Data_Analysis_Project_slondon3

Sam London

2/25/2022

```
knitr::opts_chunk$set(warning=FALSE, message=FALSE, fig.height=4, fig.width=5, fig.align='center')
library(dplyr)
library(forcats)
library(ggplot2)
library(readxl)
library(sjPlot)
```

INTRODUCTION

The data set that I will be analyzing is the Police Shooting Data set. The data, that was gathered during a study in 2015 from individuals killed by police, which consists of 3960 observations and 14 variables. Out of the 14 variables, I will be looking at the age of the victim, gender, and if they were fleeing. So my question that I will be trying to answer is: Is there a correlation with the age and gender of the victim and their choice of how to flee?

```
police <- read_excel("C:/Users/Sam/Documents/Math130/Data/police_shootings.xlsx")
```

UNIVARIANT ANALYSIS OF VARIABLES
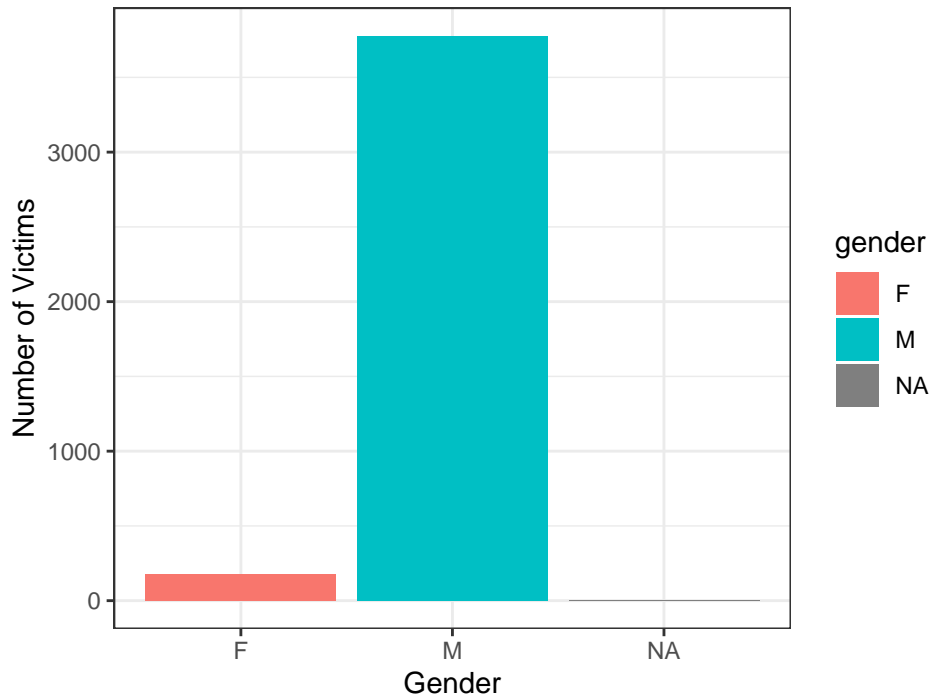
Gender

```
table(police$gender)
```

```
##
##    F    M
##  180 3777
```

The table refers to the individuals that were either male or female. Out of the data it seems that there were almost 20 times the males as victims than females.

```
ggplot(police, aes(x=gender, fill=gender)) + geom_bar() + theme_bw() + xlab("Gender") + ylab("Number of
```

Gender vs. Number of Victims

The graph shows the significance in the contrast of gender when looking at the victims. The number of males definitely surpass then those of females.

Age Range

```
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
```

The statistics show the average age of the killed victims, which is approximately 38 years old. The youngest being 6 and the oldest being 91. The average young adult being 27 and the average adult being 45. With 152 "Na's" the accuracy of the data can be questioned.

```
ggplot(police, aes(x=age)) + geom_histogram(aes(y= ..density..), col= "black", fill= "grey") + geom_den
```

