CS661 - COURSE PROJECT

INDIAN CRIME DATA VISUALIZATION PORTAL

PRESENTED BY GROUP - 5

Akshay Toshniwal - 241110005
Ansh Makwe - 241110010
Devang Agarwal - 241110019
Roshan Kumar - 241110058
Divyansh Chaurasia - 241110022
Kunal Anand - 241110040
Parjanya Aditya Shukla - 241110046
Prakhar Mandloi - 241110051

Introduction And Agenda

- **Project Goal:** To develop a user-friendly, interactive web application consolidating diverse Indian crime datasets (primarily 2001-2013) for intuitive visualization and analysis.
- **Problem:** Difficulty in accessing, integrating, and visualizing fragmented and complex crime data in India.
- **Solution:** A Python Dash/Plotly based web portal providing geographic, temporal, comparative, and thematic analysis tools.
- **Agenda:**
  - Project Overview
  - Application Story
  - Dashboard Walkthrough (Key Tabs)
  - Conclusion & Future Work

Project Objectives

- Provide a centralized platform for diverse crime statistics.
- Enable exploration across geographic, temporal, demographic, and situational dimensions.
- Facilitate comparative analysis (regions, time periods, crime types).
- Reveal patterns, hotspots, and correlations through visualization.
- Incorporate ML (clustering, forecasting) for deeper insights.

Scope & Technologies

- **Scope:**
  - Data Integration (IPC, SC/ST, Women, Children, Juvenile, Property, etc., 2001-2013).
  - Interactive Visualizations (Maps, Charts - Line, Bar, Pie, Sunburst, Sankey, Treemap, Heatmap).
  - Multi-tabbed Dash Web Application.
  - Dynamic Filtering & Interactivity.
  - ML Features (K-Means, ARIMA, Linear Regression).

Scope & Technologies

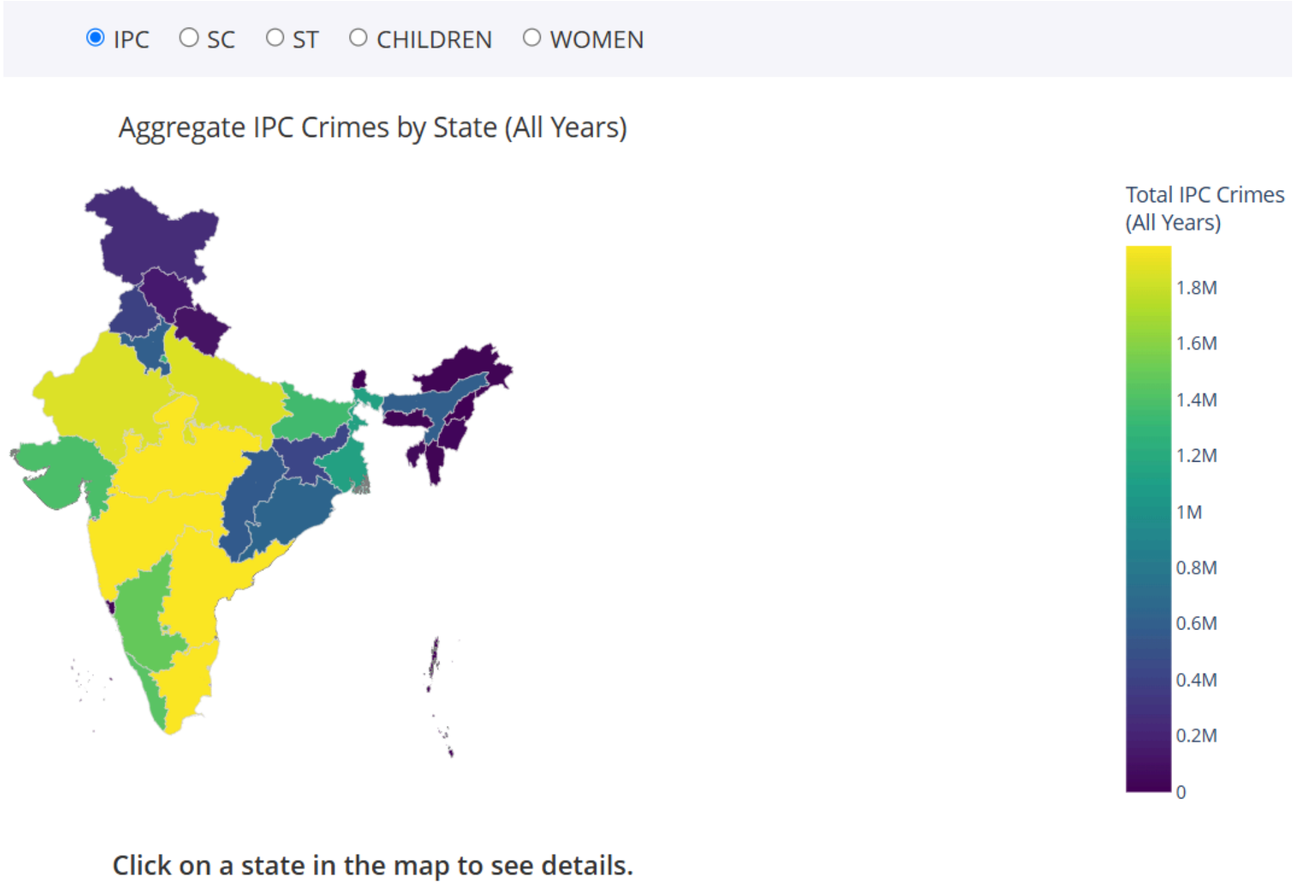
- **Technologies:**
  - Python
  - Dash & Plotly
  - Pandas & Numpy
  - Scikit-learn (KMeans, LinReg, Scaler)
  - Statsmodels (ARIMA)
  - Rapidfuzz (Name Matching)

Overall Application Story

- **Challenge:** Understanding vast, fragmented Indian crime data is difficult. Raw numbers obscure trends.
- **Our Solution:** An interactive portal unifying diverse dataset (early 2000s) into an explorable platform.
- **The Journey:**
  - Start broad (State-level map).
  - Drill down (District-level details, specific crimes).
  - Analyze trends over time (Year-wise view).
  - Investigate specifics (Place of occurrence, Victim-Offender relationships, Motives).
  - Uncover hidden patterns (Clustering districts by crime profile).
  - Explore potential futures (Forecasting).

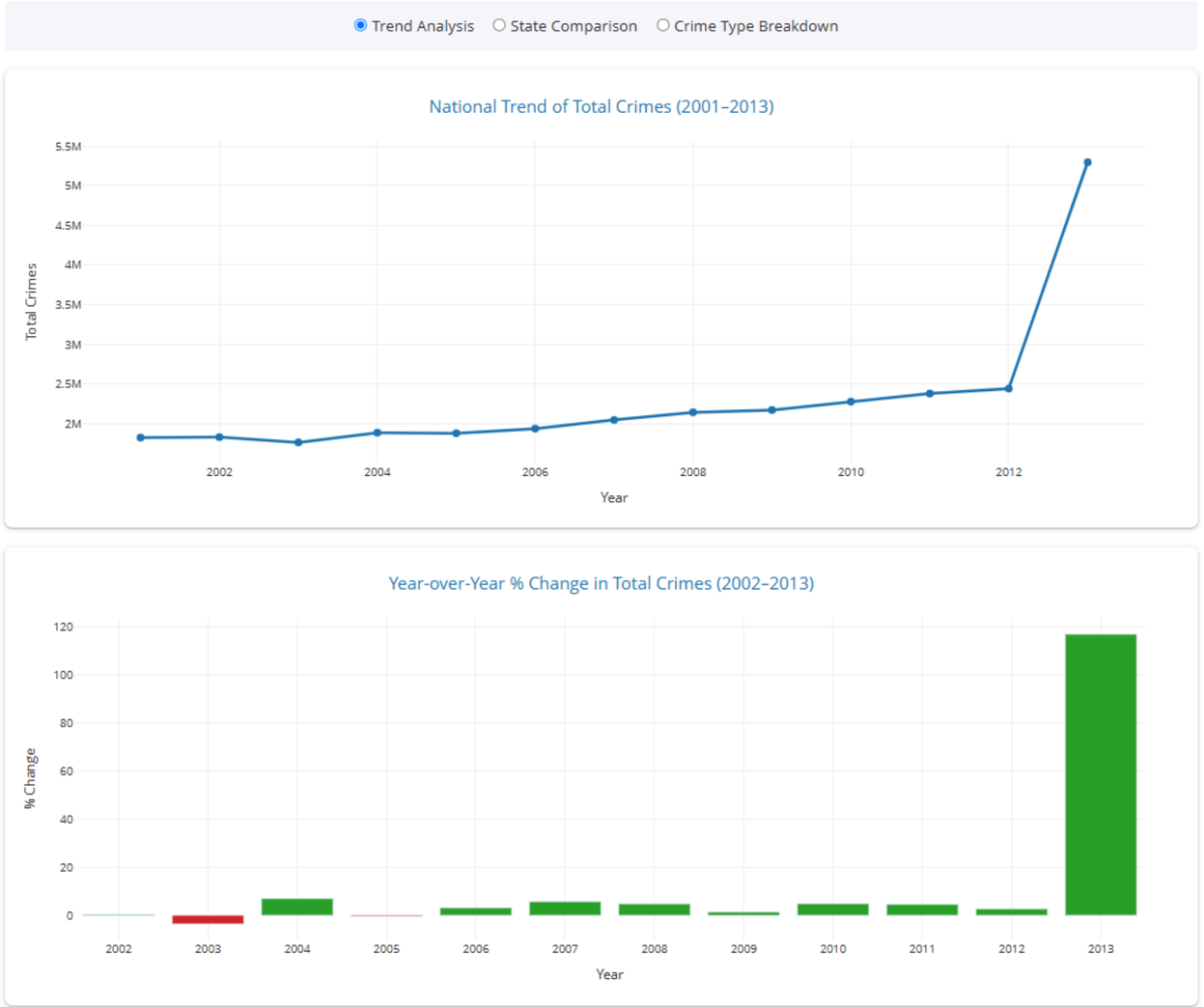
State Wise Tab

- **Core Feature:** High-level geographic overview.
- **Functionality:**
  - Interactive state-level choropleth map showing aggregated crime counts by category.
  - Click a state to drill down.
  - Analyze selected state: View district breakdown (bar chart) and yearly trend (line chart) for specific crime types/years.



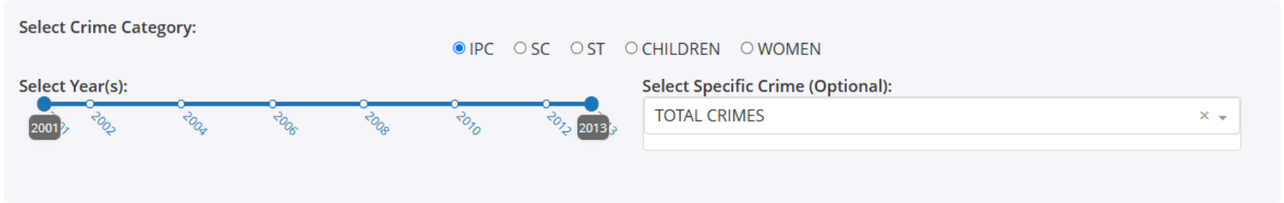
Year Wise Tab

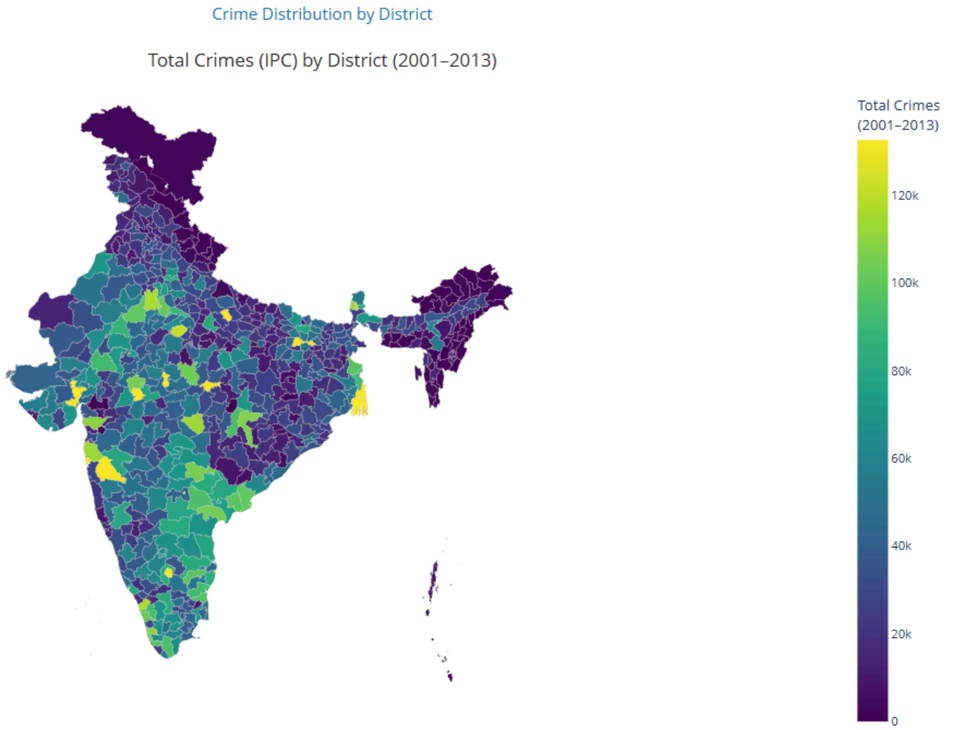
- **Core Feature:** National-level temporal analysis.
- **Functionality (Selectable Views):**
  - **Trend Analysis:** National crime trend (line) & YoY % change (bar).
  - **State Comparison:** Top states by crime count (bar) & % contribution (pie) over selected years.
  - **Crime Type Breakdown:** Proportional breakdown of crimes within a category (pie/bar).



