Creating Dashboards with Python

More specifically – plotly & Dash
Fall 2021

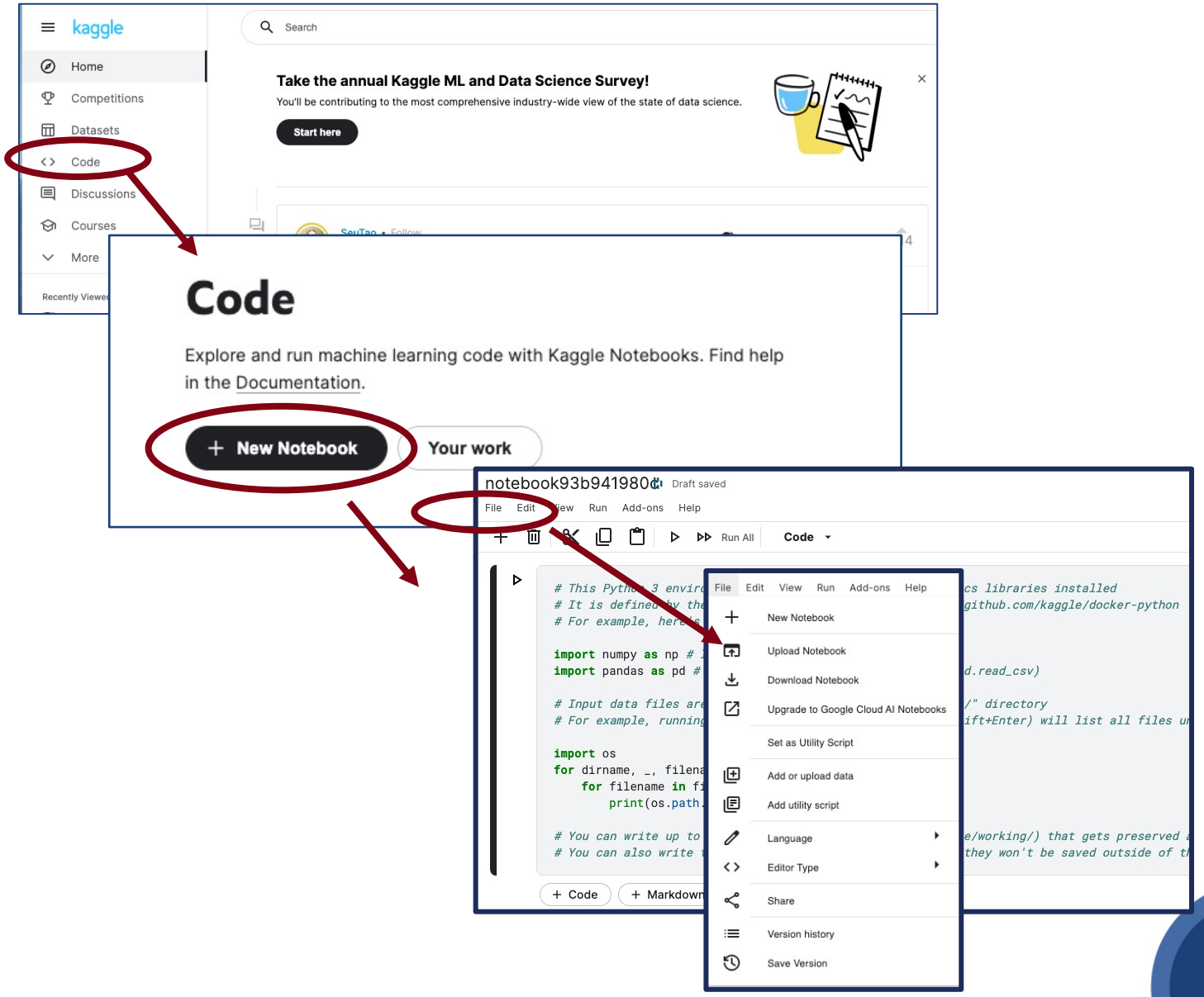
We are going to be using kaggle.com to run our python code.

If you do not already have a kaggle account, please take a few minutes now to create one. Select ‘Register’

- You can use your CDC email
- There are no cost associated with the account.
- I will be giving you code that we will upload to kaggle.

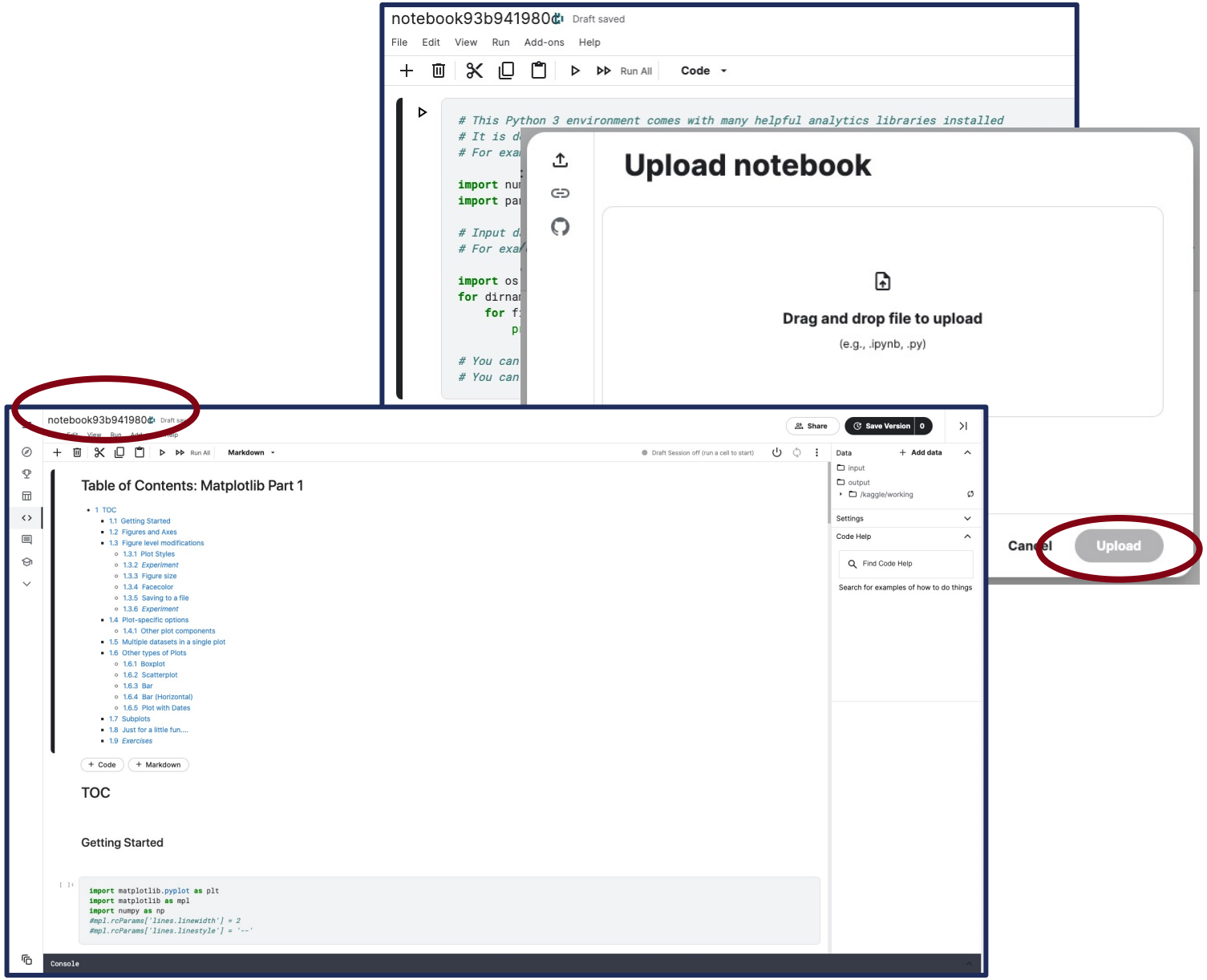
Upload a notebook

- From the Home page select **Code**.
- Select **New Notebook**.
- It will take a few seconds for the new notebook to prepare. Wait for the green text box.
- From the menu bar select File > Upload Notebook.



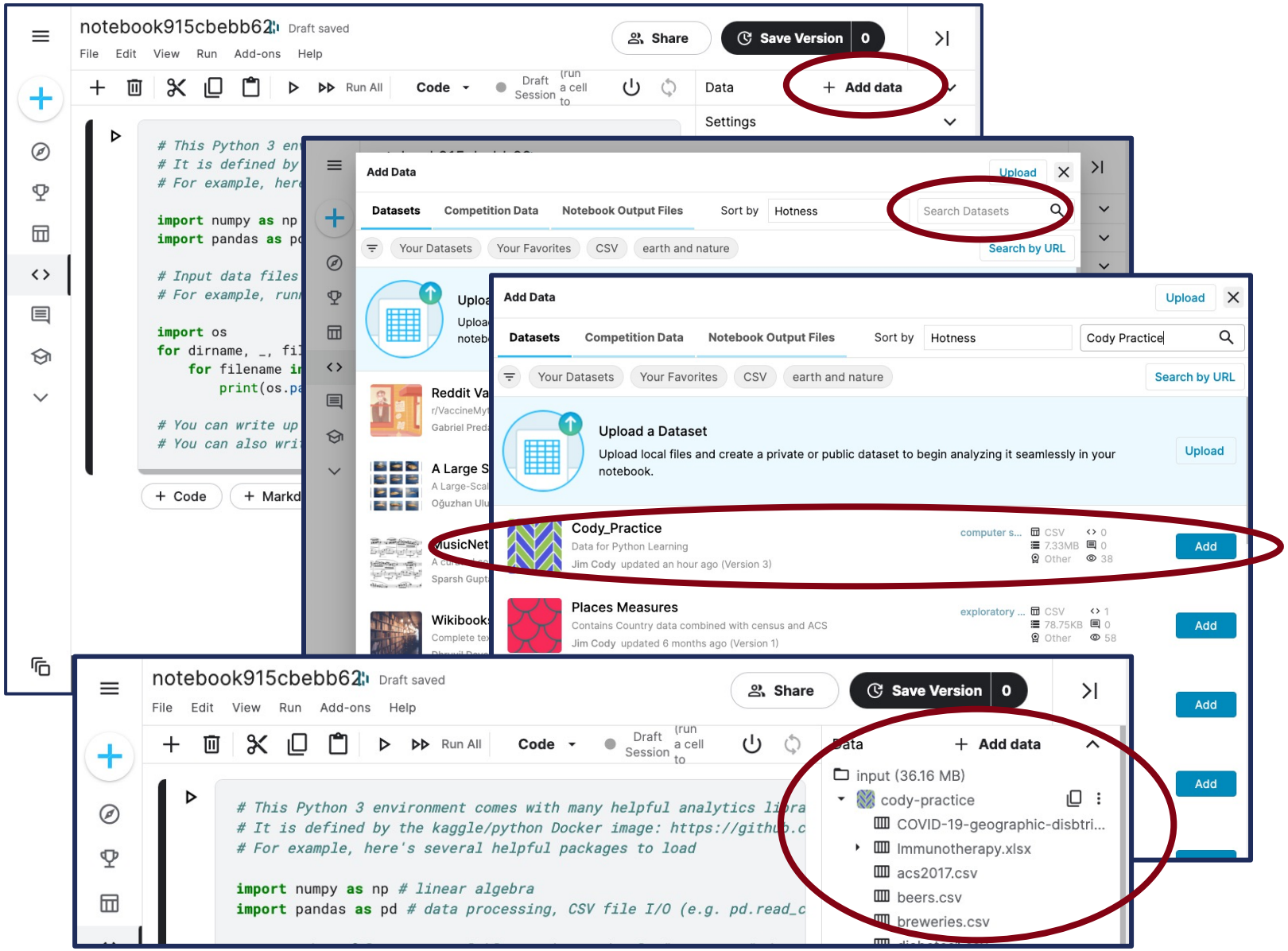
Upload a notebook

- **Browse files** to select the appropriate notebook.
- Press **Upload**.
- Double-click the notebook name to change it.



