Police Shootings and Racism in America

Connie Wu

10/12/2020

I. Introduction

On May 25, 2020, George Floyd was publicly suffocated to death by police officers in Minneapolis, Minnesota. This, along with the murder of Breonna Taylor by Louisville police, incited waves of protests against police brutality across the United States and increased the spread of Black Lives Matter content on social media. The recent boost in attention to the Black Lives Matter movement has once again brought to light the issue of racism in America and its link to police brutality, and more specifically, police use of deadly force. Several studies have already found that race does play a part in who is targeted and killed in police shootings. For example, in a recent VICE News investigation about police shootings, it was found that "Black people were shot more often and at higher rates than people of any other race." [1] Additionally, Edwards et. al performed a study regarding the effects of age, race-ethnicity, and sex on the risk of being killed by lethal force by law enforcement and similarly found that "Black men are about 2.5 times more likely to be killed by police over the life course than are white men" while "Black women are about 1.4 times more likely to be killed by police than are white women." [2] However, even with these studies and their disturbing conclusions, a poll done by AP-NORC in June 2020 found that still 39% of Americans think that police are not more likely to use lethal force against a black person than a white person. [3] Although this has decreased from an overwhelming 51% in June 2015, there is still a great deal of research that needs to be done in this area to provide even more statistical evidence backing the relationship between racism and police use of deadly force.

As a result, I have decided to build off of VICE News' study and investigate data on police shootings further to understand the roles that race of both victim and officer, as well as other factors such as their genders and whether they are carrying a weapon, play in fatal versus non-fatal police shootings. Additionally, to add predictors relating to location, I will be using data found on Kaggle detailing the gun provisions that are upheld by each state see how gun legislation affects lethal vs. non-lethal shootings. I will also add a predictor indicating whether the state in which the homicide occurred requires de-escalation training for police officers. Finally, RACE DEMOGRAPHICS SOMEHOW. This will allow me to better understand how racism has manifested itself in America's police system and determine whether current attempts to prevent police use of lethal force are effective or not.

