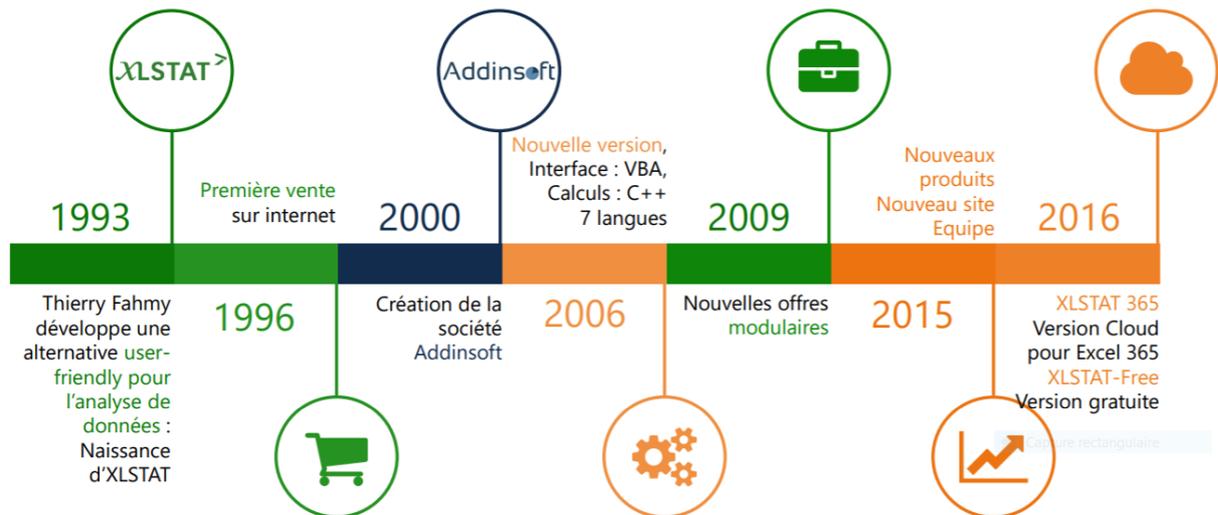


Logiciel XLSTAT

XLSTAT permet d'utiliser les techniques de statistique d'analyse de données de modélisation sans quitter Microsoft Excel

Sa particularité c'est d'être parfaitement intégré à Excel. On lance XLSTAT dans barre d'outils et menu apparaissent.



Caractéristiques du logiciel

1/ Une interface intuitive

2/ XLSTAT utilise Excel comme une interface

- de récupération des données
- et d'affichage des résultats.

3/ Tous les calculs sont réalisés en dehors d'Excel.

Quelques applications

- Marketing quantitatif
- Analyse d'enquêtes
- Analyse de portefeuille
- Modélisation de risques
- Analyse de données de survie
- Contrôle qualité Analyse sensorielle ... etc.

Sélection des données

La sélection des données se fait par colonnes

Affichage des résultats dans Excel

Dans chaque boîte de dialogue, XLSTAT propose 3 modes d'affichage des résultats :

1/ Plage : sur la même feuille des données, c'est-à-dire les résultats sont affichés dans la feuille active, à partir d'une cellule (à sélectionner) qui sera le coin supérieur gauche de la plage.

2/ Feuille : une nouvelle feuilles qui présente les calculs sur le même fichier Excel. Cette feuille sera créée par XLSTAT et porte le nom de l'analyse réalisée.

3/ Classeur : les résultats sont affichés dans un nouveau fichier (document) Excel.

Remarque :

L'affichage des résultats sur une nouvelle feuille reste la méthode la plus pratique et la plus utilisée ; et l'option de l'affichage sur un nouveau classeur est conseillée pour de gros volumes de données.

Statistiques descriptives

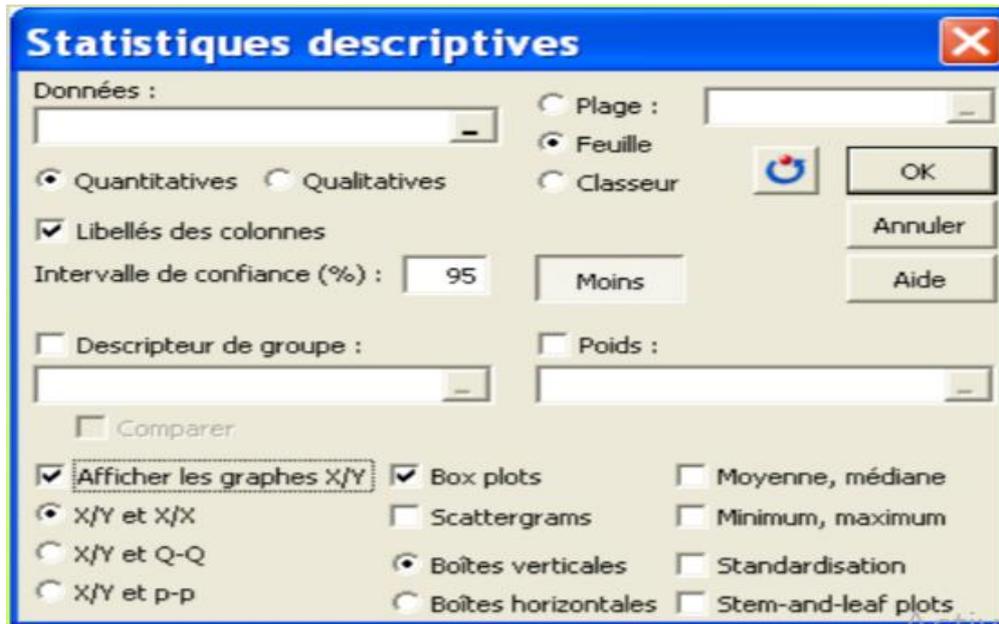
Ouverture par le menu XLSTAT : Statistiques Descriptives : Description des données



Données : sélectionner la plage des variables à décrire. S'il y a des valeurs manquantes dans une colonne, XLSTAT propose de les ignorer. En cas de refus, le traitement est abandonné.

Quantitatives / Qualitatives : choisir le type des variables

Plus : Après un clic sur le bouton Plus, la fenêtre complète s'affiche, si on a besoin d'analyses descriptives supplémentaire : box plots ... etc.



Liste des statistiques descriptives calculées dans le cas des données quantitatives

- Nbr de valeurs utilisées : nombre de valeurs effectivement utilisées dans les calculs (les valeurs manquantes ou poids nul sont exclues)
- Nbr de valeurs ignorées : nombre de valeurs manquantes ou de poids nul
- Nbr de val. min. : nombre de valeurs égales à la valeur minimale
- % de val. min. : pourcentage du nombre de valeurs égales à la valeur minimale
- Minimum : valeur minimale
- 1er quartile
- Médiane
- 3ème quartile
- Maximum : valeur maximale
- Étendue = maximum – minimum
- Total : somme des valeurs
- Moyenne
- Moyenne géométrique : Elle n'est pas définie s'il y a des valeurs négatives ou nulles.
- Moyenne harmonique : Elle n'est pas définie s'il y a des valeurs nulles.
- Aplatissement (Pearson) :
 - Valeur négative : distribution plus plate que la loi normale
 - Valeur positive : distribution plus pointue que la loi normale

- Asymétrie (Pearson) : par rapport à sa moyenne.
Valeur négative : distribution étalée vers la gauche
Valeur positive : distribution étalée vers la droite
- Aplatissement = le coefficient d'aplatissement calculé par Excel
- Asymétrie = le coefficient d'asymétrie calculé par Excel
- CV : coefficient de variation (écart-type/moyenne)
- Variance d'échantillon : variance des données
- Variance estimée : estimation de la variance d'une population dont les données constituent un échantillon (dénominateur = $n - 1$)
- Écart type d'échantillon
- Écart type estimé
- Écart absolu moyen = la moyenne des valeurs absolues des écarts de chaque valeur par rapport à la moyenne
- Écart type de la moyenne : SEM
- Écart absolu médian = la médiane des écarts à la médiane

Liste des statistiques descriptives calculées dans le cas des données qualitatives

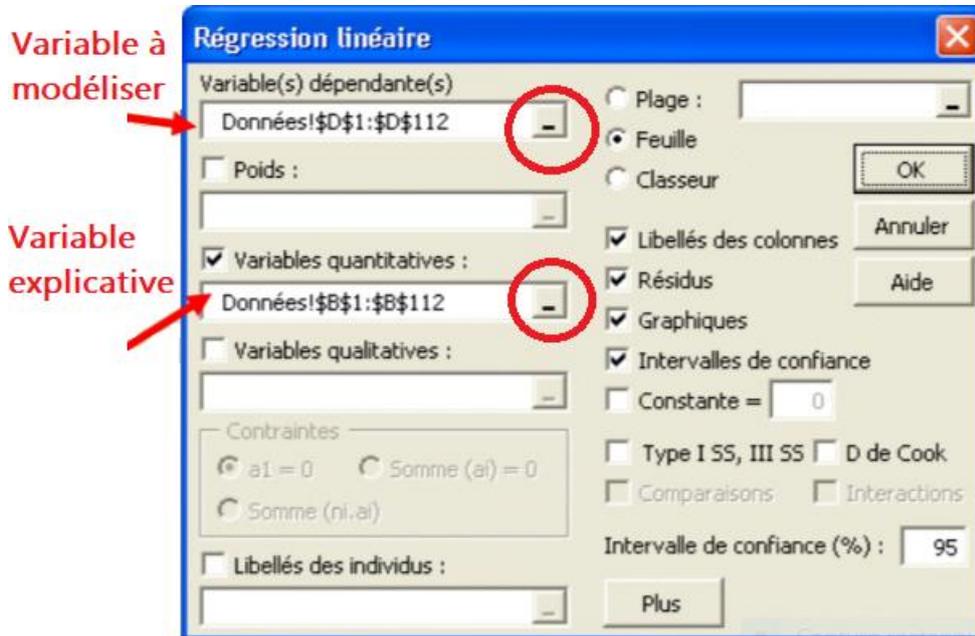
- Nbr de modalités : nombre de modalités de la variable
- Mode
- % mode : fréquence relative du mode exprimée en %
- Fréq. rel. mode : fréquence relative du mode.

Remarque : Box plots sont des graphiques " boîte à moustaches ". Non affichés s'il y a plus de 16 variables ou plus de 30 000 points.

Remarque : *Les graphiques générés par XLSTAT peuvent être modifiés comme n'importe quel autre graphique Excel.*

Régression linéaire

Ouverture par le menu XLSTAT : Modélisation des données : Régression linéaire



Exemple : On veut expliquer « le poids » par « la taille ». La variable dépendante est « le poids » (Y). La variable quantitative est « la taille » (X).

Les résultats de la régression linéaire :

Résumé pour la variable dépendante :						
Variable	Nbr. de valeurs total	Nbr. de valeurs utilisées	Nbr. de valeurs ignorées	Somme des poids	Moyenne	Ecart-type
Poids	111	111	0	11	98,878	18,616
Résumé pour les variables quantitatives :						
Variable	Moyenne	Ecart-type				
Taille	60,526	3,358				
Coefficients d'ajustement :						
R (coefficient de corrélation)	0,751					
R ² (coefficient de détermination)	0,564					
R ² aj. (coefficient de détermination ajusté)	0,560					
SCR	16614,585					
Evaluation de la valeur de l'information apportée par les variables (H0 = Y=Moy(Y)) :						
Source	ddl	Somme des carrés	Carré moyen	F de Fisher	Pr > F	
Modèle	1	21506,523	21506,523	141,094	< 0,0001	
Résidus	109	16614,585	152,427			
Total	110	38121,108				
Paramètres du modèle :						
Paramètre	Valeur	Ecart-type	t de Student	Pr > t	se inférieure	se supérieure
Constante	-153,129	21,248	-7,207	< 0,0001	-195,242	-111,016
Taille	4,164	0,351	11,878	< 0,0001	3,469	4,858
L'équation du modèle s'écrit : Poids = -153,129910179342 + 4,16361172748079*Taille Y = -153.13 + 4.16 X						

Comparaison des moyennes

Ouverture par le menu XLSTAT : Tests paramétriques : Tests t et Z pour deux échantillons

