Course guide
270215 - AD - Data Analysis

Last modified: 30/01/2024

Unit in charge: Barcelona School of Informatics
Teaching unit: 715 - EIO - Department of Statistics and Operations Research.

Degree: BACHELOR'S DEGREE IN DATA SCIENCE AND ENGINEERING (Syllabus 2017). (Compulsory subject).

Academic year: 2023 **ECTS Credits:** 6.0 **Languages:** Catalan

LECTURER

Coordinating lecturer: JAN GRAFFELMAN - JOSE ANTONIO SÁNCHEZ ESPIGARES

Others: Segon quadrimestre:
NIHAN ACAR DENIZLI - 11, 12
JOSE ANTONIO SÁNCHEZ ESPIGARES - 11, 12

PRIOR SKILLS

Knowledge of basic statistical concepts, descriptive statistics, hypothesis testing. Familiarity with the statistical software R.

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
CE1. Skillfully use mathematical concepts and methods that underlie the problems of science and data engineering.
CE2. To be able to program solutions to engineering problems: Design efficient algorithmic solutions to a given computational problem, implement them in the form of a robust, structured and maintainable program, and check the validity of the solution.
CE3. Analyze complex phenomena through probability and statistics, and propose models of these types in specific situations. Formulate and solve mathematical optimization problems.
CE4. Use current computer systems, including high performance systems, for the process of large volumes of data from the knowledge of its structure, operation and particularities.
CE8. Ability to choose and employ techniques of statistical modeling and data analysis, evaluating the quality of the models, validating and interpreting them.

Generical:
CG1. To design computer systems that integrate data of provenances and very diverse forms, create with them mathematical models, reason on these models and act accordingly, learning from experience.
CG2. Choose and apply the most appropriate methods and techniques to a problem defined by data that represents a challenge for its volume, speed, variety or heterogeneity, including computer, mathematical, statistical and signal processing methods.
CG3. Work in multidisciplinary teams and projects related to the processing and exploitation of complex data, interacting fluently with engineers and professionals from other disciplines.
CG4. Identify opportunities for innovative data-driven applications in evolving technological environments.

Transversal:
CT3. Efficient oral and written communication. Communicate in an oral and written way with other people about the results of learning, thinking and decision making; Participate in debates on topics of the specialty itself.
CT4. Teamwork. Be able to work as a member of an interdisciplinary team, either as a member or conducting management tasks, with the aim of contributing to develop projects with pragmatism and a sense of responsibility, taking commitments taking into account available resources.
CT5. Solvent use of information resources. Manage the acquisition, structuring, analysis and visualization of data and information in the field of specialty and critically evaluate the results of such management.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CT7. Third language. Know a third language, preferably English, with an adequate oral and written level and in line with the needs of graduates.

Basic:
CB2. That the students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and defense of arguments and problem solving within their area of ??study.
CB4. That the students can transmit information, ideas, problems and solutions to a specialized and non-specialized public.

TEACHING METHODOLOGY

The learning process is a combination of theoretical explanation and practical application. The theory classes are used to explain the basic scientific contents of the course, whereas the laboratory sessions work on their application to solve real-life problems.

Practicals and project form the basis for working out the transversal competences of the students, related to team-work and public presentation of results. Practicals and project also serve to integrate the different pieces of knowledge of the course.

For hands-on computer training we use the R statistical environment.

LEARNING OBJECTIVES OF THE SUBJECT

1.Exploratory Data Analysis
2.Discriminant Analysis with probabilistic hypothesis
3.Multivariate modeling
4.Time series

STUDY LOAD

Type	Hours	Percentage
Self study	90,0	60.00
Hours large group	30,0	20.00
Hours small group	30,0	20.00

Total learning time: 150 h

CONTENTS

Data preprocessing

Description:
Outliers, missing data and transformations

Principal component analysis

Description:
Multivariate description of a table of continous variables. Regression with principal components.

Factor analysis

Description:
The singular value decomposition, biplots, factor analysis

Multidimensional scaling (MDS)

Description:
Distance measures. Metric multidimensional scaling. Algorithms.

Cluster analysis

Description:
Hierarchical clustering techniques. Agglomeration methods. Ward's criterion. Dendrogram.

Correspondence analysis

Description:
Contingency tables. Row and column profiles. Independence and chi-square statistics. Simple correspondence analysis. Biplot.

Discriminant analysis

Description:
Multivariate normal distribution. Fisher's linear discriminant analysis.

Univariate time series models

Description:
Exponential smoothing, ARIMA models

Intervention analysis

Description:
Outliers, seasonal effects, intervention analysis.

ACTIVITIES

Data preprocessing

Description:
Practical on data preprocessing

Specific objectives:
1

Related competencies :
CG1. To design computer systems that integrate data of provenances and very diverse forms, create with them mathematical models, reason on these models and act accordingly, learning from experience.
CG4. Identify opportunities for innovative data-driven applications in evolving technological environments.
CG3. Work in multidisciplinary teams and projects related to the processing and exploitation of complex data, interacting fluently with engineers and professionals from other disciplines.
CE3. Analyze complex phenomena through probability and statistics, and propose models of these types in specific situations. Formulate and solve mathematical optimization problems.
CE4. Use current computer systems, including high performance systems, for the process of large volumes of data from the knowledge of its structure, operation and particularities.
CE1. Skillfully use mathematical concepts and methods that underlie the problems of science and data engineering.
CE8. Ability to choose and employ techniques of statistical modeling and data analysis, evaluating the quality of the models, validating and interpreting them.
CE2. To be able to program solutions to engineering problems: Design efficient algorithmic solutions to a given computational problem, implement them in the form of a robust, structured and maintainable program, and check the validity of the solution.
CT7. Third language. Know a third language, preferably English, with an adequate oral and written level and in line with the needs of graduates.
CT5. Solvent use of information resources. Manage the acquisition, structuring, analysis and visualization of data and information in the field of specialty and critically evaluate the results of such management.
CT3. Efficient oral and written communication. Communicate in an oral and written way with other people about the results of learning, thinking and decision making; Participate in debates on topics of the specialty itself.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CB4. That the students can transmit information, ideas, problems and solutions to a specialized and non-specialized public.
CB2. That the students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and defense of arguments and problem solving within their area of ??study.

Full-or-part-time: 12h
Theory classes: 4h
Laboratory classes: 4h
Self study: 4h

Principal component analysis

Description:
Application of principal component analysis in practical data analysis

Specific objectives:
1

Related competencies :
CG3. Work in multidisciplinary teams and projects related to the processing and exploitation of complex data, interacting fluently with engineers and professionals from other disciplines.
CG1. To design computer systems that integrate data of provenances and very diverse forms, create with them mathematical models, reason on these models and act accordingly, learning from experience.
CG4. Identify opportunities for innovative data-driven applications in evolving technological environments.
CE3. Analyze complex phenomena through probability and statistics, and propose models of these types in specific situations. Formulate and solve mathematical optimization problems.
CE4. Use current computer systems, including high performance systems, for the process of large volumes of data from the knowledge of its structure, operation and particularities.
CE1. Skillfully use mathematical concepts and methods that underlie the problems of science and data engineering.
CE8. Ability to choose and employ techniques of statistical modeling and data analysis, evaluating the quality of the models, validating and interpreting them.
CE2. To be able to program solutions to engineering problems: Design efficient algorithmic solutions to a given computational problem, implement them in the form of a robust, structured and maintainable program, and check the validity of the solution.
CT7. Third language. Know a third language, preferably English, with an adequate oral and written level and in line with the needs of graduates.
CT5. Solvent use of information resources. Manage the acquisition, structuring, analysis and visualization of data and information in the field of specialty and critically evaluate the results of such management.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CT3. Efficient oral and written communication. Communicate in an oral and written way with other people about the results of learning, thinking and decision making; Participate in debates on topics of the specialty itself.
CB4. That the students can transmit information, ideas, problems and solutions to a specialized and non-specialized public.
CB2. That the students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and defense of arguments and problem solving within their area of ??study.

Full-or-part-time: 14h
Theory classes: 4h
Laboratory classes: 4h
Self study: 6h

Factor analysis

Description:
Practical data analysis using the method

Specific objectives:
1

Related competencies :
CG3. Work in multidisciplinary teams and projects related to the processing and exploitation of complex data, interacting fluently with engineers and professionals from other disciplines.
CG1. To design computer systems that integrate data of provenances and very diverse forms, create with them mathematical models, reason on these models and act accordingly, learning from experience.
CG4. Identify opportunities for innovative data-driven applications in evolving technological environments.
CE3. Analyze complex phenomena through probability and statistics, and propose models of these types in specific situations. Formulate and solve mathematical optimization problems.
CE4. Use current computer systems, including high performance systems, for the process of large volumes of data from the knowledge of its structure, operation and particularities.
CE1. Skillfully use mathematical concepts and methods that underlie the problems of science and data engineering.
CE8. Ability to choose and employ techniques of statistical modeling and data analysis, evaluating the quality of the models, validating and interpreting them.
CE2. To be able to program solutions to engineering problems: Design efficient algorithmic solutions to a given computational problem, implement them in the form of a robust, structured and maintainable program, and check the validity of the solution.
CT7. Third language. Know a third language, preferably English, with an adequate oral and written level and in line with the needs of graduates.
CT5. Solvent use of information resources. Manage the acquisition, structuring, analysis and visualization of data and information in the field of specialty and critically evaluate the results of such management.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CT3. Efficient oral and written communication. Communicate in an oral and written way with other people about the results of learning, thinking and decision making; Participate in debates on topics of the specialty itself.
CB4. That the students can transmit information, ideas, problems and solutions to a specialized and non-specialized public.
CB2. That the students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and defense of arguments and problem solving within their area of ??study.

Full-or-part-time: 9h
Theory classes: 2h
Laboratory classes: 3h
Self study: 4h

Multidimensional scaling

Description:
Analysis of distance matrices with this method

Specific objectives:
1

Related competencies :
CG3. Work in multidisciplinary teams and projects related to the processing and exploitation of complex data, interacting fluently with engineers and professionals from other disciplines.
CG1. To design computer systems that integrate data of provenances and very diverse forms, create with them mathematical models, reason on these models and act accordingly, learning from experience.
CG4. Identify opportunities for innovative data-driven applications in evolving technological environments.
CE3. Analyze complex phenomena through probability and statistics, and propose models of these types in specific situations. Formulate and solve mathematical optimization problems.
CE4. Use current computer systems, including high performance systems, for the process of large volumes of data from the knowledge of its structure, operation and particularities.
CE1. Skillfully use mathematical concepts and methods that underlie the problems of science and data engineering.
CE8. Ability to choose and employ techniques of statistical modeling and data analysis, evaluating the quality of the models, validating and interpreting them.
CE2. To be able to program solutions to engineering problems: Design efficient algorithmic solutions to a given computational problem, implement them in the form of a robust, structured and maintainable program, and check the validity of the solution.
CT7. Third language. Know a third language, preferably English, with an adequate oral and written level and in line with the needs of graduates.
CT5. Solvent use of information resources. Manage the acquisition, structuring, analysis and visualization of data and information in the field of specialty and critically evaluate the results of such management.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CT3. Efficient oral and written communication. Communicate in an oral and written way with other people about the results of learning, thinking and decision making; Participate in debates on topics of the specialty itself.
CB4. That the students can transmit information, ideas, problems and solutions to a specialized and non-specialized public.
CB2. That the students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and defense of arguments and problem solving within their area of ??study.

Full-or-part-time: 8h
Theory classes: 2h
Laboratory classes: 2h
Self study: 4h

Clustering

Description:
Application of the method to quantitative data matrices.

Full-or-part-time: 12h
Theory classes: 4h
Laboratory classes: 4h
Self study: 4h

Correspondence Analysis

Description:
Application of the method with cross tables.

Specific objectives:
2

Related competencies :
CG2. Choose and apply the most appropriate methods and techniques to a problem defined by data that represents a challenge for its volume, speed, variety or heterogeneity, including computer, mathematical, statistical and signal processing methods.
CG3. Work in multidisciplinary teams and projects related to the processing and exploitation of complex data, interacting fluently with engineers and professionals from other disciplines.
CE3. Analyze complex phenomena through probability and statistics, and propose models of these types in specific situations. Formulate and solve mathematical optimization problems.
CE1. Skillfully use mathematical concepts and methods that underlie the problems of science and data engineering.
CE8. Ability to choose and employ techniques of statistical modeling and data analysis, evaluating the quality of the models, validating and interpreting them.
CT7. Third language. Know a third language, preferably English, with an adequate oral and written level and in line with the needs of graduates.
CT5. Solvent use of information resources. Manage the acquisition, structuring, analysis and visualization of data and information in the field of specialty and critically evaluate the results of such management.
CT4. Teamwork. Be able to work as a member of an interdisciplinary team, either as a member or conducting management tasks, with the aim of contributing to develop projects with pragmatism and a sense of responsibility, taking commitments taking into account available resources.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CT3. Efficient oral and written communication. Communicate in an oral and written way with other people about the results of learning, thinking and decision making; Participate in debates on topics of the specialty itself.

Full-or-part-time: 8h
Theory classes: 2h
Laboratory classes: 2h
Self study: 4h

Discriminant Analysis

Description:
Application of the method to empirical data sets

Specific objectives:
2

Related competencies :
CG2. Choose and apply the most appropriate methods and techniques to a problem defined by data that represents a challenge for its volume, speed, variety or heterogeneity, including computer, mathematical, statistical and signal processing methods.
CG3. Work in multidisciplinary teams and projects related to the processing and exploitation of complex data, interacting fluently with engineers and professionals from other disciplines.
CE3. Analyze complex phenomena through probability and statistics, and propose models of these types in specific situations. Formulate and solve mathematical optimization problems.
CE1. Skillfully use mathematical concepts and methods that underlie the problems of science and data engineering.
CE8. Ability to choose and employ techniques of statistical modeling and data analysis, evaluating the quality of the models, validating and interpreting them.
CT7. Third language. Know a third language, preferably English, with an adequate oral and written level and in line with the needs of graduates.
CT5. Solvent use of information resources. Manage the acquisition, structuring, analysis and visualization of data and information in the field of specialty and critically evaluate the results of such management.
CT4. Teamwork. Be able to work as a member of an interdisciplinary team, either as a member or conducting management tasks, with the aim of contributing to develop projects with pragmatism and a sense of responsibility, taking commitments taking into account available resources.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CT3. Efficient oral and written communication. Communicate in an oral and written way with other people about the results of learning, thinking and decision making; Participate in debates on topics of the specialty itself.

Full-or-part-time: 12h
Theory classes: 4h
Laboratory classes: 4h
Self study: 4h

Univariate time series models

Description:
Fitting time series models to data sets on the computer

Specific objectives:
4

Related competencies :
CE3. Analyze complex phenomena through probability and statistics, and propose models of these types in specific situations. Formulate and solve mathematical optimization problems.
CE1. Skillfully use mathematical concepts and methods that underlie the problems of science and data engineering.
CE8. Ability to choose and employ techniques of statistical modeling and data analysis, evaluating the quality of the models, validating and interpreting them.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.

Full-or-part-time: 14h
Theory classes: 4h
Laboratory classes: 4h
Self study: 6h

Intervention analysis

Description:
Application of intervention analysis to real data sets

Specific objectives:
4

Related competencies :
CE3. Analyze complex phenomena through probability and statistics, and propose models of these types in specific situations. Formulate and solve mathematical optimization problems.
CE1. Skillfully use mathematical concepts and methods that underlie the problems of science and data engineering.
CE8. Ability to choose and employ techniques of statistical modeling and data analysis, evaluating the quality of the models, validating and interpreting them.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.

Full-or-part-time: 9h
Theory classes: 2h
Laboratory classes: 3h
Self study: 4h

Practical on exploratory data analysis

Description:
Student do an exploratory analysis of a data set and hand in a questionnaire about it.