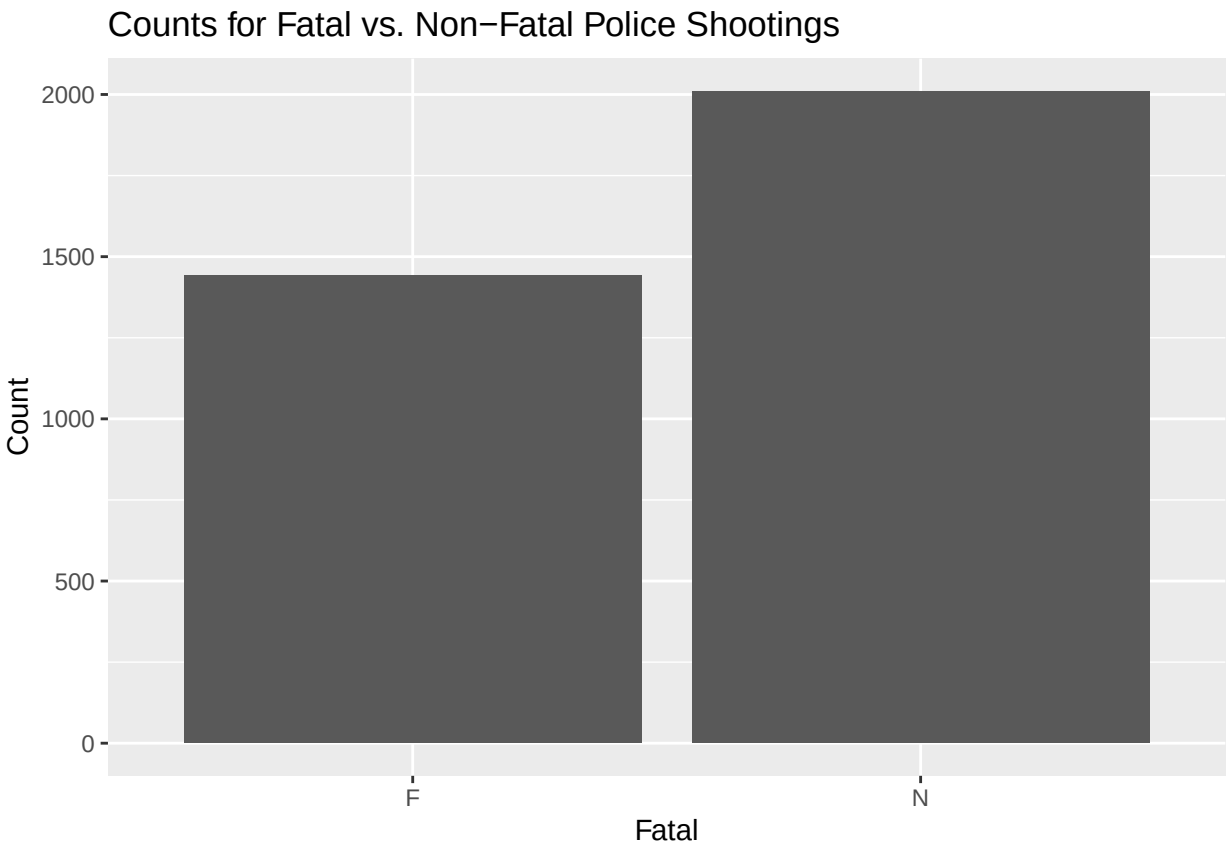
Data

To effectively build off of the study done by VICE News, I will be using the same dataset that they used. This dataset contains data on officer-involved shootings from 47 of the largest local police departments in America, and more specifically, "information on 4,117 incidents and 4,400 subjects over seven years." [1] The predictors that I will use are subject (victim) race, subject gender, officer race, officer gender, and the type of weapon the subject was carrying. The response I will be modeling will be whether the shooting was fatal or not. I will be filtering out the cases in which fatality is unknown or the subject's or the officer's race or gender is unknown from the final dataset. Additionally, there are 265 rows in which there were more than one victims of the shooting; in these cases, I will be separating each of these observations into multiple rows, with each row representing each individual victim and their associated information. I will carry out the same procedure in the cases in which there were more than one officers present at the shooting. Because

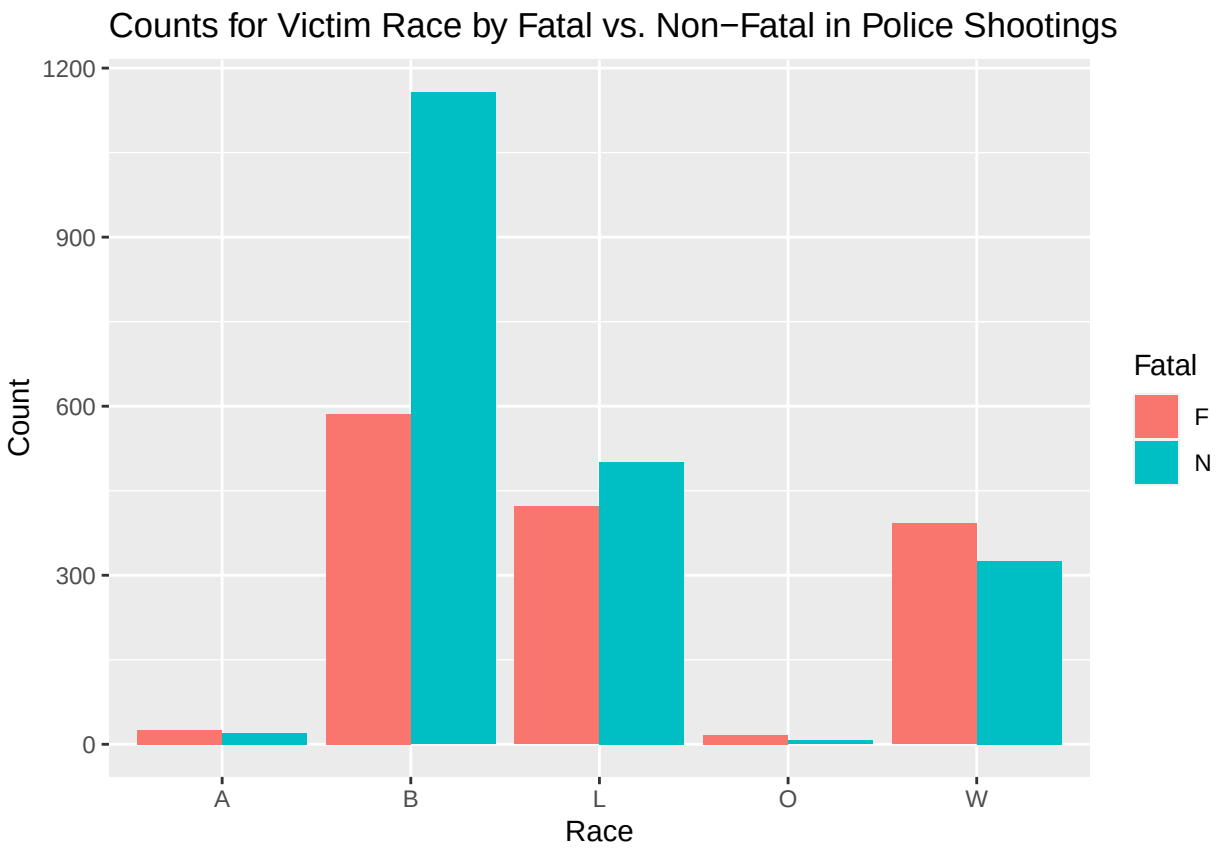
there can be multiple victims and officers recorded for a shooting event, I will also be filtering out cases in which for a shooting event, the number of observations listed for victim or officer race differs from the number of observations listed for victim or officer gender, since I can not infer the missing values in either of these columns.

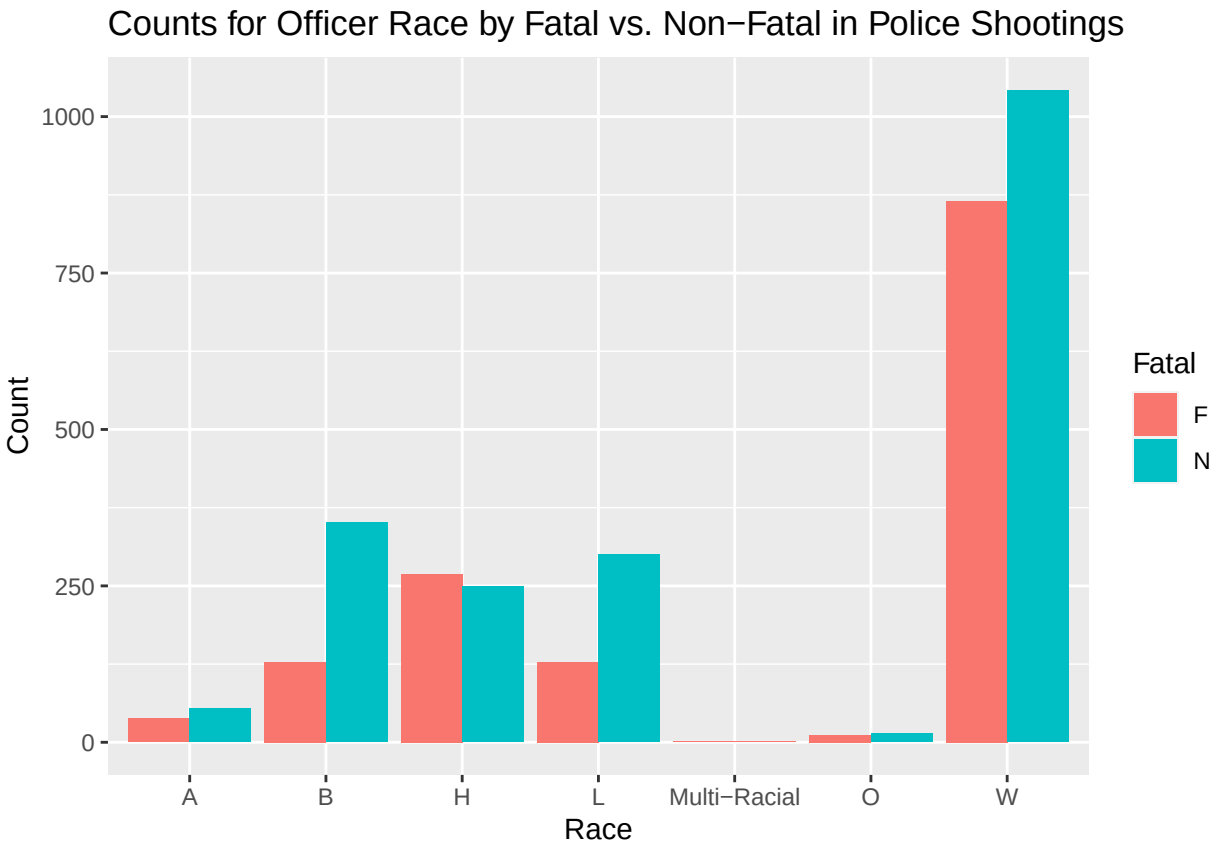
After cleaning the data, I was left with a total of 3,451 observations. I plotted some visualizations that depict the overall trends of the different predictors that I plan to use to answer my proposed research questions.

