Online Master of Science Program in Analytics

The Master of Science in Analytics gives students a thorough knowledge of techniques in the field of analytics and data science, and the ability to apply them to real-world business scenarios. Building from a core in applied statistics, math and programming, students are provided with advanced analytical training to develop their ability to draw insights from big data. This includes: machine learning and predictive analytics, deep learning, reinforcement learning, data engineering platforms, time series analysis, linear and non-linear models, statistical methods, and other sophisticated techniques for analyzing complex data.

The program is highly applied in nature, integrating business strategy, project-based learning, simulations, case studies, and specific electives addressing the analytical needs of various industry sectors. Through partnerships with key employers, the program also provides students with a client based, 2 term Capstone experience as well as access to career networks and employment pathways upon graduation.

- Program type: masters degree program
- Program structure, courses, requirements, and application ([https://grahamschool.uchicago.edu/credit/master-science-analytics/index/](https://grahamschool.uchicago.edu/credit/master-science-analytics/index/))
- Location: online (synchronous and asynchronous)
- Full-time: weekday, weekday evening, and Saturday classes (as available)
- Part-time: weekday evenings and Saturday classes (as available)
- Time to completion: 1-4 years
- Only courses with a grade of B- or better will count toward degree requirements

Minimum g.p.a. for satisfactory academic progress: 2.7

Admission criteria:
- Online application
- One transcript from each prior academic institution
- Candidate statement
- Resume or CV

Applicants who attended an international university must also:
- Satisfy English language proficiency requirement
- Provide course by course evaluation

Program requirements:
12 courses curriculum (Academic Year 2022-23 entering students)

- Foundational Skills courses (non-credit courses, 4 – 6 depending on test waivers)
- Core courses (7)
- Electives (3)
- Capstone project (2)

Foundational Courses:

Foundation courses provide the basis for our rigorous analytics degree that support the theoretical, strategic, and practical analytics studies in more advanced courses. Students with sufficient preparation may be eligible to bypass the programming course.

Pre-quarter foundational courses (non-credit):
- Based on results of a linear algebra pre-test, students may also be required to take the following non-credit course
  - Linear Algebra (online course required prior to program start)
- MSCA 31000 Introduction to Statistical Concepts
- MSCA 37016 Advanced Linear Algebra for Machine Learning
- MSCA 37006 R Workshop

Required foundational course (non-credit):
ONLINE MASTER OF SCIENCE PROGRAM IN ANALYTICS

• MSCA 37010 Programming for Analytics (recommended in first quarter)
• MSCA 37014 Python for Analytics (waived with a Python exemption test)

MSCA Core requirements:
• MSCA 31006 Time Series Analysis and Forecasting
• MSCA 31007 Statistical Analysis
• MSCA 31008 Data Mining Principles
• MSCA 31009 Machine Learning & Predictive Analytics
• MSCA 31010 Linear and Non-Linear Models

One of the following Data Engineering courses
• MSCA 31012 Data Engineering Platforms for Analytics*
• MSCA 31013 Big Data Platforms*

One of the following Leadership courses
• MSCA 31003 Leadership Skills*
• MSCA 31015 Data Science for Consulting*

MSCA Electives (subject to instructor availability):
• MSCA 32001 Financial Analytics
• MSCA 32003 Marketing Analytics
• MSCA 32007 Data Visualization Techniques
• MSCA 32009 Health Analytics
• MSCA 32013 Optimization and Simulation Methods for Analytics
• MSCA 32014 Bayesian Methods
• MSCA 32015 Digital Marketing Analytics in Theory and Practice
• MSCA 32017 Advanced Machine Learning & Artificial Intelligence
• MSCA 32018 Natural Language Processing and Cognitive Computing
• MSCA 32019 Real-Time Intelligent Systems
• MSCA 32020 Reinforcement Learning
• MSCA 32021 Machine Learning Operations
• MSCA 32023 Advanced Computer Vision with Deep Learning

Capstone project:
• MSCA 34002 Capstone 1
• MSCA 34003 Capstone 2

Non-credit workshops & short courses:
• MSCA 37001 Hadoop Workshop
• MSCA 37002 Linux Workshop
• MSCA 37013 Ethics In Big Data Analytics
• MSCA 37015 Introduction to Ethics in Data Analytics
• MSCA 37018 Next Chapter: Perspectives on Hot Topics in Machine Learning and Artificial Intelligence
• MSCA 37019 Your Career in Data Science

*Optional core courses may be taken as electives.
**Optional core courses may be taken as electives.
***The University of Chicago has applied for a license to offer online courses to students residing in California. California residents may not participate in this online program or enroll in its online courses until a California license is obtained.

Universities delivering online courses to students in California are required to charge students a fee of $2.50 per one thousand dollars ($1,000) of institutional charges to support the State's Student Tuition Recovery Fund and to provide students and prospective students with the following disclosures. Universities are not permitted to cover this charge for students.

“The State of California established the Student Tuition Recovery Fund (STRF) to relieve or mitigate economic loss suffered by a student in an educational program at a qualifying institution, who is or was a California resident while enrolled, or was enrolled in a residency program, if the student enrolled in the
institutions, prepaid tuition, and suffered an economic loss. Unless relieved of the obligation to do so, you must pay the state-imposed assessment for the STRF, or it must be paid on your behalf, if you are a student in an educational program, who is a California resident, or are enrolled in a residency program, and prepay all or part of your tuition.

You are not eligible for protection from the STRF and you are not required to pay the STRF assessment, if you are not a California resident, or are not enrolled in a residency program.

It is important that you keep copies of your enrollment agreement, financial aid documents, receipts, or any other information that documents the amount paid to the school. Questions regarding the STRF may be directed to the Bureau for Private Postsecondary Education, 1747 North Market Blvd., Suite 225, Sacramento, California, 95834, (916) 574-6900 or (888) 370-7589.

