**ADTA 5940 Concept Paper**

**Beyond Infection:**

**Predicting Mental Health Impacts in Pandemics through Data Analytics**

**By Team 1**

**Mani Sai Deeraj Marru**

**Sai Chaitanya Avasarala**

**Sai Kamala Reddy Chintaguntla**

**Introduction**

We are currently working on research that analyses the anxiety and depression that individuals experienced during the COVID-19 epidemic. Everyone had a very difficult time during this period, and several individuals experienced extreme depression or anxiety. Our goal is to examine the information we gathered throughout this time to learn more about the reasons behind people's feelings.  
  
 For our project, we are concentrating on two main concepts. First, we'll take a detailed look at the data we collected on the feelings of individuals during the COVID-19 epidemic. Our goal is to identify any patterns or explanations for why a greater number of individuals may have experienced anxiety or depression during that period. Finding patterns and determining what was happening are the main goals of this section.  
  
Our second research component involves developing an to predict people's likelihood of experiencing anxiety or depression in the event of a potential COVID-19 pandemic. To produce these predictions, this tool—which we will refer to as the model—will draw on the knowledge gathered during the first phase of our study. Our goal is to keep the general mood up even in difficult times, and by identifying these tendencies, the model might help to take action to assist individuals before they become depressed.  
  
In other words, our project's goal is to identify the reasons behind people's pandemic-related emotions and then use that information to help avoid experiencing those same emotions in the future, should a circumstance such to this one arise.

Beyond simply pandemics, we also believe that our effort may be helpful in other difficult circumstances. By knowing how to avoid depression or anxiety, we can support individuals in maintaining their mental toughness in the face of adversity.  
  
In summary, the goal of our initiative is to identify strategies for maintaining people's mental health during difficult times by using the lessons learned during the epidemic.

**Mission Statement**

Our objective is to shed light on the mental health issues that many individuals faced during the COVID-19 pandemic using data on anxiety and depression. To provide advice on how to prevent encountering such disruptions in the future, our aim is to recognize the trends and origins of these feelings. Through our research, we are developing a model to predict and mitigate similar crises' consequences on mental health. Our goal is to provide individuals with the knowledge and tools necessary to maintain mental resilience in any challenging circumstance, not just pandemics.

**EDA Questions**

1. In which age range are anxiety and depression disorders most common?
2. How does the pandemic's pattern of anxiety and depression disorders evolve over time?
3. Do any states or areas have greater than average prevalence of these disorders?

**Predictive Questions**

1. Can we predict the future incidence rates of anxiety and depressive disorders among different age groups and genders based on the current trend observed in the dataset?
2. Which demographic factors (such as age, gender, employment status) are the strongest predictors of developing anxiety and depressive disorders during different stages of the COVID-19 pandemic?
3. Based on the patterns of mental health indicators in the dataset, can we predict future healthcare utilization specifically for mental health services related to anxiety and depression?

**Data Description**

The data consists of 14374 rows and 14 columns, and the data is collected in between 04/23/2020 and 9/20/2023.

|  |  |  |
| --- | --- | --- |
| **Name** | **Data Type** | **Description** |
| Indicator | text | The specific measure being observed. |
| Group | text | The category or classification the data belongs to. |
| State | text | The geographical location in USA. |
| Subgroup | text | A further division or classification within a group. |
| Phase | text | The stage or period of the data collection or event. |
| Time Period | text | The duration over which the data was collected or applies. |
| Time Period Label | text | A descriptive name or label for the time period. |
| Time Period Start Date | calendar date | The starting date of the time period. |
| Time Period End Date | calendar date | The ending date of the time period. |
| Value | number | The numerical measurement or value associated with the indicator. |
| Low CI | number | The lower bound of the confidence interval for the value. |
| High CI | number | The upper bound of the confidence interval for the value. |
| Confidence Interval | text | The range within which the true value is expected to lie, with a given level of confidence. |
| Quartile Range | text | A statistical range dividing the dataset into four equal parts, indicating variability or distribution. |

The Household Pulse Survey was established by the US Census Bureau in collaboration with five other government agencies to collect data on the socioeconomic effects of the COVID-19 epidemic on American homes. This creative survey sought to determine how the pandemic affected areas including employment status, spending patterns, food security, housing circumstances, interruptions to schooling, and general physical and mental health.  
  
