



Universitat de Lleida

DEGREE CURRICULUM

STORYTELLING WITH DATA

Coordination: VILAPRIÑO TERRE, ESTER

Academic year 2023-24

Subject's general information

Subject name	STORYTELLING WITH DATA			
Code	14709			
Semester	1st Q(SEMESTER) CONTINUED EVALUATION			
Typology	Degree	Course	Character	Modality
	Master's Degree in Biomedical Research	1	OPTIONAL	Attendance-based
Course number of credits (ECTS)	4			
Type of activity, credits, and groups	Activity type	PRAULA		TEORIA
	Number of credits	2.5		1.5
	Number of groups	1		1
Coordination	VILAPRIÑO TERRE, ESTER			
Department	BASIC MEDICAL SCIENCES			
Important information on data processing	Consult this link for more information.			

Teaching staff	E-mail addresses	Credits taught by teacher	Office and hour of attention
MARÍN SANGUINO, ALBERTO	alberto.marin@udl.cat	1	
VAQUEIRO DE CASTRO ALVES, RUI CARLOS	rui.alves@udl.cat	,5	
VILAPRIÑO TERRE, ESTER	ester.vilapriño@udl.cat	2,5	

Learning objectives

Learning results:

- You will learn about basic data analysis and prediction techniques: dimensionality reduction, linear and logistic regression model, clusters, dendograms, neural networks and SVM.
- You will know the main data analysis and representation packages in R (ggplot2, lattice, leaflet and shiny) and Python (seaborn, pandas and numpy)
- You will know how to use them in a way adapted to your data and problem.
- You will be able to generate the right visualizations to communicate your results clearly and accurately.
- You will understand the importance of group work and cooperation among researchers.
- You will be able to formulate work objectives, plan the work, carry out the experiments, present the results obtained and draw conclusions.

Competences

CB1 Possess and understand knowledge that provides a basis or opportunity to be original in the development and / or application of ideas, often in a research context

CB2 Know how to apply the knowledge acquired and be able to solve problems in new or little-known environments within broader (or multidisciplinary) contexts related to their area of study

CB3 Being able to integrate knowledge and face the complexity of formulating judgments based on information that, being incomplete or limited, includes reflections on the social and ethical responsibilities linked to the application of their knowledge and judgments.

CB4 Know how to communicate their conclusions –and the latest knowledge and reasons that support them– to specialized and non-specialized audiences in a clear and unambiguous way

CB5 Possess the learning skills that allow them to continue studying in a way that will have to be largely self-directed or autonomous

CG3 Capacity for teamwork, leadership and decision making.

CG4 Ability to think critically and creatively with your work and that of other researchers

CG5 Ability to prepare, process and interpret the results obtained with rigor and applying the appropriate technologies

CG6 Know how to guide research towards lines of medical and translational interest (diagnosis and therapy)

CG7 Be able to present scientific reports and scientific articles that can be considered for publication in international journals.

CE4 Recognize high throughput techniques and be able to use bioinformatic data analysis tools.

CT1 Have a correct oral and written expression

CT2 Mastering a foreign language

CT3 Mastering ICT

CT4 Respect the fundamental rights of equality between men and women, the promotion of Human Rights and the values proper to a culture of peace and democratic values

Subject contents

1. **Story trends from data:** Linear Regression and Correlations.
2. **From complex data to a simpler story:** Principal Component Analysis and Factor Analysis.
3. **Group building from data:** Clusters and Dendograms.
4. **Group assignment from data:** Logistic Regression, Neural Networks, and Support Vector Machine.
5. **Story telling tools I.** A sip of R: Tydiverse, Ggplot2, Plotly, Shiny, and Leaflet.
6. **Story telling tools II.** A sip of Phyton: Seaborn, Pandas, and Numpy.

Methodology

- Active class sessions based on data analysis and student participation.
- Introduction of statistical and technical tools by lecturers.
- Intensive use of computer programs (R and Phyton).
- Students should bring their own computer for class activities (this facilitates further work at home)
- Autonomous work by students

Development plan

- The course will be developed in 20 sessions of 2 hours each. No distinction is made among practical and theoretical work.
- Depending of the situation with respect the SARS-COV-2 pandemic, some of the activites would be developed virtually.
- A specific schedule will be added in the contens section of the virtual campus as soon as we know the conditions for the next lecturing period.

Evaluation

- 20% Oral presentation of work
- 15% Class participation in discussions and data analysis
- 65% Assignments (article analysis, projects, data analysis)

Bibliography

Web links, articles, data collections, etc. will be provided during the course and made available at the virtual

campus.