

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

In [2]: population = pd.read_csv('E:/Machine learning dataset/Population and density.csv')

In [3]: population.head()
```

| Rank | State or union territory | Population(n%) | Decadal growth(2001-2011) | Rural population(n%) | Urban population(n%) | Area[16] | Density[a] | Sex ratio |     |
|------|--------------------------|----------------|---------------------------|----------------------|----------------------|----------|------------|-----------|-----|
| 0    | 1                        | Uttar Pradesh  | 199812341                 | 20.20%               | 155317278            | 44495063 | 240928     | 828       | 912 |
| 1    | 2                        | Maharashtra    | 112374333                 | 20.00%               | 61556074             | 50818259 | 307713     | 365       | 929 |
| 2    | 3                        | Bihar          | 104099452                 | 25.40%               | 92341436             | 11758016 | 94163      | 1102      | 918 |
| 3    | 4                        | West Bengal    | 91276115                  | 23.80%               | 62183113             | 29093002 | 8752       | 1029      | 953 |
| 4    | 5                        | Madhya Pradesh | 72626809                  | 16.30%               | 52557404             | 20069405 | 308245     | 236       | 931 |

```
In [4]: population.shape
Out[4]: (35, 9)
```

```
In [5]: crimes = pd.read_csv('E:/Machine learning dataset/10_Property_stolen_and_recovered.csv')

In [6]: crimes.head()
```

| Area_Name | Year                      | Group_Name | Sub_Group_Name      | Cases_Property_Recovered | Cases_Property_Stolen | Value_of_Property_Recovered | Value_of_Property_Stolen |          |           |
|-----------|---------------------------|------------|---------------------|--------------------------|-----------------------|-----------------------------|--------------------------|----------|-----------|
| 0         | Andaman & Nicobar Islands | 2001       | Burglary - Property | 3                        | Burglary              | 27                          | 64                       | 755858   | 1321961   |
| 1         | Andhra Pradesh            | 2001       | Burglary - Property | 3                        | Burglary              | 3321                        | 7134                     | 51483437 | 147019348 |
| 2         | Arunachal Pradesh         | 2001       | Burglary - Property | 3                        | Burglary              | 66                          | 248                      | 825115   | 4931904   |
| 3         | Assam                     | 2001       | Burglary - Property | 3                        | Burglary              | 539                         | 2423                     | 3722860  | 21466955  |
| 4         | Bihar                     | 2001       | Burglary - Property | 3                        | Burglary              | 367                         | 3231                     | 2327135  | 17023937  |

```
In [7]: crimes.isnull().sum()
Out[7]: Area_Name      0
Year              0
Group_Name       0
Sub_Group_Name   0
Cases_Property_Recovered  0
Cases_Property_Stolen    0
Value_of_Property_Recovered  0
Value_of_Property_Stolen  0
dtype: int64

In [45]: crimes[Year].value_counts()
Out[45]: 2001    245
2002    245
2003    245
2004    245
2005    245
2006    245
2008    245
2009    245
2010    245
2007    244
Name: Year, dtype: int64

In [13]: population.isnull().sum()
Out[13]: Rank              0
State or union territory  0
Population(n%)           0
Decadal growth\n(2001-2011)  0
Rural population\n(n%)      0
Urban population\n(n%)      0
Area[16]                 0
Density[a]                0
Sex ratio                 0
dtype: int64

In [14]: population['State or union territory'].value_counts()
Out[14]: Uttar Pradesh      1
Mizoram                  1
Tripura                  1
Meghalaya                1
Manipur[c]               1
Nagaland                 1
Goa                      1
Arunachal Pradesh        1
Sikkim                   1
Uttarakhand              1
Delhi                    1
Jammu and Kashmir        1
Puducherry               1
Chandigarh               1
Dadra and Nagar Haveli and Daman and Diu  1
Andaman and Nicobar Islands  1
Himachal Pradesh         1
Haryana                  1
Maharashtra              1
Gujarat                  1
Bihar                    1
West Bengal              1
Madhya Pradesh           1
Tamil Nadu               1
Rajasthan                1
Karnataka                1
Andhra Pradesh           1
Chhattisgarh             1
Odisha                   1
Telangana                1
Kerala                   1
Jharkhand                1
Assam                    1
Punjab                   1
Lakshadweep              1
Name: State or union territory, dtype: int64

In [15]: crimes.head()
Out[15]: 0 Andaman & Nicobar Islands 2001 Burglary - Property 3 Burglary 27 64 755858 1321961
1 Andhra Pradesh 2001 Burglary - Property 3 Burglary 3321 7134 51483437 147019348
2 Arunachal Pradesh 2001 Burglary - Property 3 Burglary 66 248 825115 4931904
3 Assam 2001 Burglary - Property 3 Burglary 539 2423 3722860 21466955
4 Bihar 2001 Burglary - Property 3 Burglary 367 3231 2327135 17023937

In [16]: crimes[Area_Name].value_counts()
Out[16]: Andaman & Nicobar Islands 70
Puducherry 70
Maharashtra 70
Manipur 70
Meghalaya 70
Mizoram 70
Nagaland 70
Odisha 70
Punjab 70
Andhra Pradesh 70
Rajasthan 70
Sikkim 70
Tamil Nadu 70
Tripura 70
Uttar Pradesh 70
Uttarakhand 70
Madhya Pradesh 70
Kerala 70
Delhi 70
Daman & Diu 70
Arunachal Pradesh 70
Assam 70
Bihar 70
Chandigarh 70
Chhattisgarh 70
Dadra & Nagar Haveli 70
West Bengal 70
Karnataka 70
Goa 70
Gujarat 70
Haryana 70
Himachal Pradesh 70
Jammu & Kashmir 70
Jharkhand 70
Lakshadweep 69
Name: Area_Name, dtype: int64

In [17]: crimes.shape
Out[17]: (2449, 8)

In [18]: # different kinds of crimes which took place in India
crimes['Group_Name'].value_counts()
Out[18]: Burglary - Property 350
Criminal Breach of Trust - Property 350
Dacoity - Property 350
Other Heads of Property 350
Robbery - Property 350
Theft - Property 350
Total Property 349
Name: Group_Name, dtype: int64

In [19]: crimes['Sub_Group_Name'].value_counts()
Out[19]: 3. Burglary 350
5. Criminal Breach of Trust 350
1. Dacoity 350
6. Other Property 350
2. Robbery 350
4. Theft 350
7. Total Property Stolen & Recovered 349
Name: Sub_Group_Name, dtype: int64

In [20]: #crimes which took place in Tamil Nadu
Tamil_Nadu = crimes.query('Area_Name=="Tamil Nadu"')
Tamil_Nadu
Out[20]: 30 Tamil Nadu 2001 Burglary - Property 3 Burglary 3227 5919 51446376 124908287
65 Tamil Nadu 2002 Burglary - Property 3 Burglary 2953 5492 45654971 115086173
100 Tamil Nadu 2003 Burglary - Property 3 Burglary 2877 4849 55810582 111017636
135 Tamil Nadu 2004 Burglary - Property 3 Burglary 2679 4147 76752113 124839785
170 Tamil Nadu 2005 Burglary - Property 3 Burglary 2568 3738 54017794 94231495
... ..
2305 Tamil Nadu 2006 Total Property 7. Total Property Stolen & Recovered 15620 17787 357486712 488594368
2339 Tamil Nadu 2007 Total Property 7. Total Property Stolen & Recovered 15374 17875 495527011 680515631
2374 Tamil Nadu 2008 Total Property 7. Total Property Stolen & Recovered 16221 19918 495235732 679476182
2409 Tamil Nadu 2009 Total Property 7. Total Property Stolen & Recovered 16645 21489 593386049 894346136
2444 Tamil Nadu 2010 Total Property 7. Total Property Stolen & Recovered 16125 21509 660311804 1317919190
70 rows x 8 columns

In [21]: Tamil_Nadu.shape
Out[21]: (70, 8)

In [22]: # count of crimes which took place in Tamil Nadu
Tamil_Nadu['Group_Name'].value_counts()
Out[22]: Burglary - Property 10
Criminal Breach of Trust - Property 10
Dacoity - Property 10
Other Heads of Property 10
Robbery - Property 10
Theft - Property 10
Total Property 10
Name: Group_Name, dtype: int64

In [23]: # Number of crimes taken place every year
Tamil_Nadu[Year].value_counts()
Out[23]: 2001 7
2002 7
2003 7
2004 7
2005 7
2006 7
2007 7
2008 7
2009 7
2010 7
Name: Year, dtype: int64

In [24]: # Total value of property stolen in Tamil Nadu
Tamil_Nadu_crimes=Tamil_Nadu['Value_of_Property_Stolen'].sum()
Tamil_Nadu_crimes
Out[24]: 14042000468

