# MAT 303 Module One Problem Set Report

Multiple Regression

Yaqub Mohamud

yaqub.mohamud@snhu.edu

Southern New Hampshire University

## Introduction

For the module one problem set, I'm examining a data collection to develop fuel efficiency prediction models. The primary focus of my research is the relationship among rear axle ratio, horsepower, and fuel efficiency (mpg). With the models and calculations we'll be looking at, we will be able to determine if a car's horsepower and rear axle ratio have an impact on its MPG. With the help of factors like horsepower and rear axle ratio, we will mostly utilize this data to predict a vehicle's mileage and focus on its MPG. The results of this study will be used to help automakers choose the best horsepower and rear axle ratio to maximize a vehicle's fuel efficiency. This study will be conducted using a multiple regression model that will enable us to determine and illustrate any relationships between MPG, horsepower, and rear axle ratio. The correlation coefficients between our independent and dependent variables will also be examined to support the analytical procedure. Then, using all the data, we will determine if horsepower and rear axle ratio statistically affect fuel consumption as expressed in miles per gallon.

## Data Preparation

There are several significant dependent and independent variables in the data set that we will be using. The data collection that we will be using contains several important independent and dependent variables. Our primary dependent variable is miles per gallon, and our independent factors of relevance are horsepower and rear axle ratio. The main variables we will be considering in this study are horsepower (hp), rear axle ratio (drat), and miles per gallon (mpg). Each variable in our dataset includes 12 rows total, with 32 rows indicating the kind of car under analysis.

A screenshot of a computer

Description automatically generated

## Multiple Regression Model

### Correlation Analysis

A graph with red dots

Description automatically generated

The graph above shows that the rear axle ratio and fuel economy are positively correlated. This suggests that when the rear axle ratio increases, cars often have higher fuel efficiency. However, the dispersed nature of the data sets suggests that while there is a general upward trend, the correlation is weak. This disparity implies that other factors may also affect fuel economy and that the rear axle ratio is not the sole factor responsible for variations in mpg.

A graph with red dots

Description automatically generated

The fuel economy and horsepower have a negative association, as the graph above illustrates. Essentially, this indicates that a car's efficiency drops as its horsepower grows. The data points' decreasing slope serves as a visual representation of this tendency. Higher horsepower cars that are typically maintained need more fuel to generate more power, which lowers fuel economy.

The correlation coefficient between rear axle ratio and mpg was 0.6812, indicating a significant positive correlation between the two variables. The results showed that horsepower and mpg had a high, negative correlation, with the correlation coefficient between the two being -0.7762.

| A matrix: 3 × 3 of type dbl | | | |
| --- | --- | --- | --- |
|  | **mpg** | **drat** | **hp** |
| **mpg** | 1.0000 | 0.6812 | -0.7762 |
| **drat** | 0.6812 | 1.0000 | -0.4488 |
| **hp** | -0.7762 | -0.4488 | 1.0000 |

### Reporting Results

The multiple regression model's general form and prediction equation for the response variable, fuel efficiency (mpg), and the predictors, rear axle ratio (drat) and horsepower (hp).

MPG = β0 + β1drat + β2hp + e

β1 and β2 are the regression coefficients for drat and hp, respectively, and B0 is the interceptor. The timing error is e. From the output, the prediction model equation is:

MPG = 10.7898 + 4.6981(drat) – 0.0517(hp)

We obtained an R-squared value of 0.7412 and an Adjusted R-squared value of 0.7233 using this model. With an R-squared score of 0.7412, an automobile's horsepower and rear axle ratio account for 74.12% of its total mpg performance. With an Adjusted R-squared value of 0.7233, the car's horsepower and rear axle ratio might be responsible for approximately 72.33% of its total fuel efficiency.

Residual standard error: 3.17 on 29 degrees of freedom

Multiple R-squared: 0.7412, Adjusted R-squared: 0.7233

F-statistic: 41.52 on 2 and 29 DF, p-value: 3.081e-09

The rear axle ratio's beta estimate was determined to be 4.6891, meaning that for every unit increase in the rear axle ratio, the vehicle's fuel efficiency rises by 4.6891 units with a given amount of horsepower of 3. However, while the rear axle ratio remains constant, the mpg drops by 0.0517 units for every unit increase in horsepower.

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 10.789861 5.077752 2.125 0.042238 \*

drat 4.698158 1.191633 3.943 0.000467 \*\*\*

hp -0.051787 0.009293 -5.573 5.17e-06 \*\*\*

A fitted value is the anticipated response that is obtained from regression model predictions. A residual is the discrepancy between the actual observed value and the value that was predicted using the same set of predictor variables as in the regression model.

*A graph with red dots

Description automatically generated*

**A graph of a normal q-q plot

Description automatically generated**

Heteroscedasticity is not evident since the residual’s graphic shows a random distribution around zero.

This random distribution around zero suggests that there is no systematic shift in variability in the mistakes and that the model is probably well-specified. Thus, it makes sense to assume homoscedasticity, or constant variance of residuals.

### Evaluating Model Significance

At a 5% significance level, we are examining the model to see if our predictors and the result variable have any meaningful relationships. In general, the null hypothesis assumes that there is no link or effect between the variables. According to the alternative hypothesis, the variables are related to one another. We may ascertain the significance of each predictor variable in the model by utilizing t-tests for the various beta coefficients.

For the rear axle ratio at 0.00467, the p-value is below the usual significance level of 0.05. We reject the null hypothesis if the test's p-value is less than 0.05. This indicates that there is enough data to draw the conclusion that the rear axle ratio significantly predicts the dependent variable. This implies that fuel economy is impacted by the rear axle ratio.

The p-value for horsepower is extremely low, 5.17e-06, which is less than 0.05. The p-value being extremely low provides strong evidence against the null hypothesis. This indicates that a highly significant predictor of fuel economy is horsepower. Accordingly, the correlation between horsepower and fuel economy is probably true and not the result of chance.

Our interval, when constructed with 95% confidence, would be (2.2610, 7.1353) for the rear axle ratio and (-0.0708, -0.0328) for horsepower. A coefficient of 7.1353 suggests a positive link between the rear axle ratio and the fuel economy. A coefficient of -0.0328 indicates a slight negative association between horsepower and the fuel economy. This is a good finding, and the intervals offer a good basis for drawing conclusions about how these factors relate to the result.

| A matrix: 3 × 2 of type dbl | | |
| --- | --- | --- |
|  | **2.5 %** | **97.5 %** |
| **(Intercept)** | 0.4047 | 21.1750 |
| **drat** | 2.2610 | 7.1353 |
| **hp** | -0.0708 | -0.0328 |

### Making Predictions Using the Model

For a car with a rear axle ratio of 3.15 and 120 horsepower, the estimated fuel efficiency is 19.37466 miles per gallon. Likewise, a car with an average mileage of 20.58 mpg has a residual of 1.12534. (12.64486, 26.10446) is the 95% prediction limit. A car with a rear axle ratio of 3.15 and 120 horsepower will, therefore, have a 95% chance of having an actual fuel economy that falls within this range. 17.57162, 21.1777 is the 95% confidence interval. This indicates that we have a 95% confidence level that the genuine mean fuel economy for vehicles with a horsepower of 120 and a rear axle ratio of 3.15 will fall within this range.

As we can see, our confidence interval is less than the predicted interval. The prediction interval will actually always have a wider range than the corresponding confidence interval since it accounts for both the variability of the response variable and the variability in parameter estimations.

|  |  |  |
| --- | --- | --- |
| *19.3747* | *12.6449* | *26.1045* |

|  |  |  |
| --- | --- | --- |
| 19.3747 | 17.5716 | 21.1777 |

## Conclusion

In conclusion, my recommendation to use the multiple regression model based on rear axle ratio (drat) and horsepower (hp) for predicting fuel efficiency (mpg) is well-supported by my analysis and the statistical tests I've conducted. I demonstrated that the model provides accurate predictions for fuel economy based on the variables drat and hp. Both the intuitive and statistically significant relationships between rear axle ratio and fuel efficiency (higher axle ratios typically correlate with better fuel efficiency) and horsepower and fuel efficiency (more horsepower slightly reduces fuel economy) are present. These relationships have been confirmed through statistical testing (such as p-values, confidence intervals, and the F-test). The F-test confirms that at least one of the predictors (drat or hp) is significantly related to fuel efficiency, which helps justify the use of the model.

Based on horsepower and rear axle ratio, my multiple regression model has shown itself to be a useful tool for forecasting fuel economy. This model offers practical insights that are important to both producers and customers due to its predicted accuracy and the statistical importance of the variables. My analysis's findings could assist buyers in making better judgments when they buy cars by allowing them to consider how horsepower and rear axle ratio impact fuel efficiency. This model might be used by manufacturers to create cars with the best possible fuel economy, assisting in the achievement of financial and environmental objectives. Automobile manufacturers might potentially utilize the model to inform design choices, guaranteeing that the most fuel-efficient vehicles are released into the market.