**Assessment of Anxiety and Depression using BDA tools and GPTmodels**

Inha University

Big Data Analysis

Kakani Vijay

May 23, 2023

Team Basic:

Kim Sol

Felester

Spencer

Ali

**Table Of Contents**

1. Introduction
2. Literature Review
3. Methodology
4. Results and Analysis
   1. Dataset Analysis
   2. Survey Analysis
5. Delegation
6. Conclusion
7. References

**Introduction**

Mental Health disorders such as anxiety and depression, have become significant issues to be considered worldwide. It affects millions of people worldwide, especially youth and causes inconvenience in daily life. Diagnosing and assessing these conditions clearly is important for effective treatment and support. Methods that are used to identify mental health disorders mostly rely on self-reporting and clinical assessments, which is time consuming and prone to human bias. However, we think that with the help of artificial intelligence and GPT models we will be able to improve the assessment process making them faster and more accurate.

The aim of this project is exploring the use of GPT models and BDA analyzing tools in diagnosis of Anxiety and Depression and further develop an efficient and reliable tool that can assist physicians in early detection and intervention. During our project we focused on analytical abilities of OpenAi’s ChatGPT, trying to take advantage of its access to the great numbers of databases and ability to interact in conversational dialogues making it suitable for analyzing data related to mental health.

The primary goal of this research is to test the effectiveness and precision of GPT models in measuring anxiety and depression. We hope to discover patterns, signs, and linguistic markers linked with various mental health problems by using large-scale pre-trained language models' knowledge and understanding of human language. This study takes a data-driven approach, drawing on existing datasets comprising anonymised text data from people with anxiety and depression.

Furthermore, the research looks into the ethical implications of using AI models for mental health assessment. The potential benefits of automated assessment must be evaluated against privacy issues, potential biases, and the significance of human monitoring in decision-making. It is critical to investigate the limitations, constraints, and ethical implications of utilizing AI-powered models in mental health assessments, allowing for a thorough assessment of their usefulness and potential hazards.

The project report will be structured as follows:

1. Literature Review: Looking through existing studies about implementation of AI in the Health Sector. Overview of the academic papers and narrowing down our topic to Anxiety and Depression,highlighting gaps and limitations of current approaches.
2. Methodology: Conducting Survey to support our research questions. Preprocessing of datasets related to our research area to identify problems and opportunities, and extract meaningful insights.
3. Results and Analysis: Presentation of Survey result analysis and Dataset analysis.
4. Delegation: Pointing out contributions of each team member during the project.
5. Conclusion: Summary of findings and future predictions about the use of our project idea.

The goal of this project report is to add to the expanding body of information about the use of AI and NLP in mental health evaluation. We think that by assessing the potential of GPT models for anxiety and depression assessment, we will pave the path for more innovative and efficient approaches to mental health screening, leading to better diagnosis, treatment, and support for people suffering from these diseases.

**Literature Review**

Before deciding on the final topic of our project all of our team members did the research about the possible implementation of GPT models in the health sector. After studying a lot of articles and academic papers we came up with a couple of options, both of which were related to psychological disorders. According to the World Health Organization (WHO) the global need for significant changes in mental health care is needed. The article1 underscores the importance of modern innovations, specifically implementation of AI to address the challenges faced by people suffering with Anxiety and Depression.

Another study conducted by the National Center for Biotechnology Information explores that AI can be the solution for identifying anxiety and depression symptoms by evaluating social media posts. The researchers use machine learning algorithms to evaluate a huge dataset of social media posts in order to illustrate the utility of automated methods for early detection and monitoring of mental health issues.

The research conducted by the Anxiety and Depression Association of America (ADAA) provides facts and statistics related to anxiety disorders. The article points out the impacts and challenges of anxiety disorder on individuals and society. It emphasizes the need for a modern approach to assessment and importance of accurate diagnosis, which is vital in constructing effective treatment.

These sources offer useful insights into the global impact of mental health diseases, the promise for AI and NLP technologies in mental health assessment, and the present challenges and gaps in traditional methodologies. The WHO publication highlights the importance of changing mental health care, whereas the NCBI study demonstrates the use of automated text classification for symptom diagnosis. The data from the ADAA shed light on the prevalence and effect of anxiety disorders, underlining the importance of appropriate assessment and intervention strategies.

By incorporating these research findings, we got a comprehensive understanding of the benefits of using GPT models for mental health assessment.

**Methodology**

To support our Project we conducted a survey, which was constructed to identify the symptoms and causes of Anxiety and Depression. Our survey Consisted of 10 questions:

1. Age, occupation, major.
2. Have you had any traumatic/depressive/anxiety experiences in the past?
3. What Symptoms do you struggle with while caught in an anxiety attack and/or depression?

and so on. Each question was designed to find out the mental state of the surveyor and identify whether he/she has anxiety or depression disorders. We mostly focused on two types of questions, causes and symptoms. We got more than 30 participants and it was enough to support our topic.

For further analysis we had to find datasets and make analysis. We used the Kaggle platform to search for datasets related to Anxiety and Depression disorders and came up with a lot of results. Among them we had to choose the most appropriate in order to gain appropriate information about our project. The dataset created by Lucas Greenweel “Depression Anxiety Stress Scales Responses” which included Questions, answers, and metadata collected from 39775 Depression Anxiety Stress Scales became the dataset of our choice.

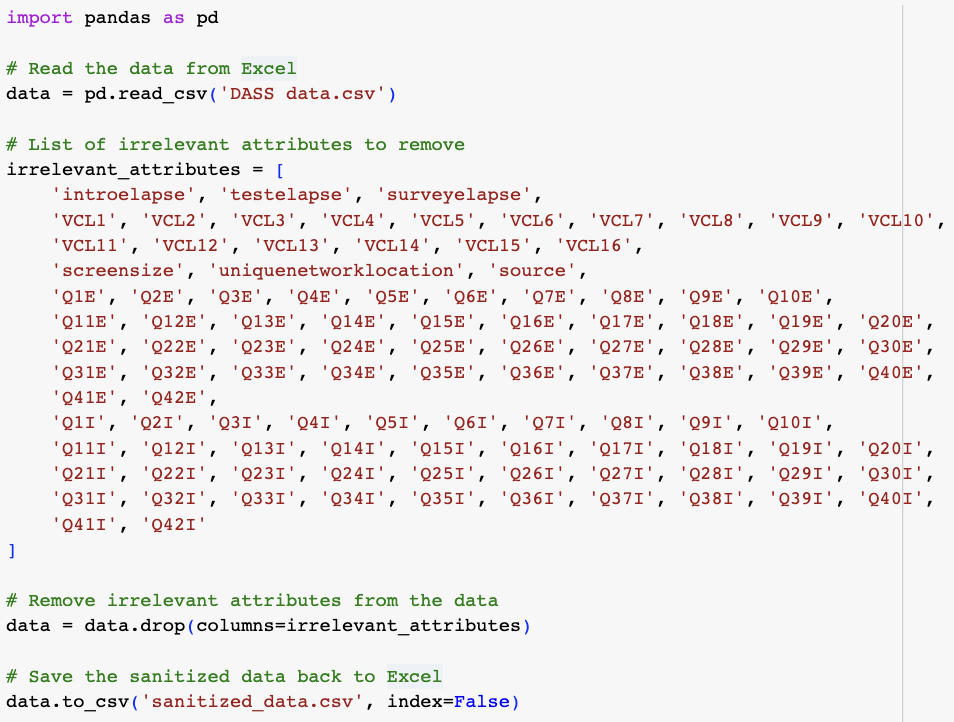
Why do we need to analyze datasets?

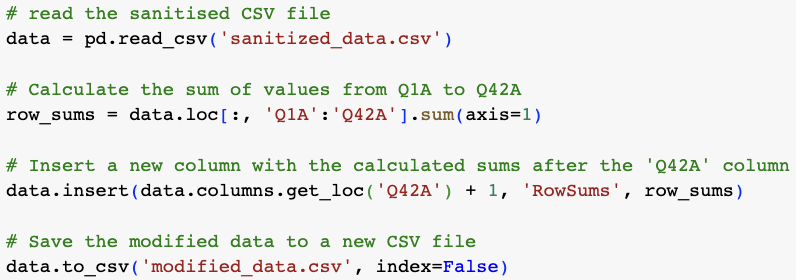
Datasets frequently include a wealth of information, and studying them helps us to extract useful insights. We can discover patterns, trends, correlations, and outliers in the data, which can help us make decisions and solve problems. These insights can be used to promote change, optimize processes, and direct strategic planning.

**Results and Analysis**

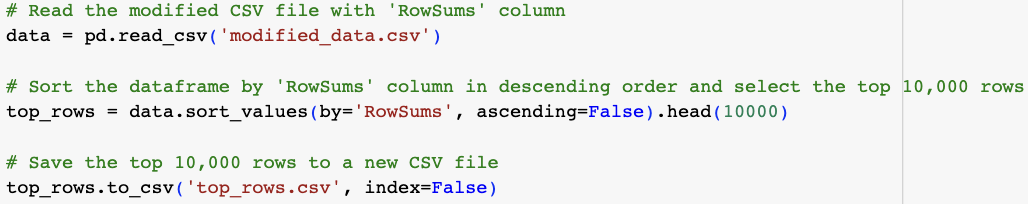
1. Dataset Analysis

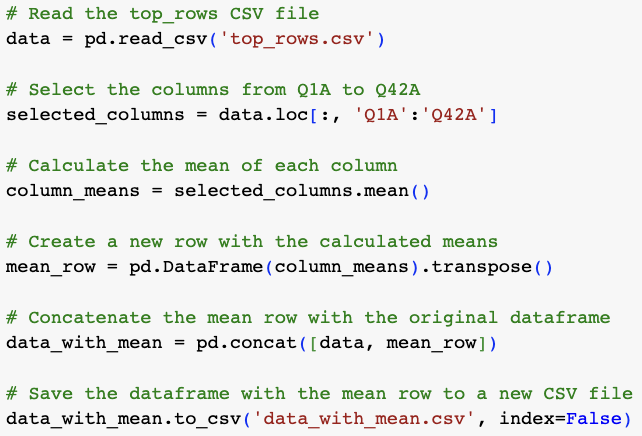
In the Analysis stage, we obtained a dataset from Kaggle that surveys participants using the Depression Anxiety Stress Scales (DASS). This questionnaire requires participants to answer 42 questions in terms of how much the particular question applies to them. Participants ranked their answers from a scale of 1 to 4, with 1 being the lowest end of how much the question resonates with them. Their total score for all 42 questions are tallied up, with higher scores being an indication of the presence of depression and/or anxiety.





We first began our analysis of this dataset by sanitizing the data. This was done by querying ChatGPT for an appropriate code to remove the redundant columns of the dataset. After the redundant columns were removed, we created a new column to sum the total scores of all 42 questions for each respondent.





Using the total scores for each participant, the top 10,000 rows were filtered to obtain the respondents with a higher likelihood of depression and/or anxiety. With the scores from these respondents that are at a higher risk of depression and/or anxiety, the mean rating of each question was calculated. This was done to determine the people at greater risk of depression and/or anxiety, which of the symptoms or experiences are more common amongst them.

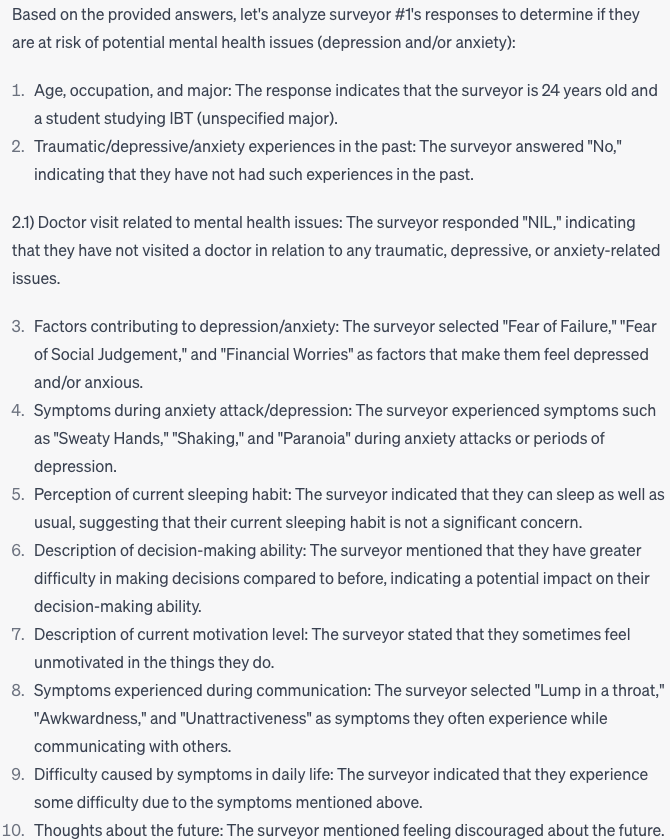
1. Survey Analysis

After we obtained the symptoms or experiences that are more common amongst people at greater risk of depression and/or anxiety, we did a matching from the dataset questions with the 10 highest means to our survey questions. This was done to see which of our survey questions are better indicators of depression and/or anxiety based on the means that we have obtained from our survey results.

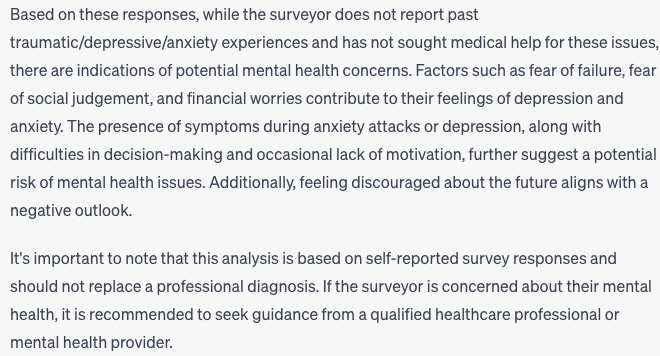
Based on the ranking of our questions as indicators of depression and/or anxiety, we instructed ChatGPT based on this ranking. We instructed it by inputting each of our surveyors' answers to determine their risk of depression and/or anxiety. Below are some examples of us doing so with ChatGPT and the results we obtained from this.

Surveyor #1

Prompt:

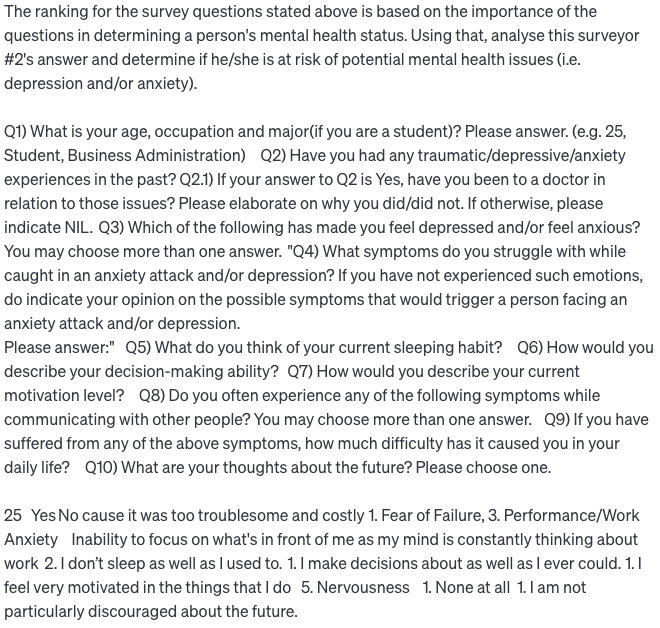


Result:

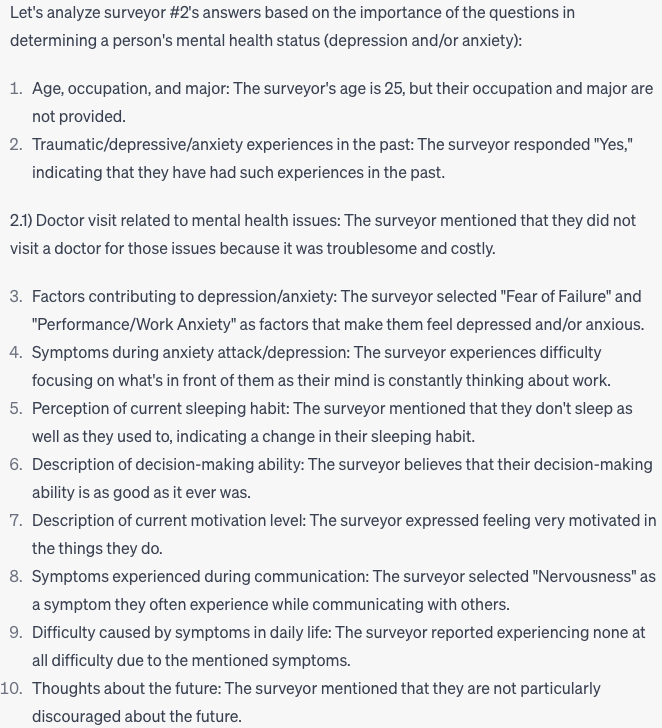
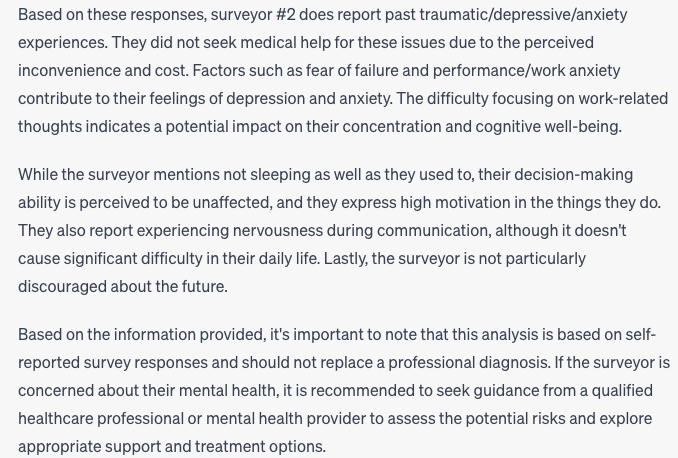


Surveyor #2

Prompt:

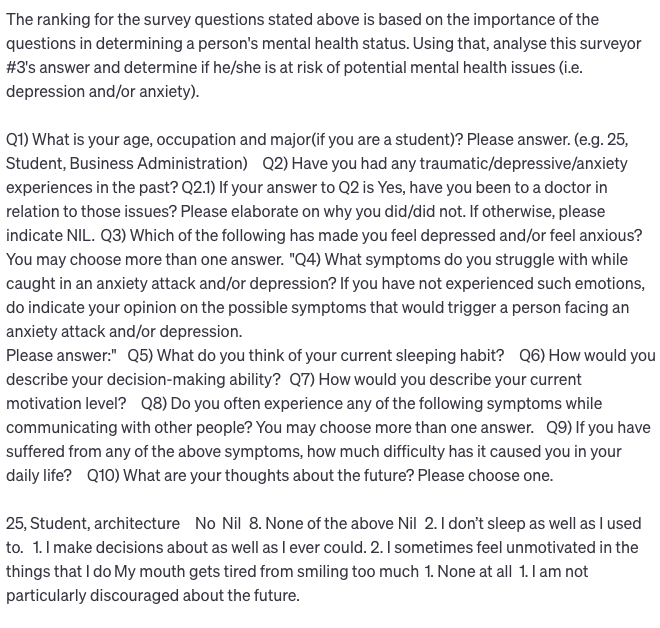


Result:

