Introduction to R
(taught by Andreas Füglistaler)

This course provides an introduction to the statistical software "R" for data analytics.

Data analytics is a broad field comprising the whole process of examining raw data. It combines data inspection, processing, description, and, its core, data analysis. "R" provides the environment for all stages of data analytics.

"R" is a free, open-source software widely used in the fields of data analytics and statistical computing. One strength of the software lies in the option of extending its functionalities via so-called packages. Packages are bundles of code, data, and documentation developed by R-users and shared with the community for solving recurring problems in data analytics. Thus, "R" is a constantly evolving programming language with currently around 15,000 packages. Further, R provides various facilities for carrying out machine learning operations and is highly compatible with other programming languages.

The course first guides you through the basic functionalities of "R" and its interface "RStudio". It then discusses how different variable types and data structures are treated and stored in "R". It covers how to read raw data into "R" and how to prepare the data for analyses. Finally, the course provides basic skills in data description and data analysis.

Objectives

- To get familiar with the software "R" and its interface "R Studio"
- To be able to prepare datasets and run first analyses

Content

- Introduction to the surface of R/Rstudio
- Data and survey design
- Reading and writing data
- Data description
- Data management (organizing and transformation of data)
- Combining and reshaping datasets
- Data analysis

Prerequisites

- None (introductory course)

Duration

- 1 day (roughly 7*45 minutes)

Evaluation

- take home exam: project work to be solved in R

ECTS

- 0.5