McCormick

Northwestern Engineering

Civil & Environmental Engineering

Data Science for Urban Systems

Civil and Environmental Engineering Department
McCormick School of Engineering

Instructor: Ying Chen
Office: TECH A224 or TC Room 214
In-person Office Hours: Thursdays 1:00 ~ 2:00 pm
On-line Office Hours: Flexible but by appointment
Email: y-chen@northwestern.edu
Website: http://sites.northwestern.edu/ych168/

Textbook:

- Think Python – How to Think Like a Computer Scientist
- Data Visualization in Python
- Machine Learning with Python Cookbook: Practical Solutions from Preprocessing to Deep Learning

Class Times: MW 10:00 ~10:50am, F 11:00 ~ 11:50am
Location: Tech MG28 and Tech A110

Software: Python

Course Description

Data Science in Urban Systems introduces most state-of-the-art data science concepts, techniques, and teaches students to select and apply the right algorithms to solve problems.

As an important component of big data processes, data management plays a critical role. Therefore, this course aims to close the gaps within students' data science knowledge toolbox. This course will help students to gain the fundamental knowledge and skills they need in data engineering such as its ecosystem, lifecycle, and tools to manage data for implementing data analytics/science and building machine learning models to solve urban systems challenges like optimizing transportation, urban design and energy portfolios.

Who should take this class?

This course is designed for students who want to learn to program in python for data science. This course guides students in work through basic Python programming

language, from basic concepts, and data processing to final data analysis using the correct python and external packages.

In order to present this idea clearly and increase relevance for engineering students, we will use applications from the urban systems areas (including transportation, urban planning, and civil and environmental engineering, but not limited to those fields). At the end of this course, students will gain a competitive edge by tapping into the power of data science.

Prior Expectations:

- Basic familiarity with computer programming in other languages at the introductory level (i.e., R, Matlab, SAS, etc.) and no prior knowledge of *Python* is required;
- Basic college-level math knowledge (probability/statistics/matrices) (such as, CIV_ENV 306).
- Since this course includes some hands-on labs, students are expected to install Python and relevant packages (recommend from Anaconda) and bring their laptop to every class.

Learning Objectives

- To provide students a solid starting point for using Data Science in their work and research;
- Students will be able to understand and use standard sequential, conditional, and iterative control structure of automated data analysis through computers;
- To familiarize students with leading tools used in modern data science practice;
- To help students understand how to manipulate data (store, query, and summarize) using a database designed to analyze structured data;
- To help students understand and use computer programming to collect, analyze and visualize data related to various urban systems challenges;

ABET Program Outcomes

- An ability to identify, formulate, and solve complex engineering problems by applying principles of engineering, science, and mathematics.
- An ability to communicate effectively with a range of audiences.
- An ability to function effectively on a team whose members together provide leadership, create a collaborative and inclusive environment, establish goals, plan tasks, and meet objectives.
- An ability to develop and conduct appropriate experimentation, analyze and interpret data, and use engineering judgment to draw conclusions.
- Ability to acquire and apply new knowledge as needed, using appropriate learning strategies.

Tools

- Jupyter Notebook (https://jupyter.org/)
- Scientific computing package Numpy (https://numpy.org)
- Using Python and SQL to create and query relational databases, in particular MySQL
- Pandas(http://pandas.pydata.org/) for handling and visualizing structured data and time series
- Machine learning packages: scikit-learn(http://scikit-learn.org/stable/documentation.html)

Tentative Schedule

It is a tentative schedule of lectures for this course. We will try to keep approximately on this schedule. (Note that we may change the agenda during the quarter.)

Schedule	Topics
Week 1	Introduction to basic programming concepts Install and run Python programs
Week 2	Sequences: Strings, Lists, and Files
Week 3	Data structures (List, Dictionary, Tuple and Set) in python
Week 4	Modules, Functions, parameters, and scientific computing: NumPy
Week 5	Database operations and Preparation: SQLite and MySQL
Week 6&7	Data Processing: Pandas; Project Version Control: Git and GitHub
Week 8	Data Visualization: Matplotlib, Plotly, Seaborn, Bokeh, and GeoPandas etc.
Week 9	Machine Learning: scikit-learn, NLTK
Week 10	Machine Learning in Urban areas
Week 11	Project Presentation

Websites for Instruction

Canvas

We will use Canvas to distribute readings, assignments, and grades.

DataCamp

I applied for a datacamp classroom for our course. You will automatically have full access to the entire course curriculum on DataCamp. This is a good source for self-learning.

Kaggle (or Github)

Sometimes, I will run the code using Kaggle notebook in class and you may consider using it for your presentation or your collaboration work for your final project.

OneDrive

I will share the large dataset with some of you depending on the project you plan to work on.

Assignments

We have five homework assignments. These assignments are mainly from the lectures. They will cover basic data querying, data preprocessing, data visualization etc. These assignments will help you understand concepts and ideas you've learned from lectures. You need to submit a report and your code at the same time.

Academic Integrity Statement: Students in this course are required to comply with the policies found in the booklet, "Academic Integrity at Northwestern University: A Basic Guide". All papers submitted for credit in this course must be submitted electronically unless otherwise instructed by the professor. Your written work may be tested for plagiarized content. For details regarding academic integrity at Northwestern or to download the guide, visit: https://www.northwestern.edu/provost/policies-procedures/academic-integrity/index.html. Any form of cheating, including improper use of content generated by artificial intelligence, constitutes a violation of Northwestern's academic integrity policy.

Late Assignment Policy: the penalty is **10%** off the grade of your project or each assignment for every additional day after the deadline.

Project

We will have a group-based class project using real data. The size of each self-assigned group is three at maximum. Each group will be assigned a case with real data and problems in the real world. Each group also can use existing online datasets or download their own datasets from online resources, like Facebook, Twitter, Yelp, etc. We expect each group could generate a technical report to show some interesting findings by running existing big data analysis algorithms. We encourage each group/student to use the dataset relevant in their fields. You need to submit a detailed technical report along with the source code.

Grading

Your final grade will be composed of the following items:

Attendance:	1% * 10 = 10%
Assignments:	10% *5 = 50%
Midterm Exam:	15% * 1 =15%
Final Project:	25% * 1 = 25%

Letter grades are assigned as follows:

Points Letter Grade	Percentage
A	100 – 90
A-	89 – 85
B+	84 – 80

B	79 – 75
B-	74 – 70
C+	69 – 65
C	64 – 60
F	Below 60

Office Hours, E-mail

Your visits, whether online or in-person, are not confined to my regular office hours. However, I prefer scheduling appointments outside of these hours via email. Even during regular office hours, please email to confirm your visit, as this helps me manage any potential conflicts. Email is my preferred communication method, as I typically respond to messages within one day of receiving them.

Accessibility Statement

Northwestern University is committed to providing the most accessible learning environment as possible for students with disabilities. Should you anticipate or experience disability-related barriers in the academic setting, please contact AccessibleNU to move forward with the university's established accommodation process (e: accessiblenu@northwestern.edu; p: 847-467-5530). If you already have established accommodations with AccessibleNU, please let me know as soon as possible, preferably within the first two weeks of the term, so we can work together to implement your disability accommodations. Disability information, including academic accommodations, is confidential under the Family Educational Rights and Privacy Act.

Religious Observance Statement

Northwestern is committed to fostering an academic community respectful and welcoming of persons from all backgrounds. To that end, the policy on academic accommodations for religious holidays stipulates that students will not be penalized for class absences to observe religious holidays. If you will observe a religious holiday during a class meeting, scheduled exam, or assignment deadline, please let me know as soon as possible, preferably within the first two weeks of class. If exams or assignment deadlines on the syllabus fall on religious holidays you observe, please reach out so that we can discuss that coursework.

Course Details Subject to Change

Please note that the specifics of this course syllabus are subject to change in the case of unforeseen circumstances. I will notify students of any changes as soon as possible. Students will be responsible for abiding by the changes.

Exceptions to Class Modality

Class sessions for this course will occur in person. Individual students will not be granted permission to attend remotely except as the result of an Americans with Disabilities Act (ADA) accommodation as determined by AccessibleNU.

Community health remains our priority. If you are experiencing symptoms of COVID-19, do not attend class and follow the steps outlined by the CDC for testing and isolation. Contact me as soon as possible to make plans to complete your coursework. Students who experience other personal emergencies should contact me as soon as possible to arrange to complete coursework. Should public health recommendations prevent in-person class from being held on a given day, I or the university will notify students.

Prohibition of Recording of Class Sessions by Students

Unauthorized student recording of classroom or other academic activities (including advising sessions or office hours) is prohibited. Unauthorized recording is unethical and may also be a violation of University policy and state law. Students requesting the use of assistive technology as an accommodation should contact AccessibleNU. Unauthorized use of classroom recordings – including distributing or posting them – is also prohibited. Under the University's Copyright Policy, faculty own the copyright to instructional materials – including those resources created specifically for the purposes of instruction, such as syllabi, lectures and lecture notes, and presentations. Students cannot copy, reproduce, display, or distribute these materials. Students who engage in unauthorized recording, unauthorized use of a recording, or unauthorized distribution of instructional materials will be referred to the appropriate University office for follow-up.

Support for Wellness and Mental Health

Northwestern University is committed to supporting the wellness of our students. Student Affairs has multiple resources to support student wellness and mental health. If you are feeling distressed or overwhelmed, please reach out for help. Students can access confidential resources through the Counseling and Psychological Services (CAPS), Religious and Spiritual Life (RSL) and the Center for Awareness, Response and Education (CARE). All Northwestern students are also eligible to access support at no cost though TimelyCare, a virtual mental health platform that provides counseling, health coaching and 24/7 on-demand services.

Additional information on the resources mentioned above can be found here:
https://www.northwestern.edu/counseling/
https://www.northwestern.edu/religious-life/
https://www.northwestern.edu/care/
https://www.northwestern.edu/studentaffairs/timelycare.html