To be eligible for STRF, you must be a California resident or enrolled in a residency program, prepaid tuition, paid or deemed to have paid the STRF assessment, and suffered an economic loss as a result of any of the following:

1. The institution, a location of the institution, or an educational program offered by the institution was closed or discontinued, and you did not choose to participate in a teach-out plan approved by the Bureau or did not complete a chosen teach-out plan approved by the Bureau.
2. You were enrolled at an institution or a location of the institution within the 120-day period before the closure of the institution or location of the institution, or were enrolled in an educational program within the 120-day period before the program was discontinued.
3. You were enrolled at an institution or a location of the institution more than 120 days before the closure of the institution or location of the institution, in an educational program offered by the institution as to which the Bureau determined there was a significant decline in the quality or value of the program more than 120 days before closure.
4. The institution has been ordered to pay a refund by the Bureau but has failed to do so.
5. The institution has failed to pay or reimburse loan proceeds under a federal student loan program as required by law or has failed to pay or reimburse proceeds received by the institution in excess of tuition and other costs.
6. You have been awarded restitution, a refund, or other monetary award by an arbitrator or court, based on a violation of this chapter by an institution or representative of an institution, but have been unable to collect the award from the institution.
7. You sought legal counsel that resulted in the cancellation of one or more of your student loans and have an invoice for services rendered and evidence of the cancellation of the student loan or loans.

To qualify for STRF reimbursement, the application must be received within four (4) years from the date of the action or event that made the student eligible for recovery from STRF.

A student whose loan is revived by a loan holder or debt collector after a period of noncollection may, at any time, file a written application for recovery from STRF for the debt that would have otherwise been eligible for recovery. If it has been more than four (4) years since the action or event that made the student eligible, the student must have filed a written application for recovery within the original four (4) year period, unless the period has been extended by another act of law.

However, no claim can be paid to any student without a social security number or a taxpayer identification number.

M.S. IN ANALYTICS COURSES

MSCA 31000. Introduction to Statistical Concepts. 000 Units.
This course provides general exposure to basic statistical concepts that are necessary for students to understand the content presented in more advanced courses in the program. The course covers theoretical distributions and the way these distributions are used to assign probabilities to events in some depth. The course also introduces students to descriptive statistical methods to explore and summarize data, methodologies for sampling units for measurement or analysis, drawing inferences on the basis of knowledge gained from samples to populations, assessing relationships between variables, and making predictions based upon relationships between variables.
Terms Offered: Autumn Spring

MSCA 31001. Research Design for Business Applications. 100 Units.
In addition to theory and experimentation, big data analytics has now emerged as an alternative way to discover new knowledge. This course covers the analytics research process from the translation of business problems into researchable questions that can be addressed by using analytics, development of data sources to address each key researchable issue, to the translation of research results back to business implications. By completing the course, students will be able to: frame a business problem; map alternative solutions to develop a plan; identify potential sources or relevant data; understand analytical principles that can be applied to design data-gathering experiments; explain the pros and cons of the selected methodology to the analytical team as well as non-analysts. Students will develop a research proposal to produce knowledge from data to address a real business problem in small steps throughout the course.
Terms Offered: Autumn Spring Summer Winter

**MSCA 31003. Leadership Skills. 100 Units.**
In Leadership Skills: Teams, Strategies, and Communications, students learn how to work effectively in teams to identify, structure, and communicate the business value of data analytics to an organization. The goals of the course are (1) to identify points in an organization that can benefit from analytics; (2) to structure analytic problems from a strategic perspective, thereby identifying business impact; (3) to develop the ability to communicate the power of analytics to others, especially senior leaders; and (4) to work in a team to accomplish these and related goals successfully. At the end of the course, students should have the ability to describe business problems that lend themselves to a data analytics approach, position these problems from the perspective of a coherent business strategy, and represent the power of analytics to a business audience. Students should also understand how to harness the powerful dynamics of a team to achieve excellence in the world of data analytics.

Terms Offered: Autumn Spring Summer Winter

**MSCA 31004. Machine Learning & Predictive Analytics. 100 Units.**
This course in advanced data mining will provide a practical, hands-on set of lectures surrounding modern predictive analytics and machine learning algorithms and techniques. It will emphasize practice over mathematical theory, and students will spend a considerable amount of class time gaining experience with each algorithm using existing packages in R, Python, and Linux libraries. The course will cover the following topics: regression and logistic regression, regularized regression including the lasso and elastic net techniques, support vector machines, neural networks, decision trees, boosted decision trees and random forests, online learning, k-means and special clustering, and survival analysis.

Terms Offered: Autumn Spring Summer Winter

**MSCA 31005. Data Mining Principles. 100 Units.**
Drawing on statistics, artificial intelligence and machine learning, the data mining process aims at discovering novel, interesting and actionable patterns in large datasets. This class will introduce the student to the fundamentals of data mining: association and sequence rules discovery, memory-based reasoning, classification and regression decision trees, comparison of data mining models, logistic models, scorecard models, and neural network models. The class follows a learn-by-doing approach in which the student will complete bi-weekly assignments using real world datasets. The student will also propose and complete a data mining research project of their own design.

Terms Offered: Autumn Spring Summer Winter

**MSCA 31006. Time Series Analysis and Forecasting. 100 Units.**
Time Series Analysis is a science as well as the art of making rational predictions based on previous records. It is widely used in various fields in today’s business settings. For example, airline companies employ time series to predict traffic volume and schedule flights; financial agencies measure market risk via stock price series; marketing analysts study the impact of a newly proposed advertisement by the sales series. A comprehensive knowledge of time series analysis is essential to the modern data scientist/analyst. This course covers important issues in applied time series analysis: a solid knowledge of time series models and their theoretical properties; how to analyze time series data by using mainstream statistical software; practical experience in real data analysis and presentation of their findings in a logical and clear way to various audiences.

