

**Advanced Methods of Data Processing and
Management**

Code: 104377
ECTS Credits: 6

Degree	Type	Year	Semester
2503758 Data Engineering	OT	4	1

Contact

Name: Daniel Franco Puentes

Email: daniel.franco@uab.cat

Teaching groups languages

You can check it through this [link](#). To consult the language you will need to enter the CODE of the subject. Please note that this information is provisional until 30 November 2023.

Teachers

Sandra Adriana Mendez

Antonio Miguel Espinosa Morales

Pedro Luis Pons Pons

Javier Panadero Martinez

Prerequisites

This subject has no PRE-REQUIREMENTS. It is recommended to have completed the following subjects:

[104357 - Computation in Cloud Environments](#)

[104365 - Data Visualisation](#)

[104358 - Development of Big Data Applications](#)

[104362 - Neural Networks and Deep Learning \(2021-22\)](#)

Objectives and Contextualisation

The objective of this course is to know the advanced methods and concepts of processing and management of massive data, both from the point of view of generation, adaptation, transmission and storage and, also, the processing and analysis to extract useful information. It is also a goal to use the right tools to carry out the work with massive data, both interactively and locally as well as batch and remote and deferred and real-time.

Competences

- Conceive, design and implement efficient and secure data storage systems.
- Prevent and solve problems, adapt to unforeseen situations and take decisions.
- Students must be capable of collecting and interpreting relevant data (usually within their area of study) in order to make statements that reflect social, scientific or ethical relevant issues.
- Students must develop the necessary learning skills to undertake further training with a high degree of autonomy.
- Work cooperatively in complex and uncertain environments and with limited resources in a multidisciplinary context, assuming and respecting the role of the different members of the group.

Learning Outcomes

1. Prevent and solve problems, adapt to unforeseen situations and take decisions.
2. Students must be capable of collecting and interpreting relevant data (usually within their area of study) in order to make statements that reflect social, scientific or ethical relevant issues.
3. Students must develop the necessary learning skills to undertake further training with a high degree of autonomy.
4. Study the adaptations made to data analysis and consultation algorithms in order for these to maintain the privacy of the entry data, of the models learnt, or of the outputs of the models used in the field of business intelligence.
5. Work cooperatively in complex and uncertain environments and with limited resources in a multidisciplinary context, assuming and respecting the role of the different members of the group.

Content

1. Advanced methods of big data processing and management
2. Reference Cloud architectures for big data management and application design: batch, streaming, decentralized.
3. Analysis of challenges and case studies with cloud-oriented and big data
4. MVP methodology for the design and development of Cloud solutions to manage big data
5. Tools for evaluating applications and big data management systems

Methodology

Following the methodology of challenge based learning, the subject is based on a coordinated set of practical works that lead to the elaboration and presentation of a proposal of technical solution of the problem posed by the actors proposing the challenge, finally contributing with some proposals of solution to solve them.

The task will be carried out in small groups of students, who can be grouped flexibly depending on the dynamics of work. The practical activity will be accompanied by a set of theoretical and methodological support sessions, as well as the tutoring of the whole learning process.

The stages of the challenge learning methodology are:

1. Discovery

The first phase involves a double recognition:

- a) on the one hand, of the field of application and the problematic of study for each one of the groups of work, coming in contact with the first actors of the subject;
- b) on the other, the tools necessary to properly carry out further research: preparation of the diagnosis, and the procedure for drawing up a Development Plan.

2. Research

Research should follow the process of analyzing the three parts of a development plan: understanding demand, supply, and needs. This stage also has a double aspect:

- a) On the one hand, the general documents of the proposed challenge are explored.
- b) On the other hand, contact is made with the necessary resources and services, conducting interviews with their managers and visiting facilities, observing users and the environment.

3. Identification of needs and proposals for improvement

The third phase is characterized by the development of a prototype for obtaining results, accurately identifying the shortcomings in each of the sectors, and coming to design proposals for improvement at an initial level for the challenge analyzed by each group.

These activities will have the reinforcement of conferences and workshops aimed at obtaining creative results.

4. Presentation of Results

Communicating the results is an essential step in the process. It will be done in three formats:

Report of the proposal
Posters
Oral presentations

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.

Activities

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Analysis, design and development of prototypes	25	1	4, 1, 3, 2, 5
Lectures on technology, methodology and the case study	20	0.8	4, 2
Type: Autonomous			
Self-study	100	4	4, 1, 3, 2, 5

Assessment

The assesment will take into account:

- a) The final results of the process of elaboration of the Design Report of the proposed solution (Written report, oral presentation and poster), where the procedure and the achievement in the resolution of the posed challenge will be evaluated.

b) The gradual learning process, based on three follow-up reports (incremental MVPs).

The subject follows a continuous learning and assessment schedule that must be followed on time. The delivery dates of the tasks must be respected. Delay in deliveries will result in a penalty.

The grade of the subject will be the average of the grades obtained in the different items evaluated. Failure to complete any of the items implies that the subject is "Not assessable". To be able to average, you must have obtained at least a 4 in each of the evaluable items.

Recovery: Recovery requires that all items requested in the assessment have been submitted.

The follow-up exercises and the poster will be recoverable items. They may be recovered, respectively, with another examination, and with the revision of the failed poster. Due to its nature, the Written Report and the oral presentations will not be retrieved.

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

Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Final report	30	2	0.08	4, 1, 3, 2, 5
MVP reports	50	2	0.08	4, 1, 3, 2, 5
Poster and oral presentation	20	1	0.04	4, 1, 3, 2, 5

Bibliography

- Dan C. Marinescu. "Cloud Computing. Theory and Practice". Morgan-Kaufmann. 2018.

-AWS Certified Cloud Practitioner Study Guide; Ben Piper, David Clinton; Sybex (14 de junio de 2019); ISBN-10: 1119490707, ISBN-13: 978-1119490708

-Infrastructure as Code; Kief Morris; O'Reilly Media; 1 edition (June 17, 2016); ISBN-10: 1491924357, ISBN-13: 978-1491924358

-Amazon Web Services in Action, 2E; Andreas Wittig, Michael Wittig; Manning Publications; Edición: 2nd edition (30 de septiembre de 2018); ISBN-10: 1617295116, ISBN-13: 978-1617295119

-Microsoft Azure Essentials - Fundamentals of Azure, 2nd Ed; Michael Collier, Robin Shahan; 2016; https://download.microsoft.com/download/6/6/2/662DD05E-BAD7-46EF-9431-135F9BAE6332/9781509302963_I

-Mastering Cloud Computing : Foundations and Applications Programming. Buyya, Rajkumar;Vecchiola, Christian;y más Elsevier Science & Technology 2013. ISBN: ISBN number:9780124114548, ISBN number:9780124095397

This is Service Design Thinking: Basics - Tools - Cases; Marc Stickdorn; BIS Publishers; ISBN 9063692560 (ISBN13: 978906369256) 2012

Design Research: Methods and Perspectives Brenda Laurel; The MIT Press; 2003; ISBN-10 : 0262122634; ISBN-13 : 978-0262122634

Gamestorming: A Playbook for Innovators, Rulebreakers, and Changemakers; Dave Gray, Sunni Brown, James Macanufo; O'Reilly Media; ISBN-10 : 0596804172 ISBN-13 : 978-0596804176; 2010

The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses; Eric Ries ; Currency; ISBN-10 : 9780307887894, ISBN-13 : 978-0307887894; 2011

Software

Visual Studio Code

Apache Spark

Redis

Power BI

Qlick

Azure Cloud

AWS Cloud

Google Cloud