

Statistiques descriptives

1. Vocabulaire statistique

Pour pratiquer les sciences statistiques, on doit collecter un grand nombre de données. Pour cela, on effectue une **enquête statistique** sur un ensemble de personnes ou d'objets.

■ DÉFINITION : Population

L'ensemble sur lequel cette enquête est effectuée est appelé la **population**.

■ DÉFINITION : Individu

Chaque élément de cette population est appelé un **individu**.

Il est rare de pouvoir interroger tous les individus d'une population, donc on en interroge un sous-ensemble, qui permet une estimation des résultats de l'enquête.

■ DÉFINITION : Échantillon

Un tel sous-ensemble est appelé un **échantillon**.

On s'intéresse à l'ensemble des résultats d'une enquête statistique pour en tirer des informations.

■ DÉFINITION : Caractère

La question posée lors d'une enquête statistique porte sur une particularité de l'ensemble étudié, appelée le **caractère**.

■ DÉFINITION : Valeurs du caractère

Les différentes réponses possibles sont les **valeurs du caractère**. On les note x_i .



Attention

Ces « valeurs » ne sont pas forcément numériques !

On peut regrouper le vocabulaire dans un tableau :

Vocabulaire ensembliste	Vocabulaire statistique	Vocabulaire probabiliste
Ensemble	Population	Ensemble des issues
Sous-ensemble	Échantillon	Événement
Élément	Individu	Issue
Propriété	Caractère	×

Exemple :

On interroge tous les élèves de troisième du lycée franco-allemand de Buc.

- Population : l'ensemble des élèves de troisième.
- Individu : un élève de troisième.
- Échantillon : la classe de troisième C.
- Caractère : le nombre de frères et sœurs.
- Valeurs du caractère : 0, 1, 2, 3, ...

Il existe différents types de caractères :

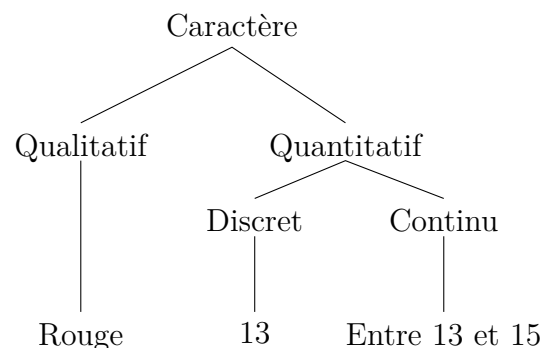
■ DÉFINITION : Caractères quantitatifs

Le caractère est dit **quantitatif** si on peut le mesurer, en associant un nombre à chaque valeur du caractère de la population. Il y a alors deux types de caractères quantitatifs possibles :

- On dit qu'il est quantitatif **discret** s'il peut prendre un nombre fini de valeurs (nombre de frères et sœurs...)
- On dit qu'il est quantitatif **continu** s'il peut prendre toutes les valeurs d'un intervalle de façon continue (taille, poids...)

■ DÉFINITION : Caractères qualitatifs

Le caractère est dit **qualitatif** si on ne peut pas le mesurer (couleur des yeux...).



L'intérêt des statistiques descriptives est de tirer des informations de l'ensemble des valeurs collectées.

■ DÉFINITION : Série statistique

On appelle **série statistique** l'ensemble des valeurs collectées.

■ DÉFINITION : Effectif

Le nombre de fois où apparaît une réponse est appelé l'**effectif** de la valeur du caractère. Les effectifs sont notés n_i .

■ DÉFINITION : Fréquence

La **fréquence** d'une valeur du caractère est le quotient de l'effectif de la valeur par l'effectif total. Elle peut s'exprimer en pourcentage comme l'exprime la propriété suivante.

■ PROPRIÉTÉ : (admise)

On a

$$f_i = \frac{\text{effectif de la valeur}}{\text{effectif total}}$$

avec $0 \leq f_i \leq 1$

En pourcentages,

$$p_i = \frac{\text{effectif de la valeur}}{\text{effectif total}} \times 100$$

avec $0 \leq p_i \leq 100$

Remarque : La somme des fréquences de toutes les valeurs de la série statistique est 1.

2. Caractères qualitatifs

Les séries statistiques peuvent être représentées grâce à des diagrammes circulaires, semi-circulaires, rectangles... C'est bien là l'unique information que l'on peut obtenir dans le cas des caractères qualitatifs. L'aire de chaque valeur du caractère devra être proportionnelle à l'effectif de cette valeur. Les fréquences permettent d'obtenir sans difficulté la part du diagramme qui devra être consacrée à chaque valeur du caractère.

On considère les données statistiques suivantes :

- Population : Élèves du lycée franco-allemand
- Échantillon : Classe de 3^e
- Caractère : Couleurs des cheveux

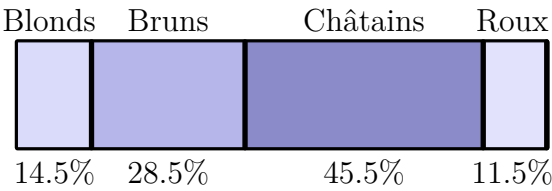
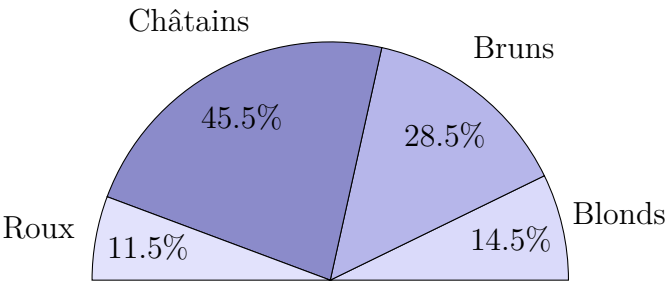
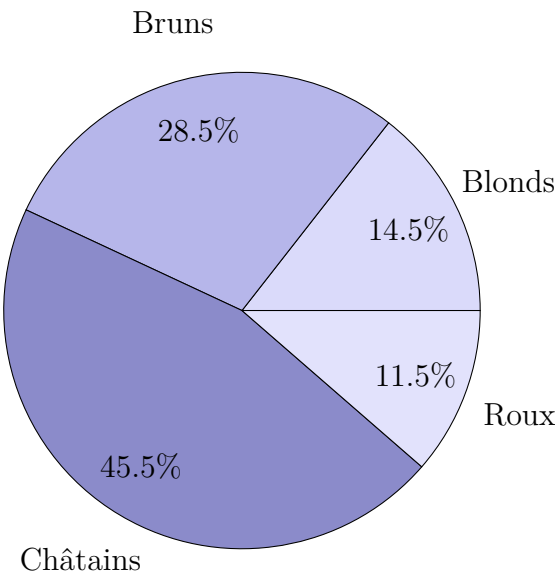
Après une enquête, on a obtenu la série statistique suivante.

Valeurs du caractère, x_i	Blonds	Bruns	Châtains	Roux	Autres	Total
Effectifs, n_i	29	57	91	23	0	200

Compléter le tableau suivant :

Valeurs du caractère, x_i	Blonds	Bruns	Châtains	Roux	Total
Effectifs, n_i	29	57	91	23	200
Fréquences f_i	$\frac{29}{200} = 0.145$	$\frac{57}{200} = 0.285$	$\frac{91}{200} = 0.455$	$\frac{23}{200} = 0.115$	$\frac{200}{200} = 1.0$
Part d'un diagramme circulaire $f_i \times 360^\circ$	52.2°	102.6°	163.8°	41.4°	360.0°
Part d'un diagramme semi-circulaire $f_i \times 180^\circ$	26.1°	51.3°	81.9°	20.7°	180.0°
Part d'un rectangle de 10 cm $f_i \times 10$	1.45 cm	2.85 cm	4.55 cm	1.15 cm	10.0 cm

