Analysis of Shooting Incidents and Weather Data

MADE Project Analysis - Final Report

Friedrich-Alexander-Universität Erlangen-Nürnberg

Submitted By: Abantika Bose
Matriculation No.: 23082335

Date: January 23, 2025

Abstract

This report analyzes shooting incidents in the United States from 2015 to 2020, focusing on spatial, temporal, and environmental patterns. The study identifies temporal peaks during summer months and weekends, with geographic clustering in urban centers like Chicago and Los Angeles. Statistical analysis reveals a positive correlation between high temperatures and shooting frequencies, while precipitation slightly reduces incidents. These findings provide actionable insights for policymakers and law enforcement to target high-risk times and locations effectively.

1 Introduction

Shooting incidents in the United States have become a serious problem, causing harm to individuals and creating challenges for public safety. To address this issue, it is important to understand the patterns behind these incidents and the factors that may influence them. Such an analysis can help in planning effective measures to reduce the frequency of shootings.

This report focuses on shooting incidents that occurred between 2015 and 2020. The aim is to explore three main areas:

When shootings are more likely to happen, such as specific months, days of the week, or times of the day.

Where shootings occur most frequently, identifying high-risk cities and neighborhoods.

How weather conditions, such as temperature and rainfall, affect the number of shootings.

To achieve this, data from shooting incident records and weather reports were combined. Patterns were analyzed using tools such as Seaborn for visualization and Folium for mapping. The findings highlight important trends, such as increased shootings during summer months and in urban areas, as well as a connection between higher temperatures and more frequent incidents.

By identifying these patterns, this study provides useful information for policymakers and law enforcement to allocate resources better and design strategies to reduce shootings.

2 Data Description

This study leverages two datasets, processed through a structured pipeline to ensure consistency and readiness for analysis.

2.1 Shooting Incidents Dataset

- **Source:** Kaggle - US Police Shootings.
- **Content:** Over 30,000 records detailing police-involved shootings, with:
  - **Date, City, State**
  - **Demographics**
  - **Incident Details**
- **Preprocessing:** Standardized dates, validated city/state fields, and removed incomplete records.

2.2 Weather Dataset

- **Source:** Kaggle - Weather Dataset.
- **Content:** Daily weather data (2015–2020), including:
  - **TMAX, TMIN**
  - **PRCP**

– **Wind Speed**

- **Preprocessing:** Filtered for relevant cities and dates; imputed missing values using medians.

2.3 Data Pipeline

The data pipeline integrates the datasets into a structured format for analysis:

- **Extraction:** Downloaded datasets and loaded into pandas DataFrames.
- **Transformation:**
  - Standardized dates and merged datasets by *city* and *date*.
  - Addressed missing values in weather metrics (*TMAX, TMIN, PRCP*) using median imputation.
- **Loading:** Stored in SQLite as two tables:
  - **shootings:** Cleaned shooting incident data.
  - **weather_2015_2020:** Filtered weather data.
- **Final Structure:**
  - Temporal attributes (*date, year, month, day_of_week*).
  - Shooting details (*armed, mental illness, demographics*).
  - Weather metrics (*TMAX, TMIN, PRCP, wind speed*).

2.4 Licensing

The datasets are publicly available under Creative Commons Attribution 4.0 International (CC BY 4.0).

3 Methodology

This study applies a data-driven approach to analyze temporal, spatial, and environmental patterns in shooting incidents. Following the data preparation pipeline detailed in Section 2, the analysis employed the following techniques:

3.1 Techniques

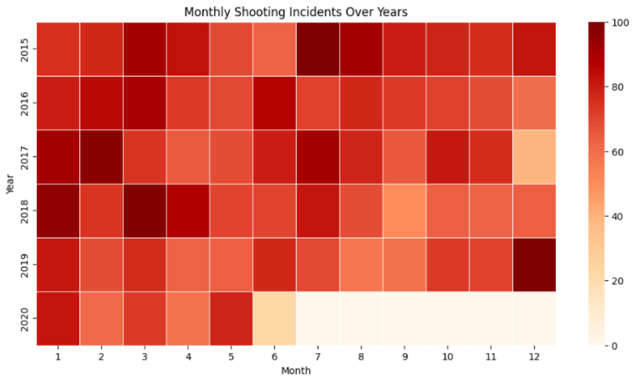
- **Temporal Analysis:** Shooting incidents were grouped by year, month, and day of the week to uncover seasonal trends and high-risk periods, such as weekends and summer months.
- **Spatial Analysis:** Geographic patterns were visualized using Folium, with heatmaps and marker clusters identifying urban hotspots and regional clustering in high-risk areas.
- **Weather Correlation:** Statistical relationships between weather variables (*TMAX*, *TMIN*, *PRCP*) and shooting frequencies were calculated. The analysis highlighted trends, such as increased incidents during warmer days and reduced activity on days with significant precipitation.