Add a Dataset to a Kaggle notebook

- Press the **Add** data text in the top right corner of the notebook. A new window will appear.
- In **Search Datasets** text box, enter *Cody Practice*.
- When the list of datasets appears, press the **Add** button next to Cody_Practice.
- The dataset – and all of its associated files - will be associated with the notebook.



Housekeeping

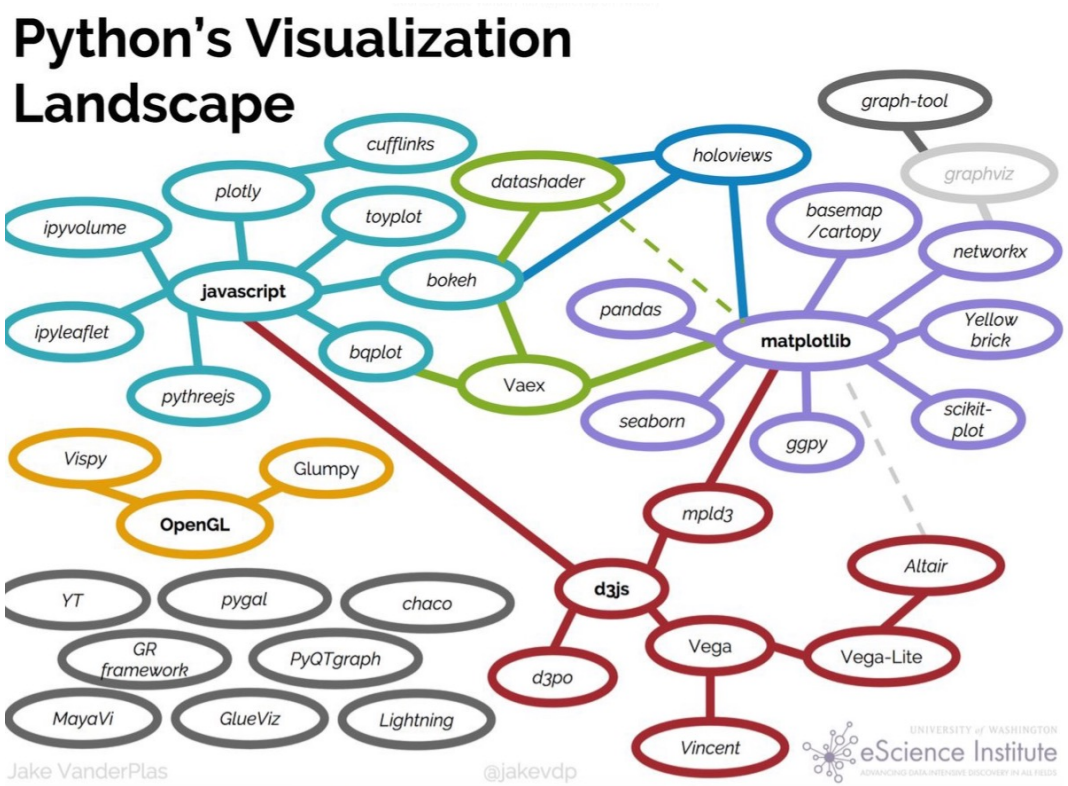
- In case of technical problems:
  - Something wrong on my end (e.g. power outage), I will send you an email.
  - Something wrong on your end, please send me a text message. 508-769-6446
  - jcodygroup@gmail.com
- We have 4 hours for each session
  - I will try to give you an opportunity to stand and stretch every hour.
  - We will take at least one 15-minute break near the halfway point.

