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

```
In [ ]:
```

```
In [ ]:  crimes_df = pd.read_csv(r'C:\Users\Vipul Parashar\Desktop\Data Analyst\Data Sets for projects\crimes_against_wo
         crimes_df
```

Out[ ]:

	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
0	0	ANDHRA PRADESH	ADILABAD	2001	50	30	16	149	34	175	0
1	1	ANDHRA PRADESH	ANANTAPUR	2001	23	30	7	118	24	154	0
2	2	ANDHRA PRADESH	CHITTOOR	2001	27	34	14	112	83	186	0
3	3	ANDHRA PRADESH	CUDDAPAH	2001	20	20	17	126	38	57	0
4	4	ANDHRA PRADESH	EAST GODAVARI	2001	23	26	12	109	58	247	0
...	...	...	...	...	...	...	...	...	...	...	...
10672	832	Lakshadweep	Lakshadweep	2014	1	0	0	1	2	0	0
10673	833	Lakshadweep	Total District(s)	2014	1	0	0	1	2	0	0
10674	834	Puducherry	Karaikal	2014	3	1	0	12	1	1	0
10675	835	Puducherry	Puducherry	2014	7	6	1	20	7	3	0
10676	836	Puducherry	Total District(s)	2014	10	7	1	32	8	4	0

10677 rows × 11 columns

```
In [ ]:  crimes_df[['Rape']].value_counts()
```

```
Out[ ]:  Rape
         0       593
         1       395
         2       296
         3       249
         12      242
                ...
         153       1
         571       1
         572       1
         573       1
         5076      1
         Name: count, Length: 481, dtype: int64
```

Data Preparation and Cleaning

```
In [ ]:  #### Let's find out if any of the columns have any Null or missing values


         total_null = crimes_df.isna().sum()
         total_null
```

```
Out[ ]:  Unnamed: 0                                              0
         STATE/UT                                               0
         DISTRICT                                               0
         Year                                                   0
         Rape                                                   0
         Kidnapping and Abduction                               0
         Dowry Deaths                                           0
         Assault on women with intent to outrage her modesty    0
         Insult to modesty of Women                             0
         Cruelty by Husband or his Relatives                    0
         Importation of Girls                                   0
         dtype: int64
```

```python
In [ ]:  crimes_df[['Rape']].sum()
```

```
Out[ ]:  Rape     619158
         dtype: int64
```

```python
In [ ]:  #crimes_df = pd.read_csv(r'C:\Users\Vipul Parashar\Desktop\Data Analyst\Data Sets for projects\crimes_against_wo
         #a = crimes_df.loc[['MADHYA PRADESH']]
         #a2 = a[['Rape']].sum()
         #a2
```

```python
In [ ]:  crimes_df.drop(['DISTRICT', 'Unnamed: 0'], axis = 1, inplace=True)
```

```python
In [ ]:  crimes_df.rename( columns = {'Kidnapping and Abduction':'Kidnapping_Abduction','Dowry Deaths':'Dowry_Deaths',
                             'Assault on women with intent to outrage her modesty':'Hurting_of_womens_modesty',
                             'Insult to modesty of Women':'Insult_to_womens_modesty',
                             'Cruelty by Husband or his Relatives':'Domestic_Cruelty',
                             'Importation of Girls':'Importation_of_Girls'}, inplace = True)
```

```python
In [ ]:  crimes_df
```

Out[ ]:

	STATE/UT	Year	Rape	Kidnapping_Abduction	Dowry_Deaths	Hurting_of_womens_modesty	Insult_to_womens_modesty
0	ANDHRA PRADESH	2001	50	30	16	149	34
1	ANDHRA PRADESH	2001	23	30	7	118	24
2	ANDHRA PRADESH	2001	27	34	14	112	83
3	ANDHRA PRADESH	2001	20	20	17	126	38
4	ANDHRA PRADESH	2001	23	26	12	109	58
...	...	...	...	...	...	...	...
10672	Lakshadweep	2014	1	0	0	1	2
10673	Lakshadweep	2014	1	0	0	1	2
10674	Puducherry	2014	3	1	0	12	1
10675	Puducherry	2014	7	6	1	20	7
10676	Puducherry	2014	10	7	1	32	8

