UAB
Universitat Autònoma de Barcelona

Data Visualisation and Modelling

Code: 43482
ECTS Credits: 6

2025/2026

Degree	Type	Year
Modelización para la Ciencia y la Ingeniería / Modelling for Science and Engineering	OP	1

Contact

Name: Pedro Puig Casado

Email: pere.puig@uab.cat

Teachers

Maria Rosa Camps Camprubi

Giulia Binotto

Juan Ramon Gonzalez Ruiz

Teaching groups languages

You can view this information at the end of this document.

Prerequisites

An elementary knowledge in Probability Theory and Statistical Inference.

Objectives and Contextualisation

Course of R . All the practical exercises will be solved using the statistical package R. This introductory course is basic for all the posterior developements.

Visualization of large-scale datasets with R. GViz, Maps and Tabplot.

Data Simulation, Boostrapping and Permutation testing. These methodologies allow a fast solution to complex statistical models without a deep knowledge of the general and classical statistical topics. They are indispensable tools in the current statistical modelling techniques. The students will complete a basic training program, including the use of an appropriate software, and they will learn how to attack several real data analysis problems.

Bayesian networks. In the opinion of many researchers one of the most significant contribution in AI in this century, are graphical structures for representing the probabilistic relationships among a large number of variables and for doing probabilistic inference with those variables, with a huge number of application fields. One of the objectives of this course is to introduce them and develop in students some skill in their use in modelling, both from a theoretical and applied point of view, with particular emphasis on the use of appropriate software.

Learning Outcomes

1. CA31 (Competence) Apply statistical and computational modelling tools to problems in the business or research field.
2. CA32 (Competence) Integrate statistical methods with other modelling tools in the context of multidisciplinary projects.
3. CA32 (Competence) Integrate statistical methods with other modelling tools in the context of multidisciplinary projects.
4. KA23 (Knowledge) Identify the programming languages and environments specific to the field of statistical modelling and Artificial Intelligence.
5. KA23 (Knowledge) Identify the programming languages and environments specific to the field of statistical modelling and Artificial Intelligence.
6. KA24 (Knowledge) Recognise Bayesian network techniques and the benefits they offer in each specific modelling area.
7. SA29 (Skill) Use specific software to solve problems in statistical modelling and data processing.
8. SA30 (Skill) Apply appropriate statistical techniques to construct models that respond to particular problems.
9. SA31 (Skill) Interpret the parameters that make up a statistical model to describe a specific situation.
10. SA32 (Skill) Interpret the results by applying a specific statistical model.
11. SA33 (Skill) Analyse the predictions obtained by applying a specific statistical model.

Content

Part 1: Introduction to R (6h)

Part 2: Visualization of large-scale datasets with R (6h)

Part 3: Bayesian Networks (13h)

1) Block 1: Basics.

2) Block 2: Causal networks and Inference in Bayesian networks.

3) Block 3: Learning Bayesian network parameters.

Part 4: Data Simulation, Boostrapping and Permutation testing (13h)

1) Permutation tests.

2) Jackknife.

3) Parametric Bootstrap.

4) Non-parametric Bootstrap.

Activities and Methodology

Title	Hours	ECTS	Learning Outcomes
Type: Directed			

Exercises	16	0.64
Lectures	38	1.52
Projects + Assigments	18	0.72
Type: Supervised		
Practical sessions	20	0.8

In this course lectures, in which the determining factor is the teacher's explanation, are the basis of the learning process. It is also very important the participation of the students, combined with practical sessions in which it is the student him/herself who must use the knowledge to solve problems.

Annotation: Within the schedule set by the centre or degree programme, 15 minutes of one class will be reserved for students to evaluate their lecturers and their courses or modules through questionnaires.

Assessment

Continous Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Daily homework	50	38	1.52	CA31, CA32, KA23, KA24, SA29, SA30, SA31, SA32, SA33
Projects	50	20	0.8	CA31, CA32, KA23, KA24, SA29, SA30, SA31, SA32, SA33

The evaluation of the course consists in a continuous assessment.

There are 4 assessments during the course, weighted as 15.8%, 15.8%, 34.2% i 34.2% corresponding to each part.

Each professor will explain his or her own type of assessment.

Part 1 assessment: Daily homework + final project (individual simple real data analysis with R).

Part 2 assessment: Daily homework + final project.

Part 3 assessment: Daily homework + delivery of some exercises + final project.

Part 4 assessment: Daily homework + delivery of some exercises done in practical sessions.

Bibliography

- Resampling methods: a practical guide to data Analysis. Phillip I. Good, 2006.
- The jackknife, the bootstrap and other resampling plans. Bradley Efron, 1982.
- Bootstrap methods and their application. A.C. Davison, D.V. Hinkley, 1997.
- "Learning Bayesian Networks" by R. E. Neapolitan, Prentice Hall Series in Artificial Intelligence, 2004.

- "Probabilistic Methods for Bioinformatics with an Introduction to Bayesian Networks" by R. E. Neapolitan, Elsevier, 2009.

Software

The R programming language will be utilized.

Groups and Languages

Please note that this information is provisional until 30 November 2025. You can check it through this link. To consult the language you will need to enter the CODE of the subject.

Name	Group	Language	Semester	Turn
(TEm) Theory (master)	1	English	first semester	afternoon