Data Analytics (MS)

PROGRAM DIRECTOR: Arthur O'Connor
CUNY School of Professional Studies
101 West 31st Street, 7th Floor
New York, NY 10001
Email Contact: Arthur O'Connor, arthur.oconnor@cuny.edu
URL: http://sps.cuny.edu/programs/ms_dataanalytics

THE PROGRAM
The online Master's Degree in Data Analytics (M.S.) prepares graduates to make sense of real-world phenomena and everyday activities by synthesizing and mining big data with the intention of uncovering patterns, relationships, and trends. Big data has emerged as the driving force behind critical business decisions. Advances in our ability to collect, store, and process different kinds of data from traditionally unconnected sources enables us to answer complex, data-driven questions in ways that have never been possible before.

Data analytics combines information management, systems thinking, quantitative methods, data modeling, data warehousing, and data mining to produce visualizations and other business intelligence models that help organizational leaders predict and evaluate best practices. For example:

- Businesses can predict future sales results by combining their customers' preference profiles with website click-stream data, social network interactions, and location data.
- Police and fire departments collaborate with emergency managers and homeland security to develop more accurate models of automotive and pedestrian traffic by using GPS data from cars, buses, taxis, and mobile phones.
- Emergency room physicians are able to reduce time to initial treatment and, as a result, patient mortality, by fusing aggregate patient histories with the results of up to the minute lab tests.

Careers in Data Analytics
Students in the online Master's Degree in Data Analytics that opt for the track in urban sustainability are trained in urban informatics, and emerge prepared for jobs that help make cities "smarter"-for example, by making the electricity grid intelligent, relieving traffic congestion, and reducing building energy consumption. Graduates of the online Master's Degree in Data Analytics are prepared for a variety of technical and managerial positions such as Data Scientist, Business Intelligence Analyst, Knowledge and Informatics Engineer, Data Analyst, Data Mining Engineer, and Data Warehousing Manager.

Admissions Criteria
In addition to the admission criteria for graduate degree programs, applicants must have earned a bachelor's degree in Computer Science, Information Systems, or another STEM field from an accredited institution. A degree in a business-related discipline will be considered on a case by case basis depending on the nature of an applicant's coursework.

Applicants must have the ability to program in a high-level computer language (e.g., Java, C++, Python). Applicants must also have a GPA of 3.0 or better. An admissions interview is required. For more information call 212.652.2869.

CURRICULUM
While the foundational courses lay out four core areas in data analytics (systems, computation, quantitative methods, and data management), the curriculum includes a breadth of cutting edge electives such as business analytics and data mining, web analytics, energy and transportation systems that provide students with options for applying analytic and informatics techniques to a host of issues that that impact the economy and our world.

Program Requirements
To complete the online Master's Degree in Data Analytics, 36 credits are required. Twenty-seven (27) credits in the core and nine credits in elective courses. For the urban sustainability track, all nine of the elective credits must be in the track.

Required Courses

- DATA 600 - Information and Systems
- DATA 602 - Advanced Programming Techniques
- DATA 604 - Simulation and Modeling Techniques
- DATA 605 - Fundamentals of Computational Mathematics

- DATA 606 - Statistics and Probability for Data Analytics
- DATA 607 - Data Acquisition and Management
- DATA 608 - Knowledge and Visual Analytics
- DATA 609 - Mathematical Modeling Techniques for Data Analytics
- DATA 698 - Analytics Master's Research Project

Electives

General Electives
- DATA 610 - Project Management Concepts
- DATA 613 - Managing Innovation and Strategy
- DATA 617 - Data Exploration and Outlier Analysis
- DATA 618 - Quantitative Finance
- DATA 620 - Web Analytics
- DATA 621 – Business Analytics and Data Mining
- DATA 622 - Machine Learning and Big Data
- DATA 624 - Predictive Analytics
- DATA 643 - Special Topics in Data Analytics
- DATA 661 - Independent Study

Specialized Electives in Urban Sustainability
- DATA 611 - Overview of Current Technologies for Sustainability
- DATA 630 - Urban Society and Sustainability
- DATA 644 - Current Topics in Urban Sustainability: Energy
- DATA 645 - Current Topics in Urban Sustainability: Transportation
- DATA 646 - Current Topics in Urban Sustainability: Complex Systems

COURSE DESCRIPTIONS

DATA 600 Information and Systems 3 Credits
Prerequisite: None
Information systems today play an important role within an organization and that role will only grow in the future as data becomes an ever more critical driver of organizational goals. This course introduces students to concepts of information systems and the role of information systems within an organization. Topics covered will include organizational structure and behavior, types of information systems, hardware and software issues, data collection tools and techniques, issues of complexity, and the relevance of information systems to larger social issues like sustainability. The course will provide a review of relevant literature and some case study discussions.

Note: This must be taken in the student's first semester.

DATA 602 Advanced Programming Techniques 3 Credits
Prerequisite: DATA 607
In this course students will learn aspects of contemporary programming that are important for data gathering and analysis, including real-time programming, GUI design, interactive database programming, service-oriented architecture, data collection with and without databases, machine learning, data mining techniques, and GIS programming. Computer security issues will also be addressed, as will overall computer architecture. Students will be required to create a working system for a large volume of data using publically available data sets.

DATA 604 Simulation and Modeling Techniques 3 Credits
Prerequisite: DATA 606
This course teaches students the basics of simulation, systems modeling, and related software applications. Topics include event-driven and agent-based simulations, such as generation of random numbers, random variates, design for simulation experiments, gathering statistics, and steady state versus transient state results. The use of combined simulation and optimization will be covered. Students will develop a contextual understanding of simulation and modeling through the use of case studies.

DATA 605 Fundamentals of Computational Mathematics 3 Credits

Prerequisites: DATA 606 and DATA 607

This course will cover basic differential and integral calculus from the viewpoint of numerical methods and some basic probability concepts. The emphasis will be on modeling and applications to a number of different fields that make use of analytics in differing ways: e.g., business, urban systems, social networks. The course will incorporate basic linear and matrix algebra. Statistical programming and modeling packages will be used throughout.

DATA 606 Statistics and Probability for Data Analytics 3 Credits

Prerequisites: None

This course covers basic techniques in probability and statistics that are important in the field of data analytics. Discrete probability models, sampling from infinite and finite populations, statistical distributions, basic Bayesian statistics, and non-parametric statistical techniques for categorical data are covered in this course. Each of these statistical concepts will be applied in a variety of real-world scenarios through the use of case studies and customized data sets.

DATA 607 Data Acquisition and Management 3 Credits

Prerequisites: None

In this course students will learn about core concepts of contemporary data collection and its management. Topics will include systems for collecting data (real time, sensors, open data sets, etc.) and implications for practice; types of data (textual, quantitative, qualitative, GIS, etc.) and sources; an overview of the use of data, including what and how much should be collected and the distinction between data, information, and knowledge from a data-centric point of view; provenance; managing data with and without databases; computer and data security; data cleaning, fusing, and processing techniques; combining data from different sources; storage techniques including very large data sets; and storing data keeping in mind privacy and security issues.

Students will be required to create a working system for a large volume of data using publically available data sets.

DATA 608 Knowledge and Visual Analytics 3 Credits

Prerequisites: DATA 602

In this course students will learn non-statistical aspects of elucidating from data its information content which leads to knowledge. Several differing visual techniques will be examined to gain this knowledge through exploratory use of visualizations as well as visualization techniques for presenting data to a variety of stakeholders. Exploratory techniques look to find patterns in the data. Finding patterns that underlie the system's characteristics when the data sets are very large or have many dimensions by reducing the dimensionality in intelligent ways is a complex task that often includes user direction. Presentation visualizations provide the viewer with useful information and knowledge since the visualizations are created with context in mind. In addition, students will learn how to integrate quantitative and qualitative data (e.g., text and narrative).

