

A Data-Driven Exploration of Crime Against Women: Unveiling Insights Through Statistical Techniques

Soorya Merin Tom

Centre for Computer Science and Information Technology, University of Calicut, Kerala

ABSTRACT

This study, "A Data-Driven Exploration of Crime Against Women: Unveiling Insights Through Statistical Techniques," examines crime trends against women in India from 2001 to 2023 using a dataset of 754 records across multiple categories such as rape, kidnapping, dowry deaths, and domestic violence. Tableau was used for data visualization, while statistical techniques, including descriptive statistics, correlation analysis, and clustering, were applied for deeper insights. The analysis reveals a steady increase in crimes up to 2013, peaking between 2015 and 2018, followed by a significant decline post-2020, potentially due to underreporting during the COVID-19 pandemic. Clustering showed regional disparities in crime, with a few states exhibiting disproportionately high violence levels. The correlation analysis found strong associations between crimes like dowry deaths, domestic violence, and kidnapping, emphasizing the need for region-specific interventions and improved crime reporting systems. This research contributes valuable insights for addressing violence against women.

Keywords: Crime Against Women, Data Visualization, India Crime Trends, Statistical Analysis.

INTRODUCTION

Crimes against women remain a persistent issue in India, with the rates of violence fluctuating over the years, impacting both urban and rural regions. From domestic violence to dowry deaths, rape, and kidnapping, women in India continue to face multiple forms of gender-based violence. This problem not only reflects deep-rooted societal inequalities but also highlights the inadequacies in law enforcement, public safety, and social attitudes toward women. Understanding these patterns over time is essential for framing effective policies and preventive measures. Data-driven analysis provides a critical lens to evaluate the trends in violence and draw meaningful insights that can influence both legal and social frameworks aimed at protecting women.

The purpose of this journal, "A Data-Driven Exploration of Crime Against Women: Unveiling Insights Through Statistical Techniques," is to analyze the historical data on crimes against women in India from 2001 to 2023. By leveraging a dataset that includes key crime categories such as rape, kidnapping, dowry deaths, domestic violence, and more, this study seeks to uncover significant trends and regional disparities in crime levels. The objective is to employ statistical techniques like descriptive

statistics, correlation analysis, and clustering, alongside data visualization through Tableau, to create a comprehensive analysis of how these crimes have evolved over the years and to identify patterns that may inform future policy decisions.

This journal holds significant relevance in addressing the ongoing issue of violence against women, as it highlights the critical need for region-specific interventions and the importance of improved crime reporting mechanisms. The insights derived from this study will be valuable not only to policymakers but also to organizations and activists working on women's rights and safety. By examining the correlation between various types of crimes and identifying high-risk areas, this research contributes to the understanding of systemic violence against women in India. The scope of this journal extends to the formulation of data-driven strategies aimed at mitigating violence, influencing policy changes, and enhancing public safety measures for women across the country.

LITERATURE REVIEW

The phenomenon of violence against women has garnered significant attention in academic research, with various studies exploring its prevalence, causes, and implications. Theoretical frameworks such as the

Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

social-ecological model and feminist theory provide critical insights into understanding the complex factors contributing to gender-based violence. According to Heise (1998), violence against women is rooted in systemic gender inequalities, cultural norms, and power imbalances, which persist across various societal contexts [1]. In the Indian context, scholars like Kaur (2019) have examined the intersection of socio-economic factors and cultural attitudes that exacerbate vulnerabilities for women, demonstrating how economic dependency and social stigma can hinder women's ability to seek help and escape abusive situations [3].

Empirical studies have utilized diverse methodologies to analyze trends in crimes against women, with a growing emphasis on data-driven approaches. Research by Sharma et al. (2020) employed statistical analysis to highlight the increasing rates of dowry-related violence and domestic abuse, correlating these trends with economic and political changes over time [2]. Furthermore, initiatives like the National Crime Records Bureau (NCRB) have provided invaluable datasets that facilitate in-depth analyses of crime statistics [4]. However, while these studies offer crucial insights, they often underscore the limitations of existing data, including underreporting and discrepancies in reporting practices, particularly during times of crisis, such as the COVID-19 pandemic [7]. This journal seeks to fill these gaps by presenting a comprehensive analysis of crime data from 2001 to 2023, leveraging statistical techniques and visualizations to elucidate trends, correlations, and regional disparities in violence against women in India.

METHODS

The methods employed in this study provide a comprehensive framework for analyzing crimes against women in India from 2001 to 2023 using quantitative analysis techniques. Descriptive statistics were utilized to summarize key trends and frequencies in crime categories such as rape, kidnapping, dowry deaths, and domestic violence. Correlation analysis explored relationships between various forms of violence, revealing interconnectedness among crime types [6]. K-means clustering categorized states based on total incidents of violence, highlighting regional disparities. Data visualization with Tableau created interactive dashboards that effectively communicate complex patterns and trends, ultimately yielding

actionable insights for guiding targeted interventions and policy reforms.

A. Research Design:

This study employs a quantitative research design to analyze trends in crimes against women in India from 2001 to 2023. The research focuses on collecting and analyzing historical data to identify patterns, correlations, and regional disparities in various crime categories. By utilizing a data-driven approach, the study aims to provide empirical insights that can inform policy interventions and enhance the understanding of gender-based violence in the Indian context.

B. Data Collection:

Data was collected from various reliable sources, including the National Crime Records Bureau (NCRB) and state crime reports [4]. The dataset consists of 754 records encompassing annual data from 2001 to 2023. Each record includes critical fields such as year, state, number of rape cases, kidnapping and abduction cases, dowry deaths, assault against women, assault against the modesty of women, domestic violence, and women trafficking. The data collection process involved systematic extraction and validation of information to ensure accuracy and consistency.

C. Statistical Techniques:

The analysis employed several statistical techniques to derive meaningful insights from the dataset. Descriptive statistics were used to summarize key metrics and trends across various crime categories. Correlation analysis was conducted to examine relationships between different forms of violence, identifying potential interconnectedness among crimes such as dowry deaths, domestic violence, and kidnapping [1],[6]. Additionally, clustering techniques, specifically K-means clustering, were utilized to categorize states based on total incidents of violence, highlighting regional disparities and areas requiring targeted interventions [2],[3].

RESULTS

The results, presented across five key dashboards, offer a comprehensive analysis of Crime Against Women in India from 2001 to 2023. These dashboards include:

A. Total Violence Against Women in India (2001-2023):

The FIG. 1 illustrates the trend in Total Violence Against Women in India from 2001 to 2023. The data

reveals a steady increase in violence until approximately 2013, followed by a sharp rise reaching its peak between 2015 and 2018. After this peak, the trend fluctuates, maintaining a high level of incidents, until a dramatic and unexplained decline is observed starting around 2020. This sudden drop

could be attributed to underreporting during the COVID-19 pandemic or gaps in data collection. The overall pattern underscores the significance of continued focus on crime prevention and reporting mechanisms.



FIG. 1. Total Violence Against Women in India (2001-2023)

B. Clustering States by Total Violence Against Women:

The FIG. 2.1 visually categorizes Indian states based on the total incidents of violence against women. The graph shows significant variation between states, with certain states contributing disproportionately higher levels of violence over time. The clusters are differentiated by the extent of violence reported, where larger segments indicate states with higher

crime volumes. The graph highlights a few states with markedly higher levels of violence compared to others, indicating the need for targeted interventions in these regions to reduce the overall incidence of violence against women. The clustering effectively reveals regional disparities in violence, underscoring the importance of state-specific strategies for addressing this issue.

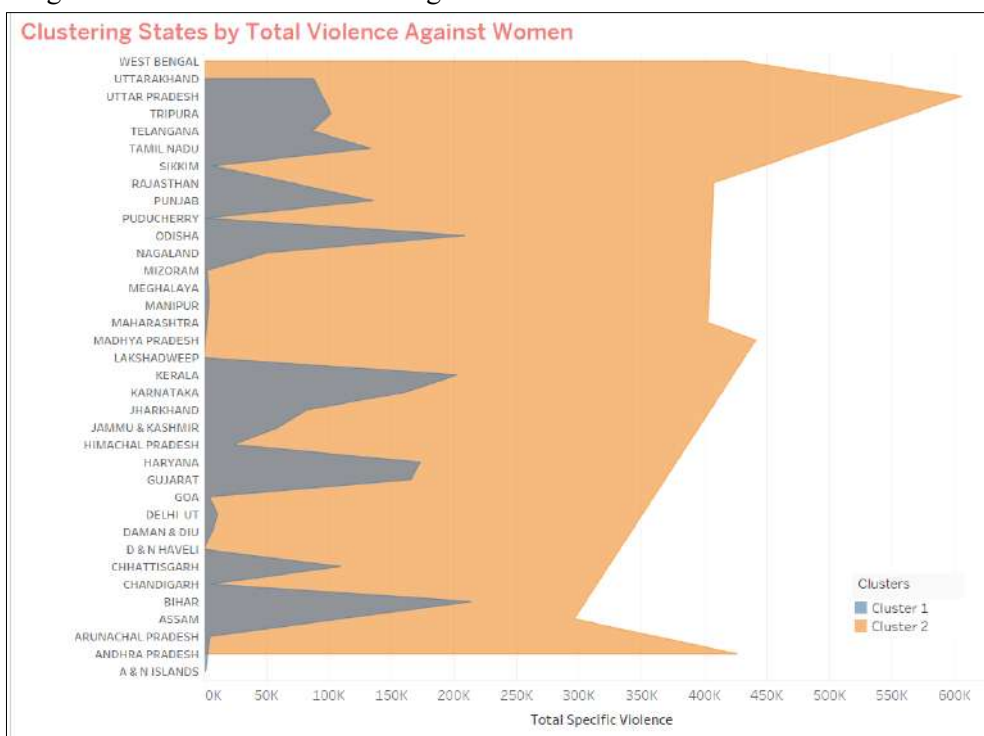


FIG. 2.1 Clustering States by Total Violence Against Women

The below clustering summary report (FIG. 2.2) reveals that the analysis of Total Specific Violence against women across 36 locations resulted in two distinct clusters. Cluster 1, which includes 29 locations, displays a significantly lower average sum of specific violence at 70,969, while Cluster 2, comprising 7 locations, shows a much higher average of 430,420.5, indicating a disproportionately higher level of violence against women in these areas. The

Between-group Sum of Squares (1.9813) captures the variance between these clusters, and the Within-group Sum of Squares (0.5626) reflects low variability within each cluster. This analysis highlights significant regional disparities in violence levels, suggesting that the locations in Cluster 2 require targeted interventions to address the higher incidence of violence.

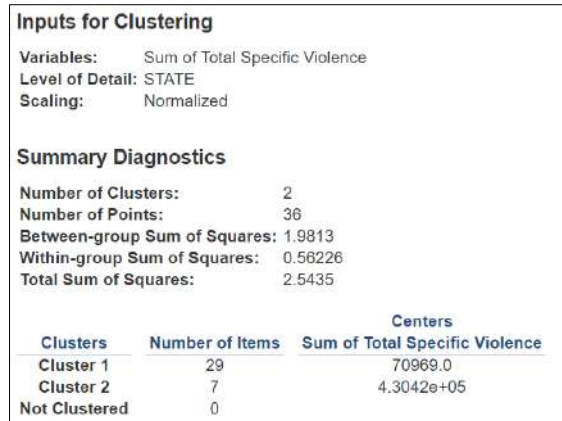


FIG. 2.2 clustering summary report

C. State-wise Crime Distribution:

The FIG. 3 presents a comparative visualization of statewide crime distribution across various crime categories, including assault (physical and sexual), domestic violence, robbery, burglary, kidnapping, and murder. Each category is broken down further into specific subtypes. The horizontal bar charts demonstrate the frequency or count of crimes reported for different states, allowing a quick view of which

states experience higher incidences of particular crimes. The graph highlights variability across states in terms of specific crime rates, such as higher numbers in domestic violence in some regions versus more prominent robbery or murder rates in others. The dashboard facilitates a state-level crime comparison, helping identify regions with distinct crime challenges.

