

Motor Trend MPG Analysis

Michael Krämer

25.10.2015

Abstract

The report shows changes on mpg for cars, especially the influence of the transmission type.

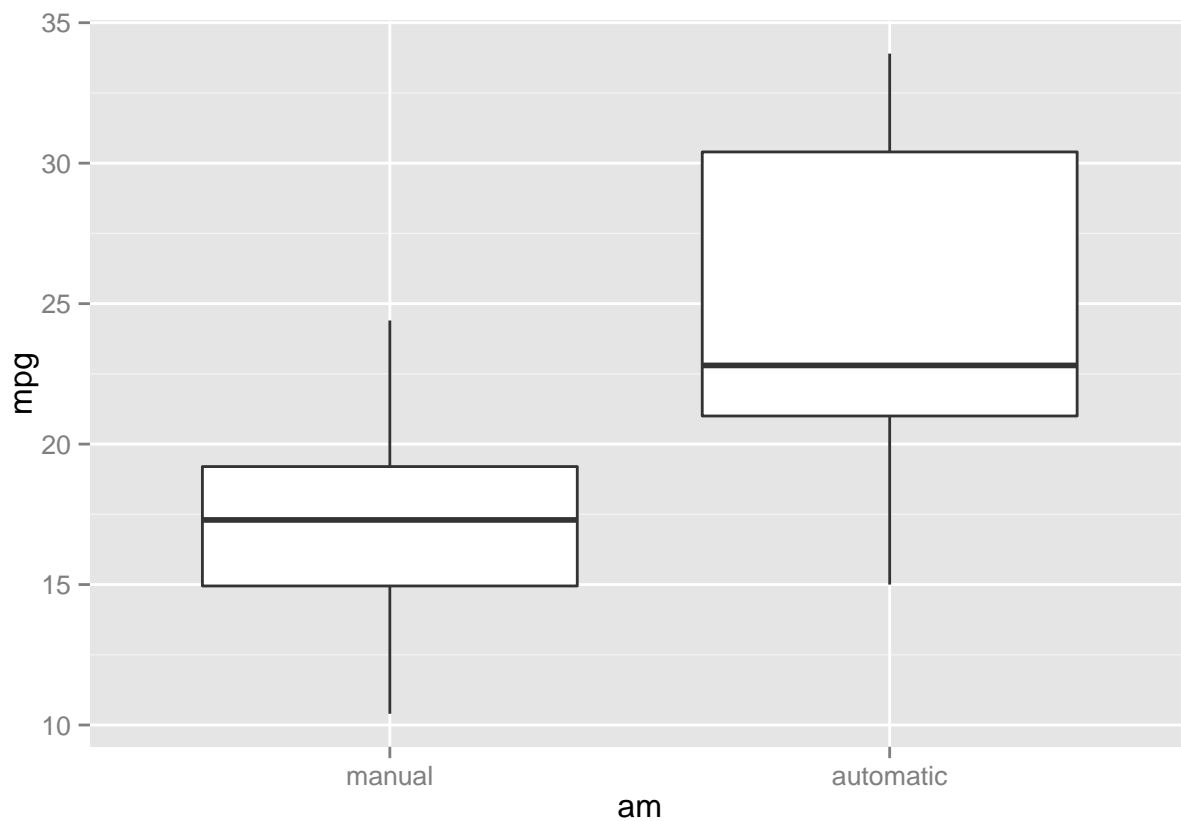
General data description

The data is the mtcars data set from the dataset library. The variables can be seen in R help.

Exploratory analysis

The interest is primarily into the mpg of the cars of the data set. Since the transmission type is of interest, too, let's have a look at this.

```
data <- mtcars %>% transform( am = factor(am, labels = c("manual", "automatic")))
g <- ggplot(data, aes(x = am, y = mpg))
g <- g + geom_boxplot()
g
```



Regression model

The coefficients of all variables against mpg.

```
summary(lm(mpg ~ ., data=data))$coefficients
```

	Estimate	Std. Error	t value	Pr(> t)
## (Intercept)	12.30337416	18.71788443	0.6573058	0.51812440
## cyl	-0.11144048	1.04502336	-0.1066392	0.91608738
## disp	0.01333524	0.01785750	0.7467585	0.46348865
## hp	-0.02148212	0.02176858	-0.9868407	0.33495531
## drat	0.78711097	1.63537307	0.4813036	0.63527790
## wt	-3.71530393	1.89441430	-1.9611887	0.06325215
## qsec	0.82104075	0.73084480	1.1234133	0.27394127
## vs	0.31776281	2.10450861	0.1509915	0.88142347
## amautomatic	2.52022689	2.05665055	1.2254035	0.23398971
## gear	0.65541302	1.49325996	0.4389142	0.66520643
## carb	-0.19941925	0.82875250	-0.2406258	0.81217871

As seen from the t-values, the weight (wt) seems to be the best candidate as a predictor for mpg. Since the transmission type is of interest, too, let's fit a model with only these 2 predictors.

```
summary(lm(mpg ~ wt + am, data=data))$coefficients
```

	Estimate	Std. Error	t value	Pr(> t)
## (Intercept)	37.32155131	3.0546385	12.21799285	5.843477e-13
## wt	-5.35281145	0.7882438	-6.79080719	1.867415e-07
## amautomatic	-0.02361522	1.5456453	-0.01527855	9.879146e-01

Unfortunately, the relation to am shows by the near-zero t value almost no relation in this regression context. I assume this has to do with the transmission type only having 2 levels. The data shows still that the cars with automatic transmission in the sample have larger mpg as seen in the figure. However, the data supports the statement that this is more due to their weight.

Hypothesis testing

Sources

The sources for the report can be found on Github at <https://github.com/mkraemerx/datasciencecoursera/tree/master/06StatisticalInference>