DATA 609 Mathematical Modeling Techniques for Data Analytics 3 Credits

Prerequisites: DATA 602 and DATA 605

In this course students will learn mathematical methods for understanding data relationships and for system optimization. Mathematical modeling techniques for representing a complex system will be presented. Topics to be covered include linear (LP) and non-linear programming (NLP); algorithmic search methods for optimization; branch and bound and dynamic programming, and their uses. Use of modeling packages will be stressed. Examples will be used from actual systems. In addition, students will be expected to explain their models, reports, and analyses in plain and easy-to-understand language.

DATA 610 Project Management Concepts 3 Credits

Prerequisite: None

Students in this course learn to plan, organize, lead, and control software projects to ensure that they meet requirements and are delivered on time and within budget. Students learn the essentials of defining requirements, scheduling, budgeting, managing complex teams and distributed work, communications, conflict resolution, and staff development.

DATA 611 Overview of Current Technologies for Sustainability 3 Credits

Prerequisite: DATA 607

This research course uses a case study format to examine the underlying technologies that offer potential for improving urban sustainability and enabling well connected and intelligent cities. Areas of study may include sensors and actuators; transportation systems; building control systems; electric power control systems; renewable energy delivery systems;

analytics and optimization for decision-making, sustainability policy, and complex systems of systems. Current papers discussing real-life examples from urban areas around the world will be used. This course ties in aspects of behavioral economics, psychology, sociology, social media, and urban design and explores the nature of human interaction with systems. Guest speakers from New York City government and industry will enrich the student experience.

DATA 613 Managing Innovation and Strategy 3 Credits

Prerequisite: None

This course has a dual focus. First, it prepares students to understand the nature of technical change in both information systems and technologies that are at the forefront of current practice. Second, the course explores current business models and product strategies that will drive market trends. Throughout the course students are responsible for analyzing how technical changes—many of which are specific to information systems—impact the populations affected by a new technology.

DATA 617 Data Exploration and Outlier Analysis 3 Credits

Prerequisite: DATA 606 and DATA 607

In this course, students will develop advanced skills in exploring and processing large sets of disparate data types. Students will perform exploratory analysis, work with imperfect data sets, apply probabilistic techniques to the characterize variables, and identify and handle outliers. In addition, students explore relationships between variables and apply appropriate transformations to these variables.

DATA 618 Quantitative Finance 3 Credits

Prerequisite: DATA 606 and DATA 607

Quantitative finance is a branch of applied mathematics concerned with calculation, modeling, and forecasting in a variety of industry segments. Professionals in this field use specialized knowledge and skills to determine value and calculate risk. Their results can play a key role in business actions such as financing, mergers, consolidations, speculation, and global expansion Students will engage in topics that include probability distributions, linear regression, stochastic calculus, Monte Carlo methods, Black-Scholes, capital asset pricing, and arbitrage pricing. Topics will be presented through academic theory and real-world examples.

DATA 620 Web Analytics 3 Credits

Prerequisites: DATA 606 and DATA 607

Organizations, both commercial and community, can benefit from deep analysis of their website interactions and mobile data. Social networks have also become a source of information for companies; search engines are an important referral mechanism. Popular social networks and other online communities provide rich sources of user information and (inter-) actions through their application programming interfaces. This data can help to identify a number of individual user preferences and behaviors, as well as fundamental relationships within the community. Search engines use algorithms to rank sites. Students will learn how to analyze social network data for types of networks, the fundamental calculations used in social networks (e.g., centrality, cohesion, affiliations, and clustering coefficient) as well as network structures and roles. Beyond social network data, students will learn about important concepts of analyzing website traffic such as click streams, referrals, keywords, page views, and drop rates. The course will touch on the fundamentals of search algorithms and search engine optimization. To provide a basic context for understanding these online user and community behaviors, students will learn about relevant social science theories such as homophily, social capital, trust, and motivations as well as business and social use contexts. In addition, this course will address ethical and privacy issues as they relate to information on the internet and social responsibility.

DATA 621 Business Analytics and Data Mining 3 Credits

Prerequisites: DATA 606 and DATA 607

This course teaches students to comb through complex business data sets to produce knowledge, and ultimately, business intelligence. Students learn the basics of business analytics. However, this course goes well beyond typical analytics for managers by including rich computational components for predictive and prescriptive analytics. Strategy and operational business contexts are provided via case studies throughout the course. Students will deal with actual business scenarios like sales, marketing, logistics, and finance. Students are expected to bring in practical problems from their own fields of interest. In addition, each student will be responsible for leading discussions in a particular application area. Teamwork is an essential part of the course.

DATA 622 Machine Learning and Big Data 3 Credits

Prerequisite: DATA 621

This course teaches students to apply advanced machine learning techniques to big data sets. Students will learn how to apply both new and previously studied techniques to large data sets in a distributed computing environment. In particular, the course will make use of the Hadoop framework and the Mahout implementation of machine learning algorithms. Students will also learn to apply basic text mining techniques as well as how to implement a basic recommender system in Hadoop.

DATA 624 Predictive Analytics 3 Credits

Prerequisite: DATA 621

This course teaches students to use advanced machine learning techniques that are focused on predictive outcomes. Topics will include time series analysis and forecasting, recommender systems, and advanced regression techniques. In addition, students will learn how to evaluate the predictions that result from these techniques, how to assess model quality, and how to improve models over time.

DATA 630 Urban Society and Sustainability 3 Credits

Prerequisite: DATA 600 and DATA 605

The course introduces students to concepts and practices of sustainability in cities. Key objectives are to review and critique how sustainability planning is being carried out, to identify the barriers and bridges to its effective implementation, and to identify the technologies and metrics of success being used to create, catalog, and understand the progress made. A related objective is to analyze the urban systems being impacted by sustainability planning and practices, and how those systems have been modeled. Furthermore, students will reflect on and discuss the impact of sustainability projects on people's lives. The course includes a review of relevant literature and extended case study discussions. Topics include: urbanization and resource utilization; society and cities; systems and the built environment; resources; environmental management; green businesses.

DATA 643 Special Topics in Data Analytics 3 Credits

Prerequisite: DATA 602 and DATA 606

This course allows the program to offer additional material on advanced and specialty topics within the Data Science field. This will be an advanced class. Emphasis will be placed on project based outcomes.

DATA 644 Current Topics in Urban Sustainability: Energy 3 Credits

Prerequisite: DATA 607

This course will cover the in detail the most up-to-date trends in energy distribution, consumption, monitoring, and conservation, including building control systems. Modeling and economic considerations will be a focal part of the course. Emphasis will be placed on software that is currently available for energy distribution, building usage, and conservation. Topics will vary, sometimes with a particular emphasis.

DATA 645 Current Topics in Urban Sustainability: Transportation 3 Credits

Prerequisite: DATA 607

The course will cover the most up-to-date trends in urban transportation systems, including both mass transit and surface transportation issues in an in-depth manner. Trends that rely on information systems, such as congestion pricing, peak demand parking, real-time transit information, and priority signaling, among others, will be considered. Emphasis will be placed on software and hardware implications.

DATA 646 Current Topics in Urban Sustainability: Complex Systems 3 Credits

Prerequisite: DATA 607

The course will cover the most up-to-date trends in urban systems and their interrelationships in an in depth manner. Emphasis will be placed on software and hardware implications.

DATA 661 Independent Study 1 - 3 Credits

Prerequisite: None

This course allows the program to offer additional material in the broad area of Information Systems after the student has gained a general background from the Prerequisites. This will be an advanced class. Topics might include: effects of internationalism on information systems (language considerations, distributed program creation techniques across time zones, etc.), cooperative information systems, security, threats, internet considerations, filtering, GUI design considerations. Emphasis will be placed on the software and hardware associated with the information systems.

DATA 698 Analytics Master's Research Project 3 Credits

Prerequisites: Permission of Academic Director

In this course, students will integrate the knowledge and skills derived from the previous classes into a real-world project. Working in small teams (that may be geographically distributed) or by themselves, students will work on designing an information system.

With the oversight of a faculty advisor, students will identify a topic, develop a research plan, conduct research, and collect and analyze data. The project may be organized in collaboration with a partner organization, for example, a local company, non-profit, or research lab.