Households were asked to participate by email and SMS in an online questionnaire format, which helped the survey accomplish its goal of delivering reliable and timely weekly data. Based on the Census Bureau's Master Address File Data, participants were chosen via a random method from dwelling units connected to at least one email address or cellphone number. One person was picked from each selected home to provide personal replies. Adjustments were performed for non-responses and to conform to the Census Bureau's demographic estimates regarding age, sex, race and ethnicity, and educational attainment in order to guarantee the representativeness of the findings. The information provided complies with the proportionate estimate presentation guidelines established by the National Centre for Health Statistics (NCHS).

**Challenges and Concerns**

1. Quality of Data: There is a significant amount of missing data, which could be a problem if not handled carefully.
2. Data collection method: As the data is collected through household pulse survey the data is not completely reliable.
3. Technical expertise: As the dataset is based on mental health indicators, medical expertise plays a vital role in making the model efficient.
4. Temporal Relevance: The timing of the collection of data impacts the trueness as the mental health condition can fluctuate over time due to various factors. The data collected at one point of time might not accurately represent conditions at another.
5. Response bias: Given that the data is derived from a home pulse survey, respondents may have a desire to withhold some information from others and provide answers that are socially acceptable.
6. Complexity of Mental Health: Anxiety and depression might also occur with other health issues which can complicate our analysis, we are not considering other mental issues which may tamper the analysis.

**References**

1. Differential relationship of observer-rated and self-rated depression and anxiety scales... by Ham, Jinsil; Kim, Hesun Erin; Kim, Jae Jin ; Seok, Jeong Ho; Kim, Eunjoo; Park, Jin Young; Lee, Boreom; Oh, Jooyoung

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10109339/>

1. Clinical presentation of psychotic experiences in patients with common...

by Wiedemann, Anna; Stochl, Jan; Russo, Debra ; Patel, Ushma; Ashford, Polly-Anna; Ali, Naima; Jones, Peter B.; Perez, Jesus <https://www.sciencedirect.com/science/article/pii/S0165032723012545?via%3Dihub>

# It’s difficult to grasp the projected deaths from Covid-19. Here’s how they compare to other causes of death By [Sharon Begley](https://www.statnews.com/staff/sharon-begley/)  and [Hyacinth Empinado](https://www.statnews.com/staff/hyacinth-empinado/)

# <https://www.statnews.com/2020/04/09/its-difficult-to-grasp-the-projected-deaths-from-covid-19-heres-how-they-compare-to-other-causes-of-death/>

# Anti-Anxiety Prescription Meds Increase Amid COVID-19 Pandemic, Report Says By [Snow Digon](https://www.ibtimes.com/reporters/snow-digon) <https://www.ibtimes.com/anti-anxiety-prescription-meds-increase-amid-covid-19-pandemic-report-says-2962093>

# Development and initial validation of the COVID Stress Scales Author links open overlay panel Steven Taylor a, Caeleigh A. Landry b, Michelle M. Paluszek b, Thomas A. Fergus c, Dean McKay d, Gordon J.G. Asmundson  <https://www.sciencedirect.com/science/article/pii/S0887618520300463?pes=vor>

# How does information technology affect individuals’ health behavior in a pandemic? Insights from mobile health application use during COVID-19 Author links open overlay panelParvin Hashemi a, Iris Reychav b, Ankur Arora c, Nour Sawaed d, Rajiv Sabherwal e, Joseph Azuri <https://www.sciencedirect.com/science/article/pii/S0747563223003564>

# Health anxiety, death anxiety and coronaphobia: Predictors of postpartum depression symptomatology during the COVID-19 pandemic links open overlay panelAna-Maria Andrei a, Rebecca Webb b, Violeta Enea <https://www.sciencedirect.com/science/article/pii/S026661382300150X>

# Anjel Vahratian, PhD1; Stephen J. Blumberg, PhD1; Emily P. Terlizzi, MPH1; Jeannine S. Schiller, MPH <https://www.cdc.gov/mmwr/volumes/70/wr/mm7013e2.htm>