CRIME PATTERN ANALYSIS AND VISUALIZATION USING OPEN DATA

NAME: S. Priyadharshini

DEPARTMENT: B.Tech Information Technology

COLLEGE: Prathyusha Engineering College

CERTIFICATE

This is to certify that Ms. S. Priyadharshini has successfully completed the project titled 'Crime Pattern Prediction in Predictive Model' submitted to Prathyusha Engineering College.

Supervisor: _____ Date: _____

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my project guide and faculty members for their constant support, valuable suggestions, and encouragement throughout the course of this project. I also extend my thanks to my family, friends, and the academic community for their continuous motivation and guidance.

ABSTRACT

This project focuses on the prediction of crime patterns using predictive modeling techniques. By leveraging machine learning algorithms, the project aims to identify crime trends, forecast potential hotspots, and provide insights for effective law enforcement strategies. The study combines open crime datasets with advanced visualization and clustering methods to highlight crime frequency, timing, and distribution. The results demonstrate how predictive models can aid in crime prevention, enhance public safety, and support policy-making.

INTRODUCTION

Crime is a major concern for every society, and predicting crime patterns has become an essential part of maintaining law and order. Predictive policing leverages data analytics, machine learning, and statistical methods to anticipate criminal activity. This project integrates open data with predictive models to develop an analytical dashboard for crime pattern prediction.

LITERATURE REVIEW

Existing studies on crime prediction highlight the importance of statistical models such as ARIMA, decision trees, and clustering methods. Recent advancements in machine learning, including Random Forests, LSTMs, and neural networks, have improved predictive accuracy. Research also emphasizes ethical considerations such as data privacy and bias reduction in predictive policing systems. Studies reviewed include applications in urban crime forecasting, hotspot detection, and real-time crime analysis.

OBJECTIVES

- To analyze open crime datasets and identify patterns.
- To apply machine learning techniques such as clustering, classification, and time-series forecasting.
- To build an interactive dashboard for visualization and prediction.
- To support law enforcement agencies with actionable insights for crime prevention.

METHODOLOGY

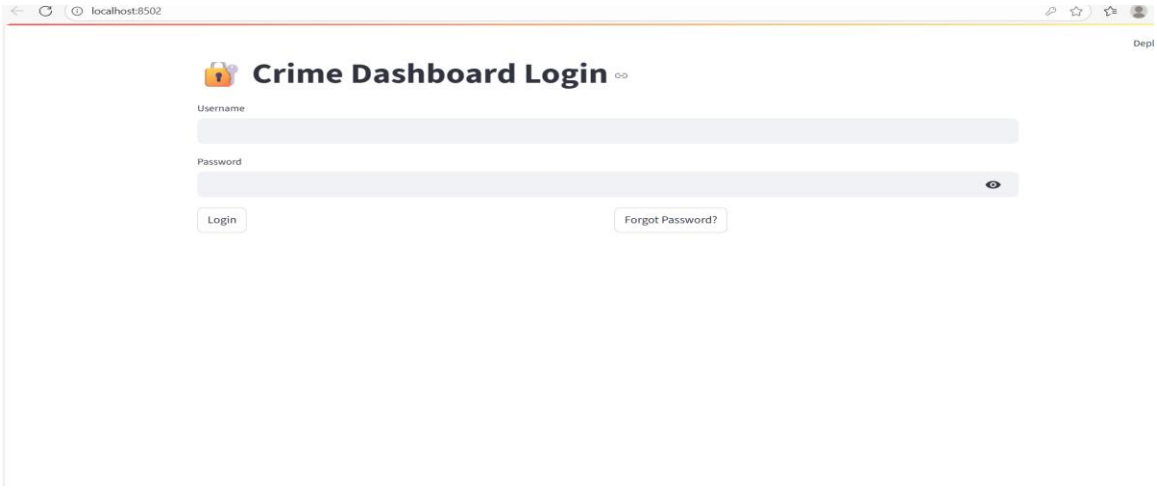
The methodology involves:

1. Data Collection: Publicly available crime datasets with attributes such as date, time, location, and category.
2. Data Preprocessing: Cleaning, handling missing values, and standardizing date-time formats.
3. Feature Engineering: Extracting features like day of week, month, and region.
4. Modeling: Applying Random Forest, LSTM, and KMeans clustering for prediction and analysis.
5. Visualization: Building dashboards using Streamlit, Plotly, and Folium for interactive exploration.
6. Evaluation: Measuring accuracy, precision, recall, and cluster quality for performance validation.

