Making Sense of The Opioid Crisis

Ray Nzaba and Shukry Zablah

May 3, 2018

Abstract

According to the U.S. Drug Enforcement Administration, "overdose deaths, particularly from prescription drugs and heroin, have reached epidemic levels". From 1999 to 2016, more than 630,000 people have died from a drug overdose. On average, 115 Americans die every day from an opioid overdose.[1]

Introduction

The opioid crisis eats up the successful entrepreneur and the jobless person alike. It is a pervasive problem that has persisted and been responsible for the deaths of thousands and the decreasing life expectancy of Americans. While it is true that every strata of society is in danger from the effects of this epidemic, its effects vary widely across age groups and gender. Our project is significant because it will be an accessible way to know more about this silent disaster. Our aim is that people that don't know about the crisis will be able to talk knowledgeably about what it is and how it affects society, and that those who already know about it will be able to correct and enhance their understanding of the numbers behind the situation.

The first wave began with increased prescribing of opioids in the 1990s[2], with overdose deaths involving prescription opioids (natural and semi-synthetic opioids and methadone) increasing since at least 1999[3]. The third wave began in 2013, with significant increases in overdose deaths involving synthetic opioids – particularly those involving illicitly-manufactured fentanyl (IMF). The IMF market continues to change, and IMF can be found in combination with heroin, counterfeit pills, and cocaine[4].

Data

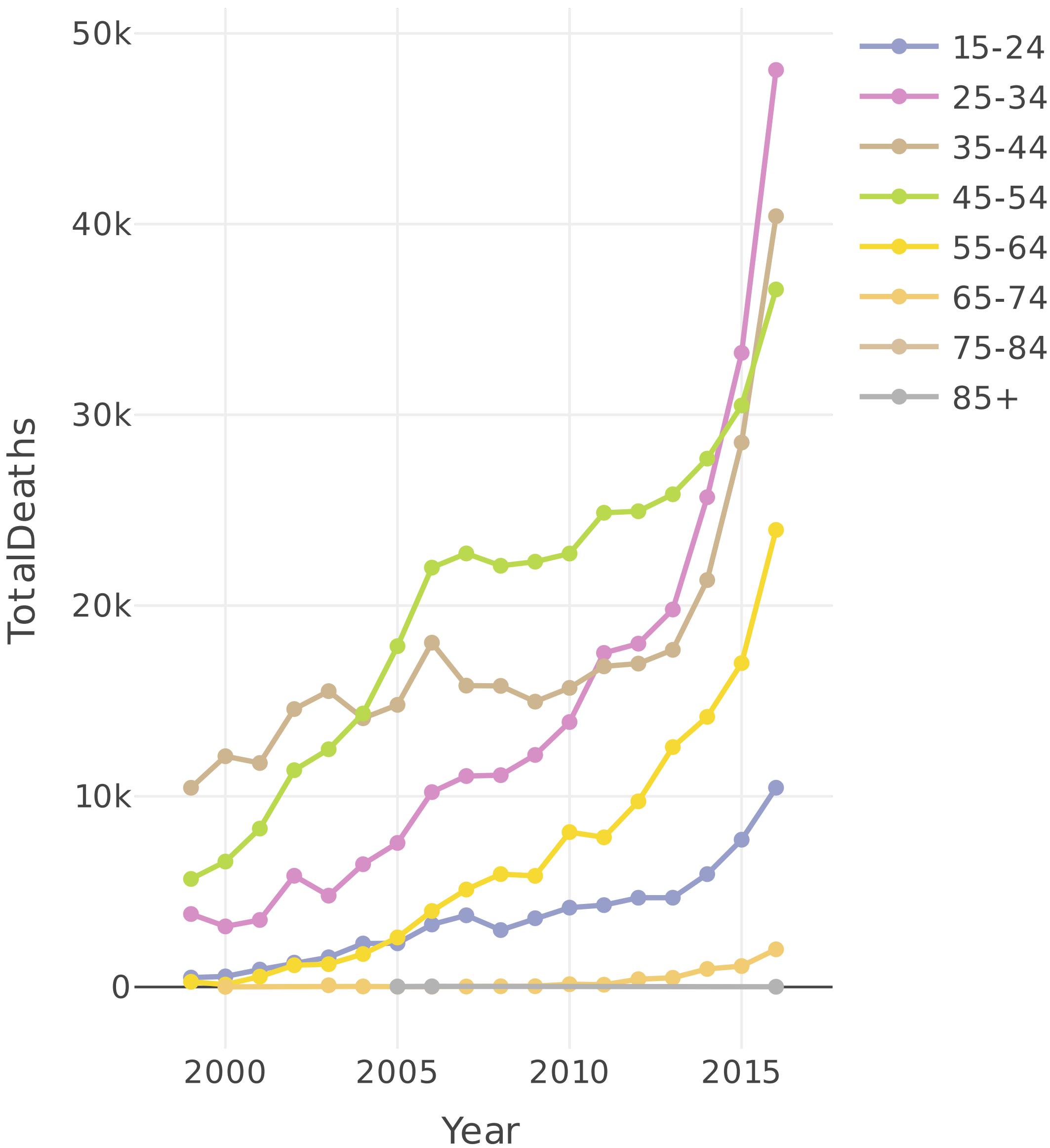
The data we used for our analysis was obtained from the following sources:

