Police-Killings

Jiayi Zhou

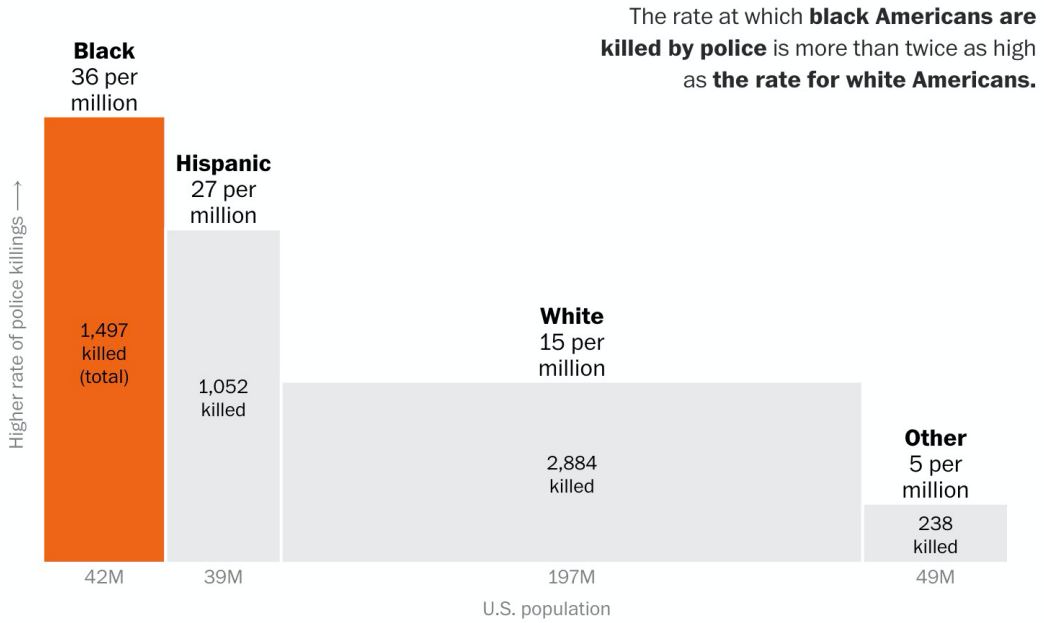
Introduction

Due to the laws of the United States, residents of all states can buy guns. Therefore, the police need to wear guns when completing daily patrols and tasks. But in recent years, more and more news reports that police shot and killed citizens. Police kill more than a thousand civilians each year in the United States, a much higher death rate than occurs in any other developed nation. It is even said that the police have abused violence and can shoot and kill when they receive threats. This project is a series of analysis on police killings.

Related Work

In Washington post, it was reported that 982 people have been shot and killed by police in the past year. Every year, the number of people killed by gunshots has an upward trend. Among the dead citizens were armed with guns, some only brought some self-defense tools, and some did not bring any lethal weapons. They all come from different races. At the same time, the report also shows a histogram, which clearly shows that the proportion of blacks shot by the police is the highest. It is even about 50% higher than that of whites.

On the Kaggle website, a person named notebook13d55efb9b analyzed the relationship between police shooting and race. In his research, he pointed out that black people are 9.5% more likely to be shot and killed, whereas white people is 9.6% less likely to be killed (The other races remain the same whether armed or unarmed).



Although each work shows that police-killings are related to race, but this police-killings is really only more related to race, or is it affected by other factors, such as the economic situation in the region. This project is to further understand the various factors that affect the number of police-killings.

