



## Statistique appliquée

Master GPRE - Université de Bretagne-Sud

Thierry Dhorne - [www.dhorne.education](http://www.dhorne.education)

16 novembre 2015

En hommage aux victimes du terrorisme  
les transparents portent un double liseré noir



## Introduction

- ❖ Données - Informations - Connaissances
- ❖ Définition de la statistique
- ❖ La statistique dans la « data science »
- ❖ Une science expérimentale !
- ❖ Statistique et ... Mathématique !
- ❖ Statistique et ... Anglais !
- ❖ La statistique est-elle anglo-saxonne ?

Données

Démarche

Approche conceptuelle de la statistique

# *Introduction*



# Données - Informations - Connaissances

## Introduction

### ❖ Données - Informations - Connaissances

- ❖ Définition de la statistique
- ❖ La statistique dans la « data science »
- ❖ Une science expérimentale !
- ❖ Statistique et ... Mathématique !
- ❖ Statistique et ... Anglais !
- ❖ La statistique est-elle anglo-saxonne ?

## Données

## Démarche

Approche conceptuelle de la statistique

- fait
  - événement ou réalité observé(e)
- donnée
  - élément permettant de caractériser un fait
- information
  - ensemble d'éléments permettant d'apprécier un fait
  - transformation compréhensible d'une donnée
  - ★ une information est (normalement ?) objective : indépendante du sujet
- connaissance
  - interprétation conceptuelle d'un ensemble d'informations
  - ★ une connaissance est (normalement ?) subjective : dépendante du sujet



# Définition de la statistique

## Introduction

- ❖ Données - Informations - Connaissances
- ❖ Définition de la statistique
- ❖ La statistique dans la « data science »
- ❖ Une science expérimentale !
- ❖ Statistique et ... Mathématique !
- ❖ Statistique et ... Anglais !
- ❖ La statistique est-elle anglo-saxonne ?

## Données

## Démarche

Approche conceptuelle de la statistique

- la statistique permet de :
  - recueillir des données caractérisant un problème
  - transformer ces données en informations
  - communiquer ces informations de façon compréhensible
  - aider à l'élaboration de connaissances ou à la prise de décision



# La statistique dans la « data science »

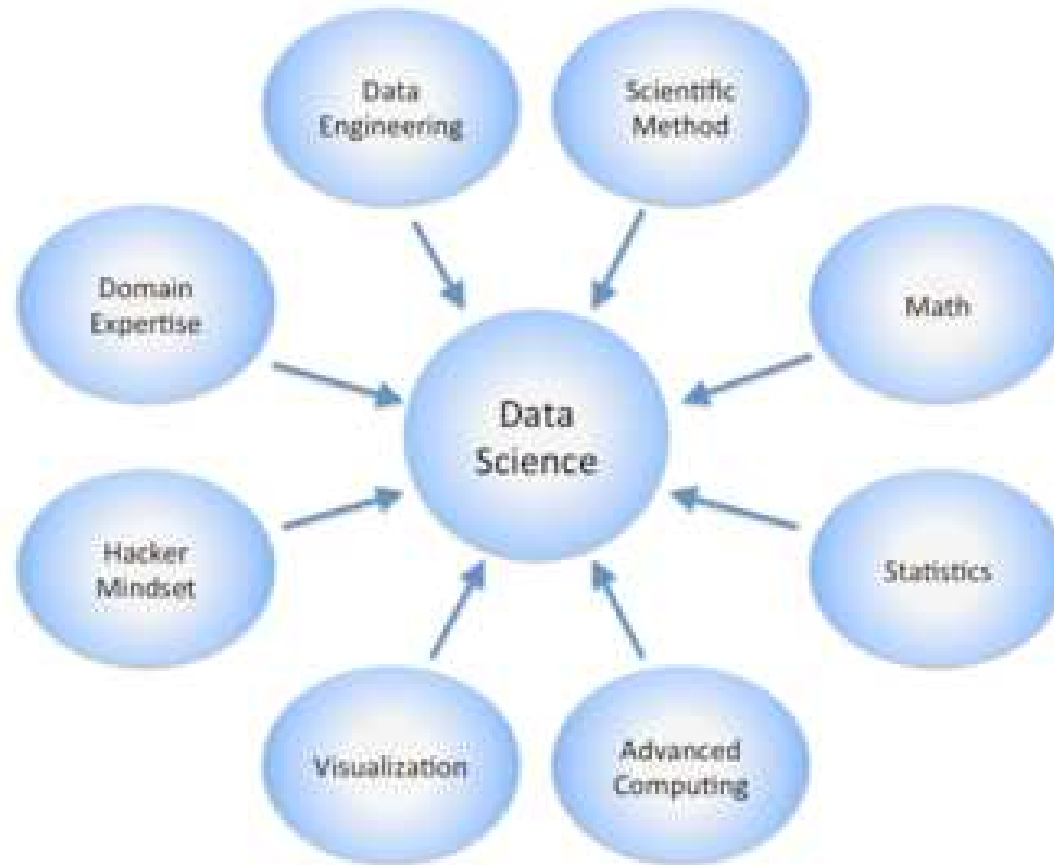
## Introduction

- ❖ Données - Informations - Connaissances
- ❖ Définition de la statistique
- ❖ La statistique dans la « data science »
- ❖ Une science expérimentale !
- ❖ Statistique et ... Mathématique !
- ❖ Statistique et ... Anglais !
- ❖ La statistique est-elle anglo-saxonne ?

## Données

## Démarche

Approche conceptuelle de la statistique





# Une science expérimentale!

## Introduction

- ❖ Données - Informations - Connaissances
- ❖ Définition de la statistique
- ❖ La statistique dans la « data science »
- ❖ Une science expérimentale !
- ❖ Statistique et ... Mathématique !
- ❖ Statistique et ... Anglais !
- ❖ La statistique est-elle anglo-saxonne ?

## Données

## Démarche

Approche conceptuelle de la statistique

## David S. Moore and George W. Cobb

Statistics, like physics and economics but unlike algebraic topology or probability theory, values mathematical understanding as a means to an end, not as an end in itself. Like physics and economics, but unlike sub-fields of mathematics, statistics has a subject matter of its own, quite apart from mathematics.

*La statistique, comme la physique et l'économie mais pas comme la topologie algébrique ou les probabilités considère la compréhension mathématique comme un moyen pour atteindre une fin et non pas comme une fin en soi. Comme la physique et l'économie et à la différence des divers champs des mathématiques, la statistique a un objet d'étude à part entière, assez éloigné des mathématiques.*



# Statistique et ... Mathématique!

## Introduction

- ❖ Données - Informations - Connaissances
- ❖ Définition de la statistique
- ❖ La statistique dans la « data science »
- ❖ Une science expérimentale !

## ❖ Statistique et ... Mathématique !

- ❖ Statistique et ... Anglais !
- ❖ La statistique est-elle anglo-saxonne ?

## Données

## Démarche

Approche conceptuelle de la statistique

## David S. Moore and George W. Cobb

It has become a truism, at least among statisticians, that while statistics is a mathematical science, it is not a subfield of mathematics. We even have aphorisms to express some ways in which our science differs from mathematics.

*Il est devenu une évidence, au moins parmi les statisticiens, que bien que la statistique soit une science mathématique, ce n'est pas une partie des mathématiques. Nous avons même des aphorismes pour exprimer quelques manières selon lesquelles notre science diffère des mathématiques.*



# Statistique et ... Anglais !

## Introduction

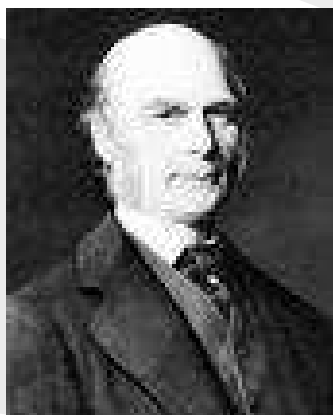
- ❖ Données - Informations - Connaissances
- ❖ Définition de la statistique
- ❖ La statistique dans la « data science »
- ❖ Une science expérimentale !
- ❖ Statistique et ... Mathématique !
- ❖ Statistique et ... Anglais !
- ❖ La statistique est-elle anglo-saxonne ?

## Données

## Démarche

Approche conceptuelle de la statistique

- pourquoi tant de références anglo-saxonnes
- le développement de la statistique à la fin du XIXème et au début du XXème siècle a essentiellement eu lieu en Grande-Bretagne (et en Irlande)
- ★ la passion de la statistique conduit à une grande considération pour les statisticiens anglo-saxons qui lui ont donné ses lettres de noblesse



Galton



Pearson



Student



Fisher

- et beaucoup d'autres





# La statistique est-elle anglo-saxonne ?

## Un point de vue... anglo-saxon

David S. Moore and George W. Cobb

We might say that mathematics is French, while statistics is resolutely Anglo-Saxon. The French, proud of a long history and high culture, are wary of aggressive Anglo-Saxon pragmatism. In the words attributed to a French diplomat, “It works in practice, but does it work in theory?”

*Nous pourrions dire que les mathématiques sont françaises, alors que les statistiques sont résolument anglo-saxonnes. Les Français, fiers d'une longue histoire et d'une haute culture, se méfient du pragmatisme agressif des anglo-saxons. Selon les mots attribués à un diplomate français « Celà marche en pratique, mais est-ce-que cela marche en théorie ? »*

### Introduction

- ❖ Données - Informations - Connaissances
- ❖ Définition de la statistique
- ❖ La statistique dans la « data science »
- ❖ Une science expérimentale !
- ❖ Statistique et ... Mathématique !
- ❖ Statistique et ... Anglais !

❖ La statistique est-elle anglo-saxonne ?

### Données

### Démarche

Approche conceptuelle de la statistique



## Introduction

---

### Données

- ❖ Organisation des données
- ❖ Types de variables
- ❖ Variables quantitatives
- ❖ Variables qualitatives
- ❖ Variables déclaratives
- ❖ Exemples

## Démarche

---

Approche conceptuelle de la statistique

---

# *Données*



# Organisation des données

Introduction

Données

❖ Organisation des données

❖ Types de variables

❖ Variables quantitatives

❖ Variables qualitatives

❖ Variables déclaratives

❖ Exemples

Démarche

Approche conceptuelle de la statistique

- les données sont habituellement organisées dans un

## Tableau individus × variables

	<i>variables</i>					
<i>individus</i>	$x_1^1$	$x_1^2$	...	$x_1^i$	...	$x_1^p$
	$x_2^1$	$x_2^2$	...	$x_2^i$	...	$x_2^p$
	⋮	⋮		⋮		⋮
	$x_r^1$	$x_r^2$	...	$x_r^i$	...	$x_r^p$
	⋮	⋮		⋮		⋮
	$x_n^1$	$x_n^2$	...	$x_n^i$	...	$x_n^p$



# Types de variables

## Introduction

### Données

❖ Organisation des données

### ❖ Types de variables

❖ Variables quantitatives

❖ Variables qualitatives

❖ Variables déclaratives

❖ Exemples

## Démarche

Approche conceptuelle de la statistique

- il existe des variables
  - quantitatives
  - qualitatives
  - déclaratives



# Variables quantitatives

## Introduction

### Données

- ❖ Organisation des données
- ❖ Types de variables
- ❖ Variables quantitatives
- ❖ Variables qualitatives
- ❖ Variables déclaratives
- ❖ Exemples

## Démarche

Approche conceptuelle de la statistique

- les variables quantitatives mesurent une quantité
  - continue (variant continûment)
  - discrète (« séparée »)
- ★ les variables entières sont un cas particulier de variables discrètes



# Variables qualitatives

## Introduction

### Données

- ❖ Organisation des données
- ❖ Types de variables
- ❖ Variables quantitatives
- ❖ Variables qualitatives
- ❖ Variables déclaratives
- ❖ Exemples

## Démarche

Approche conceptuelle de la statistique

- les variables qualitatives mesurent une qualité
- on ne peut pas quantifier
- chaque éventualité est appelée modalité
- ★ les modalités peuvent être ordonnées



