

**COURSE DATA****Data Subject**

<b>Code</b>	44952
<b>Name</b>	Big data in Economics
<b>Cycle</b>	Master's degree
<b>ECTS Credits</b>	5.0
<b>Academic year</b>	2023 - 2024

**Study (s)**

<b>Degree</b>	<b>Center</b>	<b>Acad. Period</b>
2242 - Master's Degree in Economics	Faculty of Economics	1 First term

**Subject-matter**

<b>Degree</b>	<b>Subject-matter</b>	<b>Character</b>
2242 - Master's Degree in Economics	10 - Materia instrumental	Optional

**Coordination**

<b>Name</b>	<b>Department</b>
PAVIA MIRALLES, JOSE MANUEL	110 - Applied Economics
PEIRO PALOMINO, JESUS	132 - Economic Structure

**SUMMARY**

This course introduces the student to the management and understanding of the potential of Big Data information for the analysis of issues of economic interest. The aim is, therefore, to familiarize the student with data sets or combinations of data sets whose size (volume), complexity (variability) and speed of growth (speed) hinder their capture, management, processing or analysis using conventional tools. This course will show to students the great possibilities that the use of this type of data provides for the study of topics in economics.

**PREVIOUS KNOWLEDGE**



### Relationship to other subjects of the same degree

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

### Other requirements

To successfully complete this course , it is assumed that the student has a sufficient level of mathematics, statistics and econometrics.

## COMPETENCES (RD 1393/2007) // LEARNING OUTCOMES (RD 822/2021)

### 2242 - Master's Degree in Economics

- Students should apply acquired knowledge to solve problems in unfamiliar contexts within their field of study, including multidisciplinary scenarios.
- Students should be able to integrate knowledge and address the complexity of making informed judgments based on incomplete or limited information, including reflections on the social and ethical responsibilities associated with the application of their knowledge and judgments.
- Students should communicate conclusions and underlying knowledge clearly and unambiguously to both specialized and non-specialized audiences.
- Students should demonstrate self-directed learning skills for continued academic growth.
- Students should possess and understand foundational knowledge that enables original thinking and research in the field.
- Develop time management skills for learning: skills for organisation, planning and decision making in the process of learning advanced economics.
- Desarrollar la capacidad crítica, impulsar la inquietud y el interés investigador en el ámbito de la economía, especializarse en el manejo de material bibliográfico, en la utilización de bases de datos económicas y programas matemáticos y estadísticoeconómicos, así como aprender a transmitir de forma adecuada los resultados de investigadora a través de artículos científicos y ponencias en congresos.
- Gain the capacities of abstraction and logical reasoning that are essential for the creation of economic models: ability to express oneself using formal, graphic and symbolic languages, to apply analytical and mathematical methods to economics, and to relate and manipulate concepts according to a purpose.
- Acquire linguistic and technological skills: ability to use English in the scientific field of economics and to use ICT in the field of economic study and research.
- Acquire social skills to work in a team: capacity to coordinate activities, ethical and responsible commitment, leadership and mobilisation capacity, all of which are important for economic and team management.
- Know how to promote, in academic and professional contexts, technological, social or cultural progress in a knowledge-based society that is founded on the respect for: (a) fundamental rights and the principles of equal opportunities for men and women, which involves using an inclusive and egalitarian language that promotes the visibility of women; (b) the principles of equal opportunities and universal accessibility for people with disabilities, and (c) the distinctive values of a culture of



peace and democratic values.

- Communicate orally and in writing using an inclusive and egalitarian language.
- Know the databases and bibliography necessary to carry out economic research work.
- Use the knowledge gained to identify career prospects and sources of employment, and acquire the personal skills that facilitate professional insertion and development. To that end, students should know and know how to use job search techniques and tools and consider entrepreneurship as a professional alternative.
- Know how to manage and process databases using the most appropriate and current techniques and software packages.

