Regression Model Project

Michael Simpson

February 19, 2015

Executive Summary

Using the data from the mtcars data set, a linear model was used to try and answer two main questions. First which type of transmission was better for a cars MPG, and second what is the difference in MPG between the two transmission types. Looking at the plot in appendix A it becomes clear that a manual transmission is better for MPG. Then through a linear model, we can estimate that there is a 7.25 MPG difference on average between the cars with automatic transmissions and cars with manual transmissions. Further analysis of the data revealed no meaningful outliers which may have influenced the model. While the model shows a relationship between the type of transmission and the MPG, it is by no means the only factor affecting MPG. Through a regression analysis of all the covariates another factor, weight appears to have a significant relationship with MPG as well.

Questions

"Is an automatic or manual transmission better for MPG"

"How different is the MPG between automatic manual transmission?"

Exploratory Analysis

Looking at a simple Box plot in appendix A, it appears that the transmission type has a significant effect on a cars MPG, and that cars with manual transmissions have higher MPG than those with automatic transmissions. However further analysis is required to see if other factors are influencing a cars MPG

```
head(mtcars)
```

```
lm1<- lm(mpg ~ ., data=mtcars)
lm2<- lm(mpg ~ am, data=mtcars)
summary(lm1)$coefficients
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 12.30337   18.71788  0.6573  0.51812
## cyl         -0.11144    1.04502 -0.1066  0.91609
## disp         0.01334    0.01786  0.7468  0.46349
## hp          -0.02148    0.02177 -0.9868  0.33496
## drat         0.78711    1.63537  0.4813  0.63528
## wt          -3.71530    1.89441 -1.9612  0.06325
## qsec         0.82104    0.73084  1.1234  0.27394
## vs           0.31776    2.10451  0.1510  0.88142
## am           2.52023    2.05665  1.2254  0.23399
## gear         0.65541    1.49326  0.4389  0.66521
## carb        -0.19942    0.82875 -0.2406  0.81218
```

```
a<- mtcars$am==0
m<- mtcars$am==1
ma<-mean(mtcars$mpg[a])
mm<-mean(mtcars$mpg[m])
mm-ma
```

```
## [1] 7.245
```

lm1 represents a linear models that are trying to explain the influence of all the variables on the mpg variable. From the coefficient summary it appear that am (transmission type) and wt (vehicles weight) are the two variables with the greatest influence of mpg.

```
summary(lm2)$coefficients
```

```
##             Estimate Std. Error t value  Pr(>|t|)
## (Intercept)   17.147      1.125  15.247 1.134e-15
## am             7.245      1.764   4.106 2.850e-04
```

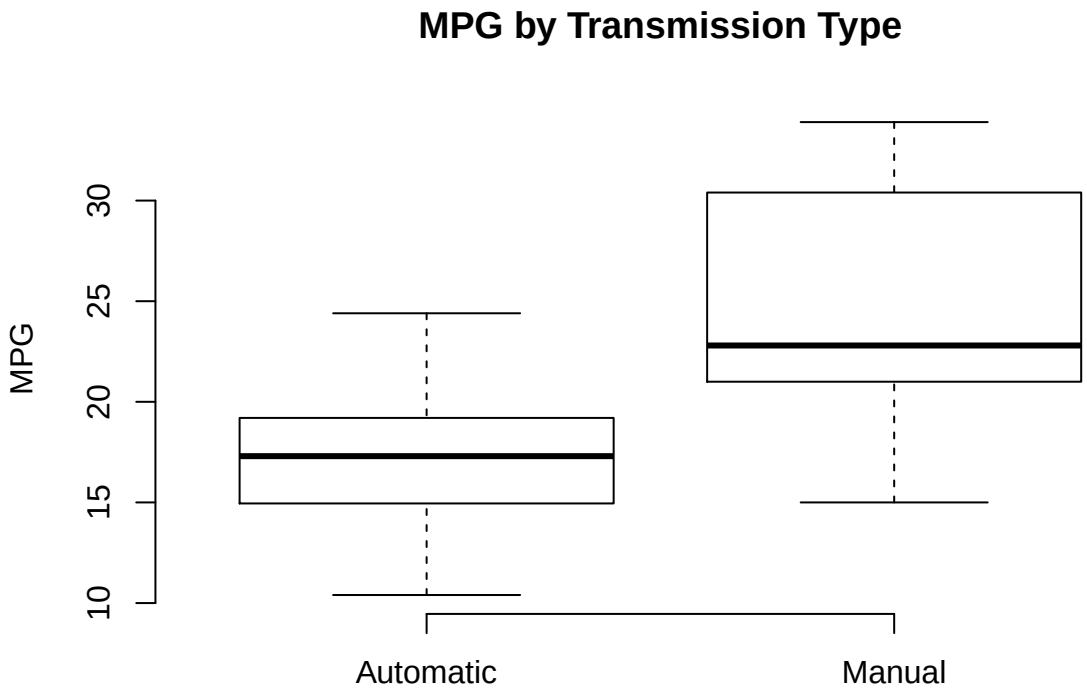
Looking at lm2 which measures the marginal influence of the transmission type on mpg the model estimates a 7.25 increase in mpg for cars with manual transmissions versus cars with automatic transmissions.

```
round(dfbeta(lm2), 2)
round(hatvalues(lm2), 3)
```

