Police Shootings and Racism in America

Connie Wu

11/09/2020

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 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 the race of both victim and officer, as well as other factors such as whether they are carrying a weapon and the total number of victims in the crime, play in fatal versus non-fatal police shootings. Additionally, to account for the varying locations of the homicides, I will be using a dataset 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, I will explore how the race demographics of each location relates to police use of deadly force. 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.

II. Data

As a basis for this study, I will be using the same dataset that VICE News 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 [(victims)] over seven years." [1] VICE News provides 34 variables in the dataset, including dates ranging from January 2010 to September 2016, 47 unique cities, subject race, subject gender, office race, officer gender, the type of weapon the subject was carrying, and whether the shooting was fatal or not. Race was broken up into 6 categories: White (non-Hispanic) (represented as W), Black (non-Hispanic) (B), Asian (A), Latino (L), Other (O), and Unknown (U). Gender was broken up into 3 categories: Male (M), Female (F), and Unknown (U). Weapon type was broken up into 5 values: “gun”, “knife”, “replica”, “other”, and “unknown”. Additionally, there were cases in which multiple victims and/or multiple officers were present in the shooting. Each of these scenarios was still represented within one row of the dataset, but semi-colons were present in the respective victim and officer columns, separating each individual’s information from each other.

Additionally, to assess the effect of different types of legislation on fatal vs. non-fatal police shootings in various states, I will be using a dataset from Kaggle containing 135 variables that detail whether a certain gun provision is absent or present in a certain year and U.S. state for 133 different gun provisions. [4] The years range from 1991 to 2017, and the gun provisions address 14 categories, some of which are dealer regulations, buyer regulations, prohibitions for high-risk gun possession, background checks, ammunition regulations, possession regulations, concealed carry permitting, assault weapons and large-capacity magazines, child access prevention, gun trafficking, and domestic violence. A 1 in the gun provision column represents a presence of the law, and a 0 indicates an absence. I will also add a column to this gun provision dataset indicating whether a state requires de-escalation training based off of information reported by APM Report in 2017, with a 1 meaning the state does require training and a 0 meaning the state does not. There were 16 states that required de-escalation training as of November 2017. [5]

Finally, to examine city demographics as a predictor, I will be using population data from the 2013 American Community Survey that VICE News had already cleaned and standardized to the shootings dataset, meaning the city names can be matched up between the two datasets. VICE News most likely provided only 2013 census data because it is the average year of all the years represented in the shootings dataset. Furthermore, using census data only from 2013 requires that we analyze this data under the assumption that there is no drastic difference in population demographics between 2010 and 2016. The census data includes 7 variables: the police department and the city it’s located in, as well as the city’s Black, Asian, Hispanic, White, and overall total population in 2013.

Data Processing

Although the VICE News dataset was already relatively clean with regards to the victim’s data, there was still quite a bit of cleaning that needed to be done for the race and gender data of the officers. To clean these columns, I replaced all values that were not “W”, “B”, “A”, “O”, or “U” in the race column with the most informed guess that I could make about what the values represented. For example, “A/PI” values were taken to represent Asian/Pacific Islander and thus were replaced with “A”, and “A/W” or other values with a “/” in them were typically taken to represent multi-racial individuals and as a result were replaced with “O.” Similar procedures were carried out in the gender column for officers for values that were not “M”, “F”, or “U”.

After this initial cleaning, I created new columns in the same dataset that represented whether any victims of race Black, White, Asian, or Other were present in the crime, respectively, and if any male or female victims were present in the crime, respectively. I also added new columns to represent the same information for officers (for each race and gender, whether each was present), and a new column to represent whether any victim involved in the shooting was shot fatally. Additionally, I made sure that each weapon type had its own column. All of these added columns had values of either 0 or 1, with 0 representing an absence of the variable and 1 a presence. Finally, I filtered out 2,063 rows that had only unknown and NA values for the fatality of the shooting or the races or genders of the victims or officers, as these rows did not provide sufficient information for my analysis, and I selected only the columns I needed, such as whether any victim of the shooting was shot fatally or not, the genders and races of the victims and officers, and the weapons that the victims were carrying, if any.

For the gun provisions dataset, I created a new column that sums up the total number of laws listed in the dataset that each state had in 2013 (for the same reason that VICE News used 2013). Additionally, as I mentioned earlier, I appended a column to this dataset representing whether a state requires de-escalation training for police officers as of November 2017. 2017 was used under the assumption that no major event occurred between September 2016 and November 2017 that caused a sudden increase in the number of states that require de-escalation training for police officers. Finally, for the census data, the only cleaning that needed to be done was extracting the state that each city was in from the department column so that the census could be easily joined with the cleaned shootings dataset by city name, as well as the gun provisions dataset by abbreviated state name.

After merging the three datasets, the final dataset consists of 2,054 observations.

Exploratory Data Analysis

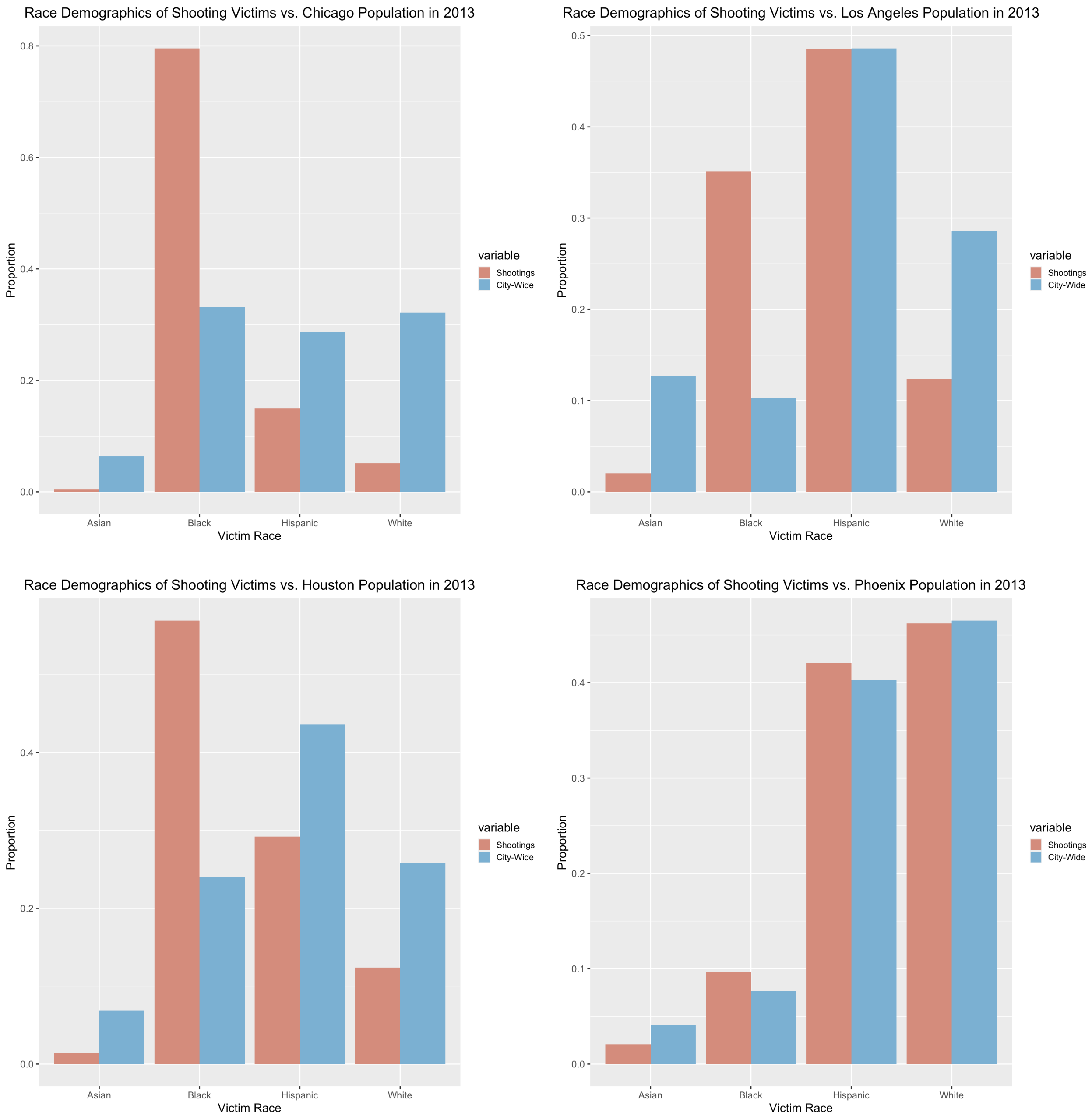

Figure 1: Race Proportions of Shooting Victims vs. City Population Demographics of the Most Popular 4 Cities in the Data

Plotted above in *Figure 1* are the race proportions of shooting victims in the top 4 most popular cities in the dataset (Chicago, Los Angeles, Houston, and Phoenix) vs. the race proportions of those cities' total populations in 2013. Because the bars representing shootings in the Black race category are taller than the bars representing city-wide population in all 4 plots, it is evident Black people are disproportionately the victims of police shootings compared to city race proportions. This can especially be seen in Houston, where Black people make up less than 25% of the city population yet more than 50% of the victim population of police shootings.

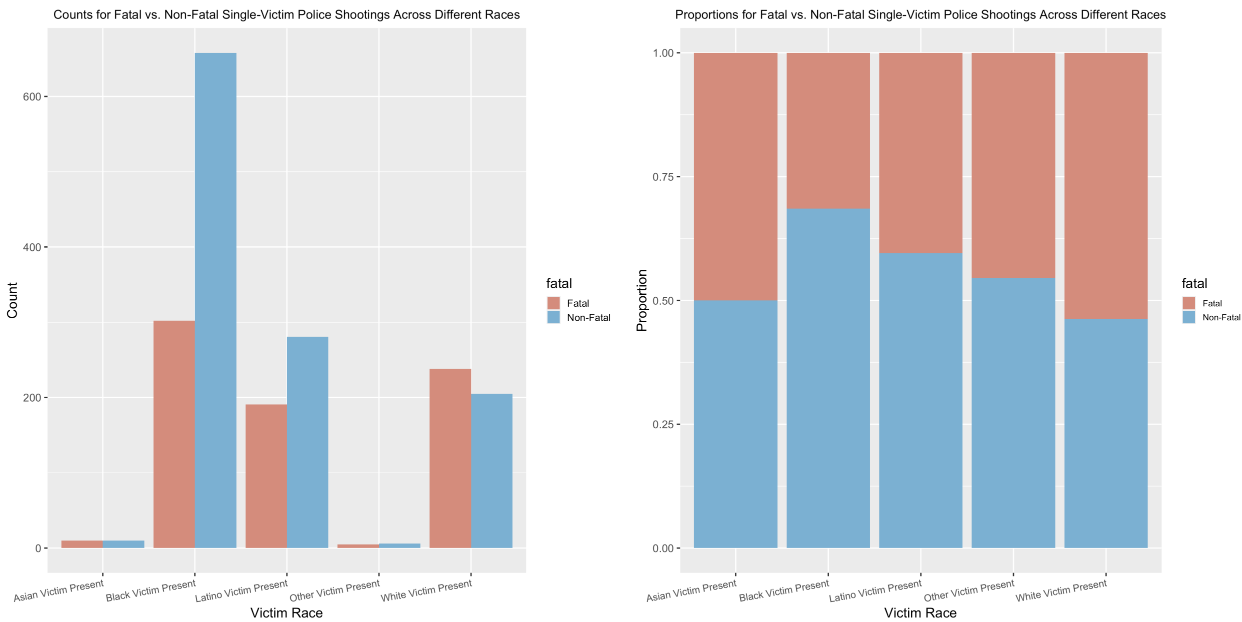

Figure 2: Counts (left) and Proportions (right) of Different Races Across Single Victim Shootings

Additionally, there appears to be 1,906 incidents in the final dataset in which there was only one victim, and 148 with multiple victims. Among all shootings with only one victim, it appears in *Figure 2* that the victim is more often Black than some other race. However, it seems that Black victims and most victims of the other race types are more likely to be shot non-fatally than fatally, whereas White victims seem more likely to be shot fatally. Finally, it is important to note that the numbers of victims that are of race types Asian and Other are relatively small. Because of this, we will combine Asian and Other race types into one single race category called "Other" for both officers and victims for our models.

