Analysis of Fatal Police Shooting Data

Meredith Lou

4/23/2022

Description of Data

I used a database from the Washington Post that records every fatal police shooting since January 1, 2015. The Post gathered this data from local news reports, law enforcement websites and social media, and independent databases. In this data, there are variables for the name of the victim, race, age, gender, manner of death, if/how they were armed, city, state, signs of mental illness, threat level, if/how they fled, latitude, and longitude.

Plan

I measured the rate of shootings by number of shootings per 1,000,000 people. I created a choropleth mapping this rate of shootings across the US, a time series of police shootings from 2015, bar graphs for variables including race, age, gender, whether fleeing or whether armed, and double deckers of race, armed, and flee and then race, age and gender.

Design

I wanted to create visualizations that addressed multiple variables associated with fatal police shootings. For categorical variables such as race, gender, age, fleeing, and armed, I created bar graphs. For a display over time I created a time series, and for geographical mapping I created a choropleth. In order to analyze multiple variables at once, I chose to create double decker displays.

```
library(tidyverse)
library(ggplot2)
library(maps)
library(vcd)
library(tidycensus)
library(lubridate)
data <- read_csv("fatal-police-shootings-data.csv")
```

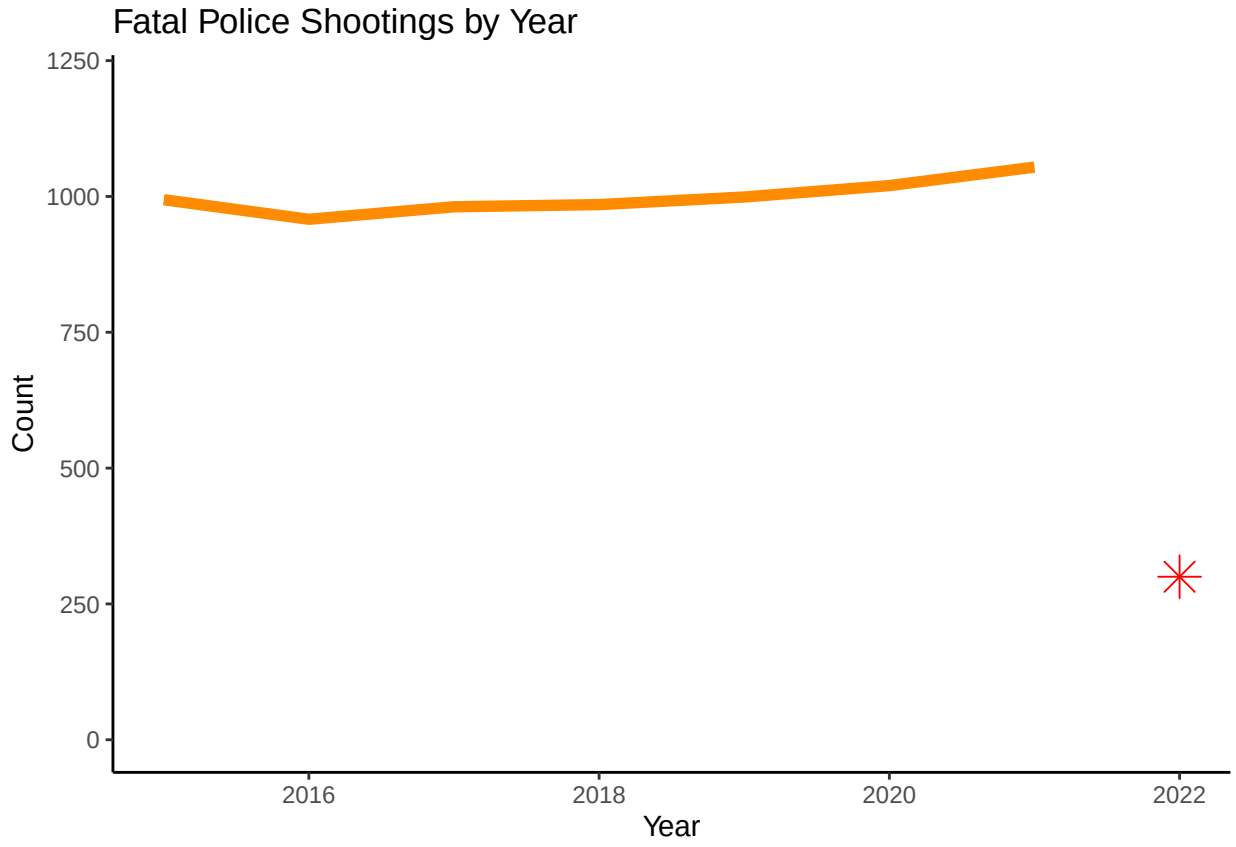
Preprocessing

A big allocation of preprocessing went to collapsing the age and armed variables into categories. I also filtered and collapsed the race variables in order to focus on the most politically relevant races: White, Black and Hispanic.

Time Series

In order to properly visualize the data for the time series, I chose to represent 2022 as a singular point to show where we are now compared to the final values of previous years.

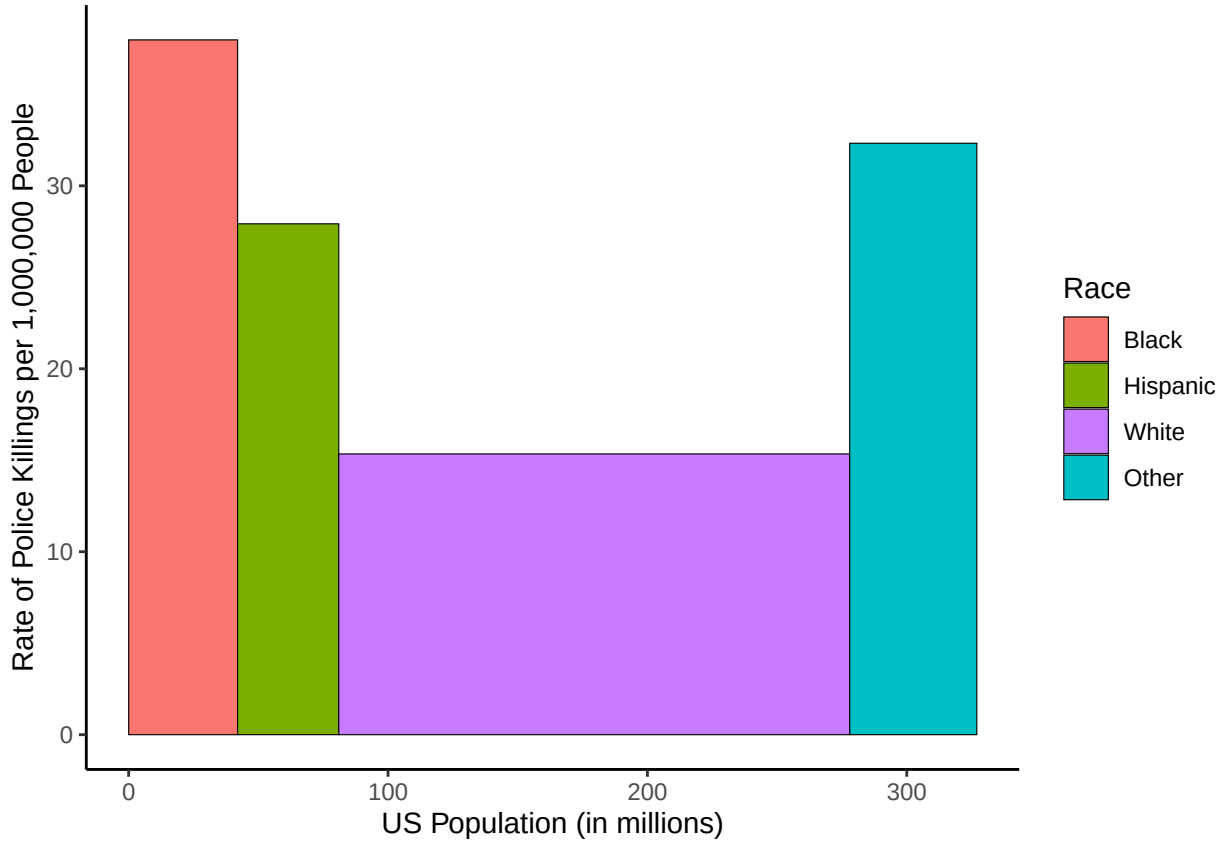
```
freqs <- aggregate(data$date, by=list(data$date), FUN=length)
freqs$Group.1 <- as.Date(freqs$Group.1, format='%m/%d/%y')
freqs$Year <- year(freqs$Group.1)
freqs <- aggregate(x~Year,freqs, FUN=sum)
freqs_2022 <- freqs[(freqs$Year == 2022), ]
freqs_all <- freqs[(freqs$Year <= 2021), ]

ggplot(freqs_all, aes(x=Year, y=x)) +
  geom_line(colour = "darkorange", linewidth=2) +
  geom_point(freqs_2022, mapping = aes(x=Year, y=x), color="red", pch = 8,
             size=5) +
  labs(title="Fatal Police Shootings by Year", y="Count") +
  ylim(0, 1200) +
  theme(legend.position="none") + theme_classic()
```



Histogram

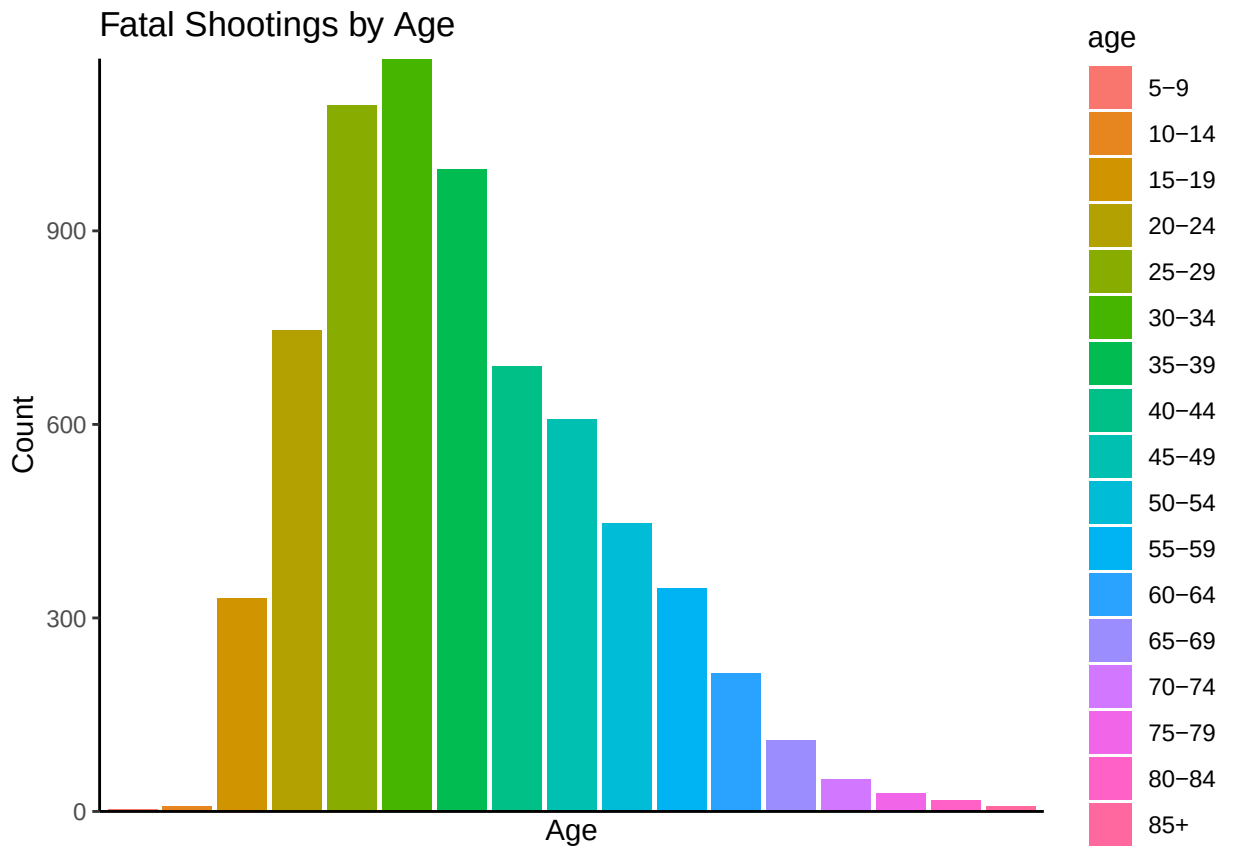
I had to create a dataframe of the population of each race in order to create visualizations that portrayed fatal shootings as a proportion of the population of each race rather than the total count.

```r
# race proportional to population size
population_data <- data.frame(race = c("Black","Hispanic","White","Other"),
                              population =
                              c(42000000,39000000,197000000,49000000))
hist <- summarize(group_by(data,race), count=n())
hist <- mutate(hist, race = fct_collapse(race, 'Black' = 'B',
                                               'Hispanic' = 'H',
                                               'White' = 'W',
                                               other_level = 'Other'))
hist[is.na(hist)] = 'Other'
hist <- summarize(group_by(hist, race), 'total' = sum(count))
race_data <- full_join(population_data,hist,by="race")
race_data$rate <- race_data$total/race_data$population*1000000
race_data$cumsumUpper <- cumsum(race_data$population)
race_data$cumsumLower <- c(0, cumsum(race_data$population))[1:nrow(race_data)]
ggplot(data=race_data) +
  geom_rect(aes(xmin=cumsumLower/1000000,xmax=cumsumUpper/1000000,
                ymin=0, ymax=rate,fill=race),
            color = "black", linewidth = .2) +
  labs(x = "US Population (in millions)", y = "Rate of Police Killings per 1,000,000 People", fill = "Ra
  theme_classic() + scale_fill_discrete(breaks=c('Black','Hispanic','White','Other'))
```

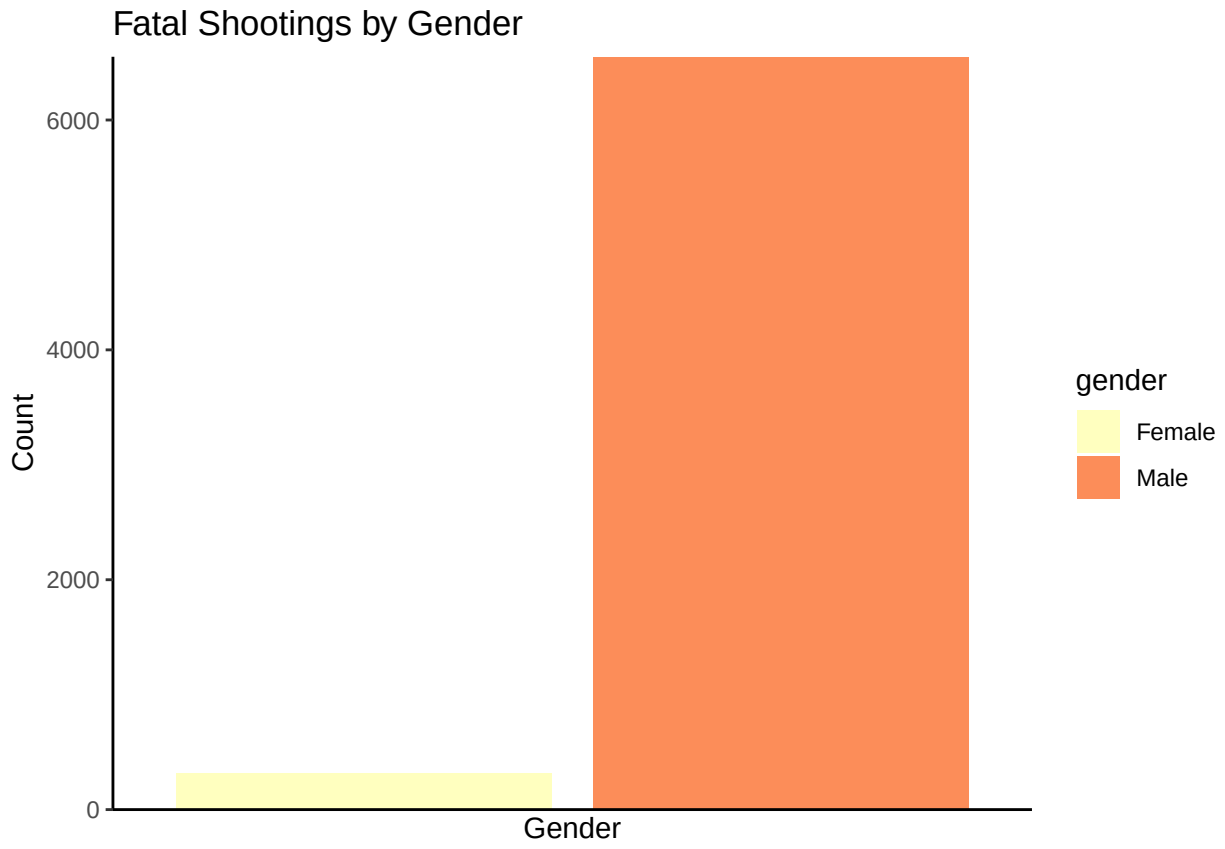



```r
#age of person shot
data$age <- as.factor(data$age)
data <- filter(data, age !="")
```

