Munich Personal RePEc Archive

Crime against Women in India (2001-2023): A Spatial Analysis Using K-Means Clustering Approach

Dutta, Chandrani and Pramanick, Amrita

Assam University

26 February 2026

Online at https://mpra.ub.uni-muenchen.de/128208/
MPRA Paper No. 128208, posted 17 Jul 2026 01:36 UTC

Crime against Women in India (2001-2023): A Spatial Analysis Using K-Means Clustering Approach

Chandrani Dutta[1]

Amrita Pramanick[2]

Abstract

According to the recent data of National Crime Records Bureau (NCRB), 51 cases of crime against women are reported every hour in India. The existing literature and recent studies on gender-based violence have shown that over the years Indian states have followed a pattern of clusters in criminal incidents. This paper studies the panel data from NCRB for twenty-three years (2001-2023) to study and identify specific features and spatial distribution of crime against women across the Indian states. The analysis is conducted on the basis of the top five reported crimes identified as Cruelty against Women by Husband and Relatives, Assault on Women, Kidnapping and Abduction, Rape and Insult to the Modesty of Women. The states with persistently high crime under each crime heads have been identified. Further, Kernel Density Estimation is run in R software which reveals the existence of clustering patterns emerging from the data. The paper further uses the K-Means Clustering method to determine how crime is clustered amongst the Indian states. The optimal value of K from the analysis has been identified as 4 – i.e., four distinct spatial clusters viz. Low Crime, Low to Moderate Crime, Moderate Crime and High Crime have been identified among the Indian states. Uttar Pradesh, Madhya Pradesh, Rajasthan, West Bengal and Maharashtra exhibit high crime rates. Further, all north-eastern states are in Cluster 1 showing persistent low crime rates except for Assam which shows up in Cluster 2. The majority of southern states like Tamil Nadu, Kerala, Karnataka, Telangana shows up in Cluster 2 with an exception in Andhra Pradesh being the only state showing up in Cluster 3.

Keywords: Crime against Women, K-Means Clustering, Kernel Density Estimation, Elbow method, spatial statistics, NCRB.

JEL Classifications: B54, J12, J16, C23, C31.

[1] Research Scholar, Department of Economics, Assam University, Email ID: chandranidutta98@gmail.com
[2] Assistant Professor, Department of Economics, Assam University, Email ID: amritapramanick87@gmail.com

1. Introduction

Over the years, despite continuous efforts and active policy interventions to reduce crime against women and promote gender equality, gender-based violence in India has consistently been a concerning issue (refer to a report on Women's safety in India: Between official data and lived reality in 2025 by Aadrita Haldar[3]).According to a Thompson Reuters Foundation Survey in 2018, India had been ranked as one of the "most dangerous country for women"[4]. According to this report, it is claimed that not much has been done to curb the gender-based violence in India. Does the situation show any improvement since? According to the Women, Peace and Security (WPS) Index 2025-26 published by the Georgetown Institute for Women, Peace and Security, India has been ranked 131 out of 181 countries, with a moderate score of 0.607.According to NCRB data, there has been a steady increase (not withstanding occasional fluctuations noticed between 2004 and 2006) in total cases registered under the Indian Penal Codes (IPC) and Special and Local Laws (SLL) in India. What is more threatening is the fact that the incidence of rape is highest among women in the age group 18-30 years. The impact of this on their incentives to work and participate in economic, political as well as high skilled work is crucial.[5]

The instances of violence can span across various forms ranging from physical and sexual assault to psychological abuse and control; this being one of the major issues highlighted in the literature on public health that is found to affect victims physically, mentally and socially (see Senthil et al. (2023)). According to Sarieddine (2018), such phenomenon occurs both within and outside the home across societies, countries and religions**;** becoming deeply ingrained socio-normative genetics of societies. Often part of patriarchal archetypes, "perpetrations" such as these are used as tools to restrain women's access to resources and participatory activities in society (see Chakraborty et.al (2018)). This is reflected in exclusion or limited access for women across various socio-economic and institutional activities. According to the World Economic Forum's Global Gender Gap Index, 2025, India is ranked at 131 among the 141 countries. This makes India also amongst the lowest ranked countries in South-Asia.

[3]See https://timesofindia.indiatimes.com/india/womens-safety-in-india-between-official-data-and-lived-reality-in-2025-year-ender/articleshow/126190067.cms. Retrieved on 16/01/2026.
[4]See the report at https://www.theguardian.com/global-development/2018/jun/28/poll-ranks-india-most-dangerous-country-for-women. Retrieved on 16/01/2026.
[5]The dual impact (on perpetrators and victims alike) of 'crime' on the incentives towards labor market participations shall be discussed in details later.

Crime in general, as is seen from the lens of the society, is an intentional act identifiable as illegal, adverse to the interests and well-being of individuals or society at large under existing laws, as well as in violation of the constitutional rights of individuals. Widely offered definitions agree that such acts are 'punishable' and identifiable as 'offence' against morality or a 'violation' of Law of the State[6]. According to World Health Organization, 'Violence Against Women' is identified as "any act of gender-based violence that results in, or is likely to result in, physical, sexual, or mental harm or suffering to women, including threats of such acts, coercion or arbitrary deprivation of liberty, whether occurring in public or in private life". As such, these acts may be identified as intentional acts of violence targeted towards women and is particularly seeking of legal attention due to the fact that women rights and status in society may be treated differently in conventional societies vis-à-vis that of men. This may lead to greater vulnerability and exclusion of women in general (see Iyer et.al. (2012)). As such, women rights receive special attention globally in order to create protection and sensitivity against gender-based violence[7].

Crime against women has also been understood and defined using several approaches; the most common among them being the Criminal Justice Approach and the Public Health Approach (see Kilpatrick, (2004)). From the Criminal Justice Approach, violence against women is defined as the subset of violent crimes targeted and committed against women like aggravated assault, rape, murder, sexual offense, assault and stalking. This approach defines violence against women strictly through the legal and penal frameworks and relies on crime statistics and reported crime for analysing instances of violence. The public health approach of crime against women or violence against women is defined as the subset of interpersonal violence. This approach of defining violence or crime against women focusses primarily on the prevention and health consequences of violence. World Health Organisation in its groundbreaking world report on Violence and Health has defined violence as the intentional use of physical force or power, threatened or actual, against one-self or another person or a community that results in or is likely to cause injury, death, psychological harm, maldevelopment or deprivation (see Kilpatrick, (2004)).

[6]For more details see Merriam Webster Dictionary and the Britannica Encyclopedia.
[7] For more refer to https://www.who.int/health-topics/violence-against-women.

According to the UN World Population Prospects 2022[8], women constitute almost 49.75% of the total world population. However, there remains a distinguishable gap between the opportunities available to men and women in various ways – in the labour market (female labour force participation stands at 48.7% compared to that of 73% among male according to ILO 2023 reports), in the form of wage-gaps and gender-specific categorization of work, access to education and skills, health care, parliamentary representations, participatory governance as well as in the relative independence to financial decision-making (see Petrongolo and Ronchi, (2020)). Becker (2010) in his book 'The Economics of Discrimination' has talked about discrimination and its economic outcomes. He showed that having discriminatory attitude or behaviour towards a certain group of people or an individual can have negative economic outcome. This approach can be applied to understand how discriminatory economic practices like unequal pay, gender-bias in work distribution can lead to devaluation of work done and its economic value for women. This leads to an environment with limited economic opportunities for women, socio-cultural rigidities, lower access to education and skill-trainings, health and finally economic deprivations for women. The perpetuation of such practices can lead to higher poverty among women referred to in the literature as the phenomenon of Feminization of Poverty.

The UN Women has defined gender-based violence as the violence which is inflicted upon women just because she is a woman[9]. Such crimes can be broadly classified as physical violence, sexual violence, psychological violence or economic violence. Crime against women is specifically classified as a 'tool' of executing gender-sensitive bias in attitudes and treatments. Historically, the site of gender discrimination has been identified as the 'family' in traditional societies. The associated outcomes have been identified to have larger social denominations in terms of access of women to ownership and inheritance of properties, land, marriage and rights to decision-making. With the development of social institutions and the labour market, the basis of gender-based discriminations have enlarged to exclusion from high skilled work, technical education, administrative and political representations, etc as is documented in Valverde (2011). This has led to work-place insecurities and access to limited opportunities among women. According to SDG 5, goals have been laid down to eradicate "inequalities in opportunities" for women. However, this can only be achievable based on a

[8]Refer to https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/wpp2022_summary_of_results.pdf

[9]For further details see https://www.unwomen.org.

multi-dimensional correction in attitudes and in the manner of treatment meted out to women – within the household as well as in other social denominations.

According to the latest NCRB 2023 Reports, the reported number of cases registered as crimes committed against women are4, 48, 211which is an estimated increase of 0.7% over 2022. The crime rate registered per lakh women population was 66.2 in 2023 which is a decrease from the earlier record of 66.4. Majority cases registered under IPC - Cruelty by Husband or His Relatives (29.8%), Kidnapping & Abduction of Women (19.8%), Assault on Women with Intent to Outrage her Modesty (18.71%), POCSO (14.8%) are indicative that gender-based violence is substantial in India. However, these are at best, underestimated figures given the fact that India suffers from a serious problem of under-reporting of crime inflicted upon women especially within households and acquaintances (see Gupta, 2014)).

More promiscuous cases may be cited as the rape cases of Unnao (2017)[10], Hathras (2020)[11], the R.G. Kar Medical College Rape Case (2024), which clearly explains the status of women's security at work place or elsewhere in India. Nonetheless, domestic violence and intimate partner violence are serious threats among Indian women– as of 2025 approximately 30% of women in India is subject to Intimate Partner Violence in their lifetime according to a report by WHO[12]. Approximately 25% of the cases registered with the National Commission for Women (NCW) in 2025-26 is listed under Domestic Violence (for more see NCW Reports). The Indian Judiciary, as burdened as it is (according to India Justice Report 2025, only 1.5 Judges are available per lakh population)[13], is also criticized for lags and delays in handlings such cases[14]. The seriousness and sensitivity of these crimes also arise out of the fact that any civil society must take responsibility towards all its citizens, especially the vulnerable and the victims. But any procedural delay or lag in responsible conduct of the Court of Law, the Administration and the Police may lead to broken chains of evidence (circumstantial) leading to acquittals and non-closure of victims' grievances.

In view of these critical aspects identified, it is pertinent to construct a detailed understanding of the nature of crime against women in India. The present study conducts an analysis of crime

[10]See https://www.lokatantra.in/articles-details/Unnao-Rape-Case
[11]Refer to https://www.bbc.com/news/world-asia-india-64820721
[12]For more see https://www.who.int/news/item/19-11-2025-lifetime-toll--840-million-women-faced-partner-or-sexual-violence.
[13]For more refer to https://indiajusticereport.org/files/English_National%20Press%20Release_India%20Justice%20Report%202025.docx.pdf.
[14]To be briefly elaborated in later sections.

against women through 2001 till 2023 (the latest available data) and its spatial features. It has been identified that 'Crime' has spatial linkages, that is, the socio-economic and institutional conditions that trigger incidence of gender-based crimes is similar among contiguous spaces (be it districts or states) which can be studied to reveal detailed structural patterns of crime behaviour across cross-section or panel data sets. This study attempts to detect the patterns of Crime against Women across Indian states using clustering technique like K-Means clustering. The analysis is done using twenty-three years data from National Crime Records Bureau (NCRB) from 2001-2023.

2. Crime against Women in India: An Overview

India has seen a gradual but steady rise in the absolute numbers of reported crimes against women (see Mongoli and Tarase, (2009; Kumar and Roshni, (2024); Sinha and Singha, (2015)). For the period under our study, i.e., 2001-2023, the data reveals an upward trend in the numbers of reported crimes under IPC and SLL as shown in Figure 1. The CAGR calculated for the 23 years is 4.8%.

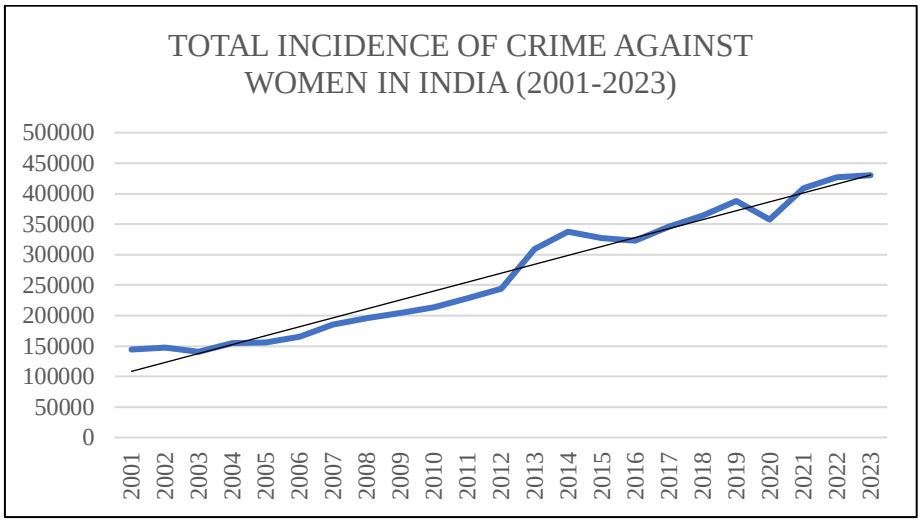


Fig.1. Total Incidence of Crime Against Women in India (2001-2023)
Source: Authors' own calculations from NCRB Data

However, there has been no specific measure to estimate the gap between actual and reported crimes. In effect, under-estimation due to under-reporting of crime against women is a serious deterrent in Indian context. The literature suggests there have been identifiable levels of bias in state-level data on crime rates in India, due to the fact in specific states socio-economic factors may generate higher levels of awareness and active reporting relative to others. An

estimate provided by a report published by Small Arms Survey (SAS) in 2015, by the Graduate Institute of International and Development Studies in Geneva, Switzerland, approximately 1 in 10 to 1 in 200 cases of rape goes unreported in India Similarly, there is ample evidence that intimate partner violence as well as violence inflicted by acquaintances and relatives go under-reported in Indian societies. This lands us in some confusion – is this rising trend to be interpreted based on the numbers available? And this may project the state with higher rate of reporting as states inflicted with higher crime and vice versa. In this paper, the authors accept the limitation existing due to under-reporting. However, for convenience and due to lack of clarity, takes the opportunity to interpret the data as per existing conventions and numbers.

The NCRB data highlights a decreasing trend in the rate of reported crimes over the years. The overall number of reported cognizable crime has declined from 60.96 lakh to 58.24 lakh cases accounting for around 4.5% decline in 2022, despite this, it is observed that the share of crime against women in total cases registered under IPC has increased over the years and during the same time crime against women have notedly increased by 4%.[15]The Year-On-Year growth of Crime against Women in India as shown in NCRB data for the duration 2001-23 is presented in Figure 2.

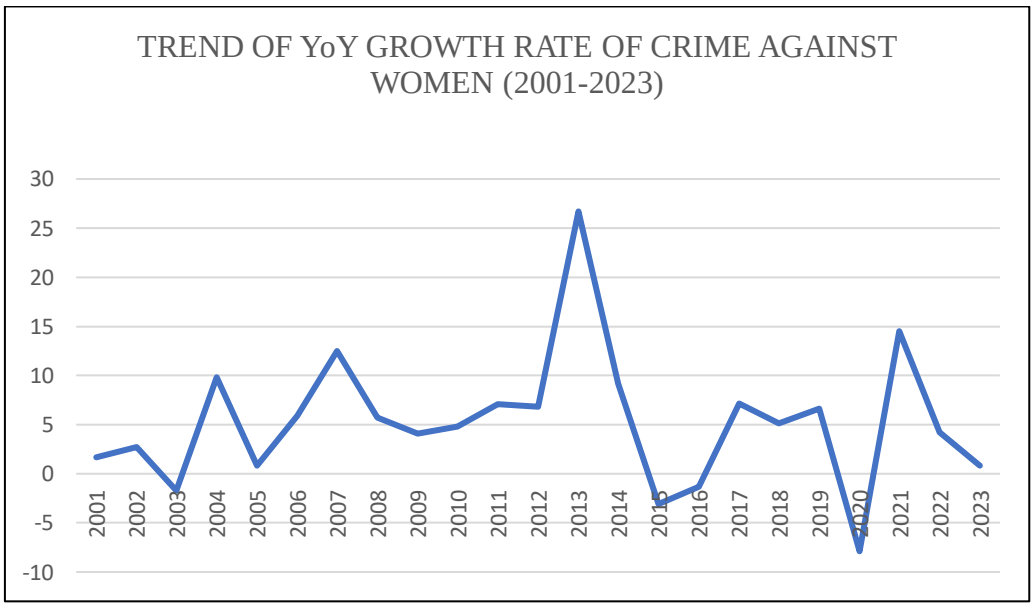


Fig.2: Year-on-Year Growth in Crime Against Women (2001-2023)
Source: Authors' own calculations from NCRB Data

[15]To see details visit https://www.data.gov.in/catalog/crime-india-2022

Figure 2 exhibits some important features: specific years (2004, 2007, 2013, 2017, 2019 and 2021) register a sudden and drastic spike in the rate of crime against women. This may be attributed to the facts that 2012 was a year of violent protests against the Nirbhaya Case in New Delhi leading to a spike in the rate of reporting and public awareness as well as protest (see Nigam, (2014)); that 2004, 2014 and 2019 were years when Lok Sabha Elections were held; the crisis due to COVID-2019; literature shows that domestic violence had drastically increased in India during the pandemic affected period 2020-2022 (see Prakash, (2021)); and many other socio-economic or demographic factors. Such correlations are documented and is found to have significant impacts on different crimes committed against women.

The recent report of National Crime Records Bureau (NCRB) 2023recordsfifteen types of crime heads related to gender-based crimes; majority cases being reported under Cruelty by Husband or Relatives (1,33,676 cases), Kidnapping and Abduction (1,33,676 cases) followed by Assault on Women with intent to outrage her modesty (83,891 cases)[16]. However, over the years NCRB has diversified and broadened its categorisations of crime heads. From 2001-2013, NCRB recoded only seven types of crime heads viz. Rape, Kidnapping and Abduction, Dowry Deaths, Cruelty by Husband and Relatives, Molestation, Eve-Teasing and Importation of Girls. In 2014-2015, another crime head was introduced- Abetment of Suicide of Women. The very next year in 2016, four more crime heads were included- Unnatural Offences, Human Trafficking, causing miscarriage without women's consent, Deaths caused by act done with intent to cause miscarriage and acid attack. Since then, from 2017-2023, NCRB has uniformly reported data under fifteen crime heads; the most recently included crime head being Murder with rape/gang rape, selling of minor girls, buying of minor girls since 2017. This gives us a broader data spectrum for deeper understanding of crime against women.

