



## CRIME AGAINST WOMEN IN INDIA : ANALYSIS OF SPATIAL DATA AND FACTORS INFLUENCING CRIME

**K.Dheenathayalan<sup>1</sup>, K.K. Savitha<sup>2</sup>**

<sup>1</sup>Assistant Professor, Department of Computer Science, Kamban College of Arts and Science, Coimbatore, Tamil Nadu, India.

<sup>2</sup>Assistant Professor, Department of Computer Applications, Bharathiar University, PG Extension and Research Centre, Erode, Tamil Nadu, India.

<sup>1</sup>[dheenamca08@gmail.com](mailto:dheenamca08@gmail.com), <sup>2</sup>[savitha.gopinath@gmail.com](mailto:savitha.gopinath@gmail.com)

**Abstract:** Data mining is the process of analyzing hidden patterns of data according to different perspectives for categorization into useful information. Detecting crime patterns will help to speed up the process of crime reduction. In crime terminology, a cluster is a group of crimes in a geographical region or a hot spot of crime. Identifying crime hotspots using the geospatial plot of crime helps to improve the efficiency of law enforcement officers. In this proposed work crime patterns are identified using k-means (KM) clustering with an enhancement by the grasshopper optimization algorithm (GOA). Crime patterns are mapped to geospatial data to plot crime hotspots. And then some of the environmental and societal factors which influence crime against women are discussed. The proposed technique is applied to official crime data acquired from National Crime Records Bureau (NCRB) India. This work is focused on analyzing and identifying crime patterns of crimes against women in India. It can also be applied to any kind of crime to analyze and predict crime patterns.

**Keywords:** crime pattern, crime hotspot, geo-spatial, clustering, spatial data.

### 1. INTRODUCTION

The Criminal Justice System in India has evolved over many years influenced by the socioeconomic and political conditions prevailing during the different phases of the history of India. The criminal justice system plays a crucial role in securing justice for the people and maintaining the unity and integrity of the nation. Failure of the criminal justice system endangers the whole civil society leading it towards a hectic circumstance. Crime analysis has become one of the most important functions within the criminal justice system. Crime analysis supports and bolsters investigators and law enforcement officers to do their jobs. A crime analyst researches crimes to identify trends and patterns of crimes. All the crimes are classified into two categories cognizable and non-cognizable. A cognizable offense or case is defined as one in which an officer-in-charge of a police station may investigate without the order of a magistrate and arrest without a warrant. Non-Cognizable crimes are defined as those which cannot be investigated by police without the order of a competent magistrate. Again these two categories are broadly classified into Indian Penal Code (IPC) and Special and Local Laws (SLL).

Crime analysis may determine the possibility of a crime pattern based on geographic

proximity, found through crime mapping. In crime terminology, a cluster is a group of crimes (high crime intensity) in a geographical region or a hotspot of crime. Crime mapping is devoted to detect crime hot spots. Crime hot spot analysis assists police departments in identifying high-crime areas, types of crime being committed, and the best way to respond. In this research work, the K-means clustering algorithm is enhanced by the grasshopper optimization algorithm and applied to identify crime patterns. Then the crime patterns are mapped to geospatial data to identify types of crimes being committed in crime hot spots.

## 2. DATASET DESCRIPTION

The official crime dataset is acquired from NCRB India's official website [ncrb.gov.in](http://ncrb.gov.in). Crime data were taken from 2000 to 2020. Since the dataset does not contain a region attribute it is manually added to the dataset for analysis purposes. These are the year-wise datasets and each dataset consists of crime attributes like region, state name, number of crime incidents, and crime head. The territory of India is divided into six regions East India, West India, North India, South India, Central India, and North East India. Crime head is otherwise known as crime type and there are seventeen different crime heads recorded such as rape, kidnapping and abduction, Assault on women, cruelty by husband or relatives, and so on. The dataset consists of 28 states and 8 union territories with the corresponding region. Two kinds of datasets are formed, one is a state-wise / crime head-wise crime dataset and which contains the state's name and union territory's name with the total number of crimes recorded under different crime heads, another one is a region-wise / crime head-wise dataset and it contains total crime incidents in regions under different crime heads. Table.1 and Table.2 are examples of datasets that were used in this work.

**Table.1. sample dataset of state/UT wise – crime head-wise total crimes**

State/UT	Total rape	Total kidnap	Total Assault	Total insult	Total cruelty
Andhra Pradesh	23502	21094	91825	55547	181419
Assam	33488	65723	43809	1026	122465
Bihar	22050	69026	10356	658	53118
Chhattisgarh	24342	15378	33716	2977	15418
Gujarat	9411	24852	18141	1809	94209
Haryana	17210	26635	21939	9874	57511
Jharkhand	26321	12768	11590	460	18924
Karnataka	17540	15851	63902	1979	50553
Kerala	30442	3319	65589	6319	73515
Madhya Pradesh	101020	49454	143510	12151	84237
Maharashtra	61681	51683	121125	22396	140061
Odisha	36910	30765	97876	6800	41896
Punjab	17511	15149	11917	1005	24106
Rajasthan	79145	60047	78928	633	215249
Tamil Nadu	13029	20919	26972	10693	29161

**Table.2. sample dataset of region wise-crime head-wise total crimes**

Region	State/UT	Total rape	Total kidnap	Total Assault	Total insult	Total cruelty
--------	----------	------------	--------------	---------------	--------------	---------------

<b>Central</b>	Chhattisgarh	24342	15378	33716	2977	15418
<b>Central</b>	Madhya Pradesh	101020	49454	143510	12151	84237
<b>East</b>	Bihar	22050	69026	10356	658	53118
<b>East</b>	Jharkhand	26321	12768	11590	460	18924
<b>East</b>	Odisha	36910	30765	97876	6800	41896
<b>East</b>	West Bengal	38770	59481	55818	5254	282191
<b>North</b>	Haryana	17210	26635	21939	9874	57511
<b>North</b>	Himachal Pradesh	4187	3699	7418	1007	5199
<b>North</b>	Punjab	17511	15149	11917	1005	24106
<b>North</b>	Rajasthan	79145	60047	78928	633	215249
<b>North</b>	Uttar Pradesh	73116	142214	111840	26789	176329
<b>North</b>	Uttarakhand	7290	5552	4777	1675	8005

### 3. KM-GOA CLUSTERING

The Grasshopper optimization algorithm is hybridized with the K-means clustering algorithm to overcome the drawback of it. The main drawback of k-means is it has a possibility of convergence to local minima because of the dependence on initial cluster centers. The Optimization algorithm plays an important role in finding global minima. So, to overcome the drawback of k-means, initial centroids are optimized using the grasshopper optimization algorithm then k-means clustering is performed to group data based on a distance measure. Here in this work, the Euclidean distance measure was used to find the similarity/dissimilarity between data points. Euclidean distance is a very familiar and frequently used distance measure to find similar data points.

In the clustering process, geographical areas and crime attributes are clustered by using the proposed clustering technique. After the clustering process, geographical groups are found and visualized using shapefiles. A shapefile is a simple, non-topological format for storing the geometric location and attribute information of geographic features. Geographic features in a shapefile can be represented by points, lines, or polygons (areas).

## 4. RESULT AND DISCUSSION

### 4.1 Cluster results

Geographical clusters according to different crimes are given below. The below tables describe the cluster results from the KM-GOA technique. Attributes of tables are Cluster (Cluster number), States/UT in the cluster (Name of states or union territories in the respective cluster), cluster size (number of data points in the cluster), and average crime incidents from 2001 to 2020. Crime data is clustered into three groups to find high-crime intensity, mid-crime intensity, and low-crime intensity areas based on total crime incidents and under different crime types.

Here are cluster results based on total crime incidents, rape crime incidents, kidnapping and abduction crime incidents, assault on women crime incidents, and cruelty by husband/relative crime incidents. Table.3 presents the details of clusters based on total crime incidents. In this table, cluster1 has six states in it and the average of crime incidents recorded in these states is 457216. So we can see that states in this cluster are high-crime-intensity areas.

Whereas cluster2 in Table.1 has ten states and is the second highest among three clusters, so states in this cluster are mid-crime intensity areas. In cluster3 there are 21 states with low averages than others. So these are the low crime intensity areas. Table.4, Table.5, Table.6, and Table.7 are similar to the Table.1 and they each represent different types of crime.

**Table.3 State/UT clusters based on total crime incidents**

Cluster	States/UT in the cluster	Cluster Size	Average crime incidents (2001-20)
Cluster-1	Andhra Pradesh, Madhya Pradesh, Maharashtra, Rajasthan, Uttar Pradesh, and West Bengal.	6	457216.3
Cluster-2	Assam, Bihar, Gujarat, Haryana, Karnataka, Kerala, Odisha, Tamil Nadu, Telangana, Delhi	10	177822.3
Cluster-3	Arunachal Pradesh, Chhattisgarh, Goa, Himachal Pradesh, Jharkhand, Manipur, Meghalaya, Mizoram, Nagaland, Punjab, Sikkim, Tripura, Uttarakhand, Andaman & Nicobar Islands, Chandigarh, Dadra & Nager Haveli, Daman & Diu, Jammu & Kashmir, Ladakh, Lakshadweep, Puducherry	21	20403.3

Table.2 represents clusters of rape crime data, and states in cluster3 are high crime intensity states. Here rape crime incident is a total of all rape-related crimes such as the attempt to rape, gang rape, custodial rape, and rape with murder. Whereas Table.3 is for kidnapping and abduction and cluster3 has high crime intensity states. Table.4 represents the assault on women; states in cluster3 are high crime intensity areas. Similarly, Table.5 represents cruelty by husband or relatives; cluster1 has high crime intensity states.

