DSC 640: WEEK 7&8

Author: Yousof Rahimian

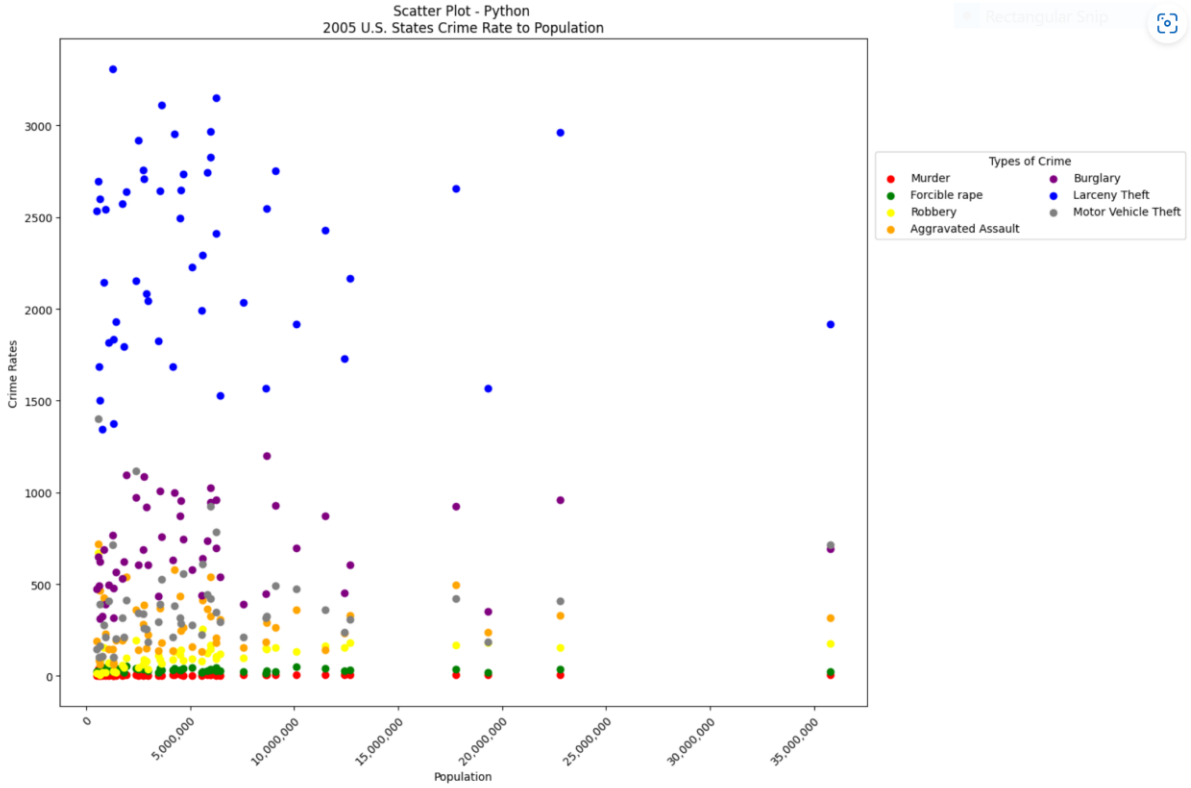
Date: 2/5/2023

Exercise 4.2 Scatter Plot, Bubble Plot, and Density Plot/Map

Scatterplot

Pyhton

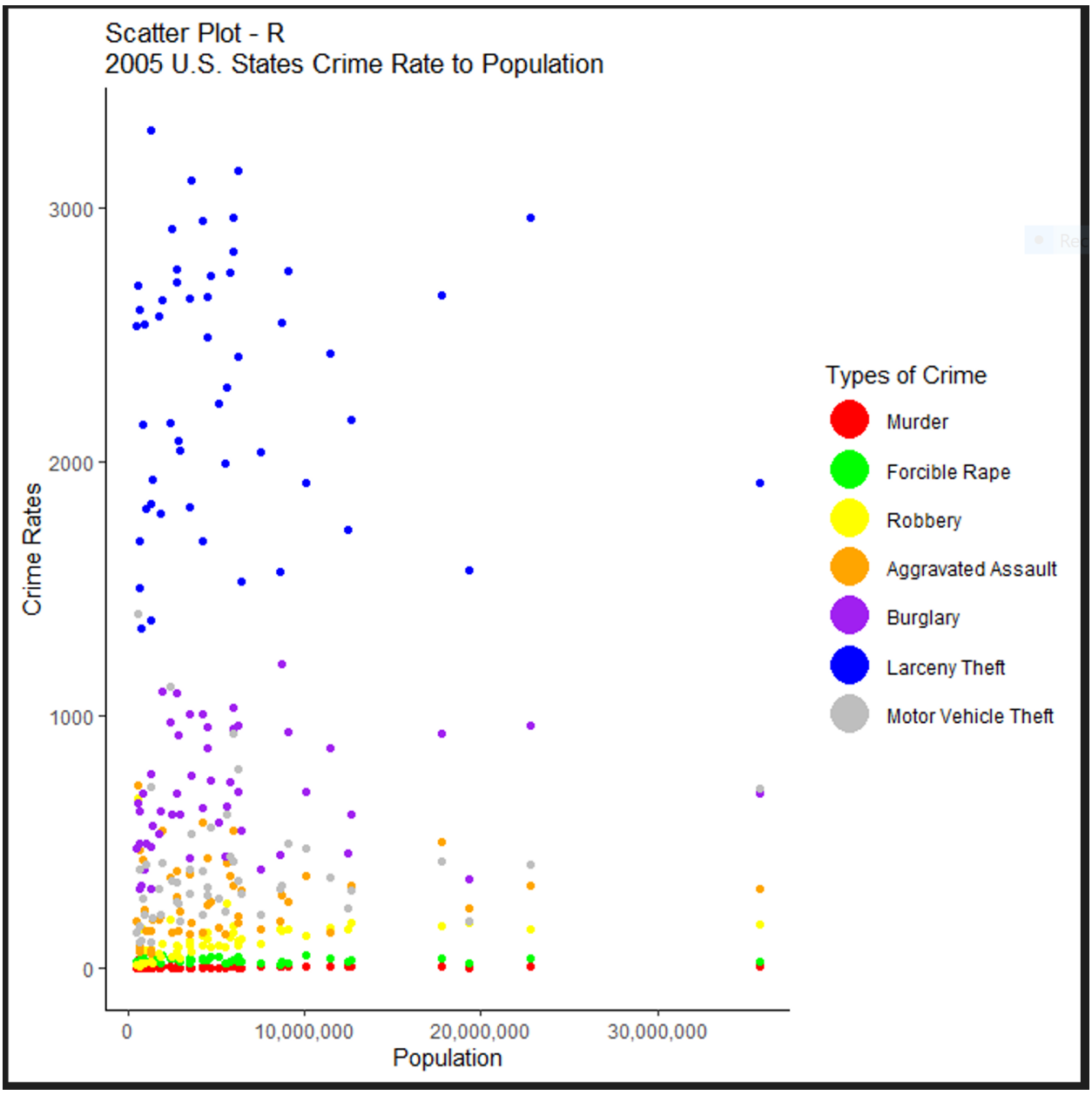
```python
# display the state crime rate based on population of each estates
# Use scatterplot

fig, ax = plt.subplots(figsize = (15, 10))

plt.scatter(states_crime['population'], states_crime['murder'], c = 'red', label = 'Murder')
plt.scatter(states_crime['population'], states_crime['forcible_rape'], c = 'green', label = 'Forcible rape')
plt.scatter(states_crime['population'], states_crime['robbery'], c = 'yellow', label = 'Robbery')
plt.scatter(states_crime['population'], states_crime['aggravated_assault'], c = 'orange', label = 'Aggravated Assault')
plt.scatter(states_crime['population'], states_crime['burglary'], c = 'purple', label = 'Burglary')
plt.scatter(states_crime['population'], states_crime['larceny_theft'], c = 'blue', label = 'Larceny Theft')
plt.scatter(states_crime['population'], states_crime['motor_vehicle_theft'], c = 'gray', label = 'Motor Vehicle Theft')

ax.get_xaxis().set_major_formatter(
    tick.FuncFormatter(lambda x, p: format(int(x), ',')))

plt.xlabel("Population")
plt.xticks(rotation = 45, ha = 'right', rotation_mode = 'anchor')
plt.ylabel("Crime Rates")
plt.title("Scatter Plot - Python \n2005 U.S. States Crime Rate to Population")

ax.legend(loc=(1.01, 0.7), ncol = 2, title = 'Types of Crime')
plt.tight_layout()

plt.show()
```



