

Curso Académico: (2023 / 2024)

Fecha de revisión: 28-04-2023

Departamento asignado a la asignatura: Departamento de Estadística

Coordinador/a: MUÑOZ GARCIA, ALBERTO

Tipo: Obligatoria Créditos ECTS : 6.0

Curso : 4 Cuatrimestre : 1

REQUISITOS (ASIGNATURAS O MATERIAS CUYO CONOCIMIENTO SE PRESUPONE)

Conocimiento de estadística multivariante básica.

OBJETIVOS

1. Conocer y utilizar técnicas estadísticas avanzadas, con soporte informático de última generación.
 2. Extraer y analizar información de grandes conjuntos de datos.
 3. Adquirir las habilidades estadísticas para el análisis de datos multivariantes socio-económicos, tales como los que aparecen en investigación de mercados.
 4. Habilidad de describir y analizar conjuntos de datos reales con las técnicas mencionadas.
 5. Capacidad de elaborar informes de resultados de los análisis sobre estudios de caso reales.
-
1. Capacidad de análisis y síntesis de la información en problemas de Data Mining.
 2. Resolución de problemas reales.
 3. Conocimiento y adiestramiento en el uso de software estadístico para la resolución de problemas reales.
 4. Razonamiento crítico y selectivo frente a la resolución de problemas.
 5. Habilidades de presentación de resultados

DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

1. Lenguaje de Programación R
 - 1.1 Tipos de datos y lectura de datos
 - 1.2 Bucles y condicionales
 - 1.3 Funciones
2. Análisis Exploratorio de Datos
 - 2.2 Paquete ggplot2
3. Técnicas de Clasificación Supervisada
 - 3.1 K-vecinos próximos
 - 3.2 Árboles de decisión
 - 3.3 La distribución Normal y el Análisis Discriminante
 - 3.4 Máquinas de vector soporte
 - 3.5 Regresión Logística
4. Técnicas de Reducción de Dimensiones y Clustering
 - 4.1 Principal Component Analysis
 - 4.2 K-means

4.3 Clustering jerárquico

5. Como escribir un informe con R-Markdown

ACTIVIDADES FORMATIVAS, METODOLOGÍA A UTILIZAR Y RÉGIMEN DE TUTORÍAS

El curso consta de 14 clases teóricas con material de apoyo disponible en aula global y 14 sesiones basadas en sesiones de prácticas computacionales. Cada semana los estudiantes dispondrán de una tutoría optativa colectiva donde podrán resolver sus dudas.

SISTEMA DE EVALUACIÓN

La evaluación continua contará el 100% de la asignatura. Los estudiantes que no superen la asignatura por evaluación continua podrán presentarse a un examen que contará el 100% de la asignatura.

La evaluación continua constará de un análisis exploratorio y un análisis predictivo de una base de datos. A la mitad del curso los estudiantes entregarán un reporte en el que mostrarán los resultados más interesantes que han obtenido en su análisis exploratorio. Este reporte tendrá un peso del 30% de la nota final de la evaluación continua. Antes del final del curso, los estudiantes completarán este trabajo exploratorio con un estudio predictivo en el que predicaran el valor de una variable de interés. Este segundo trabajo será el 30% de la evaluación continua. Los estudiantes deberán presentar y defender este segundo trabajo al final del curso. La defensa de este trabajo será el 40% restante de la evaluación continua.

Peso porcentual del Examen Final:	0
Peso porcentual del resto de la evaluación:	100

BIBLIOGRAFÍA BÁSICA

- Eva Laude, Henri Laude DATA SCIENTIST Y LENGUAJE R, Ediciones ENI, 2022

BIBLIOGRAFÍA COMPLEMENTARIA

- Peter Bruce Estadística práctica para ciencia de datos con R y Python, O'Reilly, 2022