Course guide
270950 - SIM - Statistical Inference and Modelling

Unit in charge: Barcelona School of Informatics
Teaching unit: 715 - EIO - Department of Statistics and Operations Research.
Degree: MASTER'S DEGREE IN DATA SCIENCE (Syllabus 2021). (Compulsory subject).
Academic year: 2022 ECTS Credits: 6.0 Languages: English

LECTURER
Coordinating lecturer: LIDIA MONTERO MERCADÉ
Others:
Primer quadrimestre: JOSEP FRANQUET FÀBREGAS - 12
LIDIA MONTERO MERCADÉ - 11, 12

PRIOR SKILLS
Students must have sufficient knowledge of algebra and mathematical analysis to assimilate concepts related to set algebra, numerical series, functions of real variables of one or more dimensions, derivation, and integration. Students must have taken a course in probability and statistics

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
CE10. Identify machine learning and statistical modeling methods to use and apply them rigorously in order to solve a specific data science problem
CE6. Design the Data Science process and apply scientific methodologies to obtain conclusions about populations and make decisions accordingly, from both structured and unstructured data and potentially stored in heterogeneous formats.

Generical:
CG1. Identify and apply the most appropriate data management methods and processes to manage the data life cycle, considering both structured and unstructured data
CG2. Identify and apply methods of data analysis, knowledge extraction and visualization for data collected in disparate formats

Transversal:
CT4. INFORMATION LITERACY: Capacity for managing the acquisition, the structuring, analysis and visualization of data and information in the field of specialisation, and for critically assessing the results of this management.
CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

Basic:
CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.
CB9. Possession of the learning skills that enable the students to continue studying in a way that will be mainly self-directed or autonomous.
TEACHING METHODOLOGY

The learning of the subject consists of three different phases:
1. Acquisition of specific knowledge through the study of the bibliography and the material provided by the teachers.
2. The acquisition of skills in specific techniques of data analysis, selection of the statistical modeling process and validation of the model and
3. Integration of knowledge, skills and competences (specific and transversal) through the resolution of real case studies.

In the Theory classes the fundamentals of the methodologies and techniques of the subject are exposed. Laboratory classes are used to learn the use of specific techniques for solving problems, using the appropriate computer tools, in this sense students must first repeat a problem solved by teachers and then solve one similar to the first. While the Case Studies, solved in groups and in hours of self-learning, serve to put into practice the knowledge, skills and competencies in solving real cases.

LEARNING OBJECTIVES OF THE SUBJECT

1. Know how to perform inference processes based on data and in a traditional parametric way for decision making.
2. Know how to make a report on data quality and pre-processed
3. Determination of significant characteristics aimed at numerical and categorical targets in groups of individuals
4. Estimation of parameters and interpretation of linear models of normal response
5. Validation of normal response models. Identification of unusual and influential data. Residual analysis
6. Inference of hypotheses on single and multiple parameters in normal response models
7. Estimation of parameters and interpretation of linear models of binary response
8. Validation of binary response models. Identification of unusual and influential data. Residual types
9. Inference of hypotheses on single and multiple parameters in binary response models
10. Estimation of parameters and interpretation of linear models of nominal and ordinal polytomous response
11. Validation of nominal and ordinal polytomous response models. Identification of unusual and influential data.
12. Inference of hypotheses on simple and multiple parameters in nominal and ordinal polytomous response models
13. Estimation of parameters and interpretation of linear models by counting
15. Inference of hypotheses on simple and multiple parameters in counting models
16. Know how to design factorial and fractional factorial experiments

STUDY LOAD

<table>
<thead>
<tr>
<th>Type</th>
<th>Hours</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hours large group</td>
<td>27,0</td>
<td>18.00</td>
</tr>
<tr>
<td>Self study</td>
<td>96,0</td>
<td>64.00</td>
</tr>
<tr>
<td>Hours small group</td>
<td>27,0</td>
<td>18.00</td>
</tr>
</tbody>
</table>

Total learning time: 150 h

CONTENTS

<table>
<thead>
<tr>
<th>Classical vs Fisherian inference</th>
</tr>
</thead>
</table>

**Description:**
<table>
<thead>
<tr>
<th>Data Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description:</strong></td>
</tr>
<tr>
<td>Univariate and multivariate outliers.</td>
</tr>
<tr>
<td>Missing data. Imputation procedures: deterministic, stochastic.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Normal linear models</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description:</strong></td>
</tr>
<tr>
<td>Neural network estimation of linear regression models.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Generalized linear models</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description:</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Design of Experiments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description:</strong></td>
</tr>
<tr>
<td>Factorial and fractional factorial experimental designs. Modern data analysis techniques for experimental design</td>
</tr>
</tbody>
</table>
### Classical vs Fisherian Inference

**Description:**
Know how to differentiate the conditions of applicability of the different methods of inference and know how to choose the most appropriate to the process of Data Science in hand.

Perform inference processes to draw conclusions about populations. Use p-values, confidence intervals, and permutation tests for decision-making and interpretation of analyzes in a recurring or one-time Data Science problem.

**Specific objectives:**
1

**Related competencies:**
- CE6. Design the Data Science process and apply scientific methodologies to obtain conclusions about populations and make decisions accordingly, from both structured and unstructured data and potentially stored in heterogeneous formats.
- CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.
- CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.
- CB9. Possession of the learning skills that enable the students to continue studying in a way that will be mainly self-directed or autonomous.

**Full-or-part-time:** 18h
- Theory classes: 4h
- Laboratory classes: 2h
- Self study: 12h

### Data quality

**Description:**
Problems in the quality of the data: It is a question of seeing in the Case Study the problems that present or can present the data: Inconsistencies, redundancy. Missing data. Outliers. How to make a Data Quality Report. What is the standardization of data.

**Specific objectives:**
2

**Related competencies:**
- CG2. Identify and apply methods of data analysis, knowledge extraction and visualization for data collected in disparate formats
- CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.
- CT4. INFORMATION LITERACY: Capacity for managing the acquisition, the structuring, analysis and visualization of data and information in the field of specialisation, and for critically assessing the results of this management.
- CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.

**Full-or-part-time:** 7h
- Theory classes: 2h
- Laboratory classes: 2h
- Self study: 3h
### Profiling and feature selection

**Description:**
Application of statistical inference to determine the relationships between variables present in a DB and a response variable (numerical or categorical)

**Specific objectives:**
3

**Related competencies:**
CG2. Identify and apply methods of data analysis, knowledge extraction and visualization for data collected in disparate formats
CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

CT4. INFORMATION LITERACY: Capacity for managing the acquisition, the structuring, analysis and visualization of data and information in the field of specialisation, and for critically assessing the results of this management.

**Full-or-part-time:** 5h
Theory classes: 2h
Laboratory classes: 2h
Self study: 1h

### Estimation of parameters and interpretation of linear models of normal response

**Description:**

**Specific objectives:**
4

**Related competencies:**
CG2. Identify and apply methods of data analysis, knowledge extraction and visualization for data collected in disparate formats
CG1. Identify and apply the most appropriate data management methods and processes to manage the data life cycle, considering both structured and unstructured data
CE10. Identify machine learning and statistical modeling methods to use and apply them rigorously in order to solve a specific data science problem
CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

CT4. INFORMATION LITERACY: Capacity for managing the acquisition, the structuring, analysis and visualization of data and information in the field of specialisation, and for critically assessing the results of this management.
CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.

**Full-or-part-time:** 8h
Theory classes: 2h
Laboratory classes: 2h
Self study: 4h
Validation of normal response models. Identification of unusual and influential data. Waste analysis

Description:
Elements involved in the validation of regression modeling. Influential and / or atypical values

Specific objectives:
5

Related competencies:
CG2. Identify and apply methods of data analysis, knowledge extraction and visualization for data collected in disparate formats
CG1. Identify and apply the most appropriate data management methods and processes to manage the data life cycle, considering both structured and unstructured data
CE10. Identify machine learning and statistical modeling methods to use and apply them rigorously in order to solve a specific data science problem
CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

CT4. INFORMATION LITERACY: Capacity for managing the acquisition, the structuring, analysis and visualization of data and information in the field of specialisation, and for critically assessing the results of this management.
CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.

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

Inference of hypotheses on single and multiple parameters in normal response models

Description:
Inference on parameter estimators in linear models of normal response. Confidence intervals, confidence regions. Contrasts of simple, multiple hypotheses, linear combinations. Inference about confidence interval predictions and calculations.

Specific objectives:
6

Related competencies:
CG2. Identify and apply methods of data analysis, knowledge extraction and visualization for data collected in disparate formats
CE6. Design the Data Science process and apply scientific methodologies to obtain conclusions about populations and make decisions accordingly, from both structured and unstructured data and potentially stored in heterogeneous formats.
CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.

Full-or-part-time: 6h
Theory classes: 2h
Laboratory classes: 2h
Self study: 2h
### Estimation of parameters and interpretation of linear models of binary response

**Description:**

**Specific objectives:**
7

**Related competencies:**
- CE6. Design the Data Science process and apply scientific methodologies to obtain conclusions about populations and make decisions accordingly, from both structured and unstructured data and potentially stored in heterogeneous formats.
- CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

CB9. Possession of the learning skills that enable the students to continue studying in a way that will be mainly self-directed or autonomous.

**Full-or-part-time:** 6h
- Theory classes: 2h
- Laboratory classes: 2h
- Self study: 2h

### Validation of binary response models. Identification of unusual and influential data. Type of waste

**Specific objectives:**
8

**Related competencies:**
- CG2. Identify and apply methods of data analysis, knowledge extraction and visualization for data collected in disparate formats
- CG1. Identify and apply the most appropriate data management methods and processes to manage the data life cycle, considering both structured and unstructured data
- CE6. Design the Data Science process and apply scientific methodologies to obtain conclusions about populations and make decisions accordingly, from both structured and unstructured data and potentially stored in heterogeneous formats.
- CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

CT4. INFORMATION LITERACY: Capacity for managing the acquisition, the structuring, analysis and visualization of data and information in the field of specialisation, and for critically assessing the results of this management.

CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.

**Full-or-part-time:** 6h
- Theory classes: 2h
- Laboratory classes: 2h
- Self study: 2h
### Inference of hypotheses on single and multiple parameters in binary response models

**Description:**
Inference on parameter estimators in linear models of a binary response. Confidence intervals. Contrasts of simple, multiple hypotheses, linear combinations. Inference about confidence interval predictions and calculations.

**Specific objectives:**
9

**Related competencies:**
- CG1. Identify and apply the most appropriate data management methods and processes to manage the data life cycle, considering both structured and unstructured data
- CE6. Design the Data Science process and apply scientific methodologies to obtain conclusions about populations and make decisions accordingly, from both structured and unstructured data and potentially stored in heterogeneous formats.
- CB9. Possession of the learning skills that enable the students to continue studying in a way that will be mainly self-directed or autonomous.

**Full-or-part-time:** 3h
- Theory classes: 1h
- Laboratory classes: 1h
- Self study: 1h

### Estimation of parameters and interpretation of linear models of nominal and ordinal polytomous response

**Description:**

**Specific objectives:**
10

**Related competencies:**
- CG1. Identify and apply the most appropriate data management methods and processes to manage the data life cycle, considering both structured and unstructured data
- CE10. Identify machine learning and statistical modeling methods to use and apply them rigorously in order to solve a specific data science problem
- CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

- CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.

**Full-or-part-time:** 4h
- Theory classes: 1h
- Laboratory classes: 1h
- Self study: 2h
Validación de los modelos de respuesta polítomica nominal y ordinal. Identificación de datos inusuales e influyentes

Description:
Deviance and Pearson residuals. Student residuals. Unusual and influential data indicators, by extending the indicators used in normal regression.

Specific objectives:
11

Related competencies:
CG2. Identify and apply methods of data analysis, knowledge extraction and visualization for data collected in disparate formats
CE10. Identify machine learning and statistical modeling methods to use and apply them rigorously in order to solve a specific data science problem
CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.

Full-or-part-time: 2h 30m
Theory classes: 0h 30m
Laboratory classes: 1h
Self study: 1h

Inference of hypotheses on simple and multiple parameters in nominal and ordinal polytomous response models

Description:
Inference on parameter estimators in linear polytomous response models. Confidence intervals. Simple, multiple hypothesis tests, linear combinations. Inference about predictions and confidence interval calculations.

Specific objectives:
12

Related competencies:
CG2. Identify and apply methods of data analysis, knowledge extraction and visualization for data collected in disparate formats
CG1. Identify and apply the most appropriate data management methods and processes to manage the data life cycle, considering both structured and unstructured data
CE6. Design the Data Science process and apply scientific methodologies to obtain conclusions about populations and make decisions accordingly, from both structured and unstructured data and potentially stored in heterogeneous formats.
CE10. Identify machine learning and statistical modeling methods to use and apply them rigorously in order to solve a specific data science problem
CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

Full-or-part-time: 3h
Theory classes: 1h
Laboratory classes: 1h
Self study: 1h
Estimation of parameters and interpretation of linear models by counting

Description:

Specific objectives:
13

Related competencies:
CG2. Identify and apply methods of data analysis, knowledge extraction and visualization for data collected in disparate formats
CG1. Identify and apply the most appropriate data management methods and processes to manage the data life cycle, considering both structured and unstructured data
CE10. Identify machine learning and statistical modeling methods to use and apply them rigorously in order to solve a specific data science problem
CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

CB9. Possession of the learning skills that enable the students to continue studying in a way that will be mainly self-directed or autonomous.

Full-or-part-time: 2h 30m
Theory classes: 0h 30m
Laboratory classes: 1h
Self study: 1h

Validation of counting models. Identification of unusual and influential data. Type of waste. Overdispersion diagnosis. Parametric probabilistic models

Description:
Unusual and influential data indicators. Overdispersion checking. How to overcome overdispersion.

Specific objectives:
14

Related competencies:
CG1. Identify and apply the most appropriate data management methods and processes to manage the data life cycle, considering both structured and unstructured data
CE6. Design the Data Science process and apply scientific methodologies to obtain conclusions about populations and make decisions accordingly, from both structured and unstructured data and potentially stored in heterogeneous formats.
CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.

Full-or-part-time: 2h 30m
Theory classes: 0h 30m
Laboratory classes: 1h
Self study: 1h
Inference of hypotheses on simple and multiple parameters in counting models

Description:
Inference on parameter estimators in linear models by counts. Confidence intervals. Contrasts of simple, multiple hypotheses, linear combinations. Inference on predictions and calculations of confidence intervals.

Specific objectives:
15

Related competencies:
CE6. Design the Data Science process and apply scientific methodologies to obtain conclusions about populations and make decisions accordingly, from both structured and unstructured data and potentially stored in heterogeneous formats.
CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

Full-or-part-time: 2h 30m
Theory classes: 0h 30m
Laboratory classes: 1h
Self study: 1h

Theory and practice of factorial and fractional factorial experiment design

Specific objectives:
16

Related competencies:
CG1. Identify and apply the most appropriate data management methods and processes to manage the data life cycle, considering both structured and unstructured data
CE6. Design the Data Science process and apply scientific methodologies to obtain conclusions about populations and make decisions accordingly, from both structured and unstructured data and potentially stored in heterogeneous formats.
CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.
CB9. Possession of the learning skills that enable the students to continue studying in a way that will be mainly self-directed or autonomous.

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

Specific objectives:
1, 2, 3, 4, 5, 6

Related competencies:
CG2. Identify and apply methods of data analysis, knowledge extraction and visualization for data collected in disparate formats
CG1. Identify and apply the most appropriate data management methods and processes to manage the data life cycle, considering both structured and unstructured data
CE6. Design the Data Science process and apply scientific methodologies to obtain conclusions about populations and make decisions accordingly, from both structured and unstructured data and potentially stored in heterogeneous formats.
CE10. Identify machine learning and statistical modeling methods to use and apply them rigorously in order to solve a specific data science problem
CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

CT4. INFORMATION LITERACY: Capacity for managing the acquisition, the structuring, analysis and visualization of data and information in the field of specialisation, and for critically assessing the results of this management.
CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.
CB9. Possession of the learning skills that enable the students to continue studying in a way that will be mainly self-directed or autonomous.

