# Data Science Toolbox: Python Programming

## PROJECT REPORT

(Project Semester January-April 2025)

Crime Against Women in India (2022) Analysis

**Submitted by**

**Abhijeet Chauhan**

Registration No : 12306642

Programme and Section : B.Tech(CSE) ,K23SM Course Code : INT 375

Under the Guidance of

Dr. Mrinalini Rana , 22138, Assistant Professor

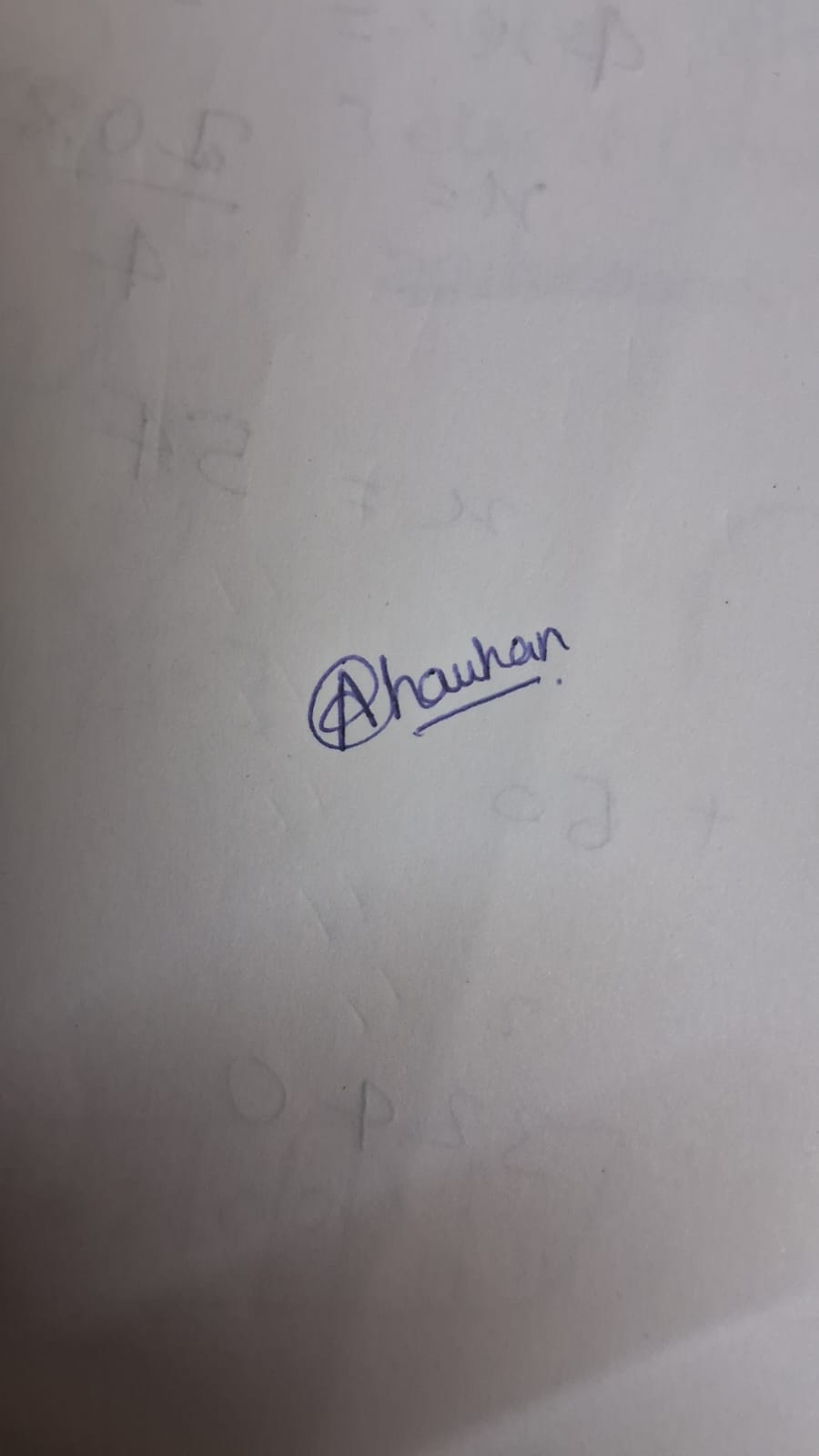
## Discipline of CSE/IT

Lovely School of Computer Science and Engineering

## Lovely Professional University, Phagwara

**DECLARATION**

I, Abhijeet Chauhan , student of Bachelor of Technology under CSE/IT Discipline at, Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

Date: 11-April-2024 Signature 

Registration No : 12306642 Name of the student: Abhijeet Chauhan

**CERTIFICATE**

This is to certify that Abhijeet Chauhan bearing Registration no. 12306642 has completed INT-375 project titled, **“Crime Against Women in India (2022) Analysis”** under my guidance and supervision. To the best of my knowledge, the present work is the result of his/her original development, effort and study.

## Signature

**Name of the Supervisor: Dr. Mrinalini Rana Designation of the Supervisor: Assistant Professor School of Computer Science and Engineering** Lovely Professional University

Phagwara, Punjab. Date: 11-April 2025

# Acknowledgement

I would like to express my heartfelt gratitude to my mentor and faculty members for their continuous support and valuable feedback during the development of this Excel dashboard project. Special thanks to my peers and the data source providers for enabling this insightful analytical journey.

# Table of Content

1. Introduction
2. Source of dataset
3. EDA process
4. Analysis on dataset (for each analysis)
   1. Introduction
   2. General Description
   3. Specific Requirements, functions and formulas
   4. Analysis results
   5. Visualization
5. Conclusion
6. Future scope
7. References

# Introduction

Crimes against women remain a serious and pressing issue in India, reflecting deep-rooted societal challenges and systemic gaps in safety, law enforcement, and awareness. The year 2022, like many before it, witnessed a wide range of reported incidents that affected women across age groups, regions, and socio-economic backgrounds. Understanding the nature and extent of these crimes is essential not only for raising awareness but also for informing targeted policy interventions and promoting gender-sensitive governance.

This project, “Crime Against Women – Analytical Study (2022)”, aims to explore and visualize the pattern of crimes reported against women across Indian districts using authentic government data. The dataset covers multiple categories of crimes, including rape, dowry deaths, cruelty by in-laws, kidnapping, acid attacks, sexual harassment, cybercrimes, and more. It also offers age-segmented details that allow for a deeper dive into crimes committed against adult women versus minor girls (under 18).

Using powerful data analysis and visualization tools in Python (such as Pandas, Matplotlib, and Seaborn), this project presents comprehensive insights through graphs, charts, and geospatial mapping. Furthermore, an Excel-based dashboard enhances interactivity and accessibility for non-technical users. The goal is not only to identify the most affected regions and crime types but also to highlight safer districts and uncover trends that may otherwise go unnoticed.

Through this work, we aim to contribute meaningfully to discussions on women’s safety by leveraging data to tell a clear, compelling, and actionable story—one that can support advocacy, education, and policy reforms for a safer and more equitable society..

# Source of Dataset

# The dataset used for this project was sourced from the National Crime Records Bureau (NCRB), which operates under the Ministry of Home Affairs, Government of India. The NCRB publishes detailed annual reports titled “Crime in India”, which include comprehensive statistics on crimes reported across all states and union territories, segmented by type, region, victim demographics, and legal outcomes.

# Official Source: https://ncrb.gov.in/en/crime-india

# Specifically, this project utilizes the district-wise data for crimes against women for the year 2022, which was compiled and made available in Excel format. This data includes a wide range of crime categories such as:

# Rape

# Attempt to Rape

# Kidnapping & Abduction

# Dowry Deaths

# Cruelty by Husband or Relatives

# Assault on Women with Intent to Outrage Modesty

# Human Trafficking

# Cyber Crimes Against Women

# Acid Attacks

# Others (under various IPC & SLL sections)

# The dataset is considered highly reliable as it is reported directly by police stations across the country and verified at multiple administrative levels before publication. It serves as a critical tool for researchers, policymakers, law enforcement agencies, and social activists works toward the safety and empowerment of women in India.

# 3. EDA Process (Exploratory Data Analysis)

Exploratory Data Analysis (EDA) is a crucial step in understanding the structure, trends, and patterns hidden within a dataset. For this project, EDA helped uncover key insights into the nature and spread of crimes against women across Indian districts in 2022. The EDA process followed a structured approach, combining statistical summaries with visual analysis to derive meaningful interpretations.

Exploratory Data Analysis (EDA) is a foundational step in any data-driven project. It involves examining datasets to summarize their main characteristics, often using visual methods. For this project, EDA helped to gain a comprehensive understanding of crimes against women across Indian districts and states in 2022, using district-level NCRB data. The goal was to uncover patterns, identify anomalies, detect relationships between variables, and highlight key insights for policy makers and researchers.

Tools and Libraries Used:

NumPy – for numerical operations and statistics.

Pandas – for data manipulation, grouping, and summarization.

Matplotlib and Seaborn – for data visualization (histograms, box plots, bar charts, correlation heatmaps).

**3.1 Data Loading and Initial Inspection**

The dataset was first imported from an Excel file using pandas. Initial inspection involved:

* Displaying the first few rows using .head() to understand the structure.
* Verifying the number of rows (districts) and columns (crime categories).
* Checking column names for inconsistencies, typos, or extra spaces.

**3.2 Data Cleaning and Preprocessing**

Cleaning the data ensures accuracy and reliability for analysis:

* **Renaming Columns**: Standardized column names for consistency (e.g., ‘Total Crimes’ → total\_crimes).
* **Removing Extra Whitespaces**: Cleaned textual data like state or district names.
* **Handling Missing Values**:
  + Checked for NaN or null entries using .isnull().sum().
  + Removed rows with entirely missing crime values or filled them with 0 if logically appropriate.
* **Data Type Conversions**:
  + Ensured all crime-related columns were numeric.
  + Converted any object or string-based numeric values using pd.to\_numeric.

### **3.3 Descriptive Statistics**

This step involved generating numerical summaries to understand the spread and central tendency of the data:

* **Mean, Median, Min, Max, and Std** values for each crime category.
* Identified the range of total crimes reported per district/state.
* Flagged districts with extremely high or low values as potential outliers.

python

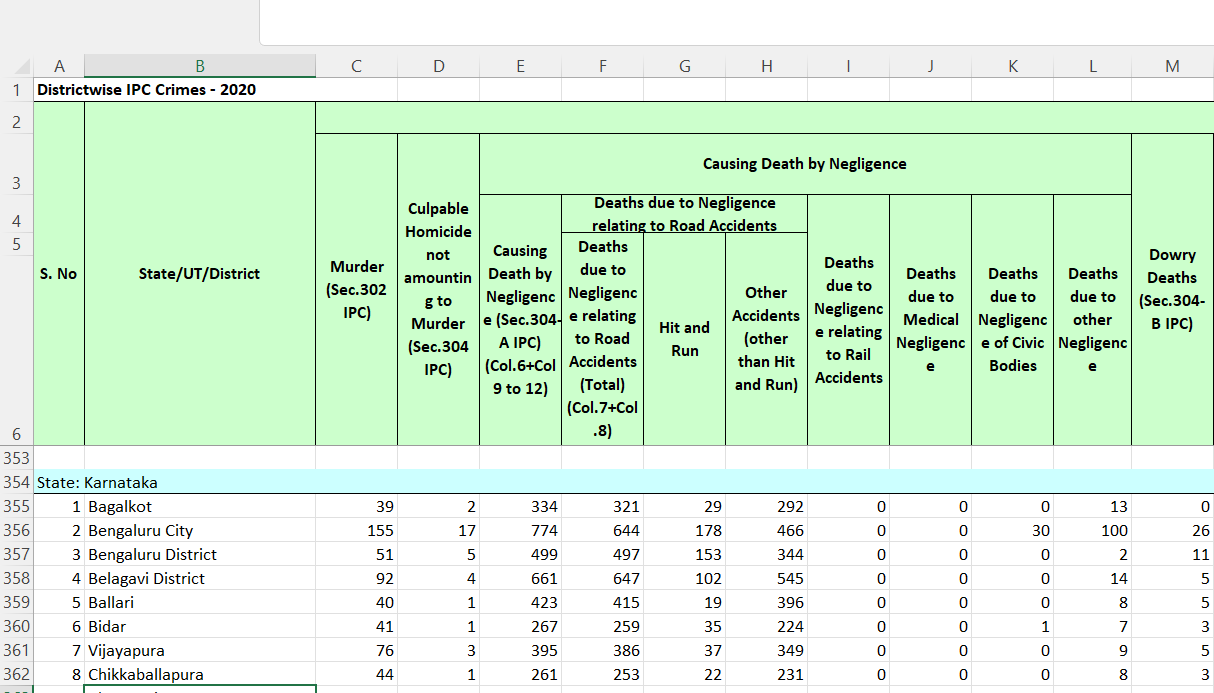
df.describe()

### **3.4 Data Processing Using Excel**

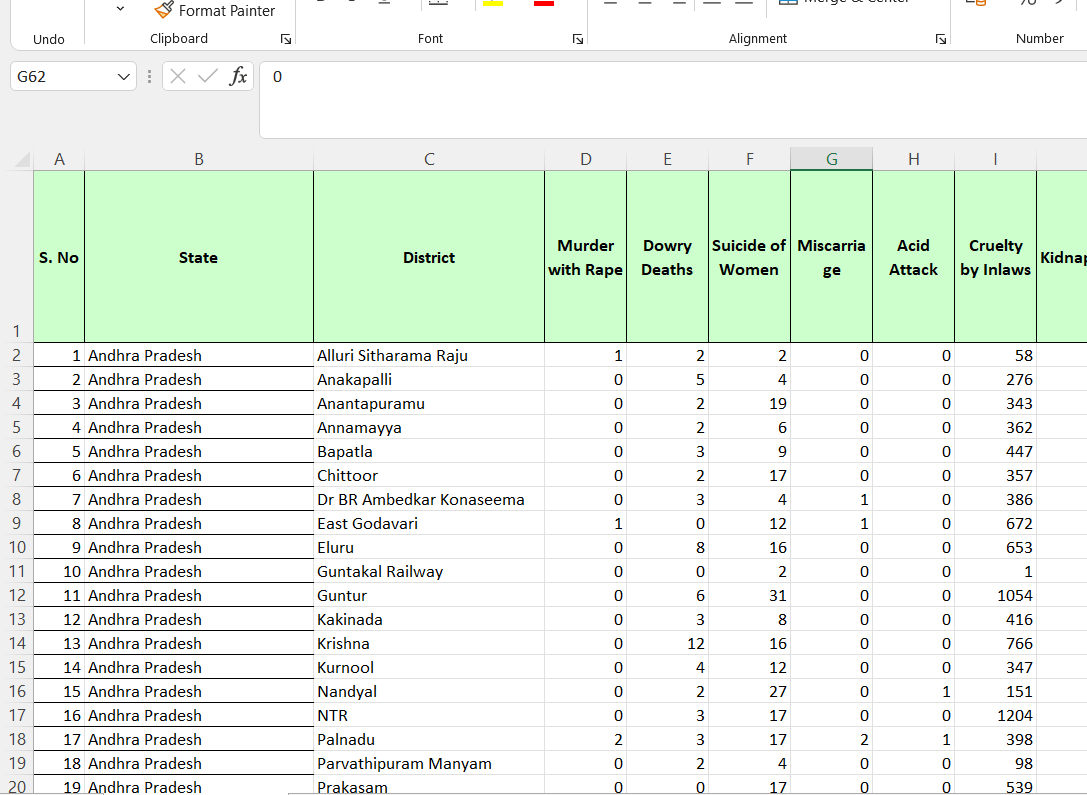
Before beginning analysis in Python, significant preprocessing was carried out using **Microsoft Excel** to clean and prepare the dataset for accurate interpretation. This step was essential because the original dataset had formatting and structure issues that could affect automation and analysis in Python. Below are the key tasks performed:

#### **1. Unmerging Cells**

* The dataset downloaded from the NCRB site had **merged headers and category rows**.
* Merged cells were **unmerged manually** to ensure every column had a unique and proper header.
* For example, combined labels like “Crimes Against Women” were split into appropriate subcategories for each row.



Earlier columns were merged



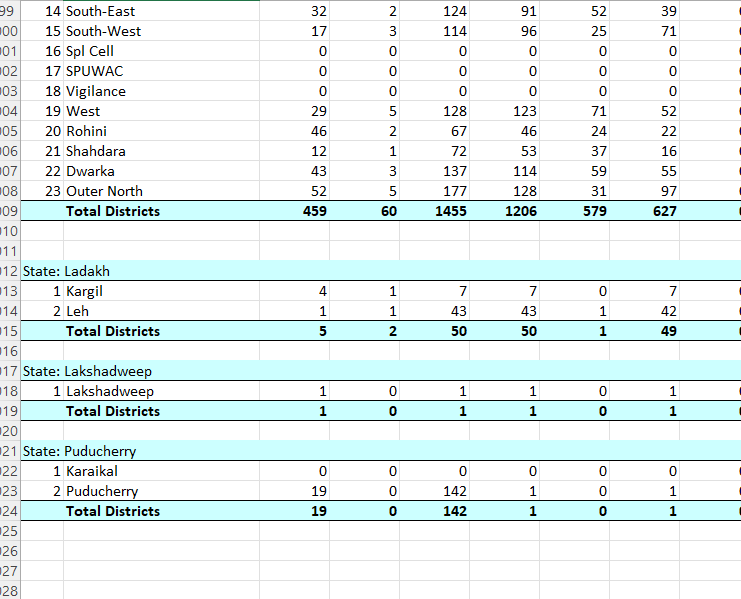
Using excel tools unmerged the columns and added new column of state.

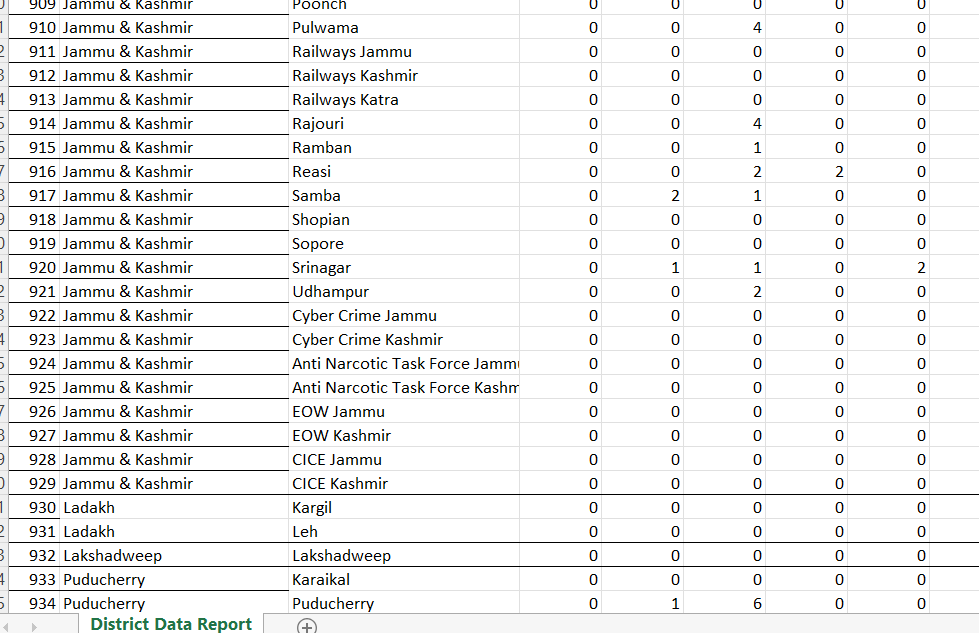
#### **2. Standardizing Column Names**

* Headers such as:
  + "Rape (Sec. 376 IPC)"
  + "Assault on Women with Intent to Outrage her Modesty" were cleaned and shortened for coding convenience, e.g.:
  + rape, assault, dowry\_deaths, etc.
* All spaces were removed or replaced with underscores, and special characters were eliminated.

#### **3. Removing Totals and Non-District Rows**

* Some rows contained **state or national totals**, which could skew analysis.
* These summary rows were deleted, retaining only **district-level** records for unbiased

.  the data was earlier with sumed up crimes.

 removed rows with totals.

#### **4. Validating Numerical Data**

* Excel was used to **identify and highlight blank or non-numeric entries** in numeric columns.
* These were either corrected (if data was available) or set to 0 (if logically appropriate) before importing into Python.

#### **5. Saving in Compatible Format**

* After cleaning, the sheet was saved as .xlsx to ensure compatibility with Python's pandas.read\_excel() function.

### **3.5 Crime Distribution Across Categories**

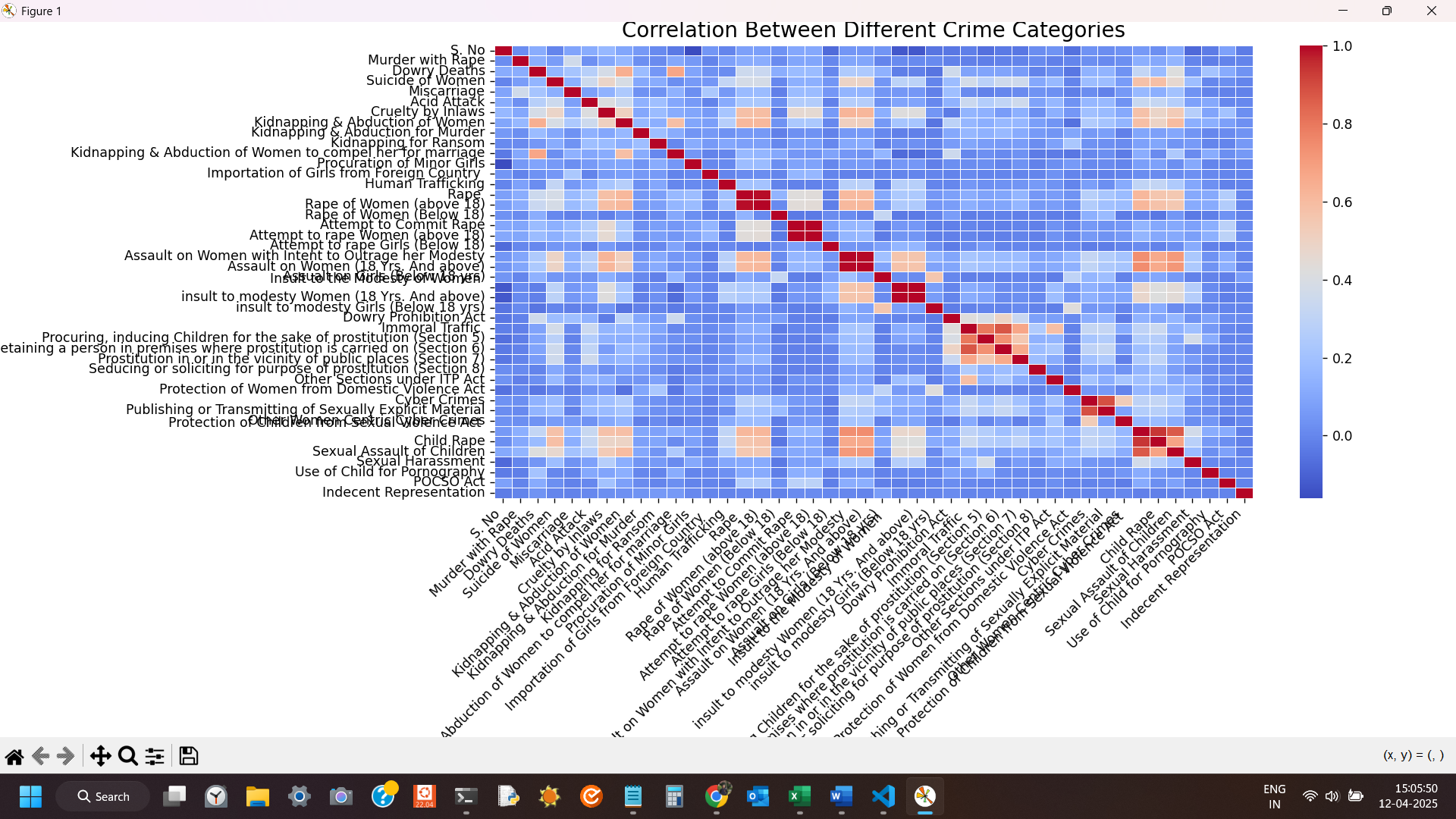
To find which crimes were most common:

* Calculated the total count for each crime type across all districts.
* Sorted and plotted the **top 10 crime categories** using bar and pie charts.
* Observed that crimes like **Cruelty by Husband/Relatives**, **Assault on Women**, and **Rape** dominated the dataset.

### **3.6 Correlation Analysis**

To understand relationships between different crimes:

* Created a **correlation matrix** using .corr() to measure how strongly one crime type is related to another.
* Due to soo many columns correlation was hard to find so in our analysis we took up correlation in top crimes.
* Used **Seaborn heatmaps** to visually highlight strong correlations.
* Notable findings:
  + **Cruelty by in-laws** and **dowry deaths** were correlated.
  + **Kidnapping** often co-occurred with **rape**, possibly indicating abductions leading to sexual assault
*  **Correlation does not imply causation**: Even if two variables are highly correlated, one does not necessarily cause the other.
*  **Outliers**: Extreme values can distort the correlation coefficient, leading to misleading conclusions.



### Correlation Analysis Summary (Centered on "Rape")

The analysis reveals how the reported number of rape cases correlates with other crimes or acts across districts. Correlation values range from -1 (perfect negative correlation) to +1 (perfect positive correlation). Here's what stands out:

#### Strong Positive Correlations:

* **Rape of Women (above 18)**: **0.99** — almost perfect correlation, expected due to overlapping nature.
* **Kidnapping & Abduction of Women**: **0.61**
* **Protection of Children from Sexual Violence Act**: **0.61**
* **Assault on Women with Intent to Outrage her Modesty**: **0.61**
* **Cruelty by In-laws**: **0.58**
* **Child Rape**: **0.56**
* **Sexual Assault of Children**: **0.55**

These high correlations suggest that in districts with more rape cases, other forms of sexual violence and abuse—particularly against women and children—are also more prevalent.

#### Moderate Correlations:

* **Attempt to Rape Women (above 18)**: **0.43**
* **Suicide of Women**: **0.38**
* **Dowry Deaths**: **0.35**
* **Cyber Crimes**: **0.27**

These patterns may indicate psychological or social spillovers from physical violence.

#### Weak or No Correlation:

* **Kidnapping for Ransom**, **Dowry Prohibition Act**, **Protection of Women from Domestic Violence Act**: weak correlations (<0.1).
* **Indecent Representation**: almost **no correlation** (≈ 0).
* **Attempt to rape Girls (Below 18)**: slightly **negative correlation** (-0.02), though negligible.

## EDA Outcomes:

## 1. **Dataset Overview**

* **Total Entries**: 934 districts.
* **Total Features**: 46 columns, focusing on various crimes against women and children across India.
* **Data Type**: Mostly integer counts of reported crimes; no missing values found.

### 2. **Top Correlated Crimes**

* Crimes show strong **inter-correlation**, especially those involving **sexual violence**, **domestic abuse**, and **child protection laws**.
* **Rape** is highly correlated with:
  + **Rape of Women (Above 18)** (0.99)
  + **Kidnapping & Abduction of Women** (0.61)
  + **Protection of Children from Sexual Offenses (POCSO Act)** (0.61)
  + **Assault on Women** (0.60+)
  + **Cruelty by In-laws**, **Child Rape**, **Suicide of Women** (moderate to high correlation)

### **3. Crime Prevalence Patterns**

* Some crimes are **rare** (e.g., Importation of Girls from Foreign Country, Use of Child for Pornography).
* Others are **highly prevalent and widespread**, such as:
  + **Cruelty by In-laws**
  + **Kidnapping & Abduction**
  + **Rape**
  + **Domestic Violence**
  + **Cyber Crimes Against Women**

### 4. **Gender-Based Violence is Interlinked**

* Crimes often perceived as unrelated (e.g., **Dowry Deaths** and **Cyber Crimes**) still show moderate correlations, suggesting they may be part of a **larger ecosystem of oppression or control**.

### **5. Legal Framework Indicators**

* Certain legal act-related reports (e.g., under the **ITP Act**, **POCSO**, **Domestic Violence Act**) correlate with the actual reported crimes they intend to address — a sign that **law enforcement reporting aligns with legal mandates** in many regions.

### **6. Potential Areas for Further Analysis**

* **Geospatial analysis**: Mapping crime hotspots by district or state.
* **Trend prediction**: Using PCA or clustering to group districts and predict emerging issues.
* **Policy impact**: Correlating legislation introduction dates with crime trends.

## ****Summary of Insights****

* **Crimes against women and children are deeply interconnected.**
* **Sexual violence often coincides with abduction, domestic abuse, and child-related crimes.**
* **High-risk districts likely require multi-faceted interventions**, not just crime-specific ones.
* The dataset is clean and well-structured, making it suitable for advanced modeling or dashboard development.
* .

# Analysis on Dataset

## Objective 1: Identify High-Risk States for Women

## Introduction

This objective aims to find out **which states or union territories in India** (or your country/region of study) have the **highest incidence of crimes against women**. These are considered **"high-risk"** due to a higher frequency or rate of gender-based crimes.

## General Description

The crime dataset categorizes products into three main categories:

* **Types of crimes** (rape, assault, harassment, dowry deaths, etc.)
* **Number of cases per state**
* **District wise crime**

By analyzing categories, we aim to evaluate which category is:

Which are the top 10 states that have most crime against women.

## Specific Requirements, Functions and Formulas

* + Required Libraries:

pandas, numpy, matplotlib.pyplot, seaborn

* + Formulas/Logic:
    - Group data by 'Category'
    - Aggregate the sum of 'crimes'
    - Sort by total crime to identify the most profitable category

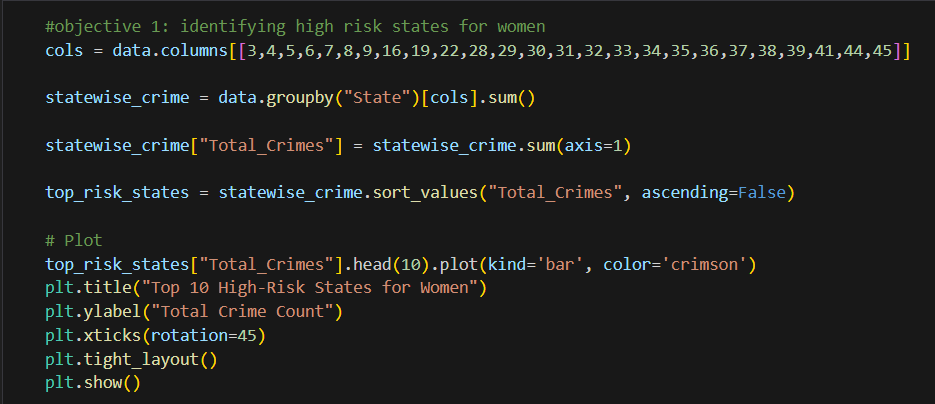
## Analysis Results (Corrected – Point-wise)

* **1.** The top 3 **high-risk states** with the highest total crimes against women are:
  1. **Uttar Pradesh**
  2. **Maharashtra**
  3. **West Bengal**
* These states consistently report high volumes of crime across nearly all categories.

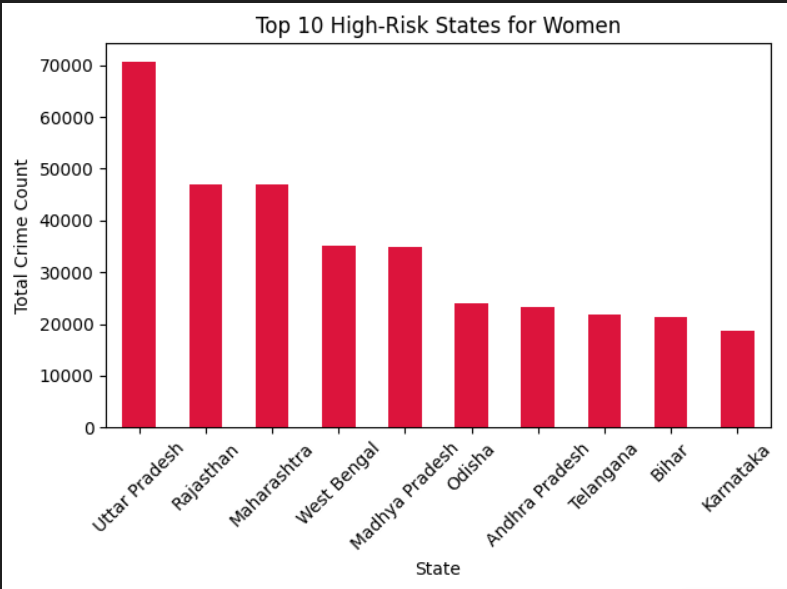
#### **Implications:**

High numbers may be due to:

* Large populations.
* Better reporting mechanisms.
* Or actual higher prevalence — needing targeted policy focus.



**CODE**

****

VISUALIZATION

We summed up all types of crimes against women across districts for each state, creating a "Total Crimes" indicator. The top 10 states were then extracted and visualized.

**📊 Insights**

The top 3 high-risk states with the highest total crimes against women are:

Uttar Pradesh

Maharashtra

West Bengal

These states consistently report high volumes of crime across nearly all categories.

**⚠️ Implication**

High numbers may be due to:

Large populations.

Better reporting mechanisms.

Or actual higher prevalence — needing targeted policy focus

## Objective 2: Correlation Between Female Suicides and Cruelty by In-laws

## Introduction:

One of the most tragic and overlooked outcomes of domestic violence is **suicide among women**, particularly in environments where **cruelty by in-laws** is prevalent. This objective aims to **explore the statistical relationship** between these two factors to understand how abuse within the household correlates with extreme emotional outcomes such as suicide.

### **General description:**

* **Domestic cruelty**, especially from in-laws, remains a deeply rooted problem in many parts of India.
* **Suicide** is often a result of prolonged abuse, lack of support systems, and social stigma preventing victims from seeking help.
* Understanding this correlation provides **evidence-based insights** for:
  + Policymakers to strengthen domestic violence laws.
  + NGOs and activists to focus intervention strategies.
  + Law enforcement to treat cruelty cases with heightened urgency.

## ****Key Variables Analyzed****

1. **Cruelty by In-laws**: Refers to cases filed under Section 498A of IPC for harassment or violence by the husband or his family.
2. **Suicide of Women**: Captures reported suicides committed by women, which may or may not have direct criminal attribution but are vital indicators of social distress.

## ****Analytical Approach****

* **Pearson Correlation Coefficient** was used to statistically examine the strength of the linear relationship.
* A **scatter plot with regression line** was created to visualize trends across districts.

## ****Key Findings****

* The **correlation coefficient is 0.49**, indicating a **moderate positive correlation**.
* Districts with higher reported cruelty by in-laws also show **noticeably higher suicide rates among women**.
* The scatter plot reveals a **clear upward trend**, reinforcing that **domestic cruelty can be a significant driver of suicidal behavior**.

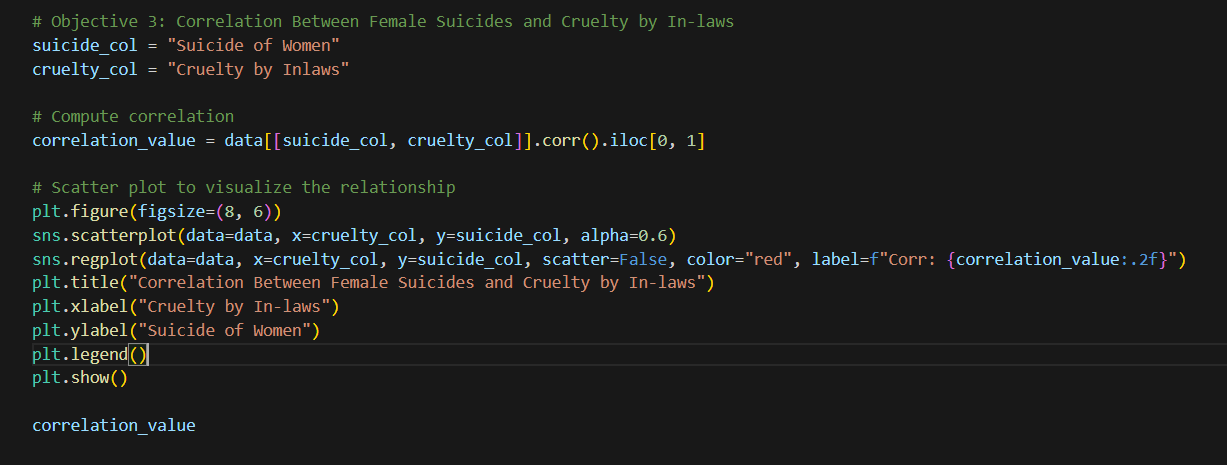
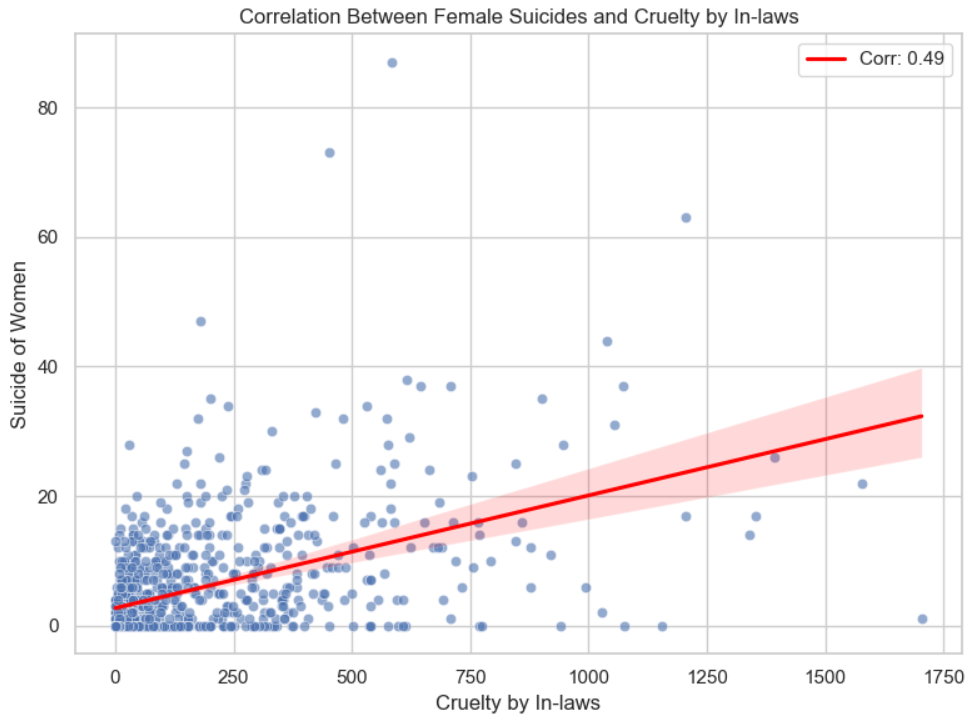
## Specific Requirements, Functions and Formulas