Age Range of Victims
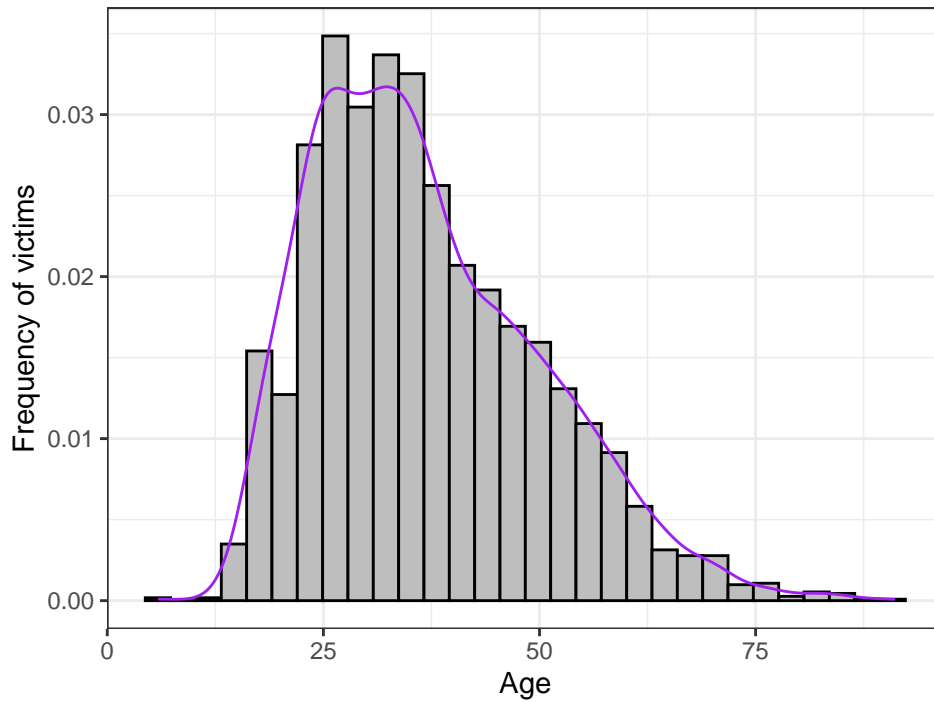
This graph show the frequency and distribution of each age group that was a victim as well as showing the individual age groups.

Fleeing Options
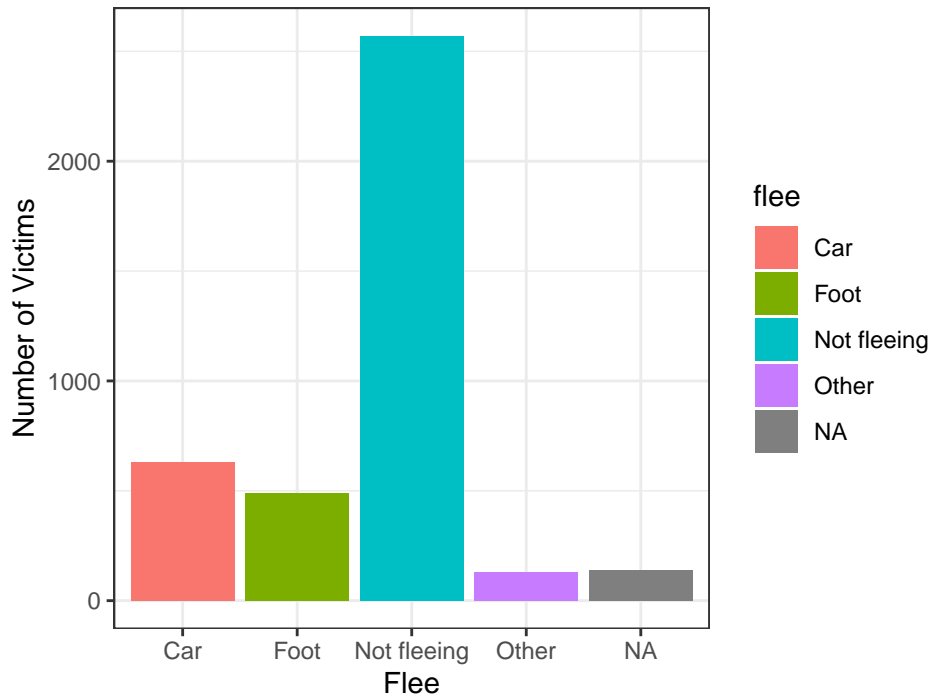
```
table(police$flee)
```

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

This table shows the options in which a victim can flee. Making it easier to understand what was the most popular option taken.

```
ggplot(police, aes(x=flee, fill=flee)) + geom_bar() + theme_bw() + xlab("Flee") + ylab("Number of Victi
```

Victims vs. Fleeing Outcome

This bar graph shows that out of the individuals that were fleeing, what mode of transportation was used to do so. It looks like the majority of those fleeing didn't actually flee, but for those that did it looks like a car was the best option.

BIVARIANT COMPARISON

Fleeing vs. Age

```
ggplot(police, aes(x=flee, y=age, fill=flee)) + geom_boxplot() + ggtitle("Age vs. Fleeing Outcome")
```

