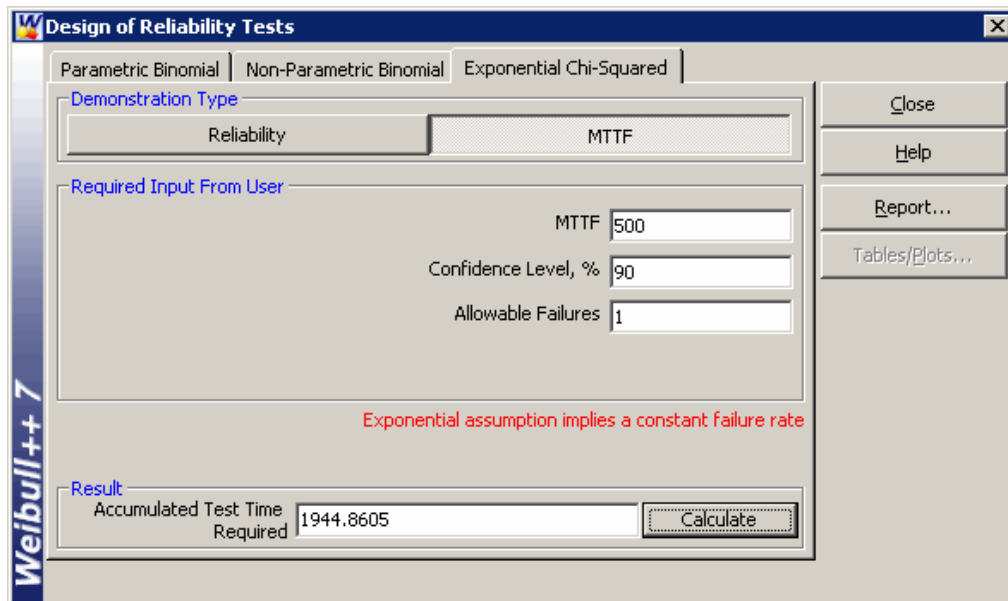
OK

Annuler

et ainsi de suite. Donc dans Minitab, le temps de test est inversement proportionnel linéairement à la taille de l'échantillon.

Ou autre exemple connu:

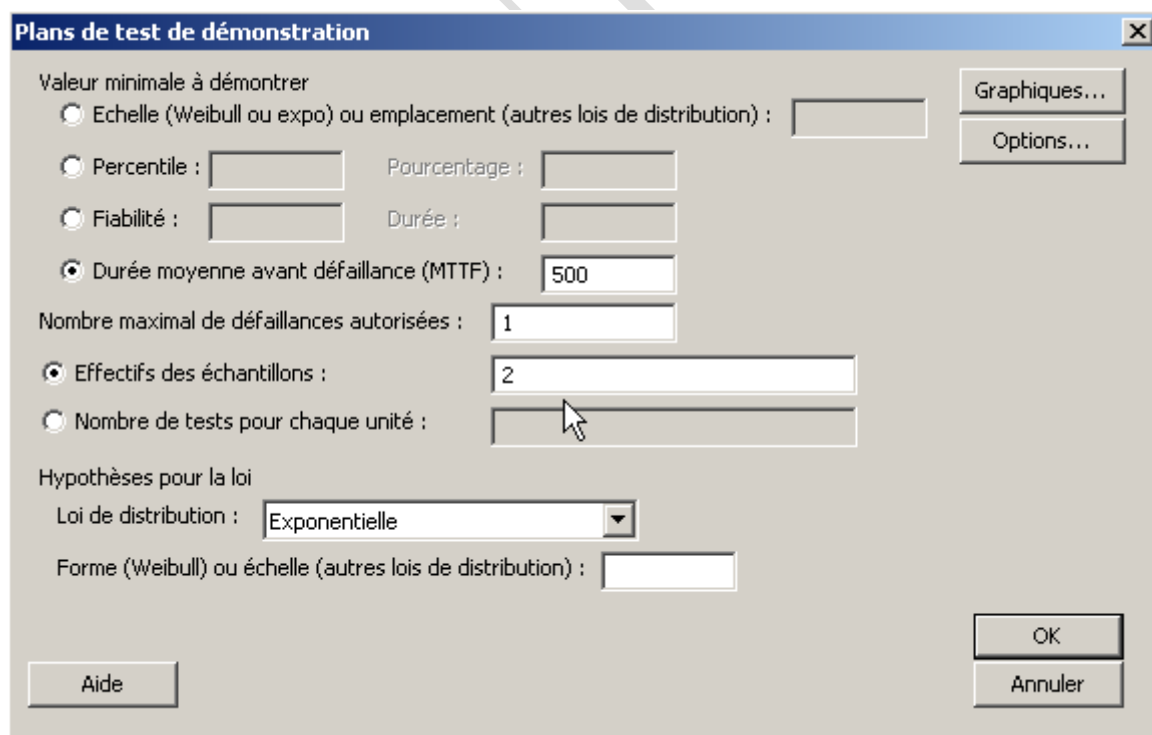
Il est demandé à ingénieur de déterminer le temps minimal de test pour démontrer que la MTTF d'un produit est au moins de 500 heures avec un niveau de confiance de 90%. Le produit est connu comme ayant une fonction de fiabilité de type exponentielle (taux de défaillance constant). Sur ces données, seule 1 défaillance est autorisée. Le temps maximum de test est alors donné par avec Weibull++:



Ce qui correspond bien à:

$$T = \frac{MTTF \cdot \chi_{CL,2(r+1)}^2}{2} = \frac{MTTF \cdot \chi_{0.9,2(1+1)}^2}{2} = 500 * CHISQ.INV(0.9;4)/2 = 1944.86$$

Avec Minitab pour reproduire ce résultat nous devons introduire qu'il y a au moins 2 échantillons (??) car il n'accepte curieusement pas qu'il y en ait qu'un seul (???):



Nous avons alors un résultat proche de Weibull++ et du calcul fait à la main:

Plan de test de fiabilité

Loi : exponentielle

Objectif de durée moyenne avant défaillance = 500, Niveau de confiance réel = 95 %

Test de défaillance	Effectif d'échantillon	Temps de test
1	2	1838.07

Voyons maintenant encore un autre scénario où le comportement de Minitab est parfaitement correct et dont nous avons fait les calculs à la main dans le cours théorique et avec Weibull++.

En supposant que le temps de défaillance d'un système suit toujours une distribution exponentielle. La période de garantie de ce système est de 1'000 heures de fonctionnement. Le fabricant est tenu de démontrer que la fiabilité du système à l'heure 1000 est au moins 95% avec un niveau de 50 % de confiance. Aucune défaillance n'est admise dans le test. Combien de temps le test devrait-t-il durer?

Nous mettons alors:

**Plans de test de démonstration**

Valeur minimale à démontrer

Echelle (Weibull ou expo) ou emplacement (autres lois de distribution) :

Percentile :  Pourcentage :

Fiabilité :  Durée :

Durée moyenne avant défaillance (MTTF) :

Nombre maximal de défaillances autorisées :

Effectifs des échantillons :

Nombre de tests pour chaque unité :

Hypothèses pour la loi

Loi de distribution :

Forme (Weibull) ou échelle (autres lois de distribution) :

Aide

Graphiques...  
Options...  
OK  
Annuler

Sans oublier de mettre dans **Options...**:

**Plans de test de démonstration - Options**

Niveau de confiance :

Aide

OK

Annuler

Ce qui nous donne:

### Plans de test de démonstration

Plan de test de fiabilité

Loi : exponentielle

Objectif de fiabilité = 0.95, Niveau de confiance réel = 50 %

Test de	Effectif	Temps
défaillance	d'échantillon	de test
0	1	13513.4

et qui est bien conforme à Weibull++:

**Design of Reliability Tests**

Parametric Binomial | Non-Parametric Binomial | Exponential Chi-Squared

**Demonstration Type**

Reliability | MTF

**Required Input From User**

Reliability (%) 95

Confidence Level, % 50

Allowable Failures 0

Time 1000

Exponential assumption implies a constant failure rate

**Result**

Accumulated Test Time Required 1.3513E+4

Calculate

Close

Help

Report...

Tables/Plots...

Weibull++ 7

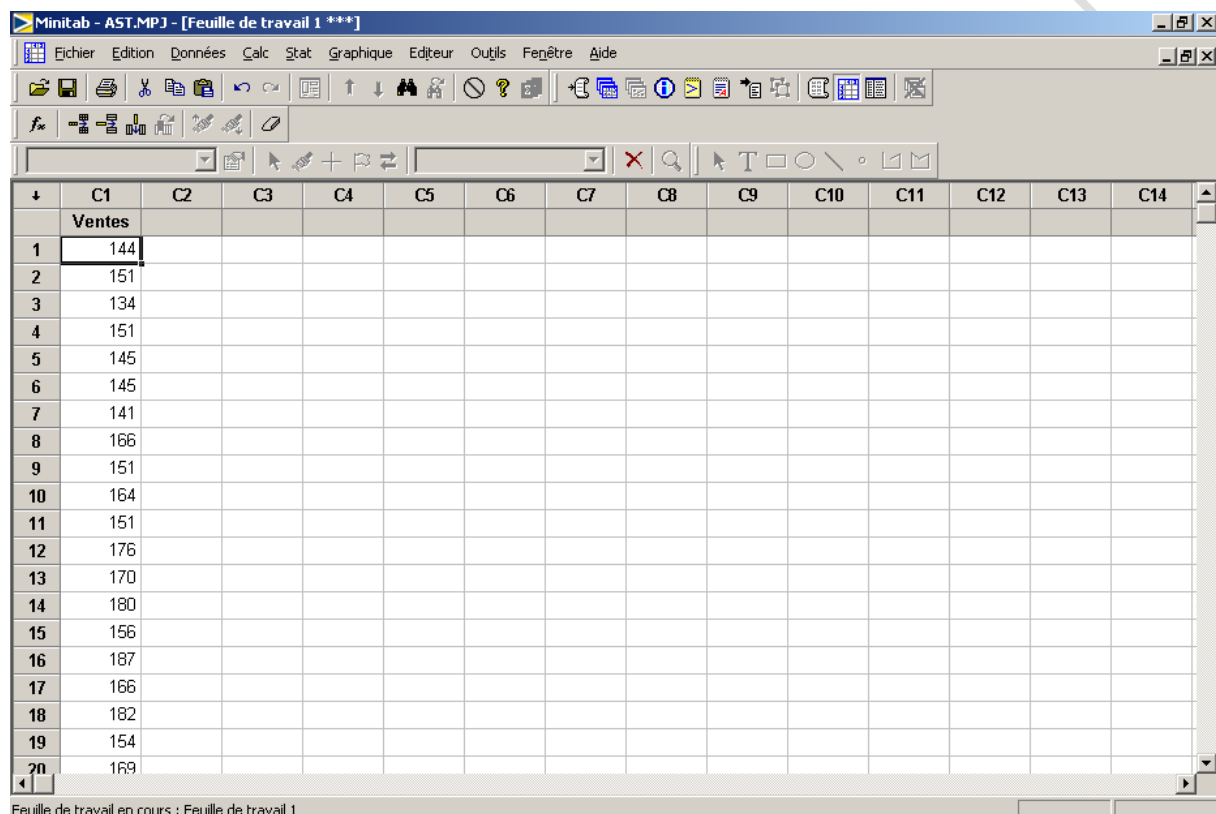


# 19. Séries temporelles

## 19.1. Exercice 200.: Générer un graphique de série chronologique (temporelle)

Minitab® Statistical Software 15.1.1.0

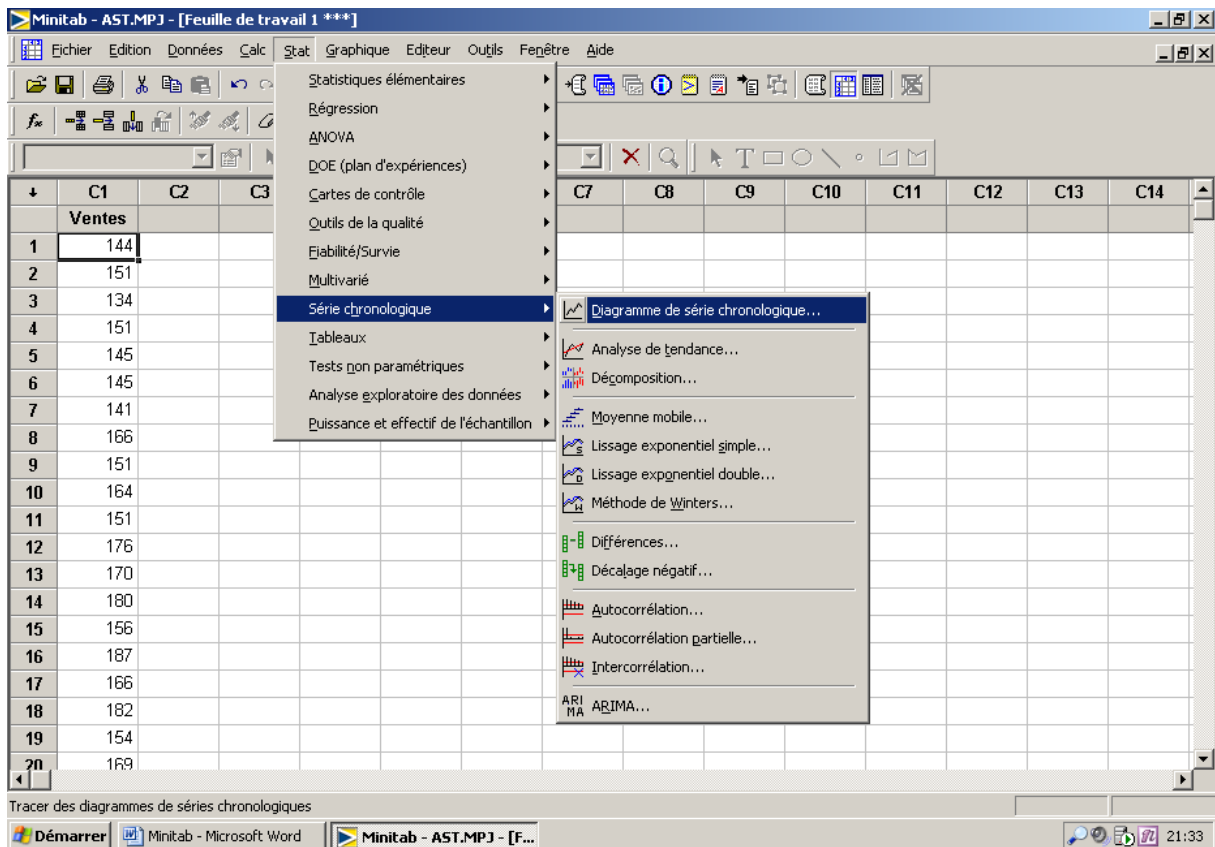
Dans Minitab® Statistical Software, ouvrez le fichier *AST.mpj*:



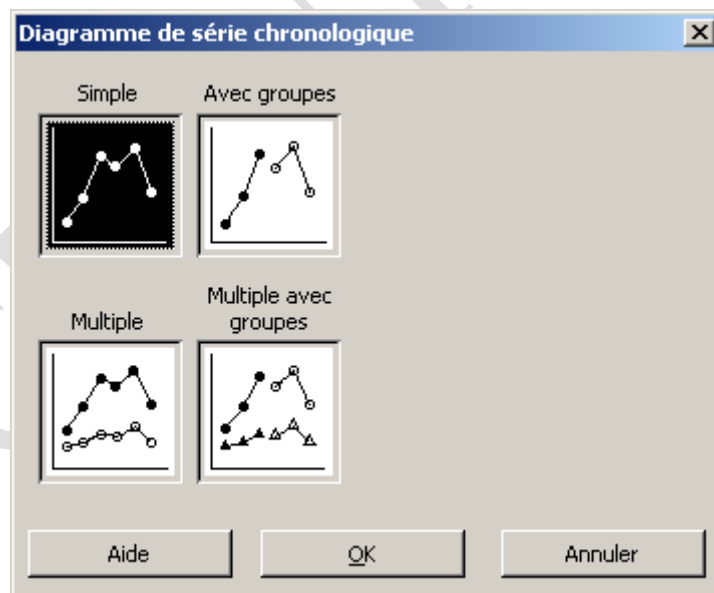
The screenshot shows the Minitab Statistical Software interface. The title bar reads "Minitab - AST.MPJ - [Feuille de travail 1 \*\*\*]". The menu bar includes "Fichier", "Edition", "Données", "Calcul", "Stat", "Graphique", "Editeur", "Outils", "Fenêtre", and "Aide". The toolbar contains various icons for file operations, editing, and analysis. The worksheet grid has columns labeled C1 through C14 and rows numbered 1 through 20. Column C1 is labeled "Ventes" and contains the following data values: 144, 151, 134, 151, 145, 145, 141, 166, 151, 164, 151, 176, 170, 180, 156, 187, 166, 182, 154, 169. The status bar at the bottom indicates "Feuille de travail en cours : Feuille de travail 1".

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Ventes													
1	144													
2	151													
3	134													
4	151													
5	145													
6	145													
7	141													
8	166													
9	151													
10	164													
11	151													
12	176													
13	170													
14	180													
15	156													
16	187													
17	166													
18	182													
19	154													
20	169													

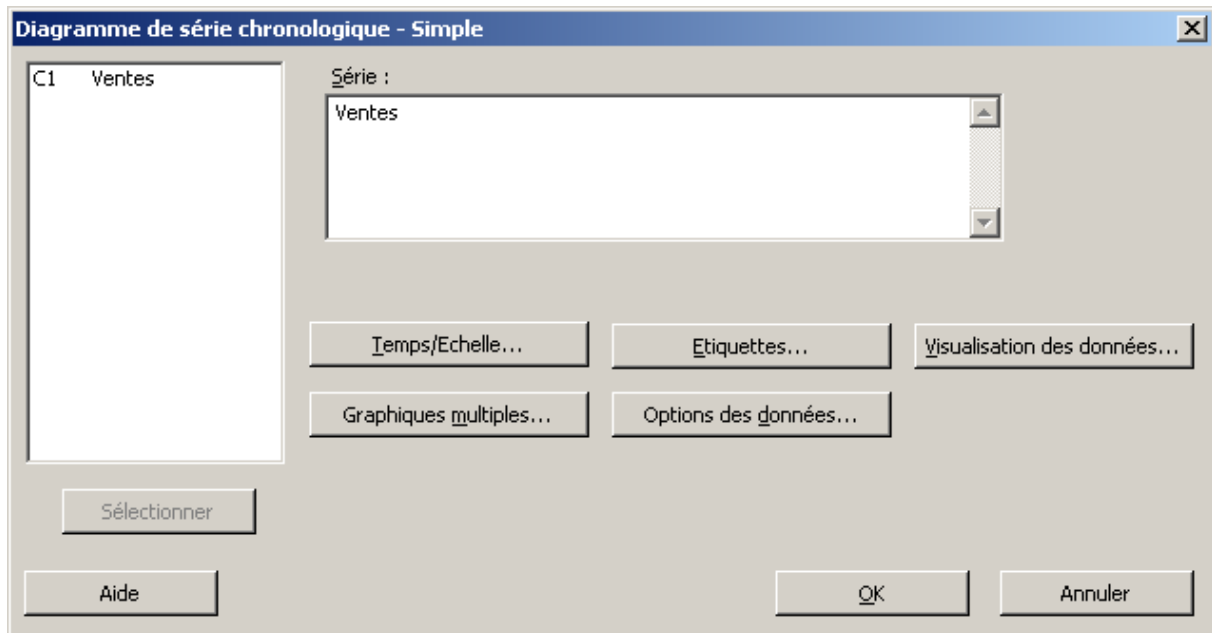
Nous allons dans le menu **Stat/Série chronologique/Diagramme de série chronologique...**:



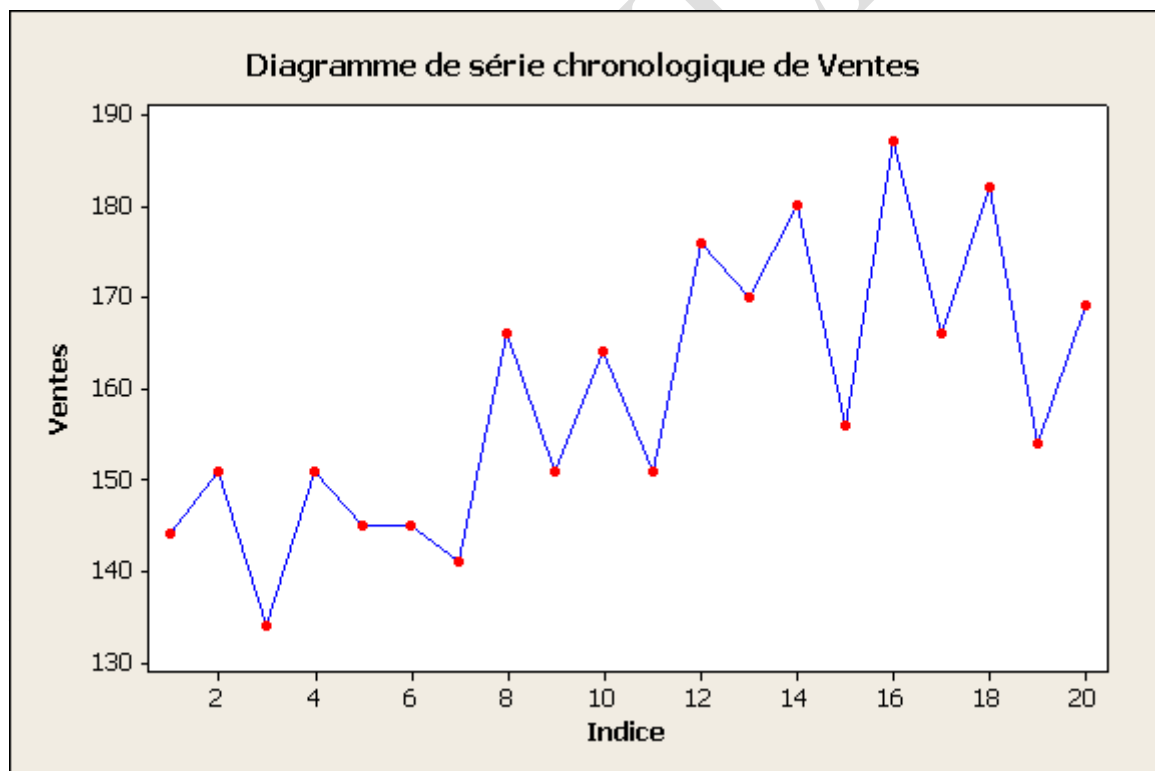
Nous prenons la première option:



Nous validons par **OK**:



Nous revalidons par **OK** pour obtenir un simple graphique (le même que celui que nous avons obtenu dans le cours Microsoft Excel):



### 19.1.1. Extraire des données d'un graphique à points quelconque

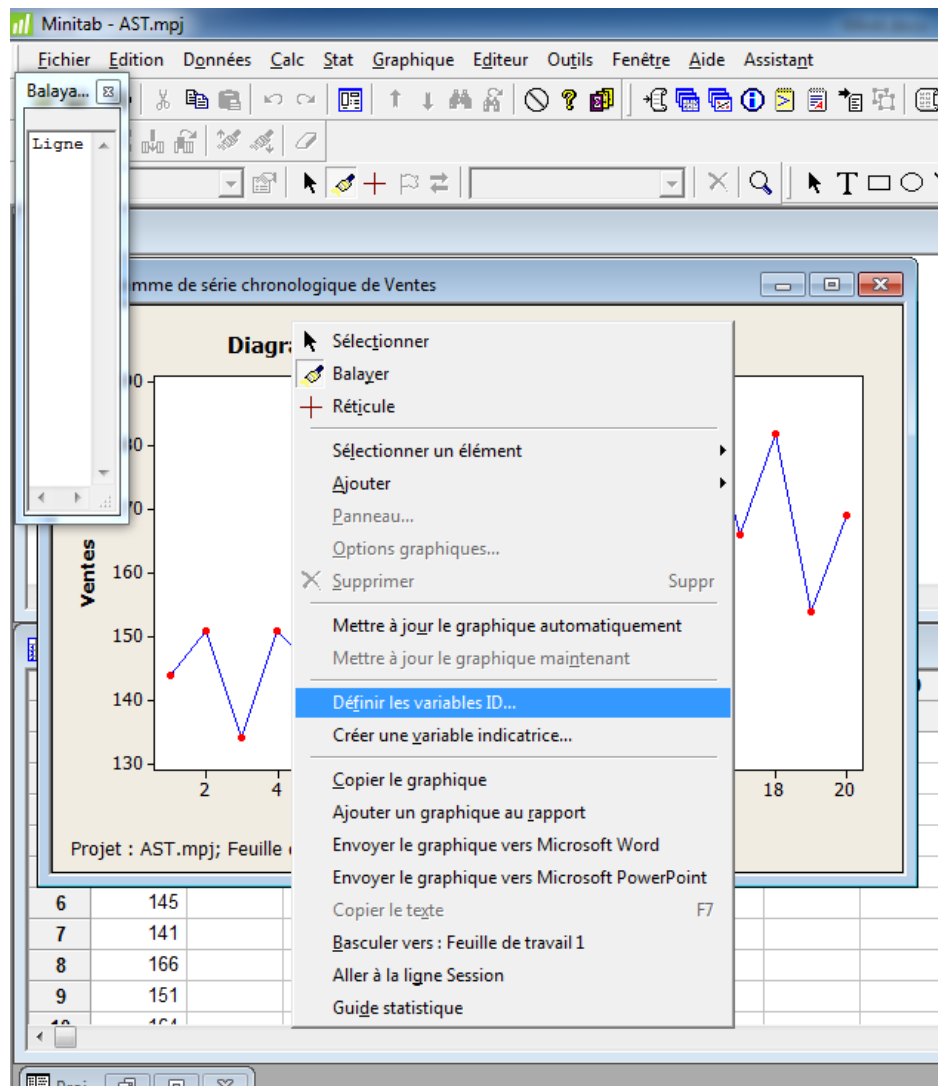
Nous allons voir ici une option qui existe presque pour tous les graphiques à points (régressions, séries chronologiques, cartes de contrôle ou autres) pour analyser un sous-ensemble de points. Pour introduire cet outil nous allons donc utiliser le cas particulier ci-dessus.

Nous faisons un clic droit sur le graphique dont nous voulons représenter un sous-ensemble de points particuliers et nous cliquons sur l'option **Balayer**:

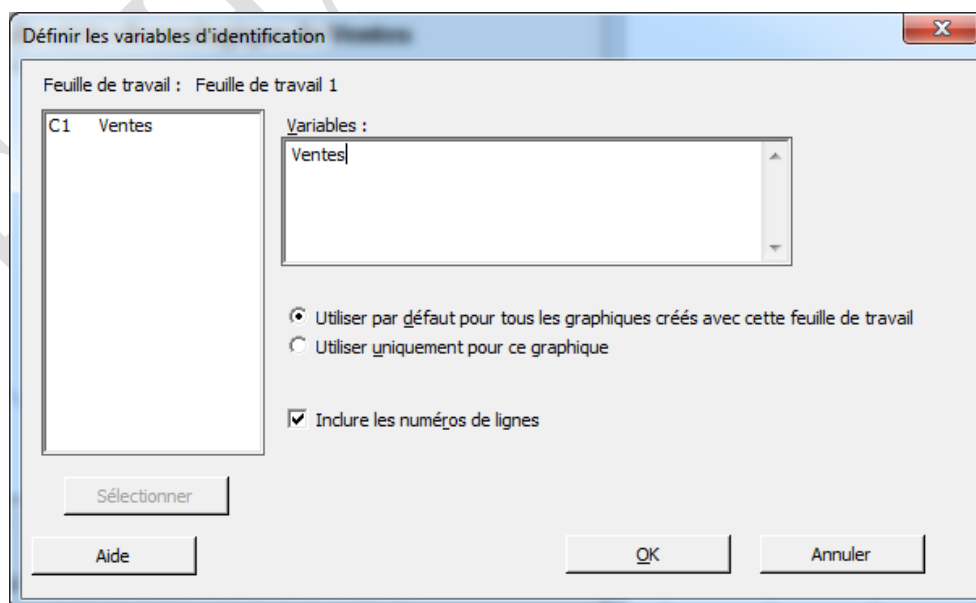
The screenshot shows the Minitab interface with a window titled "Diagramme de série chronologique de Ventes". A context menu is open over the chart, with "Balayer" selected. The chart displays sales data points connected by lines. Below the chart, a data table is visible:

Line Number	Sales Value
6	145
7	141
8	166
9	151

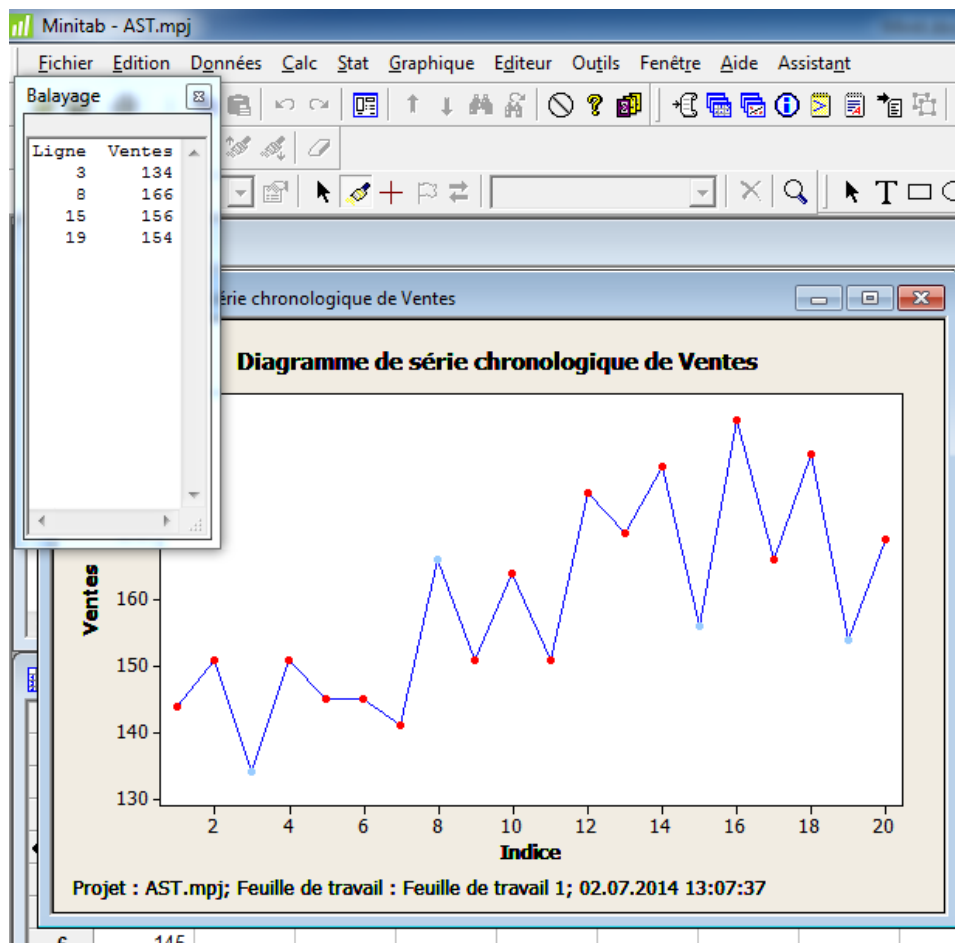
Apparaît alors une toute petite fenêtre qui nous indique le numéro de ligne des points sur lesquels nous avons cliqué mais ceci n'est pas très satisfaisant. Dès lors, pour avoir plus d'informations, nous refaisons un clic droit pour sélectionner l'option **Définir les variables ID...**:



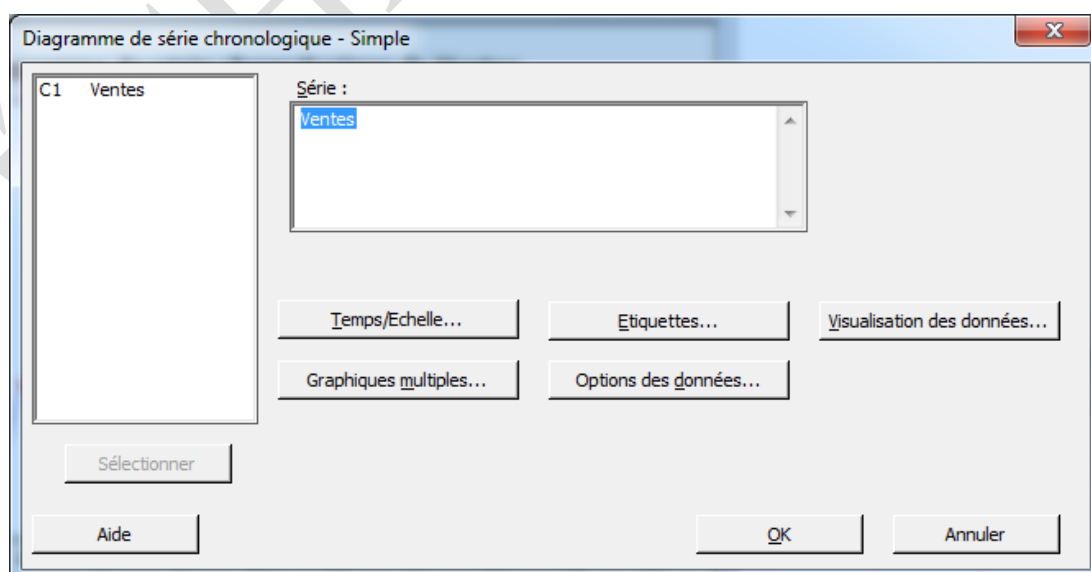
Apparaît alors la boîte de dialogue suivante où nous avons ajouté les données *Ventes*:



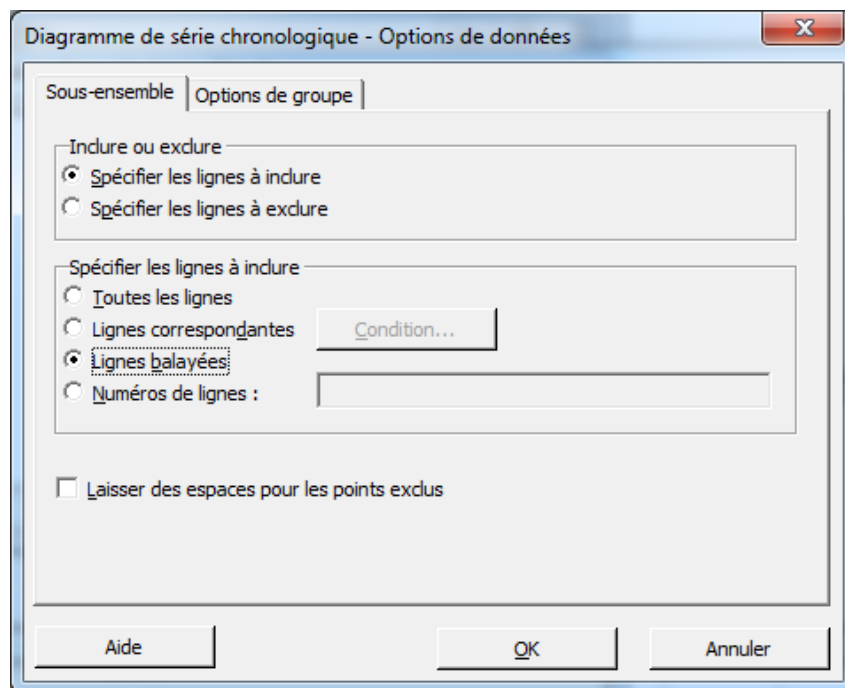
Maintenant si nous cliquons sur un seul point ou plusieurs à la suite en maintenant la touche Shift du clavier enfoncée nous verrons d'abord que les points sélectionnés passent en bleu clair et que la petite fenêtre de balayer indique les points cliqués:



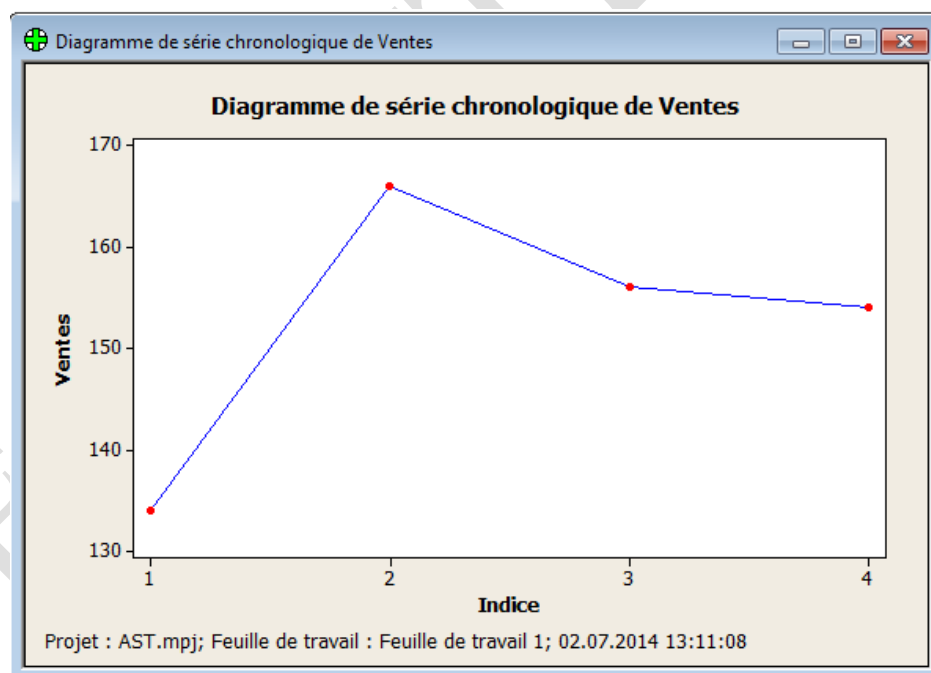
Ensuite, si nous relançons l'assistant graphique en passant par le menu **Stat/Série chronologique/Diagramme de série chronologique...**:



et que nous cliquons sur le bouton **Options des données...** pour y sélectionner l'option **Lignes balayées**:



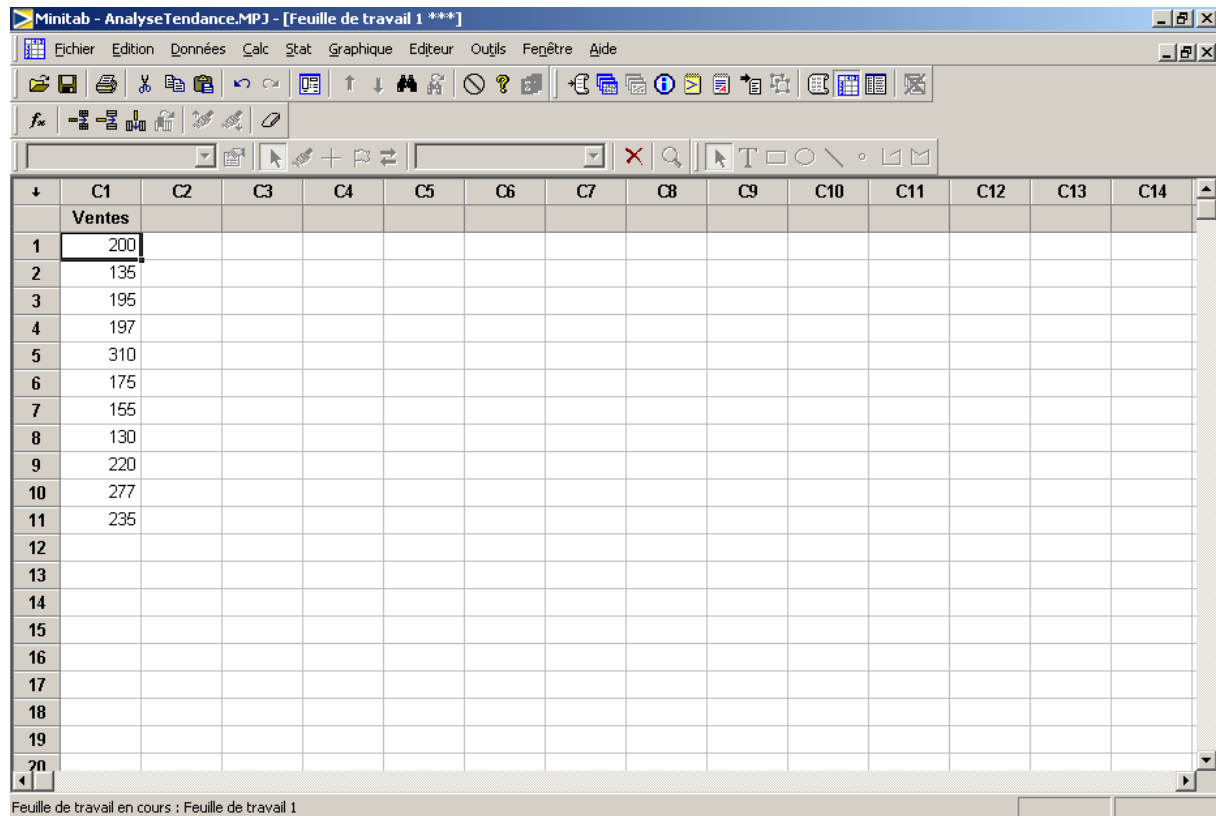
Si nous validons par **OK** nous obtenons bien le sous-ensemble des points balayés:



## 19.2. Exercice 201.: Générer un graphique d'analyse de tendance

Minitab® Statistical Software 15.1.1.0

Ouvrez le fichier *AnalyseTendance.mpj*:



The screenshot shows the Minitab software interface. The title bar reads "Minitab - AnalyseTendance.MPJ - [Feuille de travail 1 \*\*\*]". The menu bar includes "Fichier", "Edition", "Données", "Calcul", "Stat", "Graphique", "Editeur", "Outils", "Fenêtre", and "Aide". The toolbar contains various icons for file operations, editing, and analysis. The worksheet grid has columns labeled C1 through C14. Column C1 is labeled "Ventes" and contains the following data points:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
1	200													
2	135													
3	195													
4	197													
5	310													
6	175													
7	155													
8	130													
9	220													
10	277													
11	235													
12														
13														
14														
15														
16														
17														
18														
19														
20														

The status bar at the bottom indicates "Feuille de travail en cours : Feuille de travail 1".

Nous allons dans le menu **Stat/Série chronologique/Analyse de tendance...**:



Minitab - AnalyseTendance.MPJ - [Feuille de travail 1 \*\*\*]

Stat

- Statistiques élémentaires
- Régression
- ANOVA
- DOE (plan d'expériences)
- Cartes de contrôle
- Outils de la qualité
- Fiabilité/Survie
- Multivarié
- Série chronologique**
  - Diagramme de série chronologique...
  - Analyse de tendance...**
  - Décomposition...
  - Moyenne mobile...
  - Lissage exponentiel simple...
  - Lissage exponentiel double...
  - Méthode de Winters...
  - Différences...
  - Décalage négatif...
  - Autocorrélation...
  - Autocorrélation partielle...
  - Intercorrélation...
  - ARI MA ARIMA...
- Tableaux
- Tests non paramétriques
- Analyse exploratoire des données
- Puissance et effectif de l'échantillon

	C1	C2	C3
	Ventes		
1	200		
2	135		
3	195		
4	197		
5	310		
6	175		
7	155		
8	130		
9	220		
10	277		
11	235		
12			
13			
14			
15			
16			
17			
18			
19			
20			

Ajuster les lignes de tendance à l'aide d'un modèle en courbe linéaire, quadratique, de croissance ou S

Vient alors en paramétrant:

**Analyse de tendance**

C1 Ventes

Variable :

Type de modèle

Linéaire       Croissance exponentielle

Quadratique       Courbe S (logistique de Pearl-Reed)

Générer des prévisions

Nombre de prévisions :

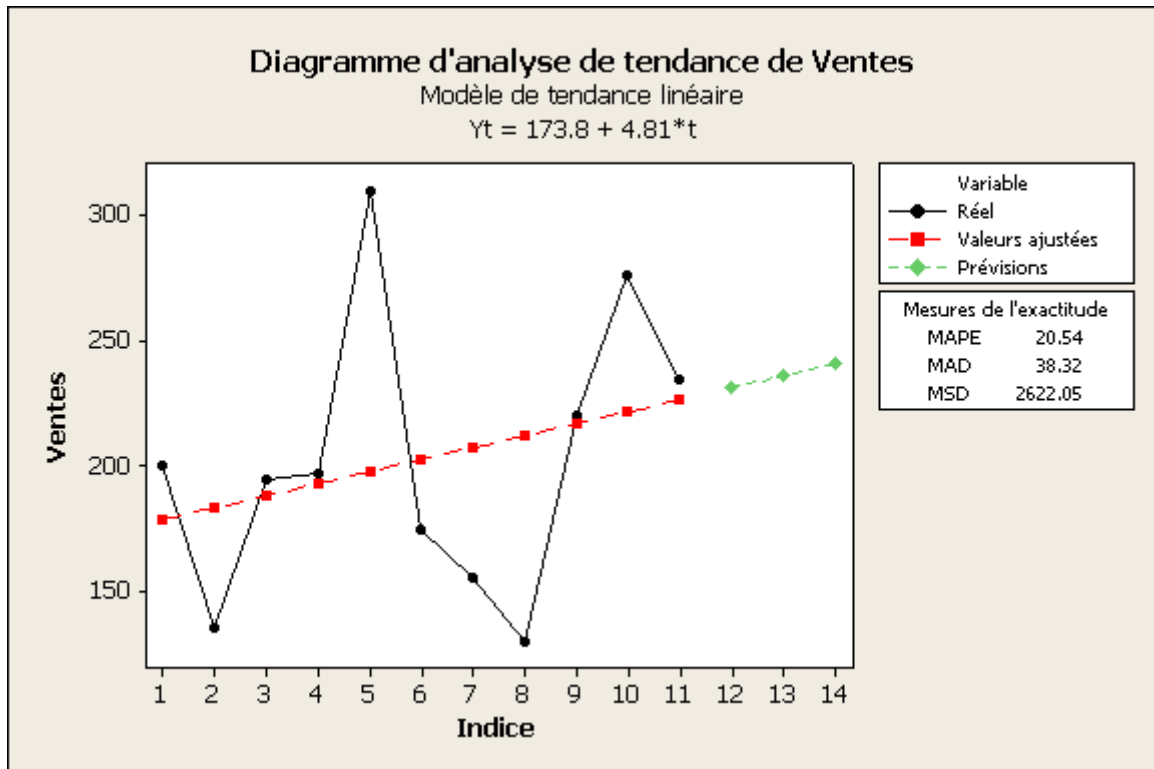
Commencer à l'origine :

Temps... Options... Stockage...

Sélectionner Graphiques... Résultats...

Aide OK Annuler

Nous validons par **OK**:

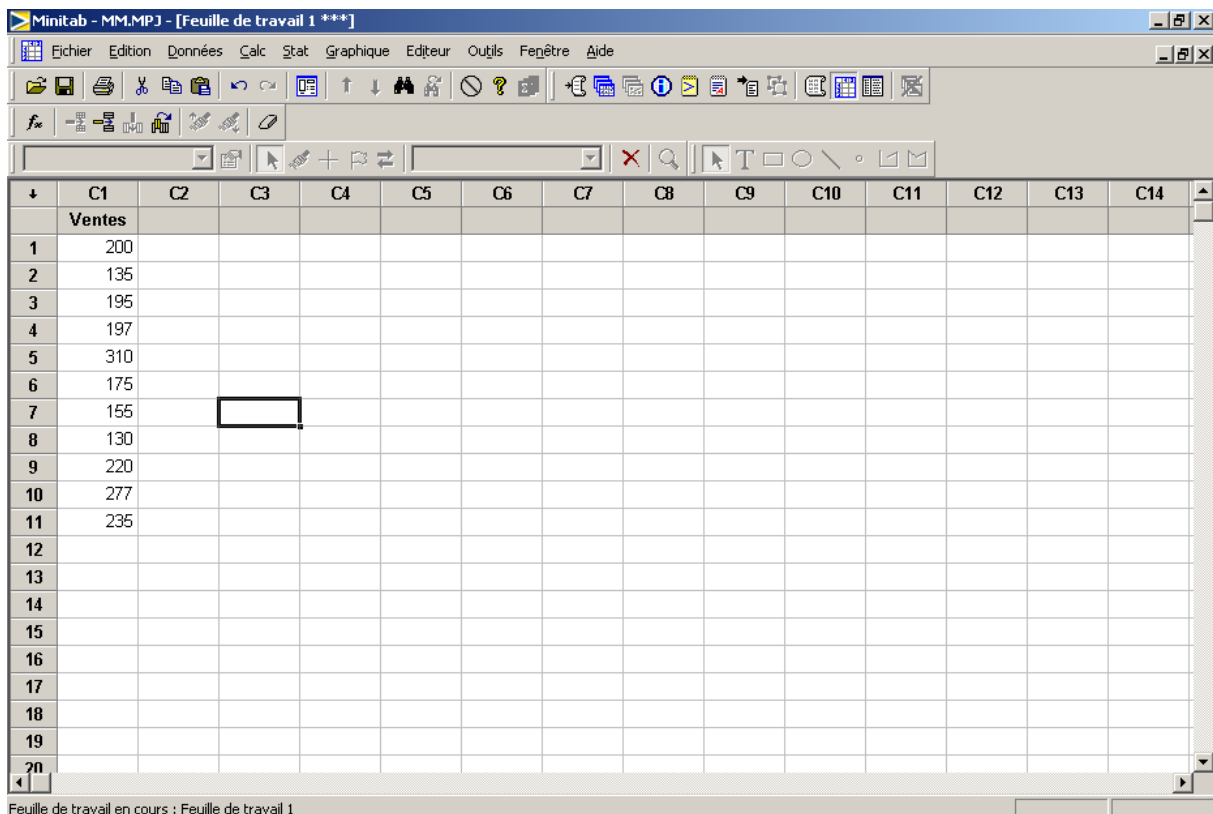


Nous obtenons donc exactement les mêmes résultats qu'avec Microsoft Excel.

## 19.3. Exercice 202.: Générer un graphique de moyennes mobiles (MM3)

Minitab® Statistical Software 15.1.1.0

Ouvrez le fichier *MM.mpj*:



The screenshot shows the Minitab Statistical Software interface. The title bar reads "Minitab - MM.MPJ - [Feuille de travail 1 \*\*\*]". The menu bar includes "Fichier", "Edition", "Données", "Calc", "Stat", "Graphique", "Editeur", "Outils", "Fenêtre", and "Aide". The toolbar contains various icons for file operations, editing, and analysis. The worksheet grid has columns labeled C1 through C14 and rows numbered 1 through 20. Column C1 is labeled "Ventures" and contains the following data values: 200, 135, 195, 197, 310, 175, 155, 130, 220, 277, 235. A small rectangular box is highlighted in the cell at row 7, column C3. The status bar at the bottom indicates "Feuille de travail en cours : Feuille de travail 1".

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Ventes													
1	200													
2	135													
3	195													
4	197													
5	310													
6	175													
7	155													
8	130													
9	220													
10	277													
11	235													
12														
13														
14														
15														
16														
17														
18														
19														
20														

Nous allons dans le menu **Stat/Série chronologique/Moyenne mobile...**:

The screenshot shows the Minitab interface with the 'Stat' menu open. The 'Série chronologique' option is selected, and the 'Moyenne mobile...' option is highlighted in the submenu. The background data table is as follows:

	C1	C2	C3
	Ventes		
1	200		
2	135		
3	195		
4	197		
5	310		
6	175		
7	155		
8	130		
9	220		
10	277		
11	235		
12			
13			
14			
15			
16			
17			
18			
19			
20			

Statistiques élémentaires  
Régression  
ANOVA  
DOE (plan d'expériences)  
Cartes de contrôle  
Outils de la qualité  
Fiabilité/Survie  
Multivarié  
Série chronologique  
Tableaux  
Tests non paramétriques  
Analyse exploratoire des données  
Puissance et effectif de l'échantillon

Diagramme de série chronologique...  
Analyse de tendance...  
Décomposition...  
Moyenne mobile...  
Lissage exponentiel simple...  
Lissage exponentiel double...  
Méthode de Winters...  
Différences...  
Décalage négatif...  
Autocorrélation...  
Autocorrélation partielle...  
Intercorrélation...  
ARIMA...

Lisser les données en établissant la moyenne des observations consécutives et des valeurs prévisionnelles

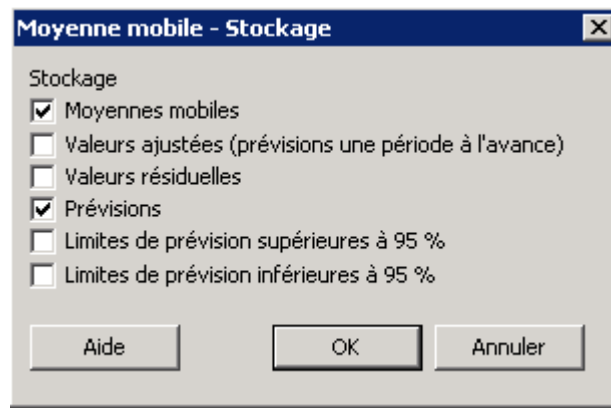
Ce qui nous donne à paramétrer:

The 'Moyenne mobile' dialog box is shown with the following settings:

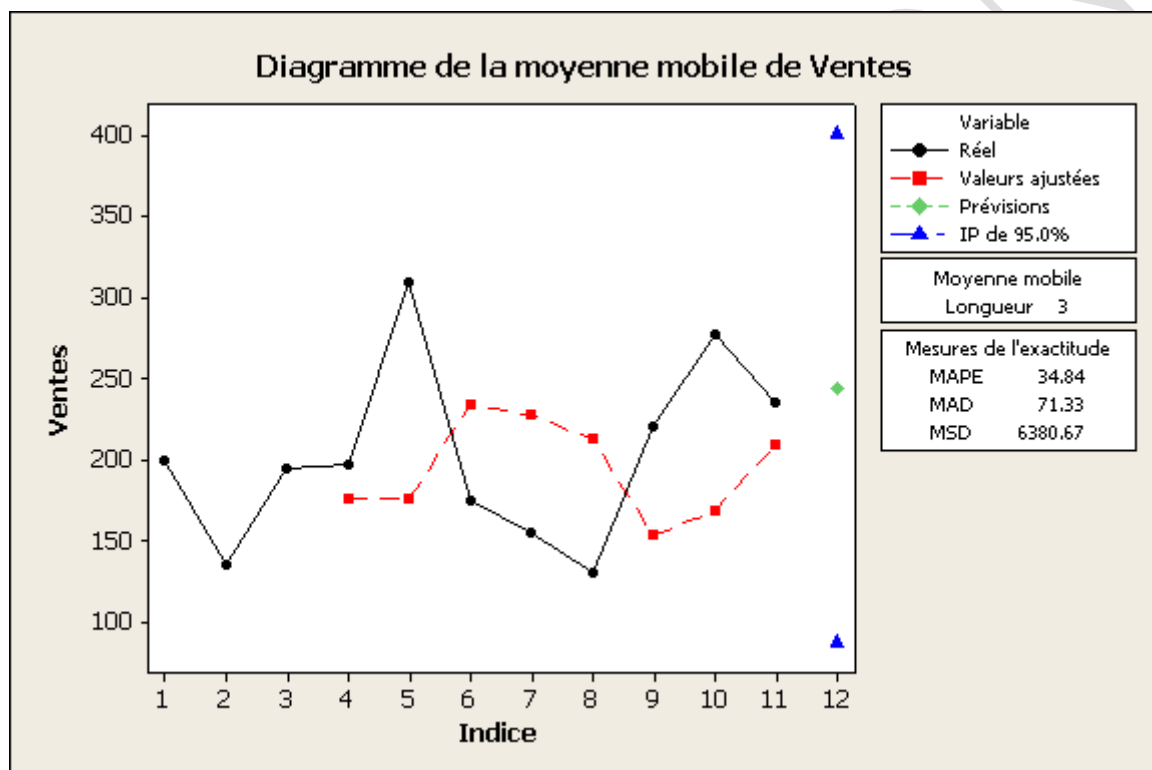
- Variable:
- Longueur MM:
- Centrer les moyennes mobiles
- Générer des prévisions
- Nombre de prévisions:
- Commencer à l'origine:

Buttons: Temps..., Options..., Stockage..., Sélectionner, Graphiques..., Résultats..., Aide, OK, Annuler

Nous allons dans le bouton **Stockage...** pour cocher:



Et nous validons par **OK** pour obtenir le même graphique qu'avec Microsoft Excel:



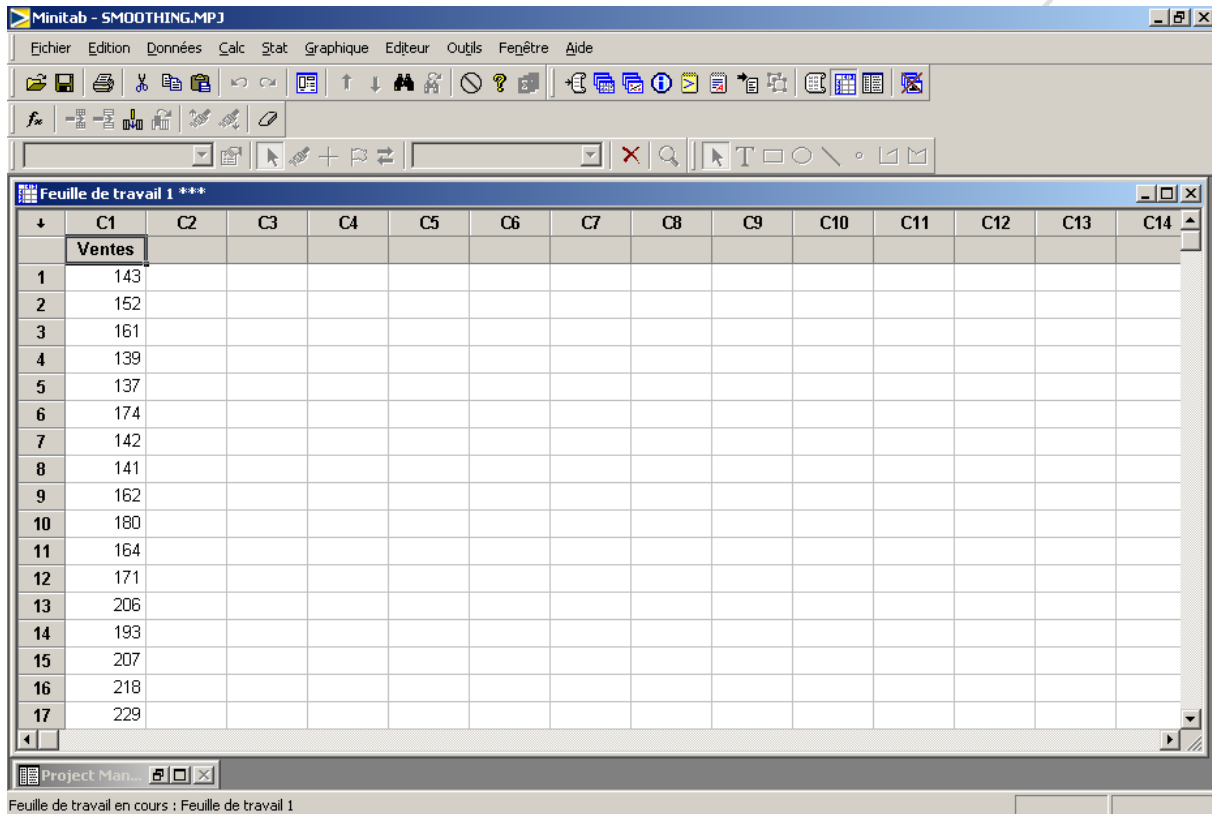
et avec les mêmes valeurs des indicateurs de qualité (MAPE, MAD et MSD) ainsi que les données suivantes dans la table de session:

↓	C1	C2	C3
	Ventes	MOY MOB1	PREVISIONS1
1	200	*	244
2	135	*	
3	195	176.667	
4	197	175.667	
5	310	234.000	
6	175	227.333	
7	155	213.333	
8	130	153.333	
9	220	168.333	
10	277	209.000	
11	235	244.000	

## 19.4. Exercice 203.: Générer un graphique de lissage exponentiel simple

Minitab® Statistical Software 15.1.1.0

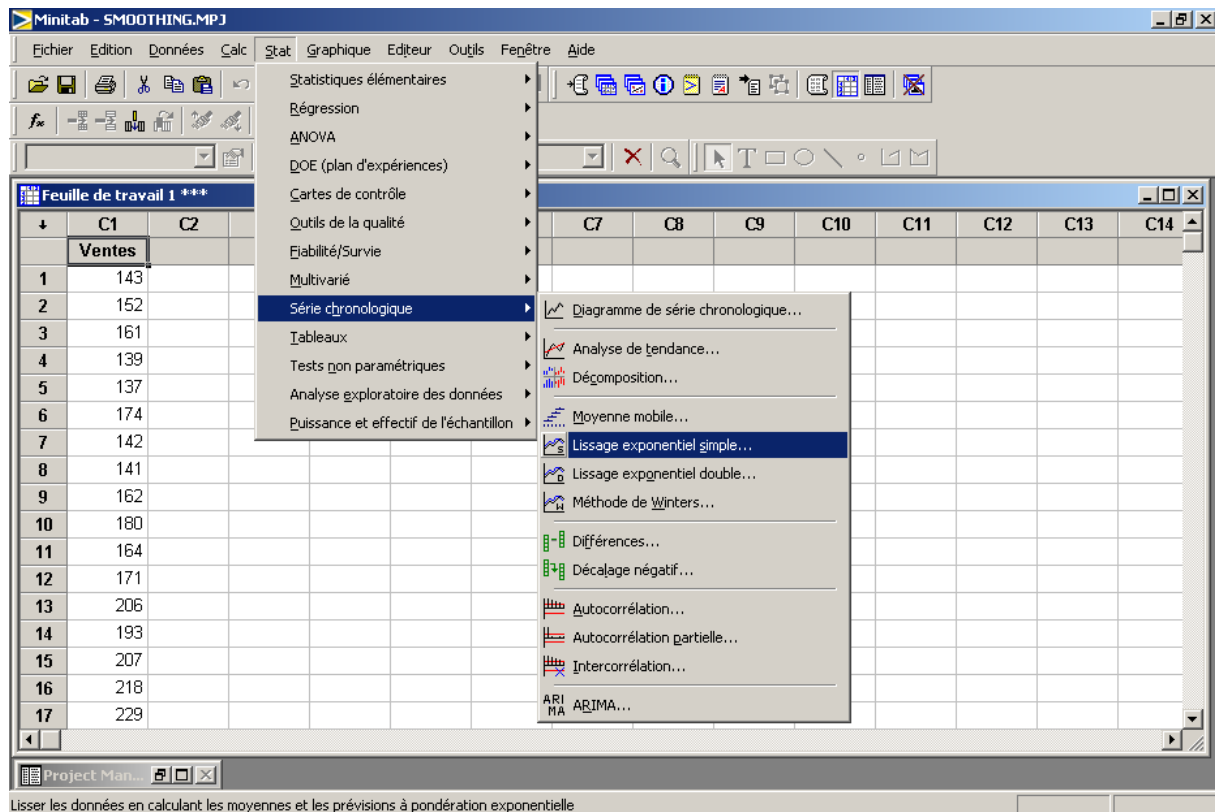
Toujours dans le but de vérifier les notions théoriques vu dans le cours théorique et calculées manuellement avec Microsoft Excel, ouvrez le fichier *Smoothing.mpj*:



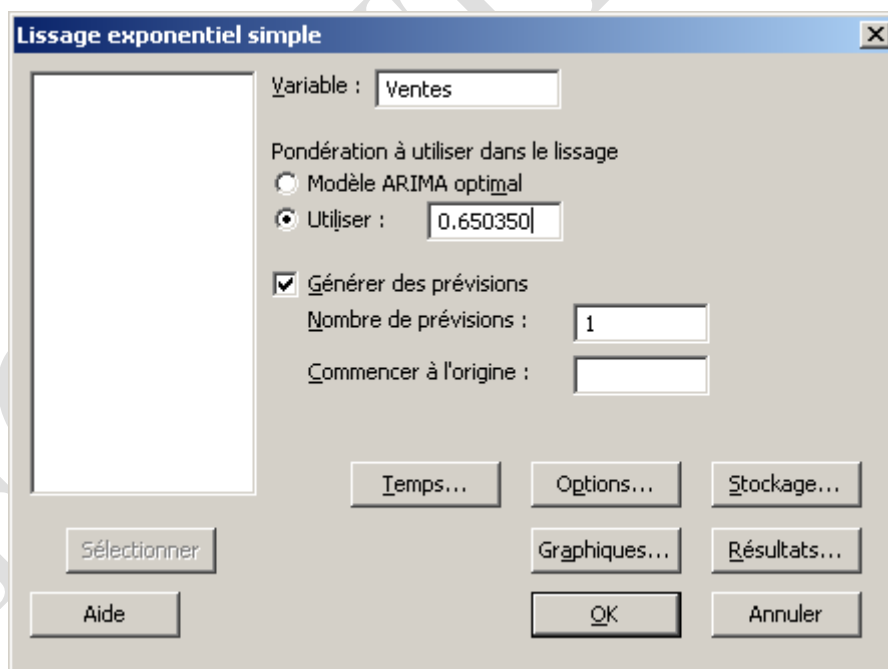
The screenshot shows the Minitab Statistical Software interface. The main window is titled "Minitab - SMOOTHING.MPJ". The menu bar includes "Fichier", "Edition", "Données", "Calc", "Stat", "Graphique", "Editeur", "Outils", "Fenêtre", and "Aide". The toolbar contains various icons for file operations, editing, and analysis. The main workspace displays a worksheet titled "Feuille de travail 1 \*\*\*". The worksheet has columns labeled C1 through C14 and rows numbered 1 through 17. The data in the "Ventes" column is as follows:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
	Ventes													
1	143													
2	152													
3	161													
4	139													
5	137													
6	174													
7	142													
8	141													
9	162													
10	180													
11	164													
12	171													
13	206													
14	193													
15	207													
16	218													
17	229													

Nous allons dans le menu **Stat/Série chronologique/Lissage exponentiel simple...**:

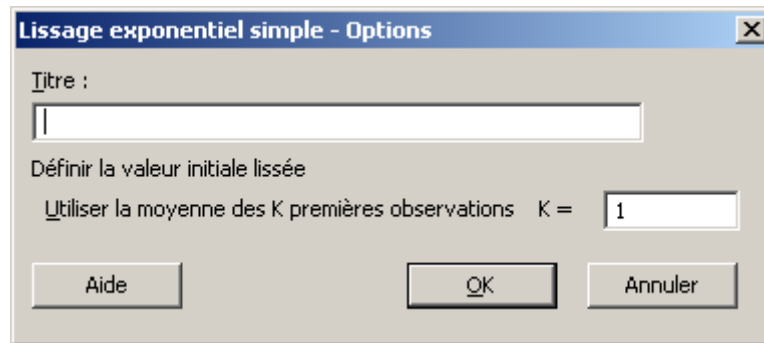


Viens alors avec les paramétrages que nous avons déterminés lors du cours Microsoft Excel:

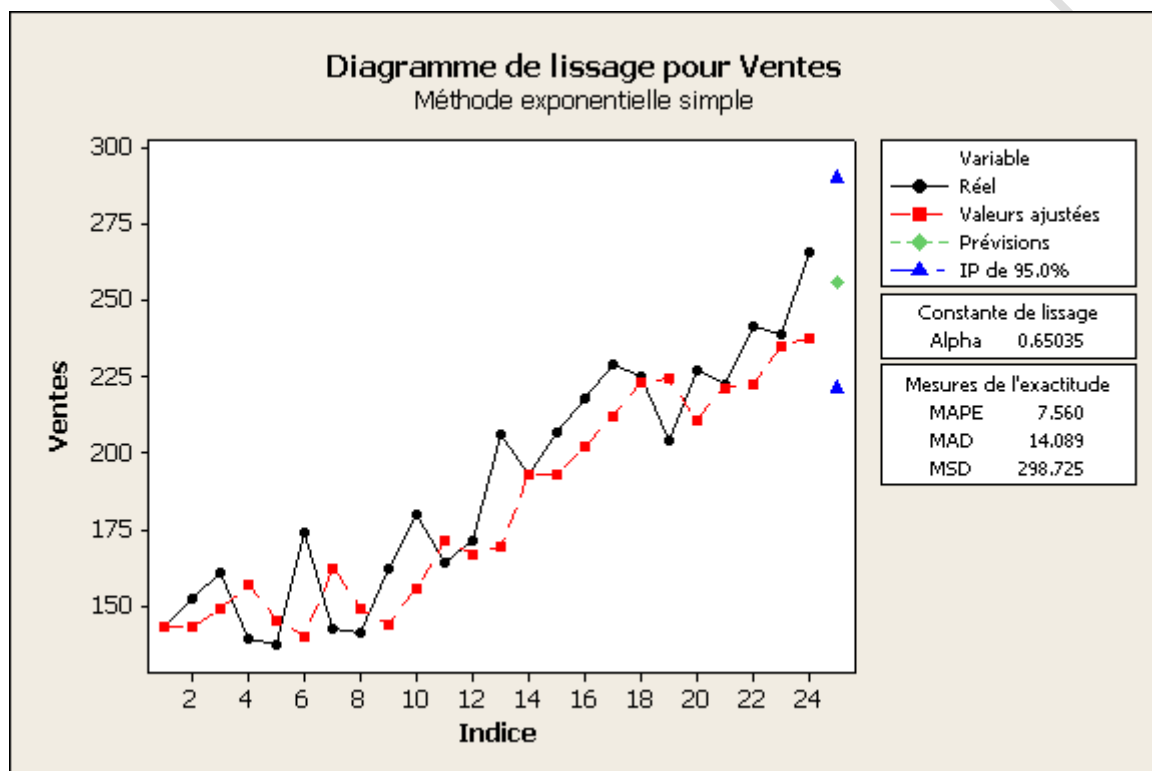


Ensuite il faut cliquer sur le bouton **Options** (si nous souhaitons retrouver le même algorithme que dans Microsoft Excel) et indiquer que nous voulons la moyenne de la première observation comme point de départ:





Ensuite nous validons deux fois par **OK** pour obtenir:



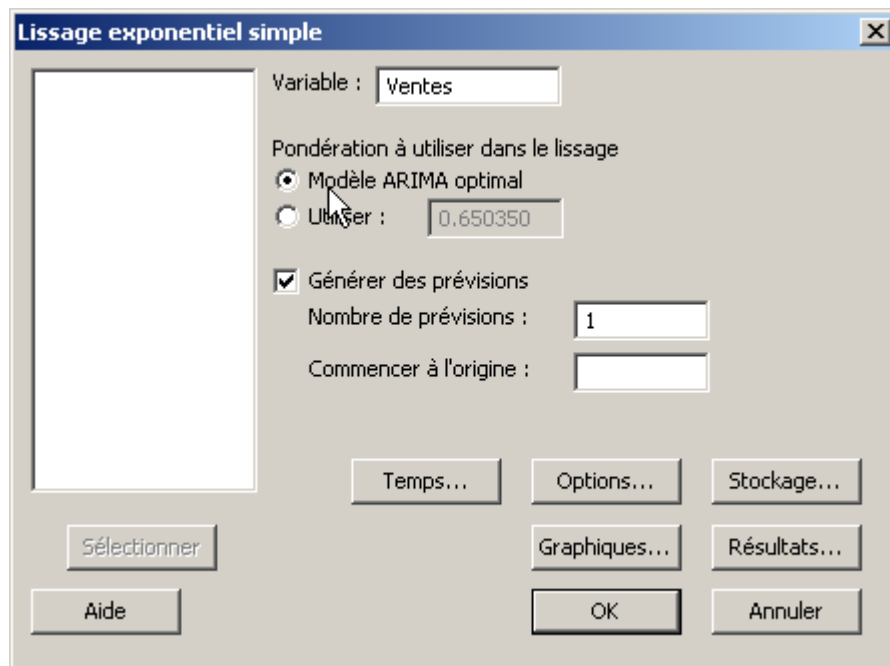
Nous obtenons les mêmes valeurs ajustées que dans Microsoft Excel (ligne rouge) mais par contre les indicateurs d'exactitude ne sont pas tout à fait identiques pour des raisons qui m'échappent.

Regardons dans la fenêtre de session ce que nous avons pour prévision:

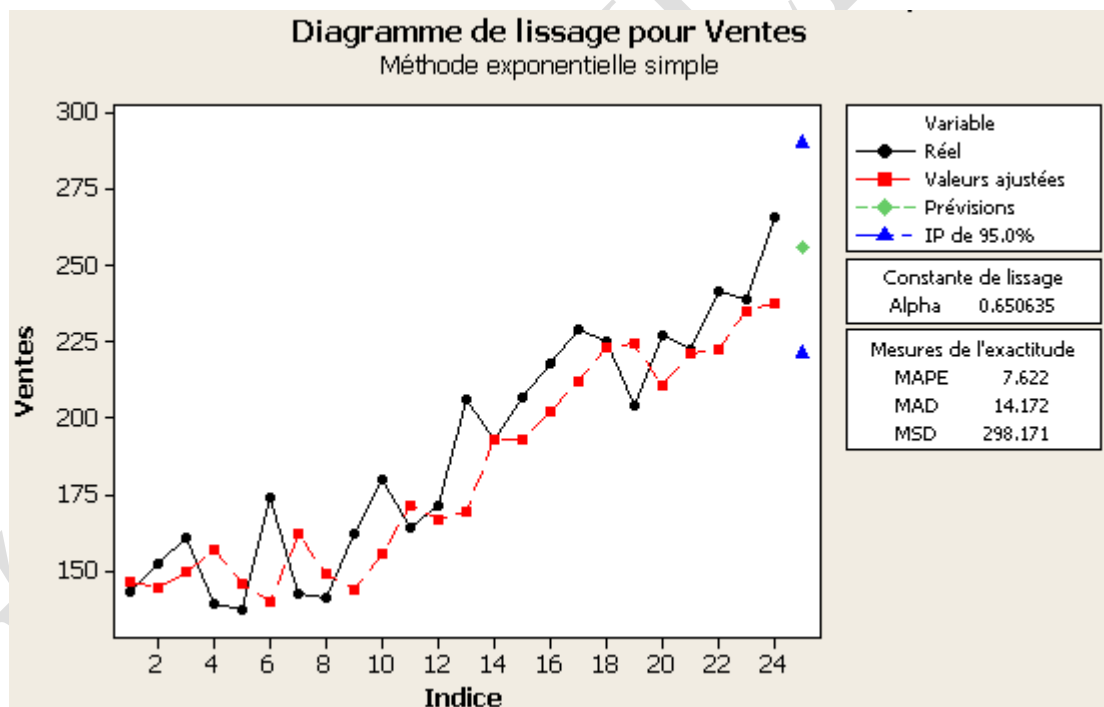
Point	Prévision	Inférieur	Supérieur
25	256.091	221.573	290.609

soit exactement la même chose que dans Microsoft Excel.

Demandons maintenant à Minitab® Statistical Software de chercher pour nous la meilleure constante de lissage:



Nous avons alors:

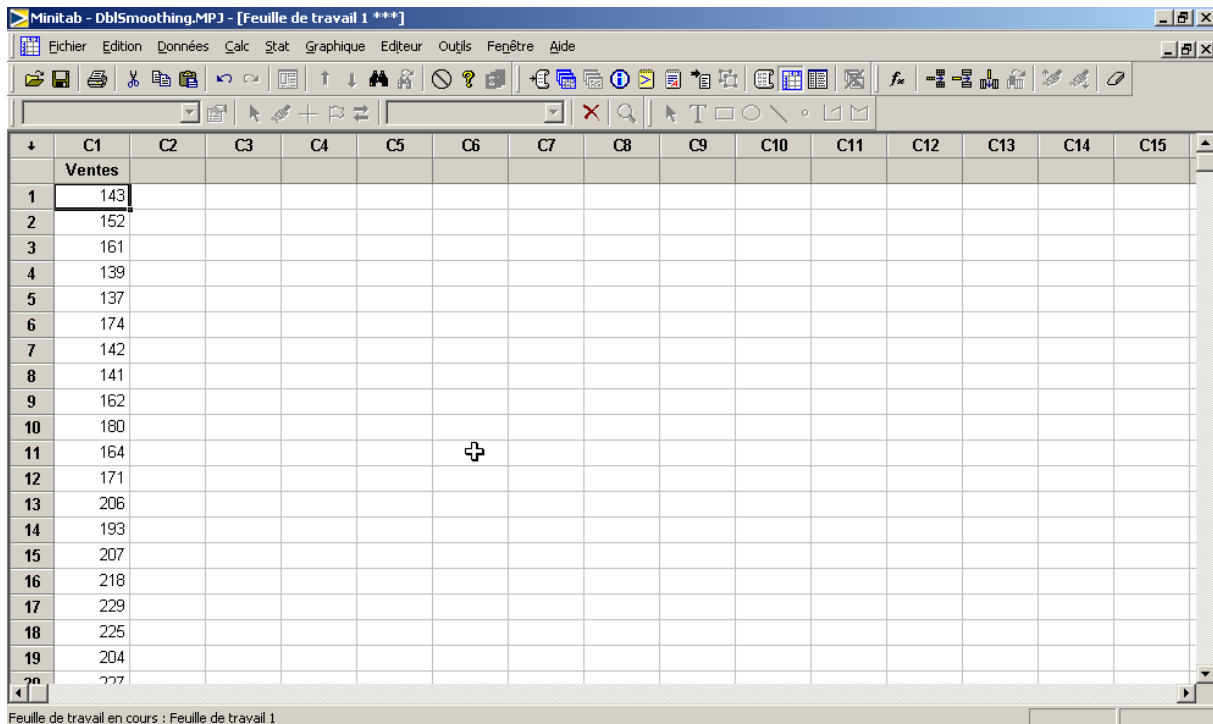


soit une constante de lissage légèrement différente et une erreur un tout petit peu plus faible. Donc Minitab® Statistical Software fait mieux que le solveur de Microsoft Excel 2003 à ce niveau-là!

## 19.5. Exercice 204.: Lissage exponentiel double selon Holt

Minitab® Statistical Software 15.1.1.0

Toujours dans le but de vérifier les notions théoriques vu dans le cours théorique et calculées manuellement avec Microsoft Excel, ouvrez le fichier *DblSmoothing.mpj* (ce sont exactement les mêmes données que pour le lissage exponentiel simple):



The screenshot shows the Minitab interface with a data table. The table has 19 rows and 15 columns. The first column is labeled 'Ventures' and contains the following values: 143, 152, 161, 139, 137, 174, 142, 141, 162, 180, 164, 171, 206, 193, 207, 218, 229, 225, 204. The columns are labeled C1 through C15. The status bar at the bottom indicates 'Feuille de travail en cours : Feuille de travail 1'.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
	Ventes														
1	143														
2	152														
3	161														
4	139														
5	137														
6	174														
7	142														
8	141														
9	162														
10	180														
11	164														
12	171														
13	206														
14	193														
15	207														
16	218														
17	229														
18	225														
19	204														
20	???														

Nous allons dans le menu **Stat/Série chronologique/Lissage exponentiel double...**:

The screenshot shows the Minitab interface with the 'Stat' menu open. The path 'Série chronologique' > 'Lissage exponentiel double...' is highlighted. The data table below shows sales figures over 20 periods.

	C1	C2	C3
	Ventes		
1	143		
2	152		
3	161		
4	139		
5	137		
6	174		
7	142		
8	141		
9	162		
10	180		
11	164		
12	171		
13	206		
14	193		
15	207		
16	218		
17	229		
18	225		
19	204		
20	227		

Statistiques élémentaires  
 Régression  
 ANOVA  
 DOE (plan d'expériences)  
 Cartes de contrôle  
 Outils de la qualité  
 Fiabilité/Survie  
 Multivarié  
**Série chronologique**  
 Tableaux  
 Tests non paramétriques  
 Analyse exploratoire des données  
 Puissance et effectif de l'échantillon

Diagramme de série chronologique...  
 Analyse de tendance...  
 Décomposition...  
 Moyenne mobile...  
 Lissage exponentiel simple...  
**Lissage exponentiel double...**  
 Méthode de Winters...  
 Différences...  
 Décalage négatif...  
 Autocorrélation...  
 Autocorrélation partielle...  
 Intercorrélation...  
 ARI  
 MA  
 ARIMA...

Lisser les données par lissage exponentiel double de Holt et par les valeurs de prévision

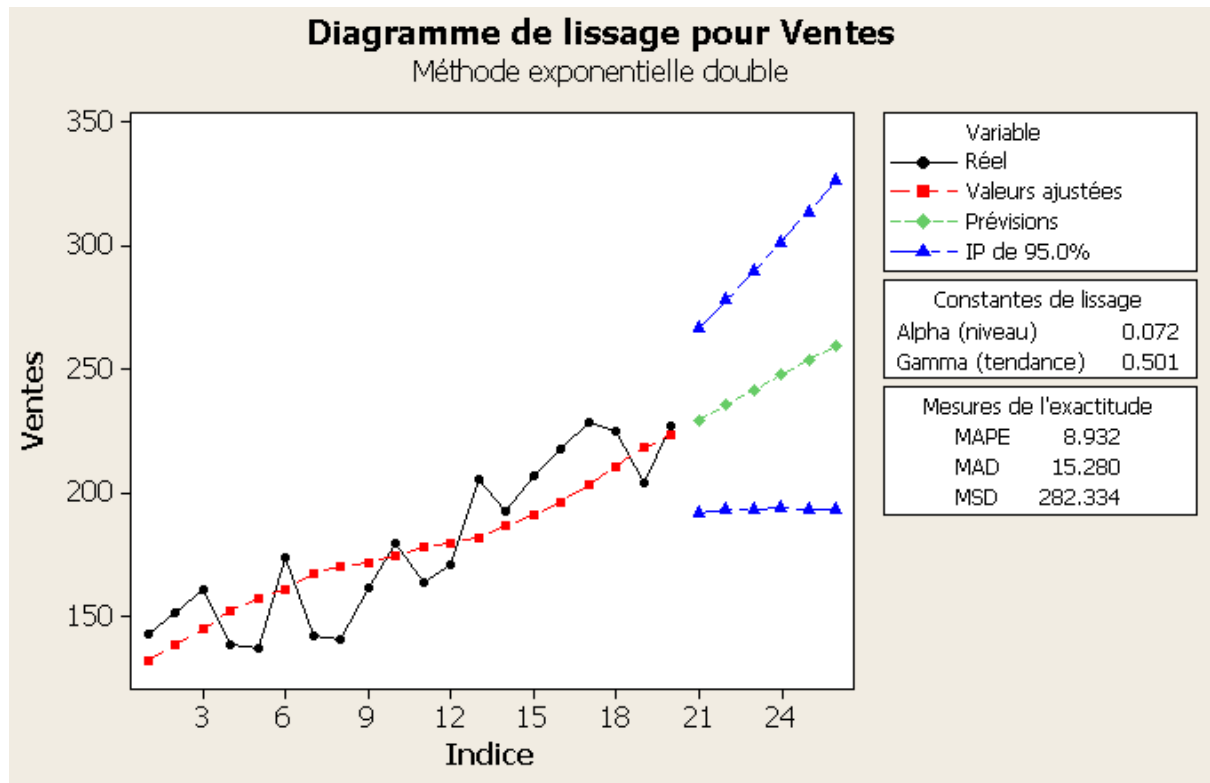
Viens alors avec les paramétrages que nous avons déterminés lors du cours Microsoft Excel:

The dialog box 'Lissage exponentiel double' is shown with the following settings:

- Variable : Ventes
- Pondérations à utiliser dans le lissage:
  - ARIMA optimal
  - Utiliser : 0.072 pour le niveau, 0.501 pour la tendance
- Générer des prévisions
  - Nombre de prévisions : 6
  - Commencer à l'origine : [ ]

Buttons: Sélectionner, Temps..., Options..., Stockage..., Graphiques..., Résultats..., Aide, OK, Annuler

et si nous validons par **OK**:



Les mesures de l'exactitude sont nettement supérieures à celle calculées dans le cours Microsoft Excel. Il faudrait bien évidemment savoir pour cela quelle est l'algorithme utilisé par Minitab® Statistical Software.

Concernant les 6 prévisions, dans la fenêtre de session, nous avons:

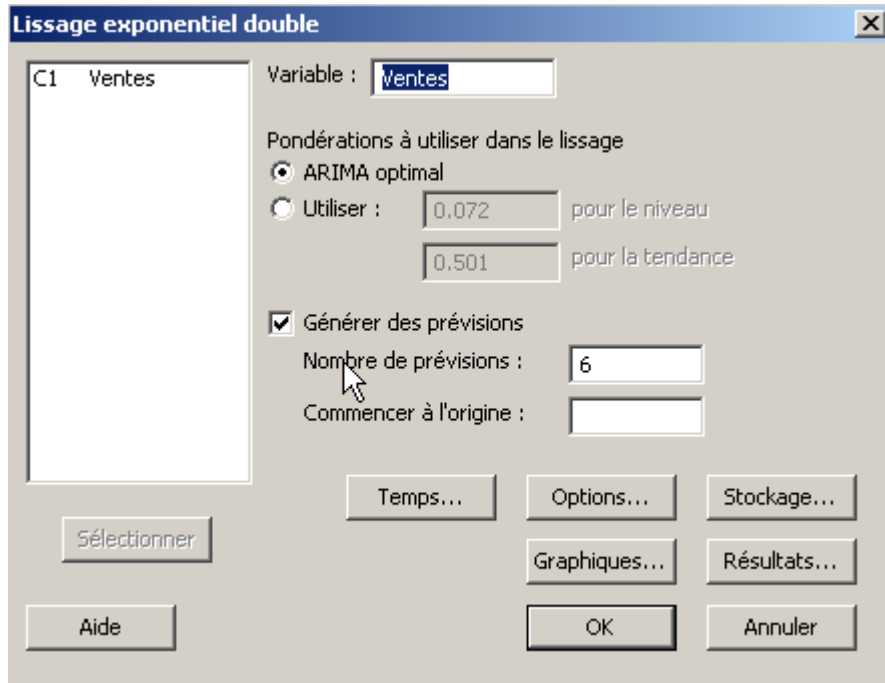
Prévisions			
Point	Prévision	Inférieur	Supérieur
21	229.614	192.180	267.049
22	235.665	193.198	278.133
23	241.716	193.692	289.741
24	247.768	193.824	301.711
25	253.819	193.701	313.936
26	259.870	193.393	326.347

à comparer avec ce que nous avons obtenu avec Microsoft Excel:

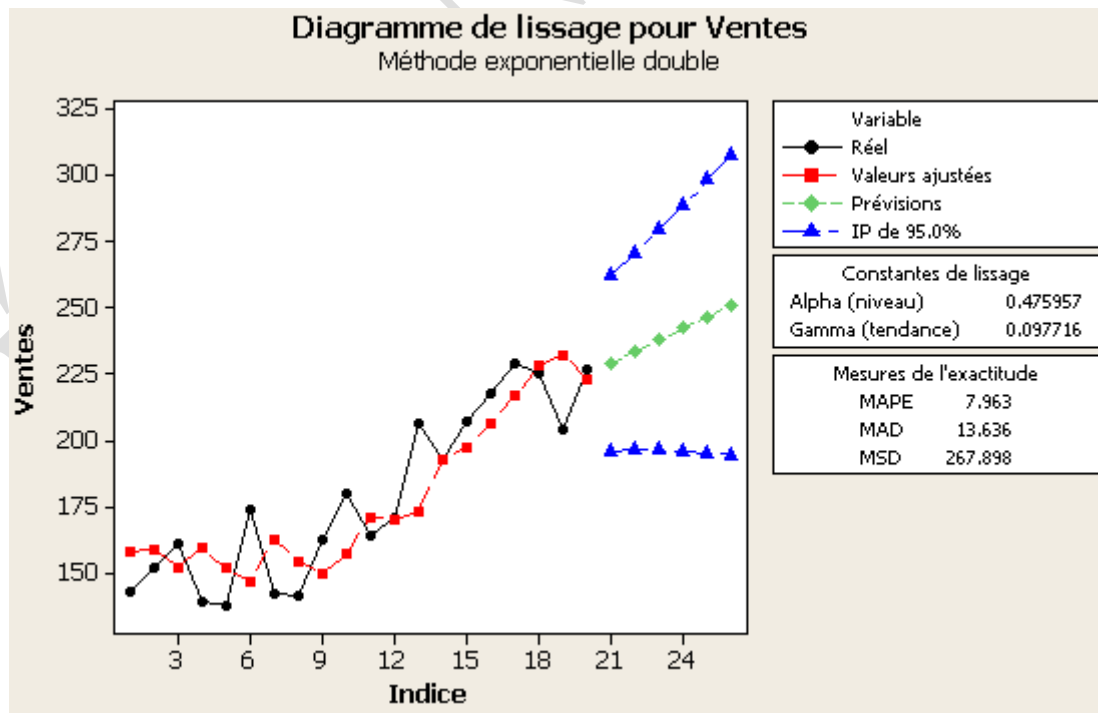
25			<b>262.47</b>	m = 1
26			<b>268.68</b>	m = 2
27			<b>274.89</b>	m = 3
28			<b>281.09</b>	m = 4
29			<b>287.30</b>	m = 5
30			<b>293.51</b>	m = 6

donc cette énorme différence est probablement due à la manière dont Minitab® Statistical Software choisit les paramètres initiaux. Nous avons par ailleurs mentionné dans le cours théorique que les résultats étaient fortement sensibles à la manière de choisir ceux-ci.

Si nous demandons à Minitab® Statistical Software de chercher pour nous les meilleures constantes de lissage:



Nous avons alors:



donc des constantes de lissage légèrement différentes à celles obtenues avec Microsoft Excel et les erreurs sont plus faibles, ce qui nous donne comme prévisions:

---

Point	Prévision	Inférieur	Supérieur
21	229.289	195.882	262.695
22	233.667	196.211	271.124
23	238.046	196.115	279.977
24	242.425	195.716	289.134
25	246.804	195.097	298.511
26	251.183	194.317	308.049

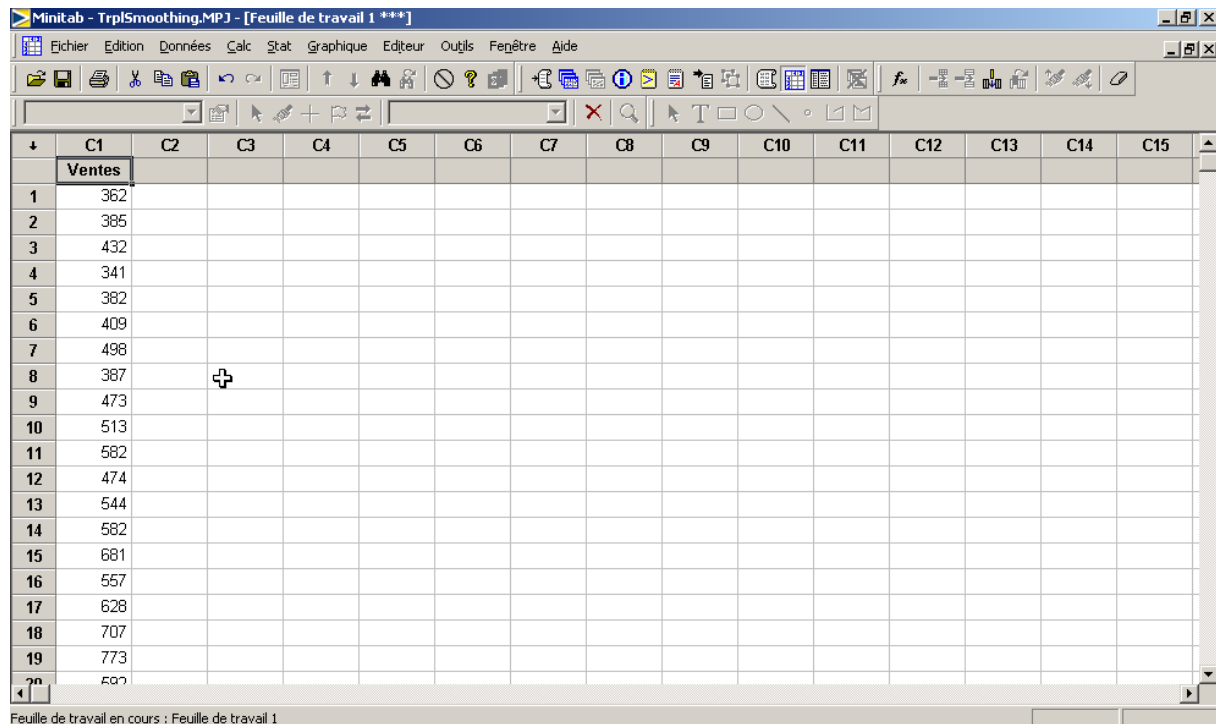
ce qui est encore plus éloigné qu'avant par rapport à ce que nous avons obtenu avec Microsoft Excel.

ÉCHANTILLON

## 19.6. Exercice 205.: Lissage exponentiel triple selon Winters

Minitab® Statistical Software 15.1.1.0

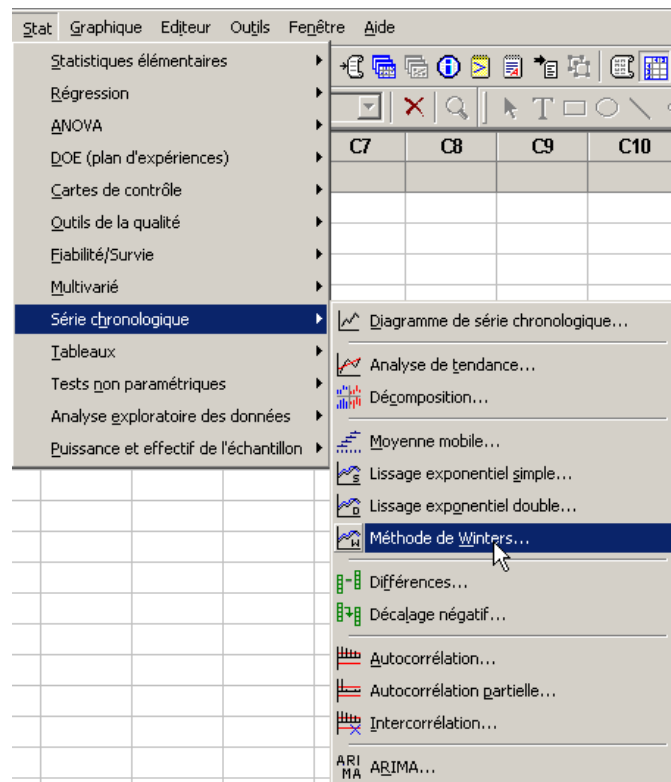
Toujours dans le but de vérifier les notions théoriques vu dans le cours théorique et calculées manuellement avec Microsoft Excel, ouvrez le fichier *TriplSmoothing.mpj*:



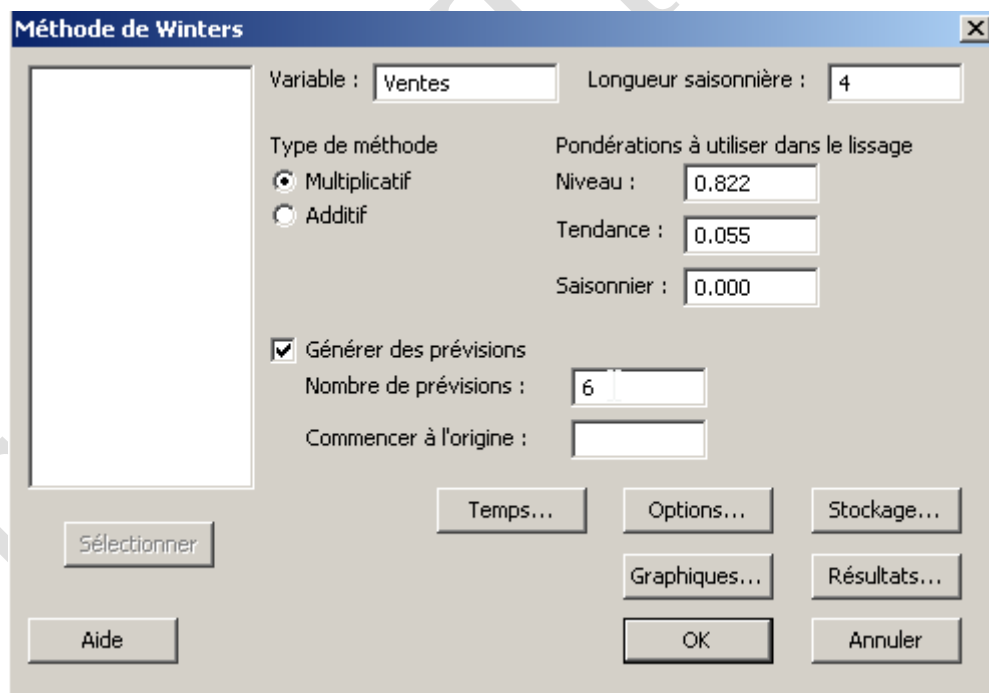
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
	Ventes														
1	362														
2	385														
3	432														
4	341														
5	382														
6	409														
7	498														
8	387														
9	473														
10	513														
11	582														
12	474														
13	544														
14	582														
15	681														
16	557														
17	628														
18	707														
19	773														
20	507														

Nous allons ensuite dans le menu **Stat/Série chronologique/Méthode de Winters...**:

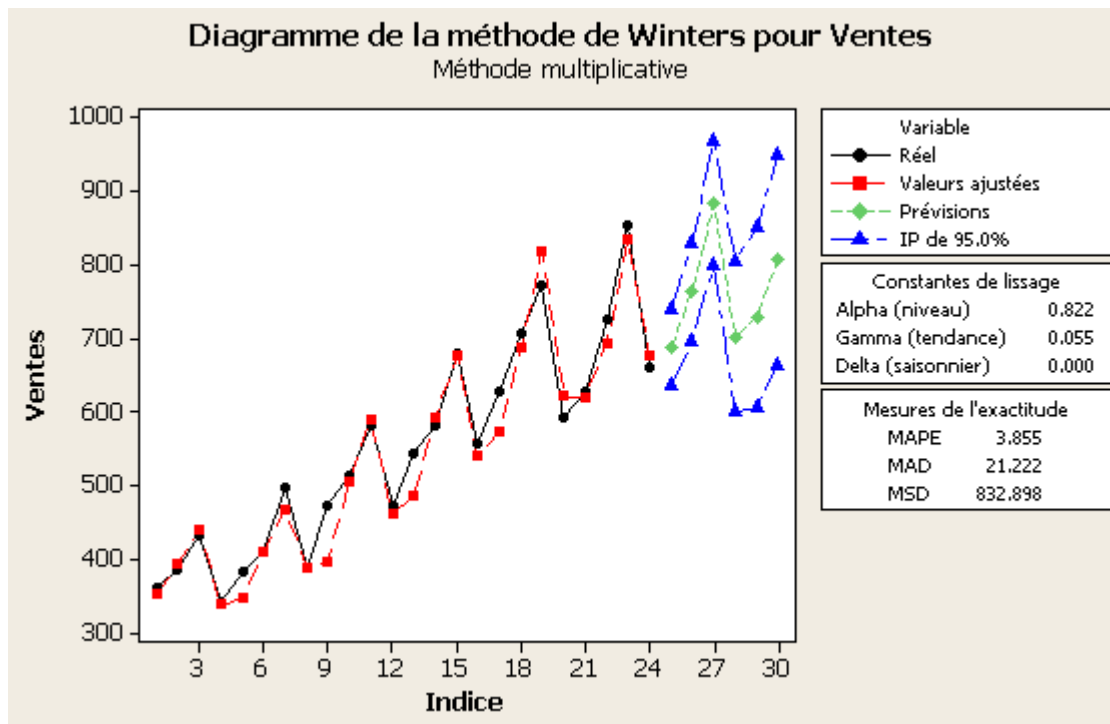




Nous avons alors en saisissant les mêmes valeurs que celles obtenues avec Microsoft Excel:



Nous cliquons sur **OK** pour obtenir d'abord le graphique suivant:



où nous pouvons déjà observer que les erreurs sont beaucoup plus grands que celles obtenus avec Microsoft Excel. Pour les prévisions, dans la fenêtre de session, nous avons:

Prévisions

Point	Prévision	Inférieur	Supérieur
25	688.642	636.649	740.635
26	763.262	696.722	829.801
27	884.774	800.533	969.016
28	703.143	599.650	806.635

A comparer avec celles obtenus dans Microsoft Excel:

25				720.26
26				781.12
27				893.41
28				718.59

Là aussi nous sommes loin du compte...

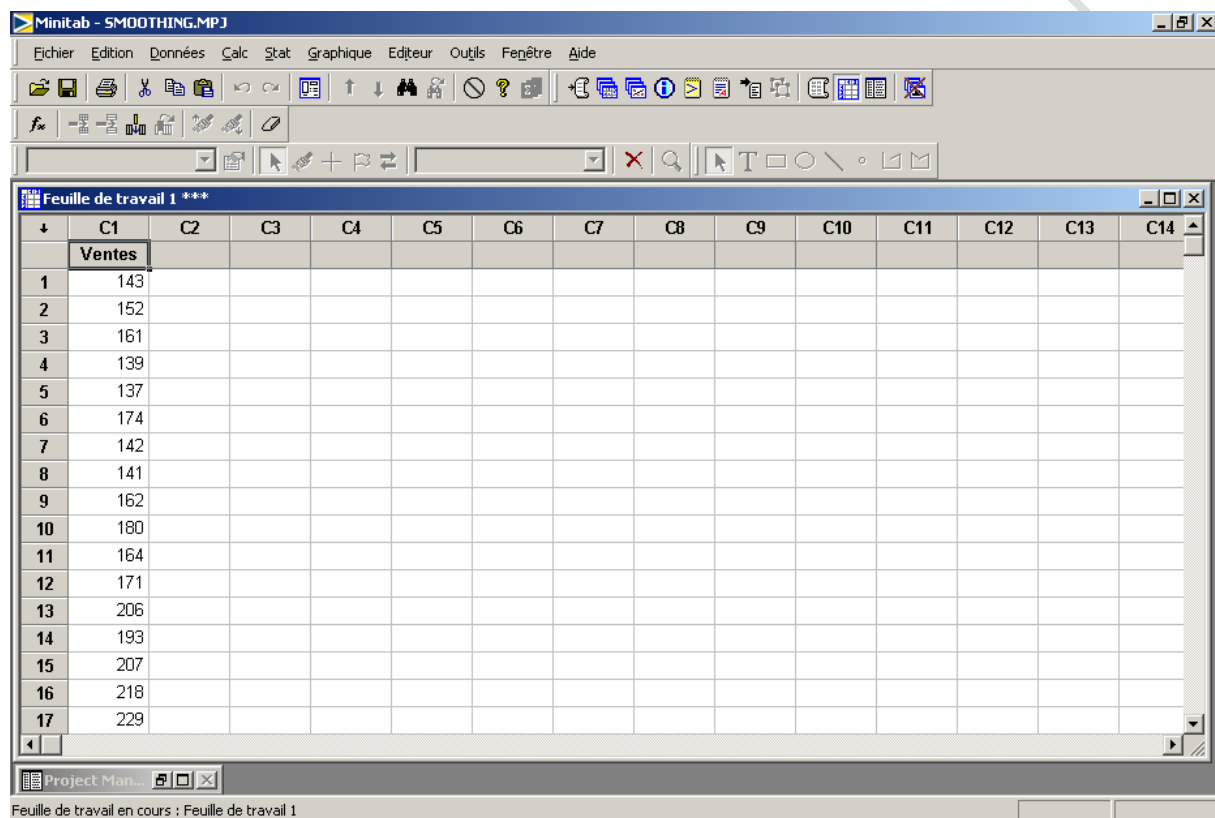
Malheureusement nous ne pouvons pas demander à Minitab® Statistical Software d'optimiser par lui-même car la méthode multiplicative de Holt-Winters n'a pas d'équivalent ARIMA.

## 19.7. Exercice 206.: ARIMA(0,1,1) et ARIMA(0,2,2)

Minitab® Statistical Software 15.1.1.0

Le but va être ici simplement de vérifier ce qui a été vu dans le cours théorique. C'est-à-dire que le lissage exponentiel simple est normalement un ARIMA(0,1,1) et le lissage exponentiel double un ARIMA(0,2,2).

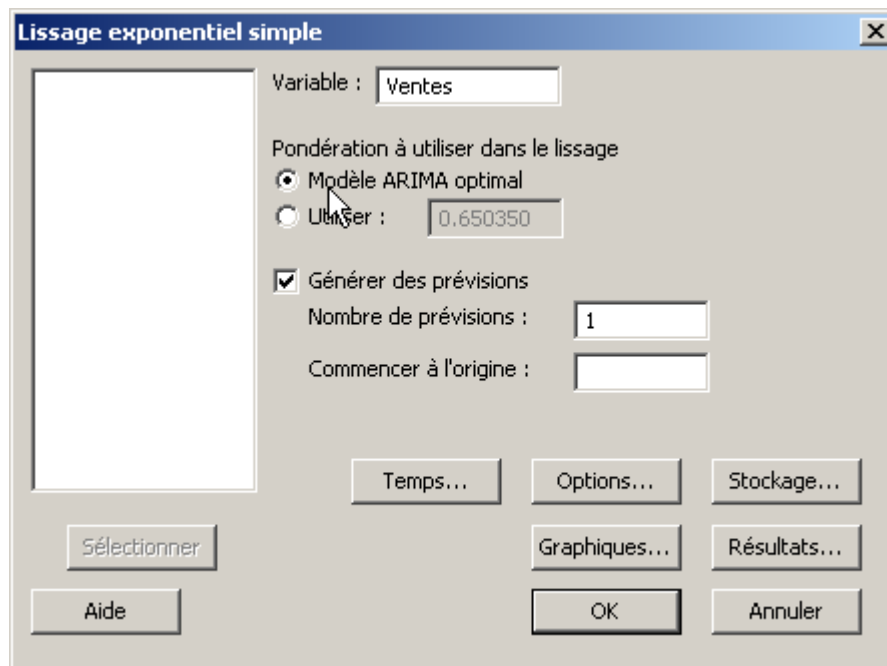
Commençons par vérifier le premier toujours avec le fichier *Smoothing.mpj*:



The screenshot shows the Minitab interface with a worksheet titled 'Feuille de travail 1 \*\*\*'. The worksheet contains a single column of data labeled 'Ventures' with the following values:

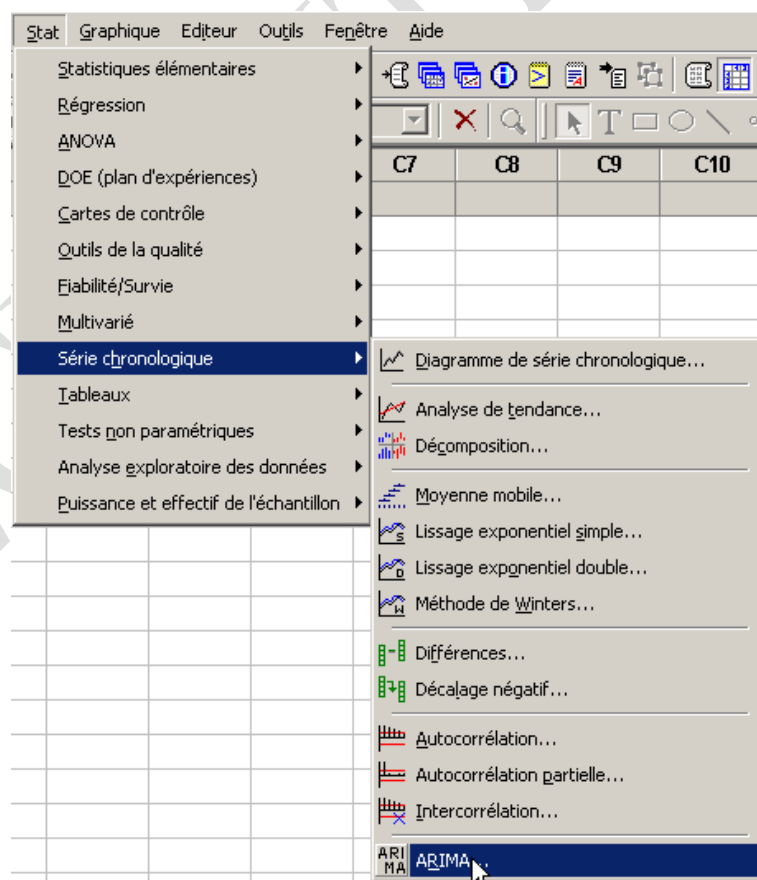
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
1	143													
2	152													
3	161													
4	139													
5	137													
6	174													
7	142													
8	141													
9	162													
10	180													
11	164													
12	171													
13	206													
14	193													
15	207													
16	218													
17	229													

Pour lequel nous avons obtenu comme prévision en cochant l'option ARIMA:



Point	Prévision	Inférieur	Supérieur
25	256.100	221.379	290.821

Maintenant, allons dans le menu **Stat/Séries chronologiques/ARIMA...**:



Nous avons alors:

et nous cliquons sur le bouton **Prévisions...**:

et nous obtenons dans la fenêtre de session:

```

Prévisions depuis la période 24

                               Limites 95%
Point  Prévision  Inférieur  Supérieur  Réel
  25    256.100    220.772    291.428

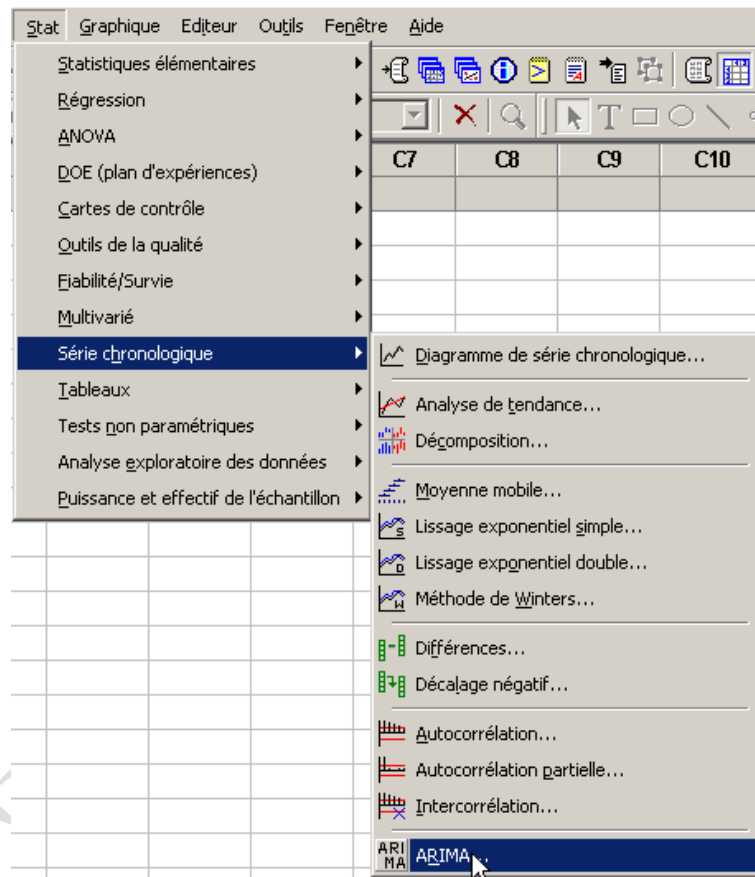
```

Soit exactement la même prévision mais pas avec les mêmes bornes (ce qui est très curieux...).

Faisons la même chose avec le lissage exponentiel double de Holt's pour lequel nous avons obtenu plus haut les prévisions suivante avec l'optimisation ARIMA:

Point	Prévision	Inférieur	Supérieur
21	229.289	195.882	262.695
22	233.667	196.211	271.124
23	238.046	196.115	279.977
24	242.425	195.716	289.134
25	246.804	195.097	298.511
26	251.183	194.317	308.049

Nous allons donc dans le menu **Stat/Séries chronologiques/ARIMA...**:



Nous mettons cette fois-ci:

ARIMA

Série :   Ajuster le modèle saisonnier

Période :

Autorégressif : Non saisonnier  Saisonnier

Différence :

Moyenne mobile :

Inclure un terme constant dans le modèle

Valeurs de départ pour les coefficients :

Sélectionner

Aide

Graphiques... Prévisions...

Résultats... Stockage...

OK Annuler

avec 6 prévisions:

ARIMA - Prévisions

Translation dans le futur :

Origine :

Stockage

Prévisions :

Limites inférieures :

Limites supérieures :

Sélectionner

Aide

OK Annuler

et nous validons par **OK** pour obtenir dans la fenêtre de session:

Point	Prévision	Limites 95%		Réel
		Inférieur	Supérieur	
21	229.289	194.295	264.282	
22	233.667	194.185	273.150	
23	238.046	193.828	282.265	
24	242.425	193.240	291.610	
25	246.804	192.437	301.171	
26	251.183	191.429	310.937	

soit exactement les mêmes prévisions mais pas avec les mêmes intervalles.

À part ce petit couac au niveau des intervalles, nous avons bien la "preuve" par la pratique que lissage exponentiel simple est un ARIMA(0,1,1) et le lissage exponentiel double de Holt's un ARIMA(0,2,2).

ÉCHANTILLON

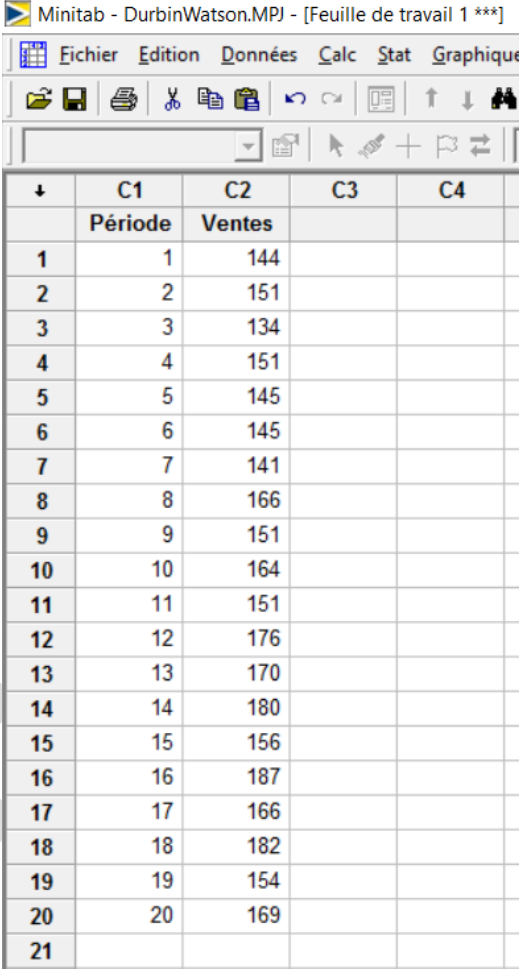


## 19.8. Exercice 207.: Test de Durbin-Watson

Minitab® Statistical Software 15.1.1.0

Nous avons introduit dans le cours théorique la statistique de Durbin-Watson. Nous n'avons pas fait de calculs à la main ou dans le tableur Microsoft Excel car c'est vraiment trop laborieux. Notre seul propose ici va être de voir comment R affiche la statistique et comparer cette dernière à Minitab.

Nous ouvrons pour cela le fichier *DurbinWatson.mpj*:

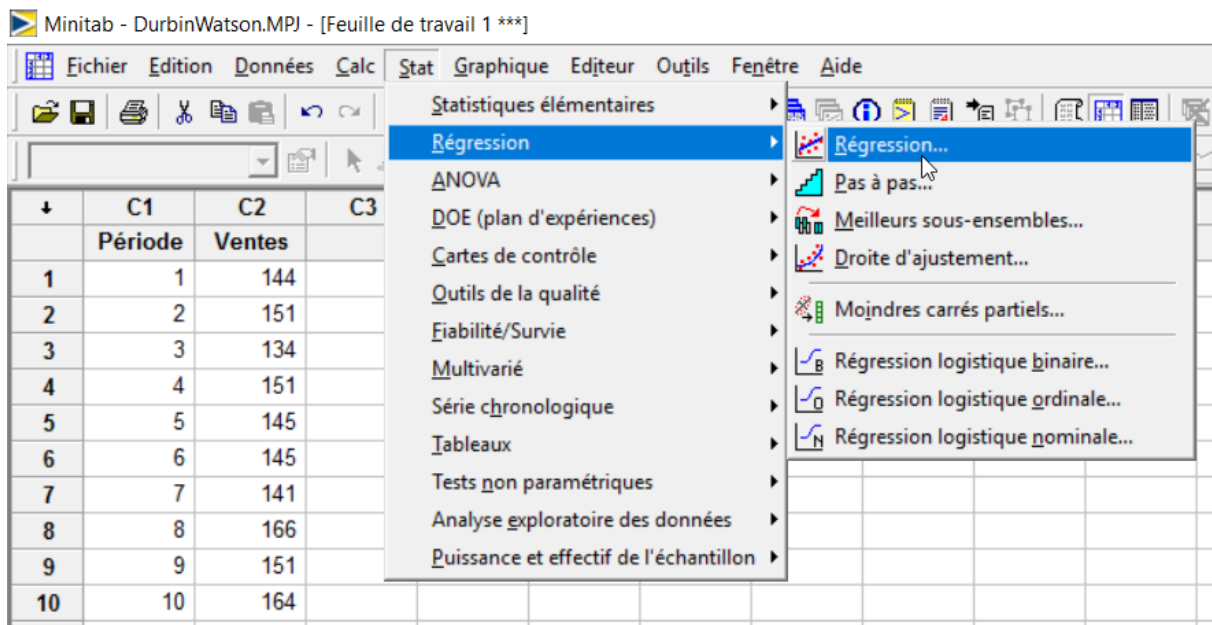


Minitab - DurbinWatson.MPJ - [Feuille de travail 1 \*\*\*]

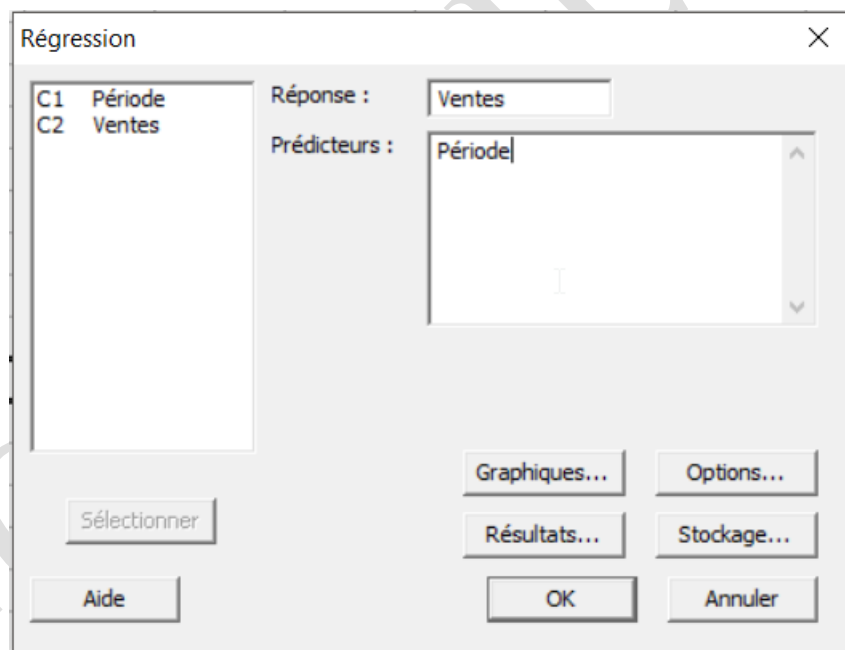
Fichier Edition Données Calc Stat Graphique

	C1	C2	C3	C4
	Période	Ventes		
1	1	144		
2	2	151		
3	3	134		
4	4	151		
5	5	145		
6	6	145		
7	7	141		
8	8	166		
9	9	151		
10	10	164		
11	11	151		
12	12	176		
13	13	170		
14	14	180		
15	15	156		
16	16	187		
17	17	166		
18	18	182		
19	19	154		
20	20	169		
21				

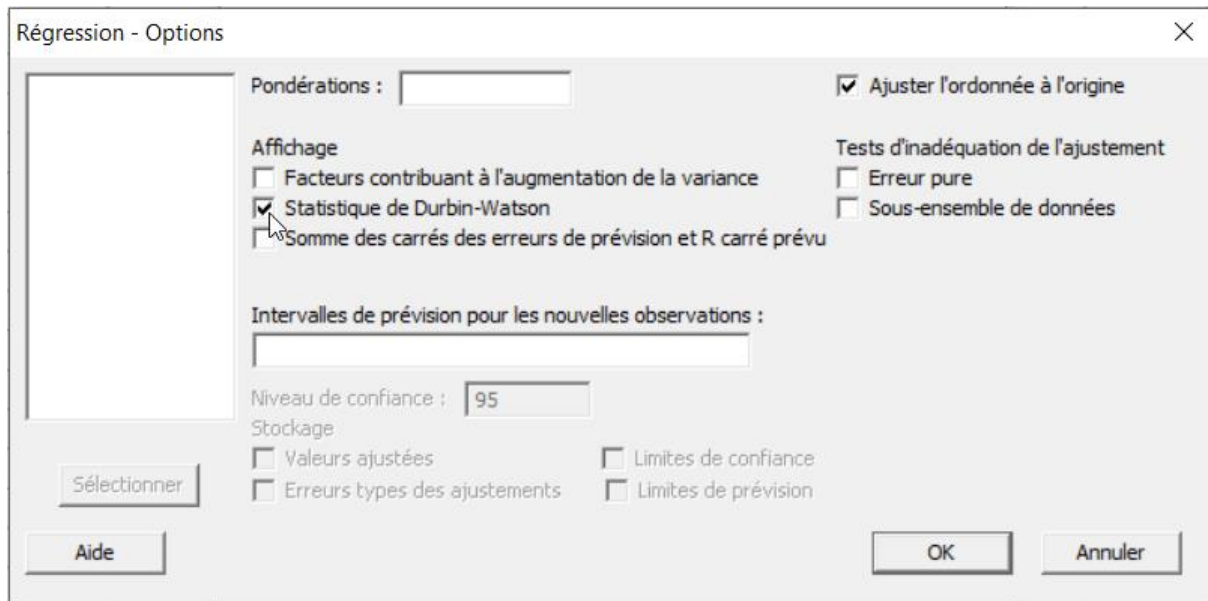
Nous allons dans le menu **Stat/Régression/Régression...**:



Nous remplissons les champs comme il convient:



et nous cliquons sur le bouton **Options...**:



pour y activer **Statistique de Durbin-Watson**. Nous validons deux fois par **OK** pour obtenir:

————— 2017-02-05 03:13:37 PM —————

Bienvenue dans Minitab, appuyez sur F1 pour obtenir de l'aide.

**Analyse de régression : Ventes en fonction de Période**

L'équation de régression est  
 Ventes = 141 + 1.77 Période

Prédicteur	Coeff	Coef ErT	T	P
Constante	140.542	5.082	27.66	0.000
Période	1.7722	0.4242	4.18	0.001

S = 10.9393 R carré = 49.2 % R carré (ajust) = 46.4 %

Analyse de variance

Source	DL	Somme des carrés	CM	F	P
Régression	1	2088.5	2088.5	17.45	0.001
Erreur résiduelle	18	2154.0	119.7		
Total	19	4242.5			

Observations aberrantes

Observation	Période	Ventes	Valeur ajustée	Valeur résiduelle	Valeur résiduelle normalisée
19	19.0	154.00	174.21	4.36	-2.01R

R indique une observation ayant une valeur résiduelle normalisée importante

Statistique de Durbin-Watson = 2.71215

Comme nous pouvons le voir:

1. On ne sait pas vraiment si la statistique de Durbin-Watson est calculée en bilatérale ou non
2. On a aucune indication quant au fait qu'on doit "accepter" ou rejeter l'hypothèse nulle excepté si on connaît par cœur les tables de DW...

Bref l'équipe de développement de Minitab aurait ici franchement pu faire mieux!

ÉCHANTILLON

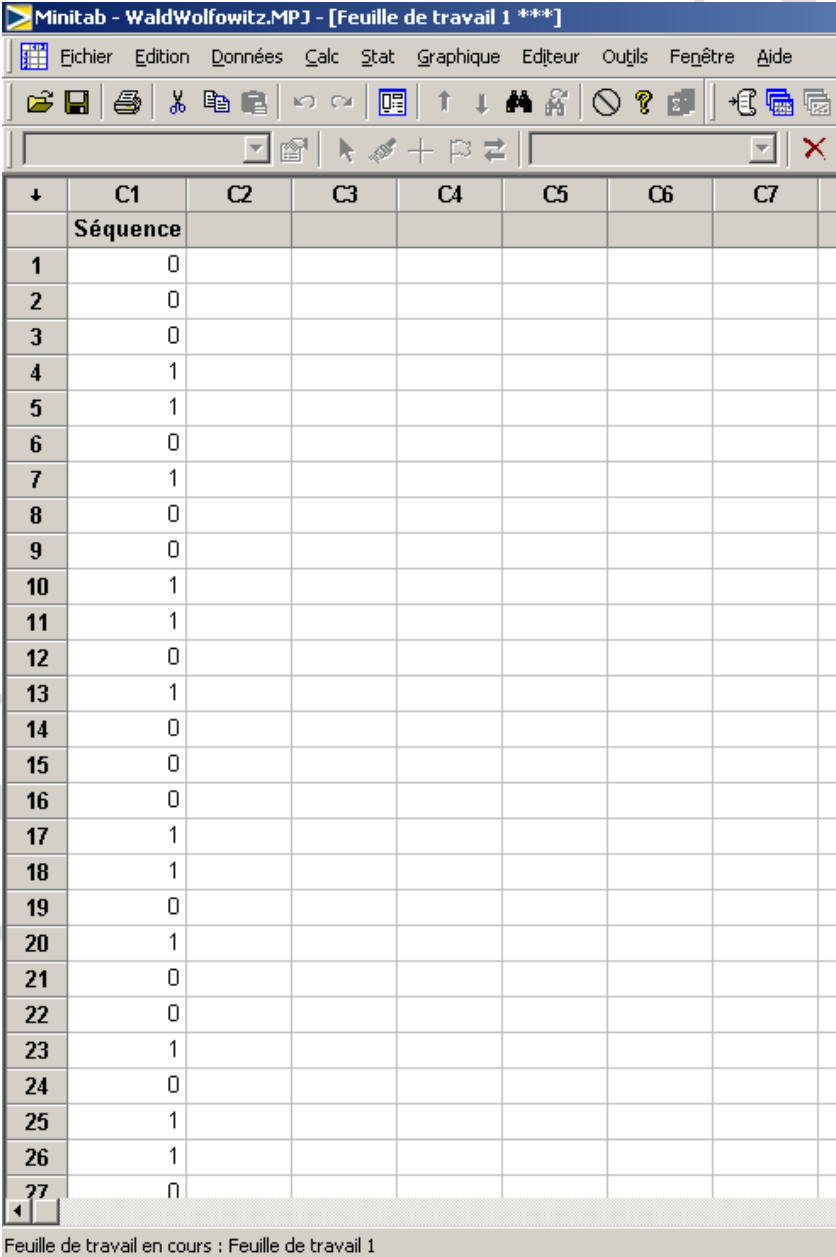
## 19.9. Exercice 208.: Test des suites (test de Wald-Wolfowitz)

Minitab® Statistical Software 15.1.1.0

Comme à l'habitude le but va être de vérifier que le logiciel utilise les résultats démontrés dans le cours théorique avec un exemple implicitement lié à une série temporelle (les "0" étant des "baisses" et les "1" étant des valeurs "haussières"). La question étant de savoir si les séquences peuvent être considérées comme aléatoires ou non statistiquement parlant.

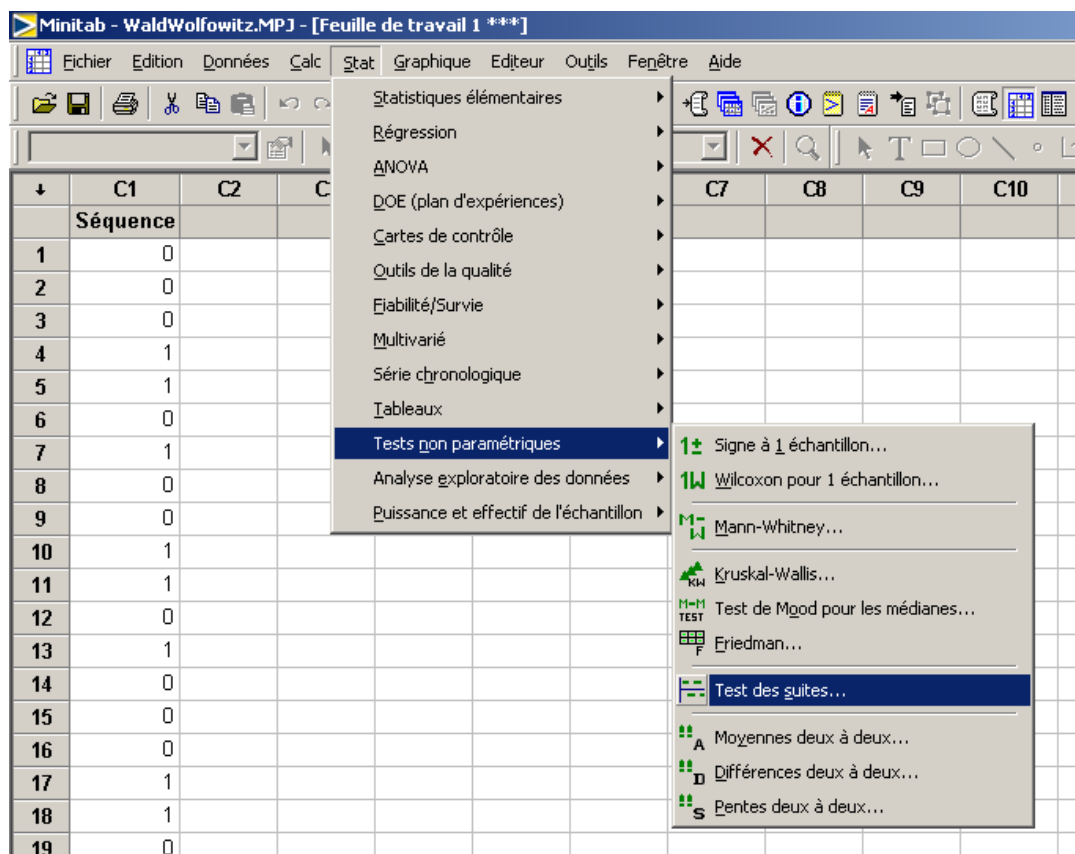
Nous devrions aussi retrouver les calculs faits à la main et à l'aide de Microsoft Excel.