About me

- **Experience:**
  - 25+ years consulting and training experience
  - Extensive work with “big data” and analytics
  - 15 years working with various data visualization tools
- **Education**
  - Ed. M., Technology, Innovation & Education, Harvard University
  - PhD Candidate, Education Policy, University of Massachusetts, Amherst

Why kaggle.com

- Ensure that we are all working with a consistent platform.
- Uses jupyter notebooks for code writing and execution.
- Many of datasets to choose from… including several shared by CDC.



matplotlib

Version 3.4.3

seaborn: statistical data visualization

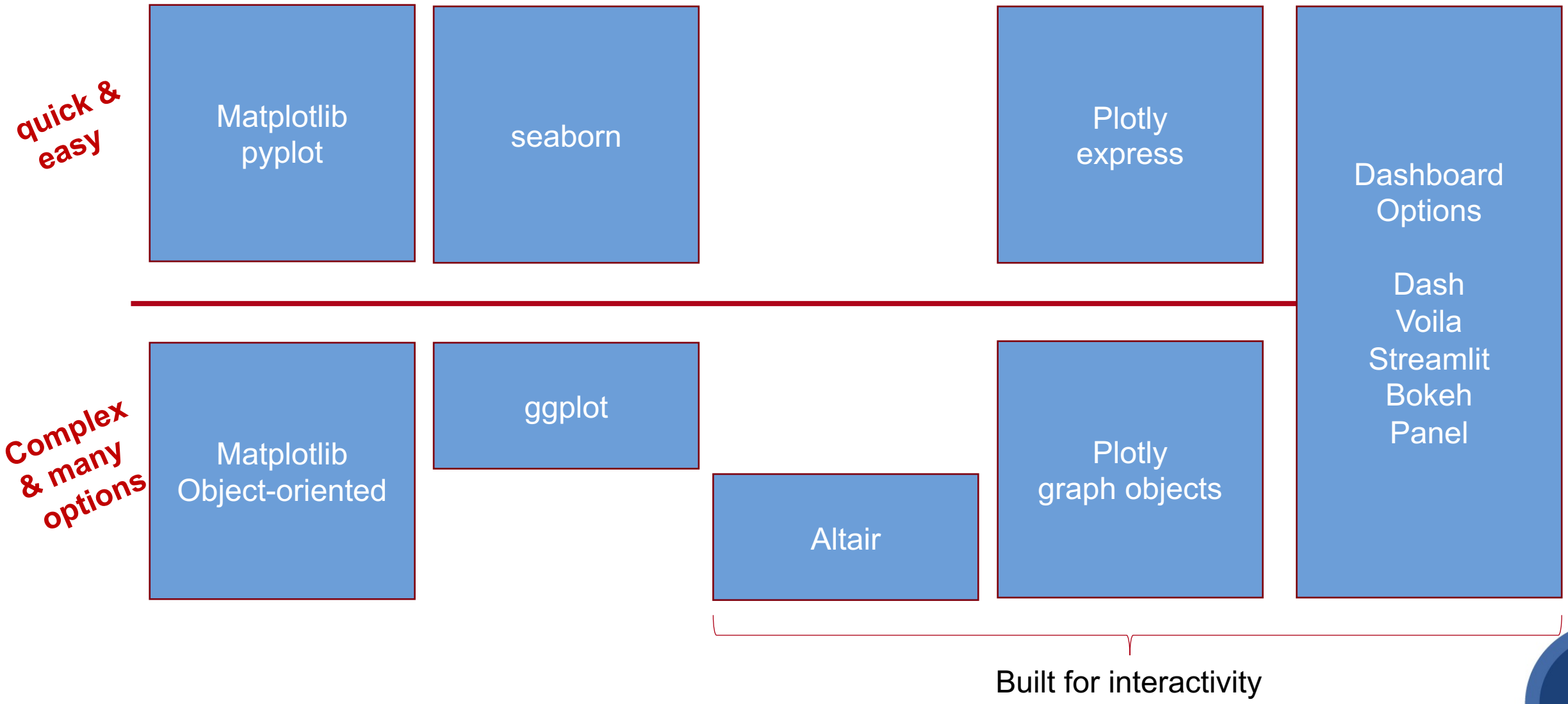
