Creating Interactive Figures with Bokeh+Python

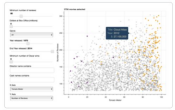
Aaron Geller

Northwestern | INFORMATION TECHNOLOGY

What is Bokeh?

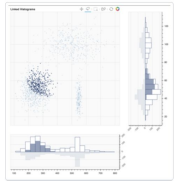
Pronounced either “boh-Kay” or “boh-Kuh”

Python’s Bokeh is a visualization library that provides tools for constructing interactive (mostly 2D) figures and apps using Python and sometimes a little Javascript. It can create interactive plots that can be embedded directly within your website.



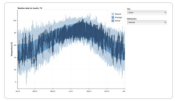
An interactive query tool for a set of IMDB data

Source code: movies



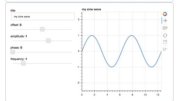
Shows axis histograms for selected *and* non-selected points in a scatter plot

Source code: selection_histogram



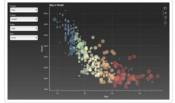
Interactive weather statistics for three cities

Source code: weather



A basic demo that has sliders for controlling a plotted trigonometric function

Source code: sliders.py



Explore the "autompg" data set by selecting and highlighting different dimensions

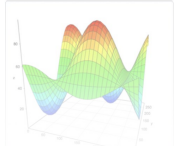
Source code: crossfilter

A reproduction of the famous Gapminder demo, with embedded video added using a custom page template

Source code: gapminder

Linked plots, summary statistics, and correlations for market data

Source code: stocks



An updating 3d plot that demonstrates using Bokeh custom extensions to wrap third-party JavaScript libraries

Source code: surface3d

[See the Bokeh gallery for some possibilities.]

Why use Bokeh?

(Why not, e.g., Plotly, D3, Tableau, etc.?)

Python — We love working in Python! And all of our scripts are already written in Python.

Interactive — We want to use custom sliders, buttons, dropdowns, etc. to manipulate our plots in real time.

Easy — We want something that is easy to get started in and doesn't require building everything from scratch.

Shareable — We'd like to be able to create interactive plots for free that can be shared easily with collaborators, posted on websites and possibly included in online journal articles.

Are there limitations?

(Yes, of course.)

Look	Bokeh has its own look and feel (toolbar, axes, fonts, available symbols, possible interactions, etc.). If you don't like it, too bad.
Big data	If you have lots of data, Bokeh may be too slow for your needs. 1000s of points are OK, but 100k (or large images) is going to push the limits.
3D	If you really want to plot 3D, you may want to look elsewhere. Bokeh can plot 3D functions, but not a 3D scatterplot (as far as I know).
Docs	Though there is some documentation and some examples online, it is limited. Other libraries (e.g., D3) have way more examples online to pull from.

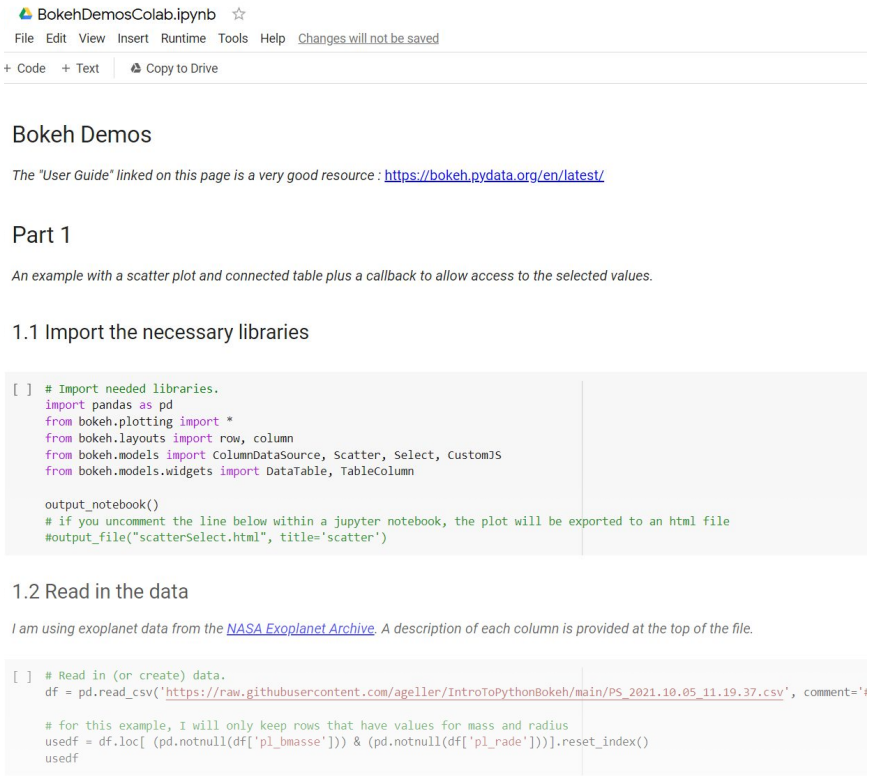
How to get started

1. Decide the ***story*** that you want to tell with your data, and how you will do it with an interactive figure
2. Prepare your ***data*** in Python (pandas works well with Bokeh)
3. Make a static ***plot*** with matplotlib first
4. Determine which "***tools***" and "***widgets***" you will want
5. ***Consult*** the gallery, google for examples, look at the docs, ask me!

Typical components

1. **ColumnDataSource** : holds a python dictionary (or a panda DataFrame) containing your data
2. **Tools** : simple manipulators (zoom, pan, select, etc.)
3. **Figure** : the plot (scatter, line, bar, etc.)
4. **Table**: a data table connected to the figure (optional)
5. **Widgets** : more sophisticated manipulators (sliders, buttons, dropdowns, etc.)
6. **Layout** : how to position the figures, tables and widgets in some grid
7. **Show** : display your plot and/or save it to an .html file

Hands-on example



[Click here to access the notebook in colab.]