

FIN-525	Financial big data				
	Challet Damien				
Cursus		Sem.	Туре	Language of	English
Financial engineering minor		Н	Opt.	teaching	English
Financial engineering		MA1, MA3	Opt.	Credits	3
Statistics		MA1, MA3	Opt.	Session Winter	
				Semester	Fall
				Exam	During the semester
				Workload	90h
				Weeks	14
				Hours	3 weekly
				Lecture	3 weekly
				Number of positions II n'est p retirer après le d	as autorisé de se de cette matière délai d'inscription.

Remark

MA3 only

Summary

The course introduces modern methods to acquire, clean, and analyze large quantities of financial data efficiently. The second part expands on how to apply these techniques and robust statistics to financial analysis, in particular to intraday data and investment strategies.

Content

Big Data

- 1. The future of storage, computing power, efficiency
- 2. Financial data sources and acquisition
- 3. Data cleaning and formatting
- 4. Efficient visualization techniques
- 5. Robust estimators for financial data
- 6. Multicore/GPU and cluster computing
- 7. Out-of-core data analysis

Application to financial data

- 1. Intraday tick-by-tick data
- 2. Dimensionality reduction

Correlation matrix cleaning with random matrix theory

Clustering of assets and days

3. Brute-force trading strategy design and selection

Non-stationary predictions

Best and worst practices

Keywords

Big Data, stylized facts, data wrangling, dimension reduction, tick-by-tick data, trading strategy, strategy selection.

Learning Prerequisites

Required courses

• Very good programming skills (required) and a first experience with R or/and Python.

• Good knowledge of the probability and statistics concepts taught in the first (two) year(s) at EPFL. This includes the Central Limit Theorem and its important applications in statistics.

Recommended courses

- Advanced statistics
- Econometrics
- Investments
- Programming with R, or Python.

Important concepts to start the course See above

Learning Outcomes

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

- Choose appropriate methods and tools to manipulate and analyze complex financial data.
- Conduct efficient data cleaning.
- Implement financial big data analysis using R and Python
- Implement proper computationally intensive strategy backtests
- Plan computing resource usage time
- Infer financial measurables with robust estimates

Transversal skills

- Collect data.
- Write a scientific or technical report.
- Demonstrate a capacity for creativity.
- Access and evaluate appropriate sources of information.
- Continue to work through difficulties or initial failure to find optimal solutions.

Teaching methods

3 hours of ex-cathedra lectures and supervised applications for 14 weeks

Expected student activities

- · Actively participating at lectures
- Completing theoretical and practical exercices during the lectures.
- Writing up of a report which demonstrates the ability to apply the concepts and tools introduced in this course.

Assessment methods

• Group projects 100%

Supervision

Assistants	Yes
Others	Assistant support envisioned depending on attendance
	Online (Skype) hours

Resources

Bibliography

Empirical properties of asset returns: stylized facts and statistical issues - Cont (2001) An Introduction to Statistical Learning - James, Witten, Hastie, Tibshirani (2013) Analysis of Financial Times Series - Tsay (2005) Financial Applications of Random Matrix Theory: A short review - Potters and Bouchaud (2009) Python for Finance: Mastering Data-Driven Finance - Hilpisch (2019)

Ressources en bibliothèque

- Financial applications of random matrix theory / Potters and Bouchaud
- Empirical properties of asset returns: stylized facts and statistical issues / Cont
- Python for Finance / Hilpisch
- Analysis of Financial Times Series / Tsay
- An Introduction to Statistical Learning / James, Witten, Hastie, Tibshirani

Moodle Link

• https://go.epfl.ch/FIN-525