STATISTIQUES EN SECONDE

I. ORGANISATION ET REPRÉSENTATION DES DONNÉES

DÉFINITIONS

- Les statistiques permettent d'étudier un caractère d'une population.
- Le nombre d'éléments de la population s'appelle l'**effectif global** (ou l'**effectif total**).
- Pour une valeur de caractère donnée, l'effectif est le nombre d'éléments correspondant à cette valeur.
- Une **série statistique** est un tableau donnant les effectifs pour chacune des valeurs possibles du caractère.

EXEMPLE 1

On fait une étude portant sur l'âge des élèves d'un lycée.

- le caractère étudié est l'âge
- la **population** est l'ensemble des élèves du lycée
- l'effectif global est le nombre d'élèves du lycée
- le tableau ci-dessous est la **série statistique** pour ce caractère dans un lycée donné :

âges (en années)	14	15	16	17	18	19	20
effectifs	3	22	65	82	59	35	2

EXEMPLE 2 : CRÉATION D'UN TABLEAU POUR UNE SÉRIE STATISTIQUE

On suppose que les notes à un contrôle dans une classe de 21 élèves sont les suivantes :

Ces données brutes sont assez peu pratiques à utiliser sous cette forme (notamment lorsqu'il y a beaucoup de valeurs).

Pour commencer on commence à trier les notes de la plus petite à la plus grande :

Ensuite, on va créer le tableau de cette série en indiquant pour chaque note son effectif c'est à dire le nombre d'élèves ayant obtenu cette note :

I	notes	2	3	5	6	8	9	12	13	14	15	16	17	18	19
	effectifs	1	1	2	1	2	2	2	3	2	1	1	1	1	1

II - MÉDIANE - QUARTILES

DÉFINITION

La **médiane** d'une série statistique est la valeur du caractère qui partage la population en deux classes de même effectif.

REMARQUE

En pratique pour trouver la médiane d'une série statistique d'effectif global n:

- On ordonne les valeurs du caractère dans l'ordre croissant.
- Si n est pair, la médiane sera la moyenne des valeurs du terme de rang $\frac{n}{2}$ et du terme de rang $\frac{n}{2}+1$.
- Si *n* est impair, la médiane sera la valeur du terme de rang $\frac{n+1}{2}$.
- Lorsque l'effectif global est élevé, il est souvent utile de calculer les effectifs cumulés pour trouver cette valeur.

EXEMPLE

Reprenons l'exemple 2 ci-dessus.

Dans cet exemple, c'est la 11ème note $(11 = \frac{21+1}{2})$ qui est la médiane. En effet, il y a 10 notes au dessous et 10 notes au dessus :

La médiane est donc 12.

Supposons qu'il n'y ait que 20 élèves (on enlève l'élève qui a eu 2) :

Il n'y a plus ici de note située "juste au milieu".

Si on choisit la 10ème note (qui est 12) il y a 9 notes en dessous et 10 notes au dessus.

Si on choisit la 11ème note (qui est 13) il y a 10 notes en dessous et 9 notes au dessus.

Dans ce cas, on prend comme médiane la moyenne de 12 et de 13 c'est à dire 12,5.

La médiane est donc 12,5.

DÉFINITIONS

- Le **premier quartile** Q1 d'une série statistique est la plus petite valeur des termes de la série pour laquelle au moins un quart des données sont inférieures ou égales à Q1.
- Le **troisième quartile** Q3 d'une série statistique est la plus petite valeur des termes de la série pour laquelle au moins trois quarts des données sont inférieures ou égales à Q3.

EXEMPLE

Reprenons l'exemple des notes ci-dessus (avec 21 élèves).

Pour le **premier quartile** il faut qu'il y ait au moins 1/4 des notes qui soient inférieures ou égales. $1/4 \times 21 = 5,25$. Le premier quartile est donc la 6ème note.

```
2; 3; 5; 5; 6; 8; 8; 9; 9; 12; 12; 13; 13; 13; 14; 14; 15; 16; 17; 18; 19
```

le premier quartile est 8.

Pour le **troisième quartile** il faut qu'il y ait au moins 3/4 des notes qui soient inférieures ou égales.3/4×21=15,75.

Le troisième quartile est donc la 16ème note.

2; 3; 5; 5; 6; 8; 8; 9; 9; 12; 12; 13; 13; 13; 14; 14; 15; 16; 17; 18; 19

le troisième quartile est 14.