

EXERCICES

ANALYSE DE DONNÉES STATISTIQUES

Exercice 1 : Pour chacune des séries statistiques suivantes, calculer la moyenne et donner la médiane, les quartiles et les déciles.

1. 36,37,38,39,40,41,42,43,44 ;
2. 1, 2, 2, 4, 4, 5, 8, 8, 8, 80, 80, 90
3. 10, 10, 10, 1100, 2500, 3000, 4000, 5000
4. Les élèves d'une section de BTS ont obtenus les notes suivantes à un devoir de mathématiques :

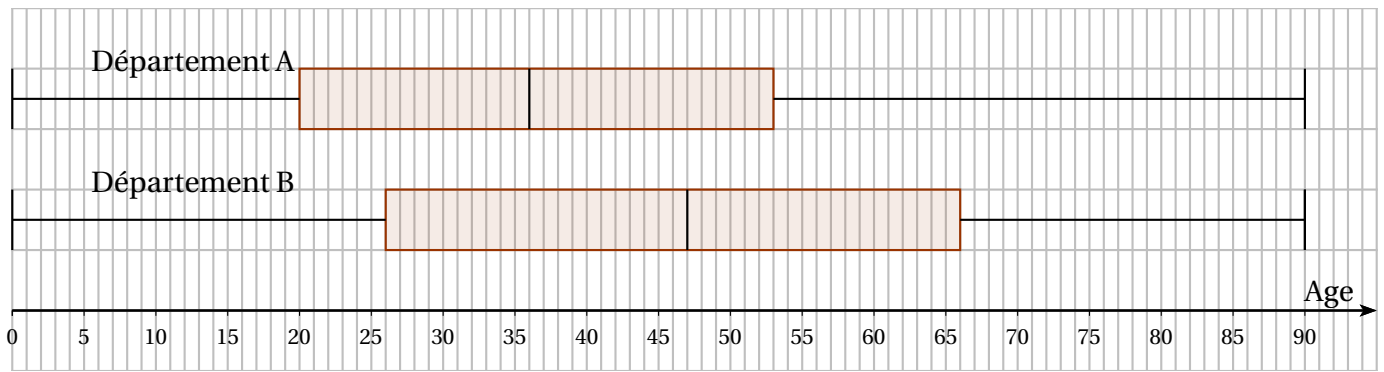
Note	8	9	10	11	12	13	14
Effectif	6	2	7	5	9	19	12

Exercice 2 : Ce tableau donne les salaires mensuels, en euros, des 185 employés d'une entreprise.

Salaire brut	1405	1480	1554	1870	2739	4215
Effectifs	40	15	81	35	9	5
Effectifs cumulés croissants						

1. Compléter le tableau ci-dessus.
2. Déterminer pour cette entreprise le salaire médian et les quartiles Q1 et Q3.
3. Représenter le diagramme en boîte de cette série.
4. Déterminer l'étendue et l'écart interquartile.
5. Déterminer le salaire moyen et l'écart type de cette série.
6. Utiliser les indicateurs précédents pour répondre aux questions suivantes et justifier :
 - a. Les salaires sont-ils dispersés ?
 - b. Si les salaires étaient égaux, quelle somme recevrait chaque employé ?
 - c. Peut-on affirmer ici que la moitié des salariés gagnent moins que la moyenne ?
 - d. Au moins 25% des employés ont un salaire inférieur à 1450€.
 - e. Au moins 50% des employés ont un salaire supérieur à 1500€.
 - f. Au moins 25% des employés ont un salaire entre 1500 et 1800€.
 - g. Au moins 25% des employés ont un salaire supérieur à 1800€.

Exercice 3 : On donne ci-après les diagrammes en boîtes représentant les âges des habitants de moins de 90 ans en 2012 pour les deux départements français, nommés ici A et B. Par lecture graphique, avec la précision permises par les diagrammes, justifier les affirmations suivantes en précisant les paramètres statistiques utilisés.



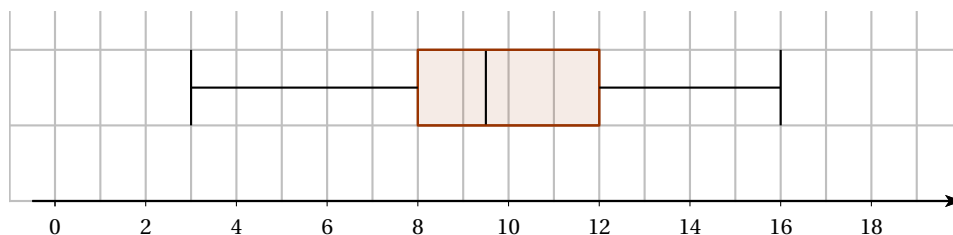
1. La proportion de personnes âgées de moins de 25 ans est inférieure dans le département B à la proportion analogue dans le département A.
2. La dispersion des âges des habitants autour de l'âge médian est plus importante pour le département B que pour le département A.

