Statistics and data science for PhD research.  
Module 2: Modelling

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

Total teaching hours: 15; lectures will take place on January-february 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 lectures delivered by the instructors and practical and PBL oriented (Project / Problems Based Learning) sessions.

Specifically, the teaching methodology will be as follows:
a) Outline the methodological needs from real data analysis,
b) Develop the theoretical model (interest will be focused on the calculation and interpretation of results and, secondarily, in demonstrating the theoretical results).
c) Analyse real data and interpret the results.

In the lab sessions, we will use the statistical software R.

Learning objectives:
The main course objective is, first, to develop the theoretical framework and, second, to implement the knowledge gained by using the statistical software R. The objectives of this module can be broken down into:

a) Be able to recognise data characteristics to identify the proper methodology to gain insights about a process.
b) Data management
c) Learn and understand the advanced modelling techniques of statistics
d) Implement the statistical knowledge gained by using the statistical software R.

The pre-requisites for the course are: ability to understand basic concepts and calculations in statistics, numerical and quantitative skills, capability to read scientific publication, 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:
- Data management and linear regression (2 hours)
- Generalized linear models (4 hours)
- Survival analysis (2 hours)
- Machine learning (2 hours)
- Bayesian inference (2)
- Intermediate R (3 hours).
The lectures will provide examples of R code that implement the methodologies introduced.

Basic bibliography:

**Complementary bibliography:**