UAB
Universitat Autònoma de Barcelona

Data Science

Code: 106946
ECTS Credits: 6

Degree	Type	Year
Management of Smart and Sustainable Cities	OB	3

Contact

Name: Xavier Miquel Armengol Fontova

Email: xaviermiquel.armengol@uab.cat

Teaching groups languages

You can view this information at the end of this document.

Prerequisites

To have completed the first-year subjects of Computer Science (Informàtica), Mathematics, Internet applications programming, and the second year subject of Databases.

Objectives and Contextualisation

This subject must allow the student to discover the existing technologies and the different ways for managing and analysing the data generated in the city on a daily basis.

Students will learn techniques for visualization, analysis and modelling of data that will allow them to generate new knowledge and intuitions from the city data.

Learning Outcomes

1. CM19 (Competence) Propose data processing solutions that take into account data privacy and security, as well as that their use respects the ethical values of an egalitarian and democratic society.
2. KM25 (Knowledge) Recognise the problems of information transmission and storage in the context of smart and sustainable cities.
3. KM26 (Knowledge) Identify and use different sources, models and databases of information generated by urban activity, as well as their operating principles, access policies and standards.
4. SM23 (Skill) Design and develop IT solutions that allow citizens distributed access to management platforms and integrated services.

Content

- Data preparation
  - Data visualization
  - Normalization
  - Unknown values

- Reduction of dimensionality
  - Feature selection
- Classification and regression (supervised techniques)
  - Linear and polynomial regression
  - Logistic regression
  - Probabilities, Naive Bayes Classifier
  - Decision trees and "random forests"
  - Hierarchical classification
- Generation of knowledge (unsupervised techniques)
  - Rules of association
  - Recommendation systems

Activities and Methodology

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Exercise sessions	12	0.48	
Theory classes	26	1.04	
Type: Supervised			
Project sessions	12	0.48	
Tutoring	5	0.2	
Type: Autonomous			
Dedication to resolve exercises	12	0.48	
Further reading and study of the material	40	1.6	
Work on practicals (projects)	37	1.48	

Data science is defined by the types of problems that it aims to solve; therefore, it will be that typology of problems that will direct the organization of all the contents.

There will be three types of activities: theory classes, solving practical exercises individually (problems) and developing projects in small teams.

1. Theory classes: The objective of these sessions is for the teacher to explain the theoretical background of the subject. For each one of the topics studied, the theory and mathematical formulation is explained, as well as the corresponding algorithmic solutions.

2. Laboratory sessions: Laboratory sessions aim to facilitate interaction and to reinforce the comprehension of the topics seen in the theory classes. During laboratory sessions we will tackle two types of activities: solving practical exercises and performing team-project follow ups and presentations.

2.1 Problems: A weekly *set of problems* to work through will be used, that require the implementation of methods seen in the theory classes. Work on the problems will be initiated in class and should be completed by each student individually at home. Students will be required to make a weekly submission of their work, that will comprise the problems portfolio.

2.2 Projects: Project sessions comprise activities related to the realization of two short projects during the semester. Students will work collaboratively on these projects in small teams. During the project sessions (1) the teacher will present and discuss the projects and possible approaches, and (2) the teams will present their final results to the class. The teams will have to design and implement a solution, manage the distribution and organization of the work to be carried out, and present final results to the teacher.

The above activities will be complemented by a system of tutoring and consultations outside class hours.

All the information of the subject and the related documents that the students need will be available at the virtual campus.

The transversal competence T01 is addressed through teamwork and collaboration during the development of the projects. The evaluation of the projects includes an oral presentation of each team, during which the students will have to present their work and explain the organization of the team.

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
Exams	40	5	0.2	CM19, KM25, KM26
Exercises deliverables	10	0	0	CM19, KM25, KM26, SM23
Project deliverables	30	0	0	CM19, KM25, KM26, SM23
Project presentations	15	1	0.04	CM19, KM25, KM26, SM23
Self-evaluation	5	0	0	

To assess the level of student learning, a formula is established that combines knowledge acquisition, the ability to solve problems and the ability to work as a team, as well as the presentation of the results obtained. This course does not include a single assessment system.

Final grade

The final grade is calculated weighted in the following way and according to the different activities that are carried out:

Final grade = 0.4 * Theory Grade + 0.1 * Exercises Grade + 0.5 * Projects Grade

This formula will be applied as long as the theory and the laboratory grades, are higher than 5. There is no restriction on the exercises grade. If doing the calculation of the formula yields >= 5 but does not reach the minimum required in any of the evaluation activities, then a final grade of 4.5 will be given.

Theory Grade

The theory grade aims to assess the individual abilities of the student in terms of the theoretical content of the subject, this is done continuously during the course through two partial exams:

Theory Grade = 0.5 * Grade Exam 1 + 0.5 * Grade Exam 2

Exam 1 is done in the middle of the semester and serves to eliminate part of the subject if it is passed.

Exam 2 is done at the end of the semester and serves to eliminate part of the subject if it is passed.

These exams aim to assess the abilities of each student in an individualized manner, both in terms of solving exercises using the techniques explained in class, as well as evaluating the level of conceptualization that the student has made of the techniques seen. In order to obtain a final pass theory grade, it will be required for the partial exam grades 1 and 2 to be both higher than 4.

Recovery exam. In case the theory grade does not reach the adequate level to pass, the students can take a recovery exam, destined to recover the failed part (1, 2 or both) of the continuous evaluation process.

Exercises Grade

The aim of the exercises is for the student to train with the contents of the subject continuously and become familiar with the application of the theoretical concepts. As evidence of this work, the presentation of a portfolio is requested in which the exercises worked out will be collated.

In order to obtain a grade for exercises, it is necessary that more than 50% of the exercises are submitted during the semester. In the contrary, the exercises grade will be 0.

Exercises Grade = Portfolio evaluation

Projects Grade

The part of projects carries an essential weight in the overall mark of the subject. Developing the projects requires that the students work in groups and design an integral solution to the defined challenge. In addition, the students must demonstrate their teamwork skills and present the results to the class.

Each of the two projects is evaluated through its deliverable, an oral presentation that students will make in class, and a self-evaluation process. The participation of students in all three activities (preparing the deliverable, presentation and auto evaluation) is necessary in order to obtain a projects grade. The grade is calculated as follows:

Project Grade X = 0.6 * Grade Deliverables + 0.3 * Grade Presentation + 0.1 * Grade Self-evaluation

If performing the above calculation yields >= 5 but the student did not participate in any of the activities (deliverable, presentation, auto evaluation), then a final grade of 4.5 will be given to the corresponding project.

Laboratory Grade = 0.5 * Grade Project 1 + 0.5 * Grade Project 2

To obtain a project grade, it will be necessary that the grades of both projects are above 4.

In case of not passing any of the projects, it will be allowed to recover it restricted to a maximum grade of 7/10.

Important notes

In this course, the use of Artificial Intelligence (AI) technologies is permitted-except in activities where otherwise indicated-as an integral part of thedevelopment of the work. In all cases, the final result must always reflect a significant contribution from the student in terms of analysis and personal reflection. The student must clearly identify which parts were generated using this technology, specify the tools used, and include a critical reflection on how these influenced the process and the final outcome of the activity. Lack of transparency in the

use of AI will be considered academic dishonesty and may result in a penalty in the activity's grade, or more severe sanctions in serious cases.

Notwithstanding other disciplinary measures deemed appropriate, and following the academic regulations in force, evaluation activities will be suspended with zero (0) whenever a student commits any academic irregularities that may alter such evaluation (for example, plagiarizing, copying, letting copy, ...). The evaluation activities qualified in this way and by this procedure will not be recoverable. If you need to pass any of these assessment activities to pass the subject, the subject will be failed directly, without the opportunity to recover it in the same course.

In case the student does not deliver any exercise solutions, does not attend any project presentation session during the laboratory sessions, and does not take any exam, the corresponding grade will be a "non-evaluable". In another case, the "no shows" count as a 0 for the calculation of the weighted average.

Starting from the second enrollment, the grade for problems and/or projects may be validated, provided it was passed with a grade equal to or higher than 6.

In order to pass the course with honours, the final grade obtained must be equal to or higher than 9 points. Because the number of students with this distinction cannot exceed 5% of the total number of students enrolled in the course, it is given to whoever has the highest final marks. In case of a tie, the results of the partial exams will be taken into account.

Bibliography

- Data Science from Scratch: First Principles with Python, Joel Grus, O'Reilly Media, 2015, 1st Ed.
- Python Data Science Handbook, Jake VanderPlas, O'Reilly Media, 2016, 1st Ed.
- Pattern Recognition and Machine Learning, Christopher Bishop, Springer, 2011
- Model-Based Machine Learning, J. Winn, C. Bishop, early access: http://mbmlbook.com/
- Computational and Inferential Thinking: The Foundations of Data Science, Ani Adhikari and John DeNero, online: https://ds8.gitbooks.io/textbook/content/

Software

For the problems and projects of the course we will use Python, and the Python: libraries NumPy, MatPlotLib, SciKit Learn, Pandas

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
(PAUL) Classroom practices	1	Catalan/Spanish	first semester	afternoon
(PAUL) Classroom practices	2	Catalan/Spanish	first semester	afternoon
(PLAB) Practical laboratories	1	Catalan/Spanish	first semester	afternoon

(PLAB) Practical laboratories	2	Catalan/Spanish	first semester	afternoon
(PLAB) Practical laboratories	3	Catalan/Spanish	first semester	afternoon
(TE) Theory	1	Catalan	first semester	afternoon