

Today's goals:

- Plotting discussion
- Baby Names Demo
- Dataviz Lab & Q&A

Discussion 3

COGS 108 Winter 2021

Plotting

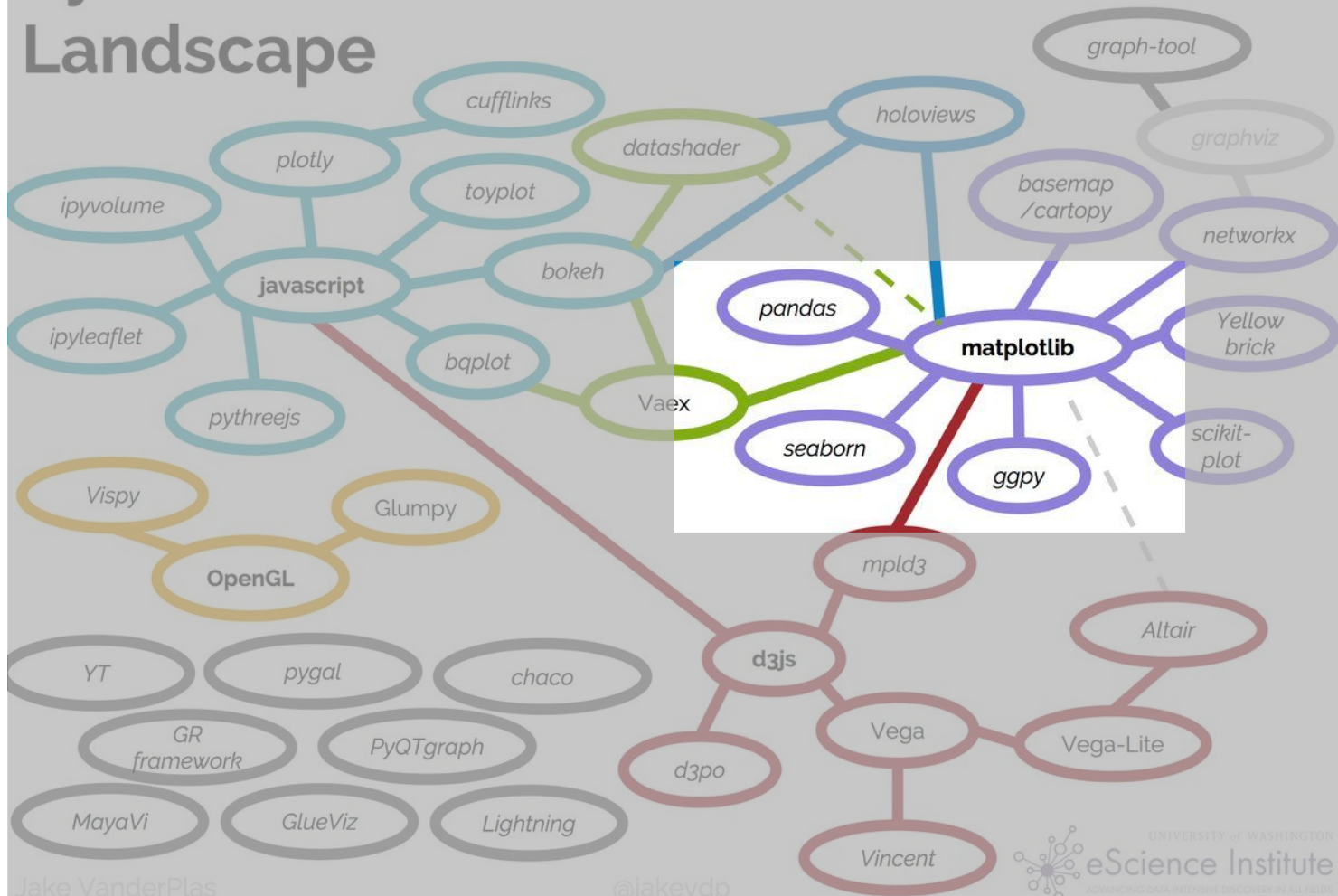
Why are there so many ways to make the same plot?

