

A3 Recap, Plotting, A4

Learning goals:

- Go over commonly asked questions for A3.
- Understand how common Python plotting libraries relate with each other.
- Walk through questions on A4

COGS 108 Winter 2020
Will McCarthy
Discussion 6

wmccarthy@ucsd.edu
OH: Fri 10a-11a on Zoom

Plotting

Why are there so many ways to make the same plot?

- All of these do the same thing:

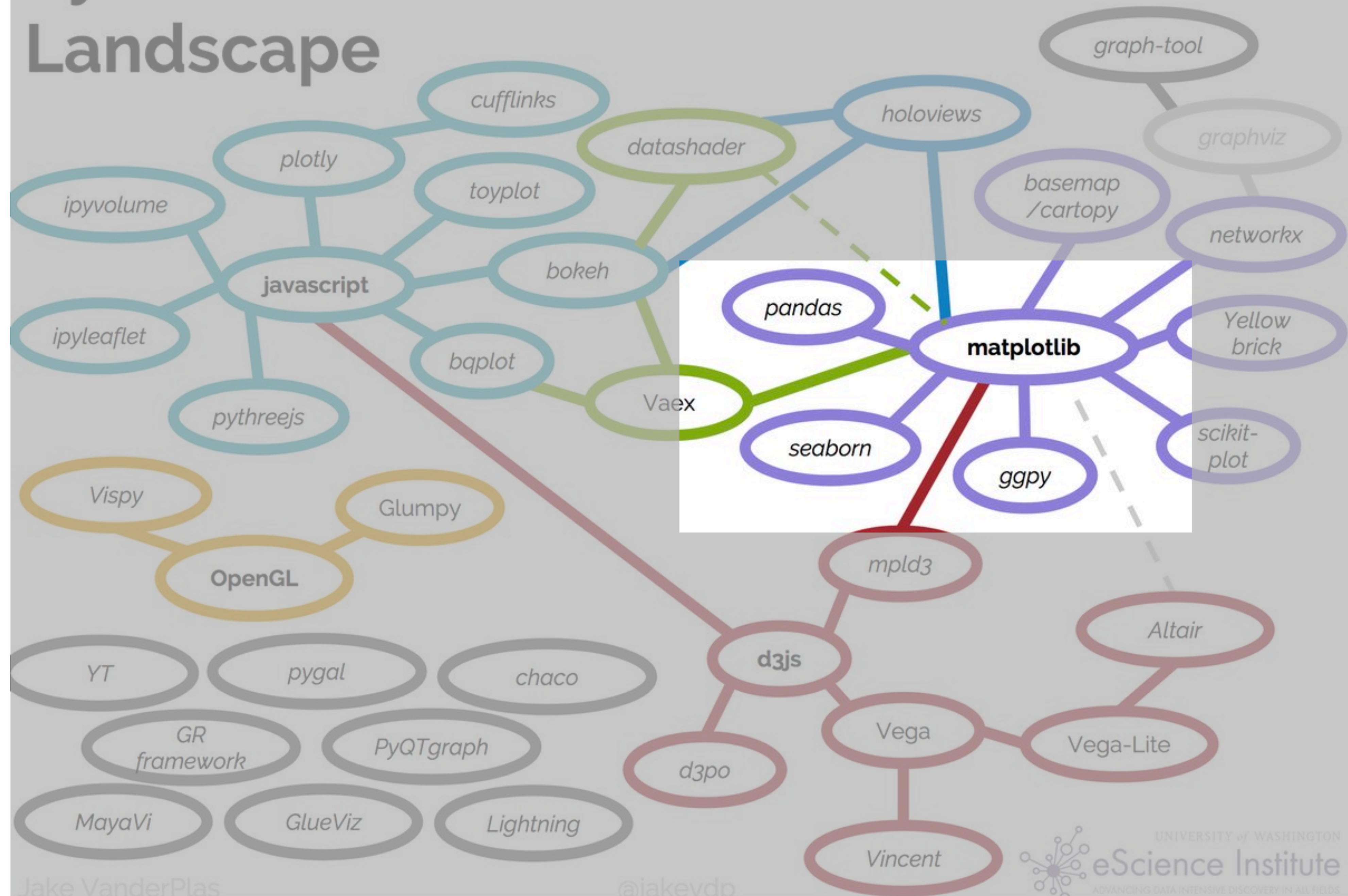
```
plt.hist(df['income10'], 25)  
df['income10'].hist(bins=25)  
df.hist('income10', bins=25)
```

- In Python, most image-based plots created using Matplotlib.

