

PROPOSED EXERCISES UNIT 7

1) (Sheet '*Farm_production*') The table shows the production costs (X) and the total income (Y), both in millions of euros, of a family farm during the last seven years.

X	Y
1.9	5.5
2.2	7.4
2.9	9.8
3.6	11.6
3.8	11.6
4.6	12.2
5.5	11.2

- A) Obtain the least squares regression line that explains Y based on X. Interpret the parameters of the estimated regression line.
- B) Analyse the goodness of fit of this model numerically and graphically.

2) (Sheet '*Population*') City planners think that the most populous cities are inhabited by residents with older ages. To investigate this relationship, they gather data about the population and the average age in 10 big cities, obtaining the results shown in the table.

City	Population (in millions)	Average age (in years)
Chicago	2,833	31,5
Dallas	1,233	30,5
Houston	2,144	30,9
Los Angeles	3,849	31,6
New York	8,214	34,2
Philadelphia	1,448	34,3
Phoenix	1,513	30,7
San Antonio	1,297	31,7
San Diego	1,257	32,5
San Jose	0,93	32,6

- A) Present the data graphically using a scatter plot (select the age as the endogenous variable). What type of association does the plot suggest?
- B) Obtain the least squares regression line that explains the average age based on the amount of population. Interpret the parameters of the estimated regression line.
- C) Analyse the goodness of fit of this model numerically.
- D) What would be the estimated average age in a city with 2.5 million of inhabitants? Is this estimation reliable?