In [49]: # Total value of property recovered in Tamil Nadu
Tamil_Nadu_Property_recovered = Tamil_Nadu['Value_of_Property_Recovered'].sum()
Tamil_Nadu_Property_recovered
Out[49]: 873172288

In [26]: Bihar = crimes.query('Area_Name=="Bihar"')
Bihar
Out[26]: 4 Bihar 2001 Burglary - Property 3 Burglary 367 3231 2327135 17023937
39 Bihar 2002 Burglary - Property 3 Burglary 371 3182 2490954 20176572
74 Bihar 2003 Burglary - Property 3 Burglary 305 2986 1719760 19821137
109 Bihar 2004 Burglary - Property 3 Burglary 277 3175 2380067 24846468
144 Bihar 2005 Burglary - Property 3 Burglary 311 3117 3584517 48489204
... ..
2279 Bihar 2006 Total Property 7. Total Property Stolen & Recovered 3182 20326 45435277 283145127
2314 Bihar 2007 Total Property 7. Total Property Stolen & Recovered 2778 19494 53230988 327832613
2348 Bihar 2008 Total Property 7. Total Property Stolen & Recovered 3414 21324 553731417 454302646
2383 Bihar 2009 Total Property 7. Total Property Stolen & Recovered 3400 23314 72473571 507723195
2418 Bihar 2010 Total Property 7. Total Property Stolen & Recovered 3180 22727 74962835 547924668
70 rows x 8 columns

In [27]: # Total property stolen in Bihar
Bihar_crimes=Bihar['Value_of_Property_Stolen'].sum()
Bihar_crimes
Out[27]: 7549162134

In [48]: # Total property recovered in Bihar
Bihar_property_recovered = Bihar['Value_of_Property_Recovered'].sum()
Bihar_property_recovered
Out[48]: 1698784766

In [28]: population.head()
```

| Rank | State or union territory | Population(n%) | Decadal growth(2001-2011) | Rural population(n%) | Urban population(n%) | Area[16] | Density[a] | Sex ratio |     |
|------|--------------------------|----------------|---------------------------|----------------------|----------------------|----------|------------|-----------|-----|
| 0    | 1                        | Uttar Pradesh  | 199812341                 | 20.20%               | 155317278            | 44495063 | 240928     | 828       | 912 |
| 1    | 2                        | Maharashtra    | 112374333                 | 20.00%               | 61556074             | 50818259 | 307713     | 365       | 929 |
| 2    | 3                        | Bihar          | 104099452                 | 25.40%               | 92341436             | 11758016 | 94163      | 1102      | 918 |
| 3    | 4                        | West Bengal    | 91276115                  | 23.80%               | 62183113             | 29093002 | 8752       | 1029      | 953 |
| 4    | 5                        | Madhya Pradesh | 72626809                  | 16.30%               | 52557404             | 20069405 | 308245     | 236       | 931 |

```
In [29]: population.tail()
```

| Rank | State or union territory | Population(n%)                           | Decadal growth(2001-2011) | Rural population(n%) | Urban population(n%) | Area[16] | Density[a] | Sex ratio |      |
|------|--------------------------|--|---------------------------|----------------------|----------------------|----------|------------|-----------|------|
| 30   | 31                       | Puducherry                               | 1247953                   | 28.10%               | 395200               | 852753   | 479        | 2598      | 1037 |
| 31   | 32                       | Chandigarh                               | 1055450                   | 17.20%               | 29991                | 1026459  | 114        | 9252      | 818  |
| 32   | 33                       | Dadra and Nagar Haveli and Daman and Diu | 585764                    | 55.10%               | 243510               | 342254   | 603        | 970       | 711  |
| 33   | 34                       | Andaman and Nicobar Islands              | 380581                    | 6.90%                | 237093               | 143488   | 8249       | 46        | 876  |
| 34   | 35                       | Lakshadweep                              | 64473                     | 6.30%                | 14141                | 50332    | 32         | 2013      | 946  |

```
In [68]: population.rename(columns={'State or union territory':'States','Rural population(n%)':'Rural_Population','Urban population(n%)':'Urban_Population','Sex ratio':'Sex_ratio'},inplace=True)

