

Statistiques descriptives

Corrigé des exercices - 2de

Corrigé détaillé des exercices du chapitre. Vérifie tes réponses étape par étape et corrige-toi.

1 Notes de Malik ★★★

- Somme : $11 + 14 + 9 + 16 + 12 + 14 + 8 + 15 + 13 = 112$; il y a 9 notes.
Moyenne = $112 \div 9 \approx \mathbf{12,44}$.
- Notes rangées : 8 ; 9 ; 11 ; 12 ; 13 ; 14 ; 14 ; 15 ; 16. $n = 9$ est impair : la médiane est la 5^e valeur, **Me = 13**.
- Étendue = $16 - 8 = \mathbf{8}$.

2 Températures de janvier ★★★

Somme : $-2 + 3 + 0 + 5 + 1 + 4 = 11$. Moyenne = $11 \div 6 \approx \mathbf{1,83\text{ °C}}$.

Rangées : $-2 ; 0 ; 1 ; 3 ; 4 ; 5$. $n = 6$ est pair : médiane = $(1 + 3) \div 2 = \mathbf{2\text{ °C}}$.

Étendue = $5 - (-2) = \mathbf{7\text{ °C}}$.

3 Frères et sœurs ★★★

- $6 + 10 + 7 + 3 + 2 = 28$.
- Moyenne = $(6 \times 0 + 10 \times 1 + 7 \times 2 + 3 \times 3 + 2 \times 4) \div 28 = 41 \div 28 \approx \mathbf{1,46}$.
- $n = 28$ est pair : on cherche les valeurs de rang 14 et 15. Effectifs cumulés : 6 ; 16 ; 23 ; 26 ; 28. Les 14^e et 15^e valeurs valent 1 et 1 (le cumul dépasse 14 pour la valeur 1 et 15 pour la valeur 1). **Me = 1**.

4 Pièces des appartements ★★★

- Fréquences : 1 : $4/50 = 8\%$; 2 : $11/50 = 22\%$; 3 : $20/50 = 40\%$; 4 : $9/50 = 18\%$; 5 : $6/50 = 12\%$.

Pièces	1	2	3	4	5
Effectif cumulé	4	15	35	44	50
Fréquence cumulée	8 %	30 %	70 %	88 %	100 %

- On lit la colonne « 3 » : **70 %** des logements ont au plus 3 pièces (35 logements sur 50).

5 Quartiles d'une série brute ★★★

On range : 17 ; 18 ; 19 ; 21 ; 22 ; 23 ; 25 ; 26 ; 27 ; 28 ; 30 ; 31. $n = 12$.

- $12 \div 4 = 3$: $Q_1 = 19$ (3^e valeur).
- $3 \times 12 \div 4 = 9$: $Q_3 = 27$ (9^e valeur).
- n pair : $Me = (23 + 25) \div 2 = 24$.

Écart interquartile : $27 - 19 = 8$. Étendue : $31 - 17 = 14$.

6 Vrai ou faux ? ★★★

- Faux.** Contre-exemple : 1 ; 2 ; 9 a pour médiane 2 et pour moyenne 4.
- Vrai.** C'est la définition de la médiane (au moins 50 % des valeurs sont $\leq Me$, et au moins 50 % sont $\geq Me$).
- Faux.** σ est une racine carrée de la variance, qui est une moyenne de carrés : il est positif ou nul.
- Faux.** Avec $y = x + 5$, on a $a = 1$: l'écart-type est $|1| \cdot \sigma = \sigma$, il ne change pas. Translater les valeurs ne change pas leur dispersion.
- Vrai.** Au moins 75 % des valeurs sont $\leq Q_3$ et au moins 75 % sont $\geq Q_1$: l'intervalle $[Q_1 ; Q_3]$ en contient au moins 50 %.

7 Lire un diagramme en boîte ★★★

- Minimum 4 ; $Q_1 = 9$; $Me = 12$; $Q_3 = 16$; maximum 25.
- Étendue : $25 - 4 = 21$. Écart interquartile : $16 - 9 = 7$.
- La boîte contient **au moins 50 %** des cyclistes.
- Non** : le diagramme ne donne que cinq valeurs de position, pas la somme des valeurs. On ne peut pas en déduire la moyenne.

