STATISTICAL LABORATORY

Applied Mathematics for Economics and Management
1st Year/1st Semester
2025/2026

CONTACT

Professor: Elisabete Fernandes
E-mail: efernandes@iseg.ulisboa.pt

https://doity.com.br/estatistica-aplicada-a-nutricao

https://basiccode.com.br/produto/informatica-basica/

PROGRAM

1. Fundamental Concepts of Statistics

2. Exploratory Data Analysis

3. Organizing and Summarizing Data

4. Association and Relationships Between Variables

5. Index Numbers

6. Time Series Analysis

BIBLIOGRAPHY

Main references:

- Triola, M. F. (2022). *Elementary statistics* (14th ed.). Pearson.
- Silvestre, A. L. (2007). *Análise de dados: Estatística descritiva*. Escolar Editora.

Complementary references:

- Flury, B., & Riedwyl, H. (1988). *Multivariate statistics: A practical approach*. Chapman & Hall.
- Newbold, P., Carlson, W. L., & Thorne, B. M. (2013). *Statistics for business and economics* (8th ed.). Pearson.
- Winston, W. L. (2016). *Microsoft Excel 2016 data analysis and business modeling*. Microsoft Press.

Package: Excel

LECTURE 1: FUNDAMENTAL CONCEPTS OF STATISTICS

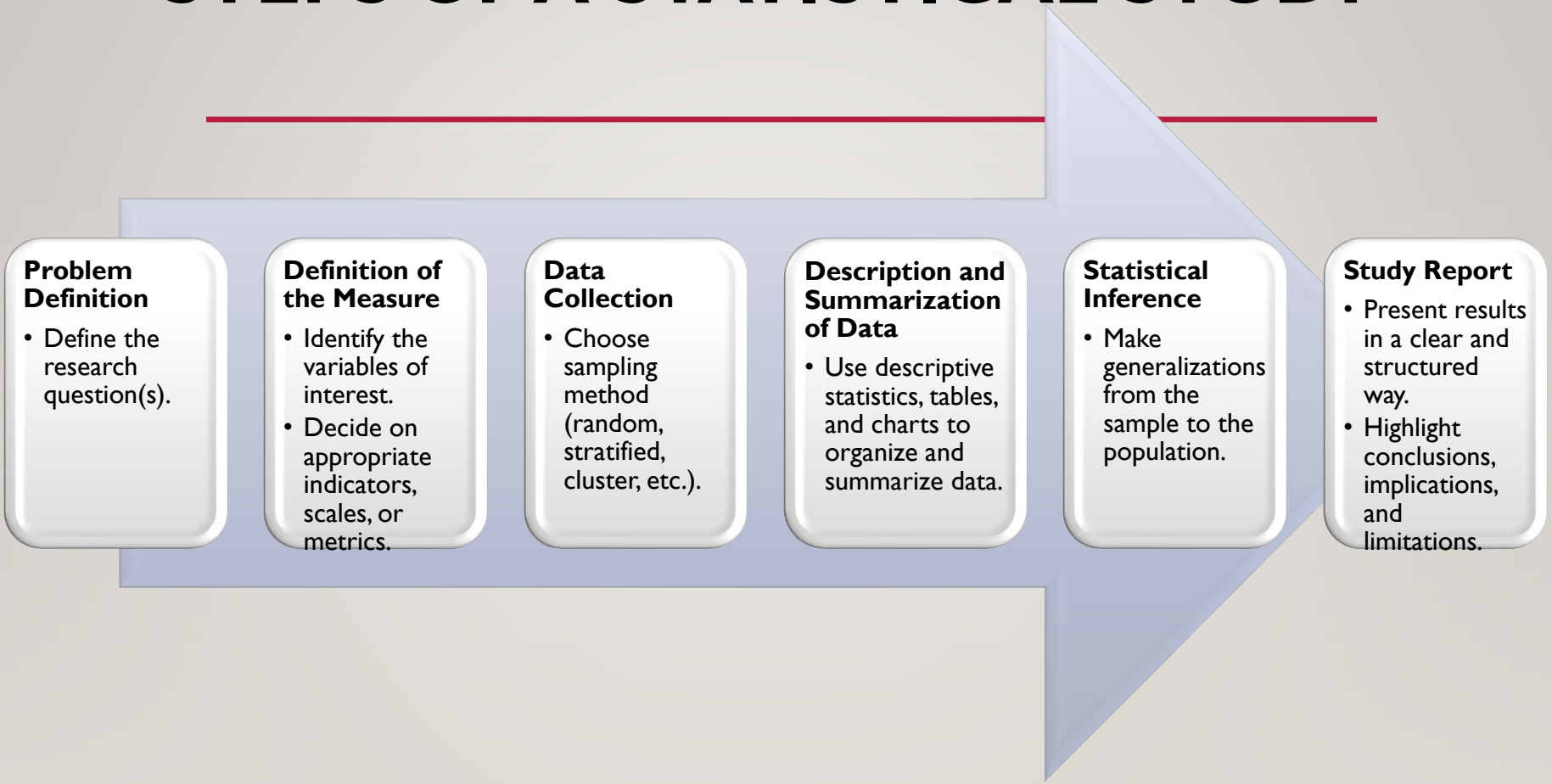
WHAT IS STATISTICS?

- The science of collecting, organizing, analysing, and interpreting data.

- Applications: **Economics, Management**, Healthcare, Social Sciences, Engineering, and more.

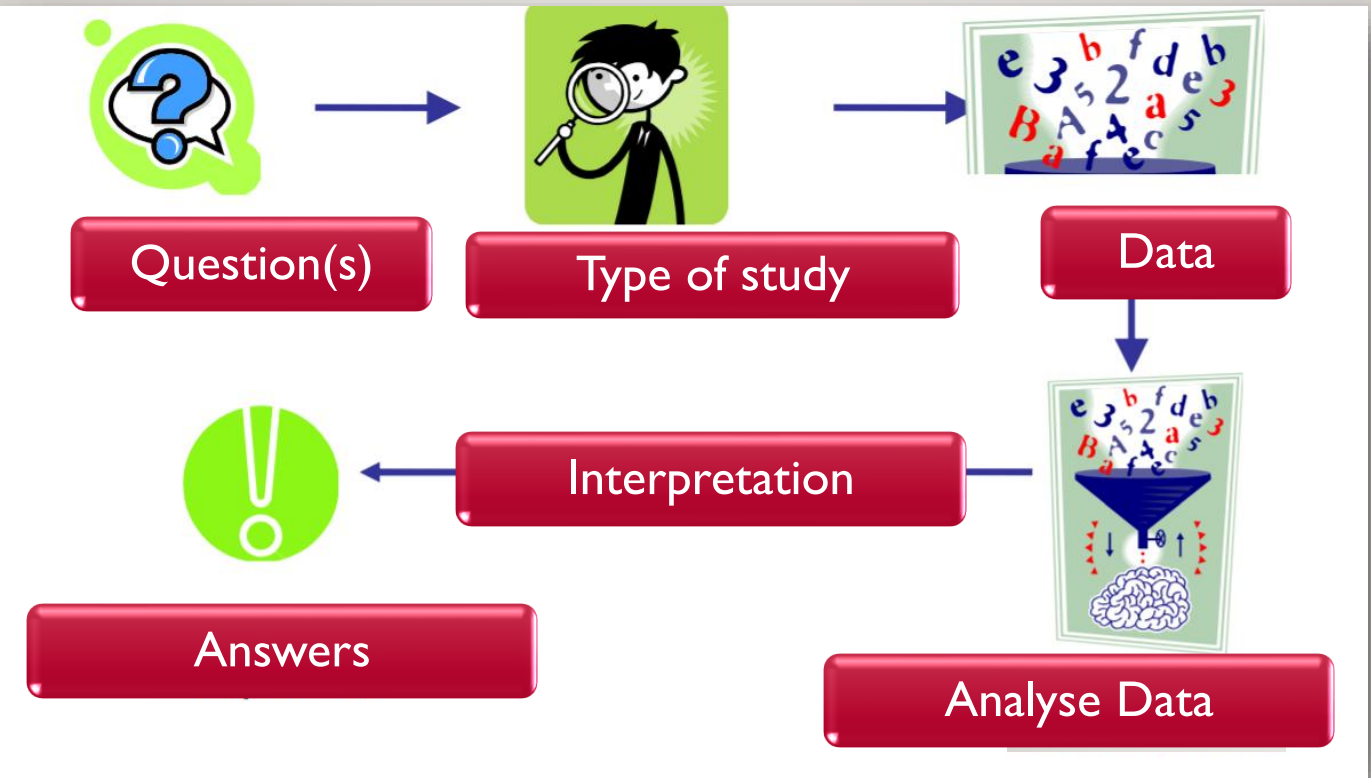
- Purpose in **Management and Economics**: support decision-making under uncertainly, identify patterns, and predict trends.

