# Analysis of Police Shootings by Demographic

## Introduction

Recently in the United States police use of deadly force has become a politically important issue in conjunction with the "Black Lives Matter" movement which is concerned with the idea that systematic racism has caused the devaluing of African American lives. In this project I will try to find insight into three questions: Are the police in the US killing a disproportionate amount of unarmed black men, where is this happening, and who is being killed? The Washington Post has been collecting data on police killings from 2015. I will use this data source to answer these questions and be able to see a visual of where this has happened.

## Datasets

To begin, I used the following datasets for my analysis:

1. "Police Shootings Since 2015" maintained by The Washington https://www.washingtonpost.com/graphics/investigations/police-shootings-database/
- This data set contains all police shootings since 2015 including location, age, gender, race, armed, mental illness, and more.
2. "United States Population Statistics" from the following website: https://www.census.gov/data/tables.html
- This dataset was used to estimate the population numbers by ethnic race.
- The Hispanic population in the United States often identifies as "White" so in many cases I combined or labeled data as "White and Hispanic"
3. "Violent Crime Statistics by Race" maintained by "National Incident Based Reporting System https://www.fbi.gov/services/cjis/ucr/nibrs

- This data set contains violent crime statistic per year per race.

4. “Percentage Over 25 Completed High School” and “Percentage of People Below. Poverty Level” https://www.kaggle.com/kwullum/fatal-police-shootings-in-the-us

## Methodology

Since 2015, the police have killed 5,424 people. 352 of those individuals were unarmed. This has happened in various locations across the United States. I used Python to visual all of the locations of unarmed deaths by police:

Between 2015 and 2020, the police in America killed the following unarmed persons:

Asian 7

African American 123

Hispanic 63

Other 9

Unknown 5

Caucasian 145

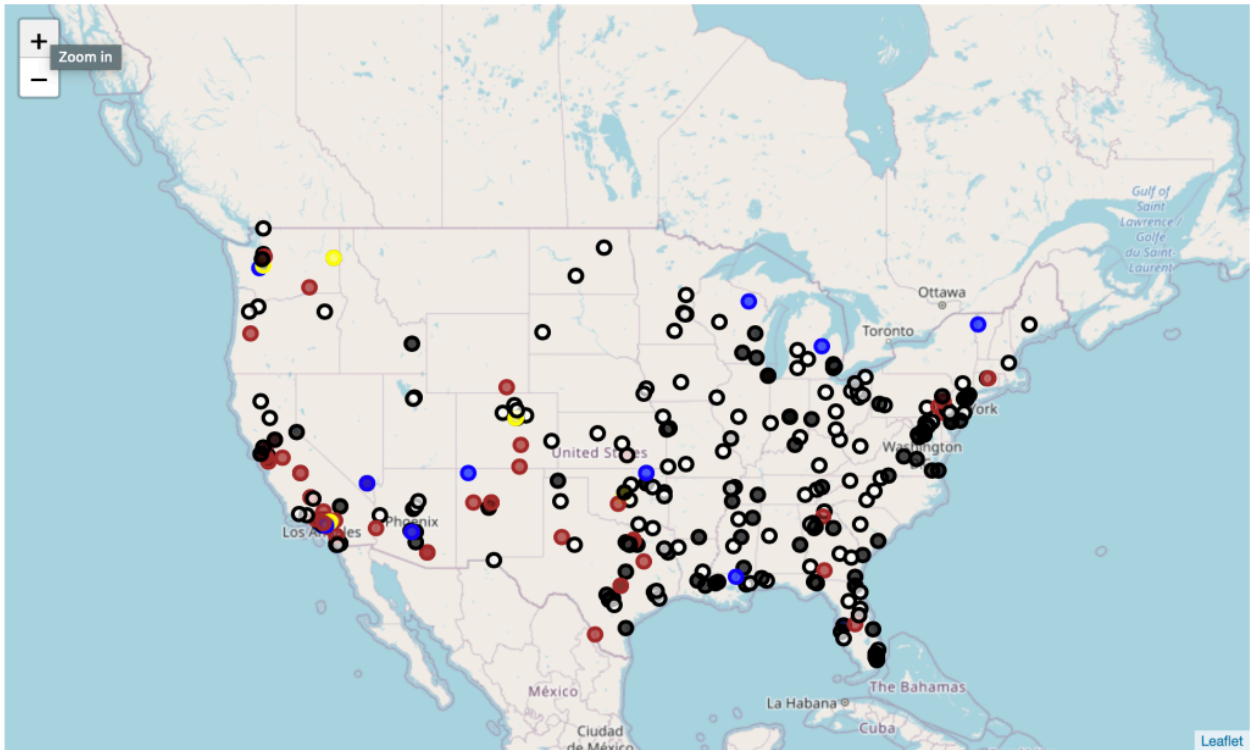
**White = Caucasian**

**Black = African American**

**Tan = Brown**

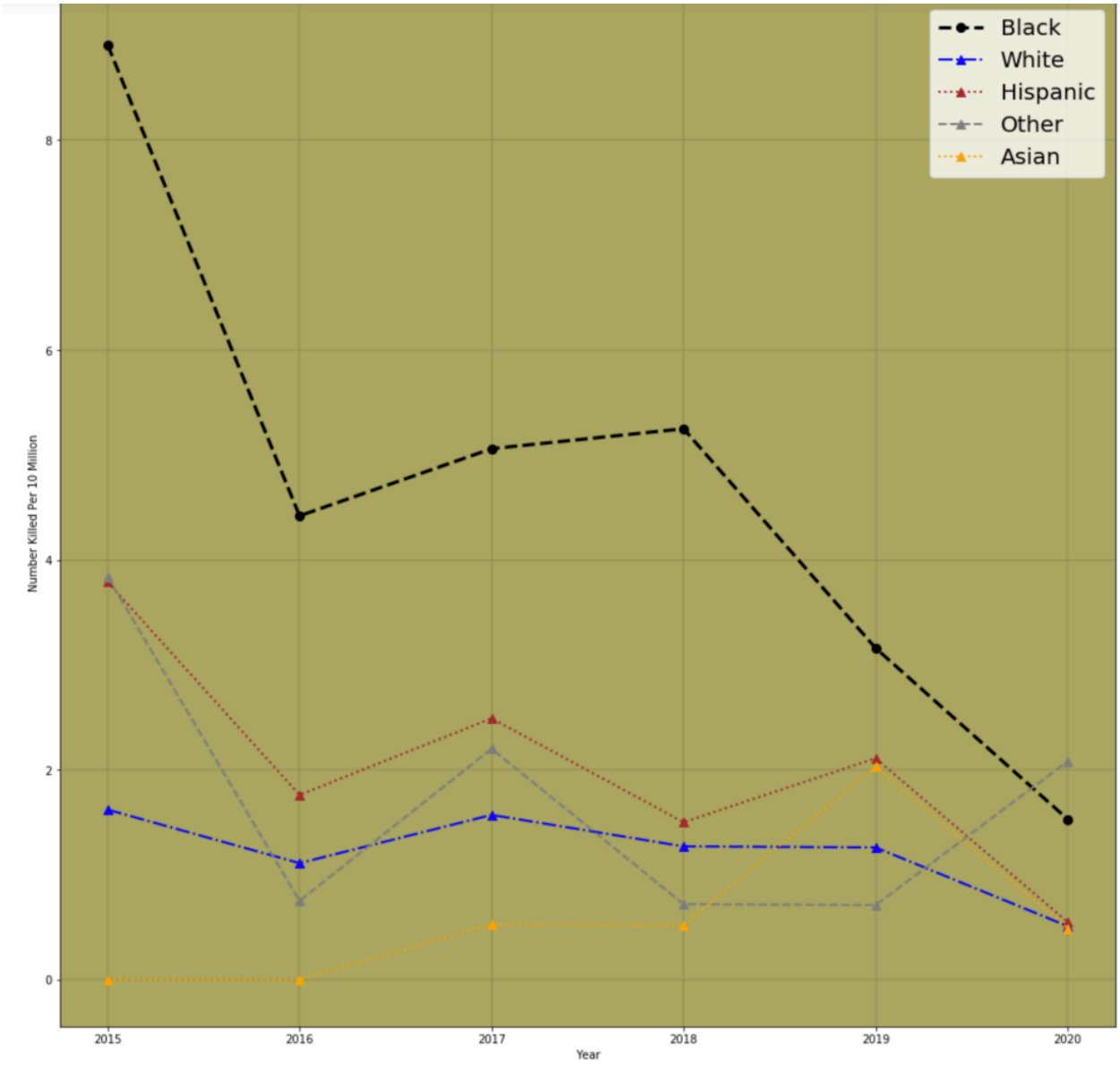
**Yellow = Asian**

**Blue= Other**




This map indicates geographic locations of police unarmed shootings. African American incidents tend to be on the eastern portion of the country with a few incidents in California and Washington state. Caucasian incidents tend to be more spread over the country. Hispanic tend to be in the south western portion of country. Asian incidents tend to be on the western seaboard.

Using population data per year per race, I calculated the amount of unarmed police killings per race:

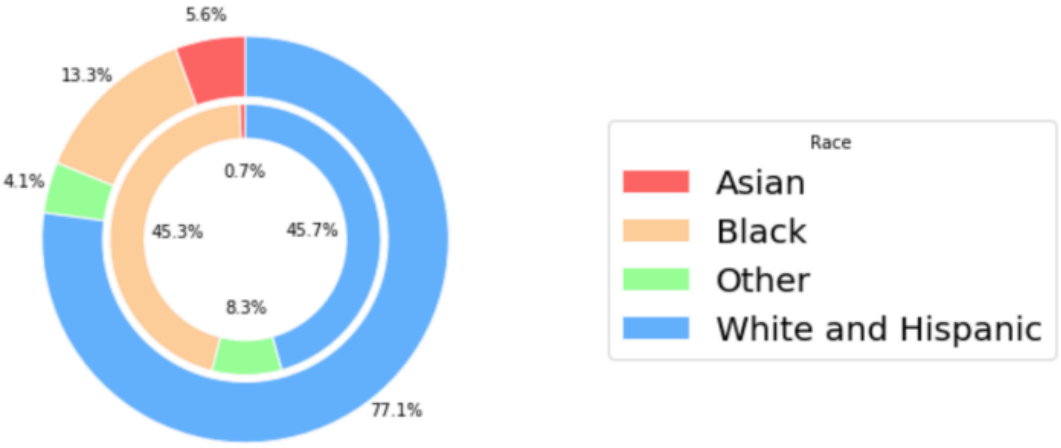



The line graph above depicts unarmed deaths by police by ethnicity and number killed per 10 million. The rate of African American's killed in 2015 was much higher than any other ethnicity, although it does decrease over time. The other ethnicities stay consistent, but this graph indicates that there are more deaths of African Americans than any other ethnicity per number of people in the country.

