Automatic vs. Manual Transmission Cars

Alex

Executive Summary

This project investigates the `mtcars` dataset which comprises fuel consumption as well as 10 aspects of automobile design and performance for 32 automobiles. The objective is to explore the relationship between these variables and miles per gallon (MPG) to determine whether manual or automatic transmission provides better mileage and quantify the difference in MPG.

Exploratory Data Analysis

The effect of transmission type on MPG was first analysed through a boxplot (Appendix A: Figure 1). This boxplot illustrates that cars with manual transmissions have a higher MPG.

A Welch Two Sample t-test also indicates that the mean MPG of cars with manual transmission (24.4mpg) is higher than that of cars with automatic transmission (17.1mpg).

Next, the effect of the 10 variables, `cyl`, `disp`, `hp`, `drat`, `wt`, `qsec`, `vs`, `am`, `gear` and `carb` on MPG were explored through a pairwise scatterplot (Appendix A: Figure 2). This pairwise scatterplot shows that the variables `cyl`, `disp`, `hp`, `drat`, `wt`, `vs` and `am` have a strong correlation with MPG.

Regression Modelling

Model 1

The first model assumes that all variables have an impact on MPG.

```
model1 <- lm(mpg ~ ., data = carsDataset)
summary(model1)
```

The adjusted R-squared value of this model is 0.7790215, indicating that the model only explains approximately 78% of the variability of the data around its mean.

Model 2

The first model was refined by applying the stepwise regression search algorithm. This was accomplished through the use of the R function `step()` to perform variable selection. As such, only useful variables that impacted MPG significantly were selected as part of the model solutioning.

```
model2 <- step(model1, direction = "both")
summary(model2)
```

The refined model narrows down the important variables to `cyl`, `hp`, `wt` and `am` as having the greatest impact on MPG. This resulted in the model achieving an adjusted R-squared value of 0.8400875 (84%).

An ANOVA of Model 2 further shows that `cyl` and `wt` are the two most important variables that have the greatest significance on MPG.

```
anova(model2)
```

```
## Analysis of Variance Table
##
## Response: mpg
##           Df Sum Sq Mean Sq F value    Pr(>F)
## cyl        2 824.78  412.39 70.9959 2.924e-11 ***
## hp         1  24.09   24.09  4.1481 0.0519844 .
## wt         1 116.39  116.39 20.0373 0.0001339 ***
## am         1   9.75    9.75  1.6789 0.2064597
## Residuals 26 151.03    5.81
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Coefficient Interpretation

```
## (Intercept)        cyl6        cyl8          hp          wt    amManual
## 33.70832390 -3.03134449 -2.16367532 -0.03210943 -2.49682942  1.80921138
```

The coefficients of Model 2 tell us that:

- An increase in the number of cylinders from 4 to 6 (cyl6) results in the car's mileage decreasing by approximately 3.0mpg.

- An increase in the number of cylinders from 4 to 8 (cyl8) results in the car's mileage decreasing by approximately 2.2mpg.

- For every increase in gross horsepower (hp), the car's mileage decreases by approximately 0.03mpg.

- For every 1000lbs increase in weight (wt), the car's mileage decreases by approximately 2.5mpg.

- A car with manual transmission (amManual) achieves better mileage as its mileage increases by approximately 1.8mpg as compared to a car with automatic transmission.

Residuals and Diagnostics

The residuals was examined through a residuals plot (Appendix A: Figure 3) and no anormalies were detected.

Diagnostics plots in Appendix A: Figure 4 reinforce these findings. The Residuals vs. Fitted diagnostic plot on the top-left illustrates randomly scattered points which verifies the independence condition, and the Normal Q-Q plot on the top-right shows that the residuals are normally distributed.

Conclusion

Through the regression analysis performed, it can be ascertained that manual transmission provides better mileage. A car with manual transmission experiences an increase of approximately 1.8mpg as compared to a car with automatic transmission.

Environment

All analysis was performed in an R Markdown (Rmd) document using RStudio 0.98.1091. It can be downloaded from this link.
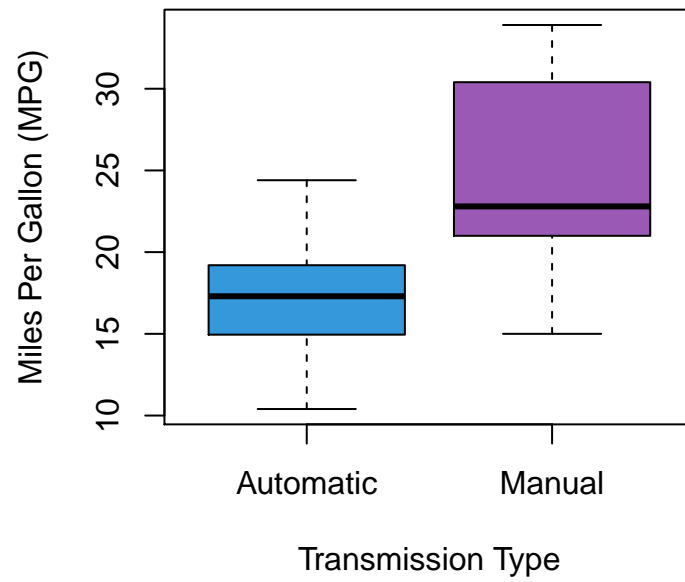
Appendix A: Plots

Figure 1: Exploratory Data Analysis Boxplot
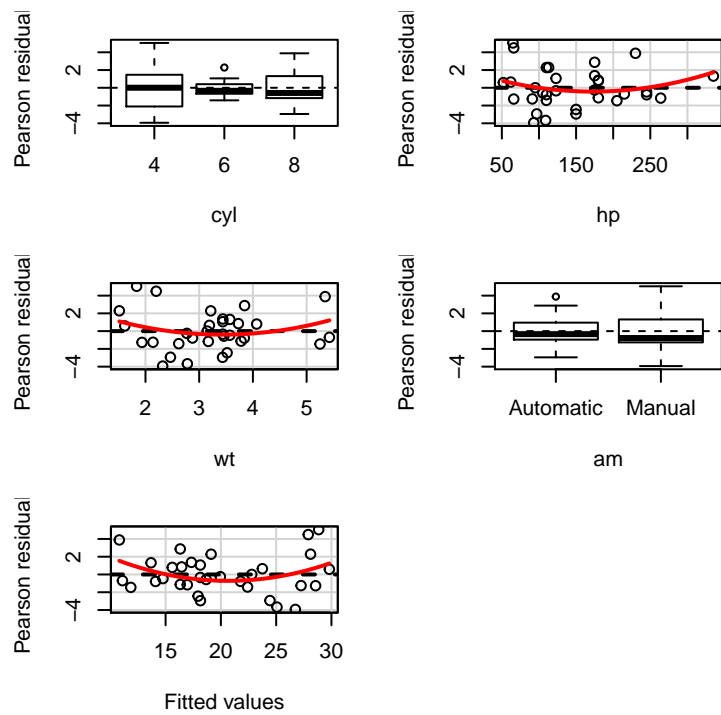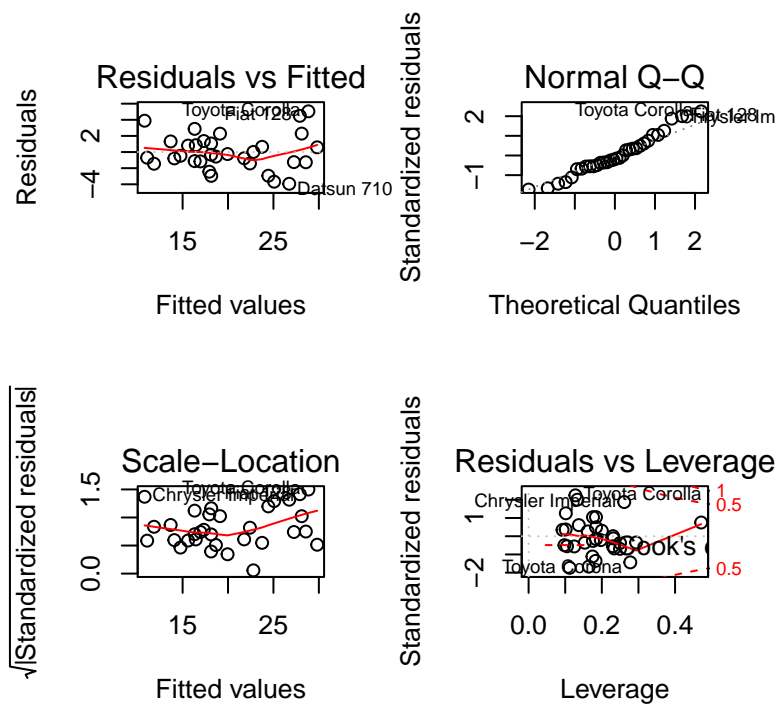
Figure 2: Exploratory Data Analysis Pairwise Scatterplot

Figure 3: Residual Plots

Figure 4: Diagnostic Plots