Similarly, as seen in *Figure 3*, Black victims make up the highest proportion of victims in shootings with multiple victims. However, in these shootings, all race types are more likely to be shot non-fatally than fatally, with Latinos having the highest proportion of being shot fatally out of all other race types. Additionally, it should be noted that there are no victims with race type "Other" in shootings where there are multiple victims.

III. Methods

Because I want to explore the relationship between the number of victims in a police shooting affects and whether the shooting is fatal or not, I decided to create two separate models: one for shootings with only one victim, and one for shootings with multiple victims. Both models are random effect logistic regression models because the range that the response variable can take in logistic regression is the most reasonable given that I am predicting whether a victim is shot fatally or not (labeled 1 and 0, respectively). Additionally, I will be using random effects to control for correlation within cities.

To create these two models, I first performed backwards selection using simple logistic regression models

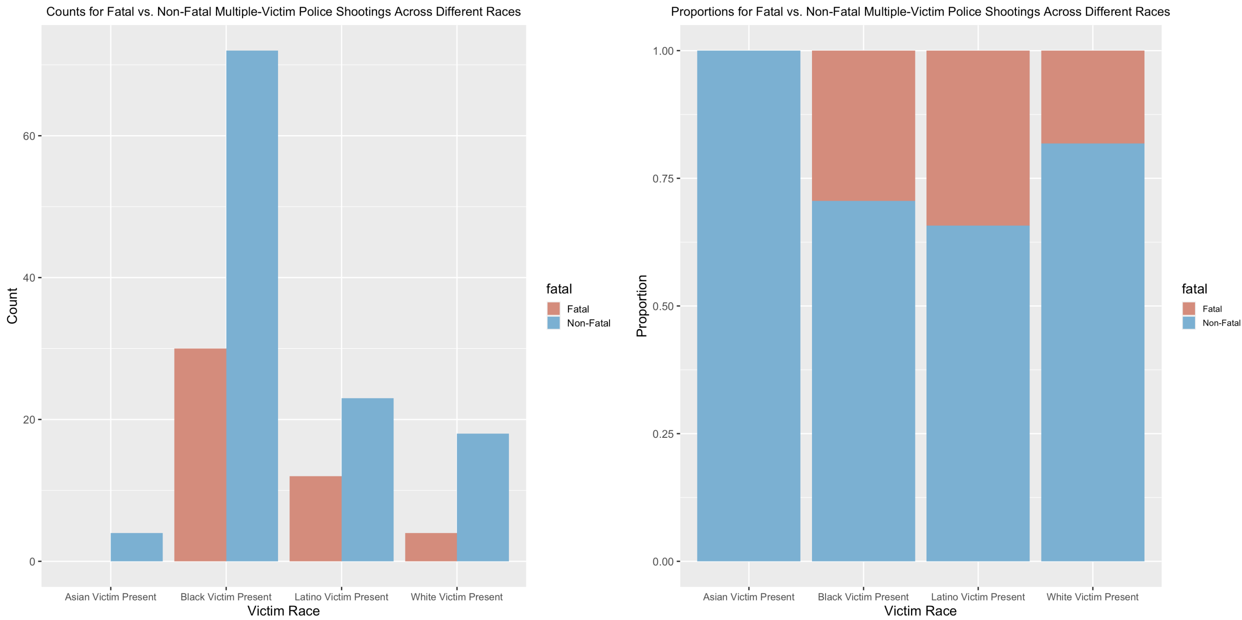

Figure 3: Counts (left) and Proportions (right) of Different Races Across Multiple Victim Shootings

on all relevant predictors in order to get a sense of which predictors are most statistically significant in prediction shooting fatality. Then, I added potential interactions that I wanted to explore one at a time and performed likelihood ratio tests with each new model iteration to check if the added interaction improved the previous model. Whether an interaction was statistically significant or not was determined by whether the likelihood ratio test's p-value was less than 0.05. Finally, I put these variables obtained from backwards selection into a random effects linear regression model with city as the random intercept and evaluated the fit of this model using cross validation, which I will describe in more detail in the Results section. From there, I tweaked the model by experimenting with additional predictors, interactions, and random effects between city and other predictors in order to obtain the best fitting model.

The final model for single-victim shootings is written out below (where p_i represents the probability of getting fatally shot for shooting i):

$$log(\frac{p_i}{(1-p_i)}) = \beta_0 + \beta_{1j}\ VictimRace_{ij} + \beta_2\ OfficerBlack_i + \beta_3\ OfficerWhite_i + \beta_4\ MultipleOfficers_i$$

$$+\beta_5\ VictimWithGun_i + \beta_6\ VictimWithOther_i + \beta_7\ VictimWithReplica_i + \beta_8\ VictimWithKnife_i$$

$$+\beta_9\ VictimUnarmed_i + b_i$$

$$b_i \sim N(0, \sigma_b^2)$$

$$VictimRace = \text{White*, Black, Latino, Other}$$

For single victims, the best fitting model was the one that used city as a random intercept and had random slopes for the SubjectRace in each of these cities.

The final model for multiple-victim shootings is written out below (where p_i represents the probability of getting fatally shot for shooting i):

$$log(\frac{p_i}{(1-p_i)}) = \beta_0 + \beta_{1j}\ VictimBlack_{ij} + \beta_{2j}\ VictimLatino_{ij} + \beta_3\ VictimOther_i$$

$$+\beta_4 \ OfficerBlack_i + \beta_5 \ OfficerLatino_i + \beta_6 \ MultipleOfficers_i + \beta_7 \ VictimWithGun_i$$

$$\beta_8 \ VictimWithOther_i + \beta_9 \ VictimWithReplica_i + \beta_{10} \ VictimWithKnife_i + b_i$$

$$b_i \sim N(0, \sigma_b^2)$$

For multiple victims, the best fitting model was the one that used city as a random intercept and had random slopes for whether a Black victim was present and whether a Latino victim was present in the shootings in each of these cities. It is important to note that none of the legislative predictors were useful in predicting the odds of someone being fatally shot in police shootings in either model.

IV. Results

Single Victims Model

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.5690089	0.1979449	-2.8745818	0.0040456
SubjectRaceB	-0.5798053	0.1573093	-3.6857657	0.0002280
SubjectRaceL	-0.5845731	0.2455106	-2.3810503	0.0172634
SubjectRaceO	-0.3777788	0.7190377	-0.5253949	0.5993087
ob	-0.3292988	0.1581182	-2.0826116	0.0372866
ow	0.2215886	0.1258798	1.7603191	0.0783537
omult	0.7424119	0.1121223	6.6214501	0.0000000
gun	0.5092389	0.1248400	4.0791314	0.0000452
weapon_other	0.0420686	0.2815327	0.1494269	0.8812168
replica	0.2117972	0.3356613	0.6309847	0.5280505
knife	1.3183828	0.2226240	5.9220162	0.0000000

In this single victim model, the β_0 coefficient, which has a value of -0.57, represents the log odds of being fatally shot by police in a 'typical' shooting. This can be defined as a shooting in which the victim is White and unarmed and there is only one officer present, who is neither Black nor White. Thus, $\beta_0 + b_i$, or the city-specific intercept, represents the log odds of being fatally shot for a White, unarmed victim in city i when there is only one officer present, who is neither Black nor White. The variabilities of the intercepts and slopes across different cities and different race groups within cities are all greater than 0, indicating that there is enough variation in the baseline odds of being shot fatally in different cities and different race groups within cities. From the fixed effect coefficients, we can see that generally, all other race types are less likely to be shot fatally than White individuals in 'typical' shootings.

The `ranef` function allows us to further quantify the difference between the average predicted log odds of being fatally shot for the general population versus that of a specific city and subject race type, other predictors remaining unchanged. This can be found in the Appendix. From this table, we can see that the city with the highest baseline odds of being shot fatally is Albuquerque, whereas the city with the lowest baseline odds of being shot fatally is Chicago. For White people in Chicago, the predicted log odds of being shot fatally is 0.76 less than that in a 'typical' shooting. However, for Black people in Chicago, the predicted log odds of being shot fatally is actually 0.28 higher than that in a 'typical' shooting. In contrast, for Latino people in Chicago, the predicted log odds of being shot fatally is 0.23 lower than that in a 'typical' shooting, and for people of other race types in Chicago, the predicted log odds of being shot fatally is 2.86 lower than that in a 'typical' shooting. We can conclude that in Chicago, Black people are disproportionately shot fatally in police shootings with singular victims. In Albuquerque, the trend is almost opposite: for Black people, the log odds of being shot fatally is 0.43 less than that in a 'typical' shooting, whereas for White people, the log odds of being shot fatally is 0.67 more. Overall, we can conclude that city plays a large role in whether a race type is fatally shot in police shootings, although the general trend is that White individuals are most likely to be shot fatally than other race types.

Multiple Victims Model

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2.0374118	9.621138e-01	-2.1176413	0.0342055
sb	-0.0063657	1.164355e+00	-0.0054671	0.9956379
sl	0.9453487	9.513395e-01	0.9937028	0.3203676
so	-19.3352867	8.585866e+03	-0.0022520	0.9982032
ob	-0.8758863	5.766520e-01	-1.5189166	0.1287835
ol	-0.4423159	4.495326e-01	-0.9839463	0.3251419
omult	1.4757664	4.499533e-01	3.2798211	0.0010387
gun	0.1546641	4.768530e-01	0.3243433	0.7456781
knife	20.3830991	1.735376e+04	0.0011746	0.9990628
weapon_other	-0.3990949	9.394298e-01	-0.4248267	0.6709630
replica	-18.5875558	9.806992e+03	-0.0018953	0.9984877

In this multiple victims model, the β_0 coefficient, which has a value of -2.03, represents the log odds of being fatally shot by police in a 'typical' shooting. This can be defined as a shooting in which the victim is White and unarmed and there is only one officer present, who is neither Black nor Latino. Thus, $\beta_0 + b_i$, or the city-specific intercept, represents the log odds of being fatally shot for a White, unarmed victim in city i when there is only one officer present, who is neither Black nor Latino. The variabilities of the intercepts and slopes across different cities and different race groups within cities are all greater than 0, indicating that there is enough variation in the baseline odds of being shot fatally in different cities and different race groups within cities. From the fixed effect coefficients, we can see that generally, the Black and Other race types are less likely to be shot fatally than White individuals in 'typical' shootings, while Latino individuals are slightly more likely to be shot fatally than these White individuals.

The output of the `ranef` function for this model can again be found in the Appendix. From this table, we can see that the city with the highest baseline odds of being shot fatally is Tucson, whereas the city with the lowest baseline odds of being shot fatally is again Chicago. For non-Black and non-Latino people in Chicago, the predicted log odds of being shot fatally is 1.04 less than that in a 'typical' shooting. However, for Black people in Chicago, the predicted log odds of being shot fatally is 2.23 higher than that in a 'typical' shooting. In contrast, for Latino people in Chicago, the predicted log odds of being shot fatally is 0.70 higher than that in a 'typical' shooting but still lower than that for Black people. We can conclude that in Chicago, Black people are disproportionately shot fatally in police shootings with multiple victims. In Tucson, the trend is almost opposite: for Black people, the log odds of being shot fatally is 1.37 less than that in a 'typical' shooting, whereas for non-Black and non-Latino people, the log odds of being shot fatally is 0.64 more. Overall, we can conclude that city plays a large role in whether a race type is fatally shot in police shootings with multiple victims, although the general trend is that Latino individuals are most likely to be shot fatally than other race types.

Differences Between Single and Multiple Victim Shootings

In both types of shooting, it was extremely important to capture correlations within cities, since there proved to be significant variations in trends between different cities. Additionally, it appears that most of the predictors remained the same between the two models, the only difference being that the presence of a Latino officer was statistically significant in multiple-victim shootings, whereas the presence of a White officer was statistically significant in single-victim shootings. Unfortunately, I was unable to add a random slope for Other race type individuals in the multiple-victim shootings because it forced the intercept not to be random anymore. Thus, I was unable to interpret the multiple-victim shootings coefficients in terms of White vs. other race types of individuals.

Some trends that remained constant between the types of shootings were that having multiple officers there increased the log odds of the shooting being fatal, whereas having a Black officer present decreased the log odds of the shooting being fatal. Additionally, the victim having a knife seemed to lead to the largest log odds of the shooting being fatal out of all types of weapons. Finally, many trends within cities remained the same among both types of shootings, one prominent example being Chicago, in which Black individuals are disproportionately shot fatally. One trend that changed between the two types of shootings was that White individuals were most likely to be shot fatally than other race types in single-victim shootings, whereas Latino individuals were most likely to be shot fatally in multiple-victim shootings.

Model Validation and Diagnostics

As I mentioned earlier, I performed 5-fold cross-validation on my models to assess model accuracy and fit. I chose to use 5 folds and 85% of the data because this is commonly used in practice, but I tried it with 90% of the data as well just to verify that my cross-validation technique was not biased for the percentage of the full data allocated for the training set. I used the `groupdata2` R package to partition our data set so that the training set contained 85% of the data from the full dataset, and then I ran 5-fold cross-validation on this training set before predicting on the test set. When assessing the performance of different models, we considered a higher average cross-validation and test set prediction accuracy to be indicative of a better model. For the single victims model, I achieved an average 5-fold cross-validation prediction accuracy of 0.66 and a test set prediction accuracy of 0.70. For the multiple victims model, I achieved an average 5-fold cross-validation prediction accuracy of 0.63, and a test set prediction accuracy of 0.40. This test set prediction accuracy was relatively low, and although adding $\frac{black_i}{white_i}$ into the model significantly increasing the predictive accuracy of the model, I decided not to include it as a predictor because I suspected it was being used as a sort of index to identify the city that the shooting occurred in, which led to such a large boost in test set prediction accuracy.

I will use a binned residual plot and QQ-plot in the next submission for model diagnostics.

Sensitivity Analysis

For sensitivity analysis, I plan to test my assumptions about my imputed values, especially in the `fatal` column. Instead of assuming ‘U’ means ‘N’, I will instead assume that it means ‘Y’ and thus will be filtering out less columns as a result. I will incorporate this analysis into my next submission.

V. Discussion and Conclusion

Overall, it appears that the most of the statistically significant predictors in single-victim vs. multiple-victim shootings overlap, meaning there is not a huge difference in these two types of shootings. Furthermore, the general trends among both types of shootings seem to be pretty similar, with the most important trend that changed among the two being that White individuals were most likely to be shot fatally than other race types in single-victim shootings, whereas Latino individuals were most likely to be shot fatally in multiple-victim shootings. Finally, it appears that legislative predictors are not significant in predicting whether a shooting will be fatal or not, whereas city, number of officers, officer race, victim weapon-type, and subject race are statistically significant. Surprisingly, White people generally have one of the highest odds ratios of being fatally shot in both types of police shootings, although this trend differs by city.

One strength of my analysis is that I am considering many variables as predictors, from individual level data to contextual predictors, such as state gun legislation. This allows me to thoroughly explore and determine whether the relationships between a wide range of potential predictors and police shooting fatality in the context of police brutality are significant in any types of shootings (single vs. multiple victims, for example). There are several limitations to this study. For example, the VICE News dataset spans over the years 2010 to 2016, and neither of my models will be exploring time series data whatsoever. Thus, the aspect of time may be playing some sort of effect on my response variable that I am not accurately capturing in my models. There are also several other confounding factors that have not been accounted for, such as how well one can shoot at baseline or what training they have received with regards to shooting. Additionally, when cleaning the data, there were several "Unknown" values in the fatality column. Unknown values were assumed to represent "No" in this study, although this may not necessarily be true, to allow for more straightforward cleaning of the data and not lose too many observations in the data. Finally, there were less than 200 cases in which there were multiple victims in my final dataset, which is a relatively small sample size. Thus, it may be necessary to take the results from the multiple-victim model with a grain of salt, especially since the test set predictive accuracy was so low.

VI. References

[1] McCann, Allison, et al. "Police Shoot Far More People than Anyone Realized, a VICE News Investigation Reveals." VICE News, VICE News, 12 Dec. 2017, news.vice.com/en_us/article/xwvv3a/shot-by-cops.

[2] Edwards, Frank, et al. "Risk of Being Killed by Police Use of Force in the United States by Age, Race–Ethnicity, and Sex." PNAS, PNAS, 20 Aug. 2019, www.pnas.org/content/116/34/16793.

[3] Press, Associated. "Sweeping Change In US Views Of Police Violence, New Poll Finds." KPBS Public Media, KPBS, 18 June 2020, www.kpbs.org/news/2020/jun/18/sweeping-change-us-views-police-violence-new-poll-/.

[4] https://www.kaggle.com/jboysen/state-firearms

[5] Gilbert, Curtis. "Most States Neglect Ordering Police to Learn De-Escalation Tactics to Avoid Shootings." Not Trained to Not Kill | APM Reports, APM Reports, 16 Sept. 2020, www.apmreports.org/story/2017/05/05/police-de-escalation-training.

[6] DeGue, Sarah et al. "Deaths Due to Use of Lethal Force by Law Enforcement: Findings From the National Violent Death Reporting System, 17 U.S. States, 2009-2012." American journal of preventive medicine vol. 51,5 Suppl 3 (2016): S173-S187. doi:10.1016/j.amepre.2016.08.027

VII. Appendix

Single-Victim Shooting

Table 3: Random Effects Coefficients for Single-Victim Shootings

grpvar	term	grp	condval	condsd
city	(Intercept)	Albuquerque	0.6725034	0.3365752
city	(Intercept)	Austin	0.3207638	0.2922127
city	(Intercept)	Boston	0.0127233	0.4652367
city	(Intercept)	CharlotteMecklenburg	-0.0382629	0.3207061
city	(Intercept)	Chicago	-0.7559945	0.2505983
city	(Intercept)	Cincinnati	0.3890076	0.3776280
city	(Intercept)	City of Miami	0.3395626	0.4077577
city	(Intercept)	Columbus	-0.3223968	0.2458863
city	(Intercept)	Dallas	-0.1218501	0.2729989
city	(Intercept)	DekalbCounty	0.3986893	0.3220665
city	(Intercept)	Denver	-0.3130670	0.3015671

grpvar	term	grp	condval	condsd
city	(Intercept)	El Paso	0.0807487	0.3579450
city	(Intercept)	Fort Worth	-0.1010622	0.3147454
city	(Intercept)	Houston	-0.4389049	0.2641939
city	(Intercept)	Indianapolis	-0.0506660	0.2954652
city	(Intercept)	Jacksonville	-0.0639388	0.4058351
city	(Intercept)	Kansas City	-0.2954982	0.3395286
city	(Intercept)	LasVegas	-0.0219129	0.4684466
city	(Intercept)	LosAngeles	-0.2005255	0.1806564
city	(Intercept)	LouisvilleMPD	-0.3697411	0.3754565
city	(Intercept)	Memphis	0.4560131	0.3747059
city	(Intercept)	MiamiDade	-0.0870951	0.4587624
city	(Intercept)	Milwaukee	-0.4241744	0.3541793
city	(Intercept)	Phoenix	0.0664581	0.2011946
city	(Intercept)	San Antonio	0.2198717	0.2906073
city	(Intercept)	San Francisco	-0.0352846	0.4610324
city	(Intercept)	Seattle	-0.0828059	0.2948878
city	(Intercept)	St. Louis	0.2625389	0.4569249
city	(Intercept)	Tampa	0.1456354	0.3940116
city	(Intercept)	Tucson	0.3229079	0.3182567
city	SubjectRaceB	Albuquerque	-0.4294709	0.2052335
city	SubjectRaceB	Austin	-0.1710587	0.1902321
city	SubjectRaceB	Boston	-0.0050195	0.2617819
city	SubjectRaceB	CharlotteMecklenburg	-0.0319859	0.2134072
city	SubjectRaceB	Chicago	0.2794469	0.1714625
city	SubjectRaceB	Cincinnati	-0.2295903	0.2295251
city	SubjectRaceB	City of Miami	-0.0926814	0.2436530
city	SubjectRaceB	Columbus	0.1157579	0.1801784
city	SubjectRaceB	Dallas	0.0108717	0.1775219
city	SubjectRaceB	DekalbCounty	-0.1153888	0.2260152
city	SubjectRaceB	Denver	0.3048440	0.1873260
city	SubjectRaceB	El Paso	-0.0697764	0.2138662
city	SubjectRaceB	Fort Worth	0.0217158	0.1963430
city	SubjectRaceB	Houston	0.1584256	0.1709962
city	SubjectRaceB	Indianapolis	0.0207495	0.1935093
city	SubjectRaceB	Jacksonville	-0.0099136	0.2392072
city	SubjectRaceB	Kansas City	0.1293862	0.2080266
city	SubjectRaceB	LasVegas	-0.0002837	0.2633057
city	SubjectRaceB	LosAngeles	0.2140127	0.1231778
city	SubjectRaceB	LouisvilleMPD	0.1936593	0.2255812
city	SubjectRaceB	Memphis	-0.2278645	0.2300353
city	SubjectRaceB	MiamiDade	0.0441805	0.2578128
city	SubjectRaceB	Milwaukee	0.1520411	0.2195395
city	SubjectRaceB	Phoenix	-0.0419912	0.1356395
city	SubjectRaceB	San Antonio	-0.0002060	0.1806117
city	SubjectRaceB	San Francisco	0.0191016	0.2624441
city	SubjectRaceB	Seattle	0.0225001	0.2065315
city	SubjectRaceB	St. Louis	-0.1035749	0.2594933
city	SubjectRaceB	Tampa	-0.0456689	0.2306841
city	SubjectRaceB	Tucson	-0.0699688	0.2011974
city	SubjectRaceL	Albuquerque	-0.7602399	0.4429097
city	SubjectRaceL	Austin	-0.1830615	0.5170735
city	SubjectRaceL	Boston	0.0021220	0.6236868

grpvar	term	grp	condval	condsd
city	SubjectRaceL	CharlotteMecklenburg	-0.2565868	0.5497977
city	SubjectRaceL	Chicago	-0.2260097	0.4173976
city	SubjectRaceL	Cincinnati	-0.3396582	0.5910160
city	SubjectRaceL	City of Miami	0.2760111	0.5609042
city	SubjectRaceL	Columbus	-0.1145223	0.5336202
city	SubjectRaceL	Dallas	-0.2180146	0.4115272
city	SubjectRaceL	DekalbCounty	0.2891612	0.5947253
city	SubjectRaceL	Denver	0.9114629	0.4538658
city	SubjectRaceL	El Paso	-0.1880526	0.5681997
city	SubjectRaceL	Fort Worth	-0.1133346	0.5395803
city	SubjectRaceL	Houston	-0.1514708	0.3815343
city	SubjectRaceL	Indianapolis	-0.0044055	0.5675559
city	SubjectRaceL	Jacksonville	-0.1974008	0.6015968
city	SubjectRaceL	Kansas City	0.0187842	0.5930398
city	SubjectRaceL	LasVegas	-0.0511044	0.6143733
city	SubjectRaceL	LosAngeles	0.6834761	0.2721438
city	SubjectRaceL	LouisvilleMPD	0.1923156	0.6118816
city	SubjectRaceL	Memphis	-0.1788293	0.6140292
city	SubjectRaceL	MiamiDade	0.0376631	0.6221316
city	SubjectRaceL	Milwaukee	-0.1520600	0.5838461
city	SubjectRaceL	Phoenix	-0.0727372	0.3307818
city	SubjectRaceL	San Antonio	0.4965537	0.3910654
city	SubjectRaceL	San Francisco	0.0216508	0.5948761
city	SubjectRaceL	Seattle	-0.0678459	0.5665332
city	SubjectRaceL	St. Louis	0.0437869	0.6235158
city	SubjectRaceL	Tampa	0.0869246	0.5986737
city	SubjectRaceL	Tucson	0.3590182	0.4549337
city	SubjectRaceO	Albuquerque	0.8076982	0.8744338
city	SubjectRaceO	Austin	0.7102946	0.9760834
city	SubjectRaceO	Boston	0.0451611	1.5558897
city	SubjectRaceO	CharlotteMecklenburg	-0.5887707	0.9424333
city	SubjectRaceO	Chicago	-2.8642890	0.6039409
city	SubjectRaceO	Cincinnati	0.6484263	1.2605085
city	SubjectRaceO	City of Miami	1.6024227	1.2079478
city	SubjectRaceO	Columbus	-1.2543259	0.8564171
city	SubjectRaceO	Dallas	-0.7903959	0.6788226
city	SubjectRaceO	DekalbCounty	1.8182463	0.8970474
city	SubjectRaceO	Denver	0.6333569	0.8973171
city	SubjectRaceO	El Paso	-0.0782035	1.2639162
city	SubjectRaceO	Fort Worth	-0.5333795	1.1159668
city	SubjectRaceO	Houston	-1.6995820	0.6124357
city	SubjectRaceO	Indianapolis	-0.1725162	1.1080391
city	SubjectRaceO	Jacksonville	-0.5650078	1.3627984
city	SubjectRaceO	Kansas City	-0.9256410	1.2811280
city	SubjectRaceO	LasVegas	-0.1636798	1.5381288
city	SubjectRaceO	LosAngeles	0.5861076	0.3170447
city	SubjectRaceO	LouisvilleMPD	-0.8525977	1.3524068
city	SubjectRaceO	Memphis	1.1571856	1.2969816
city	SubjectRaceO	MiamiDade	-0.2146631	1.5549944
city	SubjectRaceO	Milwaukee	-1.6528106	1.2079053
city	SubjectRaceO	Phoenix	0.0841470	0.5086058
city	SubjectRaceO	San Antonio	1.6129774	0.7210186

grpvar	term	grp	condval	condsd
city	SubjectRaceO	San Francisco	-0.0753934	1.4478192
city	SubjectRaceO	Seattle	-0.3917411	0.9113944
city	SubjectRaceO	St. Louis	0.9318788	1.5245406
city	SubjectRaceO	Tampa	0.6303222	1.3801569
city	SubjectRaceO	Tucson	1.6986068	0.8137966

Multiple-Victim Shooting

Table 4: Random Effects Coefficients for Multiple-Victim Shootings

grpvar	term	grp	condval	condsd
city	(Intercept)	Austin	-0.7962315	0.8809420
city	(Intercept)	CharlotteMecklenburg	-0.2865333	0.8025931
city	(Intercept)	Chicago	-1.0434559	0.2825755
city	(Intercept)	City of Miami	-0.9756777	0.9154628
city	(Intercept)	Columbus	-0.8999761	0.7281901
city	(Intercept)	Dallas	0.1400934	0.9689838
city	(Intercept)	DekalbCounty	0.3577443	0.8682077
city	(Intercept)	Denver	0.2448728	0.8668810
city	(Intercept)	Houston	-0.6596381	0.3598045
city	(Intercept)	Indianapolis	0.3626486	0.7886464
city	(Intercept)	LosAngeles	0.0820983	0.9720299
city	(Intercept)	LouisvilleMPD	0.5895608	0.8274179
city	(Intercept)	Memphis	0.1063146	0.9595129
city	(Intercept)	Milwaukee	0.3566385	0.8953294
city	(Intercept)	NewOrleans	-0.0504010	0.9897927
city	(Intercept)	Phoenix	0.2075865	0.9767241
city	(Intercept)	Seattle	0.6013662	0.8152766
city	(Intercept)	Tampa	0.3566385	0.8953294
city	(Intercept)	Tucson	0.6427279	0.8431321
city	sb	Austin	1.6995610	1.8803761
city	sb	CharlotteMecklenburg	0.6116071	1.7131400
city	sb	Chicago	2.2272631	0.6031592
city	sb	City of Miami	2.0825901	1.9540611
city	sb	Columbus	1.9210046	1.5543264
city	sb	Dallas	-0.2990303	2.0683019
city	sb	DekalbCounty	-0.7636074	1.8531948
city	sb	Denver	-0.5226826	1.8503629
city	sb	Houston	1.4080016	0.7680049
city	sb	Indianapolis	-0.7740756	1.6833706
city	sb	LosAngeles	-0.1752394	2.0748039
city	sb	LouisvilleMPD	-1.2584212	1.7661285
city	sb	Memphis	-0.2269291	2.0480863
city	sb	Milwaukee	-0.7612471	1.9110862
city	sb	NewOrleans	0.1075813	2.1127185
city	sb	Phoenix	-0.4430947	2.0848237
city	sb	Seattle	-1.2836198	1.7402128
city	sb	Tampa	-0.7612471	1.9110862
city	sb	Tucson	-1.3719066	1.7996708
city	sl	Austin	0.5379071	0.5951347

grpvar	term	grp	condval	condsd
city	sl	CharlotteMecklenburg	0.1935722	0.5422048
city	sl	Chicago	0.7049236	0.1908985
city	sl	City of Miami	0.6591349	0.6184558
city	sl	Columbus	0.6079935	0.4919407
city	sl	Dallas	-0.0946424	0.6546128
city	sl	DekalbCounty	-0.2416800	0.5865318
city	sl	Denver	-0.1654278	0.5856356
city	sl	Houston	0.4456293	0.2430718
city	sl	Indianapolis	-0.2449931	0.5327829
city	sl	LosAngeles	-0.0554629	0.6566706
city	sl	LouisvilleMPD	-0.3982874	0.5589756
city	sl	Memphis	-0.0718225	0.6482146
city	sl	Milwaukee	-0.2409330	0.6048543
city	sl	NewOrleans	0.0340492	0.6686705
city	sl	Phoenix	-0.1402385	0.6598419
city	sl	Seattle	-0.4062627	0.5507733
city	sl	Tampa	-0.2409330	0.6048543
city	sl	Tucson	-0.4342053	0.5695916

```
## [1] "Average cross-validation prediction accuracy: 0.626769230769231"
```

