

Analyse descriptive de données

Qu'est ce qu'une donnée ?

une donnée est un élément brut qui sert de point de départ à un raisonnement ayant pour objet la détermination d'une solution à un problème en relation avec cette donnée, ou bien, une description élémentaire d'une chose quelconque sans sens précis

Qu'est ce qu'une information ?

l'information, est une donnée qui a été transformée pour la rendre plus utile et significative.

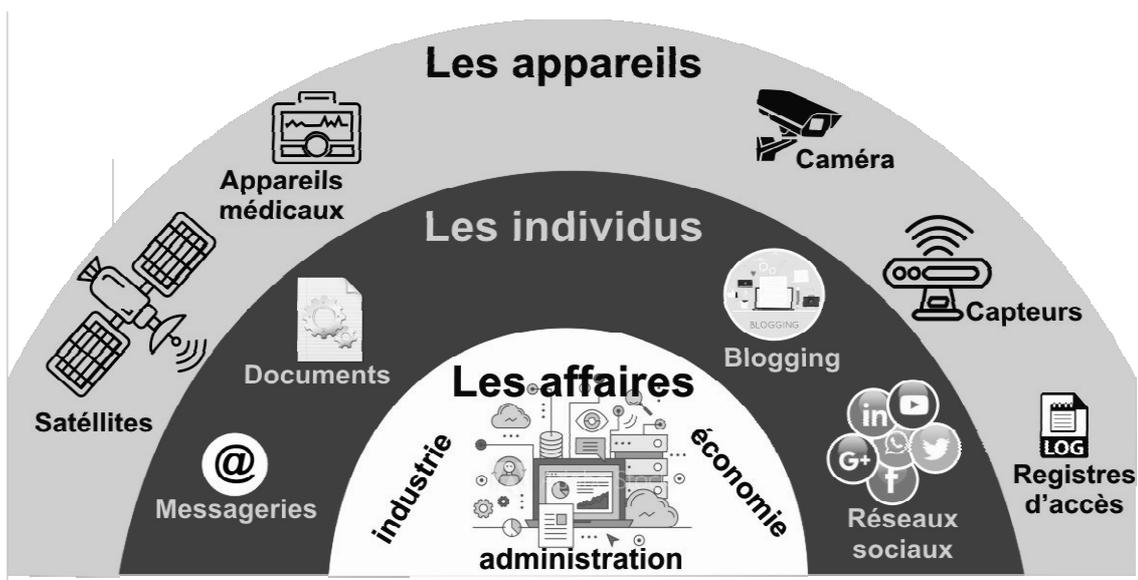
Exemple:

1. Enquête
2. Le mot "Mochila"

Une donnée devient information quand on lui associe un sens.

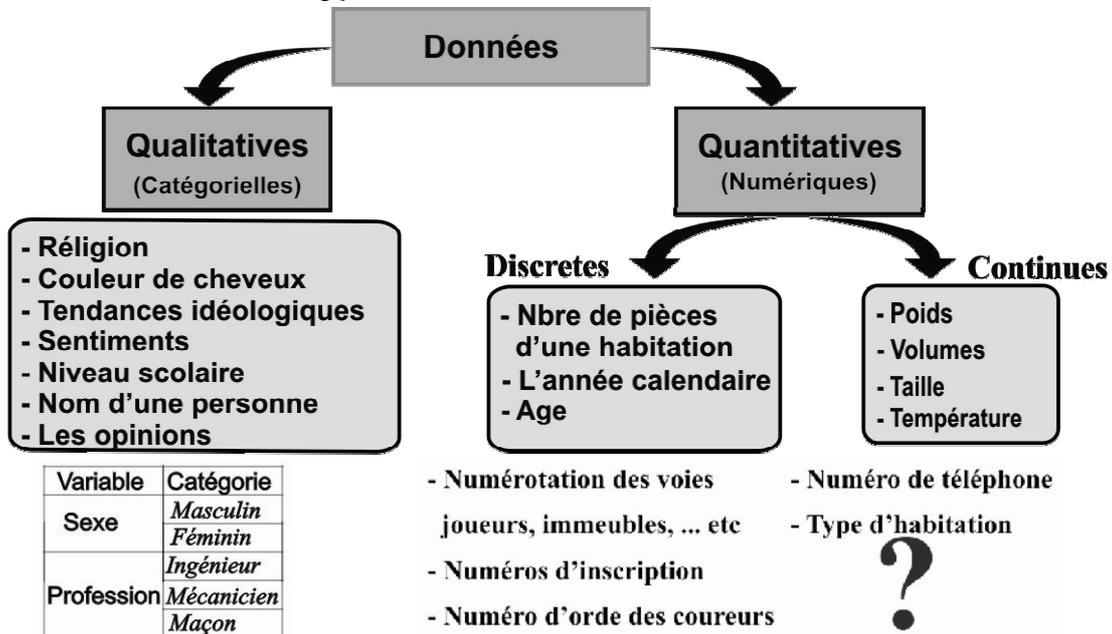
Analyse descriptive de données

Quelles sont les sources de données ?



Analyse descriptive de données

Quelles sont les types de données ?



Analyse descriptive de données

Qu'est ce qu'une variable statistique ?

Une variable est une caractéristique d'une unité statistique (un indicateur d'un phénomène) mesurable et/ou quantifiable qui peut prendre plusieurs valeurs (numérique ou une catégorie d'une classification) que ce soit par mesure ou par observation.

Exemple:

1. taille, poids, salaire, profession d'un groupe donné d'individus
2. rendement d'un ensemble de parcelles cultivées
3. température maximale et minimale, pluviométrie, ensoleillement, mesurés à un endroit donné durant un mois.

Analyse descriptive de données

Est ce que les informations sont tous kif kif ?

Bien sûr que NON Nature d'information

C'est ce que nous entendons par la nature de l'information?

Par exemple :

- Une distance de 4 kilomètres est deux fois plus longue que celle de 2 kilomètres.
- Un objet qui pèse 5 kilos a un tiers du poids d'un objet qui en pèse 15 kilos.
- Un coureur qui est arrivé premier n'a pas forcément fait trois fois mieux qu'un autre qui est arrivé troisième.
(ça diffère quand on parle du temps de course)
- Le poids total (somme) de deux objets qui pèsent 5 kilos et 10 kilos est égal à 15 kilos, mais la somme des temps de course des coureurs n'a aucun sens (ça diffère quand on parle du temps de travail ou temps d'arrêts de travail d'un bus).

Analyse descriptive de données

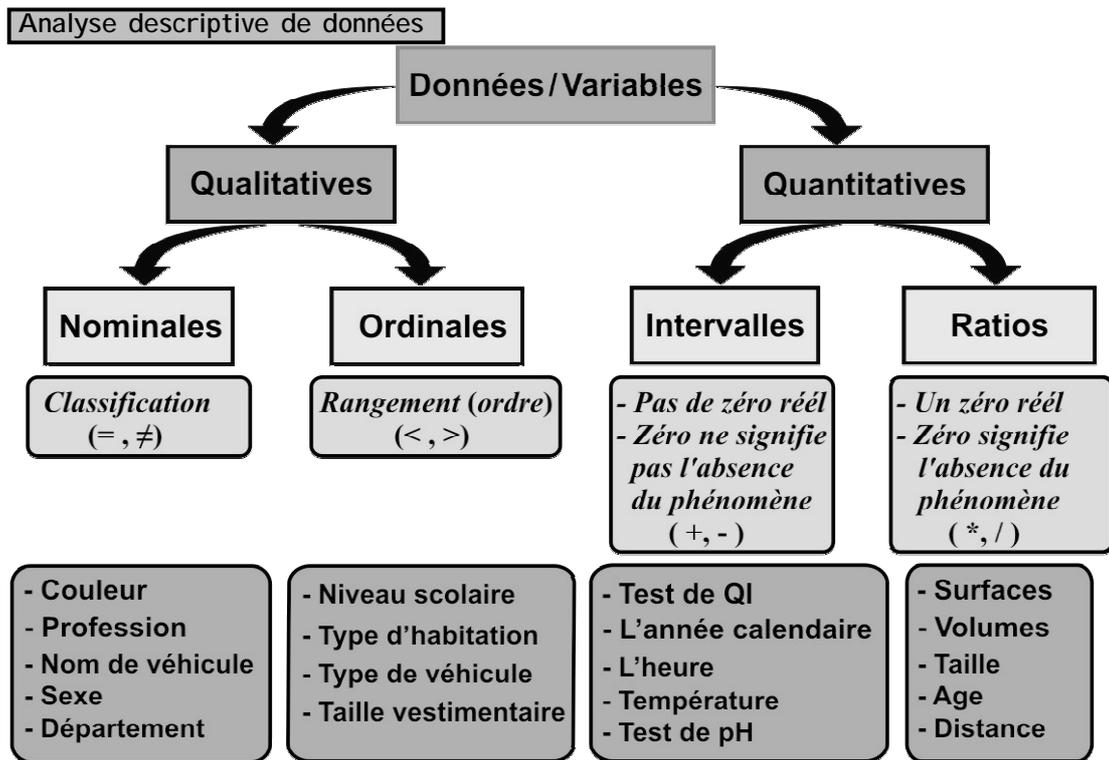
Quelles sont les échelles de mesure ?

Un système ou une échelle de mesure est une classification qui décrit la nature de l'information dans les valeurs attribuées aux variables .

Le psychologue Stanley Smith Stevens a élaboré la classification la plus connue comportant quatre niveaux, ou échelles, de mesure:

Nominal , Ordinal , Intervalle et Ratio.

Dans le même contexte, on peut aussi parler de variables mesurées dans une échelle de mesure .



Analyse descriptive de données

Quelle est l'utilité des échelles de mesures ?

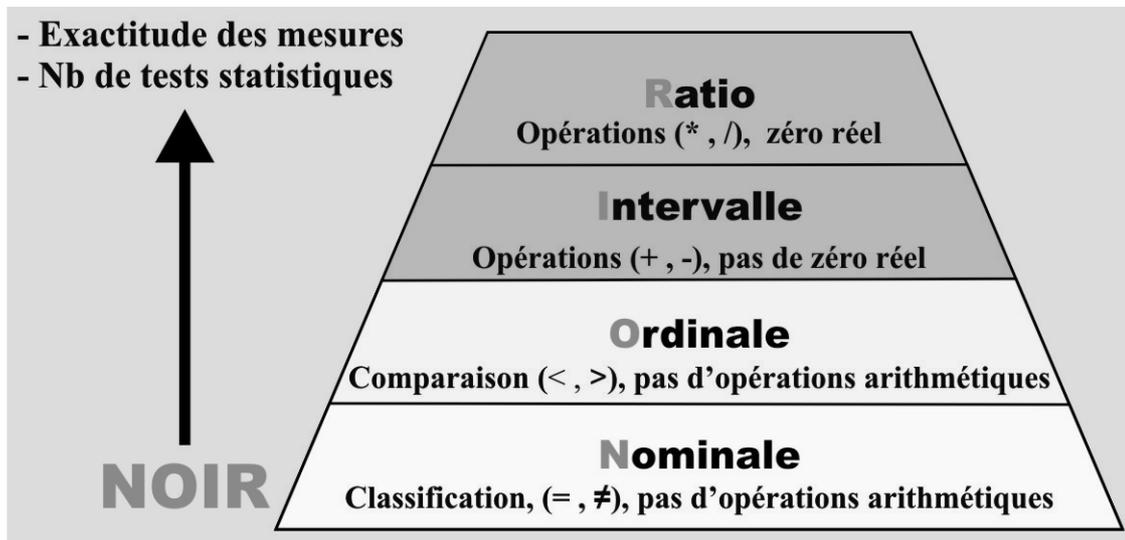
Les méthodes d'analyse de données statistiques se diffèrent en fonction des types de données. Par conséquent, il faut bien comprendre que la distinction entre les types de variables est très importante car sur laquelle on doit décider le type des analyses statistiques, des représentations graphiques ou encore des tests statistiques utilisables.

Par exemple:

Corrélation et régression, Analyse de variance, Analyse discriminante, Tableaux croisés, ...etc

Analyse descriptive de données

La pyramide des échelles de mesures



Exemple de passage des échelles de mesures

Personne	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Echelle
Taille	153	171	126	207	192	167	163	177	175	184	161	155	135	176	187	158	169	149	197	178	Ratio
Rapport avec moyenne	-16	2	-43	38	23	-2	-6	8	6	15	-8	-14	-34	7	18	11	0	-20	28	9	Intervalle
Rangement	C	M	TC	TL	L	M	M	M	M	L	M	C	TC	M	L	L	M	C	TL	M	Ordinale
Classification	No	No	An	An	No	An	No	No	No	No	No	An	No	Nominale							

Ratio → Intervalle

$$\text{Moy} = \frac{\sum T_i}{20} = 169$$

$$\text{Diff} = \text{val} - 169$$

Intervalle → Ordinale

Libellé	Marge (Intervalle)
Très Courte	Diff < -25
Courte	Diff ∈ [-25 .. -10 [
Moyenne	Diff ∈ [-10 .. 10 [
Longue	Diff ∈ [10 .. 25 [
Très Longue	Diff > 25

Ordinale → Nominale

Classe	Marge (Intervalle)
Normale	Courte
	Moyenne
	Longue
Anormale	Très Courte
	Très Longue

Analyse descriptive de données

Y-a-t-il une différence entre les variables en analyse ?

Oui, on distingue deux types de variables :

1. Variable Indépendante (VI): toute variable manipulée/assignée ou celle qui constitue la cause présumée d'un phénomène (ou d'un effet) d'étude.

2. Variable Dépendante (VD): c'est la variable observée et mesurée selon la variation d'une (des) autre(s) variable(s) indépendante(s)

Analyse descriptive de données

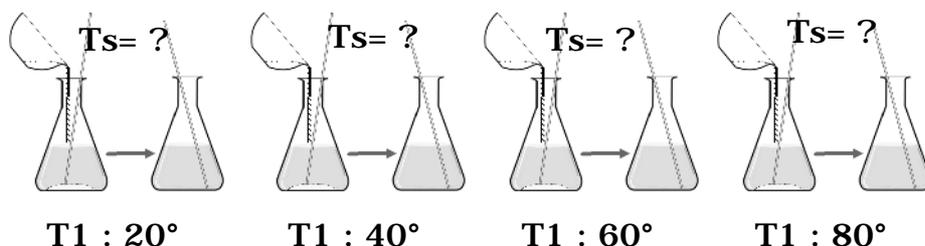
Exemple :

Etude :

Effet de la température sur la vitesse de solubilité du sel dans l'eau

Hypothèse :

La solubilité du sel dans l'eau augmente quand la température de l'eau augmente.



Variable Indépendante : Température

Variable dépendante : Vitesse de solubilité