- `plt.hist` `plt.bar` `plt.plot` etc.

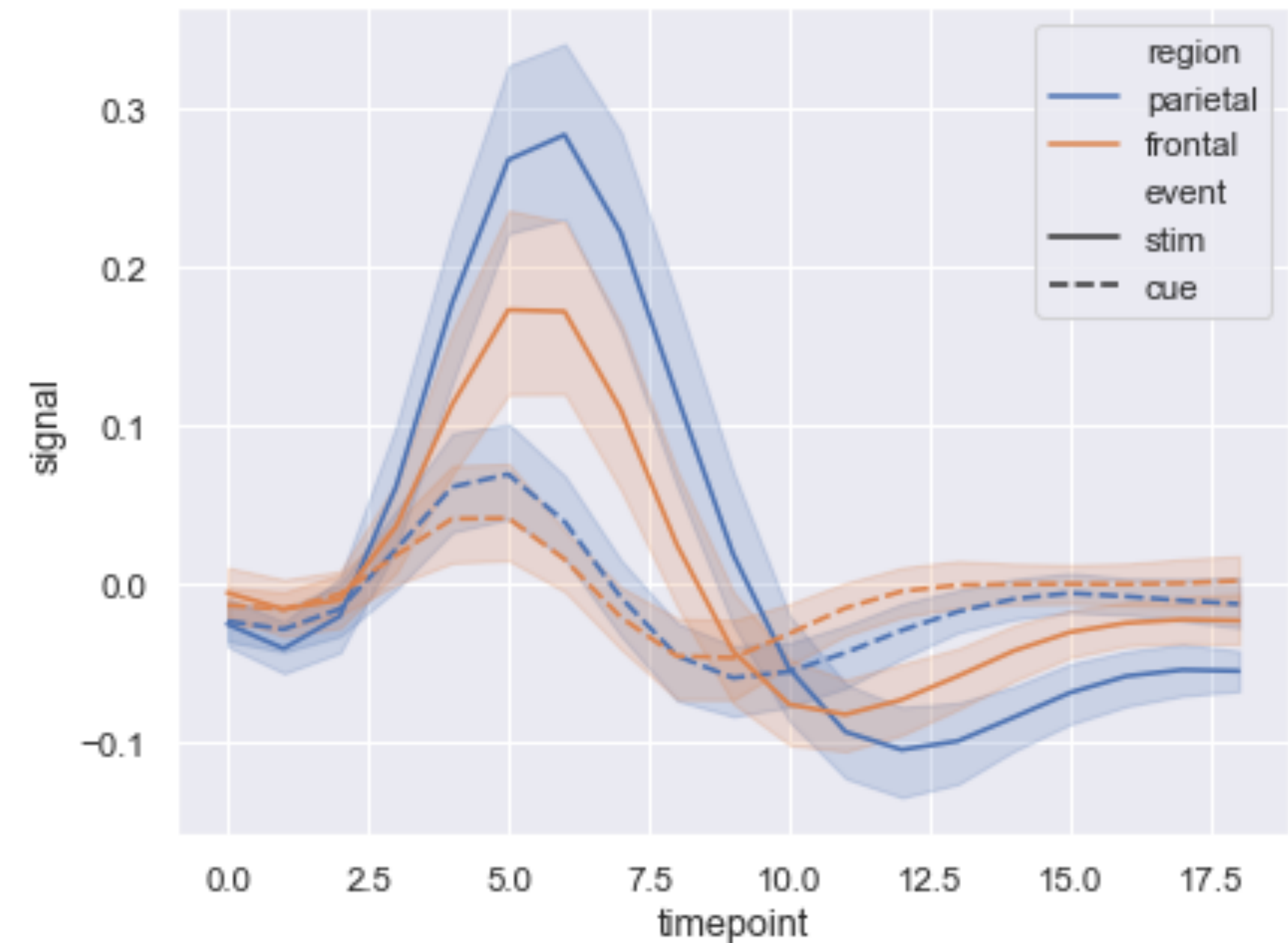
- Pandas gives shortcuts for matplotlib plots. Lines 2 and 3 are shortcuts for line 1.

Python's Visualization Landscape



Seaborn

- **My personal favorite is the seaborn library.**
- **Makes common statistical charts easy to create, like bar plots with confidence intervals.**
- **Again, seaborn is really just a bunch of shortcuts for matplotlib.**



For more details

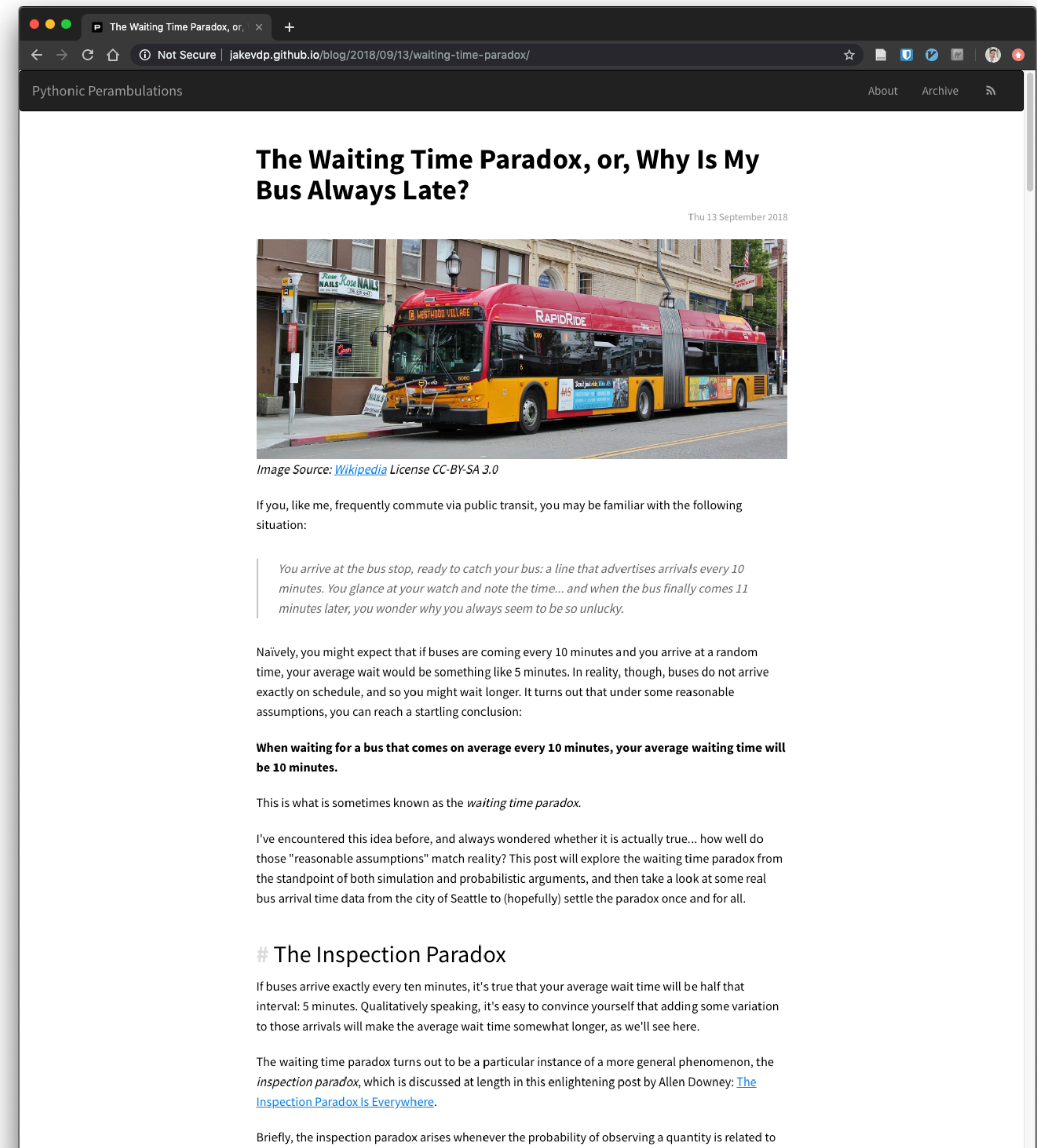
- **Making good plots is a key skill! This just scratches surface.**
- **You can get many great jobs just by being able to make informative data visualizations.**
- **For more, see Ch 6 of textbook.ds100.org.**

A4 Walkthroughs

Preview of next week

An easy way to set up a personal website using Jupyter notebooks and GitHub.

A5 question walkthroughs



A4 quick tips

- **1d: Your DF cells should have ``\n`` at the end (same for 1e)**
- **2b: Don't manually make a new Series – slice a column out of a DF**
- **2e: Use a slice with multiple boolean expressions**
- **2f: Your for loop should loop through the index of a DF**
- **3b: Don't do anything to the zip column**
- **3f: Loop through the DF's index again. Your 3-digit zip codes should be stored as strings, not ints.**