Assignment 5.2 Charts

DSC640

Taniya Adhikari

Tableau – Heat Map

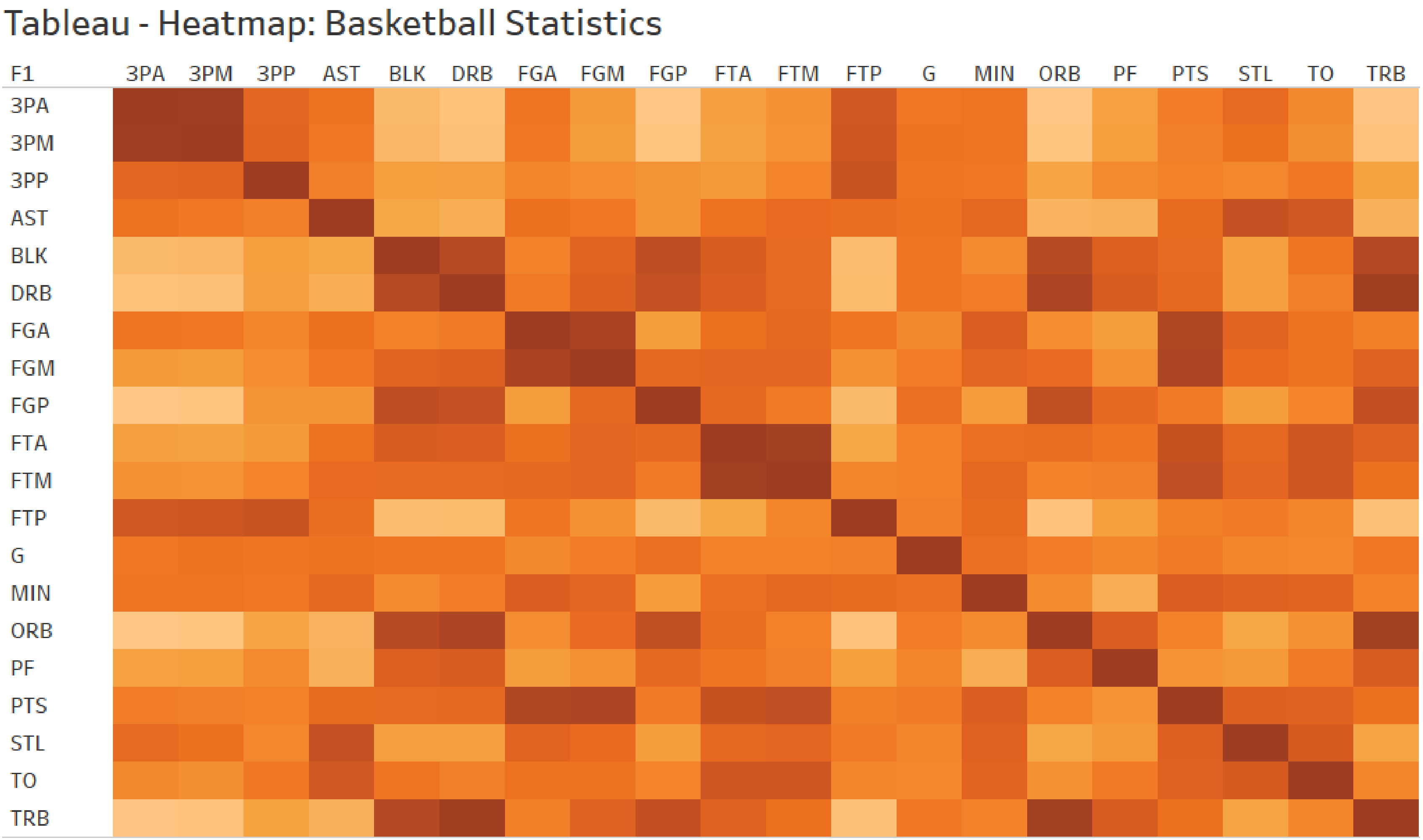


Tableau - Heatmap: Basketball Statistics

Tableau – Spatial Charts

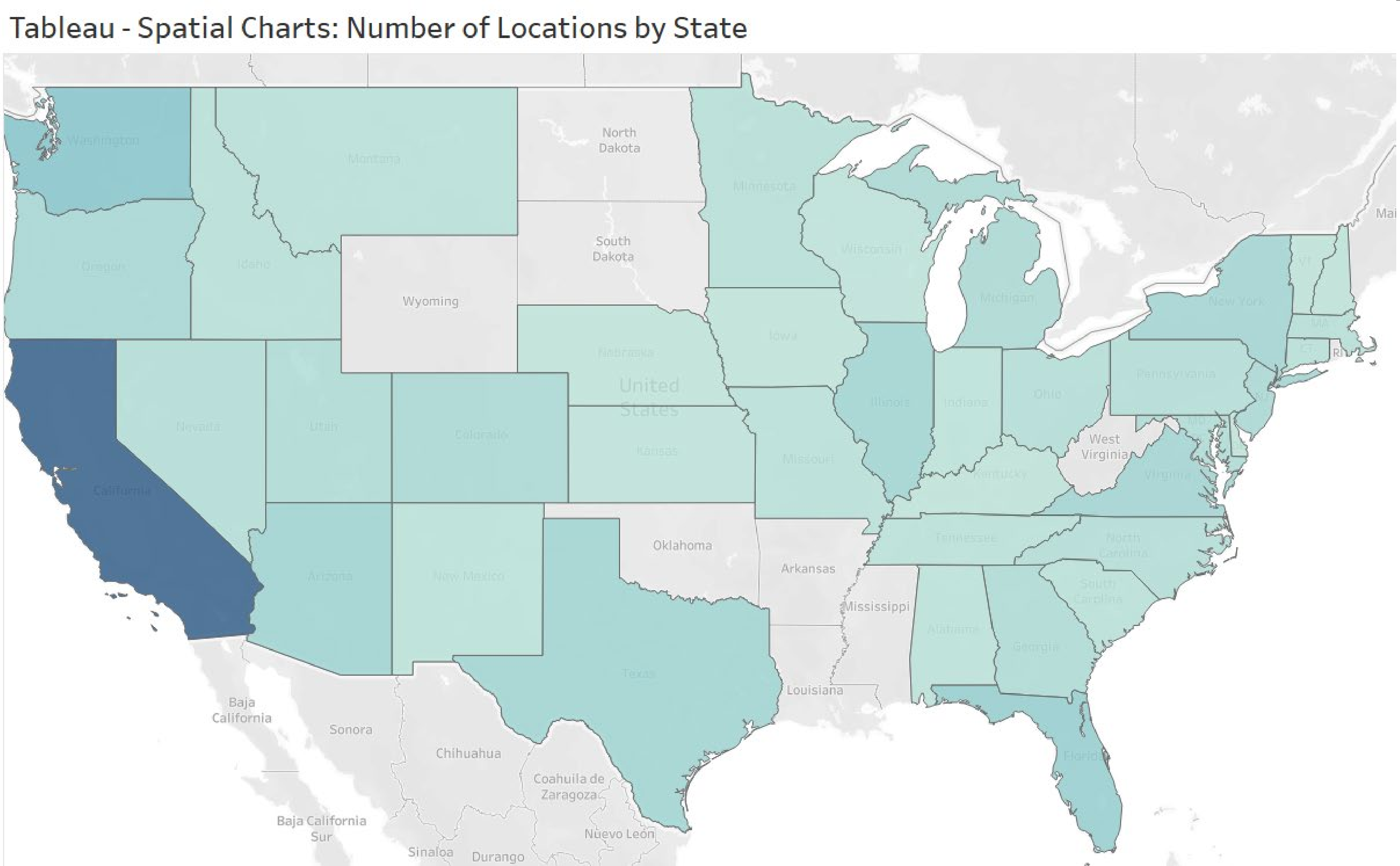


Tableau - Spatial Charts: Number of Locations by State

Tableau – Funnel Charts

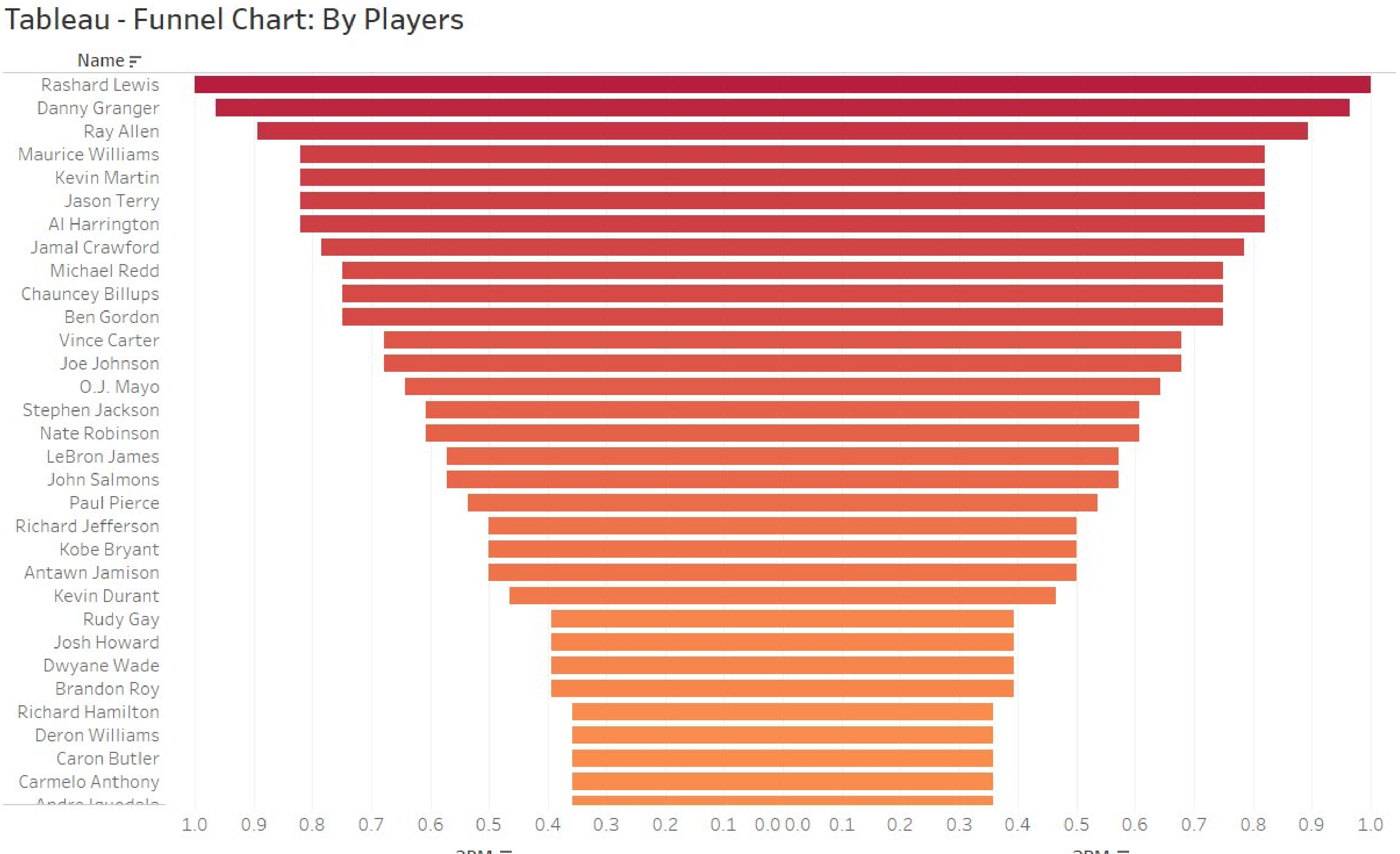
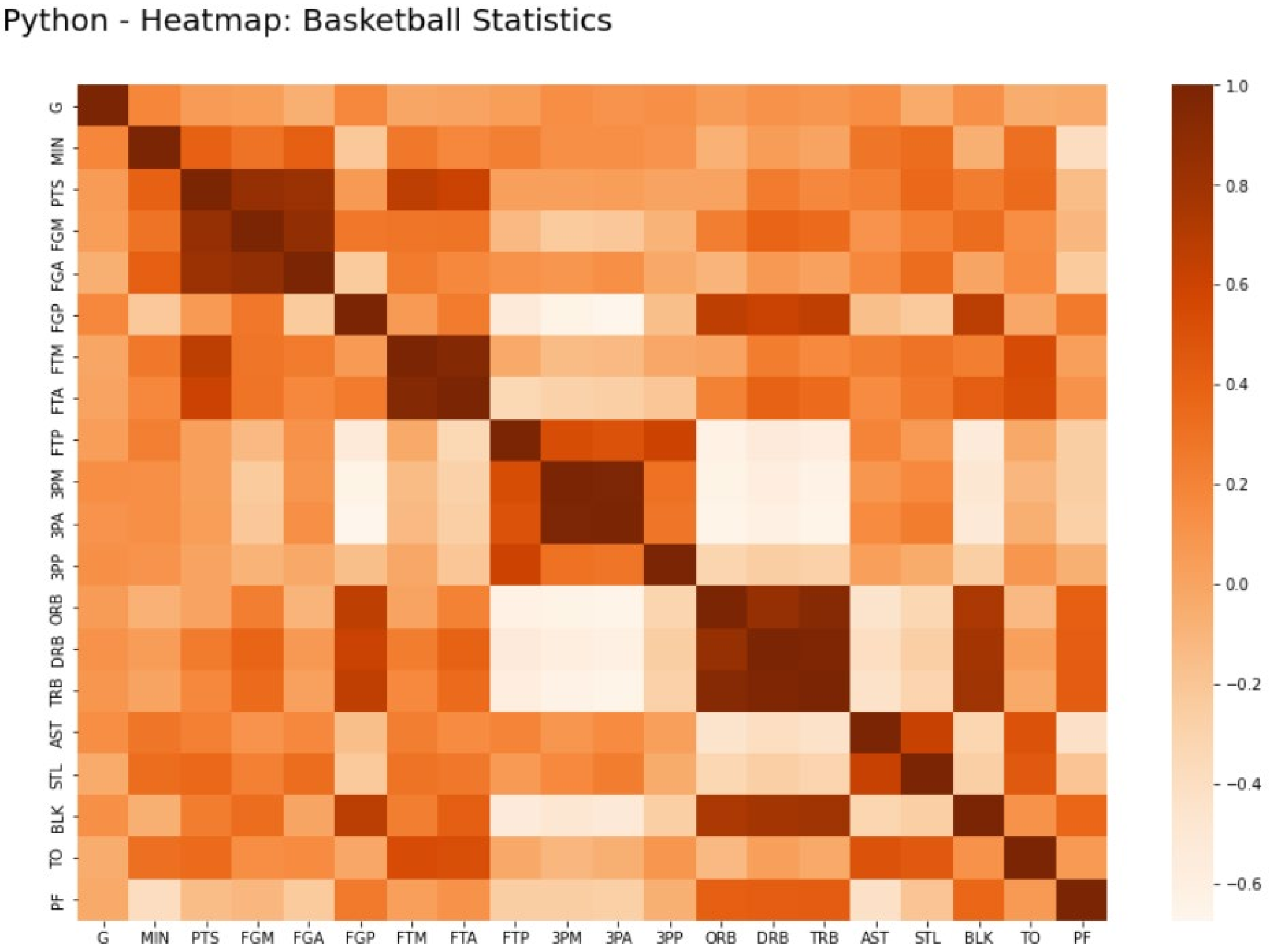


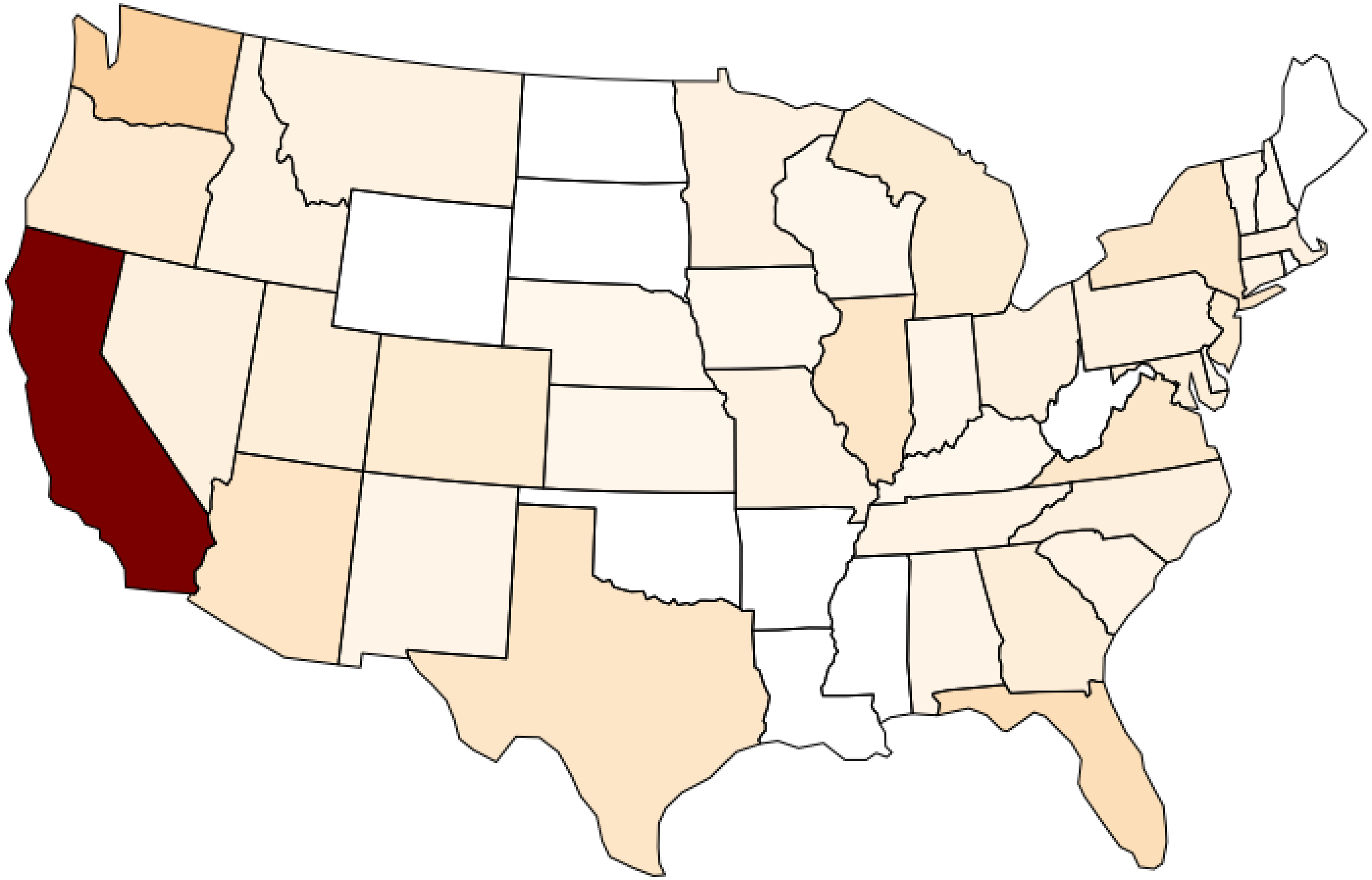
Tableau - Funnel Chart: By Players

Python – Heat Map

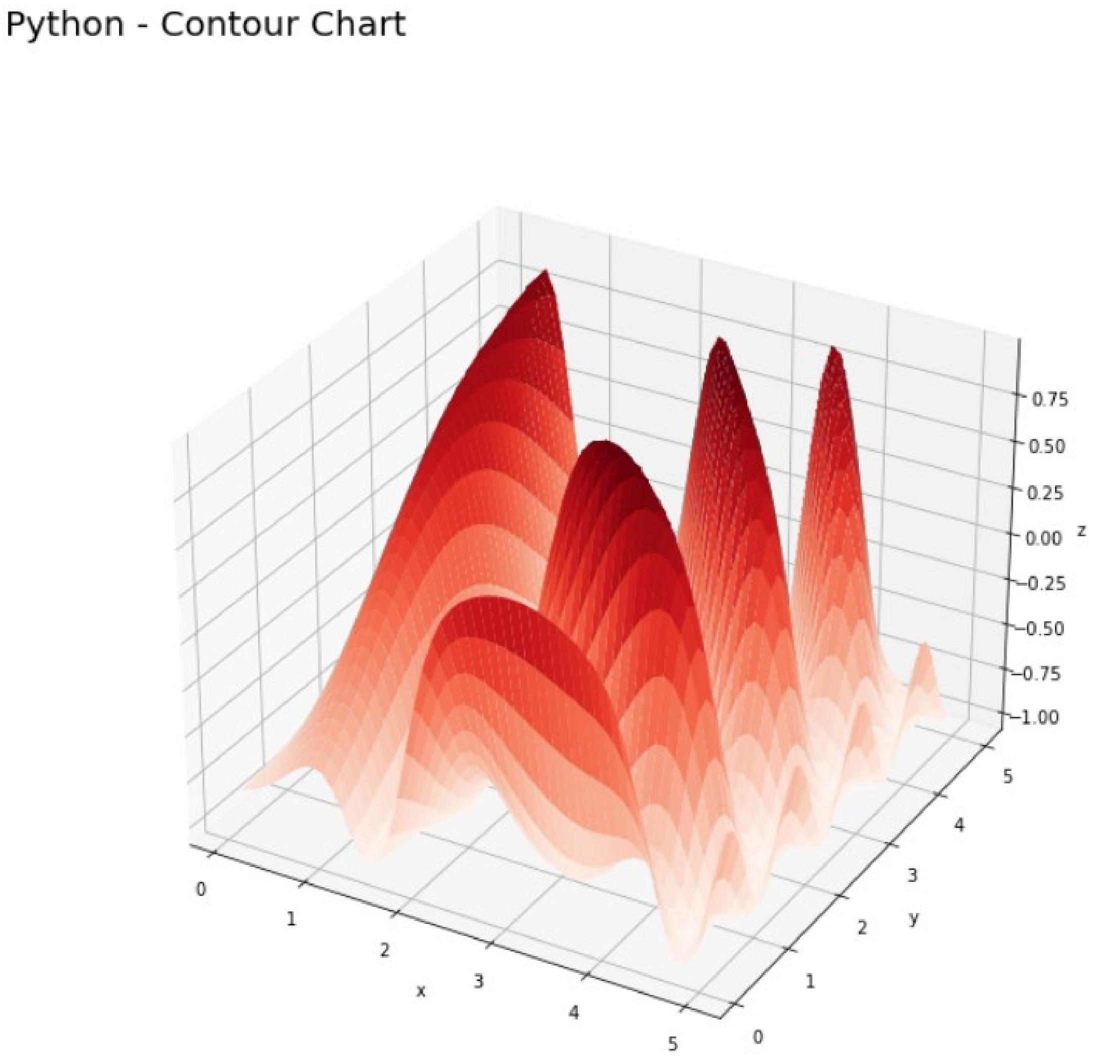


Python – Spatial Chart

Python - Spatial CHarts: Number of Locations by State

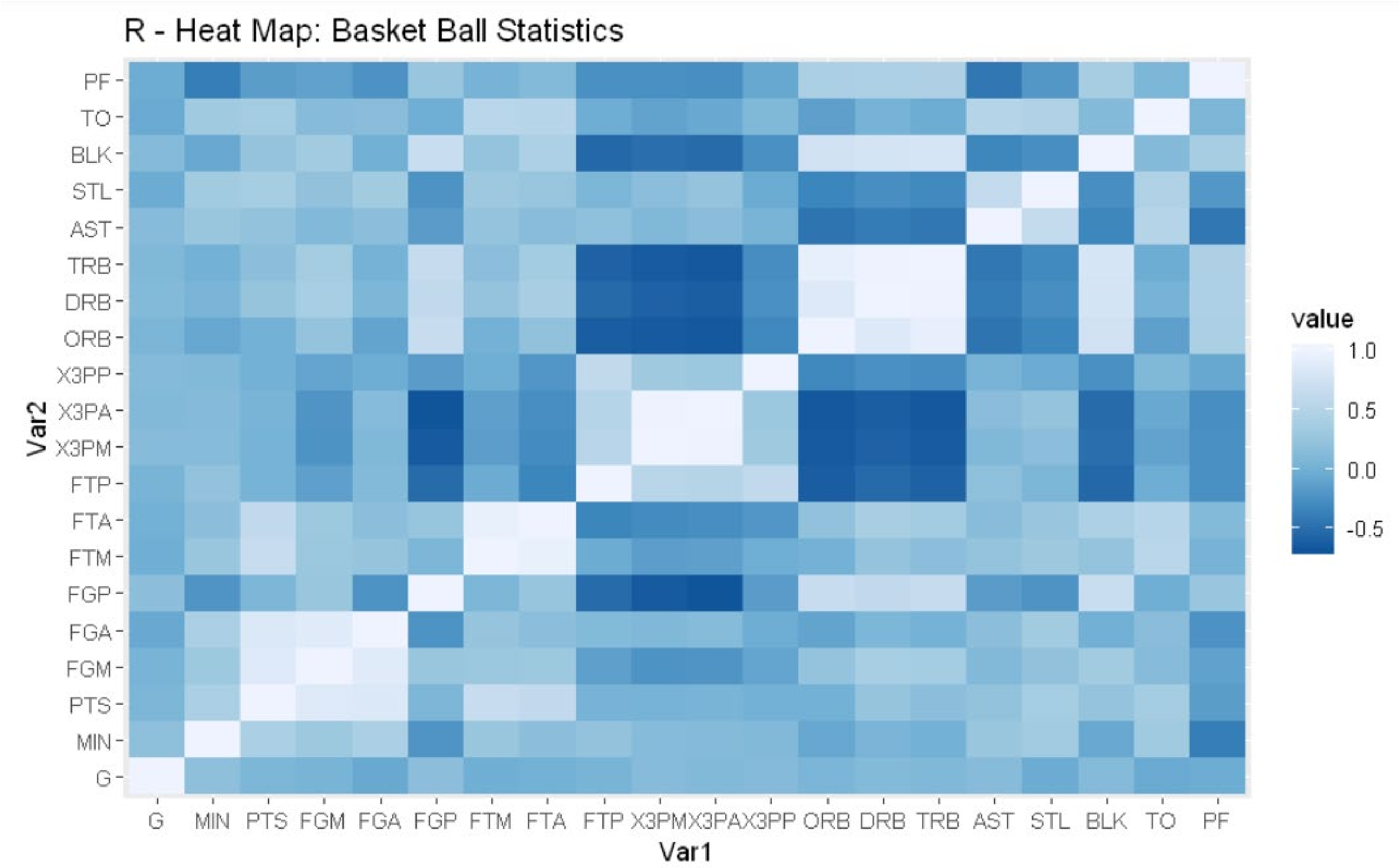


Python – Contour Chart

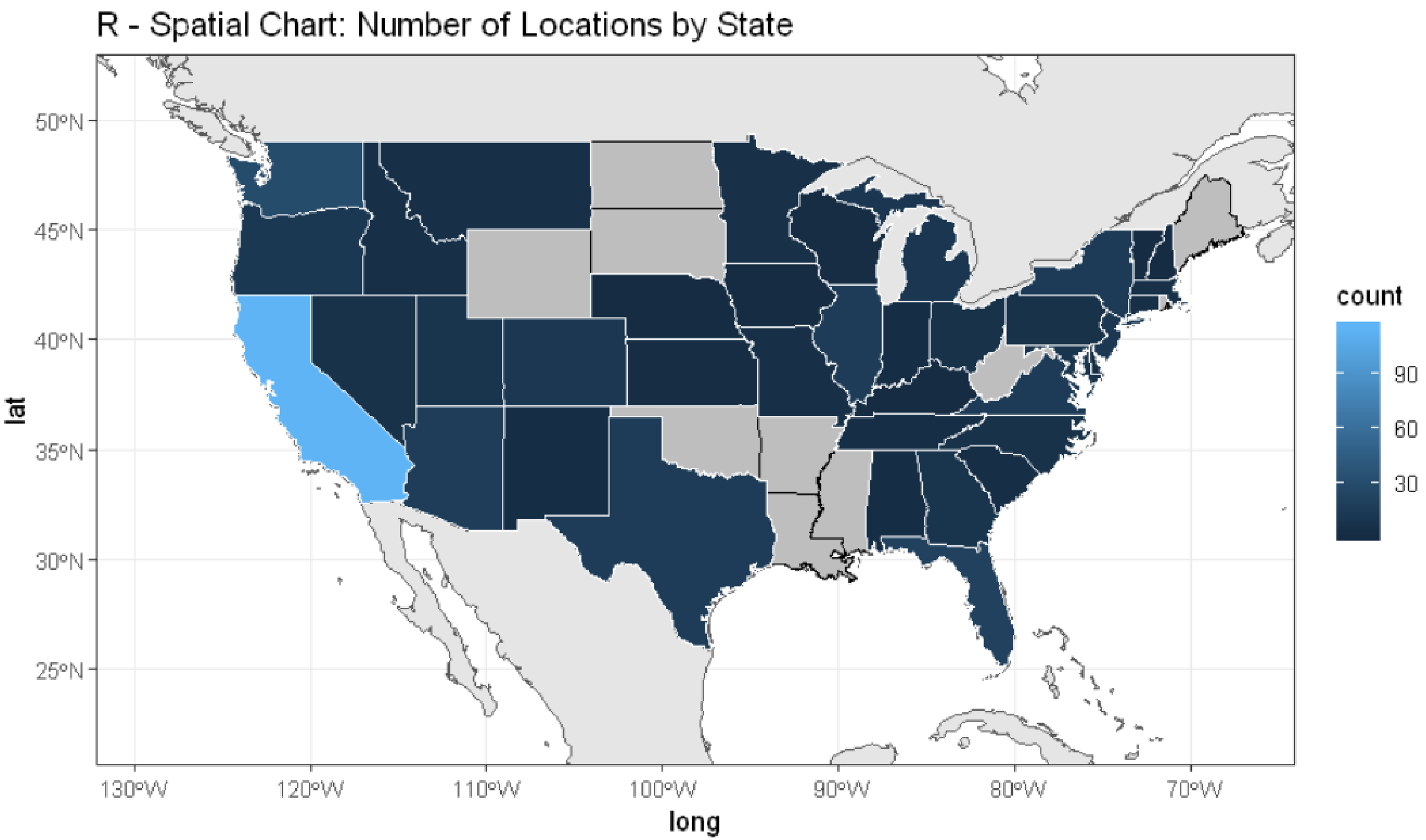


Python - Contour Chart

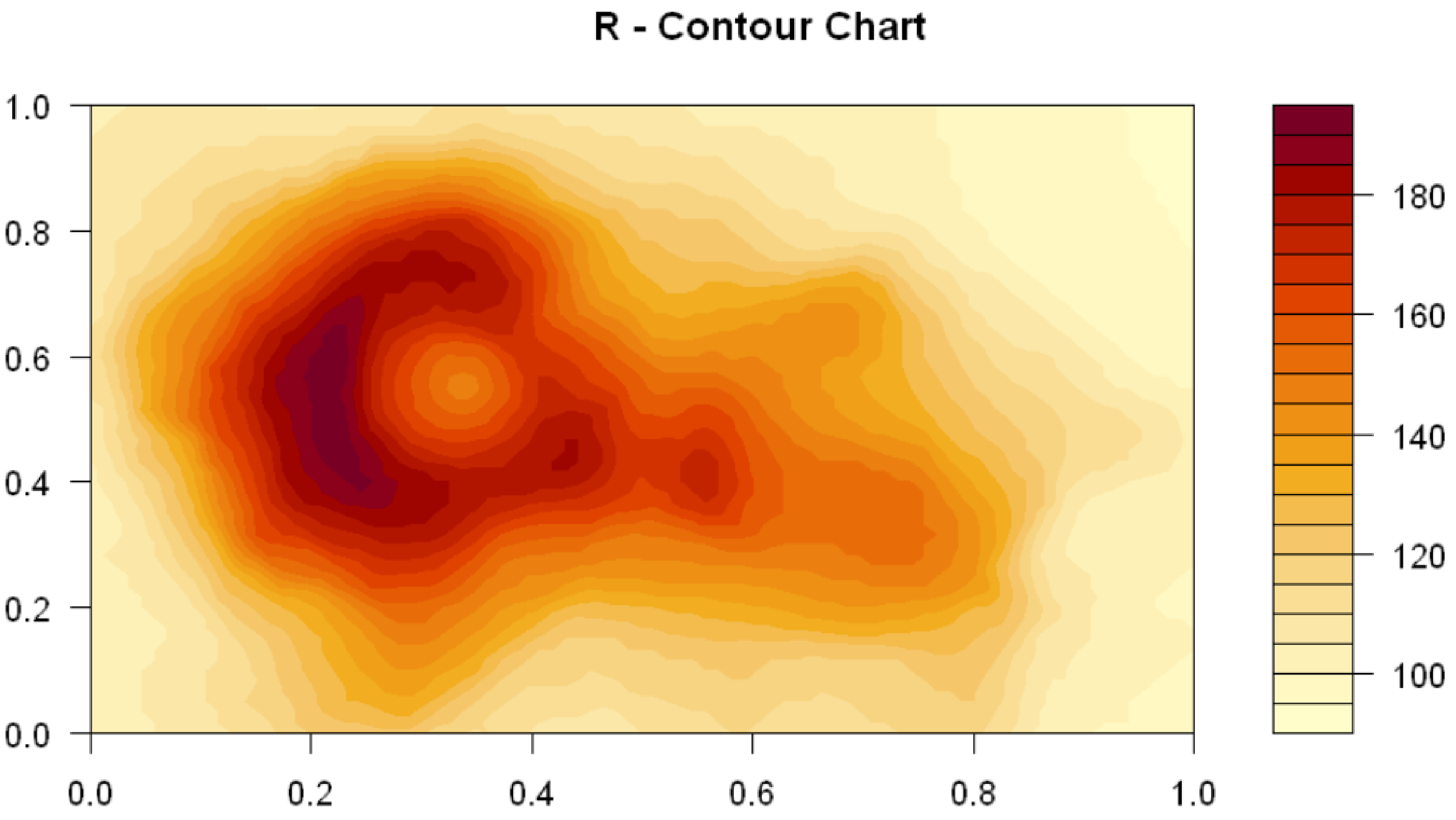
R– Heat Map


R - Heat Map: Basket Ball Statistics
Var2
PF
TO
BLK
STL
AST
TRB
DRB
ORB
X3PP
X3PA
X3PM
FTP
FTA
FTM
FGP
FGA
FGM
PTS
MIN
G
Var1
G
MIN
PTS
FGM
FGA
FGP
FTM
FTA
FTP
X3PM
X3PA
X3PP
ORB
DRB
TRB
AST
STL
BLK
TO
PF
value
1.0
0.5
0.0
-0.5

R – Spatial Chart



R – Contour Chart



Supplemental Files

- Tableau Screenshot
- Python Code
- R Code

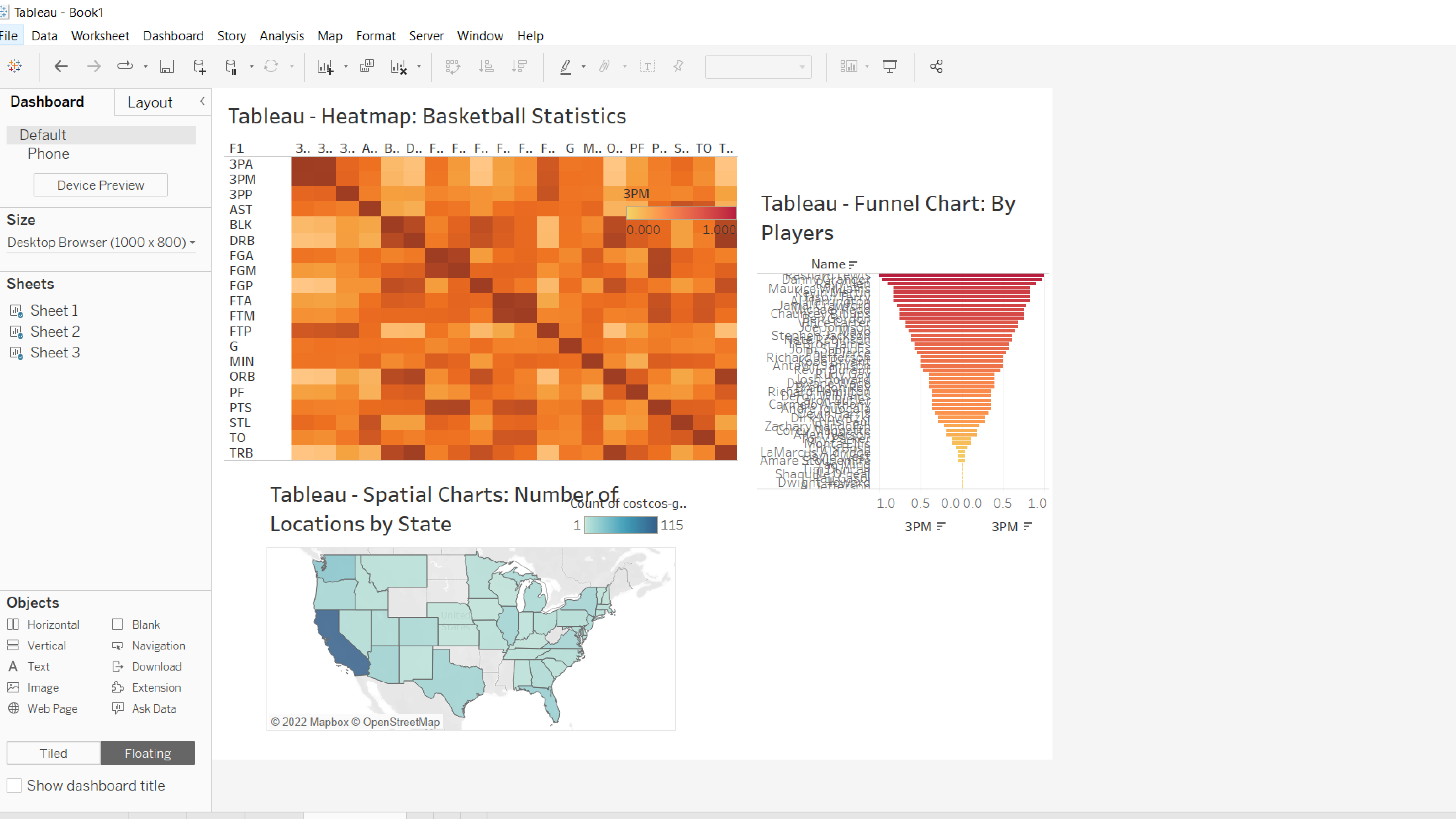


Tableau - Book1
File Data Worksheet Dashboard Story Analysis Map Format Server Window Help
Dashboard Layout
Default
Phone
Device Preview
Size
Desktop Browser (1000 x 800)
Sheets
Sheet 1
Sheet 2
Sheet 3
Objects
Horizontal
Blank
Vertical
Navigation
Text
Download
Image
Extension
Web Page
Ask Data
Tiled
Floating
Show dashboard title
Tableau - Heatmap: Basketball Statistics
F1
3.. 3.. 3.. A.. B.. D.. F.. F.. F.. F.. F.. F.. G M.. O.. PF P.. S.. TO T..
3PA
3PM
3PP
AST
BLK
DRB
FGA
FGM
FGP
FTA
FTM
FTP
G
MIN
ORB
PF
PTS
STL
TO
TRB
3PM
0.000
1.000
Tableau - Funnel Chart: By Players
Name
Rashard Lewis
Danny Granger
Maurice Williams
Jamal Crawford
Chauncey Billups
Joe Johnson
Stephen Jackson
Richard Jefferson
Antawn Jamison
Rudy Gay
Richard Hamilton
Carmelo Anthony
Zachary Randolph
Corey Maggette
LaMarcus Aldridge
Amare Stoudemire
Shaquille O'neal
Dwight Howard
1.0 0.5 0.0 0.0 0.5 1.0
3PM
3PM
Tableau - Spatial Charts: Number of Locations by State
Count of costcos-g..
1
115
United States
© 2022 Mapbox © OpenStreetMap

Python Script

Assignment 5.2: Heat Maps, Contour Charts and Spatial Charts

DSC640

Taniya Adhikari

In [2]:

```
import datetime as dt
from pathlib import Path
import math
import os
import sqlite3
import json
import geopandas as gpd
import pygeos
import pyproj
import shapely
import shapely.ops as ops
from shapely.geometry import Point, Polygon
from shapely.geometry.polygon import Polygon
from functools import partial
import geoplot as gplt
import geoplot.crs as gcrs

import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
# Data Preprocessing
from sklearn.preprocessing import MinMaxScaler
```

Heat Maps

In [4]:

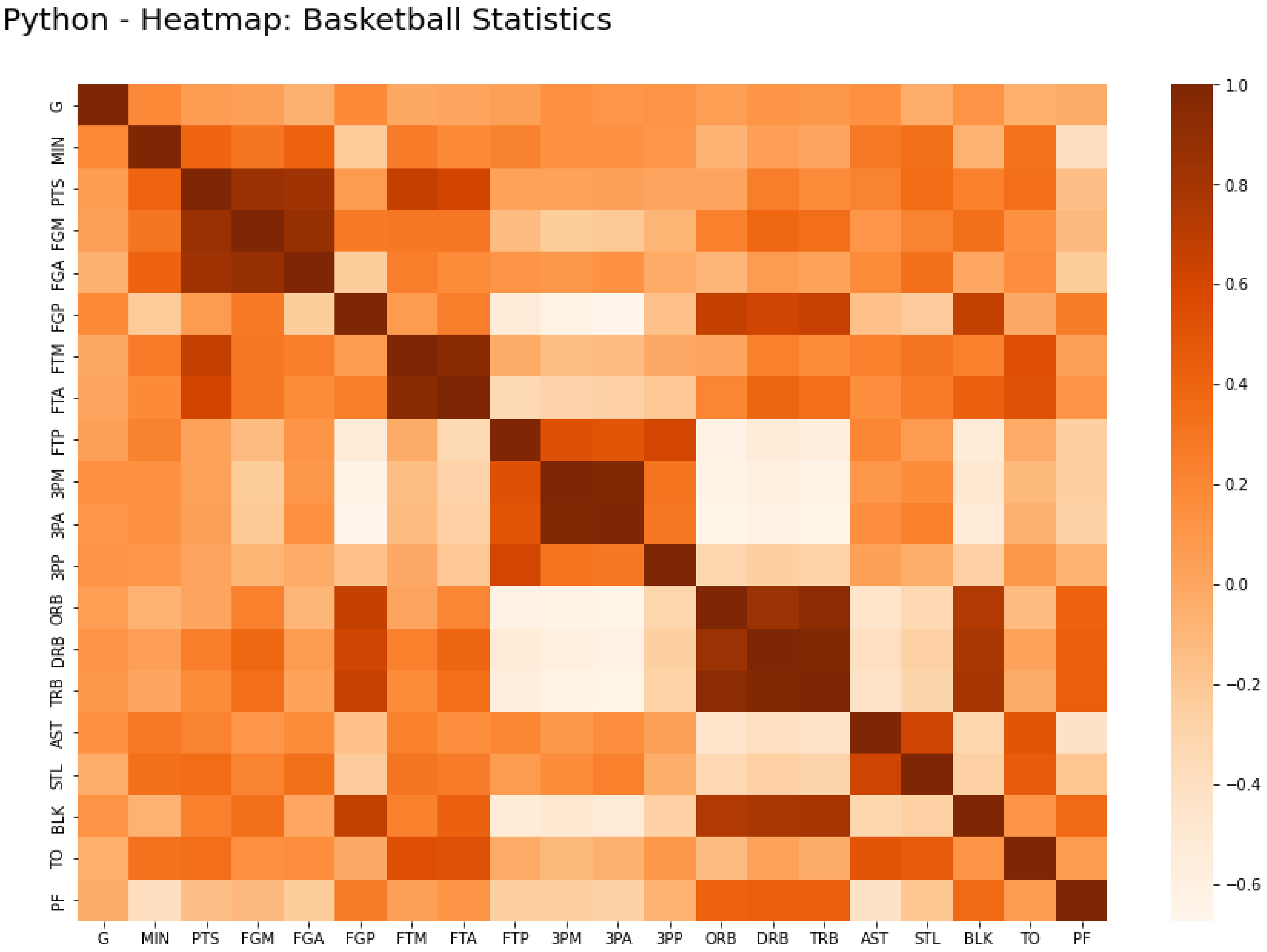
```
df = pd.read_csv("ppg2008.csv")
df.columns = df.columns.str.strip()
```

```
In [5]: df.head(5)
```

```
Out[5]:
          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 [14]:
plt.rcParams['figure.figsize'] = [15,10]
fig, ax = plt.subplots()
sns.heatmap(df.drop(['Name'], axis = 1).corr(), ax = ax, cmap="Oranges")
plt.suptitle("Python - Heatmap: Basketball Statistics",
             size=20, x=0.08, y=.95,horizontalalignment='left', verticalalignment='top')
plt.show()
```



Spatial Charts

In [8]:

```
contiguous_usa = gpd.read_file(gplt.datasets.get_path('contiguous_usa'))
```

In [9]:

```
df2 = pd.read_csv("costcos-geocoded.csv")
df2.head()
```

Out[9]:

	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 [10]:

```
location = gpd.GeoDataFrame(df2, geometry=gpd.points_from_xy(df2['Longitude'], df2['Latitude']))
```

In [11]:

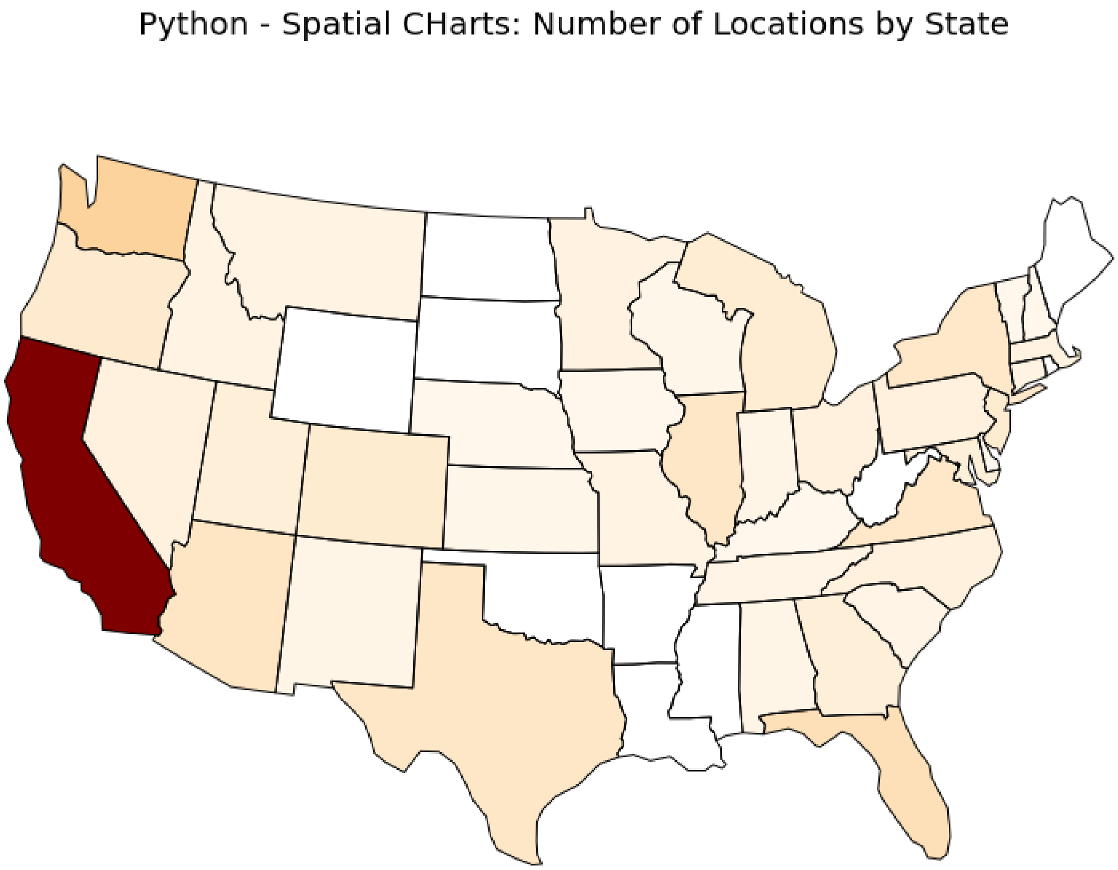
```
location2 = pd.DataFrame(location['State'].value_counts()).reset_index()
location2.columns = ['State','Count']

geodata = contiguous_usa.merge(location2, how='outer', left_on=['state'], right_on=['State'])
```

In [12]:

```
ax = gplt.polyplot(contiguous_usa,
                   projection=gcrs.AlbersEqualArea(),
                   figsize = (19,16),
                   zorder = 2)
gplt.choropleth(geodata,
                hue = geodata['Count'],
                cmap = 'OrRd',
                projection=gcrs.WebMercator(),
                ax = ax)
plt.title("Python - Spatial CHarts: Number of Locations by State", fontsize=20)
plt.show()
```

```
C:\Users\bibek\anaconda3\envs\my_env\lib\site-packages\geoplot\geoplot.py:66: UserWarning: The data being passed to "hu
e" includes null values. You probably want to remove these before plotting this data with geoplot.
  warnings.warn(
```



Contour Maps

In [15]:

```
def f(x, y):
    return -np.sin(x) ** 10 + np.cos(10 + y * x) * np.cos(x)

x = np.linspace(0.1, 5, 50)
y = np.linspace(0.1, 5, 40)

X, Y = np.meshgrid(x, y)
Z = f(X, Y)
```

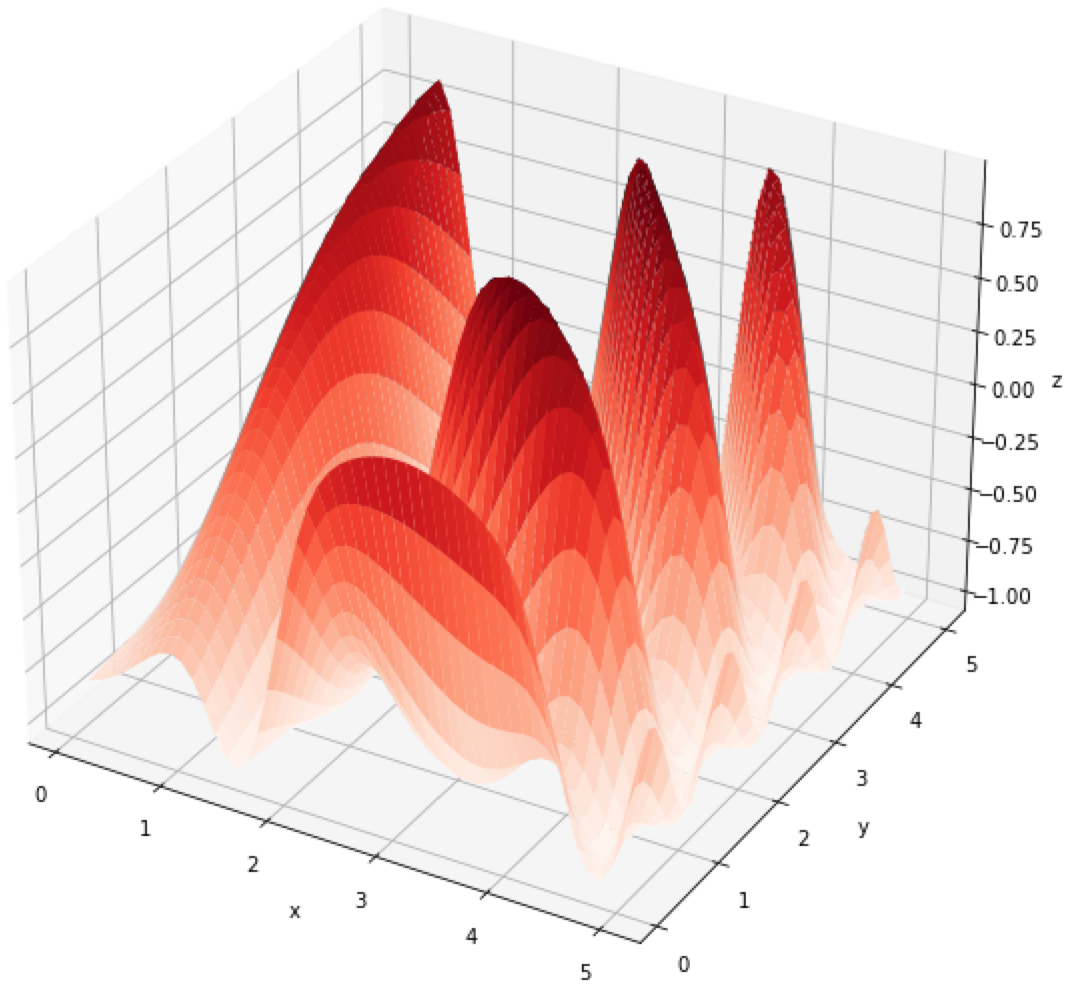
In [25]:

```
ax = plt.axes(projection='3d')
ax.contour3D(X, Y, Z, 40, cmap='Greys')
ax.plot_surface(X, Y, Z, rstride=1, cstride=1,
                cmap='Reds', edgecolor='none')
ax.set_xlabel('x')
ax.set_ylabel('y')
ax.set_zlabel('z');

plt.suptitle("Python - Contour Chart",
             size=20, x=.2,horizontalalignment='left', verticalalignment='top')

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

Python - Contour Chart



```
In [ ]:
```

R Script

Assignment 5.2: Heat Maps, Contour Charts and Spatial Charts

DSC640

Taniya Adhikari

In [60]:

```
library(ggplot2)
library(readxl)
library(scales)
library(plyr)
library(dplyr)
library(ggrepel)
library(reshape2)
```

In [32]:

```
install.packages(c("cowplot", "googleway", "ggplot2", "ggspatial", "libwgeom", "sf", "rnaturalearth", "r
```

```
Warning message:
"package 'libwgeom' is not available (for R version 3.6.3)"Warning message:
"dependency 'rjson' is not available"also installing the dependencies 'proxy', 'png', 'rgdal', 'prettymap
r', 'e1071', 'wk', 'jpeg', 'jqr', 'googlePolylines', 'isoband', 'rlang', 'rosm', 'abind', 'classInt', 's2',
'units', 'sp'



  There are binary versions available but the source versions are later:
           binary source needs_compilation
proxy      0.4-25 0.4-26              TRUE
rgdal      1.5-23 1.5-28              TRUE
e1071       1.7-6  1.7-9              TRUE
wk          0.4.1  0.6.0              TRUE
jpeg      0.1-8.1  0.1-9              TRUE
jqr         1.2.1  1.2.2              TRUE
isoband     0.2.4  0.2.5              TRUE
rlang      0.4.11  1.0.2              TRUE
s2          1.0.4  1.0.7              TRUE
units       0.7-1  0.8-0              TRUE
sp          1.4-5  1.4-6              TRUE
googleway   2.7.3  2.7.6             FALSE
ggplot2     3.3.3  3.3.5             FALSE
sf          0.9-8  1.0-6              TRUE

  Binaries will be installed

Warning message:
"package 'ggrepel' is in use and will not be installed"

package 'proxy' successfully unpacked and MD5 sums checked
package 'png' successfully unpacked and MD5 sums checked
package 'rgdal' successfully unpacked and MD5 sums checked
package 'prettymapr' successfully unpacked and MD5 sums checked
package 'e1071' successfully unpacked and MD5 sums checked
package 'wk' successfully unpacked and MD5 sums checked
package 'jpeg' successfully unpacked and MD5 sums checked
package 'jqr' successfully unpacked and MD5 sums checked
package 'googlePolylines' successfully unpacked and MD5 sums checked
package 'isoband' successfully unpacked and MD5 sums checked
package 'rlang' successfully unpacked and MD5 sums checked

Warning message:
"cannot remove prior installation of package 'rlang'"Warning message in file.copy(savedcopy, lib, recursive
= TRUE):
"problem copying C:\Users\bibek\anaconda3\envs\r-environment\Lib\R\library\00LOCK\rlang\libs\x64\rlang.dll
to C:\Users\bibek\anaconda3\envs\r-environment\Lib\R\library\rlang\libs\x64\rlang.dll: Permission denied"Wa
rning message:
"restored 'rlang'"

package 'rosm' successfully unpacked and MD5 sums checked
package 'abind' successfully unpacked and MD5 sums checked
package 'classInt' successfully unpacked and MD5 sums checked
package 's2' successfully unpacked and MD5 sums checked
package 'units' successfully unpacked and MD5 sums checked
package 'sp' successfully unpacked and MD5 sums checked
package 'cowplot' successfully unpacked and MD5 sums checked
package 'ggspatial' successfully unpacked and MD5 sums checked
package 'sf' successfully unpacked and MD5 sums checked
package 'rnaturalearth' successfully unpacked and MD5 sums checked
package 'rnaturalearthdata' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\bibek\AppData\Local\Temp\RtmpWkZbZx\downloaded_packages

installing the source packages 'googleway', 'ggplot2'

Warning message in install.packages(c("cowplot", "googleway", "ggplot2", "ggrepel", :
"installation of package 'googleway' had non-zero exit status"Warning message in install.packages(c("cowplo
t", "googleway", "ggplot2", "ggrepel", :
"installation of package 'ggplot2' had non-zero exit status"
```

In [42]:

```
install.packages("rgeos")
```

```

  There is a binary version available but the source version is later:
      binary source needs_compilation
rgeos  0.5-5  0.5-9              TRUE

  Binaries will be installed
package 'rgeos' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\bibek\AppData\Local\Temp\RtmpWkZbZx\downloaded_packages
```

In [10]:

```
df <- read.csv("ppg2008.csv")
df2 <- cor(select(df, -'Name'))
df3 <- melt(df2)
```

In [11]:

```
head(df3)
```

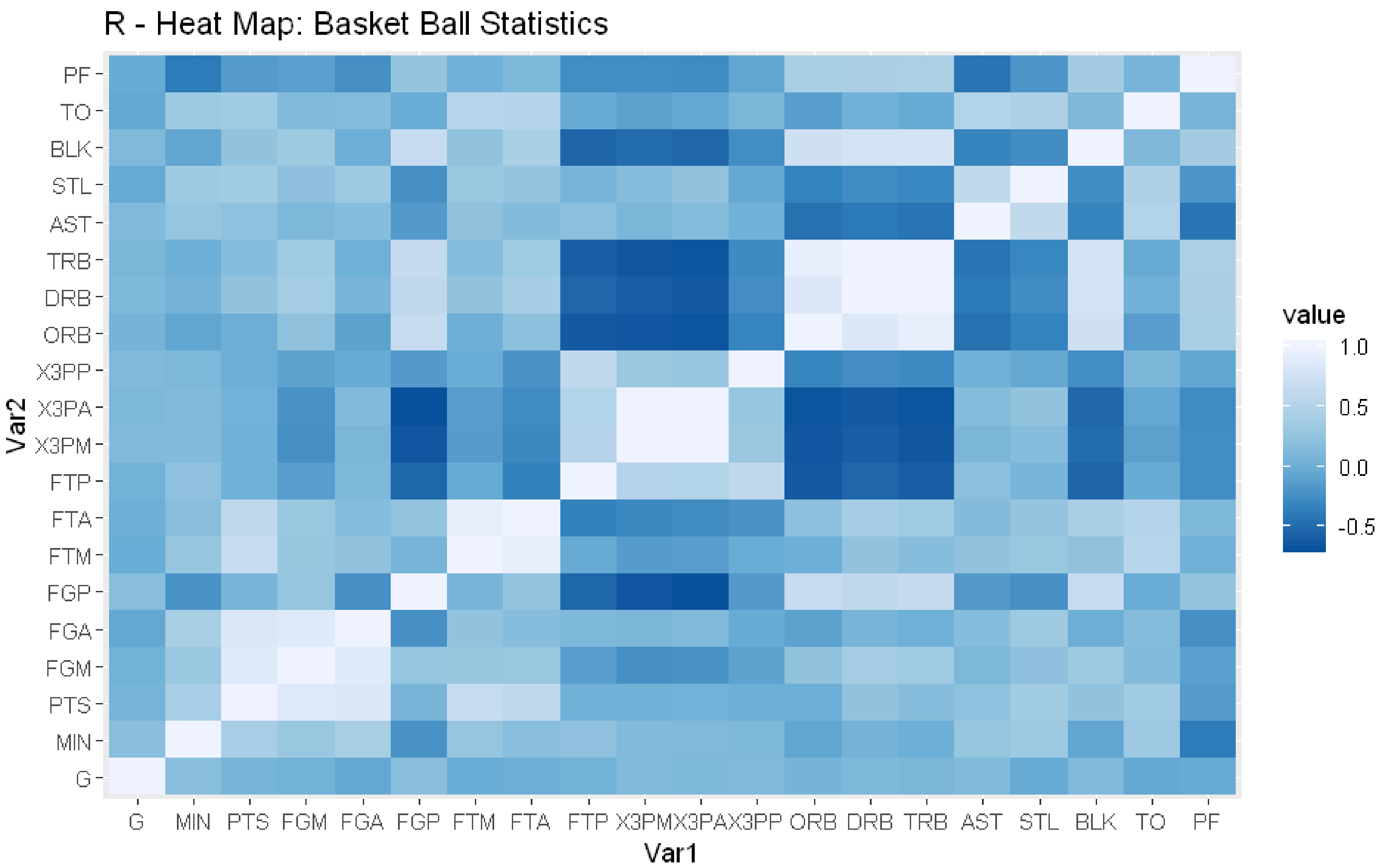
Var1	Var2	value
G	G	1.00000000
MIN	G	0.18686608
PTS	G	0.06309908
FGM	G	0.03992195
FGA	G	-0.05958051
FGP	G	0.18087541

R - HeatMap

In [16]:

```
options(repr.plot.width =8, repr.plot.height =5)
ggplot(data = df3, aes(x=Var1, y=Var2, fill=value)) +
geom_tile() +
scale_fill_distiller(palette="Blues")+
  ggtitle("R - Heat Map: Basket Ball Statistics")

```



R - Spatial Charts

In [43]:

```r
library("rnaturalearth")
library("rnaturalearthdata")
library("sf")
library("rgeos")
```

```
Loading required package: sp
rgeos version: 0.5-5, (SVN revision 640)
 GEOS runtime version: 3.8.0-CAPI-1.13.1
 Linking to sp version: 1.4-5
 Polygon checking: TRUE
```

In [44]:

```r
df <- read.csv("costcos-geocoded.csv")
```

In [45]:

```r
world <- ne_countries(scale = "medium", returnclass = "sf")
```

In [46]:

```r
counts <- df %>% count(df$State)
names(counts) <- c('region','count')
counts$region <- tolower(counts$region)
```

In [47]:

```r
MainStates <- map_data("state")
```

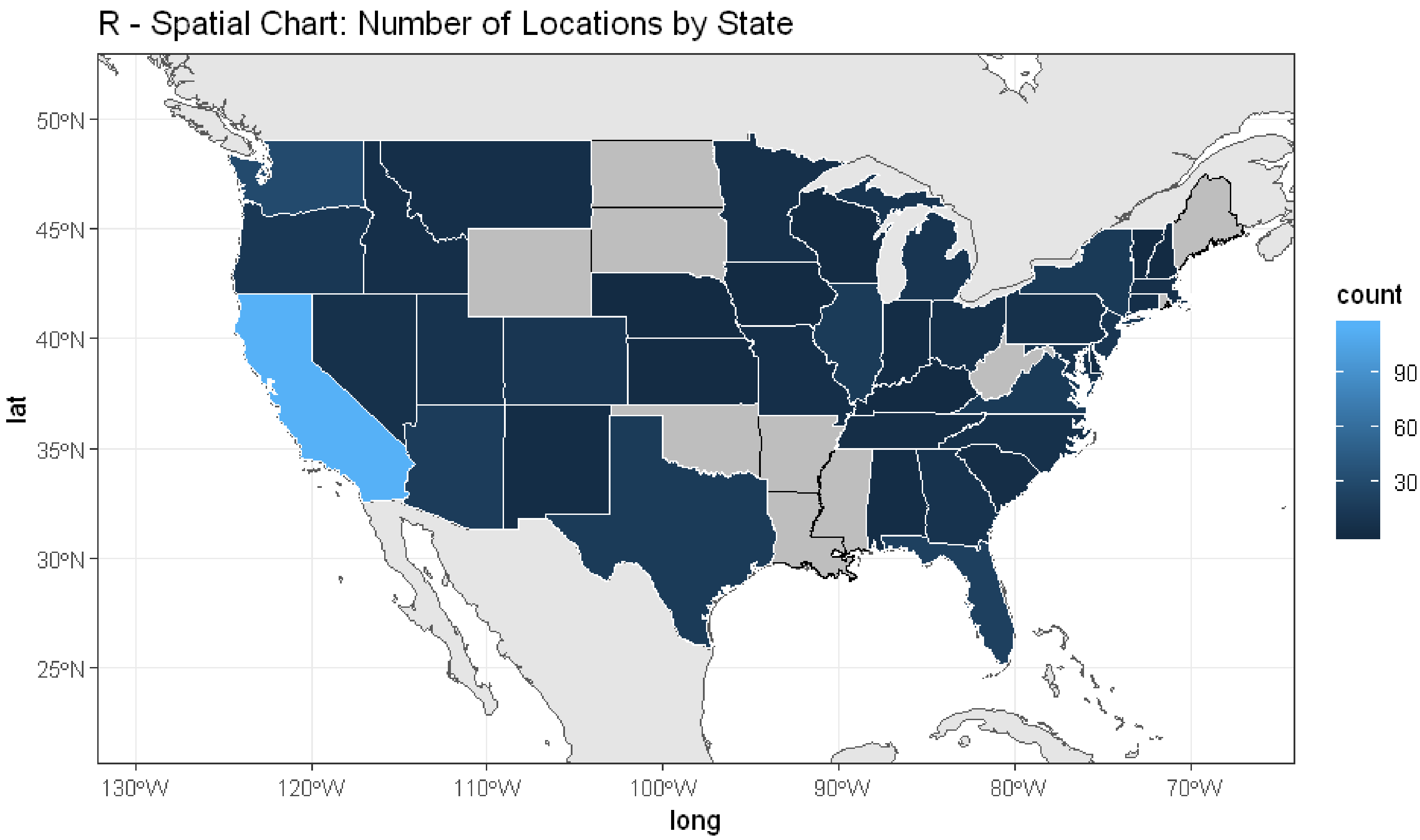
In [48]:

```r
MergedStates <-  inner_join(MainStates, counts, by = "region")
```

In [51]:

```
ggplot() +
geom_sf(data = world) +
geom_polygon( data=MainStates, aes(x=long, y=lat, group=group),
                color="black", fill="gray")+
coord_sf(xlim = c(-132.15, -64.12), ylim = c(20.65, 52.97), expand = FALSE) +
geom_polygon(data=MergedStates,
          aes(x=long, y=lat, group=group, fill = count),
          color="white", size = 0.2) +
theme_bw() +ggtitle("R - Spatial Chart: Number of Locations by State")

```



R - Contour Charts

In [ ]:

```
#install.packages("MASS")
```

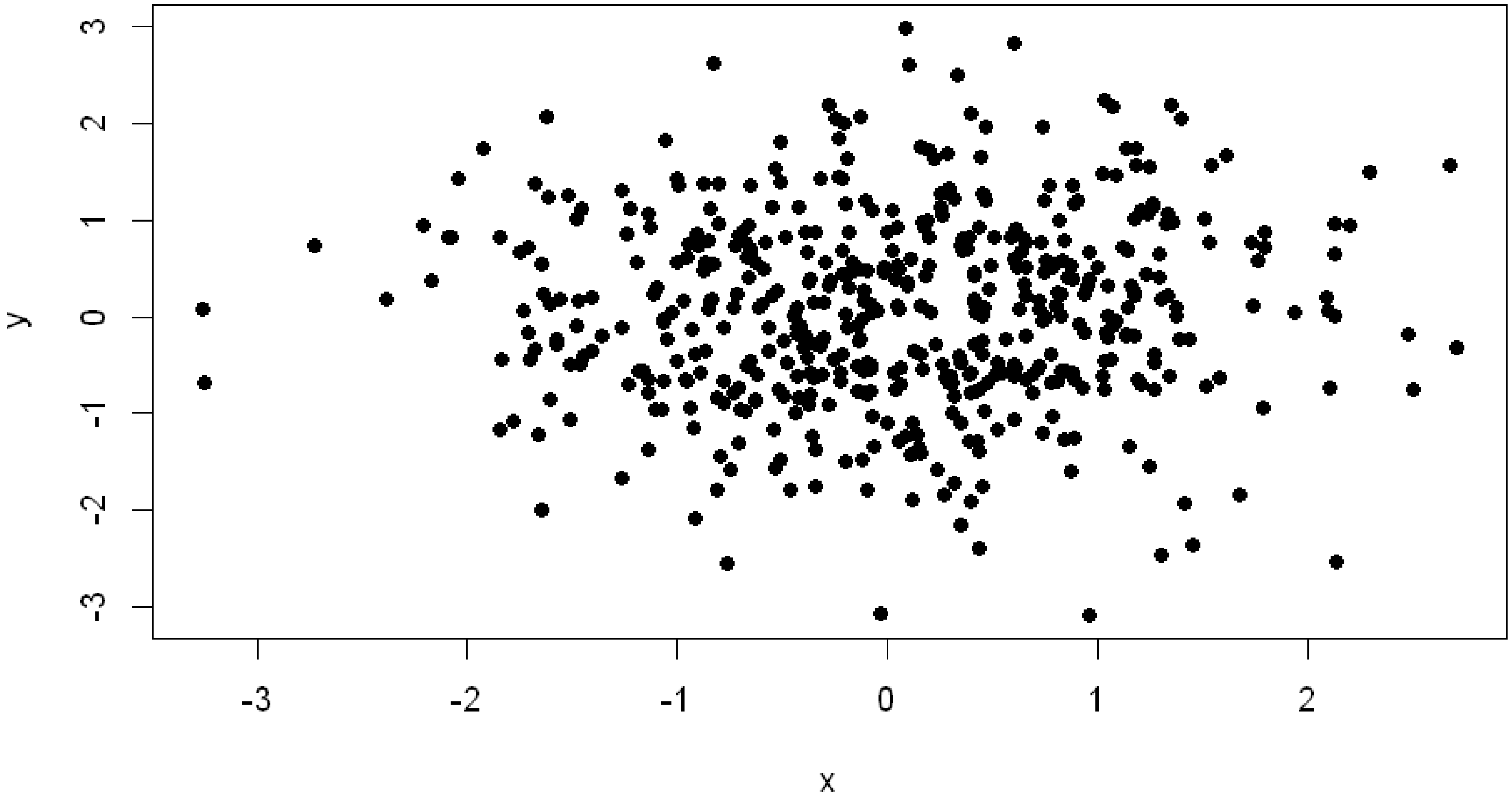
In [72]:

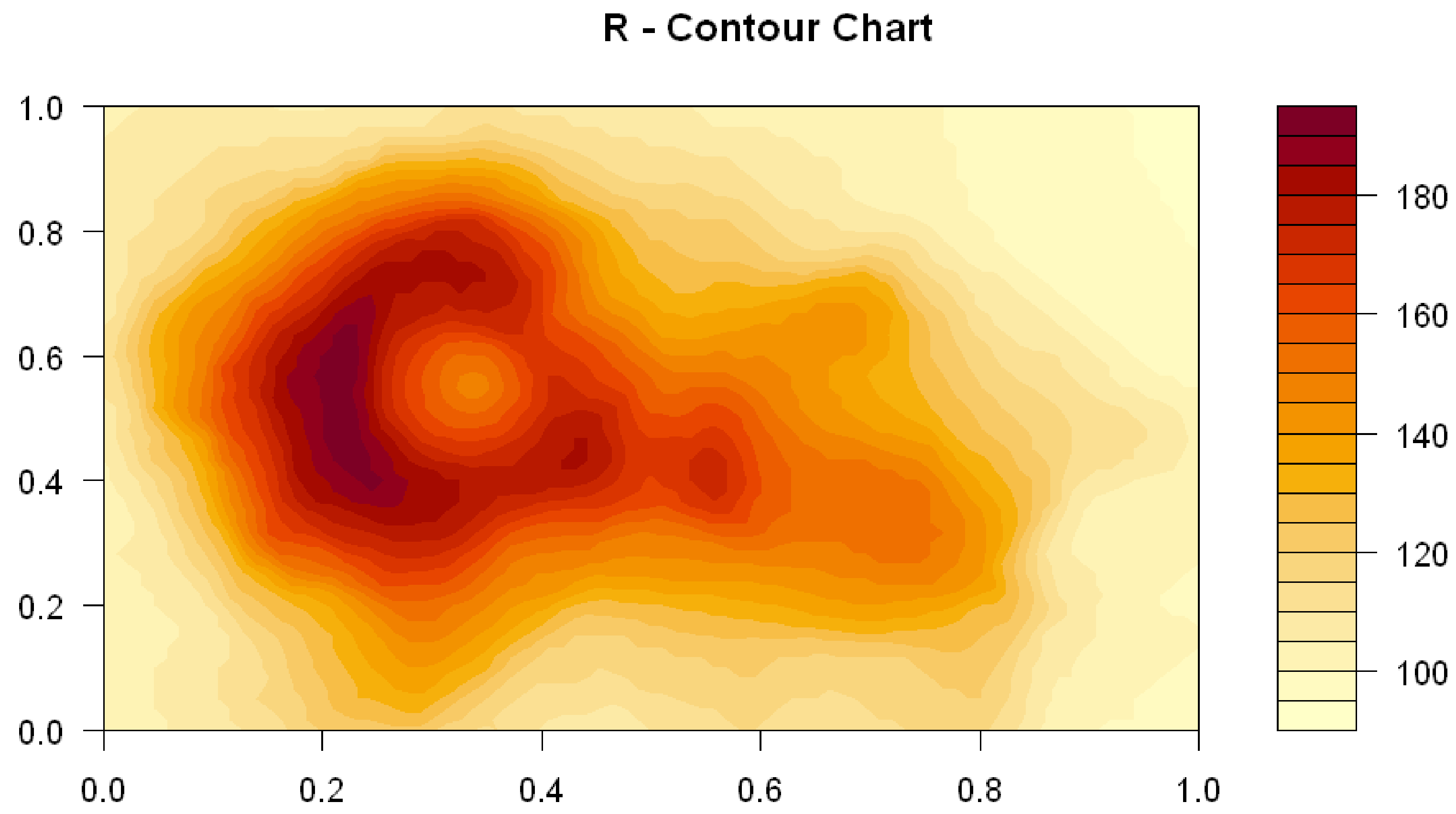
```
library(MASS)

# Data
x <- rnorm(500)
y <- rnorm(500)
z <- kde2d(x, y, n = 50)

plot(x, y, pch = 19)


filled.contour(volcano)
title("R - Contour Chart")
```


3
2
1
0
-1
-2
-3
y
-3
-2
-1
0
1
2
x


R - Contour Chart
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
180
160
140
120
100

In [ ]: