

Data Science (DATA)

COURSES

DATA 1301. INTRODUCTION TO DATA SCIENCE. 3 Hours.

This course provides an introduction to the field of data science with a high level overview of basic concepts, data types, and techniques while introducing data-informed decision making.

DATA 2301. DATA VISUALIZATION & DATA STORYTELLING. 3 Hours.

In this course, students will learn about the principles and techniques of effective data visualization and communicate data using various tools and techniques. The course will cover a variety of tools and techniques commonly used in data visualization, including data visualization software, data preparation and cleaning, and techniques for creating static and interactive visualizations. Through a combination of lectures, hands-on exercises, and individual and group projects, students will explore the principles of data visualization and gain practical skills in designing and communicating data through visual representations. Topics covered will include data exploration, charting and graphing, dashboard design, data storytelling, and best practices for visualizing data. At the end of this course, students can design and implement effective visualizations for various fields, including STEM, psychology, scientific research, and so many others.

DATA 3311. MATHEMATICS FOR DATA SCIENCE. 3 Hours.

This course covers techniques from linear algebra and probability with an emphasis on how they are used in data science. Working with real data sets will be emphasized, along with basics of Matlab or R programming. Prerequisite: MATH 1426.

DATA 3401. PYTHON FOR DATA SCIENCE 1. 4 Hours.

This is the first of a two course sequence offering the foundations of Python programming in the context of data science. It introduces the full syntax of the Python language as it overviews structured, functional, and object oriented programming methodologies. It also provides a basic conceptual understanding of computing and introduces Unix command-line tools, software employed in data science such as git and Jupyter, and Python libraries such as numpy, matplotlib, and Pandas. Prerequisite: MATH 1426 or concurrent enrollment in MATH 1426.

DATA 3402. PYTHON FOR DATA SCIENCE 2. 4 Hours.

This is the second of a two course sequence offering the foundations of Python programming in the context of data science. It reinforces concepts presented in DATA 3401 with greater depth with a focus on application to various problems in data science, while exploring the python library ecosystem. Prerequisite: DATA 3401, or consent of instructor.

DATA 3421. DATA MINING, MANAGEMENT, AND CURATION. 4 Hours.

This lecture and lab course will provide training in working with databases, including data mining techniques and principles and best practices in data management, storage, and curation. Prerequisite: DATA 3402 or concurrent enrollment in DATA 3402, or consent of instructor.

DATA 3441. STATISTICAL METHODS FOR DATA SCIENCE 1. 4 Hours.

This lecture and lab course will provide an introduction to the fundamental building blocks of advanced data analysis, with emphasis on advanced linear algebra, optimization, statistical inference, and Monte Carlo methods. Working with real data sets will be emphasized, along with basics of R programming. Prerequisite: DATA 3401 or consent of instructor.

DATA 3442. STATISTICAL METHODS FOR DATA SCIENCE 2. 4 Hours.

This lecture and lab course will provide an introduction to the principles and general methods for the analysis of categorical data. This type of data occurs extensively in both observational and experimental studies, as well as industrial applications. While some theoretical statistical detail is given, the primary focus will be on methods of data analysis. Topics include generalized regression models, logistic regression models, Poisson regression models, and multinomial regression models. Problems will be motivated from a scientific perspective. Prerequisite: DATA 3441.

DATA 3461. MACHINE LEARNING. 4 Hours.

This course introduces and surveys Machine Learning techniques and their application to various problems in data science. Prerequisite: DATA 3401, DATA 3402 or consent of instructor.

DATA 4090. UNDERGRADUATE RESEARCH. 0 Hours.

Undergraduate research experiences under supervision of faculty. Students are expected to disseminate research findings by poster or oral presentations in meetings or conferences. Students are also expected to participate in other activities as directed by the grant-funded Research Program Director.

DATA 4380. DATA PROBLEMS. 3 Hours.

This course equips Data Science students with the skills to identify, define, and explore various data science challenges. Students will engage with practical projects and expert guidance throughout the course, preparing them for the capstone project. This course includes advanced data handling, machine learning algorithms, model evaluation, and ethical considerations in data science, culminating in real-world project presentations. Prerequisite: DATA 3402, DATA 3421. DATA 3461 or current enrollment in DATA 3461, or permission of the instructor.

DATA 4381. DATA CAPSTONE PROJECT 1. 3 Hours.

This is the first of a two-semester sequence that will involve deep engagement in a team or individual project in Data Science. Presentation of written and oral reports will be required. Corequisite: DATA 4380.

DATA 4382. DATA CAPSTONE PROJECT 2. 3 Hours.

This is the second of a two-semester sequence that will involve deep engagement in a team or individual project in Data Science. Presentation of written and oral reports will be required. Prerequisite: DATA 4381.

DATA 4390. DATA SCIENCE RESEARCH. 3 Hours.

Formulation and definition of research problems, the formulation and execution of strategies of solution, and the presentation of results. Prerequisite: consent of instructor. Recommendation by other faculty encouraged.

DATA 4391. SPECIAL TOPICS IN DATA SCIENCE. 3 Hours.

Special topics in Data Science are assigned to individuals or small groups. Faculty members closely supervise the projects and assign library reference material. Small groups will hold seminars at suitable intervals. May be repeated for credit. Prerequisite: senior standing and written permission of the instructor & department chair.

DATA 4392. ADVANCED TOPICS IN DATA SCIENCE. 3 Hours.

Varies from semester to semester. New developments in Data Science, in-depth study of a topic not covered in other courses, or a special faculty expertise made available to undergraduates. May be repeated for credit as topic varies. Prerequisite: permission of instructor.

DATA 4393. HONORS THESIS/SENIOR PROJECT. 3 Hours.

Required of all students in the University Honors College. During the senior year the student must complete a thesis or a project under the direction of a faculty member in Data Science. Prerequisite: Enrollment in the University Honors College and written permission of the instructor and chair.

DATA 4394. UNDERGRADUATE RESEARCH EXPERIENCES. 3 Hours.

Research under faculty supervision and mentorship involving collaboration within a small group. The topic varies from semester to semester, is determined by the faculty teaching the course, and is announced in advance. The course promotes active learning based on inquiry, development of higher-order thinking skills, and meaningful scientific research. Prerequisite: consent of instructor.