8 Écart-type d'une petite série ★★★

Moyenne : $30 \div 5 = 6$.

Écarts : -4 ; -1 ; 0 ; 2 ; 3 ; carrés : 16 ; 1 ; 0 ; 4 ; 9 ; somme 30.

$V = 30 \div 5 = 6$. $\sigma = \sqrt{6} \approx 2,45$.

9 Variance avec des effectifs ★★★

Moyenne = $(2 \times 8 + 4 \times 10 + 8 \times 12 + 4 \times 14 + 2 \times 16) \div 20 = 240 \div 20 = 12$.

$V = [2 \times (8 - 12)^2 + 4 \times (10 - 12)^2 + 8 \times (12 - 12)^2 + 4 \times (14 - 12)^2 + 2 \times (16 - 12)^2] \div 20 = (32 + 16 + 0 + 16 + 32) \div 20 = 96 \div 20 = 4,8$.

$$\sigma = \sqrt{4,8} \approx \mathbf{2,19}.$$

10 Deux façons de calculer la variance ★★★

- Moyenne 4. Carrés des écarts : 9 ; 1 ; 1 ; 1 ; 16 ; somme 28 ; $V = 28 \div 5 = \mathbf{5,6}$.
- Carrés des valeurs : 1 ; 9 ; 9 ; 25 ; 64, de moyenne $108 \div 5 = 21,6$. Puis $21,6 - 4^2 = 21,6 - 16 = \mathbf{5,6}$.
- On trouve le même résultat, c'est la formule de König. Elle évite de calculer tous les écarts quand la moyenne n'est pas ronde.

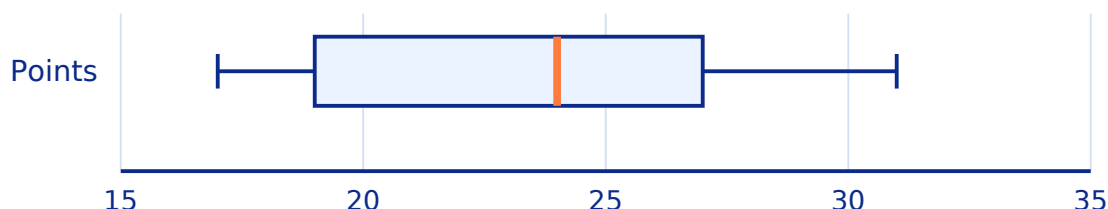
11 Quartiles avec un tableau ★★★

Points	0	1	2	3	4	5
Effectif cumulé	3	10	22	31	37	40

- $n = 40$. $40 \div 4 = 10$: Q_1 est la 10^e valeur ; le cumul vaut 10 pour la valeur 1 : $Q_1 = \mathbf{1}$. Les valeurs de rang 20 et 21 sont dans la classe de cumul 22 : toutes deux valent 2, $Me = \mathbf{2}$. $3 \times 40 \div 4 = 30$: la 30^e valeur est 3 (cumul 22 puis 31), $Q_3 = \mathbf{3}$.
- Écart interquartile = $3 - 1 = \mathbf{2}$. Moyenne = $97 \div 40 = \mathbf{2,425}$.

12 Tracer un diagramme en boîte ★★★

- (17 ; 19 ; 24 ; 27 ; 31).
- Voici le diagramme :



- Au moins **50 %** des matchs (au moins la moitié).

13 Une valeur extrême ★★★

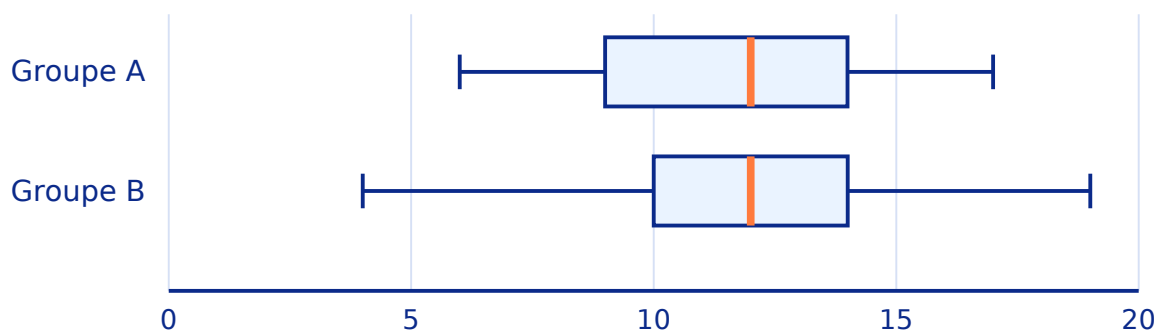
- $n = 7$: $Me = \mathbf{2,1}$; $Q_1 : 7 \div 4 = 1,75$ donc 2^e valeur, $\mathbf{1,9}$; $Q_3 : 5,25$ donc 6^e valeur, $\mathbf{2,4}$. Moyenne = $14,9 \div 7 \approx \mathbf{2,13}$. $\sigma \approx \mathbf{0,24}$.
- $n = 8$: $Me = (2,1 + 2,2) \div 2 = \mathbf{2,15}$; Q_1 (2^e valeur) = $\mathbf{1,9}$; Q_3 (6^e valeur) = $\mathbf{2,4}$. Moyenne = $26,9 \div 8 \approx \mathbf{3,36}$. $\sigma \approx \mathbf{3,27}$.
- La moyenne passe de 2,13 à 3,36 et l'écart-type de 0,24 à 3,27 : ils sont très sensibles à la valeur extrême. La médiane (2,1 → 2,15) et les quartiles bougent à peine. Pour décrire le salaire « typique », la médiane est plus fidèle.

14 Degrés Celsius et Fahrenheit ★★★

- a. Moyenne 18 ; écarts -6 ; -3 ; 0 ; 3 ; 6 ; $V = 90 \div 5 = 18$; $\sigma = \sqrt{18} \approx \mathbf{4,24\text{ }^{\circ}\text{C}}$.
- b. Avec $a = 1,8$ et $b = 32$: moyenne = $1,8 \times 18 + 32 = \mathbf{64,4\text{ }^{\circ}\text{F}}$; écart-type = $1,8 \times 4,24 \approx \mathbf{7,64\text{ }^{\circ}\text{F}}$ (le +32 ne joue pas).
- c. Valeurs en $^{\circ}\text{F}$: 53,6 ; 59 ; 64,4 ; 69,8 ; 75,2 ; moyenne 64,4 ; écart-type $\approx 7,64$: on retrouve les mêmes résultats.

15 Deux classes ★★★

- a. Groupe A : moyenne $\approx 11,55$; Me = 12 ; $Q_1 = 9$; $Q_3 = 14$; $\sigma \approx 3,06$.
Groupe B : moyenne $\approx 11,91$; Me = 12 ; $Q_1 = 10$; $Q_3 = 14$; $\sigma \approx 3,87$. (Ici $n = 11$: Q_1 est la 3^e valeur et Q_3 la 9^e.)



- b.
- c. Les positions sont très proches (médianes 12 et 12) : les deux groupes ont un niveau comparable. En revanche le groupe B est plus dispersé (écart-type 3,87 contre 3,06, étendue 15 contre 11) : il est plus hétérogène.

16 Moyenne de deux groupes ★★★

- a. Somme des notes du premier groupe : $24 \times 11,5 = 276$; du second : $16 \times 13,25 = 212$. Moyenne = $(276 + 212) \div 40 = 488 \div 40 = \mathbf{12,2}$.
- b. **Non** : la médiane dépend de la répartition des notes, pas seulement des moyennes des groupes. La moyenne est « additive », pas la médiane.

17 Données en classes ★★★

- a. Centres : 5 ; 15 ; 25 ; 35 ; 50. Moyenne estimée = $(8 \times 5 + 22 \times 15 + 35 \times 25 + 25 \times 35 + 10 \times 50) \div 100 = 2620 \div 100 = \mathbf{26,2\text{ min}}$ (valeur estimée).
- b. Effectifs cumulés : 8 ; 30 ; 65 ; 90 ; 100. Q_1 est la 25^e valeur : classe **[10 ; 20[** (cumul 8 puis 30). La médiane (50^e et 51^e valeurs) est dans **[20 ; 30[** (30 puis 65). Q_3 , 75^e valeur, est dans **[30 ; 40[** (65 puis 90).

18 Quel salaire annoncer ? ★★★

- a. Somme = $6 \times 1600 + 7 \times 1800 + 4 \times 2000 + 2 \times 2500 + 1 \times 6000 = 41200$;

moyenne = $41200 \div 20 = 2060 \text{ €}$. Cumuls : 6 ; 13 ; 17 ; 19 ; 20. Les valeurs de rang 10 et 11 valent 1800 et 1800 : **Me = 1800 €**. Q_1 (5^e valeur) = **1600 €** ; Q_3 (15^e valeur) = **2000 €**. $V = 883400$ (en €²) donc $\sigma \approx 940 \text{ €}$.

- b. Le chiffre est exact mais trompeur : 17 salariés sur 20 gagnent moins que la moyenne, car un seul salaire de 6 000 € la tire vers le haut. Le syndicat a raison sur le fond : la médiane (1800 €) décrit mieux la situation.
- c. La **médiane** et les **quartiles** (donc l'écart interquartile 400 €) : ils ne sont pas déformés par le salaire extrême.

19 Valeur manquante ★★★

- a. Moyenne = $9 + 12 + x + 14 + 10 + 13/6 = 58 + x/6 = 12$, donc $58 + x = 72$ et **$x = 14$** .
- b. Série rangée : 9 ; 10 ; 12 ; 13 ; 14 ; 14. $n = 6$: $Me = (12 + 13) \div 2 = 12,5$; $6 \div 4 = 1,5$ donc Q_1 est la 2^e valeur : **10** ; $3 \times 6 \div 4 = 4,5$ donc Q_3 est la 5^e valeur : **14**. Écarts à 12 : -3 ; -2 ; 0 ; 1 ; 2 ; 2 ; carrés : 9 ; 4 ; 0 ; 1 ; 4 ; 4 ; $V = 22 \div 6 \approx 3,67$; $\sigma \approx 1,91$.

