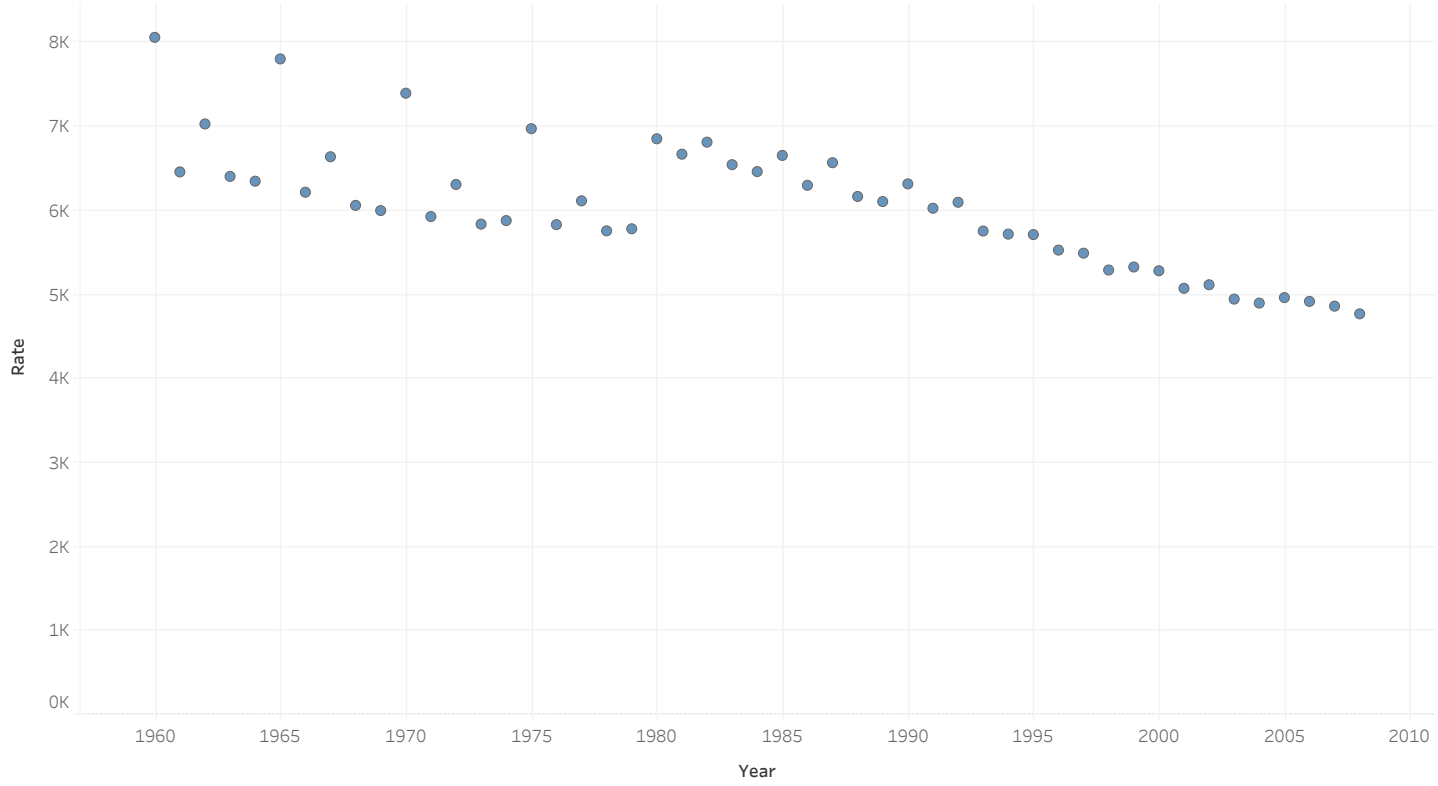
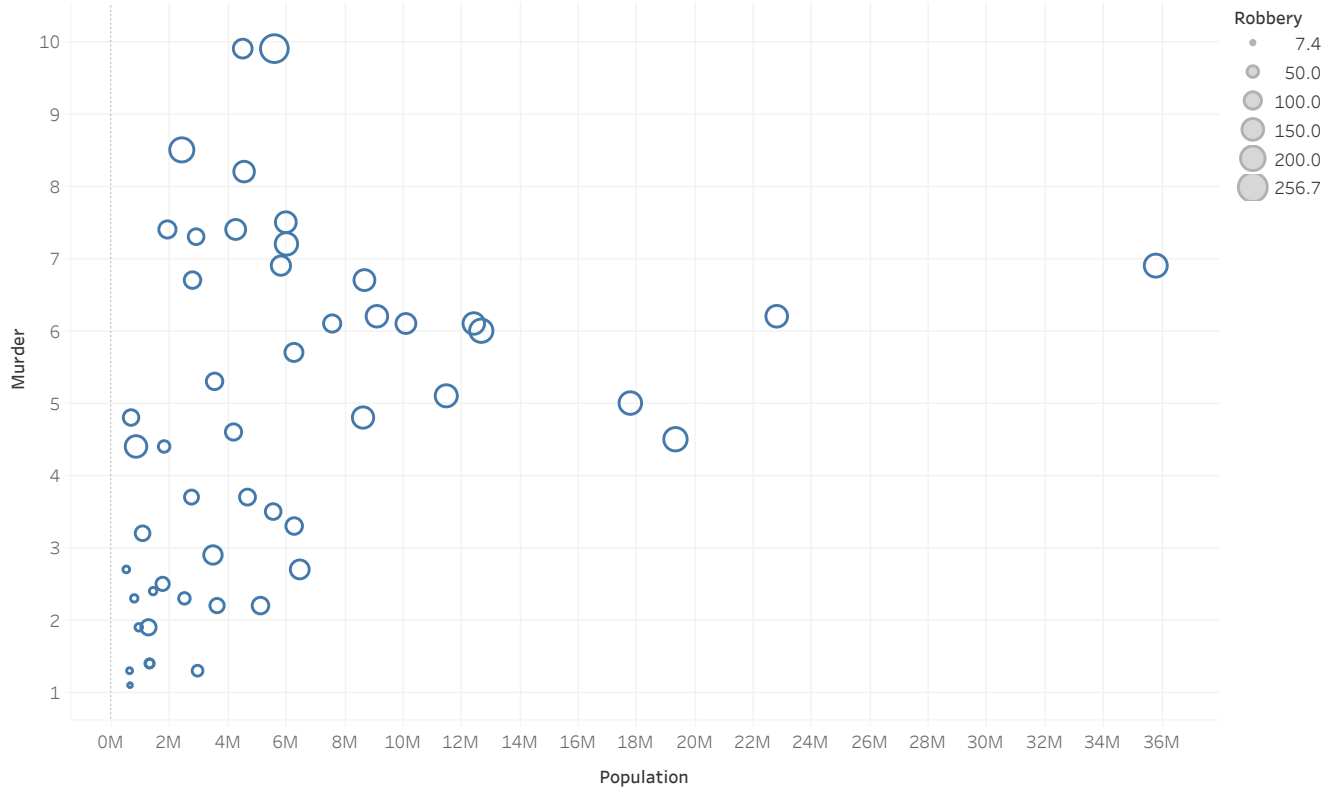
Scatter Plot W7-8



