

# Chapitre 01 : REPRÉSENTATION ET TRAITEMENT DES DONNÉES

## 1) Définitions : Population – Individu - Caractère :

Lorsque l'on mène une enquête, on s'intéresse à une **population d'individus**

**Exemples :** élèves d'une classe, pays de l'Union Européenne, animaux d'une région ...

et on étudie une propriété commune appelée un **caractère** ou une **donnée**.

**Exemples :** taille des élèves d'une classe, langue officielle des pays de l'Union Européenne, régime alimentaire des animaux d'une région ...

### Exemple :

Une Q.C.M. sur l'acquis des connaissances de 4èmes à été soumise aux **élèves de 3èmes Z** du collège de GERSTHEIM.

### Le barème était le suivant :

- +1 point par réponse correcte
- 2 points par réponse incorrecte
- 1 point pour une absence de réponse.

Ci-contre l'ensemble **des notes** obtenues :

↑  
Données

-5	-2	-20	-17	-17
4	4	-17	-20	-11
-11	4	-5	-8	7
10	-11	-2	-11	-5
-20	-14	4	1	-8
-8	10	-11	-2	-5

## 2) Définition : Étendue :

L'**étendue** d'une série **statistique** est la différence entre la plus grande et la plus petite des valeurs du caractère..

### Exemple :

Dans l'exemple précédent, les notes obtenues s'échelonnent de -20 à 10, l'étendue de cette série statistique est de :  
 $10 - (-20) = 10 + 20 = 30$ .

## 3) Définitions : Effectif – Effectif total :

L'**effectif** d'une donnée est le nombre de fois où cette donnée apparaît.

Le nombre d'individus de la population étudiée est appelée l'**effectif total**.

### Exemple :

Dans l'exemple ci-dessus, pour plus de lisibilité, il est pertinent de regrouper ces données et d'en donner les effectifs.

Notes	-20	-17	-14	-11	-8	-5	-2	1	4	7	10	<b>Effectif total</b>
Effectifs	3	3	1	5	3	4	3	1	4	1	2	<b>30</b>

L'effectif total permet de procéder à une première vérification du tri par effectif .

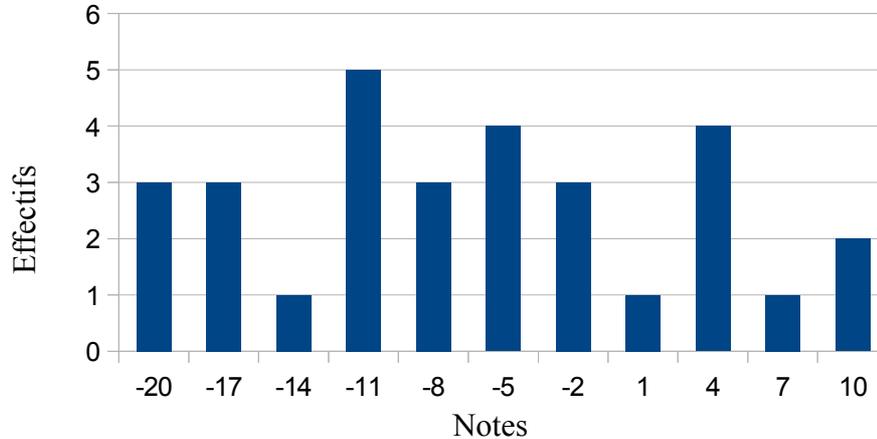
#### 4) Définition : Diagramme en barres :

Un diagramme en barres est une représentation graphique de données statistiques à l'aide de rectangles de même largeur.

Les valeurs du caractère étudié sont représentées sur l'axe horizontal, les effectifs sur l'axe vertical et à chaque valeur correspond une barre.

Les hauteurs des barres sont proportionnelles aux effectifs représentés.

##### Exemple :



#### 5) Définition : Moyenne :

La **moyenne d'une série** de valeurs est égale au quotient :

$$\text{Moyenne} = \frac{\text{Somme de toutes les valeurs}}{\text{Effectif total}}$$

##### Exemple :

Dans l'exemple précédent, pour calculer la moyenne, on peut additionner toutes les notes et diviser par le nombre total de notes

$$M = \frac{(-20) + (-20) + (-20) + (-17) + (-17) + (-17) + (-14) + \dots + (7) + (10) + (10)}{30}$$

$$M = \frac{-186}{30}$$

$$M = (-6,2)$$

##### Remarque :

La moyenne est comprise entre la **plus petite** et la **plus grande valeur** de la série.

#### 6) Définition : Moyenne pondérée :

La **moyenne pondérée d'une série** de valeurs est égale au quotient :

$$\text{Moyenne pondérée} = \frac{\text{Somme des produits des valeurs par leur effectifs}}{\text{Effectif total}}$$

##### Exemple :

Dans l'exemple précédent, le nombre de note est important. Il est judicieux d'utiliser la moyenne pondérée pour simplifier les calculs.

$$M = \frac{(-20) \times 3 + (-17) \times 3 + (-14) + \dots + 10 \times 2}{3 + 3 + 1 + 5 + 3 + 4 + 3 + 1 + 4 + 1 + 2}$$

$$M = \frac{-186}{30}$$

$$M = (-6,2)$$

## 7) Définitions : Médiane - Premier Quartile – Troisième Quartile :

On appelle **médiane** d'une série statistique rangée par ordre croissant une valeur qui la partage en deux populations de même effectif.

Le **premier quartile** d'une série statistique est la plus petite valeur  $Q_1$  telle qu'au moins 25 % des valeurs sont inférieures ou égales à  $Q_1$ .

Le **troisième quartile** d'une série statistique est la plus petite valeur  $Q_3$  telle qu'au moins 75 % des valeurs sont inférieures ou égales à  $Q_3$ .

### Exemples :

Dans l'exemple précédent, l'effectif total est de 30 élèves.

#### Calcul de la médiane :

La médiane est une valeur qui partage la série statistique en deux groupes de même effectif, ici 15. On en déduit que toute valeur comprise entre -8 et -5 est une médiane.

En effet,

- 15 élèves ont des notes inférieures ou égales à -8,
- 15 élèves ont des notes supérieures ou égales à -5.

En général on prend la demi-moyenne :  $\frac{(-8)+(-5)}{2} = -6,5$ .

#### Calcul du premier quartile et du troisième quartile :

$\frac{30 \times 25}{100} = 7,5$ . 25% de 30 valent 7,5. Le premier quartile est la 8ème valeur, soit -11.

$\frac{30 \times 75}{100} = 22,5$ . 75% de 30 valent 22,5. Le troisième quartile est la 23ème valeur, soit 1.

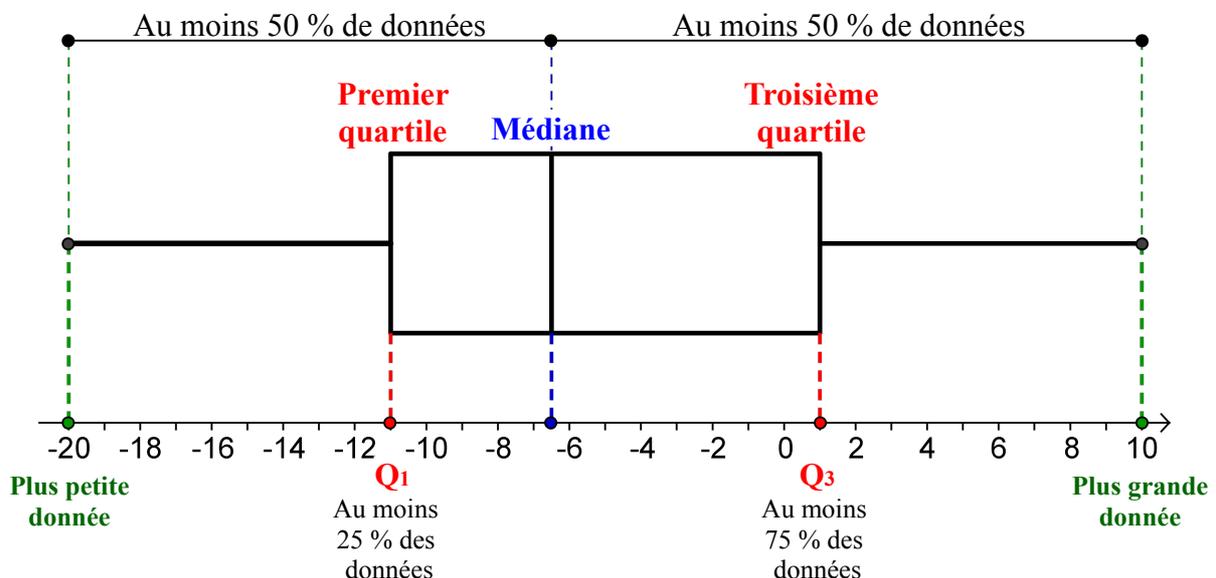
## 8) Définition : Boîte à moustache : Hors programme

Un diagramme en boîte (ou diagramme à moustaches ou diagramme à pattes) est une façon de représenter certains **paramètres** d'une série statistique.

Sur un axe, on place le **minimum** et le **maximum** de la série, le **premier quartile**, le **troisième quartile** et la **médiane**, puis on construit parallèlement à l'axe un rectangle (une boîte) de longueur l'**interquartile** et de largeur arbitraire.

### Exemple :

Représentation des données du précédent exemple par un diagramme en boîte :



### 9) Définition : Classes :

Dans le cas de nombreuses données numériques, on peut les regrouper en **classes** pour faciliter leur lecture.

#### Exemple :

En triant les données de l'exemple de l'exercice 1, on obtient :

Note	-20	-17	-14	-11	-8	-5				7	10	Total
Effectif	3	3	1	5	3	4				1	10	30

Ce tableau est à la fois long et difficile à lire et à interpréter.

Parfois il est préférable de regrouper les notes **en classes**, par exemple d'**amplitude 9**.

Note	$-20 \leq n < -11$	$-11 \leq n < -2$	$-2 \leq n < 7$	$7 \leq n$	Total
Effectif	7	12	8	3	30

On obtient l'effectif de cette classe en additionnant les effectifs des notes comprises entre -20 inclus et -11 exclus :  
 $3 + 3 + 1 = 7$ .

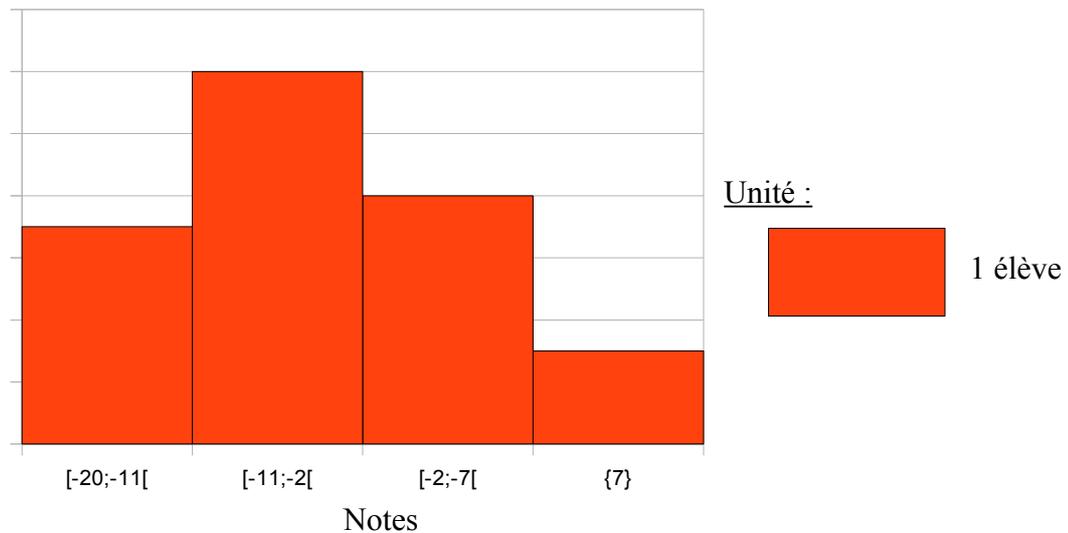
### 10) Définition : Histogramme :

On utilise un histogramme pour représenter des données numériques regroupées en classes.

L'aire des rectangles correspondant à chaque classe est proportionnelle à l'effectif de la classe.

#### Exemple :

On peut représenter les données de l'exemple de la définition 1 par un histogramme.



### 11) Définition : Fréquence :

La **fréquence** d'une donnée est le quotient de son effectif par l'effectif total.

La fréquence peut s'exprimer par une fraction, un nombre décimal ou un pourcentage.

#### Exemple :

Pour l'exemple précédent, 7 collégiens sur les 30 ont obtenu une note comprise entre -20 inclus et -11 exclus au Q.C.M.

La fréquence des collégiens qui ont obtenus une note comprise entre -20 inclus et -11 exclus au Q.C.M. est de :

$$\frac{7}{30} \approx 0,233 \text{ soit environ } 23,3 \%$$

↑ Fréquence sous forme fractionnaire
 ↑ Fréquence sous forme décimale
 ↑ Fréquence en pourcentage

#### Remarques :

- Une fréquence est comprise entre 0 et 1.
- La somme des fréquences de toutes les données est égale à 1.

#### Exercice :

1) Compléter le tableau ci-dessous :

Note	$-20 \leq n < -11$	$-11 \leq n < -2$	$-2 \leq n < 7$	$7 \leq n$	
Effectif	7	12	8	3	30
Fréquence	0,23				

Effectif total

Fréquence totale

### 12) Définition : Diagramme circulaire :

Un diagramme circulaire est une représentation graphique de données statistiques sous la forme d'un disque partagé en secteurs.

Un diagramme circulaire permet de visualiser rapidement la répartition des données.

#### Exemple :

L'angle de chaque secteur angulaire d'un diagramme circulaire est proportionnel à l'effectif correspondant. L'effectif total correspond à un angle de  $360^\circ$  ( $180^\circ$  pour les semi-circulaires).

Note	$-20 \leq n < -11$	$-11 \leq n < -2$	$-2 \leq n < 7$	$7 \leq n$	
Effectif	7	12	8	3	30
Fréquence	0,23	0,4	0,27	0,1	1
Secteur angulaire	84	144	96	36	360

↻ × 360

