SCHOOL OF DATA SCIENCE AND SOCIETY (GRAD)

The School of Data Science and Society (SDSS) is devoted to data science teaching, research, and service. Our vision is to lead by shaping the emerging field of data science with a human-centric approach to the entire data life cycle. SDSS envisions a world made healthy, safe, and prosperous for all through data-informed decisions.

The SDSS will empower a diverse community of faculty conducting research in the fundamentals and/or applications of data science. The school will train undergraduate, graduate, and professional students to be the next generation of data science leaders with the knowledge and skills to thrive in this data-driven world. The SDSS will serve the state, the nation, and society with premier data science educational programs and collaborative research directed to advance the public good.

The School of Data Science and Society website (https://datascience.unc.edu/) has additional information for prospective students.

Master of Applied Data Science Program (MADS)

The online Master of Applied Data Science (MADS) program is offered by the School of Data Science and Society, in collaboration with the School of Information and Library Science, the Department of Biostatistics in the Gillings School of Global Public Health, and the departments of computer science, mathematics, and statistics and operations research in the College of Arts and Sciences.

The program is delivered through interactive, live online classes; asynchronous lessons; in-person immersions; and a real-world, team-based capstone project. The program provides recent graduates and working professionals with a comprehensive understanding of the data life cycle; technical expertise in areas such as programming and machine learning; and opportunities to connect with industry professionals in North Carolina and beyond. Students will graduate prepared to identify and tell a story through data; work collaboratively to apply data-driven insights; and directly impact lives in their workplaces and communities.

***All MADS courses are listed as DATA 700's except for DATA 992 Capstone.

To view the full list of MADS faculty members, please visit the School of Data Science and Society Web site (https://datascience.unc.edu/mads-faculty/).

Professors

David Adalsteinsson
Ashok Krishnamurthy
Arcot Rajasekar

Associate Professors

Junier Oliva
Santiago Olivella
Danielle Szafir
Francesca Tripodi

Assistant Professors

Emily Pfaff
Shashank Srivastava

Teaching Assistant Professor

Rei Sanchez-Arias

Adjunct Professor

Chuck Pepe-Ranney

Professor of the Practice

Richard Marks

Graduate-Level Courses

DATA 710. Introduction to Applied Data Science. 3 Credits.
The first part of this course introduces various stages of the data life cycle, from defining data requirements to data creation and gathering to data fusion and data preparation to data cleaning and quality control to exploratory analytics, data interpretation, and visualization. We will explore FAIR data principles of curation, metadata, and digital preservation policies. The second part will introduce the concept of relational databases that provide storage and management for structured data.

Rules & Requirements
Grading Status: Letter grade.

DATA 715. Advanced Databases for Data Science. 3 Credits.
This course will explore intermediate-level design and implementation of database systems, emphasizing scalable, distributed systems. It will deepen students' knowledge of advanced relational database management and discuss current and emerging practices for dealing with big data and large-scale database systems. Concepts include design and implementation of relational databases, exploration of distributed data structures including graph, document, and key-value storage models and scalable and resilient query processing.

Rules & Requirements
Requisites: Prerequisite, DATA 710.
Grading Status: Letter grade.

DATA 720. Programming Methods for Data Science. 3 Credits.
This course will provide students with advanced concepts on the construction and use of data structures and their associated algorithms. Concepts covered in this course will include: abstract data types, lists, stacks, queues, trees, and graphs; sorting, searching, hashing, and an introduction to numerical error control; techniques of algorithm analysis and problem-solving paradigms using relevant programming languages and tools.

Rules & Requirements
Grading Status: Letter grade.

DATA 730. Statistical Modeling and Inference for Data Science. 3 Credits.
The course will be coding-oriented and cover concepts such as foundations in probability, including basic rules, Bayes' theorem, and basic distributions; sampling and the central limit theorem; bootstrapping, confidence intervals, hypothesis testing, and multiple testing; linear models, basic and multiple regression, inference for regression, regularization; classification, logistic regression, and tree-based methods; and prediction, model interpretation, and model evaluation.

Rules & Requirements
Grading Status: Letter grade.

DATA 735. Applied Causal Inference Methods in Data Science. 3 Credits.
This course equips participants with practical tools to estimate causal effects in real-world settings. After building a solid formal foundation, students will learn to design experiments, leverage natural experiments, and analyze observational data using modern causal inference methods. Ideal for those who want to move beyond predictive analytics in order to answer causal questions in their work.

