Crime Analyzer

User Guide for Police Intelligence Dashboard

Ferdinando Muraca, Carlo Vincenzo Stanzione

September 11, 2025

Contents

1 Executive Summary

1.1 System Overview

The Police Intelligence Dashboard is an operations-facing interface designed to quickly surface crime hotspots, priority patterns, and clear recommendations. It integrates a map-based hotspot view with key performance indicators (KPIs) and pattern insights produced by clustering and association-rule analysis.

1.2 Key Capabilities

- Hotspot visualization of spatio-temporal clusters.

- Executive KPIs summarizing clustering outputs.

- Priority patterns with plain-language summaries.

- Actionable recommendations for deployment and prevention.

- Pattern Analysis modal with global, borough, and time-bucket insights.

1.3 Target Users

- **Police operations and analysts:** For situational awareness and resource allocation.

- **Command staff:** For strategy and deployment decisions based on patterns and trends.

2 Quick Start Guide

2.1 Accessing the Dashboard

Open the Police Intelligence Dashboard by launching `Application/Police/index.html`. Ensure the images under `Documents/figures/` and `JupyterOutputs/` are present so all panels render correctly.

2.2 Daily Workflow in 3 Steps

1. **Scan hotspots:** Start from the map to see where activity concentrates and when.

2. **Read KPIs and patterns:** Use the cards and priority patterns to gauge intensity and focus areas.

3. **Zoom with Pattern Analysis:** Open the modal, choose a borough and time bucket, and review the top rules in plain language.

2.3   How to Interpret

- **Hotspot Map:** Highlights locations and time windows with the highest concentration. Use it to position patrols and define coverage.

- **KPIs:** Total volume, patterns identified, number and share of high-priority patterns. These provide an at-a-glance workload indicator.

- **Priority Patterns:** Two spotlights: *Most concentrated* (tightest cluster) and *Highest volume* (largest case count). Each includes crime type, borough, premises, time bucket, and, when available, suspect/victim modes.

- **Recommendations:** Action cues derived from the executive summary (e.g., deploy in specific premises or borough; schedule afternoon rotations).

- **Pattern Analysis Modal:** Global, Borough, and Time views list the strongest rules with strength badges (confidence, lift, support) and a short outcome sentence.

3   Dashboard Panels and Navigation

3.1   Hotspot Map

A map view of spatial clusters with temporal patterns. Use it to:

- Identify persistent hotspots.

- Note peak windows (morning, afternoon, evening, night).

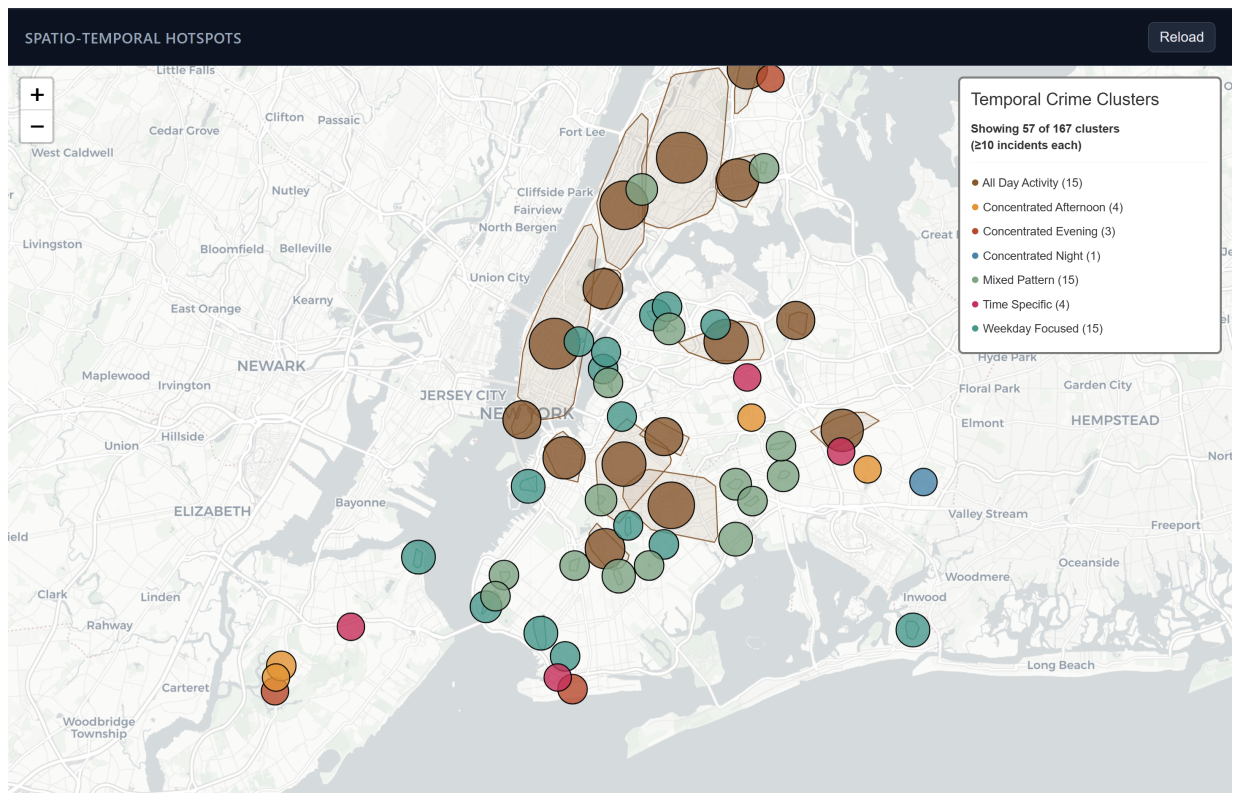- Cross-check with known events or seasonal effects.

Figure 1: Example of spatio-temporal hotspots as visualized in the dashboard map.

3.2   Executive KPIs

Cards summarize *Crimes analyzed*, *Patterns identified*, *High-priority patterns*, and *% High-priority crimes*. Rising counts suggest wider or denser clusters. A higher share of high-priority patterns signals where to focus near-term action.

3.3   Priority Patterns

Two pattern spotlights:

- **Most concentrated:** Highest density within a narrow area/time.

- **Highest volume:** Largest number of cases.

Each pattern shows crime type, borough, premises, time bucket, weekday/weekend mode, and typical suspect/victim attributes when informative.

3.4   Recommendations

Automatic cues that translate insights into actions. Examples: deploy targeted patrols over the next 7 days in afternoon windows; coordinate precincts in a specific borough; engage premises owners for prevention.

3.5 Pattern Analysis Modal

Explore the strongest association rules:

- **Views:** Global, Borough, Time.

- **Filters:** Borough chips (Manhattan, Brooklyn, Bronx, Queens, Staten Island) and time chips (Morning, Afternoon, Evening, Night).

- **Rule strength:** Each rule shows confidence (%), lift, and support; weaker rules are filtered out.

Interpreting Pattern Rules

Rules are presented as `IF (conditions) THEN (outcome)` and summarize when a given outcome becomes more likely. Read them as guidance about *context*, not causation.

- **Premises (IF):** Conditions such as borough, time bucket, premises type, and contextual factors that define a situation of interest.

- **Outcome (THEN):** The category that becomes more likely under those conditions.

- **Confidence (%):** Among incidents matching the premises, the percentage where the outcome also occurs (e.g., 72% means roughly 72 in 100).

- **Lift:** How much more likely the outcome is compared with its baseline occurrence citywide (e.g., $1.30 \Rightarrow 30\%$ more likely).

- **Support:** Prevalence of the rule in the dataset (share of all incidents matching both premises and outcome).

- **Ranking metrics:** Kulc/Score help order rules; prioritize high confidence and lift with reasonable support.

Example. `IF (BORO = MANHATTAN) AND (TIME_BUCKET = EVENING) THEN (PREMISES = RESTAURANT)`
Metrics: Conf 72%, Lift 1.30, Support 0.18.
Interpretation: In Manhattan evenings, restaurant-related incidents occur more often than baseline (about 30% higher) and cover a meaningful share of incidents. Combine with the hotspot map and KPIs to plan patrol timing and engagement with premises.

4 How the System Works

4.1 Spatio-Temporal Clustering

Incidents are grouped by geographic proximity and time-of-day windows. The method highlights compact areas that show recurring concentration. Outputs include hotspot centroids, indicative area spreads, and peak time buckets (morning, afternoon, evening, night). These inform patrol placement and scheduling decisions.

4.2   Multidimensional Clustering

Beyond location and time, the analysis considers additional attributes (e.g., premises type and contextual factors) to surface cohesive patterns. The executive layer synthesizes these clusters into KPIs (total volume, identified patterns, high-priority share) and spotlights representative patterns most relevant to operations.

4.3   Pattern Analysis

Association rules summarize condition → outcome relationships in simple sentences (e.g., *"Retail-related incidents are more likely when conditions X and Y co-occur"*). Each rule is shown with confidence, lift, and support. We filter to favor stronger evidence.

5   Data and Analysis

5.1   Data Overview

Analyses use historical complaint data enriched with temporal (hour, weekday/weekend, time bucket) and contextual features. These enable discovery of patterns across space, time, and operational contexts.

5.2   Data Sources

The analyses are grounded in reliable, comprehensive sources:

- **NYPD Complaint Data:** Historical and current crime reports from the New York City Police Department.

- **Geographic Data:** Points of Interest (POI), such as transit stations, commercial establishments, and schools.

5.3   Data Processing

To ensure accuracy, the raw data undergo rigorous preparation: cleaning, handling missing values, and integrating information from multiple sources. This is crucial for stable clustering and reliable pattern discovery.

5.4   Preprocessing and Feature Engineering

- Standardization of time features (hour bands, weekday/weekend).

- Geographic normalization for clustering stability.

- Categorical encoding for premises and other contextual fields.

5.5   Analysis Outputs

- Hotspot clusters with peak windows.

- KPI summary with high-priority share.

- Priority patterns (most concentrated, highest volume).

- Top association rules by global, borough, and time bucket views.

6   Model Validation

6.1   Cluster Quality

Configurations are assessed with complementary diagnostics and practical criteria. We use the **noise ratio score** to gauge cohesion and separation. In tandem, the **k-distance (k-dist) plot** provides an "elbow" that guides density or neighborhood parameters; we select thresholds near clear inflection points rather than chasing marginal gains. Beyond numeric indices, we prioritize *operational interpretability*: clusters should be spatially compact and contiguous, align with recognizable geography (e.g., transit hubs, retail corridors), and exhibit **temporal coherence** (clear peak windows rather than diffuse activity across all times).

