

# Police Shooting

Jannette Ortiz

2022-09-19

## Introduction

For the analysis, I will be utilizing the police shooting data set. This data was taken from the Washington Post article in 2015. There was a total of 3,960 observation and 14 variables collected. I will be exploring both of the variables “race” and “ages”.The variable “race” will determine what race have fail as a victim of police shooting. The variable “age” will determine the age of the individual. I think that exploring these data will help me find an interesting hypothesis on the age and race of those who were killed by police.

## Univariate Exploration For Each of the Variables Under Consideration

### 1. Race

```
table(shooting$race)
```

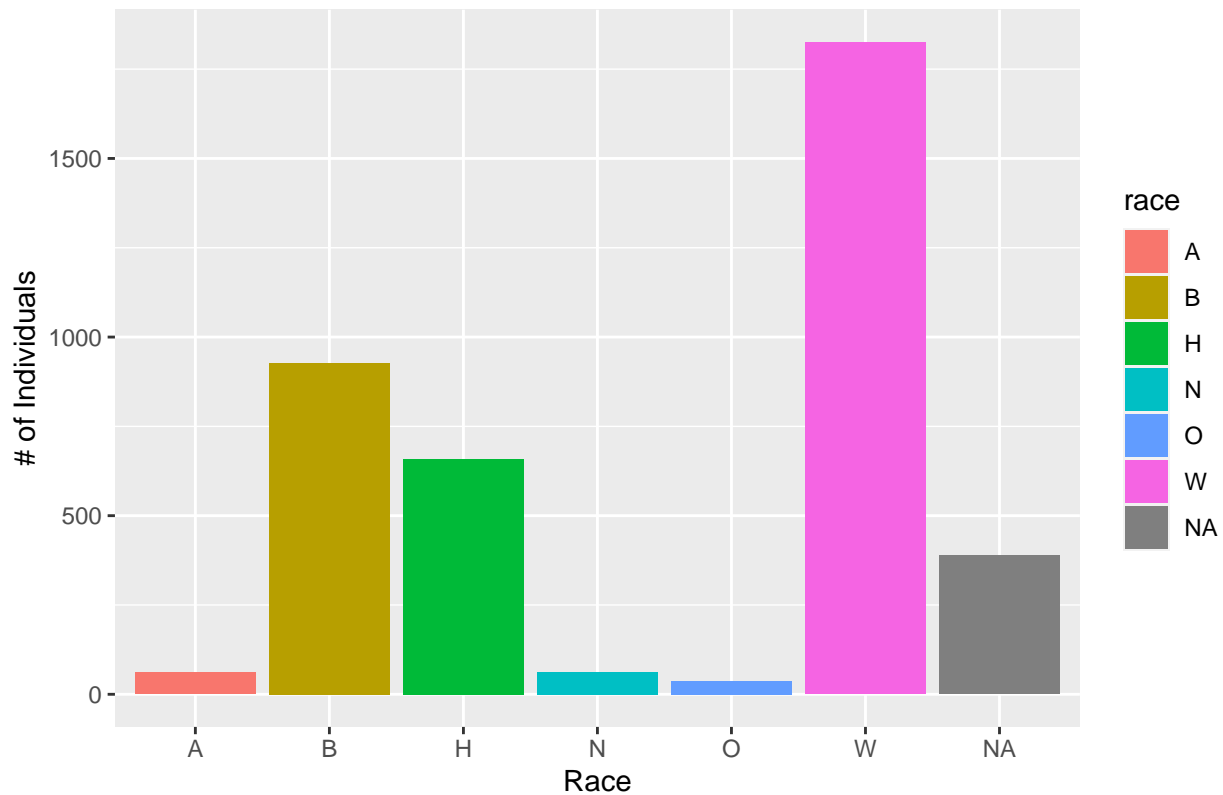
```
##  
##      A      B      H      N      O      W  
##    61    927    659     62     37   1825
```

Asian Black Hispanic Native American Other White 61 927 659 62 37 1825

In our table above we can see that White(1825), Black(927), and Hispanics(659) are the three major race that involves the individual part of the police shooting.

```
ggplot(shooting,aes(x=race, fill= race))+geom_bar()+ggtitle("Individuals by Race")+xlab("Race")+ylab("#
```

### Individuals by Race



In the bar chart above we can also see the number of individuals in each race that were involved in the police shooting. Comparing the bar chart with the table we can see how White, Black, and Hispanics had the most deaths.

## 2. Armed

```
summary(shooting$age)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   NA's  
##      6.00  27.00   35.00   36.85  45.00   91.00   152
```

Looking at the table above we can see that the mean age of the individual killed was 37 years old. The youngest individual killed was 6 yrs old. The oldest individual killed was 91 years old. There is about 152 individuals that were reported as a unknown age.

```
ggplot(shooting, aes(x=age))+geom_density()+xlab("Age of Individual")+ggtitle("Density of Individuals Aged")
```

```
## Warning: Removed 152 rows containing non-finite values (stat_density).
```



Looking at the density plot above we can see the distribution of the ages of individuals who were shot by police. As expected around the age 30-40 years old is the mean age of individuals shot.

