PRACTICE UNIT 3. DESCRIBING UNIVARIATE DATA. FRECUENCY TABLES AND GRAPHIC PRESENTATION

OBJECTIVES

The objectives of this practice are:

- Handle data obtained in a survey.
- Handle big amounts of data and use the basic options of the Excel spreadsheet.
- Present data in frequency tables and with graphics.
- Answer basic questions from the frequency table.

DATA COLLECTION

The main data to be used in this practice are data obtained from a survey of a group of students, although other data sets are also used.

In the Excel file, you will find the collected data from the survey. The sheet 'Variables' contains the description of each one of the questions that the students answered. The name given to each variable (one for each question) is also provided.

Sheet 'Data' contains the data provided by the students. It was observed that some students made mistakes when introducing their answers. The values that were detected as wrong were highlighted in different colours. In green, those that could be corrected, and in red, those for which the correction was not possible.

Sheet 'Refined data' contains the updated data, incorporating the corrections for the values highlighted in green and eliminating those in red. Moreover, for each qualitative variable, a new variable was created, in which the real values were coded with numbers.

PART A. TABULATION AND GRAPHIC PRESENTATION FROM THE FREQUENCY DISTRIBUTION

Exercise U3_A1. (Sheet '*A_Place of residence*')

Using the frequency distribution provided for the variable *Residence*:

- **a)** Describe the population, the type and the scale of measurement of this variable.
- **b)** Complete the frequency distribution of this variable (as detailed as possible).
- **c)** What percentage of students have their place of residence outside Aragón?
- **d)** Present the frequency distribution using an appropriate graph. What fact would you highlight when looking at this graph?

Exercise U3_A2. (Sheet '*A_Number of siblings*')

Using the frequency distribution provided for the variable *Siblings* (Number of siblings):

- **a)** Describe the type and the scale of measurement of this variable.
- **b)** Complete the frequency distribution of this variable (as detailed as possible).
- **c)** What is the number of students that have no brothers nor sisters? What percentage do they represent?
- **d)** What percentage of students have 2 siblings (brothers or sisters) at least?
- **e)** Present the frequency distribution using an appropriate graph. What fact would you highlight when looking at this graph?

Exercise U3_A3. (Sheet '*A_Height*')

Using the frequency distribution provided for the variable *Height* (Height of the student):

- **a)** Describe the type and the scale of measurement of this variable.
- **b)** Complete the frequency distribution of this variable (as detailed as possible).
- **c)** Complete the frequency distribution of this variable separately for the male students and the female students.
- **d)** What percentage of students measure between 160 and 180 cm height?
- **e)** What percentage of female students measure more than 180 cm height? Among the male students, what percentage measure less than 170 cm?
- **f)** Present the frequency distribution using an appropriate graph. Repeat the same type of graph for the male and for the female student's frequency distributions separately. What fact would you highlight when looking at these graphs?

Exercise U3_A4. (*Sheet 'A_Balance'*)

The following table shows the quarterly average balance (in euros) in the bank accounts of 75 clients of a bank branch:

Quarterly average balance (euros)	Number of clients
0-600	10
600-1200	15
1200-1800	35
1800-3000	10
3000-6000	5

a) Describe the population, the type and the scale of measurement of this variable.

b) Complete the frequency distribution of this variable (as detailed as possible).

c) What percentage of clients have a quarterly average balance of more than 1,800 euros?

d) Present the frequency distribution using an appropriate graph. What fact would you highlight when looking at this graph?

PART B. TABULATION AND GRAPHIC PRESENTATION WHEN THE COMPLETE DATA SET IS AVAILABLE

Exercise U3_B1. (Sheet '*B_Place of residence*')

Using the values of the variable *Residence* (Place of residence):

- **a)** Obtain the frequency distribution of this variable (as detailed as possible).
- **b)** What percentage of students have their place of residence outside Aragón?
- **c)** Present the frequency distribution using an appropriate graph. What fact would you highlight when looking at this graph?

Exercise U3_B2. (Sheet '*B_Number of siblings*')

Using the values of the variable *Siblings* (Number of siblings):

- **a)** Obtain the frequency distribution of this variable (as detailed as possible).
- **b)** What is the number of students that have no brothers nor sisters? What percentage do they represent?
- **c)** What percentage of students have 2 siblings (brothers or sisters) at least?
- **d)** Present the frequency distribution using an appropriate graph. What fact would you highlight when looking at this graph?

Exercise U3_B3. (Sheet '*B_Height*')

Using the values of the variable *Height* (Height of the student):

- **a)** Obtain the frequency distribution of this variable (as detailed as possible), grouping data in 9 classes of width 5 starting in 150.
- **b)** Obtain the frequency distribution of this variable separately for the male students and the female students, using the same grouping used in the previous question.
- **c)** What percentage of students measure between 160 and 180 cm height?
- **d)** What percentage of female students measure more than 180 cm height? Among the male students, what percentage measure less than 170 cm?
- **e)** Present the frequency distribution using an appropriate graph. Repeat the same type of graph for the male and for the female student's frequency distributions separately. What fact would you highlight when looking at these graphs?