

# Final Project: Analyzing Suicide Mortality Trends by Demographic Factors

by Joseph Kim and Jun Tianzhong

2024-04-22

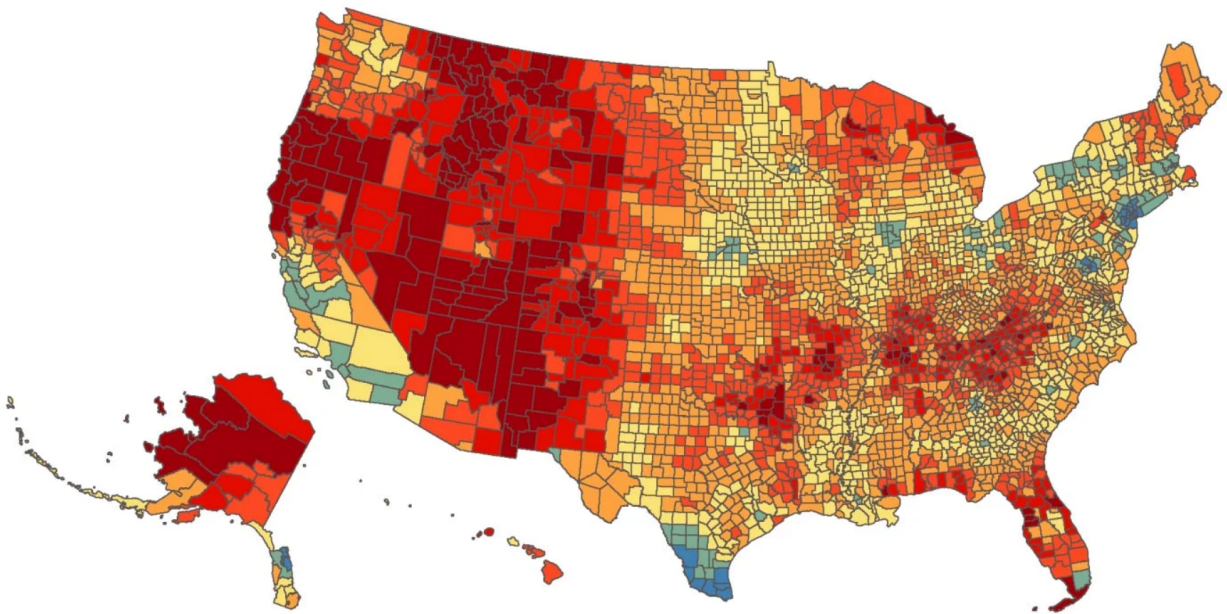


Figure 1: Suicide Rate by County in the United States in 2016.

# Abstract

## Introduction

Understanding the dynamics of suicide rates across demographic categories is paramount in addressing the complexities of mental health and societal well-being. Over the decades, the United States has witnessed fluctuations in suicide rates (increased trend), influenced by a multitude of factors including societal changes, economic conditions, and mental health awareness. The exploration of these trends sheds light on the historical patterns of suicide but also guides targeted interventions and awareness to mitigate risk factors associated with suicide.

Recent years we have seen a growing recognition of the urgency to address mental health issues, including suicide prevention, at both the nation and global level. With increasing awareness, there has been an increase in research efforts aimed to dissecting the interrelationship between demographic variables and suicide rates. This includes the examination of disparities across sex, race, Hispanic origin, and age groups, acknowledging the nuanced experiences and demographic categories that are the most vulnerable.

With a data set provided by the U.S. Department of Health & Human Services titled “Death rates for suicide, by sex, race, Hispanic origin, and age: United States”, this study seeks to contribute to this ongoing research by comprehensively analyzing death rates for suicide in the United States spanning nearly seven decades, from 1950 to 2018. By analyzing into historical data, we aim to discern long-term trends and identify significant shifts in suicide rates among different demographic groups. Moreover, we want to uncover potential underlying factors contributing to these trends, informing targeted interventions and policy initiatives tailored to address the unique needs of diverse populations like the United States. In essence, we aspire to contribute towards fostering a society that has increased mental health awareness through this research.

## Data Analysis

```
data <- read_csv("Dataset/deathRates.csv", show_col_types = F)
demoCategory <- unique(data$STUB_LABEL)
ages <- unique(data$AGE)
years <- unique(data$YEAR)
demoFactor <- unique(data$STUB_NAME)

summaryDeathRate <- data %>%
  filter(STUB_NAME %in% demoFactor[1])

deathRate.summary <- ggplot(summaryDeathRate, aes(x = YEAR, y = ESTIMATE)) +
  geom_line() +
  labs(x = "Year", y = "Deaths per 100,000 Resident Population") +
  ggtitle("Death Rate Summary")

ggsave("deathSummary.pdf", deathRate.summary)

## Saving 6.5 x 4.5 in image

srDeathRate <- data %>%
  filter(STUB_NAME %in% demoFactor[3])
```

```
deathRate.sexRace <- ggplot(srDeathRate, aes(x = STUB_LABEL, y = ESTIMATE)) +
  geom_bar(stat = "identity") +
  scale_x_discrete(labels = function(x) gsub(" ", "\n", x)) +
  labs(x = "Sex-Race", y = "Deaths per 100,000 Resident Population") +
  ggtitle("Cumulative Death Count")

ggsave("CumulativeDeath.pdf", deathRate.sexRace)
```

## Saving 6.5 x 4.5 in image

## Warning: Removed 48 rows containing missing values or values outside the scale range  
## ('geom\_bar()').

```
srhDeathRate <- data %>%
  filter(STUB_NAME %in% demoFactor[4])

hispanicOrigin<- c("Female: Hispanic or Latino: All races", "Male: Hispanic or Latino: All races")
HispanicMaleDeathRate <- srhDeathRate %>%
  filter(STUB_LABEL %in% hispanicOrigin[1])
HispanicFemaleDeathRate <- srhDeathRate %>%
  filter(STUB_LABEL %in% hispanicOrigin[2])

deathRate.hispanic <- ggplot() +
  geom_line(data = HispanicMaleDeathRate, aes(x = YEAR, y = ESTIMATE, color = "Male")) +
  geom_line(data = HispanicFemaleDeathRate, aes(x = YEAR, y = ESTIMATE, color = "Female")) +
  labs(x = "Year", y = "Deaths per 100,000 Resident Population") +
  ggtitle("Death Rate of Hispanic Origin")

ggsave("HispanicDeath.pdf", deathRate.hispanic)
```

## Saving 6.5 x 4.5 in image

## Warning: Removed 16 rows containing missing values or values outside the scale range  
## ('geom\_line()').

## Warning: Removed 16 rows containing missing values or values outside the scale range  
## ('geom\_line()').

```
nonHispanicDeathRate <- srhDeathRate %>%
  filter(!(STUB_LABEL %in% hispanicOrigin))

deathRate.nonHispanic <- ggplot(nonHispanicDeathRate, aes(x = STUB_LABEL, y = ESTIMATE)) +
  geom_bar(stat = "identity") +
  scale_x_discrete(labels = function(x) gsub(" ", "\n", x)) +
  labs(x = "Sex-Race", y = "Deaths per 100,000 Resident Population") +
  ggtitle("Non-Hispanic Origin: Cumulative Death Count")

ggsave("NonHispanicDeath.pdf", deathRate.nonHispanic)
```

## Saving 6.5 x 4.5 in image

```
## Warning: Removed 120 rows containing missing values or values outside the scale range
## ('geom_bar()').
```

```
# subset_data <- data %>%
# filter(STUB_LABEL %in% categories[1], AGE %in% ages[1])
```

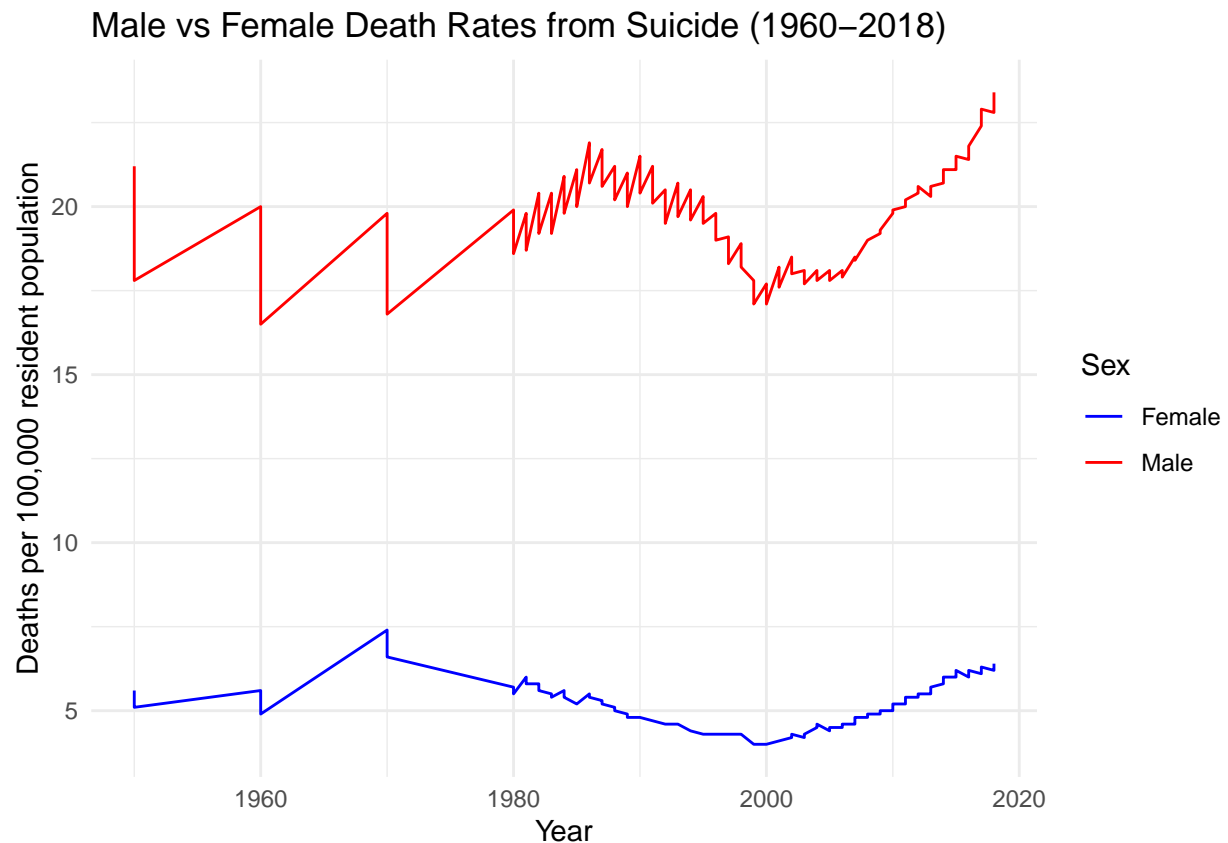
```
sexDeathRate <- data %>%
  filter(STUB_NAME %in% demoFactor[2])
male_death_rates <- sexDeathRate %>%
  filter(STUB_LABEL == "Male")
female_death_rates <- sexDeathRate %>%
  filter(STUB_LABEL == "Female")

sex_death_rate_plot <- ggplot() +
  geom_line(data = male_death_rates, aes(x = YEAR, y = ESTIMATE, color = "Male")) +
  geom_line(data = female_death_rates, aes(x = YEAR, y = ESTIMATE, color = "Female")) +
  labs(x = "Year", y = "Deaths per 100,000 resident population", color = "Sex") +
  ggtitle("Male vs Female Death Rates from Suicide (1960-2018)") +
  scale_color_manual(values = c("blue", "red")) +
  theme_minimal()

ggsave("MaleVSFemale.pdf", sex_death_rate_plot)
```

```
## Saving 6.5 x 4.5 in image
```

```
sex_death_rate_plot
```



**Conclusion**

**Team Contribution Statement**