1. The CDC. This is an online database belonging to the CDC. It contains wide ranging data for Epidemiologic Research. We tried using R packages to query the website's API, but they were not as user friendly as we hoped. Eventually we resorted to performing direct queries on the website itself, but this had its own limitations in terms of size. Forty queries were done (each had a limit of 75,000 lines) to obtain 1999 - 2016 data on drug poisoning mortality rates stratified by state and county.
2. Data world website. This dataset from the The Centers for Medicare & Medicaid Services contains the information on the individual opioid prescribing rates of health providers that participate in Medicare Part D program. This file provides data on the number and percentage of prescription claims (includes new prescriptions and refills) for opioid drugs, and contains information on each provider's name, specialty, state, and ZIP code. Available documentation can be found here.

The readin instructions can be found in the technical appendix section.

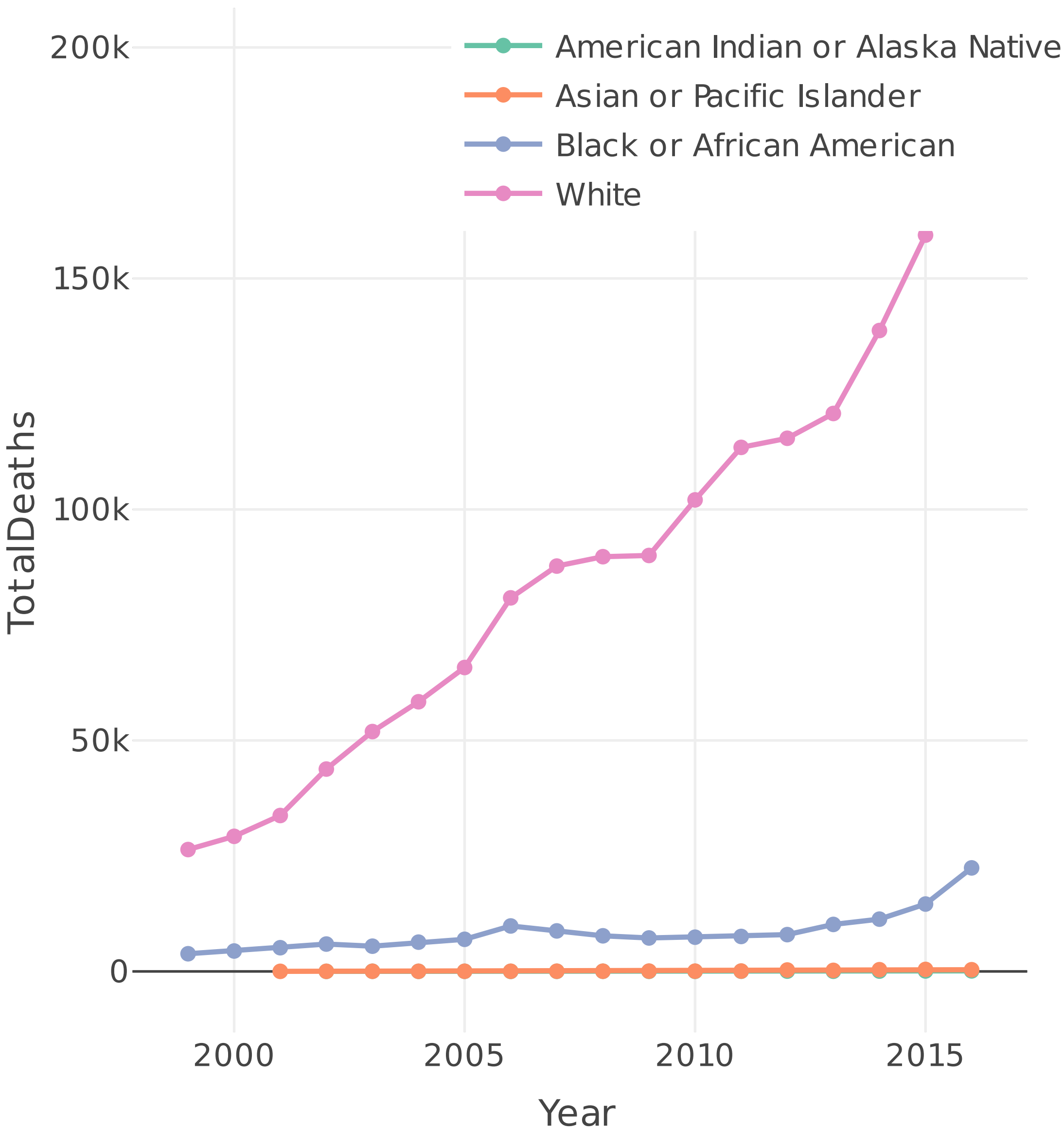
Interesting findings:

