DVA Capstone Project Report

Title: Crime Trend and Risk Analysis Across Indian States (2001–2012)

Sector: Law & Justice

Team Members:

Enrollment No.	Name	Course
2401010408	Samarth Sangtani	BTech CS & AI
2401010490	Vani Rudra	BTech CS & AI
2401020033	Kushal Sarkar	BTech CSE & DS
2401010106	Aryan Yadav	BTech CS & AI
2401010508	Vriha Dholiya	BTech CS & AI
2401010486	Vaibhav	BTech CS & AI

Institute: Newton School Of Technology (Rishihood University)

Faculty: Aayushi Vashishth

2. Executive Summary

Problem

India faces rising pressure for data-driven policing and efficient allocation of law enforcement resources. Crime distribution is uneven across states and districts, and growth patterns vary significantly over time. Decision-makers such as State Home Departments and Police Commissioners require structured analytical tools to identify hotspots, detect growth signals, and prioritize enforcement strategies.

Objective

To develop an executive decision-support dashboard that identifies:

- High-intensity hotspot districts
- Fastest-rising states by growth rate
- Crime categories driving total crime growth
- Long-term national crime trends

Approach

Using district-wise IPC crime data (2001–2012), covering 803 districts, we:

- Cleaned and standardized 8,596+ records
- Engineered KPIs including YoY growth and CAGR
- Conducted offense-share decomposition
- Built hotspot rankings
- Developed combined trend vs. growth visualizations
- Designed an interactive Google Sheets dashboard

Key Findings

- Total IPC crimes increased from approximately 4.9M in 2001 to 6.3M in 2012.
- 4,02,543 murder cases and 2,39,137 rape cases were recorded during this period.
- Top 10 districts contribute a disproportionate share of total crimes.
- Lower-volume states sometimes show highest growth rates (early-warning signals).
- Offense mix shifted over time, indicating changing intervention priorities.

Strategic Value

This dashboard transforms static historical data into a dynamic resource allocation tool, enabling targeted, cost-neutral enforcement strategies.

---

3. Sector & Business Context

Sector Overview

The Law & Justice sector is responsible for crime prevention, investigation, and public safety. Efficient resource allocation is critical due to limited personnel, funding, and infrastructure.

Decision-Maker Context

Primary stakeholders include:

- State Home Departments
- Police Commissioners
- District-level law enforcement authorities

They require:

- Identification of chronic hotspots
- Growth trend alerts
- Offense-type prioritization

Why This Problem Matters

Blanket distribution of policing resources is inefficient. Concentrated and data-prioritized allocation leads to higher operational impact and improved response times.

---

4. Problem Statement & Objectives

Problem Statement

Which districts and offense types drive IPC crime growth, and where should enforcement resources be prioritized?

Project Scope

- Time Period: 2001–2012
- Geographic Scope: 803 districts
- Crime Type: Majority crimes
- Tool Used: Google Sheets (cleaning + dashboard)

Success Criteria

- Identify persistent hotspots
- Detect fastest-growing states
- Understand offense mix evolution
- Provide actionable recommendations

---

5. Data Description

Dataset Source

Source: Kaggle – Crime in India Dataset
Link: https://www.kaggle.com/datasets/rajanand/crime-in-india

Crime-Trends-and-Hotspots-Analy…

Data Size

- Rows: 8,596
- Time Period: 2001–2012
- Districts Covered: 803

Key Variables

- State/UT
- District
- Year
- IPC Crimes
- Murder
- Rape
- Theft

- Hurt/Grievous Hurt
- Other IPC Crimes

Aggregate Metrics (2001–2012)

- Murders: 4,02,543
- Rape Cases: 2,39,137
- Total Crimes: 6,50,99,519

Data Limitations

- No population denominator (cannot compute per capita rates)
- No conviction/arrest data
- Possible reporting bias
- No socio-demographic variables

---

6. Data Cleaning & Preparation

Cleaning steps (aligned with slide 3 and cleaning documentation):

- Standardized column names
- Corrected text case formatting
- Removed duplicates
- Fixed data types
- Dropped rare/low-impact columns
- Created Total Crime KPI
- Updated district naming inconsistencies
- Merged spelling variations
- Sorted dataset for consistency

All transformations performed in Google Sheets.

