Regression Models Week 4

##Executive Summary: This report aims to investigate if there is a difference in the fuel efficiency of automatic vs manual cars. Referencing from the mtcars dataset, regression models and exploratory analyses were done to address the above aim. It is concluded that while on the whole manual cars have better fuel efficiency that automatic cars.

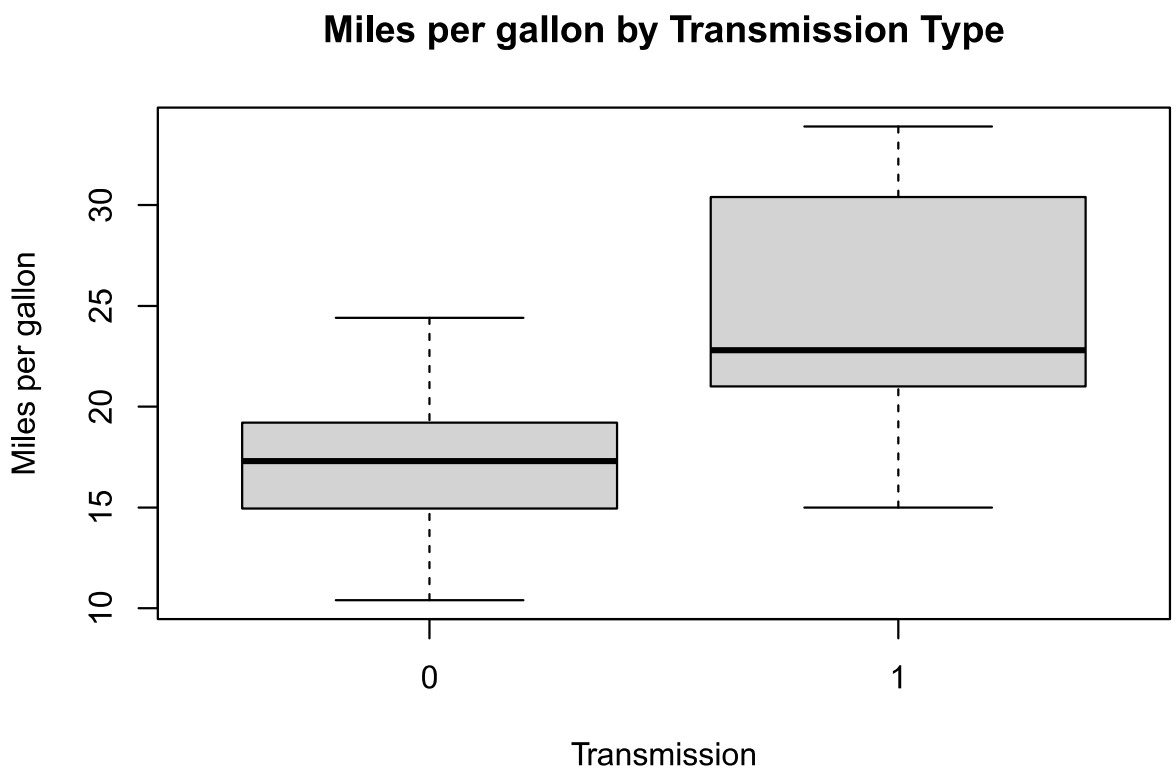
##Loading/Preparing data set

```
library(datasets)
data(mtcars)
summary(mtcars)
```

```
##       mpg             cyl             disp             hp       
##  Min.   :10.40   Min.   :4.000   Min.   : 71.1   Min.   : 52.0  
##  1st Qu.:15.43   1st Qu.:4.000   1st Qu.:120.8   1st Qu.: 96.5  
##  Median :19.20   Median :6.000   Median :196.3   Median :123.0  
##  Mean   :20.09   Mean   :6.188   Mean   :230.7   Mean   :146.7  
##  3rd Qu.:22.80   3rd Qu.:8.000   3rd Qu.:326.0   3rd Qu.:180.0  
##  Max.   :33.90   Max.   :8.000   Max.   :472.0   Max.   :335.0  
##       drat             wt             qsec             vs        
##  Min.   :2.760   Min.   :1.513   Min.   :14.50   Min.   :0.0000  
##  1st Qu.:3.080   1st Qu.:2.581   1st Qu.:16.89   1st Qu.:0.0000  
##  Median :3.695   Median :3.325   Median :17.71   Median :0.0000  
##  Mean   :3.597   Mean   :3.217   Mean   :17.85   Mean   :0.4375  
##  3rd Qu.:3.920   3rd Qu.:3.610   3rd Qu.:18.90   3rd Qu.:1.0000  
##  Max.   :4.930   Max.   :5.424   Max.   :22.90   Max.   :1.0000  
##        am              gear            carb      
##  Min.   :0.0000   Min.   :3.000   Min.   :1.000  
##  1st Qu.:0.0000   1st Qu.:3.000   1st Qu.:2.000  
##  Median :0.0000   Median :4.000   Median :2.000  
##  Mean   :0.4062   Mean   :3.688   Mean   :2.812  
##  3rd Qu.:1.0000   3rd Qu.:4.000   3rd Qu.:4.000  
##  Max.   :1.0000   Max.   :5.000   Max.   :8.000
```

##To see the broad comparison between MPG of automatic and manual cars. From the plot, it can be seen that on the whole, automatic cars have lower fuel efficiency, by 7.24 MPG.

```
boxplot(mpg ~ am, data = mtcars, xlab = "Transmission", ylab = "Miles per gallon", main="Miles per gall
```




```
aggregate(mpg~am, data = mtcars, mean)
```

```
##   am      mpg
## 1  0 17.14737
## 2  1 24.39231
```

##To further support this hypothesis, we perform a statistical analysis with a t-test. From the results, we can reject null hypothesis as we obtained a p-value of 0.000285, and conclude that there is a significant difference between fuel efficiency of manual and automatic transmissions.

```
auto <- mtcars[mtcars$am == 0,]
manual <- mtcars[mtcars$am == 1,]
t.test(auto$mpg, manual$mpg)
```

```
##
##  Welch Two Sample t-test
##
## data:  auto$mpg and manual$mpg
## 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 of x mean of y
##  17.14737  24.39231
```

##To further investigate the difference in fuel efficiency

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

```
##
## 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 the above, we can conclude that manual cars have 7.24 more MPG. However, as R^2 is 0.36, the model only accounts for 36% of variance. To perform multivariate linear regression to find relationship between weight, horsepower and MPG.

```
model <- lm(mpg~am + wt + hp, data = mtcars)
anova(m,model)
```

```
## Analysis of Variance Table
##
## Model 1: mpg ~ am
## Model 2: mpg ~ am + wt + hp
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1     30 720.90
## 2     28 180.29  2    540.61 41.979 3.745e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
summary(model)
```

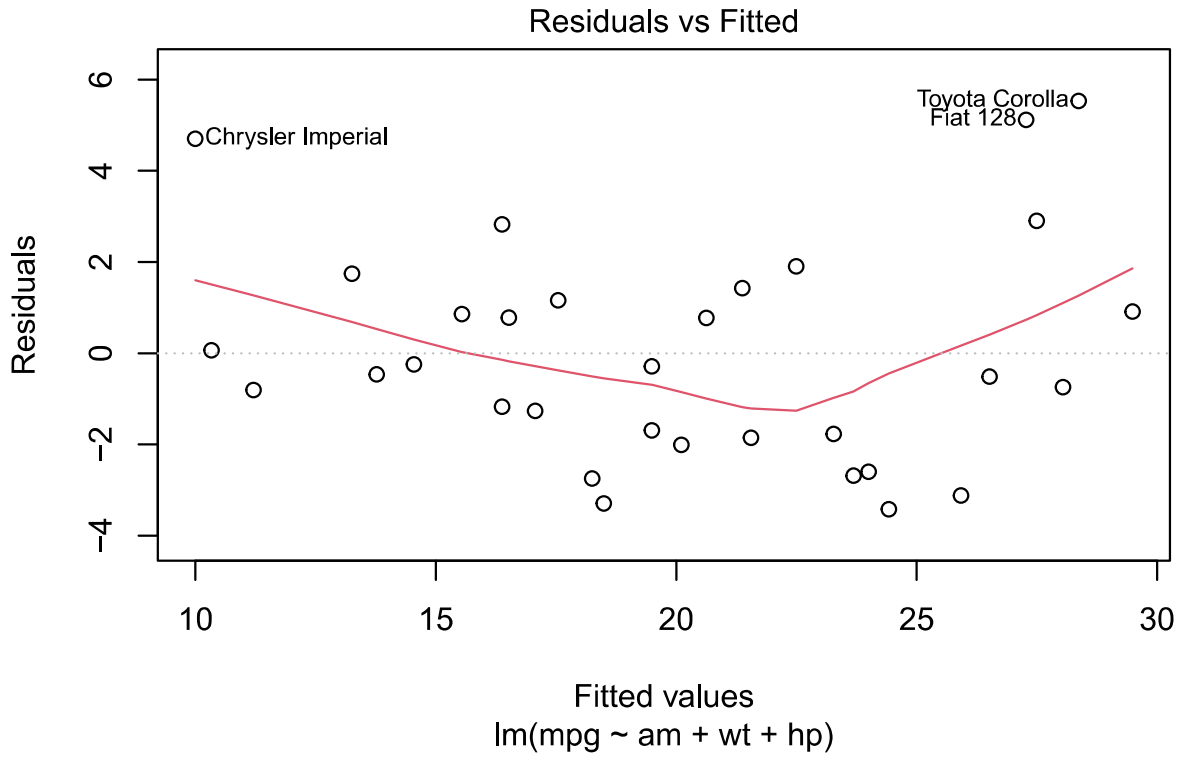
```
##
## Call:
## lm(formula = mpg ~ am + wt + hp, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.4221 -1.7924 -0.3788  1.2249  5.5317
##
## Coefficients:
```

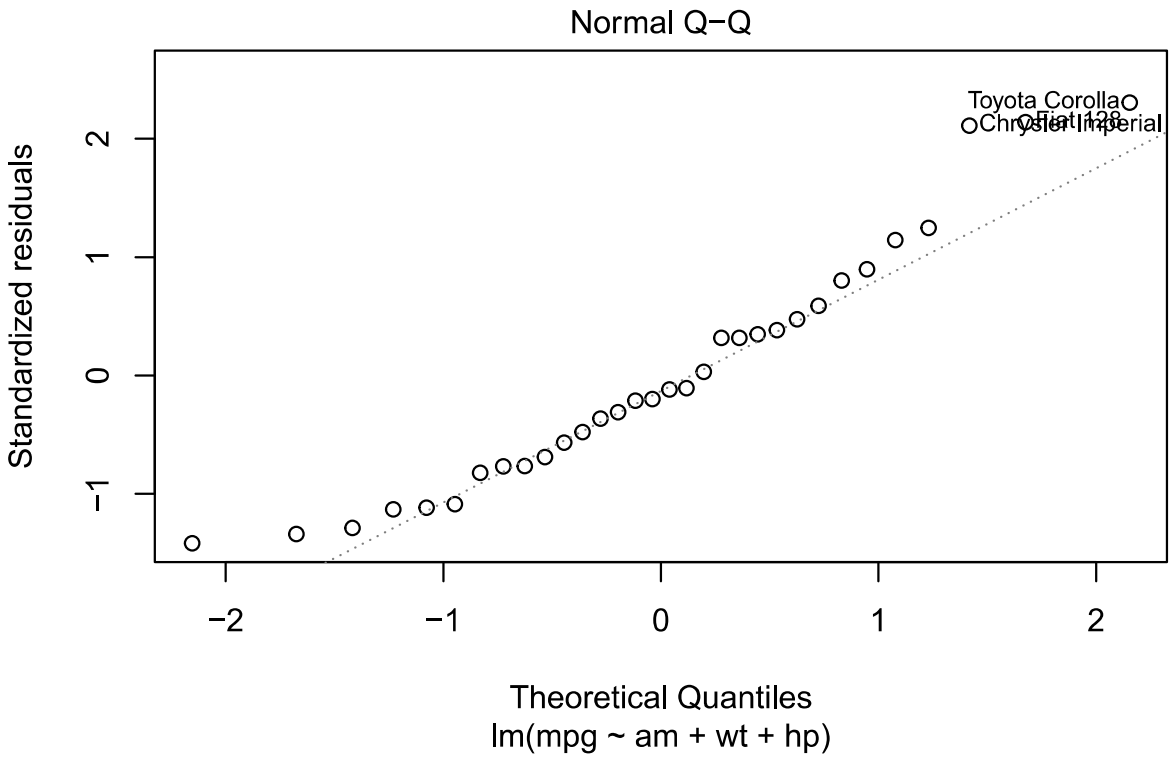
```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 34.002875   2.642659  12.867 2.82e-13 ***
## am           2.083710   1.376420   1.514 0.141268
## wt          -2.878575   0.904971  -3.181 0.003574 **
## hp          -0.037479   0.009605  -3.902 0.000546 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.538 on 28 degrees of freedom
## Multiple R-squared:  0.8399, Adjusted R-squared:  0.8227
## F-statistic: 48.96 on 3 and 28 DF,  p-value: 2.908e-11
```

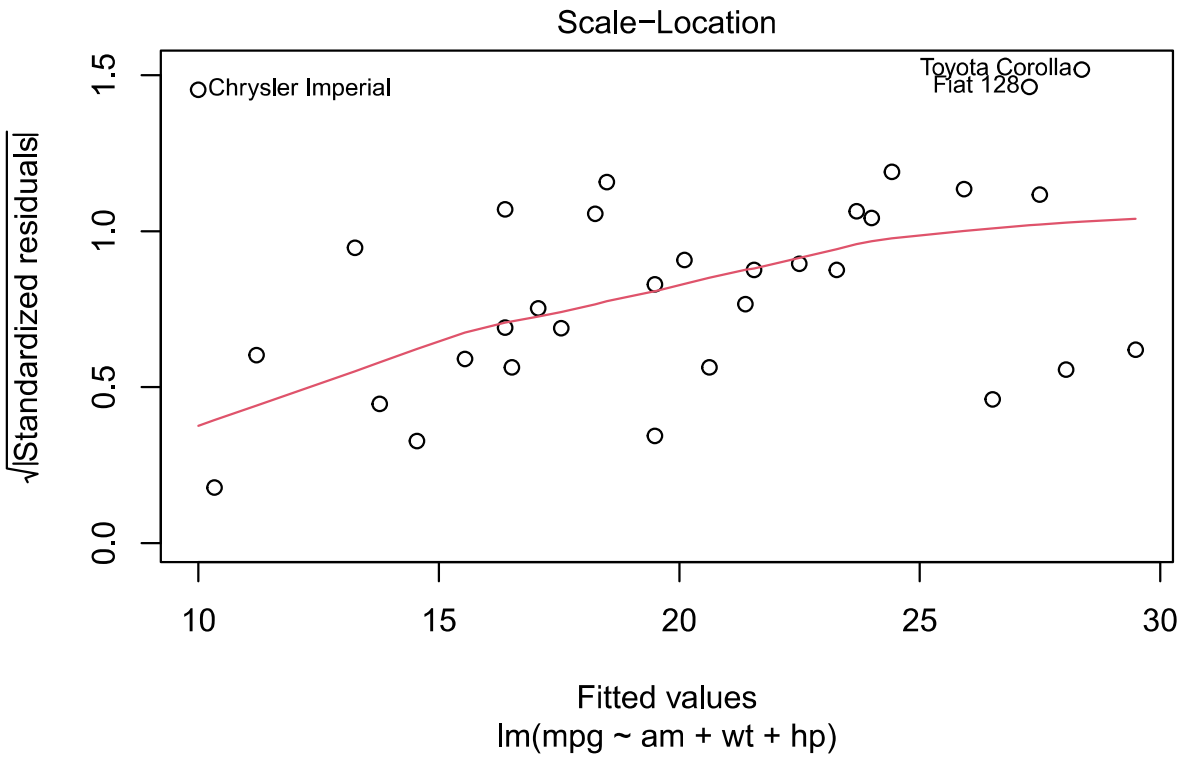
##The above model explains for 84% of the variance and it can be concluded that on the whole, manual transmissions have 1.478 more mpg than automatic.

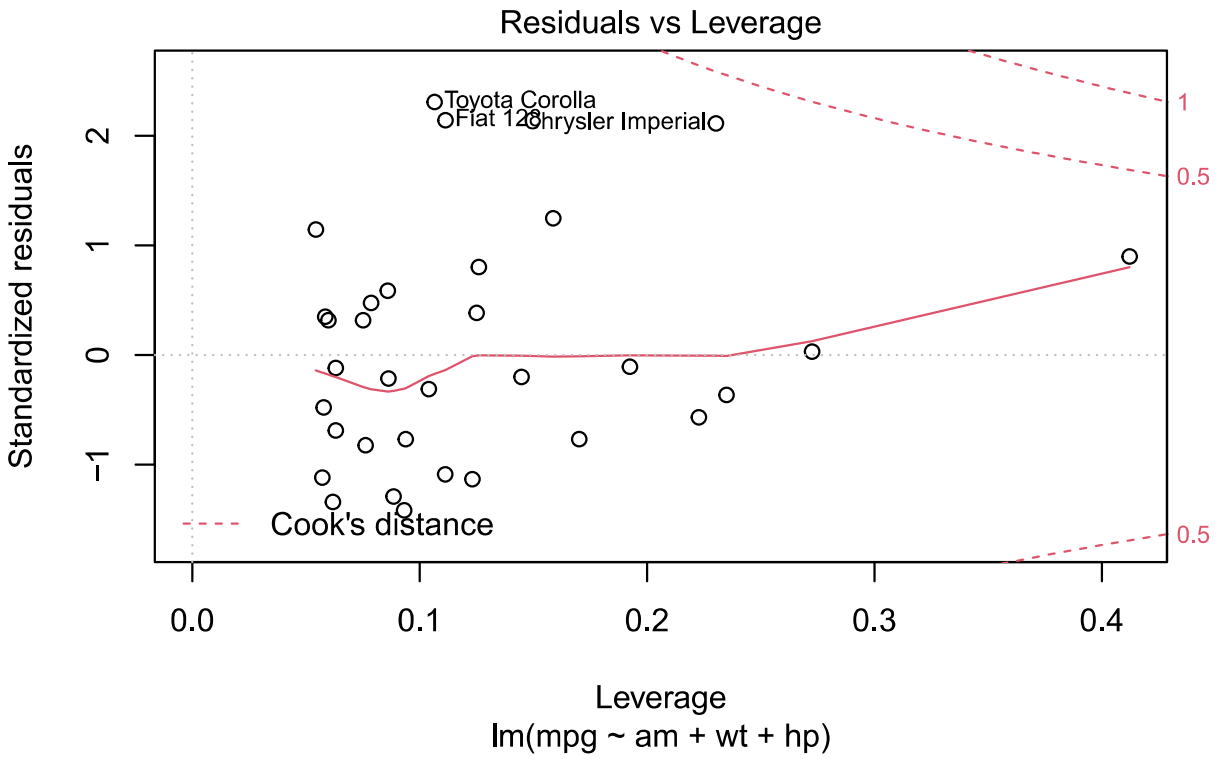
##Appendix

```
plot(model)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(mpg ~ am + wt + hp)
Toyota Corolla
Fiat 128
Chrysler Imperial


Scale−Location
Chrysler Imperial
Toyota Corolla
Fiat 128
√|Standardized residuals|
0.0
0.5
1.0
1.5
10
15
20
25
30
Fitted values
lm(mpg ~ am + wt + hp)


Residuals vs Leverage
Toyota Corolla
Fiat 128
Chrysler Imperial
1
0.5
0.5
Cook's distance
Standardized residuals
2
1
0
−1
0.0
0.1
0.2
0.3
0.4
Leverage
lm(mpg ~ am + wt + hp)