Racial Bias in Policing

Team Impressive Scientists | Ryan Hedges, Yueyi Sun, and Zejiang Shen

Racial Bias in Policing

Abstract

This project investigates the question of whether police exercise bias in their treatment of African Americans. Specifically, we attempt to answer the question of whether being an African American versus another race makes you more likely to be killed by the police under identical scenarios. The primary challenge that we faced and that historical analyses have faced is related to the available data, as the data that exists only captures people that were killed, and not the people that were not killed under similar circumstances. Through exploratory analyses and hypothesis testing, we convincingly exhibit how African Americans are killed at greater rates than people of other races. However, when we employ more comprehensive Generalized Linear Models and a Zero-Inflated Poisson Model to account for all confounding factors, we are only able to find mild evidence suggesting that explicitly being an African American makes you more likely to be killed than others.

1. Introduction

In recent years, there have been several high-profile deaths of African American citizens caused by the actions of police officers, such as Michael Brown in Ferguson, Missouri, Freddie Gray in Baltimore, Maryland, and Laquon McDonald in Chicago, Illinois [1] . The heightened press and attention of these events have spurred significant scrutiny and questioning over whether police exercise bias in their treatment of African Americans. While there are surely cases of racist and dishonest police officers who are more prone to act adversely against African Americans, is there actually a more widespread generalized bias against black people?

The great challenge for this question is the lack of data to understand the full picture. As a result, there are many conflicting as well as distrusted studies. A study by a Harvard economist found that African Americans are no more likely to be killed than people of other races during scenarios where lethal force may be deemed appropriate [2] . However, many critics take issue with the methods used in this study, specifically as it relates to the data under consideration, which is the "Achilles Heel" for analyses on this topic.

In the following investigative report, we apply a statistically rigorous approach to investigating whether police exercise bias in their treatment against African Americans.

2. Exploratory Data Analysis

In this project, we will use two datasets to analyze the potential bias against African Americans: a log of police killings and an American census dataset. In seciont 2.1 and 2.2, we will investigate and explore these two datasets. For each dataset, we explore important statistics and search for clues that could be useful for this investigation. Afterwards, we will introduce the merged dataset in section 2.3. Several visualizations will be presented to assist in understanding the datasets.

2.1 The Police Killing Data

The police killing data is sourced from *FiveThirtyEight*, a website focused on opinion poll analysis, politics, economics, and sports blogging [3] . *FiveThirtyEight* merged the original police killing information collected by the Guardian [4] with the census data from the American Community Survey. FiveThirtyEight hosts the dataset on Github, which is where our dataset was acquired.

It is important to mention that the police killing information that *FiveThirtyEight* collected represents the killings up until June 2, 2015 [5] . Therefore, less than half of 2015's killing data is included in this dataset. To be specific, our downloaded version contains 467 victims, but the total number of police killings in 2015 is 1,146 according to the Guardian [4] .

There are 467 rows and 33 features in this dataset, with one column containing the names of each victim. 34 rows contain a missing value. There are several categories of features types, including:

- Victims' demographic information, including their `name`, `age`, `gender`, and `race/ethinicity`.
- Information related to the killing incident, including:

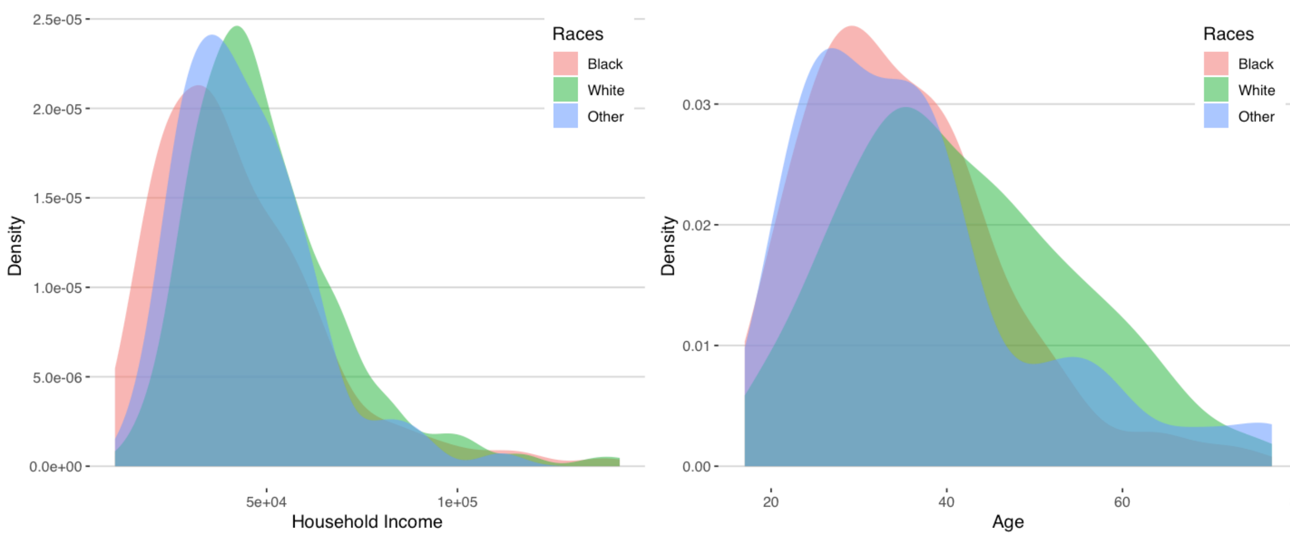
- Temporal information: `month`, `day`, and `year`
- Geolocation information: `streetaddress`, `city`, `state`, `latitude`, `longitude`, `state_fp`, etc.

- The Census information for the Tract where the victim inhabits, including race, income and education statistics.

More information about features in this dataset can be found in the Appendix A.2.

