Exercise 4.2: Scatterplots, Bubble Charts, & Density Plots/Maps

Scott Breitbach

DSC640 - 01/27/2022

Plots Using **Python**

Load Data

In [1]:
```
# Load lib raries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [2]:
```
# Load data
crimeDF = pd.read_csv("crimerates-by-state-2005.csv")
birthDF = pd.read_csv("birth-rate.csv")

# Remove Washington, DC due to outliers
crimeDF = crimeDF[crimeDF.state != 'District of Columbia']
# And remove United States due to averages
crimeDF = crimeDF[crimeDF.state != 'United States']

# Set color to Bellevue purple
color = "#4f3674"
```

Scatterplot

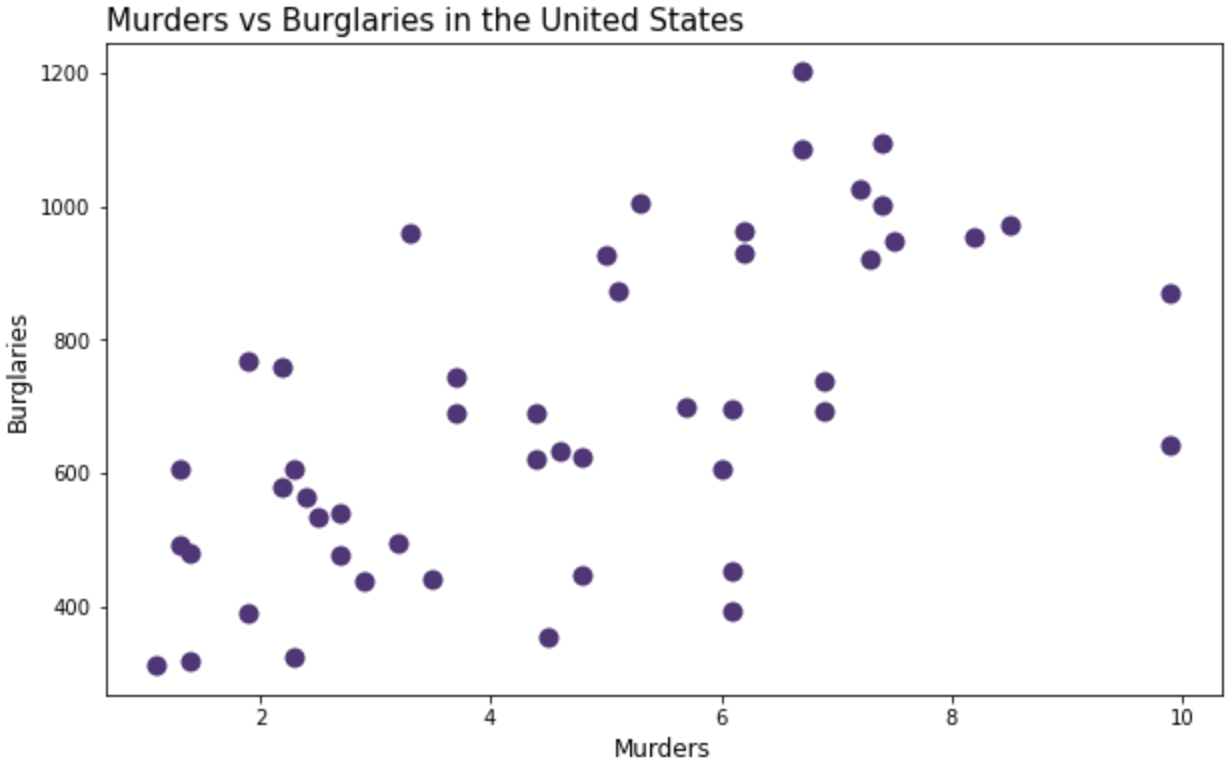
In [3]:

```
# Initialize the matplotlib figure
f, ax = plt.subplots(figsize=(10, 6))

# Create a Scatterplot
plt.plot('murder', 'burglary', data=crimeDF, linestyle='none', markersize=10,
         markerfacecolor=color, marker='o', markeredgecolor="none")

# Add chart title and labels
plt.title("Murders vs Burglaries in the United States", fontsize = 15, loc = 'left')
plt.xlabel("Murders", fontsize = 12)
plt.ylabel("Burglaries", fontsize = 12)

plt.show()
```



Bubble Chart

In [4]:

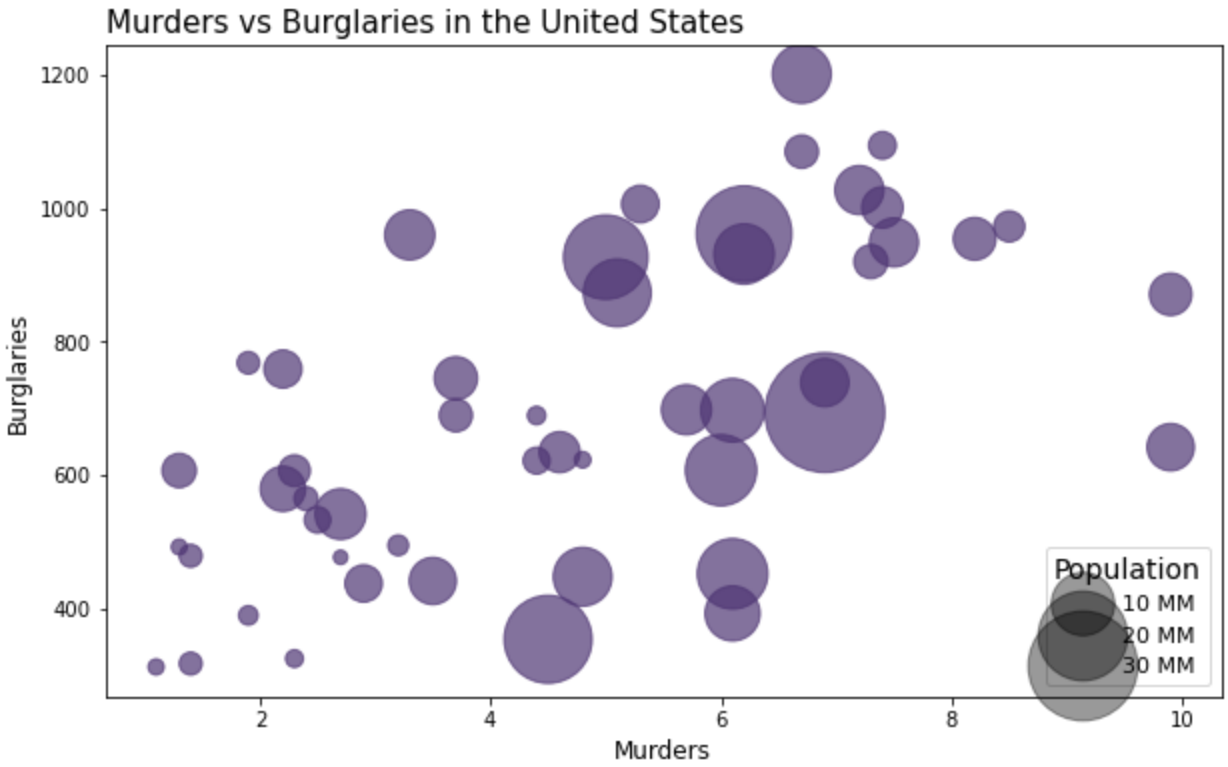
```
# Initialize the matplotlib figure
f, ax = plt.subplots(figsize=(10, 6))

# Make a bubble plot
sc = plt.scatter('murder', 'burglary', s=crimeDF.population*.0001, data=crimeDF,
                 alpha=0.7, color=color, label='Population')

# Add chart title and labels
plt.title("Murders vs Burglaries in the United States", fontsize = 15, loc = 'left')
plt.xlabel("Murders", fontsize = 12)
plt.ylabel("Burglaries", fontsize = 12)

# Set legend for population sizes
handles, labels = sc.legend_elements(prop="sizes", alpha=0.4, num=3)
labels = ["10 MM", "20 MM", "30 MM"]
legend = ax.legend(handles, labels, loc="lower right", title="Population",
                   fontsize=11, title_fontsize=14)

plt.show()
```



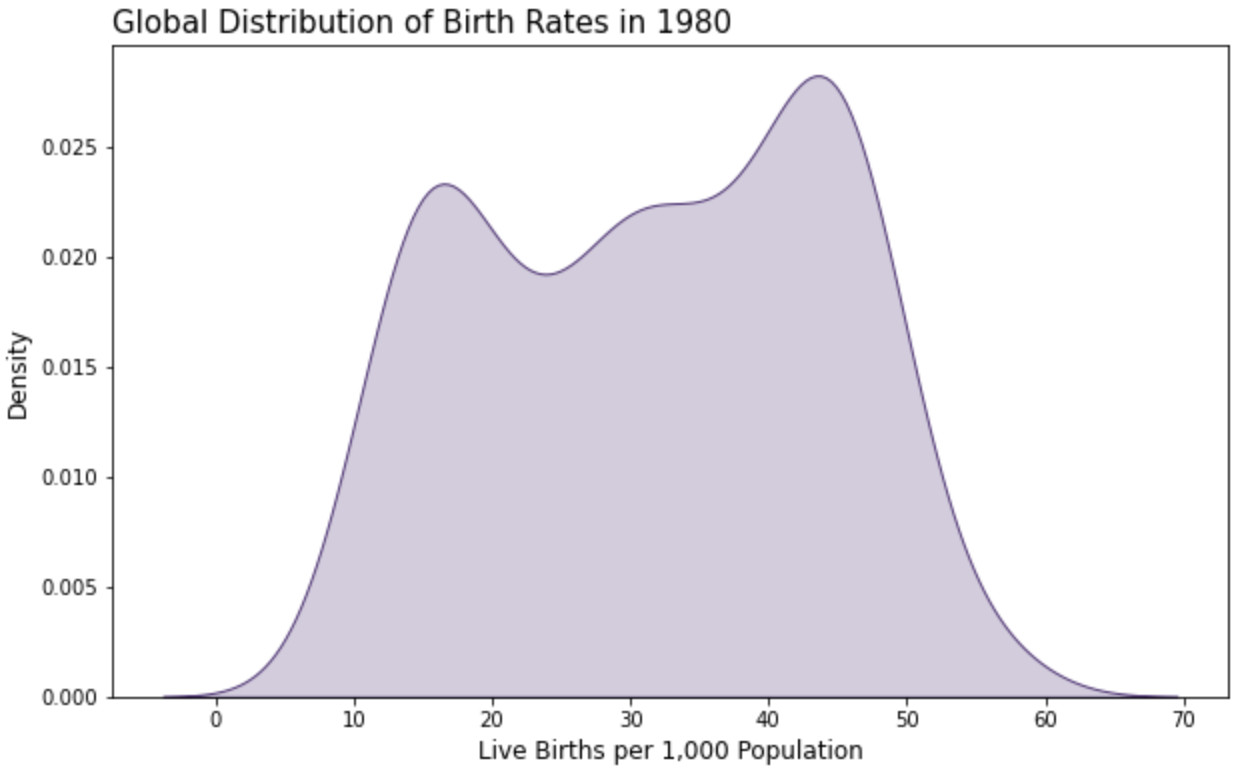
Density Plot

In [5]:
```
# Initialize the matplotlib figure
f, ax = plt.subplots(figsize=(10, 6))

# Make a density plot
sns.kdeplot(birthDF['1980'], color=color, fill=color)

# Add chart title and labels
plt.title("Global Distribution of Birth Rates in 1980", fontsize = 15, loc = 'left')
plt.xlabel("Live Births per 1,000 Population", fontsize = 12)
plt.ylabel("Density", fontsize = 12)
```

Out[5]:
```
Text(0, 0.5, 'Density')
```



In [ ]: