

IDS 560 Analytics Strategy Practice

Final Project Report

SAMHSA Data Visualization &

Sentiment Analysis

Client:

Prof. Rooshey Hasnain
Department of Disability and Human Development

Team 8

Kumud Dubey | Keerthi Raguri | Mehul Gupta | Sathwik Maddi

Spring Semester, 2023

Index

Content	Page No.
Introduction	1-5
Original Plan - Gantt Chart	5
Key milestones tasks	6
Scope Changes	7
Final Gantt Chart	7
Description of the Solution and the Final Deliverables	8 - 20
Next Steps for the Sponsor	21
Conclusion	22

Introduction

Sponsor: Prof. Rooshey Hasnain, UIC Department of Disability and Human Development, which is a part of the College of Applied Health Sciences. The department is an internationally-recognized center for the interdisciplinary study of disability, conducting scholarship and community-engaged service across the spectrum of disability, including advocacy, culture, education, health promotion, history, policy and technology.

The department's faculty and staff are committed to advancing the rights and inclusion of people with disabilities in all areas of society. They conduct research, provide training, and engage in advocacy efforts to improve the lives of people with disabilities and promote their full participation in society. They conduct research and disseminate information about disability to academicians, policymakers, businesses, government agencies, service providers and the general public. They also provide clinical and community services and offer interdisciplinary pre-service training.

The department is involved in various community projects that aim to promote the inclusion and well-being of people with disabilities. Here are a few examples of their community projects:

1. Access Living - The DHD collaborates with Access Living, a non-profit organization that advocates for the rights of people with disabilities, to promote independent living and community integration. The DHD provides support and training to Access Living staff and helps to develop disability-related policies and programs.
2. Chicago Housing Authority - The DHD partners with the Chicago Housing Authority to improve accessibility and accommodations for people with disabilities in public housing. The department provides training and technical assistance to housing authority staff and residents to promote the inclusion of people with disabilities in their communities.
3. Special Olympics - The DHD collaborates with Special Olympics to promote the inclusion of people with intellectual disabilities in sports and recreational activities. The department provides training and support to Special Olympics coaches and athletes to improve their athletic abilities and promote their social and emotional well-being.
4. Disability Rights Consortium - The DHD is a member of the Disability Rights Consortium, a coalition of disability rights organizations that advocate for disability-related policies and programs at the local, state, and national levels. The

department provides expertise and support to the consortium to advance disability rights and inclusion.

Project Overview and Business Value:

The UIC Department of Disability and Human Development can derive several valuable insights and benefits from SAMHSA data visualization and sentiment analysis. Here are some potential advantages:

Enhanced understanding of mental health trends: By analyzing SAMHSA data through data visualization techniques, the UIC Department can gain a comprehensive overview of mental health trends and patterns. This information can help identify areas of concern, such as prevalent mental health issues or demographic groups disproportionately affected by certain conditions.

Identification of high-risk populations: Sentiment analysis can be used to assess the emotional well-being and sentiments expressed by different populations regarding mental health. By analyzing sentiment across various demographics, the UIC Department can identify high-risk populations that may require targeted interventions or support.

Tailored interventions and resource allocation: Analyzing SAMHSA data through data visualization and sentiment analysis can inform the development of targeted interventions and resource allocation strategies. For example, if sentiment analysis reveals negative sentiments in a particular region or community, the UIC Department can allocate resources and design interventions to address the specific mental health challenges faced by that population.

Program evaluation and effectiveness: Data visualization can provide a clear and accessible way to monitor the effectiveness of various mental health programs and initiatives. By visualizing key metrics and outcomes, the UIC Department can assess the impact of interventions and make data-driven decisions regarding program improvement or expansion.

Advocacy and policy development: SAMHSA data visualization and sentiment analysis can serve as powerful tools for advocacy and policy development. The UIC Department can leverage these insights to support evidence-based policy recommendations, raise

awareness about mental health issues, and advocate for improved services and resources.

Collaboration and knowledge sharing: Sharing data visualizations and sentiment analysis findings can foster collaboration among different stakeholders, including researchers, policymakers, and community organizations. By providing clear and compelling visual representations of data, the UIC Department can facilitate discussions, encourage knowledge sharing, and promote a collaborative approach to addressing mental health challenges.

In summary, the key business values provided are:

- 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 on mental health.

Original Plan - Gantt Chart

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													

Key Milestone Tasks :

Planning and Data Preparation:

During week 11, the project team was focused on planning the project and preparing the data for analysis. This involved identifying the data sources and cleaning and organizing the data to ensure it is ready for analysis. We've chosen the 'mental health client-level data: 2020' dataset from SAMHSA. The team was to determine which variables they want to focus on for analysis, and ensure that the data is formatted in a way that can be easily analyzed. The key milestone for this week will be to have a clean and organized data set that is ready for analysis.

Identify key variables of interest from the dataset, Conduct exploratory data analysis, and Create visualizations

During week 12, the project team was focused on identifying the key variables of interest from the SAMHSA dataset. Once the key variables were identified, the team conducted exploratory data analysis to gain insights into the data. This involved creating summary statistics, identifying patterns and trends in the data, and using data visualization techniques such as bar charts and pie charts to present the data visually. The key milestone for this week will be to have a clear understanding of the key variables of interest, and to have created visualizations that effectively communicate the insights gained from the exploratory data analysis.

Sentiment analysis, Creating word cloud and Finalizing the project documentation

During the final week, the project team was focused on performing sentiment analysis on relevant text data from the stories. This involved using natural language processing techniques to analyze the sentiment of the text data, and to identify key themes and patterns in the data. The team then created word clouds and reports to present the sentiment analysis in a clear and concise manner. Finally, the team finalized the project documentation, including a report on the findings and recommendations based on the analysis. The key milestone for this week will be to have a comprehensive report that effectively communicates the insights gained from the analysis, and provides clear recommendations for future action.

Scope Changes

In the original plan, our task was to create a website using the UIC Red multisite network which included a UIC branded theme framework for a consistent look among all UIC department websites. And second deliverable was to create a website Using Wix for class DHD 400, which would cover the curriculum, syllabus and supplemental document. After extensive deliberation with the client, the project scope was revised. The client determined that a website with data visualization and sentiment analysis focused on SAMHSA (Substance Abuse and Mental Health Services Administration) data would better align with the principles and objectives of business analytics. This change was made to maximize the project's potential for data-driven insights and actionable recommendations. Now by visualizing the SAMHSA data and sentiment analysis of the stories we have tried to transform complex information into visual representations, making it easier for the client to understand and act upon the insights derived from the data. This will help to make informed decisions based on data as evidence. This approach reduces reliance on anecdotal evidence, leading to more effective and impactful decision-making in the mental health domain.

Final Gantt Chart

Week	Week 1	Week 2	Week 3
Task			
1. Planning and Data Preparation			
2. Identify key variables of interest from Survey			
3. Conduct exploratory data analysis			
4. Create visualizations such as bar charts and 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			

Description of the Solution and the Final Deliverables

The revised project now focuses on SAMHSA (Substance Abuse and Mental Health Services Administration) Data Visualization and sentiment analysis. The objective is to leverage business analytics techniques to analyze SAMHSA data related to mental health issues. The data will be transformed into visual representations that provide meaningful insights and support decision-making processes. Additionally, sentiment analysis will be applied to understand the emotions and attitudes expressed in the data, contributing to a deeper understanding of mental health experiences.

The project aims to deliver the following outcomes:

- Interactive Data Visualizations: Develop visually appealing and user-friendly data visualizations to effectively communicate key findings and trends from the SAMHSA data.
- Sentiment Analysis Reports: Perform sentiment analysis on the textual data to identify positive, negative, and neutral sentiments expressed by individuals regarding their mental health experiences.
- Actionable Recommendations: Derive actionable insights from the data visualizations and sentiment analysis to provide valuable guidance for the UIC Department of Disability and other stakeholders in improving mental health services and support.

