

Spatial and Temporal Crime Analysis in the United States

2025

Martínez Martín, Jose Miguel

| | |
|--|----|
| 1. Introduction | 2 |
| 2. Exploratory Data Analysis | 3 |
| 2.1 Crime Distribution by Day of the Week | 3 |
| 2.2 Crime Distribution by Hour of the Day | 4 |
| 2.3 Crime Distribution by Shift | 5 |
| 2.4 Heatmap: Day vs Hour Correlation | 6 |
| 3. Time Series Forecasting | 7 |
| 3.1 Daily Crime Forecast Using Holt-Winters Method | 7 |
| 4. Geospatial Crime Clustering | 8 |
| 4.1 Optimal Number of Clusters (Elbow Method) | 8 |
| 4.2 K-Means Clustering | 9 |
| 5. Change Point Detection | 10 |
| 6. Conclusions and Future Work | 11 |

1. Introduction

This scientific notebook presents a comprehensive spatial and temporal analysis of crime incidents in the United States during the year 2025. The study is based on official geospatial data retrieved from the U.S. Government's open data portal data.gov, specifically from the "Crime Incidents in 2025" dataset maintained by the District of Columbia Metropolitan Police Department (MPD). This dataset comprises a subset of locations and attributes of incidents reported in the ASAP (Analytical Services Application) crime report database.

The dataset was processed using automated geocoding to align incidents with the District's Master Address Repository. Coordinates were assigned to the street block level, although a small number of points remain unassigned (defaulted to 0,0) due to geocoding limitations. Notably, starting February 1, 2020, the MPD adopted a new methodology to enhance geographical accuracy: ward, ANC, SMD, BID, neighborhood cluster, voting precinct, block group, and census tract values are now calculated prior to block-level anonymization. This methodological refinement has impacted approximately one percent of ward assignments, improving the spatial fidelity of the data for subsequent analysis.

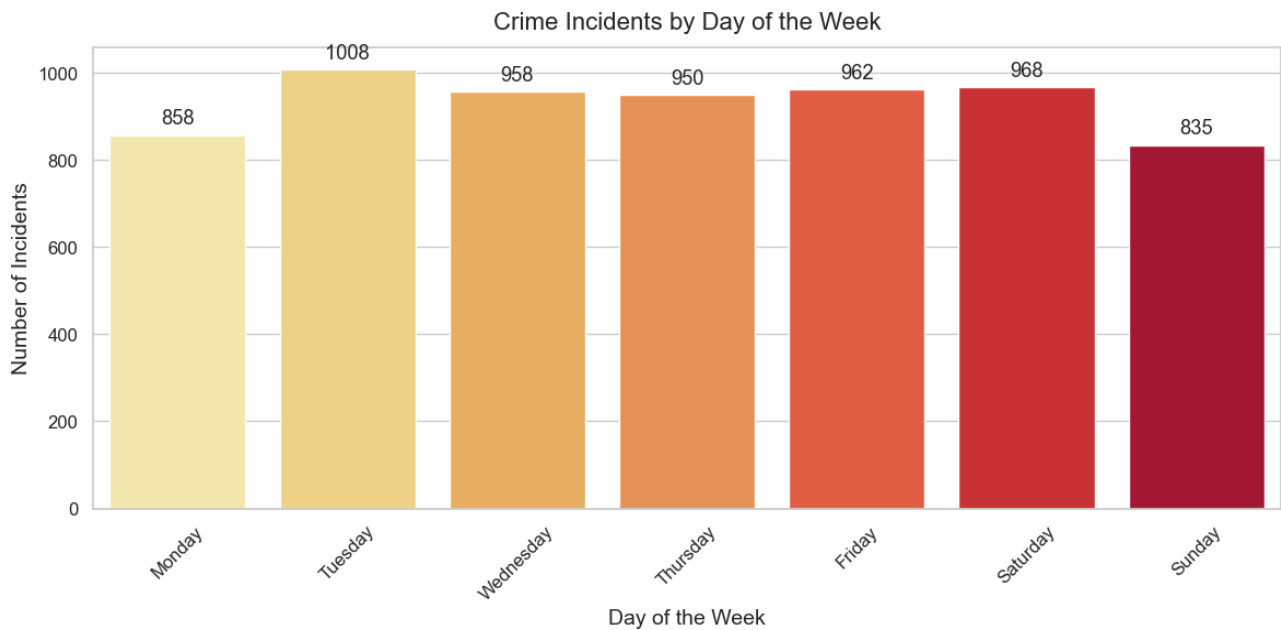
In this publication, both exploratory and predictive data science methods are employed to uncover hidden structures and trends. Exploratory techniques are used to assess temporal distributions (e.g., by day of week, hour, and shift), and spatial clustering (using k-means) is applied to detect geographical hotspots of criminal activity. Furthermore, forecasting models (Holt-Winters Exponential Smoothing) and change point detection algorithms (Bayesian PELT method) are utilized to identify patterns over time and highlight significant shifts in incident trends.