# Variables déclaratives

## Introduction

### Données

- ❖ Organisation des données
- ❖ Types de variables
- ❖ Variables quantitatives
- ❖ Variables qualitatives
- ❖ Variables déclaratives
- ❖ Exemples

## Démarche

Approche conceptuelle de la statistique

- les variables déclarative correspondent à une déclaration
- souvent des questions ouvertes
- elles peuvent être de nature
- lexicale (constituées de mots séparés : signification)
- textuelle (constituées de mots non séparables : sens)



# Exemples

## Introduction

### Données

- ❖ Organisation des données
- ❖ Types de variables
- ❖ Variables quantitatives
- ❖ Variables qualitatives
- ❖ Variables déclaratives
- ❖ Exemples

## Démarche

Approche conceptuelle de la statistique

- nous allons parler de vous
  - votre taille
  - le nombre d'enfants dans votre fratrie
  - la couleur de vos yeux
  - le type de votre appartement
  - votre sport favori
  - la perception que vous avez de l'UBS





Introduction

Données

**Démarche**

❖ Étapes de l'étude statistique

❖ Statistique et management de données

Approche conceptuelle de la statistique

# *Démarche*



# Étapes de l'étude statistique

Introduction

Données

Démarche

❖ Étapes de l'étude statistique

❖ Statistique et management de données

Approche conceptuelle de la statistique

- définition des objectifs
- définition des concepts
- définition du protocole
- collecte des données
- mise en forme et stockage des données
- traitement des données
- analyse des données
- interprétation des données
- restitution



# Statistique et management de données

Introduction

Données

Démarche

❖ Étapes de l'étude statistique

❖ Statistique et management de données

Approche conceptuelle de la statistique

- la statistique comporte des étapes d'expertise humaine (subjective)
  - l'interprétation
  - la formalisation
  - la restitution
  - qui sont (jusqu'à aujourd'hui) non automatisables
- lorsqu'il n'y a pas d'expertise mais de simples transformations de données (même intelligentes)
  - on parle de management (gestion) de données (data management)
  - ★ ordinairement avant l'analyse statistique il y a une étape de management de données
    - qui peut être coûteuse en temps
    - en particulier lorsque les données ne sont pas de bonne qualité



Introduction

Données

Démarche

Approche  
conceptuelle de la  
statistique

- ❖ Concepts de base
- ❖ Concepts de base
- ❖ Concepts de base

# *Approche conceptuelle de la statistique*



Introduction

Données

Démarche

Approche  
conceptuelle de la  
statistique

❖ Concepts de base

❖ Concepts de base

❖ Concepts de base

## Population

Ensemble des individus (unités statistiques) dont l'étude est envisagée

- terme peut être un peu impropre aujourd'hui (vient de la démographie) mais consacré
- population des français (?), population des clients d'une banque
- ★ ce concept est parfois un peu «formel»
- population des clients d'un hypermarché, population des enfants nés à Vannes en 2007
- une population peut-être (conceptuellement) de taille infinie
- la population des nombres entiers
- la population des véhicules d'un type donné



Introduction

Données

Démarche

Approche  
conceptuelle de la  
statistique

❖ Concepts de base

❖ Concepts de base

❖ Concepts de base

## Recensement

On appelle recensement une étude exhaustive de la population concernée, c'est-à-dire une étude de tous les individus la constituant. On enregistre donc toutes les variables de l'étude sur chacun des individus

- exemple du recensement de la population française jusqu'en 1999
- les recensements sont par conséquent
  - très difficiles (il faut connaître en détail la population)
  - très coûteux



Introduction

Données

Démarche

Approche  
conceptuelle de la  
statistique

❖ Concepts de base

❖ Concepts de base

❖ Concepts de base

## Sondage et échantillon

On parle de sondage lorsque l'on n'étudie qu'une partie de la population. Le sous-ensemble étudié est appelé échantillon

- un échantillon peut être obtenue de différentes manières
  - aléatoirement (par échantillonnage)
  - par sélection
  - sous contraintes