Data Analysis


Course level: Master (M1) **Track(s):** CPS2,MLDM,COSI,3DMT

ECTS Credits: 6

Course instructors: Fabrice Muhlenbach, Marc Sebban (UJM)

Education period (Dates): 1st semester **Language of instruction:** English

Expected prior-knowledge: basic mathematics and statistics

Aim and learning outcomes: This course gives the necessary mathematical background to perform data analysis using statistics, linear algebra and convex optimization. Practical sessions make use of the R-free software environment for statistical computing and graphics.

Keywords: probability, statistics, linear algebra, optimization, linear regression, PCA, clustering.

Syllabus:

- Basics in probabilities (chance experiments, random variables, moments, law of large number, …)
- Statistics (discrete and continuous distributions, estimates, Maximum Likelihood Estimation,...)
- Basics in linear algebra and in convex optimization.
- Linear/polynomial/logistic Regression (closed-form solution, batch and stochastic gradient descent)
- Principal Component Analysis
- Clustering

Organisation and timetable: Lectures (25h), tutorials (10h) and lab sessions (15h).

Form(s) of Assessment: written exam (2h, coef. 2), practical work/project (coef. 1)

Literature and study materials:

- Pattern Recognition, S theodoridis, K. Koutroumbas, 4th edition
- Introduction to Statistics and Data Analysis, R. Peck, C. Olsen, J. Devore, Brooks/Cole, 4th edition, 2010.
- Convex Optimization, Stephen Boyd & Lieven Vandenberghe, Cambridge University Press, 2012.
- On-line Machine Learning courses: https://www.coursera.org/

Additional information/Contacts:

Marc Sebban
University Jean Monnet, Saint-Etienne
E-mail: marc.sebban@univ-st-etienne.fr
Web page: http://perso.univ-st-etienne.fr/sebbanma/