1 Vocabulaire

— **Population** : c'est l'ensemble étudié.
— **Individu** : c'est un élément de la population.
— **Effectif total** : c'est le nombre total d'individus.
— **Caractère** : c'est la propriété étudiée.
Il y a deux sortes de caractères :
  — les caractères quantitatifs que l'on peut mesurer avec des nombres.
  On distingue les caractères quantitatifs **discrets** qui ne peuvent prendre qu'un nombre fini de valeurs (notes à un devoir...) et les caractères quantitatifs **continus** dont on regroupe les valeurs par intervalles (taille, durée d'écoute...).
  — les caractères qualitatifs (profession, marque de voiture ...).

Exemple. On considère l'exemple suivant (qui servira pour les prochains paragraphes) : les 17 élèves d'une classe sont séparés en deux groupes.
— Les notes obtenues à un devoir par le **groupe A** sont : 4 ; 4 ; 9 ; 9 ; 9 ; 15 ; 15 ; 17 ; 17
— Les notes obtenues à un devoir par le **groupe B** sont : 6 ; 6 ; 8 ; 8 ; 10 ; 10 ; 12 ; 12
— La population étudiée est l'ensemble ....
— Le caractère étudié est la ....
— L'effectif total est égal à ....

2 Fréquence

Exemple. Voici les notes obtenues à un contrôle sur 10 par une classe de première :

0 – 1 – 1 – 2 – 2 – 2 – 2 – 3 – 3 – 4 – 5 – 5 – 7 – 7 – 7– 8 – 8 – 10 – 10 – 10.
Compléter le tableau ci-dessous :

Note	0	1	2	3	4	···	···	···	···
Effectifs	1	2	4	2					
Effectifs cumulés	1	3	7	9					
Fréquences (en%)	5%	10%	20%	10%					
Fréquences cumulées (en%)	5%	15%	35%	45%					

3 Moyenne

On donne la série de nombres suivants :

32, 6, 18, 29, 6, 48, 50, 12, 32, 4, 50, 10, 29, 72, 32,16 ,16 , 6, 50, 50, 4, 18, 6, 10, 29, 12, 48, 6, 32, 50

Moyenne arithmétique simple

La **moyenne arithmétique simple** est égale à :

$$\frac{32+6+18+29+\cdots+6+32+50}{30} = \cdots\cdots\cdots\cdots$$

On peut aussi regrouper les nombres de l'exemple précédent dans le tableau suivant :

Nombre	4	6	10	12	16	18	29	32	48	50	72
Effectif	2	5	2	2	2	2	3	4	2	5	1
Effectif cumulé	2	7	9	11	13	15	18	22	24	29	30

Moyenne arithmétique pondérée

La **moyenne arithmétique pondérée** est égale à :

$$\frac{4\times 2+6\times 5+10\times 2+\dots+50\times 5+72\times 1}{30}=\dots\dots\dots\dots$$

Définition 3.1

On appelle moyenne de la série

valeur	x_1	x_2	$\cdots$	x_k
effectif	n_1	n_2	$\cdots$	n_k

, le réel $\overline{x}=\dfrac{n_1x_1+n_2x_2+\dots+n_kx_k}{N}$

(N représente l'effectif total et k est le nombre de valeurs prises par le caractère).

Propriété 3.2

- Si les valeurs d'une série statistique sont multipliées ou divisées par un même nombre a, la moyenne de cette série est aussi multipliée ou divisée par a.
- Si une population d'effectif N est composée d'une partie d'effectif N_A et de moyenne $\overline{x_A}$ et d'une autre partie d'effectif N_B et de moyenne $\overline{x_B}$, alors la moyenne de la population totale est telle que : $\overline{x}=\dfrac{N_A\overline{x_A}+N_B\overline{x_B}}{N}$.

Exemple.

- Pour la série du groupe A

valeur	4	9	15	17
effectif	2	3	2	2

, la moyenne est

$$\overline{x_A}=\frac{2\times 4+3\times 9+2\times 15+2\times 17}{9}=11$$

- Pour la série du groupe B

valeur	6	8	10	12
effectif	2	2	2	2

, la moyenne est

$$\overline{x_B}=\frac{2\times 6+2\times 8+2\times 10+2\times 12}{8}=9$$

La moyenne du groupe A, composée de $N_A=\cdots$ élèves, était $\overline{x_A}=\cdots$ et la moyenne du groupe B, composée de $N_B=\cdots$ élèves, était $\overline{x_B}=\cdots$. On peut en déduire que la moyenne globale de la classe est $\overline{x}=\dfrac{N_A\overline{x_A}+N_B\overline{x_B}}{N}=\dfrac{\dots\dots\dots+\dots\dots\dots}{\cdots}\approx 10,06$.

4 Médiane

Définition 4.1

La médiane d'une série statistique est la valeur qui partage le groupe étudié en deux sous groupe de même effectif chacun tels que :

- ★ tous les éléments du premier sous-groupe soient inférieurs ou égales à la médiane ;
- ★ tous les éléments du deuxième sous-groupe soient supérieurs ou égales à la médiane.

Exemples. Il y a deux cas.

1. Cas d'un nombre impair de valeurs dans une série : 1 ; 5 ; 7 ; 10 ; 11 ; 50 ; 55.
   ....
   ....
   ....
   ....
   ....
2. Cas d'un nombre pair de valeurs dans une série : 10 ; 20 ; 30 ; 35 ; 37 ; 40 ; 50 ; 60.
   ....
   ....
   ....
   ....
   ....

Exercice 1

Un professeur de SVT demande aux 29 élèves d'une classe de sixième de faire germer des graines de blé chez eux. Le professeur donne un protocole expérimental à suivre :

- mettre en culture sur du coton dans une boîte placée dans une pièce éclairée, de température entre 20 °et 25 °C ;
- arroser une fois par jour ;
- il est possible de couvrir les graines avec un film transparent pour éviter l'évaporation de l'eau.

Le tableau ci-dessous donne les tailles des plantules (petites plantes) des 29 élèves à 10 jours après la mise en germination.

Taille en cm	0	8	12	14	16	17	18	19	20	21	22
Effectif	1	2	2	4	2	2	3	3	4	4	2
Effectifs cumulés											

1. Combien de plantules ont une taille qui mesure au plus 12 cm ?
   ....
   ....
2. Calculer la moyenne de cette série. Arrondir au dixième près.
   ....
   ....
   ....
3. Déterminer la médiane de cette série et interpréter le résultat.
   ....
   ....
   ....
4. On considère qu'un élève a bien respecté le protocole si la taille de la plantule à 10 jours est supérieure ou égale à 14 cm. Quel pourcentage des élèves de la classe a bien respecté le protocole ?
   ....
   ....
   ....
5. Le professeur a fait lui-même la même expérience en suivant le même protocole. Il a relevé la taille obtenue à 10 jours de germination. Prouver que, si on ajoute la donnée du professeur à cette série, la médiane ne changera pas.
   ....
   ....
   ....

Exercice 2

On considère la série statistique donnant le SMIC horaire brut en euros de 2001 à 2011 (source : INSEE).

Année	SMIC
2011	9,40
2010	9,00
2009	8,82
2008	8,63
2007	8,44
2006	8,27
2005	8,03
2004	7,61
2003	7,19
2002	6,83
2001	6,67

1. Quelle est la médiane ?
   ....
   ....
   ....
2. Paul remarque qu'entre 2001 et 2002, l'augmentation du SMIC horaire brut est de 16 centimes alors qu'entre 2007 et 2008, elle est de 19 centimes.
   Il affirme que « le pourcentage d'augmentation entre 2007 et 2008 est supérieur à celui pratiqué entre 2001 et 2002 ». A-t-il raison ?
   ....
   ....
   ....

5 Quartiles et écart interquartile d'une série statistique

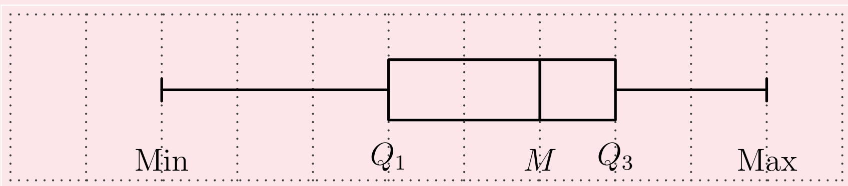
Définition 5.1

On détermine les quartiles, en écrivant les valeurs du caractère par ordre croissant de telle façon que chaque valeur apparaisse un nombre de fois égal à son effectif, et en partageant la liste en deux sous-séries de même effectif, si l'effectif total est impair, on ne tient pas compte de la médiane.
- le premier quartile Q_1 est la médiane de la sous-série inférieure ;
- le troisième quartile Q_3 est la médiane de la sous-série supérieure ;
- l'écart interquartile est défini par $Q_3 - Q_1$ et sert à mesurer la dispersion des valeurs.

Exemples. Il y a deux cas.
- Groupe A :
  - Détermination de la médiane : $4 \,;\, 4 \,;\, 9 \,;\, 9 \,;\, \overline{9} \,;\, 15 \,;\, 15 \,;\, 17 \,;\, 17$
    La médiane est $M = 9$, valeur située au milieu, car l'effectif total est impair.
  - Détermination de Q_1 et Q_3 : on partage la liste en deux sous-séries de même effectif, on ne tient pas compte de la médiane car l'effectif total est impair :
    $\underbrace{4 \,;\, \underline{4} \,;\, \underline{9} \,;\, 9}\,;\, 9 \,;\, \underbrace{15 \,;\, \underline{15} \,;\, \underline{17} \,;\, 17}$
    Les effectifs des deux sous-séries étant pairs, Q_1 est la demi-somme de la sous-série inférieure et Q_3 est la demi-somme de la sous-série supérieure :
    $Q_1 = \dfrac{\cdots + \cdots}{2} = 6,5\,;\ Q_3 = \dfrac{\cdots + \cdots}{2} = 16$
    L'écart interquartile est $Q_3 - Q_1 = \cdots - \cdots = 9,5$
- Groupe B :
  - Détermination de la médiane : $6 \,;\, 6 \,;\, 8 \,;\, \overline{8} \,;\, \overline{10} \,;\, 10 \,;\, 12 \,;\, 12$
    La médiane est $M = \dfrac{\cdots + \cdots}{2} = 9$, la demi-somme des deux valeurs situées au milieu, car l'effectif total est pair.
  - Détermination de Q_1 et Q_3 : on partage la liste en deux sous-séries de même effectif :
    $\underbrace{6 \,;\, \underline{6} \,;\, \underline{8} \,;\, 8}\,;\, \underbrace{10 \,;\, \underline{10} \,;\, \underline{12} \,;\, 12}$
    Les effectifs des deux sous-séries étant pairs, Q_1 est la demi-somme de la sous-série inférieure et Q_3 est la demi-somme de la sous-série supérieure :
    $Q_1 = \dfrac{\cdots + \cdots}{2} = 7\,;\ Q_3 = \dfrac{\cdots + \cdots}{2} = 11$
    L'écart interquartile est $Q_3 - Q_1 = \cdots - \cdots = 4$

Définition 5.2

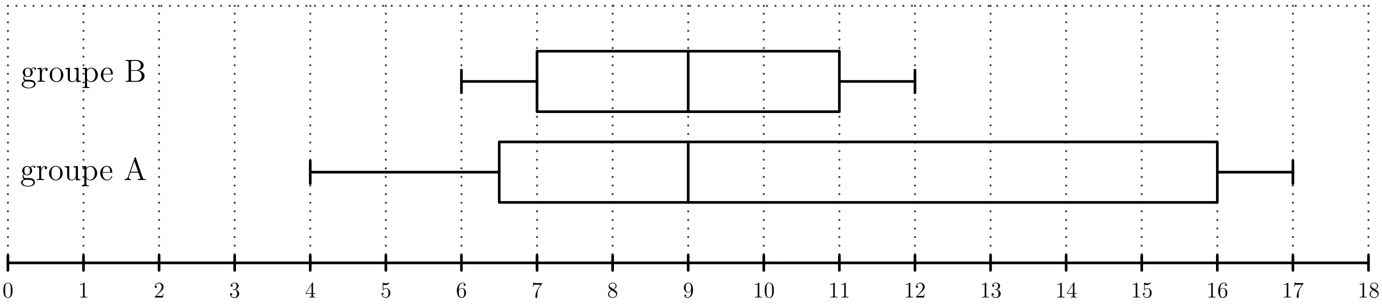
Le diagramme en boîtes d'une série statistique se construit de la façon suivante :



- 25% des valeurs du caractère sont comprises entre Min et Q_1 ;
- 25% des valeurs du caractère sont comprises entre Q_1 et M ;
- 25% des valeurs du caractère sont comprises entre M et Q_3 ;
- 25% des valeurs du caractère sont comprises entre Q_3 et Max.

Exemples. On reprend des données des exemples précédents.
- Pour le groupe A, on a $Min = \cdots$; $Q_1 = \cdots$; $M = \cdots$; $Q_3 = \cdots$; $Max = \cdots$.
- Pour le groupe B, on a $Min = \cdots$; $Q_1 = \cdots$; $M = \cdots$; $Q_3 = \cdots$; $Max = \cdots$

Les diagrammes en boîtes des deux groupes donnent,



Exercice 1

On considère la série

valeur	25	32	57	75
effectif	3	4	1	2

1. Calculer la moyenne de cette série.
   ....
   ....
2. Si on multiplie toutes les valeurs du caractère par 5, quelle est la nouvelle moyenne que l'on obtient ?
   ....

Exercice 2

Dans une classe, il y a 20 filles et 15 garçons. La taille moyenne de l'ensemble des élèves est de 1,7 m ; la taille moyenne des garçons est de 1,8 m. Quelle est la taille moyenne des filles de la classe ?
....
....
....
....
....

Exercice 3

1. On considère la série suivante (les valeurs du caractère sont écrites les unes à la suite des autres dans l'ordre croissant)

   1 ; 1 ; 4 ; 4 ; 5 ; 5 ; 6 ; 7 ; 7 ; 8 ; 8 ; 10 ; 11

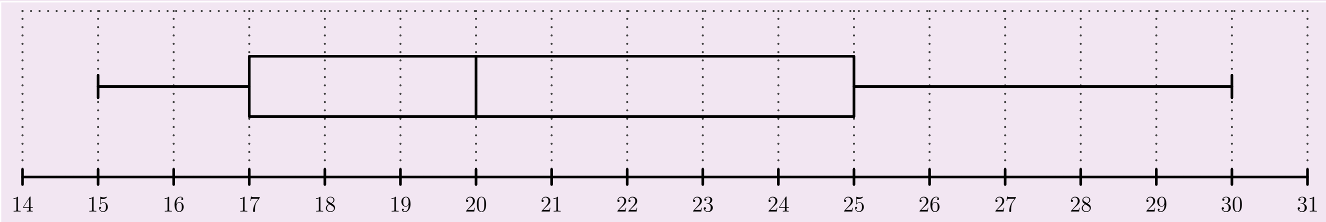
   Déterminer la médiane, le premier et le troisième quartile de cette série :
   ....
   ....
   ....
   ....
2. On considère la série suivante (les valeurs du caractère sont écrites les unes à la suite des autres dans l'ordre croissant)

   12 ; 12 ; 13 ; 15 ; 16 ; 18 ; 18 ; 21 ; 22 ; 22

   Déterminer la médiane, le premier et le troisième quartile de cette série :
   ....
   ....
   ....
   ....
   ....

Exercice 4

Le diagramme en boîte ci-dessous représente la répartition du nombre de BD vendues par jour dans une librairie.



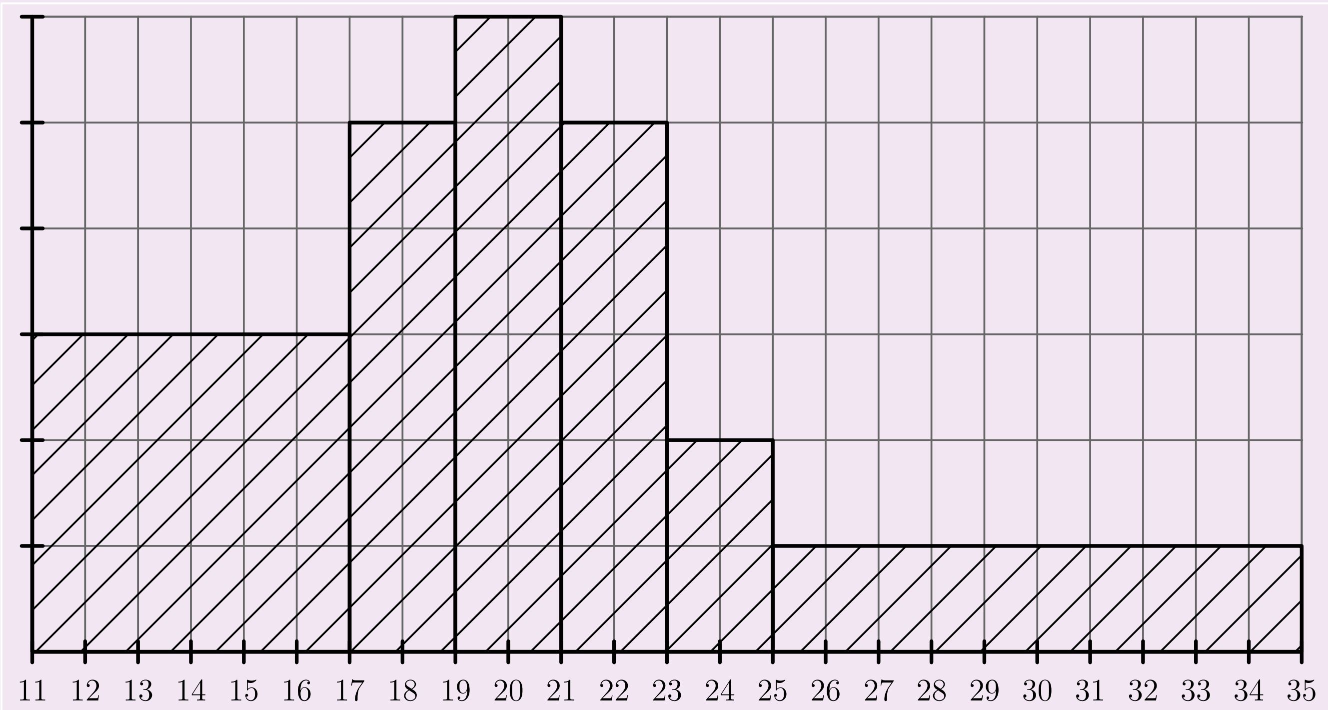
Déterminer :

1. le nombre médian de BD vendues par jour. ....
2. le pourcentage de jours où le nombre de BD vendues est compris entre 17 et 25. ....
3. le pourcentage de jours où le nombre de BD vendues est inférieur ou égal à 25 ....

Exercice 5

Dans l'histogramme ci-dessous, l'effectif correspondant à l'intervalle $[23, 25[$ est égal à 10. En déduire l'effectif correspondant aux autres intervalles grâce au tableau suivant :

valeur	[11 ; 17[	[17 ; 19[	[19 ; 21[	[21 ; 23[	[23 ; 25[	[25 ; 35[
aire						
effectif					10	



Exercice 6

On considère la série suivante :

valeur	$[2; 8[$	$[8; 10[$	$[10; 12[$	$[12; 20[$
effectif	12	6	8	16

1. Calculer la moyenne de cette série en s'aidant du tableau suivant :

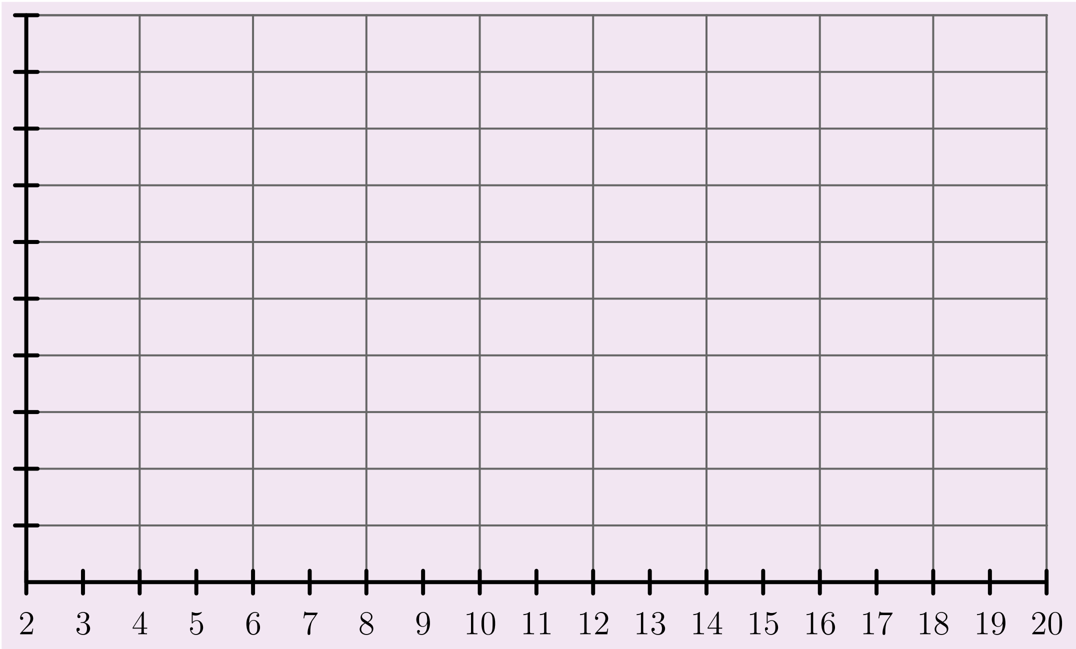
milieu				
effectif	12	6	8	16

....

Exercice 6 (suite)

2. Construire le diagramme en bâtons dans la figure ci-dessous en s'aidant du tableau suivant :

valeur	$[2;8[$	$[8;10[$	$[10;12[$	$[12;20[$
aire (1 carreau pour 1 individu)	12	6	8	16
largeur				
hauteur				



6 Écart-type d'une série statistique

Définition 6.1

— **La variance** V d'une série statistique d'effectif total N est la moyenne des carrés des écarts à la moyenne : $V = \dfrac{n_1\left(x_1-\overline{x}\right)^2+n_2\left(x_2-\overline{x}\right)^2+\cdots+n_k\left(x_k-\overline{x}\right)^2}{N}$.

— **L'écart-type** σ d'une série statistique est la racine carrée de la variance : $\sigma=\sqrt{V}$.
Il sert à mesurer la dispersion des valeurs de la série autour de la moyenne.

Exemples. On reprend les mêmes données des exemples précédents.

— Dans le groupe A, la moyenne était de 11. Donc,

Valeur	4	9	15	17
Écart à la moyenne	-7	$\cdots$	4	$\cdots$
Carré de l'écart	$\cdots$	4	16	$\cdots$
Effectif	2	3	2	2

Variance $V_A=\dfrac{2\times49+3\times4+2\times16+2\times36}{9}\approx 23{,}8$.
Écart-type $\sigma_A\approx\sqrt{23{,}8}\approx 4{,}9$.

— Dans le groupe B, la moyenne était de 9. Donc,

Valeur	6	8	10	12
Écart à la moyenne	-3	$\cdots$	1	3
Carré de l'écart	$\cdots$	1	1	$\cdots$
Effectif	2	2	2	2

Variance $V_B=\dfrac{2\times9+2\times1+2\times1+2\times9}{8}=5$.
Écart-type $\sigma_B=\sqrt{5}\approx 2{,}2$.

Exercice 1

Calculer la moyenne et l'écart-type de la série suivante :

Valeur	1	5	13	17
Effectif	2	1	3	2

Exercice 2

On considère la série suivante :

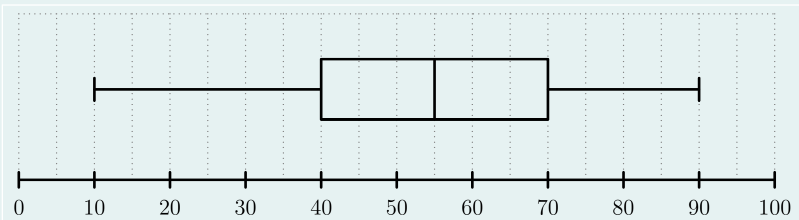
Classe	$[2;8[$	$[8;10[$	$[10;12[$	$[12;20[$
Effectif	x	6	8	16

1. Déterminer la valeur qu'il faut donner à x pour que la moyenne de la série soit égale à 9.

2. Calculer alors l'écart-type de la série.

Exercice 3

Le diagramme en boîtes d'une série est le suivant :



1. Déterminer la médiane, l'écart interquartile et les valeurs extrêmes de la série.

2. Quelle est la proportion de valeurs du caractère supérieures à 40 ?

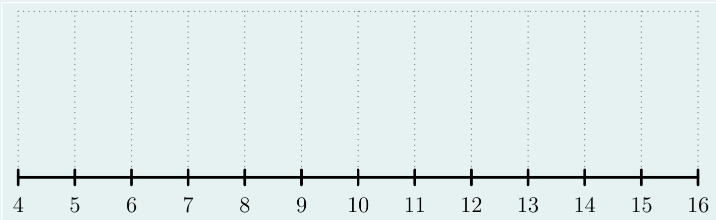
Exercice 4

On considère la série suivante.

Valeur	5	8	10	12	15
Effectif	2	1	4	3	2

1. Déterminer la médiane, le premier quartile et le troisième quartile de la série.
   ....
   ....
   ....
   ....
   ....

2. Construire ci-dessous le diagramme en boîtes de la série :



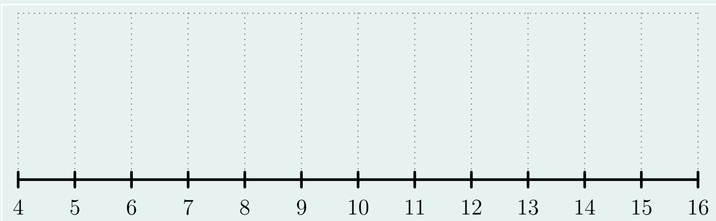
Exercice 5

On considère la série suivante.

Valeur	2	8	12	14	19
Effectif	2	1	4	2	1

1. Déterminer la médiane, le premier quartile et le troisième quartile de la série.
   ....
   ....
   ....
   ....
   ....

2. Construire ci-dessous le diagramme en boîtes de la série :



Exercice 6

Les séries suivantes donnent les précipitations moyennes mensuelles en millimètres à Nice et à Paris.

	J	F	M	A	M	J	J	A	S	O	N	D
Nice	67	83	70	70	42	37	20	38	83	109	160	97
Paris	55	48	40	45	55	60	56	65	60	52	60	52

1. Ordonner les valeurs pour les deux villes dans l'ordre croissant.
   En déduire la médiane, le premier quartile et le troisième quartile de chaque série.
   ....
   ....
   ....
   ....

Exercice 6 (suite)

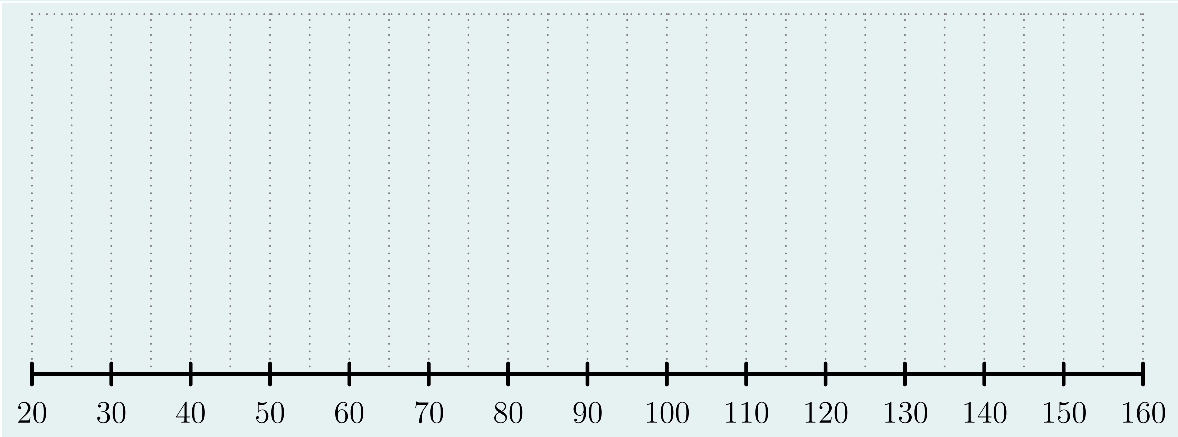
....

....

....

....

2. Construire les diagrammes en boîtes des deux séries dans le même repère ci-dessous.



3. Dans quelle ville la pluviométrie est-elle la plus régulière ?

....

....

....

Exercice 7

Une société de maintenance a en charge l'entretien de 439 machines du même type. Pendant une année, elle a noté le nombre d'interventions réalisé sur chacune de ces machines. Les données sont regroupées dans le tableau suivant.

Nombre d'interventions	1	2	3	4	5	6	7	8	9	10
Nombre de machines concernées	10	12	17	44	78	94	83	49	36	16

1. Calculer le nombre moyen d'interventions $\bar{x}$.

....

....

....

....

2. Le gérant de la société considère qu'il faudrait renouveler le parc de machines si moins de $2/3$ des valeurs de la série sont comprises dans l'intervalle $[\bar{x} - \sigma\,;\, \bar{x} + \sigma]$.
Sachant que $\sigma \approx 1,97$, que va décider le gérant ?

....

....

....

....

....

....

....

....

....

....

....

....