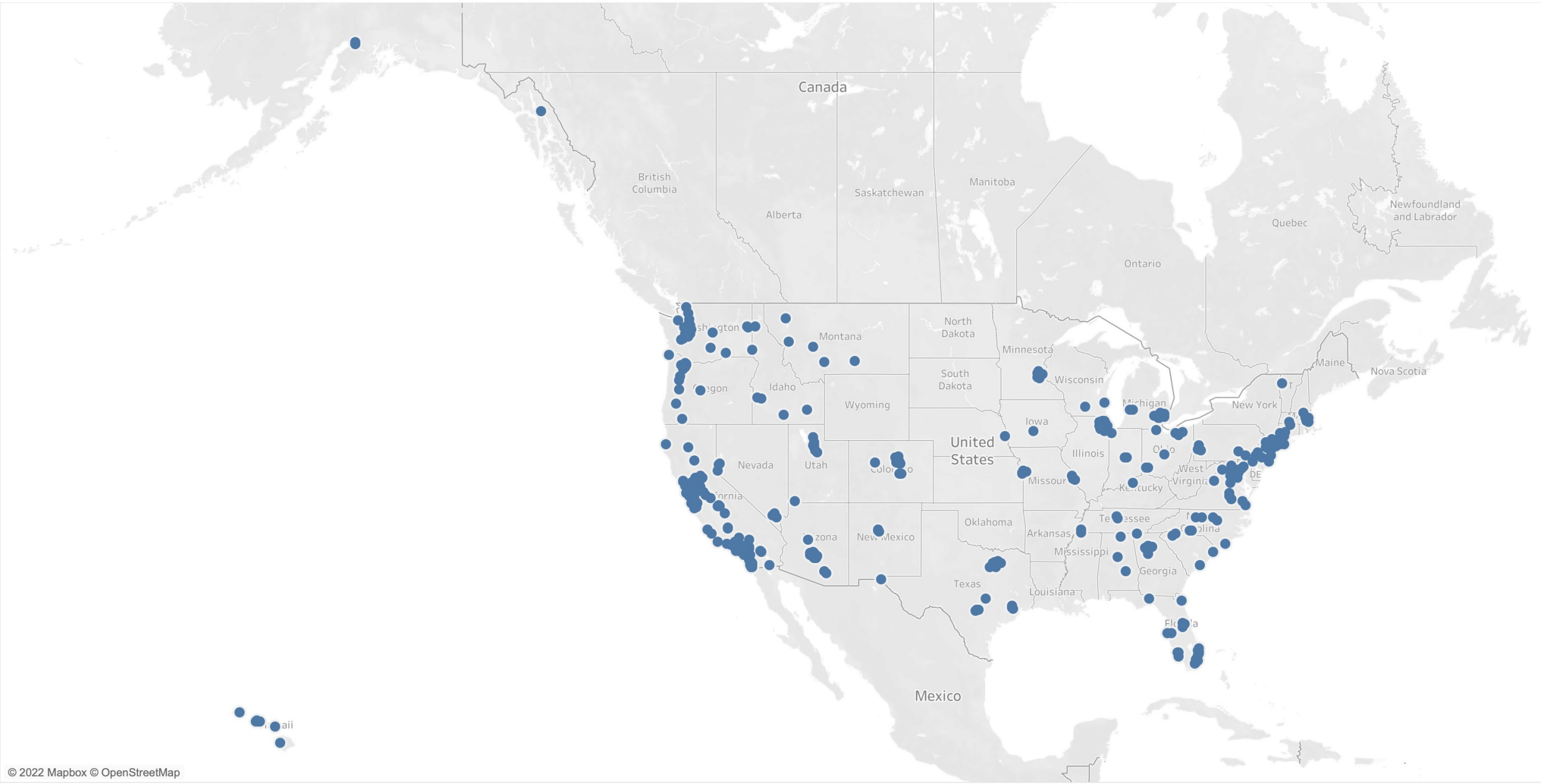
Week_09_10

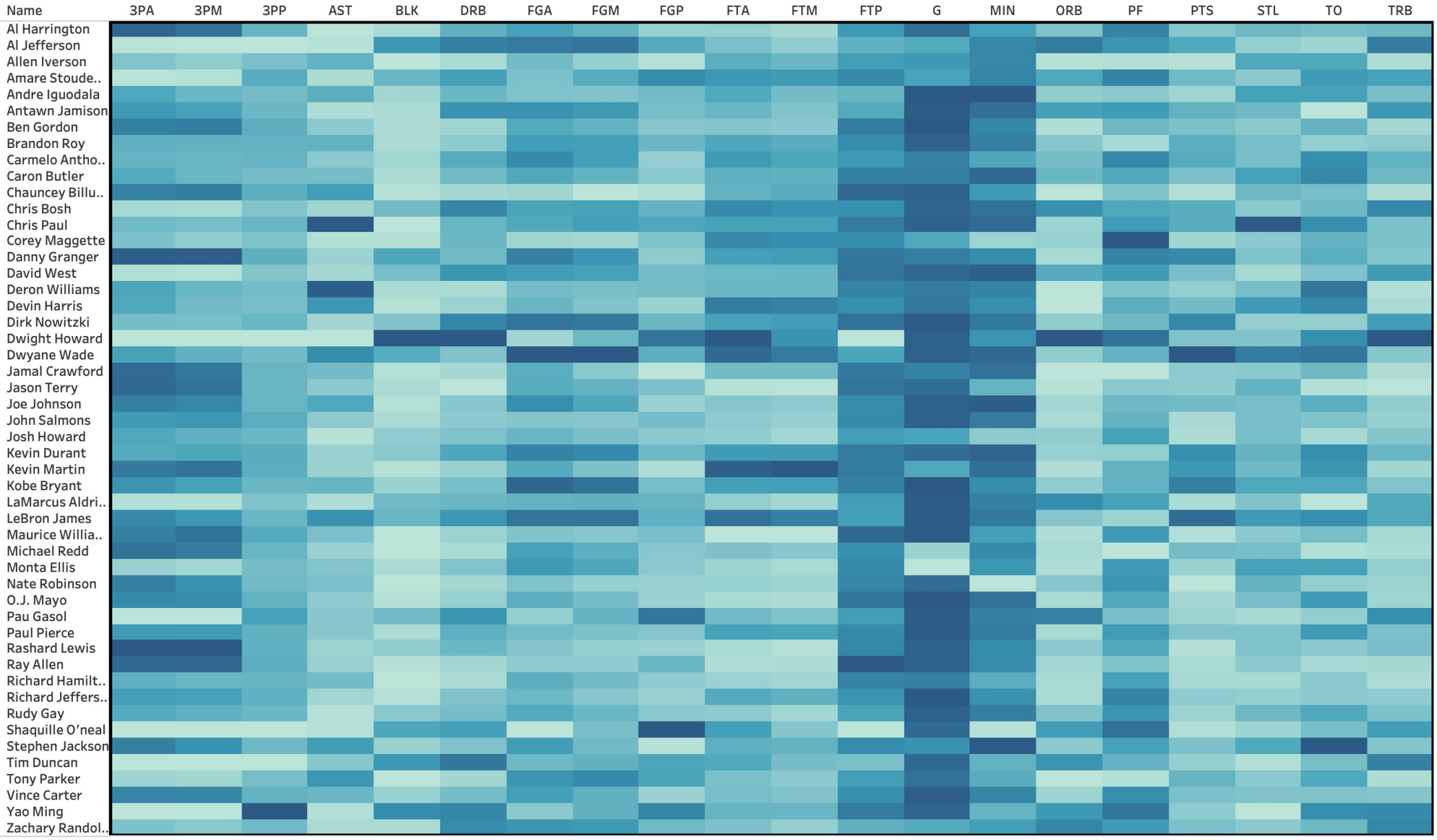
File created on: 2/13/22 1:26:33 PM EST

Spatial Chart



Map based on Longitude and Latitude.

