

Data Visualisation

Code: 104365
ECTS Credits: 6

Degree	Type	Year	Semester
2503758 Data Engineering	OB	3	2

The proposed teaching and assessment methodology that appear in the guide may be subject to changes as a result of the restrictions to face-to-face class attendance imposed by the health authorities.

Contact

Name: Enric Martí Godia
Email: Enric.Marti@uab.cat

Use of Languages

Principal working language: catalan (cat)
Some groups entirely in English: No
Some groups entirely in Catalan: Yes
Some groups entirely in Spanish: No

Teachers

Guillermo Marin Getino
Judit Chamorro Servent

Prerequisites

It is recommended that students have or review knowledge and concepts of:

- Graph analysis and information retrieval.
- Model adjustment and optimization.
- Computational learning.
- Statistical analysis.

Objectives and Contextualisation

The course aims to give students elements to find the best data visualization technique that allows better analysis:

KNOWLEDGE: At the end of the course the student must be able to:

-

Know and understand the basic systems of visualization and exploratory analysis of data.

Know and understand the main visualization models.

Know and understand current libraries for data visualization (Google Charts, JQuery plug-ins, D3.js)

SKILLS: It is intended that students acquire the following skills:

-

Given a set of data, know how to use the most appropriate visualization system and model for data visuali

-

Given a data set, know how to use one of the data display libraries to get the best graphical results.

LEARNING OUTCOMES:

1.

Know and understand the basic systems of visualization and exploratory analysis of data.

2.

Know and understand the main visualization models.

3.

Know and understand current libraries for data visualization (Google Charts, JQuery plug-ins, D3.js)

4. Know how to apply the most optimal data visualization techniques for data analysis based on a specific case.

5. To Work cooperatively.

Competences

- Search, select and manage information and knowledge responsibly.
- Students must be capable of collecting and interpreting relevant data (usually within their area of study) in order to make statements that reflect social, scientific or ethical relevant issues.
- Understand visualisation techniques for big data and be able to select the most adequate for analysis.

Learning Outcomes

1. Identify the usability and interactivity needs of a data-visualisation method and be able to produce a new version of the visualisation that improves on the above aspects.
2. Relate statistical and artificial-intelligence techniques to the visualisations that best express the underlying information.
3. Search, select and manage information and knowledge responsibly.
4. Students must be capable of collecting and interpreting relevant data (usually within their area of study) in order to make statements that reflect social, scientific or ethical relevant issues.

Content

1.

1. Visualization Introduction (2 hours)	1. V
	1. F
	2. A

2. Visual Codification Visual (2 hours)
 1. Data types and graphic types
 2. Visual channels
 3. Choose graphic information according his function
3. Visual perception and design (2 hours)
 1. Visual perception
 2. Design fundamentals
 3. Typography
 4. Colour
4. Visual precision (2 hours)
 1. Common errors in visualization
 2. Honestedat visual
 3. Bones i dolentes pràctiques en estadística
5. Data treatment to visualization (4 hours)
 1. Uncertainty i Error
 2. Dimensionality reduction
 3. Transformations

4. Data massatge
5. Computation and relevant metrics election
6. Advances visualization systems (4 hours)
 1. Multiple variables and multiple dimensions
 2. Nets
 3. Dades 3D
 4. Vector fields
 5. Scientific Visualization
 6. Mapes
7. Interaction. (4 hours)
 1. Animation
 2. Dashboards
8. Metodologies de treball (2 hores)
 1. Storytelling and data
 2. Planification strategies
 3. User centered design
9. Validación (2 hores)
 1. Usability

Methodology

LEARNING MANAGEMENT SYSTEM (LMS): Caronte (<http://caronte.uab.cat>, degree of Enginyeria de Dades (Data Engineering), subject Visualització de Dades (Data Visualization), curs VD Curs 2020-21).

The student learning process is based on two types of face-to-face activities (Theory and Seminars). Attendance at online theory and seminar classes is NOT REQUIRED, but YES HIGHLY RECOMMENDED.

THEORY

Theory classes are taught through master classes with transparencies and blackboard. All course material will be accessible through the Virtual Platform. Some of the theory sessions will be dedicated to solving exercises posed in the same classroom with the aim of delving into the most relevant concepts and teaching students to identify and analyze typical errors.

SEMINARS

In the seminar sessions we will alternate two activities: problems (7 sessions) and practices (6 sessions).

In the PROBLEMS sessions we will propose exercises to solve on the following topics:

- Simple graphs (bars, lines, boxplots, scatter, etc.) - Tool: R
- Complex graphs (parallel coordinates, small multiples, networks, etc.) - Tool: R
- Other tools (commercial and free) - Tools: Tableau Public, DataWrapper, Gephi
- Network visualization
- Map display
- Multivariable and multidimensional data graphs
- Animated and interactive graphics - Tools: R, 3D viewer
- Presentation (Titles, scales, captions and annotations) - Tools: R and design software such as Inkscape
- Dashboards - Tools: Table vs. R

In these problem sessions the exercises to be solved are proposed, which the students do with the tutoring of the teacher. All documentation (script statements, problem results) of the seminars is available on the Virtual Platform.

The PRACTICE sessions aim to train students to study and define the different views of data sets and assess the advantages and disadvantages. We will propose two practical works:

1. Treatment with a default date set. Given a date set provided by teachers, it is necessary to clean, prepare and visualize with exploratory graphs, justifying which is the best.
2. Treatment with a free date set. From a date set chosen or generated by the students, clean, prepare and visualize with exploratory graphs, justifying which is the best.

They will be carried out in groups of THREE people and for each student a dedication of 12 hours will be required in 6 sessions on line supervised by professor and 33 hours of work of the student of individual or group way of preparation of the practices, work of the student without the tutelage of the teacher and the preparation of reports.

Attendance at online internship sessions supervised by a teacher IS MANDATORY. Cases of absence will be studied in the duly justified internship sessions.

The student prepares the internship work explained in the internship statement. This preparation must be prior to attending the corresponding online internship session. The work is done by attending the online internship sessions at the established times.

The regulations for the delivery and evaluation of internships can be found in the regulations document available on the Virtual Platform. The work of internships will be evaluated and an individual exam will be held in the delivery of each internship.

Annotation: Within the schedule set by the centre or degree programme, 15 minutes of one class will be reserved for students to evaluate their lecturers and their courses or modules through questionnaires.

Activities

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Explanation and problem solving	14	0.56	3, 1, 2
Lectures	26	1.04	1, 4, 2
Resolution and practical exercises	12	0.48	3, 1, 2
Type: Supervised			
Case studies	4	0.16	3, 1, 2
Type: Autonomous			
Explanation and problem solving	30	1.2	3, 1, 2
Individual study	26	1.04	1, 4, 2
Resolution of practical exercises	30	1.2	3, 1, 2

Assessment

The subject assessment is done by continuous evaluation, based in two indicators: Theory and seminars:practices.

THEORY

There will be two individual written exams, with a weight of 50% each on the Theory Note. The first test (Par1) will be done approximately in the middle of the semester and will evaluate the theoretical concepts explained to date and the ability to solve visualization problems similar to those worked on in the seminar sessions: problems. The second test (Par2) will be carried out at the end of the semester and will evaluate the concepts explained so far, of theory and the ability to solve problems similar to those worked on in the seminar sessions: problems.

Recovery test: In case the theory note does not reach the approved, students may submit to a review of recovery on the contents of the partial examination suspended.

§

The practice assessment (notes Pr_1, Pr_2, Pr_3) is done by means of two activities:

- Assessment of the report written and the practical work done by the students (grades D1, D2), group grade.
- Individual written exam carried out on each practice (grades E1, E2).

Examen de recuperació: En cas de que no s'assoleixi no s'assoleixi el 5 en qualsevol de les proves, l'alumne té una nova oportunitat d'avaluació amb el mateix format i normes.

MARK INDICATORS:

The final grade of the subject is calculated in the following way (between square brackets the minimum marks to make half):

$$Pract = PG + PI$$

$$Par1 = ExPar1 + Pract [ExPar1 \geq 4] \text{ --- } Par2 = ExPar2$$

$$Theory\ Mark = 0,4 * (Par1) + 0,6 * (Par2) [Par1, Par2 \geq 5]$$

$$Project\ Mark = 0,4 * Prj_1 + 0,6 * Pr_2 [Pr_1, Pr_2 \geq 5]$$

$$FINAL\ SUBJECT\ MARK = 0,4 * Theory\ Mark + 0,6 * Project\ Mark$$

¶

ASSESSMENT CRITERIA:

- In order to consider approved any of the parts (theory and practices) a minimum of 5 must be obtained, that is to say Theory Mark ≥ 5 and Practice Work Mark ≥ 5 .
- PERMANENCE IN THE EXAMS: In any exam the student will not be able to leave the classroom during the first 45 minutes of the exam time, except for justified reasons or by indication of the teaching staff.
- In all retakes of assessable activities (partial or practice exams, RV project, etc.), the maximum mark will be 8.
- IMPROVING AN EXAM MARK: Students may take a retake to improve an exam score. Failure to pass the exam means waiving the previous exam mark obtained. In case it does not deliver it means that it accepts the previous exam mark obtained.
- To calculate the final grade of the subject, the theory and practices work marks must be approved.
- NOT EVALUABLE: In the case of not reaching the minimum required in any of the assessment activities, if the calculation of the final grade is equal to or greater than 5, a 4 note will be placed on the file.

- NOT PASS: In the event of not reaching the minimum required in any of the evaluation activities, if the calculation of the final grade is equal to or greater than 5, a grade of 4 will be added to the file.
- MH: The student with a grade equal to or higher than 9 points will be able to obtain an MH, at the discretion of the teacher. As the number of MHs cannot exceed 5% of the enrolled students, they will be awarded to those with the highest final grades, at the discretion of the teacher. Students who have taken all the assessment tests in the first call will be awarded.
- REPEATING STUDENTS: No approved part is approved (theory, problems, practices) from one academic year to another.
- IMPORTANT FOR ALL STUDENTS: It is important to enroll in Caronte, because the materials of the subject are published, the deliveries are made and the notes of the subject are published. You must also sign up for a group of internships. At the start of the course you must sign up for Caronte in the Interactive Graphical View course "VGI Curs 2021-22" course. If this is the first time you enter Caronte, you must provide your NIU, password, first and last name and a photo ID in JPG format. Caronte is a manager different from the UAB Virtual Campus, with the need for passwords to be the same.
- Students who have a mark equal to or greater than 9 points can obtain a MH. As the number of MH can not exceed 5% of students enrolled, students who have the highest final grades will be awarded, at the discretion of the teacher.

EVALUATION TIMETABLE:

-
- Recovery Exam: According to the academic calendar of the School of Engineering.
 - Exams and deliveries of practices: Dates in the rule of practices, that will publish in the Virtual Platform.
-

Caronte is a manager different from the UAB Virtual Campus, which means that passwords may not be the same.

Notwithstanding other disciplinary measures deemed appropriate, and in accordance with the current academic regulations, irregularities committed by a student that may lead to a variation of the qualification will be classified by zero (0). Assessment activities qualified in this way and by this procedure will not be recoverable. If it is necessary to pass any of these assessment activities to pass the subject, this subject will be suspended directly, without opportunity to recover it in the same course. These irregularities include, among others:

- The total or partial copy of a practice, report, or any other evaluation activity
- Let copy
- Present a group work not done entirely by the members of the group.
- Present as own materials produced by a third party, although they are translations or adaptations, and generally works with non-original and exclusive elements of the student.
- Have communication devices (such as mobile phones, smart watches, etc.) accessible during theoretical-practical tests (individual exams).

If you do not pass the subject due to the fact that none of the evaluation activities do not reach the required minimum level of work, the numerical note of the file will be the lowest value between 4 and the weighted average of the notes. With the exceptions that the "non-evaluable" qualification will be awarded to students who do not participate in any of the assessment activities, and that the numerical note of the file will be the lowest value between 3.0 and the weighted average of the notes in case the student has committed irregularities in an evaluation act (and therefore not approved for compensation).

SUMMING UP: Copying, copying, or plagiarizing any of the assessment activities is equivalent to a SUSPENSION with a score of less than 3.5.

Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
-------	-----------	-------	------	-------------------

Practical mark (examen and deliveries)	50%	4	0.16	3, 1, 2
Theory Mark (partial exams)	50%	4	0.16	1, 4, 2

Bibliography

SUBJECT MATERIALS (notes, slides, papers): In LMS.

BASIC BIBLIOGRAPHY:

- A. Cairo, The functional art, <http://www.thefunctionalart.com/2012/09/download-three-chapters-of-functional.html>, 2012 (darer accés: juliol 2020).
- T. Munzner, Visualization Analysis & Design, CRC Press, 2014.

C OMPLEMENTARY BIBLIOGRAPHY:

- S. Few, Information Dashboard Design: Displaying Data for At-a-glance Monitoring, 2nd edition, Analytics Press, 2013.
- A. Cairo, The truthful art. Data, charts and maps for communication, New Riders, 2016.
- A. Kirk, Data Visualization. A handbook for Data Driven Design, 2nd edition, SAGE Publications, 2019.
- C.O. Wilke, Fundamentals of Data Visualization. A Primer on Making Informative and Compelling Figures, O'Reilly Media, 2019.

Software

In this subject we will use the following software:

- Software R: R is 'GNU S', a freely available language and environment for statistical computing and graphics which provides a wide variety of statistical and graphical techniques. You can download this environment in the following link: <https://cran.rstudio.com/>.
- RStudio: Es tracta d'un framework que treballa sobre R i Python. This framework can be downloaded in different (Windows, MacOS, Linux) in the following link: <https://www.rstudio.com/products/rstudio/>.
- Tableau is a powerful data analysis and visualization tool used in the area of Business Intelligence. We have an educational license for our students to use. We can also download a free demo version in the following link: <https://www.tableau.com/es-es/why-tableau/what-is-tableau>.

All software to be used is freely accessible or offered free of charge to students with educational licenses.