10677 rows × 9 columns

```python
In [ ]:
```

Find the unique state

```python
In [ ]:  crimes_df['STATE/UT'].unique()
```

```
Out[ ]:  array(['ANDHRA PRADESH', 'ARUNACHAL PRADESH', 'ASSAM', 'BIHAR',
                'CHHATTISGARH', 'GOA', 'GUJARAT', 'HARYANA', 'HIMACHAL PRADESH',
                'JAMMU & KASHMIR', 'JHARKHAND', 'KARNATAKA', 'KERALA',
                'MADHYA PRADESH', 'MAHARASHTRA', 'MANIPUR', 'MEGHALAYA', 'MIZORAM',
                'NAGALAND', 'ODISHA', 'PUNJAB', 'RAJASTHAN', 'SIKKIM',
                'TAMIL NADU', 'TRIPURA', 'UTTAR PRADESH', 'UTTARAKHAND',
                'WEST BENGAL', 'A & N ISLANDS', 'CHANDIGARH', 'D & N HAVELI',
                'DAMAN & DIU', 'DELHI', 'LAKSHADWEEP', 'PUDUCHERRY',
                'Andhra Pradesh', 'Arunachal Pradesh', 'Assam', 'Bihar',
                'Chhattisgarh', 'Goa', 'Gujarat', 'Haryana', 'Himachal Pradesh',
                'Jammu & Kashmir', 'Jharkhand', 'Karnataka', 'Kerala',
                'Madhya Pradesh', 'Maharashtra', 'Manipur', 'Meghalaya', 'Mizoram',
                'Nagaland', 'Odisha', 'Punjab', 'Rajasthan', 'Sikkim',
                'Tamil Nadu', 'Tripura', 'Uttar Pradesh', 'Uttarakhand',
                'West Bengal', 'A&N Islands', 'Chandigarh', 'D&N Haveli',
                'Daman & Diu', 'Delhi UT', 'Lakshadweep', 'Puducherry',
                'Telangana', 'A & N Islands'], dtype=object)
```

There are lots of repeated data . so we replace them with correct one

```python
In [ ]:  # Fist we will remove all the repeated uppercase values
         def remove_uppercase(r):
             r = r['STATE/UT'].strip()
             r = r.upper()
             return r
         crimes_df['STATE/UT'] = crimes_df.apply(remove_uppercase, axis=1)
```

```python
#Now use replace function to replace the other type of repeated datas as dicussed above
crimes_df['STATE/UT'].replace("A&N ISLANDS", "A & N ISLANDS", inplace = True)
crimes_df['STATE/UT'].replace("D&N HAVELI", "D & N HAVELI", inplace = True)
crimes_df['STATE/UT'].replace("DELHI UT", "DELHI", inplace = True)
```

```python
len(crimes_df['STATE/UT'].unique())
```

Out[ ]: 36

Exploratory Analysis and Visualization

```python
#total population of women who has been a victim of the crime based on their gender.
victims_raped = crimes_df.Rape.sum()
victims_kidnapped_abducted = crimes_df.Kidnapping_Abduction.sum()
dowery_death = crimes_df.Dowry_Deaths.sum()
modesty_assault = crimes_df.Hurting_of_womens_modesty.sum()
insult_to_modesty = crimes_df.Insult_to_womens_modesty.sum()
domestic_violence = crimes_df.Domestic_Cruelty.sum()
girls_imported = crimes_df.Importation_of_Girls.sum()
```

```python
total_population_of_victim_overall = victims_raped + victims_raped + dowery_death +modesty_assault+ insult_to_mo
total_population_of_victim_overall
```

Out[ ]: 5194570

```python
fig, axes = plt.subplots(2, 3, figsize=(25, 12))

axes[0,0].set_title("Chart of rape cases in India in 2001-2014")
axes[0,0].bar(crimes_df.Year, crimes_df.Rape, color = 'black');
plt.xlabel('Year') #X-axis
plt.ylabel('Cases of Rape in India') #Y-axis

axes[0,1].set_title("Chart of Kidnapping and Abduction cases in India in 2001-2014")
axes[0,1].bar(crimes_df.Year , crimes_df.Kidnapping_Abduction , color = 'Green')
plt.xlabel('Year')
plt.ylabel('Case of Kidnapping & Abduction')

axes[0,2].set_title("Chart of Dowry death cases in India in 2001-2014")
axes[0,2].bar(crimes_df.Year, crimes_df.Dowry_Deaths, color = 'blue');
plt.xlabel('Year') #X-axis
plt.ylabel('Cases of Dowry deaths in India') #Y-axis

axes[1,0].set_title("Chart of Assault to her modesty in 2001-2014")
axes[1,0].bar(crimes_df.Year, crimes_df.Hurting_of_womens_modesty, color = 'pink');
plt.xlabel('Year') #X-axis
plt.ylabel('Cases of Assaulting a women for her modesty in India') #Y-axis

axes[1,1].set_title("Chart of Domestic Violence cases in India in 2001-2014")
axes[1,1].bar(crimes_df.Year, crimes_df.Domestic_Cruelty, color = 'orange');
plt.xlabel('Year') #X-axis
plt.ylabel('Cases of Domestic Violence in India') #Y-axis

axes[1,2].set_title("Chart of Importation of girls in India in 2001-2014")
axes[1,2].bar(crimes_df.Year, crimes_df.Importation_of_Girls, color = 'red');
plt.xlabel('Year') #X-axis
plt.ylabel('Cases of Importation of girls in India') #Y-axis
```
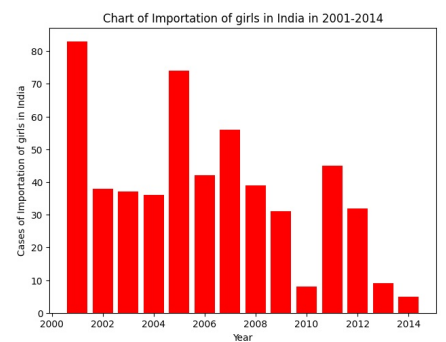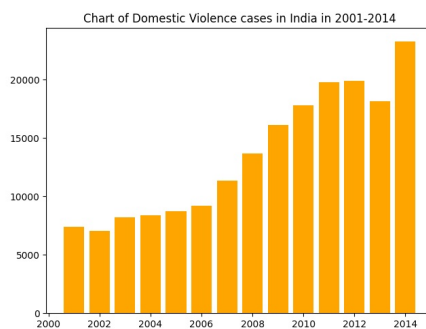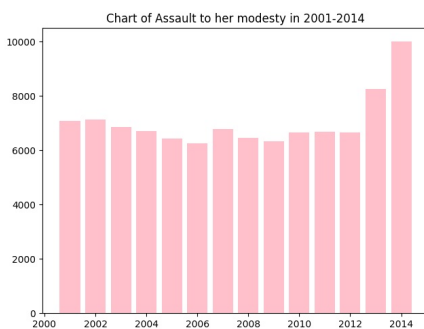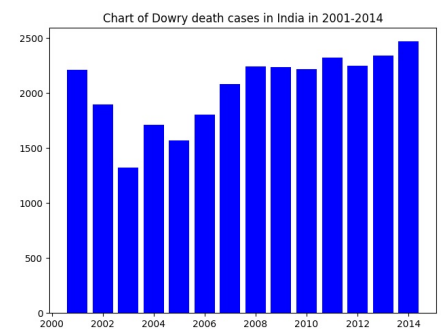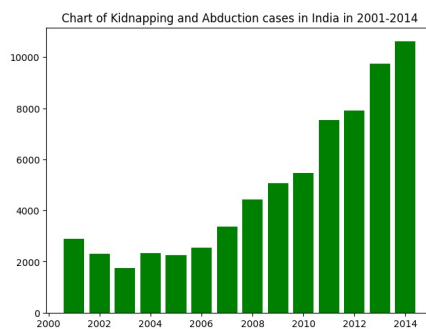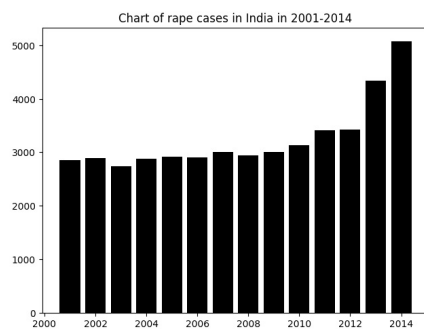
Out[ ]: Text(0, 0.5, 'Cases of Importation of girls in India')