---

7. KPI & Metrics Framework

KPIs Measured

1. Total Crimes
2. YoY Growth %
3. CAGR (2001–2012)
4. Crime by State
5. Offense Mix Share
6. Violent vs Property Crime Trends

Why These KPIs?

They directly inform:

- Resource allocation
- Tactical intervention planning
- Early-warning monitoring
- Strategic budgeting

---

8. Exploratory Data Analysis (EDA)

1. National Trend Patterns

Total crimes increased from 4.9M (2001) to 6.3M (2012).

Trend chart shows:

- Sharp spike around 2004
- Short dip in 2005
- Steady rise post-2006

This indicates structural upward pressure.

---

2. Concentration Insight

Top 10 states account for a disproportionate share of national crime totals.

This suggests:

Targeted allocation is more effective than equal distribution.

---

3. Persistent Hotspots

Certain districts (e.g., Thane, Bangalore, Pune) consistently rank high across multiple years.

These are chronic high-intensity zones requiring structural intervention.

---

4. Growth Signals

YoY growth analysis identifies fastest-rising states:

- West Bengal
- Tripura
- Assam

These are early-warning targets.

---

5. Offense Mix Shift

Offense-share decomposition shows:

- Theft and property crimes dominate totals.
- Violent crime share trends show gradual increase.
- Crime against women shows steady upward trajectory.

Intervention priorities must evolve accordingly.

---

9. Advanced Analysis & Methods

Methods Used

- YoY Growth %
- CAGR
- Offense-share decomposition

- Hotspot ranking
- Combined Trend vs YoY chart

New Understanding

- Lower-volume states can exhibit highest growth rates.
- Growth spikes do not always align with highest volume states.
- Combined trend + YoY chart enables strategic timing of intervention.

---

10. Dashboard Design

Executive View

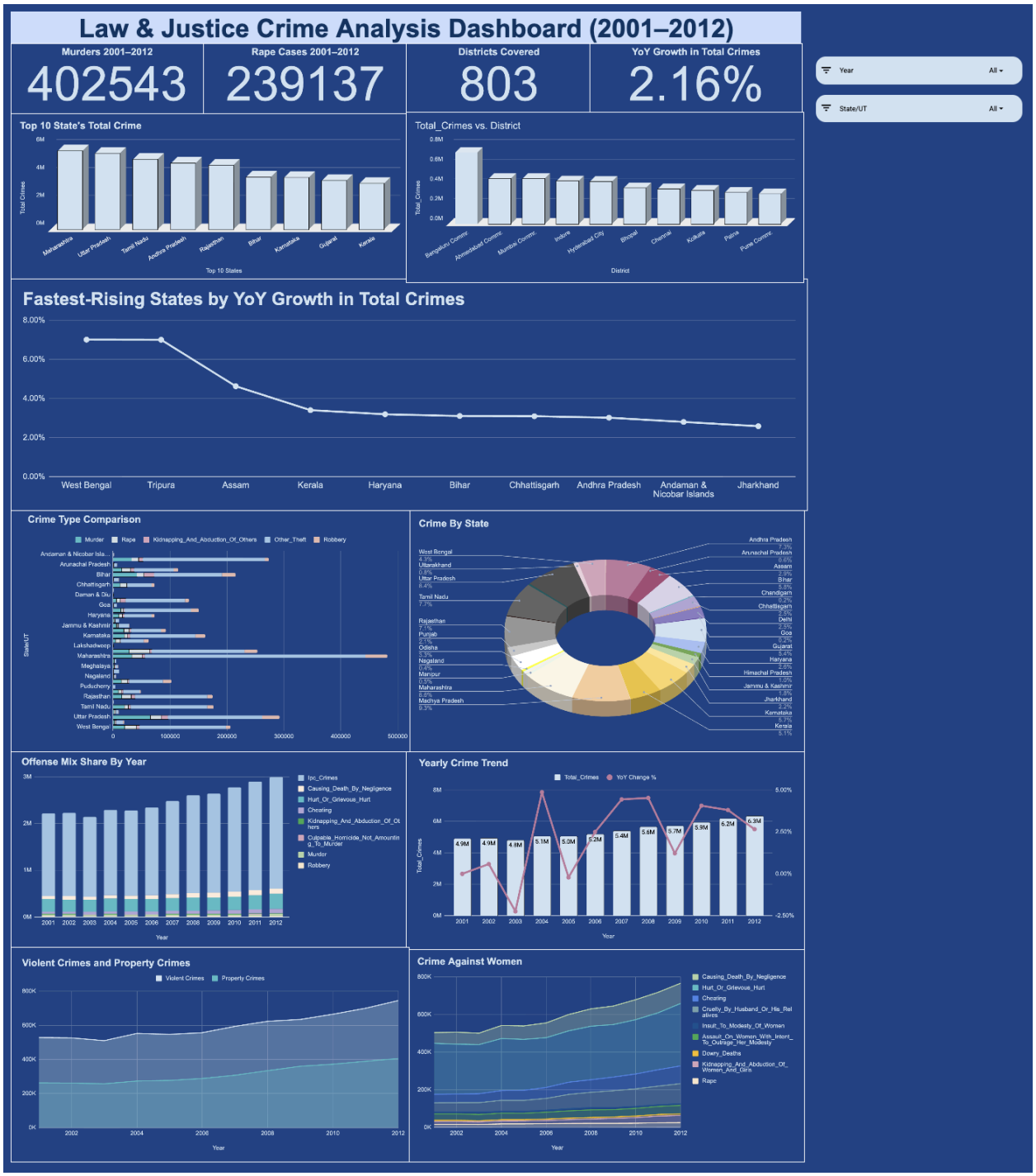
- Top-line metrics (Murders, Rape, Districts Covered, YoY Growth)
- National trend chart
- Top 10 states

