

crime-against-women

October 17, 2024

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

```
[7]: data = pd.read_csv("Downloads/crimes_against_women_2001-2014.csv")
```

```
[8]: df = data
```

```
[9]: df.head()
```

```
[9]: Unnamed: 0      STATE/UT      DISTRICT  Year  Rape  \
0          0  ANDHRA PRADESH      ADILABAD  2001    50
1          1  ANDHRA PRADESH    ANANTAPUR  2001    23
2          2  ANDHRA PRADESH    CHITTOOR  2001    27
3          3  ANDHRA PRADESH    CUDDAPAH  2001    20
4          4  ANDHRA PRADESH  EAST GODAVARI  2001    23
```

```
      Kidnapping and Abduction  Dowry Deaths  \
0                          30              16
1                          30               7
2                          34             14
3                          20             17
4                          26             12
```

```
      Assault on women with intent to outrage her modesty  \
0                                          149
1                                          118
2                                          112
3                                          126
4                                          109
```

```
      Insult to modesty of Women  Cruelty by Husband or his Relatives  \
0                          34                                          175
1                          24                                          154
2                          83                                          186
3                          38                                           57
4                          58                                         247
```

	Importation of Girls
0	0
1	0
2	0
3	0
4	0

```
[10]: df.tail()
```

```
[10]: Unnamed: 0    STATE/UT    DISTRICT    Year    Rape  \
10672      832  Lakshadweep    Lakshadweep  2014      1
10673      833  Lakshadweep  Total District(s)  2014      1
10674      834   Puducherry    Karaikal    2014      3
10675      835   Puducherry    Puducherry  2014      7
10676      836   Puducherry  Total District(s)  2014     10
```

	Kidnapping and Abduction	Dowry Deaths	\
10672	0	0	
10673	0	0	
10674	1	0	
10675	6	1	
10676	7	1	

	Assault on women with intent to outrage her modesty	\
10672	1	
10673	1	
10674	12	
10675	20	
10676	32	

	Insult to modesty of Women	Cruelty by Husband or his Relatives	\
10672	2	0	
10673	2	0	
10674	1	1	
10675	7	3	
10676	8	4	

	Importation of Girls
10672	0
10673	0
10674	0
10675	0
10676	0

```
[11]: df.columns
```

```
[11]: Index(['Unnamed: 0', 'STATE/UT', 'DISTRICT', 'Year', 'Rape',
        'Kidnapping and Abduction', 'Dowry Deaths',
        'Assault on women with intent to outrage her modesty',
        'Insult to modesty of Women', 'Cruelty by Husband or his Relatives',
        'Importation of Girls'],
        dtype='object')
```

```
[12]: df= df.drop(columns=['Unnamed: 0'])
```

```
[13]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10677 entries, 0 to 10676
Data columns (total 10 columns):
#   Column                                          Non-Null Count  Dtype
---  -
0   STATE/UT                                      10677 non-null  object
1   DISTRICT                                      10677 non-null  object
2   Year                                           10677 non-null  int64
3   Rape                                           10677 non-null  int64
4   Kidnapping and Abduction                    10677 non-null  int64
5   Dowry Deaths                                10677 non-null  int64
6   Assault on women with intent to outrage her  10677 non-null  int64
   modesty
7   Insult to modesty of Women                 10677 non-null  int64
8   Cruelty by Husband or his Relatives         10677 non-null  int64
9   Importation of Girls                        10677 non-null  int64
dtypes: int64(8), object(2)
memory usage: 834.3+ KB
```

```
[14]: df.describe()
```

```
[14]:
```

	Year	Rape	Kidnapping and Abduction	Dowry Deaths \
count	10677.000000	10677.000000	10677.000000	10677.000000
mean	2007.697949	57.989885	69.888358	20.181699
std	4.046874	214.230398	311.623450	98.276531
min	2001.000000	0.000000	0.000000	0.000000
25%	2004.000000	8.000000	6.000000	1.000000
50%	2008.000000	22.000000	20.000000	5.000000
75%	2011.000000	44.000000	49.000000	16.000000
max	2014.000000	5076.000000	10626.000000	2469.000000

```

Assault on women with intent to outrage her modesty \
count      10677.000000
mean       113.539196
std        458.903951
min         0.000000
25%        10.000000
```

