

**Course instructors:** Fabrice Muhlenbach

**Language of instruction:** English

**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.

### Course outline:

- 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

### Teaching methods

- Lectures: 25 hours
- Tutorials: 10 hours
- Lab sessions : 15 hours

### 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/>

### Expected prior-knowledge

- Basic mathematics and statistics

### Evaluation criteria

- Written exam (2h, coef. 2)
- Practical work/project (coef. 1)