Crime Against Women in India (2001-2014)

A COURSE PROJECT REPORT

MASTER OF TECHNOLOGY

INDUSTRIAL ENGINEERING AND MANAGEMENT

SUBMITTED BY:

Ayush Choudhary (M220526ME)

UNDER THE GUIDANCE OF

DR. ANUP APREM (ASSISTANT PROFESSOR)

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY

CALICUT - 673601

Table Of Content

ABSTRACT:

Violence against women has become a prominent topic of discussion in India in recent years. The country where we revere the goddess is also one of the most dangerous places for women. Over the last decade, there has been a report of a crime against a woman every two minutes. Although India has become the world's fastest growing economy, one thing has stayed unchanged: the number of crimes against women. Despite the various modern attempts to promote gender equality, violence against women remains prevalent across culture, race, geography, and religion. One can't even comprehend the severe magnitude of crimes committed against women in our country. For my Data Visualization , I want to raise awareness to a cause in which I truly believe in — Stopping Violence against Women.

Our Government and the Media have placed great focus in this issue due to continuously increasing crime rate. There many crimes committed in different states of India. This project mainly focuses on for Data analysis of the patterns in crime against women through Indian states. Data analytics is the process of analysing data sets in order to make decision about the information we have and using data visualization for the graphical representation of information and data .This helps in the proper analysis of data related to the crime issues against women and the same could provide an aid to the government in effective policy making for preventive measures. This project can be combined with machine learning to recognize crimes of different states and display them in an explicitly understandable format which could help in the generation of doable preventive measures.

This data 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.

1. INTRODUCTION

The main agenda of this article to analyze crime data by following all the steps required for the complete data analysis. The steps include Data Preparation, Data Cleaning, Data Wrangling, Feature Selection, Data Visualization & Comparison.

The Primary data is about the crimes committed against women in India. The data is being recorded from 2001 to 2014. It includes crimes like -

- Rape
- Kidnapping and Abduction
- Dowry Deaths
- Assault on Women with intent to outrage her modesty
- Insult to the modesty of Women
- Cruelty by Husband or his Relatives
- Importation of Girls

The secondary data consists of crimes against women registered between years 2019-2021. This is for a better understanding on this topic and for the comparison last decade and the present scenario.

2. METHODS

In this section we will discuss the methods used to obtain the results.

i. **The Data Set**

→ Crimes Against Women (2001-2014)–

District-wise crimes committed against Women |Open Government Data (OGD) Platform India

→ Crimes Against Women(2019-2021) –

https://ncrb.gov.in/sites/default/files/CII-2021/CII_2021Volume 1.pdf

Throughout this article, we will try to see if at all there is a reduction in crimes as the year count increases. We will visualize each column of data state-wise & year-wise, and thus explore in a much better way.

ii. **Using SQL database system (SQLite)**

The SQL database system (SQLite) is used to store & process the data and converting it to 3rd normal form. These tables are further imported in jupyter notebook to run queries .

The Database in Different normal form is as follows:

a. First Normal Form (or 1NF)

	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	ANDHRA PRADESH	ADILABAD	2001	50	30	16	149	34	175	0
1	ANDHRA PRADESH	ANANTAPUR	2001	23	30	7	118	24	154	0
2	ANDHRA PRADESH	CHITTOOR	2001	27	34	14	112	83	186	0
3	ANDHRA PRADESH	CUDDAPAH	2001	20	20	17	126	38	57	0
4	ANDHRA PRADESH	EAST GODAVARI	2001	23	26	12	109	58	247	0
...	...	...	...	...	...	...	...	...	...	...
10669	DELHI	VIGILANCE	2014	0	0	0	0	0	0	0
10670	DELHI	WEST	2014	251	546	13	484	63	416	0
10672	LAKSHADWEEP	LAKSHADWEEP	2014	1	0	0	1	2	0	0
10674	PUDUCHERRY	KARAIKAL	2014	3	1	0	12	1	1	0
10675	PUDUCHERRY	PUDUCHERRY	2014	7	6	1	20	7	3	0

10186 rows × 10 columns

Columns must have single values and unique names.

b. Second Normal Form (or 2NF)

Since the non-prime attributes are functionally dependent on the entire candidate key and non-prime attributes are directly dependent on candidate keys so 2nf is same as 3nf.

c. Third Normal Form(or 3NF)

Table 1:

	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	ADILABAD	2001	50	30	16	149	34	175	0
1	ANANTAPUR	2001	23	30	7	118	24	154	0
2	CHITTOOR	2001	27	34	14	112	83	186	0
3	CUDDAPAH	2001	20	20	17	126	38	57	0
4	EAST GODAVARI	2001	23	26	12	109	58	247	0
...	...	...	...	...	...	...	...	...	...
10181	VIGILANCE	2014	0	0	0	0	0	0	0
10182	WEST	2014	251	546	13	484	63	416	0
10183	LAKSHADWEEP	2014	1	0	0	1	2	0	0
10184	KARAIKAL	2014	3	1	0	12	1	1	0
10185	PUDUCHERRY	2014	7	6	1	20	7	3	0

10186 rows × 9 columns

Table 2:

	DISTRICT	STATE_UT
0	ADILABAD	ANDHRA PRADESH
1	ANANTAPUR	ANDHRA PRADESH
2	CHITTOOR	ANDHRA PRADESH
3	CUDDAPAH	ANDHRA PRADESH
4	EAST GODAVARI	ANDHRA PRADESH
...	...	...
10181	VIGILANCE	DELHI
10182	WEST	DELHI
10183	LAKSHADWEEP	LAKSHADWEEP
10184	KARAIKAL	PUDUCHERRY
10185	PUDUCHERRY	PUDUCHERRY

10186 rows × 2 columns

For creating tables follow queries are used:

```
create Table Discticts as
    select tabelcrime.DISTRICT ,tabelcrime.Year,tabelcrime.Rape,
    tabelcrime."Kidnapping and Abduction",tabelcrime."Dowry Deaths",
    tabelcrime."Assault on women with intent to outrage her modesty",
    tabelcrime."Insult to modesty of Women",
    tabelcrime."Cruelty by Husband or his Relatives",
    tabelcrime."Importation of Girls"
    from tabelcrime;
```

And is imported in jupyter using the import option from the local drive.

iii. **Cleaning the Dataset**

The Data Set has been cleaned using simple Python Coding. Proper plotting of graph requires extra headers to be removed. Extra headers like 'All India', 'ALL INDIA','TOTAL Crime' section has been cleaned from the dataset. Separate Cleaning codes have been written for individual states that required merging. After cleaning the Data Set it is further used to plot graphs for visualisation.

Some codes used are :-

Python codes	use
Str.upper	Converts text to uppercase
.drop	Deletes a column/row
.replace	Replaces the value in column
.unique	Returns all Unique column values

iv. **Transforming the data**

Data set is imported and as .csv using pandas and after cleaning it is transformed into required data forms to plot the required graphs and understand the scenario properly. All data was sorted by years instead of districts to get a clear picture of rate of change of cases over the years.

Following are the codes that were used in transforming the data.

Python codes	use
.groupby	Groups a column by certain value
.sum	Sums up all rows values
.reset_index	Resets the indexes

v. **Visualizing the data**

For visualising the data libraries like Matplot lib , plotnine and plotly are used . These provides a series of API, and it is suitable for making interactive mapping.

3. <u>**Exploratory Data Analysis:**</u>
<u>**Analysing each crime by Year**</u>

After grouping the rows in dataset by years we get :

	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	2001	16075	14645	6851	34124	9746	49170	114
1	2002	16373	14506	6822	33943	10155	49237	76
2	2003	15847	13296	6208	32939	12325	50703	46
3	2004	18233	15578	7026	34567	10001	58121	89
4	2005	18359	15750	6787	34175	9984	58319	149
5	2006	19348	17414	7618	36617	9966	63128	67
6	2007	20737	20416	8093	38734	10950	75930	61
7	2008	21467	22939	8172	40413	12214	81344	67
8	2009	21397	25741	8383	38711	11009	89546	48
9	2010	22172	29795	8391	40613	9961	94041	36
10	2011	24206	35565	8618	42968	8570	99135	80
11	2012	24923	38262	8233	45351	9173	106527	59
12	2013	33707	51881	8083	70739	12589	118866	31
13	2014	36735	57311	8455	82235	9735	122877	13

Summing all the cases across different years and plotting using plotly. Here plotly seems like a better option for visualisation because it looks minimal and is interactive.
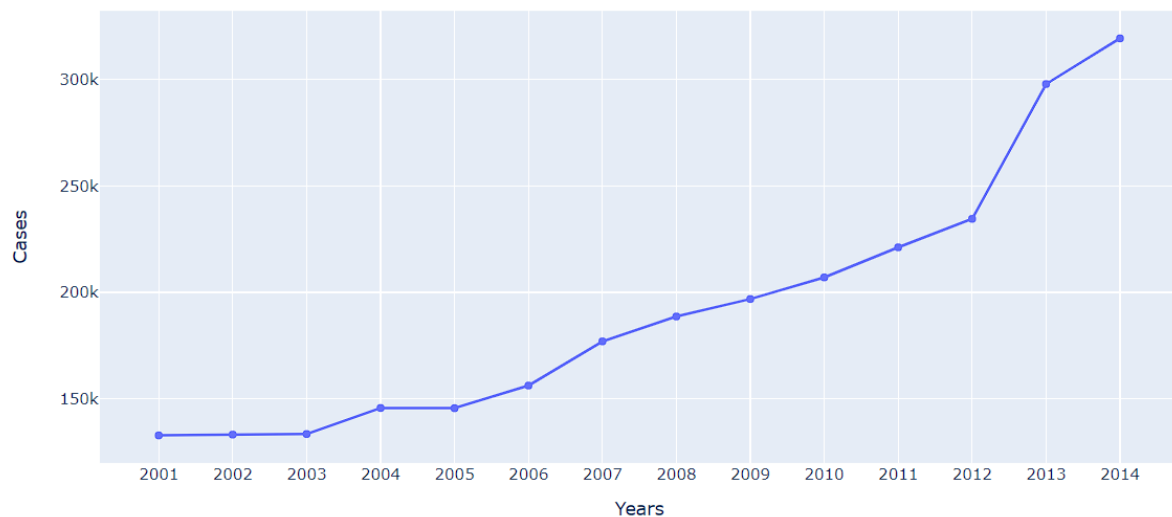

Figure 1: Year wise crimes against women in India

More than 26 lakh cases were registered between 2001 to 2014. Also, the increase in crimes can be seen between years 2012-2013 .In my opinion this is not due to the increased incidences but due to the increase in actual cases registered. This could have happened due to the increased awareness and

strict measures against these crimes after the occurrence of the **Nirbhaya Case** that happened in 2012 and shook the country entirely. This case also forced government to take strict actions against such crimes.

Did Nirbhaya Case really effect the total registered cases?

The direct correlation can be seen if we plot the total cases of rapes in each year.
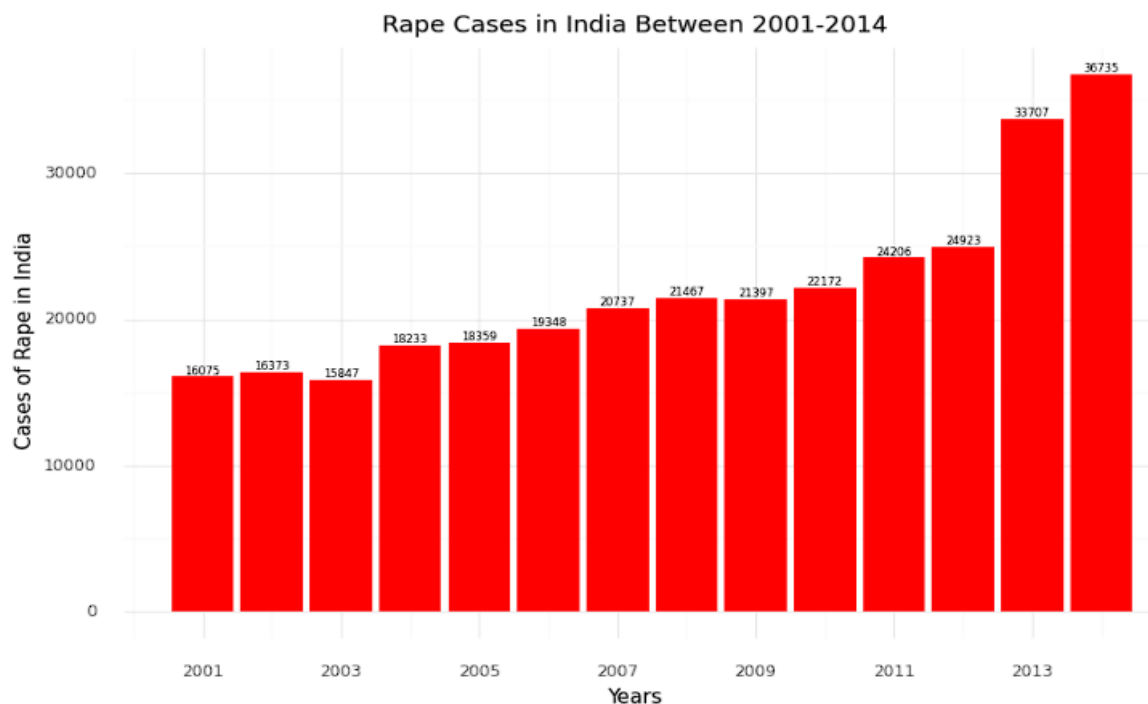

Figure 2: Year wise Rape cases against women in India

A sudden **jump of 35% in rape cases** can be clearly noticed in year 2013. The sluggishness of Government system and lack of awareness can also be noticed in the years preceding it.

