

# Jupyter: Notebooks, Lab, and Hub

Barron H. Henderson and Ben Murphy

2021-11-18

*The views expressed in this presentation are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency*

# Outline

- Meet a Jupyter Notebook
  - Explorer tab, png, csv, json, terminal, console
  - Run a notebook (COVID or AQS)
- Jupyter Lab Introduction
  - Basic overview
  - Components and understanding
  - Ready cloud Environments
- Atmos access instructions
  - How to get to access JupyterHub
  - Prepare frequently used libraries
- Walk through Air Quality System (AQS) observation notebook

Hands on!

# EPA Scientists need Jupyter Lab

- Jupyter Lab is a notebook with web-based file browser and terminals
  - Docs, code, results, and images
  - Reusable with a minimum effort
  - R, Python, Bash, even C
- Easy install windows, linux, mac
- Jupyter Notebooks are a great way to share analysis systems.
  - Easily exported to PDF or HTML for sharing
  - Easily reproducible when dependencies are clear\*



# Meet a lab notebook

- To get folks excited, we're going to walk through a notebook.
- When we're done you should be able to answer the following questions:
  - Does jupyter lab have a file explorer?
  - Can I access the Terminal in a jupyter lab?
  - Can I write documentation and math in notebooks?
  - How do I run code in a notebook?
  - Can I interactively make figures?
  - What languages work in Jupyter?
- Because not everyone here uses CMAQ, we're going to look at COVID data. (atmos://home/bhenders/Notebooks/COVID\_JHE.ipynb)

File  
Browser

The screenshot displays the JupyterLab interface. On the left, the 'File Browser' tab is active, showing a file tree for the path `home/bhenders/Notebooks/`. The file `COVID_JHE.ipynb` is selected and highlighted in blue. A blue circle highlights the folder icon in the top toolbar, and a blue line points from the text 'Make New Notebooks or Terminals (off screen)' to the '+' icon in the same toolbar. The main area shows the 'Launcher' tab, which displays a grid of application icons. A blue circle highlights the icons for 'Python 3', 'anaconda', 'Bash', 'basic38', 'C', 'gcei', 'gcpy', 'gnuplot', 'R', and 'SSH'. The bottom status bar shows 'Simple', '0 s', '1', and 'Saving completed'.

File Browser

Make New Notebooks or Terminals (off screen)

COVID\_JHE.ipynb

home/bhenders/Notebooks

Notebook

Python 3, anaconda, Bash, basic38, C, gcei, gcpy, gnuplot, R, SSH

Console

Python 3, anaconda, Bash, basic38

Simple 0 s 1 Saving completed Launcher

The screenshot shows the Jupyter Notebook interface. On the left, a sidebar contains a 'Table of Contents' section with a list of items: 'Jupyter Notebook Example', 'Installing a custom library', 'Import Libraries', 'Loading COVID Data', 'Cumulative Deaths by day in columns', 'Mix of Province/State and Country-Level data', 'Plot Cumulative Deaths', and 'Now create some maps'. A blue circle highlights the 'Table of Contents' icon in the sidebar. The main notebook area displays the 'Jupyter Notebook Example' content, which includes a title, a paragraph about visualizing COVID-19 data, a list of six examples, a data citation, and a section titled 'Installing a custom library'. A blue circle highlights the 'Run' button (a play icon) in the top toolbar. A blue oval highlights the 'Examples' list. A blue line points from the 'Installing a custom library' section in the notebook to the 'Add Documentation' section on the left. The bottom status bar shows 'Simple' mode, '0' lines, '1' column, 'anaconda | Idle', 'Saving completed', 'Mode: Command', 'Ln 1, Col 1', and 'COVID\_JHE.ipynb'.

Run a "Cell"

Table of Contents

Add Documentation

- Use Markdown
- Use LaTeX

Jupyter Notebook Example

Visualizing data from [COVID-19 Data Repository](#) by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University.

Examples:

1. Install a library for spatial plotting
2. Load data from the web
3. Add metadata and perform calculations.
4. Make plots
5. Time-series
6. Maps

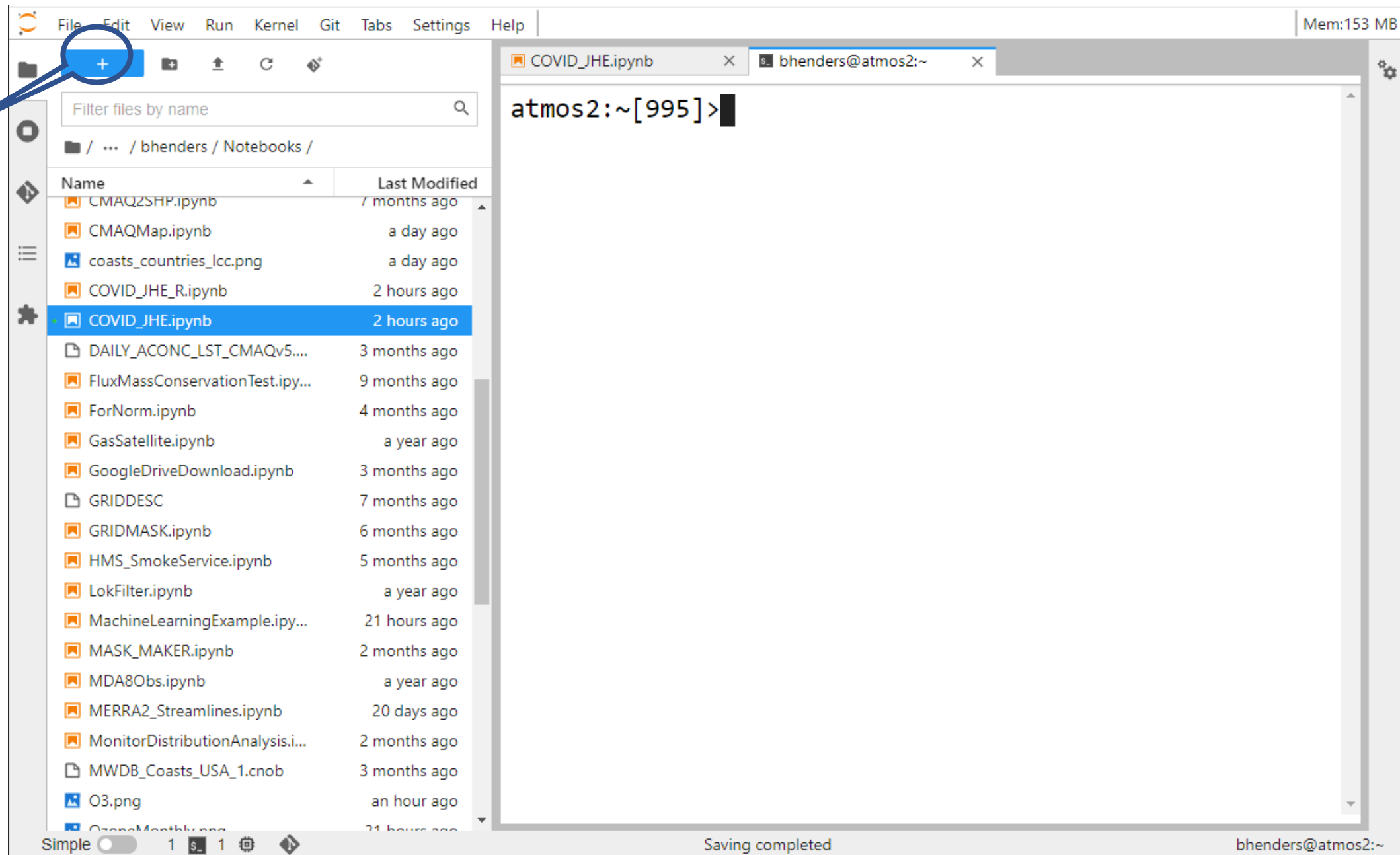
Data Citation: Ensheng Dong, Hongru Du, Lauren Gardner, An interactive web-based dashboard to track COVID-19 in real time, The Lancet Infectious Diseases, Volume 20, Issue 5, 2020, Pages 533-534, ISSN 1473-3099, doi: 10.1016/S1473-3099(20)30120-1.

### Installing a custom library

- JupyterHub on atmos comes with many libraries, but it can be useful to install your own.
- This tutorial will use NASA Computer Network Overlays (CNO) to make maps prettier.
- The python interface to NASA CNO is not installed by default.
- In this live demo, we are going to install pycno in the user space.

Simple 0 1 anaconda | Idle Saving completed Mode: Command Ln 1, Col 1 COVID\_JHE.ipynb

Make new  
Notebooks  
or Terminals



The screenshot displays the JupyterLab environment. On the left, the file browser shows a list of files and notebooks in the directory `/bhenders/Notebooks/`. The file `COVID_JHE.ipynb` is selected. A blue circle highlights the '+' button in the file browser toolbar, which is used to create new notebooks or terminals. The main area on the right shows a terminal window with the prompt `atmos2:~[995]>`. The status bar at the bottom indicates 'Saving completed' and the user `bhenders@atmos2:~`.

Name	Last Modified
CMAQ2SHP.ipynb	7 months ago
CMAQMap.ipynb	a day ago
coasts_countries_lcc.png	a day ago
COVID_JHE_R.ipynb	2 hours ago
COVID_JHE.ipynb	2 hours ago
DAILY_ACONC_LST_CMAQv5....	3 months ago
FluxMassConservationTest.ipynb	9 months ago
ForNorm.ipynb	4 months ago
GasSatellite.ipynb	a year ago
GoogleDriveDownload.ipynb	3 months ago
GRIDDESC	7 months ago
GRIDMASK.ipynb	6 months ago
HMS_SmokeService.ipynb	5 months ago
LokFilter.ipynb	a year ago
MachineLearningExample.ipynb	21 hours ago
MASK_MAKER.ipynb	2 months ago
MDA8Obs.ipynb	a year ago
MERRA2_Streamlines.ipynb	20 days ago
MonitorDistributionAnalysis.i...	2 months ago
MWDB_Coasts_USA_1.cnob	3 months ago
O3.png	an hour ago
QzssMonth.png	21 hours ago

# How do Notebooks Work?

- Users access notebooks by a “web browser” (e.g., Chrome).
- The browser accesses a Notebook “server.”
  - A server can be on your machine.
  - A server can be on atmos
  - A server can be on the cloud
- The Notebook server
  - Stores data in a Notebook file
  - Passes commands to the Kernel
- The Kernel processes commands
  - Kernels can operate in many languages
  - Python, R, CSH etc

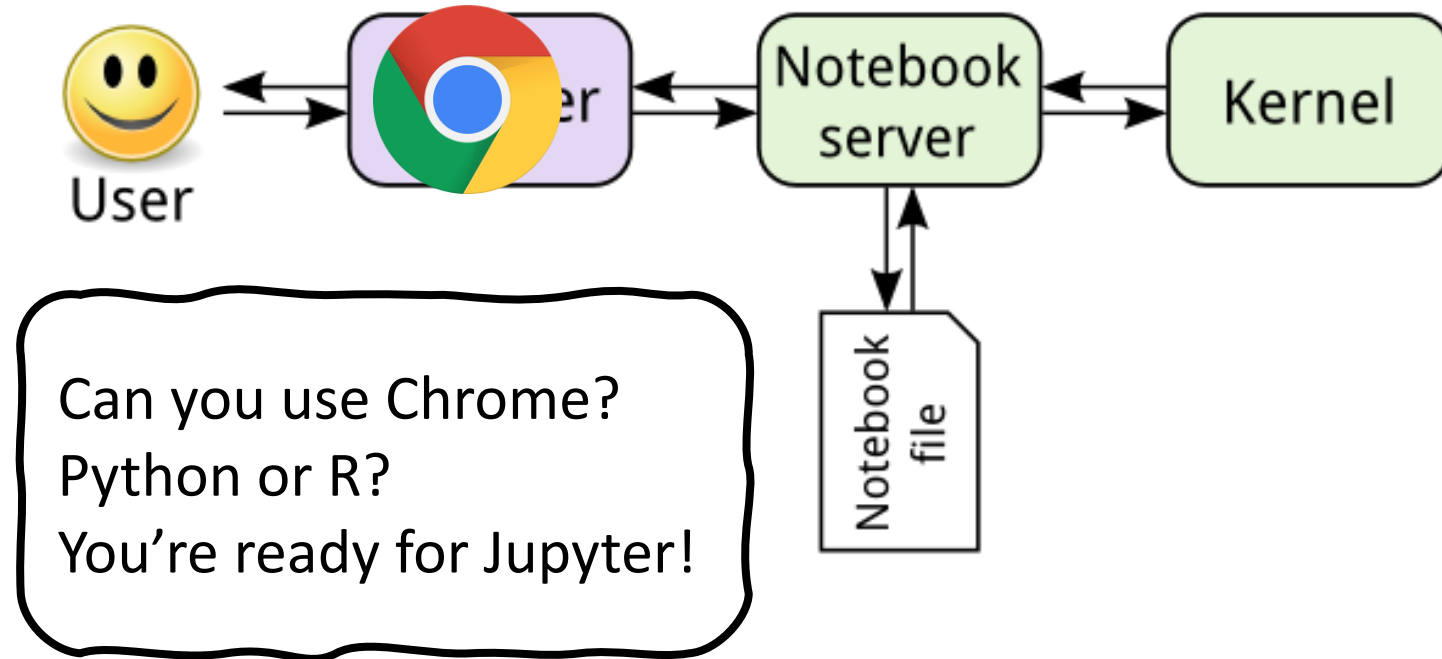


Figure source: <https://coderefinery.github.io/jupyter/interface/>



# On Windows

## Install Once

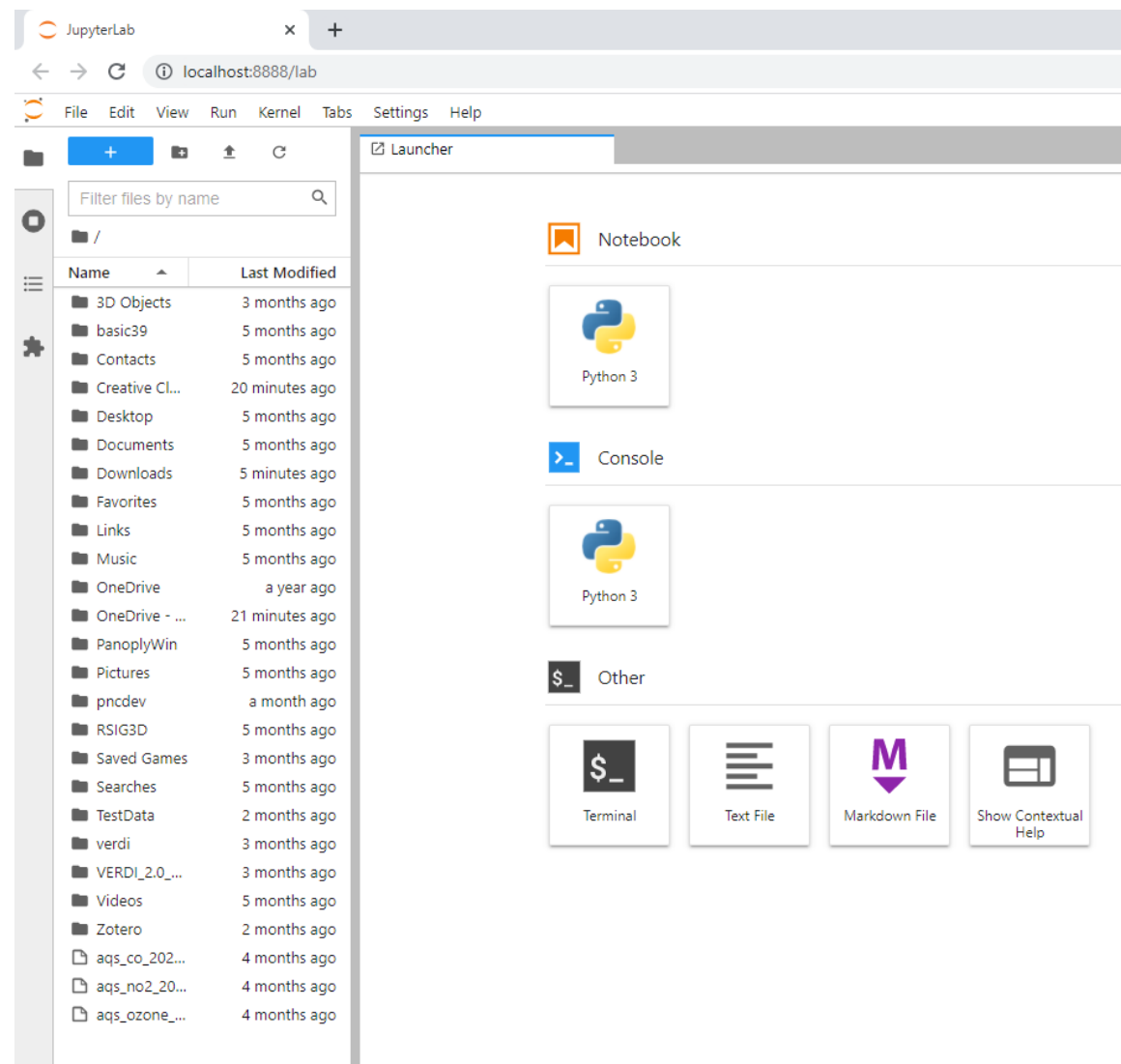
- Installer from python.org installed for user in basic39.
- `.\basic39\Scripts\pip.exe install --user jupyterlab`

