COURSE DATA

Data Subject

Code	44649
Name	Introduction to data science
Cycle	Master's degree
ECTS Credits	6.0
Academic year	2020 - 2021

Study (s)

Degree	Center	Acad. year	Period
2221 - M.U. en Ciencia de Datos	School of Engineering	1	First term

Subject-matter

Degree	Subject-matter	Character
2221 - M.U. en Ciencia de Datos	1 - Introduction to data science	Obligatory

Coordination

Name	Department
SERRANO LOPEZ, ANTONIO JOSE	242 - Electronic Engineering
SORIA OLIVAS, EMILIO	242 - Electronic Engineering

SUMMARY

This subject aims at introducing all the concepts related to Data Science as well as providing the introduction to the tools to be used in the rest of the master (R and Python). The mathematical concepts needed will be reviewed to be able to follow the rest of the modules of the master.

PREVIOUS KNOWLEDGE

Relationship to other subjects of the same degree

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

Other requirements

OUTCOMES

2221 - M.U. en Ciencia de Datos

- Students should demonstrate self-directed learning skills for continued academic growth.

- Be able to assess the need to complete their technical, scientific, language, computer, literary, ethical, social and human education, and to organise their own learning with a high degree of autonomy.

- Ability to access and manage information in different formats for subsequent analysis in order to obtain knowledge from data.

- Ser capaces de acceder a herramientas de información (bibliográficas y de empleo) y utilizarlas apropiadamente.

- Extraer conocimiento de conjuntos de datos en diferentes formatos.

- Entender la utilidad de la ciencia de datos y sus elementos asociados, así como su aplicación en la resolución de problemas, eligiendo las técnicas más adecuadas a cada problema, aplicando de forma correcta las técnicas de evaluación y, finalmente,
interpretando los modelos y resultados.

LEARNING OUTCOMES

Know the scope of data science and the different techniques and approaches that exist for it.
Know the different types of data that one can find. Knowing the different types of learning that exist.
Know the libraries of the most common programming languages for data processing by solving case studies.
Set out correctly the different stages involving the resolution of all problems of data analysis including experimental design and validation of the models.

DESCRIPTION OF CONTENTS

1. Data science definition, objectives and techniques. Programs/Languages used in data science.

2. Problem types in data science. Types of data: structured and unstructured. Types of learning: examples

3. Stages of a data analysis problem. Procedures for the evaluation of data-based models.

4. Measurements of error. Frequent problems: overfitting and imbalanced classes

5. Review of basic concepts needed for Data Science; Algebra, probability and theory of information

6. Practical introduction with examples of the use of R and Python languages, oriented to data science

WORKLOAD

ACTIVITY	Hours	% To be attended
Theoretical and practical classes	60,00	100
Development of individual work	5,00	0
Study and independent work	12,00	0
Readings supplementary material	8,00	0
Preparation of evaluation activities	20,00	0
Preparing lectures	10,00	0
Preparation of practical classes and problem	25,00	0
Resolution of case studies	10,00	0
TOTAL	**150,00**	

TEACHING METHODOLOGY

Theoretical activities. Interactive lectures about the subject with the participation of the student in the resolution of specific issues. Conducting individual evaluation questionnaires.

Practical activities. Learning through problem solving exercises and case studies through which skills about different aspects of the subject are acquired.

Work in laboratory and / or in computer room. Learning by performing activities individually or in small groups and conducted in computer rooms.

EVALUATION

- Objective test to evaluate the theory, consisting of one or more tests that contain theoretical questions (20%)
- Objective test to evaluate the laboratory work consisting of one or more tests that include practical problems (80%)

Assessment of practical activities from the preparation of works / reports, oral presentations and e-learning tools of the University (they will complement the grades above)

REFERENCES

Basic

- O Neill, D., (2013). Doing Data Science: Straight Talk from the Frontline, ed OReilly
- Foreman, J.W. (2013). Data Smart: Using Data Science to Transform Information into Insight, Ed. Wiley.
- John Braun Duncan, W., Murdoch, J. (2007). A first course in statistical programming with R. Cambridge University Press.
- Venables, V.N. (2013). An Introduction to R. http://cran.r-project.org.
- Massaron, L. (2015). Python Data Science Essentials, Packt Publishing.
- Toomey, D (2014). R for data Science, Packt Publishing.
- Ojeda, T. (2014). Practical Data Science Cookbook, Packt Publishing.
- Grolemund, G., Wickham, H. R for Data Science, OReilly, 2016

Additional

- Chen, L.M. (2015). Mathematical Problems in Data Science Theoretical and Practical Methods, Springer
- Pierson, L. (2015). Data Science for Dummies., Wiley

ADDENDUM COVID-19

This addendum will only be activated if the health situation requires so and with the prior agreement of the Governing Council

In the event of a hybrid mode of teaching (combining attendance with non-attendance) or a closure of the facilities due to health reasons that affect all or part of the classes of the course, these will be replaced preferably by synchronous non-attendance sessions following the established schedules.

VNIVERSITAT Ö ID VALÈNCIA

If the closure affects a course evaluation test, it will be replaced by a test of a similar nature that will be carried out in virtual mode through the computer tools supported by the University of Valencia.

The percentages of each assessment test will remain unchanged, as established by this guide.