Authors: Andy Bean & Carrie Harris
DIFUSE Project
Dartmouth College
May 2023

How to Move a Colab to Jupyter

Summary

Install Anaconda Navigator on your Computer

Create a virtual environment with the correct library versions for the Colab file

Download the Colab as a .ipynb, launch Jupyter, and run the code in browser on Jupyter

What is this guide for?

Colab files are .ipynb files, which work with any python notebook environment. This means we can download the colab file and run it elsewhere, such as in Jupyter notebooks. If Colab updates its packages and breaks your Colab file, you might want this guide!

Instead of using Google's hosted environment that might no longer work, we can create a local virtual environment with exactly the right packages and connect to that instead. Using Jupyter also allows you to interact with files directly on your computer (so make sure you trust the code you are running since it can touch your personal files!).

Step 1: Anaconda Navigator

Anaconda Navigator is a package manager for several coding environments. Most relevant to us, it allows us to easily create a "virtual environment" that has exactly the package versions we want for our code.

Download and install Anaconda Navigator at: https://www.anaconda.com/download

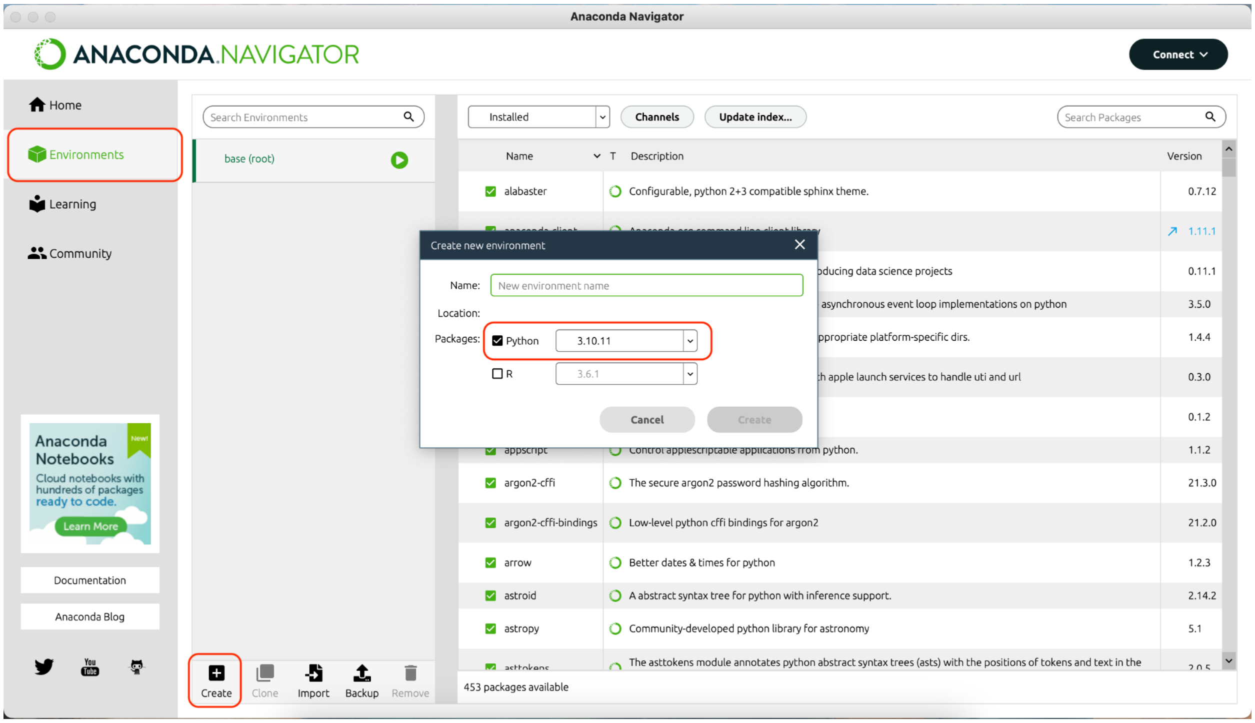
Step 2: The Virtual Environment

Create the Environment with the Right Language Version

Once Anaconda is installed, we can create the proper environment to run the Colab on. Go to “Environments” on the left, then “Create” at the bottom, and fill out the pop-up that appears with your proper version of R or Python

For the Astro-Imaging Colab, we want Python version 3.10.11

You can name the environment whatever you want, but a name with no spaces and no special characters is easier to use. E.g. “astroImagingEnv”



Create new environment

Name: New environment name

Location:

Packages: ☑ Python 3.10.11

□ R 3.6.1

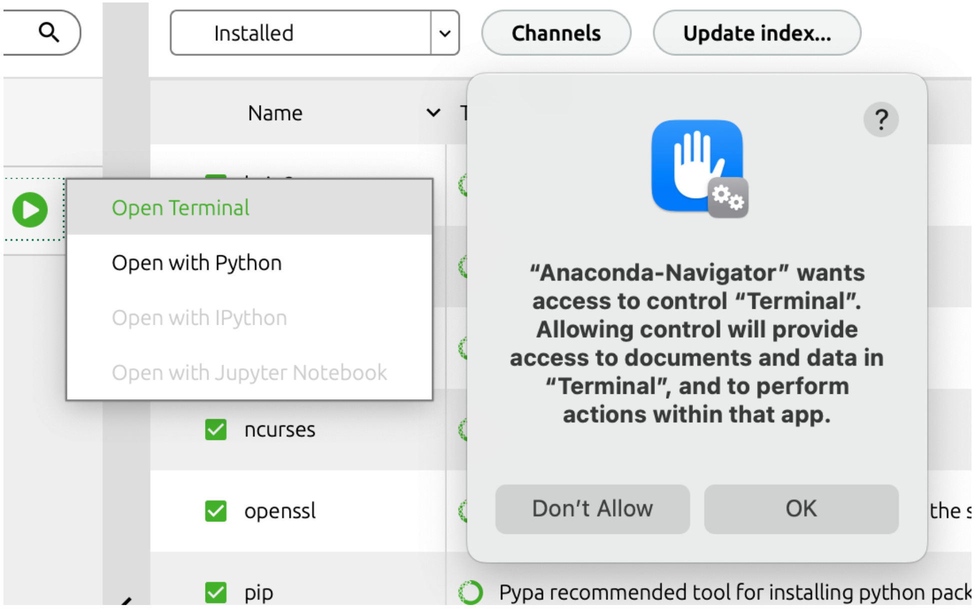
Cancel Create

Name	T	Description	Version
☑ alabaster	○	Configurable, python 2+3 compatible sphinx theme.	0.7.12
		…oducing data science projects	0.11.1
		…asynchronous event loop implementations on python	3.5.0
		…ppropriate platform-specific dirs.	1.4.4
		…h apple launch services to handle uti and url	0.3.0
			0.1.2
☑ appscript	○	Control applescriptable applications from python.	1.1.2
☑ argon2-cffi	○	The secure argon2 password hashing algorithm.	21.3.0
☑ argon2-cffi-bindings	○	Low-level python cffi bindings for argon2	21.2.0
☑ arrow	○	Better dates & times for python	1.2.3
☑ astroid	○	A abstract syntax tree for python with inference support.	2.14.2
☑ astropy	○	Community-developed python library for astronomy	5.1
☑ asttokens	○	The asttokens module annotates python abstract syntax trees (asts) with the positions of tokens and text in the	2.0.5

453 packages available

Put Packages on the Environment

Pressing “play” next to your new environment will open up a menu asking you how you want to run the environment. Select “Open Terminal” and allow Anaconda to access the computer terminal



"Anaconda-Navigator" wants access to control "Terminal". Allowing control will provide access to documents and data in "Terminal", and to perform actions within that app.

