Motor Trend Analysis

Gulab Singh

27 March 2017

Executive Summary

In this report we will analyse the dataset for a collection of manual and automatic transmission and study the relationship between a set of variables and fuel effeciency expressed in Miles per Gallon (MPG). The study is focussed in finding answers to following questions

- Is an automatic or manual transmission better for MPG
- Quantify the MPG difference between automatic and manual transmissions

Exploratory data analysis

```
library(datasets)
data("mtcars")
str(mtcars[mtcars$am=="0",])
```

```
## 'data.frame':    19 obs. of  11 variables:
##  $ mpg : num  21.4 18.7 18.1 14.3 24.4 22.8 19.2 17.8 16.4 17.3 ...
##  $ cyl : num  6 8 6 8 4 4 6 6 8 8 ...
##  $ disp: num  258 360 225 360 147 ...
##  $ hp  : num  110 175 105 245 62 95 123 123 180 180 ...
##  $ drat: num  3.08 3.15 2.76 3.21 3.69 3.92 3.92 3.92 3.07 3.07 ...
##  $ wt  : num  3.21 3.44 3.46 3.57 3.19 ...
##  $ qsec: num  19.4 17 20.2 15.8 20 ...
##  $ vs  : num  1 0 1 0 1 1 1 1 0 0 ...
##  $ am  : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ gear: num  3 3 3 3 4 4 4 4 3 3 ...
##  $ carb: num  1 2 1 4 2 2 4 4 3 3 ...
```

```
str(mtcars[mtcars$am=="1",])
```

```
## 'data.frame':    13 obs. of  11 variables:
##  $ mpg : num  21 21 22.8 32.4 30.4 33.9 27.3 26 30.4 15.8 ...
##  $ cyl : num  6 6 4 4 4 4 4 4 4 8 ...
##  $ disp: num  160 160 108 78.7 75.7 ...
##  $ hp  : num  110 110 93 66 52 65 66 91 113 264 ...
##  $ drat: num  3.9 3.9 3.85 4.08 4.93 4.22 4.08 4.43 3.77 4.22 ...
##  $ wt  : num  2.62 2.88 2.32 2.2 1.61 ...
##  $ qsec: num  16.5 17 18.6 19.5 18.5 ...
##  $ vs  : num  0 0 1 1 1 1 1 0 1 0 ...
##  $ am  : num  1 1 1 1 1 1 1 1 1 1 ...
##  $ gear: num  4 4 4 4 4 4 4 5 5 5 ...
##  $ carb: num  4 4 1 1 2 1 1 2 2 4 ...
```

There are a total of 32 observatiosn fo 11 variables, 19 observations are for automatic transmission (am=0) and 13 for mannual transmission (am=1), looking at the summary data, the mean of mpg for mannual cars is 7.24 more than for automatic cars (24.39-17.15). This is also illustrated in the boxplot included in Appendix

A, however we will need to carry out hypothesis testing and also look at the impact of variables other than transmission type on mpg.

Regression analysis

To bettwer understand the imapct of various variables, we will plot a relationship between all the variables (refer Appendix A).

Explortaory analysis suggested that the mpg for manual cars is higher as compared to automatic transmission. we will test this using t-test.

Null Hypothsis: There is no difference in MPG between Automatic and Manual transmission cars Alternate Hypothsis: There is difference in MPG between Automatic and Manual transmission cars

```
t.test(mtcars$mpg~mtcars$am, conf.level=0.95)
```

```
##
##  Welch Two Sample t-test
##
## data:  mtcars$mpg by mtcars$am
## t = -3.7671, df = 18.332, p-value = 0.001374
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.280194  -3.209684
## sample estimates:
## mean in group 0 mean in group 1
##        17.14737        24.39231
```

Since p-value is 0.001374, we ust reject the Null Hypothesis

```
linearfit <- lm(formula = mpg ~ am, data = mtcars)
summary(linearfit)
```

```
##
## Call:
## lm(formula = mpg ~ am, data = mtcars)
##
## 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
```

From above we can conclude that the fuel effecienncy for automatic cars is 17.147 and manual cars run at 17.147 + 7.245 = 24.392

since R squared value is 0.3598, this accounts for only 36% variance. We will now perform a multi variate linera regression analysis

We will let the model choose most significant paramaters.

```
automultivariate <- step(lm(data=mtcars, mpg ~ .), trace=0)
summary(automultivariate)
```

```
##
## Call:
## lm(formula = mpg ~ wt + qsec + am, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.4811 -1.5555 -0.7257  1.4110  4.6610
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   9.6178     6.9596   1.382 0.177915
## wt           -3.9165     0.7112  -5.507 6.95e-06 ***
## qsec          1.2259     0.2887   4.247 0.000216 ***
## am            2.9358     1.4109   2.081 0.046716 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.459 on 28 degrees of freedom
## Multiple R-squared:  0.8497, Adjusted R-squared:  0.8336
## F-statistic: 52.75 on 3 and 28 DF,  p-value: 1.21e-11
```

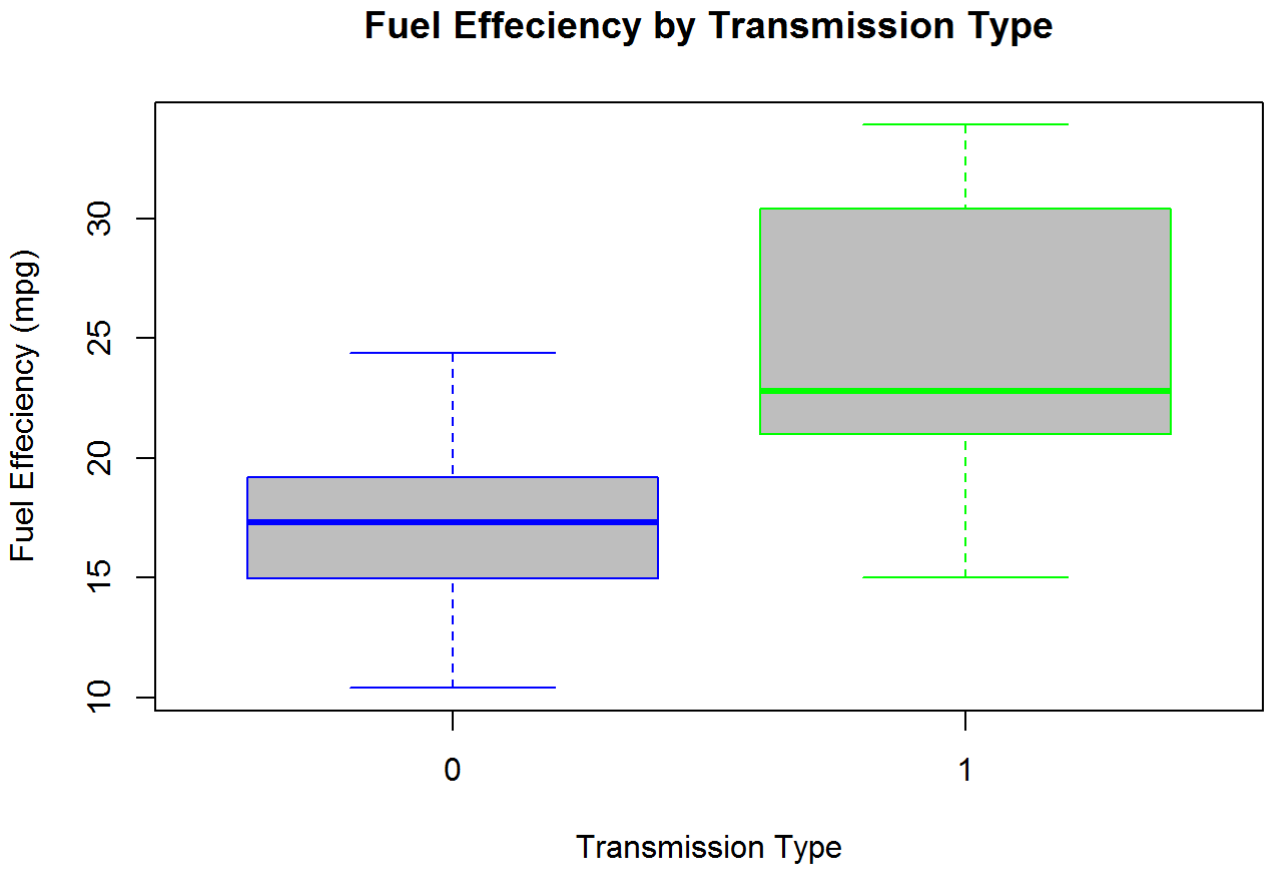
Considering wt, qsec and am explains 85% of the variance. Residual plots are included in Appendix A

