UNIVERSITÀ CATTOLICA del Sacro Cuore

The R language and environment for statistical computing

PROF. GABRIELE CANTALUPPI

Course aims and intended learning outcomes

Make participants able to implement statistical procedures in order to properly solve business and economic problems by means of the open source software R.

On successful completion of the course participants will be expected to possess:

1. knowledge of concepts and methods of coding and data visualization with R (DD1- Knowledge and understanding);
2. ability to correctly apply statistical procedures and coding with R to real economics and management problems (DD2- Applying knowledge and understanding);
3. quantitative thinking addressed to make independent judgements, driven by application of statistical methods and data visualization (DD3- Making judgements);
4. ability to present and visualize statistical results and the conclusions from them, by means of the extraction of qualitative information from data, with clarity and accuracy and in forms that are suitable for the audiences being addressed, both orally and in writing (DD4-Communication);
5. mastery of R coding procedures, rigorous reasoning and data-driven decision-making, useful for quantitative analyses in courses at the MSc curricula level, as well as for analyses required in careers in all fields involving management of data (DD5- Lifelong learning skills)..

Course content

1. Basic elements of computer programming: the R software environment.
   Data types and Data structures. Data manipulation.
   Writing functions. Conditional execution. Efficient calculation.
   Data import. Graphics production.
2. Statistics, Mathematics, Business and Economics applications.
   Location and variability measures for categorical and metric variables. Graphical representations. Contingency tables and association measures. Linear and logistic regression. Monte Carlo simulation. Graphical representations of functions in one and two variables. Optimization problems. Numerical solution of linear and non-linear systems of equations. Graphical solution of inequalities. Numerical integration.
   Analysis of a Marketing plan.
   Analysis of the 'SHIW survey on households wealth' administered by the Bank of Italy.
3. Reproducible research with knitr, markdown and Latex. Data visualization with Shiny..

Reading list

G. CANTALUPPI, An Introduction to R, EDUCatt, 2018.
G. CANTALUPPI, Exercise with R, EDUCatt, 2018.

Further readings:

G. BOARI-G. CANTALUPPI, Notes of Descriptive Statistics and Probability, EDUCatt, Milano, 2018.
G. CANTALUPPI, Computational Laboratory for Economics with R, EDUCatt, 2016.

UNIVERSITÀ CATTOLICA del Sacro Cuore

Teaching method

Lectures and lab-assignments.

Assessment method and criteria

A written exam made by 30 questions, each with three possible answers. Each correct response gives 1 point; a penalty of -0.5 points is assigned to each wrong answer; no given answers have no points and no penalty (maximum score 30 points). Available time: 60 minutes.

Writing some requested code (maximum score 10 points). Available time: 30 minutes.

Solutions of in-course homeworks can grant from 0 to 3 point bonus.

A late assignment, consisting in a production of a report with Shiny or Knitr, can grant from 0 to 2 point bonus.

A minimum score of 18 points has to be achieved in order to pass the exam.

Notes and prerequisites

Students enrolling in this course should have a basic understanding of mathematics and statistics with regard to data analysis, probability and inference, at the level of the combined courses ‘Matematica generale’/Mathematics, ‘Statistica (Analisi dei dati e probabilità)’/Statistics and ‘Statistica applicata’/‘Applied statistics and big data’. Students can attend lectures of this course simultaneously with ‘Statistica applicata’.