

## COURSE SYLLABUS

### Data in Public and Social Services

2425-2-FDS01Q028-FDS01Q034M

---

#### Aims

This module aims at teaching students how to analyze medical data (especially, data of electronic health records) through computational statistics and machine learning techniques to infer new knowledge about the conditions of patients.

#### Contents

- Dataset search and retrieval
- Data preparation and data cleaning
- Exploratory data analysis
- Unsupervised machine learning
- Supervised machine learning
- Feature ranking
- Result understanding and validation
- R programming language

#### Detailed program

- Dataset search and retrieval
- Data preparation and data cleaning
- Exploratory data analysis
- Unsupervised machine learning
- Supervised machine learning
- Feature ranking

Result understanding and validation  
R programming language

## **Prerequisites**

Basic statistics and basic machine learning  
Basic knowledge of R o Python

## **Teaching form**

7 lectures (each of them made of 2 hours for a total of 14) of frontal theory teaching.  
5 lectures (4 of two hours and 1 of one hour for a total of 9 hours) of practical exercises on the laptop computer (interactive teaching).

## **Textbook and teaching resource**

Classes slides and scientific papers mentioned during classes

## **Semester**

Second semester

## **Assessment method**

The final exam consists of:

- 1- The development of a personal scientific project, to be deployed in R analyzing medical data through the techniques learnt during the theoretical classes and during the practical classes;
- 2- The delivery of a report on the project carried out;
- 3- An oral presentation of the project carried out.

In the first component, we will assess the student's understanding on the methods, their capability to apply them in R to real medical data, and their programming skills.

In the second component, we will assess the student's capability to describe the project carried out in a written report.

In the third component, we will assess the student's capability to narrate the project carried out through an oral presentation with slides.

There are no mid-term exam tasks.

## **Office hours**

To define via email by writing to [davide.chicco\(AT\)unimib.it](mailto:davide.chicco(AT)unimib.it)

## **Sustainable Development Goals**

GOOD HEALTH AND WELL-BEING

---