

# Analyzing the Spectrum of Criminality: A Comprehensive Study of Crime Types and Rates Across Different States In India

Department of Mathematics & Statistics, IIT Kanpur

## MTH 208 Project Report (Group - 10)

---

Submitted To-

Dr. Dootika Vats

Submitted by-

- Arnab Das
- Krishanu Mukherjee
- Soumo Biswas
- Ashirvad Pawar

## Introduction

In the complex web of India's diverse society and legal system, unraveling the threads of different crimes becomes a vital task. This project, titled "Analysis of the Spectrum of Criminality: A Comprehensive Study of Crime Types and Rates Across Different States in India," is a focused effort to make sense of the intricacies of crime nationwide. India, with its rich blend of cultures, varied economies, and diverse territories, offers a unique backdrop for investigating the array of crimes.

This project embarks on a dedicated effort to delve into the complexities of crime in India. The very nature of our nation, with its multitude of cultures, economic dynamics, and geographical variations, creates a distinct landscape for the study of crimes.

## Objective

We have multiple objectives in our project. Some of our objectives are -

- To categorize and comprehensively understand the various types of crimes prevalent across different states in India. This includes traditional offenses as well as emerging challenges such as cybercrimes.
- To categorize and comprehensively understand the various types of crimes prevalent across different states in India. This includes traditional offenses as well as emerging challenges such as cybercrimes.
- To examine and understand regional disparities in crime rates, identifying hotspots and variations in criminal activities among different states.
- To scrutinize the landscape of cybercrimes, including trends, vulnerabilities, and the readiness of each state to tackle the challenges posed by criminal activities.

## Data

We have the yearwise data for 2018,2019,2020,2021 of all 37 states/union territories and the crimes have been recorded for those recorded years and we have also scrapped a crimewise dataset for all the 37 states.

- Yearwise Dataset.
- Crimewise Dataset.

Yearwise Dataset describes the total crimes,crime rates in all 37 states over the four years.

Crimewise Dataset describes different types of crime occurring in those states over the years.

## Data Collection Process

### Data Scraping :

---

We've used web scraping to collect the data. We collected the data from wikipedia.org.Then we made a R function that can easily scrap the data from the website. We have cleaned the data using various R-pacckages like tidyverse,rvest,dplyr.

## Challenges faced while obtaining data

While obtaining the data we faced certain challenges that took up much of our time. Some of these challenges were:

The data collected as .csv files contained a lot of NA values which were hard to deal with. We couldn't remove it as it resulted in a lot of data loss. So, we had to manually arrange some of the data through articles, websites etc. So, there was some manual work involved as well which took up a considerable amount of our time.

## Key Questions Of Interest

- 1.Which states has the highest crime rate in year 2018, 2019, 2020, 2021 respectively?
- 2.Which states has the lowest crime rate in year 2018, 2019, 2020,2021 respectively?
- 3.In which year a particular state has experienced most crime ?
- 4.How has the rate of change in various crimes occured in different states over the years ?
- 5.Which crime has the highest occurrence in which state?
- 6.How much the crimes are related to each other ?

## Variables Used in Analysis

We have some numerical variables in our analysis. Our variables are mentioned below-

### Categorical Variables

- States/UT Name: Names of 37 states/union territories.

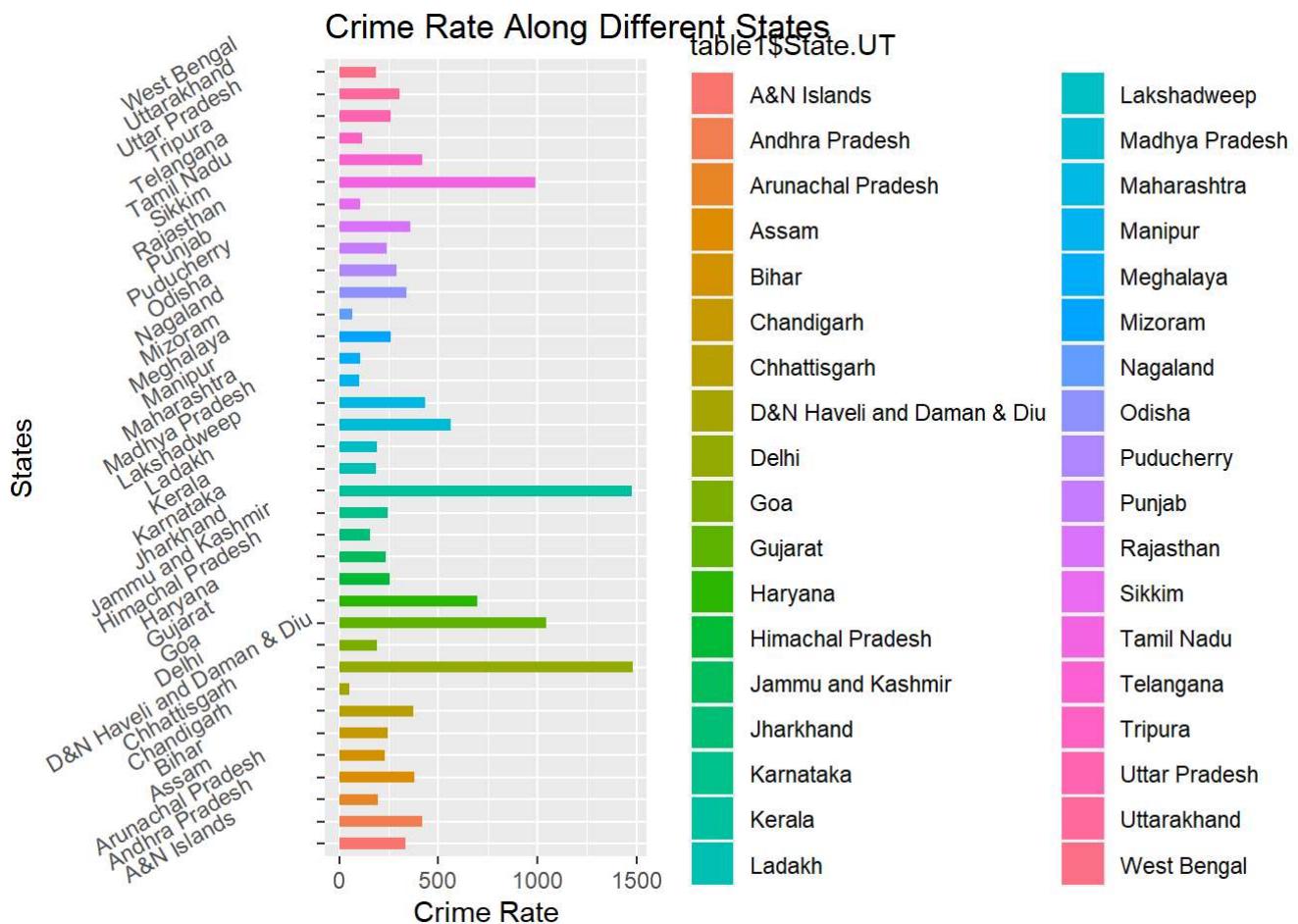
### Numerical Variables

- Crimes: Crimes like Violent Crimes , Kidnapping, MUrder, Extortion, Robbery ,Crimes Against Children.
- Total Crime: It includes the total number of crimes recorded in different crimes.
- Crime Rate : Crime Rate is calculated as crime per one lakh (100,000) of population. That is, rate of cognizable crimes .

## Important Visualizations

Firstly, we've tried to give a visual representation of our dataset to give some of our arising questions.

### Horizontal Bar Plot of States Vs Crime Rates over the years :



So, From the above bar plot we can see that in 2021 the crime rates in Chhattisgarh is the highest followed by Karnataka and Sikkim. For union Territories D&N Haveli and Daman & Diu have the higest crime rates for the year 2021.

### **TREE PLOT Across Different States in India:**

## Statewise division of registered crime cases in 2021

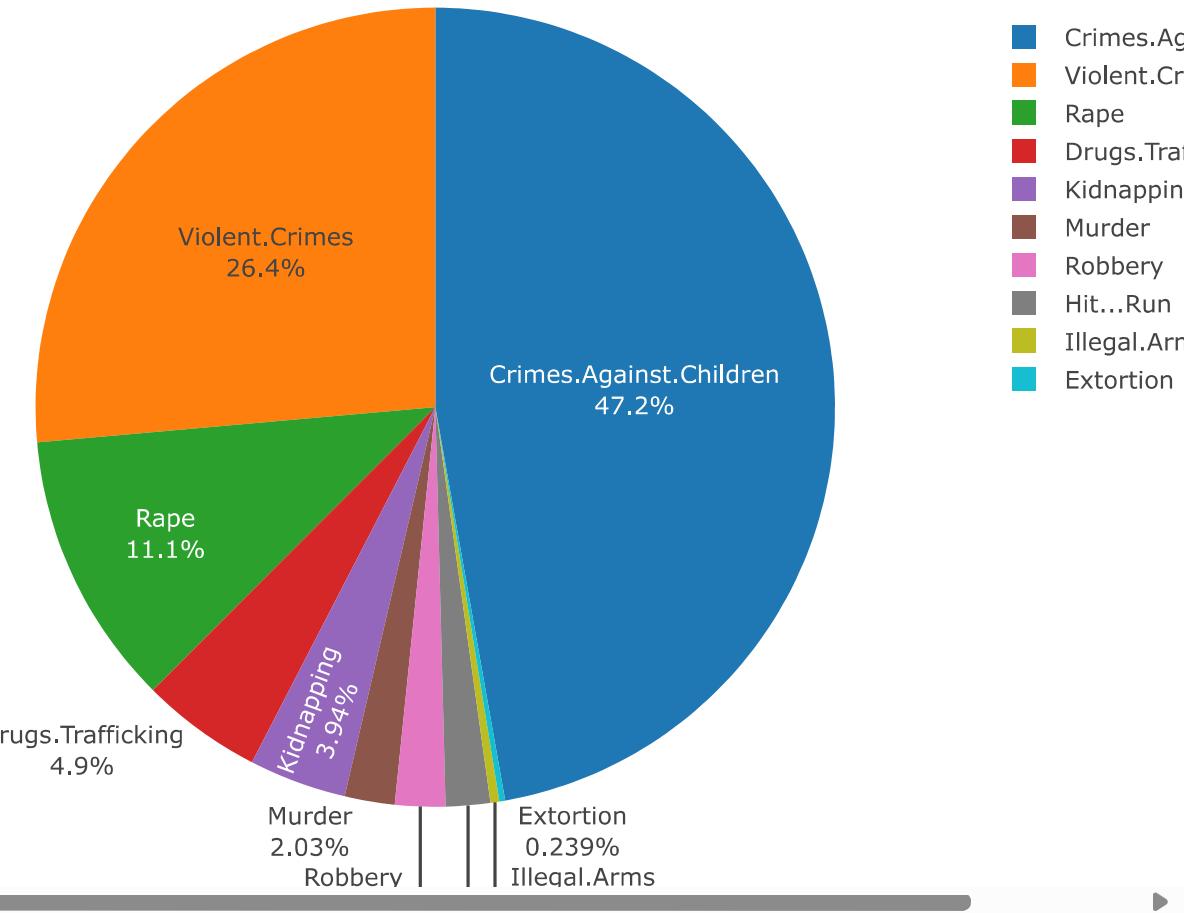


So, from the above Tree plot we see that Jharkhand has the most number of murder cases in 2021 followed by Chattisgarh, Assam, A & N Island and Haryana.

## Pie chart for crimes in Goa in 2021:

```
[1] "data.frame"
```

```
[1] "Violent.Crimes"      "Murder"
[3] "Rape"                "Kidnapping"
[5] "Crimes.Against.Children" "Extortion"
[7] "Robbery"              "Hit...Run"
[9] "Drugs.Trafficking"    "Illegal.Arms"
```

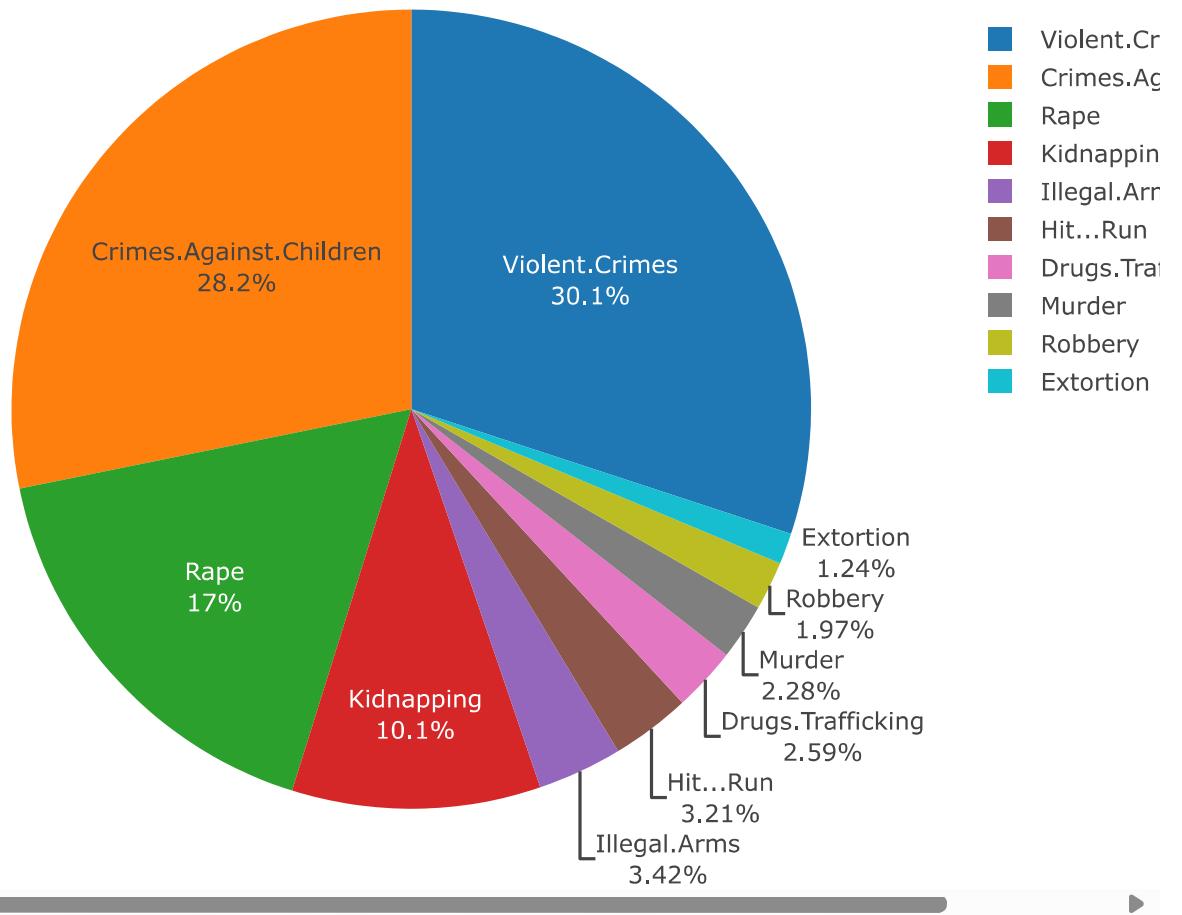


So, from the above Pie Chart it is visible that the crime against children is most frequent in Goa with a percentage of 47.2%, followed by violent crimes 26.4% and rape 11.1% with other crimes occupying the rest percentages.

## Pie chart for crimes in Rajasthan in 2021:

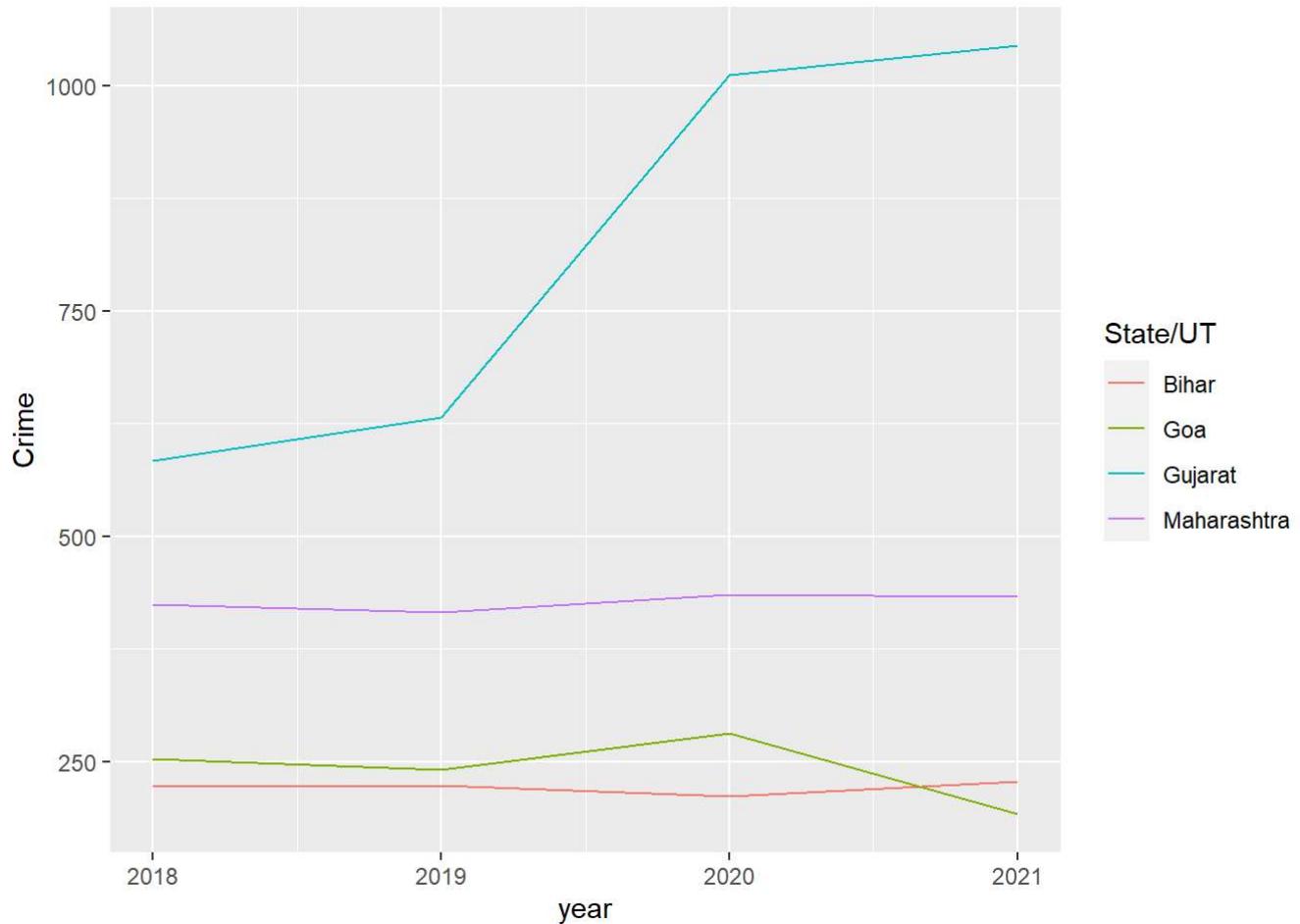
```
[1] "data.frame"
```

```
[1] "Violent.Crimes"      "Murder"
[3] "Rape"                "Kidnapping"
[5] "Crimes.Against.Children" "Extortion"
[7] "Robbery"              "Hit...Run"
[9] "Drugs.Trafficking"    "Illegal.Arms"
```



Above chart says 30.1% of total crimes in Rajasthan is in the category of violent crimes with 28.2% of crimes is against children followed by rape (17%) and kidnapping(10.1%). The rest 24% is of other types of crimes.

## Timewise Crime Analysis for Different States (2018 - 2021):



From the above multiple line plot it is clear that crimes in Gujarat has drastically increased after 2019 during COVID 19 lockdown may not be the only reason for that. Crimes in Goa has slightly decreased after late 2020s. Population of Maharashtra is far larger than Gujarat but crimes in Gujarat (per 1 lakh people) is lot more and this may indicate lack of administration in Gujarat. Crimes in Goa has decreased between 2020 and 2021.

## **Possible Biases in the data**

First of all a huge portion of crimes in India remains unreported; specially crimes against women, children. So it is nearly impossible to find a fully accurate data on crimes in India.

We don't have monthly data of crimes in each year. So, we can't conclude or carry out our analysis on a monthly scale. It damages the accuracy of the results that we have acquired so far.

Sometime a person may be accused of several charges. Again, multiple crimes can occur on a person. In these cases, crimes are counted several times which can inflate the data to an extent. This may be a serious issue while analysing the data.

## **Conclusions**

*As we've seen from the graphs, the factors indeed seem to be inter-related with each other. Moreover, we cannot generalise our assumptions to all the countries as we've already seen some examples of opposing behaviours. Furthermore, a government spending money for the welfare of its country is indeed showing out to yield positive results.*

## **References**

- <[https://en.wikipedia.org/wiki/List\\_of\\_states\\_and\\_union\\_territories\\_of\\_India\\_by\\_crime\\_rate/](https://en.wikipedia.org/wiki/List_of_states_and_union_territories_of_India_by_crime_rate/)>