Specific objectives:
1, 2, 3, 4

Related competencies :
CG2. Choose and apply the most appropriate methods and techniques to a problem defined by data that represents a challenge for its volume, speed, variety or heterogeneity, including computer, mathematical, statistical and signal processing methods.
CG3. Work in multidisciplinary teams and projects related to the processing and exploitation of complex data, interacting fluently with engineers and professionals from other disciplines.
CG1. To design computer systems that integrate data of provenances and very diverse forms, create with them mathematical models, reason on these models and act accordingly, learning from experience.
CG4. Identify opportunities for innovative data-driven applications in evolving technological environments.
CE3. Analyze complex phenomena through probability and statistics, and propose models of these types in specific situations. Formulate and solve mathematical optimization problems.
CE4. Use current computer systems, including high performance systems, for the process of large volumes of data from the knowledge of its structure, operation and particularities.
CE1. Skillfully use mathematical concepts and methods that underlie the problems of science and data engineering.
CE8. Ability to choose and employ techniques of statistical modeling and data analysis, evaluating the quality of the models, validating and interpreting them.
CE2. To be able to program solutions to engineering problems: Design efficient algorithmic solutions to a given computational problem, implement them in the form of a robust, structured and maintainable program, and check the validity of the solution.
CT7. Third language. Know a third language, preferably English, with an adequate oral and written level and in line with the needs of graduates.
CT5. Solvent use of information resources. Manage the acquisition, structuring, analysis and visualization of data and information in the field of specialty and critically evaluate the results of such management.
CT4. Teamwork. Be able to work as a member of an interdisciplinary team, either as a member or conducting management tasks, with the aim of contributing to develop projects with pragmatism and a sense of responsibility, taking commitments taking into account available resources.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CT3. Efficient oral and written communication. Communicate in an oral and written way with other people about the results of learning, thinking and decision making; Participate in debates on topics of the specialty itself.
CB4. That the students can transmit information, ideas, problems and solutions to a specialized and non-specialized public.
CB2. That the students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and defense of arguments and problem solving within their area of ??study.

Full-or-part-time: 18h
Guided activities: 3h
Self study: 15h

Project

Description:
Students realize, in couples, a complete multivariate study of a certain dataset using the techniques they studied during the course, and hand in a written report about it.

Specific objectives:
1, 2, 3, 4

Related competencies :
CG2. Choose and apply the most appropriate methods and techniques to a problem defined by data that represents a challenge for its volume, speed, variety or heterogeneity, including computer, mathematical, statistical and signal processing methods.
CG3. Work in multidisciplinary teams and projects related to the processing and exploitation of complex data, interacting fluently with engineers and professionals from other disciplines.
CG1. To design computer systems that integrate data of provenances and very diverse forms, create with them mathematical models, reason on these models and act accordingly, learning from experience.
CG4. Identify opportunities for innovative data-driven applications in evolving technological environments.
CE3. Analyze complex phenomena through probability and statistics, and propose models of these types in specific situations. Formulate and solve mathematical optimization problems.
CE4. Use current computer systems, including high performance systems, for the process of large volumes of data from the knowledge of its structure, operation and particularities.
CE1. Skillfully use mathematical concepts and methods that underlie the problems of science and data engineering.
CE8. Ability to choose and employ techniques of statistical modeling and data analysis, evaluating the quality of the models, validating and interpreting them.
CE2. To be able to program solutions to engineering problems: Design efficient algorithmic solutions to a given computational problem, implement them in the form of a robust, structured and maintainable program, and check the validity of the solution.
CT7. Third language. Know a third language, preferably English, with an adequate oral and written level and in line with the needs of graduates.
CT5. Solvent use of information resources. Manage the acquisition, structuring, analysis and visualization of data and information in the field of specialty and critically evaluate the results of such management.
CT4. Teamwork. Be able to work as a member of an interdisciplinary team, either as a member or conducting management tasks, with the aim of contributing to develop projects with pragmatism and a sense of responsibility, taking commitments taking into account available resources.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CT3. Efficient oral and written communication. Communicate in an oral and written way with other people about the results of learning, thinking and decision making; Participate in debates on topics of the specialty itself.
CB4. That the students can transmit information, ideas, problems and solutions to a specialized and non-specialized public.
CB2. That the students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and defense of arguments and problem solving within their area of ??study.

Full-or-part-time: 16h
Guided activities: 3h
Self study: 13h

Exam concering basic concepts

Description:
There are two exams related to the theoretical concepts of the course.

Specific objectives:
1, 2, 3, 4

Related competencies :
CG2. Choose and apply the most appropriate methods and techniques to a problem defined by data that represents a challenge for its volume, speed, variety or heterogeneity, including computer, mathematical, statistical and signal processing methods.
CG3. Work in multidisciplinary teams and projects related to the processing and exploitation of complex data, interacting fluently with engineers and professionals from other disciplines.
CG1. To design computer systems that integrate data of provenances and very diverse forms, create with them mathematical models, reason on these models and act accordingly, learning from experience.
CG4. Identify opportunities for innovative data-driven applications in evolving technological environments.
CE3. Analyze complex phenomena through probability and statistics, and propose models of these types in specific situations. Formulate and solve mathematical optimization problems.
CE4. Use current computer systems, including high performance systems, for the process of large volumes of data from the knowledge of its structure, operation and particularities.
CE1. Skillfully use mathematical concepts and methods that underlie the problems of science and data engineering.
CE8. Ability to choose and employ techniques of statistical modeling and data analysis, evaluating the quality of the models, validating and interpreting them.
CE2. To be able to program solutions to engineering problems: Design efficient algorithmic solutions to a given computational problem, implement them in the form of a robust, structured and maintainable program, and check the validity of the solution.
CT7. Third language. Know a third language, preferably English, with an adequate oral and written level and in line with the needs of graduates.
CT5. Solvent use of information resources. Manage the acquisition, structuring, analysis and visualization of data and information in the field of specialty and critically evaluate the results of such management.
CT4. Teamwork. Be able to work as a member of an interdisciplinary team, either as a member or conducting management tasks, with the aim of contributing to develop projects with pragmatism and a sense of responsibility, taking commitments taking into account available resources.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CT3. Efficient oral and written communication. Communicate in an oral and written way with other people about the results of learning, thinking and decision making; Participate in debates on topics of the specialty itself.
CB4. That the students can transmit information, ideas, problems and solutions to a specialized and non-specialized public.
CB2. That the students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and defense of arguments and problem solving within their area of ??study.

Full-or-part-time: 16h 30m
Guided activities: 2h
Self study: 14h 30m

GRADING SYSTEM

The student's final grade for the course is based on grades obtained for weekly homework assignments (25%), a partial exam half-way the course (25%), a final exam covering the second half of the course (25%) and a project (25%).

Each weekly assignments consists of resolving a questionnaire. These assigments aim at consolidating knowledge of the techniques exposed in the theoretical sessions. The assignments require analysis of datasets in the statistical environment R.

A project is carried out by a group of two students, and students have to show they can resolve problems with the techniques they have learned during the course. Each group hands in a written report about their project at the end of the course.

The two exams will be programmed according to the calendar of the faculty, and evaluate if students have assimilated the basic concepts of the material of the course.

For the resit exam, the student can choose to do a re-examination of only the first partial (25%), or of only the second partial (25%), or of both partials (50%). The re-evaluation thus represents at most 50% of the final course grade.

BIBLIOGRAPHY

Basic:
- Manly, B.F.J.; Navarro, J.A. Multivariate statistical methods: a primer. 4th ed. Boca Raton: CRC Press, Taylor & Francis Group, 2017. ISBN 9781498728966.
- Johnson, R.A.; Wichern, D.W. Applied multivariate statistical analysis. 6th ed. Harlow, Essex: Pearson, 2014. ISBN 9781292024943.
- Peña, D. Análisis de datos multivariantes. Madrdi [etc.]: McGraw-Hill, cop. 2002. ISBN 9788448136109.
- Cuadras, C.M. Nuevos métodos de análisis multivariante. CMC Ediciones, 2012.
- Shumway, R.H.; Stoffer, D.S. Time series analysis and its applications: with R examples. 4th ed. Springer, 2017. ISBN 9783319524511.
- Graffelman, Jan. Course slides for Multivariate Analysis (in English).

Complementary:
- Mardia, K.V; Kent, J.T; Bibby, J.M. Multivariate analysis. Academic Press, 1979. ISBN 0124712509.
- Anderson, T.W. An introduction to multivariate statistical analysis. 3rd ed. Wiley, 2003. ISBN 0471360910.
- Aluja, T.; Morineau, A. Aprender de los datos: el análisis de componentes principales: una aproximación desde el Data Mining. EUB, 1999. ISBN 8483120224.
- Box, G.E.P.; Jenkins, G.M.; Reinsel, G.C.; Ljung, G.M. Time series analysis: forecasting and control. 5th ed. Wiley, 2016. ISBN 9781118675021.
- Peña, D. Análisis de series temporales. 2a ed. Madrid: Alianza, 2010. ISBN 9788420669458.
- Brockwell, P.J.; Davis, R.A. Time series: theory and methods. 2nd ed. Springer-Verlag, 1991. ISBN 9781441903198.