

## Course guide

# 480022 - EMSD - Fundamentals of Applied Statistics and Sustainability and Development Measurement

Last modified: 12/06/2023

<b>Unit in charge:</b>	Barcelona School of Civil Engineering	
<b>Teaching unit:</b>	715 - EIO - Department of Statistics and Operations Research.	
<b>Degree:</b>	MASTER'S DEGREE IN SUSTAINABILITY SCIENCE AND TECHNOLOGY (Syllabus 2013). (Compulsory subject).	
<b>Academic year:</b> 2023	<b>ECTS Credits:</b> 5.0	<b>Languages:</b> Catalan, Spanish

## LECTURER

---

<b>Coordinating lecturer:</b>	Gibert Oliveras, Carina
<b>Others:</b>	Primer quadrimestre: XAVIER ANGERRI TORREDEFLOT - 10Q1 CARINA GIBERT OLIVERAS - 10Q1

## PRIOR SKILLS

---

Basics on computer science, probability and univariate statistics (bachelor course).

Recommended introductory texts:

- Bass, R.F. (2013), Undergraduate probability. 62 pp. Available at: [bass.math.uconn.edu/elemprob.pdf](http://bass.math.uconn.edu/elemprob.pdf) [Last accessed March 2017]
- Grinstead, C. M., Snell, J.L. (1997), Introduction to Probability (2 Edition), American Mathematical Society. 510 pp. ISBN 978-0821807491. Chapters 1- 10. Available at: [www.dartmouth.edu/chance/teaching\\_aids/books\\_articles/probability\\_book/book](http://www.dartmouth.edu/chance/teaching_aids/books_articles/probability_book/book) [Last accessed March 2017]
- Venables, W.N., Smith, D.M. & R Core Team (2017) "An Introduction to R", v3.3.3. Chapters 1-8, 12-13. Available at: [cran.r-project.org/doc/manuals/r-release/R-intro.pdf](http://cran.r-project.org/doc/manuals/r-release/R-intro.pdf) [Last accessed March 2017]

Introductory and Open Access R and RStudio courses:

- [www.udemy.com/r-basics/#instructor-5678748](http://www.udemy.com/r-basics/#instructor-5678748)
- [www.datacamp.com/courses/free-introduction-to-r](http://www.datacamp.com/courses/free-introduction-to-r)

Playlist on youtube:

- [www.youtube.com/playlist?list=PLqijB62qLSh0YB3sXP4r2QpnIZJEzqdx3](http://www.youtube.com/playlist?list=PLqijB62qLSh0YB3sXP4r2QpnIZJEzqdx3)

## REQUIREMENTS

---

None

## DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

---

### Specific:

CE04. The ability to apply, critically and effectively, conceptual frameworks, data collection and processing techniques, applied statistics, mathematical modelling, systems analysis, geographic information systems, information and communication technologies and industrial ecology to meeting the challenges of sustainability and sustainable development.

CE06. The capacity to apply the methods and tools used in the identification, information management, planning, management, execution and evaluation of programmes and projects in the fields of sustainability and environmental management to specific problems in a collaborative manner.

CE12. Design, develop, apply and evaluate conceptual frameworks, theories, methodologies and techniques typical of ICT in contexts of promoting sustainable development and sustainability.

CE03. The ability to critically analyse theories and perspectives on the traits and properties of the geosphere and biosphere that facilitate and frame the development of socio-environmental systems, as well as the main challenges posed by climate change.

CE13. The ability to apply, critically analyse results and assess valorisation theories, approaches and methods in the fields of food and rural development and agricultural, water, energy, building construction, transport and spatial engineering.

### Generical:

CG03. The ability to analyze, evaluate and synthesize, critically, new and complex ideas and promote, within academic and professional, scientific, technological, social or cultural knowledge society contexts.

CG02. Develop and / or implement innovative ideas in a research context by identifying and formulating hypotheses and by submitting to prove objectivity, consistency and viability.

