Are we living in a depressive society?*

Trends in symptoms of depressive disorder in the US population from May 2020 to March 2024

Chay Park

April 12, 2024

This study explores how socio-cultural factors like age, sex, race/ethnicity, and education level influenced depressive disorder in the U.S. during and after the COVID-19 pandemic, analyzing data from 2020 to 2024. Using advanced statistical methods in R, it uncovered significant variations in mental health impacts across different demographic groups. The research reveals a complex relationship between socio-cultural factors and mental health, highlighting the increased vulnerability of certain populations during global crises. The findings contribute to the development of targeted public health measures to alleviate mental health issues amid and following pandemics, enhancing our understanding of the impact of socio-cultural factors on mental well-being in times of crisis.

1 Introduction

Depression and anxiety are not only prevalent but escalating issues in the modern world, affecting individuals across the globe regardless of demographic factors (World Health Organization, 2021). These mental health challenges have been further exacerbated by the COVID-19 pandemic, which led to widespread social and economic upheaval (Brooks et al., 2020). While various studies have explored the impact of the pandemic on mental health, there remains a significant gap in understanding the interactions between socio-cultural factors and the prevalence of anxiety and depressive disorders during this period (Pfefferbaum & North, 2020).

This paper addresses this gap by analyzing data from the U.S. Census Bureau's Household Pulse Survey, initiated in response to the pandemic to assess its effects on American households, including mental wellness (U.S. Census Bureau, 2020). Our analysis focuses on the correlation between the prevalence of anxiety or depressive disorders and socio-cultural factors such as age, sex, race/ethnicity, and education level during the pandemic and post pandemic years of

*Code and data are available at: https://github.com/Chay-HyunminPark/Anxiety_depressive_disorder.

2020 to 2024. Specifically, the estimand of this study is the difference in the rates of symptoms of anxiety or depressive disorders across individuals, correlated with socio-cultural factors including age, sex, race/ethnicity, and education level.

Employing a quantitative methodology, this paper utilizes R (R Core Team 2020) for analysis. We utilized packages like ggplot2 (Wickham et al., 2016) and gridExtra (Murrell, 2021) for data analysis and visualization. Our findings reveal significant trends and disparities in the rates of anxiety and depressive disorders across different socio-cultural groups, highlighting the intricate relationship between these factors and mental health outcomes during the pandemic.

The importance of this research lies in its potential to inform public health policies and interventions, aiming to mitigate the adverse effects of such global crises on mental health (Xiong et al., 2020). By understanding the specific socio-cultural factors that contribute to increased vulnerability to anxiety and depressive disorders, targeted strategies can be developed to support affected populations more effectively.

The paper is structured to first present the trend analysis of anxiety and depressive disorder rates in the U.S. from 2020 to 2024, followed by an examination of how these rates vary across different demographic groups. Subsequent sections discuss the statistical methods and results for each research question. It opens up the findings and let the world explores their broader implications for mental health research and policy. The conclusion summarizes the key insights, acknowledges the limitations of the study, and suggests avenues for future research.

The study focuses on answering the following research questions:

- What is the trend in US anxiety and depressive disorder rates from 2020 to 2024?
- How do depressive rates vary across different age groups?
- How do depressive rates differ based on race and ethnicity?
- How do depressive rates differ based on sex?
- How do adepressive rates differ based on education level?
- Is there codependency of social connection on one's depression level?

2 Data

—- A thorough discussion of measurement, relating to the dataset, is provided in the data section. —- ## Methodology

The U.S. Census Bureau, in collaboration with five federal agencies, launched the Household Pulse Survey to produce data on the social and economic impacts of Covid-19 on American households. The Household Pulse Survey was designed to gauge the impact of the pandemic on employment status, consumer spending, food security, housing, education disruptions, and dimensions of physical and mental wellness (National Center for Health Statistics, 2024). The data on the trend in US anxiety and depressive disorder rates, measured as ~ from 2020 to 2024, was sourced from the Centers for Disease Control and Prevention (CDC). Alongside with National Estimate in which covers the whole population, trends in anxiety and depressive disorder rates by population subgroup, specifically categorized by age, race and ethnicity, sex, education groups were gathered from the National Center for Health Statistics (NCHS). The data is freely available at https://data.cdc.gov/NCHS/Indicators-of-Anxiety-or-Depression-Based-on-Repor/8pt5-q6wp/about_data, with raw files located in the inputs/data on the GitHub repository.

2.1 Features

In this paper, data validation is conducted using an R script file named '03-test_data' to ensure the accuracy of the dataset obtained directly from the source. The paper utilizes both 'Indicators_of_Anxiety_or_Depression_Based_on_Reported_Frequency_of_Symptoms_During_Last_7_Days_2 and 'Lack_of_Social_Connection' CSV files. Then by the characteristics of the group, the data sets are separately saved as 'Age_subgroup_trends' where it only saves the rows with 'By Age' under the 'Group' column. Similarly, 'Educ_subgroup_trends' are the dataset only with the rows 'By Education' under the 'Group' column to identify the patterns of each population subgroup with classification of highest educational attainment. All CSV files are saved as parquet file as Parquet's columnar storage format. As opposed to row-based formats like CSV or Excel, Parquet allows it to efficiently handle analytical queries by accessing only the necessary columns of data, which reduces storage reads, improves query performance, and lowers storage costs.

The column 'Value' is used to denote the symptom rate on the y-axis. Column 'Time Period End Date' is used to denote the range of the period, a date for the x-axis.

3 Results

3.1 Trend in US Anxiety and Depressive Disorder Rates

Both Anxiety and Depressive Disorder showing the similar trends over the period.

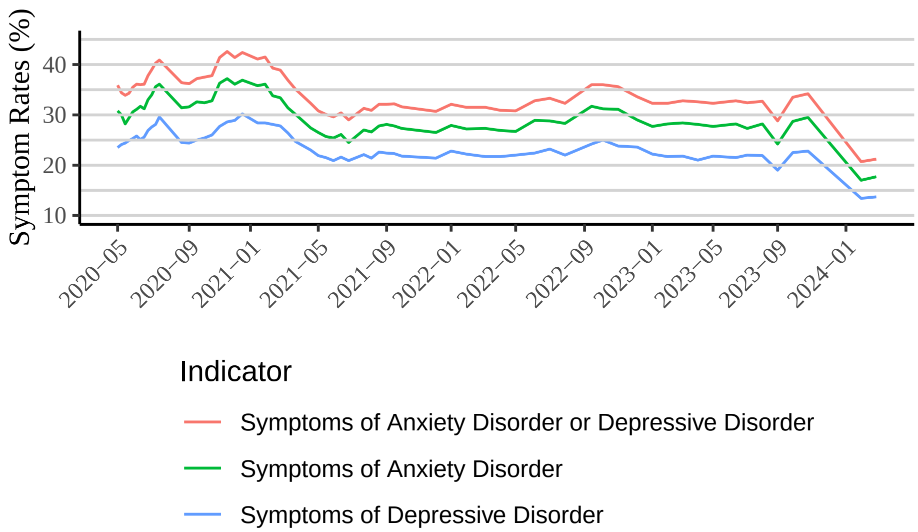


Figure 1: Trend in US Anxiety and Depressive Disorder Rates

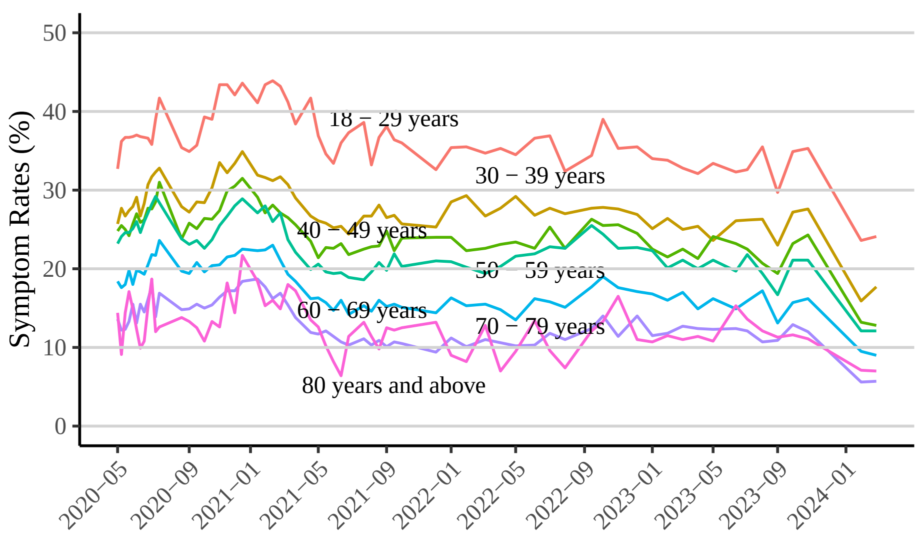


Figure 2: Depressive Disorder Trend by Population Subgroup by Age

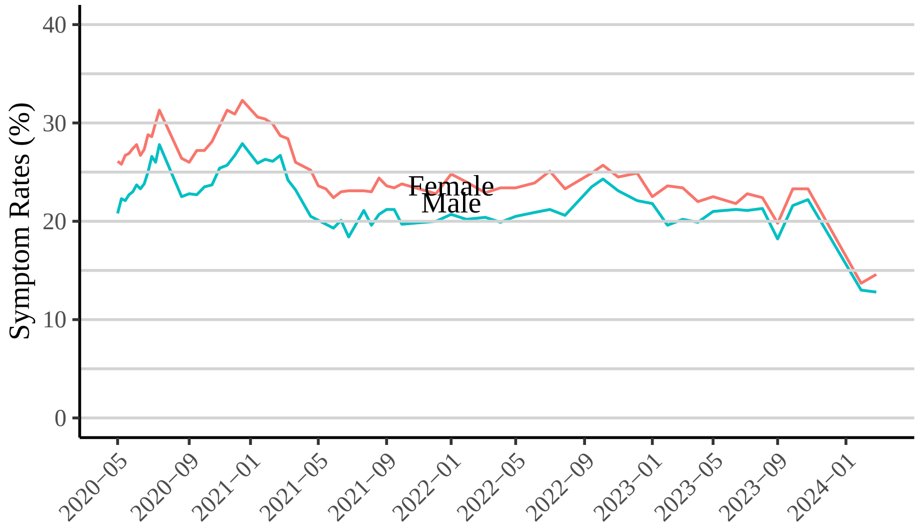


Figure 3: Depressive Disorder Trend by Population Subgroup by Sex

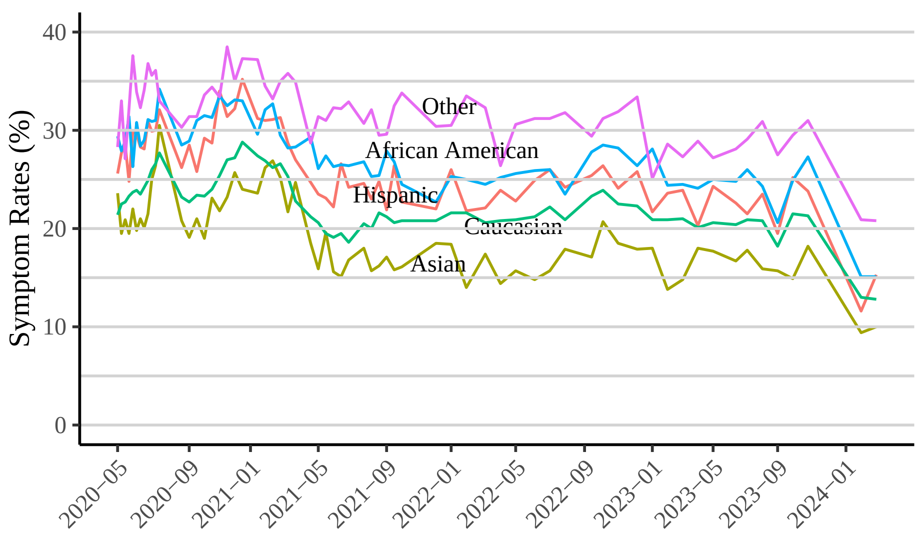


Figure 4: Depressive Disorder Trend by Population Subgroup by Race

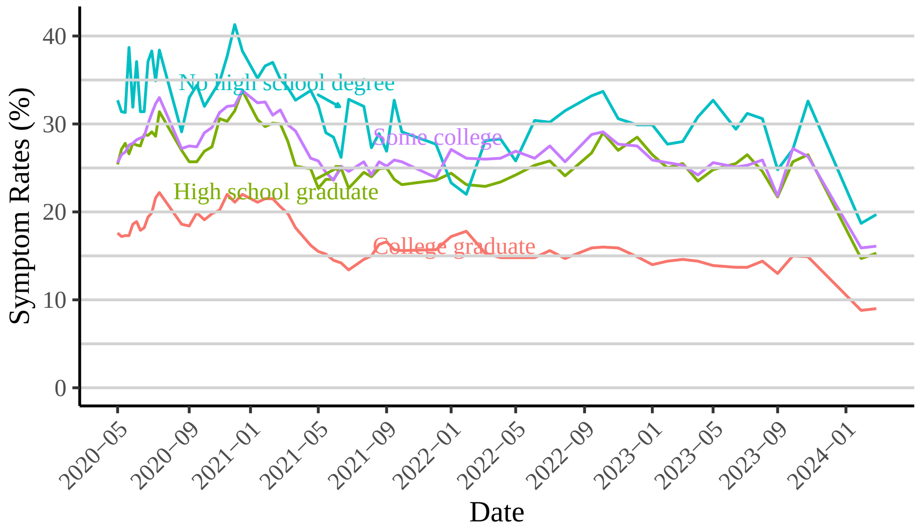


Figure 5: Depressive Disorder Trend by Population Subgroup by Education

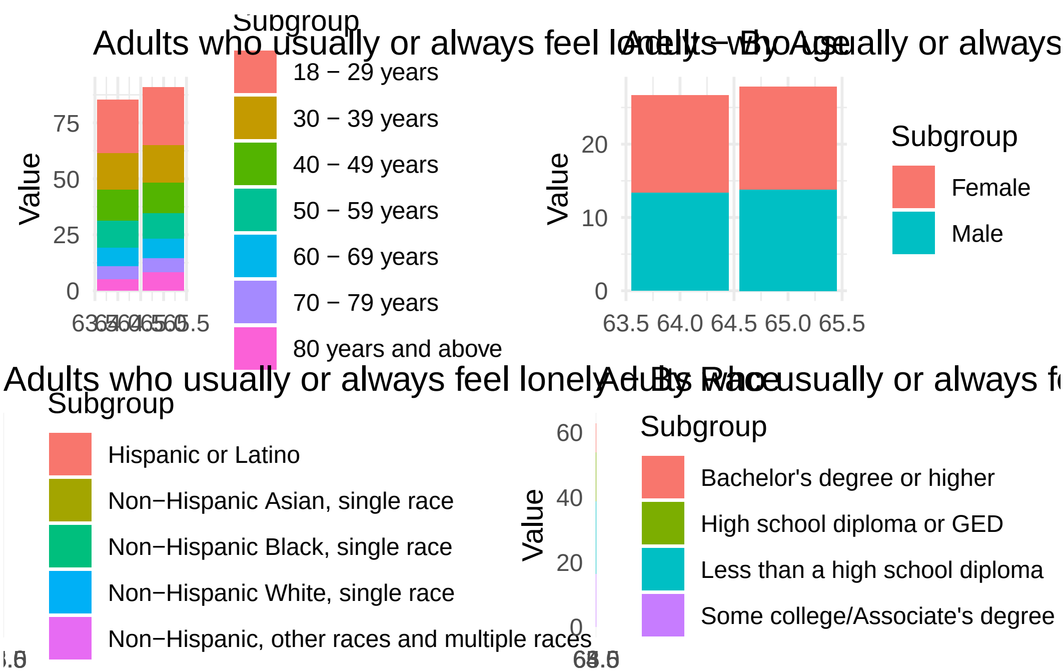
3.2 Trends in Depressive Disorder Rates by Population Subgroup

