

Statistique descriptive

Chapitre 1 : Qu'est-ce que la statistique ?



Jaouad Madkour

Faculté d'Économie et de Gestion de Tanger · 27 août 2026

Prérequis

Le mot et la chose

À quoi elle sert en économie

Descriptive et inférentielle

Le fil rouge du cours

Ce qu'il faut retenir

Prérequis

Aucun cours préalable

Le programme du baccalauréat : manipuler une somme, un pourcentage, une fraction, et lire un graphique. Rien de plus.

Les notions mobilisées

La notation $\sum$ est introduite au premier chapitre ; il suffit de savoir additionner. Aucune notion de probabilité n'est employée.

Repris depuis le début

Moyenne, médiane, variance, corrélation : **tout est défini ici**, et c'est ce cours qui sert de socle à trois autres — probabilités, analyse des données, économétrie.

Il décrit un ensemble de données observées. **Il n'en tire aucune conclusion sur une population** : c'est le rôle de la statistique inférentielle, et la distinction est faite dès le premier chapitre.

Le mot et la chose

Statistique

La **statistique** est l'ensemble des méthodes qui permettent de **recueillir, résumer, représenter** et **interpréter** des données chiffrées portant sur un ensemble d'individus.

L'étymologie n'est pas anodine

Le mot vient de l'italien *statista*, l'homme d'État. La statistique naît au XVII^e siècle comme la « science de l'État » : recenser la population, mesurer les récoltes, évaluer la matière imposable.

Elle est née d'un besoin de **gouverner**, c'est-à-dire de décider sans pouvoir tout observer. C'est encore exactement sa fonction.

Le problème qu'elle résout

Un problème très concret

Une entreprise emploie 500 salariés. Le directeur des ressources humaines dispose d'un fichier de 500 lignes et de six colonnes.

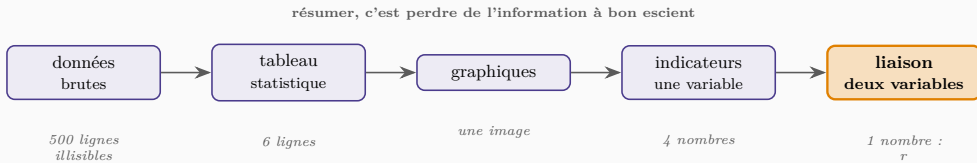
Question du directeur général : « Alors, où en est-on sur les salaires ? »

Il est impossible de répondre en lisant les 500 lignes. Il faut **résumer**.

Le geste fondamental

Résumer, c'est **perdre volontairement de l'information** pour rendre le reste utilisable.

500 nombres deviennent un tableau de six lignes, puis un graphique, puis quatre indicateurs, puis un seul.



À quoi elle sert en économie

Mesurer

Aucune grandeur économique ne s'observe directement. Le **PIB**, l'**inflation**, le **chômage** n'existent pas dans la nature : ce sont des **constructions statistiques**.

L'indice des prix à la consommation résulte d'un panier de biens, de pondérations, d'un relevé mensuel de milliers de prix. Le taux de chômage dépend d'une définition — celle du Bureau international du travail — qui décide qui est « chômeur » et qui ne l'est pas.

Changer la méthode change le chiffre. C'est pourquoi un économiste doit savoir comment un indicateur est fabriqué avant de le commenter.

Comparer

Deux régions, deux entreprises, deux périodes. Comparer exige des grandeurs **rendues comparables** : par tête, en volume, à prix constants, en parité de pouvoir d'achat.

C'est un travail statistique, et il précède toute analyse économique.

Décider

Fixer un prix, dimensionner un stock, accorder un crédit, calibrer une politique publique. Toute décision suppose une estimation, et toute estimation suppose une méthode.

Le chiffre n'est jamais neutre

Un même jeu de données peut donner deux messages opposés selon l'indicateur choisi.

Nos 500 salariés : le salaire **moyen** est de 7 158 dirhams, le salaire **médian** de 6 770. Un syndicat citera le second, une direction le premier. Aucun des deux ne ment.

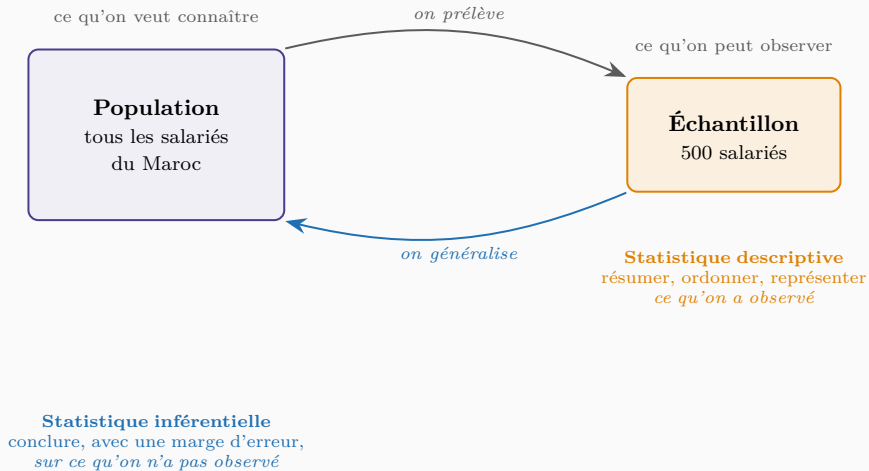
Savoir lequel des deux répond à la question posée est précisément l'objet de ce cours.

