MA in Computational Social Science

Faculty Directors

- James Evans, Sociology
- Marc Berman, Psychology

Executive Committee

- Luc Anselin, Sociology
- Kathleen Cagney, Sociology
- Guanglei Hong, Comparative Human Development
- Ali Hortaçsu, Economics
- Leslie M. Kay, Psychology
- Howard Nusbaum, Psychology
- John Padgett, Political Science
- Stephen W. Raudenbush, Sociology

Affiliated Faculty

- Wilma Bainbridge, Psychology
- Stéphane Bonhomme, Economics
- Magne Mogstad, Economics
- James T. Sparrow, History
- Rochelle Terman, Political Science

Executive Director and Assistant Dean

- Chad Cyrenne

Associate Senior Instructional Professor and Associate Director

- Jean Clipperton

Assistant Senior Instructional Professor and Associate Director of Undergraduate Studies

- Jon Clindaniel

Senior Instructional Professor in Economics and Co-Director of MACSS-Econ

- Victor O. Lima

Assistant Senior Instructional Professor in Economics and Co-Director of MACSS-Econ

- Min Sok Lee

Assistant Instructional Professors

- Sabrina Nardin (MACSS)
- Zhao Wang (MACSS)

Teaching Fellow’s

- Pedro Arroyo
- Shilin Jia
- Sanja Miklin

Senior Director of Career Services and Professional Development

- Shelly Robinson

Employer Relations Manager

- Lauren Sheely

Director of Graduate Admissions and Enrollment

- Sparkle Dalphinis

Student Affairs Administrators
MA in Computational Social Science

- Sabrina Biggus (MACSS)
- Emma Rubenstein (MACSS-Econ)

Events Coordinator, MA Programs
- Jana Chiu
Budget and Programs Administrator
- Daniel Duhon

GENERAL INFORMATION

The Master of Arts in Computational Social Science (http://macss.uchicago.edu/) is a two-year program of graduate study. It has a structured curriculum, with a total of 18 required and elective courses tailored to the disciplinary track a student follows. Students submit an article-length MA thesis in their second year, modeled on a professional journal article.

The program aims to produce leading social scientists in each of the core social science fields—economics, sociology, political science, psychology, history, and anthropology—producing competitive PhD applicants, well-trained in computational approaches, who have mastered the research and analytical skills necessary to make important contributions. MACSS also offers a concentration in economics which is operated jointly with the Department of Economics. Students receive close mentorship from the program’s Faculty Directors, academic staff, and members of our Executive and Affiliated Faculty to design a customized curriculum, define an area of scholarly research, and write a MA Thesis.

They receive full professional support from the director of career services, with biweekly workshops, career planning, and employer recruitment.

Finally, all MA students may participate in an optional summer practicum between their first and second years, with internships drawn from academic and professional organizations. International students have three years of STEM work eligibility after they graduate.

PROGRAM REQUIREMENTS AND COURSE WORK

GENERAL MACSS CURRICULUM

In their first year, all MA students will complete:

- The Computational Math camp or the Econ Math camp and the Computing Fundamentals Boot Camp, which run for three weeks in August / September.
- A three-course core in Perspectives (Perspectives on Computational Analysis, Perspectives on Computational Modeling, and Perspectives on Computational Research).
- A two-course core in Computer Science with Social Science Applications in the Fall and Winter, and Large-Scale Computing for the Social Sciences in the Spring.
- Three graduate social science electives that will vary, depending on the student’s prior training and intended disciplinary path.
- Minor exceptions: (1) students who do not place out on our statistics exam will take an introductory course in computational statistics instead of a social science elective in the Fall quarter; (2) students may take more advanced programming instead of the two courses in Computer Science with Social Science Applications if they place out on a computer science placement exam.

In their second year, all MA students will complete:

- Three advanced courses in computational methods, drawn from any graduate department or professional school of the University.
- Three graduate social science electives.
- Three graduate courses that the student may select, from any UChicago department or professional school, where the student meets the minimum prerequisites.
- Students may register for the MA Research Commitment up to three times and use this as elective credit. This option allows students to allocate more time to their MA thesis.

Outside of their coursework, all MA students are expected to attend our weekly Computation Workshop, where advanced scholars and invited guests present drafts of their research for critique and discussion.

MACSS-ECON CURRICULUM

In their first year, all MACSS-Econ students will complete:

- Computational Math Camp (MACS 33000), which will run for three weeks in September.
- Computing Fundamentals Boot Camp (MACS 30120), which will also run for three weeks in September.
• A Statistics placement exam, offered during orientation week.
• A Computer Science placement exam, offered during orientation week.
• A two-course core in Perspectives on Computational Analysis and Computational Modeling, the first of which must be taken autumn quarter.
• A three-course core in Foundations of Quantitative Analysis. (Up to 2 courses may be waived via statistics placement exam)

In their second AND/OR first year, all MACSS-Econ students will complete:

• Econ Math Camp (ECON 30400) which will run for three weeks in September.
• A three-course core in Computer Science with Social Science Applications. (Up to 2 courses may be waived via computer science placement exam)
• Two courses to satisfy the Foundations in Economic Theory requirement.
• Two courses to satisfy the Foundations in Empirical Analysis/Computational Methods requirement.
• Six graduate economics/computational method electives that will vary, depending on the student’s prior training and intended disciplinary path. (This number may increase if student earns placement on CS or stats exams)
• Students may register once for the MA Research Commitment and use this as an elective credit. This option allows students to allocate more time to their MA thesis.

Outside of their coursework, all MA students are expected to attend our weekly Computation Workshop, where advanced scholars and invited guests present drafts of their research for critique and discussion.

ADMISSION

MACSS applicants must meet the formal requirements of the Graduate Social Sciences Division which are listed here (https://socialsciences.uchicago.edu/admissions/application-materials/).

MACSS offers tuition scholarships at the time of admission. Some financial need-based tuition scholarships may be available through an application process after prospective students are admitted.

Applicants to the Division of the Social Sciences and the University of Chicago who do not meet waiver criteria must submit proof of English language proficiency. Please contact our Dean of Students Office with any questions: ssd-admissions@uchicago.edu

HOW TO APPLY

The application with instructions and deadlines, is available online at: https://socialsciences.uchicago.edu/admissions/how-to-apply (https://socialsciences.uchicago.edu/admissions/how-to-apply/). Current students in the University of Chicago College may be eligible to apply through UChicago Advanced Scholars instead.

Any questions about MACSS can be directed to Sabrina Biggus our Student Affairs Administrator, at 773-702-8304 or sbiggus@uchicago.edu.

Please also visit our website: https://macss.uchicago.edu (https://macss.uchicago.edu/)

Computational Social Science Courses

MACS 30000. Perspectives on Computational Analysis. 100 Units.
Massive digital traces of human behavior and ubiquitous computation have both extended and altered classical social science inquiry. This course surveys successful social science applications of computational approaches to the representation of complex data, information visualization, and model construction and estimation. We will reexamine the scientific method in the social sciences in context of both theory development and testing, exploring how computation and digital data enables new answers to classic investigations, the posing of novel questions, and new ethical challenges and opportunities. Students will review fundamental research designs such as observational studies and experiments, statistical summaries, visualization of data, and how computational opportunities can enhance them. The focus of the course is on exploring the wide range of contemporary approaches to computational social science.

MACS 30005. Graduate Practicum in the Social Sciences. 100 Units.
Students interested in gaining additional applied experience or writing an alternative or public facing thesis should consider taking the Graduate Practicum in the Social Sciences course. The practicum is a hands-on experiential course designed to enable students to apply and expand their knowledge in a career pathway based on their unique interests and improve their technical and applied writing skills, all while providing a useful service to a Chicago-based community partner. During the course, students work individually or in small teams to research a career pathway and identify and address issues/needs faced by a local community partner of their choice. Each student or team receives guidance from their instructor and community partner throughout the project. The experience culminates in a final project (report and formal presentation to the community partner and/or class).
Equivalent Course(s): INRE 30005, MAPS 30005

MACS 30100. Perspectives on Computational Modeling. 100 Units.
This is a core-course for the MACSS program and it requires Python programming experience (for non-MACSS students, please email the instructor for consultation). This course will teach fundamental skills of applying statistical machine learning models in computational social science tasks. It will focus on understanding the strengths and weaknesses of modern machine learning algorithms as well as their applications in real-world tasks. Topics will include the key techniques in standard machine learning pipelines: data processing (e.g., data representation, feature selection), classification models (e.g., decision trees, logistic regression, naive bayes), regression models (e.g., linear regression), model evaluation (e.g., cross-validation, confusion matrix, precision, recall, and f1 for classification models; RMSE and Pearson correlation for regression models), and error analysis (e.g., data imbalance, bias-variance tradeoff, interpret model performance).

MACS 30111. Principles of Computing 1: Computational Thinking for Social Scientists. 100 Units.
This course is the first in a three-quarter sequence that teaches fundamentals of computational thinking to students in the social sciences. Lectures in the class will cover topics such as functions, data structures, as well as classes and objects. Assignments will give students the opportunity to practice these basic computing concepts using the Python programming language and get familiar with computational logic in real-world tasks.

MACS 30112. Principles of Computing 2: Data Management for Social Scientists. 100 Units.
This course is the second in a three-quarter sequence that teaches computational thinking and programming skills to students in the social sciences. Specifically, this course equips students with a fundamental toolkit for working with social science data. Students will learn the basics of web-scraping, relational databases, record linkage, data cleaning, modeling, visualization, and data structures. The programming language of the course is Python.

MACS 30113. Principles of Computing 3: Big Data and High Performance Computing for Social Scientists. 100 Units.
Computational social scientists increasingly need to grapple with data that is too big and code that is too resource intensive to run on a local machine. Using Python, students in this course will learn how to effectively scale their computational methods beyond their local machines—optimizing and parallelizing their code across clusters of CPUs and GPUs, both on-premises and in the cloud. The focus of the course will be on social scientific applications, such as: accelerating social simulations by several orders of magnitude, processing large amounts of social media data in real-time, and training machine learning models on economic datasets that are too large for an average laptop to handle.

MACS 30120. Computing Fundamentals Boot Camp. 000 Units.
This boot camp focuses on introducing fundamental open-source tools for producing reproducible, computational research. Topics include the basics of Python programming, working on the Linux command line, as well as using Git/GitHub for version control. The course assumes no prior exposure to these topics and serves as preparation for MACS 30121 “Computer Science with Social Science Applications 1.”

MACS 30121. Computer Science with Social Science Applications 1. 100 Units.
This course is the first in a three-quarter sequence that teaches computational thinking and essential skills to students in the social sciences. Lectures in the class will cover topics such as functions, data structures, classes and objects, as well as recursion. Assignments will give students the opportunity to practice these computing concepts using the Python programming language and get familiar with computational logic in real-world tasks. Topics will include the key techniques in standard machine learning pipelines: data processing (e.g., data imbalance, bias-variance tradeoff, interpret model performance).

MACS 30122. Computer Science with Social Science Applications 2. 100 Units.
This course is the second in a three-quarter sequence that teaches computational thinking and programming skills to students in the social sciences. Specifically, this course equips students with a fundamental toolkit for working with social science data. Students will learn the basics of web-scraping, relational databases, record linkage, data cleaning, modeling, visualization, and data structures. The programming language of the course is Python. This is an accelerated introductory course that is designed for advanced beginner programmers.

