

Chapter 10: Regression Analysis: Estimating Relationships Problems
Book Title: Business Analytics: Data Analysis and Decision Making
Printed By: Rita Manuel-Byrd (rmanuelbyrd@gmail.com)
© 2017 South-Western, Cengage Learning

Problems

Note: Student solutions for problems whose numbers appear blue are available at <http://www.cengage.com/decisionsciences/albright/ba/6e>.

Level A

1. Explore the relationship between the selling prices (Y) and the appraised values (X) of the 148 homes in the file [P02_11.xlsx](#) by estimating a simple linear regression model. Interpret the standard error of estimate s_e and R^2 and the least squares line for these data.
 - a. Is there evidence of a *linear* relationship between the selling price and appraised value? If so, characterize the relationship. Is it positive or negative? Is it weak or strong?
 - b. For which of the three remaining variables, the size of the home, the number of bedrooms, and the number of bathrooms, is the relationship with the home's selling price *stronger*? Justify your choice with additional simple linear regression models.
2. The file [P02_10.xlsx](#) contains midterm and final exam scores for 96 students in a corporate finance course. Each row contains the two exam scores for a given student, so you might expect them to be positively correlated.
 - a. Create a scatterplot of the final exam score (Y) versus the midterm score (X). Based on the visual evidence, would you say that the scores for the two exams are strongly related? Is the relationship a linear one?
 - b. Superimpose a trend line on the scatterplot, and use the option to display the equation and the R^2 value. What does this equation indicate in terms of predicting a student's final exam score from his or her midterm score? Be specific.
 - c. Run a regression to confirm the trend-line equation from [part b](#). What does the standard error of estimate say about the accuracy of the prediction requested in [part b](#)?
3. A company produces electric motors for use in home appliances. One of the company's production managers is interested in examining the relationship

Problems

Level A

16. A trucking company wants to predict the yearly maintenance expense (Y) for a truck using the number of miles driven during the year (X_1) and the age of the truck (X_2 , in years) at the beginning of the year. The company has gathered the data given in the file [P10_16.xlsx](#), where each observation corresponds to a particular truck.

a. Estimate a multiple regression equation using the given data. Interpret each of the estimated regression coefficients. Why is the magnitude of the Miles Driven coefficient so much lower than the magnitude of the Age of Truck coefficient? Is it because Miles Driven is not as important in predicting Maintenance Expense?

b. Interpret the standard error of estimate s_e and R^2 for these data.

17. DataPro is a small but rapidly growing firm that provides electronic data-processing services to commercial firms, hospitals, and other organizations. For each of the past several months, DataPro has tracked the number of contracts sold, the average contract price, advertising expenditures, and personal selling expenditures. These data are provided in the file [P10_17.xlsx](#). Using the number of contracts sold as the dependent variable, estimate a multiple regression equation with three explanatory variables. Interpret each of the estimated regression coefficients, the standard error of estimate, and R^2 .

18. An antique collector believes that the price received for a particular item increases with its age and with the number of bidders. The file [P10_18.xlsx](#) contains data on these three variables for several recently auctioned comparable items.

a. Estimate a multiple regression equation using the given data. Interpret each of the estimated regression coefficients. Is the antique collector correct in believing that the price received for the item increases with its age and with the number of bidders?

- d. According to the estimated regression equation, what is the difference between the mean number of worker-hours required for erecting boilers with steam drums and those with mud drums?
- e. Given the estimated regression equation, predict the number of worker-hours needed to erect a utility-field, steam-drum boiler with a capacity of 550,000 pounds per hour and a design pressure of 1400 pounds per square inch. How accurate is your prediction?
- f. Given the estimated regression equation, predict the number of worker-hours needed to erect an industrial-field, mud-drum boiler with a capacity of 100,000 pounds per hour and a design pressure of 1000 pounds per square inch. How accurate is your prediction?

28. Suppose that a regional express delivery service company wants to estimate the cost of shipping a package (Y) as a function of cargo type, where cargo type includes the following possibilities: fragile, semifragile, and durable. Costs for several randomly chosen packages of approximately the same weight and same distance shipped, but of different cargo types, are provided in the file [P10_28.xlsx](#).

- a. Estimate an appropriate multiple regression equation to predict the cost of shipping a given package.
- b. Interpret the estimated regression coefficients. You should find that the estimated intercept and slope of the equation are sample means. Which sample means are they?
- c. According to the estimated regression equation, which cargo type is the *most* costly to ship? Which cargo type is the *least* costly to ship?
- d. How well does the estimated equation fit the given sample data? How do you think the model's goodness of fit could be improved?
- e. Given the estimated regression equation, predict the cost of shipping a package with semifragile cargo.

29. The file [P10_11.xlsx](#) contains annual observations (in column B) of the American minimum wage. The basic question here is whether the minimum wage has been growing at roughly a *constant* rate over this period.

- a. Create a time series graph for these data. Comment on the observed behavior of the minimum wage over time.