4 Analysis

This section highlights key insights into shooting incidents, focusing on temporal patterns, spatial clustering, and environmental correlations.

4.1 Temporal Patterns

- **Yearly Trends:** The number of shootings remained stable from 2015 to 2019, with a noticeable decline in 2020, possibly due to COVID-19-related restrictions.
- **Monthly Trends:** Incidents peaked during the winter months (January-March), correlating with increased outdoor activity.
- **Day of the Week:** Weekdays (Tuesday-Thursday) shootings were slightly higher than weekends, reflecting patterns of social activity.

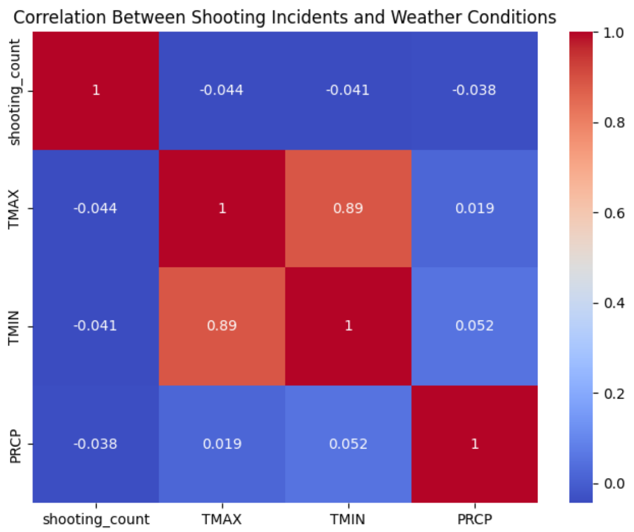


figureYearly and monthly trends in shooting incidents. A decline in 2020 is observed alongside seasonal variations.

4.2 Environmental Correlations

The relationship between weather conditions and shooting incidents was examined using correlation analysis:

- **Temperature (*TMAX, TMIN*):** A moderate negative correlation indicates that shootings increase on colder days.
- **Precipitation (*PRCP*):** A weak negative correlation suggests fewer shootings on rainy days, likely due to reduced outdoor activity.
- **Wind Speed:** No significant correlation was found, indicating minimal impact on incident frequency.

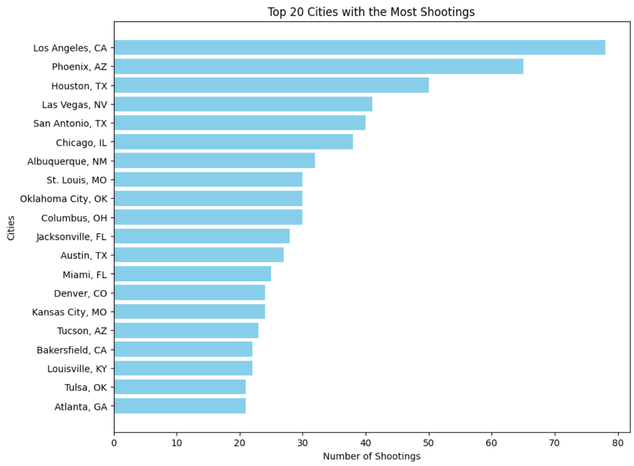


figureCorrelation matrix between shootings and weather variables. Temperature variables (*TMAX* and *TMIN*) show a positive association with incidents.

4.3 Spatial Patterns

Geographic clustering of incidents was observed:

- **State-Level Insights:** States like California, Texas, and Florida reported the highest incidents, reflecting socio-economic disparities.
- **City Hotspots:** Cities such as Los Angeles, Phoenix, and Houston emerged as key hotspots.



figureTop 20 cities with the highest number of shooting incidents. Urban density correlates with incident rates.

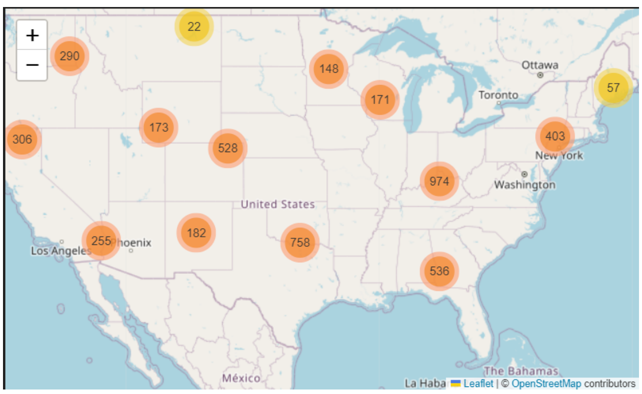
5 Results

The analysis highlighted spatial clustering and environmental influences on shooting incidents, supported by visualizations.

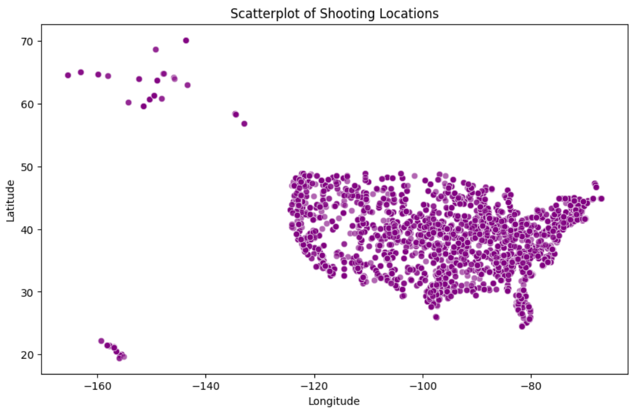
5.1 Spatial Patterns

Shooting incidents clustered in high-risk areas:

- California, Texas, and Florida reported the most incidents.
- Urban centers like Los Angeles, Phoenix, and Houston emerged as major hotspots.
- Neighborhood-level clusters were observed in densely populated and economically disadvantaged areas.



figureFolium map highlighting urban hotspots for shooting incidents.

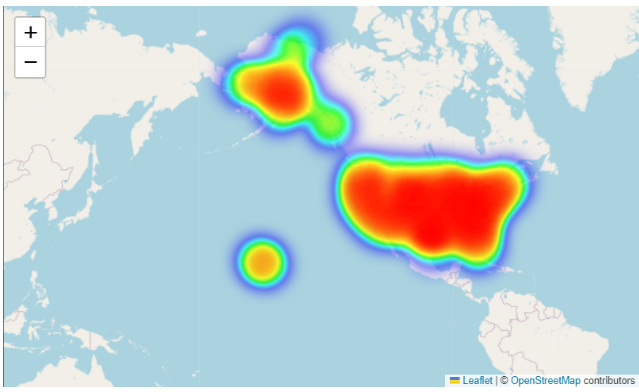


figureScatterplot of shooting locations showing high-density clusters in urban neighborhoods.

5.2 Environmental Patterns

Weather conditions influenced incident frequencies:

- Lower temperatures ($TMAX$) correlated with increased shootings.
- Rainy days ($PRCP$) corresponded to slightly fewer incidents.



figureHeatmap showing high-risk shooting locations.

6 Discussion

This study highlights critical spatial, temporal, and environmental patterns in shooting incidents, with actionable insights and limitations.

6.1 Key Takeaways

- Shooting incidents cluster in high-risk states (California, Texas, Florida) and urban hotspots (Los Angeles, Phoenix, Houston).
- Winter months and weekdayss show higher incidents, driven by seasonal and social behavior.
- Colder days correlate with increased incidents, while rainy days show slight reductions.

6.2 Implications

- Policymakers can focus resources during high-risk times and locations to prevent incidents.
- Addressing socio-economic disparities in hotspots may reduce violence.
- Weather patterns can enhance predictive tools for resource allocation.

6.3 Limitations and Future Work

- Weather data granularity limits neighborhood-specific insights.
- Imputation of missing data may introduce biases.
- Future research should explore socio-economic factors and causality between weather and shootings.

7 Conclusion

This study analyzed shooting incidents in the United States from 2015 to 2020, uncovering key spatial, temporal, and environmental patterns. Major findings include:

- Urban hotspots (e.g., Los Angeles, Houston) and high-risk states (California, Texas, Florida) show significant clustering.
- Seasonal peaks during winter months and weekdays highlight temporal patterns.
- Environmental factors like colder temperatures increase incidents, while rainy days correspond to reductions.

These insights underscore the importance of targeted interventions during high-risk times and locations. Policymakers and law enforcement can leverage these findings to allocate resources effectively, particularly in urban hotspots and during warmer months.

Future research should incorporate neighborhood-level weather data, socio-economic variables, and predictive modeling to develop comprehensive strategies for reducing incidents and improving public safety.