# Manual or automatic transmission for better MPG?

Luis Fernando Agottani

03/09/2019

# Summary

The objective of this document is to compare the consume of fuel between automatic and manual transmission with de dataset mtcars. The variables of this data set are:

[, 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 (1000 lbs) [, 7] qsec 1/4 mile time [, 8] vs Engine (0 = V-shaped, 1 = straight) [, 9] am Transmission (0 = automatic, 1 = manual) [,10] gear Number of forward gears [,11] carb Number of carburetors

#Exploratory Data Analysis

```
head(mtcars)
```

```
##                   mpg cyl disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
## Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
## Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
## Valiant           18.1   6  225 105 2.76 3.460 20.22  1  0    3    1
```

For this analysis, will be compared all the variables to analyse which one has more influence in the value of mpg.

```
cor(mtcars)[1,]
```

```
##        mpg        cyl       disp         hp       drat         wt
##  1.0000000 -0.8521620 -0.8475514 -0.7761684  0.6811719 -0.8676594
##       qsec         vs         am       gear       carb
##  0.4186840  0.6640389  0.5998324  0.4802848 -0.5509251
```

We can see a strong correlation with mpg, cyl, disp, wt and hp.

Convert the transmission variable in factor “Automatic” and “Manual”.

```
mtcars$am   <- factor(mtcars$am,labels=c("Automatic","Manual"))
```

It is possible to conclude in APPENDIX A.1 that the manual cars have higher mpg values than automatic. Now it is important to verify if the tranmission is a key factor for the influence in mpg or other variable is more important.

Simple regression, just with transmission and mpg.

```
smr<- lm(mpg~am, mtcars)

summary(smr)
```

```
##
## 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 ***
## amManual       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
```

As we can see, the Manual transmission has higher mgp values, but, the R-squared indicates 0,3385, what says that the tramission only represents 33,85% of the mpg results.

```
amr<- lm(mpg~., mtcars)

summary(amr)
```

```
##
## Call:
## lm(formula = mpg ~ ., data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.4506 -1.6044 -0.1196  1.2193  4.6271
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 12.30337   18.71788   0.657   0.5181
## cyl         -0.11144    1.04502  -0.107   0.9161
## disp         0.01334    0.01786   0.747   0.4635
## hp          -0.02148    0.02177  -0.987   0.3350
## drat         0.78711    1.63537   0.481   0.6353
## wt          -3.71530    1.89441  -1.961   0.0633 .
## qsec         0.82104    0.73084   1.123   0.2739
## vs           0.31776    2.10451   0.151   0.8814
## amManual     2.52023    2.05665   1.225   0.2340
## gear         0.65541    1.49326   0.439   0.6652
## carb        -0.19942    0.82875  -0.241   0.8122
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.65 on 21 degrees of freedom
## Multiple R-squared:  0.869,  Adjusted R-squared:  0.8066
## F-statistic: 13.93 on 10 and 21 DF,  p-value: 3.793e-07
```

As we can see, there are strong correlations in value of P in mpg, disp, wt and hp.

Using these variables we can get a better fit.

```
bmr<- lm(mpg~am + cyl + disp + wt + hp, mtcars)

summary(bmr)
```

```
##
## Call:
## lm(formula = mpg ~ am + cyl + disp + wt + hp, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.5952 -1.5864 -0.7157  1.2821  5.5725
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 38.20280    3.66910  10.412 9.08e-11 ***
## amManual     1.55649    1.44054   1.080  0.28984
## cyl         -1.10638    0.67636  -1.636  0.11393
## disp         0.01226    0.01171   1.047  0.30472
## wt          -3.30262    1.13364  -2.913  0.00726 **
## hp          -0.02796    0.01392  -2.008  0.05510 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.505 on 26 degrees of freedom
## Multiple R-squared:  0.8551, Adjusted R-squared:  0.8273
## F-statistic:  30.7 on 5 and 26 DF,  p-value: 4.029e-10
```

This is a better fit for, with R-squared equal to 82,73%.

In APPENDIX A.2 is possible to analyse the correlation between the variables in the best fit model.

In APPENDIX A.3 is possible to analyse the residuals of the best fit model.