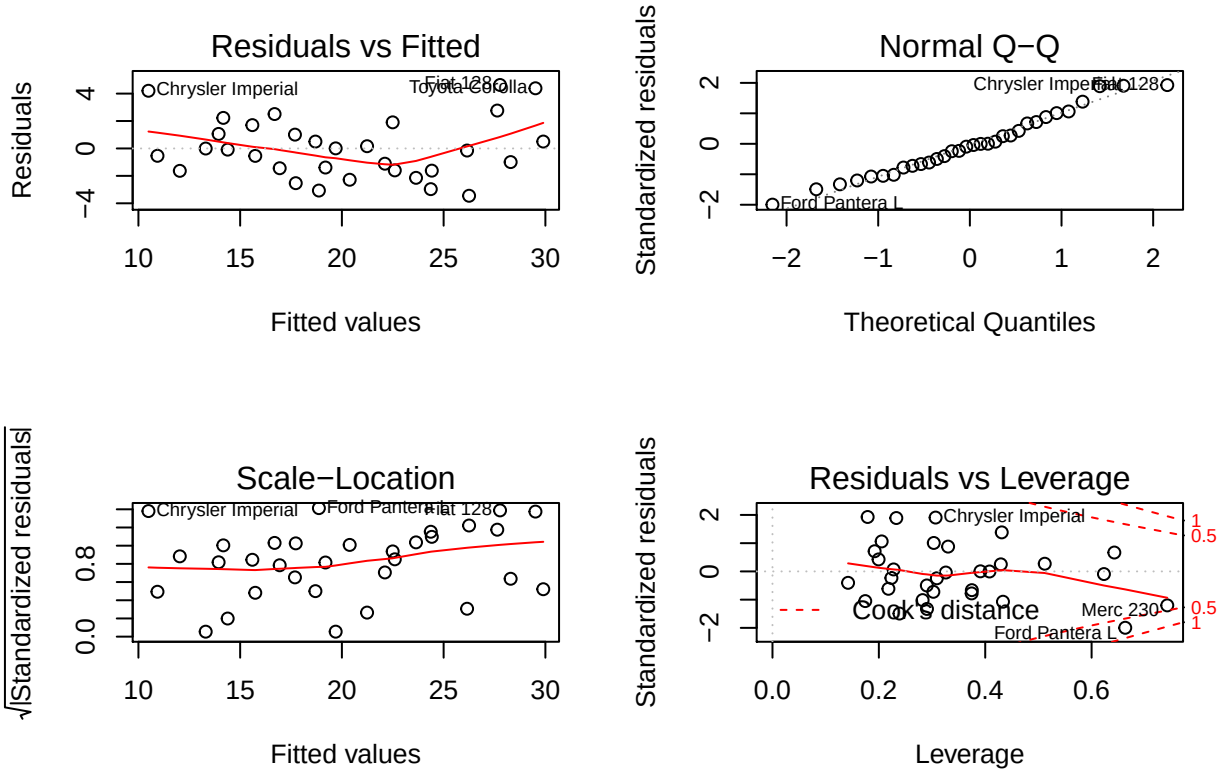
Looking at the dfbeta and hatvalue diagnostics, as well as the plots from appendix B, did not reveal any highly influential points from the data set indicating that the model is not being influenced by an outlier.

Appendix

A



B


Residuals vs Fitted
Residuals
Fitted values
Chrysler Imperial
Fiat 128
Toyota Corolla
Normal Q−Q
Standardized residuals
Theoretical Quantiles
Ford Pantera L
Scale−Location
√|Standardized residuals|
Residuals vs Leverage
Cook's distance
Leverage
Merc 230