

Instructions: This project is worth 20 out of 100 total points for exam 2. This is an individual project. You are only permitted to seek help or clarification from me. Submit your TYPED responses and graphs in hard copy form. Use proper English grammar, spelling and punctuation.

The following data are from the website <http://lib.stat.cmu.edu/DASL/Datafiles/BreastCancer.html>. The data consist of breast cancer mortality rates and mean temperature for several regions in Europe. Read the description then answer the questions.

Reference: A.J. Lea. (1965). New Observations on Distribution of Neoplasms of Female Breast in Certain Countries. *British Medical Journal*, 1, 488-490.

Description: Data contains the mean annual temperature (in degrees F) and Mortality Index for neoplasms of the female breast. Data were taken from certain regions of Great Britain, Norway, and Sweden.

Number of cases: 16

Variable Names:

1. Mortality: Mortality index for neoplasms of the female breast
2. Temperature: Mean annual temperature (in degrees F)

The Data:

Mortality	Temperature
102.5	51.3
104.5	49.9
100.4	50
95.9	49.2
87	48.5
95	47.8
88.6	47.3
89.2	45.1
78.9	46.3
84.6	42.1
81.7	44.2
72.2	43.5
65.1	42.3
68.1	40.2
67.3	31.8
52.5	34

Questions

1. Use R Commander (or some other computer software of your choice) to obtain a presentation-quality scatterplot of these data. Turn in this scatterplot. Does the relationship between the two variables appear to be linear?
2. Use R Commander to calculate the best fit regression line and correlation coefficient. Give the fitted line in the form $y = a + bx$.
3. Interpret the slope of the regression line. Specifically, indicate what the slope tells you quantitatively about the relationship between temperature and mortality.
4. Interpret the correlation coefficient.
5. Predict the mortality for a region having mean temperature 62 degrees Fahrenheit.
6. Do these data indicate that higher temperatures *cause* breast cancer mortality? Why or why not?
7. If the temperature is converted to degrees Celsius, does the correlation coefficient change? Does the fitted regression line change? Give reasons for your response.