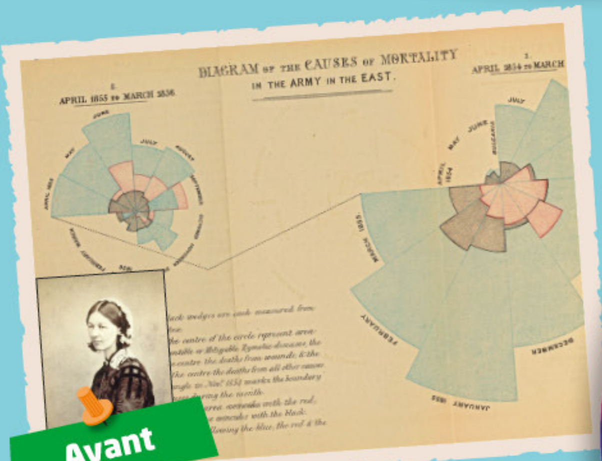
12 Statistique descriptive



Avant

▶ Au 19e siècle, les statistiques et représentations graphiques produites par l'infirmière Florence Nightingale (1820-1910) ont contribué à faire chuter de façon spectaculaire la mortalité des soldats britanniques pendant la guerre de Crimée (1853-1856).

À présent

▶ Au 21e siècle, les statistiques médicales produites à l'aide de puissants ordinateurs aident les chercheurs à prolonger et à améliorer la vie de nombreux patients.

Les capacités travaillées dans ce chapitre

	Exercices
• Calculer une moyenne pondérée.	1, 3, 9 à 16
• Utiliser la linéarité de la moyenne.	17 à 20
• Déterminer la médiane et les quartiles d'une série.	2, 4, 5, 7, 21 à 32
• Utiliser des indicateurs de dispersion (étendue, écart interquartile, écart-type).	6, 8, 33 à 41
• Décrire les différences entre deux séries statistiques.	42 à 46
• Lire et comprendre une fonction écrite en langage Python renvoyant la moyenne m, l'écart-type s et la proportion d'éléments appartenant à $[m-2s\,;\,m+2s]$.	54 à 56, 66, 67

1 Indicateurs de tendance centrale

La médaille Fields récompense tous les 4 ans un mathématicien ou une mathématicienne de moins de 40 ans. Le tableau ci-dessous présente l'âge des 60 lauréats de la médaille Fields.

Âge	27	28	30	31	32	33	34	35	36	37	38	39
Effectif	1	3	5	4	3	4	6	5	7	7	8	7
Effectif cumulé croissant	1	4										

1 Déterminer l'âge moyen des 60 lauréats. *Arrondir au dixième.*

2 **a)** Combien de lauréats étaient âgés de 30 ans ou moins ?
Ce nombre est appelé **effectif cumulé croissant** de la valeur 30.
b) Compléter la ligne des effectifs cumulés croissants.

3 **a)** Déterminer la médiane de cette série. Interpréter le résultat.
b) Quelle est la plus petite valeur de la série telle qu'au moins un quart des lauréats ont un âge inférieur ou égal à cette valeur ? Cette valeur, notée Q_1, est le **premier quartile** de la série.
c) En 2014, l'iranienne Maryam Mirzakhani fut la 1re femme à recevoir la médaille Fields. Elle avait 37 ans.
Charlotte affirme : « Au moins les trois quarts des lauréats ont un âge inférieur ou égal à 37 ans. »
A-t-elle raison ? Le nombre 37 est le **troisième quartile,** noté Q_3, de la série.

2 Un indicateur de dispersion : l'écart-type

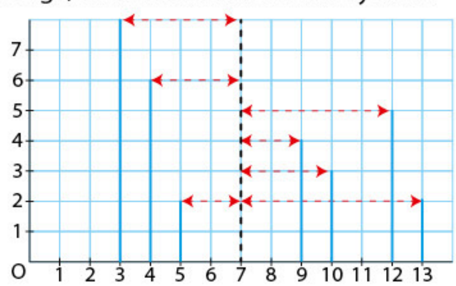
Chaque jour du mois de septembre, Cédric a relevé la distance qu'il a parcourue lors de son footing.

Distance (en km)	3	4	5	9	10	12	13
Effectif	8	6	2	4	3	5	2

1 Vérifier que la moyenne de cette série est égale à 7 km.

2 On a représenté ci-contre la série (en bleu), ainsi que les écarts (en rouge) entre les valeurs et la moyenne.

Distance (en km)	3	4	5	9	10	12	13
Écart	−4						
Carré de l'écart	16						
Effectif	8	6	2	4	3	5	2



a) Recopier le tableau et compléter la ligne « Écart ».
b) Pour n'avoir que des nombres positifs, on calcule le carré de chaque écart. Compléter cette ligne du tableau.
c) Calculer la moyenne V des carrés des écarts pondérée par les effectifs. *Arrondir au dixième.*
On dit que V est la **variance** de la série des distances.
Dans quelle unité est exprimée la variance V ?
d) Calculer $s = \sqrt{V}$. *Arrondir au dixième.* On dit que s est l'**écart-type** de la série des distances, il mesure la dispersion des données autour de la moyenne. Dans quelle unité est exprimé s ?

Cours

1 Indicateurs de tendance centrale

A La moyenne pondérée

Définition

La **moyenne pondérée** de la série statistique (x_i) ci-contre est le nombre réel, noté $\boldsymbol{m}$ (*ou* $\overline{x}$), tel que :

$$\boldsymbol{m = \frac{n_1x_1 + n_2x_2 + \ldots + n_px_p}{N}} \text{ ou } \boldsymbol{m = f_1x_1 + f_2x_2 + \ldots + f_px_p}.$$

Valeur	x_1	x_2	…	x_p
Effectif	n_1	n_2	…	n_p
Fréquence	f_1	f_2	…	f_p

Effectif total : $N = n_1 + n_2 + \ldots + n_p$

Remarque : pour une série regroupée en classes, on obtient une **valeur approchée** de la moyenne en prenant pour x_i les centres des classes.

B Linéarité de la moyenne

Propriété

a et b désignent des nombres réels.
Si la série (x_i) a pour moyenne m, alors la série $(ax_i + b)$ a pour moyenne $\boldsymbol{M = am + b}$.

x_i	x_1	x_2	…	x_p
$ax_i + b$	$ax_1 + b$	$ax_2 + b$	…	$ax_p + b$
Effectif	n_1	n_2	…	n_p

Démonstration

$$M = \frac{n_1(ax_1 + b) + n_2(ax_2 + b) + \ldots + n_p(ax_p + b)}{n_1 + n_2 + \ldots + n_p} = \frac{a(n_1x_1 + n_2x_2 + \ldots + n_px_p) + b(n_1 + n_2 + \ldots + n_p)}{n_1 + n_2 + \ldots + n_p}$$

$$M = a \times \frac{n_1x_1 + n_2x_2 + \ldots + n_px_p}{n_1 + n_2 + \ldots + n_p} + b \times \frac{n_1 + n_2 + \ldots + n_p}{n_1 + n_2 + \ldots + n_p} = am + b$$

C La médiane et les quartiles

Définition

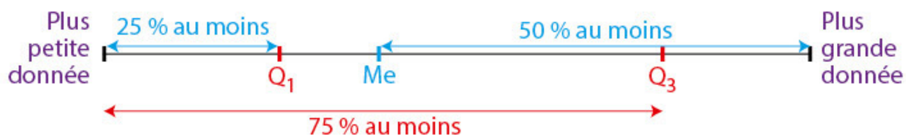
La **médiane** d'une série de N valeurs **rangées par ordre croissant** est le nombre Me, tel qu'**au moins** 50 % des valeurs de la série lui sont inférieures ou égales et au moins 50 % lui sont supérieures ou égales.

- Si N est **impair**, Me est la valeur centrale.
- Si N est **pair**, Me est la demi-somme des deux valeurs centrales.

Définitions

Les valeurs d'une série statistique étant **rangées par ordre croissant** :

- **le premier quartile** est la plus petite valeur Q_1 de la série telle qu'**au moins 25 %** des valeurs de la série sont **inférieures ou égale à Q_1** ;
- **le troisième quartile** est la plus petite valeur Q_3 de la série telle qu'**au moins 75 %** des valeurs de la série sont **inférieures ou égale à Q_3**.



2 Indicateurs de dispersion

A Étendue et écart interquartile

Définitions

- **L'étendue** est la différence entre la plus grande et la plus petite valeur d'une série.
- **L'écart interquartile** est la différence $Q_3 - Q_1$ entre les troisième et premier quartiles.

B Écart-type

Par la suite, on considère la série statistique (S) définie par le tableau ci-contre. Son effectif total est $N = n_1 + n_2 + \ldots + n_p$.

Valeur	x_1	x_2	$\ldots$	x_p
Effectif	n_1	n_2	$\ldots$	n_p

Définitions

- La **variance** de la série statistique (S) est le nombre réel positif, noté V, tel que :

$$\mathbf{V} = \frac{n_1(x_1 - m)^2 + n_2(x_2 - m)^2 + \ldots + n_p(x_p - m)^2}{\mathbf{N}}$$

- L'**écart-type** est le nombre réel positif, noté s (ou σ) tel que $\mathbf{s} = \sqrt{\mathbf{V}}$.

Remarques :

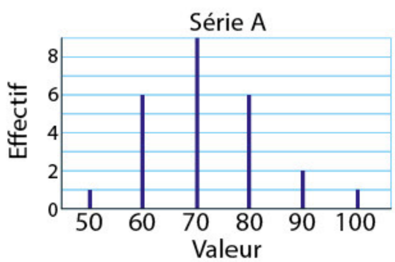
- L'écart-type mesure la dispersion des valeurs autour de la moyenne : plus il est grand, plus les valeurs sont éloignées de la moyenne.
- L'écart-type est exprimé dans **la même unité** que celle du caractère étudié.
- Il est souvent intéressant de calculer la proportion des valeurs de la série appartenant à l'intervalle $[m - 2s\,; m + 2s]$, c'est-à-dire distantes de moins de deux écarts-types de la moyenne m de la série.

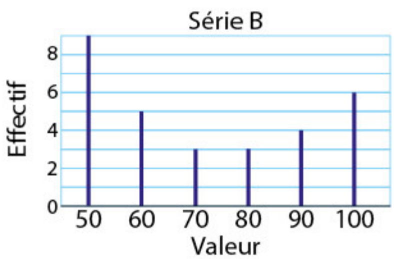
C Décrire les différences entre deux séries

On résume souvent une série statistique par un couple associant un indicateur de tendance centrale à un indicateur de dispersion ; généralement on choisit (**médiane** ; **écart interquartile**) ou (**moyenne** ; **écart-type**). La donnée d'un tel couple permet de comparer deux séries.

Exemple

On se propose de comparer les séries A et B étudiant le même caractère dans deux populations différentes.





La moyenne des deux séries est 72. On remarque cependant sur les diagrammes en bâtons que les valeurs de la série B semblent plus dispersées que celles de la série A.

On détermine ensuite le couple (**moyenne** ; **écart-type**) pour chaque série. L'observation faite sur les diagrammes est confirmée par la comparaison des écarts-types de chaque série.

	Effectif total	Moyenne	Écart-type
Série A	25	$m_A = 72$	$s_A \approx 11{,}3$
Série B	30	$m_B = 72$	$s_B \approx 19{,}4$

Acquérir des automatismes

EXERCICES RÉSOLUS

1 Calculer une moyenne pondérée

→ Cours 1. A et B

Le tableau ci-dessous donne les salaires mensuels nets des employés d'une PME.

Salaire (en €)	1 100	1 200	1 500	1 800	2 200	2 800
Nombre d'employés	6	10	9	14	7	4

a) Calculer le salaire mensuel net moyen des employés de cette PME.

b) L'entreprise décide une augmentation de 100 € du salaire mensuel net de chaque employé. Calculer le nouveau salaire mensuel net moyen des employés de l'entreprise.

Solution

a) L'effectif total de cette série est $6+10+9+14+7+4=50$.

$$m=\frac{6\times 1100+10\times 1200+9\times 1500+14\times 1800+7\times 2\,200+4\times 2\,800}{50}=\frac{83\,900}{50}=1678$$

Le salaire mensuel net moyen dans cette PME est 1 678 €.

b) Chaque valeur de la série est augmentée de 100 €. La moyenne des salaires de l'entreprise est donc augmentée de 100 €. Le nouveau salaire mensuel net moyen est alors de 1 778 €.

Il est inutile de calculer les salaires augmentés pour obtenir leur moyenne, grâce à la linéarité de la moyenne.

2 Déterminer la médiane d'une série

→ Cours 1. C

Le 1er juillet 2018 est entrée en vigueur la limitation de vitesse à 80 km·h^{-1} sur certaines routes secondaires. Ce tableau donne les vitesses mesurées pendant 10 min consécutives sur l'une de ces routes. Déterminer la médiane de cette série. Interpréter le résultat.

Vitesse (km·h^{-1})	76	77	79	80	82	87	97
Nombre de véhicules	11	8	11	13	5	6	1

Solution

L'effectif total est impair : $55=2\times 27+1$, donc la médiane est la 28e vitesse.

$11+8=19$ et $19+11=30$. De la 20e à la 30e, les valeurs sont des 79, donc Me $=79$ km·h^{-1}.

Ainsi, au moins 50 % des véhicules contrôlés roulaient à 79 km·h^{-1} ou moins (ou 79 km·h^{-1} ou plus).

On cumule les effectifs jusqu'à atteindre ou dépasser pour la 1re fois 28.

EXERCICES D'APPLICATION DIRECTE

Sur le modèle de l'exercice résolu 1

3 Dans un magasin, on a relevé les prix de mugs.

Prix (en €)	3	6	7	8,50
Nombre de mugs	3	7	4	6

a) Calculer le prix moyen des mugs dans ce magasin.

b) Le magasin décide une baisse de 10 % des prix. Calculer le prix moyen des mugs après cette réduction.

Sur le modèle de l'exercice résolu 2

4 La série considérée est celle de l'exercice 3.

a) Déterminer le prix médian. Interpréter le résultat.

b) 5 mugs supplémentaires sont vendus 9,50 €. Déterminer la médiane de la nouvelle série. Interpréter le résultat.

EXERCICES RÉSOLUS

5 Déterminer les quartiles d'une série

→ Cours 1. C

On considère à nouveau la série de l'exercice résolu 1.

Salaire (en €)	1 100	1 200	1 500	1 800	2 200	2 800
Nombre d'employés	6	10	9	14	7	4

a) Déterminer les quartiles de cette série de salaires mensuels nets.

b) Interpréter les résultats dans le contexte de l'exercice.

Solution

a) • $\frac{1}{4} \times 50 = 12{,}5$ donc le 1er quartile est le 13^{e} salaire de la série. Ainsi $Q_1 = 1200$ €.

• $\frac{3}{4} \times 50 = 37{,}5$ donc le 3^{e} quartile est le 38^{e} salaire de la série. Ainsi $Q_3 = 1800$ €.

b) Au moins 25 % des employés de cette PME ont un salaire inférieur ou égal à 1 200 €.
Au moins 75 % des employés de cette PME ont un salaire inférieur ou égal à 1 800 €.

> Pour déterminer Q_1, on calcule $\frac{N}{4}$ et le rang de Q_1 est le nombre entier immédiatement supérieur ou égal à $\frac{N}{4}$.

> Pour déterminer Q_3, on calcule $\frac{3N}{4}$ et le rang de Q_3 est le nombre entier immédiatement supérieur ou égal à $\frac{3N}{4}$.

6 Déterminer l'écart-type d'une série

→ Cours 2. B

Lors d'un contrôle qualité, on a compté le nombre de vis contenues dans 30 sachets.

Nombre x_i de vis par sachet	18	19	20	21	22	23
Nombre de sachets	2	9	13	2	1	3

a) Calculer le nombre moyen m de vis par sachet.

b) Déterminer la variance V dans un tableau et en déduire l'écart-type s. *Arrondir au centième.*

c) Combien de sachets contiennent un nombre de vis appartenant à l'intervalle $[m - 2s\ ;\ m + 2s]$?

Solution

a) $m = \frac{2 \times 18 + 9 \times 19 + 13 \times 20 + 2 \times 21 + 1 \times 22 + 3 \times 23}{30} = \frac{600}{30} = 20$

b)

x_i	18	19	20	21	22	23
$x_i - m$	−2	−1	0	1	2	3
$(x_i - m)^2$	4	1	0	1	4	9

$V = \frac{2 \times 4 + 9 \times 1 + 13 \times 0 + 2 \times 1 + 1 \times 4 + 3 \times 9}{30}$

donc $V \approx 1{,}67$. $s = \sqrt{V}$ donc $s \approx 1{,}29$.

c) On détermine les bornes de l'intervalle : $m - 2s \approx 17{,}4$ et $m + 2s \approx 22{,}6$.
$2 + 9 + 13 + 2 + 1 = 27$ donc 27 sachets ont un nombre de vis appartenant à $[m - 2s\ ;\ m + 2s]$.

> Ici, le calcul à la main de la variance permet de mieux comprendre le fonctionnement de la formule. Par la suite, il sera plus rapide d'utiliser la calculatrice (voir p. XIII, XV ou XVII) pour calculer les indicateurs d'une série.

EXERCICES D'APPLICATION DIRECTE

Sur le modèle de l'exercice résolu 5

7 La série considérée est celle de l'exercice 6.

a) Déterminer les quartiles de cette série de nombres de vis par sachet.

b) Interpréter les résultats dans le contexte de l'exercice.

Sur le modèle de l'exercice résolu 6

8 La série considérée est celle de l'exercice 5.

a) Déterminer la variance V dans un tableau et en déduire l'écart-type s. *Arrondir au centième.*

b) Combien d'employés ont un salaire appartenant à l'intervalle $[m - 2s\ ;\ m + 2s]$?

Calcul mental → Exercez-vous sur la plateforme Hyperbole
+ de détails page 4

Moyenne pondérée

→ Cours 1. A

Questions Flash

9 Sur la carte d'un restaurant figurent 10 desserts répartis selon quatre prix.

Prix (en €)	4	5	6	8
Effectif	3	4	2	1

Calculer mentalement le prix moyen d'un dessert.

10 L'âge moyen des joueurs de l'équipe de France de football championne du monde en 2018 était de 26 ans. Flora affirme : « Cela signifie qu'une moitié de l'équipe avait moins de 26 ans et l'autre moitié plus de 26 ans. » Que peut-on en penser ?

11 Le tableau ci-dessous donne la répartition des salaires annuels, exprimés en milliers d'euros, dans une entreprise.

Salaire	15	17	20	24	36	42
Effectif	4	6	10	14	5	1

Calculer le salaire annuel moyen dans l'entreprise.

