Final Report

Hening Cui, Jialiang Hua, Congyang Xie, Wenhan Bao, Qihang Wu

Dec 9 2021

Motivation

In the wake of George Floyd's murder, a number of protests arisen across the nation. This motivates us to investigate innocent deaths under police violence and the association of confounders with excessive police force. Although the majority encounters that meet death with law enforcement officials present a lethal threat to officers or society, there are still people that involved in the police action meets death with a quantity not to be ignored. This project analyses the data on people involved in a police operation and results in fatalities. Despite the assessment of overuse of police force across geographical factors, this project also focuses on socioeconomic factors, political factors and periodical patterns.

Questions

- What

- How

- When

Data

The main data source for our project is from Kaggle and we also combine it from other sources including the World Population Review, Statista, and the U.S. Bureau of Labor Statistics. After data cleaning and wrangling, our main dataset contains 21178 observations and 13 variables. Some key variables are listed as following:
- `unique_id` is a key variable and unique in this dataset
- `age`, `age_bin`, `gender`, and `race` represent the subject's demographic characteristics
- `date_of_death` represents the date of the subject's death
- `city`, `state`, `latitude`, and `longitude` represent the location of the death
- `highest_force` represents the cause of death
- `intended_force` represents the intentional use of force by police