

## IV. R Markdown (presentation)

Data Science Lab, University of Copenhagen

August 2025

This document does not include much, because we will simply illustrate some features along the way. You are encouraged to work with the exercises which include explanations.

The presentation (but not this document, as just explained) will cover the following:

- Installation of the **rmarkdown** package
- 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
- Example: `table1`
- Output formats: html, docx (Word), pdf
- Good practice: See the list in the end of the document with exercises

---

### Some simple commands

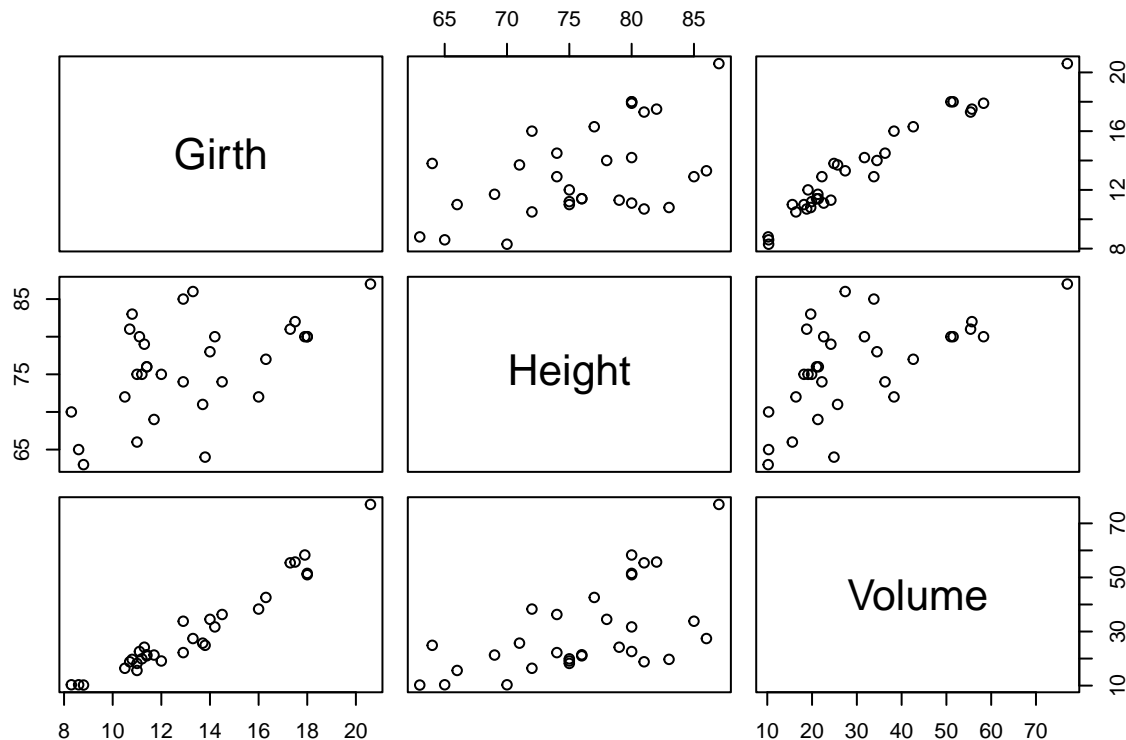
```
11 * 5
```

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

```
sqrt(25)
```

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

```
plot(trees)
```



## Run code without knitting: Results shown inline or in Console

We recommend that you run and adopt your R code without knitting all the time. You can do that as usual (e.g. with Ctrl-Enter) or you can click the “Play” button to the far left in the R chunk. You can choose whether the results are shown in the Console/Plot window (numerical results/graphs) or inline in the Rmd file: *Tools* → *Global options* → *R Markdown* → *Show output inline for all R Markdown documents* (tick if off or not)

## 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 1: table1

The html output format plays well together with certain facilities for table generation. As an example, 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 once again. (I inserted an option such that we don't get messages about loading of packages).

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

We then make an unstratified table:

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

	Overall
	(N=147035)
<b>size</b>	
Mean (SD)	4150 (88900)
Median [Min, Max]	0 [0, 14500000]
<b>time</b>	
Mean (SD)	0.954 (14.2)
Median [Min, Max]	0 [0, 1880]

And finally a table stratified after machine name:

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

	cs18	kermit	piglet	pluto	tweetie	Overall
	(N=16822)	(N=39157)	(N=41307)	(N=18396)	(N=31353)	(N=147035)
<b>size</b>						
Mean (SD)	5980 (100000)	4470 (103000)	3830 (98300)	3950 (77400)	3330 (46000)	4150 (88900)
Median [Min, Max]	0 [0, 6360000]	0 [0, 14500000]	0 [0, 14200000]	0 [0, 8670000]	0 [0, 4660000]	0 [0, 14500000]
<b>time</b>						
Mean (SD)	1.21 (26.8)	0.957 (13.0)	0.823 (8.51)	1.26 (17.1)	0.804 (9.20)	0.954 (14.2)
Median [Min, Max]	0 [0, 1750]	0 [0, 1380]	0 [0, 597]	0 [0, 1880]	0 [0, 1210]	0 [0, 1880]

Notice that the `table1` function generates nice html-output, but less nice pdf output unless the package **kableExtra** is also installed. The Word output is even worse if the package **flexTable** is not installed. You get a message about that if you have not installed the relevant package.

---

End of presentation.