VNIVERSITAT DE VALÈNCIA

Course Guide
36494 Big Data Programming and Management

COURSE DATA

DATA SUBJECT

Code: 36494
Name: Big Data Programming and Management
Cycle: Undergraduate Studies
ECTS Credits: 6
Academic year: 2025-26

STUDY (S)

Degree	Center	Acad. year	Period
1316 - Degree in Economics	Facultat d'Economia	4	First quarter

SUBJECT-MATTER

Degree	Subject-matter	Character
1316 - Degree in Economics	Economía Internacional y Territorio	ELECTIVES

COORDINATION

PEREZ VAZQUEZ PEDRO JOSE

SUMMARY

The subject is taught in the first semester of the fourth year of the Degree in Economics. It is an optional subject with a total teaching load of 6 credits (150 hours). It is included in the subject Data analysis and in the mention in International, Regional and Urban Economics, although due to its content it may be of interest to all Grade students.

The course aims to initiate students in the field of data science by presenting them with an overview of topics, and to expose them, perhaps for the first time in the career, to programming and to a branch of statistics that is emerging with great force such as Machine Learning.

The course will begin with a topic in which the objectives of the course would be presented, together with the terms and processes that justify the subject (Big data, Data Science, Machine Learning, ...). Examples of successful application of these new techniques in the field of Economics would be presented, to finally highlight the importance of reproducible research and the role that free software plays in it.

Theme 2 presents the rudiments of the programming language R. This basic knowledge will be reinforced during the course by working on different topics and applications. In theme 3, several cases of tabular data analysis will be developed, the most typical format in Social Sciences, with which the programming knowledge initiated in theme 2 will be consolidated, now applied to concrete tasks related to the handling

of tabular data and diverse topics of analysis.

Theme 4 is an introduction to the basic concepts of Machine Learning (training set, validation set, etc.), to then present the basic ML techniques or algorithms and more related to what students have already seen in Econometrics, now presented from a different perspective, more oriented towards prediction.

The first four topics constitute the fundamental part of the subject that ends with an additional topic in which, once the students have reached a certain fluency in the handling of R, different topics are presented in the field of Data Science, for example, spatial data and its representation, textual data (for example, Twitter) and linked data (Wikidata).

PREVIOUS KNOWLEDGE

RELATIONSHIP TO OTHER SUBJECTS OF THE SAME DEGREE

There are no specified enrollment restrictions with other subjects of the curriculum.

OTHER REQUIREMENTS

No prerequisites. However, it is highly recommended that students have a basic knowledge of computer science, as well as a minimum knowledge of Statistics and Econometrics.

COMPETENCES / LEARNING OUTCOMES

1316 - Degree in Economics

Apply the principles of economic analysis (rational decision) to the diagnosis and resolution of problems.

Be able to collect and analyse information.

Be able to learn autonomously.

Be able to use ICTs.

Have decision-making skills and be able to apply knowledge to practice.

Understand and apply the scientific method, which involves formulating hypotheses, deducing verifiable results and contrasting them with empirical and experimental evidence.

DESCRIPTION OF CONTENTS

1. Data Science and the Big Data Era

2. An Introduction to R for Data Science

3. Data Wrangling of tabular data

4. Introduction to Machine Learning

5. Other types of data

WORKLOAD

PRESENCIAL ACTIVITIES

Activity	Hours
Theory	30,00
Classroom practices	30,00
Total hours	**60,00**

NON PRESENCIAL ACTIVITIES

Activity	Hours
Attendance at other activities	0,00
Individual or group project	40,00
Independent study and work	20,00
Preparation of lessons	20,00
Preparation for assessment activities	10,00
Resolution of case studies	0,00
Total hours	**90,00**

TEACHING METHODOLOGY

The teaching methodology consists of:

- Participatory masterclass to present the basic theoretical contents

- Practical classes with problem solving, case studies, application of techniques, oral presentations, debates...

- Autonomous supervised work based on reading and evaluating reports, carrying out exercises and/or projects individually and/or in teams.

- Independent study of the students

- Written and/or oral tests.

EVALUATION

The evaluation of the students' learning will be carried out through:

1. A final exam or test, not necessarily written, consisting of both theoretical and practical questions.

2. Continuous evaluation of the student based both on their participation and involvement in the teaching-learning process and on the evaluation of the practical activities developed by the student during the course. These practical activities may consist of the preparation of papers, memoirs, oral presentations, ...

Continuous assessment will account for 40% of the total grade. By their very nature, continuous evaluation activities are NOT recoverable and their grade will be maintained for the second call. In order to pass the subject, the minimum grade in the global computation must be 5 points out of 10. In case of not carrying out the continuous evaluation tasks, the student will only be able to obtain the points of the final test (6 maximum), and would need to obtain a 5 out of 6 in this exam to pass the subject.

The specific criteria and processes that will be used for the evaluation, as well as their specific numerical weighting, will depend on the number of students finally enrolled and will be adequately publicised at the beginning of the course.

REFERENCES

- Grolemund, G., Wickham, H. (2019): R para Ciencia de Datos. https://es.r4ds.hadley.nz/
- Grolemund, G., Wickham, H. (2016): R for Data Science, OReilly. https://r4ds.had.co.nz/
- Peng, R. D. (2016): Exploratory Data Analysis with R. Lean Publishing. https://bookdown.org/rdpeng/exdata/
- James, G., Witten, E., Hastie T. y Tibshirani, R. (2015): An Introduction to Statistical Learning with

applications in R. http://www-bcf.usc.edu/~gareth/ISL/

- Ayala, G. (2015): Estadística básica. https://www.uv.es/~ayala/docencia/nmr/nmr13.pdf
- Casas, P. (2019): Data Science Live Book. https://livebook.datascienceheroes.com/
- Coll, V. y Pérez P. J. (2017): Curso de introducción a R. https://www.uv.es/pjperez/curso_R/index.html
- Gil Bellosta, C. G. (2018): R para profesionales de los datos: una introducción. https://www.datanalytics.com/libro_r/
- Santana, A. y Hernández, C. M. (2016): R4ULPGC: Introducción a R. http://www.dma.ulpgc.es/profesores/personal/stat/cursoR4ULPGC/index.html