BASIC DETAILS:

Subject:	VISUALIZACIÓN DE DATOS		
Id.:	33711		
Programme:	DOBLE GRADO EN FARMACIA Y BIOINFORMÁTICA. PLAN 2019		
Module:	BIOINFORMÁTICA		
Subject type:	OBLIGATORIA		
Year:	3	**Teaching period:**	Primer Cuatrimestre
Credits:	3	**Total hours:**	75
Classroom activities:	32	**Individual study:**	43
Main teaching language:	Inglés	**Secondary teaching language:**	Castellano
Lecturer:	ALCAINE OTIN, ALEJANDRO (T)	**Email:**	lalcaine@usj.es

PRESENTATION:

In the digital era, the data volume grows exponentially, and its curation, analysis and visualization have become a relevant aspect for extracting and visualizing meaningful information. This course introduces to the student the key aspects of visual design, aesthetics and vector graphics formats. The student will learn tools for data curation and generation of static and dynamic visualizations using Python. The student will handle visualizer and data structure, transforming them into effective visualization and information. Finally, specific tools and libraries for Bioinformatics will be introduced such as hive plots, circular genomic plots and tree plots.

PROFESSIONAL COMPETENCES ACQUIRED IN THE SUBJECT:

General programme competences	G01	Use learning strategies autonomously for their application in the continuous improvement of professional practice.
	G02	Perform the analysis and synthesis of problems of their professional activity and apply them in similar environments.
	G03	Cooperate to achieve common results through teamwork in a context of integration, collaboration and empowerment of critical discussion.
	G04	Reason critically based on information, data and lines of action and their application on relevant issues of a social, scientific or ethical nature.
	G05	Communicate professional topics in Spanish and / or English both orally and in writing.
	G06	Solve complex or unforeseen problems that arise during the professional activity within any type of organisation and adapt to the needs and demands of their professional environment.
	G07	Choose between different complex models of knowledge to solve problems.
	G09	Apply information and communication technologies in the professional field.
	G10	Apply creativity, independence of thought, self-criticism and autonomy in the professional practice.
Specific programme competences	E02	Develop the use and programming of computers, databases and computer programs and their application in bioinformatics.
	E03	Apply the fundamental concepts of mathematics, logic, algorithmics and computational complexity to solve problems specific to bioinformatics.
	E04	Program applications in a robust, correct, and efficient way, choosing the paradigm and the most appropriate programming languages, applying knowledge about basic algorithmic procedures and using the most appropriate types and data structures.
	E05	Implement well-founded applications, previously designed and analysed, in the characteristics of the databases.
	E06	Apply the fundamental principles and basic techniques of intelligent systems and their practical application in the field of bioinformatics.
	E07	Apply the principles, methodologies and life cycles of software engineering to the development of a project in the field of bioinformatics.
	E12	Apply the principles and techniques of protein computational modelling to predict their biological function, their activity or new therapeutic targets (Structural Bioinformatics, Computational Toxicology).
	E13	Apply omics technologies for the extraction of statistically significant information and for the creation of relational databases of biodata that can be updated and publicly accessible to the scientific community.
	E14	Use programming languages, most commonly used in the field of Life Sciences, to develop

		and evaluate techniques and/ or computational tools.
	E15	Infer the evolutionary history of genes and proteins through the creation and interpretation of phylogenetic trees.
	E16	Plan linkage and association studies for medical and environmental purposes.
	E17	Induce complex relationships between samples by applying statistical and classification techniques.
	E18	Apply statistical and computational methods to solve problems in the fields of molecular biology, genomics, medical research and population genetics.
	E21	Apply computational and data processing techniques for the integration of physical, chemical and biological concepts and data for the description and/ or prediction of the activity of a substance in a given context.

PRE-REQUISITES:

The course will be delivered in English language. Academic reading and writing skills are expected from the students. Theory lectures will be completed with programming examples and practices/ projects will require programing in Python, so students should have general programming knowledge. Basic knowledge of mathematics and statistics are also required.

SUBJECT PROGRAMME:

Subject contents:

1 - Introduction to data visualization
1.1 - Basic concepts of visual design.
1.2 - Figure design considerations and strategies.
1.3 - Image formats.
2 - Tools and libraries for data visualization
2.1 - Review of Python
2.2 - Data preparation with Python
2.3 - Introduction to MatPlotLib and Seaborn
2.4 - Interactive graphics with Bokeh
3 - Basic data representations
3.1 - Comparison plots
3.2 - Relations plots
3.3 - Composition plots
3.4 - Distribution plots
4 - Bioinformatic plots
4.1 - Python libraries for Bioinformatics
4.2 - Hive plots, Pylogenetic Trees, Graphs and Circular genomic plots

Subject planning could be modified due unforeseen circumstances (group performance, availability of resources, changes to academic calendar etc.) and should not, therefore, be considered to be definitive.

TEACHING AND LEARNING METHODOLOGIES AND ACTIVITIES:

Teaching and learning methodologies and activities applied:

Magistral lectures will be used to explain the different aspects of the subject and encouraged to be highly dynamic and interactive with visual examples and codes. Small exercises will be solved during class in order to consolidate the concepts.

The subject is highly practical, therefore the Magistral lectures schedule will be altered with **workshop sessions** where the students consolidate and practice the subject concepts mixing problem-based and project-based learning approaches.

Additionally, a **practical session** where the students will put in practice the concepts of the subject using a project-based learning approach.

The subject requires a high practical effort from the student, and it is important to follow the concepts and exercises during the presential lectures. Additionally, the students will have via PDU many proposed exercises by the teacher with small tasks and challenges for autonomous learning. As a matter of that, the lecturer will be available to students during the tutorial schedule to help them in all matters concerning the course.

Student work load:

Teaching mode	Teaching methods	Estimated hours
Classroom activities	Master classes	16
	Workshops	5
	Laboratory practice	2
	Assessment activities	2
	Workshop webinars	5
	Tutorials	2
Individual study	Individual study	12
	Individual coursework preparation	18
	Recommended reading	6
	Individual exercises	5
	Collaborative activities	2
	Total hours:	75

ASSESSMENT SCHEME:

Calculation of final mark:

Individual coursework:	30	%
Final exam:	50	%
Continuous assessment:	15	%
Test:	5	%
TOTAL	100	%

*Las observaciones específicas sobre el sistema de evaluación serán comunicadas por escrito a los alumnos al inicio de la materia.

BIBLIOGRAPHY AND DOCUMENTATION:

Basic bibliography:

WILKE, Claus O. Fundamentals of Data Visualization: O'Reilly Media, 2019
DÖBLER, M. The Data Visualization Workshop: Packt Publishing 2020

Recommended bibliography:

TUFTE, E.R. The Visual display of quantitative information: Graphics Press 1983
JOLLY, K. Hands-on data visualization with bokeh: Packt Publishing 2018
BASSI, S. Python for Bioinformatics: Chapman & Hall/ CRC 2010

Recommended websites:

Python documentation	https://www.python.org/doc/
MatPlotLib documentation	https://matplotlib.org
Seaborn documentation	https://seaborn.pydata.org
Bokeh documentation	https://docs.bokeh.org/en/latest/index.html
BioPython documentation	https://biopython.org
Cognitive Class MOOC: Applied Data Science with Python	https://cognitiveclass.ai/learn/data-science-with-python