

# A tour of data viz in Python



Éléonore Mayola



@EleonoreMayola | github.com/eleonore9  
Elle-est-au-nord.com

PYCON UK | Cardiff | Sept 2019

github.com/eleonore9/  
tour\_dataviz\_python



- Exhaustive list of Python tools for data viz: [pyviz.org/tools.html](http://pyviz.org/tools.html)
- Libraries mentioned:

matplotlib.org  
pandas.pydata.org  
plotly  
bokeh.org  
altair-viz.github.io

- Geospatial libraries for Python: **Geopandas, Folium, Cartopy, gmplot, ipyleaflet, geoviews**

PYCON UK | Cardiff | Sept 2019

## matplotlib

"Matplotlib tries to make easy things easy and hard things possible."

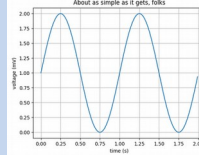
```
import matplotlib
import matplotlib.pyplot as plt
import numpy as np
```

```
# Data for plotting
t = np.arange(0.0, 2.0, 0.01)
s = 1 + np.sin(2 * np.pi * t)
```

```
fig, ax = plt.subplots()
ax.plot(t, s)

ax.set(xlabel='time (s)', ylabel='voltage (mV)',
       title='About as simple as it gets, folks')
ax.grid()
```

```
fig.savefig("test.png")
plt.show()
```

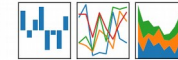


- Complex or customised plots
- Syntax can become tricky

PYCON UK | Cardiff | Sept 2019

## pandas

$$y_i = \mu_i + \epsilon_i \quad \beta = \mu_i$$



"[...] high-performance, easy-to-use data structures and data analysis tools for the Python programming language."

```
In [1]: %matplotlib inline
```

```
In [2]: import matplotlib
import matplotlib.pyplot as plt
matplotlib.style.use('seaborn')
import numpy as np
import pandas as pd
```

```
In [3]: df = pd.DataFrame(np.random.randn(1000, 4),
                        columns=list('ABCD'))
```

```
df = df.cumsum()
```

```
plt.figure()
```

```
df.plot()
```

```
plt.show()
```

```
<Figure size 576x396 with 0 Axes>
```



- Plotting during data analysis
- Not the most aesthetic → try Seaborn

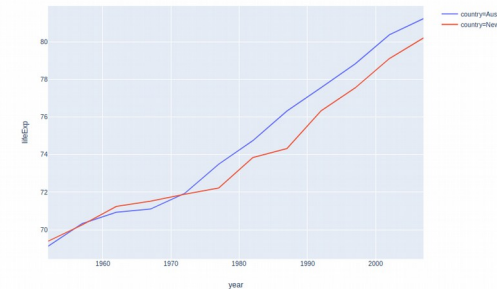
PYCON UK | Cardiff | Sept 2019



High-level, declarative charting library with over 30 chart types, including scientific charts, 3D graphs, statistical charts, SVG maps, financial charts, and more.

```
In [1]: import plotly.express as px
```

```
In [2]: Gapinder = px.data.gapinder().query("continent='Oceania'")
fig = px.line(Gapinder, x='year', y='lifeExp', color='country')
fig.show()
```



- Notebook or plain html report
- Maintained by a private company

"Altair is a declarative statistical visualization library for Python, based on Vega and Vega-Lite."



PYCON UK | Cardiff | Sept 2019

- Interactive plots and maps
- One main maintainer

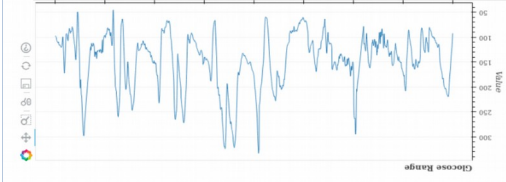


"Bokeh is an interactive visualization library for Python that enables beautiful and meaningful visual presentation of data in modern web browsers."



PYCON UK | Cardiff | Sept 2019

```
In [1]: from bokeh.plotting import figure
from bokeh.io import output_notebook, show
# first, to import output notebook, show
# bokeh.sample_data.download()
# bokeh.sample_data
# import bokeh
from bokeh.sample_data import data
# reduce data size to one week
week = data.loc['2010-10-01': '2010-10-08']
p = figure(x轴='date', y轴='glucose', title='Glucose Range', plot_width=800)
p.grid(x轴='date', y轴='glucose')
p.xaxis.axis_label = 'date'
p.yaxis.axis_label = 'glucose'
p.line(week.index, week['glucose'])
show(p)
```



- Interactive plots or dashboards
- Tricky to create dashboard apps