Dataset

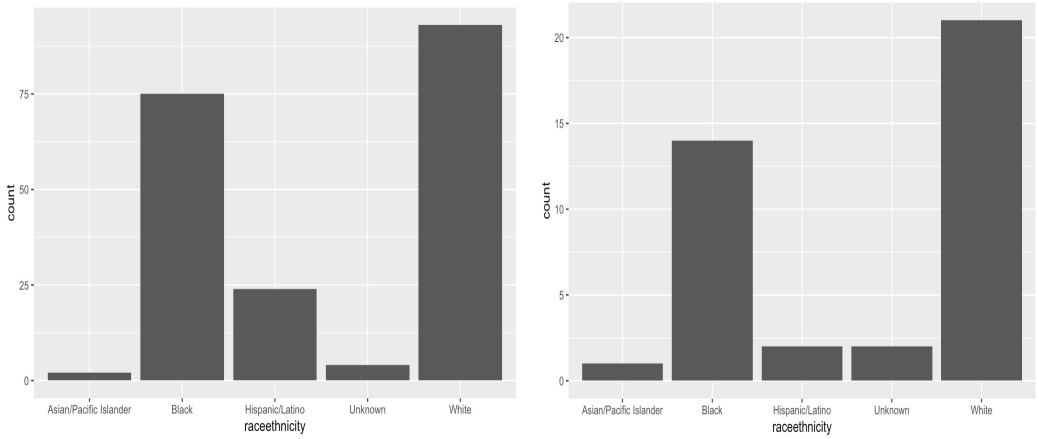
The main data source for this study is Where Police Have Killed Americans in 2015. This data includes 34 different details about the deceased. For the initial prepossessing, I removed the year, month, day, street address, lawenforcementagency, city, latitude and longitude which are not related with my analysis. We set 52 states to 4 different regions which is Northeast Region, Midwest Region, South Region, West Region. Take a rough look at the overall situation.

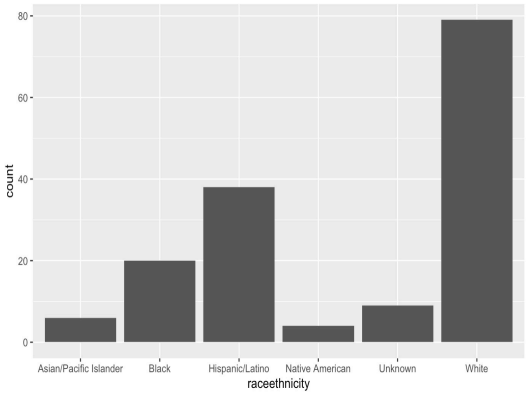
region <chr>	n <int>
South	198
West	156
MidWest	73
NorthWest	40

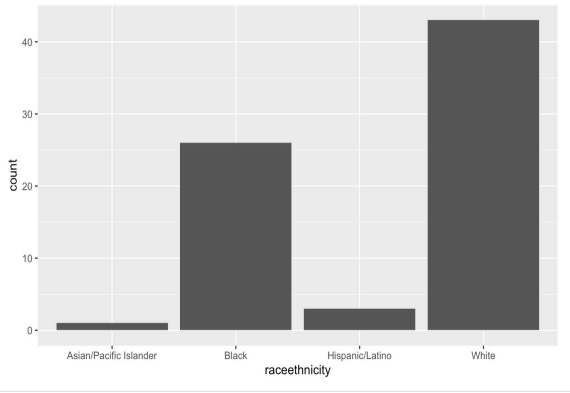
We can clearly see from the table that the South region has more police-killings, which is even about 5 times that of the Northwest region, which has the least number. From this table, we can see that there should be some factors that affect the number of different regions.

Experiments & Results

In response to the above data, I conducted a more in-depth analysis. As can be seen from the histogram in the figure below, in fact, the number of citizens killed by the police in each region is about the same according to race classification, and it does not show that regions with more police-killings have more blacks injured., So we have no way to make a conclusion. Different races can affect the value of police-killings.