Also, in total these had been an increase of more than **56%** in rape cases from 2001 to 2014.

Awareness in women's is directly proportional to the number of rape complaints registered against the wrongdoers . Does this apply other crimes too?

The similar trend can be seen in other two cases as well.

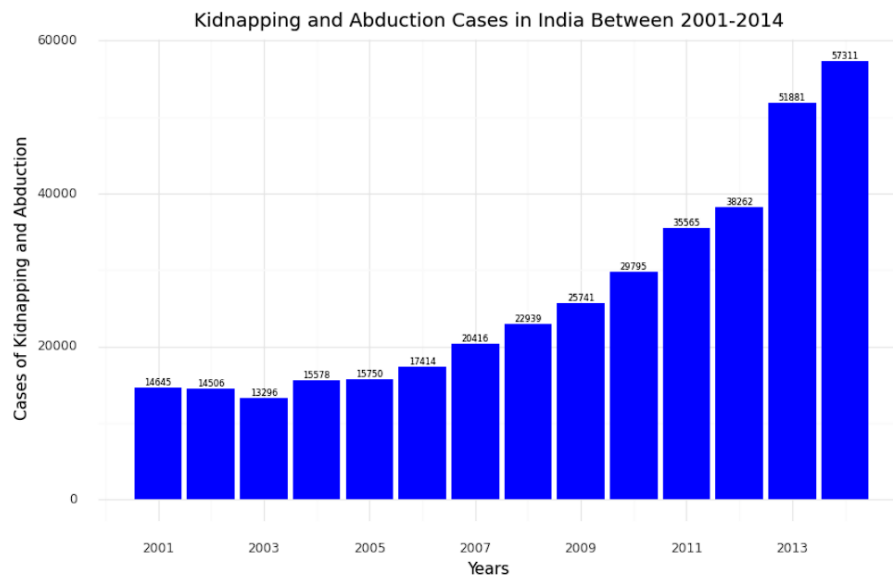They can be plotted using plotnine as follows.

Figure 3: Year wise Kidnapping and Abduction cases against women in India

The abduction cases can be seen to rise by almost 7.5% per year till 2013 but a sudden hike of 35% in year 2014 from 2013.
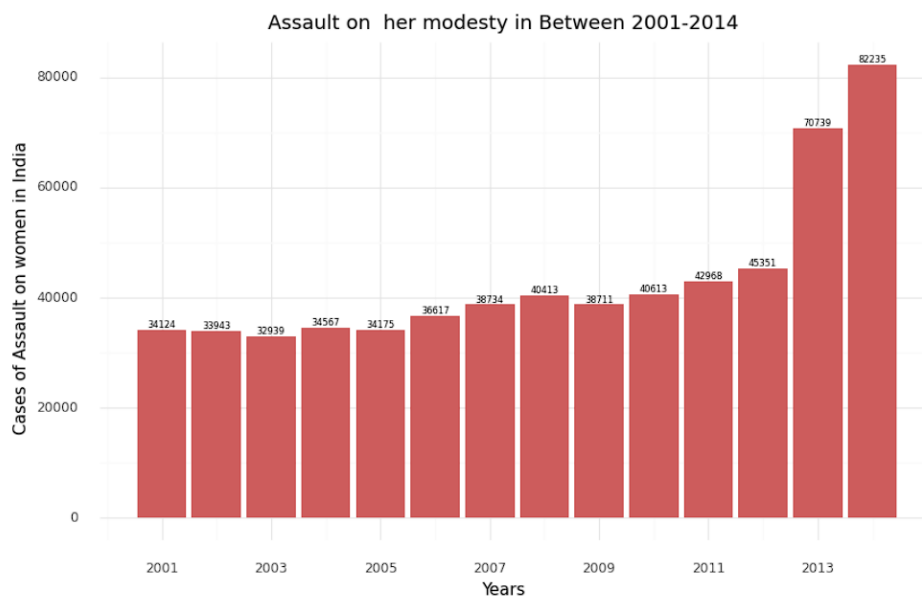

Figure 4: Year wise Assault on her modesty cases between 2001-2014

Similarly, an enormous increase of 56% can be seen in year 2014 for cases of Assault on women with intent to outrage her modest.

So it can be concluded that awareness and strict measures against criminals acts like a motivating factor for women to actually file cases.
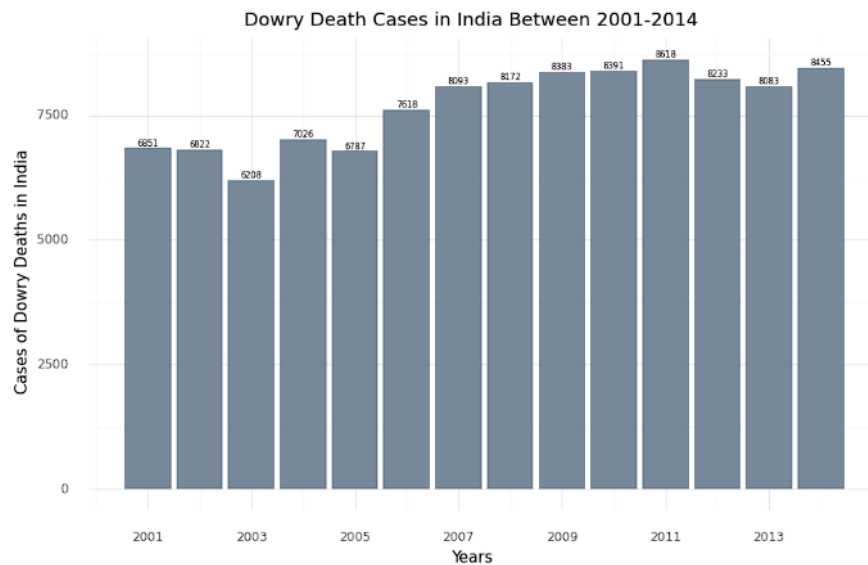
Some other crimes are visualized as follows:

Figure 5: Dowry Death cases between 2001-2014

By looking at this graph one can question have Indians really became more literate ? Because although the literacy rate of India is increased to 75%, there is not much change in dowry cases.
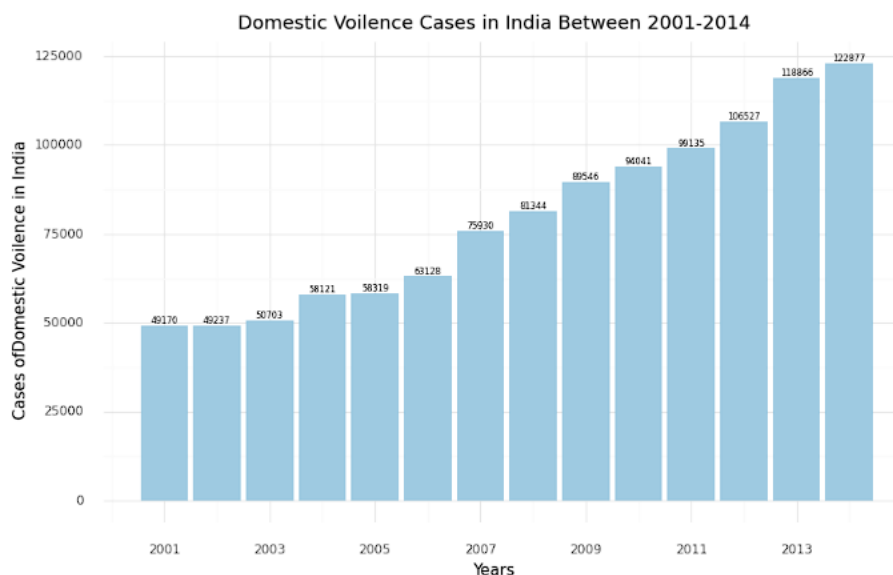

Figure 6: Domestic Violence cases between 2001-2014

Domestic Violence cases have become more than 1.5x in 14 years. These incorporate most crimes against women.

But are there any particular crimes that have reduced over these years?

11

Importation of girls has seen a significant drop in these years. Due to strict laws importation of girl is the only crime which had reduced considerably. There was a reduction of 88% from 2001 to 2014.
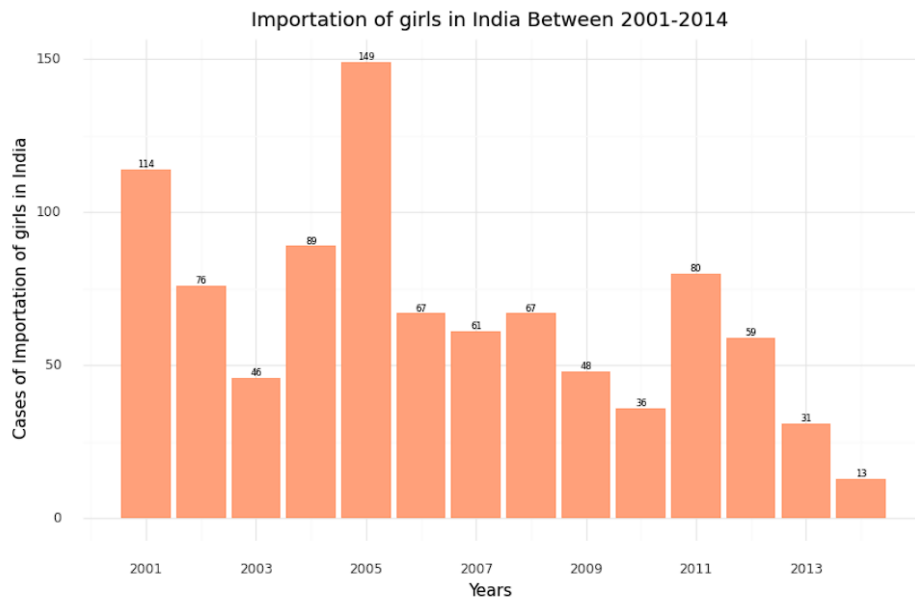

Figure 7: Importation of girls in India between 2001-2014

Let's plot the above graphs in one to see the rate of all the crimes over the years using line plot and pie chart.
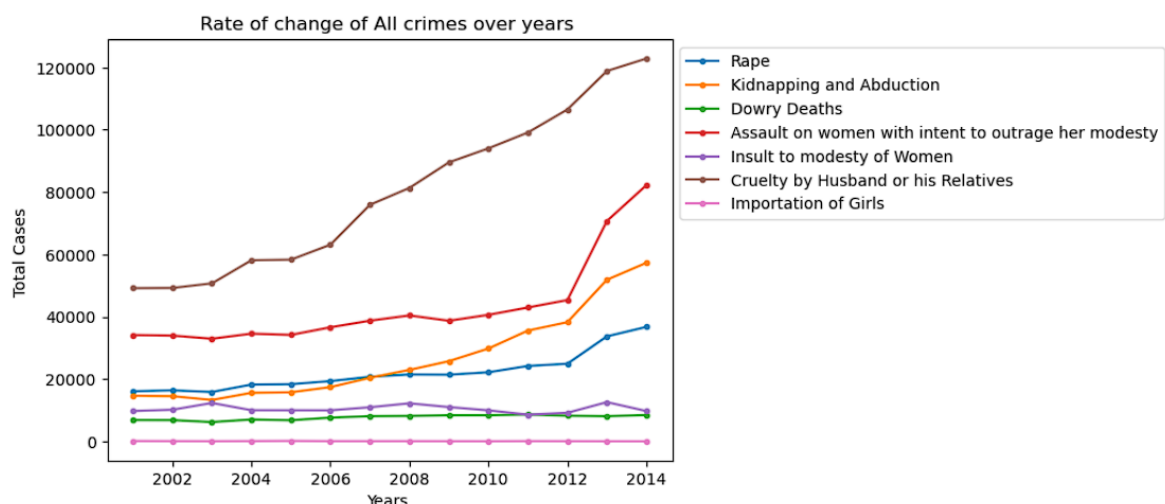

Figure 8: All crimes between 2001-2014

This plot easily helps in visualizing the increase in registered crimes in 2013. We can say that crimes were being committed before the 2012 incident but

women had lack of faith in governance system, and hence these crimes were never registered .

Highest and least reported crimes

Using Count function in pandas and after sorting the values we get.

	Count
Cruelty by Husband or his Relatives	1116944
Assault on women with intent to outrage her modesty	606129
Kidnapping and Abduction	373099
Rape	309579
Insult to modesty of Women	146378
Dowry Deaths	107740
Importation of Girls	936

Plotting A pie chart in plotly:

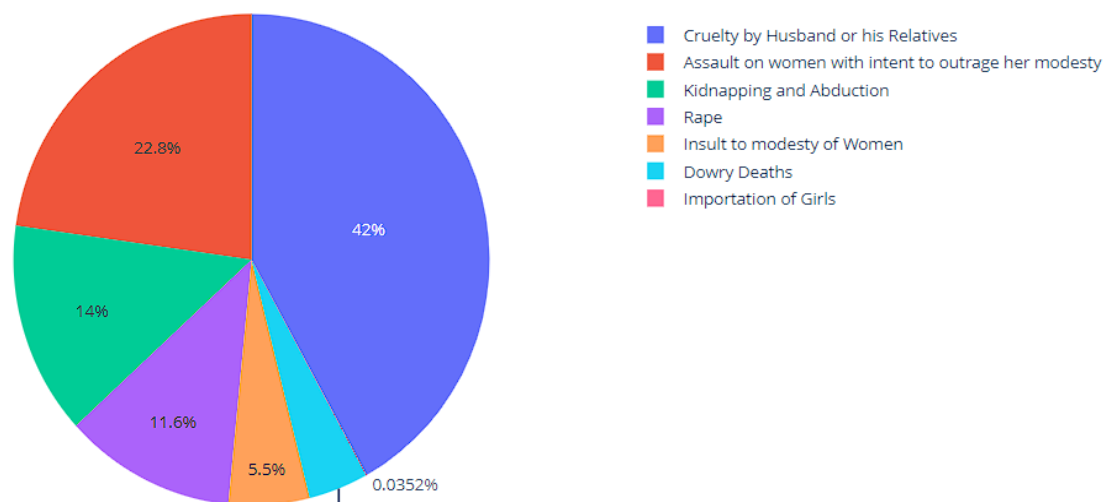Percentage of Each Crime between 2001 - 2014

Figure 9: Proportion of crimes between 2001-2014

Cruelty by Husband or his Relatives i.e., **Domestic Violence was the highest reported** crime between 2001-2014 with more than 11 lakh cases which is **almost half of the crimes** committed against women in these years. It is followed by assaults on women to outrage her modesty. One reason to this trend can be arrange marriages done without the will of women.

Analysing crimes by State/UT

Analysing the data based on State/UT so, as to find out which State/UT is the best and which state is worst to live in for women.

What are the top States/Union Territories with the highest number of crime cases in these years?

On grouping, the data is based on State/UT.

STATE/UT	Total Cases
ANDHRA PRADESH	300355
UTTAR PRADESH	291199
WEST BENGAL	268988
RAJASTHAN	235093
MADHYA PRADESH	233543
MAHARASHTRA	214792
ASSAM	129682
KERALA	112588
GUJARAT	110153
BIHAR	104188

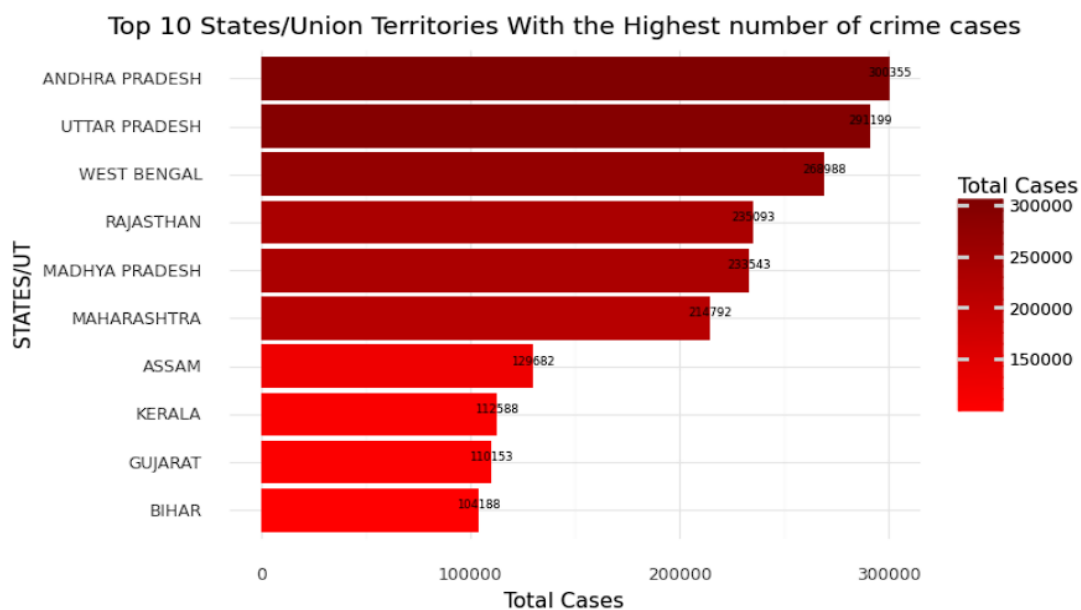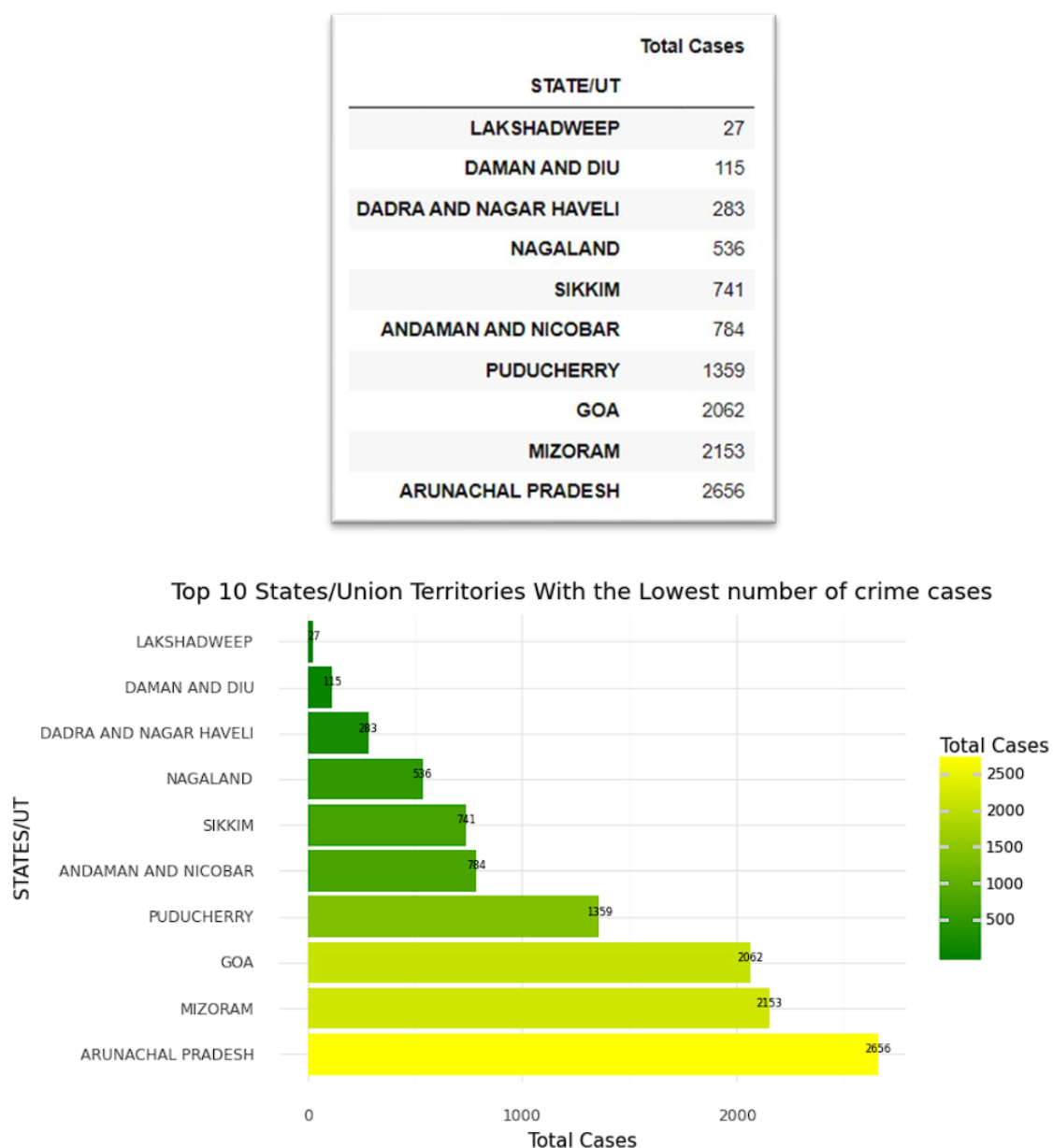On plotting the graph using plotnine we get :

Figure 10: States with highest number of crimes (2001-2014)

Andhra Pradesh has the highest number of crime cases and Uttar Pradesh is the second highest. The results obtained after visualization imply high vulnerability of crimes in states of Andhra Pradesh, Uttar Pradesh, Madhya Pradesh, Rajasthan and Maharashtra

Note: I have already merged Andhra Pradesh and Telangana since they were same before 2014, without the merging of Andhra Pradesh and Telangana, Uttar Pradesh will be the state with the highest number of cases.

What are the top States/Union Territories with the lowest number of crime cases in these years?

STATE/UT	Total Cases
LAKSHADWEEP	27
DAMAN AND DIU	115
DADRA AND NAGAR HAVELI	283
NAGALAND	536
SIKKIM	741
ANDAMAN AND NICOBAR	784
PUDUCHERRY	1359
GOA	2062
MIZORAM	2153
ARUNACHAL PRADESH	2656

Figure 11: States with Lowest number of crimes (2001-2014)

Lakshadweep has the lowest crime rate. It can be noticed that number of crimes registered strongly depends on the size of State/UT.

Out of seven types of crime, which is the States/UT having maximum cases of each type?

	STATE / UT
Rape	MADHYA PRADESH
Kidnapping and Abduction	UTTAR PRADESH
Dowry Deaths	UTTAR PRADESH
Assault on women with intent to outrage her modesty	MADHYA PRADESH
Insult to modesty of Women	ANDHRA PRADESH
Cruelty by Husband or his Relatives	WEST BENGAL
Importation of Girls	BIHAR

Which is the States/UT having minimum cases of each type?

	STATE / UT
Rape	LAKSHADWEEP
Kidnapping and Abduction	LAKSHADWEEP
Dowry Deaths	LAKSHADWEEP
Assault on women with intent to outrage her modesty	LAKSHADWEEP
Insult to modesty of Women	DAMAN AND DIU
Cruelty by Husband or his Relatives	LAKSHADWEEP
Importation of Girls	ANDAMAN AND NICOBAR

Can we say that Madhya Pradesh is the worst state for women while Lakshadweep is the best? This can't be said with surety because it may happen that law enforcement in small UTs is less cases are registered as complaints and vice versa.

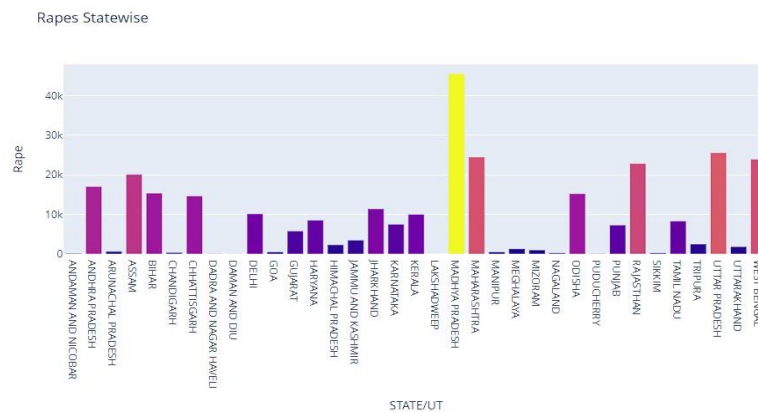**Let's visualise the above scenario by plotting :**

Figure 12: Number of Rapes state wise

In the above graph, the yellow bar shows highest number of Rapes in the state of Madhya Pradesh and Lakshadweep has the lowest rate of Rapes.
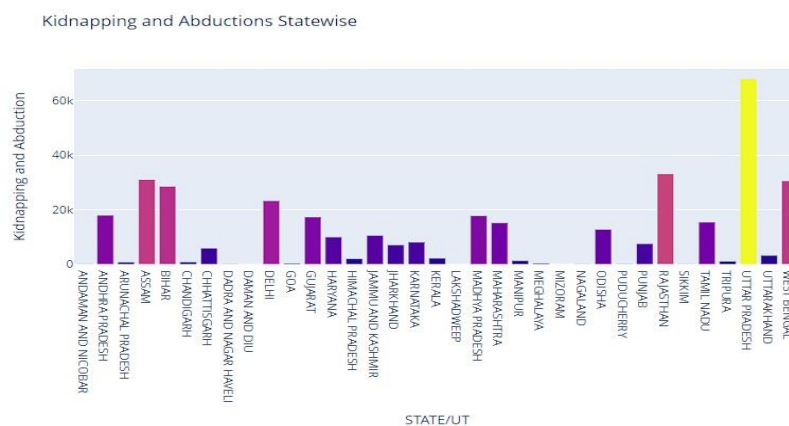