In the plot below we can see the drug overdoses through time divided by age group:



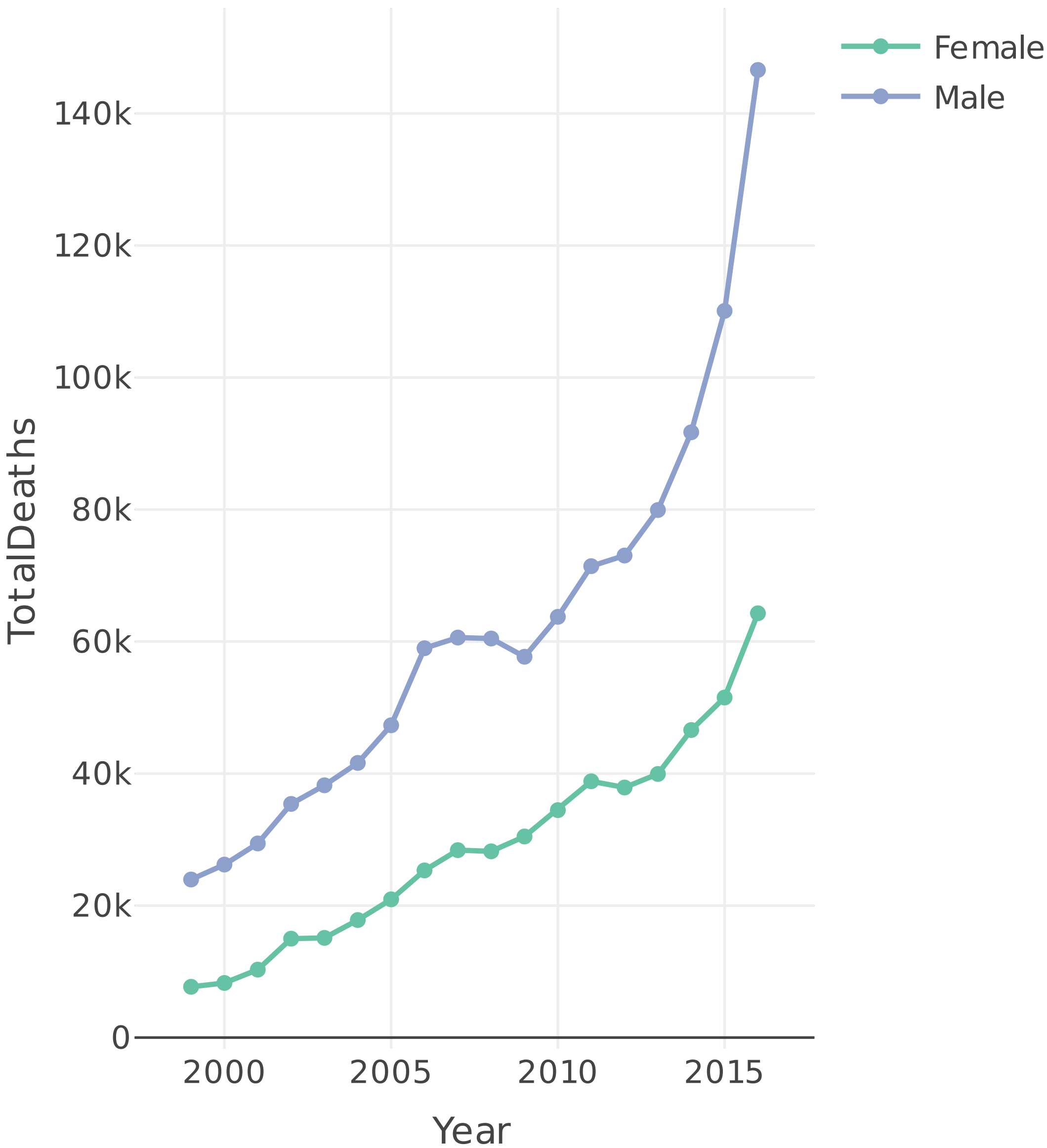
We can see how early in the 2000s, the age groups that were most affected by drug related deaths were individuals that were 45 years or older. As the crisis expanded, and with the introduction of synthetic opioids such as Fentanyl in the early 2010s, the 25-34 age group became the most affected age group. Generally, all the age groups have experienced an upward trend, but this younger age group's change is a good insight to keep in mind when we are finding ways to implement policies to address the drug crisis.

