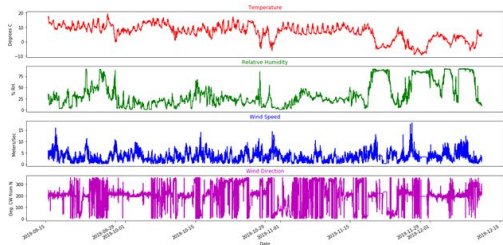
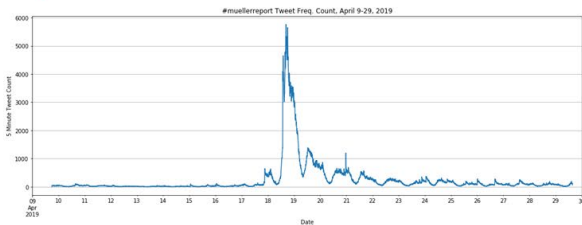
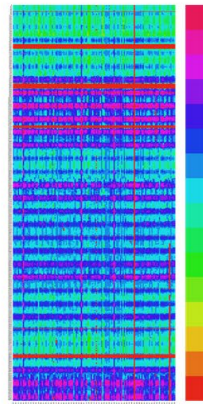
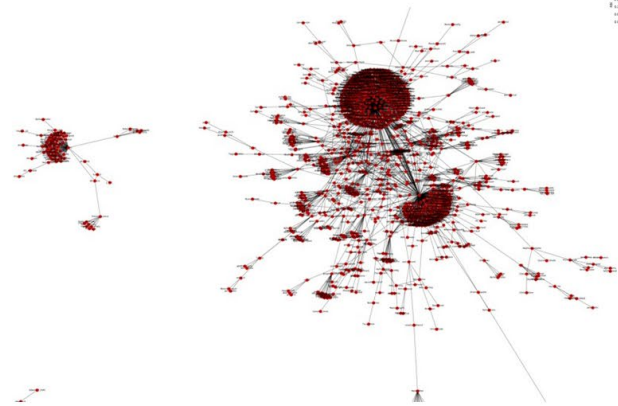
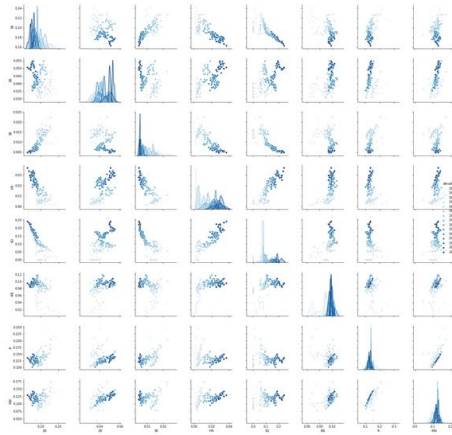


# Data Visualization With Python Using Jupyter Notebooks

Jeff Sale

SDSC Learning Design Technologist  
XSEDE ECSS Visualization Consultant  
XSEDE Workforce Development

[https://github.com/sdsc-hpc-training-org/data\\_vis\\_with\\_python](https://github.com/sdsc-hpc-training-org/data_vis_with_python)









# Acknowledgements

- Chris Myers\*, Senior Research Associate, [Center for Advanced Computing](#), Adjunct Professor, Dept. of Physics, Cornell University
- Susan Mehringer\*, Associate Director, Consulting, Cornell University Center for Advanced Computing, & XSEDE L3 for Training, WBS 2.1.2
- Kate Starbird, Associate Professor, Human Centered Design & Eng., U. of Washington
- Jen Zeimke, Associate Professor, Political Science, John Carroll Univ.
- Mary Thomas and Bob Sinkovits, SDSC User Training



**\*Data Science With Python Tutorial**  
Cornell Virtual Workshop, <https://cvw.cac.cornell.edu/>

*XSEDE Tutorial*  
*Python for Data Science*

Cornell Virtual Workshop

**Part 1: <https://cvw.cac.cornell.edu/PyDataSci1/>**

**Part 2: <https://cvw.cac.cornell.edu/PyDataSci2/>**

# Commands to Get Started

Install Bokeh and Holoviews

```
conda install -c pyviz holoviews bokeh
```

Install Datashader

```
conda install datashader
```

Clone the Holoviews and DataShader Repos

```
git clone https://github.com/holoviz/holoviews.git
```

```
git clone https://github.com/holoviz/datashader.git
```

# Commands to Get Started

Download the Data for this Webinar

**wget**

**[https://education.sdsc.edu/hpc\\_training\\_series\\_2021/python\\_vis\\_webinar\\_data.zip](https://education.sdsc.edu/hpc_training_series_2021/python_vis_webinar_data.zip)**

Run the shell script to reserve an Expanse compute node and launch a Jupyter Notebook

**/cm/shared/apps/sdsc/galileo/galileo.sh launch -j notebook -A **abc123** -p compute -n 1 -M 8 -t 00:30:00 --conda-env base**

# The Jupyter Project

“Project Jupyter exists to develop open-source software, open-standards, and services for interactive computing across dozens of programming languages.”

<https://jupyter.org/>



# The Jupyter Menagerie



## Notebook Widgets

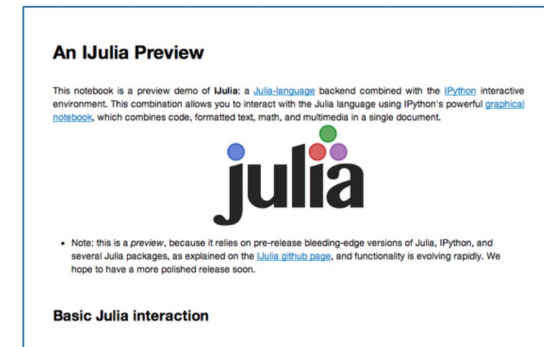
IPython



IRuby



IJulia



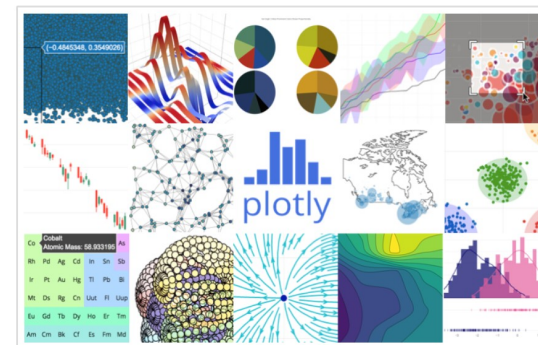
Data Visualization with Lightning



Interactive data visualization with Bokeh



Interactive plots with Plotly





# Python's Visualization Landscape

The diagram illustrates the Python visualization ecosystem, showing the relationships between various libraries and frameworks. The central node is **matplotlib**, which is connected to **holoviews**, **pandas**, **seaborn**, **ggpy**, **basemap / cartopy**, **networkx**, **Yellow brick**, **scikit-plot**, **Altair**, **Vega**, **Vega-Lite**, **Vincent**, **d3js**, **mpld3**, **d3po**, **Vaex**, **datashader**, **bokeh**, **toyplot**, **plotly**, **ipyvolume**, **ipyleaflet**, **pythreejs**, **bqplot**, **graph-tool**, **graphviz**, **Vispy**, **Glumpy**, **OpenGL**, **YT**, **pygal**, **chaco**, **GR framework**, **PyQTgraph**, **MayaVi**, **Glueviz**, and **Lightning**. The connections are color-coded: teal for JavaScript-related libraries, purple for Matplotlib-related libraries, green for Vaex-related libraries, red for D3.js-related libraries, and grey for other libraries.

