



Universitat de Lleida

GUIA DOCENT

STORYTELLING WITH DATA

Coordinació: VILAPRIÑO TERRE, ESTER

Any acadèmic 2021-22

Informació general de l'assignatura

Denominació	STORYTELLING WITH DATA			
Codi	14709			
Semestre d'impartició	1R Q(SEMESTRE) AVALUACIÓ CONTINUADA			
Caràcter	Grau/Màster	Curs	Caràcter	Modalitat
	Màster Universitari en Investigació Biomèdica	1	OPTATIVA	Presencial
Nombre de crèdits assignatura (ECTS)	4			
Tipus d'activitat, crèdits i grups	Tipus d'activitat	PRAULA		TEORIA
	Nombre de crèdits	2.5		1.5
	Nombre de grups	1		1
Coordinació	VILAPRIÑO TERRE, ESTER			
Departament/s	CIÈNCIES MÈDIQUES BÀSIQUES			
Distribució càrrega docent entre la classe presencial i el treball autònom de l'estudiant	60% in class and 40% autonomous work by students			
Informació important sobre tractament de dades	Consulteu aquest enllaç per a més informació.			
Idioma/es d'impartició	English			
Distribució de crèdits	All activities will involve practical work by students under teacher supervision. We shall use publicly available data for introducing concepts, strategies and analytical methods.			

Professor/a (s/es)	Adreça electrònica professor/a (s/es)	Crèdits impartits pel professorat	Horari de tutoria/lloc
MARÍN SANGUINO, ALBERTO	alberto.marin@udl.cat	1	
VAQUEIRO DE CASTRO ALVES, RUI CARLOS	rui.alves@udl.cat	1	
VILAPRIÑO TERRE, ESTER	ester.vilapriño@udl.cat	2	

Informació complementària de l'assignatura

Data processing, analysis and representation are crucial for a correct interpretation of biomedical research results. This is more evident when dealing with results from data-intensive and highly complex research that integrates heterogeneous data from different sources with varying quality.

Therefore, a biomedical researcher should know the methodologies and tools available to facilitate organization, verification, analysis and representation of data in order to obtain the appropriate evidence to support the conclusions of an study.

This course will provide the student with this knowledge and train him in choosing the methodology and tools appropriate to his research.

Objectius acadèmics de l'assignatura

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.

Competències

CB1 Posseir i comprendre coneixements que aportin una base o oportunitat de ser originals en el desenvolupament i / o aplicació d'idees, sovint en un context de recerca

CB2 Saber aplicar els coneixements adquirits i tenir capacitat de resolució de problemes en entorns nous o poc coneguts dins de contextos més amplis (o multidisciplinaris) relacionats amb la seva àrea d'estudi

CB3 Ser capaç d'integrar coneixements i enfrontar-se a la complexitat de formular judicis a partir d'una informació que, sent incompleta o limitada, inclogui reflexions sobre les responsabilitats socials i ètiques vinculades a l'aplicació dels seus coneixements i judicis

CB4 Saber comunicar les conclusions -i els coneixements i raons últimes que les sustenten- a públics especialitzats i no especialitzats d'una manera clara i sense ambigüitats

CB5 Posseir les habilitats d'aprenentatge que els permetin continuar estudiant d'una manera que haurà de ser en gran manera autodirigida o autònoma

CG3 Capacitat de treball en equip, lideratge i presa de decisions.

CG4 Capacitat de pensament crític i creatiu amb el seu treball i el d'altres investigadors

CG5 Capacitat de preparar, processar i interpretar els resultats obtinguts amb rigor i aplicant les tecnologies apropiades

CG6 Saber orientar la investigació a línies d'interès mèdic i traslacional (diagnòstic i teràpia)

CG7 Ser capaç de presentar memòries científiques i articles científics que puguin ser considerats per a la seva publicació en revistes internacionals

CE4 Reconèixer les tècniques d'alt rendiment (high throughput) i ser capaç d'utilitzar les eines bioinformàtiques d'anàlisi de dades.

CT1 Tenir una correcta expressió oral i escrita

CT2 Dominar una llengua estrangera

CT3 Dominar les TIC

CT4 Respectar els drets fonamentals d'igualtat entre homes i dones, a la promoció dels Drets Humans i als valors propis d'una cultura de pau i de valors democràtics

Continguts fonamentals de l'assignatura

1. **Basic data description:** tables, graphics, confidence intervals, and related tools that are important to understand.
2. **Story telling tools I.** A sip of R: Tydiverse, Ggplot2, Plotly, Shiny, and Leaflet.
3. **Story trends from data:** Linear Regression and Correlations.
4. **From complex data to a simpler story:** Principal Component Analysis and Factor Analysis.
5. **Group building from data:** Clusters and Dendograms.
6. **Group assignment from data:** Logistic Regression, Neural Networks, and Support Vector Machine.
7. **Story telling tools II.** A sip of Python: Seaborn, Pandas, and Numpy.

Eixos metodològics de l'assignatura

- 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 Python).
- Students should bring their own computer for class activities (this facilitates further work at home)
- Autonomous work by students

Pla de desenvolupament de l'assignatura

- 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.

Sistema d'avaluació

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

Bibliografia i recursos d'informació

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