Exploratory Data Analysis

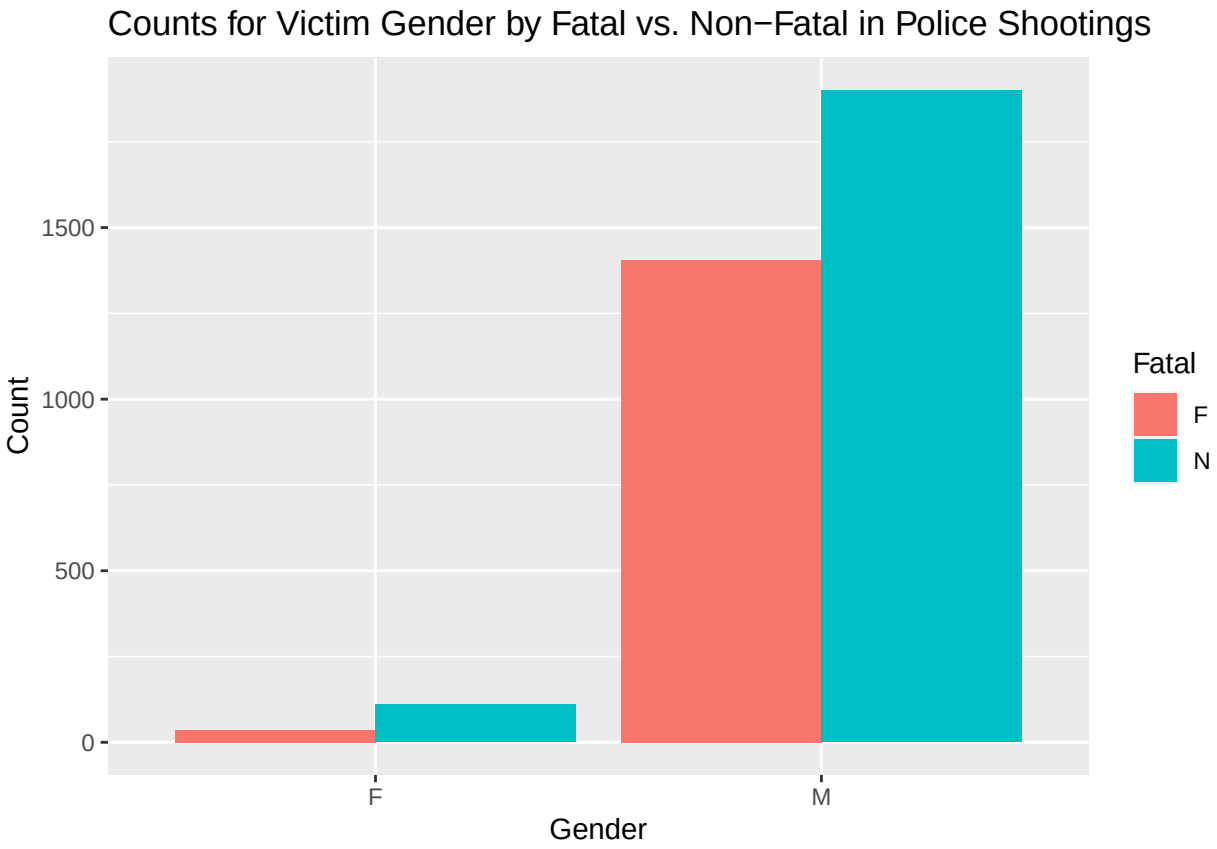


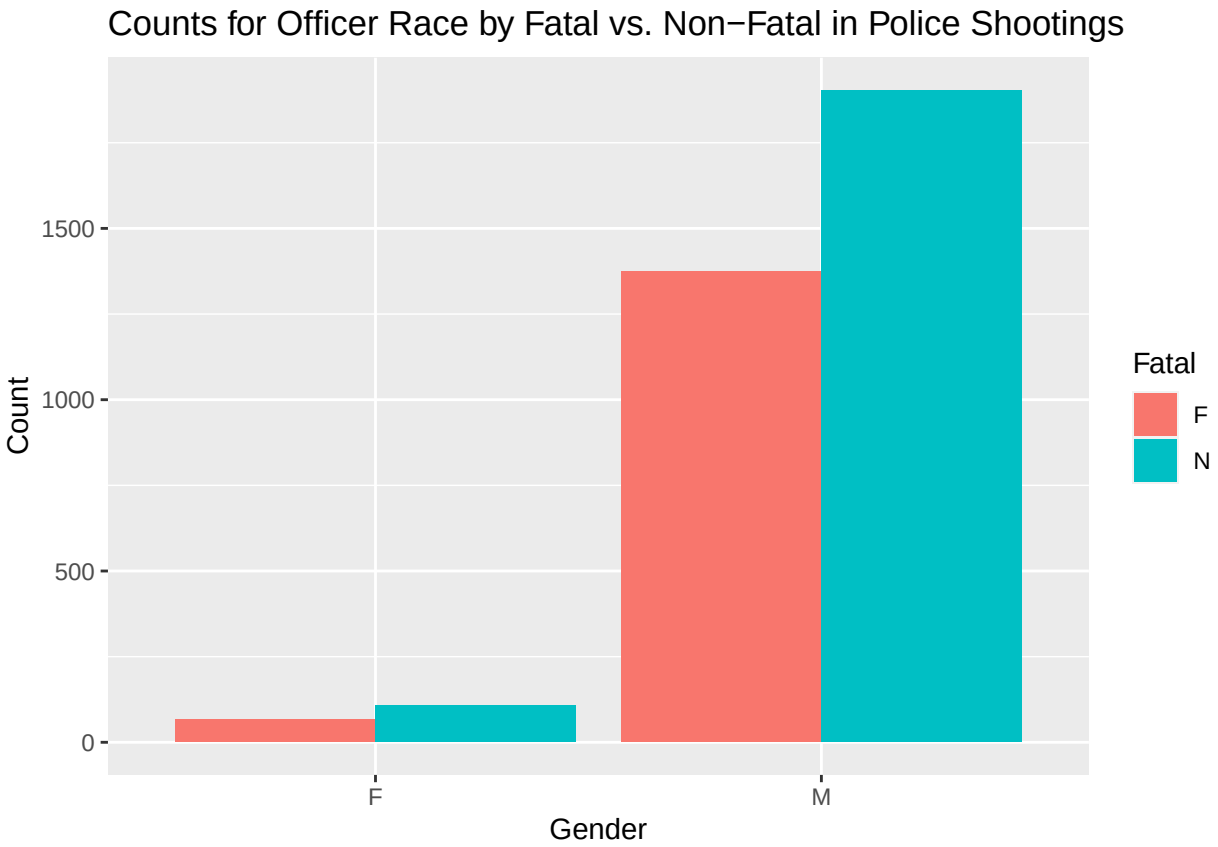
In this dataset, it is important to keep in mind that there is a higher number of non-fatal shooting events than fatal shooting events.


Counts for Victim Race by Fatal vs. Non−Fatal in Police Shootings
Count
1200
900
600
300
0
A
B
L
O
W
Race
Fatal
F
N

