# Teaching guide

## IDENTIFICATION DETAILS

| | |
|---|---|
| Degree: | Business Analytics |
| Field of Knowledge: | Social and Legal Science |
| Faculty/School: | Law, Business and Governance |
| Course: | BIG DATA II: STORAGE ANALYSIS |
| Type: | Optional |
| ECTS credits: | 3 |
| Year: | 2 |
| Code: | 5349 |
| Teaching period: | Fourth semester |
| Area: | Big Data |
| Module: | Disciplinary Training |
| Teaching type: | Classroom-based |
| Language: | English |
| Total number of student study hours: | 75 |

| Teaching staff | E-mail |
|---|---|
| Alfonso Antolínez García | |

## SUBJECT DESCRIPTION

The subject of Big Data II: Storage Analysis informs the student in the characteristics of data storage in the era of Big Data, entering in turn into their different storage systems. During the development of the classes, first of all, the characteristics of big data storage will be studied. It will continue with a reference to the characteristics of storage in Business Intelligence systems to finally continue with the study of the different storage systems in Big Data.

## GOAL

Acquire significant and critical learning in the detailed knowledge of the different architectures, methodologies, techniques and solutions of massive data storage (BigData), their processes, entities, design parameters and associated technological implications.

## PRIOR KNOWLEDGE

It is recommended to have studied the subject Big Data I: Infrastructure, of the Degree in Business Analysis.

## COURSE SYLLABUS

1. Types of data storage depending on their nature: Digital Data, Digital Store, Database/DBMS.
2. Types of storage depending on the architectures of networks and systems: local/on premise, Cloud.
3. Mass data storage techniques: scale-out, scale-up.
4. Distributed storage systems: types, implicit risks, design parameters, efficiency, availability and reliability (MTBF, MTTR, Failure Rate, etc.)
5. Data Warehouse: nature, processes, types and architecture. MDX syntax and language.
6. Hadoop Distributed File System (HDFS). MapReduce architecture.
7. Storage ecosystems in BigData. Datalakes.

## EDUCATION ACTIVITIES

The methodology followed in this subject is aimed at achieving a significant learning by the student of the concepts

and fundamental techniques of the subject. For this reason, exhibition and interactive sessions are combined with the students, with practical sessions and presentations of results / conclusions of the same, both individually and in groups.
In this way, student participation and student-teacher and student-student interaction are achieved as a way to promote collaborative learning and self-learning capacity. In some cases, the student will have to make in class the presentation of the main conclusions of his study or work, which will allow the exchange of knowledge and experiences between students.
Priority will be given to the pedagogical techniques of Problem-Based Learning (ABP) and "Flipped-Learning". The face-to-face work will be completed with autonomous work by the student, in some cases developed in a group, so that cooperative learning is encouraged.
Finally, in order to facilitate the student's access to the materials and the planning of their work, as well as communication with the teacher and the rest of the students, LMS platform will be used: Virtual Classroom (CANVAS), which is a learning platform that offers different electronic resources to complement, in a very significant way, the student's learning.All the study and work carried out by the student will be supervised and guided by the teacher through tutoring, individually or in a group.

## DISTRIBUTION OF WORK TIME

| CLASSROOM-BASED ACTIVITY | INDEPENDENT STUDY/OUT-OF-CLASSROOM ACTIVITY |
|---|---|
| 30 hours | 45 hours |
| Expository lesson 11h<br>Practical classes 11h<br>Tests/Practices/Works 8h | Study and individual work 30h<br>Group work 15h |

## SKILLS

### Basic Skills

Students must have demonstrated knowledge and understanding in an area of study that is founded on general secondary education. Moreover, the area of study is typically at a level that includes certain aspects implying knowledge at the forefront of its field of study, albeit supported by advanced textbooks

Students must be able to apply their knowledge to their work or vocation in a professional manner and possess skills that can typically be demonstrated by coming up with and sustaining arguments and solving problems within their field of study

Students must have the ability to gather and interpret relevant data (usually within their field of study) in order to make judgments that include reflections on pertinent social, scientific or ethical issues

Students must be able to convey information, ideas, problems and solutions to both an expert and non-expert audience

Students must have developed the learning skills needed to undertake further study with a high degree of independence

**General Skills**

Capacity for organising, systematization and planning in identifying problems, levers and models in the context of big data.

Capacity for analysing data on a large scale from different sources: audiovisual, textual and numerical.

**Specific skills**

Know and understand the basic concepts of Big Data and its most characteristic elements.

Understand and use advanced statistics, data storage, relational databases, non-relational databases and big data management systems tools to an analyst level.

## LEARNING RESULTS

Understand the characteristics of data storage in the Big Data era.

Gain insight into the evolution of data storage in analytics systems.

Get to know the different Big Data data storage systems.

## LEARNING APPRAISAL SYSTEM

The continuous evaluation system includes four types of tests: Final exam (50%) + Resolution of exercises (20%) + Individual/Group Case Studies (20%) + Class Participation (10%)
•[1] Final theoretical-practical written exam: presents a weight of 50% in the final grade. The format of the same may contain short questions, development questions, resolution of practical assumptions and / or questions type of test of different typology: multiple answer, single answer, True / False, etc.
•[2,3] Class tests, practices, resolution of practical cases and other works related to the subject both individual and group: presents a weight of 40% in the final grade (distributed as follows: resolution of exercises (20%); preparation, resolution of individual/group ABP case studies and presentations (20%)
•[3] Class participation, forum interaction, cooperative learning attitude and involvement in learning (Flipped-Learning): it has a weight of 10% in the final grade.

The weighted score of the continuous evaluation shall be a value between 0 and 10 and shall be calculated as follows: 0,5*[1]+ 0,2*[2]+ 0,2*[3]+ 0,1*[4] . In the first three tests [1],[2], [3] it is necessary to obtain a minimum of 5 points out of 10 to be able to pass the subject.
The students who do not take the continuous evaluation of the subject and those students who are exempt from the obligation to attend class, either by second registration in the subject or successive, or by having express authorization from the Direction of the Degree, will be evaluated by the computation of: a theoretical-practical examination (70%) that combines all the contents and skills described in this didactic guide. The format of such a test will be similar to that stated above as [1]; and for an Individual Job (30%).
Recovery in extraordinary call: Students who have not reached the minimum grade in the ordinary evaluation may apply to the extraordinary call, evaluating all the contents and skills as described in the previous section (continuous evaluation).
The condition of Not Presented in the ordinary/extraordinary call will correspond to the non-presentation by the student to the theorico-practical tests.
IMPORTANT NOTE: Plagiarism behaviors, as well as the use of illegitimate means in the evaluation tests, will be sanctioned in accordance with the provisions of the Evaluation Regulations and the University's Coexistence Regulations.

## BIBLIOGRAPHY AND OTHER RESOURCES

### Basic


edited by Ian Foster [and four others]. Big data and social science :a practical guide to methods and tools /  2017.

José Manuel Ortega Candel. Big data, machine learning y data science en Python /  Madrid :RA-MA Editorial,2022.

Marie Lowman. A practical guide to analytics for governments :using big data for good /  2017.

Wesley W. Chu (ed.). Data Mining and Knowledge Discovery for Big Data: Methodologies, Challenge and Opportunities /  Los Ángeles :Springer,2014.