12 Calculer le nombre moyen de fautes de frappe dans un texte tapé pendant un stage de formation.

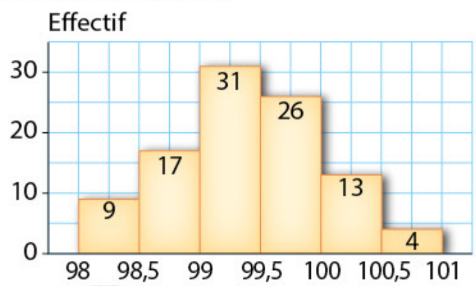
Fautes	0	1	2	3	4	5	6	7
Effectif	8	12	18	26	16	10	6	4

13 Déterminer la moyenne de cette série.

Valeur	54	59	68	88
Fréquence	0,17	0,19	0,48	0,16

14 Dans une ville, 20 % des familles ont un enfant, 40 % ont deux enfants, 25 % ont trois enfants, 10 % ont quatre enfants et 5 % ont cinq enfants.
Calculer le nombre moyen d'enfants par famille.

15 En utilisant les centres des classes, calculer une valeur approchée de la moyenne de la série ci-dessous. *Arrondir au centième.*



16 On compte le nombre de bonbons vendus dans des paquets de 100 g.

Nombre de bonbons	38	39	40	41	42	43	44	45	46
Nombre de paquets	14	34	78	120	350	145	21	10	2

Arrondir les résultats au dixième.
a) Calculer la moyenne de cette série.
b) En déduire la masse moyenne d'un bonbon.

Linéarité de la moyenne

→ Cours 1. B

Questions Flash

17 La série suivante a pour moyenne 82,8 :
71 – 80 – 83 – 88 – 92
Calculer mentalement la moyenne de chaque série.
a) 171 – 180 – 183 – 188 – 192
b) 7 100 – 8 000 – 8 300 – 8 800 – 9 200
c) 713 – 803 – 833 – 883 – 923

18 Sans calculatrice, déterminer la moyenne de la série.
a) 703 – 704 – 708 – 710 – 715
b) 0,037 – 0,041 – 0,043 – 0,060 – 0,069

19 On considère la série suivante :
10 – 14 – 14 – 11 – 11 – 3 – 11 – 10 – 21 – 10
a) Calculer la moyenne *m* de cette série.
b) Mathieu affirme : « Si on retranche *m* à chaque valeur de la série, la moyenne reste inchangée. » Mathieu a-t-il raison ? Justifier.

20 En 2014, Morgane a planté neuf pommiers dans son jardin. Durant l'été 2018, elle a relevé la masse (en kilogramme) de pommes sur chaque arbre :

9,7 – 11,2 – 12,4 – 13,9 –15 –
14,7 –16,1 –17,1 –14,1
a) Calculer la masse moyenne de pommes sur un arbre.
b) L'été suivant, la production de chaque arbre a augmenté de 50 %. Calculer la masse moyenne de pommes sur un arbre en 2019.
c) Morgane a lu que la production de chaque arbre augmente de 20 kg entre la 5^{e} et la 6^{e} année.
Quelle masse moyenne de pommes sur un arbre peut-elle prévoir en 2020 ?

Médiane et quartiles

→ Cours 1. C

Questions flash

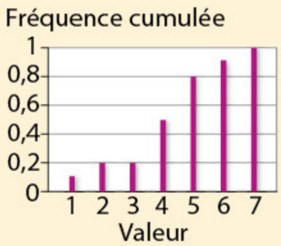
21 Voici les notes obtenues par des élèves lors d'un test sur 100.

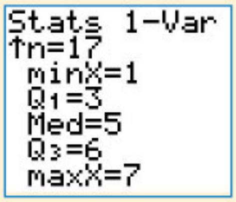
15 – 36 – 44 – 46 – 54 – 54 – 55 – 56
62 – 64 – 73 – 74 – 74 – 83 – 85 – 95

Déterminer mentalement :

a) la médiane ; **b)** les quartiles.

22 Expliquer pourquoi ce graphique ne peut pas correspondre à l'écran de la calculatrice.





Pour les exercices 23 à 25, déterminer la médiane, le premier quartile et le troisième quartile de la série.

23

Valeur	2	4	6	8	10
Effectif	7	10	4	5	4

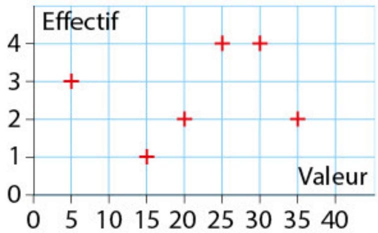
24

Valeur	3	5	11	17	19
Effectif	5	10	9	4	3

25

Valeur	7	12	14	27	32
Fréquence	0,05	0,2	0,4	0,15	0,2

26 Utiliser les effectifs cumulés pour déterminer la médiane et les quartiles de la série représentée.



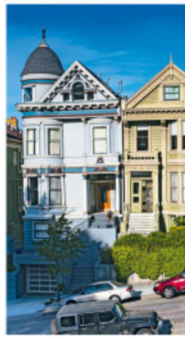
27 Leïla a relevé le nombre de magazines publicitaires reçus par jour dans sa boîte aux lettres, sur une période de deux semaines. Elle a obtenu cette série :

5 – 1 – 3 – 7 – 5 – 5 – 4 – 0 – 1 – 4 – 6 – 3 – 2 – 10

a) Déterminer la médiane et les quartiles.

b) Le dernier jour, Leïla a en fait reçu 7 magazines et non 10. Cela entraîne-t-il des modifications sur les résultats précédents ? Si oui, lesquelles ?

28 Dans une ville des États-Unis, le nombre de véhicules par foyer est réparti de la façon suivante :



Nombre de véhicules	0	1	2	3	4
Nombre de foyers	267	3 402	19 203	20 471	1 657

1. Déterminer la médiane et les quartiles de cette série.

2. Compléter les phrases suivantes :

a) « Au moins … % des foyers possèdent trois véhicules au plus » ;

b) « Au moins 50 % des foyers possèdent … véhicules ou moins ».

29 Le tableau ci-contre donne la part, en pourcentage, des surfaces agricoles dans les régions de France.

Auvergne-Rhône-Alpes	44,4
Bourgogne-Franche-Comté	54,4
Bretagne	63,5
Centre-Val de Loire	61,3
Corse	47,5
Grand Est	53,0
Hauts-de-France	68,6
Île-de-France	48,4
Normandie	70,1
Nouvelle-Aquitaine	49,9
Occitanie	47,9
Pays de la Loire	68,5
Provence-Alpes-Côte d'Azur	25,3

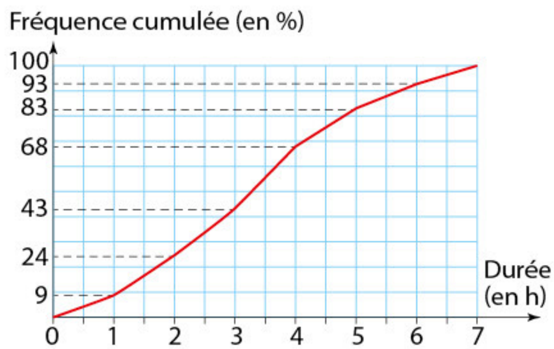
1. Déterminer la médiane et les quartiles de cette série.

2. Compléter chaque phrase.

a) « Dans plus de 50 % des régions, la part des surfaces agricoles est supérieure à …. »

b) « Dans plus de 75 % des régions, la part des surfaces agricoles est inférieure à …. »

30 On a demandé à des familles leur temps de connexion quotidienne à Internet, en heure.
Voici la courbe des fréquences cumulées (croissantes) de cette série de valeurs.



a) Quel pourcentage de ces familles se connectent entre 4 h et 5 h par jour ?

b) Lire graphiquement la médiane et les quartiles de cette série.

c) Interpréter les résultats obtenus à la question **b)**.

31 Le tableau ci-dessous indique les capacités des disques durs, en Go, des ordinateurs d'un magasin.

Go	80	160	250	320	500	800	1 000	1 150
Effectif	2	9	11	7	5	2	4	3

a) Déterminer la médiane de cette série.
b) Déterminer le premier quartile Q_1 et le troisième quartile Q_3.
c) Estimer le pourcentage d'ordinateurs dont la capacité est inférieure ou égale à Q_3.
d) Calculer le pourcentage d'ordinateurs dont la capacité est inférieure ou égale à Q_3.
Arrondir au centième.
e) Calculer le pourcentage d'ordinateurs dont la capacité appartient à l'intervalle $[Q_1 ; Q_3]$.
Arrondir au centième.

32 Le tableau suivant donne la répartition des salaires, en euro, dans une entreprise.

Salaire	900	1 100	1 300	1 500	1 700	1 900	2 100
Effectif	12	10	20	18	12	13	5

Salaire	2 300	2 500	2 700	2 900	3 100	3 300	3 500
Effectif	3	12	5	7	10	0	6

Salaire	3 700	3 900	4 100	4 300	4 500	**Total**
Effectif	5	0	0	0	1	**139**

a) Calculer le salaire moyen S, en euro, dans cette entreprise.
Arrondir à l'unité.
b) Dresser le tableau des effectifs cumulés croissants. En déduire le salaire médian.
c) Quel indicateur de position utiliserait-on pour présenter l'entreprise sous son meilleur jour si l'on était le directeur ? un observateur impartial ?

Étendue et écart interquartile

→ Cours 2. A

Questions flash

33 Voici des indicateurs d'une série : min = 27, $Q_1 = 31$, Me = 34, $m = 35$, $Q_3 = 42$, max = 53.
Calculer mentalement l'étendue et l'écart interquartile de la série.