50%	34.000000
75%	85.000000
max	10001.000000

	Insult to modesty of Women	Cruelty by Husband or his Relatives \
count	10677.000000	10677.000000
mean	27.419313	209.224314
std	167.806797	905.664362
min	0.000000	0.000000
25%	0.000000	11.000000
50%	2.000000	50.000000
75%	12.000000	144.000000
max	4970.000000	23278.000000

	Importation of Girls
count	10677.000000
mean	0.175330
std	2.228637
min	0.000000
25%	0.000000
50%	0.000000
75%	0.000000
max	83.000000

[15]: # Q1 Create a Dataset containing the 10 highest reported rape cases in India, in the span of years 2001-2014.

```
top10 = df.nlargest(10, 'Rape')
rape_cases_dataset = top10[['STATE/UT', 'DISTRICT', 'Year', 'Rape']]
rape_cases_dataset
```

	STATE/UT	DISTRICT	Year	Rape
10244	Madhya Pradesh	Total District(s)	2014	5076
9426	Madhya Pradesh	ZZ TOTAL	2013	4335
10445	Rajasthan	Total District(s)	2014	3759
10595	Uttar Pradesh	Total District(s)	2014	3467
10291	Maharashtra	Total District(s)	2014	3438
8611	MADHYA PRADESH	TOTAL	2012	3425
7810	MADHYA PRADESH	TOTAL	2011	3406
9628	Rajasthan	ZZ TOTAL	2013	3285
7025	MADHYA PRADESH	TOTAL	2010	3135
9472	Maharashtra	ZZ TOTAL	2013	3063

[16]: # Q2 Create a Dataset containing the 10 highest reported deaths caused by Dowry cases in India, in the span of years 2001-2014.

```
top_10_dowry_deaths = df.nlargest(10, 'Dowry Deaths')
```

```
dowry_deaths_dataset = top_10_dowry_deaths[['STATE/UT', 'DISTRICT', 'Year',
↪ 'Dowry Deaths']]
dowry_deaths_dataset
```

```
[16]:
```

	STATE/UT	DISTRICT	Year	Dowry Deaths
10595	Uttar Pradesh	Total District(s)	2014	2469
9760	Uttar Pradesh	ZZ TOTAL	2013	2335
8132	UTTAR PRADESH	TOTAL	2011	2322
8938	UTTAR PRADESH	TOTAL	2012	2244
5796	UTTAR PRADESH	TOTAL	2008	2237
6563	UTTAR PRADESH	TOTAL	2009	2232
7342	UTTAR PRADESH	TOTAL	2010	2217
650	UTTAR PRADESH	TOTAL	2001	2211
5040	UTTAR PRADESH	TOTAL	2007	2076
1366	UTTAR PRADESH	TOTAL	2002	1893

```
[17]: # Q3 Create a dataset containing the 10 highest reported Domestic violence
↪ cases in India, in the span of years 2001-2014.

top_10_domestic_violence_cases = df.nlargest(10, 'Cruelty by Husband or his
↪ Relatives')
domestic_violence_dataset = top_10_domestic_violence_cases[['STATE/UT',
↪ 'DISTRICT', 'Year', 'Cruelty by Husband or his Relatives']]
domestic_violence_dataset
```

```
[17]:
```

	STATE/UT	DISTRICT	Year	\
10640	West Bengal	Total District(s)	2014	
8982	WEST BENGAL	TOTAL	2012	
8172	WEST BENGAL	TOTAL	2011	
9804	West Bengal	ZZ TOTAL	2013	
7381	WEST BENGAL	TOTAL	2010	
6602	WEST BENGAL	TOTAL	2009	
10445	Rajasthan	Total District(s)	2014	
9628	Rajasthan	ZZ TOTAL	2013	
9050	Andhra Pradesh	ZZ TOTAL	2013	
5835	WEST BENGAL	TOTAL	2008	

	Cruelty by Husband or his Relatives
10640	23278
8982	19865
8172	19772
9804	18116
7381	17796
6602	16112
10445	15905
9628	15094
9050	15084

5835

13663

```
[58]: # Q4 Create a Dataset containing the 10 highest reported Importation cases in
      ↪India, in the span of years 2001-2014.
top_10_importation_cases = df.nlargest(10, 'Importation of Girls')
importation_dataset = top_10_importation_cases[['STATE/UT', 'DISTRICT', 'Year',
      ↪'Importation of Girls']]
importation_dataset
```

```
[58]:
```

	STATE/UT	DISTRICT	Year	Importation of Girls
115	BIHAR	TOTAL	2001	83
3013	BIHAR	TOTAL	2005	74
3597	WEST BENGAL	TOTAL	2005	61
3590	WEST BENGAL	NADIA	2005	60
4486	BIHAR	TOTAL	2007	56
3005	BIHAR	SAMASTIPUR	2005	48
7810	MADHYA PRADESH	TOTAL	2011	45
3746	BIHAR	TOTAL	2006	42
102	BIHAR	PATNA	2001	39
5378	JHARKHAND	TOTAL	2008	39

```
[69]: # Q5 Find out the total number of cases, in the span of 2001-2014 under each
      ↪category, state-wise.
```

```
total_cases_statewise = df.groupby('STATE/UT').agg({
    'Rape': 'sum',
    'Kidnapping and Abduction': 'sum',
    'Dowry Deaths': 'sum',
    'Assault on women with intent to outrage her modesty': 'sum',
    'Insult to modesty of Women': 'sum',
    'Cruelty by Husband or his Relatives': 'sum',
    'Importation of Girls': 'sum'
}).reset_index()

total_cases_statewise
```

```
[69]:
```

	STATE/UT	Rape	Kidnapping and Abduction	Dowry Deaths	\
0	A & N ISLANDS	218	154	12	
1	A & N Islands	64	34	6	
2	A&N Islands	54	24	2	
3	ANDHRA PRADESH	26958	29872	12430	
4	ARUNACHAL PRADESH	1000	1116	4	
..	
67	UTTARAKHAND	2756	4668	1812	
68	Uttar Pradesh	13034	40726	9608	
69	Uttarakhand	996	1816	162	
70	WEST BENGAL	41574	43546	10344	

71	West Bengal	6302	17612	1964
----	-------------	------	-------	------

	Assault on women with intent to outrage her modesty \	
0		428
1		82
2		90
3		103998
4		1568
..		...
67		3090
68		31816
69		1276
70		45742
71		21166

	Insult to modesty of Women	Cruelty by Husband or his Relatives \
0	86	242
1	10	28
2	16	18
3	72262	238014
4	36	334
..
67	3038	8284
68	108	38504
69	152	1472
70	3108	261336
71	2240	82788

	Importation of Girls
0	0
1	0
2	0
3	34
4	0
..	...
67	2
68	0
69	0
70	228
71	26

[72 rows x 8 columns]

[44]: ''' Q6 Find out the top 5 states, where the maximum number of cases has been reported in TOTAL in the span of 2001-2014, each category-wise.
For the "Rape" case
For Kidnapping and abduction case