# CONCLUSION

# Is an automatic or manual transmission better for MPG?

The manual transmission usually have a higher mpg value than automatic, but the type of the transmission isn't the most influent variable in mpg value, the cyl, disp, hp, and wt have a stronger correlation.

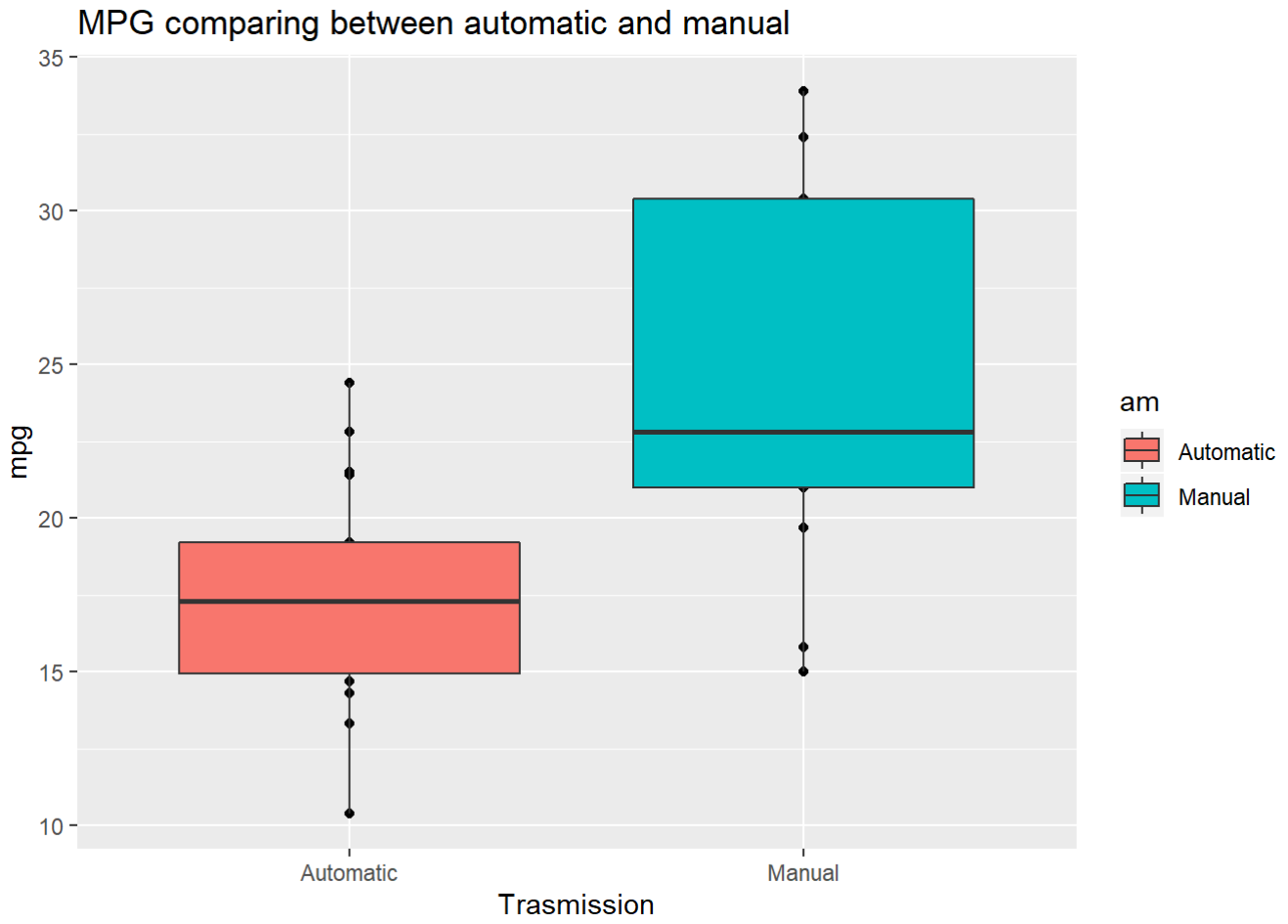
# Quantify the MPG difference between automatic and manual transmissions?

