NESARC Report

Preface

After reviewing various datasets, I decided to focus on the NESARC report for my analysis. This dataset seemed the most interesting and comprehensive, and it would allow me to delve deeply into my area of interest: regional data related to major depression. As I began working with the data, I found myself particularly interested in exploring this topic further.

Research Question

Which region has the highest prevalence of major depression?

Hypothesis

After reviewing the codebook, I have identified enough relevant data to pursue my hypothesis that states with less sunshine may have higher rates of depression. To confirm the accuracy of this hypothesis, further research and analysis will be necessary. My ultimate aim in this research is to reach conclusions that are backed by the data and that shed light on the relationship between sunshine and depression rates in different states.

Codebook

I have chosen to use seven variables for my dataset: IDNUM, SEX, AGE, REGION, S4AQ1, S4AQ2, and S4AQ53. I believe this is the best approach given the information I have. However, I am starting to realize that the data is largely qualitative and may not provide any useful insights into my hypothesis. I may consider adding additional variables as I progress through the project in order to extract useful data.

NESARC Report

IDNUM	UNIQUE ID NUMBER WITH NO ALPHABETICS
43093	1-43093. Unique Identification Number
SEX	SEX
18518	1. Male
24575	2. Female
AGE	AGE
43079	18-97. Age in years
14	98. 98 years or older
REGION	CENSUS REGION
8209	1. Northeast
8991	2. Midwest
16156	3. South
9737	4. West
S4AQ1	EVER HAD 2-WEEK PERIOD WHEN FELT SAD, BLUE, DEPRESSED, OR DOWN MOST OF THE TIME
12785	1. Yes
29416	2. No
892	9. Unknown
S4AQ2	EVER HAD 2-WEEK PERIOD WHEN DIDN'T CARE ABOUT THINGS USUALLY CARED ABOUT
10533	1. Yes
31618	2. No
942	9. Unknown
S4AQ53	FELT TROUBLED BECAUSE OF WAY YOU FELT/OFTEN WISHED COULD GET BETTER
7096	1. Yes
1632	2. No
89	9. Unknown
34276	BL. NA. worst period did not meet sympton criteria for major depression

Literature Review

I conducted a search on Google Scholar for studies on depression by state and found several relevant studies. However, I noticed that many of these studies do not place a strong emphasis on regional data. The CDC has data on current depression rates by state, which range from 4.8% in North Dakota to 14.8% in Mississippi [1]. One study used Google Trends data to map population depression and found evidence of seasonal patterns [2]. Another factor that I had not considered is that it may not be a lack of sunshine specifically, but rather a deficiency in Vitamin D, that contributes to depression [3]. As we delve deeper into the data, we will be better able to begin answering questions about these potential relationships.

References

1. Centers for Disease Control and Prevention (CDC). Current depression among adults---United States, 2006 and 2008. MMWR Morb Mortal Wkly Rep. 2010 Oct 1;59(38):1229-35. PMID: 20881934. https://pubmed.ncbi.nlm.nih.gov/20881934/
2. Wang A, McCarron R, Azzam D, Stehli A, Xiong G, DeMartini J. Utilizing Big Data From Google Trends to Map Population Depression in the United States: Exploratory Infodemiology Study JMIR Ment Health 2022;9(3):e35253 URL: https://mental.jmir.org/2022/3/e35253 DOI: 10.2196/35253