For **multidimensional clustering**, evaluation focuses on internal separation/cohesion , *parsimony* in the number of segments, and **interpretability** of the resulting profiles. We prefer compact, well-separated segments that cover a meaningful share of incidents and exhibit stable composition under light resampling.
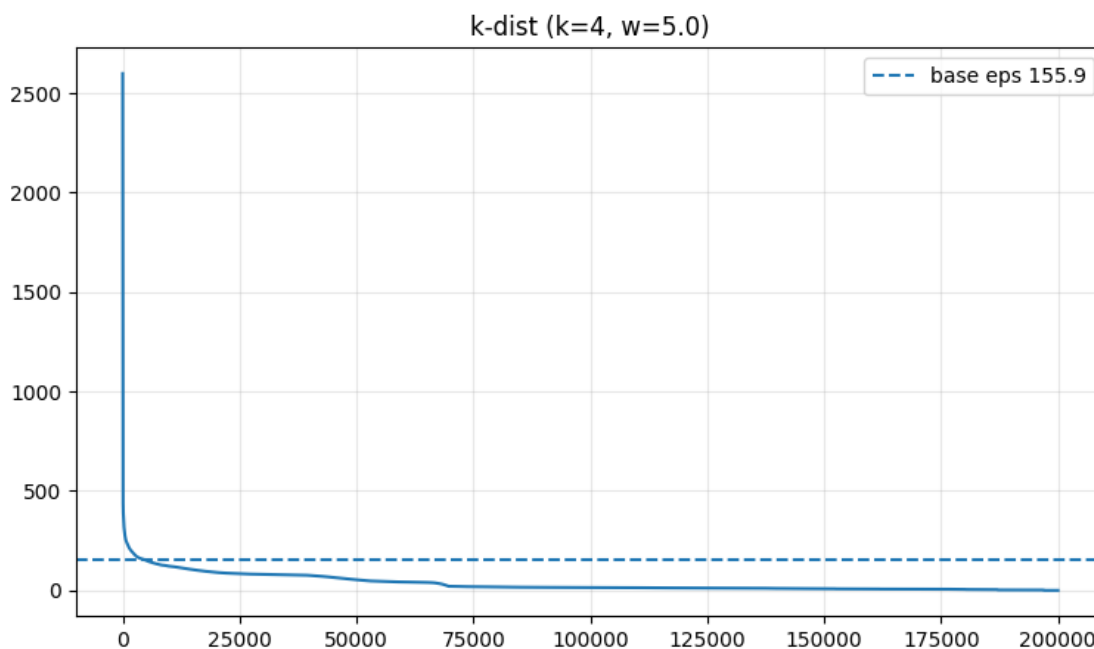

Figure 2: k-dist (k-distance) plot used to select density/neighborhood parameters for clustering.

6.2   Rule Strength

Displayed rules are screened by thresholds (confidence, lift, support) to prioritize reliable associations and reduce noise.

7   Responsible Use and Ethics

7.1   Ethical Principles

- **Non-individualization:** Use insights to guide locations, times, and premises for prevention; do not infer individual culpability.

- **Fairness:** Demographic attributes are descriptive of historical data and must not be used for targeting or discriminatory enforcement.

- **Transparency:** Communicate that associations indicate context, not causation; present rule strengths (confidence, lift, support).

- **Community engagement:** Incorporate local knowledge and community input when interpreting patterns.

7.2   Operational Guidance

- Combine dashboard insights with officer expertise and recent incident reports before acting.

- Review outputs periodically to detect drift or unintended effects; update tactics accordingly.

- Document key deployment decisions and their rationale when driven by pattern insights.

8   System Limitations

8.1   Known Constraints

- **Historical basis:** Patterns reflect past data and may shift with events, policy changes, or seasonality.

- **Data quality:** Missing or noisy attributes can affect clustering and rules.

- **Granularity:** Time buckets and spatial resolution trade off detail and stability.

- **Interpretation:** Associations are correlational; use with operational judgment.

8.2   Risk Mitigation

- Cross-check with recent incident logs and officer observations before acting.

- Note unusual events (holidays, disruptions) that may temporarily shift conditions.

- Prefer tactics that are reversible and review outcomes regularly to refine deployment.

9   FAQ

How often should we revisit the dashboard?

Daily for hotspot scanning; weekly for reviewing pattern rules and adjusting tactics.

What indicates immediate action?

A surge in high-priority share, a new dense hotspot, or consistent strong rules around a premises/time.

Can we compare boroughs quickly?

Yes. Use the Pattern Analysis modal's borough view and read the strongest rules per borough.

How do we choose shift times?

Align patrol windows with hotspot peak time buckets and the recommendations panel.

Do rules target individuals?

No. Rules describe context patterns; they must not be used for individual targeting.

10   Future Enhancements

- Near-real-time refresh and alerting when thresholds are exceeded.

- Seasonal and event-aware models to adapt to city calendars.

- Comparative views across weeks and precincts with trend lines.

- What-if tools to test deployment changes against recent patterns.

11   Conclusion

The Police Intelligence Dashboard unifies hotspots, executive indicators, and explainable patterns into a single operational view. Used alongside officer expertise and community knowledge, it supports timely, targeted prevention and efficient allocation of resources.