The Bokeh Visualization Library

plotnine 0.8.0 API Gallery Tutorials Site Page

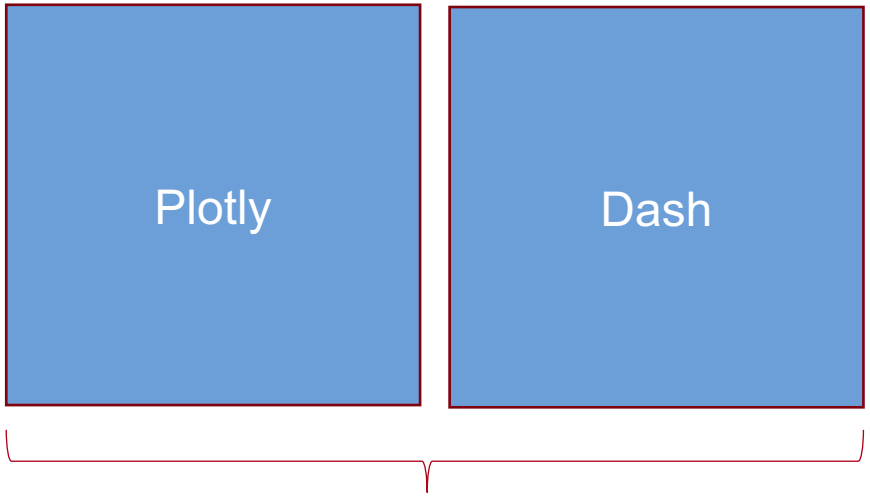
A Grammar of Graphics for Python

Altair: Declarative Visualization in Python

Visualization packages



Our tasks

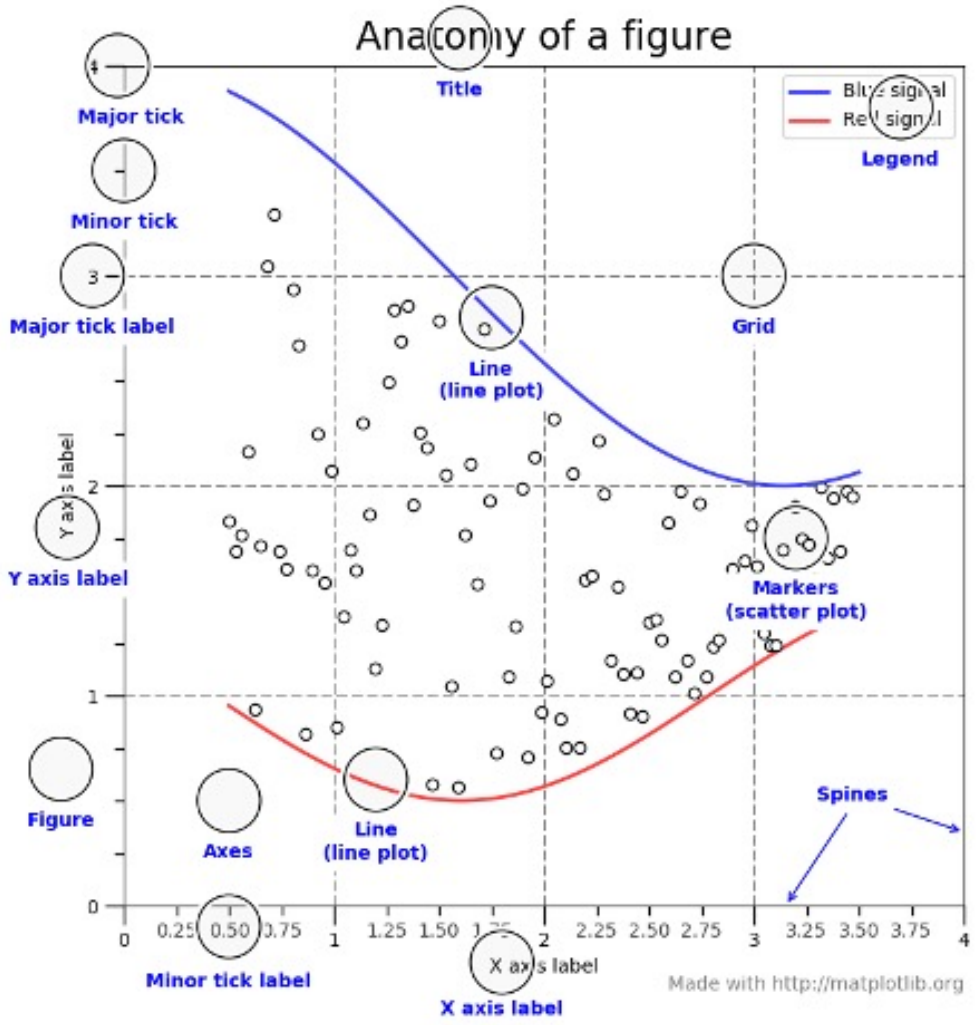


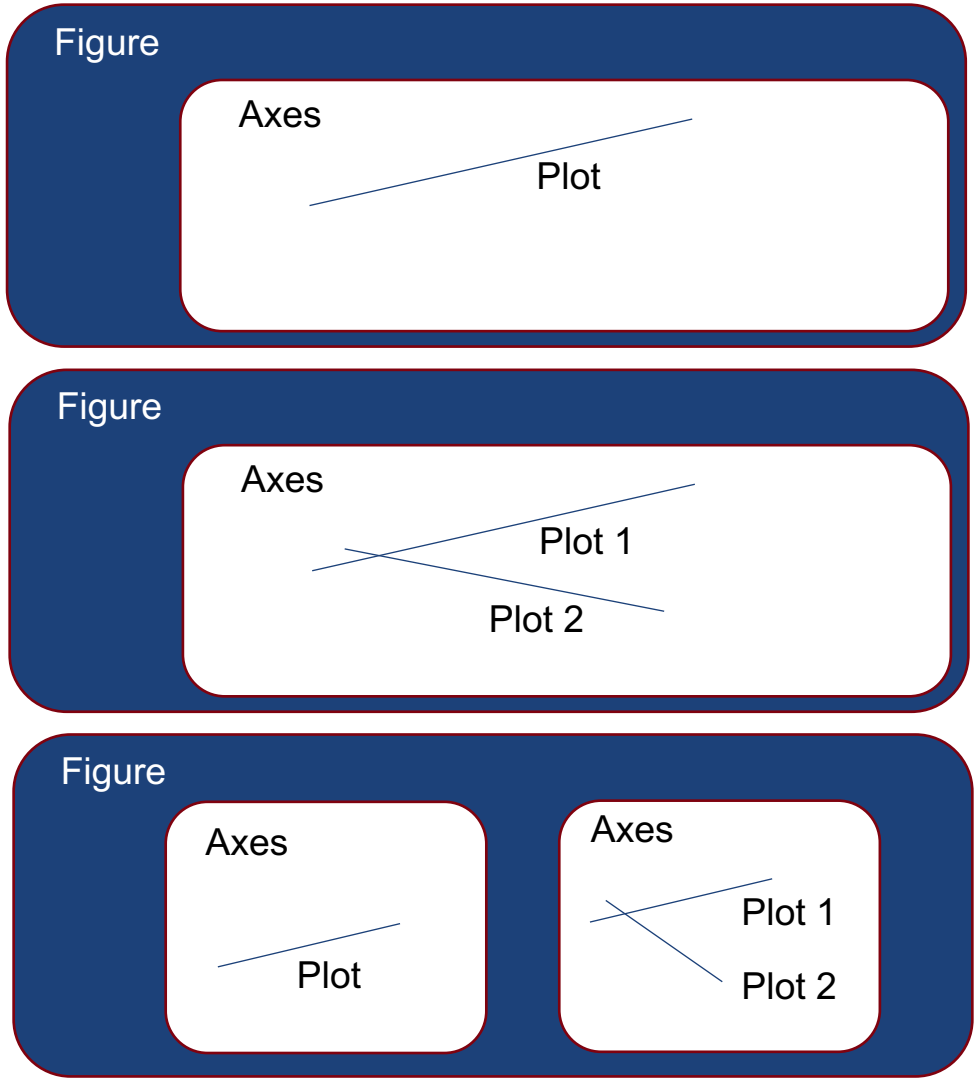
Python
(notebooks
ipywidgets)

git
heroku

and through
kaggle

Matplotlib/Seaborn: Two big concepts to keep in mind

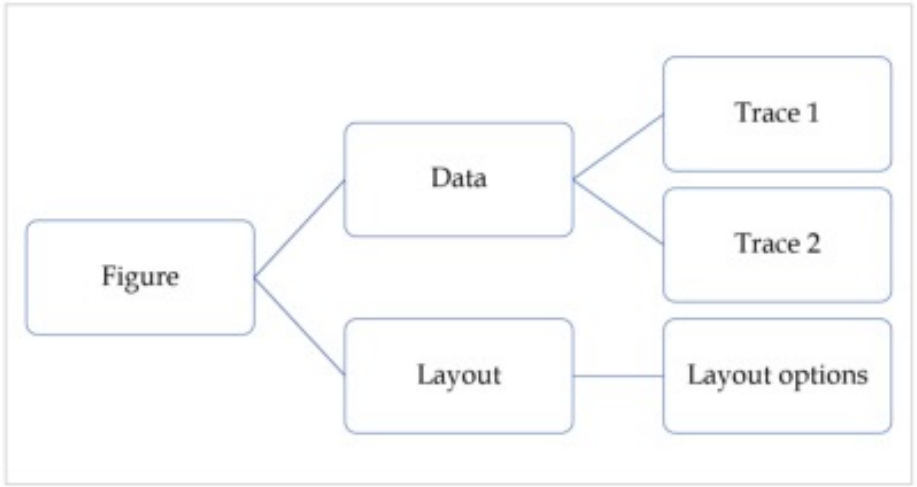




plotly graph objects

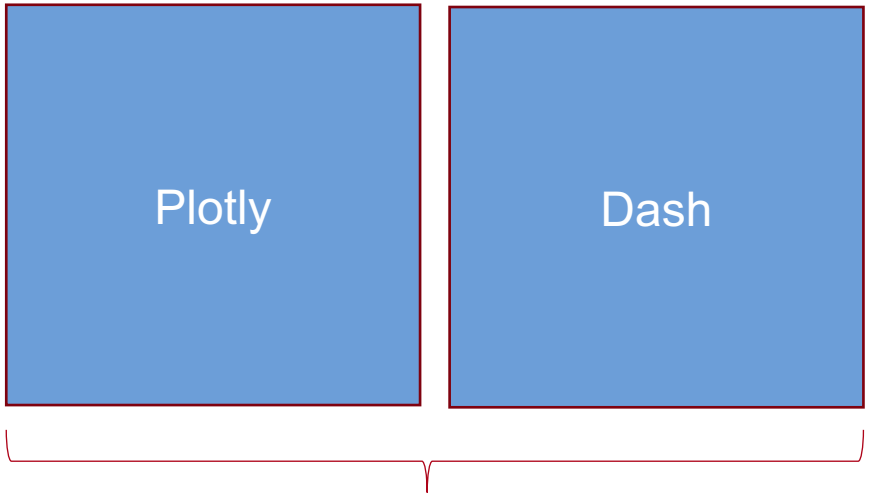
The **plotly.graph_objs module** is the most important module that contains all of the class definitions for the objects that make up the plots you see. Following graph objects are defined:

- Figure,
- Data,
- Layout,
- Different graph traces like **Scatter**, **Box**, **Histogram** etc.



All graph objects are dictionary- and list-like objects used to generate and/or modify every feature of a Plotly plot.

Day 2 - Dash



Python
(notebooks
ipywidgets)

git
heroku

and through
kaggle