

Asude Berber - Regression Model Course - Project Assignment

Asude Berber

2023-09-04

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
#cars examples skip
```

```
#Starting....
```

```
#Executive Summary for my project and purposes: I will examine the mtcars data and collection cars. For
```

```
#First I will process the data and make summary
```

```
#There are two groups: Automatic and Manual ; so 0 = automatic, 1 = manual
```

```
library(ggplot2)
```

```
library(dplyr)
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
## filter, lag
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
## intersect, setdiff, setequal, union
```

```
library(ggfortify)
```

```
data(mtcars)
```

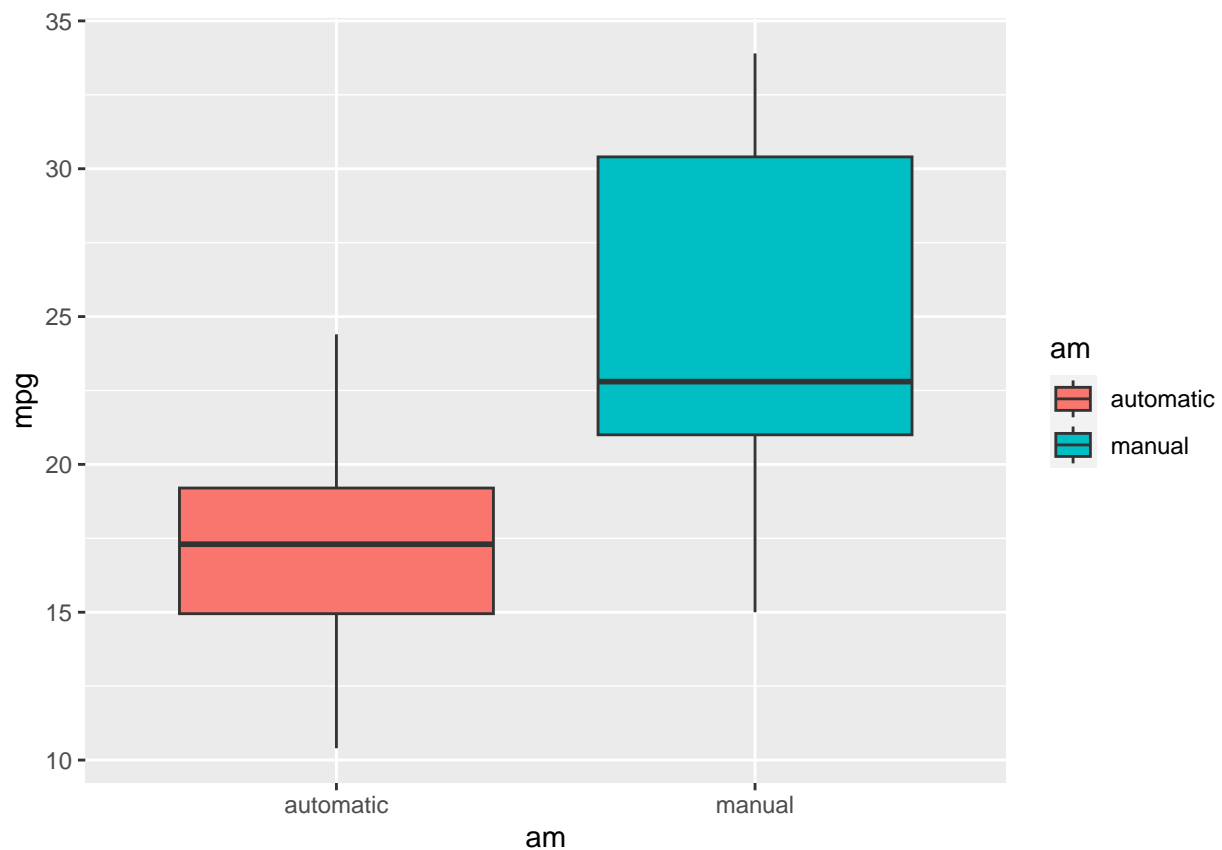
```
dim(mtcars)
```

```
## [1] 32 11
```

```
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
```

```
library(ggplot2)
mtcars$am <- as.factor(mtcars$am)
levels(mtcars$am) <- c("automatic", "manual")
p <- ggplot(mtcars, aes(am, mpg))
p + geom_boxplot(aes(fill = am))
```



#here you can see that manual has higher mpg usage compare to the automatic

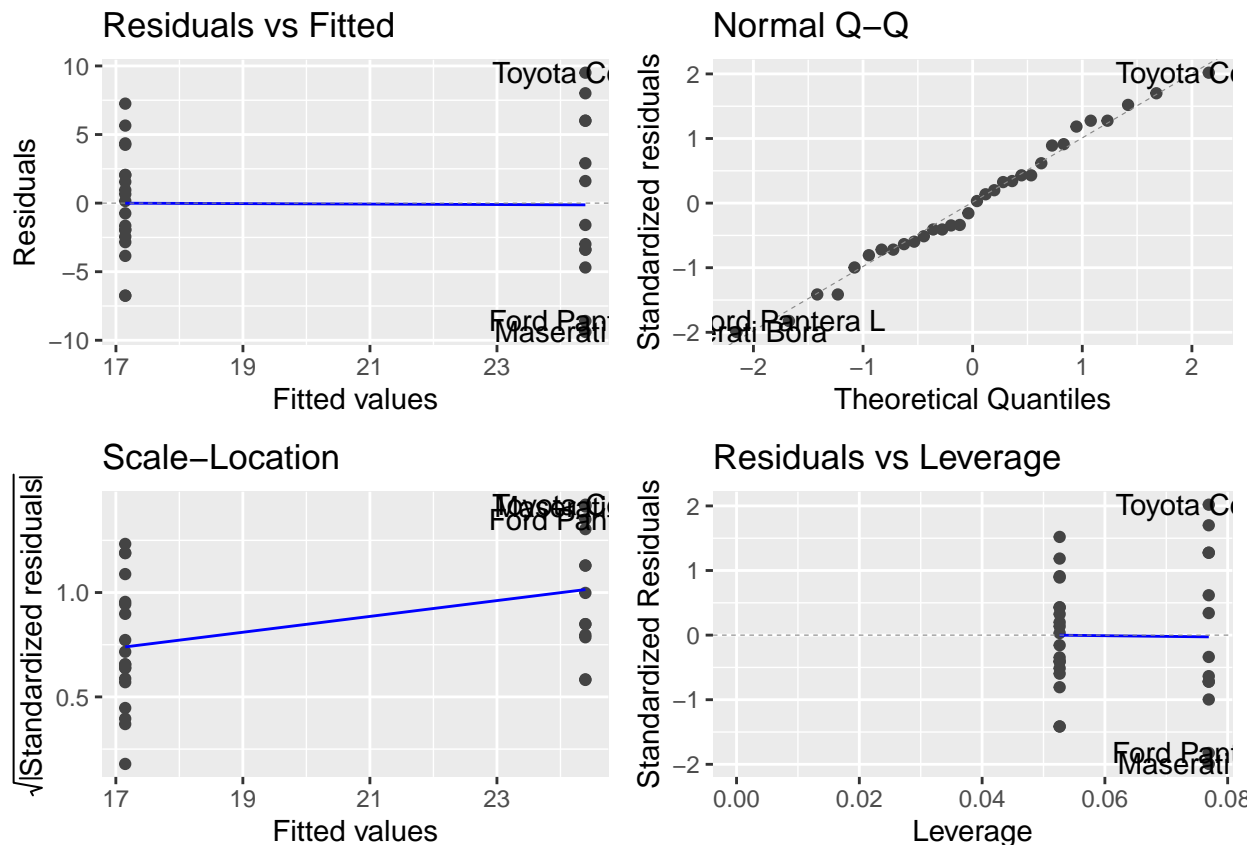
#Now I will look at the correlation of mpg cars

#for model selection

```
#I will create linear model
lmfit <- lm(mpg ~ am, mtcars)
summary(lmfit)
```

```
##
## 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
```

```
#p value is so low , std error 1 is not okay.
autoplot(lmfit)
```



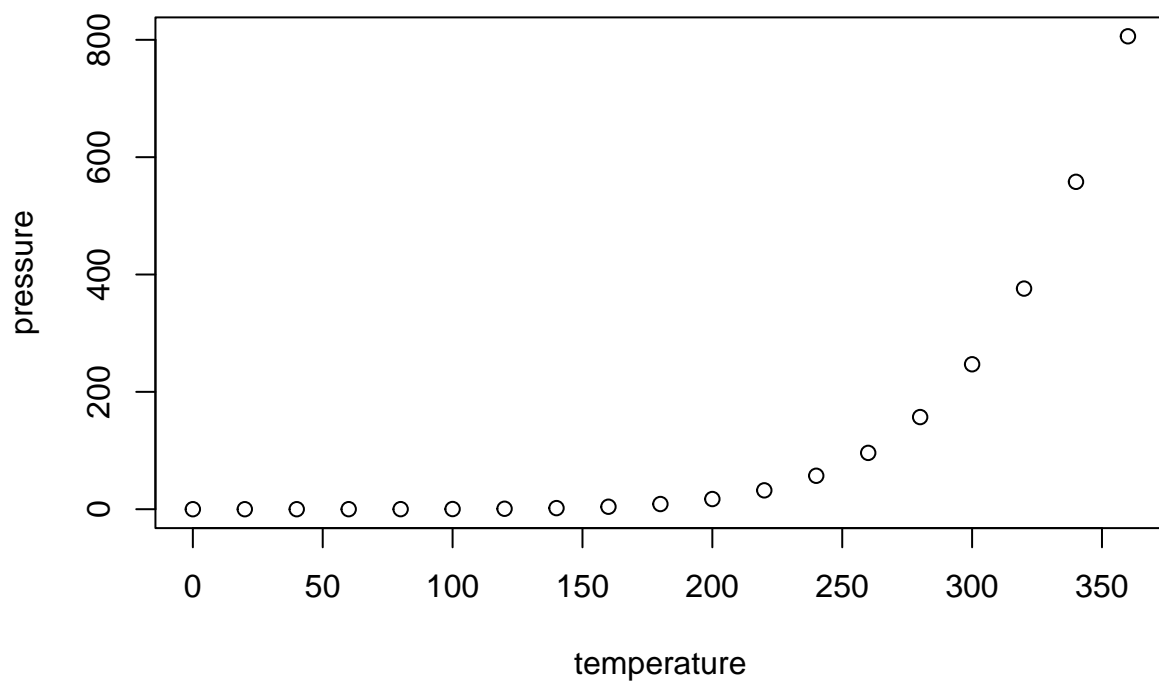
```
#only normal plot seems okay
```

```
#So for the answer , definitely automatic is better for less usage of MPG
```

```
#for quantification between automatic and manual -> R squared 0.3598 . So low so for our model is not g
```

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.