### Transversal:

1. EFFECTIVE USE OF INFORMATION RESOURCES: Managing the acquisition, structuring, analysis and display of data and information in the chosen area of specialisation and critically assessing the results obtained.

CT3. TEAMWORK: Being able to work in an interdisciplinary team, whether as a member or as a leader, with the aim of contributing to projects pragmatically and responsibly and making commitments in view of the resources that are available.

### Basic:

CB9. That students can communicate their conclusions-and the knowledge and rationale underpinning these, to specialist and non-specialist audiences clearly and unambiguously.

CB10. That students have the learning skills to allow them to continue studying in a way that will have to be largely autodirigido or autonomous.

## TEACHING METHODOLOGY

---

During the development of the subject the following teaching methodologies will be used:

Master class or conference (EXP): exhibition of knowledge by lecturers through master classes or by external people through guest lectures.

Problem solving and case study (PR): collaborative resolution of exercises, conducting group discussions and group dynamics, with the lecturer and other students in the classroom; presentation in the classroom of an activity to be done individually or in small groups.

Practical work in laboratory or workshop (TP): designing, measuring, verifying, etc. : and presentation of the results in oral or written form, individually or in reduced groups.

Directed theoretical-practical work (TD): activity or exercise either theoretical or practical, developed in the classroom, individually or in small groups, with the lecturer supervision.

Project, activity or work of limited scope (PR): learning by doing, individually or in group, of a work of reduced complexity or extension, applying knowledge and presenting results.

Extensive project or work (PA): learning based on designing, planning and working groups of a project or work of wide complexity or extension, applying and widen knowledge and writing a report which reflects the approach of the work and the results and conclusions.

Evaluation Activities (AV).

Along the course, the following training activities will be used:

Face-to-face

Theoretical classes and conferences (CTC): knowing, understanding and synthesizing the knowledge presented by the lecturers through master classes or by lecturers.

Practical classes (CP): participate in the collaborative problem solving, as well as in group discussions and dynamics, with the teacher and other students in the classroom.

Laboratory or workshop practices (L / T): understanding the operation of equipments, specifications and documentation; design-making, measurements, verifications, etc. ; and presenting the written results or orally either individually or in small groups.

Oral Presentations (PS): present in the classroom an activity carried out individually or in small groups.

Assistance of theoretical and practical work (TD): carry out in the classroom an activity or exercise either theoretical or practical, individually or in small groups, with the advice of the lecturer.

Non-presential

Carrying out a project or work of wide scope (AP): designing, planning and carrying out individually or in groups a project or work of wide complexity or extension, applying and expanding knowledge and writing a report where the approach is explained and the results and conclusions.

## LEARNING OBJECTIVES OF THE SUBJECT

Statistics is the oldest science of extracting information from data and has been used from ancient times to understand reality and to build models that allow

a better understanding and useful predictions for decision-making. In this course, statistical principles will be introduced from the perspective of the support they provide to the analysis of problems related to sustainability. The course follows a basically applied approach and will focus on real problem solving with basic statistical methods.

The aim of the course is to get the students familiar with the basic statistical tools for dealing with socio-environmental databases, being able to:

- 1) select the relevant data to provide answers to a specific question
- 2) prepare data properly to be analyzed, with a strong emphasis on data cleaning and preprocessing
- 3) identify the most appropriate statistical modeling methods for a given problem, according to the structure of the available data, the objectives of the study and the subsequent uses of the model results
- 4) construct correct statistical models from the data
- 5) validate the models obtained and make a critical interpretation of the results from a both technical point of view and contextualizing the results in the framework problem

The course pretends that students:

get familiar with statistical methodology as a basic scheme for extracting relevant information from complex phenomena (such as environmental or related to sustainability) from data

become able to apply the knowledge obtained in class in the analysis of a real data set in an integrated way (taking advantage of open data sources) and also entering into the mechanisms of teamwork

This course established methodological bases of wide utility in the observation of reality and informed decision-making and also prepares to address more complex models that are presented in the course socio-environmental data science

to finish the subject, the student:

Develops and applies concepts and theories of mathematics and statistics applied with originality to solving challenges of sustainability and development, identifying and formulating hypotheses or innovative ideas and subjecting them to a test of objectivity, coherence and viability.