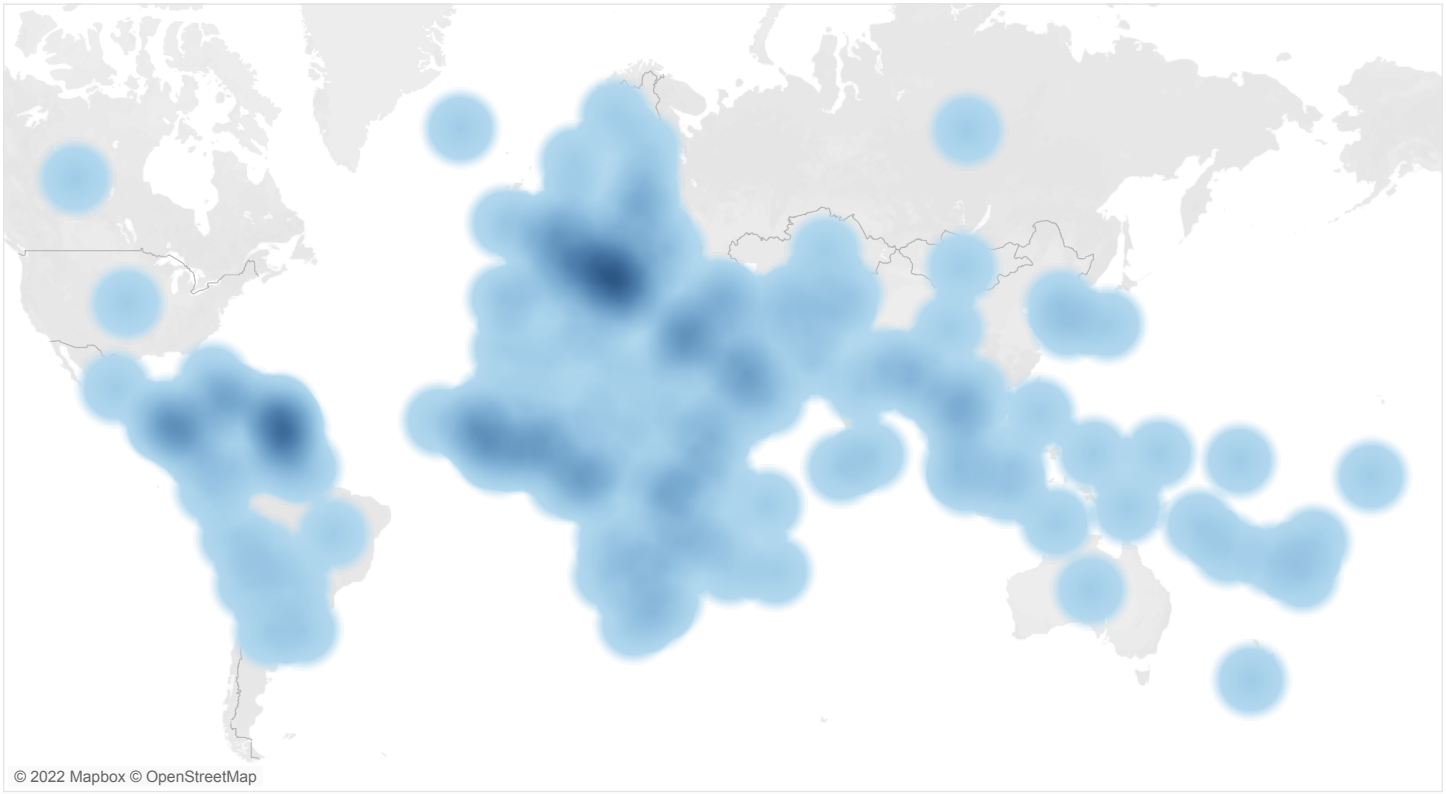
The plot of sum of Rate for Year.

Bubble Chart



Population vs. Murder. Size shows Robbery as an attribute. The data is filtered on State, which has multiple members selected.

Density Map



Map based on Longitude (generated) and Latitude (generated). Details are shown for Country. The data is filtered on Expectancy, which ranges from 42 to 83.

In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

In [13]:
```
df=pd.read_csv("birth-rates-yearly.csv")
df
```

Out[13]:

	year	rate
0	1960	36.400
1	1961	35.179
2	1962	33.863
3	1963	32.459
4	1964	30.994
...	...	...
9865	2004	30.123
9866	2005	30.067
9867	2006	30.027
9868	2007	29.987
9869	2008	29.930

9870 rows × 2 columns

In [19]:
```
avg = df.groupby('year')[['rate']].mean().reset_index()
avg
```

Out[19]:

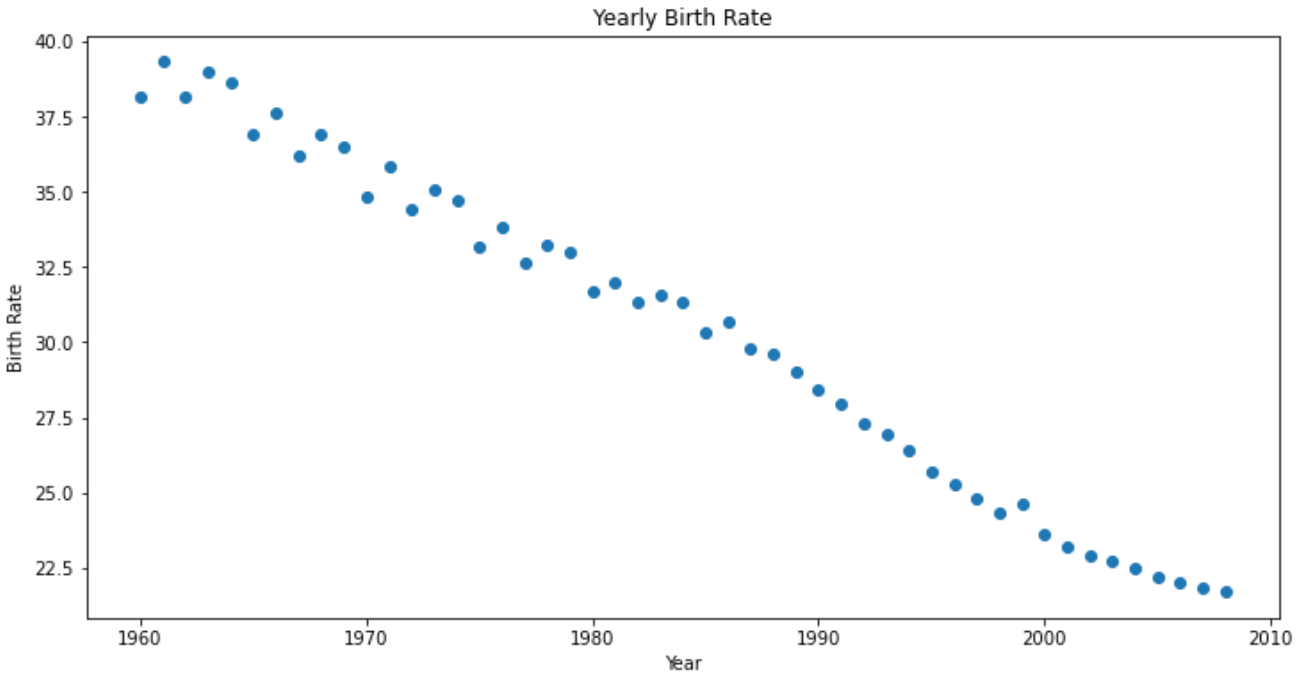
	year	rate
0	1960	38.143298
1	1961	39.309189
2	1962	38.139809
3	1963	38.983646
4	1964	38.636152
5	1965	36.927179
6	1966	37.605933
7	1967	36.212692

8	1968	36.880024
9	1969	36.501018
10	1970	34.828550
11	1971	35.851927
12	1972	34.412160
13	1973	35.090506
14	1974	34.718621
15	1975	33.149989
16	1976	33.832035
17	1977	32.632433
18	1978	33.209006
19	1979	32.968371
20	1980	31.671848
21	1981	32.008343
22	1982	31.340386
23	1983	31.559597
24	1984	31.313783
25	1985	30.332996
26	1986	30.669408
27	1987	29.801436
28	1988	29.589116
29	1989	29.018110
30	1990	28.397586
31	1991	27.972156
32	1992	27.289441
33	1993	26.961650
34	1994	26.417772
35	1995	25.679301
36	1996	25.302754
37	1997	24.793530
38	1998	24.327411

39	1999	24.606884
40	2000	23.630850
41	2001	23.215782
42	2002	22.885726
43	2003	22.731459
44	2004	22.513075
45	2005	22.198215
46	2006	21.997903
47	2007	21.842214
48	2008	21.722674

In [20]:

```
plt.figure(figsize=(12, 6))
plt.scatter(avg['year'], avg['rate'])           # scatter plot showing actual
plt.xlabel('Year')
plt.ylabel('Birth Rate')
plt.title('Yearly Birth Rate')
plt.show()
```



In [31]:

```
df1=pd.read_csv('crimerates-by-state-2005.csv')
df1=df1.drop([0,9])
df1
```

Out[31]:

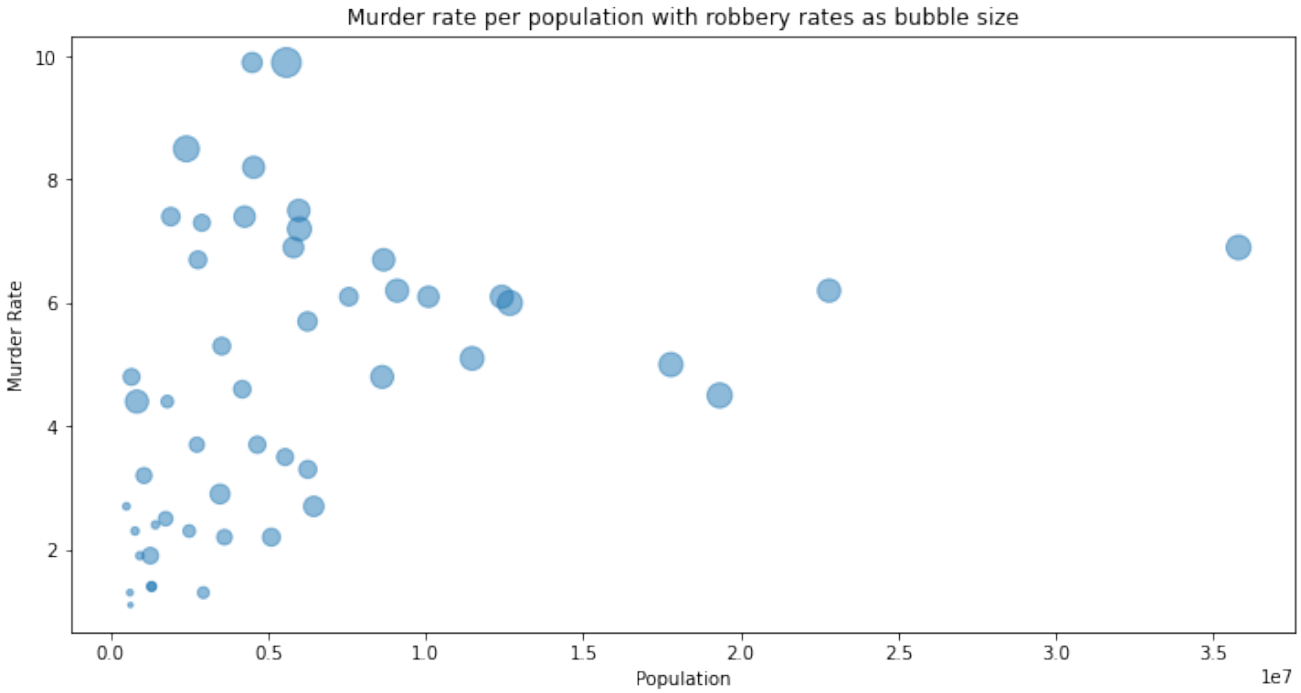
	state	murder	forcible_rape	robbery	aggravated_assault	burglary	larceny_theft

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
6	Colorado	3.7	43.4	84.6	264.7	744.8	2735.2
7	Connecticut	2.9	20.0	113.0	138.6	437.1	1824.1
8	Delaware	4.4	44.7	154.8	428.2	688.9	2144.0
10	Florida	5.0	37.1	169.4	496.6	926.3	2658.3
11	Georgia	6.2	23.6	154.8	264.3	931.0	2751.1
12	Hawaii	1.9	26.9	78.5	147.8	767.9	3308.4
13	Idaho	2.4	40.4	18.6	195.4	564.4	1931.7
14	Illinois	6.0	33.7	181.7	330.2	606.9	2164.8
15	Indiana	5.7	29.6	108.6	179.9	697.6	2412.0
16	Iowa	1.3	27.9	38.9	223.3	606.4	2042.7
17	Kansas	3.7	38.4	65.3	280.0	689.2	2758.1
18	Kentucky	4.6	34.0	88.4	139.8	634.0	1685.8
19	Louisiana	9.9	31.4	118.0	435.1	870.6	2494.5
20	Maine	1.4	24.7	24.4	61.7	478.5	1832.6
21	Maryland	9.9	22.6	256.7	413.8	641.4	2294.3
22	Massachusetts	2.7	27.1	119.0	308.1	541.1	1527.4
23	Michigan	6.1	51.3	131.8	362.9	696.8	1917.8
24	Minnesota	2.2	44.0	92.0	158.7	578.9	2226.9
25	Mississippi	7.3	39.3	82.3	149.4	919.7	2083.9
26	Missouri	6.9	28.0	124.1	366.4	738.3	2746.2
27	Montana	1.9	32.2	18.9	228.5	389.2	2543.0
28	Nebraska	2.5	32.9	59.1	192.5	532.4	2574.3
29	Nevada	8.5	42.1	194.7	361.5	972.4	2153.9
30	New Hampshire	1.4	30.9	27.4	72.3	317.0	1377.3
31	New Jersey	4.8	13.9	151.6	184.4	447.1	1568.4
32	New Mexico	7.4	54.1	98.7	541.9	1093.9	2639.9

33	New York	4.5	18.9	182.7	239.7	353.3	1569.6
34	North Carolina	6.7	26.5	145.5	289.4	1201.1	2546.2
35	North Dakota	1.1	24.2	7.4	65.5	311.9	1500.3
36	Ohio	5.1	39.8	163.1	143.4	872.8	2429.0
37	Oklahoma	5.3	41.7	91.0	370.5	1006.0	2644.2
38	Oregon	2.2	34.8	68.1	181.8	758.6	3112.2
39	Pennsylvania	6.1	28.9	154.6	235.0	451.6	1729.1
40	Rhode Island	3.2	29.8	72.1	146.1	494.2	1816.0
41	South Carolina	7.4	42.5	132.1	579.0	1000.9	2954.1
42	South Dakota	2.3	46.7	18.6	108.1	324.4	1343.7
43	Tennessee	7.2	36.4	167.3	541.9	1026.9	2828.1
44	Texas	6.2	37.2	156.6	329.8	961.6	2961.7
45	Utah	2.3	37.3	44.3	143.4	606.2	2918.8
46	Vermont	1.3	23.3	11.7	83.5	491.8	1686.1
47	Virginia	6.1	22.7	99.2	154.8	392.1	2035.0
48	Washington	3.3	44.7	92.1	205.8	959.7	3149.5
49	West Virginia	4.4	17.7	44.6	206.1	621.2	1794.0
50	Wisconsin	3.5	20.6	82.2	135.2	440.8	1992.8
51	Wyoming	2.7	24.0	15.3	188.1	476.3	2533.9

In [33]:

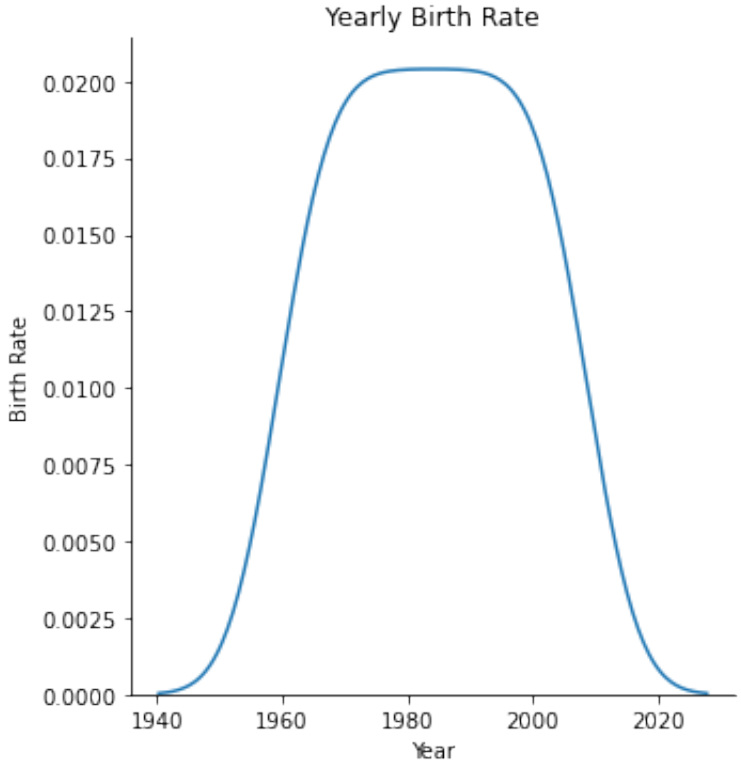
```
plt.figure(figsize=(12, 6))
plt.scatter(df1['population'], df1['murder'], df1['robbery'], alpha=0.5)
plt.xlabel('Population')
plt.ylabel('Murder Rate')
plt.title('Murder rate per population with robbery rates as bubble size')
plt.show()
```



In [36]:
```
import seaborn as sns
```

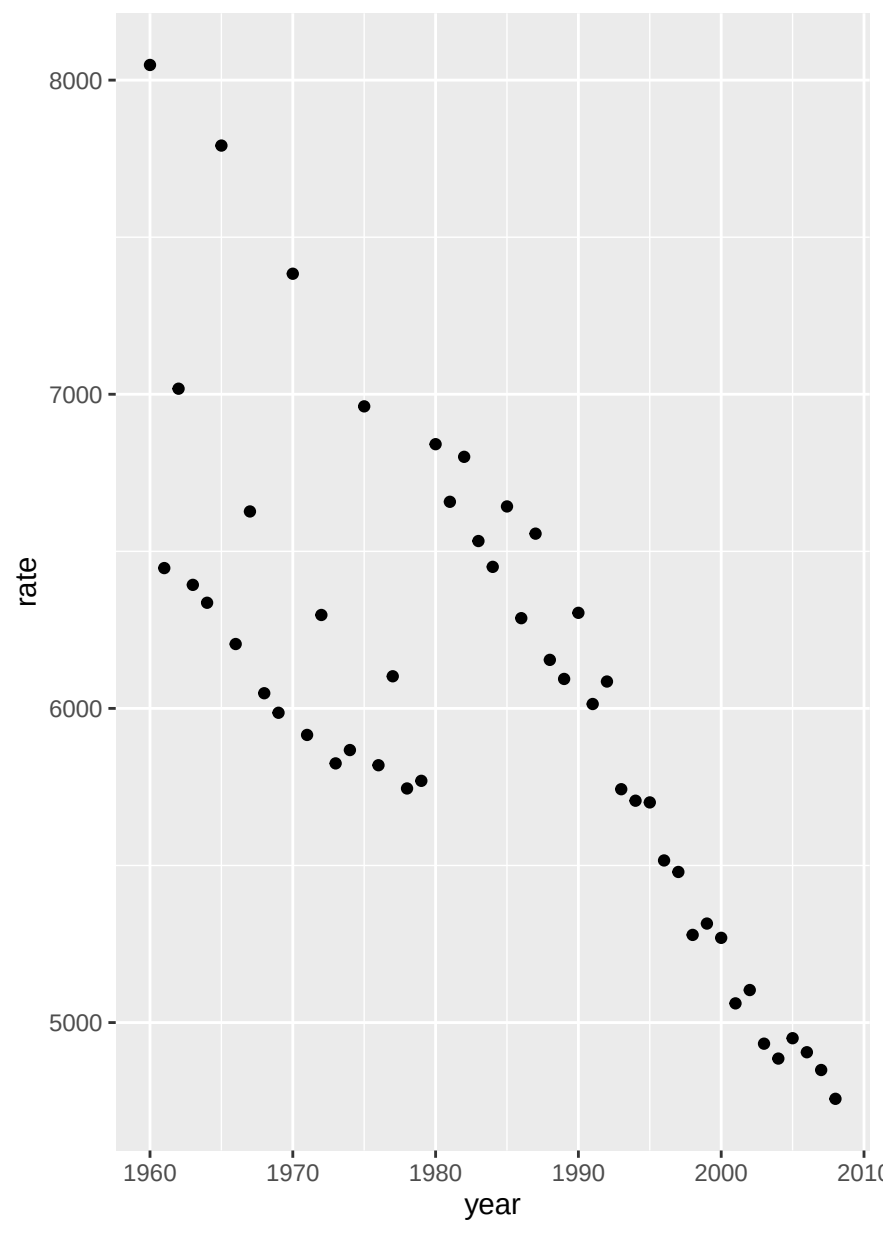
In [44]:
```
plt.figure(figsize=(12, 6))
sns.displot(avg, x="year", kind="kde")
plt.xlabel('Year')
plt.ylabel('Birth Rate')
plt.title('Yearly Birth Rate')
plt.show()
```