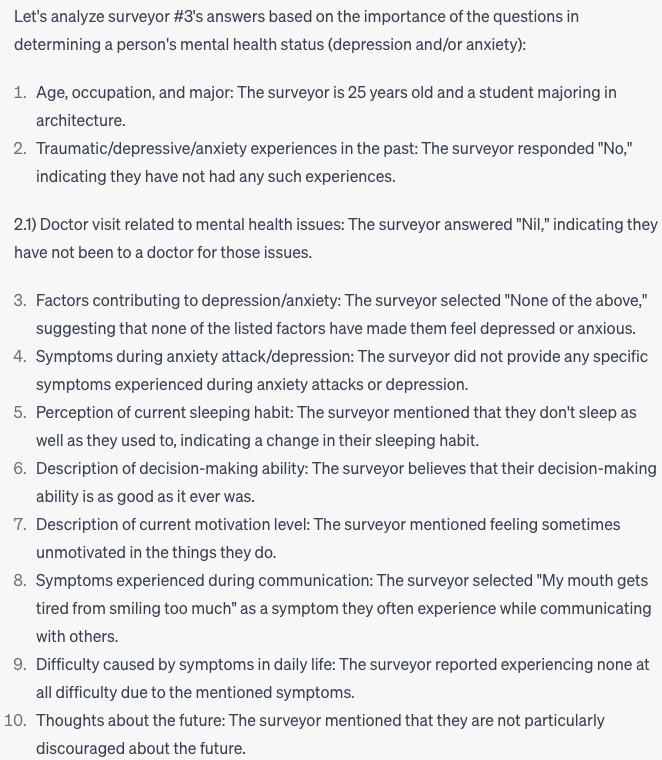
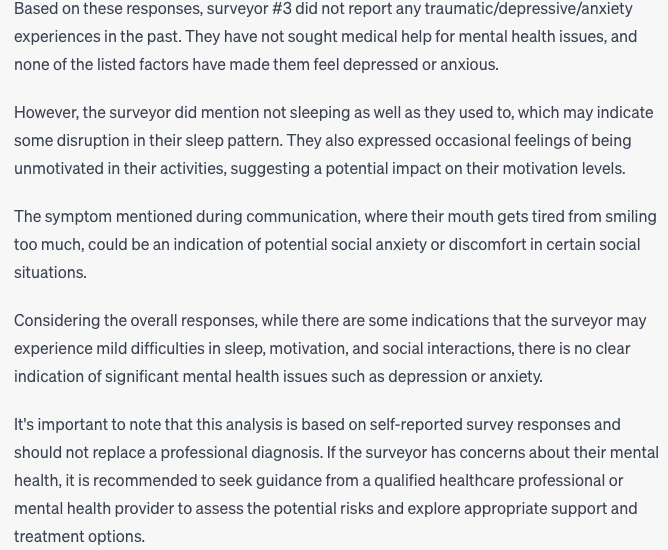
 

Surveyor #3

Prompt:

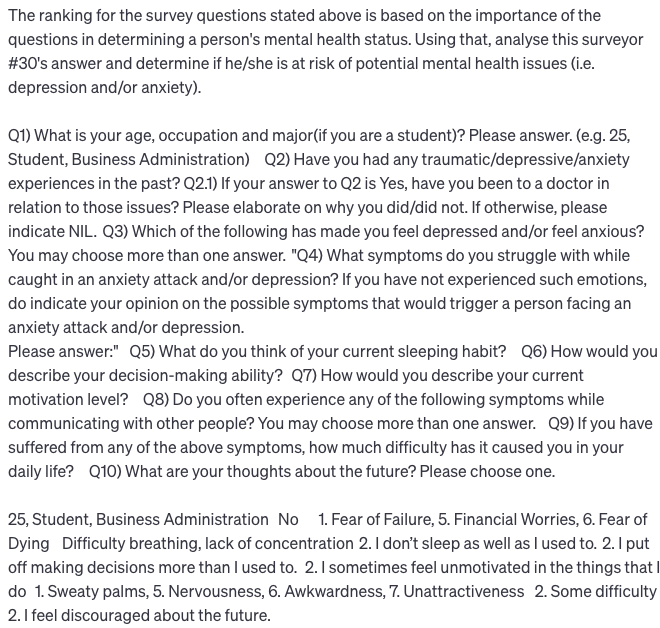


Result:

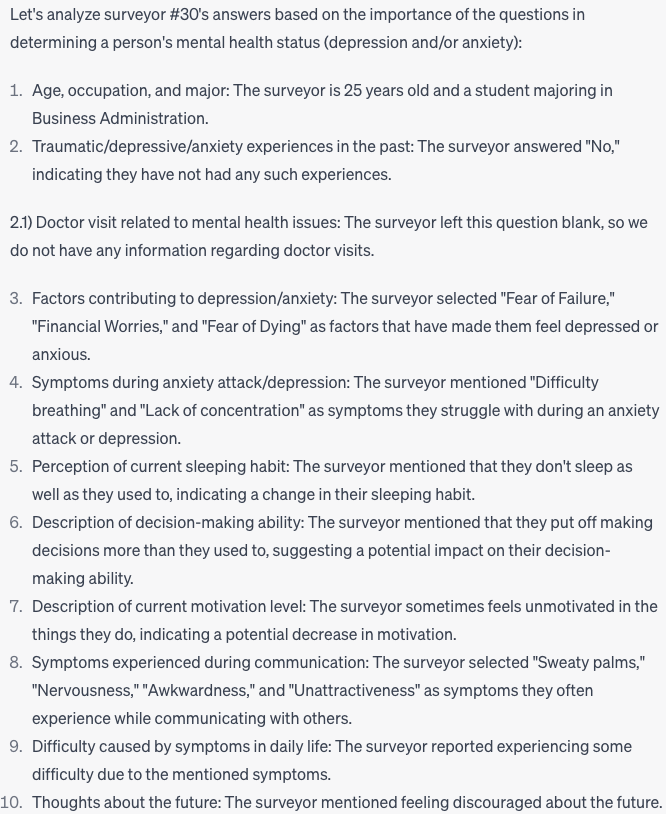
 

Surveyor #30

Prompt:



Result:

**Delegation**

During our project all of our team members put great effort in conducting research, surveys and all other parts. We all together worked on the survey questions, then we divided parts in order to have higher productivity.   
Spencer and Felester worked on Dataset and Survey result analysis. Also they worked on codebooks and with Kim Sol did DASS data. Spencer and Felested created jupiter notebook with all needed codes.

Ali was responsible for finalizing the written report, but all other members also took part. Kim Sol Prepared PPT slides and worked on documents.

Distribution part was divided as following:

* Dataset and survey analyzing with python codes - Lee Lang Feng Felester, Spencer Lim Wei Yang
* PPT slides - Sol Kim
* Final documents - Sol Kim, Abdullaev Fikratali
* Oral presentation - All group members

**References**

1.World Health Organization. (2022, June 17). WHO highlights urgent need to transform mental health and mental health care [Press release]. Retrieved from <https://www.who.int/news/item/17-06-2022-who-highlights-urgent-need-to-transform-mental-health-and-mental-health-care>

2.Cuijpers, P., Karyotaki, E., Weitz, E., Andersson, G., Hollon, S. D., & van Straten, A. (2019). The effects of psychotherapies for major depression in adults on remission, recovery, and improvement: A meta-analysis. Clinical Psychology Review, 70, 105-123. <https://doi.org/10.1016/j.cpr.2019.03.003>

3.Anxiety and Depression Association of America. (n.d.). Facts & Statistics. Retrieved from <https://adaa.org/understanding-anxiety/facts-statistics>

4.Rudin, C., & Tiller, J. (2021). The opportunities and challenges of using artificial intelligence in the healthcare sector. Nature Medicine, 27(1), 5-7. <https://doi.org/10.1038/s41591-021-01614-0>

5.World Health Organization. (2021, June 28). WHO issues first global report on AI in health and six guiding principles for its design and use [Press release]. Retrieved from <https://www.who.int/news/item/28-06-2021-who-issues-first-global-report-on-ai-in-health-and-six-guiding-principles-for-its-design-and-use>

Dataset: <https://www.kaggle.com/datasets/lucasgreenwell/depression-anxiety-stress-scales-responses?select=codebook.txt>