



14 – 18 AUGUST 2017 / UNIVERSITY OF COPENHAGEN

# DATA SCIENCE WITH R

Do you want to analyse data in a structured, documented and well organized way? Then you want to learn R – and RStudio – a competitive and modern data science environment and programming language. With an unprecedented back catalogue of packages, R is extremely versatile, and it has statistical methods available for most tasks. R is, moreover, suitable for automatizing boring repetitive tasks, for making sure that your analyses are correct and reproducible, and for customizing your analysis to your needs.

## WHAT YOU WILL LEARN

You will learn to build a complete data analysis pipeline in R. This includes learning R programming techniques for:

- Data import from multiple sources
- Data manipulation and visualization
- Modeling
- Automatic and interactive report generation

In addition to the technical programming skills, you will also learn a conceptual framework for data analysis, where all steps of the data analysis are automatized via a programmatic pipeline.

## COURSE CONTENT

The course is based on RStudio and a collection of modern R packages. The focus will be on learning to exploit the full potential of these tools, which can serve as an infrastructure for almost any perceivable data analysis in R. Generalized additive models will be treated as a non-trivial example of how to build a predictive regression model in R.

## Core elements:

- RStudio: An integrated development environment for R, which supports interactive data analysis, building of data analysis pipelines, and R software development
- Tidyverse: A framework and collection of R packages centered on the concept of tidy data
- Generalized additive models: A flexible but interpretable and easy-to-use prediction model
- Visualization: High-quality figures are created from structured specifications using the R package ggplot2
- Reproducible analysis: Automatic and reproducible reports are written and generated using R Markdown
- Interactive communication: Reactive web-applications for interactive presentations of data and analyses are written using Shiny

## Other tools/methods and topics:

- R as a programming language
- Calling compiled code and Rcpp
- R package development
- Organisation of R code and version control
- Other statistical models e.g. mixed models, survival models, time series or sparse regression models

**“Fantastic program and faculty. Best update I have had in years. Great value for money.”**

*Former participant on a Copenhagen Summer University*



#### PARTICIPANTS

The course is for:

- People with some experience in data analysis i.e. SAS, Matlab or Python, but with no or limited experience with R
- People with an IT background with some experience in relation to programming and/or databases, but limited or no experience with data analysis and data modeling
- People with some experience using classic R and an interest to come up-to-date

R is a programming language and the course takes a programmatic approach to data analysis. To get the full benefit from the course the participants should therefore be interested in and willing to program. The statistical and mathematical prerequisites are limited, but participants should know about mean, variance and simple linear regression.

R ([www.r-project.org](http://www.r-project.org)) and RStudio ([www.rstudio.com](http://www.rstudio.com)) are open source and available free of charge. Participants are expected to bring a laptop with these programs installed.

#### COURSE DATES

5 days, 14 – 18 August 2017, 9:00 – 16:30 at the University of Copenhagen, Frederiksberg Campus.

#### COURSE DIRECTORS

**Niels Richard Hansen**, Professor, Department of Mathematical Sciences, University of Copenhagen  
**Anders Tolver**, Associate Professor, Department of Mathematical Sciences, University of Copenhagen

#### TEACHING MATERIAL

Participants will receive a copy of the book *R for Data Science* (2016) by Garrett Grolemund and Hadley Wickham.

#### COURSE FEE

EUR 2,600/DKK 19,000 excl. Danish VAT. Fee includes teaching, course materials, all meals during the course.

FOR MORE INFORMATION AND REGISTRATION:  
[copenhagensummeruniversity.ku.dk](http://copenhagensummeruniversity.ku.dk)

