

My First Markdown Document

DataTeka

2018-06-20

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

Material from today's workshop can be found on our GitHub page.

When analysing data, a starting point is to examine the characteristics of each individual variable in the data set. The way to proceed depends upon the type of variable being examined. The variables can be one of two broad types: Attribute variable: has its outcomes described in terms of its characteristics or attributes; Measured variable: has the resulting outcome expressed in numerical terms.

Including R Code

When you click the Knit button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
install.packages("dplyr", repos = "http://cran.us.r-project.org")

##
## The downloaded binary packages are in
## /var/folders/71/96w85flx3yl928r2hzpfwvd00000gp/T//RtmpvruRpT/downloaded_packages
install.packages("ggplot2", repos = "http://cran.us.r-project.org")

##
## The downloaded binary packages are in
## /var/folders/71/96w85flx3yl928r2hzpfwvd00000gp/T//RtmpvruRpT/downloaded_packages
install.packages("gapminder", repos = "http://cran.us.r-project.org")

##
## The downloaded binary packages are in
## /var/folders/71/96w85flx3yl928r2hzpfwvd00000gp/T//RtmpvruRpT/downloaded_packages
summary(cars)

##      speed      dist
## Min.   : 4.0    Min.   :  2.00
## 1st Qu.:12.0    1st Qu.: 26.00
## Median :15.0    Median : 36.00
## Mean   :15.4    Mean   : 42.98
## 3rd Qu.:19.0    3rd Qu.: 56.00
## Max.   :25.0    Max.   :120.00

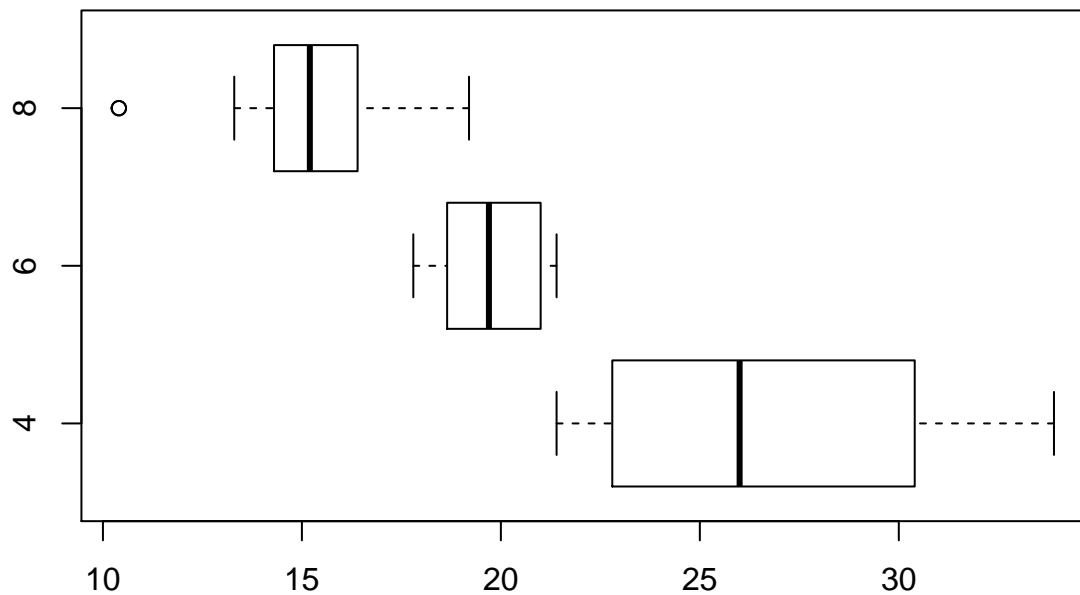
cars[1:10, ]

##      speed dist
## 1         4    2
## 2         4   10
```

```
## 3      7      4
## 4      7     22
## 5      8     16
## 6      9     10
## 7     10     18
## 8     10     26
## 9     10     34
## 10    11     17
```

Including Plots

You can also embed plots by setting `echo = FALSE` to the code chunk to prevent printing of the R code that generates the plot. For example:



Including Mathematical Equations

Let us fit the following model $lifeExp = b_0 + b_1pop + b_2gdpPercap$ which we write using the LaTeX.

```
library(gapminder)
m1 <- lm(gapminder$lifeExp ~ gapminder$pop + gapminder$gdpPercap)
summary(m1)
```

```
##
## Call:
## lm(formula = gapminder$lifeExp ~ gapminder$pop + gapminder$gdpPercap)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -82.754  -7.745   2.055   8.212  18.534
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.365e+01  3.225e-01  166.36  < 2e-16 ***
## gapminder$pop  9.728e-09  2.385e-09   4.08  4.72e-05 ***
```

```
## gapminder$gdpPercap 7.676e-04 2.568e-05 29.89 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.44 on 1701 degrees of freedom
## Multiple R-squared:  0.3471, Adjusted R-squared:  0.3463
## F-statistic: 452.2 on 2 and 1701 DF,  p-value: < 2.2e-16
```

What do we think of this model?

Let's discuss it next time we meet up.