

**FICHA IDENTIFICATIVA****Datos de la Asignatura**

Código	36494
Nombre	Programación y Manejo de Datos en la era del Big Data
Ciclo	Grado
Créditos ECTS	6.0
Curso académico	2020 - 2021

Titulación(es)

Titulación	Centro	Curso	Periodo
1316 - Grado de Economía	Facultad de Economía	4	Primer cuatrimestre

Materias

Titulación	Materia	Carácter
1316 - Grado de Economía	26 - Economía Internacional y Territorio	Optativa

RESUMEN

La asignatura se imparte en el primer semestre del cuarto curso del Grado en Economía. Es una asignatura optativa con una carga lectiva total de 6 créditos (150 horas). Está incluida en la materia Análisis de datos y en la mención en Economía Internacional, Regional y Urbana, aunque por su contenido puede resultar de interés para todos los estudiantes del Grado.

La asignatura pretende iniciar a los estudiantes en el campo de la ciencia de datos presentándoles una panorámica de tópicos, y exponerles, quizás por primera vez en la carrera, a la programación y a una rama de la estadística que está emergiendo con gran fuerza como es el Machine Learning.

El curso comenzará con un tema en el que se presentarían los objetivos del curso, junto con los términos y procesos que justifican la asignatura (Big data, Ciencia de Datos, Machine Learning, ...). Se presentarían ejemplos de aplicación exitosa de estas nuevas técnicas en el campo de la Economía, para finalmente resaltar la importancia de la investigación reproducible y el papel que juega el software libre en ella.

El tema 2 se presentan los rudimentos del lenguaje de programación R. Estos conocimientos básicos se irán reforzando durante el curso trabajando en diferentes tópicos y aplicaciones. En el tema 3, se desarrollarán diversos casos de análisis de datos tabulares, el formato más típico en Ciencias Sociales, con los que se irán asentando los conocimientos de programación iniciados en el tema 2, ahora aplicados a tareas concretas relacionados con el manejo de datos tabulares y diversos tópicos de análisis.

El tema 4 es una introducción a los conceptos básicos del Machine Learning (training set, validation set, etc.), para luego presentar las técnicas o algoritmos de ML básicos y más relacionados con lo que los estudiantes ya han visto en Econometría, ahora presentados desde una óptica diferente, más orientada a la predicción



Los cuatro primeros temas constituyen la parte fundamental de la asignatura que finaliza con un tema adicional en el que, una vez los estudiantes han alcanzado cierta soltura en el manejo de R, se presentan diferentes tópicos en el campo de la Ciencia de Datos, por ejemplo, datos espaciales y su representación, datos textuales (por ejemplo, Twitter) y datos enlazados (Wikidata).

CONOCIMIENTOS PREVIOS

Relación con otras asignaturas de la misma titulación

No se han especificado restricciones de matrícula con otras asignaturas del plan de estudios.

Otros tipos de requisitos

Sin requisitos previos. No obstante, es muy recomendable que los estudiantes posean conocimientos básicos de informática, así como que posean unos conocimientos mínimos de Estadística y Econometría.

COMPETENCIAS

1316 - Grado de Economía

- Capacidad para la búsqueda y análisis de información.
- Capacidad para la toma de decisiones aplicando los conocimientos a la práctica.
- Capacidad de aprendizaje autónomo.
- Capacidad de manejar las tecnologías de la información.
- Aplicar los principios del análisis económico (decisión racional) al diagnóstico y resolución de problemas.
- Comprender y aplicar el método científico, consistente en formular hipótesis, deducir resultados comprobables y contrastarlos con la evidencia empírica y experimental.

RESULTADOS DE APRENDIZAJE

- Búsqueda de información y conocimiento de fuentes estadísticas. Manejo de datos estadísticos y de bases de datos para el estudio de la economía.
- Abstracción y procesamiento de la información para su posterior utilización en paquetes informáticos.
- Identificación, clasificación, razonamiento, argumentación e interpretación de las relaciones entre variables económicas.
- Reconocimiento de un problema económico a partir de la observación de la realidad.
- Extracción de información útil a partir de datos y modelos cuantitativos.
- Sintetizar, abstraer y aportar conclusiones o implicaciones económicas a partir del análisis realizado.
- Programación para el análisis de datos.



