

# STATISTIQUES : Vocabulaire et définitions

Une étude statistique consiste à recueillir des données, à organiser ces données sous forme adaptée de tableaux ou de représentations graphiques, à traiter ces informations afin de permettre une meilleure lecture et une analyse objective.

## I. LIRE UN TABLEAU.

Ce tableau donne la répartition des élèves demi-pensionnaires et externes selon les classes.

	6 <sup>ème</sup>	5 <sup>ème</sup>	4 <sup>ème</sup>	3 <sup>ème</sup>	Total
DEMI-PENSIONNAIRES	84	85	72	37	<b>278</b>
EXTERNES	78	96	91	64	<b>329</b>

Pour lire un tableau, on utilise à chaque fois le croisement d'une ligne et d'une colonne.

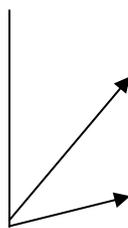
### **Exemple :**

Au croisement (intersection) de la ligne « **demi-pensionnaires** » et de la colonne « **5<sup>ème</sup>** », on trouve le nombre d'élèves « de 5<sup>ème</sup> **et** demi-pensionnaires » : 85 est l'**effectif** des élèves de 5<sup>ème</sup> demi-pensionnaires.

## II. DEFINITIONS.

**Population** : ensemble des personnes ou des objets (ex : barre de métal, les apprentis d'une classe...) sur lesquels porte l'étude statistique.

**Caractère** : propriété étudiée. (ex : couleur des yeux, nombre de cv de la voiture, âge des personnes)



**Qualitatif** (ce n'est pas un nombre). **Il exprime une qualité.**  
( ex : musique préférée des élèves, couleur du pull ).

**Quantitatif** (c'est un nombre). **Il exprime une quantité.**  
( ex : le poids, la taille des élèves....)

**Effectifs** : L' **effectif** est le nombre correspondant à une valeur ou à une classe.

L'**effectif total** est le nombre total de personnes ou d'objets étudié.

**Classe** : C'est un intervalle de valeurs, où toutes les valeurs peuvent être prises.

Dans certains cas on étudie un caractère quantitatif **continu**, les valeurs peuvent être toutes prises (ex : quelle taille fais-tu ? On aura des réponses pour toutes les valeurs entre 140 cm et 190 cm) alors ces valeurs sont regroupées en classes ou intervalles [a ;b[.

**Maximum** d'une série statistique à caractère quantitatif : est la valeur maximale que peut prendre le caractère.

**Minimum** d'une série statistique à caractère quantitatif : est la valeur minimale que peut prendre le caractère.

**III. FREQUENCES.** C'est le rapport d'un effectif sur l'effectif total.

### **Exemple :**

162 élèves sur 607 sont des élèves de 6<sup>ème</sup>. La **fréquence** des élèves de 6<sup>ème</sup> parmi les élèves du collège est  $\frac{162}{607} \approx 0,267 = 26,7\%$ . Le tableau des fréquences est donc le suivant :

	6 <sup>ème</sup>	5 <sup>ème</sup>	4 <sup>ème</sup>	3 <sup>ème</sup>	Total
FREQUENCE	26,7%	29,8%	26,8%	16,7%	<b>100%</b>

On peut l'exprimer par un **quotient**, un **nombre décimal**, ou un **pourcentage**.

### POURCENTAGES (EXEMPLES) :

Appliquer (prendre) un pourcentage :

Combien font 20% de 540 élèves ?

$$\frac{20}{100} \times 540 = 108 \text{ élèves.}$$

Car

Nombre	%
	20 %
540	100 %

Calculer un pourcentage :

Quel pourcentage représentent 81 élèves dans un groupe de 540 ?

$$\frac{81}{540} \times 100 = 15\%$$

Car

Nombre	%
81	
540	100 %

### IV. DIAGRAMMES STATISTIQUES. REPRESENTATIONS

#### Par un diagramme en bâtons (ou en barres):

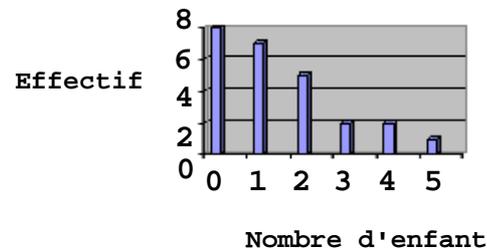
On l'utilise généralement pour des caractères qualitatifs ou quantitatifs discrets. (où le caractère ne peut prendre que des valeurs précises).

La hauteur des bâtons est proportionnelle à l'effectif correspondant. Les bâtons ne sont pas accolés les uns aux autres.

Ex : combien d'enfants dans la famille ?

Chaque valeur du nombre d'enfant est représenté par un bâton dont la longueur est proportionnelle au nombre d'enfant

Diagramme à bâtons



#### Par un diagramme à secteurs circulaires (ou camembert)

On l'utilise indifféremment pour des caractères qualitatifs ou quantitatifs.

Un diagramme à secteurs circulaires est composé de secteurs dont la mesure de l'angle est proportionnelle à l'effectif (ou à la fréquence).

Ex : Quelle est votre situation familiale ?

Déterminons l'angle représentant le nombre de personnes célibataire.

**Le disque entier (360°) représente l'effectif total (25)**

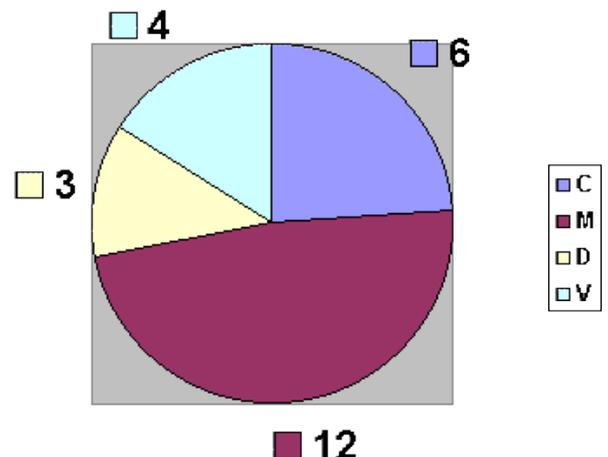
Nous cherchons par quel angle (x) sera représenté l'effectif des célibataires : 6

**tableau de proportionnalité:**

effectif	25	6
angle	360	x

**Donc (produit en croix) :**  $\frac{360 \times 6}{25} = \frac{72 \times 6}{5} = 86,4^\circ$

On fait la même chose avec les effectifs de mariés (M), de divorcés (D) et veufs (V)



## Par un histogramme:

L'histogramme est plutôt employé pour des caractères quantitatifs continus (**où le caractère peut prendre n'importe quelle valeur dans un intervalle donné**).

L'aire des rectangles est proportionnelle à l'effectif de chaque intervalle (appelé aussi classe).

Chaque rectangle est accolé au précédent.

Ex : quel âge avez-vous?

Ici, l'histogramme représente les classes d'âge, où on impose:

- 0,5 cm de haut représente 1 personne.
- une classe a pour largeur 2 cm

Il reste donc à déterminer les longueurs des rectangles

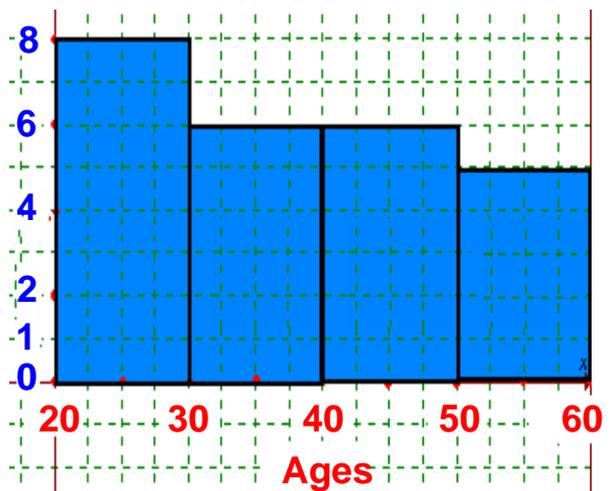
On peut s'aider en remplissant le tableau :

classe	effectif	largeur	hauteur
[20 ; 30[	8	2 cm	4 cm
[30 ; 40[	6	2 cm	3 cm
[40 ; 50[	6	2 cm	3 cm
[50 ; 60[	5	2 cm	2,5 cm

[30 ; 40[ : signifie personnes d'age compris entre 30 et 40 ans, (30 ans inclus et 40 ans exclu).

