Regression Models - Transmission and MPG

Ryan Wissman

Friday, March 20, 2015

Executive Summary

Fuel efficienty and transmission type are both very important and greatly debated factors when selecting a new car. This report examines the relationship between tranmission type and fuel economy to determine if there is any MPG benefit to purchasing a car of either type transmission. The data used in this report is from the 1974 Motor Trend US magazine.

Exploring the Data

First the mtcars data is loaded and some brief exploratory statistics are discovered.

```
data(mtcars); attach(mtcars)
head(mtcars,1) #Examine how the data is structured in mtcars
```

```
##           mpg cyl disp  hp drat   wt  qsec vs am gear carb
## Mazda RX4  21   6  160 110  3.9 2.62 16.46  0  1    4    4
```

```
mean_a <- mean(mpg[am=="0"]) #Automatic Transmission mean
mean_m <- mean(mpg[am=="1"]) #Manual Transmission mean
```

Quickly examining the data to determine the means we find that the average fuel economy among automatic cars is 17.14 MPG whereas the average among manual transmissions is 24.39 MPG. Furthermore, according to a boxplot of the data (*see* **Figure 1**) we could guess that the fuel efficiency of a manual transmission is greater than that of an automatic transmission. The average and median is MPG for manual transmission is distinctly higher than that of automatic transmissions. However, we cannot yet make a conclusion on based on this chart alone. First we will need to determine if a relationship does exist by using regression.

Regression Models

First we try a linear model using mpg as the outcome and transmission type (variable am, "0" denotes automatic whereas "1" denotes manual) as the predictor.

```
model_am <- lm(mpg ~ am)
summary(model_am)
```

```
##
## Call:
## lm(formula = mpg ~ am)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.3923 -3.0923 -0.2974  3.2439  9.5077
##
```

```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   17.147      1.125  15.247 1.13e-15 ***
## am             7.245      1.764   4.106 0.000285 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.902 on 30 degrees of freedom
## Multiple R-squared:  0.3598, Adjusted R-squared:  0.3385
## F-statistic: 16.86 on 1 and 30 DF,  p-value: 0.000285
```

Using this model we can determine that this model, while significant, can only explain about 36% (Multiple R-squared value of 0.3598) of the variance in MPG. Therefore we shoudl try another model to examine the other variables in the mtcars dataset.

Accoding to the matrix scatterplot of all the variables in mtcars (*see* **Figure 2**) the there are a number of other variables that show significant correlation with MPG. Therefore exparimentation is necessary to create a model with the highest multiple R-squared value.

Appendix

Data and figures that accompany the report.

Figure 1: Boxplot summarizing both automatic and manual transmission types relative to MPG

```
boxplot(mpg ~ am, xlab="Transmission Type", ylab="Miles per Gallon (MPG)", xaxt="n",
        main="Transmission and MPG",
        col=c("green", "yellow"))
axis(1, at=1:2, labels=c("Automatic", "Manual"))
```

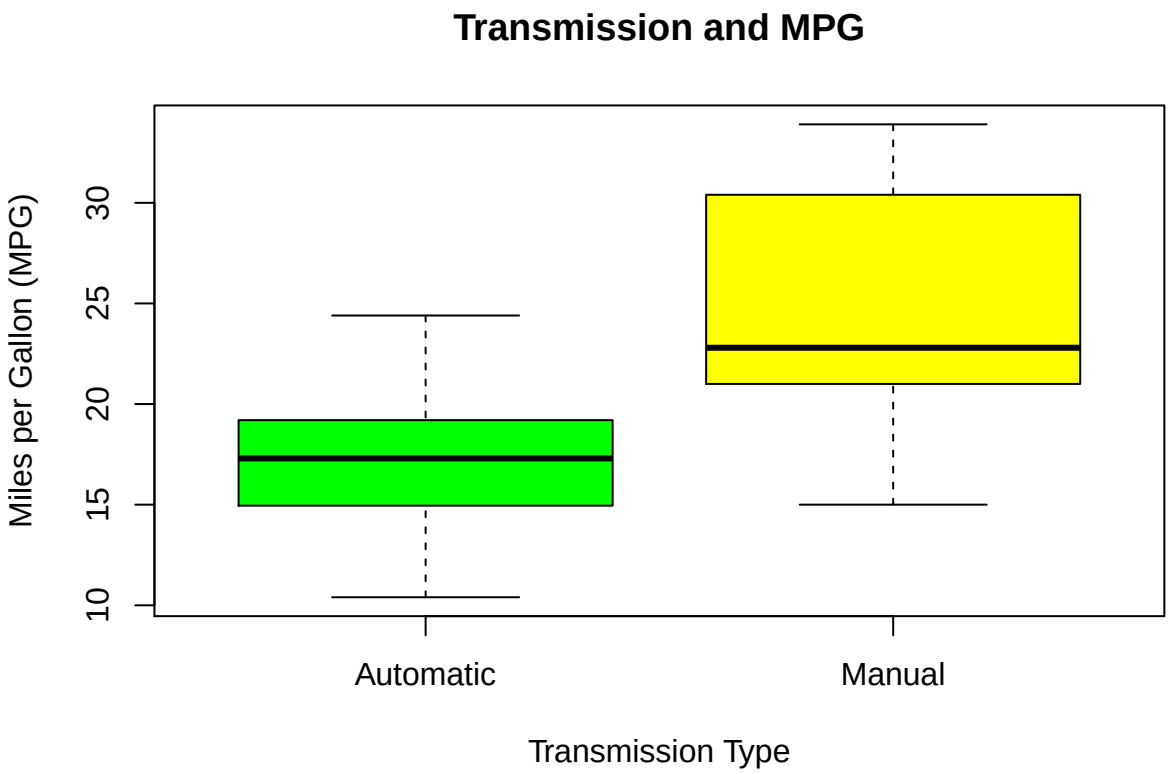

Figure 2:

Pairs chart covering the variances between many different variables in the mtcars dataset.