

# Regression

Markos Hay

12/04/2020

## Motor Trend Car Road Analysis

### Executive Summary

You work for Motor Trend, a magazine about the automobile industry. Looking at a data set of a collection of cars, they are interested in exploring the relationship between a set of variables and miles per gallon (MPG) (outcome). They are particularly interested in the following two questions:

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

### Data Processing

The data was extracted from the 1974 Motor Trend US magazine, and comprises fuel consumption and 10 aspects of automobile design and performance for 32 automobiles (1973–74 models).

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

It consists of 32 observations on 11 variables.

[, 1] mpg Miles/(US) gallon [, 2] cyl Number of cylinders [, 3] disp Displacement (cu.in.) [, 4] hp Gross horsepower [, 5] drat Rear axle ratio [, 6] wt Weight (lb/1000) [, 7] qsec 1/4 mile time [, 8] vs V/S [, 9] am Transmission (0 = automatic, 1 = manual) [,10] gear Number of forward gears [,11] carb Number of carburetors

### Is an automatic or manual transmission better for MPG?

For Automatic

```
summary(mtcars[mtcars$am==0,])
```

##	mpg	cyl	disp	hp
##	Min. :10.40	Min. :4.000	Min. :120.1	Min. : 62.0
##	1st Qu.:14.95	1st Qu.:6.000	1st Qu.:196.3	1st Qu.:116.5
##	Median :17.30	Median :8.000	Median :275.8	Median :175.0
##	Mean :17.15	Mean :6.947	Mean :290.4	Mean :160.3
##	3rd Qu.:19.20	3rd Qu.:8.000	3rd Qu.:360.0	3rd Qu.:192.5
##	Max. :24.40	Max. :8.000	Max. :472.0	Max. :245.0

```
##      drat      wt      qsec      vs      am
## Min.   :2.760  Min.   :2.465  Min.   :15.41  Min.   :0.0000  Min.   :0
## 1st Qu.:3.070  1st Qu.:3.438  1st Qu.:17.18  1st Qu.:0.0000  1st Qu.:0
## Median :3.150  Median :3.520  Median :17.82  Median :0.0000  Median :0
## Mean   :3.286  Mean   :3.769  Mean   :18.18  Mean   :0.3684  Mean   :0
## 3rd Qu.:3.695  3rd Qu.:3.842  3rd Qu.:19.17  3rd Qu.:1.0000  3rd Qu.:0
## Max.   :3.920  Max.   :5.424  Max.   :22.90  Max.   :1.0000  Max.   :0
##      gear      carb
## Min.   :3.000  Min.   :1.000
## 1st Qu.:3.000  1st Qu.:2.000
## Median :3.000  Median :3.000
## Mean   :3.211  Mean   :2.737
## 3rd Qu.:3.000  3rd Qu.:4.000
## Max.   :4.000  Max.   :4.000
```

For Manual

```
summary(mtcars[mtcars$am==1,])
```

```
##      mpg      cyl      disp      hp      drat
## Min.   :15.00  Min.   :4.000  Min.   : 71.1  Min.   : 52.0  Min.   :3.54
## 1st Qu.:21.00  1st Qu.:4.000  1st Qu.: 79.0  1st Qu.: 66.0  1st Qu.:3.85
## Median :22.80  Median :4.000  Median :120.3  Median :109.0  Median :4.08
## Mean   :24.39  Mean   :5.077  Mean   :143.5  Mean   :126.8  Mean   :4.05
## 3rd Qu.:30.40  3rd Qu.:6.000  3rd Qu.:160.0  3rd Qu.:113.0  3rd Qu.:4.22
## Max.   :33.90  Max.   :8.000  Max.   :351.0  Max.   :335.0  Max.   :4.93
##      wt      qsec      vs      am      gear
## Min.   :1.513  Min.   :14.50  Min.   :0.0000  Min.   :1  Min.   :4.000
## 1st Qu.:1.935  1st Qu.:16.46  1st Qu.:0.0000  1st Qu.:1  1st Qu.:4.000
## Median :2.320  Median :17.02  Median :1.0000  Median :1  Median :4.000
## Mean   :2.411  Mean   :17.36  Mean   :0.5385  Mean   :1  Mean   :4.385
## 3rd Qu.:2.780  3rd Qu.:18.61  3rd Qu.:1.0000  3rd Qu.:1  3rd Qu.:5.000
## Max.   :3.570  Max.   :19.90  Max.   :1.0000  Max.   :1  Max.   :5.000
##      carb
## Min.   :1.000
## 1st Qu.:1.000
## Median :2.000
## Mean   :2.923
## 3rd Qu.:4.000
## Max.   :8.000
```

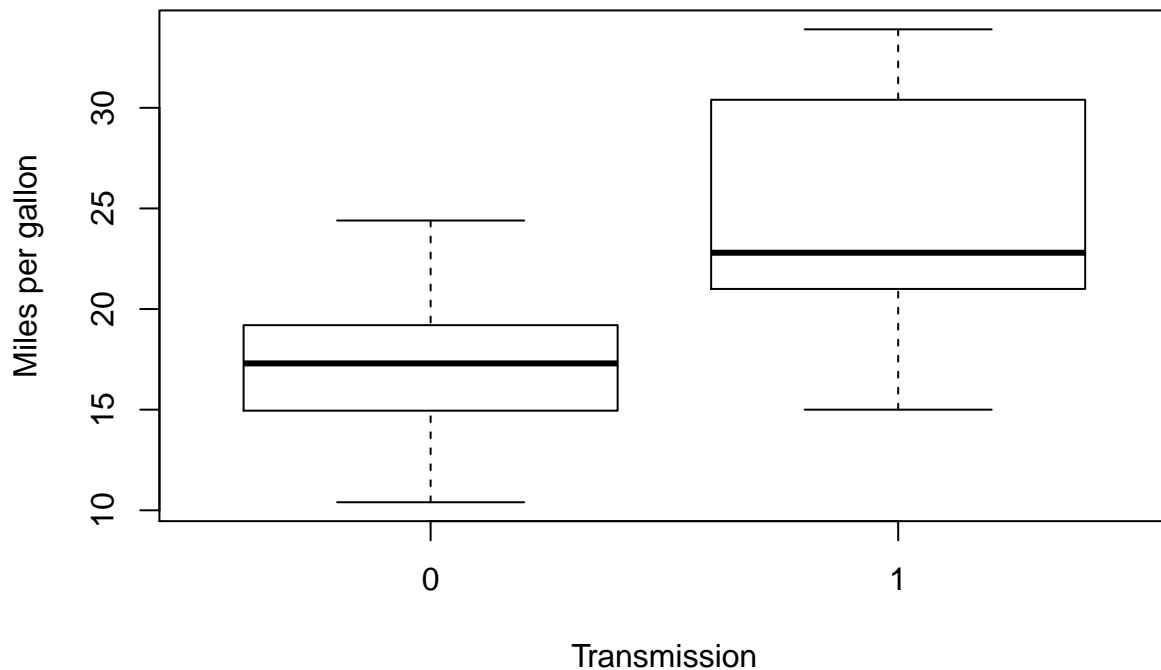
Hence, the mean of mpg is greater for manual (at 24.4) than automatic (at 17.1).

Investigating further..

**Quantify the MPG difference between automatic and manual transmissions.**

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

## Miles per gallon by Transmission Type



Manual (represented by 1) has a higher mean for mpg than automatic (represented by 0).

## Hypothesis Testing

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

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

The mean transmission for manual is 7.24mpg higher than automatic. Let  $\alpha=0.5$ .

```
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
```

Since  $p\text{-value} = 0.001374$ , we reject the null hypothesis. There is a major difference between mpg of manual and automatic transmissions.

```
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
```

Since  $p\text{-value} = 0.001374$ , we reject the null hypothesis. There is a major difference between mpg of manual and automatic transmissions.