***Start a server any time...***

```
C:\Users\BHenders>.\basic39\Scripts\jupyter.exe lab
```

```
[I 2021-11-16 20:39:33.950 ServerApp] jupyterlab | extension was successfully linked.
[W 2021-11-16 20:39:33.982 ServerApp] The 'min_open_files_limit' trait of a ServerApp instance
NoneType None.
[I 2021-11-16 20:39:34.715 ServerApp] nbclassic | extension was successfully loaded.
[I 2021-11-16 20:39:34.732 LabApp] JupyterLab extension loaded from c:\users\bhenders\basic39
ab
[I 2021-11-16 20:39:34.732 LabApp] JupyterLab application directory is C:\Users\BHenders\ba
[I 2021-11-16 20:39:34.737 ServerApp] jupyterlab | extension was successfully loaded.
[I 2021-11-16 20:39:34.737 ServerApp] Serving notebooks from local directory: C:\Users\BHe
[I 2021-11-16 20:39:34.737 ServerApp] Jupyter Server 1.8.0 is running at:
[I 2021-11-16 20:39:34.737 ServerApp] http://localhost:8888/lab?token=3ab2554914c937288f55b
[I 2021-11-16 20:39:34.737 ServerApp] http://127.0.0.1:8888/lab?token=3ab2554914c937288
ac
[I 2021-11-16 20:39:34.737 ServerApp] Use Control-C to stop this server and shut down all
(ation).
[C 2021-11-16 20:39:34.815 ServerApp]

To access the server, open this file in a browser:
file:///C:/Users/BHenders/AppData/Roaming/jupyter/runtime/jpserver-12384-open.html
Or copy and paste one of these URLs:
http://localhost:8888/lab?token=3ab2554914c937288f55be1967ad88ea09117cb30d982aac
http://127.0.0.1:8888/lab?token=3ab2554914c937288f55be1967ad88ea09117cb30d982aac
```



This also works on atmos... we would all need our own “server”, our own “port”, and SSH tunnels... not easy...

# On Mac or Linux

## Install Once

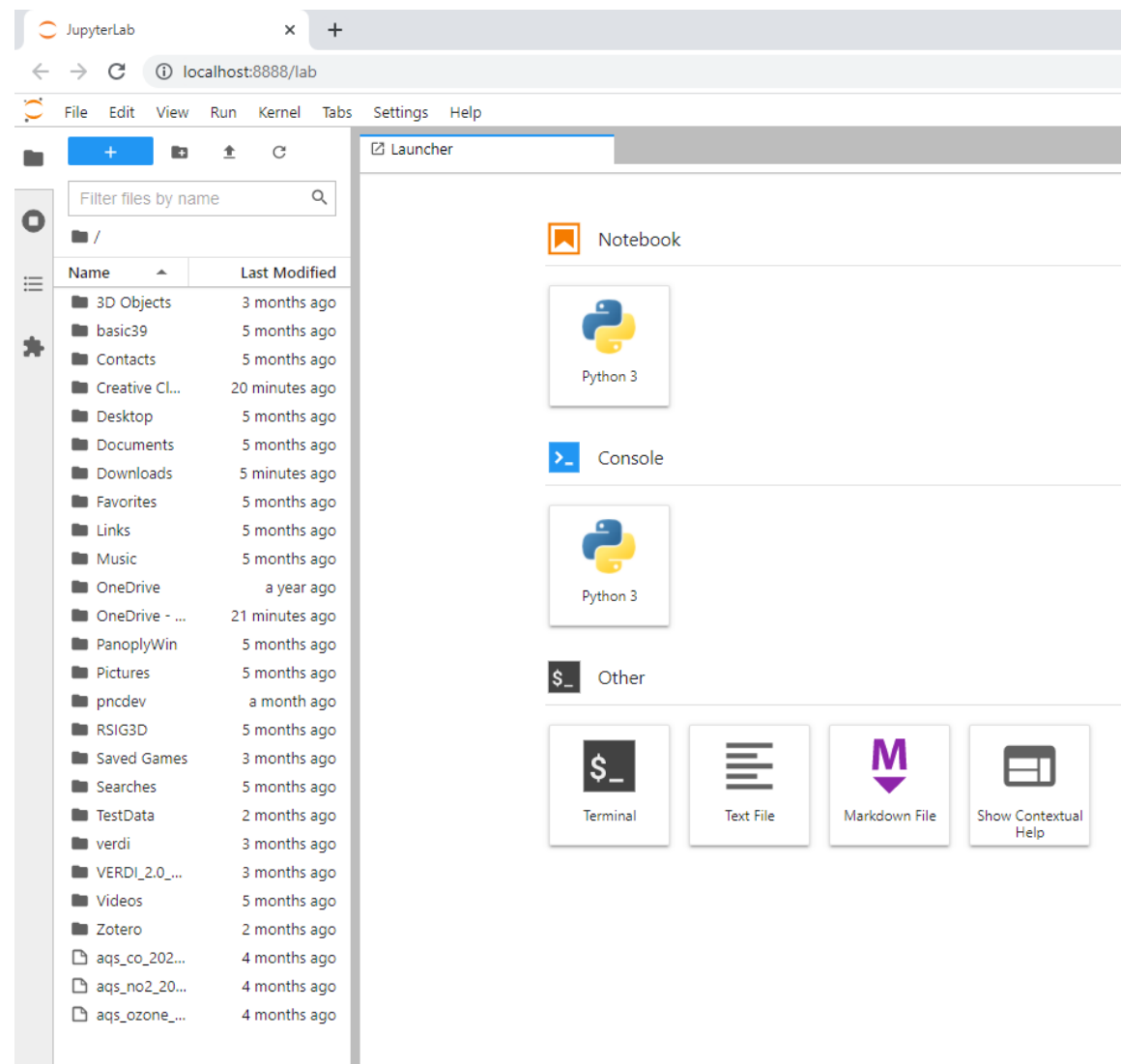
- Installer from python.org installed for user in basic39.
- ``python -m pip install --user jupyterlab``

***Start a server any time...***

```
$ jupyter lab
```

```
[I 2021-11-16 20:39:33.950 ServerApp] jupyterlab | extension was successfully linked.
[W 2021-11-16 20:39:33.982 ServerApp] The 'min_open_files_limit' trait of a ServerApp instance
NoneType None.
[I 2021-11-16 20:39:34.715 ServerApp] nbclassic | extension was successfully loaded.
[I 2021-11-16 20:39:34.732 LabApp] JupyterLab extension loaded from c:\users\bhenders\basic39\
ab
[I 2021-11-16 20:39:34.732 LabApp] JupyterLab application directory is C:\Users\BHenders\ba
[I 2021-11-16 20:39:34.737 ServerApp] jupyterlab | extension was successfully loaded.
[I 2021-11-16 20:39:34.737 ServerApp] Serving notebooks from local directory: C:\Users\BHer
[I 2021-11-16 20:39:34.737 ServerApp] Jupyter Server 1.8.0 is running at:
[I 2021-11-16 20:39:34.737 ServerApp] http://localhost:8888/lab?token=3ab2554914c937288f55b
[I 2021-11-16 20:39:34.737 ServerApp] http://127.0.0.1:8888/lab?token=3ab2554914c937288
ac
[I 2021-11-16 20:39:34.737 ServerApp] Use Control-C to stop this server and shut down all k
(ation).
[C 2021-11-16 20:39:34.815 ServerApp]

To access the server, open this file in a browser:
file:///C:/Users/BHenders/AppData/Roaming/jupyter/runtime/jpserver-12384-open.html
Or copy and paste one of these URLs:
http://localhost:8888/lab?token=3ab2554914c937288f55be1967ad88ea09117cb30d982aac
http://127.0.0.1:8888/lab?token=3ab2554914c937288f55be1967ad88ea09117cb30d982aac
```



This also works on atmos... we would all need our own “server”, our own “port”, and SSH tunnels... not easy...

