

Scenarios and Uses of Big Data

Code: 104749
ECTS Credits: 6

2025/2026

Degree	Type	Year
Interactive Communication	OB	4

Contact

Name: Cristina Pulido Rodriguez
Email: cristina.pulido@uab.cat

Teaching groups languages

You can view this information at the [end](#) of this document.

Prerequisites

It is recommended to have completed the course "Introduction to Big Data" in order to better understand and carry out this course.

Objectives and Contextualisation

The aim of the course is to study and practice the handling of large databases, complemented by real use cases. Various techniques and functions will be explored to enable students to clean and analyze databases regardless of their size. The course also embraces the open data paradigm as a working system with multiple possibilities for practical application.

Competences

- Act with ethical responsibility and respect for fundamental rights and duties, diversity and democratic values.
- Act within one's own area of knowledge, evaluating sex/gender-based inequalities.
- Determine and plan the technological infrastructure necessary for the creation, storage, analysis and distribution of interactive multimedia and social-networking products.
- Introduce changes in the methods and processes of the field of knowledge to provide innovative responses to the needs and demands of society.
- Manage time efficiently and plan for short-, medium- and long-term tasks.
- Promote and launch new products and services based on massive-scale mining and analysis of data from the Media.
- Search for, select and rank any type of source and document that is useful for creating messages, academic papers, presentations, etc.
- Students must be capable of applying their knowledge to their work or vocation in a professional way and they should have building arguments and problem resolution skills within their area of study.
- Students must be capable of communicating information, ideas, problems and solutions to both specialised and non-specialised audiences.

- Students must develop the necessary learning skills to undertake further training with a high degree of autonomy.
- Take account of social, economic and environmental impacts when operating within one's own area of knowledge.

Learning Outcomes

1. Analyse a situation and identify its points for improvement.
2. Communicate using language that is not sexist or discriminatory.
3. Critically analyse the principles, values and procedures that govern the exercise of the profession.
4. Cross-check information to establish its veracity, using evaluation criteria.
5. Describe the infrastructure needed to store big data.
6. Differentiate between the various types of existing architectures for working with big data.
7. Distinguish the salient features in all types of documents within the subject.
8. Evaluate the impact of problems, prejudices and discrimination that could be included in actions and projects in the short or medium term in relation to certain people or groups.
9. Explain the characteristics of the infrastructure needed to recover big data.
10. Explain the explicit or implicit deontological code in your area of knowledge.
11. Explain the infrastructure needed to process big data.
12. Extract large volumes of data from social networks and the new digital media in particular.
13. Identify situations in which a change or improvement is needed.
14. Identify the social, economic and environmental implications of academic and professional activities within one's own area of knowledge.
15. Plan and execute academic projects in the field of big data.
16. Propose new methods or well-founded alternative solutions.
17. Propose projects and actions that are in accordance with the principles of ethical responsibility and respect for fundamental rights and obligations, diversity and democratic values.
18. Propose projects and actions that incorporate the gender perspective.
19. Propose viable projects and actions to boost social, economic and environmental benefits.
20. Share experiences with the group as a path to learning, in order to work subsequently in multidisciplinary groups.
21. Solve basic problems in big data.
22. Submit course assignments on time, showing the individual and/or group planning involved.
23. Weigh up the risks and opportunities of both one's own and other people's proposals for improvement.

Content

1. Introduction to R and RStudio
2. Fundamentals of descriptive statistics
3. Types of databases
4. Open data acquisition and web scraping
5. Data cleaning and transformation
6. Text processing with regular expressions
7. Data visualization
8. Web application development with RShiny

Activities and Methodology

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Lab practices	18	0.72	3, 1, 20, 2, 4, 5, 7, 6, 10, 11, 9, 12, 14, 13, 15, 23, 22, 16, 17, 18, 19, 21, 8
SEMINARS	15	0.6	1, 4, 5, 7, 6, 11, 12, 13, 8
THEORETICAL SESSIONS	15	0.6	3, 2, 5, 7, 11, 9, 8
Type: Supervised			
ASSESSMENT	8.5	0.34	20, 2, 4, 5, 7, 12, 15, 22, 17, 18, 21, 8
TUTORING	8.5	0.34	1, 20, 2, 14, 13, 15, 23, 22, 16, 17, 18, 19, 21, 8
Type: Autonomous			
OTHER ACTIVITIES (study time; practice preparation; seminar preparation, readings, etc.)	50	2	3, 1, 2, 4, 5, 7, 6, 10, 11, 9, 12, 14, 13, 15, 22, 17, 18, 21, 8

The methodology is based on the following activities:

Theoretical sessions: Introduction theoretical sessions to concepts

Laboratory practices: individual or team works in which practical activities are carried out with one task with time limit. Students must apply knowledge, distribute time and prepare the submission within the classroom and in the hours spent in practice under the professor's guidance.

