The Average Difference in Death Age Due to Police Violence Between African Americans and Caucasian Americans

Weilin Yin

2022-04-16

Abstract
Police shooting has turned into a rising social issue, particularly for African Americans. Numbers of literature investigated the relationship between police shooting and various factors (e.g., race, age, income, etc.). However, none of them addresses the connection between death age of police killing victims and personal characteristics. Not only is conflict between blacks and police a troubling social issue, but the fact that more young Americans are dying as a result of police use of force is also an issue that deserves attention. In this study, this social issue is studied by utilizing linear regression methods and supervised machine learning methods. The main finding is that being a black male suspect would be killed by the police use of brutal force 10 years earlier than being a white female suspect. Another finding that both linear regression model and machine learning method yield is that being an armed suspect or a suspect with mental illness would be killed at a later age compared to suspects are not armed or with no mental health issue. Nevertheless, this study suffers from a few limitations, and potential future improvements are also discussed.

Keywords
Police Killing, African Americans, Linear Regression, Machine Learning, Econometrics

1. Introduction

Police are one of the main offices in a city in light of the fact that the request for the general public is kept up with by them. As society creates and develops, the force of these specialists increases. Notwithstanding, police shooting has turned into a rising social issue, particularly for African Americans. At last, the passing of George Floyd in Minneapolis, Minnesota set off the flare-up of cross-country fights police savagery (Deliso, 2021). Since society's concerns and questions about police going for the kill are cumulating, the authority released the information of police utilization of power to reduce the social pressure on the government. Henceforth, huge loads of research have been done and populated the literature in the police shooting area. The vast majority of studies use whether cops use force as the outcome of interest (e.g., Fyfe, 1982a; Geller and Karales, 1981; Milton, Halleck, Lardner, and Abrecht, 1977). Additionally, there are quantities of studies center around the impact of race and the impact of age. Discoveries guides out that African Americans has less possibility toward be shot contrasted with different races (Worrall et al., 2018) and police will in general utilize force against more youthful hoodlums over the old ones (Reiss, 1972; Terrill, 2005; Terrill and Mastrofski, 2002). Nevertheless, there are restricted papers studying the connection between police viciousness casualties' ages and the races of casualties. Alongside the deterioration of the relationship between youthful Americans and police (Solomon, 2016), it is additionally vital to be aware of the importance of race on affecting police decision of the utilization of merciless power. In this paper, the distinction between the death age of Black individuals who were shot absurdly and the death age of White individuals was explored. In view of the audit of current literature, my theory is that Black individuals are bound to pass on from police shootings at an early age than White individuals. The rest of this paper is organized by the following sections: **Data** where explains the source of data and shows summaries of data, **Methods** where discussed the statistical models used and its assumption as well as the use of machine learning method, **Results** which demonstrates the findings from different approaches. **Future Steps** and **Conclusion** points out the limitation of this study, summarizes findings, provide a guideline of improving this study.

2. Data

The data used in this study comes from the University of Southern California (USC) and Campaign Zero. The data from USC records all deaths that happen when police are present or that are caused by police. And the data from Campaign Zero documents all police killing through the use of firearm or not. Both datasets contain individual-level data. Datasets were merged to gain better insight. Since the interest of research is about the death age, only suspects that are recorded to be shot to death by the police are kept. The outcome variable (Y) is the age variable which documents the age of people killed due to police shootings. The main independent variables (X) are race. Other independent variables are gender, state, year, mental illness issue, armed or not, geographic area.

2.1 Summary Statistics

	age	gender	race_imputed	mental_illness	armed	geographic_area
count	2273.000000	2273.000000	2273.000000	2273.000000	2273.000000	2273.000000
mean	37.996480	0.951606	0.369116	0.230532	0.886934	0.507699
std	13.662493	0.214645	0.482672	0.421266	0.316744	0.500051
min	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	27.000000	1.000000	0.000000	0.000000	1.000000	0.000000
50%	36.000000	1.000000	0.000000	0.000000	1.000000	1.000000
75%	47.000000	1.000000	1.000000	0.000000	1.000000	1.000000
max	93.000000	1.000000	1.000000	1.000000	1.000000	1.000000

Table 1: Summary Statistics for numerical variables

The Table 1 shows the summary statistics for numeric variables. In the merged dataset, there are 2773 observations in total and the average age among all people who died from police shooting is approximately 37.99 years old. The range starts from a 1-year-old child to 93 years old. The median age is 36 years old, and it is less than the mean which indicates that the distribution is skewed to the right. For the gender variable, the mean is around 0.9516 which means that there are 95.16% of police shooting victims are male as male is represented by 1 and 0 for female. Then the race_imputed variable is a binary variable (1=Black; 0=White) and its mean is 0.3691 which represents that 36.91% of people who were shot to death are African Americans. One thing to notice is that, in this variable, when a person's race cannot be determined, the imputed race is filled. Thus, when analyzing the relationship between death age and race later, the results might suffer from this limitation. For the mental_illness variable, it is a binary variable (1=have mental illness; 0=no mental illness) and its mean is 0.2305 approximately. This implies that 23.05% of victims are experiencing mental illness when killed by police. Then, the armed variable is also binary (1=armed; 0=not armed). The mean of this variable is around 0.8869 which means that 88.69% of people who died from police shooting were armed with a weapon. Lastly, geographic_area variable records whether the event happens in urban area (geographic_area = 1) or not (geographic_area = 0). The mean is 0.5077 which means that 50.77% of police shooting events happen in urban area.

There are also categorical variables that are not applicable for summary statistics. The state initial is recorded in the dataset, for example, CA for California. Moreover, this summary table does not include numerical variables that are meaningless to be summarized on, such as latitude, longitude, and year.

To further understand the data, some exploratory data analyses are conducted. Figure 1 demonstrates the histogram of people's age when people were shot to death by police officers. The reason I choose this variable to plot is that it is the outcome variable that this study is investigating. Understanding if there is a relationship between death age and people's race is important for the public to protect a certain group of people from racist police shootings. From this histogram, the peak happens around age 30 and the graph is slightly skewed to the right. This means that younger people are more likely to be shot to death than the elders. Also, the number of deaths increases dramatically from 18 years old to 30 years old. The reason for such a rapid increase could be that younger people are more active and reckless to engage in criminal activities. Thus, when facing the police, they tend to make decisions that threaten police officers' safety, which causes the occurrence of a police shooting.

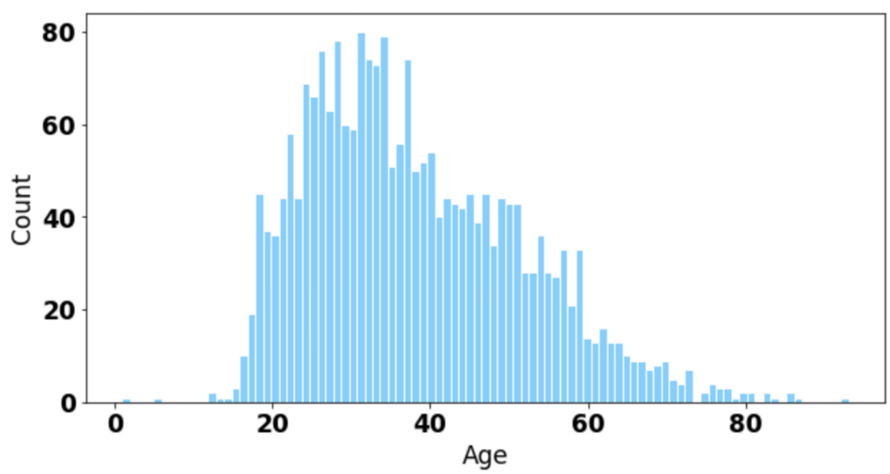

Figure 1: Histogram for the death age of people due to police shooting

Another interesting visualization (Figure 2) is the bar plot for counting the number of deaths per 100k citizens due to police use of brutal force in each State in the US. This variable is demonstrated because it is suspected that there are differences in criminal activity frequency across states in the US. The death counts are divided by 100000 to calculate the death count per

100k citizens so that the difference in population of states is normalized. From this graph, California (CA) has the most police shootings that cause deaths per 100k civilian. The second highest police shooting occurrence is in Texas (TX) and the third highest is Florida (FL). Other states have relatively fewer occurrences of police shootings. The three states that have high police shooting cases might be due to gang activities and conflicts so that police shoot more frequently when arresting criminals.

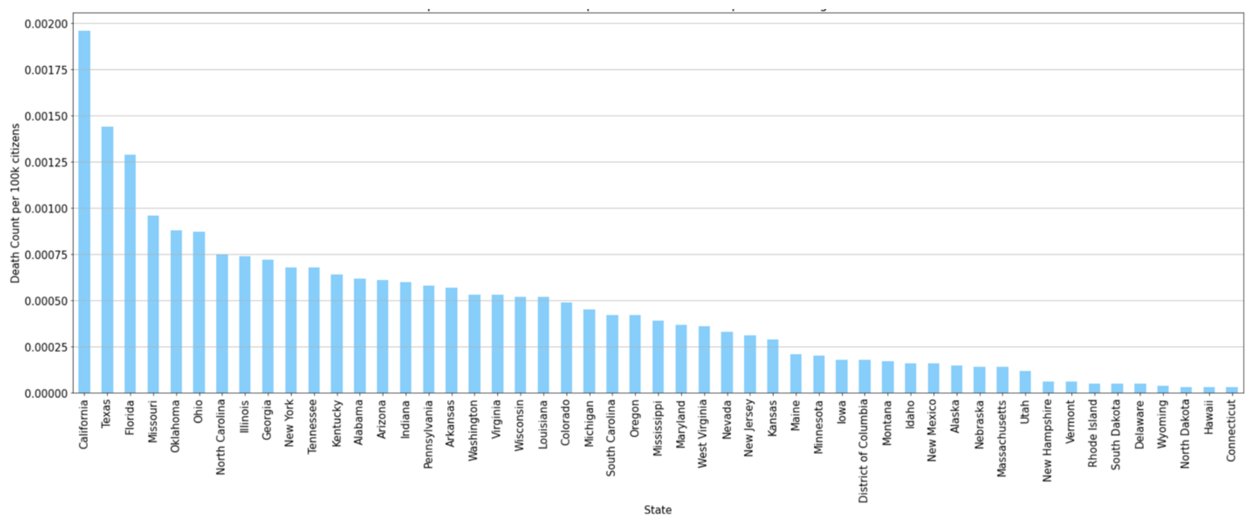

Figure 2: Bar plot for number of deaths per 100k citizens from police shooting in each State

As we have seen the distribution of age of all casualties of police violence, the distributions of age of different victim groups are also informative. Figure 3 demonstrates the distribution of age of victims for the Blacks and the Whites. It is clear to see that most African American victims die around the age of 25 whereas most Caucasian American victims die around the age of 32. Also, we can observe that the density at the peak age for black victims is higher than the density at the peak age for white victims. As the data records death cases from 2013 to 2020 and during this period the conflict between African American community and police had become more severe, especially after the death of George Floyd, thus these conflicts could be the reason of our observed difference in age.