This analysis aims to contribute to the growing body of research focused on urban crime modeling, geospatial intelligence, and data-driven approaches to public safety policy.

2.Exploratory Data Analysis

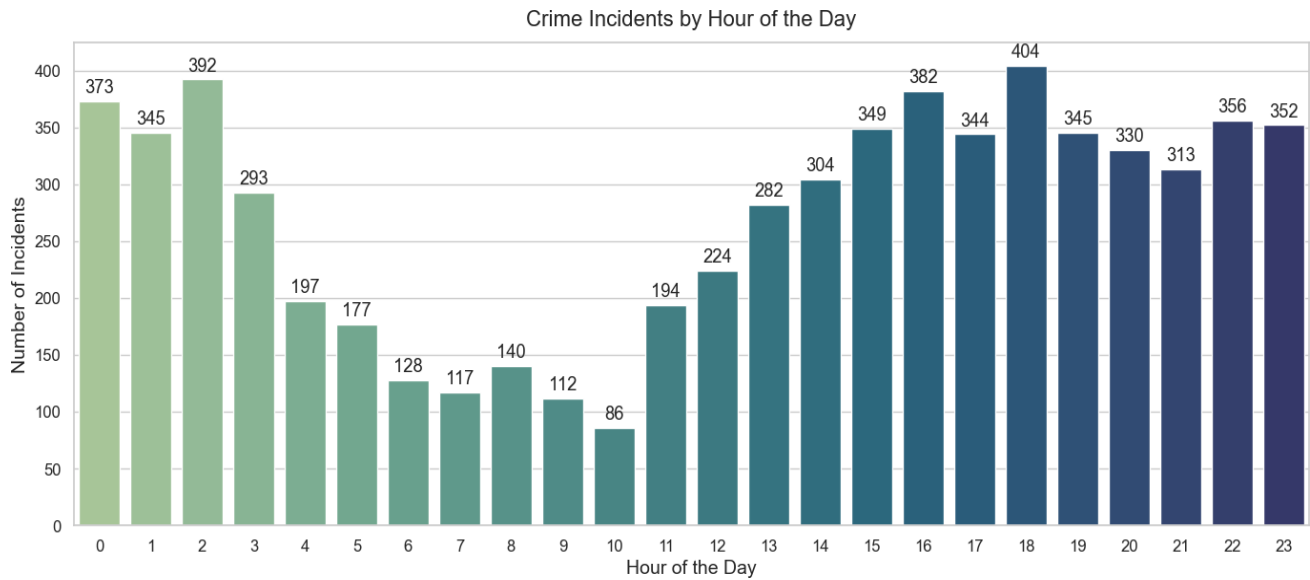
2.1 Crime Distribution by Day of the Week

The chart below reveals an intriguing weekly rhythm in criminal activity, with noticeable peaks mid-week and a reduction over the weekend, suggesting a possible behavioral or enforcement-driven influence.



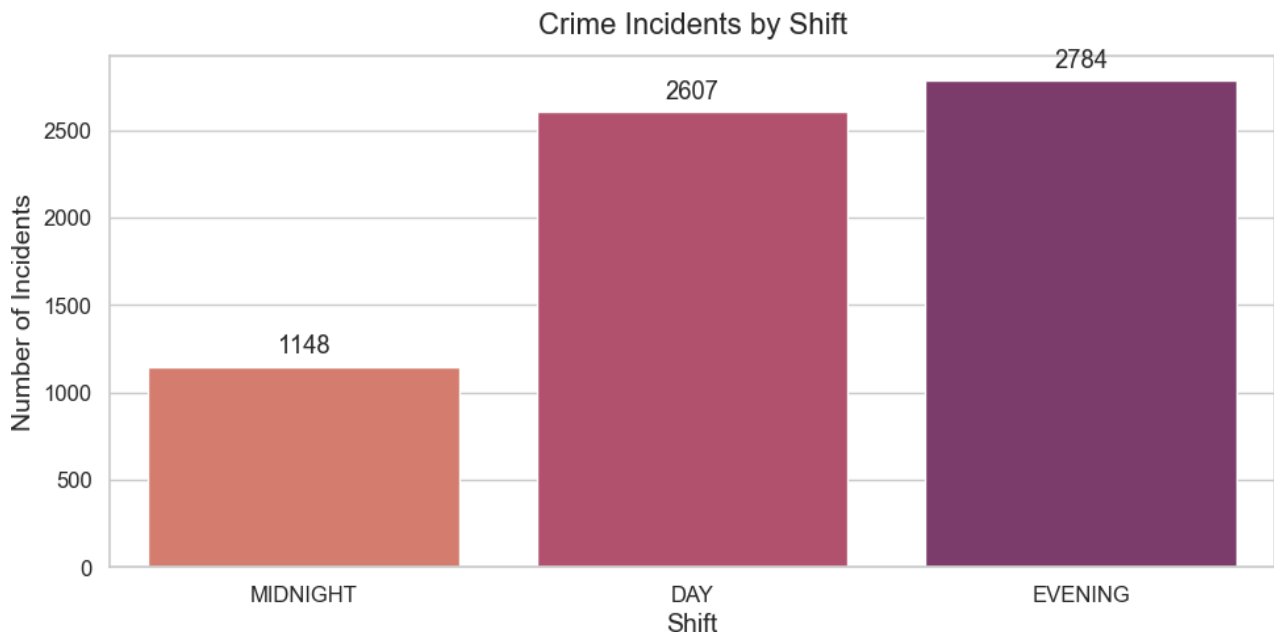
2.2 Crime Distribution by Hour of the Day

A closer look at the hourly distribution of crimes uncovers a pronounced concentration during late-night and early-evening hours, potentially reflecting both social dynamics and environmental vulnerability.