34 Déterminer mentalement l'étendue de cette série.

Valeur	152	234	148	78	204	128
Fréquence	0,21	0,135	0,15	0,22	0,03	0,255

35 Voici les scores obtenus par Lou et Alan à des parties d'un jeu vidéo.
Lou : 700 – 1 810 – 980 – 625 – 1 090 – 975 – 1 350 – 1 400 – 885 – 1 200
Alan : 720 – 1 540 – 800 – 950 – 1 100 – 1 625 – 790 – 1 335 – 915 – 1 010 – 1 480
Pour quel joueur :
a) l'étendue est-elle la plus grande ?
b) l'écart interquartile est-il le plus grand ?

36 Dans un laboratoire d'astrophysique, un détecteur a relevé les durées d'attente, en h, entre les réceptions successives des particules captées.
Déterminer Q_1, Q_3 et l'écart interquartile.

75 – 265 – 225 – 402 – 35
105 – 411 – 346 – 159 – 229
62 – 256 – 431 – 177 – 56
144 – 354 – 178 – 386 – 294

Variance et écart-type

→ Cours 2. B

Questions flash

37 Calculer mentalement la moyenne puis l'écart-type de cette série.

Valeur	4	5	6	7
Effectif	4	3	2	1

38 Dans chaque cas, calculer la moyenne et l'écart-type de la série.
a) La série comporte 40 termes égaux à 15.
b) La série comporte 20 termes égaux à 10 et 20 termes égaux à 20.

39 Nikola note le nombre de buts qu'il marque par match dans son équipe de handball.

Nombre de buts par match	3	4	5	6	7	8	9	10
Nombre de matchs	1	2	7	10	8	6	4	2

Déterminer en appliquant les formules du cours :
a) la moyenne pondérée ; **b)** l'écart-type.
Arrondir au dixième si besoin.

40 a) Saisir les données de la série ci-dessous à la calculatrice.

Valeur	8,7	9,9	10,4	10,7	12,1	12,7
Effectif	1	3	4	6	5	2

b) Donner sa moyenne m et son écart-type s.
Arrondir au centième.

c) Calculer le pourcentage des données de la série qui se trouvent dans chaque intervalle :
- $[m - s \,;\, m + s]$
- $[m - 2s \,;\, m + 2s]$

Arrondir à l'unité.

41 Cinquante-quatre concurrents participent à un « triathlon découverte ». Leurs résultats sont donnés dans le tableau ci-dessous.

Temps (en min)	[30 ; 35[	[35 ; 40[	[40 ; 45[	[45 ; 50[
Effectif	1	16	8	12

Temps (en min)	[50 ; 55[	[55 ; 60[	[60 ; 65[
Effectif	5	9	3

a) Déterminer, en prenant les centres des classes, une valeur approchée arrondie à l'unité :
- du temps moyen m d'un concurrent ;
- de l'écart-type s de la série.

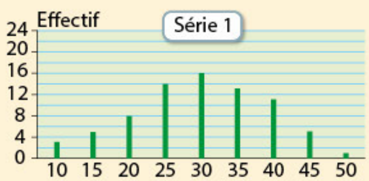
b) Le couple $(m\,;\,s)$ semble-t-il pertinent pour résumer cette série ?

Comparaison de séries

→ Cours 2. C

Questions flash

42 Voici les diagrammes en bâtons de deux séries.

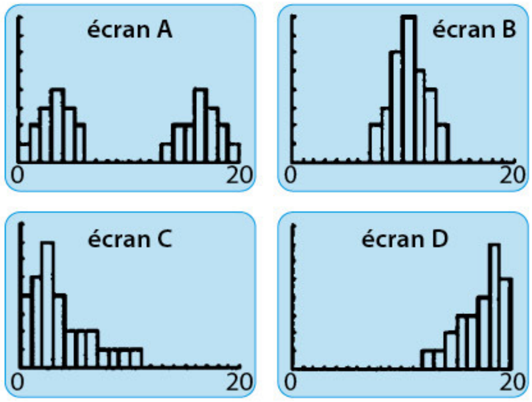


Par lecture graphique :

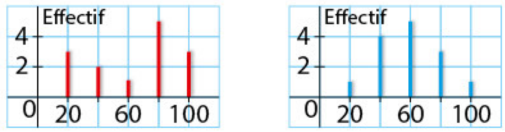
a) estimer la moyenne de chaque série ;

b) conjecturer laquelle des deux séries a le plus grand écart-type.

43 Sans faire de calcul, décrire les différences entre les quatre séries représentées ci-dessous.



44 a) Par lecture graphique, conjecturer la série qui a le plus grand écart-type.



b) Avec la calculatrice, déterminer la moyenne et l'écart-type de chaque série. *Arrondir au dixième.*

45 Corentin relève pendant 14 jours la température à 10 h du matin dans son jardin.
Voici ses résultats :

Température (en °C)	18	19	20	21	23	25	27
Nombre de jours	1	1	1	5	3	2	1

1. Calculer la température moyenne à 10 h du matin sur cette période.

2. Il compare son relevé à celui de l'année précédente sur la même période.

Température (en °C)	10	12	20	22	24	25	30
Nombre de jours	1	1	2	2	3	4	1

a) Vérifier que la température moyenne des deux séries est la même.

b) Déterminer alors l'écart-type de chaque série et commenter les résultats trouvés. *Arrondir au dixième.*

46 Voici les résultats de deux groupes de candidats à un test de compétence en anglais.

Note	0	1	2	3	4	5	6	7	8	9	10
Effectif groupe A	0	2	2	4	6	12	14	10	3	2	1
Effectif groupe B	3	5	5	6	7	8	9	9	7	6	6

En utilisant des indicateurs statistiques appropriés, décrire les différences entre ces deux séries.

QCM interactifs

Pour se tester

47 Dans chaque cas, donner la réponse exacte sans justifier.

Une société de location possède 80 véhicules. On a relevé, au cours de l'année 2018, le nombre de jours d'immobilisation de chaque véhicule.

Nombre de jours d'immobilisation	1	2	3	4	5	6	7	8
Effectif	5	4	10	20	20	8	8	5

		A	B	C	D
1	Le nombre moyen (arrondi au dixième) de jours d'immobilisation d'un véhicule est …	4,5	4,6	5	367
2	La médiane de cette série est …	supérieure à sa moyenne	inférieure à sa moyenne	égale à sa moyenne	on ne peut pas savoir
3	L'écart-type (arrondi au dixième) de cette série est …	3,1	2	1 931	1,8
4	L'écart interquartile est …	2	4	7	8

48 Dans chaque cas, donner la (ou les) réponse(s) exacte(s) sans justifier.

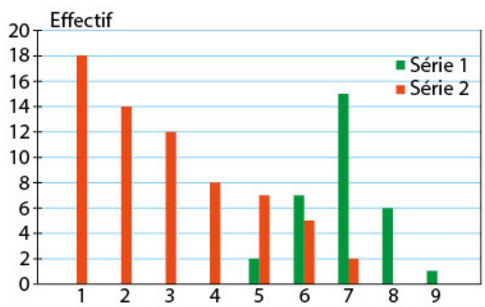
On a relevé le nombre de buts marqués lors des 37 premières journées du championnat de Ligue 1 (2017-2018).

Nombre de buts	14	15	17	21	23	24	25	26	27	28	29	30	31	32	34
Nombre de journées	1	1	1	1	1	6	3	2	1	2	6	1	4	3	4

		A	B	C	D
1	La médiane de la série est telle que …	Me = 27,5	Me = 28	Me = 29	au moins 50 % des valeurs sont supérieures ou égales à Me
2	Le 1er quartile de la série est tel que …	$Q_1 = 9{,}25$	$Q_1 = 21$	$Q_1 = 24$	au moins 75 % des valeurs sont supérieures ou égales à Q_1
3	Au moins 25 % des valeurs de la série sont …	inférieures ou égales à 28	inférieures ou égales à 24	supérieures ou égales à 31	comprises entre 24 et 31

49 Dans chaque cas, dire si l'affirmation est vraie ou fausse en justifiant.

On a représenté ci-contre deux séries statistiques.



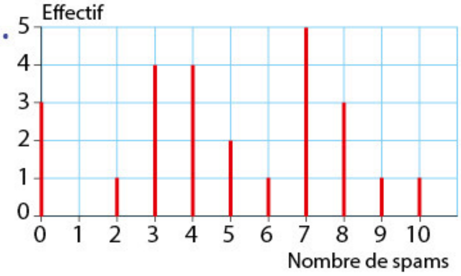
Affirmations :

1. La moyenne de la série 1 est supérieure à celle de la série 2.
2. L'écart-type de la série 1 est inférieur à celui de la série 2.
3. Pour la série 1, la médiane et la moyenne sont proches.
4. Pour la série 2, la médiane est supérieure à la moyenne.

Vérifiez vos réponses : p. 346

Nathan live!

50 Dresser et interpréter un tableau d'effectifs



Le graphique ci-contre illustre le nombre de spams reçus en un jour dans les boîtes aux lettres électroniques des élèves d'une classe.

1. Dresser le tableau des effectifs correspondant.

2. Quel est l'effectif total de la classe ?

3. a) Combien d'élèves ont reçu plus de 7 spams ?

b) Combien d'élèves ont reçu au plus 5 spams ?

AIDE

3. Il ne faut pas confondre « plus de 2 » (c'est-à-dire 3 ou 4 ou …) et « au plus 2 » (c'est-à-dire 0 ou 1 ou 2).

51 Déterminer la médiane et les quartiles d'une série

Voici une série statistique.

Valeur	35	37	38	39	41	42	43	45
Effectif	5	5	7	10	10	2	9	5
Effectif cumulé croissant								

a) Reproduire le tableau et compléter la ligne des effectifs cumulés croissants. Quel est l'effectif total de la série ?

b) Déterminer le rang de la médiane. En déduire la valeur de la médiane.

c) Déterminer le rang du premier puis du troisième quartile.

En déduire les valeurs de Q_1 et Q_3.

AIDE

d) Le rang de Q_1 est le nombre entier qui suit immédiatement $\frac{N}{4}$. On trouve 14 pour ce rang, la valeur de Q_1 est celle de la 14^{e} donnée.

52 Interpréter médiane et quartiles

Les données ci-dessous concernent l'âge des 120 patients du service cardiologie d'un hôpital :

$$\text{min} = 20 \qquad Q_1 = 53 \qquad Q_3 = 78 \qquad \text{max} = 100$$

Pour chacune des affirmations suivantes, dire si elle est vraie ou fausse ou si on ne peut pas savoir.

a) « Trois quarts des patients ont un âge inférieur ou égal à 78 ans. »

b) « L'âge médian des patients de ce service est de 60 ans. »

c) « Au moins 91 patients ont strictement plus de 53 ans. »

d) « Au moins 30 patients de ce service ont entre 20 et 53 ans inclus. »

AIDE

b) Me n'est pas nécessairement le centre des intervalles [min ; max] ou [Q_1 ; Q_3].

53 Interpréter l'écart interquartile et l'écart-type

Voici des informations sur la taille (en cm) des élèves de Seconde d'un lycée.

	m	s	Q_1	Me	Q_3
Taille des filles	164	4	159	163	168
Taille des garçons	174	9	165	176	181

a) Calculer l'écart interquartile de chacune de ces deux séries.

b) Comparer alors la dispersion des données de ces séries autour de Me.

c) Comparer la dispersion des données de ces séries autour de m.

AIDE

Pour résumer une série on utilise les couples :
(médiane ; écart interquartile)
ou (moyenne ; écart-type).

EXERCICE RÉSOLU

54 Lire et comprendre une fonction écrite en langage Python

On étudie la fonction G écrite ci-contre en langage Python, où le paramètre « serie » est une liste.
On considère la série L = [2, 1, 6, 3].
On se propose d'exécuter pas à pas ce programme afin de déterminer le résultat renvoyé par G(L).

```
1 def G(serie):
2     N=len(serie)
3     T=0
4     for i in range(N):
5         T=T+serie[i]
6     return T
```

a) À la ligne 2, quelle est la valeur renvoyée par len(L) ?
b) À la ligne 5, que représente L[i] ?
c) Compléter le tableau ci-dessous de suivi des variables.
Quelle est alors la valeur renvoyée par G(L) ?

i	╳	0	1	2	3
T	0				

d) De façon plus générale, expliquer le rôle de cette fonction G.
e) Saisir ce programme.
Dans la console, saisir L puis exécuter G(L).
Vérifier le résultat obtenu à la question **c)**.

Solution

a) len(L) renvoie le nombre de termes de la liste L, donc $N = 4$.
b) L[i] représente le terme de rang i dans la liste L.
Ainsi $L[0] = 2$, $L[1] = 1$, $L[2] = 6$, $L[3] = 3$.
c)

i	╳	0	1	2	3
T	0	2	3	9	12

Pour $i = 0$, $T = 0 + 2 = 2$
Pour $i = 1$, $T = 2 + 1 = 3$
Pour $i = 2$, $T = 3 + 6 = 9$
Pour $i = 3$, $T = 9 + 3 = 12$

La valeur renvoyée par G(L) est 12.
d) De façon plus générale, cette fonction G renvoie la somme des termes de la liste « serie ».
e) On retrouve bien le résultat obtenu à la question **c)**.

```
>>> L=[2,1,6,3]
>>> G(L)
12
```

À VOTRE TOUR

55 **a)** Quel est le rôle de la fonction H ci-dessous, faisant appel à la fonction G de l'exercice 54 ?

```
def H(serie):
    m=G(serie)/len(serie)
    return m
```

b) Compléter le programme de l'exercice 54 par ces trois lignes puis exécuter H(L).
Quel résultat obtient-on ?
c) Exécuter H(M) pour :
M = [8, 6, 5, 8, 9, 3, 10, 5, 7, 1, 7, 9, 4, 5, 6].
Quel résultat obtient-on ?

56 **a)** Que doit-on écrire dans les cadres afin que cette fonction renvoie la moyenne de « serie » ?

```
def Moyenne(serie):
    N=[      ]
    T=0
    for i in range(N):
        T=T+[      ]
    m=[   ]
    return m
```

b) Saisir ce programme et exécuter Moyenne(M) pour :
M = [8, 6, 5, 8, 9, 3, 10, 5, 7, 1, 7, 9, 4, 5, 6].
Quel résultat obtient-on ?

EXERCICE RÉSOLU

57 Expérimenter avec le tableur

1. On effectue une série de lancers d'un dé à six faces. Les résultats figurent ci-contre dans la plage A2:B7.
On se propose de calculer certains indicateurs de cette série de lancers.

a) Réaliser cette feuille de calcul et calculer :
- l'effectif total dans la cellule B8 ;
- les produits nx dans la plage C2:C7 et leur total en C8 ;
- la moyenne m dans la cellule B10.

b) Calculer les produits $n(x-m)^2$ dans la plage D2:D7 et leur total en D8.

c) Calculer la variance V en B11 et l'écart-type s en B12.

	A	B	C	D
1	Face x	Effectif n	Produit nx	$n(x-m)^2$
2	1	1		
3	2	3		
4	3	1		
5	4	2		
6	5	0		
7	6	3		
8	Total			
9				
10	Moyenne m			
11	Variance V			
12	Ecart-type s			

2. Modifier les effectifs de la plage B2:B7 (tout en conservant dix lancers) de façon à obtenir l'écart-type :

a) le plus petit possible ;

b) le plus grand possible.

Solution

1. a) • Dans la cellule B8, on saisit $\boxed{\text{=SOMME(B2:B7)}}$, l'effectif total est 10.

• Dans la cellule C2, on saisit $\boxed{\text{=A2*B2}}$ et on recopie jusqu'en C7, puis dans la cellule C8, on saisit $\boxed{\text{=SOMME(C2:C7)}}$.

• Dans la cellule B10, on saisit $\boxed{\text{=C8/B8}}$, la moyenne de la série est 3,6.

b) Dans la cellule D2, on saisit $\boxed{\text{=B2*(A2-\$B\$10)\^{}2}}$ et dans la cellule D8, on saisit $\boxed{\text{=SOMME(D2:D7)}}$.

c) Dans B11, on saisit $\boxed{\text{=D8/B8}}$, la variance est 3,24.
Dans B12, on saisit $\boxed{\text{=RACINE(B11)}}$, l'écart-type est 1,8.

	A	B	C	D
1	Face x	Effectif n	Produit nx	$n(x-m)^2$
2	1	1	1	6,76
3	2	3	6	7,68
4	3	1	3	0,36
5	4	2	8	0,32
6	5	0	0	0
7	6	3	18	17,28
8	Total	10	36	32,4
9				
10	Moyenne m	3,6		
11	Variance V	3,24		
12	Ecart-type s	1,8		

2. a) Le plus petit écart-type que l'on peut obtenir est 0, lorsque tous les lancers sont égaux.

b) Le plus grand écart-type que l'on peut obtenir est 2,5 avec 5 lancers égaux à 1 et 5 lancers égaux à 6.

À VOTRE TOUR

58 On lance dix fois un dé à six faces et on obtient les résultats ci-contre.

	A	B
1	Face x	Effectif n
2	1	3
3	2	0
4	3	3
5	4	1
6	5	2
7	6	1

1. Réaliser une feuille de calcul permettant de déterminer la moyenne et l'écart-type de la série.

2. Quel est l'effet sur l'écart-type :

a) de l'ajout d'un 3 ?

b) de l'ajout d'un 6 ?

59 On considère la série ci-dessous.

Valeur	2	7	12	15	22	31	35	42
Effectif	3	2	4	1	5	1	6	3

1. Réaliser une feuille de calcul permettant de calculer la moyenne et l'écart-type de cette série.

2. Dans chaque cas, rechercher l'effet sur l'écart-type lorsque :

a) on ajoute 10 à chaque valeur de la série ;

b) on multiplie par 10 chaque valeur de la série.

Démontrer et raisonner

60 Découvrir un nouveau symbole

Le symbole Σ (sigma) s'utilise pour désigner la somme de plusieurs termes.

Par exemple, $\sum_{i=1}^{n} x_i = x_1 + x_2 + \ldots + x_n$.

> **Méthode**
> Pour transformer une expression telle que $\sum_{i=1}^{n} kx_i$, on peut l'écrire sans Σ et ensuite utiliser les règles de calcul algébrique.

1. Démontrer chacune des égalités suivantes où k désigne un nombre réel.

a) $\sum_{i=1}^{n} kx_i = k\sum_{i=1}^{n} x_i$ **b)** $\sum_{i=1}^{n} k = n \times k$

c) $\sum_{i=1}^{n} (x_i + y_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i$

2. Nathan affirme : « $\sum_{i=1}^{3} x_i y_i = \left(\sum_{i=1}^{3} x_i\right) \times \left(\sum_{i=1}^{3} y_i\right)$ »

A-t-il raison ? Justifier.

61 Démontrer le théorème de König-Huygens

On considère la série ci-contre.

Valeur	x_1	x_2	…	x_p
Effectif	n_1	n_2	…	n_p

On note son effectif total N, sa moyenne m et sa variance V.

