UAB
Universitat Autònoma de Barcelona

Data Visualisation and Modelling

2026/2027

Code: 43482
Credits: 6

Degree programme	Type	Course
Modelling for Science and Engineering	OP	1

Contact lecturer

Name : Alejandra Cabaña Nigro

Email : anaalejandra.cabana@uab.cat

Teaching staff

Juan Ramón Gonzalez Ruiz

Group languages

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

Prerequisites

An elementary knowledge in Probability Theory and Statistical Inference and programing language R.

Objectives

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

Bayesian networks. Bayesian networks constitute, according to many researchers, one of the most important contributions to modern Artificial Intelligence. They are probabilistic models that represent the conditional dependence relationships between a set of variables. These models combine a graphical structure, represented by a directed acyclic graph, with a set of conditional probabilities that determine their behavior. Bayesian networks can be learned from data using Machine Learning techniques and have proven to be very useful tools in numerous scientific and technological fields. Among their applications, classification stands out, although they are also used for prediction, diagnosis, and decision-making under uncertainty. In this section, the fundamental concepts of Bayesian networks will be introduced, as well as the main methods of probabilistic inference and parameter and structure learning. Finally, their use as probabilistic classifiers in various real-world problems will be studied. The emphasis will be on both understanding the theoretical foundations and their correct practical application through the use of real data and specialized R packages.

Data Simulation, Boostrapping and Permutation testing. The main goal is to train students in using the plug-in principle and resampling methods to perform precise statistical inference when underlying distributions are unknown. Students learn to estimate bias and standard errors and construct advanced confidence intervals that outperform traditional asymptotic normality, particularly with censored, dependent, or non-normal data.

Furthermore, the course explores the design of permutation tests to achieve exact significance levels by conditioning on observed data. Finally, the course covers theoretical foundations which are essential for evaluating second-order accuracy and the consistency of these methods for complex functionals.

Learning outcomes

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

Contents

Part 1: Visualization of large-scale datasets with R (6h)

Part 2: Bayesian Networks (16h)

1) Basics.

2) Causal networks and Inference in Bayesian networks.

3) Learning Bayesian network parameters ans structure.

4) Bayesian networks as classifiers.

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

1) Permutation tests.

2) Parametric Bootstrap.

3) Non-parametric Bootstrap.

Learning activities and methodology

Title	Hours	ECTS	Learning outcomes
Lectures	38	1.52	
Exercises	16	0.64	
Practical sessions	20	0.8	
Projects + Assigments	18	0.72	

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

Continuous assessment activities

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

The evaluation of the course consists in a continuous assessment.

There will be 3 assessments during the course, weighted as 16%, 42% and 42% corresponding to each blo.

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

Part 1 assessment: Daily homework + final project.

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

Part 4 assessment: Deelivery problems and exercises done in practical sessions.

Use of Artificial Intelligence (AI)

In this course, the use of Artificial Intelligence (AI) technologies is allowed only for support tasks, such as:

- bibliographic or information searches,
- text or code corrections,
- translations.

Students must clearly identify which parts were generated using AI, specify the tools used, and include a critical reflection on how these tools influenced both the process and the final result of the activity.

Lack of transparency in the use of AI in assessed activities will be considered a breach of academic integrity and may lead to partial or full grade penalties, or more serious sanctions in severe

cases.

Use of Artificial Intelligence (AI)

In this course, the use of Artificial Intelligence (AI) technologies is allowed only for support tasks, such as:

- bibliographic or information searches,
- text or code corrections,
- translations.

Students must clearly identify which parts were generated using AI, specify the tools used, and include a critical reflection on how these tools influenced both the process and the final result of the activity.

Lack of transparency in the use of AI in assessed activities will be considered a breach of academic integrity and may lead to partial or full grade penalties, or more serious sanctions in severe cases.

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

R Core Team (2026). R: A language and environment for statisticalcomputing. R Foundation for Statistical Computing, Vienna, Austria.

URL https://www.R-project.org/

Course groups and languages

The information provided is provisional until November 30. After this date, you will be able to consult the language of each group through this link. To access the information, you will need to enter the course CODE

Type of teaching	Group	Language	Semester	Shift
(TEm) Theory (master)	1	English	first semester	afternoon