

```
In [1]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
In [7]: df=pd.read_csv("Capstone project phase 2.csv")
df
```

Out[7]:

	State	Year	Population of Each state	Litracy rate	Area in Sq Km	Type of Crime	Total Crimes	MURDER	ATTE MUR
0	Andhra Pradesh	2001	75728400	66.40	1,62,975	MURDER	130089	1555	
1	Arunachal Pradesh	2001	1098328	66.95	83,743	ATTEMPT TO MURDER	2342	53	
2	Assam	2001	26638600	73.18	78,438	CULPABLE HOMICIDE NOT AMOUNTING TO MURDER	36877	481	
3	Bihar	2001	82879910	69.82	94,163	RAPE	88432	3419	
4	Chhattisgarh	2001	20834530	71.04	1,35,192	CUSTODIAL RAPE	38460	529	
...	
319	Tamil Nadu	2012	635963102	81.33	1,30,058	DOWRY DEATHS	200474	2954	
320	Tripura	2012	32659810	88.75	1,12,077	ASSAULT ON WOMEN WITH INTENT TO OUTRAGE HER MO...	6264	79	
321	Uttar Pradesh	2012	179673604	69.78	2,40,928	INSULT TO MODESTY OF WOMEN	198093	4811	
322	Uttarakhand	2012	89449107	79.64	53,483	CRUELTY BY HUSBAND OR HIS RELATIVES	8882	161	
323	West Bengal	2012	86571309	78.08	88,752	IMPORTATION OF GIRLS FROM FOREIGN COUNTRIES	161427	2854	
324 rows × 37 columns									

```
In [8]: print(df.head())
```

	State	Year	Population of Each state	Litracy rate \
0	Andhra Pradesh	2001	75728400	66.40
1	Arunachal Pradesh	2001	1098328	66.95
2	Assam	2001	26638600	73.18
3	Bihar	2001	82879910	69.82
4	Chhattisgarh	2001	20834530	71.04

	Area in Sq Km	Type of Crime	Total Crimes
0	1,62,975	MURDER	130089
1	83,743	ATTEMPT TO MURDER	2342
2	78,438	CULPABLE HOMICIDE NOT AMOUNTING TO MURDER	36877
3	94,163	RAPE	88432
4	1,35,192	CUSTODIAL RAPE	38460

	MURDER	ATTEMPT TO MURDER	CULPABLE HOMICIDE NOT AMOUNTING TO MURDER
...	\		
0	1555	1555	136
...			
1	53	53	3
...			
2	481	481	40
...			
3	3419	3419	250
...			
4	529	529	45
...			

	ARSON	HURT/GREVIOUS HURT	DOWRY DEATHS \
0	872	34947	420
1	13	466	0
2	441	5805	59
3	502	7544	859
4	215	5477	70

	ASSAULT ON WOMEN WITH INTENT TO OUTRAGE HER MODESTY \
0	3544
1	78
2	850
3	562
4	1763

	INSULT TO MODESTY OF WOMEN	CRUELTY BY HUSBAND OR HIS RELATIVES \
0	2271	5791
1	3	11
2	4	1248
3	21	1558
4	161	840

	IMPORTATION OF GIRLS FROM FOREIGN COUNTRIES	CAUSING DEATH BY NEGLIGENCE \
0	7	740
0		
1	0	
0		
2	0	201
0		
3	83	240
6		
4	0	68
9		

	OTHER IPC CRIMES	TOTAL IPC CRIMES
0	34344	130089
1	618	2342
2	9315	36877
3	36667	88432
4	15790	38460

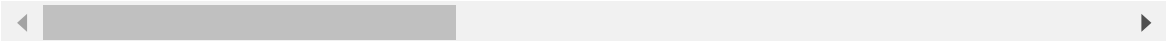
[5 rows x 37 columns]

```
In [19]: df["Capstone project phase 2.csv"] = (df['Total Crimes'] / df['Population o
df
```

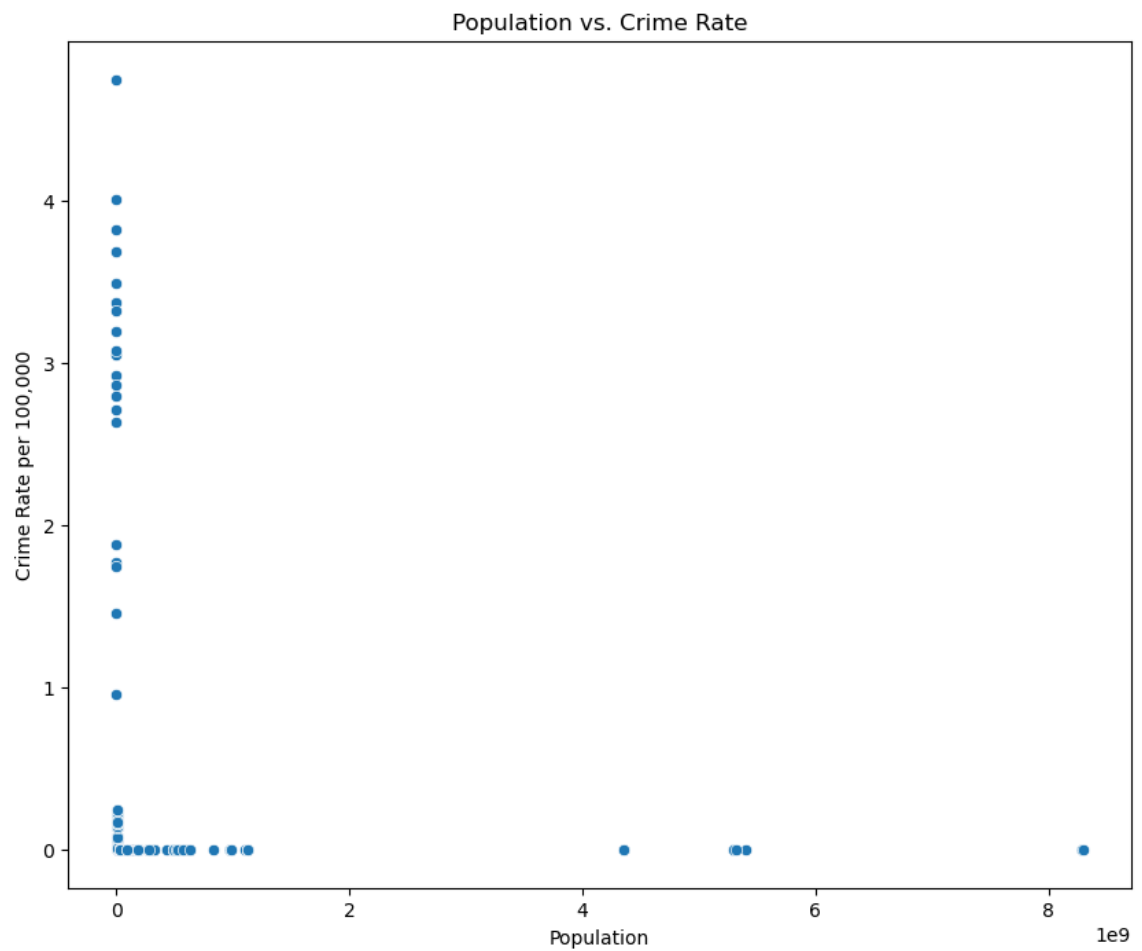
Out[19]:

	State	Year	Population of Each state	Litracy rate	Area in Sq Km	Type of Crime	Total Crimes	MURDER
0	Andhra Pradesh	2001	75728400	66.40	1,62,975	MURDER	2.995465e-04	1555
1	Arunachal Pradesh	2001	1098328	66.95	83,743	ATTEMPT TO MURDER	1.767627e+00	53
2	Assam	2001	26638600	73.18	78,438	CULPABLE HOMICIDE NOT AMOUNTING TO MURDER	1.950839e-03	481
3	Bihar	2001	82879910	69.82	94,163	RAPE	1.553323e-04	3419
4	Chhattisgarh	2001	20834530	71.04	1,35,192	CUSTODIAL RAPE	4.252635e-03	529
...
319	Tamil Nadu	2012	635963102	81.33	1,30,058	DOWRY DEATHS	7.794034e-07	2954
320	Tripura	2012	32659810	88.75	1,12,077	ASSAULT ON WOMEN WITH INTENT TO OUTRAGE HER MO...	1.798087e-04	79
321	Uttar Pradesh	2012	179673604	69.78	2,40,928	INSULT TO MODESTY OF WOMEN	3.415201e-05	4811
322	Uttarakhand	2012	89449107	79.64	53,483	CRUELTY BY HUSBAND OR HIS RELATIVES	1.241031e-05	161
323	West Bengal	2012	86571309	78.08	88,752	IMPORTATION OF GIRLS FROM FOREIGN COUNTRIES	2.488022e-04	2854

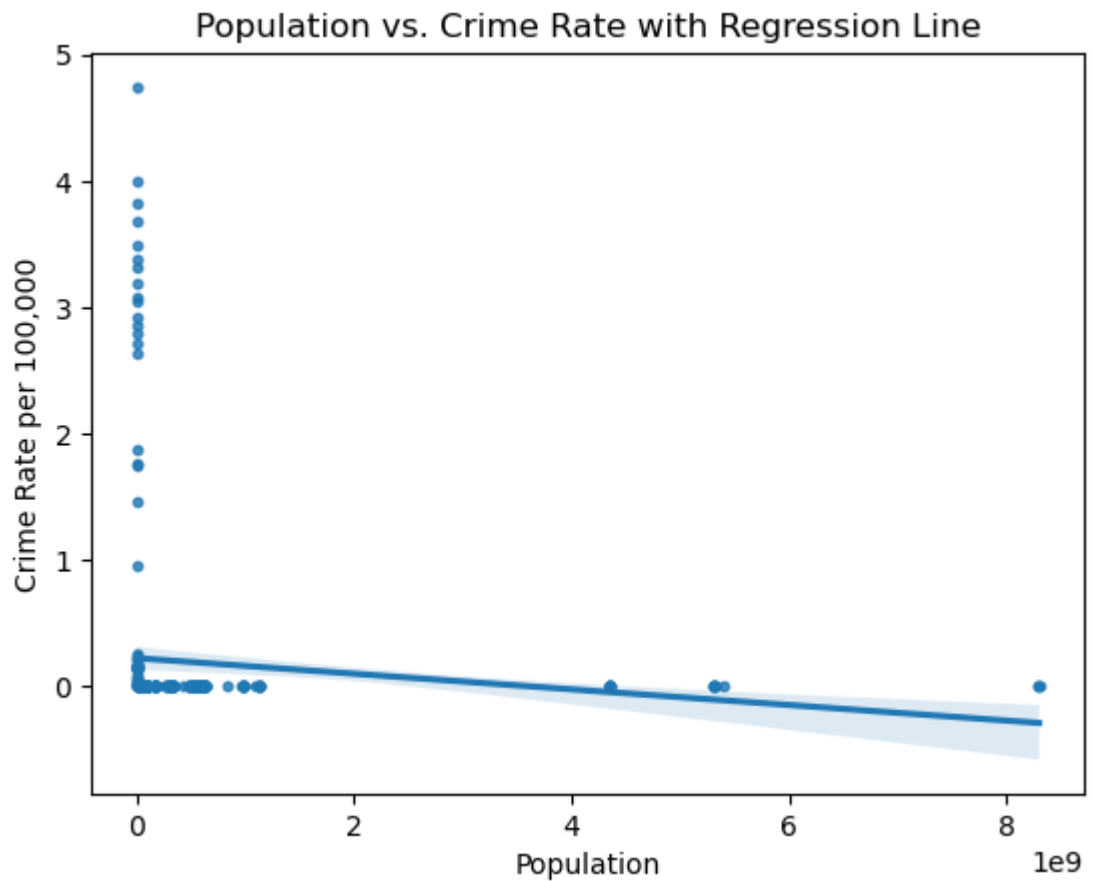
324 rows x 38 columns



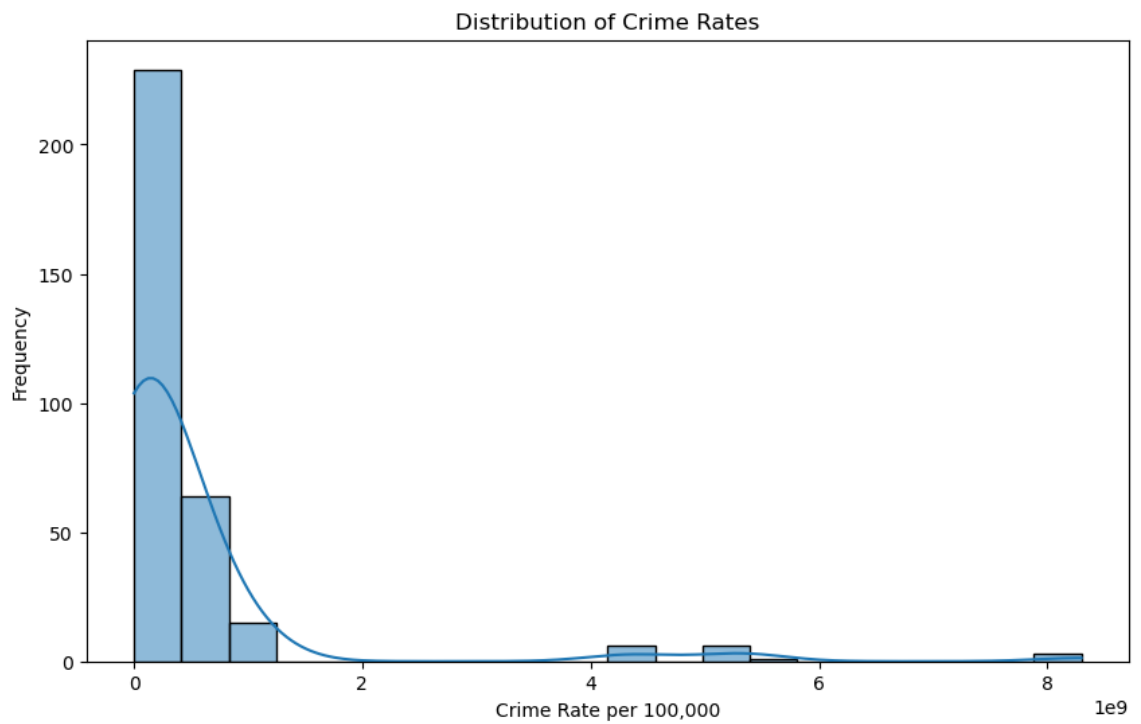
```
In [20]: plt.figure(figsize=(10, 8))
sns.scatterplot(x='Population of Each state', y='Total Crimes' 'Total Crime',
plt.title('Population vs. Crime Rate')
plt.xlabel('Population')
plt.ylabel('Crime Rate per 100,000')
plt.show()
```



```
In [21]: # Explore the relationship using a regression plot
sns.regplot(x='Population of Each state', y='Total Crimes', data=df, scatter_kws={'s': 100})
plt.title('Population vs. Crime Rate with Regression Line')
plt.xlabel('Population')
plt.ylabel('Crime Rate per 100,000')
plt.show()
```



```
In [22]: plt.figure(figsize=(10, 6))
sns.histplot(df['Population of Each state'], bins=20, kde=True)
plt.title('Distribution of Crime Rates')
plt.xlabel('Crime Rate per 100,000')
plt.ylabel('Frequency')
plt.show()
```



In []: