M2 RECHERCHE EN ECONOMÉTRIE
M2 RESEARCH IN ECONOMETRICS

S1 – *Advanced Econometrics* *1*
S1 – *Economics of Networks* *2*
S1 – *Machine Learning and Statistical Learning* *3*
S1 – *Advanced Macroeconomics* *4*
S1 – *Advanced Microeconomics* *5*
S1 – *Predictive methods* *7*
S1 – *Non-Parametric Methods in Econometrics* *9*
S1 – *Advanced Corporate Finance* *10*
S1 – *Stochastic Process and Application* *11*
S1 – *Research Questions* *12*
S1 – *Professionalisation Workshops* *13*
S1 – *Automatic Model Selection Methods* *14*
S2 – *Time Series* *15*
S2 – *Transition and Duration Models* *16*
S2 – *Interpretability and Causality in Machine Learning* *18*
S2 – *Research Methodology* *19*

Cours spécifiques à l'option Magistère Economie, Data Science et Finance

S1 – *Big Data Tools (MAG)* *20*
S1 – *Machine Learning and New Data (MAG)* *21*
S1 – *End-of-Studies Project (MAG)* *22*
S2 – *Topics in Data Science (MAG)* *23*
S2 – *Projects in Data Science (MAG)* *24*

Dernière mise à jour / last update: juin 2025.

Advanced Econometrics
Econométrie approfondie

COURSE LANGUAGE
English

TEACHER
Emmanuel FLACHAIRE – emmanuel.flachaire@univ-amu.fr

COURSE DESCRIPTION AND OBJECTIVES
The goal of this course is to present advanced methods in econometrics for distributional analysis, regression and classification models. The course will present theoretical foundations and underlying intuition of each method, as well as several empirical examples.

COURSE OUTLINE
1. Resampling Methods
   - Pseudo-random generator
   - Monte Carlo experiments
   - Bootstrap and permutation tests
2. Nonparametric Econometrics
   - Density estimation
   - Regression splines
   - Finite mixture models
3. Econometrics and Machine Learning
   - Philosophy and general principle
   - Resampling-based methods and algorithms
   - Misspecification detection

ORGANIZATION
Semester: S1
Teaching Hours: 24 h of lectures

BIBLIOGRAPHY AND TEXTBOOKS
Ahamada et Flachaire (2011) Non-Parametric Econometrics, Oxford University Press.
Efron et Hastie (2016) Computer Age Statistical Inference, Cambridge University Press.

Dernière mise à jour / last update: juin 2025.

Economics of Networks
Economie des réseaux

COURSE LANGUAGE
English

TEACHER
Romain FERRALI – romain.ferrali@univ-amu.fr

COURSE DESCRIPTION AND OBJECTIVES
This course introduces a fast-growing field of research: the economics of social networks, from both theoretical and empirical perspectives. The course covers three main themes: network description, network formation, and behavior on networks. The approach combines theoretical rigor with empirical applications.

COURSE OUTLINE
Describing networks (theory and empirics)
Network formation (theory and empirics)
Behaviour on networks (diffusion, network games, learning)

KEY PROFESSIONAL SKILLS UPON GRADUATION
Formally describe networks
Understand key questions in network economics
Solve standard theoretical network models
Identify statistical challenges in network analysis and appropriate techniques

ORGANIZATION
Semester: S1
Teaching Hours: 24 h of lectures
Examination Method: Final exam + Written essay

BIBLIOGRAPHY AND TEXTBOOKS
Jackson, Matthew O. 2008. *Social and Economic Networks*. Princeton University Press
Newman, Mark. 2018. *Networks*. Oxford University Press
Bramoullé, Yann; Galeotti, Andrea; Rogers, Brian (eds.). *The Oxford Handbook of the Economics of Networks*. Oxford University Press
Watts, Duncan. 2004. *Six Degrees: The Science of a Connected Age*. W. W. Norton & Company

MANDATORY PREREQUISITES
Basic mathematical tools
Microeconomics knowledge

RECOMMENDED PREREQUISITES
Familiarity with formal economic models
Basic statistics

KEYWORDS
Social networks, network economics, graph theory, diffusion, learning

Dernière mise à jour / last update: septembre 2024.

Machine Learning and Statistical Learning
Machine learning et statistical learning

COURSE LANGUAGE
English

TEACHER
Pierre MICHEL – pierre.michel@univ-amu.fr

COURSE DESCRIPTION AND OBJECTIVES
This course introduces the fundamental concepts of machine learning and statistical learning. Its primary goal is to facilitate a comprehensive understanding of how machine learning algorithms operate. Students will explore a variety of widely used methods for supervised tasks. The course emphasizes practical implementation by providing hands-on experience with both synthetic and real-world datasets.

COURSE OUTLINE
1. Introduction
2. Estimation of the Parameters: an Optimization Problem
3. Regression Tasks
4. Classification Tasks
5. Explainable Machine Learning
6. Machine Learning, Ethics, and Fairness

KEY PROFESSIONAL SKILLS UPON GRADUATION
Understanding how machine learning algorithms work
Understanding the basic theory underlying machine learning
Being able to code (in Python or R – choice is up to the students) simple machine learning algorithms.

ORGANIZATION
Semester: S1
Teaching Hours: 24 h of lectures, sessions alternating theoretical presentations and applications.
Comment: The applications will be carried out on personal computers.

BIBLIOGRAPHY AND TEXTBOOKS
Berk, R. A. (2016). Statistical Learning from a Regression Perspective. Springer Texts in Statistics. doi:10.1007/978-3-319-44048-4
Charpentier, A. (2024). Insurance, Biases, Discrimination and Fairness. Springer
James, G, Witte, D., Hastie, T., Tibshirani, R. (2023). An Introduction to Statistical Learning with Applications in R. Second Edition. Springer
Murphy, K. (2012). Machine Learning: A Probabilistic Perspective. The MIT Press. ISBN: ISBN: 9780262018029

Dernière mise à jour / last update: juin 2025.

Advanced Macroeconomics
Macroéconomie approfondie

COURSE LANGUAGE
English

TEACHER
Cécilia GARCIA PEÑALOSA – cecilia.garcia-penalosa@univ-amu.fr

COURSE DESCRIPTION AND OBJECTIVES
The aim of the course is to explore the branch of macroeconomics concerned with economic growth and development. The course seeks to make students acquire two types of skills. First, we will examine in detail the core models and learn to solve macro-models that address development questions. Second, the course will attempt to answer a number of questions on growth by discussing the literature that addresses a particular question.
The course consists of two sections. The first part will examine the seminal work in the field. We will study the various mechanisms that will result in sustained long-run growth -learning-by-doing, investments in infrastructure, education, and firms' R&D decisions– and analyse the role played by externalities and increasing returns to scale. We will see that a crucial implication of these growth models is that the equilibrium growth rate is not socially optimal, and that a laissez-faire economy can grow either too slowly or too fast. We will also examine the causes of economic development, and why poverty traps may emerge. The explanations proposed include the theory of "the big push", whereby increasing returns to scale can result in poverty traps and hence explain why certain economies remain underdeveloped. We will then address the role of "threshold effects" in education, their implications for development, and the importance of the distribution of wealth.
The second part of the course will examine several topics to provide an overview of the literature covering them.
There will also be three problem sessions.

COURSE OUTLINE
Part I
- Towards endogenous growth
- Poverty Traps
- The Big Push
- Threshold Effects
- Income Distribution and Macroeconomics
- Technological change
- Expanding product variety
- Quality ladders
- General purpose technologies

Part II
- The new growth evidence
- Trade and Growth
- Institutions and development
- Competition and Growth

KEY PROFESSIONAL SKILLS UPON GRADUATION
Develop modelling skills in macroeconomics, understanding the causes of long-term growth and of why some countries remain in poverty traps.

BIBLIOGRAPHY AND TEXTBOOKS
Barro, R. and X. Sala-i-Martín, *Economic Growth*, 2004.
Aghion, P. and P. Howitt, *Endogenous Growth Theory*, MIT Press 1998.
Aghion, P. and P. Howitt, *The Economics of Growth* , MIT Press 2008.
Aghion, P. and S. Durlauf (eds.) *Handbook of Economic Growth,* North Holland 2005.

ORGANIZATION
Semester: S1
Teaching Hours: 24 h of lectures
Comment: Lecture notes and other material are available on AMeTICE.
Examination Method: Problem sets + Final written exam

Dernière mise à jour / last update: juin 2025.

Advanced Microeconomics
Microéconomie approfondie

COURSE LANGUAGE
English

TEACHER
Renaud BOURLES – renaud.bourles@centrale-med.fr
Frédéric DEROIAN – frederic.deroian@univ-amu.fr

COURSE DESCRIPTION AND OBJECTIVES
The theory of incentives primarily addresses situations of asymmetric (or decentralized) information in economic interactions, particularly when the objectives of the parties involved are conflicting. It provides tools to tackle regulatory issues typically absent from general equilibrium models and to analyse the inner workings of firms in greater detail. Examples include the provision of public goods (when the government lacks complete information about preferences) and task delegation, whether by a manager to a worker or by a government to a firm managing a natural monopoly. In such cases, unobserved actions or private information about costs or valuations cause deviations from classical models, often invalidating welfare theorems and preventing efficiency. This inefficiency arises from the need to offer "informational rents" to the party holding the private information.

Other prominent applications of incentive theory include optimal taxation, price discrimination, auctions, and insurance. Asymmetric information typically falls into two categories: hidden information and hidden action. In the first case, one party has incomplete knowledge about certain characteristics of the relationship, such as production costs, consumer willingness to pay, or risk levels. These scenarios are modelled as adverse selection, where failing to distinguish between different "types" often leaves the uninformed party dealing with the least desirable ones. In the case of hidden action, the asymmetry concerns a choice—commonly referred to as effort—made by the informed party. Here, the challenge is to design incentives that align the effort of the informed party with the objectives of the uninformed party, a situation known as moral hazard.

In this course, we focus on the Principal-Agent framework, which assumes (i) two parties: one informed and one uninformed, and (ii) that the Principal makes a take-it-or-leave-it offer to the Agent. We abstract away issues of bargaining (game-theoretic considerations) and contract enforceability, assuming all agreements are binding. The course covers the basic models of adverse selection and moral hazard, their main applications, and key extensions. It concludes with an exploration of the limits of incentive theory, including countervailing incentives and behavioural considerations.

COURSE OUTLINE
I. Hidden information: screening and signalling
  1. A classic example: recall on second degree price discrimination
  2. Mechanism design and revelation principle
  3. A more general model of adverse selection
  4. Applications and extensions
    i. Credit rationing
    ii. Regulation of natural monopolies
    iii. Delegation and audit
  5. Signalling models
    iv. The basic problem: market for lemons
    v. Education as a Signal
    vi. Application to corporate finance
  6. Dynamic aspects: renegotiation and commitment

II. Hidden action: The issue of moral hazard
  1. Core model - basic insights
    i. First and second best
    ii. Risk neutrality
    iii. Limited liability
  2. Extending the core-model
    i. more than two outcomes
    ii. more than two levels of effort
    iii. a kick look at continuous outcomes
  3. Applications
    i. Financial contracts
    ii. Sharecropping
    iii. Efficiency wages

iv. Insurance markets

4. Extensions
   i. Does payment scheme destroy intrinsic motivation?
   ii. Relational contracts

BIBLIOGRAPHY AND TEXTBOOKS

Bolton, P. and Dewatripont, M., *Contract Theory*, MIT Press.
Laffont, J.-J. and D. Martimort, D., *The Theory of Incentives -- The Principal-Agent Model*, Princeton University Press.
Salanié, B., *The Economics of Contracts: A Primer*, MIT Press.

ORGANIZATION

Semester: S1
Teaching Hours: 24 h of lectures
Comment: 12 h with R. Bourlès Bourlès (http://renaud.bourles.perso.centrale-marseille.fr) and 12 h with F. Deroïan (https://sites.google.com/view/fredericderoian). Handouts are available on AMeTICE.
Examination Method: Mid-term exam (1/3 of the grade): comment a paper related to the empirics of incentives + Final written exam (2/3 of the grade)

Dernière mise à jour / last update: juin 2025.

Predictive methods
Méthodes de prévisions

COURSE LANGUAGE
English

TEACHER
Christophe MULLER – christophe.muller@univ-amu.fr

COURSE DESCRIPTION AND OBJECTIVES
The main objective of this course is to provide the students with a synthetic framework so that they can theoretically understand and efficiently apply the main forecast and prediction techniques for most applied econometric problems. An important aim is to make students able to think by themselves when facing an applied econometric problem to solve.

COURSE OUTLINE
I. Introduction to Economic Predictions
- Uncertainty and Information
- Cross Sections, Time Series and Panel Data
- Notions of Prediction and Forecast
- Reminders of probability, algebra and optimisation

II. Expectations and Regressions
- Reminders of linear regression
- Consistency and Asymptotic Normality
- Simulators

III. Predictions
- Statistical decision theory
- Linear Predictors
- Evaluating Predictive Power
- Identification
- Nonlinear predictions: example of the nonlinear Tobit model

IV. Dynamic Forecasts without Time Series Models
- Sequences
- Time operators, moving averages and autocorrelations
- Extrapolation and smoothing
- Seasonality correction and trends

V. Dynamic Forecasts with Univariate Time Series Models
- Forecasts with stationary ARIMA models
- Evaluating forecasts

VI. Dynamic Forecasts with Multivariate Time Series Models
- VAR models
- Forecasting under non-stationarity
- Panel data

VII. Poll Surveys and Experts
- Inference for statistical surveys based on random samples
- Quotas and redressing
- Expert advice
- Structural models

VIII. Predictions with Big Data and Machine Learning
- Machine learning for predictions
- Variable selection and LASSO
- Neural networks
- Artificial intelligence

IX. Empirical Applications
- Empirical examples
- The command forecast in Stata
- The forecasting package in R
- Poll internet survey
- Treatment effect and external validity
- Volatility forecast
- Use of prediction statistics by stakeholders

KEY PROFESSIONAL SKILLS UPON GRADUATION

The targeted skills are an understanding of the reasonings made when using econometric techniques. At the end of the course, the students should be able to make their own informed methodological decisions in empirical applications, notably for prediction and forecasts.

ORGANIZATION

Semester: S1
Teaching Hours: 24 h of lectures
Examination Method: Report of prediction analysis

BIBLIOGRAPHY AND TEXTBOOKS

Elliott, G. and A. Timmerman, Economic Forecasting, Journal of Economic Literature, February 2008.
Ghysels, E. and M. Marcellino, Applied Economic Forecasting Using Time Series Methods, Oxford University Press, 2018.
Gouriéroux, C. and A. Monfort, Time Series Econometrics, Cambridge University Press, 1996.
Hyndman, R.J., Forecasting: Principles and Practice, 2024.
Hamilton, J.D., Time Series Analysis, Princeton University Press, 1994.
Montgomery, D.C., M. Kulahci and C.L. Jennings, Time Series Forecasting, 2024IIU
Stata Manual, 2024
Wooldridge, J.W., Introduction Econometrics: A Modern Approach, South-Western College Publishing, 2012.

MANDATORY PREREQUISITES

Linear econometrics, basic mathematics of the optimisation, linear algebra, Stata software, basic time series econometrics.

KEYWORDS

Econometrics, Identification, Inference, Estimation & Prediction, Forecasts, Econometric models, Big Data.

Dernière mise à jour / last update: septembre 2024.

Non-Parametric Methods in Econometrics
Méthodes non paramétriques en économétrie

COURSE LANGUAGE
English

TEACHER
TBA

COURSE DESCRIPTION AND OBJECTIVES
Non-Parametric methods are statistical techniques that do not require to specify functional forms for objects being estimated. Instead, they let the data itself play and inform the resulting model in a particular manner. Such methods are becoming increasingly popular for applied data analysis; they are best suited to situations involving large data sets for which the number of variables involved is manageable. These methods are often deployed after common parametric specifications are found to be unsuitable for the problem at hand, particularly when formal rejection of a parametric model based on specification tests yields no clues as to the direction in which to search for an improved parametric model.
The job market understood the importance of the non/semi-parametric methods and almost any serious software contains the principal techniques in this area. We illustrate the different models and techniques with R and Matlab. First, R because of the huge number of packages from CRAN, and secondly Matlab because it is the easiest environment for programming arrays in econometrics (and typically all objects are arrays in applied econometrics). Both are very representative of the job market. Each lecture will be accompanied by numerical examples and small programming tutorials.

KEY PROFESSIONAL SKILLS UPON GRADUATION
To master the concepts specific to nonparametric methods for modelling the conditional expectation and the conditional variance. Estimation and forecasting in i.i.d and non i.i.d. simulated or real data sets.
Be able to use the nonparametric estimation for driving more structural models.
To use the basic packages from R and the toolboxes from Matlab for nonparametric and semiparametric modelling.

ORGANIZATION
Semester: S1
Teaching Hours: 24 h of lectures

BIBLIOGRAPHY AND TEXTBOOKS
Given the wide content of the course, there is no single textbook for this course, therefore the lecture handouts are self-contained. However, the content is covered in some classic textbooks, such as:

Jeff RACINE: “Nonparametric Econometrics, A Primer”, Foundations and Trends in Econometrics, Vol. 3, No 1 (2008).
Adrian PAGAN and Aman ULLAH, “Nonparametric Econometrics”, Cambrdige University Press, (1999).
B. P. SILVERMAN: “Density Estimation for Statistics and Data Analysis”, Chapman Hall, (1986).
Wolfgang HÄRDLE: “Applied Nonparametric Regression”, Cambridge University Press; Revised edition (1992).
Denis BOSQ: “Nonparametric Statistics for Stochastic Processes: Estimation And Prediction”, Springer-Verlag New York Inc, 2nd ed. (2013).
Ziyue LIU and Wensheng GUO: “Data Driven Adaptive Spline Smoothing”, Statistica Sinica 20 (2010).

Dernière mise à jour / last update: septembre 2024.

Advanced Corporate Finance
Finance d'entreprise approfondie

COURSE LANGUAGE
English

TEACHER
TBA

COURSE DESCRIPTION AND OBJECTIVES
The aim of this course is to provide the fundamentals of empirical research in corporate finance, so that students can move into the full range of research areas within corporate finance.

PLAN DU COURS / COURSE OUTLINE
The course will cover the following topics:
- tools and databases used in empirical corporate finance
- research in corporate finance (classical topics such as : financing and payout policies, M&A and takeover defenses, corporate ownership and governance structures)
- research in entrepreneurial finance and private equity
- research in behavioral corporate finance

ORGANIZATION
Semester: S1
Teaching Hours: 24 h of lectures

BIBLIOGRAPHY AND TEXTBOOKS
Denis, D. J. (2024). Handbook of Corporate Finance. Edward Elgar Publishing Ltd.
Eckbo, B. E. (2007). Handbook Of Corporate Finance. North-Holland.
Eckbo, B. E., Phillips, G. M., & Sorensen, M. (2023). Handbook of the Economics of Corporate Finance: Private Equity and Entrepreneurial Finance. North-Holland.
Malmendier, U. (2018). Behavioral Corporate Finance✶. In B. D. Bernheim, S. DellaVigna, & D. Laibson (Eds.), Handbook of Behavioral Economics: Applications and Foundations 1 (Vol. 1, pp. 277–379). North-Holland. https://doi.org/10.1016/bs.hesbe.2018.08.001

Dernière mise à jour / last update: septembre 2024.

Stochastic Process and Application
Processus Stochastique et Application

COURSE LANGUAGE
English

TEACHER
Mathieu FAURE – mathieu.faure@univ-amu.fr

COURSE DESCRIPTION AND OBJECTIVES
The course aims at presenting stochastic processes and their application to economics and finance.

COURSE OUTLINE
1. Introduction: conditional expectation and integral theory
2. Stochastic processes (incl. martingales)
3. Discrete martingales and applications (incl. Cox-Ross-Rubinstein model)
4. Continuous martingales and Brownian motion
5. Application to finance: the Black and Scholes model
6. Application to economics: dynamics of natural resources

KEY PROFESSIONAL SKILLS UPON GRADUATION
Understand the definition and basic properties of stochastic processes
Know how to apply them to model economic and financial phenomena.

ORGANIZATION
Semester: S1
Teaching Hours: 24 h of lectures

BIBLIOGRAPHY AND TEXTBOOKS
Steven Shreve, Stochastic Calculus for Finance I: The Binomial Asset Pricing Model, 2005.
Steven Shreve, Stochastic Calculus for Finance II: Continuous-Time Models , 2005.
Steven Roman, Introduction to the Mathematics of Finance: From Risk Management to Options Pricing, 2004.

Dernière mise à jour / last update: septembre 2024.

Research Questions
Questions de recherche

COURSE LANGUAGE

English

TEACHER

Practitioners

COURSE DESCRIPTION AND OBJECTIVES

This course is organized as a series of 12 lectures of 2 hours each during which faculty members will introduce students to a series of modern research topics.

ORGANIZATION

Semester: S1
Teaching Hours: 24 h of lectures

Dernière mise à jour / last update: juin 2025.

Professionalisation Workshops
Ateliers de professionnalisation

COURSE LANGUAGE

English

COURSE DESCRIPTION AND OBJECTIVES

This workshop is designed to guide students in their transition from academic training to the job market. **Participation in all activities is mandatory.**

It combines several complementary components:

- **Afterworks** (on campus or online), where companies and institutions introduce themselves to students, share insights into their missions, and discuss opportunities for collaboration.
- A **Career Day**, organized in two parts: first, recent graduates present their career paths, current positions, and how their training helped them enter the job market; second, a large recruitment fair brings together around 50 local, national, and international companies and institutions to offer internships and job opportunities.
- A course entitled *"Building a Strong Application"*, providing practical tools and strategies for professional integration.
  It is divided into two parts:
  - First part (lecture): Preparing for interviews (best practices, preparation methods, and self-presentation); searching for an internship or a job abroad (application strategies, networks, and resources); negotiating salaries (key principles for successful negotiation).
  - Second part (workshops): Small-group sessions offered to M2 students, focusing on CV writing and mock interview practice.

Together, these activities give students concrete experience, direct contact with employers, and essential skills to confidently approach their future careers.

KEY PROFESSIONAL SKILLS UPON GRADUATION

By the end of the workshop, students will possess the essential skills to approach the job market with confidence. They will know how to present themselves effectively, understand recruiters' expectations in France and abroad, and activate a professional network. Through lectures and practical workshops, they will be able to prepare strong applications, succeed in interviews, and conduct salary negotiations with assurance.

ORGANIZATION

Semester: S1

Teaching Hours: 10 h of tutorials

Dernière mise à jour / last update: juin 2025.

Automatic Model Selection Methods
Modèles de réduction de l'information

COURSE LANGUAGE
English

TEACHER
Sullivan HUE – sullivan.hue@univ-amu.fr

COURSE DESCRIPTION AND OBJECTIVES
The objective of this course is to introduce quantitative methods allowing to reduce information. These methods cover different fields of statistics and are based on classical econometric methods (OLS, MLE) or classificatory or principal component methods. The goal is to study methods to do automatic variable selection in large-scale problems and to apply them to real data.

COURSE OUTLINE
- Classification methods
- Statistical factor models
- Lasso methods
- The so-called « General to Specific » method (Hendry, Gets or Autometrics Methodology)

KEY PROFESSIONAL SKILLS UPON GRADUATION
Understanding new methods
Application on real data
Learning new tools or econometric softwares dedicated to the reduction of information

ORGANIZATION
Semester: S1
Teaching Hours: 24 h of lectures

BIBLIOGRAPHY AND TEXTBOOKS
Doornik, J.A. and Hendry, D.F. (2015). Statistical model selection with "Big Data", Cogent Economics & Finance, vol 3, n°1, 1-15.
Hendry, D.F. and Doornik, J.A. (2014). Empirical Model Discovery and Theory Evaluation. Automatic Selection Methods in Econometrics. The MIT Press.
Richard A. Johnson and Dean W. Wichern, Applied Multivariate Statistical Analysis, Pearson.

Dernière mise à jour / last update: juin 2025.

Time Series
Séries temporelles

COURSE LANGUAGE
English

TEACHER
Gilles DUFRENOT – gilles.dufrenot@univ-amu.fr

COURSE DESCRIPTION AND OBJECTIVES
The objective of this course is to provide the students with some tools that are necessary for those who will be confronted in their future careers with functions requiring knowledge in quantitative finance and quantitative economics analysis. The course includes theoretical and applied aspects using R software.

COURSE OUTLINE
1.- Extreme distributions and copula analysis
2.- Regime-switching non-linear models: Markov-switching and STAR models
3.- Long-memory models

KEY PROFESSIONAL SKILLS UPON GRADUATION
Acquire complex tools for analyzing time series
Learn how to use the R programming language to estimate nonlinear models
Learn how to identify unusual behavior in series: regime shifts, hysteresis, extreme events.

ORGANIZATION
Semester: S1
Teaching Hours: 24
Examination Method: Final Exam + Work in group

BIBLIOGRAPHY AND TEXTBOOKS
G. Dufrénot and T. Matsuki (eds), 2021, Recent Econometric Techniques for Macroeconomics and Financial Data, Springer Verlag (see the chapter written by Aditi Chaubal).
Elements of copula Modelling with R (https://copula.r-forge.r-project.org/book/)
P. Robinson (ed), 2023, Time series with long-memory, Oxford University Press (https://academic.oup.com/book/51958)

MANDATORY PREREQUISITES
Students must have basic knowledge of time series: dynamic modelling, ARIMA models, stochastic and deterministic stationarity. In addition, knowledge of programming in R is required. Familiarity with elementary probability distributions and laws is also essential.

RECOMMENDED PREREQUISITES
Knowledge of spectral analysis is recommended for analysing time series in the frequency domain (spectral density, power spectrum, spectral windows).

KEYWORDS
Markov-switching, Copula, ARFIMA, Long-Memory, nonlinearity

Dernière mise à jour / last update: juin 2025.

Transition and Duration Models
Modèles de transitions et de durées

COURSE LANGUAGE
English

TEACHER
Christian SCHLUTER – christian.schluter@univ-amu.fr

COURSE DESCRIPTION AND OBJECTIVES
Students will study models of transitions and durations and learn how to estimate these using real-world data.
This course is an introduction to modelling transitions into a state of interest (such as the transition into employment from unemployment) and durations (such as unemployment, survival of patients after medical treatment or firms after a financial crash, time-to-default of loans, or time-to-purchase, criminal recidivism). Time-to-event or survival analysis are alternative labels. We start with the basic building blocks (Poisson processes, Markovian transitions, Markov chains, hazard models). Since duration data might be censored (individuals might still be in the state of interest at the end of the observation window), classic ordinary least squares (OLS) is invalid, and we develop appropriate methods for estimation. Unobserved heterogeneity introduces fundamental identification challenges (duration dependence v. dynamic sorting) that are discussed in detail. Finally, we consider how recent machine learning methods have been adapted for such censored duration data (such as Random Survival Forests).
Throughout all methods will be illustrated using examples in R and python, and we will replicate several papers from the established empirical literature. Several exercise sets will help students deepen their understanding of the theory.

COURSE OUTLINE
(I) Introduction to Poisson and counting processes
- Counting processes and the Poisson process
- Exponentially distributed inter-arrival times / durations
- The Poisson process and Markov chains
- Poisson regressions
- The Piece-Wise Exponential (PWE) model and estimation using a GLM

(II) Introduction to Markov processes
- Transitions, and the Chapman-Kolmogorov equation
- Classification of states and first passage or hitting times
- The invariant distribution
    - Markov's (or the ergodic) theorem
- Examples in theory and practice
    - State-space modelling: MC approximation to an AR(1) process (Rouwenhorst method)
    - State-space modelling using the Poisson process.
    - Unemployment transitions
    - Google's PageRank

(III) Duration and survival analysis: Hazard models
- Survival functions: The Kaplan-Meier estimator, the log-rank test
- Hazards, and the Proportional Hazard (PH) model
    - Maximum likelihood estimation (flow and stock samples)
- The Mixed Proportional Hazard (MPH) model, identification challenges
- The PH model and grouped data
- Cox's Partial Likelihood
- Machine Learning and Survival Analysis
    - Training a PH model
    - Random Survival Forests

KEY PROFESSIONAL SKILLS UPON GRADUATION
The students will master the theory of transition and duration models, learn how to estimate such models in practice using real-world data, and understand and address the empirical challenges that arise in empirical work.

ORGANIZATION
Semester: S2
Teaching Hours: 24 h of lectures
Examination Method: Research Project + Exam.

BIBLIOGRAPHY AND TEXTBOOKS

G. James et al. (2021): An Introduction to Statistical Learning, Chapter 11.
Wooldridge (2002), Chapters 19 and 20,
van den Berg, Chapter 55, Handbook of Econometrics.

MANDATORY PREREQUISITES

Basic econometrics.

KEYWORDS

Transition models, duration models, Poisson process, hazard and survival functions.

Dernière mise à jour / last update: juin 2025.

Interpretability and Causality in Machine Learning
Interprétabilité et causalité en Machine learning

COURSE LANGUAGE
English

TEACHER
Emmanuel FLACHAIRE – emmanuel.flachaire@univ-amu.fr
Sullivan HUE – sullivan.hue@univ-amu.fr

COURSE DESCRIPTION AND OBJECTIVES
This course is divided into two parts: (1) understanding why machine learning algorithms are not interpretable and how to use interpretable methods to interpret black boxes; (2) introducing recent advances in causal machine learning.

COURSE OUTLINE
1. Interpretability
   - Interpretable models
   - Feature effects methods: PDP, ICE and ALE
   - Surrogate models
   - Shapley Values based methods
   - IML limits
   - Inherently interpretable models with high performance
2. Causality
   - Directed Acyclic Graphs (DAG) and potential outcomes approaches
   - Regression and double orthogonalization
   - IPW and AIPW estimators
   - Post-Lasso and Post-double Lasso
   - Double/debiased machine learning
   - Generic machine learning for CATE

ORGANIZATION
Semester: S2
Teaching Hours: 24 h of lectures

BIBLIOGRAPHY AND TEXTBOOKS
Molnar (2025) Interpretable machine learning
Chernozhukov et al (2024) Causal ML book
Facure (2023) Causal inference for the brave and true
GHaillac and L'Hour (2025) Machine learning for econometrics

Dernière mise à jour / last update: juin 2025.

Research Methodology
Méthodologie de la recherche

COURSE LANGUAGE
English

TEACHER
Christian SCHLUTER – christian.schluter@univ-amu.fr
Miriam TESCHL – Miriam.teschl@ehess.fr

COURSE DESCRIPTION AND OBJECTIVES
The first part of the course is made of lectures devoted to the epistemology of economics and to the description of today's organization of academic economics. This course discusses methodological and epistemological questions concerning the practice and understanding of economics. The aim is to have a better understanding of how we see, understand and intervene is this world as economists.
The second set of lectures consists of three parts: research design, which will focus on good research design, and ways to credibly communicate the validity of the chosen approach; how to write a referee report, with the objective of the referee report being to *evaluate* the contributions of the submitted paper, ending in a judgement of whether the paper could eventually be published in the target journal, it involves evaluating the relevance as well as the quality of the scientific content and of the writing; academic writing.

KEY PROFESSIONAL SKILLS UPON GRADUATION
Capacity to understand research papers, capacity to construct a research proposal, key communication skills, critical thinking.

ORGANIZATION
Semester: S2
Teaching Hours: 24 h of lectures

BIBLIOGRAPHY AND TEXTBOOKS
A.F. Chalmers, 2013, What is this thing called science?, Hackett Publishing Company
Milton Friedman, 1953 The Methodoloy of Positive Economics, reprinted in Hausman, Daniel (ed.), 2008, The Philosophy of Economics: An Anthology, Cambridge University Press, pp. 154 - 187.
Martin Hollis and Edward Nell, 1975, Rational Economic Man: A philosophical critique of neoclassical economics, Cambridge University Press
Thomas Kuhn, 1996 The Structure of Scientific Revolution, University of Chicago Press

KEYWORDS
Academic writing, communication skills, research design, methodology, epistemology, science, model, falsificationism

Cours spécifique Magistère
Dernière mise à jour / last update: juin 2025.

Big Data Tools (MAG)
Outils des Big Data (MAG)

COURSE LANGUAGE

English in Marseille

TEACHER

Hervé MIGNOT – practitioner from Equancy

COURSE OUTLINE

1. Hadoop. HDFS. MapReduce. Stockage et calculs distribués. Déploiement d'un cluster.
2. Préparation, stockage et traitement des big data : Pandas, Hive and Pig
3. Data visualisation avec matplotlib & seaborn
4. Alternatives : solutions propriétaires, bases NoSQL, ElasticSearch

ORGANIZATION

Semester: S1
Teaching Hours: 24 h of lectures
Comment: Class exclusive for Magistere students.

Cours spécifique Magistère
Dernière mise à jour / last update: juin 2025.

Machine Learning and New Data (MAG)
Machine learning et nouvelles données (MAG)

COURSE LANGUAGE
English in Marseille

TEACHER
Quentin LIPPMANN – quentin.lippmann@univ-amu.fr

COURSE DESCRIPTION AND OBJECTIVES
This course proposes to study the processing and analysis of unstructured data, and more specifically textual data and images.

COURSE OUTLINE
This course is divided in two parts of 12 hours each. The first part covers text as data. The second is about image as data.

Part 1 – Text as Data

Block 1 – Foundations of NLP

Students will learn about the complete document-pre-processing pipeline, beginning with tokenisation and the construction of n-grams. They will create Bag-of-Words representations and apply TF–IDF weighting to highlight discriminative terms. We will then move to distributed word representations through word embeddings, extract entities with Named-Entity Recognition, and analyse sentence structure by performing dependency parsing.

Block 2 – Large Language Models

Students will learn about the transformer architecture and its self-attention mechanism, compare pre-training with fine-tuning, and experiment with in-context learning. They will study Reinforcement Learning from Human Feedback as an alignment method and practice prompt-engineering patterns to steer model behaviour. We will tackle hallucination and explore retrieval-augmented generation as a mitigation strategy.

Part 2 – Image as Data

Block 1 – Image Fundamentals

Students will learn about digital image representation and colour spaces, then examine convolution operations—kernel size, stride, padding—and their effect on the receptive field. They will study activation and pooling layers for feature extraction and understand bounding-box regression. Anchor-based and anchor-free object-detection strategies will be compared.

Block 2 – Facial Analysis, Segmentation & Generative AI

Students will learn about classical Haar cascades versus modern RetinaFace detectors for face localisation. They will map facial landmarks, build embedding-based recognition pipelines, and evaluate systems using FAR, FRR, ROC curves, and demographic-bias checks. Promptable segmentation models will be introduced, followed by diffusion-based generative models for image synthesis.

All the concepts are applied and illustrated in Python applications.

KEY PROFESSIONAL SKILLS UPON GRADUATION
To learn how to process and analyse textual data
To learn how to process and analyse images

ORGANIZATION
Semester: S1
Teaching Hours: 24 h of lectures
Comment: Class exclusive for Magistere students.

Cours spécifique Magistère
Dernière mise à jour / last update: juin 2025.

End-of-Studies Project (MAG)
Projet de fin d'études (MAG)

COURSE LANGUAGE
French in Marseille

TEACHER
A teacher + a practitioner

COURSE DESCRIPTION AND OBJECTIVES
The end-of-studies project is carried out in collaboration with a company from October to March. This project enables students to carry out operational engineering work in data science and to compare theory with applications in the professional world.

KEY PROFESSIONAL SKILLS UPON GRADUATION
To be able to tackle a data science problem and write a report to answer it.
To know how to work as a team and to meet a set of specifications.

ORGANIZATION
Semester: S1
Comment: Class exclusive for Magistere students. Bimonthly meetings with supervisors, and autonomous work between meetings.
Examination Method: Project + Defense

Cours spécifique Magistère
Dernière mise à jour / last update: juin 2025.

Topics in Data Science (MAG)
Sujets en Data Science (MAG)

COURSE LANGUAGE
English in Marseille

TEACHER
Pierre MICHEL – pierre.michel@univ-amu.fr
Christophe HURLIN – practitioner

COURSE DESCRIPTION AND OBJECTIVES
This course aims to raise students' awareness of topical issues in data science.

COURSE OUTLINE
1. Conformal prediction
   a. Introduction and theoretical foundations
   b. Conformal prediction for regression
   c. Conformal prediction for classification
2. Algorithmic fairness
   a. Introduction to algorithmic fairness
   b. Framework for fairness in machine learning
   c. Measuring algorithmic fairness
   d. Testing for algorithmic fairness
   e. Mitigating algorithmic biases

KEY PROFESSIONAL SKILLS UPON GRADUATION
Understand how to make conformal prediction for regression and classification
Understand algorithmic fairness, and how to measure it, test it and mitigate its biases.

ORGANIZATION
Semester: S2
Teaching Hours: 24 h of lectures
Comment: Class exclusive for Magistere students.

Cours spécifique Magistère
Dernière mise à jour / last update: juin 2025.

Projects in Data Science (MAG)
Projets en Data Science (MAG)

COURSE LANGUAGE
English in Marseille

TEACHER
Pierre MICHEL – pierre.michel@univ-amu.fr
Christophe HURLIN – practitioner

COURSE DESCRIPTION AND OBJECTIVES
This course is complementary to the course of « Topics in data science ». The goal of this course is to make students work on projects related to the topics studied in the other course.

KEY PROFESSIONAL SKILLS UPON GRADUATION
To be able to tackle a data science problem and write a report to answer it.

ORGANIZATION
Semester: S2
Teaching Hours: 24 h of lectures
Comment: Class exclusive for Magistere students.