
Academic Year: (2022 / 2023)

Review date: 18-05-2022

Department assigned to the subject: Statistics Department

Coordinating teacher: UCAR MARQUES, IÑAKI

Type: Compulsory ECTS Credits : 6.0

Year : 1 Semester : 1

REQUIREMENTS (SUBJECTS THAT ARE ASSUMED TO BE KNOWN)

Introduction to Programming with R (19151)

OBJECTIVES

Core Competences:

- Having and understanding the knowledge that provides a basis or opportunity to be original in the development and/or application of ideas, often in a research context.
- Students know how to apply their acquired knowledge and problem-solving skills in new or unfamiliar settings within broader (or multidisciplinary) contexts related to their field of study.
- Students are able to integrate knowledge and to face the complexity of making judgments based on information that, being incomplete or limited, includes reflections on the social and ethical responsibilities linked to the application of their knowledge and judgments.
- Students know how to communicate their conclusions and the knowledge and ultimate reasons behind them to specialised and non-specialised audiences in a clear and unambiguous way.
- Students have the learning skills that will enable them to continue studying in a way that will be largely self-directed or autonomous.

General Competences:

- Ability to understand and analyze the main global social theories and how they are changing with the application of computational tools.
- Ability to identify, define and formulate social science problems and solve them using computational techniques. This includes the ability to assess all the factors involved, not only technical but also legal.
- Ability to compile and analyze existing knowledge in the different areas of computational social sciences, and to propose possible solutions to the problems raised.
- Ability to apply theoretical and methodological knowledge of computational social sciences to the analysis and resolution of specific cases and empirical problems.
- Ability to address issues raised under new or unfamiliar environments, within the context of computational social sciences.

Specific Competences:

- Ability to use computational tools specific to the computational social sciences at an advanced level.

Learning Outcomes:

- Knowledge of the main tools in computational social sciences.
- Knowledge of data programming structures and procedures.
- Ability to import tabular data in a variety of formats with the R programming language.
- Ability to work with remote databases.
- Ability to prepare, clean, transform and enrich tabular data for further modeling and visualization with R and SQL programming languages.

DESCRIPTION OF CONTENTS: PROGRAMME

1. R programming
 - 1.1. Basic data types and operations
 - 1.2. Conditionals and control flow
 - 1.3. Loops and vectorial programming
 - 1.4. Functions
 - 1.5. Understanding errors and debugging

2. Data handling
 - 2.1. Importing and exporting tabular data
 - 2.2. Data filtering and reordering
 - 2.3. Basic Exploratory Data Analysis
 - 2.4. Data aggregation. The apply family
 - 2.5. Joining data frames
 - 2.6. Long and wide formats

3. R Ecosystem
 - 3.1. Packages and CRAN
 - 3.2. Important contributors and future trends

4. Tidyverse
 - 4.1. Pipe operator
 - 4.2. tidyr package
 - 4.3. dplyr package

5. Advanced data types
 - 5.1. Factors
 - 5.2. Dates and times: lubridate package
 - 5.3. Strings: regular expressions and stringr package

LEARNING ACTIVITIES AND METHODOLOGY

Training Activities:

- Theoretical-practical classes
- Tutorials
- Group work
- Individual student work
- Partial and final examinations

Teaching Methods:

- Presentations in the professor's lecture room with computer and audiovisual support, in which the main concepts of the subject are developed and a bibliography is provided to complement the students' learning.
- Resolution of practical cases, problems, etc. raised by the professor, either individually or in a group.

ASSESSMENT SYSTEM

- Participation in the class (10%)
- Individual or group work done during the course (40%)
- Final exam (50%)

% end-of-term-examination:	50
% of continuous assessment (assignments, laboratory, practicals...):	50

BASIC BIBLIOGRAPHY

- Hadley Wickham R for Data Science, O'Reilly, 2017

ADDITIONAL BIBLIOGRAPHY

- Chester Ismay and Albert Y. Kim Statistical Inference via Data Science: a Modern Dive into R and the tidyverse, Chapman & Hall, 2022
- Steph Locke Data Manipulation in R, Locke Data, 2017

