Exploration 7A: Correlation Activity

Instructions: Using data \_le C02 Homes.xls [.rda], answer each question below.

1. Compute the mean and standard deviation for each of the following numerical variables:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Taxes | Year | Acres | Size | Value | Price |
| Mean |  |  |  |  |  |  |
| Standard Deviation |  |  |  |  |  |  |

1. Using the mean as a model, how much would you say the typical single-family home costs in this market?
2. How reliable is your estimate?
3. Using a table of correlations, calculate the correlation coefficient (r) for the following pairs of variables:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Taxes | Style | Bath | Bed | Rooms | Year | Acres | Size | Value |
| Price |  |  |  |  |  |  |  |  |  |

1. Based on the correlation coefficients, which of the above variables seems to have the MOST effect on the PRICE of a house? Which as the LEAST effect?
2. Generate a scatterplot that describes the relationship between PRICE and SIZE. Which variable is the independent variable (should be on the x-axis)? Which variable is the dependent variable (on the y-axis)? What is the Correlation for this relationship? Your scatterplot should look something like figure 7.11.
3. Draw a vertical line on the above chart to represent the MEAN for SIZE.

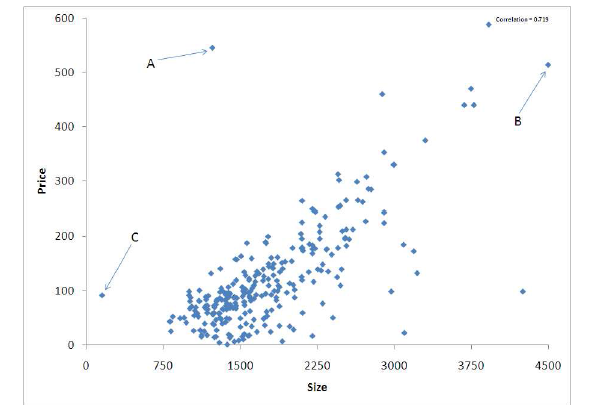


Figure 7.11: Scatterplot showing home price versus size.

1. Draw a horizontal line on the above chart to represent the MEAN for PRICE.
2. Hover your mouse over the points marked A, B and C on the chart to determine the values for PRICE and SIZE at each point. Then fill in the table below to estimate the correlation.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Point | Size | Price | z-score for Size (X) | z-score for Price (Y) | Total Contribution to the Numerator of Correlation |
| A |  |  |  |  |  |
| B |  |  |  |  |  |
| C |  |  |  |  |  |