R getting started - session 1

Alina Ferecatu
(with contributions from Pieter Schoonees and Andreas Alfons)
Rotterdam School of Management
Erasmus University

9/2/2020

Documents and software

Have the **latest version** of:

- R: https://CRAN.R-project.org
- RStudio Desktop: https://www.rstudio.com
- Installation instructions on Canvas

Today's lecture

Getting started with R:

- Overview of R ecosystem
- Load data
- Descriptive statistics
- Graphics

Overview of R

About R

- Open source (free) environment for statistical computing
- One of the most popular data science tools worldwide
- Runs on Linux/Unix, Mac OS X, Microsoft Windows
- Fully developed and easy-to-use programming language
- Extensible by community contributed packages
  - 13500+ packages on CRAN, Bioconductor and Github

What is R used for?

1. Collecting data
   - Scrape from the web, import from databases, ...
2. Preparing, exploring and cleaning data
   - Data wrangling, exploratory data analysis, plotting
3. Modelling
   - Regression, segmentation, machine learning, custom methods, ...
4. Model evaluation
   - Assessing model quality
5. Reporting results

- Writing (dynamic) reports, visualization, creating interactive web applications, ...

And more (including everything SPSS can do)

Who uses R

- Google
- Twitter
- Facebook
- New York Times
- John Deere
- Deloitte
- Credit Suisse
- Novartis
- eBay
- Ford Motor Company
- Kickstarter
- Uber
- Airbnb
- Booking.com
- Bank of America
- McKinsey & Company
- FourSquare
- ...

You too?

RStudio: four panels

- Top left: Script editor (if open)
- Bottom left: R console
- Top right: *Two tabs*
  - Environment: list of objects used in the session
  - History: allows to re-run previous commands
- Bottom right: Five tabs
  - Files: browse through filles on the computer
  - Plots: graphics are displayed here
  - Packages: list of installed packages
  - Help: R help files are displayed here
  - Viewer: local web content created in the session

Example session 1

Please open RStudio, and open example1.R in the script editor

- Available on Canvas

- Execute the line in which your cursor is with Ctrl / Cmd + Enter

- You can also type the command directly in the console, but it is better to store your commands separately in a script (reproducibility)

Script files in RStudio

Create a new script file: File ! New file -> R Script

Save script file:

- Keyboard shortcut: *Ctrl / Cmd + S*
- File -> Save As. . . and enter the file name in the dialog

—> Use file extension .R

Open existing script file:

- File -> Open File. . . and select the script file in the dialog
- Click the Files tab in the lower right panel, navigate to the script file and click on it

Some details

- R is case sensitive
- The + prompt means R is waiting for you to complete the command
- Press Esc in the console to cancel the command being evaluated
- Use the Tab key for code completion
- Remember to close your parentheses ()