On se propose de démontrer l'égalité :

$$V = \frac{1}{N}\sum_{i=1}^{n} n_i x_i^2 - m^2$$

1. a) Développer l'expression $(x_i - m)^2$.

b) En utilisant les propriétés établies à la question **1** de l'exercice **60**, montrer que :

$$\sum_{i=1}^{n} n_i (x_i - m)^2 = \sum_{i=1}^{n} n_i x_i^2 - 2m\sum_{i=1}^{n} n_i x_i + Nm^2$$

c) En déduire que :

$$\frac{1}{N}\sum_{i=1}^{n} n_i (x_i - m)^2 = \frac{1}{N}\sum_{i=1}^{n} n_i x_i^2 - \frac{2m}{N}\sum_{i=1}^{n} n_i x_i + m^2$$

2. a) Écrire m et V en utilisant le symbole Σ.

b) En déduire que $V = \frac{1}{N}\sum_{i=1}^{n} n_i x_i^2 - m^2$.

3. Calculer la variance V de la série ci-dessous à l'aide du théorème de König-Huygens.

x_i	14	16	17	18	19	21
n_i	2	1	3	4	5	5

Utiliser des indicateurs statistiques

62 Dans une classe, il y a 7 internes, 12 demi-pensionnaires et 10 externes.
La moyenne en mathématiques du premier trimestre est 11,2 pour les internes, 10,8 pour les demi-pensionnaires et 10,4 pour les externes.
Calculer la moyenne de la classe. *Arrondir au dixième.*

63 Une entreprise produit des dosettes de café. Une dosette doit théoriquement avoir une masse de 8 g.

Pour vérifier la qualité de la production, on prélève un échantillon de 50 dosettes et on relève la masse réelle de chacune.

7,8	8,3	8,5	7,9	8,5	8,2	8,1	8,1	8,3	7,6
7,8	8	7,6	7,9	8,3	8,5	7,7	8,2	8	8,4
8,2	8,1	8	8,3	7,7	7,9	8,3	8,5	7,8	8,5
8	8	7,7	7,7	8,5	8	7,7	8,1	8,4	8,1
7,8	8,3	8,4	7,9	7,7	8,2	8,1	7,6	8,4	8

1. a) Calculer la moyenne de cet échantillon.

b) Déterminer l'écart-type s de cet échantillon.

2. La qualité de la production est jugée satisfaisante si les trois conditions suivantes sont satisfaites :

- la masse moyenne d'une dosette de l'échantillon diffère de la masse théorique de moins de 0,1 g ;
- au moins 95 % des données de l'échantillon appartiennent à l'intervalle $[8 - 2s\,;\,8 + 2s]$;
- au moins 65 % des données de l'échantillon appartiennent à l'intervalle $[8 - s\,;\,8 + s]$.

À partir de cet échantillon, la qualité de la production paraît-elle satisfaisante ?

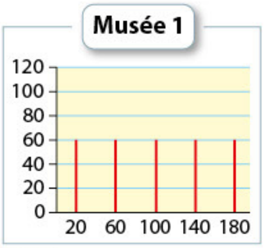
64 Une entreprise a effectué une étude sur la durée de réalisation d'un projet par ses employés.

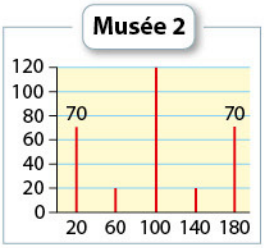
a) Déterminer la moyenne m et l'écart-type s de cette série. *Arrondir au centième.*

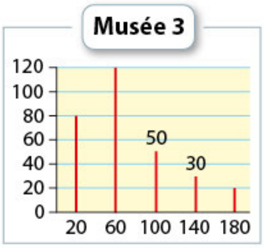
b) Déterminer l'intervalle $[m - 2s\,;\,m + 2s]$ et calculer la proportion en pourcentage de projets dont la durée de réalisation est dans cet intervalle.
Arrondir au dixième.

Nombre de jours	Nombre de projets
15	3
17	12
20	30
22	61
23	76
25	95
28	81
29	63
32	35
33	10
35	5

65 Ces diagrammes en bâtons illustrent la fréquentation de trois musées sur 300 jours de l'année. Le nombre de visiteurs est indiqué en abscisse et le nombre de jours en ordonnée.







1. Pour chaque musée, déterminer :

a) la médiane et l'écart interquartile ;

b) la moyenne et l'écart-type.

2. Indiquer le couple d'indicateurs qui semble le plus approprié pour résumer chaque série.

Pour les exercices 66 à 68, on utilise la fonction écrite ci-dessous en langage Python qui renvoie la moyenne de la liste « serie ».

```
def Moyenne(serie):
    N=len(serie)
    T=0
    for i in range(N):
        T=T+serie[i]
    m=T/N
    return m
```

66 Algo Voici une fonction écrite en langage Python qui renvoie l'écart-type de la liste « serie ».

```
from math import *

def Ecarttype(serie):
    N=len(serie)
    m=Moyenne(serie)
    SC=0
    for i in range(N):
        SC=SC+(serie[i]-m)**2
    s=sqrt(SC/N)
    return s
```

a) Que représente chacune des variables SC et N ?

b) Saisir ce programme et exécuter Ecarttype(M), pour la série :

$$M = [1, 3, 4, 4, 5, 6, 10]$$

Quel résultat obtient-on ?

67 Algo L'exercice utilise la fonction Ecarttype et la liste M de l'exercice 66.

a) Que représente C dans la fonction suivante ?

```
def Proportion(serie):
    N=len(serie)
    m=Moyenne(serie)
    s=Ecarttype(serie)
    C=0
    for i in range(N):
        if m-2*s<=serie[i]<=m+2*s:
            C=C+1
    p=C/N
    return p
```

b) Définir le rôle de la fonction Proportion.

c) Saisir ce programme et exécuter Proportion(M). Quel résultat obtient-on ?

68 Algo Voici un programme écrit en langage Python.

```
from random import *

def Lancer(n):
    L=[]
    for i in range(n):
        L=L+[randint(1,6)]
    return L
```

a) Quel est le rôle de la fonction Lancer ?

b) Saisir le programme et exécuter Lancer(100). On note A la série renvoyée.

c) Déterminer la moyenne et l'écart-type de la série A. Comparer les résultats obtenus avec ceux d'autres élèves.

69 Algo Voici un programme écrit en langage Python.

```
from random import *

def SommeDes(n):
    L=[]
    for i in range(n):
        X=randint(1,6)+randint(1,6)
        L=L+[X]
    return L
```

a) Quel est le rôle de la fonction SommeDes ?

b) Saisir le programme et exécuter SommeDes(100). On note B la série renvoyée.

c) Déterminer la proportion en pourcentage des valeurs de la série B appartenant à l'intervalle $[m - 2s \,;\, m + 2s]$, où m est la moyenne et s est l'écart-type de la série B.

Comparer le résultat obtenu avec ceux d'autres élèves.

Comparer deux séries

70 Un fabricant de macarons teste une machine permettant de déposer sur les plaques de cuisson les disques qui formeront les gâteaux.

Ces disques, après cuisson, doivent avoir 3 cm de diamètre pour satisfaire le bon conditionnement des macarons.
Voici les diamètres en cm des disques après cuisson pour deux fournées : l'une réalisée à l'aide de la machine, l'autre réalisée manuellement.

Fournée 1

2,8	2,7	2,7	2,8	3,2	3,9	4	3,1
3	4	3	3,1	2,8	3,5	3	3,1
2,1	2,2	3	3,4	3	2,5	2,4	2,9

Fournée 2

3,1	2,7	3,2	2,8	2,9	3	3	2,8
2,9	3	2,7	3,2	3,1	3,1	3,4	3
2,7	2,9	3	3,5	2,9	2,9	2,7	3

a) Calculer la moyenne et l'écart-type de chaque série. *Arrondir au centième.*
b) Quelle série semble correspondre à celle réalisée à l'aide de la machine ?

71 On a prélevé des échantillons de 60 pots chez deux producteurs de confiture d'abricot.
L'un de ces producteurs utilise des méthodes artisanales qui donnent des taux de sucre, en pourcentage, assez hétérogènes. L'autre producteur utilise des procédés industriels qui assurent des taux de sucre, en pourcentage, plus homogènes.

Série 1

Taux de sucre (%)	55	56	57	58	59	60	61	62	63	64	65
Effectif	0	0	1	3	15	18	16	5	2	0	0

Série 2

Taux de sucre (%)	55	56	57	58	59	60	61	62	63	64	65
Effectif	1	2	2	5	9	12	11	9	6	2	1

a) Calculer la moyenne et l'écart-type de chaque série. *Arrondir au centième.*
b) Associer chaque série au producteur correspondant.
c) Pour chaque série, calculer le pourcentage de données appartenant à l'intervalle $[m-s\ ;\ m+s]$ puis à l'intervalle $[m-2s\ ;\ m+2s]$.

72 On a relevé les distances parcourues par deux joueurs de football durant le dernier championnat.

Joueur 1

Distance (en km)	8	8,5	9	9,5	10	10,5
Nombre de matchs	8	7	3	10	2	8

Joueur 2

Distance (en km)	8	8,5	9	9,5	10	10,5
Nombre de matchs	0	9	14	7	7	1

a) Déterminer la médiane et les quartiles de chacune des deux séries.
b) En quoi ces joueurs se distinguent-ils ?

S'entraîner à la logique → p. XXI

73 Quantificateurs (1)

Dans chaque cas, dire si la proposition est vraie ou fausse. Justifier.
a) Il existe une série statistique de moyenne et médiane égales.
b) Quelle que soit la série statistique, la médiane appartient à l'intervalle $[Q_1\ ;\ Q_3]$.

74 Quantificateurs (2)

On considère la série de notes :
7 – 8 – 10 – 10 – 11 – 13 – 14 – 14 – 16 – 20
On note m sa moyenne et s son écart-type.
Dans chaque cas, dire si la proposition est vraie ou fausse. Justifier la réponse.
a) Il existe une valeur a de la série telle que :
$$a > m + s$$
b) Pour toute valeur b de la série, $|b - m| < 2s$.

75 Implication réciproque

v est une valeur d'une série statistique dont on connaît deux indicateurs : $\mathrm{Me} = 14$ et $Q_1 = 9$.
On considère les propositions suivantes :
• A : « $v > 14$ » • B : « $v \geqslant 9$ »
• C : « Au moins 50 % des valeurs de la série sont inférieures ou égales à v. »
• D : « Moins de 75 % des valeurs de la série sont strictement supérieures à v. »
1. Pour chacune des propositions suivantes, dire si elle est vraie ou fausse.
a) A implique C. **b)** A implique D.
c) B implique D.
2. Envisager la réciproque de chacune des propositions ci-dessus et dire si elle est vraie ou fausse.

76 Retrouver une donnée manquante

Problème ouvert

Raisonner **Calculer**

Voici une série statistique où x désigne un nombre réel.

Valeur	8	10	15	20	x	40
Effectif	2	2	1	3	1	1

On augmente chaque valeur de la série de 40 % et on calcule la moyenne de la série obtenue.
On obtient 25,62.
Quelle est la valeur manquante x de la série initiale ?

77 Découvrir un indicateur

Calculer **Communiquer**

Ce tableau donne le débit mensuel moyen (en $m^3 \cdot s^{-1}$) de la Seine à Paris et du Rhône à Beaucaire.

Mois	Seine	Rhône	Mois	Seine	Rhône
Janvier	510	2 296	Juillet	112	1 230
Février	545	2 050	Août	94	1 148
Mars	445	2 280	Septembre	99	1 427
Avril	232	1 673	Octobre	124	1 066
Mai	229	1 668	Novembre	244	1 591
Juin	157	1 558	Décembre	309	1 378

1. a) Pour chacune des deux séries, calculer la moyenne et l'écart-type. *Arrondir au dixième.*
b) Quel fleuve semble le plus capricieux ?
2. a) Relativiser la valeur de l'écart-type s à la moyenne m en calculant le coefficient de variation $\frac{s}{m}$ de chaque série.
b) Que peut-on déduire de la comparaison de ces coefficients ?

78 Prendre des initiatives

Calculer **Communiquer**

Lors d'une étude clinique, on a constitué deux groupes, A et B, de patients de plus de 30 ans.
On a donné un médicament contre le cholestérol au groupe A et un placebo au groupe B.
Voici les résultats d'analyse du taux de cholestérol en $g \cdot L^{-1}$ pour chaque groupe.

Groupe A

2,4	2,2	2,1	2,4	2,7	3	2,1	2,3
2,7	2,2	2	2,1	2,4	2,2	3	3,2
2,1	2,5	2,5	3	2,9	2,9	2,4	2,3

Groupe B

2,4	2,7	2,3	2,4	2,8	3,2	2,8	3
2,5	3	3,2	3,1	2,8	3	2,5	2,5
2,4	2,5	3	3,1	3,1	2,9	2,7	2,7

Comparer l'efficacité du médicament à celle du placebo.

79 Réduire les disparités

Raisonner **Calculer**

On donne ci-contre la répartition des salaires mensuels bruts (en euro) au sein d'une entreprise.
Le PDG souhaite valoriser le travail de ses employés en augmentant les salaires tout en réduisant les disparités.
Pour cela, il hésite entre deux méthodes.

Salaire	Effectif
[1200 ; 1500[	201
[1500 ; 1700[	213
[1700 ; 2000[	350
[2000 ; 2500[	500
[2500 ; 3000[	233
[3000 ; 3500[	51
[3500 ; 7000[	9

Méthode 1 : réduire de 3 % chaque salaire, puis ajouter 100 €.
Méthode 2 : faire évoluer 10 salariés de chaque classe (sauf la dernière) à celle supérieure.
Étudier les conséquences de chaque méthode :
- sur les salaires moyens ;
- sur les écarts-types.

80 Algo Écrire une fonction en langage Python

Modéliser **Raisonner**

Voici un programme écrit en langage Python.

```
def Minimum(L):
    N=len(L)
    mini=L[0]
    for i in range(1,N):
        if L[i]<mini:
            mini=L[i]
    return mini
```

1. a) On considère la série F = [5, 20, 3, 12].
Quelles sont les valeurs successives prises par la variable mini lorsqu'on exécute Minimum(F) ?
b) Quel est le rôle de la fonction Minimum ?
2. Écrire une fonction Maximum qui renvoie le maximum d'une série.
3. Écrire une fonction Etendue qui renvoie l'étendue d'une série.

81 Déterminer des effectifs

Narration de recherche

Chercher **Communiquer**

Rédiger les différentes étapes de la recherche, sans omettre les fausses pistes et les changements de méthode.

Problème

Valeur	−1	0	1	3	5	Total
Effectif	…	…	…	…	…	25

Reproduire et compléter ce tableau afin que :
- la moyenne soit égale à 2 ;
- la médiane soit égale à 1 ;
- les quartiles soient $Q_1 = 0$ et $Q_3 = 3$.

82 Study the Brexit referendum

Calculer **Communiquer**

In 2016, British people were asked: "Should the United Kingdom remain a member of the European Union or leave the European Union?"
Here are the results of this referendum.

Country	Valid votes	Leave	Remain
England	28 455 402	53,38 %	46,62 %
Northern Ireland	790 149	44,22 %	55,78 %
Scotland	2 679 513	38,00 %	62,00 %
Wales	1 626 919	52,53 %	47,47 %

Calculate the results, as a percentage, of this referendum for the United Kingdom.

83 Découvrir de nouveaux quantiles

Raisonner **Calculer**

Définition : Pour tout nombre entier i compris entre 1 et 9, le décile D_i d'une série ordonnée est la plus petite valeur de la série telle qu'au moins $(10 \times i)$ % des valeurs sont inférieures ou égales à D_i.

Le tableau ci-contre donne différents indicateurs sur la répartition des salaires mensuels nets, en euro, des hommes et des femmes en France, en 2015.

Déciles	Hommes	Femmes
D_1	1 262	1 171
D_2	1 427	1 288
D_3	1 573	1 396
D_4	1 728	1 512
D_5	1 906	1 650
D_6	2 130	1 830
D_7	2 451	2 073
D_8	2 996	2 432
D_9	3 990	3 149

1. Pour chacune des affirmations suivantes, indiquer si elle est vraie ou fausse.
Justifier

a) Au moins 90 % des hommes gagnent 3 990 € ou moins.

b) La médiane de la série des salaires des hommes est supérieure ou égale à 1 906 €.

c) Le troisième quartile de la série des salaires des femmes est égal à 2 450 €.

d) Moins de 20 % des hommes et plus de 30 % des femmes gagnent moins de 1 400 €.

2. a) Calculer, pour chaque série, la valeur des rapports $\frac{D_9}{D_1}$, $\frac{\text{Me}}{D_1}$ et $\frac{D_9}{\text{Me}}$.

b) Interpréter les résultats.

84 Étudier une transformation

Chercher **Calculer**

Ce tableau donne la répartition des salaires, en euro, dans une entreprise.

Salaire x_i	Effectif n_i
$[1000 ; 1500[$	300
$[1500 ; 2000[$	55
$[2000 ; 3500[$	35
$[3500 ; 4500[$	10

Suite à des négociations, le nouveau salaire est donné par la formule $y_i = 0{,}98x_i + 90$.
Étudier les conséquences pour les salariés :

- sur les salaires moyen et médian ;
- sur l'évolution des salaires en pourcentage ;
- sur les indicateurs de dispersion.

85 Imaginer une stratégie

Raisonner **Calculer**

Multiplier chaque valeur de cette série par un même nombre a, puis leur ajouter un même nombre b afin que la moyenne soit égale à 10 et l'écart-type égal à 3 :

30 – 45 – 45 – 45 – 52,5 – 52,5

Des Défis

86 Calculer des moyennes

Hugo est pâtissier, il fabrique et vend des mille-feuilles. En 2018, à partir du 1^er^ avril, le nombre de mille-feuilles vendus chaque mois est la moyenne des nombres vendus au cours des trois mois précédents.
Il en a vendu 2017 en octobre, 2018 en novembre et 2019 en décembre.
Combien en a-t-il vendu au total au cours de l'année 2018 ?

(d'après Rallye de Bourgogne)

87 Calculer l'âge de la professeure

Dans une classe de lycée, l'âge moyen, en comptant la professeure de mathématiques, est 18,193.
Sans elle, l'âge moyen des élèves est 16,966.
Les âges sont tous des nombres entiers mais les moyennes sont données à l'aide d'arrondis au millième.
Quel est l'âge de la professeure de mathématiques ?

(d'après Rallye de Bourgogne 2013)

QCM interactifs

QCM Bilan

88 Dans chaque cas, donner la réponse exacte en justifiant.

Pour les questions 1 à 5, on considère la série ci-contre.
On note Me sa médiane, Q_1 et Q_3 ses quartiles.

