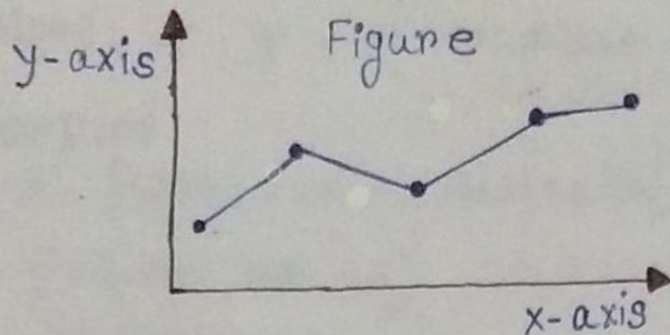# Data Viz Cheat Sheet

Matplotlib is a python 2D plotting library that produces figures in a variety of formats.

Matplotlib
workflow

The basic steps to creating plots with matplotlib are prepare Data, Plot, customize plot, save plot and show plot.

```
import matplotlib.pyplot as plt
```

Example with lineplot

Prepare data
```
x = [2017, 2018, 2019, 2020, 2021]

y = [43, 45, 47, 48, 50]
```
Plot & customize plot
```
plt.plot (x, y, marker = 'o', linestyle = '__',
color = 'g', label = 'USA')
plt.xlabel ('years')
plt.ylabel ('population (M)')
plt.title ('years vs population')
plt.legend (loc = 'lower right')
plt.yticks ([41, 45, 48, 51])
```
Save plot
```
plt.savefig ('example.png')
```

Show plot
```python
plt.show()
```
Marker: '.' , 'o' , 'v' , '<' , '>'

Line styles: '_' , '__' , '-.' , ':'

Colors: 'b' , 'g' , 'r' , 'y' #blue, green, red, yellow

Barplot
```python
x = ['USA', 'UK', 'Australia']
y = [40, 50, 33]
plt.bar(x,y)
plt.show()
```

Piechart
```python
plt.pie(y, labels=x, autopct='%.of%%')
plt.show()
```

Histogram
```python
ages = [15, 16, 17, 30, 31, 32, 35]
bins = [15, 20, 25, 30, 35]
plt.hist(ages, bins, edgecolor='black')
plt.show()
```

Boxplot
```python
ages = [15, 16, 17, 30, 31, 32, 35]
plt.boxplot(ages)
plt.show()
```

Scatterplot
```python
a = [1, 2, 3, 4, 5, 4, 3, 2, 5, 6, 7]
b = [7, 2, 3, 5, 5, 7, 3, 2, 6, 3, 2]
plt.scatter(a,b)
plt.show()
```

Subplots

Add the code below to make multiple plots with 'n' number of rows and columns.

```python
fig, ax = plt.subplots(nrows = 1,
                       ncols = 2,
                       sharey = True,
                       figsize =(12,4))
```

Plot & Customize Each Graph

```python
ax[0].plot(x, y, color ='g')
ax[0].legend()
ax[1].plot(a,b, color = 'r')
ax[1].legend()
plt.show()
```

Seaborn

workflow

```python
import seaborn as sns
import matplotlib.pyplot as plt
import pandas as pd
```

Lineplot

```python
plt.figure(figsize = (10,5))
flights = sns.load_dataset("flights")
may_flights = flights.query("month=='may'")
ax = sns.lineplot(data=may-flights,
                  x = "year"
                  y = "passengers")
ax.set(xlabel ='x', ylabel ='y',
       title ='my-title, xticks =[1,2,3])
ax.legend(title = 'my-legend,
          title-fontsize = 13)
plt.show()
```

Barplot

```
tips = sns.load_dataset("tips")
ax = sns.barplot(x="day",
                 y="total_bill",
                 data=tips)
```

Histogram

```
Penguins = sns.load_dataset("Penguins")
sns.histplot(data = penguins,
             x= "flipper_length_mm")
```

Boxplot

```
tips = sns.load_dataset("tips")
ax = sns.boxplot(x= tips["total_bill"])
```

Scatterplot

```
tips = sns.load_dataset("tips")
Sns.scatterplot(data= tips,
                x = "total_bill",
                y = "tip")
```

Figure aesthetics

```
sns.set_style('darkgrid') # styles
sns.set_palette('husl', 3) # Palettes
sns.color_palette('husl') # colors
```

Fontsize of the axes title, x and y labels, tick labels and legend:

```
plt.rc('axes', titlesize = 18)
plt.rc('axes', labelsize = 14)
plt.rc('xtick', labelsize = 13)
plt.rc('ytick', labelsize = 13)
plt.rc('legend', fontsize = 13)
plt.rc('font', size = 13)
```