**Table.4 Crime type: Rape / attempt to rape / rape with murder / custodial rape / gang rape**

Cluster	States in cluster	Cluster Size	Average crime incidents (2001-2020)
Cluster 1	Andhra Pradesh, Assam, Bihar, Chhattisgarh, Haryana, Jharkhand, Karnataka, Kerala, Odisha, Punjab, Telangana, West Bengal, Delhi	13	24766
Cluster 2	Arunachal Pradesh, Goa, Gujarat, Himachal Pradesh, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim, Tamil Nadu, Tripura, Uttarakhand, Andaman & Nicobar Islands, Chandigarh, Dadra & Nager Haveli, Daman & Diu, Jammu & Kashmir, Ladakh, Lakshadweep, Puducherry.	20	2723.947
Cluster 3	Madhya Pradesh, Maharashtra, Rajasthan, Uttar Pradesh	4	78740.5

**Table.5 Crime type: Kidnapping and Abduction**

Cluster	States in cluster	Cluster Size	Average crime incidents (2001-2020)
Cluster 1	Assam, Bihar, Madhya Pradesh, Maharashtra, Rajasthan, West Bengal, Delhi	7	57286.71
Cluster 2	Andhra Pradesh, Arunachal Pradesh, Chhattisgarh, Goa, Gujarat, Haryana, Himachal Pradesh, Jharkhand, Karnataka, Kerala, Manipur, Meghalaya, Mizoram, Nagaland, Odisha, Punjab, Sikkim, Tamil Nadu, Telangana, Tripura, Uttarakhand, Andaman & Nicobar Islands, Chandigarh, Dadra & Nager Haveli, Daman & Diu, Jammu & Kashmir, Ladakh, Lakshadweep, Puducherry	29	8137.25
Cluster 3	Uttar Pradesh	1	142214

**Table.6 Crime type: Assault on woman with intent to outrage her modesty**

Cluster	States in cluster	Cluster Size	Average crime incidents (2001-2020)
Cluster 1	Assam, Chhattisgarh, Karnataka, Kerala, Tamil Nadu, Telangana, West Bengal, Delhi	8	44486.500
Cluster 2	Arunachal Pradesh, Bihar, Goa, Gujarat, Haryana, Himachal Pradesh, Jharkhand, Manipur, Meghalaya, Mizoram, Nagaland, Punjab, Sikkim, Tripura, Uttarakhand, Andaman & Nicobar Islands, Chandigarh, Dadra & Nager Haveli, Daman & Diu, Jammu & Kashmir, Ladakh, Lakshadweep, Puducherry	23	5500.591
Cluster 3	Andhra Pradesh, Madhya Pradesh, Maharashtra, Odisha, Rajasthan, Uttar Pradesh.	6	107517.333

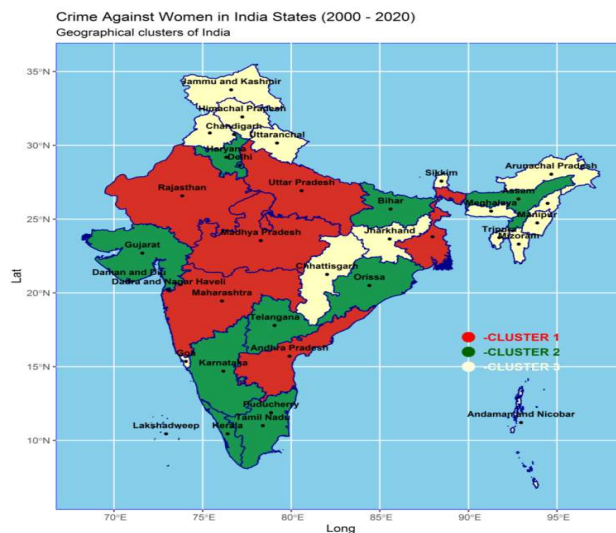
**Table.7 Crime type: Cruelty by husband/relatives**

Cluster	States in cluster	Cluster Size	Average crime incidents (2001-2020)
Cluster 1	Andhra Pradesh, Maharashtra, Rajasthan, Uttar Pradesh, West Bengal	5	199049.800
Cluster 2	Assam, Bihar, Gujarat, Haryana, Karnataka, Kerala, Madhya Pradesh, Odisha, Telangana, Delhi	10	67635.100
Cluster 3	Arunachal Pradesh, Chhattisgarh, Goa, Himachal Pradesh, Jharkhand, Manipur, Meghalaya, Mizoram, Nagaland, Punjab, Sikkim, Tamil Nadu, Tripura,	22	5417.143

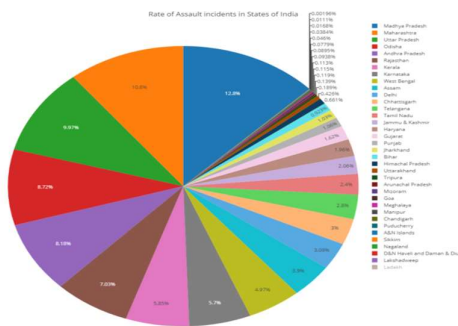
Uttarakhand, Andaman & Nicobar Islands, Chandigarh, Dadra & Nager Haveli, Daman & Diu, Jammu & Kashmir, Ladakh, Lakshadweep, Puducherry		
---	--	--

Fig.1 is the pictorial representation of the Table.1, it shows geographical clusters based on total crime incidents from 2000- 2020. There are three clusters as mentioned in Table.1. State territories in red color are in cluster-1 of high crime intensity areas. Territories in green color are cluster-2 of mid-crime intensity areas, and territories in pale yellow color are cluster-3 of low-crime intensity areas. Fig.2, Fig.3, Fig.4, and Fig.5 are pie charts which show the percentage of crime incidents recorded under different crime types in all states.

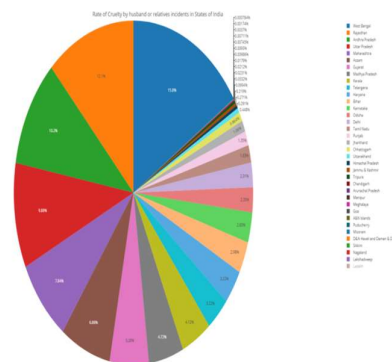
We can easily identify the states with a high percentage of crime incidents under each crime head. In Fig.2 it shows that 12.8 % of assault on women crime is recorded in Madhya Pradesh state, 10.8% of crimes are recorded in Maharashtra and 9.97% of crimes are recorded in Uttar Pradesh followed by other states of India. Whereas Fig.3 shows that 15.8 % of cruelty by husband or relatives crime is recorded in West Bengal state, 12.1% of crimes are recorded in Rajasthan and 10.2% of crimes are recorded in Andhra Pradesh followed by other states of India.



**Fig.1 Geographical clusters of India based on crime against women (2000-2020)**

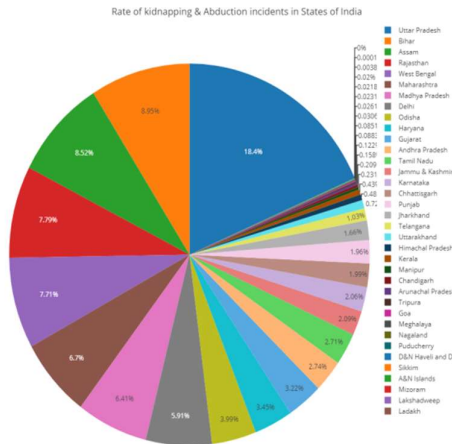


**Fig.2 Percentage of assault on women incidents**

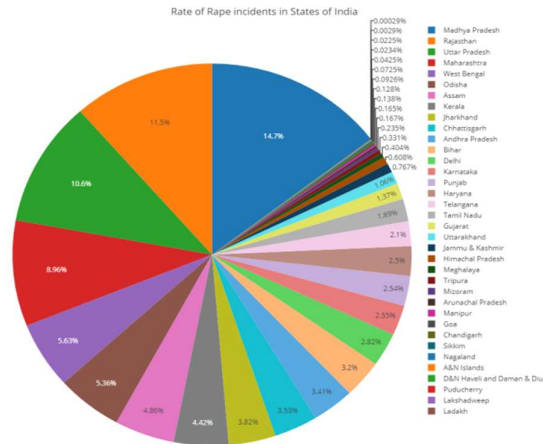


**Fig.3 Percentage of cruelty by husband incidents**

Fig.4 shows that 18.4 % of kidnapping and abduction crimes are recorded in Uttar Pradesh state, 8.95% of crimes are recorded in Bihar and 8.52% of crimes are recorded in Assam followed by other states of India. Whereas Fig.5 shows that 14.7 % of rape crime incidents are recorded in Madhya Pradesh state, 11.5% of rape crimes are recorded in Rajasthan and 10.6% of rape crimes are recorded in Uttar Pradesh followed by other states of India.



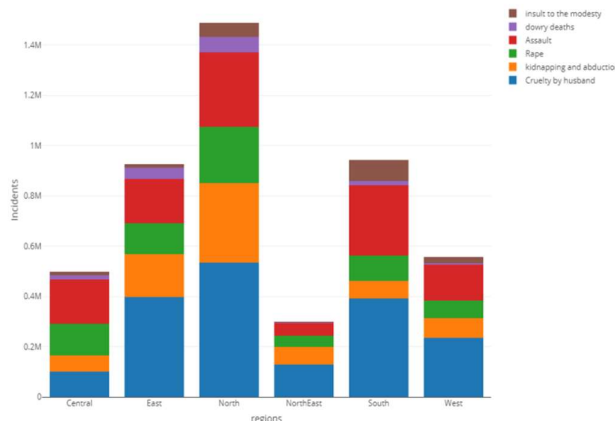
**Fig.4 Percentage of kidnapping and abduction incidents**



**Fig.5 Percentage of rape incidents**

**4.2 Region-wise crime percentage**

Fig.6 is a stacked bar chart; each bar represents a region of India. It depicts six dangerous crimes, cruelty by husband or relatives, kidnapping and abduction, rape, assault, dowry deaths, and the insult to the modesty of women. It can be seen that the northern region of India is the high crime intensity region among the six regions. The number of crimes registered under cruelty by husband/relatives, kidnapping and abduction, rape, assault on women, and dowry deaths is high in the northern region of India. The number of crimes registered under insult to the modesty of women is higher in southern regions of India than in other regions.



