Análisis y Visualización de Datos

Curso Académico: (2021 / 2022) **Fecha de revisión: 24-06-2021**

Departamento asignado a la asignatura:

Coordinador/a: MUÑOZ GARCIA, ALBERTO

Tipo: Obligatoria Créditos ECTS : 3.0

Curso : 1 Cuatrimestre : 1

REQUISITOS (ASIGNATURAS O MATERIAS CUYO CONOCIMIENTO SE PRESUPONE)

Son convenientes nociones de Estadística.

OBJETIVOS

-Que los estudiantes posean las habilidades de aprendizaje que les permitan continuar estudiando de un modo que habrá de ser en gran medida autodirigido o autónomo.
- Capacidad de relacionar teoría y práctica, de manera que puedan aplicar conceptos y soluciones a contextos organizativos específicos.
- Capacidad de comunicación oral y escrita de los diagnósticos y soluciones que propongan
- Capacidad de trabajo en equipo
- Saber conseguir información, analizarla y extraer conclusiones de ella
- Saber desarrollar alternativas y reglas de evaluación y selección de las mismas

DESCRIPCIÓN DE CONTENIDOS: PROGRAMA

1: Introduction and Descriptive Statistics
- 1.1 Introduction to the course.
- 1.2 Introduction to R. Basics, arithmetic with R, variable assignment. Basic data types in R.
- 1.3 Vectors, matrices, factors, data frames.
- 1.4 Reading and writing data in R.

2: Exploring categorical and numerical data data.
- 2.1 Bar charts, contingency tables, counts, proportions, piecharts.
- 2.2 Histograms, boxplots, visualizing in higher dimensions.

3: Numerical Summaries.
- 3.1 Measures of center. Median, median, quartiles and quantiles.
- 3.2 Measures of variability. Variance, standard deviation, IQR.
- 3.3 Shape and transformations.
- 3.4 Outliers.

4. Case Study for lessons 1-3.

5. Multivariate Data
- 5.1 Description of multivariate data.
- 5.2 Covariance, correlation, distances.
- 5.3 Visualization of multivariate data: scatterplots, bubble plots, etc.

6. Principal Component Analysis for visualization
- 6.1 Introduction and main ideas.
- 6.2 Implementing PCA in R.
- 6.3 Case Study.

7. Cluster Analysis for data exploration
- 7.1 Introduction and main ideas.
- 7.2 Hierarchical Methods.
- 7.3 Partitioning Methods.
- 7.4 Case study.

8. Linear Regression
- 8.1 Univariate Case.

8.2 Multivariate Case.
8.3 Case Study

9. Introduction to Tidyverse.
9.1 Data wrangling
9.2 Data Visualization: ggplot2
9.3 Grouping and summarizing.

10. Final Real case study.

ACTIVIDADES FORMATIVAS, METODOLOGÍA A UTILIZAR Y RÉGIMEN DE TUTORÍAS

ACTIVIDADES FORMATIVAS
Clase teórica (15 horas)
Clases prácticas (15 horas)
Clases complementarias (5 horas)
Tutorías (10 horas)
Trabajo en grupo e individual del estudiante.

METODOLOGÍAS DOCENTES
-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.
- Resolución de casos prácticos, problemas, etc.¿ planteados por el profesor de manera individual o en grupo
- Elaboración de trabajos e informes de manera individual o en grupo

SISTEMA DE EVALUACIÓN

%50 Final Exam: It consists of a individual analysis of a data set chosen by the student or a case study proposed by the teacher.
%10 Homework. There are a maximum of 9 homework assignments. Includes participation in activities during the course.
%40 Team work. Team work consists of a statistical analysis of a business case study.

Peso porcentual del Examen Final:	50
Peso porcentual del resto de la evaluación:	50

BIBLIOGRAFÍA BÁSICA

- Antony Unwin Graphical Data Analisis with R, CRC Press, 2015
- Robert I. Kabacoff R in action. Data analysis and graphics with R, Manning, 2015

BIBLIOGRAFÍA COMPLEMENTARIA

- Brian Everitt, Torsten Hothorn An introduction to Applied Multivariate Analysis with R, Springer, 2011
- Chris Chapman, Elea McDonnell Feit R for Marketing Research and Analytics, Springer, 2015
- James E. Monogan III Political Analysis using R, Springer, 2015
- Peter Dalgaard Introductory Statistics with R, 2 Ed, Springer, 2008