Let's take a look at the same trend but divided by race:



A clear conclusion from this figure is the fact that the opioid crisis predominantly affects the White population of the United States. However, the fact that the Black community's deaths have more than sixtupled since 1999 cannot be overlooked. It is important to consider how the different races are affected by the problem.

Now let's consider how the problem affects males and females.



Over the years, the problem has been affecting both genders. The male sex has stayed at a constant level above the female sex with roughly twice as many deaths.

The problem is clearly a national problem, but it affects the young adult age group the most, and the White race the most. Is the problem worse in certain regions of the country? Below is a map that shows the top 10 counties that were most affected by the epidemic in 2016.

The midwest is the area that has been severely hit by the opioid epidemic.

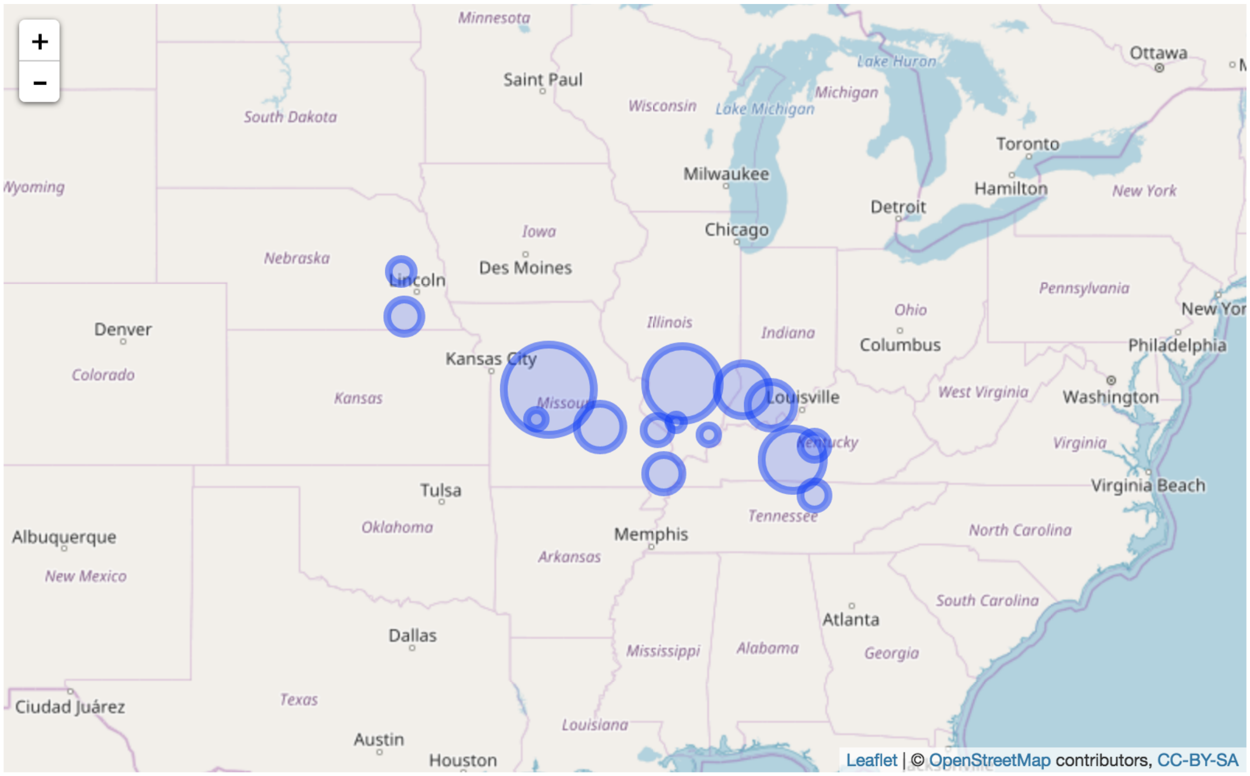
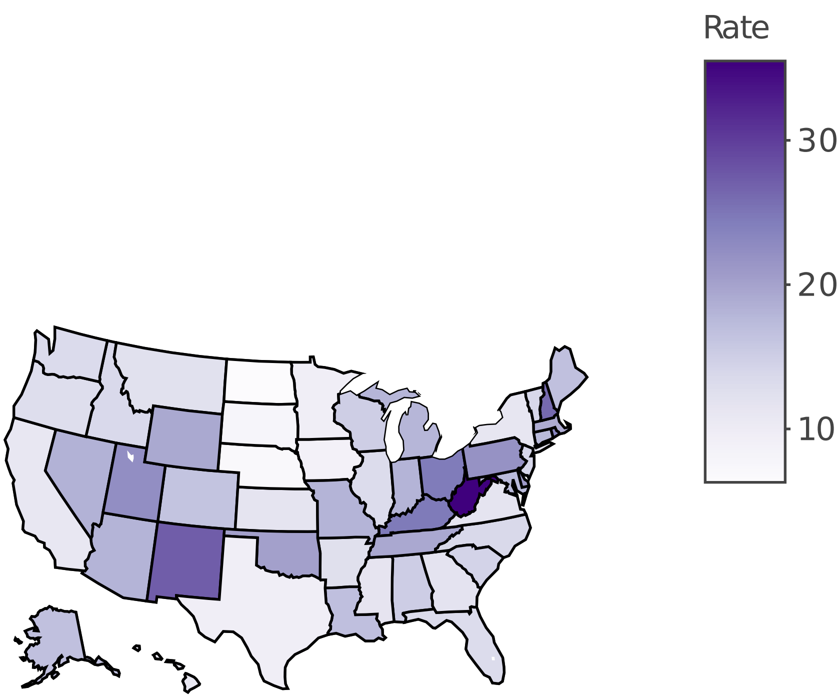


Figure 1:

Average death rates breakdown



The choropleth map above is shaded in propotion to the number of age adjusted deaths in all states in 2014.

Some of the cities with the highest Opioid prescription rates were Sacramento (CA), Piedmont (SD) and Stringtown (OK). The rates and prescribers tables from the appendix show the breakdown of prescription rates that we reported in the 2014 case we looked at in our Shiny application.

Limitations

Inasmuch as we did a really good job in honing in our research on deaths caused by drug poisoning and overdoses, we feel like the opioid crisis is a multifaceted problem involving policing and a whole drug trafficking and distribution industry. If we had more time and resources, the next step in this project would have been to obtain data that would ideally would enable us to get an enriched perspective on the problem using new found figures from drug prescription and arrest information.

Conclusion

Our final product is a tool aided by a series of navigation buttons that enable one to change between different slides in an intuitive way as they break down this sensitive topic and advance through the loop around the slideshow containing different reactive elements, plots and pictures showing the severity of the opioid epidemic. Our app is posted on the internet and can be used by anyone as a tool to inform and educate.

Technical Appendix

```r
#Imports necessary packages
library(tidyverse)
library(ggplot2)
library(readr)
library(stringr)
library(maps)
library(ggmap)
library(mapproj)
library(leaflet)
library(httr)
library(plotly)
library(leaflet)
library(plotly)
library(RColorBrewer)
```

```r
#Reads in the data that is grouped by Year, County, Drug Cause, and Age Groups for 1999-2016
YearCountyCauseAgeGroupsAllYears <- as_tibble()
for (year in 1999:2016) {
  YearCountyCauseAgeGroupsAllYears <- rbind(YearCountyCauseAgeGroupsAllYears,
                                            read_csv(paste0("../Data/WonderData/YearCountyCauseAgeGroups
}
```

```r
#Reads in the data that is grouped by Year, County, Drug Cause, and Gender for 1999-2016
YearCountyCauseGenderAllYears  <- as_tibble()
for (part in 1:9) {
  YearCountyCauseGenderAllYears <- rbind(YearCountyCauseGenderAllYears,
                                         read_csv(paste0("../Data/WonderData/YearCountyCauseGender10Sta
}
```

```r
#Reads in the data that is grouped by Year, County, Drug Cause, and Race for 1999-2016
YearCountyCauseRaceHomicideAllYears  <- as_tibble()
for (part in 1:6) {
  YearCountyCauseRaceHomicideAllYears <- rbind(YearCountyCauseRaceHomicideAllYears,
                                         read_csv(paste0("../Data/WonderData/YearCountyCauseRaceHomicid
}

#Reads in the data that is grouped by Year, County, Drug Cause, and Race for 1999-2016
YearCountyCauseRaceAllYears <- rbind(read_csv("../Data/WonderData/YearCountyCauseRaceUndeterminedAllYea
                                     read_csv("../Data/WonderData/YearCountyCauseRaceSuicidesAllYears.c
                                     read_csv("../Data/WonderData/YearCountyCauseRaceUnintentional1999-
                                     read_csv("../Data/WonderData/YearCountyCauseRaceUnintentional2008-
                                     YearCountyCauseRaceHomicideAllYears)
```

```r
X2015_Gaz_counties_national <- read_csv("../Data/WonderData/2015_Gaz_counties_national.csv")
```

```r
PState <- read_csv("../Data/Drug_Poisoning_Mortality_by_State__United_States.csv")
