**INTD 8065 Data Analysis for Cancer Research**

**Homework #6**

A suspicious set of data found in Internet. (Dept. of Mathematical and Statistical Sciences. University of Alberta, Canada). The data were originally published in A.J. Lea (1965), British Medical Journal, Volume 1, pp. 488-490.

The data was presumably obtained in different regions of England, Norway and Sweden for mean annual temperature (in °F) and mortality index for neoplasms of the female breast (deaths for 100000 females)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Mortality rate (M) | 102.5 | 104.5 | 100.4 | 95.9 | 87.0 | 95.0 | 88.6 | 89.2 |
| Temperature ( T) | 51.3 | 49.4 | 50.0 | 49.2 | 48.5 | 47.8 | 47.3 | 45.1 |
|  |  |  |  |  |  |  |  |  |
| Mortality rate (M) | 78.9 | 84.6 | 81.7 | 72.2 | 65.1 | 68.1 | 67.3 | 52.5 |
| Temperature ( T) | 46.3 | 42.1 | 44.2 | 43.5 | 42.3 | 40.2 | 31.8 | 34.0 |

1. Plot the data
2. Fit the regression
3. Which parameter estimates are significant?
4. How much of the variance is explained by the regression?
5. Plot the residuals. What is your diagnostics?
6. Find the confidence intervals of the regression line for the following values of T: 35,40,45,50
7. Find the prediction interval of the regression line for the same values of previous question.
8. What is your overall conclusion?