Statistics 101 – Homework 4

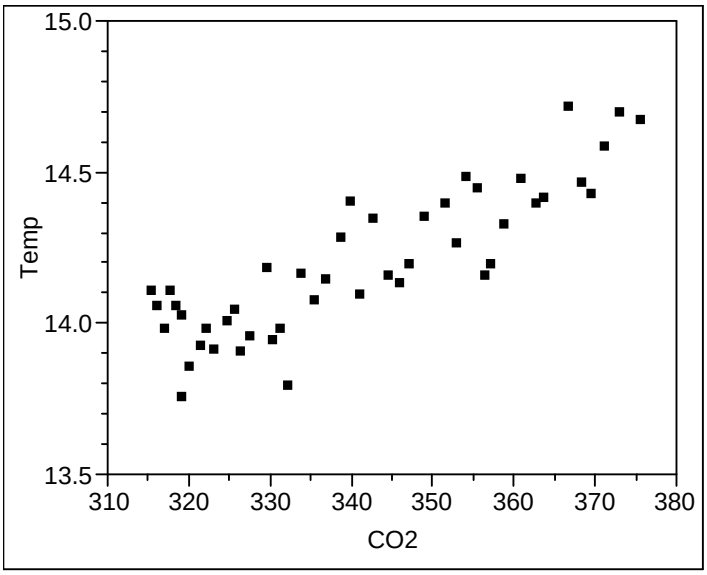
Due Wednesday, February 18, 2009

Homework is due on the due date at the end of the lab.

Reading: February 9 – February 13 Chapters 8 & 9
February 16 – February 18 Chapter 10

Assignment:

1. Global warming is a hot topic these days. One of the factors that may explain increases in global temperatures is the amount of carbon dioxide in the atmosphere. Annual atmospheric carbon dioxide (CO2) concentrations measured as parts per million by volume (ppmv) are derived from air samples collected at Mauna Loa Observatory in Hawaii. Additionally the annual average global temperatures (degrees Celsius) are recorded. The data plotted below are for years from 1958 through 2003.



a) Answer the questions Who? and What? for this problem.
b) The value of the correlation for these data is most likely to be?
–0.95 –0.67 –0.31 +0.03 +0.55 +0.86
Explain the reason for your choice.
c) Write a sentence or two explaining what this correlation means for these data. Write about CO2 concentrations and global temperatures rather than about correlation coefficients.
d) If the temperature were recorded in degrees Fahrenheit rather than degrees Celsius, how would the correlation change? Explain your answer.
e) A news reporter looking at these data writes; "It is clear that increasing CO2 in the atmosphere causes global temperature to increase." Explain why this is not an appropriate conclusion from these data alone.

2. The October 2005 issue of Consumer Reports (CR) has an article on the difference between Environmental Protection Agency (EPA) mileage ratings (miles per gallon) and mileage results (miles per gallon) obtained by Consumer Reports in test-track driving. The EPA mileage ratings are the ones posted on new car sales stickers. Below are mileage values for 10 different vehicle types.

Vehicle type	1	2	3	4	5	6	7	8	9	10
EPA mileage	15	21	13	48	22	17	20	13	26	18
CR mileage	9	13	8	26	11	10	12	8	17	11

a) Answer the questions Who? and What? for this problem.
b) Plot the data. Use EPA mileage as the explanatory variable and CR mileage as the response.
c) Compute the mean and standard deviation for the EPA mileage. Round final answers to 2 decimal places.
d) Compute the mean and standard deviation for the CR mileage. Round final answers to 2 decimal places.
e) The correlation between the EPA mileage and the CR mileage is r=0.9839. Explain in words what this correlation means.
f) Compute the estimate of the slope for the least squares regression line. Round final answer to 4 decimal places.
g) Give an interpretation of the estimated slope within the context of the problem.
h) Compute the estimate of the intercept for the least squares regression line. Round final answer to 4 decimal places.
i) Give the equation of the least squares regression line. Use this equation to predict the Consumer Reports mileage for a vehicle the EPA rates at 30 miles per gallon.
j) How would you describe vehicle 4? Choose all that apply: outlier in regression, high leverage value, influential value. Explain your choice(s) briefly.
k) Give the value of R^2 for this regression. Give an interpretation of this value within the context of the problem.

3. The December 2003 issue of Kiplinger's Personal Finance published data on the 2004 model year cars and trucks. The weight (pounds) and EPA estimated highway mileage (mpg) for 57 cars are displayed below. The data, Fuel Economy Data, are on the main Stat 101 L course page (www.stat.iastate.edu/courses/stat101L.html) under Homework 4. Follow the instructions in the JMP Guide to download/open the data set. Use JMP to look at the distribution of Highway mpg and the relationship between Weight and Highway mpg. Use the JMP output to help you answer the questions below. Be sure to attach the JMP output to your assignment.

a) Describe the distribution of Highway mpg values. Make sure to include in your description the five number summary, the mean and standard deviation, and the shape of the histogram. Are there any outliers?
b) Describe the scatter plot of Highway mpg versus Weight. Give the regression equation for predicting Highway mpg from Weight. Use the regression equation to predict the Highway mpg for a car weighing 3000 pounds. Give an interpretation of the estimated slope. Give the value of R^2 and an interpretation of this value. Finally, describe the plot of residuals versus weight values and make note of any potential problems with the regression.