UAB
Universitat Autònoma de Barcelona

2022/2023

Big Data Analysis and Visualisation

Code: 104750
ECTS Credits: 6

Degree	Type	Year	Semester
2503873 Interactive Communication	OT	4	2

Contact

Name: Adrian Padilla Molina

Email: adrian.padilla@uab.cat

Use of Languages

Principal working language: catalan (cat)

Some groups entirely in English: No

Some groups entirely in Catalan: Yes

Some groups entirely in Spanish: No

Prerequisites

- Basic knowledge of Microsoft Excel or OpenOffice Calc.
- Python knowledge acquired in previous subjects.
- A clear desire to learn data-oriented Python.

Objectives and Contextualisation

The main objective of the course is to provide students with the basic skills and competencies to be able to manage, analyze and visualize large volumes of structured information. For this, we will make an introduction to the Python programming language, oriented exclusively to working with data. We will emphasize on the methods of data mining, cleaning and transformation of information for its subsequent analysis. For this we will focus on libraries such as Pandas and Numpy. For the analysis and visualization section we will dive into Tableau Desktop.

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

- Theoretical Framework.
- Basic infrastructure for handling big data.
- Data typologies.
- Data sources (APIs, webscrapping, data mining).
- Data Storage.
- Preparation of the work environment: programming IDE (Pycharm)
- Introduction to data-oriented Python
- Cleaning and preparing data for analysis
- Introduction to Tableau
- Data visualization and analysis
- Elaboration of a project based on data

The calendar will be available on the first day of class. Students will find all information on the Virtual Campus: the description of the activities, teaching materials, and any necessary information for the proper follow-up of the subject.

Methodology

theoretical and practical sessions

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			
Laboratory sessions	33	1.32	20, 5, 6, 11, 9, 12, 15, 19, 21
Theory and guided hands-on learning	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
Type: Supervised			
Group Final Project	50	2	3, 1, 2, 4, 7, 6, 12, 13, 15, 22, 16, 17, 18, 19, 21, 8
Type: Autonomous			
Hours of study	27	1.08	15, 21
Laboratory practices preparation	25	1	1, 20, 6, 12, 13, 19, 21

Assessment

- Group work is mandatory for all students. Without this element it is not possible to pass the subject.
- Those students who do NOT pass the evaluation have the option of taking the final exam, which will take place on the last day of class. This exam will be theoretical and practical. For the final exam it is mandatory to have presented the group work.

Inthe 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.

Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Attendance and Participation	10%	0	0	20, 15, 21
Classrom exercises	30%	0	0	4, 5, 7, 6, 11, 9, 13, 15
Final group Project	50%	0	0	3, 1, 20, 2, 4, 5, 7, 6, 10, 11, 9, 12, 14, 13, 15, 23, 22, 16, 17, 18, 19, 21, 8
Oral presentation in the classroom	10%	0	0	1, 2, 7, 14, 22

Bibliography

The bibliography will be made available to students on the first day of class.

Software

- Tableau
- Gephi
- Pycharm
- Microsoft Exccel / OpenOffice Calc