

Regression Models Peer Assessment

Emanuel Calvo

November 21, 2015

Context

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

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.

Executive Summary

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). This dataset is included in the `datasets` library.

Exploratory data analysis

We can observe that the transmission systems come from different populations, as the p-value over a t-test give us 0.0013736 between Miles per Galon and Transmissions. Also the mean of each group has 7.2449393 of difference.

All the correlations against mpg:

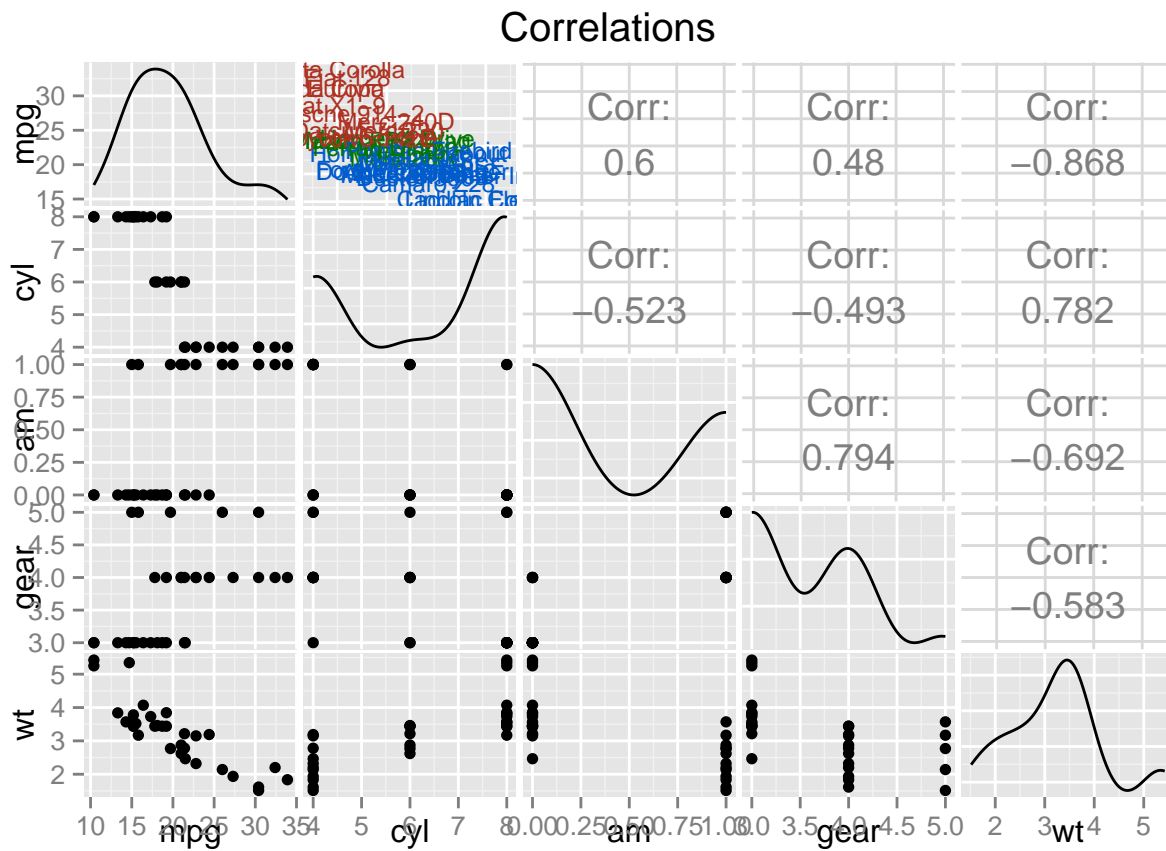
```
cor(mtcars)[1,]
```

```
##      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
```

Let's consider the following figure with the correlations among a selected set of columns from the dataset:

```
g <- ggpairs(mtcars[,c("mpg","cyl","am", "gear","wt")], title = "Correlations" )
plot <- ggplot2::ggplot(mtcars, ggplot2::aes(x=wt, y=mpg, label=rownames(mtcars)))
plot <- plot +
  ggplot2::geom_text(ggplot2::aes(colour=factor(cyl)), size = 3) +
  ggplot2::scale_colour_discrete(l=40)
g <- putPlot(g, plot, 1, 2)
```

