THE UNIVERSITY OF CHICAGO
HARRIS SCHOOL OF PUBLIC POLICY

PPHA 30538: Data and Programming for Public Policy II – Python Programming

Dr. Jeffrey A. Levy
levyjeff@uchicago.edu
Keller 3101

Head TA: TBD
TAs: TBD

Autumn Quarter 2023
Tuesday September 26th – Thursday November 30th

Section 1: M, W 1:30 PM – 2:50 PM* Keller 0021
Section 2: M, W 3:00 PM – 4:20 PM* Keller 0021
Section 3: T, Th 9:30 PM – 10:50 PM Keller 0021
Section 4: T. Th 11:00 PM – 12:20 PM Keller 0021

Office Hours for the professor can be reserved at this Calendly link, while remote lab sessions with the TAs are listed on the Canvas Zoom page.

Course Description

What exactly separates "beginner" code from more advanced code?

What are the benefits of moving away from "beginner" code if it got the job done in PPHA 30537, or in a summer internship?

How do I take these new programming skills and actually perform original research with them, as opposed to simply completing specified tasks?

These are some of the questions we will seek to answer, as students who have completed PPHA 30537 continue to build on their data analysis skills with the Python programming language. The first part of this course will expand and deepen our skills with Python. The second part will focus on a final research project, where students will pursue a research question of their choosing (alone or in groups of up to 3) in order to showcase their research programming skills.

Prerequisites

This course requires a passing grade in PPHA 30537: Data and Programming for Public Policy I – Python Programming. As the courses are designed to build on each other, exceptions to this rule are not possible.

** Due to the quarter staring on Tuesday, the first Monday class will meet at the scheduled time on Friday of week 1.*

Modes of Engagement

Instruction for this class will have four primary elements, mirroring those used in PPHA 30537:

- New content will be introduced in **asynchronous lectures** posted to Canvas around noon on the day before scheduled class time. I aim to keep these below 30 minutes in length.
- The scheduled lecture times will be used as **live labs**, in which we delve deeper into the content introduced in the lectures, and work through examples in groups.
- Weekly optional **office hours** for the professor and TAs, where individuals can get one-on-one help with questions.
- An optional **discussion board** for questions and discussions outside of office hours and class.

Learning Objectives

This course will build directly on the material covered in PPHA 30537. We will assume a grasp of the Python skills from the previous class at the start, so that we can focus on practical applications to research. Whereas the goals of the first class were to introduce Python as a tool for data analysis and to prepare students for internship-level policy research positions, the goals of this course will be to:

- Go from applying Python to solve research questions, to applying Python professionally in a way that supports code maintenance, collaboration, efficiency, and readability.
- Deepen existing skills.
- Broaden into new skills that require a higher level of Python proficiency, such as creating a web dashboard.
- Prepare for the post-graduation job market in data and programming related areas.

Assessment and Grading

Your progress in the learning objectives will be assessed in two ways:

In-class quizzes (10%) – Each class will have a brief (5 minute, 2-3 question) quiz on Canvas that will cover a core skill or concept from lecture. These will serve as attendance and provide important instant feedback on the material.

Take-home assignments (45%) – Weeks 1-7 will have coding assignments that ask students to use class concepts to solve research programming questions. Assignments will test your ability to work on a question with a starting place and a broad goal, mimicking real-world research tasks wherever possible.

Final project (45%) – The capstone of both Data and Programming I and II, you will have a chance to showcase all of your Python skills in a research project of your choice, working alone or in groups of up to 3.

This class requires a 60% or above to pass and is not curved. All passing grades will use the following intervals:

A	[95% - 100%]	B+	[85% - 90%)	B-	[60% - 80%)
A-	[90% - 95%)	B	[80% - 85%)		

Class Policies

No **attendance** is taken, but quizzes can only be completed in the classroom. Absences must be cleared with the head TA.

Assignments and the project **must be turned in** using GitHub and Gradescope, similar to last quarter. General feedback according to an assignment-specific rubric will be provided through Gradescope approximately one week after the due date.

Regrade requests must be submitted on Gradescope with a (polite) explanation, which will then be re-evaluated by the original grader. Continued disagreement may be escalated to the head TA first, and finally to the professor. All regrade requests may result in a full regrade and potentially a higher or lower score. See the Gradescope Guidelines document on Canvas for additional important details.

Every student has **two 12-hour late tokens** available to them during the quarter. Those extensions will be automatically applied to any late take-home assignments and require no excuse to be given. These extensions are used in complete blocks of time – e.g. turning in an assignment 12 hours and 30 minutes late will use two tokens. Once your late tokens are used up for the quarter, all assignments will be penalized at a rate of 5% per 12-hour block. These tokens are intended to cover ordinary illness, family events, and so on – only issues of sufficient magnitude that academic affairs is involved in the discussion can qualify for exceptions. Once solutions have been posted to the class (generally Wednesday), no further assignments may be turned in. Late tokens may not be used on the final project.

See the **academic integrity policy** and the general **grading rubric** on the Canvas course page.

Support

Your mental and physical health is important. As graduate students, I recognize that you are all under immense pressure to achieve academic excellence alongside maintaining personal and often professional lives. Please take care of yourselves and each other, and speak to me if, for any reason, you are having difficulty keeping up with the course. Many other sources of support are available, including:

Find the Harris Student Affairs office here.

Learn more about accommodations for students with disabilities here.

See the Harris academic support programs, including tutoring and code labs, here.

Software and Resources

There are no required textbooks, as Python is extremely well supported online. I expect students will primarily be using the official Python documentation and StackOverflow, which will be discussed in class. The text Python for Data Analysis 3rd Edition by Wes McKinney may be helpful, but will not be referenced directly in class.

There are two pieces of software that are required for this class, both of which are free:

- The Anaconda Python distribution (or alternative)
- The GitHub Desktop application

Course Outline

Homework 1: Given Thursday Sep 28, due Monday Oct 9 by midnight
Homework 2: Given Monday Oct 9, due Monday Oct 23 by midnight
Homework 3: Given Monday Oct 23, due Monday Nov 6 by midnight
Homework 4: Given Monday Nov 6, due Monday Nov 13 by midnight
Final Project: Due Wednesday Dec 6 by midnight

Week 1 (Sep 26-Sep 29): Homework 1 material

- Introduction, final project discussion
- Code generalization and organization, functions, style

Week 2 (Oct 2-Oct 5): Homework 1 material

- Difficult and irregularly-shaped data

Week 3 (Oct 9-Oct 13): Homework 2 material

- Data visualization

Week 4 (Oct 16-Oct 20): Homework 2 material

- Spatial data

Week 5 (Oct 23-Oct 27): Homework 3 material

- PDF parsing and text and language processing

Week 6 (Oct 30-Nov 3): Homework 3 material

- Text and language processing

Week 7 (Nov 6-Nov 10): Homework 4 material

- Dashboards with Shiny

Week 8 (Nov 13-Nov 17)

- Group project presentations

Thanksgiving Break (Nov 20-Nov 24)

- No class

Week 9 (Nov 27-Nov 30)

- Virtual environments
- Code samples and job prep