Final Project: Policing data

YOUR NAME HERE

2025-04-17

Abstract

Add your abstract here

GUIDELINES: For those not completing an original data analysis for the final project, you will use data on police-involved fatalities in the US. Use the files mpv_1_3_24.csv and pop_st.csv.

1) Using the mpv and pop_st files, create a joined table that provides race-specific estimates of police-involved mortality as deaths per 100,000 population per state and year. Provide basic descriptive statistics to summarize the distribution of this measure.

2) Visualize the distribution of this mortality rate. Consider appropriate ways to visualize heterogeneity across states and across racial and ethnic groups.

3) Identify states with consistently high mortality levels for particular groups. Do you believe that these states have elevated mortality levels compared to their peers? Any concerns?

4) Using the population portion of the data, compute measures of racial population composition at the state-level (i.e. proportion of the total population that is Black, or the proportion that is Latine). Visualize and describe this measure.

5) Consider and describe an appropriate regression model for policy mortality using racial population composition as a predictor.

6) Estimate your preferred model. Provide appropriate output and visuals and explain your findings.

7) Using what you now know about regression, interpret the meaning of your findings and provide a discussion of assumptions, limitations, and potential directions for future research.