Python Plots

In [34]:
```
# Import libraries
import pandas as pd
import matplotlib.pyplot as plt
import squarify
import numpy as np
from scipy.stats import kde
```

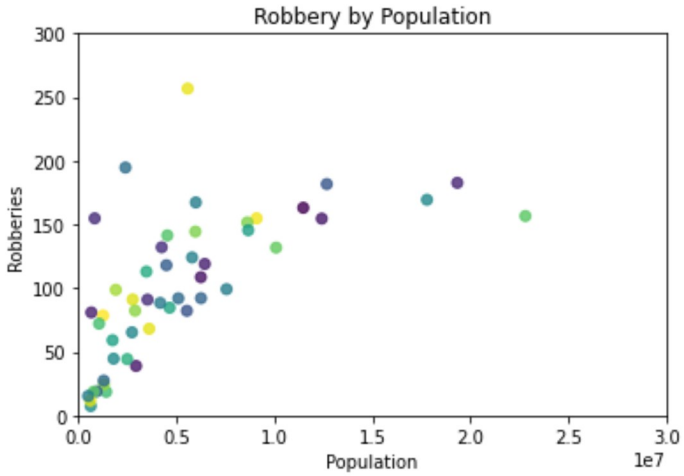
In [3]:
```
# Import Data
crime_df = pd.read_csv('crimerates-by-state-2005.csv')
crime_df.head()
```

Out[3]:

	state	murder	forcible_rape	robbery	aggravated_assault	burglary	larceny_theft	motor_vehicle_theft	p
0	United States	5.6	31.7	140.7	291.1	726.7	2286.3	416.7	2
1	Alabama	8.2	34.3	141.4	247.8	953.8	2650.0	288.3	
2	Alaska	4.8	81.1	80.9	465.1	622.5	2599.1	391.0	
3	Arizona	7.5	33.8	144.4	327.4	948.4	2965.2	924.4	
4	Arkansas	6.7	42.9	91.1	386.8	1084.6	2711.2	262.1	

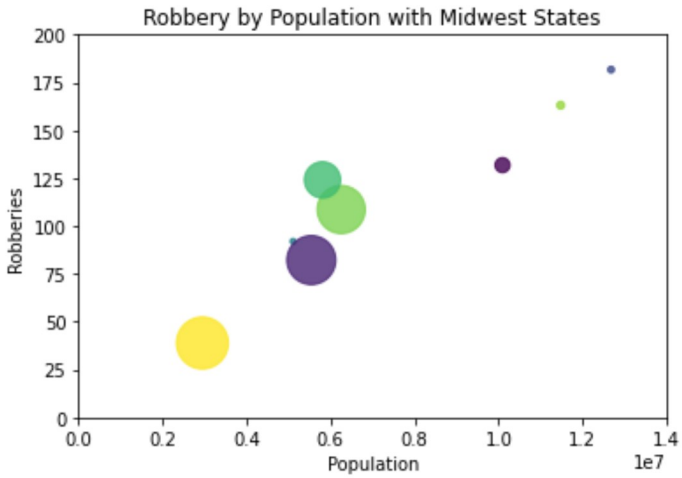
Scatter Plot

In [26]:
```
colors = np.random.rand(52)
plt.scatter(crime_df['population'], crime_df['robbery'],c=colors, alpha=0.8, label=
crime_df['state'])
plt.xlim([0,30000000])
plt.ylim([0,300])
plt.title("Robbery by Population")
plt.xlabel("Population")
plt.ylabel('Robberies')
plt.show()
```



Bubble Plot

In [35]:
```
midwest_states = ['Illinois', 'Indiana', 'Iowa', 'Michigan', 'Minnesota', 'Missouri
', 'Ohio', 'Wisconsin']
midwest_crime_df = crime_df[crime_df['state'].isin(midwest_states)]
area = (30 * np.random.rand(8)) ** 2
colors = np.random.rand(8)
plt.scatter(midwest_crime_df['population'], midwest_crime_df['robbery'],c=colors, a
lpha=0.8, s=area, label=crime_df['state'])
plt.xlim([0,14000000])
plt.ylim([0,200])
plt.title("Robbery by Population with Midwest States")
plt.xlabel("Population")
plt.ylabel('Robberies')
plt.show()
```



Density Plot

In [41]:
```
nbins = 300
k = kde.gaussian_kde([crime_df['population'], crime_df['robbery']])
xi, yi = np.mgrid[crime_df['population'].min():crime_df['population'].max():nbins*1
j, crime_df['robbery'].min():crime_df['robbery'].max():nbins*1j]
zi = k(np.vstack([xi.flatten(), yi.flatten()]))
plt.pcolormesh(xi, yi, zi.reshape(xi.shape))
plt.xlim([0,30000000])
plt.ylim([0,300])
plt.title("Robbery by Population")
plt.xlabel("Population")
plt.ylabel('Robberies')
plt.show()
```