Heatmap



TRB
2.40 13.80

FGM
5.200 10.800

TO
1.500 3.900

FGA
11.20 22.00

STL
0.400 2.800

DRB
1.900 9.600

PTS
17.20 30.20

BLK
0.100 2.900

PF
1.400 3.800

AST
1.400 11.000

ORB
0.400 4.300

3PP
0.000 1.000

MIN
29.900 39.800

3PM
0.000 2.800

G
25 82

3PA
0.000 7.000

FTP
0.5940 0.9520

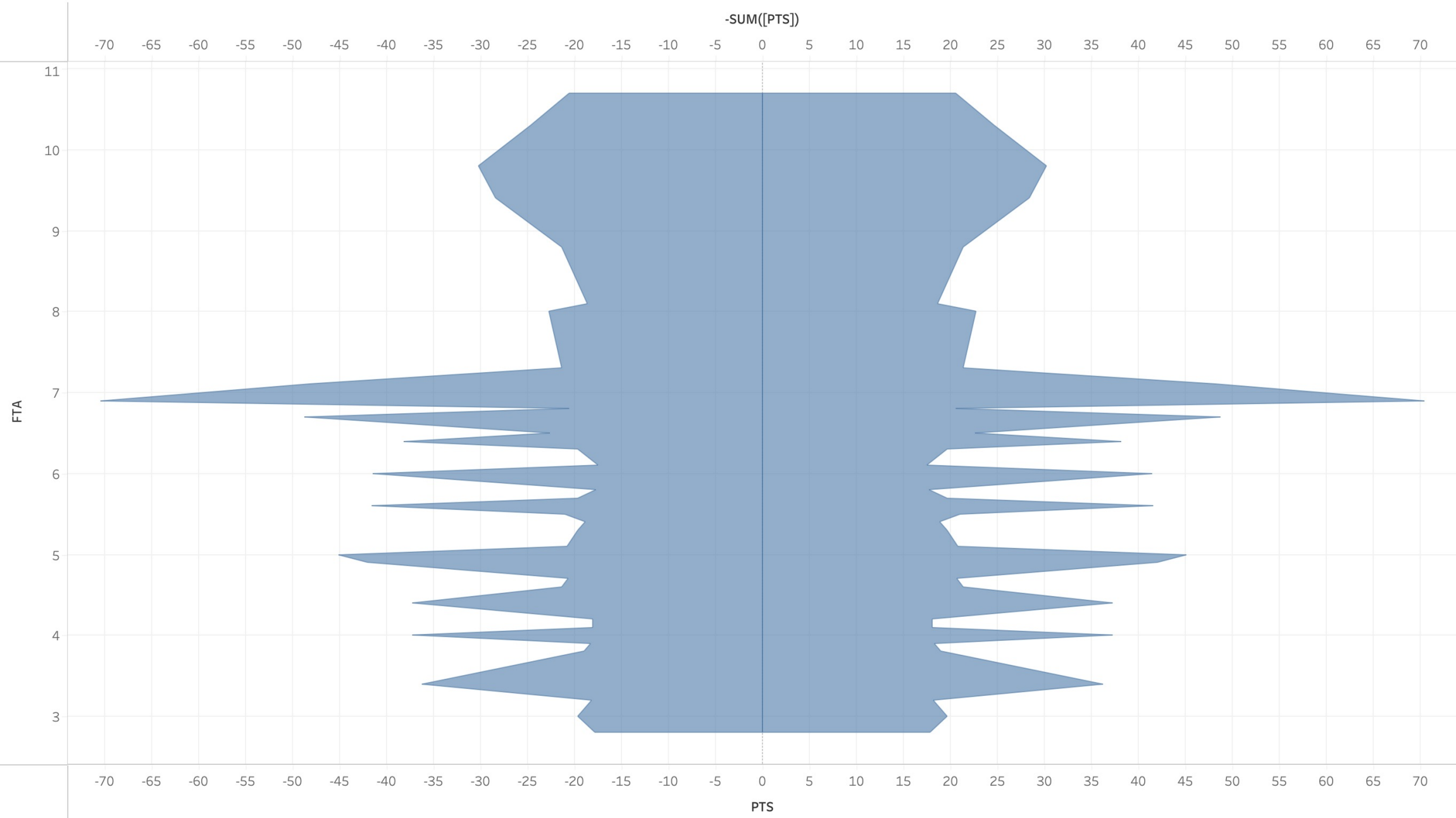
FTM
2.600 9.000

FTA
2.800 10.700

FGP
0.4100 0.6090

3PA, 3PM, 3PP, AST, BLK, DRB, FGA, FGM, FGP, FTA, FTM, FTP, G, MIN, ORB, PF, PTS, STL, TO and TRB (color) broken down by Name.

Violin



The plots of PTS and -SUM([PTS]) for FTA. Color shows details about PTS and -SUM([PTS]).

Week 9&10 - Exercise 5.2

Name: Madhukar Ayachit

Date: 09 Feb 2022

Class: DSC-640

Assignment: scatterplots, bubble charts, and density plots/maps

In [36]:
```
## import required libraries
import pandas as pd
import numpy as np

import matplotlib.pyplot as plt
import matplotlib
import plotly.graph_objects as go
import plotly.express as px
import seaborn as sns

## for spatial chart
#import geopandas as gpd
#import geoplot as gplt
#import geoplot.crs as gcrs

import folium as flm


import warnings
warnings.filterwarnings("ignore")

## pip install geopandas
## pip install plotly
## conda install -c conda-forge geoplot   via command power shell in anacon
```

In [37]:

```
## read the data from csv file into a dataframe

df1 = pd.read_csv('Data/ex5-2/ppg2008.csv', header = 0)
df1_temp = df1

df1.head()
```

Out[37]:

	Name	G	MIN	PTS	FGM	FGA	FGP	FTM	FTA	FTP	...	3PA	3PP	ORB	DRB	TRB
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
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
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
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
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

5 rows × 21 columns

In [38]:

```
## read the data from csv file into a dataframe

df2 = pd.read_csv('Data/ex5-2/costcos-geocoded.csv')

df2.head()
```

