**Project VAWStats: Analysing “Violence Against Women” in India**

*Prepared in the partial fulfilment of the Summer Internship Program on Data analysis*

AT



*Under the guidance of*

***Mrs. Meenakshi, APSSDC***

***Mrs. Narmada Mani, APSSDC***

*Submitted by* ***Vaishnavi Ratnam Movva***

***AP21110010169***

***SRM University AP***

*(August, 2023)*

## ACKNOWLEDGEMENT

I would like to express my heartfelt gratitude to all those who have contributed to the successful completion of my summer internship project at **Andhra Pradesh Skill Development Corporation (APSSDC)**. This opportunity has been an enriching and transformative experience for me, and I am truly thankful for the support, guidance, and encouragement I have received along the way.

First and foremost, I extend my sincere regards to Mrs Meenakshi and Mrs Narmada Mani, my supervisors and mentors, for providing me with valuable insights, constant guidance, and unwavering support throughout the duration of the internship. Their expertise and encouragement have been instrumental in shaping the direction of this project.

I would like to thank the entire team at **Andhra Pradesh Skill Development Corporation (APSSDC)** for fostering a collaborative and innovative environment. The camaraderie, knowledge sharing, and feedback I received from my colleagues significantly contributed to the development and success of this project.

In conclusion, I am honoured to have been a part of this internship program, and I look forward to leveraging the skills and knowledge gained to contribute positively to future endeavours.

Thank you.

Sincerely,

Vaishnavi Ratnam Movva, AP21110010169

vaishnaviratnam\_m@srmap.edu.in

# ABSTRACT

Analysing crime against women in India is of utmost importance due to its profound impact on society and human rights. Through comprehensive data study and rigorous analysis, this research aims to shed light on the alarming trends, identify vulnerable regions, and understand the underlying factors contributing to such crimes. By unravelling patterns and demographics associated with these offenses, this study seeks to inform policymakers, law enforcement, and social activists to develop targeted interventions and policies aimed at combating this pressing issue. With a data-driven approach, the research hopes to pave the way for a safer, more inclusive society, empowering women and ensuring their fundamental right to live free from violence and discrimination.

## TABLE OF CONTENTS

### S.No Content Pg.No

[1. Introduction 5](#_Toc13047)

[2. System Requirements 6](#_Toc13048)

[3. Uses of data analysis library 7](#_Toc13049)

[4. Implementation and architecture 8](#_Toc13050)

[5. Results 13](#_Toc13051)

[6. Advantages 15](#_Toc13052)

7.Conclusion……………………………………………………………………………. 16

8.References……………………………………………………………………………..17

# INTRODUCTION

Crime against women is a deeply concerning and pervasive issue in India that poses significant challenges to societal well-being and gender equality. The plight of women facing violence, harassment, and discrimination has drawn widespread attention and sparked a pressing need for data-driven analysis to comprehend the gravity of the problem. This research embarks on a comprehensive data study to investigate the multifaceted aspects of crime against women in India.

This study endeavours to unravel critical insights into the patterns, trends, and contributing factors of these crimes. By employing advanced statistical techniques and geospatial analysis, we aim to identify crime hotspots and demographic correlations, providing policymakers, law enforcement, and social advocates with the knowledge needed to implement targeted interventions and policy reforms. With a focus on the human rights and dignity of women, this research seeks to make a significant contribution to the ongoing efforts to combat crime against women and foster a safer and more equitable society.

In the context of my project on crime against women in India, data analysis plays a critical role in understanding the magnitude and nature of this pressing social issue.

# 

**SYSTEM REQUIREMENTS**

**Hardware requirements:**

Windows 7 or Greater Version

Stable Internet Connection

A good amount of RAM (8 GB or more)

**Software requirements:**

Python 3.0

Python IDE (Jupyter Notebook)

CRIME DATA SET

Relevant Python Libraries (NumPy, Pandas, Seaborn, Matplotlib, etc.)

**USES OF DATA ANALYSIS LIBRARY**

Python has become one of the most popular programming languages for data science due to its versatility, ease of use, and rich ecosystem of libraries specifically designed for data analysis, manipulation, visualization, and machine learning.

**Pandas**: Pandas is a powerful library for data manipulation and analysis. It provides data structures like Series and Data Frame, allowing easy handling of tabular data .

**NumPy**: A fundamental library for numerical computing in Python, enabling high-performance operations on arrays and matrices.

**Pandas:** A versatile data manipulation library that offers powerful data structures and tools for data analysis and cleaning in Python.

**Seaborn:** A Python data visualization library that facilitates creating informative and visually appealing statistical graphics.

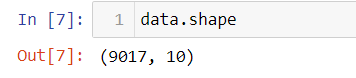
**Matplotlib:** Matplotlib is a popular plotting library in Python. It provides a wide variety of 2D and limited 3D plotting options and is highly customizable for creating publication-quality plots.

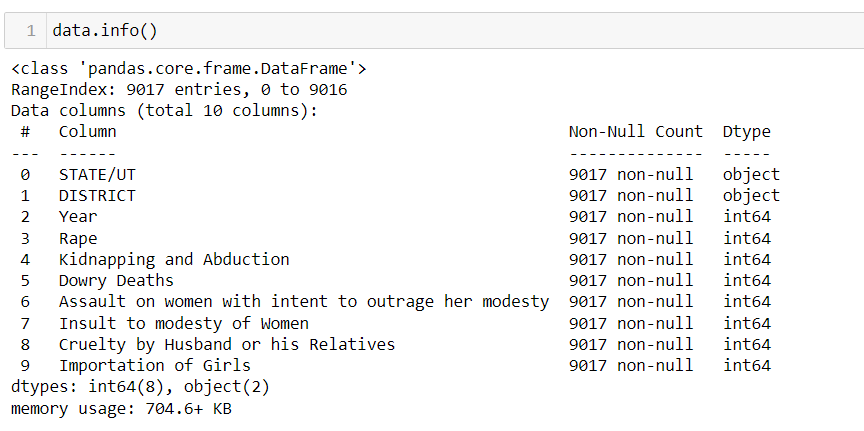
**ARCHITECHTURE**

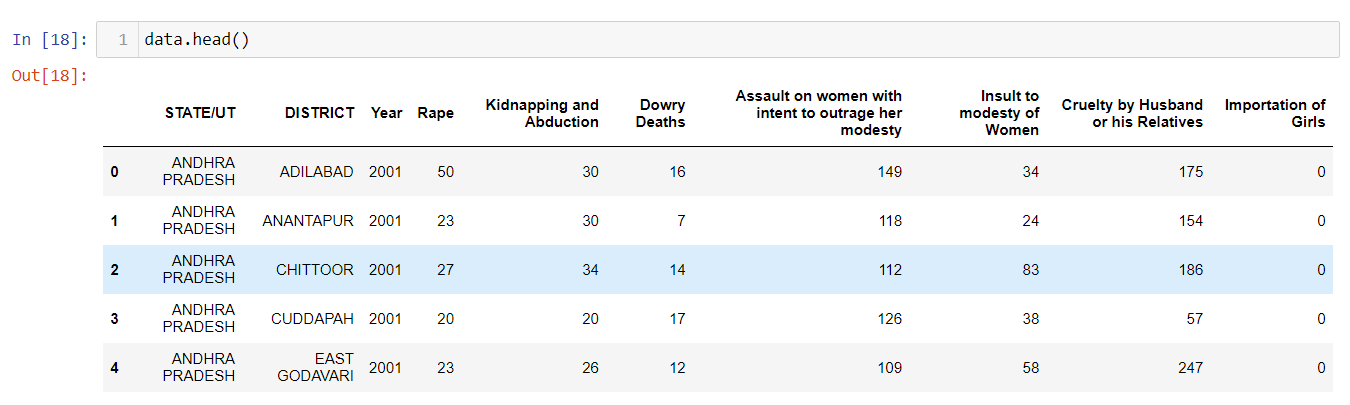