Exercice 4 : Dans un lycée, on étudie les moyennes trimestrielles du premier trimestre de deux classes appelées respectivement Jaune et Rouge.

1. Les 25 élèves de la classe Jaune ont obtenu les moyennes trimestrielles suivantes au premier trimestre :

Notes	3	4	5	7	10	11	12	13	14	15	16	18
Effectifs	1	1	1	2	5	2	5	3	1	1	1	1
Effectifs cumulés croissants												

- a. Déterminer la médiane Me , le premier quartile $Q1$ et le troisième quartile $Q3$ de cette série statistique.
 - b. Représenter le diagramme en boîte correspondant.
 - c. Corriger l'une des deux phrases suivantes afin qu'elles soient toutes les deux vraies :
 - ↪ au moins 25% des élèves de la classe Jaune n'ont pas la moyenne
 - ↪ au moins 75% des élèves de la classe Jaune ont la moyenne
 - d. Calculer la moyenne trimestrielle de la classe Jaune (obtenu en faisant la moyenne des moyennes individuelles)
2. On donne le diagramme en boîte suivant de la classe Rouge :

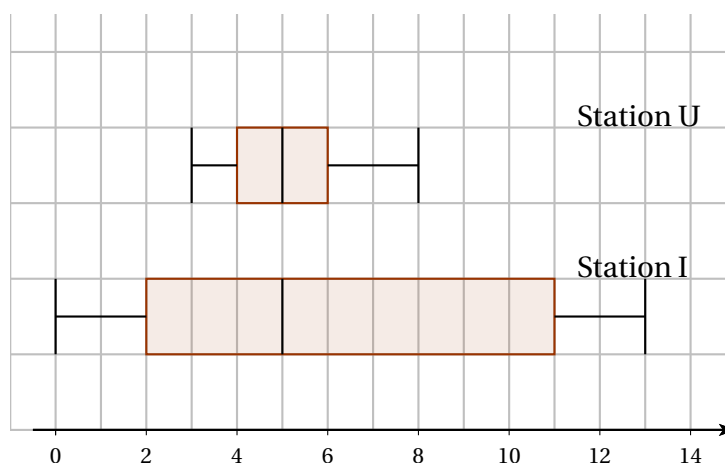


Parmi les affirmations suivantes, lesquelles sont vraies, fausses ou indécidables (ie que l'on ne peut pas conclure avec les éléments connus). **Justifier**

- au moins 50% des élèves de la classe Rouge ont une note comprise entre 10 et 12
- au moins 50% des élèves de la classe Rouge ont une note inférieure ou égale à 10
- au moins 75% des élèves de la classe Rouge ont une note inférieure ou égale à 12
- environ 12.5% des élèves de la classe Rouge ont une note supérieure ou égale à 14
- les notes de la classes Jaune sont plus dispersée que celle de la classe Rouge
- le pourcentage d'élèves ayant une note inférieure ou égale à 11 est plus petit dans la classe Rouge que dans la classe Jaune

Exercice 5 : Dans une région de l'est de la France, la pollution atmosphérique est contrôlée quotidiennement, heure par heure, par un réseau de 21 stations de mesures. Parmi celles-ci, nous considérons la station notée U, qui est située en zone urbaine et la station I, en zone industrialisée et la station R, en zone rurale de moyenne montagne.

- Pour cette partie, on compare les mesures obtenues aux stations U et I pour le dioxyde de soufre, un des polluants mesurés. Les concentrations de ce polluants sont exprimées en millionnièmes de gramme par mètre cube d'air (noté $\mu\text{g}/\text{m}^3$). Les diagrammes en boîtes dessinés ci-après concernent les mesures horaires du polluants aux stations U et I, pour la journée du 16 novembre 2006. Les extrémités du diagrammes correspondent aux valeurs minimales et maximales.



- Pour chacune des deux stations, indiquer la médiane et calculer l'écart interquartile, ainsi que l'étendue de la série de mesures.
 - Indiquer par lecture graphique et en précisant les paramètres statistiques utilisés, sur quelle(s) station(s) ce jour là :
 - ↪ la dispersion des mesures a été la plus importante ?
 - ↪ la moitié des mesures au moins ont été inférieures ou égales à 5 ?
 - ↪ 75% des mesures au moins ont été inférieures ou égales à 6 ?
- Dans cette partie, on considère la station R.
Le tableau ci-dessous donne les relevés horaires, pour la même journée du 16 novembre 2006 en ce qui concerne le polluant ozone. Les concentrations sont exprimées en millionnièmes de gramme par mètre cube d'air.

Heure	1	2	3	4	5	6	7	8	9	10	11	12
Concentration	78	79	77	59	57	65	65	67	68	67	59	54
Heure	13	14	15	16	17	18	19	20	21	22	23	24
Concentration	64	68	72	74	72	72	76	77	76	74	77	76

3. Quelles sont les valeurs minimale et maximale de cette série ?
4. Déterminer la médiane et les quartiles de cette série.
5. Calculer la moyenne de cette série.
6. Construite le diagramme en boîte de cette série.