# Assignment 3

Due date for this assignment is at 12.00 pm on 21st Feb 2018

Submit your answer in hard-copy (hand written or word document print out). If you are not plotting the graph by hand, print and attach all your graphs. If you solve any of the questions with R, email a softcopy of your .r file to me latest by 12.00 pm on 21st Feb 2018.

1. In a data analysis study conducted by personnel at the Statistics Consulting Centre at Virginia Tech, two different materials, alloy A and alloy B, were compared in terms of breaking strength. Alloy B is more expensive, but it should certainly be adopted if it can be shown to be stronger than alloy A. The consistency of performance of the two alloys should also be taken into account. Random samples of beams made from each alloy were selected, and strength was measured in units of 0.001-inch defection as a fixed force was applied at both ends of the beam (The smaller measured value would suggest a stronger alloy). Twenty specimens were used for each of the two alloys. The data are given as below:

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | 88 | 79 | 84 | 89 | 81 | 87 | 80 | 78 | 83 | 81 | 83 | 82 | 79 | 86 | 85 | 82 | 85 | 87 | 80 | 85 |
| B | 75 | 77 | 86 | 84 | 80 | 76 | 85 | 79 | 77 | 78 | 78 | 83 | 76 | 80 | 81 | 81 | 78 | 78 | 82 | 80 |

   a) Construct a quantile-quantile plot for each sample. Is it reasonable to assume that the two samples each came from a normal population?
   b) Assume that the samples are from normal populations; conduct a hypothesis test to compare the variances of the two samples. Use level of significance of 0.05.
   c) Is alloy B stronger than alloy A? Conduct a hypothesis test to answer this question. Use level of significance of 0.01.

2. Chemical and mechanical engineers often need to know the vapour pressure of water for specific temperatures. Physical chemistry suggests that the vapour pressure should follow an exponential relationship to the inverse of the temperature. Specifically, let $p$ be the vapour pressure and $T$ be the temperature, the Clausius Clapeyron equation states that natural log of $p$ is directly proportion to negative value of the inverse of $T$, that is

$$ln(p) \propto -\frac{1}{T}$$

   The following table lists the vapour pressures of water for various temperature (in Kelvin) from $0^oC$ to $100^oC$.

| Temperature (K) | Vapor pressure (mm Hg) |
|---|---|
| 273 | 4.6 |
| 283 | 9.2 |
| 293 | 17.5 |
| 303 | 31.8 |
| 313 | 55.3 |
| 323 | 92.5 |
| 333 | 149.4 |
| 343 | 233.7 |
| 353 | 355.1 |
| 363 | 525.8 |
| 373 | 760.0 |

a) Draw a scatter plot of the data. Comment on your scatter plot.
b) By referring to the Clausius Clapeyron equation, suggest a reasonable model for the vapour pressure over a wide range of temperatures.
c) Apply your suggested transformation from part (b) on the data set and produce a scatter plot of the transformed data. Fit the regression model to the transformed data and plot this regression line on the same page with the transformed data.
d) Verify the goodness of fit of your model in part (c).
e) Convert the prediction equation in part (c) back to the original metric and plot this curve on the same graph with the original data.

3. Icicles form on building eaves and trees in very cold environments. Scientists wanted to know more about the growth of these icicles. An icicle was grown in a cold chamber at -11$^{o}$C with no wind and a water flow of 11.9 mg/sec. The length (in cm) of the icicle was measured as a function of time (in minutes). The data is given below:

| Time (min) | Length (cm) |
|---|---|
| 10 | 0.6 |
| 20 | 1.8 |
| 30 | 2.9 |
| 40 | 4 |
| 50 | 5 |
| 60 | 6.1 |
| 70 | 7.9 |
| 80 | 10.1 |
| 90 | 10.9 |
| 100 | 12.7 |
| 110 | 14.4 |
| 120 | 16.6 |
| 130 | 18.1 |
| 140 | 19.9 |
| 150 | 21 |
| 160 | 23.4 |
| 170 | 24.7 |
| 180 | 27.8 |

a) Verify that linear model is not appropriate for this dataset.

b) Explore the data given above and determine on an appropriate regression model for it. Fit your choice of the regression model to the data and determine the goodness of fit. You may need to apply transformation to the data set in order to obtain a suitable regression model. You may repeat the process of exploration, transformation, model fitting and model checking for more than one time, either iteratively or simultaneously. But you only need to present your best fit model in the hardcopy submission. However, please keep your R code for all the transformation, model fitting and model checking, that you have applied (if any) in you final .r file submission.