Plotly and Cufflinks

# Set up

* Plotly is an open-source interactive visualization library.
* Cufflinks connects plotly to pandas.
* Plotly is a business as well and the business makes money by hosting visualizations and dashboards. You can host locally without paying for hosting.
* You need to import plotly and cufflinks from the command prompt with pip install plotly and pip install cufflinks respectively.
* from plotly.offline import download\_plotlyjs allows you host it locally
* To use these, import
* numpy,
* pandas
* cufflinks as cf,
* from plotly.offline import download\_plotlyjs, init\_notebook\_mode, plot, iplot.
* To make sure it works in your notebook, call init\_notebook\_mode(connected = True)
* To allow you use cufflinks offline, use cf.go\_offline()
* Make sure the version is at least 1.10

# Usage

* Instead of the typical df.plot() which produces a line graph of the data frame, you simply add an ‘i’ to make it df.iplot() and you get an interactive line plot.

## Histplot

* To produce other kinds of plots, simply pass in the kind argument in e.g., df.iplot(kind = ‘hist’) or for a histplot.

## Scatter plot

* You can create a scatter plot with df.iplot(kind = ‘scatter’, x = ‘a’, y = ‘b’). by default, lines connect all the points, to remove, pass in argument mode = ‘markers’ e.g., df.iplot(kind = ‘scatter’, x = ‘a’, y = ‘b’, mode = ‘markers’).
* You can pass in other scatter plot customization arguments such as size etc.

## Bar plot

* You can create a bar plot with df.iplot(kind = ‘bar’, x = ‘a’, y = ‘b’)
* You can use an aggregate call such as mean, sum, etc. if the data does not fit a bar plot by default with df.t().iplot(kind = ‘bar’, x = ‘a’, y = ‘b’) where t is the aggregate.

## Surface plot

* You can create a surface plot with df.iplot(kind = ‘surface’) which creates a 3 D surface representation of the data that suits the data frame.
* You can customize the surface plot by passing in the arguments needed.

## Spread plot

* You can also create a spread plot type which is common in the financial markets with df.iplot(kind = ‘spread’). This creates a line plot of the values and the spread between them in a spread plot.

## Bubble plot

* You can create a bubble plot with df.iplot(kind = ‘bubble, x = ‘a’, y = ‘b’, size = ‘c’) where c is the column that determines the size of the bubble.
* Bubble plots are common for population analysis, GDP etc. especially by multinational organizations.

## Scatter matrix plot

* A scatter matrix plot is similar to seaborn’s pair plot. It creates a plot of all the numerical data in a data frame against one another.
* You can create a scatter matrix with df.scatter\_matrix()