# What if I can't install?

What if I don't \*want\* to install?

Use  jupyterhub

JupyterHubs give anyone access to cloud computing services.

- The interface is the same in the cloud as on your machine.
- Available all kinds of places
  - Google Colab (integrated GitHub)
  - Amazon Sage Maker
  - Microsoft Azure (discontinued?)
  - mybinder.org (integrated GitHub)
  - Atmos and EPA DMAP...
- When your data is on the same machine... WOW!

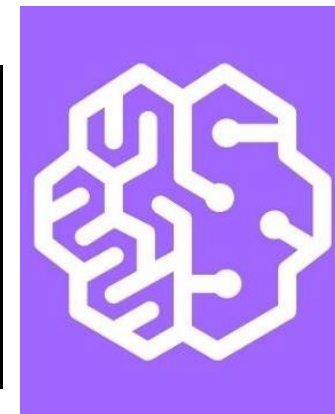
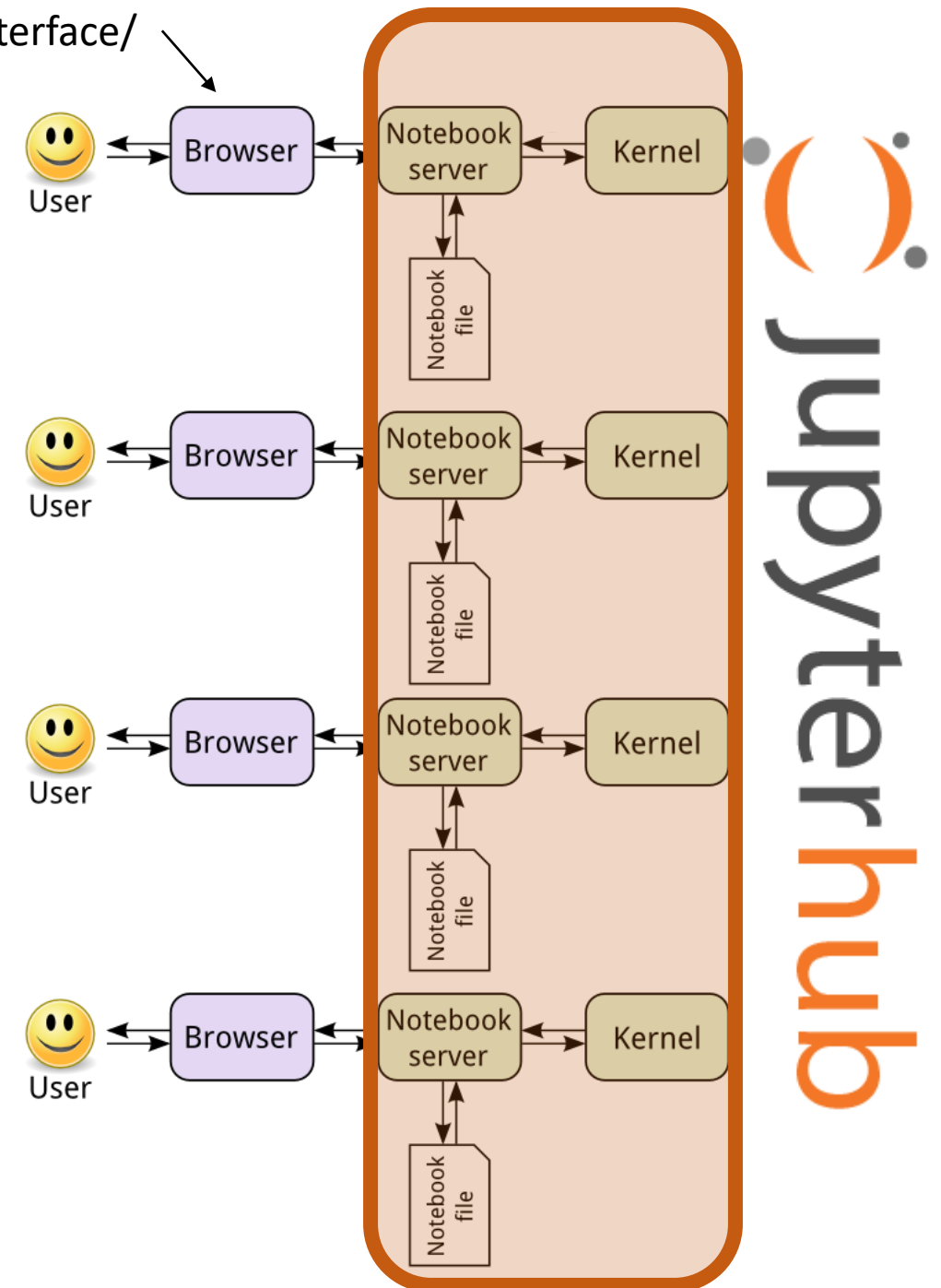


Figure source: <https://coderefinery.github.io/jupyter/interface/>



Figure source: <https://geohackweek.github.io/Introductory/05-Jupyter-tutorial/>



# Cloud Options at EPA

## **Amazon-based Data Management and Analytics Platform (DMAP)**

- Spinup your own Jupyter Server
- Access computational resources and data on AWS
- Up to \$100/month approved
- Requires an account request to OMS
- Currently notebooks not labs, but basically the same.
- May expand to Amazon Sagemaker

## **High Performance Computing (HPC) on atmos**

- JupyterHub is already running on atmos.
- Access login nodes and computational nodes.
- Access data on atmos
- Already paid for.
- Requires an account that many of you already have... ask your branch chief or [emvl\\_help@epa.gov](mailto:emvl_help@epa.gov).

\*Google Colab requires a Google account

\*Mybinder.org requires an account

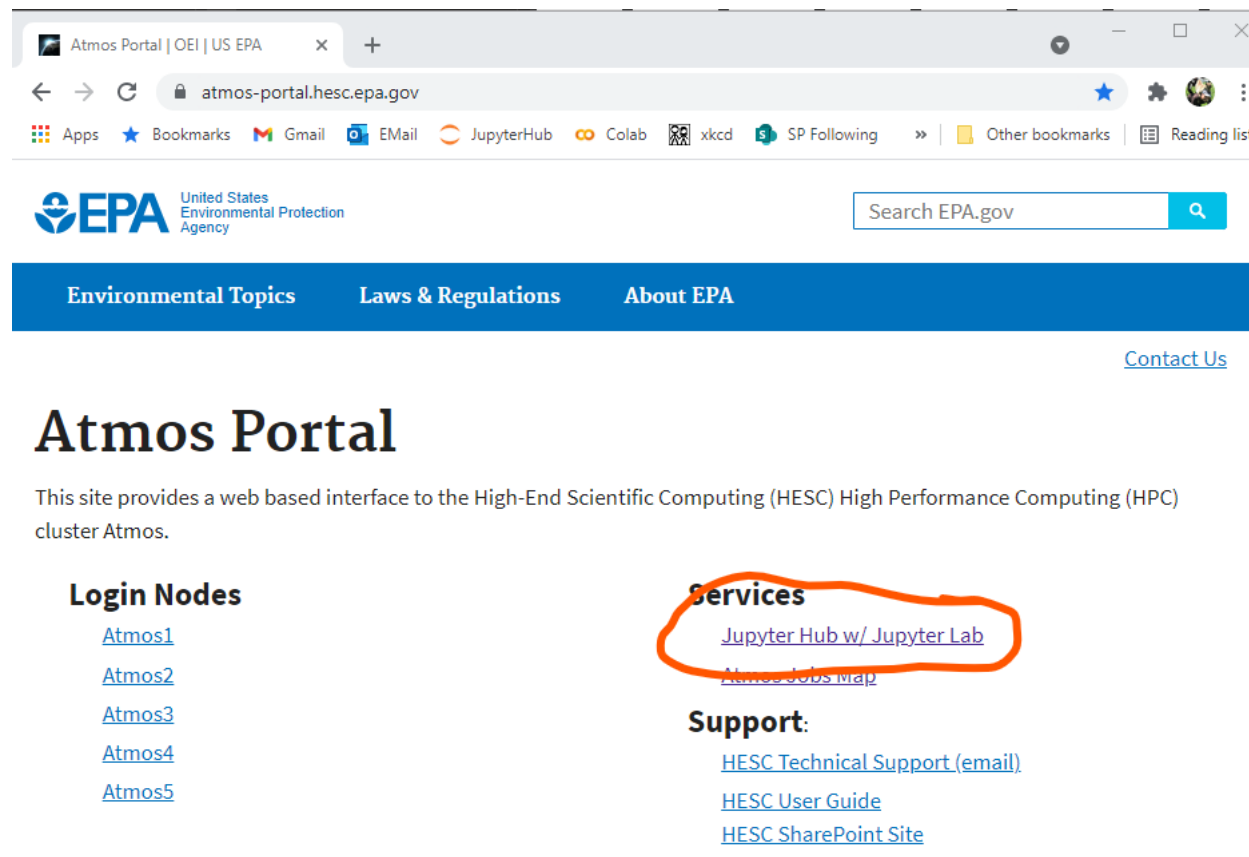
# Atmos Instructions

- The next few examples assume you have access to atmos.
- If you do not, you can use mybinder.org or Google Colab to explore publicly available notebooks.
  - <https://colab.research.google.com/>
  - For the CMAQ crew [http://github.com/barronh/pseudo-netcdf\\_examples](http://github.com/barronh/pseudo-netcdf_examples)
- If you have an atmos account, and you already know how to access JupyterHub, consider exploring
  - /home/bhenderson/Notebooks/
    - a. AQS\_Pregenerated\_MonthMean.ipynb
    - b. BasicCMAQEval.ipynb
    - c. CMAQ\_Ozone\_Evaluation.ipynb
    - d. MachineLearningExample.ipynb
    - e. MASK\_MAKER.ipynb
- The rest of us are going to access atmos Jupyter Hub for the first time.



# Access Jupyter Hub from the atmos-portal

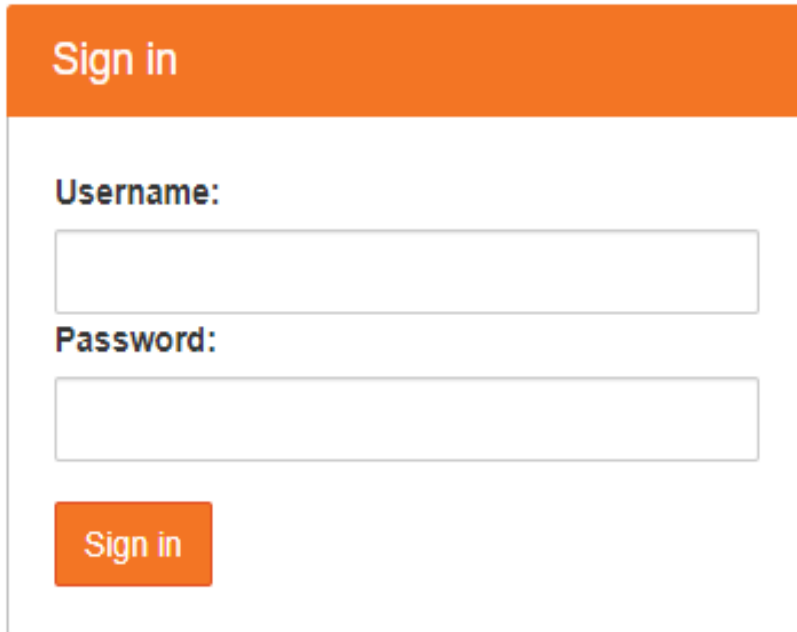
1. Connect to the EPA Network or VPN
2. Navigate a web browser to <https://atmos-portal.hesc.epa.gov/>
3. Click on “Jupyter Hub w/ Jupyter Lab”





# Logging in and starting a “user server”

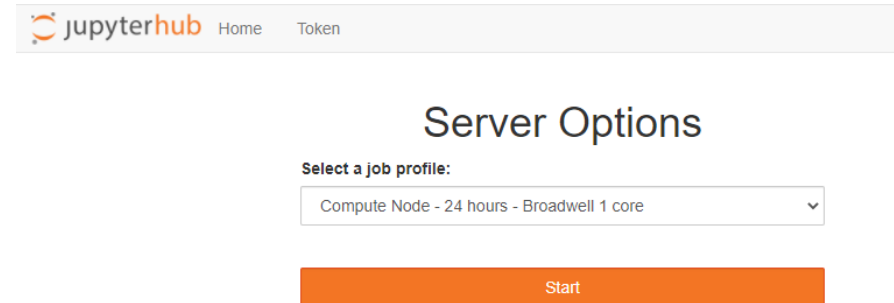
4. Log in with your Lan ID and password (not necessary every time)



A sign-in form with an orange header bar containing the text "Sign in". Below the header, there are two input fields: "Username:" and "Password:". The "Username:" field is a simple text box, while the "Password:" field is a text box with a small eye icon on the right side. Below the password field is an orange button with the text "Sign in".

5. Normally, select a “Compute Node – 24 hours - Broadwell 1 core” and click start.

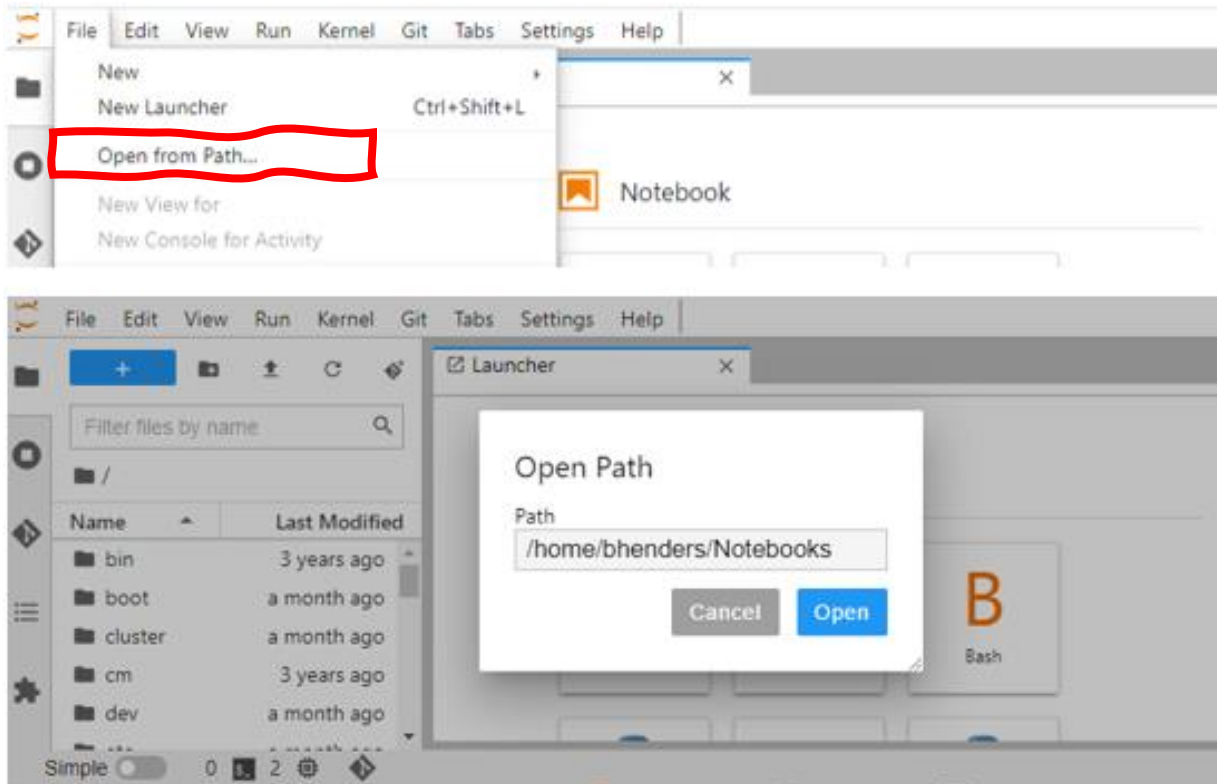
- Select a singlepe node
- Select a debug node for 4 hours
- Select a compute node for 1 days.



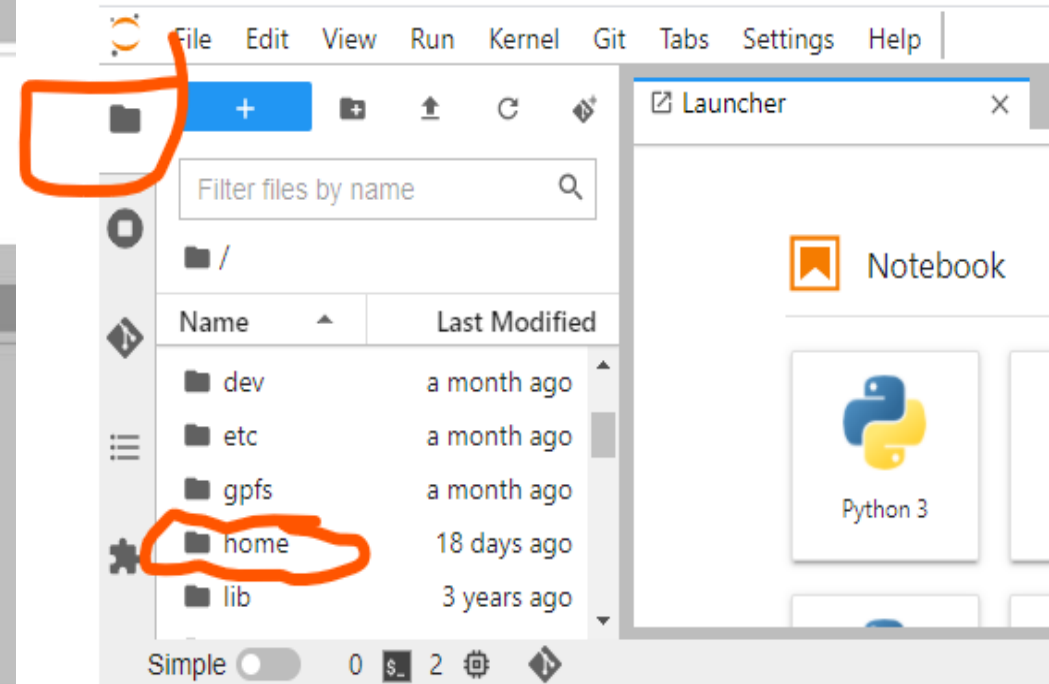
The JupyterHub interface showing the "Server Options" page. At the top, there is a navigation bar with the JupyterHub logo, "Home", and "Token". Below the navigation bar, the title "Server Options" is displayed. Underneath the title, there is a section labeled "Select a job profile:" followed by a dropdown menu. The dropdown menu is open, showing the selected option "Compute Node - 24 hours - Broadwell 1 core". Below the dropdown menu is an orange button with the text "Start".

## 6. Navigate to a folder with a Notebook

a) Select “File” and “Open from Path...”



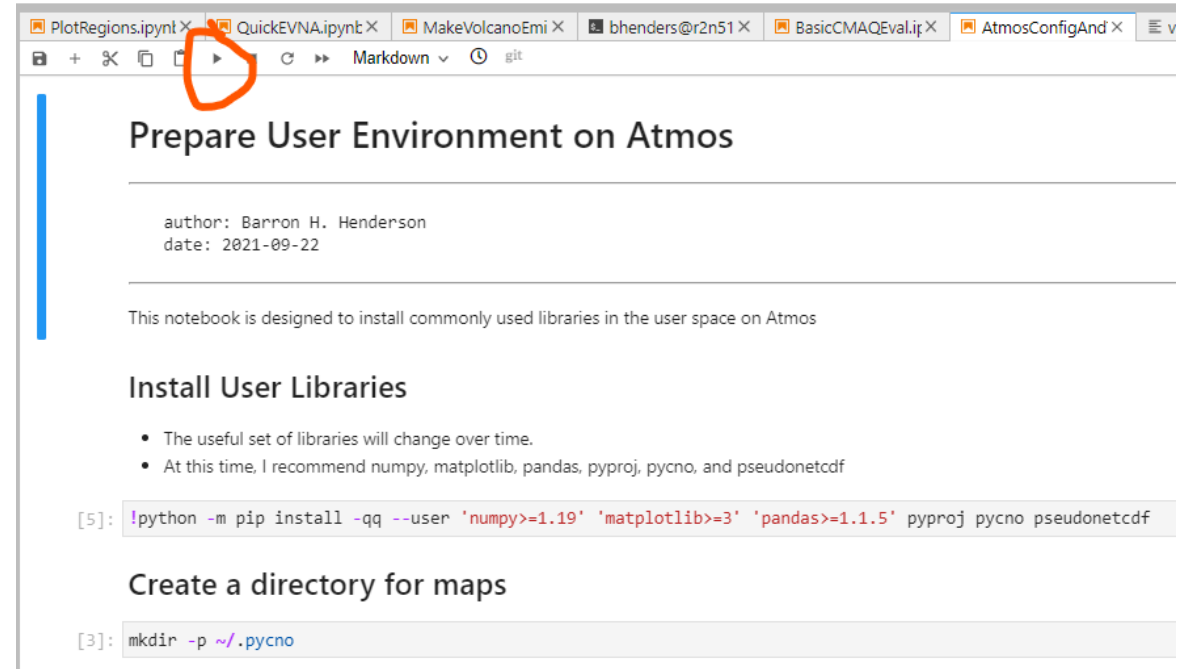
b) Or navigate by clicking in the file browser (folder on left)



# 7. Example notebook and system configuration

7. For your first example, navigate to /home/bhenders/Notebooks, then double click on 0AtmosConfigAndTest.ipynb notebook. This will open the Notebook.

- This is your first time, and this notebook helps to update or install a few libraries
- Optionally, Use File “Save Notebook As” and save it in your own user space (/home/<username>). The notebook that is open is the newly saved notebook.
- Click the play button once for each cell (e.g., see [1] in next figure).
- The primary purpose of the notebook is to install libraries that are known to work



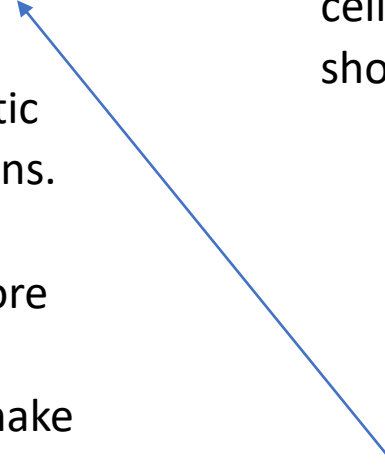
# Explore and make other notebooks

8. You can open any of the notebooks in /home/bhenders/Notebooks and follow a similar process to steps 6 and 7. The three notebooks below are intended to be updated as needed so that they can be used as tutorials.

- a. AQS\_Pregenerated\_MonthMean.ipynb a simple observation analysis system***
- b. BasicCMAQEval.ipynb performs a simplistic CMAQ evaluation against AQS observations.
- c. MachineLearningExample.ipynb
- d. CMAQ\_Ozone\_Evaluation.ipynb has a more detailed evaluation of a year.
- e. MASK\_MAKER.ipynb is a simple tool to make arbitrary masks from shapefiles

9. Lastly, any time you make a new notebook:

- a. Choose the anaconda kernel. This will ensure you have access to important scientific libraries.
- b. Always add “%matplotlib inline” in the first cell. On Atmos, this ensures plots will be shown



We're going to do this one together.

# Live Examples

- A repository of fully functional runnable code.
- Oriented toward solving CMAQ types of problems.
- Right now, the CMAS Forum Python Channel allows us to link to these sorts of things
  - [cmascenter.org](https://forum.cmascenter.org/c/python/) -> Help -> CMAS Forum -> Category Python
  - <https://forum.cmascenter.org/c/python/>