Results

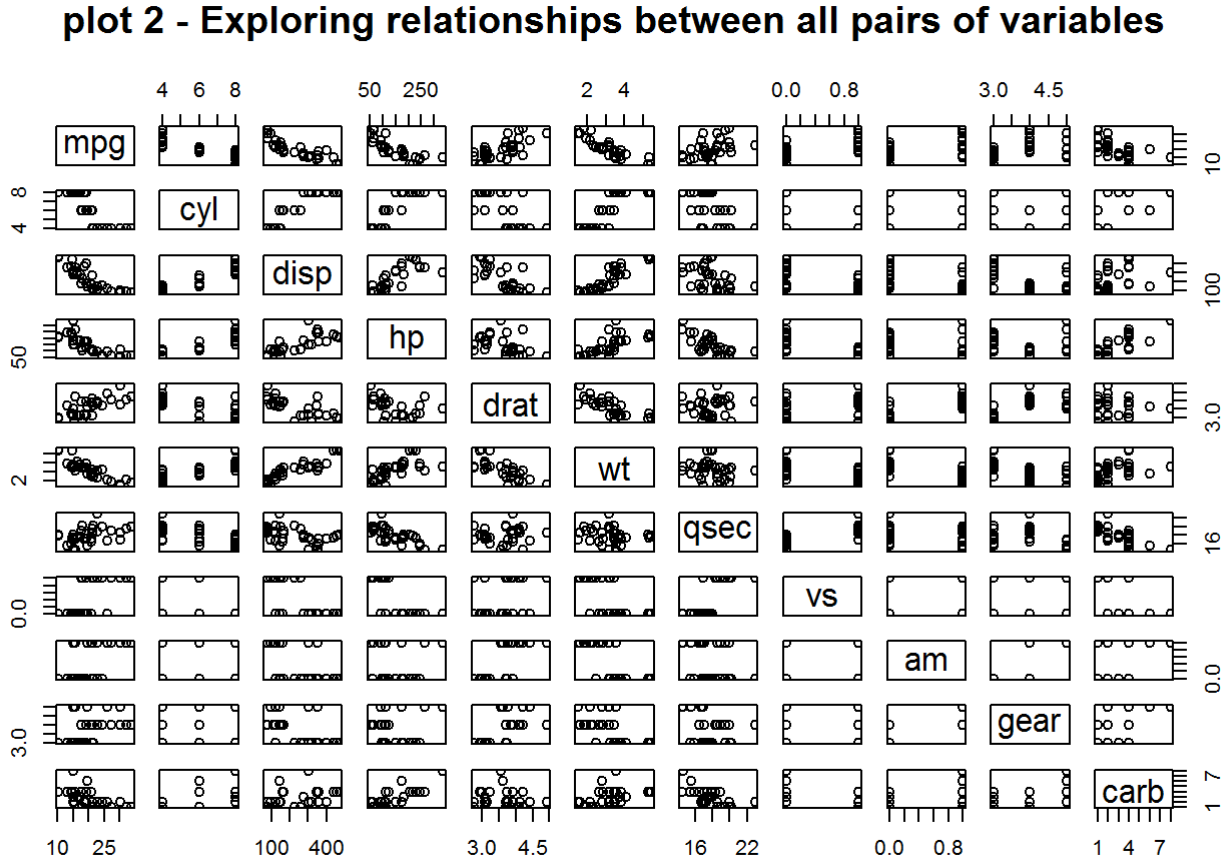
Car Weight and Acceleration must be considered to study impact of transmission type (Automatic or Manual) on the fuel effeciency (MPG) of the car

Appendix A

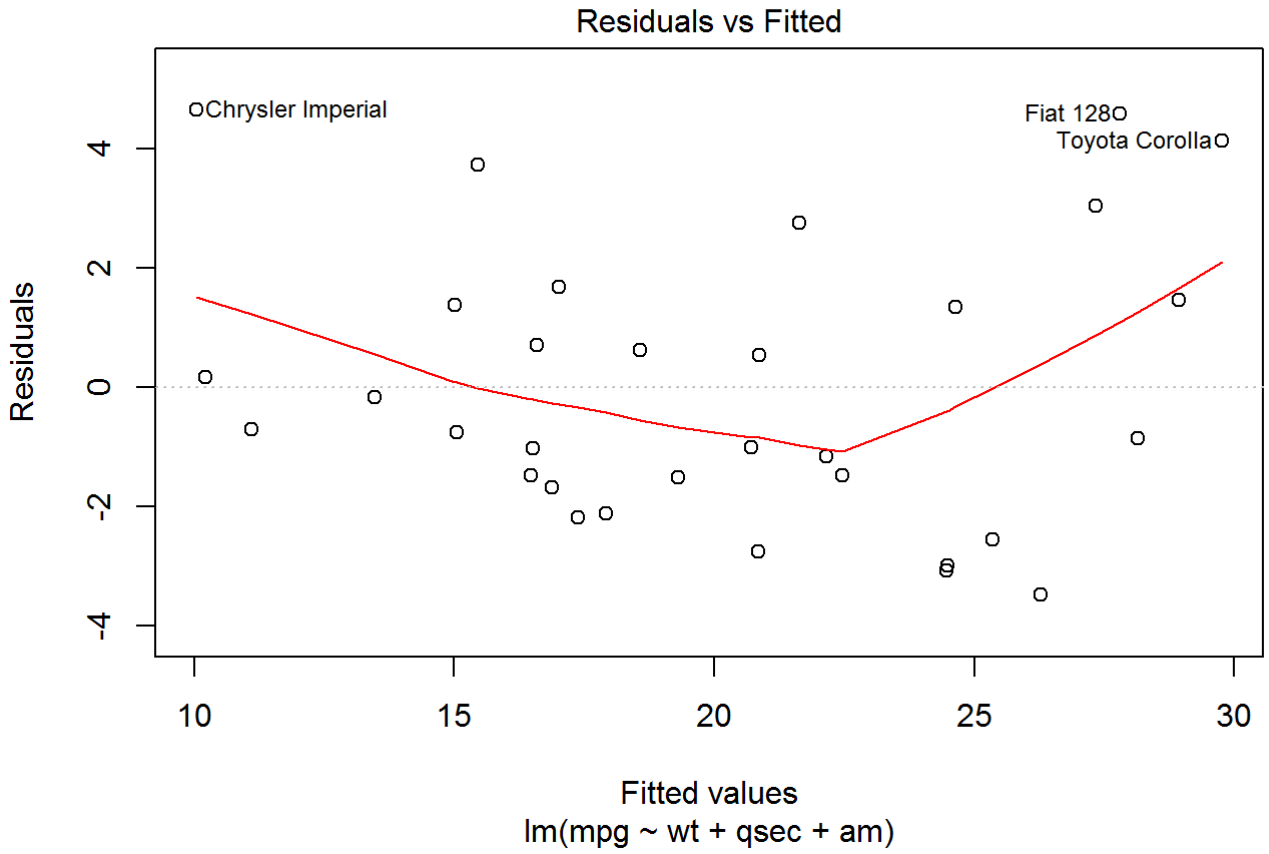
```
boxplot(mpg ~ am, data = mtcars, xlab = "Transmission Type", ylab = "Fuel Effeciency (mpg)",
col="gray",border = c("blue","green"), main="Fuel Effeciency by Transmission Type")
```