Le 40 ans sera compté dans la classe : [40 ; 50[

### Personnes



## DIFFERENTS MODES DE CALCULS D'UNE MOYENNE

### MOYENNES SIMPLES :

Un élève a obtenu les notes suivantes au bac : (sans tenir compte de coefficients)

Matière	Français	Mathématiques	Histoire	Anglais	Espagnol
Note	12	10	11	8	5

Si on calcule la moyenne « simple », on trouve :

$$M = \frac{12 + 10 + 11 + 8 + 5}{5} = 9,2 : \text{L'élève échoue à son bac !}$$

### MOYENNES PONDÉRÉES :

Un élève a obtenu les notes suivantes au bac : (avec coefficients)

Matière	Français	Mathématiques	Histoire	Anglais	Espagnol
Note	12	10	11	8	5
Coefficient	4	4	2	2	1

Si on calcule la **moyenne pondérée** (c'est à dire en appliquant un « poids », un coefficient à chaque note) on trouve :

$$M = \frac{12 \times 4 + 10 \times 4 + 11 \times 2 + 8 \times 2 + 5 \times 1}{4 + 4 + 2 + 2 + 1} = \frac{131}{13} = 10,1 : \text{L'élève obtient son bac !}$$

### Remarque :

On veut calculer une approximation de la taille moyenne des élèves de groupe, mais on a regroupé les tailles en classes.

Taille	$1,20 \leq T < 1,30$	$1,30 \leq T < 1,40$	$1,40 \leq T < 1,50$	$1,50 \leq T < 1,60$	$1,60 \leq T < 1,70$
Effectif	5	7	13	9	6

Autre écriture des classes :

Taille	[1.20 ; 1.30[	[1.30 ; 1.40[	[1.40 ; 1.50[	[1.50 ; 1.60[	[1.60 ; 1.70[
Effectif	5	7	13	9	6

Il faut remplacer chaque classe par son centre :

Taille	1,25	1,35	1,45	1,55	1,65
Effectif	5	7	13	9	6

Calculons maintenant la moyenne pondérée :

$$M = \frac{1,25 \times 5 + 1,35 \times 7 + 1,45 \times 13 + 1,55 \times 9 + 1,65 \times 6}{5 + 7 + 13 + 9 + 6} = \frac{58,4}{40} = 1,46 \text{ m}$$

# MOYENNE, MEDIANE et ETENDUE

Cette série statistique porte sur l'âge des joueurs de l'équipe de France championne d'Europe en 2000 :

Age	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	TOTAL
Effectif	0	1	2	0	1	0	4	0	4	3	1	2	2	0	1	1	22

Ce qui signifie qu'il y a :

- 1 joueur âgé de 21 ans.
- 2 joueurs âgés de 22 ans.
- 4 joueurs âgés de 26 ans.
- Aucun joueur âgé de 33 ans...

## a. Valeur MOYENNE de la série statistique :

C'est la somme de tous les âges (en n'oubliant pas de prendre en compte l'effectif de chaque âge), divisé par l'effectif total :

$$M = \frac{(21 \times 1) + (22 \times 2) + (24 \times 1) + (26 \times 4) + (28 \times 4) + (29 \times 3) + (30 \times 1) + (31 \times 2) + (32 \times 2) + (34 \times 1) + (35 \times 1)}{22}$$

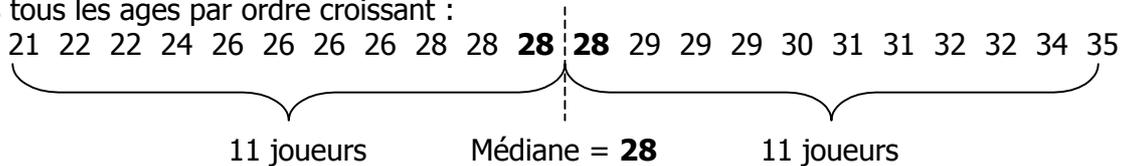
$$M = \frac{21 + 44 + 24 + 104 + 112 + 87 + 62 + 30 + 64 + 34 + 35}{22}$$

$$M = \frac{617}{22} \approx 28$$

La moyenne d'âge de l'équipe est donc d'environ 28 ans.

## b. Valeur MEDIANE de la série statistique :

C'est la valeur (de l'âge) qui se trouve au MILIEU de la série, qui la partage en deux séries d'effectif égal. Réécrivons tous les âges par ordre croissant :



La médiane de cette série statistique est de 28 ans.

## Remarques :

- Dans cet exemple, la médiane est égale à la moyenne. C'est un hasard, même si ces deux valeurs sont généralement très proches.
- Dans le cas où l'effectif de la série est impair, la « ligne de partage » est située juste sur une valeur : C'est la valeur médiane.
- Dans le cas où l'effectif de la série est pair (dans notre exemple), la « ligne de partage » est située juste entre deux valeurs de la série. Si ces deux valeurs sont différentes, on prend leur moyenne pour valeur médiane.

## c. Étendue d'une série statistique :

C'est la différence entre la valeur (de l'âge) la plus haute et la valeur la plus basse.

$$35 - 21 = \mathbf{14 \text{ ans}}$$