

1 Notions de base

1.1 Contexte

Les statistiques descriptives ont pour but de décrire de manière synthétiques les caractéristiques d'un ensemble d'individus. Les individus peuvent être des personnes, mais aussi des objets ou des faits.

1.2 Vocabulaire

Population : ensemble de tous les individus que l'on souhaite étudier.
Chaque élément d'une population s'appelle une **unité statistique**.

Echantillon : ensemble des individus que l'on interroge, ou dont on connaît les informations.

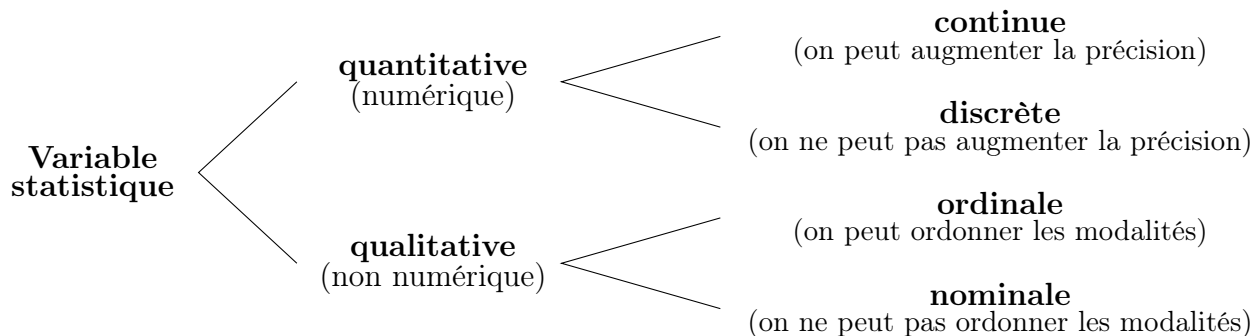
Si l'échantillon est égal à la population entière, on parle d'un **recensement**. Si l'échantillon n'est formé que d'une partie de la population, on parle d'un **sondage**.

Variable statistique : caractéristique que l'on souhaite étudier.

Une variable statistique est dite **quantitative** si les valeurs qu'elle peut prendre sont des nombres. Dans le cas contraire, elle est **qualitative**, et ses valeurs s'appellent des **modalités** ou des **catégories**.

Si les valeurs que peut prendre une variable statistique quantitative peuvent être listées (par exemple 0, 1, 2, 3, ...) et leur précision ne peut pas être augmentée, on dit que la variable est **discrète**. Si la précision des valeurs peut être modifiée (par exemple 2.3 ou 2.31 ou 2.314), on dit qu'elle est **continue**.

Si les valeurs que peut prendre une variable statistique qualitative peuvent être ordonnées (mise dans un certain ordre, par exemple *jamais, parfois, souvent, toujours*), la variable est dite **ordinaire**. Dans le cas contraire, elle est **nominale**.



Exemples

- a) On demande à 200 passagers pris au hasard dans un aéroport de donner leur nationalité.

Population :
Echantillon :
Unité statistique :
S'agit-il d'un recensement ou d'un sondage ?
Variable statistique :
Type de variable :
Modalités / catégories ou valeurs :

- b) Durant tout le mois d'avril, on mesure la température maximale de la journée à la Tour-de-Peilz.

Population :
Echantillon :
Unité statistique :
S'agit-il d'un recensement ou d'un sondage ?
Variable statistique :
Type de variable :
Modalités / catégories ou valeurs :

- c) On demande à 5'000 familles résidant en Suisse le nombre d'enfants dans leur foyer.

Population :
Echantillon :
Unité statistique :
S'agit-il d'un recensement ou d'un sondage ?
Variable statistique :
Type de variable :
Modalités / catégories ou valeurs :

- d) A la fin d'une période de révision, un enseignant demande à ses élèves de répondre à la question suivante :

Cette période de révision a-t-elle été utile ? *pas du tout* *un peu* *beaucoup*

Population :
Echantillon :
Unité statistique :
S'agit-il d'un recensement ou d'un sondage ?
Variable statistique :
Type de variable :
Modalités / catégories ou valeurs :

1.3 Echelles de mesure

On distingue quatre types d'échelles de mesure.

Echelle nominale

Exemple : 0 (homme) 1 (femme)

- s'utilise pour une variable qualitative nominale
- sert uniquement à différencier les catégories d'une variable ($=$, \neq)
- ne permet pas d'établir une relation d'ordre ($<$, $>$)
- ne permet pas d'effectuer d'opération arithmétique ($+$, $-$, \cdot , $:$)

Echelle ordinale

Exemple : 0 (pas du tout) 1 (un peu) 2 (beaucoup)

- s'utilise pour une variable qualitative ordinale ou quantitative
- sert à différencier les catégories d'une variable ($=$, \neq)
- permet d'ordonner les catégories ($<$, $>$)
- ne permet pas d'effectuer d'opération arithmétique ($+$, $-$, \cdot , $:$)

Echelle d'intervalle

Exemple : les degrés Celsius

- s'utilise pour une variable quantitative
- sert à différencier les valeurs d'une variable ($=$, \neq)
- permet d'ordonner les valeurs ($<$, $>$)
- permet d'effectuer des additions / soustractions, mais pas des multiplications / divisions
- la valeur 0 ne signifie pas l'absence de la caractéristique

Echelle de rapport

Exemple : nombre d'enfants dans une famille : 0, 1, 2, ...

- s'utilise pour une variable quantitative
- sert à différencier les valeurs d'une variable ($=$, \neq)
- permet d'ordonner les valeurs ($<$, $>$)
- permet d'effectuer les quatre opérations arithmétiques ($+$, $-$, \cdot , $:$)
- la valeur 0 signifie une absence de la caractéristique mesurée

Exemple

Voici quatre manières différentes de mesurer la consommation de cigarettes dans un questionnaire :

1. Fumez-vous des cigarettes ?
1. Oui 2. Non

Echelle :

2. A quelle fréquence fumez-vous des cigarettes ?
0. Jamais 1. Rarement 2. Occasionnellement 3. Régulièrement

Echelle :

3. En moyenne, combien de cigarettes fumez-vous par jour ?
1. 0 2. De 1 à 5 3. de 6 à 10 4. de 11 à 20 5. Plus de 20

Echelle :

4. Combien de cigarettes fumez-vous par jour ?

.....

Echelle :

1.4 Tableau de distribution

Une fois les données récoltées, on les regroupe par modalité dans un tableau de distribution.

Exemples

- a) On a demandé à tous les élèves d'une classe quelle était leur matière préférée parmi les matières suivantes : français, anglais, maths et musique.

musique	français	français	anglais	français
anglais	musique	maths	musique	musique
musique	musique	musique	français	français
français	anglais	musique	anglais	maths

Tableau de distribution :

Répartition des selon

matière	nombre d'élèves	fréquence
Total		

- b) On a demandé à ces mêmes élèves leur dernière note d'anglais.

5	4.5	3.5	5	6	3.5	4	2.5	4	4.5
4	4.5	4.5	4.5	3	4	4.5	5	3.5	4

Tableau de distribution :

Répartition des selon

	fréquence
Total	

c) Enfin, on a demandé à ces élèves leur taille en centimètre.

172 157 162 156 167 179 173 173 178 160
 168 171 165 166 184 170 165 164 160 175

.....

Tableau de distribution :

Répartition des selon

Total		

Aide pour le choix du nombre de classes : Table de Sturges

Nombre de données	Nombre approximatif de classes
Entre 10 et 22	5
Entre 23 et 44	6
Entre 45 et 90	7
Entre 91 et 180	8
Entre 181 et 360	9
Entre 361 et 720	10

(Formule : $1 + \log_2(n)$)

2 Représentations graphiques

2.1 Représentation graphique d'une variable qualitative

Répartition des 20 élèves d'une classe selon leur matières préférées (exemple p.5).

Diagramme à rectangles

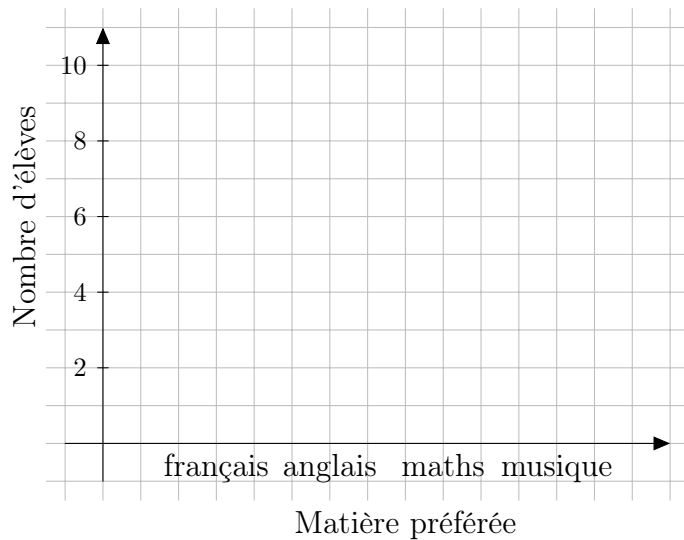


Diagramme circulaire

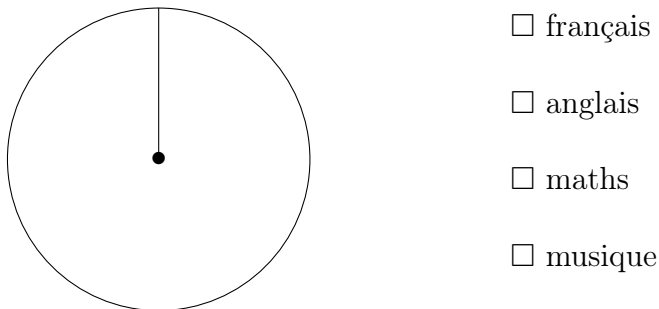
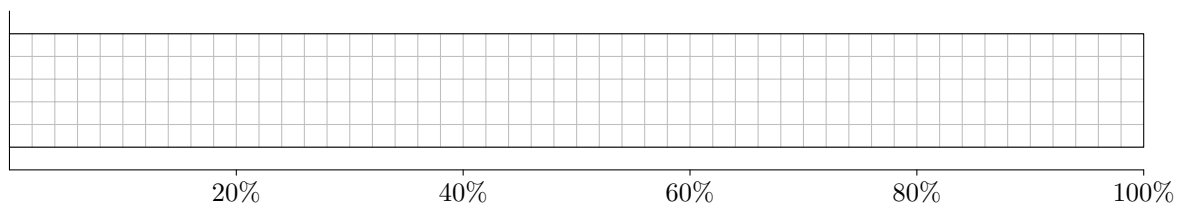


Diagramme linéaire

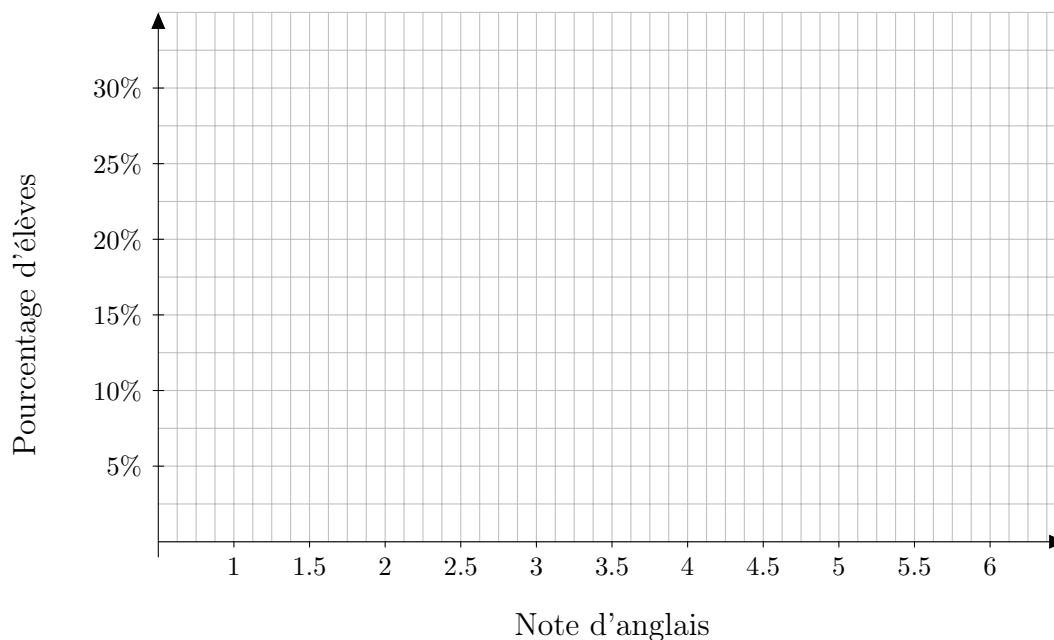
français anglais maths musique



2.2 Représentation graphique d'une variable quantitative discrète

Répartition des 20 élèves d'une classe selon leur note d'anglais (exemple p.5).

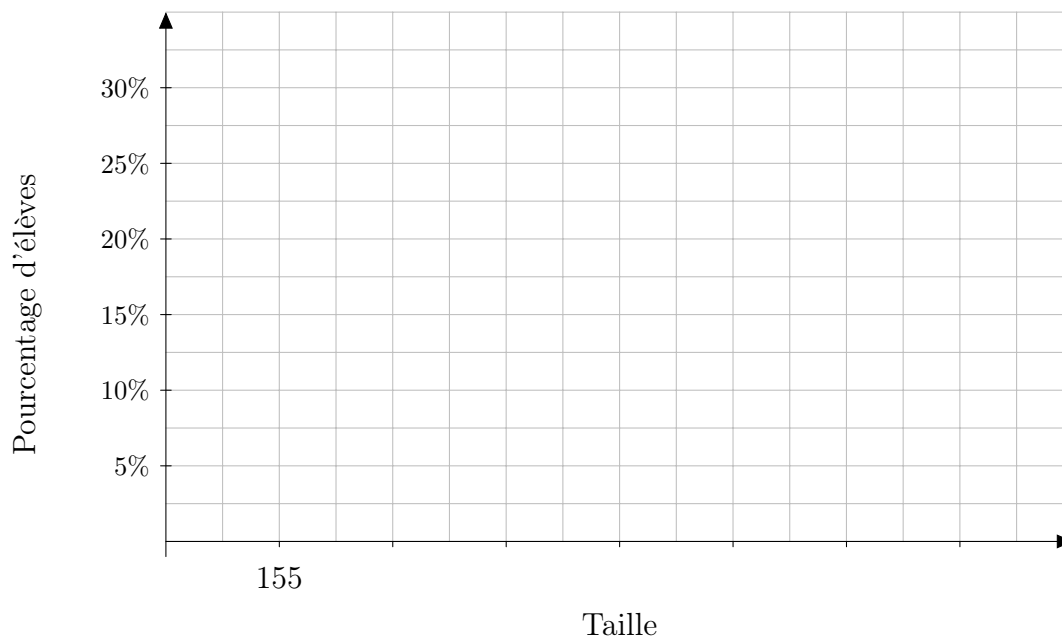
Diagramme en bâtons (ou rectangles)



2.3 Représentation graphique d'une variable quantitative continue

Répartition des 20 élèves d'une classe selon leur taille (exemple p.6).