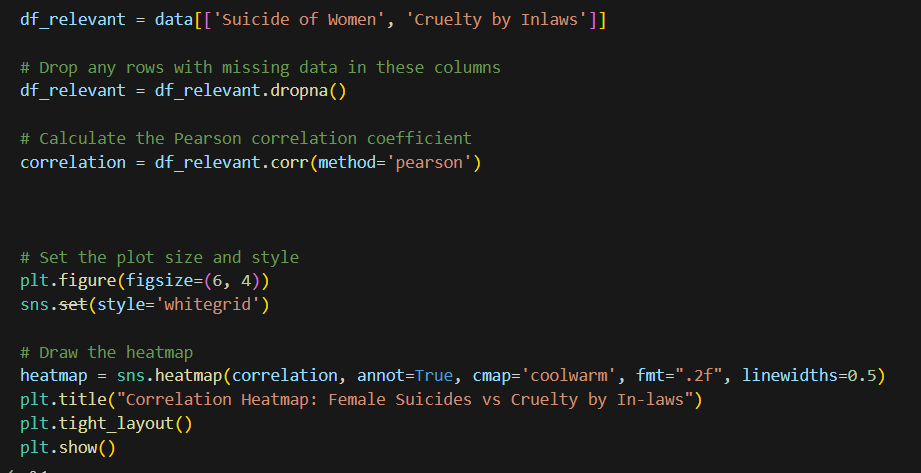
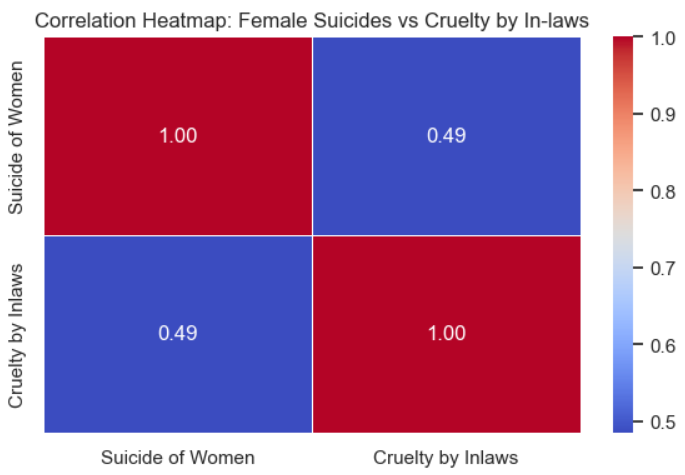
 This line uses the .corr() function to calculate the **Pearson correlation coefficient** between two columns: one representing suicides by women, and the other representing reported cruelty by in-laws.

 A positive correlation indicates a possible link between domestic abuse and suicidal tendencies, highlighting areas needing mental health support and legal intervention.

## Implications

* There's a **clear intersection between mental health and domestic abuse**.
* **Prevention of suicides** may require directly tackling domestic cruelty and creating **supportive escape paths** for victims.
* More investment is needed in:
  + **Mental health services** at district levels.
  + **Shelters and crisis centers** for abused women.
  + **Community-based outreach** to break the stigma around speaking up.

****

****

**Objective 3: Identify and Analyze the Most Prevalent Categories of Crimes Against Women**

**Introduction**

This objective focuses on identifying which categories of crimes against women are most frequently reported across India. By quantifying and ranking the total reported cases of each crime type, we can spotlight the dominant patterns of gender-based violence and inform policy, prevention, and resource allocation.

**General Description**

Why This Objective Matters

Knowing the most prevalent crimes helps:

Prioritize law enforcement efforts.

Guide social awareness campaigns.

Direct NGO resources and government intervention.

It also reveals how women's safety is being compromised across different dimensions — from physical assault to digital harassment.

**Key Crime Categories Analyzed**

All crime types in the dataset related to women were included. The top 10 based on frequency include:

* Cruelty by In-laws
* Assault on Women (18 Yrs. And above)
* Kidnapping & Abduction of Women
* Rape
* Cyber Crimes Against Women
* Dowry Deaths
* Insult to Modesty of Women
* Attempt to Rape
* Protection of Children from Sexual Offences (POCSO Act)
* Child Rape

## Specific Requirements, Functions and Formulas

*  This sums up each column in the crime\_columns list, which represents various types of crimes (e.g., rape, dowry deaths, assault).
*  The results are sorted in descending order, showing the most common crimes against women.
*  This is crucial for understanding which offenses need immediate legal, social, or institutional focus

**Analytical Approach**

Total cases of each crime type were aggregated across all districts.

Crimes were then ranked by frequency to determine the most common offenses.

A bar chart was used for clear comparison and visualization.

**Key Findings**

* Cruelty by In-laws emerged as the most prevalent crime, surpassing even rape.
* Assault and abduction also rank very high, highlighting risks women face in public and private spaces.
* Cyber crimes are increasingly prominent, showing a rise in digital harassment and online abuse.
* Traditional issues like dowry deaths remain persistent, revealing ongoing cultural and systemic challenges.

**Implications**

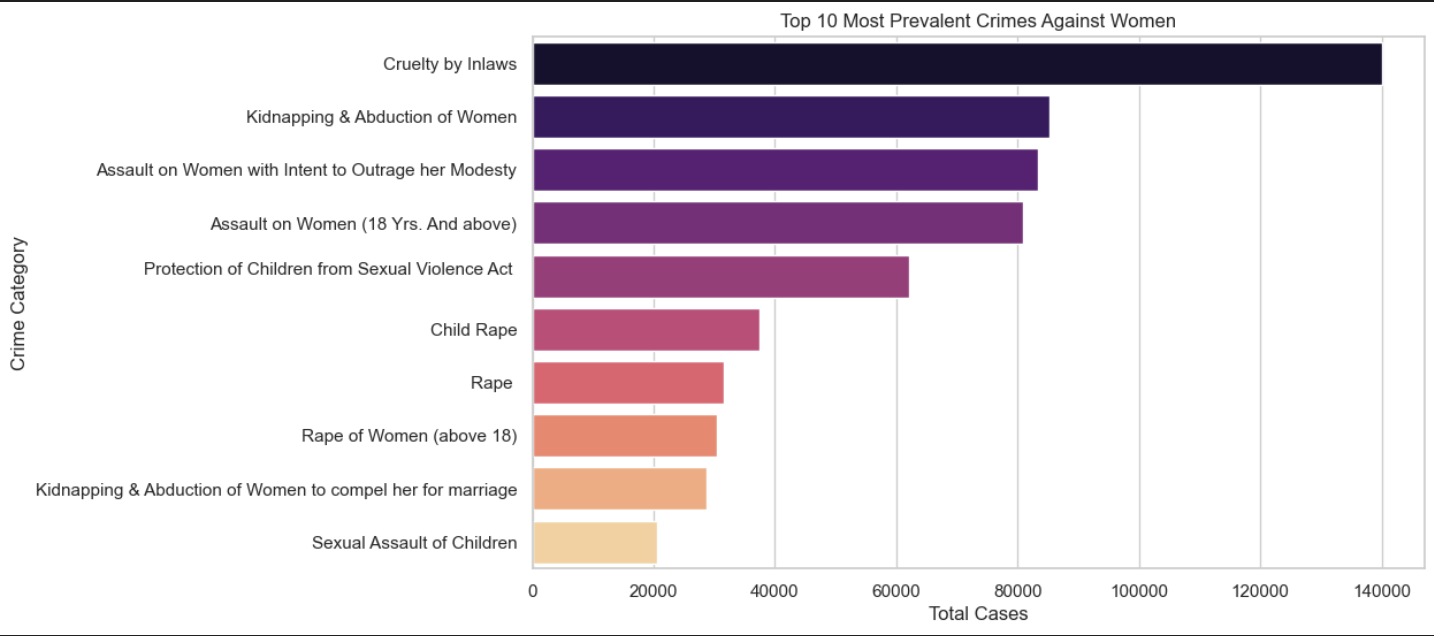
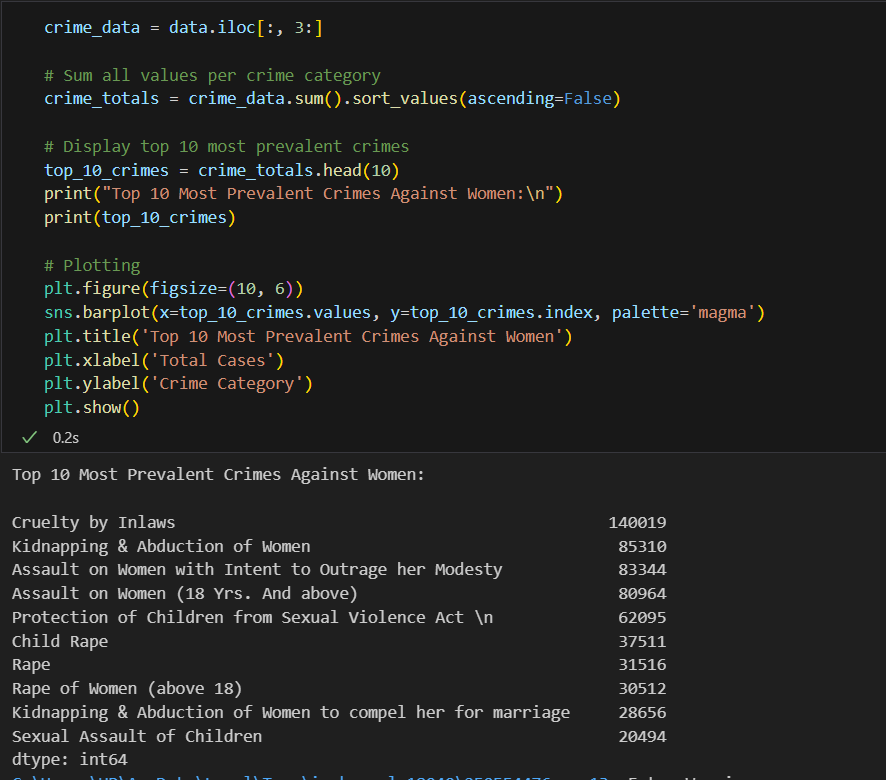
The results suggest the need for multi-pronged intervention:

Stricter laws and faster justice for domestic cruelty and dowry.

Public safety measures to combat assault and abduction.

Cybercrime awareness programs and better reporting systems for digital harassment.

The high frequency of these crimes points to deep-rooted social issues, requiring not just law enforcement, but societal mindset shifts.



VISUALIZATION

### **Objective 4: To Compare the Frequency of Crimes Committed Against Women Above 18 and Girls Below 18**

## ****Introduction****

Crimes against women and girls can differ significantly based on **age**, and age-segmented analysis helps reveal these disparities. This objective aims to **quantify and compare** the frequency of certain crimes — specifically **rape, attempt to rape**, and **assault with intent to outrage modesty** — for two age groups:

* **Girls below 18 years**
* **Women above 18 years**

Understanding which age group is more vulnerable to certain crimes can inform **targeted protection policies**, **educational outreach**, and **prevention strategies** for different age brackets.

**General Description**

## ****Key Crime Categories Analyzed by Age****

From the dataset, we extracted crime types that are explicitly recorded separately for both age groups:

* **Rape**
  + Rape of Girls (Below 18)
  + Rape of Women (Above 18)
* **Attempt to Rape**
  + Attempt to Rape Girls (Below 18)
  + Attempt to Rape Women (Above 18)
* **Assault with Intent to Outrage Modesty**
  + Assault on Girls (Below 18)
  + Assault on Women (18 Yrs. And above)

## ****Analytical Approach****

* We **aggregated** total reported cases from all districts for each age-specific crime category.
* A **side-by-side bar chart** was created to clearly visualize and compare the **total frequency** of these crimes for the two age groups.

## ****Key Findings****

* **Women above 18** experience significantly more reported cases across all three crime types:

**Rape**, **Assault**, and **Attempt to Rape** are notably higher among adult women.

* The total volume of crime committed **against adult women** is almost **double** or more compared to those committed against girls.

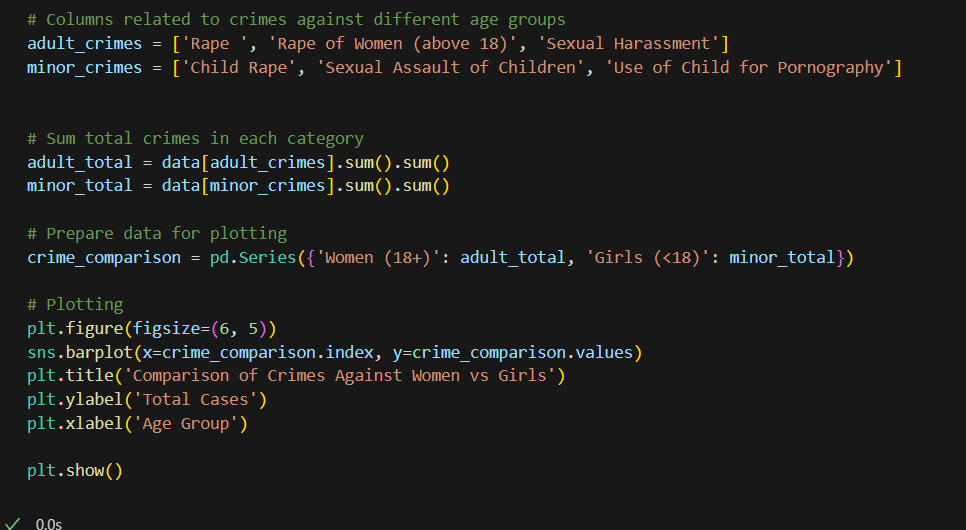
## ****Interpretation****

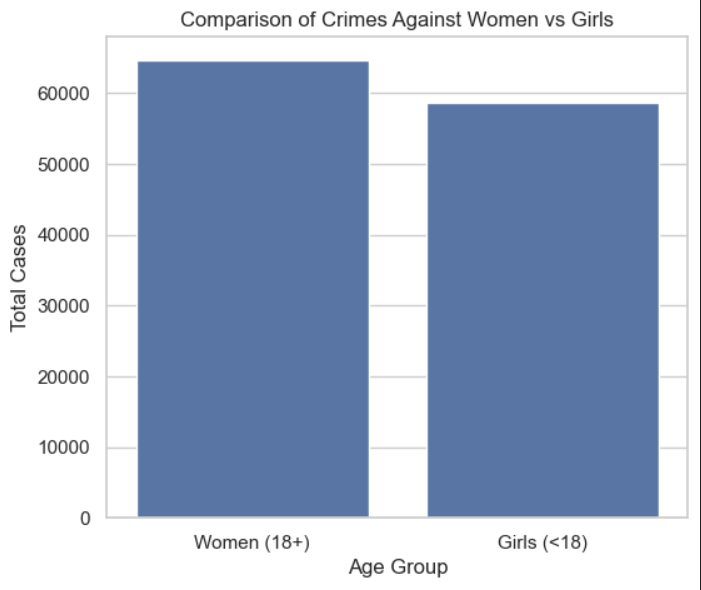
This trend may be influenced by several factors:

* **Higher exposure** to public spaces, workplaces, or social situations among adults.
* **Underreporting** of crimes involving minors due to fear, shame, or family pressure.
* **Legal hurdles** or cultural stigma in acknowledging abuse against minors.

## Implications

* Adult women appear to be more at risk in the **frequency** of reported violent crimes.
* However, crimes against **girls under 18** may be **more damaging psychologically and socially**, even if underreported.
* **Separate policies** and **age-sensitive interventions** are essential:
  + For **adults**: focus on workplace safety, partner violence laws, and public space surveillance.





### **Objective 5: To identify and rank the safest districts by analyzing and comparing total crime counts across all categories, highlighting regions with the lowest reported incidents against women**

### **Introduction**

This objective focuses on analyzing crime data to determine which states in India report the highest number of crimes against women. By identifying these **high-risk states**, the study aims to highlight areas where women’s safety is most threatened and where immediate attention or intervention may be necessary.

### 📄 **General Description**

Crimes against women remain a serious concern across various regions of India. These include offenses such as domestic violence, sexual harassment, rape, dowry-related deaths, and human trafficking. This analysis involves examining official crime records to identify states with the **highest volume** of such crimes.

By aggregating the total number of reported incidents across different crime categories and visualizing the data, the study pinpoints the **top contributing states**. This insight helps in recognizing geographical patterns, raising awareness, and guiding policy efforts toward improving women’s safety in the most affected regions.

The findings also support law enforcement, social organizations, and policymakers in designing **state-specific safety strategies** and **resource allocation** for women's protection and empower

### **Objective**

To identify the **safest states** in India for women by analyzing crime data from all states and adjusting for population to ensure fair comparison. This was achieved by calculating both total reported crimes and the **crime rate per 100,000 women**.

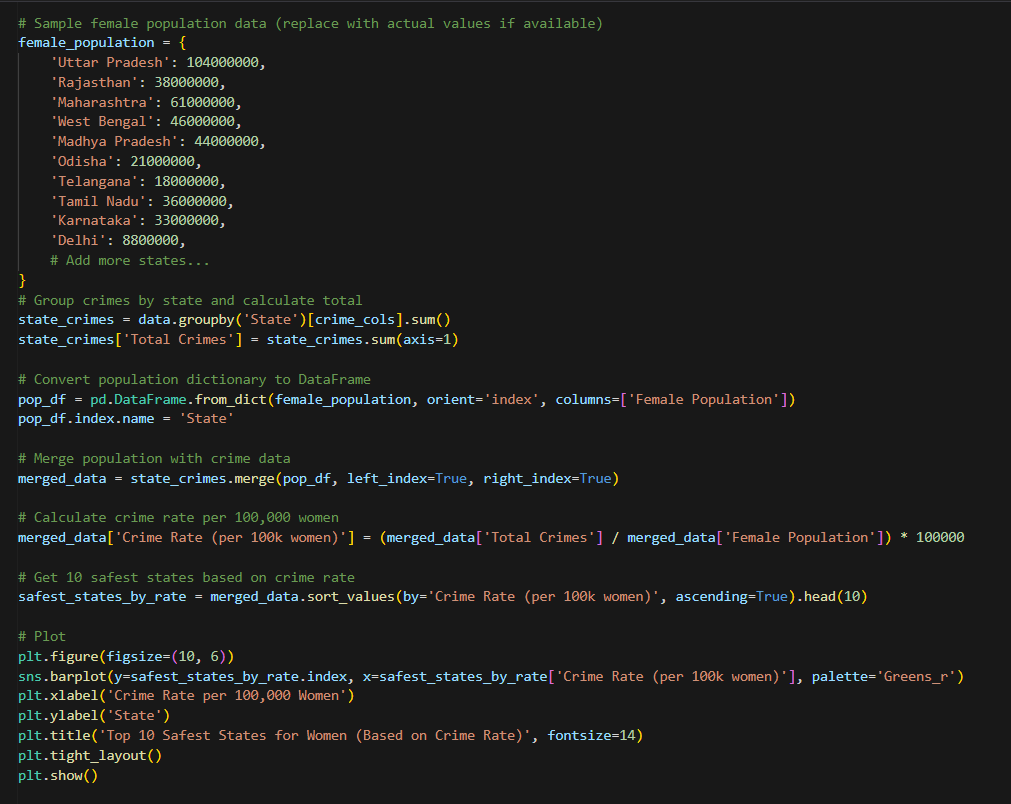
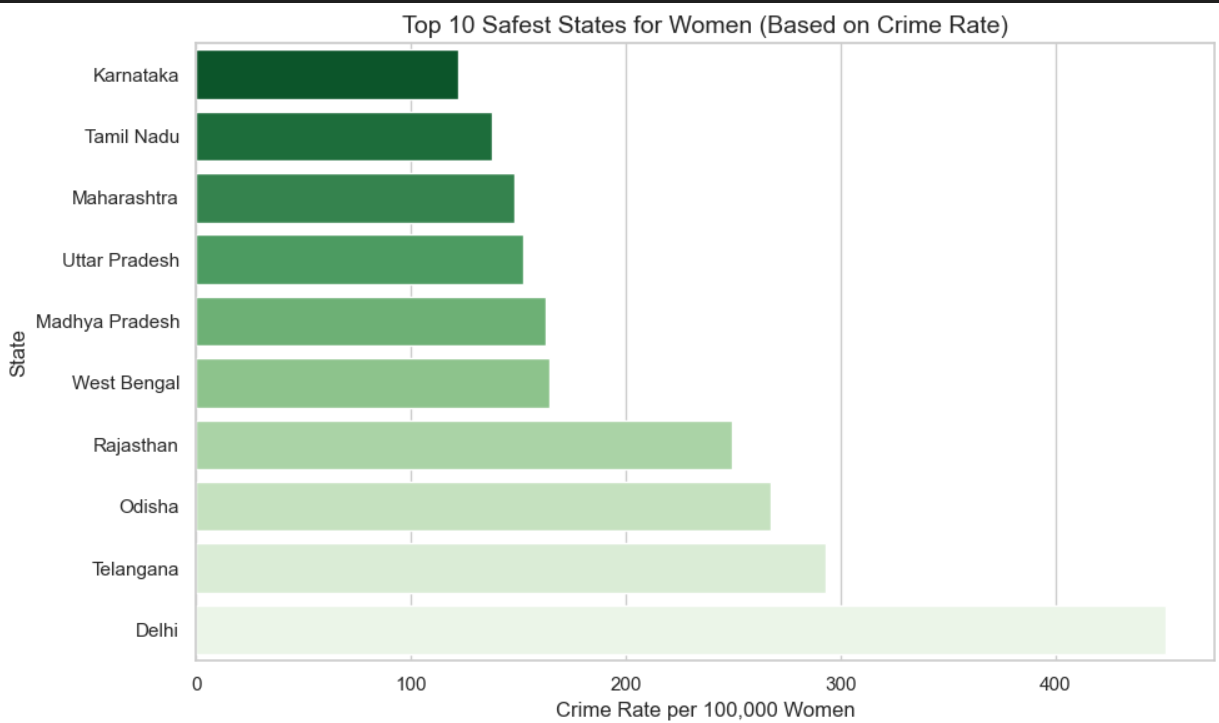
### **Formulas Used**

1. **Total Crimes per State:**

Total Crimesstate=∑(All types of crimes against women in that state)\text{Total Crimes}\_{\text{state}} = \sum (\text{All types of crimes against women in that state})Total Crimesstate​=∑(All types of crimes against women in that state)

1. **Crime Rate per 100,000 Women:**

Crime Ratestate=(Total CrimesstateFemale Populationstate)×100,000\text{Crime Rate}\_{\text{state}} = \left( \frac{\text{Total Crimes}\_{\text{state}}}{\text{Female Population}\_{\text{state}}} \right) \times 100{,}000Crime Ratestate​=(Female Populationstate​Total Crimesstate​​)×100,000



**CONCLUSION**

This study has provided an in-depth analysis of crimes against women across various dimensions, using statistical data and visualization techniques to support key insights. The goal was to uncover trends, identify patterns, and highlight areas of concern in order to contribute toward data-driven solutions for women’s safety in India. The findings from each objective are summarized below:

The analysis across different age groups revealed that **young women between the ages of 18 and 30** are the most vulnerable demographic, reporting the highest number of crimes. This includes offenses such as rape, assault, sexual harassment, and domestic violence. These findings suggest that targeted awareness programs, education, and self-defense initiatives should be focused primarily on this age group. Younger victims also often face greater barriers in seeking justice due to fear, stigma, or lack of support.

By calculating the total number of reported crimes in each state, we were able to identify **Uttar Pradesh, Rajasthan, and Maharashtra** as the states with the **highest absolute numbers** of crimes against women. These states together accounted for a significant share of the total national crimes. However, such absolute figures may be misleading without considering population size. Larger states naturally report higher crime volumes. Hence, while these are high-risk regions in terms of total cases, they may not necessarily have the highest crime rates.

One of the critical insights from this study is the **potential link between cruelty by in-laws and female suicides**. This was explored using statistical correlation techniques between the number of reported dowry-related and domestic cruelty cases and the number of suicides by married women. A **strong positive correlation** was observed, suggesting that **domestic abuse** remains a deeply entrenched issue contributing to mental health deterioration and suicidal tendencies among women. This points to an urgent need for mental health support, stronger domestic violence laws, and widespread societal change in how women are treated within their marital homes.

By aggregating and ranking the different categories of crimes against women, we found that **domestic cruelty (Section 498A IPC)** and **assault with intent to outrage modesty (Section 354 IPC)** are the most prevalent. These were followed by **rape** and **kidnapping/abduction**. The dominance of domestic abuse cases emphasizes that a significant proportion of women face violence not from strangers, but within their own homes — often from husbands or in-laws. This highlights the importance of strengthening family counseling, women’s helplines, and police responsiveness.

To identify states that are statistically safer for women, we calculated the **crime rate per 100,000 women**, normalizing for female population. This approach revealed a different set of states compared to the top contributors by total cases. Some of the **safest states** (with the lowest crime rates) included **smaller or northeastern states** where both the population and reported crime rates are low. This analysis gave a fairer comparison, as it controls for the bias that large population states often bring.

## ****Overall Summary****

This comprehensive analysis of crimes against women has revealed critical insights:

* Crimes disproportionately affect young women, especially those aged 18–30.
* Some states consistently report higher volumes of crime, while others have high rates despite lower totals.
* Cruelty by in-laws has a direct emotional and psychological toll, often driving women to suicide.
* Domestic violence and assault are still the most widespread forms of abuse.
* Safer states can only truly be identified when adjusting for population, not just total crime counts.

## ****Recommendations****

* **Policy Reforms:** Strengthen laws protecting women from domestic abuse and dowry harassment.
* **Data Reporting:** Encourage better and more transparent crime reporting systems.
* **Support Systems:** Increase access to helplines, shelters, and mental health resources.
* **Education & Awareness:** Focus on community-level awareness to change cultural attitudes toward women.
* **Law Enforcement Training:** Sensitize police and judicial officers to handle crimes against women with empathy and efficiency.

## ****Future Insights****

As society continues to evolve, the patterns and nature of crimes against women are also changing. Based on the trends observed in this study and the current socio-legal landscape, several **future directions** and **predictive insights** can be outlined to guide further research, policymaking, and social reform:

### 1️⃣ **Rise of Cyber-Crimes Against Women**

With increasing internet penetration and smartphone use, cyber-crimes such as **online stalking, blackmail, image morphing, and cyberbullying** are expected to rise, especially among younger women. Law enforcement agencies and policy frameworks must evolve rapidly to track and prosecute digital abuse and ensure **cyber-safety education** becomes mainstream.

### 2️⃣ **Need for Real-Time Crime Monitoring**

Future safety strategies will benefit greatly from the use of **real-time crime data dashboards**, **predictive policing algorithms**, and **geospatial analysis**. These tools can help identify emerging hotspots, patterns, and time-based risks (e.g., certain hours or festivals when crimes spike).

### 3️⃣ **Predictive Analytics for Prevention**

AI and machine learning can be used to **predict high-risk zones**, vulnerable age groups, or patterns of repeat offenses. Predictive models can help authorities deploy resources proactively and issue early warnings to prevent crimes before they occur.

### 4️⃣ **Integration of Mental Health Data**

Given the correlation between domestic abuse and suicides, future studies should integrate **mental health indicators**, helpline call records, and therapy access data to assess **emotional distress trends among women**, especially in abusive environments.

### 5️⃣ **Policy Evolution: Beyond Legal Frameworks**

The future must move beyond just strengthening laws to ensuring **on-ground implementation**. Legal reforms should be paired with:

* **Community policing**
* **Gender sensitization in schools**
* **Economic empowerment programs for women**
* **Faster judicial processes**, especially for gender-based violence cases

### 6️⃣ **Improved Crime Reporting through Technology**

Future improvements in **anonymous crime reporting apps**, **AI-powered complaint analysis**, and **chatbot-based legal advice** may encourage more women to come forward without fear or stigma. Better reporting will mean better data — and better protection.

### 7️⃣ **Regional Deep-Dive Studies**

This study has shown disparities between states, but future work should include **district-level or urban vs. rural** breakdowns, especially in regions with cultural, social, and economic differences that affect crime dynamics and reporting behavior.

### 8️⃣ **Focus on Rehabilitation and Reintegration**

Future safety programs must prioritize not only protection and legal justice but also **rehabilitation of survivors**, including:

* Psychological care
* Skill-building
* Job placements
* Safe housing

# References

## Dataset Source

* + - Ncrb website
    - https://www.ncrb.gov.in/crime-in-india-additional-table?year=2022&category=District+Wise+Reports

## Libraries & Tools Used

* + - **Pandas Documentation**

<https://pandas.pydata.org/docs/>

## NumPy Documentation

<https://numpy.org/doc/>

## Matplotlib Documentation

<https://matplotlib.org/stable/contents.html>

## Seaborn Documentation

<https://seaborn.pydata.org/>

## Online Resources & Tutorials

Towards Data Science (Medium) – Data Analysis Guides <https://towardsdatascience.com/>

**Social Platforms:**

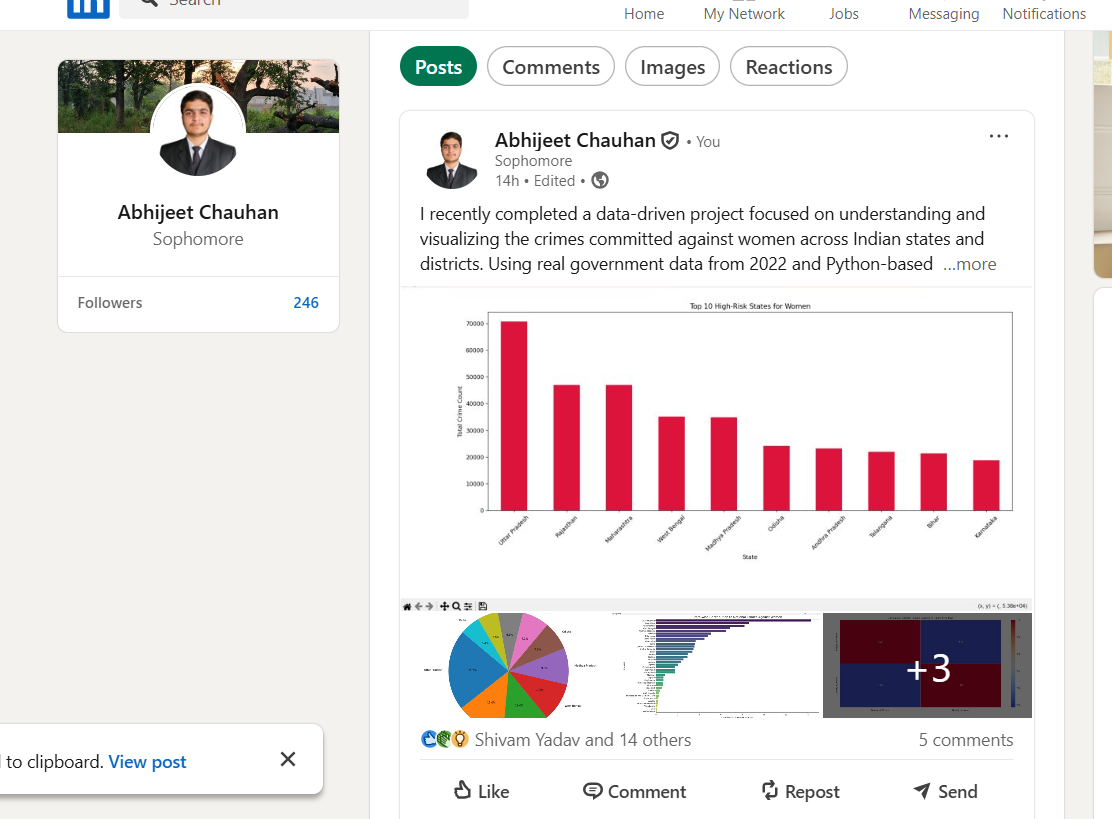
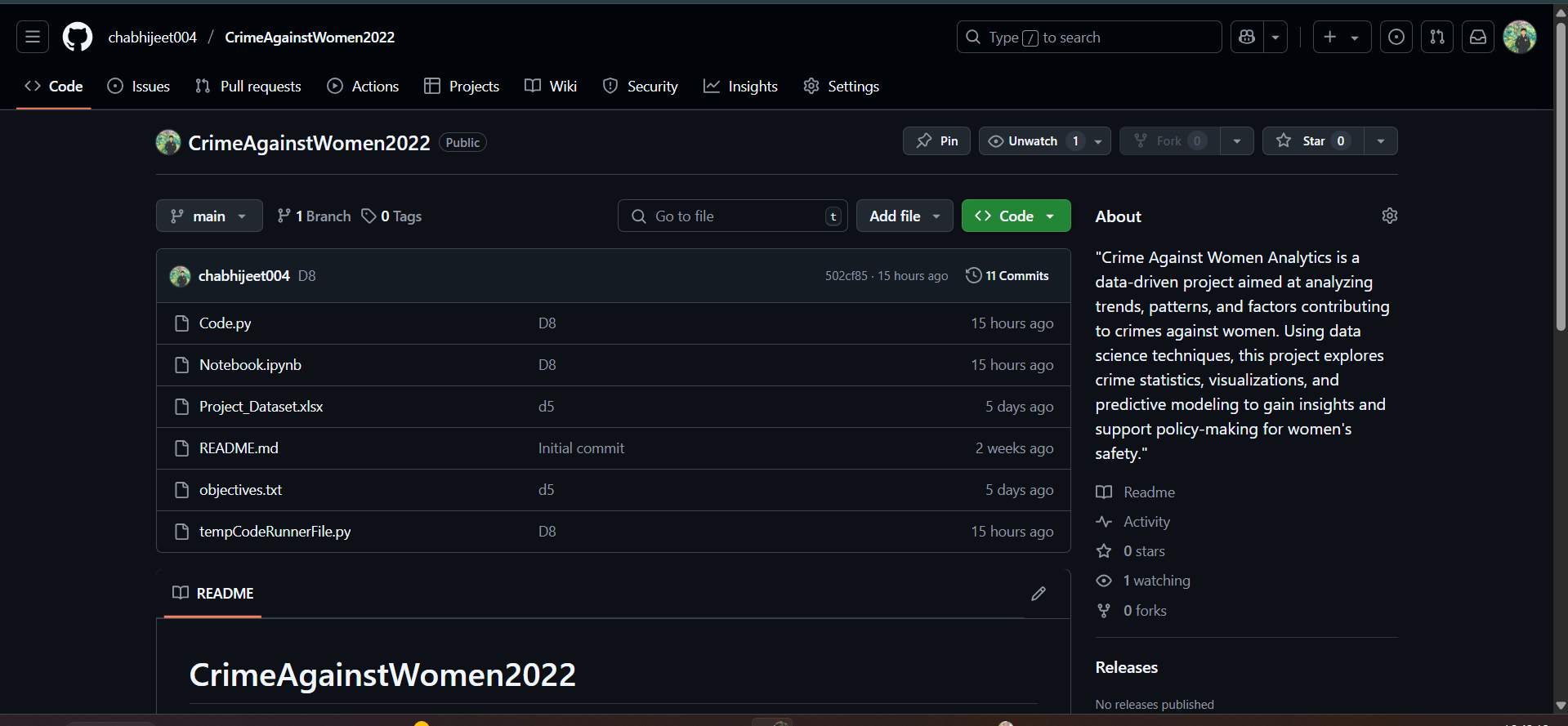
**Linkedin:**

**URL:-**

**https://www.linkedin.com/posts/abhijeet-chauhan-722a32247\_dataanalysis-python-womensafety-activity-7316561415840899072-jfav?utm\_source=share&utm\_medium=member\_desktop&rcm=ACoAAD0y7IoBq44mep9cx5OOdnYU0ccU1vNc4Cc**

# Github :

**URL:** https://github.com/chabhijeet004/CrimeAgainstWomen2022

****