

Examining Premature Death at the County Level: The Role of Alcohol, Income, and Education

Mason Caskey, Pilar Gibson, Erin Melia, Makenna Rodriguez

Summer Undergraduate Research Experience (SURE) 2023

Introduction

Our lifespans are possibly impacted by challenges such as **poor education**, **excessive drinking**, and **income inequality**. We wish to investigate how all of these factors are related to cumulative premature death in the US.

Data Collection at the County Level

Our dataset is the 2023 County Health Rankings & Roadmaps (CHR&R) compiled by the University of Wisconsin Population Health Institute. Some of the variables are missing more than others.

VARIABLE	MISSING	PERCENT OF TOTAL COUNTIES
Homicides	1816	57.8%
Firearm Fatalities	871	27.7%
Suicides	709	22.6%
Years Premature Death	60	1.9%
Median Household Income	2	0.1%
Excessive Drinking	2	0.1%
High School	0	0.0%

Figure 1. Many counties are missing homicide or suicide data.

We focused specifically on premature deaths and these health behavior factors at the county level.

Does Excessive Drinking Correlate with Premature Death?

Alcohol misuse is the seventh leading risk factor for premature death [2], and in our dataset, we found that excessive drinking use is **negatively associated** with premature death among adults aged 20 to 64 years.

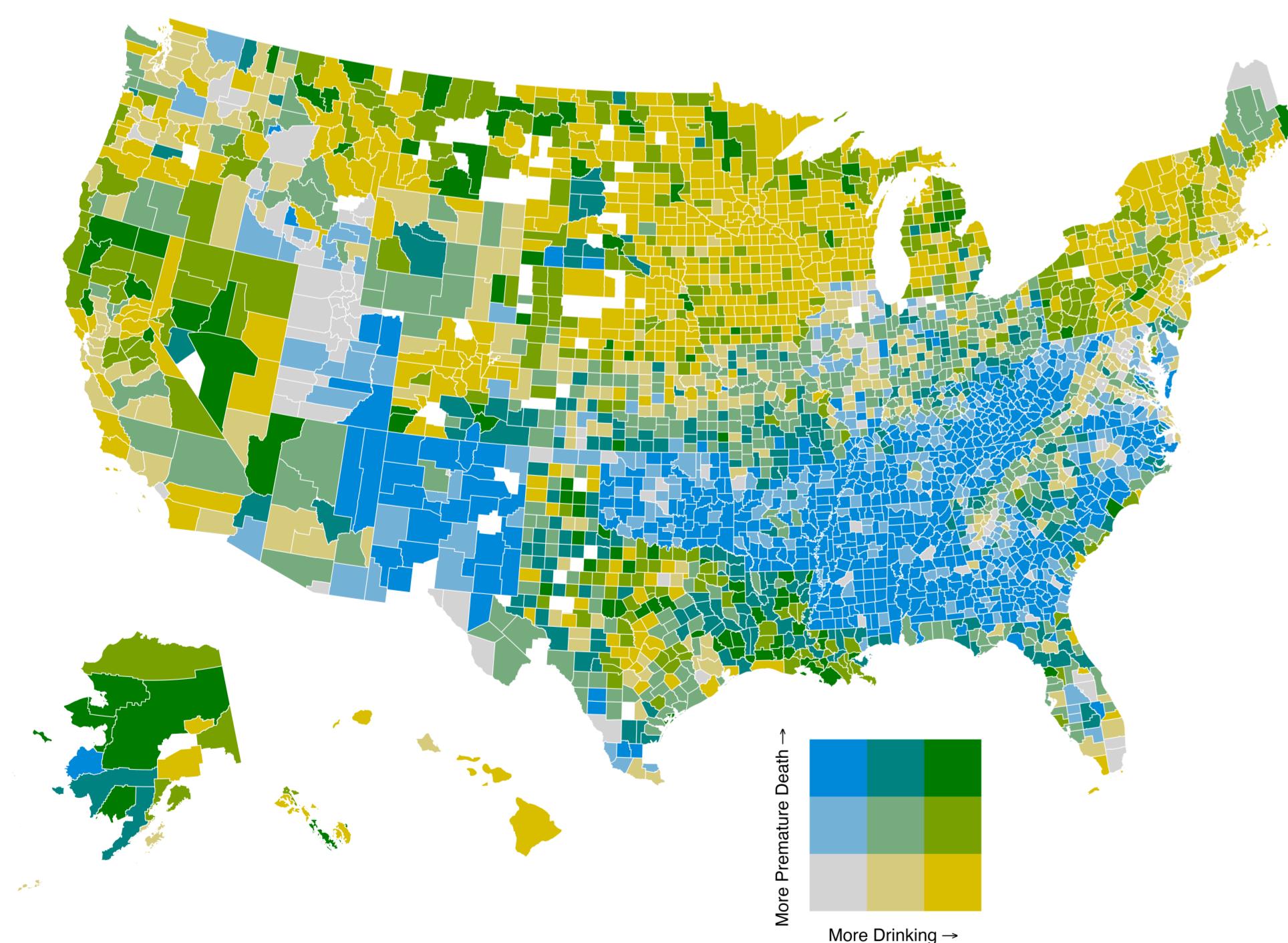


Figure 2. Bivariate Chloropleth of Excessive Drinking and Premature Deaths in the United States

Figure 2 presents the percentage of excessive drinking in U.S. counties. We hypothesized that states with a high level of binge drinking would lead to more premature deaths, but the data seems to describe an opposite association.

Higher Median Household Income, Lower Premature Death

We hypothesized that the states with the **highest household median income** would have the **fewest years of life lost** to premature death.

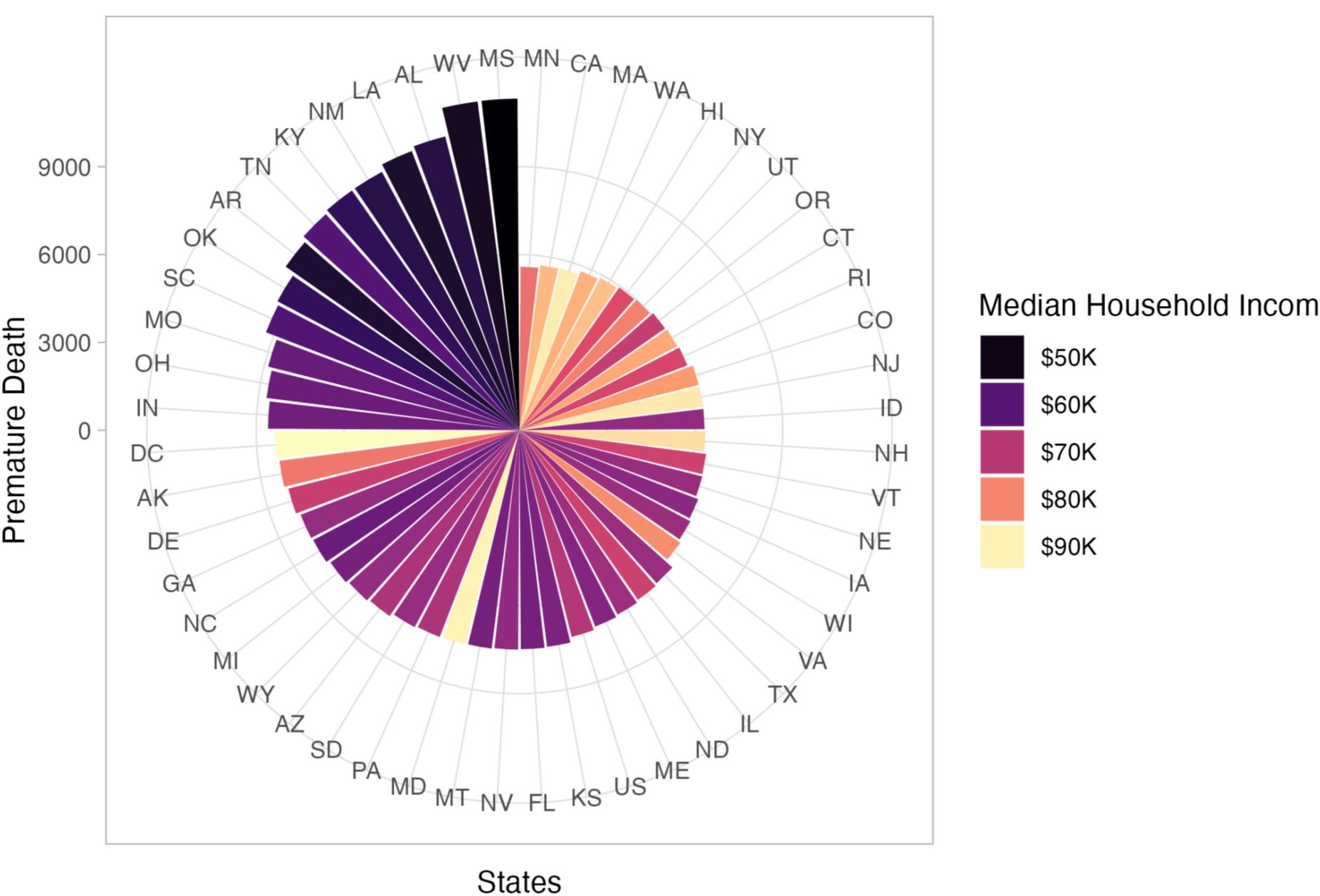


Figure 3. Rose Plot of Premature Deaths and Median Household Income by State

Generally speaking, moving clockwise from MN, the rose plot shows a general gradient from high incomes and low premature death to low incomes and high premature death.

Suicides

A major driver of premature death is suicides, which has significant impacts on many peoples lives[1]. In addition to the analysis into median household income, we hypothesized a possible interaction with suicide rate, such that suicide rates will be **higher**, in counties that are **older** and have a **lower median household income**.

Limitations of High School Completion Rates

Past research has demonstrated that education is highly correlated to life expectancy's and quality of life, and these education levels often differ by state and county.

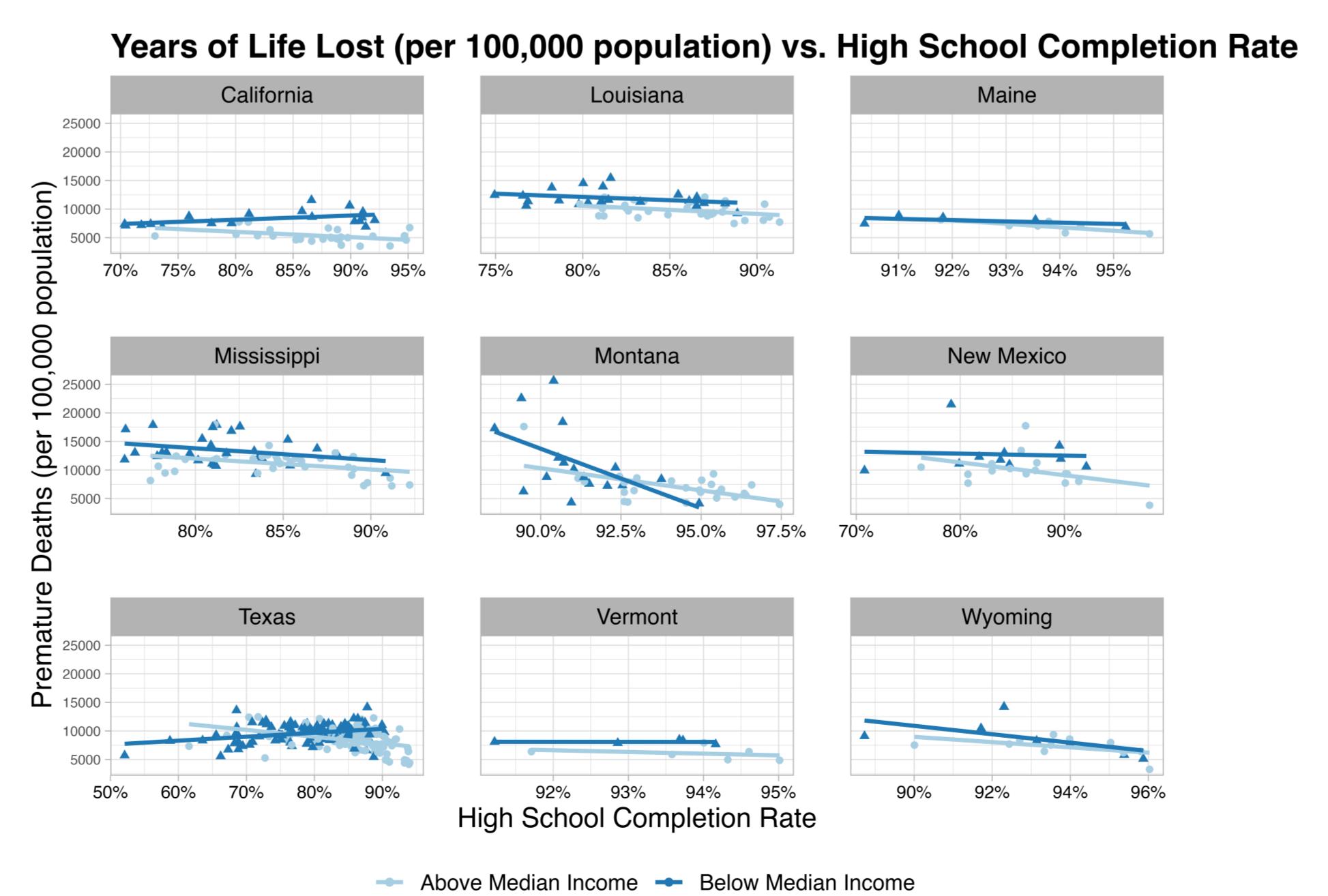


Figure 4. High School Completion Rate versus Premature Deaths

Our original hypothesis was that counties with **higher high school completion rates** would have **fewer years** of life lost to premature death.