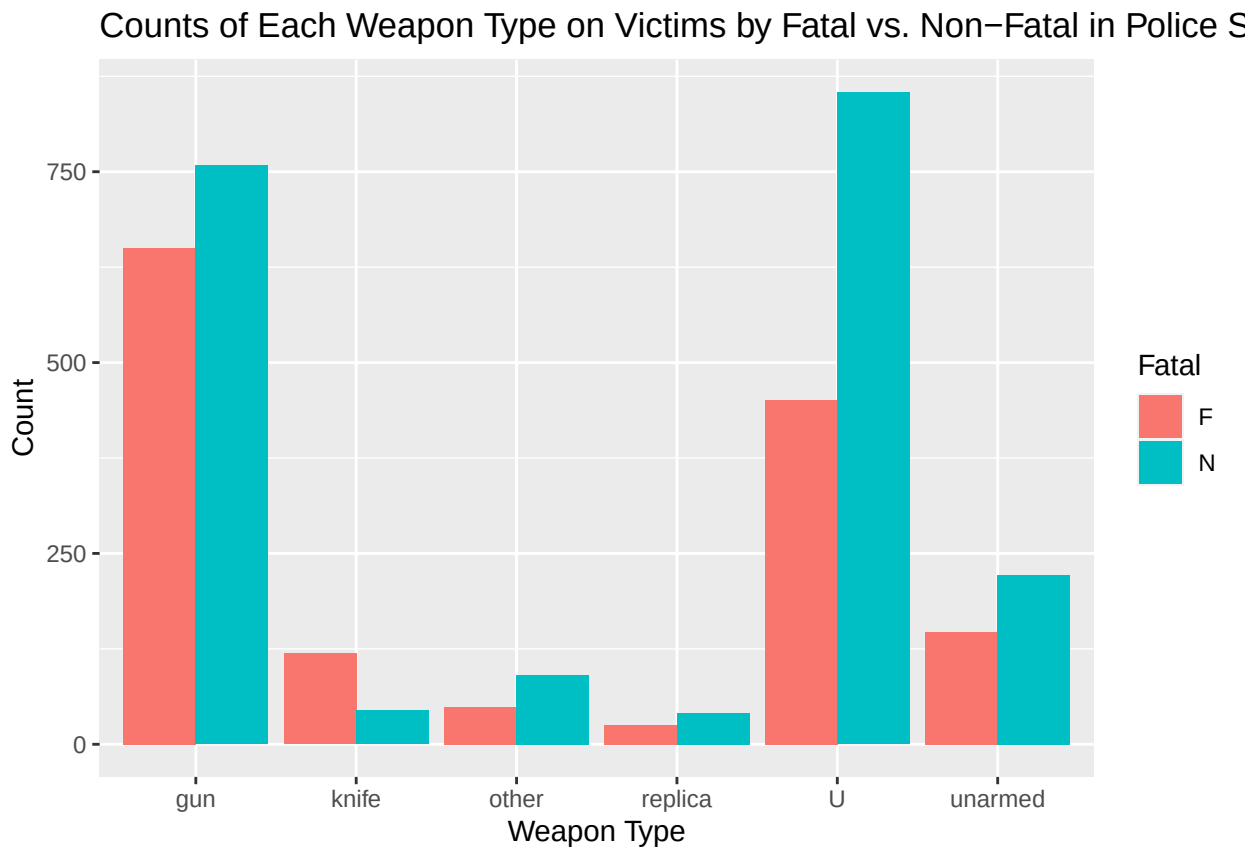


In the plots above, we can see the counts for both victim and officer race in the final dataset split by fatal versus non-fatal shootings. In this dataset, A stands for Asian, B stands for Black, H stands for Hispanic, L stands for Latino, O stands for other, and W stands for White. Already, we can see that in both fatal and non-fatal cases, Black is the race with the highest count for the victim category, whereas White is the race with the highest count for the officer category. It is also interesting to note that the number of Black victims is higher in non-fatal cases than in fatal cases, while the number of white victims is higher in fatal cases than in non-fatal cases.


Counts for Victim Gender by Fatal vs. Non−Fatal in Police Shootings
Count
1500
1000
500
0
F
M
Gender
Fatal
F
N

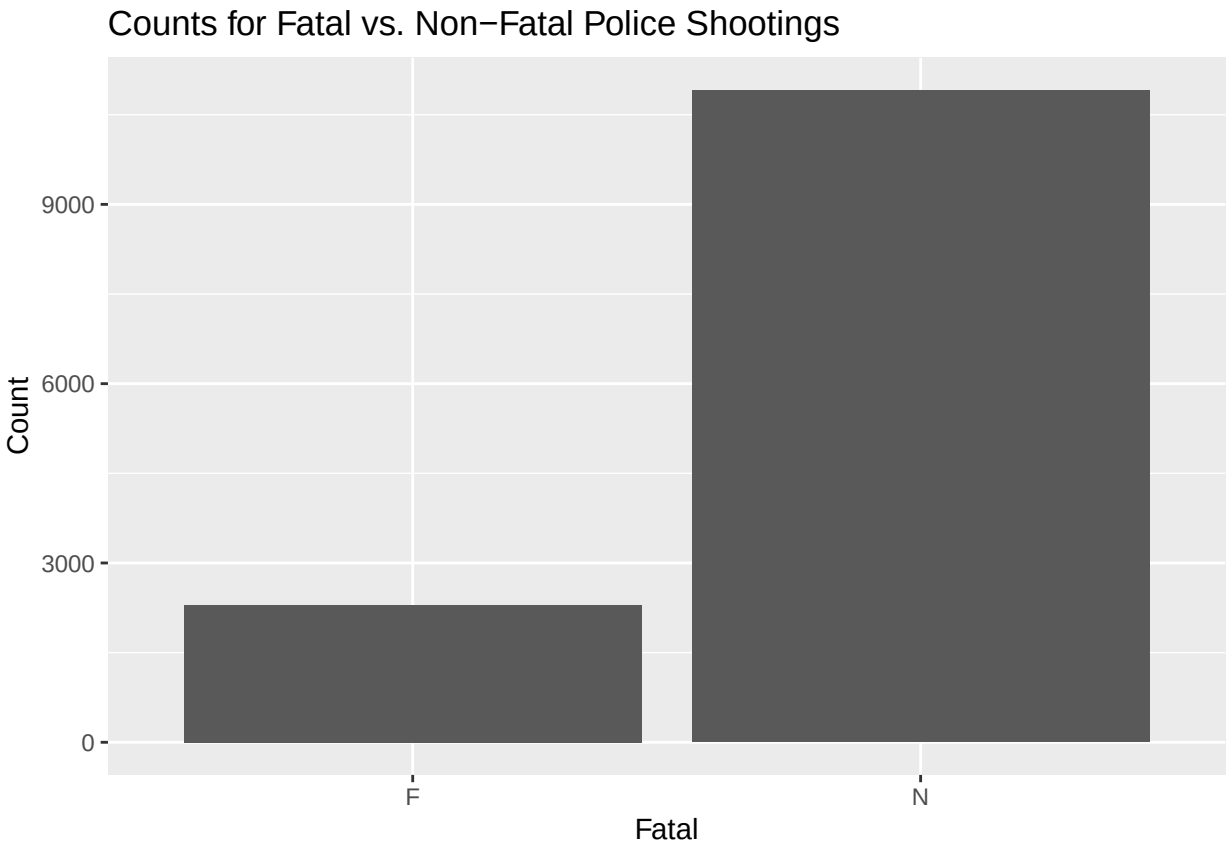


Additionally, it is evident from the graphs above that male is the gender with the highest count in the cases of victims and officers for both fatal and non-fatal shootings.

