THE UNIVERSITY OF CHICAGO
DATA SCIENCE INSTITUTE

DATA/ STAT 11800: Introduction to Data Science I
Spring 2025
University of Chicago

Class Time: Tuesdays and Thursdays 09:30 AM- 10:50 AM
Classroom: Ryerson Phys Lab 251
Units/ Credits: 100 units

Instructor: Dr. Kriti Sehgal
Email: ksehgal@uchicago.edu
Office: Ryerson Phys Lab 260

Office Hours:
Time: 11:00 AM- 12:00 PM on Tuesdays and Thursdays,
and by appointment (Zoom or in-person)
Location: Ryerson Phys Lab 257 G

Course Description: In today's world, data is everywhere, shaping decisions in science, business, and everyday life. This course provides an accessible introduction to the fundamental concepts and tools of data science. Students will learn how to collect, store, and analyze data using Python and Jupyter Notebooks, gaining hands-on experience in data manipulation and statistical inference. Beyond technical skills, we will critically examine the role of data in society, discussing questions of bias, ethics, and responsible data use. Through lectures, coding exercises, labs, and projects, students will develop problem-solving skills that apply across disciplines.
No prior programming or statistics experience is required, but curiosity and a willingness to engage with data are essential. This course serves as the gateway to the Data Science major and minor and prepares students for further study in data analysis.

Prerequisites: There are no prerequisites for this course.

Textbook/ Materials:
1. Introduction to Data Science I & II
   https://ds1.datascience.uchicago.edu/intro.html
2. (Optional) Python Data Science Handbook by Jake VanderPlas, O'Reilly Media
   https://jakevdp.github.io/PythonDataScienceHandbook/

Software: Students will need access to Jupyter Notebooks to complete assignments and view lectures. You must have a computer with Python 3 and Jupyter Notebooks installed or access to Google Colab to open and edit .ipynb files.

Course Objectives:

1. Introduce students to the data science pipeline.
2. Develop students' abilities to be informed and critical readers of quantitative, data-based arguments.
3. Enable students to perform data analysis using Python.
4. Develop students' abilities to integrate code, background information, results and interpretation in a Jupyter Notebook to communicate data analyses effectively.
5. Help students gain flexible problem-solving and programming skills applicable to a large variety of problems independently.
6. Understand foundational concepts of probability and statistics.
7. Develop skills to apply probability and statistics for data analysis.

Student Learning Outcomes:

1. Students will be able to code and plot in Python using Jupyter Notebooks and packages such as NumPy, Pandas, and Matplotlib.
2. Students will be able to think critically about the use and collection of data.
3. Students will be able to clean, filter, group, and visualize datasets and use these methods to better understand and explain their data.
4. Students will be able to use simulation to study difficult mathematical or computational problems.
5. Students will gain an understanding of statistical methods including hypothesis testing, confidence intervals, and bootstrap.
6. Students will be able to use concepts of statistical inference to engage with research questions.
7. Students will be informed and critical readers of quantitative data-based arguments.

Course Communication: We will use Ed and Canvas for all class-related questions and discussions. Students should post questions on Ed rather than emailing, as this helps the entire class. Many students may have the same question, and posting on Ed allows instructors and TAs to provide a shared response that benefits everyone. Additionally, questions posted on Ed often receive faster replies since both instructors and TAs can see and answer your question.
All class announcements will be made during lectures, on Canvas, or on Ed. Any announcement posted on Canvas or Ed will be considered known to all students within one business day. To stay informed, students are expected to check Ed and Canvas regularly for updates.

Assessments:

- Labs (ungraded): Labs provide hands-on practice with the concepts covered in class. While they are not graded, they are a valuable tool for understanding the concepts we cover. Students are encouraged to complete labs during weekly TA office hours, where they can collaborate with peers and receive support from the TAs as needed.

- Homework: These assignments will be posted on Canvas. Students are encouraged to bring questions to TA or instructor office hours, but all submitted work must be completed independently.

- Quizzes: There will be short in-class quizzes (10-15 minutes each) weekly to check comprehension of key topics. The best way to prepare for quizzes is by actively engaging with labs and homework.

- Midterm Project: Students will select a dataset (CSV file) of personal interest, write code to analyze the data, and produce a report (maximum of four pages, including references) summarizing their findings.

- Final Exam: The course will conclude with a cumulative, in-person final exam during finals week. To ensure fairness, students must take the exam as scheduled and cannot reschedule it early or late. If a student has an unavoidable conflict, they must provide official documentation to request an alternative arrangement.

Grading Policy: Your course grade will be calculated as follows:

- Quizzes: 15%
- Homework: 30%
- Midterm Project: 25%
- Final Exam: 30%

Bonus Opportunity: The textbook *Introduction to Data Science I & II* was written by the Data Science faculty at UChicago. This is the first release and chapters for DATA 119 are still under revision. Students can earn up to 2% extra credit by providing thoughtful, constructive feedback on the chapters covered in this course. Feedback may include suggestions for clarity, identifying errors, or proposing ways to improve explanations. To submit feedback, please visit instructor's office hours or email your comments to ksehgal@uchicago.edu
Whether students obtain full extra credit or partial credit will be at the discretion of the instructor and will be based on the quality of the comments and the effort demonstrated by the student.

Late Policy: All assessments must be submitted on time, as late submissions will not be accepted, and extensions will not be granted.
However, to accommodate unforeseen circumstances, the grading policy includes built-in flexibility: the lowest homework and quiz grades will be dropped. This allowance is intended to account for personal challenges, technical issues, or other unexpected situations, so students are encouraged to use it wisely.
While individual extensions cannot be granted to ensure fairness for all, students facing difficulties are encouraged to reach out. Please note that no additional homework or quiz grades will be dropped, and no further extensions will be provided.

Use of Generative AI policy: You are not permitted to use ChatGPT, Google Bard, or any similar large language models for homework assignments. Doing so will be considered a violation of academic integrity. This restriction is in place because it is essential that you develop the ability to understand and write code independently. Mastering these skills now will enable you to use AI tools effectively and responsibly in the future.
However, for your midterm project, you may use AI tools with limitations. While you are allowed to incorporate them into your workflow, you must not simply copy and paste responses. Using AI is optional, but if you choose to use these tools, you must properly cite their contributions. This is an opportunity to learn how to engage with AI tools critically and ethically. Failure to cite AI-generated content will be treated as a violation of the University of Chicago's Academic Honesty and Plagiarism policy.

Academic Integrity: Acting with academic integrity means, in brief, not submitting the statements, work, or ideas of others as one's own. Students are expected to comply with University regulations regarding honest work. If you are in doubt about what constitutes academic dishonesty, speak with me before the assignment is due. Failure to maintain academic integrity on an assignment will result in a penalty befitting the violation, up to and including failing the course and further University sanctions. For more information, consult the student manual: https://studentmanual.uchicago.edu/academic-policies/academic-honesty-plagiarism/

Accessibility: The University of Chicago is committed to ensuring equitable access to our academic programs and services. Students with disabilities who have been approved for the use of academic accommodations by Student Disability Services (SDS) and need a reasonable accommodation(s) to participate fully in this course should follow the procedures established by SDS for using accommodations. Timely notifications are required to ensure that your accommodations can be implemented. Please meet with me to discuss your access needs in this class after you have completed the SDS procedures for requesting accommodations. For more information, please visit: https://disabilities.uchicago.edu/

Mental Health: Student Wellness' Mental Health professional staff members work with students to resolve personal and interpersonal difficulties, many of which can affect the academic experience. These include conflicts with or worry about friends or family, concerns about eating or drinking patterns, and feelings of anxiety and depression. See:
https://wellness.uchicago.edu/mental-health/

Diversity and Inclusion: I value diversity and inclusion. I am committed to a climate of mutual respect and full participation. My goal is to create learning environments that are usable, equitable, inclusive, and welcoming. If there are aspects of the instruction or design of this course that result in barriers to your inclusion, achievement, or the accurate assessment of your learning, please notify me as soon as possible. Students with disabilities are also welcome to contact Student Disability Services to discuss a range of options to removing barriers in the course, including official accommodations.

Bias Reporting: The University has a process through which students, faculty, staff and community members who have experienced or witnessed incidents of bias, prejudice or discrimination against a student can report their experiences to the University's Bias Education and Support (BEST) team.
See: https://diversityandinclusion.uchicago.edu/resources/reporting-incidents/

Accommodations based upon sexual assault: The University is committed to offering reasonable academic accommodations to students who are victims of relationship or sexual violence, regardless of whether they seek criminal or disciplinary action. If a student comes to us to discuss or disclose an instance of sexual assault, sex discrimination, sexual harassment, dating violence, domestic violence or stalking, or if we otherwise observe or become aware of such an allegation, we will keep the information as private as we can, but as faculty members of University of Chicago, we are required to immediately report it to a Department Chair or Dean or directly to the University's Title IX Coordinator. If you would like to speak with the Title IX Coordinator directly, Bridget Collier can be reached at bcollier@uchicago.edu. Additionally, you can report incidents or complaints by contacting UCPD at (773)702-8181 or your local law enforcement agency. For more information:
https://studentmanual.uchicago.edu/university-policies/the-university-of-chicago-policy-on-title-ix-sexual-harassment/