Universidad de Navarra

Data Science (MIDI)

Guía docente 2025-26

PRESENTATION

Overview:

"*Data Science*" is an expanding multidisciplinary field that, as its name suggests, focuses on the study of data. Its main objective is to **extract hidden patterns from the data** and generate new insights that could **guide rational decision-making**. To achieve this, a variety of techniques are used, such as statistics, predictive models or Machine Learning algorithms, along with programming technique

This course therefore has **two pillars**. On one hand, the **mathematical/statistical foundation** required to approach the computational, methodological, and mathematical problems derived from the study of data, and essential for taking objective and scientific decisions. The other pillar is **programming**, which is crucial for carrying out the exploration, transformation, and subsequent analysis of data, following the previously mentioned statistical and mathematical standards.

Although the impact of this discipline certainly extends to all aspects of the modern world, in this course, we will focus on its potential application in the **field of health sciences**, and more specifically in the research and development of new drugs.

- **Degree**: DATA SCIENCE
- **Module**: Modulo I: Drug Reseach. Topic: Drug Discovery
- **ECTS**: 3 ETCs
- **Course/semester**: 1er semestre
- **Type**: Mandatory
- **Lecturers**: Zinnia Parra Guillén, y María Jesús Garrido Cid
- **Language**: English
- **Location, schedule**: SEE CALENDAR

COMPETENCES

CB6. To possess and understand knowledge that provides a foundation or opportunity to be original in the development and/or application of ideas, often in a research context.

CB9. The students know how to communicate their conclusions and the knowledge and ultimate reasons that support them to specialized and non-specialized audiences in a clear and unambiguous way.

CB10. The students will have acquired learning abilities that will permit them to continue studying in a self-directed and autonomous manner.

CG4. Identify and know how to create strategies and actions aimed to achieve the planned objectives and concrete the required resources accordingly, in the field of the pharmaceutical company.

CG5. Learn the current techniques and trends related to drug research, development and innovation.

CG6. Acquire critical ability to make the necessary decisions and adapt to new situations that may arise in the field of pharmaceutical and related companies.

Universidad de Navarra

CE1. Design new drugs and innovative therapies for the pharmaceutical industry by applying specialized techniques in chemoinformatics and bioinformatics.

PROGRAM

Theory

Topic 1. Introduction to Data Science

Topic 2. Data transformation and visualization

Topic 3. Descriptive statistics

Topic 4. Statistical inference

Topic 5. Predictive models

Practice

Introduction to R programming language (Practical computer sessions)

EDUCATIONAL ACTIVITIES

3 ECTS x 25 = 75 h

The course consists of in-person activities (1.28 ECTS, 32 hours) and personal activities (1.72 ECTS, 43 hours) as follows:

- Theory (in-person): 16 h
- Practices (in-person) : 14 h
- Personal work: 43 h
- Written evaluation: 2h

ASSESSMENT

ORDINARY CALL

The evaluation of the course will be based on the following criteria:

- A written exam consisting of the resolution of a questionnaire that includes theoretical and practical content.: 60%
- Group work: 30 %
- Class participation: 10 %

The final grade will be the weighted average of each component, with a minimum score of 4 out of 10 required on the theoretical exam.

Universidad de Navarra

SECOND CALL

- Students who do not pass the regular exam (minimum score of 5 out of 10) will be required to retake the exam, including both theoretical and practical aspects.

Special Needs

Students with special educational needs must contact the Faculty/School's Study Coordination Office in advance to obtain authorization for accommodations (for example, more time for exams). This authorization must be sent by the student to the professor. It is recommended that this be done at the beginning of the semester.

ATTENTION

Please note that any attempt at fraud, copying, plagiarism, or other irregular behavior constitutes a serious infraction, as defined in Title IV "Rules of Academic Discipline for Students" within the System of Rules on Coexistence at the University of Navarra (Rules on Coexistence UNAV).

OFFICE HOURS

Dra. Zinnia Parra Guillén: zparra@unav.es

- 0F14. Edificio de Ciencias. Planta 0

Dr. José Ignacio Fernández de Trocóniz: itroconiz@unav.es

- 0F12. Edificio de Ciencias. Planta 0

Dra. María Jesús Garrido Cid: mgarrido@unav.es

- 0F13. Edificio de Ciencias. Planta 0

No pre-defined schedule. Arrange an appointment with the lecturer by email.

BIBLIOGRAPHY & RESOURCES

https://cran.r-project.org/

Main resources

Required material for the following of the course:

- Lecturers' slides (this material will be uploaded to ADI)
- Instructions for conducting the practical sessions with the computer. These will be uploaded to ADI before the sessions take place.

Recommending readings

Universidad de Navarra

- Free manual: W. N. Venables, D. M. Smith and the R Core Team. An Introduction to R. (Notes on R: A Programming Environment for Data Analysis and Graphics). (https://cran.r-project.org/doc/manuals/R-intro.pdf)
- Free manual: John Verzani. simpleR: Using R for Introductory Statistics. (http://cran.r-project.org/doc/contrib/Verzani-SimpleR.pdf)
- Book: Dalgaard, Peter. Introductory Statistics with R (Springer) 2008. Localízalo en la Biblioteca
- Book: Wickham, H, Çetinkaya-Rundel M and G Grolemund. *R for Data Science (2e) : Import, Tidy, Transform, Visualize and Model Data*. O'Reilly Media, Inc. (https://r4ds.hadley.nz/)
- Kuhn M and SIlge J. Tidy Modeling with R: A framework for Modeling in the Tidyverse (1e) O'Reilly Media, Inc.(https://www.tmwr.org/