Figure 13: total kidnapping and abduction state wise

In the above graph, the yellow bar shows highest number of kidnapping and abduction in the state of Uttar Pradesh and Lakshadweep has the lowest rate of Kidnapping and abduction.
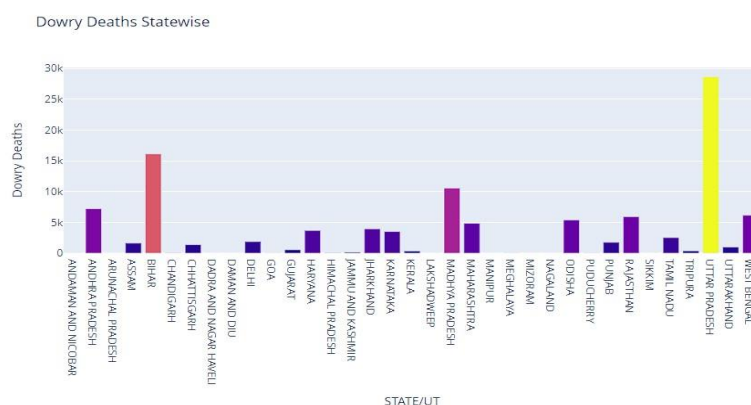

Figure 13: total number of Dowry deaths state wise

In the above graph, the yellow bar shows highest number of deaths due to Dowry issues in the state of Uttar Pradesh and Lakshadweep has the lowest rate of Dowry deaths
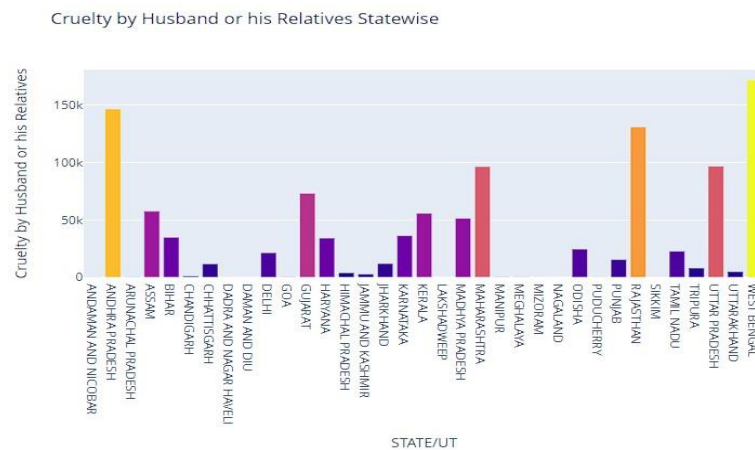

**Figure 14: total number of cruelties by relatives state wise**

In the above graph, the yellow bar shows highest number of Domestic Violence in the state of West Bengal and the blue bar shows Lakshadweep having the lowest rate of deaths due to cruelty by relatives.

Specific state analysis :

Plotting crime distribution of each state/UT would make the code lengthy and data visualisation complicated, so I have defined a function that plots the pie chart for any specific State/UT mentioning all the crimes in that place. To access different plots, enter the state/UT name in the input box.

Enter Name of State/UT : kerala

Total Crime Rate Distribution for KERALA

18
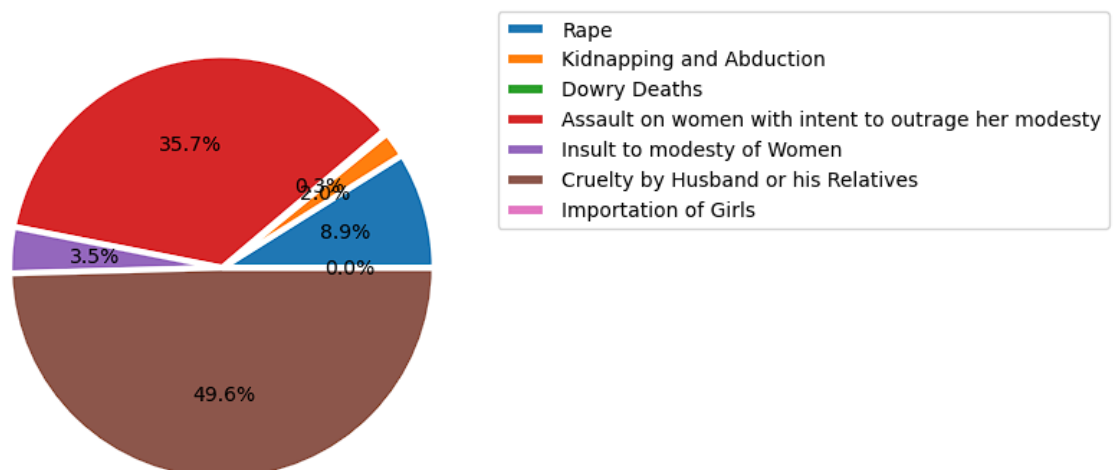
Comparing data from (2001-2014) with data from (2019-2021)

Importing the secondary data as csv and cleaning, we get:

	STATES/UT	2019	2020	2021
0	ANDAMAN AND NICOBAR	135	143	169
1	ANDHRA PRADESH	36140	34860	38617
2	ARUNACHAL PRADESH	317	281	366
3	ASSAM	30025	26352	29046
4	BIHAR	18587	15359	17950
5	CHANDIGARH	515	301	343
6	CHHATTISGARH	7689	7385	7344
7	D&N HAVELI AND nDAMAN & DIU	82	61	99
8	DELHI	13395	10093	14277
9	GOA	329	219	224
10	GUJARAT	8799	8028	7348
11	HARYANA	14683	13000	16658
12	HIMACHAL PRADESH	1636	1614	1599
13	JAMMU & KASHMIR	3069	3405	3937
14	JHARKHAND	8760	7630	8110
15	KARNATAKA	13828	12660	14468
16	KERALA	11462	10139	13539
17	LAKSHADWEEP	38	15	9
18	MADHYA PRADESH	27560	25640	30673
19	MAHARASHTRA	37144	31954	39526
20	MANIPUR	266	247	302
21	MEGHALAYA	558	568	685
22	MIZORAM	170	172	176
23	NAGALAND	43	39	54
24	ODISHA	23183	25489	31352
25	PUDUCHERRY	95	113	153
26	PUNJAB	5886	4838	5662
27	RAJASTHAN	41550	34535	40738
28	SIKKIM	125	140	130
29	TAMIL NADU	5934	6630	8501
30	TRIPURA	1070	874	807
31	UTTAR PRADESH	59853	49365	56083
32	UTTARAKHAND	2541	2846	3431
33	WEST BENGAL	29859	36439	35884

Comparing this data by calling data from year 2014 and merging it.

