Untitled

Ion Scerbatiuc

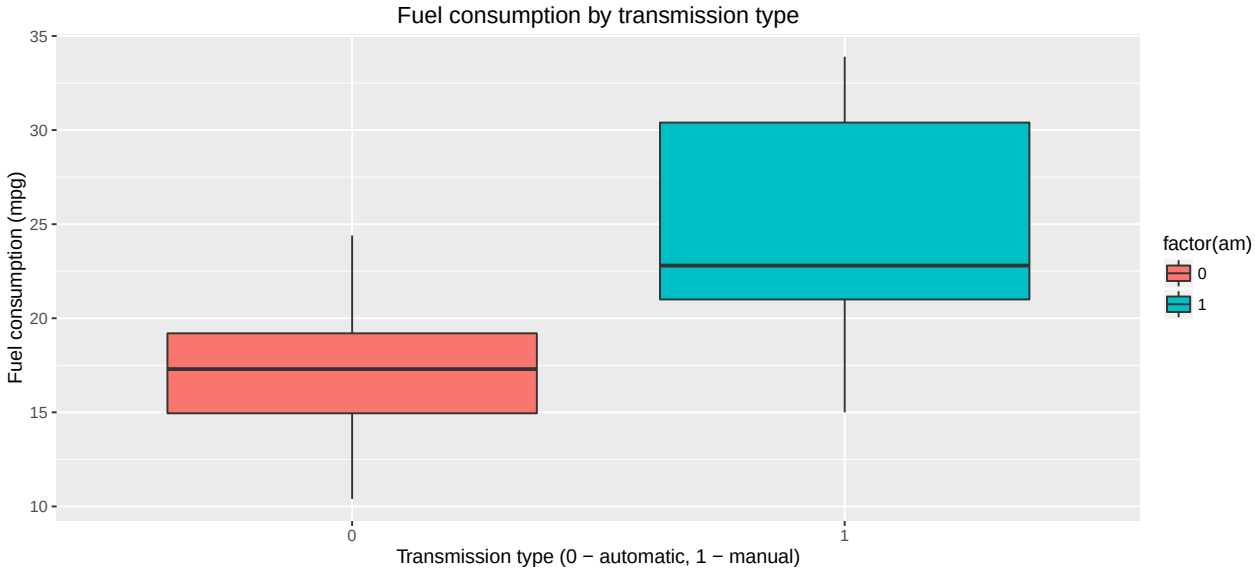
February 25, 2016

TODO: Executive Summary http://www.businesswritingblog.com/business_writing/2013/05/write-better-executive-summaries.html http://unilearning.uow.edu.au/report/4bi1.html

```
library(datasets)
library(ggplot2)
library(dplyr)
library(xtable)
options(xtable.comment = FALSE)
data(mtcars)
```

Let's first explore what is the measured fuel consumption by transmission type.

```
ggplot(mtcars, aes(x = factor(am), y = mpg, fill = factor(am))) +
    geom_boxplot() +
    xlab("Transmission type (0 - automatic, 1 - manual)") +
    ylab("Fuel consumption (mpg)") +
    ggtitle('Fuel consumption by transmission type')
```



From the plot above we can conclude that automatic cars are consuming on average more fuel than manual transmission cars. But is that significant? Let's dig deeper to find that out.

Let's fit a linear model whit mpg as the outcome and the transmission type as a factor variable.

```
fit <- lm(mpg ~ factor(am), mtcars)
print(xtable(summary(fit)$coefficients), include.rownames=FALSE, sanitize.text=identity)
```

By inspecting the coefficients table we can conclude that manual cars have an expected mileage `7.25` mpg higher than the expected mileage of automatic cars. The result is also statistically significant because the p-value for is lower than `0.05`.

Estimate	Std. Error	t value	Pr(>|t|)
17.15	1.12	15.25	0.00
7.24	1.76	4.11	0.00

The result is suprising. Why would automatic cars consume more fuel than manual cars? Are there any variables that we are missing from the model? Knowing how cars work I would suspect the mileage to be corelated to the number of cylinders (engine size) and the weight of the car.