

Mathematics

Data Visualization and Data Science

IDENTIFICATION

CODE : GEN-5-S1-EC-AVD
ECTS : 2.0

HOURS

Lectures : 8.0 h
Seminars : 4.0 h
Laboratory : 17.0 h
Project : 0.0 h
Teacher-student
contact : 29.0 h
Personal work : 6.0 h
Total : 35.0 h

ASSESSMENT METHOD

The evaluation is done through a data science project [see GEN-5-PROJ-S1].

Special case of exchange students who are only registered in GEN-5-AVD-S1: the evaluation of this course is done through an oral defense.

TEACHING AIDS

Course handout

TEACHING LANGUAGE

French

CONTACT

M. ROUSSET François
francois.rousset@insa-lyon.fr

AIMS

- When faced with a data set, be able to choose an analysis method
- Be able to conduct a data analysis in Python language
- Be able to interpret the results of a data analysis

CONTENT

FACTORIAL ANALYSIS

1. Principal component analysis (PCA)
2. PCA and hierarchical ascending classification
3. Other methods: factorial correspondence analysis, multiple correspondence analysis factor analysis, multiple correspondence analysis, mixed data factor analysis

TIME SERIES

1. Introduction
2. Characterization of univariate time series
3. Models for the prediction of time series
4. The ARIMA-SARIMA family
5. The case of multivariate series
6. Opening : neural network approach

BIBLIOGRAPHY

Escofier B., Pagès J. [2016] Analyses factorielles simples et multiples : Cours et études de cas. 5e édition, Dunod, Paris.

PRE-REQUISITE

Basics of linear algebra