	0	1
0	119	79
1	985	597

```
## [1] "Test set prediction accuracy: 0.402247191011236"
```

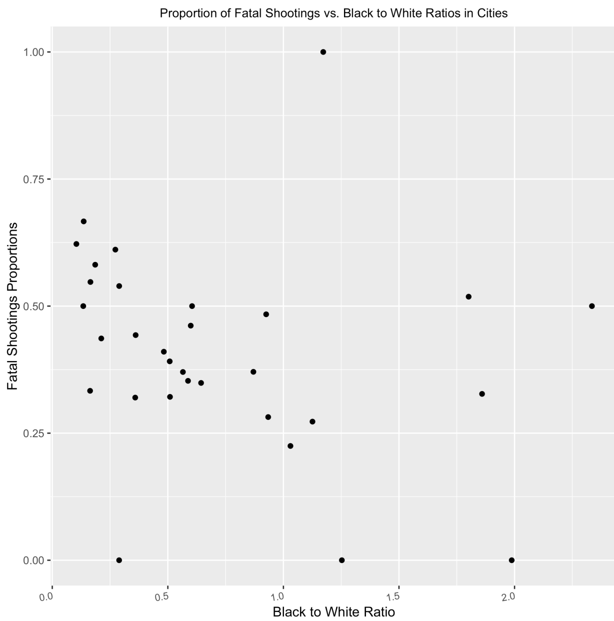

Figure 4: Proportion of Fatal Shootings vs. Black to White Ratios in Cities in Dataset (left) and vs. Black Population Percentages in Cities in Dataset (right)

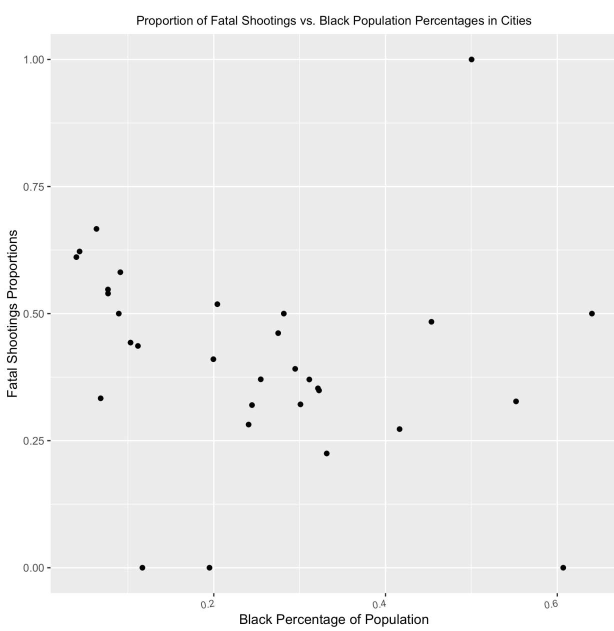

Figure 5: Proportion of Fatal Shootings vs. Black to White Ratios in Cities in Dataset (left) and vs. Black Population Percentages in Cities in Dataset (right)

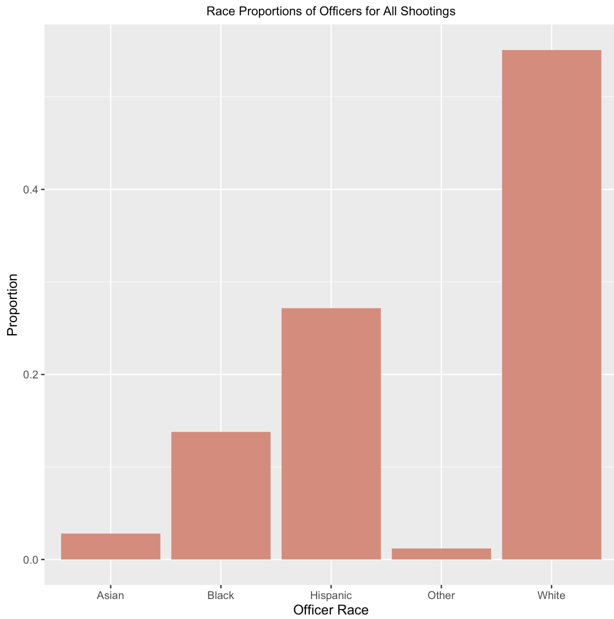

Figure 6: Race Proportions of Officers For All Shootings (left) and For Fatal vs. Non-Fatal Police Shootings Where At Least One Black Victim was Present (right) and

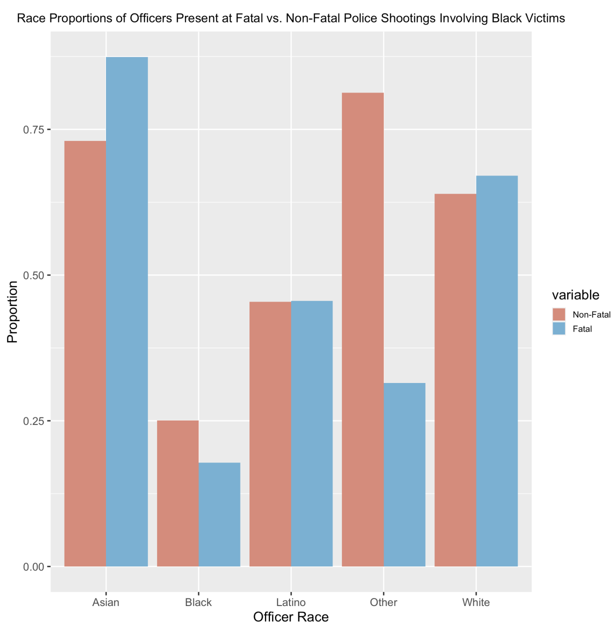

Figure 7: Race Proportions of Officers For All Shootings (left) and For Fatal vs. Non-Fatal Police Shootings Where At Least One Black Victim was Present (right) and