

Data Science Online Master's Course Descriptions

Data Science: Orientation

Delivered live during the required campus-based launch immersion in South Bend, this course provides an introduction and overview of the processes of solving data science problems, and a foundational treatment of data standards, management, integration, annotation, and archiving. The course focuses on enhancing the students' conceptual understanding of information, data, and databases.

Probability & Statistics for Data Science

This first-semester course builds the statistical foundations for further work in data science, with a specific focus on statistical thinking in data collection, data quality analysis, probability theory, statistical inference, and modeling. Fall Semester. (3 credits)

Systems and Technologies: R

This course outfits students with the technical and practical skills required for working with modern data systems and technologies. Students learn how to use the R programming language for data manipulation, data cleaning, visualization, and exploratory data analysis. Students will build on the skills developed in this course throughout the program. Fall Semester. (1.5 credits)

Systems and Technologies: Python

This course introduces students to the Python programming language and its application in Data Science. Students learn the practical aspects of data manipulation and cleaning with Python and are introduced to libraries designed for data exploration and modeling. Students will build on the skills developed in this course throughout the program. Fall Semester. (1.5 credits)

Introduction to Data Science

Building on the quantitative foundations established in the first semester, this course introduces students to the entire process and lifecycle of data science, including data acquisition, data visualization, data quality analysis, relevant machine learning methods, communicating results, aspects of deploying and monitoring the models, and the ethical considerations in managing and processing data. Throughout the course, students implement and experiment with the concepts and methods of the data science process, and apply them to real-world datasets. Spring Semester. (3 credits)

Linear Models

This course trains students in applied linear regression modeling. Beginning with an introduction to fundamental concepts in regression model building and inference, the course then delves into advanced techniques such as ridge regression and lasso. Spring Semester. (3 credits)

Databases & Data Security

Calibrated to data science applications, this course focuses on effective techniques in designing relational databases and retrieving data from them using both SQL and R. It provides an introduction to relational databases, including topics such as relational calculus and algebra, integrity constraints, distributed databases, and data security. Students are introduced to database technologies utilized in industry, such as NoSQL, graph databases, and Hadoop. The course also introduces students to the fundamental concepts of cybersecurity and privacy relevant to data science. Summer Semester. (3 credits)

Storytelling & Communications for Data Scientists

This course is designed to develop communication skills for data scientists working in industry and business contexts. Students master the art of clear, effective, and engaging scientific and technical communications, with attention to the business necessity of translating complex technical subjects into actionable insights for a lay audience. Students identify and analyze rhetorical situations in technical discourse communities, assist them in defining their purpose in writing/presenting information, and teach them to design materials and deliver presentations that are properly targeted and appropriately styled. Summer Semester. (1.5 Credits)

Ethics and Policy in Data Science

Data-informed decision making has created new opportunities, e.g. personalized marketing and recommendations, but also expands the set of possible risks, e.g. privacy, security, etc.; this is especially true for businesses collecting, storing, and analyzing human data. Organizations need to consider the "should we?" question with regard to data and analytics, and not just be concerned with "can we?". In this course, we will explore ethical frameworks, guidelines, codes, and checklists, and also consider how they apply to all phases of the data science process. Existing research ethics standards provide a necessary but insufficient foundation when doing data science and analytics. Together, we will wrestle with the rapidly-changing capabilities, conflicts, and desires that emerge from new data practices. Upon completion of the course, you will be able to identify and balance: what an organization wants to do from a business perspective, can do from technical and legal perspectives, and should do from an ethical perspective. Summer Semester. (1.5 Credits)

Behavioral Data Science

Behavioral Data Science provides students the opportunity to explore sources and types of behavioral data and empowers students to select and use appropriate tools for finding answers to questions about human behavior. Students will work with a variety of data models and theories, like factor analysis, item response theory, centroid clustering models, recommender systems, and topic models using a wide variety of data (e.g., traffic violations, crime, and video game data). Fall Semester. (3 credits)

Statistical Learning for Data Science

This course focuses on advanced statistical learning methods and will build on earlier material on model building and machine learning. Topics covered include classification (discriminant analysis, Bayesian inference, density estimation), tree-based and ensemble methods (random forest, boosting, bagging), support vector machines, neural networks, unsupervised learning (principal component analysis, nearest neighbor, k-means clustering, hierarchical clustering). Fall Semester. (2 credits)

Data Visualization

This course focuses on methods of visualizing data for exploration, reporting, and monitoring tools, such as dashboards. Students are introduced to computational tools for building interactive graphics as well as commercial visualization software. The role of visualization in storytelling will be emphasized. Fall Semester. (1 credit)

Data Science Now: Industry, Cases, and Projects

This course teams groups of students with industry partners to solve real data science problems. Data Science Now asks students to solve data science problems in an integrated fashion as a simulation of the live conditions of work as a professional data scientist. Student teams carry out all steps of the data science process: data acquisition, modeling, analysis, and communication of results. Spring Semester. (3 credits)

Generalized Linear Models

This course examines extensions and generalizations of the linear regression model. Specifically, methods for fitting and evaluating logistic, multinomial, and count response models are presented using examples from a wide variety of fields. Bootstrapping, cross-validation and penalized estimation are woven throughout the coverage. Spring Semester. (2 credits)

Time Series and Forecasting

Focusing on applications, in this course students study time series models and computational techniques for model estimation, model diagnostics and forecasting. Spring Semester. (1 credit)

Optional Immersion Weekend

This intensive experience will be held in person and will combine instruction and project work with industry roundtables, site visits, and professional topics suited to the location. These weekends are held on the Notre Dame campus and other locations around the United States.