Analyzing The Trend And Correlation Between Crime Rates And School Dropouts In India Using Power BI

by

Evelyn Anna Alex

Under the Supervision of Dr. Victer Paul

Data Analytics and Business Decisions Research Group, IIIT Kottayam

Table of Contents

Abstract

Crime and education are two vital social issues that are closely intertwined in the Indian context. Rising crime rates threaten public safety, while high school dropout rates reflect educational disengagement and socio-economic vulnerabilities. Although these issues have been studied independently, the relationship between dropout rates and crime remains underexplored in Indian datasets. This study aims to examine national and state-level patterns in crime and dropout rates from 2013 to 2022, and to investigate whether meaningful correlations exist between these two variables.

The study uses crime statistics from the National Crime Records Bureau (NCRB) and dropout data from UDISE+, supplemented with information from government portals. Data preprocessing was performed using Excel and Power Query to normalize state names, handle missing values, and standardize year formats. A Power BI dashboard was developed to visualize trends and patterns through various interactive visualizations. Statistical correlation techniques, including Pearson and Spearman coefficients, were applied to test the associations observed in the visualizations.

Chapter 1: Introduction

1.1 Background

Education plays a pivotal role in shaping social development and promoting equitable growth. In India, where education retention remains a challenge, dropout rates are often viewed as indicators of socio-economic stress [1]. At the same time, crime rates reflect societal safety and the effectiveness of governance. Although education and crime are frequently studied independently, examining their intersection can provide valuable insights into the vulnerabilities faced by communities [1]. Systematic analysis of the relationship between school dropouts and crime using state-level datasets in India remains limited. In the era of data-driven governance, tools such as Power BI enable policymakers to visualize, analyze, and interpret large-scale data for informed decision-making.

1.2 Problem Statement

Understanding the social dynamics behind crime and education is essential for effective policymaking in India. Educational dropouts are often connected to socio-economic vulnerabilities that can lead to or be worsened by crime [2]. However, the link between education dropout rates and increasing crime levels remains insufficiently explored in Indian data [3]. This project aims to analyze national and regional trends in school dropout rates and crime statistics using publicly available government data. By utilizing Power BI, we strive to identify patterns, regional disparities, and correlations over time. These insights can help design targeted interventions to improve education retention and reduce crime. The project promotes data-driven governance by visually clarifying the complex relationships between education and crime.

1.3 Research Objectives

The main objectives of this study are:

- To examine temporal and spatial patterns of crime rates and school dropout rates in India.
- To analyze the relationship between dropout rates and specific social issues such as juvenile crime, child marriage, and trafficking.
- To assess whether female literacy levels influence crimes against women.
- To design and implement a Power BI dashboard for data visualisation.

1.4 Scope of the Study

The study uses data from 2013 to 2022 across Indian states and union territories. Crime data is drawn from NCRB reports, while education dropout statistics are obtained from UDISE+ and supplementary portals. The analysis is restricted to state-level data and focuses on correlations rather than establishing causal relationships.

1.5 Significance of the Study

This research is significant for multiple reasons. First, it provides an integrated view of two critical socio-economic concerns, crime and education. Second, it equips policymakers with a data-driven framework for designing targeted interventions. Third, by combining statistical analysis with visualization, the study offers a practical demonstration of how analytical dashboards can enhance transparency and governance. [4]

1.6 Organization of the Paper

- Chapter 1 introduces the background, objectives, scope, and significance of the study.
- Chapter 2 presents a literature review covering crime trends, dropout studies, global and Indian perspectives on the education-crime nexus, and the role of data analytics in policymaking.
- Chapter 3 details the methodology, including data sources, preprocessing techniques, and the analytical framework implemented in Power BI.
- Chapter 4 reports the experimental setup, dashboard development, results, and discussions of trends, correlations, and insights.
- Chapter 5 concludes with a summary of findings, contributions, limitations, and future directions.

Chapter 2: Literature Study

2.1 Crime Trends and Socio-economic Factors in India

Recent patterns in Indian crime data reveal a strong connection between property-related offences such as theft and burglary and underlying socio-economic conditions. Poverty stands out as the most significant factor, as more households fall below the poverty line, incidents of property crime rise sharply, suggesting that economic hardship often compels individuals toward unlawful activity [5]. Regions marked by rapid but uneven income growth also experience elevated theft rates, indicating that economic disparity may foster resentment and a sense of opportunity among those left behind.

Education appears to counteract these trends. Communities with higher school enrollment and better access to quality education consistently report lower levels of crime [6]. This indicates that the role of education is not only in improving employment prospects but also in reinforcing social norms that discourage criminal behaviour.

The effectiveness of the legal system further influences crime rates. Weak enforcement, slow judicial processes, and limited access to justice contribute to an environment where offenders perceive a low risk of punishment, increasing the likelihood of repeat offences [7]. In contrast, stronger legal institutions tend to deter criminal activity by raising the certainty of consequences [8].

2. 2 Educational Dropout Rates and Social Impact

Higher education in India has grown substantially to meet the demands of a dynamic economy, yet dropout rates continue to challenge its effectiveness. Students discontinue their studies due to a combination of socioeconomic disparities, academic unpreparedness, limited institutional support, psychological stress, and external pressures [9]. Comparative analysis with global contexts shows that while financial constraints and academic readiness are common issues, India's socioeconomic diversity and cultural emphasis on education present distinct challenges [9].

The social impact of dropout rates is substantial. School dropouts contribute to a shortage of skilled labor, resulting in a workforce that is not prepared to meet the demands of a growing economy [10]. This skill gap affects national productivity and reduces India's global competitiveness [11]. An uneducated population contributes less to economic output and increases reliance on social welfare programs, placing a financial strain on public resources. Social consequences include increased vulnerability to exploitation, early marriages, particularly among girls, and the perpetuation of poverty [9]. These outcomes reinforce cycles of inequality and hinder inclusive development.

Dropout rates in higher education reflect broader systemic issues that extend beyond academic institutions. The loss of educational opportunity affects not only individual futures but also the social and economic fabric of the nation [9]. As dropout rates persist, the challenges of building an equitable, skilled, and resilient society remain unresolved.

2.3 Relationship Between Education and Crime – Global and Indian Studies

The connection between education and crime has been widely examined across different countries. Globally, education is often considered a protective factor against criminal behavior. It promotes values such as discipline, cooperation, and future-oriented thinking, which can reduce the likelihood of engaging in unlawful activities [12]. In many developed nations, higher levels of education are associated with lower rates of violent and property crimes [13]. Educational environments also provide structure and social integration, which are important in deterring youth from criminal paths.

In India, the relationship between education and crime is more complex. While broader access to education has contributed to reductions in certain types of crime, there is evidence that higher educational attainment may be linked to an increase in non-violent offenses, such as economic and cybercrimes [14]. This suggests that education does not uniformly deter criminal behavior and may, in some cases, enable access to more sophisticated forms of crime. Socioeconomic factors including poverty, inequality, and population density, also influence crime patterns and interact with education levels in significant ways [15].

India's unique social and economic landscape further shapes this relationship. Rapid urbanization, uneven development, and challenges within the legal system contribute to rising crime rates. While education remains a vital tool in promoting lawful behavior and social stability, its impact is influenced by broader demographic and institutional conditions [7]. Understanding these interactions is essential for developing informed policies that address both the causes and consequences of crime.

2.4 Role of Data Analytics and Visualization in Policy-making

Data analytics has become an essential part of modern policy-making, enabling governments to make decisions based on evidence rather than assumptions [16]. By analyzing large volumes of data, policymakers can identify trends, forecast outcomes, and design targeted interventions in areas such as public health, education, and infrastructure. This approach improves the precision and effectiveness of government initiatives while promoting transparency and accountability [16].

In India, the growing emphasis on digital governance has accelerated the use of data analytics in public administration. Platforms like the National Data and Analytics Platform and initiatives under Digital India have improved governance efficiency [17]. These tools have helped address challenges such as urban congestion, resource distribution, and public health management, making governance more responsive and inclusive.

Data visualization plays a critical role in translating complex datasets into actionable insights. Interactive dashboards, geospatial mapping, and visual storytelling enable decision-makers to grasp patterns and communicate findings effectively and quickly [18]. As India continues to navigate rapid socio-economic changes, the integration of visualization with analytics will be key to designing adaptive policies that reflect ground realities and serve diverse communities [18].

2.5 Gap in Existing Research/Need for the Work

Studies addressing the relationship between education and crime often face structural limitations in the available datasets. Crime statistics and dropout rates are recorded by academic year, yet they are typically labelled using only the starting year of that cycle. This labelling convention creates ambiguity in temporal alignment, making it difficult to maintain consistency when comparing data across different sources or tracking changes over time.

Another constraint involves the availability of literacy rate data, which is confined to recent years. Earlier records are either missing or inaccessible, restricting the scope of longitudinal analysis. This limitation reduces the ability to incorporate foundational educational indicators when examining broader patterns related to dropout rates and crime statistics.

Chapter 3: Proposed Work

3.1 Overview of the Study Design

This study explores the relationship between education and crime in India by analyzing trends in crime rates and dropout rates across states using Microsoft Power BI. It follows a descriptive analytical design that integrates official datasets on crime, education, and socioeconomic factors to examine how variations in educational attainment and literacy levels may influence criminal activity. The approach combines quantitative data analysis with visual exploration, allowing both spatial and temporal comparisons across different regions and years.

The design emphasizes the identification of regional disparities, long-term trends, and possible linkages between high dropout or low literacy rates and increased crime incidence. Through Power BI, raw data from multiple government sources are transformed into an interactive dashboard, facilitating correlation analysis and enabling users to interpret complex relationships between social variables visually. This integration of data visualization and statistical interpretation enhances the clarity and accessibility of insights, supporting evidence-based understanding of how educational outcomes may affect crime patterns in India.

3.2 Data Sources and Description

This study explores the relationship between education and crime across Indian states over the period 2013 to 2022, utilizing secondary data sourced from official government repositories. The primary datasets were obtained from the National Crime Records Bureau (NCRB), which provides detailed statistics on reported crimes by category and region, and the Unified District Information System for Education Plus (UDISE+), which offers comprehensive data on school dropout rates [19] [20]. These sources were selected for their reliability, consistency, and national coverage.

3.2.1 Crime Data – NCRB Reports

State-wise crime data were obtained from the National Crime Records Bureau (NCRB) for the period 2013 to 2022. The dataset includes annual statistics on various categories of offences, including theft, murder, hurt, crimes against women, kidnapping and abduction, economic offences, child trafficking, child marriage, and juvenile crimes. In addition to these crime categories, the data provide demographic details related to juveniles, such as the juvenile population, apprehension rates, and levels of educational attainment. Juvenile education is classified into four categories: illiterate, up to primary, above primary but below matric, and matric or higher secondary and above [19].

These data form the basis for analyzing the prevalence and demographic composition of crime across states and union territories. The inclusion of education-related variables for juveniles enables the study to explore potential correlations between low educational attainment and higher rates of juvenile delinquency.

3.2.2 Education Dropout Data – UDISE+ Statistics

Educational indicators were obtained from UDISE+, the national information system maintained by the Department of School Education. The dataset provides annual state-wise and gender-wise statistics on dropout trends at the primary, upper primary, and secondary levels. In the compiled dataset, variables such as Primary Girls, Primary Boys, Upper Primary Girls, Upper Primary Boys, Secondary Girls, and Secondary Boys reflect the extent to which students discontinued schooling before completing the respective stages [20].

These indicators contribute to understanding patterns of educational discontinuity across regions. Integrating them with crime data helps examine how dropout trends may relate to variations in criminal activity across states and union territories.

3.2.3 Supplementary Sources – Government Portals

Additional variables, such as population density and literacy rate, were referenced from reports available on official government portals. These include the Open Government Data (OGD) Platform India, the Ministry of Statistics and Programme Implementation (mospi.gov.in), and the Directorate General of Education (dge.gov.in). Population data were used to calculate relevant ratios. Literacy data, available in a state-wise and gender-wise format, were used to provide context for interpreting educational outcomes and regional disparities.

These supplementary indicators support the analytical framework by accounting for broader environmental and structural factors that may influence both education and crime. Combined with the NCRB and UDISE+ datasets, they contribute to a more comprehensive understanding of patterns across Indian states.

3.3 Data Preprocessing Methodology

The preparation of the datasets involved several structured steps to ensure that the information was consistent, accurate, and ready for analysis within Power BI. The process involved converting raw reports into structured formats, cleaning and normalizing the data, managing missing values, and integrating datasets across states and years.

3.3.1 Data Cleaning and Normalization

The data were initially collected from government reports such as UDISE+ educational statistics and NCRB crime records. As most reports were available in PDF or tabular formats, they were first imported into Microsoft Excel for preliminary processing. This step involved organizing the data into structured tables and preparing it for further analysis.

Subsequent data cleaning addressed inconsistencies such as irregular state or union territory names, non-standardized year fields, and duplicate entries. State names and year formats were standardized to enable accurate relational mapping across datasets. Non-relevant records were excluded, and variable names were made consistent to ensure clarity and seamless integration within Power BI.

3.3.2 Handling Missing Values and Aggregation

To enable cross-state comparison, all variables were converted to ratios. This ensured consistency in scale and allowed meaningful comparison across regions with differing population sizes and educational profiles.

Missing or incomplete records were minimal and were mainly observed in smaller union territories or in recent education statistics. These gaps were handled using methods such as averaging values from adjacent years and by removing blank rows. This approach maintained data continuity without affecting the overall trend or accuracy of the analysis.

3.3.3 Merging Crime and Dropout Datasets

Following the cleaning and normalization process, all datasets were consolidated and converted into CSV format for structured processing. These CSV files were then imported into Microsoft Power BI for further refinement and integration. Within Power BI, redundant columns were removed, and selected fields were unpivoted to facilitate flexible analysis.

The datasets were merged using state and year as key fields, resulting in a unified analytical model that incorporated both educational and crime-related indicators. Analytical measures were also created within Power BI to calculate overall rates and enable meaningful comparisons across indicators. Each record corresponded to one state or union territory for a specific year. This integrated dataset formed the basis for visual and statistical analysis.

3.4 Analytical Framework

Power BI was chosen for its ability to integrate large multi-year datasets and convert them into interactive, visual, and data-driven insights. The combined data model, built using the state and year fields as relational keys, supported comparative analysis of crime, education, and social indicators. The dashboard was designed to offer a unified analytical environment that connects crime patterns with educational outcomes, enabling relational insights within a single interface.

3.4.1 Power BI Dashboard Architecture

Building on these visualization elements, the Power BI dashboard was structured to provide a clear, organized, and interactive view of the combined crime and education data. The first section presented an overview of overall crime rates and dropout rates across Indian states, allowing a regional understanding of the areas most affected by either social or educational challenges. Dropout data were analyzed by gender to identify differences between male and female students and to highlight regions with the greatest gender disparities.

Another segment of the dashboard analyzed the educational background of arrested juveniles, offering insight into the role of education in juvenile crime. The relationship between dropout rates and juvenile crime was then explored, showing how educational discontinuation might relate to youth involvement in criminal activity. Additional visuals examined patterns in girls' dropout rates across years and educational categories, as well as their connection to social issues such as child marriage and child trafficking. Literacy indicators were also analyzed to assess the differences between male and female literacy levels, while a specific comparison between female literacy and crimes against women was conducted to study how educational attainment may influence vulnerability to gender-based violence.

The Power BI dashboard integrated educational and crime-related indicators into a single analytical platform, enabling comparative analysis across states and demographic groups.

3.4.2 Visualization Elements

The Power BI dashboard utilized a range of visualization elements to present data in an accessible and interpretable manner. Geographical maps illustrated regional variations in crime and dropout rates, while time-series visuals showed how these indicators evolved over the study period. Categorical comparisons were made through stacked and clustered visuals that displayed gender-based dropout patterns, literacy rates, and the educational background of arrested juveniles.

Pie charts were used to show proportional distributions across categories, and interactive filters allowed users to focus on specific years, states, or indicators. Through these visual elements, the dashboard transformed raw data into a cohesive analytical model, enabling deeper insights into the relationship between education and crime in India.

3.5 Proposed Analytical Model for Correlation Analysis

The proposed model quantitatively examines the relationship between educational and crime variables across Indian states and years. While Power BI provides a dynamic platform for exploring trends and spatial patterns, statistical correlation analysis formally evaluates the strength and direction of associations between dropout rates, literacy, and different crime categories. This framework helps examine whether differences in education levels are connected to juvenile arrests and social issues like child marriage and trafficking.

3.5.1 Statistical Correlation Techniques

Both Pearson and Spearman correlation coefficients were computed using Python's pandas and scipy.stats libraries [21]. Pearson's correlation measured linear associations between continuous variables, such as dropout and overall crime rates, under the assumption of approximate normality. Spearman's rank correlation was selected as a complementary measure, suitable for monotonic relationships and robust to potential violations of linearity or normality assumptions.

Analyses were conducted for multiple variable pairs, including dropout rate versus overall crime rate, dropout rate versus juvenile apprehensions, and dropout rate versus child marriage and trafficking rates. For each pair, correlation coefficients and p-values were calculated to assess significance. By integrating Power BI visualizations with statistical analysis, the model examines how differences in education relate to crime patterns across states, years, and demographic groups.

3.5.2 Possible Causality Considerations

While correlation analysis identifies associations between educational and crime indicators, it does not imply causation. Observed relationships between dropout rates and crime may be influenced by a range of socioeconomic and structural factors, including poverty, urbanization, unemployment, regional disparities, and the effectiveness of education and law enforcement systems. As a result, correlations should be interpreted as indicative rather than definitive, and by themselves may not provide a complete or accurate explanation of the patterns observed.

The associations identified are influenced by multiple interacting factors, and correlation analysis alone cannot fully capture the complexity of these relationships. The model is therefore designed as an exploratory framework to identify patterns.

Chapter 4: Experimental Setup, Results and Discussion

4.1 Implementation Environment (Power BI, Excel, Python)

The experimental environment for this study comprised Power BI, Microsoft Excel, and Python, each serving distinct roles in data processing, analysis, and visualization. Excel was used for preliminary data cleaning, organization, and creating structured tables from datasets collected from government reports, including UDISE+ educational statistics and NCRB crime records. Measures and calculations, such as normalization and derived indicators, were also prepared in Excel before importing the data into Power BI.

Python was employed for data preprocessing and advanced analysis. Libraries such as pandas and scipy.stats enabled efficient handling of large datasets, while exploratory data analysis identified trends, correlations, and outliers. Python scripts also automated repetitive tasks and ensured consistency in the calculated measures, facilitating smooth integration with Power BI.

Power BI served as the main platform for dashboard development, integrating processed data from Excel and Python. Features such as Power Query, DAX calculations, and interactive visualizations were used to create dashboards that supported data-driven insights into the relationships between educational outcomes and crime indicators.

4.2 Dashboard Development Process

The dashboard development followed a structured workflow aligned with the research objectives. Initially, cleaned datasets were imported into Power BI, and relationships between tables were established to allow coherent analysis across multiple indicators. Calculated measures, including dropout rates and juvenile apprehended rates, were defined using DAX to enable precise and dynamic analysis.

Visualization design focused on clarity and interactivity. Interactive filters and slicers allowed dynamic exploration of data by year, region, and indicator type, while multiple visual elements were combined into a cohesive dashboard interface. The final dashboards highlighted key trends and relationships, providing actionable insights into the interplay between educational outcomes and crime statistics.

4.3 Visual Analysis of Trends

4.3.1 Temporal Trends in Dropout and Crime Rates

Crime Rate vs. Dropout Rate

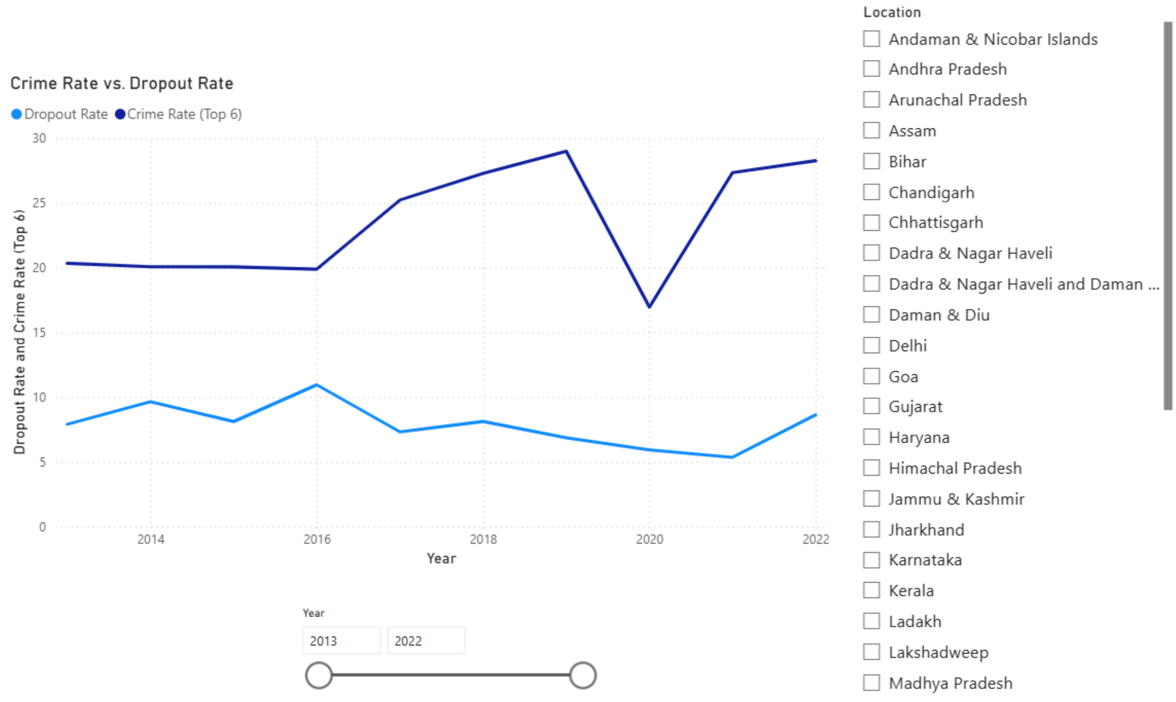

Figure 4.3.1.1

The graph shows how the overall dropout rate and crime rate have changed between 2013 and 2022. The dropout rate, shown in light blue, remains fairly steady through most years with slight rises around 2016 and 2022. The crime rate, shown in dark blue, remains higher overall and shows a noticeable increase after 2016, followed by a drop around 2020, then a rise in 2022. Both span a similar range of years but differ in intensity, with fluctuations across the period.

Crime Rates by Year and Location

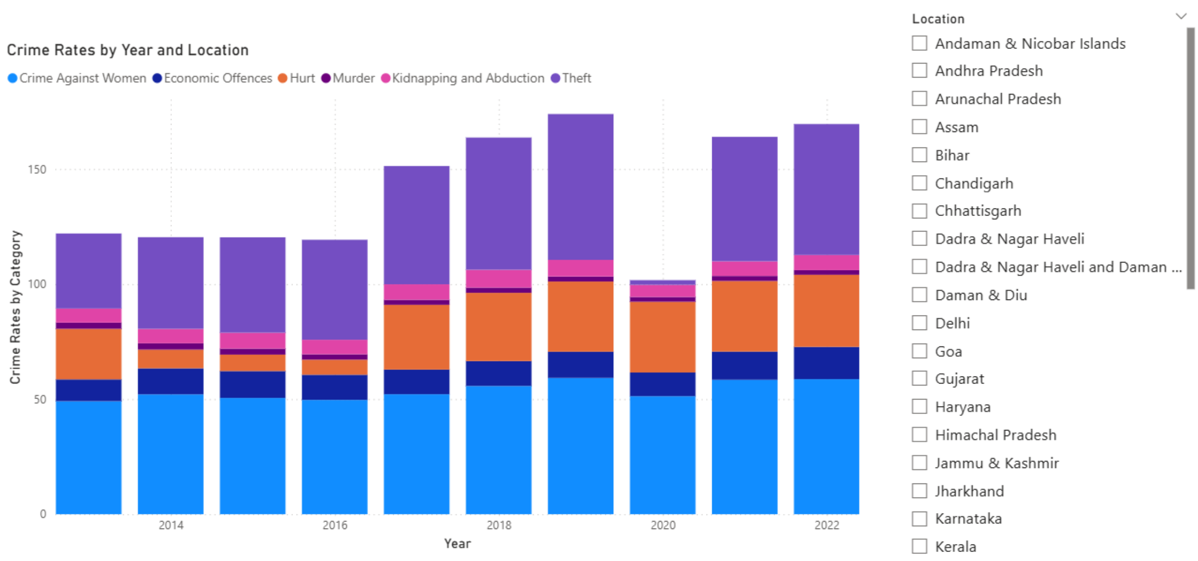

Figure 4.3.1.2

The figure presents trends in total crime rates by category from 2013 to 2022. Reported crimes exhibit variation over the years, with a general increase observed after 2017. Among the different categories, theft and crimes against women consistently represent the largest proportion, while murder and bodily harm constitute comparatively smaller shares.

