Analysis of Law Enforcement Transparency in Fatal Police Shootings in the United States from 2015 to 2024*

My subtitle if needed

Zhixi(Arno) Qu

November 27, 2025

First sentence. Second sentence. Third sentence. Fourth sentence.

1 Introduction

In recent years, fatal shootings by U.S. police officers during law enforcement operations have consistently drawn public attention and have been widely debated. The large number of such cases has prompted society and the public to focus not only on the situations under which police may use fatal force, but also on the extent to which these cases have been recorded and made public. Under this background, body cameras have been used as a key tool for police departments to maintain public trust and provide professional, fair policing services. In social surveys, both citizens and police officers have recognized the importance of body cameras in law enforcement (2020). However, Equipping police officers with body cameras does not guarantee that all police incidents will be recorded and made public, especially in extreme cases such as fatal police shootings. In actual policing, whether officers activate their body cameras is often influenced by individual factors and situations. While body cameras may be more commonly used during certain types of enforcement activities, their usage rates are notably lower in specific scenarios or when interacting with particular groups. This may lead to unequal levels of policing transparency and accountability for police officers or victims across different populations and situations.

Current reports and studies on body-worn cameras are limited to the ability of wearing them to enhance law enforcement transparency (2024). However, what truly concerns society is whether these cameras are genuinely and fairly activated to improve police transparency and

*Code and data are available at: https://github.com/ZhixiQu/US-Fatal-Police-Shootings.git public trust. To address this gap, this paper analyzes changes in U.S. police fatal shooting rates using U.S. police fatal shooting data and US Census Data. Most importantly, uing the police fatal shooting data to analyse whether different policing scenarios and demographic groups systematically influence body camera activated status during fatal police shootings. Specifically, I utilize U.S. Police Fatal Shooting Database From The Washington Post (2025) combined with two Census data from U.S. Census Bureau during 2010 and 2025 (2021) (2025). First, I provide an overview of changes in U.S. police fatal shooting rates from 2015 to 2024. Besides, I will treat the probability of body camera was activated as the dependent variable in police-involved fatal incidents and fit a Bayesian Logistic Regression model. This model will include victim race, gender, age grouping variables established based on the victim's age, whether the victim was armed, whether the victim fled, whether the victim had mental health issues, and the year as covariates. Furthermore, an interaction term between victim race and whether the victim was armed will be included to examine systemic associations between different situations and demographic characteristics and body camera activated status. Overall, this study quantifies police transparency in fatal shootings to complement the debates on policing transparency and accountability.

Empirical findings indicate that both the rate of fatal police shootings per million people and the activation rate of body-worn cameras during such incidents showed an overall upward trend from 2015 to 2024. the most important finding is that the activation of body-worn cameras exhibits systemic variation across specific demographic groups. When victims have mental health conditions, body-worn cameras are more likely to be activated. Furthermore, older individuals are less likely to be recorded by body-worn cameras during incidents compared to those who are minors. Incomplete data such as when the victim's race is unknown which correlates with lower body camera activation probality. The interaction between victim race and whether the victim was armed was not statistically significant in this study, it indicates there is no stable and significant systemic differences in body camera usage across racial groups in armed scenarios within currently available public fatality data. Overall, Overall, this study found uneven increases in law enforcement transparency over time and their distribution across specific groups of victims. However, due to limitations of the data, all findings are presented as descriptive correlations rather than causal inferences.

The structure of the paper is divided into four main sections. Section 2 discusses the selected data, data pre-processing, and key variables for analysis. Section 3 discusses the model comparison process that ultimately determined the final model for this study's analysis. Section 4 presents the model results, trends observed in the data, and variables associated with whether body cameras were activated. Section 5 draws conclusions from the models and relates them to real-world scenarios to evaluate the reliability of these conclusions. Additionally, it describes the limitations and weaknesses of this study and suggests directions for future improvements.

2 Data

2.1 Overview

In this analysis, I merged the three datasets into two separate datasets. The preliminary study (1) Change in the rate of fatal force caused by U.S. police per million population. Combined police fatality incidents with U.S. Census data to analyze changes in the rate of fatal police shootings during 2015 and 2024. The main focus (2) Analyzed differences in body camera activation probabilty across different groups of victims and scenarios in U.S. fatal police shootings. All data cleaning and analysis were performed in the statistical programming language R (R Core Team, 2023), primarily utilizing packages such as tidyverse (Wickham et al., 2019), lubridate (2024), readxl (2025), and forcats (Wickham, 2023) for data organization and visualization.

The U.S. police fatality data for this study originates from The Washington Post's "data police shooting" database. Since 2015, researchers at The Washington Post have continuously documented fatal shootings by researchers at The Post. The database has two versions. Version 2 data was migrated from Version 1 in 2022 and is continuously updated, with the latest update occurring in May 2025. After this point, Version 1 data will no longer be updated, so I decided to use Version 2 data. The other two data sources are state population total data from the U.S. Census Bureau: NST-EST2020 (covering 2010–2020) and NST-EST2024 (covering 2020–2024), which provide annual population estimates for each state as of July 1st. After filtering, these population data were used to calculate the fatal police shootings from 2015 to 2024. Similar datasets exist that provide nationwide statistics on police use-of-force incidents through collaboration between the FBI and law enforcement agencies. These include police-involved fatalities (FBI 2024), but crucially, data reporting by local agencies is voluntary, leading to volunteer bias. Besides, the absence of body camera data made this dataset incompatible with my primary research focus, so I finally did not use it.

2.2 US police fatality rate and body camera activated rate

For two population census datasets from the U.S. Census Bureau, I extracted annual population estimates for each state as of July 1st from 2015 to 2019 in NST-EST2020, because the 2020 census data in this dataset is less accurate compared to the NST-EST2024 dataset. From the NST-EST2024 dataset, I extracted annual population estimates for each state as of July 1st from 2020 to 2024. Additionally, I mapped the state names in both files to their two letter state abbreviations. The final dataset is named "state_year_pop" and contains annual population estimates for 50 states and the District of Columbia from 2015 to 2024. It includes three new variables: state, year, and pop.

Each row in the original U.S. police fatality dataset organized by The Washington Post represents a victim shot by police in a fatal shooting case. It includes the date and location of the incident, along with several characteristics of the victim and the situation. First, dates were converted to Gregorian calendar years, and to limit the analysis within the 2015–2024 period. To avoid duplicate counting, only unique incident IDs were kept. Death counts were calculated based on state*year. Finally, the fatal case data and annual population data were merged to create state_year_rates, and the police fatality rate per million people was calculated. This dataset includes the following variables: - state: Two letter state abbreviation. - year: Gregorian calendar year. (2015–2024) - death: Number of fatal police shootings in that state during that year. - pop: Total population estimation of that state during that year. - rate_per_million: Police fatality rate per million residents.

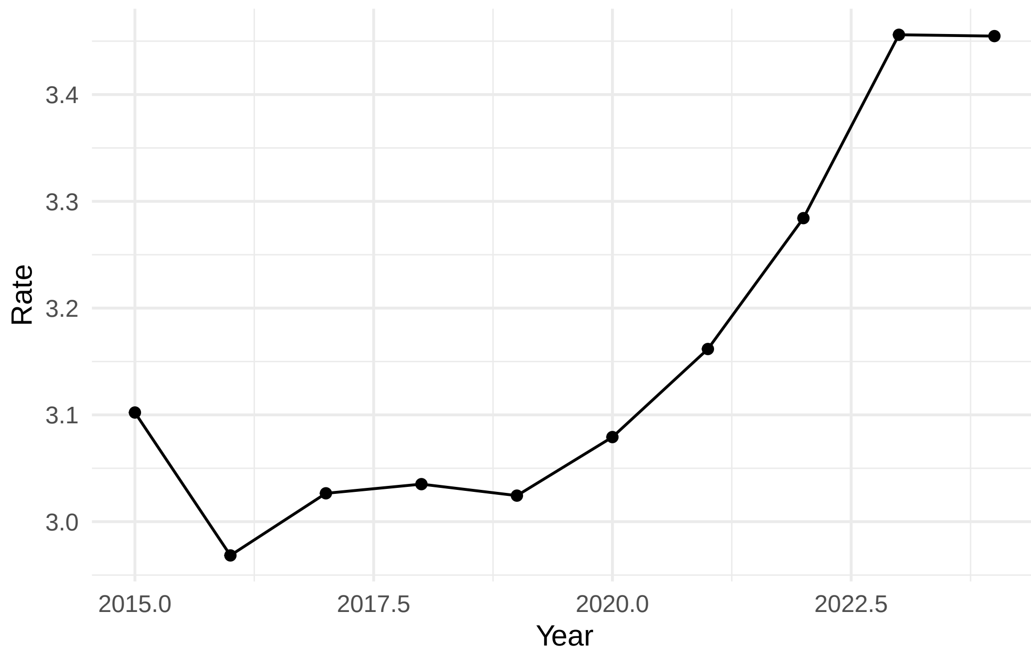

Figure 1: U.S. Fatal police shooting rate over time per million (rate)

