**IDS462 HW5**

Download the SAS dataset attached for this assignment, and see the problem descriptions below. The are 2 problems some with multiple parts in this assignment.

Your SAS code/program should start with the following:

/\* IDS462 HW5 \*/

/\* Name of student1 \*/

and if this is a team of more than 1 student contributed, then add the names of all students as

/\* Name of student2 \*/

….

Submit only one SAS code/program to solve all problems given, and make sure to put comments within in the beginning of each problem/part to show what you’re solving, for ex **\*Prob1-Part1;**

**Problem 1:**

The csv file attached provided the following information for 392 different car models:

Variables SAS variable names

Cylinders cyl

Displacement disp

Horse power HP

Weight wt

Acceleration accel

Miles per Gallon MPG

1. Create a SAS data-set, name it “cardata” with the variables names listed as above.
2. Determine an equation that can be used to predict MPG.
3. Which independent variables are insignificant? Why?
4. After eliminating the insignificant independent variables, run the regression with the rest of the variables:
   1. Provide statistical model equation
   2. Provide your analysis on the data accuracy
   3. Plot residuals versus predicted values. Are the errors randomly distributed?

You may use any selection model, FORWARD, BACKWARD, etc. Here’s an example on how to add: / selection=forward slentry=**0.1** P R;

**Problem 2:**

We want to predict students’ success in college by a battery of tests. Graduating seniors volunteer to take our test battery; and their final grade point average (GPA) is recorded. Using a MAXR technique, develop a prediction equation for the GPA using the test battery results.

Data is as follows with the numerical fields GPA HS\_GPA BOARD IQ, where hs\_gpa stands for the high school GPA.

3.9 3.8 680 130

3.9 3.9 720 110

3.8 3.8 650 120

3.1 3.5 620 125

2.9 2.7 480 110

2.7 2.5 440 100

2.2 2.5 500 115

2.1 1.9 380 105

1.9 2.2 380 110

1.4 2.4 400 110