# MAT 303 Module One Problem Set Report

Multiple Regression

Luisa Heisler

luisa.heisler@snhu.edu

Southern New Hampshire University

## Introduction

The statistical analyses that are being performed here is to determine the relationship between a vehicles weight and horsepower and the effect it has on the fuel efficiency of the vehicle. The data set being used is a set of 32 different vehicles obtain from the car marker to study the fuel economy of vehicles. The results can be used to determine how the car marker could get a nice balance of weight and horsepower while still offering great fuel economy. Example they could shave some weight and add a little more horsepower and still have the same fuel economy or vice versa lower the horsepower and add some weight, or even lower both and offer a more fuel-efficient vehicle. The type of analyses that we will be running for this problem set are, to create a multiple regression model for fuel economy vs weight and fuel economy vs horsepower. So, then we can find the fitted values and residuals and plot them against one and generate a q-q plot to test assumptions of normality of the residuals and find the confidence interval. We will use all this information to then determine if the weight and horsepower of a vehicle have a statistically significant effect on fuel economy.

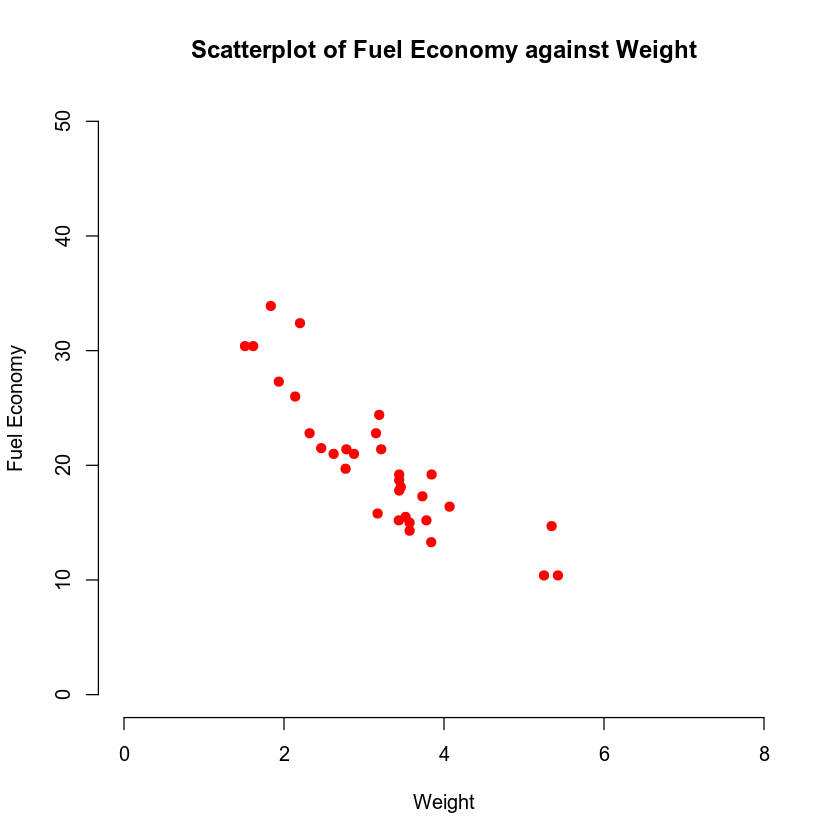
## Data Preparation

The important variables in this data set when comparing weight and horsepower to fuel economy are weight (wt in the code), horsepower (hp in the code), quarter mile time (qsec in the code), and fuel economy in miles (US) gallon (mpg in the code). The focus will be on the weight, horsepower, and fuel economy in miles for this problem set. There are 12 columns presents in the data set one column for each variable and 32 rows for each vehicle.

## Multiple Regression Model

### Correlation Analysis

The scatterplot below shows the relationships of fuel economy against weight. This scatterplot shows as the weight increases, the fuel economy decreases and that is a negative correlation. Based off the above scatterplot and the Pearson’s product-moment correlation of -0.8677, this indicates that there is a strong negative correlation.

****

### Reporting Results

**E(Y)=β0+ β1 x1+ β2 x2**

* Above is the general form of the multiple regression model. In this model, Y is the fuel efficiency (mpg), x1 is the weight, x2 is the horsepower. The equation will look like this;

**E(Y) = 19.7462 – 5.0480x1 + 0.9292x2**

For this equation R2 is 0.8264 and the adjusted R2 is 0.8144. Because R2 is 0.8264, that means that is fits the model regression equation that presented and adjusted R2 being 0.8144 means that the addition variable input was adding value to the multiple regression model. The beta estimate shows for every 5.0480 (this is in 1,000 lbs. \*5.0480 = 5,048lbs) of additional weight for the horsepower shows that for every 0.9292 of horsepower, it will decrease the fuel efficiency(mpg). A fitted value is the predicted value of the dependent variable (mpg) for data points from the data set. The residual is the difference between the actual value and the predicted value of the dependent variable(mpg). Based on the plots the normality of the residuals and the fitted plot value are there will be varied predicted values as the plots show very little correlation.

### Evaluating Model Significance

Null Hypothesis (H0)

Alternative Hypothesis (H1)

The F statistics = 69.03

The P-value = 9.395e-12

The P-value is 9.395e-12 is less than 0 and that means it has less than the significant level of 0.05, and that means the Null hypothesis is rejected and concludes that the multiple regression model is significant. If testing the Null hypothesis with the weight being significant at 5%,

H 0 : β 1 = 0 or

### H 1 : β 1 ≠ 0

The t-value is -10.430 and the P-value = 2.52e-11 and again it is less than 0. Again, since the P-value is less than the 5% level of significance, the Null hypothesis is rejected, and the weight is significant.

If testing the Null and Alternative hypothesis with the horsepower being significant at 5%,

H 0 : β 1 = 0 or

H 1 : β 1 ≠ 0

The t-value = 3.760 and the P-value = 0.000765 with again the value being less than 0. Again, since the P-value is less than the 5% level of significance, the Null hypothesis is rejected and the horsepower is significant.

### Making Predictions Using the Model

*Make predictions using the regression model. Address the following questions in your analysis:*

* *What is the predicted fuel efficiency (miles per gallon) for a car that has a rear axle ratio of 3.15 and a horsepower of 120? Suppose that this car achieves an average of 20.5 miles per gallon. What is the residual for this observation?*
* *What is the 95% prediction interval for the car identified in the previous question? Interpret this interval.*
* *What is the 95% confidence interval for the car identified in the previous question? Interpret this interval.*
* *Why is the prediction interval wider than the confidence interval?*

I had some trouble with the equations on the above question. I would love some help figuring it out and seeing what I missed when I was looking for the information in my Codio assignment.

## Conclusion

This model shows a higher accuracy where it represents how weight and horsepower changes the fuel economy of the vehicles. We do have the assumptions that the sample size is sufficiently large. We can conclude with the statement that higher the weight and the horsepower; fuel economy decreases and vice versa. There is strong correlation between weight, horsepower, and fuel economy. This model does have higher importance. This model will give the car manufacturer the ability to forecast and make cars as per the consumer needs. The company can furthermore invest in high power cars and a high fuel economy car for consumers as per the specific which will yield in optimized car and increase in revenue.