Week 7 & 8 Assignment - Python

Name: Kesav Adithya Venkidusamy
Course: DSC640 - Data Presentation and Visualization
Instructor: Catherine Williams

These two weeks we are going to be focused on heat maps, spatial charts, and contour charts and using various tools to create these visualizations. You must consolidate all the charts into ONE document with each chart labeled with the type of chart and technology - for example: Python - Bar Chart. Failure to label and consolidate the charts will resort in points being taken off or a 0 for the assignment.

Sample Datasets (click on the Downloads tab.)

You may also download them directly from this link: Exercise 4.2 Datasets (click the link to download a folder containing the datasets.)

You need to submit:
1 heat map, 1 spatial chart and 1 funnel or violin chart using Tableau or PowerBI
1 heat map, 1 spatial chart and 1 contour chart using Python
1 heat map, 1 spatial chart and 1 contour chart using R

1 heat map, 1 spatial chart and 1 contour chart using Python

In [1]:
```
## Importing libraries required for this exercise
import pandas as pd
import numpy as np
import squarify
import matplotlib.pyplot as plt
%matplotlib inline
import plotly.express as px
import seaborn as sns
import matplotlib
import plotly.graph_objects as go
```

Read Input datasets

In [2]:
```
## Reading the costco data
costco_df = pd.read_csv('costcos-geocoded.csv')
```

```
costco_df.head()
```

Out[2]:

	Address	City	State	Zip Code	Latitude	Longitude
0	1205 N. Memorial Parkway	Huntsville	Alabama	35801-5930	34.743095	-86.600955
1	3650 Galleria Circle	Hoover	Alabama	35244-2346	33.377649	-86.812420
2	8251 Eastchase Parkway	Montgomery	Alabama	36117	32.363889	-86.150884
3	5225 Commercial Boulevard	Juneau	Alaska	99801-7210	58.359200	-134.483000
4	330 West Dimond Blvd	Anchorage	Alaska	99515-1950	61.143266	-149.884217

In [3]:

```
## Reading ppg dataset
ppg2008_df = pd.read_csv('ppg2008.csv')
ppg2008_df.head()
```

Out[3]:

	Name	G	MIN	PTS	FGM	FGA	FGP	FTM	FTA	FTP	...	3PA	3PP	ORB	DRB	TRB	AST	STL	BLK	TO	PF
0	Dwyane Wade	79	38.6	30.2	10.8	22.0	0.491	7.5	9.8	0.765	...	3.5	0.317	1.1	3.9	5.0	7.5	2.2	1.3	3.4	2.3
1	LeBron James	81	37.7	28.4	9.7	19.9	0.489	7.3	9.4	0.780	...	4.7	0.344	1.3	6.3	7.6	7.2	1.7	1.1	3.0	1.7
2	Kobe Bryant	82	36.2	26.8	9.8	20.9	0.467	5.9	6.9	0.856	...	4.1	0.351	1.1	4.1	5.2	4.9	1.5	0.5	2.6	2.3
3	Dirk Nowitzki	81	37.7	25.9	9.6	20.0	0.479	6.0	6.7	0.890	...	2.1	0.359	1.1	7.3	8.4	2.4	0.8	0.8	1.9	2.2
4	Danny Granger	67	36.2	25.8	8.5	19.1	0.447	6.0	6.9	0.878	...	6.7	0.404	0.7	4.4	5.1	2.7	1.0	1.4	2.5	3.1

5 rows × 21 columns

In [4]:

```
ppg2008_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 50 entries, 0 to 49
Data columns (total 21 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   Name    50 non-null     object
 1   G       50 non-null     int64
 2   MIN     50 non-null     float64
 3   PTS     50 non-null     float64
 4   FGM     50 non-null     float64
```

```
 5   FGA     50 non-null     float64
 6   FGP     50 non-null     float64
 7   FTM     50 non-null     float64
 8   FTA     50 non-null     float64
 9   FTP     50 non-null     float64
 10  3PM     50 non-null     float64
 11  3PA     50 non-null     float64
 12  3PP     50 non-null     float64
 13  ORB     50 non-null     float64
 14  DRB     50 non-null     float64
 15  TRB     50 non-null     float64
 16  AST     50 non-null     float64
 17  STL     50 non-null     float64
 18  BLK     50 non-null     float64
 19  TO      50 non-null     float64
 20  PF      50 non-null     float64
dtypes: float64(19), int64(1), object(1)
memory usage: 8.3+ KB
```

