Which is better for MPG - automatic or manual

Executive Summary

While there are many factors contributing to a car's miles per gallon(MPG) indicator, such as weight, horse power, and the number of cylinders. Today we will focus mainly on the transmission type - automatic and manual. Based on the linear regression model, it can be concluded that the manual cars run 7.24 more miles per gallon than automatic on average. Therefore, the transmisiion is better for MPG than the automatic one.

Data Prepossessing

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).

```
data(mtcars)
View(mtcars)
mtcars$cyl<-factor(mtcars$cyl)
mtcars$vs<-factor(mtcars$vs)
mtcars$gear<-factor(mtcars$gear)
mtcars$carb<-factor(mtcars$carb)
mtcars$am<-factor(mtcars$am,labels=c("automatic","manual"))
```

Below is abbreviations for the 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;

Compare means of transmission types

```
t.test(mtcars$mpg~mtcars$am)
```

```
##
##  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 automatic    mean in group manual
##                17.14737                24.39231
```

A p-value less than 0.01 proves that there exists a difference on MPG between automatic and manual transmission. It could be easily told that the advantage of automatic cars is 7.24 MPG.

Model Building

```
model1<-lm(mpg~.,data=mtcars)
model2<-step(model1,derection="backward",k=2,trace=0)
```

At first we construct a variables-all-included model, then we use step() to eliminate variables that are not significant.

```
summary(model2)
```

```
##
## Call:
## lm(formula = mpg ~ cyl + hp + wt + am, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.9387 -1.2560 -0.4013  1.1253  5.0513
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 33.70832    2.60489  12.940 7.73e-13 ***
## cyl6        -3.03134    1.40728  -2.154  0.04068 *
## cyl8        -2.16368    2.28425  -0.947  0.35225
## hp          -0.03211    0.01369  -2.345  0.02693 *
## wt          -2.49683    0.88559  -2.819  0.00908 **
## ammanual     1.80921    1.39630   1.296  0.20646
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.41 on 26 degrees of freedom
## Multiple R-squared:  0.8659, Adjusted R-squared:  0.8401
## F-statistic: 33.57 on 5 and 26 DF,  p-value: 1.506e-10
```

The probability of coefficients suggest that cylinders, horsepower, weight are significant variables(smaller than 0.05). and R square for the selected model is 0.84.

Residual Analysis

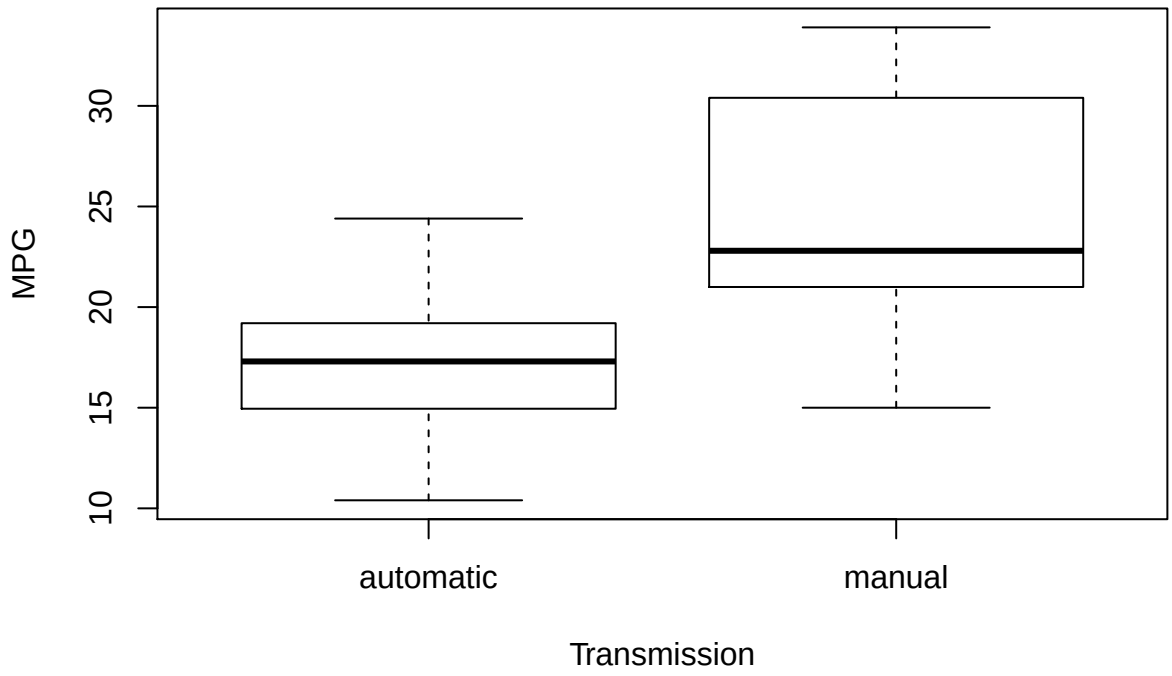
Residuals plots can be viewed in Appendix. It is suggested that residuals are scattered randomly. QQ plot illustrates that residuals are normally distributed.

