Introduction to R

Christian Dudel

April 5, 2021

Course instructors

-

Zoom

- Please use your preferred name consistently
- Please mute your microphone when not speaking
- Feel free to turn off video
- Feel free to interrupt me any time (unmute first)...
- ...for instance, if you cannot hear me well...
- ...but be aware that there might be some delay
- I might miss things you write in the Zoom chat

Materials

Materials will be available from GitHub, also mirrored on OSF:

- https://github.com/christiandudel/IIPS2021
- https://osf.io/dnm8y/

What will be covered in this part of the course?

- Software: R (and RStudio)
- Mostly basic things in these programs
- We will not cover many things
- We will not go deep

Prerequisites

- Basic demographic knowledge (e.g., you know what a ‘rate’ is)
- Basic statistical knowledge (e.g., you know what a ‘mean’ is)
- First experince using statistical software (Stata, Excel, SAS, SPSS, R, . . .)

Contact

- Email: dudel@demogr.mpg.de
- Twitter: @c_dudel
- Website: christiandudel.com

Exercises

- For each session, there will be some voluntary exercises to solve
- If you have any questions regarding the exercises you can post them on Slack
- Solutions will be available online (GitHub/OSF)
- These voluntary exercises have to be distinguished from the (mandatory!) assignment

What is R?

- R is an open source statistical programming language
- First release in 1995
- Used for data analysis and statistical programming

Why use R?

- Free, open source
- Can easily be extended
- More than 16,000 packages available
- Commonly used in both science and industry
- Tons of R-related materials: Books, journals, conferences, forums, tutorials...
- Many methods are already implemented in R

Why use RStudio?

- RStudio comes on top of R
- RStudio is a tool to use R more efficiently
- Features:
  - Syntax highlighting, code folding
  - Project management (e.g., GitHub)
  - Markdown support
  - ...

Disclaimer

- RStudio is not the only IDE/editor for R (ESS, RKWward, Tinn-R, . . .)
- R can be used in many different ways
- Example: base R vs tidyverse vs data.table vs specialized packages
- I do things in certain ways, and my teaching follows that
- This does not mean that my solutions are the only or the best way to do things

What do you need to get started?

- R: https://cran.r-project.org/
- R-Studio: https://www.rstudio.com/