

Course guide

270100 - ADEI - Data Analysis and Information Exploitation

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 INFORMATICS ENGINEERING (Syllabus 2010). (Optional subject).

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

LECTURER

Coordinating lecturer: XAVIER ANGERRI TORREDEFLOT - JOSEP FRANQUET FÀBREGAS - BHUMIKA ASHVINBHAI PATEL - LIDIA MONTERO MERCADÉ

Others: Primer quadrimestre:
XAVIER ANGERRI TORREDEFLOT - 10
BHUMIKA ASHVINBHAI PATEL - 10

Segon quadrimestre:
XAVIER ANGERRI TORREDEFLOT - 10
JOSEP FRANQUET FÀBREGAS - 10

PRIOR SKILLS

Students must have completed a course in probability and statistics and a course on business and economic environment

REQUIREMENTS

- Prerequisite PE

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:

CSI2.1. To demonstrate comprehension and apply the management principles and techniques about quality and technological innovation in the organizations.

CSI2.3. To demonstrate knowledge and application capacity of extraction and knowledge management systems .

Generical:

G3. THIRD LANGUAGE: to know the English language in a correct oral and written level, and accordingly to the needs of the graduates in Informatics Engineering. Capacity to work in a multidisciplinary group and in a multi-language environment and to communicate, orally and in a written way, knowledge, procedures, results and ideas related to the technical informatics engineer profession.

G9. PROPER THINKING HABITS: capacity of critical, logical and mathematical reasoning. Capacity to solve problems in her study area. Abstraction capacity: capacity to create and use models that reflect real situations. Capacity to design and perform simple experiments and analyse and interpret its results. Analysis, synthesis and evaluation capacity.

TEACHING METHODOLOGY

Learning the course consists of three distinct phases:

1. Acquisition of specific knowledge through the study of literature and material provided by teachers.
2. The acquisition of skills in specific techniques of data analysis and exploitation of information and
3. Integration of knowledge, skills and competencies (specific and generic) by solving a real Case Study.

In theory classes serve to expose the foundations of methodologies and techniques of the subject.

The laboratory classes are used to learn the use of specific techniques for solving problems, using appropriate informatics tools, in this sense, students first must repeat the problem solved previously by the teachers and then solve a similar one.

While the case study, are settled in groups in selflearning hours, and serves to put into practice the knowledge, skills and competences in solving a real case of ADEI.

LEARNING OBJECTIVES OF THE SUBJECT

1. Learn how to identify the three levels of decision making in a company
2. Control Quality
3. Control of discrete indicators
4. Determining the drivers of continuous response.
5. Diagnosis of a statistical model
6. Modelling of discrete choices
7. Modelling the propensity
8. Analysis of databases. Determination of the significant characteristics of groups of individuals.
9. Concept and measurement of intangibles in a company
10. Multivariate information visualization
11. Clustering
12. Modelling intangibles. Models for consumer satisfaction
13. Statistical tools for support decision making
15. Continuous process control
16. Learn how to make a report on data quality

STUDY LOAD

Type	Hours	Percentage
Hours large group	30,0	20.00
Hours small group	30,0	20.00
Guided activities	6,0	4.00
Self study	84,0	56.00

Total learning time: 150 h

CONTENTS

Bloc1: Levels of corporate decision

Block 2: Summary description and data quality

Block 3: Statistical Modeling

Block 4: Multivariate Data Analysis and intangible measurement

Block 5: Clustering and profiling

ACTIVITIES

Presentation of the Case of Study

Specific objectives:

1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16

Related competencies :

G3. THIRD LANGUAGE: to know the English language in a correct oral and written level, and accordingly to the needs of the graduates in Informatics Engineering. Capacity to work in a multidisciplinary group and in a multi-language environment and to communicate, orally and in a written way, knowledge, procedures, results and ideas related to the technical informatics engineer profession.

G9. PROPER THINKING HABITS: capacity of critical, logical and mathematical reasoning. Capacity to solve problems in her study area. Abstraction capacity: capacity to create and use models that reflect real situations. Capacity to design and perform simple experiments and analyse and interpret its results. Analysis, synthesis and evaluation capacity.

Full-or-part-time: 14h

Guided activities: 2h

Self study: 12h

Quiz Blocks 4 and 5

Specific objectives:

1, 8, 9, 10, 11, 12

Related competencies :

G9. PROPER THINKING HABITS: capacity of critical, logical and mathematical reasoning. Capacity to solve problems in her study area. Abstraction capacity: capacity to create and use models that reflect real situations. Capacity to design and perform simple experiments and analyse and interpret its results. Analysis, synthesis and evaluation capacity.

Full-or-part-time: 14h

Guided activities: 2h

Self study: 12h

Quiz blocks 2 and 3

Specific objectives:

1, 2, 4, 5, 6, 16

Related competencies :

G9. PROPER THINKING HABITS: capacity of critical, logical and mathematical reasoning. Capacity to solve problems in her study area. Abstraction capacity: capacity to create and use models that reflect real situations. Capacity to design and perform simple experiments and analyse and interpret its results. Analysis, synthesis and evaluation capacity.

Full-or-part-time: 14h

Guided activities: 2h

Self study: 12h

Block 1. Levels of corporate decision

Description:

It presented the three levels of decision making in companies. What are the main business processes and how is stored the generated data.

Specific objectives:

1

Full-or-part-time: 5h

Theory classes: 2h

Laboratory classes: 2h

Self study: 1h

Block 2. Description and quality control data

Description:

Problems in data quality: This is seen in the Case Study or problems that may present data: inconsistency, redundancy. Missing data. Outliers. How do I report data quality. What is the standardization of data.

Specific objectives:

16

Related competencies :

G9. PROPER THINKING HABITS: capacity of critical, logical and mathematical reasoning. Capacity to solve problems in her study area. Abstraction capacity: capacity to create and use models that reflect real situations. Capacity to design and perform simple experiments and analyse and interpret its results. Analysis, synthesis and evaluation capacity.

Full-or-part-time: 8h

Theory classes: 2h

Laboratory classes: 2h

Self study: 4h

Block 2. Data visualisation

Description:

Type of Data Collection and applicability to operational control. Indicators common in continuous process control

Specific objectives:

2, 13, 16

Related competencies :

G3. THIRD LANGUAGE: to know the English language in a correct oral and written level, and accordingly to the needs of the graduates in Informatics Engineering. Capacity to work in a multidisciplinary group and in a multi-language environment and to communicate, orally and in a written way, knowledge, procedures, results and ideas related to the technical informatics engineer profession.

G9. PROPER THINKING HABITS: capacity of critical, logical and mathematical reasoning. Capacity to solve problems in her study area. Abstraction capacity: capacity to create and use models that reflect real situations. Capacity to design and perform simple experiments and analyse and interpret its results. Analysis, synthesis and evaluation capacity.

Full-or-part-time: 7h

Theory classes: 2h

Laboratory classes: 2h

Self study: 3h

Block 3. Statistical Modeling

Specific objectives:

4, 13

Related competencies :

G3. THIRD LANGUAGE: to know the English language in a correct oral and written level, and accordingly to the needs of the graduates in Informatics Engineering. Capacity to work in a multidisciplinary group and in a multi-language environment and to communicate, orally and in a written way, knowledge, procedures, results and ideas related to the technical informatics engineer profession.

G9. PROPER THINKING HABITS: capacity of critical, logical and mathematical reasoning. Capacity to solve problems in her study area. Abstraction capacity: capacity to create and use models that reflect real situations. Capacity to design and perform simple experiments and analyse and interpret its results. Analysis, synthesis and evaluation capacity.

Full-or-part-time: 12h

Theory classes: 2h

Laboratory classes: 2h

Self study: 8h

Block 3. Training the model

Specific objectives:

4

Related competencies :

G9. PROPER THINKING HABITS: capacity of critical, logical and mathematical reasoning. Capacity to solve problems in her study area. Abstraction capacity: capacity to create and use models that reflect real situations. Capacity to design and perform simple experiments and analyse and interpret its results. Analysis, synthesis and evaluation capacity.

