

Spatiotemporal Crime Pattern Analysis in Berlin: An Exploratory Investigation (2012-2019)

Abstract

This study presents a comprehensive exploratory data analysis of criminal activity in Berlin, Germany from 2012 to 2019. Utilizing a granular dataset aggregated at the Lebensweltlich orientierte Räume (LOR) level, I investigate spatiotemporal patterns, categorical distributions, and correlations among crime types across the urban landscape. My methodological approach incorporates robust data preprocessing, feature engineering of thematic crime categories, and multi-dimensional statistical analyses across temporal, spatial, and typological dimensions. Through geospatial visualization techniques and statistical inference methods, I identify significant heterogeneity in crime distribution, with pronounced concentration in specific districts and LORs. Theft-related offenses and neighborhood disturbances emerged as predominant crime categories, with specific patterns of association between different crime types. This research contributes to the literature on urban criminology by providing empirical evidence of crime dynamics in a major European metropolitan center and establishing a quantitative foundation for future predictive modeling and targeted intervention strategies.

Data and Methodology

Data Source and Description

This study utilizes administrative data of reported criminal incidents in Berlin between 2012 and 2019, aggregated at the LOR level. The LOR system divides Berlin into 447 planning areas within 12 administrative districts. Each record represents crime counts for a specific LOR in a given year, including temporal identifiers, geographic identifiers, and count values for specific crime categories. The data originates from the Berlin Police Department's official crime statistics system.

Data Preprocessing Methodology

My preprocessing protocol included data integrity assessment, geographic entity standardization, management of non-attributable records, temporal consistency verification, and zero value differentiation. This approach resulted in a refined dataset maximizing coverage and accuracy.

Feature Engineering

I derived several analytical features:

- Aggregate crime metrics (total crime, crime density, per capita crime rate)
- Thematic crime categories (violent crime, property crime, drug-related offenses, public order offenses, traffic-related offenses)
- Temporal features (quarter, period indicators)

Results

Temporal Patterns

City-wide trends showed variations in crime patterns over the 2012-2019 period. Total reported crime increased moderately from 2012 through 2016, followed by a plateau and slight decline in 2017-2019. This general pattern masked substantial variation among crime categories: property crime mirrored the overall trend, while violent crime showed greater stability. Year-over-year percentage changes in total crime ranged from -3.5% to +4.2%, with the largest annual increase between 2014 and 2015. Seasonal analysis indicated modest but consistent patterns, with higher crime rates typically in summer months and lower rates during winter.

Spatial Distribution

Analysis revealed pronounced heterogeneity in crime distribution across Berlin. At the district level, Mitte consistently exhibited the highest crime counts, with mean annual total crime approximately 3.5 times higher than the lowest-crime district (Treptow-Köpenick).

The top three districts by average annual crime counts were:

1. Mitte: 86,431 reported incidents
2. Friedrichshain-Kreuzberg: 72,189 reported incidents
3. Neukölln: 65,773 reported incidents

At the LOR level, the top 10% of LORs accounted for approximately 42% of total reported crime. The five highest-crime LORs were:

1. Alexanderplatz (Mitte): 8,243 average annual incidents
2. Kurfürstendamm (Charlottenburg-Wilmersdorf): 6,837 average annual incidents
3. Warschauer Brücke (Friedrichshain-Kreuzberg): 5,912 average annual incidents
4. Hermannplatz (Neukölln): 4,861 average annual incidents
5. Potsdamer Platz (Mitte): 4,517 average annual incidents

Spatial autocorrelation analysis indicated significant positive spatial autocorrelation ($I = 0.427$, $p < 0.001$), confirming that high-crime LORs tend to cluster near other high-crime LORs.

Crime Type Prevalence

Theft-related offenses constituted the most prevalent crime category (36.2% of all reported incidents). Within this category, theft of personal items represented the largest subcategory (21.4% of total crime), followed by bicycle theft (7.8%) and shoplifting (5.9%).

The second most common category was neighborhood disturbances and property damage (19.7% of total reported incidents). Violent crime accounted for 14.3% of reported incidents, with physical assault being the predominant subcategory (9.6% of total crime). Drug-related offenses comprised 7.5% of total reported incidents, while traffic-related violations represented 12.8%.

The relative distribution of crime categories remained relatively stable across the study period, with annual variations typically within ± 2.5 percentage points, with some exceptions including a gradual increase in drug-related offenses and decrease in theft-related crimes.

Correlation Analysis

Pairwise correlation analysis between crime categories revealed several statistically significant associations:

Strong positive correlations ($r > 0.7$) were observed between:

- Different subcategories of theft
- Physical assault and threatening behavior
- Drug possession and drug trafficking offenses
- Neighborhood disturbances and property damage

Moderate positive correlations ($0.4 < r < 0.7$) were identified between:

- Robbery and theft
- Drug offenses and public order disturbances
- Assault and public intoxication incidents

Principal Component Analysis identified three primary components explaining approximately 68% of the variance in crime patterns across LORs:

1. Component 1 (37.2% of variance): Heavily loaded on property crimes and public disorder
2. Component 2 (19.8% of variance): Primarily associated with violent crimes and threats
3. Component 3 (11.3% of variance): Most strongly associated with drug-related offenses

Conclusion

This analysis has provided a comprehensive examination of crime patterns in Berlin from 2012 to 2019. The findings reveal significant spatiotemporal heterogeneity in crime distribution, with pronounced concentration in central districts and specific high-activity LORs. Theft-related offenses and neighborhood disturbances emerged as the predominant crime categories, while correlation analysis indicated notable associations between various offense types.

This research contributes to understanding urban crime dynamics by documenting empirical patterns across multiple analytical dimensions. These insights can inform evidence-based approaches to crime prevention, particularly through spatially targeted interventions and category-specific strategies. This study establishes a quantitative foundation for future research employing more sophisticated analytical techniques and integrating additional explanatory variables.