WHERE DO POLICE KILL?

Spatial Predictors of Police Lethality in São Paulo City

INTRODUCTION

Violence remains a pressing issue in Brazil. In 2023, the country recorded 38,595 homicides, with **São Paulo being one of the states with the most murders recorded**. While violence in Brazil is often attributed to organized crime, **a substantial portion stems from the State.** In 2023 alone, 504 people were murdered by the police in São Paulo State, an average of **1.3 victims per day.** These figures raise questions about the nature of state violence: **where** it happens, under **what conditions**, and **who** is affected the most.

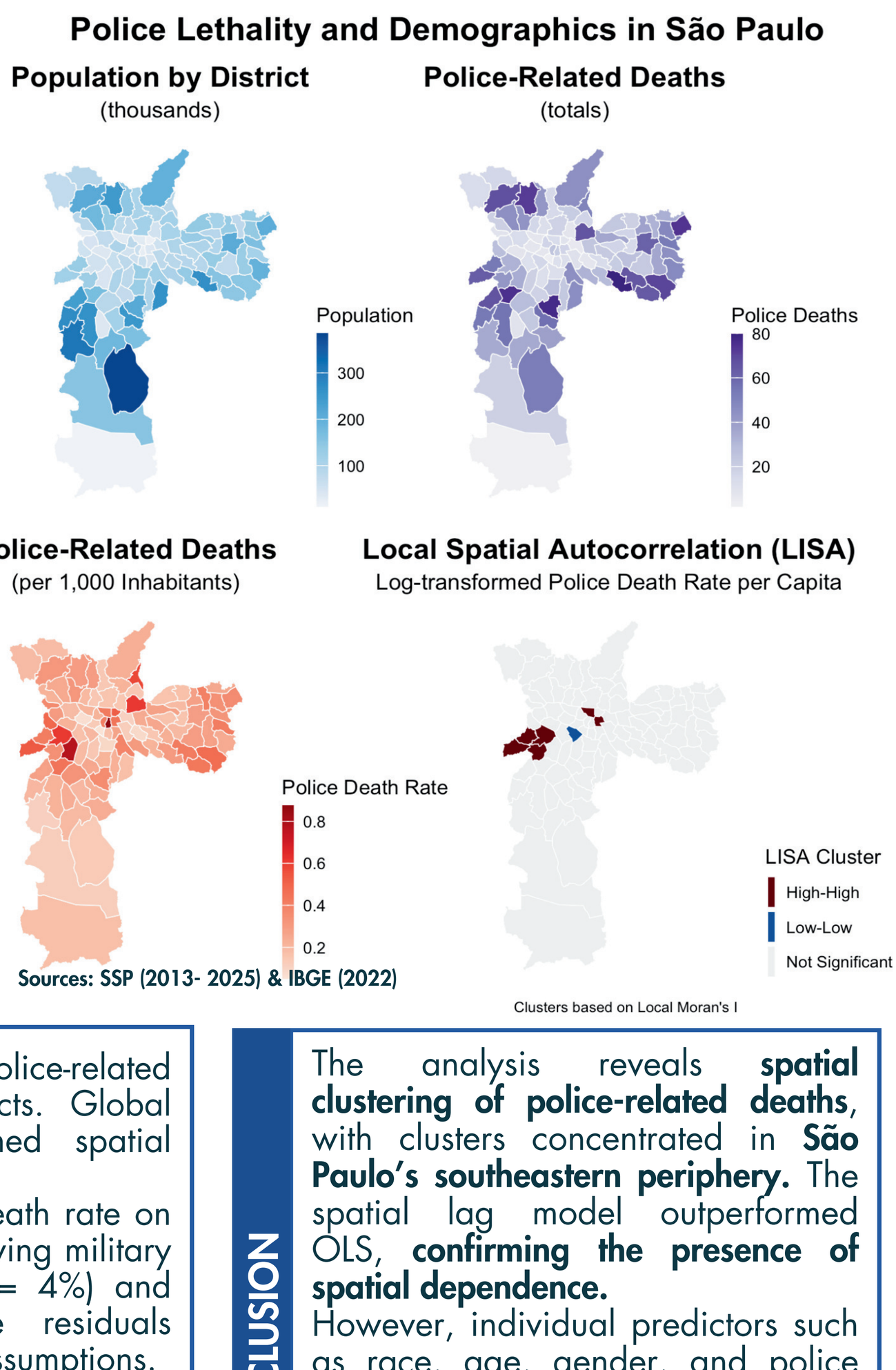
OBJECTIVE

This research aims to analyze the **spatial distribution of police killings** in São Paulo city and identify the factors that contribute to police lethality in the city.

METHODOLOGY

A spatial analysis of police-related deaths in São Paulo (2013–2025) was conducted using the following approach:

- **Data Cleaning:** Filtering the data to include police-related homicides in the city of São Paulo, removing duplicates and missing values.
- **Spatial Aggregation:** Individual deaths were geocoded and aggregated to administrative districts.
- **Variable Construction:** Created district-level variables. Death counts were normalized by population and log-transformed to reduce skewness.
- **Exploratory Spatial Analysis:** Local Moran's I (LISA) was applied to detect spatial clusters of high and low police lethality and to visualize the results through cluster maps.
- **Spatial Regression:** Comparing Ordinary Least Squares (OLS), Spatial Lag Model (SLM), and Spatial Error Model (SEM) specifications. Residual spatial autocorrelation was tested using Moran's I.



RESULTS

- **LISA:** Significant High-High clusters of police-related deaths were identified in peripheral districts. Global Moran's I (0.217, $p < 0.001$) confirmed spatial dependence.
- **OLS Model:** Regressing the log-transformed death rate on victim-level indicators (race, age, sex, % involving military police) yielded low explanatory power (R^2 = 4%) and significant spatial autocorrelation in the residuals (Moran's I = 0.18, $p < 0.01$), violating OLS assumptions.
- **Spatial Lag Model:** Accounting for spatial dependence ($\rho = 0.40$, $p < 0.001$) improved model performance, but none of the victim-level predictors were statistically significant.
- **Extended Spatial Lag Model:** Adding socio-economic variables (% favela area, income per capita, population density) improved model fit (AIC = –1418.3), yet no individual variable reached significance.

CONCLUSION

The analysis reveals **spatial clustering of police-related deaths**, with clusters concentrated in **São Paulo's southeastern periphery.** The spatial lag model outperformed OLS, **confirming the presence of spatial dependence.**

However, individual predictors such as race, age, gender, and police affiliation were not statistically significant.

Future analyses should incorporate **contextual variables** such as other crime indicators and the presence of organized crime to better explain spatial patterns of police lethality.

AUTHORS

Monserrat López, Rodrigo Dornelles,
Luis Fernando Ramírez

GitHub Repository

Hertie School