Full-or-part-time: 7h

Theory classes: 2h

Laboratory classes: 2h

Self study: 3h

Blok 3. Validation of statistical modeling

Description:

Elements involved in the validation of regression modeling. Values ☐☐influential and / or outliers

Specific objectives:

5

Related competencies :

G9. PROPER THINKING HABITS: capacity of critical, logical and mathematical reasoning. Capacity to solve problems in her study area. Abstraction capacity: capacity to create and use models that reflect real situations. Capacity to design and perform simple experiments and analyse and interpret its results. Analysis, synthesis and evaluation capacity.

Full-or-part-time: 7h

Theory classes: 2h

Laboratory classes: 2h

Self study: 3h

Bloc 3. Statistical Modeling of binary variables

Specific objectives:

5, 6

Related competencies :

G9. PROPER THINKING HABITS: capacity of critical, logical and mathematical reasoning. Capacity to solve problems in her study area. Abstraction capacity: capacity to create and use models that reflect real situations. Capacity to design and perform simple experiments and analyse and interpret its results. Analysis, synthesis and evaluation capacity.

Full-or-part-time: 12h

Theory classes: 2h

Laboratory classes: 2h

Self study: 8h

Block 4. Multivariate Data Analysis

Specific objectives:

9, 10

Related competencies :

G9. PROPER THINKING HABITS: capacity of critical, logical and mathematical reasoning. Capacity to solve problems in her study area. Abstraction capacity: capacity to create and use models that reflect real situations. Capacity to design and perform simple experiments and analyse and interpret its results. Analysis, synthesis and evaluation capacity.

Full-or-part-time: 6h

Theory classes: 2h

Laboratory classes: 1h

Self study: 3h

Block 4. Principal Component Analysis

Specific objectives:

9, 10

Related competencies :

G9. PROPER THINKING HABITS: capacity of critical, logical and mathematical reasoning. Capacity to solve problems in her study area. Abstraction capacity: capacity to create and use models that reflect real situations. Capacity to design and perform simple experiments and analyse and interpret its results. Analysis, synthesis and evaluation capacity.

Full-or-part-time: 9h

Theory classes: 2h

Laboratory classes: 2h

Self study: 5h

Blok 4. Measurement of intangibles

Specific objectives:

9, 10

Related competencies :

G9. PROPER THINKING HABITS: capacity of critical, logical and mathematical reasoning. Capacity to solve problems in her study area. Abstraction capacity: capacity to create and use models that reflect real situations. Capacity to design and perform simple experiments and analyse and interpret its results. Analysis, synthesis and evaluation capacity.

Full-or-part-time: 7h

Theory classes: 2h

Laboratory classes: 2h

Self study: 3h

Blok 4. Practice of Principal Component Analysis

Description:

Practice Principal Component Analysis, interpretation of the representations obtained. Positioning of the supplementary information.

Specific objectives:

9, 10, 13

Related competencies :

G3. THIRD LANGUAGE: to know the English language in a correct oral and written level, and accordingly to the needs of the graduates in Informatics Engineering. Capacity to work in a multidisciplinary group and in a multi-language environment and to communicate, orally and in a written way, knowledge, procedures, results and ideas related to the technical informatics engineer profession.

G9. PROPER THINKING HABITS: capacity of critical, logical and mathematical reasoning. Capacity to solve problems in her study area. Abstraction capacity: capacity to create and use models that reflect real situations. Capacity to design and perform simple experiments and analyse and interpret its results. Analysis, synthesis and evaluation capacity.

Full-or-part-time: 7h

Theory classes: 2h

Laboratory classes: 2h

Self study: 3h

Blok 5. Clustering

Specific objectives:

11

Related competencies :

G9. PROPER THINKING HABITS: capacity of critical, logical and mathematical reasoning. Capacity to solve problems in her study area. Abstraction capacity: capacity to create and use models that reflect real situations. Capacity to design and perform simple experiments and analyse and interpret its results. Analysis, synthesis and evaluation capacity.

Full-or-part-time: 7h

Theory classes: 2h

Laboratory classes: 2h

Self study: 3h

Block 4. Practice of Clustering

Description:

Presentation of the k-means and hierarchical methods.

Specific objectives:

11, 13

Related competencies :

G3. THIRD LANGUAGE: to know the English language in a correct oral and written level, and accordingly to the needs of the graduates in Informatics Engineering. Capacity to work in a multidisciplinary group and in a multi-language environment and to communicate, orally and in a written way, knowledge, procedures, results and ideas related to the technical informatics engineer profession.

G9. PROPER THINKING HABITS: capacity of critical, logical and mathematical reasoning. Capacity to solve problems in her study area. Abstraction capacity: capacity to create and use models that reflect real situations. Capacity to design and perform simple experiments and analyse and interpret its results. Analysis, synthesis and evaluation capacity.

Full-or-part-time: 6h

Theory classes: 2h

Laboratory classes: 1h

Self study: 3h

Block 5. profiling

Specific objectives:

8, 13

Related competencies :

G3. THIRD LANGUAGE: to know the English language in a correct oral and written level, and accordingly to the needs of the graduates in Informatics Engineering. Capacity to work in a multidisciplinary group and in a multi-language environment and to communicate, orally and in a written way, knowledge, procedures, results and ideas related to the technical informatics engineer profession.

G9. PROPER THINKING HABITS: capacity of critical, logical and mathematical reasoning. Capacity to solve problems in her study area. Abstraction capacity: capacity to create and use models that reflect real situations. Capacity to design and perform simple experiments and analyse and interpret its results. Analysis, synthesis and evaluation capacity.

Full-or-part-time: 8h

Theory classes: 2h

Laboratory classes: 2h

Self study: 4h

GRADING SYSTEM

The evaluation of the course integrates the three phases of learning process: knowledge, skills and competencies.

The knowledge is assessed by two quizzes, in the middle and last week of the course. If you fail this exam, students may have a final resit. (score T).

The skills assessed from the delivery from 2 to 5 practices relating to the course case study. Each of the blocks 2 and 3 involve a practice that students will perform either individually or in groups of 2, the same for blocks 4 and 5. The average of the scores comes out the L score.

The case study as a whole exercise will be evaluated based on the oral presentation (score P).

In the presentation of case study that generic skills will be assessed. In any case, the presentation of the case study is compulsory.

The final grade will be obtained weighing the three scores: Final Mark = $0.4P + 0.3L + 0.3T$.

Generic skills will be assessed on the scale: Fail, Pass, Good and Very good (D,C,B and A).

To assess the competence on English, it will be required to have written in English the report on the Case Study, moreover at the beginning of the presentation, the student must do an outline of the work in English as well. Regarding the reasoning competence, it will be assessed from the answers given to the presentation of the Case Study.

BIBLIOGRAPHY

Basic:

- Husson, F.; Lê, S.; Pagès, J. Exploratory multivariate analysis by example using R. Second edition. Boca Raton: CRC Press, Taylor & Francis Group, 2017. ISBN 9781315301860.
- Wickham, H. Ggplot2: elegant graphics for data analysis. 2nd ed. Springer, 2016. ISBN 9783319242774.
- Aluja, T.; Morineau, A. Aprender de los datos: el análisis de componentes principales: una aproximación desde el Data Mining. EUB, 1999. ISBN 8483120224.
- Fox, J. Applied regression analysis and generalized linear models. 3rd ed. SAGE, 2015. ISBN 9781452205663.
- Fox, J.; Weisberg, S. An R companion to applied regression. 3rd ed. SAGE Publications, Inc., 2019. ISBN 9781544336473.

RESOURCES

Hyperlink:

- <https://www.r-project.org/>- <https://www.rstudio.com/>