RMReport

DrMuhsin

Monday, September 21, 2015

Executive Summary

Using the Mtcars data (gathered by Moto Trends), the correlation between miles per gallon of automobiles with car transmission will be analyzed and observed. Among the key question in this report are:

- Which one between an automatic or manual transmission have better MPG?
- What is the MPG difference between automatic and manual transmissions?

Summary of data

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). The qualitative variables such as number of cylinders and gears were converted to factors are as follow (refer to Appendix):

```
data(mtcars)
library(knitr)

## Warning: package 'knitr' was built under R version 3.1.3

opts_chunk$set(echo = FALSE)
opts_chunk$set(fig.width = 5)
```

Exploratory analysis

Using the boxplot, the differences between automatic and manual transmission in terms of MPG has been plotted. A clear indication that manual transmission produces more mpg compared to automatic transmission.

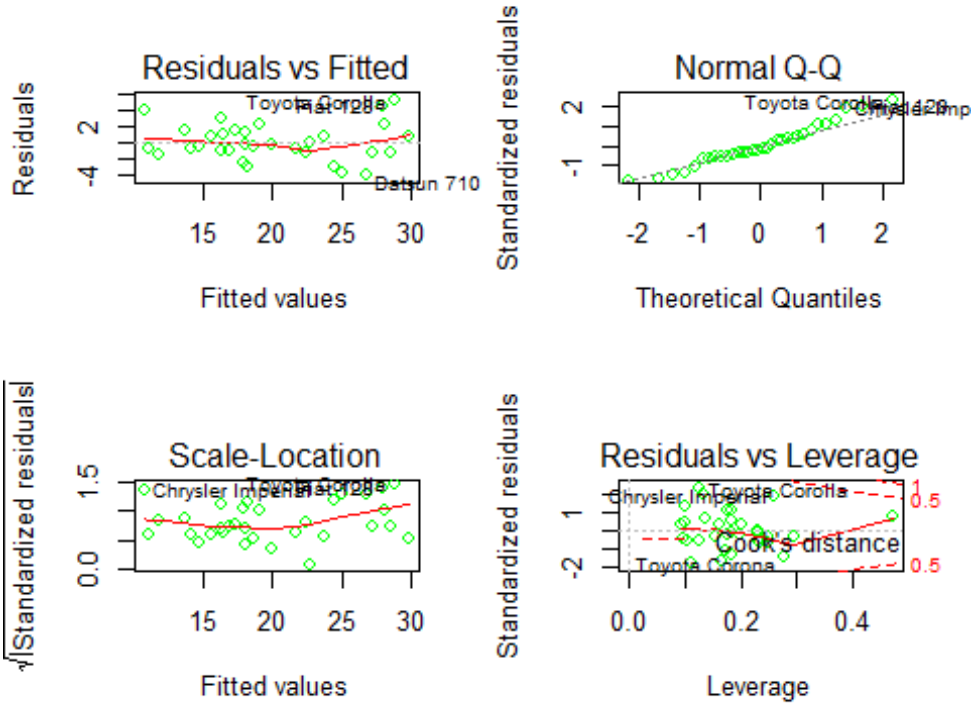
Figure 2 shows pairwise graph that indicates other variables that might have a correlation with MPG. There is a linear relationship between MPG and each of cyl, disp, hp, drat, wt, qsec, vs, am. The covariance was also computed (Table 1) between every variable and the positive values were noted (qsec = 0.419, vs = 0.664, am = 0.600, gear = 0.480). Then a linear model was fit on all the variables to determine which variables should be used in the final models. In Table 2 the summary from this model is shown. The lowest p values were taken (i.e. wt = 0.063, am = 0.234, qsec = 0.274) due to their high significance in predicting MPG.

Model

From the initial model, covariance test and visually inspecting the pairwise graph the following variables stood out in particular: qsec, vs, am, wt and gear. Next a stepwise model process was used in order to obtain the most significant predictors to be used. This is done by using the step function which creates multple regression models with different variables and produces list of the best predictors. As shown in figure 5, the most significant predictors in determining the MPG are cyl, hp, wt and am. The summary for this model is show in Table 3, in particular the forumla is given as: lm(formula = mpg ~ cyl + hp + wt + am, data = mtcars). This selection model yielded an R squared value of 84% (Table 4) meaning that very high percentage of variation is explained by the regression model. Next, the new model was compared with a basic model that only uses transmission type as its predictor. A p-value of 1.688e-08 was obtained (Table 5). This value is miniscule which means that the added predictors are significant to improving the model's accuracy.

Diagnostics

The residuals from the final model are plotted below.



- The Residuals vs Fitted plot shows no pattern between the residuals and fitted values indicating that this regression model is well fit.
- The QQ plot shows that the points line up as expected meaning that the distribtion is normal and our model predictions are accurate.

- In both the Scale-Location plot and the Residuals vs Leverage plots, the points are in a group with none too far from the center indicating no point had too much leverage.

Statistical Inference

A Two Sample t-test was conducted between the different transmission types. The null hypothesis that transmission types don't have an effect on the MPG is discarded for a p-value greater than 0.05. The results are shown in table 6. The p-value of 0.001374 and difference in means show that manual transmission has significantly more MPG than automatic.

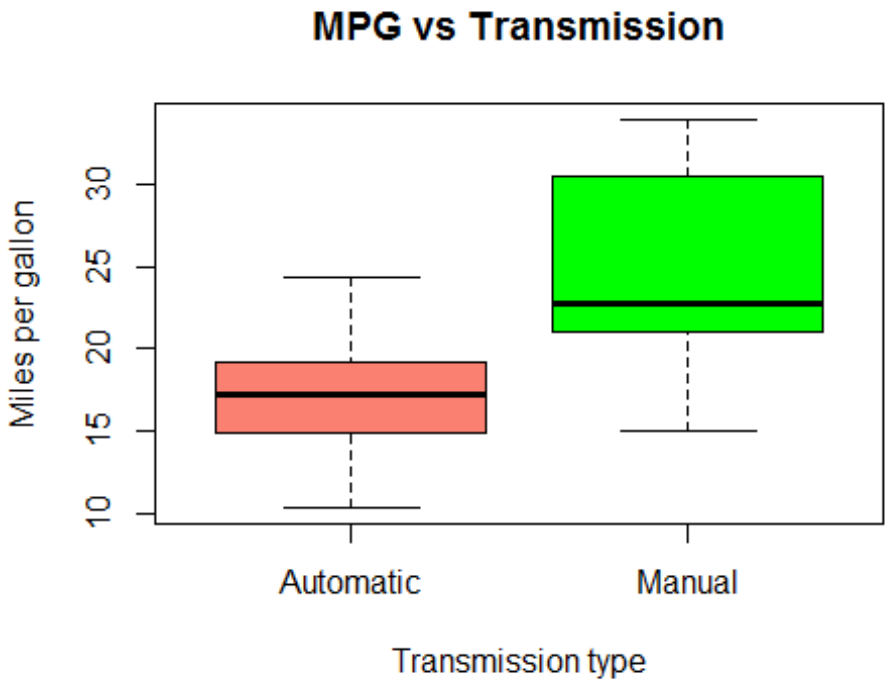
Conclusions

Interestingly the type of transmission of a car have a significant effect on its fuel consumption. Manual transmission have a better miles per gallon (1.81MPG) more than automatic transmission on average.

Appendix

Description of variables

- mpg Miles/(US) gallon
- cyl Number of cylinders
- disp Displacement (cu.in.)
- hp Gross horsepower
- drat Rear axle ratio
- wt Weight (lb/1000)
- qsec Time to drive ¼ mile
- vs V or ordinary engine
- am Transmission (0 = automatic, 1 = manual)
- gear Number of forward gears
- carb Number of carburetors



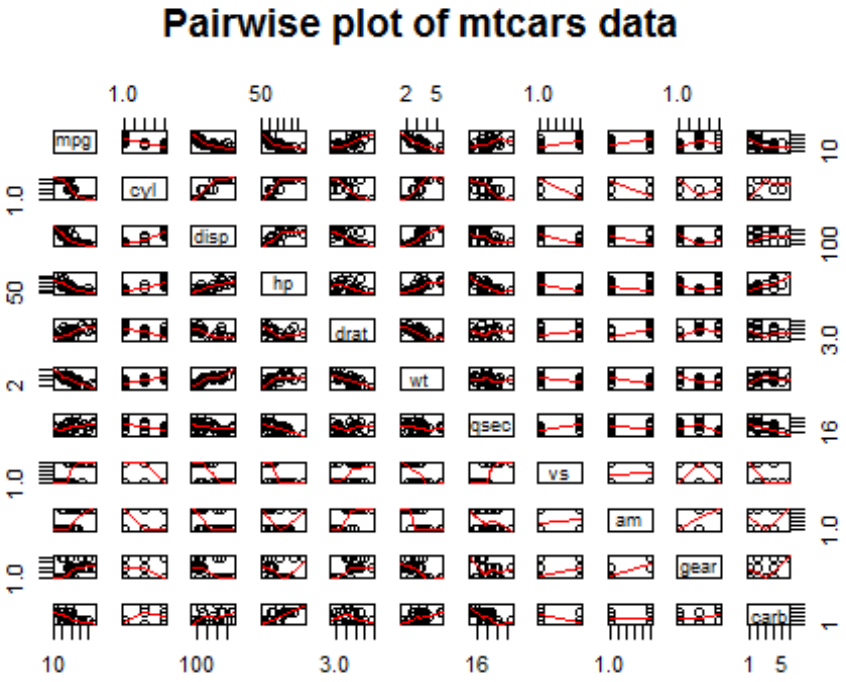


Table 1

```
##      mpg       cyl       disp         hp      drat         wt      qsec
## mpg    1 -0.852162 -0.8475514 -0.7761684 0.6811719 -0.8676594 0.418684
```

```
##            vs        am      gear       carb
## mpg 0.6640389 0.5998324 0.4802848 -0.6067431
```

Table 2

```
## (Intercept)        cyl6        cyl8        disp          hp        drat
## 23.87913244 -2.64869528 -0.33616298  0.03554632 -0.07050683  1.18283018
##          wt        qsec         vs1    amManual       gear4       gear5
## -4.52977584  0.36784482  1.93085054  1.21211570  1.11435494  2.52839599
##       carb2       carb3       carb4       carb6       carb8
## -0.97935432  2.99963875  1.09142288  4.47756921  7.25041126
```

Table 3

```
##                Estimate Std. Error   t value     Pr(>|t|)
## (Intercept) 33.70832390 2.60488618 12.940421 7.733392e-13
## cyl6        -3.03134449 1.40728351 -2.154040 4.068272e-02
## cyl8        -2.16367532 2.28425172 -0.947214 3.522509e-01
## hp          -0.03210943 0.01369257 -2.345025 2.693461e-02
## wt          -2.49682942 0.88558779 -2.819404 9.081408e-03
## amManual     1.80921138 1.39630450  1.295714 2.064597e-01
```

Table 4

```
## (Intercept)        cyl6        cyl8          hp          wt    amManual
## 33.70832390 -3.03134449 -2.16367532 -0.03210943 -2.49682942  1.80921138
```

Table 5

```
## [1]           NA 1.688435e-08
```

Table 6

```
##
##  Welch Two Sample t-test
##
## data:  mpg by 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
```