

IV. R Markdown (presentation)

Center for Health Data Science, University of Copenhagen

15 March, 2023

This document does not include much, because we will simply review some features along the way. You are encouraged to work with the exercises (which include explanations) and ask us questions either during the course or later.

The R Markdown cheat sheet

The R Markdown tutorial by R Studio

The exercise you will do with R markdown will cover the following:

- Open new markdown document, look at its structure. Knit!
- Change some text, headers, etc.
- R chunks: Knit, run code without knitting, good practice
- Options in R chunks
- Tables
- Other output formats: docx, pdf, latex (Using the dropdown menu by the Knit button)
- Good practice: See the list in the end of the document with exercises

some more text formatting

You can easily use *italic* text and **bold** text. You can also mark text as `inline code`.

Some simple commands

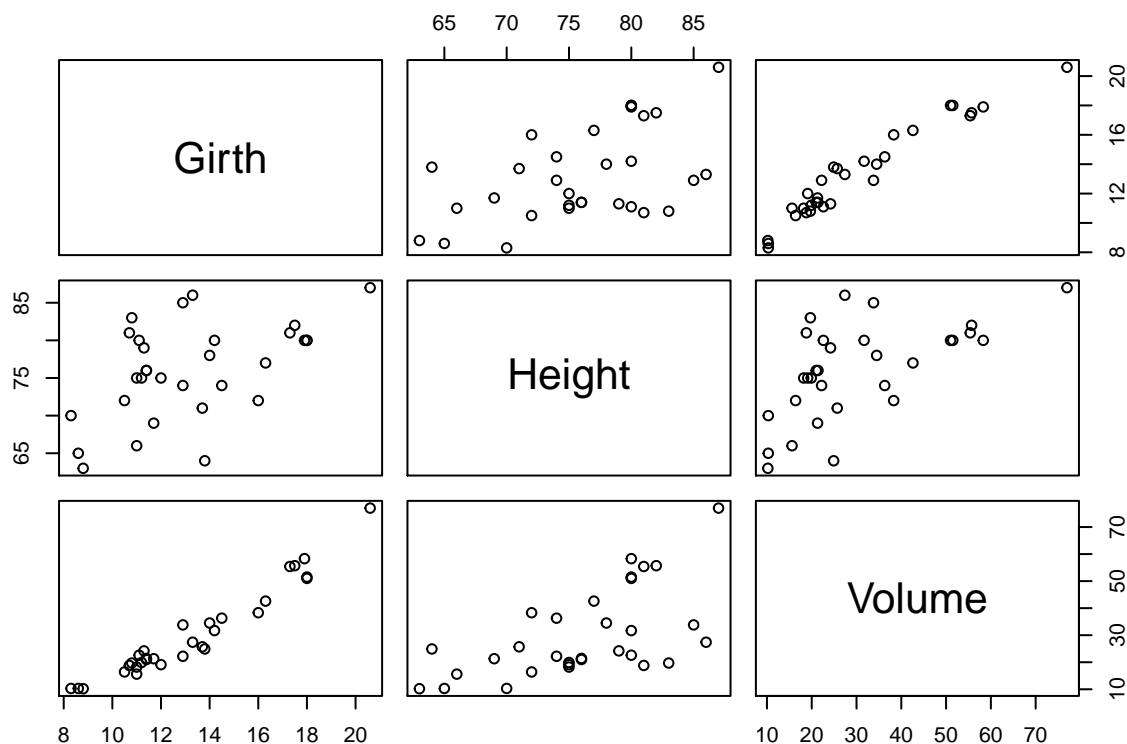
```
11 * 5
```

```
## [1] 55
```

```
sqrt(25)
```

```
## [1] 5
```

```
plot(trees)
```



Options in R chunks

It is possible to control if code and/or output is shown in the knitted document.

First, some code with the default settings (both code and output shown):

```
reg <- lm(Volume ~ Girth, data=trees)
summary(reg)$coefficients
```

```
##           Estimate Std. Error  t value    Pr(>|t|)
## (Intercept) -36.943459   3.365145 -10.97827 7.621449e-12
## Girth        5.065856   0.247377  20.47829 8.644334e-19
```

Then exactly the same code, but now with the code suppressed. This done with the option `echo=FALSE` (not visible in the output). The easiest thing is to insert such options via the small wheel in the upper right corner of the R chunk.

```
##           Estimate Std. Error  t value    Pr(>|t|)
## (Intercept) -36.943459   3.365145 -10.97827 7.621449e-12
## Girth        5.065856   0.247377  20.47829 8.644334e-19
```

Example: table1

The html output format plays well together with certain facilities for table generation. There is a function called `table1` which easily generates a table of statistics for variables of a dataset, possibly stratified after other variables in the dataset. The `table1` function is in a package with the same name.

We first (install and) load the package and import the **downloads** data. (I inserted an option such that we don't get messages about loading of packages). Note: You may need to adjust the path for the file `downloads.xlsx` to import the dataset. You can do this manually or via the "Files" tab in the bottom right window of Rstudio.

```
# install.packages("table1")
library(table1)
library(readxl)
downloads <- read_excel("./data/downloads.xlsx")
```

We then make an unstratified table (first) and a table stratified by machine name (second):

```
# Unstratified table
table1(~ size + time, data=downloads)
```

```
## [1] "<table class=\"Rtable1\">\n<thead>\n<tr>\n<th class='rowlabel firstrow lastrow'></th>\n<th clas
```

```
# Stratified table
table1(~ size + time | machineName, data=downloads)
```

```
## [1] "<table class=\"Rtable1\">\n<thead>\n<tr>\n<th class='rowlabel firstrow lastrow'></th>\n<th clas
```

Notice that the `table1` function generates `html-output` only, and it cannot be used with MS Word or pdf output. Therefore, look at the corresponding `html` file to see the content.