# Chapitre 15
# Statistiques descriptives

## Objectifs

- Connaître les indicateurs de tendance centrale d'une série statistique : moyenne pondérée.
- Connaître la propriété de linéarité de la moyenne.
- Connaître les indicateurs de dispersion : écart interquartile, écart type.
- Savoir décrire verbalement les différences entre deux séries statistiques, en s'appuyant sur des indicateurs ou sur des représentations graphiques données.

## Culture scientifique

**Gertrude Cox** (1900-1978) est une statisticienne américaine.
Elle a marqué le domaine de la statistique, de la science des enquêtes et l'avancement des méthodes de sondage. Elle a également promu leur utilisation dans des domaines variés comme l'agriculture et l'industrie. Première femme à obtenir un doctorat en statistiques aux États-Unis, elle a ouvert la voie aux femmes dans les sciences.

### Et sinon, dans la vraie vie ?

Le domaine d'application des statistiques est vaste et touche à tous les aspects de la vie quotidienne (politique, social, économique...). Par exemple, dans le secteur du pari sportif, on utilise les statistiques pour établir les côtes.

# Cours et applications directes

Compléments numériques
↘ Sur Sacado via votre ENT
À consulter dans "Livre numérique"
en indiquant le numéro de page : 204

## A Moyenne d'une série statistique

**↘ Définition**

La **moyenne pondérée**, notée $\overline{x}$, d'une série statistique ayant $k$ valeurs $x_1$ ; $x_2$ ; ... ; $x_k$ d'effectifs respectifs $n_1$ ; $n_2$ ; ... ; $n_k$ est égale à : $\overline{x} = \dfrac{n_1 \times x_1 + n_2 \times x_2 + \ldots + n_k \times x_k}{n_1 + n_2 + \ldots + n_k}$

Lorsque la série est donnée par classes (ou tranches), les $x_i$ sont les milieux des classes.

> Exercice résolu
>
> Un élève obtient les notes : $8$, $11$ et $13$ de coefficient respectif $1$, $2$ et $3$.
>
> Calculer sa moyenne $\overline{n}$. $\overline{n} = \dfrac{1 \times 8 + 2 \times 11 + 3 \times 13}{1 + 2 + 3} = \dfrac{69}{6} = 11{,}5$

## B Moyenne et fréquence

**Propriété**

La **moyenne pondérée** $\overline{x}$ d'une série statistique ayant $k$ valeurs $x_1$ ; $x_2$ ; ... ; $x_k$ de **fréquences** respectives $f_1$ ; $f_2$ ; ... ; $f_k$ est égale à : $\overline{x} = f_1 \times x_1 + f_2 \times x_2 + \ldots + f_k \times x_k$

> Exercice résolu
>
> $15\%$ des élèves ont $15$ ans, $63\%$ des élèves ont $16$ ans, $22\%$ des élèves ont $17$ ans.
>
> Calculer la moyenne d'âge $\overline{a}$ des élèves. $\overline{a} = 0{,}15 \times 15 + 0{,}63 \times 16 + 0{,}22 \times 17 = 16{,}07$ ans

## C Linéarité

**Propriété**

Lorsque toutes les valeurs d'une série de moyenne $\overline{x}$ sont transformées par une fonction affine $x \mapsto mx + p$, la moyenne de la nouvelle série est $m \times \overline{x} + p$.

> Exercice résolu
>
> Une machine fabrique des billes de $5\ cm$ de diamètre moyen.
> Pour des raisons de coût, le tourneur modifie les réglages en utilisant la fonction $f : x \mapsto 0{,}9\,x + 0{,}4$.
> Quel est le diamètre moyen $\overline{d}$ obtenu après les réglages ? $\overline{d} = 0{,}9 \times 5 + 0{,}4 = 4{,}9\ cm$

## D Médiane d'une série statistique

**↘ Définition**

Dans une série ordonnée, on appelle **médiane** un nombre qui partage cette série en deux séries de même effectif.
La médiane peut être une valeur qui n'appartient pas à la série.

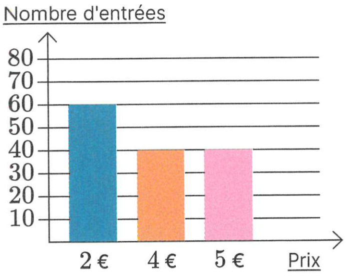
- Si l'effectif total $N$ est **impair**, la médiane est **la valeur centrale** de la série ordonnée.
- Si l'effectif total $N$ est **pair**, la médiane est **la valeur moyenne des deux valeurs centrales** de la série ordonnée.

> Exercice résolu
>
> Quelle est la médiane de la série de valeurs suivantes : $6$-$9$-$6$-$10$-$13$-$10$-$20$ ?
>
> Etape 1 : ordonner les valeurs dans l'ordre croissant : $6$-$6$-$9$-$10$-$10$-$13$-$20$.
> Etape 2 : compter l'effectif total $N = 7$.
>
> $7$ est impair donc la médiane est la $4^e$ valeur de la série ordonnée : $\underbrace{6 - 6 - 9}_{3\ valeurs} - 10 - \underbrace{10 - 13 - 20}_{3\ valeurs}$
> La médiane est donc $10$.
>
> Il y a autant de valeurs à gauche qu'à droite de la médiane dans la série ordonnée, soit $3$ valeurs.

**Compléments numériques**

↘ *Sur Sacado via votre ENT*
À consulter dans "Livre numérique"
en indiquant le numéro de page : 205

**1** La piscine municipale pratique trois tarifs : $2$ €, $4$ € et $5$ €.
On a représenté la répartition des entrées d'une journée.
Calculer le prix moyen $\overline{p}$ d'une entrée payée par les usagers.
Arrondir au centime près.




**2** Lors d'un concours, les correcteurs choisissent $10$ copies au hasard et regardent les notes : $2$-$5$-$7$-$3$-$1$-$9$-$10$-$6$-$7$-$4$.

**a.** Calculer la moyenne $\overline{x}_1$ de cet échantillon.

**b.** Les correcteurs décident d'augmenter les notes en les multipliant par $1,5$ puis en ajoutant $3$ points.
Calculer la moyenne $\overline{x}_2$ du nouvel échantillon de deux façons différentes.

**3** Pour payer ses impôts, Tom a versé neuf mensualités de $286$ € et trois mensualités de $337$ €.
Calculer le montant moyen mensuel de son impôt. Arrondir à l'euro près.

**4** On a relevé les températures, en °C, sur $24$ h dans la région de Hyères (Var) lors du mois de novembre.
Les températures relevées, classées par ordre croissant, sont les suivantes :
$5$-$6$-$6$-$7$-$8$-$8$-$8$-$8$-$9$-$10$-$12$-$12$-$14$-$15$-$15$-$16$-$18$-$18$-$18$-$19$-$19$-$20$-$21$-$22$-$22$.
Déterminer la température médiane du mois de novembre.

**5** À l'aide du tableau ci-dessous, déterminer l'âge moyen $\overline{a}$ et l'âge médian $a_{med}$ des élèves d'un collège.

| Tranche d'âge | $[10 ; 11[$ | $[11 ; 13[$ | $[13 ; 14[$ | $[14 ; 18[$ |
|---|---|---|---|---|
| Effectif | 19 | 112 | 125 | 134 |

Compléments numériques
↘ Sur Sacado via votre ENT
À consulter dans "Livre numérique" en indiquant le numéro de page : 206

## E Quartiles

**↘ Définition**

**Le premier quartile**, noté $Q_1$, est la plus petite valeur de la série telle qu'au moins $\frac{1}{4}$ des valeurs, soit au moins $25\%$ des valeurs de la série sont inférieures ou égales à $Q_1$. $Q_1$ **est une valeur de la série**.

**Le troisième quartile**, noté $Q_3$, est la plus petite valeur de la série telle qu'au moins $\frac{3}{4}$ des valeurs, soit au moins $75\%$ des valeurs de la série sont inférieures ou égales à $Q_3$. $Q_3$ **est une valeur de la série**.

Exercice résolu

Calculer $Q_1$ et $Q_3$ de la série donnée : $9-3-4-14-6-14-15-17-11-2-11-10-7-11-8$

Etape 1 : Ordonner la série par ordre croissant. $2-3-4-6-7-8-9-10-11-11-11-14-14-15-17$ soit $15$ valeurs.
Etape 2 : Calculer les quartiles. $15$ n'est pas un multiple de $4$.

· $\frac{1}{4} \times 15 = 3{,}75$ donc $Q_1$ est la $4^e$ valeur de la série ordonnée, donc $Q_1 = 6$.

· $\frac{3}{4} \times 15 = 11{,}75$ donc $Q_3$ est la $12^e$ valeur de la série ordonnée, donc $Q_3 = 14$.

## F Étendue d'une série

**↘ Définition**

**L'étendue** d'une série statistique est l'écart entre la plus grande valeur ($max$) et la plus petite valeur ($min$) de la série : $e = max - min$.

Interprétation

L'étendue permet de mesurer la dispersion des données d'une série statistique.
Comme elle ne dépend que des valeurs extrêmes, l'étendue ne donne pas d'indication sur la répartition des autres données.

## G Écart interquartile d'une série

**↘ Définition**

**L'écart interquartile** d'une série statistique de premier quartile $Q_1$ et de troisième quartile $Q_3$ est égal à la différence $Q_3$ - $Q_1$.

Interprétation

L'écart interquartile d'une série mesure la dispersion autour de la moyenne ; il n'est pas influencé par les valeurs extrêmes.
L'intervalle entre les quartiles contient au moins $50\%$ des valeurs de la série.

## H Variance et écart type

**↘ Définitions**

· **La variance** $V$ d'une série statistique est la moyenne des carrés des écarts entre les valeurs et la moyenne.
Pour une série de moyenne $\overline{x}$ dont les valeurs sont $x_1$ ; $x_2$ ; ... ; $x_k$ et les effectifs correspondants sont $n_1$ ; $n_2$ ; ... ; $n_k$ on a :

$$V = \frac{n_1 \times (x_1 - \overline{x})^2 + n_2 \times (x_2 - \overline{x})^2 + \ldots + n_k \times (x_k - \overline{x})^2}{n_1 + n_2 + \ldots + n_k}$$

· **L'écart type** $\sigma$ (se lit "sigma") d'une série statistique de variance $V$ est : $\sigma = \sqrt{V}$.
L'écart type se calcule généralement à l'aide des modules "statistiques" des calculatrices.

**Remarque**

· L'écart type exprime la dispersion des valeurs d'une série statistique autour de sa moyenne.
· Les valeurs extrêmes influencent l'écart type.
· L'écart type est exprimé dans la même unité que les valeurs de la série.

**Compléments numériques**

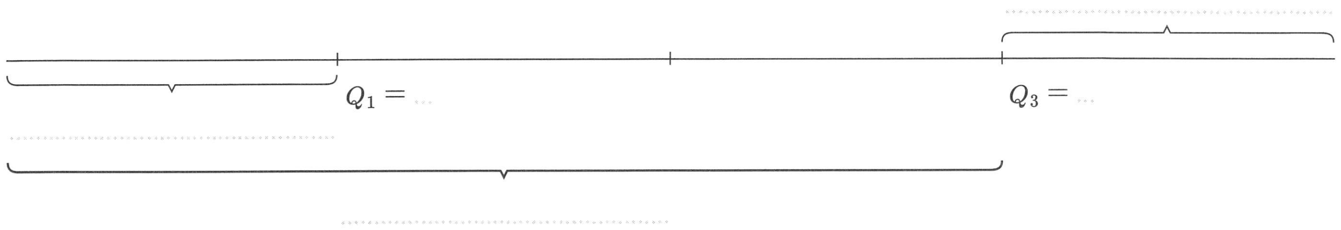
↘ *Sur Sacado via votre ENT*
À consulter dans "Livre numérique"
en indiquant le numéro de page : **207**

**6** Un distributeur automatique de café propose des expressos. Le tableau qui suit indique la masse de café utilisée (en grammes) pour chaque expresso d'une série de $30$ expressos.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 81 | 82 | 85 | 83 | 83 | 82 | 87 | 84 | 85 | 84 |
| 84 | 81 | 83 | 86 | 84 | 80 | 80 | 79 | 87 | 85 |
| 81 | 82 | 85 | 87 | 79 | 80 | 86 | 89 | 83 | 89 |

Déterminer le premier quartile et le troisième quartile. Légender le schéma ci-dessous avec les résultats obtenus.

.........................................................................................................................................




**7** Une entreprise fabrique des boules de Noël. La machine est calibrée pour que le diamètre des boules soit de $12{,}5\ cm$. On a relevé les diamètres de $180$ boules.

| Diamètre en $mm$ | 120 | 121 | 122 | 123 | 124 | 125 | 126 | 127 | 128 | 129 | 130 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Effectifs | 19 | 18 | 16 | 25 | 17 | 15 | 10 | 9 | 22 | 9 | 20 |

Déterminer la médiane et l'écart interquartile.

.........................................................................................................................................

.........................................................................................................................................

.........................................................................................................................................

**8** Voici une série de $10$ températures, en °C, relevées dans $10$ villes françaises : $11$-$12$-$13$-$7$-$7$-$10$-$8$-$5$-$15$-$7$.
Déterminer les indicateurs suivants :

$\overline{t} =$ .............................. , l'étendue $e =$ .............................. , $\sigma =$ ..............................

médiane $=$ .............................. , $Q_1 =$ .............................. , $Q_3 =$ .............................. , $I = Q_3 - Q_1$ ..............................

**9** On considère les deux séries de données suivantes : Série 1 : $42$-$12$-$35$-$17$-$28$-$19$-$10$
Série 2 : $36$-$14$-$35$-$17$-$28$-$19$-$15$

Calculer la moyenne, l'écart type, la médiane et l'écart interquartile de chaque série.
Décrire chaque série et les comparer.

.........................................................................................................................................

.........................................................................................................................................

.........................................................................................................................................

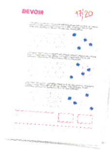
.........................................................................................................................................

# Exercices | **Parcours 1**

**Compléments numériques**
↘ *Sur Sacado via votre ENT*
À consulter dans "Livre numérique"
en indiquant le numéro de page : **208**

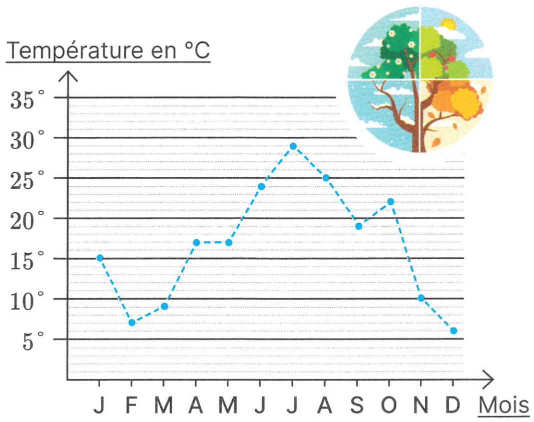
**1** Voici une série de $18$ notes sur $20$ :
$2$-$3$-$7$-$7$-$8$-$10$-$11$-$11$-$11$-$12$-$14$-$14$-$16$-$17$-$18$-$18$-$19$-$20$.
Calculer la moyenne de cette série. Arrondir au centième.

**2** Voici les notes des élèves de la classe au dernier devoir : $6$-$11$-$8$-$19$-$7$-$14$-$16$-$15$-$17$-$18$-$6$-$15$-$5$-$11$-$15$-$16$.
Déterminer la médiane et donner une interprétation de sa valeur.



**3** Au diplôme national du brevet, la moyenne des $25$ élèves de la 3ᵉ A est $13{,}5$ et la moyenne des $29$ élèves de la 3ᵉ B est $13$.
Calculer la moyenne $\overline{m}$ des deux classes réunies. Arrondir au centième.

**4** Le 15 de chaque mois, à midi, Mathis a relevé la température durant toute une année.
Il a représenté l'ensemble des températures relevées sur ce graphique.
Déterminer l'étendue puis la moyenne annuelle et la médiane de cette série.




**5** Un commerçant a relevé les montants des achats en euros des $10$ derniers clients :
$115{,}34$ - $122{,}52$ - $113{,}05$ - $78{,}38$ - $105{,}27$ - $130{,}33$ - $98{,}15$ - $95{,}27$ - $105{,}35$ - $87{,}95$.
Déterminer les indicateurs suivants :

$\overline{m} =$ ……………… , l'étendue $e =$ ……………… , $\sigma =$ ………………

médiane $=$ ……………… , $Q_1 =$ ……………… , $Q_3 =$ ……………… , $I = Q_3 - Q_1$ ………………

# Exercices | **Parcours 2**

**Compléments numériques**
*Sur Sacado via votre ENT*
À consulter dans "Livre numérique"
en indiquant le numéro de page : **209**

**1** Dans une boulangerie, la masse affichée de la baguette est $250$ grammes.
Lors d'un contrôle, un agent a relevé la masse de $120$ baguettes. Voici ses résultats :

| Masse des baguettes ($g$) | $248$ | $249$ | $250$ | $251$ | $252$ |
|---|---|---|---|---|---|
| Fréquence | $21\%$ | $19\%$ | $22\%$ | $21\%$ | $17\%$ |

a. Calculer la masse moyenne d'une baguette.

b. Calculer l'écart type de la série.

**2** La capacité vitale est le volume d'air maximal pouvant être expiré après une inspiration maximale.
Sur un échantillon de $17$ personnes, on a mesuré la capacité vitale (en litres). Voici la liste des résultats :
$4{,}15$ - $4{,}48$ - $5{,}24$ - $4{,}8$ - $4{,}95$ - $4{,}05$ - $4{,}3$ - $4{,}7$ - $5{,}51$ - $4{,}58$ - $4{,}12$ - $5{,}7$ - $4{,}85$ - $5{,}05$ - $4{,}65$ - $4{,}7$ - $4{,}28$

a. À l'aide de la calculatrice, déterminer l'étendue, la moyenne et l'écart type de cette série. Arrondir au centilitre près.

b. Déterminer la médiane, les quartiles et l'écart interquartile de cette série.

**3** Dans un immeuble, on a relevé les surfaces (en $m^2$) de tous les appartements :
Quelle série est la plus dispersée autour de sa moyenne ?

Appartements de 1 à 3 pièces

| Surfaces ($m^2$) | $[0\,;20[$ | $[20\,;40[$ | $[40\,;60[$ | $[60\,;100[$ |
|---|---|---|---|---|
| Effectifs | $17$ | $15$ | $7$ | $9$ |

Appartements de 4 à 6 pièces

| Surfaces ($m^2$) | $[40\,;60[$ | $[60\,;80[$ | $[80\,;100[$ | $[100\,;140[$ |
|---|---|---|---|---|
| Effectifs | $16$ | $17$ | $24$ | $11$ |

**4** Paul étudie une série statistique. Il calcule la moyenne et obtient $\overline{x} = 6$.
Le professeur annonce que la valeur de la moyenne qu'il fallait trouver est égale à $\overline{x} = 7{,}5$.
Paul s'aperçoit qu'il a oublié une valeur en utilisant seulement $7$ valeurs.

a. Justifier, sans calculer, si la valeur oubliée est inférieure ou supérieure à $6$.

b. Calculer la valeur manquante.

# Algorithmique, TICE & logique

**Compléments numériques**
↘ *Sur Sacado via votre ENT*
À consulter dans "Livre numérique"
en indiquant le numéro de page : **212**

**1** Paul a obtenu les notes $6$, $10$ et $12$ lors de ses trois premiers contrôles de mathématiques. Il a codé un programme Python qui calcule sa moyenne.

a. Quel est le rôle de la fonction **note_coeff** (ci-dessous) ?

b. Compléter le code pour qu'il renvoie la moyenne de Paul.

```
def note_coeff(n,c):
      return n*c

ok, nc, co = True, 0, 0
while ok :
   note = .........(input("Quelle est la note ?"))
   coeff = .........(input("Quel est le coefficient ?"))
   nc = nc + note_coeff(note,coeff)
   co = co + ............
   suite = input("Saisir une autre note ? O/N ")
   if suite == "N" :
      ok = False
if co > 0 : print(round(nc/co,2))
else :
   print("Saisir au moins une note")
```

**2** On considère la série statistique donnée par le tableur ci-contre. L'objectif de cet exercice est de déterminer la moyenne et l'écart type de cette série.

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | valeur xi | 0 | 1 | 2 | 3 | 4 | 5 | |
| 2 | effectif ni | 1 | 7 | 37 | 66 | 65 | 24 | 200 |
| 3 | xi*ni | | | | | | | |
| 4 | ni*(xi-moy)^2 | | | | | | | |

a. À quoi correspond la valeur $200$ ? Comment l'obtenir ?

b. Quelle formule faut-il saisir en **B3** puis étirer vers la droite ?

c. Quelle formule faut-il saisir en **H3** pour obtenir la moyenne de la série ? Donner sa valeur.

d. Quelle formule faut-il saisir en **B4** puis étirer vers la droite ?

e. Comment obtenir l'écart type en **H4** ? Donner sa valeur.

f. Créer ce fichier pour tester vos réponses.