prescription<- read_csv("../Data/Opioid analgesic prescriptions dispensed from US retail pharmacies, Q4
zip_codes <- read_csv("../Data/zip_codes_states.csv")
stateabbr <- read.csv("https://raw.githubusercontent.com/plotly/datasets/master/2011_us_ag_exports.csv"
load("../Data/prescriber.Rda")
```

```r
#cleans up the data that has age groups
ageGroupsData <- YearCountyCauseAgeGroupsAllYears %>%
```

```r
  mutate(percentTotalDeaths = parse_number(`% of Total Deaths`),
         crudeRateLower95Confint = parse_number(`Crude Rate Upper 95% Confidence Interval`),
         crudeRateUpper95Confint = parse_number(`Crude Rate Lower 95% Confidence Interval`),
         crudeRate = parse_number(`Crude Rate`),
         population = parse_number(Population)) %>%
  separate(County, c("County", "State"), sep = ",") %>%
  left_join(X2015_Gaz_counties_national, by = c("County" = "NAME")) %>%
  rename(drugAlcoholInducedCause = `Drug/Alcohol Induced Cause`,
         ageGroups = `Ten-Year Age Groups Code`,
         latitude = INTPTLAT,
         longitude = INTPTLONG) %>%
  select(Year, County, State, drugAlcoholInducedCause, ageGroups, Deaths, population, crudeRate, crudeRa
  mutate(State = str_trim(State),
         ageGroups = ifelse(ageGroups == "1", "<1", ageGroups),
         ageGroups = factor(ageGroups, levels = c("<1", "1-4", "5-14", "15-24", "25-34", "35-44", "45-5
```

```r
#cleans up the data that has gender
genderData <- YearCountyCauseGenderAllYears %>%
   mutate(percentTotalDeaths = parse_number(`% of Total Deaths`),
         crudeRateLower95Confint = parse_number(`Crude Rate Upper 95% Confidence Interval`),
         crudeRateUpper95Confint = parse_number(`Crude Rate Lower 95% Confidence Interval`),
         crudeRate = parse_number(`Crude Rate`),
         population = parse_number(Population),
         ageAdjRate = parse_number(`Age Adjusted Rate`),
         gender = as.factor(Gender))%>%
  separate(County, c("County", "State"), sep = ",") %>%
  left_join(X2015_Gaz_counties_national, by = c("County" = "NAME")) %>%
  rename(drugAlcoholInducedCause = `Drug/Alcohol Induced Cause`,
         latitude = INTPTLAT,
         longitude = INTPTLONG) %>%
  select(Year, County, State, drugAlcoholInducedCause, Deaths, population, crudeRate, crudeRateLower95C
  mutate(State = str_trim(State))
```

```r
#cleans up the data that has race
raceData <- YearCountyCauseRaceAllYears %>%
  mutate(percentTotalDeaths = parse_number(`% of Total Deaths`),
         Year = parse_number(Year),
         crudeRateLower95Confint = parse_number(`Crude Rate Upper 95% Confidence Interval`),
         crudeRateUpper95Confint = parse_number(`Crude Rate Lower 95% Confidence Interval`),
         crudeRate = parse_number(`Crude Rate`),
         population = parse_number(Population),
         ageAdjRate = parse_number(`Age Adjusted Rate`),
         ageAdjRateLowerConfint = parse_number(`Age Adjusted Rate Lower 95% Confidence Interval`),
         ageAdjRateUpperConfint = parse_number(`Age Adjusted Rate Upper 95% Confidence Interval`),
         race = as.factor(Race),
         Deaths = parse_number(Deaths))%>%
  separate(County, c("County", "State"), sep = ",") %>%
  left_join(X2015_Gaz_counties_national, by = c("County" = "NAME")) %>%
  rename(drugAlcoholInducedCause = `Drug/Alcohol Induced Cause`,
         latitude = INTPTLAT,
         longitude = INTPTLONG) %>%
  select(Year, County, State, drugAlcoholInducedCause, Deaths, population, crudeRate, crudeRateLower95C
  mutate(State = str_trim(State))
```

```r
#Select and rename relevant columns
PState_wrangled <-PState %>%
  select(State,Year,Sex,`Age Group`,`Race and Hispanic Origin`,`Crude Death Rate` ,`Age-adjusted Rate`)
  mutate(Age_Group=`Age Group`,Race=`Race and Hispanic Origin`,Crude_Death_Rate=`Crude Death Rate`,Age_A
  select(-`Age Group`,-`Race and Hispanic Origin`,-`Crude Death Rate`,-`Age-adjusted Rate`)
```

```r
#Parse and make column names more readable
presc<-prescription %>%
  mutate(Yearly_totals_all_opioid_analgesics = as.integer(`Yearly totals (All Opioid Analgesics)`),
         Yearly_totals_HOTMF = as.integer(`Yearly totals (H+O+T+M+F)`),
        Yearly_totals_HO = as.integer(`Yearly totals (H+O)`,
        Yearly_totals_ER_LA_opioid_analgesics = as.integer(`Yearly totals (ER/LA Opioid Analgesics)`)))
```

```r
#top opioid prescribers
prescribers<-prescriber %>%
  separate(`Opioid Prescribing Rate`, c("Opioid_Prescribing_Rate","junk"), sep = "%") %>%
  mutate(Opioid_Prescribing_Rate=as.integer(Opioid_Prescribing_Rate)) %>%
  select(-junk,-NPI) %>%
  group_by(`Specialty Description`) %>%
  summarize(count = n(),Opioid_Prescribing_Rate=sum(Opioid_Prescribing_Rate)/count) %>%
  drop_na() %>%
  arrange(desc(Opioid_Prescribing_Rate)) %>%
  head(4)
```

```r
#opioid prescription rates by zip code
rates<-prescriber %>%
  separate(`Opioid Prescribing Rate`, c("Opioid_Prescribing_Rate","junk"), sep = "%") %>%
  mutate(Opioid_Prescribing_Rate=as.integer(Opioid_Prescribing_Rate),
         NPPES_zip_code=as.character(`NPPES Provider Zip Code`)) %>%
  select(-junk,-NPI,-`NPPES Provider Zip Code`) %>%
  group_by(NPPES_zip_code) %>%
  summarize(count=n(),Opioid_Prescribing_Rate=sum(Opioid_Prescribing_Rate)/count) %>%
  arrange(desc(Opioid_Prescribing_Rate))
```

```r
#clean table containing state abbreviations
stateabbr <- stateabbr %>%
  select(code,state)
```

```r
#Consolidate zipcode and prescribing rate data, death and drug info
rates<-rates %>%
  left_join(zip_codes,by=c("NPPES_zip_code" = "zip_code")) %>%
  na.omit() %>%
  group_by(state) %>%
  summarize(count=n(), Opioid_Prescribing_Rate=sum(Opioid_Prescribing_Rate)/count) %>%
  arrange(desc(Opioid_Prescribing_Rate)) %>%
  head(4)
```

```r
#Wrangle state death records to display on map
PS<-PState_wrangled %>%
  filter(Year==2014) %>%
  group_by(State) %>%
  summarize(count=n(),Crude_Death_Rate=sum(Crude_Death_Rate)/count,
            Age_Adjusted_Rate=sum(Age_Adjusted_Rate)/count) %>%
  filter(State!="United States") %>%
  left_join(stateabbr,by=c("State" = "state")) %>%
  select(State,Crude_Death_Rate,code,Age_Adjusted_Rate) %>%
```

```
  mutate(code=replace(code, State=="California", "CA"),Crude_Death_Rate=round(Crude_Death_Rate, 1),
         Age_Adjusted_Rate=round(Age_Adjusted_Rate, 1))%>%
  left_join(rates,by=c("code" = "state"))%>%
  mutate(Opioid_Prescribing_Rate=round(Opioid_Prescribing_Rate,1))
```

```
ageGroupsData %>%
  filter(ageGroups != "NS", Year == 2016, drugAlcoholInducedCause == "Drug poisonings (overdose) Uninte
  group_by(Year, ageGroups) %>%
  summarize(numObs = n(), TotalDeaths = sum(Deaths)) %>%
  plot_ly(x = ~ageGroups, y = ~TotalDeaths, type = "bar", name = "bargraph")
```

```
raceData %>%
  filter(Year == 2016, drugAlcoholInducedCause == "Drug poisonings (overdose) Unintentional (X40-X44)")
  group_by(Year, race) %>%
  summarize(numObs = n(), TotalDeaths = sum(Deaths)) %>%
  plot_ly(x = ~race, y = ~TotalDeaths, type = "bar", color = ~race)
```

```
genderData %>%
  filter(Year == 2016, drugAlcoholInducedCause == "Drug poisonings (overdose) Unintentional (X40-X44)")
