Motor Trend

Hugo Janssen

Monday, July 13, 2015

You work for Motor Trend, a magazine about the automobile industry. Looking at a data set of a collection of cars, they are interested in exploring the relationship between a set of variables and miles per gallon (MPG) (outcome). They are particularly interested in the following two questions:

"Is an automatic or manual transmission better for MPG" "Quantify the MPG difference between automatic and manual transmissions"

Question

Take the mtcars data set and write up an analysis to answer their question using regression models and exploratory data analyses.

Your report must be:

Written as a PDF printout of a compiled (using knitr) R markdown document. Brief. Roughly the equivalent of 2 pages or less for the main text. Supporting figures in an appendix can be included up to 5 total pages including the 2 for the main report. The appendix can only include figures. Include a first paragraph executive summary.

Did the student do some exploratory data analyses? Did the student fit multiple models and detail their strategy for model selection? Did the student answer the questions of interest or detail why the question(s) is (are) not answerable? Did the student do a residual plot and some diagnostics? Did the student quantify the uncertainty in their conclusions and/or perform an inference correctly?

Executive Summary

For figures, please refer to Appendix 1: Figures.

Exploratory Analysis

Sneak peek:

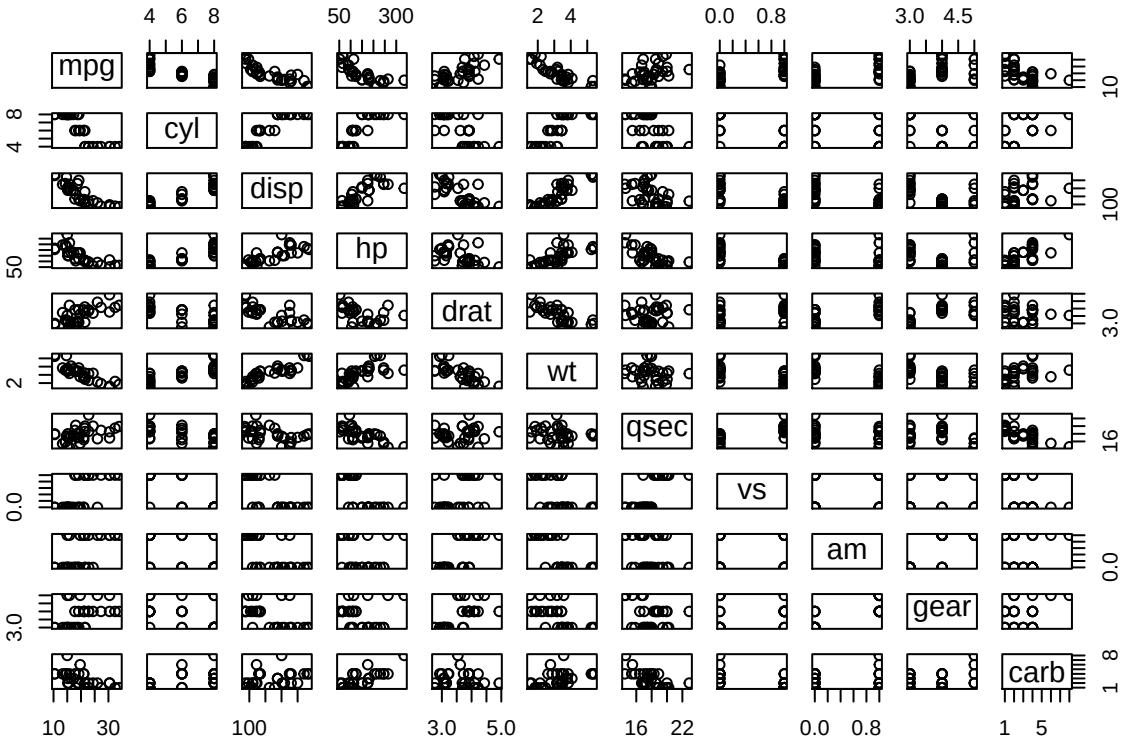
```
head(mtcars, 2)
```

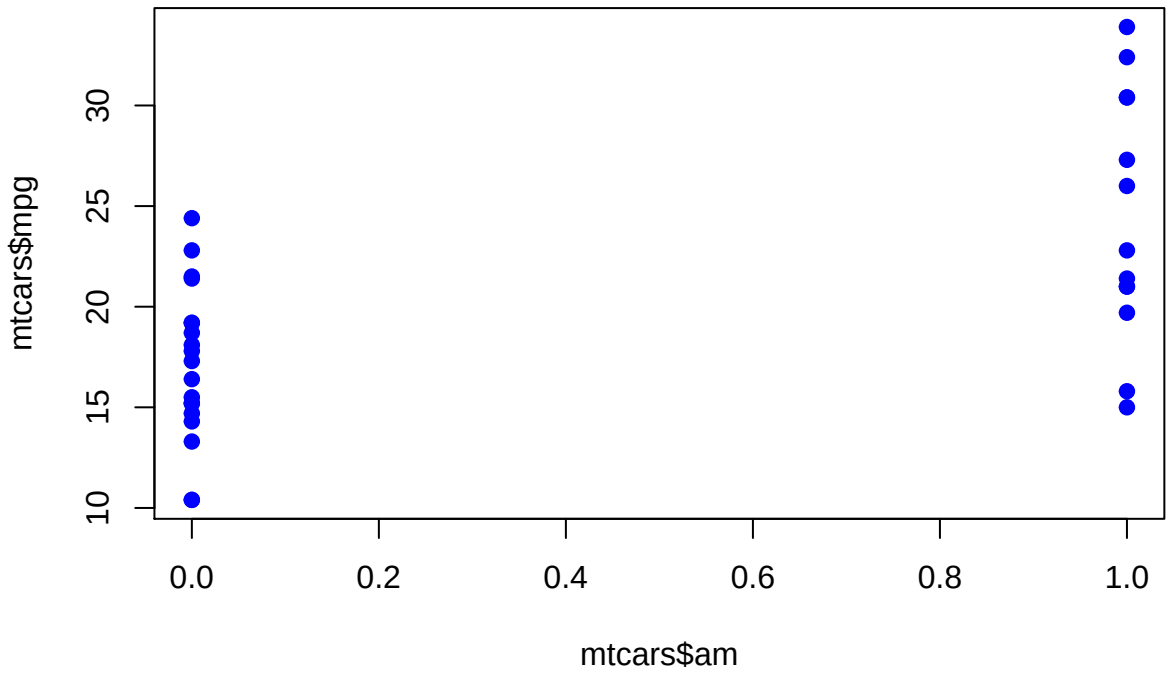
```
##               mpg cyl disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4      21   6  160 110  3.9 2.620 16.46  0  1    4    4
## Mazda RX4 Wag  21   6  160 110  3.9 2.875 17.02  0  1    4    4
```

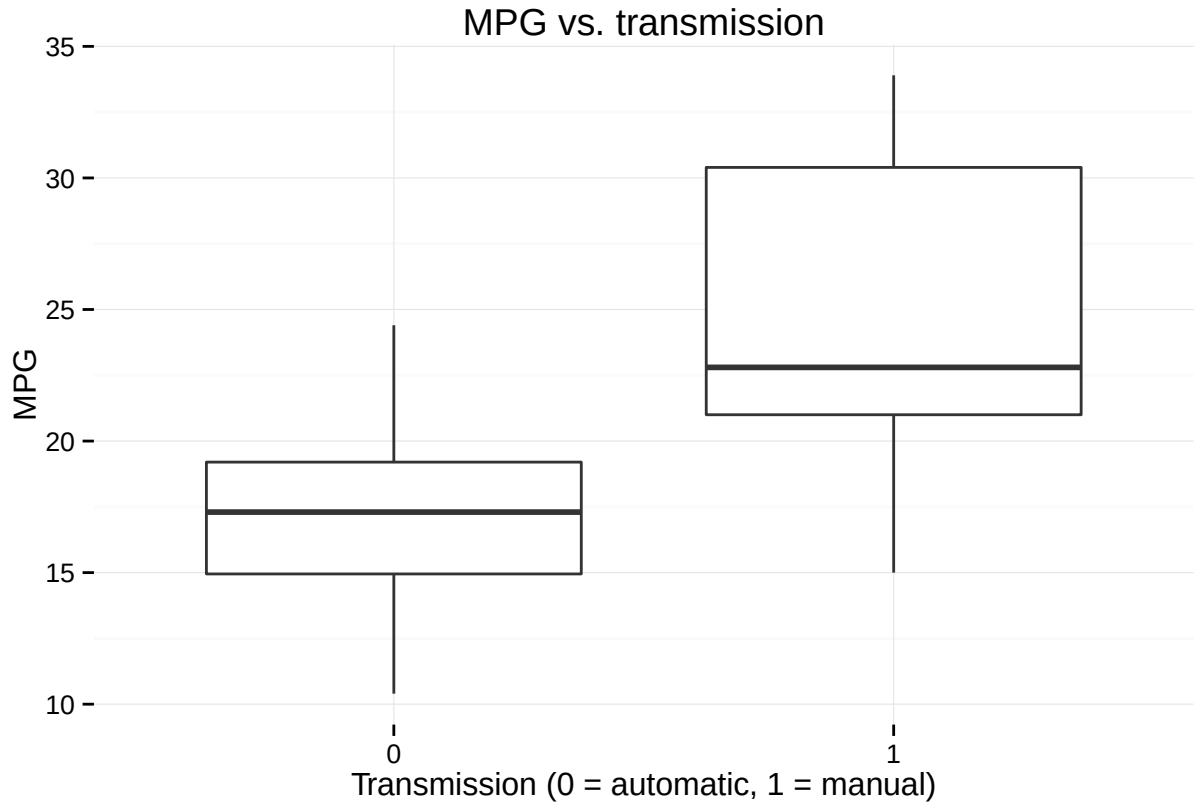
See [figure 1][001]

Conclusion

You can also embed plots, for example:


mpg
cyl
disp
hp
drat
wt
qsec
vs
am
gear
carb


mtcars$mpg
mtcars$am
10
15
20
25
30
0.0
0.2
0.4
0.6
0.8
1.0

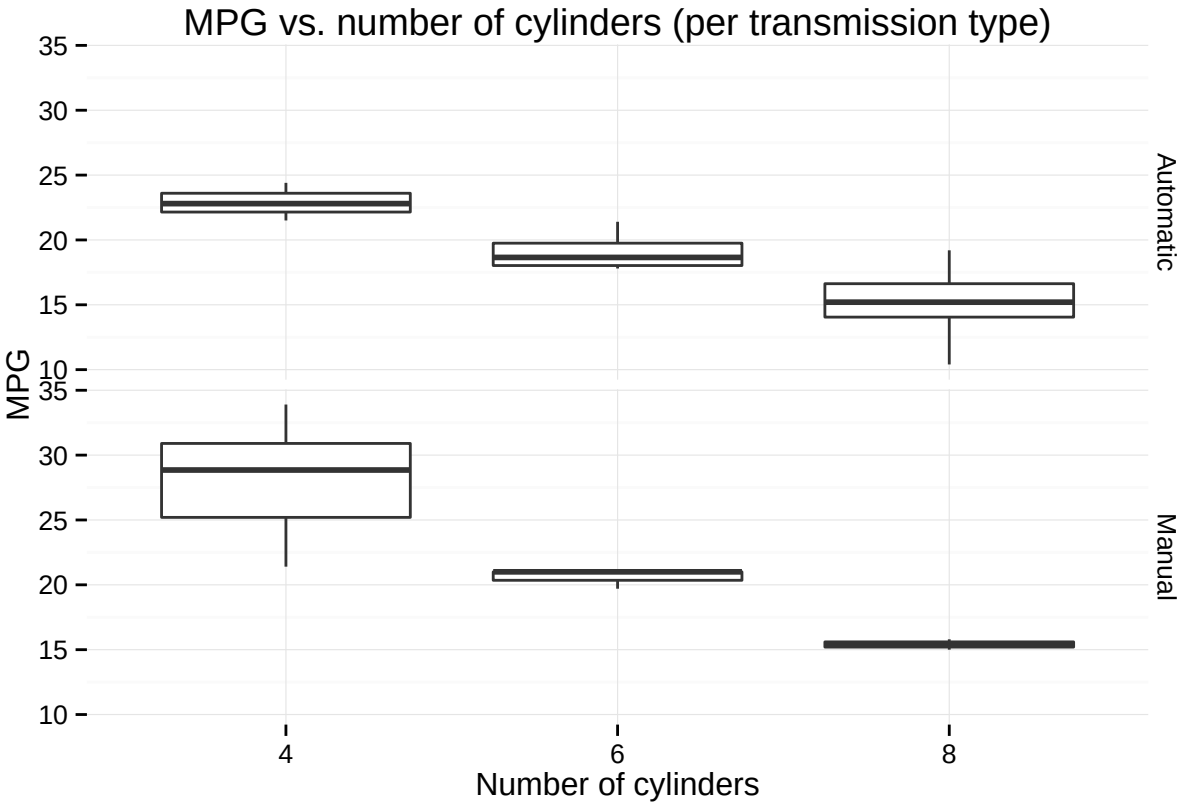


Appendix 1: Figures

[001]Figure 1. Boxplot of MPG vs. number of cylinders per transmission type

```
# Prepare vector for labelling panels
labels <- c("0" = "Automatic", "1" = "Manual")

ggplot(mtcars, aes(x=factor(cyl), y=mpg)) +
    facet_grid(am ~ ., labeller = labeller(am = labels)) +
    geom_boxplot() +
    labs(title = "MPG vs. number of cylinders (per transmission type)") +
    labs(x = "Number of cylinders") +
    labs(y = "MPG") +
    theme_minimal(base_family = "")
```


MPG vs. number of cylinders (per transmission type)
MPG
35
30
25
20
15
10
Automatic
Manual
4
6
8
Number of cylinders