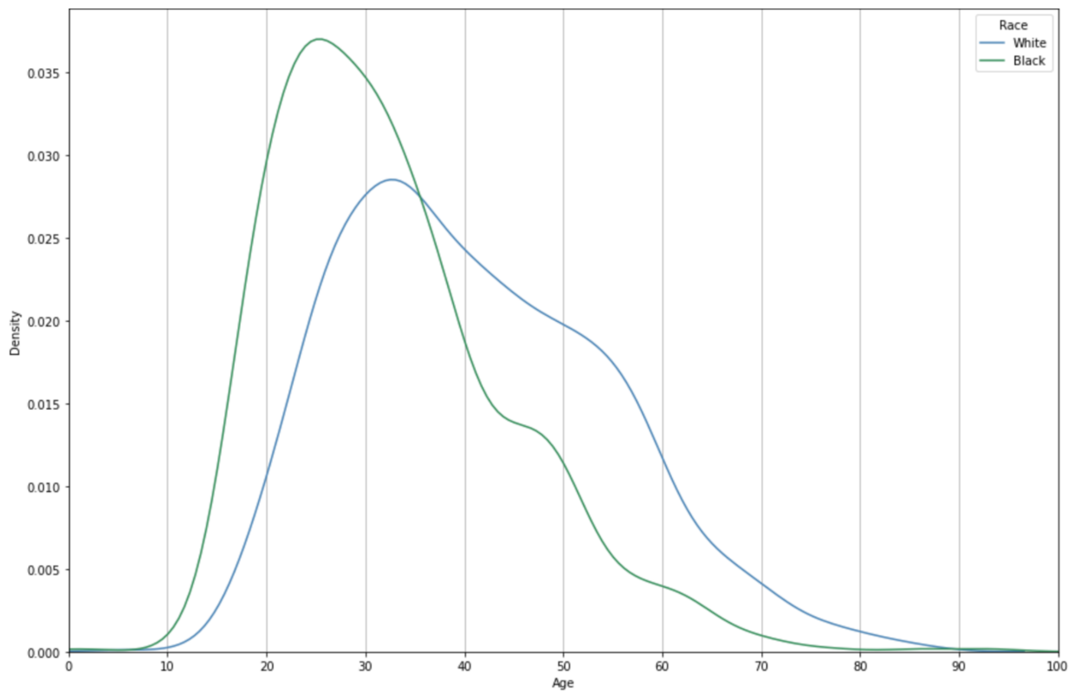

Figure 3: Police Shooting Victims Death Age Distribution

After diving a little deep in the individual-level data of police shooting, it is important to know more about the data from another dimension. Using the US State census map data and the latitude and longitude recorded for each victim in the data, a set of maps of the distribution of the place where police killing occurs can be generated. Figure 5 and Figure 6 in Appendix shows the distribution of place where police shootings occur and the heat map of the counts of police shooting cases in each State, respectively. In figure 5, we observed that most of shootings happened in the middle east part or east part of the US as well as the west coast. Points are sparse for the States that are in the middle region of the US. Figure 6 ends up with consistent pattern with figure 5 that the color of these States is more solid than other States, for example, States in the middle north area, indicating higher death rate per 100k citizens. Another pattern that is obvious to observe through the heatmap is that California has the highest death count per 100k citizens and Texas is the second and lastly Florida is the third highest. This observation matches the result that we saw in the Bar plot for number of deaths per 100k citizens in each State. One potential reason that can explain this pattern is gang activities. Especially for California and Florida, the underground gang activities are strongly associated with popular hip-hop music culture.

Another potential factor that may affects the frequency of police shooting in each State is the population of the State. It is intuitive to consider population as the influential factor because having larger population means higher chances of capturing illegal activities. Hence, using web-scraping ethically, a map visualization of population size of each State and a correlation graph of population and the number of drug arrests are produced. Figure 7 demonstrate the size of each State in the US by the size of circles. we can see that most of the States I just mentioned have relatively large population size. Thus, we can see the correlation that the larger the State's population the higher the number of police killing cases. After having this correlation in mind, from figure 8, we can see that there is a positive relationship between the population size of a state and the number of drug arrests. Thus, we can conclude a correlation that higher population size leads to higher frequency of illegal activities and subsequently leads to higher frequency of police killing.

3. Methods

In this study, a set of linear regression analysis is conducted. I choose to use linear regression is firstly because my outcome variable (age) is a continuous variable and secondly, I believe that the effects of my predictors are linear to age because from existing literature researchers have been using linear regressions to analyze the relationship. All the variables I use to fit linear regressions are listed below.

- Outcome Variable

    - Age: the death age of the suspect who is being shot to death by police
- Independent Variables
    - Race: the race of the suspect, a binary variable that 0 represents white and 1 represents black
    - Gender: the gender of the suspect, a binary variable that 0 represents female and 1 represents male
    - Mental illness: whether suspect has mental illness, a binary variable that 1 represents suspect has mental illness otherwise 0
    - Armed: whether suspect is armed with weapon, a binary variable that 1 represents armed 0 otherwise
    - Geographic region: the geographic region in which the police shooting takes place, a binary variable that 1 represents urban and 0 represents rural

Since race is the main independent variable that this study is interested in, it needs to be in the model. I believe that gender is also an important factor since male suspects in general would cause more threats than women suspects. Mental illness is also included because a suspect with mental health issue would be more dangerous as he/she might not be conscious enough to behave rationally and then police needs to use force to arrest him/her. This would cause higher chance of police killing. Next, suspects who are armed with weapon can be a crucial factor that influence police's decision since the threats from suspects are magnified, thus I included it as one of the predictors. Lastly, I included the geographic region variable since from previous analysis that population size has a positive effect of police shooting and urban area usually has larger population than rural area. Hence geographic region would influence the police killing to some extents.

In order to get unbiased estimates of the effects of each predictor, four linear regression model assumptions need to be satisfied:

- Age is related to the coefficients of predictors in a linear relationship.
- Error is uncorrelated.
- Error variance is constant (homoskedasticity).
- Error follows a Normal distribution.

There are 6 proposed models:

- $age = \beta_0 + \beta_1 \text{race} + \mu$
- $\text{age} = \beta_0 + \beta_1 \text{race} + \beta_2 \text{gender} + \beta_3 \text{mental illness} + \beta_4 \text{armed} + \beta_5 \text{geographic region} + \mu$
- $age = \beta_0 + \beta_1 \text{race} + \beta_2 \text{gender} + \beta_3 \text{mental illness} + \beta_4 \text{armed} + \beta_5 \text{geographic region} + \beta_6 \text{race:gender} + \mu$
- $age = \beta_0 + \beta_1 \text{race} + \beta_2 \text{mental illness} + \beta_3 \text{armed} + \beta_4 \text{mental illness:armed} + \mu$
- $age = \beta_0 + \beta_1 \text{race} + \beta_2 \text{gender} + \beta_3 \text{geographic region} + \mu$
- $age = \beta_0 + \beta_1 \text{race} + \beta_2 \text{gender} + \mu$

To assess these assumptions, residual plots and QQ-plots are generated for each proposed model. By inspecting the residual plots, we observe that points are randomly distributed around 0 which means that the first three assumptions are satisfied. In terms of the last assumption, the QQ-plots for these models show that there is a little skewness in error since some points in the QQ-plot do not follow the 45-degree line, but the amount is insignificant thus we are good to continue.

In all the 6 model specifications, $\beta 0$ is the intercept of the linear trend on the y-axis. other βs are the effect of the corresponding variable on the death age when holding everything else constant. Lastly, the μ in each model represents the random error.

For the first simple linear regression model, I only use the race variable as the predictor since it can directly model the effects of race to the death age. However, this simple model suffers from huge biases as it may omit other important variables that can explain the outcome variable.

For the second model, I include all the predictors because it is the model that can capture all the effects of all variables that potentially can explain the death age.

For the third model, I added an interaction term between race and gender as I suspect that there might be a significant effect of some combination of race and gender on the death age.

For the fourth model, I include the race, mental illness, armed and an interaction between mental illness and armed because I suspect that there might be a significant effect of having mental illness and weapon on the death age.

For the fifth model, I include the race, gender, and geographic region because I would like to examine if gender and geographic region can explain the difference in death age well.

For the last model, I only include the race and gender because I am wondering if race and gender are good enough to explain the death age.
Other than linear regression, one of the most popular supervised machine learning methods is applied --- Decision Tree. First of all, the objective function here is the square error loss function

$$L(y, y_{pred}) = \sum_{N}(y - y_{pred})^2$$

where N represents the total number of observations, y represents the true death age of the suspect, and y_{pred} represents the predicted death age. Thus, in order to get the best model, we tend to optimize the objective function (e.g., minimize the loss function). In other words, we would like to minimize the prediction error, that is minimizing the error when predicting the death age of the suspect.

Since to prevent overfitting, I added a regularization term to the loss function so that the model would get penalized if it has too many predictors and the loss function would get larger but the goal is to minimize the loss function. Thus, the new objective function is as follows:

$$L(y, y_{pred}) = \sum_{m=1}^{|T|}\sum_{N}(y - y_{pred})^2 + \alpha|T|$$

where |T| is the absolute value of the number of terminal nodes of the decision tree. This loss function will have higher value if there are many terminal nodes, which means the decision tree is too complex and it is not preferred. α is a hyperparameter that we need to tune, and it decides how hard the model would be penalized for having too many terminal nodes. If α is large, then the loss function would be increased by a large value which is proportional to the number of terminal nodes in the decision tree and vice versa. It is crucial to tune an appropriate αα because it is neither desirable to penalize too hard, so the final model is too simple nor to have little penalty so that the final model is too complicated and has the risk of overfitting. In both cases, the model would not perform well on new data.

4. Results

4.1 Linear Regression

From Table 2, we can see the estimates of coefficients for each model. To decide the best model, models are selected based on Adjusted R-squared, AIC, and BIC and we want to select a model that has highest adjusted R-squared while having the smallest AIC and BIC. The adjusted R-squared is a way to quantify how much variation the model predictors can explain while penalizing having large set of predictors. AIC and BIC are criteria that penalize having many predictors in the model. The adjusted R-squared of Model 3 is 0.12 and is slightly larger than the ones of other models. The AIC and BIC of model 3 are 18056.29 and 18096.40 respectively which are smaller than the ones of other models. Hence, Model 3 is strongly preferred.

	Model 1	Model 6	Model 5	Model 2	Model 4	Model 3
Intercept	41.12***	42.58***	43.76***	40.36***	37.70***	37.63***
	(0.34)	(1.25)	(1.27)	(1.44)	(0.97)	(1.62)
race_imputed	-8.47***	-8.44***	-6.90***	-6.23***	-7.82***	3.53
	(0.57)	(0.57)	(0.64)	(0.64)	(0.57)	(2.76)
gender		-1.54	-1.67	-1.89		0.91
		(1.28)	(1.27)	(1.26)		(1.48)
mental_illness				3.41***	2.46	3.33***
				(0.65)	(1.90)	(0.65)
armed				2.92***	2.72***	3.03***
				(0.86)	(0.98)	(0.86)
geographic_area			-3.21***	-3.22***		-3.21***
			(0.61)	(0.61)		(0.61)
armed:mental_illness					1.00	
					(2.02)	
race_imputed:gender						-10.19***
						(2.81)
R-squared	0.09	0.09	0.10	0.12	0.10	0.12
R-squared Adj.	0.09	0.09	0.10	0.11	0.10	0.12
F-Statistic	223.53	112.51	84.97	59.49	66.09	52.04
AIC	18126.32	18126.87	18101.70	18067.46	18095.09	18056.29
BIC	18137.78	18144.05	18124.62	18101.84	18123.74	18096.40
No. observations	2273	2273	2273	2273	2273	2273

Standard errors in parentheses.
* p<.1, ** p<.05, ***p<.01

Table 2: Results of 6 linear regression models

Thus, the final model is

$$age = 37.63 + 3.53(race) + 0.91(gender) + 3.33(mental\ illness) + 3.03(armed) - 3.21(geographic\ area) - 10.19(race{:}\ gender)$$

From the final model, we can conclude the following things:

- the average death age of a white female suspect who has no mental illness and not armed in rural area is 37.63
- when holding everything else constant, being a black suspect would die around 3.53 years later than a white suspect
- when holding everything else constant, being a male suspect would die around 0.91 years later than a female suspect
- when holding everything else constant, a suspect having mental illness would die around 3.33 years later than a suspect that has no mental illness
- when holding everything else constant, a suspect who is armed would die around 3.03 years later than a suspect who is not armed
- when holding everything else constant, a suspect who is in urban area would die around -3.21 years earlier than a suspect who is in rural area
- when holding mental illness, armed, and geographic region constant, a black male suspect would die around 10.19 years earlier than a white female suspect.

However, it is important to keep in mind that the estimated coefficients for race and gender are not statistically significant, thus it has less meaning. From this model, we can observe some findings that are aligned with previous analysis. For example, suspects die at a younger age in urban area than rural area from police violence and this is consistent with the findings in state-level analysis. On the other hand, it is surprising to see that having mental illness or suspects are armed with weapon would in fact delay the death age of police killing. Lastly, one important finding is that if keep everything else constant then a suspect who is male and black would die around 10.19 years earlier than a white female suspect. The reason why it is important is that there exists a significant effect of the combination of race and gender on the death age. This finding is intuitive because in daily life Male African Americans have severer conflicts with police and suffer from stereotype issue. From the result of regression model, we understood the relationship between death age and being a male African Americans better.

4.2 Decision Tree

Figure 4 shows the results of decision tree on predicting the death age for people with different characteristics. For each node of the tree, the first line represents the rule of splitting, except the terminal nodes. For example, for the root node, the race_imputed <= 0.5 means when race is white (race_imputed = 0) because race variable is a binary variable. In this decision tree, all of the independent variables are binary variables. The second line displays the squared prediction error based on the corresponding split. For instance, for the root node, the squared error is 185 which means the prediction error is $\sqrt{186} = 13.6$ and it means there are 13.6 years error between the true death age and predicted death age. The third line shows the number of observations that are eligible for the split. In the same example, sample = 2273 in the first root means there are 2273 observations are available to be split based on the condition in the first line. Lastly, we have the value which is simply the predicted value for death age.