For cases of deaths due to dowry
For Domestic Violence case
For the Importation of Girls case '''

```
top_5_rape = df.groupby('STATE/UT')['Rape'].sum().nlargest(5).reset_index()
top_5_kidnapping = df.groupby('STATE/UT')['Kidnapping and Abduction'].sum().
    ↪nlargest(5).reset_index()
top_5_dowry_deaths = df.groupby('STATE/UT')['Dowry Deaths'].sum().nlargest(5).
    ↪reset_index()
top_5_domestic_violence = df.groupby('STATE/UT')['Cruelty by Husband or his_
    ↪Relatives'].sum().nlargest(5).reset_index()
top_5_importation = df.groupby('STATE/UT')['Importation of Girls'].sum().
    ↪nlargest(5).reset_index()

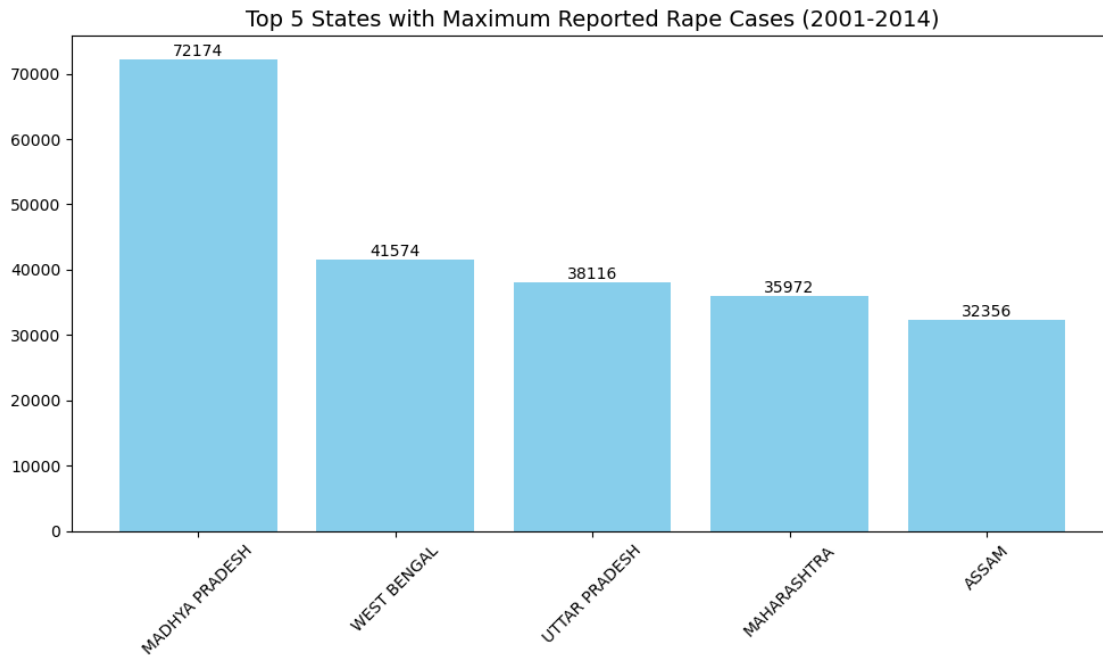
print(top_5_rape)

plt.figure(figsize=(10, 6))
bars = plt.bar(top_5_rape['STATE/UT'], top_5_rape['Rape'], color='skyblue')
plt.title('Top 5 States with Maximum Reported Rape Cases (2001-2014)',
    ↪fontsize=14)
plt.xlabel('')
plt.ylabel('')
plt.xticks(rotation=45)

for bar in bars:
    yval = bar.get_height()
    plt.text(bar.get_x() + bar.get_width() / 2, yval, int(yval), ha='center',
    ↪va='bottom')

plt.tight_layout()
plt.show()
```

	STATE/UT	Rape
0	MADHYA PRADESH	72174
1	WEST BENGAL	41574
2	UTTAR PRADESH	38116
3	MAHARASHTRA	35972
4	ASSAM	32356



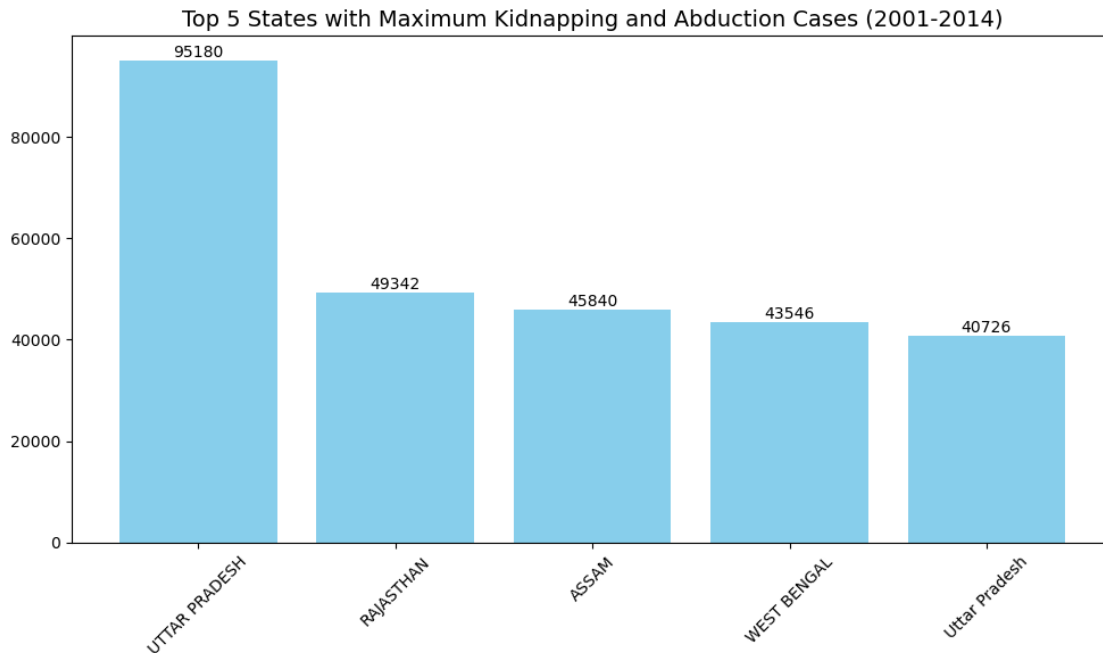
```
[46]: print(top_5_kidnapping)

plt.figure(figsize=(10, 6))
bars = plt.bar(top_5_kidnapping['STATE/UT'], top_5_kidnapping['Kidnapping and
↳Abduction'], color='skyblue')
plt.title('Top 5 States with Maximum Kidnapping and Abduction Cases
↳(2001-2014)', fontsize=14)
plt.xlabel('')
plt.ylabel('')
plt.xticks(rotation=45)

for bar in bars:
    yval = bar.get_height()
    plt.text(bar.get_x() + bar.get_width() / 2, yval, int(yval), ha='center',
↳va='bottom')

plt.tight_layout()
plt.show()
```

	STATE/UT	Kidnapping and Abduction
0	UTTAR PRADESH	95180
1	RAJASTHAN	49342
2	ASSAM	45840
3	WEST BENGAL	43546
4	Uttar Pradesh	40726



```
[48]: print(top_5_dowry_deaths)

plt.figure(figsize=(10, 6))
sns.barplot(x='STATE/UT', y='Dowry Deaths', data=top_5_dowry_deaths,
            palette='Blues_d')
plt.title('Top 5 States with Maximum Dowry Deaths (2001-2014)', fontsize=14)
plt.xlabel('', fontsize=12)
plt.ylabel('', fontsize=12)
plt.xticks(rotation=45)

for index, value in enumerate(top_5_dowry_deaths['Dowry Deaths']):
    plt.text(index, value, int(value), ha='center', va='bottom')

plt.tight_layout()
plt.show()
```

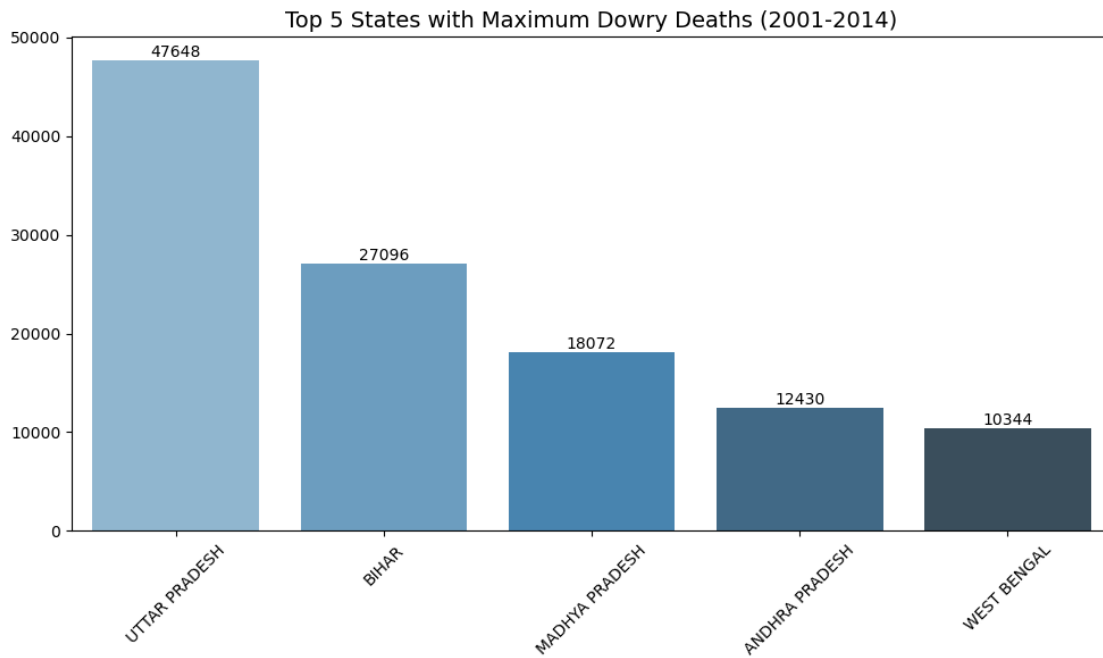
	STATE/UT	Dowry Deaths
0	UTTAR PRADESH	47648
1	BIHAR	27096
2	MADHYA PRADESH	18072
3	ANDHRA PRADESH	12430
4	WEST BENGAL	10344

C:\Users\Abc\AppData\Local\Temp\ipykernel_9840\671217628.py:4: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in

v0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effect.

```
sns.barplot(x='STATE/UT', y='Dowry Deaths', data=top_5_dowry_deaths,
palette='Blues_d')
```



```
[52]: print(top_5_domestic_violence)

plt.figure(figsize=(10, 6))
bars = sns.barplot(x='STATE/UT', y='Cruelty by Husband or his Relatives',
    ↪data=top_5_domestic_violence, palette='Blues_d')
plt.title('Top 5 States with Maximum Domestic Violence Cases (2001-2014)',
    ↪fontsize=14)
plt.xlabel('', fontsize=12)
plt.ylabel('', fontsize=12)
plt.xticks(rotation=45)

for bar in bars.patches:
    yval = bar.get_height()
    plt.text(bar.get_x() + bar.get_width() / 2, yval, int(yval), ha='center',
    ↪va='bottom')

plt.tight_layout()
plt.show()
```

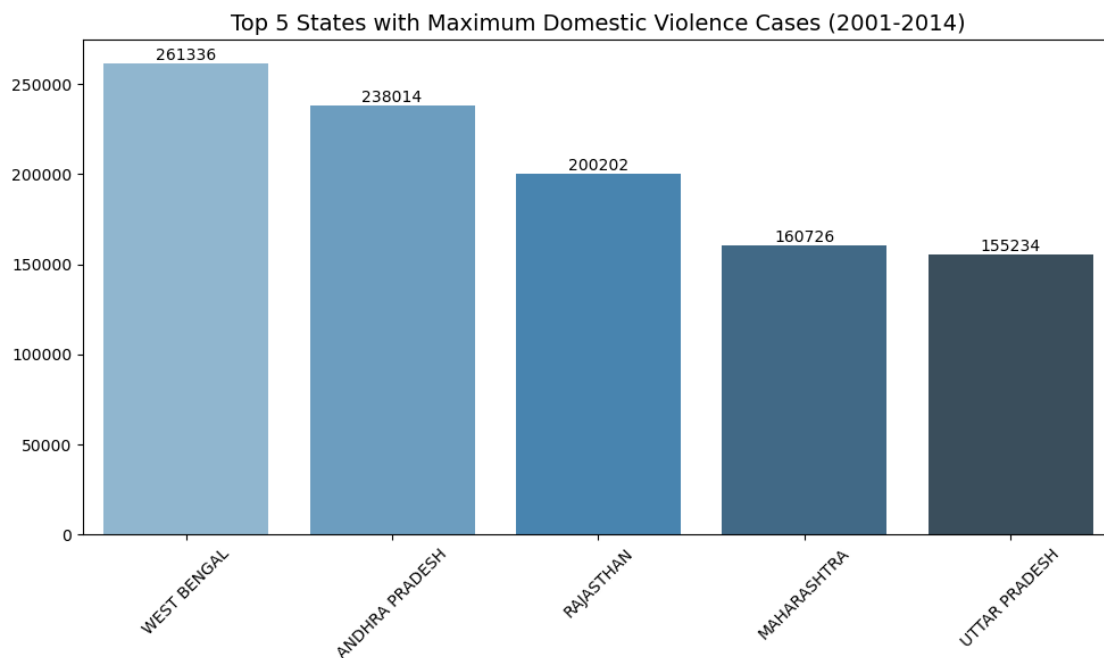
STATE/UT Cruelty by Husband or his Relatives

0	WEST BENGAL	261336
1	ANDHRA PRADESH	238014
2	RAJASTHAN	200202
3	MAHARASHTRA	160726
4	UTTAR PRADESH	155234

C:\Users\Abc\AppData\Local\Temp\ipykernel_9840\1245862534.py:4: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effect.

```
bars = sns.barplot(x='STATE/UT', y='Cruelty by Husband or his Relatives',
data=top_5_domestic_violence, palette='Blues_d')
```



```
[56]: print(top_5_importation)

plt.figure(figsize=(10, 6))
bars = sns.barplot(x='STATE/UT', y='Importation of Girls',
↳data=top_5_importation, palette='Blues_d')
plt.title('Top 5 States with Maximum Importation of Girls Cases (2001-2014)',
↳fontsize=14)

plt.xlabel('')
plt.ylabel('')
plt.xticks(rotation=45)
```

```

for bar in bars.patches:
    yval = bar.get_height()
    plt.text(bar.get_x() + bar.get_width() / 2, yval, int(yval), ha='center',
    ↪va='bottom')

plt.tight_layout()
plt.show()

```

	STATE/UT	Importation of Girls
0	BIHAR	880
1	JHARKHAND	290
2	WEST BENGAL	228
3	MADHYA PRADESH	120
4	KARNATAKA	94

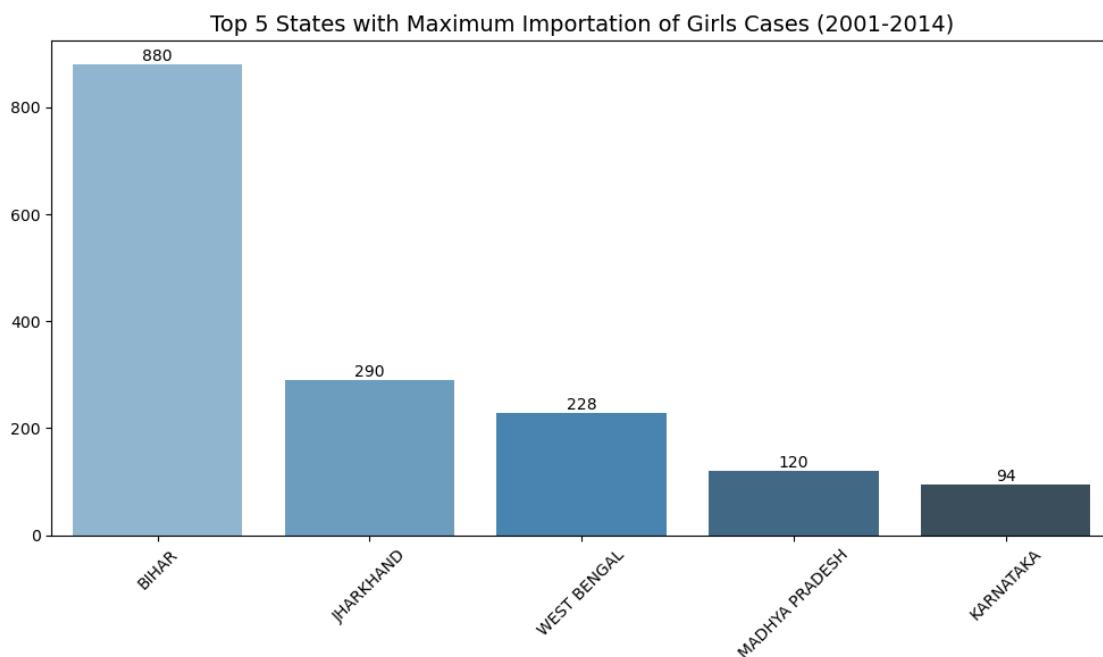
C:\Users\Abc\AppData\Local\Temp\ipykernel_9840\1222582682.py:4: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable to `hue` and set `legend=False` for the same effect.

```

bars = sns.barplot(x='STATE/UT', y='Importation of Girls',
data=top_5_importation, palette='Blues_d')

```



```
[126]: # Q7 Which state has featured in both the lists of "Maximum number of rape
↪cases" and "Maximum number of Importation cases".

top_5_rape = df.groupby('STATE/UT')['Rape'].sum().nlargest(5).reset_index()
top_5_importation = df.groupby('STATE/UT')['Importation of Girls'].sum().
↪nlargest(5).reset_index()
common_state = set(top_5_rape['STATE/UT']).
↪intersection(set(top_5_importation['STATE/UT']))

print("'Maximum number of Rape cases' and 'Maximum number of Importation cases':
↪", common_state)
```

'Maximum number of Rape cases' and 'Maximum number of Importation cases':
{'MADHYA PRADESH', 'WEST BENGAL'}

```
[128]: # Q8 Which state has featured in both the lists of "Maximum number of rape
↪cases" and "Maximum number of Deaths due to Dowry cases".

top_5_rape = df.groupby('STATE/UT')['Rape'].sum().nlargest(5).reset_index()
top_5_dowry_deaths = df.groupby('STATE/UT')['Dowry Deaths'].sum().nlargest(5).
↪reset_index()
common_state = set(top_5_rape['STATE/UT']).
↪intersection(set(top_5_dowry_deaths['STATE/UT']))

print("'Maximum number of Rape cases' and 'Maximum number of Deaths due to
↪Dowry cases':", common_state)
```

'Maximum number of Rape cases' and 'Maximum number of Deaths due to Dowry cases': {'MADHYA PRADESH', 'WEST BENGAL', 'UTTAR PRADESH'}

```
[60]: # Some Extra Analysis
# Trend Analysis Over Years

yearly_trends = df.groupby('Year').sum().reset_index()

# Visualization

plt.figure(figsize=(12, 6))
for crime in ['Rape', 'Kidnapping and Abduction', 'Dowry Deaths',
↪'Assault on women with intent to outrage her modesty',
↪'Insult to modesty of Women', 'Cruelty by Husband or his
↪Relatives',
↪'Importation of Girls']:
    plt.plot(yearly_trends['Year'], yearly_trends[crime], label=crime)

plt.title('Trend of Crimes Against Women Over Years')
```

```
plt.xlabel('Year')
plt.ylabel('Number of Cases')
plt.legend()
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```