Analysis demonstrate that with the manual transmission the mpg value increses 7.245. But, when included more variable in the comparassion, cyl, disp, hp and wt, the value decrease for 1,55, and shows us that the other variables has more impact, like wt, with -3,30.

# APPENDIX

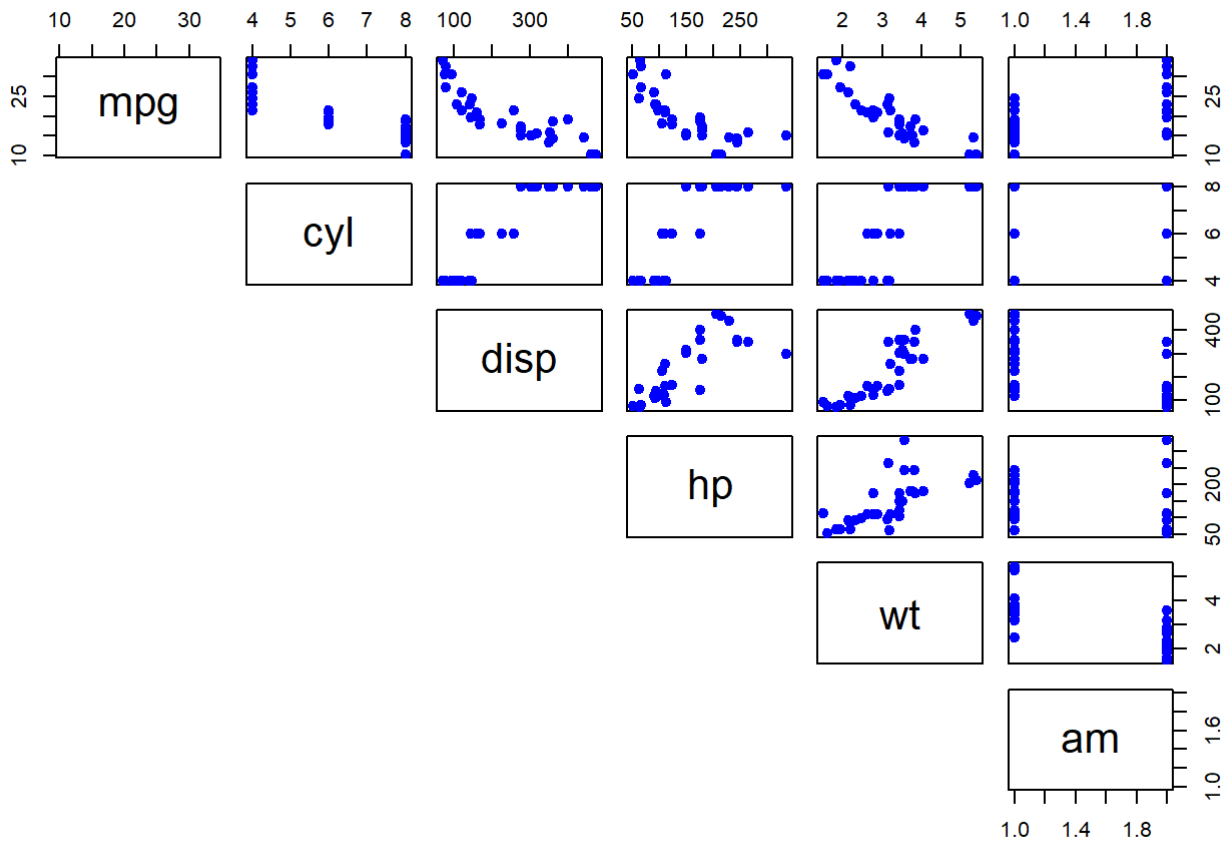
A.1 - Boxplot to compare mpg with Automatic and Manual transmission.

```
library(ggplot2)

g<- qplot(am, mpg, data=mtcars, main="MPG comparing between automatic and manual",xlab="Trasm
ission", ylab="mpg")

g+geom_boxplot(aes(fill = am))
```



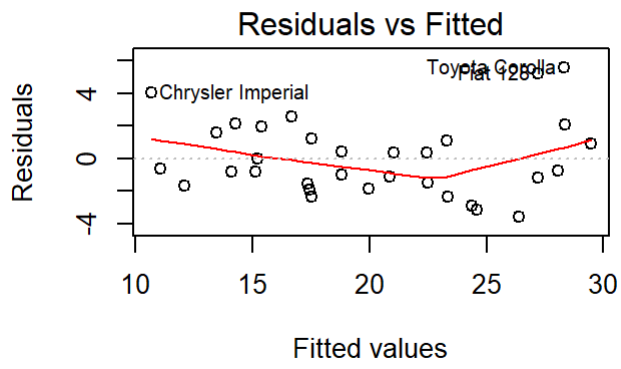

A.2 - Correlation table.

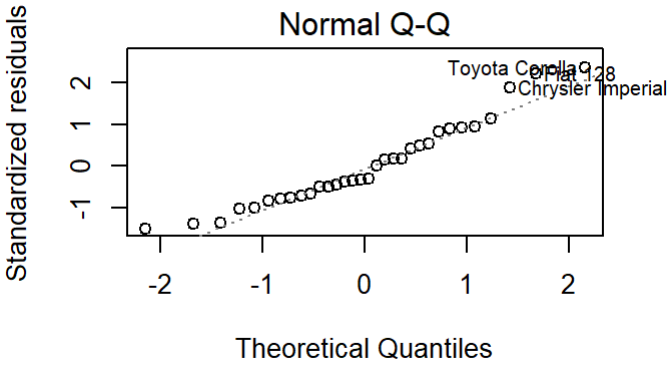
```
pairs(mtcars[,c(1,2,3,4,6,9)], pch = 19, col = "blue", lower.panel=NULL)
```

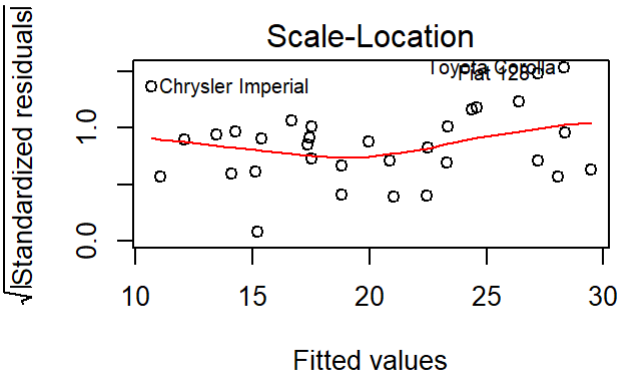


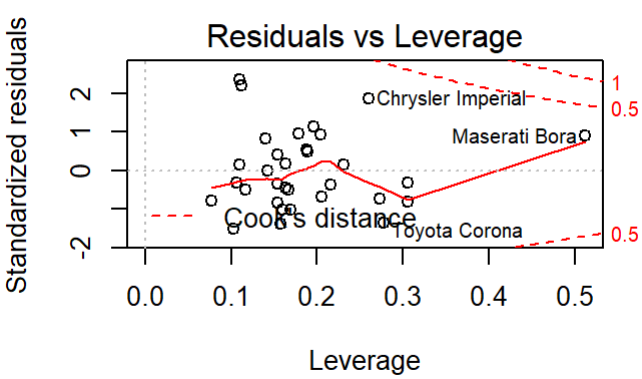

A.3 - Residual plot.

```
par(mfrow = c(2, 2))
plot(bmr)
```


Residuals vs Fitted
Residuals
Fitted values
Chrysler Imperial
Toyota Corolla
Fiat 128
10
15
20
25
30
-4
0
4


Normal Q-Q
Standardized residuals
Theoretical Quantiles
Toyota Corolla
Fiat 128
Chrysler Imperial
-2
-1
0
1
2


Scale-Location
√|Standardized residuals|
Fitted values
Chrysler Imperial
Toyota Corolla
Fiat 128
10
15
20
25
30
0.0
1.0


Residuals vs Leverage
Standardized residuals
Leverage
Chrysler Imperial
Maserati Bora
Toyota Corona
Cook's distance
1
0.5
0.5
0.0
0.1
0.2
0.3
0.4
0.5