

Assignment 4.2: Scatter Plots, Bubble Charts & Density Plot/Map.**Date: 2/3/2023**

```
In [7]: # Import libraries
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import chart_studio.plotly as py
import cufflinks as cf
import seaborn as sns
import plotly.express as px
```

```
In [8]: # Read world population data
dirData = 'ex4-2/'
f_crime = 'crimerates-by-state-2005.csv'

dir_crime = dirData+f_crime
crime = pd.read_csv(dir_crime)
crime_state = crime[crime['state']!='United States']

print(crime_state.head())
```

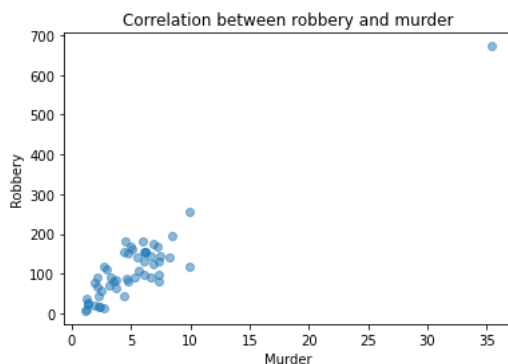
	state	murder	forcible_rape	robbery	aggravated_assault	burglary	\
1	Alabama	8.2	34.3	141.4	247.8	953.8	
2	Alaska	4.8	81.1	80.9	465.1	622.5	
3	Arizona	7.5	33.8	144.4	327.4	948.4	
4	Arkansas	6.7	42.9	91.1	386.8	1084.6	
5	California	6.9	26.0	176.1	317.3	693.3	

	larceny_theft	motor_vehicle_theft	population
1	2650.0	288.3	4545049
2	2599.1	391.0	669488
3	2965.2	924.4	5974834
4	2711.2	262.1	2776221
5	1916.5	712.8	35795255

1. Scatter plot - Python

```
In [9]: # Create a scatter plot showing correlation between murder and robbery

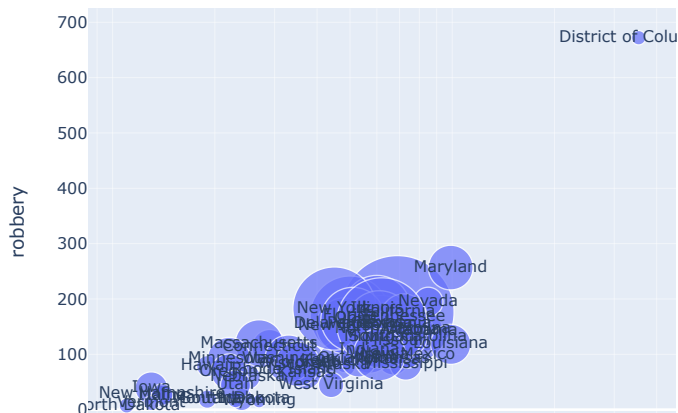
plt.scatter(x=crime['murder'], y=crime['robbery'],alpha=0.5)
plt.title('Correlation between robbery and murder')
plt.xlabel('Murder')
plt.ylabel('Robbery')
plt.show()
```



2. Bubble Chart - Python

```
In [30]: fig = px.scatter(crime_state, x='murder', y='robbery', size='population', text='state',
                        log_x=True, size_max=60,
                        labels={'Murder', 'Robbery'},
                        title='Murder vs Robbery by Population',
                        width=600, height=500)

fig.show()
```



3. Density plot - Python

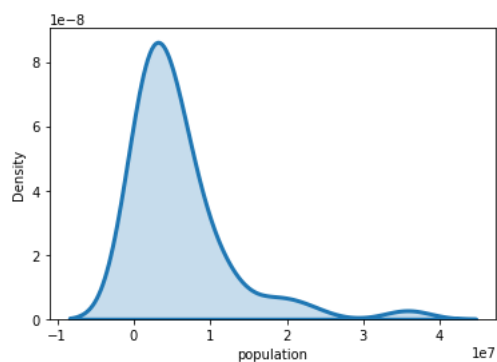
```
In [24]: # Plotting distribution of population

sns.distplot(crime_state['population'], hist=False, kde=True,
             kde_kws={'shade': True, 'linewidth': 3})
```

/Users/anjanibonda/opt/anaconda3/lib/python3.9/site-packages/seaborn/distributions.py:2619: FutureWarning:

`distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `kdeplot` (an axes-level function for kernel density plots).

Out[24]: <AxesSubplot:xlabel='population', ylabel='Density'>



End of code