Python Script

Assignment 3.2: Tree Maps, Area Chart and Stacked Area Chart

DSC640

Taniya Adhikari

In [1]:
```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
from numerize import numerize
import matplotlib.ticker as ticker
from matplotlib.ticker import FuncFormatter
import plotly.express as px
```

In [2]:
```
df = pd.read_csv('unemployement-rate-1948-2010.csv')
```

In [3]:
```
df['Period'].replace({"M01": "Jan", "M02": "Feb","M03": "Mar","M04": "Apr","M05": "May","M06": "Jun",
                      "M07": "Jul", "M08": "Aug","M09": "Sep","M10": "Oct","M11": "Nov","M12": "Dec"}, inplace=True)
```

In [4]:
```
df.describe()
```

Out[4]:

	Year	Value
count	746.000000	746.000000
mean	1978.584450	5.666488
std	17.957638	1.567909
min	1948.000000	2.500000
25%	1963.000000	4.525000
50%	1979.000000	5.500000
75%	1994.000000	6.600000
max	2010.000000	10.800000

In [5]:
```
df[df['Year'] ==2010]
```

Out[5]:

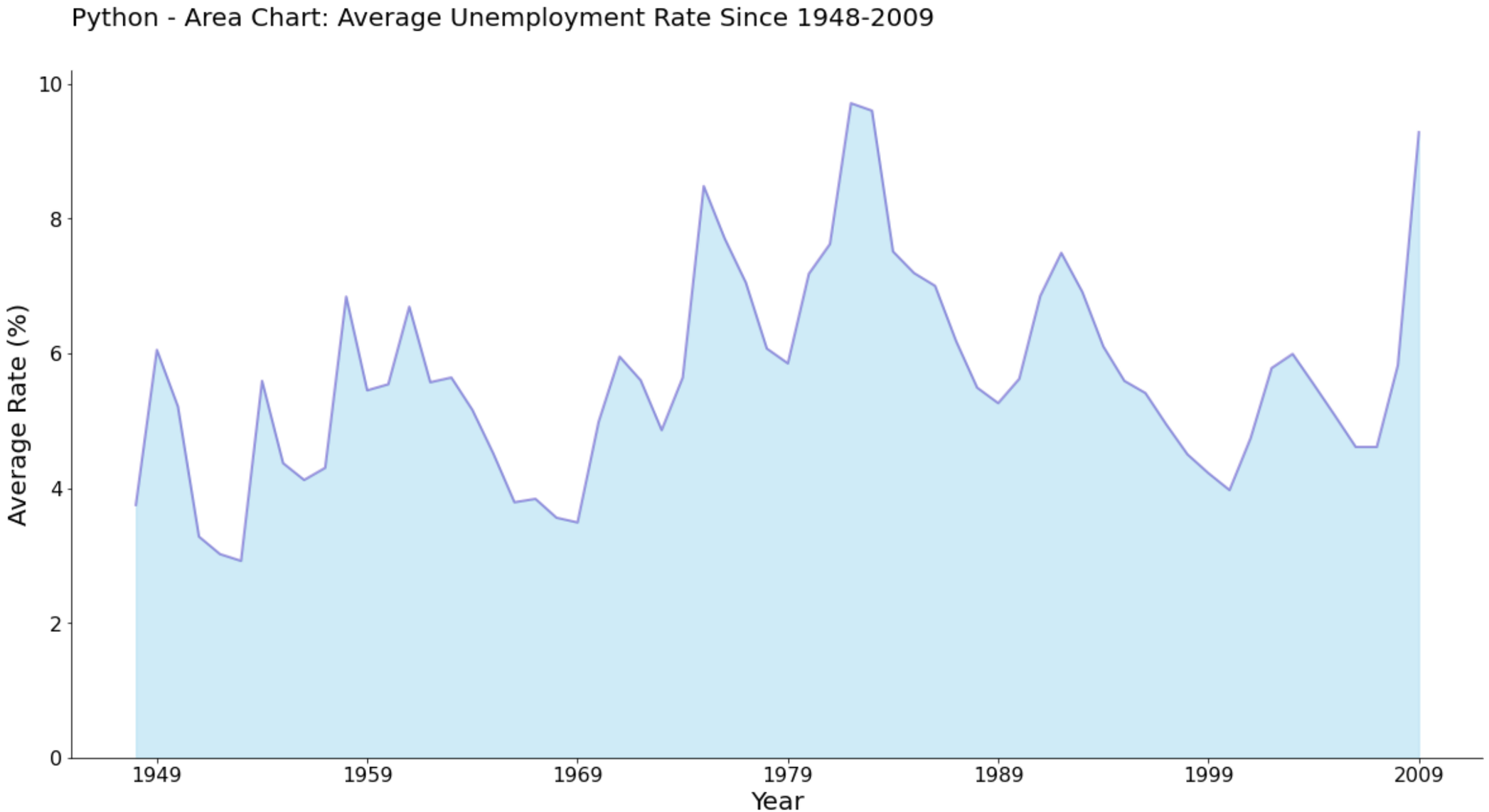
	Series id	Year	Period	Value
744	LNS14000000	2010	Jan	9.7
745	LNS14000000	2010	Feb	9.7