Tests t et Z pour 2 échantillons

Données dans différentes colonnes
 Données dans une colonne

Echantillon 1 :
 Echantillon 2 :

Seuil de signification :

indépendants appariés

Plage :
 Feuille Classeur

Libellés des colonnes
 Graphique

Test t de Student
 Test Z

OK Annuler Aide

Résultats de comparaison des moyennes

comparaison nage forcée - Microsoft Excel

XLSTAT 7.5.2 - Tests t et Z pour 2 échantillons - le 15/11/2017 à 06:08:56

Echantillon 1 : classeur = comparaison nage forcée.xlsx / feuille = Feuil1 / plage = \$B\$4:\$B\$8 / 5 lignes et 1 colonne
 Echantillon 2 : classeur = comparaison nage forcée.xlsx / feuille = Feuil1 / plage = \$D\$4:\$D\$8 / 5 lignes et 1 colonne
 Aucune donnée manquante détectée
 Seuil de signification : 0,05

Statistiques descriptives :

Echantillon	Fréquence	Moyenne	Variance	Ecart-type	type de la mo	Minimum	remier Quartil	Médiane	isième quart	Maximum
♂Témoïn	5	111,200	822,700	28,683	12,827	76,000	86,000	105,000	139,500	149,000
♂Plante	5	33,000	1082,000	32,894	14,711	6,000	12,000	23,000	59,000	90,000

Test t de Student pour données indépendantes / test bilatéral :

On fait l'hypothèse que les variances théoriques sont égales

Intervalle de confiance à 95,00% de la différence des moyennes : 33,192 à 123,208

t (valeur obse	4,007
t (valeur critiq	2,306
ddl	8
p-value bilaté	0,004
Alpha	0,05

Conclusion :

Au seuil de signification Alpha=0,050 on peut rejeter l'hypothèse nulle d'égalité des moyennes. Autrement dit, la différence entre les moyennes est significative.

Feuil1 M-TN F-TN M-TE F-TE M-TI F-TI Tests t et Z 2 échantillons

Comparaison des Variances

Comparaison des variances de k échantillons

Données dans différentes colonnes
 Données dans une colonne

Données :

Groupes :

Seuil de signification :

Plage :
 Feuille
 Classeur

Libellés des colonnes
 Test de Bartlett
 Test de Levene

Moyenne

OK Annuler Aide

Résultats de comparaison des variances

	B	C	D	E	F	G	H	I	J	K	L
	Echantillon	Fréquence	Moyenne	Variance	Ecart-type	type de la mo	Minimum	remier Quartil	Médiane	ioisième quart	Maximum
	♂Témoin	5	111,200	822,700	28,683	12,827	76,000	86,000	105,000	139,500	149,000
	♂Insecticide	5	66,400	2796,300	52,880	23,649	22,000	22,500	67,000	110,000	152,000
	♂Plante	10	35,100	587,433	24,237	7,664	6,000	18,000	28,500	52,000	90,000
	Test de Bartlett :										
	Khi ² (valeur ob	3,677									
	Khi ² (valeur cr	5,991									
	ddl	2									
	p-value unilat	0,159									
	Alpha	0,05									
	Conclusion :										
	Au seuil de signification Alpha=0,050 on ne peut pas rejeter l'hypothèse nulle d'égalité des variances.										
	Autrement dit, l'inégalité des variances n'est pas significative.										
	Test de Levene :										
	F (valeur obse	1,074	Nb : Test de Levene est le test le plus Robuste								
	F (valeur critic	4,619									
	ddl 1	2									
	ddl 2	17									
	p-value unilat	0,364									
	Alpha	0,05									
	Conclusion :										
	Au seuil de signification Alpha=0,050 on ne peut pas rejeter l'hypothèse nulle d'égalité des variances.										
	Autrement dit, l'inégalité des variances n'est pas significative.										

Autres analyses sur XLStat:

Ouverture par le menu XLSTAT : Analyse de données : Analyse en Composantes Principales (ACP), Analyse Factorielle Discriminante (AFD), Analyse Factorielle des Correspondances (AFC), Analyse des Correspondances Multiples (ACM ou AFCM), Classification Ascendante Hiérarchique (CAH), Nuées dynamiques (k-means), ... etc.