

# ACE 592 SAE: Data Science for Applied Economics

Spring 2024

Class Meets: TR, 4:00 - 5:20  
Instructor: Professor Jared Hutchins  
Contact: jhtchns2@illinois.edu

320 Mumford  
Office: 312 Mumford  
Office Hours: TR, 2:30-3:30pm  
(or by appointment)

TA: Paavani Sachdeva  
Contact: paavani2@illinois.edu  
Credit Hours: 4

Office: Zoom  
Office Hours: M,W, 10am-11am

## Course Description

The ability to obtain, process, and analyze data using coding and algorithms has become an absolutely essential skill for anybody engaged in economic analysis. In the digital age, there is more data available than ever before on human behavior: from analyzing an elected official's opinion from Twitter to identifying a farmer's crop choices through satellite images. For those engaged in academic research in economics, these new data sources have drastically expanded the number of research questions that can be answered and revolutionized how we answer classic questions in the field. For those working outside of academia, the ability to process and analyze large datasets to provide insight has become an essential skill for producing useful analysis on a problem or question.

The goal of this course is to teach students in applied economics how to use data science tools and workflow for answering questions in economics. Using the python programming language and git version control, we will cover obtaining data via scraping and APIs, processing and cleaning data using python, and analyzing data via data visualization and basic machine learning techniques. The course will broadly cover the basics of text, spatial, and numeric data with an emphasis on their uses in analyzing economic questions and conducting research.

Data science can be broadly defined as “an approach to data analysis with a foundation in code and algorithms.” The main goals of data science include drawing useful conclusions from large and diverse datasets through exploration, prediction, and inference. Data science is fundamentally different than most previous approaches to data analysis because:

- It uses **diverse data sources**, such as text, image, spatial, or numeric.

- There is an emphasis on **work flow** that uses coding and version control to maximize reproducibility and transparency.
- It uses **new tools and approaches** for answering questions about data.

## Course Objectives

In light of this, by the end of the course we will be able to:

1. Obtain and process text, image, and numeric data using Python.
2. Analyze data using basic data visualization and machine learning in Python.
3. Construct a git repository and collaborate on a research project on Github.
4. Document code and communicate results using Jupyter notebooks.

To achieve these objectives, we will learn three main tools:

- **Python**, an object-oriented and general purpose programming language that can be used for reading, processing, and analyzing data.
- **Git**, a version control software for documenting and collaboration, as well as its online platform Github.
- **Jupyter notebooks**, a development environment for interactive programming that supports markdown.

## Course Delivery

We will meet **in-person** in Mumford 320 on Tuesdays and Thursdays 4:00 - 5:20 pm. For those interested in recorded lectures to review, there is a catalog of the previous two years of lectures than can be made available through MediaSpace. Office hours for the instructor can be in person in Mumford 312 or over Zoom.

The TA's office hours will be over Zoom on Wednesdays and Fridays from 10-11 am. Special discussion sections to go over specific topics or review for homeworks will take place on the Friday office hours slot over Zoom.

Zoom links for office hours can be found in the Syllabus Appendix.

## Prerequisites

There is no formal prerequisite for this course, but to be the most successful in this course the student *should be at least a masters student, should have a baseline knowledge of econometrics and statistics, and should have some experience coding (e.g. STATA, SAS, Matlab)*, though experience coding Python is not assumed for this course.

Since much of this course is learning how to write code, students should note that **to learn to code, the best way is to do it. A lot.** Part of learning how to code will also be learning what resources exist to help you overcome problems, so while the lectures and homeworks exist as guides **how much you learn is a function of how much effort you put into practicing and supplementing your existing knowledge.**

## Grading and Assignments

Grading will be based on two components:

- **Five assignments** due throughout the semester (60 points total).
  - Submissions must be in the form of **Jupyter Notebook**, submitted to GitHub Classroom (links in Syllabus Appendix).
  - Group work is encouraged, but write ups **must be individual**.
  - There is one additional homework at the end of the semester, Homework 4, which is 10 extra credit points.
- One **final analysis project and presentation**, done in groups (15 points GitHub Repo + 15 points write-up + 10 points presentation = 40 points).
  - The project must answer an applied economics question using data.
  - Groups must be formed by **Feb 16**, and the topic has to be approved by me by **March 23**.
  - Submission will be 1) a presentation done at the end of the semester, 2) a Github repository for your group's project, and 3) a write up of the findings of your project.

## Grading Rubric

Assignment	Points	Due Date
Homework 0	5	February 8
Homework 1	15	February 29
Homework 2	20	March 28
Homework 3	20	April 18
Final Presentation	10	May 7 - May 9
Final Project	30	May 9
Total Points	100	
Homework 4 (Extra Credit)	10	May 9

# Resources

There is no textbook for this course, as most of the learning in this course is self-guided to some extent. As you come across problems in solving assignments and the like, these resources can help you overcome issues.

## Big Picture Resources

- [Learn Python](#) (good for basics)
- [Python for Data Science](#) (most useful one in my opinion)
- [Coding for Data course](#)
- [Inferential Thinking](#), a Berkeley course on data science:
- [Fundamentals of Data Visualization](#)
- [Python Graph Gallery](#)

## Specific Resources

- Stack Overflow for very specific, coding problems.
- [DataCamp Cheat Sheets](#), for git, Jupyter, and various python operations.

# Rough Course Schedule

Subject to change given needs of the semester.

Week	Module	Main Topics
1	Introduction	- Python, git, and Jupyter basics
2		- Pandas, numpy, matplotlib
3		- requests, API basics
4		- <b>Homework 0</b>
5	Text as Data	- Python processing text
6		- HTML parsing and scraping, basic NLP
7		- <b>Homework 1</b>
8	Images as Data	- Image editing, spatial data types
9		- Spatial statistics, basic mapping
10		- <b>Homework 2</b>
11	Numbers as Data	- Advanced pandas and numpy operations
12		- Distributed programming, scaling and parallelizing;
13		- <b>Homework 3</b>
14	Data Analysis	- Advanced visualization
15		- Unsupervised learning, feature generation
16		- Supervised learning, cross-validation
17	Presentations	

## Absence Policy

There is no penalty for absences in the class, but it is expected that students regularly show up to participate and be active participants in class. Students who are occasionally not able to attend for reasons of sickness or other special situations are encouraged to talk to the

instructor and make use of lecture materials that are being posted to keep pace with the class.

## **Academic Integrity**

Academic dishonesty may result in a failing grade. Every student is expected to review and abide by the Academic Integrity Policy: <https://studentcode.illinois.edu/article1/part4/1-401/> .

Ignorance is not an excuse for any academic dishonesty. It is your responsibility to read this policy to avoid any misunderstanding. Do not hesitate to ask the instructor(s) if you are ever in doubt about what constitutes plagiarism, cheating, or any other breach of academic integrity.

## **Family Educational Rights and Privacy Act (FERPA)**

Any student who has suppressed their directory information pursuant to Family Educational Rights and Privacy Act (FERPA) should self-identify to the instructor to ensure protection of the privacy of their attendance in this course. See <http://registrar.illinois.edu/ferpa> for more information on FERPA.

## **Mental Health**

Significant stress, mood changes, excessive worry, substance/alcohol misuse or interferences in eating or sleep can have an impact on academic performance, social development, and emotional wellbeing. The University of Illinois offers a variety of confidential services including individual and group counseling, crisis intervention, psychiatric services, and specialized screenings which are covered through the Student Health Fee. If you or someone you know experiences any of the above mental health concerns, it is strongly encouraged to contact or visit any of the University's resources provided below. Getting help is a smart and courageous thing to do for yourself and for those who care about you.

- Counseling Center (217) 333-3704
- McKinley Health Center (217) 333-2700
- National Suicide Prevention Lifeline (800) 273-8255
- Rosecrance Crisis Line (217) 359-4141 (available 24/7, 365 days a year)

If you are in immediate danger, call 911.

\*This statement is approved by the University of Illinois Counseling Center.

## Community of Care

As members of the Illinois community, we each have a responsibility to express care and concern for one another. If you come across a classmate whose behavior concerns you, whether in regards to their well-being or yours, we encourage you to refer this behavior to the Student Assistance Center (217-333-0050 or <http://odos.illinois.edu/community-of-care/referral/>). Based on your report, the staff in the Student Assistance Center reaches out to students to make sure they have the support they need to be healthy and safe.

Further, as a Community of Care, we want to support you in your overall wellness. We know that students sometimes face challenges that can impact academic performance (examples include mental health concerns, food insecurity, homelessness, personal emergencies). Should you find that you are managing such a challenge and that it is interfering with your coursework, you are encouraged to contact the Student Assistance Center (SAC) in the Office of the Dean of Students for support and referrals to campus and/or community resources.

## Students with Disabilities

To obtain disability-related academic adjustments and/or auxiliary aids, students with disabilities must contact the course instructor as soon as possible and provide the instructor with a Letter of Academic Accommodations from Disability Resources and Educational Services (DRES). To ensure that disability-related concerns are properly addressed from the beginning, students with disabilities who require assistance to participate in this class should apply for services with DRES and see the instructor as soon as possible. If you need accommodations for any sort of disability, please speak to me after class, or make an appointment to see me or see me during my office hours. DRES provides students with academic accommodations, access, and support services. To contact DRES, you may visit 1207 S. Oak St., Champaign, call 217-333-1970, e-mail [disability@illinois.edu](mailto:disability@illinois.edu) or visit the DRES website at <http://www.disability.illinois.edu/>. Here is the direct link to apply for services at DRES, <https://www.disability.illinois.edu/applying-services>.

## Disruptive Behavior

Behavior that persistently or grossly interferes with classroom activities is considered disruptive behavior and may be subject to disciplinary action. Such behavior inhibits other students' ability to learn and an instructor's ability to teach. A student responsible for disruptive behavior may be required to leave class pending discussion and resolution of the problem and may be reported to the Office for Student Conflict Resolution (<https://conflictresolution.illinois.edu>; [conflictresolution@illinois.edu](mailto:conflictresolution@illinois.edu); 333-3680) for disciplinary action.

## Emergency Response Recommendations

Emergency response recommendations and campus building floor plans can be found at the following website: <https://police.illinois.edu/em/run-hide-fight/>. I encourage you to review

this website within the first 10 days of class.

## **Religious Observances**

Illinois law requires the University to reasonably accommodate its students' religious beliefs, observances, and practices in regard to admissions, class attendance, and the scheduling of examinations and work requirements. Students should complete the Request for Accommodation for Religious Observances form should any instructors require an absence letter in order to manage the absence. In order to best facilitate planning and communication between students and faculty, students should make requests for absence letters as early as possible in the semester in which the request applies.

## **Sexual Misconduct Reporting Obligation**

The University of Illinois is committed to combating sexual misconduct. Faculty and staff members are required to report any instances of sexual misconduct to the University's Title IX Office. In turn, an individual with the Title IX will provide information about rights and options, including accommodations, support services, the campus disciplinary process, and law enforcement options.

A list of the designated University employees who, as counselors, confidential advisors, and medical professionals, do not have this reporting responsibility and can maintain confidentiality, can be found here: [wecare.illinois.edu/resources/students/#confidential](https://wecare.illinois.edu/resources/students/#confidential).

Other information about resources and reporting is available here: [wecare.illinois.edu](https://wecare.illinois.edu).