

Séquence n°12
ORGANISATION DE DONNÉES : STATISTIQUES

I. Introduction : de l'importance des statistiques actuellement ?

Exemples de questions d'actualités :

- **Climat** : assiste-t-on à un réchauffement de la planète ?
- **Santé** : faut-il encore vacciner les enfants contre la variole ?
- **Paris sportifs** : une activité à risques ?
- **Qualité industrielle** : comment faire pour être « sûr » que dans un lot de 1000 piles électriques vendues, il y en a au moins 995 qui fonctionnent correctement ?
- **Météo** : fera-t-il beau dimanche ?
- **Population** : quel pays aura le plus d'habitant en 2020 ?
- **Jeux** : un dé qui affiche 241 fois la face « 6 » sur 1 000 lancers est-il truqué ?

II. Organiser des données dans un tableau : effectifs et fréquences

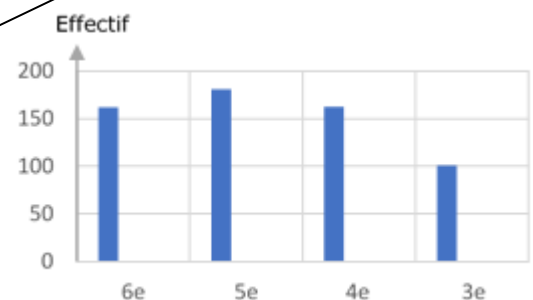
1. Effectifs

Le tableau ci-contre donne la répartition des **effectifs** des élèves dans un collège dont l'**effectif total** est de 607 élèves.

| | 6 ^{ème} | 5 ^{ème} | 4 ^{ème} | 3 ^{ème} | Total |
|-----------|------------------|------------------|------------------|------------------|-------|
| Effectifs | 162 | 181 | 163 | 101 | 607 |

Exemple L'effectif des 5^{èmes} est de 181 élèves.

Graphique On peut représenter ces effectifs par un diagramme en barres où la hauteur de chaque barre est proportionnelle à l'effectif qu'elle représente.



2. Fréquences (et pourcentages)

Définition

$$\text{fréquence} = \frac{\text{effectif}}{\text{effectif total}}$$

| | 6 ^{ème} | 5 ^{ème} | 4 ^{ème} | 3 ^{ème} | Total |
|------------|------------------|------------------|------------------|------------------|-------|
| Effectifs | 162 | 181 | 163 | 101 | 607 |
| Fréquences | 0,267 | 0,298 | 0,268 | 0,167 | 1 |

Exemple

La fréquence des élèves de 3^{ème} dans ce collège est 0,167 environ car $\frac{101}{607} \approx 0,167$

Remarques

× Une fréquence est un nombre compris entre 0 et 1.

×
$$\text{fréquence en \%} = \frac{\text{effectif}}{\text{effectif total}} \times 100$$

On obtient alors :

| | 6 ^{ème} | 5 ^{ème} | 4 ^{ème} | 3 ^{ème} | Total |
|-------------------|------------------|------------------|------------------|------------------|-------|
| Effectifs | 162 | 181 | 163 | 101 | 607 |
| Fréquences (en %) | 26,7 | 29,8 | 26,8 | 16,7 | 100 |

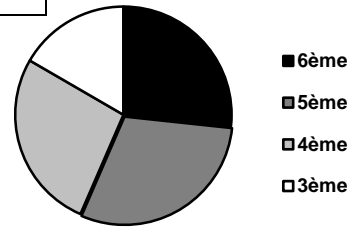
Exemple

Le pourcentage de 4^{ème} dans ce collège est 26,8 % environ car $\frac{163}{607} \times 100 \approx 26,8$

Graphique On peut représenter ces effectifs ou fréquences par un diagramme circulaire où la mesure de chaque angle est proportionnelle à l'effectif ou la fréquence qu'il représente.

$$\text{Mesure d'un angle (en degrés)} = \frac{\text{effectif}}{\text{effectif total}} \times 360$$

| | 6 ^{ème} | 5 ^{ème} | 4 ^{ème} | 3 ^{ème} | Total |
|---------------------|------------------|------------------|------------------|------------------|-------|
| Effectifs | 162 | 181 | 163 | 101 | 607 |
| Mesures (en degrés) | 96 | 107 | 97 | 60 | 360 |



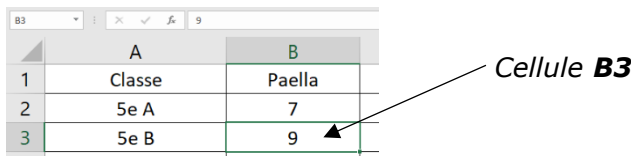
Exemple L'angle représentant les 5^{èmes} mesure 107° environ car $\frac{181}{607} \times 360 \approx 107$

3. Utiliser un tableur pour organiser des données

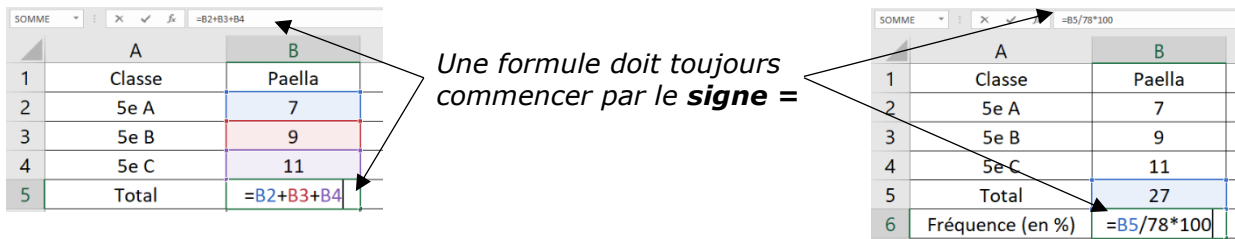
Un tableur est un logiciel qui permet d'effectuer un grand nombre de **calculs répétitifs**. Il se présente sous forme de plusieurs tableaux appelés « **feuilles de calcul** » : il sert ainsi à organiser et à trier des données. Chaque case de ce tableau est appelée « **cellule** ».

Avec un tableur, on peut :

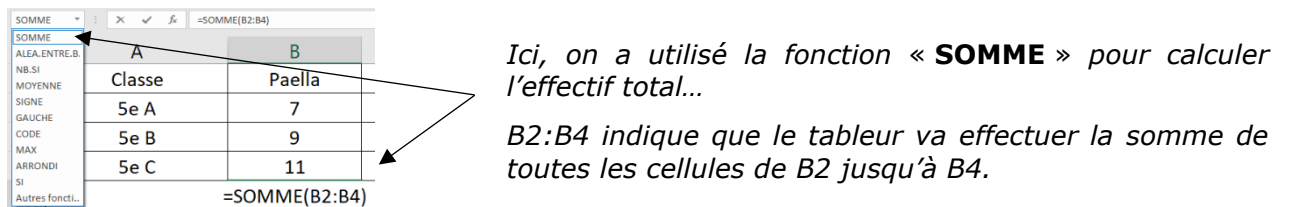
- Repérer une **cellule** grâce à la colonne et la ligne correspondantes :



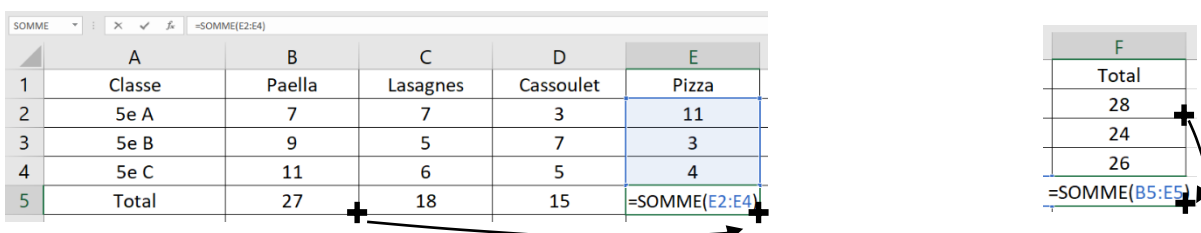
- Effectuer un calcul à partir d'une **formule** qui dépend d'autre(s) cellule(s) :



- Utiliser des « **fonctions** » du tableur pour effectuer des calculs :



- Utiliser la **poignée de recopie automatique** pour effectuer des **calculs répétitifs** :



Le tableur a modifié automatiquement la formule de B5 jusqu'à E5.

Le tableur a modifié automatiquement la formule de F2 jusqu'à F5.

III. Graphiques : comment choisir une bonne représentation ?

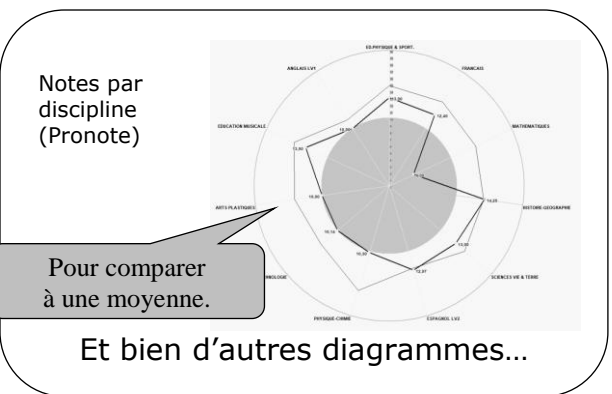
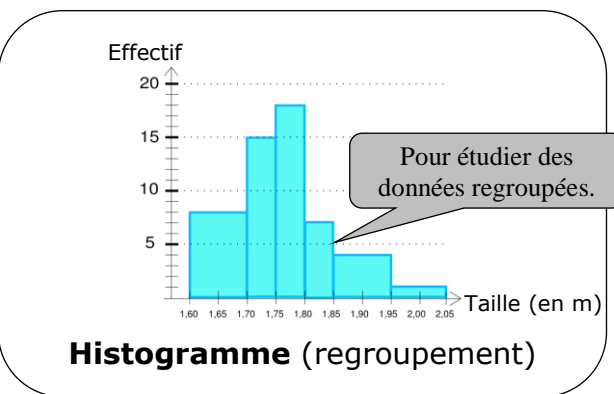
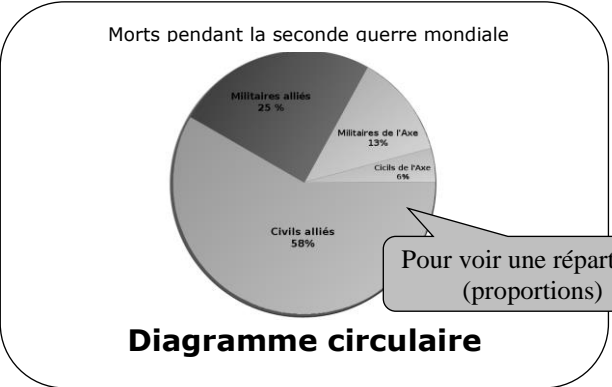
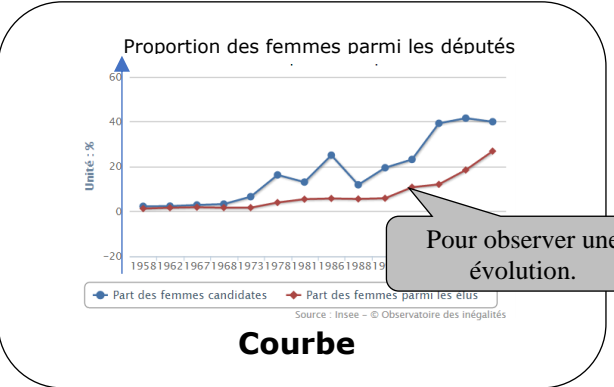
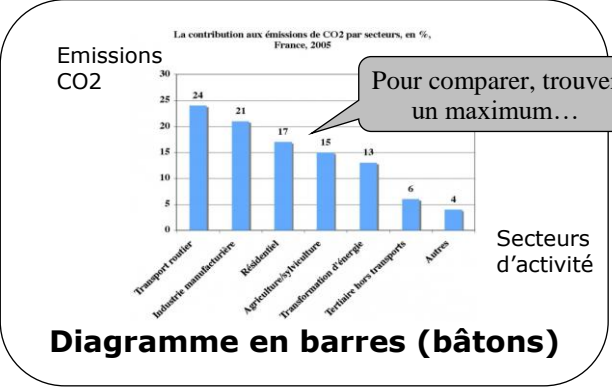
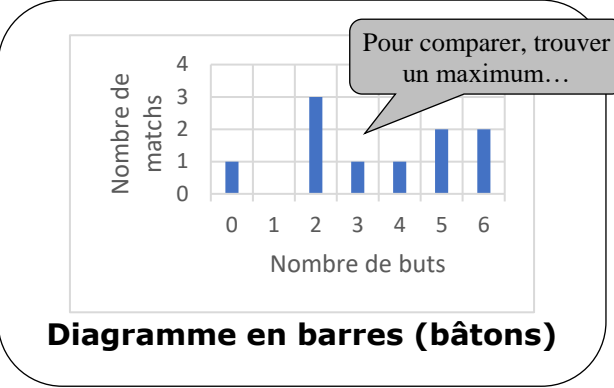
On peut représenter des données avec plusieurs graphiques selon :

- ce que l'on souhaite observer : le **caractère** ;
- ce que l'on souhaite montrer grâce au graphique : pour comparer des **effectifs**, pour répartir des **fréquences** ou pour **regrouper des données** (histogramme).

Quel type de **caractère** dois-je étudier et observer ?

Caractères numériques
 (nombres ordonnés > axe des abscisses)
 Exemples : âge, temps ou durée, masse, notes d'un contrôle, prix ou salaires, etc.

Caractères qualitatifs
 Exemples : couleur des yeux, boisson préférée, langue parlée, secteurs d'activités, jours de la semaine, etc.



IV. Caractéristiques de position d'une série statistique

1. Moyenne pondérée

Définition La **moyenne** d'une série de données est égale à la somme de toutes les données divisée par l'effectif total de la série.

Remarque Attention, la moyenne doit prendre en compte toutes les données... Et donc prendre en compte tous les effectifs !

EXERCICE TYPE 1 Déterminer la taille moyenne pour les 10 personnes suivantes

| | | | | | |
|---------------|------|------|------|------|-------|
| Taille (en m) | 1,70 | 1,75 | 1,80 | 1,85 | Total |
| Effectif | 3 | 4 | 2 | 1 | 10 |

Solution $1,70 \times 3 + 1,75 \times 4 + 1,80 \times 2 + 1,85 \times 1 = 17,55$
 $17,55 \div 10 = 1,755$

La taille moyenne de ces 10 personnes est **environ 1,76 m.**

Attention, la moyenne doit prendre en compte toutes les données...
Il faut **prendre en compte tous les effectifs !**

2. Découvrir la médiane d'une série

Définition La **médiane** d'une série ordonnée est **une valeur** telle qu'il y ait **autant de valeurs inférieures** ou égales **que de valeurs supérieures** ou égales.

EXERCICE TYPE 2 Déterminer les médianes des séries de notes suivantes.

- série A : 13, 13, 20, 15, 18, 19, 15

- série B : 17, 14, 3, 16, 5, 17

Solution Pour déterminer une médiane, **il faut d'abord ordonner la série.**

- série A : $13 \leq 13 \leq 15 \leq \boxed{15} \leq 18 \leq 19 \leq 20$.
 $\xleftarrow{3 \text{ notes}} \quad \xrightarrow{3 \text{ notes}}$

La médiane de cette série A est **15**.

- série B : $3 \leq 5 \leq \boxed{14 \leq 16} \leq 17 \leq 17$.
 $\xleftarrow{3 \text{ notes}} \quad \xrightarrow{3 \text{ notes}}$

La médiane de cette série est entre 14 et 16. Par habitude, on prendra alors la valeur centrale : la médiane de cette série B est donc **15**.

Remarque **La moyenne et la médiane ne sont pas identiques !**

Dans l'exemple ci-dessus, les séries A et B ont la même médiane mais pas la même moyenne :

Moyenne de la série A : $(13+13+20+15+18+19+15) \div 7 \approx 16,1$

Moyenne de la série B : $(17+14+3+16+5+17) \div 6 = 12$