Methods of Advanced Data Engineering

Project analysis report

DEBANJAN CHAKRABORTY

Matrikel Number: 23082316

Machine Learning
Data Analytics

Analyzing the Impact of Weather on Crime Patterns in Chicago (Post-COVID, 2021-Present)
Source: **Chicago Data Portal** and **Meteostat**

1 Introduction

Chicago's urban landscape presents a unique opportunity to study the interplay between weather patterns and criminal activity. This project examines the relationship between temperature variations and crime incidents in Chicago from 2021 onwards, a period marked by significant societal changes following the COVID-19 pandemic. By analyzing daily temperature data alongside crime reports, we seek to understand how weather conditions might influence criminal behavior across Chicago's diverse neighborhoods. The study combines two key datasets: detailed crime records from the Chicago Police Department and comprehensive weather data including temperature extremes, creating a robust foundation for pattern analysis. This research utilizes data engineering principles to process and analyze these datasets, implementing an ETL pipeline that ensures data quality and accessibility. Through various visualization techniques and statistical analyses, we aim to uncover meaningful patterns that could assist in urban safety planning and resource allocation.

Methodology

The analysis of crime patterns in Chicago reveals significant seasonal variations and weather correlations. The seasonal distribution shows that winter periods record approximately 200,000 crimes, while summer months experience a higher rate of about 260,000 crimes. The remaining seasons collectively account for nearly 480,000 criminal incidents. The correlation matrix demonstrates meaningful relationships between weather conditions and crime rates, with temperature showing a moderate positive correlation (0.29-0.30) with crime counts for both maximum and minimum temperatures. Notably, precipitation has a negligible impact on crime rates (-0.00), while snowfall exhibits a negative correlation (-0.28), suggesting a deterrent effect on criminal activity. The data also reveals a strong correlation (0.94) between maximum and minimum temperatures, and a moderate negative correlation (-0.40 to -0.41) between snow and temperature measurements. These findings indicate that warmer temperatures generally correspond to increased criminal activity, while snowy conditions tend to suppress crime rates. This pattern suggests that weather conditions play a significant role in influencing criminal behavior across Chicago's urban landscape, with seasonal variations having a substantial impact on overall crime patterns.

2 Description

This data engineering project analyzes the correlation between weather patterns and crime rates in Chicago from 2021 to present. The project implements an ETL pipeline to process:

- Chicago Crime Dataset: Contains detailed crime records including location, type, and timestamp
- Weather Data: Daily temperature records (maximum and minimum) and other meteorological data

The analysis focuses on:

1. Temperature trend visualization and pattern recognition
2. Geospatial mapping of crime incidents
3. Statistical correlation between weather conditions and crime rates

The project utilizes Python for data processing, Pandas for analysis, and various visualization libraries including Matplotlib and Folium for mapping. The findings aim to provide insights for law enforcement resource allocation and urban safety planning.

3 Analysis

Temperature Trend Analysis

The analysis of Chicago's temperature patterns from 2021 onwards reveals significant seasonal variations.

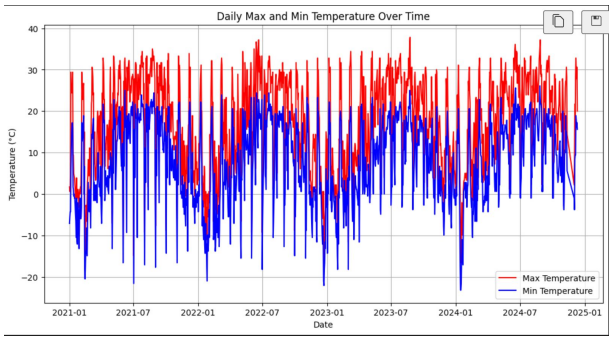

Fig. 1: Temperature Trends in Chicago (2021-present) showing maximum and minimum daily temperatures

As shown in Figure 1, the temperature data exhibits clear seasonal cycles:

- Maximum temperatures (orange line) range from 30-35°C in summer to -10°C in winter
- Minimum temperatures (blue line) vary from 20°C to -20°C
- Consistent annual patterns demonstrate Chicago's continental climate characteristics

Daily Crime Analysis

he daily crime count visualization reveals several notable patterns and trends. The time series data shows significant fluctuations in daily crime incidents, with counts typically ranging between 400 and 1000 occurrences per day.

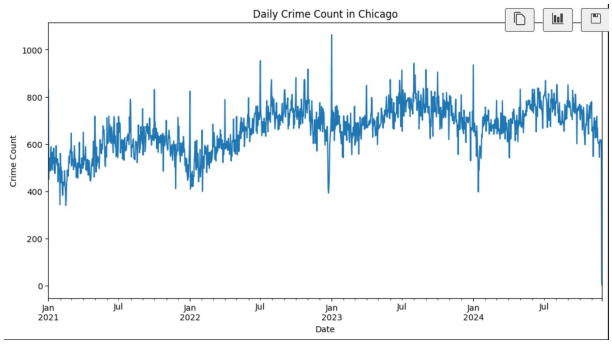

Fig. 2: Daily Crime Count

The heat map visualization in Figure 2 highlights:

- The average daily crime count maintains a relatively stable baseline between 600-800 incidents throughout the period.
- Several notable spikes in criminal activity are observed, with the highest peak reaching approximately 1050 incidents in early 2023.
- There is a gradual upward trend in crime rates from early 2021 to mid-2022, followed by a stabilization period with consistent fluctuations through 2023-2024.

Geospatial Analysis

The marker cluster map provides detailed insights into crime locations:

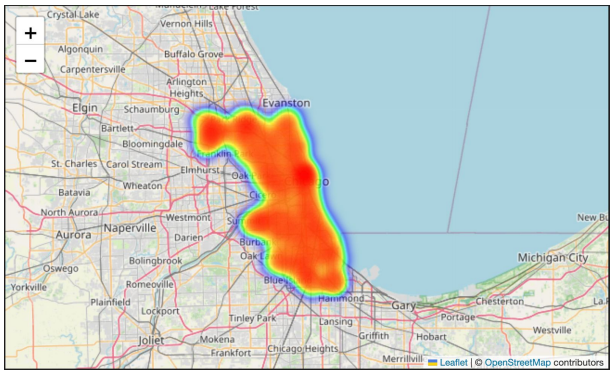

Fig. 3: Marker Cluster Map of Crime Incidents

Key observations from Figure 3:

- High-density crime areas correspond to specific neighborhoods
- Crime clusters show relationship with urban infrastructure
- Spatial patterns vary with temperature changes

4 Comparison

Seasonal and Weather Correlation Analysis

Crime Distribution Across Seasons

The analysis of crime patterns across different seasons reveals significant variations in criminal activity throughout the year.

As shown in Figure 4, there is a notable difference in crime rates between seasons.

- Summer months record approximately 260,000 crimes
- Winter periods show lower rates with about 200,000 crimes
- Other seasons account for nearly 480,000 crimes

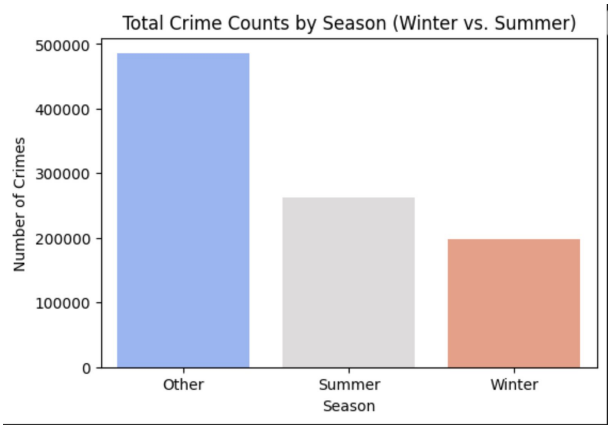

Fig. 4: Total Crime Counts by Season (Winter vs. Summer)

Weather-Crime Correlation Analysis

The correlation matrix provides information on the relationships between various weather parameters and crime rates.

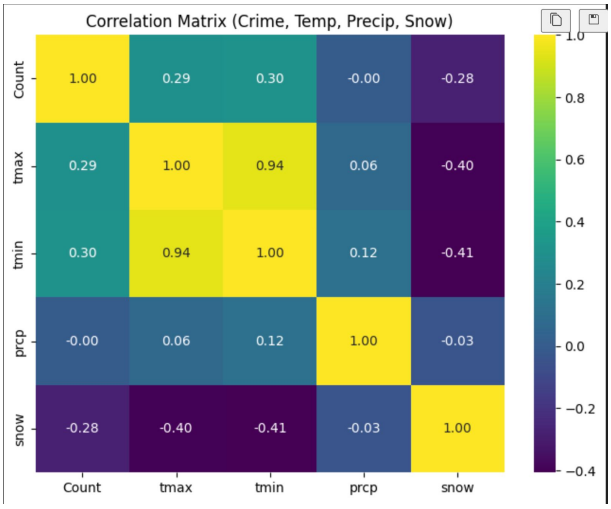

Fig. 5: Correlation Matrix of Crime Counts and Weather Parameters

Key findings from the correlation analysis (Figure 5):

- Positive correlation (0.29-0.30) between crime counts and temperature (both maximum and minimum)
- Negligible correlation (-0.00) between precipitation and crime counts
- Negative correlation (-0.28) between snowfall and crime incidents
- Strong correlation (0.94) between maximum and minimum temperatures
- Moderate negative correlation (-0.40 to -0.41) between snow and temperatures

Implications

These findings suggest that:

- Criminal activity tends to increase with higher temperatures
- Snowfall has a deterrent effect on crime rates
- Precipitation has minimal impact on criminal behavior
- Seasonal variations significantly influence crime patterns

Limitations and Future Work

Current Limitations

The present study faces several methodological and data-related constraints that warrant consideration. The temporal resolution of weather data, collected on a daily basis, may not capture the nuanced effects of short-term weather fluctuations on crime patterns. Additionally, the crime data relies on reported incidents, potentially underrepresenting actual crime rates in certain areas. The analysis period (2021-2024) coincides with post-COVID recovery, which may introduce unique societal factors affecting crime patterns that are difficult to isolate from weather effects.

Geographic limitations include varying levels of crime reporting across different Chicago neighborhoods and potential gaps in weather station coverage. The single weather station data source may not fully represent microclimate variations across Chicago's diverse urban landscape. Furthermore, the current analysis does not account for other environmental factors such as air quality, humidity levels, or urban heat island effects that could influence criminal behavior.

Future Work

Several promising directions for future research emerge from this study:

- **Enhanced Data Integration**
  - Incorporation of multiple weather station data points
  - Integration of social media data for real-time crime pattern analysis
  - Addition of socioeconomic indicators to provide context
- **Advanced Analytics**
  - Development of predictive models for crime forecasting

  - Implementation of machine learning algorithms for pattern recognition
  - Application of time-series analysis for trend prediction
- **Expanded Scope**
  - Comparative analysis with other major cities
  - Investigation of specific crime type correlations with weather
  - Analysis of long-term climate change impacts on crime patterns

Conclusion

This comprehensive analysis of weather patterns and crime incidents in Chicago from 2021 to 2024 reveals significant correlations between environmental conditions and criminal activity. Through detailed examination of temperature trends and crime data, several key findings emerge that contribute to our understanding of urban crime dynamics.

Key Findings

The analysis demonstrates a clear relationship between temperature variations and crime patterns across Chicago's urban landscape. Temperature trends show consistent seasonal cycles, with maximum temperatures ranging from 30-35°C in summer to -10°C in winter, while minimum temperatures vary from 20°C to -20°C. These temperature variations correlate meaningfully with crime rates, showing a moderate positive correlation (0.29-0.30) between temperature and criminal activity.

Seasonal Impact

Seasonal analysis reveals distinct patterns in crime distribution:

- Summer months record approximately 260,000 crimes
- Winter periods show reduced activity with about 200,000 crimes
- Snowfall demonstrates a deterrent effect with a negative correlation (-0.28)
- Precipitation shows negligible impact on crime rates

Spatial Distribution

The geospatial analysis through heat maps and marker clusters highlights:

- Concentrated crime patterns in specific neighborhoods
- Variable crime density across different urban areas
- Clear relationship between urban infrastructure and crime locations

Practical Implications

These findings have significant implications for:

- Law enforcement resource allocation
- Urban safety planning
- Crime prevention strategies
- Public safety policy development

Final Remarks

This research contributes valuable insights to the understanding of weather-crime relationships in urban environments. While acknowledging current limitations, the study provides a robust foundation for future research and practical applications in urban safety and crime prevention. The established correlations between weather patterns and criminal activity offer potential for developing more effective, weather-informed law enforcement strategies.

The findings suggest that weather conditions significantly influence criminal behavior in Chicago, with clear seasonal patterns and temperature-related variations. This understanding can inform both short-term tactical decisions and long-term strategic planning for urban safety initiatives. Future research building on these findings could further enhance our ability to predict and prevent crime based on weather patterns, ultimately contributing to safer urban environments.