Motor Trend Final

Harsh Reddy

November 15, 2017

Created with knitr

Summary

This report analyzed the correlation between the transmission type and miles per gallon (MPG). The report set out to determine which transmission type produces a higher MPG. We used mtcars dataset for this analysis. A t-test between automatic and manual transmission vehiclesshows that manual transmission vehicles have a approximately 7.3 greater MPG than automatic transmission vehicles.Once fitted with multiple linear regressions, analysis showed that the manual transmission contributed less towards the MPG, only an improvement of nearly 1.8 MPG. Other weight, variables, horsepower, and number of cylinders contributed more towards the overall MPG of vehicles.

Data

Load the dataset and convert categorical variables to factors.

```
library(ggplot2)
data(mtcars)
head(mtcars, n=3)
dim(mtcars)
mtcars$cyl <- as.factor(mtcars$cyl)
mtcars$vs <- as.factor(mtcars$vs)
mtcars$am <- factor(mtcars$am)
mtcars$gear <- factor(mtcars$gear)
mtcars$carb <- factor(mtcars$carb)
attach(mtcars)
```

Exploratory Analysis

See Appendix Fig I Exploratory Box graph that compares Automatic and Manual transmission MPG.The graph tells us that there is a huge increase in MPG for vehicles with a manualtransmission vs automatic.

Statistical Inference

T-Test transmission type and MPG

```
testResults <- t.test(mpg ~ am)
testResults$p.value
```

```
## [1] 0.001373638
```

The T-Test rejects the null hypothesis that the difference between transmission types is 0.

```
testResults$estimate
```

```
## mean in group 0 mean in group 1
##        17.14737        24.39231
```

The difference estimate between the 2 transmissions is 7.24494 MPG in favor of manual.

Regression Analysis

Fit the full model of the data

```
fullModelFit <- lm(mpg ~ ., data = mtcars)
summary(fullModelFit)  # results hidden
summary(fullModelFit)$coeff  # results hidden
```

As the above coefficients have a p-value less than 0.05 we cannot conclude which variables are more statistically significant.

Backward selection to determine which variables are most significant

```
stepFit <- step(fullModelFit)
summary(stepFit) # results hidden
summary(stepFit)$coeff # results hidden
```

The new model has 4 variables (number of cylinders, weight, horsepower, transmission). The Rsquared value of 0.866 confirms that this model provides information regarding the 87% of the variance in MPG. The p-values are also statistically significantly because they have a p-value less than 0.05. The coefficients conclude that raising the number of cylinders from 4 to 6 with decrease the MPG by 3.03. Further increasing thecylinders to 8 with reduce the MPG by 2.16. Increasing the horsepower is reduces MPG 3.21 for every 100 horsepower. Weight decreases the MPG by 2.5 for each 1000 lbs increase. A Manualtransmission improves the MPG by 1.81.

Residuals & Diagnostics

Residual Plot **See Appendix Figure II**

The plots conclude:

1. The randomness of the Residuals vs. Fitted plot supports the assumption of independence
2. The points of the Normal Q-Q plot following closely to the line conclude that the distribution of residuals is normal
3. The Scale-Location plot random distribution confirms the constant variance assumption
4. Since all points are within the 0.05 lines, the Residuals vs. Leverage concludes that there are no outliers

```
sum((abs(dfbetas(stepFit)))>1)
```

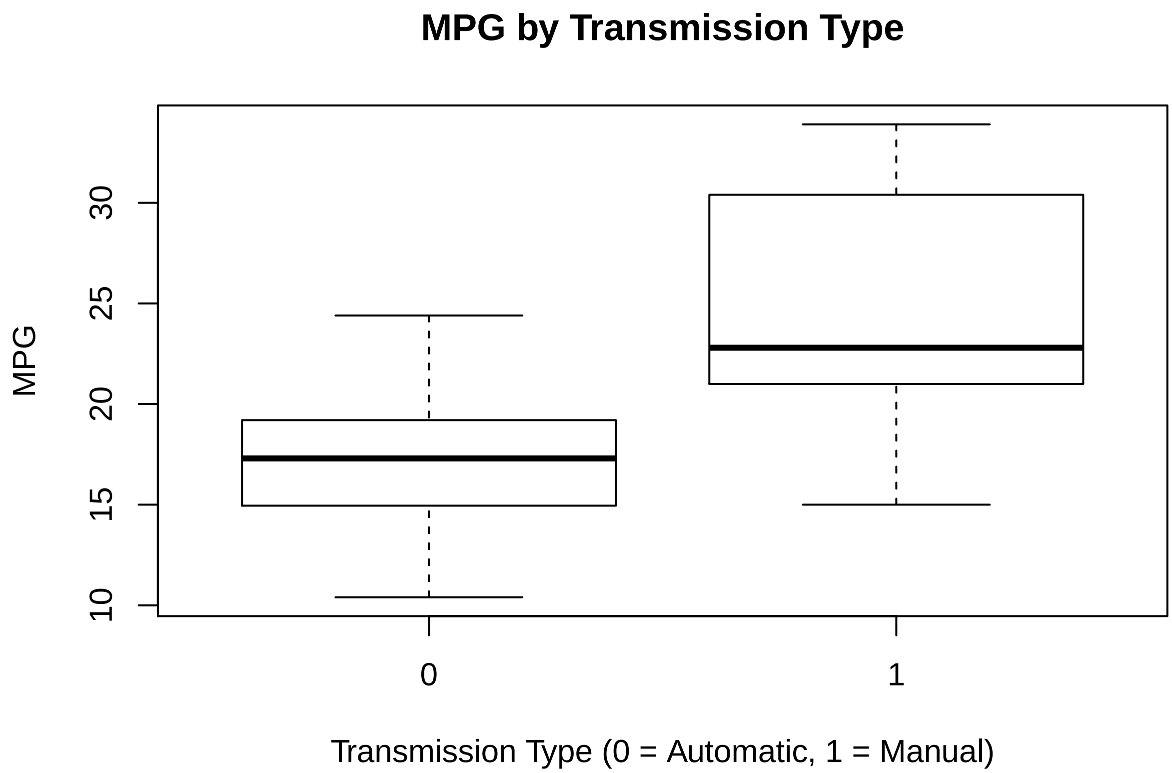
```
## [1] 0
```

The Conclusion

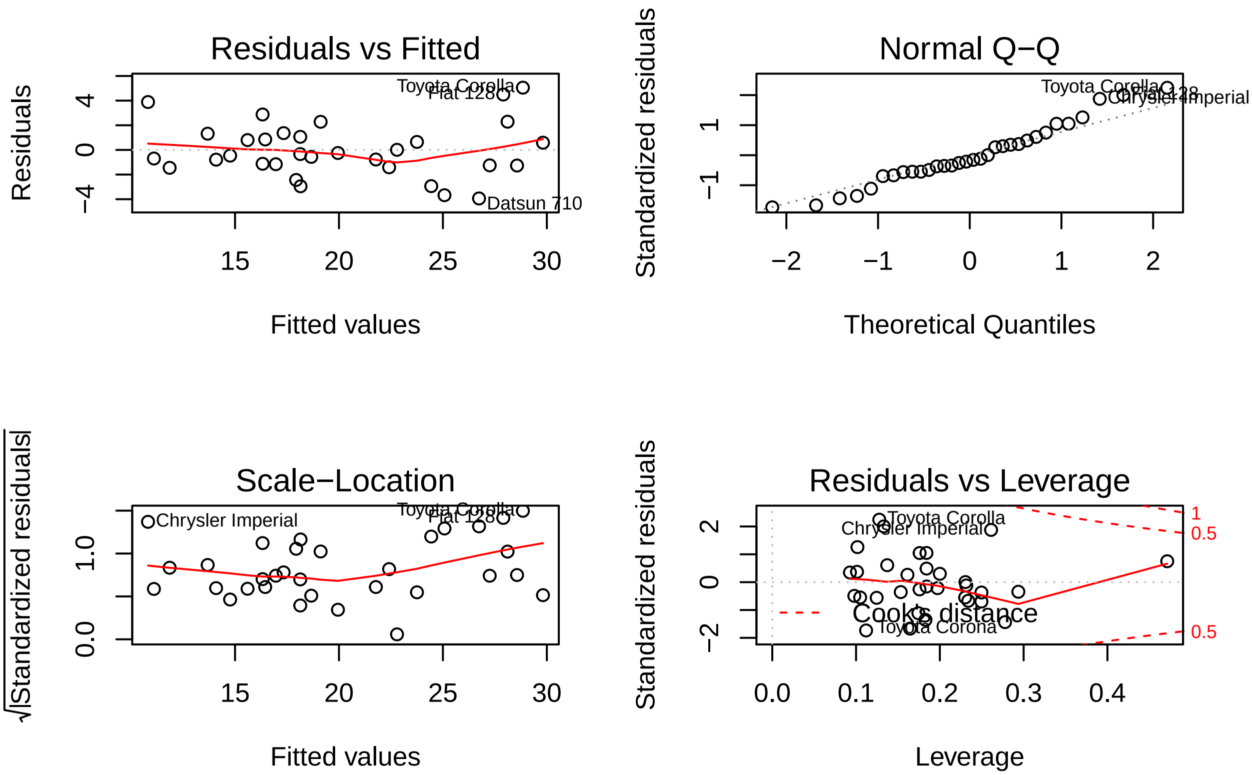
There is a variation in MPG due to the transmission type. A manual transmission will result in a minute MPGboost. But it does seem that weight, horsepower, & number of cylinders are more statistically significantwhile determining the MPG.

Appendix Fig

I


MPG by Transmission Type
MPG
10
15
20
25
30
0
1
Transmission Type (0 = Automatic, 1 = Manual)

II


Residuals vs Fitted
Residuals
Fitted values
Toyota Corolla
Fiat 128
Datsun 710
Normal Q−Q
Standardized residuals
Theoretical Quantiles
Chrysler Imperial
Scale−Location
√|Standardized residuals|
Residuals vs Leverage
Leverage
Cook's distance
Toyota Corona
1
0.5