International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 10 Issue 2, Mar-Apr 2026 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Geographical Variability and Socio-Demographic Determinants of Crimes Against Women: Evidence from West Bengal (2017-2022)

Prosenjit Murmu[1]*, Sonali Sahoo[2], Indrajit Murmu[3], Bijoy Hansda[4]

[1]Assistant Professor, Department of Geography, Egra S.S.B. College, Purba Medinipur, West Bengal, India
[2]Independent Researcher, Egra, Purba Medinipur, West Bengal, India
[3]Department of Education, Rabindra Bharati University (LSC), Bankura, West Bengal, India
[4]Department of Geography, Ranchi University, Jharkhand, India
*Corresponding Author: Prosenjit Murmu

ABSTRACT

This study examines the socio-demographic factors and spatial-temporal dynamics of crimes against women (CAW) in West Bengal from 2017-2022. A multi-stage analytical framework is used, incorporating deviation analysis, standardized residuals, chi-square contributions, Cramér's V, and Spearman's rank correlation with regression, using district-level data from the National Crime Records Bureau (NCRB) and socio-economic indicators from various sources. The findings show that crime incidence is highly heterogeneous with notable temporal and spatial fluctuations. According to the Pareto principle, districts such as South 24 Parganas, Jalpaiguri, and North 24 Parganas consistently show up as hotspots, collectively contributing more than 80% of the chi-square value. Sharp oscillations in some districts are highlighted by standardized residuals, suggesting erratic crime dynamics. While female literacy and unemployment have little explanatory power, correlation-regression analysis reveals a significant correlation between higher crime rates and female education beyond the tenth grade, multi-dimensional poverty, urbanization, and sex ratio. While underreporting probably hides vulnerability in poorer, less literate areas, higher crime rates in urbanized and better-educated districts may partially reflect increased awareness and reporting of these crimes. The results highlight how education, gender parity, poverty, and urbanization interact to shape CAW. The policy implications focus on geographically focused interventions in hotspot districts to enhance gender parity, increase access to education, reduce poverty and fortify reporting and justice systems. Underreporting biases and temporal mismatches in data sources are examples of these limitations. To improve comprehension and policy responsiveness, future research should include spatial clustering, qualitative case studies and predictive modelling.

How to cite this paper: Prosenjit Murmu | Sonali Sahoo | Indrajit Murmu | Bijoy Hansda "Geographical Variability and Socio-Demographic Determinants of Crimes Against Women: Evidence from West Bengal (2017-2022)" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-10 | Issue-2, April 2026, pp.784-797, URL: www.ijtsrd.com/papers/ijtsrd101472.pdf

IJTSRD101472

KEYWORDS: *Crimes Against Women, Spatial-Temporal Dynamics, Socio-Demographic Correlates, West Bengal.*

1. INTRODUCTION

Crimes against women (CAW) represent a persistent violation of human rights and a critical barrier to social development. India exhibits particularly concerning trends owing to its vast population and complex socio-cultural dynamics. CAW encompasses physical, sexual, psychological, and economic harm, manifesting in forms such as intimate partner violence, sexual assault, domestic cruelty, human trafficking, dowry-related offences, and emerging challenges such as cyber harassment (Bohra et al., 2015; Devakunchari et al., 2019; Mir, 2022). Globally, approximately one-third of women experience physical or sexual violence during their lifetime, underscoring the pervasive nature of this systemic issue, which intersects with social, economic, and gender inequalities and impedes progress toward Sustainable Development Goals, notably SDG-5 (gender equality) and SDG-16 (peace,

justice, and strong institutions) (Islam & Saikia, 2025).

India's extensive legislative framework, including constitutional provisions, the Indian Penal Code, the Protection of Women from Domestic Violence Act (2005), and specialized laws addressing sexual offences, dowry violence, and workplace harassment, has not been sufficient to curb the rising incidence of CAW (Pooja et al., 2024; Siddiqui, 2026; Saxena, 2025). The National Crime Records Bureau (NCRB) data reveal a near doubling of reported cases from 228,650 in 2011 to 445,256 in 2022, with marked spatial disparities, such as Delhi's rate of 145 per 100,000 women compared to Nagaland's 5 per 100,000 (Pooja et al., 2024). However, underreporting is endemic, with estimates indicating that up to 90% of molestation and rape cases remain unregistered due to stigma, distrust in law enforcement, and procedural barriers (Sharma et al., 2022; S. Gupta et al., 2022).

The existence and diversity of CAW in India are deeply rooted in ingrained patriarchal social structures that assign men dominant roles and enforce women's conformity to the ideals of obedience and sacrifice (Kumar, 2015; Siddiqui, 2026). Historical analyses trace the decline in women's status from relative equality in ancient India to institutionalized gender discrimination during medieval times, exemplified by practices such as Sati and child marriage (Kalaiyarasi, 2015). Socio-economic factors, including poverty, low female education, substance abuse, and inadequate institutional responses, compound the complexity of this issue (Pooja et al., 2024; Islam & Saikia, 2025). Existing empirical studies have predominantly focused on specific crime categories, such as rape and domestic violence, often neglecting forms such as molestation, abduction, and human trafficking, and frequently relying on national averages that obscure critical regional, urban-rural, and temporal variations (Gupta et al., 2022; Siddiqui, 2026; Kumari et al., 2025). Key knowledge gaps remain regarding the socio-economic determinants, psychological profiles of perpetrators, underreporting mechanisms, and effectiveness of legal interventions (Ghooi & Deshpande, 2013; Joseph et al., 2017; Kumari et al., 2025). For example, despite Kerala's economic development, violence against women persists, indicating that development alone does not fully explain CAW patterns (Joseph et al., 2017).

West Bengal exemplifies these complexities within the Indian context, contributing disproportionately to national crime statistics against women, despite constitutional and legislative protections. The state, comprising 7.5% of India's female population, recorded over 35,000 criminal cases against women, with the highest incidence of dowry-related crimes and human trafficking nationally (Chatterjee et al., 2023; Das, 2021). Districts such as Murshidabad, South 24 Parganas, North 24 Parganas, and Jalpaiguri exhibit persistently high crime concentrations, with rates escalating from below 9 per 100,000 female population in 1981 to over 131 in Murshidabad by 2011 (Prasad, 2018). Crimes in West Bengal encompass physical, psychological, sexual, verbal, and economic dimensions and occur within domestic, workplace, and public spheres (Banerjee, 2017; Das, 2021). Despite relatively lower overall crime rates than the national averages, socio-legal barriers, including social stigma, family preservation priorities, and women's historical exclusion from political participation, hinder effective reporting and access to justice, resulting in pervasive underreporting (Dhawan & Bhasin, 2024; Das, 2021).

This present study investigates the geographical variability and socio-demographic determinants of CAW in West Bengal by testing hypotheses on spatial distribution, temporal volatility, and socio-demographic correlates at the district level. It employs a multi-stage analytical framework that integrates descriptive, inferential, and comparative techniques. Deviation and standardized residual analyses highlighted districts with excess or deficit reporting, whereas chi-square contributions ranked their relative importance. In addition, Spearman's rank correlation and regression were used to assess the association between crime rates and socio-demographic variables. This structured and reproducible methodology ensures rigor and transparency by addressing spatial and temporal dimensions to inform evidence-based and geographically targeted policy interventions.

2. Hypotheses:

$H0_1$: CAW in West Bengal is uniformly and proportionally distributed across districts, with the observed deviations attributable to random variation.

$H1_1$: CAW exhibits significant spatial variation and temporal volatility, with certain districts disproportionately contributing to state-level variation.

$H0_2$: District-level variations in CAW do not deviate significantly from the statistically expected confidence limits (±2, ±3, ±4), and no localized anomalies are present.

$H1_2$:District-level variations in CAW generally conform to the expected confidence limits (±2, ±3,

±4), but significant localized deviations can be detected through residual analysis.

$H0_3$: All districts contribute equally to the overall chi-square statistic and crime burden.

$H1_3$: A small cluster of districts (Pareto concentration) accounts for over 80% of the statistical variation, enabling geographically targeted intervention.

$H0_4$: No statistically significant correlation exists between CAW and socio-demographic variables.

$H1_4$: CAW is positively associated with socio-demographic variables.

3. Objectives:

1. To analyze and quantify the extent of district-wise deviations in C A W from statistically expected values in West Bengal (2017-2022), thereby testing spatial heterogeneity and temporal volatility.
2. To examine the temporal volatility of crime patterns across districts in West Bengal (2017–2022) and identify statistically significant localized anomalies through residual analysis.
3. To rank districts by their contribution to the overall variation through chi-square analysis and identify concentrated crime hotspots, validating the Pareto concentration hypothesis.
4. To determine the multi-factorial socio-demographic determinants of CAW by distinguishing between significantly correlated and non-significant variables to develop a nuanced socio-economic risk profile that enables targeted policy interventions.

4. Methodology

4.1. Study Design and Data Sources

This study utilized a multi-stage analytical framework that integrated descriptive, inferential, and comparative methodologies to investigate spatial and temporal variations in CAW across districts in West Bengal. It further examines the relationship between CAW and selected socio-demographic variables, elucidating the interaction between crime incidence and broader structural conditions. The methodology incorporates deviation analysis to identify systematic district-level reporting anomalies, residual analysis to detect statistically significant deviations, chi-square contribution analysis to quantify relative district importance, Cramér's V coefficient to evaluate the strength of association between categorical crime distributions and district characteristics, and correlation-regression analysis to explore associations with socio-demographic indicators. This structured approach ensures transparency, reproducibility, and adherence to scientific standards while acknowledging the inherent limitations of analyses.

The study employs secondary data on CAW, as well as socio-economic and demographic indicators, obtained from multiple reliable sources. District-specific CAW incidents for the period 2017–2022 were acquired from the National Crime Records Bureau (NCRB), Government of India. To provide context for these data, the projected female population for 2021 was estimated using population data from the Census of India (2001 and 2011). Demographic and educational indicators were sourced from the National Family Health Survey (NFHS 5, 2019–21), including the sex ratio (number of females per 1000 males), female literacy rate, and proportion of females with education beyond the 10th grade. The head count ratio (of multi-dimensional poverty) for 2015–16 and the projected urbanization levels for 2021 were derived from Pandit and Halder (2023). Additionally, the unemployment rate for 2015, measured as the number of unemployed persons per 1000 population, was obtained from the Labour & Employment Statistics database, Bureau of Applied Economics and Statistics, Government of West Bengal. Together, these data sources enabled a multi-dimensional analysis of CAW across districts in West Bengal.

4.2. Data Processing and Normalization

Crime rates (CR_{ij}) were calculated per 100,000 female population using the following formula:

$$CR_{ij} = \frac{CAW_{ij}}{FPOP_{ij}} \times 100,000$$

Where CAW_{ij} denotes the incidence of crimes against women in year i of district j, and (FPOPij) represents the female population.

To ensure comparability across heterogeneous indicators with different scales, all socio-demographic variables were normalized using min-max scaling.

$$X'_{ij} = \frac{X_{ij} - X_{i(min)}}{X_{i(max)} - X_{i(min)}}$$

Where X'_{ij} denotes the normalized crime incidence for the i^{th} year in the j^{th} district, X_{ij} represents the actual crime incidence value, and $X_i(min)$ and $X_i(max)$ represent the minimum and maximum values of crime incidence for the i^{th} year across all districts, respectively. This transformation produces values between 0 and 1, enabling a meaningful comparative analysis. Besides, descriptive statistics (Mean, SD, CV) were computed to assess central tendency, variability, and distributional asymmetry before proceeding to statistical analysis.

4.3. Statistical Analyses

4.3.1. Deviation Analysis

To identify districts with systematic excess or deficit crime reporting, expected frequencies were derived from two-way contingency tables (districts × years) using the following equation:

$$E_{ij} - \frac{O_i \times O_j}{O_n}$$

Where O_i denotes the row total (district cases across years), O_j represents the column total (year cases across districts), and O_n denotes the grand total. The percentage deviation from the expected frequency for each district-year cell was calculated as follows:

$$D_{ij} = \frac{O_{ij} - E_{ij}}{E_{ij}} \times 100$$

This allowed us to identify districts reporting either excess or deficit cases relative to their expectations. To capture long-term trends, we computed the six-year mean deviation, highlighting structural imbalances in district-level reporting patterns.

4.3.2. Residual Analysis

Standardized residuals (r_{ij}) were derived using the following conventional formula:

$$r_{ij} - \frac{O_{ij} - E_{ij}}{\sqrt{E_{ij}}}$$

This adjustment measures deviations relative to the expected frequencies, thereby permitting meaningful comparisons across districts with different case volumes. Statistical significance was assessed using standardized residual thresholds, where values exceeding ±2, ±3, and ±4 were considered significant at the 95%, 99%, and 99.99% confidence levels, respectively. In this case, positive standardized residuals indicate districts that consistently recorded more cases than expected, whereas negative values indicate under-representation.

4.3.3. Chi-Square Contributions Analysis

District-wise chi-squared (χ^2_{ij}) contributions were computed to quantify the relative importance of each district in shaping the overall variance.

$$\chi^2_{ij} = \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Additionally, Cramér's V coefficient was used to assess the strength of the association between the categorical crime distributions and district characteristics.

$$V = \sqrt{\frac{X^2}{n(k-1)}}$$

Where, X^2 is the total chi-square value, n is the total number of observations, and k is the smaller number between the number of rows or columns.

The values of V were interpreted using specific thresholds, where 0.0-0.1 indicated a negligible or very weak association, 0.0-10.3 reflected a weak or small association, 0.3-0.5 denoted a moderate or medium association, and values greater than 0.5 represented a strong or large association. This measure provides a concise quantification of the relationship between district characteristics and crime incidence, complementing the chi-square analysis with an interpretable index of the strength of association.

4.3.4. Correlation and Regression Analysis

Spearman's rank correlation coefficient (ρ) was employed to explore the associations between crime rates and socio-demographic variables, chosen for its ability to accommodate non-normal distributions. The coefficient is calculated as follows:

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

The Statistical significance of the correlations was evaluated using the following test statistic:

$$t = \rho \sqrt{\frac{n-2}{1-\rho^2}}$$

Subsequently, variables that demonstrated significant correlations (p<0.05) were further examined through least-squares regression to estimate explanatory power (adjusted R^2 values).

5. Limitations

The current study integrated district-level CAW data with socio-demographic indicators; however, certain limitations arise from the temporal and structural characteristics of the datasets. First, owing to the lack of contemporary socio-demographic data, some indicators are older than the crime data. For instance, unemployment data are restricted to 2015, whereas crime data span 2017–2022, creating a temporal imbalance that may affect comparability. Second, the female projected population for 2021 was estimated using decadal growth rates from the 2001 and 2011 censuses, applying the compound annual growth rate (CAGR) method. Although robust, this projection may not fully capture recent demographic shifts. Third, indicators such as the head count ratio of multi-dimensional poverty (HCR-MDP) and urbanization (URB) were derived from NFHS-4 (2015–16) but published in secondary sources (Pandit & Halder, 2023), introducing an additional layer of estimation. Although the National Family Health

Survey (NFHS)-5 provides reliable measures of the sex ratio, literacy, and schooling, reliance on different survey rounds for separate indicators may introduce such inconsistencies. However, these limitations do not undermine the validity of the analysis but highlight the need for cautious interpretation and acknowledgment of data heterogeneity across time and data sources.

6. Results:

6.1. Spatial and Temporal Patterns of Crime Against Women:

6.1.1. District-Level Variation and Percentage Deviation:

The spatial-temporal pattern of CAW in West Bengal (table 1 and figure 1) reveals pronounced heterogeneity across districts, highlighting alternating clusters of high positive deviation (e.g., Dakshin Dinajpur in 2017 and 2021, Jalpaiguri in 2020–2022, and Malda in 2018-2019) alongside persistent negative deviations in districts such as Bankura, Bardhaman, and Nadia. This spatial volatility is supported by the descriptive statistics, where the standard deviation remained consistently high (10.83%-16.65%), and the coefficients of variation exceeded 25% in 2017 and 2020, indicating strong inter-district disparities. The bar chart (figure 1-G) emphasizes that while most districts oscillated around the baseline with marginal mean deviations (close to zero), Kolkata and South 24 Parganas exhibited sustained positive deviations, and Jalpaiguri, Dakshin Dinajpur, Purba Medinipur, and Nadia districts showed negative deviations, underscoring localized vulnerability. Importantly, the rebound of positive deviations in 2021-2022 following widespread negative deviations in 2018-2019 suggests a possible post-pandemic resurgence in reported cases. Collectively, these findings demonstrate that CAW in West Bengal is not uniformly distributed but is characterized by sharp spatial contrasts and temporal shifts.

Table 1. Descriptive Statistic and Percentage Deviation of CAW (2017–2022)

Districts	2017	2018	2019	2020	2021	2022	Mean
Bankura	-4.58	-11.28	-11.28	-1.90	7.27	18.31	-0.58
Bardhaman	-13.20	-9.48	-9.48	16.52	6.40	4.40	-0.81
Birbhum	-3.14	6.18	6.18	-6.03	-0.72	-0.93	0.26
Cooch Behar	10.92	-14.57	-14.57	-7.93	11.76	11.94	-0.41
Dakshin Dinajpur	18.92	-39.92	-39.92	0.03	36.40	15.35	-1.52
Darjeeling	0.07	6.69	6.69	-6.82	-4.00	-0.47	0.36
Hooghly	12.32	3.16	3.16	-8.23	-5.09	-2.60	0.45
Howrah	13.77	-0.67	-0.67	-3.36	-2.37	-5.13	0.26
Jalpaiguri	-24.98	-28.14	-28.14	23.97	20.80	24.84	-1.94
Kolkata	6.33	18.89	18.89	-9.69	-18.17	-9.76	1.08
Malda	-33.04	23.32	23.32	-12.89	-8.49	10.99	0.53
Murshidabad	23.75	2.51	2.51	-1.80	-13.02	-10.23	0.62
Nadia	-12.49	-15.41	-15.41	18.47	13.64	4.61	-1.10
North 24 Parganas	25.37	-14.09	-14.09	1.21	1.30	-0.60	-0.15
Paschim Medinipur	-1.08	-11.07	-11.07	-0.30	33.95	-14.42	-0.66
Purba Medinipur	-16.12	-17.73	-17.73	19.22	10.83	14.03	-1.25
Purulia	19.19	7.37	7.37	-7.54	-9.53	-12.25	0.77
South 24 Parganas	-16.88	27.31	27.31	-10.93	-14.26	-6.51	1.01
Uttar Dinajpur	5.73	-9.00	-9.00	5.51	2.60	2.15	-0.33
Mean	0.57	-4.00	-4.00	0.39	3.65	2.30	-0.18
SD	16.36	16.65	16.65	10.89	14.72	10.83	0.85
CV	28.61	4.17	4.17	27.63	4.04	4.71	4.76

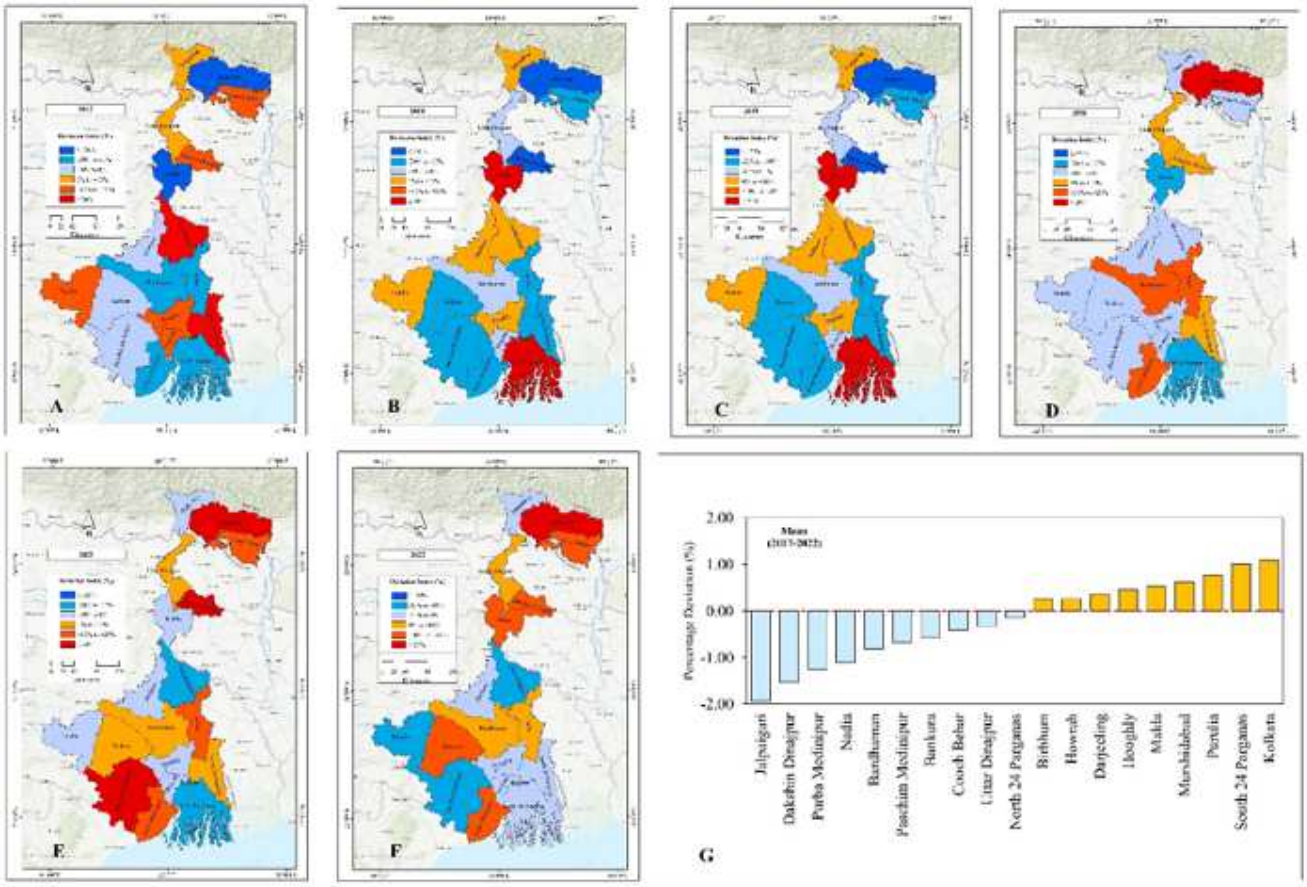


Figure 1: (A-F) shows the spatial distribution of the percentage deviation in CAW in West Bengal for 2017-2022, and (G) shows the six-year (2017-2022) mean deviation.

6.1.2. Identification of Significant Localized Anomalies:

The standardized residuals (r_i) (table 2 and figure 2) provide a robust measure of the statistical significance of deviations in CAW across districts. It reveals distinct spatial clustering of extreme residuals, and districts such as Jalpaiguri, Malda, and South 24 Parganas consistently exhibit highly significant deviations (both positive and negative), often crossing the 95–99.9% confidence thresholds. For instance, Jalpaiguri recorded strongly negative residuals in 2017–2019 ($r_i \leq -10.70$, 99.9% confidence), followed by sharp positive residuals in 2020–2022 ($r_i \geq +10.10$, 99.9% confidence), indicating a dramatic reversal in crime incidence. Similarly, South 24 Parganas displayed extreme oscillations with a large negative residual in 2017 ($r_i = -12.11$) and a highly positive spike in 2018 ($r_i = +19.41$), reflecting unstable crime dynamics. In contrast, districts such as Birbhum, Darjeeling, and Purulia maintained residuals within a non-significant range, suggesting relative stability around the expected values. The six-year mean residuals across districts were generally non-significant ($< \pm 2$), except for Jalpaiguri and South 24 Parganas. Jalpaiguri showed negative residuals (r_i = -2.31, 95% confidence), indicating a lower than expected incidence, whereas South 24 Parganas displayed positive residuals ($r_i = 2.23$, 95% confidence) and emerged as a statistically significant hotspot. These findings confirm that CAW in West Bengal is not only unevenly distributed but also statistically significant in certain districts.

Table 2. District-wise Standardized Residuals of CAW (2017–2022)

Districts	Standardized Residual (r_i)						
	2017	2018	2019	2020	2021	2022	Mean
Bankura	-0.97	-2.37	-2.37	-0.44	1.66	4.12	-0.38
Bardhaman	-5.79	-4.12	-4.12	7.87	3.02	2.04	-1.10
Birbhum	-0.75	1.45	1.45	-1.55	-0.18	-0.23	0.19
Cooch Behar	3.48	-4.60	-4.60	-2.74	4.04	4.03	-0.40
Dakshin Dinajpur	3.62	-7.57	-7.57	0.01	7.50	3.11	-0.90
Darjeeling	0.02	2.32	2.32	-2.59	-1.51	-0.17	0.39
Hooghly	4.61	1.17	1.17	-3.34	-2.05	-1.03	0.52
Howrah	5.32	-0.26	-0.26	-1.41	-0.99	-2.10	0.31

Jalpaiguri	-9.59	-10.70	-10.70	9.99	8.59	10.10	-2.31
Kolkata	2.75	8.13	8.13	-4.57	-8.50	-4.49	1.46
Malda	-12.54	8.77	8.77	-5.31	-3.47	4.42	0.63
Murshidabad	11.46	1.20	1.20	-0.94	-6.77	-5.23	0.93
Nadia	-5.44	-6.64	-6.64	8.72	6.38	2.12	-1.48
North 24 Parganas	17.01	-9.35	-9.35	0.88	0.94	-0.42	-0.30
Paschim Medinipur	-0.41	-4.20	-4.20	-0.13	14.00	-5.85	-0.79
Purba Medinipur	-5.98	-6.52	-6.52	7.74	4.32	5.51	-1.44
Purulia	3.64	1.39	1.39	-1.55	-1.95	-2.46	0.45
South 24 Parganas	-12.11	19.41	19.41	-8.51	-11.01	-4.95	2.23
Uttar Dinajpur	1.49	-2.31	-2.31	1.55	0.73	0.59	-0.27
Total	-0.18	-14.81	-14.81	3.66	14.77	9.11	-2.26

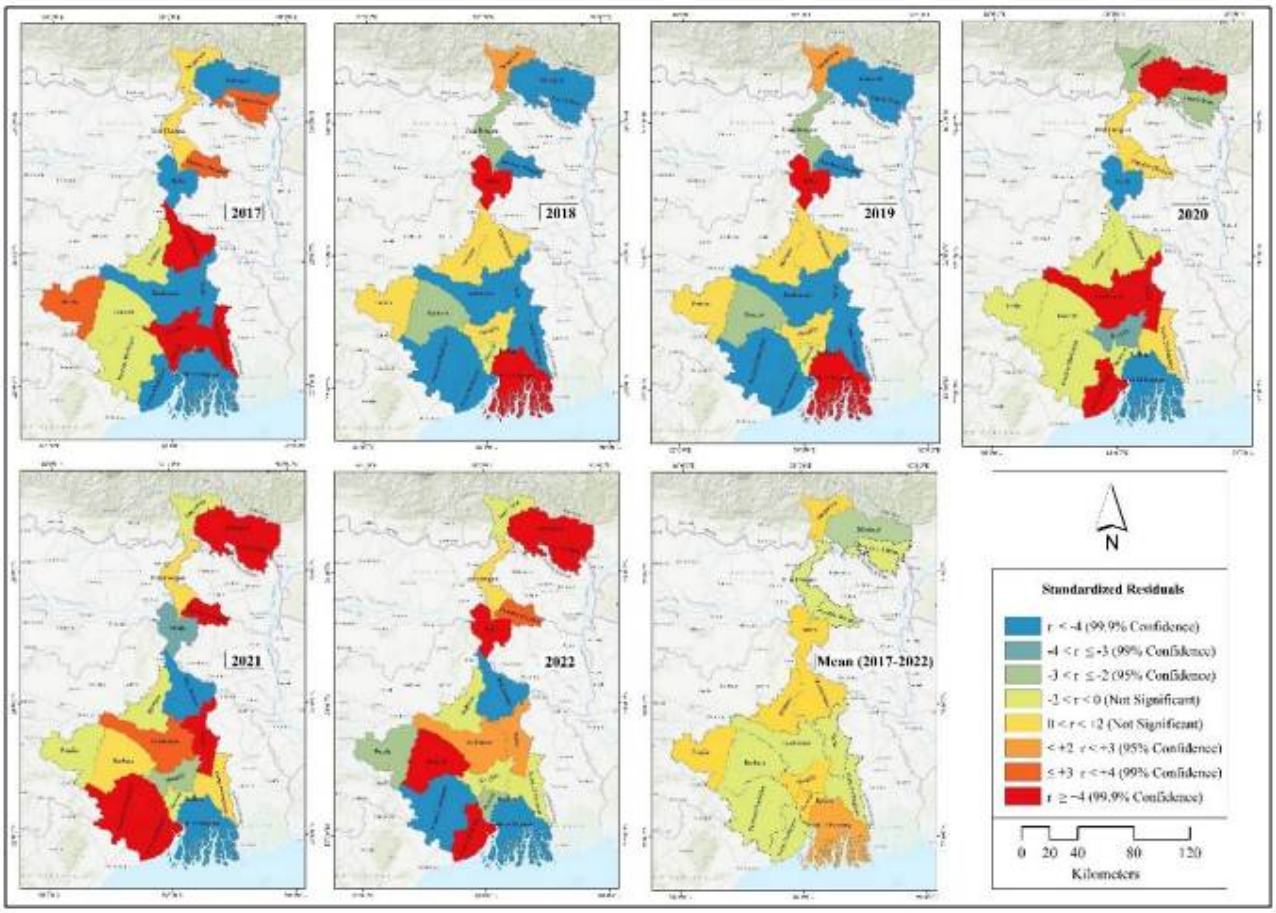


Figure 2: District-wise Localized Anomalies of crime against women in West Bengal, 2017–2022.

6.2. Quantification of District Contributions and Key Hotspots Identification:

6.2.1. Chi-Square Test of Spatial and Temporal Variation:

To understand the contribution of each district to the overall variation in CAW, district-level chi-square contributions were calculated, and the results are summarized in table 3.

Table 3. Chi-Square Contribution and District Ranking of CAW in West Bengal (2017-2022)

Districts	Chi-square (χ^2)							Rank	Remarks
	2017	2018	2019	2020	2021	2022	Total		
South 24 Parganas	146.61	376.56	376.56	72.47	121.26	24.47	1117.94	1	Very high contribution
Jalpaiguri	91.99	114.46	114.46	99.79	73.84	102.00	596.55	2	High contribution
North 24 Parganas	289.26	87.49	87.49	0.78	0.88	0.18	466.08	3	High contribution
Malda	157.36	76.85	76.85	28.19	12.04	19.52	370.81	4	Moderate contribution
Paschim Medinipur	0.17	17.65	17.65	0.02	196.00	34.22	265.71	5	Moderate contribution
Kolkata	7.58	66.16	66.16	20.91	72.24	20.17	253.23	6	Moderate contribution
Nadia	29.54	44.05	44.05	76.02	40.76	4.51	238.93	7	Moderate contribution

Purba Medinipur	35.80	42.46	42.46	59.94	18.69	30.37	229.72	8	Moderate contribution
Murshidabad	131.42	1.44	1.44	0.89	45.78	27.35	208.31	9	Moderate contribution
Dakshin Dinajpur	13.13	57.31	57.31	0.00	56.24	9.68	193.66	10	Low contribution
Bardhaman	33.56	16.98	16.98	61.92	9.13	4.17	142.74	11	Low contribution
Cooch Behar	12.12	21.18	21.18	7.53	16.29	16.25	94.57	12	Low contribution
Hooghly	21.23	1.37	1.37	11.17	4.20	1.06	40.39	13	Very low contribution
Howrah	28.30	0.07	0.07	1.99	0.97	4.41	35.80	14	Very low contribution
Bankura	0.95	5.64	5.64	0.19	2.76	16.98	32.16	15	Very low contribution
Purulia	13.26	1.92	1.92	2.41	3.79	6.05	29.35	16	Very low contribution
Darjeeling	0.00	5.39	5.39	6.73	2.27	0.03	19.81	17	Very low contribution
Uttar Dinajpur	2.21	5.34	5.34	2.41	0.53	0.35	16.19	18	Very low contribution
Birbhum	0.56	2.11	2.11	2.42	0.03	0.05	7.28	19	Very low contribution
Total	1015.05	944.44	944.44	455.77	677.71	321.83	4359.24		

The chi-square test confirmed substantial spatial and temporal variations in CAW across districts. The observed statistic ($\chi^2 = 4847.90$, d.f. = 90) was far greater than the critical value of 113.15 at the 5% significance level. This confirms that the distribution of CAW was not proportional across districts and years but was highly uneven in nature. The p-value (< 0.0001) reinforces the conclusion of strong statistical significance, demonstrating that variations in CAW is unlikely to be due to chance and reflect genuine differences in vulnerability at the district level.

In addition to the chi-square test, Cramér's V was applied to evaluate the strength of this association and calculated using the following formula:

$$V = \sqrt{\frac{4847.90}{198911 \times \min(19-1),(6-1)}} = 0.070$$

This effect size (0.070) indicates a weak but statistically significant association between the district-level distribution and crime incidence. In the setting of a very large sample (N = 198,911), the results suggest that while the variation is highly significant, the strength of the relationship remains modest. This means that CAW were not evenly distributed but were concentrated disproportionately in certain districts, with a few areas contributing heavily to the overall variation. These results highlight the presence of localized hotspots and underline the importance of district-specific factors in shaping vulnerability patterns.

6.2.2. Districts' Ranking by Chi-Square Contribution and Hotspots Identification:

The combined evidence from table 3 and figure 3 underscores the disproportionate influence of a few districts in shaping the overall chi-square distribution of CAW. South 24 Parganas alone contributed more than one-fourth of the total chi-square value ($\chi^2 = 1117.94$), marking it as the most statistically significant hotspot. Jalpaiguri ($\chi^2 = 596.55$) and North 24 Parganas ($\chi^2 = 466.08$) also emerged as high-contribution districts, accounting for nearly 60% of the cumulative chi-square statistic. The cumulative percentage curve in figure 3 demonstrates that the top eight districts (South 24 Parganas, Jalpaiguri, North 24 Parganas, Malda, Paschim Medinipur, Kolkata, Nadia, and Purba Medinipur) collectively explain over 80% of the total chi-square contribution, thereby meeting the Pareto threshold of concentration. In contrast, districts such as Birbhum, Uttar Dinajpur, Darjeeling, Purulia, Bankura, and Howrah contributed minimally, with values below 40, indicating a negligible influence on the overall pattern. This concentration effect highlights that CAW is statistically driven by a limited set of districts, whereas the majority of regions remain relatively stable. Temporal variation further revealed that South 24 Parganas dominated contributions in 2018–2019, whereas Paschim Medinipur and Jalpaiguri surged in 2020–2021, reflecting shifting hotspots.

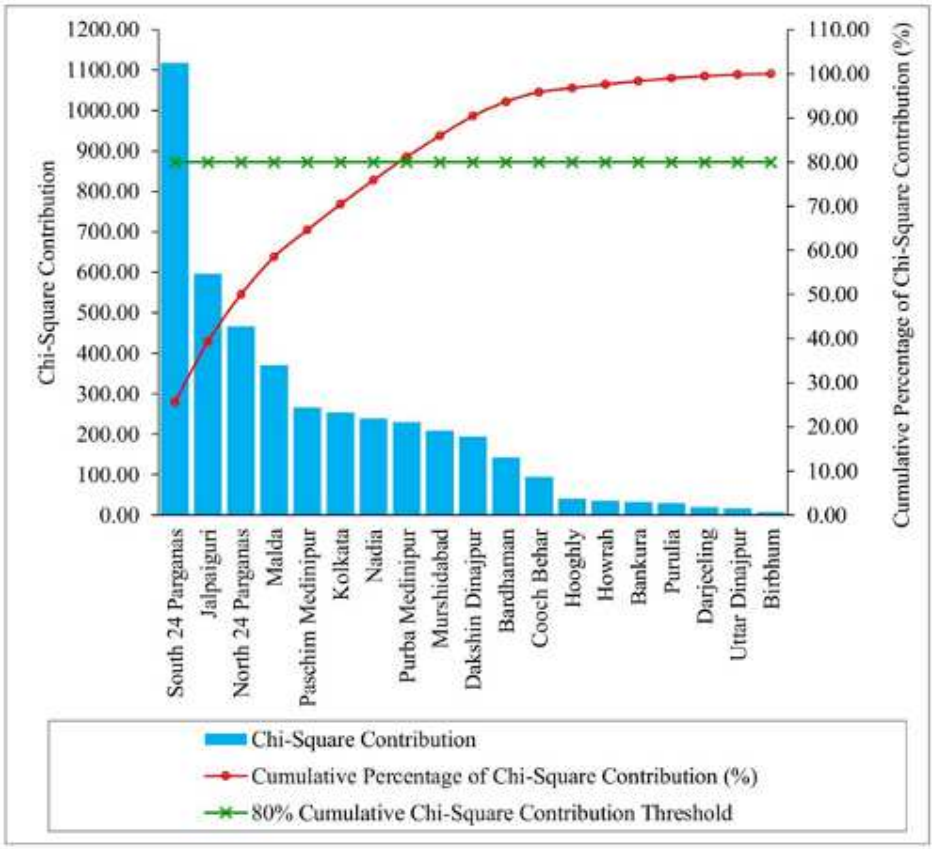


Figure 3: District-wise chi-square contributions and cumulative percentage of CAW in West Bengal (2017–2022).

6.3. Socio-Demographic Drivers of Crimes Against Women:

6.3.1. District-Level Socio-Demographic Profiles and Normalized Indicators

District-level socio-demographic indicators reveal pronounced disparities across West Bengal (table 4). The normalized values show that Kolkata ranks highest in the mean rate of CAW (0.57), achieved maximum literacy (1.00), female schooling beyond the 10th grade (1.00), and urbanization (1.00). In contrast, Purulia records the lowest literacy (0.00) and the highest poverty (1.00). In addition, it is also notable that Darjeeling ranks 2nd in the mean rate of CAW (0.51) and highest in unemployment (1.00), reflecting its distinct socio-demographic profile. These normalized indicators emphasize that crime hotspots are not exclusively linked to poverty; rather, they are shaped by the multi-dimensional interplay of urbanization, education, and socio-economic exposure. Districts with higher educational attainment and urbanization often report higher CAW, suggesting that modernization and awareness may increase reporting, while structural poverty continues to exacerbate vulnerability.

Table 4 presents both actual magnitudes and normalized positions that highlight absolute extremes alongside relative standings. This dual representation of the data strengthens the analytical basis for subsequent correlation and regression assessments, ensuring that both structural conditions and relative district positions are captured in evaluating the socio-demographic drivers of CAW.

Table 4: Actual and Normalized Values of Crime Against Women and Socio-Demographic Variables by District, West Bengal

Districts	Y	Y (N)	$\bar{Y}$	$\bar{Y}$ (N)	X1	X1 (N)	X2	X2 (N)	X3	X3 (N)	X4	X4 (N)	X5	X5 (N)	X6 (N)	$\bar{X}$ (N)
Bankura	30.64	0.08	24.72	0.01	1053.00	0.65	68.30	0.27	27.90	0.12	50.00	0.59	5.00	0.03	9.33	0.00
Bardhaman	53.80	0.29	49.18	0.25	1022.00	0.45	73.35	0.46	32.20	0.26	40.00	0.33	28.00	0.34	42.89	0.37
Birbhum	32.04	0.09	30.86	0.07	1033.00	0.52	70.80	0.37	25.80	0.05	50.00	0.59	39.00	0.49	17.10	0.09
Cooch Behar	83.33	0.56	71.04	0.46	1058.00	0.68	79.20	0.68	26.70	0.08	44.00	0.44	12.00	0.12	11.40	0.02
Dakshin Dinajpur	52.74	0.28	43.63	0.19	1106.00	1.00	74.30	0.50	30.50	0.20	41.00	0.36	3.00	0.00	15.10	0.06
Darjeeling	132.15	1.00	126.71	1.00	1029.00	0.49	77.00	0.60	41.30	0.55	27.00	0.00	77.00	1.00	46.50	0.41

Hooghly	51.76	0.27	50.71	0.26	1082.00	0.84	77.40	0.62	37.00	0.41	33.00	0.15	37.00	0.46	43.70	0.38
Howrah	59.91	0.35	60.27	0.35	1011.00	0.38	80.50	0.73	40.10	0.51	37.00	0.26	26.00	0.31	76.40	0.74
Jalpaiguri	96.89	0.68	74.07	0.49	1038.00	0.55	73.60	0.47	33.90	0.31	41.00	0.36	37.00	0.46	36.90	0.30
Kolkata	87.90	0.60	92.96	0.67	954.00	0.00	87.60	1.00	55.40	1.00	28.00	0.03	34.00	0.42	100.00	1.00
Malda	78.93	0.52	67.86	0.43	1062.00	0.71	72.30	0.42	30.00	0.19	48.00	0.54	11.00	0.11	19.80	0.12
Murshidabad	57.42	0.32	61.04	0.36	1103.00	0.98	67.60	0.25	24.20	0.00	45.00	0.46	17.00	0.19	27.00	0.19
Nadia	79.57	0.52	72.59	0.47	1071.00	0.77	76.20	0.57	28.30	0.13	28.00	0.03	16.00	0.18	34.40	0.28
North 24 Parganas	90.63	0.62	87.01	0.61	1013.00	0.39	85.50	0.92	40.70	0.53	27.00	0.00	34.00	0.42	60.20	0.56
Paschim Medinipur	43.01	0.19	47.96	0.24	1089.00	0.89	70.90	0.37	26.20	0.06	48.00	0.54	12.00	0.12	14.20	0.05
Purba Medinipur	63.37	0.38	53.04	0.28	1031.00	0.51	77.00	0.60	30.90	0.21	34.00	0.18	24.00	0.28	13.00	0.04
Purulia	21.82	0.00	23.73	0.00	1065.00	0.73	61.00	0.00	26.90	0.09	66.00	1.00	9.00	0.08	15.40	0.07
South 24 Parganas	117.26	0.87	119.70	0.93	1055.00	0.66	85.60	0.92	34.90	0.34	42.00	0.38	11.00	0.11	35.40	0.29
Uttar Dinajpur	44.85	0.21	41.90	0.18	1093.00	0.91	65.40	0.17	29.30	0.16	56.00	0.74	21.00	0.24	12.50	0.03
Total	1278.01		1198.98		19968.00		1423.55		622.20		785.00		453.00		631.22	
Min.	21.82		23.73		954.00		61.00		24.20		27.00		3.00		9.33	
Max.	132.15		126.71		1106.00		87.60		55.40		66.00		77.00		100.00	
Mean	67.26		63.10		1050.95		74.92		32.75		41.32		23.84		33.22	
Median	59.91		60.27		1055.00		74.30		30.50		41.00		21.00		27.00	
SD	28.69		27.71		36.53		6.80		7.38		10.24		16.81		23.87	
CV	0.43		0.44		0.03		0.09		0.23		0.25		0.70		0.72	

Notes: Y = Crime rate per 100,000 females (2022), $\bar{Y}$ = Mean crime rate per 100,000 females (2017-22), X1 = Sex ratio per 1000 males (2019–21), X2 = Female literacy rate (2019–21, %), X3 = Female schooling >10th level (2019-21, %), X4 = Head count ratio of multi-dimensional poverty (2015–16, %), X5 = Unemployment rate per 1000 persons (2014), X6 = Urbanization (Projected, 2021, %), $\bar{X}$ = Composite mean of normalized socio-demographic variables, and (N) = Normalized values (0-1)

6.3.2. Associations Between Crimes Against Women and Socio-Demographic Factors

This subsection examines how socio-demographic factors shape the incidence of CAW across districts, highlighting both the strength of their associations and the extent to which these variables explain observed variations in crime rates.

Table 5: Correlation between Crimes Against Women (CAW, 2022) and Socio-Demographic Variables

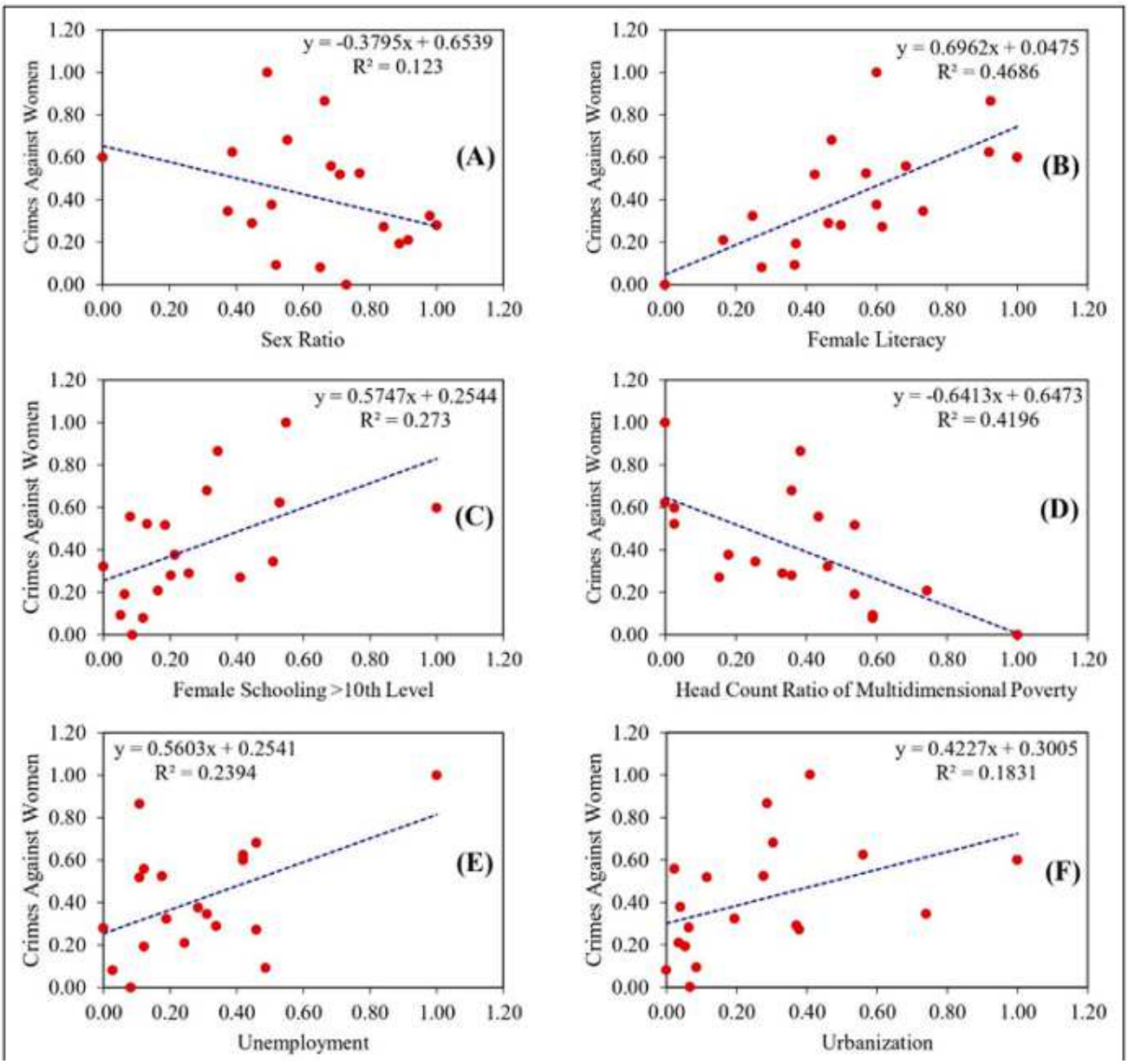
Variables	Spearman's ρ	t-value	p-value	Significance Level
Female Literacy	-0.42	-1.93	0.0700	Not significant ($p \geq 0.05$)
Female Schooling >10th Level	0.72	4.33	0.0005	Highly significant ($p < 0.001$)
Head Count Ratio of Multi-dimensional Poverty	0.61	3.18	0.0045	Significant at 1% level ($p < 0.01$)
Sex Ratio	-0.67	-3.76	0.0015	Significant at 1% level ($p < 0.01$)
Unemployment Rate	0.33	1.42	0.1700	Not significant ($p \geq 0.05$)
Urbanization	0.53	2.59	0.0220	Significant at 5% level ($p < 0.05$)
Composite Socio-demographic Mean	0.48	2.23	0.0400	Significant at 5% level ($p < 0.05$)
Mean CAW (2017-22) and Composite Socio-demographic Mean	0.51	2.46	0.0200	Significant at 5% level ($p < 0.05$)

The combined application of Spearman's rank correlation (table 5) and least-squares regression (figure 4) provide a nuanced understanding of the socio-demographic determinants of CAW. Female schooling beyond the 10th level emerged as the most consistent predictor, showing a strong monotonic association ($\rho = 0.72$, $p < 0.001$) and explaining 27% of the variance ($R^2 = 0.273$). This convergence highlights the role of advanced

education in shaping awareness, reporting, and exposure to gender-based crimes. In contrast, female literacy reveals a moderate and non-significant correlation (ρ = -0.42), yet regression explains nearly half the variance (R^2 = 0.469). This inconsistency suggests that literacy interacts with other socio-demographic factors or functions as a threshold variable, in which its impact becomes evident only beyond certain levels of education.

Structural poverty (Head Count Ratio) demonstrated robust explanatory strength across both methods, with a strong positive correlation (ρ = 0.61, $p < 0.01$) and R^2 = 0.420. Thus, poverty emerges as a fundamental determinant, reinforcing the hypothesis that socio-economic deprivation amplifies vulnerability to CAW. The sex ratio showed a strong negative monotonic relationship (ρ = -0.67, $p < 0.01$), but the regression explained only 12% of the variance (R^2 = 0.123), indicating non-linear or threshold effects. Moderate predictors, such as urbanization and the composite socio-demographic mean, exhibited consistent but weaker associations (ρ = 0.53 and 0.48; R^2 = 0.183 and 0.326), reflecting modernization and aggregated influences. Conversely, unemployment (ρ = 0.33 $p \geq 0.05$) showed weak, non-significant correlations, and regression explained 23% of the variance (ρ = 0.33 ; R^2 = 0.2394). A moderate correlation was found between the mean CAW and the composite socio-demographic mean, with the regression model explaining 43% of the variance.

Overall, the synthesis identified female schooling and poverty as the most robust CAW determinants, supported by both correlation and regression analyses. The sex ratio exerts a significant influence but with non-linear dynamics, whereas urbanization and composite indices provide moderate explanatory strength. This dual-method approach enhances interpretive rigor by revealing convergences and divergences that deepen our understanding of the socio-demographic drivers of CAW.



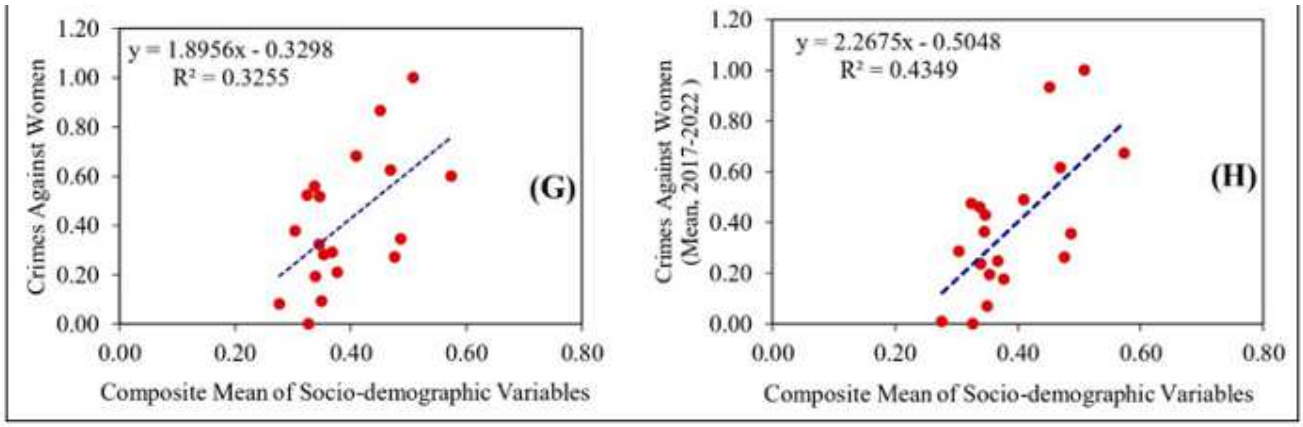


Figure 4: Scatter plots showing the correlation between crimes against women and socio-demographic variables: (A) Sex ratio (2019–21), (B) Female literacy (2019–21), (C) Female schooling >10th level (2019–21), (D) Head count ratio of multi-dimensional poverty (2015–16), (E) Unemployment rate (2015), (F) Urbanization (2021, projected), (G) Composite mean of socio-demographic variables, and (H) Mean crimes against women (2017–22) with Composite mean of socio-demographic variables.

7. Discussion:

The findings demonstrate that CAW in West Bengal is characterized by sharp spatial contrasts and high temporal volatility. Districts such as Dakshin Dinajpur, Jalpaiguri, and Malda exhibited alternating positive and negative deviations, reflecting localized vulnerability and unstable crime dynamics. The standardized residual analysis confirmed statistically significant deviations in Jalpaiguri, Malda, and South 24 Parganas, underscoring their role as hotspots. In contrast, districts such as Birbhum, Darjeeling, and Purulia remained relatively stable, with residuals close to expected values. The temporal shift from widespread negative deviations in 2017–2019 to positive clusters in 2021–2022 may reflect post-pandemic reporting patterns or socio-political changes that influenced crime visibility.

The chi-square contributions revealed that a handful of districts disproportionately shaped the overall crime landscape of the state. South 24 Parganas alone accounted for more than one-fourth of the total chi-square value, and Jalpaiguri and North 24 Parganas emerged as major contributors. The cumulative percentage curve confirmed that the top eight districts explained over 80% of the total chi-square contribution, consistent with the Pareto principle of concentration. This concentration effect highlights that state-level crime statistics are driven by a limited set of districts, whereas the most regions remain relatively stable. These findings emphasize the importance of geographically prioritizing interventions over uniform state-wide policies.

The integration of socio-demographic profiles revealed that crime against women is not solely linked to poverty but is also associated with education exposure and urbanization. Districts with higher levels of female schooling and urbanization, such as Kolkata and North 24 Parganas, reported higher crime rates, which may reflect increased awareness, empowerment, and reporting mechanisms, rather than a higher incidence alone. Purulia, with the lowest literacy and highest poverty rates, recorded the lowest crime rate, suggesting that underreporting may mask actual vulnerability. Regression analysis confirmed that female schooling explained the largest share of variation, followed by urbanization and sex ratio, while poverty contributed modestly and unemployment was largely insignificant. These results underscore the multi-factorial nature of CAW, which are shaped by structural inequalities and demographic imbalances.

8. Policy Implications

This evidence highlights the need for targeted criminological and sociological inquiries in hotspot districts. Policy emphasis should focus on improving gender balance, expanding educational opportunities, and alleviating poverty while recognizing the role of urbanization and reporting mechanisms in shaping crime statistics. Concentrated interventions in South 24 Parganas, Jalpaiguri, and North 24 Parganas could substantially alter the state's crime landscape. Simultaneously, efforts must be made to strengthen reporting systems in districts with low literacy and high poverty, where under-reporting may conceal actual vulnerability.

9. Future Research Directions

Future research could advance this study by conducting qualitative case studies in high-contribution districts to uncover the contextual factors that shape crime patterns and by integrating variables such as law enforcement capacity and victim support infrastructure to enrich the explanatory frameworks. Spatial clustering analysis may help identify contiguous high-risk zones and offer insights into the

regional dynamics of vulnerability. Finally, predictive modelling using machine learning techniques applied to normalized variables could enhance the ability to forecast emerging trends and inform targeted interventions in the future.

10. Conclusion

This study demonstrates that CAW in West Bengal between 2017-2022 is marked by pronounced spatial heterogeneity and temporal volatility. Districts such as South 24 Parganas, Jalpaiguri, and North 24 Parganas have emerged as persistent hotspots, contributing disproportionately to the overall crime. Chi-square analysis confirmed that a small number of districts accounted for more than 80% of the cumulative contribution, underscoring the concentration of crime in specific regions. Simultaneously, the standardized residuals revealed sharp oscillations in certain districts, highlighting unstable crime dynamics and the need for localized monitoring.

The integration of socio-demographic profiles further revealed that CAW is multi-factorial. Female schooling beyond the 10th level, poverty, sex ratio, and urbanization were significantly associated with crime rates, whereas literacy and unemployment showed limited explanatory power. Regression analysis confirmed that educational exposure and gender imbalance were the strongest predictors, although their overall explanatory power remained modest. These findings suggest that higher crime rates in urbanized and better-educated districts may reflect increased awareness and reporting rather than incidence alone, whereas under reporting remains a concern in poorer and less literate regions.

References

[1] Banerjee, D. (2017). Graph of Violence over Women in West Bengal & India: the possible way out. *IOSR Journal of Humanities and Social Science*, *22*(01), 20–25. Https://doi.org/10.9790/0837-2201062025

[2] Bohra, N., Sharma, I., Srivastava, S., Bhatia, M. S., Chaudhuri, U., Parial, S., Sharma, A., Kataria, D. (2015). Violence against women. *Indian Journal of Psychiatry 57*(Suppl 2), S333-S338. https://doi.org/10.4103/0019-5545.161500

[3] Chatterjee, S., Patel, R., & Kar, A. (n.d.). *Crime against women: survey of crimes committed against women in west Bengal in the recent time*. *Gradiva Review Journal, 9*(5), 90-102. https://www.385582715_Crime_against_women_survey_of_crimes_committed_against_women_in_west_Bengal_in_the_recent_time

[4] Das, S. (2021). Critical Analysis on Crime Against Women in West Bengal based Upon Socio-Political Transformation. *International Journal Of Engineering Research & Technology (IJERT) ,10*(11), 147-157. https://doi.org/10.5281/zenodo.18513620

[5] Devakunchari, R., Hackett, M., Thaikkat, R., Sureka, R., Mathur, A., Anshu, Mukherjee, C., Rustagi, P., Krishnaji, N., Ropmay, D., & A. (2019). Crimes Against Women in India using Regression. *International Journal of Innovative Technology and Exploring Engineering*, *8*(6S4), 1460–1463. https://doi.org/10.35940/ijitee.f1297.0486s419

[6] Dhawan, N. B., & Bhasin, H. (2024). Combating domestic violence in West Bengal, India: Gendered norms and legal regulations. *International Journal of Educational Research Open*, *7*, 100352. Https://doi.org/10.1016/j.ijedro.2024.100352

[7] Ghooi, R. B., & Deshpande, S. R. (2013). Violence Against Women in India: A Case for Research in Tackling the Menace. *SSRN Electronic Journal*. Https://doi.org/10.2139/ssrn.2272617

[8] Gupta, S., Sahoo, P. K., & Paltasingh, K. R. (2022). Does development deter crime against women? Panel evidence from India. *Journal of Business and Socio-Economic Development*, *2*(1), 19–33. Https://doi.org/10.1108/jbsed-11-2021-0146

[9] International Institute for Population Sciences (IIPS) & Ministry of Health and Family Welfare, Government of India. (2021). National Family Health Survey (NFHS-5), 2019–21: District Fact Sheets, West Bengal. Mumbai: IIPS. https://www.nfhsiips.in/nfhsuser/publication.php

[10] Islam, J. A., & Saikia, M. (2025). Exploring Crime Against Women in India: socio-economic and spatial dimensions. *Society Register*, *9*(2), 51–72. Https://doi.org/10.14746/sr.2025.9.2.03

[11] Joseph, G., Javaid, S. U., Andres, L. A., Chellaraj, G., Solotaroff, J. L., & Rajan, S. I. (2017). Underreporting of Gender-Based Violence in Kerala, India: An Application of the List Randomization Method. *Policy Research Working Paper 8044, World Bank group, South Asia Region*. Https://doi.org/10.1596/1813-9450-8044

[12] Kalaiyarasi, R. (2015). Violence against Women in India. *IOSR Journal Of Humanities And Social Science (IOSR-JHSS), 20*(2), 51-55. https://www.iosrjournals.org/iosr-jhss/papers/Vol20-issue2/Version-3/H020235155.pdf

[13] Kumaar, S. S. and Verdelli, L. (2015). Crime Against Women - An Indian Scenario [Master Thesis, École Polytechnique De L'universite François Rabelais De Tours]. In *École Polytechnique De L'universite François Rabelais De Tours*. http://memoires.scd.univ-tours.fr/epu_da/local/2015_m2ri_shakthe_sharavana%20kumaar.pdf

[14] Kumari, S., Kumar, N., & Sharma, A. (2025). Mapping the dynamics of crime against women in India: a spatio-temporal analysis. *BELGEO* [Online], *1*. https://doi.org/10.4000/133q5

[15] Mir, M. A. (2022). Analysing Crime against Women in India: A Comprehensive Study. *International Journal for Multidisciplinary Research (IJFMR)*, *4*(4), 631–634. https://www.ijfmr.com/papers/2022/4/5930.pdf

[16] National Crime Records Bureau (2017-2022), "Crime in India", Ministry of Home Affairs, Government of India, New Delhi. https://www.ncrb.gov.in/crime-in-india-all-previous-publications.html

[17] Pandit, K., and Halder, S. (2023). Explaining Crime against Women in the Districts of West Bengal: An Analytical and Empirical Analysis, *In International Journal of Law Management & Humanities*, *6*(4), 111-125. https://ijlmh.com/paper/explaining-crime-against-women

[18] Pooja, B. S., Guddattu, V., & Rao, K. A. (2024). Crime against women in India: district-level risk estimation using the small area estimation approach. *Frontiers in Public Health*, *12*, 1362406. https://doi.org/10.3389/fpubh.2024.1362406

[19] Prasad, G. (2018). Analysis of women victimization in West Bengal. *International Journal of Research in Social Sciences*, *8*,(7), 831-835. http://www.ijmra.us

[20] Registrar General & Census Commissioner, India. (2001 & 2011). Census of India 2001 & 2011: West Bengal, Series 20. Government of India. https://censusindia.gov.in/nada/index.php/catalog/Population+by+district+West+Bengal

[21] Saxena, T. (2025). *Crimes against women in India: trends, challenges, and policy responses – SPRF*. https://sprf.in/crimes-against-women-in-india-trends-challenges-and-policy-responses/

[22] Sharma, T., Devi, S., & Rupa (2022). A Case Study of Crime Against Women In India and Its Impact. *International Advanced Research Journal in Science, Engineering and Technology, 9*(4), 323-325. https://iarjset.com/wp-content/uploads/2022/05/IARJSET.2022.9449.pdf

[23] Siddiqui, A. (2026). Temporal Trends and Regional Variation in Major Crimes against Women in India: Evidence from NCRB Data, 2017–2023, Preprint (Version 1) available at Research Square. https://doi.org/10.21203/rs.3.rs-8764214/v1