Operational View

- Fastest-rising states
- Offense mix by year
- Crime type comparison
- District hotspot ranking

Interactivity

- Year filter
- State filter



11. Insights Summary

1. National crime exhibits structural upward trend.

2. Top 10 states dominate overall volume.
3. Certain districts are chronic hotspots.
4. Lower-volume states can be early-warning signals.
5. Crime against women shows consistent growth.
6. Property crimes dominate total counts.
7. Growth volatility indicates tactical monitoring need.
8. Concentrated allocation improves efficiency.

---

12. Recommendations

1. Prioritize Top 10 Districts

Allocate ~60% enforcement resources to high-impact hotspots.

2. Establish Quarterly Monitoring Dashboard

Track trend vs YoY change for rapid tactical response.

3. Implement Early-Warning System

Focus predictive monitoring on fast-growing states (West Bengal, Punjab, Odisha).

4. Pilot Data-Driven Allocation

Test in Thane Bangalore, and Pune for 6 months.

5. Shift Strategy by Offense Type

Target theft prevention in high-volume districts and violence reduction in concentrated hotspots.

---

13. Estimated Impact

Targeted enforcement in hotspot districts could:

- Enable cost-neutral reallocation
- Improve efficiency compared to blanket deployment

---

14. Limitations

- Historical data only (2001–2012)
- No real-time data feed
- Reporting bias possible
- No contextual socio-economic variables

---

15. Next Steps

- Integrate recent year data
- Add population denominator
- Incorporate arrest/outcome data
- Implement near-real-time feed
- Run controlled pilot measurement

---

16. Conclusion

This project converts historical crime data into a strategic decision-support system. By integrating volume, growth, concentration, and offense-mix analysis, it provides a structured framework for data-driven policing and targeted resource allocation.

The dashboard bridges the gap between raw data and executive action.

---

17. Appendix

- Cleaning Log (as documented)
  cleaned
- Pivot Tables

- CAGR Calculations
- Dashboard Formula References

---

18. Contribution Matrix

Team Member	Dataset & Sourcing	Cleaning	KPI & Analysis	Dashboard	Report Writing	PPT	Overall Role
Samarth Sangtani	Support	Support	Co-Lead	Co-Lead	Support	Lead	Project Lead & Strategy Coordinator
Vani Rudra	Lead	Support	Lead	Lead	Support	Co-Lead	Data Sourcing & Analytics
Kushal Sarkar	Support	Lead	Co-Lead	Lead	Lead	Support	Data Cleaning & Design
Aryan Yadav		Support	Co-Lead	Co-Lead			Analytics & Visualization
Vriha Dholiya	Support		Co-Lead	Support		Co-Lead	Visualization & Presentation
Vaibhav			Support	Support			Support

Declaration: We confirm that the above contribution details are accurate and verifiable through version history and submitted artifacts.

Team Signature Block:-
Samarth Sangtani, Vani Rudra, Kushal Sarkar, Aryan Yadav, Vriha Dholiya, Vaibhav