The National Commission for Women (NCW) reports for 2025-26 suggests that approximately 5910 cases of Domestic Violence are registered out of the total 23692 cases, which constitutes a 25% of all complaints registered with NCW. Assault (13.8%), Dowry Harassment (14.3%), Criminal Intimidation (13.3%) ranks next among the other crime heads listed with NCW that received significant numbers of complaints registered during 2025-26. The other important feature to notice is the state-wise distribution of crime recorded as per complaints registered with NCW; the state-wise list of states reported with highest number of complaints received by NCW for the last five years is as follows:

[16]Refer to https://www.ncrb.gov.in/uploads/files/1CrimeinIndia2023PartI1.pdf

States	Complaints received by NCW 2025-26	Complaints received by NCW 2024-25	Complaints received by NCW 2023-24	Complaints received by NCW 2022-23	Complaints received by NCW 2021-22
Uttar Pradesh	5061	5375	15762	17048	15445
Delhi	1467	1590	2387	2822	3361
Maharashtra	1246	1182	1445	1336	1460
Bihar	848	934	1271	1360	1450
Madhya Pradesh	698	703	1148	1162	1103
West Bengal	614	509	605	614	755
Haryana	576	597	1142	1291	1397
Rajasthan	573	516	981	1003	1117

Table 1: States and UTs Reported with Highest Crime against Women as per Complaints received by NCW during 2025-26
Source: Portal for Statistical Overview of Received Complaints (NCW)

Noticeably, Uttar Pradesh has consistently reported highest incidence of crime against women, along with Delhi, Haryana and Rajasthan. This constitutes the Northern belt of India with highest reported crimes against women. Though, there is a decline in reported crimes among these States and UTs, the numbers remain alarmingly high in these states. According to Bhattacharya et.al. (2021) incidence of crime follows identifiable spatial patterns. There is much evidence for the observation in the fact that Uttar Pradesh, Punjab, Haryana, Madhya Pradesh, Rajasthan have consistently recorded higher crimes relative to other states (see Mukherjee et al., (2001); Sharma, (2023)). On the other hand, North-Eastern states of Arunachal Pradesh, Sikkim, Manipur, Mizoram, Nagaland were found to have very low crime rates (see Petrongolo and Ronchi, (2020); Maity and Roy, (2021); Chakraborty et.al (2021); Goswami, (2019); Aggarwal et.,al (2023); Murmu, (2023); Dev and Shaheen, (2014); Vicente, Goicoa and Fernandez, (2020); Jha, (2015); Chauhan and Baraik, (2016)).

The other feature to note is incidence of crime is higher among specific states and UTs: Delhi, Maharashtra, Karnataka and West Bengal. One of the prime reasons may be attributed to the existence of a metropolitan city in these states. The higher incidence of crime in these states and UTs, specifically Delhi, involves a complex combination of various socio-economic and demographic factors including higher density of population, significant number of migrant population and rapid increase in the number of slums (see Kumari et., al (2025)). Also, other factors such as the rate of urbanization, migration and modernization, existence of industrial clusters etc. determine the rate of incidence (see Lamba (2025)).

Crime against women is a worldwide phenomenon but the visibility of these crimes differs substantially across countries, within countries as well as across cultures (see Chernoff and Simon, (2000)). The countries which are economically developed with a higher Gross Domestic Product had higher crime rates than the less developed countries (see Kersten, (1993), Zhao, (2000)). The 1993 U.N World conference on Human Rights have defined violence against women as the act of gender-based violence resulting in the harm of women. However, the definition of crime or violence against women has varied across different countries and times. With development and progress, definitions have been reviewed and homogeneity in practices has been established across economies with the definitions offered by United Nations and World Health Organization. For instance, crime against women has been mainly represented as domestic violence in United Kingdom during the 1990s. Highly industrialised countries like United States, Finland and Netherlands had the largest increase in cases of crime against women; Germany and Japan have witnessed a gradual decline in the rate of rape cases (see Yodanis, (2004), Kulwicki, (2002)). Developed countries have a mixed trend of crime against women; whereas some countries are witnessing a decline, others are having a steady rate of reported crimes. The LDCs on the other hand, have seen a steady increase in the crime rates against women (see Kangaspunta and Marshall, (2012); Hadidi, (1996); Heidensohn, (1975); Johnson, (2001); Niaz, (2003); Wilson, (2014); Nasir and Hyder, (2003)).

Is such a pattern identifiable among Indian states???

STATES	PCNSDP QUARTILE	CRIME RATE QUARTILE
Andhra Pradesh	Q2(Low to Moderate)	Q3 (MODERATE)
Arunachal Pradesh	Q2(Low to Moderate)	Q3 (MODERATE)
Assam	Q1(Low)	Q4 (High)
Bihar	Q1(Low)	Q1 (LOW)
Chhattisgarh	Q2(Low to Moderate)	Q3 (MODERATE)
Goa	Q4(High)	Q3 (MODERATE)
Gujrat	Q4(High)	Q1 (LOW)
Haryana	Q4(High)	Q4 (High)
Himachal Pradesh	Q3(Moderate)	Q2 (LOW TO MODERATE)
Jammu and Kashmir	Q1(Low)	Q2 (LOW TO MODERATE)
Jharkhand	Q1(Low)	Q2 (LOW TO MODERATE)
Karnataka	Q4(High)	Q2 (LOW TO MODERATE)
Kerala	Q3(Moderate)	Q4 (High)
Madhya Pradesh	Q1(Low)	Q4 (High)
Maharashtra	Q3(Moderate)	Q3 (MODERATE)
Manipur	Q1(Low)	Q1 (LOW)
Meghalaya	Q1(Low)	Q1 (LOW)
Mizoram	Q2(Low to Moderate)	Q1 (LOW)
Nagaland	Q2(Low to Moderate)	Q1 (LOW)
Odisha	Q2(Low to Moderate)	Q4 (High)
Punjab	Q3(Moderate)	Q2 (LOW TO MODERATE)
Rajasthan	Q2(Low to Moderate)	Q4 (High)
Sikkim	Q4(High)	Q1 (LOW)
Tamil Nadu	Q3(Moderate)	Q1 (LOW)
Telangana	Q3(Moderate)	Q3 (MODERATE)
Tripura	Q2(Low to Moderate)	Q3 (MODERATE)
Uttar Pradesh	Q1(Low)	Q2 (LOW TO MODERATE)
Uttarakhand	Q4(High)	Q2 (LOW TO MODERATE)
West Bengal	Q1(Low)	Q4 (High)

Table 2: Comparison between PCNSDP at 2011-12 base price and Crime Rates during 2011-2023
Source: Author's own calculation from PCNSDP and NCRB data

Table 2 shows the comparison between the crime levels of different states along with their Per-capita Net State Domestic Product at 2011-12 base price. The analysis helps in understanding whether Indian states exhibit similar patterns of development and crime as seen globally. The comparison across Indian states shows no clear monotonic relationship. Both rich states and poor states fall in high and low crime quartile. The high-income states like Goa, Gujarat, Haryana, Karnataka, Sikkim and Uttarakhand falls across all the quartiles of crime level. Similarly, low-income states like Assam, Bihar, Jammu and Kashmir, Jharkhand, Manipur, Meghalaya, Uttar Pradesh, West Bengal and Madhya Pradesh also indicates mixed patterns. The states with higher GDP like Haryana and Kerala having higher crime rates can be explained

by the modernisation hypothesis which states that as states become richer and moves towards modernisation with better institutions, reporting of crime also may increase due to factors like urban exposure and increased female mobility[17]. Various literatures have also indicated towards the under reporting of crime in low-income states which may be a reason behind Bihar being in both low income and low crime quartile, (see Soares (2004)). The mixed patterns among Indian states shows that development and growth alone cannot be a sufficient variable to reduce crime levels in a state.

2.1. Crime and the Economy:

Violence against women has been and remains "one of the world's most persistent and under-addressed human rights crises" as per reports of World Health Organization. With growing economies across the globe, questions have been raised about how crimes hamper economic growth and development. For any substantial amount of development to take place, the society must be provided with a feeling of security and for the security to be ensured, the level of crime in any society must be brought to minimal. Study conducted by (see Balogun et.al (2014)) in Nigeria found that crime surpasses all stratifications of the society affecting both the Haves and the Have-nots. Studies have confirmed that the level of crime negatively impacts the performance of the economy. Higher levels of crime in the economy discourage foreign investment, reduce the competitiveness of the firms and relocate the resources creating uncertainty and inefficiency. A model studied by (see Detotto and Otranto (2010)) has identified that an average rise in the crimes by 1% reduces the real economic growth by 0.00040% in a month and by 0.00022% in a year. A similar study conducted by (see Debnath and Das (2017)) studied the relationship between crime and economic welfare in India between the time period 1982-2013. Their study found that the relationship between these two variables is time dependent. Their results also identified a positive impact of economic growth on violent crime in the long run, but a rather negative impact in the short run. There is sufficient evidence from the literature that an inverted U-shaped relationship exists between crime and economic development (for more refer to Chakraborty et. al (2021)). This study highlights the positive direct impact of urbanization on people's tendencies to resort to drug abuse leading to an increase in crime against women. However, after a certain point of time, employment increases engaging more population in work and improving income levels, thus lowering crime.

[17] To refer, visit https://www.encyclopedia.com/law/legal-and-political-magazines/modernization-and-crime

2.2. The Relevance of Spatial Investigation into Crime

Understanding the distribution of crime across region has been a core area for criminological studies. Studies show that crime is not randomly distributed, rather it is shaped by the socio-economic and institutional contexts in which it takes place, forming hotspots (refer to Anselin et al., (2000), Sampson et al., (1997)). Different theoretical contribution of criminological studies like Becker's economic theory of crime shows how individual choice of committing a crime is influenced by situational and institutional factors such as 'lenient and strict law enforcement' (see Becker (1968) and Bratingham and Bratingham, (2017)). Empirical studies further provide evidence of crime concentration; a phenomenon described as the "Law of Concentration" and reveals spillover effects indicating that neighbourhood areas tend to have similar levels of crime (see Weisburd, (2015), Baller et al., (2001)).

3. Analytical Approaches in Understanding Spatial Patterns of Crime against Women

To get a better understanding of the spatial patterns and the structure of crime, literatures have used analytical and machine learning methodology like K-Means Clustering and Kernel Density Estimation. Kernel Density Estimation is considered to be a well-known and one of the most effective approaches to estimate the Probability Density Function of a given set of observation and it estimates the density function of a particular set of observation automatically (see Chen, (2017)). Out of all the probability density function, Kernel Density Estimation is found to be the best method that shows the entire distribution of the probability mass across the X random variable (see Węglarczyk, (2018)). A detailed tutorial study done by (see Hu et.,al (2018)) highlights that the Kernel Density Estimation in a spatial analysis is an effective way of locating and visualising the hotspots of crime; however, it has certain limitations like it fails to capture the temporal aspect of crime for which Spatio-Temporal Kernel Density Estimation can be a more appropriate methodology. Kounadiet.,al (2020) in their paper did a systematic review of around 32 research paper and found that Kernel Density Estimation is a frequent used method for spatial crime forecasting and the studies that have used this methodology have focused primarily on predicting crime hotspots. The review also found that this methodology is one of the top performing models for spatial crime forecasting where it yields the highest Prediction Accuracy Index (PAI) when compared to other hot spot mapping techniques. Hart et al., (2014) found in their study that Kernel Density Estimation affects the

predictive accuracy of crime hotspots. It not only shows the clustering pattern of crime but also smoothens out the data from discrete observation.

K-Means Clustering is a clustering tool that aims to part n numbers of data into k number of pre-determined groups; the fundamental idea behind K-Means clustering model is to find out k mean values known as centroids around which the data representing similar characteristics are clustered; however the K-Means algorithm follows an iterative process which aims to reduce the variance within the cluster, thus fitting the most appropriate observations into a single cluster (see Burkardt, (2009)). K- Means Clustering have been repeatedly used in the clustering of different types of crime. Kodinariya (2013) showed that there are multiple ways in which the optimal number of clusters can be identified viz. Elbow Method, Rule of Thumb, Information Criterion Approach, Cross Validation, Using the Silhouette and Theoretic Approach. The study further shows that in case of elbow method; there can be ambiguity in identify the cluster points and there may be multiple elbow points. K- Means Clustering have been repeatedly used in the clustering of different types of crime. A study by Agarwal et.,al (2013) uses K- means clustering analysis to track the changes of crime in England from 1990-2011. Using this method, the entire dataset was grouped into four clusters and the results of the analysis showed that the rate of homicide has shown a mixed pattern throughout the years. Joshi et.al (2017) in their paper used Rapid Miner tool and K-Means clustering. They generated three clusters viz. Low, Medium and High to identify the crimes which are highest in different cities of Australia. A similar kind of study was done by Krisnendu et., (2020) where they used elbow method to determine the optimal number of clusters for their k-means Clustering Model. Initially, eight clusters were randomly selected for the study as the starting centroids. Elbow method was then applied and the kink in the elbow plot was determined at k=8. After finding the optimal number of clusters, K-Means clustering model was run.

With this background in mind, this paper sets out to identify the top five crimes committed against women in India from 2001-2023 and investigate into the nature of its prevalence across states. The paper also aims to check if crime against women in India exhibits clustering among the states.

4. Data and Methodology

For this study, twenty-three years state-level data of crime against women from 2001-2023 has been taken from National Crime Records Bureau. To determine the top five crimes committed

against women over these twenty-three years and to identify the worst performing states under these crime heads, descriptive statistics have been used. Aggregated number of reported crimes has been calculated and their respective share in the reported crime has been identified.

To check the pattern of clustering among the observations over the years, K-Means clustering method has been used. K-Means Clustering method primarily aims to group *n* number of observations into *k* number of clusters in a way that each of the observations in each cluster belongs to the nearest mean. The objective function of this model is to minimise the distance between the observations and its cluster centroid; thus, reducing the within cluster sum of square (WSS).

The formula for this model is;

$$\min_s \sum_{i=1}^{k} \sum_{x \in S_i} \| x - \mu_i \|^2$$

where,

min= minimization of the total within cluster sum squared

S_i= set of data points in cluster i

μ_i=Centroid or mean of the cluster i

$\|x-\mu_i\|^2$= Euclidean distance squared between observation x and centroid of cluster i

The model is run in software R where K-means algorithm uses the iterative optimisation process. In order to get a valid result from running this model in R, the observations need to be scaled. That has been done using the formula-

$$z = \frac{\mathrm{x} - \bar{\mathrm{x}}}{\mathrm{s}}$$

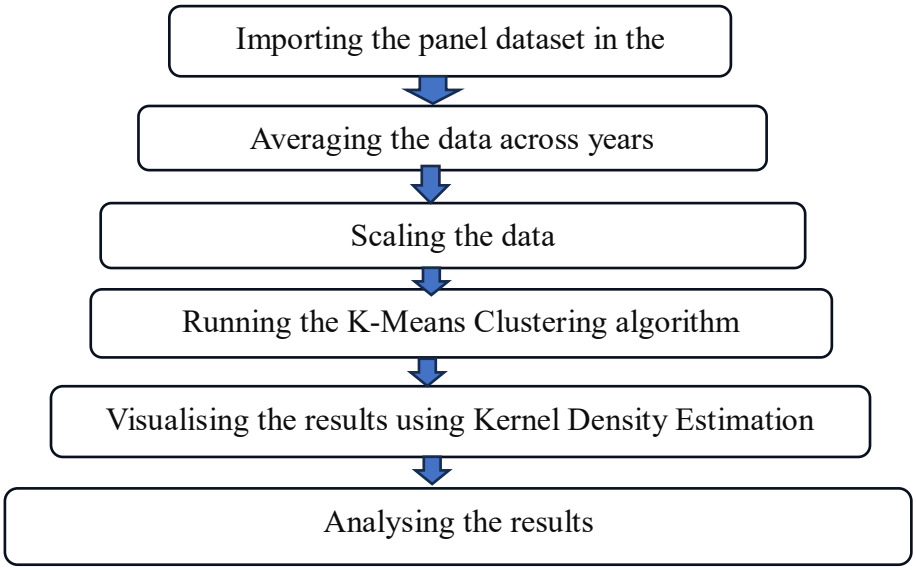
where,

x= the observation

$\bar{x}$= mean of the x observations

S= standard deviation of the x observations

To run the K-Means algorithm in R, the following steps have been used-



To identify the optimal number of clusters for this study Elbow method has been employed for efficient interpretation of the observations. Further Kernel Density Estimation is applied on the scaled value of crime using a Gaussian Kernel with an automatic bandwidth selection which is found by default in the R software in ggplot2 package. The Kernel Density can be expressed as,

$$\widehat{f_h}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right), where\ the\ Kernel\ Gausian(K) is,$$

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

$$\widehat{f_h}(x) = estimated\ probability\ density\ function\ at\ point\ x$$

$$n = number\ of\ observations$$

$$h = smoothing\ parameter\ or\ bandwidth$$

$$xi = separate\ data\ point$$

5. Results and Discussion

5.1. Identification and Study of Top Five Crime Heads (2001-2023)

From the year 2001-2023, NCRB has published the data of reported crime on several crime heads. However, to maintain a uniformity of data in this study, only those crime heads have been considered which have been reported commonly throughout the ten years. These crime heads are-

1. Rape (Sec. 376 IPC)
2. Kidnapping & Abduction (Sec.363-369,371-373 IPC)
3. Dowry Deaths (Sec.304B IPC)
4. Cruelty By Husband and Relatives/Torture (Sec.498A IPC)
5. Molestation /Assault on women with an intent to outrage her modesty (Sec.354 IPC)
6. Eve-Teasing (Sexual Harassment)/Insult to the modesty of women (Sec.509 IPC)

Table 3: Commonly Reported Crime Heads (2001-2023)
Data Source: National Crime Records Bureau

State	Rape	Kidnapping & Abduction	Dowry Deaths	Cruelty By Husband and Relatives	Assault on women	Insult to the modesty of women
Andhra Pradesh	24835	29073	8107	205278	107747	60595
Arunachal Pradesh	1266	1358	6	826	1776	141
Assam	35134	78038	3151	141794	51887	2360
Bihar	23453	100013	25798	57511	11617	1718
Chhattisgarh	26053	21097	2021	17238	37320	4395
Goa	1137	831	24	351	1686	499
Gujrat	12479	28324	644	97193	20182	1900
Haryana	21297	35847	5874	72827	29825	9465
Himachal Pradesh	5062	4720	76	5516	8901	1386
Jammu and Kashmir	5912	15312	183	4436	21351	4529
Jharkhand	21439	16445	6336	20991	15013	1261
Karnataka	13519	20896	5262	56316	81998	2387
Kerala	21230	3930	454	85659	79404	8390
Madhya Pradesh	75015	70313	15690	105110	160108	14251
Maharashtra	48427	77900	6718	166659	155338	25523
Manipur	996	1917	7	442	1277	90
Meghalaya	2073	848	26	437	1469	271
Mizoram	1240	33	4	137	1419	120
Nagaland	369	252	3	38	223	56
Odisha	28599	46859	8330	56375	129423	8941
Punjab	13278	19673	2382	27864	13917	1207
Rajasthan	67537	80120	9976	263546	105971	1821
Sikkim	457	218	4	72	487	38
Tamil Nadu	12254	22383	2891	31281	31113	8633
Telangana	7794	11166	1894	70482	37653	8444
Tripura	3285	2136	604	11217	5187	201
Uttar Pradesh	61739	181947	49757	227599	141234	23969
Uttarakhand	6413	7348	1538	10193	6755	1638
West Bengal	33335	79997	10308	509155	63259	7167
TOTAL	574237	954372	168056	3086146	1328369	202181

Table 4: Aggregate Number of reported IPC crime against women from 2001-2023
Data Source: National Crime Records Bureau

Table 4 shows the cumulative reported number of crimes under each crime head from 2001-2023. It is evident that the top 5 crimes out the six crime heads are Cruelty against Women by Husband and Relatives, Assault on Women, Kidnapping and Abduction, Rape and Insult to the modesty of women. To understand the nature of the data, descriptive statistics have been calculated to see the pattern of distribution of the data-

Variable	Observation	Mean	Std.Deviation	Min	Max
Cruelty by Husband	650	3267.68	4500.55	0	23278
Assault	665	1989.99	2687.357	0	14853
Kidnapping and Abduction	653	1471.436	2373.92	0	15381
Rape	653	881.6371	1050.99	0	6337
Insult	649	325.12	675.48	0	4970

Table 5: Descriptive statistics of the top five crime heads
Source: Authors' own calculations from NCRB Data

The descriptive statistics of the five crime heads indicates that the reporting of crime is highly uneven across states and years. Higher mean or average value across the crime heads suggests substantial variation in the average reporting of the crimes with cruelty against women being the most frequently reported crime. The higher value of standard deviation shows that the values of reported crime hugely deviate across states over the years implying strong variations in the reporting of crimes.

To find how much these crime heads have contributed to the total IPC crime, proportions of each crime head has been calculated-

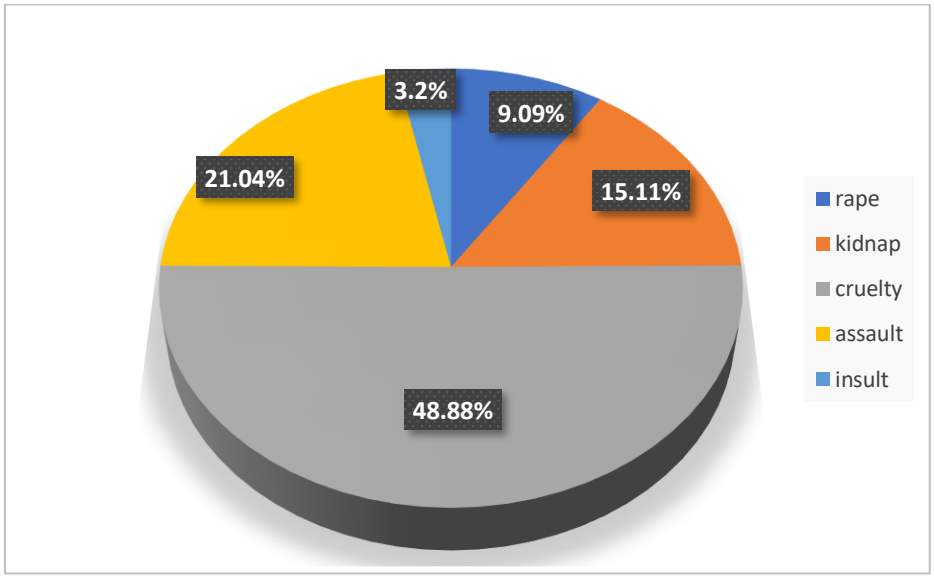


Fig 3: % share of top five reported IPC crime
Source: Authors' own calculations from NCRB Data

Fig 1 shows that Cruelty by Husband and Relatives constitutes 48.8% of the total IPC crime reported, Assault on Women constitutes 21.04%, Kidnapping and Abduction is 15.1%, Rape is 9.09% and Insult to the Modesty of Women makes up to 3.2% of the total reported IPC crime.

The calculation of the aggregate crime heads, descriptive statistics and the proportional contribution of these crime heads to the total IPC crimes confirms that the top five reported crime heads from 2001-2023 are-

Cruelty By Husband and Relatives/Torture (Sec.498A IPC),
Molestation /Assault on women with an intent to outrage her modesty (Sec.354 IPC),
Kidnapping & Abduction (Sec.363-369,371-373 IPC),
Rape (Sec. 376 IPC)
Eve-Teasing (Sexual Harassment)/Insult to the modesty of women (Sec.509 IPC),

Table 6: Top 5 reported IPC crime against women from 2013-2022

The states reported with highest incidence of crime has been listed as below for each of the above 5 crime heads. It is observed that specific states such as Madhya Pradesh, Andhra Pradesh, Maharashtra and Uttar Pradesh have been most frequently found to be reported as exhibiting highest incidences of cases across the above five crime heads based on the average annual reported crimes during the time-period 2001-2023. However, the reasons cannot be nailed down as of now – mostly citable reasons may be urbanization, migration, per capita net state GDP, literacy rates and sex ratio, female workforce participation rate (Maity and Roy, 2020; Chakraborty et., al, 2018).

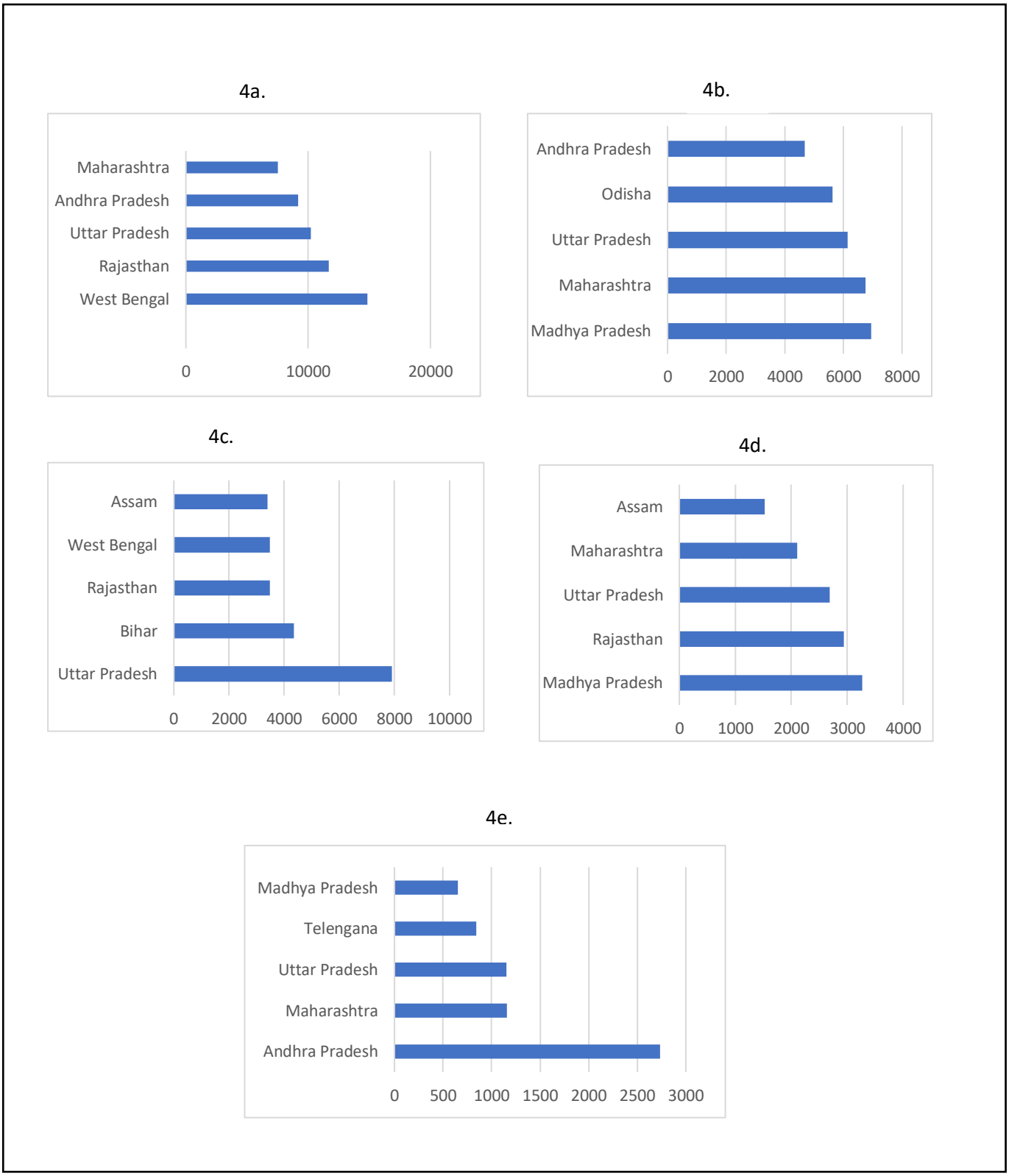


Fig 4. States reporting highest incidence of crime: a) Cruelty by Husband and Relatives, b) Assault on Women, c) Kidnapping and abduction, d) Rape, d) Insult to the Modesty of Women
Source: Author's own calculation from NCRB data

5.2. Kernel Density Estimation

Kernel Density Estimation is a non-parametric method used to estimate the probability density function of a random variable. To get a better understanding of the nature and distribution of the major crimes across states, Kernel Density Estimation (KDE) approach was undertaken. The results of KDE for the top crimes are reported below-

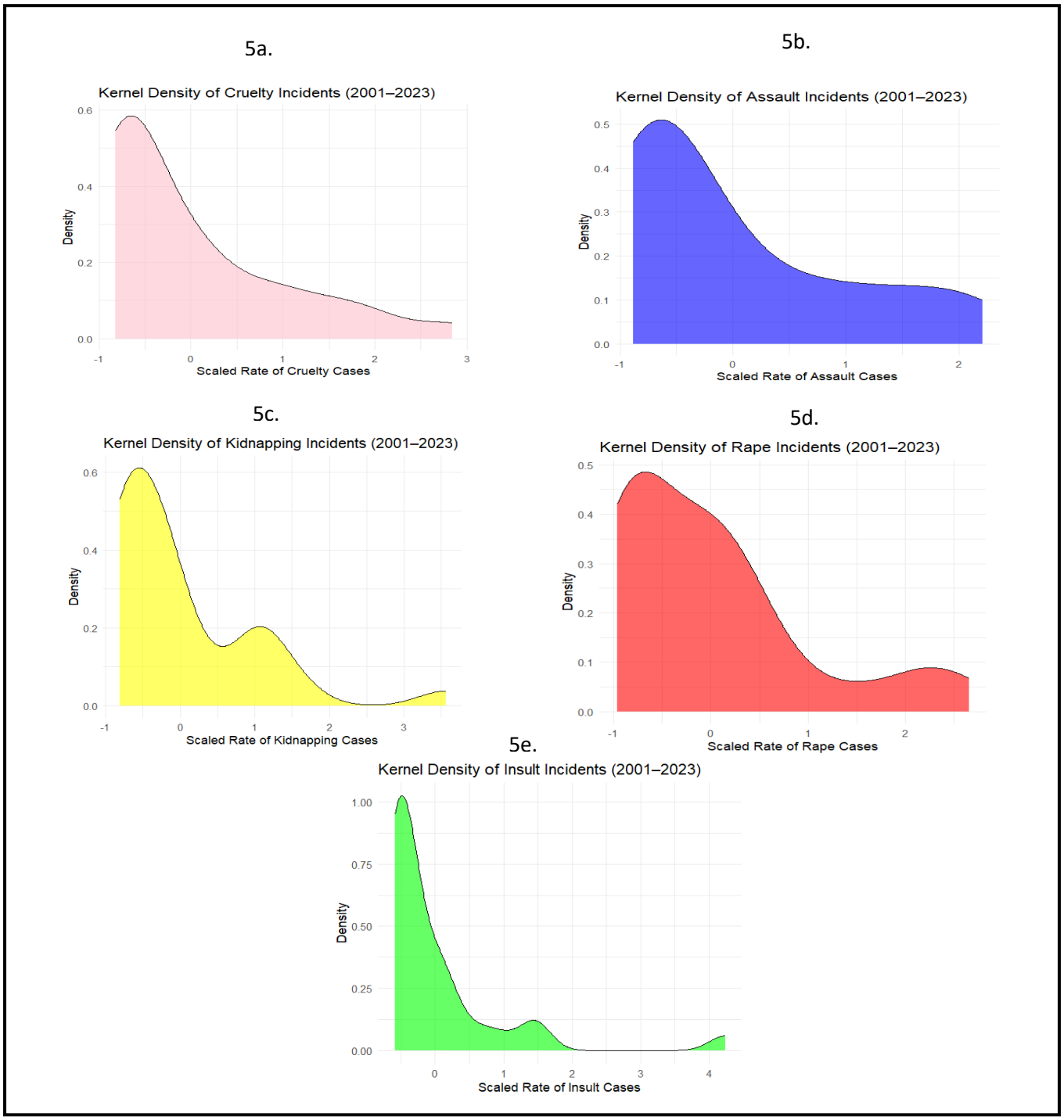


Fig 5: Results for Kernel Density Estimation: a) Cruelty by Husband and Relatives, b) Assault on Women, c) Kidnapping and abduction, d) Rape, e) Insult to the Modesty of Women
Source: Author's own calculation from NCRB data

5a. The x-axis represents the standardized or scaled rate of cruelty incidents (2001-2023) and the y-axis represents the Probability Density Function. The KDE shows that the distribution of cruelty incident is rightly skewed and the peak occurs below 0. The occurrence of peak density below 0 implies that most of the states experience cruelty rate below or close to average. The tail of the curve is long and extends to the right showing that there is a smaller group of state which experience significantly higher number of cruelty cases. The curve further extends to the right with a smooth and slight bump around 2 standard deviations implying that there is less variation in the reporting across state or years in the distribution of cruelty cases.

5b. For assault cases, the Kernel Density shows a rightward skewed distribution with a long tale. The curve reaches its peak before 0 indicating that maximum states report assault cases below the average level. The curve then declines slowly in density showing that a smaller number of states or years have recorded higher number of assault cases extending up to 2 standard deviations from the average level. The curve however has a smooth decline indicating a continuous variation among the observations rather than having discrete or bi-modal peaks.

5c. The Kernel Density for Kidnapping and Abduction shows a rightward skewed multi-modal distribution. The first peak takes place below 0 indicating that most of the states or years have recorded kidnapping cases below average rate. The second peak takes place at around 1 showing that there is a group of state which has recorded kidnapping cases 1 standard deviation above the average level. The curve then declines and extends rightward with a long tail and a slight bump beyond 3 standard deviations implying the presence of extreme outlier among the observations.

5d. The KDE for rape is a unimodal right-skewed distribution. The curve consists of a peak and a slight bump afterwards with a long tail rightwards. The first peak is below zero which signifies that most states have had below average rate of rape cases which are clustered around -0.5 to 0. The curve then slows downwards with a slight bump occurring at 2 standard deviations away from the average level which indicates that there exists a cluster of outlier states which have recorded extremely high rate of rape across the years.

5e. The KDE for cases of insult is leptokurtic meaning that it has a sharp peak near the average value or 0. The distribution has a sharp peak just below zero with multiple peaks, first one at $z \approx -0.5$, the second one at $z \approx 1.5$ and the third bump at $z \approx 4$. The first peak below

0 signifies that majority of the states have reported dowry cases below the average level. The second bump around z≈1.5 indicates that a group of state exist with moderately higher number of insult cases. The curve then has a long tail rightwards with a possible third minor bump at z≈4 signifying the presence of few extreme outlier states.

5.3. A Comparative Analysis Amongst the Kernel Distributions of Major Crime Heads

After individual analysis of kernel estimation of the major crime heads separately, a comparative analysis is done clubbing all the kernel estimations to get an understanding of how each crime head has performed when compared to each other.

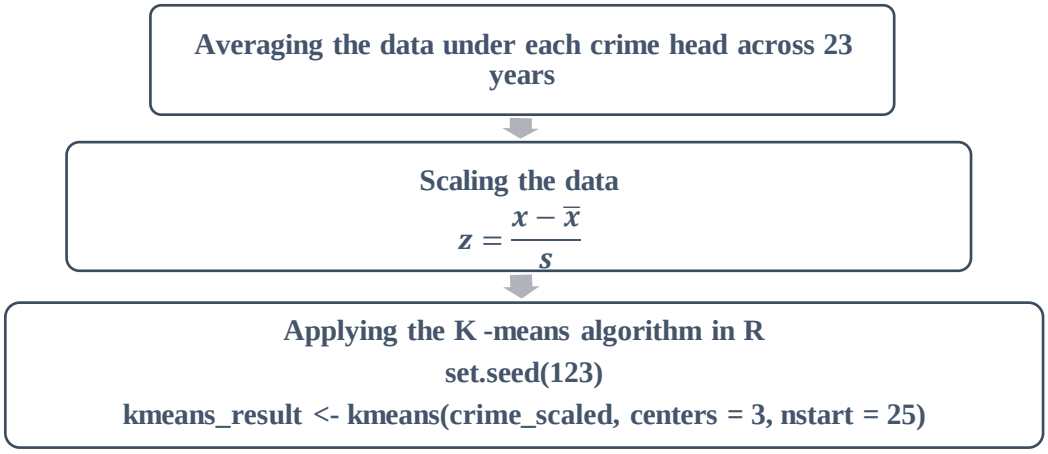
CRIME HEADS	PEAKS	SHAPE OF THE DISTRIBUTION	NATURE OF THE DISTRIBUTION
Cruelty by Husband and Relatives	z≈ -0.25	Rightward-Skewed	One strong peak below the average with a smooth distribution away from the average level. This indicates lesser variations across states above average and an even distribution of states with low to moderate rates of crime.
Assault against Women	z≈ -0.5	Rightward-Skewed	Single peak just below the average level and a smooth right tail indicating a more uniform distribution of states at above average reported crime levels.
Kidnapping and Abduction	z≈ -0.5, z≈ 1, z≈3.5	Rightward-Skewed and multi-modal	Multi-modal distribution of reported crimes. There are three identifiable clusters – specifically peaks are observed at below average crime rates, and two above average rates of incidences.
Rape	z≈ -0.5, z≈2.5	Rightward-Skewed	Strong bimodal distribution observed with one peak at below average rate of incidence. The other peak (is a softer peak) noticed at higher-than-average rate of incidence.
Insult to the Modesty of Women	z≈ -0.5, z≈1.5, z≈4.5	Leptokurtic with multiple bumps	The distribution of state is clustered around below average (-0.5) and two peak at above average (1.5, 4.5). the existence of three peaks indicates the presence of three distinguishable clusters.

Table 7: Comparative analysis of the kernel distribution of five crime heads

The comparative analysis of the Kernel Density Estimation indicates that most of the states have reported crime below the average level. However, under each crime heads, there are local clustering indicating that there are group of states that have reported crimes much higher than that of the average level. The bimodal distribution with two peaks and local bumps along with the leptokurtic distribution with multiple peaks implies the presence of outliers and local clusters in the distribution of crime against women in the last twenty-three years.

5.4 K-Means Clustering

K-Means Clustering is a machine learning method which is effectively used upon a dataset to group the similar kinds of data points into k clusters, where each of the data point with the nearest mean belong to the same group. The following steps have been undertaken to get desirable results from K-Means model-



Initially, 3 clusters were taken randomly viz. low crime, moderate crime and high crime. At k=3, total within cluster sum squares (WSS) was 46.7. To check the efficiency of clustering, r^2 was calculated and was found out to be 0.66. Theoretically, if r^2 value lies below 0.5 it indicates weak clustering, for moderate clustering it lies between the range of 0.5-0.7 and any r^2 value above 0.7 indicates good clustering. To find out the optimal number of clusters for the dataset of 667 observations, elbow method was run-

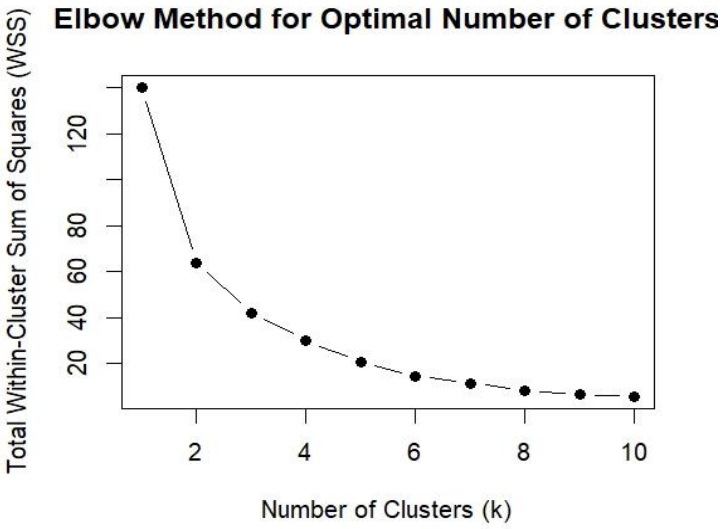


Fig 6: Elbow Plot for Optimal Cluster
Source: Authors' own calculations from NCRB Data