Terms Offered: Autumn Spring Summer Winter
Prerequisite(s): MSCA 31007: Statistical Analysis

**MSCA 31007. Statistical Analysis. 100 Units.**
This course provides a comprehensive and practical introduction to statistical data analysis. The statistical techniques taught in this course will enable students to analyze complex datasets and formulate and solve real-world problems to facilitate data-driven decisions. Throughout the course, students will learn concepts and fundamentals of statistical inference and regression analysis by studying theory, developing intuition, and working through several practical examples. Students will become proficient in interpreting standard regression output and conducting model selection and validation. Students will also learn the statistical programming language used to construct examples and homework exercises. Examples will be constructed using R. Students will have many opportunities to apply the new concepts to real data and develop their own statistical routines. The course also addresses the importance of quality control and reproducibility when conducting research and developing work product.

Terms Offered: Autumn Spring Summer Winter
Prerequisite(s): MSCA 31000: Introduction to Statistical Concepts MSCA 37016: Advanced Linear Algebra for Machine Learning

**MSCA 31008. Data Mining Principles. 100 Units.**
Drawing on statistics, artificial intelligence and machine learning, the data mining process aims at discovering novel, interesting and actionable patterns in large datasets. This class will introduce the student to the fundamentals of data mining: association and sequence rules discovery, memory-based reasoning, classification and regression decision trees, comparison of data mining models, logistic models, scorecard models, and neural network models. The class follows a learn-by-doing approach in which the student will complete bi-weekly assignments using real world datasets. The student will also propose and complete a data mining research project of their own design.

Terms Offered: Autumn Spring Summer Winter
Prerequisite(s): MSCA 31007: Statistical Analysis

**MSCA 31009. Machine Learning & Predictive Analytics. 100 Units.**
This course in advanced data mining will provide a practical, hands-on set of lectures surrounding modern predictive analytics and machine learning algorithms and techniques. It will emphasize practice over mathematical theory, and students will spend a considerable amount of class time gaining experience with each algorithm using existing packages in R, Python, and Linux libraries. The course will cover the following topics: regression and logistic regression, regularized regression including the lasso and elastic net techniques, support vector machines, neural networks, decision trees, boosted decision trees and random forests, online learning, k-means and special clustering, and survival analysis.

Terms Offered: Autumn Spring Summer Winter
Prerequisite(s): MSCA 31008: Data Mining Principles, MSCA 31010: Linear and Non-Linear Models; MSCA 37014: Python for Analytics

**MSCA 31010. Linear and Non-Linear Models. 100 Units.**
This course concentrates on the following topics: Review of statistical inference based on linear model, extension to the linear model by removing the assumption of Gaussian distribution for the output (Generalized Linear Model), extension to the linear model by allowing a correlation structure for the model residuals (mixed effect models), and extension of the linear model by relaxing the requirement that inputs are combined linearly.
(nonparametric regression, regime switches). Course emphasizes applications of these models to various fields and covers main steps of building analytics from visualizing data and building intuition about their structure and patterns to selecting appropriate statistical method to interpretation of the results and building analytical model. Topics are illustrated by data analysis projects using R. Familiarity with R at some basic level is not a requirement but recommendation. Students can pick up the programming language by following the descriptions of the examples.

Terms Offered: Autumn Spring Summer Winter
Prerequisite(s): MSCA 31007: Statistical Analysis

MSCA 31011. Statistical Analysis Review. 000 Units.
Lab / TA review session to supplement MSCA 31007: Statistical Analysis

MSCA 31012. Data Engineering Platforms for Analytics. 100 Units.
Effective data engineering is an essential first step in building an analytics-driven competitive advantage in the market. Modern data engineering platforms reduce manual data preparation by automating processes, which in turn, enable companies to focus on deriving efficiencies in data processing to develop impactful business insights. This course provides students with a thorough understanding of the fundamentals of data engineering platforms, for both operational and analytical use cases, while gaining hands-on expertise in building these platforms in a way to develop analytical solutions effectively. Students will have the opportunity to construct both relational and analytical databases on the cloud or on premise from real-life datasets while using programmatic or configuration driven data pipelines. By the end of the course, students will be able to design and implement an end-to-end data engineering platform capable of supporting sustainable analytics solutions.

Terms Offered: Autumn Spring Summer Winter

MSCA 31013. Big Data Platforms. 100 Units.
This course teaches students how to approach Big Data and large-scale machine learning applications. While there is no single definition of Big Data and multiple emerging software packages exist to work with Big Data, we will cover the most popular approaches. Students will learn the Big Data infrastructure, including Linux, Massive parallelization and Distributed Computing, and how to apply both Hadoop and Spark map-reduce concepts for clustering, similarity search, web analytics and classification. During the course, we will cover the applications of NoSQL systems, such as JSON stores, object storage and Elasticsearch. The cloud computing section of the course will focus on virtualization and container orchestration, including virtual machines, dockers and Kubernetes. During the course students will gain hands-on expertise leveraging Hive, Pig, Python and PySpark for Big Data applications in client-server environment. For MSCA students in the 12-course curriculum who wish to take Big Data Platforms as a core course, instead of MSCA 31012 Data Engineering Platforms: Certain technical skills and knowledge are required to be successful in this course. Required skills may be found for MSCA 31013 Big Data Platforms at https://professional.uchicago.edu/find-your-fit/masters/master-science-analytics/curriculum. Contact the program administration for further details.