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

```
## Analysis of Variance Table
##
## Model 1: mpg ~ am
## Model 2: mpg ~ am + wt + hp + cyl
##   Res.Df  RSS Df Sum of Sq    F    Pr(>F)
## 1      30 720.9
## 2      27 170.0  3      550.9 29.166 1.274e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The final model is below:

```
summary(model)
```

```
##
```

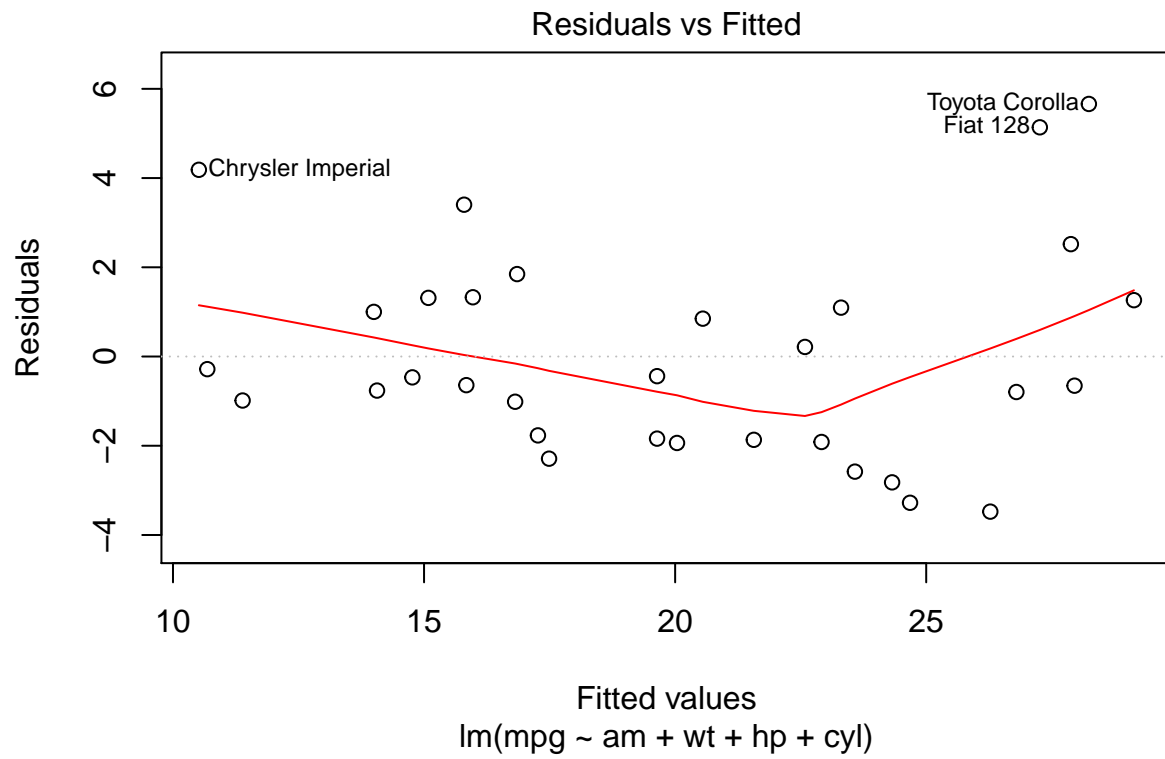
```
