PROPOSED EXERCISES UNIT 6

1) Sheet '*Sports-Gender*' shows the bivariate joint frequency distribution of the variables *Sport* (how often the student participates in sporting activities) and *Gender* (Male/Female) for the students of the course Statistics I (course 2014-2015) that answered the survey conducted with them.

a) Calculate the bivariate relative joint frequency distribution. What is the number of students that participate rarely in sporting activities and are female students? And the percentage of students that practice sport regularly and are male students?

b) Present the bivariate joint frequency distribution graphically.

c) Obtain the frequency distributions of the gender variable given how often the student participates in sporting activities. Which profiles, row or column, are these? What is the percentage of female students among the students that practice sport frequently?

d) Obtain the frequency distributions of the Sport variable given the gender. Which profiles, row or column, are these? Among the male students, what is the percentage of students that practice sport regularly?

e) Present the row and column profiles graphically. What can be deduced when looking at them?

f) Analyse if the variables 'Sport' and 'Gender' are statistically independent.

2) The table shows information referring to the salaries of the workers of a company installed in five autonomous communities (sheet '*Salary-Community*').

X/Y	Castilla León	Castilla La Mancha	Extremadura	Andalucía	Valencia
500 – 1000	12	26	12	12	26
1000 – 3000	50	23	25	50	36
3000 – 4500	50	36	18	35	35
4500 – 6000	18	15	5	20	12

a) What is the average salary of the workers of this company? Is this value representative of the whole set of workers?

b) Which Autonomous Community has the greatest number of workers?

c) In which autonomous community is the distribution of salaries more homogenous, in 'Castilla-León' or in 'Castilla-La Mancha'?

d) Obtain the frequency distributions of the Autonomous Communities given the salary. Which profiles, row or column, are these? What is the percentage of workers in Extremadura among those with a salary over 4500 euros?

e) Present the previous distributions graphically. What can be observed in this graphic presentation?

f) Obtain the bivariate relative joint frequency distribution. Using these values, analyse if the variables *Salary* and *Autonomous Community* are statistically independent? Present the joint frequency distribution graphically.

g) Obtain the distributions of the salaries for each Autonomous Community. Which profiles, row or column, are these? Present them graphically. What can be observed in this graphic presentation?