

TD 13

Statistiques univariées.

❑ 01 Le nombre de sénateurs par département de France métropolitaine en 2024 est résumé dans le tableau suivant (<https://www.senat.fr>) :

Nombre de sénateurs	1	2	3	4	5	6	7	8	11	12
Nombre de départements	8	35	24	9	8	6	3	1	1	1

1. Déterminer à l'aide de python la moyenne de cette série.
2. Déterminer la médiane, le premier et le troisième quartile de cette série.
3. Tracer la boîte à moustaches représentant cette série.
4. Déterminer, à l'aide de python, la variance et l'écart-type cette série.

❑ 02 On donne une série par le tableau suivant où la première colonnes contient les modalités et la seconde les fréquences cumulées croissantes :

modalité	fcc
-1	0.08
0	0.15
2	0.32
4	0.64
5	0.75
10	1

1. Tracer la courbe des fréquences cumulées croissantes de cette série.
2. Déterminer la médiane, le premier et le troisième quartile de cette série.
3. Déterminer le quantile d'ordre 0.10.
4. Sachant que la population étudiée a un effectif de 300, résumer la série par un tableau modalités/effectifs.

5. En déduire, à l'aide de python, la moyenne et l'écart-type de cette série.

❑ 03 On a téléchargé les fichiers `chomeurs_catA_milliers.csv` et `emplois_vacants.csv` dans le dossier `documentscsv`.
Le fichier `chomeurs_catA_milliers.csv` contient (notamment) le nombre de millier de chômeurs de catégorie A par trimestre en France et le fichier `emplois_vacants.csv` contient le nombre d'emplois vacants⁽¹⁾ par trimestre en France (source : <https://dares.travail-emploi.gouv.fr/>).

1. Décrire les deux séries statistiques (population, caractère étudié, effectif).
2. Déterminer
 - (a) les moyennes et écart-types,
 - (b) les médianes,
 - (c) les boîtes à moustaches.

3. Commenter les résultats précédents.
Remarque : la commande `tableau["colonne"][i:j]` renvoie une colonne constituée des valeur de la colonne "colonne" du dataframe `tableau` de la ligne `i` à la ligne `j-1`.

❑ 04 Démontrer la formule de Koenig.

❑ 05 Soit $n \in \mathbb{N}^*$ et $x = (x_1, x_2, \dots, x_n)$ une série statistique réelle.

1. Soient $a \in \mathbb{R}$ et $y = (ax_1, ax_2, \dots, ax_n)$. Déterminer la variance de y en fonction de celle de x .
2. Soient $b \in \mathbb{R}$ et $z = (x_1 + b, x_2 + b, \dots, x_n + b)$. Déterminer la variance de z en fonction de celle de x .

(1) : Les emplois dits « vacants » sont des postes libres, nouvellement créés ou inoccupés, ou encore occupés et sur le point de se libérer, pour lesquels des démarches actives sont entreprises pour trouver le candidat convenable. Le recrutement souhaité peut correspondre à un contrat à durée indéterminée (CDI), un contrat à durée déterminée (CDD), ou à un emploi saisonnier, même de courte durée. Les données couvrent les établissements des entreprises de 10 salariés ou plus du champ privé, situés en France métropolitaine et dans les départements-régions d’Outremer (Drom) hors Mayotte. Sont exclus du champ l’agriculture, l’intérim, les particuliers employeurs et les emplois publics.