Course guide
2703119 - DV - Data Visualization

Last modified: 30/01/2026

Unit in charge: Barcelona School of Informatics
Teaching unit: 1022 - UAB - (ANG) pendent.

Degree: BACHELOR'S DEGREE IN BIOINFORMATICS (Syllabus 2024). (Compulsory subject).

Academic year: 2025 **ECTS Credits:** 6.0 **Languages:** English

LECTURER

Coordinating lecturer:

Others:

PRIOR SKILLS

Basic knowledge in R and familiarity with RStudio are prerequisites.

LEARNING RESULTS

Knowledges:
K1. Recognise the fundamental principles of biology, from the cellular to the organismic scale, and how they relate to current knowledge in bioinformatics, data analysis and machine learning, achieving an interdisciplinary vision with an emphasis on biomedical applications.
K2. Identify statistical and computational methods and mathematical models that can be used to solve problems in molecular biology, genomics, medical research and population genetics.
K7. Analyse sources of valid and reliable scientific information to determine the state of the art of a bioinformatics problem and how to tackle it.

Skills:
S4. Develop specific problem-solving tools for the interpretation of biological and biomedical data, including complex visualisations.
S5. Communicate information, ideas, problems and solutions from bioinformatics and computational biology to a general audience.
S7. Implement programming methods and data analysis based on the development of working hypotheses within the area of study.
S8. Make and defend reasoned decisions when solving problems in biology and, in appropriate fields, the health sciences, computer science and experimental sciences.
S9. Exploit biological and biomedical information to transform it into knowledge, in particular by extracting and analysing information from databases to solve new biological and biomedical problems.

Competences:
C2. Recognise the complexity of the economic and social phenomena typical of a welfare society and relate welfare to globalisation, sustainability and climate change in order to use techniques, technology and principles of economics and sustainability in a balanced and compatible way.
C3. Communicate orally and in writing with others in English about learning outcomes, thought processes and decision making.
C4. Work as part of a multidisciplinary team, whether as a team member or in a leadership role, to contribute to the development of projects (including business and research projects) with pragmatism, a sense of responsibility and ethical principles, undertaking commitments with due regard to the resources available.

TEACHING METHODOLOGY

During theoretical sessions, the professor will present concepts dynamically, using examples and solving practical casses in class. During practical sessions, students will independently work on hands-on exercises, with supervision and assistance from the professor as needed. Both theory and practical lessons require a laptop.

LEARNING OBJECTIVES OF THE SUBJECT

1.Visualize, manipulate and extract biological data
2.Know existing techniques and computational tools in a particular field
3. Evaluate what is the most suitable technical and/or computational tool in every situation

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

- Basic tools: Grammar of Graphics (ggplot2)

Description:
Theoretical sessions on perception, visual illusions, the Grammar of Graphics by applying ggplot2, as well as specialized libraries and advanced visualizations.

- Interactive visualizations using htmlwidgets and Shiny

Description:
Learn how to create interactive visualizations with htmlwidgets packages and Shiny applications.

- Principal component analysis (PCA)

Description:
Explore techniques for visualizing complex data and dimensionality reduction (PCA).

- Non-linear projections: t-SNE and UMAP

Description:
Application of t-SNE and UMAP methods for data reduction.

ACTIVITIES

Practical lessons

Description:
Fill RMarkdown document / application-based

Specific objectives:
1, 2, 3

Full-or-part-time: 22h
Practical classes: 22h

Home assignments

Description:
Conceptual / synthesis-based / application-based

Specific objectives:
1, 2, 3

Full-or-part-time: 53h
Practical classes: 8h
Self study: 45h

Mid-term exam

Description:
Conceptual / synthesis-based / application-based

Specific objectives:
1, 2, 3

Full-or-part-time: 22h
Guided activities: 2h
Self study: 20h

Final exam

Description:
Conceptual / synthesis-based / application-based

Specific objectives:
1, 2, 3

Full-or-part-time: 28h
Guided activities: 3h
Self study: 25h

Theoretical lectures

Description:
Conceptual / synthesis-based / application-based

Specific objectives:
1, 2, 3

Full-or-part-time: 25h
Theory classes: 25h

GRADING SYSTEM

The evaluation of the subject will be structured as follows:

1. Active participation in class (10%): Weekly assessment of the participation in the theoretical and practical sessions, including discussions, activities, and brief quizzes.

2. Assignments (40%): Group activities will be evaluated through four assignments per main unit.

3. Midterm exam (20%): theory-practical exam to evaluate the concepts acquired during the first block of the subject.

4. Final exam (30%): theory-practical exam covering all concepts.

Both midterm and final exams are done using a computer.

The weighted grade of the midterm exam and the final exam requires a minimum score of 3.5 out of 10 to consider the other parts of the evaluation. A final grade of at least 5 out of 10 is required to pass the course. Plagiarism or cheating will result in course failure and potential disciplinary actions.

A student is considered to have taken the subject if he/she takes the final exam. If the student has taken the subject but has failed, then the student may take the re-evaluation exam and in this case the grade of the subject will be 40% home assignments, 10% participation and 50% recovery exam.

BIBLIOGRAPHY

Basic:
- Tufte, Edward R. The Visual display of quantitative information. 14th printing. Cheshire, Connecticut: Graphics Press, cop. 1983. ISBN 096139210X.
- Wilkinson, Leland; Wills, Graham. The Grammar of graphics [on line]. 2nd ed. New York: Springer Science, cop. 2005 [Consultation: 03/03/2025]. Available on: https://link-springer-com.recursos.biblioteca.upc.edu/book/10.1007/0-387-28695-0. ISBN 1-280-46066-0.
- Chang, Winston. R graphics cookbook : practical recipes for visualizing data [on line]. 2nd ed. Beijing: O'Reilly, 2018 [Consultation: 03/03/2025]. Available on: https://ebookcentral-proquest-com.recursos.biblioteca.upc.edu/lib/upcatalunya-ebooks/detail.action?pq-origsite=primo&docID=5568320. ISBN 9781491978597.
- McCandless, David. Information is beautiful. New ed., rev., recalculated and reimagined. London: William Collins, 2012. ISBN 9780007492893.
- Wickham, Hadley. Ggplot2 : elegant graphics for data analysis [on line]. 2nd ed. Dordrecht ; New York: Springer, 2009 [Consultation: 03/03/2025]. Available on: https://ebookcentral-proquest-com.recursos.biblioteca.upc.edu/lib/upcatalunya-ebooks/detail.action?pq-origsite=primo&docID=4546676. ISBN 9783319242774.

RESOURCES

Hyperlink:
- http://ggplot2.tidyverse.org/reference/- http://r4ds.had.co.nz/data-visualisation.html- https://distill.pub/2016/misread-tsne/- https://liorpachter.wordpress.com/2014/05/26/what-is-principal-component-analysis/- https://towardsdatascience.com/11-dimensionality-reduction-techniques-you-should- https://umap-learn.readthedocs.io/en/latest/basic_usage.html- https://www.youtube.com/watch?v=fSgEeI2Xpdc