
Curso Académico: (2023 / 2024)

Fecha de revisión: 22/07/2023 13:57:32

Departamento asignado a la asignatura: Departamento de Estadística

Coordinador/a: MARIN DIAZARAQUE, JUAN MIGUEL

Tipo: Obligatoria Créditos ECTS : 3.0

Curso : 1 Cuatrimestre : 1

REQUISITOS (ASIGNATURAS O MATERIAS CUYO CONOCIMIENTO SE PRESUPONE)

Ninguno.

OBJETIVOS

El estudiante adquirirá los siguientes conocimientos:

1. Dominio del lenguaje de programación R y del entorno de trabajo R-studio.
2. Dominio de las distintas estructuras de datos.
3. Análisis exploratorio de datos y presentación de resultados a través de técnicas de visualización de datos.
4. Conocimiento de los principales paquetes de análisis de datos de R.
5. Saber como realizar una simulación adecuadamente.
6. Acelerar los programas realizados mediante programación en paralelo.
7. Encontrar errores y cuellos de botella en el código y generar informes.

DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

1. Fundamentos de Programación I.

El entorno R-studio. Tipos de datos (Matrices, Listas, Factores, Data Frames,...) y sus operaciones. Estructuras de control. Funciones.

2. Fundamentos de Programación II.

Estructuras avanzadas de datos. Lectura y almacenamiento de datos.

3. Visualización de datos.

El paquete ggplot2.

4. Introducción a algunos paquetes útiles en R.

Paquetes MASS, Caret, dplyr y data.table

5. Simulaciones.

6. Programación en paralelo.

7. Debugging, Profiling y presentación de los resultados con Rmarkdown.

ACTIVIDADES FORMATIVAS, METODOLOGÍA A UTILIZAR Y RÉGIMEN DE TUTORÍAS

Las clases consisten en una mezcla de exposiciones sobre los conceptos fundamentales de la asignatura y la presentación de casos prácticos mediante el uso de software R. Se espera que los estudiantes traigan sus propios portátiles para experimentar con el código durante las clases.

* Actividades formativas

- AF1: Clase teórica.
- AF2: Clase práctica.
- AF5: Tutorías.
- AF6: Trabajo en grupo.
- AF7: Trabajo individual.
- AF8: Pruebas de evaluación presenciales.

* Metodologías docentes

- MD1: Exposiciones en clase del profesor con soporte de medios informáticos y audiovisuales, en las que se desarrollan los conceptos principales de la materia y se proporciona la bibliografía para complementar el aprendizaje de los alumnos.
- MD2: Lectura crítica de textos recomendados por el profesor de la asignatura: Artículos de prensa, informes, manuales y/o artículos académicos, bien para su posterior discusión en clase, bien para ampliar y consolidar los conocimientos de la asignatura.
- MD3: Resolución de casos prácticos, problemas, etc. planteados por el profesor de manera individual o en grupo.
- MD4: Exposición y discusión en clase, bajo la moderación del profesor de temas relacionados con el contenido de la materia, así como de casos prácticos.
- MD5: Elaboración de trabajos e informes de manera individual o en grupo.

SISTEMA DE EVALUACIÓN

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

La asignatura será evaluada a través de la entrega de cuatro prácticas cortas que se entregarán a lo largo de las semanas del curso. También tendrán que entregar un trabajo basado en un guion que será accesible en Aula Global, consistente en el análisis de una base de datos real usando las técnicas de programación aprendidas durante el curso.

BIBLIOGRAFÍA BÁSICA

- Felicidad Marques Asension R en profundidad. Programación, gráficos y estadística, RC, 2017
- Fox, J. Using the R Commander: A Point-and-click Interface for R, CRC Press., 2016
- Irizarry, R.A. Introduction to data science: data analysis and prediction algorithms with R, Boca Raton, Florida. CRC Press, 2020
- Wickham, H., & Grolemund, G. (2016) R for data science: import, tidy, transform, visualize, and model data, O'Reilly Media, Inc., 2016

RECURSOS ELECTRÓNICOS BÁSICOS

- . R for Data Science: <https://r4ds.had.co.nz/index.html>
- . R Programming for Data Science: <https://bookdown.org/rdpeng/rprogdatascience>
- . R Programming for Data Science: <https://rafalab.github.io/dsbook>