In the terminal window that opens, we are going to install the packages we need. The command format for this is generally:

```
conda install package
```

The list of libraries we need to install is below:

```
pandas v1.5.3
matplotlib v3.7.1
ipywidgets v7.7.2
scikit-learn v1.2.2
statsmodels v0.13.5
seaborn v0.12.2
appnope v0.1.2
matplotlib-inline v0.1.6
```

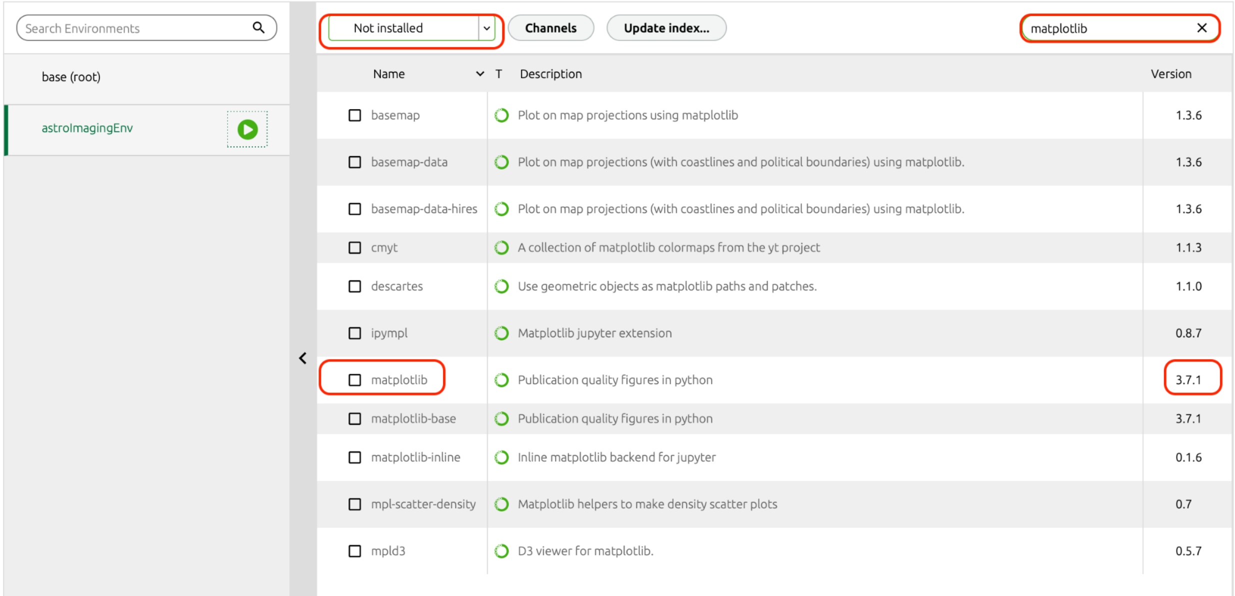
A few other libraries are necessary to run these libraries (e.g. numpy, scipy), and those will be automatically installed by installing these main ones. Those other necessary libraries are called "dependencies."

So type the following commands into your terminal and press enter to execute each command, one at a time:

```
conda install pandas
conda install matplotlib
conda install -c esri ipywidgets
conda install scikit-learn
conda install statsmodels
conda install seaborn
conda install appnope
conda install matplotlib-inline
```

After running each command, the console may say that several other packages are also needed. If the console asks you to confirm with [y]/ n, type “y” and hit enter. This is just to confirm that you agree to the packages being installed. Let the console take its time to install your packages!

If this method is a little daunting, you can also install packages through the Anaconda UI. Change your packages list to show “not installed” and then use the search bar to find a package. If that package is the right version, click the check box next to it and then press apply in the bottom right to install it. HOWEVER, we need an older version of ipywidgets, so that one has to be installed with the console command, or by adding the channel “esri” to the channels options to find and select ipywidgets version 7.7.2 instead of version 8.



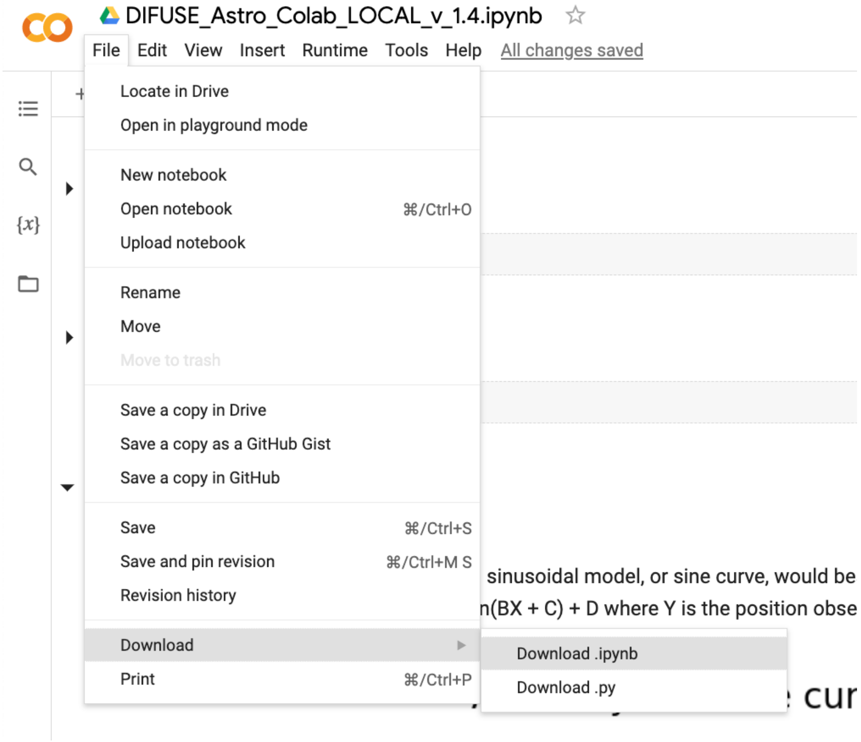
Name	T	Description	Version
□ basemap	○	Plot on map projections using matplotlib	1.3.6
□ basemap-data	○	Plot on map projections (with coastlines and political boundaries) using matplotlib.	1.3.6
□ basemap-data-hires	○	Plot on map projections (with coastlines and political boundaries) using matplotlib.	1.3.6
□ cmyt	○	A collection of matplotlib colormaps from the yt project	1.1.3
□ descartes	○	Use geometric objects as matplotlib paths and patches.	1.1.0
□ ipympl	○	Matplotlib jupyter extension	0.8.7
□ matplotlib	○	Publication quality figures in python	3.7.1
□ matplotlib-base	○	Publication quality figures in python	3.7.1
□ matplotlib-inline	○	Inline matplotlib backend for jupyter	0.1.6
□ mpl-scatter-density	○	Matplotlib helpers to make density scatter plots	0.7
□ mpld3	○	D3 viewer for matplotlib.	0.5.7

Step 3: Open the Colab in Jupyter

This next step requires “jupyter,” but that was automatically installed with Anaconda, so we don’t need to do anything! (If jupyter is giving you trouble, you may need to uninstall and reinstall it from the home page of Anaconda navigator)

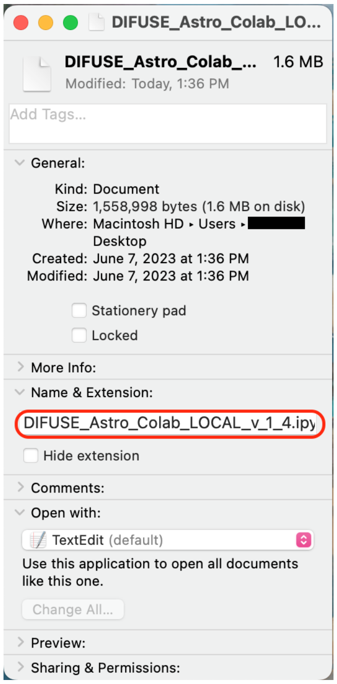
Download the .ipynb file

First, we need to fetch the file we want to run. In the Colab file, go to “File”, then “Download,” then select “Download .ipynb” and move the downloaded file to wherever you would like to run it from.



sinusoidal model, or sine curve, would be
n(BX + C) + D where Y is the position obse

Sometimes, .ipynb files get downloaded and the computer interprets them as .txt files. If this happens, they won’t run in Jupyter. You can fix this to going to the file and deleting the “.txt” extension. If that doesn’t work, try right clicking on the file and going to “Get Info” and then editing the extension in the pop-up that appears.



DIFUSE_Astro_Colab_LO...

DIFUSE_Astro_Colab_... 1.6 MB

Modified: Today, 1:36 PM

Add Tags...

General:

Kind: Document
Size: 1,558,998 bytes (1.6 MB on disk)
Where: Macintosh HD ▸ Users ▸ [redacted]
Desktop
Created: June 7, 2023 at 1:36 PM
Modified: June 7, 2023 at 1:36 PM

□ Stationery pad
□ Locked

More Info:

Name & Extension:

DIFUSE_Astro_Colab_LOCAL_v_1_4.ipy

□ Hide extension

Comments:

Open with:

TextEdit (default)

Use this application to open all documents like this one.

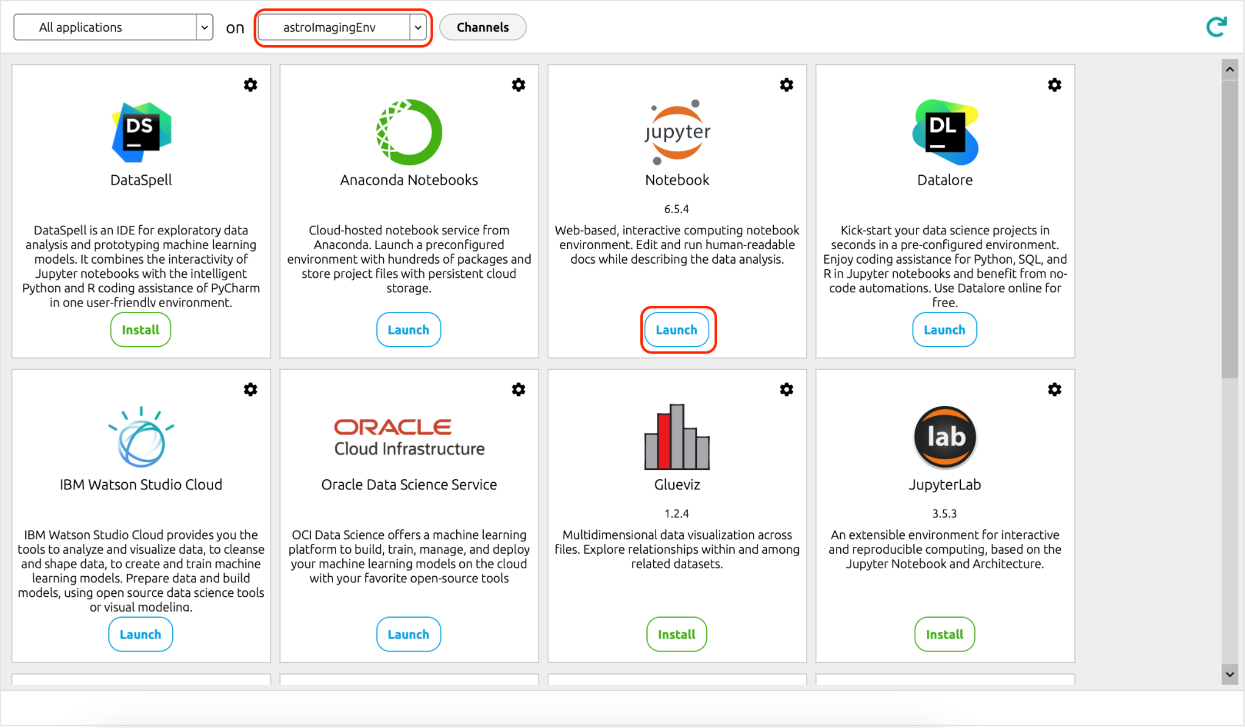
Change All...

Preview:

Sharing & Permissions:

Launch Jupyter

Now we can return to the home page of Anaconda and launch Jupyter. Simply confirm that you are in the right environment in the dropdown at the top (should be the new environment you created in step 2), and then pressing "Launch."



All applications on astroImagingEnv Channels

DataSpell

DataSpell is an IDE for exploratory data analysis and prototyping machine learning models. It combines the interactivity of Jupyter notebooks with the intelligent Python and R coding assistance of PyCharm in one user-friendly environment.

Install

Anaconda Notebooks

Cloud-hosted notebook service from Anaconda. Launch a preconfigured environment with hundreds of packages and store project files with persistent cloud storage.

Launch

Notebook

6.5.4

Web-based, interactive computing notebook environment. Edit and run human-readable docs while describing the data analysis.

Launch

Datalore

Kick-start your data science projects in seconds in a pre-configured environment. Enjoy coding assistance for Python, SQL, and R in Jupyter notebooks and benefit from no-code automations. Use Datalore online for free.

Launch

IBM Watson Studio Cloud

IBM Watson Studio Cloud provides you the tools to analyze and visualize data, to cleanse and shape data, to create and train machine learning models. Prepare data and build models, using open source data science tools or visual modeling.

Launch

Oracle Data Science Service

OCI Data Science offers a machine learning platform to build, train, manage, and deploy your machine learning models on the cloud with your favorite open-source tools

Launch

Glueviz

1.2.4

Multidimensional data visualization across files. Explore relationships within and among related datasets.

Install

JupyterLab

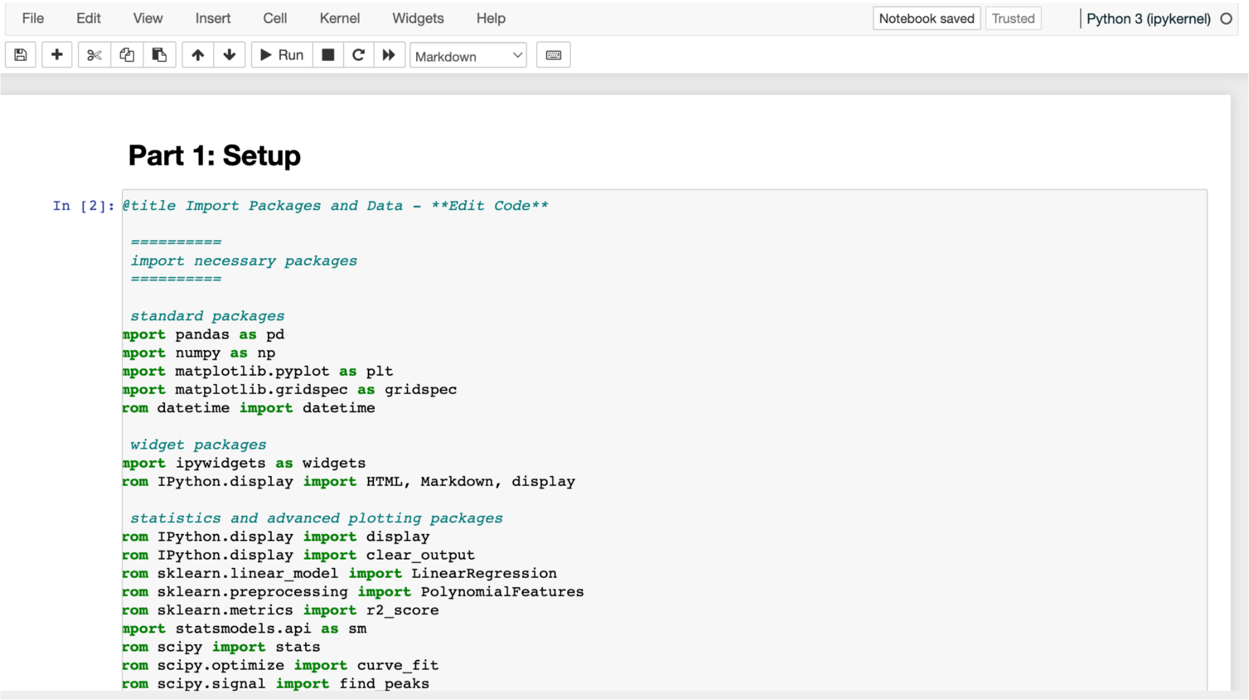
3.5.3

An extensible environment for interactive and reproducible computing, based on the Jupyter Notebook and Architecture.

Install

Jupyter Notebook is automatically installed with Anaconda, so you shouldn't need to install it, but if you do, the "launch" button will say "install" instead. Press install and let Anaconda work.

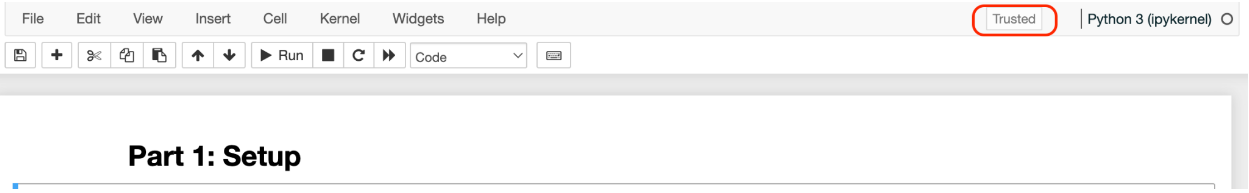
When Jupyter launches, it will create a local host in a browser tab that you have open. Navigate to your .ipynb file in this host tab and click it to open it in the Jupyter Notebooks viewer. The Jupyter viewer looks similar to the Colab viewer, but doesn't hide the code blocks as easily.



Part 1: Setup

```
In [2]: @title Import Packages and Data - **Edit Code**

         ==========
         import necessary packages
         ==========

         standard packages
        mport pandas as pd
        mport numpy as np
        mport matplotlib.pyplot as plt
        mport matplotlib.gridspec as gridspec
        rom datetime import datetime

         widget packages
        mport ipywidgets as widgets
        rom IPython.display import HTML, Markdown, display

         statistics and advanced plotting packages
        rom IPython.display import display
        rom IPython.display import clear_output
        rom sklearn.linear_model import LinearRegression
        rom sklearn.preprocessing import PolynomialFeatures
        rom sklearn.metrics import r2_score
        mport statsmodels.api as sm
        rom scipy import stats
        rom scipy.optimize import curve_fit
        rom scipy.signal import find_peaks
```

Before you can run the code, you will need to "trust" the notebook. If the button in the top right of your screen does not say "Trusted," click it and agree to the popup that occurs to trust the notebook and allow Jupyter to run it.



Part 1: Setup

Now, you can use "Cell" and then "Run all" to run all cells as you would per the "how to Colab" instructions. To run an individual cell, click on its text and you should see a blue bar on the left side highlighting the cell you have selected. Then you can press the run button, or "Cell" → "Run Cells" to run the selected cell(s).

Notes:

A few notes:

- You'll now need to upload your files to the same place that the code itself is running from. So if the .ipynb file is on your desktop, also put the data file on your desktop
- Some packages specific to google colab can no longer be used, so those library calls and related comments will need to be commented out. If you are a module creator, consider making a separate .ipynb file without any of the google colab libraries!
- To exit out, you can quit your process from the terminal or the "home page" that appears when Jupyter is launched
- If you need to quit and connect again, you do not need to reinstall any packages, just skip to step 3 to launch your runtime!
- MOST IMPORTANT for DIFUSE module creators: Colab uses ipywidgets version 7.7.X, but Anaconda automatically provides version 8.0.X, and 8.0 is NOT backwards compatible (same thing goes for ipython). If your module is using ipywidgets and/or ipython, you will need to tell users to install an older version, such as the version in the "esri" channel:
  - Use either the different install command: `conda install -c esri ipywidgets`
  - Or instruct users to add "esri" to their available channels in the package search for Anaconda, and then select version 7.7.2 that way