Chart of rape cases in India in 2001-2014

Chart of Kidnapping and Abduction cases in India in 2001-2014

Chart of Dowry death cases in India in 2001-2014

Chart of Assault to her modesty in 2001-2014

Chart of Domestic Violence cases in India in 2001-2014

Chart of Importation of girls in India in 2001-2014

Cases of Importation of girls in India

Year

```
In [ ]:  counts_df = crimes_df.groupby('STATE/UT')[['Rape']].sum()
         counts_df
```

	Rape
STATE/UT	
A & N ISLANDS	336
ANDHRA PRADESH	32150
ARUNACHAL PRADESH	1316
ASSAM	40190
BIHAR	30758
CHANDIGARH	770
CHHATTISGARH	29308
D & N HAVELI	132
DAMAN & DIU	60
DELHI	20312
GOA	1062
GUJARAT	11644
HARYANA	17110
HIMACHAL PRADESH	4674
JAMMU & KASHMIR	7038
JHARKHAND	22826
KARNATAKA	15056
KERALA	20030
LAKSHADWEEP	20
MADHYA PRADESH	90996
MAHARASHTRA	48974
MANIPUR	1068
MEGHALAYA	2642
MIZORAM	2070
NAGALAND	562
ODISHA	30480
PUDUCHERRY	208
PUNJAB	14656
RAJASTHAN	45684
SIKKIM	570
TAMIL NADU	16660
TELANGANA	1958
TRIPURA	5060
UTTAR PRADESH	51150
UTTARAKHAND	3752
WEST BENGAL	47876

There are two things to be concluded from the above bar chart -

1) The cases have incraesed over the years.

2) 2014 has been the year, where violance against women was reported the maximum, under each of the cases.

```
In [ ]: crimes_df1=crimes_df.groupby('Year')[['STATE/UT']].count()
        crimes_df1
```

	STATE/UT
Year	
2001	716
2002	719
2003	728
2004	729
2005	733
2006	740
2007	743
2008	761
2009	767
2010	779
2011	791
2012	811
2013	823
2014	837

In [ ]:
```python
crimes_df.drop(['Hurting_of_womens_modesty', 'Insult_to_womens_modesty'],axis=1 , inplace=True)
```

In [ ]:
```python
crimes_df
```

Out[ ]:

	STATE/UT	Year	Rape	Kidnapping_Abduction	Dowry_Deaths	Domestic_Cruelty	Importation_of_Girls
0	ANDHRA PRADESH	2001	50	30	16	175	0
1	ANDHRA PRADESH	2001	23	30	7	154	0
2	ANDHRA PRADESH	2001	27	34	14	186	0
3	ANDHRA PRADESH	2001	20	20	17	57	0
4	ANDHRA PRADESH	2001	23	26	12	247	0
...	...	...	...	...	...	...	...
10672	LAKSHADWEEP	2014	1	0	0	0	0
10673	LAKSHADWEEP	2014	1	0	0	0	0
10674	PUDUCHERRY	2014	3	1	0	1	0
10675	PUDUCHERRY	2014	7	6	1	3	0
10676	PUDUCHERRY	2014	10	7	1	4	0

10677 rows × 7 columns

It gives us the conclusion that overall(cases under each category), from every state, has increased with time, not decreased.