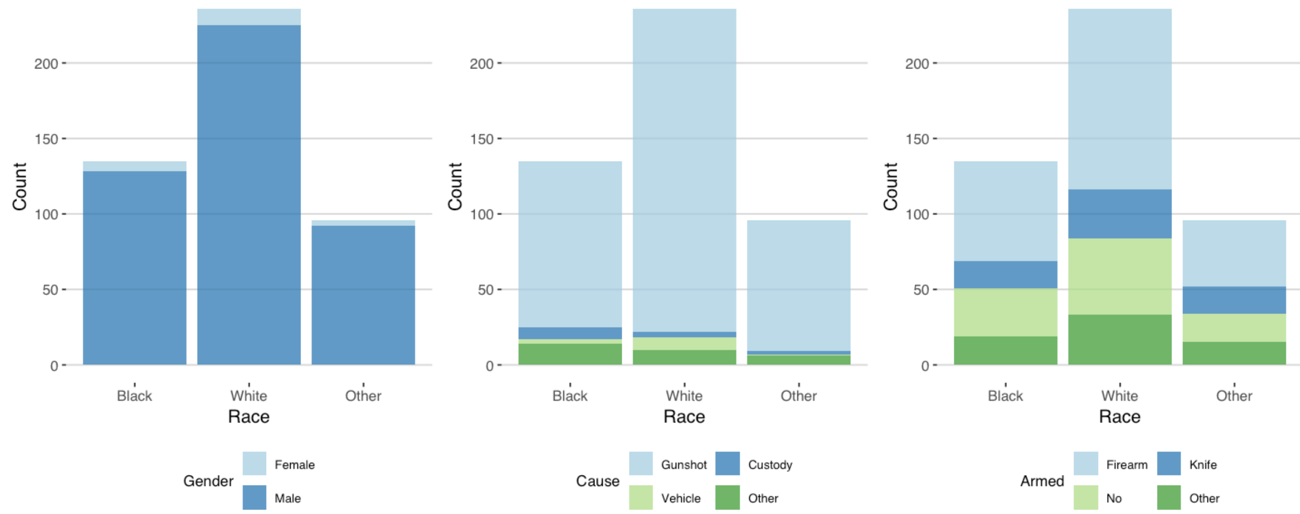
The `race/ethinicity` covariate originally had six categorical values. However, values different from `Black` and `White` only occupy 20.6% of the total samples. Therefore, we merged them into the same group which we labeled `Other`.

In Figure 1, we visualize the density of household income (`h_income`) and age (`age`) across different races of victims. In the left subgraph, we observe that the mean income for killed black people is lower than that of white people. In the right subgraph, the age distribution of African Americans and people with `Other` race categories is similar. Further, white victims tend to be older than non-white victims.



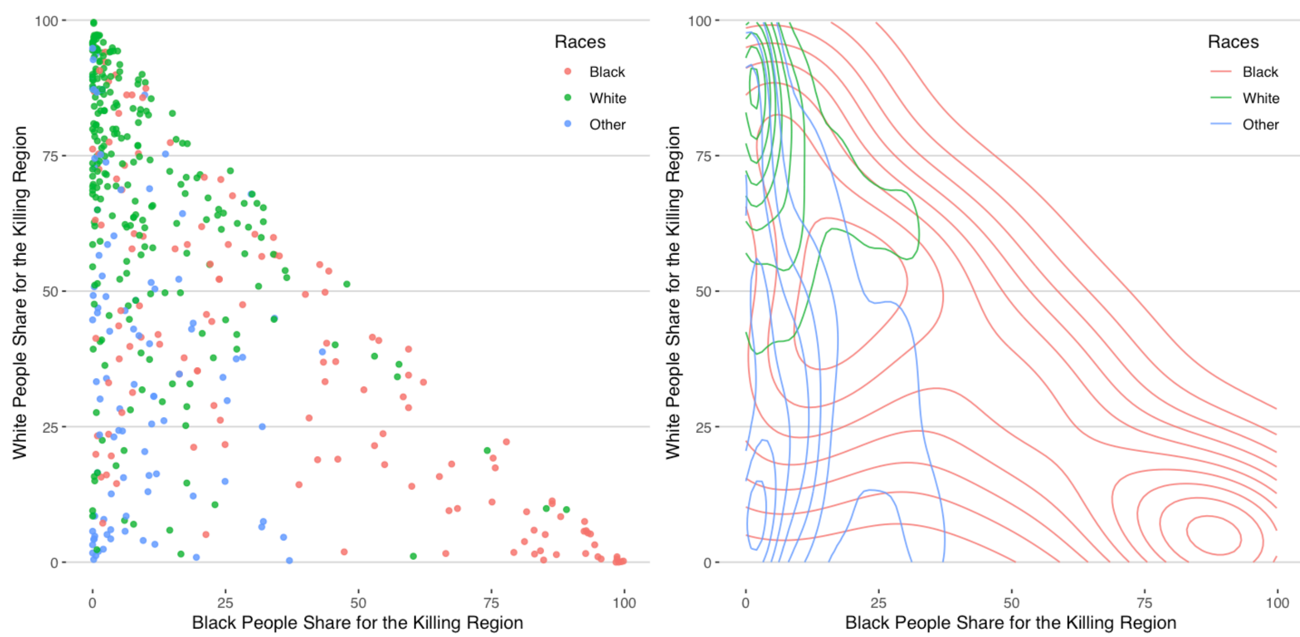
▲ Figure 1: The Household Income and Age Distribution among Different Races

In Figure 2, the `gender`, `Cause`, and `Armed` covariates are broken down by race. From the first two subplots, we conclude that the victims are mostly male, and the leading cause of a killing is a gunshot. In the third subplot, we identify three primary items that victims are armed with: Firearm, Knife, and None. (Note: the original number of categories of `race/ethinicity`, `Cause`, and `Armed` are 6, 5 and 8, respectively. We merged some categories to balance the population sizes in each category for the purposes of aiding the efficiency of the visualizations.)



▲ Figure 2: The `Gender`, `Cause`, and `Armed` Distributions for Victims among Different Races

Figure 3 displays an exceptionally interesting trend. In the left subplot, each data point represents an individual killed by the police while the x and y axis represent the `black_share` and `white_share` of their county population, respectively. The right subplot displays the same data in the format of a density contour plot. Intuitively, we should expect to observe a higher density of white people killed in the top left corner where counties are mostly white and a smaller density of white people killed in the bottom left and bottom right corner. Similarly, `black` and `other` races are expected to exhibit a higher density in the bottom left and bottom right corners, respectively. However, these expectations appear to only be true for the `white` and `other` races, and not for the `black` race. The killing pattern for African Americans appears to be less dependent on the underlying county population. This is best illustrated by the contour plot where the `black` race is distributed across the entire graph. While we are not yet accounting for any confounding factors at this stage in the analysis, these plots hint at the potential bias against African Americans.



▲ Figure 3: The distribution of Black People Share and White People Share for a Victim.

2.2 The Census Dataset

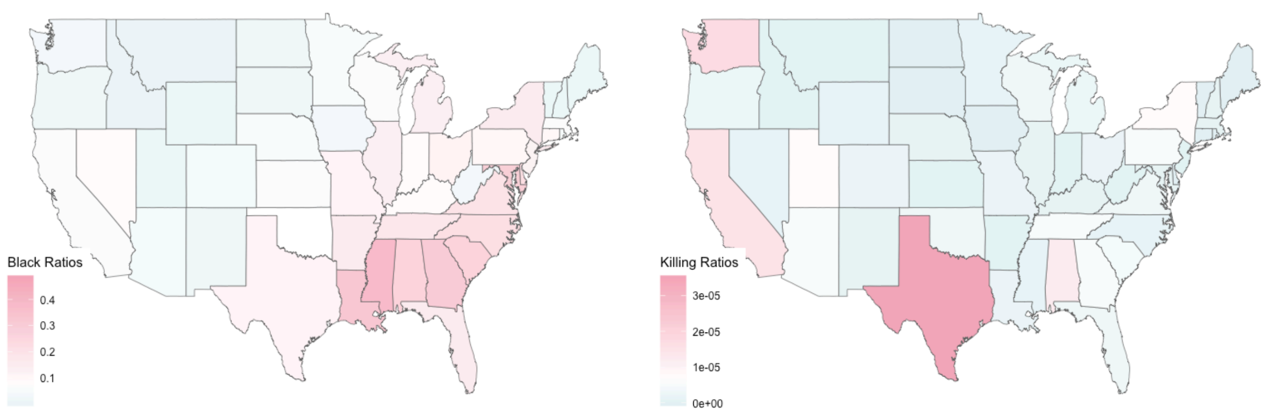
The second dataset to be used is the **2015 American Community Survey 5-year Census** estimates found here. The dataset contains 37 variables on 74,001 tracts (or towns) in the United States. The information contained in this table includes population, gender, race, income, and employment ratios and counts.

For the purposes of our study, the census dataset is used to generate two new aggregated datasets, one summarized to the state level and a second summarized to the county level. For features like `White_pop`, we calculate the overall ratio by $(\sum \text{White Population}) / (\sum \text{Total Population})$for each state and county. After such preprocessing, the resulting aggregated tables include 52 rows for each state and roughly 2,000 rows representing each county.

2.3 The Merged Dataset

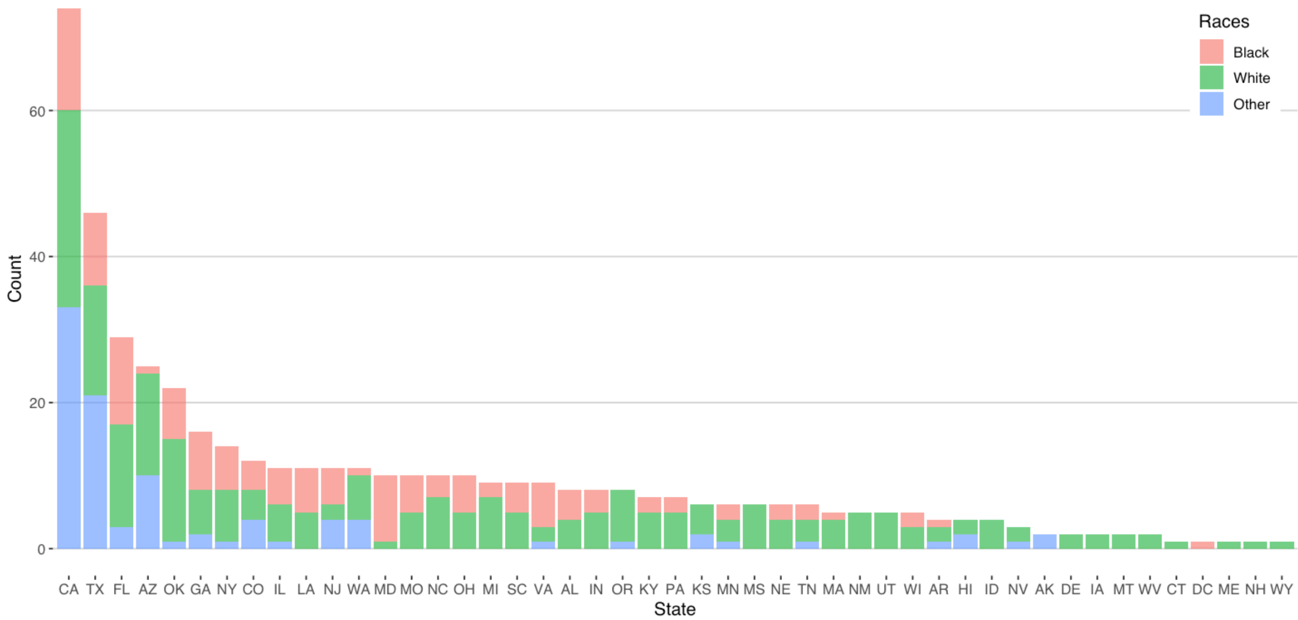
The final datasets to be used for the modeling will include elements from both of the previously introduced datasets. The summarized state and county-level census datasets will be used as the feature sets. The killings dataset is then used to generate the target variables, which is the total counts of people killed at both the state and county level. These kill counts are then merged onto the state and county-level tables, respectively.

Figure 4 helps us visualize important features as they relate to the geography of the United States. The left subplot displays how African American population proportions are larger on the eastern half of the country, and especially in the deep south. The right subplot captures police killings per capita. The killing ratio is the number of people killed divided by the total population in that state. This plot is characterized by a few states that jump out signifantly ahead of the rest of the country, notably Texas, California, Washington and Alabama.



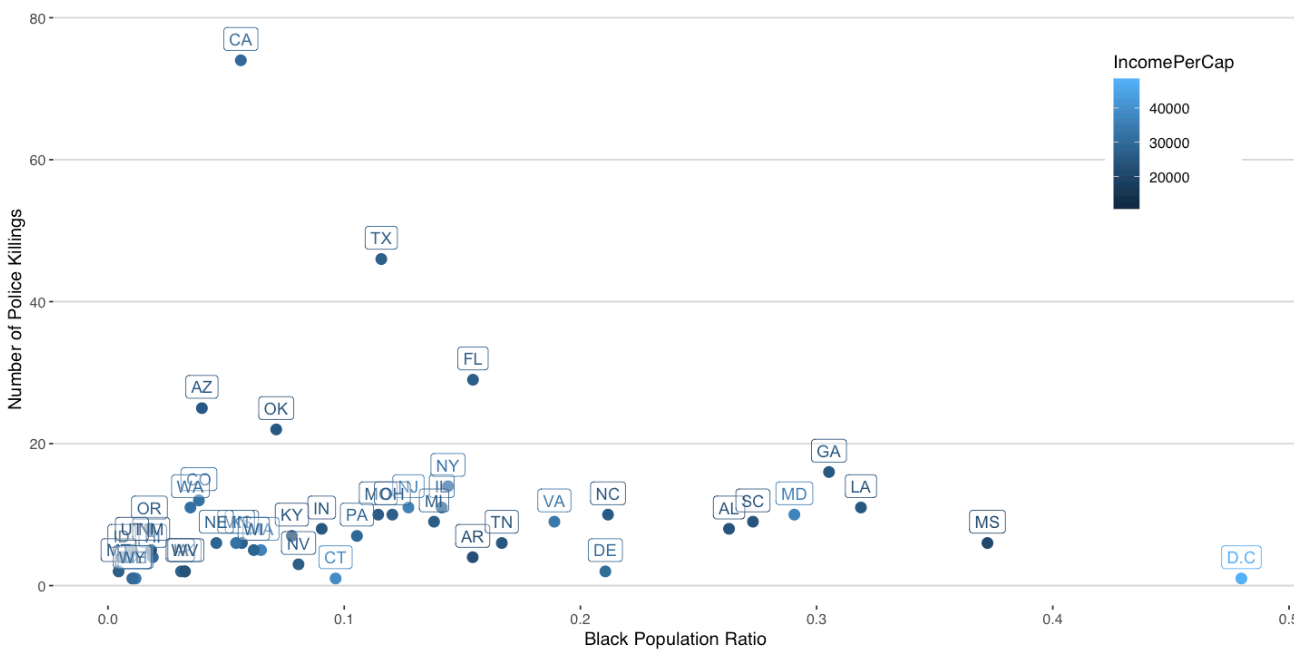
▲ Figure 4: The Black People Ratio and Police Killing Counts for each State

