Data analysis

ECTS
3 crédits

Hourly volume
62h

Introducing

Objectives

At the end of this module, the student will have understood and be able to explain (main concepts):

-The main steps of a data science analysis: preparation, visualization & exploration, prediction, interpretation.
-The main methods in data exploration.
-The main concepts / dangers of statistical learning.
-The main methods of statistical learning on vector data, requiring little expert knowledge / tuning.
-The functioning of R and Python software for data science.

At the end of this module, the student should be able to:

-Solve simple exercises about the underlying mathematical theory.
-Put in action the data science methodology on case studies with R and Python.
-Criticize the assumptions and results, summarize the main conclusions.

Necessary prerequisites

Statistics: descriptive statistics
Probability: random vectors, probability distribution, Bayes law, multivariate normal distribution.
Algebra: vector spaces, Euclidean spaces, matrix calculus, eigenvalue decomposition.
Geometry / mecanics: barycenter, inertia, Huygens formula.

Practical info

Location(s)

Toulouse