DATA ANALYTICS

Description: Data analytics involves predicting and evaluating business decisions for all functional business areas with an analytic mindset that is crucial in data-driven environments. The data analytics module introduces quantitative methods used to analyze business data and to develop data-based insight for strategic decision making. With exposure to cutting-edge analytics software, the module will cover data mining techniques, sentiment analysis, text mining and predictive analytics that exploit the massive amount of data generated from business interactions and social exchanges. The data analytics knowledge and skill set will be valuable for a wide range of for-profit and non-profit organizations in business, government, education, and other settings.

Courses:

1) **Business Analytics 53:716:502**
   This course provides an overview of business analytics, emphasizing how firms implement data-driven decision making. Students will learn an array of statistical concepts and different analytics methodologies, use spreadsheet modeling and learn through a mix of lectures, cases, practice problems and class discussions. The course will also provide opportunities to apply these concepts hands-on using a language such as "R". An important goal of the course is to make students understand and implement fact-based decision making and to enable them to become comfortable in gathering and analyzing data in order to make managerial decisions. Topics include exploratory data analysis, sampling, hypothesis testing, regression modeling, experimental design, analysis of variance, text mining, web analytics, and social media analytics. Case studies and assignments will introduce students to various contemporary business applications and innovative use of these ideas. *Prerequisite: Quantitative Skills 53:135:502 unless waived.*

2) **Data Management & Business Intelligence 53:623:517**
   This course focuses on the design and management of the data resources of an organization and the extraction of business intelligence from the data for managerial decision making. The basic concepts and techniques of data management and mining data will be examined with real-world examples and cases to place these techniques in proper context. Data management module covers the fundamentals of robust data models, management of data resources and the retrieval of data using SQL/NoSQL. Business Intelligence module focuses on the concepts and applications of various data mining models such as classification and clustering, and provides an overview of machine learning. The course delivers adequate technical detail with hands-on training, while emphasizing the interpretation, organizational and implementation issues relevant to managers. Topics include conceptual data modeling, database management, structured query language, business intelligence, classification, clustering, and introduction to machine learning. *Prerequisite: Business Analytics 53:716:502.*

3) **Big Data Analytics & Visualization 53:716:535**
   The course provides students with an in-depth introduction to the Hadoop ecosystem, which is an environment used by companies to store and manipulate "Big Data" of a size and scale that cannot be fit into traditional databases. Students will be exposed to tools such as Sqoop, Pig and Hive - for ETL (extract-transform-load) operations within Hadoop. The course also provides exposure to state of the art data mining algorithms for clustering, classification and collaborative filtering (a technique used by businesses to recommend products to customers). Participants will build machine learning applications with Apache Spark. In the second half of the

course, students will learn information design, specifically tactics for visualizing Big Data that can consist of temporal, spatial, topical or network relationships among entities. *Prerequisite: Data Management & Business Intelligence 53:623:517. Also, students should have familiarity with the Python programming language before taking this course.*

4) **Social Media Analytics & Sentiment Analysis 53:716:540**
The course enables students to ingest Big Data from APIs for social media platforms such as Twitter. After assembling data from social media, students learn to analyze the data to gain business insights. Concepts for the analysis of social media, such as community detection and assignment, node centrality, information diffusion and opinion formation will be presented. The course also addresses analyzing unstructured text information using open source software. As an application of text analytics, the course explores the issues of analyzing product reviews and opinion mining. Students will learn the process for sentiment extraction, opinion mining and recognizing opinion spam. *Prerequisite: Data Management & Business Intelligence 53:623:517. Also, students should have familiarity with Python programming language before taking this course.*