The goal of the SAMHSA Data Visualization and Sentiment Analysis Project is to use business analytics to obtain knowledge and improve decision-making on mental health issues. The project's focus on SAMHSA data intends to give the UIC Department of Disability and Human Development useful information and analysis.

Deliverable 1: SAMHSA Data Visualization

SAMHSA's National Mental Health Services Survey (NMHSS) is an annual survey that collects information on the characteristics and use of mental health treatment facilities in the United States. The survey is conducted by the Substance Abuse and Mental Health Services Administration (SAMHSA) and is used to inform policy decisions, identify gaps in service provision, and monitor trends in mental health service use.

The NMHSS collects data from a variety of mental health treatment facilities, including inpatient and outpatient mental health centers, residential treatment centers, and partial

hospitalization programs. The survey covers a wide range of topics, including staffing and personnel, treatment modalities, patient demographics, and payment sources.

The NMHSS is a valuable resource for researchers, policymakers, and mental health advocates. The data collected in the survey can be used to identify disparities in mental health service use, track trends in mental health treatment, and inform efforts to improve access to care for individuals with mental health disorders. The survey data is available to the public and can be accessed through the SAMHSA website.

The analysis and visualization of SAMHSA's NMHSS data can provide valuable insights for the Department of Disability (DHD) at the University of Illinois at Chicago (UIC). By analyzing the data, the DHD can identify trends and gaps in mental health service provision, as well as assess the effectiveness of existing policies and programs. The visualization of this data can help communicate these findings in a clear and accessible way, enabling stakeholders to make informed decisions about mental health policy and practice.

Visualization Results:

Fig. 1: Our visualization suggests that depression is a common mental illness among the Asian population in the United States comprising 28.1%, i.e. approx 25,500 out of 91,00, followed by Schizophrenia, Trauma & Stress, and Anxiety.

Fig. 2: Depression is a prevalent condition among people in the United States, irrespective of race, with 24.8%, i.e. approx 1.7 million out of 6.9 million, followed by Trauma & Stress, and Anxiety.

We can see that Schizophrenia is more common among Asian people.

Depression is the primary condition, for both genders among the population. There is a reversal in trend in regards to trauma and stress related disorders and schizophrenia among females and males, with more females having the prior and less of the latter, and more males having schizophrenia and less trauma related issues.

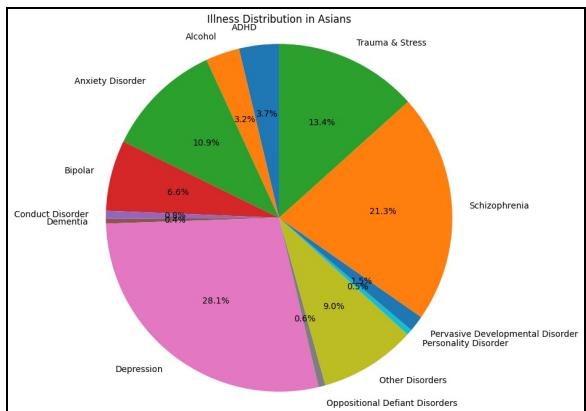


Fig. 1

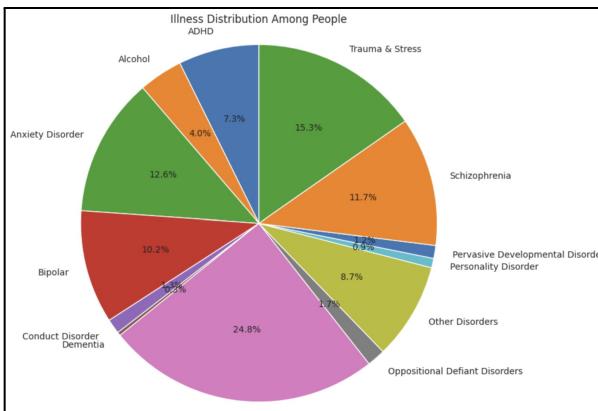


Fig. 2

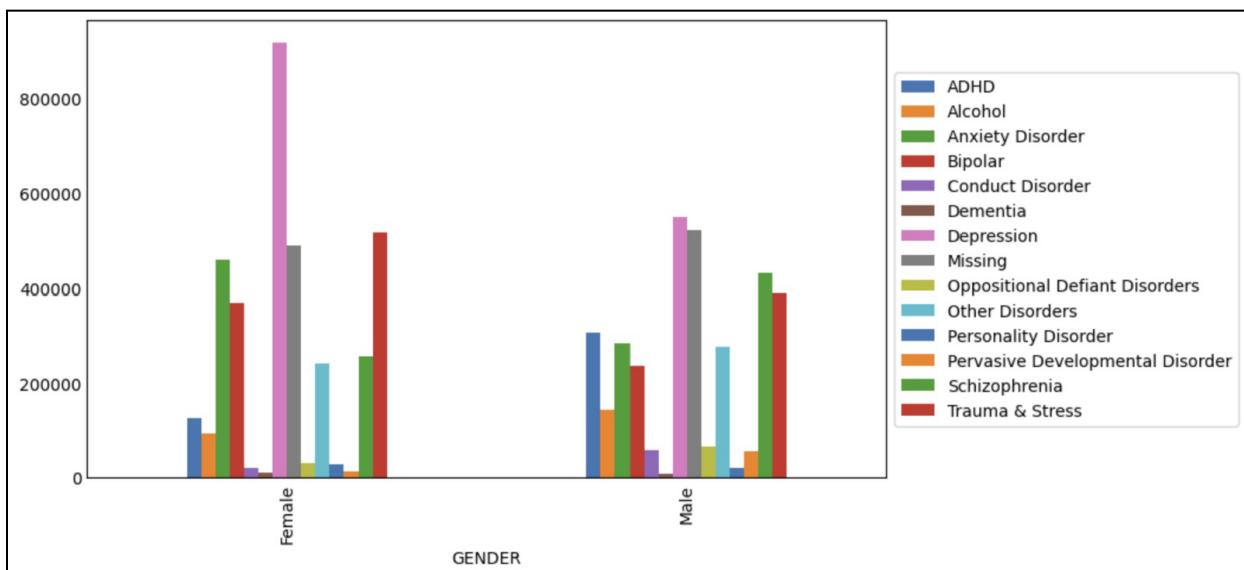


Fig. 3



Fig. 4

Fig. 4 shows the illnesses across gender among Asians. It is apparent that depression again is the most widespread issue for both genders, while it is much higher among Asian females, as it was for the entire population. As expected, alcohol related issues are higher among males, which is inline with alcohol consumption trends, which is also higher for men. The graph is a stacked bar chart plotted using the Plotly library.

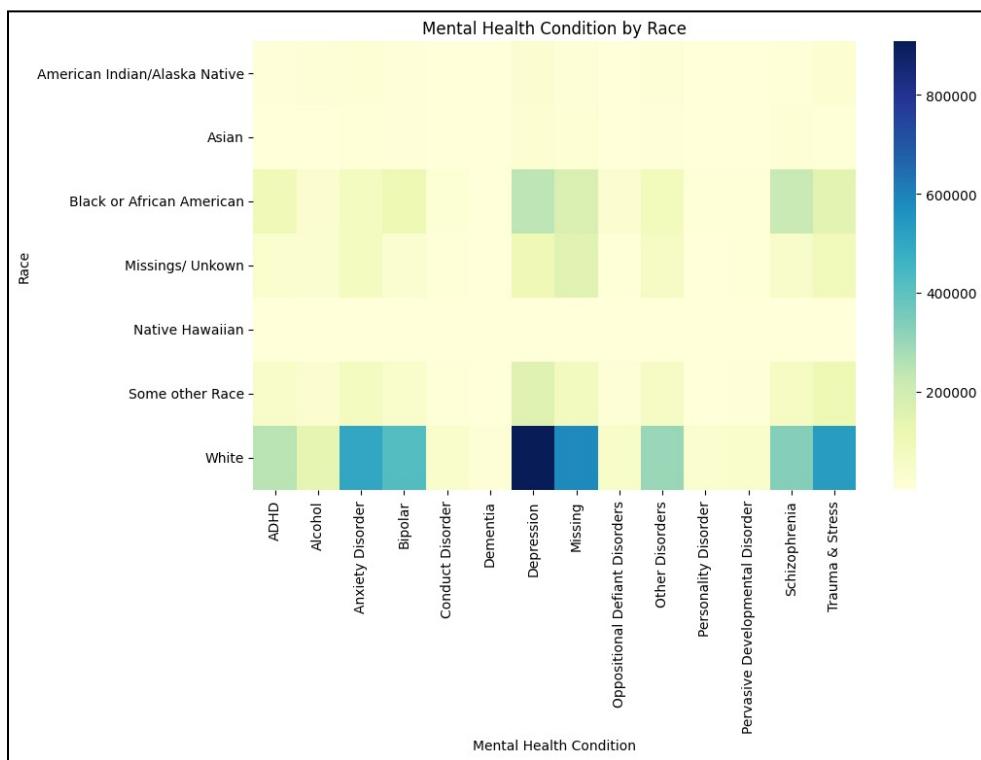


Fig. 5

Fig. 5 is a heat map showing the illness distribution among the entire population across various races, which shows that the most prevalent mental disease across all genders is depression, among caucasians.

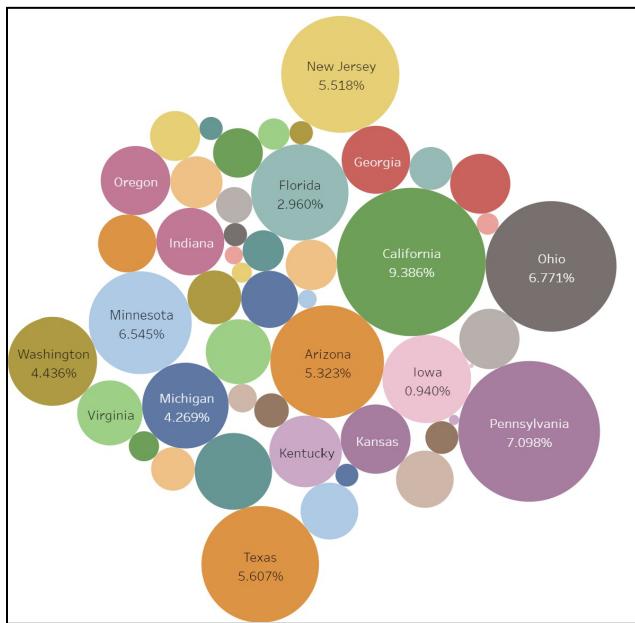


Fig. 6

California leads the United States in the most reported mental illness cases, accounting for 9.3%, i.e., approximately 641,700 people.

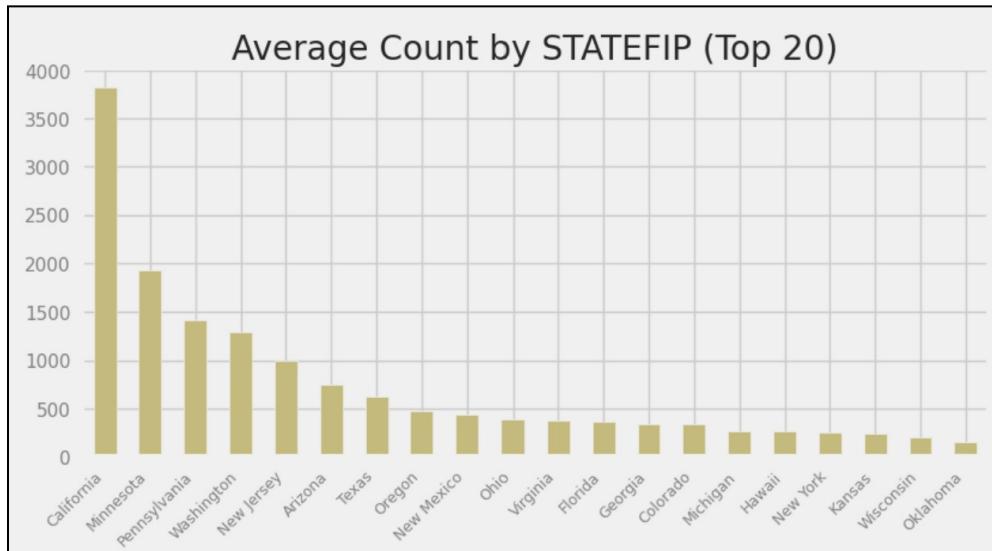


Fig. 7

Fig. 7 is an interesting visualization, which shows the average of the number of diagnoses for various mental illnesses across different states. It can be seen that the average number of diagnoses for any kind of mental health issue in California in this study is approx. 3800.

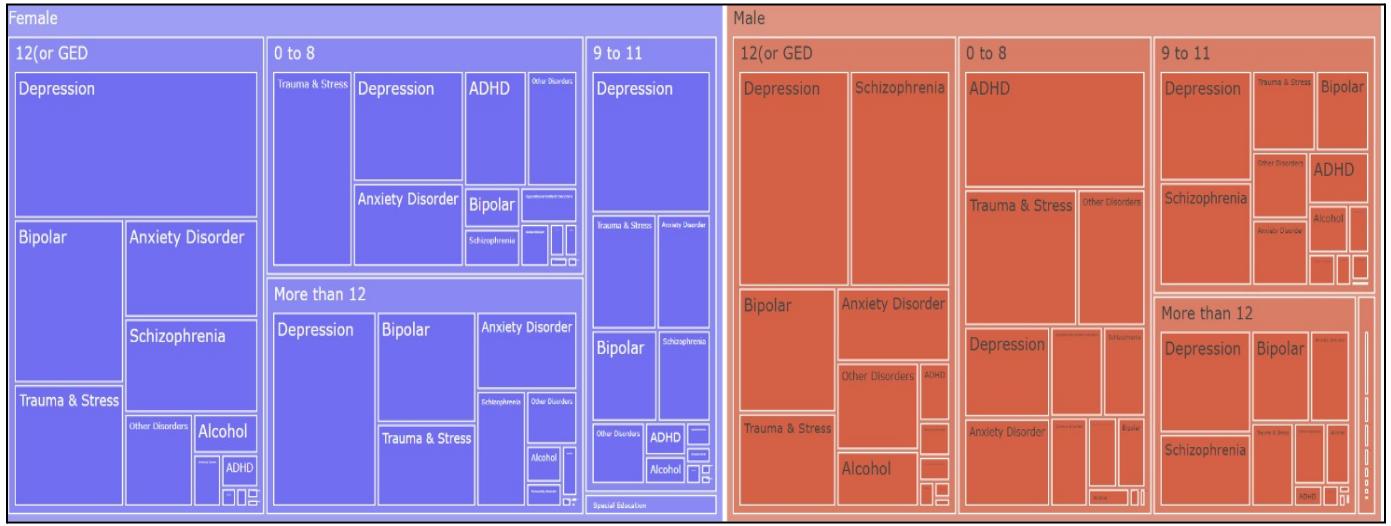


Fig. 8

Fig. 8 shows the variation in mental illness prevalence among both the genders across different education levels in the entire population. One observation from this visualization is that as the education level drops, trauma and stress becomes increasingly widespread. This may go to show that people with early childhood trauma and stress drop out of school, explaining the lower education levels achieved.

