**SMDA – Assignment 2**

**Instructions:**

1. Last date of submission is 09 December 2022.

**Q1.** In an attempt to assess total daily travel taxes in various cities, the Global Business Travel Association conducted a study of daily travel taxes on lodging, rental car, and meals (GBTA Foundation website, October 30, 2012). The data contained in the DATAfile named **TravelTax** are consistent with the findings of that study for business travel to Chicago. Assume the population standard deviation is known to be $8.50 and develop a 95% confidence interval of the population mean total daily travel taxes for Chicago.

**Q2.** The International Air Transport Association surveys business travelers to develop quality ratings for transatlantic gateway airports. The maximum possible rating is 10. Suppose a simple random sample of 50 business travelers is selected and each traveler is asked to provide a rating for the Miami International Airport. The ratings obtained from the sample of 50 business travelers is provided in the datafile **MIAMI**.

Develop a 95% confidence interval estimate of the population mean rating for Miami.

**Q3.** The mean length of a work week for the population of workers was reported to be 39.2 hours (Investor’s Business Daily, September 11, 2000). Suppose that we would like to take a current sample of workers to see whether the mean length of a work week has changed from the previously reported 39.2 hours.

a. State the hypotheses that will help us determine whether a change occurred in the mean length of a work week.

b. Suppose a current sample of 112 workers provided a sample mean of 38.5 hours. Use a population standard deviation σ = 4.8 hours. What is the p-value?

c. At α = .05, can the null hypothesis be rejected? What is your conclusion?

d. Repeat the preceding hypothesis test using the critical value approach.

**Q4.** Joan’s Nursery specializes in custom-designed landscaping for residential areas. The estimated labor cost associated with a particular landscaping proposal is based on the number of plantings of trees, shrubs, and so on to be used for the project. For cost estimating purposes, managers use two hours of labor time for the planting of a medium sized tree. Actual times from a sample of 10 plantings during the past month follow (times in hours).

1.7 1.5 2.6 2.2 2.4 2.3 2.6 3.0 1.4 2.3

With a .05 level of significance, test to see whether the mean tree-planting time differs from two hours.

a. State the null and alternative hypotheses.

b. Compute the sample mean.

c. Compute the sample standard deviation.

d. What is the p-value?

e. What is your conclusion?

**Q5.** The Heldrich Center for Workforce Development found that 40% of Internet users received more than 10 e-mail messages per day (USA Today, May 7, 2000). A similar study on the use of e-mail was repeated in 2002.

a. Formulate the hypotheses that can be used to determine whether the proportion of Internet users receiving more than 10 e-mail messages per day increased.

b. If a sample of 425 Internet users found 189 receiving more than 10 e-mail messages per day, what is the p-value?

c. At α = .05, what is your conclusion?

**Q6.** Microsoft Outlook is the most widely used e-mail manager. A Microsoft executive claims that Microsoft Outlook is used by at least 75% of Internet users. A sample of Internet users will be used to test this claim.

a. Formulate the hypotheses that can be used to test the claim.

b. A Merrill Lynch study reported that Microsoft Outlook is used by 72% of Internet users (CNBC, June 2000). Assume that the report was based on a sample size of 300 Internet users. What is the p-value?

c. At α = .05, should the executive’s claim of at least 75% be rejected?

**Q7.** An auto manufacturing company wanted to investigate how the price of one of its car models depreciates with age. The research department at the company took a sample of eight cars of this model and collected the following information on the ages (in years) and prices (in hundreds of dollars) of these cars.



a. Construct a scatter diagram for these data. Does the scatter diagram exhibit a linear relationship between ages and prices of cars?

b. Find the regression line with price as a dependent variable and age as an independent variable.

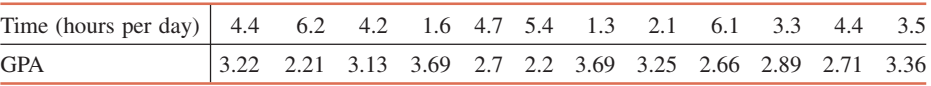
c. Give a brief interpretation of the values of a and b calculated in part b.

d. Plot the regression line on the scatter diagram of part a and show the errors by drawing vertical lines between scatter points and the regression line.

e. Predict the price of a 7-year-old car of this model.

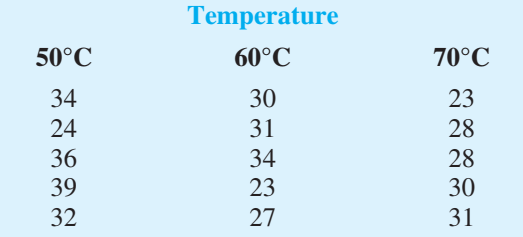
f. Estimate the price of an 18-year-old car of this model. Comment on this finding.

**Q8.** The following table contains information on the amount of time spent each day (on average) on social networks and the Internet for social or entertainment purposes and the grade point average for a random sample of 12 college students.

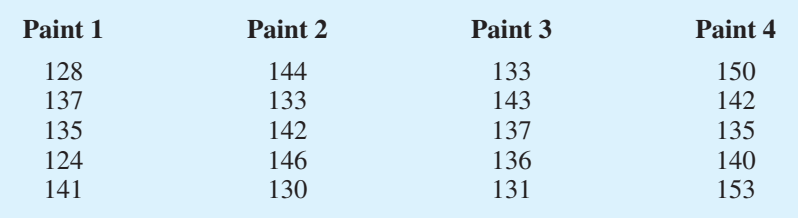


Compute the coefficient of determination, and give a brief interpretation of it. What percentage of the variation in GPA is explained by the least squares regression line of GPA on time? What percentage is not explained?

**Q9.** To study the effect of temperature on yield in a chemical process, five batches were produced at each of three temperature levels. The results follow. Construct an analysis of variance table. Use a .05 level of significance to test whether the temperature level has an effect on the mean yield of the process.



**Q10.** Four different paints are advertised as having the same drying time. To check the manufacturer’s claims, five samples were tested for each of the paints. The time in minutes until the paint was dry enough for a second coat to be applied was recorded. The following data were obtained.



At the a = .05 level of significance, test to see whether the mean drying time is the same for each type of paint.