3.2.1 Population subgroup by age

3.2.2 Population subgroup by sex

```
# Define the path to the data file and load it
social_data <- read_parquet(file = here::here("outputs/data/Social_Connection.parquet"))

# Filter data for the indicator "Adults who usually or always feel lonely" and "By Age" subg
social_age_data <- social_data %>%
  filter(Indicator == "Adults who usually or always feel lonely", Group == "By Age")

# Plot the stacked bar chart
graph_age <- ggplot(social_age_data, aes(x = `Time.Period`, y = Value, fill = Subgroup)) +
  geom_bar(stat = "identity", position = "stack") +
  labs(title = "Adults who usually or always feel lonely - By Age",
       x = NULL,
       y = "Value") +
  theme_minimal()


# Filter data for the indicator "Adults who usually or always feel lonely" and "By Sex" subg
social_sex_data <- social_data %>%
```

```
  filter(Indicator == "Adults who usually or always feel lonely", Group == "By Sex")

# Plot the stacked bar chart
graph_sex <- ggplot(social_sex_data, aes(x = `Time.Period`, y = Value, fill = Subgroup)) +
  geom_bar(stat = "identity", position = "stack") +
  labs(title = "Adults who usually or always feel lonely - By Sex",
       x = NULL,
       y = "Value") +
  theme_minimal()

# Filter data for the indicator "Adults who usually or always feel lonely" and "By Race/Hispa
social_race_data <- social_data %>%
  filter(Indicator == "Adults who usually or always feel lonely", Group == "By Race/Hispanic

# Plot the stacked bar chart
graph_race <- ggplot(social_race_data, aes(x = `Time.Period`, y = Value, fill = Subgroup)) +
  geom_bar(stat = "identity", position = "stack") +
  labs(title = "Adults who usually or always feel lonely - By Race",
       x = NULL,
       y = "Value") +
  theme_minimal()

# Filter data for the indicator "Adults who usually or always feel lonely" and "By Education"
social_educ_data <- social_data %>%
  filter(Indicator == "Adults who usually or always feel lonely", Group == "By Education")

# Plot the stacked bar chart
graph_educ <- ggplot(social_educ_data, aes(x = `Time.Period`, y = Value, fill = Subgroup)) +
  geom_bar(stat = "identity", position = "stack") +
  labs(title = "Adults who usually or always feel lonely - By Education",
       x = NULL,
       y = "Value") +
  theme_minimal()

# Arrange graphs in single pane
grid.arrange(graph_age, graph_sex, graph_race, graph_educ, nrow = 2, ncol = 2)
```



4 Model

4.1 Model set-up

4.1.1 Model justification

5 Discussion

What is done in this paper? What is something that we learn about the world? What is another thing that we learn about the world?

5.1 Findings

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

5.2 Ethical Implication

5.3 Limitation

– What are some weaknesses of what was done? – The survey was designed to meet the goal of accurate and timely weekly estimates. Hence, it was conducted by an internet questionnaire, with invitations to participate sent by email and text message (National Center for Health Statistics, 2024). Internet questionnaires face several significant challenges. One of the primary concerns is self-selection bias, where the survey tends to attract respondents who are particularly interested in the topic, which can skew the results and limit their generalizability. Additionally, these surveys may not reach a diverse demographic due to varying levels of internet access across different population segments, thus excluding certain groups, such as older adults or lower-income individuals, and potentially leading to a lack of representativeness. Respondents might also misunderstand or misinterpret questions without the opportunity for real-time clarification, compromising the accuracy of the data collected. Privacy and security concerns are prevalent too, as individuals may be hesitant to share personal or sensitive information online for fear of data breaches or doubts about the survey platform's security measures. Lastly, the validity and honesty of responses are at risk; while the anonymity of online surveys can encourage candidness, it might also lead to less thoughtful or dishonest answers, affecting the reliability of the survey outcomes (Dillman, Smyth, & Christian, 2014; Morling, 2018).

5.4 Future Research

– What is left to learn or how should we proceed in the future? – For future research, survey can further ask what are the factors that affect their mood.

Appendix

A Additional data details

B Model details

B.1 Posterior predictive check

B.2 Diagnostics

References

Alexander, Rohan. 2023. "Telling Stories with Data." Chapman; Hall/CRC. https://tellingstorieswithdata.com/.

Brooks, Samantha K., Rebecca K. Webster, Louise E. Smith, Lisa Woodland, Simon Wessely, Neil Greenberg, and Gideon James Rubin. 2020. "The Psychological Impact of Quarantine and How to Reduce It: Rapid Review of the Evidence." *The Lancet* 395 (10227): 912–20. https://doi.org/10.1016/S0140-6736(20)30460-8.

Friendly, Michael, Chris Dalzell, Martin Monkman, and Dennis Murphy. 2020. *Lahman: Sean "Lahman" Baseball Database*. https://CRAN.R-project.org/package=Lahman.

Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé Iii, and Kate Crawford. 2021. "Datasheets for Datasets." *Communications of the ACM* 64 (12): 86–92.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

Health Statistics, National Center for. 2024. "Indicators of Anxiety or Depression Based on Reported Frequency of Symptoms During Last 7 Days." United States: National Center for Health Statistics. 2024. https://data.cdc.gov/NCHS/Indicators-of-Anxiety-or-Depression-Based-on-Repor/8pt5-q6wp/about_data.

Horst, Allison Marie, Alison Presmanes Hill, and Kristen B Gorman. 2020. *Palmerpenguins: Palmer Archipelago (Antarctica) Penguin Data*. https://doi.org/10.5281/zenodo.3960218.

Murrell, Paul. 2021. *R Graphics Cookbook: Practical Recipes for Visualizing Data*. O'Reilly Media.

Pfefferbaum, Betty, and Carol S. North. 2020. "Mental Health and the Covid-19 Pandemic." *New England Journal of Medicine* 383 (6): 510–12. https://doi.org/10.1056/NEJMp2008017.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

U.S. Census Bureau. 2020. *Household Pulse Survey: Measuring Social and Economic Impacts During the Covid-19 Pandemic*. https://www.census.gov/programs-surveys/household-pulse-survey.html.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

World Health Organization. 2021. "Depression and Other Common Mental Disorders: Global Health Estimates." https://www.who.int/news-room/fact-sheets/detail/depression.

Xiong, Jiaqi, Orly Lipsitz, Flora Nasri, Leanna M. W. Lui, Hartej Gill, Laura Phan, David Chen-Li, et al. 2020. "Impact of COVID-19 Pandemic on Mental Health in the General Population: A Systematic Review." *Journal of Affective Disorders* 277: 55–64. https://doi.org/10.1016/j.jad.2020.08.001.