**Reading the Data Set:**

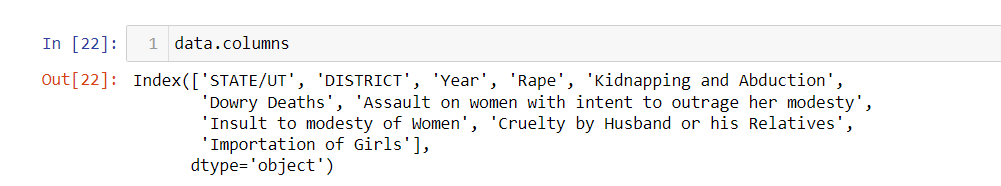
****

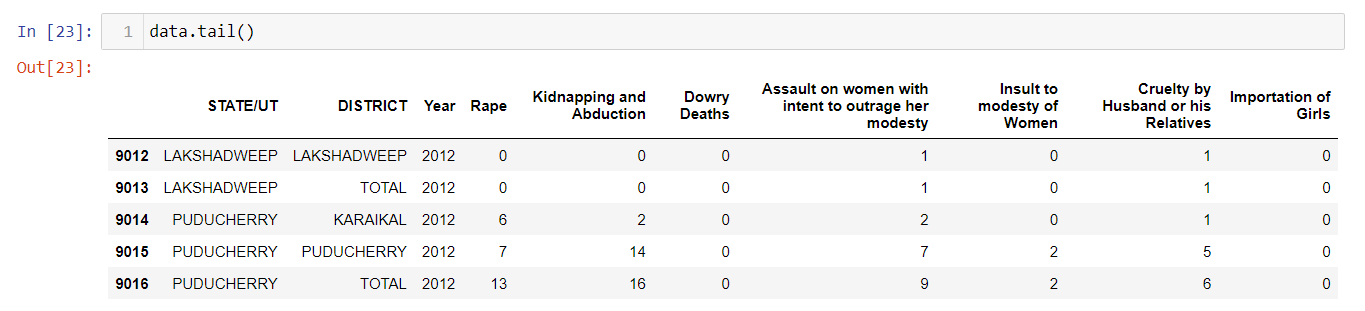
**Applying basic functions:**

****

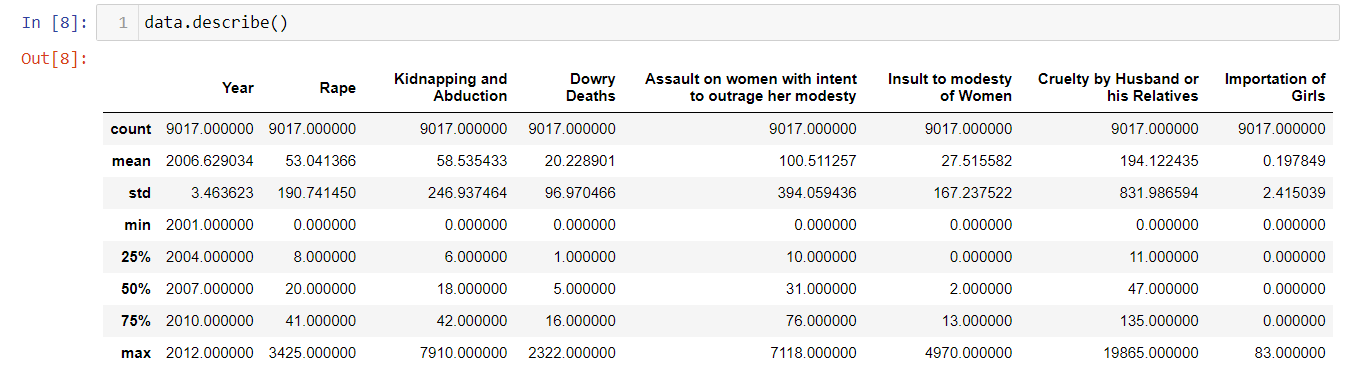
****

****

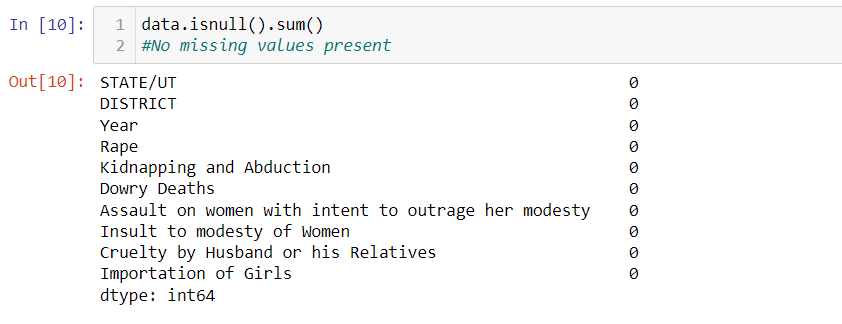
****

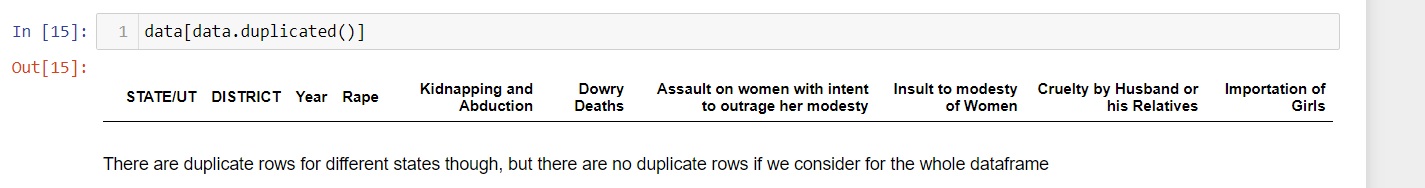
****

**Summary statistics using describe ():**

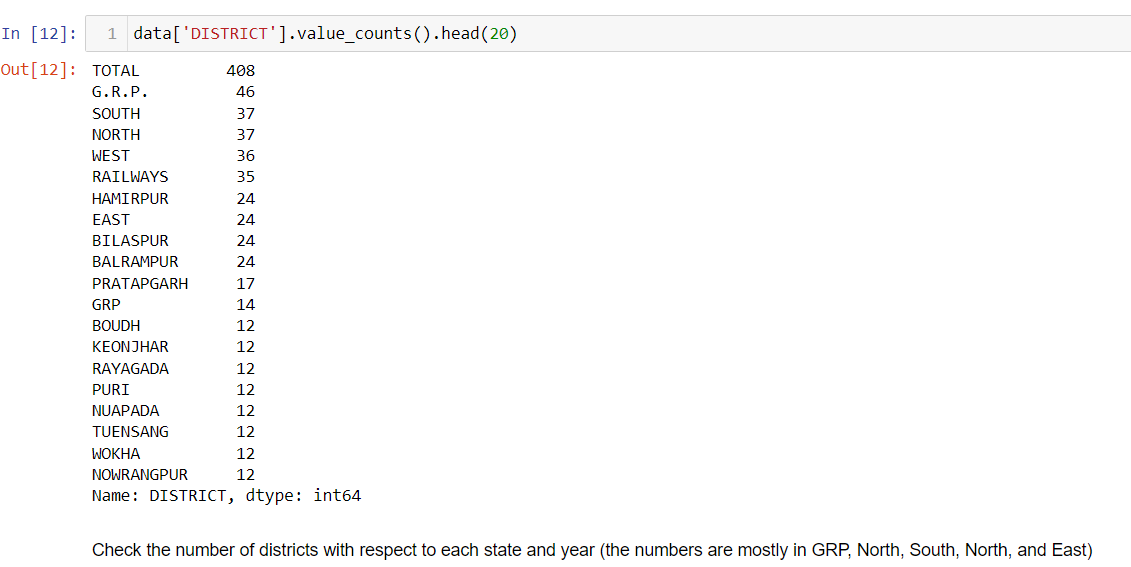
****

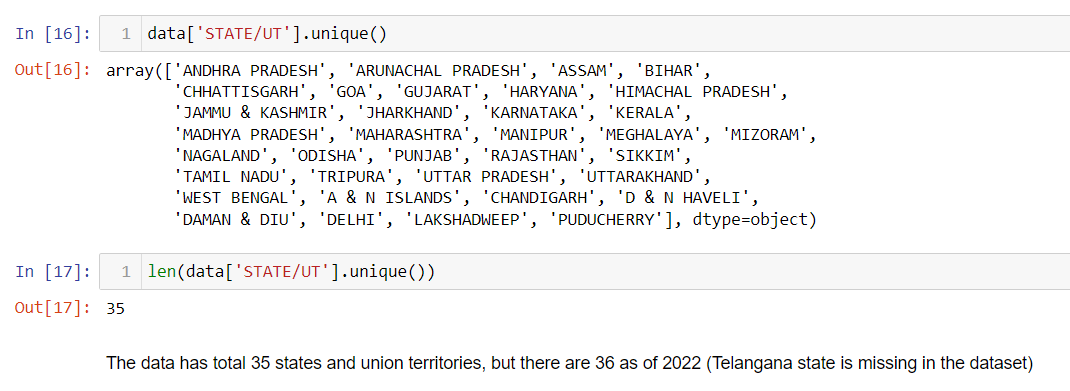
**DATA CLEANING:**

****

****

**RANDOM FINDINGS:**

****

****

**DATA VISUALISATION:**

The heatmap is created using seaborn to observe the correlation between the variables. The created correlation matrix shows the single correlation between each feature with other features on the dataset.

If 2 variables are correlated to each other by:

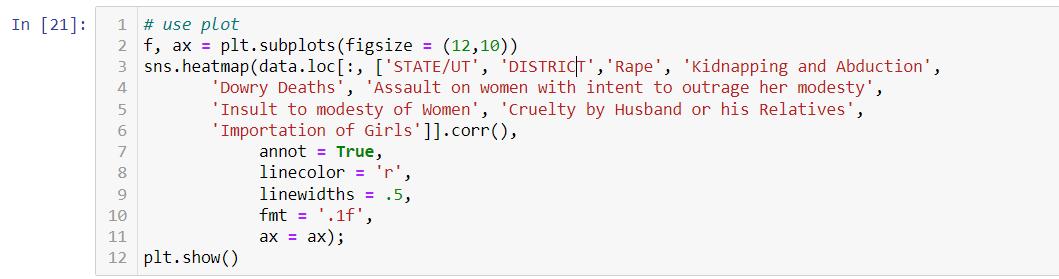
0.0 to 0.3, they are weakly correlated,

0.3 to 0.6, they are moderately correlated,

0.6 to 0.9, they are strongly correlated,

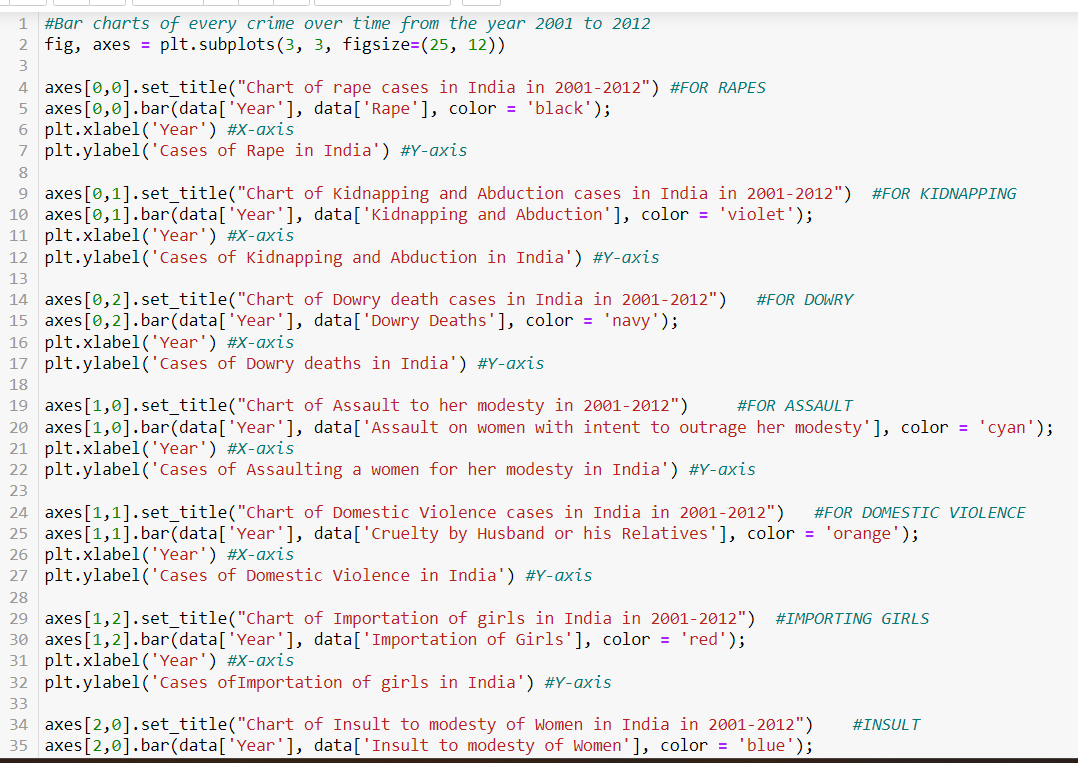
finally, >0.9, they are very strongly correlated

Positive and negative indicates whether the variables are directly or inversely related e.g. a correlation of -0.7 between 2 variables denote that if one variable increases, the other decreases strongly (as defined in the list above)