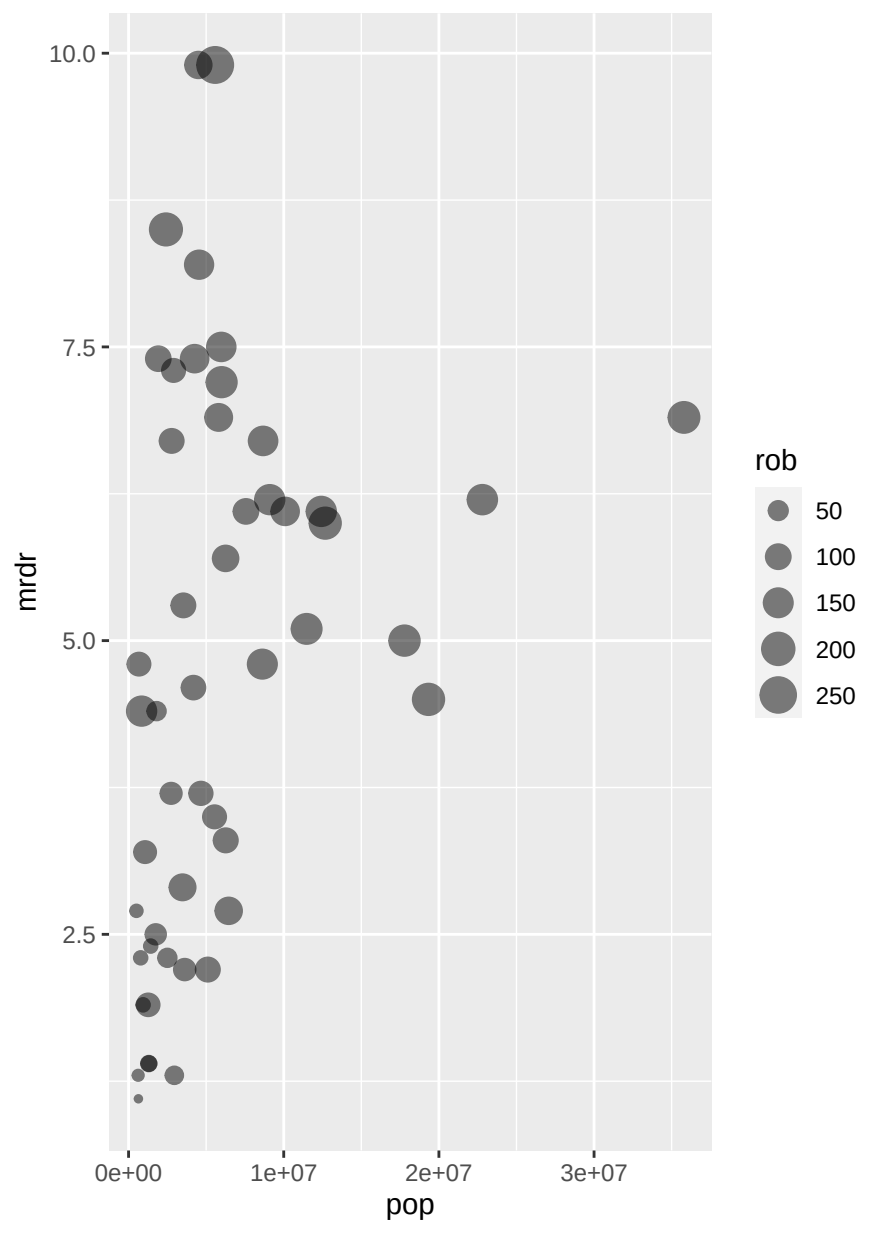
```
<Figure size 864x432 with 0 Axes>
```

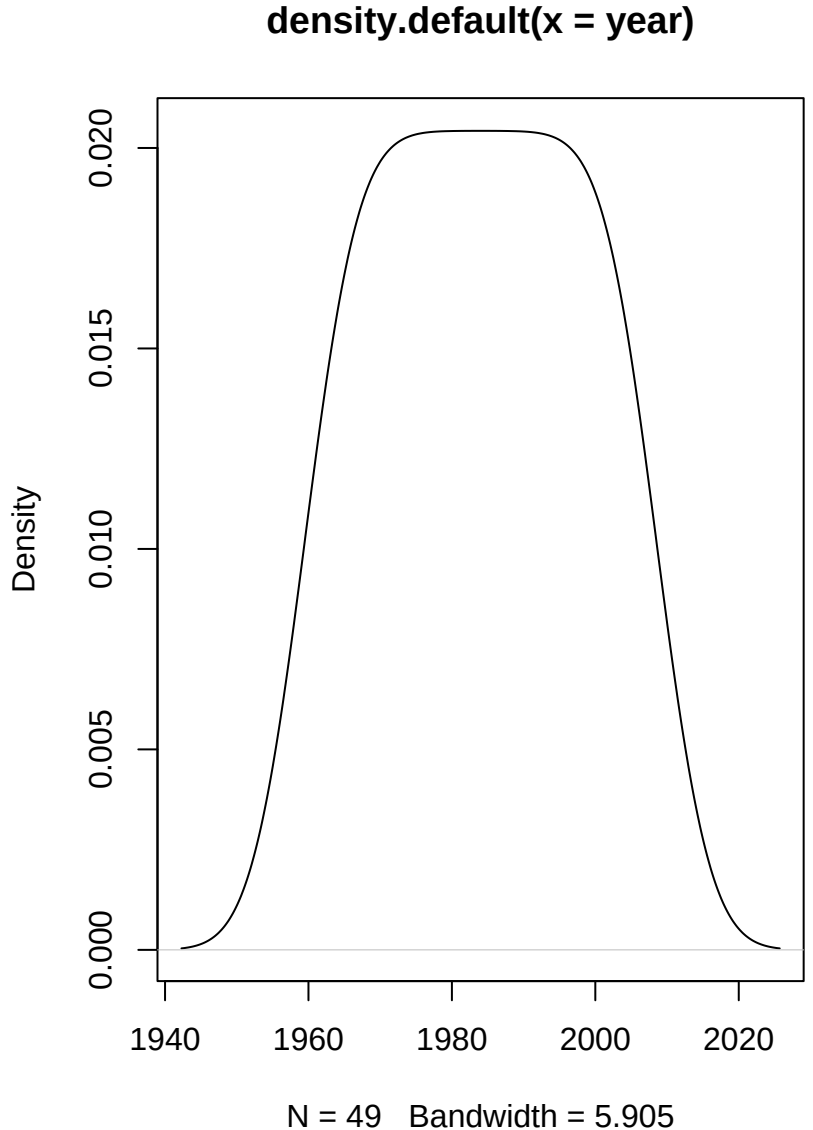


In [ ]:

```
library(ggplot2)
getwd()
setwd("~/Documents/DSC 640")
library(readr)
df <- read_csv("birth-rates-yearly.csv")
df1 <-read_csv("crimerates-by-state-2005.csv")
View(df)
View(df1)
library("dplyr")
avg <- df %>% group_by(year)%>% summarise(rate = sum(rate),.groups = 'drop')
View(avg)
year<-avg$year
rate<-avg$rate
ggplot(avg, aes(x=year, y=rate)) + geom_point()
df2<-df1[-c(1),]
View(df2)
df3<-df2[-c(9),]
View(df3)
pop<-df3$population
mrdr<-df3$murder
rob<- df3$robbery
ggplot(df3, aes(x=pop, y=mrdr, size = rob)) + geom_point(alpha=0.5)
d <- density(year)
plot(d)
```


8000
7000
6000
5000
rate
1960
1970
1980
1990
2000
2010
year


10.0
7.5
5.0
2.5
mrdr
0e+00
1e+07
2e+07
3e+07
pop
rob
50
100
150
200
250


density.default(x = year)
Density
0.000
0.005
0.010
0.015
0.020
1940
1960
1980
2000
2020
N = 49 Bandwidth = 5.905