SAMHSA Data Visualization & Sentiment Analysis |Team 8

Client:
Prof. Rooshey Hasnain
Department of Disability and Human Development

Kumud Dubey | Keerthi Raguri | Mehul Gupta | Sathwik Maddi
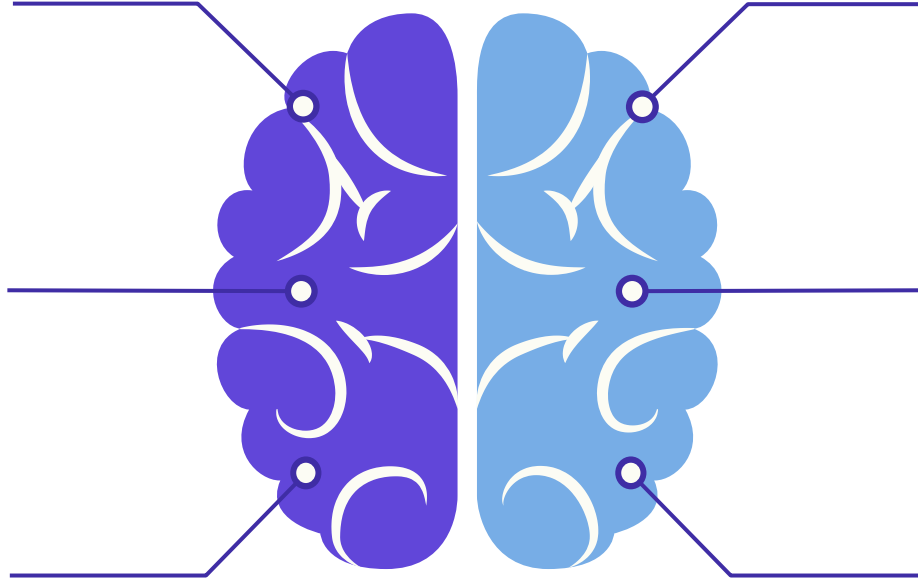
Table of Contents

INTRODUCTION

APPLIED HEALTH SCIENCES

Department of Disability and Human Development

The Department of Disability and Human Development (DHD) at UIC is renowned academic unit that provides undergraduate and graduate degree programs in the field of disability studies.

It conducts research and disseminate information about disability to various stakeholders.

It is involved in various community projects that aim to promote the inclusion and well-being of people with disabilities

INITIAL SOW VS UPDATED SOW

Creating a digital repository that showcases textual data of experiences that have been collected from people with disabilities and mental health issues from the Asian community to the general public, department partners, and sponsors

Performing analytics on data on mental health issues and treatments, obtained from SAMHSA - Substance Abuse and Mental Health Services Administration to gain insights on the state and extent of mental health issues prevalent among various ethnic groups in the US.

The objective is to raise awareness on the hardships and discrimination faced by these individuals in day to day life and showcase the work the department and community partners are doing in addition to attracting more sponsorship and funding to their programs.

Sentiment analysis of stories gathered by students of the DHD from individuals in the local community that have shared their experiences with disabilities and mental illnesses

BUSINESS VALUE
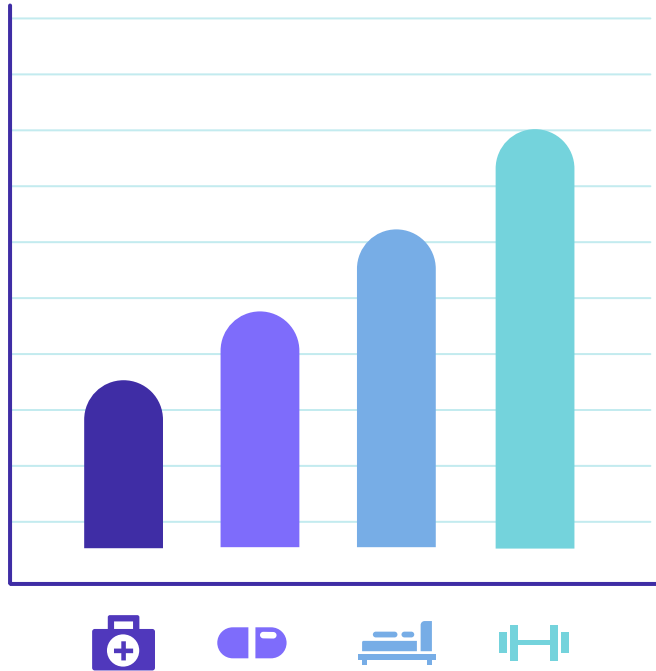
Enhanced understanding of mental health trends

Identification of high-risk mental issues

Advocacy and policy development

Collaboration and knowledge sharing

Integrate the visualizations into Prof. Rooshey's publications based on mental health

Original Plan

Week	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12	Week 13
Task													
1. Client Touchbase	■	■											
2. Understanding Project Statement		■	■										
3. Define Deliverables			■	■									
4. Identify appropriate Platform (UIC RED)					■	■							
5. Research and gather data							■	■					
6. Install and configure software								■	■				
7. Create user interface design										■	■		
8. Develop front-end										■	■		
9. Perform testing										■	■	■	
10. Populate digital repository												■	
11. Launch digital repository													■

Scope Changes

1

The original project **did not fully leverage** the analytical skills and techniques associated with the field of business analytics

2

Maximise the potential of the project by incorporating data visualisation

3

A comprehensive view of mental health data, facilitating **informed** decision-making

4

Engage stakeholders, including policymakers, researchers, clinicians, and the general public.

Underutilized Analytics

Maximize Insights

Better Decision

Stakeholder Engagement

Final Gantt Chart

(APR 12 - APR 27)

Week Task	Week 1	Week 2	Week 3
1. Planning and Data Preparation	🟩		
2. Identify key variables of interest from the Survey	🟩	🟩	
3. Conduct exploratory data analysis		🟩	
4. Create visualizations such as bar charts, pie charts		🟩	🟩
5. Perform sentiment analysis on relevant text data from the stories			🟩
6. Create word cloud or reports to present sentiment analysis			🟩
7. Finalize the project documentation			🟩

Dataset

```
df.columns

Index(['YEAR', 'AGE', 'EDUC', 'ETHNIC', 'RACE', 'GENDER', 'SPHSERVICE',
       'CMPSERVICE', 'OPISERVICE', 'RTCSERVICE', 'IJSSERVICE', 'MH1', 'MH2',
       'MH3', 'SUB', 'MARSTAT', 'SMISED', 'SAP', 'EMPLOY', 'DETNLF', 'VETERAN',
       'LIVARAG', 'NUMMHS', 'TRAUSTREFLG', 'ANXIETYFLG', 'ADHDFLG',
       'CONDUCTFLG', 'DELIRDEMFLG', 'BIPOLARFLG', 'DEPRESSFLG', 'ODDFLG',
       'PDDFLG', 'PERSONFLG', 'SCHIZOFLG', 'ALCSUBFLG', 'OTHERDISFLG',
       'STATEFIP', 'DIVISION', 'REGION', 'CASEID'],
      dtype='object')
```

National Mental Health Services Survey (N-MHSS) 2020 from SAMHSA which is an annual survey that collects data on the services and characteristics of all known mental health treatment facilities **in the 50 states.**

The provided dataset **contains various variables** from the SAMHSA (Substance Abuse and Mental Health Services Administration) National Mental Health Services Survey

Tools Used for Analysis
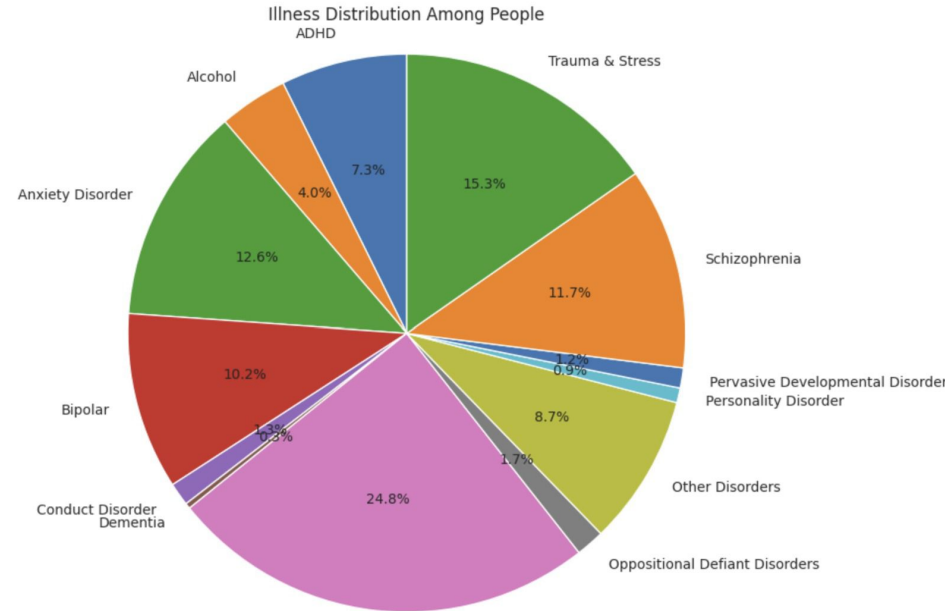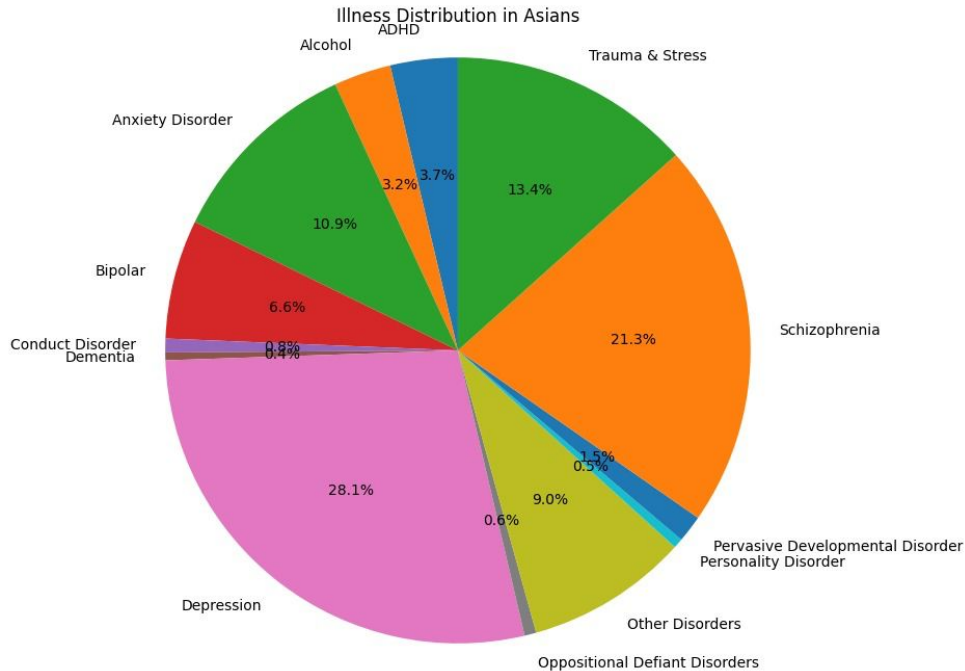
Tools used:

- ***Tableau***
- ***Python***

Libraries used:

- ***Matplotlib*** (Visualization)
- ***Plotly*** (Visualization)
- ***Seaborn*** (Visualization)
- ***NLTK*** (Sentiment Analysis)
- ***Vader Lexicon*** (Sentiment Analysis)

Deliverable 1 : SAMHSA Data Visualizations

Illness Distribution in Asians

- Trauma & Stress — 13.4%
- Schizophrenia — 21.3%
- Pervasive Developmental Disorder
- Personality Disorder — 1.5%, 0.5%
- Other Disorders — 9.0%
- Oppositional Defiant Disorders — 0.6%
- Depression — 28.1%
- Conduct Disorder — 0.8%
- Dementia — 0.4%
- Bipolar — 6.6%
- Anxiety Disorder — 10.9%
- Alcohol — 3.2%
- ADHD — 3.7%

Illness Distribution Among People

- Trauma & Stress — 15.3%
- Schizophrenia — 11.7%
- Pervasive Developmental Disorder — 1.2%
- Personality Disorder — 0.9%
- Other Disorders — 8.7%
- Oppositional Defiant Disorders — 1.7%
- 24.8%
- Conduct Disorder — 1.3%
- Dementia — 0.3%
- Bipolar — 10.2%
- Anxiety Disorder — 12.6%
- Alcohol — 4.0%
- ADHD — 7.3%

- Distribution of illnesses in the surveyed population are - ***depression***, ***schizophrenia***, ***trauma & stress induced disorders*** and ***anxiety***.

Most prevalent illnesses across *genders* in the population.
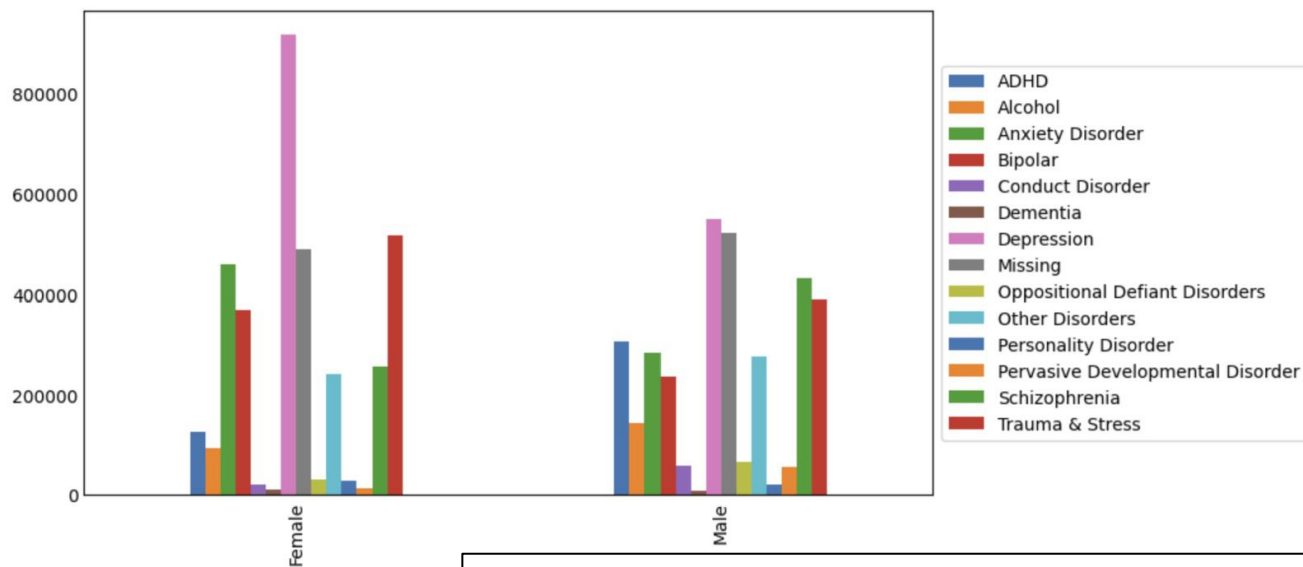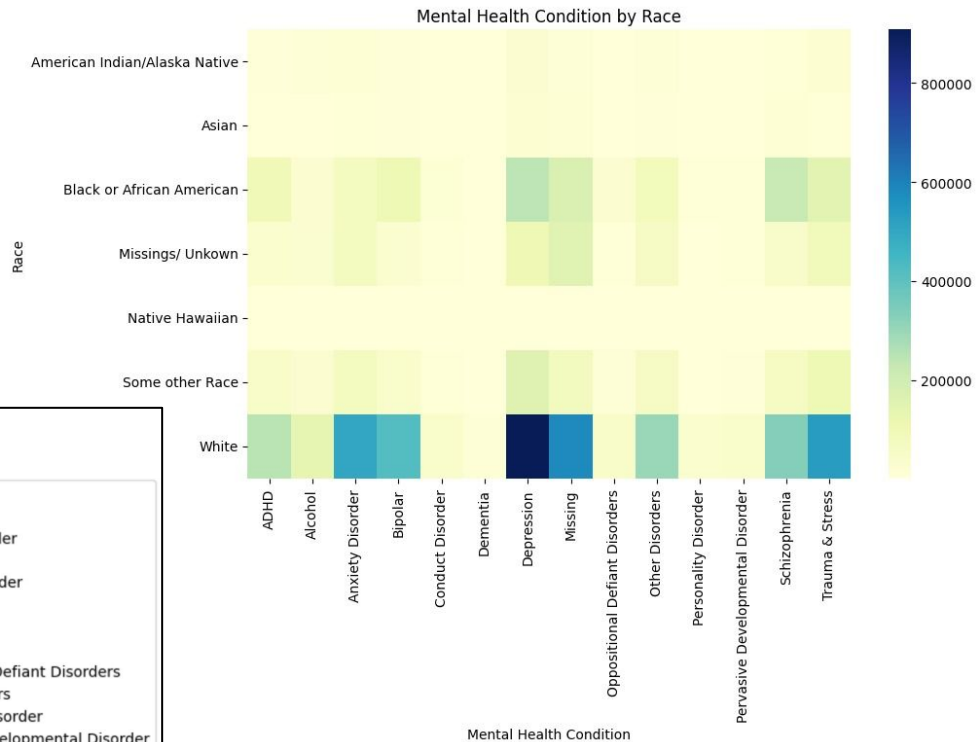