

Statistiques et tableur

Récupérer le fichier SIO1stats.ods présent dans Dossup/...

Le copier sur son propre poste.

Le renommer en monNomSIO1stats.ods

L'ouvrir, de préférence avec Libre Office.

En fin de séance, déposer son fichier monNomSIO1stats.ods dans Dossup...

1. Statistiques à une variable

Travail sur la feuille : stats1var

Cette feuille contient des données de l'INSEE⁴ concernant la population de France en 2019.

Calcul de la moyenne et de l'écart-type de l'ensemble

1. Entrer une formule dans la case F8 ; recopier cette formule vers le bas (jusqu'à F108) pour obtenir tous les produits des âges par l'effectif de l'ensemble correspondant.
2. Dans F111, calculer la somme de la colonne F.
3. Dans G2, faire afficher l'âge moyen de l'ensemble (utiliser F111...)
4. Pour calculer l'écart-type de l'ensemble, entrer une formule dans G8, et copier cette formule vers le bas (jusqu'à G108) : les carrés des écarts à la moyenne (de l'ensemble) multipliés par les effectifs correspondants seront ainsi calculés.
5. Dans G111, calculer la somme de la colonne G.
6. Dans K2, faire afficher l'écart-type de l'ensemble (utiliser G111...)

Calcul de la moyenne et de l'écart-type des hommes

Utiliser les colonnes H et I et procéder comme ci-dessus.

L'âge moyen des hommes devra apparaître dans G3 et l'écart-type dans K3.

Calcul de la moyenne et de l'écart-type des femmes

Utiliser les colonnes J et K et procéder comme ci-dessus.

L'âge moyen des femmes devra apparaître dans G4 et l'écart-type dans K4.

4. <https://www.insee.fr/fr/statistiques/1892086?sommaire=1912926>, consulté le 19 septembre 2019.

2. Statistiques à deux variables

Travail sur la feuille : stats2var

Cette feuille contient les indices des prix à la consommation (base 2015) pour l'ensemble des ménages de France⁵.

C'est une série chronologique : on peut la voir comme une série à deux variables : x est le numéro du mois et y l'indice des prix.

On veut calculer le coefficient de corrélation linéaire r et l'équation $y = ax + b$ de la droite d'ajustement obtenue par la méthode des moindres carrés « à la main » (sans utiliser les fonctions prédéfinies du tableur).

Voici les formules permettant de calculer r , a et b , pour des séries (x_i) et (y_i) de taille n :

$$a = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} = \frac{COV(x,y)}{V(x)} \quad b = \bar{y} - a\bar{x} \quad r = \frac{COV(x,y)}{\sigma_x \sigma_y}$$

$COV(x,y)$ est la covariance de x et y , $V(x)$ est la variance de x , $\sigma_x = \sqrt{V(x)}$ est l'écart-type de la série x .

1. Sous le tableau, faire apparaître le nuage de points (insérer/diagramme/...)
2. Comme dans la première partie, utiliser le tableur pour calculer \bar{x} , \bar{y} , σ_x , σ_y , $COV(x,y)$, et enfin, r , a et b .
3. Retrouver r , a et b à l'aide des fonctions prédéfinies du tableur.

5. <https://www.insee.fr/fr/statistiques/3532441?sommaire=3530678> consulté le 19 septembre 2019.