Efficiently applies techniques and instruments specific to mathematics and statistics applied to sustainability and development challenges using open development computer tools.

Integrates and critically analyzes the result of using mathematical and statistical models in the definition of solutions and strategies of sustainability and development.

## STUDY LOAD

Type	Hours	Percentage
Self study	80,0	64.00
Hours small group	9,0	7.20
Hours medium group	12,0	9.60
Hours large group	24,0	19.20

**Total learning time:** 125 h

## CONTENTS

### (ENG) Sustainability indicators and open data

**Description:**

(ENG) Indicators and sustainability indices. International information and open data

**Specific objectives:**

Identify mechanisms to assess the sustainability levels of a system or ecosystem and how to take advantage from open data

**Related activities:**

Open data for sustainability R Practice.

Teamwork

**Related competencies :**

CE12. Design, develop, apply and evaluate conceptual frameworks, theories, methodologies and techniques typical of ICT in contexts of promoting sustainable development and sustainability.

CT4. EFFECTIVE USE OF INFORMATION RESOURCES: Managing the acquisition, structuring, analysis and display of data and information in the chosen area of specialisation and critically assessing the results obtained.

CB10. That students have the learning skills to allow them to continue studying in a way that will have to be largely autodirigido or autonomous.

**Full-or-part-time:** 5h

Theory classes: 2h

Laboratory classes: 1h

Self study : 2h

### Introduction to data processing and descriptive statistics

**Description:**

(ENG) Data types

Descriptive Statistics: Univariate Statistics, Bivariate Statistics.

Fundamentals of pre-processing and data cleaning.

Automatic reporting

**Specific objectives:**

Learn how to build a relevant dataset for a given problem.

Prepare data properly to ensure correct analysis.

Ensure that data is treated according to its nature and typology correctly.

Build scripts for descriptive statistics and automatic reporting.

Synthesize the main characteristics of a dataset from descriptive analysis

**Related activities:**

Teamwork

Session in R

**Related competencies :**

CE04. The ability to apply, critically and effectively, conceptual frameworks, data collection and processing techniques, applied statistics, mathematical modelling, systems analysis, geographic information systems, information and communication technologies and industrial ecology to meeting the challenges of sustainability and sustainable development.

CE12. Design, develop, apply and evaluate conceptual frameworks, theories, methodologies and techniques typical of ICT in contexts of promoting sustainable development and sustainability.

CT3. TEAMWORK: Being able to work in an interdisciplinary team, whether as a member or as a leader, with the aim of contributing to projects pragmatically and responsibly and making commitments in view of the resources that are available.

**Full-or-part-time:** 12h

Theory classes: 12h

## Statistical inference

### Description:

Fundamentals of Probability  
Statistical sampling, Concept of Statistic  
Estimation and Confidence Intervals  
Statistical reasoning and hypothesis testing  
Fisher Permutation Tests

### Specific objectives:

Understand the concept of random sample, statistics, uncertainty management, reasoning in hypothesis testing, interpretation of p-value.  
Know the conditions of applicability of the different statistics and tests  
Know the principles of Fisher permutation tests and the situations in which they add value with respect to classical inference  
Know how to choose the appropriate statistics or test in a specific real situation, apply it appropriately and correctly interpret the results in real problems  
Know the instructions of R to select samples from a database, calculate hypothesis tests and confidence intervals in different conditions and apply them to team practice

