Setting up Computational Infrastructure for SI 649

1. Installing Anaconda

For python, we recommend using Anaconda, which installs both python and a useful package manager (conda) for installing python packages. If you already have Anaconda installed, you can skip this step, otherwise, download and install it from here: https://www.anaconda.com/products/individual
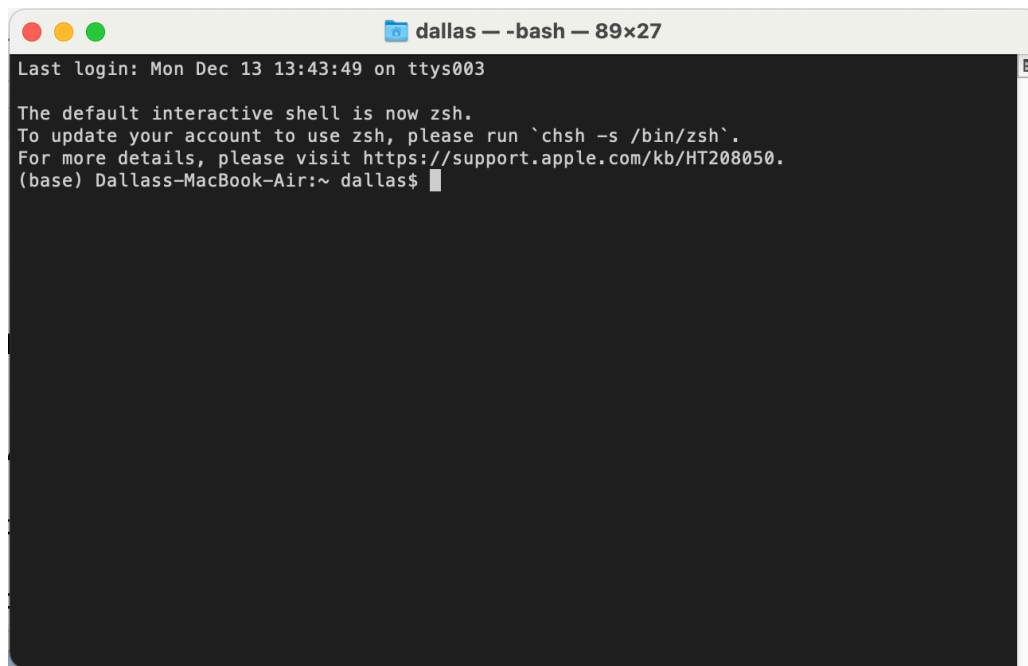
2. Creating a new python environment for this course

2.1. Open the shell:

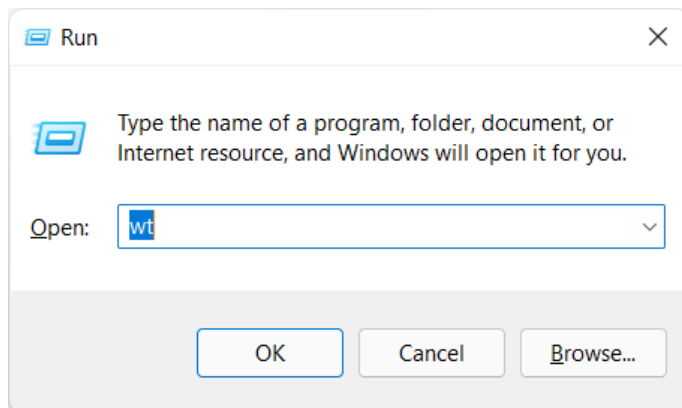**(Option A) On a Mac, search for the Terminal app:**
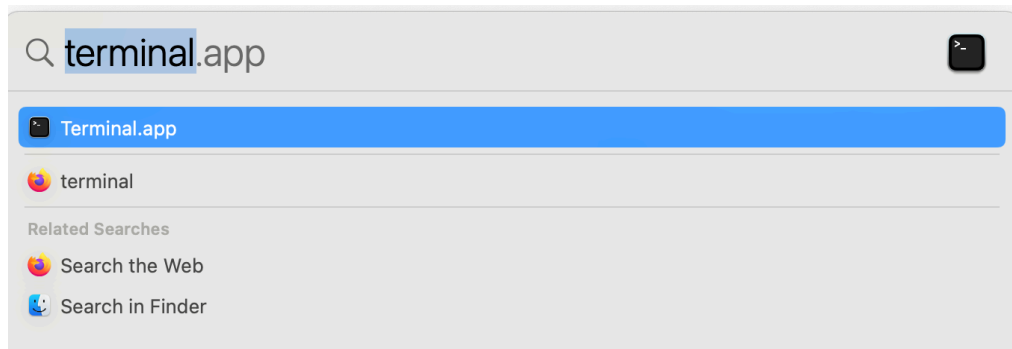
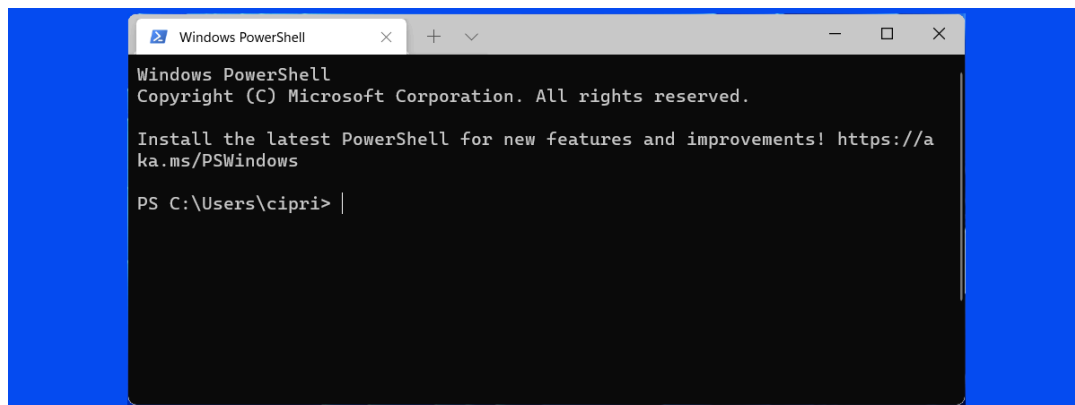This should open a window that looks something like this:

(Option B): On a Windows machine:
Use *Windows key + R* to get the "Run" window.

Type in "wt" and click "OK", which should open a terminal:

(if "wt" doesn't work, try "cmd")

2.2. Verify conda is working (and optionally update it)

In the terminal, run:
`conda --version`
which should print
`conda 4.13.0`
[or some other version]

If you already had an older version of conda installed, you may want to update it using:
`conda update -n base conda`
(then hit 'y' and enter)

2.3. Create a new environment for this class

To create a new environment (specific to the packages we'll use in this class), run:
`conda create -n si649 python=3`

This should display something like the following (hit 'y' and enter):

```
Proceed ([y]/n)? y


Downloading and Extracting Packages
tzdata-2021e       | 112 KB   | ######################################## | 100%
ca-certificates-2021 | 115 KB | ######################################## | 100%
python-3.10.0      | 12.8 MB  | ######################################## | 100%
ncurses-6.3        | 856 KB   | ######################################## | 100%
Preparing transaction: done
Verifying transaction: done
Executing transaction: done
#
# To activate this environment, use
#
#     $ conda activate si649
#
# To deactivate an active environment, use
#
#     $ conda deactivate

(base) Dallass-MacBook-Air:~ dallas$
```

2.4. Activate the environment by running:

`conda activate si649`

which should change the command prompt slightly, as below:

```
(base) Dallass-MacBook-Air:~ dallas$ conda activate si649
(si649) Dallass-MacBook-Air:~ dallas$
```

   ^ *Make sure your environment name is shown here*

2.5. Install some relevant packages.

Run these two commands, one at a time (each of which may take several minutes):

`conda install -c conda-forge altair vega_datasets vl-convert-python`

`conda install -c conda-forge jupyter numpy scipy pandas panel`

(For each one, hit 'y' and enter when prompted)

Both should eventually show something like:

```
Preparing transaction: done
Verifying transaction: done
Executing transaction: - Enabling notebook extension jupyter-js-widgets/extension...
      - Validating: OK

done
(si649) Dallass-MacBook-Air:~ dallas$
```

3. Starting a jupyter notebook (python) session

3.1 Open a terminal

If you have not already done so, open a terminal (as in 2.1 above)

3.2 Activate your class environment

If you have not already done, so activate your environment (as in 2.4 above), by running:
```
conda activate si649
```
or whatever you have called your environment. This will load the packages you have already installed.

## 3.3. Start an interactive notebook session, by running:
```
jupyter notebook
```

This should open up a jupyter notebook interface in your web browser (as shown below, perhaps with slightly different aesthetics due to these being from an older version):
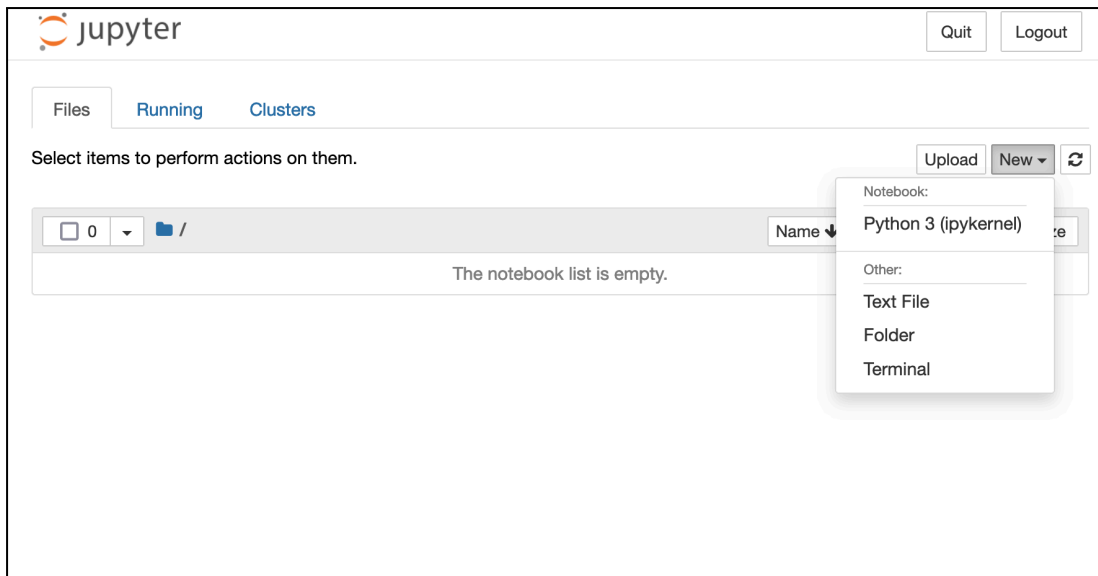
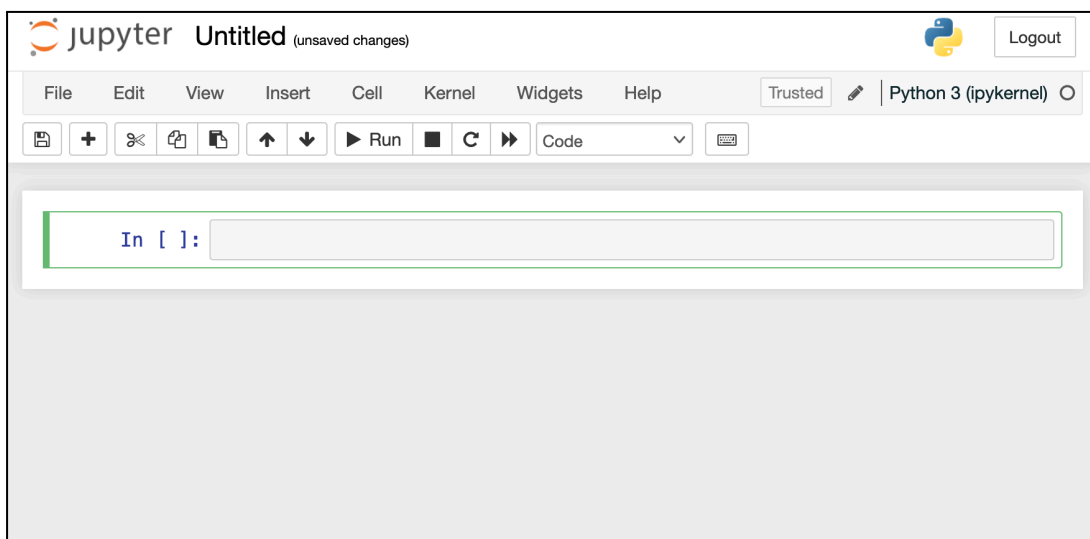*[to shut this down, select the terminal window, and press Ctrl+C]*
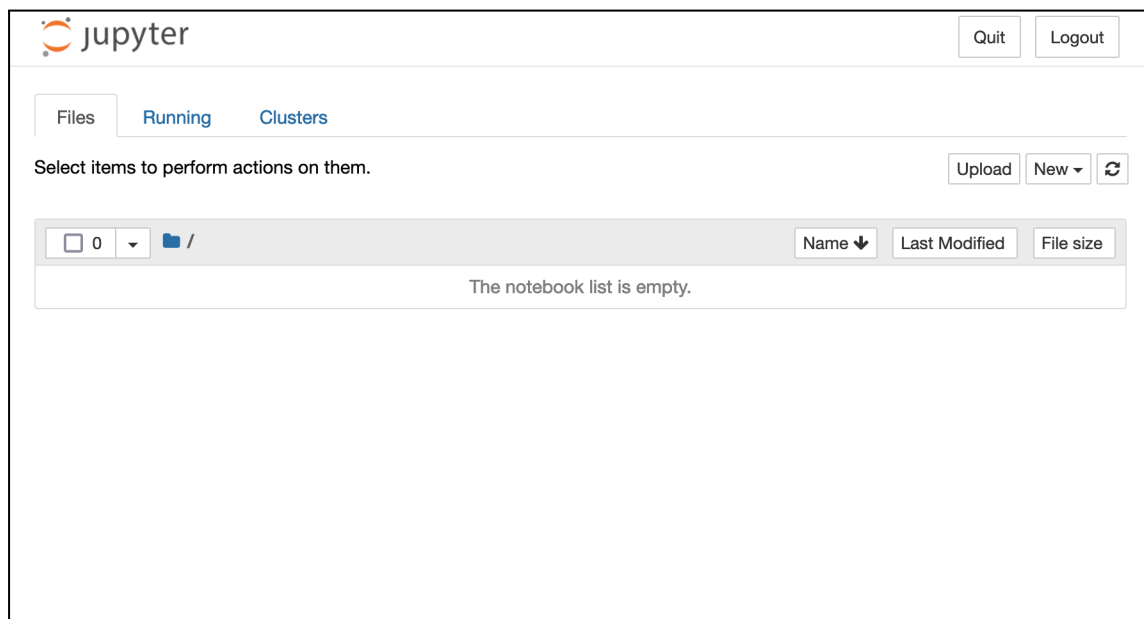


If you are only seeing a blank webpage, try using (CTRL + Shift + R) to refresh the page.

## 3.4 Create a new notebook

To create a new notebook, click the "New" button, then "Python 3"

which should open a new tab that looks something like this:



This is an interactive python session where you can run python code in cells. Changes (and outputs) will be automatically saved, and you can reload your work later by opening up a new jupyter notebook, navigating to the same location, and opening the same file (in jupyter notebook).

Try entering the following code in the notebook cell, and confirm that you are able to import altair, and that it is version 5.4.1:

```
import altair as alt
alt.__version__
```

This should print '5.4.1', or similar. If you have a much older version (because you are using an already existing environment), you may run into problems, and it is recommended you follow the steps above to create a new environment with an up to date version of altair.

If you switch back to the previous tab, you should see the notebook listed:
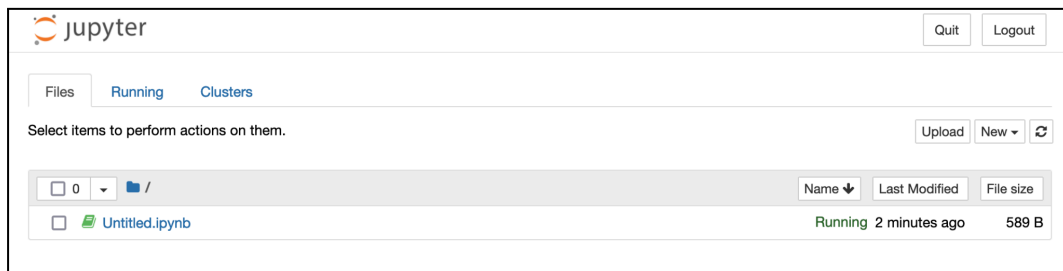
You can rename the notebook, which sits as a file on your computer.

If you close the window, and shut down jupyter, you can later re-open them and return to your work, by following the above steps.

4. Using Google Colaboratory

Another option is to use Google Colab, although you should also make sure you are able to get altair running on your own computer. Google Colab is very similar to jupyter notebooks, but it is especially useful for collaboration or sharing notebooks. They also don't require creating an environment in the same way. However, for most individual work, we strongly recommend using anaconda and jupyter notebooks, to keep everything local.
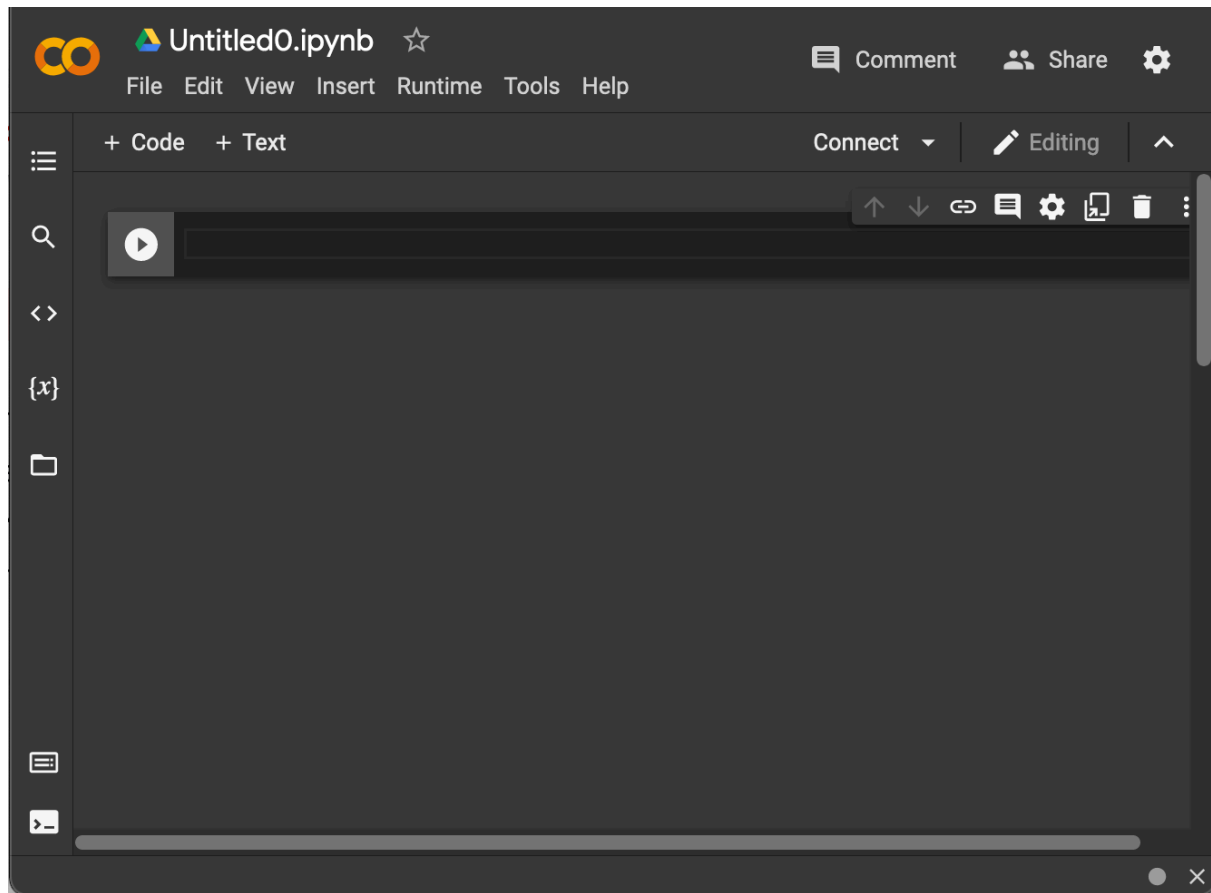
To open a colab notebook, go to:
https://colab.research.google.com/

and create a New Notebook:
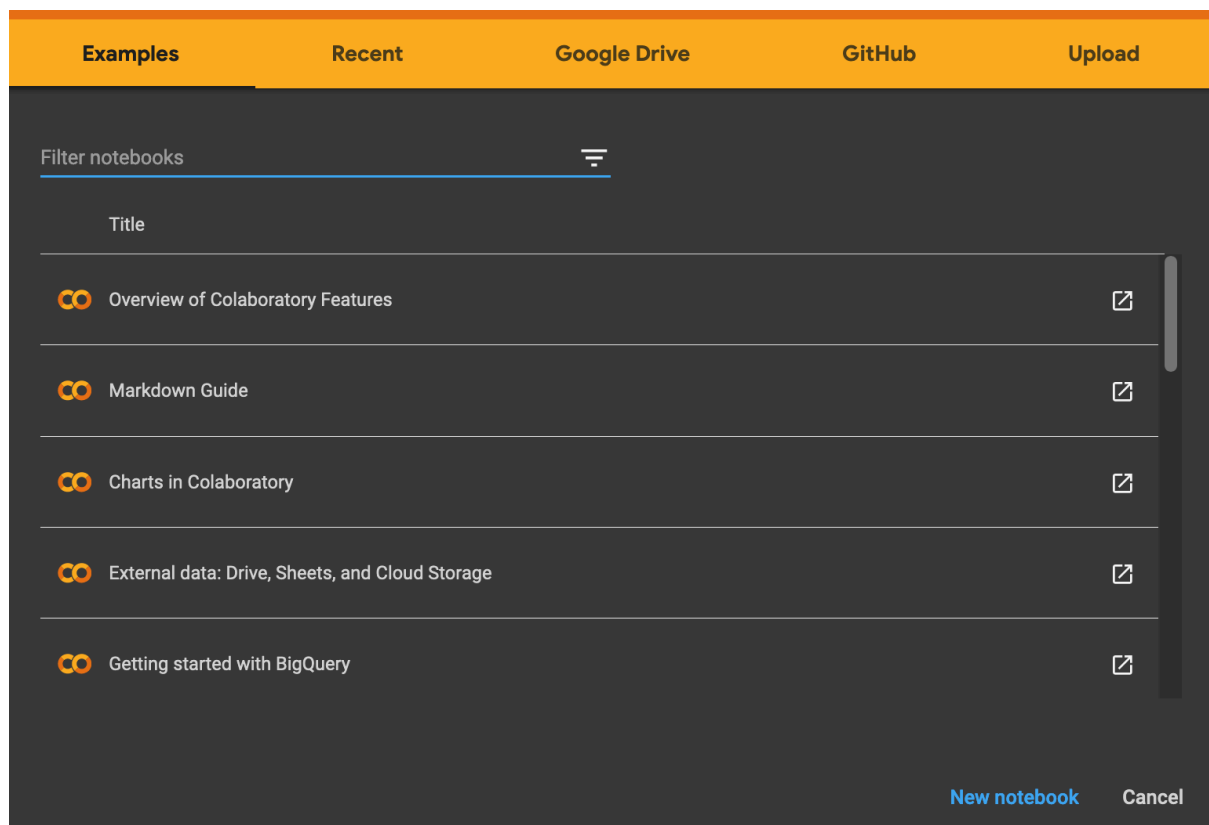
which should give you an interface like this:

You can work with these notebooks similarly to how you use Google docs (e.g., save, copy, share, etc.).