DATA SCIENCE (DS-UA)

DS-UA 100 Survey in Data Science (4 Credits)
Typically offered Fall and Spring
Data science is a relatively new discipline that is radically reshaping our world. This course is a one-semester tour of data science highlights for non-majors. Specifically, the course will start with brief introductions to programming in Python, basic probability and statistics, and causal inference. We will then see examples of how these tools are applied in various real-world contexts. There is no assumed background in either math or programming. The lectures for this course will focus on covering conceptual aspects of the tools of data science and demonstrations of how data science can be used. The associated laboratory sections will focus specifically on direct application of what is being learned in lectures using Python code, with guided demonstrations by the section leader(s). Restrictions: not open to students who are enrolled in, or have completed for credit, DS-UA 111 and/or 112; not open to students who have declared: the major and minor in Data Science; the major in Computer and Data Science; or the major in Data Science and Mathematics. This course does NOT count toward the requirements of these majors and minors.
Grading: CAS Graded
Repeatable for additional credit: No
Prerequisites: Not open to certain major/minors and not open to anyone who has taken DS-UA 111 and 112.

DS-UA 111 Principles of Data Science I (4 Credits)
Typically offered Fall and Spring
Restricted to students who intend to major or minor in Data Science or to major in either Computer and Data Science or Data Science and Mathematics. (All other students should enroll in Survey in Data Science.) Principles of Data Science I is the first course in a two part sequence that introduces foundational concepts in data science, with a focus on statistical principles. In this course, students will develop programming skills in Python. Building on these programming skills, students will learn the process of both drawing conclusions from data and making predictions. Students will also explore related ethical, legal, and privacy issues. This course lays the groundwork for the next course of the sequence. Formerly titled Data Science for Everyone (the content of the course has not changed).
Grading: CAS Graded
Repeatable for additional credit: No

DS-UA 112 Principles of Data Science II (4 Credits)
Typically offered Fall and Spring
Restricted to students who intend to major or minor in Data Science or to major in either Computer and Data Science or Data Science and Mathematics. (All other students should enroll in DS-UA 100 Survey in Data Science.) Principles of Data Science II builds upon the concepts introduced in Data Science I and shifts focus toward machine learning. In this course, we will cover the principles of machine learning in the domains of supervised learning, unsupervised learning and reinforcement learning. We will illuminate these principles in terms of their mathematical foundations, their implementation in code as well as practical applications. Specifically, we cover classical prediction and classification methods such as random forests or support vector machines as well as neural network approaches to these problems. Students will tie what they learned in this class together in a capstone project that incorporates these methods. Finally, we aim to touch on current developments in machine learning, such as generative AI. Formerly titled Principles of Data Science (the content of the course has not changed).
Grading: CAS Graded
Repeatable for additional credit: No
Prerequisites: DS-UA 111.

DS-UA 201 Causal Inference (4 Credits)
Typically offered Fall and 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
Prerequisites: DS-UA 112 and (MATH-UA 234 or MATH-UA 235 or MA-UY 2224) and restricted to Majors/Minors.

DS-UA 202 Responsible Data Science (4 Credits)
Typically offered Spring
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
Prerequisites: DS-UA 112.

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)
Typically offered Summer term
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 205 Machine Learning for Climate Change (4 Credits)
Typically offered Spring term of odd numbered years
This course will introduce you to both modern machine learning techniques and their domain-specific application toward climate change problems. Lectures will be organized around published research papers. Unless otherwise noted, in the second half of each class, there will be a lecture that provides the necessary background needed to read the assigned paper. In the first half of the next class, we will have a group discussion on the assigned paper. Paper discussions will be complemented by coding assignments wherein students will apply machine learning techniques to climate-related data. Skills gained in the course will be applied to a final course project designed by the students.
Grading: CAS Graded
Repeatable for additional credit: No

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)
Typically offered Fall and Spring
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
Prerequisites: DS-UA 112 and (MATH-UA 185 or MATH-UA 334 or MA-UY 2224 as co-requisites) and (CSCI-UA 473 as a co-requisite) and restricted to Majors/Minors.

DS-UA 9111 Principles of Data Science I (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
Prerequisites: Academic Level not equal to Senior, Academic Plan not equal to UCLIBAAA and not Equal to Seniors and Juniors cannot have taken DS-UA 112.

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
Prerequisites: DS-UA 112.