Plaquette 4 – Diagrammes, boîtes à moustaches et lecture critique

Statistiques descriptives — Seconde

Corrections détaillées

Boîte à outils

Résumé en cinq nombres et boîte à moustaches.

$$\text{min} \quad Q_1 \quad \text{Méd} \quad Q_3 \quad \text{max}$$

— la **boîte** va de Q_1 à Q_3 : elle contient la **moitié centrale** des données ;
— le trait intérieur marque la **médiane** ;
— les **moustaches** vont jusqu'au minimum et au maximum.

Zone du diagramme	Proportion des données
de min à Q_1	environ 25 %
de Q_1 à Méd	environ 25 %
de Méd à Q_3	environ 25 %
de Q_3 à max	environ 25 %

Repérage des valeurs aberrantes. On considère comme suspecte toute valeur hors de

$$[\,Q_1 - 1{,}5 \times (Q_3 - Q_1)\ ;\ Q_3 + 1{,}5 \times (Q_3 - Q_1)\,]$$

Diagramme circulaire. L'angle est proportionnel à l'effectif :

$$\text{angle}_i = \frac{n_i}{N} \times 360° \qquad \text{et réciproquement} \qquad n_i = \frac{\text{angle}_i}{360°} \times N$$

Sur un diagramme **semi-circulaire**, on remplace 360° par 180°.

Trois pièges de lecture à connaître.

— **Axe tronqué** : un axe vertical qui ne part pas de 0 exagère visuellement les écarts.
— **Effectif contre fréquence** : deux groupes de tailles différentes ne se comparent qu'en **fréquences**.
— **Aire contre hauteur** : dans un histogramme à classes inégales, c'est l'**aire** qui porte l'effectif.

Correction de l'exercice 1 – Construire une boîte à moustaches

La série est déjà ordonnée et compte $N = 12$ valeurs. L'effectif étant pair, la médiane est la demi-somme des 6^{e} et 7^{e} valeurs :

$$\text{Méd} = \frac{11 + 12}{2} = 11{,}5 \text{ s}$$

Q_1 est la plus petite valeur ayant au moins 25 % des valeurs avant elle ou à son rang : rang $\frac{12}{4} = 3$, soit 6 s. Q_3 est au rang $\frac{3 \times 12}{4} = 9$, soit 14 s. Le résumé est donc 3 ; 6 ; 11,5 ; 14 ; 25.

La boîte va de 6 à 14 s. La moustache gauche mesure 3 s, la droite 11 s : la série est étalée vers les grandes valeurs, à cause notamment du temps de 25 s.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

Réponse. 3 ; 6 ; 11,5 ; 14 ; 25 — boîte de 6 à 14 s, moustache droite bien plus longue.

Correction de l'exercice 2 – Lire une boîte à moustaches

Les quartiles partagent l'effectif en quatre parts d'environ 25 % chacune. Au-dessus de $Q_3 = 2\,400$ € se trouve environ un quart des salariés :

$$0{,}25 \times 80 = 20 \text{ salariés}$$

Entre $Q_1 = 1\,650$ € et $Q_3 = 2\,400$ € se trouve la moitié des salariés, soit environ 40.

On dit « environ », car les quartiles ne sont exacts qu'à un rang près. La boîte ne mesure que 750 € alors que l'étendue atteint 3 900 € : la moitié des salaires tient dans moins d'un cinquième de l'étendue, et la moustache droite, de 2 800 €, révèle quelques salaires très élevés. La moyenne, elle, n'est pas lisible sur une boîte à moustaches.

Réponse. Environ 20 salariés au-dessus de 2400 €, et 40 entre 1650 et 2400 €.

Correction de l'exercice 3 – Comparer deux boîtes

Les deux classes ont la même médiane, 11 : la moitié des élèves a moins de 11 dans chacune. On compare la dispersion par l'écart interquartile, largeur de la boîte :

$$Q_3 - Q_1 = 13 - 9 = 4 \text{ pour A} \qquad 16 - 7 = 9 \text{ pour B}$$

Les étendues vont dans le même sens : 14 pour A, 18 pour B.

La classe B est donc nettement plus hétérogène : la moitié centrale de ses notes s'étale sur 9 points, plus de deux fois plus que dans A. A est un groupe homogène de niveau moyen ; B mélange des élèves très solides (jusqu'à 20) et un quart d'élèves sous 7, pour qui une remédiation s'impose.

Réponse. Même médiane 11, mais écart interquartile 4 pour A contre 9 pour B : B est bien plus hétérogène.

Correction de l'exercice 4 – Détecter une valeur aberrante

On applique la règle de l'énoncé. L'écart interquartile vaut $Q_3 - Q_1 = 30 - 20 = 10$, et 1,5 fois cet écart vaut 15. Les seuils sont donc

$$Q_1 - 15 = 5 \qquad \text{et} \qquad Q_3 + 15 = 45$$

Une valeur hors de $[5\,;45]$ est suspecte. La valeur 6 est dans l'intervalle, de justesse ; la valeur 52 dépasse 45 : elle est **suspecte**.

Une valeur suspecte n'est pas supprimée d'office : on vérifie d'abord s'il s'agit d'une erreur de saisie ou d'une mesure réelle. Sur une boîte à moustaches, on la dessine en point isolé, la moustache s'arrêtant à la dernière valeur non suspecte.

Réponse. Intervalle $[5\,;45]$: seule la valeur 52 est suspecte.

Correction de l'exercice 5 – Du diagramme en bâtons aux indicateurs

L'effectif total vaut $9 + 18 + 14 + 6 + 3 = 50$. Le **mode** est la valeur la plus fréquente : 1 frère ou sœur, avec 18 élèves (le mode est une valeur, pas un effectif).

Pour la médiane, l'effectif est pair : on cherche les valeurs de rangs 25 et 26. Les effectifs cumulés valent 9 ; 27 ; 41 ; 47 ; 50, donc les rangs 10 à 27 correspondent à la valeur 1 : Méd = 1.

$$\overline{x} = \frac{0 \times 9 + 1 \times 18 + 2 \times 14 + 3 \times 6 + 4 \times 3}{50} = \frac{76}{50} = 1{,}52$$

La moyenne dépasse la médiane : les familles nombreuses la tirent vers le haut.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s'expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

Réponse. Mode 1, médiane 1, moyenne $\frac{76}{50} = 1{,}52$.

Correction de l’exercice 6 – Construire un diagramme circulaire

Dans un diagramme circulaire, les angles sont proportionnels aux effectifs, et le tour complet, 360°, représente les 40 élèves : un élève correspond à $\frac{360}{40} = 9°$.

Transport	Effectif	Angle
à pied	12	108°
bus	18	162°
vélo	6	54°
voiture	4	36°

Contrôle indispensable : $108 + 162 + 54 + 36 = 360°$. Les fréquences correspondantes sont 30 %, 45 %, 15 % et 10 % ; on peut aussi utiliser $1\,\% \leftrightarrow 3{,}6°$. Le bus occupe presque un demi-disque.

$$\text{angle} = \frac{\text{effectif}}{\text{effectif total}} \times 360°$$

Réponse. 108° ; 162° ; 54° ; 36° — de somme 360°.

Correction de l’exercice 7 – Remonter d’un angle à un effectif

On inverse la proportionnalité : l’effectif d’un secteur est la fraction du tour qu’il occupe, multipliée par l’effectif total.

$$n = \frac{81}{360} \times 200 = 0{,}225 \times 200 = 45 \text{ personnes}$$

Contrôle : $\frac{45}{200} \times 360 = 81°$.

Sur un diagramme semi-circulaire, les 200 personnes se partagent 180°, soit 0,9° par personne : ce groupe occuperait $45 \times 0{,}9 = 40{,}5°$, la moitié de 81°. Lire un demi-disque avec la formule du disque entier diviserait l’effectif par deux.

Réponse. 45 personnes ; sur un semi-disque, ce groupe occuperait 40,5°.

Correction de l’exercice 8 – Le piège de l’axe tronqué

La hausse réelle des ventes vaut $1\,040 - 980 = 60$ unités, soit un taux d’évolution

$$\frac{1\,040 - 980}{980} \approx 0{,}061 = 6{,}1\,\%$$

Mais l’axe ne commence pas à 0 : il va de 970 à 1 050. Sur le graphique, la première barre a une hauteur de $980 - 970 = 10$ unités, la dernière $1\,040 - 970 = 70$: la dernière paraît 7 fois plus haute, alors que le rapport réel des ventes est $\frac{1\,040}{980} \approx 1{,}06$.

Le lecteur a été trompé par l’**axe tronqué**. Avec un axe partant de 0, les trois barres seraient presque égales. La tendance est bien à la hausse, mais modérée : environ 3 % par mois.

Réponse. La hausse réelle n’est que de 6,1 % : l’axe tronqué de 970 à 1050 multiplie l’écart apparent par 7.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s’expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

Correction de l’exercice 9 – Effectifs ou fréquences ?

Les effectifs bruts favorisent B (80 nageurs contre 60), mais les deux lycées n’ont pas la même taille. Pour comparer, on calcule les **fréquences** :

$$f_A = \frac{60}{400} = 15\,\% \qquad f_B = \frac{80}{1\,000} = 8\,\%$$

Proportionnellement, la natation est presque deux fois plus pratiquée dans A. Avec la fréquence de A, le lycée B compterait $0{,}15 \times 1\,000 = 150$ nageurs.

Le titre confond effectif et proportion. Un titre honnête serait : « B a le plus grand groupe de nageurs, A la plus forte participation ». Deux populations de tailles différentes se comparent en fréquences.

Réponse. 15 % dans A contre 8 % dans B : c’est A qui domine en proportion.

Correction de l’exercice 10 – Des cumuls à la boîte

Les quartiles et la médiane se lisent là où la fréquence cumulée atteint 25 %, 50 % et 75 % : $Q_1 = 15$, Méd $= 19$ et $Q_3 = 24$. Le maximum est atteint à 100 % : 35.

Le minimum n’est pas donné exactement : on sait seulement que 10 % des mesures sont inférieures ou égales à 12. En prenant 12 comme valeur basse connue, le résumé est

$$12\ ;\ 15\ ;\ 19\ ;\ 24\ ;\ 35$$

La boîte a pour largeur $24 - 15 = 9$. Chaque quart contient environ 15 mesures. La moustache droite (11) est bien plus longue que la gauche (3 au plus connu) : la série est étalée vers le haut.

Réponse. 12 ; 15 ; 19 ; 24 ; 35 — boîte de largeur 9.

Correction de l’exercice 11 – Un résumé ne détermine pas la série

Avec 9 valeurs rangées, la médiane est la 5^{e} valeur, Q_1 la 3^{e} (rang $\frac{9}{4} = 2{,}25$, arrondi au rang supérieur) et Q_3 la 7^{e} (rang 6,75). Le résumé impose donc le squelette

$$2\ ;\ ?\ ;\ 5\ ;\ ?\ ;\ 10\ ;\ ?\ ;\ 14\ ;\ ?\ ;\ 20$$

où chaque « ? » est un entier compris entre ses voisins. Par exemple 2 ; 3 ; 5 ; 7 ; 10 ; 12 ; 14 ; 16 ; 20, et 2 ; 4 ; 5 ; 9 ; 10 ; 13 ; 14 ; 18 ; 20.

Ces deux séries ont le même résumé, mais des moyennes différentes : $\frac{89}{9} \approx 9{,}89$ et $\frac{95}{9} \approx 10{,}56$. Une boîte à moustaches résume une série, elle ne la détermine pas.

Réponse. Par exemple $2; 3; 5; 7; 10; 12; 14; 16; 20$ et $2; 4; 5; 9; 10; 13; 14; 18; 20$ — même résumé, moyennes 9,89 et 10,56.

Correction de l’exercice 12 – Quartiles d’une série de 20 valeurs

L’effectif vaut $N = 20$. La médiane est la demi-somme des 10^{e} et 11^{e} valeurs, $\frac{11 + 11}{2} = 11$. Q_1 est la valeur de rang $\frac{20}{4} = 5$, soit 8, et Q_3 celle de rang 15, soit 14.

$$4\ ;\ 8\ ;\ 11\ ;\ 14\ ;\ 19$$

L’écart interquartile vaut 6 et l’étendue 15. La somme des notes, 223, donne une moyenne de 11,15, très proche de la médiane. La médiane est au centre de la boîte ($11 - 8 = 3 = 14 - 11$) et les moustaches ont des longueurs voisines (4 et 5) : la série est presque symétrique.

Réponse. 4 ; 8 ; 11 ; 14 ; 19 (moyenne 11,15).

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s’expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

Correction de l’exercice 13 – Barres empilées

On calcule d’abord les totaux annuels, 80, 90 et 100 inscrits, puis la part de chaque activité dans le total.

Année	Sport	Musique	Théâtre
2024	50 %	37,5 %	12,5 %
2025	50 %	31,1 %	18,9 %
2026	52 %	24 %	24 %

En effectif, de 2024 à 2026, le sport progresse de $\frac{52-40}{40} = 30\,\%$, la musique recule de 20 % et le théâtre bondit de 140 %. En part du total, le sport est presque stable : sa hausse suit celle du total (+25 %). La musique perd beaucoup de terrain et le théâtre double sa part.

Un diagramme empilé lit mal les segments du milieu, dont les deux bords bougent : des barres groupées seraient plus lisibles.

Réponse. Sport +30 %, musique −20 %, théâtre +140 % ; en part du total : 50 → 52 %, 37,5 → 24 %, 12,5 → 24 %.

Correction de l’exercice 14 – Trois machines à comparer

On compare d’abord le centrage par la médiane : 100 pour M_1 et M_2, 101 pour M_3, décalée. Puis la dispersion par l’écart interquartile : 1 pour M_1, 0,4 pour M_2 et 3 pour M_3.

— M_2 : tout son résumé, de 99 à 101, tient dans la tolérance : toutes les tiges sont conformes.
— M_1 : sa boîte [99,5 ; 100,5] est dans la tolérance, mais ses moustaches en sortent : une partie des tiges extrêmes est rebutée.
— M_3 : sa médiane est déjà à la limite 101, donc au moins la moitié des tiges mesure 101 mm ou plus.

Classement : M_2, puis M_1, puis M_3. Même recentrée de −1 mm, M_3 garderait une boîte de largeur 3, plus large que la tolérance [99 ; 101] : elle doit être révisée, pas seulement réglée.

Réponse. M_2 (tout conforme), puis M_1 (quelques rebuts), puis M_3 (décalée **et** dispersée).

Correction de l’exercice 15 – Synthèse : enquête sur les temps d’écran

a) Avec $N = 24$ valeurs rangées, la médiane est la demi-somme des 12^{e} et 13^{e} : $\frac{95+100}{2} = 97{,}5$ min. Q_1 est la 6^{e} valeur, 70 min, et Q_3 la 18^{e}, 130 min. Résumé : 30 ; 70 ; 97,5 ; 130 ; 300.

b) La somme vaut 2 660 min, donc $\overline{x} = \frac{2\,660}{24} \approx 110{,}8$ min, nettement au-dessus de la médiane : quelques très gros temps d’écran tirent la moyenne vers le haut.

c) L’écart interquartile vaut 60 ; avec la règle des 1,5 écart interquartile, le seuil haut est

$$Q_3 + 1{,}5 \times 60 = 130 + 90 = 220$$

Seule la valeur 300 le dépasse. Sans elle, la moyenne tomberait à $\frac{2\,360}{23} \approx 102{,}6$ min, la médiane à 95 min.

d) On retient la **médiane** : « la moitié des élèves passe plus d’une heure et demie par jour devant un écran ». 15 élèves sur 24 sont sous la moyenne, qui ne représente pas

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s’expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).

l’élève typique.

Réponse. a) 30 ; 70 ; 97,5 ; 130 ; 300 · b) $\overline{x} \approx 110{,}83 >$ Méd · c) seule la valeur 300 est aberrante · d) la médiane.

mathlvl.fr — Tous droits réservés. Toute reproduction, diffusion, revente ou exploitation commerciale, même partielle, est strictement interdite.
Tout contrevenant s’expose à des poursuites en vertu du Code de la propriété intellectuelle (art. L.335-2 et suivants).