Python for Scientists & Engineers: An Introduction to Programmatic Data Analysis

Hosted by Alexandra Gunderson (alexandra@unifai.dev) and Rebecca Trojanowski (rtrojanowski@bnl.gov)

Option 1: Use Google Colab

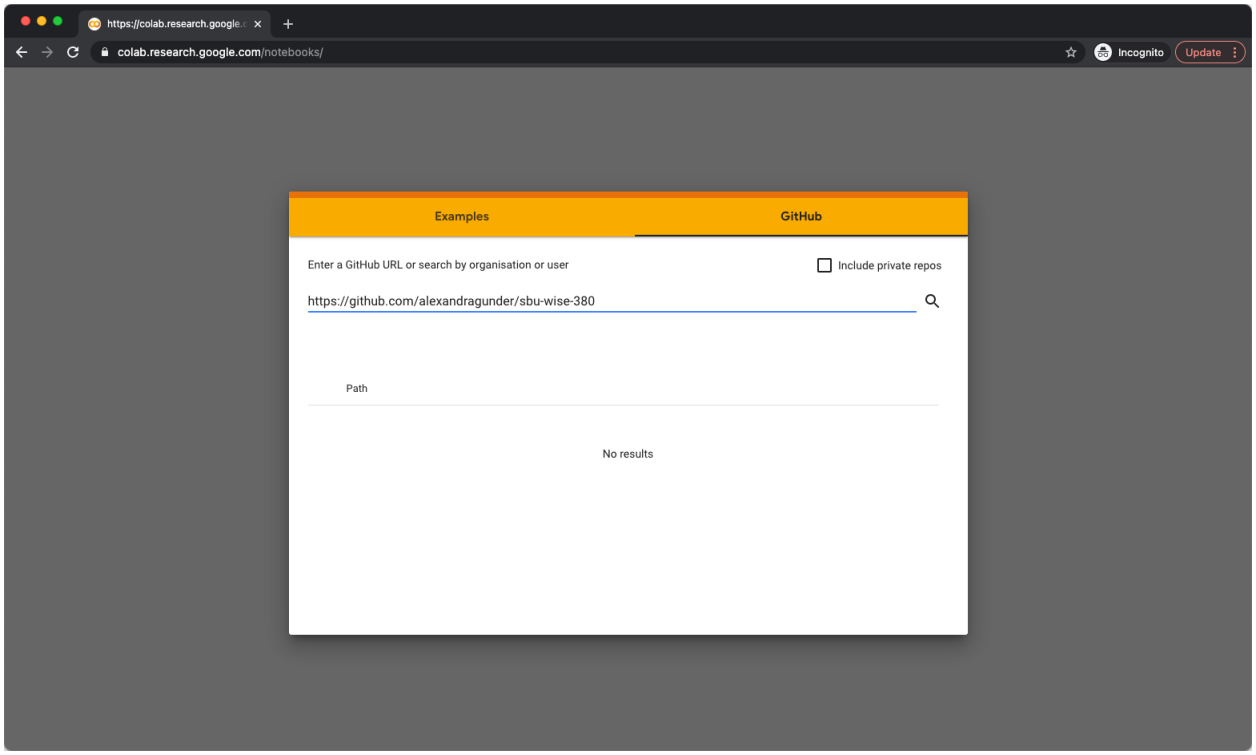
Google Colaboratory, or 'Colab' for short, allows you to write and execute Python in your browser without having to download anything.

GET STARTED

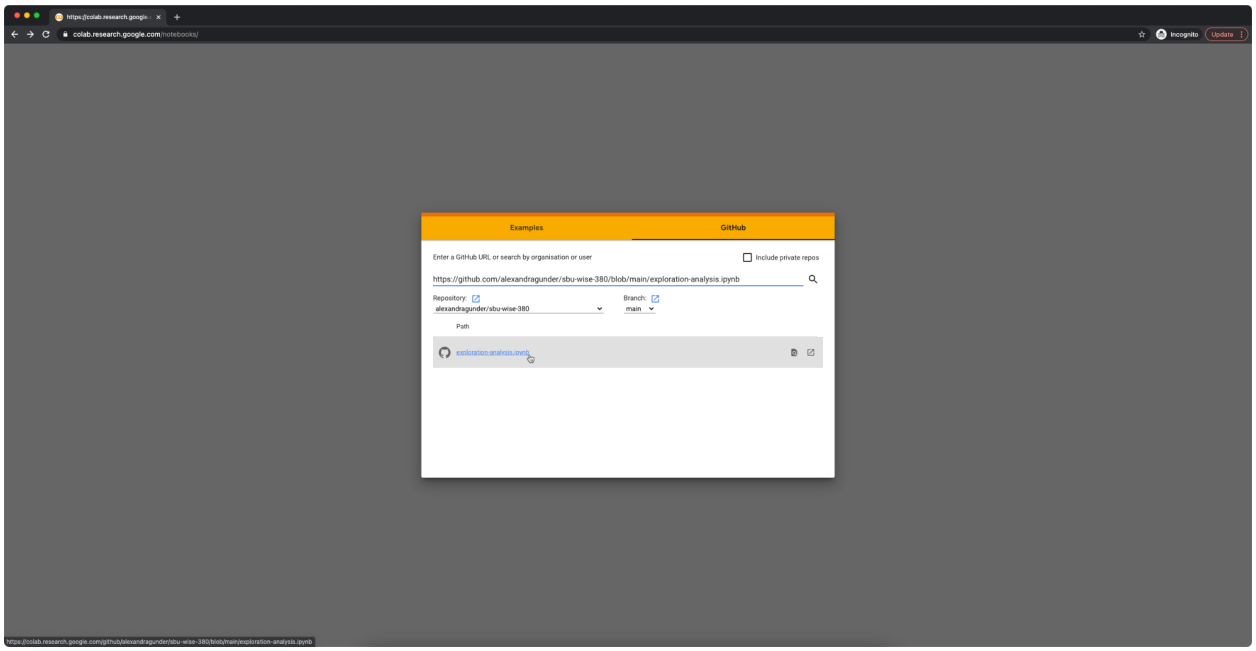
1. Go to → https://colab.research.google.com/notebooks/
2. **Click on GitHub.**
3. **Paste this repository** beside the search bar