Uncertainty

The conclusion is reached based on a limited number of observations(32). If more samples are provided, the support will be more solid.

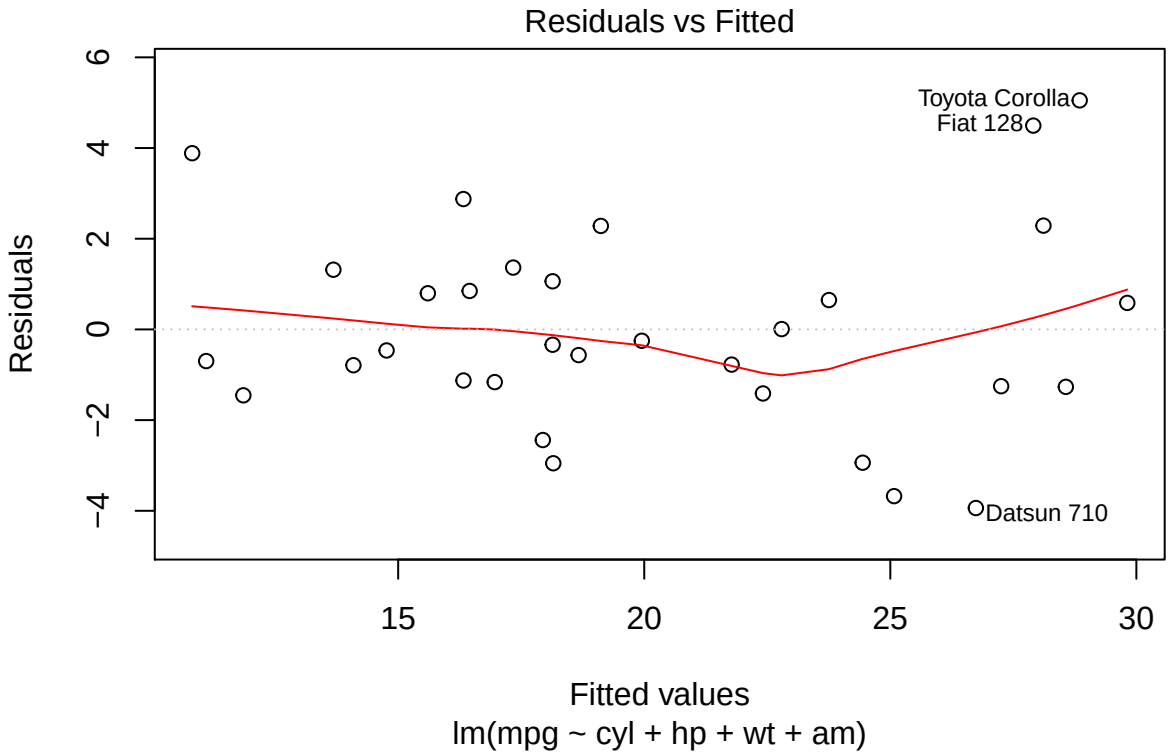
Appendix

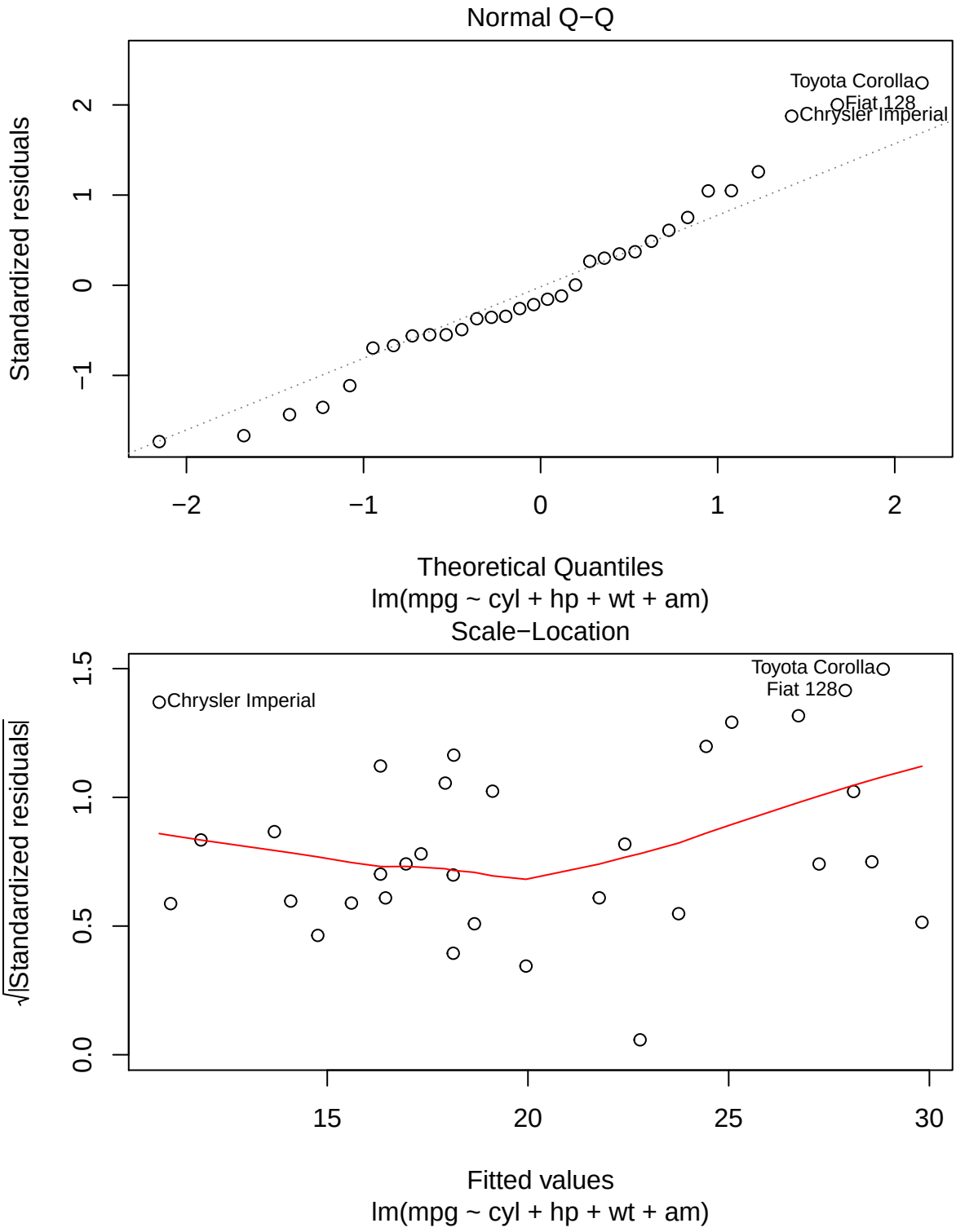
```
boxplot(mpg~am,data=mtcars,xlab="Transmission",ylab="MPG")
```

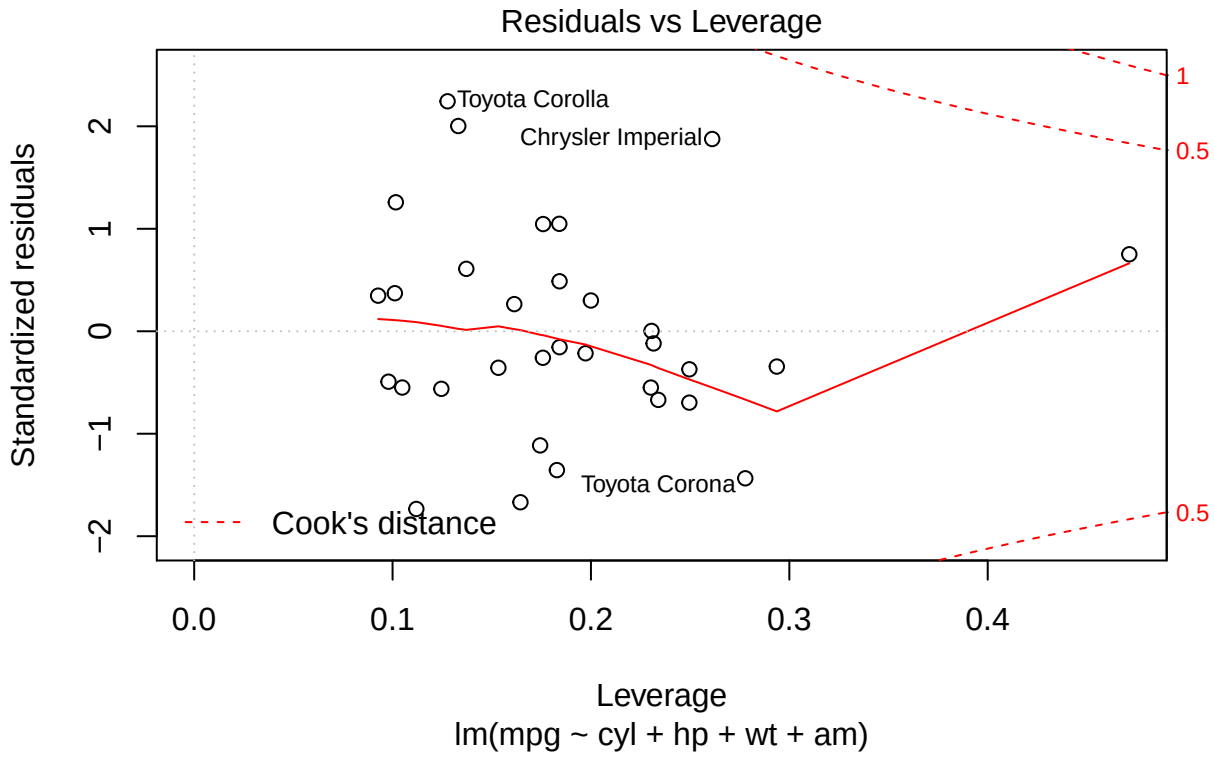


Boxplot of automatic and manual transmission cars

```
plot(model2)
```




Normal Q−Q
Toyota Corolla
Fiat 128
Chrysler Imperial
Standardized residuals
−2
−1
0
1
2
Theoretical Quantiles
lm(mpg ~ cyl + hp + wt + am)
Scale−Location
Toyota Corolla
Fiat 128
Chrysler Imperial
√|Standardized residuals|
0.0
0.5
1.0
1.5
15
20
25
30
Fitted values
lm(mpg ~ cyl + hp + wt + am)



Residual plots of fitted model