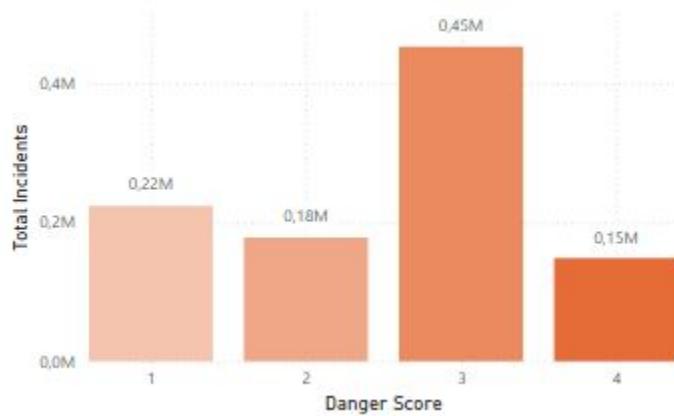


Locating Critical Threats in the Noise of Los Angeles

Identifying High-Harm Hotspots by Decoupling Frequency from Severity: A Spatiotemporal Analysis (2020-2024).

Author: Federica Brasca
Contact: f.brasca3@campus.unimib.it

1M incidents, but only a fraction are truly high-harm.
Project goal: make those critical threats visible.



Main research question

When and where in Los Angeles does recorded **crime** pose the greatest harm, combining how often incidents happen with how serious they are?

- How do **crime volume** and **danger score** distribute across the day's **time windows** (morning, afternoon, evening, night) and across LA areas?
- We usually think that the night is the most dangerous time window of the day. In LA, is the **night disproportionately dangerous**?
- What is the typical **composition of incidents** and **how does the sheer volume of low-harm incidents obscure areas of critical risk**?

Why these research questions matter

- **Resource Efficiency:** Police patrols are allocated based on call volume but in order to avoid misallocations, a danger score would make the difference (see [here](#)*).
- **The "Danger Matrix":** Distinguishing between high-frequency petty theft and low-frequency violent crime changes how a "safe neighborhood" is defined.

Intended audience and use cases

This analysis is imagined for **policy analysts** and **city planners** who need an evidence-based picture of when and where crime harm is concentrated, **LAPD leadership** who want a high-level view of patterns in recorded incidents, **journalists** and **citizens** interested in how danger is distributed in time and space in Los Angeles.

* Analysis confirms that current patrol allocation prioritizes volume over harm, evidenced by the fact that police are dispatched to 88% of all calls despite 62.6% being non-criminal nuisances and less than 8% involving violence. Consequently, the report concludes that police are "inappropriate responders" for this administrative volume and recommends restricting their involvement to situations posing "imminent threats".

About Data

The **dataset** used for this project is: *Los Angeles Police Department (LAPD) – Crime Data from 2020 to Present* ([link](#)) released under the **license CCO 1.0** (Public Domain Dedication). Data can be freely reused for any purpose.

The City of Los Angeles Open Data Portal ([data.lacity.org](#)) warns about known quality limitations, including potential **coding inaccuracies from paper transcription** and **masked/missing/rounded coordinate data** used for privacy.

To ensure the **integrity of the analysis** despite these issues:

- I focused on **aggregated patterns** rather than on single incidents, so occasional coding errors have **less impact**.
- I utilized the higher-level **Area field** instead of relying on exact locations to avoid over-interpreting noisy spatial detail.

About Data

I recognized other **several limitations**, particularly the **high volume of missing values** in descriptive fields and the challenges posed by **incomplete 2024 reporting** due to a system transition.

To mitigate these issues, I undertook **comprehensive data cleaning, standardizing** key features, dealing with missingness through **imputation**, and carefully **normalizing** the time-series data to account for the major methodological shift and ensure accurate comparisons.

Overall, I treat the LAPD crime data as one imperfect but useful measurement: it allows me to build a relative picture of crime harm across time and space, **not an exact “truth” about all crime in Los Angeles**.

