

Academic Year: (2023 / 2024)

Review date: 28-04-2023

Department assigned to the subject: Business Administration Department

Coordinating teacher: MUÑOZ GARCIA, ALBERTO

Type: Compulsory ECTS Credits : 3.0

Year : 1 Semester : 1

REQUIREMENTS (SUBJECTS THAT ARE ASSUMED TO BE KNOWN)

Statistical notions are convenient.

OBJECTIVES

Acquire the ability to gather information, analyse it and draw conclusions from it, using specialised data mining software and real case studies. Acquisition of the ability to relate theory and practice, so that they can apply concepts and solutions to specific organisational contexts.

DESCRIPTION OF CONTENTS: PROGRAMME

- 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.

LEARNING ACTIVITIES AND METHODOLOGY

ACTIVIDADES FORMATIVAS

Theory (15 hours)
Practices (15 hours)
Complementary tutoring classes (5 hours)
Office Hours (10 horas)
Group Work and Individual Work

ASSESSMENT SYSTEM

%50 Class participation: There are a maximum of 5 homework assignments. Includes participation in activities during the course.

%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.

% end-of-term-examination:	50
% of continuous assessment (assignments, laboratory, practicals...):	50

BASIC BIBLIOGRAPHY

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

ADDITIONAL BIBLIOGRAPHY

- 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