Collective Logic Lab - Running Jupyter to work with data

Notes:

- *Documentation is done by Diya Hamada (dhamada@asu.edu) for M2 Mac.*
- *The example shown is forage_model_CLL.ipynb for the Honey Bee Project's analysis of foraging behaviors, available in the lab's GitHub*
- *Last updated: Jan 29th, 2024*
- *Editing link to this Google Doc: https://docs.google.com/document/d/1ZdfKKkZ9LerUyaIPJk8KVDj26dshBabdMP_sWvcRy28/edit?usp=sharing*

Step 1: Download relevant data and files to your computer.

For the Honey Bee Project, this is in Zenodo:

2018 data: https://zenodo.org/records/6045860

2019 data: https://zenodo.org/record/7298798

This may take a while! When it is complete, move all data into a single folder. That can go wherever you want on your computer, just be sure it is easy to find later.

This is also a good opportunity to download what you need from the Collective Logic Lab GitHub repository (ask Dr. Daniels for an invite if you do not have access already). I recommend keeping everything in the same folder.
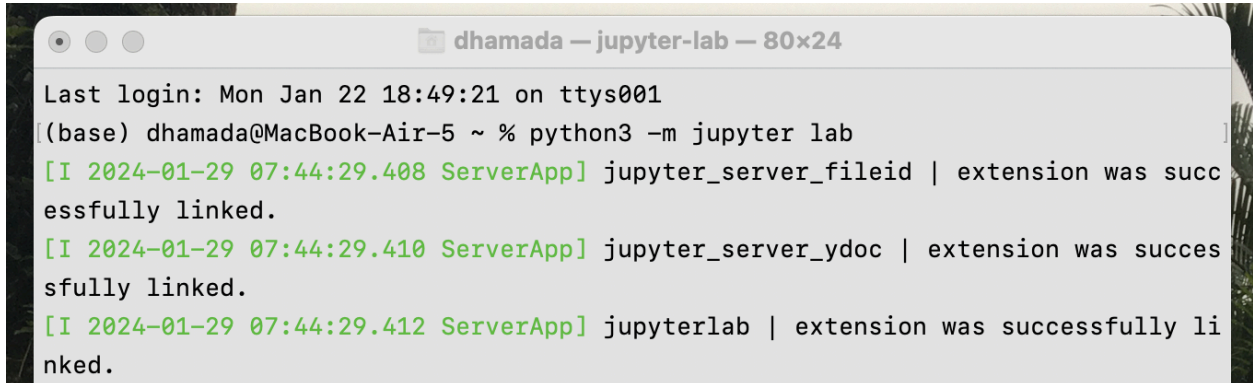
Step 2: Open Jupyter Lab

You will need to have python3 on your computer. If you need help with that, I recommend checking this link out: https://realpython.com/installing-python/

You will also need Jupyter Lab. Here's another helpful link:

https://jupyterlab.readthedocs.io/en/stable/getting_started/installation.html
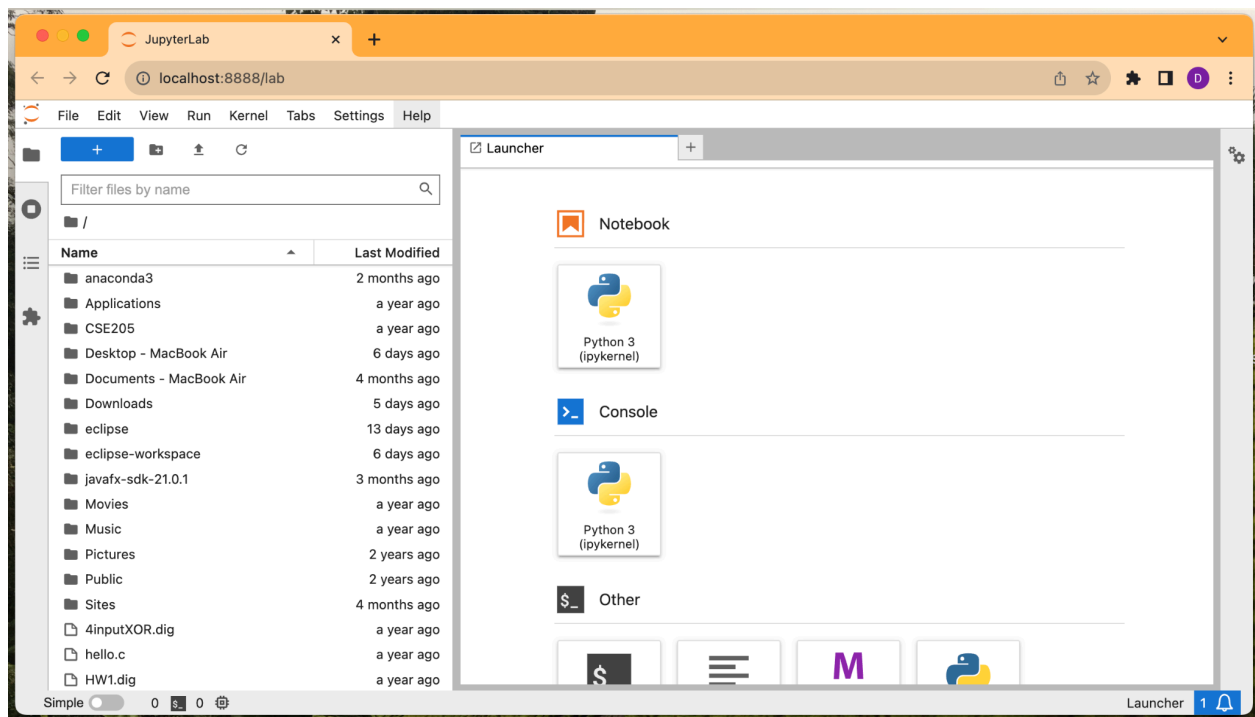
I use my Terminal to open Jupyter Lab with the command "**python3 -m jupyter lab**" as shown below.


