

## Course Syllabus-Modern Statistical Computing

---

**Language of Instruction:** English

**Professor:** Vladimir ZAIATS PROTCHENKO

**Professor's Contact and Office Hours:** [vladimir.zaiats@upf.edu](mailto:vladimir.zaiats@upf.edu), send an e-mail for an appointment

**Course Contact Hours:** 45 hours

**Recommended Credit:** 6 ECTS credits (international students) / 5 ECTS (UPF students)

**Weeks:** 4

**Course Prerequisites:** An introductory course on probability and statistics is basic for enrolment to this course. For UPF students, the compulsory requirement is the Probability and Statistics course of the second year in the studies of ECO/ADE/IBE.

**Language Requirements:** Recommended level in the European Framework B2 (or equivalent: Cambridge Certificate if the teaching language is English, DELE or 3 semesters in the case of Spanish).

### Course Description:

Statistical computing is a highly sought-after analytical data analysis skill both in professional and research environments. This course presents modern graphical displays and data manipulation methods, interactive and reproducible reporting, emphasizing statistical methods and computation related to regression and classification methods such as linear, generalized linear and non-linear models. The concepts are introduced in R (<https://www.rproject.org>) one of the leading programming languages for statistical computing. Knowledge of R is highly-valued by companies in many sectors for positions related to data science, quantitative analysis, or finance. R is also an indispensable tool in most research fields, including Economics, Finance, Marketing, Biomedicine, etc. R provides a rich set of off-the-shelf data analysis tools, and the possibility to design our own data processing and analyses. R runs in all operating systems (Windows, Mac, Unix-like) and is a free open-source language that is enhanced by an extensive list of user-contributed packages. The purpose of this course is to introduce students to statistical computing, including flexible regression data analysis methods, and to advanced R skills. The idea is that students learn by doing. Therefore, there is a strong applied emphasis, all concepts are driven by examples discussed in class, where students are given the code to reproduce them. Students will become skilled in applications of elementary statistical methods, with an emphasis on data exploration, graphics, and programming. Focus will

also be placed on opportunities to enhance the learning experience in other statistical courses.

### **Learning Objectives:**

At the end of the course:

- Everybody will be able to use a fundamental data analysis tool for quantitative analytical methods
- Programming, data handling, exploratory data analysis, linear/generalised linear/non-linear regression, summarising data, effective graphics, model-free computational methods (bootstrap, permutation tests, cross-validation)
- Preparing notebooks to automatically perform quantitative analyses and create reports in formats such as pdf and html, with interactive elements

### **Course Workload:**

The course is constituted by lectures and practices. Students should be prepared to work with computer codes on a permanent basis.

### **Methods of Instruction:**

The course includes both lectures and practice with laptop computers. The teaching philosophy is that students learn by doing.

Classroom sessions are normally split into a lecture and a practice part.

**Students are required to attend classes with their own laptops.**

### **Method of Assessment:**

Homework + Class contribution: 20 percent

Controls, in-class exercises: 40 percent

Final project: 40 percent

The Final Project (in groups of 2 students) requires a report to be submitted of up to 10 typed pages (not counting appendices). Students will select their projects from topics of their own interest (accepted by the course instructors) and will make a brief oral presentation at the end of the course.

**Absence Policy:**

Attending class is mandatory and will be monitored daily by professors. The impact of absences on the final grade is as follows:

Absences	Penalization
Up to three (3) absences	No penalization.
Four (4) absences	1 point subtracted from final grade (on a 10 point scale)
Five (5) absences	2 points subtracted from final grade (on a 10 point scale)
Six (6) absences or more	The student receives an INCOMPLETE for the course

The BISS attendance policy does not make a distinction between justified and unjustified absences. All absences—whether due to common short-term illnesses or personal reasons—are counted toward the total amount and cannot be excused. Therefore, students are responsible for managing all their absences.

Only in cases of longer absences—such as hospitalization, prolonged illness, traumatic events, or other exceptional situations—will absences be considered for exceptions with appropriate documentation. The Academic Director will review these cases on an individual basis.

Students must inform the Instructor and the International Programs Office promptly via email if serious circumstances arise.

**Classroom Norms:**

- No food or drink is permitted.
- There will be a ten-minute break during the class.
- Students must come to class fully prepared.

**Course Contents:****Week 1**

Introduction to R  
Graphical displays (ggplot)

*The intellectual property rights of this course belong to the instructor  
The exploitation rights of this course belong exclusively to Universitat Pompeu Fabra*

Data handling

## **Week 2**

Programming basics.

Statistical models: linear, generalized linear models. Non-linear regression.

Optimisation.

## **Week 3**

Computational inference methods: bootstrap, permutation tests, cross-validation.

Model comparison techniques.

## **Week 4**

Advanced reports: interactive plots, dashboards.

Methods of flexible data analysis.

**Required Readings:** The professor will assemble a coursepack/or indicate mandatory textbooks.

## **Recommended bibliography:**

Students are encouraged, but not limited, to consult the following sources on their own.

Wickam, H., Golemund, G. *R for the Data Science*. O'Reilly, <https://r4ds.had.co.nz>.

Lander, Jared P. *R for Everyone: Advanced Analysis and Graphics*. Boston etc.: Pearson Education, Inc, 2017.

Last update, February 2025