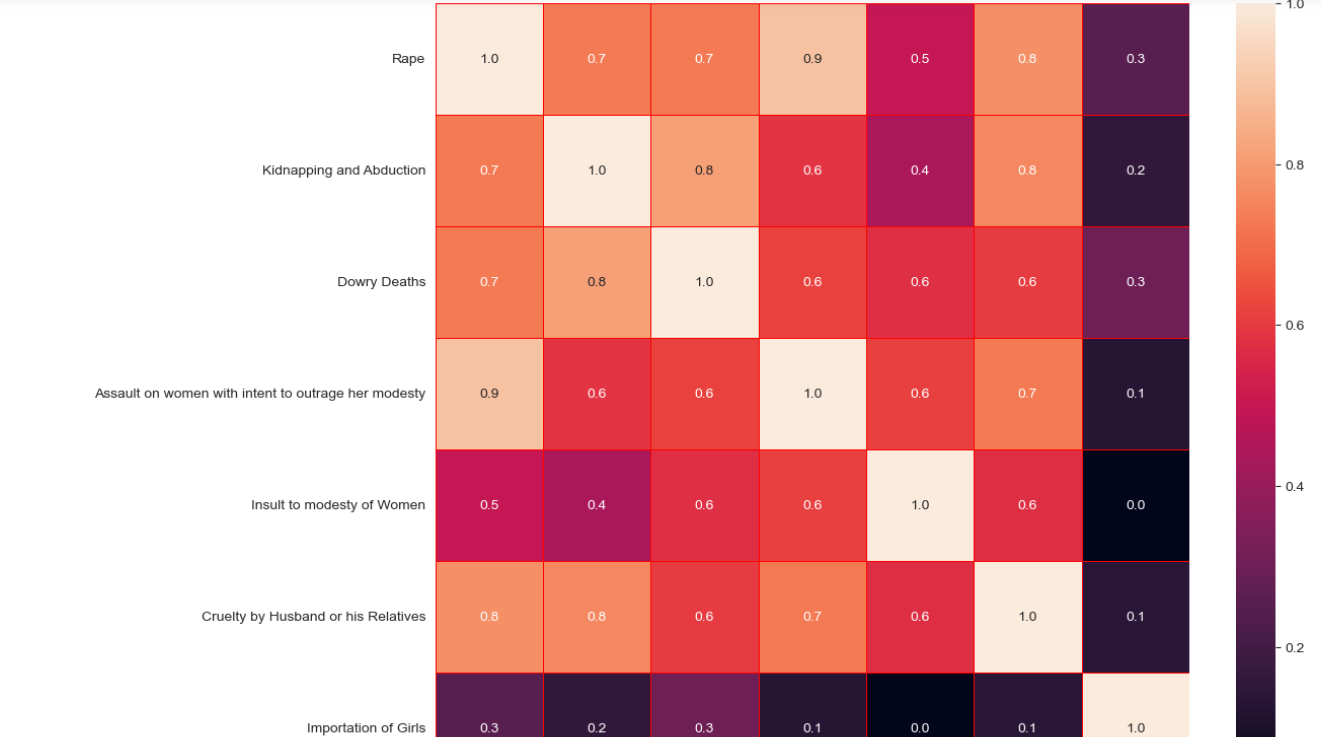
Similarly, I have also constructed bar graphs and outlier detection using boxplot method. See the results section below for more details.

Here is the code for bar graphs representation

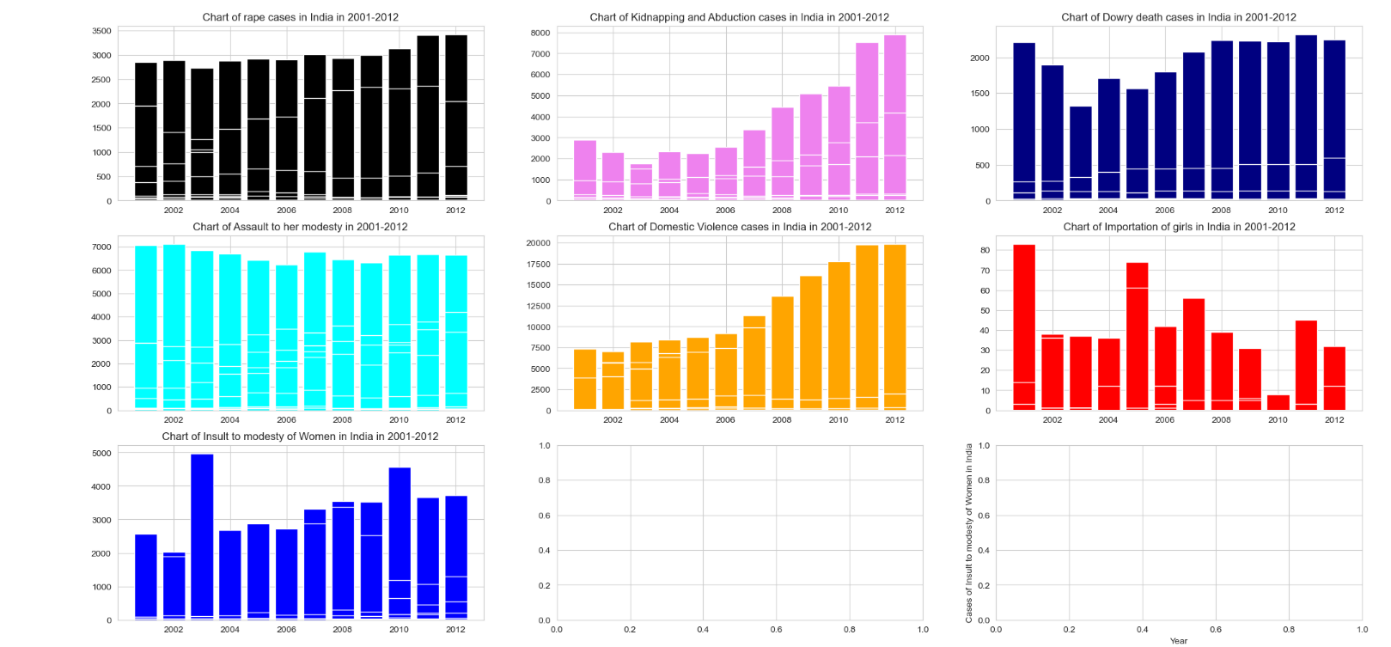


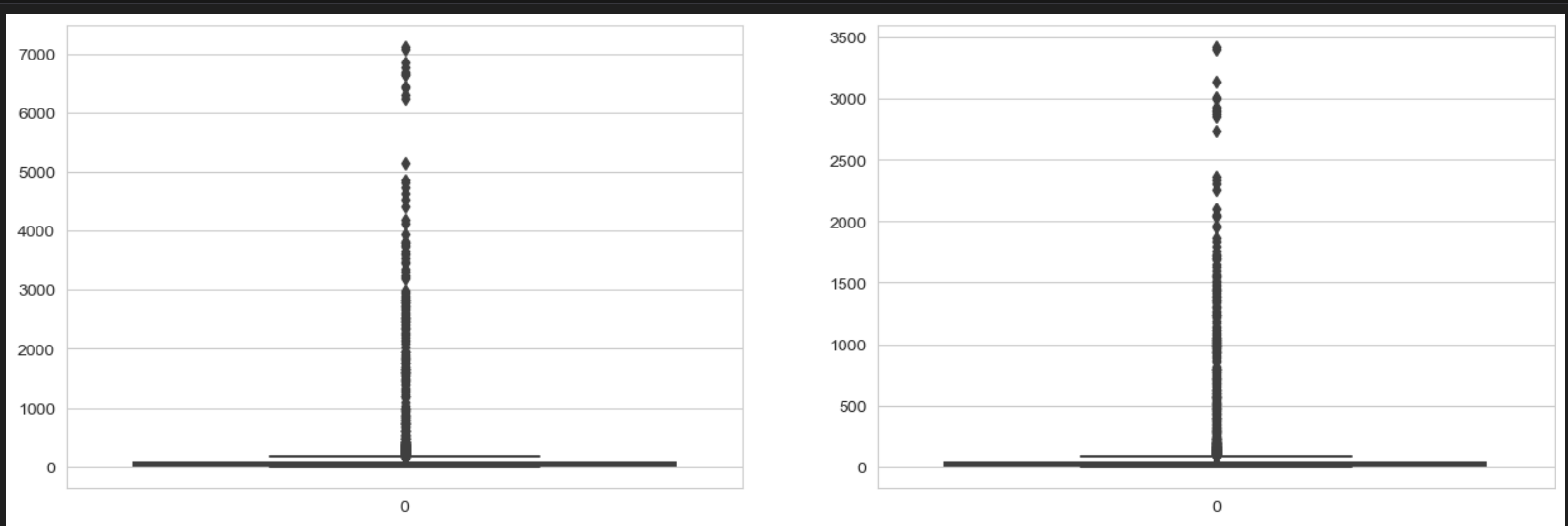
# RESULTS

The heat map for the taken data set is as follows

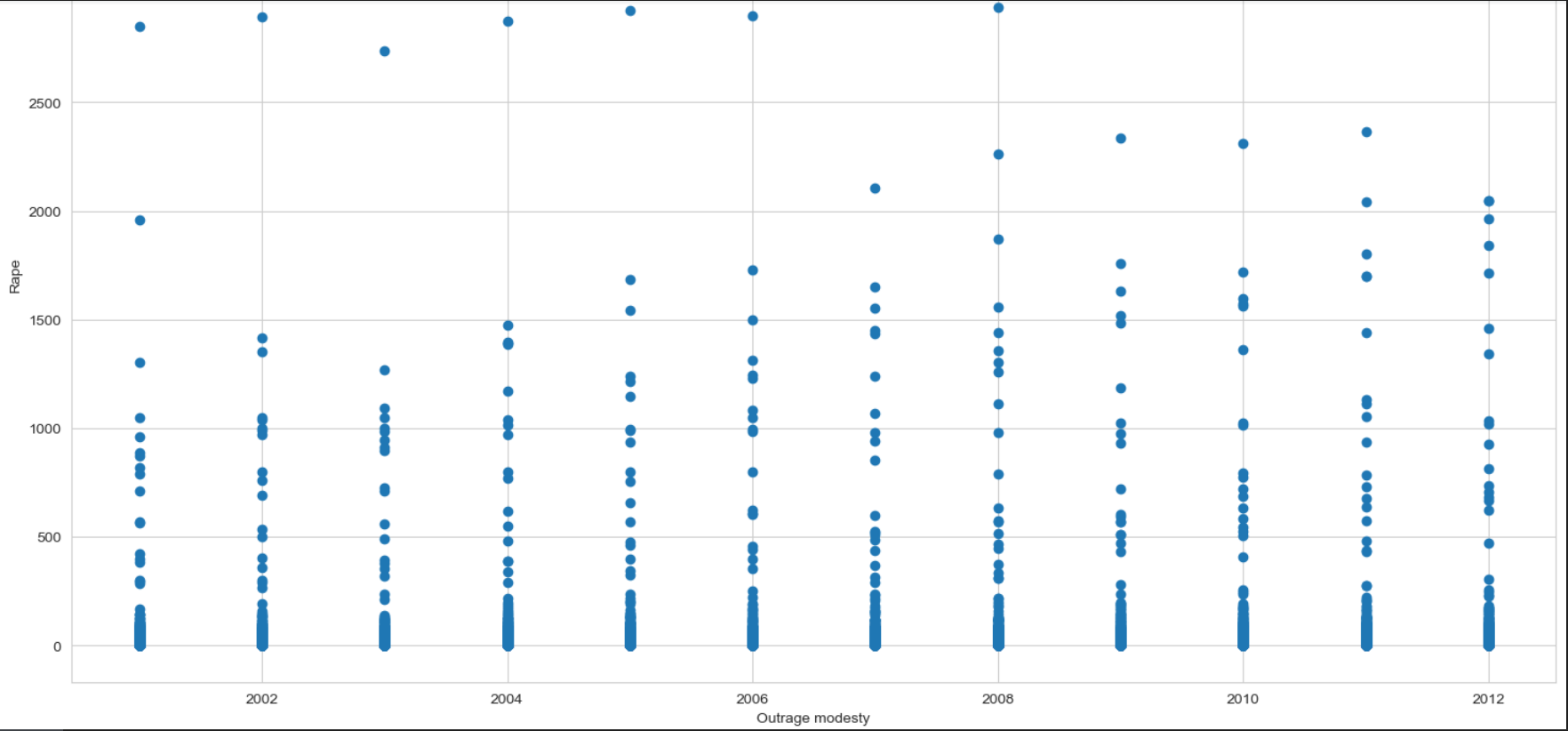


The bar graph representation of every crime is as follows





Scatter plot :



# ADVANTAGES

Analysing crime rates in India through data analysis offers several advantages that can contribute to a better understanding of the issue and aid in devising effective strategies to address it. Some of the advantages of your project on data analysis of crime rates in India include:

**Evidence-Based Insights:** Data analysis provides evidence-based insights into crime trends, patterns, and demographics. This helps in understanding the scale and nature of crime against women in India, enabling policymakers and law enforcement agencies to make informed decisions.

**Identifying Hotspots:** By analysing crime data spatially, you can identify crime hotspots and areas with higher crime rates. This information allows for targeted resource allocation and focused interventions to address crime in the most affected regions.

**Understanding Crime Types:** Data analysis helps in categorizing different types of crimes against women, such as domestic violence, rape, dowry-related offenses, and sexual harassment. This understanding allows for tailored strategies to combat each type effectively.

**Evaluating Impact of Policies:** My project can assess the impact of existing policies and interventions on crime rates. This evaluation is crucial in refining and implementing more effective measures to address the issue.

**Raising Awareness:** Data analysis and visualization can present a clear picture of the severity of crime against women, thus raising awareness among the public, media, and policymakers. This can lead to greater attention and advocacy for tackling the issue.

**Cross-Comparisons:** Data analysis allows for cross-comparisons with other countries or regions, providing valuable insights into India's standing in addressing crime against women on a global scale.

**CONCLUSION**

In conclusion, the data analysis of crime rates in India, specifically focusing on crime against women, has offered valuable insights and evidence to address this critical societal issue. Through comprehensive data study and rigorous analysis, we have gained a deeper understanding of the nature, patterns, and geographic distribution of these crimes. The identification of crime hotspots and vulnerable demographic groups has provided a foundation for targeted interventions and resource allocation.

The findings of this project have served as an evidence-based guide for policymakers and law enforcement agencies in formulating strategies and policies aimed at combating crime against women effectively. By evaluating the impact of existing initiatives, we have identified areas for improvement and proposed evidence-driven approaches to enhance women's safety and rights.

Moreover, the project's data visualizations and reports have played a pivotal role in raising public awareness and generating advocacy for the cause. Through meticulous data cleaning, preprocessing, and exploratory data analysis (EDA), we gained valuable insights into the distribution of crime types, regional variations, and potential gender-based disparities

**REFERENCES**

1) <https://www.geeksforgeeks.org/>

2)<https://www.tutorialspoint.com/python/index.htm>

3) <https://www.kaggle.com/>