This data is collated from https://data.gov.in (https://data.gov.in). It has state-wise and district level data on the various crimes committed against women between 2001 to 2014. Some crimes that are included are Rape, Kidnapping and Abduction, Dowry Deaths etc.

```
In [2]: import pandas as pd
        import numpy as np
        import matplotlib.pyplot as plt
        import seaborn as sns
```

readingb the csv file

```
In [3]: crimes_df = pd.read_csv("crimes_against_women_2001-2014.csv")
```

```
In [4]: crimes_df.head()
```

Out[4]:

	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	Imp
0	ANDHRA PRADESH	ADILABAD	2001	50	30	16	149	34	175	
1	ANDHRA PRADESH	ANANTAPUR	2001	23	30	7	118	24	154	
2	ANDHRA PRADESH	CHITTOOR	2001	27	34	14	112	83	186	
3	ANDHRA PRADESH	CUDDAPAH	2001	20	20	17	126	38	57	
4	ANDHRA PRADESH	EAST GODAVARI	2001	23	26	12	109	58	247	

let us find total number of row and column

```
In [8]: crimes_df.shape
```

Out[8]: (10677, 10)

In [9]: `crimes_df.columns`

Out[9]: Index(['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')

In [10]: `crimes_df.dtypes`

Out[10]:
```
STATE/UT                                                 object
DISTRICT                                                 object
Year                                                      int64
Rape                                                      int64
Kidnapping and Abduction                                  int64
Dowry Deaths                                              int64
Assault on women with intent to outrage her modesty       int64
Insult to modesty of Women                                int64
Cruelty by Husband or his Relatives                       int64
Importation of Girls                                      int64
dtype: object
```

In [12]: `crimes_df.describe()`

Out[12]:

	Year	Rape	Kidnapping and Abduction	Dowry Deaths	Assault on women with intent to outrage her modesty	Insult to modesty of Women	H his
count	10677.000000	10677.000000	10677.000000	10677.000000	10677.000000	10677.000000	10(
mean	2007.697949	57.989885	69.888358	20.181699	113.539196	27.419313	;
std	4.046874	214.230398	311.623450	98.276531	458.903951	167.806797	(
min	2001.000000	0.000000	0.000000	0.000000	0.000000	0.000000	
25%	2004.000000	8.000000	6.000000	1.000000	10.000000	0.000000	
50%	2008.000000	22.000000	20.000000	5.000000	34.000000	2.000000	
75%	2011.000000	44.000000	49.000000	16.000000	85.000000	12.000000	
max	2014.000000	5076.000000	10626.000000	2469.000000	10001.000000	4970.000000	23;

In [13]: `crimes_df.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10677 entries, 0 to 10676
Data columns (total 10 columns):
 #   Column                                         Non-Null Count  Dty
pe
---  ------                                         --------------  ---
--
 0   STATE/UT                                       10677 non-null  obj
ect
 1   DISTRICT                                       10677 non-null  obj
ect
 2   Year                                           10677 non-null  int
64
 3   Rape                                           10677 non-null  int
64
 4   Kidnapping and Abduction                       10677 non-null  int
64
 5   Dowry Deaths                                   10677 non-null  int
64
 6   Assault on women with intent to outrage her modesty  10677 non-null  int
64
 7   Insult to modesty of Women                     10677 non-null  int
64
 8   Cruelty by Husband or his Relatives            10677 non-null  int
64
 9   Importation of Girls                           10677 non-null  int
64
dtypes: int64(8), object(2)
memory usage: 834.3+ KB
```

data preparation and cleaning

In [19]:
```python
# is there null values in dataset

overall_crime = crimes_df.isna().sum()
overall_crime.;[]
```

Out[19]:
```
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
```

In [ ]: