

IS 445 - ACG/ACU: Data Visualization - Fall 2023

Assignment 4

Student: Christopher Mujjabi

Date: October 16, 2023

Instructions

Compare Visualization Engines

Your assignment is to pick four of the following possibilities and write up a set of comparisons for constructing the same visualization. You must evaluate matplotlib, vega-lite and bqplot, and you can choose one of the following in addition: D3, Bokeh, Plotly, R/RStudio. These should total roughly half a page per engine.

1. **Matplotlib**

What is the license for the software?

The Matplotlib license is based on the PSF (Python Software Foundation) license and only uses BSD (Berkeley Software Distribution) compatible code. The BSD License is a permissive open source software, which means that users can freely use, modify and redistribute the software in open-source and proprietary projects with minimal restrictions.

What is the focus of the software?

Matplotlib focuses on creating static, animated and interactive visualizations in Python. Using matplotlib, one can generate publication-quality plots, make customized visual styles and layouts, make interactive visualizations that can zoom, pan and update and can be exported in different file formats.

Does it have interactivity, and how easy is it?

matplotlib is mostly used to generate static visualizations; however, since matplotlib can embed JupyterLab, it is possible to use the Jupyter extension to add widgets and other functionality that can transform the static plots into a more interactive visualization.

What are the pros and cons of using it?

Some of the pros for using matplotlib include the wide support community given that research scientists and data analytics communities widely use it. Therefore, new users can access a lot of information to simplify the learning journey. It is also easily accessible, given the free-access license. Matplotlib is also highly customizable, allowing users to manipulate and customize charts in different styles.

The cons of using this software include the steep learning curve for new or novice users, especially users who are new or have never used Python before. Users need to learn Python to understand the matplotlib syntax better.

2. Vega-lite

What is the license for the software?

Vega-lite is licensed under the BSD 3-Clause "New" or "Revised" License, which is similar to the BSD 2-Clause License. However, it has a 3rd clause that prohibits users from using the name of the copyright holder or its contributors to promote derived products without written consent.

What is the focus of the software?

Vega-lite is a high-level visualization grammar for generating interactive visualization using JSON syntax. Vega-lite also serves as a declarative format for creating visualization.

Does it have interactivity, and how easy is it?

Vega-lite excels in interactivity by offering a straightforward and concise way to specify interactive features using a JSON-based specification. It supports interactive features such as tooltips that display information when you hover over the data points, zooming to explore specific areas of the plot and panning to navigate large datasets.

What are the pros and cons of using it?

The pros for Vega-lite are the ability to embed the Vega-lite interactive visualization and deploy it in a JSON format on GitHub pages. I think this makes graphical sharing and data presentation very easy. Also, vega-lite visualizations can be exported in various ways, including PGN, HTML, SVG, PDF and JSON. The other vega-lite component I personally liked was the vega-lite documentation/guides, which explicitly explains all the features that can be used to customize the visualization.

However, I personally found the syntax very difficult to use since I wasn't familiar with Javascript.

3. Bqplot

What is the license for the software?

Bqplot is licensed under the Apache License 2.0. This permissive license allows users to use, modify, distribute and create derivative work using the software without additional permissions.

What is the focus of the software?

The Bqplot library was built on top of javascript and Python library ipywidgets, and it is mainly used to build fully interactive web applications in a Python environment using fewer lines of code.

Does it have interactivity, and how easy is it?

Bqplot specializes in generating highly interactive web-based visualizations in Python through jupyter notebooks using numerous tools such as IPywidgets. For example, we can use bqplot to generate colored maps, add tooltips to hover over specific areas/regions on the map and obtain summarized information about that region.

What are the pros and cons of using it?

Some of the pros include the rich set of widgets that can be used for interactivity and how biplot requires fewer lines of code to generate interactive plots. Also, since it requires less coding, it is more user-friendly, particularly for novice users. One of the cons would be the limitation to use bqplot outside of jupyter notebook.