Age vs. Fleeing Outcome
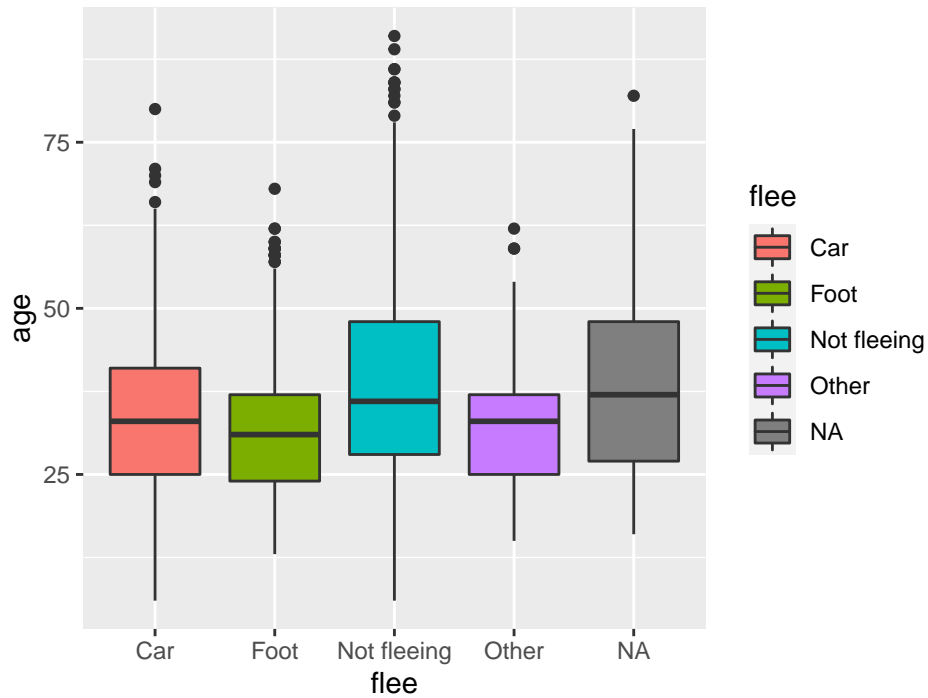
The correlation between age and fleeing outcomes does not seem to be of any significant value. As shown on the graph the individuals that were involved in the police shooting seemed to have equal amount of fleeing options, but the fleeing outcome for not fleeing was higher than the others.

Fleeing vs. Gender

```
table(police$gender, police$flee) %>% prop.table(margin = 1) %>% round(3)
```

```
##
##      Car  Foot Not fleeing Other
##   F 0.224 0.046       0.713 0.017
##   M 0.162 0.133       0.671 0.034
```

```
plot_xtab(police$gender, police$flee, margin = 'row', show.total = FLASE) + scale_fill_brewer(palette =
```
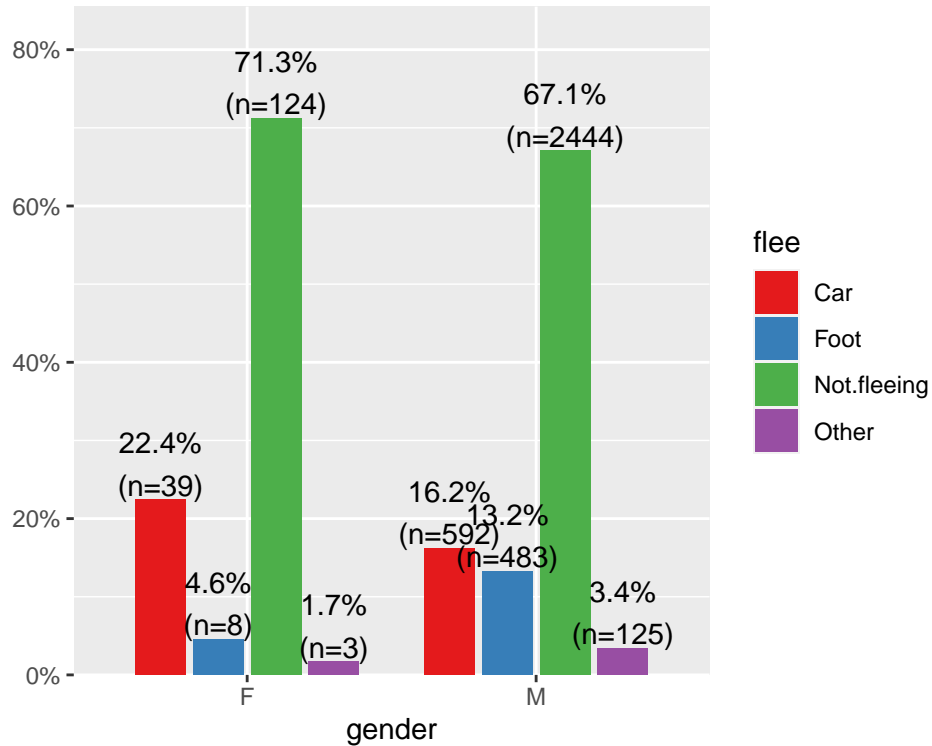
This graph shows what type of flee outcome happened to the two different genders in the shooting. The pattern for both genders seem to be the same with the most picked being not fleeing to car, foot, than other. you can see that gender and the fleeing options do not correlate with one another and work independently.

CONCLUSION

From my results, there seems to be no correlation between the age and gender of the victim and how the chose to flee from the scene. The results concluded that the majority, whether male or female, decided or didn't flee from the scene. So there is no significance in the correlation of gender, age, and fleeing options.