

---

**Academic Year: ( 2023 / 2024 )****Review date: 20-07-2023**

---

**Department assigned to the subject: Statistics Department****Coordinating teacher: UCAR MARQUES, IÑAKI****Type: Electives ECTS Credits : 3.0****Year : 1 Semester : 2**

---

## REQUIREMENTS (SUBJECTS THAT ARE ASSUMED TO BE KNOWN)

Programming in R  
Advanced Programming

## OBJECTIVES

The student will acquire the following knowledge:

- Knowledge of techniques for automatic presentation of results in reports.
- Ability to develop Shiny applications.
- Knowledge of the tidyverse environment.
- Knowledge of the tidymodels environment.

## DESCRIPTION OF CONTENTS: PROGRAMME

This course covers several tools for streamlining the consulting pipeline in R: from data wrangling to presentation of results, passing through a fast statistical modeling. The focus is placed on seeing the main features of many different packages and solutions.

1. Advanced R Markdown for reporting
  - 1.1. Advanced topics on R Markdown
  - 1.2. Writing good reports
  - 1.3. Customized presentations
  - 1.4. Other documents and topics
2. Shiny applications
  - 2.1. Main paradigm
  - 2.2. Examples of simple applications
  - 2.3. Reactions and appearance
  - 2.4. More advanced applications
  - 2.5. flexdashboard
  - 2.6. Other topics
3. Data wrangling in the tidyverse I
  - 3.1. dplyr
  - 3.2. tidyr
  - 3.3. readr
  - 3.4. tibble
  - 3.5. Other packages
4. Data wrangling in the tidyverse II
  - 4.1. stringr
  - 4.2. forcats
  - 4.3. lubridate and hms
  - 4.4. glue
  - 4.5. purrr
  - 4.6. Other packages
5. Fast modeling using AutoML
  - 5.1. Introduction to AutoML
  - 5.2. Explainability

- 5.3. Examples in regression
- 5.4. Examples in binary classification
- 5.5. Examples in multiclass classification
- 6. Fast modeling with tidymodels I
  - 6.1. broom
  - 6.2. rsample
  - 6.3. parsnip
  - 6.4. yardstick
  - 6.5. Other packages
- 7. Fast modeling with tidymodels II
  - 7.1. recipes
  - 7.2. workflows
  - 7.3. tune
  - 7.4. infer
  - 7.5. Other packages

The program is subject to modifications due to the course development and/or academic calendar.

## LEARNING ACTIVITIES AND METHODOLOGY

The classes consist of a mixture of lectures on the software and its practical use. The statistical language R is used. Students are expected to bring their own laptops to experiment with the code during the lectures.

### \* Training activities

- AF1: Theoretical lesson.
- AF2: Practical lesson.
- AF5: Tutorials.
- AF6: Group work.
- AF7: Individual work.
- AF8: On-site evaluation tests.

### \* Teaching methodologies

- MD1: Class lectures by the professor with the support of computer and audiovisual media, in which the main concepts of the subject are developed and the bibliography is provided to complement the students' learning.
- MD2: Critical reading of texts recommended by the professor of the subject: press articles, reports, manuals and/or academic articles, either for later discussion in class, or to expand and consolidate the knowledge of the subject.
- MD3: Resolution of practical cases, problems, etc. posed by the teacher individually or in groups.
- MD4: Presentation and discussion in class, under the moderation of the professor of topics related to the content of the subject, as well as case studies.
- MD5: Preparation of papers and reports individually or in groups.

## ASSESSMENT SYSTEM

<b>% end-of-term-examination:</b>	0
<b>% of continuous assessment (assignments, laboratory, practicals...):</b>	100

The evaluation in the ordinary call is done entirely by continuous evaluation. This is done by a mixture of:

- (a) a set of practical exercises;
- (b) active participation in lessons.

The continuous evaluation grade (in the scale 0-10) is

$$\min(A + 0.10 * B, 10),$$

where

- A (in the scale 0-10) is the grade of the practical exercises;
- B (in the scale 0-10) is the grade corresponding to (b).

<b>% end-of-term-examination:</b>	0
<b>% of continuous assessment (assignments, laboratory, practicals...):</b>	100

Students who have not followed the continuous evaluation may take a final exam in the ordinary call with a value of 60% of the final grade.

The grade in the extraordinary call is established by the delivery of a set of practical exercises.

Further details are provided in Aula Global. The evaluation is subject to modifications due to the course development and/or academic calendar.

#### BASIC BIBLIOGRAPHY

- Hadley, W. and Golemund, G. R for Data Science, O'Reilly, 2017
- Xie, Y., Allaire, J.J., and Golemund, G. R Markdown, CRC Press/Chapman & Hall, 2019

#### BASIC ELECTRONIC RESOURCES

- Kuhn, M. and Silge, J. . Tidy Modeling with R: <https://www.tmw.org/>