Simple Linear Regression and Correlation: Exercises 16.1, 16.7, 16.28, and 16.100

Multiple Regression: Exercises 17.2 and 17.5

Directions:

-Submit your answers in a Microsoft Excel workbook, with each problem on a separate worksheet.

-Label each tab in the workbook with the exercise number.

-Highlight the answers in yellow and provide an interpretation in a text box.

-Cite any sources using the APA format on a separate page

**16.1**

The term regression was originally used in 1885 by Sir Francis Galton in his analysis of the relationship between the heights of children and parents. He formulated the “law of universal regression,” which specifies that “each peculiarity in a man is shared by his kinsmen, but on average in a less degree.” (Evidently, people spoke this way in 1885.) In 1903, two statisticians, K. Pearson and A. Lee, took a random sample of 1,078 father–son pairs to examine Galton’s law (“On the Laws of Inheritance in Man, I. Inheritance of Physical Characteristics,” Biometrika 2:457–462). Their sample regression line was Son’s height = 33.73 + .516 × Father’s height

a. Interpret the coefficients.

b. What does the regression line tell you about the heights of sons of tall fathers?

c. What does the regression line tell you about the heights of sons of short fathers?

**16.7**

Florida condominiums are popular winter retreats for many North Americans. In recent years, the prices have steadily increased. A real estate agent wanted to know why prices of similar-sized apartments in the same building vary. A possible answer lies in the floor. It may be that the higher the floor, the greater the sale price of the apartment. He recorded the price (in $1,000s) of 1,200 sq. ft. condominiums in several buildings in the same location that have sold recently and the floor number of the condominium.

a. Determine the regression line.

b. What do the coefficients tell you about the relationship between the two variables?

**16.28 (Refer to Exercise 16.6)**

**16.28**

a. What is the standard error of estimate? Interpret its value.

b. Describe how well the memory test scores and length of television commercial are linearly related.

c. Are the memory test scores and length of commercial linearly related? Test using a 5% significance level.

d. Estimate the slope coefficient with 90% confidence.

**(16.6)**

In television’s early years, most commercials were 60 seconds long. Now, however, commercials can be any length. The objective of commercials remains the same—to have as many viewers as possible remember the product in a favorable way and eventually buy it. In an experiment to determine how the length of a commercial is related to people’s memory of it, 60 randomly selected people were asked to watch a 1-hour television program. In the middle of the show, a commercial advertising a brand of toothpaste appeared. Some viewers watched a commercial that lasted for 20 seconds, others watched one that lasted for 24 seconds, 28 seconds, …, 60 seconds. The essential content of the commercials was the same. After the show, each person was given a test to measure how much he or she remembered about the product. The commercial times and test scores (on a 30-point test) were recorded.

a. Draw a scatter diagram of the data to determine whether a linear model appears to be appropriate.

b. Determine the least squares line.

c. Interpret the coefficients.

**16.100** *Conduct all tests of hypotheses at the 5% significance level*

The president of a company that manufactures car seats has been concerned about the number and cost of machine breakdowns. The problem is that the machines are old and becoming quite unreliable. However, the cost of replacing them is quite high, and the president is not certain that the cost can be made up in today’s slow economy. To help make a decision about replacement, he gathered data about last month’s costs for repairs and the ages (in months) of the plant’s 20 welding machines.

a. Find the sample regression line.

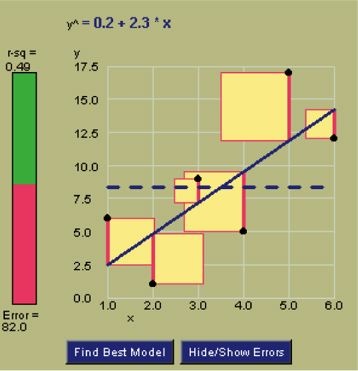
b. Interpret the coefficients.

c. Determine the coefficient of determination, and discuss what this statistic tells you.

d. Conduct a test to determine whether the age of a machine and its monthly cost of repair are linearly related.

e. Is the fit of the simple linear model good enough to allow the president to predict the monthly repair cost of a welding machine that is 120 months old? If so, find a 95% prediction interval. If not, explain why not.

**17.2 Change the slope (if necessary) so that the line is horizontal.**



What is the y-intercept?

**17.5**

Drag the mouse to change the slope to .5. What is the sum of squared errors?