

Exploring Suicide and Drug Overdose Rates Across Different Population Groups

Stat 184

Ziyue Han, Lingjie Dai, Samarth Patel

2024-12-13

1 Introduction

Health disparities, particularly those concerning mental health and substance abuse, are among the most critical challenges facing public health systems. Suicide and drug overdoses are leading causes of death in the United States, and their prevalence differs across demographic groups. Examining these disparities provides insights into the populations most at risk and highlights areas for intervention.

This report investigates three research questions regarding suicide and drug overdose rates:

1. Which gender-race combinations show the highest vulnerability to either suicide or drug overdose?
2. What are the differences in suicide and drug overdose rates among various age groups?
3. Is there a significant difference in suicide rates between male and female across racial group?

We will present background information, describe the methods used for analysis, summarize the collected data, and discuss our findings and their implications.

2 Methodology

2.1 Data Provenance

The datasets used in this research were found from Data.gov, the U.S. government's open data platform that provides access to a wide range of datasets collected by federal agencies. Specifically, the datasets on drug overdose death rates and suicide death rates were provided by the National Center for Health Statistics (NCHS), a division of the Centers for Disease Control and Prevention (CDC). These datasets were created to track trends in mortality rates across different demographic groups and to inform public health strategies. This information provides the foundation for analyzing disparities and identifying groups that may be disproportionately affected. These datasets are aligned with FAIR principles—they are findable, accessible, interoperable, and reusable—and provide the foundation for demographic analyses.

2.2 Attributes for Analysis

Our analysis focuses on five key attributes: Gender, Age, Race, Suicide Rate, and Overdose Rate. These attributes are directly available in the datasets and provide the foundation for examining patterns and disparities. Gender and Race provide demographic context, while Suicide_Rate and Overdose_Rate represent the number of deaths per 100,000 residents for suicides and drug overdoses, respectively.

2.3 Analytical Approach

To address the research questions, we used descriptive statistics to summarize suicide and overdose rates across demographic groups, visualization (e.g., bar chart, heatmap) to explore patterns and disparities, and correlation analysis to evaluate potential relationships between suicide and drug overdose rates.

3 FAIR Principle

The datasets used in this study follow the FAIR principles, meaning they are findable, accessible, interoperable, and reusable. They are easy to find on Data.gov, a platform that organizes and shares public data. The data can be downloaded in simple formats like CSV, which works with most tools for analysis. It also includes clear information about how the data was collected and structured, making it easier for others to use and build on. Following these principles ensures the data is easy to understand, share, and use for future research.

4 Data Exploration

4.1 Summary Statistics

4.2 Visualizations

5 Results

5.1 Research Question 1

5.2 Research Question 2

5.3 Research Question 3

6 Discussion

7 References

Data.gov. (n.d.). Death Rates for Suicide by Sex, Race, Hispanic Origin, and Age. <https://catalog.data.gov/dataset/death-rates-for-suicide-by-sex-race-hispanic-origin-and-age-united-states-020c1>

Data.gov. (n.d.). Drug Overdose Death Rates by Drug Type, Sex, Age, Race, and Hispanic Origin. <https://catalog.data.gov/dataset/drug-overdose-death-rates-by-drug-type-sex-age-race-and-hispanic-origin-united-states-3f72f>