Figure 5 conveys the total kill count by state, broken out by the race of the victims. The state kill rankings are roughly in alignment with the general state population sizes, with the exceptions for Arizona and Oklahoma, which each have disproportionately large number of police killings. Oppositely, Pennsylvania has a disproportionately small number of police killings. In regards to the race distribution of people killed within each state, there are a few observations. The 'Other' category is signifantly present in only California, Texas, and Arizona, likely due to their close proximity to the southern border, and thus larger Latino populations. One state with a massively disproportionate number of black killings is Maryland.



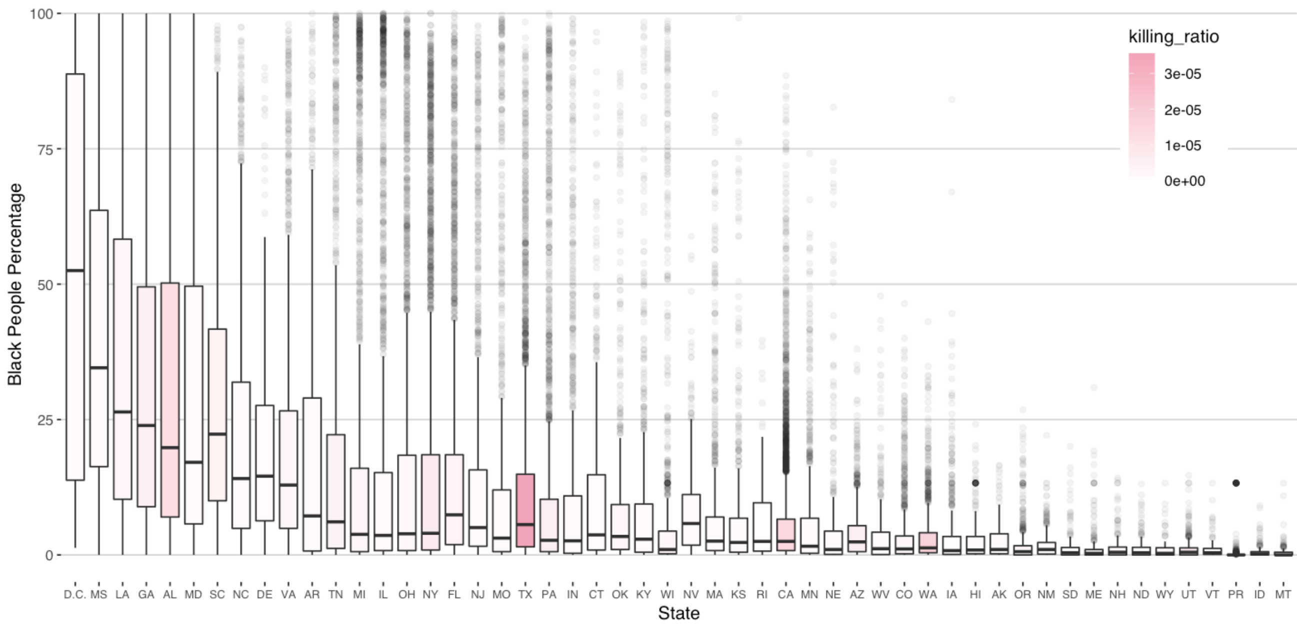
▲ Figure 5: The Distribution for Police Killing Counts for Differnet States

In Figure 6, we examine the correlation between police killings (y axis) and the black population ratio by state (x axis). The darkness of the blue shade of each point indicates the per capita income for the corresponding state. We do not observe a clear positive correlation between the black population ratio and police killings. CA and D.C. appear to be potential outliers, with exceptionally large police killing numbers and high African American population ratios, respectively. We will carefully handle these samples during our modeling process. In regards to Income Per Capita, D.C. appears to potentially be an outlier, however there does not appear to be an overall correlation between income and killing rate.



▲ Figure 6: The Relationship between Killing Counts and Black Population Ratio for each State

Figure 7 also presents the relationship between the black population percentage and overall police killing frequency. If a bias is present in the police's handling of African Americans, a larger black population ratio should yield a higher overall killing ratio (when accounting for other factors, which is not captured here). However, the color scheme of the barplot, which denotes the killing ratio, suggests that total killing frequency isn't notably correlated with black population percentages.



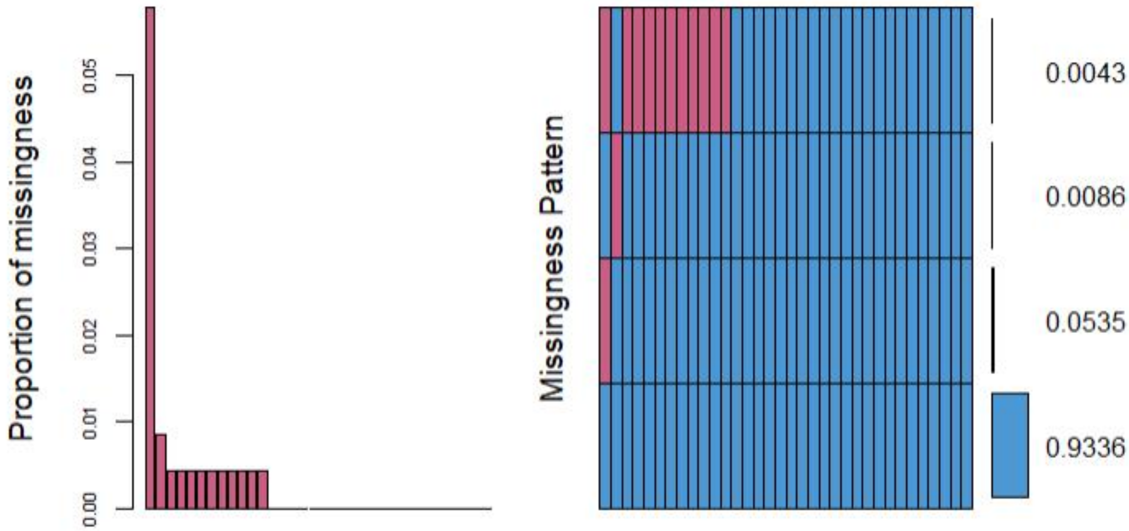
▲ Figure 7: The Black Ratio Boxplot for each State and the Killing Ratio

2.4 Missing Data Handling

For the Police Killings dataset, there are 34 observations with missing variables.

Missing percentage: (largest 5 variables)

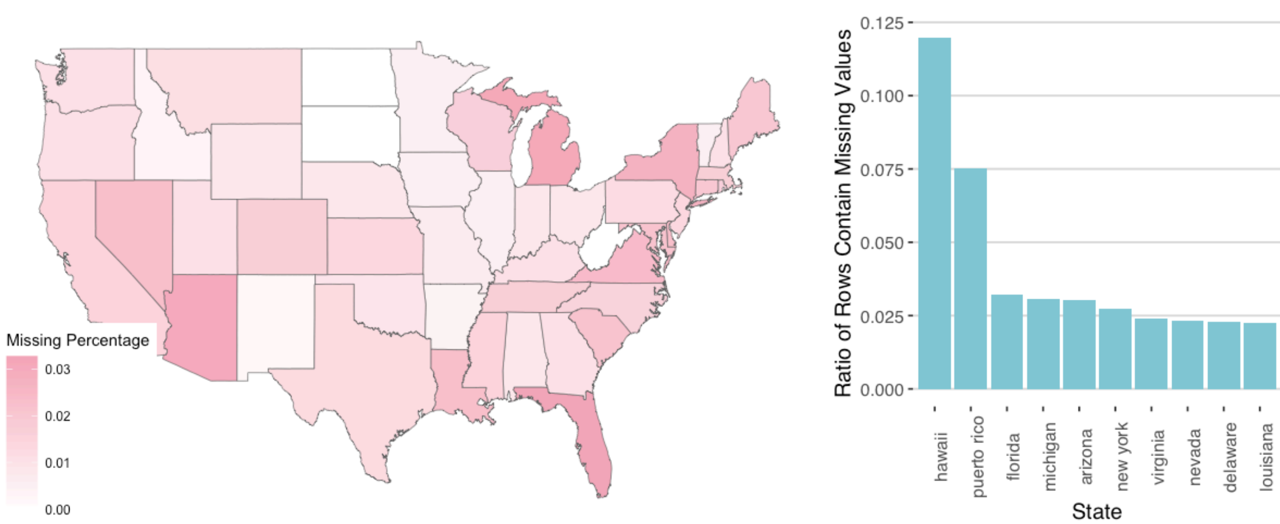
Variable	Percentage
county_bucket	5.78%
streetaddress	0.86%
share_white	0.43%
share_black	0.43%
share_hispanic	0.43%



▲ Figure 8: The Pattern of Missingness

From the table and Figure 8 above, we exhibit how there are many observations containing missing variables in the killings dataset. However, there are no missing values in the state or county fields, which are the only fields that are used for the purpose of joining the kill counts to our census data. Therefore, there is no need to impute the missing values in the police killing dataset.

For the census data, there are 1274 observations with missing variables. As the proportion of missing data is small, we impute the mean values for all missing data. To ensure our analysis is valid, we need to check that there is no significant bias in the dataset we use. Therefore, we need to confirm that missingness occurs randomly for each state. Figure 9 displays the plots of the proportion of randomness for each state.



▲ Figure 9: The Missingness for Each State

Hawaii has the largest percent of missingness: 11.97%, and Puerto Rico has the second largest percent of missingness: 7.51%. All the other states have similar portions of missingess, with ratios below 3.3%. Therefore, it is safe to impute the missing data using the mean of the variables.

3. Assumptions

Before we dive into the statistical testing and modeling section, it is important to highlight a few assumptions that are being made in order for our methods to be valid. These assumptions all relate to the limitations of the data we are using.

The first assumption is that the proportion of African Americans present in each state and each county is also representative of the proportion that are confronted, or interacted with, by police. In the following modeling section, the African American population percentages of each state and county will be used as our covariate of interest. The question we are asking is: **Does a greater portion of African Americans in an area lead to more police killings, after controlling for other variables?** We are making the assumption that the population race distribution is the same as the race distribution that the police will have interactions with, when holding all other variables constant.

The second assumption is that states and counties are independent of one another. In reality, we know this is probably not the case, as bordering states and counties will likely have similar situations that cannot be captured by our data, and thus will likely have correlated residuals.

The last assumption is that the **police killings that took place between January and May of 2015, as captured by our data, are independent of the killings that took place during the remaining months of 2015**.

4. Hypothesis Testing

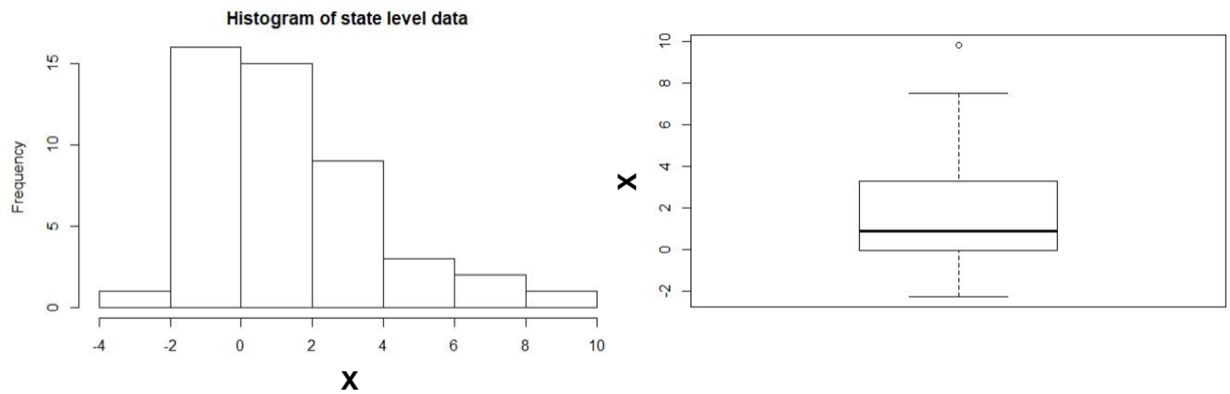
In this section, we will run one sample t-tests to detect a difference in the killing rate of black Americans compared to non-black Americans.

We use a one-sample t-test to check whether the mean of the differences between the number of actual black people being killed by police and the expected number of black people being killed (defined by the total number of people being killed by police multiplied by the black proportion in each state) are greater than 0.

Hypothesis: $H_0 : \mu_1 = 0$ vs $H_1 : \mu_1 > 0$ where:

- $\mu_1 = \text{Mean}(N_b - N \times P_b)$,
- N_b is the number of black people being killed by police,
- N: the number of people being killed by police,
- P_b: the proportion of black people in the state.

The sample data is numeric and continuous. The normality can be checked by the histogram in Figure 10. From the boxplot in Figure 10, the only outlier does not raise significant concern. Therefore, the assumptions of the hypothesis test are satisfied.



$$X = N_b - N \times P_b$$

▲ Figure 10: The Histogram and Boxplot of the sample data

From the one sample t-test result, the p-value is 4.939×10^{-6} and the confidence interval is $(1.128, \infty)$ which indicates there is bias against African Americans in police shootings, without accounting for other variables.

5. Modeling

For this analysis, our objective is to assess whether being an African American increases one's likelihood of being killed by the police. The great challenge of this investigation is that there are no data on the individuals that were not killed and there are so many confounding factors that correlate with both race as well as one's likelihood to be killed. Therefore, the purpose of the modeling section is to account for these confounding factors and determine if race impacts the probability of being killed, when holding all other variables constant.

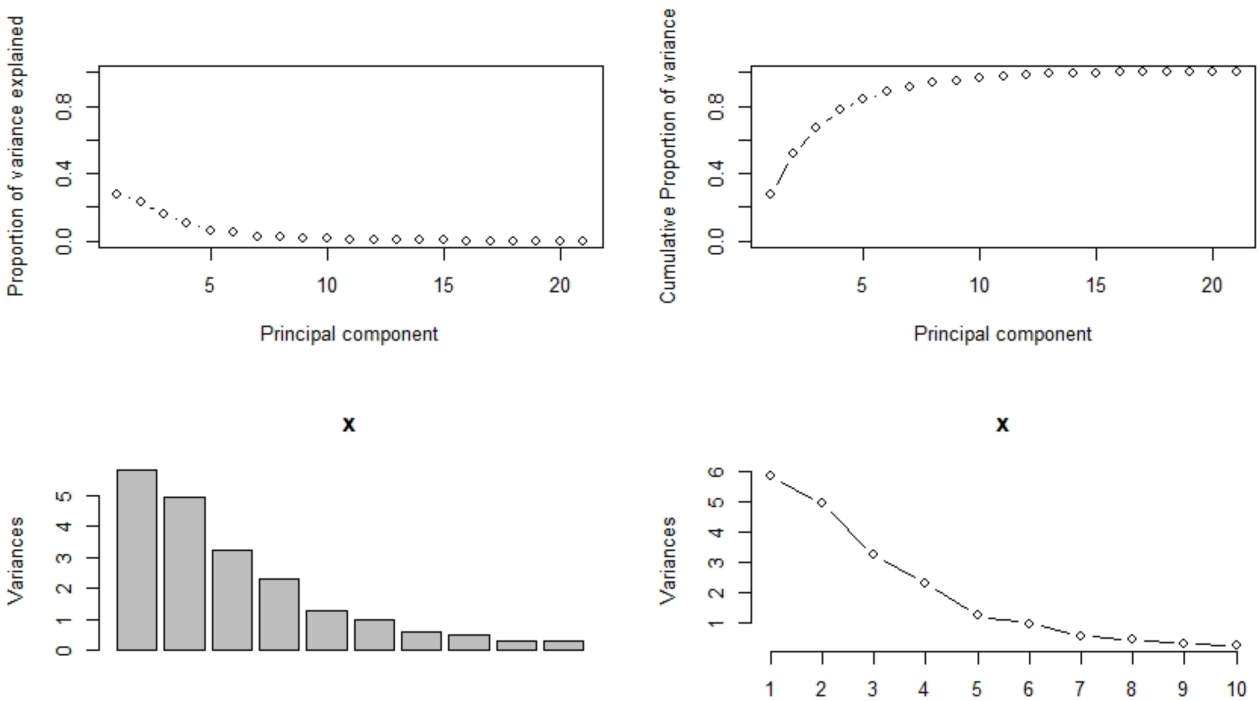
Our null hypothesis will be that being an African American does not directly impact the likelihood of being killed. That being said, based on our intuition and findings from both our exploratory analysis and our initial hypothesis testing, we believe that we will find enough evidence to reject the null hypothesis and claim that being an African American does increase the likelihood of being killed.

Three different modeling approaches will be used to investigate this question and they will all exploit the following logic: When predicting the likelihood of being killed, if the inclusion of an **African American** covariate is significant and improves the predictive power of an already comprehensive model, this exposes a bias indicating that being African American affects your probability of being killed. If the African American covariate is insignificant to the model, then this suggests that there is no bias present.

5.1. The Poisson Model (State Level)

The first modeling approach to be explored is to apply a **Poisson Regression** model to predict the number of total people killed in a given **state**.

Given that this model uses data that is aggregated to the state level, the resulting dataset only contains 52 observations. Since there are 28 total potential covariates, there are very few degrees of freedom in a full model, which could introduce instability. To account for this, before the fitting of our model, we reduced the dimensions of our feature space by conducting **Principal Component Analysis**. The PCA was conducted while excluding all race-related variables. This process resulted in 12 new features, while retaining the majority of our information. Figure 11 conveys how the first 12 principle components explain more than 98 percent of the variance in the data.



▲ Figure 11 The PCA results

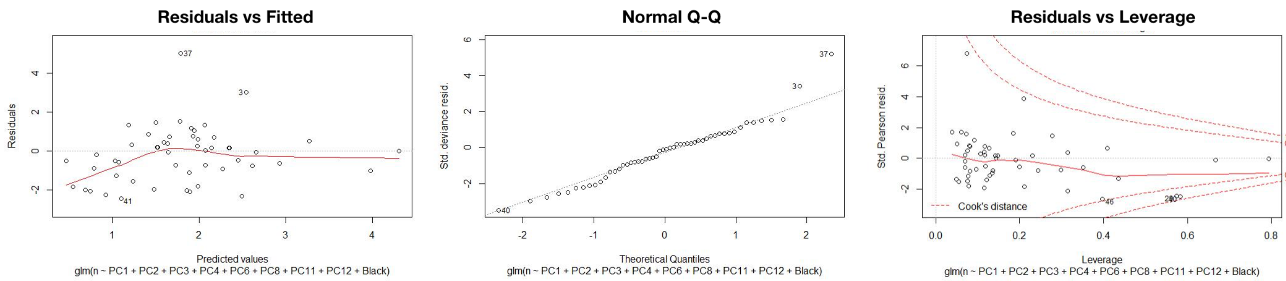
A stepwise forward and backward selection process is then used to construct an optimal model using the 12 principal components, while still ignoring all the race variables. Afterwards, this same model is refit, except now the **African American Population Ratio** covariate is added into the model. In doing so, the model's predictive power improves as measured by AIC, and the P-value on the African American ratio is significant to the 90% confidence level. Note that since PCA was conducted on the original feature matrix, the model coefficients are uninterpretable. This result loosely supports the idea that police exhibit bias against African Americans. However, this result is not convincing, especially since the model is based off of only 52 state observations, which could be unstable, as a single state can have an undesireably large significance on the results.

Significance of Covariates:

Covariate	P-Value
Intercept	< 2e-16
PC1	0.00394
PC2	0.61070
PC3	0.00103
PC4	< 2e-16
PC6	4.39e-14
PC8	0.05056
PC11	0.00909

PC12	0.00261
Black	0.06314

Diagnostics: In the residuals versus fitted plot, we observe a slightly quadratic pattern. This is a reasonable result due to the small sample size. Also, this pattern is not of great concern because our aim is to detect the possible bias against African Americans, as opposed to finding the most predictive model. The assumption of normality appears to be confirmed by the QQPlot. Lastly, there are two points that appear to be mildly influential, however none of them have a Cook's distance greater than 1.



▲ Figure 12: Diagonostic Plots for Poisson Regression

The following two models address some of the limitations of this state-aggregated Poisson Regression.

5.2. The Logistic Model (County Level)

There are a couple concerns with using state-aggregated data to fit a model, as was done in the previous section. One of them is regarding the small sample size of 52 state-observations. This small sample size exposes the model to significant vulnerability to the influence of any individual state. Additionally, when aggregating data to the state level, many potentially helpful data points that are representing smaller towns and cities get smoothed over and their usefulness is washed out when they are averaged in with an entire state. Therefore, we are potentially suffocating helpful datapoints and relationships during this aggregation process.

To address these concerns, a second modeling approach is attempted where the data is aggregated at the county level, as opposed to the state level. The benefit is that there are now almost 2,000 observations. When aggregating at the county level, the majority of the counties have zero killings. Therefore, for this model, a logistic regression is used, and the target variable is now binary. We set the target to be zero if there are no killings in the county and to one if there is at least one killing. Because we have enough observations at the county level, we do not need to use PCA to reduce the dimensions of the feature space.

Similar to the first model, a forward and backward stepwise regression approach is used where the race variables are excluded. Once this model is fit, the **African American Population Ratio** covariate is introduced. This method led to no improvement in the predictability of the overall model, as measured by AIC, and the P-value of the Black Indicator indicates insignificance. When examining the other covariates included in the model, many of the original variables that would have been expected to be significant, like Unemployment and Poverty, are not present. The one expected variable that has an unsurprising coefficient is the **% Citizen** covariate, which indicates that counties with a greater portion of non-citizens are more likely to experience a police killing, when holding all other variables constant.

Covariate Results:

Covariate	Coefficient	P-Value
Intercept	-2.61	0.1342
TotalPop	38.38	< 2e-16
SelfEmployed	-13.13	0.0007
Production	-6.67	0.0003
Citizen	-3.32	0.0313
Drive	4.21	0.00384
Carpool	8.21	0.0015
Black	0.42	0.4853

We checked the mean of the model residuals with a one sample t-test. The null hypothesis is that the true mean of residuals is equal to zero and the alternative hypothesis is that the true mean is not equal to zero. There is no need to check the assumptions of the t-test as the sample size is large enough this time. The p-value is equal to 0.003 with a 95-percent confidence interval of the mean between -0.592 and -0.120, which indicates the residuals are not centered around 0. In other words, the logistic model can not predict the target variable while meeting the required model assumptions.

While this logistic model suggests no racial bias against African Americans, the upcoming model introduces an additional enhancement, which allows us to capture the total killing counts by county, as opposed to a binary target variable.

5.3. The Zero-Inflated Poisson Model (County Level)

As previously mentioned, the solution to the small sample size in the original state-aggregated Poisson model was to aggregate to the county level instead of the state level. In doing so, since the majority of the observations now had no killings, the Poisson distribution was less applicable, and therefore a logistic model was explored. However, in using a logistic regression, we concede the ability to include killing counts in the analysis. The **Zero-Inflated Poisson Regression** model is the answer to this concern. The Zero-Inflated Poisson Regression model is made to predict a target variable that includes count numbers, but also carries an excess proportion of zeros. This model can be thought of as a composition of a Poisson Regression and a Logistic Regression which interact to model and represent the counts and the excess zeros, respectively.

Similar to the prior two models, a backwards and forwards stepwise approach was used to construct optimal models while ignoring the race variables. Once this model is complete, the **African American Population Ratio** covariate is introduced. In doing so, the predictability of the model, as measured by AIC, improves by a mild degree, and the P-value for the African American indicator is equal to 0.08. Interestingly, a few of the covariates that are present in our model have coefficients indicating impacts of the opposite direction of what was expected, notably Poverty Rate, Citizenship Rate, and Office Employment Rate. This could be due to multicollinearity within the model, or potentially a nuance related to the Zero-Inflated Poisson Regression model. That being said, the sign of the Black Ratio covariate is as expected.

Count model coefficients (poisson with log link):

Covariate	Coefficient	P-Value
Intercept	-8.2392	< 2e-16

TotalPop	3.1239	9.53e-15
Professional	6.0421	2.54e-09
Service	7.4529	0.00048
Office	13.1554	5.09e-07
Black	0.9083	0.07708

Zero-Inflation model coefficients (binomial with logit link):

Covariate	Coefficient	P-Value
Intercept	-8.2392	1.81e-06
Citizen	3.1239	1.54e-05
Pov	-13.13	0.00642
SelfEmployed	-6.67	2.97e-05

Lastly, as was conducted with the logistic model, we run a one-sample t-test to check that the mean of the residuals is centered around zero. The p-value of the t-test is 0.967 and the 95 percent confidence interval of the residuals is (-0.014,0.015). The confidence interval and the p-value indicate that the mean of the residuals is centered around zero. This result exhibits how the Zero-Inflated Poisson Regression is more robust and can provide a more trustable conclusion than the logistic model.

This modeling result provides mild support of the idea that police exercise bias against African Americans.

6. Limitations

There were a few limitations to this study that mostly relate to the available data, which also likely contributed to our mostly indeterminate results.

The largest limitation is that of the individual-level data that we had, it only captured people that were killed. In order to best understand the difference in how African Americans versus other races are treated by police, it would be ideal if data was available that captured individuals who were not killed. For example, if a dataset existed at the individual-level containing information on every police encounter in 2015, we would be able to construct a significantly more statistically sound test.

An additional data limitation is that our dataset only contained police killings representing the first five months of 2015. This presents a challenge in that our sample size is significantly smaller than it should be and that we are forced to make the assumption that the killing data distributions representing these first five months are also representative of what happened in the final seven months. In other words, we are making the assumption that the two time periods are independent of one another.

One final data limitation worth noting is that during exploratory analysis, it was discovered that the duplicate variables that existed across both datasets were often significantly misaligned. This raises a concern about the integrity or the accuracy of our data.

7. Conclusion

In this investigative statistical report, we used census data and a dataset capturing the individual police killings of 2015 to determine whether police exhibit bias in how they treat African Americans compared to individuals of other races. A major factor to this analysis was the data that was not available, and the limitations that this presented.

To account for our data limitations, we employed a modeling technique where we predicted the total number of people to be killed, with and without a race factor, to determine if race was impactful after accounting for all other variables. We tested three different modeling approaches using this logic, and two out of the three models resulted in only a mildly significant result for the **African American** covariate, while the 3rd model exhibited insignificance. Therefore, while there exists mild evidence that police exhibit bias against African Americans, we cannot definitively reject the null hypothesis that being an African American directly impacts the likelihood that you are killed. We are unable to confirm our intuition and suspicions that were raised during our exploratory analysis and hypothesis testing. More killings data, along with a more comprehensive dataset on individuals who were not killed, would be very helpful in arriving on a more definitive conclusion in the future.

A. Appendix

A.1 Code

We have open sourced our code in Github. The link to our Github Repository is https://github.com/2020proj/police-killing-analysis.

A.2 Feature Description

We list the name and meanings for the features in the used dataset.

Table A.1 The Features and Their Meanings of the Police Killing Dataset [5]

Feature Name	Description	Source
`name`	Name of deceased	Guardian
`age`	Age of deceased	Guardian
`gender`	Gender of deceased	Guardian
`raceethnicity`	Race/ethnicity of deceased	Guardian
`month`	Month of killing	Guardian
`day`	Day of incident	Guardian
`year`	Year of incident	Guardian
`streetaddress`	Address/intersection where incident occurred	Guardian
`city`	City where incident occurred	Guardian
`state`	State where incident occurred	Guardian
`latitude`	Latitude, geocoded from address	
`longitude`	Longitude, geocoded from address	
`state_fp`	State FIPS code	Census

county_fp	County FIPS code	Census
tract_ce	Tract ID code	Census
geo_id	Combined tract ID code	
county_id	Combined county ID code	
namelsad	Tract description	Census
lawenforcementagency	Agency involved in incident	Guardian
cause	Cause of death	Guardian
armed	How/whether deceased was armed	Guardian
pop	Tract population	Census
share_white	Share of pop that is non-Hispanic white	Census
share_bloack	Share of pop that is black (alone, not in combination)	Census
share_hispanic	Share of pop that is Hispanic/Latino (any race)	Census
p_income	Tract-level median personal income	Census
h_income	Tract-level median household income	Census
county_income	County-level median household income	Census
comp_income	h_income / county_income	Calculated from Census
county_bucket	Household income, quintile within county	Calculated from Census
nat_bucket	Household income, quintile nationally	Calculated from Census
pov	Tract-level poverty rate (official)	Census
urate	Tract-level unemployment rate	Calculated from Census
college	Share of 25+ pop with BA or higher	Calculated from Census

Table A.2 The Features and Their Meanings of the 2015 Census Dataset [6]

Feature	Description
CensusTract	Census tract ID
State	State, DC, or Puerto Rico
County	County or county equivalent
TotalPop	Total population
Men	Number of men
Women	Number of women
Hispanic	% of population that is Hispanic/Latino
White	% of population that is white
Black	% of population that is black
Native	% of population that is Native American or Native Alaskan
Asian	% of population that is Asian
Pacific	% of population that is Native Hawaiian or Pacific Islander
Citizen	Number of citizens
Income	Median household income ($)
IncomeErr	Median household income error ($)

IncomePerCap	Income per capita ($)
IncomePerCapErr	Income per capita error ($)
Poverty	% under poverty level
ChildPoverty	% of children under poverty level
Professional	% employed in management, business, science, and arts
Service	% employed in service jobs
Office	% employed in sales and office jobs
Construction	% employed in natural resources, construction, and maintenance
Production	% employed in production, transportation, and material movement
Drive	% commuting alone in a car, van, or truck
Carpool	% carpooling in a car, van, or truck
Transit	% commuting on public transportation
Walk	% walking to work
OtherTransp	% commuting via other means
WorkAtHome	% working at home
MeanCommuteMean commute time (minutes	MeanCommuteMean commute time (minutes)
EmployedNumber of employed (16+)	EmployedNumber of employed (16+)
PrivateWork	% employed in private industry
PublicWork	% employed in public jobs
SelfEmployed	% self-employed
FamilyWork	% in unpaid family work
Unemployment	Unemployment rate (%)

B. Reference

1. Wbur. (2016, July 11). After 9 High-Profile Police-Involved Deaths Of African-Americans, What Happened To The Officers? Retrieved from https://www.wbur.org/hereandnow/2016/07/11/america-police-shooting-timeline↩

2. NBER. (2018, January). An Empirical Analysis of Racial Differences in Police Use of Force https://www.nber.org/papers/w22399↩

3. Wikipedia, FiveThirtyEight, https://en.wikipedia.org/wiki/FiveThirtyEight. ↩

4. Jon SwainePeople killed by police in the US, https://www.theguardian.com/us-news/ng-interactive/2015/jun/01/the-counted-police-killings-us-database↩↩

5. Fivethirtyeight. (2018, February 09). Fivethirtyeight/data. Retrieved from https://github.com/fivethirtyeight/data/tree/master/police-killings↩↩

6. MuonNeutrino. (2019, March 03). US Census Demographic Data. Retrieved from https://www.kaggle.com/muonneutrino/us-census-demographic-data↩