**Homework-7**

|  |
| --- |
| Team#: \_\_\_  Team Member-1: Ishmam rahman - Member’s Contribution (in %) 33  Team Member-2: Bhagwat dhakal Member’s Contribution (in %) 33  Team Member-3: Luis Fuentas Member’s Contribution (in %) 3 |
| **Submission**   1. Answer the following questions 2. **Problem-1 and Problme-2:** Rename this word file to “HW7-YourTeam#” (e.g., HW7-Team1.docx) 3. **Problme-1 and Problem-2:** Rename your R file ““HW7\_YourTeam#” (e.g., HW7\_Team1.R) 4. Upload the files to the blackboard system |

Use **Auto** dataset that is available in the homework folder and write an R script to answer the following questions:

**Problem-1: Simple Regression [30 points]**

1. Plot **mpg** as the response and **horsepower** as the predictor. Use the **abline()** function to display the least squares regression line **[5 points]**.

**Answer:**

1. Use the lm() function to perform a simple linear regression with **mpg** as the response and **horsepower** as the predictor **[5 points]**. Use the summary() function to print the results and answer the following questions:
2. Is there a relationship between the predictor and the response? **[5 points]**

**Answer:** Yes

1. How strong is the relationship between the predictor and the response? **[5 points]**

**Answer:** The R-squared value is 0.6059 which means the model explains 60.59 % of the variation in mpg.

1. Is the relationship between the predictor and the response positive or negative? **[5 points]**

**Answer:** The relationship is negative.

1. What is the predicted mpg associated with a horsepower of 98? **[5 points]**

**Answer:** The predicted mpg for a horsepower of 98 is 24.47

**Problem-2: Multiple Regression [70 points]**

1. Produce a scatterplot matrix which includes all of the variables in the data set. You will need to exclude the name variable, which is qualitative. **[10 points]**

**Answer:**

1. Compute the matrix of correlations between the variables using the function cor(). Comments on the results. Specifically, comment about strength of each correlation and if it is a positive or negative correlation. **[10 points]**

**Answer:**

1. Use the lm() function to perform a multiple linear regression with **mpg** as the response and all other variables except name as the predictors **[10 points]**. Use the summary() function to print the results and answer the following questions:
2. Is there a relationship between the predictors and the response? **[10 points]**

**Answer:**

1. Which predictors appear to have a statistically significant relationship to the response? **[10 points]**

**Answer:**

1. What does the coefficient for the year variable suggest? **[10 points]**

**Answer:**

1. Use the \* symbol to fit linear regression models with interaction effects. Do any interactions appear to be statistically significant? **[10 points]**

**Answer:**