Full-or-part-time: 8h
Guided activities: 2h
Self study: 6h

Final Exam

Specific objectives:
7, 8, 9, 10, 11, 12, 13, 14, 15, 16

Related competencies:
CG2. Identify and apply methods of data analysis, knowledge extraction and visualization for data collected in disparate formats
CG1. Identify and apply the most appropriate data management methods and processes to manage the data life cycle, considering both structured and unstructured data
CE6. Design the Data Science process and apply scientific methodologies to obtain conclusions about populations and make decisions accordingly, from both structured and unstructured data and potentially stored in heterogeneous formats.
CE10. Identify machine learning and statistical modeling methods to use and apply them rigorously in order to solve a specific data science problem
CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

CT4. INFORMATION LITERACY: Capacity for managing the acquisition, the structuring, analysis and visualization of data and information in the field of specialisation, and for critically assessing the results of this management.
CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.
CB9. Possession of the learning skills that enable the students to continue studying in a way that will be mainly self-directed or autonomous.

Full-or-part-time: 10h
Guided activities: 2h
Self study: 8h
Linear Model Assignment

Specific objectives:
2, 3, 4, 5, 6

Related competencies:
CG2. Identify and apply methods of data analysis, knowledge extraction and visualization for data collected in disparate formats
CG1. Identify and apply the most appropriate data management methods and processes to manage the data life cycle, considering both structured and unstructured data
CE6. Design the Data Science process and apply scientific methodologies to obtain conclusions about populations and make decisions accordingly, from both structured and unstructured data and potentially stored in heterogeneous formats.
CE10. Identify machine learning and statistical modeling methods to use and apply them rigorously in order to solve a specific data science problem
CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

CT4. INFORMATION LITERACY: Capacity for managing the acquisition, the structuring, analysis and visualization of data and information in the field of specialisation, and for critically assessing the results of this management.
CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.

Full-or-part-time: 20h
Self study: 20h

Generalized Linear Model Assignment

Specific objectives:
7, 8, 9, 10, 11, 12, 13, 14, 15

Related competencies:
CG2. Identify and apply methods of data analysis, knowledge extraction and visualization for data collected in disparate formats
CG1. Identify and apply the most appropriate data management methods and processes to manage the data life cycle, considering both structured and unstructured data
CE6. Design the Data Science process and apply scientific methodologies to obtain conclusions about populations and make decisions accordingly, from both structured and unstructured data and potentially stored in heterogeneous formats.
CE10. Identify machine learning and statistical modeling methods to use and apply them rigorously in order to solve a specific data science problem
CT5. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

CT4. INFORMATION LITERACY: Capacity for managing the acquisition, the structuring, analysis and visualization of data and information in the field of specialisation, and for critically assessing the results of this management.
CB6. Ability to apply the acquired knowledge and capacity for solving problems in new or unknown environments within broader (or multidisciplinary) contexts related to their area of study.
CB9. Possession of the learning skills that enable the students to continue studying in a way that will be mainly self-directed or autonomous.

Full-or-part-time: 20h
Self study: 20h
GRADING SYSTEM

The evaluation of the subject integrates the three phases of learning described: knowledge, skills and competences.

Knowledge is assessed by two exams conducted in the middle (T1, weight 1/3) and during the week of final exams of the course (T2, weight 2/3). In case of failing the partial exam, the student may repeat it as an extension of the final exam (note T).

The skills will be evaluated from the delivery of 2 practices, as well as the transversal competences. Each of the blocks 1, 2 and 3 for the first practice (P1) and 4 and 5 for the second (P2) will involve a practice that the student must do individually or in groups of 2. The average of the marks gives the mark P.

The Final Grade (NF) is calculated:

Partial Exam (T1, 1/3) and Final Exam (T2, 2/3).
Practice 1 (P1) and Practice 2 (P2)
P: Practice Note P = (P1 + P2) / 2.
T: Theory Note = Max (T2, (T1 + 2T2) / 3).
NF: Final Grade = 0.6T + 0.4P.

BIBLIOGRAPHY

Basic:

Complementary:

RESOURCES

Hyperlink:
- https://www.rstudio.com/
- https://www.r-project.org/