

Commencer avec XLStatistics

Bruno Fischer Colonimos

XLStatistics est un logiciel d'analyse des données développé en utilisant Excel par Rodney Carr. C'est un add-in pour Excel, comprenant de nombreux classeurs Excel, et faisant appel au langage de programmation VBA.

Il est disponible (gratuitement pour un usage non-professionnel) à l'adresse :

<http://www.deakin.edu.au/~rodneyc/XLStatistics/>

Ce texte se veut un guide de démarrage rapide. Il faut remarquer que :

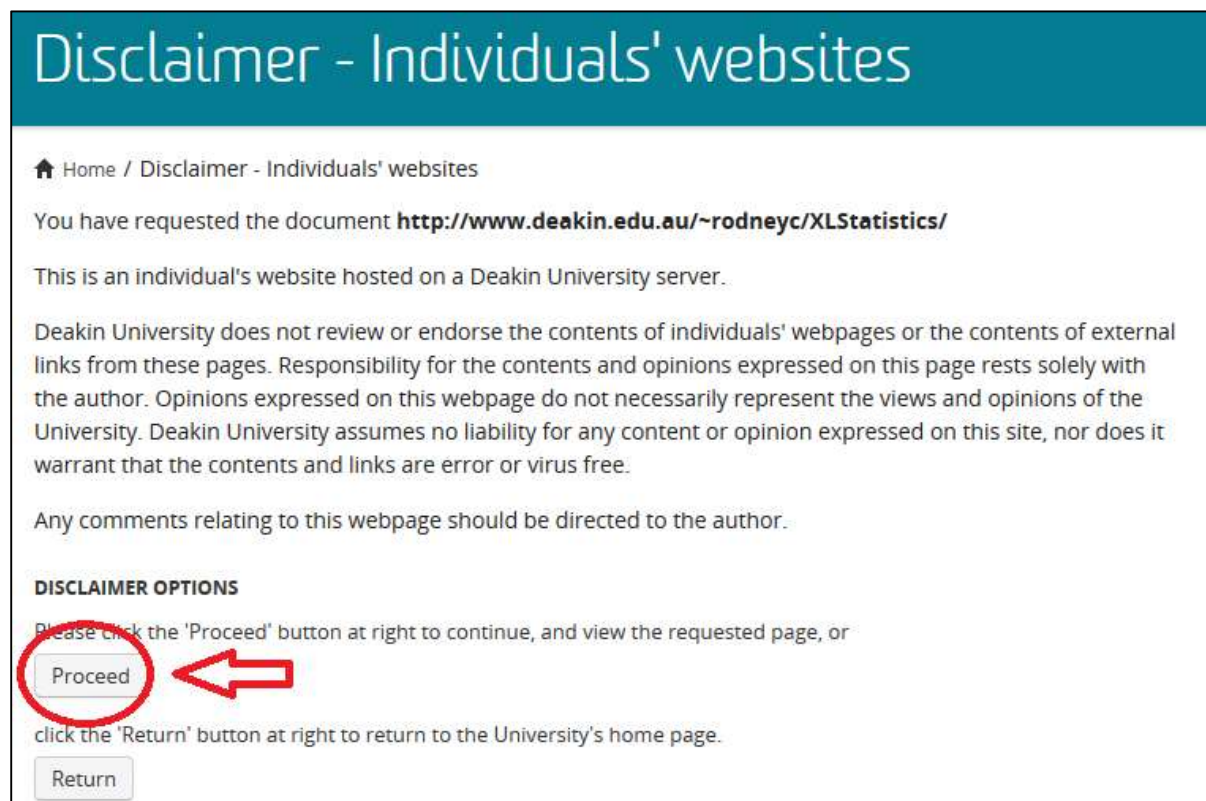
1. Le principe d'organisation du logiciel est particulièrement simple.
2. Rodney fournit, dans le fichier téléchargeable XLS6.zip,
 - a. Une vidéo (relative, il est vrai, à une ancienne version du logiciel), mais qui reste utile
 - b. Un manuel.
3. Il y a également une vidéo sur Youtube. (également relative à une ancienne version du logiciel), mais utile.

Installation

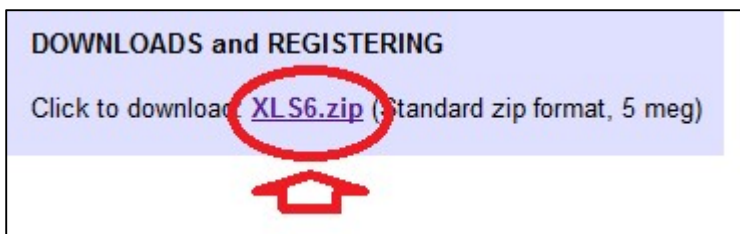
Il n'y a pas véritablement d'installation. XLStatistics est très discret et ne modifie votre machine en aucune manière (et donc, pour « désinstaller », il suffit d'effacer tout). Il y a juste une marche à suivre pour que ça fonctionne.

Téléchargement

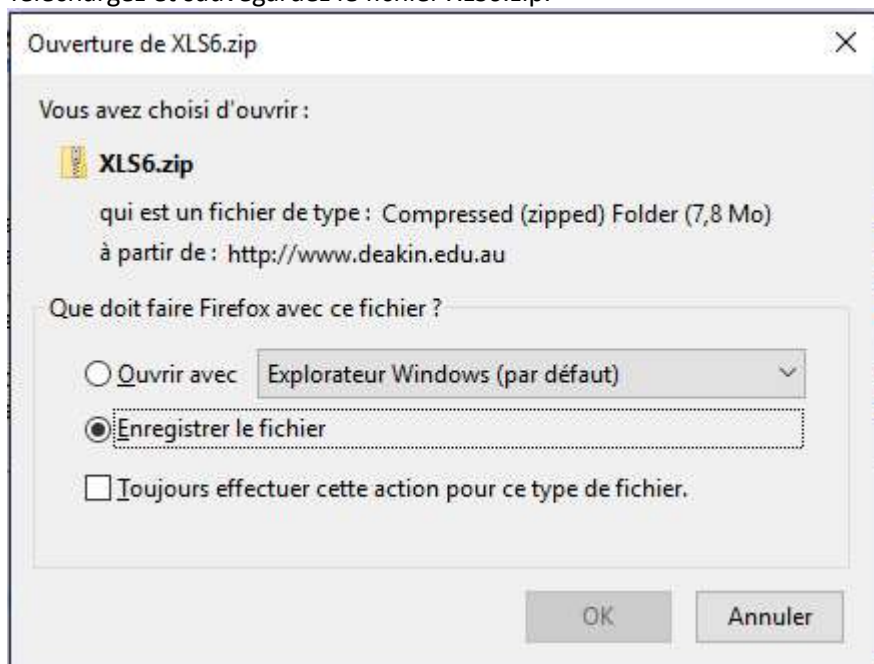
Cliquez sur le lien <http://www.deakin.edu.au/~rodneyc/XLStatistics/>. Vous trouverez sans doute sur votre chemin une page d'avertissement de l'université de Rodney. Cliquez sur le bouton « Proceed »



Vous êtes alors dirigés vers la page XLStatistics. Lisez la page et téléchargez le fichier .zip en cliquant sur le lien :



Téléchargez et sauvegardez le fichier XLS6.zip.

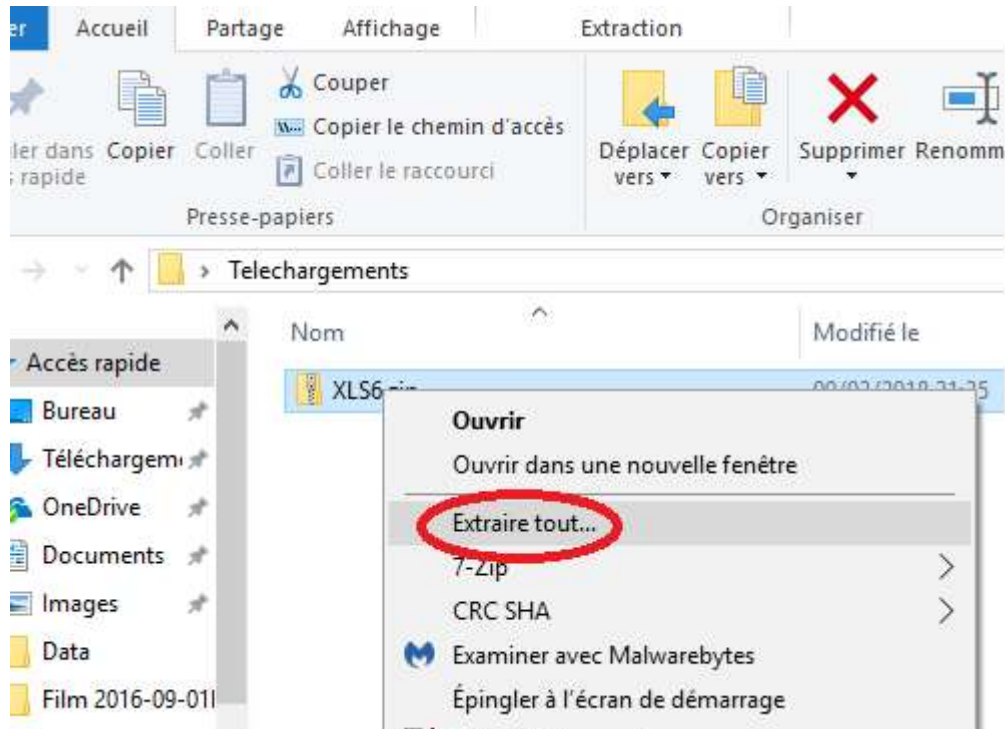


Préparation du logiciel

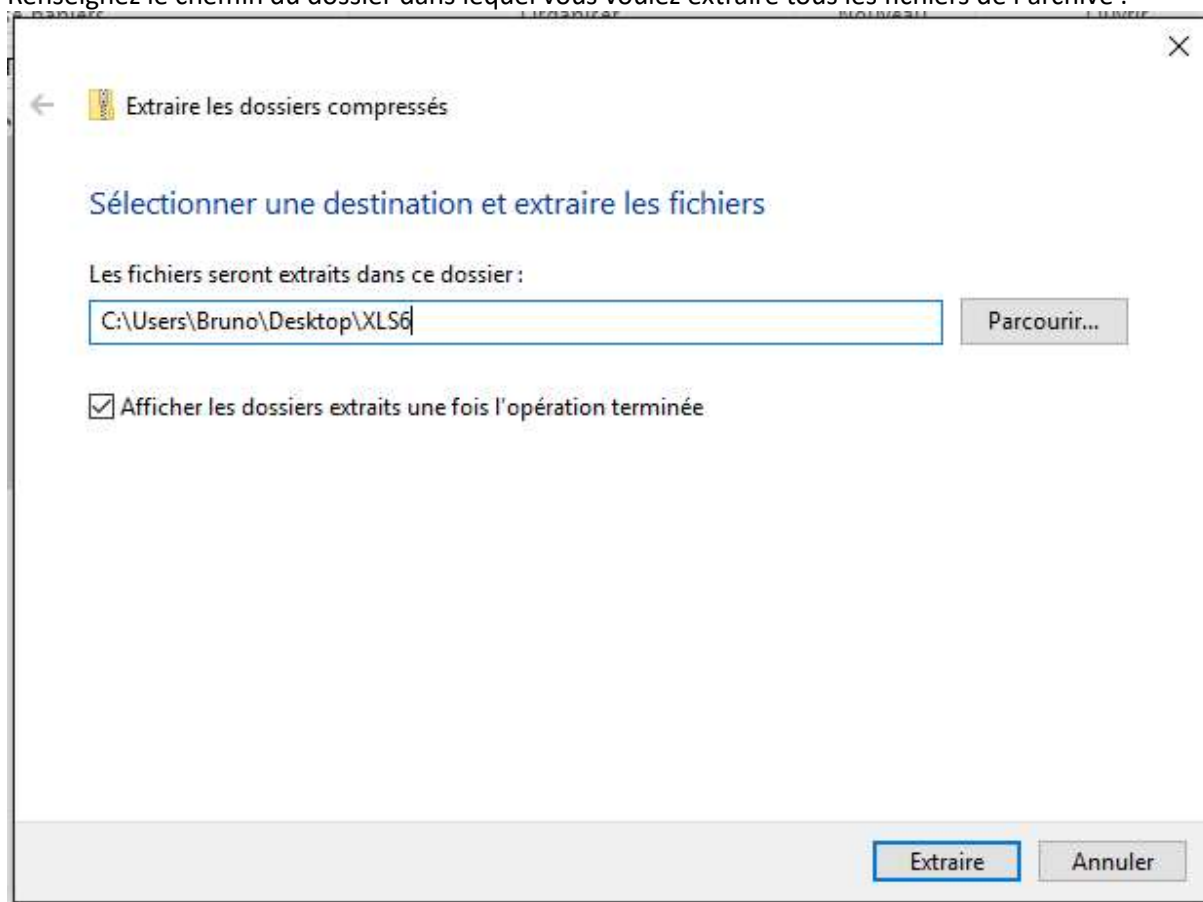
Une fois le fichier XLS6.zip téléchargé sur votre ordinateur, il sera, selon votre configuration, soit dans le répertoire « Téléchargements », soit dans le répertoire que vous avez choisi pendant le téléchargement.

Ne tentez PAS d'ouvrir quoi que ce soit en ouvrant directement le zip. Cela ne marchera pas, ou plutôt, ça risque de sembler marcher, sans que vous ne puissiez rien en faire par la suite. Au contraire :

1. Localisez le sous-répertoire dans lequel se trouve le fichier .zip
2. Faites un clic droit sur le fichier et choisissez « Extraire tout... » dans le menu contextuel

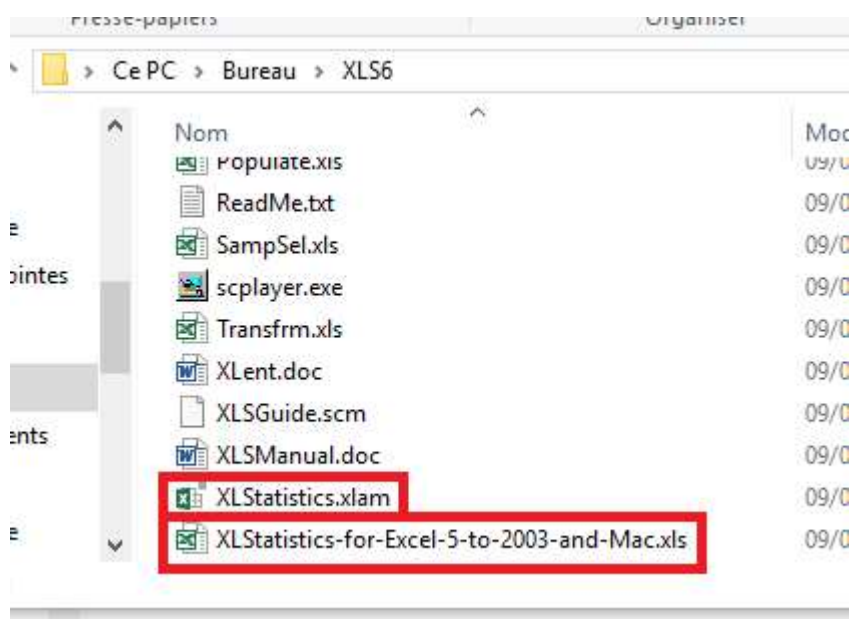


3. Renseignez le chemin du dossier dans lequel vous voulez extraire tous les fichiers de l'archive :



Une fois les fichiers extraits, vous obtenez dans le dossier de destination une centaine de fichiers Excel. Descendez jusqu'à la fin de la liste et ouvrez le fichier suivant :

- Si vous avez un PC : **XLStatistics.xlam**
- Si vous avez un Mac : **XLStatistics-for-Excel-5-to-2003-and mac.xls**



Première ouverture

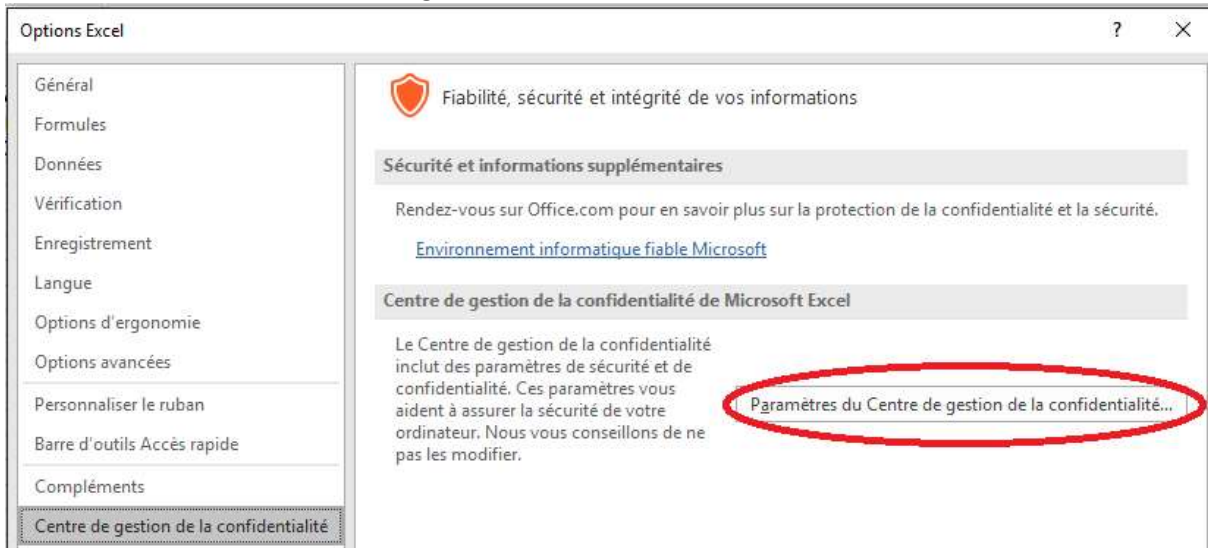
Les paramètres de sécurité de la suite Office bloqueront peut-être l'ouverture du fichier, car :

- Le fichier provient d'internet
- Il contient des macros.

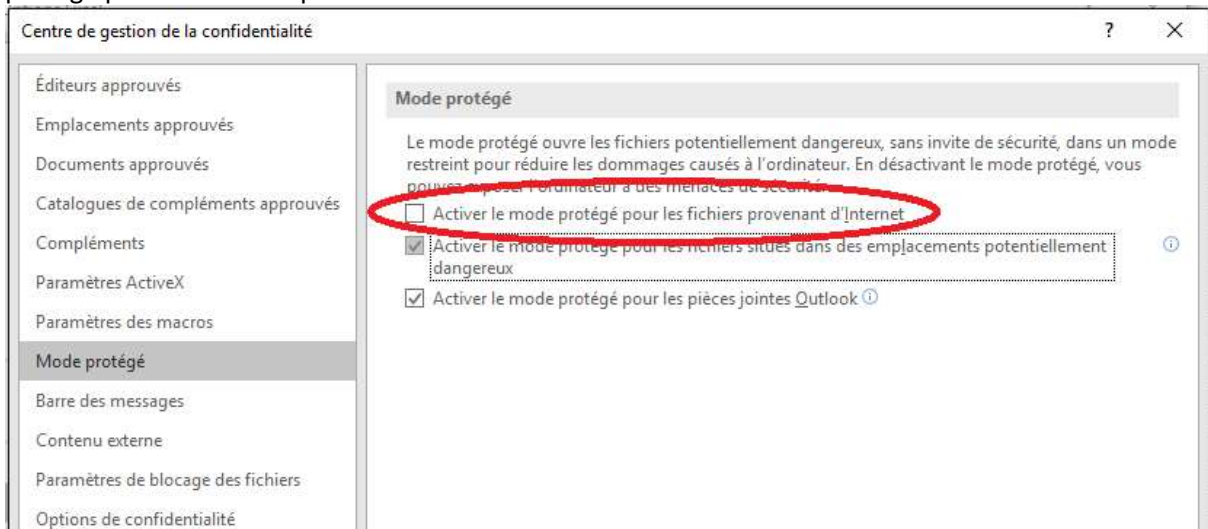
Ce sera encore plus probable si vous tentez d'ouvrir le fichier en double-cliquant dessus. Il ne se passera rien.

Si vous avez un Mac, je ne sais pas. Il faudra expérimenter. Sur PC, une procédure qui marche pour moi est la suivante :

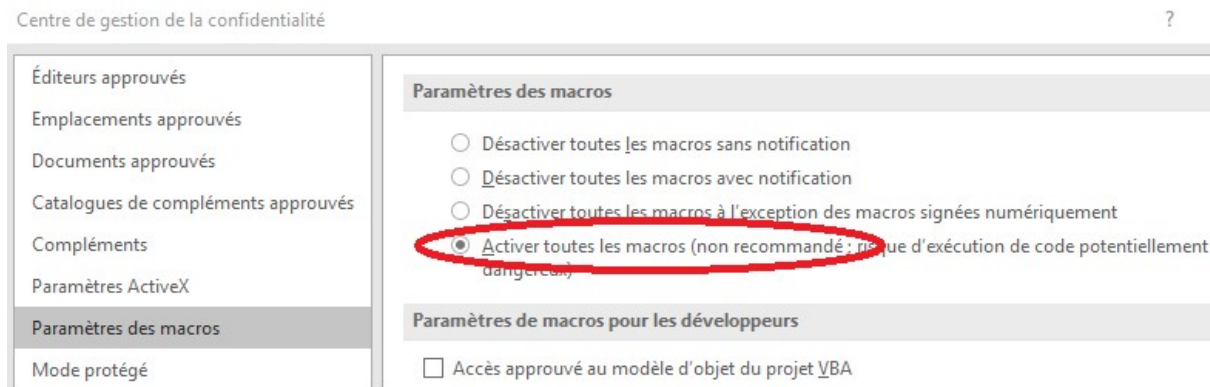
1. Ouvrez Excel
2. Allez au menu Fichier/Options et choisir l'onglet « Centre de gestion de la confidentialité », puis cliquez sur le bouton « Paramètres du centre de gestion de la confidentialité »,



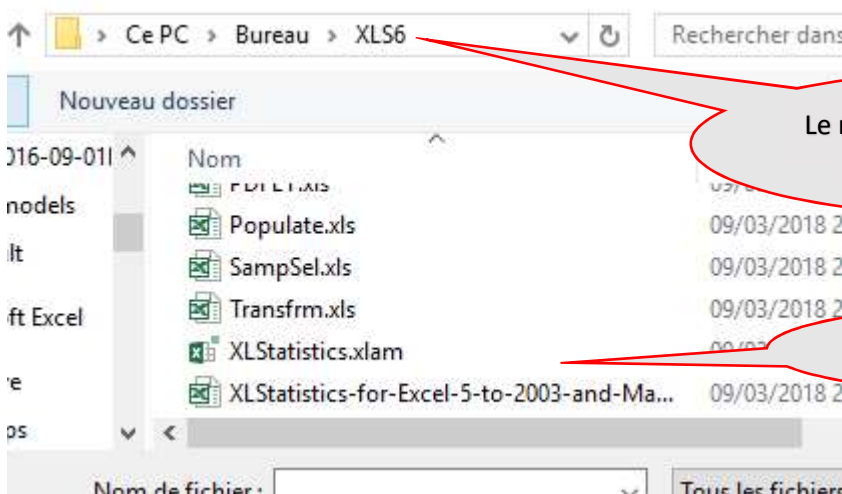
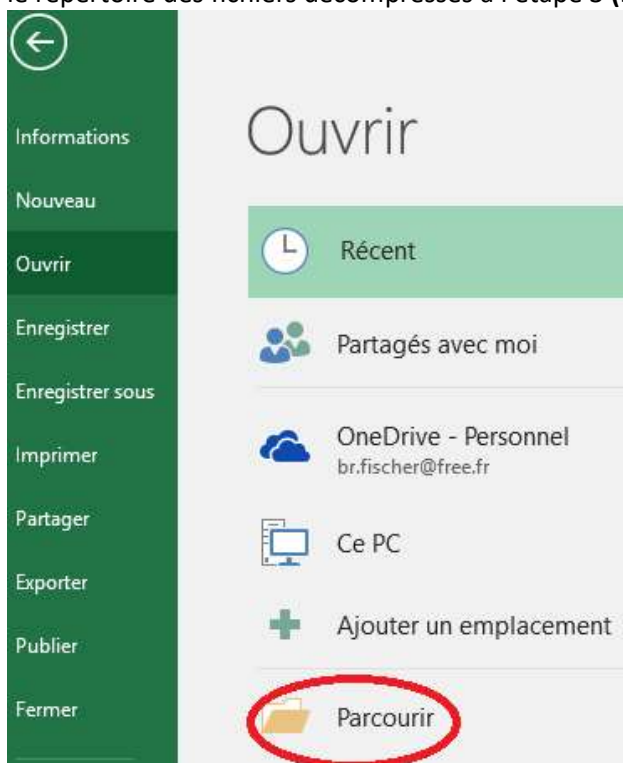
3. Dans la boîte de dialogue suivante, choisissez « Mode Protégé » et décochez la case « Activer le mode protégé pour les fichiers provenant d'Internet »



4. Choisissez maintenant l'onglet « Paramètres des macros » et choisissez « Activer toutes les macros » (c'est temporaire. Il faudra après repositionner ce réglage sur « Désactiver toutes les macros avec notification »)



5. Ouvrez maintenant le fichier XLStatistics.xlam (PC) **A PARTIR D'EXCEL** : Fichier/Ouvrir/Parcourir, puis trouvez le répertoire des fichiers décompressés à l'étape 3 (**PAS le fichier .zip !**) et ouvrez XLStatistics.xlam.



Le répertoire que vous avez choisi pour décompresser le zip

Le fichier à ouvrir

6. Vérifiez que le fichier a bien été ouvert et les macros activées. Un onglet de ruban supplémentaire nommé XLStatistics doit avoir été ajouté vers la droite du ruban, comme ça :



7. Maintenant (si ça a marché), vous pouvez remettre les options de sécurité que vous avez modifiées dans leur état initial. Sur ma machine, les ouvertures suivantes de XLStatistics ne posent plus de problème (Vous verrez seulement une boîte de dialogue vous demandant s'il faut vraiment activer les macros. Cliquez alors sur le bouton « Activer »).

Cette façon de faire fonctionne pour moi. J'espère qu'elle fonctionnera aussi pour vous. En tout cas, la clé est sans nul doute dans la manipulation des options de sécurité. Si la procédure présentée ne suffit pas, on peut alors essayer les « Emplacements approuvés »...

Commencer avec XLStatistics

Principe d'organisation du logiciel et mise en route

XLStatistics est essentiellement un jeu de classeurs Excel. Chaque classeur contient un ensemble d'outils qui sont utiles dans une situation d'analyse donnée.

Dans une large mesure, les outils d'analyse à utiliser sont déterminés par :

- Le nombre de variables à analyser simultanément
- Le type de chacune de ces variables.

XLStatistics distingue seulement 2 types de variables :

- **Qualitative** c.a.d. **Categorical** (en anglais)
- **Quantitative** c.a.d. **Numerical**.