District Wise Tab

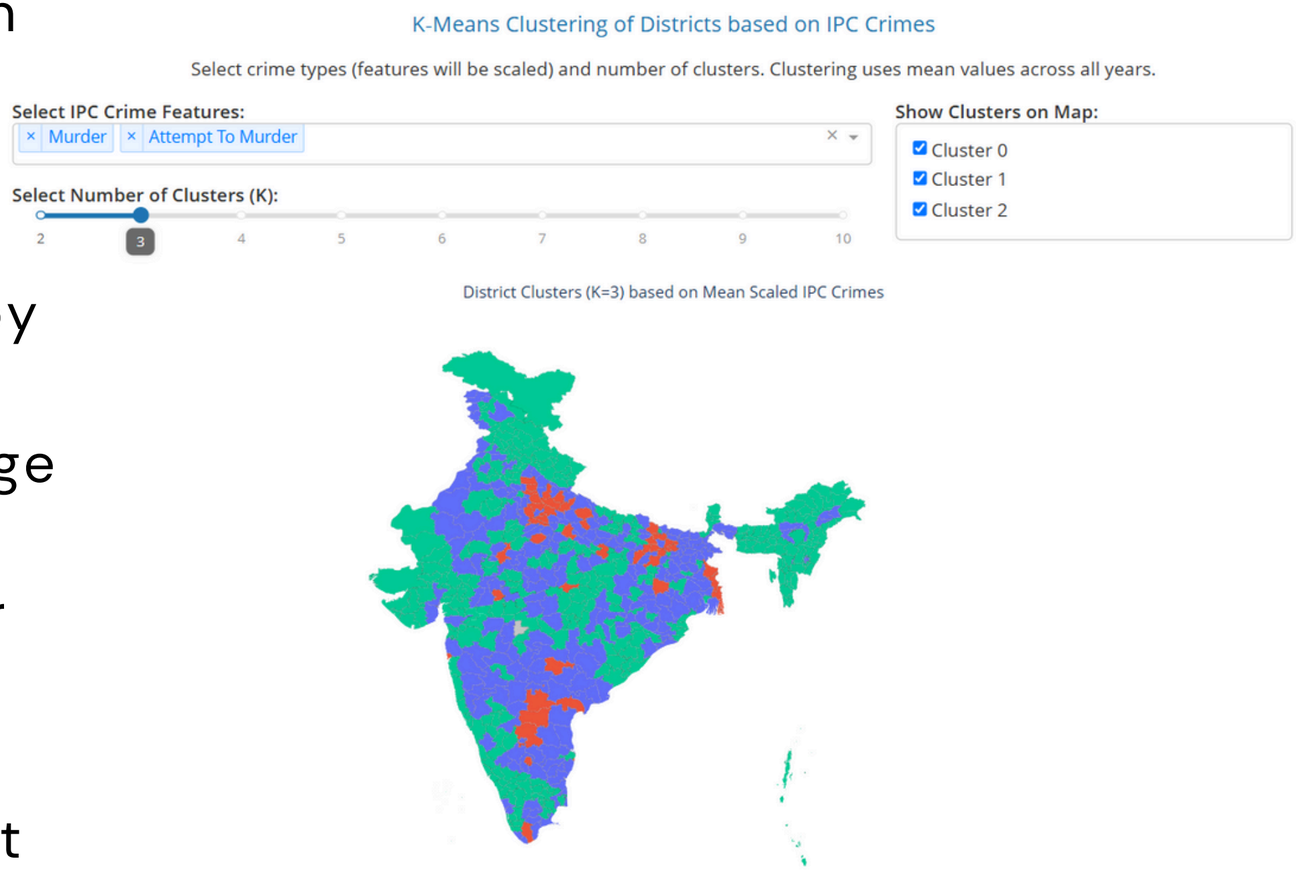
- **Core Feature:** Granular district-level geographic analysis.
- **Functionality:**
  - Interactive district-level choropleth map for selected category, crime type, and year range.
  - Click a district for detailed trend (line) & crime breakdown (pie).
  - Compare multiple selected districts (line/bar charts).
  - Identify top N districts (hotspots) for specific crimes (bar chart).





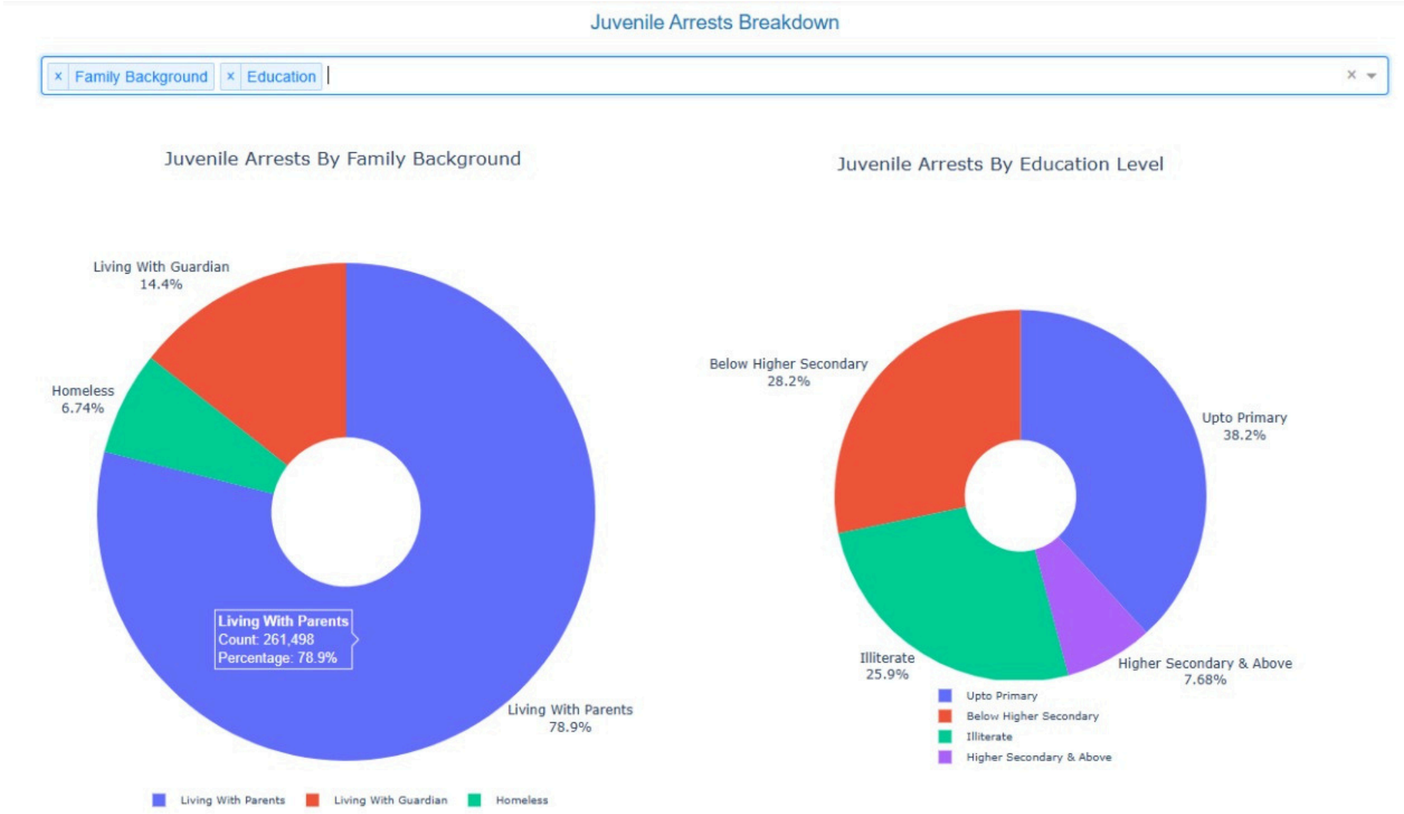
Crime Clusters (IPC) Tab

- **Core Feature:** Identify districts with similar IPC crime profiles using K-Means.
- **Functionality:**
  - **Cluster Map:** Districts colored by cluster assignment.
  - **Centroid Chart:** Compare average crime profiles of clusters.
  - **Silhouette Plot:** Evaluate cluster quality.
  - **Trend/Forecast Lines:** Show historical trend + ARIMA forecast for each cluster.



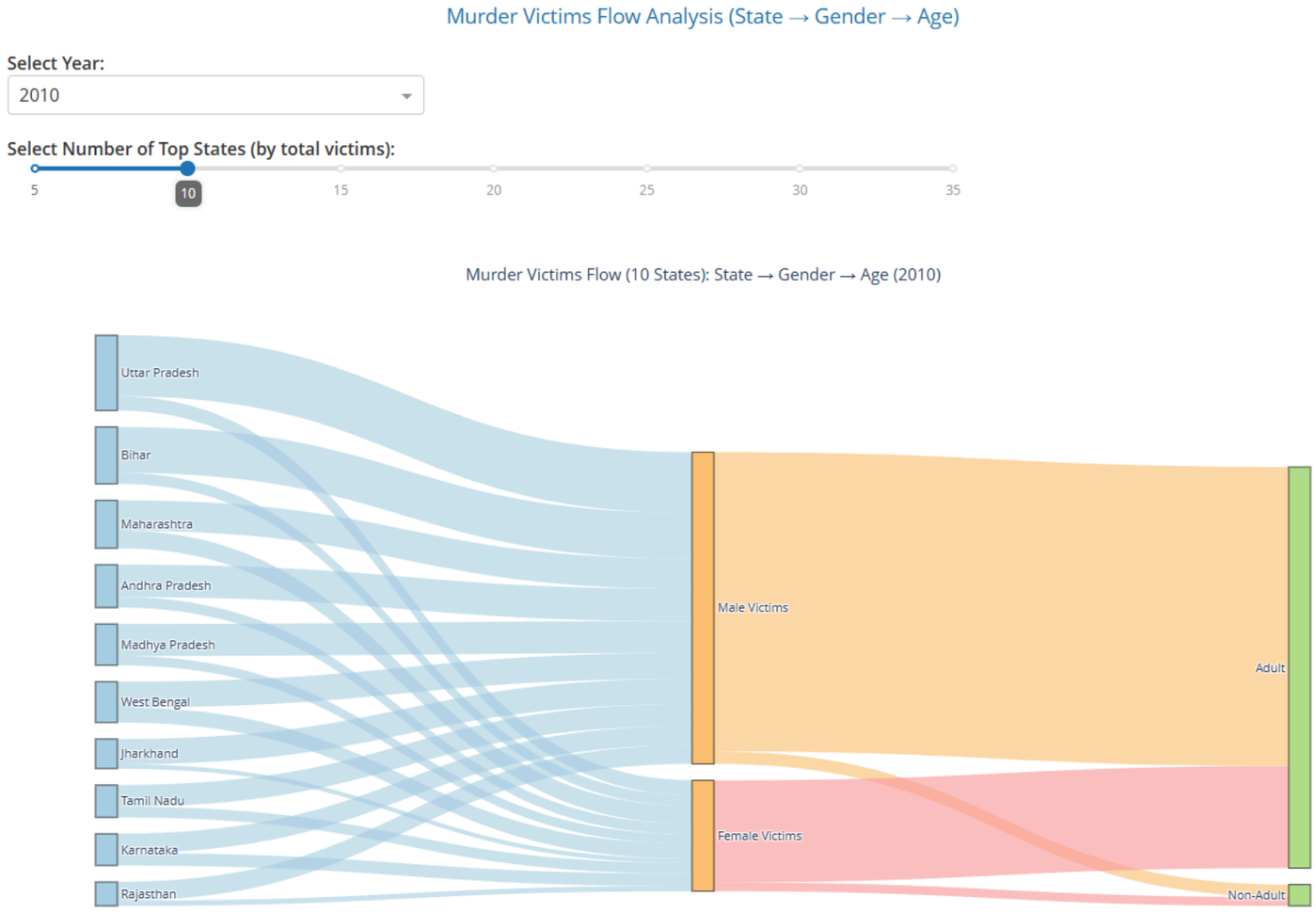
Juvenile Plots Tab

- **Core Feature:** Show background profiles of arrested juveniles.
- **Functionality:**
  - Select categories (Education, Economic Setup, Family, Recidivism).
  - Generates Pie chart for each selected category showing proportional breakdown (aggregated nationally/over time).



