

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?

- ➊ Collecting data
 - Scrape from the web, import from databases, ...
- ➋ Preparing, exploring and cleaning data
 - Data wrangling, exploratory data analysis, plotting
- ➌ Modelling
 - Regression, segmentation, machine learning, custom methods, ...
- ➍ Model evaluation
 - Assessing model quality
- ➎ Reporting results
 - Writing (dynamic) reports, visualization, creating interactive web applications, ...

And more (including everything SPSS can do)

Who uses R

- Google

- NYT

Now only one column You too?

Slide with Plot

