MATH130 final project

Anne Sullivan

February 21, 2019

library(readxl)
library(rmarkdown)
library(knitr)
library(forcats)
library(dplyr)
library(scales)
library(ggplot2)

fatalpoliceshootings <- read\_excel("/Users/anniesullivan/MATH130/fatal-police-shootings-data.xlsx", sheet=1, col\_names=TRUE)

### Intro:

#### This data set includes every recorded fatal shooting committed by American police officers upon civilians until the date of January 1 2015. The data includes variables such as race of deceased, situation of the shooting, whether the deceased was carried a weapon or had mental health issues etc.

### Univariant description of variables (age, flee):

#### In this section we will explore the relationship of fatal police shootings between the two variables age and whether or not the civilian fleed upon interaction with the police officer. To get acquainted with the data I have first included a small exerpt of the data set. From here we will isolate the variable age and then flee and explore more in depth information regarding the variables.

str(fatalpoliceshootings)

## Classes 'tbl\_df', 'tbl' and 'data.frame': 3960 obs. of 14 variables:
## $ id : num 3 4 5 8 9 11 13 15 16 17 ...
## $ name : chr "Tim Elliot" "Lewis Lee Lembke" "John Paul Quintero" "Matthew Hoffman" ...
## $ date : POSIXct, format: "2015-01-02" "2015-01-02" ...
## $ manner\_of\_death : chr "shot" "shot" "shot and Tasered" "shot" ...
## $ armed : chr "gun" "gun" "unarmed" "toy weapon" ...
## $ age : num 53 47 23 32 39 18 22 35 34 47 ...
## $ gender : chr "M" "M" "M" "M" ...
## $ race : chr "A" "W" "H" "W" ...
## $ city : chr "Shelton" "Aloha" "Wichita" "San Francisco" ...
## $ state : chr "WA" "OR" "KS" "CA" ...
## $ signs\_of\_mental\_illness: logi TRUE FALSE FALSE TRUE FALSE FALSE ...
## $ threat\_level : chr "attack" "attack" "other" "attack" ...
## $ flee : chr "Not fleeing" "Not fleeing" "Not fleeing" "Not fleeing" ...
## $ body\_camera : logi FALSE FALSE FALSE FALSE FALSE FALSE ...

### Age:

##### The minimum age recorded was 6 years and the maximum age was 91. The average age of those fatally shot was roughly 37. Of the 3960 observations 152 were not applicable when it came to the variable age.

summary(fatalpoliceshootings$age)

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

boxplot(fatalpoliceshootings$age, horizontal=TRUE, main= "Dispersal of ages of deceased", xlab="age")



##### This graph plots the recorded ages of the deceased claimed by police shottings, the youngest being 6, the average being 37 and the eldest being 91. As indicated the majority of the deceased is roughly between the ages of 25 and 45.

### Flee:

##### This table shows us the ways in which those of the deceased fleed from left to right: 631 fled by car, 491 fled by foot, 2570 did not flee and 128 fled by other means not specified.

table(fatalpoliceshootings$flee)

##
## Car Foot Not fleeing Other
## 631 491 2570 128

ggplot(fatalpoliceshootings, aes(x=forcats::fct\_infreq(flee))) + geom\_bar() + xlab("flee")



#### This graph shows us the relationship of the number of deceased who fled and the means by which they fled, it also indicates that a majority of the deceased did not flee.

### Bivariate Comparison:

#### In the following section we will inspect visually the correlation between age and flee and how these factors may have played a roll in the fatal shooting.

ggplot(fatalpoliceshootings, aes(x=flee, y=age, fill=flee)) + geom\_violin(alpha=.1) + geom\_boxplot(alpha=.5, width=.2)

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

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



#### This graph plots flee on the x-axis and age on the y-axis. We can see that the average age of those who flee is roughly 27, and a mjority of the deceased did not flee.

### Conclusion:

#### From the above listed data presented by graphs, data tables and summaries we can review the data regarding fatal shootings of civilians by on duty police officers in the U.S. We scrutinized two univariables: age and flee. From the univariable age we notice that the youngest age recorded was 6 and the eldest age recorded was 91; there were 152 recorded NA’s, and the mean was 37. We can reference the first graph for a conclusion of this data. The univariable flee shows the methods of fleeing that were recorded: by Car, Foot, Not Fleeing, other, or NA. Of the methods of fleeing the majority or average of those did not flee from police. We can reference the second graph for a summary of this data. From the bivariables I chose (age, flee) we can see that a majority of the deceased who were shot did not flee and the average age of the deceased was 37, we can reference the final graph for a visual of this information.