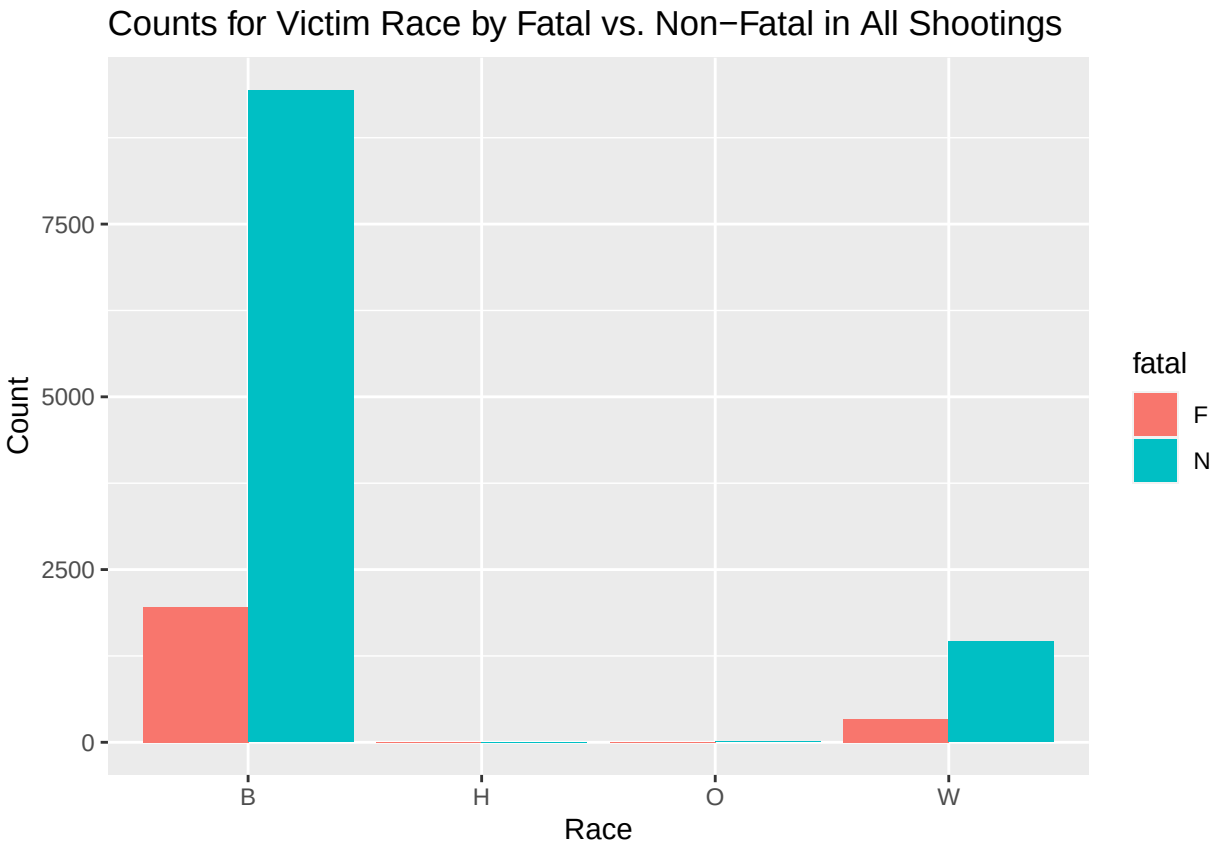


Finally, we can see that in cases of police shootings, victims have guns the most often in fatal cases and unknown weapons in non-fatal cases, whereas they have gun replicas the least often in both cases.

As I mentioned above, in order to draw comparisons between the demographics of those involved in police shootings versus non-police shootings, I will be using data collected by the Police Data Initiative on the shootings in two cities (Philadelphia and Cincinnati). I will only be using data from these two cities because obtaining data on both fatal and non-fatal shooting crime in America is extremely difficult, since non-fatal shootings are "generally lumped-in with other types of assault in national data collections." [2] The Police Data Initiative also suggests using general shooting data from Rochester, New York. However, after looking into this dataset, I found that it did not specify whether each shooting was fatal or non-fatal. As a result, I have decided to leave Rochester out of my final dataset so that I can primarily focus on identifying trends in victim and officer demographics in the context of fatal versus non-fatal shootings. The main predictors I will use from the Philadelphia and Cincinnati datasets will be race and sex, both of which refer to the victims of the shootings. I will combine these selected variables from the two datasets into one final dataset. Once again, I will be using fatality of the shooting as the response variable.



In this dataset, it is important to keep in mind that again, there is a significantly higher number of non-fatal shooting events than fatal shooting events.


Counts for Victim Race by Fatal vs. Non−Fatal in All Shootings
Count
7500
5000
2500
0
B
H
O
W
Race
fatal
F
N

Furthermore, it appears that Black is the race with the highest count and male is the gender with the highest count for victims of all shootings in these two cities, which is similar to what we discovered for the police shooting data. However, it does not seem that this trend in race and gender varies too much between fatal and non-fatal shootings.

II. Methods

Data Processing

Model

III. Preliminary Results

References

1) Vice
2) https://www.pnas.org/content/pnas/116/34/16793.full.pdf
3) https://www.kpbs.org/news/2020/jun/18/sweeping-change-us-views-police-violence-new-poll-/
4) http://govred.com/blog/deescalation-training-state-requirements/