In [69]: population

In [69]: population
Out[69]: 0 Rank 1 States 2 Population(n%) 3 Decadal growth(2001-2011) 4 Rural 5 Urban 6 Area[16] 7 Density[a] 8 Sex_ratio
1 0 1 Uttar Pradesh 199812341 20.20% 155317278 44495063 240928 828 912
2 1 2 Maharashtra 112374333 20.00% 61556074 50818259 307713 365 929
3 2 3 Bihar 104099452 25.40% 92341436 11758016 94163 1102 918
4 3 4 West Bengal 91276115 23.80% 62183113 29093002 8752 1029 953
5 4 5 Madhya Pradesh 72626809 16.30% 52557404 20069405 308245 236 931
6 5 6 Tamil Nadu 72147030 15.60% 37229990 34817440 130098 585 996
7 6 7 Rajasthan 68548437 21.30% 51500352 17048095 342239 201 928
8 7 8 Karnataka 61095297 15.80% 37489335 23825962 191791 319 973
9 8 9 Gujarat 60439692 19.30% 34694609 25745083 196024 308 919
10 9 10 Andhra Pradesh 49577103 11.00% 34965693 14610410 162968 303 993
11 10 11 Odisha 41974218 14.00% 34970562 7003656 155707 269 979
12 11 12 Telangana 35003674 13.58% 21395099 13608605 112077 312 988
13 12 13 Kerala 33400601 4.90% 17471135 15934061 38863 859 1084
14 13 14 Jharkhand 32988134 22.40% 25059073 7933061 79714 414 948
15 14 15 Assam 31205576 17.10% 26807034 438542 78438 397 954
16 15 16 Punjab 27743338 13.90% 17344192 10399146 50362 550 895
17 16 17 Chhattisgarh 25545198 22.60% 19607961 5937237 135191 189 991
18 17 18 Haryana 25351462 19.90% 16509359 8842103 44212 573 879
19 18 19 Uttarakhand 10086292 12.90% 6176050 3049338 53483 189 963
20 19 20 Himachal Pradesh 6864502 18.80% 6176050 888562 55673 123 972
21 20 21 Tripura 3673917 14.80% 2712464 961453 10486 350 960
22 21 22 Meghalaya 2966899 27.90% 2371439 595450 22429 132 989
23 22 23 Manipur[c] 2570390 18.60% 1793875 776515 22327 122 992
24 23 24 Nagaland 1978502 -0.6% 1407536 570966 16579 119 931
25 24 25 Goa 1458545 8.20% 551731 906814 3702 394 973
26 25 26 Arunachal Pradesh 1383727 20.00% 1066358 317369 83743 17 938
27 26 27 Mizoram 1097206 23.50% 525435 571771 21081 82 876
28 27 28 Delhi 16787941 12.90% 459999 135578 7096 66 890
29 28 29 Jammu and Kashmir 12541032 41.40% 9100860 3432972 222236 15 889
30 29 30 Rajasthan 342239 1247953 28.10% 399520 852753 479 2598 1037
31 30 31 Puducherry 1247953 17.20% 29991 1026459 114 9252 818
32 31 32 Dadra and Nagar Haveli and Daman and Diu 585764 55.10% 243510 342254 603 970 711
33 32 33 Andaman and Nicobar Islands 380581 6.90% 237093 143488 8249 46 876
34 33 34 Lakshadweep 64473 6.30% 14141 50332 32 2013 946

In [64]: # Population of Bihar state
Bihar = population.query('States=="Bihar"')
Bihar
Out[64]: 2 3 Bihar 104099452 25.40% 92341436 11758016 94163 1102 918

In [39]: # Top 10 states with the highest population
top_10_populated_states=population.groupby(by=['State or union territory']).max()[['Population(n%)']].sort_values(by=['Population(n%)'],ascending=False).reset_index()
top_10_populated_states.head(10)
Out[39]: 0 State or union territory 1 Population(n%)
1 Maharashtra 112374333
2 Bihar 104099452
3 West Bengal 91276115
4 Madhya Pradesh 72626809
5 Tamil Nadu 72147030
6 Rajasthan 68548437
7 Karnataka 61095297
8 Gujarat 60439692
9 Andhra Pradesh 49577103

In [46]: plt.figure(figsize=(18,10))
plt.title('Top 10 populated states',fontweight='bold')
sns.barplot(x='State or union territory',y='Population(n%)',data=top_10_populated_states,linewidth=2,edgecolor='black')
plt.xlabel('States',fontweight='bold')
plt.ylabel('Population',fontweight='bold')
Text(0, 0.5, 'population')
Out[46]: 
Top 10 populated states

```