

The Influence of Socioeconomic Factors on Mental Health Outcomes

Introduction

Mental health has emerged as a significant global concern in recent years, encompassing various challenges faced by individuals. The impact of socioeconomic factors on mental well-being has garnered considerable attention, revealing a complex relationship between the two. Individuals with lower socioeconomic status (SES) tend to experience higher rates of mental disorders, encounter barriers in accessing mental health services, and often suffer from increased psychological distress. Conversely, individuals with higher SES generally exhibit lower rates of mental disorders, possess better access to resources and support, and enjoy stronger social networks. Understanding the influence of socioeconomic factors on mental health outcomes is crucial for the development of targeted interventions and policies. This overview aims to provide a foundation for further exploration of the connection between mental health and socioeconomic factors, emphasizing the need to comprehend and address these factors to improve mental health outcomes globally.

This leads to our research question: *What specific socioeconomic factors among adults in the United States have a profound impact on mental health outcomes?*

To answer this question, we will look at the 2023 National Health Interview Survey provided by the CDC. This is a dataset consisting of a broad range of health topics collected through personal household interviews. Specifically, to ensure low variability in results, we will look at adult interviews.

Dataset Description

This dataset has 29522 rows of information for 647 attributes. Each row represents an adult interview conducted and their responses to certain questions.

Link to dataset: <https://www.cdc.gov/nchs/nhis/2023nhis.htm>

Below is an explanation of each attribute:

[Attribute Description.pdf](#)

For the purposes of this study, we are classifying for the attributes listed as “brief mental health assessment” which includes the PHQ41_A, PHQ42_A, PHQ44_A, and PHQ44_A attributes. These attributes ask questions related to certain behaviors such as loss of interest and anxiety within the past 2 weeks. It also asks for different medical conditions that a individual may have had in the past to try to correlate that with their mental health.

Classifying for this attribute will be useful because we can directly see the impact of specific socioeconomic factors on the mental health of the individual and use this information to

The Influence of Socioeconomic Factors on Mental Health Outcomes

create conclusions on infrastructure/public policy plans to alleviate mental health crises. For example, if the level of education affects mental health outcomes, then we can suggest more plans to keep people in education for longer. When it comes to different health conditions, if a certain cancer or illness leads to a higher correlation of mental health problems, then we would know the relative target area to help stop and find solutions to mitigate this issue and lower overall mental health problems.

Preprocessing Plans

- Plans for preprocessing
- Get rid of irrelevant attributes

Most of the data is numeric. We need to focus on a few things before we have good data:

- Some data has missing values, we can fix that by using either mode or mean as a substitution and seeing the correlation. If it is still pretty poor, we will just remove the instance.

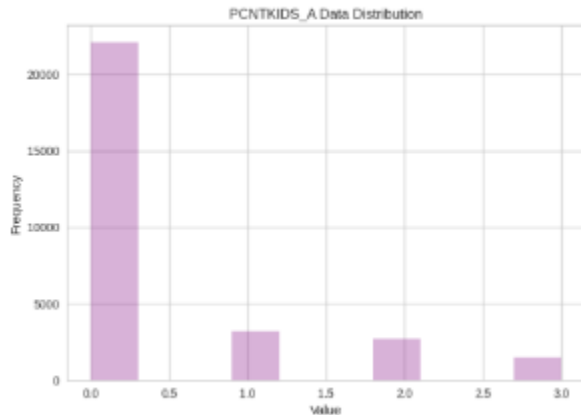
MHA: Brief mental health assessment

Variable #	Question #	Variable Name	Source Variables	Description	Type	Location	Length
1	MHA.0020.00.4	PHQ41_A		How often little interest in things, past 2 weeks	Num	518	1
2	MHA.0030.00.4	PHQ42_A		How often feeling down, past 2 weeks	Num	519	1
3	Recode	PHQ2SCREEN_A	PHQ41_A; PHQ42_A	PHQ-2 screener result	Num	520	1
4	MHA.0040.00.4	PHQ43_A		How often felt nervous/anxious/on edge, past 2 weeks	Num	521	1
5	MHA.0050.00.4	PHQ44_A		How often can't stop/control worrying, past 2 weeks	Num	522	1
6	Recode	GAD2SCREEN_A	PHQ43_A; PHQ44_A	GAD-2 screener result	Num	523	1

- These attributes are similar to each other such as feeling down and having little interest in things for the past 2 weeks and most of these can be combined, we will probably take an average or add these values together since they are all numbers.
- What we can do for those 4 values, is that they are a range from 1-9 and then if it has anything between 2-4 that might imply they have a mental health disorder. If they don't have 2-4 but do have 1 then we put as no mental health disorder and we put -1 for the ones where they didn't answer or didn't want to answer.

The Influence of Socioeconomic Factors on Mental Health Outcomes

- Some attributes just have one small range of values and the rest are just really small/insignificant. We will normalize these values.



- These skewed ranges will be further determined/analyzed by normalization which we will normalize by dividing each value by the max value in the dataset. This will give us a better sense of each range is importance and allow us to determine things further.

ATTRIBUTE SELECTION:

- If we get rid of around 550-600 attributes out of the 670 attribute selection in weka, we can use a baseline correlation of around 0.7-0.8 but we can increase or decrease.

Remove values from missing attributes