EPFL

MATH-517	**Statistical computation and visualisation**
	Mhalla Linda

Cursus	Sem.	Type
Ing.-math	MA1, MA3	Opt.
Mathématicien	MA1, MA3	Opt.
Statistics	MA1, MA3	Opt.

Language of teaching	English
Credits	5
Withdrawal	Unauthorized
Session	Winter
Semester	Fall
Exam	During the semester
Workload	150h
Weeks	14
Hours	**4 weekly**
Lecture	2 weekly
Exercises	2 weekly
Number of positions	

It is not allowed to withdraw from this subject after the registration deadline.

Summary

The course will provide the opportunity to tackle real world problems requiring advanced computational skills and visualisation techniques to complement statistical thinking. Students will practice proposing efficient solutions, and effectively communicating the results with stakeholders.

Content

- Modern statistical computing environments (e.g., R, Rstudio and Python)

- Overview of other software (e.g., MATLAB, Julia)

- Aids to efficiency and reproducibility (e.g., GitHub, Markdown, Jupyter)

- Data management, wrangling, and ethics

- Statistical graphics (grammar, good practices, applications, and examples)
- EM algorithm and applications
- Kernel density estimation and smoothing
- Resampling methods for uncertainty assessment (bootstrap, jackknife, cross-validation), with applications to regression, time series and dependent data

- Markov chain Monte Carlo techniques (Gibbs sampler, Metropolis-Hastings algorithm, Hamiltonian Monte Carlo, convergence diagnostics) and software (e.g., Stan)

- Other methods for Bayesian inference (e.g., importance sampling, INLA, AGHQ, ...)

Keywords

Statistical computation, data visualisation, data wrangling, resampling methods, EM algorithm, Bayesian inference

Learning Prerequisites

Required courses

- Probability and statistics
- Linear models

Learning Outcomes

By the end of the course, the student must be able to:

- Plan complex visualisation and computational tasks
- Perform complex visualisation and computational tasks
- Implement reproducible computational solutions to statistical problems in modern environments and platforms

Transversal skills

- Take feedback (critique) and respond in an appropriate manner.
- Communicate effectively with professionals from other disciplines.
- Demonstrate the capacity for critical thinking
- Identify the different roles that are involved in well-functioning teams and assume different roles, including leadership roles.

Teaching methods

Two lecture hours per week, two hours of exercises and support on mini-projects

Expected student activities

Students will work on mini-projects in teams

Assessment methods

Contrôle continue

Supervision

Office hours	No
Assistants	Yes
Forum	No

Resources

Bibliography

Wickham H. & Grolemund G. (2017) R for Data Science
Bootstrap Methods and their Application
An Introduction to Statistical Learning

Ressources en bibliothèque

- Bootstrap Methods and their Application / Davison
- An Introduction to Statistical Learning / Gareth
- R for Data Science / Wickham

Moodle Link

- https://go.epfl.ch/MATH-517

Prerequisite for

Applied Statistics (MATH-516)