Motor Trend MPG Data Analysis

Jody9678

2023-09-26

Created with knitr

Executive Summary This report analyzes the relationship between transmission type (manual or automatic) and miles per gallon (MPG) as a means to answer two questions: Which transmission type is better for MPG? What is the MPG difference between automatic and manual transmissions?

The `mtcars` dataset was used for this analysis. A simple t-test between automatic and manual transmission vehicles indicates that vehicles with manual transmissions are far more efficient, achieving 7.24 more miles per gallon than their automatic counterparts. A deeper dive through the application of multiple linear regressions provided a more meaningful view of the data and a more concise conclusion; in that the transmission type contributed less significantly to actual fuel efficiency (manual transmissions over automatic by a mere 1.81 MPG) with factors of weight, horsepower, and number of cylinders contributing a great deal more to the overall fuel efficiency of vehicles. Manual transmission is the overall winner on efficiency, but size and power change the picture significantly.

Load 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 Figure I** 'Exploratory Box graph comparing Automatic and Manual transmission MPG.'

The graph leads us to believe that there is a significant improvement in MPG for vehicles with a manual transmission 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 transmissions.

Regression Analysis Fit the full model of the data

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

Since none of the 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 statistically significant

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

The new model has 4 variables (cylinders, horsepower, weight, transmission). The R-squared value of 0.8659 confirms that this model explains almost 87% of the variance in MPG. The p-values also are statistically significantly because they have a p-value less than 0.05. The coefficients conclude that increasing the number of cylinders from 4 to 6 with decrease the MPG by 3.03, with an additional drop 0f 2.16 MPG with 8 cylinder vehicles. Efficiency drops by 3.21 MPG for every 100 horsepower added. Every 1000lb increase costs 2.5 MPG. Taking all of this into account, manual transmissions offer a marginal 1.81 MPG improvement over automatic transmissions.

Residuals & Diagnostics **See Appendix Figure II** 'Residual Plots including all available variables.'

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
```

Conclusion There is a difference in MPG based on transmission type, with manual transmissions offering a slight improvement in fuel efficiency. However, it seems that weight, horsepower, & number of cylinders are more statistically significant when determining MPG.

Appendix Figures

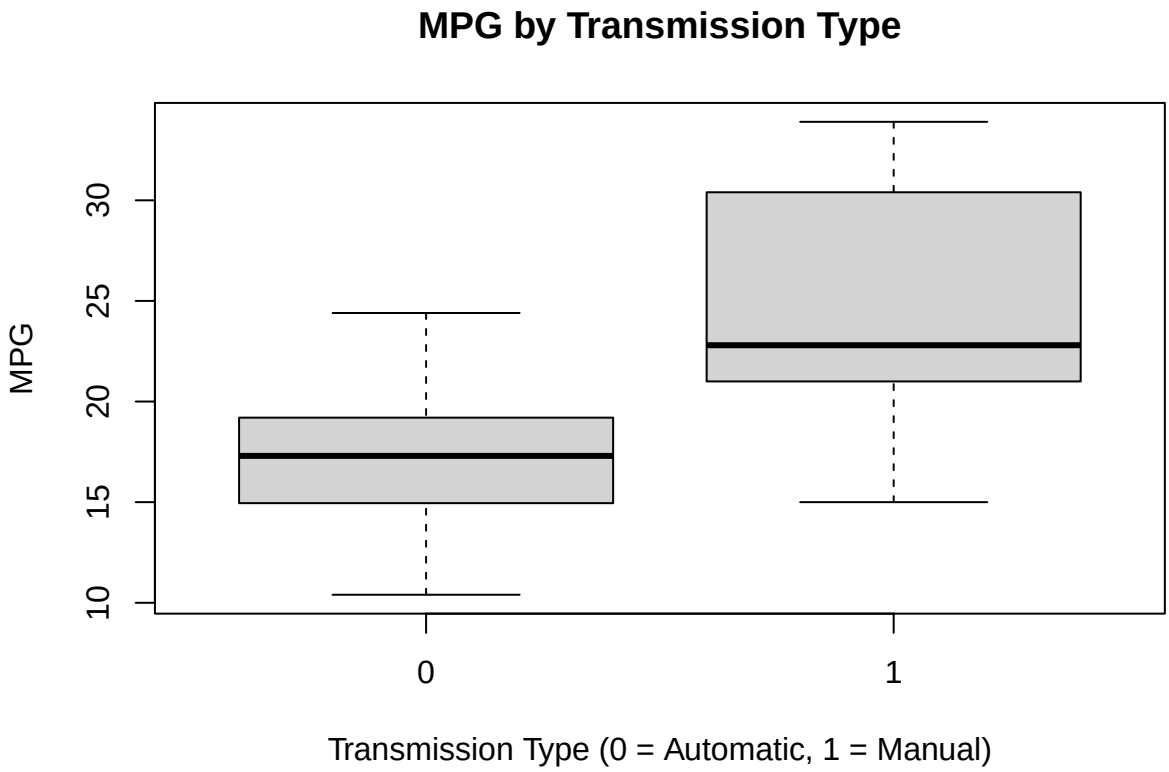

Figure I
'Exploratory Box graph comparing Automatic and Manual transmission MPG.'

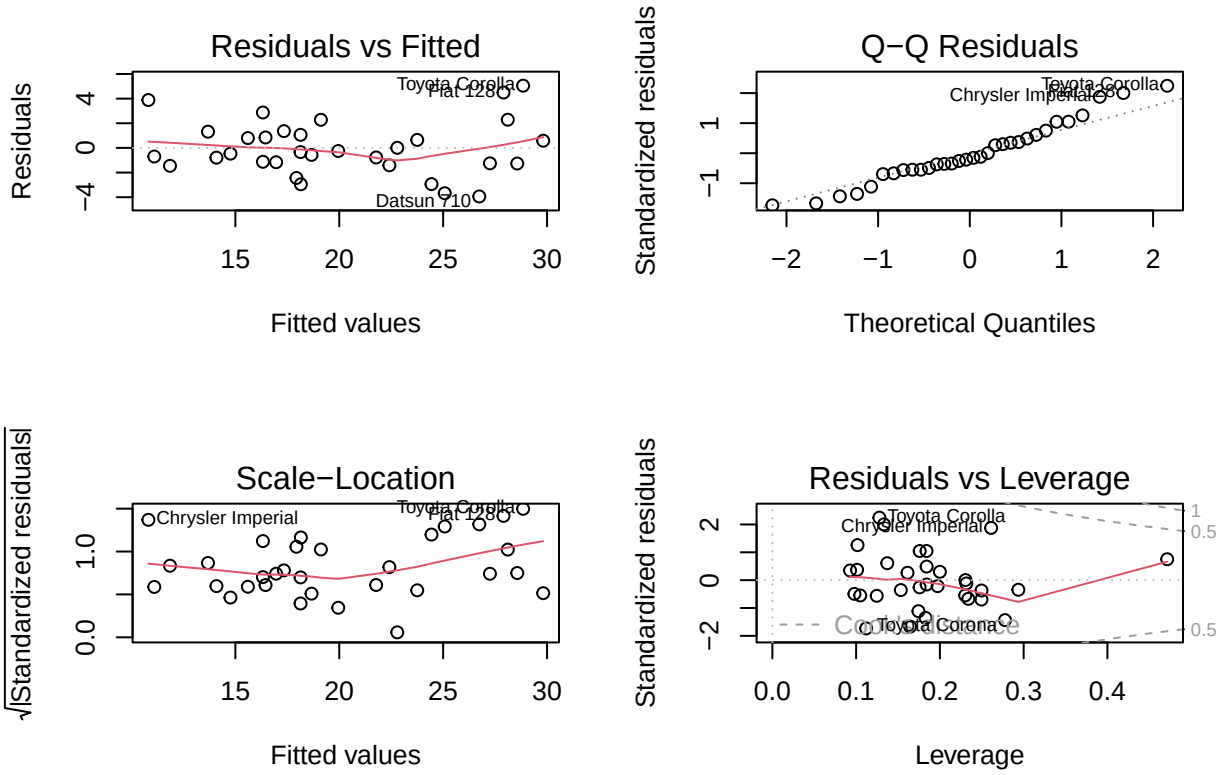

Figure II
'Residual Plots including all available variables.'