**Homework 3**

The purpose of this assignment is to gain some practice of variable selection in R. This assignment is to apply variable selection tools to a simulated training dataset (training.txt) and compare the resultant models by comparing their prediction errors on an independently simulated test dataset (test.txt). The data include one response Y and twenty predictors X1 to X20. You need to submit the .pdf file including your code and analysis results.

The tasks of this homework are:

* Fit a linear regression on the whole **training data** and summarize the fitted model.
* Apply at least three variable selection methods (sequential techniques, AIC/BIC, cross validation, lasso, etc.) to find **a smaller model after variable selection**.
* Compare the three models from the above analysis by comparing their **prediction errors** on the independently simulated test data.
* Clearly summarize your analysis and results.
* Submit the .pdf file (generated from a .rmd file, same as the lab materials) to be graded.

This assignment is due on 12/16/2019, by 9:45 am. The total grade of this review is 20 points. The grading rubrics of this assignment is as follows.

* Fit a linear regression on the original data and summarize the fitted model (4 points).
* Application of each variable selection method to the training data (4 points \* 3).
* Comparison of the three models using prediction errors (4 points).