

Macro complémentaire « **Boxplot.xla** »

Olivier MARTIN omartin@inapg.inra.fr

Installation :

1. Copier le fichier **Boxplot.xla** dans le dossier des macros complémentaire d'Excel. Ce dossier est identifiable en recherchant la localisation des fichiers d'extension **.xla** sur la machine (**Démarrer/Rechercher : *.xla**)
2. Lancer Excel
3. Dans le menu **Outils/Macros complémentaires**, cocher la macro **Box Plot**

Exécution :

La macro complémentaire Box Plot contient une procédure automatique qui crée dans le menu **Outils** une commande d'exécution **Box Plot...** Lorsque la macro complémentaire est désactivée cet élément de menu est supprimé.

Description :

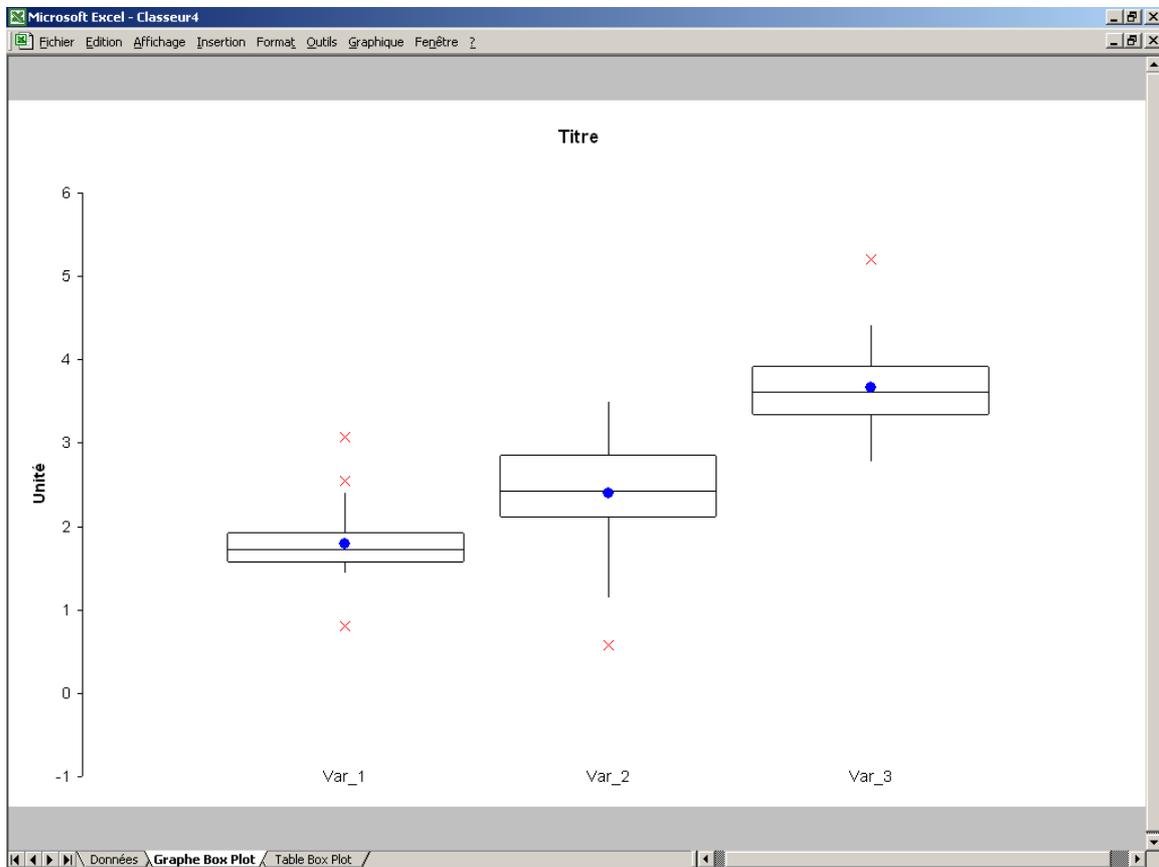
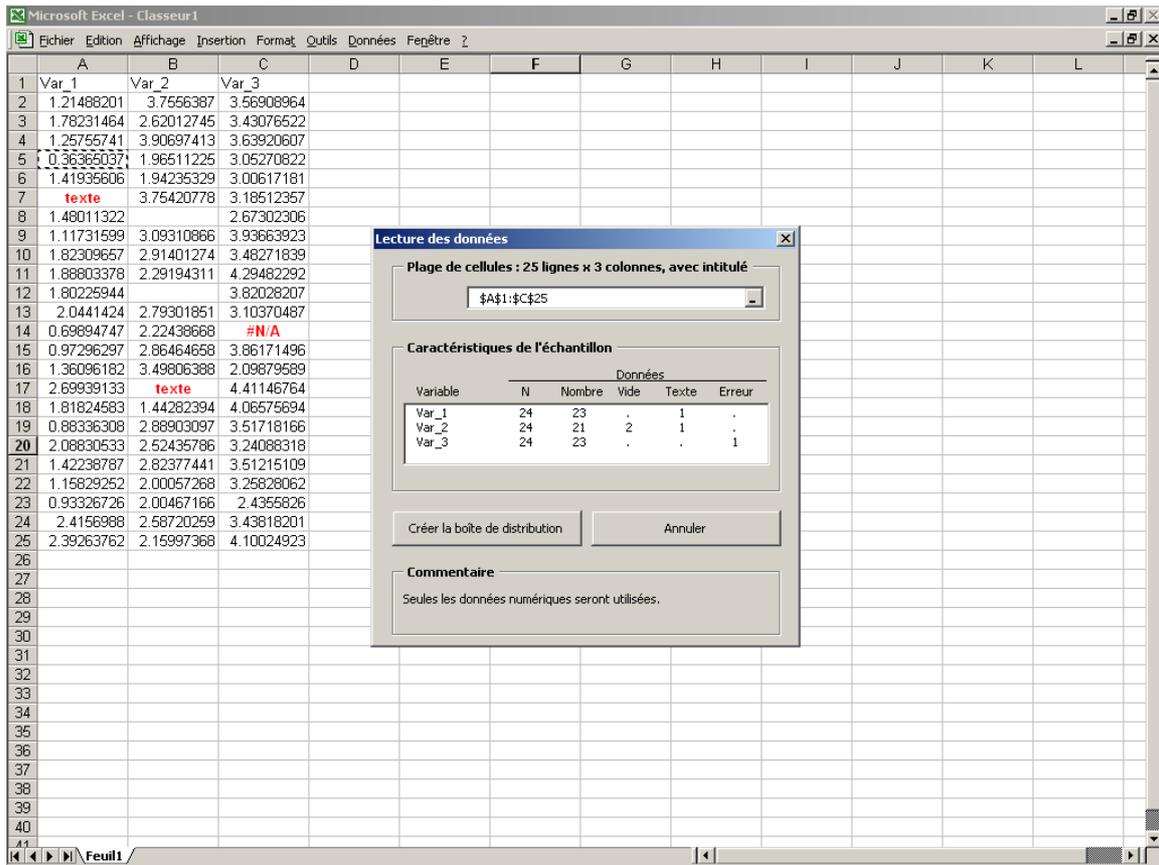
La macro complémentaire Box Plot construit une boîte de distribution verticale de une ou plusieurs variables. A l'exécution, la procédure recherche automatiquement une plage de données (en colonne) dans le classeur actif, détecte la présence d'intitulés sous forme de texte, et compile l'échantillon en caractérisant les données lues (Nom de variables, nombre de cellules, données numériques, données textuelles, données manquantes, valeurs d'erreur). L'utilisateur peut modifier la plage de cellules proposées par défaut en indiquant la plage à prendre en compte. Le bouton **Annuler** ferme la boîte de dialogue. Le bouton **Créer la boîte de distribution** lance la procédure. Ce bouton n'est pas accessible si la plage de cellules ne contient pas de donnée numérique.

A l'exécution, la procédure ouvre un nouveau classeur contenant 3 feuilles :

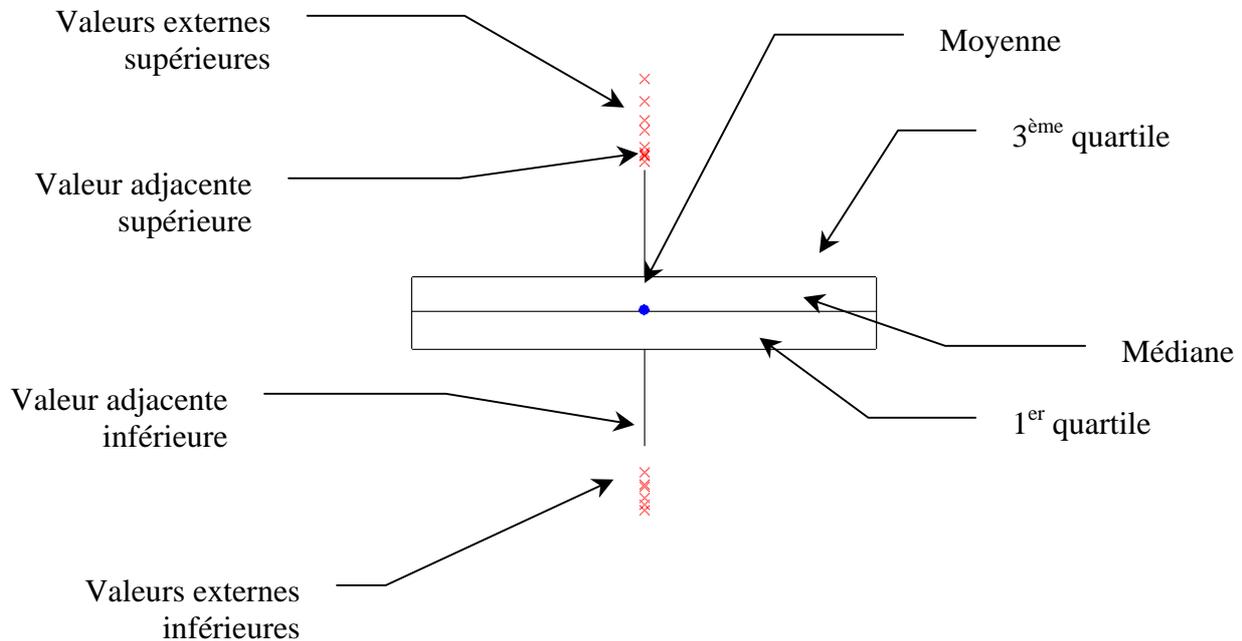
La feuille **Données** contient une copie de la plage de cellule limitée aux données numériques.

La feuille **Table Box Plot** contient les valeurs calculées à partir de l'échantillon nécessaires à la construction du graphe Box Plot.

La feuille **Graphe Box Plot** contient le graphique de la boîte de distribution (Par défaut un des titres du graphique et de l'axe des ordonnées est inclus).



La boîte de distribution pour chaque variable est construite selon la méthode donnée par W.S. Cleveland (In : *Visualizing Data*. Hobart Press, Summit, New Jersey, U.S.A, 1st edition, 1993).

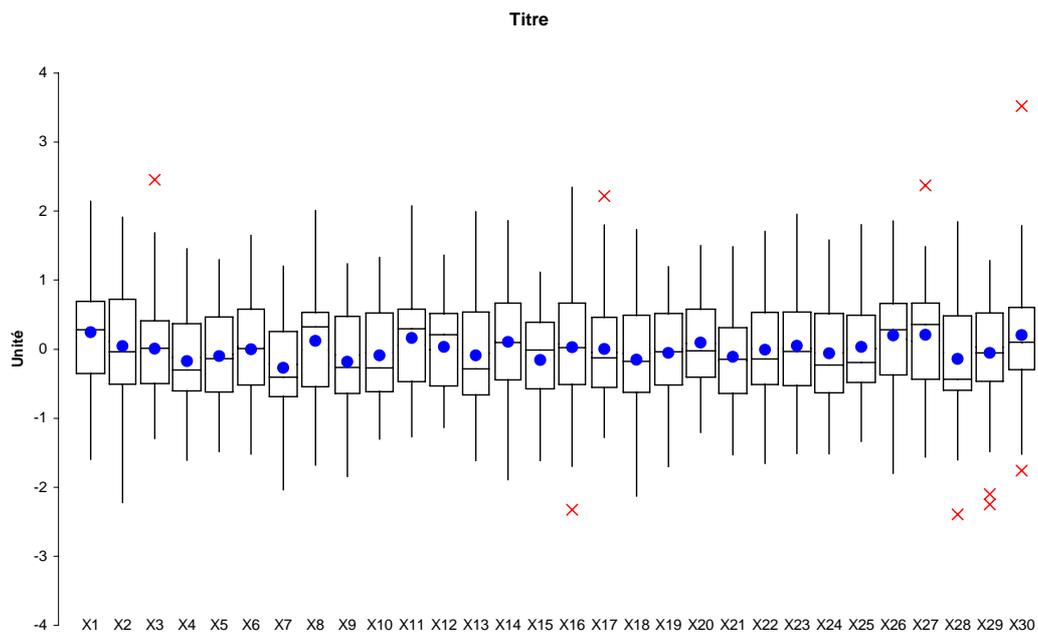
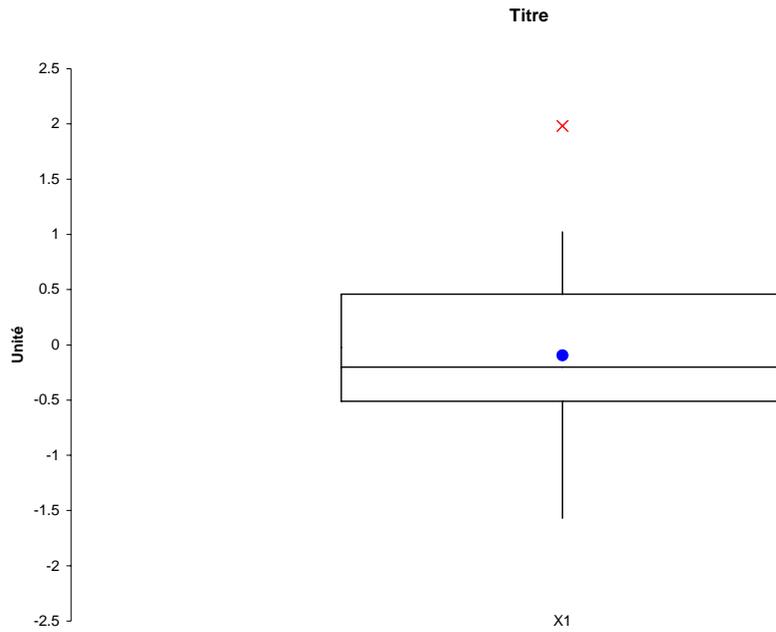


Le cadre de la boîte de distribution est délimité par les 1^{er} et 3^{ème} quartiles (Q1 et Q3) de la distribution (50% des données se situent donc dans le rectangle). La médiane est représentée par une ligne horizontale dans la boîte et la moyenne par un cercle plein bleu.

Les extrémités des « moustaches » ou « pattes » de la boîte de distribution correspondent aux valeurs adjacentes supérieures et inférieures de la distribution, i.e. respectivement la plus grande valeur inférieure ou égale à $Q3 + 1.5 \times (Q3 - Q1)$ et la plus petite valeur supérieure ou égale à $Q1 - 1.5 \times (Q3 - Q1)$.

Les valeurs externes supérieures et inférieures, représentées par des croix rouges, sont les valeurs respectivement supérieures à la valeur adjacente supérieure et inférieure à la valeur adjacente inférieure.

Enfin la largeur de la boîte est proportionnelle au nombre de données numériques pour la variable considérée.



Ci-dessus deux exemples de boîtes de distribution d'une seule ou de trente variables.

La construction « manuelle » d'une boîte de distribution avec Excel est détaillée dans **L'analyse statistique des données. Apprendre, comprendre et réaliser avec Excel. Editions Ellipses, 2005.**