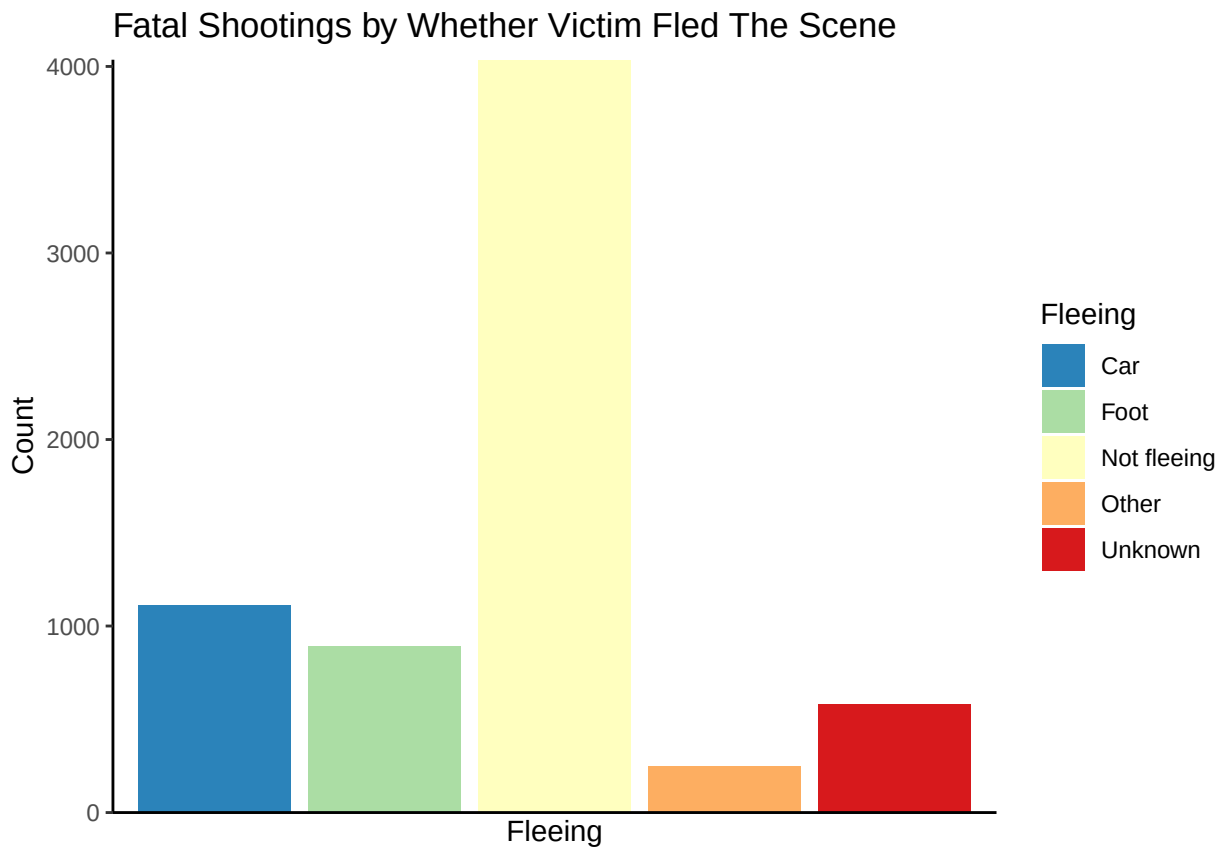
```
age_data <- mutate(data,
                   age=fct_collapse(age,"5-9"=c("6","8"),
                                    "10-14"=c("12","13","14"),
                                    "15-19"=c("15","16","17","18","19"),
                                    "20-24"=c("20","21","22","23","24"),
                                    "25-29"=c("25","26","27","28","29"),
                                    "30-34"=c("30","31","32","33","34"),
                                    "35-39"=c("35","36","37","38","39"),
                                    "40-44"=c("40","41","42","43","44"),
                                    "45-49"=c("45","46","47","48","49"),
                                    "50-54"=c("50","51","52","53","54"),
                                    "55-59"=c("55","56","57","58","59"),
                                    "60-64"=c("60","61","62","63","64"),
                                    "65-69"=c("65","66","67","68","69"),
                                    "70-74"=c("70","71","72","73","74"),
                                    "75-79"=c("75","76","77","78","79"),
                                    "80-84"=c("80","81","82","83","84"),
                                    "85+"=c("86","88","89","91","92")))

ggplot(data=age_data) +
  geom_bar(aes(x=age,fill=age)) +
  scale_x_discrete(labels=NULL, breaks=NULL) +
  scale_y_continuous(expand=c(0,0)) +
  theme_classic() +
  labs(title="Fatal Shootings by Age", x="Age", y="Count")
```



```
#gender
data <- mutate(data,gender=fct_recode(gender,"Male"="M",
                                      "Female"="F"))
data <- filter(data,gender !="")
ggplot(data=data) +
  geom_bar(aes(x=gender,fill=gender)) +
  scale_fill_brewer(palette="Spectral",direction=-1) +
  scale_x_discrete(labels=NULL,breaks=NULL) +
  scale_y_continuous(expand=c(0,0)) +
  theme_classic() +
  labs(title="Fatal Shootings by Gender",x="Gender",y="Count")
```

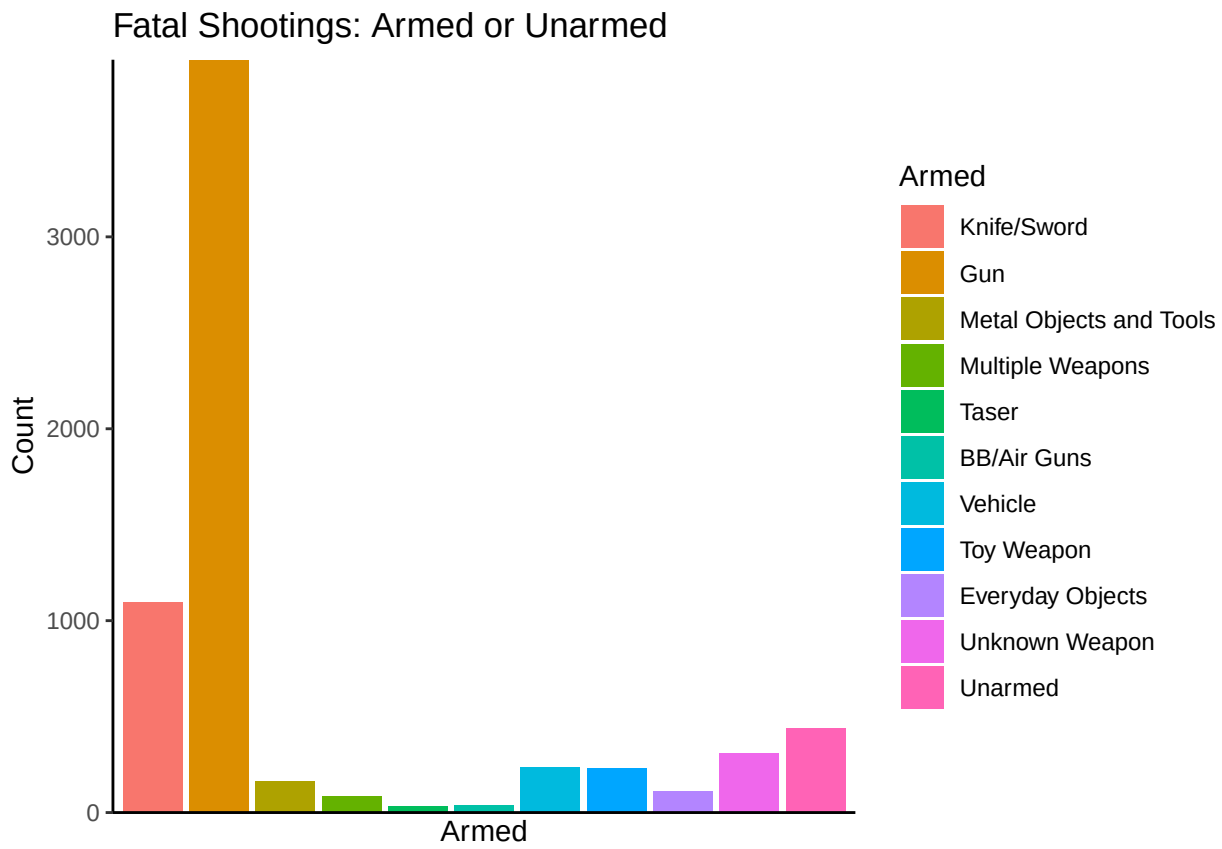



```
#fleeing vs not fleeing
flee_data <- data[,"flee"]
flee_data[is.na(flee_data)] = "Unknown"
ggplot(data=flee_data) + geom_bar(aes(x=flee,fill=flee)) +
  scale_x_discrete(labels=NULL,breaks=NULL) +
  scale_y_continuous(expand=c(0,0)) +
  theme_classic() +
  labs(title="Fatal Shootings by Whether Victim Fled The Scene",
       x="Fleeing",y="Count") +
  scale_fill_brewer(name="Fleeing", palette="Spectral", direction=-1)
```



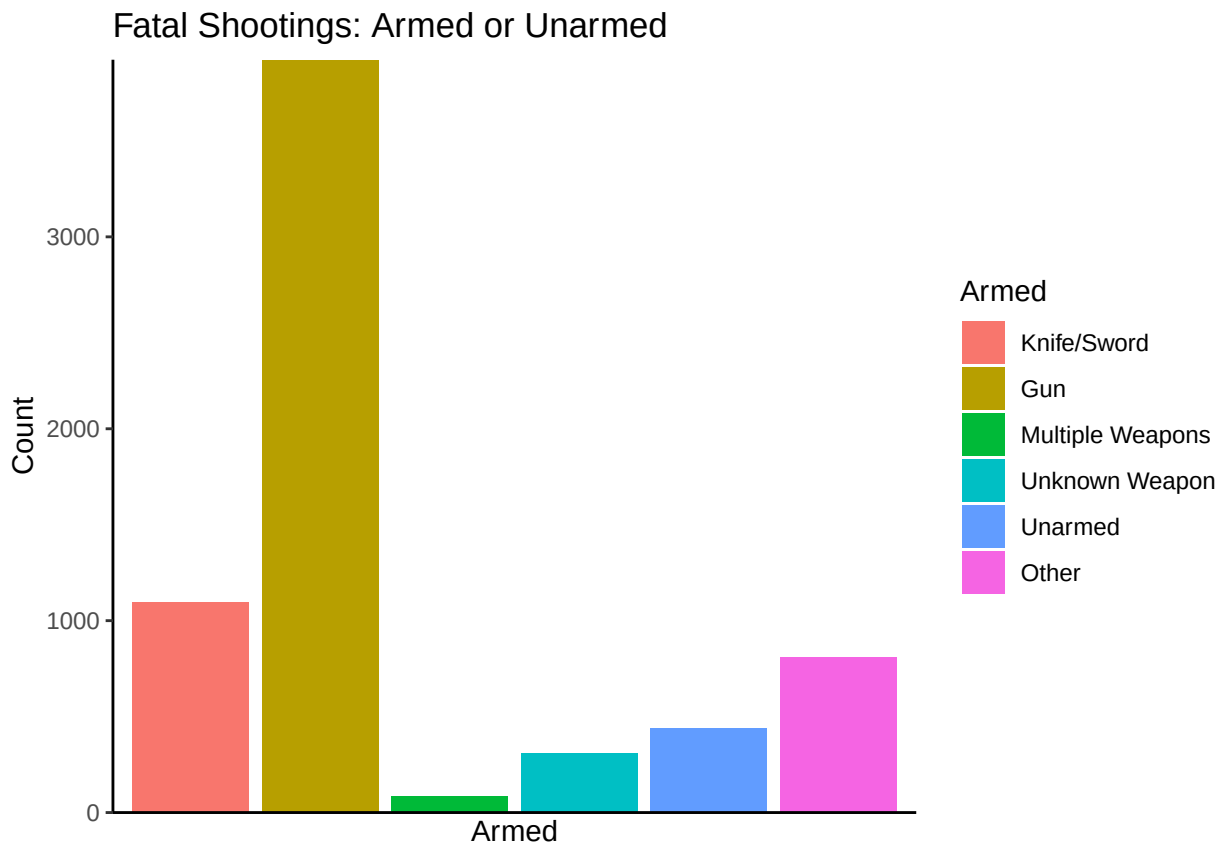

```
#armed vs not armed
data <- filter(data,armed !="")
armed_data <- mutate(data,armed=fct_collapse(armed,"Gun"=c("gun"),
                    "Knife/Sword"=c("knife","sword","machete","box cutter",
                                    "straight edge razor","chain saw",
                                    "samurai sword"),
                    "Metal Objects and Tools"=c("shovel","hammer",
                                                "hatchet","metal object",
                                                "metal pole","metal pipe",
                                                "metal hand tool",
                                                "blunt object",
                                                "metal stick",
                                                "sharp object",
                                                "hand torch","ax",
                                                "pole","pick-axe",
                                                "pitchfork","baton",
                                                "crowbar","pipe",
                                                "incendiary device",
                                                "ice pick","tire iron",
                                                "spear","wrench","metal rake",
                                                "chainsaw","grenade"),
                    "Taser"=c("Taser"),
                    "BB/Air Guns"=c("BB gun","bean-bag gun","pellet gun",
                                    "air pistol","Airsoft pistol","crossbow",
                                    "bow and arrow"),
                    "Toy Weapon"=c("toy weapon"),
```

```
                    "Vehicle"=c("vehicle","motorcycle"),
                    "Everyday Objects"=c("nail gun","screwdriver",
                                         "lawn mower blade","flagpole",
                                         "cordless drill","meat cleaver",
                                         "carjack","contractor's level",
                                         "railroad spikes","stapler",
                                         "beer bottle","binoculars",
                                         "baseball bat","garden tool",
                                         "flashlight","piece of wood",
                                         "scissors","glass shard",
                                         "pepper spray","air conditioner",
                                         "chair","walking stick",
                                         "rock","barstool","bottle",
                                         "microphone","wasp spray",
                                         "brick","chain","pen","oar",
                                         "nail gun","fireworks"),
                    "Unknown Weapon"=c("undetermined","unknown weapon",
                                       "claimed to be armed"),
                    "Multiple Weapons"=c("guns and explosives",
                                         "hatchet and gun",
                                         "machete and gun",
                                         "pole and knife",
                                         "gun and knife",
                                         "gun and sword",
                                         "gun and car",
                                         "vehicle and gun",
                                         "BB gun and vehicle",
                                         "vehicle and machete",
                                         "gun and vehicle",
                                         "baseball bat and fireplace poker",
                                         "baseball bat and bottle",
                                         "gun and machete",
                                         "pole and knife",
                                         "car, knife and mace",
                                         "machete and hammer",
                                         "knife and vehicle",
                                         "baseball bat and knife"),
                    "Unarmed"=c("unarmed")))
armed_data <- mutate(armed_data,
                     armed=fct_relevel(armed,"Knife/Sword","Gun",
                                       "Metal Objects and Tools",
                                       "Multiple Weapons",
                                       "Taser","BB/Air Guns","Vehicle","Toy Weapon",
                                       "Everyday Objects","Unknown Weapon",
                                       "Unarmed"))
ggplot(data=armed_data) + geom_bar(aes(x=armed,fill=armed)) +
  scale_x_discrete(labels=NULL,breaks=NULL)+
  scale_y_continuous(expand=c(0,0)) +
  theme_classic() +
  labs(title="Fatal Shootings: Armed or Unarmed",
       x="Armed",y="Count", fill = "Armed")
```




```
#armed vs armed more collapsed
armed_data2 <- mutate(data,
                      armed = fct_collapse(armed,
                                           "Gun"=c("gun"),
                                           "Knife/Sword"=c("knife","sword",
                                                           "machete",
                                                           "box cutter",
                                                           "straight edge razor",
                                                           "chain saw",
                                                           "samurai sword"),
                                           "Other"=c("shovel","hammer","hatchet",
                                                     "metal object","metal pole",
                                                     "metal pipe",
                                                     "metal hand tool",
                                                     "blunt object",
                                                     "metal stick","sharp object",
                                                     "hand torch","ax","pole",
                                                     "pick-axe","pitchfork",
                                                     "baton","crowbar","pipe",
                                                     "incendiary device","ice pick",
                                                     "tire iron","spear",
                                                     "wrench","metal rake",
                                                     "chainsaw","grenade","Taser",
                                                     "BB gun","bean-bag gun",
                                                     "pellet gun","air pistol",
                                                     "Airsoft pistol","crossbow",
```

```
                                                "bow and arrow","toy weapon",
                                                "vehicle","motorcycle",
                                                "nail gun","screwdriver",
                                                "lawn mower blade",
                                                "flagpole","cordless drill",
                                                "meat cleaver","carjack",
                                                "contractor's level",
                                                "railroad spikes","stapler",
                                                "beer bottle","binoculars",
                                                "baseball bat","garden tool",
                                                "flashlight","piece of wood",
                                                "scissors","glass shard",
                                                "pepper spray",
                                                "air conditioner","chair",
                                                "walking stick","rock",
                                                "barstool","bottle",
                                                "microphone","wasp spray",
                                                "brick","chain","pen","oar",
                                                "nail gun","fireworks"),
                                    "Unknown Weapon"=c("undetermined",
                                                       "unknown weapon",
                                                       "claimed to be armed"),
                                    "Multiple Weapons"=c("guns and explosives",
                                                         "hatchet and gun",
                                                         "machete and gun",
                                                         "pole and knife",
                                                         "gun and knife",
                                                         "gun and sword",
                                                         "gun and car",
                                                         "vehicle and gun",
                                                         "BB gun and vehicle",
                                                         "vehicle and machete",
                                                         "gun and vehicle",
                                                         "baseball bat and fireplace poker",
                                                         "baseball bat and bottle",
                                                         "gun and machete",
                                                         "pole and knife",
                                                         "car, knife and mace",
                                                         "machete and hammer",
                                                         "knife and vehicle",
                                                         "baseball bat and knife"),
                                    "Unarmed"=c("unarmed")))
armed_data2 <- mutate(armed_data2,
                      armed=fct_relevel(armed,
                                    "Knife/Sword","Gun","Multiple Weapons",
                                    "Unknown Weapon","Unarmed","Other"))
ggplot(data=armed_data2) +
  geom_bar(aes(x=armed,fill=armed)) +
  scale_x_discrete(labels=NULL,breaks=NULL) +
  scale_y_continuous(expand=c(0,0)) +
  theme_classic() +
  labs(title="Fatal Shootings: Armed or Unarmed",x="Armed",y="Count", fill =
         "Armed")
```

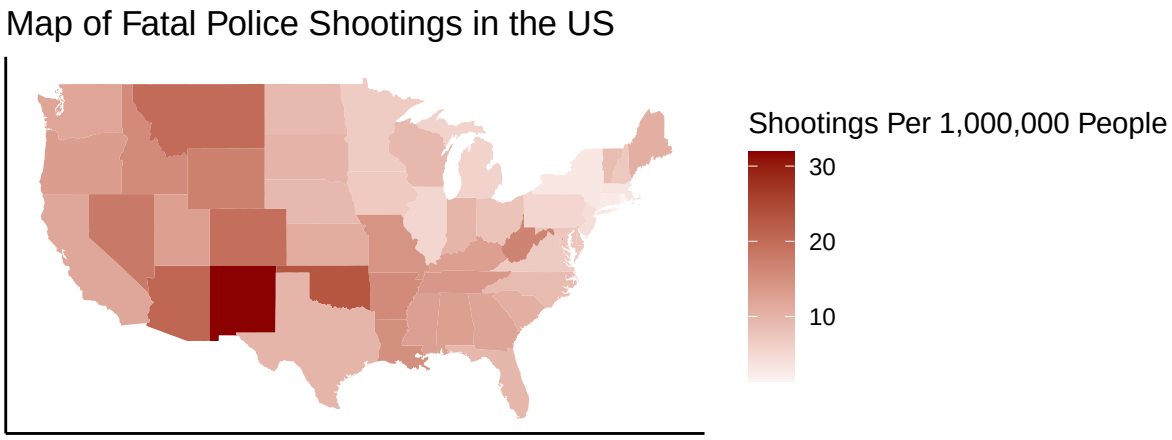


Choropleth

In order to find rate of shootings per 1,000,000 for each state, I had to use a Census API to get estimates for population breakdowns by state.

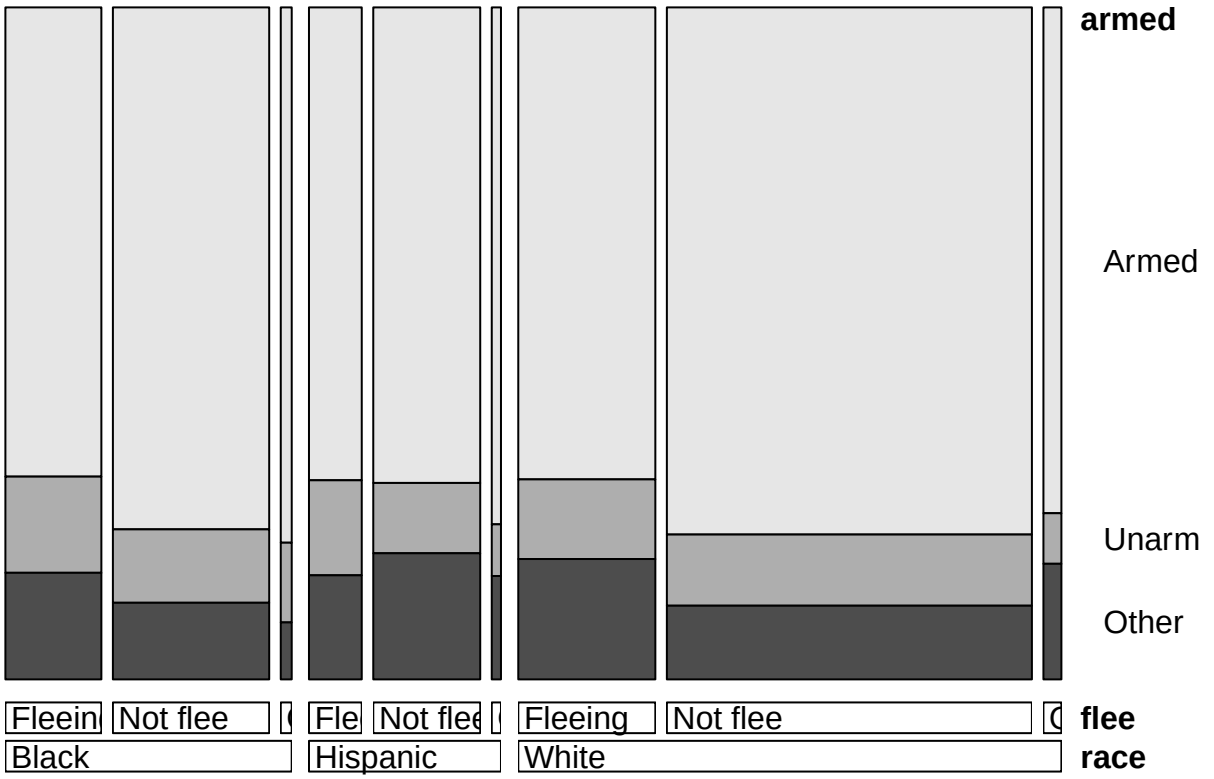
```
#turn abbreviations into names
data <- mutate(data, state = fct_recode(state, "district of columbia" = 'DC'))
data <- mutate(data, state = ifelse(state != "district of columbia",  tolower(state.name[match(state,sta
stateCounts <- count(data, state)
population <- get_estimates(geography = "state", product = "characteristics", breakdown = "SEX", year =
population <- mutate(population, NAME = tolower(NAME))
population <- summarize(group_by(population, NAME), 'population' = sum(value))
fullState <- left_join(stateCounts, population, by=c("state"="NAME"))
fullState <- mutate(fullState, proportion = n/population*1000000)
all_states <- map_data("state")
stateData <- inner_join(all_states, fullState, by=c("region" = "state"))

ggplot() +
  geom_polygon(data=stateData,
               aes(x = long, y = lat, group = group, fill =proportion)) +
  coord_map() +
  labs(x = "", y = "", title = "Map of Fatal Police Shootings in the US") +
  theme_classic() +
  theme(axis.ticks.y = element_blank(), axis.text.y = element_blank(),
        axis.ticks.x = element_blank(), axis.text.x = element_blank()) +
```

```
  scale_fill_gradient2(name = "Shootings Per 1,000,000 People", low =
                         "whitesmoke", high = "darkred")
```

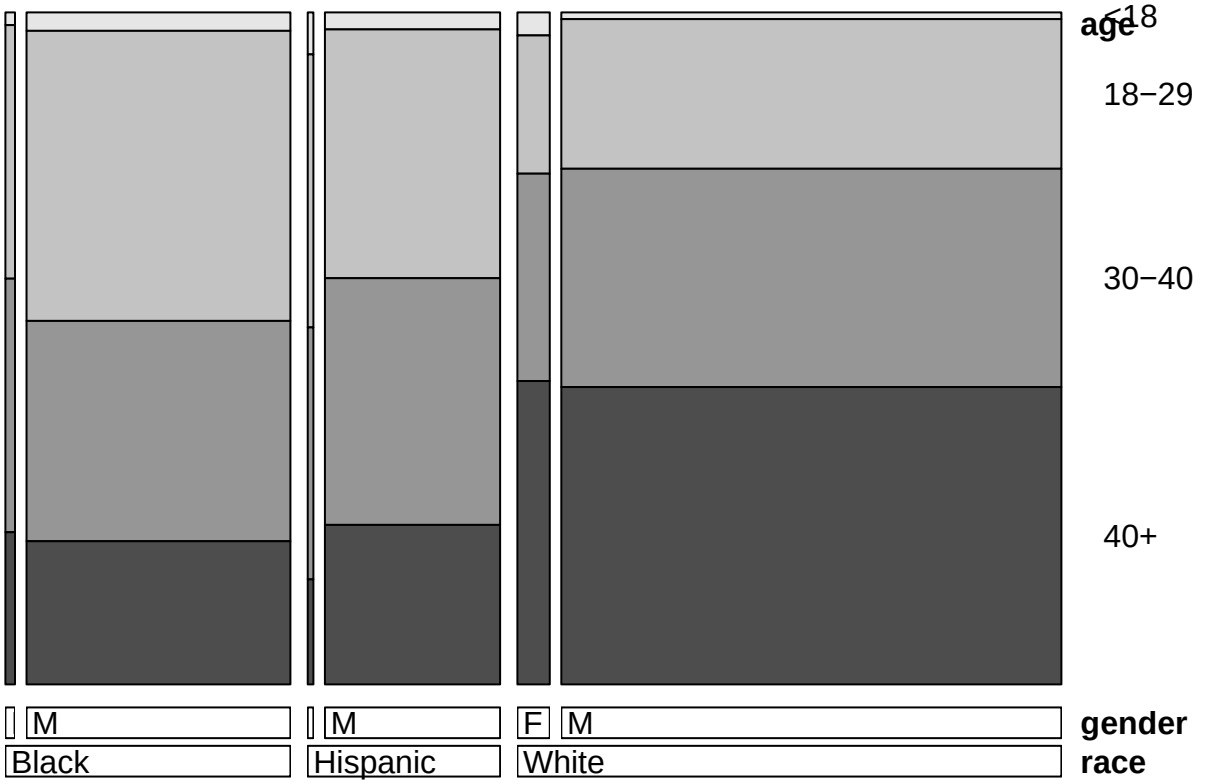


Doubledecker

```
one <- select(data, race, armed, flee)
one <- filter(one, race == 'W'| race == 'B'| race == 'H')
one <- mutate(one, race = as.factor(race), armed = as.factor(armed), flee = as.factor(flee))
one <- mutate(one, race = fct_recode(race, 'White' = 'W','Black' = 'B','Hispanic' = 'H'))
one <- mutate(one, armed = fct_collapse(armed, 'Armed' = c('gun','knife','sword'),
                                                'Unarm' = c('unarmed', 'toy weapon'),
                                                other_level = 'Other'))
one <- mutate(one, flee = fct_collapse(flee, 'Fleeing' = c('Car','Foot'),
                                              'Not flee' = 'Not fleeing',
                                              other_level = 'Other'))
one_table <- xtabs(~race+flee+armed, data=one)
doubledecker(one_table, spacing=spacing_highlighting(start = unit(0.2, "lines"), rate = 1.5))
```




```
two <- select(data, race, age, gender)
two <- filter(two, race == 'W'| race == 'B'| race == 'H')
two <- mutate(two, race = as.factor(race), age = as.factor(age), gender = as.factor(gender))
two <- mutate(two, gender = fct_recode(gender, 'M' = 'Male', 'F' = 'Female'))
two <- mutate(two, race = fct_recode(race, 'White' = 'W',
                                           'Black' = 'B',
                                           'Hispanic' = 'H'))
two <- mutate(two, age = fct_collapse(age,
                                      '<18' = c('6','8','12','13',
                                                '14','15','16','17'),
                                      '18-29' = c('18','19','20','21','22',
                                                  '23','24','25','26','27',
                                                  '28','29'),
                                      '30-40' = c('30','31','32','33',
                                                  '34','35','36','37',
                                                  '38','39', '40'),
                                      other_level = '40+'))
two_table <- xtabs(~race+gender+age, data=two)
doubledecker(two_table)
```



Conclusion

Our final analysis shows a bleak truth of the reality that we live in: Black and Hispanic young men face a higher chance to be fatally shot by the police, despite being a smaller proportion of the US population. The choropleth shows that New Mexico shows a high rate of fatal police shootings. The histogram that was proportional to the racial population size demonstrates how as a race, Black and Hispanic people are killed at a disproportionate rate. Further breakdowns of gender show that young men are more likely to be killed. We also see that an overwhelming majority of fatal police shootings occur when the victim is not fleeing, bringing about questions of use of force and de-escalation tactics. The saddest fact lies in our doubledecker of race, age, and gender: the highest proportion of fatal police shootings in boys under 18 were in the Black population.