

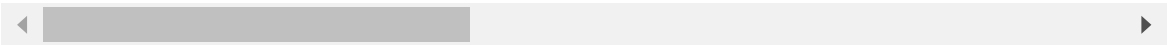
```
In [1]: import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings('ignore')
```

```
In [3]: df = pd.read_csv('Capstone project phase 2.csv')
df
```

Out[3]:

	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 [4]: # Display the first few rows of the dataset  
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

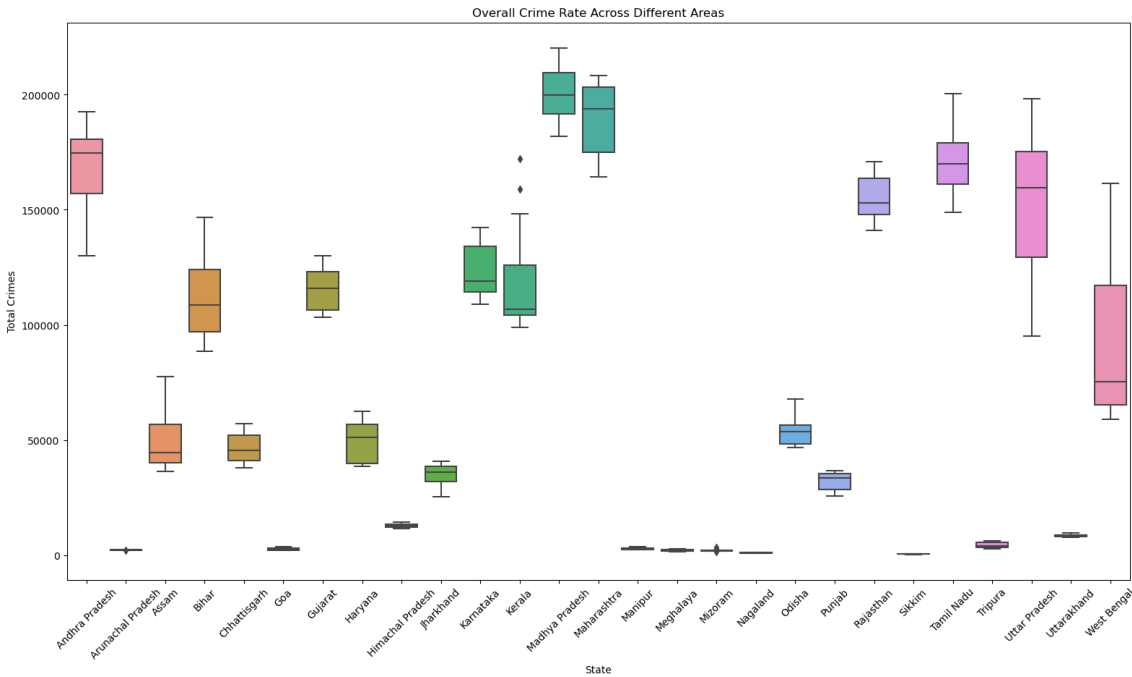
	INSULT TO MODESTY OF WOMEN	CRUELTY BY HUSBAND OR HIS RELATIVES \
0	2271	5791
1	3	11
2	4	1248
3	21	1558
4	161	840

	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 [9]: # Create a boxplot to visualize the distribution of overall crime rates across different areas
plt.figure(figsize=(19, 10))
sns.boxplot(x='State', y='Total Crimes', data=df)
plt.title('Overall Crime Rate Across Different Areas')
plt.xlabel('State')
plt.ylabel('Total Crimes')
plt.xticks(rotation=45)
plt.show()
```



```
In [11]: mean_crime_by_area = df.groupby('State')['Total Crimes'].mean().reset_index  
print(mean_crime_by_area)
```

	State	Total Crimes
0	Andhra Pradesh	168248.416667
1	Arunachal Pradesh	2304.333333
2	Assam	49813.666667
3	Bihar	112191.083333
4	Chhattisgarh	46752.250000
5	Goa	2670.916667
6	Gujarat	115481.250000
7	Haryana	49608.583333
8	Himachal Pradesh	12912.333333
9	Jharkhand	35195.916667
10	Karnataka	123421.916667
11	Kerala	119788.250000
12	Madhya Pradesh	201147.500000
13	Maharashtra	189453.000000
14	Manipur	2922.666667
15	Meghalaya	2104.083333
16	Mizoram	2178.833333
17	Nagaland	1094.416667
18	Odisha	53995.500000
19	Punjab	32010.916667
20	Rajasthan	154659.666667
21	Sikkim	583.333333
22	Tamil Nadu	171681.333333
23	Tripura	4394.500000
24	Uttar Pradesh	154839.500000
25	Uttarakhand	8600.333333
26	West Bengal	93275.333333

```
In [ ]:
```