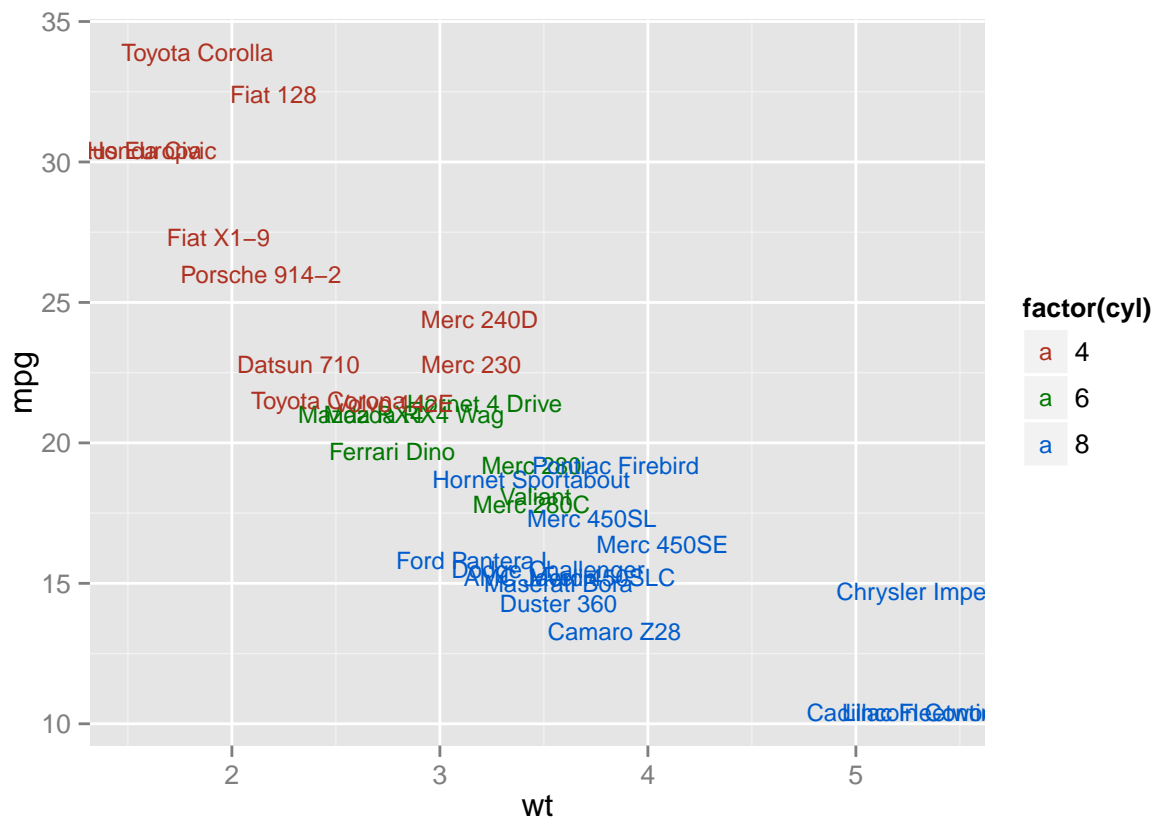
g



```
Model <- lm(mpg ~ ., data=mtcars)
```

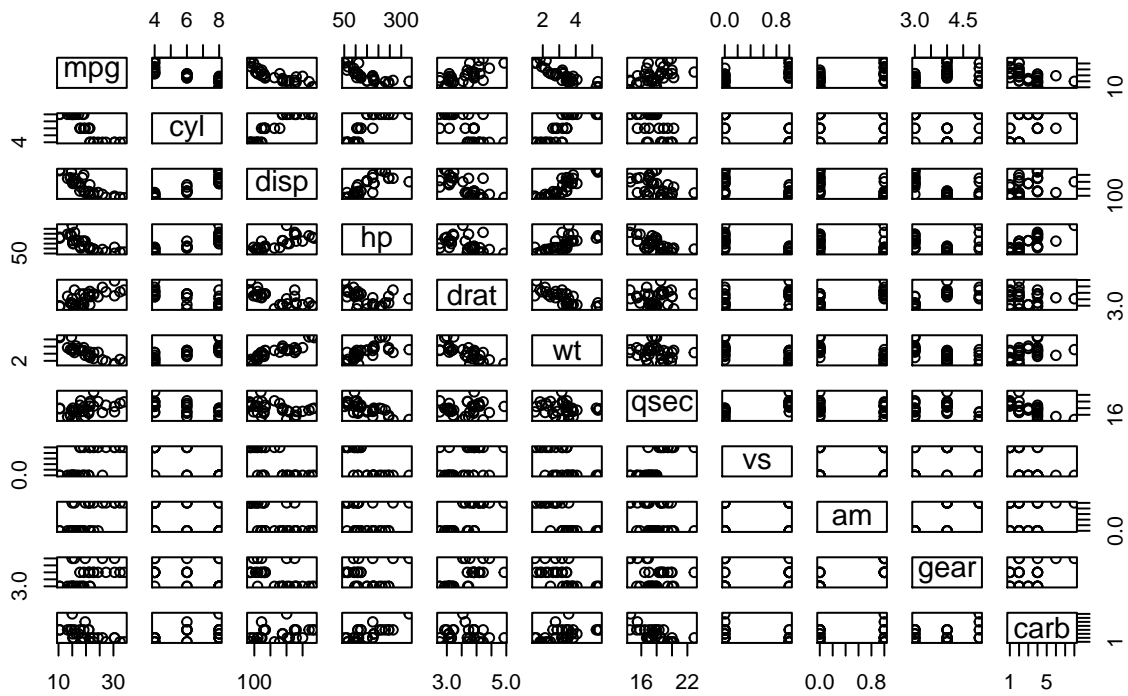
```
g2 <- ggplot2::ggplot(mtcars, ggplot2::aes(x=wt, y=mpg, label=rownames(mtcars))) +
  ggplot2::geom_text(ggplot2::aes(colour=factor(cyl)), size = 3) +
  ggplot2::scale_colour_discrete(l=40)
```

g2



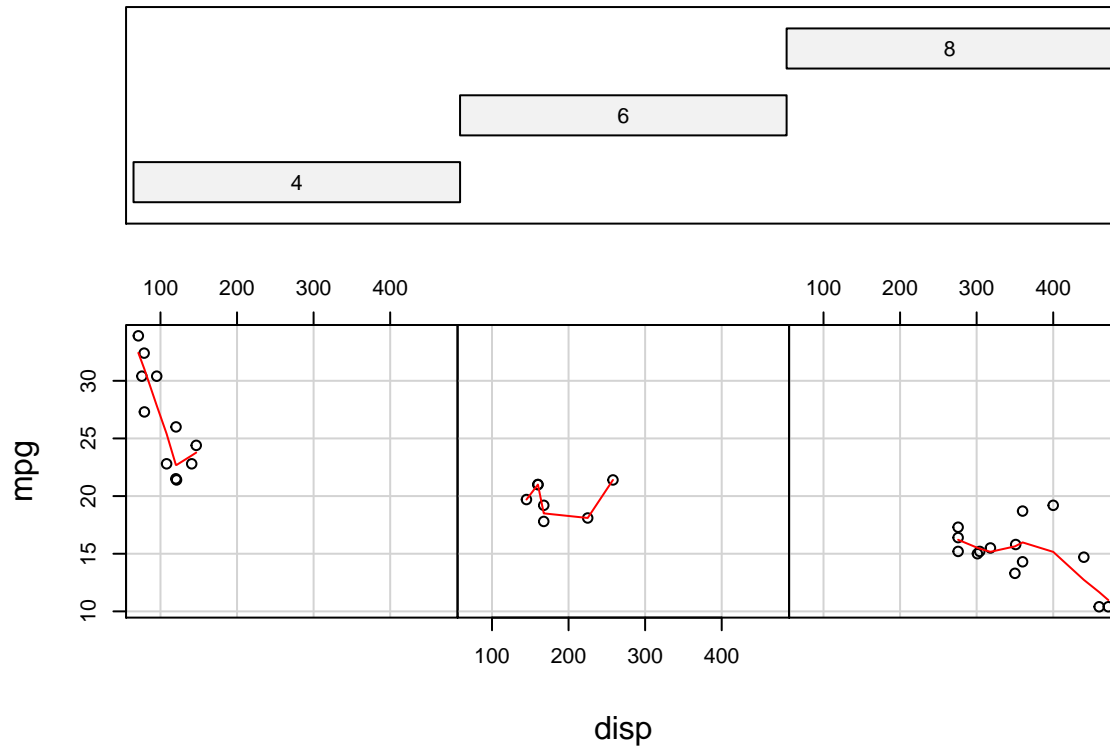
```
pairs(mtcars, main = "mtcars data")
```

mtcars data

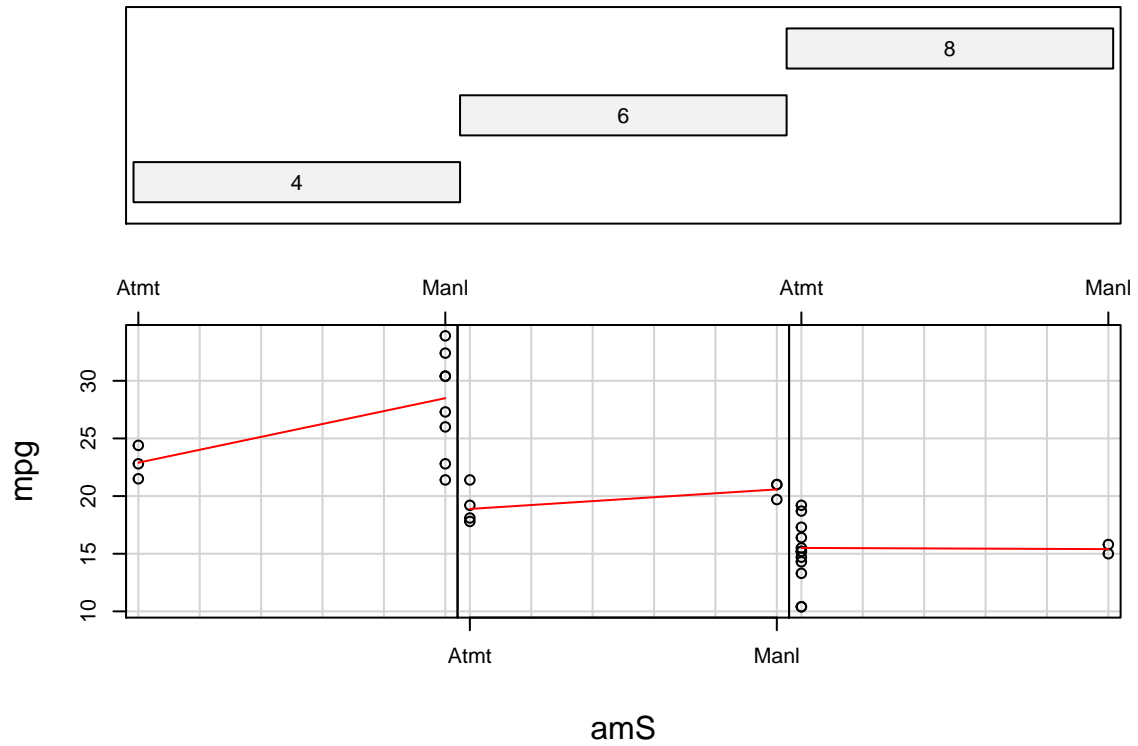


```
coplot(mpg ~ disp | as.factor(cyl), data = mtcars,
       panel = panel.smooth, rows = 1)
```

Given : `as.factor(cyl)`

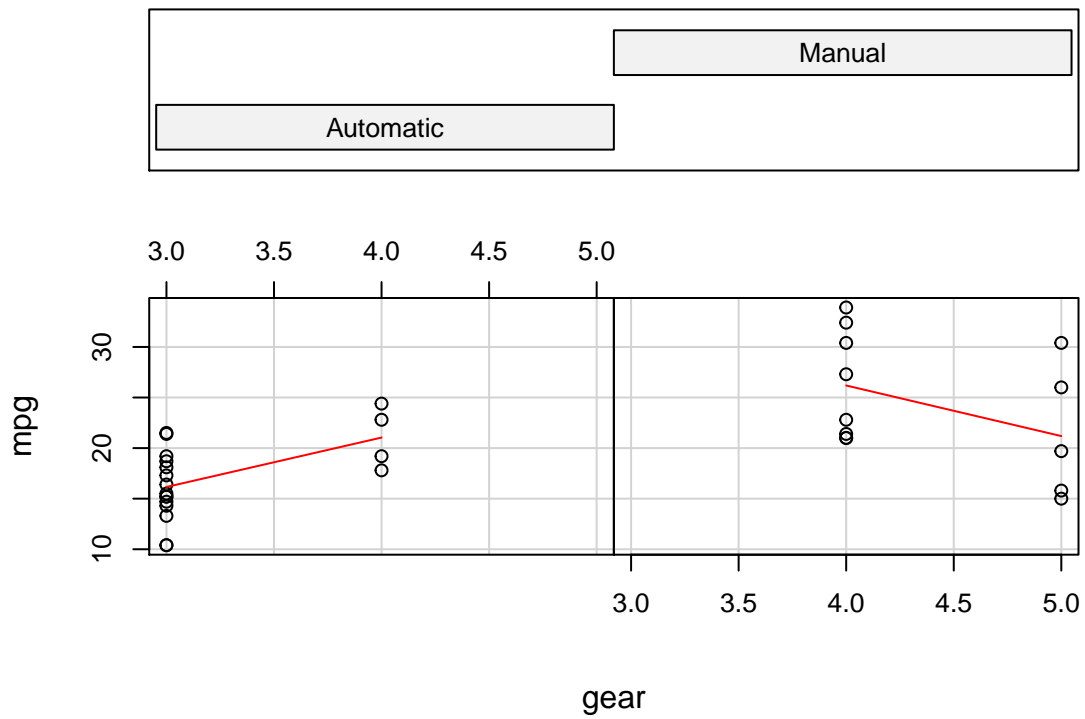


Given : `as.factor(cyl)`

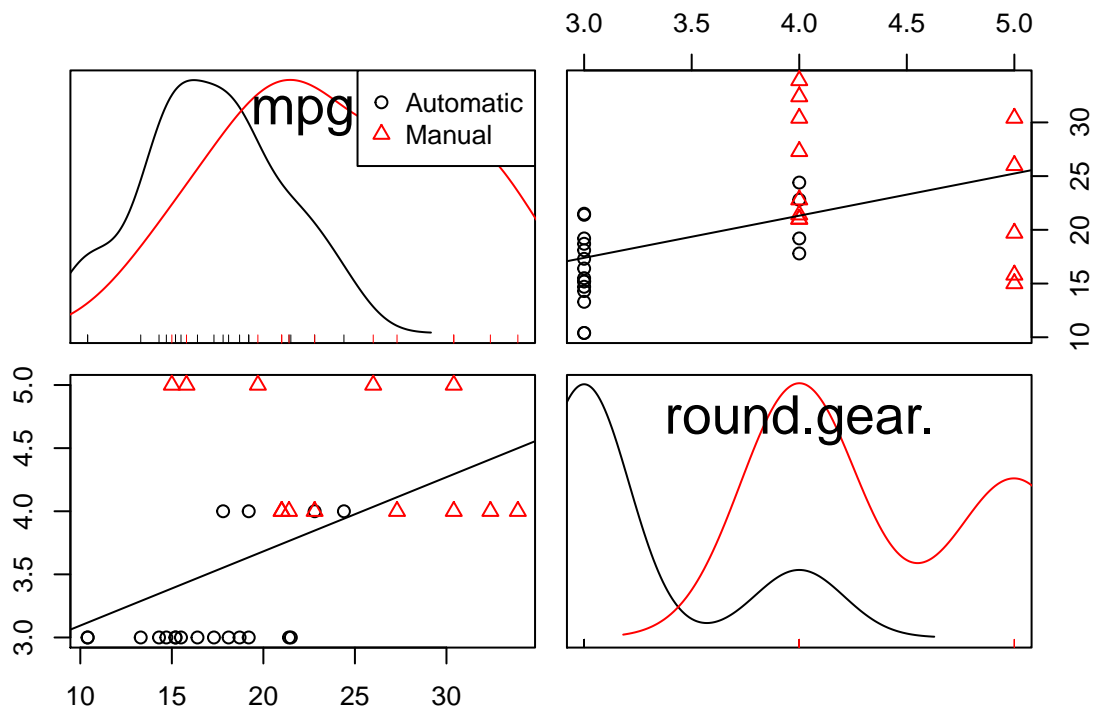


```
coplot(mpg ~ gear | amS, data = mtcars,  
       panel = panel.smooth, rows = 1)
```

Given : amS



```
spm(~ mpg + round(gear) | amS, data = mtcars, smoother=FALSE)
```



Quantifying the MPG difference between automatic and manual transmissions