### Related activities:

Sessions in R  
Problem Solving  
Teamwork

### Related competencies :

CE04. The ability to apply, critically and effectively, conceptual frameworks, data collection and processing techniques, applied statistics, mathematical modelling, systems analysis, geographic information systems, information and communication technologies and industrial ecology to meeting the challenges of sustainability and sustainable development.  
CE12. Design, develop, apply and evaluate conceptual frameworks, theories, methodologies and techniques typical of ICT in contexts of promoting sustainable development and sustainability.  
CT3. TEAMWORK: Being able to work in an interdisciplinary team, whether as a member or as a leader, with the aim of contributing to projects pragmatically and responsibly and making commitments in view of the resources that are available.  
CB9. That students can communicate their conclusions-and the knowledge and rationale underpinning these, to specialist and non-specialist audiences clearly and unambiguously.

### Full-or-part-time: 18h

Theory classes: 18h

## (ENG) x

### Description:

x  
Basic methods of statistical modeling

(CAT) Simple linear regression.

Model Estimation and validation .

Multiple linear regression.

Generalized linear regression, logistic regression.

Analysis of variance. Covariance Analysis. Selection of models.

**Specific objectives:**

Identify in which situations the different models presented are suitable  
 Know the principles of model fitting and validation of the different models.  
 Know and apply the principle of parsimony in modeling.  
 Correctly interpret modeling results and model quality indicators.  
 Apply the different models to real cases and make a critical assessment of the results

**Related activities:**

Problem solving  
 R session  
 TeamWork

**Related competencies :**

CG02. Develop and / or implement innovative ideas in a research context by identifying and formulating hypotheses and by submitting to prove objectivity, consistency and viability.  
 CG03. The ability to analyze, evaluate and synthesize, critically, new and complex ideas and promote, within academic and professional, scientific, technological, social or cultural knowledge society contexts.  
 CE04. The ability to apply, critically and effectively, conceptual frameworks, data collection and processing techniques, applied statistics, mathematical modelling, systems analysis, geographic information systems, information and communication technologies and industrial ecology to meeting the challenges of sustainability and sustainable development.  
 CE12. Design, develop, apply and evaluate conceptual frameworks, theories, methodologies and techniques typical of ICT in contexts of promoting sustainable development and sustainability.  
 CE06. The capacity to apply the methods and tools used in the identification, information management, planning, management, execution and evaluation of programmes and projects in the fields of sustainability and environmental management to specific problems in a collaborative manner.  
 CT3. TEAMWORK: Being able to work in an interdisciplinary team, whether as a member or as a leader, with the aim of contributing to projects pragmatically and responsibly and making commitments in view of the resources that are available.  
 CT4. EFFECTIVE USE OF INFORMATION RESOURCES: Managing the acquisition, structuring, analysis and display of data and information in the chosen area of specialisation and critically assessing the results obtained.  
 CB10. That students have the learning skills to allow them to continue studying in a way that will have to be largely autodirigido or autonomous.  
 CB9. That students can communicate their conclusions-and the knowledge and rationale underpinning these, to specialist and non-specialist audiences clearly and unambiguously.

**Full-or-part-time: 9h**

Theory classes: 9h

## ACTIVITIES

### Working teams creation, planning, work distribution.

**Description:**

Students form working groups to develop the team practice, develop the planning of the project and distribution of work

**Specific objectives:**

Practice team working in diversity, planning and projecting a medium-term team work intended to practice the techniques seen along the course on real data

**Material:**

Teacher's instructions on the characteristics of the group

Teamwork materials on the course page

Subject calendar

Support for Gantt diagrams design

**Delivery:**

Deliverable with the composition of the working team according to the statement of the practice, working plan and grid of tasks and responsibilities distribution

**Full-or-part-time:** 2h 30m

Laboratory classes: 1h 30m

Self study: 1h

### (ENG) 2. Dataset selection for teamwork.

**Description:**

Each working group will use open data sources and other data sources of interest to select a dataset to will develop the teamwork.

**Specific objectives:**

Reliable use of information resources.

Set a topic of application of interest for the group that supports active learning throughout the course.

**Material:**

Opendata links list provided by the teachers

Specifications of dataset (to guarantee the successfull execution of the entire practical work)

**Delivery:**

Deliverable according to the practice statement.

**Full-or-part-time:** 3h

Laboratory classes: 1h

Self study: 2h



### (ENG) 3. Teamwork

**Description:**

Team project for the entire course

**Specific objectives:**

Students apply the methods seen in class to real data, critically analyze the results, manage the operation of the team, interpret the results, write a report, synthesize the results in an oral presentation and acquire skills to discuss the topic. of the practice and the decisions taken in the elaboration of the solution

**Material:**

Statement of teamWork

Slides Course

Bibliography

Scripts in R from the course page

**Delivery:**

At the end of the course. As indicated in the statement (additional intermediate delivery)

**Full-or-part-time:** 22h

Theory classes: 11h

Self study: 11h

### (ENG) 4 Oral Presentation

**Description:**

Oral presentation by groups on the intermediate and final results of the practice

**Specific objectives:**

Assess the ability to synthesize, oral and non-verbal communication critical thinking and technical skills

**Material:**

the team practical work

**Delivery:**

According to assignment document

**Full-or-part-time:** 10h

Practical classes: 4h

Self study: 6h

#### (ENG) 5 Discussion

**Description:**

Open discussions following the collaborative quiz solving or the oral presentations o practical teamworks

**Specific objectives:**

To evaluate the knowledge integration skills and the topic understanding of students, as well as their critical thinking achievements

**Material:**

Two modalities:

Solved reference quiz

Oral presentations of practical works

**Delivery:**

Not deliverable. It is an activity where the teacher observes

**Full-or-part-time:** 2h

Theory classes: 2h

#### (ENG) 6 Quiz

**Description:**

Short questionnaires to be solved in classs after a session (no more than 3)

**Specific objectives:**

To evaluate the level of knowledge acquired by the student.

To generate feedback to the student about his evolution

To fix the most relevant knowledge pills

To give immediate feedback to students about concept errors, thus contributing to long-term learning

**Material:**

The topic corresponding to that class

**Delivery:**

Solved questionnaire. These activities are followed by collaborative evaluation practices and discussion

**Full-or-part-time:** 1h 30m

Theory classes: 1h 30m

#### (ENG) Collaborative problem Solving

**Description:**

Collaborative problem solving throughout the class (no more than 3)

**Specific objectives:**

To practice the application of the contents of the course toproblem solving of controlled difficulty. Easier to solve than the real problems found in the teampractical work

**Material:**

Problem statement

Slides and bibliography

The Rpackage and datasets files

**Delivery:**

Problem solved

**Full-or-part-time:** 1h 30m

Theory classes: 1h 30m

### (ENG) 8 Practice with the R package

**Description:**

Practice the use of R to apply the methods presented in theoretical classes to the dataset of the team practical work

**Specific objectives:**

Learning by doing  
Team work

**Material:**

Slides, references  
R Scripts from the course website  
Data from the team practical work

**Delivery:**

According to practical work instructions

**Full-or-part-time:** 17h

Laboratory classes: 6h

Self study: 11h

### (ENG) 9 Active learning evaluation activities

**Description:**

Self-evaluation activities for the problems or quiz while the teacher solves them publicly in class  
Eventually can be followed by pairs evaluation techniques

**Specific objectives:**

To provide feedback to the student about his learning level allowing him to fix the important knowledge pills, as well as to recognize the concept errors immediately after the problem solving, to correct them and to fix the correct learning

**Material:**

the problem or questionnaire solved

**Delivery:**

the problem or questionnaire solved

**Full-or-part-time:** 1h

Practical classes: 1h

### 110 Theoretical lessons, case study and examples

**Description:**

Conceptual development of the course and presentation of case studies and didactical examples

**Specific objectives:**

Course topics disclosure

**Material:**

Slides, References, Videos and other material indicated by the teacher

**Delivery:**

None

**Full-or-part-time:** 15h 30m

Theory classes: 15h 30m

## GRADING SYSTEM

---

- EV1 Written Knowledge Control Test (PE). 30%
- EV2 Oral knowledge control test (PO). 10%
- EV3 Work done throughout the course (TR). 50%
- EV4 Attendance and participation in classes and laboratories (AP). 10%
- EV5 Performance and quality of group work (TG). 10%

## EXAMINATION RULES.

---

The instructions for the development of activities will be specified in the same statement.