Dropout Rates by Gender and Year

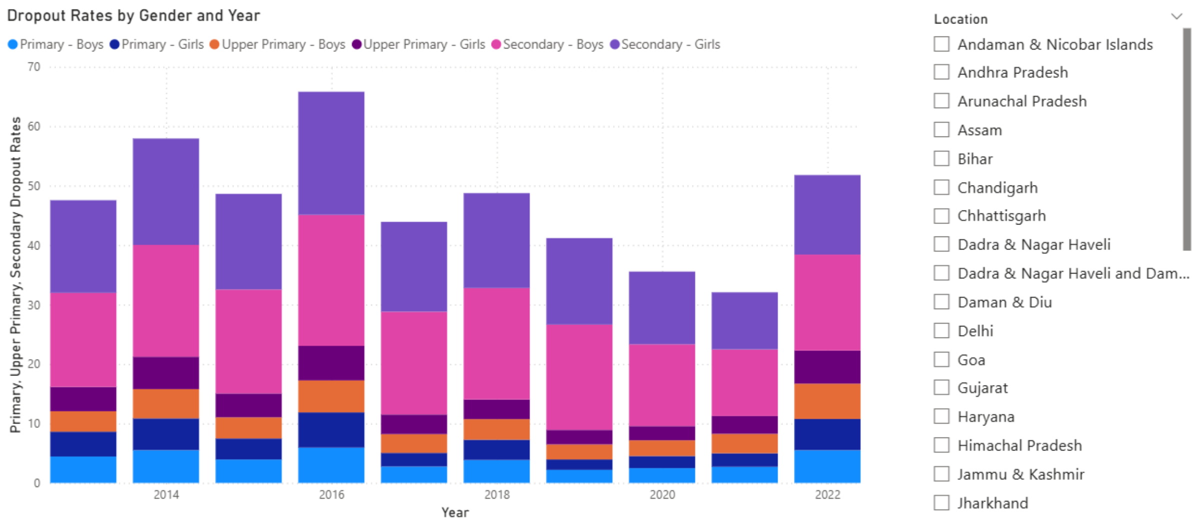

Figure 4.3.1.3

The figure also illustrates dropout rates for boys and girls across primary, upper primary, and secondary education levels during the same period. Dropout levels vary by year and gender, with secondary-level dropouts comprising the most substantial share in most years. Peak dropout rates occur around 2016, followed by a moderate decline in subsequent years before rising again in 2022. Across all education levels, dropout rates for girls are generally higher, particularly in upper primary and secondary education.

Education Background of Arrested Juveniles by Year and Location

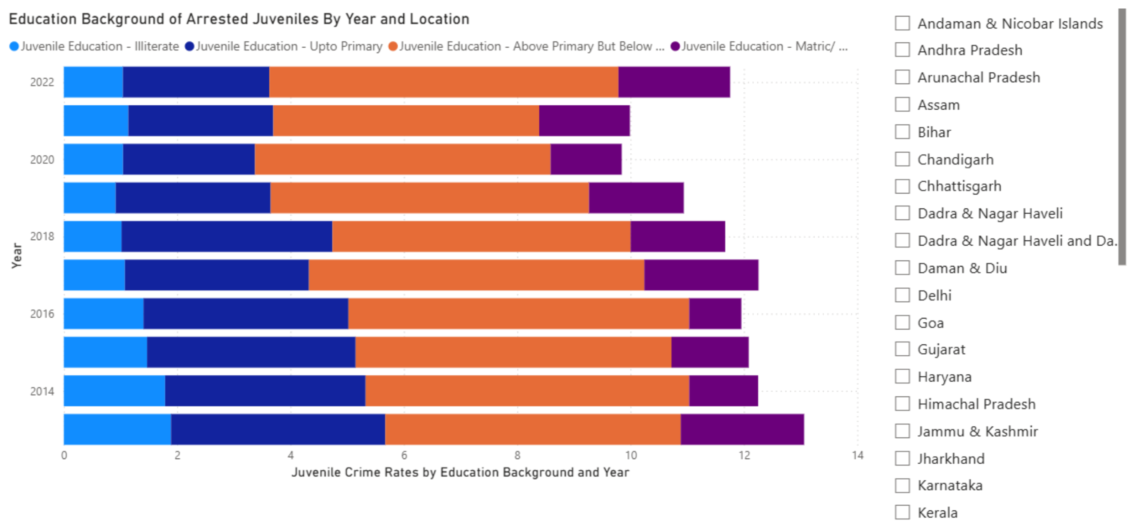

Figure 4.3.1.4

From 2014 to 2022, juveniles with lower levels of education, especially those who are illiterate or have only completed primary school, consistently made up the largest share of arrests. The proportion of juveniles with education above primary but below matric increased steadily during this period. Those with matric-level education remained the least represented in the data. The chart includes filters for viewing data by Indian states and union territories.

Dropout Rate vs. Juvenile Crime Rate by Year

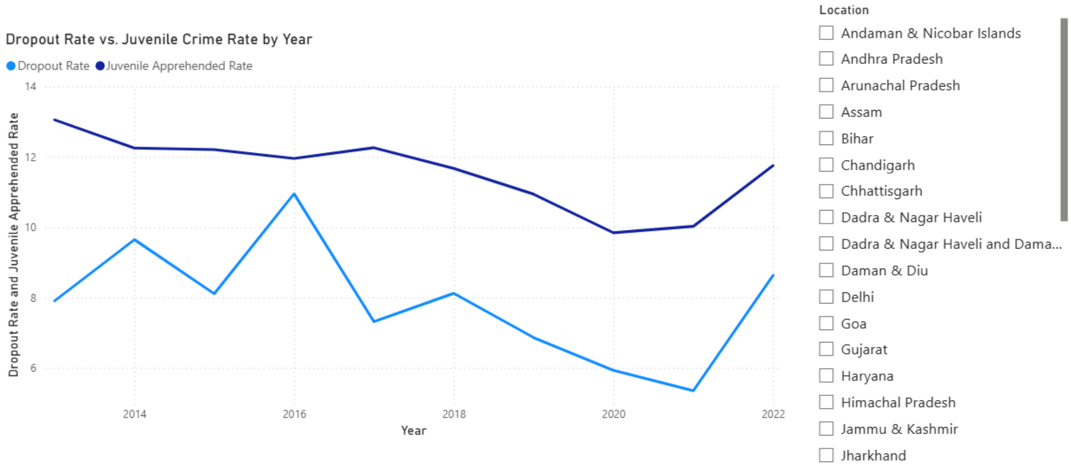

Figure 4.3.1.5

Between 2012 and 2022, the dropout rate in India steadily declined, starting from a higher level in 2012 with noticeable reductions around 2016 and 2020. The juvenile apprehension rate remained relatively stable during this period, with minor increases around 2015 and 2022. The graph includes location filters for viewing trends across different states and union territories.

4.3.2 Regional Disparities and Hotspot Identification

Overall Crime Rate by Location

Crime Type	Crime Rate
Crime Against Women	53.71
Theft	44.30
Hurt	22.49
Juvenile Apprehended Rate	11.59
Economic Offences	11.30
Kidnapping and Abduction	6.56
Murder	2.33
Child Trafficking	0.13
Child Marriage	0.11
Total	**16.95**

Year

2013 2022

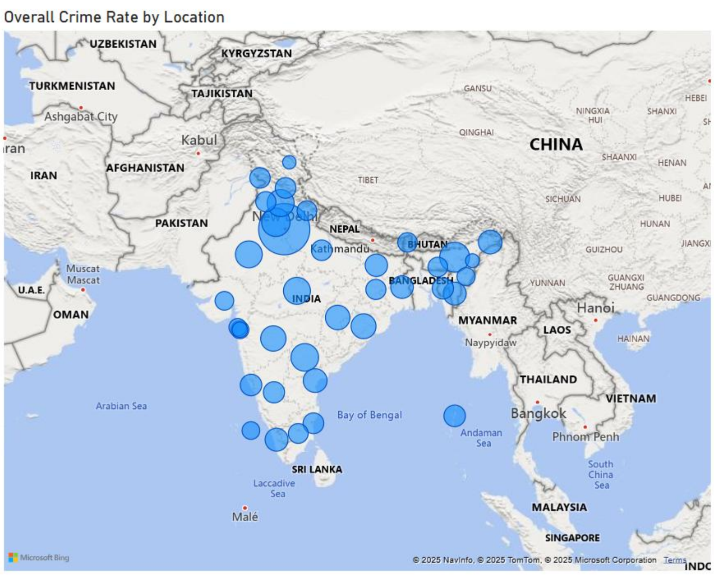

Figure 4.3.2.1

This figure illustrates the overall distribution of crime rates across locations. Among these, crimes against women account for the largest share, followed by theft and hurt. Other categories, such as economic offences, juvenile apprehensions, and kidnapping or abduction, show moderate levels, while incidents like murder, child trafficking, and child marriage remain relatively lower in comparison. Among all major cities in India, Delhi consistently records the highest number of reported crimes.

Overall Dropout Rate by Location

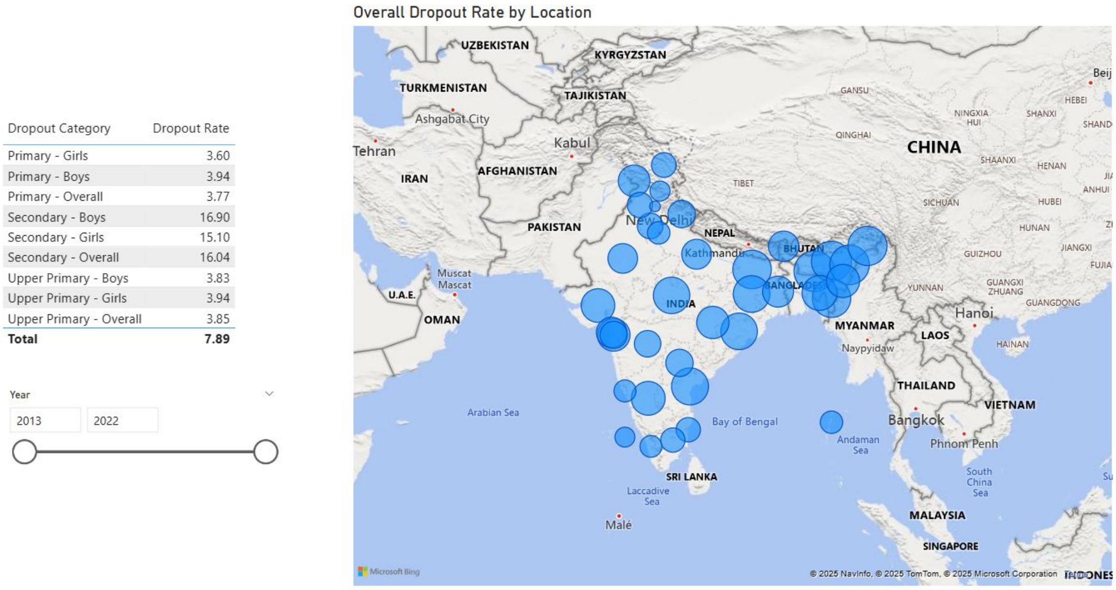

Figure 4.3.2.2

This figure presents the overall dropout rates across different locations. Secondary-level dropouts represent the highest proportion, with both boys and girls showing higher rates compared to the primary and upper primary levels. In contrast, dropout levels at the primary stage remain considerably lower across most regions. The visualization suggests regional disparities in educational continuity, with certain states showing more pronounced dropout concentrations at the secondary level.

Crime Rate vs. Dropout Rate by Location

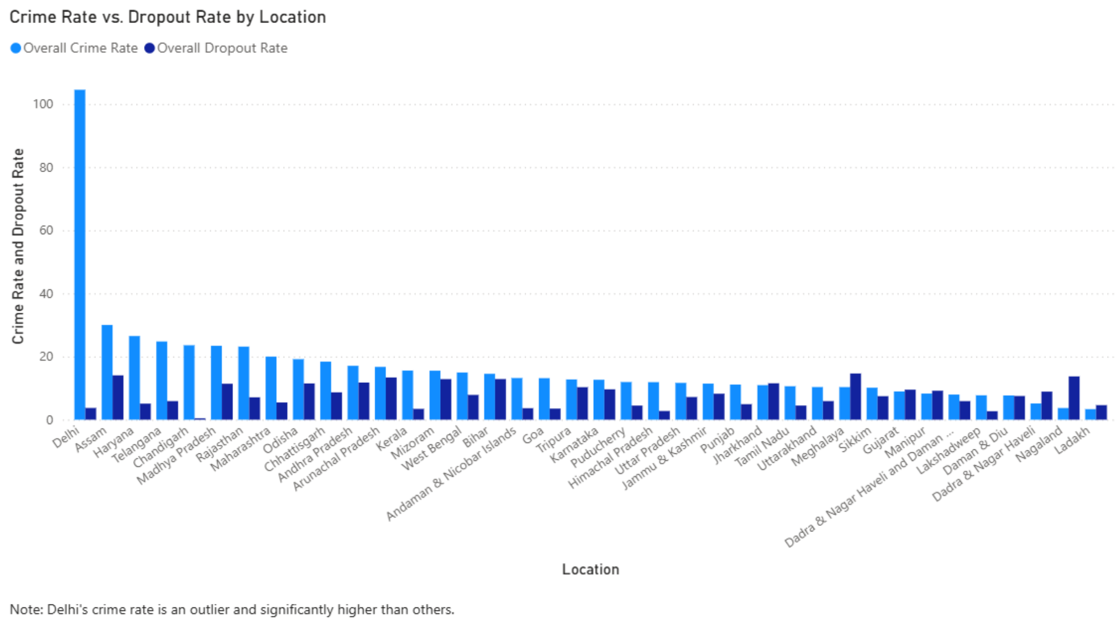

Figure 4.3.2.3

This figure compares the overall crime rate and dropout rate across locations to identify possible patterns between educational disengagement and crime prevalence. Delhi stands out with a markedly higher crime rate than all other regions, indicating a significant outlier that affects overall comparability. Other states such as Assam, Haryana, and Telangana also record relatively higher crime rates, while their dropout levels vary. Several regions show moderate rates for both indicators, suggesting diverse social and structural contexts influencing these outcomes.

Crime Rate vs. Dropout Rate by Location (Excluding Delhi)

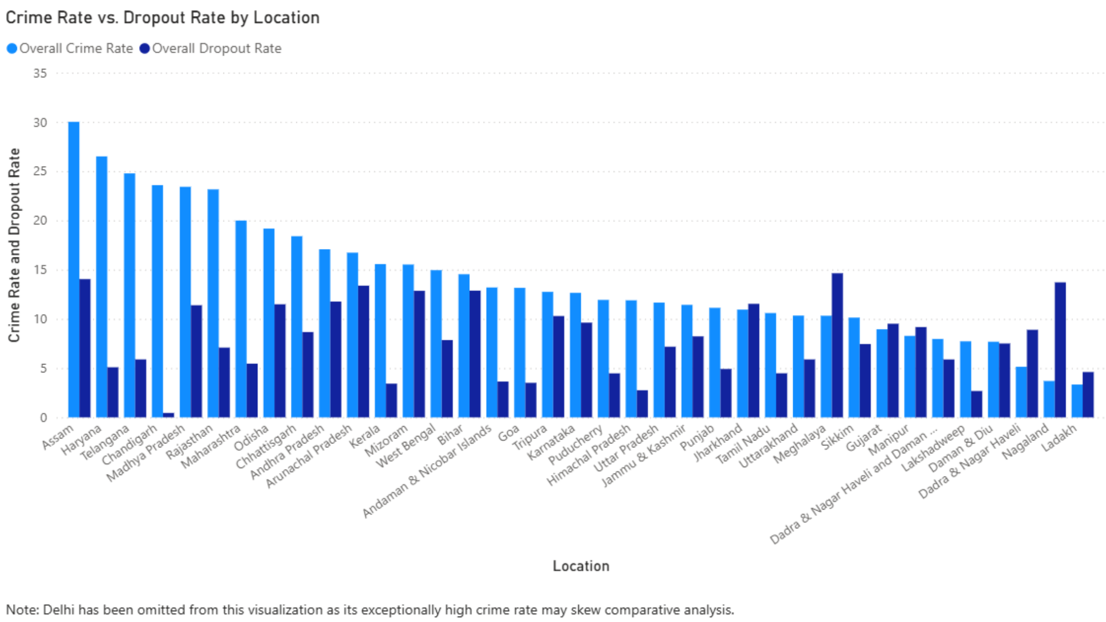

Figure 4.3.2.4

Delhi, which exhibited an exceptionally high crime rate in the previous visualization, has been excluded here to enhance the clarity of comparisons among the remaining regions. Without Delhi skewing the scale, patterns emerge more distinctly. States like Assam, Haryana, and Telangana continue to show elevated crime levels, though their dropout rates differ. Meanwhile, several other states display moderate values for both indicators, pointing to a range of underlying social and structural factors shaping these outcomes.

4.3.3 Gender and Age-based

Primary, Upper Primary, Secondary Dropout Rates of Girls by Year

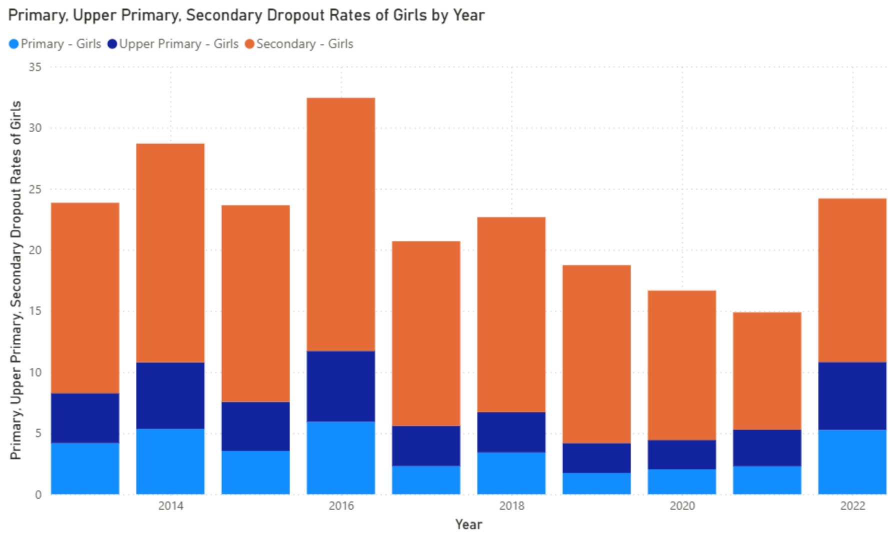

Figure 4.3.3.1

The graph shows the annual dropout rates for girls in primary, upper primary, and secondary education from 2013 to 2022. The rates vary across years and educational levels. Secondary education consistently shows higher dropout percentages compared to primary and upper primary. The highest recorded dropout rates occurred in 2016, while the lowest were observed in 2021. Each year presents a distinct distribution across the three categories.

Literacy Rate by Gender

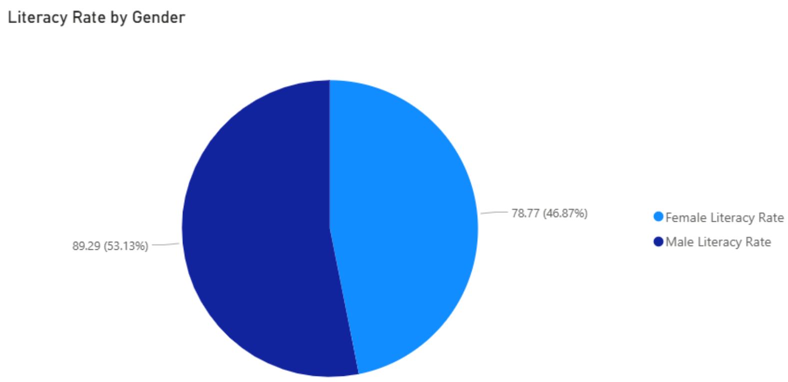

Figure 4.3.3.2

The data is divided into two categories representing male and female literacy rates. The male literacy rate is recorded at 89.29 percent, while the female literacy rate stands at 78.77 percent. Each category also reflects its respective share of the total literate population, with males comprising 53.13 percent and females accounting for 46.87 percent.

Crime against Women vs. Female Literacy Rate by Location (Sorted by Crime Rate)

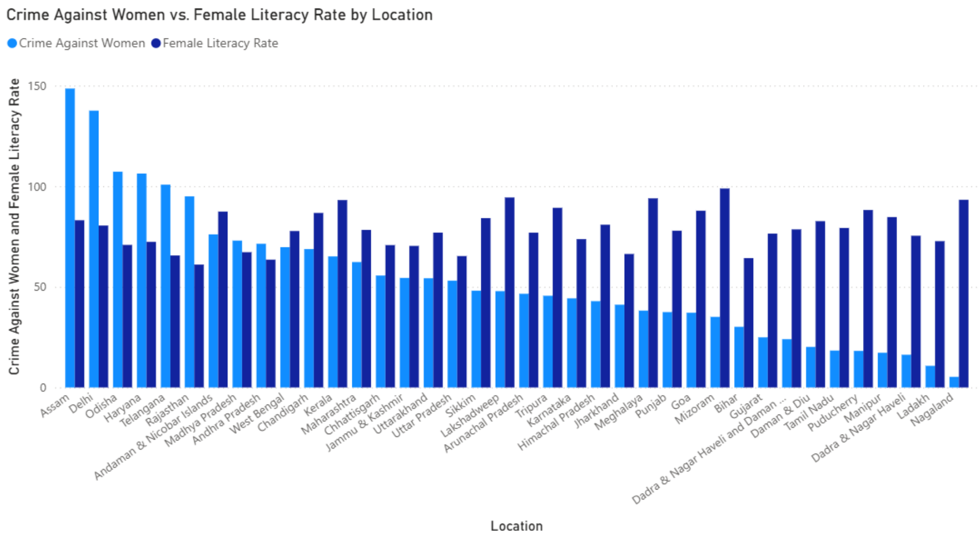

Figure 4.3.3.3

The graph compares crime rates against women with female literacy levels across selected states and union territories. Delhi records the highest crime rate among all regions. Literacy rates vary independently, with some areas like Tripura and Arunachal Pradesh showing higher crime rates alongside lower literacy levels. States such as Ladakh and Nagaland report low crime rates.

Crime against Women vs. Female Literacy Rate by Location (Sorted by Literacy Rate)

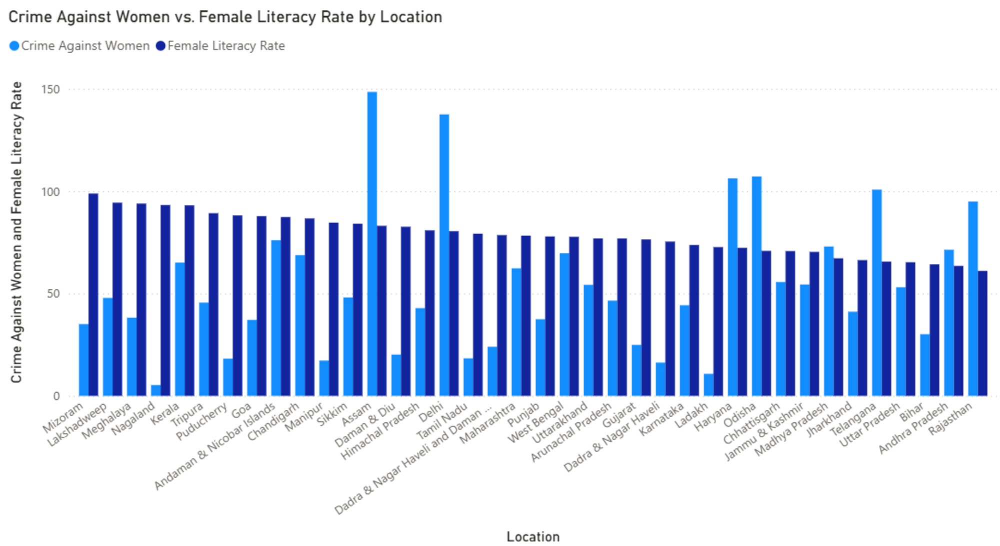

Figure 4.3.3.4

This graph displays crime rates against women alongside female literacy levels across various regions, sorted in descending order by female literacy rate. Mizoram ranks highest in literacy, followed by Lakshadweep, both showing moderate crime rates compared to other regions. Rajasthan appears at the lower end of the literacy spectrum and records a relatively high rate of crimes against women.

