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 were chosen as they provide a clear and measurable framework to investigate disparities in public health outcomes. Gender and Race are essential demographic indicators that help identify populations at risk based on societal and systemic factors. Age allows us to assess how vulnerabilities change across different life stages, such as adolescence, middle adulthood, and elderly years. Suicide_Rate and Overdose_Rate are expressed as deaths per 100,000 individuals, offering a standardized metric for comparison across groups. Together, these attributes allow for a multidimensional analysis, capturing the complexity of health disparities and enabling targeted recommendations for intervention.

2.3 Analytical Approach

To answer the research questions, we used different methods to analyze the data. For visual patterns and comparisons, we created charts like bar graphs and heatmaps to show the differences in suicide and overdose rates across groups. For the third research question—examining differences in suicide rates between males and females across racial groups—we performed a correlation and significance test to see if the differences are statistically meaningful. This approach helps us understand which groups are most affected and how these issues vary by demographic factors.

3 FAIR Principle

The datasets used in this study follow the FAIR principles, which make them easy to find, use, and share. They are available on Data.gov, a platform that organizes data with clear titles and descriptions so anyone can search for and access them. The data can be downloaded in formats like CSV, which work with many tools like Excel, R, and Python. It is also structured consistently, making it easy to combine with other datasets or use for different kinds of analysis. Detailed information about the data, including how it was collected and its limitations, is provided, ensuring that others can understand and reuse it for their own research. Following these principles helps ensure the data is clear, accessible, and useful for future studies.

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