On obtient les diagrammes suivants :



3. Caractères quantitatifs

A. Caractères quantitatifs discrets

C'est dans le cas des caractères quantitatifs discrets que l'on peut obtenir le plus d'informations. Nous allons définir les caractéristiques que vous devez savoir donner. On travaillera systématiquement avec l'exemple suivant :

- Population : Élèves du lycée franco-allemand
- Échantillon : Classe de première ES
- Caractère : Nombre de frères et sœurs

Comme d'habitude, on écrit la série statistique sous forme d'un tableau :

Nombre de frères et sœurs, noté x_i	Nombre d'élèves, notés n_i	Fréquences en %, notées p_i	Effectifs cumulés croissants	Produits $n_i \times x_i$
0	2	6.7	2	0
1	11	36.7	13	11
2	11	36.7	24	22
3	3	10.0	27	9
4	1	3.3	28	4
5	2	6.7	30	10
Total	30	100.0	×	56

Vous pouvez compléter le tableau de l'exemple précédent par une colonne dans laquelle vous calculez la fréquence, exprimée en pourcentages. On notera cette colonne les p_i , et on utilise la relation donnée dans l'introduction.

a. Étendue

■ DÉFINITION : Étendue

L'**étendue** d'une série statistique est le nombre positif égal à la différence entre la plus grande valeur du caractère et la plus petite valeur du caractère.

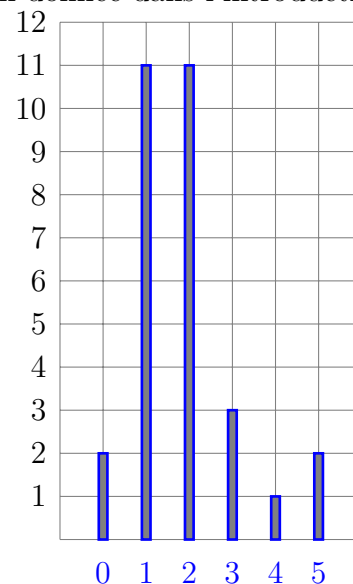
Exemple :

Dans l'exemple précédent, l'étendue est $5 - 0 = 5$.

b. Diagramme en bâtons des effectifs

Il est possible de construire le diagramme en bâtons des effectifs. Avec l'exemple précédent, on obtient le diagramme ci-contre.

On remarque que la hauteur des bâtons est égale ou proportionnelle à l'effectif.



c. Mode

■ DÉFINITION : Mode

On appelle **mode** d'une série statistique une valeur du caractère qui a le plus grand effectif.

Exemple :

Dans l'exemple précédent, les modes sont 1 et 2.

d. Effectifs cumulés croissants

Les effectifs cumulés croissants sont un outil qui permet de répondre à certaines questions. Il convient d'apprendre à les calculer pour compléter notre exemple précédent. Pour cela, nous allons d'abord travailler sur l'exemple suivant :

Exemple :

On considère les 150 notes obtenues par les élèves de cinq classes de première lors d'un contrôle.

Pour connaître le nombre d'élèves qui ont obtenu une note inférieure ou égale à une certaine note, on complète le tableau des effectifs par une ligne dans laquelle on indique les effectifs cumulés croissants.

Notes	6	8	9	10	11	12	14	15	16
Effectifs	3	12	18	24	30	36	12	9	6
Effectifs cumulés croissants	3	$3 + 12 = 15$	$3 + 12 + 18 = 33$ ou encore $15 + 18 = 33$	57	87	123	135	144	150

Par exemple, l'effectif cumulé de la note 8 est la somme des effectifs de la note 6 et de la note 8. De même, l'effectif cumulé de la note 9 est la somme des effectifs de la note 6, de la note 8 et de la note 9, etc... Ainsi, on sait que 33 élèves ont obtenu une note inférieure ou égale à 9.

