

Introduction to R (Two-day-course)

R is considered the world's most powerful programming language for statistical computing, machine learning and data visualization and is also known for a thriving community of users, developers and contributors. Its popularity and user numbers are rising steadily.

This course provides an introduction to R and its most popular integrated development environment (IDE) RStudio. We cover both Base R functionality and some additional packages that make using R more convenient. Examples we look at include data on rock music history and the sinking of the titanic.

Topics include:

Introduction to R and RStudio

- Some background information on R
- How and why to use R together with RStudio
- Working in the console (command line)
- Data structures in R: vectors, matrices, arrays, lists, data frames, factors
- Atomic data types: logical, numeric, character
- Using R scripts
- Functions in R

Data preparation in R

- Base R and recommended packages
- Importing data from and exporting to various data formats (R, .csv, .txt, Excel, other software for statistical data analysis)
- Creating, transforming and deleting variables; sorting and filtering data
- Describing data sets, obtaining summary statistics

Statistical Analysis

- Correlation analysis, tests of significance
- Regression analysis
- Decision trees

Communicating results: Visualization and Reports

- Introduction to Base R graphics
- Introduction to ggplot2
- Reporting using R Markdown