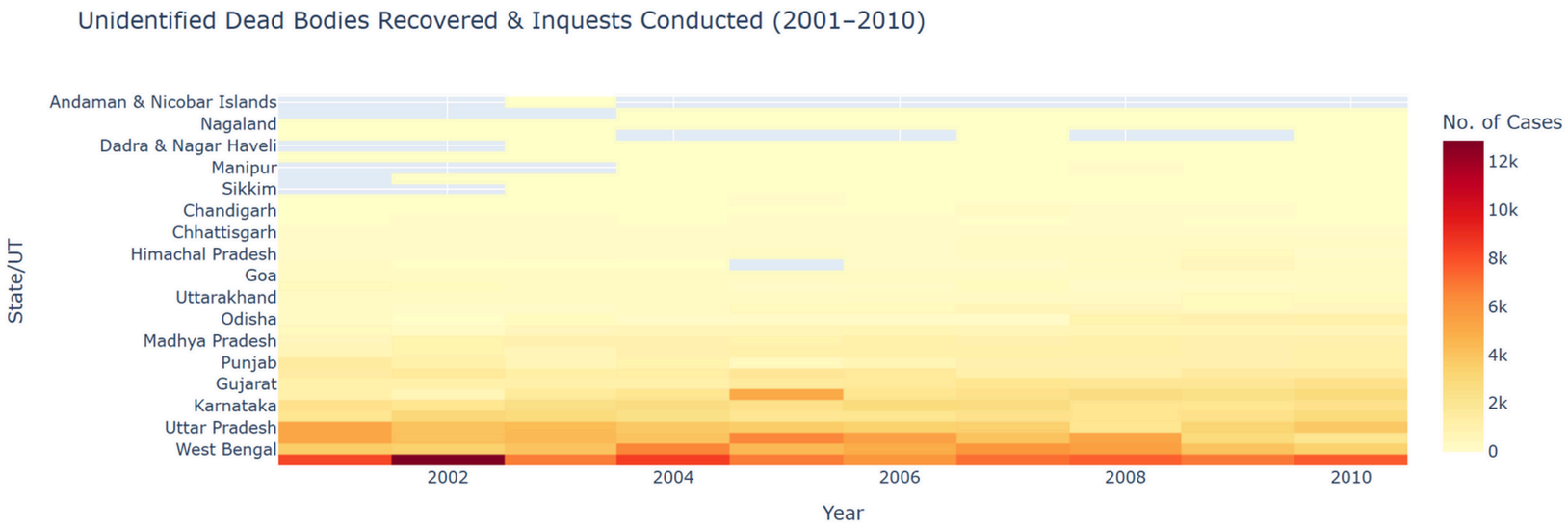
Murder Victims Flow Tab

- **Core Feature:** Visualize flow of murder victims through demographics.
- **Functionality:**
  - Sankey diagram showing flow: Top N States -> Gender -> Age Group for a selected year.
  - Link thickness represents victim counts.



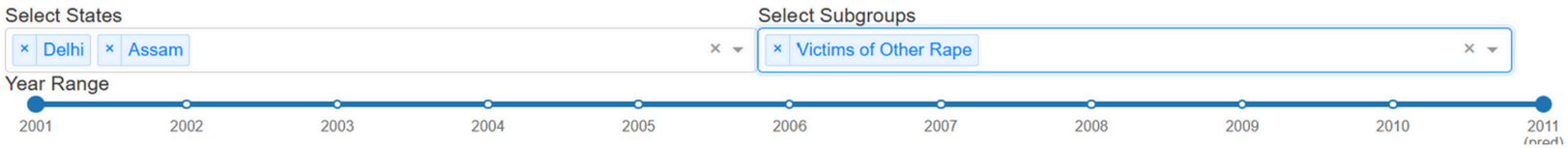
Unidentified Bodies Heatmap Tab

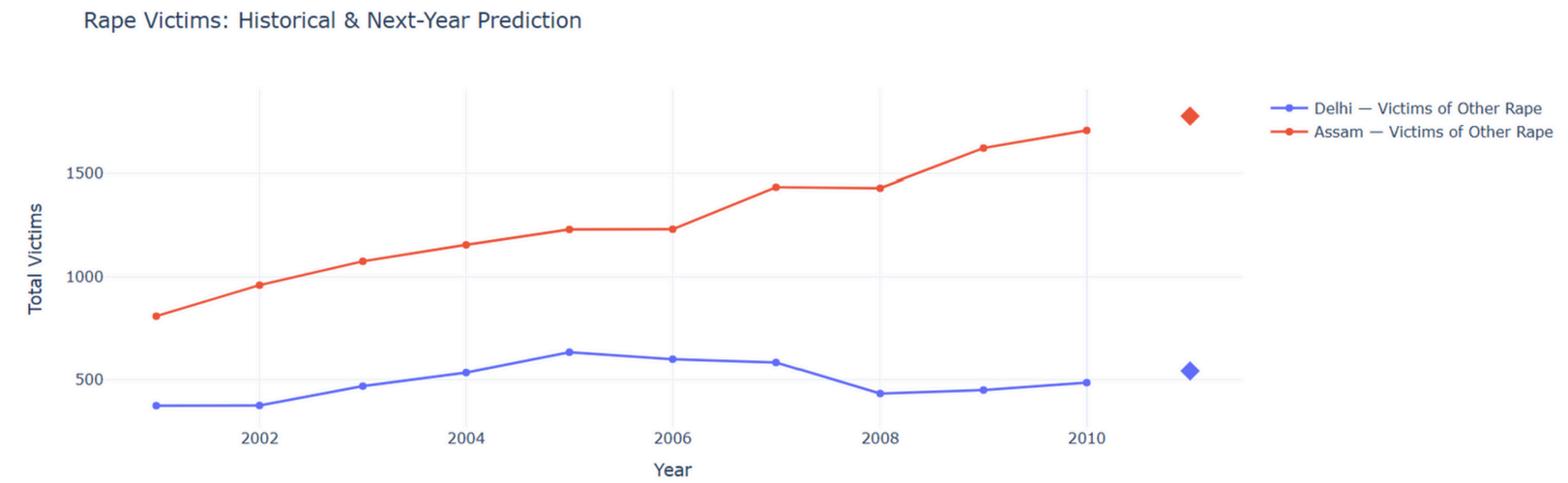
- **Core Feature:** Visualize unidentified body recoveries (2001–2010).
- **Functionality:**
  - Static Heatmap: State/UT vs. Year.
  - Color intensity shows count of recoveries/inquests.



Rape Victims Trend & Prediction Tab

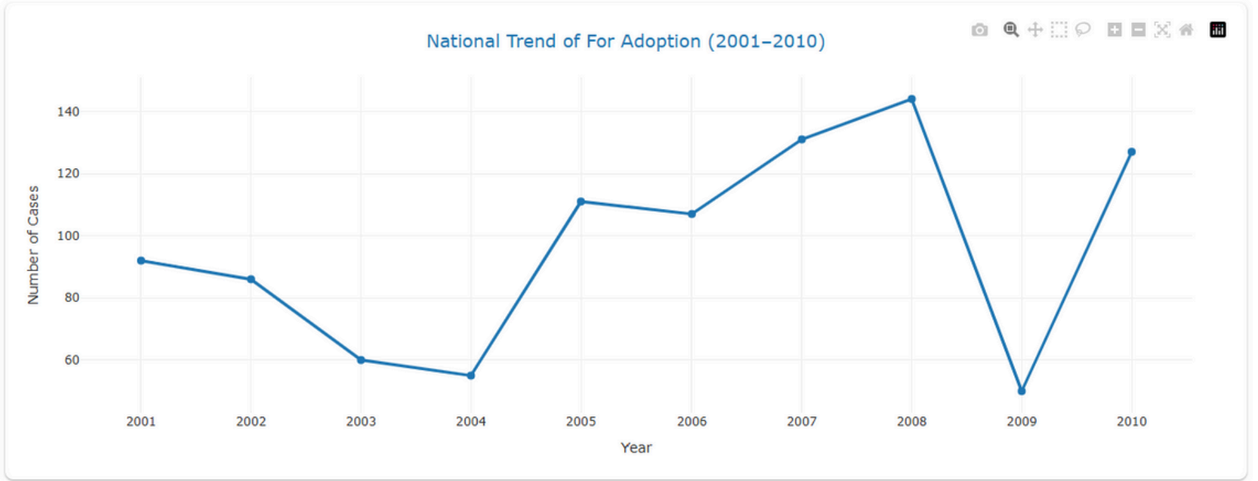
- **Core Feature:** Analyze historical rape victim trends + simple prediction.
- **Functionality:**
  - Line chart showing trends for selected state(s)/subgroup(s).
  - Includes 1-year prediction point (Linear Regression) if year range selected.

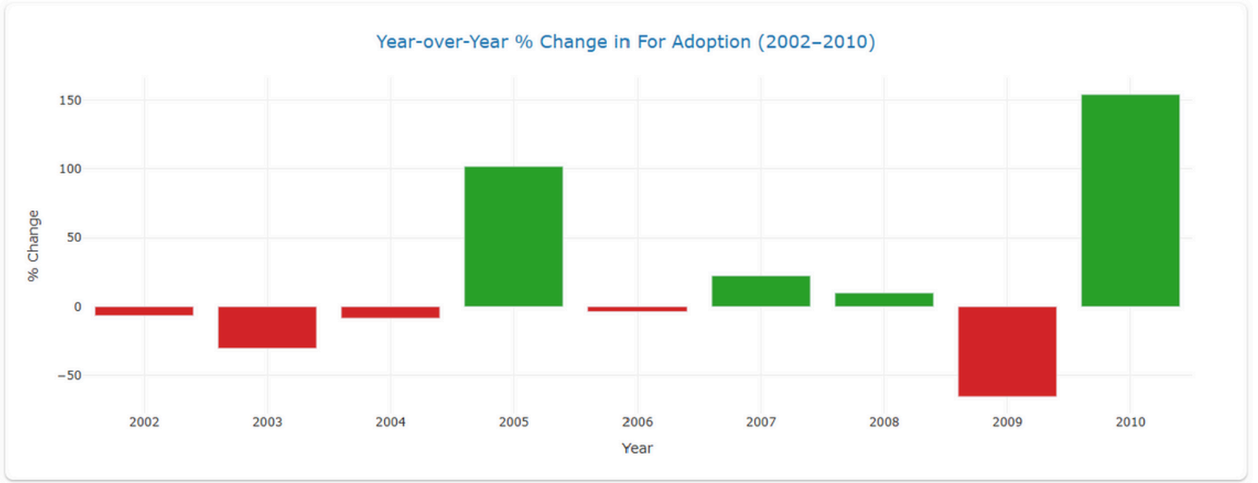




Kidnappings & Abductions Tab

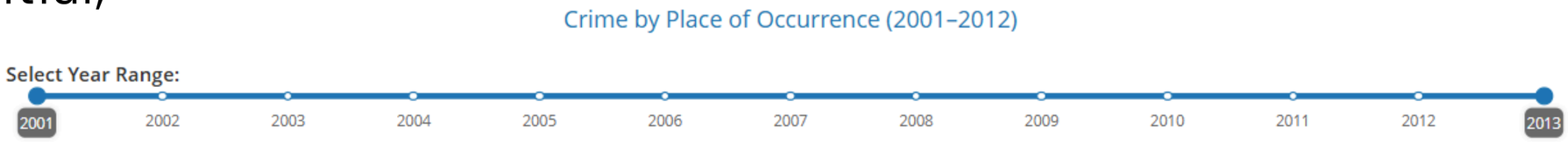
- **Core Feature:** Deep dive into kidnapping motives and victimology.
- **Functionality (Selectable Views):**
  - **Trend:** National trend/YoY change for total or specific purpose.
  - **State Comparison:** Top states (counts) by purpose.
  - **Purpose Profile:** Heatmap comparing % breakdown of motives across states.
  - **Victim Demographics:** Age/Gender breakdown for selected purpose/location.

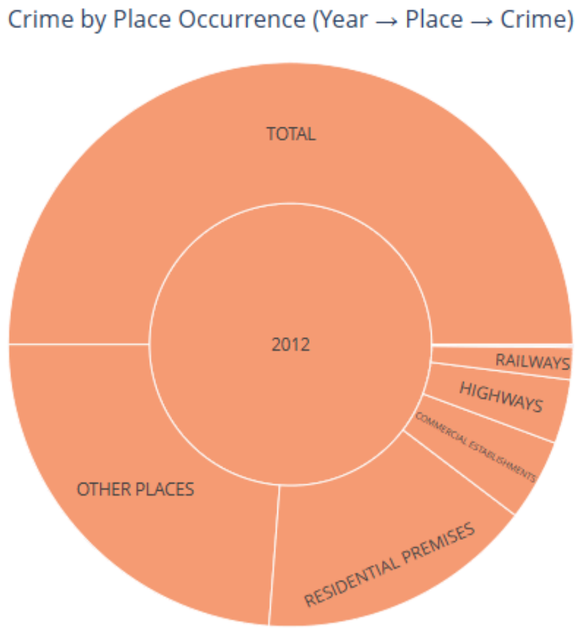




Place Occurrence Tab

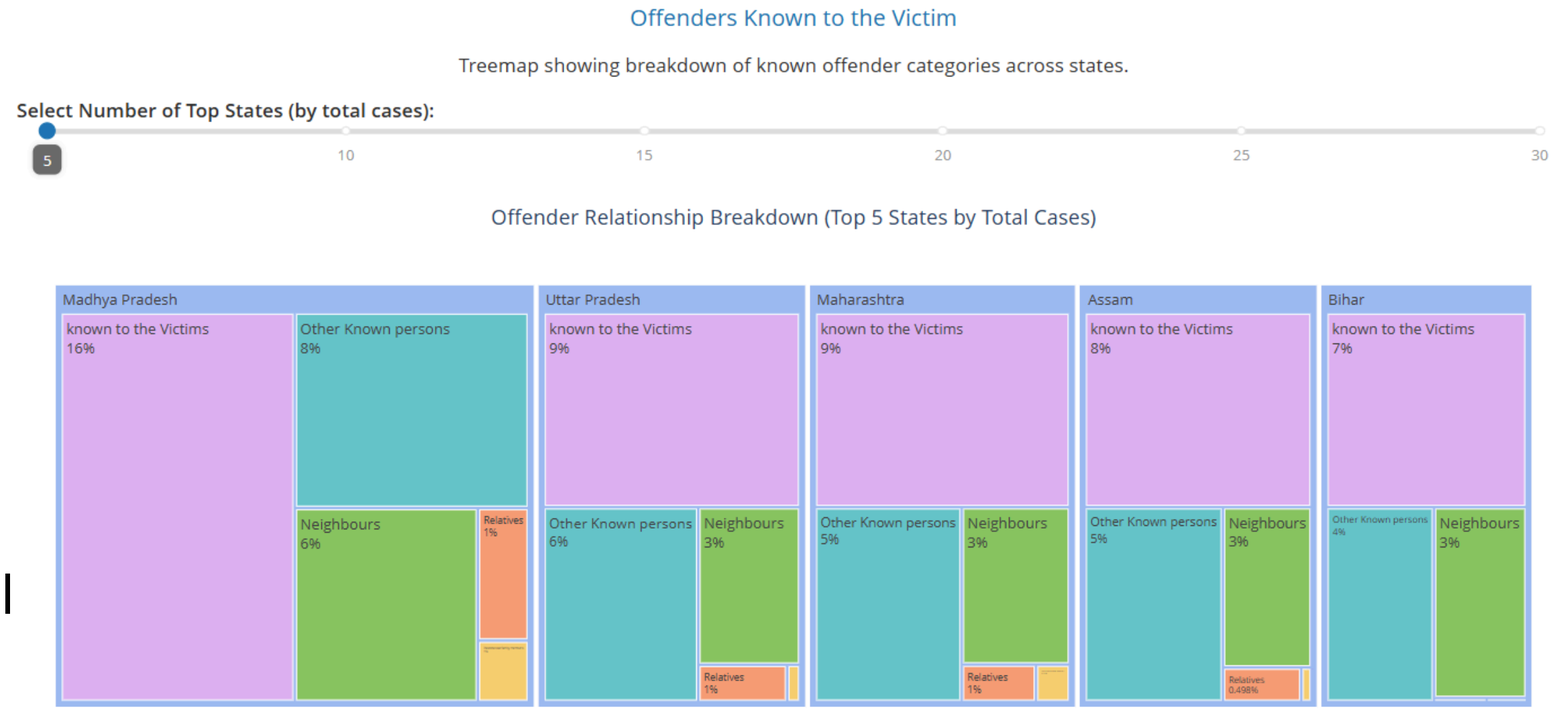
- **Core Feature:** Analyze crime locations (e.g., Residential, Highway).
- **Functionality:**
  - **Sunburst chart:** Hierarchical breakdown (Year -> Place -> Crime).
  - **Small Multiples Bar Chart:** Crime trends over selected years, faceted by place type.





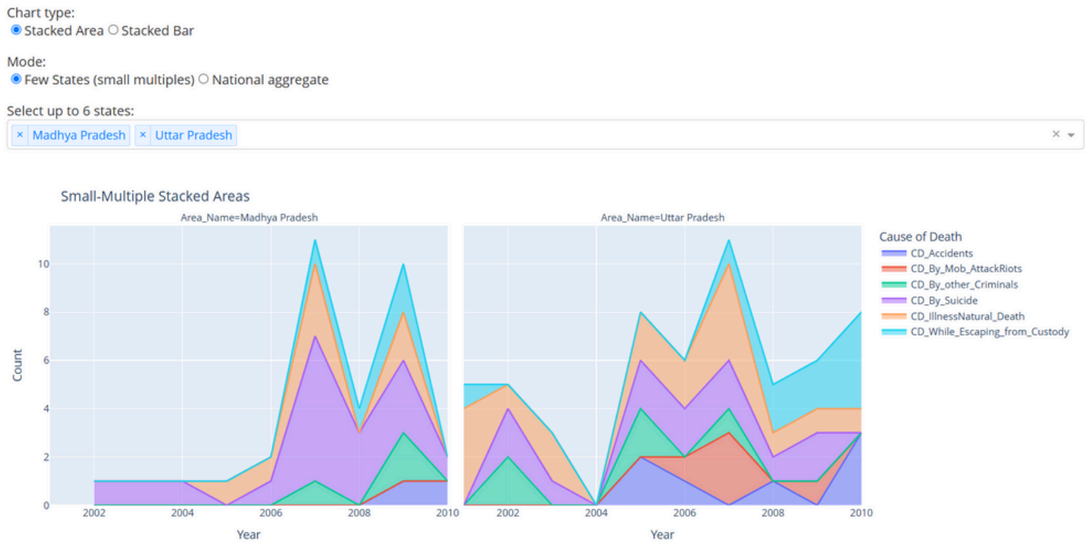
Offender Relationships Tab

- **Core Feature:** Explore relationship between victims and known offenders.
- **Functionality:**
  - Treemap showing breakdown by relationship type (Relatives, Neighbors, etc.) within Top N states.
  - Rectangle size proportional to case count.

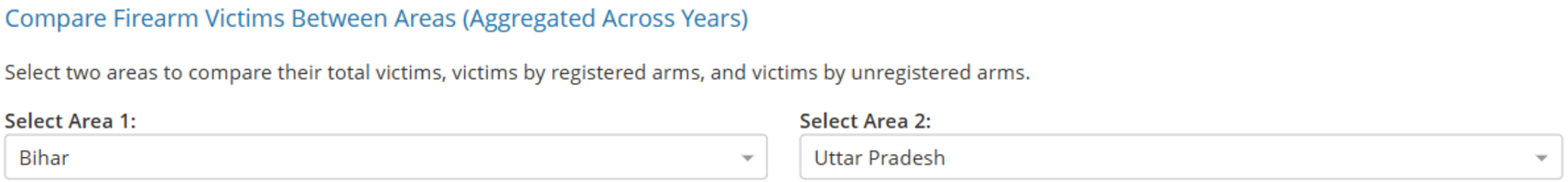


Custodial Deaths Plots Tab

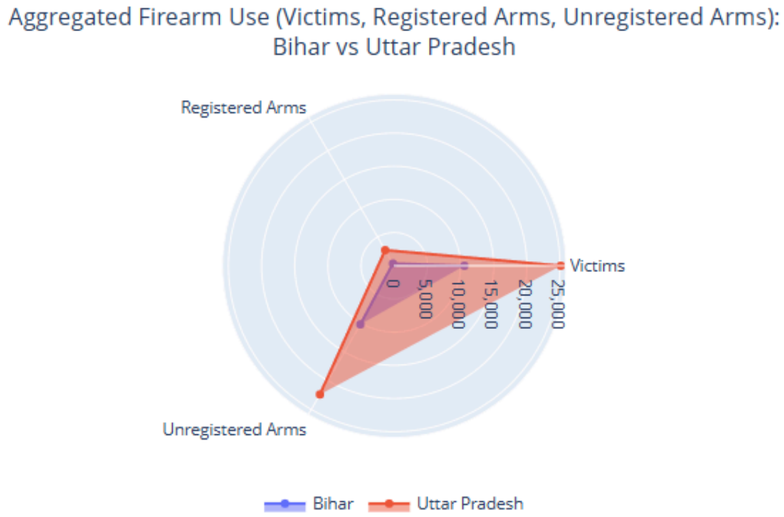
- **Core Feature:** Analyze trends and causes of deaths in custody.
- **Functionality:**
  - Visualize national trend or compare selected states (Small Multiples).
  - Stacked Area/Bar charts showing breakdown by cause (Suicide, Illness, etc.) over time.



Area Comparison Tab

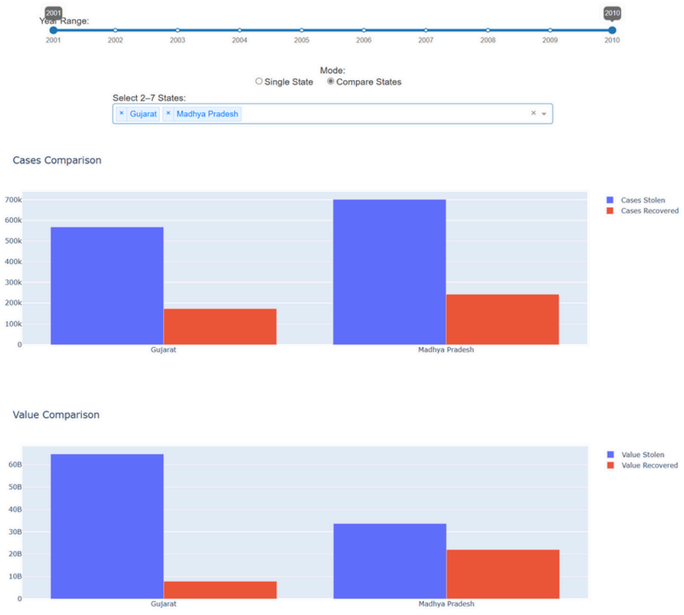


- **Core Feature:** Direct comparison of two States/UTs on firearm use in murder.
- **Functionality:**
  - Select two areas using dropdowns.
  - Radar chart compares Total Victims, Victims by Registered Arms, Victims by Unregistered Arms (aggregated over all years).



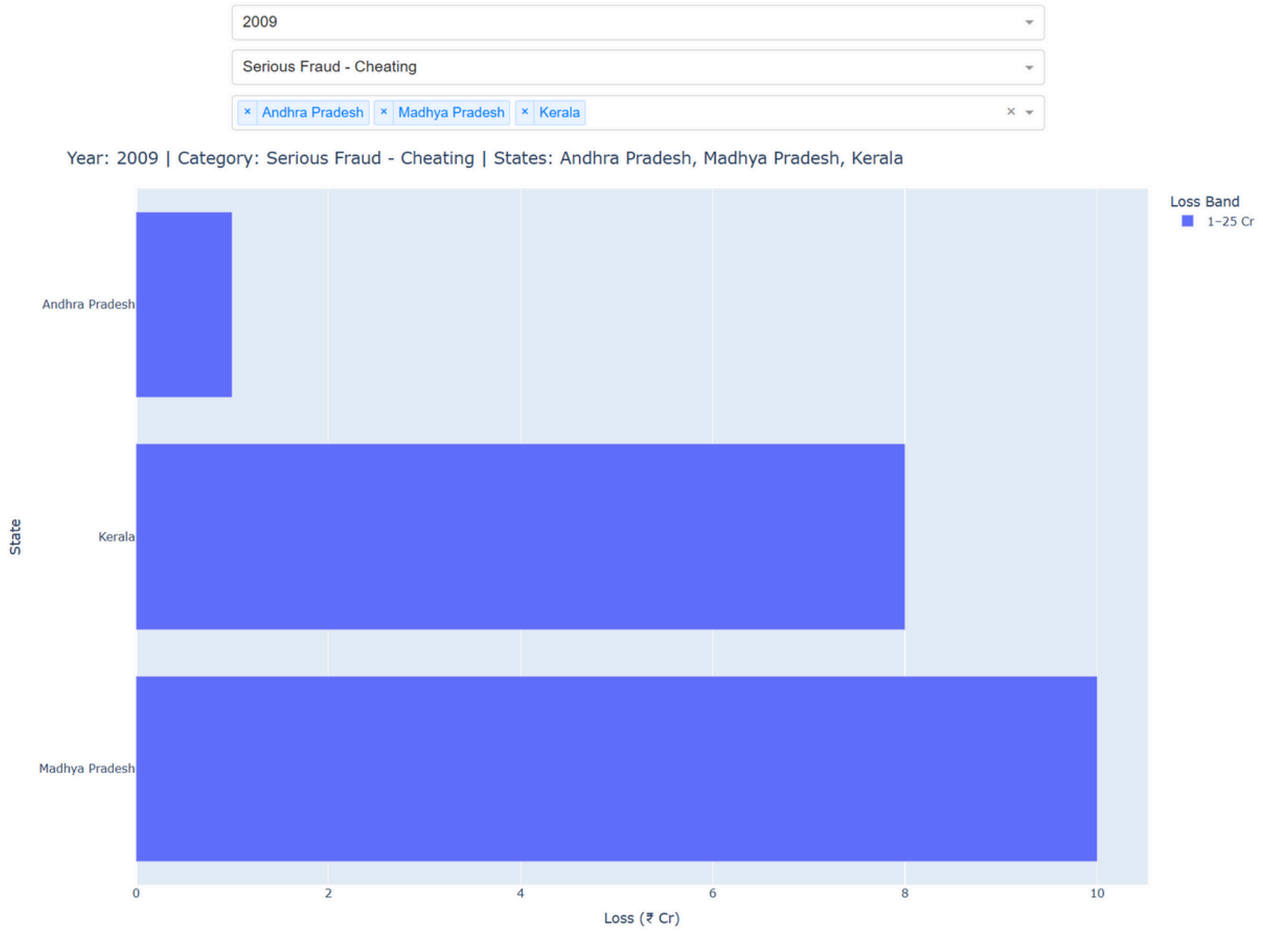
Property Stolen Analysis Tab

- **Core Feature:** Compare property stolen vs. recovered (value & cases).
- **Functionality (Selectable Modes):**
  - Single State: Trend lines for stolen/recovered value and cases over time.
  - Compare States: Grouped bar charts comparing total stolen/recovered value and cases across selected states.



Serious Fraud Losses Tab

- **Core Feature:** Visualize reported serious fraud cases by loss magnitude.
- **Functionality:**
  - Horizontal Stacked Bar Chart: States vs. Case Count.
  - Segments show breakdown by loss band (e.g., 1-25 Cr, >100 Cr).
  - Filter by Year, Fraud Group, State(s).



Limitations

- **Data Timeliness:** Primarily uses data up to 2012/2013; may not reflect current situation.
- **Data Quality:** Potential inconsistencies, under-reporting, missing values in source data.
- **Aggregation Level:** Mainly State/District; masks intra-unit variations.
- **Reliance on Reported Crime:** Actual rates may differ; definitions can change.
- **Fuzzy Name Matching:** Imperfect, potential for mismatches.
    -

Future Work

- **Update Data:** Incorporate more recent crime statistics.
- **Expand Datasets:** Add socio-economic indicators, conviction data.
- **Advanced ML:** More sophisticated forecasting, clustering, anomaly detection.
- **Granular Geography:** Sub-district analysis (if data permits).
- **User Uploads:** Allow users to integrate their own relevant data.

Conclusion

- Successfully integrated diverse Indian crime dataset (2001-2013) into an interactive Dash/Plotly portal.
- Empowers exploration across geographic, temporal, and thematic dimensions.
- Utilizes various visualizations (maps, charts, Sankey, Treemap, etc.) to make complex data accessible.
- Integrates ML (Clustering, Forecasting) for deeper pattern discovery.
- Fosters data-informed understanding of crime in India.

Thank You