After running the command, many lines of text should run across your Terminal, and then this tab should pop up in your browser:

Step 3: Open a Notebook and import data

If you already have a Notebook you would like to run, navigate to it in the bar on the left. Otherwise, use the tabs at the top of the screen and go to File > New > Notebook.

Here is my Notebook, forage_model_CLL.ipynb. For those working on the Honey Bee Project, I highly recommend keeping the data, Notebooks, and other files from GitHub in the same folder for convenience, as shown on the left.
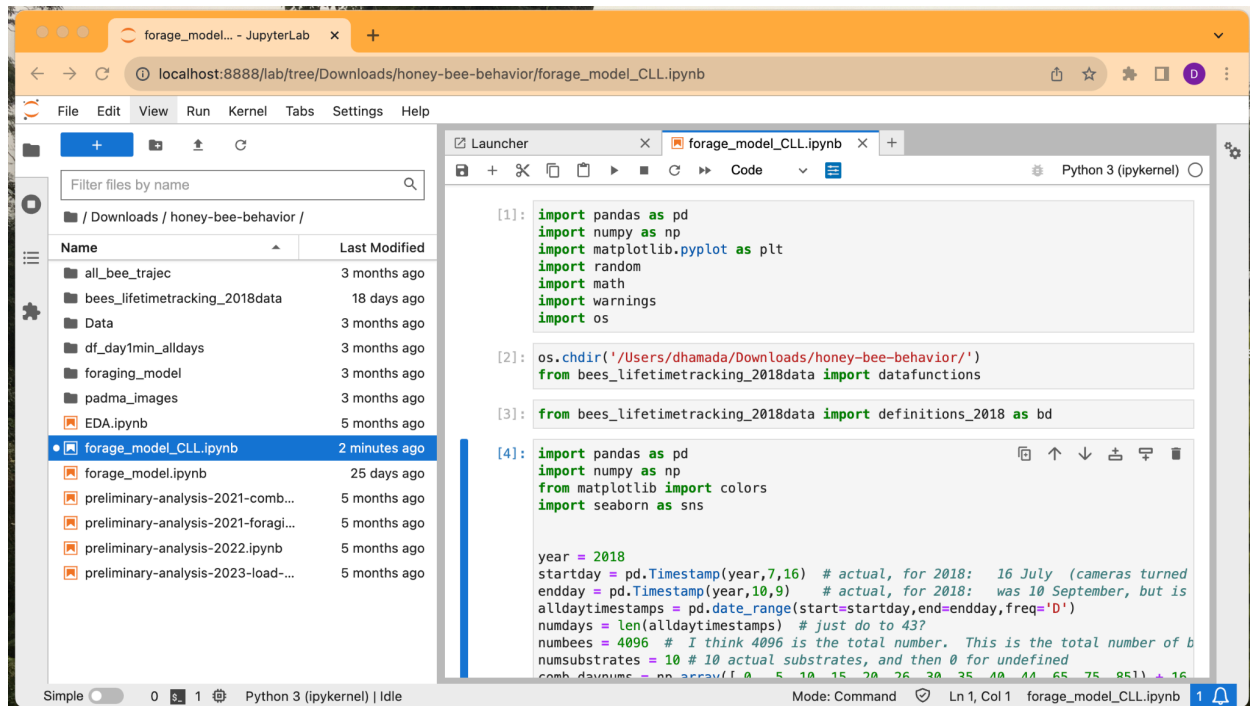
To import data or functions, start with "import os" and then use os.chdir('FILEPATH") to navigate to the folder with your project files as shown in cells 1 and 2 of my Notebook. For Mac users, FILEPATH can be found or checked by locating the folder in Finder, clicking once to highlight it in blue, and holding the "option" key.

Finally, use the keywords "from" and "import" as shown in cells 2 and 3 for each file to be imported (datafunctions and definitions_2018 are both .py files but note that I have dropped their file types).

"as" can be added (cell 3) if the file has an excessively long name and you'd like to call it something nice and short while you work.

If there are no errors when you run an "import" command, it's worked!

More Notes:

An easy alternative to Jupyter Lab would be Google Colab. No setup or downloads are required, and files on your computer can be easily uploaded and read—extremely similar UI to Jupyter Notebook.

IMPORTANT: If you are using definitions_2018, you need to delete the "freq" parameter from line 20. New versions of pandas do not support it anymore. You can double-click your definitions_2018.py in Jupyter Lab, edit it, save it (command+s), and run your Notebook with no issues.

Line 17 path also needs to be changed to location of cohort data file.

If you are running forage_model_CLL.ipynb, in line 40 near the end, "df.append(temp)" needs to be changed to "df = pd.concat([df, temp])"