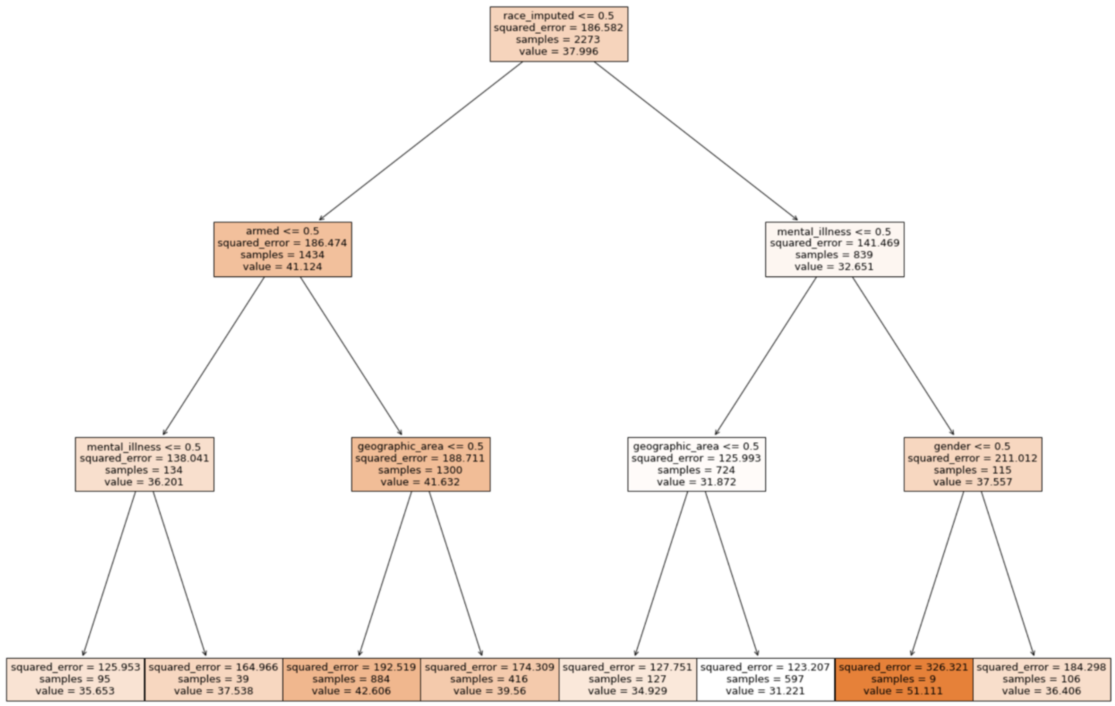

Figure 4: Visualization of the decision tree of predicting the death age for different people

The conclusion we can draw from the decision tree is as follows:

- The predicted death age of an unarmed white suspect with no mental illness is 35.7 years old while 37.5 for an unarmed white suspect with mental illness.
- The predicted death age of an armed white suspect in the rural area is 42.6 years old while 39.6 for the same suspect in the urban area.
- The predicted death age of a black suspect with no mental illness in the rural area is 34.9 years old while 31.2 for the same suspect in the urban area.
- The predicted death age of a black female suspect with mental illness is 51.1 years old while 36.4 for a black male suspect with mental illness.
    - It is noticeable that the squared error for the prediction of a black male suspect with mental illness is 326 which is much larger than other squared error. Thus, the prediction may not be reliable.

Here we can observe a pattern that armed suspects have larger death age than unarmed ones which is consistent with the result of linear regression model. Also, the decision tree shows a pattern that suspects with mental health issue would have higher death age compared to suspects with no mental illness. This finding is aligned with the result in linear regression, and thus the pattern can be justified. Other than the similar results from decision tree and linear regression, there are a few extra information we can observe only in decision tree. Firstly, decision tree shows the squared error for each prediction which researchers can see the performance for specific prediction, while linear regression only displays the mean squared error which only provide a general performance evaluation. Another advantage of decision tree over linear model is that people can follow the rules to understand the relationship between death age and each predictor. In linear regression there is no visualization to display the process of predicting age based on these predictors. Moreover, in decision tree, we can see the number of observations for each splitting group. The benefit of this is that we can determine whether the poor performance of the prediction is due to the lack of training sample or due to model itself.

5. Future Steps

There are 3 ways to further improve the analysis. Firstly, collect more individual-level data to run linear regression model to reduce the omitted variable bias. Then using AIC, BIC, ANOVA test to reduce the insignificant predictors to make better prediction. The data at hand is not rich enough to do it. Secondly, using more advanced machine learning algorithms (e.g., Bagging, Random Forest) to make more unbiased predictions. Lastly, more appropriate models should be implemented to account for the random effect in the data (e.g., Linear Mixed Model) since the observations can be grouped based on the states and neighbourhoods.

6. Conclusion

There have been rising questions about the use of force of police officers. And the conflicts between African Americans have been worsen after a few incidences of police killing. Several findings are established. Firstly, by plotting the density curve, we found that the death age of Black people due to police killing is younger than the White people. To be more specific, African Americans' death ages are concentrated around 25 years old whereas Caucasian Americans' death ages are concentrated around 33 years old. Secondly, I found that most states in the middle east part of the US, California, Texas, and Florida have higher police killing frequency than other states, for instance, some states that are in the middle part of the US. Thirdly, by using HTML-based web scraping and folium maps, I created visualizations for the population and the number of drug arrest cases for each state. Hence, I found that the population size and the drug abuse severity have positive correlations with number of police killing cases. Lastly, by running multiple linear regressions and applying supervised machine learning method, we found that a black male suspect's death age due to police shooting is around 10 years earlier than a white female suspect. Also, both decision tree method and linear regression prove that being an armed suspect or a suspect with mental illness would have older death age compared to unarmed suspect or a suspect with no mental illness respectively and prove that in urban area suspect would die at a younger age than in the rural area.

However, the conclusions suffer from the following limitations. Firstly, the regression models are not large enough to account for all potential influential factors on the death age, thus there exists omitted variable bias. Secondly, the linear regression model does not account for clustering effect as observations can be grouped together. Hence, in the future steps, a linear mixed model is more appropriate to use. Thirdly, modelling the relationship of interest by a single decision tree is not rousted to bias, thus in future, a random forest or bagging is strongly recommended.

Reference

Deliso, M. (2021). Timeline: The impact of George Floyd's death in Minneapolis and beyond. ABC News. Retrieved April 14, 2022, from https://abcnews.go.com/US/timeline-impact-george-floyds-death-minneapolis/story?id=70999322

Fyfe, J. J. (1982a). Blind justice: Police shootings in Memphis. The Journal of Criminal Law and Criminology, 73, 707-722.

Geller, W. A., & Karales, K. J. (1981). Shootings of and by Chicago police: Uncommon crises—Part I: Shootings by Chicago police. The Journal of Criminal Law and Criminology, 72, 1813-1866.

James, L., James, S. M., & Vila, B. J. (2016). The reverse racism effect. Criminology & Public Policy, 15, 457-479.

Klinger, D., Rosenfeld, R., Isom, R. D., & Deckard, M. (2016). Race, crime, and the micro-ecology of deadly force. Criminology and Public Policy, 15, 193-222.

Milton, C. H., Halleck, J. W., Lardner, L., & Abrecht, G. L. (1977). Police use of deadly force. Washington, DC: Police Foundation.

Nix, J., Campbell, B. A., Byers, E. H., & Alpert, G. P. (2017). A bird's eye view of civilians killed by police in 2015: Further evidence of shooter bias. Criminology & Public Policy, 16, 1-32.

Reiss, A. J., Jr. (1972). The police and the public. New Haven, CT: Yale University Press.

Solomon, D. (2016, September 27). The intersection of policing and Race. Center for American Progress. Retrieved April 14, 2022, from https://www.americanprogress.org/article/the-intersection-of-policing-and-race/

Terrill, W. (2005). Police use of force: A transactional approach. Justice Quarterly, 22, 107-138.

Terrill, W., & Mastrofski, S. D. (2002). Situational and officer-based determinants of police coercion. Justice Quarterly, 19, 215-248.

Worrall et al. (2018). Exploring bias in police shooting decisions with real ... Retrieved April 14, 2022, from https://journals.sagepub.com/doi/10.1177/0011128718756038