STEPS OF A STATISTICAL STUDY



Purpose: Transform raw data into meaningful information.

STEPS OF A STATISTICAL STUDY: VISUAL REPRESENTATION



MEANINGS OF "STATISTICS"

Scientific Discipline

Measure:

Numerical summaries that describe characteristics of a dataset.

Examples: mean, variance, and percentages.

Data:

Synonym for numerical information in specific areas.

Examples: heath statistics, industrial statistics, and employment statistics.

POPULATION AND SAMPLE

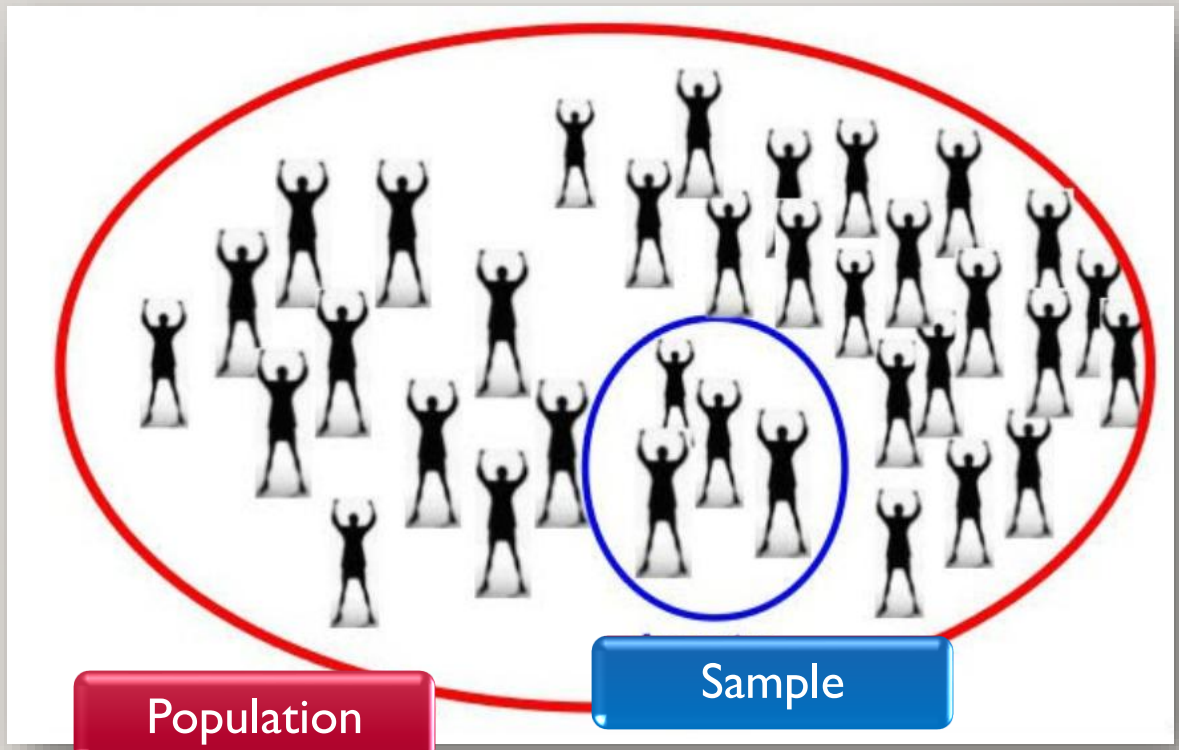
Population/universe: All elements of interest in a study or research.

Types of Population: real vs hypothetical; finite vs infinite.

Sample: A representative subset of a population, used to draw conclusions about the whole.

Example: all customers of a company (population) vs 50 selected customers (sample).

POPULATION AND SAMPLE: VISUAL REPRESENTATION



https://sites.google.com/site/estatisticabasicacc/conteudo/

OBJECTIVES OF STATISTICAL ANALYSIS

Data summarization and reduction

- Organize and condense large datasets into understandable forms
- **Examples**: frequency tables, charts, and summary measures (mean, median, standard deviation)

Inference to other datasets

- Make predictions or generalizations about a population based on sample data
- **Example**: estimating total sales based on a sample of transactions

Identification of relationships between datasets

- Discover correlations, associations, or causal links between variables
- **Example**: analyzing the relationship between marketing spend and customer acquisition

Dimensionality reduction

- Simplify data by reducing the number of variables while retaining essential information
- **Example:** principal component analysis in multivariate data

Classification and discrimination

- Assign data points to categories or groups based on characteristics
- **Example**: categorizing customers into segments (loyal, occasional, new)

Data clustering

- Group similar data points together to identify patterns or natural groupings
- **Example**: grouping products based on sales patterns or customer behavior

TYPES OF STATISTICS

- **Descriptive Statistics**: organizes and summarizes data using tables, graphs, and measures.
- **Inferential Statistics**: draws conclusions about a population from a sample through estimation and hypothesis testing.
- **Example**: average revenue (descriptive) vs sales forecast (inferential).

IMPORTANCE OF SAMPLING STUDIES

Types of studies: census vs sampling (advantages and disadvantages)

Sampling: more efficient, less costly, and less time-consuming than surveying the entire population.

Census: covers entire population, but is expensive, time-consuming, and often impractical.

Example: surveying 1,000 households from a city (sampling) vs. All households in the entire country (census).

DATA COLLECTION

Sources: surveys, interviews, administrative databases, sensors, and company records.

Processes: coding, recording, and validating data.

Data Coding: transforming answers into numerical or categorical codes.
Data Recording: storing information systematically (Excel, SPSS, SQL).
Data Validation: checking consistency, completeness, and accuracy before analysis.

TYPES OF DATA

1. Cross-Sectional Data

Observations for multiple units collected at **one point in time** (one or more variables).

Example: survey of customer satisfaction across 100 stores in January 2025.

2. Time Series Data

Observations collected **over time** for a single unit (one or more variables).

Example: monthly sales revenue of a company from January 2020 to December 2024.

3. Panel Data (Longitudinal Data)

Combines **cross-sectional and time series data.** Observations for multiple units over multiple periods.

Example: annual income of 500 households from 2018 to 2024.

STATISTICAL UNIT, PARAMETER, AND STATISTICS

STATISTICAL UNIT: Each element of the population or sample.

Sample: $(x_1, x_2, \ldots, x_n)$

PARAMETERS: Characteristics of the population.

Example: population mean μ and population standard deviation σ.

STATISTICS: Measures calculated from a sample.

Example: sample mean $\bar{x}$ and sample standard deviation s.

RANDOM EXPERIMENT

A **random experiment** is a process or action whose outcome **cannot be predicted with certainty** in advance, even under identical conditions.

Characteristics:

- Multiple possible outcomes
- Repeatable under same conditions
- Set of all possible outcomes = **Sample Space**

Examples:

- Roll a Die → Outcomes: 1, 2, 3, 4, 5, 6
- Coin Toss (Coin Flip) → Outcomes: Heads or Tails
- Selecting a Random Employee → Outcome: Age of selected person

VARIABLES

Variable: A characteristic or property **observed in a random experiment.**

Random variable: Theoretical concept that assigns values to the outcomes of a random experiment.

Examples: weekly sales and number of defective items.

Empirical variable: Observed in practice, based on collected data.

Examples: age, weight, and number of products sold.

Random Variable (X) vs **Empirical Variable (x)**

Sample: $(x_1, x_2, \ldots, x_n)$

LEVELS OF MEASUREMENT

Nominal

- Categories without order.
- **Examples:** gender and nationality.

Ordinal

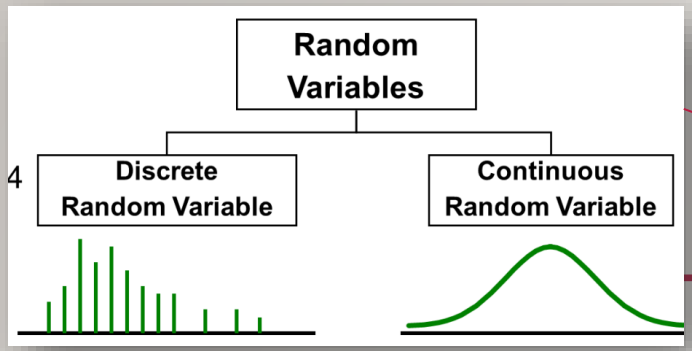
- Ordered categories.
- **Example:** education level.

Interval

- Differences between values are meaningful, but there is **no absolute zero**.
- **Example:** temperature in Celsius °C (0 °C ≠ no temperature).

Ratio

- An **absolute zero exists**, and ratios are meaningful.
- **Examples:** income, age, and weight. (0 kg = no weight).



CLASSIFICATION OF VARIABLES

$$Y = ax + b$$

Number of Values

Discrete Variables:

- Take finite or countable values.
- **Examples:** Number of children and number of defective items.

Continuous Variables:

- Can take infinite values within a range.
- **Examples:** Height and weight.

Explanatory Orientation

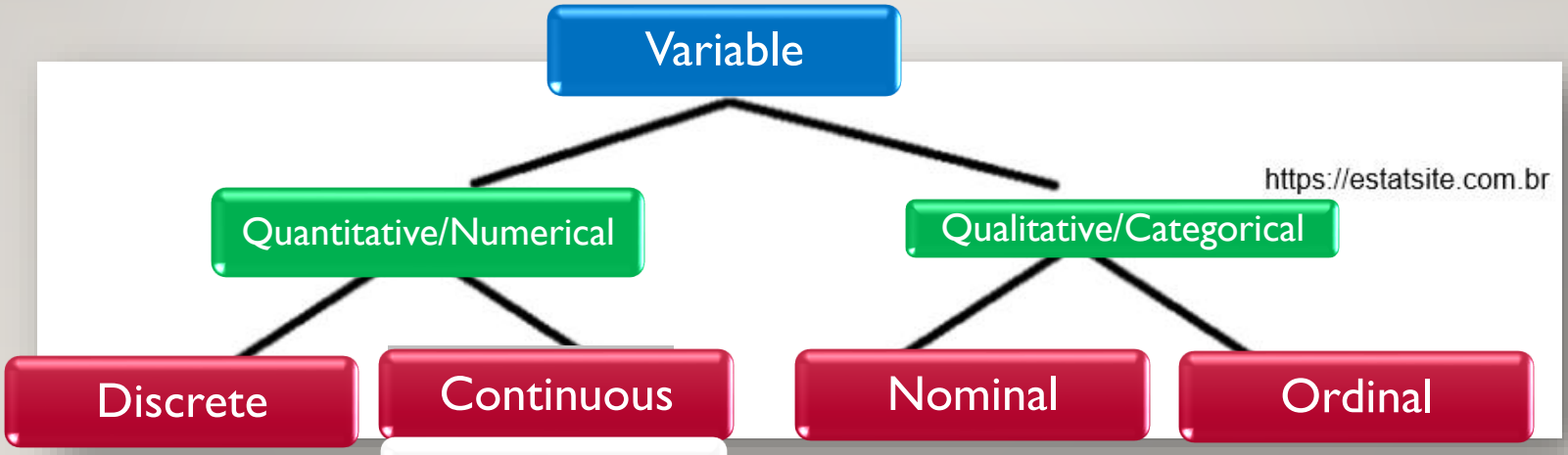
Explanatory Variable (Independent / Predictor Variable)

- A variable used to explain or predict changes in another variable.
- Represents the potential cause, influence, or input.
- Usually placed on the **x-axis** in graphs.
- **Examples:** marketing expenditure, study hours, and product price.

Explained Variable (Dependent / Response Variable)

- The variable whose variation we want to understand or predict.
- Represents the effect, outcome, or result.
- Usually placed on the **y-axis** in graphs.
- **Examples:** sales revenue, exam score, and demand for a product.

CLASSIFICATION OF VARIABLES: VISUAL REPRESENTATION



Example: Number of children

Interval or Ratio

Example: Height and weight

Example: Gender

Example: Satisfaction level

EXERCISE 1.1

1.1 A mortgage company randomly samples accounts of their time-share customers. State whether each of the following variables is categorical or numerical. If categorical, give the level of measurement. If numerical, is it discrete or continuous?

a. The original purchase price of a customer's time-share unit
b. The state (or country) of residence of a time-share owner
c. A time-share owner's satisfaction level with the maintenance of the unit purchased (1: very dissatisfied to 5: very satisfied)
d. The number of times a customer's payment was late

Newbold et al (2013)

EXERCISE 1.1: SOLUTION

✅ Answers:

a. **Numerical, Continuous (Ratio)** → Purchase price, any value, true zero.

b. **Categorical (Nominal)** → State/country, categories with no order.

c. **Categorical (Ordinal)** → Satisfaction scale 1–5, ordered categories.

d. **Numerical, Discrete (Ratio)** → Number of late payments, count values.

EXERCISE 1.2

1.2 Visitors to a supermarket in Singapore were asked to complete a customer service survey. Are the answers to the following survey questions categorical or numerical? If an answer is categorical, give the level of measurement. If an answer is numerical, is it discrete or continuous?

a. Have you visited this store before?
b. How would you rate the level of customer service you received today on a scale from 1 (very poor) to 5 (very good)?
c. How much money did you spend in the store today?

Newbold et al (2013)

EXERCISE 1.2: SOLUTION

✅ Answers:

a. **Categorical (Nominal)** → Yes/No, no natural order.

b. **Categorical (Ordinal)** → Rating scale 1–5, ordered categories.

c. **Numerical, Continuous (Ratio)** → Money spent, decimal values possible, true zero.

EXERCISE 1.5

1.5 A number of questions were posed to a random sample of visitors to a London tourist information center. For each question below, describe the type of data obtained.

a. Are you staying overnight in London?
b. How many times have you visited London previously?
c. Which of the following attractions have you visited?

Tower of London
Buckingham Palace
Big Ben
Covent Garden
Westminster Abbey

d. How likely are you to visit London again in the next 12 months: (1) unlikely, (2) likely, (3) very likely?

Newbold et al (2013)

EXERCISE 1.5: SOLUTION

✅ Answers:

a. Are you staying overnight in London?

- **Categorical (Nominal)** → Yes/No, no inherent order.

b. How many times have you visited London previously?

- **Numerical, Discrete (Ratio)** → Count of visits, zero possible, only integer values.

c. Which of the following attractions have you visited?

- **Categorical (Nominal, Multiple Response)** → Each attraction is a yes/no question; categories with no order.

d. How likely are you to visit London again in the next 12 months?

- **Categorical (Ordinal)** → Likert-type scale (e.g., very unlikely → very likely), ordered categories.

FORMAL REPRESENTATION OF DATA

$(x_1, x_2, x_3, ..., x_n)$ ou $x_i (i = 1, 2, ..., n)$

n observations of one variable

n observations of two variables

$[(x_1, y_1), (x_2, y_2), (x_3, y_3), ..., (x_n, y_n)]$

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}$$

n observations of p variables

Contingency tables

Movies Attended	Gender		Total
	Men	Women	
0	20	40	60
1	40	30	70
2 or more	10	10	20
Total	70	80	150

E.g. A survey of 150 adults classified each as to gender and the number of movies attended last month. Each respondent is classified according to two criteria—the number of movies attended and gender.

WHAT IS AN EMPIRICAL STUDY?

1. Based on **direct observation or data collection.**

2. Evidence-based: uses real data from experiments, questionnaires, interviews, or observations.

3. Analytical purpose: identify patterns, relationships, or effects.

4. Systematic approach: rigorous collection, processing, and interpretation.

5. Replicable: methods can be repeated by other researchers.

6. Example: Measuring teaching methods' impact on student performance.

EMPIRICAL RESEARCH IN ECONOMICS AND MANAGEMENT

- **Data sources:** surveys, experiments, administrative records, and financial databases.

- **Importance**:

  Understanding market behavior

  Supporting decision-making

  Evaluating policies

  Forecasting trends

EMPIRICAL WORK PROCEDURE

Steps:

- 1. Define research question.
- 2. Collect relevant data.
- 3. Organize and process data.
- 4. Apply descriptive and inferential statistical methods.
- 5. Interpret and report results.

SOURCES OF INFORMATION

1. Primary sources: original data from experiments, surveys, interviews, or observations.

2. Secondary sources: previously collected and published data (books, articles, reports, databases).

3. Tertiary sources: compilations and summaries (encyclopedias, handbooks, indexes).

4. Key point: Choosing reliable sources ensures **accuracy, relevance, and credibility.**

EXAMPLES OF RESEARCH USING NUMERIC DATABASES

- **Statistical databases**: Eurostat, World Bank, IBGE – population, economic, and social indicators
- **Financial and market databases**: Bloomberg, Thomson Reuters, Yahoo Finance – company financials, stock prices, market trends
- **Survey and panel data**: Longitudinal surveys on consumer behavior, employment, health (e.g., Eurobarometer, OECD)
- **Experimental/observational numeric datasets**: Controlled experiments, lab studies, field measurements
- **Purpose**: Analyze trends, correlations, and causal relationships quantitatively for evidence-based conclusions

EXAMPLES OF DATABASES IN ECONOMICS & MANAGEMENT

World Bank Open Data

Eurostat

IMF (International Monetary Fund)

OECD Data

National Statistics Offices (e.g., INE Portugal)

THANKS!

Questions?