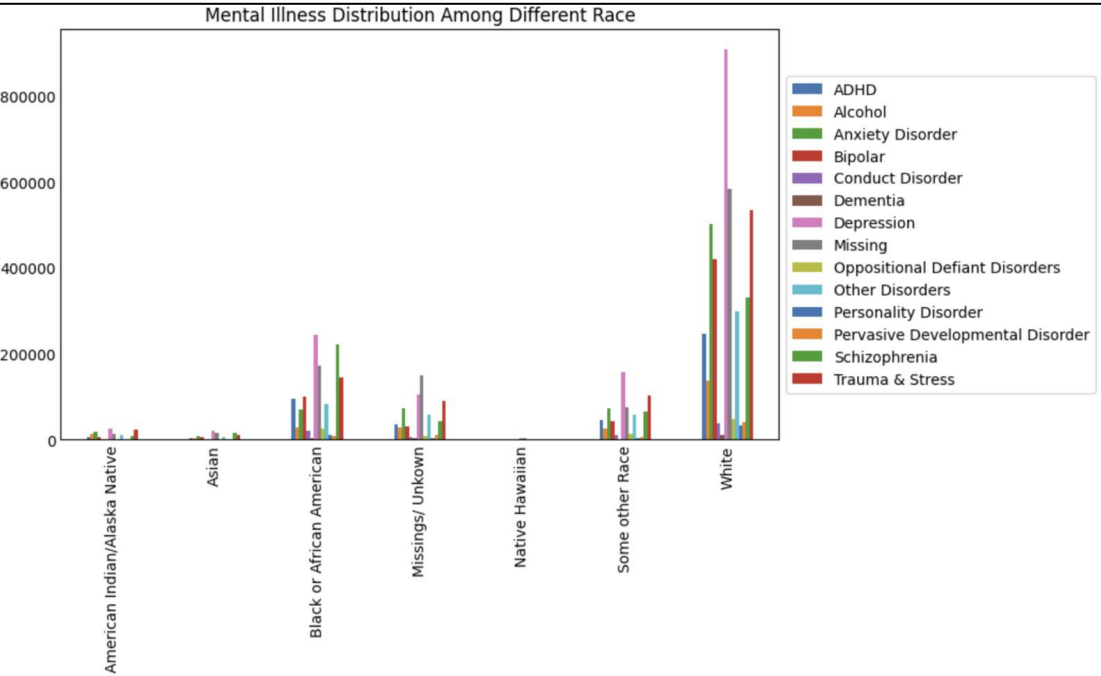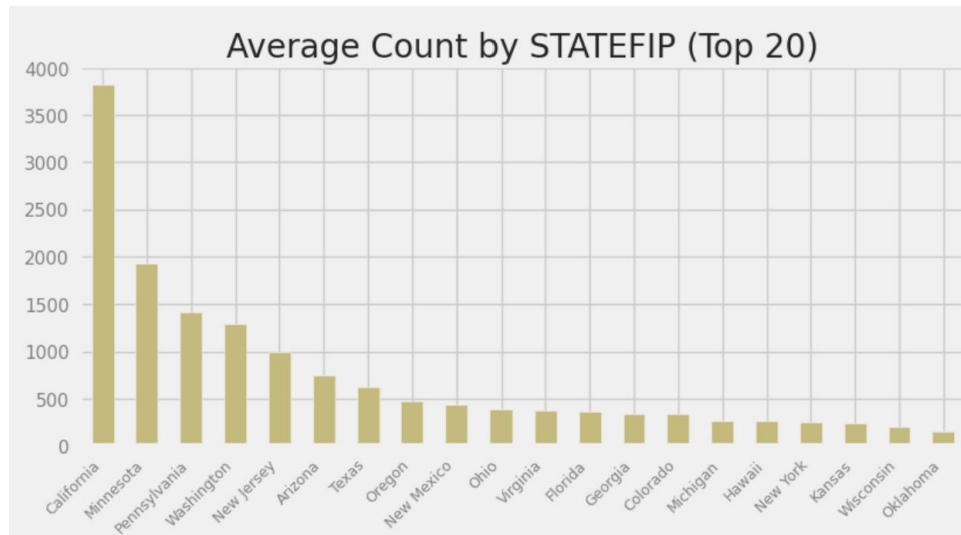
Illnesses across males/females among *asians*.
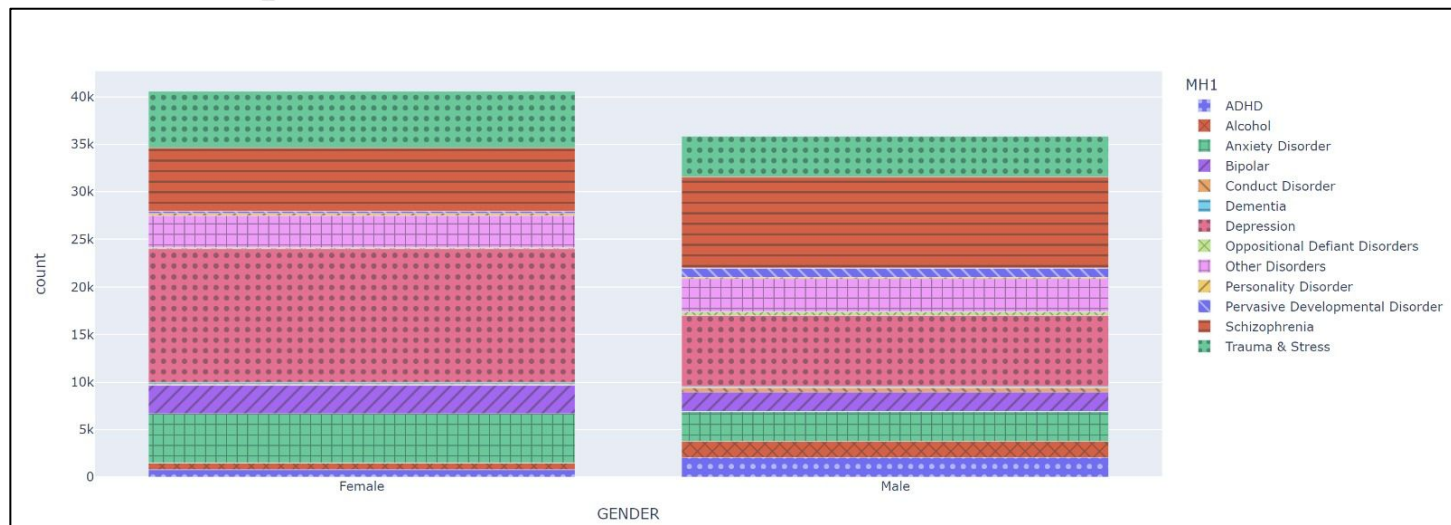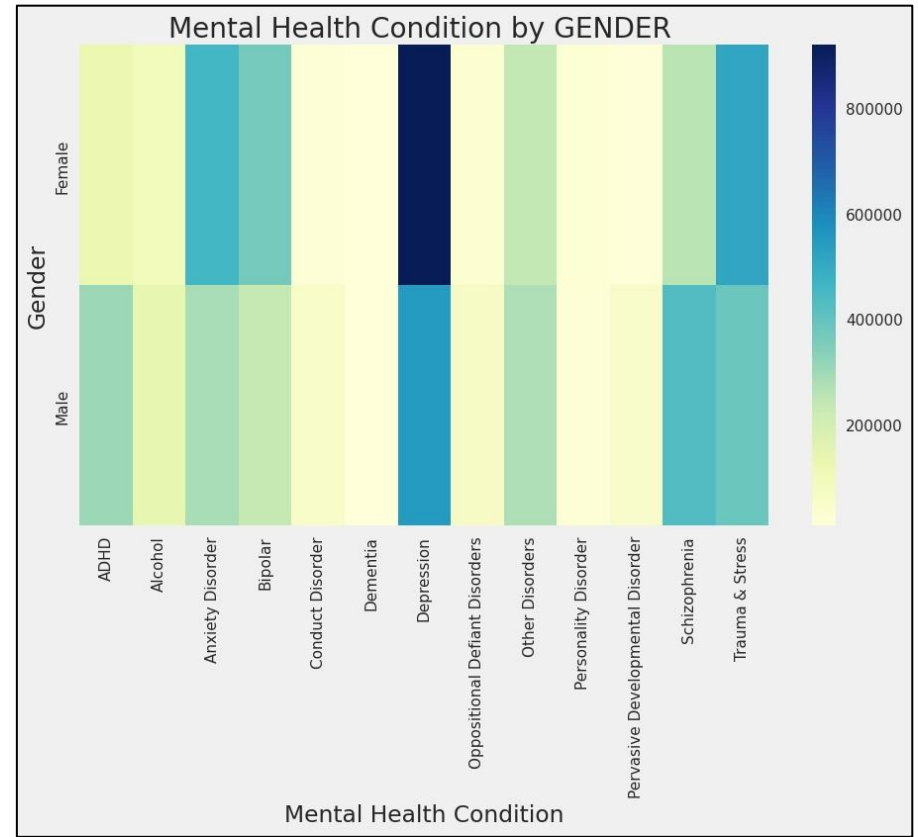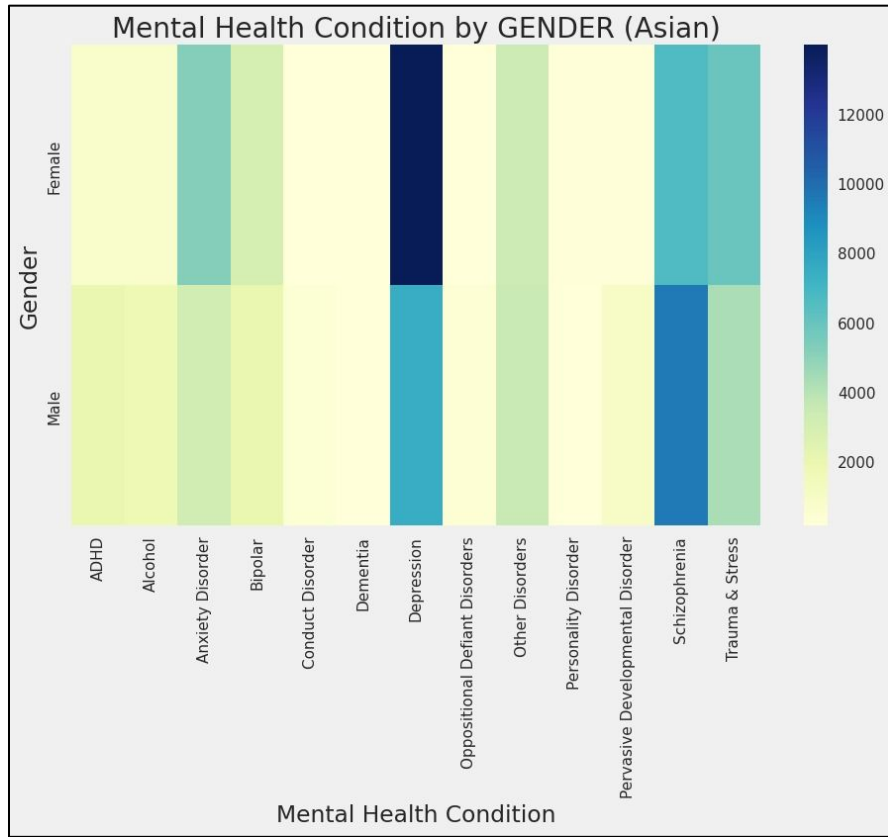
Heat maps - distribution of illnesses across *genders*.
Findings - widespread presence of *schizophrenia* (which is hard to diagnose) among *asians*.