**JavaScript** (teal): **ipyvolume**, **ipyleaflet**, **pythreejs**, **plotly**, **bokeh**, **toyplot**, **bqplot**.

**Matplotlib** (purple): **holoviews**, **pandas**, **seaborn**, **ggpy**, **basemap / cartopy**, **networkx**, **Yellow brick**, **scikit-plot**, **Altair**, **Vega**, **Vega-Lite**, **Vincent**, **d3js**, **mpld3**, **d3po**.

**Vaex** (green): **datashader**, **bokeh**, **toyplot**, **bqplot**.

**D3.js** (red): **d3po**, **Vega**, **Vega-Lite**, **Vincent**.

**Other** (grey): **graph-tool**, **graphviz**, **Vispy**, **Glumpy**, **OpenGL**, **YT**, **pygal**, **chaco**, **GR framework**, **PyQTgraph**, **MayaVi**, **Glueviz**, **Lightning**.

**Footer:** Jake VanderPlas, @jakevdp, UNIVERSITY OF WASHINGTON eScience Institute

# Bokeh

“Bokeh is an interactive visualization library for modern web browsers. It serves as a web-based front end to matplotlib using JavaScript behind the scenes to run in a browser”.

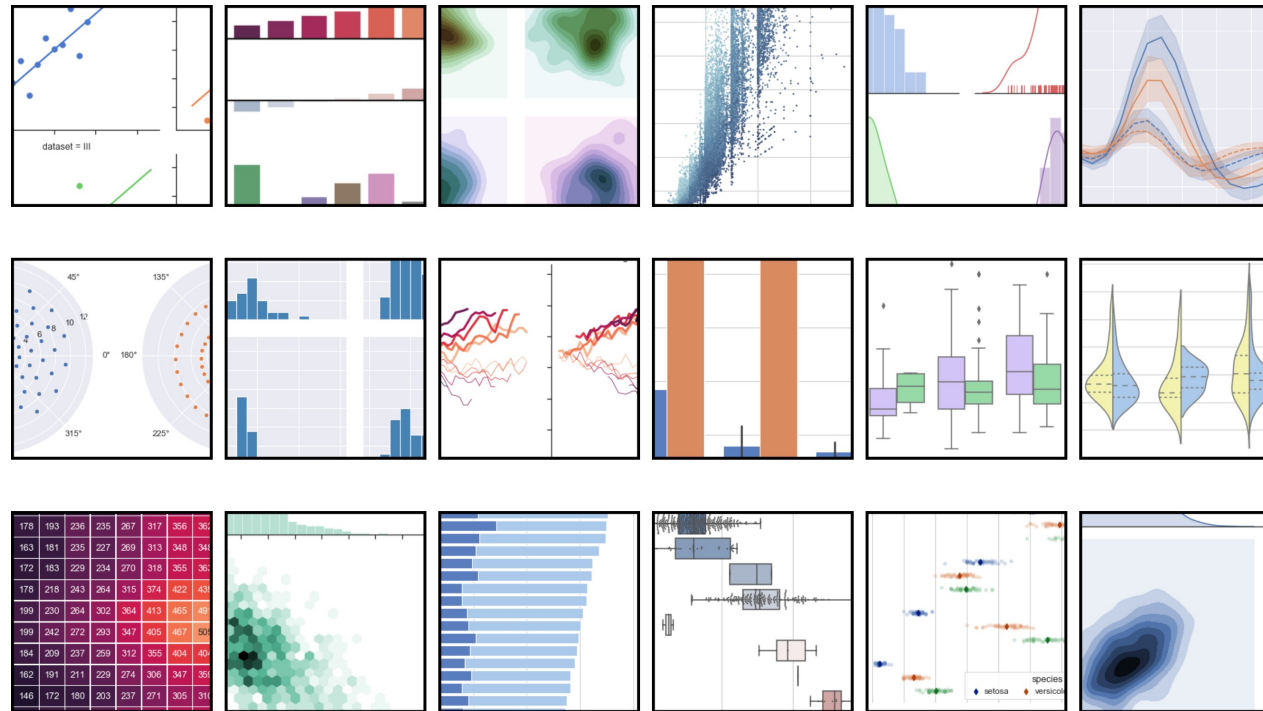
<https://docs.bokeh.org/en/latest/index.html>



# Seaborn

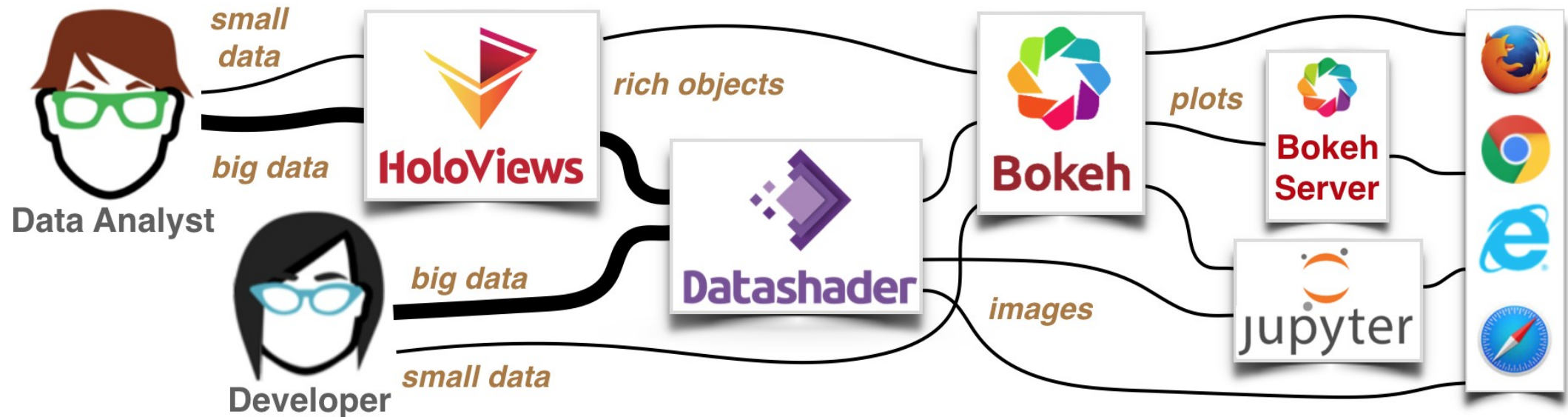
Seaborn is a Python data visualization library based on [matplotlib](https://matplotlib.org/). It provides a high-level interface for drawing attractive and informative statistical graphics.

<https://seaborn.pydata.org/>





# HoloViz



# Ways to Access Jupyter on XSEDE Systems

- TACC Visualization Portal
  - <https://vis.tacc.utexas.edu/>
- SDSC Expanse\*
  - [https://education.sdsc.edu/training/interactive/202012\\_running\\_jupyter\\_notebooks\\_on\\_expanse/index.html](https://education.sdsc.edu/training/interactive/202012_running_jupyter_notebooks_on_expanse/index.html)
  - <https://hpc-training.sdsc.edu/notebooks-101/notebook-101.html>
- PSC Bridges 2
  - <https://www.psc.edu/user-resources/software/jupyter>

\*Run new script, galileo.sh, using this command:

```
/cm/shared/apps/sdsc/galileo/galileo.sh launch -j notebook -A abc123 -p compute -n 1 -M 8 -t 00:30:00 --conda-env base
```

# Visualization Tutorials

- Scientific Visualization with VisIt, Amit Chourasia, SDSC  
Director of Visualization Services
  - <http://users.sdsc.edu/~amit/scivis-tutorial/>
- Hyperglyph Visualization, Jeff Sale
  - [https://www.iluvdata.org/antz/toroids/tutorials/intro\\_lessons/index.html](https://www.iluvdata.org/antz/toroids/tutorials/intro_lessons/index.html)



*Let's just  
jump into it!*