4.3.4 Education and Social Vulnerabilities

Child Marriage and Child Trafficking Crime Rates by Year

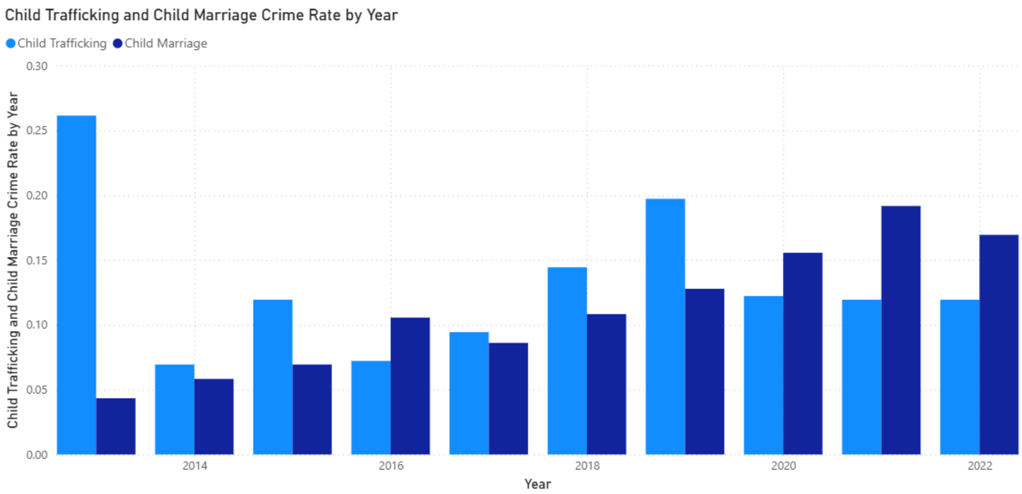

Figure 4.3.4.1

The graph displays annual crime rates for child trafficking and child marriage from 2013 to 2022. Child trafficking recorded its highest rate in 2013, followed by a significant increase in 2019. Child marriage rates remained lower throughout but showed a gradual rise after 2018. By 2022, both crime rates had risen and were closely aligned. The data reflects year-on-year fluctuations across both categories.

Dropout Rates, Child Trafficking and Child Marriage

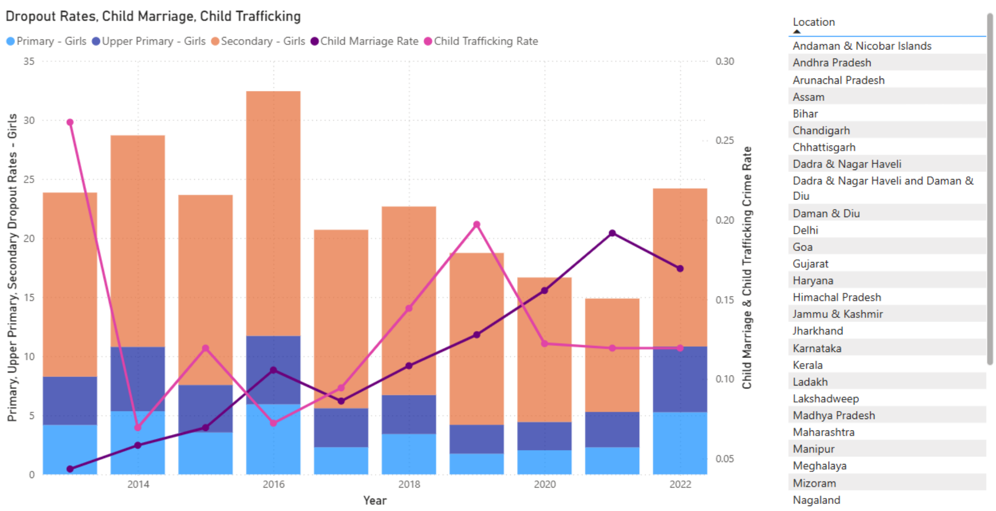

Figure 4.3.4.2

The figure combines a stacked-bar display of girls' dropout rates (primary, upper-primary, secondary) with line plots of child-marriage and child-trafficking rates for Indian states between 2014 and 2022. The left y-axis represents dropout levels, while the right y-axis shows the two crime indicators. Over the examined period, dropout levels exhibit a steady decline, whereas child-marriage trends upward modestly and child-trafficking declines sharply before stabilising. A sidebar lists the 36 states and union territories, allowing direct regional comparison of education outcomes and gender-related crimes.

Child Marriage, Child Trafficking and Dropout Rates by Year

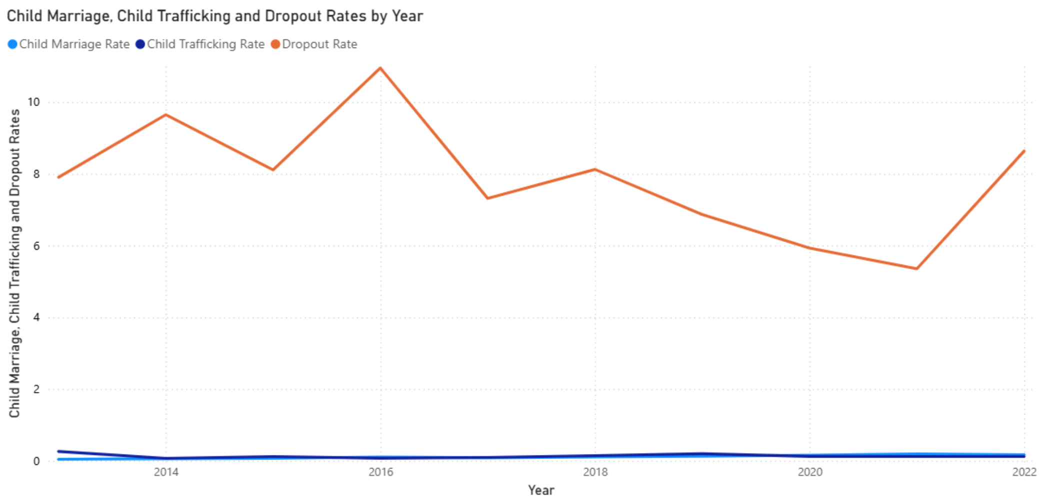

Figure 4.3.4.3

The graph shows annual rates of child marriage, child trafficking, and school dropouts from 2013 to 2022. Dropout rates fluctuated significantly, peaking around 2016 and dipping in 2021 before rising again. In contrast, child marriage and trafficking rates remained relatively low and stable throughout the period. All three indicators are plotted to highlight comparative trends over the decade.

Normalized Child Marriage, Child Trafficking and Dropout Rates by Year

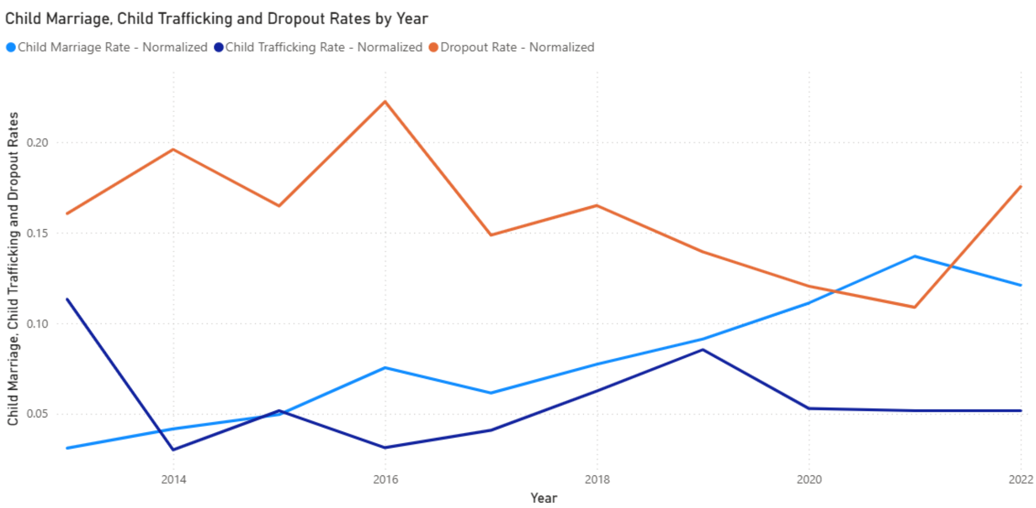

Figure 4.3.4.4

This graph is the normalized version of the previous one, showing child marriage, child trafficking, and dropout rates from 2013 to 2022. Normalization allows for direct comparison of trends across variables with different scales. The dropout rate remains consistently higher and more volatile, peaking in 2016 and again in 2022. Child marriage shows a steady rise until 2021, while child trafficking fluctuates, with a sharp peak in 2019 followed by a decline.

4.4 Correlation Analysis Results

Pearson and Spearman correlation analyses were used to examine the relationships among the study variables. These methods allowed assessment of both linear and rank-based associations. The correlation coefficients and corresponding p-values are reported for each variable pair to indicate the direction and statistical significance of the observed relationships.

4.4.1 Dropout Rates vs. Juvenile Crimes with insights

Correlation analysis was used to examine the association between dropout rates and juvenile apprehension rates. The Pearson correlation coefficient was 0.669, with a p-value of 0.0344, indicating a statistically significant positive linear relationship. his result suggests that increases in dropout rates are associated with corresponding increases in juvenile crime rates. The Spearman correlation coefficient was 0.467 with a p-value of 0.1739. This indicates a positive monotonic relationship, but the result is not statistically significant.

The findings indicate that the linear correlation measured by Pearson is statistically significant. The results show that higher dropout rates are associated with higher juvenile crime rates within the observed data.

4.4.2 Crimes Against Women vs Female Literacy with insights

The analysis examined the relationship between crimes against women and female literacy rates using Pearson and Spearman correlation methods. The Pearson correlation coefficient was -0.243 with a p-value of 0.1414, indicating a negative linear relationship that is not statistically significant. The Spearman correlation coefficient was -0.285 with a p-value of 0.0826, also indicating a negative monotonic relationship that is not statistically significant.

These results show that, within the observed data, higher female literacy rates are associated with lower crime rates against women, but the relationships are not statistically significant under either method.

4.4.3 Dropouts and Child Marriage/Trafficking Incidents with insights

The correlation between dropout rates and child marriage was assessed using Pearson and Spearman methods. The Pearson coefficient was -0.489 with a p-value of 0.152, and the Spearman coefficient was -0.401 with a p-value of 0.25. Both results indicate a moderate negative relationship, but neither is statistically significant.

For dropout rates and child trafficking, the Pearson coefficient was -0.341 with a p-value of 0.335, and the Spearman coefficient was -0.445 with a p-value of 0.197. These results also show a weak to moderate negative correlation without statistical significance, suggesting that other contextual factors may influence the observed patterns.

4.5 Discussion of Key Findings

The analyses showed clear patterns among the variables. Dropout rates and juvenile crimes had a statistically significant positive correlation, meaning higher dropout levels were linked with more juvenile apprehensions. Female literacy and crimes against women showed negative correlations, but these were not significant. Dropout rates and child marriage or trafficking also showed moderate negative correlations without significance.

Overall, the outcomes appear to be shaped by broader social, economic, and cultural factors. These influences likely shape the observed patterns more than single indicators alone.

Chapter 5: Conclusion

5.1 Summary of Findings

The study found a statistically significant positive correlation between dropout rates and juvenile crimes, indicating that higher dropout levels were linked with increased juvenile apprehensions. Other relationships, such as female literacy with crimes against women and dropout rates with child marriage or trafficking, were negative but not statistically significant.

The descriptive analysis highlighted key patterns. Delhi consistently recorded the highest crime rates, standing out as a major outlier. Crimes against women and theft accounted for the largest share of reported offences, while murder and bodily harm were comparatively lower. Dropout levels appeared to influence juvenile crime rates, and child marriage and trafficking, though lower overall, showed fluctuating trends across the years.

5.2 Contribution of the Study

This study highlights the link between dropout rates and juvenile crime, offering statistical evidence of their association. It adds to existing research by combining correlation analysis with descriptive patterns, showing how education outcomes intersect with broader social issues.

5.3 Limitations

Relationships between dropout rates and crime may be influenced by broader socioeconomic and structural factors such as poverty, urbanization, unemployment, regional disparities, and the effectiveness of education and law enforcement. These findings should be viewed as indicative rather than definitive, with the model serving as an exploratory framework that cannot fully capture the complexity of these interactions.

5.4 Future Work

Future studies should use methods that go beyond correlations to explore causal links between education and crime. Including more socioeconomic and cultural factors, and combining statistical analysis with qualitative insights, would provide a fuller understanding of these relationships.

References

[1] A. Sanwaria, "Poverty and Crime: Impact of socioeconomic factors on crime in India," vol. 05, no. 01, 2018.

[2] A. Senthivel, D. Madurai, M. L. Valan and G. E. Richards, "The ripple effect of school dropout: exploring the nexus between school dropout, deviance, and subsequent arrest amongst children in India," *Journal of Criminological Research, Policy and Practice,* vol. 10, no. 2, 2024.

[3] E. Rizzo, "Education and Crime in India: a District Level Analysis," 2017. [Online]. Available: https://www.isid.ac.in/~epu/acegd2017/papers/ElisaRizzo.pdf.

[4] R. Dubbudu, "Beyond Dashboards: How India Can Move from Data Displays to Data Dialogue in the AI era," 2 November 2025. [Online]. Available: https://www.linkedin.com/pulse/beyond-dashboards-how-india-can-move-from-data-displays-dubbudu-pn40c.

[5] A. Bharadwaj, "Is poverty the mother of crime? Empirical evidence of the impact of socioeconomic factors on crime in India," 2014. [Online]. Available: https://www.econstor.eu/bitstream/10419/146581/1/792110218.pdf.

[6] B. Rakshit and Y. Neog, "Does higher educational attainment imply less crime? Evidence from the Indian states," 2020. [Online]. Available: https://ideas.repec.org/a/eme/jespps/jes-05-2019-0218.html.

[7] B. Kishan, "EVALUATING THE EFFECTIVENESS OF INDIA'S CRIMINAL JUSTICE MECHANISMS," 31 July 2025. [Online]. Available: https://eprajournals.com/IJSR/article/17208.

[8] Lawful Legal, " A Critical Examination of Punishment in the Indian Legal System," 12 August 2024. [Online]. Available: https://lawfullegal.in/a-critical-examination-of-punishment-in-the-indian-legal-system/.

[9] M. Solanki, "Exploring The Factors Affecting Dropout Rates In Higher Education Institutions In India," 1 January 2025. [Online]. Available: https://ijabs.niilmuniversity.ac.in/exploring-the-factors-affecting-dropout-rates-in-higher-education-institutions-in-india/.

[10] Smile Foundation, "School Dropouts Interfering with Desired Progress in India," 4 December 2024. [Online]. Available: https://www.smilefoundationindia.org/blog/school-dropouts-interfering-with-desired-progress-in-india/.

[11] P. R. Kuwar, "Impact of Education on Economy," January 2021. [Online]. Available: https://www.researchgate.net/publication/361261179_Impact_of_Education_on_Economy.

[12] G. A. Crews, "Education and Crime," 2009. [Online]. Available: https://scholarworks.utrgv.edu/cgi/viewcontent.cgi?article=1064&context=cj_fac.

[13] L. Lochner, "Chapter 9 - Education and crime," 2020. [Online]. Available: https://www.sciencedirect.com/science/chapter/edited-volume/abs/pii/B9780128153918000094.

[14] W. Groot and H. Maassenvandenbrink, "The effects of education on crime," February 2010. [Online]. Available: https://www.researchgate.net/publication/46528773_The_effects_of_education_on_crime.

[15] L. Sugiharti, R. Purwono, M. A. Esquivias and H. Rohmawati, "The Nexus between Crime Rates, Poverty, and Income Inequality: A Case Study of Indonesia," 2023. [Online]. Available: https://www.mdpi.com/2227-7099/11/2/62.

[16] G. M. S, "THE ROLE OF DATA ANALYTICS IN ECONOMIC POLICY MAKING," September 2017. [Online]. Available: https://www.jetir.org/papers/JETIR1709144.pdf.

[17] Digital India, "Digital India: Power to Empower," [Online]. Available: https://www.digitalindia.gov.in/about-us/.

[18] Expressions Limitless, "Visualizing India: Unveiling Data's Secrets in a Digital Age," 3 November 2023. [Online]. Available: https://www.linkedin.com/pulse/visualizing-india-unveiling-datas-secrets-digital-4ntkf?trackingId=nOGRMrobUuVI6PGV88Ee9A%3D%3D.

[19] NCRB, "Crime in India Year Wise," [Online]. Available: https://www.ncrb.gov.in/crime-in-india-year-wise.html.

[20] UDISE+, "Reports," [Online]. Available: https://dashboard.udiseplus.gov.in/.

[21] M. Stojiljković, "NumPy, SciPy, and pandas: Correlation With Python," [Online]. Available: https://realpython.com/numpy-scipy-pandas-correlation-python/#pearson-correlation-coefficient.