Statistics and data science for PhD research.
Module 3: Research project

Department in charge: 715 - EIO - Department of Statistics and Operations Research

Total teaching hours: 15; lectures will take place on May-June in sessions of 3 hours, 4 sessions face to face and 1 of the five sessions online.

Language: English

Coordinating lecturers: Nuria Perez Alvarez, Victor Peña

Teaching methodologies:
The course has a light theoretical part, and a practical PBL-oriented (Project / Problems Based Learning) part. Specifically, the teaching methodology will be as follows:

a) Outline the methodological needs from real data analysis
b) Understand the possible theoretical models and selection of the proper/s one/s
c) Return to the data to perform the data management, the analysis and interpretation of results
d) Scientific communication of the results.

Labs sessions will be in R.

Learning objectives:
During the course, the instructors will train and guide the process of research, data analysis and results communication. The students can propose their own research projects, which may be based on their own research. If a student wishes to follow the course without proposing a project, the teachers can provide a project and a data set to work with.

Data analysis will be done using the statistical software R.

The learning objectives are:
a) Understand numerical results and develop critical thinking skills
b) Be able to recognise data characteristics to identify the proper methodology to gain insights about a process
c) Be able to formulate scientific questions
d) Implement the statistical knowledge gained by using the statistical software R
e) Communicate scientific results

The pre-requisites for the course are: ability to understand advanced concepts in statistics, ability to perform simple data analyses, numerical and quantitative skills, capability to read scientific publications, and intermediate level of English.

Continuous evaluation. During the course, exercises and tests are going to be assessed and scored. A minimum of 80% attendance is required.

Course contents:
The specifics of the projects will depend on the students’ interest. In all cases, we expect the students to go through the following workflow:

1. Formulate a research question and, potentially, design an experiment.
2. Collect, clean, and validate the data
3. Choose an implement appropriate statistical methodology
4. Interpret and communicate the results, identifying strengths and weaknesses of their analysis

At the end of the course, students will present their projects to the instructors and fellow students.

Basic bibliography:
Complementary bibliography: