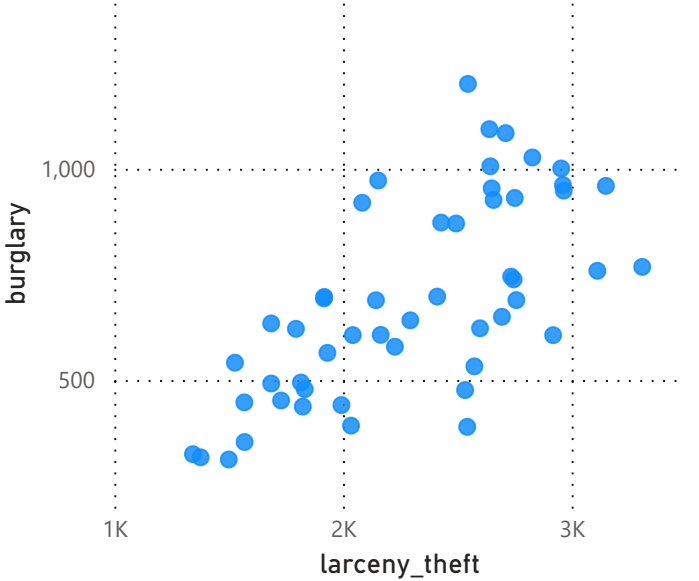
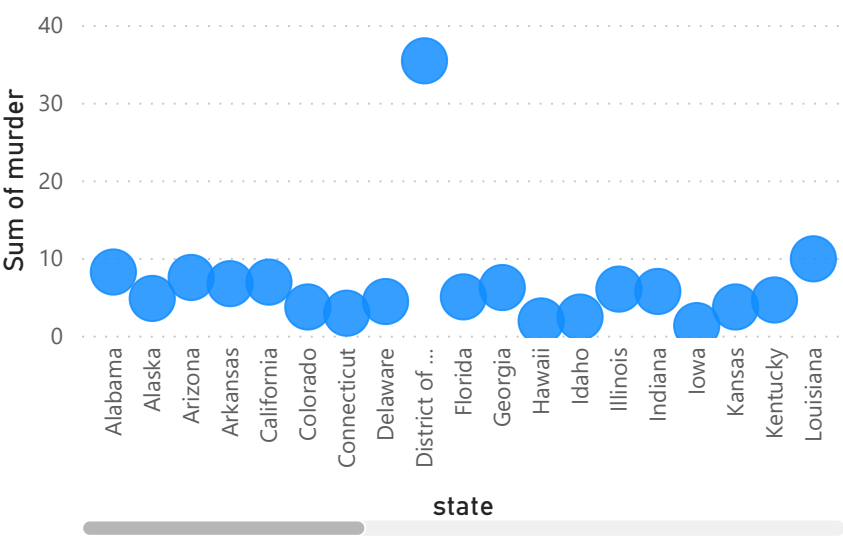


larceny\_theft and burglary scatterplot in powerbi

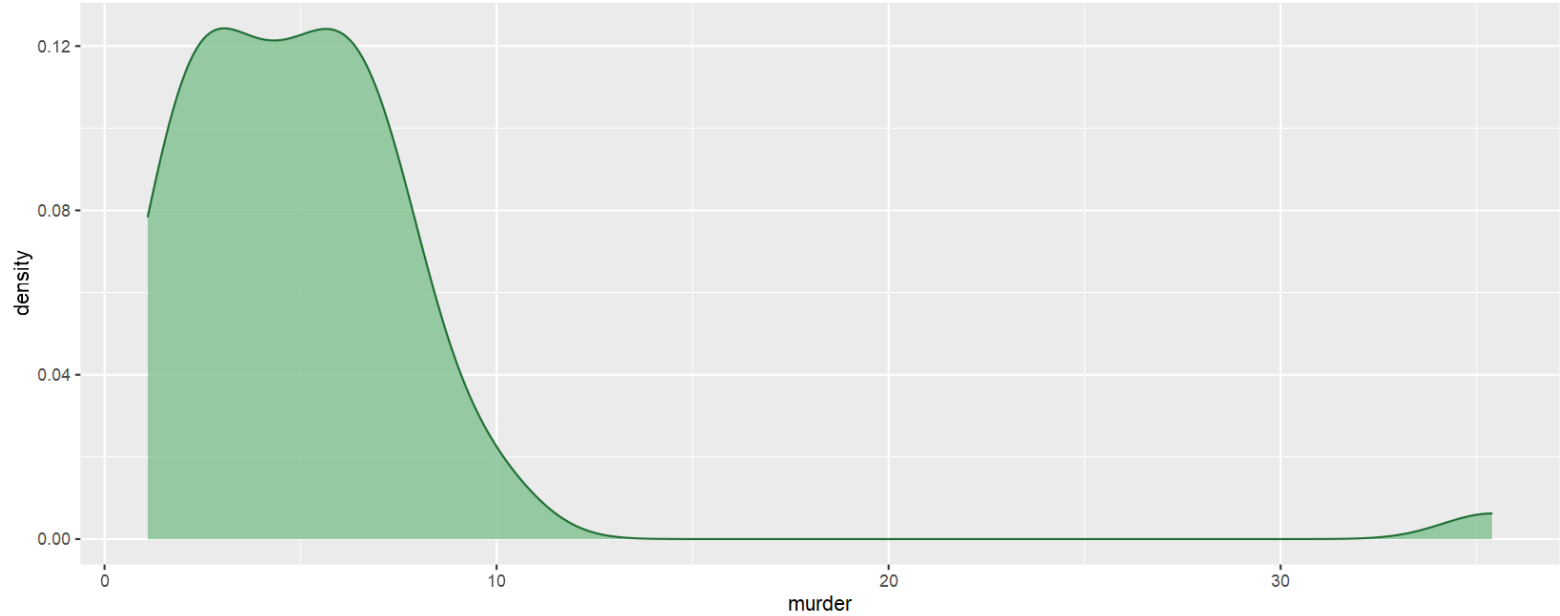


Sum of Murders and Population by State bubble chart in powerbi



Sum of murder and population

Density Plot of Murders in PowerBi



# Weeks 5 & 6 in R

Moshe Burnstein

2023-07-02

Scatterplot in R

```
crimerates_df <- read.csv('crimerates-by-state-2005.csv')
head(crimerates_df)
```

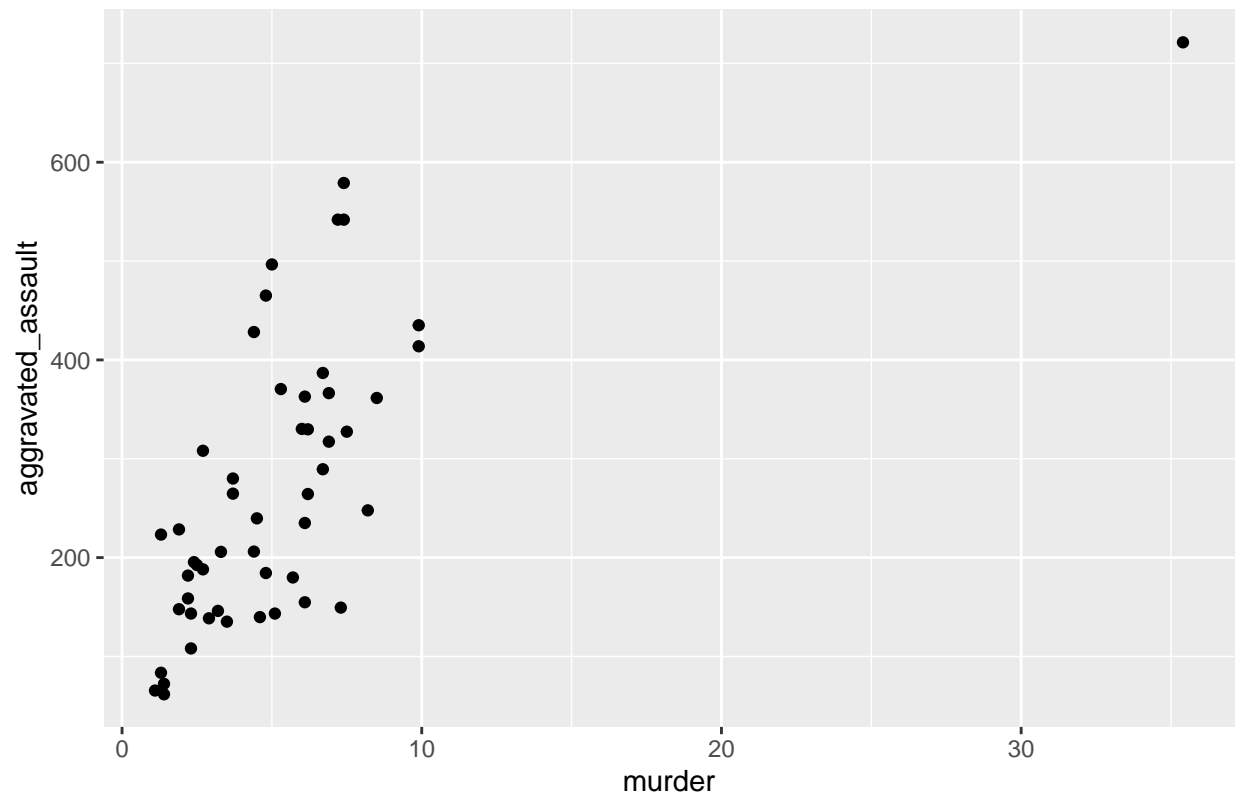
```
##           state murder forcible_rape robbery aggravated_assault burglary
## 1 United States   5.6           31.7   140.7           291.1     726.7
## 2      Alabama   8.2           34.3   141.4           247.8     953.8
## 3       Alaska   4.8           81.1    80.9           465.1     622.5
## 4      Arizona   7.5           33.8   144.4           327.4     948.4
## 5     Arkansas   6.7           42.9    91.1           386.8    1084.6
## 6    California   6.9           26.0   176.1           317.3     693.3
##  larceny_theft motor_vehicle_theft population
## 1         2286.3           416.7  295753151
## 2         2650.0           288.3   4545049
## 3         2599.1           391.0    669488
## 4         2965.2           924.4   5974834
## 5         2711.2           262.1   2776221
## 6         1916.5           712.8  35795255
```

```
crimerates_df <- crimerates_df[-1, ]
```

Plot murder vs aggravated assault

```
library(ggplot2)
scatter_murder_assault <- ggplot(crimerates_df, aes(x=murder, y=aggravated_assault)) +
  geom_point() + labs(title = 'Scatterplot of Murder vs Aggravated Assault in R')
scatter_murder_assault
```

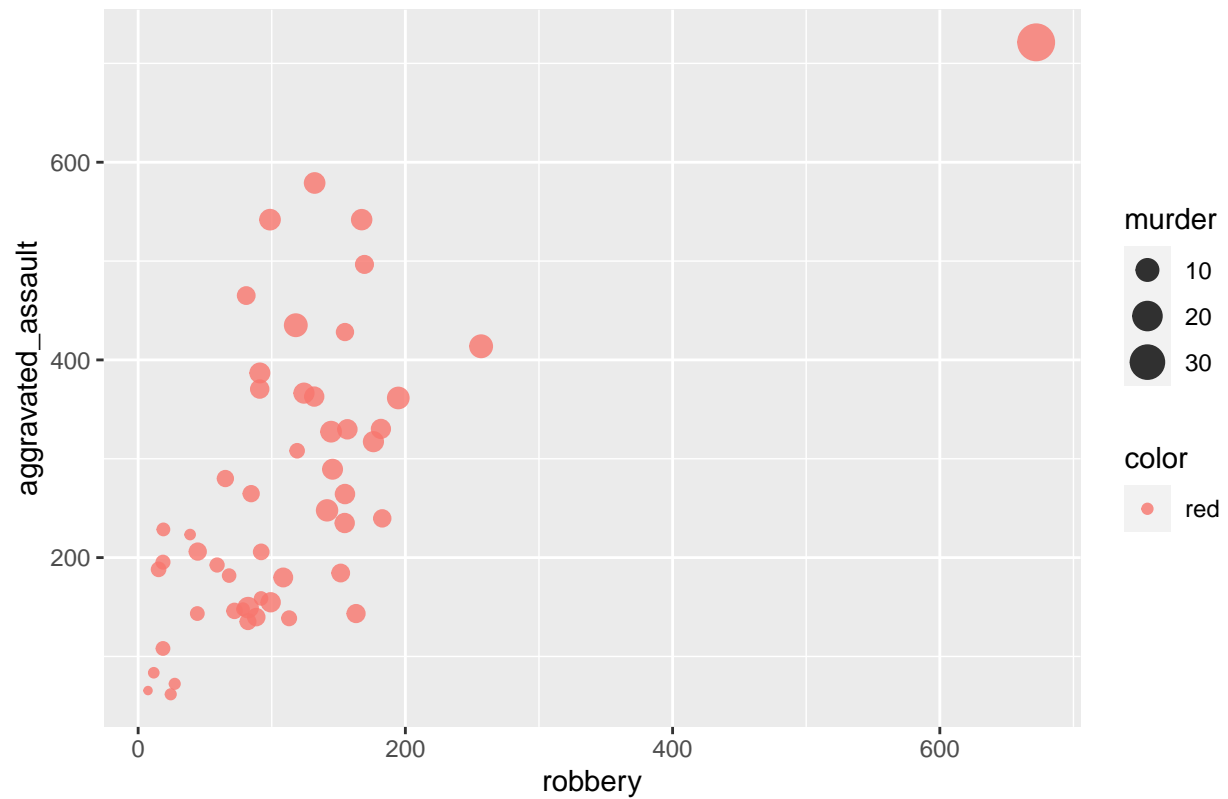
Scatterplot of Murder vs Aggravated Assault in R



Bubble chart in R

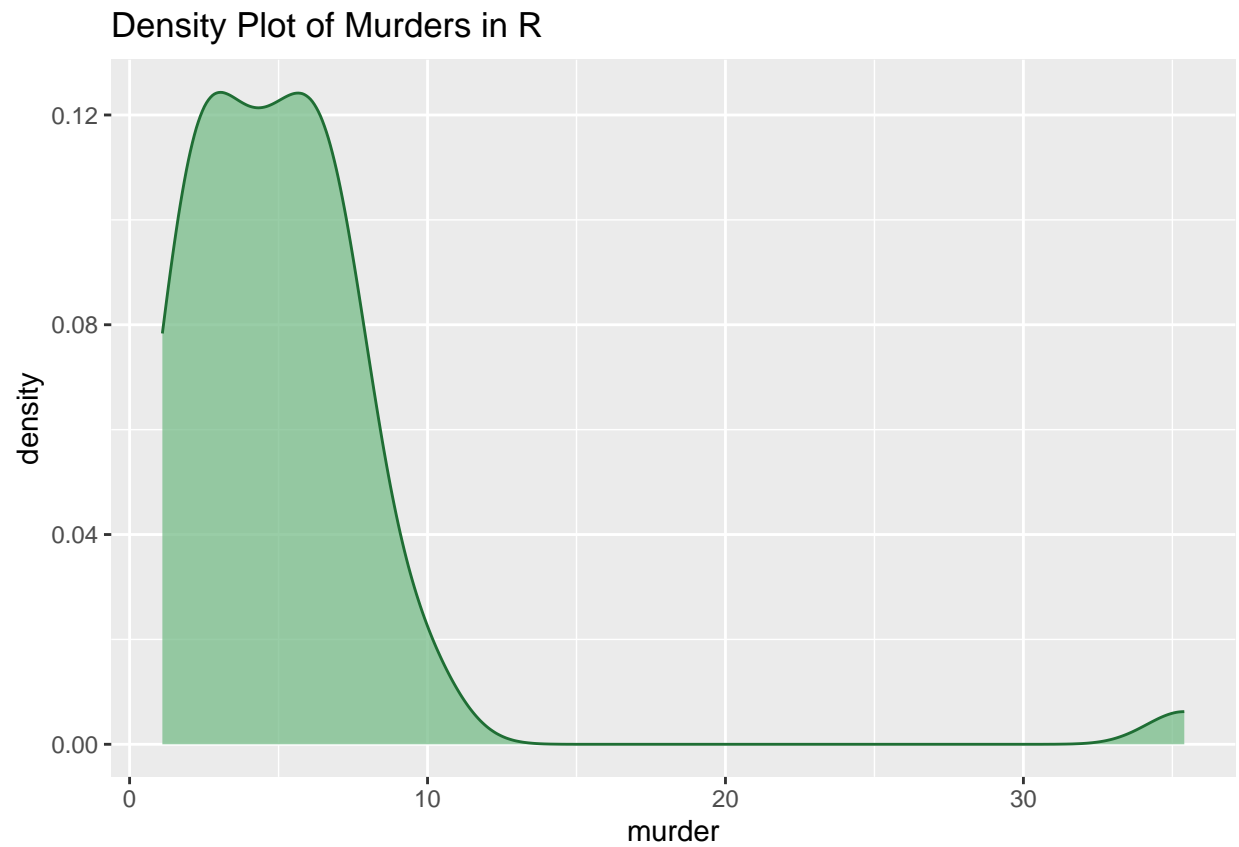
```
library(ggplot2)
color <- 'red'
bubble_crimes <- ggplot(crimerates_df, aes(x=robbery, y=aggravated_assault, size=murder,
                                             color=color)) + geom_point(alpha=0.8) +
  labs(title = 'Bubble Chart Robbery vs Aggravated Assault by Murder')
bubble_crimes
```