Les commandes du ruban :

Dans le ruban « XLStatistics », la commande à invoquer est nommée en fonction de cette classification. Par exemple :

- S'il s'agit d'analyser 1 variable **quantitative** seule (Numerical), il faudra cliquer sur la commande « **1Num** »
- S'il s'agit d'analyser 1 variable **qualitative** seule (Categorical), il faudra cliquer sur la commande « **1Cat** »
- S'il s'agit d'analyser 2 variables **qualitatives** ensemble (Categorical), il faudra cliquer sur la commande « **2Cat** »
- S'il s'agit d'analyser 1 variable **quantitative** (Numerical) et 1 variable **qualitative** ensemble (Categorical), il faudra cliquer sur la commande « **1Num1Cat** »
- Etc.

Une fois appelée, la commande chargera les données dans le classeur qui convient et ouvrira celui-ci.

Les Données :

Les données doivent être dans un classeur Excel, et être organisées de façon « Tidy »¹, c.a.d. :

- Chaque colonne contient les valeurs d'une variable
- Chaque ligne contient les données relatives à un individu.
- Des individus de types différents sont classés dans des tables différentes.

De plus, la première ligne de la base de données doit contenir les noms des variables.

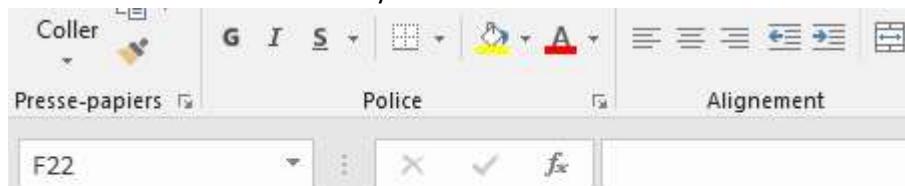
Il est également souhaitable que la base de donnée débute sur la première ligne de la feuille Excel qui la contient.

¹ Wickham, H (2014). *Tidy data*. The Journal of Statistical Software 59

Exemple :

Téléchargez sur Ipagora Le classeur Cfood mod - BFC.xls (dans les documents généraux).

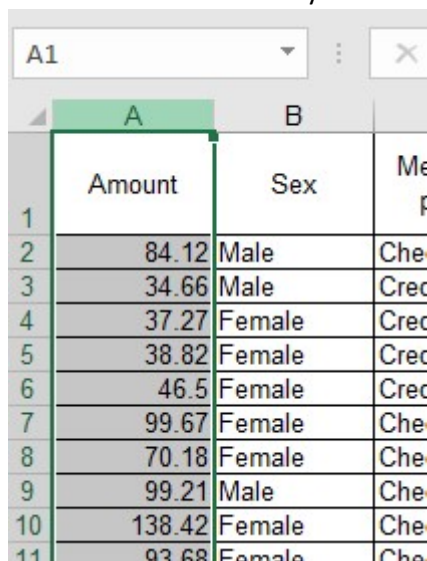
Ouvrez le fichier Excel. Vous voyez ceci :



	A	B	C	D	E
	Amount	Sex	Method of paymt	Number of Items purchased	Number of visits last month
1					
2	84.12	Male	Check	1	0
3	34.66	Male	Credit Card	9	3
4	37.27	Female	Credit Card	1	0
5	38.82	Female	Credit Card	10	2
6	46.5	Female	Credit Card	8	6
7	99.67	Female	Check	5	1
8	70.18	Female	Check	11	6
9	99.21	Male	Check	18	4

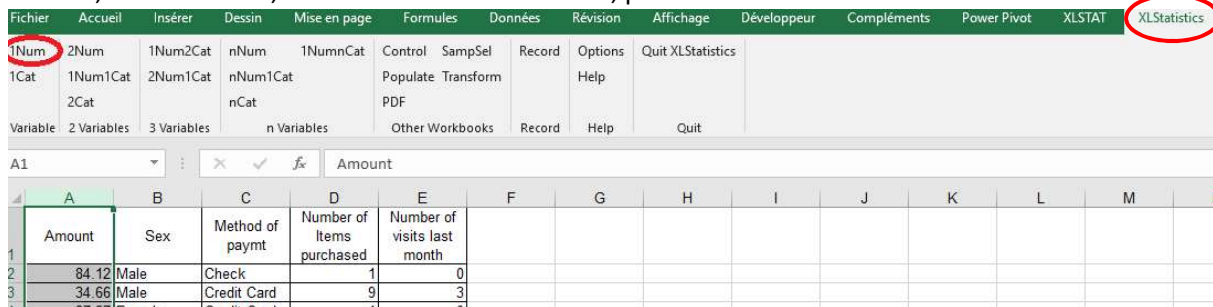
Les données sont bien organisées comme prévu.

A. Si nous voulons analyser la variable « Amount » (seule), sélectionnons d'abord la colonne :



	A	B	C
	Amount	Sex	Method of paymt
1			
2	84.12	Male	Check
3	34.66	Male	Credit Card
4	37.27	Female	Credit Card
5	38.82	Female	Credit Card
6	46.5	Female	Credit Card
7	99.67	Female	Check
8	70.18	Female	Check
9	99.21	Male	Check
10	138.42	Female	Check
11	93.68	Female	Check

Ensuite, dans le ruban, sélectionnons « XLStatistics, puis « 1Num »



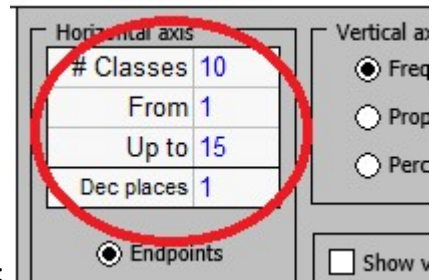
Fichier	Accueil	Insérer	Dessin	Mise en page	Formules	Données	Révision	Affichage	Développeur	Compléments	Power Pivot	XLSTAT	XLStatistics
1Num	2Num	1Num2Cat	nNum	1NumnCat	Control	SampSel	Record	Options	Quit XLStatistics				
1Cat	1Num1Cat	2Num1Cat	nNum1Cat		Populate	Transform		Help					
2Cat			nCat		PDF								
Variable	2 Variables	3 Variables	n Variables		Other Workbooks		Record	Help	Quit				

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
	Amount	Sex	Method of paymt	Number of Items purchased	Number of visits last month									
1														
2	84.12	Male	Check	1	0									
3	34.66	Male	Credit Card	9	3									

Un nouveau classeur s'ouvre alors, qui contient ce qu'il faut pour l'analyse.

Dans chaque classeur d'analyse :

- On peut saisir des données dans les cellules en bleu. Par exemple, pour paramétrer l'histogramme



proposé par le logiciel, on peut changer les valeurs ici :

- On peut obtenir des explications en double-cliquant sur les cellules en rouge, comme
- On peut naviguer d'une fonctionnalité à l'autre en utilisant les flèches bleues et

Frequency Chart

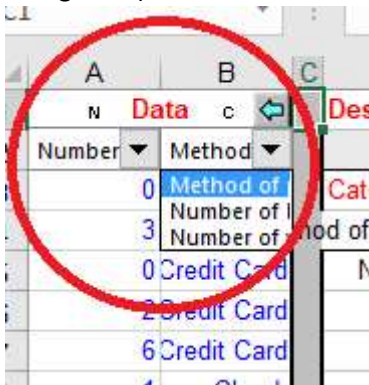
B. Si nous voulons analyser le lien entre Method of Payment (qualitative) et Number of visits (quantitative), on sélectionne un bloc de colonnes qui contient ces deux colonnes, puis on appelle la commande 1Num1Cat :

Fichier	Accueil	Insérer	Dessin	Mise en page	Formules	Données	Révision	Affichage	Développeur	Compléments	Power Pivot	XLSTAT	XLStatistics
1Num	2Num	1Num2Cat	nNum	1NumnCat	Control	SampSel	Record	Options	Quit XLStatistics				
1Cat	1Num1Cat	2Num1Cat	nNum1Cat	nCat	Populate	Transform	Record	Help					
Variable	2 Variables	3 Variables	n Variables		PDF	Other Workbooks	Record	Help	Quit				

	A	B	C	D	E	F	G	H	I	J	K	L	M
	Amount	Sex	Method of paymt	Number of Items purchased	Number of visits last month								
1													
2	84.12	Male	Check	1	0								
3	34.66	Male	Credit Card	9	3								
4	37.27	Female	Credit Card	1	0								
5	38.82	Female	Credit Card	10	2								
6	46.5	Female	Credit Card	8	6								
7	99.67	Female	Check	5	1								
8	70.18	Female	Check	11	6								
9	99.21	Male	Check	18	4								
10	138.42	Female	Check	32	2								
11	93.68	Female	Check	9	6								

On obtient un message d'avertissement : « Too many categories », qui vient du fait que les variables ne sont pas dans le bon ordre.

A gauche en haut, on utilise les zones de listes pour affecter chaque variable à la bonne colonne (N = Numerical, C = Categorical)



Voilà. Il vous reste à expérimenter un peu. Choisissez un ensemble de données que vous connaissez et essayez de voir ce que vous pouvez en faire avec XLStatistics.

Remarque : L'usage de chaque classeur est rappelé dans la feuille « Information » de chaque classeur.

Bon travail.