  group_by(Year, gender) %>%
  summarize(numObs = n(), TotalDeaths = sum(Deaths)) %>%
  plot_ly(x = ~gender, y = ~TotalDeaths, type = "bar", color = ~gender)
```

```
ageGroupsData %>%
  group_by(ageGroups, Year) %>%
  filter(drugAlcoholInducedCause == "Drug poisonings (overdose) Unintentional (X40-X44)"|
           drugAlcoholInducedCause == "All other drug-induced causes"|
           drugAlcoholInducedCause == "Drug poisonings (overdose) Suicide (X60-X64)"|
           drugAlcoholInducedCause == "Drug poisonings (overdose) Undetermined (Y10-Y14)") %>%
  summarize(numObs = n(), TotalDeaths = sum(Deaths)) %>%
  plot_ly(x = ~Year, y = ~TotalDeaths, type = 'scatter', color = ~ageGroups, mode = "lines+markers")
```

```
raceData %>%
  group_by(race, Year) %>%
  filter(drugAlcoholInducedCause == "Drug poisonings (overdose) Unintentional (X40-X44)"|
           drugAlcoholInducedCause == "All other drug-induced causes"|
           drugAlcoholInducedCause == "Drug poisonings (overdose) Suicide (X60-X64)"|
           drugAlcoholInducedCause == "Drug poisonings (overdose) Undetermined (Y10-Y14)") %>%
  summarize(numObs = n(), TotalDeaths = sum(Deaths)) %>%
  plot_ly(x = ~Year, y = ~TotalDeaths, type = 'scatter', color = ~race, mode = "lines+markers")
```

```
genderData %>%
  group_by(gender, Year) %>%
  filter(drugAlcoholInducedCause == "Drug poisonings (overdose) Unintentional (X40-X44)"|
           drugAlcoholInducedCause == "All other drug-induced causes"|
           drugAlcoholInducedCause == "Drug poisonings (overdose) Suicide (X60-X64)"|
           drugAlcoholInducedCause == "Drug poisonings (overdose) Undetermined (Y10-Y14)") %>%
  summarize(numObs = n(), TotalDeaths = sum(Deaths)) %>%
  plot_ly(x = ~Year, y = ~TotalDeaths, type = 'scatter', color = ~gender, mode = "lines+markers")
```

```
myTemp <- ageGroupsData %>%
  group_by(County) %>%
  summarize(NumObs = n(), SumDeaths = sum(Deaths), avgLong = mean(longitude), avgLat = mean(latitude)) 
  arrange(desc(SumDeaths)) %>%
  head(25)
```

```r
myTemp

leaflet(myTemp) %>%
  addTiles() %>%
  addCircles(lng = ~avgLong, lat = ~avgLat, radius = ~SumDeaths/100) %>%
  setView(lng = -91.39, lat = 38.42, zoom = 5)
```

```r
#Exploratory visual showing trends in single states
PState_wrangled %>%
  filter(State=="Wyoming") %>%
ggplot(.,aes(x=Year,y=Crude_Death_Rate,color=Sex))+geom_line()+facet_wrap(~Race)
```

```r
# give state boundaries a white border
l <- list(color = toRGB("white"), width = 2)
```

```r
# specify some map projection/options for plotly map
g <- list(
  scope = 'usa',
  projection = list(type = 'albers usa'),
  showlakes = TRUE,
  lakecolor = toRGB('white')
)
```

```r
#Customize tooltip for death rates and prescriptions map
PS$hover <- with(PS, paste(State,'<br>',"Crude death rate:", Crude_Death_Rate,'<br>',"Opioid Prescribing
```

```r
#Make death rates and prescriptions plot
plot_geo(PS, locationmode = 'USA-states') %>%
  add_trace(
    z = ~Age_Adjusted_Rate, text = ~hover, locations = ~code,
    color = ~Age_Adjusted_Rate, colors = 'Purples'
  ) %>%
  colorbar(title = "Rate") %>%
  layout(
    title = 'Average death rates breakdown ',
    geo = g
  )
```

```r
#Shows crude death rates by state
br_down<-PState_wrangled %>%
  filter(State=="United States") %>%
  group_by(State)
```

```r
#Countrywide breakdown by race
br_down %>%
  filter(Sex=="Female",Age_Group=="All Ages") %>%
ggplot(.,aes(x=Year,y=Crude_Death_Rate))+geom_line()+facet_wrap(~Race)
```