Ce que vous devez pouvoir faire à la fin du semestre

1. Construire un tableau statistique à partir de données brutes.
2. Choisir la représentation graphique qui convient au type de variable.
3. Calculer et **interpréter** les indicateurs de tendance centrale, de dispersion, de forme et de concentration.
4. Étudier la liaison entre deux variables et l'ajuster par une droite.
5. **Reconnaître un chiffre mal construit** — et c'est peut-être le plus utile.

Descriptive et inférentielle

La distinction fondamentale



Statistique descriptive

Elle **résume et représente** les données dont on dispose. Elle ne prétend rien dire au-delà.

Ses conclusions sont **certaines** : si l'on a bien calculé, le salaire moyen de ces 500 salariés **est** de 7 158 dirhams. Il n'y a pas de marge d'erreur.

Statistique inférentielle

Elle **généralise** de l'échantillon observé à la population dont il est issu.

Ses conclusions sont **probables**, jamais certaines, et elles s'accompagnent toujours d'une marge d'erreur et d'un niveau de confiance.

La phrase qui les sépare

- Descriptive : « Dans **cette** entreprise, le salaire moyen est de 7 158 dirhams. »
- Inférentielle : « Le salaire moyen dans **le secteur** est compris entre 6 900 et 7 400 dirhams, avec 95 % de confiance. »

La première décrit ce qu'on a vu. La seconde parie sur ce qu'on n'a pas vu.

Le tableau de comparaison

Les deux démarches		
	Descriptive	Inférentielle
Sur quoi elle porte	les données observées	la population entière
Ce qu'elle produit	des constats	des estimations et des tests
Nature des conclusions	certaines	probables
Outil central	moyenne, écart-type, graphique	intervalle de confiance, test
Suppose	rien	un échantillon aléatoire
Notation	$\bar{x}, s$	μ, σ

La notation, à retenir dès maintenant

Les grandeurs calculées **sur l'échantillon** se notent en lettres latines : $\bar{x}$, s , r .

Les grandeurs **de la population**, inconnues, se notent en lettres grecques : μ , σ , ρ .

Cette convention vaut dans toute la statistique. Elle rappelle en permanence ce qu'on a mesuré et ce qu'on cherche à atteindre.

Pourquoi la descriptive vient en premier

Trois raisons, dans cet ordre d'importance :

1. **L'inférentielle repose sur elle.** Un intervalle de confiance est construit à partir d'une moyenne et d'un écart-type : il faut savoir les calculer et les lire.
2. **Elle se passe d'hypothèses.** Aucune loi de probabilité, aucun échantillonnage aléatoire. Elle est vraie par construction.
3. **Elle vient toujours en premier dans la pratique.** Aucun statisticien ne teste avant d'avoir regardé ses données. Le premier réflexe est toujours de décrire.

Le piège symétrique

Attention cependant à la tentation inverse : **décrire un échantillon et parler comme si l'on décrivait la population.**

« Le salaire moyen au Maroc est de 7 158 dirhams » — non : le salaire moyen **de ces 500 salariés** l'est. Passer de l'un à l'autre demande l'inférentielle, et une justification.

Le fil rouge du cours

Le fichier salaires

500 salariés d'une entreprise, six colonnes :

Variable	Nature	Valeurs
identifiant	repère	1 à 500
salaire	quantitative continue	de 2 810 à 19 050 dirhams
anciennete	quantitative discrète	de 0 à 35 ans
diplome	qualitative ordinale	Bac, Bac+2, Bac+5
region	qualitative nominale	Tanger, Casablanca, Rabat, Autres
sexe	qualitative nominale	F, H

Pourquoi un seul jeu de données

Chaque indicateur du cours sera calculé sur **ces mêmes 500** salariés. Vous verrez donc, à chaque chapitre, le même nuage de nombres révéler une propriété de plus.

À la fin, vous saurez tout dire de ce fichier — et vous saurez le refaire sur n'importe quel autre.

Un premier coup d'œil, sans aucun calcul

Question	Réponse
Combien de salariés?	500
Salaire le plus bas	2 810 dirhams
Salaire le plus haut	19 050 dirhams
Écart entre les deux	facteur 6,8

Trois nombres, et l'on sait déjà que la distribution est **très étalée**. Tout le reste du cours va préciser cette impression.

Ce qu'il faut retenir

Cinq points

1. La statistique **recueille, résume, représente et interprète** des données.
2. Résumer, c'est **perdre de l'information à bon escient** — jamais par négligence.
3. En économie, elle sert à **mesurer** des grandeurs qui n'existent que par leur définition, à **comparer** et à **décider**.
4. La **descriptive** décrit ce qu'on a observé : conclusions certaines. L'**inférentielle** généralise : conclusions probables.
5. Notation : lettres latines pour l'échantillon, lettres grecques pour la population.

La suite

Avant de calculer quoi que ce soit, il faut nommer précisément ce sur quoi l'on travaille : population, individu, variable, modalité.

C'est le chapitre 2, et ce vocabulaire commande tout le reste — notamment le choix des graphiques et des indicateurs autorisés.