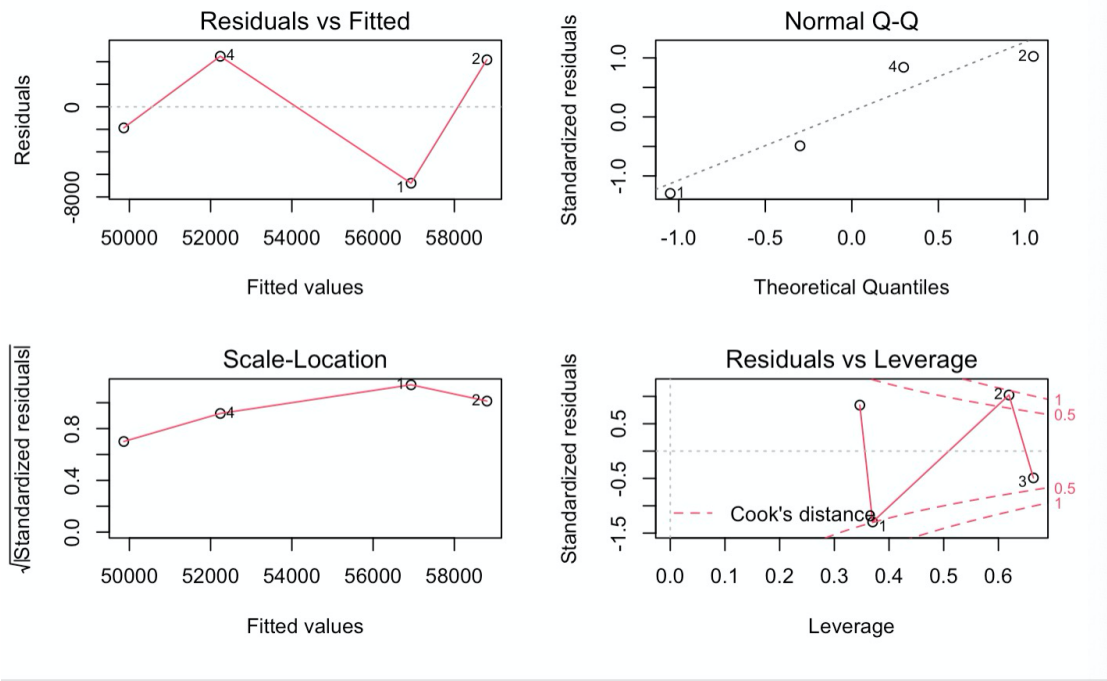
Then I analyzed an economic situation in each region, mainly looking at an average of County-level median household income in each region. The value that can more reflect the economic situation of a region is household income. Here I used a simple linear regression model to see the relationship between the two.

```
Call:
lm(formula = mean_county_income ~ number, data = df)

Residuals:
    1     2     3     4
-6784  4176 -1875  4484

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 61065.33    6935.64   8.805   0.0127 *
number        -56.58      52.25  -1.083   0.3921
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6599 on 2 degrees of freedom
Multiple R-squared:  0.3696,    Adjusted R-squared:  0.05438
F-statistic: 1.173 on 1 and 2 DF,  p-value: 0.3921
```



Because of the limited amount of data, in fact, we have no way to see the image of the simple linear regression we have done intuitively. However, it can be seen from the conclusion that p-value is equal to 0.3921 and greater than 0.05 that this model is actually not that reasonable. The number of police-killings is not very related to income between regions.

Conclusion & Discussion

From the beginning we read that the number of police killings in the Washington post has been increasing every year. And the number of blacks killed is much larger than the number of whites and other races. Including the analysis of this notebook13d55efb9b also said that blacks are more likely to be killed by the police. But in fact, from my own model, at least in 2015, the factor of this race is not very big. No matter it is South region, West region, Midwest region and Northwest region, the number and proportion of police-killing of each race are not much different. After that, I also conducted an analysis on whether this number depends on the economic situation of different regions. Our simple linear regression also shows that their correlation is not very large. The number difference between each region depends more on the total population of the region. However, in 2021, although there are many news reports with more police killing cases, does the number of such cases really differ between each race? Or does this ratio get bigger and bigger with the years? This is something we can analyze further. Although the amount of data used in this project is not very large, it can be seen from a simple analysis that in 2015, the number of police killings was stable. Of course, if this number decreases every year, then society will become more harmonious.

Reference

Police Shootings and Race

https://www.kaggle.com/suadadaism/police-shootings-and-race

Franklin E. Zimring ‘Police Killings as a Problem of Governance’ Volume: 687 issue: 1, page(s): 114-123

police-shootings-database
https://www.washingtonpost.com/graphics/investigations/police-shootings-database/