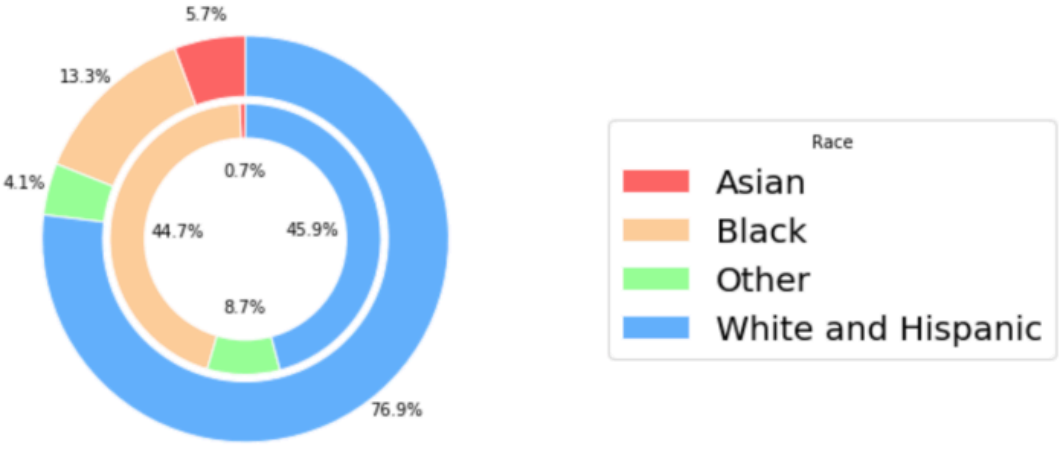
**Note: 2020 is a partial year so its numbers are lower than other years.**

The next important step is to analyze the amount of violent crimes recorded as compared the population numbers per year:

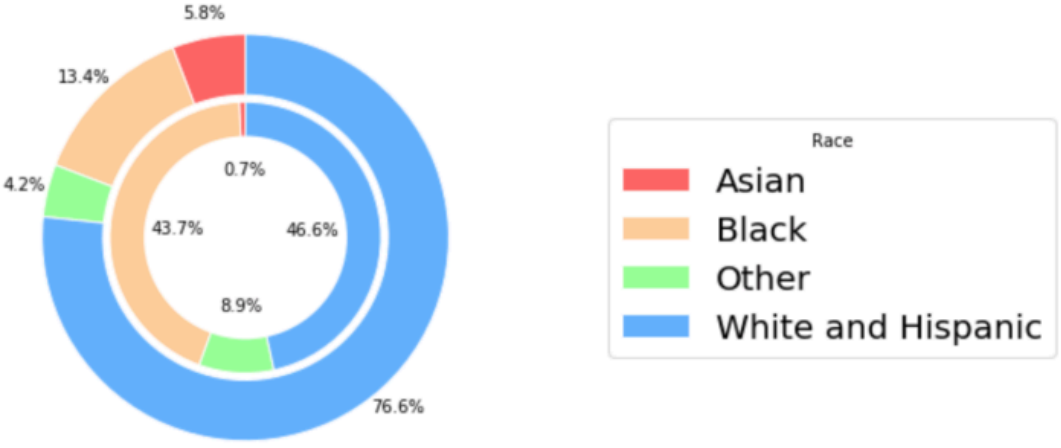
2015: Outer Circle: % Population by Race, Inner Circle % Violent Crime by Race

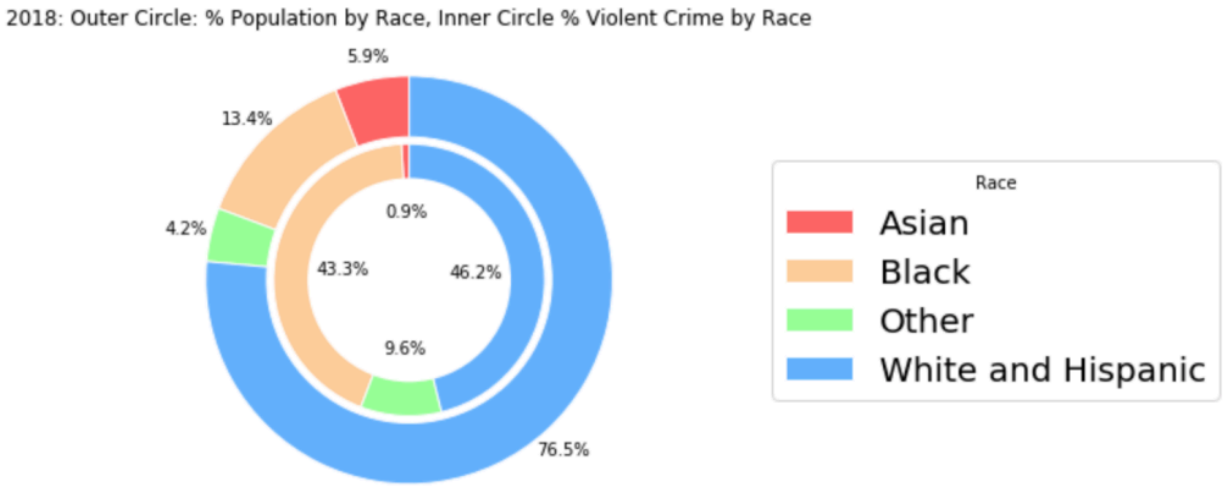



2016: Outer Circle: % Population by Race, Inner Circle % Violent Crime by Race




2017: Outer Circle: % Population by Race, Inner Circle % Violent Crime by Race





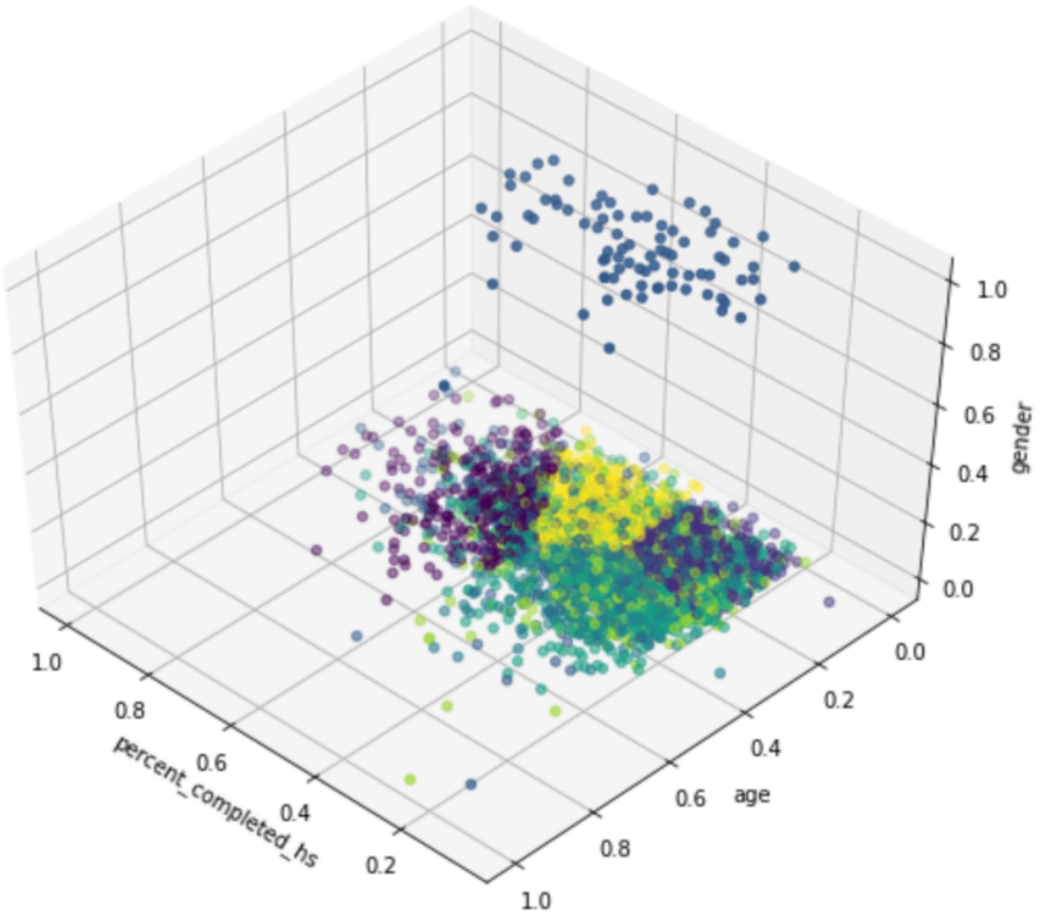

Year over year analysis shows that the African American population in the United States has a disproportionate amount of violent crimes compared to other races. This most likely plays a factor in the amount of police shootings of African Americans.

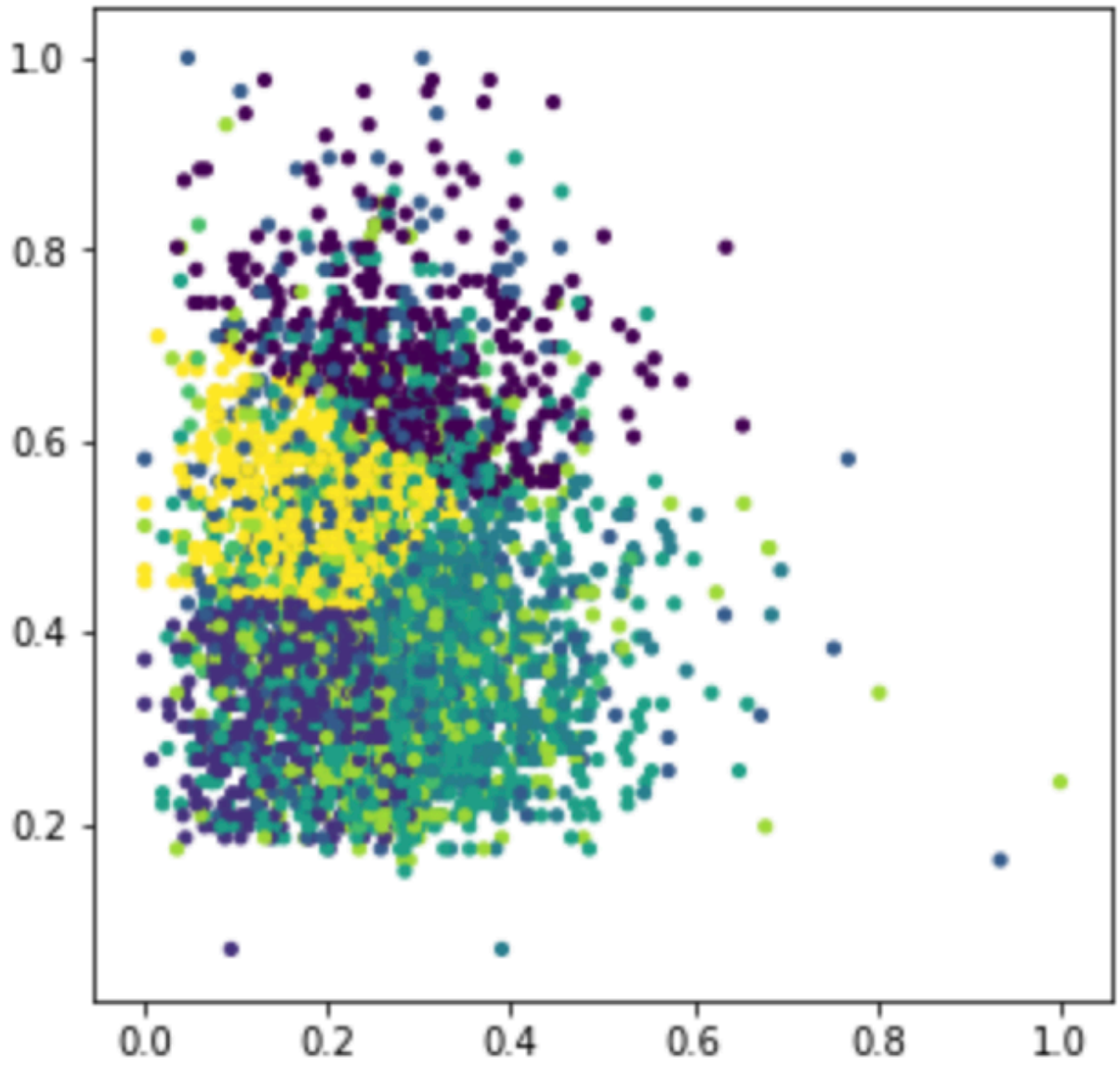
Now that the unarmed killings by police have been analyzed, its time to look at the overall police shooting data set and compare to a few sociological factors. To do this I took the police killings dataset and merged it with the poverty level and education level datasets. Once this was completed, I used the K-Means clustering algorithm to attempt to find insightful information based on this new dataset. I found that “8” was the optimal cluster number based on the silhouette score of .773. Using K-Means I was able to find the following data clusters:

; double click to hide

| | percent_completed_hs | poverty_rate | age | gender | Asian | Black | Hispanic | None | Other | White |
|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 0.842495 | 0.280831 | 0.375392 | 1.000000 | 0.0 | 1.000000 | 0.0 | 0.0 | 0.000000 | 0.0 |
| **1** | 0.860813 | 0.242530 | 0.462973 | 1.000000 | 0.0 | 0.000000 | 0.0 | 0.0 | 0.000000 | 1.0 |
| **2** | 0.803777 | 0.256963 | 0.391233 | 0.973333 | 0.0 | 0.000000 | 1.0 | 0.0 | 0.000000 | 0.0 |
| **3** | 0.844979 | 0.249322 | 0.499965 | 1.000000 | 0.0 | 0.000000 | 0.0 | 0.0 | 0.109756 | 0.0 |
| **4** | 0.854000 | 0.243571 | 0.455032 | 0.000000 | 0.0 | 0.000000 | 0.0 | 0.0 | 0.000000 | 1.0 |
| **5** | 0.867174 | 0.269254 | 0.364678 | 0.942029 | 0.0 | 0.000000 | 0.0 | 1.0 | 0.000000 | 0.0 |
| **6** | 0.845880 | 0.234987 | 0.428692 | 0.951807 | 1.0 | 0.000000 | 0.0 | 0.0 | 0.000000 | 0.0 |
| **7** | 0.843302 | 0.260718 | 0.412242 | 0.000000 | 0.0 | 0.811321 | 0.0 | 0.0 | 0.056604 | 0.0 |

I also used Python data visualizations to analyze the data:


1.0
0.8
0.6
0.4
0.2
0.0
gender
0.0
0.2
0.4
0.6
age
0.8
1.0
1.0
0.8
0.6
0.4
0.2
percent_completed_hs




Overall Report Results:

1. African Americans are killed more than any other race in the United States by the police. African American violent crimes account for 45% of all crime but African Americans account for only 13% of the population. Compared with other groups this is disproportionate.
2. Individuals who are killed by the police tend to be male, middle-aged. They also tend to come from areas with increased povery levels and less education.

3. African American shootings by police tend to happen in densely populated areas, White tend to happened all over the United States, and Hispanic deaths tend occur on the west coast.

**Discussion**

Based on the observations in this report, African Americans are killed more by the police but also commit a disproportionate amount of violent crimes. These two observations are most likely related. It's not easy or realistic to be able to prove that the police are disproportionally killing African Americans because of racism. Racism is sociological complex problem.

**Conclusion**

Based on the findings of this report, decreasing poverty, increasing education, and decreasing violent crime will reduce the amount of police killings for African Americans and all people in the United States.

Github Link to my Notebook and Datasets:

https://github.com/sn3kyJ3di/DS-Capstone_Project/tree/master/FinalProject