## Bivariate Exploration Between Two Variables of Interest: Race and Age

Race vs. Age

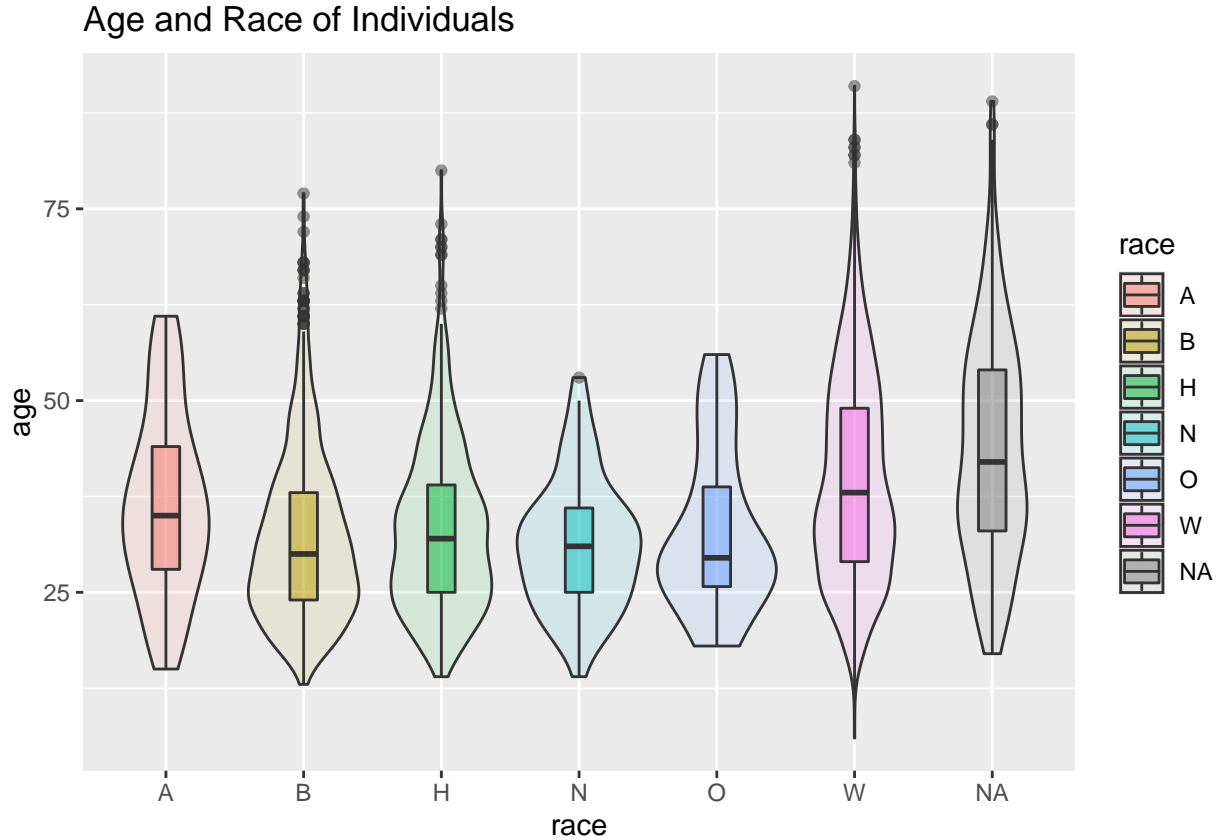
```
by_race<-group_by(shooting,race)
summarise(by_race,avg_age_and_race=mean(age,na.rm=TRUE))
```

```
## # A tibble: 7 x 2
##   race avg_age_and_race
##   <chr>         <dbl>
## 1 A             36.5
## 2 B             32.2
## 3 H             33.3
## 4 N             31.2
## 5 O             32.9
## 6 W             39.7
## 7 <NA>         43.4
```

```
ggplot(shooting, aes(x=race,y=age, fill=race))+geom_violin(alpha=.1)+geom_boxplot(alpha=.5, width=.2)+g
```

## Warning: Removed 152 rows containing non-finite values (stat\_ydensity).

## Warning: Removed 152 rows containing non-finite values (stat\_boxplot).



With the boxplot on top we can see the relationship between age and race. We can see that in our three major race White, Black, and Hispanic they all fall under the age of 25-50 years old. All the rest of the races also land in between 25-50 years old.

## Conclusion

From our data above we are shown that our mean age of an individual killed was 37 years old. According to my results there seems to be a correlation of the majority group killed by police officers are White(1825), Black(927), and Hispanics(659) from the age of 25 to 50 years old. Looking at my data above there must be a correlation of the age and race group that are targeted the most. The Officer involve with the shooting is racially profiling a specific age group of individuals.