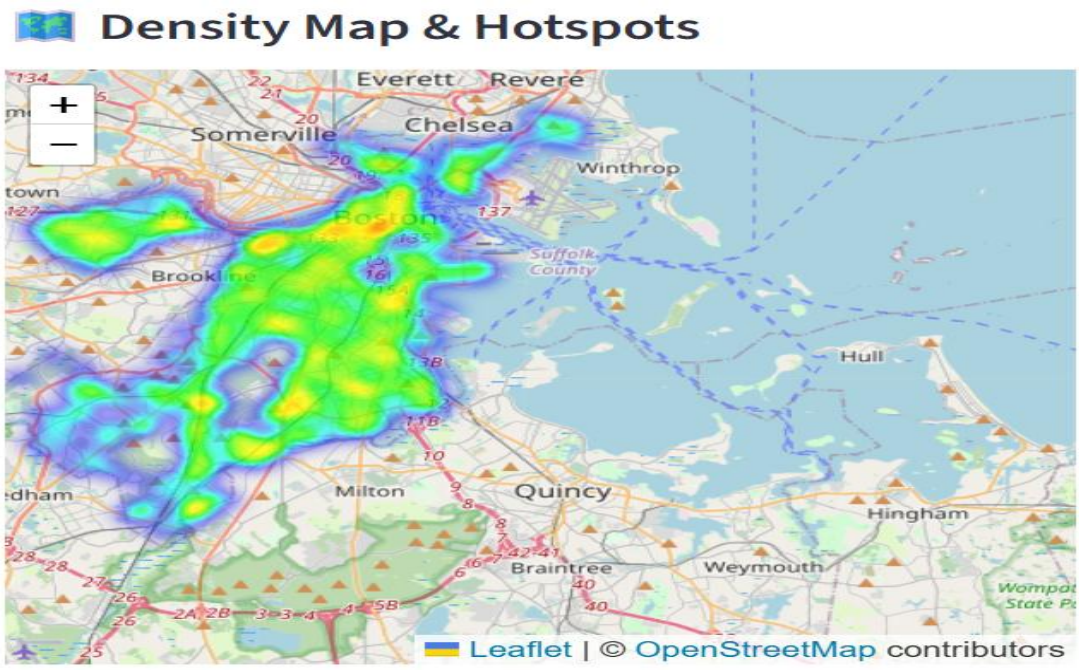
RESULTS

The results of the project include several visualizations and predictive outputs:
- Heatmaps showing crime density across regions.
- Daily and monthly crime trends visualized with line and area charts.
- Bar and pie charts showing crime distribution by category.
- Clustering analysis identifying regions with similar crime volumes.
- Predictive models providing classification of crime categories and time-series forecasting.

Sample visualization results and charts were generated using the uploaded dataset.




localhost:8502
Deploy
Data
Upload CSV
Drag and drop file here
Limit 200MB per file • CSV
Browse files
crime...
50.9MB
Privacy
Anonymize (drop PII, round coords)
Map Options
Max points (sampling)
2500
500
10000
Show markers
Advanced Crime Analytics & Governance Suite
Filters
Date Range
2015/01/07 – 2018/12/08
Crime Category
All
Region
All
Records
259469
Categories
66
Regions
13
Peak Day
Friday
Tip: Use fewer markers and the heatmap for faster rendering on large files.
Density Map & Hotspots
Everett
Revere
Somerville
Chelsea
Winthrop


Density Map & Hotspots
Everett
Revere
Somerville
Chelsea
Winthrop
Boston
Suffolk County
Brookline
Hull
Milton
Quincy
Hingham
Braintree
Weymouth
Leaflet | © OpenStreetMap contributors

Crime Analytics Report

Total Records: 259469

Unique Categories: 66

Regions: A1, A15, A7, B2, B3, C11, C6, D14, D4, E13 ...

Date Range: 2015-01-07 to 2018-12-08

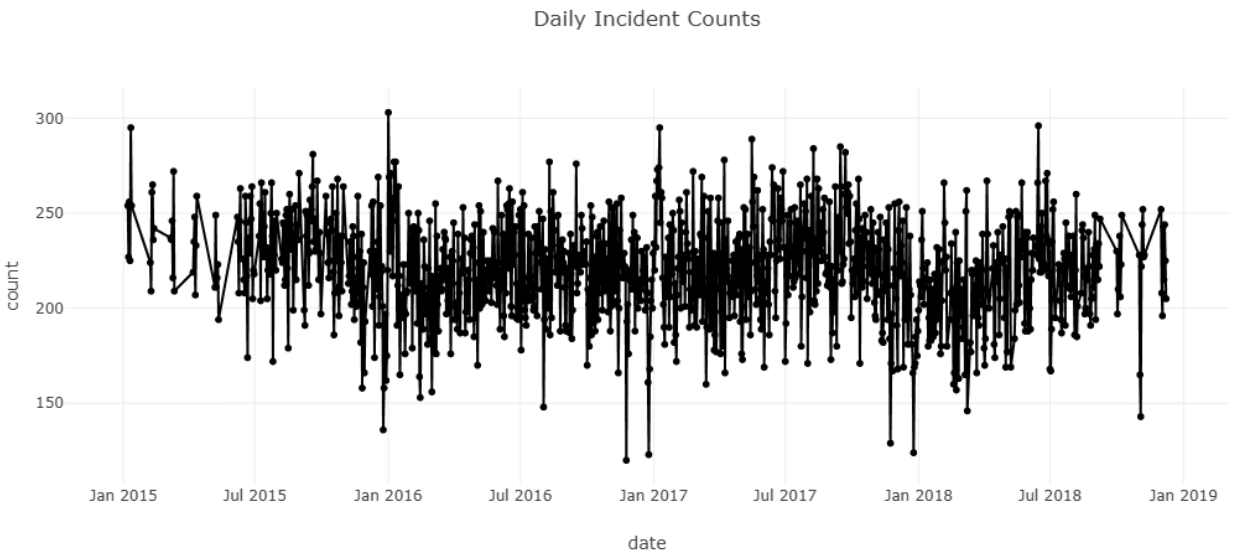
Key Insights

- Most common category: Motor Vehicle Accident Response

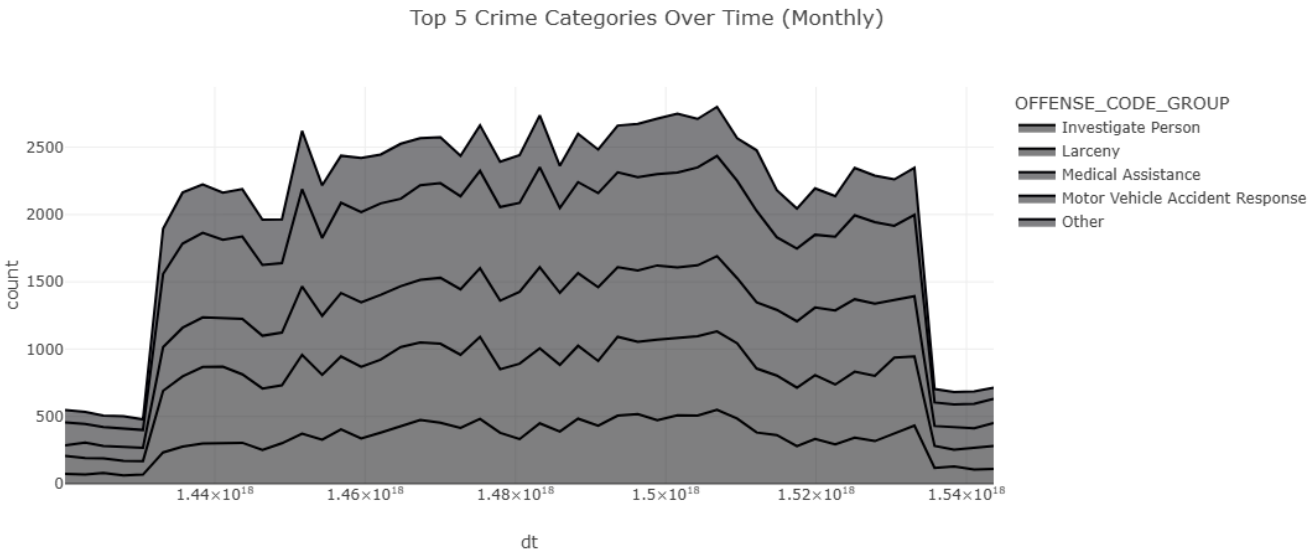
- Region with most incidents: B2

- Peak day of week: Friday

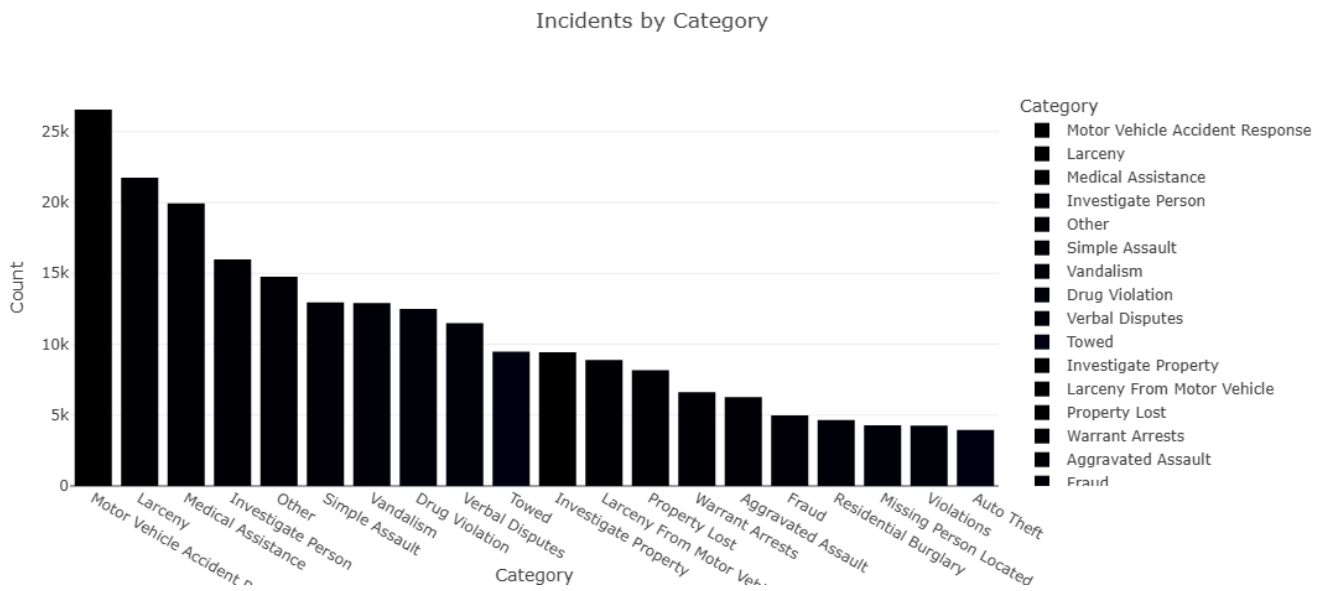
Daily Trend



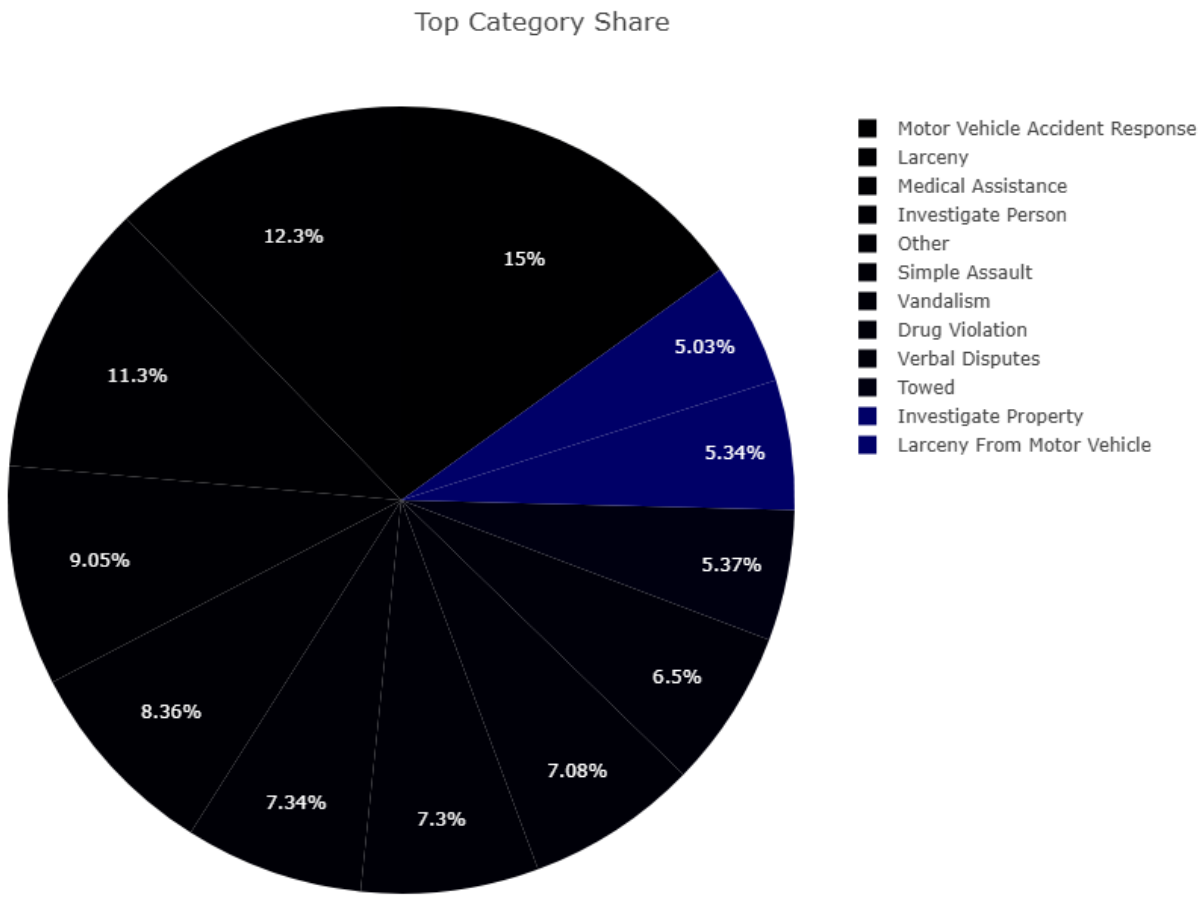
Monthly Series (Top 5)



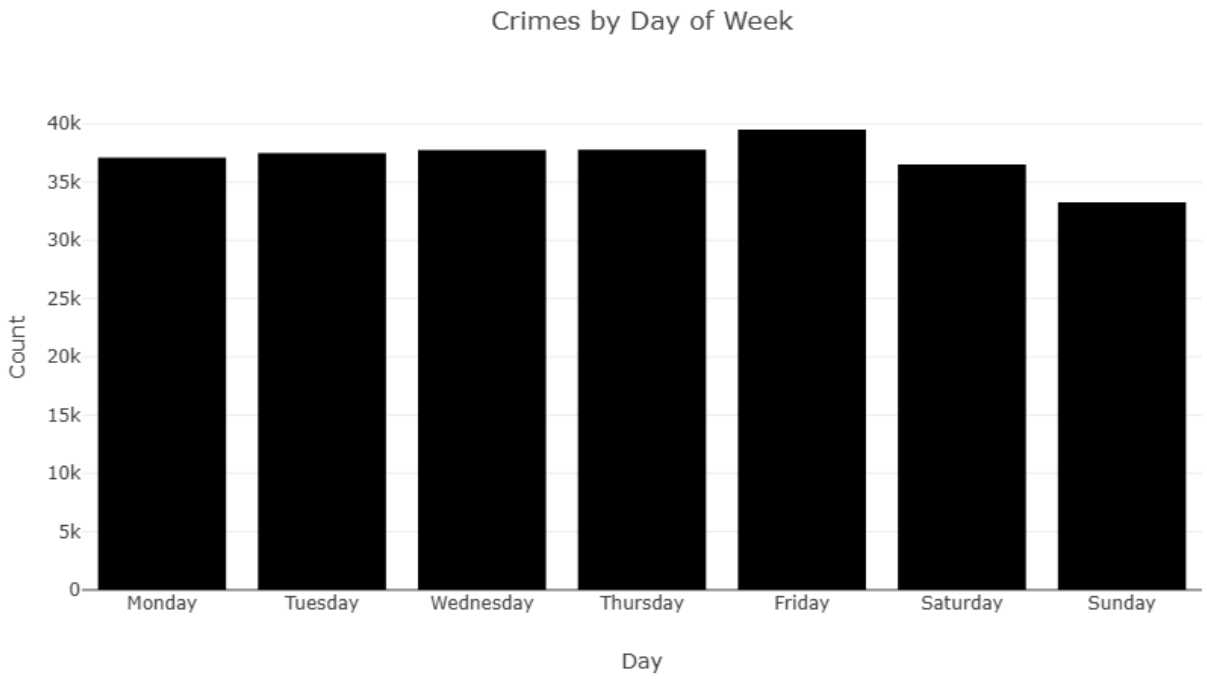
By Category



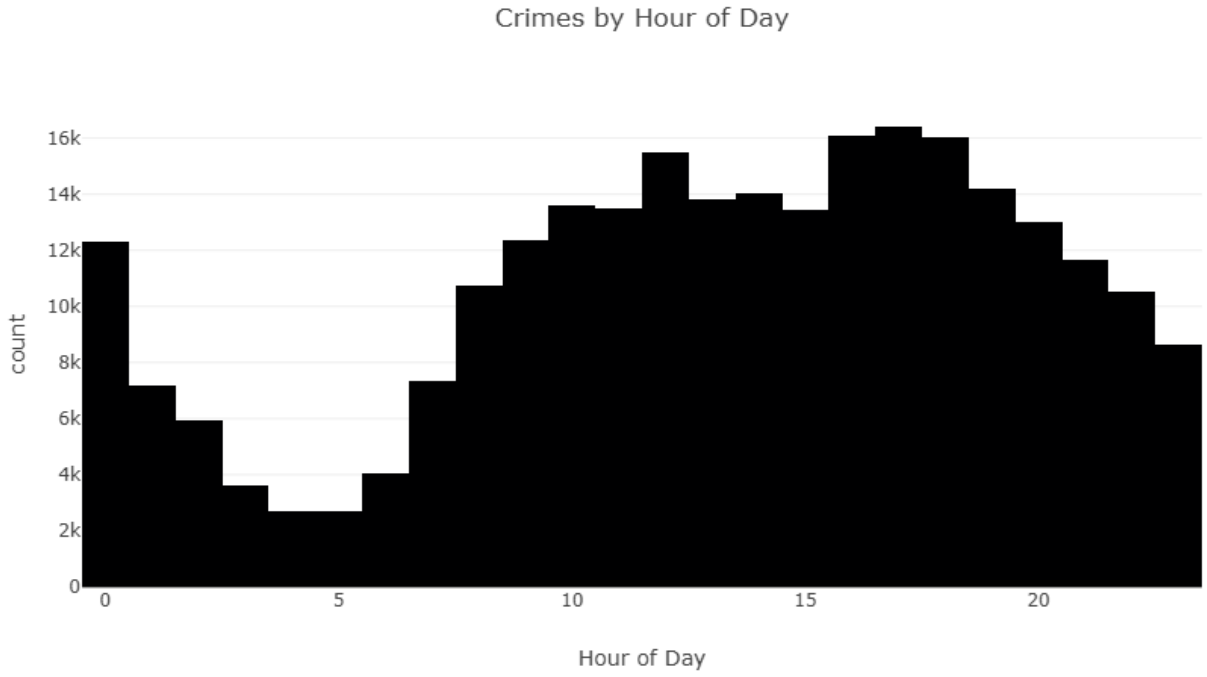
Category Share



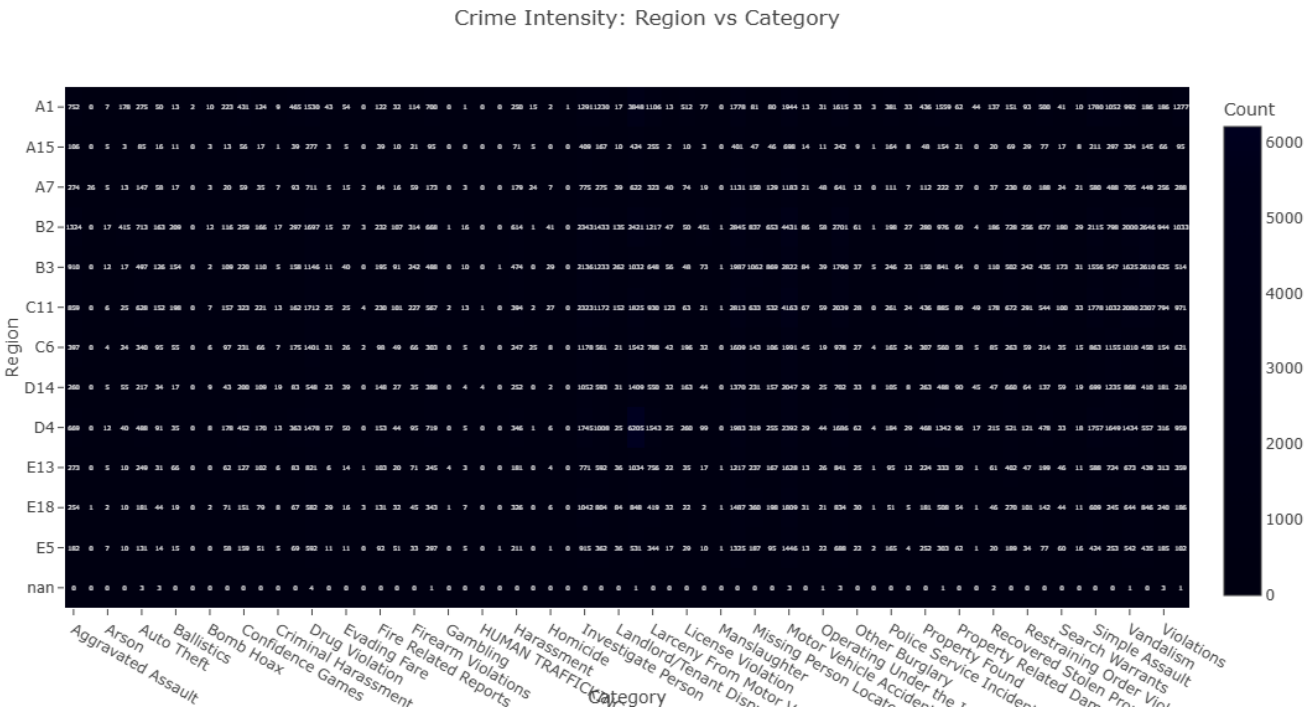
By Day of Week



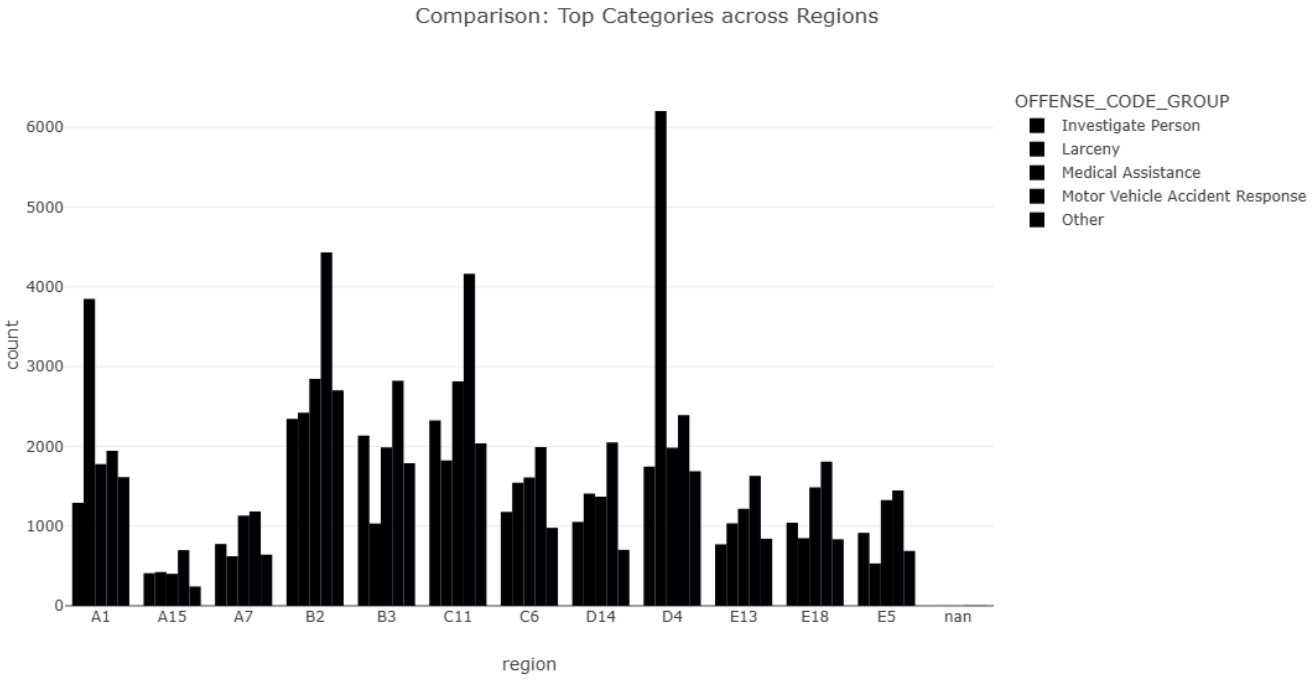
By Hour



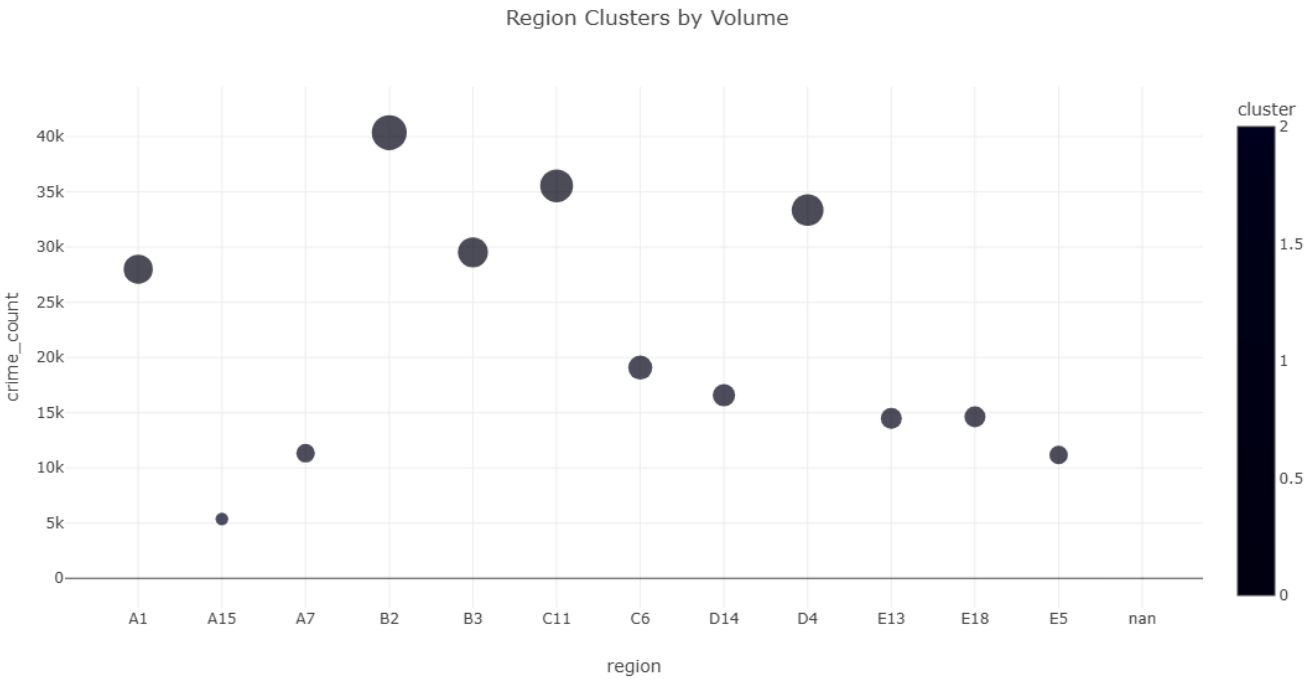
Region vs Category Heatmap

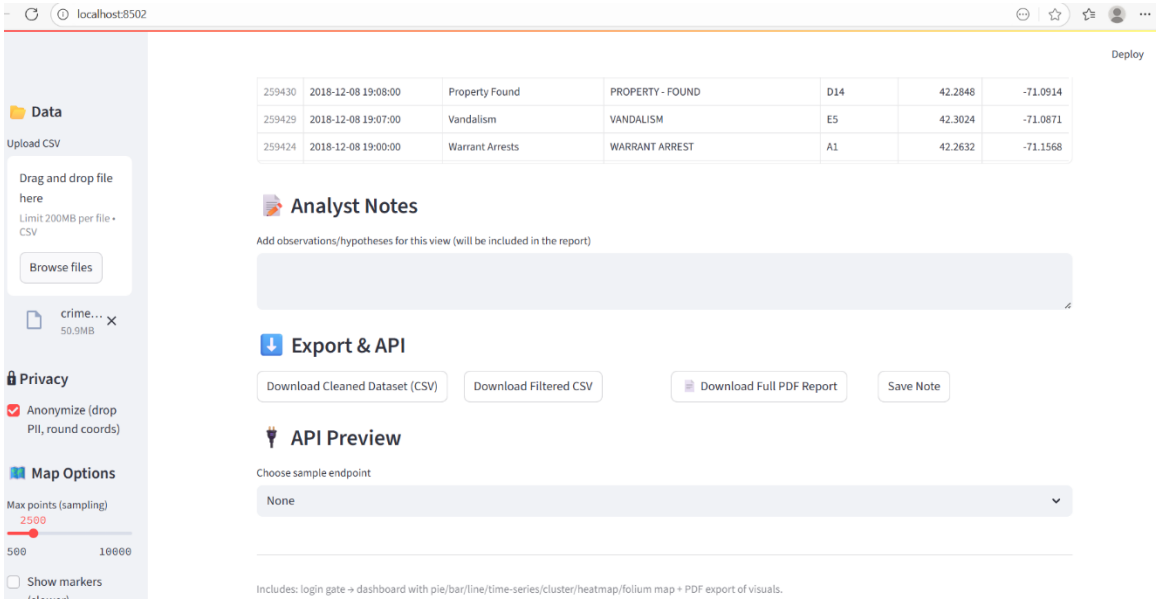


Regions Comparison (Top Cat)



Region Clusters


Region Clusters by Volume
crime_count
40k
35k
30k
25k
20k
15k
10k
5k
0
A1
A15
A7
B2
B3
C11
C6
D14
D4
E13
E18
E5
nan
region
cluster
2
1.5
1
0.5
0



ANALYSIS

The analysis reveals that crime patterns follow distinct temporal and spatial distributions. Certain regions exhibit consistently high crime rates, while others remain relatively safe. Time-series models captured seasonal variations, with crime peaks during weekends and late evenings. Clustering highlighted groups of regions with similar crime frequencies, aiding targeted interventions. Machine learning models achieved significant accuracy in classification tasks, validating the predictive capability of the approach.

CONCLUSION

This project demonstrates that predictive modeling can significantly enhance crime pattern analysis. By integrating statistical and machine learning models, the system provides valuable insights into crime trends. The dashboard offers a practical tool for law enforcement to visualize and forecast crime, enabling proactive decision-making. Future work may involve incorporating real-time data streams, advanced deep learning models, and fairness-aware algorithms to ensure ethical and unbiased predictions.

REFERENCES

1. Machine Learning Approaches for Crime Prediction (ResearchGate, 2023)

2. Predictive Policing and Big Data Analytics (ScienceDirect, 2022)

3. Crime Forecasting Using Time Series Models (IEEE, 2021)

4. LSTM Networks for Crime Trend Prediction (Springer, 2024)

5. Data Mining in Criminal Investigations (Elsevier, 2020)

6. Spatial Analysis of Crime Hotspots (Taylor & Francis, 2023)

7. Deep Learning Applications in Crime Prediction (arXiv, 2024)

8. Ethical Concerns in Predictive Policing (Science, 2022)

9. Random Forest Models for Crime Classification (MDPI, 2021)

10. AI in Public Safety and Governance (Springer, 2023)

11. Crime Data Visualization and Dashboards (JIER, 2022)

12. Clustering Techniques for Crime Pattern Detection (INFORMS, 2023)

13. Predictive Analytics in Law Enforcement (ResearchGate, 2024)

14. Urban Crime Forecasting with Neural Networks (Elsevier, 2025)

15. Real-Time Crime Prediction Systems (IEEE Access, 2024)