

TP31 – Math et Stat

Resource : stats.xlsx

1. Dans la feuille **stat**,
 - a. Calculer la moyenne μ (**AVERAGE**), la médiane (**MEDIAN**), le minimum (**MIN**), le maximum (**MAX**), le mode (**MODE**) et l'écart type (**STDEV**) de la variable X.
 - b. Vérifier que l'écart type n'est autre que la racine (**SQRT**) de la variance calculée par la somme (**SUM**) suivante :

$$\text{Var}(X) = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$$

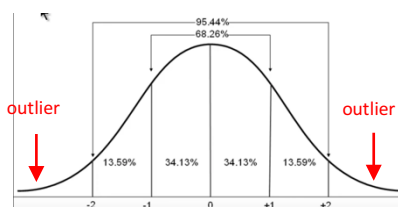
$$\sigma = \sqrt{\text{Var}(X)}$$

(Remarque : **x** à la puissance **p** s'écrit **x^p**)

- c. Détecter les valeurs aberrantes ('outliers') à 2σ de la moyenne. Utiliser les formules de comparaison suivantes :

$$\text{outlier} < \mu - 2 * \sigma$$

$$\mu + 2 * \sigma < \text{outlier}$$



2. Dans la feuille **pourcentatg**, calculer le pourcentage d'augmentation des naissances entre deux années et le taux de naissance dans chaque ville par 100000 habitants.
 - a. Pourcentage = **(NOUVEAU - ANCIEN) / ANCIEN**
 - b. Taux = **(ÉVÉNEMENTS / POPULATION) * UNITÉ**
3. Dans la feuille **histogram**, afficher l'histogramme de la variable 'data'. Utiliser les 'bins' (intervalles de fréquences) spécifiés dans la variable 'bins'.
 - a. Activer le Add-in Analysis ToolPack : tools->Excel Add-ins -> cocher 'Analysis ToolPack'
 - b. Data-> 'Data Analysis' -> Histogram
 - c. Sélectionner les cellules d'entrée (Input Range) et les cellules des 'bins' (bins Range) et choisir les cellules de sortie (Output Range)
 - d. Détecter les outliers à 2σ de la moyenne.
4. Dans la feuille **regression**, tracer la droite descriptive de la variable Y
 - a. Tracer Y en fonction de X. Insert -> X Y scatter -> scatter
 - b. Sélectionner les points -> click droit -> Add Trend line -> Linear
 - c. Afficher l'équation de la droite ($Y = aX + b$)
 - d. Quelle est la prédiction de la valeur $X = 1$