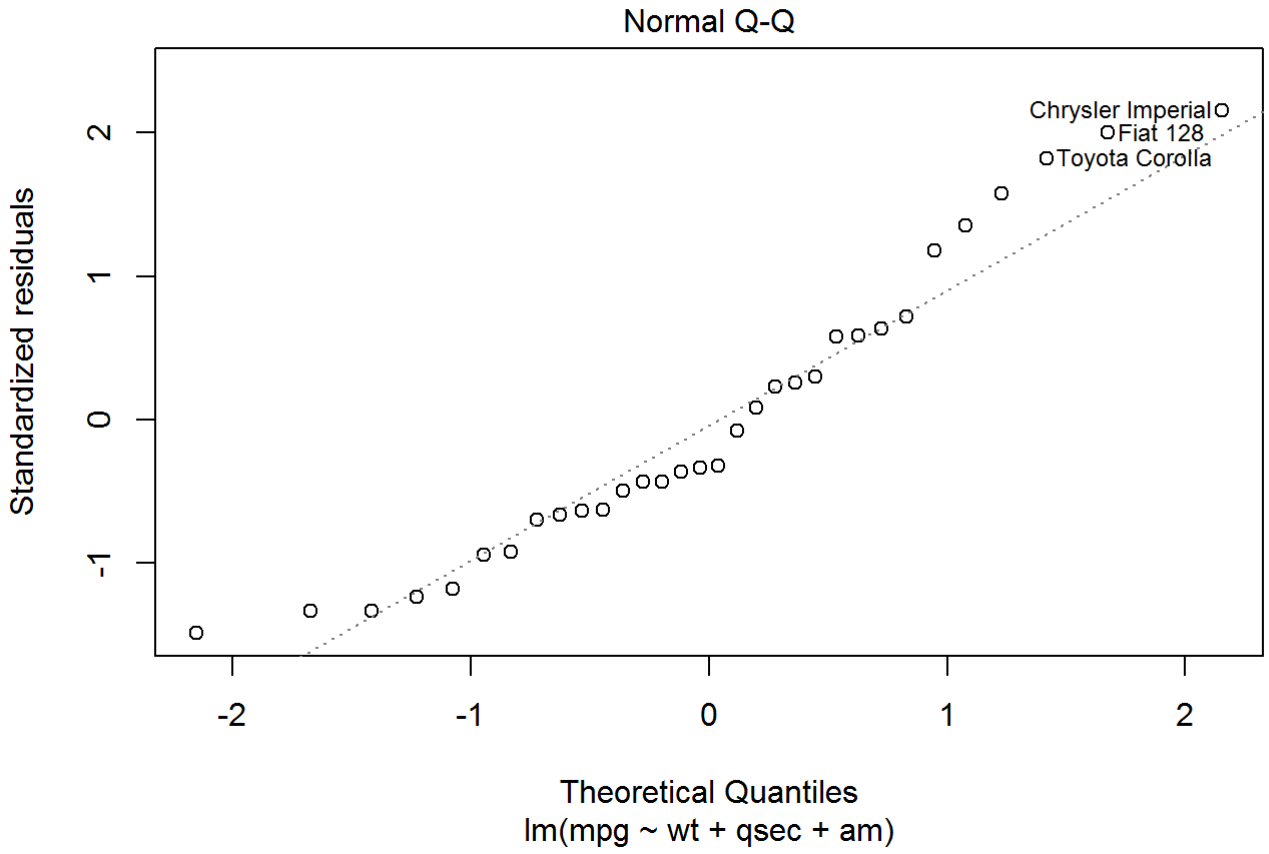


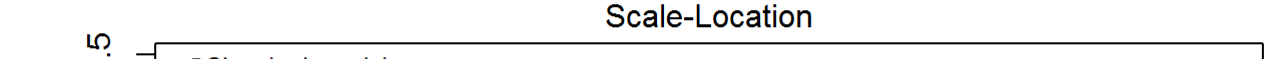

```
pairs(mpg ~ ., data = mtcars, main="plot 2 - Exploring relationships between all pairs of var
iables")
```

