

Vocabulaire

Lorsqu'on réalise une enquête, on est amené à étudier des **caractères** (propriétés) propres à chaque **individu**.

L'ensemble des individus est appelé la **population**.

Le caractère peut être **quantitatif** : nombres (la taille, l'âge, le temps passé devant l'ordinateur...) ou **qualitatif** sinon (la couleur des cheveux, les sports pratiqués...).

Le nombre total d'individus de la population est appelé **effectif total**.

- **Effectif d'une valeur** : nombre de fois où cette valeur apparaît dans la liste.
- **Série de données** : ensemble des valeurs et des effectifs.

Indicateur de position : la moyenne

A quoi elle sert : indicateur le plus simple pour résumer l'information fournie par un ensemble de données statistiques.

Interprétation : toutes les valeurs auraient pu lui être égales, on aurait eu la même somme totale

La **moyenne dite « simple »** d'une série de valeurs est le nombre obtenu en additionnant ces valeurs et en divisant le résultat par le nombre de valeurs.

$$M = \frac{\text{somme des données}}{\text{effectif total}}$$

La **moyenne pondérée** d'une série de valeurs est le nombre obtenu en additionnant les produits de ces valeurs par leurs coefficients et en divisant le résultat par la somme des coefficients.

$$M = \frac{\text{somme des produits des données par leurs effectifs}}{\text{effectif total}}$$

Fréquence

On appelle **fréquence d'une valeur** le quotient de son effectif par l'effectif total.

$$\text{Fréquence} = \frac{\text{Effectif}}{\text{Effectif total}}$$

Remarque :

- c'est un nombre entre 0 et 1
- la somme des fréquences d'une série statistique vaut toujours 1.

Gestion des données

Indicateur de position : la médiane

La médiane est nombre qui **partage une série** rangée dans l'**ordre croissant en deux** séries de **même effectif**.

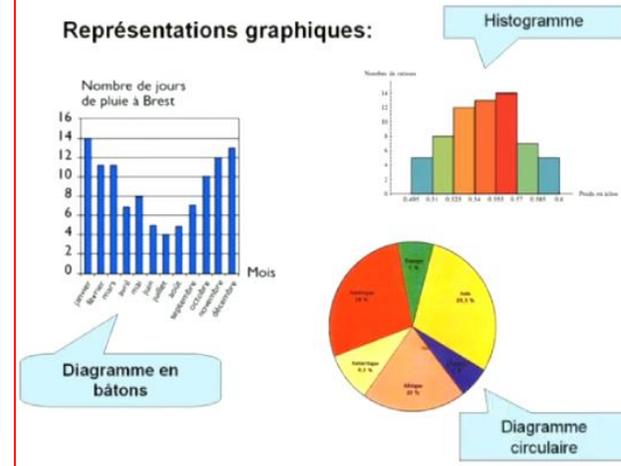
A quoi elle sert : donner des informations sur la répartition des valeurs dans la série.

Interprétation : au moins 50% de valeurs sont situées sous cette valeur médiane

Pour déterminer une **médiane** il faut :

- trier les données par ordre croissant
- compter le nombre de valeurs de la série :
 - si le nombre est **impair**, alors on prend pour médiane : le nombre correspondant à la valeur centrale de cette série
 - si le nombre est **pair**, alors on prend pour médiane : choisit souvent la demi-somme des deux « valeurs centrales ».

Représentations graphiques:



Indicateur de dispersion : l'étendue

L'étendue d'une série statistique est la **différence** entre la valeur la plus grande et la valeur la plus petite de cette série.

Interprétation : c'est l'écart maximum entre les valeurs de la série

Série regroupée en classe et histogramme

Parfois, on décide de regrouper les données en classes c'est-à-dire en regroupant les valeurs par groupe de même amplitude.

Dans un histogramme, l'**aire** des rectangles doit être **proportionnelle** aux effectifs. Comme ici les classes ont la même amplitude, la hauteur des rectangles est proportionnelle aux effectifs.