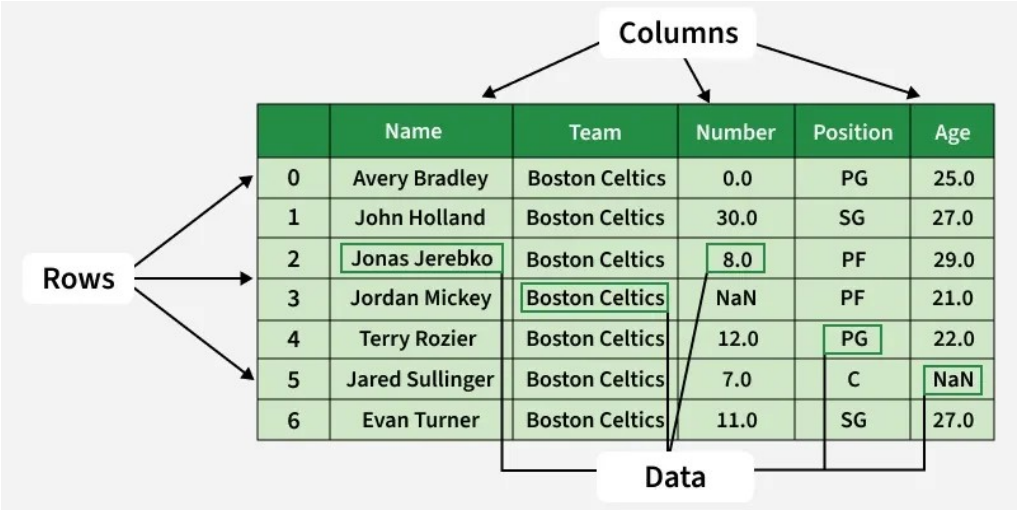
COMP 150:
Pandas and Plotting in Python
Week 10, Lecture 14

Agenda

1. Review - Pandas
2. Visualizing Data
    a. In general
    b. Pitfalls
    c. Options:
        i. Plotly
        ii. Seaborn
        iii. Matplotlib
    d. LLMs - a How-To

Pandas - Quick Review

- Data library
- DataFrames look like tables
- To use:
  - import pandas
  - Read a CSV using pandas.read_csv("filename.csv")
  - Index using data["column_name"]
  - Filter using itself: data[data["column_name"] == value]



Visualizing Data

Why use Python?

- Common
  - Lots of resources out there
  - Great on a resume!
- Fast
  - Becomes easy to use
  - Performance is better than, say, Excel
- Professional, customizable
  - Go as simple or detailed as you please!
- Interactive
  - It's code - let it behave like code

Visualizing Data

Pitfalls:

- Performance
  - Take care to filter ahead of time
  - Use the right framework
- Errors
  - Mistakes in the code = mistakes in the graphs
  - Build self-checking in as you go
  - Take notes

Visualizing Data: Options

Plotly

Seaborn

Matplotlib

Visualizing Data: Framework Options

Plotly

- Works directly with Pandas
- Simple, performant, effective
- Customizable -- to a point
- Interactive

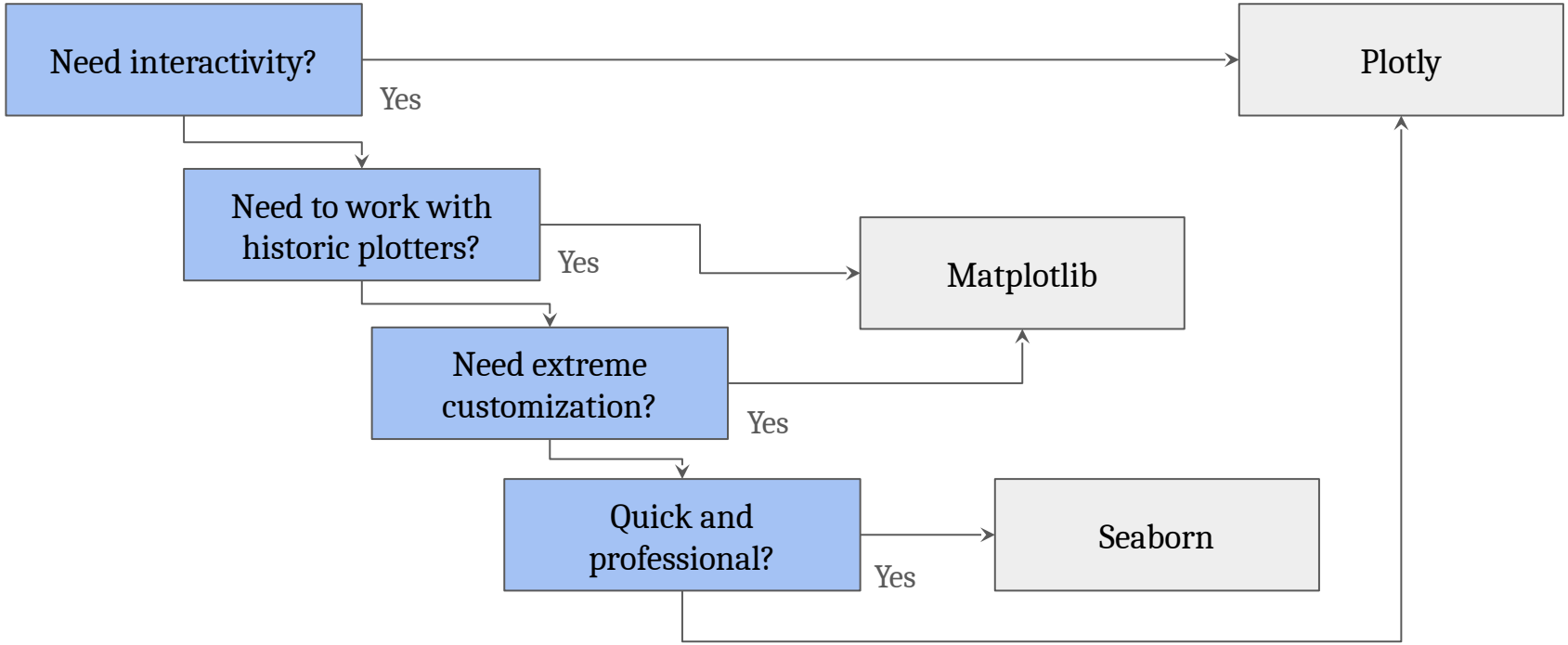
Seaborn

- Works directly with Pandas
- Complex and rich, not as performant
- Out-of-the-box professional
- Highly customizable, but hard to use
- Static

Matplotlib

- Historic
- Looks very similar to MatLab
- Extremely customizable, but not polished
- Difficult to make professional
- Static

Visualizing Data: Options



Plotly

- Express
  - import plotly.express as px
  - Works directly with DataFrames
  - Quick and easy
  - Customization options are vast, *but*
  - Not suitable for significant customization
- Graph Objects
  - import plotly.graph_objects as go
  - Does *not* work directly with DataFrames
  - Does work with lists, dictionaries
  - Extremely detailed
  - Very suitable for customization

Plotly Express

Common:

- Histogram
- Bar
- Line
- Scatter
- 2D Heatmap
- Box/Violin

Plotly Express Histogram

https://plotly.com/python/histograms/

Plotly Express Bar

https://plotly.com/python/bar-charts/

Plotly Express Line

https://plotly.com/python/line-charts/

Plotly Express Scatter

https://plotly.com/python/line-and-scatter/

Plotly Express 2D Heatmap

https://plotly.com/python/2D-Histogram/

Plotly Express Box/Violin

https://plotly.com/python/box-plots/

https://plotly.com/python/violin/