## BIBLIOGRAPHY

---

### Basic:

- Gibert, K. Descriptiva i inferència. Barcelona: UPC (distributed by Lecturers), 2009.
- Fox, J. An R and S-Plus companion to applied regression. Thousand Oaks (Calif.) [etc.]: Sage, 2002. ISBN 0761922792.
- Maindonald, J.; Braun, W.J. Data analysis and graphics using R: an example-based approach. 3rd ed. Cambridge: Cambridge University, 2010. ISBN 9780521762939.
- Barnett, V. Sample survey: principles and methods. 3rd ed. London: John Wiley & Sons, 2002. ISBN 0340763981.
- Tan, P.-N.; Steinbach, M.; Karpatne, A.; Kumar, V. Introduction to data mining. 2nd ed. New York: Pearson Education, 2019. ISBN 9780133128901.
- Ross, S.M. Introduction to probability and statistics for engineers and scientists. 5th ed. Oxford: Academic Press, 2014. ISBN 9780123948113.
- Dekking, F.M.; Kraaikamp, C.; Lopuhaä, H.P. A modern introduction to probability and statistics: understanding why and how [on line]. London: Springer-Verlag, 2005 [Consultation: 21/02/2023]. Available on: <https://link.springer.com/book/10.1007/1-84628-168-7>. ISBN 9781846281686.
- Ugarte, M.D.; Militino, A.F.; Arnholt, A.T. Probability and statistics with R. Second Edition. Boca Raton: Taylor and Francis Group, 2016. ISBN 978-1466504394.
- Rao, J.N.K.; Molina, I. Small area estimation [on line]. Second Edition. Hoboken, New Jersey: Wiley, 2015 [Consultation: 03/10/2018]. Available on: <https://onlinelibrary.wiley.com/doi/book/10.1002/9781118735855>. ISBN 9781118735787.
- Kottegoda, N.T.; Rosso, R. Applied statistics for civil and environmental engineers [on line]. 2nd ed. Oxford: Wiley-Blackwell, 2009 [Consultation: 17/10/2018]. Available on: <https://ebookcentral.proquest.com/lib/upcatalunya-ebooks/detail.action?docID=428240>. ISBN 9781405179171.
- Berthouex, P.M.; Brown, L.C. Statistics for environmental engineers. Second Edition. Boca Raton: CRC Press LLC, 2002. ISBN 1566705924.

## RESOURCES

---

### Hyperlink:

- United Nations Development Program. Human Development Report, (annual report).. [hdr.undp.org](http://hdr.undp.org)
- Alkire, S.; Foster, J.E.; Seth, S.; Santos, M.E.; Ballón, P. & Roche, J. M. (2015). Multidimensional poverty measurement and analysis / a book from the Oxford Poverty and Human development Initiative (OPHI). University of Oxford.. [multidimensionalpoverty.org](http://multidimensionalpoverty.org)
- UNICEF (2017) MICS6 Tools, Multiple Indicator Cluster Surveys (MICS).. [mics.unicef.org/tools](http://mics.unicef.org/tools)
- United Nations (2007) Indicators of sustainable development: guidelines and methodologies (Third Edition), ISBN: 978-9211045772. [www.un.org/esa/sustdev/natlinfo/indicators/guidelines.pdf](http://www.un.org/esa/sustdev/natlinfo/indicators/guidelines.pdf)