- Importación, limpieza, visualización y análisis exploratorio de datos tabulares

DESCRIPCIÓN DE CONTENIDOS

1. Ciencia de datos y la era del Big Data

2. Una introducción a R para Ciencia de datos

3. Manejo y análisis exploratorio de datos tabulares

4. Introducción al Machine Learning

5. Otros tipos de datos

VOLUMEN DE TRABAJO

ACTIVIDAD	Horas	% Presencial
Clases de teoría	30.00	100
Prácticas en aula	30.00	100
Elaboración de trabajos en grupo	20.00	0
Elaboración de trabajos individuales	20.00	0
Estudio y trabajo autónomo	20.00	0
Preparación de actividades de evaluación	10.00	0
Preparación de clases de teoría	10.00	0
Preparación de clases prácticas y de problemas	10.00	0
TOTAL	150.00	

METODOLOGÍA DOCENTE

La metodología docente consiste en:

- Lección magistral participativa para presentar los contenidos teóricos básicos
- Clases prácticas con resolución de problemas, estudios de casos, aplicación de técnicas, presentaciones orales, debates...



- Trabajo autónomo supervisado basado en lectura y valoración de informes, realización de ejercicios y/o proyectos individualmente y/o en equipo.
- Estudio independiente de los estudiantes
- Realización de pruebas escritas y/u orales.

EVALUACIÓN

La evaluación del aprendizaje de los alumnos se realizará mediante:

1. Un examen o prueba final, no necesariamente escrito, que constará tanto de cuestiones tanto teóricas como prácticas.
2. Evaluación continua del estudiante basado tanto en su participación e implicación en el proceso de enseñanza-aprendizaje como en la evaluación de las actividades prácticas desarrolladas por el estudiante durante el curso. Estas actividades prácticas pueden consistir en elaboración de trabajos, memorias, exposiciones orales, ...

La evaluación continua supondrá el 40% de la nota total. Por su propia naturaleza, las actividades de evaluación continua son NO recuperables y su nota se mantendrá para la segunda convocatoria. Para superar la asignatura la calificación mínima en el cómputo global ha de ser de 5 puntos sobre 10. En caso de no realizar las tareas de evaluación continua, el estudiante sólo podrá obtener los puntos de la prueba final (6 como máximo), y necesitaría obtener un 5 sobre 6 en dicho examen para aprobar la asignatura.

Los criterios y procesos específicos que se utilizarán para la evaluación, así como su ponderación numérica concreta, estarán en función del número de estudiantes finalmente matriculados y se publicitarán adecuadamente al comenzar el curso

REFERENCIAS

Básicas

- Golemund, G., Wickham, H. (2019): R para Ciencia de Datos. <https://es.r4ds.hadley.nz/>
- Golemund, 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/>

**Complementarias**

- 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>

ADENDA COVID-19

Para el curso 2020-2021 está previsto que la docencia de esta asignatura sea presencial, tanto en teoría como en práctica, por tanto, se seguirá todo lo establecido en esta Guía Docente salvo en los siguientes dos puntos:

- **Evaluación:** siguiendo las recomendaciones actuales por parte de la UV, la evaluación continua aportará un 60% de la nota final, mientras que el examen final aportará el restante 40%.
- **Tutorías:** siguiendo las recomendaciones actuales por parte de la UV, las tutorías se realizarán de forma on-line.

No obstante, si hubiesen cambios en la situación sanitaria se adecuarán los métodos docentes y de evaluación a la nueva situación. Las medidas concretas de adaptación dependerán de la severidad de las nuevas condiciones sanitarias y del porcentaje de clases ya impartidas, informándose de ellas con la máxima rapidez posible a los estudiantes.

La asignatura tiene una orientación eminentemente práctica y computacional, de forma que es necesario que, tanto en el modulo práctico como en el teórico, los estudiantes tengan acceso a ordenadores. Sin embargo, la situación sanitaria hace que actualmente la docencia de la asignatura esté programada en un aula sin ordenadores. Es por ello, que los estudiantes deberán venir a clase con su propio ordenador, o utilizar los sistemas de préstamo que la UV pone a disposición de estos.