## LEARNING OUTCOMES (RD 1393/2007) // NO CONTENT (RD 822/2021)

The expected results of learning this subject are as follows:

- Understand the potential of Big Data information for the analysis of issues of economic interest.
- Know the main methods of extraction and treatment of Big-data.
- Communication: summary of results through visualization, stories and interpretable summaries.
- Knowledge of basic statistical or machine learning techniques.
- Application of statistical or automatic learning to economic problems.
- Interpret both databases and economic models. Be familiar with microeconomic techniques in applied economics.
- Use the R statistical package to solve practical cases that use the different algorithms explained in the theoretical part of the program.

## DESCRIPTION OF CONTENTS

### 1. 1

#### 1. THEORY UNIT 1: INTRODUCTION TO STATISTICAL (MACHINE) LEARNING

1. A brief tour for the machine (statistical) learning world.
2. Supervised, unsupervised and semi-supervised learning. Active learning.
3. Regression and classification.
4. Bootstrap.
5. RMarkdown.



## 2. 2

### 2. THEORY UNIT 2: LINEAR MODELS FOR REGRESSION. REGULARIZATION

1. Linear models and regularization.
2. Ridge, lasso and net elastic regression.
3. Nearest neighbours.
4. Model assessment and selection.
5. High dimensional problems,  $p \gg N$ .

## 3. 3

### 3. THEORY UNIT 3: CLASSIFICATION

1. Regression models for classification.
2. Decision trees. Bagging and Random Forest.
3. (Gradient) Boosting.
4. Support Vector Machines.
5. Neural Networks and Deep Learning.

## 4. 4

### 4. THEORY UNIT 4: BIG DATA SOURCES

1. How big is big?
2. Open data.
3. Goggle trends.
4. Administrative data.
5. Real time data.

## 5. 5

### 5. APPLICATIONS

1. Applications of big data in several economic fields: labor market, entrepreneurship, economic growth, trade, urban economics, economic policy, energy economics, among others.



## WORKLOAD

ACTIVITY	Hours	% To be attended
Theory classes	40,00	100
Classroom practices	10,00	100
Study and independent work	75,00	0
<b>TOTAL</b>	<b>125,00</b>	

## TEACHING METHODOLOGY

The development of the subject is basically structured around sessions where the theoretical contents will be continuously exemplified with practical cases and practical sessions where the student will practice and test the practical cases of the classes with examples. Depending on the type of session (theoretical or practical), one teaching method or another will be chosen.

In the theoretical part of the sessions, the main contents of the topics that make up the subject will be exposed, introducing the relevant elements and concepts and contextualizing them to the different fields of application of data analysis, the economic environment and its application to the world of the economy.

In the practical sessions, the teacher will propose to the students situations (real or fictitious) or practical cases that they must solve with the application of techniques and the use of appropriate computer programs, making oral presentations or debates, where appropriate, individually and/or or in a team. Likewise, projects and situations will be proposed that the students must solve.

In the applications part of the course, students will study academic papers using big data techniques. Students will prepare presentations of these works.

## EVALUATION

The course has two parts: theory (40%) and applications (60%). It is compulsory to obtain at least 4 points (out of 10) in each part to compute the average grade and pass the course.

In both parts there will be continuous assessment activities and a final exam.

## REFERENCES





### Basic

- Efron, B. and Hastie, T. (2016). Computer Age Statistical Inference: Algorithms, Evidence and Data Science. Cambridge University Press.

Hastie, T., Tibshinari, R. and Friedman J. (2009). The Elements of Statistical Learning. Data mining, Inference, and Prediction. Springer.

James, G., Witten, D., Hastie, T. and Tibshirani, R. (2013). An Introduction to Statistical Learning, with Application in R. Springer.

### Additional

- Knaflic, C.L. (2015). Storytelling with Data: A data visualization guide for business professionals, Wiley.

Wickham, H. y Grolemund, G. (2017). R for Data Science. O'Reilly Media, Inc.

Xie, Y., Allaire, J.J. y Grolemund, G. (2018). R Markdown: The definitive guide, Chapman & Hall/CRC.