- All of these do the same thing:

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

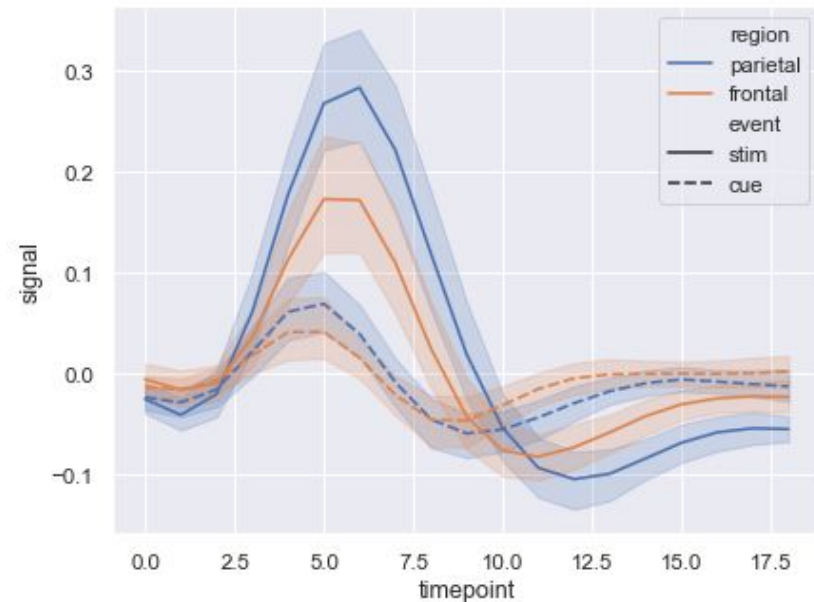
- In Python, most image-based plots created using Matplotlib.
 - **plt.hist** **plt.bar** **plt.plot** etc.
- Pandas gives shortcuts for matplotlib plots. Lines 2 and 3 are shortcuts for line 1.

Python's Visualization Landscape



Seaborn

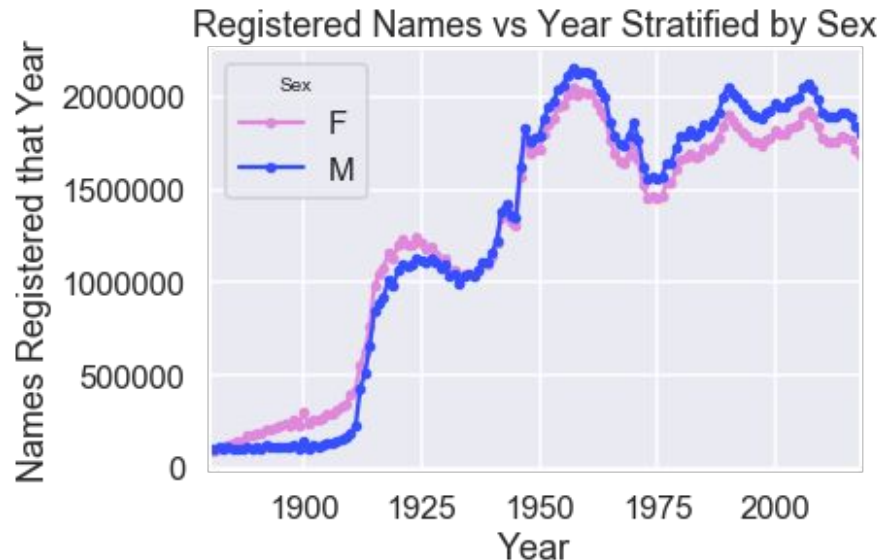
- My personal favorite is the seaborn library.
- Makes common statistical charts easy to create, like bar plots with confidence intervals.
- Again, seaborn is really just a bunch of shortcuts for matplotlib.



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.

pandas and dataviz demo: Baby Names



Demo notebook (BabyNames.ipynb) and data are available on datahub

Your Turn: Work on
D3_dataviz