   https://github.com/alexandragunder/sbu-wise-380/blob/main/exploration-analysis.ipynb

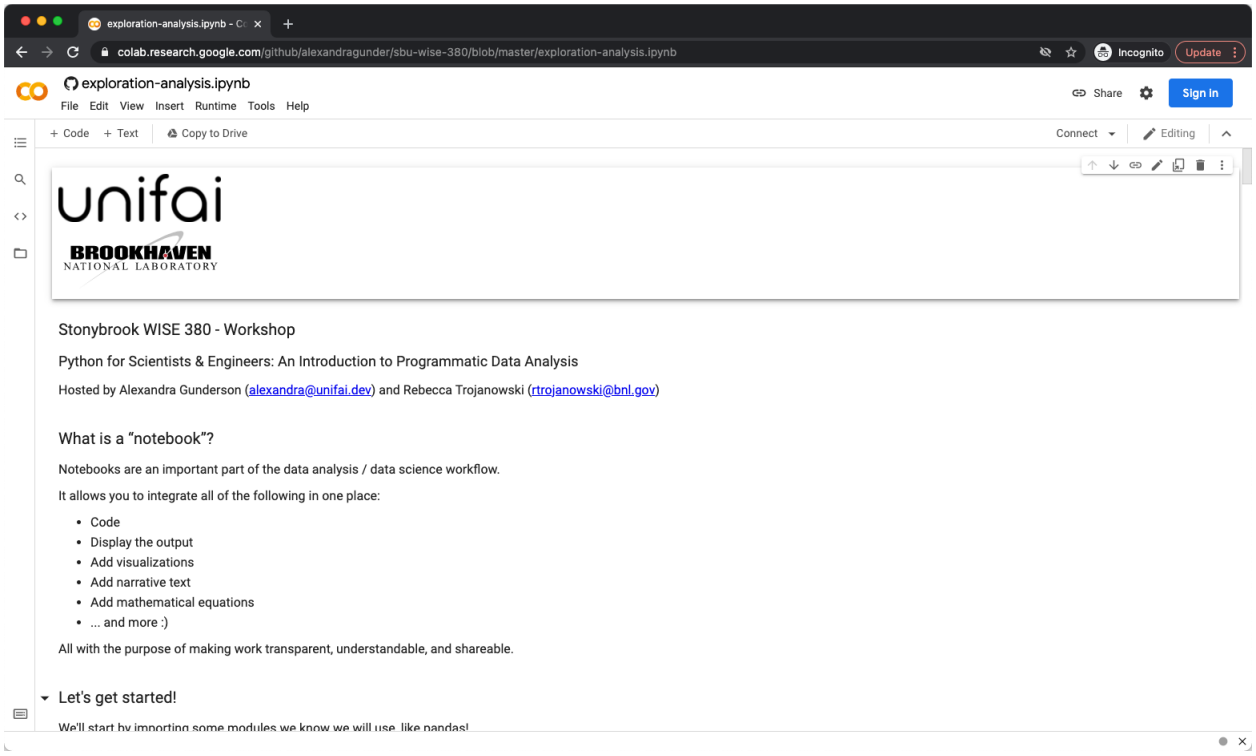
You should see a view similar to this:



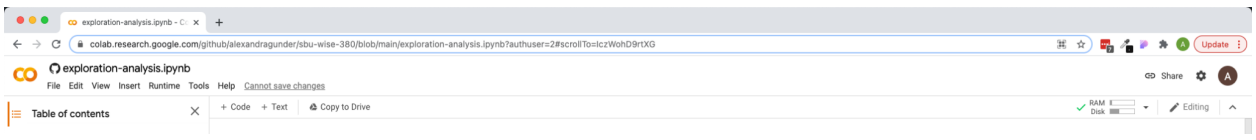
4. **Press search or press enter.** This should display a view like the following (or bring you directly to the notebook).



5. **Click on the notebook,** exploration_analysis.ipynb. This will bring you to the following view.



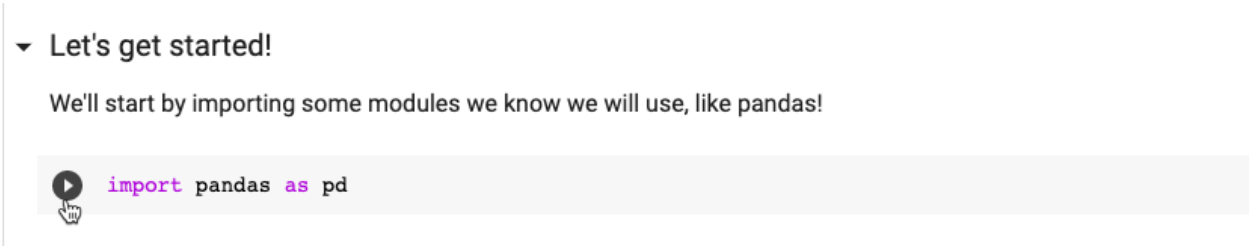
6. Click **Connect** in the top right corner. **Note: If you are not signed in to Google, you will be prompted to sign in.**
7. Once connected, you should see a view similar to this one.



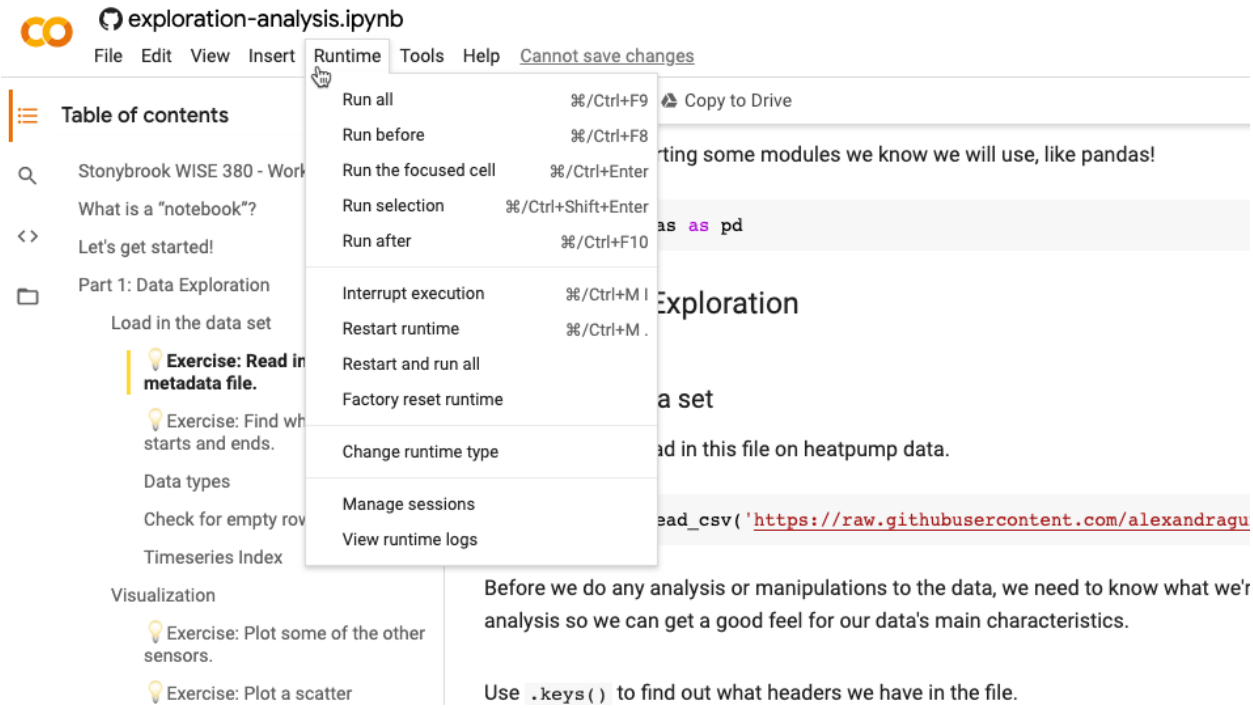
NOTE → In some cases, there might not be available resources and it might show "Allocating" for a long time. In this case, you can refresh the page and try again in a few minutes. If this still doesn't work, you might have to download Python directly. In this case, follow the instructions in the section after this.

8. **Running the code**.

You can run each cell individually by pressing the "play" button beside the cell:

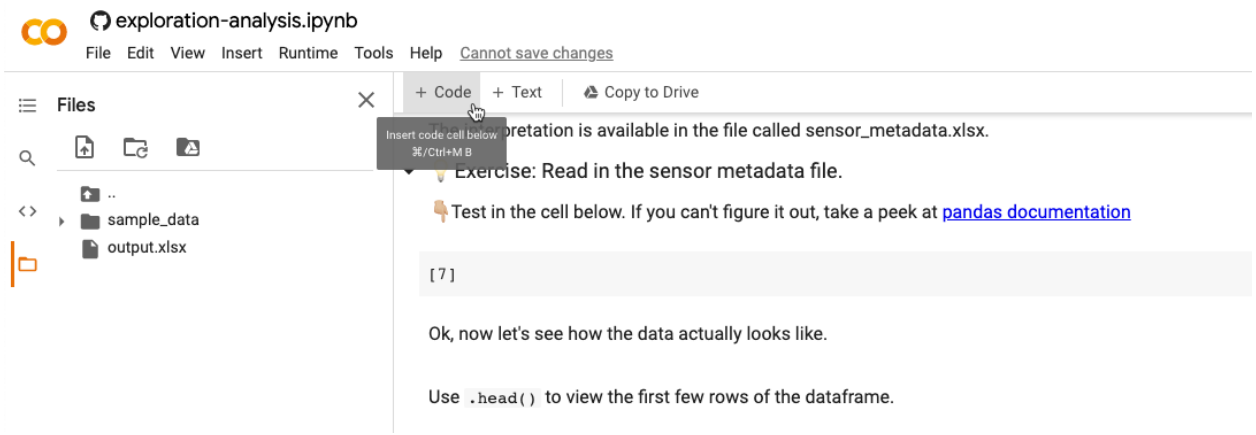


Alternatively, you can run using keyboard shortcuts, or from the menu. See here.

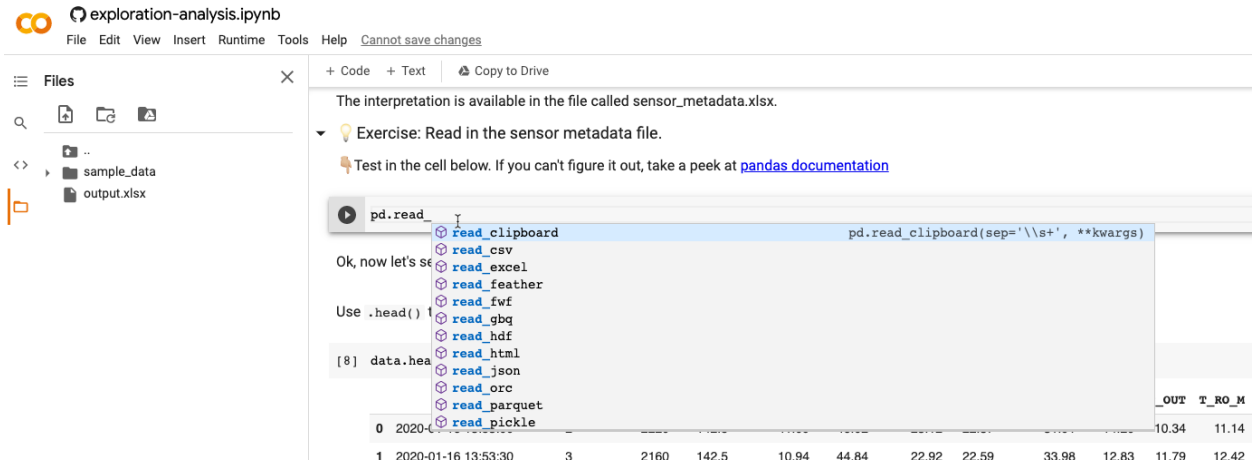


9. **Writing your own code.**

To write code, you can add cells by clicking +Code or using the keyboard shortcut.



You can also edit directly in the cells that are already available.



You'll notice that once you start typing, you get prompted with some suggestions for easy code completion :)

Option 2: Download Python

GET STARTED

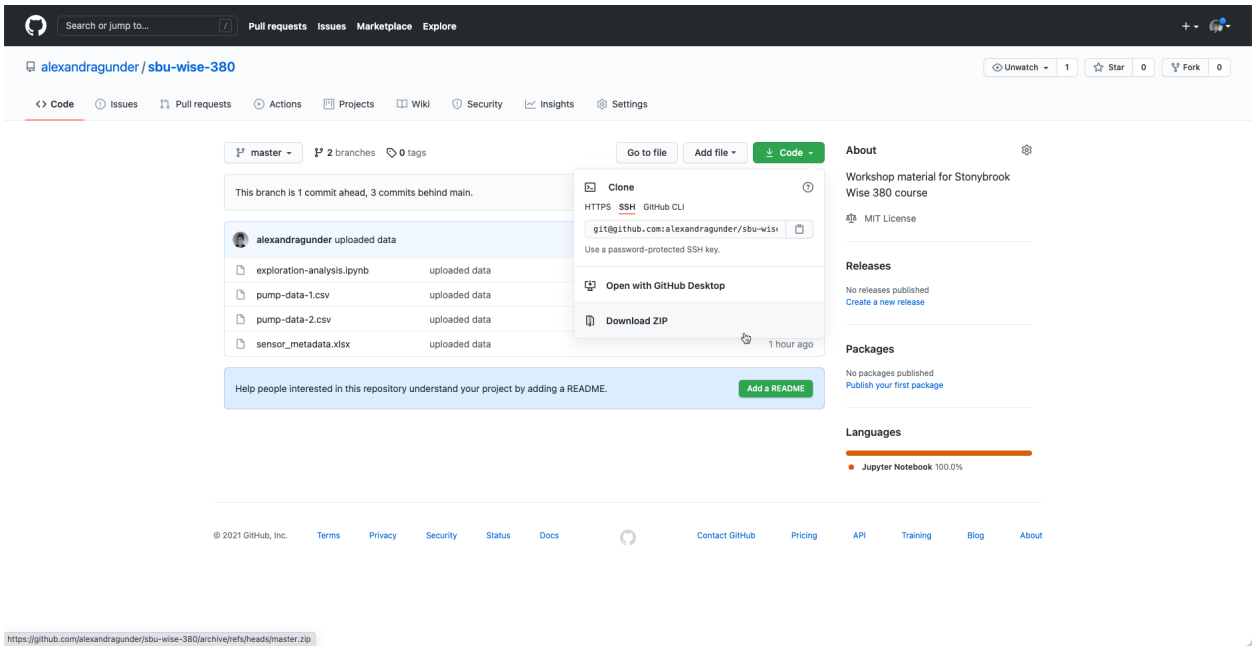
To install python, you can install Anaconda (a Python distribution with recommended libraries for machine learning) from here → https://www.anaconda.com/products/individual

Follow the instructions for downloading according to your operating system.

DOWNLOADING THE DATA

Since you are using your own local version of Python, you will also need to download the data.

1. Go to this link https://github.com/alexandragunder/sbu-wise-380/tree/master

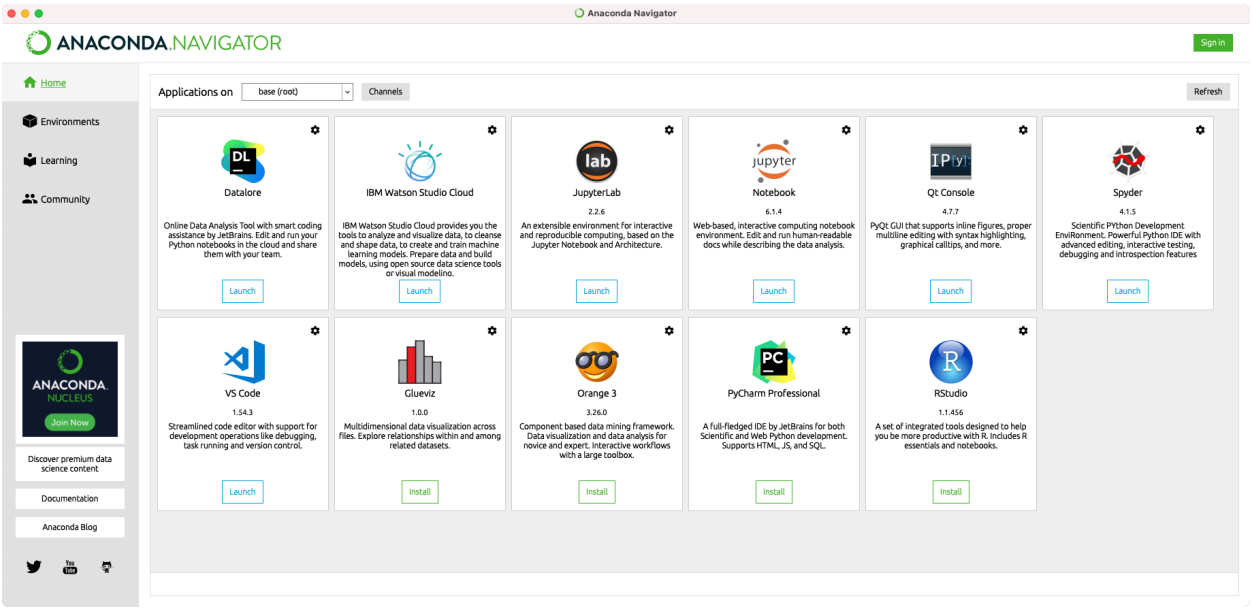


https://github.com/alexandragunder/sbu-wise-380/archive/refs/heads/master.zip

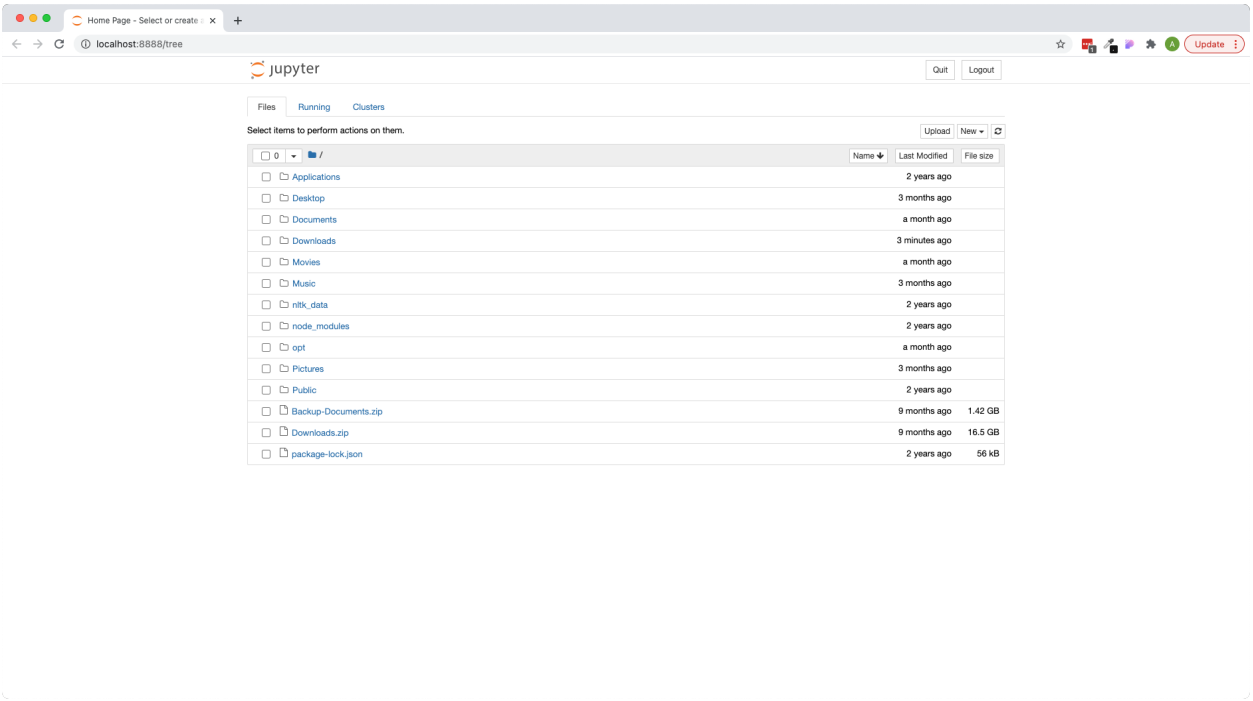
2. Click on the Code button and "Download Zip"
3. Unzip the folder and put it somewhere you'll remember :)

GETTING STARTED

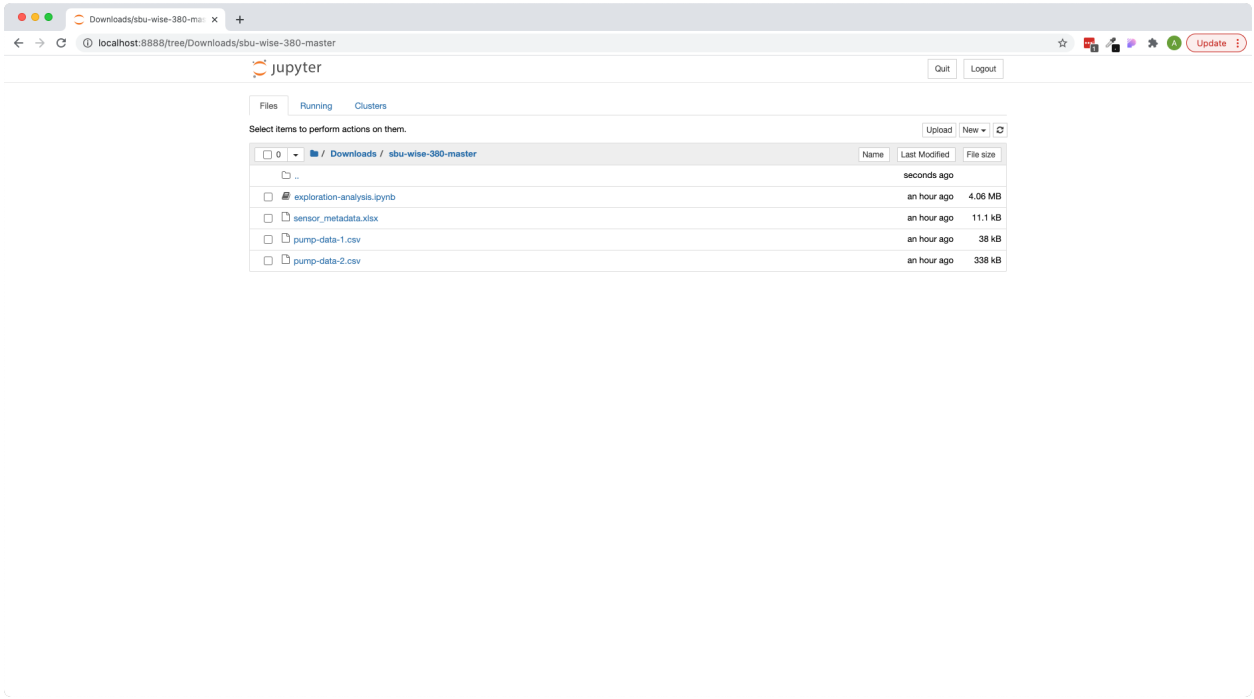
1. Open Anaconda Navigator. This should bring you to somewhere like this:



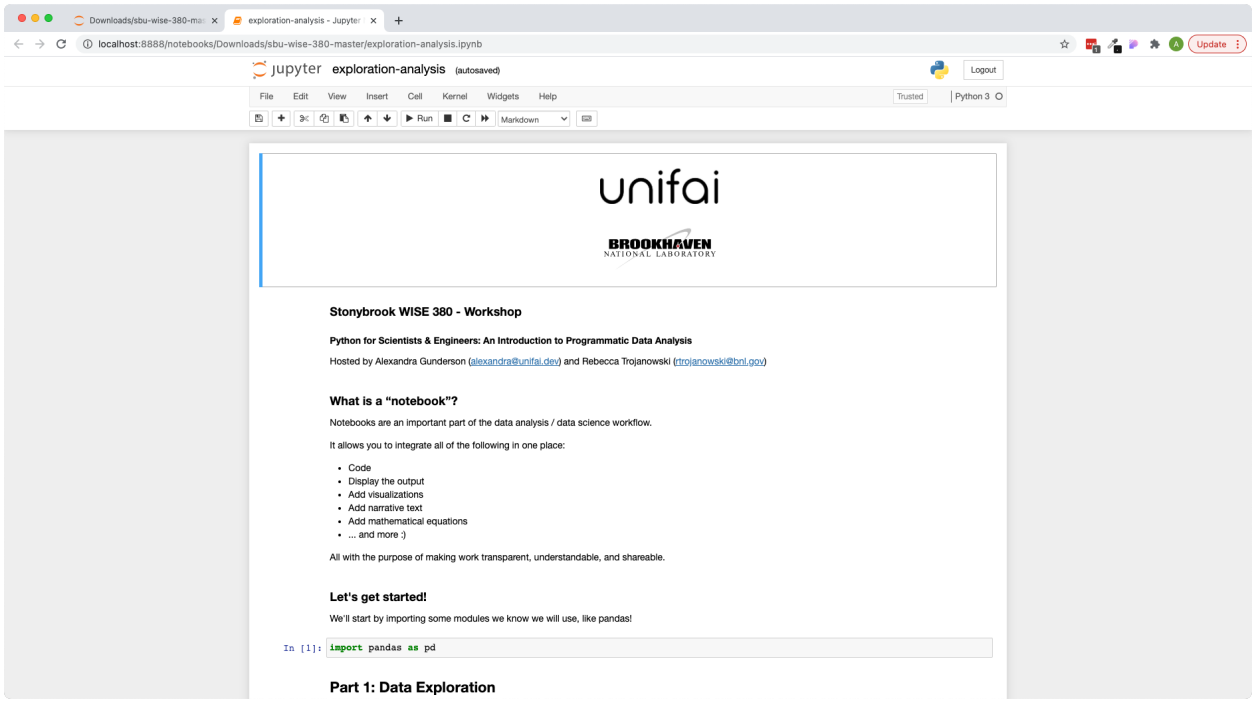
2. Click Launch under Jupyter Notebook. This should open a browser window that looks like this.



3. Navigate to the folder you saved the files we downloaded earlier.



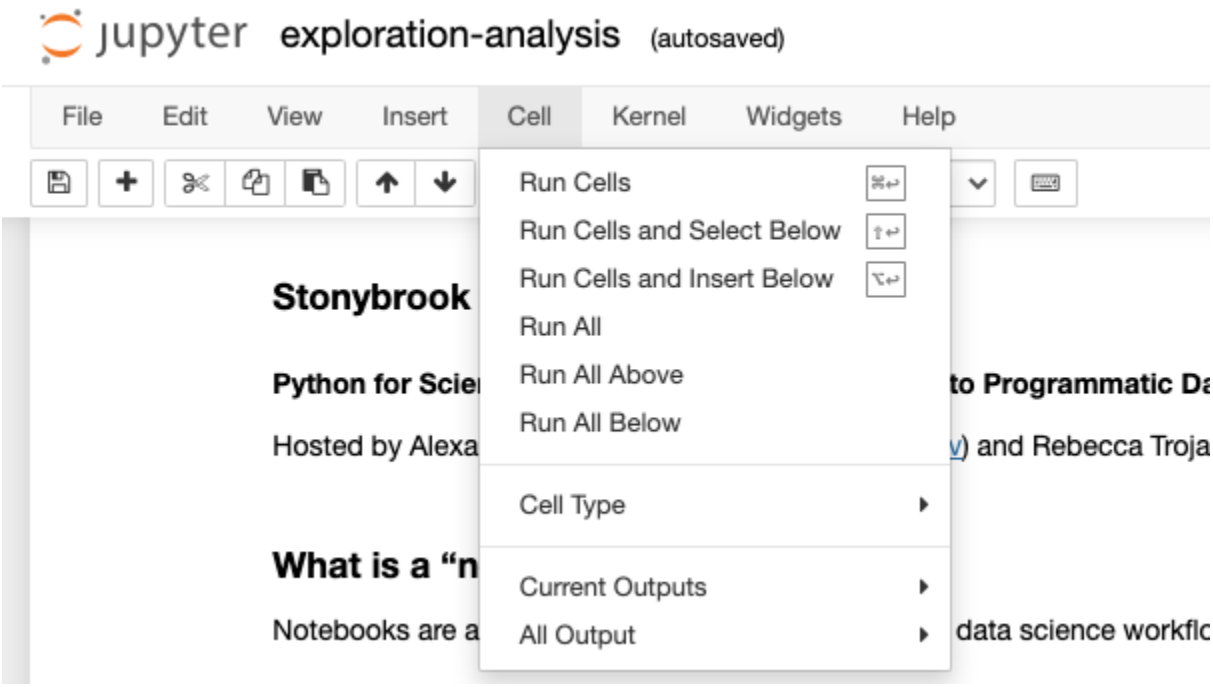
4. Click on exploration_analysis.ipynb



5. Woohoo! If you've made it this far, congratulations :) Now we're ready to run the code. You have a couple of options for running the code.

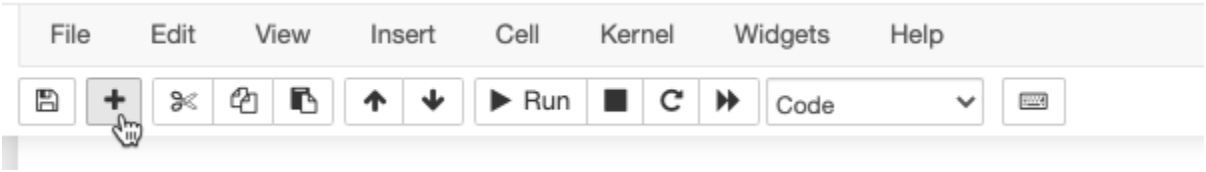
   To run a cell, click in it and press "Run"

For other alternatives, look to the Cell menu, where you can also find keyboard shortcuts.

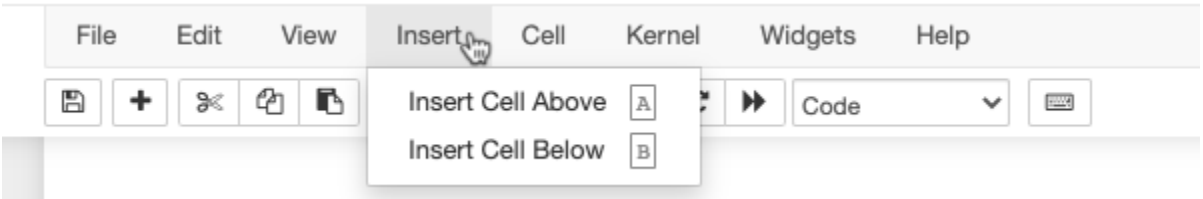


6. To edit cells, you can click directly in the cell and type.

If you want to add additional cells, click the "+" button



Alternatively, click Insert on the menu and see your options.



7. If you come across any error message, such as "ModuleNotFoundError", this is because the library hasn't been installed yet. Don't fret.

```
---------------------------------------------------------------------------
ModuleNotFoundError                       Traceback (most recent call last)
```

To install libraries, you can do it directly in the code, by typing

```
!conda install NAME-OF-LIBRARY
```

If that doesn't work, send us a message and we will figure it out. Sometimes packages aren't available from conda, so you can also try replacing `conda` with `pip`. In other cases, it is sometimes the name of the package in the package management system is named differently.

Once you have successfully installed the package, you can rerun the code and it should work without that error.

NOTE → We haven't gone through any Python best practices. When you run Python code, there are a lot of things to consider with versioning (which Python version, which version of each module, etc.) if you want to have truly reproducible code. To do this, we often create what is called "environments". You can read more about that here → https://realpython.com/python-virtual-environments-a-primer/