Chapitre 8 – Statistiques

1. Indicateurs

1a. Moyenne et médiane

La **moyenne** d'une série statistique $x_1 ; x_2 ; \dots ; x_n$ est la **somme des valeurs, divisée par l'effectif** n.

$$\text{Moyenne } = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

La **médiane** d'une série statistique **rangée dans l'ordre croissant** est le « **nombre du milieu** », tel que 50% des valeurs de la série sont inférieures à la médiane.

- Si l'effectif n est pair, la médiane est la moyenne de la $\frac{n}{2}$ – ième valeur et de la $\frac{n}{2}+1$ – ième.
- Si l'effectif n est impair, la médiane est la $\frac{n+1}{2}$ – ième valeur.

Exemple 1 Un restaurant a obtenu les sept notes suivantes sur dix : 3; 10; 4; 9; 10; 4; 2.
Déterminer la note moyenne, et la note médiane.

Exemple 2 Un élève a obtenu une série de trois notes 6 ; 11 ; 13 en mathématiques.
a. Déterminer la moyenne et la médiane de cette série.
L'élève obtient deux nouvelles notes : 10 et 17 et obtient une nouvelle série de notes : 6 ; 10 ; 11 ; 13 ; 17.
b. Sans refaire de calcul, la moyenne et la médiane ont-elles changé ? Justifier.

Exemple 3 On donne une série de nombres correspondant à la taille, en cm de membres d'une équipe de basket : 196 ; 209; 198; 176; x. On sait de plus que la taille moyenne des joueurs est de 196 cm.
Écrire une égalité permettant de trouver le nombre x, puis déterminer ce nombre.

Exemple 4 Dans la capture d'écran de tableur ci-contre, on cherche à calculer la somme et la moyenne des valeurs.
a. Indiquer la formule à rentrer dans la cellule D1.
b. Indiquer la formule en D2. Donner deux possibilités.
c. Quelle est la médiane de cette série ?

	A	B	C	D
1	18		Somme des valeurs	
2	16		Moyenne des valeurs	
3	9			
4	12			
5	17			
6	14			

Exemple 1

On commence par la moyenne : $\dfrac{3+10+4+9+10+4+2}{7} = \dfrac{42}{7} = \mathbf{6}$

Pour la médiane, on range la série dans l'ordre croissant : 2 ; 3 ; 4 ; 4 ; 9 ; 10 ; 10
Ensuite, la médiane correspond à la note « du milieu », c'est-à-dire **4**.

Exemple 2

a. La moyenne est $\dfrac{6+11+13}{3} = \dfrac{30}{3} = \mathbf{10}$

La série est déjà rangée dans l'ordre, donc la médiane est la note « du milieu », c'est-à-dire **11**.
b. La « note du milieu » est restée la même, donc **la médiane n'a pas changé**.
En revanche, les deux nouvelles notes sont supérieures ou égales à la moyenne précédente, donc **la moyenne a augmenté**.

Exemple 3

On peut écrire l'égalité suivante : $\frac{196+209+198+176+x}{5} = 196$

C'est donc une équation, on multiplie les deux membres par 5 :

$196 + 209 + 198 + 176 + x = 196 \times 5$

$\Leftrightarrow 779 + x = 980$

$\Leftrightarrow x = 980 - 779 = \mathbf{201}$

Le joueur manquant a donc une taille de 201 cm.

Exemple 4

a. Pour effectuer une somme, on entre « **=SOMME(A1 : A6)** ».

b. On peut aussi calculer une moyenne avec « **=MOYENNE(A1 : A6)** »
ou remarquer qu'il s'agit de diviser la somme obtenue en question précédente par l'effectif. On peut donc écrire « **=D1/6** ».

c. On range la série dans l'ordre croissante : 9 ; 12 ; 14 ; 16 ; 17 ; 18
Ici, la série a un effectif pair, il y a donc deux « nombres du milieu »,
ce sont 14 et 16. La médiane est la moyenne de ces deux nombres, c'est-à-dire **15**.

1b. Quartiles

Dans une série statistique **d'effectif n**, rangée dans l'ordre croissant :

• **le 1er quartile**, noté Q_1 , est la plus petite valeur de la série telle qu'au moins 25% des valeurs lui soient inférieures ou égales. Il correspond à la valeur de **rang $\frac{1}{4}n$ arrondi à l'entier supérieur**.

• **le 3ème quartile**, noté Q_3, est la plus petite valeur de la série telle qu'au moins 75% des valeurs lui soient inférieures ou égales. Il correspond à la valeur de **rang $\frac{3}{4}n$ arrondi à l'entier supérieur**.

• l'écart interquartile est la différence $Q_3 - Q_1$.

• Il existe également le 1er décile (corrrespondant à 10% des valeurs) et le 9ème décile (90% des valeurs).

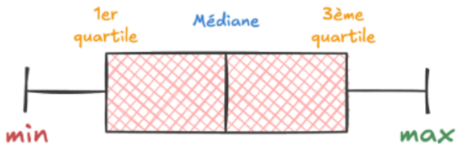
Exemple 1 En Physique, on a demandé à 13 groupes d'élèves de mesurer la résistance d'un conducteur.
Voici leurs résultats : 14,7 ; 18,4 ; 43,5 ; 43,7 ; 44,8 ; 44,8 ; 44,9 ; 45,1 ; 45,1 ; 45,2 ; 45,2 ; 46,3 ; 46,4
Déterminer la médiane et les quartiles.

Exemple 2 Les indicateurs statistiques d'une série de notes d'un contrôle noté sur 20 sont données :

Minimum	1er quartile	Médiane	3ème quartile	Maximum
2	5	10	12	17

a. Quelle proportion d'élèves a obtenu 12/20 ou moins ?
b. Quelle proportion d'élèves a obtenu 10/20 ou plus ?

Remarque Les indicateurs peuvent être résumés sous la forme d'une « boîte à moustaches ».



Exemple 1

La série est, fort heureusement, déjà rangée dans l'ordre croissant.

• On compte 13 valeurs, donc la valeur « du milieu » est la 7ème valeur.
La **médiane est 44,9**.

• *Pour déterminer à quelle valeur correspondent les quartiles, on divise 13 par 4 et on arrondit **toujours à l'entier supérieur**.*

$\frac{13}{4} = 3,25 \approx 4$ donc le 1er quartile correspond à la 4ème valeur, c'est-à-dire **43,7**.

• *Pour le troisième quartile, on calcule trois quarts de 13.*

$\frac{13}{4} \times 3 = 9,75 \approx 10$ donc le 3ème quartile est la 10ème valeur, c'est-à-dire **45,2**.

Exemple 2

a. Cela correspond au 3ème quartile, donc **au moins 75%** des élèves ont obtenu 12/20 ou moins.

b. Cela correspond à la médiane, donc **au moins 50%** des élèves ont obtenu 10/20 ou plus.

1c. Effectifs et coefficients

Quand **des effectifs ou des coefficients entiers** sont donnés, on peut faire **comme si la valeur était répétée**.

Exemple 1 Un élève a obtenu la note de 13/20 a un devoir coefficienté 2 et la note de 16/20 à un devoir coefficienté 1. Quelle est sa moyenne ?

Exemple 2 Voici les notes obtenus dans une classe lors d'un contrôle de mathématiques. Quelle est la note médiane, et quels sont les quartiles ?

Note	7	10	12	14
Nombre d'élèves	5	4	8	10

Exemple 3 Un élève a obtenu trois notes coefficient 1 en mathématiques : 10 ; 13 et 12.
Il reste encore un devoir coefficient 2 et l'élève aimerait avoir une moyenne de 15.
Peut-il y arriver, et si possible, combien doit-il obtenir ?

Exemple 1

On fait comme si le 13/20 avait été obtenu deux fois.

On calcule : $\frac{13+13+16}{3} = \frac{42}{3} = \mathbf{14}$.

Exemple 2

Il faut d'abord calculer l'effectif : $5 + 4 + 8 + 10 = 27$.

- La médiane correspond donc à la 14ème valeur, qui est un **12**.
- $\frac{27}{4} = 6{,}75 \approx 7$ donc le 1er quartile correspond à la 7ème valeur, c'est-à-dire **10**.
- $\frac{27}{4} \times 3 = 20{,}25 \approx 21$ donc le 3ème quartile est la 21ème valeur, c'est-à-dire **14**.

Exemple 3

Il s'agit de résoudre l'équation : $\frac{10+13+12+2x}{5} = 15$

C'est une équation, on multiplie les deux membres par 5 :

$10 + 13 + 12 + 2x = 15 \times 5$

$\Leftrightarrow 35 + 2x = 75$

$\Leftrightarrow 2x = 40$

$\Leftrightarrow x = \mathbf{20}$

Il faut donc que l'élève obtienne un 20/20.

2. Statistiques à deux variables

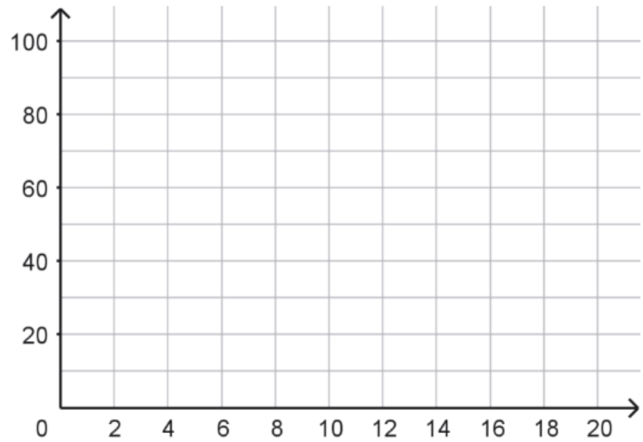
2a. Nuage de points

On considère deux séries statistiques $x_1 ; x_2 ; \ldots ; x_n$ et $y_1 ; y_2 ; \ldots ; y_n$ de même effectif. L'ensemble des points $(x_1 ; y_1)$; $(x_2 ; y_2)$; … ; $(x_n ; y_n)$ est le nuage de points associé.
Le point moyen est celui dont l'abscisse et l'ordonnée sont la moyenne des abscisses $x_1 ; x_2 ; \ldots ; x_n$ et des ordonnées $y_1 ; y_2 ; \ldots ; y_n$.

Exemple Le tableau suivant présente l'évolution du budget publicitaire et du chiffre d'affaires d'une société au cours des 6 dernières années :

Budget publicitaire (en milliers d'€)	8	10	12	14	16	18
Chiffre d'affaires (en milliers d'€)	40	55	55	70	75	95

Représenter le nuage de points correspondant, ainsi que le point moyen M.



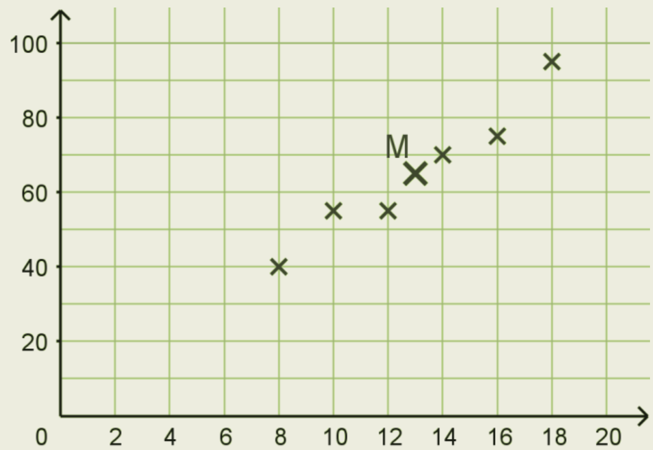
Pour le point moyen, on calcule la moyenne des abscisses :

$$\frac{8+10+12+14+16+18}{6} = \frac{78}{3} = \mathbf{13}$$

et la moyenne des ordonnées :

$$\frac{40+55+55+70+75+95}{6} = \frac{390}{6} = \mathbf{65}$$

Ainsi, $\boldsymbol{M}(\mathbf{13} ; \mathbf{65})$.

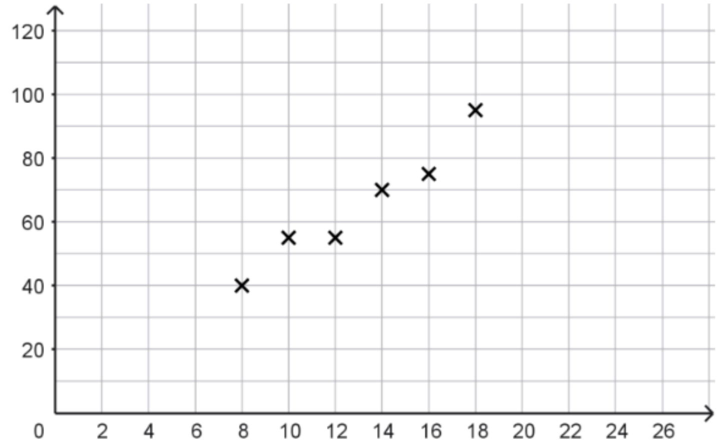


2b. Ajustement affine

Lorsque les points semblent alignés, on peut construire une droite d'ajustement affine, qui passe au plus près des points.
On peut s'en servir pour des interpolations ou extrapolations.

Exemple On reprend le nuage de points précédemment obtenu, qui correspond au chiffre d'affaires d'une société en fonction de son budget publicitaire, le tout en milliers d'euros.

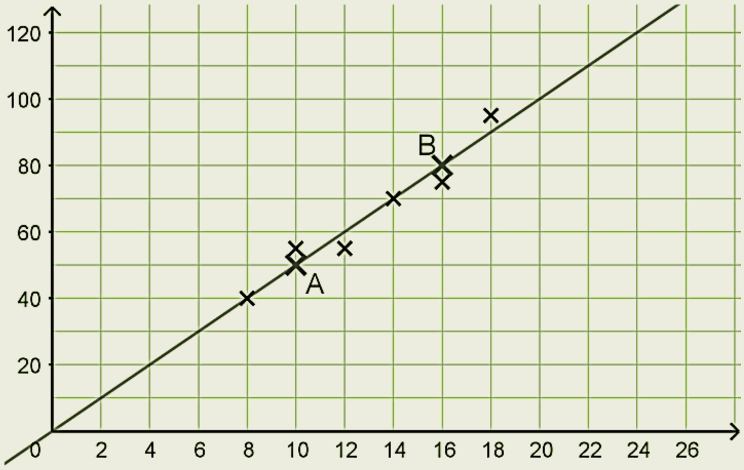
1. Placer les points $A(10;50)$ et $B(16;80)$ dans le repère, puis tracer la droite (AB). On choisit cette droite comme droite d'ajustement affine.
2. Vérifier que cette droite a pour équation $y = 5x$.



3. À l'aide de la droite, estimer le chiffre d'affaires correspondant à un budget publicitaire de 11 000€. *Il s'agit d'une **interpolation** (essayer de prévoir des valeurs situées entre les valeurs existantes).*

4. Toujours à l'aide de la droite, estimer le chiffre d'affaires correspondant à un budget publicitaire de 24 000€. *Il s'agit d'une **extrapolation** (prévoir des valeurs dans le prolongement des valeurs existantes).*

2. On vérifie bien que $50 = 5 \times 10$ et que $80 = 5 \times 16$, donc les deux points A et B appartiennent à la droite d'équation $y = 5x$.



3. L'énoncé nous demande de chercher y pour $x = 11$.
On calcule $y = 5 \times 11 = 55$,
donc la société peut s'attendre à un **chiffre d'affaires de 55 000€.**
On pouvait aussi lire cette réponse sur le graphique.

4. L'énoncé nous demande de chercher y pour $x = 24$.
On calcule $y = 5 \times 24 = 120$,
donc la société peut s'attendre à un **chiffre d'affaires de 120 000€.**