1. BASIC DATA

Subject	Project: Big Data II
Degree	Degree in Data Science
School/Faculty	School of Science, Engineering and Design
Course	3º
ECTS (European Credit Transfer and Accumulation System)	9 ECTS
Nature	Mandatory
Language(s)	English
Mode	Face-to-face
Semester	Semester 6

2. INTRODUCTION

The subject "Project: Big Data II" is a mandatory subject within the planning of the teachings of the Degree in Data Science of the European University. The subject introduces the student to the world of data processing and modelling projects, helping them to understand the development of data projects in a professional environment.

In the subject, students must learn fundamental concepts of data extraction, as well as its modelling, making use of the programming skills already acquired in previous courses, both in Python and SQL. In addition, we will raise the new problem of creating a Dashboard, helping students better understand their future role in the business world.

3. LEARNING SKILLS AND COMPETENCIES

Basic competencies:

- CB1 - That students have demonstrated possessing and understanding knowledge in an area of study that starts from the basis of general secondary education, and is usually found at a level that, although supported by advanced textbooks, also includes some aspects that involve knowledge from the forefront of their field of study.
- CB2 - That students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and Defence of arguments and the resolution of problems within their area of study.

- CB3 - That students have the ability to collect and interpret relevant data (usually within their area of study) to make judgements that include a reflection on relevant topics of a social, scientific or ethical nature
- CB4 - That students can transmit information, ideas, problems and solutions to both a specialised and non-specialised audience
- CB5 - That students have developed those learning skills necessary to undertake further studies with a high degree of autonomy

Cross-cutting competencies:

- CT1 - Ethical values: Ability to think and act according to universal principles based on the value of the person who are aimed at their full development and that entails commitment to certain social values.
- CT3 - Teamwork: Ability to integrate and collaborate actively with other people, areas and/or organisations to achieve common objectives.
- CT4 - Written communication / Oral communication: Ability to transmit and receive data, ideas, opinions and attitudes to achieve understanding and action, being oral the one that is done through words and gestures and, written, through writing and / or graphic supports.
- CT5 - Analysis and problem solving: Be able to critically evaluate information, break down complex situations into their constituent parts, recognise patterns, and consider other alternatives, approaches and perspectives to find optimal solutions and efficient negotiations.
- CT6 -Adaptation to change: Being able to accept, value and integrate different positions, adapting one's own approach as the situation requires it, as well as working effectively in situations of ambiguity.
- CT7 -Leadership: Be able to guide, motivate and guide other people, recognising their abilities and skills to effectively manage their development and common interests.
- CT8 -Entrepreneurial spirit: Ability to assume and carry out activities that generate new opportunities, anticipate problems or suppose improvements.

Specific competencies:

- CE9 - Ability to apply the criteria and mechanisms of security evaluation and certification, as well as the current legislation on personal data, privacy and fundamental rights of people.
- CE11 - Ability to apply computational learning techniques to design and implement applications and systems that use them, including those dedicated to automatic extraction of information and knowledge from large volumes of data.
- CE17 - Ability to organise, manage and defend a project in the field of data science.

Learning outcomes:

- RA1 - Carry out a team project in the field of Big Data applied to a real case that includes machine learning techniques, and the visualisation tools to be used, taking into account the relevant security and legislation aspects.
- RA2 - Prepare and interpret technical documentation on large-volume data systems.
- RA3 - Collect information related to the theme of the project and carry out an analysis and synthesis of it that allows you to understand the field of application.
- RA4 - Apply agile methodologies for the development of projects, planning the tasks to be performed, those responsible for each one, the delivery times and the planned products.

- RA5 - Hold work meetings where the student will reflect on his position and that of his colleagues by applying critical reasoning, proposing innovative solutions and developing arguments in an objective way that allow consensual decisions to be made.
- RA6 - Adapt to new situations by rethinking the starting hypotheses and reformulating them to address the final objective in the most appropriate way.
- RA7 - Generate a well-structured report that includes the extraction of conclusions including social and ethical implications.
- RA8 - Make a presentation of the results obtained to a specialised audience.
- RA9 - Defend the quality of a project for a non-specialised audience by making very short-time explanatory videos

The table below shows the relationship between the competencies developed in the subject and the learning outcomes that are pursued:

Competencies	Learning outcomes
CB1, CB2,CT3, CT4,CT6,CT7,CT8 CE9, CE11, CE17	RA1 - Carry out a team project in the field of Big Data applied to a real case that includes machine learning techniques, and the visualisation tools to be used, taking into account the relevant security and legislation aspects.
CB3, CB4 CT4 CE9, CE11, CE17	RA2 - Prepare and interpret technical documentation on large-volume data systems.
CB1 CT4, CT5 CE9, CE11, CE17	RA3 - Collect information related to the theme of the project and carry out an analysis and synthesis of it that allows you to understand the field of application.
CB1, CB2 CT3, CT4, CT7 CE9, CE11, CE17	RA4 - Apply agile methodologies for the development of projects, planning the tasks to be performed, those responsible for each one, the delivery times and the planned products.
CB4 CT1, CT3, CT5, CT7, CT8 CE9, CE11, CE17	RA5 - Hold work meetings where the student will reflect on his position and that of his colleagues by applying critical reasoning, proposing innovative solutions and developing arguments in an objective way that allow consensual decisions to be made.
CB5 CT5, CT6 CE9, CE11, CE17	RA6 - Adapt to new situations by rethinking the starting hypotheses and reformulating them to address the final objective in the most appropriate way.
CB5 CT4, CT5 CE9, CE11, CE17	RA7 - Generate a well-structured report that includes the extraction of conclusions including social and ethical implications.
CB4 CT4 CE9, CE11, CE17	RA8 - Make a presentation of the results obtained to a specialised audience.
CB4, CB5 CT6, CT7, CT8 CE9, CE11, CE17	RA9 - Defend the quality of a project for a non-specialised audience by making very short-time explanatory videos

4. CONTENTS

The subject covers the following contents:

- Machine learning environments for data analysis.
- Commercial tools for data visualisation.
- Security and Legislation.

5. TEACHING-LEARNING METHODOLOGIES

Below are the types of teaching-learning methodologies that will be applied:

- **Master Class:** Presentations made by the teacher with the necessary technological tools for maximum understanding of the concepts taught.
- **Case method:** Students acquire the knowledge by mastering the instrumentation they will need in their profession. It involves learning by doing.
- **Cooperative learning:** Students learn to collaborate with other people (other students and teachers) to solve in a creative, integrative and constructive way the questions and problems identified from the cases raised, using the knowledge and material resources available.
- **Problem-based learning:** Problems will arise with the aim of students solving them by working as a team or individually.
- **Activities based on laboratory teachings:** More autonomous, individual and group jobs, with search for information, written synthesis and debates and public defence of works.
- **Field Experiences:** Presentations by external professionals, visits to companies

6. TRAINING ACTIVITIES

Below, the types of training activities that will be carried out and the student's dedication in hours to each of them are identified:

Face-to-face mode:

Training activity	Number of hours
Master classes and practical seminars	25,5
Problem solving	7,5
Case studies and field studies	7,5
Integrative Projects	92,5
Debate and colloquium	12
Learning contract (definition of interests, needs and objectives)	3
Self-study	52,5
Tutoring	21,5
Face-to-face knowledge tests	3
TOTAL	**225**

7. EVALUATION

Below, the evaluation systems are listed, as well as their weight on the total grade of the subject:

Face-to-face mode:

Evaluation system	Weight
A. Tests to evaluate attitudes (attitude evaluation headings, class participation)	10%
B. Self-evaluation and co-evaluation tests (collaborative activities, co-evaluation activities)	15%
C. Project reports (integrating projects, final degree work)	40%
D. Exhibitions and project defences (integrating projects, final degree work)	35%

In the Virtual Campus, when you access the subject, you will be able to consult in detail the evaluation activities you must carry out, as well as the delivery dates and the evaluation procedures of each of them. Keep in mind that the evaluation procedures of each of the different activities may be specific and two activities do not have to be weighted with the same weight, and/or the evaluation criteria/rubrics may be different. For each of the activities, both the evaluation criteria and the weighting of these within the training activity block will be specified.

The evaluation process is based on the personal work of each student and presupposes the authenticity of the authorship and the originality of the exercises performed. The lack of authenticity in the authorship or originality of the evaluation tests; copying or plagiarism are irregular behaviours that can have academic and disciplinary consequences.

This subject can only be overcome from continuous evaluation. The weighted average of each of the continuous evaluation notes of each of the blocks of training actions becomes the final grade of the subject.

The subject will not be passed if each of the training blocks that make up the subject are not passed. In this case, there will be no average score between blocks to compensate.

7.1. Ordinary call

To pass the subject in ordinary call you must obtain a grade greater than or equal to 5.0 out of 10.0 in the final grade (weighted average) of the subject.

In any case, it will be necessary for you to obtain a score greater than or equal to 5.0 in the final test, so that it can average with the rest of the activities.

7.2. Extraordinary call

To pass the subject in extraordinary call you must obtain a grade greater than or equal to 5.0 out of 10.0 in the final grade (weighted average) of the subject.

In any case, it will be necessary for you to obtain a score greater than or equal to 5.0 in the final test, so that it can average with the rest of the activities.

The activities not passed in ordinary call must be delivered after having received the corrections corresponding to them by the teacher, or those that were not delivered (they will have a penalty of 0.7 on the note taken). The teacher can also establish some complementary activities that would have to be delivered on the indicated date. The teacher will personally spend enough time with those students who have failed in ordinary to establish what activities, jobs, etc. they have to do for extraordinary.

8. BIBLIOGRAPHY

The recommended literature is indicated below.
The teacher will expand this bibliography in each of the modules.

- Python for Data Analysis. Data Wrangling with Pandas, NumPy, and IPython (2017, O'Reilly)
- Bruce, P., Bruce, A., & Gedeck, P. (2020). Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python. (2020, O'Reilly)