

## **APPLIED DATA ANALYTICS**

**Academic Year:** 2023/2024

**Semester:** 2<sup>nd</sup>

**ECTS:** 7.5

**LEVEL OF CURRICULAR UNIT: UNDERGRADUATE** (1st cycle, as defined in the Framework of Qualifications for the European Higher Education Area)

---

### **INSTRUCTOR(S)**

Antonio Fidalgo (theory lectures); Miguel Salema (practical sessions)

### **CONTACTS AND OFFICE HOURS**

Email: [antonio.fidalgo@ucp.pt](mailto:antonio.fidalgo@ucp.pt), [misalema@ucp.pt](mailto:misalema@ucp.pt)

Office: 5322A

Office hours: office hours by appointment

### **BIOGRAPHY**

Antonio Fidalgo is an Affiliate Professor at the CATÓLICA-LISBON School of Business & Economics. He has a M.A. in Economics (Universitat Pompeu Fabra, Spain) and a Ph.D. in Economics (Lausanne University, Switzerland). He was a Lecturer at Fresenius University of Applied Sciences, (Germany) and taught at Boston University (USA) and the University of Magdeburg (Germany). His research focuses on long run economic development, employing a quantitative empirical approach and informed by economic theory. Outside academia, he worked as a consultant in data analysis.

---

### **COURSE OVERVIEW**

This course introduces quantitative methods used to extract information from data with the ultimate goal of improving managerial decisions. The course's approach is twofold. First, it develops concepts and methods that are useful in the current data-oriented business environment: from data collection to statistical models of data interpretation. Second, it introduces software tools to carry out analyses of real-world datasets and to report results.

---

### **LEARNING OBJECTIVES**

This intermediate course offers a modern approach to applied statistics with a three-fold objective.

- To build and expand the students' knowledge of the fundamental concepts in statistics in order to carry sound empirical analyses. These concepts include, for instance, sample selection issues, causality claims, tests of hypotheses, errors in variables and proxies, data collection and tidying, reproducible research, etc.
- To present and apply the most common techniques for regression analysis as well as classification problems.
- To introduce students to the free software R as a tool for data analysis. The structure and the scope of this introduction to R will follow and complement the treatment of the theoretical concepts. The aim is to allow students to carry statistical analyses on real-world data as well as to produce and learn from simulated data.

---

## TEACHING AND LEARNING METHODOLOGY

The learning methodology is based on class lectures and homework. The lectures describe a structured set of selected relevant concepts along with illustrative examples. Students are invited to participate in class discussions as a way to successfully master the material. The homework consists in assignments based on the topics covered in the lectures and requiring coding in the software R.

---

## REQUIRED BACKGROUND

Students must have attended an introductory class in statistics covering, among others, fundamentals concepts in descriptive statistics, probability theory and random variables.

---

## ASSESSMENT

Assignments and/or quizzes	group and/or individual grade	33.3%
Midterm exam	individual grade	33.3%
Endterm exam	individual grade	33.3%

Necessary conditions to be waived from the final exam:

- Weighted average  $\geq 9.5$ , grade  $\geq 7.5$  in both the midterm and the endterm exams, and grade  $\geq 9.5$  in at least one of them.

Necessary requisites to be admitted to the final exam:

- Weighted average  $\geq 7.5$ , and average of the midterm and the endterm exams  $\geq 6.5$ .

Midterm and Endterm: Each test covers half of the class contents. Concepts that were developed in the first half and used again in the second half are part of the content for the endterm. The tests are closed book and closed notes. For the midterm, students may bring one A4, two-sided sheet with formulas and notes. For the endterm, students may bring two A4, two-sided sheets with formulas and notes.

Final Exam: The final exam will cover the entire course's program. The exam is closed book and closed notes. Students may bring two A4, two-sided sheets with formulas and notes. The final grade of students admitted to the final exam will be the weighted average of the continuous evaluation (50%) and the final exam (50%). A final exam grade below 9.5 will result in the student failing the course. Students with grades between 9.5 and 12 in the final exam will have this grade as final grade if it is higher than the weighted average indicated above.

Class Participation: Students are encouraged to actively engage with the course and attend and participate in both theoretical and practical classes. Class attendance may be monitored. A student's class participation and engagement with the course can be considered in border situations.

Grade improvements: Students who pass the course may take the final exam to in order to improve their grade. The final grade of these students will be the weighted average of the continuous evaluation (50%) and the final exam (50%). Students have 15 minutes at the beginning of the final exam to decide whether they want to take it. Once this decision is made, it is final, and the student's exam will be graded.

---

## COURSE CONTENT

1. Random sampling and sampling distributions
2. Further statistical concepts (sample bias, endogeneity, regression to the mean, etc.)
3. Introduction to R, RStudio and qmd files

4. Model construction
5. Hypothesis testing
6. Regression analysis
7. Classification
8. Further statistical models and concepts (time allowing)

---

## BIBLIOGRAPHY

### Required readings:

Lecture notes (online, including pdf version).

### Recommended books:

Diez, D., Çetinkaya-Rundel, M., & Barr, C. (2019), *OpenIntro Statistics*. <https://www.openintro.org/book/os>

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Sage publications.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2023). *An introduction to statistical learning*. New York: Springer. (pdf free version online)

### Online materials:

(TBA, if any)

**Extra Costs (case studies, platforms...):** None.

---

## ADDITIONAL RESOURCES

Bootstrap is a curated collection of resources, techniques, and personal development tools from academic sources, thought-leaders, and well-established productivity practices. [bootstrap - Productivity & Study Resources | CATÓLICA-LISBON \(ucp.pt\)](#)

---

## CODE OF CONDUCT AND ETHICS

Católica Lisbon School of Business and Economics is a community of individuals with diverse backgrounds and interests who share certain fundamental goals. A crucial element to achieve these goals is the creation and maintenance of an atmosphere contributing to learning and personal growth for everyone in the community. The success of CATÓLICA-LISBON in attaining its goals and in maintaining its reputation of academic excellence depends on the willingness of its members, both collectively and individually, to meet their responsibilities.

Along with all the other members of our community, students are expected to follow professional standards and CATÓLICA-LISBON standards of Academic Integrity. Some details should be mentioned here: Please arrive on time for class with uninterrupted attendance for the duration of the class. Signing attendance sheet for anyone else in the class constitutes fraud and a violation of the CLSBE code of conduct. Use of computers and other electronic devices during the class is not allowed, unless expressly requested by the instructor of the course. Students who persistently act in a disruptive and disrespectful manner during the class session may be invited to leave.

Students are expected to behave at all times according to the fundamental principles of academic integrity, including honesty, trust, fairness, respect, and responsibility. In particular,

- a. In **individual graded assignments** of any type, students may not collaborate with others or use any materials without explicit permission from the instructor of the course;
- b. In **group assignments** and reports, all students listed as authors should have performed a substantial amount of work for that assignment;

- c. It is dishonest to fabricate or falsify data in experiments, surveys, papers, reports or other circumstances; fabricate source material in a bibliography or “works cited” list; or provide false information in other documents in connection with academic efforts;
- d. **Plagiarizing**, i.e. “to steal and pass off the ideas or words of another as one’s own and or to use another’s production without crediting the source” (Merriam-Webster Dictionary) is an Academic Integrity breach. It can be avoided by using proper methods of documentation and acknowledgement. Visit this guide for additional resources on how to avoid plagiarism in your written submissions <http://en.writecheck.com/plagiarism-guide>
- e. In **exams** students must not receive or provide any unauthorized assistance. During an examination, students may use only material and items authorized by the faculty. Use of smartwatches or other communication devices is not permitted during the exam.

Academic integrity breaches will be dealt with in accordance with the school’s code of Academic Integrity: <https://www.clsbe.lisboa.ucp.pt/system/files/assets/files/academicintegritycode.pdf>

---

## APPENDIX

The empirical analyses will be carried on a computer. Under request, students should be ready to bring a personal laptop to class.

