**Assignment administered by team 12: Polynomial Regressions and Step-Functions**

BUAD 5082 – Spring 2019

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**1. Objectives**

The purpose of this assignment is to provide you with some experience working with the poly() and cut() and functions and several of its supporting functions.

**2. What You Will Need**

Access to a Windows computer with R

**3. Due Date**

The solutions will be posted a week from today, February 18th

**4. Note on Collaboration**

This is a Category C assignment. Specifically, you may work with others or receive help from the instructor on this assignment. You must, however, turn in your own paper. You may not divide the work with others or copy other student’s work.

**5. Preliminaries**

As the first statement in your script file, enter rm(list=ls())

And set the random seed to 1.

Import dataset Auto from the packet “ISLR”

**Problem 1**: **Polynomial Regressions- Moving Beyond Linearity**

* 1. Fit a polynomial regression on Auto data, using Horsepower as an explanatory variable to explain MPG.
  2. Choose degrees of polynomial (print F-statistic using ANOVA).
  3. Print summary of the coefficients, describe the null hypothesis and the reason why you chose the specific degree of polynomial.
  4. Plot the data and the polynomial fit displaying the x-axis and y-axis names and the title with the degree of polynomial chosen.

**Problem 2: Predictions**

a. Fit a logistic regression to predict weather a car has MPG over 25 or not. Hint: create a dummy variable that assigns value of 1 to those who have MPG>25 and 0 to those who don’t.

b. Run the glm model created in the above step through the “prediction” function.

1. Calculate the Confidence Interval using the logit function.
2. Plot the predictions with the associated Confidence Intervals labeling the x-axis and y- axis and the title.
3. Comment of the findings or patterns that you observe in the graph.

**Problem 3: Step-Function**

* 1. Using the “cut” function create the bins or regions in the “Horsepower” variable (choose 6 number of cuts, Hint: look at the horsepower summary to come up with the cuts) \* cross validation using glm.
  2. Fit the “lm” function on the cut Horsepower variable to predict the MPG. Hint: lm(MPG~cut(Horsepower, “constant”), data=Auto)
  3. Display the coefficients using the coef(summary(fit)) function.
  4. Plot the step-function fit.