**Fig.6 Region-wise crimes**

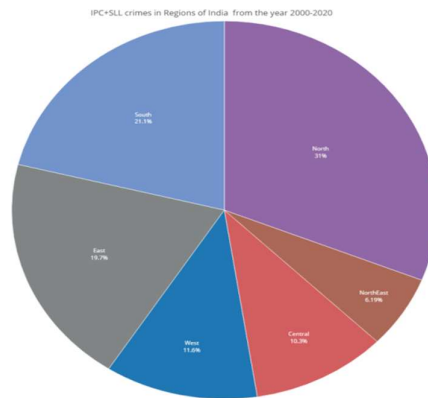
There were nearly 50 lakh cases reported under the crime against women category from the year 2000-2020. This is the sum of IPC and SLL crimes registered in different regions of India. The percentage distribution of crimes among regions of India is analyzed. 31% of the



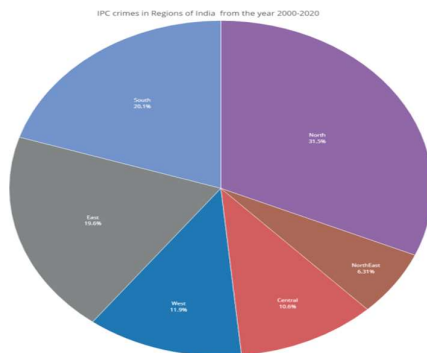
total crimes are registered in the northern region of India, followed by the southern region-21.1%, eastern region-19.7%, western region-11.6%, central region-10.3%, and north-east region-6.19%. Under IPC offenses, 31.5% of the total crimes are registered in the northern region of India itself, followed by the southern region- 20.1%, eastern region-19.6%, western region-11.9%, central region-10.6%, and north-east region-6.31%. Under SLL offenses, 48.7% of the total crimes are registered in the southern region of India, followed by the eastern region-23.7%, northern region-18.5%, western region-4.35%, north-east region-3.18%, and central region-1.57%.

**Table16. Percentage distribution of crimes registered in different regions of India**

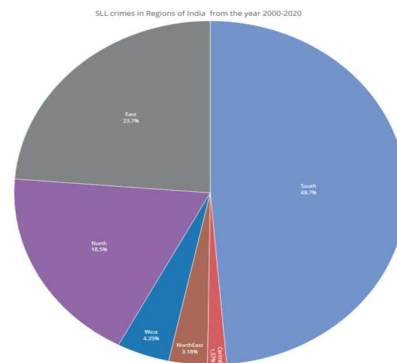
Crime group / Region	Central	East	North	North-East	South	West
IPC crimes	10.6%	19.6%	<b>*31.5%</b>	6.31%	20.1%	11.9%
SLL crimes	1.57%	23.7%	18.5%	3.18%	<b>*48.7%</b>	4.35%
IPC+SLL	10.3%	19.7%	<b>*31.0%</b>	6.19%	21.1%	11.6%



**Fig. 7(a). Percentage of IPC+SLL crimes in regions of India**



**Fig.7 (b). Percentage of IPC crimes in regions of India**



**Fig.7(c). Percentage of SLL crimes in regions of India**

### 4.3 Factors that influence crime against women

Crime is the result of social, environmental, and cultural causes. Crime in India has been on the rise of over the past several years. In particular, in recent years, it has been increasingly perceived as unsafe for women due to crime against women. There are socio-economic factors i.e. poverty rate, illiteracy, unemployment, demographic factor i.e. population and other environmental factors i.e. family environment, childhood life, alcohol and



drugs, temperature, and air pollution that trigger an increase in crime and violent activities against women. Already many studies conducted previously by various authors to find out the relationship between crime and socioeconomic factors. In this part of this paper, we have discussed the relationship between the above-said factors on crime against women in India.

Here, correlation analysis is used to find the relationship between crime and other socioeconomic and demographic metrics. Correlation analysis calculates the level of change in one metric due to the change in the other metric. Correlation between two metrics could be positive correlation or negative correlation or neutral correlation. The Correlation coefficient is comprised between -1 and 1. -1 or closer indicates there is a strong negative correlation between the metrics, 0 indicates no correlation between metrics and 1 or closer indicates a strong positive correlation between metrics.

#### 4.3.1 Population and crime

The Population growth of India increased dramatically since the beginning of the 20th century. India is predicted to become the most populated country in the world by 2030. On the other hand, crime against women in India had shown an upward trend since 2001. As per the population projections report prepared by the NCP (National Commission on Population) for the period 2011 – 2036, the population of India is expected to increase from 121.1 crores to 152.2 crores during the period 2011-2036 an increase of 25.7 percent in twenty- five years at the rate of 1.0 percent annually. As a consequence, the density of the population will increase from 368 to 463 persons per square kilometer. Since we want to know the impact of the population on the crime rate against women in India, we will look at the population changes and crime incidents in all states and union territories of India from 2001 to 2020.

**Table8. The correlation coefficient between population and crime**

Demographic Area	Correlation coefficient	Metrics
All India	0.97	Year-wise total population and total crime (2001-2020)
States and union territories	$\geq 0.82$ (all states and UTs)	State/UT & Year-wise total population and total crime (2001-2020)

Table.3 presents the correlation coefficient between India's population and crime incidents against women. Relationship between all India population and total crime incidents, relationship between the state/UT population and total crime incidents in state/UT are statistically significant since coefficient (0.97 and 0.82) values are closer to one. Hence these metrics are strongly correlated.

#### 4.3.2 Alcohol/drugs and crime

Alcohol and drug have the highest relationship with aggressive crimes. Addicts are forced to commit a crime to acquire drugs. An allegedly addicted alcoholic husband used to beat his wife and children almost daily. Sometimes wife/child is beaten to death because of the aggression of an alcoholic. Alcohol is the most common substance used in cases of 'date rape' and is often a factor involved in sexual assault and rape, increasing both the likelihood of committing sexual violence as well as becoming a sexual assault victim. Table.9 presents the correlation coefficient (0.97) between total crimes against women and total drug and alcohol

cases registered. And this value indicates a strong correlation between them, on the second row of the table presents the correlation coefficient between total cases registered under cruelty by a husband or his relatives with total drug and alcohol cases. It has a correlation coefficient value of  $\geq 0.57$  for all states and union territories. It indicates a positive correlation and the relationship is statistically significant.

**Table 9 The correlation coefficient between drug/alcohol and crime**

Demographic Area	Correlation coefficient	Metrics
All India	0.97	Year-wise total NDPS & Alcohol cases registered and total crime (2001-2020)
States and union territories	$\geq 0.57$ (all states and UTs)	State/UT & Year wise total NDPS & Alcohol cases registered and cruelty by husband or his relatives (2001-2020)

#### 4.3.3 Internet usage and crime

Though there were statistically significant positive correlations between the number of internet users and sexual crime rates the number of internet users is not the sole cause of sexual crimes. The advent of the ‘internet’ and ‘the world wide web’ in India has significantly influenced pornography distribution and access. Misconceptions about pornography and its adverse effects can cause the development of wrong concepts about sexuality. Table.10 presents the correlation between the metrics of the total number of internet users and total rape incidents and total internet users and cyber crime against women. These two relationships are statistically significant and strongly correlated.

**Table10. The correlation coefficient between internet usage and crime**

Demographic Area	Correlation coefficient	Metrics
All India	0.75	Year-wise total internet users and total rape incidents (2001-2020)
All India	0.80	Year-wise total internet users and cybercrime (2016-2020)

#### 4.3.4 Unemployment and crime

Table.11 illustrates the relationships between unemployment and crime against women. The unemployment effect on total crimes in India, the unemployment effect on total cases registered under kidnap and abduction, and the unemployment of states and UTs effect on total crime of states and UTs are positively correlated and statistically significant. Unemployment tends to propel persons into crime which enable them to survive economically, and also that unemployment tends to increase the anomie among the unemployed that is related to criminal behavior.

**Table11. The correlation coefficient between unemployment and crime**

Demographic Area	Correlation coefficient	Metrics
------------------	-------------------------	---------

All India	0.79	Year-wise total unemployment and total crime (2001-2020)
All India	0.57	Year-wise total unemployment and total kidnap & abduction cases (2001-2020)
States and union territories	$\geq 0.80$ (all states and UTs)	State/UT & Year-wise total unemployment and crime (2001-2020)

#### 4.3.5 Poverty and crime

Poverty is caused by lack of employment which leads to children resorting to crime at an early age as a source of income. India's national MPI (Multidimensional Poverty Index) captures multiple and simultaneous deprivations faced by households across the three macro dimensions of health, education, and living standards. Table.12 presents the relationship between the MPI of states/UTs and kidnapping and abduction crimes in India. The impact of MPI on kidnapping and abduction crimes is moderately positive and statistically significant but not the sole cause of such crimes.

**Table12. The correlation coefficient between MPI and crime**

Demographic Area	Correlation coefficient	Metrics
States and union territories	0.58	Poverty index (2019-2021) and kidnapping & abduction (2019-2021)

#### 4.3.6 Illiteracy and crime

Illiteracy has a direct impact on the high unemployment rate in India, unemployment causes poverty, and poverty leads to people committing the crime as a source of income. Table.13 depicts the correlation between illiteracy and a total crime against women in India. Though this relationship is statistically non-significant somehow illiteracy as a social factor indirectly affects the crime rate in India.

**Table13. The correlation coefficient between illiteracy and crime**

Demographic Area	Correlation coefficient	Metrics
States and union territories	$< 0.02$	Illiteracy rate (2019-2021) and crime incidents (2019-2021)

#### 4.3.7 Police strength and crime

Table.14 depicts the relationship between the metrics of total police strength of states/UTs and total crime incidents. In this scenario, the linear correlation between the strength of police and total crime should be negatively correlated as crime rates should decrease as police strength increases. Since it is a positive value the police strength does not affect the crime rate but how police are used is more important in crime deterrence.

**Table14. The correlation coefficient between police strength and crime**

Demographic Area	Correlation coefficient	Metrics
------------------	-------------------------	---------

States and union territories	> 0.75	Total police strength (2001-2020) and total crime incidents (2001-2020)
------------------------------	--------	---

#### 4.3.8 Ambient temperature and crime

Hot temperatures increase aggression by directly increasing feelings of hostility and indirectly increasing aggressive thoughts. Here the relationship between average annual temperatures (in °C) of all states and UTs of India and the annual total crime of states and UTs of India is measured. Table.15 depicts that the relationship is positive, statistically significant, and moderately correlated.

**Table15. The correlation coefficient between ambient temperature and crime**

Demographic Area	Correlation coefficient	Metrics
States and union territories	> 0.55	The temperature in °C (2001-2020) and total crime incidents (2001-2020)

These are some common factors that are discussed in many studies and under deep research to find out the truth behind the increasing trend in crime in India.

#### CONCLUSION

The main objectives of this research work are to find geospatial patterns of crimes committed against women in India and to find the relationship between demographic, socioeconomic, and environmental factors and crime rates in India. For the purpose of finding geospatial patterns, the KM-GOA technique is applied to crime datasets from the year 2000-2020 acquired from NCRB India. Firstly, India states/UTs are clustered into three groups according to the average crime incidents from the year 2000-2020 to find out the high crime intensity, mid-crime intensity, and low crime intensity areas. Among three clusters, the one has highest value of average crime incidents is noted as a crime hotspot. Secondly, the same clustering approach is followed to find high crime intensity areas for each crime type i.e. rape, kidnap & abduction, cruelty by husband, assault on women and insult to the modesty of women. In both the cases, northern states of India reported high number of crimes against women than other states. Nearly 31% of total crimes are committed in northern states of India and 22% of crimes are committed in southern states of India. Indian Territory is divided into six regions namely, central, east, north, north-east, south, and west India. Crimes committed in these six regions were analyzed. North India region takes first place as the total number of crimes registered under the crime against women act, particularly violent crimes against women i.e. rape, kidnap & abduction, cruelty, and assault. In south India, the highest number of crimes registered was under the insult to the modesty of women act.

It was observed from the data that the population, alcohol/drugs, unemployment, internet usage and crime rate in India has significant relationship. These variables directly affect the rate of crime committed against women in India. It was observed that poverty and ambient temperature are positively correlated to crime rate and the relationship is modest. The illiteracy rate and police strength depict a negative relationship with the crime rate. Both of them do not affect crime rate.

However, apart from these, some other factors must be considered. They are family

environment, conviction rate, and pendency in police cases. To reduce the crime rate in India, we must focus on controlling population density, eradicating alcohol and drugs, improving employment, promoting education level, proper use of the internet, decreasing poverty rate, better climate control and effectively using police strength.

#### **FUTURE SCOPE**

This work could be extended to other factors like the ratio of the female and male population in states, gender-wise illiteracy rate, the male and female ratio in alcohol and drug consumption, family environment, conviction rate, and pendency in police cases.

#### **LIMITATIONS**

The data used in this work is mainly dependent on the data published by the National Crime Records Bureau, the census of India, and the National Family Health Survey.

#### **ACKNOWLEDGEMENT**

We would like to thank the officials in the Ministry of Home Affairs, police, lawyers, Statistical Branch of National Crime Records Bureau, for sharing meaningful data on NCRB official website that helped us in conducting the research.

#### **REFERENCES**

- [1]. Leonard D.Eron, L.Rowell Huesmann, Monroe M. and Leopold O (1972), Does Television Violence Cause Aggression?, American Psychologist.
- [2]. Steven Raphael and Rudolf Winter-Ebmer (2001), Identifying the Effect of Unemployment on Crime, Journal of Law and Economics, vol. XLIV.
- [3]. Craig A. Anderson (2001), Heat and Violence. Current Directions in Psychological Science, Volume-10, Number-1. American Psychological Society. The University of Chicago.
- [4]. Darshan Kumar Jha (2015), Geography of Rape Crime in India: A Spatial Analysis of Official Data, International Journal of Research in Social Sciences, Volume 5, Issue 5, ISSN: 2249-2496.
- [5]. Lawrence McClendon and Natarajan Meghanathan (2015), Using Machine Learning Algorithms to Analyze Crime Data, Machine Learning and Applications: An International Journal, Vol.2, No.1.
- [6]. Hossain S, Abtahee A, Kashem I, Hoque M.M and Sarker I.H. Crime Prediction Using Spatio-Temporal Data. Computing Science, Communication and Security. COMS2 (2020). CCIS, vol 1235. Springer, Singapore. [https://doi.org/10.1007/978-981-15-6648-6\\_22](https://doi.org/10.1007/978-981-15-6648-6_22).
- [7]. Anneleen Rummens, Wim Hardyns and Lieven Pauwels (2017), The use of predictive analysis in spatiotemporal crime forecasting: Building and testing a model in an urban context, Applied Geography, <http://dx.doi.org/10.1016/j.apgeog.2017.06.011>.
- [8]. James J. Nolan (2004), Establishing the statistical relationship between population size and UCR crime rate: Its impact and implications, Journal of Criminal Justice 32, 547–555.
- [9]. Keivan Kianmehr and Reda Alhajj (2006), Crime Hot-Spots Prediction Using Support Vector Machine, IEEE.
- [10]. Kounadi, O, Alina Ristea, Adelson Araujo and Michael Leitner. A systematic review on spatial crime forecasting. Crime Sci 9, 7 (2020). <https://doi.org/10.1186/s40163-020-00116-7>.

- [11]. John Braithwaite (1975), Population Growth and Crime, Australian and New Zealand Journal of Criminology. DOI: 10.1177/000486587500800107.
- [12]. Faryad Parhiz, Asghar Zarrabi, Abolfazl Meshkini (2017), Analyzing the Relationship between Crime and Population Intensity (12th Municipality Area of Tehran as a Case Study), European Online Journal of Natural and Social Sciences, Vol. 6, No.1(s) Special Issue on Economic and Social Progress. ISSN 1805-3602.
- [13]. Kalantari, M., Ghezelbash, S., & Yaghmaei, B. (2010). Geographical Investigation of Crime HotSpots in Zanjan city (Case Study: Drug-Related Crimes). Human Geography Research Quarterly, 74.
- [14]. Joel Mathew Jose, Crime Is A Result of Social and Environmental Factors and Not On Biological Factors (2021), International Journal Of Current Advanced Research, ISSN: O: 2319-6475, ISSN: P: 2319-6505, volume 10; issue 02. Page no. 23939-23941. DOI: <http://dx.doi.org/10.24327/ijcar.2021.23940.4743>.
- [15]. Nasim Ahamed Mondal, Mithun Mog and Kacho Amir Khan, Alcohol Consumption in India: An Analysis of IHDS Data (2018), American Research Journal of Humanities and Social Sciences. ISSN (Online): 2378-7031, Volume 4, 15 pages.
- [16]. P.Thivakar (2019), A Study On Alcohol Consumption In India, International Journal of Research and Analytical Reviews, Volume 6, Issue 2, E-ISSN 2348-1269, P- ISSN 2349-5138.
- [17]. Vania Ceccato (2014), The nature of rape places, Journal of Environmental Psychology. <http://dx.doi.org/10.1016/j.jenvp.2014.05.006>.
- [18]. R. Karl Hanson and Kelly E. Morton-Bourgon (2006), The Characteristics of Persistent Sexual Offenders: A Meta-Analysis of Recidivism Studies, journal of consulting and clinical psychology, Vol. 73, No. 6, 1154–1163, DOI: 10.1037/0022-006X.73.6.1154.
- [19]. Michael Flood and Bob Pease (2009), Factors Influencing Attitudes to Violence Against Women, Trauma Violence & Abuse, Vol. 10, No. 2, April,125-142, DOI: 10.1177/1524838009334131.
- [20]. Jessica Saini and Vagisha Srivastava (2019), Impact of Population Density and Literacy Levels on Crime in India, IEEE – 45670.
- [21]. Suresh Bada Math, Biju Viswanath, Ami Sebastian Maroky, Naveen C. Kumar, Anish V. Cherian and Maria Christine Nirmala (2014), Sexual Crime in India: Is it Influenced by Pornography?, Indian Journal of Psychological Medicine, Apr – Jun, Vol 36, Issue 2, DOI:10.4103/0253-7176.130976.
- [22]. Basayya M Hosurmath (2018), Impact of Internet on Rising Crime Rates, International Journal of Creative Research Thoughts, Volume 6, Issue 1, February, ISSN: 2320-2882.
- [23]. Pankhuri Thukral and Vanshika Kainya (2022), How Social Media Influence Crimes, Indian Journal of Law and Legal Research, Volume 4, Issue 2, ISSN: 2582-8878.

#### **Author's profile**



**Mr.K.Dheenathayalan**, received B.Sc degree in Computer Science, M.C.A and M.Phil degree in Computer Science from Bharathiar University, Coimbatore, Tamil Nadu in 2008, 2011 and 2014 respectively. He is currently working as Assistant Professor in Department of Computer Science, Kamban College of Arts and Science, Coimbatore. He has 7 years of experience as Assistant Professor. He is currently pursuing Ph.D in Computer Science in Bharathiar University. His current research interests include Data Mining and Bio-informatics.



**Dr.K.K.Savitha** received B.Sc, MCA and M.Phil degrees. She received Ph.D in Computer Science from Anna University, Chennai, in 2013. She is currently working as an Assistant Professor in the Department of Computer Applications, Bharthiar University PG Extension Centre, Erode. Her research interest includes mobile computing, Cloud Computing and Soft Computing. She is a Life member of ISTE, CSI.