Methodology

Python was the primary software used for **processing and cleaning**. Key libraries included **pandas** for data manipulation, **missingno** for initial data quality assessment, **NumPy** for numerical stability, and **matplotlib** and **seaborn** for static visualizations. The cleaned and transformed data was intended for further **interactive dashboard** creation in **Power BI**.

A **feature engineering** step allowed me to calculate an ordinal 'Danger Score' (1-4) based on the legal severity (Part I/II) and weapon involvement to measure crime harm.

As anticipated, I employed year-**normalization** for aggregated metrics to mitigate skew from the incomplete 2024 reporting period (due to the NIBRS system migration) and dropped 2025 data to avoid considering an incomplete year*.

A **categorization** step allowed me to group detailed crime descriptions into broader categories (CrimeGroups), and convert Hour to TimeOfDay buckets (e.g., Night, Morning).

* This project is being conducted in November 2025 so the year is not fully completed.

Methodology

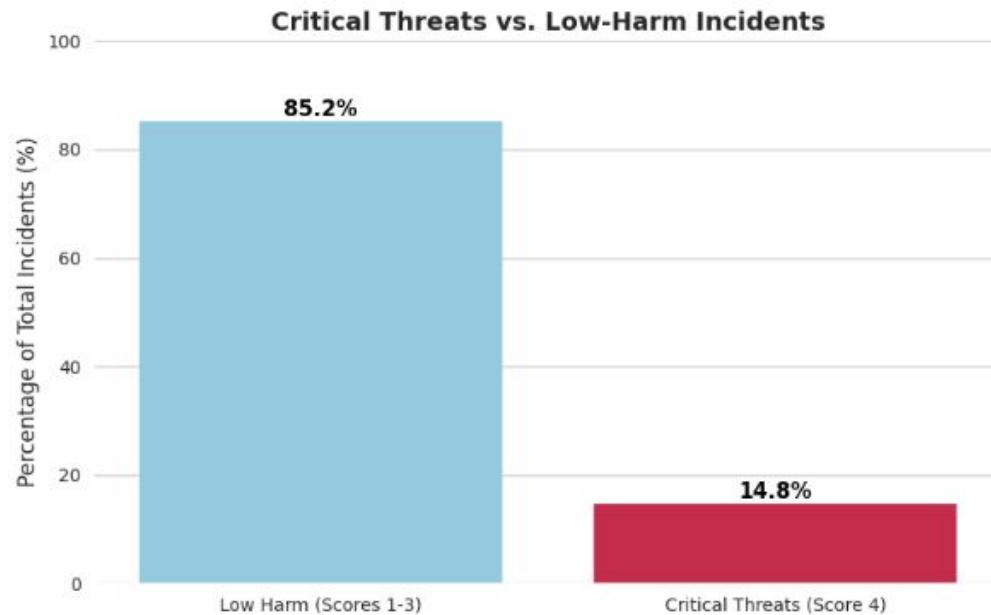
Insights were derived through **descriptive statistics** (e.g., volume counts, distribution of the Danger Score), **comparative analysis** (comparing high-harm vs. low-harm crime peaks), **segmentation (cluster analysis)** and **trend analysis** visualized via **heatmaps**, **density plots** and **interactive graphs** on Power BI.

The analysis is **fully replicable** as the data is from a public source and the entire process, including data cleaning, transformation, and visualization code, is documented and shared in a Colab Notebook (inspectionable [here](#)). Power BI interactive dashboards can be inspected by request.

Volume Creates the Noise: low-harm crimes dominate 85% of all incidents

Police patrols respond to any type of call but a lot of them are nuisances.

By setting the threshold for critical threats strictly at **Score 4** (major crime threatening people's lives), we find that: **85.2%** of all reported incidents fall into the **low harm** (Scores 1-3) category. Only **14.8%** of incidents are defined as **critical threats** (Score 4).



The Primary Problem: noisy volume hides harm

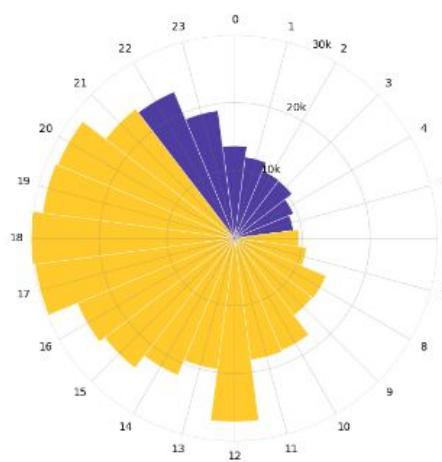
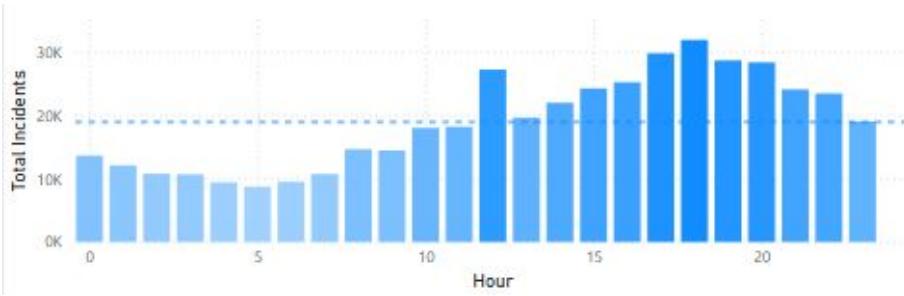
Standard reporting treats every incident the same. A stolen bike and an armed robbery both appear as "1 crime" in a map or daily briefing. **This hides critical threats.**

In the collected data, the **noisy layer** (Scores 1-3) is dominated by **high-volume property crime**: theft/larceny, vehicle crime, burglary, minor assault, and vandalism. But in the **critical layer** (Score 4) over a third of incidents are armed assaults batteries, another quarter are robberies, and a large share involve weapons or domestic violence. **The incidents that dominate volume are not the incidents that truly endanger people.**

### Composition of Low Harm Incidents (Scores 1-3) ###			
	CrimeGroup	Low_Harm_Count	Contribution_to_Low_Harm_Percent
0	THEFT / LARCENY	209189	24.45
1	VEHICLE CRIME	124619	14.57
2	BURGLARY	114118	13.34
3	ASSAULT / BATTERY	101662	11.88
4	PROPERTY DAMAGE / ARSON / VANDALISM	88362	10.33

### Composition of Critical Threats Incidents (Score 4) ###			
	CrimeGroup	Critical_Count	Contribution_to_Critical_Percent
0	ASSAULT / BATTERY	54572	36.70
1	ROBBERY / PERSONAL THEFT	38837	26.12
2	WEAPONS	16633	11.19
3	DOMESTIC VIOLENCE	12656	8.51
4	BURGLARY	11591	7.79

A day-night comparison of crime volume and harm



Standard reporting often leads to another problem: the **misperception** that crime is purely a daytime or evening problem because that is when volume peaks...

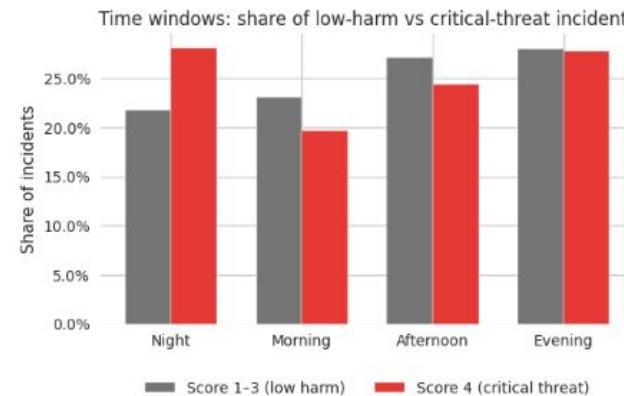
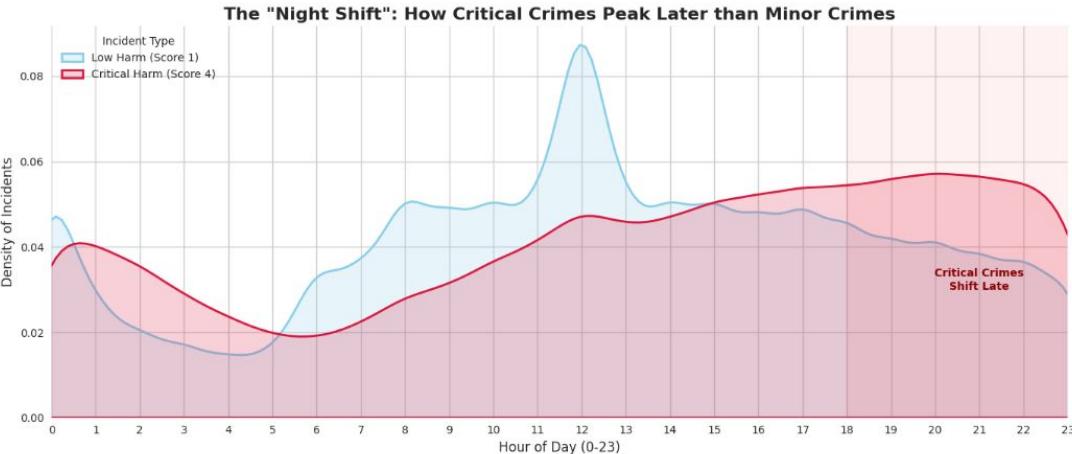
In reality, this volume is driven by **opportunity crimes** (theft, burglary) that happen when people are awake and moving.

The Reality of Risk:

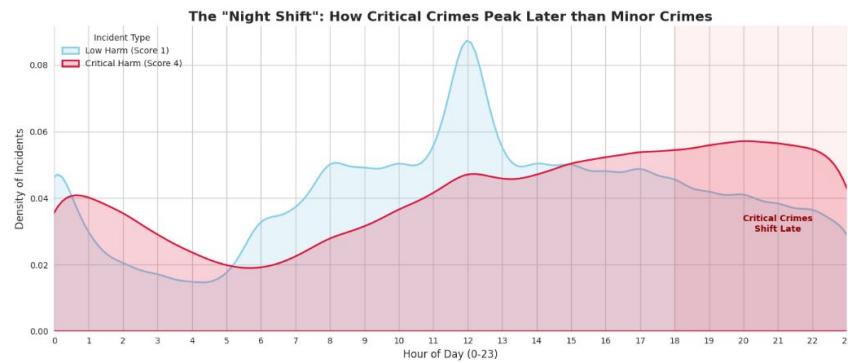
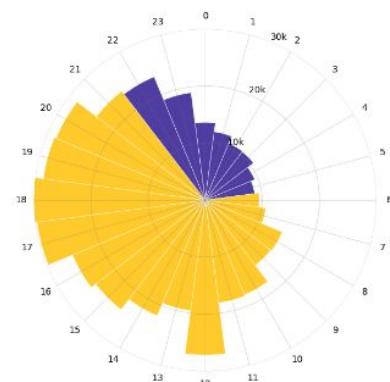
The analysis reveals that volume hides risk.

Minor crimes follow the city's **waking rhythm**, dropping to near zero by early morning (2 AM - 5 AM).

Critical crimes stay **dangerously active during the night**. While the overall volume is low, the probability of an incident being a Level 4 (Homicide, Aggravated Assault) hits its **relative peak**.



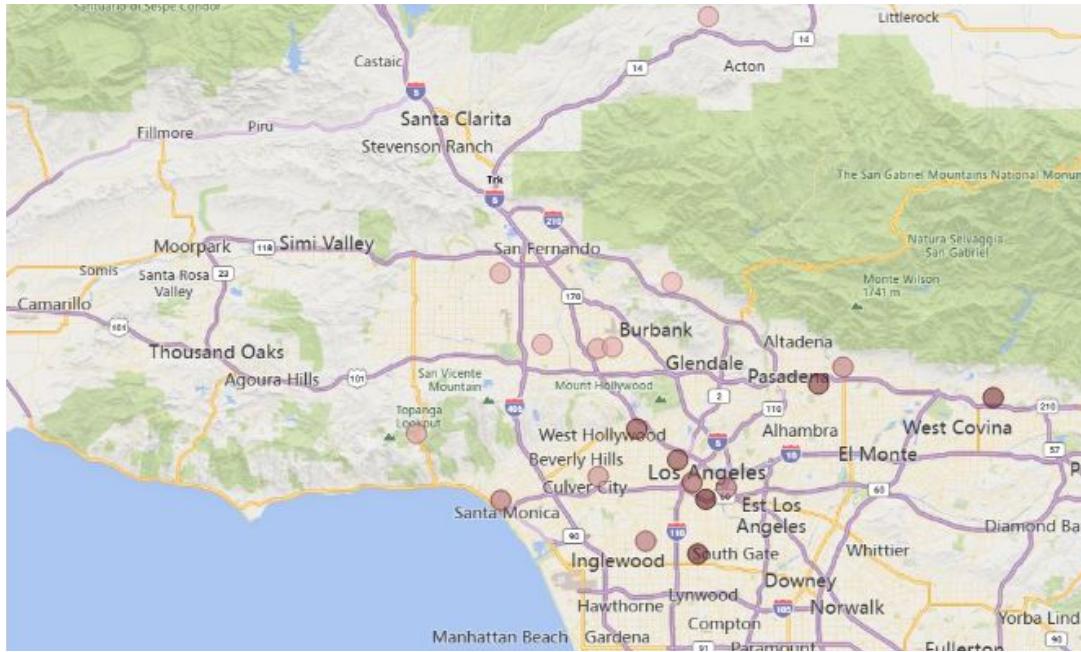
Petty criminals sleep at night. Critical threats do not.



Conclusion: While volume drops significantly at night, the danger per incident spikes. The data proves that the 'Night Shift' (fewer but more dangerous crimes) is not a time for lower staffing, but for a different type of resource: **shifting from administrative handling in the day to tactical readiness after 17 PM until early morning.**

Areas where crime harm clusters

HAVING IDENTIFIED WHEN CRITICAL THREATS PEAK, LET'S NOW EXAMINE WHERE THEY CONCENTRATE



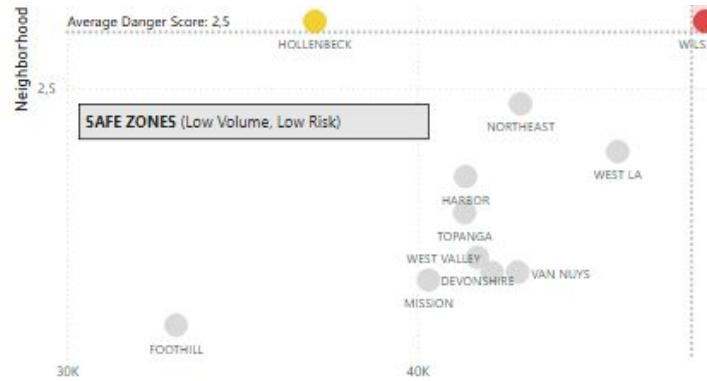
Crime happens everywhere in Los Angeles. To have a better geospatial insight on where crime is really dangerous in LA, we have to decouple frequency from severity.

The Risk Matrix



Traditional maps hide the truth by treating all crimes equally. By plotting Volume (x-axis) against Harm Score (y-axis), it is possible to segment Los Angeles into actionable tactical quadrants.

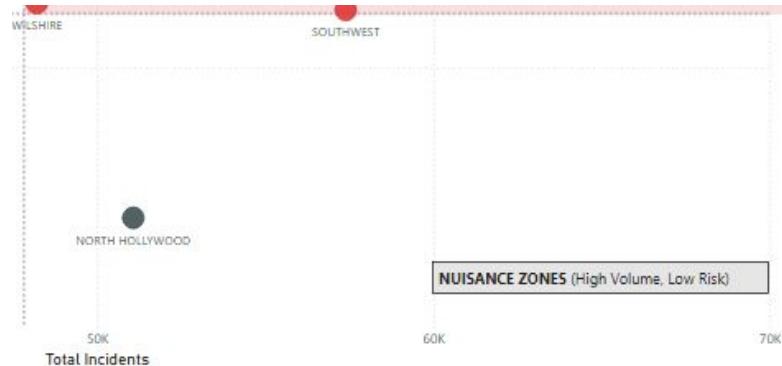
The “Safe” Zone: bottom-left quadrant



Areas such as Foothill and Mission have both **low incident volume** and **low severity scores**. Weapons are often associated with minor crimes. These areas require **minimal police intervention**, allowing resources to be safely redeployed to higher-risk zones without compromising public safety.

Strategy: Routine patrol and maintenance.

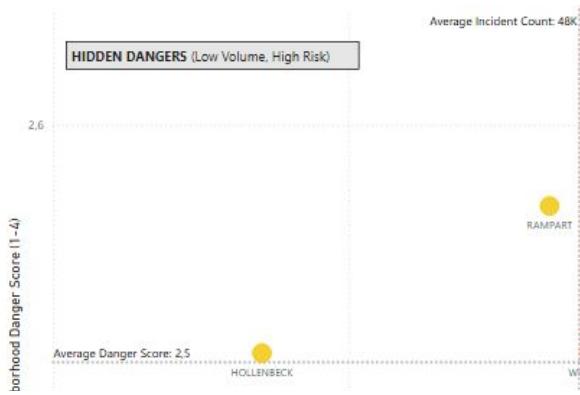
The “Nuisance” Zone: bottom-right quadrant



Areas like North Hollywood have **high call volume** but a **lower average harm score** (dominated by theft/burglary). Weapons are often associated with minor crimes.

Strategy: Administrative reporting & community support.

The "Hidden Dangers": top-left quadrant



Areas like Hollenbeck and Rampart appear "safe" on volume maps due to **lower counts**, but they sit **high on the harm axis**. This signals a **dense cluster of violent crime** that volume-based allocation misses. Weapons are often associated with major crimes.

Strategy: Targeted investigation.

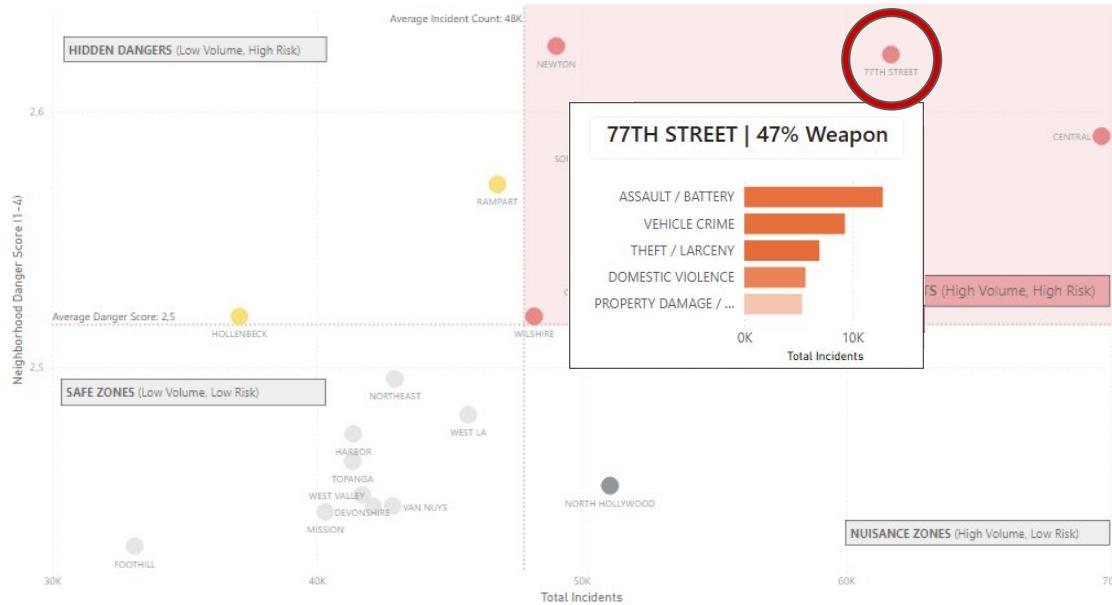
The “Critical” Zone: top-right quadrant



This quadrant represents the convergence of **high frequency** and **high severity**. These hotspots represent the **most significant threat to public safety**, where calls for service are not only frequent but statistically far more likely to involve violence. Weapons are often associated with major crimes.

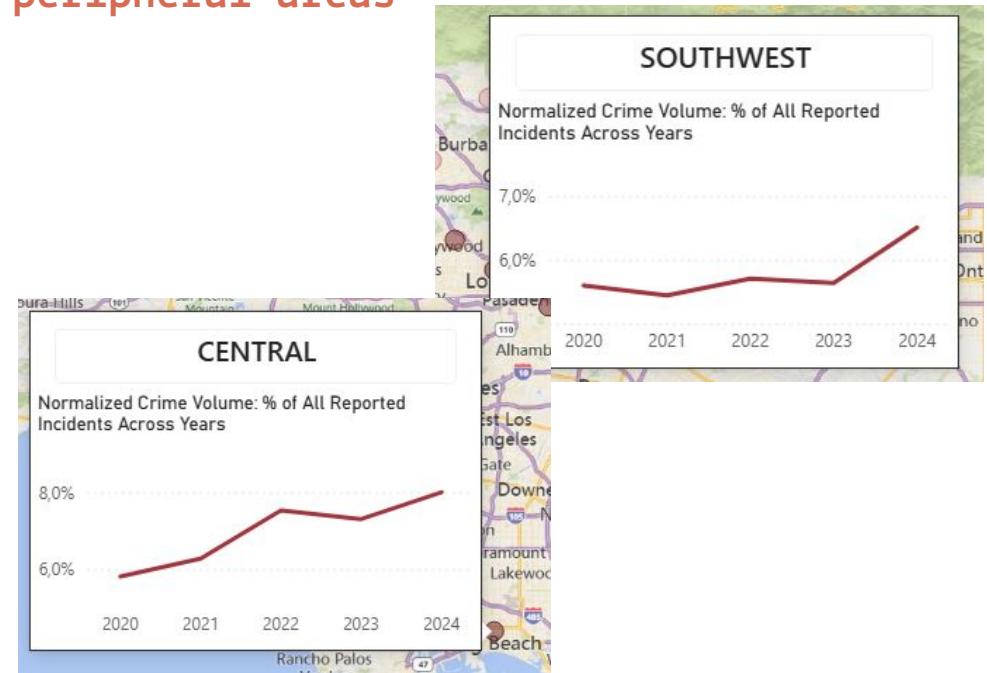
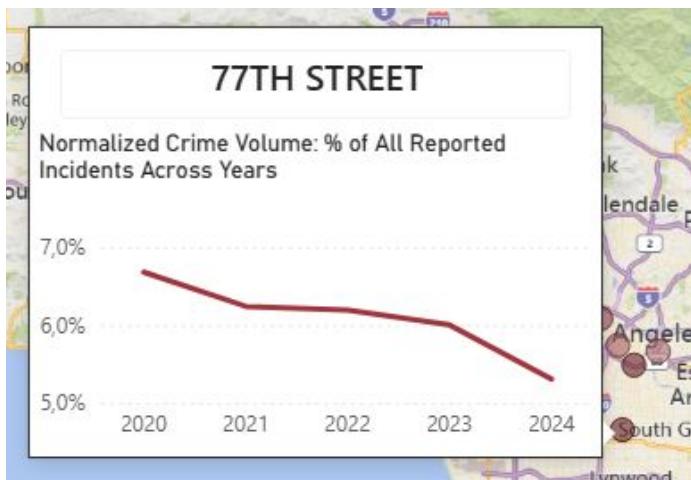
Strategy: Maximum tactical readiness and rapid response.

Keep your eyes on...



In **77th Street**, **high volume** collides with **maximum severity**. Here 47% of incidents involve a weapon: drastically higher than the 33% city average.
This is the most dangerous area of Los Angeles.

Active peripheral areas



Crime in the **77th Street** area is steadily declining. While it remains the most dangerous area overall, the fastest-growing concern is in **Southwest** and **Central**, meaning those areas need closer monitoring and stronger prevention efforts as well.

LICENCE



Slides are shared with the following license:

'Locating Critical Threats in the Noise of Los Angeles' by Federica Brasca, is licensed under a CC BY-NC-SA 4.0