Ouvrez le fichier. Vous aurez alors 37 lignes de données:

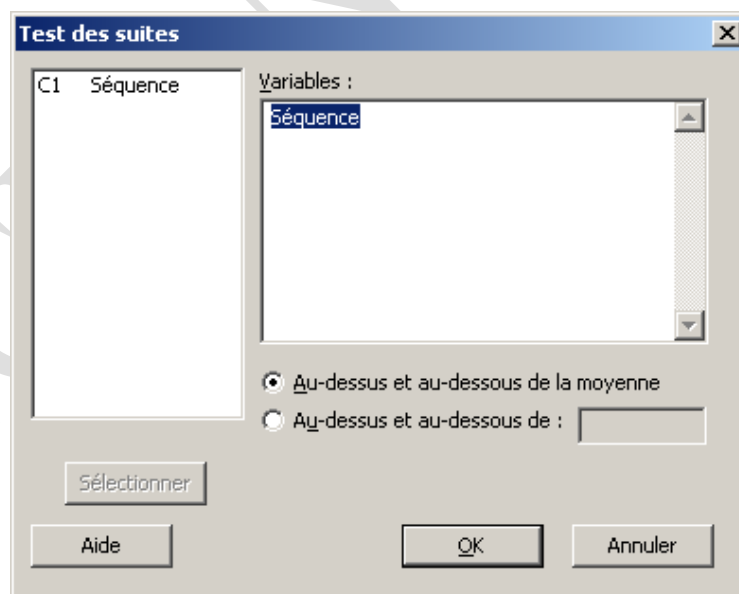


	C1	C2	C3	C4	C5	C6	C7
	Séquence						
1	0						
2	0						
3	0						
4	1						
5	1						
6	0						
7	1						
8	0						
9	0						
10	1						
11	1						
12	0						
13	1						
14	0						
15	0						
16	0						
17	1						
18	1						
19	0						
20	1						
21	0						
22	0						
23	1						
24	0						
25	1						
26	1						
27	0						

Nous allons ensuite dans le menu **Stat/Tests non paramétriques/Test de suites...**:

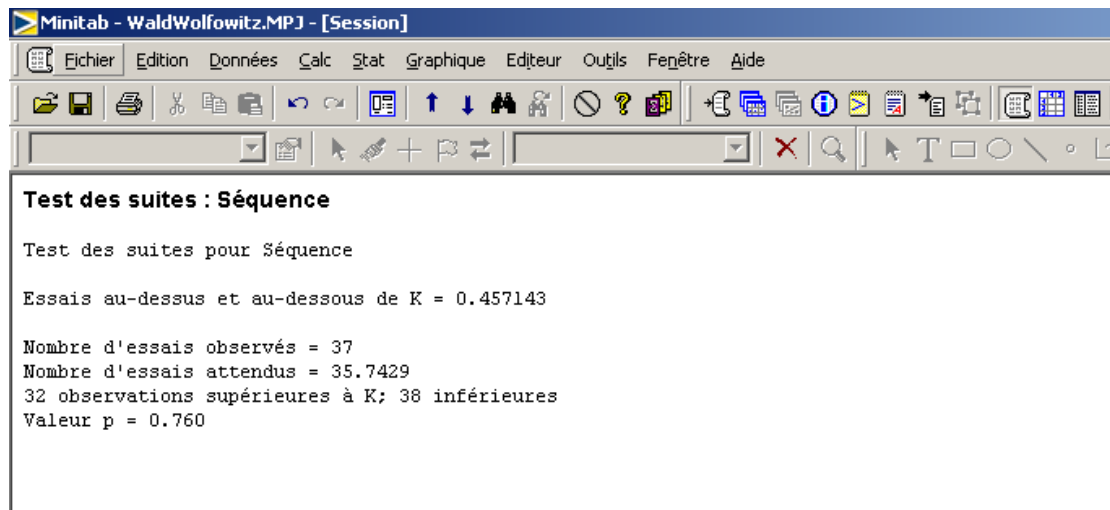


Nous avons alors:



Remarquez la possibilité en plus d'avoir le test, le fait que Minitab va compter pour nous le nombre de points au-dessus et en-dessous de la moyenne (c'est accessoire mais bon...).

Nous validons par **OK** et obtenons:



Nous retrouvons donc les mêmes valeurs que celles calculées à la main. Ce qui conforte nos démonstrations mathématiques et nous permet d'arriver à la même conclusion.

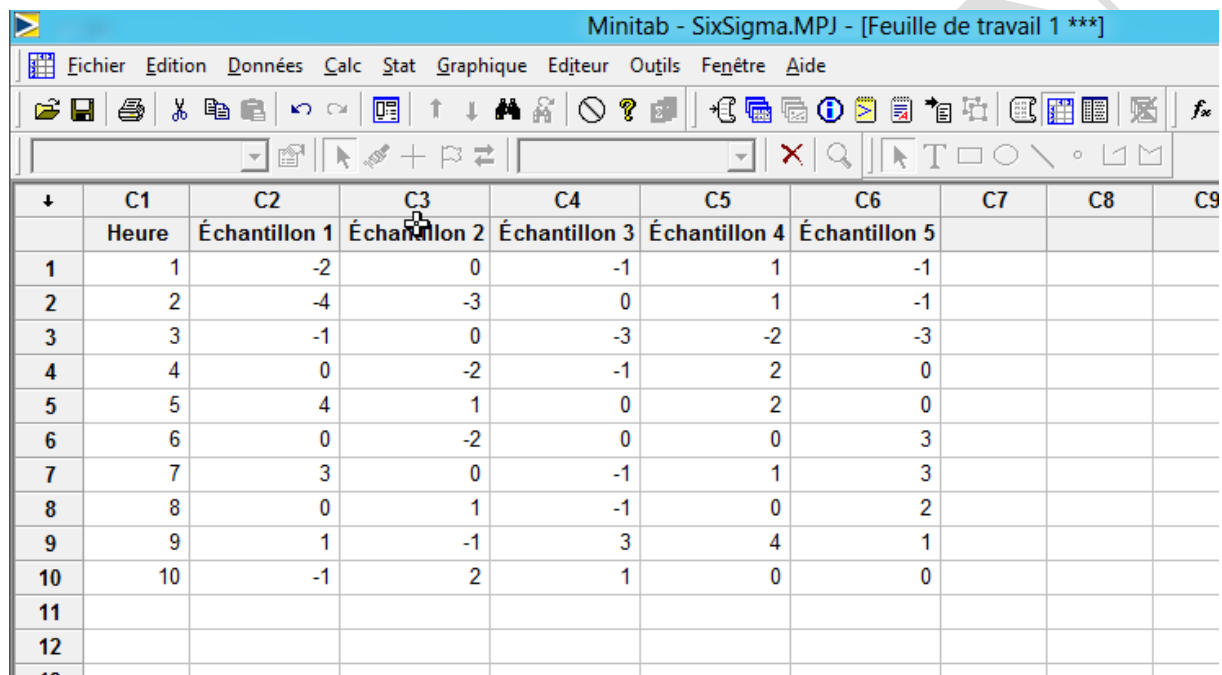
## 20. Macros/Scripting

### 20.1. Exercice 209.: Convertir des actions en macro

Minitab® Statistical Software 15.1.1.0

Nous allons voir ici comment il est facile dans Minitab de convertir une série de manipulations faites avec la souris en une macro.

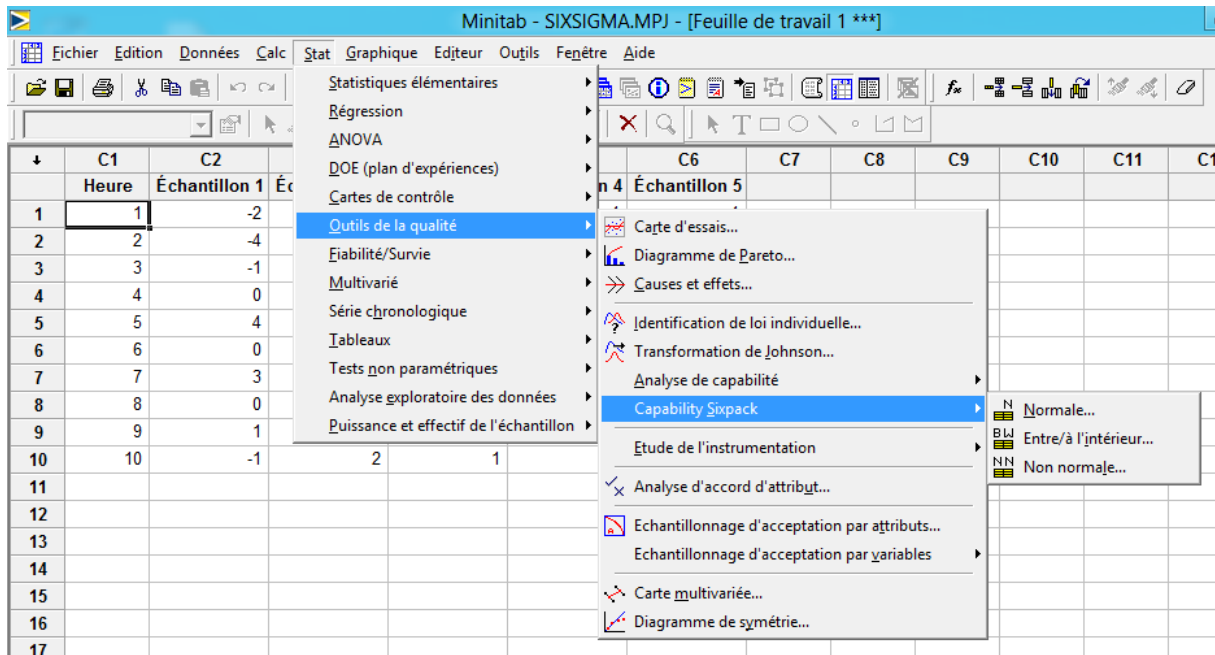
Nous partons du même jeu de données que dans le cours théorique:



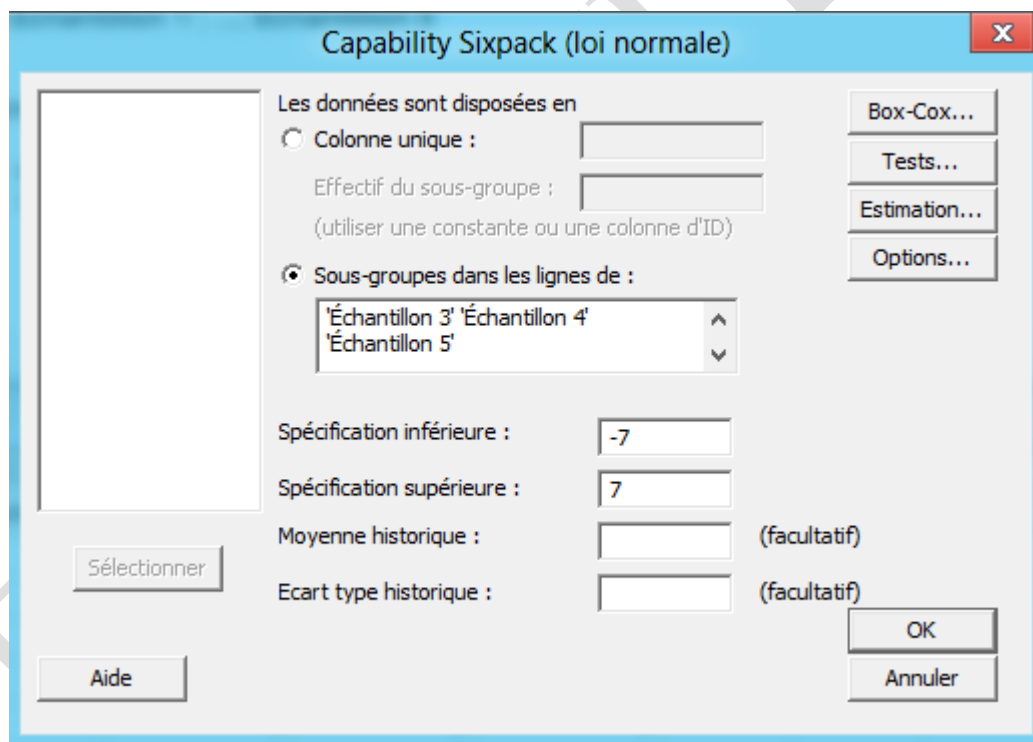
	C1	C2	C3	C4	C5	C6	C7	C8	C9
	Heure	Échantillon 1	Échantillon 2	Échantillon 3	Échantillon 4	Échantillon 5			
1	1	-2	0	-1	1	-1			
2	2	-4	-3	0	1	-1			
3	3	-1	0	-3	-2	-3			
4	4	0	-2	-1	2	0			
5	5	4	1	0	2	0			
6	6	0	-2	0	0	3			
7	7	3	0	-1	1	3			
8	8	0	1	-1	0	2			
9	9	1	-1	3	4	1			
10	10	-1	2	1	0	0			
11									
12									

Ensuite, nous allons dans la menu **Stat/Outils de la qualité/Capability Sixpack/Normale...**:

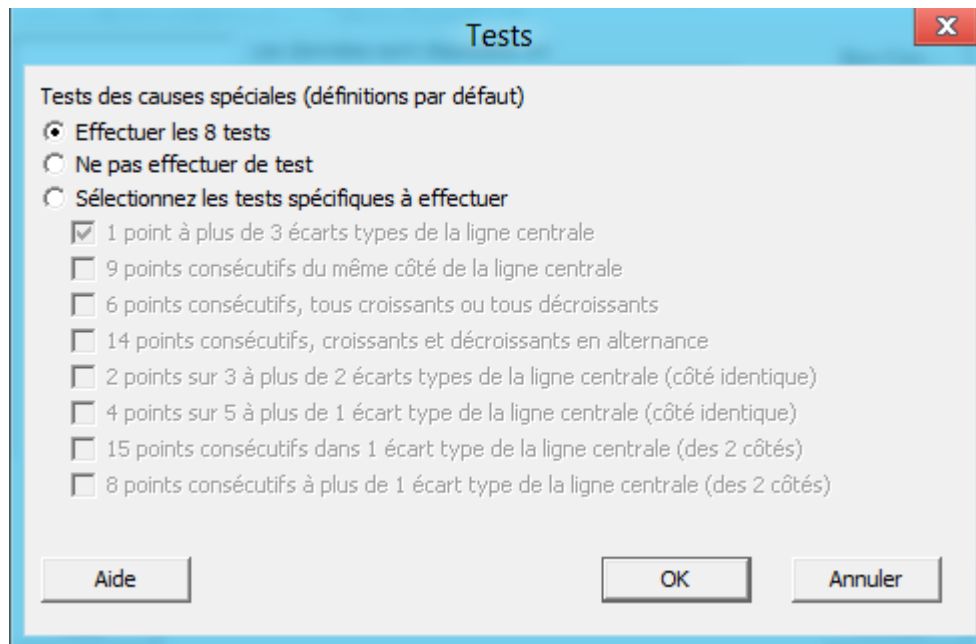




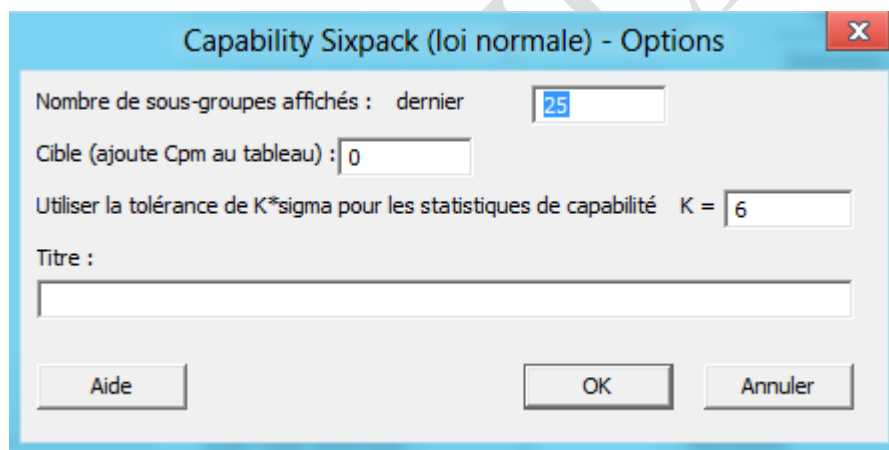
Et nous mettons alors:



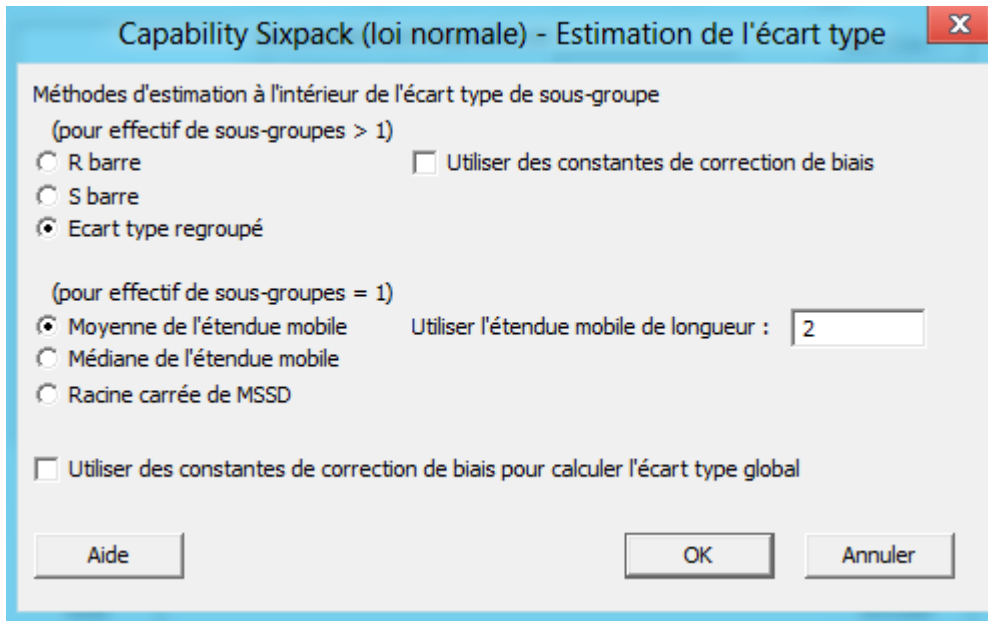
Dans le bouton **Tests...** nous prenons tous les tests de la WECO:



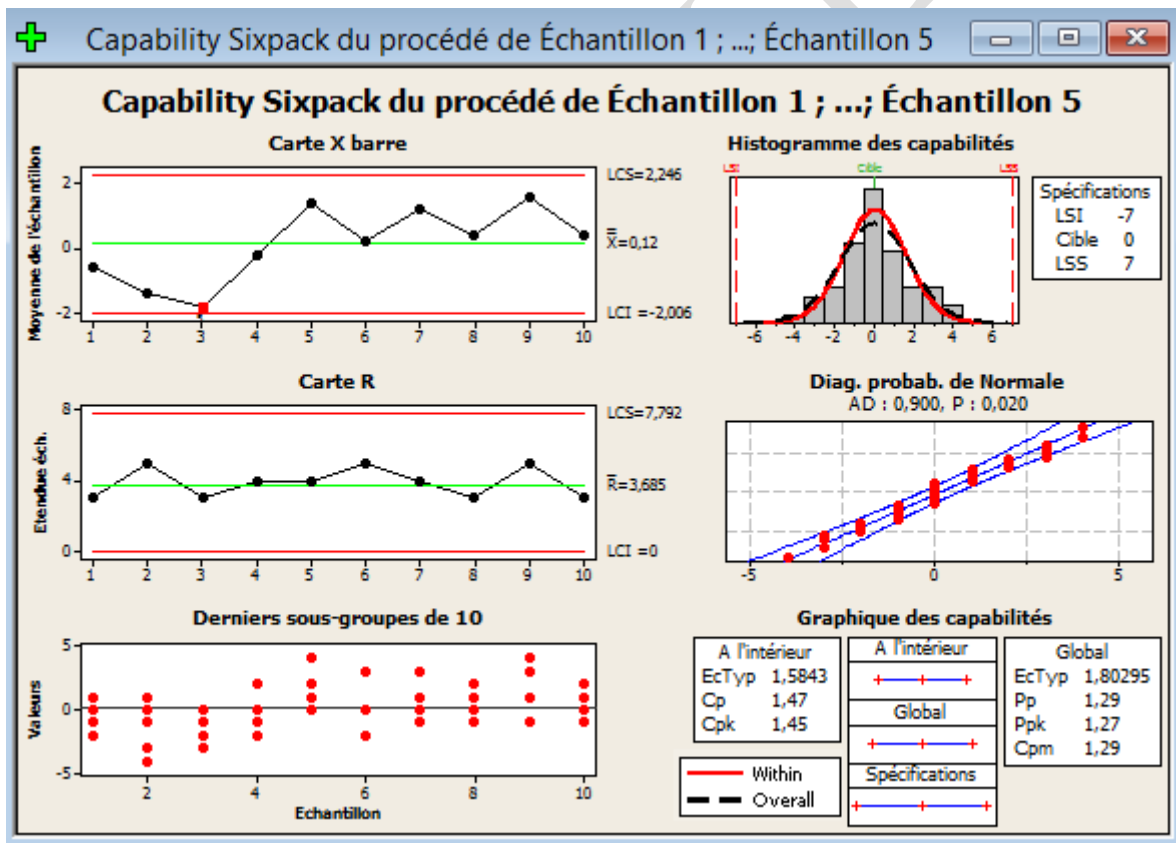
Dans le bouton **Options...** nous prenons (toujours pour être conforme aux calculs faits à la main dans le cours théorique):




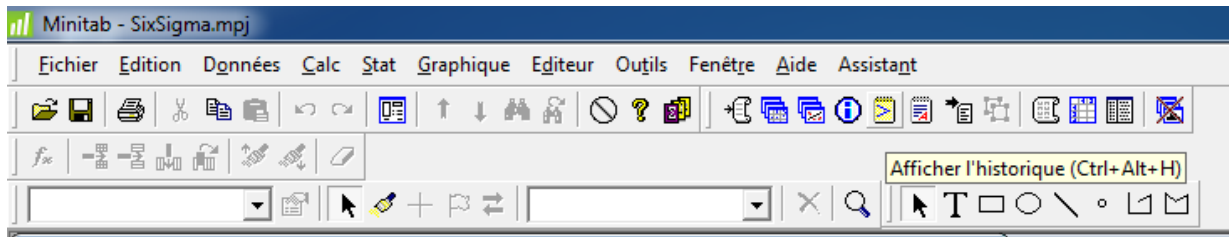
Comme les échantillons sont petit (inférieur à 10), nous devrions prendre la carte R barre mais si nous voulons retrouver les calculs faits dans le cours théorique, il nous faudra laisser l'option **Ecart Type regroupé** active dans le bouton **Estimation...** (car dans Minitab les indices de Capacité et non pas seulement la carte de contrôle dépendent de la manière dont l'écart-type est calculé):



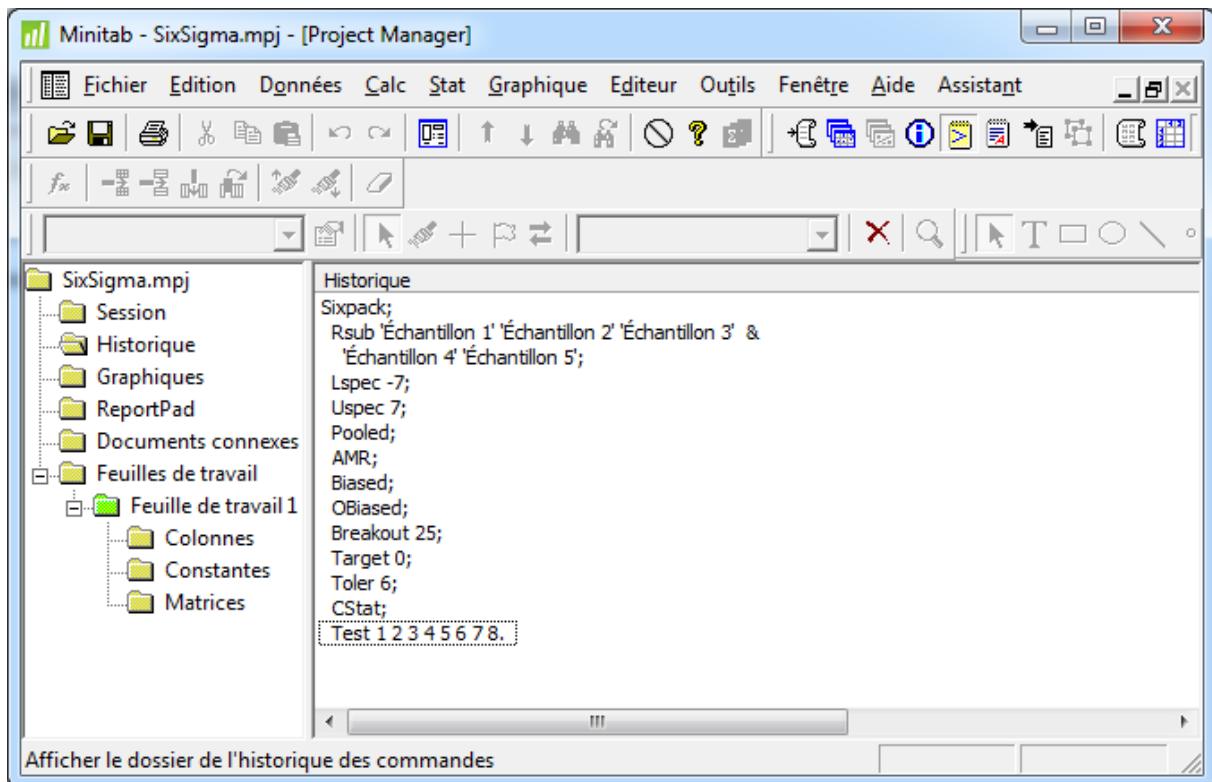
Nous validons tout ce beau monde par **OK** pour obtenir:



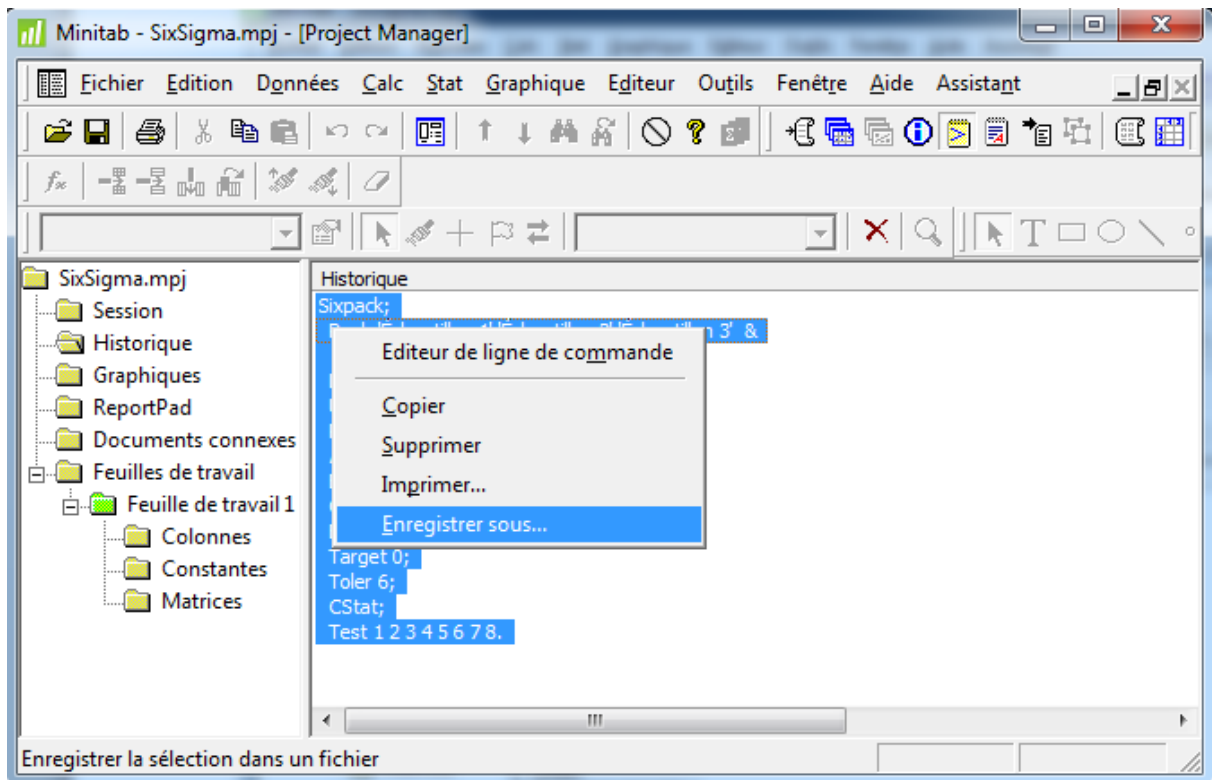
Une fois ceci fait nous cliquons sur le bouton Afficher l'historique  visible sur la barre d'outils ci-dessous:



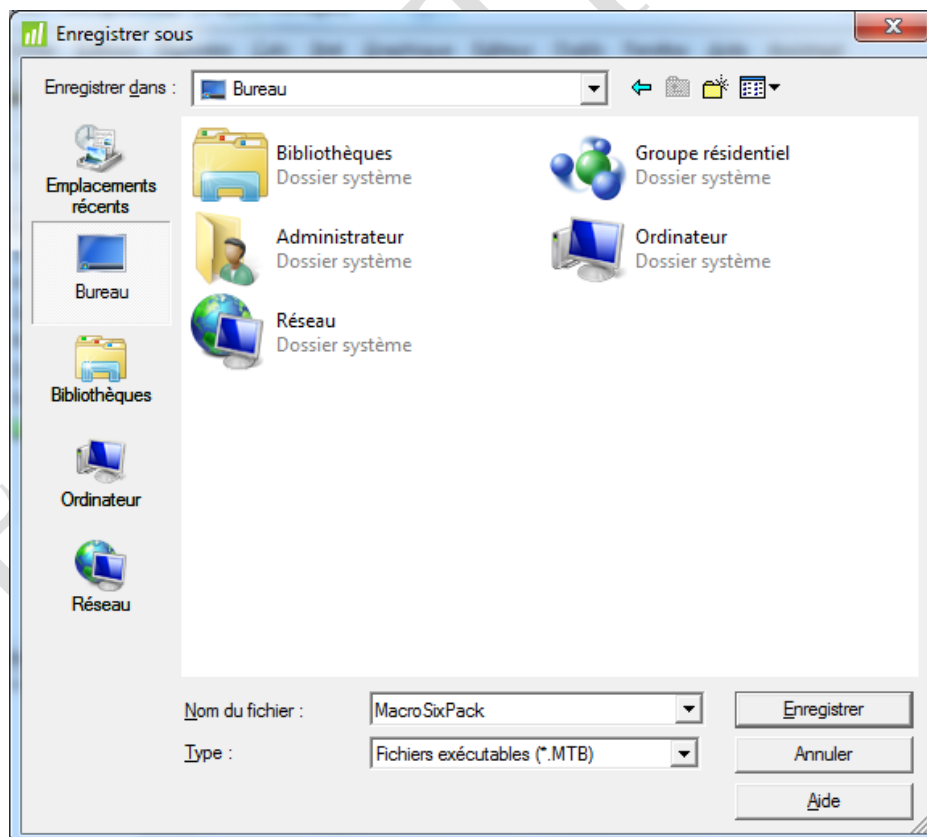
Apparaît alors l'historique des commandes suivant:



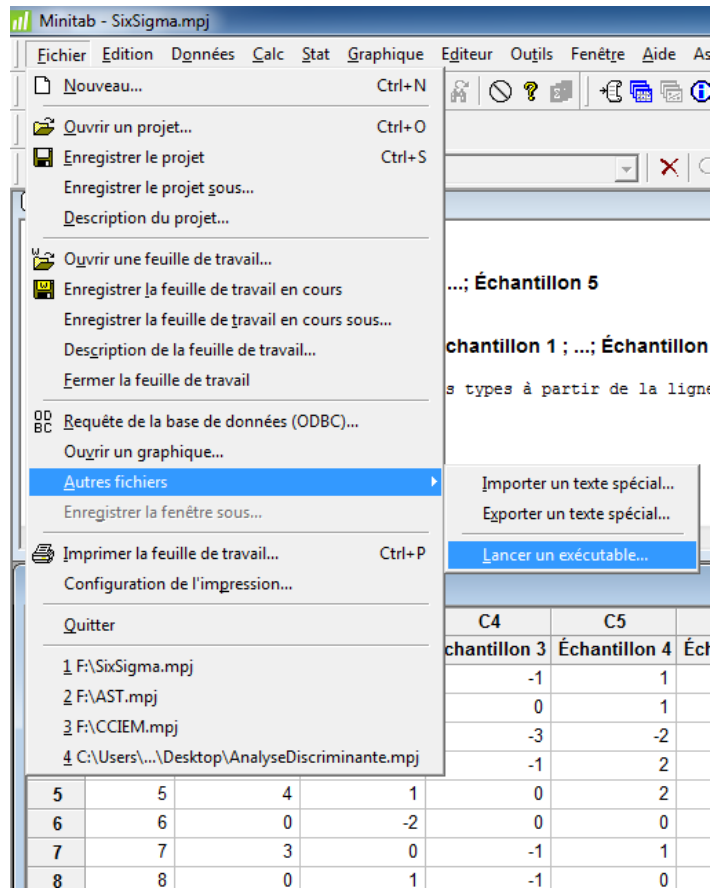
Nous faisons un **Ctrl+A** et un clic-droit pour choisir l'option **Enregistrer sous...**:



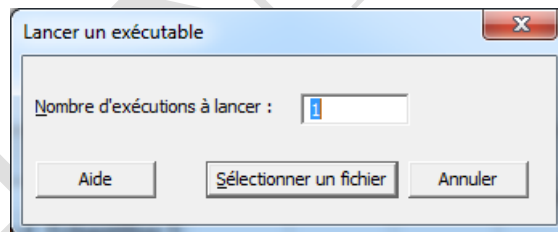
Nous prenons garde à choisir comme type de fichier un \*.MTB et validons par **Enregistrer**:



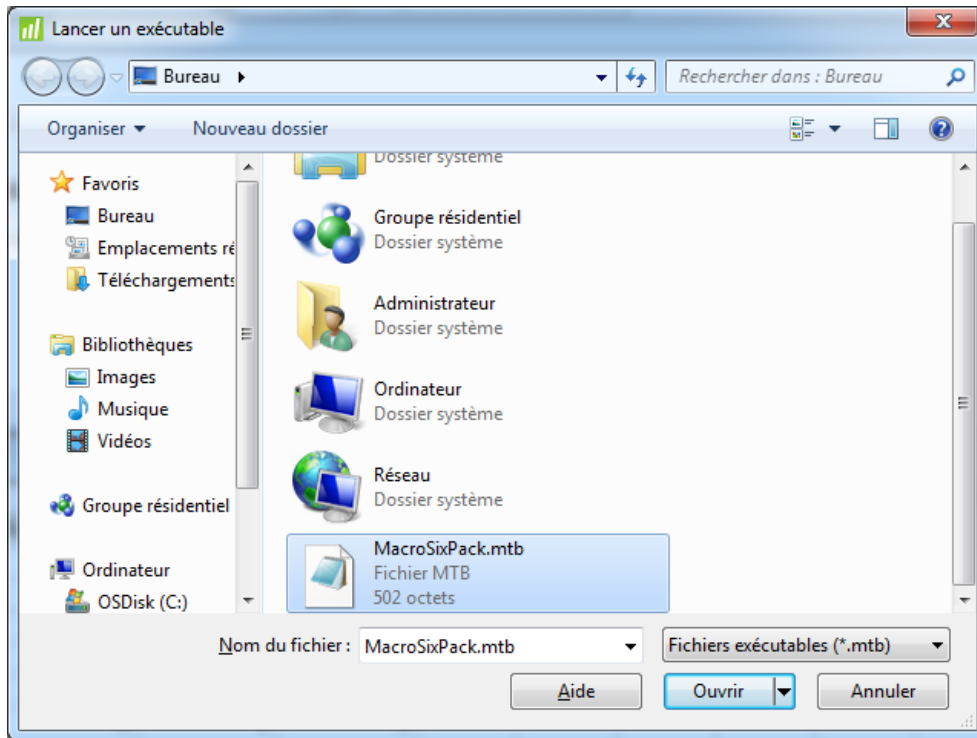
Une fois ceci fait, nous pouvons ré exécuter plus tard cette macro en allant dans le menu **Fichier/Autres fichiers/Lancer un exécutable...**:



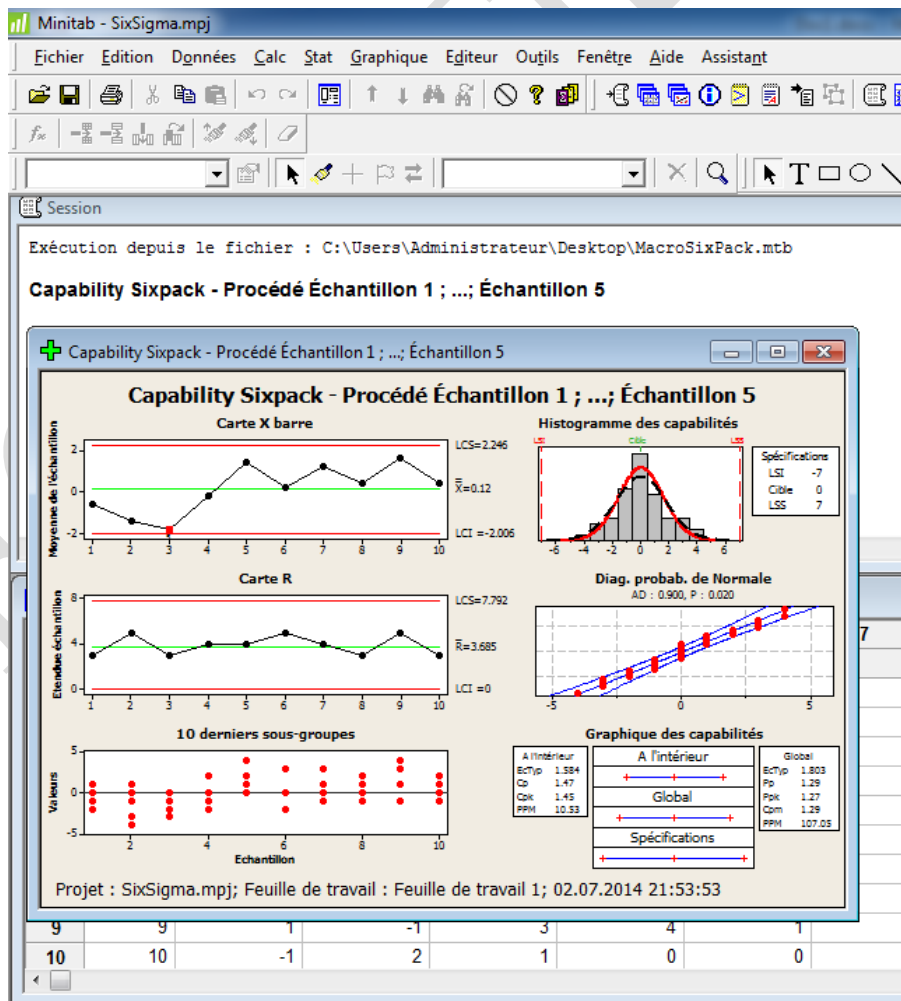
Il vient alors:



Nous cliquons sur **Sélectionner un fichier** et il vient:



et au moment que nous cliquons sur **Ouvrir**, la macro est exécutée et nous obtenons bien:



ÉCHANTILLON



## 20.2. Exercice 210.: Intégration d'une macro téléchargée

Minitab® Statistical Software 15.1.1.0

Nous avons démontré dans le cours de statistique théorique un certain nombre de tests très importants qui n'existent pas encore à ce jour dans Minitab.

Le but ici est de faire un exemple d'intégration d'une macro prise sur le site internet de Minitab et qui permet d'exécuter le test de McNemar que nous avons démontré dans le cours théorique pour vérifier si nous retombons sur les mêmes résultats que ceux calculés avec Microsoft Excel.

Le lien direct vers cette macro particulière se trouve ici:

<http://www.minitab.com/en-US/support/macros/default.aspx?action=code&id=130>

```

macro
#####
#
# Name:          MCNEMAR.MAC
# Version:       Release 15 and 16
# Written by:    Eli Walters and Eduardo Santiago 7/23/2007
# Modified:      6/1/2009 and 10/14/2011
#
# This macro performs McNemar's test on either (1) a 2x2 table
# stored in two columns of the active worksheet or (2) two columns
# of raw data in the active worksheet.  Input constant f indicates
# the form of the data (1 = table, 2 = raw data).  For example, if a
# 2x2 table is stored in the first two cells of columns C1 and C2,
# then type:
#
# %mcnemar 1 C1 C2;
# correction;
# confidence 0.99.
#
# The output will provide the chi-square statistic with a continuity
# correction (to provide the standard chi-square statistic omit the
# subcommand for correction), p-value for the test and a 99% confidence
# interval for the difference of the observed marginal proportions
# (if the confidence subcommand is omitted a 95% confidence interval will
# be created by default).
#
#####
# Neither Minitab, Inc. nor the author(s) of this MACRO makes any claim
# of or offers any Warranty whatsoever with regard to the accuracy of
# this MACRO or its suitability for use.  Minitab, Inc. and the author(s)
# of this MACRO each hereby disclaims any Warranty and/or liability with
# respect thereto.
#
#####

mcnemar f x y;
correction;
confidence conf.

mcolumn x y x2 y2 labels xlab ylab cellct deltap lowci upci k obserx obsery dummy civals cint
mconstant f a b c d delta chi k1 pval xval yval bc conf typex typey ct phr1 phr2 phr3 phr4
phr5 tconf tlowci tupci xlen ylen val prob

default conf=0.95
mreset
noecho
brief 0
notitle

```

```

if conf >= 1 or conf <= 0
  note
  note **Error** The confidence level is a number between 0 and 1, exclusive.
  note
  exit
endif

if f=1          # Data has been entered in the worksheet as a 2x2 table

# Check to verify that the data is numeric

  dtype x typex
  dtype y typey
  n x xval
  n y yval

  brief
  if typex = 0 or typey = 0 or xval ne 2 or yval ne 2
    note
    note **Error** The columns for this table need to be numeric and need to have only two
rows each.
    note
    exit
  endif

  let a = x(1)
  let b = y(1)
  let c = x(2)
  let d = y(2)

elseif f=2      # Raw data has been entered in the worksheet

  stats;
  by x;
  gval xlab.

  stats;
  by y;
  gval ylab.

  n xlab xval
  n ylab yval

  n x xlen
  n y ylen

  brief

  if xlen ne ylen
    note
    note **Error** Columns must be of equal length.
    note
    exit
  endif

  if xval ne 2 or yval ne 2
    note
    note **Error** Each column must have exactly two distinct levels.
    note
    exit
  endif

  if xlab ne ylab
    note
    note **Error** Each column must have the same two distinct levels.
    note
    exit
  endif
  brief 0

  # If the column lengths are not the same at this point, the macro will be automatically
terminated.

  dtype x typex
  dtype y typey

```

```

if typex = 0 and typey = 0
  copy x y x2 y2;
  Exclude;
  Where "x = "" "" or y = "" """.
elseif typex ne 0 and typey ne 0 and typex = typey
  copy x y x2 y2;
  Exclude;
  Where "x = '*' or y = '*'".
endif

n x2 obserx
n y2 obsery

brief 1
if obserx ne obsery
  note
  note **Error** Since the data is matched, each column must have the same number of
observations.
  note
  exit
endif
brief 0

# This assumes the raw data is text or numeric.

let ct = count(x2)
Set dummy
1( 1 : 1 / 1 )ct
End.
Statistics dummy;
By x2 y2;
Sums cellct.
let a = cellct(1)
let b = cellct(2)
let c = cellct(3)
let d = cellct(4)

else
  brief 1
  note
  note **Error** The first input parameter can only take values of 1 or 2.
  note
  exit
endif

let bc=b+c

brief
if bc < 10
  note
  note **Warning** Test statistic may not be approximated well by the chisquare
distribution in this case.
  note
  note A sign test may be more appropriate.
  note
endif
brief 0

if correction = 0
  let chi = (b-c)**2/(b+c)
else
  let chi = ((abs(b-c)-1)**2)/(b+c)
endif

cdf chi k1;
chisquare 1.

let pval = 1-k1

let delta = (c-b)/(a+b+c+d)
let prob = 1-(1-conf)/2
invcdf prob k;
normal 0 1.
let lowci = delta - k*sqrt(((a+b+c+d)*(b+c) - (b-c)**2)/(a+b+c+d)**3)
let upci = delta + k*sqrt(((a+b+c+d)*(b+c) - (b-c)**2)/(a+b+c+d)**3)

brief 1
if correction = 0

```

```

    mtitle "McNemar's Test for Marginal Homogeneity"
else
    mtitle "McNemar's Test for Marginal Homogeneity with Continuity Correction"
endif

let conf=conf*100

let civals[1]=conf
let civals[2]=lowci
let civals[3]=upci

text civals civals

let tconf=left(civals[1],2)
let tlowci=civals[2]
let tupci=civals[3]

kkset phr1 "The "
kkset phr2 "% CI for delta is: "
kkset phr3 "("
kkset phr4 ","
kkset phr5 ")."

kkcat phr1 tconf phr1
kkcat phr1 phr2 phr1
kkcat phr3 tlowci val
kkcat val phr4 val
kkcat val tupci val
kkcat val phr5 val
kkcat phr1 val val
let cint = val

name chi "Chi-Square Statistic:"
name pval "P-value:"

print chi pval
write cint


write "where delta represents the difference of the observed marginal proportions."
endmacro

```

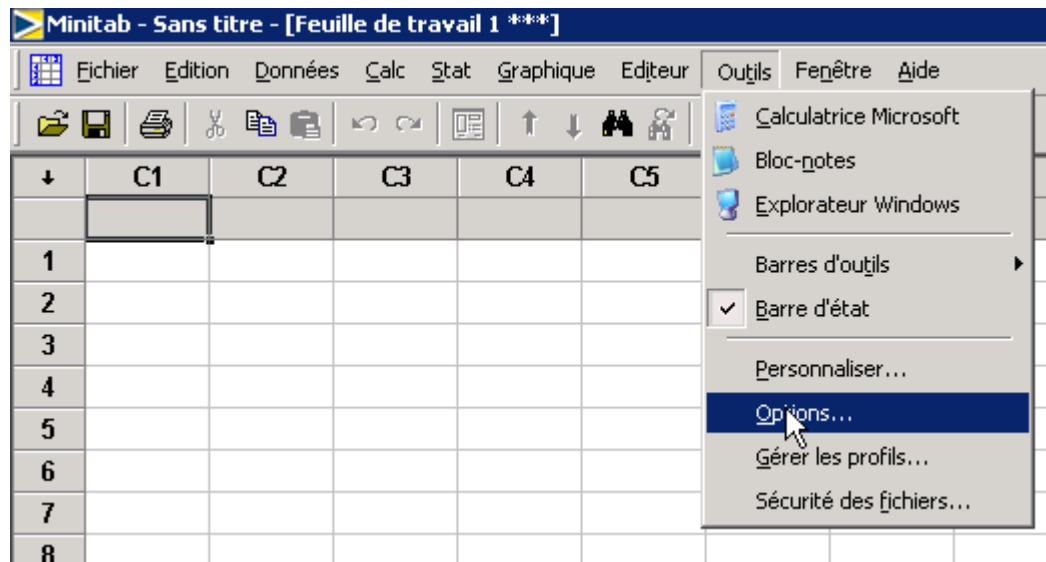
Cependant, pour rationaliser le travail, le mieux est de télécharger le fichier \*.zip contenant toutes les macros proposées par Minitab en suivant la procédure indiquée:

## Télécharger les macros

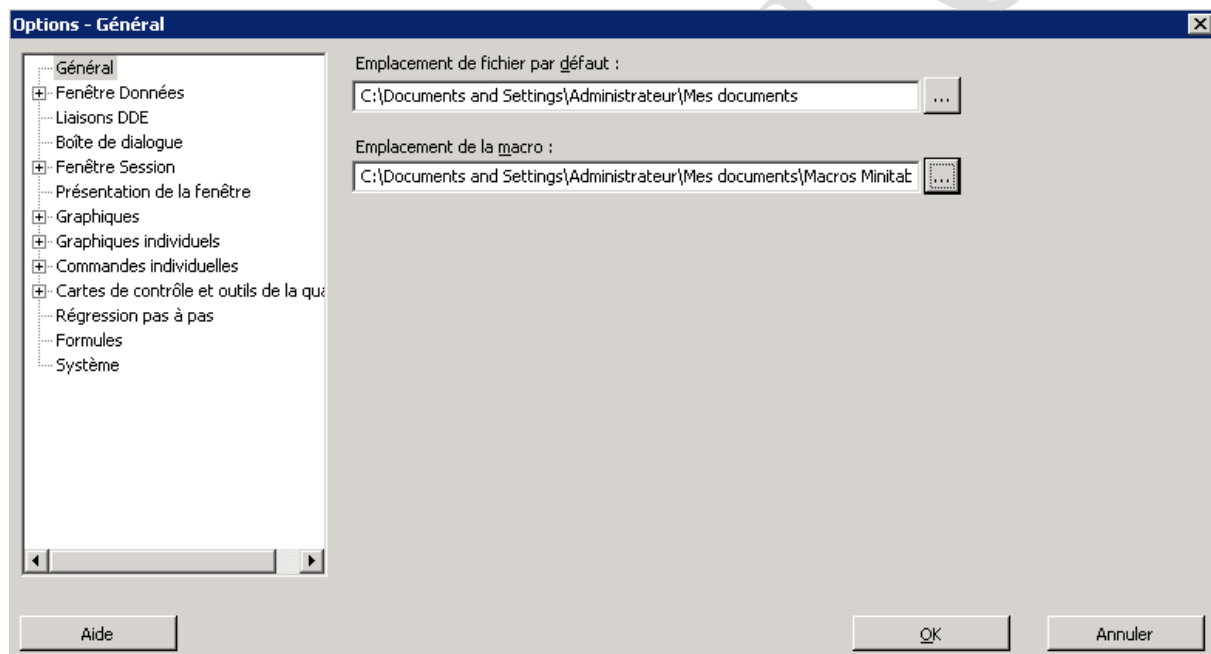
### Instructions de téléchargement

1. Dans le dossier **Mes documents**, créez un dossier où vous enregistrerez les fichiers de macros. Appelez ce nouveau dossier **Macros Minitab**.
2. Cliquez sur **Télécharger maintenant** .
3. Enregistrez le fichier zip dans le dossier **Macros Minitab** créé à l'étape 1.
4. Extrayez ou décompressez le contenu du fichier zip dans le dossier créé à l'étape 1.
5. Lancez Minitab.
6. Sélectionnez **Outils > Options > Général**. Dans **Emplacement de la macro**, accédez au dossier **Macros Minitab** que vous avez créé à l'étape 1. Minitab recherchera d'abord les macros dans ce dossier lorsque vous lancerez une macro depuis Minitab.
7. Si des macros sont enregistrées dans d'autres dossiers, pensez à les déplacer dans ce dossier **Macros Minitab**.

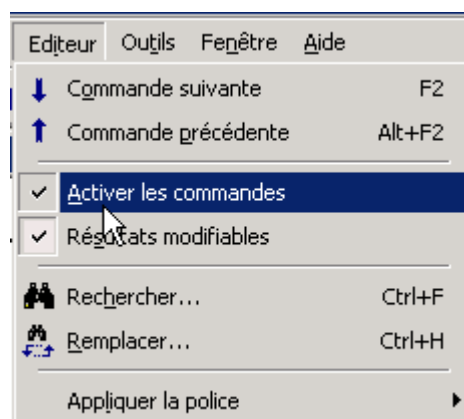
Donc nous faisons ce qui est écrit à partir de l'étape 6 (les étapes précédentes n'étant que l'application des bases élémentaires de Windows):



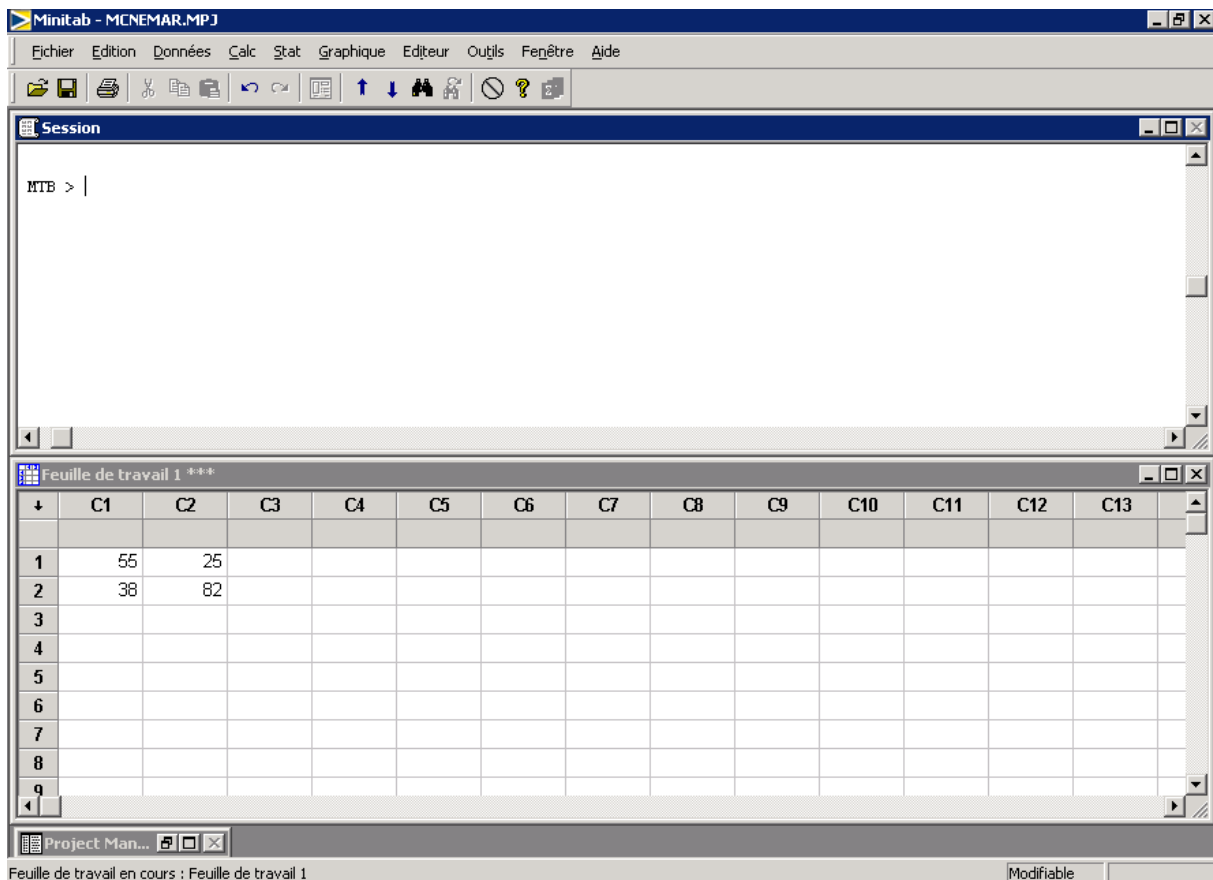
et:



Ensuite, nous allons de le menu **Editeur/Activer les commandes**:



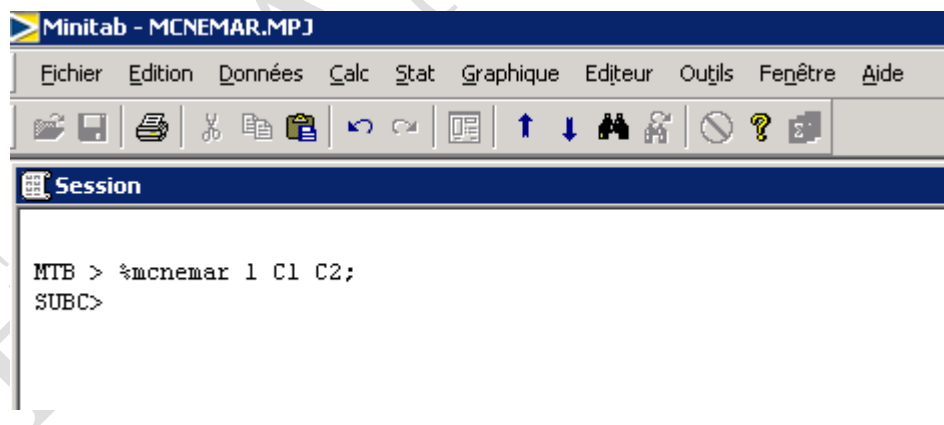
Nous avons alors **MTB>** qui apparaît:



The screenshot shows the Minitab interface with the Session window open. The Session window displays the prompt `MTB > |`. Below the Session window is a data table with columns C1 through C13 and rows 1 through 9. The data in the table is as follows:

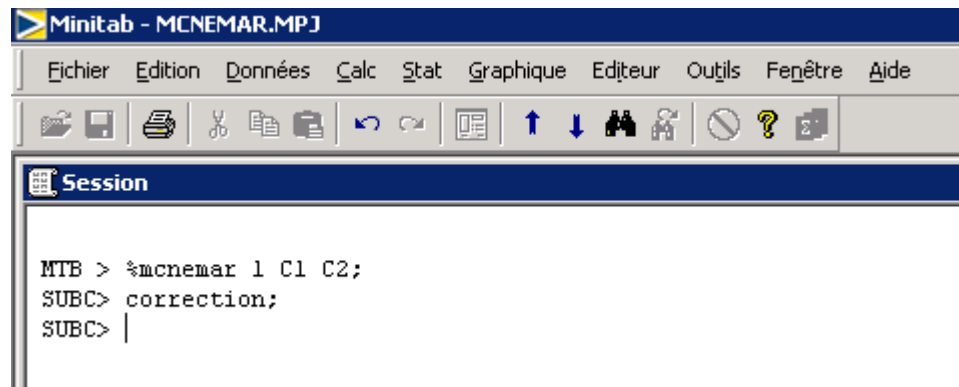
	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
1	55	25											
2	38	82											
3													
4													
5													
6													
7													
8													
9													

Nous tapons dans le cas qui nous intéresse d'abord `%mcnemar 1 C1 C2;` et nous validons par entrée:



The screenshot shows the Minitab interface with the Session window open. The Session window displays the command `MTB > %mcnemar 1 C1 C2;` and the response `SUBC>`.

et ensuite nous tapons **correction;** et nous validons pas entrée (SUBS signifie subcommand):



et nous finissons avec **confidence 0.95**. puis nous validons par entrée:

```

Minitab - MCNEMAR.MPJ
Fichier Edition Données Calc Stat Graphique Editeur Outils Fenêtre Aide
Session
MTB > %mcnemar 1 C1 C2;
SUBC> correction;
SUBC> confidence 0.95.
Exécution depuis le fichier : C:\Documents and Settings\Administrateur\Mes documents\Macros Minitab\MCNEMAR.MAC

McNemar's Test for Marginal Homogeneity with Continuity Correction

Chi-Square Statistic:    2.28571
P-value:                0.130570

The 95% CI for delta is: (-0.0123, 0.1423).

where delta represents the difference of the observed marginal proportions.

```

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
1	55	25											
2	38	82											
3													
4													
5													

Nous retrouvons presque les mêmes valeurs que dans le cours de méthodes numériques où nous avons démontré mathématiquement la démarche et l'origine du test. Nous voyons donc que Minitab applique la correction du Yule.

Pour exécuter le test sans la correction, il aurait fallu utiliser la commande suivante:

```

MTB>%mcnemar 1 C2 C2;
SUBC>0;
SUBC>confidence 0.95.

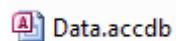
```

## 20.3. Exercice 211.: Macro automation MS Office Access-Minitab-MS Office Word

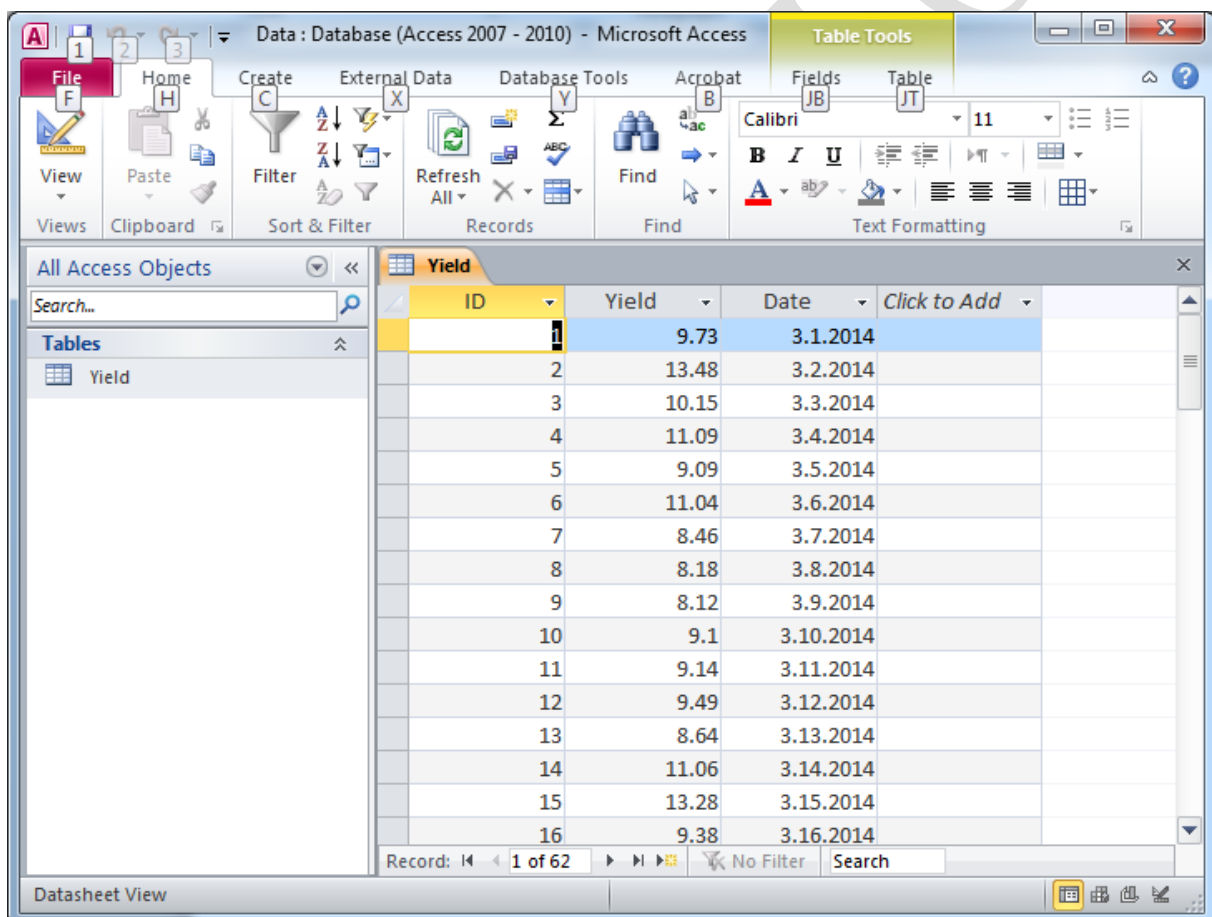
Minitab® Statistical Software 16.1.1.0

Nous allons voir ici comment faire une macro Minitab qui se connecte à une base de données Microsoft Office Access et fait quelques analyses triviales d'une table de celle-ci par une requête SQL intermédiaire et qui ensuite est appelée par une routine VBA dans Microsoft Office Word pour générer des rapports automatiques:

Le fichier Microsoft Access sera le suivant et sera placé à la racine du disque C:\ :



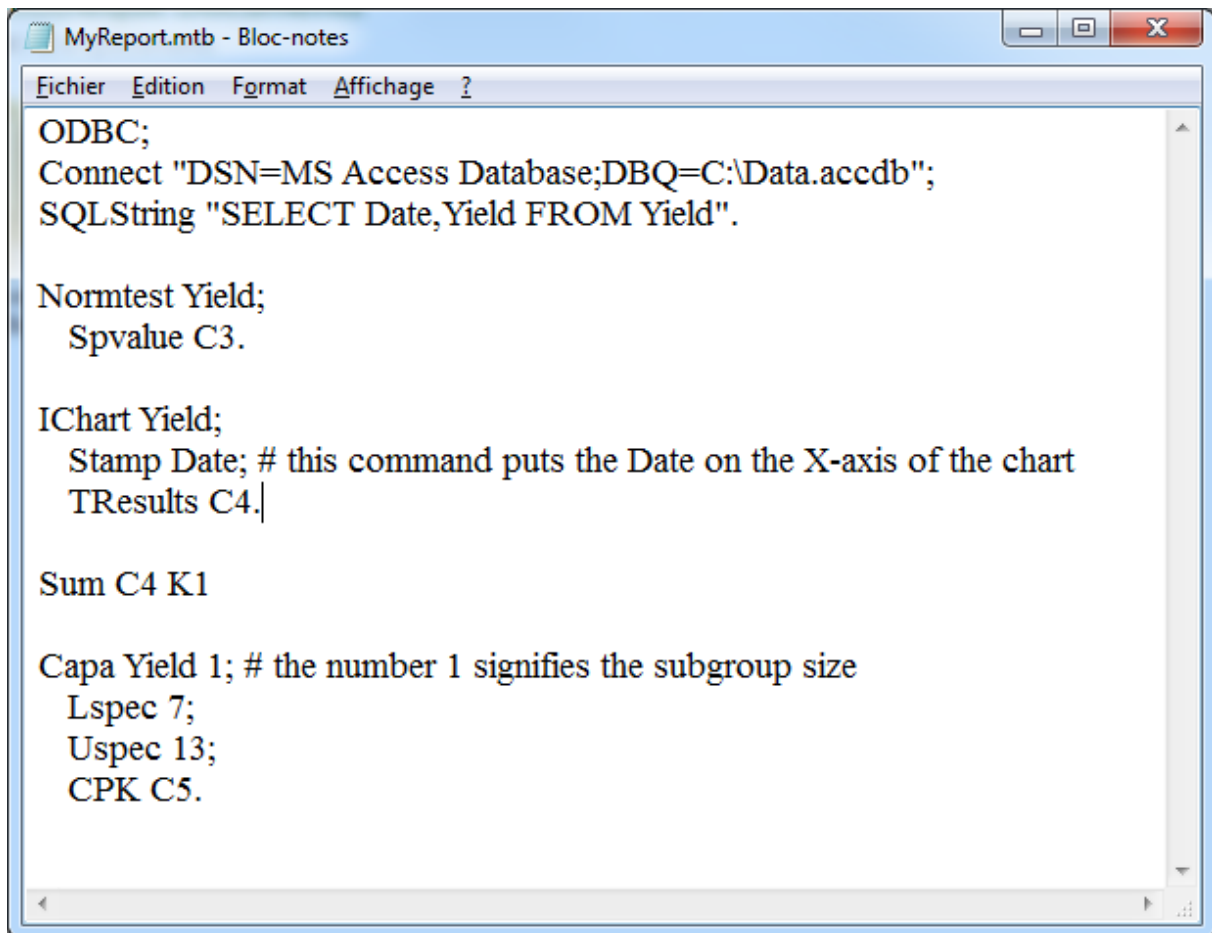
Son contenu est on ne peut plus standard:



ID	Yield	Date	Click to Add
1	9.73	3.1.2014	
2	13.48	3.2.2014	
3	10.15	3.3.2014	
4	11.09	3.4.2014	
5	9.09	3.5.2014	
6	11.04	3.6.2014	
7	8.46	3.7.2014	
8	8.18	3.8.2014	
9	8.12	3.9.2014	
10	9.1	3.10.2014	
11	9.14	3.11.2014	
12	9.49	3.12.2014	
13	8.64	3.13.2014	
14	11.06	3.14.2014	
15	13.28	3.15.2014	
16	9.38	3.16.2014	

Ensuite, avec le bloc-notes, nous écrivons la macro suivante:





```
MyReport.mtb - Bloc-notes
Fichier  Edition  Format  Affichage  ?
ODBC;
Connect "DSN=MS Access Database;DBQ=C:\Data.accdb";
SQLString "SELECT Date,Yield FROM Yield".


Normtest Yield;
  Spvalue C3.

IChart Yield;
  Stamp Date; # this command puts the Date on the X-axis of the chart
  TResults C4.

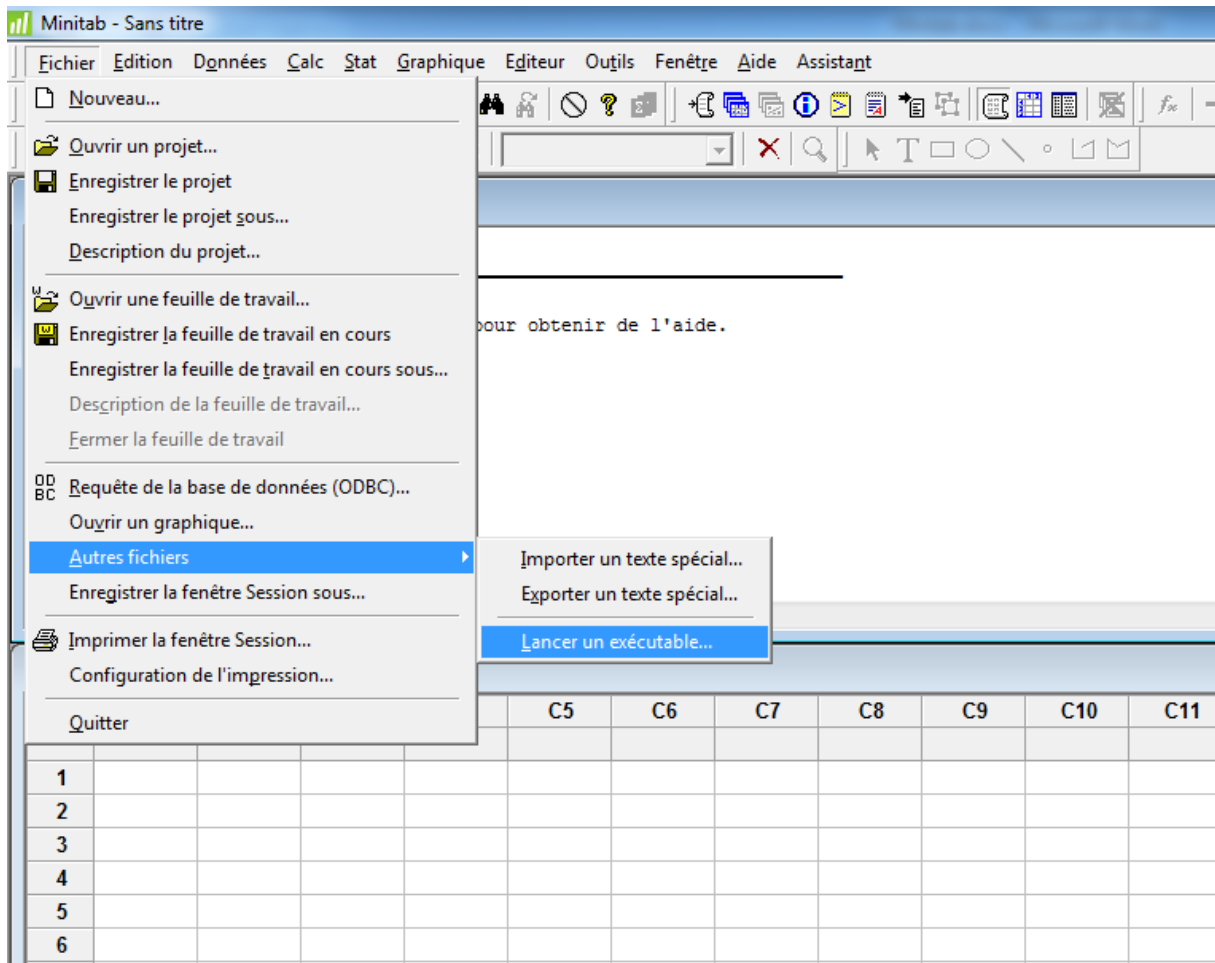
Sum C4 K1

Capa Yield 1; # the number 1 signifies the subgroup size
  Lspec 7;
  Uspec 13;
  CPK C5.
```

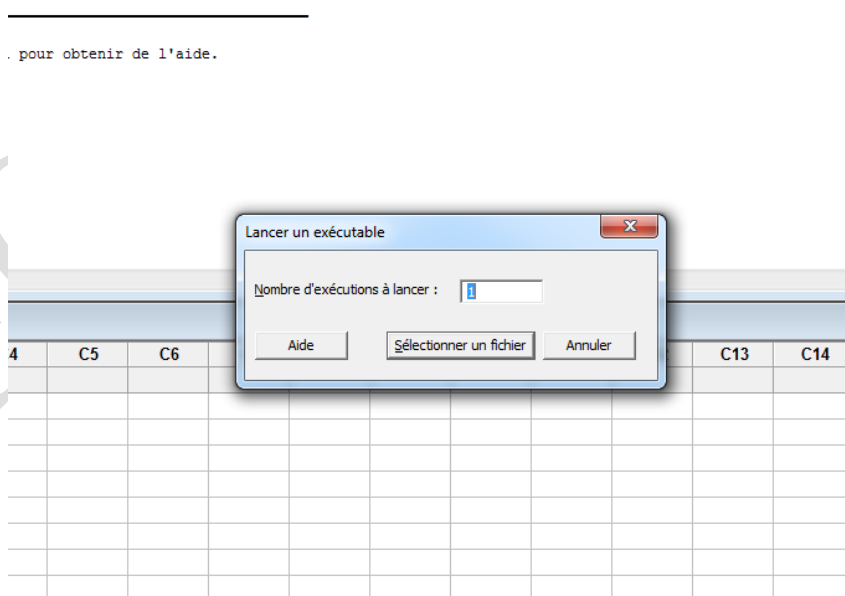
que nous enregistrons sous le nom *MyReport.mtb*:

 MyReport.mtb

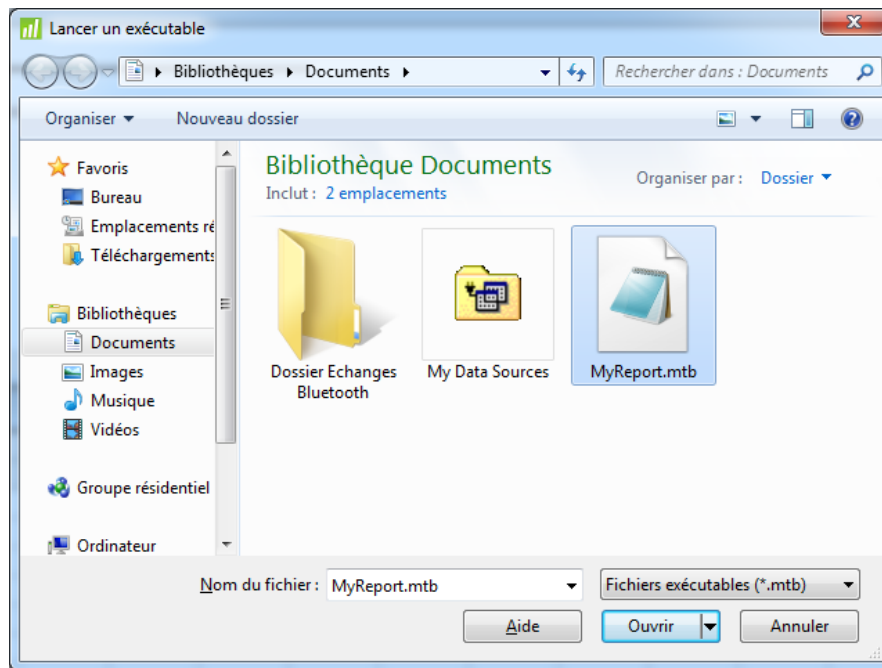
Ensuite, dans Minitab, nous allons dans le menu **Fichier/Autres Fichiers/Lancer un exécutable...**:



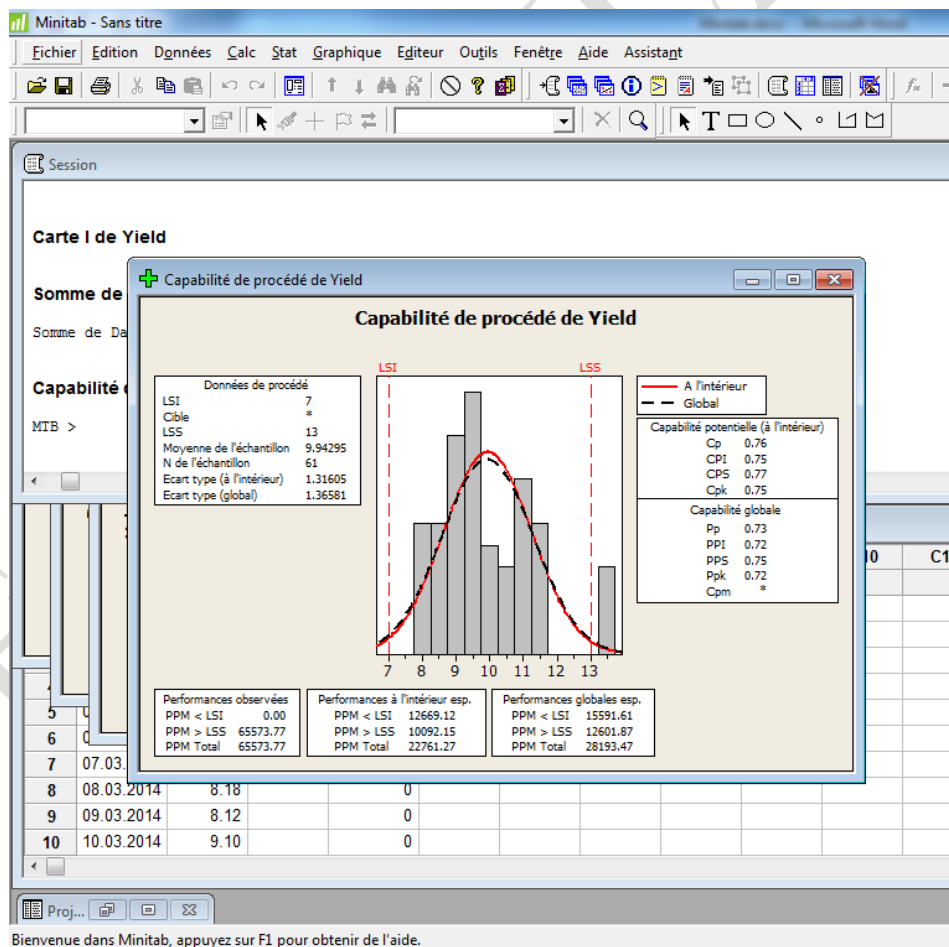
Il vient alors:



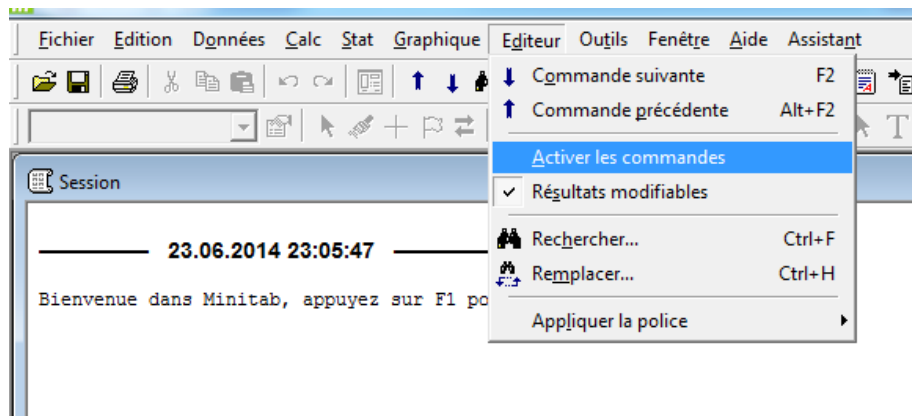
Nous cliquons sur **Sélectionner un fichier:**



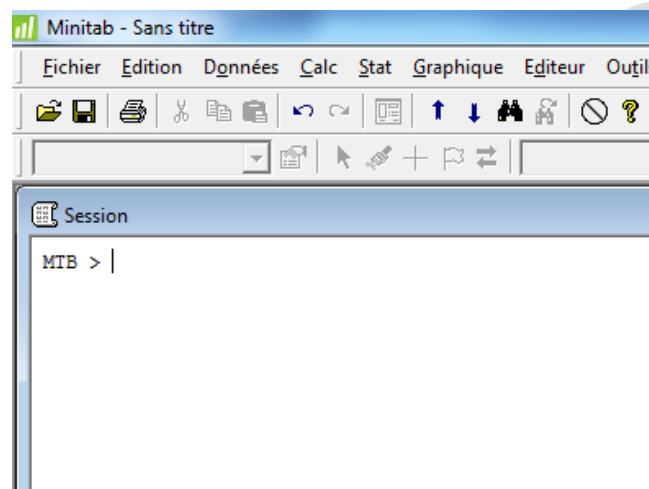
Nous prenons notre petit script et cliquons sur **Ouvrir**:



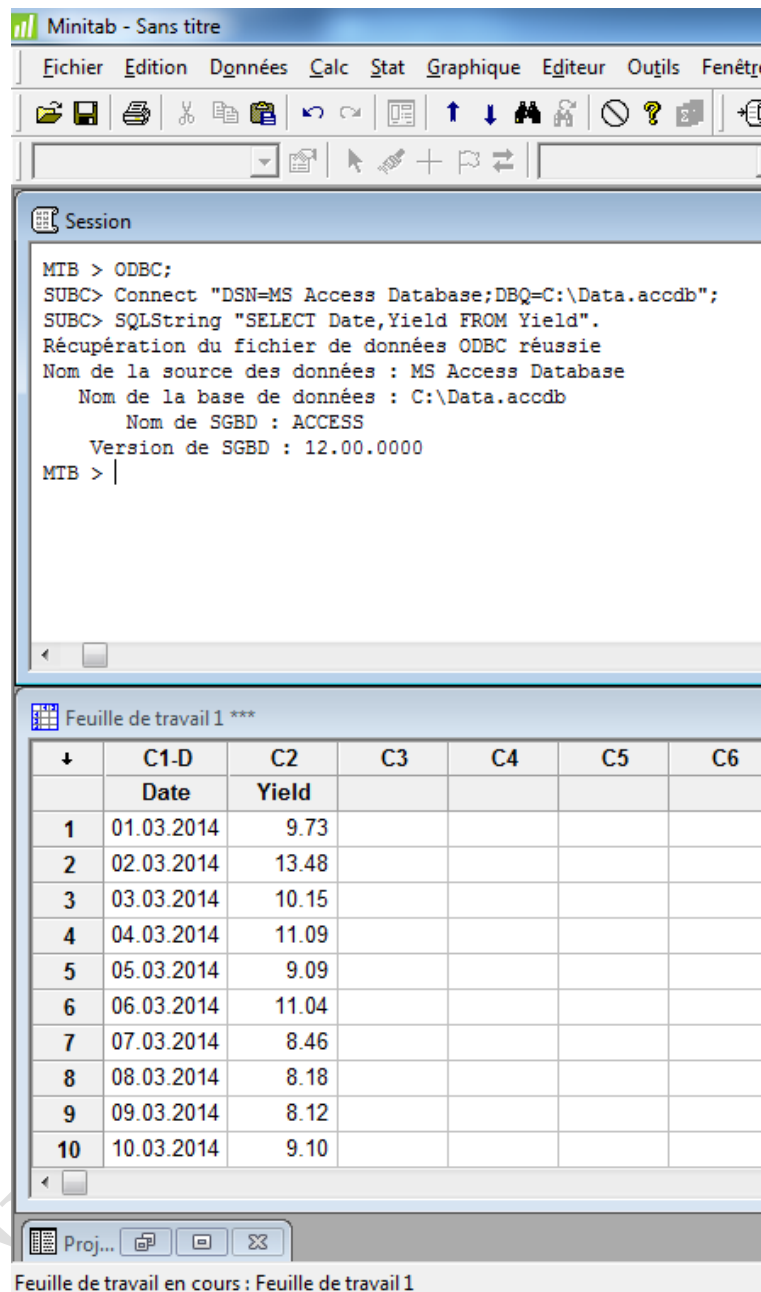
Nous pouvons exécuter cela pas à pas en allant dans Minitab dans le menu **Editeur/Activer les commandes**:



Il vient alors:



On peut alors copier/coller les lignes de code une par une et valider par *Entrée* sur le clavier. Voici des captures d'écran des étapes majeures:



The screenshot shows the Minitab interface. The 'Session' window displays the following text:

```
MTB > ODBC;
SUBC> Connect "DSN=MS Access Database;DBQ=C:\Data.accdb";
SUBC> SQLString "SELECT Date,Yield FROM Yield".
Récupération du fichier de données ODBC réussie
Nom de la source des données : MS Access Database
  Nom de la base de données : C:\Data.accdb
  Nom de SGBD : ACCESS
  Version de SGBD : 12.00.0000
MTB > |
```

The 'Feuille de travail 1 \*\*\*' window shows a table with the following data:

	C1-D	C2	C3	C4	C5	C6
	Date	Yield				
1	01.03.2014	9.73				
2	02.03.2014	13.48				
3	03.03.2014	10.15				
4	04.03.2014	11.09				
5	05.03.2014	9.09				
6	06.03.2014	11.04				
7	07.03.2014	8.46				
8	08.03.2014	8.18				
9	09.03.2014	8.12				
10	10.03.2014	9.10				

The status bar at the bottom indicates 'Feuille de travail en cours : Feuille de travail 1'.

Les deux prochaines commandes donnent (observez la troisième colonne de la feuille de données):

The screenshot shows the Minitab interface. The Session window contains the following text:

```

SUBC> Connect "DSN=MS Access Database;DBQ=C:\Data.accdb";
SUBC> SQLString "SELECT Date,Yield FROM Yield".
Récupération du fichier de données ODBC réussie
Nom de la source des données : MS Access Database
  Nom de la base de données : C:\Data.accdb
  Nom de SGBD : ACCESS
  Version de SGBD : 12.00.0000
MTB > Normtest Yield;
SUBC> Spvalue C3.
    
```

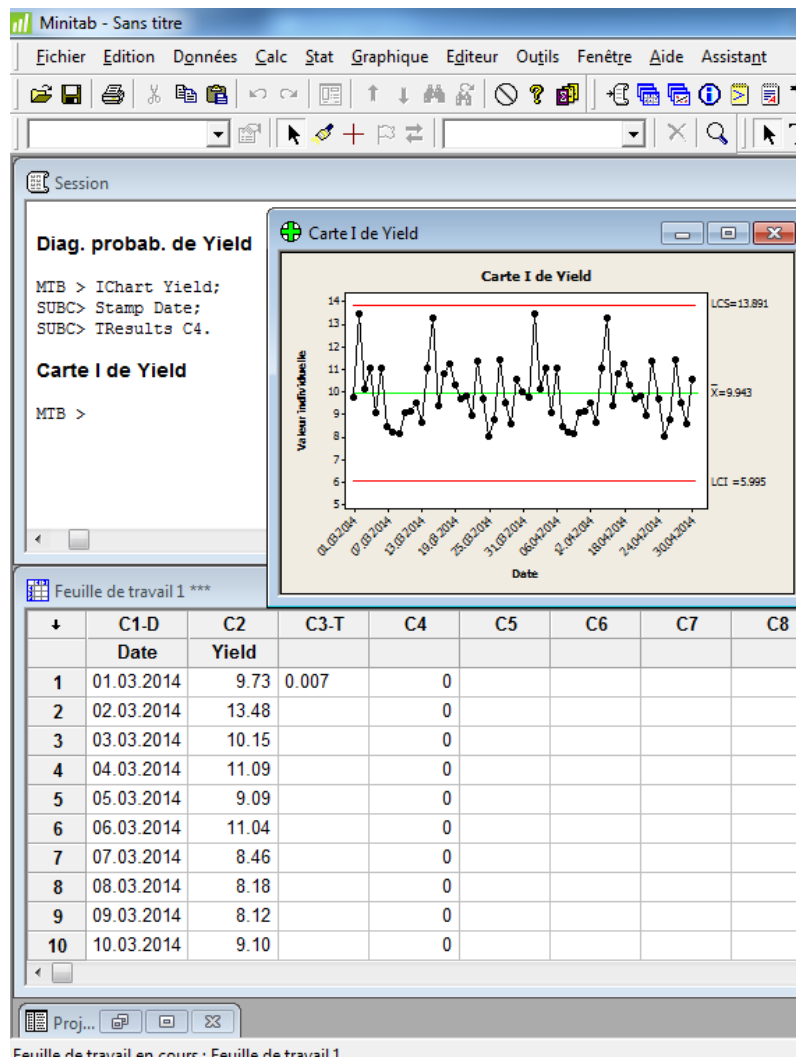
Below the session window, a normal probability plot titled "Diag. probab. de Yield" is displayed. The plot shows data points following a normal distribution line. A statistics box on the right provides the following values:

Moyenne	9.943
Ectyp	1.366
N	61
AD	1.092
Valeur de P	0.007

At the bottom, a worksheet window titled "Feuille de travail 1 \*\*\*" shows a table with the following data:

↓	C1-D Date	C2 Yield	C3-T
1	01.03.2014	9.73	0.007
2	02.03.2014	13.48	
3	03.03.2014	10.15	
4	04.03.2014	11.09	
5	05.03.2014	9.09	
6	06.03.2014	11.04	

Les trois prochaines commandes donnent (observez la quatrième colonne):



Ensuite on peut faire la prochaine commande:

```

Minitab - Sans titre
Fichier Edition Données Calc
[Icons]
Session
MTB > IChart Yield;
SUBC> Stamp Date;
SUBC> TResults C4.

Carte I de Yield

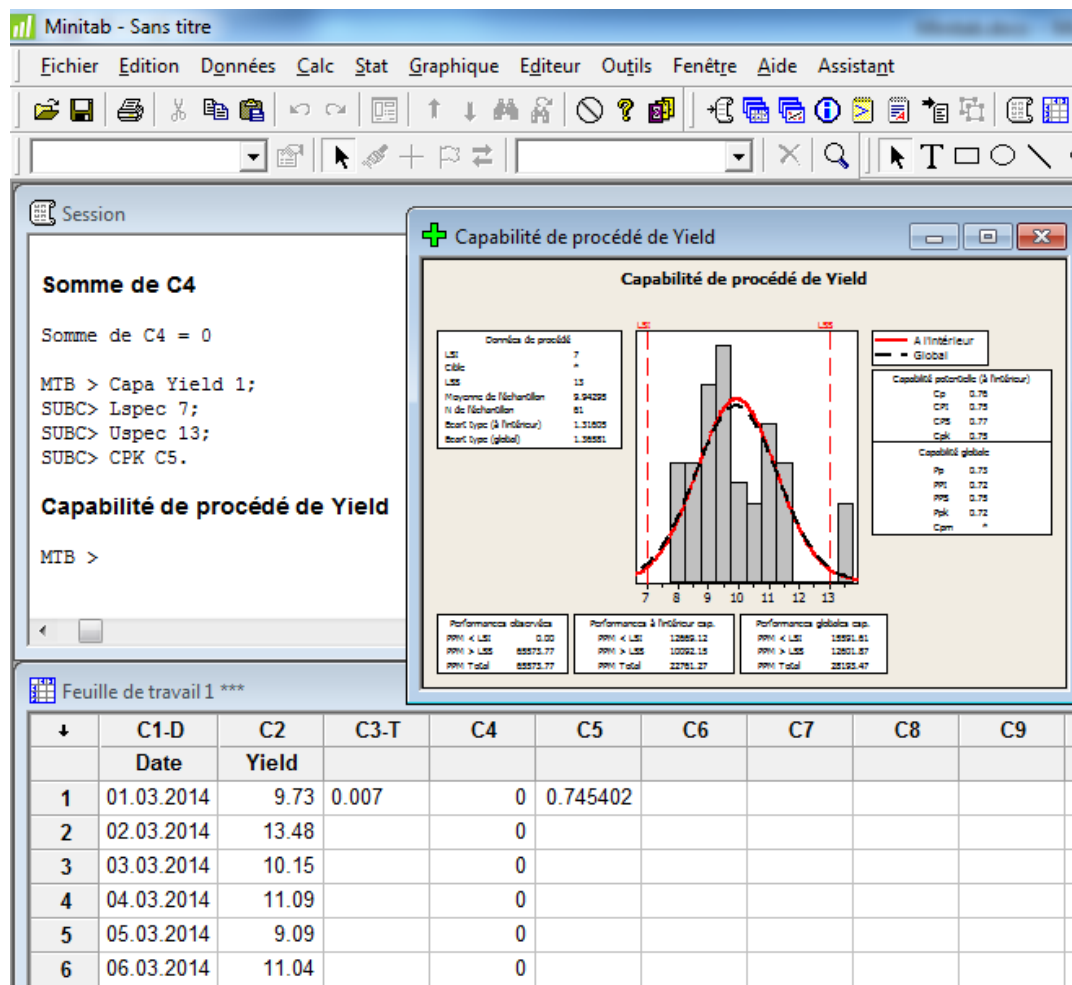
MTB > Sum C4 K1

Somme de C4

Somme de C4 = 0

MTB > |
    
```

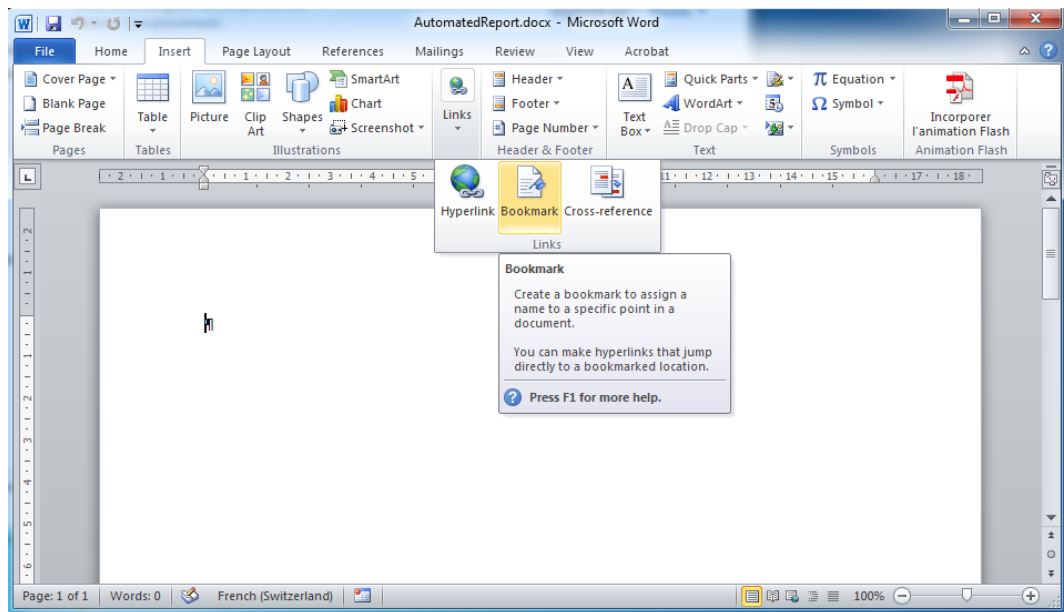
Et enfin les dernières commandes:



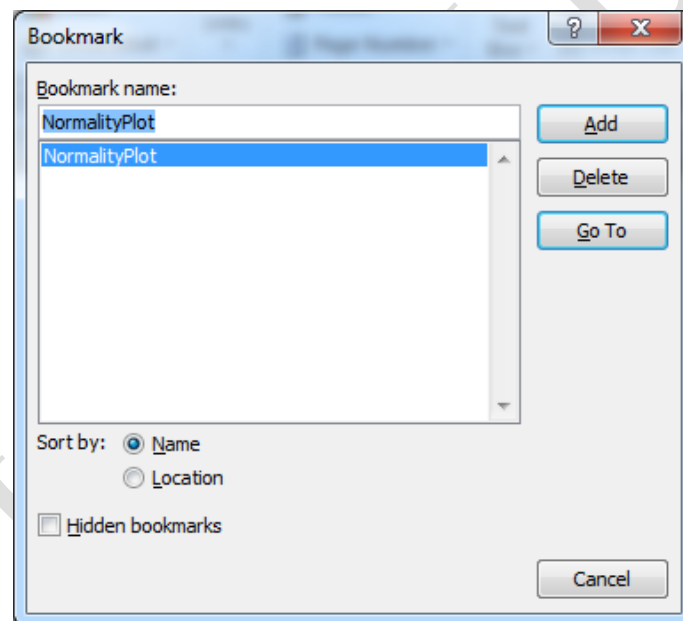
Passons maintenant à la partie VBA!

Nous ouvrons Microsoft Word (la version 2010 x86 dans le cas présent) et nous allons dans le ruban **Insert/Links/Bookmark** (pour décider de la position d'import automatique du graphique et d'une phrase de conclusion (cas classique dans les formations VBA MS Word):

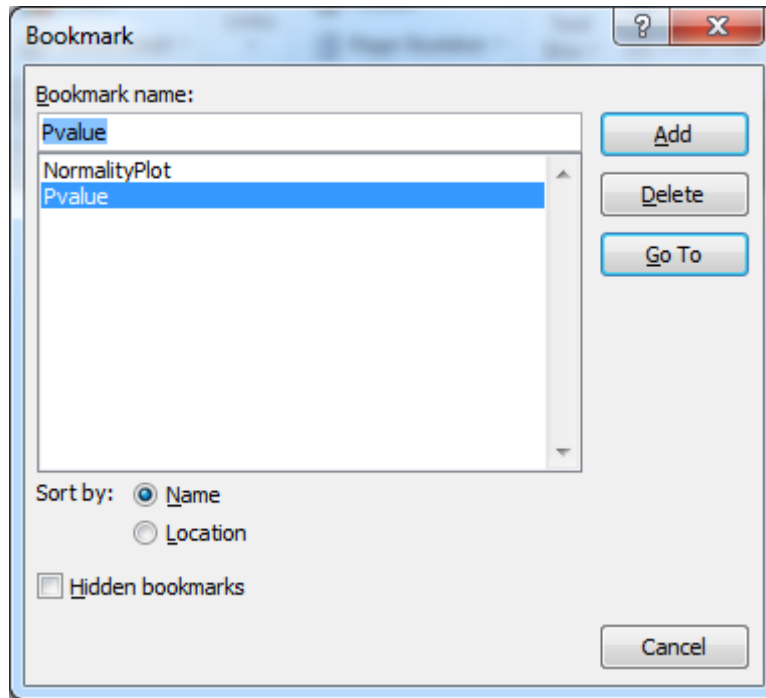




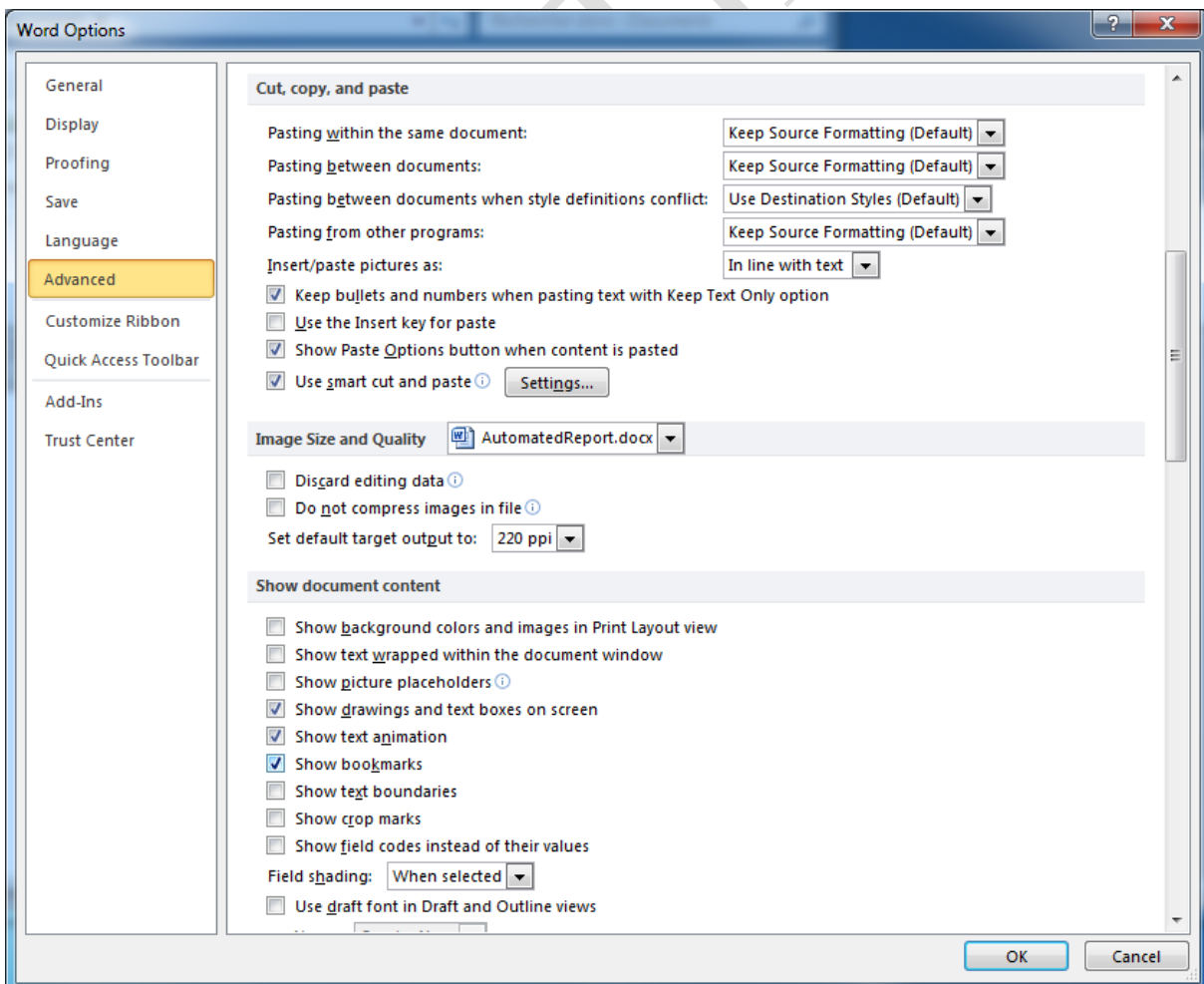
Nous créons un premier bookmark (signet) pour le graphique à l'endroit de notre choix:



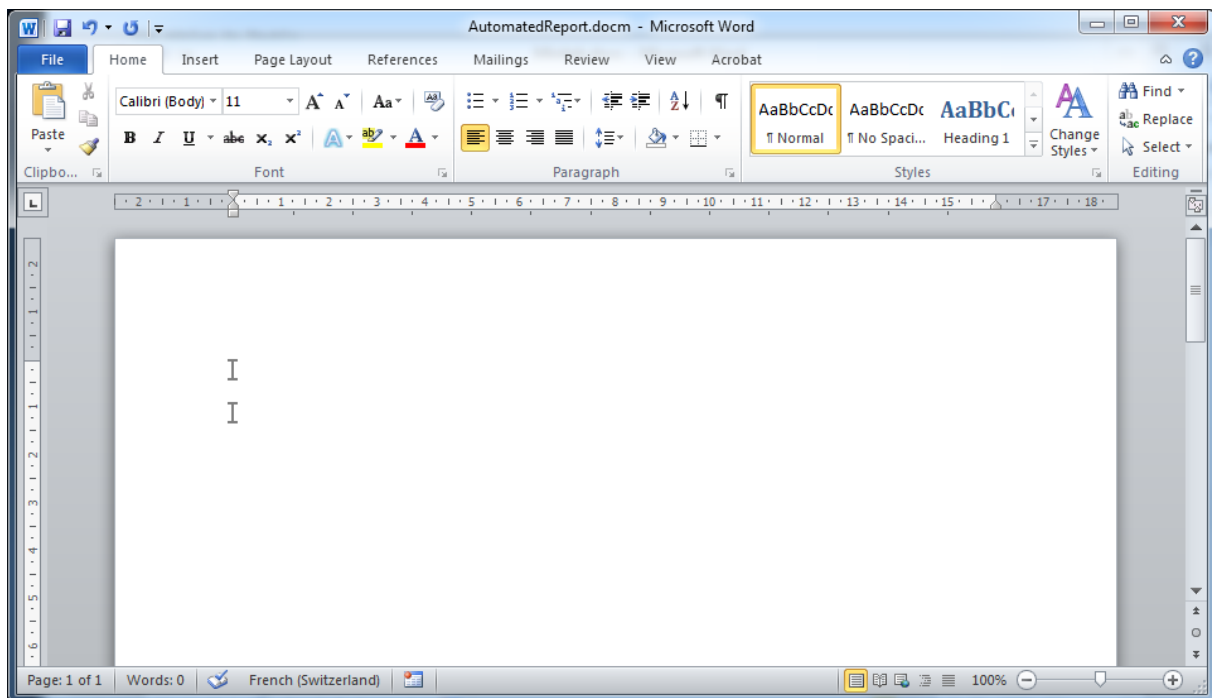
Nous faisons un retour à la ligne dans Microsoft Word en ensuite nous rajoutons un deuxième bookmark :



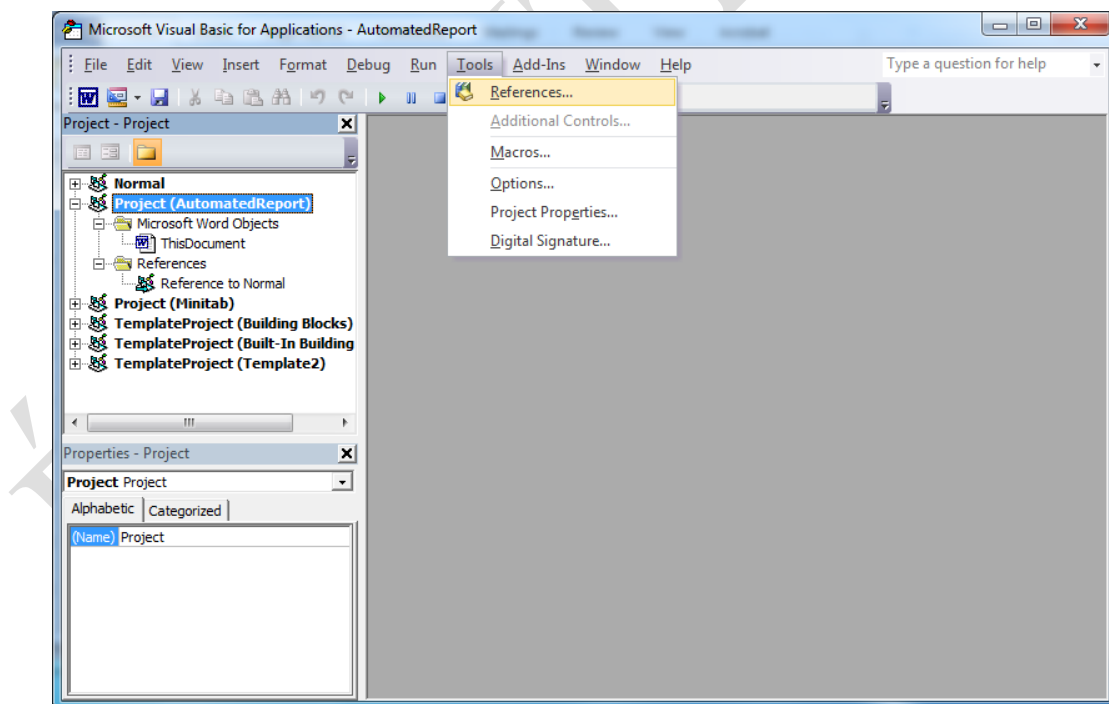
Si vous ne voyez pas les signets insérés dans MS Word c'est normal car c'est le comportement par défaut. Pour changer cela (rappel cours MS Word de base), allez dans les options du logiciel et dans la section **Advanced** cochez **Show bookmarks**:



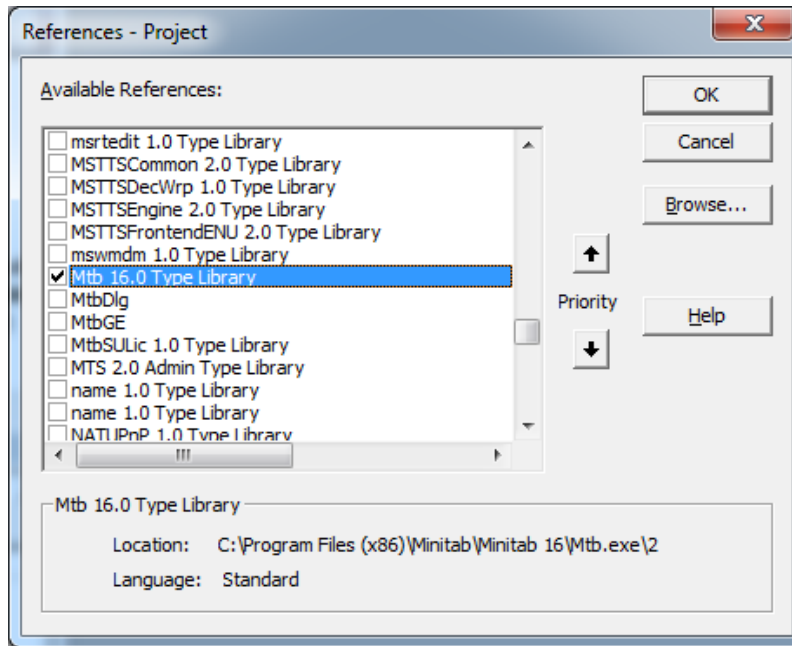
Vous verrez alors les bookmarks (signets) dans la page (petits traits gris verticaux):



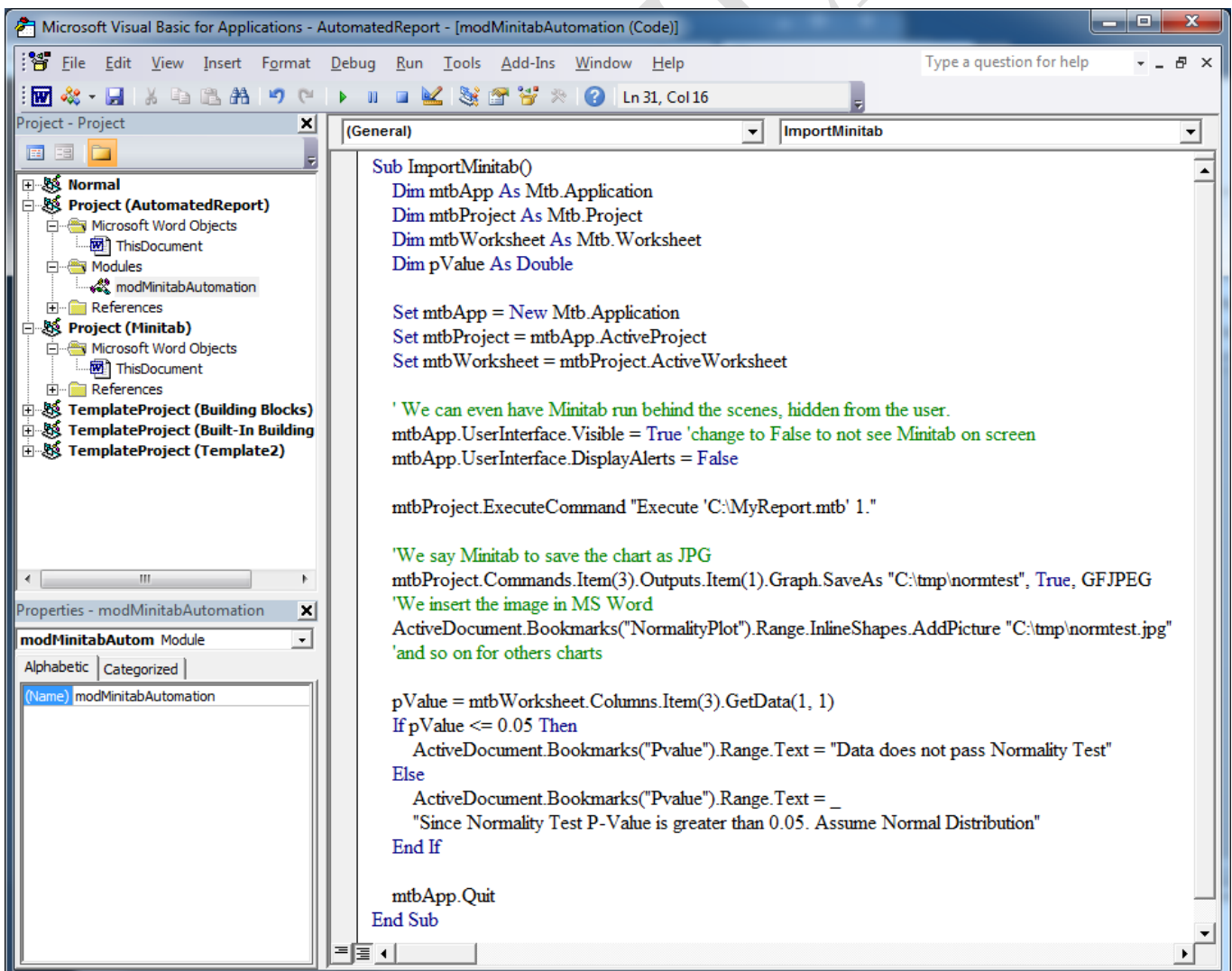
Ensuite, nous allons ans l'éditeur VBA (Alt+F11) et allons dans le menu **Tools/References** pour ajouter la référence de la bibliothèque Minitab:



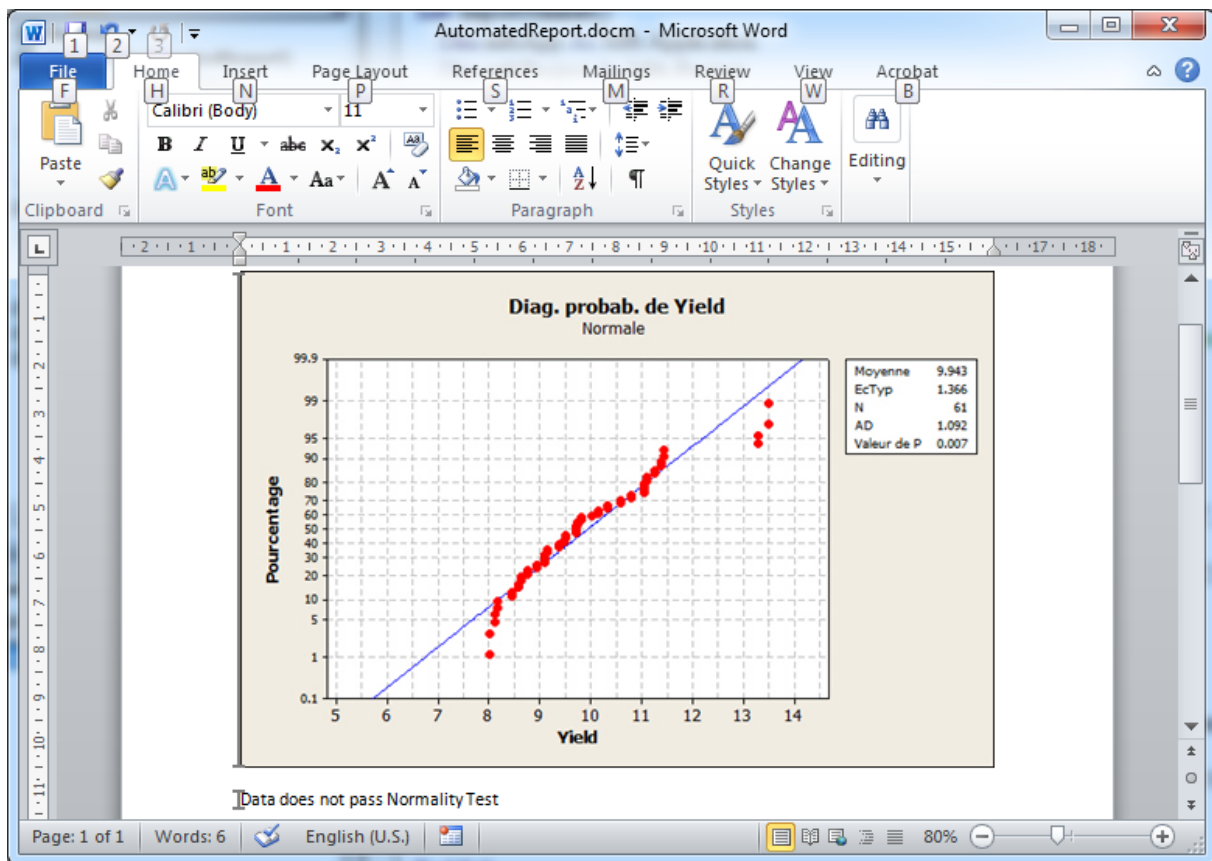
Nous voyons bien que Minitab est dans la liste et nous le cochons:



Ensuite, dans un module préalablement créé (voir un cours de base sur le VBA) nous écrivons le code suivant:



Et si nous l'exécutons nous obtenons:



Voilà donc comment il est possible d'automatiser l'analyse de données avec Minitab et ainsi d'éviter de se tromper dans des manipulations humaines ou même en externalisant le développement à des experts qui automatiseront les conclusions des analyses pour vous.

## Conclusion

Voilà nous sommes arrivés à la fin de ce livre. Ce dernier n'est bien évidemment pas encore terminé et nécessite aussi de nombreuses mises à jour et compléments d'informations mais constitue toutefois probablement une bonne référence francophone pour tout praticien des statistiques industrielles de niveau LMD (Licence, Master, Doctorat) et tout débutant à Minitab ou à la méthodologie Six Sigma qui souhaiterait démarrer du bon pied et ce tout en étant informé des bonnes pratiques d'utilisation.

Nous rappelons, pour ceux qui auraient sauté la lecture de l'introduction, que ce livre est surtout un compagnon au livre théorique *Opera Magistris* et aux formations dispensées par l'auteur de ce livre électronique, et que bien que le livre soit gratuit, les exercices d'accompagnements (fichiers de données) sont eux payants!

# 21. Index

<b>A</b>	<b>C</b>
AFE (avec rotation OrthoMax)..... 865	Calculatrice ..... 71
AFE (avec rotation VariMax)..... 862	calculs arithmétiques élémentaires ..... 70
AFE (sans rotation) ..... 846	capabilité (Normale) ..... 715
afficher ou masquer les colonnes..... 58	capabilité (Poisson)..... 718
alpha de Cronbach ..... 817	Carte barre..... 110
Analyse d'accord d'attribut ..... 793, 801	Carte d'essai de l'instrumentation ..... 789
Analyse d'accord de concordance ..... 793	Cartes de contrôle..... 614
Analyse de la capabilité (Normale)..... 715, 740	Aide au choix de la carte P ..... 693
Analyse de la capabilité (Poisson) ..... 718	Aide au choix de la carte U..... 699
Analyse en composantes principales ..... 836	CC aux attributs C (non-conformités rares d'assemblages)..... 620
Analyse factorielle exploratoire (avec rotation OrthoMax) ..... 865	CC aux attributs NP (comptage par proportions) ... 617
Analyse factorielle exploratoire (avec rotation VariMax) ..... 862	CC aux attributs P (proportions)..... 614
Analyse factorielle exploratoire (sans rotation) ..... 846	CC aux attributs U (Carte C normalisée en pourcentages) ..... 623
Analyse linéaire discriminante ..... 831	CC de mesures chronologiques pondérées CUSUM671
Analyse SixPack Normale ..... 746	CC de mesures chronologiques pondérées EWMA avec échantillons ..... 675
Analyses graphiques des interactions et effets principaux ..... 417	CC de mesures chronologiques pondérées EWMA avec mesures individuelles..... 678
ANCOVA (Analyse de la Covariance) ..... 473	CC de mesures individuelles par étendue mobile (I- Embarre-EM) ..... 662
ANOVA	CC de mesures individuelles par étendue mobile (I- Xbarre-EM) ..... 657, 668
Analyse de la variance à deux facteurs fixes (ANOVA- 2) avec répétitions (et sans interactions) ..... 455	CC de mesures individuelles par étendue mobile (I- Xbarre-EM) combinées..... 665
Analyse de la variance à deux facteurs fixes (ANOVA- 2) sans répétitions ..... 444	CC de mesures R-barre R-barre sous-groupes ..... 647
Analyse de la variance à un facteur fixe (ANOVA-1) désempilé ..... 434	CC de mesures R-barre R-barre sous-groupes combinées ..... 654
Analyse de la variance à un facteur fixe (ANOVA-1) empilé..... 437	CC de mesures S-barre S sous-groupes ..... 630
Analyse de la variance non paramétrique de Friedman ..... 383	CC de mesures X-barre R-barre sous-groupes..... 650
ANOVA carré gréco-latin ..... 453	CC de mesures X-barre S-barre sous-groupes ..... 633
ANOVA carré latin ..... 452	CC de mesures X-barre S-barre sous-groupes combinées ..... 637
ANOVA équilibrée..... 432	CC de mesures X-barre S-barre sous-groupes combinées par dates ..... 640
ANOVA Imbriquée/Hiérarchisée (nested ANOVA) . 459	CC de mesures X-barre S <sub>p</sub> sous-groupes ..... 645
ANOVA totalement emboîtée ..... 459	CC de X Barre-R et R Barre-R avec phases ..... 681
Graphique des effets principaux ..... 418	CC des événements rares G ..... 688
Graphique des interactions ..... 419	CC des événements rares T ..... 685
Graphique multivarié ..... 431	CC multivariée T de Hotelling ..... 702
MANOVA ..... 477	CC par attributs P' de Laney ..... 692
Modèle linéaire général ..... 448, 452	CC par attributs U' de Laney..... 698
Test de Kruskal-Wallis (ANOVA non paramétrique)470	coefficient de variation ..... 207
Test de Tukey ..... 439	commentaires et métadonnées..... 48
ANOVA paramétriques ou non paramétriques ..... 412	communalité ..... 849
Autocorrélation ..... 600	courbe ROC ..... 576
autocorrélation partielle ..... 602	Covariables..... 1004
	Covariance ..... 551
	Criblage ..... 903
<b>B</b>	
Box-Cox ..... 728, 951	

<b>D</b>		Graphique des interactions..... 419
data mining ..... 817		graphique des symétrie ..... 213
date/heure ..... 40		Graphique des valeurs résiduelles de la régressions .. 482
Definitive Screening Design..... 903		graphique en secteurs ..... 113
Deming ..... 587		graphique histogramme..... 131
d'empiler les données ..... 95		Graphique histogramme ..... 129
dendrogramme ..... 821		<b>Graphique matriciel diagonal</b> ..... 145
description ..... 47		Graphique nuage de points..... 135
désempiler des données ..... 97		Graphique nuage de points en 3D ..... 153
Diagramme de Pareto ..... 178		graphique récapitulatif ..... 122
Diagrammes factoriels ..... 1036		graphiques multiples ..... 133
Différence de proportions sur 2 échantillons		grille de fond des graphiques..... 44
indépendants..... 335		
Différence de proportions sur une même population 333		
d'import de données ..... 80		
droite de Henry ..... 36		
<b>E</b>		
Ecart type regroupé..... 645		
échantillonnage ..... 805, 808, 811		
Effectif de l'échantillon pour l'estimation ..... 267		
Effet de rapport..... 1036		
étendue ..... 207		
Étude de la cohérence avec l'alpha de Cronbach ..... 817		
<b>Étude de linéarité de l'instrumentation et biais</b> ..... 539		
Étude de l'instrumentation de type I..... 752		
Étude de R&R de l'instrumentation (croisée). pour		
données continues ..... 758		
Étude de R&R de l'instrumentation (emboîtée). pour		
données continues ..... 773		
Étude de R&R de l'instrumentation (emboîtée). pour		
données étendue ..... 783		
<b>F</b>		
fenêtre de session ..... 38		
feuilles ..... 49		
Fiabilité..... 1187		
Format conditionnel..... 107		
fouille de données ..... 817		
<b>G</b>		
génération de variables aléatoires ..... 100		
graphique à barre..... 119		
Graphique à bulles ..... 142		
graphique à points ..... 116		
Graphique boîte à moustaches ..... 156		
Graphique branches et feuilles ..... 182		
Graphique CDF empirique ..... 138		
graphique dans les marges..... 149		
Graphique de contour ..... 188		
Graphique de loi de probabilité ..... 172, 198		
Graphique de probabilités..... 168		
Graphique de surface en 3D..... 185		
Graphique de valeurs individuelles ..... 125		
Graphique des effets principaux ..... 418		
		Histogramme ..... 102
		Historique ..... 67
		<b>H</b>
		<b>I</b>
		Identification de loi individuelle..... 377
		importer des données d'une base Microsoft Office
		Access..... 87
		Intervalle de confiance de la médiane (via test du signe)
		..... 389
		Intervalle de confiance de la moyenne (test t à un 1
		échantillon) ..... 303
		Intervalle de confiance de la proportion..... 327
		Intervalle de confiance de l'écart-type ..... 347
		intervalle de tolérance ..... 814
		<b>K</b>
		Kappa de Cohen (analyse d'accord d'attribut) ..... 801
		k-Means ..... 825
		K-Moyennes en groupes ..... 825
		Kurtosis ..... 210
		<b>L</b>
		la liste des fichiers connexes à nos analyses ..... 52
		langue ..... 34
		leptokurtique ..... 211
		lien log-linéaire ..... 590
		Lignes de contour
		isolignes ..... 189
		loadings..... 849
		loi Normale ..... 101
		<b>M</b>
		machine learning ..... 817
		Macros ..... 1256
		Automations..... 1272
		Convertir actions en macros ..... 1260
		médiane ..... 207
		mesokurtique..... 211
		métadonnées des fichiers..... 45
		mettre à jour un graphique..... 42





reproductibilité/répétabilité (R&R) emboîtée.....	772
reproductibilité/répétabilité (R&R) étendue.....	782
Résolution d'un test Z en bilatéral .....	279

**S**

saturations .....	849
scinder un fichier.....	63
Scripting .....	1256
Séries temporelles.....	1217
Diagramme de série chronologique .....	1217
Skewness.....	210
split-plot ANOVA .....	464
statistiques descriptives .....	206
statistiques multivariées .....	817
Statistiques non-paramétriques.....	383
Survie .....	1187
Diagramme de rémission.....	1190
Diagramme de survie.....	1190
symboles d'état de graphiques .....	42

**T**

tableau de contingence.....	224
Taille d'échantillon (effectif) d'un test p (proportion) à 1 échantillon en bilatéral.....	323
Taille d'échantillon (effectif) d'un test p (proportion) à 2 échantillons en bilatéral .....	330
Taille d'échantillon (effectif) d'un test t à 1 échantillon en bilatéral .....	298
Taille d'échantillon (effectif) d'un test Z en bilatéral.....	276
Techniques de régressions et interpolations .....	479
Test binomial exact .....	333
Test d'ajustement de Poisson .....	258
Test d'ajustement du Khi-deux.....	253
Test d'Anderson-Darling .....	36
Test d'association du khi-deux .....	231
Test de Dixon.....	265
Test de Dunnett.....	364
Test de Fisher d'égalité des variances .....	350
Test de Grubbs .....	262
Test de Kaplan-Meier censuré à droite .....	1187
Test de la médiane .....	386
Test de la somme des rangs de (Wilcoxon)Mann- Withney pour deux échantillons indépendants.....	403
Test de la somme des rangs signés de Wilcoxon pour 1 échantillon.....	406

Test de la somme des rangs signés de Wilcoxon pour 2 échantillons appariés .....	409
Test de Levene d'égalité de deux variances.....	359
Test de Levene et Bartlett d'égalité des variances d'une ANOVA .....	361
Test de l'homoscédasticité des résidus.....	487
Test de Mantel-Haenszel-Cochran .....	245
Test de Mood (test des médianes).....	393
Test de normalité.....	163
Test de normalité d'Anderson-Darling (ie Agostino- Stephens) .....	376
Test de normalité de Kolmogorov-Smirnov .....	369
Test de normalité de Lilliefors .....	369
Test de normalité de Ryan-Joiner (ie Shapiro-Wilk)....	373
Test de Poisson à 1 échantillon.....	338
Test de Poisson à 2 échantillons .....	344
Test de Tukey.....	439
test d'équivalence.....	318
Test des valeurs aberrantes .....	261
Test du khi-2 de la variance .....	347
Test du Khi-deux .....	227
Test du signe .....	389
Test du signe binomial (test de la médiane) .....	386
Test exact de Fisher pour les tableaux 2x2 .....	238
Test non-paramétrique ANOVA de Friedman .....	383
Test t de Student pour données appariées.....	312
Test t-Student de différence de la moyenne bilatéral d'un échantillon .....	307
Test t-Student hétéroscédastique bilatéral d'égalité de la moyenne (test t de Welch).....	317
Test t-Student homoscédastique bilatéral d'égalité de la moyenne .....	309
Test Z à 1 échantillon .....	281
Tests d'Anderson-Darling multiples .....	377
TOST student.....	318
Transformation de Box-Cox .....	728, 951
Transformation de Johnson .....	735
Trier à plat.....	221
Trier les colonnes .....	61
typage de colonnes .....	57

**V**

V carré de Cramer .....	241
valeur modale .....	207
valeurs propres .....	76
vecteurs propres .....	78
version .....	33