1. Compléter le tableau en remplissant la ligne des effectifs cumulés croissants.
2. Utilisez le tableau pour répondre aux questions suivantes :
 - Combien d'élèves ont obtenu une note strictement inférieure à 10 ? 33
inférieure ou égale à 12 ? 123
 - 144 élèves ont obtenu une note inférieure ou égale à une certaine note.
Quelle est cette note ? 15
 - Combien d'élèves ont obtenu une note supérieure ou égale à 9 ? $135 = 150 - 15$

Exemple :

Dans l'exemple des frères et sœurs, complétez dans le tableau la colonne des effectifs cumulés croissants pour répondre à la questions :

- Combien d'élèves ont moins de 2 frères et sœurs ? 17, c'est-à-dire $30 - 13$
- Combien d'élèves ont au plus 4 frères et sœurs ? 28, c'est-à-dire $30 - 2$
- Combien d'élèves ont plus de 3 frères et sœurs ? 3, c'est-à-dire $30 - 27$.

e. Moyenne

■ DÉFINITION :

La **moyenne** d'une série statistique est égale au quotient de la somme de ses valeurs par l'effectif total.

Exemple :

Mélanie et Julie lancent deux dés et notent le total des points obtenus :

10 ; 12 ; 5 ; 2 ; 3 ; 6 ; 7 ; 11

Calculer la moyenne des points obtenus. $\frac{10 + 12 + 5 + 2 + 3 + 6 + 7 + 11}{8} = \frac{56}{8} = 7$

■ PROPRIÉTÉ :

La moyenne est la note qui, substituée à chaque valeur du caractère de la série, redonne la même somme.

Exemple :

Lily joue au Volley-ball. Elle a participé à tous les matchs de la saison. La capitaine de son équipe a noté le nombre de points marqués par Lily lors de chaque match :

3 - 2 - 5 - 5 - 6 - 0 - 3 - 4 - 3 - 6 - 6 - 4 - 2 - 5 - 4 - 4 - 4 - 6 - 4 - 4

- Quel(le) est...
 - la population étudiée ? Les matchs de Voley-ball auxquels Lily a participé.
 - le caractère étudié ? Les points marqués par Lily.
 - l'effectif total de la série ? 20
 - l'étendue de la série ? $6 - 0 = 6$
- Compléter le tableau suivant :

Nombre de buts	0	1	2	3	4	5	6
Effectif	1	0	2	3	7	3	4
Effectifs cumulés croissants	1	1	3	6	13	16	20

- Lily fait le calcul :

$$\frac{0 + 1 + 2 + 3 + 4 + 5 + 6}{7} = 3$$

Expliquer pourquoi ce nombre n'est pas le nombre moyen de points de Lily.

Lily a marqué plus souvent certaines quantités de points que d'autres.

- Une autre joueuse effectue le calcul ci-dessous.

$$\frac{0 + 2 + 2 + 3 + 3 + 3 + 4 + 4 + 4 + 4 + 4 + 4 + 4 + 5 + 5 + 5 + 6 + 6 + 6 + 6}{20}$$

Combien trouve-t-elle ? $\frac{80}{20} = 4$.

Que représente ce nombre ? Ce nombre signifie que Lily a marqué en moyenne 4 points lors des matchs de Volley-ball de la saison.

- La capitaine de l'équipe calcule plus simplement cette moyenne. Comment a-t-elle fait ?

$$\frac{0 \times 1 + 2 \times 2 + 3 \times 3 + 4 \times 7 + 5 \times 3 + 6 \times 4}{20} = \frac{80}{20} = 4$$

■ DÉFINITION :

La **moyenne pondérée** d'une série statistique est égale à la somme des produits de chaque valeur du caractère par son effectif, divisée par l'effectif total.

Exemple :

Dans l'exemple des frères et sœurs, compléter la dernière colonne du tableau en effectuant le produit de chaque valeur du caractère par son effectif, puis calculer le nombre moyen de frères et sœurs des élèves de la classe de 1^{ère} ES interrogée. On divise 56 par 30 pour obtenir 1,87.

Remarque : Une moyenne n'est pas nécessairement égale à l'une des valeurs de la série. Elle prend en compte les valeurs extrêmes d'une série.

■ PROPRIÉTÉ :

Une moyenne est toujours comprise entre la plus petite valeur du caractère et la plus grande valeur du caractère.

Dans la suite, on considère les notes obtenues en mathématiques par Manon, Julien et Séverine au dernier trimestre. Les DM sont coefficient 1, les interrogations écrites coefficient 3 et les contrôles coefficient 5. Cela signifie que la note du contrôle compte cinq fois plus qu'une note de DM, ou qu'une note de contrôle équivaut à 5 fois la même note en DM ! Pour comprendre l'importance des coefficients, observons l'exemple suivant :

Exemple :

Manon a obtenu

9 au contrôle, 20 à l'interrogation écrite, 7 et 8 aux DM.

Julien a obtenu

12 au contrôle, 11 à l'interrogation écrite et 13 et 14 aux DM.

et Séverine a obtenu

15 au contrôle, 4 à l'interrogation écrite, 16 et 17 aux DM.

Compléter les tableaux de la feuille suivante avec les valeurs de chacune des séries.

Pour chaque élève, répondre aux questions suivantes :

- Pour Manon :
 - Étendue : $20 - 7 = 13$
 - Moyenne : $\frac{120}{10} = 12$
- Pour Julien :
 - Étendue : $14 - 11 = 3$
 - Moyenne : $\frac{120}{10} = 12$
- Pour Séverine :
 - Étendue : $17 - 4 = 13$
 - Moyenne : $\frac{120}{10} = 12$

Quelle appréciation écrire sur le bulletin de chacun ?

Vous constatez que les élèves n'ont pas le même profil. Leur professeur ne leur mettra donc pas la même appréciation. Comment se rendre compte de cela grâce aux statistiques ? Vous verrez les années à venir des solutions pour répondre à cette question. La *médiane* est une réponse...

Manon :

Notes, notées x_i	Coefficients, notés n_i	Fréquences, notées f_i	Effectifs cumulés croissants	Produits $n_i \times x_i$
7	1	0, 1	1	7
8	1	0, 1	2	8
9	5	0, 5	7	45
20	3	0, 3	10	60
Total	10	1	×	120

Julien :

Notes, notées x_i	Coefficients, notés n_i	Fréquences, notées f_i	Effectifs cumulés croissants	Produits $n_i \times x_i$
11	3	0, 3	3	33
12	5	0, 5	8	60
13	1	0, 1	9	13
14	1	0, 1	10	14
Total	10	1	×	120

Séverine :

Notes, notées x_i	Coefficients, notés n_i	Fréquences, notées f_i	Effectifs cumulés croissants	Produits $n_i \times x_i$
4	3	0, 3	3	12
15	5	0, 5	8	75
16	1	0, 1	9	16
17	1	0, 1	10	17
Total	10	1	×	120

f. Médiane

■ DÉFINITION :

Une **médiane** est un nombre qui sépare **une série ordonnée** en deux séries de même effectif.

On peut donc dire que

- la moitié des valeurs de la série sont inférieures à la médiane.
- la moitié des valeurs de la série sont supérieures à la médiane.

■ PROPRIÉTÉ :

- Si l'effectif total de la série est impair, une médiane est la valeur de la série qui, une fois ordonnée, est au milieu de cette série. Soit N l'effectif total de la série. Une médiane est la valeur du terme de rang $\frac{N+1}{2}$ de la série ordonnée.
- Si l'effectif total de la série est pair, une médiane est un nombre qui se trouve entre les deux valeurs du milieu de la série. C'est, par exemple, la moyenne de ces deux valeurs de la série, c'est-à-dire la moyenne des valeurs $x_{\frac{N}{2}}$ et $x_{\frac{N}{2}+1}$.

Exemple :

Un élève a obtenu les résultats suivants pendant le trimestre :

$$11 - 16 - 18 - 8 - 12$$

- Quelle est la moyenne de la série ? $\frac{11 + 16 + 18 + 8 + 12}{5} = \frac{65}{5} = 13$.
- Quelle est une médiane de la série ? Comment interpréter ce nombre ? On classe d'abord la série dans l'ordre croissant, et on regarde la valeur centrale.
 $8 - 11 - \mathbf{12} - 16 - 18$. Une médiane est donc 12. Il y a donc 50% des notes obtenues par l'élève qui sont inférieures à 12 et 50% des notes obtenues par l'élèves qui sont supérieures à 12.

Exemple :

Un élève a obtenu les résultats suivants pendant le trimestre :

$$11 - 16 - 18 - 8 - 12 - 13$$

- Quelle est la moyenne de la série ? $\frac{11 + 16 + 18 + 8 + 12 + 13}{6} = \frac{78}{6} = 13$.
- Quelle est une médiane de la série ? Comment interpréter ce nombre ? On classe d'abord la série dans l'ordre croissant, et on regarde la moyenne des valeurs centrales. $8 - 11 - \mathbf{12} - \mathbf{13} - 16 - 18$. Une médiane est donc $\frac{12 + 13}{2} = 12,5$. Il y a donc 50% des notes obtenues par l'élèves qui sont inférieures à 12,5 et 50% des notes obtenues par l'élèves qui sont supérieures à 12,5.

Exemple :

Voilà les séries des notes de Séverine, Julien et Manon. Ordonnez chacune des séries de notes, et déterminez une médiane de chaque série.

- Manon : $7 - 8 - 9 - 9 - 9 - 9 - 9 - 9 - 20 - 20 - 20$; une médiane est $\frac{9 + 9}{2} = 9$.
- Julien : $11 - 11 - 11 - 12 - 12 - 12 - 12 - 12 - 13 - 14$; une médiane est $\frac{12 + 12}{2} = 12$
- Séverine : $4 - 4 - 4 - 15 - 15 - 15 - 15 - 15 - 16 - 17$; une médiane est $\frac{15 + 15}{2} = 15$.

Exemple :

Dans l'exemple des frères et sœurs, quelle est une médiane du nombre de frères et sœurs dans la classe interrogée ? Comment interpréter ce résultat ? Une médiane de la série est 2. Il y a donc 50% des élèves qui ont moins de 2 frères et sœurs et 50% des élèves qui ont plus de 2 frères et sœurs.

Pouvez-vous trouver une médiane d'une série à l'aide des effectifs cumulés croissants ?

Oui ! Par exemple, dans le cas des frères et sœurs, les valeurs d'indice 15 et 16 sont atteintes pour 2.

Donc la médiane est bien $\frac{2+2}{2} = 2$ frères et sœurs.

Remarque : Pour déterminer une médiane d'une série, il faut l'ordonner. Une médiane est un indicateur de position qui ne dépend pas des valeurs extrêmes.

Exemple :

Déterminez une médiane de la série de buts marqués par Lily. Que signifie ce nombre ?

0 - 2 - 2 - 3 - 3 - 3 - 4 - 4 - 4 - 4 - 4 - 4 - 4 - 5 - 5 - 5 - 6 - 6 - 6 - 6

Une médiane est $\frac{4+4}{2} = 4$. Donc dans 50% des cas, Lily a marqué moins de 4 buts, et dans 50% des cas, Lily a marqué plus de 4 buts.

g. Quartiles

■ DÉFINITION :

Soit $x_1 \leq x_2 \leq \dots \leq x_N$ une série statistique d'effectif total N dont les valeurs sont classées par ordre croissant.

- Le premier quartile Q_1 de la série est la plus petite valeur de la série telle qu'au moins 25 % des valeurs de la série lui soient inférieures ou égales.
- Le troisième quartile Q_3 de la série est la plus petite valeur de la série telle qu'au moins 75 % des valeurs de la série lui soient inférieures ou égales.

Remarque : Le deuxième quartile Q_2 correspond à la médiane.

Remarque : Pour tout calcul de médiane ou quartile, il convient d'abord d'ordonner les valeurs de la série par ordre croissant.

■ MÉTHODE :

En pratique, pour déterminer Q_1 , on calcule $\frac{1}{4}N$ puis on considère la première valeur d'indice supérieur ou égale à $\frac{1}{4}N$: cette valeur est le premier quartile. De même pour le troisième quartile avec $\frac{3}{4}N$

■ DÉFINITION :

On appelle **intervalle inter-quartile** l'intervalle $[Q_1 ; Q_3]$. Il regroupe environ 50% des valeurs de la série autour de leur médiane. L'**écart inter-quartile** est $Q_3 - Q_1$, c'est-à-dire l'amplitude de l'intervalle précédent.

Exemple :

Voici la série des salaires mensuels dans une certaine entreprise :

Salaire x_i (en euros)	1500	2000	3000	10 000
Effectif n_i	500	296	201	3

Calculer la moyenne, la médiane, les quartiles et l'écart inter-quartile de cette série. Interpréter.

$$\bar{x} = \frac{1\,500 \times 500 + 2\,000 \times 296 + 3\,000 \times 201 + 10\,000 \times 3}{500 + 296 + 201 + 3} = \frac{1\,975\,000}{1\,000} = 1\,975 \text{ euros.}$$

$$\frac{N}{4} = 250; \quad \frac{N}{2} = 500; \quad \frac{N}{2} + 1 = 501 \text{ et } \frac{3N}{4} = 750.$$

Une fois les valeurs rangées par ordre croissant, la valeur du caractère d'indice 250 est 1 500, celle d'indice 500 est 1 500, celle d'indice 501 est 2 000, celle d'indice 750 est 2 000.

Donc le premier quartile est 1 500 euros, la médiane est $\frac{1500 + 2000}{2} = \frac{3500}{2} = 1\,750$ euros et le troisième quartile est 2 000 euros.

Ainsi, 25 % des salariés gagnent au plus 1500 euros, 50 % des salariés gagnent au plus 1750 euros et 75 % des salariés gagnent au plus 2 000.

L'intervalle inter-quartile est [1500 ; 2000]. Ainsi 50 % des salaires sont compris entre 1500 et 2000 euros.

g. Écart-type

On observe les résultats en mathématiques de deux élèves, Alice et Bastien, au cours d'un trimestre.

- Alice : 11 / 9 / 13
 - Bastien : 12 / 3 / 18
- On constate que $\overline{x_A} = \overline{x_B} = 11$.

Notes	11 / 9 / 13	12 / 3 / 18
Moyenne	11	11
Écarts par rapport à la moyenne	0 / -2 / +2	+1 / -8 / +7
Moyenne des écarts par rapport à la moyenne	0	0
Valeurs absolues des écarts par rapport à la moyenne	0 / 2 / 1	1 / 8 / 7
Moyenne des valeurs absolues des écarts par rapport à la moyenne	$\frac{4}{3} \approx 1,33$	$\frac{16}{3} \approx 5,33$
Carrés des écarts par rapports à la moyenne	0 / 4 / 4	1 / 64 / 49
Moyenne des carrés des écarts par rapport à la moyenne	$\frac{8}{3} \approx 2,67$	38
Racine carrée de la moyenne des carrés des écarts par rapport à la moyenne	$\sqrt{\frac{8}{3}} \approx 1,63$	$\sqrt{38} \approx 6,16$

■ DÉFINITION :

On appelle **variance** le nombre $V_x = \frac{\sum n_i (x_i - \overline{x})^2}{\sum n_i}$.

■ DÉFINITION :

On appelle **écart-type** le nombre $\sigma_x = \sqrt{V_x} = \sqrt{\frac{\sum n_i (x_i - \overline{x})^2}{\sum n_i}}$.

