UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Programming Language Laboratory

2627-1-E4104B008

Learning objectives

The course aims to provide students with practical skills in using the R language for data analysis and representation. Starting from the working environment and fundamental data structures, the course will guide students through techniques for importing, manipulating, and joining datasets, up to the production of effective graphical visualizations. Upon completion of the course, students will be able to use R and the main libraries of its ecosystem to independently tackle real-world data analysis problems.

Contents

- The R working environment (RStudio, projects, scripts)
- Data structures in R: scalars, vectors, data frames
- Data import and project organization best practices
- Data frame manipulation: exploration, selection, filtering, missing values
- Joining data frames: row-binding and merge/join
- Data visualization with ggplot2

Detailed program

- Introduction to R and RStudio
- Working environment: console, scripts, workspace
- R as a calculator: arithmetic operators and mathematical functions
- Objects and variables: creation, inspection and deletion
- R projects and reproducible folder organization
- Data structures in R: everything is a vector

    * Scalars
    * Vectors: creation, indexing, operations
    * Arithmetic progressions and sequences
- Data frames: structure and creation
- Importing data from CSV files
- Exploring a data frame: structure, dimensions and descriptive statistics
- Extracting columns and row subsets
- Missing values: identification, counting and removal
- Concatenating data frames by rows
- Merging data frames: inner join and outer join
- The grammar of graphics: the ggplot2 package
- Structure of a ggplot2 chart: data, aesthetics, geometries
- Main chart types and customization

Prerequisites

None.

Teaching methods

- Frontal lecture
- Hands-on session in a computer science laboratory
- Final test simulation (to introduce the student to the test platform and help them gain knowledge about their preparation level)

Assessment methods

The examination consists of a written part held in a computer lab. The examination requires to solve exercises by means of a program developed in R.
The nature of the written test allows to verify the specific knowledge of the student, together with its ability to build a logic path to tackle and solve a data analysis problem.

Facultative oral exam (on request of the teacher or of the students). The oral exam can both increase or decrease the overall evaluation.

Textbooks and Reading Materials

All the material presented during the lectures will be published on this web site.

Semester

II semester

Teaching language

Italian

Sustainable Development Goals

QUALITY EDUCATION