Analytics Jumpstart

Introduction

Mike Schriefer, Instructor/ Data Engineer at Stratasan

Dalila Paquiot, Data Scientist at XSOLIS

Julia Romano, Integration Engineer at PlayMaker Health

Mary van Valkenburg, Data Science + Analytics Program Manager

NASHVILLE SOFTWARE SCHOOL

- Your name
- The place you call home
- Something new in the last (pandemic) year

NASHVILLE SOFTWARE SCHOOL

Classroom *guidelines*

- Ask lots of questions
- Help each other; learn from each other
- Get comfortable with discomfort. Making mistakes, figuring them out, and then correcting them is part of the learning process
- After working through the assignment, form your own ideas and do your own exploration beyond what has been suggested

NASHVILLE SOFTWARE SCHOOL

Class format

1. Concepts/Code Lecture
2. Coding practice
3. Interactive with instruction team and other students!

The weekly targets are for help in pacing lectures. Feel free to work ahead, but be aware you may need to use Google or the pandas API!!

NASHVILLE SOFTWARE SCHOOL

Goals for the class

- Get hands-on experience of what it might be like to work as a data analyst or data scientist
- Get an idea of whether or not this might be a good fit for a career
- Learn some tools to help you on personal analysis projects
- Make discoveries and have fun

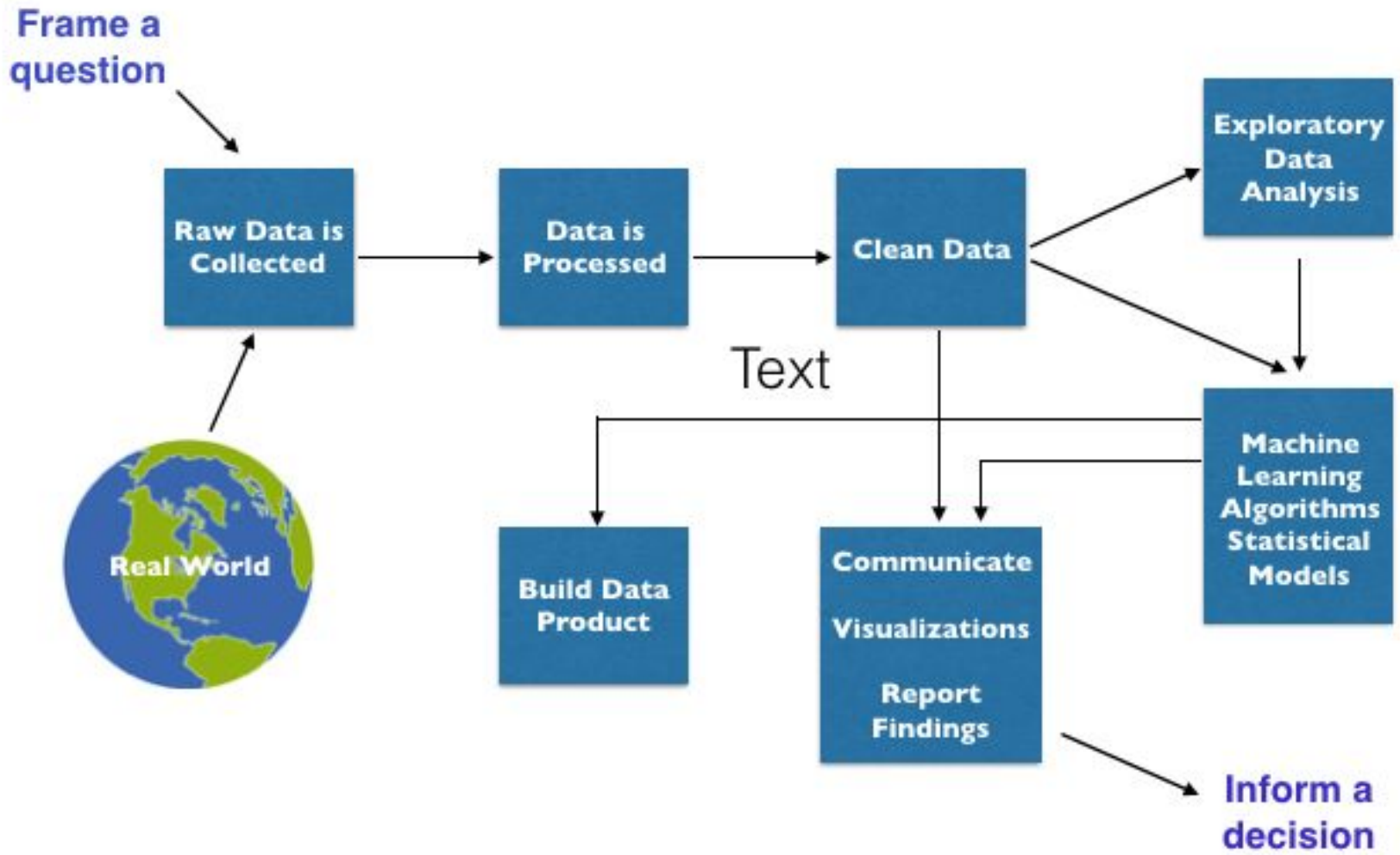
NASHVILLE SOFTWARE SCHOOL

Goals for today

- Get an idea of what to expect
- Define the Data Science Process – *a basic mental model for analysis*
- Folder Setup
- Jupyter Notebook walkthrough/orientation
- Analysis Guide walkthrough
- Get started on the project!

NASHVILLE SOFTWARE SCHOOL

The Data Science Process

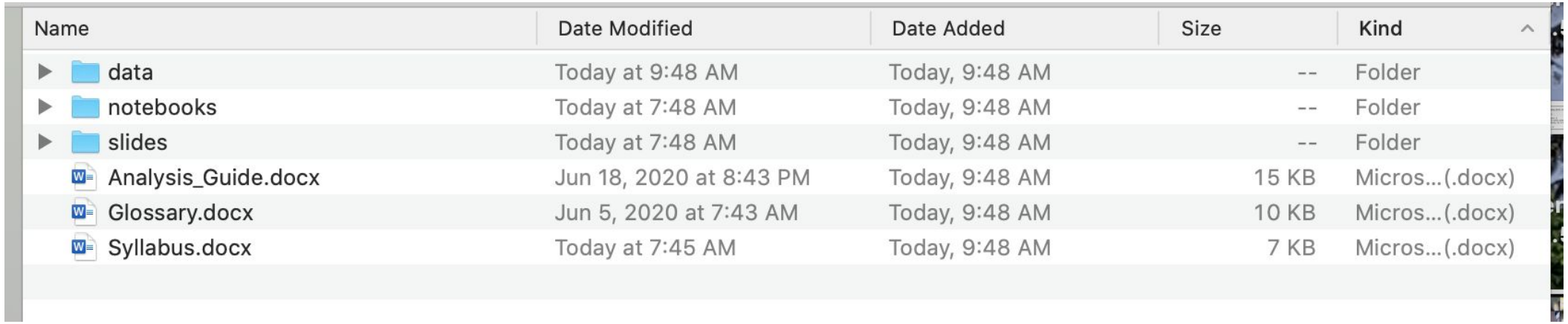


NASHVILLE SOFTWARE SCHOOL

Organize your files in the structure shown

We will share a Zip file you can unzip and move to where you want (Documents, Desktop, etc)

Do not put your files on the cloud (One Drive, iCloud)

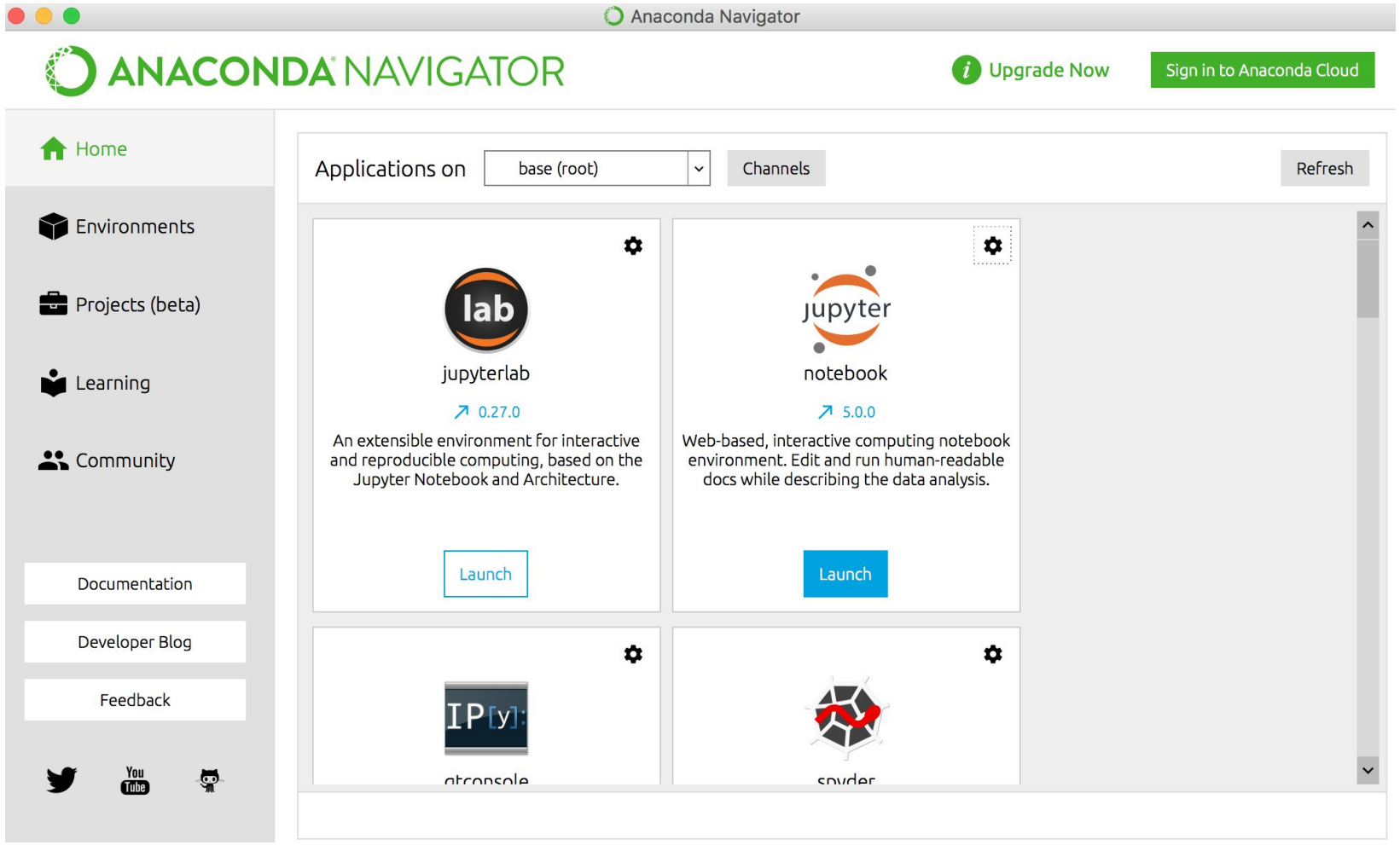


Name	Date Modified	Date Added	Size	Kind
data	Today at 9:48 AM	Today, 9:48 AM	--	Folder
notebooks	Today at 7:48 AM	Today, 9:48 AM	--	Folder
slides	Today at 7:48 AM	Today, 9:48 AM	--	Folder
Analysis_Guide.docx	Jun 18, 2020 at 8:43 PM	Today, 9:48 AM	15 KB	Micros...(.docx)
Glossary.docx	Jun 5, 2020 at 7:43 AM	Today, 9:48 AM	10 KB	Micros...(.docx)
Syllabus.docx	Today at 7:45 AM	Today, 9:48 AM	7 KB	Micros...(.docx)

NASHVILLE SOFTWARE SCHOOL

Orientation to Jupyter Notebook

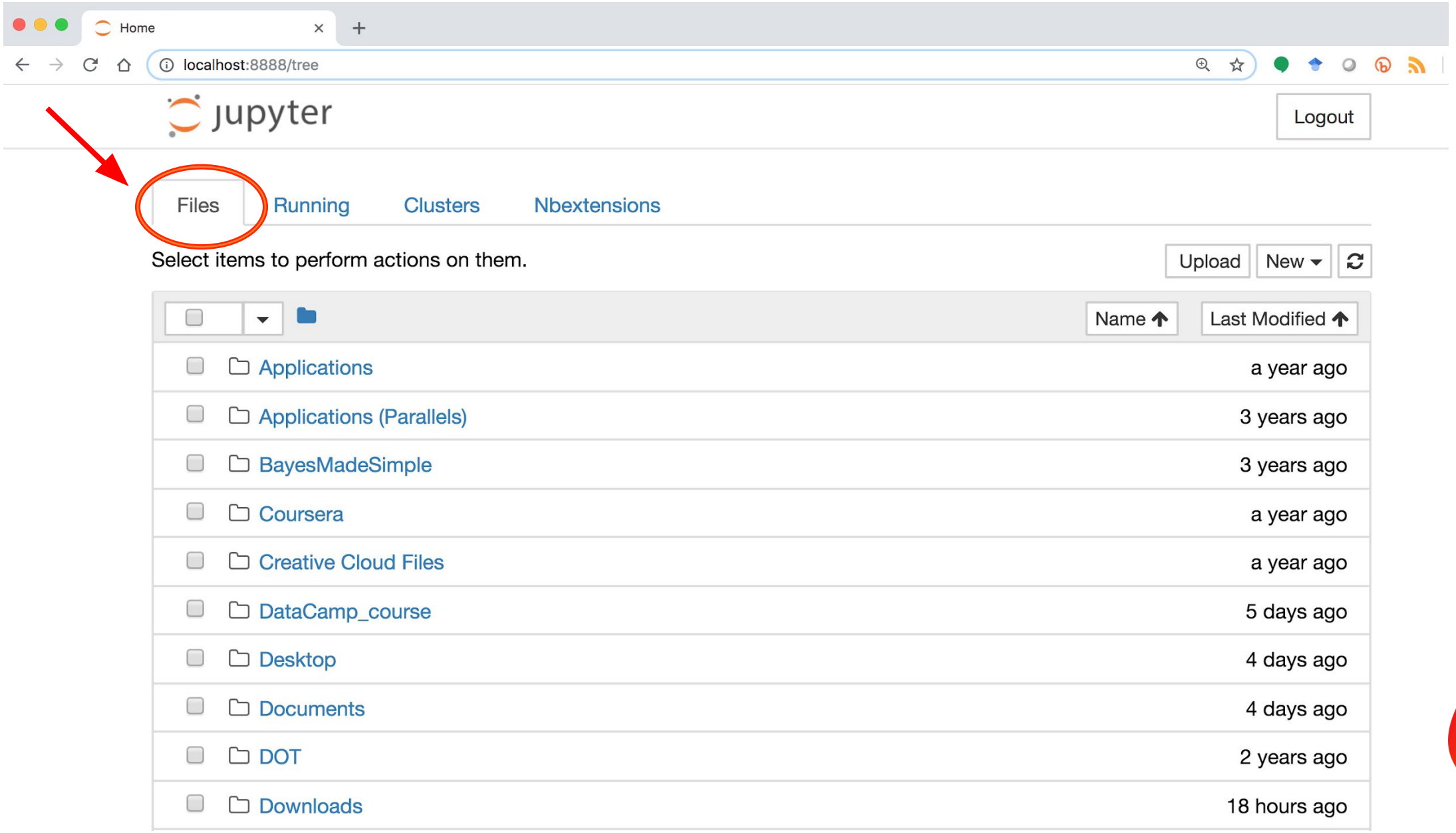
Open Anaconda Navigator, install, and launch Jupyter Notebook



A new tab will open in your default browser. It's not actually connecting to the internet, just running on your machine

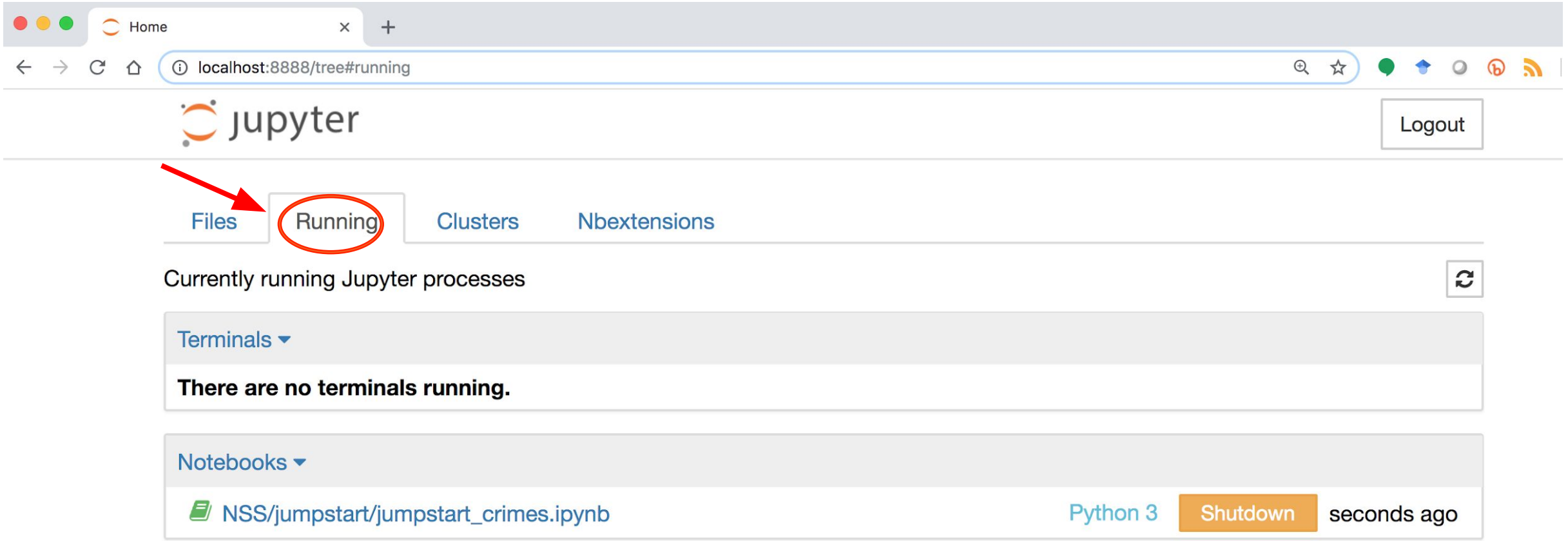
NASHVILLE SOFTWARE SCHOOL

You will see the file structure on your computer and can navigate as normal



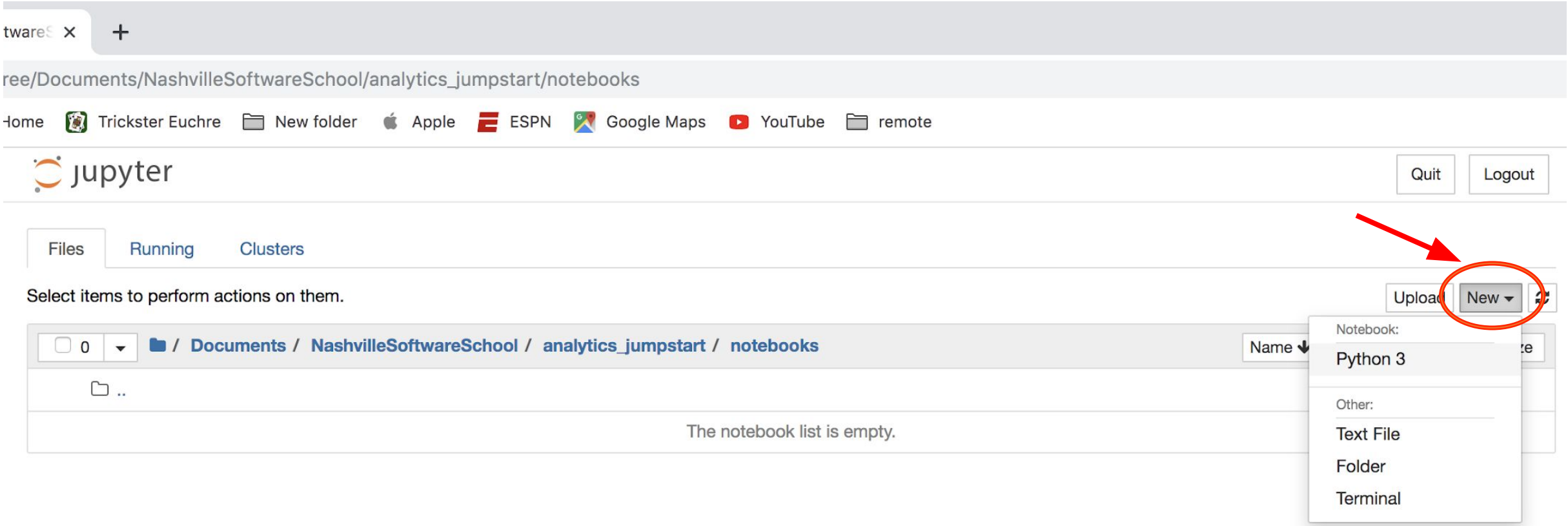
NASHVILLE SOFTWARE SCHOOL

See what notebooks are already running (should be empty if just opening Jupyter)



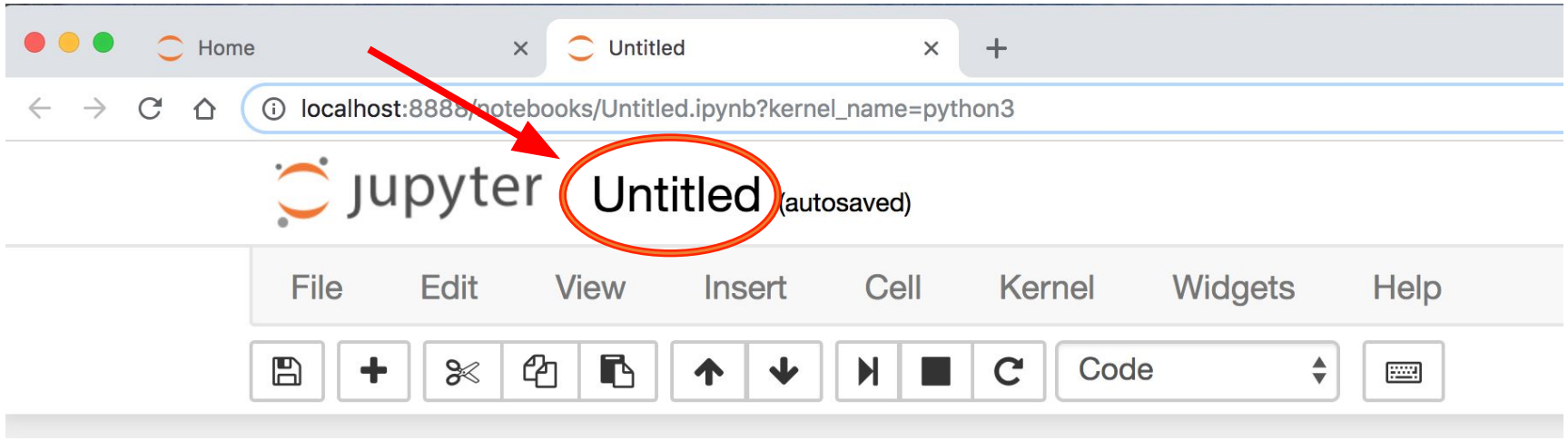
NASHVILLE SOFTWARE SCHOOL

Navigate to analytics_jumpstart/notebooks and create a new Python 3 notebook

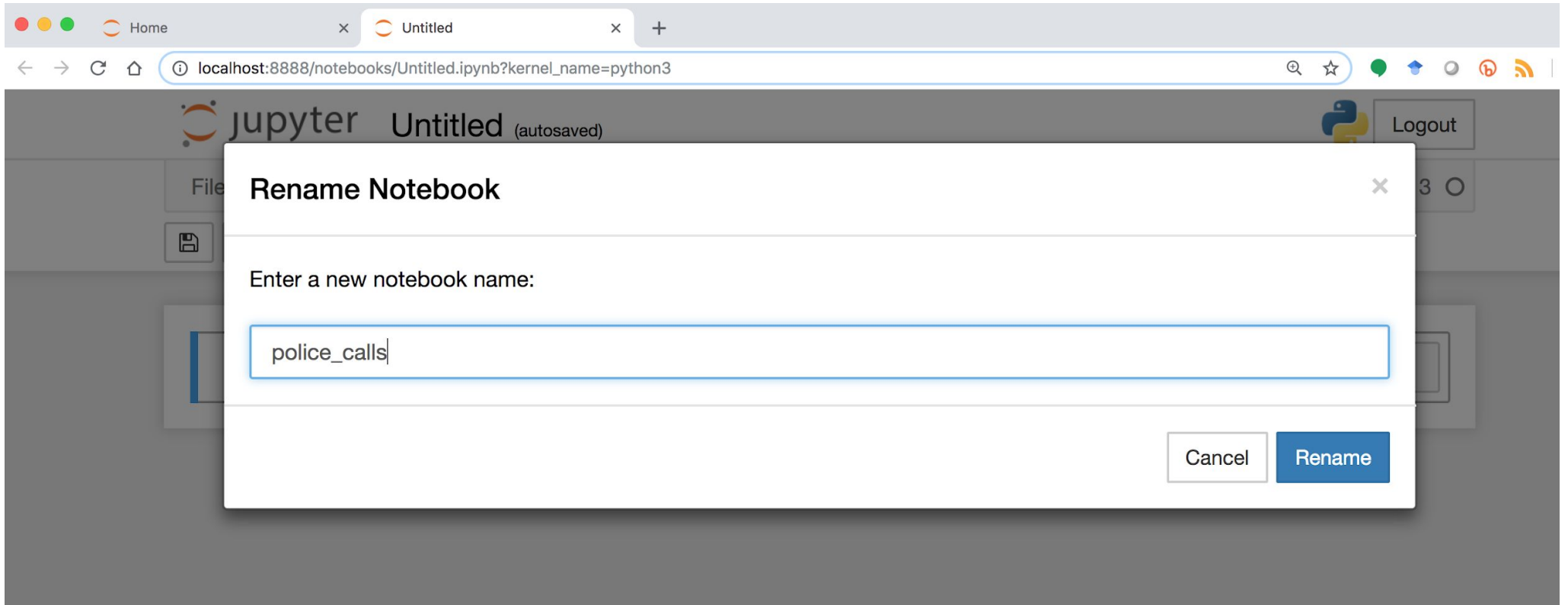


NASHVILLE SOFTWARE SCHOOL

Title the notebook

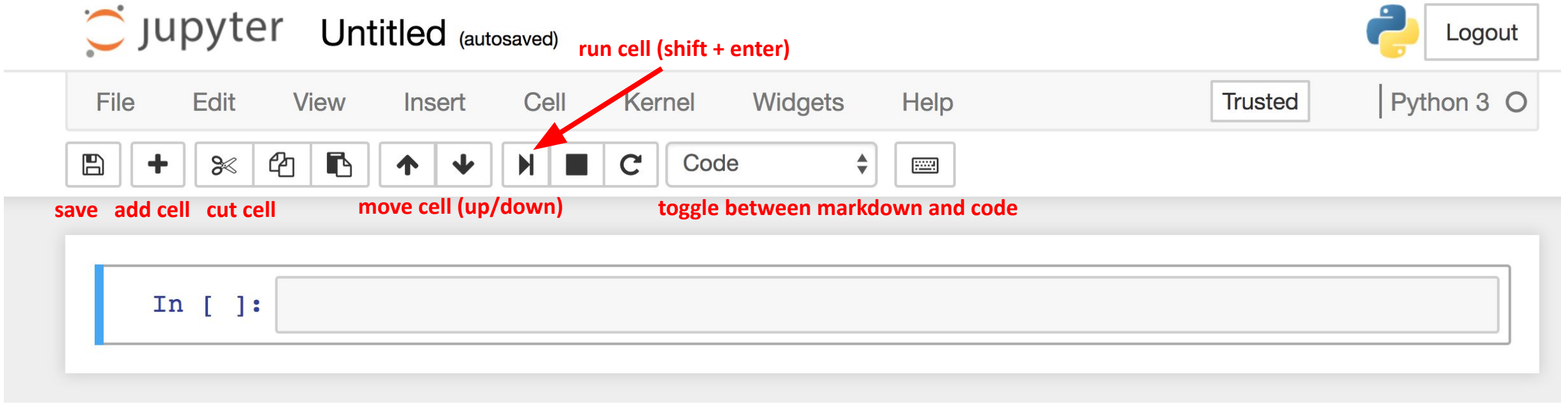


- **meaningful**
- **no spaces**



NASHVILLE SOFTWARE SCHOOL

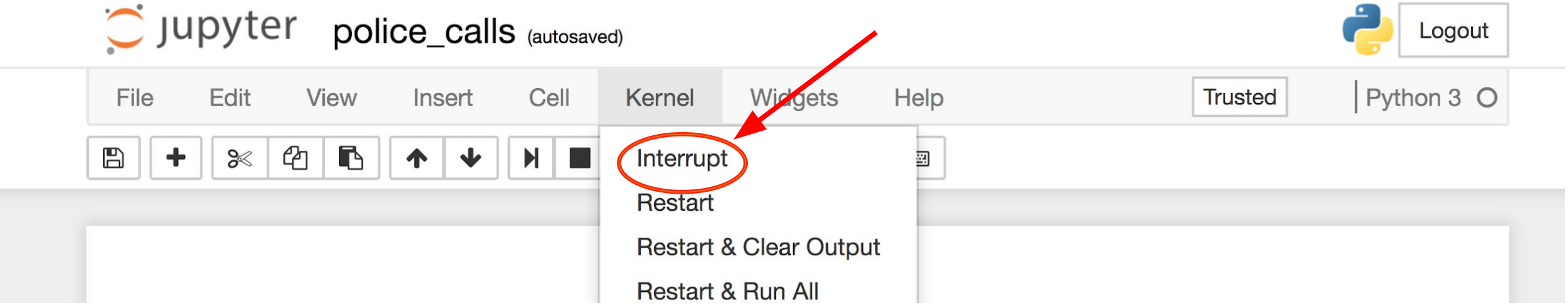
Useful buttons (and shortcuts) for running code and moving cells around



NASHVILLE SOFTWARE SCHOOL

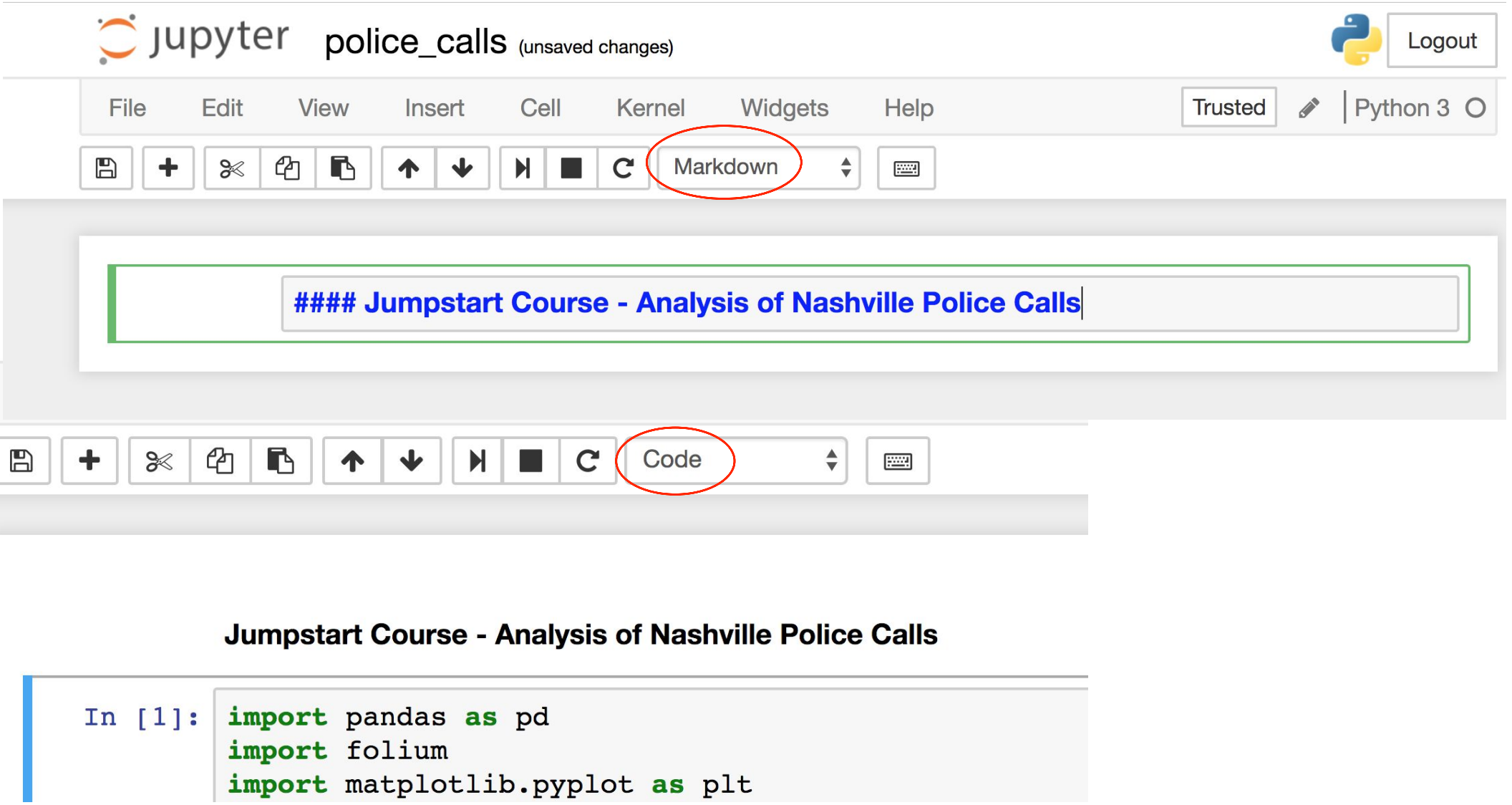
If your code is taking too long or is giving unexpected results, try restarting the kernel

Each time you run a piece of code in a Jupyter Notebook, that process is saved to a kernel. All the inputs, outputs, variables, etc. are saved. Even if you modify or delete a cell, the first time it was run was saved. This can sometimes lead to strange results. Restarting the kernel will clear out the memory so you can start fresh. Closing and opening the notebook will also do this. But remember to rerun all your code after you restart the kernel!



NASHVILLE SOFTWARE SCHOOL

Change the format of the cell to add notes or run code



NASHVILLE SOFTWARE SCHOOL

Next up:

- launch Jupyter
- walk through analysis guide