Rules & Requirements
Requisites: Prerequisites, DATA 710 and DATA 730.
Grading Status: Letter grade.

DATA 740. Governance, Bias, and Ethics in Data Science and Artificial Intelligence. 3 Credits.
We will explore the foundational concepts of ethics in data science and AI. This overview will set the stage for a deep understanding of what ethical frameworks mean in practice, providing students the opportunity to create actionable examples. By focusing on a wide variety of case studies throughout a myriad of industries and settings, this class will develop leaders who can effectively integrate and leverage data science solutions while ensuring responsible use of data.

Rules & Requirements
Requisites: Prerequisite, DATA 710.
Grading Status: Letter grade.

DATA 750. Mathematical Tools for Data Science. 3 Credits.
This course will present the mathematical intuition, theory, and techniques driving the numerical computation methods used for processing and analyzing data in various real-life problems. Topics include dimensionality reduction; linear and non-linear approximation; frequency and wavelet analysis; and a glimpse into the mathematics of deep neural networks, classification, large-scale and high-performance numerical computing, and visualization.

Rules & Requirements
Requisites: Prerequisite, DATA 710.
Grading Status: Letter grade.

DATA 760. Visualization and Communication in Data Science. 3 Credits.
This course will provide students with a foundational understanding of visual perceptional and data visualization design practices, provide instruction on using visualization for tasks such as exploratory analysis and storytelling to support both data-driven discovery and communication. The class will focus hands-on experiences with commonly used data science tools and technologies.

Rules & Requirements
Requisites: Prerequisite, DATA 710.
Grading Status: Letter grade.

DATA 766. Leading Research Teams. 3 Credits.
Graduate students will lead groups of four to five undergraduate students to complete a project for an external client. Clients from an array of industries, as well as those from within UNC, will submit project requests that are feasible within a 12-week semester. The course will offer graduate students an authentic learning experience to develop management skills and professional attributes, such as teamwork, communication, and project management, in preparation for workforce entry.

Rules & Requirements
Requisites: Prerequisites, STOR 664, STOR 665 and STOR 765, or Instructor's Permission.
Grading Status: Letter grade.
Same as: STOR 766.

DATA 780. Machine Learning. 3 Credits.
This course will be an introductory course to machine learning (ML). The course will cover core principles of artificial intelligence for statistical inference and pattern analysis. Topics will include probability distributions; graphical models; optimization, maximum likelihood estimation, and regression; classification; cross validation; generalization and overfitting; neural networks; nonparametric estimators; clustering; autoencoders; generative models; and kernel methods. Applications in tabular, image, and textual data for supervised and unsupervised learning tasks also will be covered.

Rules & Requirements
Requisites: Prerequisites, DATA 710 and DATA 720 and DATA 730.
Grading Status: Letter grade.

DATA 785. Deep Learning. 3 Credits.
Deep learning fundamentals and applications with emphasis on their broad applicability to problems across a range of disciplines. Topics include regularization, optimization, convolutional networks, sequence modeling, generative learning, instance-based learning, and deep reinforcement learning. Students will complete several substantive programming assignments in PyTorch and Keras.

Rules & Requirements
Requisites: Prerequisites, DATA 720, DATA 730, and DATA 780.
Grading Status: Letter grade.

DATA 890. Special Topics in Data Science. 3 Credits.
The course goal is to expose graduate students in any UNC department to a broad range of topics in the theory and applications of data science. Students will learn about current and emerging methods and techniques in data science to advance individual research efforts and facilitate inter-disciplinary collaboration. Open to graduate students only and by permission only.

Rules & Requirements
Repeat Rules: May be repeated for credit; may be repeated in the same term for different topics; 9 total credits. 3 total completions.
Grading Status: Letter grade.

DATA 992. Master's (Non-Thesis). 3 Credits.
Team based project in final term of program.

Rules & Requirements
Requisites: Prerequisites, DATA 710, DATA 715, DATA 720, DATA 730, DATA 750, and DATA 780.
Repeat Rules: May be repeated for credit.

DATA 993. Master's Research and Thesis. 3 Credits.
Rules & Requirements
Repeat Rules: May be repeated for credit.

DATA 994. Doctoral Research and Dissertation. 3 Credits.

Contact Information

Program Director
Emma Dehne
emma.dehne@unc.edu