

I. Vocabulaire de base

- Conduire une **Étude statistique** : C'est étudier comment sont répartis certains aspects, certaines particularités, d'un groupe.
 - ⇒ étude de la répartition des moyennes de maths des élèves de 3^e de Haute-Savoie ;
 - ⇒ étude de la répartition des couleurs des voitures françaises de moins de 15 000 € ; ...
- **Population** : C'est le groupe sur lequel porte l'étude. Ce peut être un groupe de personnes, d'animaux, d'objets. La population est composée d'**individus**.
 - ⇒ les élèves de 3^e de Haute-Savoie ;
 - ⇒ les grenouilles du lac de Thyez ;
 - ⇒ les voitures de moins de 15 000 € ; ...
- **Caractère** : C'est la particularité, la *caractéristique* qu'on étudie. Le caractère peut être :
 - quantitatif**, s'il correspond à une valeur numérique,
 - ⇒ âge ; moyenne en Maths ; taille ; nombre d'absences ; consommation aux 100 km ; ...
 - qualitatif**, si on ne peut pas le *quantifier* (i.e. lui attribuer une valeur numérique).
 - ⇒ sport préféré ; ville du domicile ; couleur de la carrosserie ; ...
- **Série statistique** : C'est la liste des résultats obtenus après recueil des données. Pour conduire l'étude, les données doivent être classées et présentées de façon claire. Dans le cas d'un grand nombre de données, un traitement informatique s'impose.
- **Effectif** d'une valeur: C'est le nombre de fois que la valeur apparaît dans la série statistique, c'est-à-dire le nombre d'individus pour qui le caractère correspond à cette valeur.
 - ⇒ élèves de 3^e de haute savoie de 15 ans : 2 000 ;
 - ⇒ voitures françaises de couleur blanche : 400 000 ;
 - ⇒ grenouilles de deux ans : 700 ; ...
- **Fréquence** : C'est le quotient de l'effectif d'une valeur par l'effectif total. On l'exprime souvent en pourcentage, en multipliant la valeur trouvée par 100.

$$\Rightarrow \text{fréquence} = \frac{\text{effectif}}{\text{effectif total}} ; \text{fréquence(en \%)} = \frac{\text{effectif}}{\text{effectif total}} \times 100$$

On arrondit en général la fréquence au millième près (ou à 0,1 % près).

 - ⇒ élèves absents moins de 5 demi-journées : 38 %
 - ⇒ grenouilles atteignant 5 ans : 26 %

II. Indicateurs statistiques

1. Indicateurs de position : moyenne et médiane

- La **moyenne** d'une série statistique est obtenue en divisant la somme des valeurs du caractère par le nombre total de valeurs.

- Exemple : pour la série de notes suivante : 10 ; 8 ; 12 ; 13 ; 15 ; 10 ; 8 ; 13 ; 12 ; 10 ; 10 ; 13

$$M = \frac{10+8+12+13+15+10+8+13+12+10+10+13}{12} \approx$$

- Si (comme ici) certaines valeurs apparaissent plusieurs fois, on peut utiliser les effectifs. La même série de notes peut être présentée ainsi :

note	8	10			
effectif	2				

Et la moyenne se calcule ainsi :

$$M = \frac{2 \times 8 + \quad \times 10 + \quad \times \quad + \quad \times \quad + \quad \times}{\quad + \quad + \quad + \quad +} \approx$$

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

La série précédente, ordonnée, s'écrit : 8 ; 8 ; 10 ;

Sa médiane est _____.

$$\begin{array}{c} \text{au moins 25\%} \\ \text{des effectifs} \\ \underbrace{8; 8; 10} \end{array} \cdot \begin{array}{c} \text{médiane} \\ \downarrow \\ 10; 10; 10 \end{array} \cdot \begin{array}{c} \underbrace{12; 12; 13} \\ \text{au moins 75\%} \\ \text{des effectifs} \end{array} \cdot 13; 13; 15$$

- On définit de même les **quartiles** d'une série statistique :

Le premier quartile est _____, le troisième quartile est _____.

2. Indicateur de dispersion : l'étendue

- L'**étendue** d'une série statistique est la différence entre les deux valeurs extrêmes du caractère.

- Étendue de la série précédente :