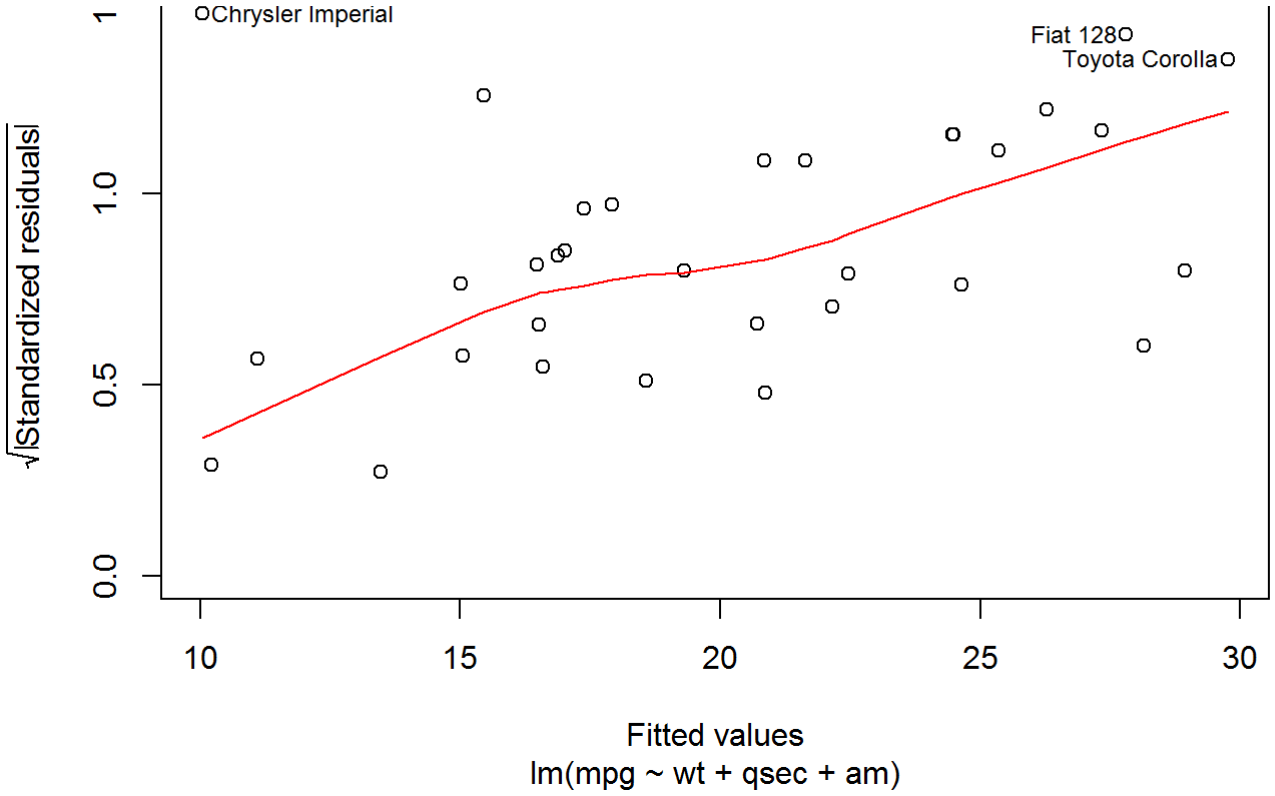


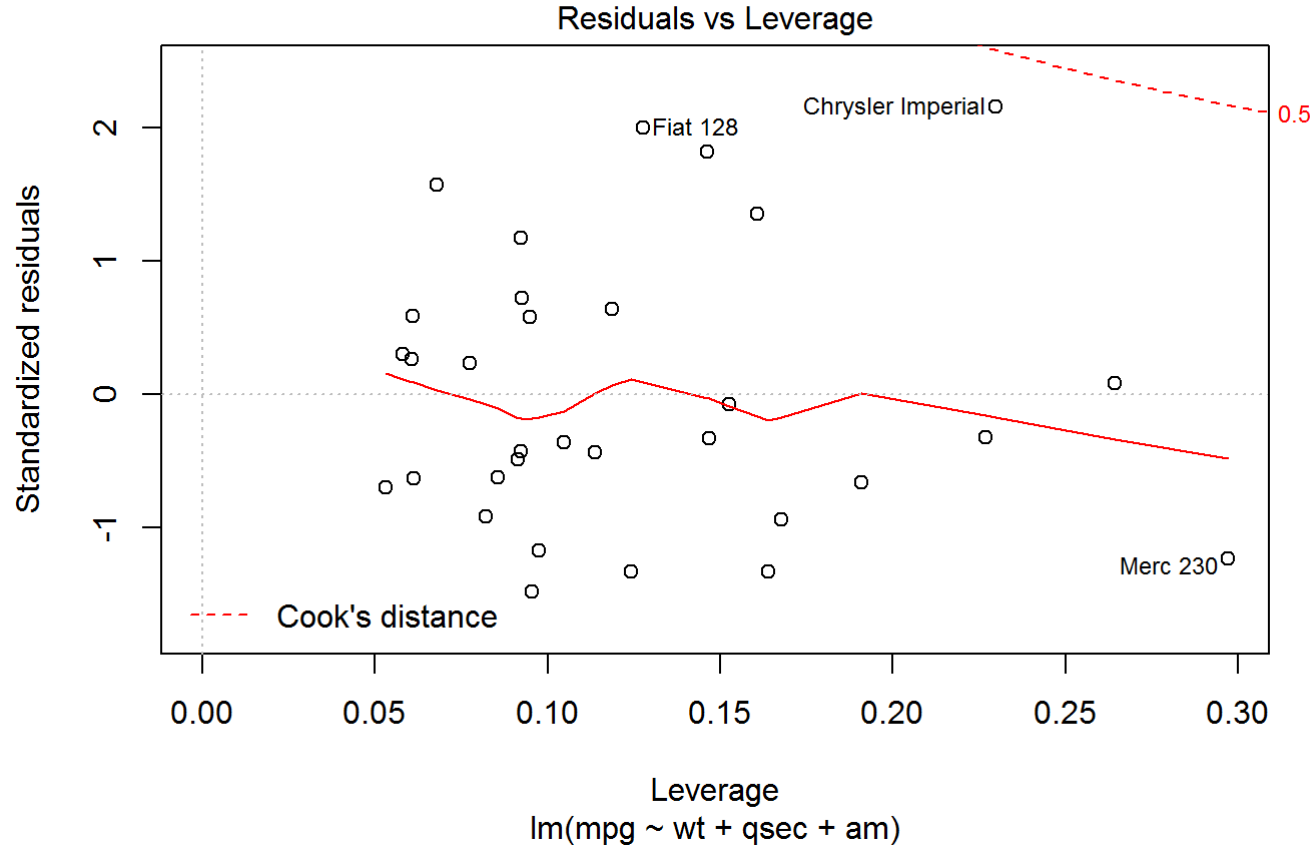
```
plot(automultivariate)
```


Residuals vs Fitted
Chrysler Imperial
Fiat 128
Toyota Corolla
Residuals
4
2
0
-2
-4
10
15
20
25
30
Fitted values
lm(mpg ~ wt + qsec + am)


Normal Q-Q
Chrysler Imperial
Fiat 128
Toyota Corolla
Standardized residuals
2
1
0
-1
-2
-1
0
1
2
Theoretical Quantiles
lm(mpg ~ wt + qsec + am)


Scale-Location


Chrysler Imperial
Fiat 128
Toyota Corolla
√|Standardized residuals|
0.0
0.5
1.0
1
10
15
20
25
30
Fitted values
lm(mpg ~ wt + qsec + am)


Residuals vs Leverage
Fiat 128
Chrysler Imperial
0.5
Merc 230
Cook's distance
Standardized residuals
-1
0
1
2
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Leverage
lm(mpg ~ wt + qsec + am)