The optimal number of clusters for this study was determined using the Elbow Method by finding out the Total Within Sum Squared (WSS), Between Sum Squared (BSS) and r^2. The elbow plot shows that the optimal number of clusters at the point where the curve kinks towards the x-axis. Subsequently K-Means algorithm was run again to check WSS at k=4, 5, 6, 7, 8 and 9. Also, r^2 value was calculated at each K level to check the effectiveness of clustering. The results found are shown below-

No. of K	Total Within Sum Squared (WSS)	Between Sum Squared (BSS)	R^2
3	46.7	93.29	0.66
4	32.39	107.6	0.76
5	24.54	115.45	0.82
6	19.61	120.38	0.85
7	15.66	124.33	0.88
8	11.1	128.9	0.92
9	11.1	128.9	0.92

Table 8: WSS, BSS, r^2 at different levels of K
Source: Author's own findings running codes in R

The results show that the WSS is decreasing at a decreasing rate whereas BSS is increasing at a decreasing rate and both becomes zero at k =9. At K=3, WSS is 46.7 which drops sharply by

14.31 when K moves from 3 to 4. Parallelly, BSS also increases substantially from 93.29 to 107.6 when K shifts from 3 to 4. This increase in the level of K also increases the effectiveness of the clustering from 0.66 to 0.76. Beyond K=4, WSS drops at a decreasing rate, BSS increases at a decreasing rate and r^2 increases but the gains become smaller indicating the major structural break at K=4 beyond which the elbow curve starts flattening gradually and finally becomes 0 at K=9. With these results in consideration, the optimal number of K for this study was taken as 4. The four clusters were subsequently labelled as 'Low Crime', 'Low to Moderate Crime', 'Moderate crime' and 'High Crime'. The results obtained are-

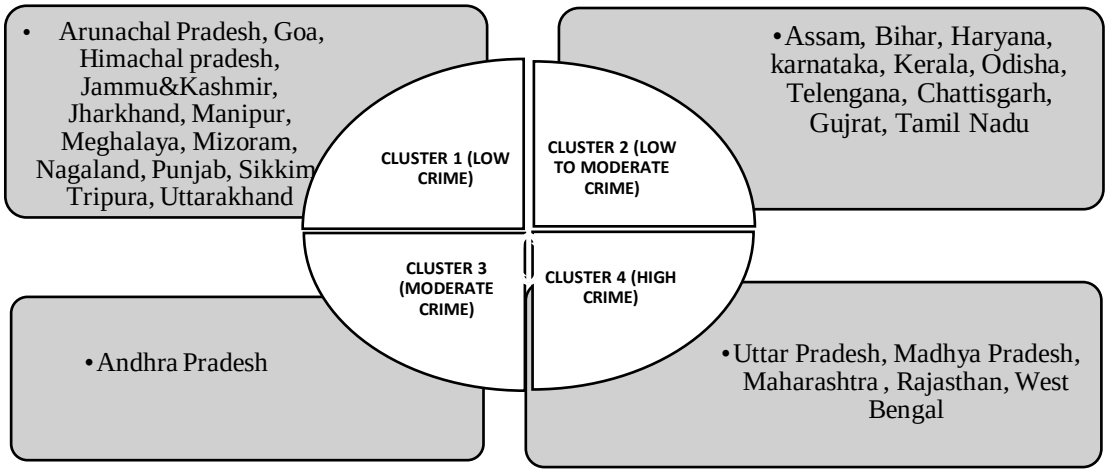


Fig 7: Results from K-Means Clustering
Source: Authors' own calculations from NCRB Data

The results from K-Means Clustering clarifies the implications made by Kernel Density Estimation and the report of National Commission of Women (NCW). The results from Kernel Density Estimation showed that all the crime heads have the highest peak near the average value or close to 0 which indicated that most of the states are clustered at a lower level of crime. The smaller bumps or peaks implied towards the presence of a smaller number of states having relatively higher number of reported crime and the sharp right tail in each of the crime head showed that there are states with extreme high number of crimes and outliers. The clustering pattern showed by K-Means Clustering coincides with the implications of Kernel Density Estimation in a way that it shows that most of the states including the North-Eastern states fall in cluster 1 indicating lower level of crimes. A relatively smaller number of states like the South Indian states of Telangana, Kerala, Karnataka, Orissa and Tamil Nadu fall in cluster 2 i.e., Low to moderate crime, Andhra Pradesh is the only state that falls in Cluster 3 indicating Moderate Crime and Uttar Pradesh, West Bengal, Rajasthan, Maharashtra and Madhya Pradesh fall in Cluster 4 implying that these states have the highest number of reported crimes from 2001-

2023. The states identified in Cluster 4 (high Crime) also closely correspond with the states reporting highest number of cases of crimes against women to the National Commission for Women in the last five years. This overlap between the results signifies that K-Means clustering have effectively captured the regions with persistent higher crime. To see how Per-capita Net State domestic Product might impact crime rate across India, Indian states were classified into quartile based on their average PCNSDP and compared with the results of K-Means clustering. The results indicate towards heterogeneous relationship where several states with high PCNSDP like Goa, Gujarat, Karnataka, Haryana have reported low or low to moderate crime, while other states with low PCNSDP such as West Bengal, Rajasthan, Uttar Pradesh have consistently reported higher levels of crime.

However, anomalies in the result including Bihar and Haryana having low to moderate crime level indicates the under-reporting of crimes in these states (Gupta, 2014); Various studies show that around 86% of women in Bihar were unaware of the helplines meant for reporting domestic violence, and of the women who were aware, 80% did not know how to report against domestic crimes (Kumari, 2024). One of the various reasons behind this under-reporting can be explained with the help of social norm theory which states that individual behaviour is driven strongly by the thinking and behavioural patterns of the other members of the social group (Berkowitz, 2005). This theory works on the core idea that actions are influenced by attitude, so behaviour will not change until and unless the underlying beliefs and norms are restructured (Ball and Paluck, 2010). Thus, norms theory can be used to understand India's cultural values which influences women to under-report crimes like domestic violence, assault and rape; and given the high prevalence of patriarchy in states like Bihar and Haryana, norm theory explains that many women may not even realise the actual rate of domestic violence (Kumari, 2024).

6. Conclusion

The study has thus provided meaningful insights into the study of spatial analysis of crime in a sense that it has been able to indicate towards the presence of clustering behaviour among the states in cases of reported crime against women in India. Methodologies like kernel density estimation and k-means clustering has identified the highest reporting states under different crime heads. Further, it was observed that the major crime against women have followed a right skewed distribution. The kernel density estimation showed that most of the states display crime rate below average while a smaller group of state and specifically Uttar Pradesh, Maharashtra, West Bengal, Rajasthan and Madhya Pradesh have repeatedly reported higher incidences of

crime. This implies the presence of local clusters of crime, unequal distribution and outliers in the observation.

The results from kernel density estimation showed disparities in the observation through its unimodal and bimodal pattern of curves, implying hotspots of crime and spatial concentration of crime as discussed in the study of Weisburd (2004). The clustering of states in the K-means clustering and the findings from the study reveals pattern of crimes similar to that of literatures already reviewed in the study such as Singh et al., (2020) and Bhattacharya et al., (2021) who also observed regional clustering and neighbouring effect in criminal activities. Over the twenty-three years, smaller group of states have consistently reported higher number of crimes which indicates the presence of hotspots and crime clustering as put forwarded by Anselin et al., (2000). This study thus emphasises the importance of spatial policy targeting and resource allocation.

Reference

1. Agarwal, J., Nagpal, R., & Sehgal, R. (2013). Crime Analysis Using K-means Clustering. *International Journal of Computer Applications*, *83*(4).

2. Anselin, L., Cohen, J., Cook, D., Gorr, W., & Tita, G. (2000). Spatial Analyses of Crime. *Criminal justice*, *4*(2), 213-262.

3. Baller, R. D., Anselin, L., Messner, S. F., Deane, G., & Hawkins, D. F. (2001). Structural Covariates of US County Homicide Rates: Incorporating Spatial Effects. *Criminology*, *39*(3), 561-588.

4. Balogun, T. F., Okeke, H., & Chukwukere, C. I. (2014). Crime Mapping in Nigeria Using GIS. *Journal of Geographic Information System*, *6*(5), 453-466.

5. Becker, G. S. (2010). The Economics of Discrimination. University of Chicago press.

6. Berkowitz, A. D. (2005). An Overview of the Social Norms Approach. *Changing the culture of college drinking: A socially situated health communication campaign*, *1*, 193-214.

7. Bhattacharyya, A., Haldar, S. K., & Banerjee, S. (2021). Determinants of Crime against Women in India: A Spatial Panel Data Regression analysis. *Millennial Asia*, *13*(3), 411–441. Available at https://doi.org/10.1177/09763996211003379.

8. Bhattacharyya, R. (2015). Understanding the Spatialities of Sexual Assault Against Indian Women In India. *Gender, Place & Culture*, *22*(9), 1340-1356.

9. Brantingham, P. L., & Brantingham, P. J. (2017). Environment, Routine, And Situation: Toward A Pattern Theory of Crime. In *Routine activity and rational choice* (pp. 259-294). Routledge.

10. Chakraborty, Chandrima; Afreen, Anam; and Pal, Dipyaman (2021). Crime against Women in India: A State Level Analysis. *Journal of International Women's Studies*, 22(5), 1-18.

11. Chakraborty, T., Mukherjee, A., Rachapalli, S. R., & Saha, S. (2018). Stigma of Sexual Violence and Women's Decision to Work. *World Development*, *103*, 226-238.

12. Chauhan, R., & Baraik, V. K. (2016). Mapping Crime Against Women in India: Spatio-Temporal Analysis, 2001-2012. *International Journal of Law and Political Sciences*, *10*(6), 2243-2254.

13. Chen, Y. C. (2017). A Tutorial on Kernel Density Estimation and Recent Advances. *Biostatistics & Epidemiology*, *1*(1), 161-187.

14. Chrisler, J. C., & Ferguson, S. (2006). Violence Against Women as a Public Health Issue. *Annals of the New York Academy of Sciences*, *1087*(1), 235-249.

15. Debnath, A., & Das, S. (2017). Relationship Between Crime and Economic Affluence in India: An Empirical Study From 1982 To 2013. *Journal of Criminological Research, Policy and Practice*, *3*(1), 27-37.

16. Detotto, C., & Otranto, E. (2010). Does Crime Affect Economic Growth? *Kyklos*, *63*(3), 330-345.

17. Dev, K (2014). A Spatial and Temporal Analysis of Crime Against of Women in India. In *Human Rights International Research Journal* [Journal-article]. https://www.researchgate.net/publication/366670483.

18. Dutta, P. (2017). Deprivation of Women in Northeast India: An Exploratory Study. In *Inequality, Poverty and Development in India: Focus on the North Eastern Region* (pp. 223-236). Singapore: Springer Singapore.

19. Gupta, A. (2014). Reporting And Incidence of Violence Against Women in India. *Rice Institute*, *130*.

20. Hadidi, M. (1996). The National Institute of Medicine: The Role of the Forensic Medicine in Uncovering Crimes Committed Against Women.

21. Hart, T., & Zandbergen, P. (2014). Kernel Density Estimation and Hotspot Mapping: Examining the Influence of Interpolation Method, Grid Cell Size, And Bandwidth on Crime Forecasting. *Policing: An International Journal of Police Strategies & Management*, *37*(2), 305-323.

22. Heidensohn, F. (1975). The Imprisonment of Females. In *The Use of Imprisonment* (pp. 43-56). Routledge.

23. Hu, Y., Wang, F., Guin, C., & Zhu, H. (2018). A Spatio-Temporal Kernel Density Estimation Framework for Predictive Crime Hotspot Mapping and Evaluation. *Applied geography*, *99*, 89-97.

24. Iyer, L., Mani, A., Mishra, P., & Topalova, P. (2012). The Power of Political Voice: Women's Political Representation and Crime in India. *American Economic Journal: Applied Economics*, *4*(4), 165-193.

25. Jha, D. (2015). Geography Of Rape Crime in India: A Spatial Analysis of Official Data, *International Journal of Research in Social Sciences*, *5*(5).

26. Johnson, H. (2001). Enhancing Knowledge on Violence Against Women. *Publication Series-European Institute for Crime Prevention and Control*, 198-213.

27. Joshi, A., Sabitha, A. S., & Choudhury, T. (2017, October). Crime Analysis Using K-Means Clustering. In *2017 3rd International conference on computational intelligence and networks (CINE)* (pp. 33-39). IEEE.

28. Kangaspunta, K., & Marshall, I. H. (2012). Trends In Violence Against Women: Some Good News and Some Bad News. In *The international crime drop: New directions in research* (pp. 103-133). London: Palgrave Macmillan UK.

29. Kersten, J. (1993). Crime And Masculinities in Australia, Germany and Japan. *International Sociology*, *8*(4), 461-478.

30. Kilpatrick, D. G. (2004). What Is Violence Against Women: Defining and Measuring the Problem. *Journal of interpersonal violence*, *19*(11), 1209-1234.

31. Kodinariya, T. M., & Makwana, P. R. (2013). Review On Determining Number of Cluster In K-Means Clustering. *International Journal*, *1*(6), 90-95.

32. Kounadi, O., Ristea, A., Araujo Jr, A., & Leitner, M. (2020). A Systematic Review on Spatial Crime Forecasting. *Crime science*, *9*(1), 7.

33. Krishnendu, S. G., Lakshmi, P. P., & Nitha, L. (2020, March). Crime Analysis and Prediction Using Optimized K-Means Algorithm. In *2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 915-918). IEEE.

34. Kulwicki, A. D. (2002). The Practice of Honor Crimes: A Glimpse of Domestic Violence in The Arab World. *Issues in mental health nursing*, *23*(1), 77-87.

35. Kumar, D. (2025). Urbanization And Crimes Against Women in India: A Spatial and Statistical Analysis. *International Journal of Science and Research (IJSR), 14*(3), Article SR25324145034.

36. Kumar, K. A., & Roshni, S. Exploratory Data Analysis on "Crime against Women in India from year 2001–2021".

37. Kumari, R. Domestic Violence & Crime Against Women: A Curse for Civilized Society in The Context of Bihar & Government's Initiative for Women's Safety.

38. Kumari, S., Kumar, N., & Sharma, A. (2025). Mapping The Dynamics of Crime Against Women in India: A Spatio-Temporal Analysis. *Belgeo. Revue belge de géographie*, (1).

39. Lamba, N. (2025). Changing Urban Realities and Women's Vulnerability: A Metropolitan Analysis of Crimes Against Women Using Crime in India Reports, 2021–2023. *Journal of South Asian Research*, *3*(1), 205-226.

40. MacQueen, J. (1967). Multivariate Observations. In *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1, pp. 281-297).

41. Mangoli, R. N., &Tarase, G. M. (2009). Crime Against Women in India: A Statistical Review. In *International Journal of Criminology and Sociological Theory* (Vol. 2, Issue 2, pp. 292–302).

42. Mukherjee, C., RustagiP., Krishnaji N, Crime Against Women in India-Analysis of Offocial Statistic, *Economic and Political Weekly*, Oct. 27 - Nov. 2, 2001, Vol. 36, No. 43 (Oct.27 -Nov. 2, 2001).

43. Nasir, K., &Hyder, A. A. (2003). Violence Against Pregnant Women in Developing Countries: Review of Evidence. *The European Journal of Public Health*, *13*(2), 105-107.

44. Niaz, U. (2003). Violence Against Women in South Asian Countries. *Archives of women's mental health*, *6*, 173-184.

45. Nigam, S. (2012). Violence, Protest and Change: A Socio-Legal Analysis of Extraordinary Mobilization After The 2012 Delhi Gang Rape Case, 197-221.

46. Pal, S. (2020). *Essays On Social Norms and Status of Women in Northeastern India* (Doctoral dissertation).

47. Paluck, E. L., & Ball, L. (2010). Social Norms Marketing to Reduce Gender Based Violence. *IRC Policy Briefcase*.

48. Parzen, E. (1962). On Estimation of a Probability Density Function and Mode. *The annals of mathematical statistics*, *33*(3), 1065-1076.

49. Petrongolo, B., & Ronchi, M. (2020). A Survey of Gender Gaps Through the Lens of The Industry Structure and Local Labor Markets.

50. Prakash, V. (2021). *COVID-19-Impact: Increased Violence Against Women in India. Kardan Journal of Social Sciences and Humanities, 4*(1), 57–67.

51. Rosenblat, M. (1956). Remarks On Some Nonparametric Estimates of a Density Function. *Ann. Math. Stat*, *27*, 832-837.

52. Sarieddine, M. (2018). Violence Against Women: A Phenomenological Perspective. *ClinExpPsychol*, *4*(01), 1-5.

53. Sen, A. (2000). Development As Freedom. *Development in Practice-Oxford-*, *10*(2), 258-258.

54. Senthil, N., &Vajiram, J. (2023). The Misuse of Law by Women in India-Constitutionality of Gender Bias. *arXiv preprint arXiv:2307.14651*.

55. Sharma, M., & Kumar, S. (2023). Regional Dimension of Crime against Women in India. *Annals of the National Association of Geographers, India*, *43*(1), 110-127.

56. Sinha, A., & Singha, P. (2015). Crime Against Women in India: Trend and Regional Variation. *International Journal of Advanced Research*, 1403–1406. https://www.researchgate.net/publication/334771275.

57. Soares, R. R. (2004). Crime Reporting as a Measure of Institutional Development. *Economic Development and cultural change*, *52*(4), 851-871.

58. Syakur, M. A., Khotimah, B. K., Rochman, E. M., & Satoto, B. D. (2018, April). Integration K-Means Clustering Method and Elbow Method for Identification of The Best Customer Profile Cluster. In *IOP conference series: materials science and engineering* (Vol. 336, p. 012017). IOP Publishing.

59. Valverde, M. E. R. (2011). Gender-Based Political Harassment and Violence: Effects on The Political Work and Public Roles of Women. *NEW SOLUTIONS: A Journal of Environmental and Occupational Health Policy*, *20*(4), 527-535.

60. Vicente, G., Goicoa, T., Fernandez-Rasines, P., & Ugarte, M. D. (2020). Crime Against Women in India: Unveiling Spatial Patterns and Temporal Trends of Dowry Deaths in The Districts of Uttar Pradesh. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *183*(2), 655-679.

61. Węglarczyk, S. (2018). Kernel Density Estimation and Its Application. In *ITM web of conferences* (Vol. 23, p. 00037). EDP Sciences.

62. Weisburd, D., Bushway, S., Lum, C., & Yang, S. M. (2004). Trajectories Of Crime at Places: A Longitudinal Study of Street Segments in The City Of Seattle. *Criminology*, *42*(2), 283-322.

63. Wilson, T. D. (2014). Violence Against Women in Latin America. *Latin American Perspectives*, *41*(1), 3-18.

64. Yodanis, C. L. (2004). Gender Inequality, Violence Against Women, And Fear: A Cross-National Test of The Feminist Theory of Violence Against Women. *Journal of interpersonal violence*, *19*(6), 655-675.

65. Zhao, Y. (2000). Domestic Violence in China: In search of legal and social responses. *Ucla Pac. Basin Lj*, *18*, 211.

Important Links:

https://ncwapps.nic.in/frmComp_stat_Overview.aspx