In [ ]:
```python
# contaning top highest rape case

max_rape_case = crimes_df.sort_values('Rape',ascending = False).head(15)
max_rape_case
```

	STATE/UT	Year	Rape	Kidnapping_Abduction	Dowry_Deaths	Domestic_Cruelty	Importation_of_Girls
10244	MADHYA PRADESH	2014	5076	5688	733	6451	0
9426	MADHYA PRADESH	2013	4335	2873	776	4988	7
10445	RAJASTHAN	2014	3759	4421	408	15905	0
10595	UTTAR PRADESH	2014	3467	10626	2469	10471	0
10291	MAHARASHTRA	2014	3438	2457	279	7696	1
8611	MADHYA PRADESH	2012	3425	1127	743	3988	6
7810	MADHYA PRADESH	2011	3406	1088	811	3732	45
9628	RAJASTHAN	2013	3285	4047	453	15094	1
7025	MADHYA PRADESH	2010	3135	1030	892	3756	5
9472	MAHARASHTRA	2013	3063	1874	320	8542	0
9760	UTTAR PRADESH	2013	3050	9737	2335	8781	0
4735	MADHYA PRADESH	2007	3010	701	742	3294	0
6250	MADHYA PRADESH	2009	2998	841	858	3983	1
5485	MADHYA PRADESH	2008	2937	736	805	3185	0
3259	MADHYA PRADESH	2005	2921	604	739	2989	3

```python
max_importation_case = crimes_df.sort_values('Importation_of_Girls',ascending = False).head(15)
max_importation_case
```

	STATE/UT	Year	Rape	Kidnapping_Abduction	Dowry_Deaths	Domestic_Cruelty	Importation_of_Girls
115	BIHAR	2001	888	518	859	1558	83
3013	BIHAR	2005	1147	929	1014	1574	74
3597	WEST BENGAL	2005	1686	1039	446	6936	61
3590	WEST BENGAL	2005	148	97	48	545	60
4486	BIHAR	2007	1555	1260	1172	1635	56
3005	BIHAR	2005	28	4	40	73	48
7810	MADHYA PRADESH	2011	3406	1088	811	3732	45
3746	BIHAR	2006	1232	1084	1188	1689	42
102	BIHAR	2001	27	11	90	152	39
5378	JHARKHAND	2008	791	499	266	851	39
831	BIHAR	2002	1040	744	927	1577	38
1553	BIHAR	2003	985	674	909	1880	37
2421	JHARKHAND	2004	797	178	275	588	36
970	JHARKHAND	2002	797	178	275	588	36
2281	BIHAR	2004	1390	997	1029	2679	35

```python
max_dowry_death = crimes_df.sort_values('Dowry_Deaths' , ascending= False).head(5)
max_dowry_death
```

	STATE/UT	Year	Rape	Kidnapping_Abduction	Dowry_Deaths	Domestic_Cruelty	Importation_of_Girls
10595	UTTAR PRADESH	2014	3467	10626	2469	10471	0
9760	UTTAR PRADESH	2013	3050	9737	2335	8781	0
8132	UTTAR PRADESH	2011	2042	7525	2322	7121	0
8938	UTTAR PRADESH	2012	1963	7910	2244	7661	0
5796	UTTAR PRADESH	2008	1871	4439	2237	8312	0

We can see that the maximum Dowry death is in 2014 from the Uttar Predesh State with number of reports is 2469.

```python
max_domestic_violance_cases = crimes_df.sort_values('Domestic_Cruelty', ascending = False).head(10)
max_domestic_violance_cases
```

