final project

austin\_testa

February 19, 2019

## R Markdown

library(ggplot2)

Introduction; For this project I decided to analyze the police shooting data that was created by the FBI and the CDC in 2015. The data was made to keep track of the rising levels of fatal police shootings. I will be analyzing the race, age, and gender of the victims of the shootings.

library(readxl)
police <- read\_excel("C:/Users/Austin/Desktop/police.data.xlsx", sheet=1, col\_names=TRUE)
police[1:10,1:5]

## # A tibble: 10 x 5
## id name date manner\_of\_death armed
## <dbl> <chr> <dttm> <chr> <chr>
## 1 3 Tim Elliot 2015-01-02 00:00:00 shot gun
## 2 4 Lewis Lee Lembke 2015-01-02 00:00:00 shot gun
## 3 5 John Paul Quintero 2015-01-03 00:00:00 shot and Tasered unarmed
## 4 8 Matthew Hoffman 2015-01-04 00:00:00 shot toy weap~
## 5 9 Michael Rodriguez 2015-01-04 00:00:00 shot nail gun
## 6 11 Kenneth Joe Brown 2015-01-04 00:00:00 shot gun
## 7 13 Kenneth Arnold Buck 2015-01-05 00:00:00 shot gun
## 8 15 Brock Nichols 2015-01-06 00:00:00 shot gun
## 9 16 Autumn Steele 2015-01-06 00:00:00 shot unarmed
## 10 17 Leslie Sapp III 2015-01-06 00:00:00 shot toy weap~

summary(police$armed)

## Length Class Mode
## 3960 character character

This set was to analyze wether or not the victims were armed,and with what weapon but the list of differentweapons was too large to make into a graph.

this next data set that I looked at was the ages of the victims.

summary(police$age)

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

This graph was to show the distribution of the victims ages. The average age was 36 years old, and the median ws 34.

ggplot(police, aes(x=age)) + geom\_density()

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



summary(police$race)

## Length Class Mode
## 3960 character character

ggplot(police, aes(x=gender)) + geom\_bar()



The graph above was analyzing gender of the victims and about 90 percent of the victims were male.

boxplot(depression$cesd, horizontal = TRUE, main=“Distribution of depression”, xlab=“cesd”)

This graph was focusing on the race of the victims. I analyzed and made a bar chart to show which race was most commonly shot and to see if there is any trends in the data. This chart shows that the majority of the victims were wight and the next largest group was the african american.

police <- police[!is.na(police$race),]
ggplot(police,aes(x=race, fill = race)) + geom\_bar(position = "dodge") + ggtitle("Race")



ggplot(police, aes(x=race,y=age,fill=race)) + geom\_boxplot() + ggtitle("Race vs. Age")

## Warning: Removed 47 rows containing non-finite values (stat\_boxplot).

 This boxplot shows the comparison between race and age in the police shooting data set. From this data I can back up what I said earlier, that regaurdless of race the average age is around 35 years old.

Conclusion: In conclusion the data shows that the majority of the victims are white and secondly black. Also the victims seem to vary in ages from 25 to 35 and are predominantly males.