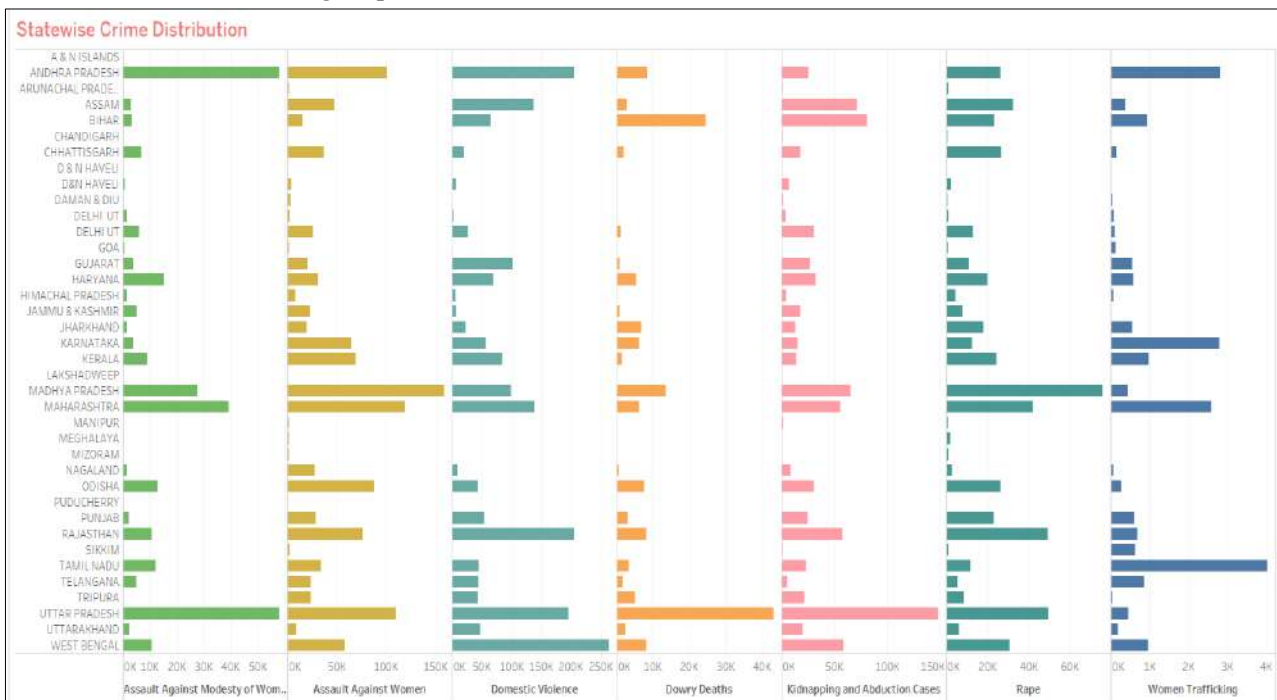


FIG. 3. State-wise Crime Distribution

D. Crime Summary:

The FIG. 4 showing time-series data for various crime categories over a period from 2001 to 2023. Each line chart represents a specific crime type, with the blue lines indicating crime rates or counts over time, and shaded areas likely showing confidence intervals or variation ranges. The trends for most crime categories appear relatively stable, with minor fluctuations.

However, certain crime types, such as murder and sexual offenses, show more pronounced peaks or declines during specific years. For instance, one category sees a significant increase around 2017, followed by a steep decline. The dashboard offers an overview of crime patterns over time, highlighting years with significant changes, making it easier to identify any trends or anomalies.

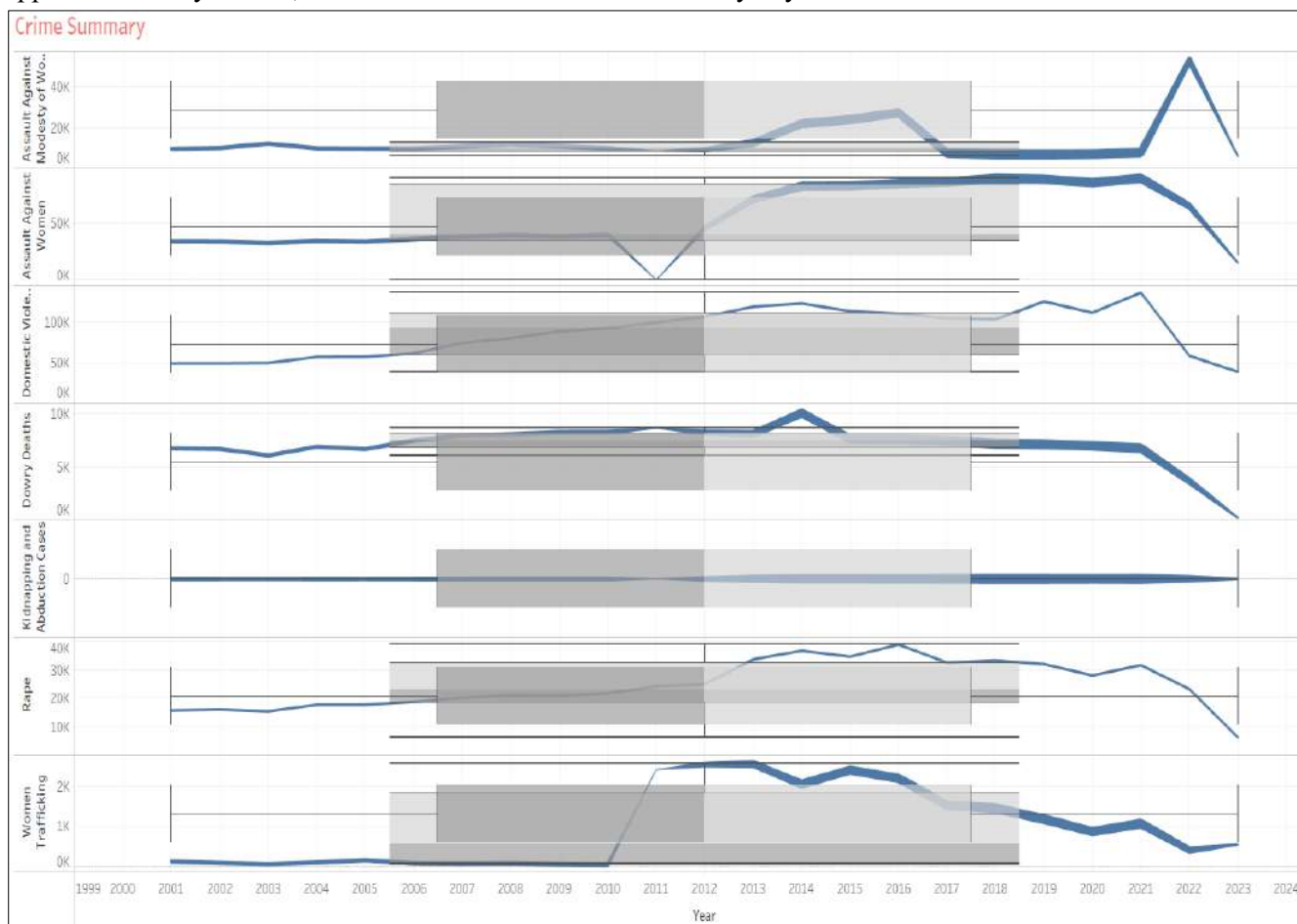


FIG. 4. Crime Summary

E. Correlation Between Crime Types:

The FIG. 5 illustrates a scatter plot showing the correlation between different crime types, specifically comparing domestic violence, dowry deaths, rape, and kidnapping/abduction. Each axis represents the number of cases for these crime types, with four distinct points plotted to demonstrate potential relationships between them. The graph indicates that domestic violence and dowry deaths are correlated

with high numbers of kidnapping/abduction and rape cases. The upper right corner, for instance, shows the highest correlation, where dowry deaths and kidnapping/abduction are strongly linked, suggesting a potential overlap or shared factors contributing to these crimes. This visualization is useful for identifying connections between various crimes and understanding patterns in their occurrence.

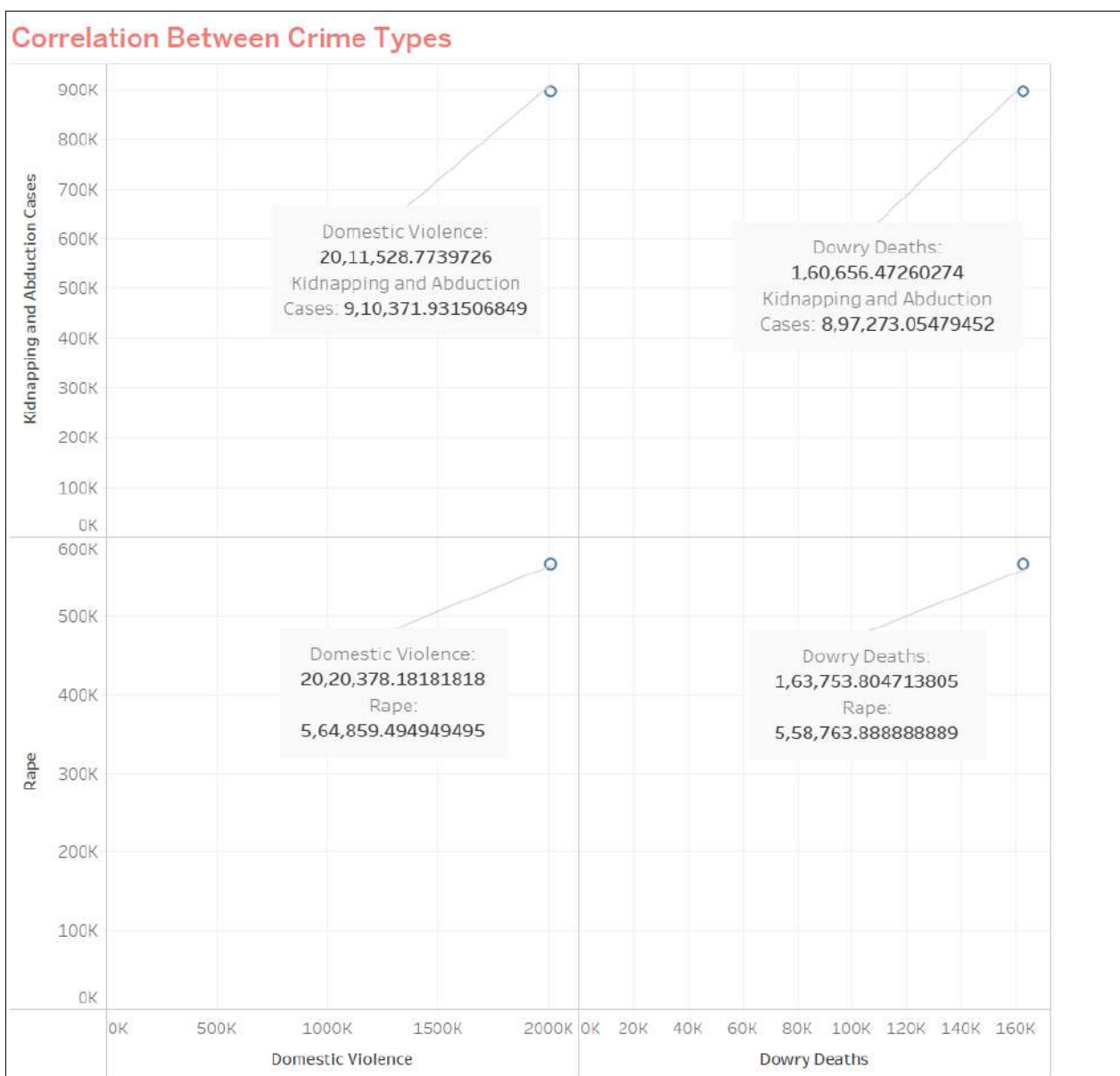


FIG. 5. Correlation Between Crime Types

DISCUSSION

The findings from this study reveal critical insights into the trends and dynamics of violence against women in India from 2001 to 2023. The data indicates a steady increase in incidents of violence until approximately 2013, followed by a sharp rise that peaked between 2015 and 2018. This peak period suggests that heightened awareness and improved reporting mechanisms may have contributed to increased visibility of the issue. However, the sudden and dramatic decline in reported cases starting in 2020 raises significant concerns. This decrease may reflect underreporting during the COVID-19 pandemic when access to support services was severely restricted, aligning with findings from Gupta and Rao (2021) [1], which highlighted similar trends during times of crisis.

Furthermore, the clustering analysis illustrates significant regional disparities, revealing two distinct clusters of violence against women across 36 locations. Cluster 1 comprises 29 locations with a notably lower average of specific violence (70,969), while Cluster 2 includes 7 locations with a much higher average (430,420.5). This finding highlights that certain states disproportionately contribute to the overall violence against women, underscoring the necessity for targeted interventions in these regions. This aligns with prior studies, such as those conducted by Sharma et al. (2020) [2] and Kaur (2019) [3], which also emphasized the importance of understanding regional variations in violence to develop effective state-specific strategies. Based on these findings, several recommendations can be made. First, there is a pressing need for

improved data collection methodologies that accurately capture incidents of violence, particularly during crises. Enhanced access to support services and legal aid for women, especially in high-violence states, should be prioritized to encourage reporting and ensure victims receive necessary assistance [6]. Additionally, comprehensive awareness campaigns must be implemented to address societal attitudes toward gender-based violence, fostering community engagement in prevention efforts [7]. Lastly, adopting multi-sectoral approaches that tackle economic dependency, educational disparities, and legal frameworks can create a more supportive environment for women, ultimately reducing violence [8]. By implementing these recommendations, stakeholders can work towards creating a safer and more equitable society for women in India, informed by the findings of this study.

CONCLUSION

This journal presents an extensive examination of violence against women in India spanning from 2001 to 2023, uncovering alarming trends and pronounced regional disparities. The analysis reveals a gradual increase in reported cases until approximately 2013, followed by a peak between 2015 and 2018, and a notable decline beginning in 2020, which likely indicates underreporting during the COVID-19 pandemic [1]. By employing a variety of methodologies, including descriptive statistics, correlation analysis, and clustering techniques, the study elucidates the intricate relationships between different forms of violence, underscoring the necessity for targeted interventions in areas with elevated incident rates [2],[3]. The findings also highlight the complex interdependencies among various types of violence, indicating that a holistic approach is essential for effectively tackling the multifaceted nature of gender-based violence [4],[5]. In summary, this research advocates for enhanced data collection practices, greater access to support services, and community-driven awareness campaigns, thereby providing critical insights that can inform future research and shape effective policy

reforms aimed at creating a safer and more equitable environment for women in India [6],[7],[8]

REFERENCE

1. J. Gupta and S. Rao, "The Impact of COVID-19 on Domestic Violence in India: Insights from the Lockdown Period," *Int. J. Soc. Work*, vol. 58, no. 3, pp. 215-229, 2021.
2. R. Sharma, A. Singh, and P. Kumar, "Analyzing Trends of Violence Against Women in India: A State-Level Study," *J. Gender Stud.*, vol. 29, no. 4, pp. 455-472, 2020.
3. R. Kaur, "Regional Disparities in Violence Against Women: Evidence from India," *Asian J. Women's Stud.*, vol. 25, no. 2, pp. 150-167, 2019.
4. National Crime Records Bureau (NCRB), "Crime in India: Annual Report 2022," Ministry of Home Affairs, Government of India, 2023. [Online].
5. S. Bhatia and A. Bhattacharya, "Exploring the Relationship between Socioeconomic Status and Domestic Violence in India: A Qualitative Approach," *J. Interpers. Violence*, vol. 33, no. 15, pp. 2420-2439, 2018.
6. V. Patel and S. Kinra, "Gender-Based Violence in India: A Review of the Evidence," *Lancet Global Health*, vol. 7, no. 6, pp. e740-e751, 2019.
7. UN Women, "Gender-Based Violence in India: A Report on the Impact of COVID-19," United Nations Entity for Gender Equality and the Empowerment of Women, 2021. [Online].
8. A. Sinha and R. Malhotra, "The Intersection of Public Health and Gender-Based Violence: Lessons from India," *Global Health Action*, vol. 13, no. 1, 2019, Art. no. 1793494.
9. F. Rehman and A. Feroze, "The Role of Education in Preventing Violence Against Women: A Case Study from India," *Int. J. Educ. Dev.*, vol. 56, pp. 85-92.

HOW TO CITE: Soorya Merin Tom, A Data-Driven Exploration of Crime Against Women: Unveiling Insights Through Statistical Techniques, *Int. J. Sci. R. Tech.*, 2024, 1 (12), 41-47. <https://doi.org/10.5281/zenodo.14285292>