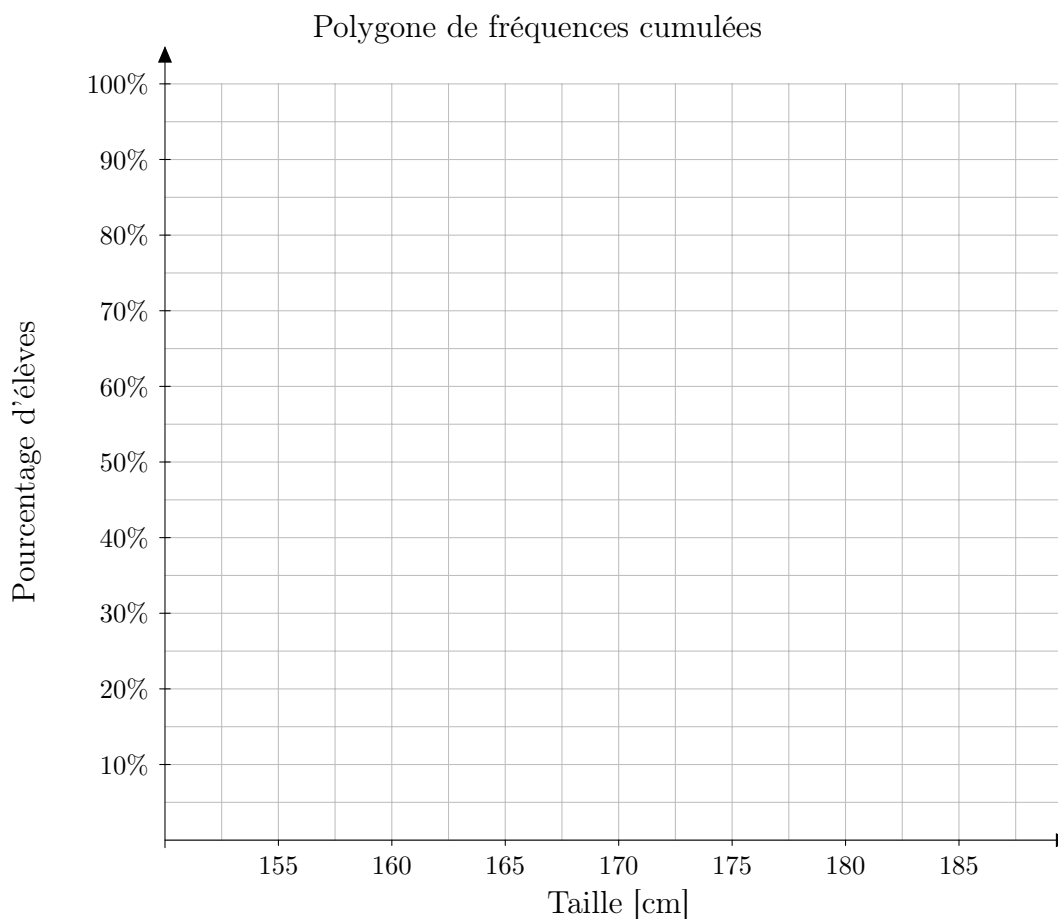
Histogramme et polygone de fréquences



Polygone de fréquences cumulées (ou courbe de fréquences cumulées)

Répartition des élèves selon leur taille

taille	effectif	fréquence	fréquence cumulée
[155 ; 160[2	10%	
[160 ; 165[4	20%	
[165 ; 170[5	25%	
[170 ; 175[5	25%	
[175 ; 180[3	15%	
[180 ; 185[1	5%	
Total	20	100%	



30% des élèves mesurent

.....% des élèves mesurent moins de 1.75 mètre.

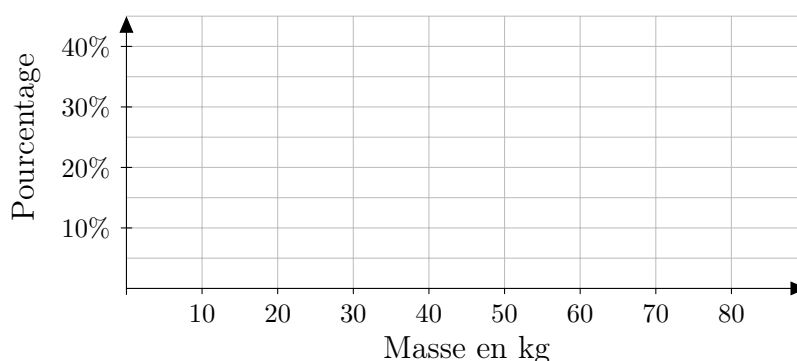
5% des élèves mesurent au moins

Histogramme dans le cas de classes inégales

On veut représenter par un histogramme la répartition des roches dans une rocaille de fleurs selon leur masse. Cette répartition est donnée par le tableau suivant :

masse en kg	pourcentage de roches
[10 ; 30[10%
[30 ; 40[25%
[40 ; 50[35%
[50 ; 80[30%
Total	100%

Histogramme obtenu en utilisant directement les fréquences données dans le tableau :

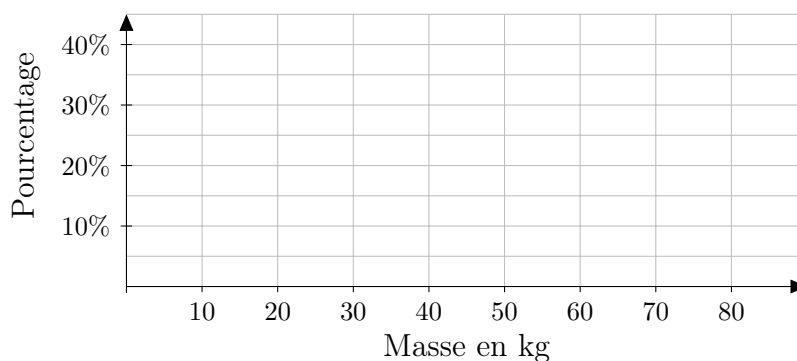


Cet histogramme donne-t-il une information correcte ou faussée?

Principe de proportionnalité des aires :

.....

Histogramme correct (respectant le principe de proportionnalité des aires) :



Remarque importante

Comme les pourcentages **réels** ne peuvent plus être lus directement sur le graphique, on les note au dessus de chaque rectangle.