

Instructions for Installation and Environment Setup

Overview

Over the next few days, we'll use Python to query, transform, and analyze datasets. Before we can use Python, we will do a little work to set up the software needed on your machines. The instructions below will walk you through the steps you need to take.

- Download the Code
- Install Anaconda
- Open Anaconda Navigator and run Jupyter
- Open a terminal

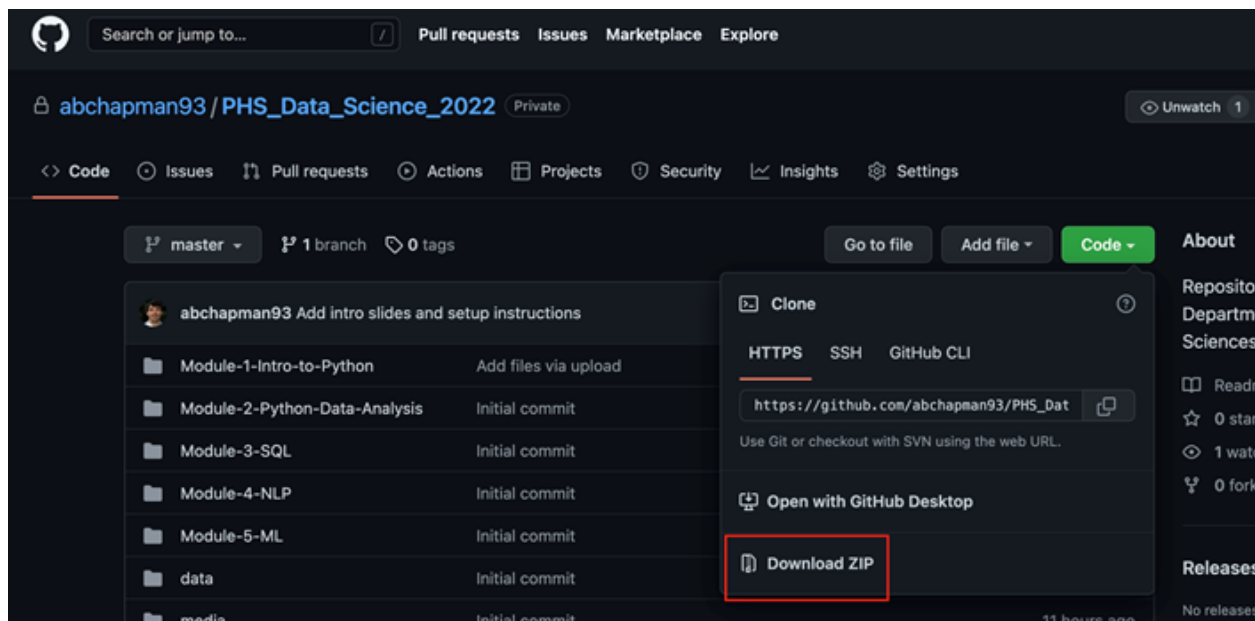
1. Download the Code

The code for this class is stored in a remote GitHub repository. You can either download the code directly as a zip file or clone it using GitHub Desktop.

Link: https://github.com/abchapman93/PHS_Data_Science_2022

a. Download as a .zip file

Go to the GitHub repository in the link above. Click the green "Code" button and then select "Download ZIP". Unzip the file somewhere on your machine that you'll be able to find later (we'll come back to it after installing Anaconda).

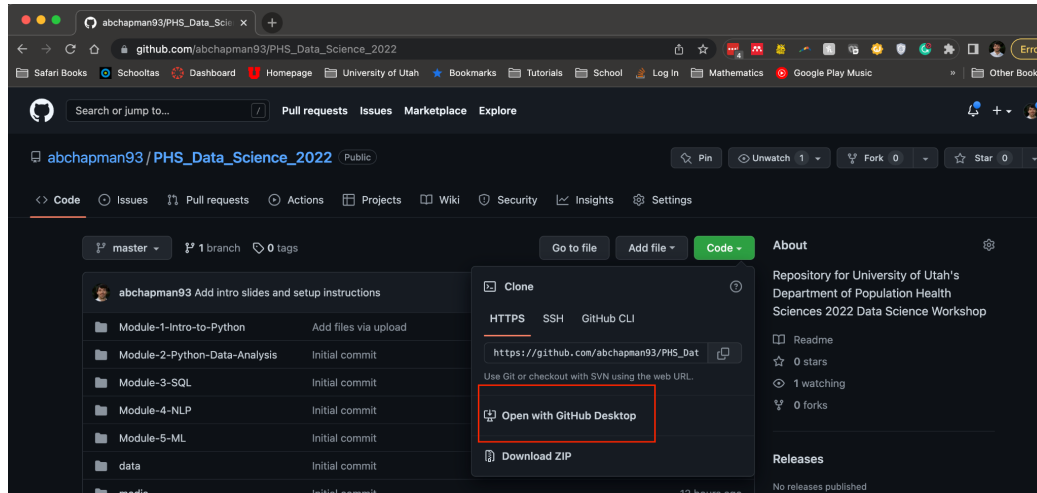


b. Clone using GitHub Desktop

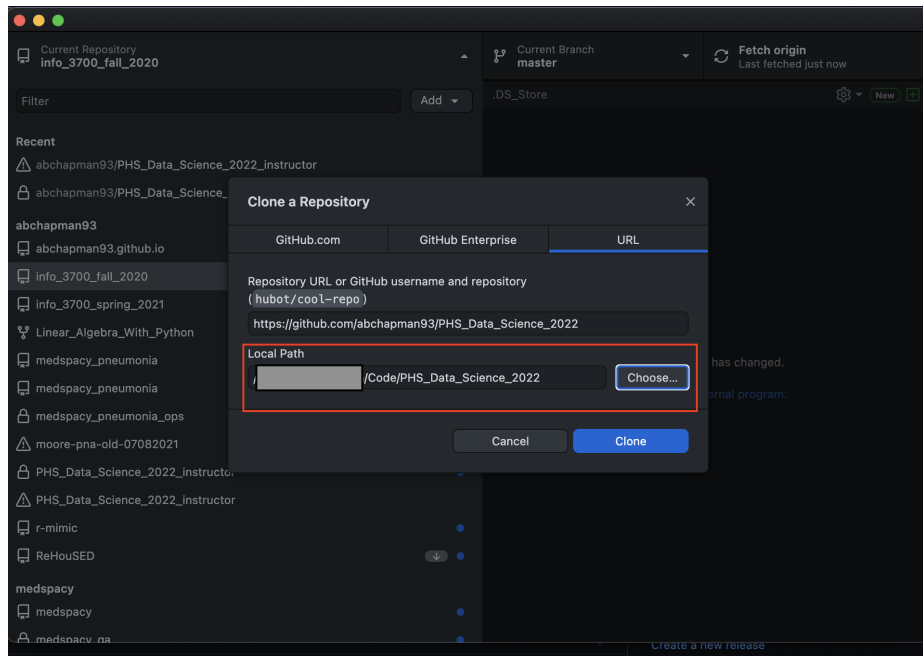
We won't spend too much time going over git in the class, but if you're familiar with git you can clone the repository directly.

First, download and install [GitHub Desktop](#).

Click the green "Code" button and select "Open with GitHub Desktop".



This should launch GitHub Desktop on your machine and ask you where to put the code. Choose somewhere on your machine that you'll be able to find later (we'll come back to it after installing Anaconda). For example, in the screenshot below I'll be saving the code to a folder called "Code/PHS_Data_Science_2022"



2. Install Anaconda

Anaconda is an open-source distribution of Python. When you install Anaconda, it will include the Python software and a number of additional tools to use it.

1. Go to the Anaconda website: <https://www.anaconda.com/distribution/> Select your platform (Windows, Mac, or Linux)
2. Click the "Download" button to download the installer (see screenshot below). Make sure it says Python 3.9 like in this screenshot.

Individual Edition is now

ANACONDA DISTRIBUTION

The world's most popular open-source Python distribution platform



3. Run Jupyter Notebook

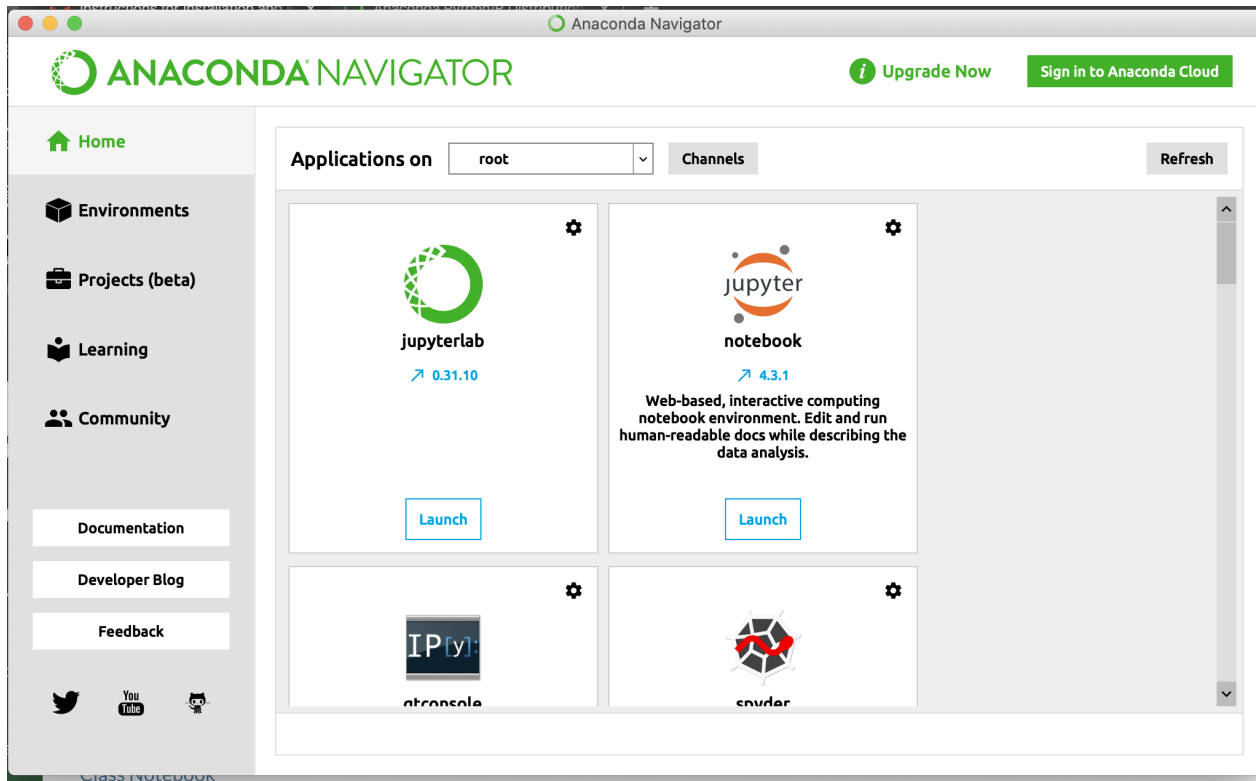
Once you've installed Anaconda, there will be several new applications on your computer. There are two main ways to interact with Anaconda:

- A visual GUI
- A command-line interface

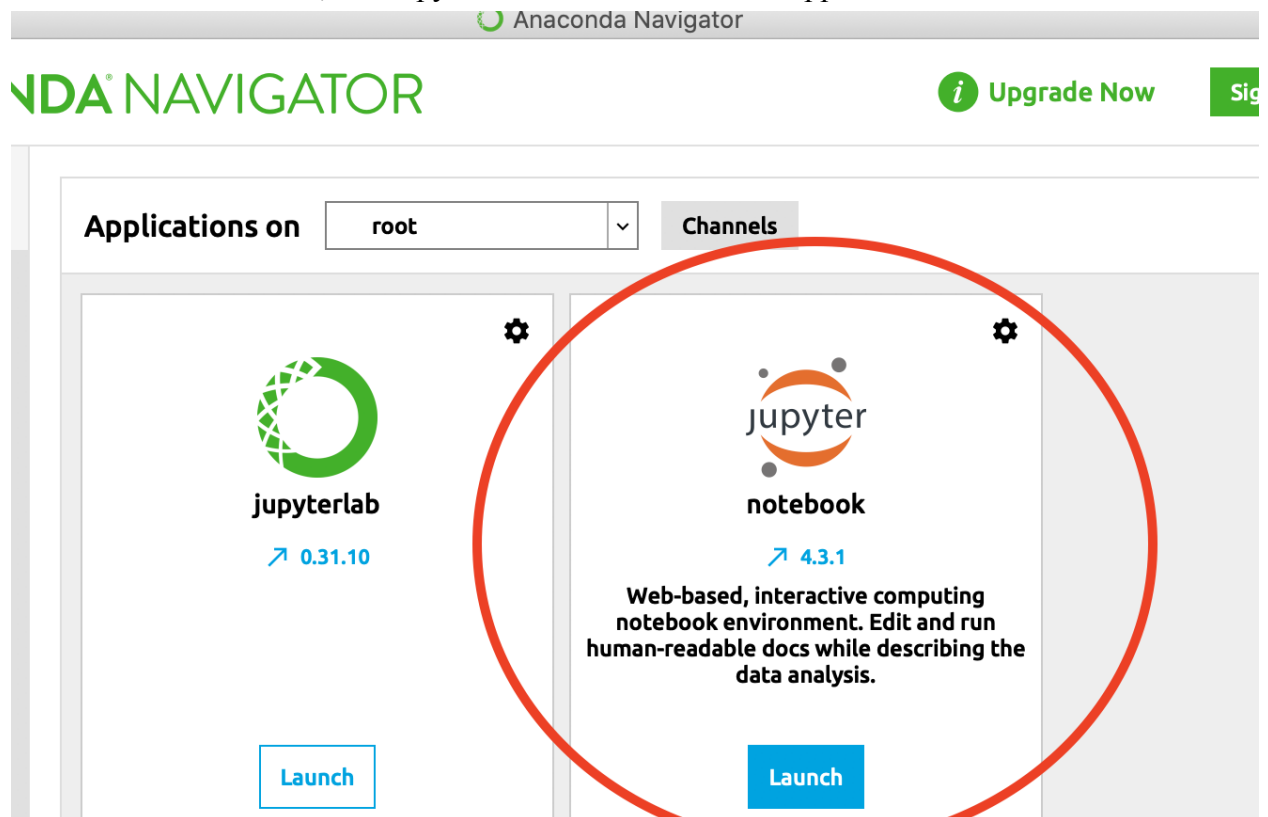
We'll start with the GUI since that's easier if you're not used to command-line prompts.

a. Anaconda Prompt

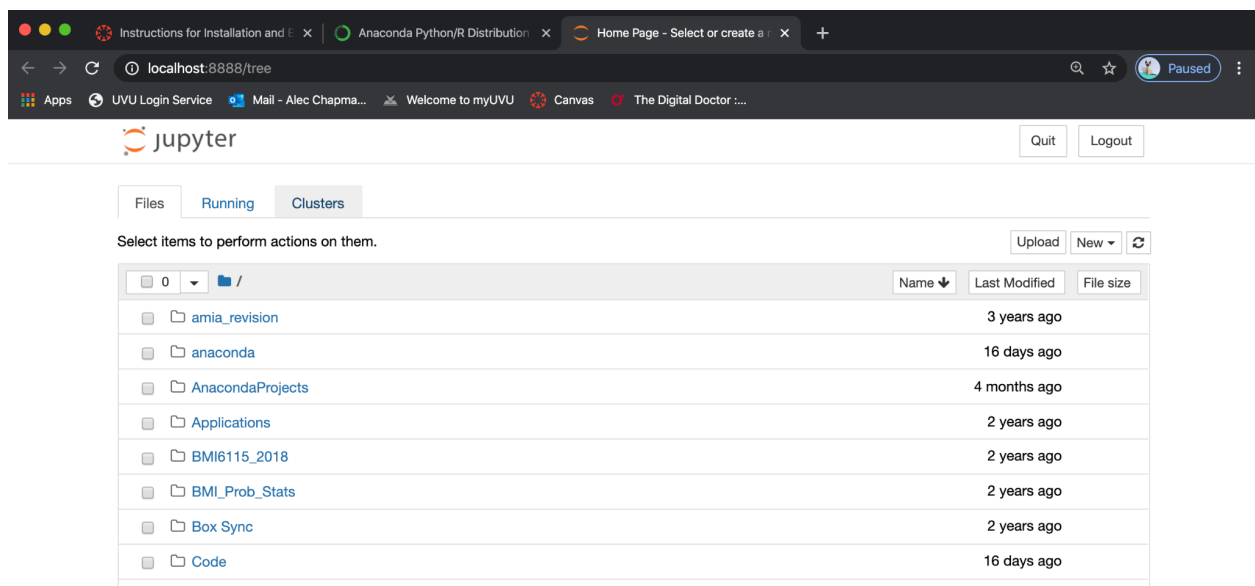
1. Open the app Anaconda Navigator. That should launch a window like this one



- This window shows the applications which come installed with Anaconda. The first app we'll look at is notebook, aka Jupyter Notebook. Launch this application:



- It might take a minute, but this will eventually launch your web browser and will look like this:



You are now successfully running Jupyter Notebook! We'll learn more in class next week of how to use Jupyter. You can exit the application by clicking "Quit" in the upper right-hand corner.

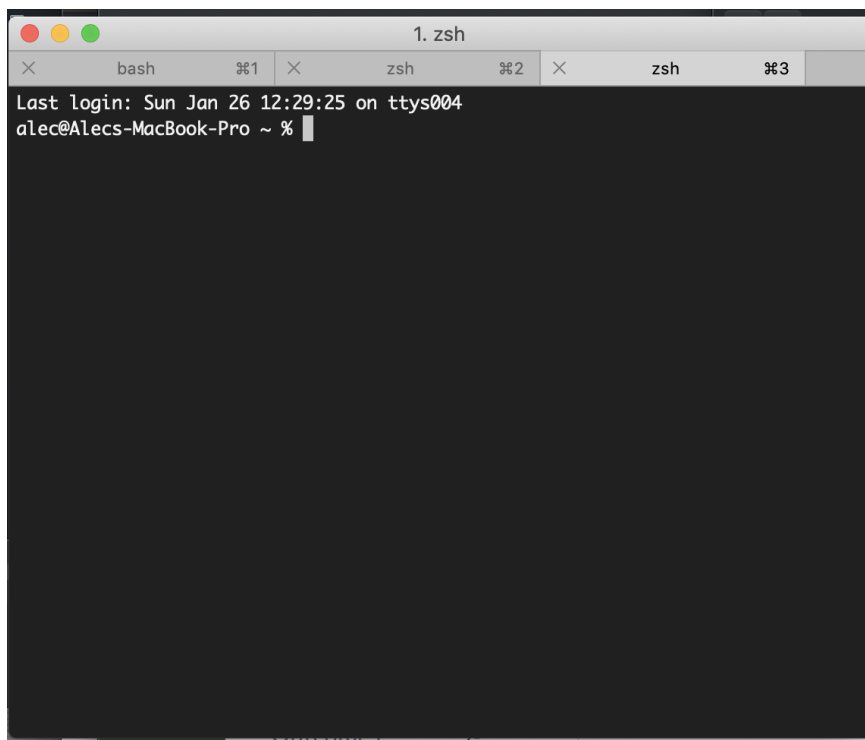
b. With a terminal

We will sometimes have to use a command-line prompt to run a few basic commands. If you've installed Anaconda, it shouldn't be too difficult. The applications will be different depending on if you're using Mac or Windows.

Instructions for Mac

Open the application "Terminal"

1. This should open a window like this one:



2. You can start Jupyter Notebook from the terminal rather than Anaconda Navigator by typing: `jupyter notebook`
3. This should launch Jupyter in a web browser. To stop the app, go back to the terminal and hold down "Ctl+c"

Instructions for Windows

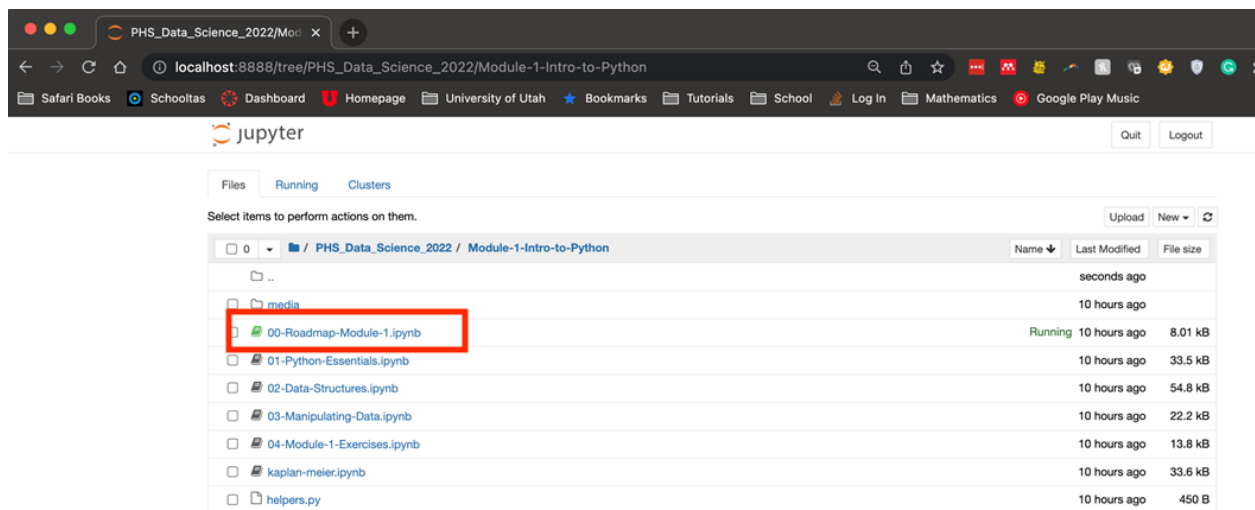
1. Open the application "Anaconda Prompt" (this can also be launched from Anaconda Navigator)

2. This will now behave similarly to the "Terminal" application used by Mac
3. You can start Jupyter Notebook from the terminal rather than Anaconda Navigator by typing: `jupyter notebook`
4. This should launch Jupyter in a web browser. To stop the app, go back to the terminal and hold down "Ctl+c"

4. Open a notebook in the class folder

Once you've started Jupyter notebook, you should be able to open and run the code for the class.

In the browser where Jupyter Notebook is running, navigate to the location you saved the code in Step #1. Go into the Module 1 folder and open the file "00-Roadmap-Module-1.ipynb".



That should then open the following file:

00-Roadmap-Module-1 - Jupyter

localhost:8888/notebooks/PHS_Data_Science_2022/Module-1-Intro-to-Python/00-Roadmap-Module-...

Safari Books

Schools

Dashboard

Homepage

University of Utah

Bookmarks

Tutorials

School

Log In

Mathematics

Google Play Music

Other Bookmarks

jupyter

00-Roadmap-Module-1 (autosaved)

File

Edit

View

Insert

Cell

Kernel

Widgets

Help

Python 3 (ipykernel)

+

%

↺

↻

⬆


⬇

▶ Run

⏏

⌂

▶ Markdown



The banner features a photograph of a modern university building with a large mountain in the background. To the right of the image, the text reads 'University of Utah Population Health Sciences Data Science Workshop'. Further right is the University of Utah Health logo, which consists of a stylized red 'U' and the word 'HEALTH' in red, with 'UNIVERSITY OF UTAH' in smaller black text below it.

Module 1

Welcome to the PHS Data Science Workshop! This class will use **Jupyter Notebooks** for writing code, running it on data, and viewing the results. Each day will start with a **Roadmap** that will outline the plans for the day.

The general outline for each day will be:

1. Introduction of today's topic and review what we learned the day before
2. Learn the skills of the day by completing educational notebooks
3. Apply our new skills to a use case of pneumonia

If you've gotten this far, then you should be good to go! If you've hit any snags, let me know and I'll help debug.