R

```r
#| label: mutatestates
crimes_df <- melt(states_crime, id.vars = c('state', 'population'), variable.name = 'crimes')
crimes_df
```

```r
#| label: scatterplot
#| echo: false
fig <- crimes_df %>%
  ggplot(aes(x = population, y = value, color = crimes)) +
  geom_point() +
  ggtitle("Scatter Plot - R \n2005 U.S. States Crime Rate to Population") +
  guides(size = 'none') +
  xlab("Population") +
  ylab("Crime Rates") +
  scale_x_continuous(labels = scales::comma) +
  guides(color = guide_legend(title = "Types of Crime", override.aes = list(size = 8))) +
  scale_color_manual(values = c("murder" = "red",
                                "forcible_rape" = "green",
                                "robbery" ="yellow",
                                "aggravated_assault" = "orange",
                                "burglary" = "purple",
                                "larceny_theft" = "blue",
                                "motor_vehicle_theft" = "gray"),
                     labels = c("Murder", "Forcible Rape", "Robbery", "Aggravated Assault",
                                "Burglary", "Larceny Theft", "Motor Vehicle Theft"))
fig
```


Scatter Plot - R
2005 U.S. States Crime Rate to Population
Crime Rates
3000
2000
1000
0
0
10,000,000
20,000,000
30,000,000
Population
Types of Crime
Murder
Forcible Rape
Robbery
Aggravated Assault
Burglary
Larceny Theft
Motor Vehicle Theft

Tableau

Bubble Plot

Python

In [7]:

```
fig, ax = plt.subplots(figsize = (20, 3))

plt.scatter(us_crime['murder'], us_crime['state'],
            c = 'purple', label = 'Murder', s = us_crime['murder'] * 5,
           alpha = 0.5)
plt.scatter(us_crime['forcible_rape'], us_crime['state'],
            c = 'green', label = 'Forcible rape', s = us_crime['forcible_rape'] * 2,
           alpha = 0.5)
plt.scatter(us_crime['robbery'], us_crime['state'],
            c = 'yellow', label = 'Robbery', s = us_crime['robbery'] * 2,
           alpha = 0.5)
plt.scatter(us_crime['aggravated_assault'], us_crime['state'],
            c = 'orange', label = 'Aggravated Assault', s = us_crime['aggravated_assault'] * 2,
           alpha = 0.5)
plt.scatter(us_crime['burglary'], us_crime['state'],
            c = 'blue', label = 'Burglary', s = us_crime['burglary'] * 2,
           alpha = 0.5)
plt.scatter(us_crime['larceny_theft'], us_crime['state'],
            c = 'red', label = 'Larceny Theft', s = us_crime['larceny_theft'] * 2,
           alpha = 0.5)
plt.scatter(us_crime['motor_vehicle_theft'], us_crime['state'],
            c = 'gray', label = 'Motor Vehicle Theft', s = us_crime['motor_vehicle_theft'] * 2,
           alpha = 0.5)

plt.xlabel("Crime Rates")
plt.title("Bubble Chart - Python \n2005 U.S. Crime Rate")

crimes_list = ["Murder", "Forcible Rape", "Robbery", "Aggravated Assault", "Burglary",
               "Larceny Theft", "Motor Vehicle Theft"]
color_list = ["purple", "green", "yellow", "orange", "blue", "red", "gray"]
legend_list = []
for i in range(0, len(crimes_list)):
    legend_list.append(mpatches.Patch(color = color_list[i],
                               alpha = 0.5,
                               label = crimes_list[i]))
ax.legend(handles = legend_list, loc = (1.05, 0), title = 'Types of Crime')

plt.tight_layout()

plt.show()
```

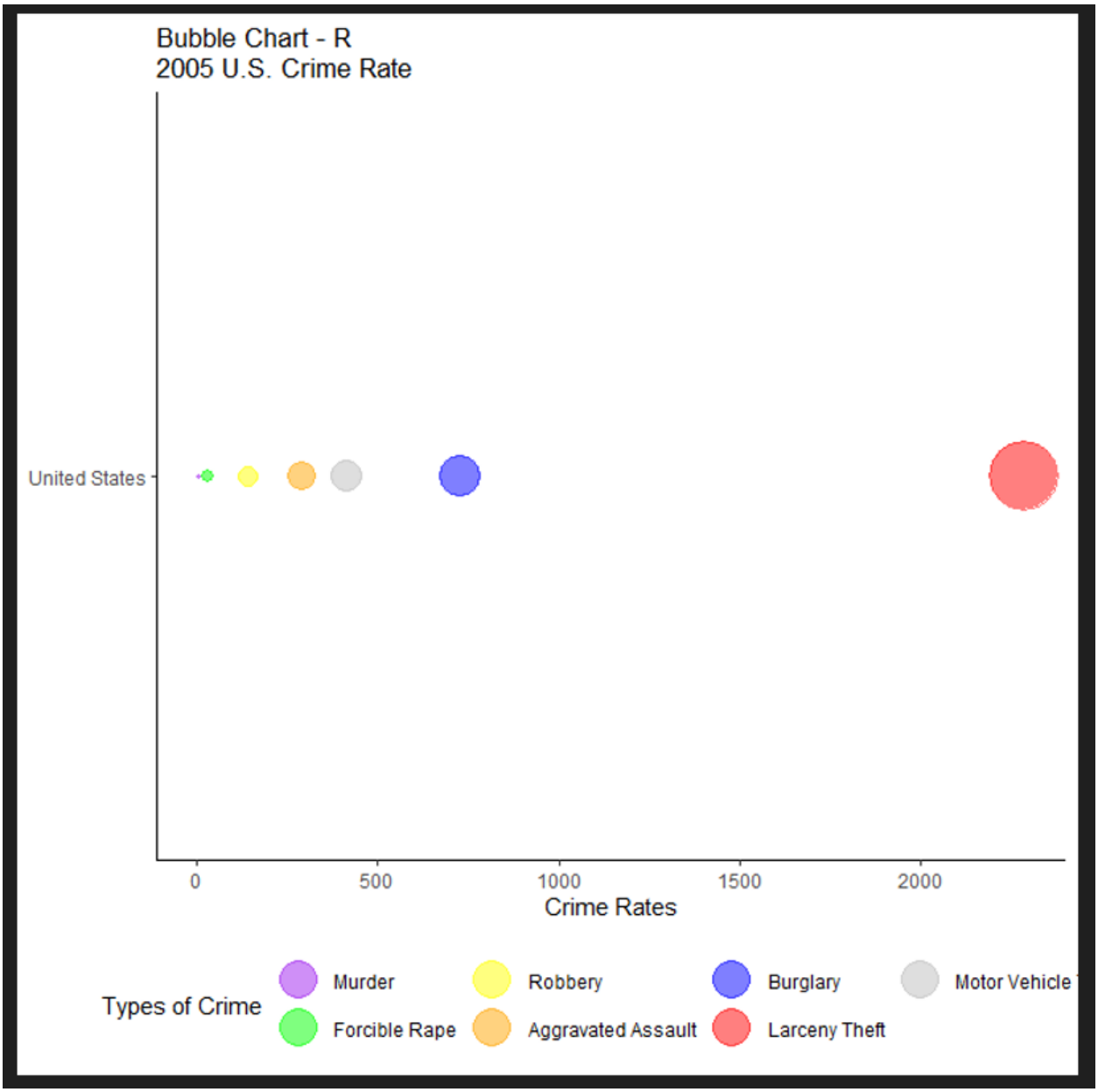
R

```r
### 2. Bubble Chart

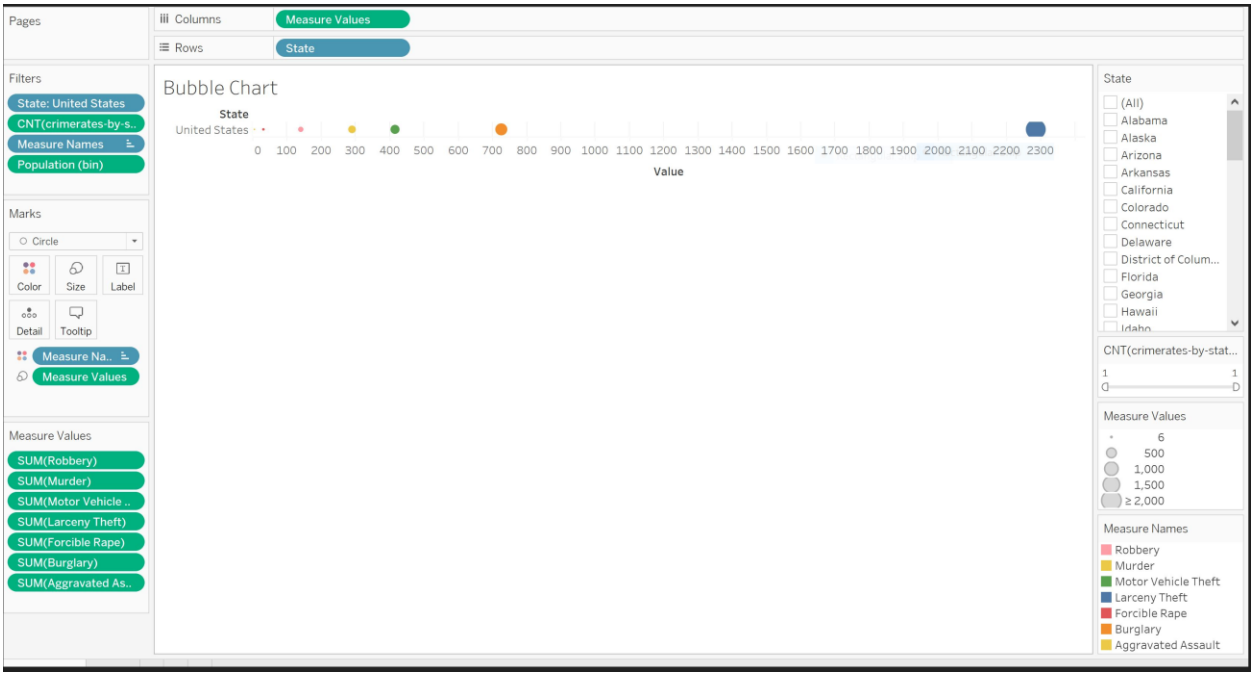
```{r}
#| label: uscrime
us_crime <- crimerate_df[crimerate_df$state == 'United States', ]
us_crime
```

```{r}
#| label: mutateus
uscrimes_df <- melt(us_crime, id.vars = c('state', 'population'), variable.name = 'crimes')
uscrimes_df
```

```{r}
#| label: bubblechart
#| echo: false
fig <- uscrimes_df %>%
  ggplot(aes(x = value, y = state)) +
  geom_point(aes(color = crimes, size = value), alpha = 0.5) +
  scale_size(range = c(0.8, 15), guide = 'none') +
  guides(color = guide_legend(title = "Types of Crime", override.aes = list(size = 8))) +
  scale_color_manual(values = c("murder" = "purple",
                                "forcible_rape" = "green",
                                "robbery" ="yellow",
                                "aggravated_assault" = "orange",
                                "burglary" = "blue",
                                "larceny_theft" = "red",
                                "motor_vehicle_theft" = "gray"),
                     labels = c("Murder", "Forcible Rape", "Robbery", "Aggravated Assault",
                                "Burglary", "Larceny Theft", "Motor Vehicle Theft")) +
  ggtitle("Bubble Chart - R \n2005 U.S. Crime Rate") +
  xlab("Crime Rates") +
  ylab(NULL) +
  theme(legend.position="bottom")
fig
```
```


Bubble Chart - R
2005 U.S. Crime Rate
United States
0
500
1000
1500
2000
Crime Rates
Types of Crime
Murder
Robbery
Burglary
Motor Vehicle
Forcible Rape
Aggravated Assault
Larceny Theft

Tableau



Density Plot/Map

Python

```
density = gaussian_kde(states_crime_df['burglary'])

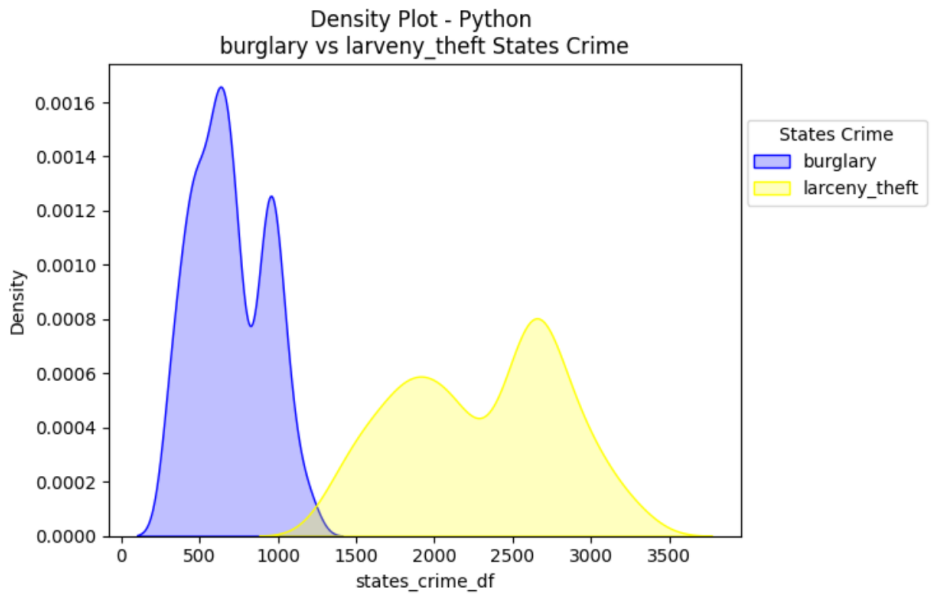
fig, ax = plt.subplots()

sb.kdeplot(states_crime_df['burglary'] , bw_method = 0.3 , fill = True, color = 'blue', label = 'burglary')
sb.kdeplot(larceny_theft_df['larceny_theft'] , bw_method = 0.3 , fill = True, color = 'yellow', label = 'larceny_theft')

plt.title('Density Plot - Python \nburglary vs larveny_theft States Crime')
plt.xlabel('states_crime_df')

ax.legend(loc=(1.01, 0.7), ncol = 1, title = 'States Crime')

plt.show()
```

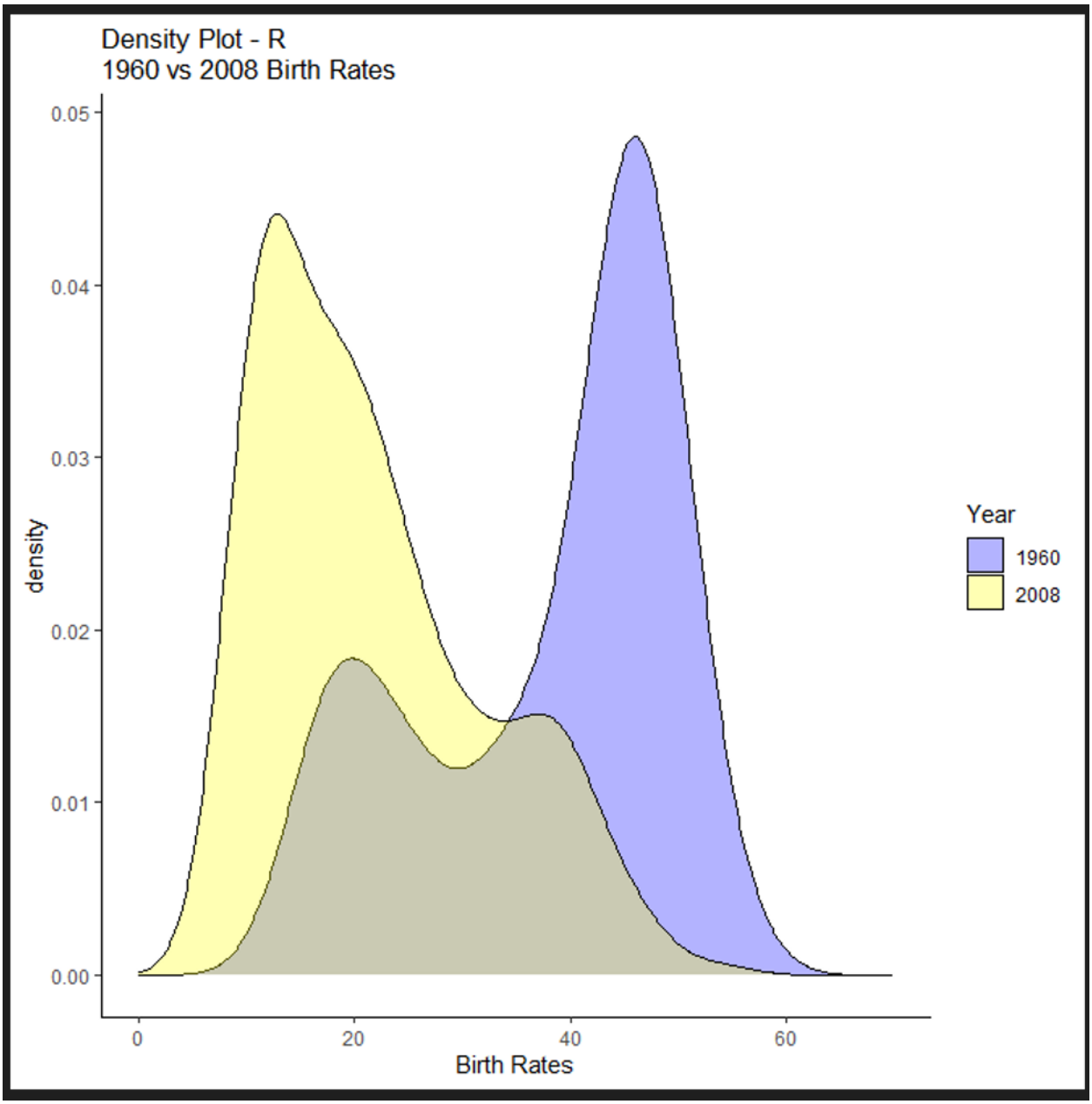


R

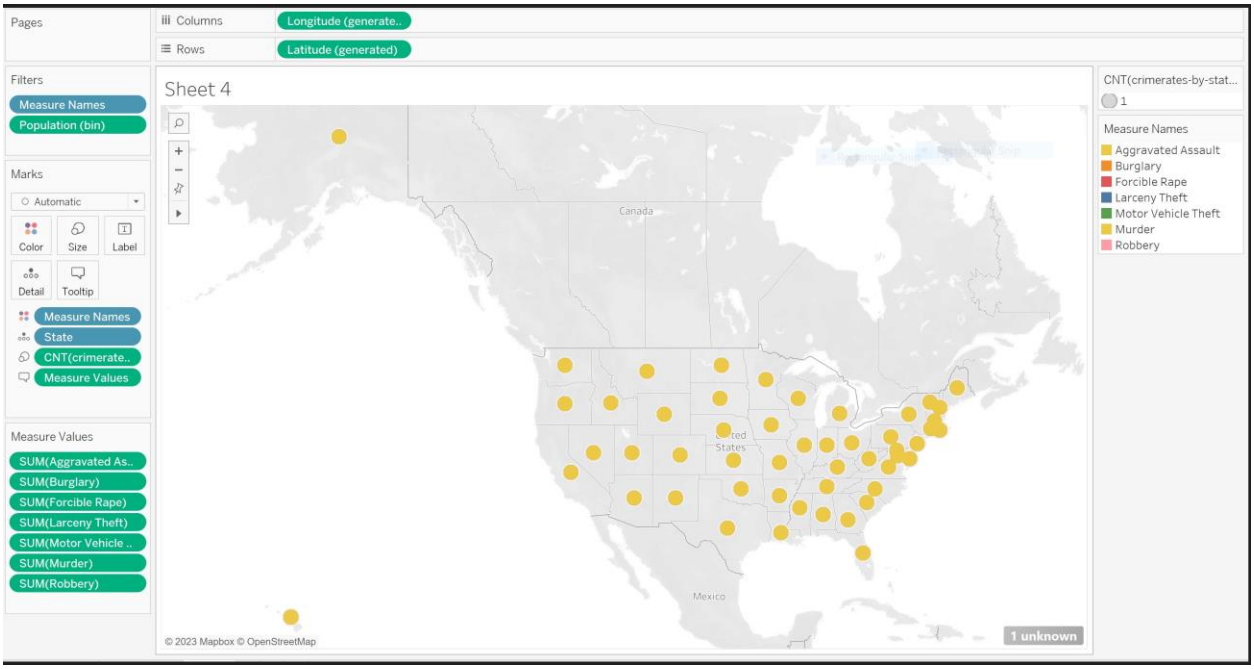
```
### 3. Density Plot

births_df <- birthrate_df %>%
  filter(year == 2008 | year == 1960)
births_df <- births_df[order(births_df$year), ]
births_df$year <- as.character(births_df$year)
births_df
```

```{r}
#| label: densityplot
#| echo: false
fig <- births_df %>%
  ggplot(aes(x = rate, fill = year)) +
  geom_density(alpha = 0.3) +
  ggtitle("Density Plot - R\n1960 vs 2008 Birth Rates") +
  labs(fill = 'Year') +
  scale_x_continuous(name="Birth Rates", limits=c(0, 70)) +
  scale_fill_manual(values = c("blue", "yellow"))
fig
```


Density Plot - R
1960 vs 2008 Birth Rates
0.05
0.04
0.03
0.02
0.01
0.00
density
0
20
40
60
Birth Rates
Year
1960
2008

Tableau


Pages
Columns
Longitude (generate..
Rows
Latitude (generated)
Filters
Measure Names
Population (bin)
Marks
Automatic
Color
Size
Label
Detail
Tooltip
Measure Names
State
CNT(crimerate..
Measure Values
Measure Values
SUM(Aggravated As..
SUM(Burglary)
SUM(Forcible Rape)
SUM(Larceny Theft)
SUM(Motor Vehicle ..
SUM(Murder)
SUM(Robbery)
Sheet 4
Canada
Mexico
© 2023 Mapbox © OpenStreetMap
1 unknown
CNT(crimerates-by-stat...
1
Measure Names
Aggravated Assault
Burglary
Forcible Rape
Larceny Theft
Motor Vehicle Theft
Murder
Robbery