DEPARTMENT OF MEDICAL EPIDEMIOLOGY AND BIOSTATISTICS

C8F2958, Introduction to R, 1.5 credits (hec)

Introduktionskurs i R, 1,5 högskolepoäng

Third-cycle level / Forskarnivå

Approval

This syllabus was approved by the The Committee for Doctoral Education on 2023-11-10, and was last revised on 2024-09-12. The revised course syllabus is valid from autumn semester 2024.

Responsible department

Department of Medical Epidemiology and Biostatistics, Faculty of Medicine

Prerequisite courses, or equivalent

Biostatistics I: Introduction for epidemiologists or corresponding courses.

Purpose & Intended learning outcomes

Purpose

The purpose of this course is to introduce students to using the R statistical software to perform basic to intermediate statistical data analysis in a replicable manner.

Intended learning outcomes

After successfully completing this course, students are expected to be able to:
- explain basic concepts of the R language and environment, the online- and offline sources of documentation for R, and basic concepts of data management and workflow in a standard statistical analysis,
- run a standard statistical analysis interactively within the R environment,
- formalize and document such a standard analysis as a stand-alone R script,
- produce graphical representations, as part of reporting their analysis,
- interpret their scripts for potential simplifications via functional implementation,
- find, install and compare extension packages for unfamiliar statistical applications.

Course content

The course will cover the basic elements of a standard statistical workflow: reading data into R; pre-processing and quality assessment of data via numerical and graphical methods; descriptive statistics via summary measures, tabulations and graphics; basic statistical inference in terms of significance testing and confidence intervals; specification, fitting & diagnosis of regression models; exporting and reporting results from the previous steps.

The course includes an introduction to the Rstudio integrated development environment to provide a common framework for interactive and scripted analysis. The extensibility of the R system will be demonstrated by example.

Forms of teaching and learning

Theoretical concepts and background will be covered via presentations, demonstrations, live exercises and discussions. Students will practice the application of these ideas in individual and small-group lab exercises with support from qualified teaching assistants. Formative assessment will be integrated via quizzes and lab reviews.

Language of instruction

The course is given in English

Grading scale

Pass (G) /Fail (U)

Compulsory components & forms of assessment

Compulsory components

The individual examination (summative assessment) is compulsory.

Forms of assessment

Students will perform an open-book examination based on practical application of the concepts presented during the course to realistic data sets and problems.

Students who do not pass the examination will be offered a second examination within two months from the end of the course (excluding academic holidays).

Course literature

1. Lecture notes and code examples as provided by the course
2. Official R manuals and documentation available within the software
3. Selected material from the contributed documentation available from the Complete R Archive Network at https://cran.r-project.org/other-docs.html