Random Forest Modelling

We additionally built a random forests predictive model to predict premature death levels. Figure 5 shows the accuracy of our model.

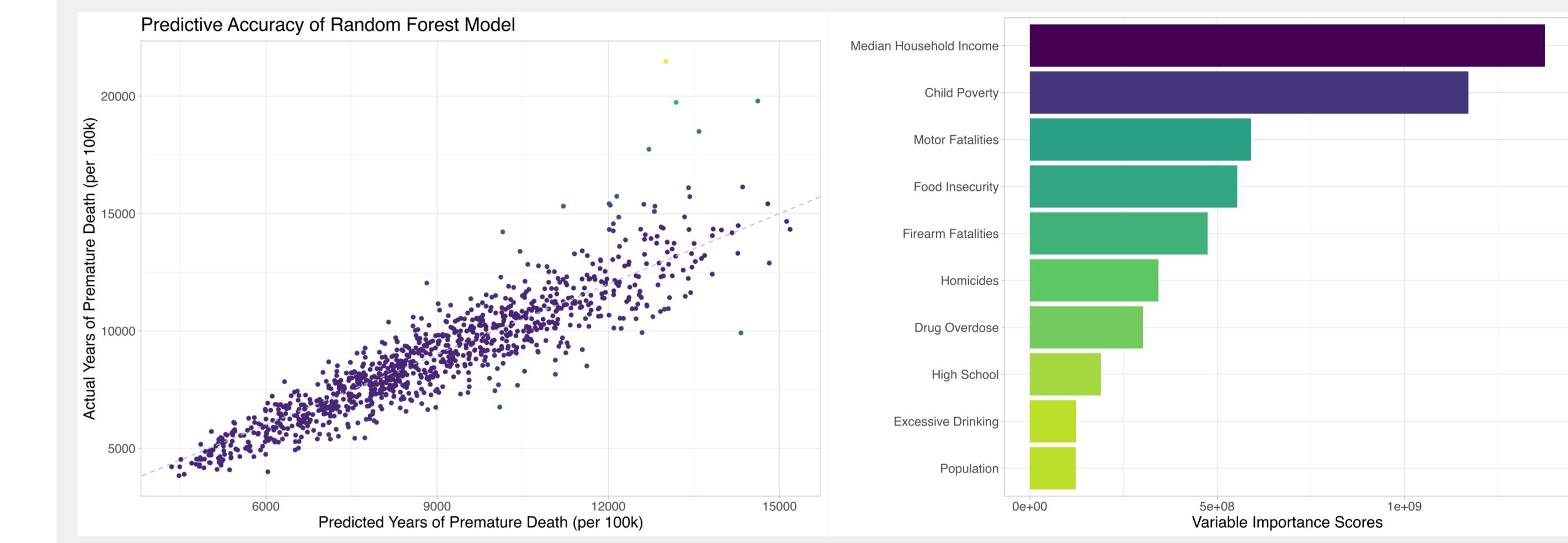


Figure 5. Predicted Premature Deaths versus Actual

Figure 6. Variable Importance

We built a random forest model, with hyperparameters of 450 trees and 5 randomly chosen variables considered at each split point.

Figure 6 depicts the most important of our randomly selected variables. **Variable importance scores** are highest for **median household income**, **child poverty**, and **motor fatalities**.

Conclusions

1. **States with greater rates of excessive drinking also had lower levels of premature death.** We found some regions had high premature death rates despite low drinking levels, specifically in the Bible Belt. We conjectured that this could be due to the role of religion and contraception laws.
2. **Premature death and median household income have a negative correlation.** Potential reasoning for this trend includes lack of access to quality healthcare and job security due to a poor median household income.
3. **On average, counties with older populations experience more suicides, and counties above the median suicide rate have significantly lower median household incomes.** After finding that both age distributions and median household incomes for counties highly correlate to suicide rates, we better understand how we may drive suicide rates down by focusing on government and private sector resources.
4. **Our plots of high school completion rates and premature death rates failed to show a consistent trend.** We believe data that can better capture the quality of the exposure of students to education would have a more clear negative association with premature deaths.

To reduce premature mortality, we recommend interventions that may address the **negative effects of low income, excessive drinking, and high school completion**. Reducing the drivers of premature death allows us to promote health equity in a time where health outcomes can drastically vary by race or economic status.

References

- [1] Centers for Disease Control and Prevention. Suicide. Centers for Disease Control and Prevention, 2023.
- [2] National Institute on Alcohol Abuse and Alcoholism.