## Call:
## lm(formula = mpg ~ am + wt + hp + cyl, data = mtcars)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.4765 -1.8471 -0.5544  1.2758  5.6608
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  36.14654    3.10478   11.642 4.94e-12 ***
## am           1.47805    1.44115    1.026  0.3142
## wt          -2.60648    0.91984   -2.834  0.0086 **
## hp           -0.02495    0.01365   -1.828  0.0786 .
## cyl          -0.74516    0.58279   -1.279  0.2119
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.509 on 27 degrees of freedom
## Multiple R-squared:  0.849, Adjusted R-squared:  0.8267
## F-statistic: 37.96 on 4 and 27 DF, p-value: 1.025e-10
```

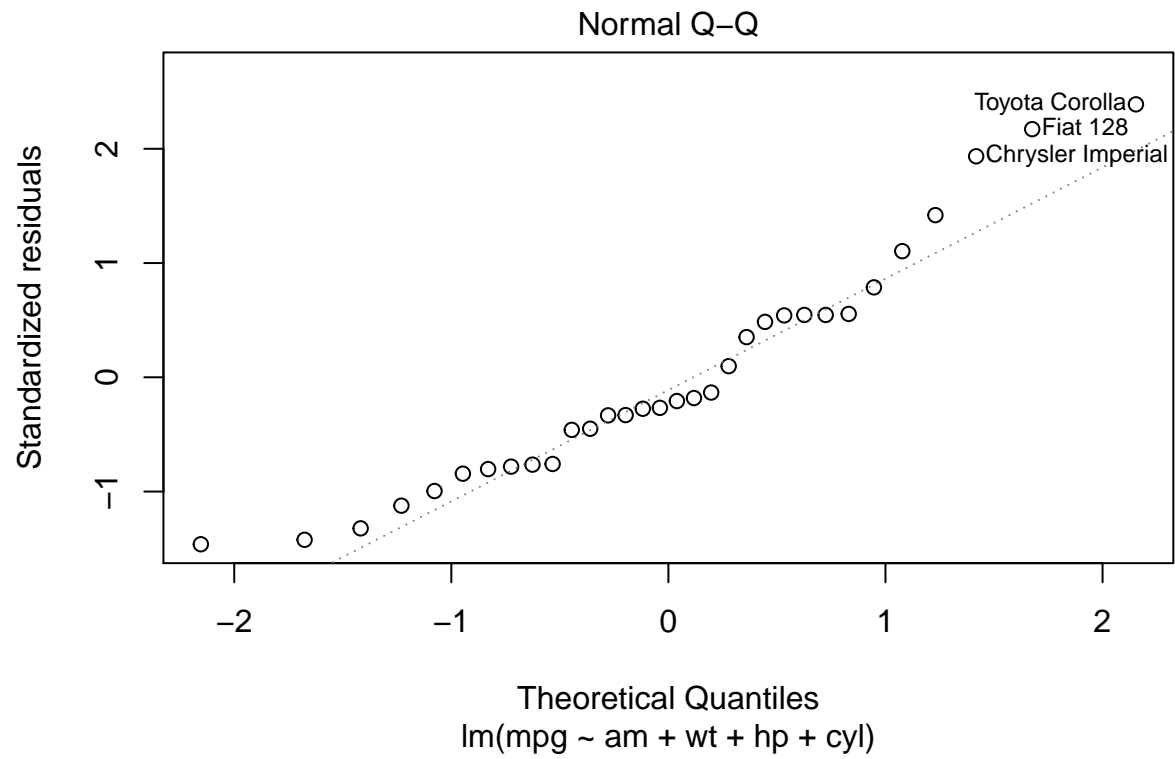
## Concolusion

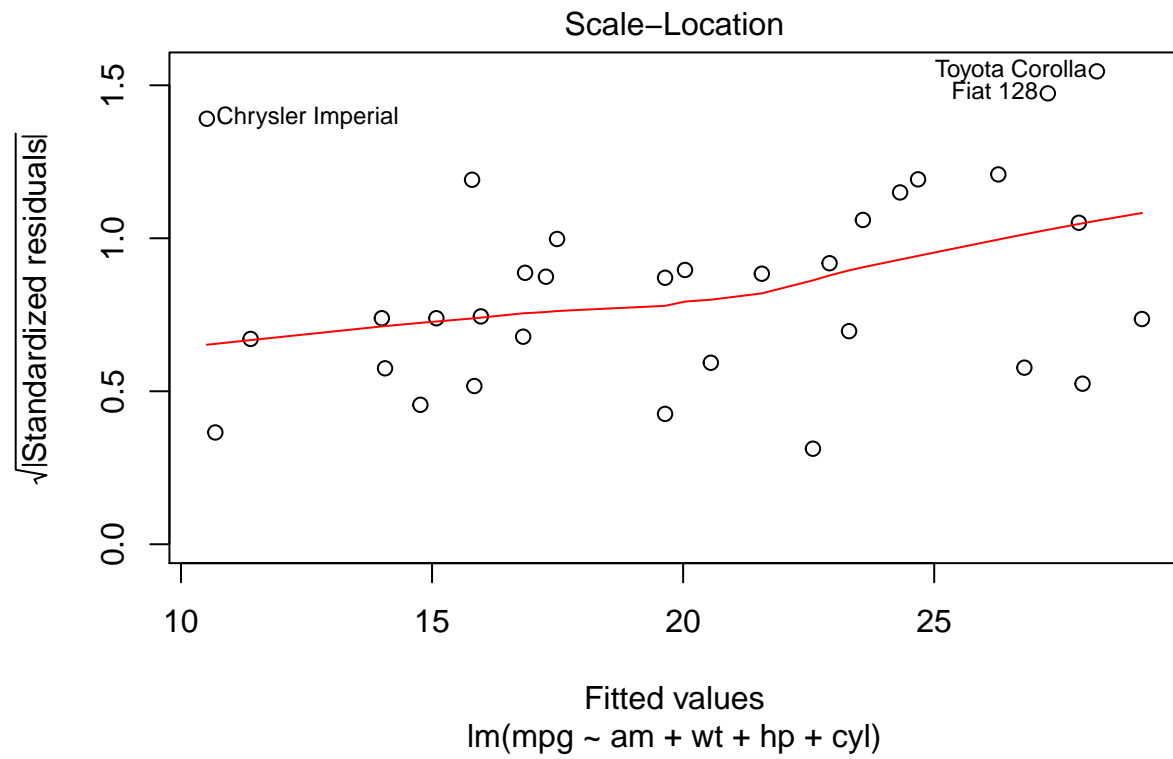
This model explains 84.9% of the variance. It may be concluded that on average, manual transmissions have 1.478 more mpg than automatic.

## Appendix

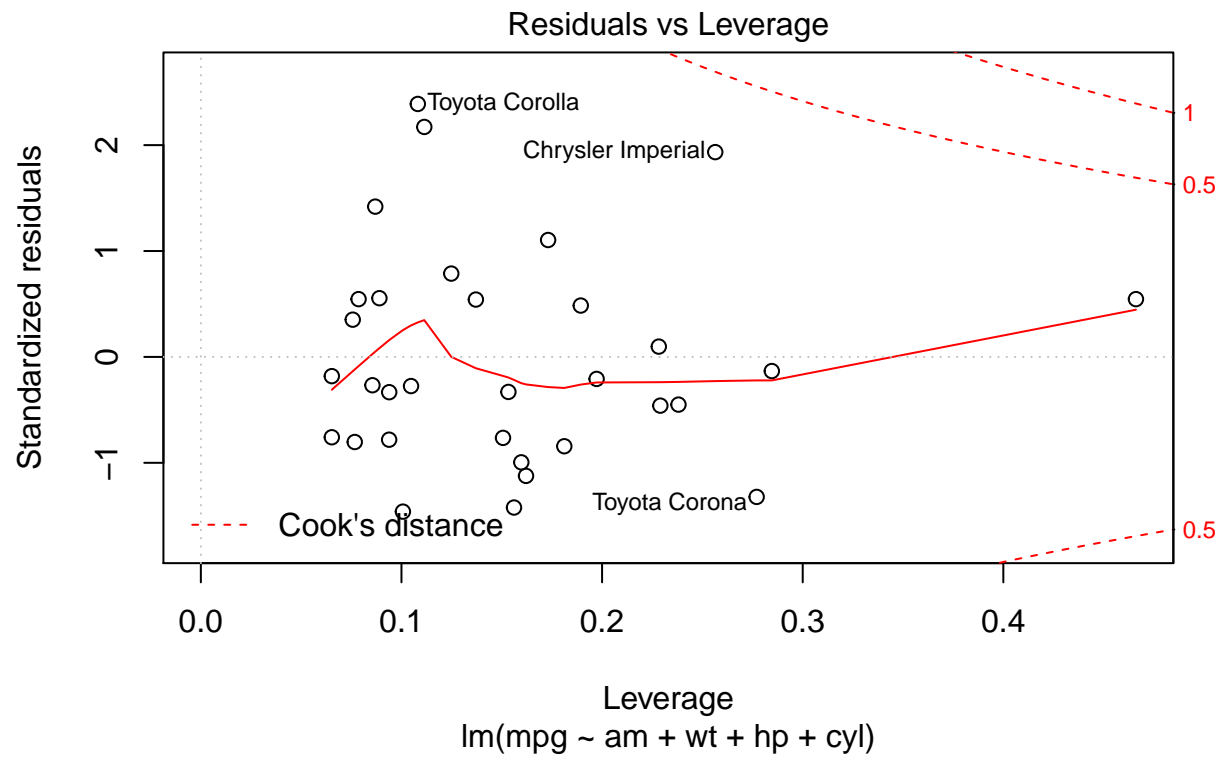
```
plot(model)
```











Hence, the residuals are normally distributed, and homoskedastic.