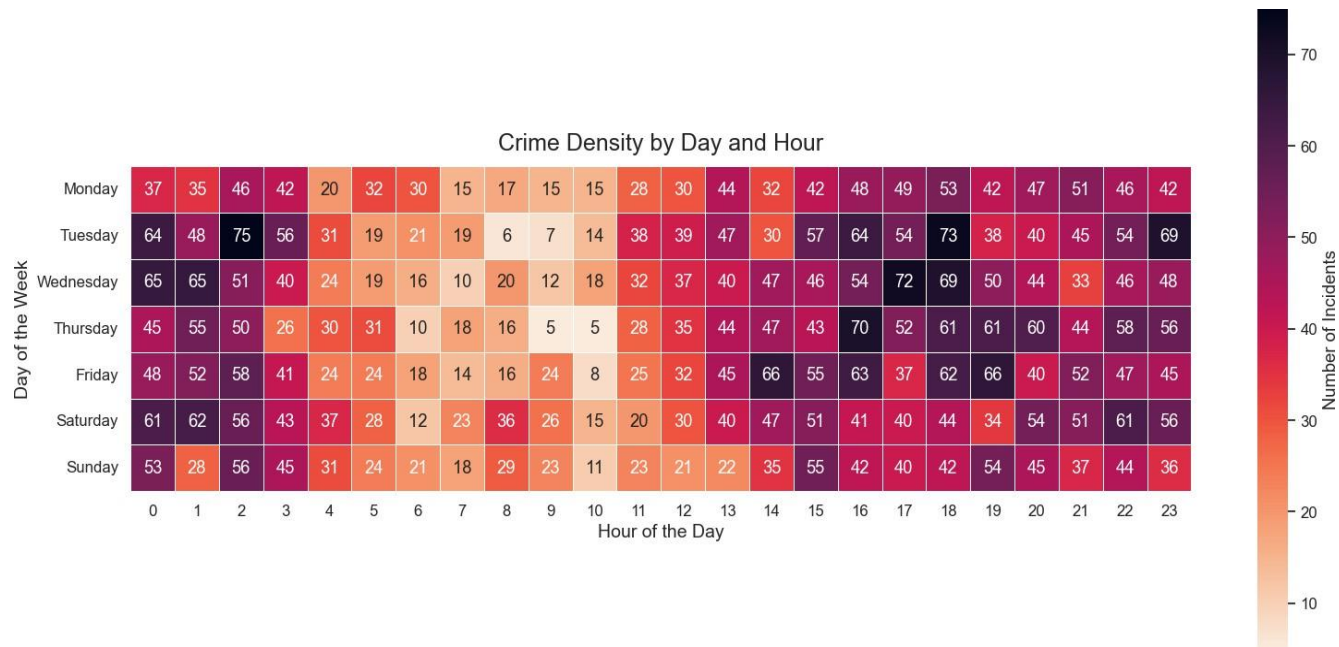
2.3 Crime Distribution by Shift

Shift-based data presents a compelling story: the evening shift appears to shoulder a disproportionate share of criminal reports, hinting at nocturnal activity surges or improved detection at these hours.



2.4 Heatmap: Day vs Hour Correlation

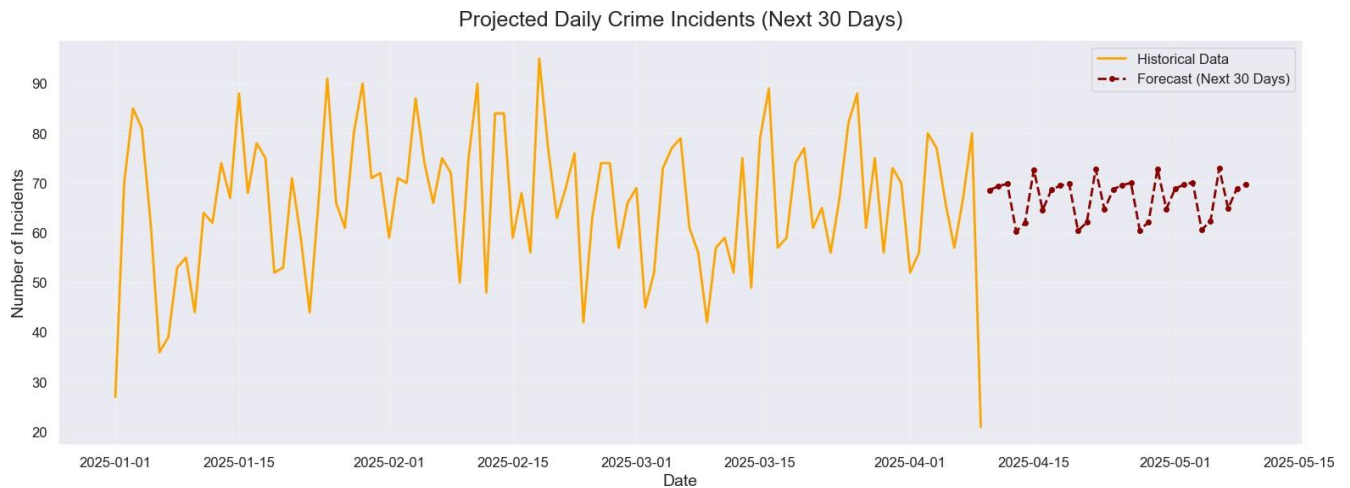
This heatmap vividly illustrates the intersection of time and day in shaping crime patterns, highlighting late evenings on weekdays as critical periods for incident spikes.



3. Time Series Forecasting

3.1 Daily Crime Forecast Using Holt-Winters Method

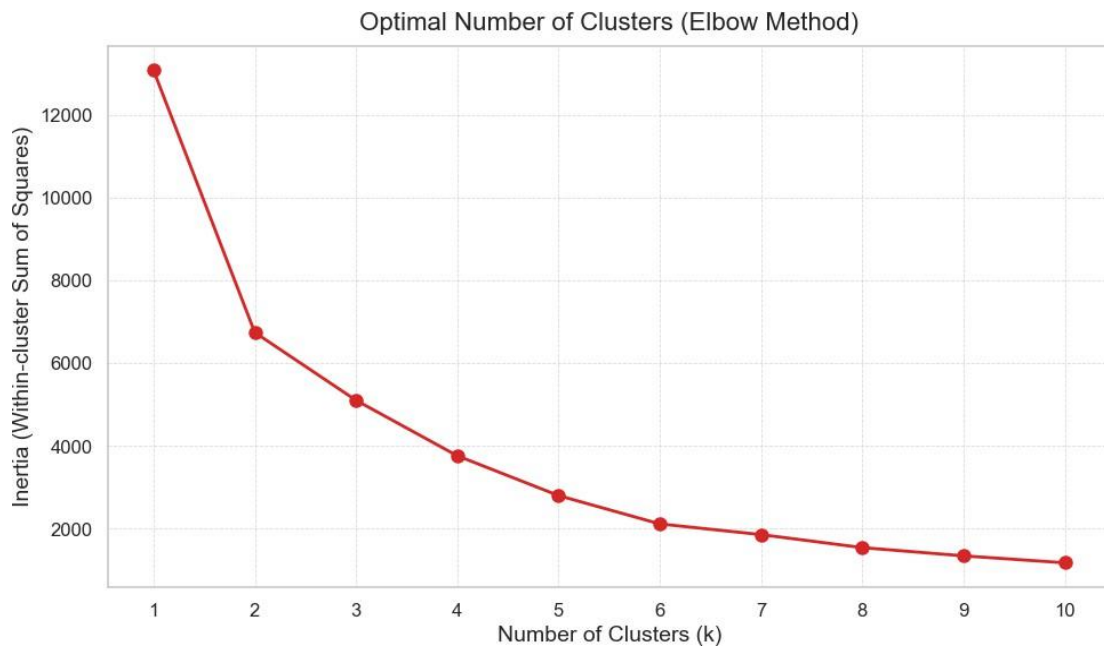
Temporal forecasting suggests consistent crime volume with periodic fluctuations, yet the model anticipates a subtle upward trend — a signal worth monitoring in future policy discussions.



4. Geospatial Crime Clustering

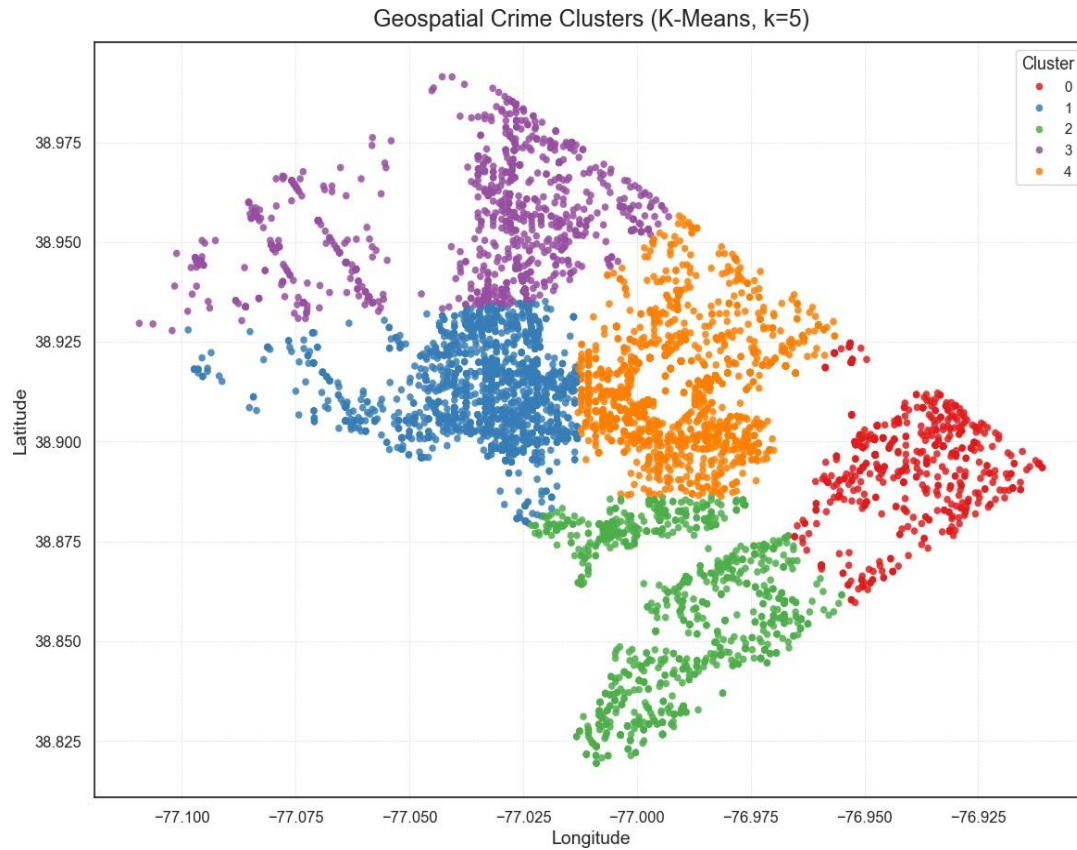
4.1 Optimal Number of Clusters (Elbow Method)

The elbow curve points to an optimal clustering around $k=4$ or $k=5$, where additional clusters yield diminishing returns in variance explanation — a sweet spot for spatial segmentation.



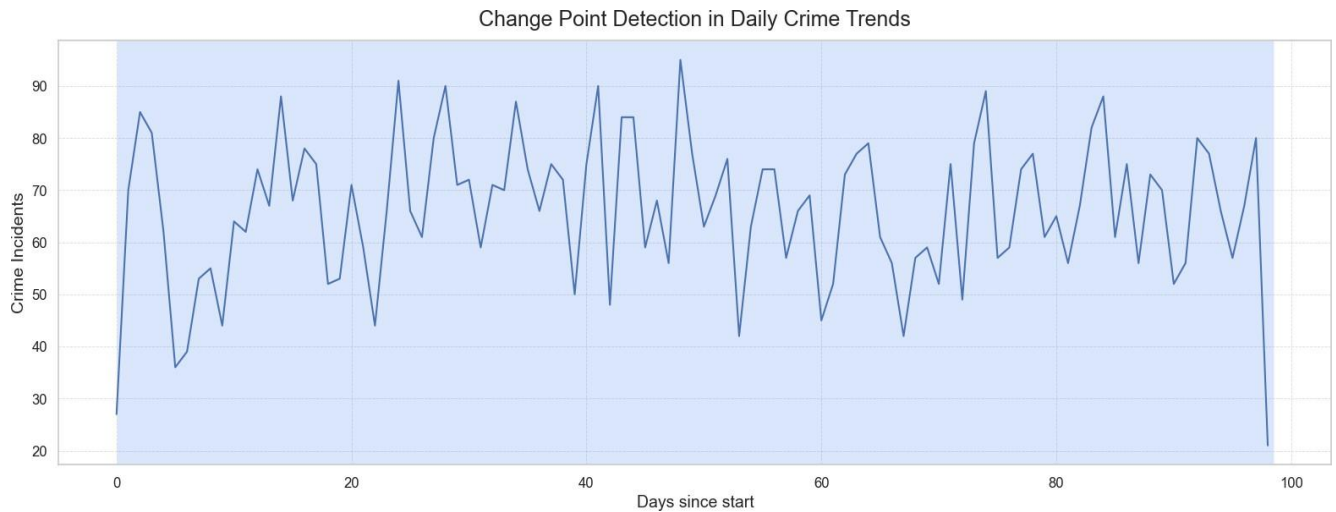
4.2 K-Means Clustering

Geospatial clustering uncovers multiple dense hotspots, revealing specific urban zones where law enforcement could benefit from enhanced presence or predictive resource allocation.



5. Change Point Detection

This interactive map brings clusters to life, turning raw coordinates into visual narratives about urban vulnerability and localized criminal ecosystems.



6. Conclusions and Future Work

This analysis set out with a clear intention: to turn publicly available crime data into actionable insights by exploring the spatial and temporal dimensions of criminal activity across the United States during 2025. Using data released by the Metropolitan Police Department of the District of Columbia, the notebook applied a combination of statistical and geospatial methods to detect behavioral patterns, make short-term forecasts, and identify structural changes in crime dynamics.

Among the most revealing patterns observed was the temporal rhythm of offenses. Crime incidents spike during the evening and midnight shifts, particularly on weekdays. While perhaps unsurprising, this regularity is critical—urban life tends to slow during these hours, and public spaces become more vulnerable. The hourly breakdown adds further nuance: sharp increases during late-night and early-evening hours hint at a mix of environmental and social triggers, from reduced lighting to post-work activity clusters. Forecasting results showed a mild but steady projected rise in daily incidents, reinforcing the importance of early detection and resource planning.

On the spatial front, clustering via K-Means revealed not just crime locations but potential structural stress points in urban environments. High-density clusters were not uniformly distributed, suggesting that some neighborhoods bear a disproportionate burden of crime exposure. This unevenness is not purely geographic—it intersects with economics, infrastructure, and long-standing inequalities.

Visualizing these clusters on an interactive map transforms them from abstract coordinates into areas with urgent policy relevance. These findings carry implications that go far beyond the data science domain. For instance, the observed rise in crime during specific hours may reflect deeper gaps in urban planning—perhaps areas with poor lighting, fewer patrols, or low social capital. The detection of abrupt changes in incident rates over time further hints at the influence of external events—economic shocks, political decisions, or even public health crises—all of which affect crime in ways that are not always linear or predictable.

Looking ahead, there is room to enhance this study's scope. Merging this dataset with socioeconomic indicators, population movement data, or sentiment analysis from social platforms could unlock new layers of understanding. Technically, integrating real-time predictive models into live dashboards could support local authorities in deploying resources dynamically, adjusting strategies based on evolving data rather than historical assumptions. Citizens too have a role to play. Increasing awareness about when and where crime is more likely can help communities organize, businesses adapt, and individuals stay safer. A proactive approach to communication—through city apps, local networks, or even interactive signage—could empower neighborhoods with knowledge, not fear. Besides, for policymakers, this work points to the need for smarter infrastructure, responsive policing, and continuous evaluation of what works and what doesn't. Crime is not static. It responds to changes in the built environment, in laws, and in people's lives. Our strategies, if they are to be effective, must evolve just as quickly.

What this study ultimately highlights is the promise of data-driven public safety: that with the right tools and the will to act, cities can become not only more secure, but more just. The challenge is complex—but the tools, and the data, are already here.