Bubble Chart Robbery vs Aggravated Assault by Murder



Density plot in R

```
library(ggplot2)
density_murder_plot <- ggplot(crimerates_df, aes(x=murder)) + geom_density(fill='#77bd89',
                                                                              color='#1f6e34',
                                                                              alpha=0.75) +
  labs(title = 'Density Plot of Murders in R')
density_murder_plot
```



Merge pdfs for exercise 3.2

```
library(pdftools)
```

```
## Using poppler version 22.04.0
```

```
# pdf_combine(c())
```

# Scatterplot in Python

```
In [1]: import pandas as pd
crimerates_df = pd.read_csv('crimerates-by-state-2005.csv')
```

```
In [2]: crimerates_df.info()
```

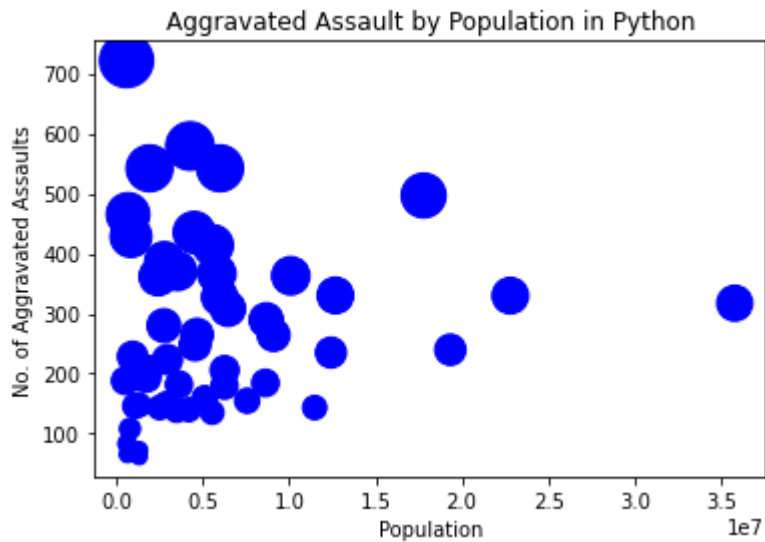
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 52 entries, 0 to 51
Data columns (total 9 columns):
#   Column                Non-Null Count  Dtype
---  -
0   state                 52 non-null    object
1   murder               52 non-null    float64
2   forcible_rape        52 non-null    float64
3   robbery              52 non-null    float64
4   aggravated_assault    52 non-null    float64
5   burglary             52 non-null    float64
6   larceny_theft         52 non-null    float64
7   motor_vehicle_theft   52 non-null    float64
8   population            52 non-null    int64
dtypes: float64(7), int64(1), object(1)
memory usage: 3.8+ KB
```

```
In [3]: # Remove 1st row which includes entire U.S.
crimerates_df = crimerates_df.tail(-1)
crimerates_df.head()
```

```
Out[3]:
```

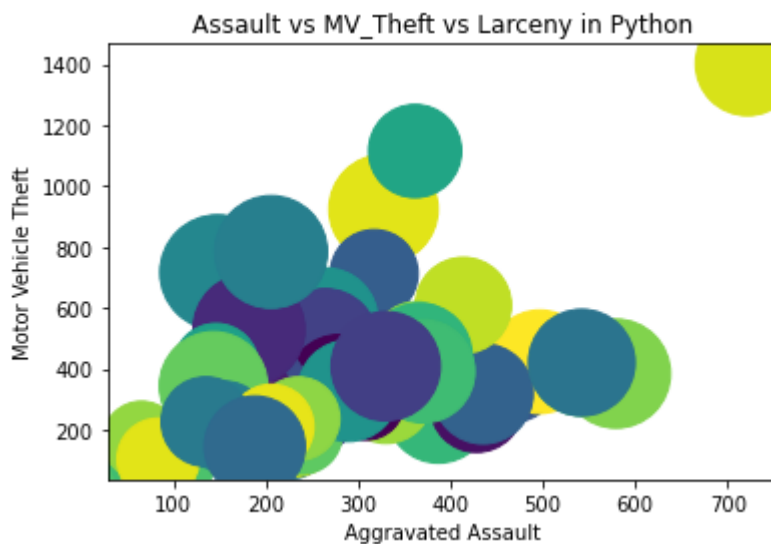
	state	murder	forcible_rape	robbery	aggravated_assault	burglary	larceny_theft	motor_vehic
1	Alabama	8.2	34.3	141.4	247.8	953.8	2650.0	
2	Alaska	4.8	81.1	80.9	465.1	622.5	2599.1	
3	Arizona	7.5	33.8	144.4	327.4	948.4	2965.2	
4	Arkansas	6.7	42.9	91.1	386.8	1084.6	2711.2	
5	California	6.9	26.0	176.1	317.3	693.3	1916.5	

```
In [4]: import matplotlib.pyplot as plt
x = crimerates_df['population']
y = crimerates_df['aggravated_assault']
plt.scatter(x = x, y = y, s = y, c = 'blue')
plt.title('Aggravated Assault by Population in Python')
plt.xlabel('Population')
plt.ylabel('No. of Aggravated Assaults')
plt.show()
```



## Bubble Chart in Python

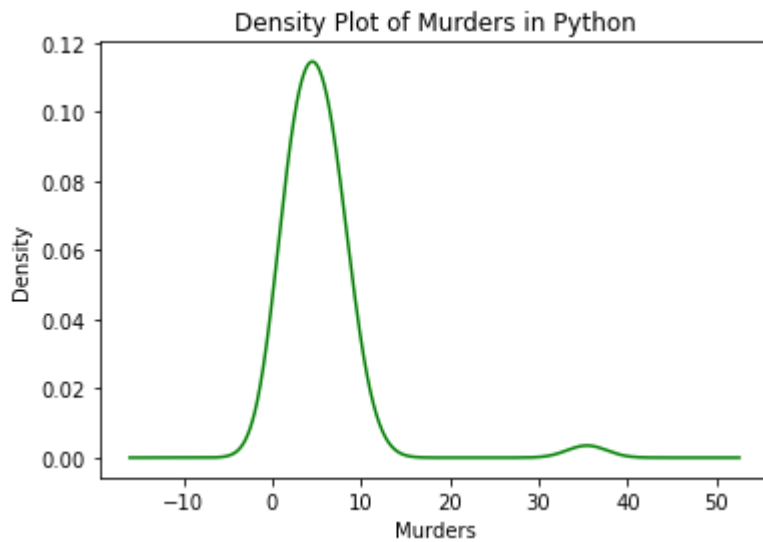
```
In [5]: import matplotlib.pyplot as plt
import numpy as np
x = crimerates_df['aggravated_assault']
y = crimerates_df['motor_vehicle_theft']
z = crimerates_df['larceny_theft']
colors = np.random.rand(51)
plt.scatter(x, y, s=z, c = colors)
plt.title('Assault vs MV_Theft vs Larceny in Python')
plt.xlabel('Aggravated Assault')
plt.ylabel('Motor Vehicle Theft')
plt.show()
```



## Density Plot in Python

```
In [6]: crimerates_df.murder.plot.density(color='green')
plt.title('Density Plot of Murders in Python')
```

```
plt.xlabel('Murders')  
plt.show()
```



In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: