EDA-Base-Script

PSTAT 296A

2025-10-14

Import packages

```r
library(tidyverse)
library(ggplot2)
library(dplyr)
library(tseries)
library(forecast)
library(ggfortify)
```

Background Research

(source: https://www.sciencedirect.com/science/article/pii/S0955395919300180)

Each wave is driven by a surge in popularity for a certain type of opioid substance, followed by restricter regulations and monitoring programs that drive death rates to go down a little, before the cycle repeats again. Each wave has a different target demographic and will be found to be more prevalent in a certain age group, geographic area, gender, etc.

Wave 1 (~ 1990s - 2010): Rise of prescription opioids including Oxycontin

Wave 2 (~ 2010 - 2013): Rise of heroin

Wave 3 (~ 2013 - present): Rise of synthetic opioids like Fentanyl

EDA - Age

```r
# Analyze age distribution (US opioid rate age.csv)
age_data <- read.csv("Data/US opioid rate age.csv")

# Check unique values of age bins
unique(age_data$age)
```

```
##  [1] "15-19 years" "20-24 years" "25-29 years" "30-34 years" "35-39 years"
##  [6] "40-44 years" "45-49 years" "50-54 years" "55-59 years" "60-64 years"
```

```r
# Plot the distribution of opioid death rates by age
ggplot(age_data, aes(x = year, y = val, color = age)) +
  geom_line(linewidth = 1.2) +
```

```
  geom_point(size = 1.8) +
  labs(
    title = "Opioid Mortality Rates by Age Group (1990-2000)",
    x = "Year",
    y = "Deaths per 100k",
    color = "Age Group"
  ) +
  theme_minimal(base_size = 14) +
  theme(
    legend.position = "right",
    plot.title = element_text(face = "bold")
  )
```

**Opioid Mortality Rates by Age Group (1990–2000)**
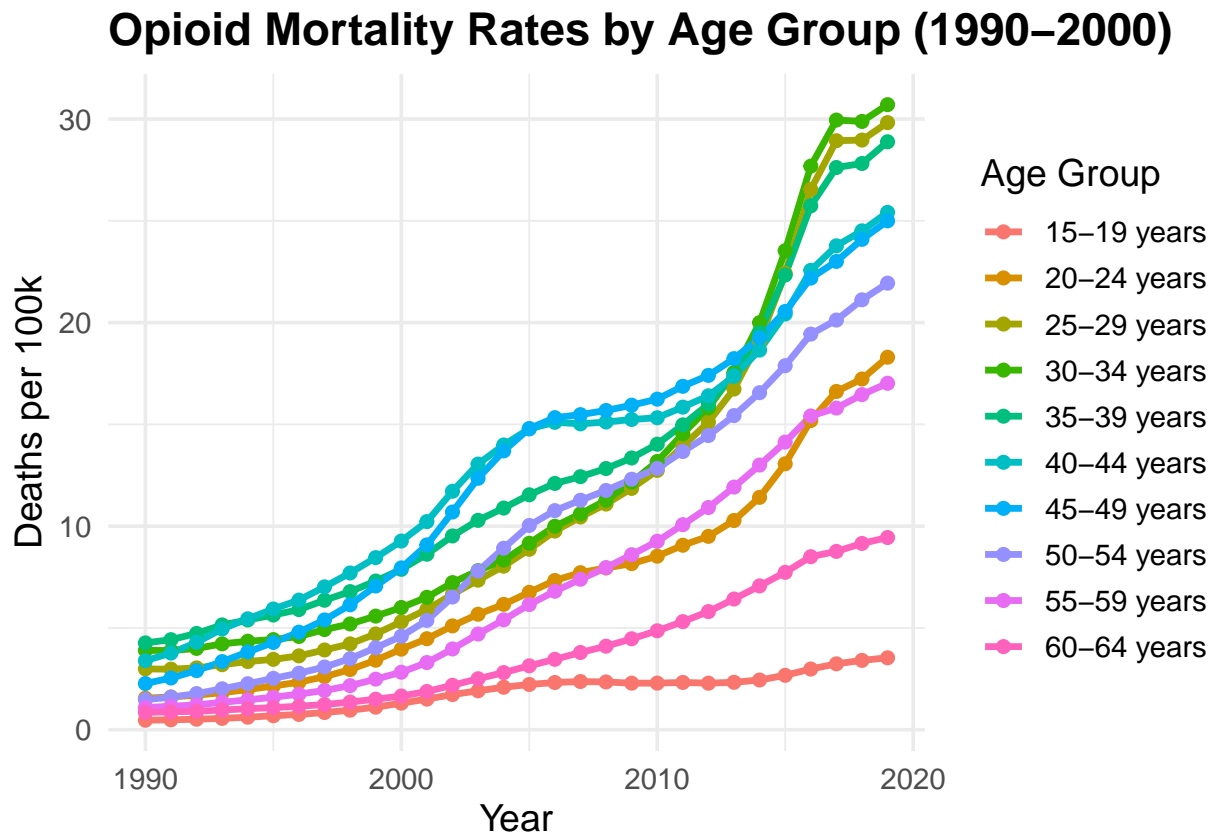
Ages from 25-39 years old show a similar trend with the highest opioid death rates. Overall, opioid death rate is trending up since 1990.

```
# Add highlights around the confidence interval
ggplot(age_data, aes(x = year, y = val, color = age, fill = age)) +
  geom_ribbon(aes(ymin = lower, ymax = upper), alpha = 0.2, color = NA) +
  geom_line(linewidth = 1.1) +
  geom_point(size = 1.5) +
  labs(
    title = "Opioid Mortality Rates by Age Group (1990-2000)",
    x = "Year",
    y = "Deaths per 100k",
    color = "Age Group",
    fill = "Age Group"
```

```
  ) +
  theme_minimal(base_size = 14) +
  theme(
    legend.position = "right",
    plot.title = element_text(face = "bold")
  )
```

Opioid Mortality Rates by Age Group (1990–2000)
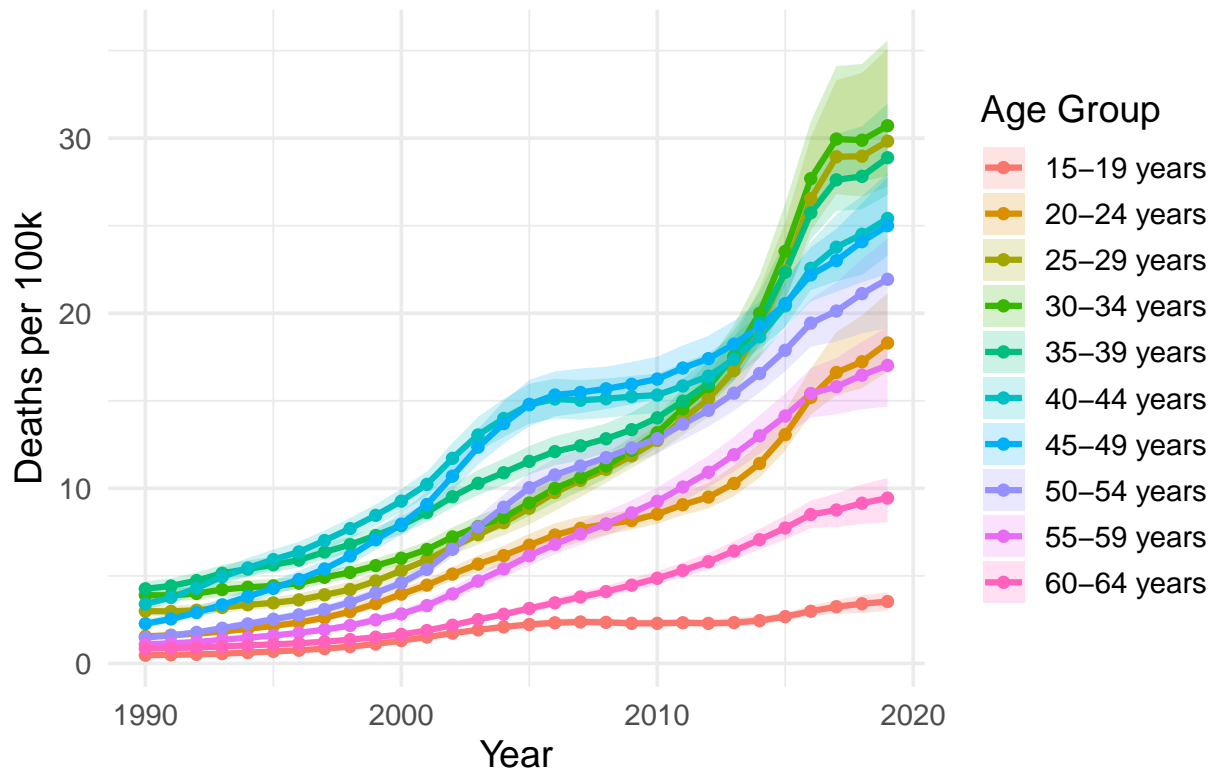
There is much more variability for opioid death rates for ages 25-39, especially in the more recent years.

```
# Compute year-to-year percent change per age group
age_change <- age_data %>%
  group_by(age) %>%
  arrange(year, .by_group = TRUE) %>%
  mutate(
    pct_change = 100 * (val - lag(val)) / lag(val)
  ) %>%
  ungroup()

ggplot(age_change, aes(x = year, y = pct_change, color = age)) +
  geom_line(linewidth = 1) +
  geom_point(size = 1.5) +
  geom_hline(yintercept = 0, linetype = "dashed", color = "gray50") +
  labs(
    title = "Year-to-Year Percent Change in Opioid Mortality Rate by Age Group",
    x = "Year",
    y = "Percent Change (%)",
    color = "Age Group"
```

```
) +
theme_minimal(base_size = 14) +
theme(
  legend.position = "right",
  plot.title = element_text(face = "bold")
)
```

Year–to–Year Percent Change in Opioid Mortality Rat