Regression Models project

In this document we are exploring variables of mtcar dataset. We are specially interested in the relationship between miles per gallon and car transmission. For the study we are using a linear regression model.

Load and preprocess data

We first load mtcars data and preprocess it. We are replacing 0 and 1 values by the proper transmission label (0 = automatic, 1 = manual)

```
#load data
data(mtcars)

#change the number by a label
mtcars$am[mtcars$am==0] <- "automatic"
mtcars$am[mtcars$am==1] <- "manual"
mtcars$am <- as.factor(mtcars$am)
```

Exploratory analysis and regression

In Fig. 1 and Fig 2 in the appendix we show some plots explaining the relation between mpg and transmission. As we can see in Fig. 2 it seems that in general manual transmission have higher MPG value.

```
fit<-lm(mtcars$mpg~mtcars$am)
summary(fit)
```

```
##
## Call:
## lm(formula = mtcars$mpg ~ mtcars$am)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -9.392 -3.092 -0.297  3.244  9.508
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)        17.15       1.12   15.25  1.1e-15 ***
## mtcars$ammanual     7.24       1.76    4.11  0.00029 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.9 on 30 degrees of freedom
## Multiple R-squared:  0.36,   Adjusted R-squared:  0.338
## F-statistic: 16.9 on 1 and 30 DF,  p-value: 0.000285
```

Since manual coefficient is positive we expect to have a higher mpg value. In fact we predict to have 7.24 more mpg value if transmission is automatic.

Diagnosis

In Fig 3. (in the appendix) we show the plot of the residuals. We cannot see any pattern, so we deduce that linear regression is enough good and there is no need to search for other models (Polynomial regression, logistic regression...).

For estimating how much higher we would expect to be MPG value having a manual transmission car:

```
sumCoef.coef<-summary(fit)$coefficient[1,1]
sumCoef.stdError<-summary(fit)$coefficient[1,2]

intrv<-sumCoef.coef+c(-1,1)*qt(.975,df=fit$df)*sumCoef.stdError
intrv
```

```
## [1] 14.85 19.44
```

We can conclude with 95% confidence the mpg will be between 14.85 and 19.44 higher if the transmission is manual.

Summary

To sum up, at overall having a manual transmission car we can cover more miles per each gallon. With this study we have reached to the conclusion that we should expect about 7.24 more miles.
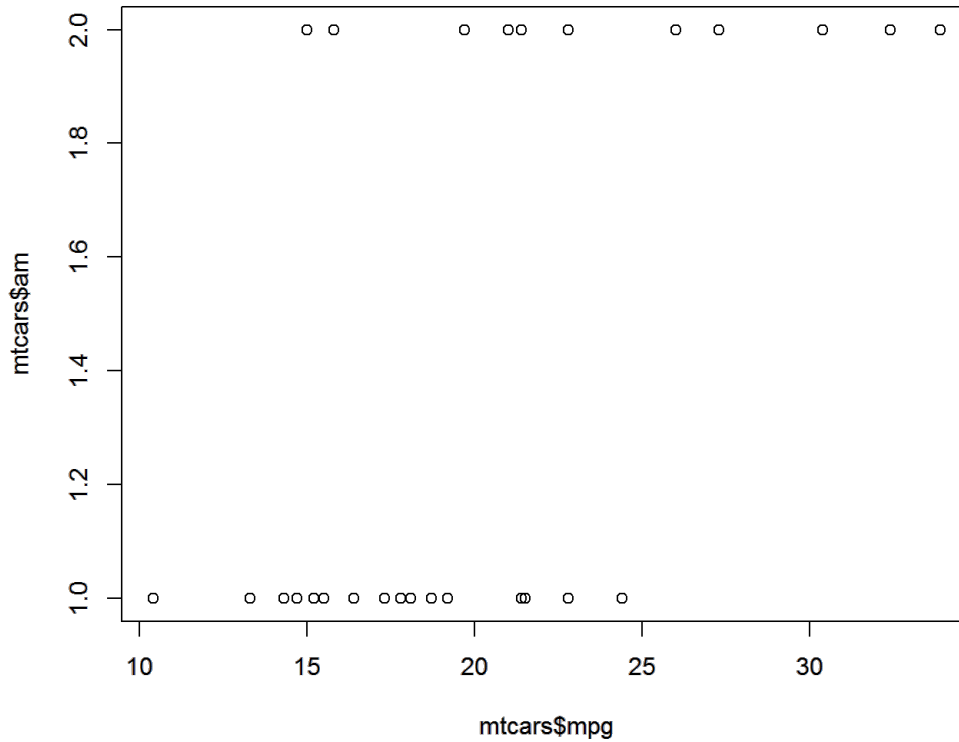
Appendix

```
plot(mtcars$mpg,mtcars$am)
```


Fig. 1: Plot of mpg vs transmission

```
boxplot(mtcars$mpg~mtcars$am)
```
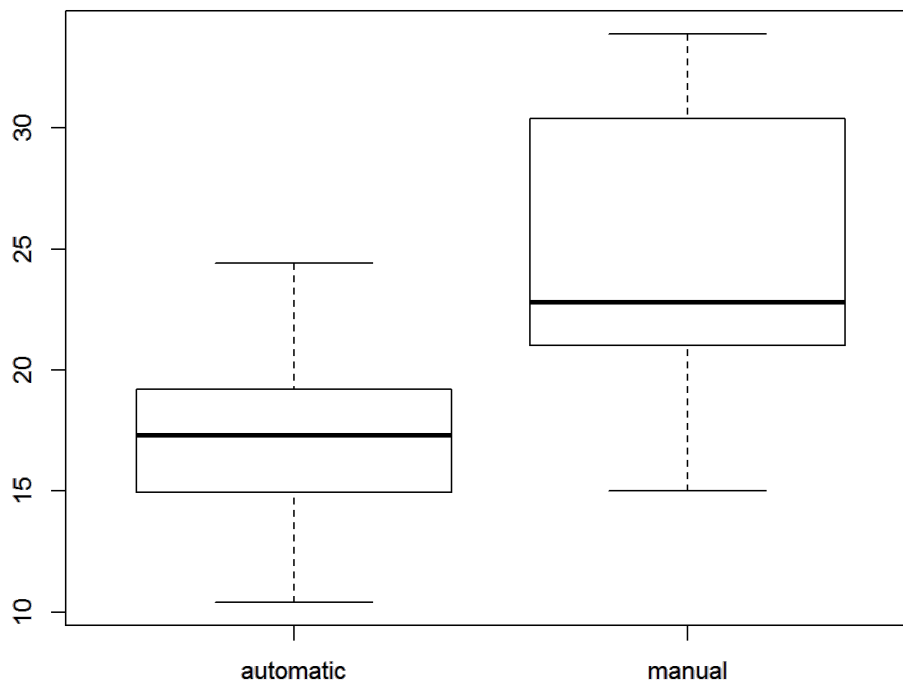
Fig. 2: Boxplot of mpg distribution
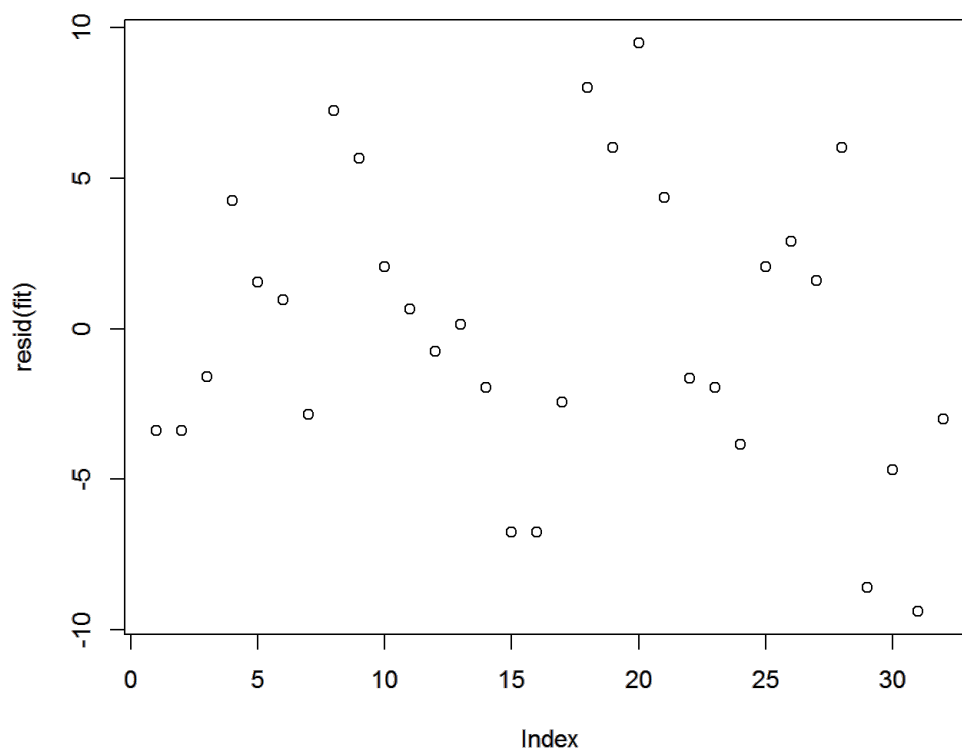
conditioned to transmission

```
plot(resid(fit))
```


Fig. 3: Plot of residuals of linear

model