Data analysis and visualization using R
The R toolbox

Michiel Noback

January 2020

Why do statistical programming?

Complete reproducibility and flexibility

- ▶ Keeping an exact log of all your mouse clicks in Excel may be problematic. . .
- ▶ Redoing an analysis with adjustments is easy
- ▶ Store the analysis, not its output

The R toolbox

Overview

This presentation will introduce you to a toolbox that will serve you well during your data quests.

- ▶ The R programming language
- ▶ The R studio IDE (Integrated Development Environment)
- ▶ R Markdown as documenting and reporting tool
- ▶ swirl as training environment

Tool 1: R itself

Figure 1: R logo

The R programming language

Wikipedia:
"R is a programming language and software environment for statistical computing and graphics. The R language is widely used among statisticians and data miners for developing statistical software and data analysis"
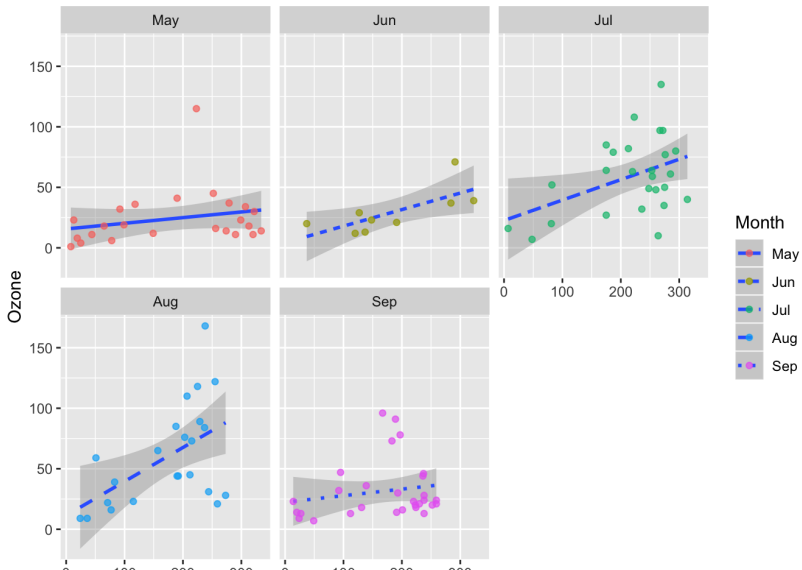
R is a programming language

- ▶ R is a programming language much like any other.
- ▶ It knows about character data, numbers, lists, functions
- ▶ It has an interactive mode

But

- ▶ It has embedded advanced graphical support
- ▶ It has extensive functionality for statistical work

R graphics showcase

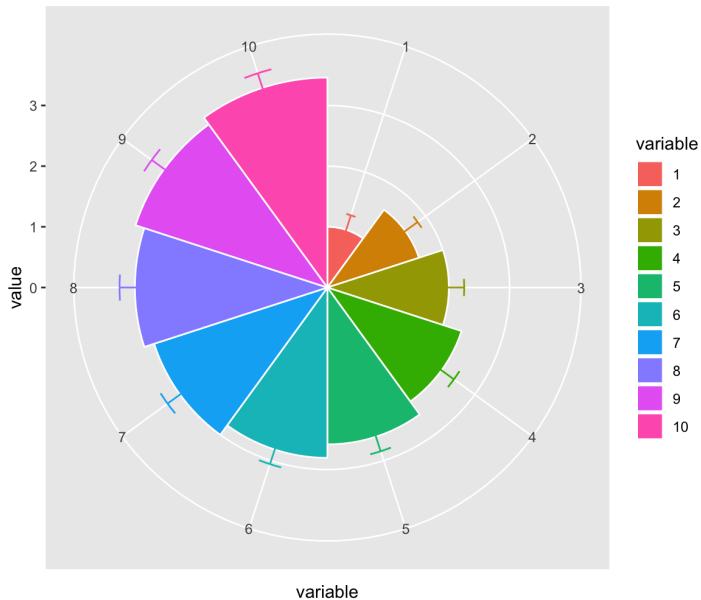Below follow some figures that you will be able to make at the end of this smester
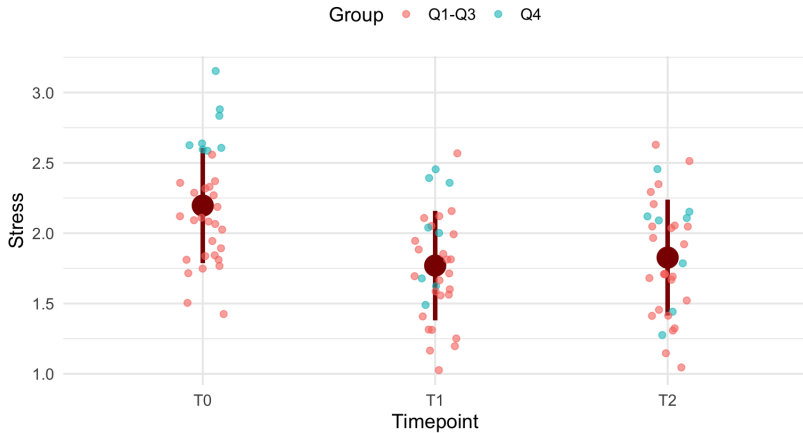
Figure 3: R logo

Figure 4: jitter

Tool 2: RStudio

Figure 5: RStudio logo

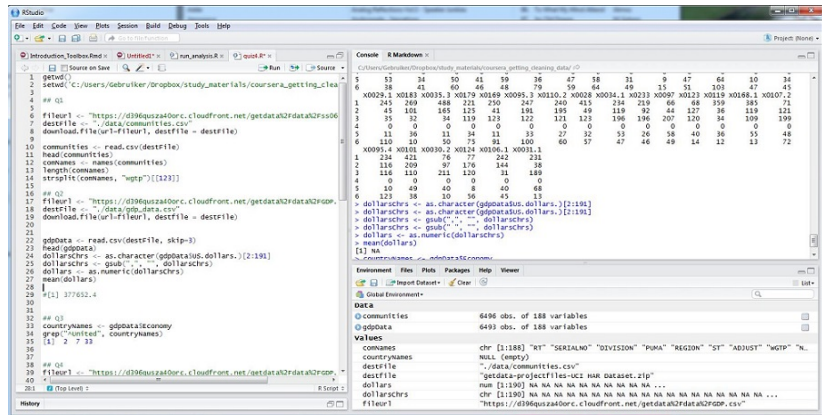# RStudio is the IDE for R

Figure 6: RStudio screen

Tool 3: RMarkdown

Figure 7: RStudio logo

RMarkdown is a great reporting tool

In RMarkdown, you can combine regular text and figures with embedded R code that will be executed to generate a final document.

You can use it to

▶ create reports in word, pdf or html
▶ create presentations (such as this one)
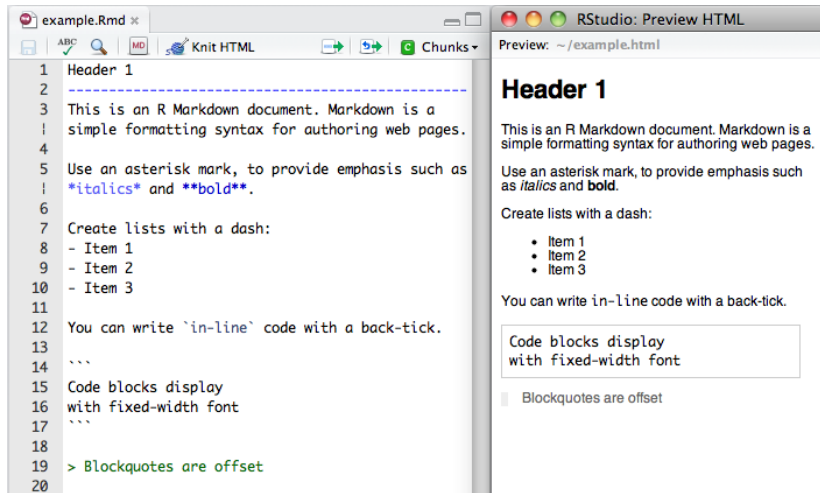▶ create entire websites (as the next course shows)

RMarkdown is really basic

Figure 8: RMarkdown

source http://rmarkdown.rstudio.com/

Figure 9: SWIRL

swirl

- Swirl is a very nice tool to train your R brain
- It is a framework for interactive lessons in R
- I recommend doing those!
- See swirlstats.com

The end

- This concludes the overview of the toolbox and demo use case
- In the next presentation we'll cover some basic R and then come back to visit R markdown