		A	B	C	D
1	Me est telle que …	Me = 12,5	Me = 12	si la valeur 15 devient 17, Me change	si on ajoute une valeur quelconque à la série, Me change
2	Me est …	$Q_3 - Q_1$	comprise entre Q_1 et Q_3	égale à 50 %	égale à $Q_1 + 0{,}25$
3	La moyenne m de cette série est telle que …	$m = 12$	$m = \dfrac{9+10+12+13+14+15}{6}$	si la valeur 9 devient 7, m ne change pas	$m >$ Me
4	La proportion des valeurs de la série strictement inférieures à Q_3 est …	25 %	strictement inférieure à 50 %	strictement inférieure à 75 %	au moins 75 %
5	L'écart interquartile est …	supérieur à l'étendue	égal à 4	inférieur à l'écart-type	égal à 50 %

Pour les questions 6 à 8 :
- on note m_1 et s_1 la moyenne et l'écart-type de la série K : 10 ; 15 ; 21 ; 34 ; 51 ; 59 ;
- on note m_2 et s_2 la moyenne et l'écart-type de la série L : 1 000 ; 1 500 ; 2 100 ; 3 400 ; 5 100 ; 5 900.

		A	B	C	D
6	On peut affirmer que …	$m_1 = m_2$	$m_1 + m_2 = 100$	$100m_1 = m_2$	$m_1 + 100 = m_2$
7	On peut affirmer que …	$s_1 = s_2$	$s_1 = 10s_2$	$10s_1 = s_2$	$100s_1 = s_2$
8	Si on multiplie par 2 chaque valeur de la série K, alors …	le troisième quartile ne change pas	l'écart interquartile est multiplié par 2	l'étendue est multipliée par 4	la médiane ne change pas

Pour les questions 9 et 10, on lance un dé cubique dont les faces sont numérotées de 1 à 6.
On note m la moyenne et s l'écart-type de la série des lancers.

		A	B	C	D
9	Alors …	s peut être strictement négatif	s peut être égal à 0	s peut être supérieur à m	l'intervalle $[m-2s\ ;\ m+2s]$ ne peut pas contenir 100 % des valeurs
10	Alors …	l'étendue de la série peut être strictement supérieure à 5	s peut être supérieur à 5	m peut être égale à 3	la médiane de la série ne peut pas être égale à 3,5

Vérifiez vos réponses p. 346 pour avoir votre note (considérez 1 point par réponse juste).

Exploiter ses compétences

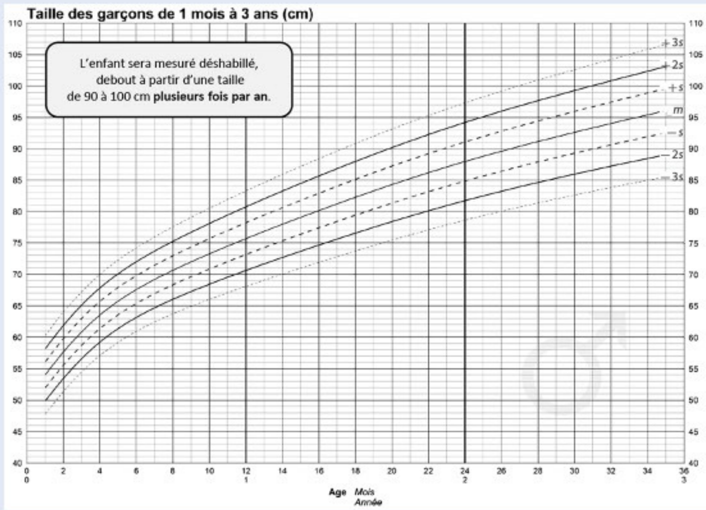
89 Courbes de croissances

La situation problème

La surveillance de la croissance physique d'un enfant se fait grâce, notamment, aux courbes de taille qui sont dans son carnet de santé. Elles permettent aux parents et au médecin de faire un suivi individuel de l'enfant.

Utiliser les différentes informations pour savoir si la proportion de garçons ayant une taille dans l'intervalle $[m-2s\ ;\ m+2s]$ varie en fonction de l'âge.

DOC 1 Courbes de croissance



DOC 2 Répartition (en pourcentage) des garçons par taille, à différents âges

Âge	Taille (en cm) : $[68 ; 70[$	$[70 ; 72[$	$[72 ; 74[$	$[74 ; 76[$	$[76 ; 78[$	$[78 ; 80[$	$[80 ; 82[$	$[82 ; 84[$	$[84 ; 86[$	$[86 ; 88[$	$[88 ; 90[$	$[90 ; 92[$	$[92 ; 94[$	$[94 ; 96[$	$[96 ; 98[$	$[98 ; 100[$	$[100 ; 102[$	$[102 ; 104[$	$[104 ; 106[$
12 mois	1	6	17	28	31	13	3	1	0	0	0	0	0	0	0	0	0	0	0
18 mois	0	0	0	1	6	13	28	25	17	8	2	0	0	0	0	0	0	0	0
24 mois	0	0	0	0	0	1	2	8	15	24	26	14	7	2	1	0	0	0	0
30 mois	0	0	0	0	0	0	0	0	2	12	9	19	25	17	11	4	1	0	0
36 mois	0	0	0	0	0	0	0	0	0	1	2	5	13	19	22	21	10	5	2

90 Libérez Caroline !

La situation problème

Caroline aime jouer au jeu de plateau ForMath. Au cours d'une partie, un joueur peut être envoyé en prison. Caroline se demande combien de lancers sont nécessaires à un joueur prisonnier pour se libérer.

Utiliser les différentes informations pour aider Caroline à étudier le nombre de lancers nécessaires.

DOC 1 Règle du jeu ForMath

Un joueur reste prisonnier tant qu'il n'a pas obtenu un 6 au lancer de dé.

DOC 2 Deux fonctions Python

```python
from random import *

def NombreLancers():
    Nb=1
    De=randint(1,6)
    while De!=6:
        Nb=Nb+1
        De=randint(1,6)
    return Nb

def Serie(N):
    L=[]
    for i in range(N):
        L=L+[NombreLancers()]
    return L
```

91 Le prix de l'électricité dans l'Union européenne

La situation problème

Utiliser les différentes informations pour étudier l'impact des impôts et taxes sur la dispersion des tarifs de l'électricité au sein de l'Union européenne.

DOC 1 La formation des prix

Le prix de vente Toutes Taxes Comprises (TTC) de l'électricité est obtenu en additionnant le prix de vente sans impôts, la Taxe sur la Valeur Ajoutée (TVA) et le montant des autres impôts.

DOC 2 Les tarifs d'électricité dans l'Union européenne

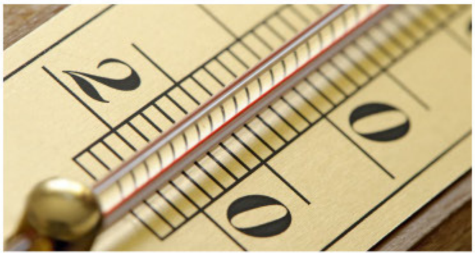
en centimes d'euro par kWh	Prix de vente sans impôts	TVA	Autres impôts
Lituanie	7,77	1,92	1,38
Bulgarie	8,19	1,64	0,00
Slovaquie	8,38	2,40	3,64
Hongrie	8,93	2,41	0,00
Danemark	9,22	6,02	14,86
Roumanie	9,45	2,06	1,38
Estonie	9,50	2,20	1,49
Pologne	9,50	2,71	2,30
Croatie	10,12	1,42	0,82
Lettonie	10,40	2,75	2,67
Finlande	10,64	3,10	2,25
Portugal	10,80	4,17	7,33
Grèce	10,90	1,60	3,70
Slovénie	11,05	2,91	2,17
France	11,32	2,58	3,66
Pays-Bas	11,52	2,70	1,34
Luxembourg	11,70	1,20	3,28
Autriche	12,18	3,29	4,31
Rép. tchèque	12,18	2,59	0,11
Malte	12,99	0,65	0,00
Suède	13,05	3,99	2,89
Italie	13,26	1,90	5,64
Royaume-Uni	13,44	0,88	4,24
Allemagne	13,83	4,87	11,78
Chypre	14,20	2,76	1,30
Espagne	17,12	3,78	0,87
Belgique	17,90	4,71	6,16
Irlande	18,65	2,80	2,10

92 Évolution du climat en France

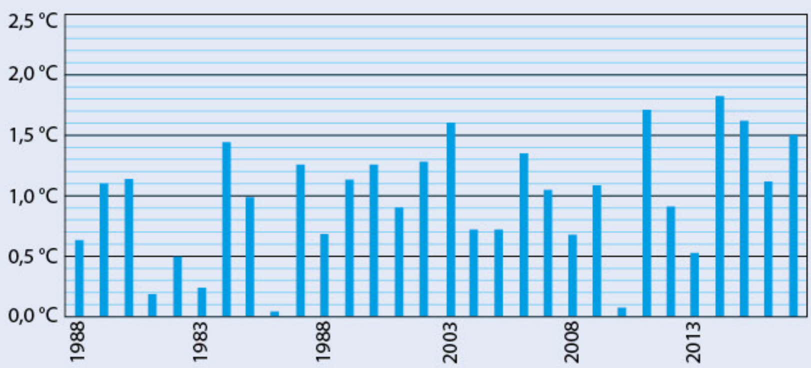
La situation problème

On étudie l'évolution de la température moyenne en France métropolitaine depuis 1900. La température moyenne de référence est 11,8 °C (moyenne sur la période 1961 - 1990).

Utiliser les différentes informations pour comparer les températures de la période 1988-2017 avec celles de la période 1900-1987.



DOC 1 Évolution des écarts par rapport à la température moyenne de référence depuis 1988



DOC 2 Indicateurs

Voici quelques indicateurs statistiques de la série des écarts par rapport à la température de référence sur la période 1900-1987.

Min	– 1,2 °C
Q_1	– 0,4 °C
Med	– 0,2 °C
Q_3	– 0,1 °C
Max	0,8 °C