	STATE/UT	Year	Rape	Kidnapping_Abduction	Dowry_Deaths	Domestic_Cruelty	Importation_of_Girls
10640	WEST BENGAL	2014	1466	4976	501	23278	4
8982	WEST BENGAL	2012	2046	4168	593	19865	12
8172	WEST BENGAL	2011	2363	3711	510	19772	0
9804	WEST BENGAL	2013	1685	3830	481	18116	9
7381	WEST BENGAL	2010	2311	2764	507	17796	8
6602	WEST BENGAL	2009	2336	2187	506	16112	5
10445	RAJASTHAN	2014	3759	4421	408	15905	0
9628	RAJASTHAN	2013	3285	4047	453	15094	1
9050	ANDHRA PRADESH	2013	1635	1595	492	15084	0
5835	WEST BENGAL	2008	2263	1907	451	13663	5

With our analysis the maximum domestic violance happened again in 2014 from West Bangal .

total number of cases in 2001-2014 under each category, state wise.

In [ ]:
```
counts_df1 = crimes_df.groupby('STATE/UT')[['Rape', 'Kidnapping_Abduction', 'Dowry_Deaths','Domestic_Cruelty',
counts_df1
```

STATE/UT	Rape	Kidnapping_Abduction	Dowry_Deaths	Domestic_Cruelty	Importation_of_Girls
A & N ISLANDS	336	212	20	288	0
ANDHRA PRADESH	32150	34504	13844	280906	34
ARUNACHAL PRADESH	1316	1470	6	476	0
ASSAM	40190	62074	3268	115300	22
BIHAR	30758	57086	32206	69770	904
CHANDIGARH	770	1682	90	2080	0
CHHATTISGARH	29308	11808	2758	23436	12
D & N HAVELI	132	224	2	90	0
DAMAN & DIU	60	44	6	76	0
DELHI	20312	46586	3758	42834	2
GOA	1062	640	38	532	0
GUJARAT	11644	34670	1108	146468	0
HARYANA	17110	20016	7372	68414	4
HIMACHAL PRADESH	4674	4116	112	7796	0
JAMMU & KASHMIR	7038	21164	294	5390	0
JHARKHAND	22826	14186	7896	23910	298
KARNATAKA	15056	16262	7016	72706	94
KERALA	20030	4452	700	111626	0
LAKSHADWEEP	20	2	0	14	0
MADHYA PRADESH	90996	35608	21090	102816	134
MAHARASHTRA	48974	30368	9696	193202	6
MANIPUR	1068	2606	6	578	0
MEGHALAYA	2642	670	36	460	8
MIZORAM	2070	30	8	134	6
NAGALAND	562	190	2	32	2
ODISHA	30480	25588	10782	49206	36
PUDUCHERRY	208	306	56	234	0
PUNJAB	14656	15096	3524	30840	4
RAJASTHAN	45684	66278	11854	262200	14
SIKKIM	570	180	4	108	0
TAMIL NADU	16660	30908	5060	45524	30
TELANGANA	1958	1422	578	12738	0
TRIPURA	5060	2202	752	16086	0
UTTAR PRADESH	51150	135906	57256	193738	6
UTTARAKHAND	3752	6484	1974	9756	2
WEST BENGAL	47876	61158	12308	344124	254

Q. Top 5 states, where maximum numbers of cases has been reported in TOTAL of 2001-2014, each category wise.

```
counts_df1.sort_values('Rape',ascending=False).head(10)
```

STATE/UT	Rape	Kidnapping_Abduction	Dowry_Deaths	Domestic_Cruelty	Importation_of_Girls
MADHYA PRADESH	90996	35608	21090	102816	134
UTTAR PRADESH	51150	135906	57256	193738	6
MAHARASHTRA	48974	30368	9696	193202	6
WEST BENGAL	47876	61158	12308	344124	254
RAJASTHAN	45684	66278	11854	262200	14
ASSAM	40190	62074	3268	115300	22
ANDHRA PRADESH	32150	34504	13844	280906	34
BIHAR	30758	57086	32206	69770	904
ODISHA	30480	25588	10782	49206	36
CHHATTISGARH	29308	11808	2758	23436	12

Madhya Pradesh has reported the highest number of rape cases in TOTAL in span of 2001-2014.

For Kidnapping case -

```
counts_df1.sort_values('Kidnapping_Abduction',ascending=False).head(5)
```

STATE/UT	Rape	Kidnapping_Abduction	Dowry_Deaths	Domestic_Cruelty	Importation_of_Girls
UTTAR PRADESH	51150	135906	57256	193738	6
RAJASTHAN	45684	66278	11854	262200	14
ASSAM	40190	62074	3268	115300	22
WEST BENGAL	47876	61158	12308	344124	254
BIHAR	30758	57086	32206	69770	904

Uttar Pradesh has reported the highest number of Kidnapping Cases

```
counts_df1.sort_values('Dowry_Deaths',ascending=False).head(5)
```

STATE/UT	Rape	Kidnapping_Abduction	Dowry_Deaths	Domestic_Cruelty	Importation_of_Girls
UTTAR PRADESH	51150	135906	57256	193738	6
BIHAR	30758	57086	32206	69770	904
MADHYA PRADESH	90996	35608	21090	102816	134
ANDHRA PRADESH	32150	34504	13844	280906	34
WEST BENGAL	47876	61158	12308	344124	254

```
counts_df1.sort_values('Importation_of_Girls',ascending=False).head(5)
```

STATE/UT	Rape	Kidnapping_Abduction	Dowry_Deaths	Domestic_Cruelty	Importation_of_Girls
BIHAR	30758	57086	32206	69770	904
JHARKHAND	22826	14186	7896	23910	298
WEST BENGAL	47876	61158	12308	344124	254
MADHYA PRADESH	90996	35608	21090	102816	134
KARNATAKA	15056	16262	7016	72706	94

```
counts_df1.sort_values('Dowry_Deaths',ascending=False).head(5)
```

STATE/UT	Rape	Kidnapping_Abduction	Dowry_Deaths	Domestic_Cruelty	Importation_of_Girls
UTTAR PRADESH	51150	135906	57256	193738	6
BIHAR	30758	57086	32206	69770	904
MADHYA PRADESH	90996	35608	21090	102816	134
ANDHRA PRADESH	32150	34504	13844	280906	34
WEST BENGAL	47876	61158	12308	344124	254

Once again Uttar Pradesh has reported the highest number of Dowry death also

```
max_importation_and_rape_case = max_importation_case.merge(max_rape_case)
max_importation_and_rape_case
```

	STATE/UT	Year	Rape	Kidnapping_Abduction	Dowry_Deaths	Domestic_Cruelty	Importation_of_Girls
0	MADHYA PRADESH	2011	3406	1088	811	3732	45

Madhya Pradesh who has maximun cases reported in both categories i.e Rape and Importation of girls .

```
max_dowery_and_rape_case = max_dowry_death.merge(max_rape_case)
max_dowery_and_rape_case
```

	STATE/UT	Year	Rape	Kidnapping_Abduction	Dowry_Deaths	Domestic_Cruelty	Importation_of_Girls
0	UTTAR PRADESH	2014	3467	10626	2469	10471	0
1	UTTAR PRADESH	2013	3050	9737	2335	8781	0

CONCLUSION

By Analysing these data we concluded that

```
    1) In 2014 ,the highest crime were reported under all the categories

    2) We found the top highest cases reported ever Where Madhya Pradesh having highest number
 of cases of rape in 2014,
        Uttar Pradesh having highest cases in Dowry death in 2014,
        West Bengal having highest cases in Domestic Violance in 2014 and
        Bihar having the highest cases in importaion of girls in 2011.

    3)We also merged the data in two different cases.
```

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js