PROPOSERS
Prof. Davide Gentilini

TITLE: SummeR School: Introduction to R for Statistical Analysis

EDUCATIONAL AIM
Data management and analysis is one of the most important and most critical factors in many work areas; proprietary tools and software are often expensive or limited in their application. The course aims to provide participants with the necessary and sufficient knowledge to introduce and use the language and potential of R in their work. R can be defined as an environment for statistical analysis and at the same time a language and a software. It is a very powerful and widely used open source tool for statistical analysis of data, moreover, being a real programming language, it contains the potential to independently create and develop various applications useful for manipulation, management and analysis of any type of data. In fact, its main features include simplicity in data management and manipulation, the availability of a suite of tools for calculations on vectors, matrices and other complex operations, access to a vast set of integrated tools and functions developed by others and made available for statistical analysis, the production of numerous particularly flexible graphic potentials, the possibility of using a real object-oriented programming language that allows the use of conditional and cyclic structures, as well as functions created by the user. Offering an elaborate introduction to programming with R, this course aims to intercept the needs of the participants by focusing in particular on some fundamental aspects such as data manipulation and management, their analysis through the identification of the most appropriate statistical test and the visualization of data and results using the graphic potential made available by R. The course involves the use of numerous "datasets" and examples that may be familiar to the various areas of interest in order to facilitate participants in understanding and applying the knowledge acquired.

LANGUAGE: English

NUMBER OF HOURS (CFU) / LESSONS
The course will last at least 20 hours (5 CFU), and will be divided into 5 meetings. Individual meetings can also be attended.

COURSE PERIOD
22–26 June 2026 (summeR school). The calendar may be subject to change

TEACHER: Dr. Barbara Tarantino, Dr. Davide Sacco & Prof. Davide Gentilini

SCIENTIFIC COMMITTEE, Dr. Barbara Tarantino, Dr. Davide Sacco & Prof. Davide Gentilini
ORGANIZING COMMITTEE Department of Nervous System and Behavior Sciences

POTENTIAL INTERESTED DOCTORATES
The course aims to be transversal and to provide useful skills in every area that has to do with data and the need to process and manage them. For this reason, the course aims to be useful to any type of doctorate, in particular the macro-area of Life Sciences.

LEARNING VERIFICATION METHOD
The degree of learning will be tested at the end of the course by subjecting the participants to a test. The verification test will consist of a dataset to analyze and a series of questions related to the dataset to answer.

On line REGISTRATION FORM: https://forms.gle/UmXK6DMswL7ZdmsZ6

PROGRAM
Activities Calendar and Program

22 June 2026 – 9.00-13.00
Title: "The general programming environment R and the main objects"
The objective of the lesson is to introduce the participants to the R environment, illustrating its main functions, architecture and presenting the main objects.
• Introduction of the Course and Generalities
• R installation and configuration
• Preliminary notions on its architecture on packages and functions
• Use of packages
• Use of the Help
• The main objects
• The vectors
• Vector Assignment (Arithmetic, Logic, Character Vectors)
• Operations and functions for working with vectors
• Matrices and Arrays
• Operators and functions for the Matricial calculation
• Lists and Data Frames
• Operators and functions for working with Data Frames and Lists
• Examples
• Practical exercise

23 June 2026 – 9.00-13.00
Title: "Data import and data manipulation with R"
The aim of the lesson is to teach participants to import data of various types and formats into the R environment. The lesson also aims to teach "data description" or to use R to view, manage, extract and manipulate data.
• Import and cleanup of data sets
• Data import functions based on their nature and format
• Data display and manipulation functions
• Data Quality Control
• Treatment of missing data
• Imputation
• Methods of reclassification of variables
• Functions for data Manipulation

24 June 2026 – 9.00-13.00
Title: "The Graphics with R"
The lesson will cover the graphical functions of R so that course participants acquire the basic knowledge to be able to view their data and results independently.
• The graphic environment
• Basic Package
• ggplot2 package
• The graphic parameters
• Types of graphs and functions
• Classification of data type and discussion on how to display
• Examples
• Practical exercise

25 June 2026 – 9.00-13.00
Title: "Basic statistics with R"
The objective of the lesson is to go into the merits of descriptive / inferential statistics and provide the essential tools to be able to choose and apply the most appropriate statistical method.
• Study design principles and data types
• Univariate / bivariate descriptive measures
• Statistical tests: Parametric tests on mean, difference between means, proportion, difference between proportions and correlation Nonparametric tests on mean, difference between means, difference between medians, proportion, difference between proportions and correlation, Association test for 2x2 and JxK tables

- Examples
- Practical exercise

26 June 2026 – 9.00-13.00
Title: "Basic statistics with R"
- Analysis of variance
- Simple and multiple linear regression
- Logistic regression
- Principal component analysis
- Examples
- Practical exercise