Appendix

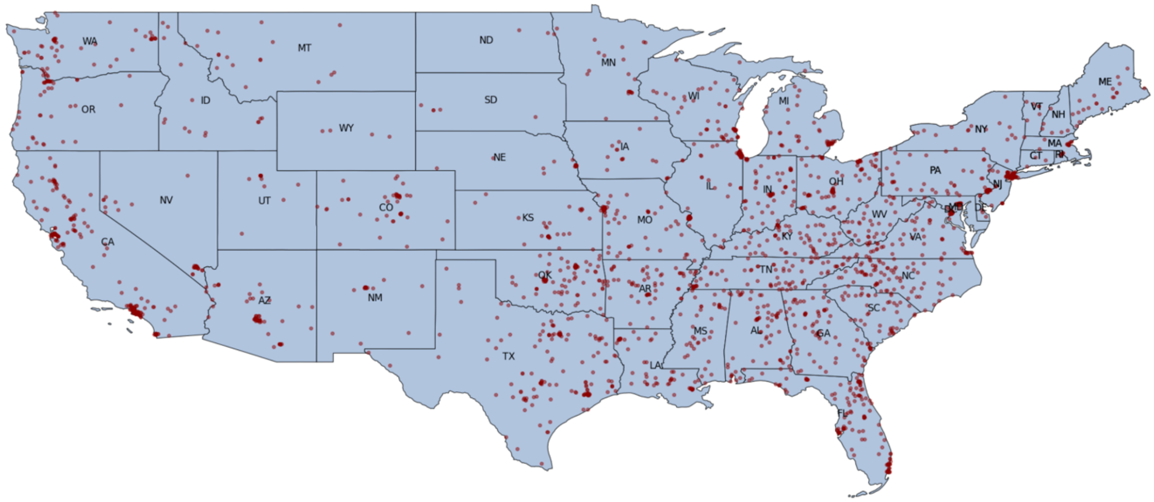

Figure 5: Distribution of the locations of police shooting

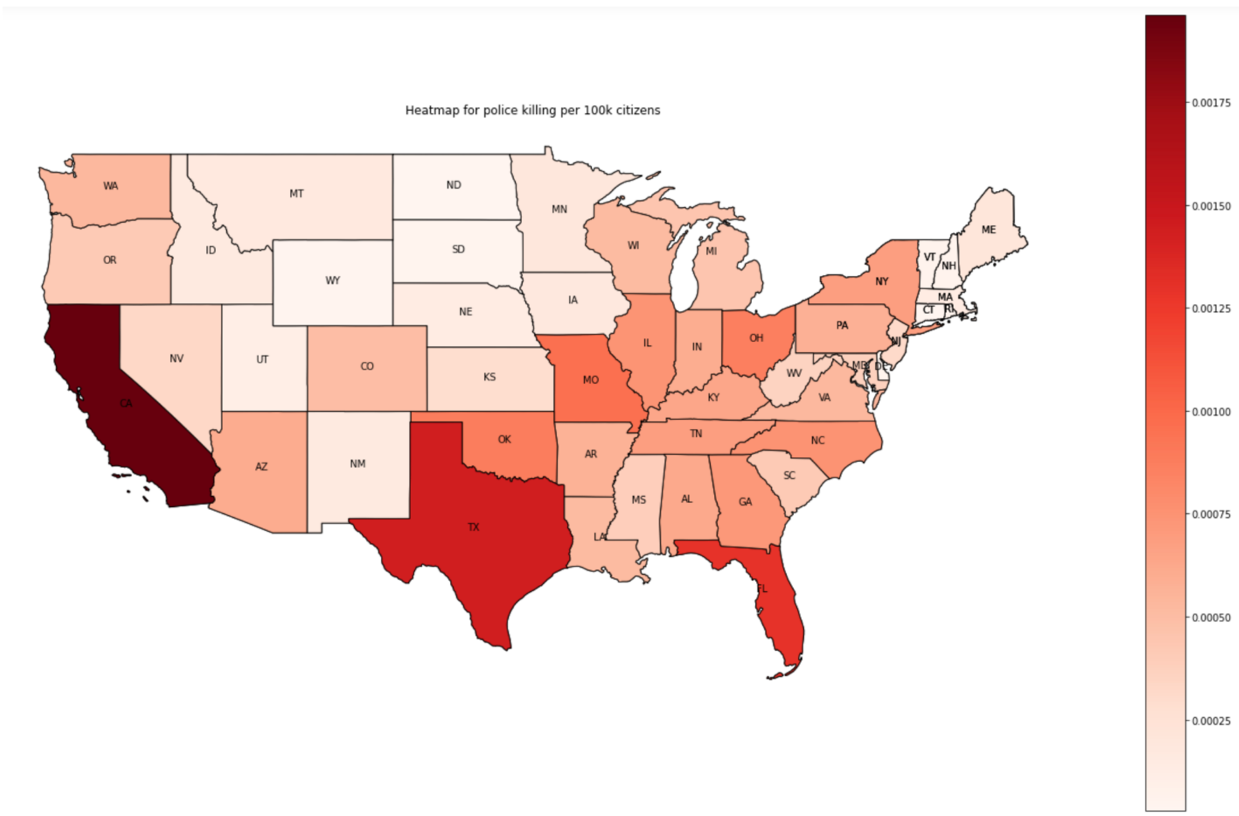

Figure 6: Heatmap for police killing per 100k citizens

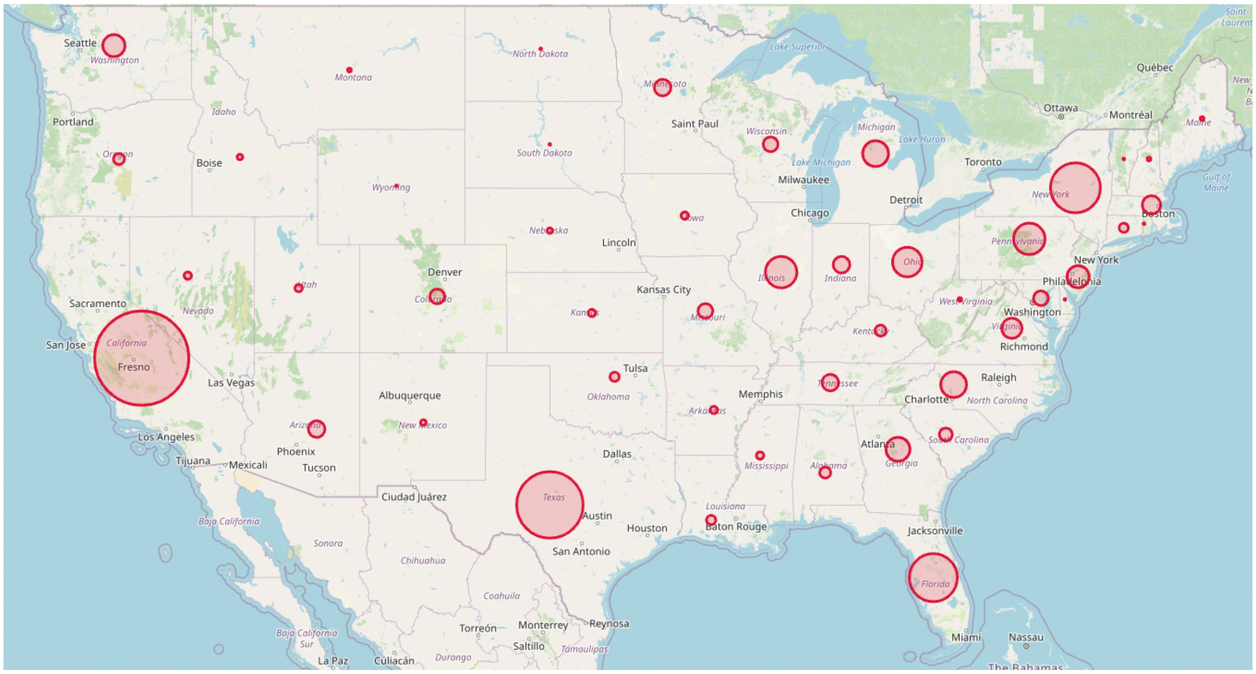

Figure 7: The size of the States in the USA

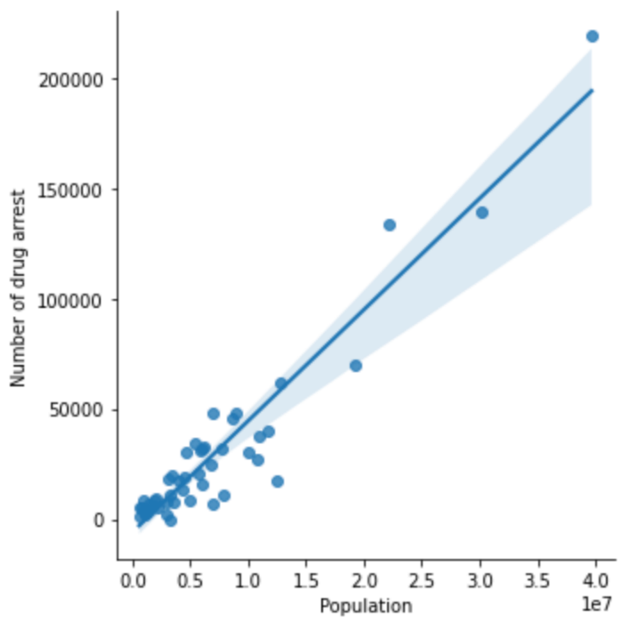

Figure 8: Correlation between number of drug arrest and population size of States in US