Out[38]:

	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 [39]:
```
## For Heatmap data - we will drop Name column and use the new temp datafra

df1_temp.drop(columns = ['Name  '], inplace = True)

df1_temp.head()
```

Out[39]:

	G	MIN	PTS	FGM	FGA	FGP	FTM	FTA	FTP	3PM	3PA	3PP	ORB	DRB	TRB	AST
0	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
1	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
2	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
3	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
4	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

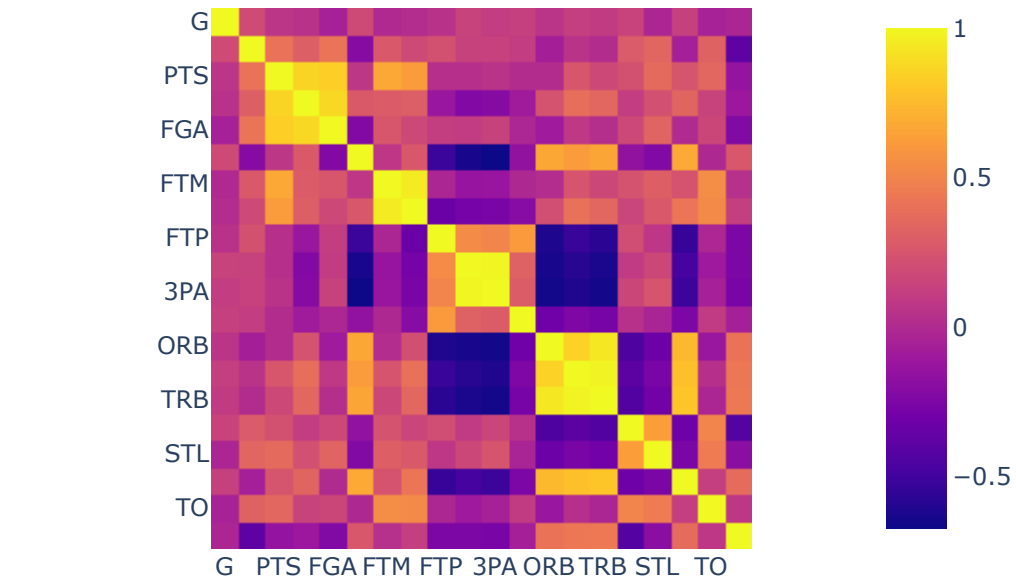
Heatmap

In [40]:
```
df1_temp.corr()
```

Out[40]:

	G	MIN	PTS	FGM	FGA	FGP	FTM	FTA	F
G	1.000000	0.186866	0.063099	0.039922	-0.059581	0.180875	-0.011056	0.009983	0.0394
MIN	0.186866	1.000000	0.403883	0.297420	0.413101	-0.218011	0.265562	0.181878	0.2238
PTS	0.063099	0.403883	1.000000	0.852261	0.828392	0.067161	0.667786	0.611409	0.0304
FGM	0.039922	0.297420	0.852261	1.000000	0.871049	0.268118	0.281119	0.292160	-0.1283
FGA	-0.059581	0.413101	0.828392	0.871049	1.000000	-0.233212	0.249641	0.173999	0.1134
FGP	0.180875	-0.218011	0.067161	0.268118	-0.233212	1.000000	0.069158	0.254235	-0.5240
FTM	-0.011056	0.265562	0.667786	0.281119	0.249641	0.069158	1.000000	0.946365	-0.0325
FTA	0.009983	0.181878	0.611409	0.292160	0.173999	0.254235	0.946365	1.000000	-0.3343
FTP	0.039469	0.223845	0.030490	-0.128330	0.113445	-0.524097	-0.032534	-0.334383	1.0000
3PM	0.138762	0.134413	0.028281	-0.233191	0.097139	-0.637242	-0.144279	-0.285762	0.5297
3PA	0.106132	0.130734	0.044990	-0.215467	0.135048	-0.673994	-0.130957	-0.267314	0.4985
3PP	0.124935	0.106795	0.007157	-0.090980	-0.023221	-0.161036	-0.018012	-0.208422	0.6041
ORB	0.054876	-0.073281	0.007552	0.230899	-0.096734	0.662932	0.011582	0.210608	-0.6151
DRB	0.115330	0.049165	0.249159	0.380277	0.076602	0.614529	0.240176	0.396417	-0.5308
TRB	0.095680	0.005783	0.173261	0.344299	0.020818	0.654599	0.165459	0.342049	-0.5797
AST	0.142441	0.275151	0.220795	0.106198	0.172407	-0.160528	0.231774	0.155229	0.2005
STL	-0.033838	0.331295	0.361697	0.217720	0.327577	-0.228928	0.299572	0.261716	0.0672
BLK	0.125587	-0.066443	0.236802	0.330224	-0.000694	0.675726	0.234402	0.421167	-0.5412
TO	-0.051218	0.316904	0.343304	0.140205	0.157998	-0.016609	0.539411	0.524913	-0.0241
PF	-0.028372	-0.387609	-0.150233	-0.117188	-0.235652	0.258301	0.032845	0.115829	-0.2545

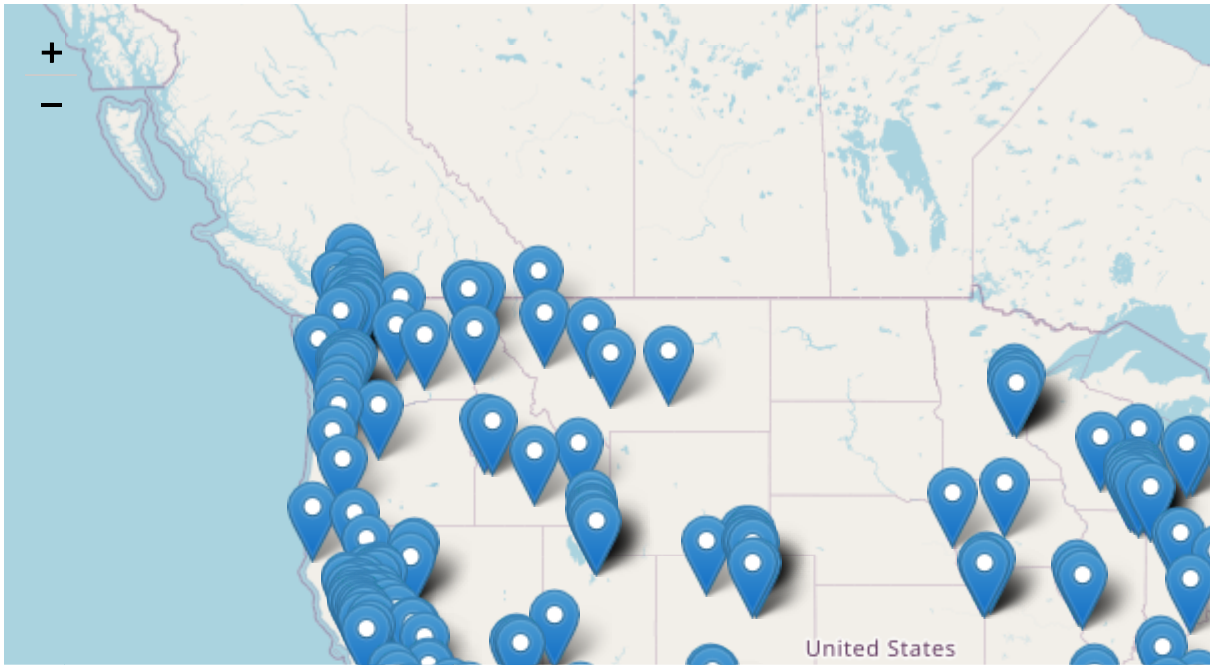
In [41]:
```
fig = px.imshow(df1_temp.corr(), title = "Heatmap (correlations)")

fig.update_layout(
    autosize=False,
    width=550,
    height=450)

fig.show()
```



Spatial chart

In [46]:
```
## start with location for united states

m = flm.Map(location = [40, -95],
            zoom_start = 4)

## iterate through the Costco Locations Dataframe and use Latitude and Logi

for idx, row in df2.iterrows():
    flm.Marker([row.Latitude, row.Longitude]).add_to(m)
## display map with Costco Locations

m
```

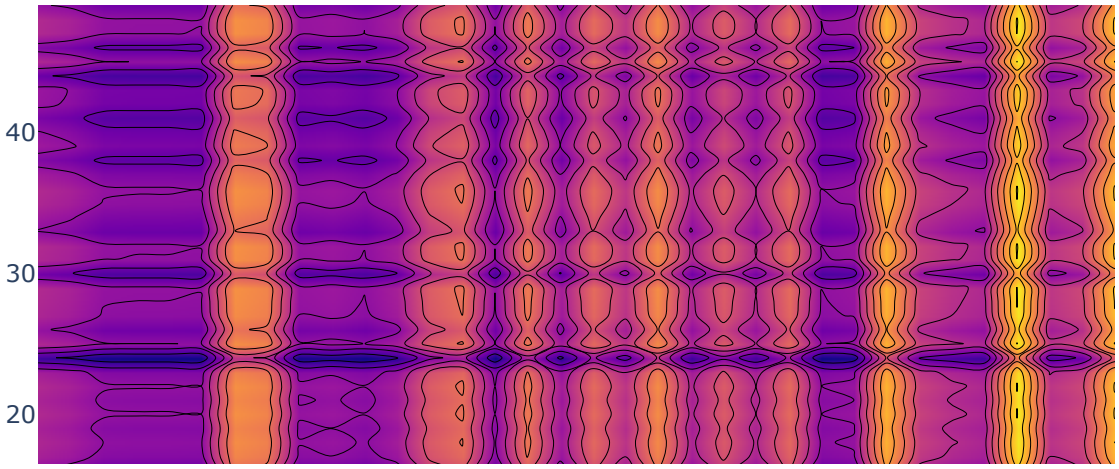
Out[46]:



Contour chart

In [43]:
```
## plot Contour chart

## create 2D array of data input
[x, y] = np.meshgrid(df1['FTM'], df1['FTA'])

##z = x**2 + y**2

z = np.cos(x / 2) + np.sin(y / 4)


## convert the numpy array to a list, so that it is compatible with plotly

z = z.tolist()
```

In [44]:
```
#import plotly.graph_objects as go

fig = go.Figure(data= go.Contour(
    z = z,
        # heatmap gradient coloring is applied between each contour level
        contours_coloring = 'heatmap' # can also be 'lines', or 'none
    )
)

fig.update_layout(
    title="Python Contour Plot"
)

fig.show()
```



In [ ]:

tree maps, area charts, and stacked area charts

Code

Hide

```
# setting current working diirectory
setwd("/Users/madhukarayachit/DSC640")
```

Hide

```
#Load in libraries

#Load libraries
library(readxl)
library(ggplot2)
library(tidyverse)
library(plotly)
library(dbplyr)
library(scales)
library(plotly)
library(maps)
```

Hide

```
# load data


costco <- read.csv("Data/ex5-2/costcos-geocoded.csv")
ppg <- read.csv("Data/ex5-2/ppg2008.csv")

nba
```

MIN <dbl>	PTS <dbl>	FGM <dbl>	FGA <dbl>	FGP <dbl>	FTM <dbl>	FTA <dbl>	FTP <dbl>	X3PM <dbl>	X3PA <dbl>
29.9	17.2	6.1	13.9	0.437	3.4	4.0	0.841	1.7	5.2
36.7	17.5	6.1	14.6	0.417	4.8	6.1	0.781	0.5	1.7
36.2	17.7	6.1	13.8	0.439	2.8	3.4	0.836	2.8	7.0
35.3	17.7	5.2	12.4	0.418	5.3	5.8	0.913	2.1	5.0
35.0	17.8	6.5	13.9	0.467	2.6	2.8	0.912	2.3	5.2
30.1	17.8	6.8	11.2	0.609	4.1	6.9	0.595	0.0	0.0
31.9	18.0	6.8	15.1	0.451	3.3	4.2	0.782	1.1	3.2
37.1	18.1	7.4	15.3	0.484	3.2	4.1	0.781	0.1	0.3

MIN <dbl>	PTS <dbl>	FGM <dbl>	FGA <dbl>	FGP <dbl>	FTM <dbl>	FTA <dbl>	FTP <dbl>	X3PM <dbl>	X3PA <dbl>
36.3	18.2	6.3	13.2	0.480	3.0	3.2	0.952	2.5	6.2
37.5	18.3	6.5	13.8	0.472	3.6	4.4	0.830	1.6	3.8

1-10 of 50 rows | 1-10 of 18 columns Previous **1** 2 3 4 5 Next

Hide

```
costco
```

Address <chr>	City <chr>	State <chr>	Zip.Code <chr>	Latitud <db
1205 N. Memorial Parkway	Huntsville	Alabama	35801-5930	34.743
3650 Galleria Circle	Hoover	Alabama	35244-2346	33.377
8251 Eastchase Parkway	Montgomery	Alabama	36117	32.363
5225 Commercial Boulevard	Juneau	Alaska	99801-7210	58.359
330 West Dimond Blvd	Anchorage	Alaska	99515-1950	61.143
4125 DeBarr Road	Anchorage	Alaska	99508-3115	61.210
3911 Highway 69	Prescott	Arizona	86301-6717	34.548
3901 West Costco Drive	Tucson	Arizona	85741-2864	32.326
6255 East Grant Road	Tucson	Arizona	85712-5834	32.252
17550 N. 79th Ave.	Glendale	Arizona	85308-8711	33.643