Terms Offered: Autumn Spring Summer Winter

MSCA 31015. Data Science for Consulting. 100 Units.
The demand for analytics and data-driven decision making creates a market demand for expertise driven leadership - evidenced in knowledgeable consultants that bring data science and results-driven impact to clients. The successful data science leader / consultant brings an uncommon combination of deep business acumen, data literacy, leading edge methodology experience, inspirational team leadership, client communication management and organizational change skills. Successful consultants rely on a variety of consulting tools to diagnose organizational problems, identify solutions and deliver those solutions. The Data Science for Consulting course will enable students: 1) Understanding the structure of consulting organizations and engagements 2) Developing data science solutions to enterprise problems through employing traditional consulting frameworks and best practice tools. 3) Practicing successful project delivery through effective data discovery, communication, influential team leadership and client relationship management.

Instructor(s): Gregory Green, Marco Serrato, Donald Patchell Terms Offered: Autumn Spring Summer Winter
Prerequisite(s): Restricted to MSCA & MSAP students only.

MSCA 32001. Financial Analytics. 100 Units.
This course concentrates on the following topics: review of financial markets and assets traded on them; main characteristics of financial analytics: returns, yields, volatility; review of stochastic models of market price and their statistical representations; concept of arbitrage, elements of arbitrage pricing approach; principles of volatility analyses, implied vs. realized volatility; correlation, cointegration and other relationships between various financial assets; market risk analytics and management of portfolios of financial assets. The course puts special emphasis on covering main steps of building analytics from visualizing data and building intuition about their structure and patterns to selecting appropriate statistical method to interpretation of the results and building analytical models. Topics are illustrated by data analysis projects using R. Basic familiarity with R is a requirement.

Terms Offered: Summer Winter
Prerequisite(s): MSCA 31007: Statistical Analysis
MSCA 32003. Marketing Analytics. 100 Units.
(Data Science for Algorithmic Marketing) This course focuses on data science methods and algorithms for that are used to develop marketing strategies, and create a link between marketing, customer behavior and business outcome. The course will focus on analytical techniques organized according to the Strategic Marketing Process. The course would cover algorithms for competitive analysis and market sizing, market segmentation, targeted marketing via database marketing, design of new products, market sizing & forecasting via diffusion models, real time product positioning, algorithmic marketing in the digital world, pricing and promotions, marketing effectiveness and ROI. The course will use a combination of lecture, in-class discussions, and group work.
Terms Offered: Autumn Spring
Prerequisite(s): MSCA 31007: Statistical Analysis

MSCA 32004. Risk Analytics. 100 Units.
This course teaches analytical tools commonly employed in the areas of credit and insurance risk. In the area of credit risk, students are at the end of the course should be able to: Understand the business problems and their challenges in the consumer credit risk analytics, design and apply analytical approaches tailored to each problem, and identify and address the underlying assumptions in the designed approaches. In the area of insurance risk, students should be able to: Understand various risks related to the insurance business, in particular the underwriting or pricing risks, quantify and price an individual insurance risk exposure and construct customer segmentation by using statistical and actuarial approaches, and assess company's overall risk management performance at the portfolio level.
Terms Offered: Summer
Prerequisite(s): MSCA 31007: Statistical Analysis

MSCA 32007. Data Visualization Techniques. 100 Units.
In today’s data driven enterprise, data storytelling using effective visualization strategies is an essential skill for analytics practitioners in almost every field to explore and present data. This course focuses on modern data visualization technologies, tools, and techniques to convert raw data into actionable information. Modern data visualization tools are at the forefront of the ”self-service analytics” architectures which are decentralizing analytics and breaking down IT bottlenecks for business experts. Moreover, with its foundations rooted in statistics, psychology, and computer science, data visualization shows you how to better understand the data, present clear evidence of your findings to your intended audience and tell engaging data stories through charts and graphics. This course is designed to introduce data visualization as a medium of effective communication using strategic storytelling, and the basis for interactive information dashboards.
Terms Offered: Winter

MSCA 32009. Health Analytics. 100 Units.
Given the breadth of the field of health analytics, this course will provide an overview of the development and rapid expansion of analytics in healthcare, major and emerging topical areas, and current issues related to research methods to improve human health. We will cover such topics as security concerns unique to the field, research design strategies, and the integration of epidemiologic and quality improvement methodologies to operationalize data for continuous improvement. Students will be introduced to the application of predictive analytics to healthcare. Students will understand factors impacting the delivery of quality and safe patient care and the application of data-driven methods to improve care at the healthcare system level, design approaches to answering a research question at the population level, become familiar with the application of data analytics to impacting care at the provider level through Clinical Decision Systems, and understand the process of a Clinical Trail.
Terms Offered: Winter

MSCA 32010. Linear Algebra and Matrix Analysis. 100 Units.
The objective of this course is to provide students a strong foundation on linear equations and matrices. On completion of this course, students will be able to formulate, apply and interpret systems of linear equations and matrices, interpret data analytics problems in elementary linear algebra, and demonstrate understanding of various applications using linear transformations.
Terms Offered: Autumn Spring

MSCA 32013. Optimization and Simulation Methods for Analytics. 100 Units.
This course introduces students to how optimization and simulation techniques can be used to solve many real-life problems. It will cover two classes of optimization methods. First class has been developed to optimize real, non-simulated systems or to find the optimal solution of a mathematical model. The methods that belong to this class include liner programming, quadratic programming and mixed-integer programming. Second class of methods has been developed to optimize a simulation model. The difference with the classical mathematical programming methods is that the objective function (which is the function to be minimized or maximized) is not known explicitly and is defined by the simulation model (computer code). The course will demonstrate multiple approaches to build simulation models, such as discrete event simulations and agent-based simulations. Then, it will show how stochastic optimization and heuristic approaches can be used to analyze the simulated system and design a sequence of computational experiments that allow to develop a basic understanding of a particular simulation model or system through exploration of the parameter space, to find robust plausible behaviors and conditions and robust near-optimal solutions that are not prone to being unstable under small perturbations.
Terms Offered: Summer Winter
MSCA 32014. Bayesian Methods. 100 Units.
Bayesian inference is a method of learning in which Bayes’ theorem is used to combine the previous knowledge with the new evidence in the data to form an improved posterior knowledge. Another name for such methods is probabilistic inference. Probabilistic Bayesian models form the foundation of the most modern algorithms of Machine Learning and Artificial Intelligence. The focus of this course is an introduction to Bayesian approach. Many methods learned by students in Statistical Analysis, Linear and Nonlinear Models, Data Mining and Machine Learning will be reviewed from the point of view of probabilistic inference. We will look at hierarchical, mixture, robust, and non-parametric Bayesian models and learn how to use them in practical applications. Content will include using probabilistic models to make business decisions under uncertainty, analyze causation in the data, use probabilistic inference to assess risk of black swan events, account for uncertainty in project management and other applications. Students will learn necessary facts of probability theory, Bayesian reasoning, Markov chain Monte Carlo using JAGS, STAN and PyMC. The course contains large number of interactive demonstrations, workshops with examples through which the lecturer shares his own hands-on experience with the students.
Terms Offered: Autumn Spring
Prerequisite(s): MSCA 31010: Linear and Non-Linear Models

MSCA 32015. Digital Marketing Analytics in Theory and Practice. 100 Units.
Successfully marketing brands today requires a well-balanced blend of art and science. This course introduces students to the science of web analytics while casting a keen eye toward the artful use of numbers found in the digital space. The goal is to provide marketers with the foundation needed to apply data analytics to real-world challenges they confront daily in their professional lives. Students will learn to identify the web analytic tool right for their specific needs; understand valid and reliable ways to collect, analyze, and visualize data from the web; and utilize data in decision making for their agencies, organizations or clients. By completing this course, students will gain an understanding of the motivations behind data collection and analysis methods used by marketing professionals; learn to evaluate and choose appropriate web analytics tools and techniques; understand frameworks and approaches to measuring consumers’ digital actions; earn familiarity with the unique measurement opportunities and challenges presented by New Media; gain hands-on, working knowledge of a step-by-step approach to planning, collecting, analyzing, and reporting data; utilize tools to collect data using today’s most important online techniques: performing bulk downloads, tapping APIs, and scraping webpages; and understand approaches to visualizing data effectively.
Terms Offered: Autumn

MSCA 32017. Advanced Machine Learning & Artificial Intelligence. 100 Units.
Since the era of big data started, challenges associated with data analysis have grown significantly in different directions: First, the technological infrastructure had to be developed that can hold and process large amounts of data from different sources and of multiple not always well formalized formats. Second, data analysis methods had to be reviewed, selected and modified to work in distributed computational environments like combinations of in-house clusters of servers and cloud. But the biggest challenge of all is learning to think differently in order to ask new types of questions that could not be answered by analyses of less complex data streams with less complex technological infrastructure. In recent years significant progress has been achieved in creating technological ecosystems for big data analysis. Innovative technologies such as open source projects MapReduce, Hadoop, Spark, Storm, Kafka, TensorFlow, H2O, etc. allowed us to look at depths of data unseen before. We have now growing number of sources and educational courses introducing these new tools. It appeared little more difficult to develop new data analysis methods appropriate for the new data ecosystems. There are some new interesting ideas, there is significant
Terms Offered: Summer Winter
Prerequisite(s): Required: MSCA 31009: Machine Learning & Predictive Analytics Recommended: MSCA 37011 Deep Learning & Image Recognition

MSCA 32018. Natural Language Processing and Cognitive Computing. 100 Units.
Extracting actionable insights from unstructured text and designing cognitive applications have become significant areas of application for analytics. Students in this course will learn foundations of natural language processing, including: concept extraction; text summarization and topic modeling; part of speech tagging; named entity recognition; semantic roles and sentiment analysis. For advanced NLP applications, we will focus on feature extraction from unstructured text, including word and paragraph embedding and representing words and paragraphs as vectors. For cognitive analytics section of the course, students will practice designing question answering systems with intent classification, semantic knowledge extraction and reasoning under uncertainty. Students will gain hands-on expertise applying Python for text analysis tasks, as well as practice with multiple IBM Watson services, including: Watson Discovery, Watson Conversation, Watson Natural Language Classification and Watson Natural Language Understanding.
Terms Offered: Autumn Spring
Prerequisite(s): MSCA 31008: Data Mining Principles

MSCA 32019. Real-Time Intelligent Systems. 100 Units.
Developing end-to-end automation and intelligent systems is now the most advanced area of application for analytics. Building such systems requires proficiency in programming, understanding of computer systems,
as well as knowledge of related analytical methodologies, which are the skills that this course aims to teach to students. The course focuses on python and is tailored for students with basic programming knowledge in Python. The course is partially project based. During the first three sessions, we will review basic python concepts and then learn more advanced python and the ways to use Python to handle large data flows. The later sessions are project based and will focus on developing end-to-end analytical solutions in the following areas: Finance and trading, blockchains and crypto-currencies, image recognition, and video surveillance systems.

Terms Offered: Autumn
Prerequisite(s): MSCA 31007: Statistical Analysis Recommended: MSCA 37014: Python for Analytics

MSCA 32020. Reinforcement Learning. 100 Units.
This course is an introduction to reinforcement learning, also known as neuro-dynamic programming. It discusses basic and advanced concepts in reinforcement learning and provides several practical applications. Reinforcement learning refers to a system or agent interacting with an environment and learning how to behave optimally in such environment. An environment typically includes time, actions, states, uncertainty and rewards. Reinforcement learning combines neuro networks and dynamic programming to find an optimal behavior or policy of the system or agent in complex environment setting. Neuro networks approximations are used to circumvent the well-known ‘curse of dimensionality’ which have been a barrier to solving many practical applications. Dynamic programming is the key learning mechanism that the system or the agent uses to interact with the environment and improve its performance. Students will master key learning techniques and will become proficient in applying these techniques to complex stochastic decision processes and intelligent control.

Terms Offered: Spring Summer Winter
Prerequisite(s): MSCA 31007 Statistical Analysis

MSCA 32021. Machine Learning Operations. 100 Units.
The objective of this course is two-folds - first, to understand what Machine Learning Operations (MLOps) is and why it is a key component in enterprise production deployment of machine learning projects. Second, to expose students to software engineering, model engineering and state-of-the-art deployment engineering with hands-on platform and tools experience. This course crosses the chasm that separates machine learning projects/ experiments and enterprise production deployment. It covers 3 pillars in MLOps: software engineering such as AutoML and A/B experimentation; and deployment engineering such as docker containers and model monitoring. The course focuses on best practices in the industry that are critical to enterprise production deployment of machine learning projects. Having completed this course, a student understands the machine learning lifecycle and what it takes to go from ideation to operationalization in an enterprise environment. Furthermore, students get exposure to state-of-the-art MLOps platforms such as allegro (https://allegro.ai/), xpresso (https://abzooba.com/xpresso-ai), Dataiku (https://www.dataiku.com/), LityxIQ (https://lityx.com/), DataRobot (https://www.datarobot.com/), AWS Sagemaker (https://aws.amazon.com/sagemaker/), and technologies such as gitHub, Jenkins, slack, docker, and kubernetes.

Instructor(s): Arnab Bose Terms Offered: Autumn
Prerequisite(s): MSCA 31009: Machine Learning & Predictive Analytics

MSCA 32022. Applied Data Science in Fintech and Cryptocurrencies. 100 Units.
This class explores Data Science methodologies used within the Fintech industry. Each class will focus on a specific business use case within Finance. Each use case will have a case study framework, where students will learn to balance business requirements with technical solutions. This course will include Fintech topics in Commercial Banking, Investment Banking and Cryptocurrencies. Commercial Banking topics include Credit Lending, Fraud Detection, and Customer Service Chatbots. Investment Banking topics include NLP for Stock Prediction and Portfolio Management. Cryptocurrency topics include NFTs, DeFi (Decentralized Finance), Digital Currencies, Portfolio Management, and Anti-Money Laundering. The case study framework will allow students to practice Stakeholder Management and Design Thinking, while refining skills in effective Storytelling and Data Visualization.

Instructor(s): John Navarro Terms Offered: Autumn Spring
Prerequisite(s): MSCA 31009: Machine Learning & Predictive Analytics

MSCA 32023. Advanced Computer Vision with Deep Learning. 100 Units.
Computer vision is the field of computer science that focuses on creating digital systems that can process, analyze, and make sense of visual data in the same way that humans do. Deep learning is a subset of machine learning and a branch of Artificial Intelligence (AI). It involves the training, deployment, and application of large complex neural network architectures to solve cutting-edge problems. Deep Learning has become the primary approach for solving cognitive problems such as Computer Vision and Natural Language Processing (NLP) and has had a massive impact on various industries such as healthcare, retail, automotive, industrial automation, and agriculture. This course will enable students to build Deep Learning models and apply them to computer vision tasks such as object recognition, detection, and segmentation. Students will gain an in-depth understanding of the Deep Learning model development process, tools, and frameworks. Although the focus of the course will primarily be computer vision, students will work on both image and nonimage datasets during class exercises and assignments. Students will gain hands-on experience in popular libraries such as Tensorflow, Keras, and PyTorch. Students will also learn to apply state of the art models such as ResNet, EfficientNet, RCNNs, YOLO, Vision Transformers, etc. for computer vision and work on datasets such as CIFAR, ImageNet, MS COCO, and MPII Human Poses.
Instructor(s): Ashish Pujari Terms Offered: Autumn Spring Summer Winter
Prerequisite(s): Know your computer (Setting environment variables, Using the Mac/PC terminal, traversing applications/folders, updating security preferences). Familiarity with data engineering and cloud computing. MSCA 37010 Programming for Analytics, MSCA 37014 Python for Analytics, MSCA 31009 Machine Learning and Predictive Analytics. ANLT-MS, ANLP-MS, and GSCP students only.

MSCA 32024. Data Science for Algorithmic Marketing. 100 Units.
This course focuses on marketing science methods and algorithms. The course will expose and immerse students in the marketing science algorithms. The course would cover algorithms for undertaking competitive analysis in the digital landscape, market segmentation, mining databases for effective digital marketing, design of new digital and traditional products, forecasting sales and product diffusion, real time product positioning, intra omni-channel optimization and inter omni-channel resource allocation, and pricing across both omni-channel marketing effectiveness and ROI. The course will use a combination of lecture, in-class discussions, group assignments, and a final group project. The course lays special emphasis on algorithms. Hence it draws heavily from the fields of optimization, machine-learning based recommendation systems, association rules, consumer choice models, Bayesian estimation, experimentation and analysis of covariance, advanced visualization techniques for mapping brand perceptions, and analysis of social media data using advanced NLP techniques.
Instructor(s): Anil Chaturvedi Terms Offered: Autumn Winter
Prerequisite(s): MSCA 31007 Statistical Analysis
Note(s): ANLP-MS, ANLT-MS and GSCP-MS only

MSCA 32025. Supply Chain Optimization. 100 Units.
Big Data continues to grow exponentially in our large-scale transactional world where 100,000s of SKUs and millions of customers are interacting with 1:1 offers that include differential pricing, shipping timing/costs and even made to order "custom" product configurations. These consumer behaviors are quickly advancing the availability of new data and techniques within the discipline of Data Science. This elective course will give students the opportunity to apply their skills in data visualization, data mining tools, predictive modeling, and advanced optimization techniques to address Supply Chain challenges. The course focuses on the use of Advanced Predictive Modeling, Machine Learning, AI and other Data Science insight and activation tools are to automate and optimize the performance of the Supply Chain. Students will also learn how to optimize the performance of the Supply Chain from the lens of multiple related disciplines including: Sales Forecasting, Warehousing/Inventory Management, Promotion, Pricing, Logistics Network Optimization, Freight Cost Management, Manufacturing, Retail POS Information, Ecommerce, Consumer Data, and Product Design/ Packaging. After completing this course, you will be prepared to work in any of the numerous specialty areas possible in the world of Supply Chain Management.
Instructor(s): Gizem Aydin Terms Offered: Autumn Spring Summer Winter
Prerequisite(s): MSCA 31008, MSCA 37014, and for ANLT-MS, ANLP-MS, GSCP students only

MSCA 34000. Capstone Project Implementation. 100 Units.
The capstone project implementation course is an independent study offered during the second quarter of the three-quarters long capstone process. With the guidance of a faculty member, student teams implement the capstone proposal written as part the Research Design for Business Applications course completed during the first quarter. Teams engineer an analytical solution and develop insights from data that would address the problem posed by the client industry partner. Prerequisites: MSCA 31001 Research Design for Business Applications and Approved Capstone Proposal.

MSCA 34001. Capstone Project Writing. 100 Units.
Capstone Project writing is the last course in which teams complete the capstone process by writing a report and developing a presentation that describe the analytical solution they devised to address a problem posed by their client industry partners. Teams submit the report to the program as well as the client partner and present their findings in the MSCA Capstone Showcase at the end of the quarter.
Terms Offered: Autumn Spring Winter
Prerequisite(s): MSCA 31001 Research Design for Business Applications and Approved Capstone Proposal.

MSCA 34002. Capstone I. 100 Units.
The overarching goal of this course is to take students two steps closer to being “Complete Data Scientists”. The first step is by letting students manage and solve a real data science project with real clients and real problems. Students will complete the design of their Capstone Projects, and begin the implementation. The second step is by exposing them to data science methodologies in the absence of pre-existing data - by exposing them to quantitative methodologies in optimally designing data collection tasks. This course covers the Business analytic process from the translation of business problems and opportunities into questions that can be addressed by using data science, development of analytical plans including methodologies and data to address these issues, and initial implementation of these analytical plans.
Instructor(s): Anil Chaturvedi, Gregory Green Terms Offered: Autumn Spring Summer Winter
Prerequisite(s): Prerequisites: 6 Core Courses Completed from MSCA Curriculum. Restricted to MScA students completing the 12-course program curriculum.

MSCA 34003. Capstone II. 100 Units.
The Capstone 2 class is designed to: 1) Provide students maximum flexibility in the latter stages of their Capstone project to work heavily with their Capstone advisors in concluding the execution of the analytic methodology
and any client/sponsor deliverables for the project. 2) Provide maximum support to students in the curation and delivery of key project communications: a) Formal research paper. b) Formal business presentation of project details, value, findings and recommendations. c) Live presentation by the team in Capstone Showcase including question/answer session with a judging panel.

Instructor(s): Donald Patchell

Terms Offered: Autumn Spring Summer Winter

Prerequisite(s): Required: MSCA 34002 Capstone 1. Restricted to MScA students completing the 12-course program curriculum.

**MSCA 37001. Hadoop Workshop. 000 Units.**

This short course is designed to provide a brief, practical introduction to working with data on a Hadoop cluster. The course is aimed at students with no prior knowledge of Hadoop. Topics covered include loading data into Hadoop cluster, using Hive HQL and using Pig script language. Course includes live demos and tutorials so students should complete exercises in class. Students who complete the course will acquire skills to be able to take further studies in Big Data and Text Analytics course.

**MSCA 37002. Linux Workshop. 000 Units.**

This short practical course is designed to provide a brief introduction to Linux operating system. It is aimed at students with no prior knowledge of Linux. Topics covered include uploading files to Linux account, working with files in Linux and managing processes in Linux shell. The course includes live demos and tutorials. Students who complete this tutorial course will acquire skills to be able to take further studies in Big Data and Text Analytics course.

**MSCA 37003. Python Workshop. 000 Units.**

This short course is designed to provide a brief introduction to Python programming language to students with no prior knowledge of Python. Topics covered in the course include Python data types, reading/writing data files, flow control in Python and working with Python modules. The course also introduces Spyder and Jupyter GUIs. Students who complete this introductory course should be able to write and execute simple Python scripts and take further studies in Big Data and Text Analytics course.

**MSCA 37004. RCC Workshop. 000 Units.**

TBD

**MSCA 37005. Tableau Workshop. 000 Units.**

TBD

**MSCA 37006. R Workshop. 000 Units.**

This one-day workshop is an introduction to the essential concepts and techniques for the statistical computing language R. Topics covered include the R and RStudio environment, arithmetic, basic data structure, importing and exporting data, visualization, and basic statistics. No prior R or programming experience is required.

**MSCA 37010. Programming for Analytics. 000 Units.**

This course introduces the essential general programming concepts and techniques to a data analytics audience without prior programming experience. The goal is to equip the students with the necessary programming skill to be successful in the other courses in the MSCA program. Topics covered include: boolean, numbers, loops, function, debugging, R’s specifics (such as list, data frame, factor, apply, RMarkdown), Python’s specifics (such as NumPy, Pandas, Jupyter notebook), version control, and docker. Examples are drawn from the problems and programming patterns often encountered in data analysis. It will use the programming language R in the first part of the course and Python in the second part.

Terms Offered: Autumn Spring

**MSCA 37011. Deep Learning & Image Recognition. 000 Units.**

This course in Deep Learning and Image Recognition will provide a practical, hands-on set of lectures on Deep Learning and Image Processing tools and techniques. It will emphasize practice over advanced mathematical theory, and students will spend a considerable amount of class time gaining experience on Neural Networks and their applications in Python and other open source libraries.