Using .head feature of pandas to call only 10 worst state we get

	STATE/UT	2014	2019	2020	2021
11	HARYANA	8657	14683	13000	16658
4	BIHAR	12553	18587	15359	17950
3	ASSAM	18790	30025	26352	29046
18	MADHYA PRADESH	27986	27560	25640	30673
24	ODISHA	13891	23183	25489	31352
33	WEST BENGAL	36438	29859	36439	35884
1	ANDHRA PRADESH	28133	36140	34880	38617
19	MAHARASHTRA	25447	37144	31954	39526
27	RAJASTHAN	30510	41550	34535	40738
31	UTTAR PRADESH	35667	59853	49385	56083

using plotly to make the stacked bar graph from the above table:
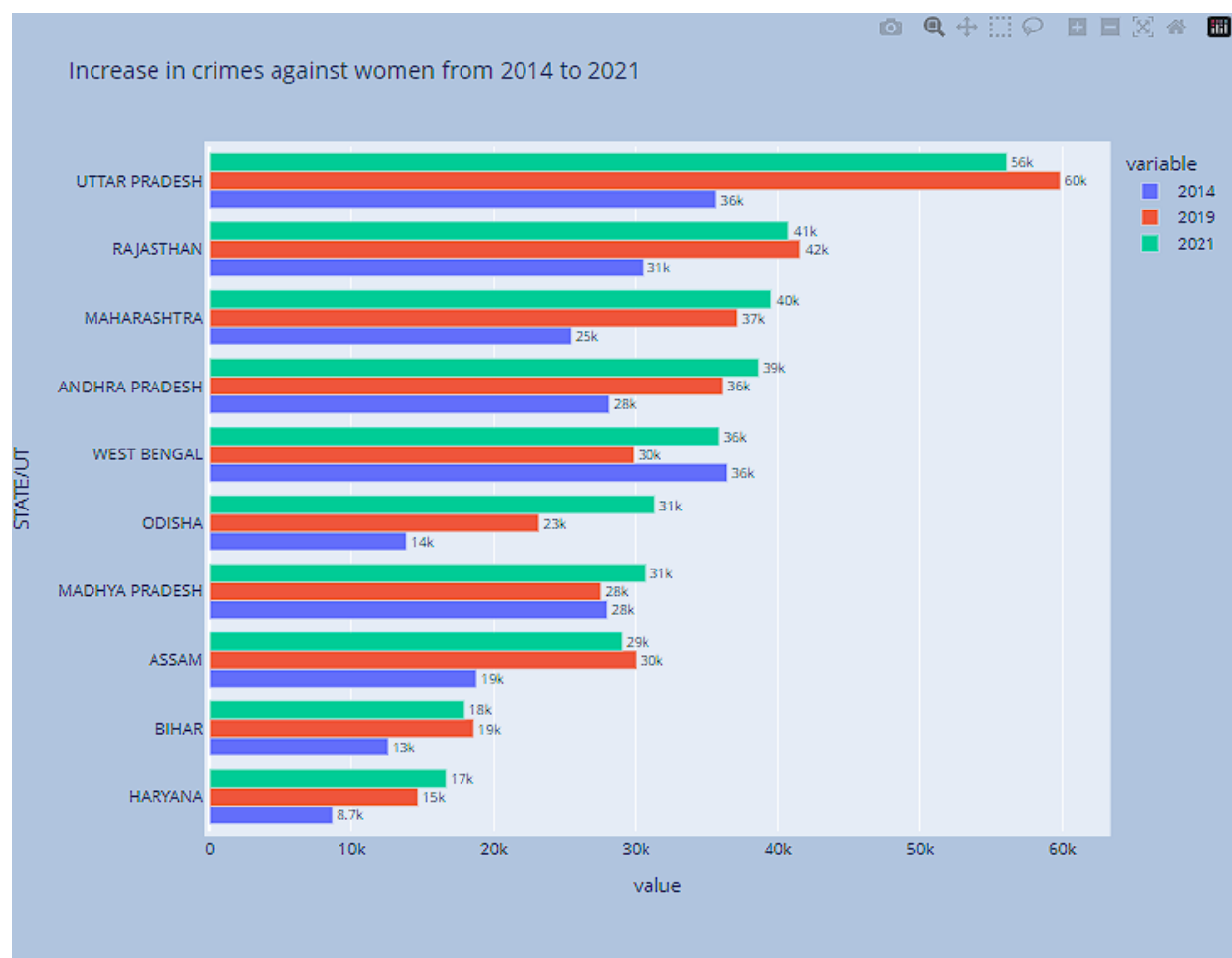

Figure 1Comparing Total Cases in 2014 ,2019 and 2021

From the comparison we can see that cases in Uttar Pradesh have continue rise at the same rate while states like Andhra Pradesh and Madhya Pradesh have worked towards reducing the crimes against women.

4. <u>**CONCLUSION**</u>

With the rapid development of technology, Data Visualization is the most fundamental approach review the historical data and compare It to the present scenario and learn from the other states who have improved. In our case Uttar Pradesh can benefit from the policies made by Andhra Pradesh and Madhya Pradesh in order to reduce crimes against women.

   We can also conclude that many crimes are not being reported and no one has the record about it. To reduce such crimes awareness campaigns should be done to increase the awareness among women.

5. <u>**FUTURE PROSPECTS**</u>

The Crime rate case study shows that the Exploratory Data Analysis using python and its libraries works well in the data analysis of different crimes state-wise and year-wise. Crimes against women is a problem increasing rapidly and women are becoming more and more prone to rapes, kidnapping, assault, and molestation. This is a problem which can be prevented if actions are taken at the right time. The law and order for states having highest rate of crimes should be upgraded. Social Awareness steps should be taken by our government for reducing these crimes, educating people to respect the dignity of women.