



PRACTICE UNIT 6 – DESCRIBING BIVARIATE DATA. FREQUENCY DISTRIBUTIONS AND GRAPHIC PRESENTATIONS.

OBJECTIVES

The objectives of this practice are:

- Strengthen the use of the Excel spreadsheet.
- Present bivariate data in frequency tables: joint, marginal and conditional distributions.
- Present bivariate data graphically: clustered and stacked bar charts, frequency polygons.
- Describe bivariate data numerically: calculation of numerical measures.
- Analyse the statistical dependence or independence between two variables.

EXERCISE U6_1. *Bivariate descriptive analysis of two variables, one qualitative and the other quantitative* (sheet *Height-Gender*)

Sheet '*Height-Gender*' shows the bivariate joint frequency distribution of the variables *Height* and *Gender* for the students that answered the survey.

- a) In which group of students is the height more homogeneous, in the male or in the female student's group?
- b) Obtain an appropriate location measure for the variable "Gender"
- c) Obtain the bivariate relative frequency distribution.
- d) Analyse if the variables *Height* and *Gender* are statistically independent.
- e) Obtain the distributions of the student's gender given student's height. Which profiles, row or column, are these? Present these profiles graphically.
- f) Obtain the distributions of the student's height given the gender. Which profiles, row or column, are these? Present these profiles graphically.
- g) Among the female students, what is the percentage of them whose height is greater than 1.80m?



EXERCISE U6_2. Bivariate descriptive analysis of two quantitative variables (sheet *Height-Weight*)

Sheet '*Height-Weight*' shows the bivariate joint frequency distribution of the variables *Height* and *Weight* for the students that answered the survey. The two variables have been grouped into classes.

- a) Calculate the average weight of the students using the values of the frequency table. Is this value representative of the behaviour of the variable?
- b) Among the tallest students (height greater than 1.84m), what is the most frequent weight?
- c) What is the height that 20% of the students with a weight smaller than 69kg have more than?
- d) Obtain the row and column profiles and interpret them in the context of the data. Present one of these graphically.
- e) Are the variables *Height* and *Weight* statistically independent?
- f) Calculate the bivariate relative joint frequency distribution.