Terms Offered: Autumn Spring

Prerequisite(s): MSCA 31008: Data Mining or MSCA 31009: Machine Learning

**MSCA 37013. Ethics In Big Data Analytics. 000 Units.**

Big data and analytics methodology are enormously beneficial to individuals, corporations, human services organizations, and government. Big data has dramatically improved how we address critically important global challenges, climate change and disease prevention. It has influenced our political life and generated enormous corporate profit. However, the use of huge datasets and data analytical methods raises an array of challenging ethical questions, including: How who owns big data? Are there implications to its sale or transfer? Are there limits to its commercial and public policy use? Is the right of an individual to privacy a thing of the past? Are algorithms inherently biased? Who or what is liable when machines make decisions? In “Ethics in Big Data Analytics” we first explore the impact of data analytics on society and corporations/organizations to establish a framework for understanding ethical challenges. We then investigate ethical issues associated with data collection, storage, transfer/sale, analysis, and visualization. We study bias in algorithms, machine learning, and artificial intelligence. By the conclusion of this short course, students should be able (a) to explain why ethics is important to their work as data analysts/data scientists; (b) to express verbally and in writing the import
of a specific ethical challenge a corporation or organization might confront; and (c) construct and present an argument related to how big data should or should not be used in this situation.

**MSCA 37014. Python for Analytics. 000 Units.**
This course in python starts with introduction to the python programming language basic syntax and environment. It methodically builds up the learner's experience from the level of simple python statements and expressions to writing succinct, efficient and fast Python expressions and package the code in methods and classes. In general, the course is geared toward developing a data science's toolbox such as data importing, cleaning and preparation and covers a number of machine learning algorithms. However the course expands beyond these skills as it stresses upon the importance of some of Python's most unique and powerful features and serves as an introduction to object oriented programming and Python Classes.

**MSCA 37015. Introduction to Ethics in Data Analytics. 000 Units.**
This 3-hour course is an introduction to ethical issues surrounding data analytics, machine learning, and artificial intelligence. As a stand-alone offering, the course has no formal syllabus outlining weekly topics, reading, and assignments. The goal of the course is to offer students a workable, introductory understanding of current ethical challenges they will face in their careers as data science professionals. Ethical and policy-related concepts the course explore include the notion of privacy; data, discrimination, and disparate impact; and algorithmic bias. The course also presents a rudimentary overview of the current regulatory environment in the United States and European Union. Finally, the course introduce students to methods of ethical argument. Understanding these methods will help students communicate a point of view on the ethics of decisions that may be consequential to a business's success. This is a pass/fail course with no advanced reading or assignments. However, the instructor will provide students with a bibliography of materials at the conclusion of the course.

**Terms Offered: TBD**

**MSCA 37016. Advanced Linear Algebra for Machine Learning. 000 Units.**
An advanced linear algebra course focused on the theoretical foundations and applications of linear algebra for machine learning. Upon completion of this course, students will be provided a strong foundation of theoretical linear algebra and linear analysis topics essential for the development of core machine learning and data mining concepts. In addition, various real-life applications of linear algebra for data analytics will be demonstrated.

Instructor(s): Shaddy Abado, Lian Huan Ng, Arnab Bose Terms Offered: Autumn Spring

Prerequisite(s): Successful completion of Undergraduate level coursework in Linear Algebra. Successful completion of the Advanced Linear Algebra for Machine Learning pretest exam with a passing grade, or, successful completion of online Coursera course as outlined and recommended by MSCA program. MSCA & MSAP Students Only.

**MSCA 37017. Advanced Research. 000 Units.**
This course will focus on professional development needs of the data scientist as they work to advance in their career through this masters program. The core areas of research will focus in on how to discover your personal strengths and passions, explore the broad array of jobs that data scientists advance through, and also focus on the companies that may be the best fit for the next stages of each student's career. The focus developed through this research will help prepare the student for the data science marketplace, present themselves confidently, and accelerate the students professional successful career search process.

Instructor(s): Gregory Green Terms Offered: Autumn

Prerequisite(s): Restricted to MSCA and MSAP students only.

**MSCA 37018. Next Chapter: Prospections on Hot Topics in Machine Learning and Artificial Intelligence. 000 Units.**
This 3-week course is designed to help data analytics specialists to stay on top of the most influential developments in the areas of Data Science, Machine Learning and Artificial Intelligence. It covers the newest topics appearing in the curriculum of MScA at University of Chicago as well as foresights of invited leading industry specialists about main trends in the field of analytics. The students attend the course remotely. Activities include live presentations, workshops, individual and group projects and prerecorded videos for asynchronous learning. Upon completion of the course students receive a document certifying completion.

Instructor(s): Yuri Balasanov Terms Offered: Autumn

Prerequisite(s): Required Prerequisites: MSCA 31009: Machine Learning & Predictive Analytics Recommended: MSCA 31013: Big Data Platforms; MSCA 32017: Advanced Machine Learning & Artificial Intelligence Restricted to MSCA and MSAP students only.

**MSCA 37019. Your Career in Data Science. 000 Units.**
This course will help you navigate your career in data science and land a job that fits your needs and desires. It starts with taking a deeper discovery into who you are, clarifying what you want to do with your career, and navigating the market to find the right company and job match.

Instructor(s): Jennifer Schmidt, Gregory Green Terms Offered: Autumn Spring Winter. Winter, Spring, Autumn 2021, Spring 2022

Prerequisite(s): Restricted to MSCA & MSAP students, and MSCA Alumni Scholars only.

**MSCA 40100. Analytics Practicum. 000 Units.**
Analytics Practicum is part of the co-operative educational agreement between MSCA program and employers that provides off-campus work authorization for international students to pursue internships. The internships
must meet the requirement that students archive at least five learning objectives of the course. The learning objectives are about students developing or sharpening their skills in applying analytical tools to solve real life problems.

Terms Offered: Autumn Spring Summer Winter