Prevalence of mental illnesses by *race*.

Mental Health Condition by Race

Mental Illness Distribution Among Different Race

Prevalence of depression across *states/geographies*.

Prevalence of mental illnesses across *genders* & *education levels*. Notice *trauma* & *stress* as education levels lower.

Deliverable 2 : Sentiment Analysis

What's That?

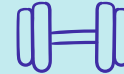Uses Machine learning and NLP techniques to identify and extract information from text

Goal: Determine the overall sentiment or emotional tone behind the text

Why we use it?

- To find out the emotional tone behind our stories
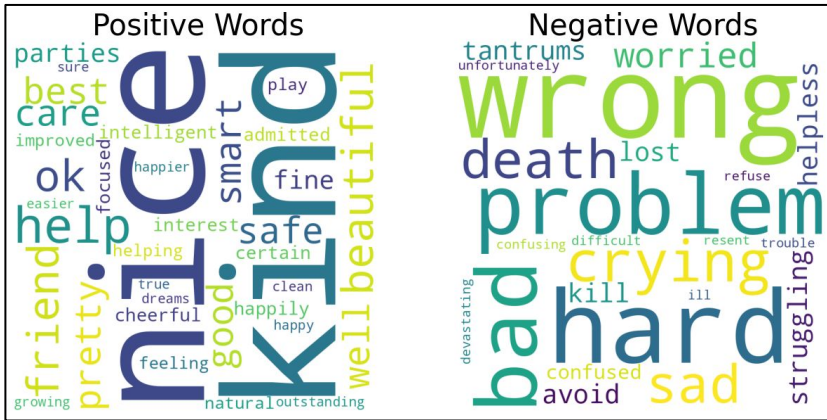- How people perceive the mentally ill patients

- Visualizing the positive and negative words

How we used Sentiment Analysis

- WordCloud

  Ashley

- Polarity Scores
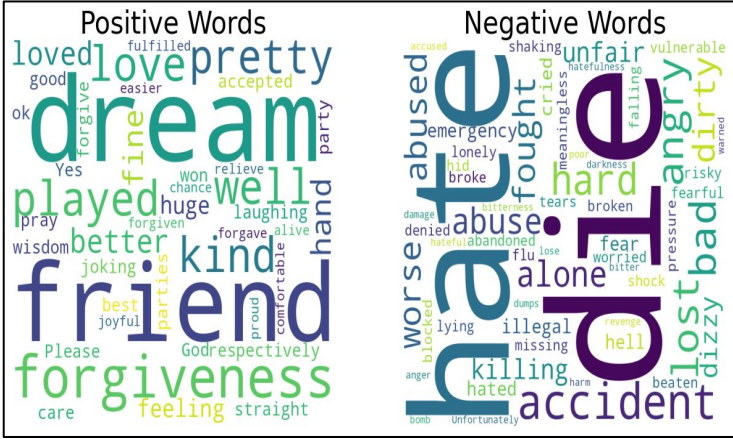
{'neg': 0.079, 'neu': 0.807, 'pos': 0.114, 'compound': 0.9965}

- Emotional Tone

The emotional tone of the text is positive.

● WordCloud

● Polarity Scores

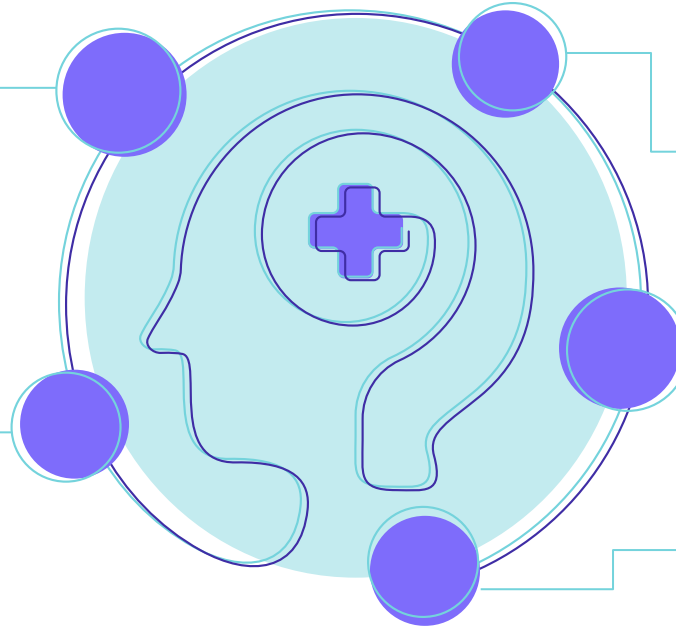{'neg': 0.174, 'neu': 0.75, 'pos': 0.076, 'compound': -0.9998}

● Emotional Tone

The emotional tone of the text is negative.

Conclusion

Impact: Provided the DHD with material they could use to support their research, community work and which helps them have more informed conversations with key stakeholders such as policymakers, etc.

Challenges: Data availability.

DEFINE and CEMENT deliverables as soon as possible.
-Decision making.
-Measure progress.
-Manage expectations.

Refined our client communication skills.

Practical application of knowledge gained from the course.