1. Python - Heat Map

Plotting Heat Map for ppg test result

In [34]:

```
ppg2008_sub_df = ppg2008_df.copy()
ppg2008_sub_df.set_index('Name', inplace = True)
ppg2008_sub_df.head()
```

Out[34]:

Name	G	MIN	PTS	FGM	FGA	FGP	FTM	FTA	FTP	3PM	3PA	3PP	ORB	DRB	TRB	AST	STL	BLK	TO	PF
Dwyane Wade	79	38.6	30.2	10.8	22.0	0.491	7.5	9.8	0.765	1.1	3.5	0.317	1.1	3.9	5.0	7.5	2.2	1.3	3.4	2.3
LeBron James	81	37.7	28.4	9.7	19.9	0.489	7.3	9.4	0.780	1.6	4.7	0.344	1.3	6.3	7.6	7.2	1.7	1.1	3.0	1.7
Kobe Bryant	82	36.2	26.8	9.8	20.9	0.467	5.9	6.9	0.856	1.4	4.1	0.351	1.1	4.1	5.2	4.9	1.5	0.5	2.6	2.3
Dirk Nowitzki	81	37.7	25.9	9.6	20.0	0.479	6.0	6.7	0.890	0.8	2.1	0.359	1.1	7.3	8.4	2.4	0.8	0.8	1.9	2.2
Danny Granger	67	36.2	25.8	8.5	19.1	0.447	6.0	6.9	0.878	2.7	6.7	0.404	0.7	4.4	5.1	2.7	1.0	1.4	2.5	3.1

In [39]:

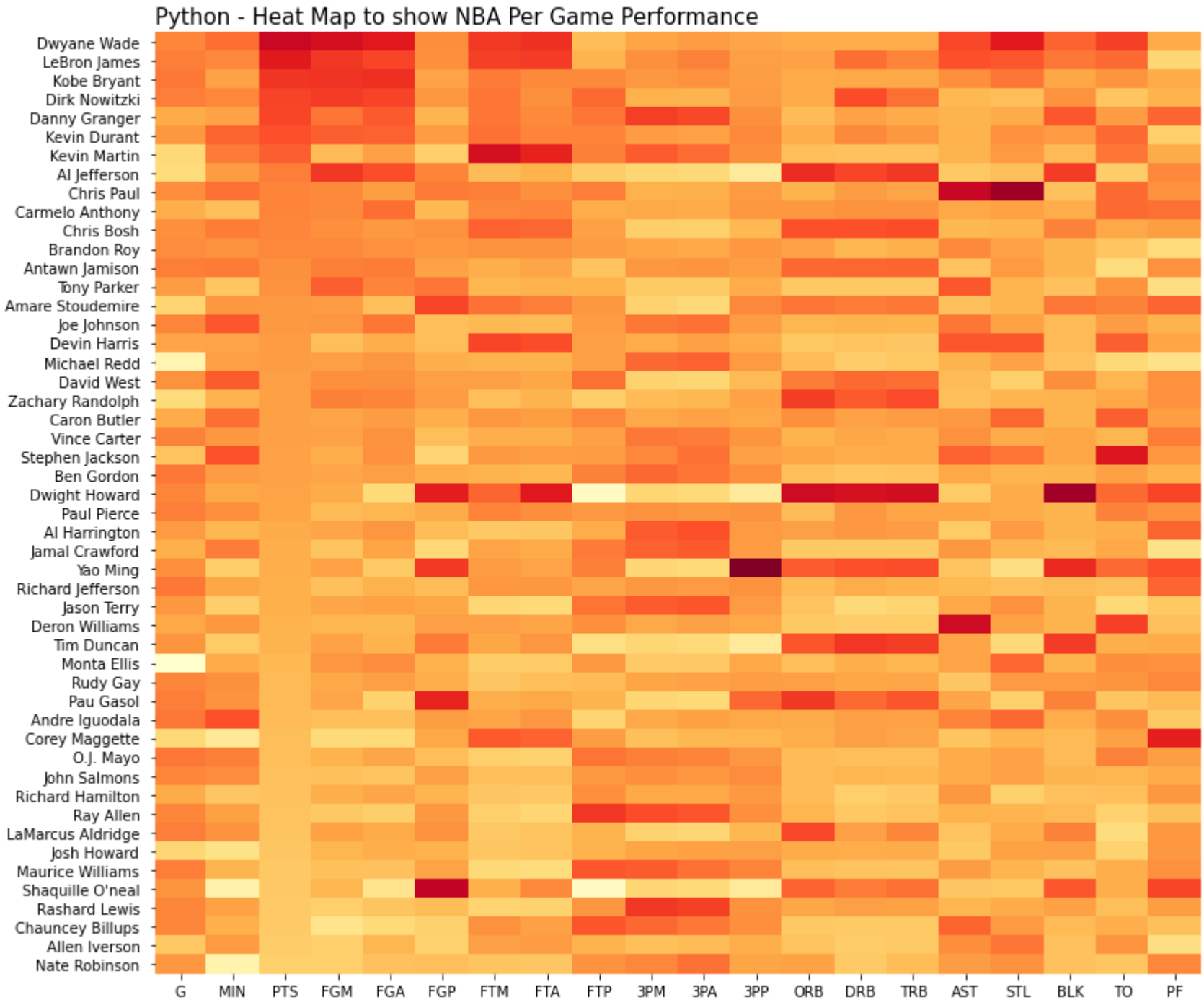
```
# Normalize columns
ppg2008_sub_norm = (ppg2008_sub_df - ppg2008_sub_df.mean())/ppg2008_sub_df.std()
```

In [51]:
```
# Initialize the matplotlib figure
f, ax = plt.subplots(figsize=(13, 12))

# Create heatmap of normalized data
sns.heatmap(ppg2008_sub_norm, cmap='YlOrRd', cbar=False)

# Add chart title and labels
plt.title("Python - Heat Map to show NBA Per Game Performance", fontsize = 15, loc = 'left')
plt.ylabel("")

plt.show()
```


Python - Heat Map to show NBA Per Game Performance
Dwyane Wade
LeBron James
Kobe Bryant
Dirk Nowitzki
Danny Granger
Kevin Durant
Kevin Martin
Al Jefferson
Chris Paul
Carmelo Anthony
Chris Bosh
Brandon Roy
Antawn Jamison
Tony Parker
Amare Stoudemire
Joe Johnson
Devin Harris
Michael Redd
David West
Zachary Randolph
Caron Butler
Vince Carter
Stephen Jackson
Ben Gordon
Dwight Howard
Paul Pierce
Al Harrington
Jamal Crawford
Yao Ming
Richard Jefferson
Jason Terry
Deron Williams
Tim Duncan
Monta Ellis
Rudy Gay
Pau Gasol
Andre Iguodala
Corey Maggette
O.J. Mayo
John Salmons
Richard Hamilton
Ray Allen
LaMarcus Aldridge
Josh Howard
Maurice Williams
Shaquille O'neal
Rashard Lewis
Chauncey Billups
Allen Iverson
Nate Robinson
G
MIN
PTS
FGM
FGA
FGP
FTM
FTA
FTP
3PM
3PA
3PP
ORB
DRB
TRB
AST
STL
BLK
TO
PF

Python - Spatial Chart

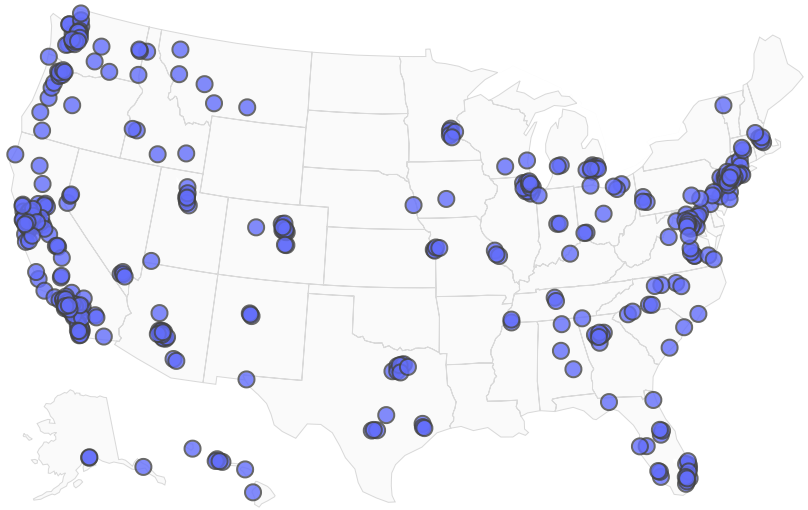
In [5]:

```
fig = go.Figure(data=go.Scattergeo(
        locationmode = 'USA-states',
        lon = costco_df['Longitude'],
        lat = costco_df['Latitude'],
        text = costco_df['Address'],
        mode = 'markers',
        marker = dict(
            size = 8,
            opacity = 0.8,
            reversescale = True,
            autocolorscale = False,
            symbol = 'circle',
            line = dict(
                    width = 1,
                    color = 'rgba(102, 102, 102)'
                    ),
            colorscale = 'Blues',
            cmin = 0,
            colorbar_title = "Costco Store Locations"
            )))

fig.update_layout(
            title = 'Python - Spatial Chart to show Costco Locations',
            geo = dict(
                    scope = 'usa',
                    projection_type = 'albers usa',
                    showland = True,
                    landcolor = "rgb(250, 250, 250)",
                    subunitcolor = "rgb(217, 217, 217)",
                    countrycolor = "rgb(217, 217, 217)",
                    countrywidth = 0.5,
                    subunitwidth = 0.5
                    ),
            )

fig.show()
```

Python - Spatial Chart to show Costco Locations



Python - Contour Chart

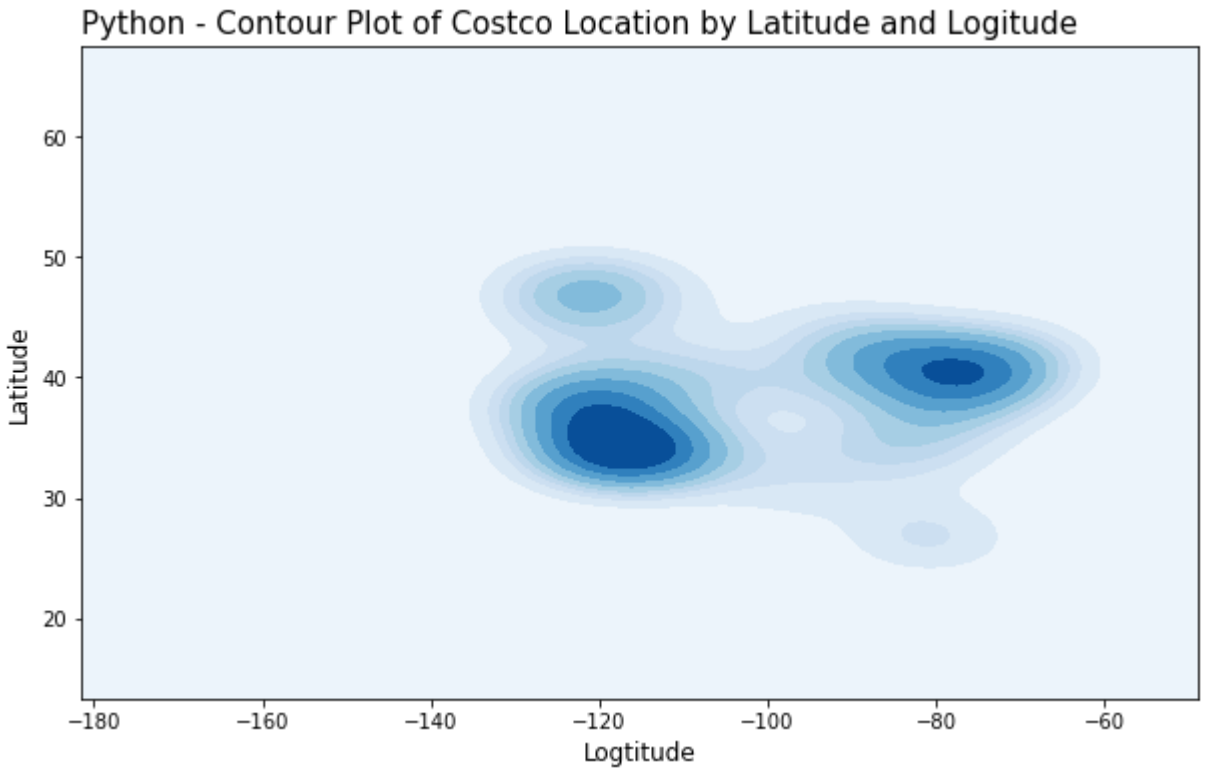
In [66]:

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

sns.kdeplot(x=costco_df.Longitude, y = costco_df.Latitude,
               cmap = "Blues", shade = True, thresh=0)

## Add chart title and labels
plt.title("Python - Contour Plot of Costco Location by Latitude and Logitude",
            fontsize = 15,loc = 'left')
plt.xlabel('Logtitude',fontsize =12)
plt.ylabel('Latitude', fontsize = 12)

plt.show()
```


Python - Contour Plot of Costco Location by Latitude and Logitude
Latitude
60
50
40
30
20
−180
−160
−140
−120
−100
−80
−60
Logtitude

In [ ]: