# Disparities in Police Killings: Socioeconomic, Demographic, and Political Correlates

2024-11-22

## Disparities in Police Killings: Socioeconomic, Demographic, and Political Correlates

### Abstract

Police killings have become one of the top political and social issues in the United States. The present study investigates the socioeconomic, demographic, and political factors that affect police killings across various counties in the United States. Using a database of cases involving police killings, socioeconomic measures, and Poisson regression analyses, we evaluate the influence of income, poverty, unemployment, education, state political affiliation, armed status, and circumstances of death. The results establish significant disparities in police-related deaths with specific regard to racial and economic factors, while political affiliation of the states and whether the victim was armed emerged as two of the most powerful predictors. These results further speak to the pressing need for structural reforms that alleviate inequalities and work to reduce police killings going forward.

### Introduction

Police killings in the United States have been one of the most scrutinized law enforcement practices in concern with racial inequity, economic vulnerability, and political influence. The spate of highly publicized police killings has underlined the need for further understanding about how systemic factors contribute to the incidents. Literature commonly sticks to an individual case or broad racial disparities without tending to the interplay between socioeconomic conditions, the political context, and incident-specific circumstances. This paper attempts to address this gap by investigating the underlying factors that drive police killings, county by county. In particular, it looks at the role of economic indicators-eye-level median income, poverty, and unemployment rates-alongside educational attainment, state-level political affiliation, and situational factors like armed status and cause of death. It is within these dimensions that the research hopes to contribute to actionable insight into the systemic drivers of police killings, to inform broader discussions about the reform of law enforcement.

**Research Questions:**

1. What are the key socioeconomic, demographic, and political factors associated with the frequency of police killings at the county level?
2. How do armed status and the cause of death influence the likelihood of police killings?
3. What role does state-level political affiliation (Red, Blue, or Unaffiliated) play in shaping disparities in police killings?

### Data and Methods

This paper relies on a newly compiled dataset on police killings at the county level in the United States, the data is extracted from 538 . Socio-economic variables include median income, poverty rate, rate of unemployment, and educational attainment. Other demographic controls include race/ethnicity and age. Incident-specific controls are armed status and cause of death. State political affiliation was then categorized

as Red, Blue, or Unaffiliated based on the voting pattern in the most recent presidential elections. Data cleansing was done to ensure reliability and usability. Quantitative variables of age, poverty rate, median income, and household income were transformed into appropriate formats. Missing values in the data, especially in regards to key variables of income, were dealt with by removing incomplete cases. To facilitate regression modeling, categorical variables of armed status and cause of death were reclassified as factors. Data aggregation was done at the county level to enable regional analysis of police-related deaths.

While considering the socioeconomic, political, and situational factors, the number of police killings was modeled using Poisson regression. This analytical method is particularly appropriate in the context of the present study, as it efficiently handles count data, while allowing for an exploration of how the predictor variables impact event frequency. The independent variables integrated into this model included median income, poverty rate, unemployment rate, educational level, political state affiliation, armed status, and cause of death.

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

## Warning: NAs introduced by coercion
## Warning: NAs introduced by coercion
## Warning: NAs introduced by coercion
## Warning: NAs introduced by coercion
## Warning: NAs introduced by coercion
```

## Data Analysis

**Disparities in Cause of death by State Political Affiliation**

```
## The following errors were returned during `modify_header()`:
## x For variable `raceethnicity` (`state_affiliation`) and "estimate", "p.value",
##   "conf.low", and "conf.high" statistics: FEXACT error 7(location). LDSTP=18570
##   is too small for this problem, (pastp=16.7249, ipn_0:=ipoin[itp=407]=2076,
##   stp[ipn_0]=17.5824). Increase workspace or consider using
##   'simulate.p.value=TRUE'
```

Demographic and socioeconomic data of individuals dissimilarly characterized and involved in police killings, identify three political categories: blue states (Democratic Party), red states (Republican Party), and unaffiliated states (independent party). Independent variables analyzed for this study included age, gender, race/ethnicity, personal income (p_income), household income (h_income), and income at the county level (county income).

In terms of age, individuals killed, by police, appeared to share similar median ages within the political category groups. In Blue states, the median age was 36 years (IQR, 27-45), which also matched that of the Red states, (36 years, IQR, 28-45). In Unaffiliated states, the median was slightly younger at 34 years (IQR, 27-44). The p-value for the comparison was 0.7, implying a failure to reject the null hypothesis: there was no statistical significance among the political groups regarding age.

The gender distribution was also not significantly different across the political groups; a majority were male-95% of people killed were from Blue states, 97% from Red states, and 93% from Unaffiliated ones. Among females, although lower, the representation was stable at 4.6% in Blue states, 2.9% in Red states, and 7.1% in Unaffiliated. The p-value for gender was 0.2, suggesting no significant difference in the gender distribution between political groups.

| Variable | N | **Blue** N = 151[1] | **Red** N = 173[1] | **Unaffiliated** N = 14 |
|---|---|---|---|---|
| **age** | 461 | 36 (27, 45) | 36 (28, 45) | 34 (27, 44) |
| **gender** | 465 | | | |
|   Female | | 7 (4.6%) | 5 (2.9%) | 10 (7.1%) |
|   Male | | 144 (95%) | 168 (97%) | 131 (93%) |
| **raceethnicity** | 465 | | | |
|   Asian/Pacific Islander | | 6 (4.0%) | 2 (1.2%) | 2 (1.4%) |
|   Black | | 41 (27%) | 49 (28%) | 43 (30%) |
|   Hispanic/Latino | | 30 (20%) | 24 (14%) | 13 (9.2%) |
|   Native American | | 1 (0.7%) | 0 (0%) | 3 (2.1%) |
|   Unknown | | 9 (6.0%) | 3 (1.7%) | 3 (2.1%) |
|   White | | 64 (42%) | 95 (55%) | 77 (55%) |
| **p_income** | 465 | 22,969 (17,901, 30,723) | 22,080 (18,029, 27,385) | 22,750 (19,031, 27,91 |
| **h_income** | 465 | 49,318 (35,507, 63,052) | 39,561 (32,034, 53,244) | 42,027 (33,370, 53,89 |
| **county_income** | 465 | 55,909 (51,939, 72,112) | 47,335 (41,393, 51,574) | 49,669 (42,334, 54,92 |

[1]Median (Q1, Q3); n (%)

[2]Kruskal-Wallis rank sum test; Pearson's Chi-squared test

The racial breakdown by states suggests that in all three types of state political affiliation, the representation of Black individuals is about the same: 27% in Blue states, 28% in Red states, and 30% in Unaffiliated states. White individuals accounted for 42% in Blue states and 55% in both Red and Unaffiliated states, while other minority racial/ethnic groups include Hispanic/Latino, Asian/Pacific Islander, and Native American formed only minor proportions, with no statistically significant difference among the three states. The p-value for race/ethnicity was 0.1, indicating no statistically significant differences between any of the comparisons made for those groups based on race or ethnicity.

When we consider personal income (p_income), Blue states have a median p_income of $22,969 (IQR: $17,901-30,723), Red states $22,080 (IQR: $18,029-27,385), and Unaffiliated states at $22,750 (IQR: $19,031-27,917); and again, in this case, no marked difference across state political affiliations was found.

The p-value for personal income was found to be 0.3, thereby indicating that there were no significant variations in this domain between the three political groups. Statistically significant differences did exist with respect to household income (h_income). Blue states again have a higher median household income level of $49,318 (IQR: $35,507-63,052), Red states at $39,561 (IQR: $32,034-53,244), and Unaffiliated states at $42,027 (IQR: $33,370-53,891). The p-value for household income was less than 0.001, thus indicating that differences regarding household income levels among the three political groups were significant. In summary, county income (county_income) follows almost the same trend. In Blue states, the median income stands at $55,909 (IQR: $51,939-72,112) against $47,335 (IQR: $41,393-51,574) in Red states and $49,669 (IQR: $42,334-54,923) in Unaffiliated states. The p-value for county income was again less than 0.001, indicating a statistically significant difference among the political affiliations, where, as already demonstrated, Blue states exhibited the highest county income.

**Disparities in Cause of death by Race**

```
## `summarise()` has grouped output by 'raceethnicity'. You can override using the
## `.groups` argument.
```

## Disparities in Cause of Death by Race/Ethnicity
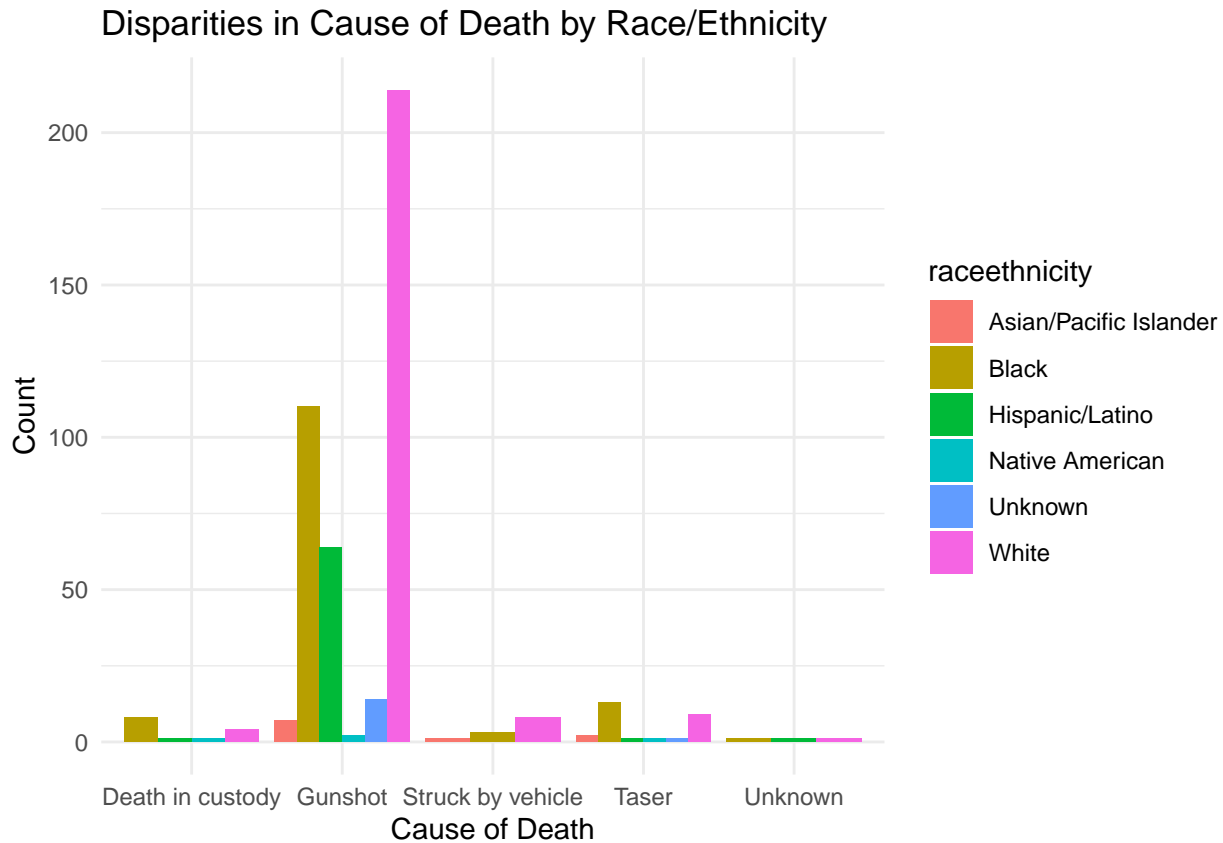


The distribution of causes of death among individuals involved in police killings was analyzed across different racial and ethnic groups. The causes of death in the dataset included Death in Custody, Gunshot, Struck by Vehicle, Taser, and Unknown.

The most important cause of death recorded was that by gunshot, showing a stark divergence by race. At least 200 deaths were assigned to White individuals as being caused by gunshot, followed by Black and Hispanic/Latino individuals. Gunshot deaths occurred most frequently in the Black and White population, accounting for 27% of gunshot deaths and being also 42% in the White population. The stark difference in gunshot fatalities among the races illustrates how this might impact police interactions in regard to different racial/ethnic groups, where the White group exhibits highest casualties.

The second most common cause of death was struck by vehicle, which was particularly prominent among White individuals, accounting for more than 50 deaths. Hispanic/Latino individuals and Black individuals were also impacted by deaths caused by vehicles, but to a lesser extent. This cause of death, though still considerable, was less prevalent than gunshot fatalities across all groups. Native American individuals, while a small proportion of the dataset, showed some representation in this category as well, with Native Americans representing 2.1% of the deaths due to vehicles in Unaffiliated states.

Deaths in custody were less frequent overall, and this category saw some variation across racial groups. Black individuals were represented in 28% of deaths in custody, a significant proportion considering the smaller share of Black individuals in the overall dataset, indicating potential systemic issues related to police conduct in detainment scenarios. White individuals also accounted for a notable number of deaths in custody, comprising 42% of this cause.

Taser-induced fatalities revealed distinct divergence in the racial composition: among those dead, a significant share were Black individuals. That notwithstanding, the White made up the highest share of taser-dependent fatalities, although that was considerably less than gunshot-related ones. The amount of Black individuals in the Taser category against other causes points toward racial disparity and functional target of certain endemic nature within some communities. There too were a number of Unknown causes; an quite sizeable
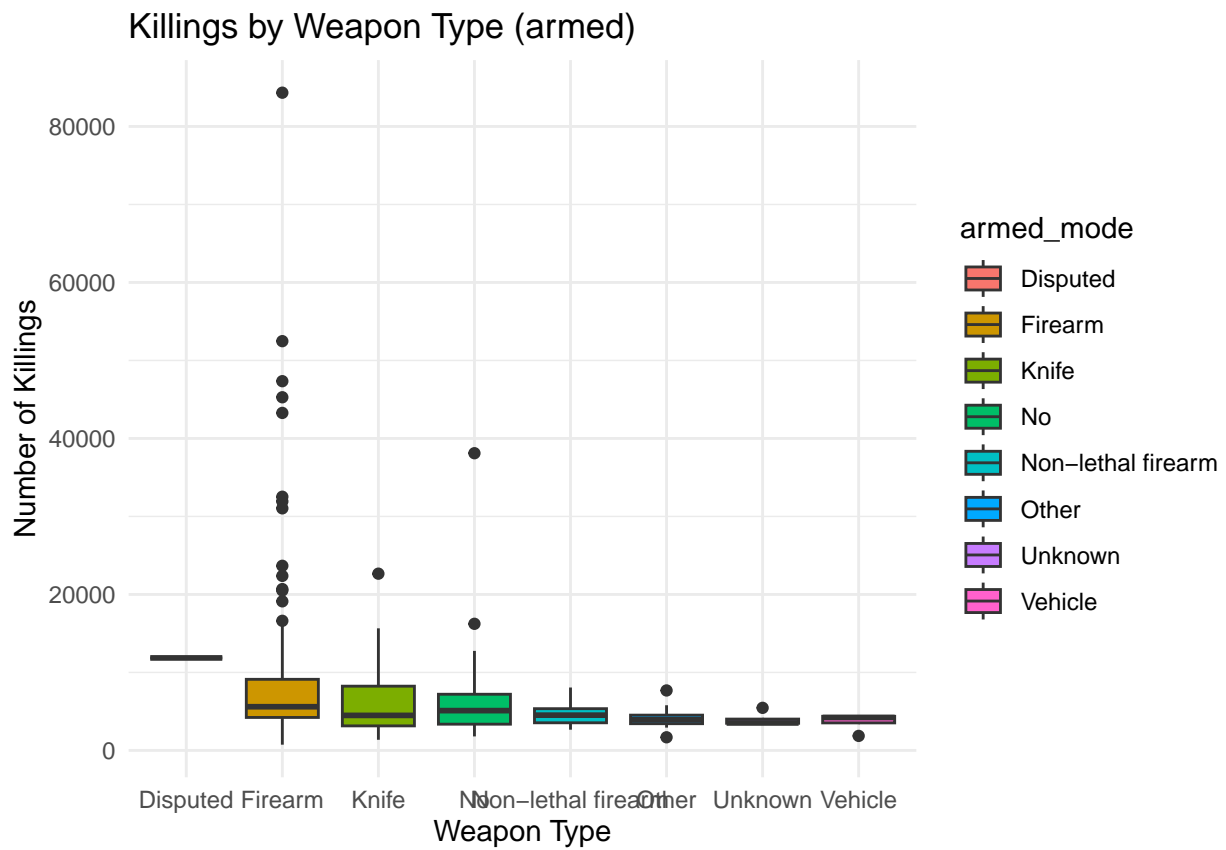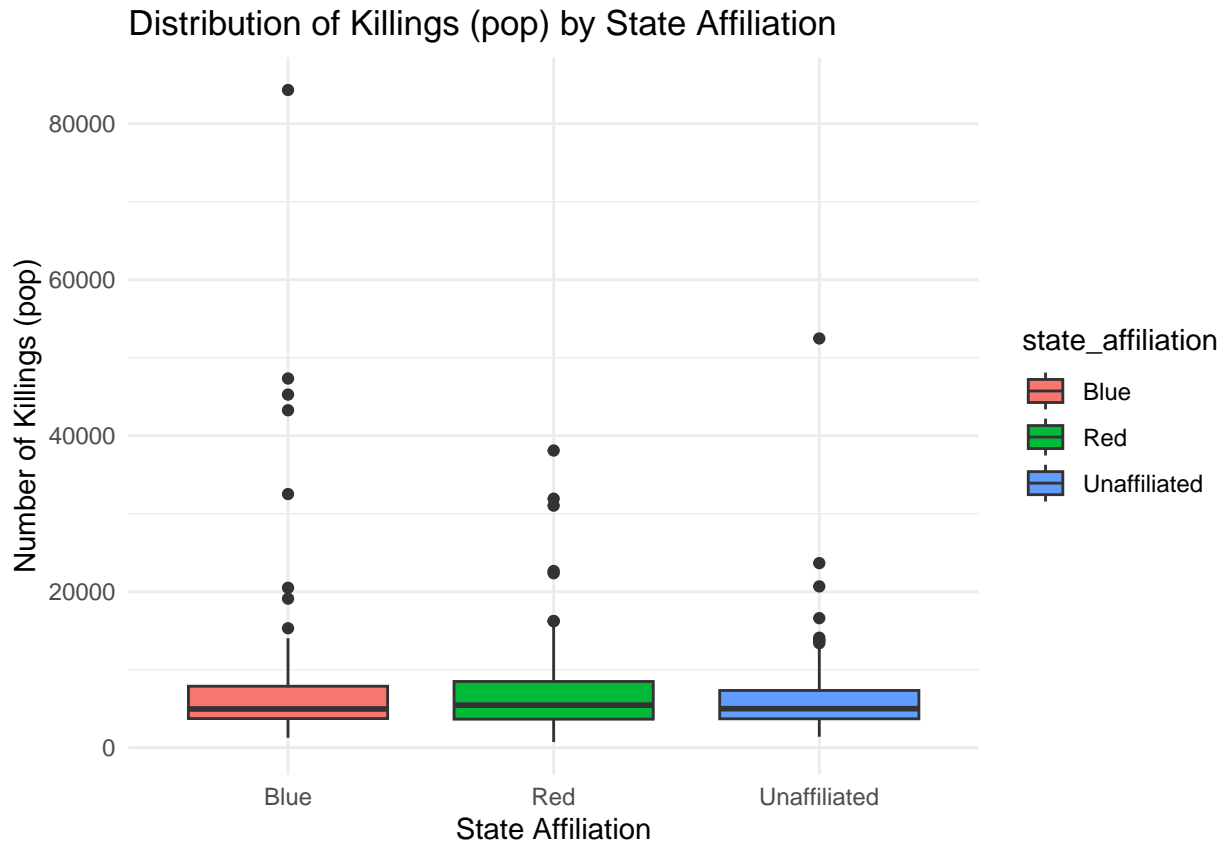
number of them were recorded for every single racial group. In some cases, extremes of ambiguous reasons for the individual's death or inadequate documentation may well be at work, by arming such into this category.
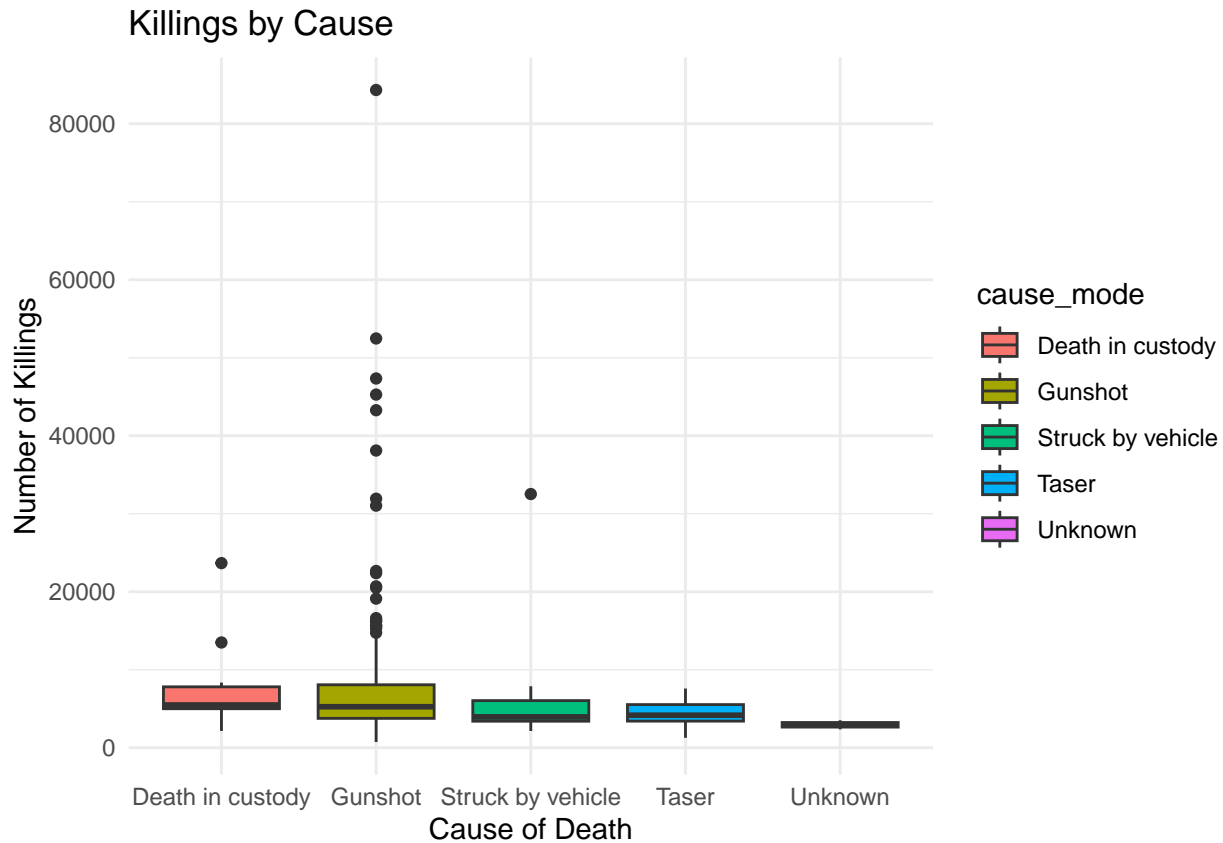
**What are the key socioeconomic, demographic, and political factors associated with the frequency of police killings at the county level?**

```
## `summarise()` has grouped output by 'county_id'. You can override using the
## `.groups` argument.
```

```
## `summarise()` has grouped output by 'county_id'. You can override using the
## `.groups` argument.
```

```
##
## Call:
## glm(formula = killings ~ median_income + poverty_rate + unemployment_rate +
##     college_graduates + state_affiliation + armed_mode + cause_mode,
##     family = poisson, data = county_data)
##
## Coefficients:
##                               Estimate Std. Error  z value Pr(>|z|)
## (Intercept)                   8.617e+00  1.104e-02  780.880  < 2e-16 ***
## median_income                 8.523e-06  6.390e-08  133.371  < 2e-16 ***
## poverty_rate                  8.809e-03  9.956e-05   88.476  < 2e-16 ***
## unemployment_rate            -5.311e-01  1.497e-02  -35.478  < 2e-16 ***
## college_graduates            -6.357e-03  6.989e-03   -0.910 0.363048
## state_affiliationRed         -1.256e-01  1.808e-03  -69.466  < 2e-16 ***
## state_affiliationUnaffiliated -2.364e-01  1.841e-03 -128.385  < 2e-16 ***
## armed_modeFirearm            -3.309e-02  9.310e-03   -3.555 0.000379 ***
## armed_modeKnife              -3.729e-01  9.482e-03  -39.332  < 2e-16 ***
## armed_modeNo                 -3.864e-01  9.436e-03  -40.947  < 2e-16 ***
## armed_modeNon-lethal firearm -7.434e-01  1.032e-02  -72.012  < 2e-16 ***
## armed_modeOther              -6.528e-01  1.055e-02  -61.901  < 2e-16 ***
## armed_modeUnknown            -7.609e-01  1.237e-02  -61.518  < 2e-16 ***
## armed_modeVehicle            -8.956e-01  1.240e-02  -72.215  < 2e-16 ***
## cause_modeGunshot             1.226e-01  3.517e-03   34.847  < 2e-16 ***
## cause_modeStruck by vehicle   8.538e-02  5.436e-03   15.706  < 2e-16 ***
## cause_modeTaser              -4.398e-01  5.057e-03  -86.966  < 2e-16 ***
## cause_modeUnknown            -5.398e-01  1.383e-02  -39.016  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 1503753  on 300  degrees of freedom
## Residual deviance: 1305970  on 283  degrees of freedom
## AIC: 1309159
##
## Number of Fisher Scoring iterations: 5
```

## Distribution of Killings (pop) by State Affiliation



## Killings by Weapon Type (armed)

## Killings by Cause



The Poisson regression model reveals several significant predictors influencing the number of killings. Increased Median Income was positively associated with the number of killings (Estimate = 0.000008523, p < 0.0001). This indicates that when the median income of a county increases, a corresponding increase in the number of killings is also likely, though the effect size is rather small. The finding reinforces the notion that income-related factors could be contributing to killing rates in these areas.

Another significant predictor, Poverty Rate, was positively associated with the number of killings (Estimate = 0.008809, p < 0.0001). In particular, when there are more killings, one noted, higher poverty rates generally lead to higher incidences of murder. The social environment, such as poverty, could well be responsible for the high levels of violence. A significant p-value backs up the strength of the relationship.

Unemployment Rate showed a negative association with killings (Estimate = -0.5311, p < 0.0001), suggesting that counties with higher unemployment rates tend to have fewer killings. At first glance, this may seem odd because, in more impoverished communities with high unemployment, a host of social factors or interventions may combine to decrease killings, or the police forces may be diff, in such counties.

The variable College Graduates did not exhibit a statistically significant relationship with the number of killings (Estimate = -0.006357, p = 0.363). This suggests that the incidence killing is not significantly affected by the percentage of college graduates in a county, at least in the context of controlling for other variables in the model.

Another notable factor influencing police killings was State Affiliation. Red states (Republican-controlled) and Unaffiliated states have fewer killings than Blue states (Democratic-controlled). The coefficient for Red states was negative: they exhibited fewer killings (Estimate = -0.1256, p < 0.0001), along with a reduction in killings among counties in Unaffiliated states (Estimate = -0.2364, p < 0.0001). Such results suggest that political affiliation may influence the rate of police killings.

There were several categories under Armed Mode that had significant statistically negative associations with killings. In particular, compared to other modes, gun-wielding (Estimate = -0.03309, p = 0.000379),

knife-wielding (Estimate = -0.3729, p < 0.0001), and other weapon-bearing (Estimate = -0.8956, p < 0.0001) groups are associated with fewer killings. It could reflect differences in how police respond to or interact in such situations, although such patterns would require further inspection.

A notable effect was also exhibited by the Cause of Death variable. While significantly positive in relation to killings, gunshot deaths explained a higher incidence of killings (estimate = 0.121626, p<.0001). Similar relations proved to work with vehicle deaths (estimate = 0.08538, p < 0.0001), although ones involving tasers provided a negative correlation (estimate = -0.4398, p < 0.0001) and those of unknown cause an even more pronounced effect (estimate = -0.5398, p < 0.0001), suggesting a lesser likelihood of fatal outcomes from these forms of death compared to that inflicted by gunshot wounds.

The model goodness of fit was checked with deviance statistics. It produced a null deviance of 1,503,753 with 300 degrees of freedom and residual deviance of 1,305,970 with 283 degrees of freedom, indicating that the model has good explanatory power for the number of killings. The AIC value of 1,309,159 suggests a close fit of the model to the data.

## Results

Descriptive statistics revealed significant disparities in the causes of death by race/ethnicity. Gunshots emerged as the predominant cause across all racial groups, with notable overrepresentation among Whites and Blacks. Other causes, such as being struck by a vehicle or taser use, were less common but showed disparities across racial groups.

Regression analysis revealed several significant findings. Higher median income was associated with an increased number of police killings, potentially reflecting urbanization or increased interactions in wealthier areas. Similarly, higher poverty rates were positively correlated with killings, consistent with the notion that economically vulnerable communities experience greater police interactions and, consequently, higher risks. Interestingly, unemployment rates were negatively associated with killings, a counterintuitive finding that warrants further investigation into potential confounding variables.

State political affiliation also played a significant role. Compared to Blue states, Red and Unaffiliated states exhibited significantly fewer police killings, suggesting that regional differences in law enforcement practices or population density may influence outcomes. Armed status emerged as a critical factor in predicting killings. Individuals armed with knives, non-lethal firearms, or other objects were less likely to be killed than those in the reference category. Notably, unarmed individuals exhibited the lowest likelihood of being killed, highlighting potential issues with proportionality in police responses. Cause of death was another significant predictor. Gunshot-related deaths were strongly associated with higher killing counts, while deaths involving tasers or unknown causes were less likely. This finding underscores the disproportionate use of lethal force in police encounters.

## Discussion

These findings underscore the complex interplay between socioeconomic, political, and situational factors in police killings. Economic disparities grew to become the main driving force in the experiment, with both high-income and high-poverty areas functioning at an elevated killing rate. This twinning may reflect different mechanisms at play: higher urbanization and police presence in affluent realms versus systemic marginalization and hyper-policing in economically disadvantaged neighborhoods.

State political affiliation also has an influence on the frequency of killings. Blue states had more than Red or Unaffiliated states. This could be attributable to different reporting practices, population density, or law enforcement strategies. Future research should disentangle these factors and further consider how political contexts shape policing practices.

The documentation of armed status and cause of death describes acute insights regarding the proportionality of police responses. The much lower incidence rate of killings for unarmed persons spurs discussion regarding the force needing use in many cases. So does the high volume of gunshot deaths, which suggests a need for alternative de-escalation strategies and less-than-lethal alternatives.

Racial disparities in causes of death, notably the disproportionate numbers of black and white persons in gunshot-related killings, serve to amplify systemic inequities in law enforcement practices. Accordingly, such disparities must be addressed in a multifaceted manner that incorporates training for officers, strengthening the various accountability mechanisms, and introducing community engagement strategies to build trust.

**Conclusion**

This study provides a comprehensive analysis of the factors influencing police killings in the United States and reveal rather stark disparities associated with socioeconomic conditions in tandem with state political affiliation and situational factors. The findings underscore the urgent need for systemic reforms that target the root causes for police violence; recommendations for policy changes include targeted intervention to economically vulnerable communities, enhanced de-escalation training for law enforcement, and action to decrease racial disparity in policing outcomes.

Future research should elaborately investigate qualifying aspects of police interactions to augment all quantitative evidence thereof. Also, investigating a potential overdispersion in count data within an alternative modelling approach would be helpful, such as in the case of a Negative Binomial regression.

# Appendix: All code for this report

```r
knitr::opts_chunk$set(echo = TRUE)
# Load necessary libraries
library(dplyr)
library(ggplot2)

# Load data
police_data <- read.csv("police_killings.csv", encoding = "ISO-8859-1")
data <- read.csv("police_killings.csv", encoding = "ISO-8859-1")

# Convert columns to numeric where necessary
police_data$age <- as.numeric(police_data$age)
police_data$share_white <- as.numeric(police_data$share_white)
police_data$share_black <- as.numeric(police_data$share_black)
police_data$share_hispanic <- as.numeric(police_data$share_hispanic)
police_data$p_income <- as.numeric(police_data$p_income)
police_data$h_income <- as.numeric(police_data$h_income)




# Handle missing values - Remove rows with missing income data
police_data_clean <- police_data[!is.na(police_data$h_income), ]

# Categorize states into Red, Blue, and Unaffiliated
# Define Red, Blue, and Unaffiliated states (simplified example; adjust as needed)
red_states <- c("AL", "TX", "FL", "MS", "TN", "SC", "LA", "OK", "AR", "KY", "WV", "ND", "SD", "WY", "MT"
blue_states <- c("CA", "NY", "WA", "MA", "IL", "NJ", "OR", "MD", "CT", "HI", "VT", "DE", "RI")
unaffiliated_states <- setdiff(unique(police_data_clean$state), c(red_states, blue_states))

# Create a new column for state affiliation
police_data_clean$state_affiliation <- ifelse(police_data_clean$state %in% red_states, "Red",
                                        ifelse(police_data_clean$state %in% blue_states, "Blue", "Unaffili

# Ensure the new column is a factor
```

```r
police_data_clean$state_affiliation <- factor(police_data_clean$state_affiliation)
library(gtsummary)

# Example: Summarizing police killings dataset with key variables (age, gender, raceethnicity, income)
table2 <- police_data_clean |>
  tbl_summary(
    include = c(age,gender,  raceethnicity, p_income, h_income, county_income),  # Variables to summari
    by = state_affiliation,  # Split the table by state
    missing = "no"  # Do not list missing data separately
  ) |>
  add_n() |>  # Add column with total number of non-missing observations
  add_p() |>  # Test for a difference between groups (state in this case)
  modify_header(label = "**Variable**") |>  # Update column header
  bold_labels()  # Bold labels for clarity

# View the summary table
table2



# Count causes of death by race/ethnicity
cause_race <- data %>%
  group_by(raceethnicity, cause) %>%
  summarise(count = n()) %>%
  ungroup()

# Visualize
ggplot(cause_race, aes(x = cause, y = count, fill = raceethnicity)) +
  geom_bar(stat = "identity", position = "dodge") +
  labs(title = "Disparities in Cause of Death by Race/Ethnicity",
       x = "Cause of Death",
       y = "Count") +
  theme_minimal()

# Aggregate data to county level
county_data <- police_data_clean %>%
  group_by(county_id, state_affiliation) %>%
  summarise(
    killings = n(),
    median_income = mean(h_income, na.rm = TRUE),
    poverty_rate = mean(as.numeric(pov), na.rm = TRUE),
    unemployment_rate = mean(urate, na.rm = TRUE),
    college_graduates = mean(college, na.rm = TRUE)
  )

# Check for missing values in the aggregated data
county_data <- na.omit(county_data)


# Convert 'armed' and 'cause' to factors
police_data_clean$armed <- as.factor(police_data_clean$armed)
police_data_clean$cause <- as.factor(police_data_clean$cause)
```

```r
# Aggregate data to county level
county_data <- police_data_clean %>%
  group_by(county_id, state_affiliation) %>%
  summarise(
    killings = sum(pop),  # Use 'pop' to represent killings
    median_income = mean(h_income, na.rm = TRUE),
    poverty_rate = mean(as.numeric(pov), na.rm = TRUE),
    unemployment_rate = mean(urate, na.rm = TRUE),
    college_graduates = mean(college, na.rm = TRUE),
    armed_mode = names(sort(table(armed), decreasing = TRUE)[1]),  # Most frequent value
    cause_mode = names(sort(table(cause), decreasing = TRUE)[1])   # Most frequent value
  )

# Handle missing values
county_data <- na.omit(county_data)

# Convert 'armed_mode' and 'cause_mode' to factors
county_data$armed_mode <- as.factor(county_data$armed_mode)
county_data$cause_mode <- as.factor(county_data$cause_mode)

# Fit Poisson regression model
poisson_model <- glm(killings ~ median_income + poverty_rate + unemployment_rate +
                       college_graduates + state_affiliation +
                       armed_mode + cause_mode,
                 data = county_data,
                 family = poisson)

# Display regression summary
summary(poisson_model)



# Visualize the effect of state affiliation on killings
ggplot(county_data, aes(x = state_affiliation, y = killings, fill = state_affiliation)) +
  geom_boxplot() +
  labs(title = "Distribution of Killings (pop) by State Affiliation",
       x = "State Affiliation",
       y = "Number of Killings (pop)") +
  theme_minimal()

# Visualize the effect of 'armed' and 'cause' on killings
ggplot(county_data, aes(x = armed_mode, y = killings, fill = armed_mode)) +
  geom_boxplot() +
  labs(title = "Killings by Weapon Type (armed)",
       x = "Weapon Type",
       y = "Number of Killings") +
  theme_minimal()

ggplot(county_data, aes(x = cause_mode, y = killings, fill = cause_mode)) +
  geom_boxplot() +
  labs(title = "Killings by Cause",
       x = "Cause of Death",
       y = "Number of Killings") +
```

```
  theme_minimal()
```

```
theme_minimal()
```