Seminars: individual or teamwork in which more extensive practical activities are carried out and with tasks open to student creativity. There are no limited time in the classroom, but deadlines for submission. Students must apply knowledge, distribute time and prepare submissions by starting their work within the classroom, but continue it in the form of activities supervised by the professor's team.

Course final work: practical group assessment exercise in which students must solve, during course development, a practical application problem linked to the subject's objectives. Students must raise the problem and perform the four processes to provide a solution based on large amounts of data: search, extraction, analysis and publication of data report including a proposal for a decision based on the information collected and analysed.

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
PRACTICAL SESSIONS	40%	15	0.6	3, 1, 20, 2, 4, 5, 7, 6, 10, 11, 9, 12, 14,

				13, 15, 23, 22, 16, 17, 18, 19, 21, 8
STUDENT'S PARTICIPATION	10%	8	0.32	3, 1, 20, 2, 4, 5, 7, 6, 10, 11, 9, 12, 14, 13, 15, 23, 22, 16, 17, 18, 19, 21, 8
SUBMISSION AND PRESENTATION OF THE COURSE PROJECT	50%	12	0.48	1, 20, 2, 5, 6, 11, 9, 12, 13, 15, 23, 22, 16, 17, 18, 19, 21, 8

Activity A. Course project and oral presentation (group) . 50% of the final grade.

Activity B: Laboratory Practice. 40% of the final grade.

Activity C. Student participation. 10% of the final grade.

To approve the subject, it is necessary to get a minimum approval note (5,0) in activities A and B.

RE-EVALUATION:

In the last two weeks of the course, students who have not pass the course can participate in a re-evaluation process consisting of a theoretical test and a practical exercise. Students must have done at least 2/3 of the total course practices (activities grouped on B) and must have obtained an average grade equal to or greater than 3.5 (and less than 5) in all evaluation activities.

PLAGIARISM:

In the event that the student performs any irregularity that may lead to a significant variation of an evaluation act, this evaluation act will be graded with 0, regardless of the disciplinary process that could be instructed. In the event, that several irregularities occur in the evaluation acts of the same subject, the final grade for this subject will be 0.

The use of AI

Students are allowed to use artificial intelligence; however, the material provided on the course's virtual campus already contains the necessary knowledge to complete the assigned tasks without the need to consult external sources. Nevertheless, if phrases such as "Here is the text you requested" or other expressions are detected that suggest the text was copied and pasted directly from an AI tool-and thus the work was not reviewed before submission-the assignment or project will receive a grade of 0.

Bibliography

Fernández-Avilés, Gema; Montero, José-María; et al. (2024) Fundamentos de ciencia de datos con R. Editorial McGraw-Hill. Disponible a: <https://cdr-book.github.io/index.html>

Casas Roma, Jordi (2019) *Big data: análisis de datos en entornos masivos*. Barcelona: Editorial UOC.

Duran, Xavier (2019). *El imperio de los datos: el big data, la privacidad y la sociedad del futuro*. PUV Publicacions, Universitat de València: Càtedra de Divulgació de la Ciència, UCC+i, Unitat de Cultura Científica i de la Innovació, Universitat de València.

Dur Lahoz-Beltrá, Rafael (2019). En las entrañas del big data: una aproximación a la estadística. Emse Edapp, S.L.

Fuller A. (2012). The White Book of Big Data. The definitive guide to the revolution in business analytics. Fujitsu. <https://www.fujitsu.com/rs/Images/WhiteBookofBigData.pdf>

Software

This is mostly a practical course, thus the required software is the usual one for the performance of capture, processing and analysis tasks in different formats.

Specifically, the following tools are required:

Data analysis software: Excel or similar

Data visualisation software: Infogram - Datawrapper - Flourish

Multimedia editing software: Wordpress - Blogger - Wix

As the subject will carry out practical sessions during all its activities, it is recommended that students (if possible) always bring their laptop to the sessions.

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
(PLAB) Practical laboratories	61	Catalan	first semester	afternoon
(PLAB) Practical laboratories	62	Catalan	first semester	afternoon
(TE) Theory	6	Catalan	first semester	afternoon