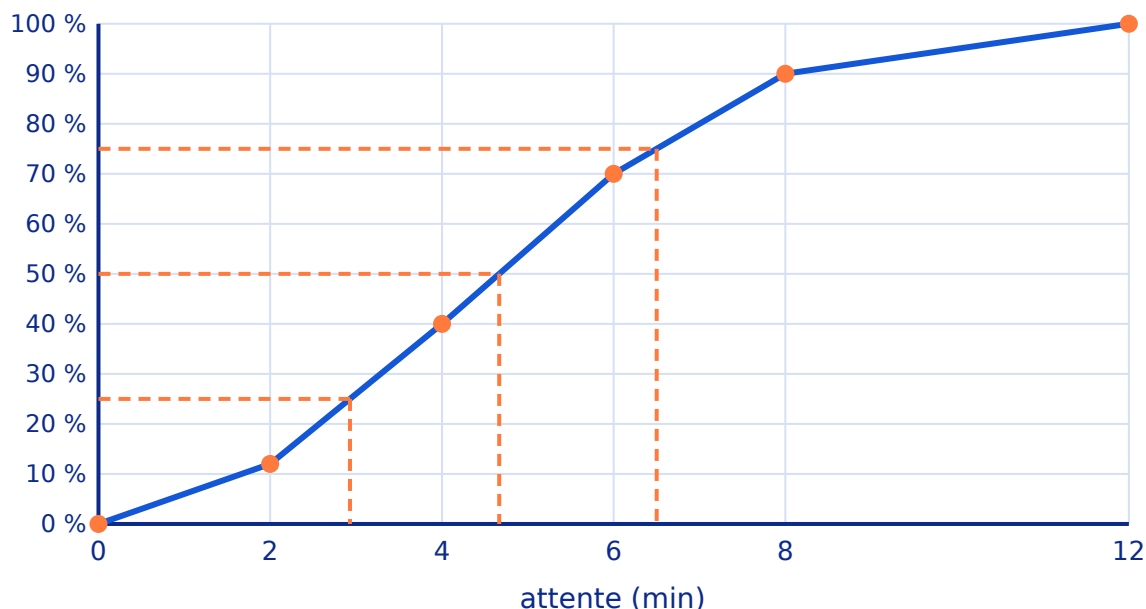
20 Ajouter une valeur ★★★

- a. Moyenne = $50 \div 5 = 10$. Écarts : -4 ; -2 ; 0 ; 2 ; 4 ; carrés : 16 ; 4 ; 0 ; 4 ; 16 ; $V = 40 \div 5 = 8$.
- b. Moyenne = $(50 + 10) \div 6 = 10$: inchangée, car on ajoute une valeur égale à la moyenne. L'écart de la nouvelle valeur est 0 : la somme des carrés reste 40 mais elle est divisée par 6 : $V = 40 \div 6 \approx 6,67$: la variance diminue.
- c. Moyenne = $(50 + 20) \div 6 = 70 \div 6 \approx 11,67$. En calculant $V = (\text{moyenne des carrés}) - (\text{moyenne})^2 = 156,67 - 136,11 \approx 20,56$, donc $\sigma \approx 4,53$: la valeur éloignée augmente fortement la dispersion.

21 Polygone des fréquences cumulées ★★★

Les points ont pour ordonnées 12 ; 40 ; 70 ; 90 ; 100 % (effectifs 6 ; 14 ; 15 ; 10 ; 5, cumulés 6 ; 20 ; 35 ; 45 ; 50).

- a. Entre 4 min (40 %) et 6 min (70 %), la courbe est un segment. Pour 50 % : $4 + 2 \times (10 \div 30) \approx 4,67 \text{ min}$. Graphiquement : un peu moins de 5 min.
- b. Q_1 (25 %) : entre 2 min (12 %) et 4 min (40 %) : $2 + 2 \times (13 \div 28) \approx 2,93 \text{ min}$. Q_3 (75 %) : entre 6 min (70 %) et 8 min (90 %) : $6 + 2 \times (5 \div 20) = 6,5 \text{ min}$.
- c. Pour 5 min : entre 40 % et 70 %, au milieu : $40 + 30 \div 2 = 55 \text{ %}$ environ.



22 Deux lanceurs de javelot ★★★

- $A : 378 \div 6 = 63$; $B : 378 \div 6 = 63$. Même moyenne : **63 m**.
- A : carrés des écarts 1 ; 1 ; 0 ; 4 ; 4 ; 0, $V = 10 \div 6 \approx 1,67$, $\sigma_A \approx \mathbf{1,29 \text{ m}}$. B : carrés 64 ; 49 ; 1 ; 25 ; 9 ; 16, $V = 164 \div 6 \approx 27,33$, $\sigma_B \approx \mathbf{5,23 \text{ m}}$.
- À moyennes égales, A est beaucoup plus régulier ($\sigma_A \approx 1,29$ contre 5,23). Tous ses lancers dépassent 60 m, alors que B a deux lancers sous 60 m. On choisit **A**.

23 Démonstrations ★★★

- $\Sigma(x_i - \bar{x}) = \Sigma x_i - n \cdot \bar{x} = n \cdot \bar{x} - n \cdot \bar{x} = \mathbf{0}$, car $\Sigma x_i = n \cdot \bar{x}$.
- $\Sigma y_i = a \Sigma x_i + n \cdot b$ donc, en divisant par n : $\bar{y} = a \cdot \bar{x} + b$.
- $y_i - \bar{y} = a \cdot x_i + b - a \cdot \bar{x} - b = a(x_i - \bar{x})$. Donc $(y_i - \bar{y})^2 = a^2(x_i - \bar{x})^2$. En faisant la moyenne : $V(y) = a^2 \cdot V(x)$. Enfin $\sigma(y) = \sqrt{a^2 \cdot V(x)} = |a| \cdot \sigma$.