Figure 1 shows the time trend of fatal police shootings rate across the United States from 2015 to 2024. It is evident that the rate of fatal police shootings in the U.S. is generally increasing. Although there was a short period of decline between 2015 and 2020, the decrease was not significant and did not appear long. Notably, from 2020 to May 2022, the rate of police fatal shootings experienced rapid and continuous growth and no clear decrease observed after late 2022. This indicates that in high-risk cases the risk of people when interacting with police has increased in recent years and reflects why society continues to focus on fatal police shootings.

2.3 Law enforcement transparency analysis data

The second analytical dataset is also constructed based on the above U.S. police fatality data, used to model whether body cameras were activated during fatal police shooting. When handling missing values, I only deleted entire rows of observations for categorical variables or critical variables. For other variables, I retained the sample by labeling missing values as "Unknown" whenever possible. Specifically, if the missing value pertained to whether the body camera was turned on (the dependent variable) or age (used to construct the crucial covariate of age groups), arbitrary definitions could undermine the analysis's credibility, so such observations were excluded from the sample. However, for variables such as race, gender, weapon information, and mental health status, missing data were coded as "Unknown." These observations were retained in charts and models, significantly reducing data loss while ensuring the accuracy and credibility of analytical results. The following variables were constructed:

The second analytical dataset is also constructed based on the above U.S. police fatality data, used to model whether body cameras were activated during fatal police incidents. When processing missing values, I only deleted entire rows of observations for categorical variables or key variables. For other variables, I kept the sample by using "Unknown" where possible. Specifically, if the missing value was whether the body camera was activated (the dependent variable) or age (used to construct the important covariate of age group), arbitrary definitions could undermine the credibility of the analysis, so these observations were excluded from the sample. However, for variables such as race, gender, weapon information, and mental health status, missing data were coded as "Unknown." These observations were kept in charts and models, significantly reducing data loss while ensuring the accuracy and credibility of analytical results. The following variables were constructed: - body_camera_binary: Assigns 1 when the body camera is activated, and 0 when body camera was off. Since missing data for the dependent variable would cause model fitting to fail, I chose to delete the missing value. - race: The race of victim. In the original data, some records had multiple race categories separated by semicolons. To simplify analysis, I kept only the first race category, treating it as the primary category for that individual. Missing values were defined as "Unknown." I preserved the original data values such as "W" means the race of victim is white. - gender: Gender of victim. The value of original data are male, female, or non-binary. Missing values were defined as "Unknown". - age_group: Victim's age group at the time of case. First, raw ages were converted to numeric values, then grouped into five intervals: under 18, 18–29, 30–44, 45–59, and 60+. Observations with missing age (unable to be grouped) were excluded from the analysis sample. - whether_mental_ill: Assigns "yes" if the victim had a history of mental health issues or was experiencing mental distress at the time of the shooting. Assigns 'no' if the original variable indicates no mental illness association. If the value is missing, it is defined as "Unknown" to keep these observations and mark the information as missing. - armed: Whether the victim had a high-threat weapon. If the victim was unarmed during the incident, there was no evidence confirming weapon possession or the victim held only a non-functional firearm, this variable is assigned as "no". If the victim's weapon status is missing, it is assigned as "Unknown." In all other cases, it is assigned as "yes". - flee: Whether the victim fled. If the victim did not flee assigned “no”. If any form of flight occurred, value is yes. Since there is no missing value so do not need further step. In addition to the above constructor variables, I have kept the following identification and context variables: - id: A unique identifier for each fatal police shooting incident. - state: The two-letter abbreviation of the state where the incident occurred. - year: The year the incident occurred.

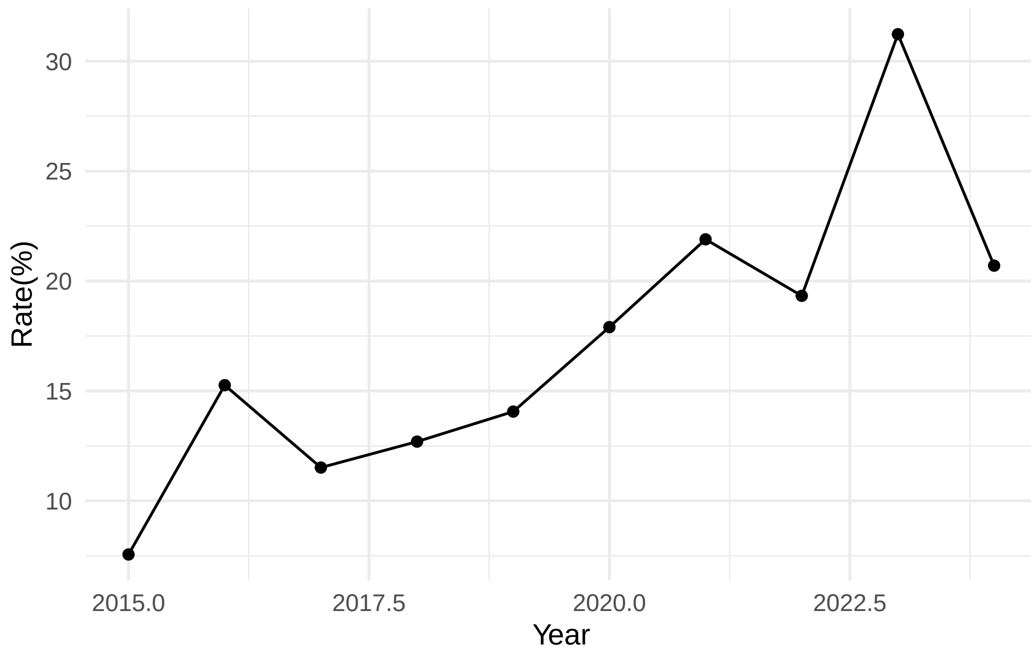

Figure 2: Active Rate of Body Cameras in Fatal Police Shooting(US 2015–2024)

Figure 2 shows the proportion of body cameras that were activated during fatal police shootings over the same period. Overall, the activation rate of body cameras has fluctuated but shown a significant increase. Moreover, the peak activation rate over the past decade has been approximately 32%. This indicates that the use of body cameras in U.S. police fatal shootings is increasing but remains unstable, potentially due to insufficient officer training or the environment of the police operation. Furthermore, the fact that the proportion of body camera activated in fatal shootings remains below 50% suggests significant debate over police transparency in such cases over the past decade. Notably, the increase in proportion of body camera activated between 2015 and 2017 and the decrease in fatal shootings rate indicate a potential negative correlation between body camera deployment and fatal shooting rates. Therefore, researching factors influencing body camera usage provides empirical evidence to enhance public trust in police and develop more effective regulatory and training strategies.

2.4 Predictor variables

Add graphs, tables and text.

Use sub-sub-headings for each outcome variable and feel free to combine a few into one if they go together naturally.

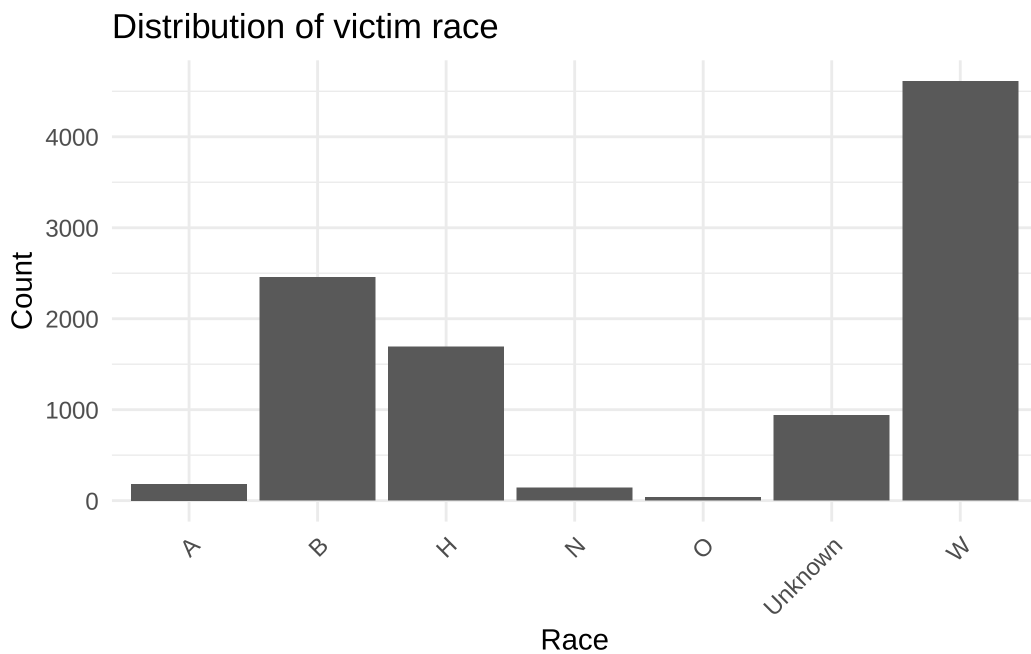

Figure 3: Active Rate of Body Cameras in Fatal Police Shooting(US 2015–2024)

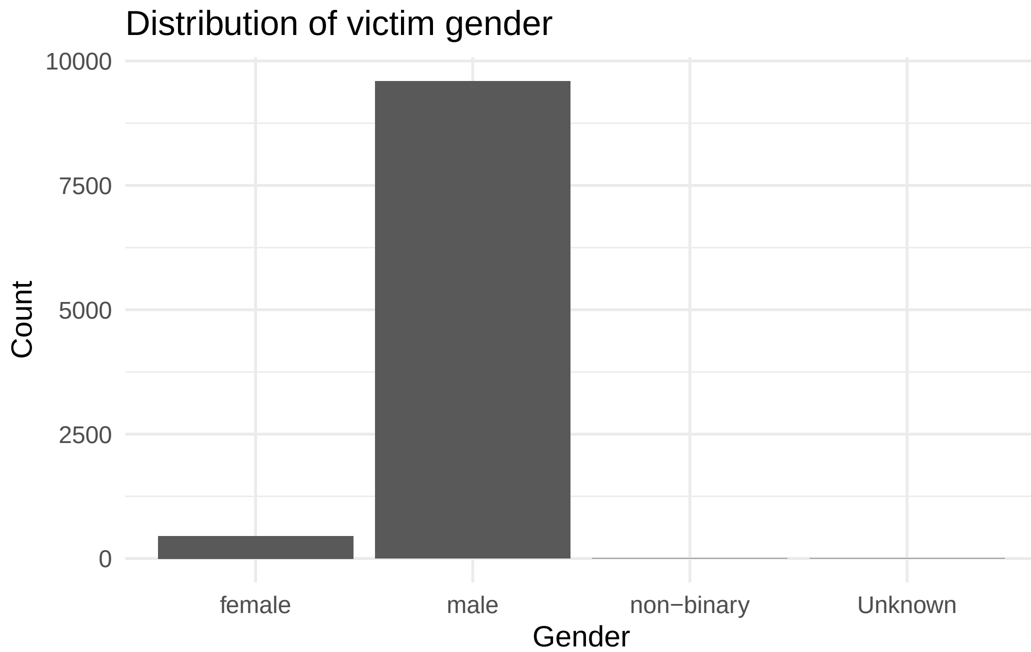

Figure 4: Active Rate of Body Cameras in Fatal Police Shooting(US 2015–2024)

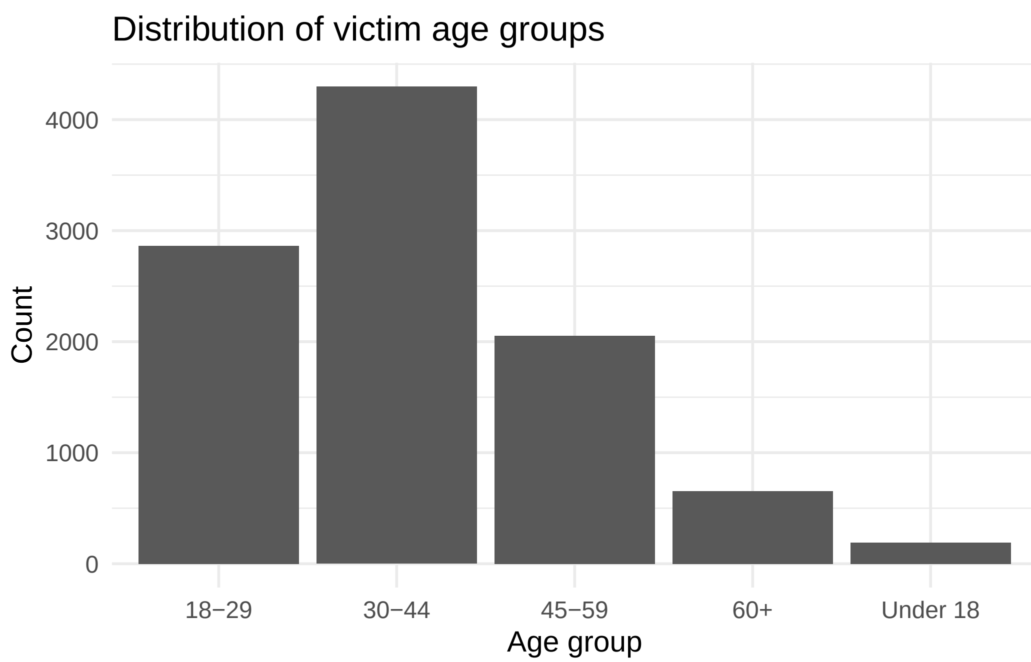

Figure 5: Active Rate of Body Cameras in Fatal Police Shooting(US 2015–2024)

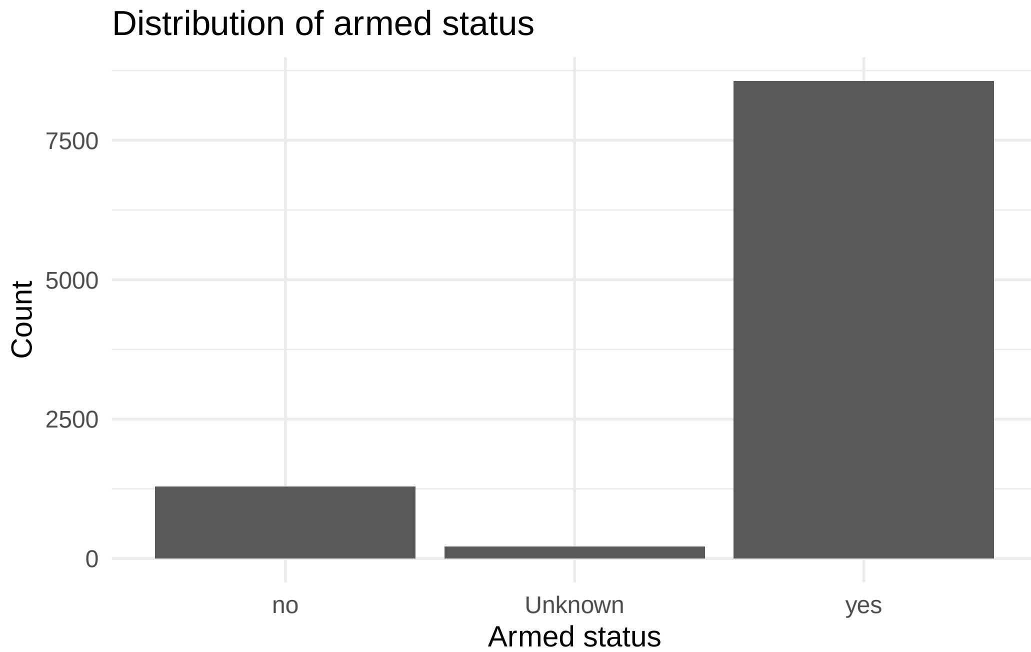

Figure 6: Active Rate of Body Cameras in Fatal Police Shooting(US 2015–2024)

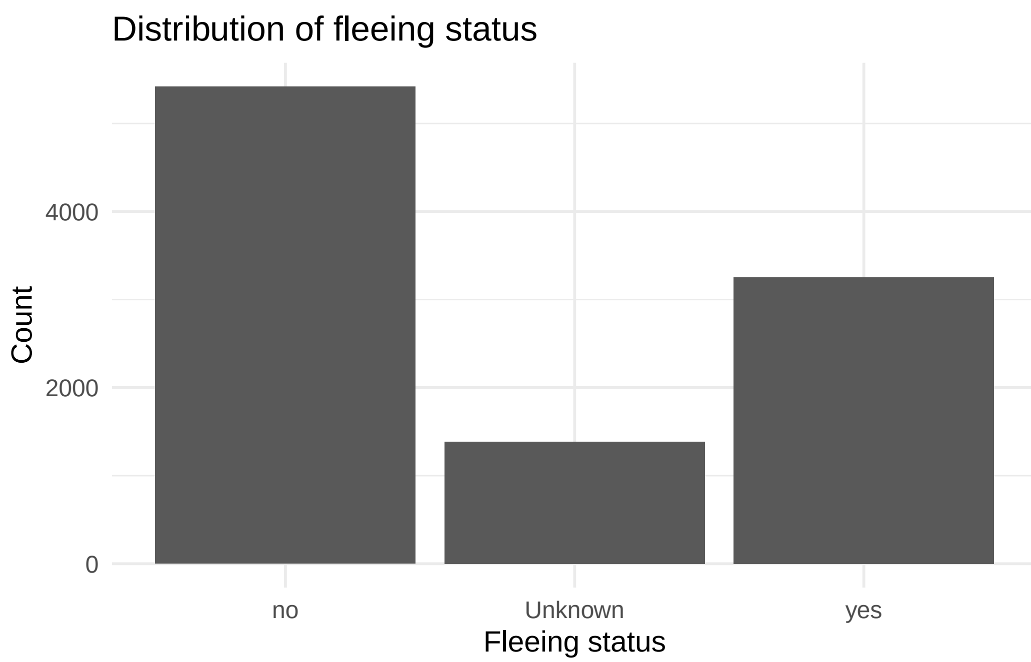

Figure 7: Active Rate of Body Cameras in Fatal Police Shooting(US 2015–2024)

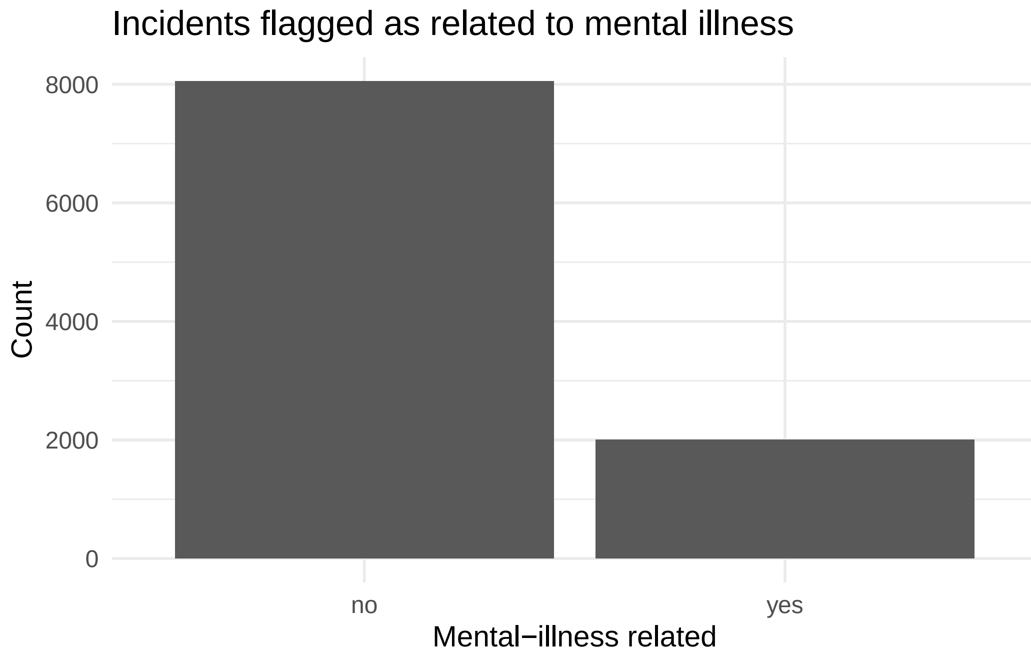

Figure 8: Active Rate of Body Cameras in Fatal Police Shooting(US 2015–2024)

3 Model

The goal of our modelling strategy is twofold. Firstly,…

Here we briefly describe the Bayesian analysis model used to investigate… Background details and diagnostics are included in **?@sec-model-details**.

3.1 Model set-up

Define y_i as the number of seconds that the plane remained aloft. Then β_i is the wing width and γ_i is the wing length, both measured in millimeters.

$$y_i \mid p_i \sim \text{Bernoulli}(p_i)$$

$$\begin{aligned}\log\left(\frac{\hat{p}_i}{1-\hat{p}_i}\right) = \beta_0 + \beta_1 \,\text{race}_i + \beta_2 \,\text{gender}_i + \beta_3 \,\text{age}_i + \beta_4 \,\text{armed}_i + \beta_5 \,\text{flee}_i \\ + \beta_6 \,\text{mental\_health}_i + \beta_7 \,\text{year}_i + \beta_8 \,\text{race}_i \times \text{armed}_i\end{aligned}$$

$$\begin{aligned}
\beta_0 &\sim \text{Normal}(0, 2.5) \\
\beta_1 &\sim \text{Normal}(0, 2.5) \\
\beta_2 &\sim \text{Normal}(0, 2.5) \\
\beta_3 &\sim \text{Normal}(0, 2.5) \\
\beta_4 &\sim \text{Normal}(0, 2.5) \\
\beta_5 &\sim \text{Normal}(0, 2.5) \\
\beta_6 &\sim \text{Normal}(0, 2.5) \\
\beta_7 &\sim \text{Normal}(0, 2.5) \\
\beta_8 &\sim \text{Normal}(0, 2.5)
\end{aligned}$$

where,

- $\hat{p}$ represents the probability that the body camera were activated.
- β_0 represents the intercept term of this logistical regression. It is the probability that body cameras being activated when an individual belongs to the reference group for each categorical variable and the year is 2015.
- β_1 is the coefficient corresponding to race of the victim.
- β_2 is the coefficient corresponding to gender of the victim.
- β_3 is the coefficients corresponding to age group of the person, Age group is categorical variable grouped by the victim's age. The categories are: Under 18, "18–29" "30–44" "45–59" "60+".
- β_4 is the coefficients corresponding to whether the victim is armed.
- β_5 is the coefficients corresponding to whether the victim is fleeing.
- β_6 is the coefficients corresponding to whether the victim had a history of mental health issues.
- β_7 is the coefficients corresponding to the year the event occurred.
- β_8 is the coefficients corresponding to the interaction term between the race of victim and whether the victim were armed.

We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022). We use the default priors from `rstanarm`.

3.1.1 Model justification

We expect a positive relationship between the size of the wings and time spent aloft. In particular…

We can use maths by including latex between dollar signs, for instance θ.

4 Results

Our results are summarized in **?@tbl-modelresults**.

5 Discussion

5.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

5.2 Second discussion point

Please don't use these as sub-heading labels - change them to be what your point actually is.

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Appendix

A Additional data details

B References

Body-Worn Cameras-Toronto Police Service.(2020). https://www.tps.ca/use-technology/body-worn-cameras/ Sponsored content: Body-Worn cameras build transparency and trust for law enforcement across the nation. (2024). POLITICO. https://www.politico.com/sponsored/2024/06/body-worn-cameras-build-transparency-and-trust-for-law-enforcement-across-the-nation/ washingtonpost/data-police-shootings: The Washington Post is compiling a database of every fatal shooting in the United States by a police officer in the line of duty since 2015. (2025). https://github.com/washingtonpost/data-police-shootings US Census Bureau. (2021, October 8). State population totals: 2010-2020. Census.gov. https://www.census.gov/programs-surveys/popest/technical-documentation/research/evaluation-estimates/2020-evaluation-estimates/2010s-state-total.html US Census Bureau. (2025, May 28). State Population Totals and Components of Change: 2020-2024. Census.gov. https://www.census.gov/data/tables/time-series/demo/popest/2020s-state-total.html Use-of-Force. (2024, June 10). Federal Bureau of Investigation. https://www.fbi.gov/how-we-can-help-you/more-fbi-services-and-information/ucr/use-of-force

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "rstanarm: Bayesian applied regression modeling via Stan." https://mc-stan.org/rstanarm/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.