2211CS010124
CH. Hanuma Pavani

CrimeAnalysisIndia_2002-2012

1. Dataset Description

1.1 Source: District-wise Crimes in India (2001–2012) dataset obtained from Kaggle (public government dataset).

1.2 Columns:

- State/UT – Name of the state or union territory.
- District – Name of the district.
- Year – Year of crime record (2001–2012).
- Murder – Number of murder cases reported.
- Attempt to Murder – Number of attempted murder cases.
- Rape – Number of rape cases reported.
- Kidnapping and Abduction – Number of kidnapping and abduction cases.
- Dacoity – Number of dacoity (armed robbery) cases.
- Robbery – Number of robbery cases.
- Burglary – Number of burglary cases.
- Theft – Number of theft cases.
- Auto Theft – Number of vehicle theft cases.
- Other IPC Crimes – Number of other crimes under Indian Penal Code (IPC).
- Total Cognizable Crimes – Total number of cognizable crimes reported.

1.3 Data Quality:

- No missing values after cleaning.
- Clean and consistent data types.
- Balanced structure with district-level and state-level distribution.
- Covers 640+ districts across multiple states over 12 years, ensuring both temporal and geographic diversity

2. Operations Performed

2.1 Data Cleaning & Exploration

- Verified dataset integrity — no missing or null values found after loading.
- Checked unique values in categorical columns (State/UT, District, Year).
- Ensured numerical columns (Murder, Rape, Theft, etc.) contain valid non-negative integers.
- Summarized numerical columns using statistical measures — mean, median, standard deviation, and IQR to understand distribution.
- Identified districts and states with maximum and minimum crime counts.

2.2 Descriptive Analytics

- State-wise total IPC crimes calculated and visualized (bar chart).
- Year-wise crime trend analyzed to observe rise or fall in total crimes (line chart).
- Distribution of total crimes across states visualized (histogram).
- Share of top 5 states in total crimes represented (pie chart).
- Trends of selected crimes such as Murder, Rape, and Theft shown over years.

2.3 Relationship Analysis

- Comparison of major crime types (Murder, Rape, Theft) across years.
- Identification of top states contributing highest number of crimes.
- Observation of yearly variation patterns among different crime categories.
- Highlighted correlations and visual relationships using trend and distribution plots.

3. **Key Insights**

3.1 Crime Distribution Overview

• Crime rates vary significantly across Indian states and union territories.
• States like Uttar Pradesh, Maharashtra, and Madhya Pradesh recorded the highest total IPC crimes.
• Northeastern and smaller states contributed minimal crime counts.
• Overall increase in total crimes observed over the 2001–2012 period.

3.2 Yearly Crime Trends

• Gradual rise in total IPC crimes from 2001 to 2012.
• Certain years (e.g., 2006–2010) showed sharper increases, possibly linked to population growth or reporting improvements.
• Trendline indicates consistent upward movement across the decade.

3.3 State-Level Insights

• Uttar Pradesh consistently reported the highest number of IPC crimes across years.
• Maharashtra and Madhya Pradesh followed closely with significant contributions.
• Southern states like Tamil Nadu and Andhra Pradesh showed moderate but steady trends.
• Union Territories recorded comparatively fewer crimes.

3.4 Crime Type Trends

• Theft and Rape cases showed a steady increase, reflecting urbanization and awareness.
• Murder rates remained relatively stable over the years.
• Variations across crime categories highlight the shifting nature of reported offenses.

3.5 Overall Observations

• Data shows strong regional disparities in crime rates.
• Top 5 states together accounted for a major share of total IPC crimes.
• The increasing trend across most categories emphasizes the need for targeted policy interventions.
• Visual patterns make it easy to identify areas and years with high crime incidence.

4. **Recommendations**

4.1 Crime Prevention & Awareness

• Launch community awareness programs in states with high crime rates to promote public safety and reporting.
• Strengthen local policing and surveillance systems, especially in urban and high-density areas.
• Encourage education and rehabilitation programs to reduce repeat offenses.

4.2 Policy & Law Enforcement

• Implement stricter law enforcement measures in high-crime states like Uttar Pradesh and Maharashtra.
• Allocate additional resources to states showing consistent crime growth over the years.
• Regularly review and update IPC laws to address emerging crime patterns such as cyber and economic crimes.

4.3 Data-Driven Policing

• Utilize predictive analytics to identify potential crime hotspots using historical data trends.
• Develop dashboards for real-time monitoring of crime data by state and year.
• Integrate PySpark-based analysis tools into police record systems for automated insights.

4.4 Regional Focus

• Focus on regions with rapid increases in crime rates for targeted interventions.
• Support smaller states and union territories with limited law enforcement infrastructure.
• Encourage inter-state collaboration to tackle cross-border criminal activities.

4.5 Future Analytical Opportunities

• Extend the analysis by including socio-economic indicators (e.g., literacy rate, unemployment) to understand crime causes.
• Apply clustering or correlation models to identify patterns among different crime types.
• Use forecasting techniques to predict future crime trends and aid in preventive policymaking.