1-10 of 417 rows Previous **1** 2 3 4 5 6 … 42 Next

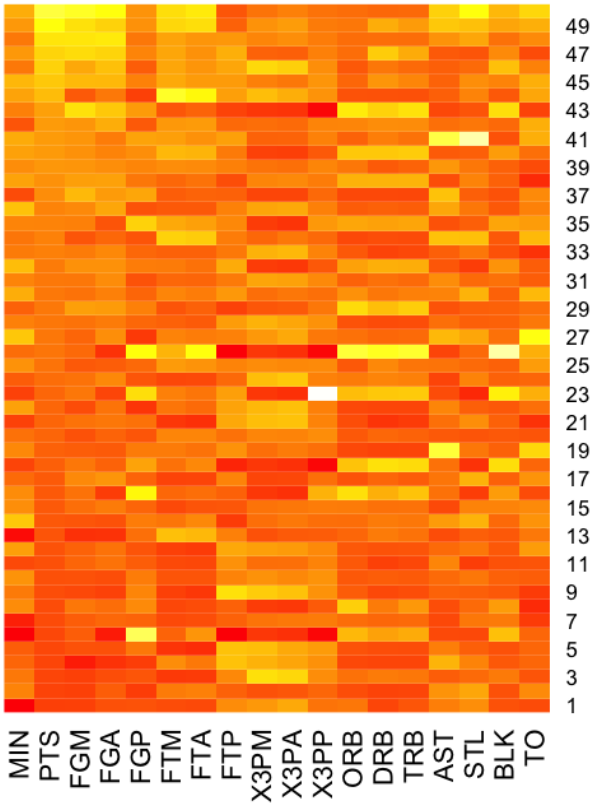
Hide

```
NA
```

Heat map

Hide

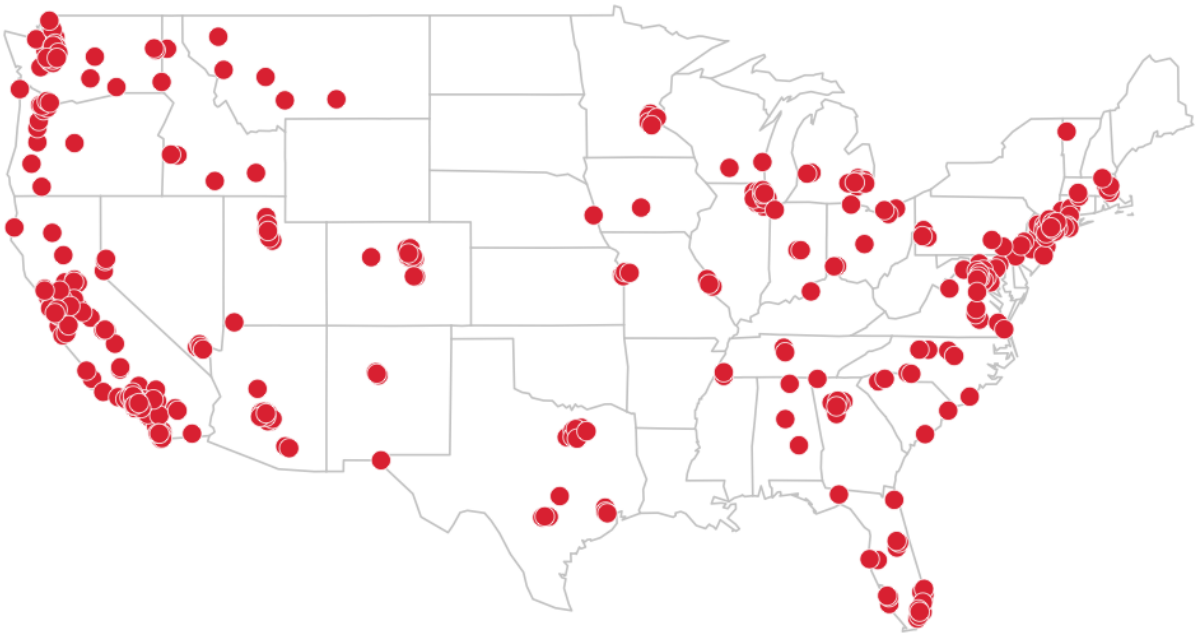
```
#nba.PTS <- nba[order(nba$PTS),]
#row.names(nba) <-nba$Name
#nba <- nba[,2:20]
nba_matrix <- data.matrix(nba)
nba_heatmap <- heatmap(nba_matrix, Rowv=NA, Colv=NA, col = heat.colors(256), scale="colu
mn", margins=c(5,10))
```



Spatial chart

Hide

```
#Close up of states
map(database = 'state', col = '#cccccc')
symbols(costco$Longitude, costco$Latitude, bg = '#e2373f', fg = '#ffffff',
        lwd = 0.5, circles = rep(1, length(costco$Longitude)),
        inches = 0.05, add = TRUE)
```



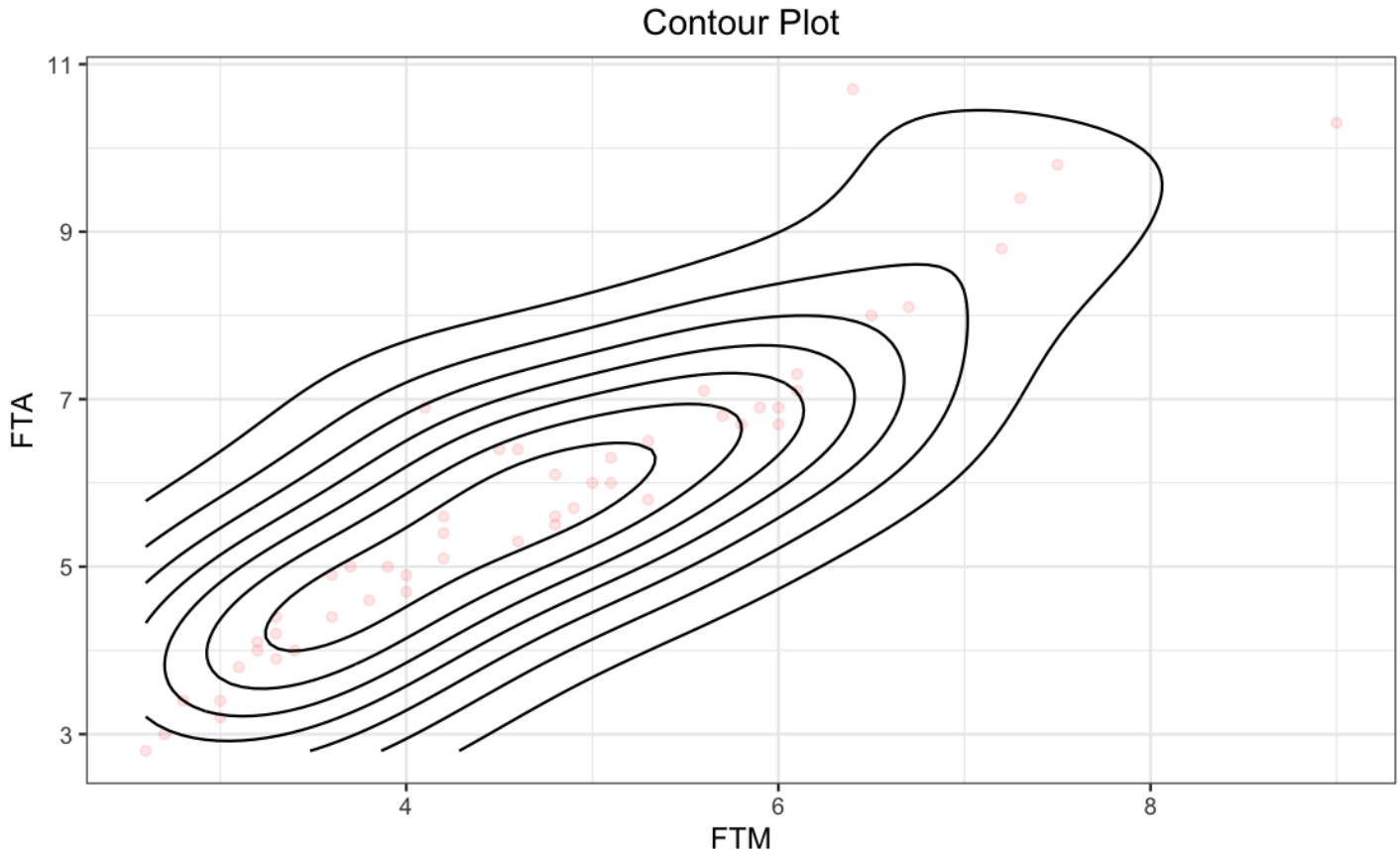
Hide

```
#View with Hawaii and Alaska included
map(database = 'world', col = '#cccccc')
symbols(costco$Longitude, costco$Latitude, bg = '#e2373f', fg = '#ffffff',
        lwd = 0.5, circles = rep(1, length(costco$Longitude)),
                       inches = 0.05, add = TRUE)
```

Contour plot

Hide

```
#Using ggplot

ggplot(nba, aes(x=FTM, y=FTA))+
  theme_bw()+
  geom_point(alpha=0.1, col='red')+
  geom_density2d(color='black')+
  ggtitle('Contour Plot')+
  theme(plot.title = element_text(hjust = 0.5))+
  labs(x='FTM', y='FTA')
```


Contour Plot
FTA
FTM
11
9
7
5
3
4
6
8