MACS 30123. Large-Scale Computing for the Social Sciences. 100 Units.
Computational social scientists increasingly need to grapple with data that is too big and code that is too resource intensive to run on a local machine. Using Python, students in this course will learn how to effectively scale their computational methods beyond their local machines—optimizing and parallelizing their code across clusters of CPUs and GPUs, both on-premises and in the cloud. The focus of the course will be on social scientific applications, such as: accelerating social simulations by several orders of magnitude, processing large amounts of social media data in real-time, and training machine learning models on economic datasets that are too large for an average laptop to handle.
Equivalent Course(s): PLSC 30123, MAPS 30123

MACS 30124. Computational Analysis of Social Processes. 100 Units.
How does the human social and cultural world develop and change? The focus of this course is on introducing computational methods for studying the evolution of phenomena over time, alongside relevant theories for
interpreting these processes from fields such as History, Anthropology, and Sociology. Students will gain hands-on experience using the Python programming language to harness a diverse set of digital data sources, ranging from satellite images to social media posts. Additionally, they will learn to employ computational approaches, such as simulation and dynamic topic modeling, to study social processes over a variety of different time scales: from the short term (changes in social media network structures over the course of the past week), to longer term (the evolution of English language discourse over the past 100 years), to deep time scales (long-term settlement pattern dynamics over the past 10,000 years).

Equivalent Course(s): MAPS 30124

MACS 30133. Machine Learning for Political Analysis. 100 Units.
This is an intermediate-to-advanced introduction to the mathematical and computational aspects of the core statistical and machine learning techniques. The goal is to equip students with a knowledge of the theoretical and practical aspects of four groups of machine learning methods which are widely used in applied research: (1) dimension reduction (PCA, MDS, and their extensions) (2) classification methods (SVM, Bayes classifiers, and other classification methods) (3) clustering procedures and density estimation (K-means, FMM, non- and semi-parametric Bayesian methods) (4) categorical data analysis (with brief introduction to probabilistic graphical models). The course includes applications in Political Science, such as FMM to estimate fraud in elections, PCA to construct indices to measure democracy, and text classification.

Equivalent Course(s): MAPS 30133

MACS 30150. Perspectives on Computational Modeling for Economics. 100 Units.
In this course students will learn several computational methodologies and tools to solve, simulate, and analyze models that are the backbone of current macroeconomic analysis. While learning the relevant computational methods is the main objective, the theoretical economic aspects of the model will be stressed and the students will be required to apply their economic knowledge and skills to interpret and analyze the results. We will examine non-stochastic and stochastic general equilibrium models, both under local and global approximations. The main part of the course will deal with representative agent models, but a significant part will be devoted to introducing students to the solution of heterogeneous agent models as well.

Equivalent Course(s): ECMA 31140

MACS 30200. Perspectives on Computational Research. 100 Units.
This course focuses on applying computational methods to conducting social scientific research through a student-developed research project. Students will identify a research question of their own interest that involves a direct reference to social scientific theory, use of data, and a significant computational component. The students will collect data, develop, apply, and interpret statistical learning models, and generate a fully reproducible research paper. We will identify how computational methods can be used throughout the research process, from data collection and hiding, to exploration, visualization and modeling, to the final communication of results. The course will include modules on theoretical and practical considerations, including topics such as epistemological questions about research design, writing and critiquing papers, and additional computational tools for analysis.

MACS 30250. Perspectives on Computational Research for Economics. 100 Units.
This course focuses on scaling up computational approaches to social science analysis and modeling with big data in context of opportunities afforded by high performance and cloud computing. We will begin by exploring various data structures encountered in social science research, how to deal with large or complex data storage and streaming data, and how to factor considerations of computational complexity into their analyses. We will also study social science applications of parallel computing, both on stand-alone machines and in supercomputing environments, to carry out complex computations. Students will learn to carry out parallel I/O and parallel computation on their own machines and on a cluster. We will also address API construction and access, and explore cloud configurations for social science research designs. We will also help students construct web-based outward facing data, analysis and visualization portals. Students will efficiently gather, structure, perform and present analysis on large-scale social science data. This course will be specifically tailored to students concentrating in Economics.

MACS 30300. Data Science Clinic I. 100 Units.
In order to enroll in this class, students must first submit an application and be matched with a project. Visit the Data Science Clinic site for application deadlines, how to apply, and information session details: bit.ly/ds-clinic. The Data Science Clinic partners with public interest organizations to leverage data science research and technology to address pressing social and environmental challenges. The Clinic also provides students with exposure to real-world projects and problems that transcend the conventional classroom experience including: working with imperfect datasets, applying models and algorithms to real-world data, navigating security and privacy issues, communicating results to a diverse set of stakeholders (e.g., industry, public interest, government agencies), and translating information into actionable insights, policy briefs and software prototypes. The Clinic is an experiential project-based course where students work in teams as data scientists with real-world clients under the supervision of instructors. Students will be tasked with producing key deliverables, such as data analysis, open source software, as well as final client presentations, and reports.

Equivalent Course(s): PPHA 30581, DATA 27100, CAPP 30300, MPCS 57300

MACS 30301. Introduction to Bayesian Statistics. 100 Units.
The goal of this course is to give students an overview of the theory and methods for data analyses using the Bayesian paradigm. Topics include: (1) foundations of Bayesian inference, (2) development of Bayesian
models and prior choices (3) analytical and simulation techniques for posterior estimation (4) model choice and
diagnostics (5) sensitivity analysis, (6) an introduction to Monte Carlo Markov Chain (MCMC) simulations (7)
intro to commonly used Bayesian estimation packages (R/JAGS/Bugs) (8) application of Bayesian analysis in real
world and Political Science problems.
Equivalent Course(s): MAPS 30301

MACS 30405. Exploring Cultural Space. 100 Units.
The class will put a special emphasis on the construct of space in cultural analysis. Spatial models have been
prevalently used in quantitative studies of culture and ideology, for instance, most famously in Pierre Bourdieu’s
analysis of French cultural fields. With the development of big data and machine learning, there has also been
bourgeoning advancement in its methodology. In the first five weeks of the lectures and discussions, we will
cover the foundational social theories and most commonly-used statistical/computational methods in the studies
of cultural space. We will ask and try to answer: what is a cultural space? What are its dimensions? What is
its topology? What social processes take place in it? Major statistical techniques, such as principal component
analysis, correspondence analysis, and latent class analysis as well as recent advances in computational text
analysis and neural-embedding models, will be introduced. The second half of the class will be devoted to
empirical studies and student projects. Some prior programming experience will be helpful but not required.
Undergraduate students are admitted with the consent of the instructor. Every student is expected to submit an
empirical study or extensive literature review at the end of the course.
Equivalent Course(s): MACS 20500, SOCI 20584

MACS 30500. Computing for the Social Sciences. 100 Units.
This is an applied course for social scientists with little-to-no programming experience who wish to harness
growing digital and computational resources. The focus of the course is on learning the basics of programming
and on generating reproducible research. Topics include coding concepts (e.g., data structures, control structures,
functions, etc.), data visualization, data wrangling and cleaning, version control software, exploratory data
analysis, etc. Students will leave the course with basic computational skills implemented through many methods
and approaches to social science; while students will not become expert programmers, they will gain the
knowledge of how to adapt and expand these skills as they are presented with new questions, methods, and
data. The course will be taught in R.
Equivalent Course(s): PLSC 30235, MAPS 30500, ENST 20550, PSYC 30510, SOCI 40176, MACS 20500, SOSC
26032, CHDV 30511, SOCI 20278

MACS 30501. Computational Anthropology. 100 Units.
This course exposes students to the methods and data of Computational Anthropology-the systematic,
computational study of the human species, past and present. Such methods have been essential in recent years
for simulating human behavior in different cultures and economic systems, uncovering ancient demographic
changes that still have an influence into the present day, preserving cultural heritage, and much more.
Anthropological data allows social science researchers to evaluate long term trends in the human condition,
across a variety of cultures, with a unique combination of material, textual, and structured database data.
Students will have the opportunity to evaluate state of the art approaches in computational anthropology and
learn how to apply these methods to their own social scientific research agendas using open anthropological
datasets and the Python programming language.

MACS 30519. Spatial Cluster Analysis. 100 Units.
This course provides an overview of methods to identify interesting patterns in geographic data, so-called spatial
clusters. Cluster concepts come in many different forms and can generally be differentiated between the search
for interesting locations and the grouping of similar locations. The first category consists of the identification of
extreme concentrations of locations (events), such as hot spots of crime events, and the location of geographical
concentrations of observations with similar values for one or more variables, such as areas with elevated disease
incidence. The second group consists of the combination of spatial observations into larger (aggregate) areas
such that internal similarity is maximized (regionalization). The methods covered come from the fields of spatial
statistics as well as machine learning (unsupervised learning) and operations research. Topics include point
pattern analysis, spatial scan statistics, local spatial autocorrelation, dimension reduction, as well as spatially
explicit hierarchical, agglomerative and density-based clustering. Applications range from criminology and
public health to politics and marketing. An important aspect of the course is the analysis of actual data sets by
means of open source software, such as GeoDa, R or Python.
Equivalent Course(s): GISC 20519, ENST 20519, SOCI 20519, GISC 30519, SOCI 30519

MACS 30550. From Data to Manuscript in R. 100 Units.
This course tackles the basic skills needed to build an integrated research report with the R programming
language. We will cover every step from data to manuscript including: Using R’s libraries to clean up and re-
format messy datasets, preparing data sets for analysis, running statistical tools, generating clear and attractive
figures and tables, and knitting those bits of code together with your manuscript writing. The result will be a
reproducible, open-science friendly report that you can easily update after finishing data collection or receiving
comments from readers. Never copy-paste your way through a table again! The R universe is large, so this course
will focus specifically on: The core R libraries, the tidyverse library, and R Markdown. Students will also learn
about the use of GitHub for version control.
Equivalent Course(s): MAPS 30550, PSYC 20550, CHDV 20550, CHDV 30550, PSYC 30550
MACS 30617. Organizational Analysis. 100 Units.
Organizations - NGOs, corporations, social movement organizations, governments, etc. - impact almost every aspect of social life; in addition, organizations have become some of the most significant actors in modern society. The course will provide a grounding in the sociological literature on how organizations function as well as the dynamics that govern both their internal structures and how they interface with society. We will cover rational, ecological, and resource-based approaches, as well as others. We will study organizations in local and global contexts, their role in economic production, their impact on members and non-members, as well as public policy. Throughout, we will engage questions pertaining to where organizations come from, how they function, when they 'succeed' and 'fail', as well as their social consequences. At the completion of the course, students will apply the concepts covered in class to a final project.
Equivalent Course(s): SOCI 20585, SOCI 30337, MACS 20617, PBPL 23002, MAPS 30617

MACS 31300. AI Applications in the Social Sciences. 100 Units.
Artificial Intelligence (AI) describes algorithms constructed to reason in uncertain environments. This course provides an introduction to AI applications in the social sciences. Driven by the rapid increase in accessible big data documenting social behavior, AI has been applied to: increase effective diagnosis and prediction under different conditions, improve our understanding of human interaction, and increase the effectiveness of data management in different social and human services. Random forests and neural networks are among the most frequent AI methods used for prediction, while natural language processing and computer vision contribute to understanding decision-making and improving service provision. We begin with careful consideration for what AI can achieve and where current limitations exist by looking at a variety of real-world applications. We will focus on three core sections: search, representation, and uncertainty. In each section, we will explore major approaches, representational techniques and core algorithms. We will examine the trade-offs between model structure and the algorithmic constraints that this structure implies. The course is driven by hands-on exercises with AI algorithms written in Python. At the end of the term, you should be able to apply and tweak these algorithms to accommodate your own data and research interests.

MACS 33000. Computational Math Camp. 000 Units.

MACS 33001. Mathematics and Statistics for Computational Social Science. 100 Units.
This course aims to provide students with a core understanding of mathematics and statistics for computational social science. Students who complete this course should be prepared to take more advanced computational methods courses. Completion of the Computational Math Camp in September is recommended, but not required.

MACS 33002. Introduction to Machine Learning. 100 Units.
This course requires Python programming experience. The course will train students to gain the fundamental skills of machine learning. It will cover knowledge and skills of running with computational research projects from a machine learning perspective, including the key techniques used in standard machine learning pipelines: data processing (e.g., data cleaning, feature selection, feature engineering), classification models (e.g., logistic regression, decision trees, naive bayes), regression models (e.g., linear regression, polynomial regression), parameter tuning (e.g., grid-search), model evaluation (e.g., cross-validation, confusion matrix, precision, recall, and F1 for classification models; RMSE and Pearson correlation for regression models), and error analysis (e.g., data imbalance, bias-variance tradeoff). Students will learn simple and efficient machine learning algorithms for predictive data analysis as well as gain hands-on experience by applying machine learning algorithms in social science tasks. The ultimate goal of this course is to prepare students with essential machine learning skills that are in demand both in research and industry.
Equivalent Course(s): MAPS 33002, PLSC 43505, MACS 23002

MACS 34000. Data Mining and Data Visualization for the Social Sciences. 100 Units.
This course introduces students to techniques for extracting and communicating knowledge from data. In the first half, students study visualizations as a method for summarizing information and reporting analysis and conclusions in a compelling format. This introduces the ideas and methods of data visualization, with emphasis on both why you are doing something as well as how to produce optimal visualizations. In the second half, students are introduced to the rapidly developing world of data mining. Focus will be on knowledge discovery and pattern recognition in the context of social science problem solving. From partitioning and anomaly detection to text clustering, high-dimensional mining, and deep learning, students will be given a thorough introduction to prominent techniques for exploring and discovering patterns in data. Throughout the course, class sessions will combine lecture, coding challenges, and computational problem solving to encourage wide engagement with the techniques using the R programming language.
Equivalent Course(s): MACS 24000

MACS 35000. MA Research Commitment. 100 Units.
Student Initiated research and writing for the MA research component.

MACS 35001. Structured MA Research Commitment. 100 Units.
Student initiated research and writing for the MA research component.

MACS 36000. Computational Methods Using Online Social Media Data. 100 Units.
This course will discuss a broad range of computational social science topics that leverage large-scale data from online communication platforms to gain insights into social issues. We will start from collecting and processing data from social media platforms (e.g., Twitter, IMDB, Airbnb, Yelp), and then introduce computational research
topics that include but are not limited to: sentiment analysis, deceptive marketing, recommendation system, fake news detection, spam detection, bot detection, demographic inference, public health, political attitude analysis, personality and behavior analysis, and cyberbullying. We will use version control techniques (e.g., git, Github) to keep track of the class projects. The ultimate goal of this course is to provide a broad introduction of computational social science research areas and train students to be familiar with the pipelines of doing computational research.

Equivalent Course(s): MAPS 36001, MACS 26000

MACS 36043. The Aesthetics of Artificial Intelligence. 100 Units.
With the emergence of generative AI tools such as ChatGPT, DALL-E, and Midjourney, the production of computer-generated content has become accessible to a wide range of users and use cases. Knowledge institutions are particularly challenged to find adequate responses to changing notions of authorship as the mainstreaming of "artificial" texts, audio-visual artifacts, and code is transforming our paradigms of communication in real-time. This course offers a survey of scholarship from the nascent field of critical AI studies to investigate the impact of AI, machine learning, and big data on knowledge production, representation, and consumption. In addition to theoretical discussions, we will conduct research-creation experiments aimed at documenting and evaluating emerging methods of AI-augmented content creation across text, image, and sound. Prospective students should demonstrate a substantial interest in media art and design and its connections to digital humanities, critical theory, and pedagogy. Experience with artistic and/or engineering practice is a plus. Please submit a 300 word max statement of interest to uhl@uchicago.edu by 12/22 in order to be considered for enrollment.

Equivalent Course(s): KNOW 36043, ANTH 26043, ANTH 36043, CHSS 36043, MAAD 12043, CMST 36043

MACS 37000. Thinking with Deep Learning for Complex Social & Cultural Data Analysis. 100 Units.
A deluge of digital content is generated daily by web-based platforms and sensors that capture digital traces of human communication and connection, and complex states of society, culture, economy, and the world. Emerging deep learning methods enable the integration of these complex data into unified social and cultural "spaces" that enable new answers to classic social and cultural questions, and also pose novel questions. From the perspective of deep learning, everything can be viewed as data-novels, field notes, photographs, lists of transactions, networks of interaction, theories, epistemic styles-and our treatment examines how to configure deep learning architectures and multi-modal data pipelines to improve the capacity of representations, the accuracy of complex predictions, and the relevance of insights to substantial social and cultural questions. This class is for anyone wishing to analyse textual, network, image or arbitrary structured and unstructured data, especially in concert with one another to solve complex social and cultural analysis problems (e.g., characterize a culture; predict next year's ideology).

Equivalent Course(s): SOCI 36032

MACS 40000. Economic Policy Analysis with Overlapping Generation Models. 100 Units.
This course will study economic policy questions ideally addressed by the overlapping generations (OG) dynamic general equilibrium framework. OG models represent a rich class of macroeconomic general equilibrium model that is extremely useful for answering questions in which inequality, demographics, and individual heterogeneity are important. OG models are used extensively by the Joint Committee on Taxation, Congressional Budget Office, and Department of the Treasury. This course will train students how to set up and solve OG models. The standard nonlinear global solution method for these models-time path iteration--is a fixed point method that is similar to but significantly different from value function iteration. This course will take students through progressively richer versions of the model, which will include endogenous labor supply, nontrivial demographics, bequests, stochastic income, multiple industries, non-balanced government budget constraint, and household tax structure.

MACS 40100. Big Data and Society. 100 Units.
The massive explosion of information produced by computers and sophisticated computational methods capable of harnessing this data to generate inferences has led to an increasingly data-driven society. Businesses, governments, and individuals seek to leverage this data to develop and market products, formulate policy, and improve the human condition. Computational approaches to decision making have become increasingly prevalent in domains such as criminal justice, education, employment, finance, and politics. While decision making based on data mining and algorithms has the capacity to improve society, critics argue that these approaches strengthen socioeconomic class divisions, constitute an invasion of privacy, or violate the civil rights of minority groups. This course will survey some of the major uses of big data in society and assess the potential ethical, moral, and legal implications of these models.

MACS 40101. Social Network Analysis. 100 Units.
This course introduces students to concepts and techniques of Social Network Analysis ("SNA"). Social Network Analysis is a theoretical approach and a set of methods to study the structure of relationships among entities (e.g., people, organizations, ideas, words, etc.). Students will learn concepts and tools to identify network nodes, groups, and structures in different types of networks. Specifically, the class will focus on a number of social network concepts, such as social capital, homophily, contagion, etc., and on how to operationalize them using network measures, such as centrality, structural holes, and others.

Equivalent Course(s): MACS 20101, SOCI 40248
MACS 40200. Structural Estimation. 100 Units.
Structural estimation refers to the estimation of model parameters by taking a theoretical model directly to the data. (This is in contrast to reduced form estimation, which often entails estimating a linear model that is either explicitly or implicitly a simplified, linear version of a related theoretical model). This class will survey a range of structural models, then teach students estimation approaches including the generalized method of moments approach and maximum likelihood estimation. We will then examine the strengths and weaknesses of both approaches in a series of examples from the fields of economics, political science, and sociology. We will also learn the simulated method of moments approach. We will explore applications across the social sciences.

MACS 40236. Panel Data Spatial Econometrics. 100 Units.
This course covers econometric methods specifically geared to deal with the presence of spatial dependence and spatial heterogeneity in panel data models, i.e., models based on data with both a cross-sectional and time series dimension. Such data are increasingly common in many areas of empirical social science research. The main objectives of the course are to gain insight into the way spatial effects can be incorporated into panel data regression model specifications, what are the proper methods to carry out specification tests and to estimate such models, and how the results should be interpreted in terms of the implied dynamics across space and over time. Special attention is paid to the application to spatial models of generic statistical paradigms, such as fixed and random effects, maximum likelihood and quasi-maximum likelihood estimation, the generalized method of moments, and semi-parametric estimation. An important aspect of the course is an emphasis on computation and leveraging open source software tools such as R and Python to carry out estimation and simulation.
Equivalent Course(s): SOCI 40236

MACS 40300. Open Research Methods. 100 Units.
The purpose of this course is to give students experience in the broad set of skills and tools for managing, collaborating on, and contributing to open source research projects. Transparency and replicability of research have received renewed emphasis in recent years due to the increased prevalence and sophistication of empirical and computational methods as well as the increased availability of large high frequency data sources. This course focuses on the open source programming languages of Python and R, but the principles could be applied to projects using any language. The course will present the common open source software development workflow as an efficient structure for collaborative academic research. We will learn Git and GitHub basic tools and methods. We will practice multiple levels of documentation ranging from in-code docstrings to full PDF and HTML documentation tools. Students will implement continuous integration testing and regression testing in their own open source repositories. And students will learn how to set an environment with specific library and package versions. We will also discuss methods for anonymizing proprietary data or creating synthetic datasets that can be used by the general public.

MACS 40400. Computation and the Identification of Cultural Patterns. 100 Units.
Culture is increasingly becoming digital, making it more and more necessary for those in both academia and industry to use computational strategies to effectively identify, understand, and (in the case of industry) capitalize on emerging cultural patterns. In this course, students will explore interdisciplinary approaches for defining and mobilizing the concept of “culture” in their computational analyses, drawing on relevant literature from the fields of Anthropology, Psychology and Sociology. Additionally, they will receive hands-on experience applying computational approaches to identify and analyze a wide range of cultural patterns using the Python programming language. For instance, students will learn to identify emerging social movements using social media data, predict the next fashion trends, and even decipher ancient symbols using archaeological databases.
Equivalent Course(s): CHDV 40404, PSYC 40460, MACS 20400, MAPS 40401

MACS 40500. Computational Methods for American Politics. 100 Units.
In this class, students will be introduced to several computational techniques aimed at exploring, understanding, and diagnosing substantive American political phenomena. Rather than focus on derivations and proofs of models, the main focus of the course will be applying and diagnosing model fit, along with computation and application in R. The goal of the class is twofold: first, to offer students a methodological toolbox to tackle complex questions of interest in the social sciences. The second goal, then, is to prepare students for applied quantitative research, offering modern data science techniques and computational training in the service of understanding and predicting American political behavior in a range of contexts. The course will be a combination seminar/applied, where we will read and discuss the latest developments as well as classical works related to a week’s topic, but also apply the concepts in R.
Equivalent Course(s): CHDV 40500, PLSC 20525, PLSC 40525

MACS 40550. Agent-Based Modeling. 100 Units.
Social science problems often have so many details and moving parts that it can be difficult for researchers to gain traction without models. In this course, we explore agent-based modeling approaches to understand these social science problems including cooperation and the development of culture. Agent-based models enable us to build an understanding from the bottom up, starting with simple assumptions and analyzing how patterns emerge at a larger scale. Through the term, we’ll cover the fundamentals of modeling, including basic principles of model design, data extraction, and canonical examples like Conway’s Game of Life, Schelling’s segregation model, and Boids/flocking. The course is balanced between social science readings and applications and hands-on coding. It culminates in a final project consisting of an agent-based model designed by students to apply to a social science phenomenon.
Equivalent Course(s): MACS 20550

MACS 40600. More Computing for the Social Sciences. 100 Units.
This is an applied course for social scientists expanding on computational approaches to reproducible research via programming. It extends on the training in MACS 30500 to cover intermediate and advanced techniques for core data science tasks such as data wrangling, visualization, modeling, and communication. Exact topics will vary, but may include items such as interactive visualizations and web applications, package and API development, functional programming, code profiling and optimization, etc.

MACS 40700. Data Visualization. 100 Units.
Social scientists frequently wish to convey information to a broader audience in a cohesive and interpretable manner. Visualizations are an excellent method to summarize information and report analysis and conclusions in a compelling format. This course introduces the theory and applications of data visualization. Students will learn techniques and methods for developing rich, informative and interactive, web-facing visualizations based on principles from graphic design and perceptual psychology. Students will practice these techniques on many types of social science data, including multivariate, temporal, geospatial, text, hierarchical, and network data. These techniques will be developed using a variety of software implementations such as R, ggplot2, D3, and Tableau.
Equivalent Course(s): MACS 20700

MACS 40800. Unsupervised Machine Learning. 100 Units.
A full understanding of data structure is not always possible, nor are tidy labeled data always available to researchers. With an applied focus, this course will cover prominent unsupervised machine learning techniques such as clustering, partitioning, dimension reduction, and deep learning for discovering latent, non-random structure in data. Further, mechanics involved in unsupervised machine learning will also be covered, such as measuring distance, visualization, and methods of validation. Where appropriate, we will also cover best practices in functional programming.
Equivalent Course(s): PLSC 40825, PLSC 20825, MAPS 40800

MACS 40900. Studying 'Social Problems': Theory and Methods. 100 Units.
What does it mean for something to be a ‘social problem?’ How do particular ‘social problems’ emerge and how do they stop being ‘problems’? This course answers these questions from the perspective of the sociology of social problems and introduces students to different tools and data sources they can use to better understand a social problem, or a variety of other phenomena, that they might be interested in. In the first part of the course, we will cover sociological theories of ‘social problems’ and read a selection of case studies. In the second part of the course, we will survey different kinds of data sources (Twitter feeds, newspaper and congressional records, article databases, various publicly available datasets, etc.) and discuss how you can best leverage them to study specific ‘social problems.’ By the end of the class, each student will have produced an extensive report on a topic of interest. As such, the class is particularly well-suited for students doing independent research, such as working on their BA or MA. The course does not assume any previous knowledge, beyond basic proficiency with the Internet and software such as Excel. However, the instructor will orient parts of the class towards students who have some programming background, in order to emphasize the utility of computational approaches.
Equivalent Course(s): MAPS 40900, SOCI 30560, CHDV 20900, SOCI 20560, MACS 20900

MACS 41200. Advanced Machine Learning. 100 Units.
This is an intermediate-to-advanced introduction to the mathematical and computational aspects of the core statistical and machine learning techniques. The goal is to equip students with a knowledge of the theoretical and practical aspects of four groups of machine learning methods which are widely used in applied research: (1) dimension reduction (PCA, MDS, and their extensions) (2) classification methods (SVM, Bayes classifiers, and other classification methods) (3) clustering procedures (K-means, FMM, non- and semi-parametric Bayesian methods) (4) categorical data analysis (with brief introduction to probabilistic graphical models). The course includes some applications in Political Science, such as FMM to estimate fraud in elections, PCA to construct indices to measure democracy, and text classification.

MACS 41300. Computational Methods for Comparative Politics. 100 Units.
Comparative Politics is one of the most traditional areas in Political Science. In this course, students are exposed to some of the methodological challenges of studying politics from a comparative perspective. The course draws on canonical substantive and methodological debates in Comparative Politics and discusses some modern machine learning, latent variable analysis, and computational methods to overcome some of those difficulties. With instructor guidance, students will have the opportunity to develop their project and apply computational methods to study a topic of their choice in comparative politics.

MACS 49800. Research Experience: Psychology Lab. 000 Units.
All MAPSS and MACSS students with a confirmed Psychology lab placement will be pre-registered. Students should contact Sabrina Biggus (sbiggus@uchicago.edu) to specify the lab placement they have. This course does not count as one of the three for-credit courses you must take each quarter to maintain full-time status in your MA program.
Equivalent Course(s): MAPS 49800
MACS 50000. Computational Social Science Workshop. 000 Units.
High performance and cloud computing, massive digital traces of human behavior from ubiquitous sensors, and a growing suite of efficient model estimation, machine learning and simulation tools are not just extending classical social science inquiry, but transforming it to pose novel questions at larger and smaller scales. The Computational Social Science (CSS) Workshop is a weekly event that features this work, highlights associated skills and data, and explores the use of CSS in the world. The CSS Workshop alternates weekly between research workshops and professional workshops. The research workshops feature new CSS work from top faculty and advanced graduate students from UChicago and around the world, while professional workshops highlight useful skills and data (e.g., machine learning with Python’s scikit-learn; the Twitter firehose API) and showcase practitioners using CSS in the government, industry and nonprofit sectors. Each quarter, the CSS Workshop also hosts a distinguished lecture, debate and dinner, and a student conference.

MACS 51000. Introduction to Causal Inference. 100 Units.
This course is designed for graduate students and advanced undergraduate students from the social sciences, education, public health science, public policy, social service administration, and statistics who are interested in quantitative research and are interested in studying causality. The goal of this course is to equip students with basic knowledge of and analytic skills in causal inference. Topics for the course will include the potential outcomes framework for causal inference; experimental and observational studies; identification assumptions for causal parameters; potential pitfalls of using ANCOVA to estimate a causal effect; propensity score based methods including matching, stratification, inverse-probability-of-treatment-weighting (IPTW), marginal mean weighting through stratification (MMWS), and doubly robust estimation; the instrumental variable (IV) method; regression discontinuity design (RDD) including sharp RDD and fuzzy RDD; difference in difference (DID) and generalized DID methods for cross-section and panel data, and fixed effects model. Intermediate Statistics or equivalent such as STAT 224/PBHS 324, PP 31301, BUS 41100, or SOC 30005 is a prerequisite. This course is a prerequisite for "Advanced Topics in Causal Inference" and "Mediation, moderation, and spillover effects."
Equivalent Course(s): CHDV 20102, SOCI 30315, STAT 31900, PBHS 43201, PLSC 30102, CHDV 30102

MACS 51100. Computational Social Science Skills Workshop. 000 Units.
Modern social scientific research designs often require individuals to have advanced computational skills and the ability to write programs that implement the research tasks. This workshop teaches participants a range of computational tools and methods within open-source programming languages (e.g. R, Python, Julia). Workshop topics will vary throughout the quarter and have differing prerequisites (purely introductory, intermediate, advanced training, etc.).

MACS 52000. Advanced Topics in Causal Inference. 100 Units.
This course provides an in-depth discussion of selected topics in causal inference that are beyond what are covered in the introduction to causal inference course. The course is intended for graduate students and advanced undergraduate students who have taken the intro course and want to extend their knowledge in causal inference. Topics include (1) alternative matching methods, randomization inference for testing hypothesis and sensitivity analysis; (2) marginal structural models and structural nested models for time-varying treatment; (3) Rubin Causal Model (RCM) and Heckman’s scientific model of causality; (4) latent class treatment variable; (5) measurement error in the covariates; (6) the M-estimation for the standard error of the treatment effect for the use of IPW; (7) the local average treatment effect (LATE) and its problems, sensitivity analysis to examine the impact of plausible departure from the IV assumptions, and identification issues of multiple IVs for multiple/one treatments; (8) Multi-level data for treatment evaluation for multilevel experimental designs and observational designs, and spilt-over effect; (9) Nonignorable missingness and informative censoring issues. Equivalent Course(s): CHDV 40102, SOCI 40202

MACS 54000. Introduction to Spatial Data Science. 100 Units.
Spatial data science consists of a collection of concepts and methods drawn from both statistics and computer science that deal with accessing, manipulating, visualizing, exploring and reasoning about geographical data. The course introduces the types of spatial data relevant in social science inquiry and reviews a range of methods to explore these data. Topics covered include formal spatial data structures, geovisualization and visual analytics, rate smoothing, spatial autocorrelation, cluster detection and spatial data mining. An important aspect of the course is to learn and apply open source GeoDa software. Equivalent Course(s): GISc 20500, CEGU 20253, GISc 30500, ENST 20253, SOCI 20253, SOCI 30253

MACS 55000. Spatial Regression Analysis. 100 Units.
This course covers statistical and econometric methods specifically geared to the problems of spatial dependence and spatial heterogeneity in cross-sectional data. The main objective of the course is to gain insight into the scope of spatial regression methods, to be able to apply them in an empirical setting, and to properly interpret the results of spatial regression analysis. While the focus is on spatial aspects, the types of methods covered have general validity in statistical practice. The course covers the specification of spatial regression models in order to incorporate spatial dependence and spatial heterogeneity, as well as different estimation methods and specification tests to detect the presence of spatial autocorrelation and spatial heterogeneity. Special attention is paid to the application to spatial models of generic statistical paradigms, such as Maximum Likelihood, Generalized Methods of Moments and the Bayesian perspective. An important aspect of the course is the application of open source software tools such as R, GeoDa and PySal to solve empirical problems. Equivalent Course(s): GEOG 40217, SOCI 40217
MACS 60000. Computational Content Analysis. 100 Units.
A vast expanse of information about what people do, know, think, and feel lies embedded in text, and more of the contemporary social world lives natively within electronic text than ever before. These textual traces range from collective activity on the web, social media, instant messaging and automatically transcribed YouTube videos to online transactions, medical records, digitized libraries and government intelligence. This supply of text has elicited demand for natural language processing and machine learning tools to filter, search, and translate text into valuable data. The course will survey and practically apply many of the most exciting computational approaches to text analysis, highlighting both supervised methods that extend old theories to new data and unsupervised techniques that discover hidden regularities worth theorizing. These will be examined and evaluated on their own merits, and relative to the validity and reliability concerns of classical content analysis, the interpretive concerns of qualitative content analysis, and the interactional concerns of conversation analysis. We will also consider how these approaches can be adapted to content beyond text, including audio, images, and video. We will simultaneously review recent research that uses these approaches to develop social insight by exploring (a) collective attention and reasoning through the content of communication; (b) social relationships through the process of communication; and (c) social state
Equivalent Course(s): SOCI 40133, CHDV 30510

MACS 95000. Computation MA Internship. 000 Units.
All MACSS students who have completed three academic quarters of full-time course work in our MA program are eligible to participate in the Computational Social Science Internship Program. Any interested persons must speak with Career Services, have an approved external employer, complete a petition from our Student Affairs Administrator, and enroll in this non-credit field research course. The course will appear on your transcript, and will be evaluated on a pass/fail basis, in consultation with the employer. Note that MACS 95000 does not count against your other curricular requirements.