Exemple :

On considère une série statistique des notes de 28 élèves d'une classe, que l'on regroupe dans le tableau suivant :

x_i	n_i	$n_i x_i$	$n_i (x_i - \bar{x})^2$	$n_i x_i^2$
2	2	4	128	8
6	7	42	112	252
10	9	90	0	900
14	9	126	144	1764
18	1	18	64	324
Σ	28	280	448	3248

La moyenne de la série statistique est :

$$\bar{x} = \frac{\sum n_i x_i}{\sum n_i} = \frac{280}{28} = 10$$

La variance de cette série statistique est :

$$V_x = \frac{448}{28} = 16$$

L'écart-type de cette série statistique est :

$$\sigma_x = \sqrt{\frac{448}{28}} = 4$$

■ PROPRIÉTÉ : Lemme de König-Huygens

L'écart-type est la racine carrée de la différence entre moyenne des carrés et le carré de la moyenne, c'est-à-dire $\sigma_x = \sqrt{\overline{x^2} - \bar{x}^2}$

Exemple :

Dans l'exemple précédent, on rappelle que $\bar{x} = 10$. On calcule $\overline{x^2}$ grâce au tableau.

$$\overline{x^2} = \frac{\sum n_i x_i^2}{\sum n_i} = \frac{3248}{28}$$

Alors

$$\sigma_x = \sqrt{\overline{x^2} - \bar{x}^2} = \sqrt{\frac{3248}{28} - \left(\frac{280}{28}\right)^2} = 4$$

On peut vérifier grâce à la calculatrice :

rad STATISTICS			rad STATISTICS		
Data	Graph	Stats	Data	Graph	Stats
		V1/N1			V1/N1
Number of data points	n	28	Mean	μ	10
Minimum	Min	2	Standard deviation	σ	4
First quartile	Q1	6	Variance	σ^2	16
Median	Med	10	Sample mean	$\bar{x}$	10
Third quartile	Q3	14	Sample std deviation	s	4.073401
Maximum	Max	18	Sample variance	s ²	16.59259
Range	R	16	Sum of values	Σx	280
Interquartile range	IQR	8	Sum of squared values	Σx^2	3248
			Mode 1	Mod1	10

B. Caractères quantitatifs continus

Pour les caractères quantitatifs continus, on sépare les valeurs du caractère en intervalles de même taille.

■ DÉFINITION :

Les intervalles dans lesquels on regroupe les valeurs sont appelés *classes*.

La taille de chaque classe est appelée **l'amplitude de la classe**.

On considère l'enquête suivante :

- Population : Élèves du lycée franco-allemand
- Échantillon : Classe de première ES
- Caractère : Taille, en cm

On a obtenu les résultats suivants :

156 - 159 - 168 - 161 - 156 - 176 - 174 - 162 - 158 - 161 - 174 - 161 - 151 - 166 - 167 - 176 - 153 - 173 - 166 - 175 -
166 - 161 - 175 - 160 - 174 - 166 - 161 - 175 - 172 - 160

Regrouper cette série statistique par classes d'amplitude 5 cm et calculer les fréquences en % arrondies à l'unité :

Tailles, x_i	$150 \leq t < 155$	$155 \leq t < 160$	$160 \leq t < 165$	$165 \leq t < 170$	$170 \leq t < 175$	$175 \leq t < 180$	Total
Effectifs, n_i	2	4	8	6	5	5	30
Fréquences, f_i	0.067	0.133	0.267	0.2	0.167	0.167	1.0
Fréquences, en %	6.7	13.3	26.7	20.0	16.7	16.7	100.0

Histogramme des effectifs

On peut associer aux classes un **histogramme**, dans lequel la légende est **indispensable**.

Si les classes ont toutes la même amplitude, alors la hauteur des rectangles est proportionnelle à l'effectif de chaque classe, comme dans la figure de gauche ci-dessous. Sinon, dans cet exemple, en regroupant les deux premières classes, on obtiendrait la représentation de droite :

