

STATISTICAL LABORATORY

Lecture 2: Exploratory Data Analysis (Practice)

Yiu Lim LUI, Allen

Lisbon School of Economics & Management
Universidade de Lisboa

`allenlui@iseg.ulisboa.pt`

`https://sites.google.com/view/allenlui`

Applied Mathematics for Economics and Management
1st Year / 1st Semester – 2026/2027

Exercise 1.9

What is the general procedure underlying empirical research in the fields of Economics and Management?

Exercise 1.9 (Answer)

Empirical work in economics and management follows a systematic procedure consisting of the following steps:

- 1 Define research question: Clearly establish what is being investigated.
- 2 Collect relevant data: Gather data from sources such as surveys, experiments, administrative records, and financial databases.
- 3 Organize and process data: Handle data coding, systematic recording, validation, and cleaning.
- 4 Apply descriptive and inferential statistical methods: Use quantitative techniques to analyze patterns, relationships, or effects.
- 5 Interpret and report results: Draw conclusions and report findings.

Exercise 1.10

Provide examples of databases available for conducting empirical research in the fields of Economics and Management.

Exercise 1.10 (Answer)

We can categorize available databases in economics and management into two main groups:

Economic Databases:

- Eurostat: Official European Union statistics.
- World Bank Open Data: Global development metrics.
- INE (Portugal): National statistical datasets.
- OECD Data: Economic and structural policies.

Financial Databases:

- Bloomberg / Refinitiv: Real-time financial market data.
- S&P Capital IQ / Orbis: Corporate reporting and accounts.
- Yahoo Finance: Market indices and stock data.

Exercise 2.1

For quality-control purposes, the product is shipped in boxes of six units. A simple random sample without replacement of 40 boxes is drawn from the lot. For each box, the number of defective units is recorded. The observed sample is:

2, 0, 1, 3, 0, 3, 6, 4, 5, 3, 0, 1, 0, 3, 2, 0, 1, 0, 2, 3, 1, 0, 3, 0, 1, 0, 0, 1, 5, 6, 0, 2, 5, 1, 0, 3, 0, 2, 3, 0

- a) Define the statistical unit and the variable under analysis, and classify the variable.
- b) Present the frequency table.
- c) Provide a graphical representation of the data.

Exercise 2.1 (a) (Answer)

- a) Statistical Unit: Box of 6 units; Variable: number of defective products per box; Classification of the variable: Quantitative discrete variable (since it takes numerical integer values representing counts).

Exercise 2.1 (b) (Answer)

Let x_i represent the number of defective units, n_i the absolute frequency, f_i the relative frequency, N_i the cumulative absolute frequency, and F_i the cumulative relative frequency.

Table: Frequency Distribution of Defective Units per Box

x_i	n_i	f_i	N_i	F_i
0	14	0.350	14	0.350
1	7	0.175	21	0.525
2	5	0.125	26	0.650
3	8	0.200	34	0.850
4	1	0.025	35	0.875
5	3	0.075	38	0.950
6	2	0.050	40	1.000
Total	40	1.000	–	–

Exercise 2.1 (c) (Answer)

