

TD Statistique en Python.

A Statistique uni-varié

Pour cette partie vous traiterez tous les exercices dans un même fichier nommé "univarie"

On notera : X la liste des valeurs x_i classé dans l'ordre croissant et N la liste des effectifs de chacune de ces valeurs.

Dans un premier temps, on testera les exercices ci-dessous à l'exemple ci-dessous. Les notions mathématiques mathématiques ainsi que les valeurs à obtenir sont sur le document "Notions et exemples pour le TD Statistique en Python.". Vous avez aussi à votre disposition un glossaire.

Exemple : Liste de notes obtenues pour une promotion de 120 étudiants.

x_i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
n_i	3	5	6	5	6	7	7	10	17	20	25	21	23	12	10	5	7	5	3	2	1

Exercice 1. Écrire une fonction **Moyenne** qui prend comme argument la liste des X et la liste des effectifs N et renvoie la **moyenne** de la série.

Exercice 2. Écrire une fonction **effectcumul** qui prend comme argument la liste des effectifs N et renvoie la **liste des effectifs cumulés croissants** de la série.

Exercice 3. Écrire une fonction **Mediane** qui prend comme argument la liste des X et la liste des effectifs et renvoie la **médiane** de la série.

```
def Mediane(X,N):
    i = 0
    m = 0
    P = effectcumul(N)
    total = P[len(N)-1]
    while m < total//2:
        m = m+N[i]
        i = i+1
    return X[i-1]
```

Exercice 4. Écrire une fonction **Quart1** qui prend comme argument la liste des X et la liste des effectifs et renvoie le **premier quartile** de la série.

Exercice 5. Écrire une fonction **Quart3** qui prend comme argument la liste des X et la liste des effectifs et renvoie le **troisième quartile** de la série.

Exercice 6. Écrire une fonction **Variance** qui prend comme argument la liste des X et la liste des effectifs et renvoie le **variance** de la série.

Exercice 7. Écrire une fonction **Ecartype** qui prend comme argument la liste des X et la liste des effectifs et renvoie le **l'écart type** de la série.

B Statistique bi-varié

Pour cette partie vous traiterez tous les exercices dans un même fichier nommé "bivarie"

Exemple 1. On considérera dans cette partie l'exemple ci-dessous pour tester les procédures :

Un négociant a fait mener une étude visant à déterminer à quel prix maximal ses clients sont prêts à acheter un certain produit. Les résultats sont regroupés dans le tableau suivant :

Prix maximal en euros du produit : x_i	5	10	15	20	25	30
Nombre d'acheteurs potentiels sur une journée : y_i	84	58	30	19	7	4

On notera X la liste des prix et Y la liste constituée du nombre d'acheteurs potentiels correspondant pour une journée.

Exercice 8. Écrire une fonction **Nuage** qui prend comme argument la liste des X et la liste des Y et renvoie le **graphique** formé du **nuage de points** de la série. (Pour cette exercice, vous devrez charger la librairie "matplotlib" par "import matplotlib.pyplot as plt")

Finir par "plt.show()" pour observer le nuage de point.

```
def Nuage(X,Y):
    plt.plot(X,Y,"r.")
    plt.show()
```

Exercice 9. Écrire une fonction **moyenne1** qui prend comme argument une liste X et renvoie la **moyenne** de la série X . Déterminer à l'aide de cette fonction les coordonnées du point moyen.

Exercice 10. Écrire une fonction **ecart1** qui prend comme argument une liste X et renvoie l'**écart type** de la série X . Déterminer à l'aide de cette fonction l'écart type des séries X et Y .

Exercice 11. Écrire une fonction **covariance** qui prend comme argument la liste des X et la liste des Y et renvoie la **covariance** de la série.

Exercice 12. Écrire une fonction **coef** qui prend comme argument la liste des X et la liste des Y et renvoie le **coefficient de variance** de la série.

Exercice 13. Écrire une fonction **coefdir** qui prend comme argument la liste des X et la liste des Y et renvoie le **coefficient directeur** de la droite de régression obtenue par la méthode des moindres carrés de y en x .

Exercice 14. Écrire une fonction **coeford** qui prend comme argument la liste des X et la liste des Y et renvoie l'**ordonnée à l'origine** de la droite de régression obtenue par la méthode des moindres carrés de y en x .

Exercice 15. Écrire une fonction **nuageaffine** qui prend comme argument la liste des X et la liste des Y et renvoie le **graphique** formé du nuage de points ainsi que de la droite de régression de y en x .

Exercice 16. Écrire une fonction **nuageaffine2** qui prend comme argument la liste des X et la liste des Y et renvoie le **graphique** formé du nuage de points ainsi que les droites de régression de y en x et x en y .

C Exercices d'application des procédures ci-dessus.

Exercice 17. Appliquer les procédures ci-dessus à la série :

Le tableau suivant résume les résultats obtenus par les élèves d'une classe de première lors d'un devoir.

Notes	2	4	5	7	8	9	10	11	12	14	15	18
Effectifs	1	2	3	2	3	4	5	3	3	2	1	1

Utiliser les procédures ci-dessus pour résoudre les exercices ci-dessous (pour les nuages de points vous ne suivrez pas les instructions bien évidemment) :

Exercice 18. .

Le tableau suivant recense, par clinique, le nombre de postes du personnel non médical en fonction du nombre de lits de la clinique.

Clinique	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}
Nombre de lits : x_i	122	177	77	135	109	88	185	128	120	146	100
Nombre de postes : y_i	205	249	114	178	127	122	242	170	164	188	172

1. Représenter le nuage de points associé à la série statistique (X, Y) dans le plan rapporté à un repère orthogonal en prenant pour unités graphiques : 1 cm pour 20 lits en abscisse et 1 cm pour 50 postes en ordonnée.
2. Déterminer le coefficient de corrélation linéaire entre X et Y et donner une équation de la droite de régression de Y en X par la méthode des moindres carrés.
3. Une clinique possède 35 lits. En utilisant les résultats obtenus en 2°, déterminer combien devrait-elle embaucher de personnel occupant un poste non médical à temps plein ?
4. En réalité, cette clinique dispose de 60 postes. Calculer les différences entre le nombre de poste réel et le nombre de poste théorique obtenu précédemment. Quel pourcentage cette différence représente-t-elle par rapport à la situation théorique ?

Exercice 19. .

Le tableau suivant recense, par clinique, le nombre de postes du personnel non médical en fonction du nombre de lits de la clinique.

Clinique	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}
Nombre de lits : x_i	122	177	77	135	109	88	185	128	120	146	100
Nombre de postes : y_i	205	249	114	178	127	122	242	170	164	188	172

1. Représenter le nuage de points associé à la série statistique (X, Y) dans le plan rapporté à un repère orthogonal en prenant pour unités graphiques : 1 cm pour 20 lits en abscisse et 1 cm pour 50 postes en ordonnée.
2. Déterminer le coefficient de corrélation linéaire entre X et Y et donner une équation de la droite de régression de Y en X par la méthode des moindres carrés.
3. Une clinique possède 35 lits. En utilisant les résultats obtenus en 2°, déterminer combien devrait-elle embaucher de personnel occupant un poste non médical à temps plein ?
4. En réalité, cette clinique dispose de 60 postes. Calculer les différences entre le nombre de poste réel et le nombre de poste théorique obtenu précédemment. Quel pourcentage cette différence représente-t-elle par rapport à la situation théorique ?