

# DATA SCIENCE (DS-UA)

---

## DS-UA 111 Data Science for Everyone (4 Credits)

Data Science for Everyone is a foundational course that prepares students to participate in the data-driven world that we are all experiencing. It develops programming skills in Python so that students can write programs to summarize and compare real-world datasets. Building on these data analysis skills, students will learn how to draw conclusions and make predictions about the data. Students will also explore related ethical, legal, and privacy issues.

**Grading:** CAS Graded

**Repeatable for additional credit:** No

## DS-UA 112 Principles of Data Science (4 Credits)

Principles of Data Science offers the fundamental principles and techniques of data science. Students will develop a toolkit to examine real world examples and cases to place data science techniques in context, to develop data-analytic thinking, and to illustrate that proper application is as much an art as it is a science. In addition, students will gain hands-on experience with the Python programming language and its associated data analysis libraries. Students will also consider ethical implications surrounding privacy, data sharing, and algorithmic decision making for a given data science solution.

**Grading:** CAS Graded

**Repeatable for additional credit:** No

**Prerequisites:** DS-UA 111.

## DS-UA 201 Causal Inference (4 Credits)

Causal Inference provides students with the tools for understanding causation, i.e., the relationship between cause and effect. We will start with the situation in which you are able to design and implement the data gathering process, called the experiment. We will then define causation, identify preconditions required for A to cause B, show how to design perfect experiments, and discuss how to understand threats to the validity of less-than-perfect experiments. In this course, we will cover experimental design and then turn to those careful approaches, where we will consider such approaches as quasi-experiments, regression discontinuities, differences in differences, and contemporary advanced approaches.

**Grading:** CAS Graded

**Repeatable for additional credit:** No

## DS-UA 202 Responsible Data Science (4 Credits)

The first wave of data science focused on accuracy and efficiency: on what we can do with data. The second wave is about responsibility: what we should and should not do. Accordingly, this technical course tackles the issues of ethics and responsibility in data science, including legal compliance, data quality, algorithmic fairness and diversity, transparency of data and algorithms, privacy, and data protection. An important feature of this course is its holistic treatment of the data science lifecycle, beginning with data discovery and acquisition, through data cleaning, integration, querying, analysis, and result interpretation.

**Grading:** CAS Graded

**Repeatable for additional credit:** No

## DS-UA 203 Machine Learning for Language Understanding (4 Credits)

This course covers widely-used machine learning methods for language understanding—with a special focus on machine learning methods based on artificial neural networks—and culminates in a substantial final project in which students write an original research paper in AI or computational linguistics. If you take this class, you'll be exposed only to a fraction of the many approaches that researchers have used to teach language to computers. However, you'll get training and practice with all the research skills that you'll need to explore the field further on your own. This includes not only the skills to design and build computational models, but also to design experiments to test those models, to write and present your results, and to read and evaluate results from the scientific literature.

**Grading:** CAS Graded

**Repeatable for additional credit:** No

## DS-UA 204 Practical Training (2-4 Credits)

Provides data science students with an opportunity to apply the knowledge gained in their course work to practical problems in industry. This course is for majors and minors only.

**Grading:** CAS Pass/Fail

**Repeatable for additional credit:** Yes

## DS-UA 300 Special Topics in Data Science (4 Credits)

Topics and prerequisites vary by semester

**Grading:** CAS Graded

**Repeatable for additional credit:** No

## DS-UA 301 Advanced Topics in Data Science (4 Credits)

Advanced Topics in Data Science exposes students to two specialized topics within Data Science: Examples of topics include time series, deep learning, and other advanced machine learning topics. Students will learn the theoretical underpinnings of advanced data science techniques, as well as engage in hands-on activities to build a practical toolkit.

**Grading:** CAS Graded

**Repeatable for additional credit:** Yes

## DS-UA 9111 Data Science for Everyone (4 Credits)

Data Science for Everyone is a foundational course that prepares students to participate in the data-driven world that we are all experiencing. It develops programming skills in Python so that students can write programs to summarize and compare real-world datasets. Building on these data analysis skills, students will learn how to draw conclusions and make predictions about the data. Students will also explore related ethical, legal, and privacy issues.

**Grading:** CAS Graded

**Repeatable for additional credit:** No

## DS-UA 9201 Causal Inference (4 Credits)

*Typically offered Spring*

Causal Inference provides students with the tools for understanding causation, i.e., the relationship between cause and effect. We will start with the situation in which you are able to design and implement the data gathering process, called the experiment. We will then define causation, identify preconditions required for A to cause B, show how to design perfect experiments, and discuss how to understand threats to the validity of less-than-perfect experiments. In this course, we will cover experimental design and then turn to those careful approaches, where we will consider such approaches as quasi-experiments, regression discontinuities, differences in differences, and contemporary advanced approaches.

**Grading:** CAS Graded

**Repeatable for additional credit:** No