In [6]:
```
df.drop([744, 745], inplace=True)
```

In [7]:
```
avg = df.groupby('Year').mean()
avg.reset_index(level=0, inplace=True)
avg = avg.round({'Value': 2})
```

In [8]:
```
plt.rcParams['figure.figsize'] = [20,10]
fig, ax = plt.subplots()

plt.fill_between(avg['Year'], avg['Value'], color="skyblue", alpha=0.4)
plt.plot(avg['Year'], avg['Value'], color="Slateblue", alpha=0.6, linewidth=2)

plt.tick_params(labelsize=16)
plt.xticks([1949, 1959, 1969, 1979, 1989, 1999, 2009])
plt.suptitle("Python - Area Chart: Average Unemployment Rate Since 1948-2009",
```

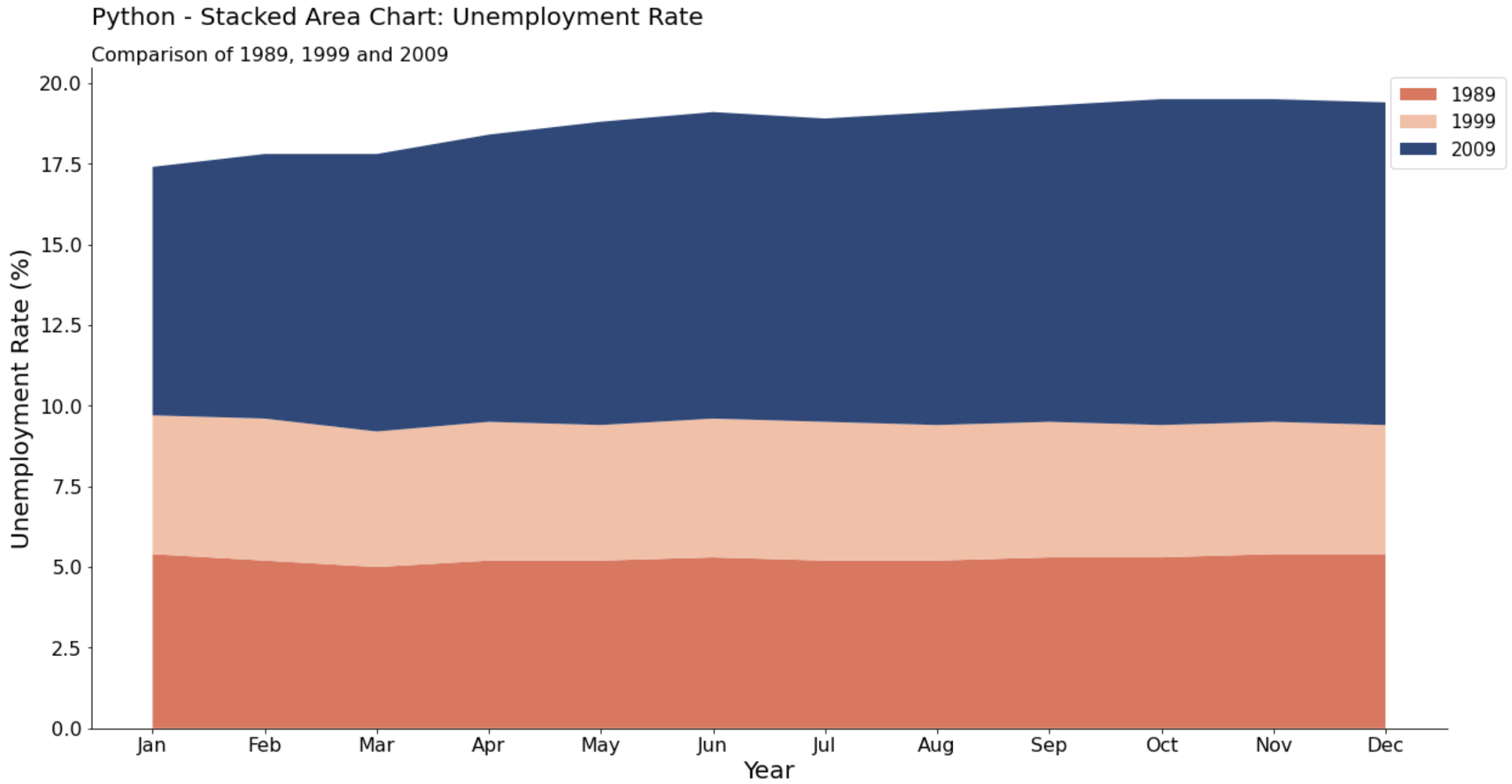
```
            size=20, x=.125, y=.95,horizontalalignment='left', verticalalignment='top')

plt.xlabel('Year', size=20)
plt.ylabel('Average Rate (%)', size=20)
plt.ylim(bottom=0)

right_side = ax.spines["right"]
right_side.set_visible(False)
top = ax.spines["top"]
top.set_visible(False)

plt.show()
```



In [9]:

```
y_1999 = df[df['Year']==1999]
y_2009 = df[df['Year']==2009]
y_1989 = df[df['Year']==1989]
```

In [25]:
```
plt.rcParams['figure.figsize'] = [20,10]
fig, ax = plt.subplots()

plt.stackplot(y_1989['Period'], y_1989['Value'],y_1999['Value'], y_2009['Value'], labels=['1989','1999','2009'],
              colors =['#D87860', '#f0c0a8', '#304878'])

plt.tick_params(labelsize=16)
plt.suptitle("Python - Stacked Area Chart: Unemployment Rate",
             size=20, x=.125, y=.95,horizontalalignment='left', verticalalignment='top')

plt.title("Comparison of 1989, 1999 and 2009", size=16, loc='left')
plt.xlabel('Year', size=20)
plt.ylabel('Unemployment Rate (%)', size=20)
plt.ylim(bottom=0)

right_side = ax.spines["right"]
right_side.set_visible(False)
top = ax.spines["top"]
top.set_visible(False)
plt.legend(bbox_to_anchor=(1.05, 1), loc='upper right', borderpad=0.5, fontsize=15)

plt.show()
```



In [11]:

```
#binning method for confidence of fire.
bins = [2.5,4.0,6.0,8.0,11]
labels = ['Between 2.5-4.0%', 'Between 4.1-6.0%', 'Between 6.1-8.0%', 'Above 8.1%']
avg['Threshold'] = pd.cut(avg['Value'], bins=bins, labels=labels)
avg['Threshold'].fillna('Between 2.5-4.0%', inplace=True)
```

In [12]:
```
avg.rename(columns={"Value": "YearlyAverage"}, inplace=True)
```

In [13]:
```
fig = px.treemap(avg,
                 path=['Threshold', 'Year'],
                 values='YearlyAverage',
                 color='YearlyAverage',
                 color_continuous_scale='GnBu')

fig.update_layout(
    uniformtext=dict(minsize=10, mode='show'),
    margin = dict(t=50, l=25, r=25, b=25),
    title_text = 'Python - Tree Map: Average Unemployment Rate Distribution from 1948-2009',
    title_font_family="Arial",
    title_font_size = 22,
    title_font_color="black",
    title_x=0.03,
)
fig.data[0].textinfo = 'label+text+value'
fig.layout.hovermode = False
fig.show()
```

```
C:\Users\bibek\anaconda3\envs\my_env\lib\site-packages\plotly\express\_core.py:1637: FutureWarning: The frame.append method is deprecated and will be removed f
rom pandas in a future version. Use pandas.concat instead.
  df_all_trees = df_all_trees.append(df_tree, ignore_index=True)
C:\Users\bibek\anaconda3\envs\my_env\lib\site-packages\plotly\express\_core.py:1637: FutureWarning: The frame.append method is deprecated and will be removed f
rom pandas in a future version. Use pandas.concat instead.
  df_all_trees = df_all_trees.append(df_tree, ignore_index=True)
```

Python - Tree Map: Average Unemployment Rate Distribution from 1948-2009