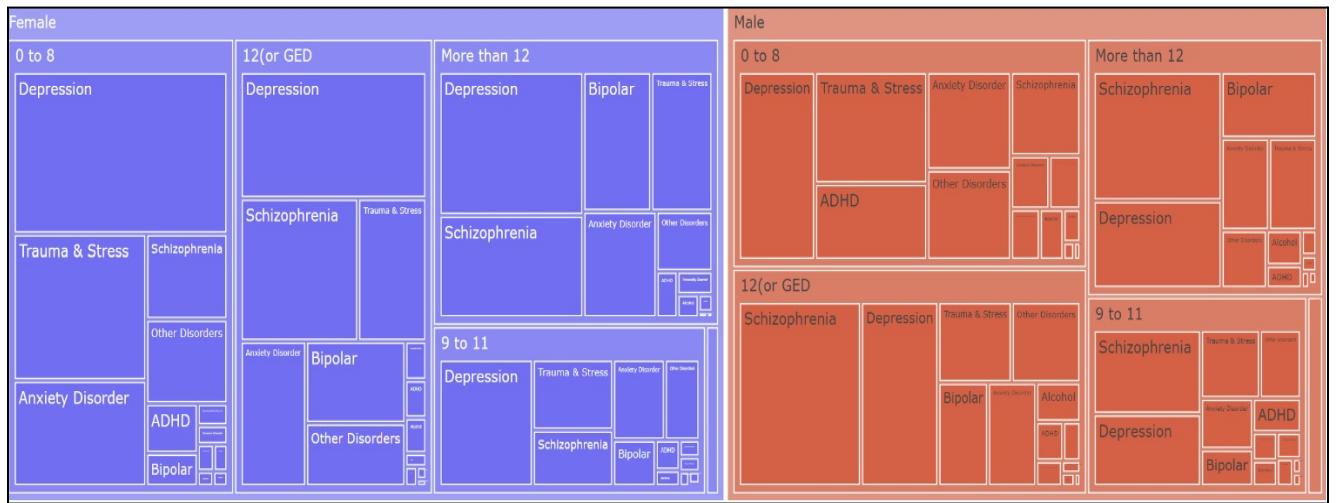


Fig. 9

Fig. 9 shows the same as the previous figure, but for the Asian population. It can be seen that for Asian men, schizophrenia is more prevalent than for the entire population. This may imply that among men, Asian men are more susceptible to schizophrenia.

Top illnesses reported by RACE:

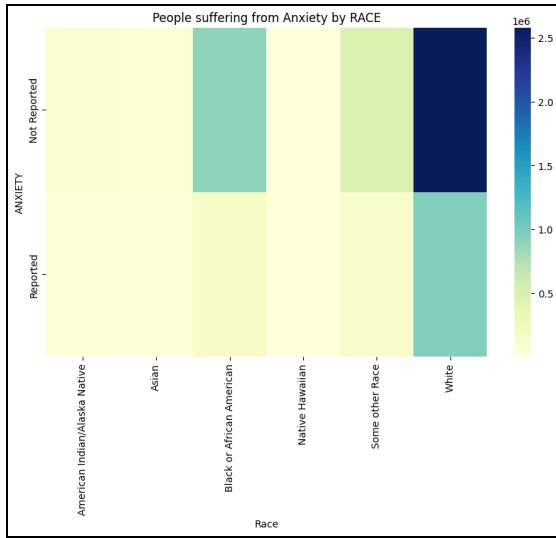


Fig. 10

Fig. 10 shows the diagnoses of anxiety among various racial groups in the country. It can be seen that it is highest among the white population, followed by the black / african american community.

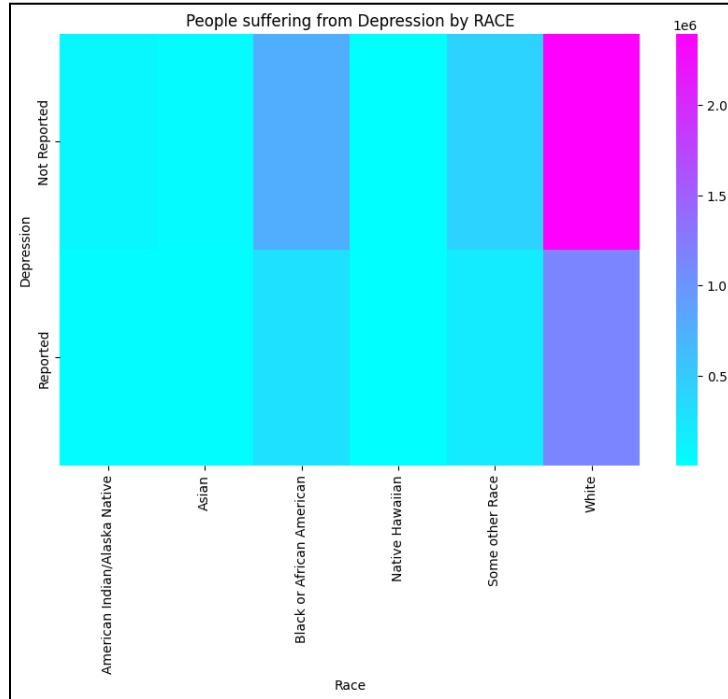


Fig. 11

Fig. 11 shows the heatmap of people suffering from depression by Race, what was reported or not. It can be seen high cases of depression reported in White, African American and some other races

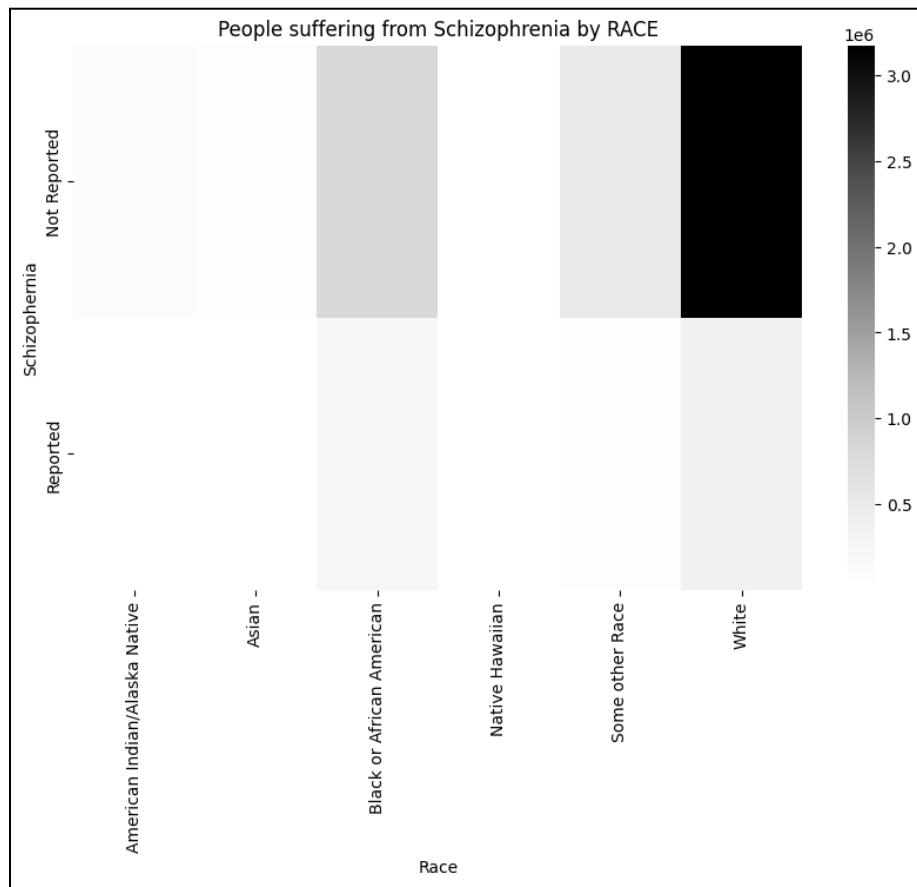


Fig. 12

Fig. 12 shows the heatmap of people suffering from Schizophrenia by Race, what was reported or not. High cases of depression reported in White, African American and some other races. Asian too have high rate of Schizophrenia given the portion of the population the data

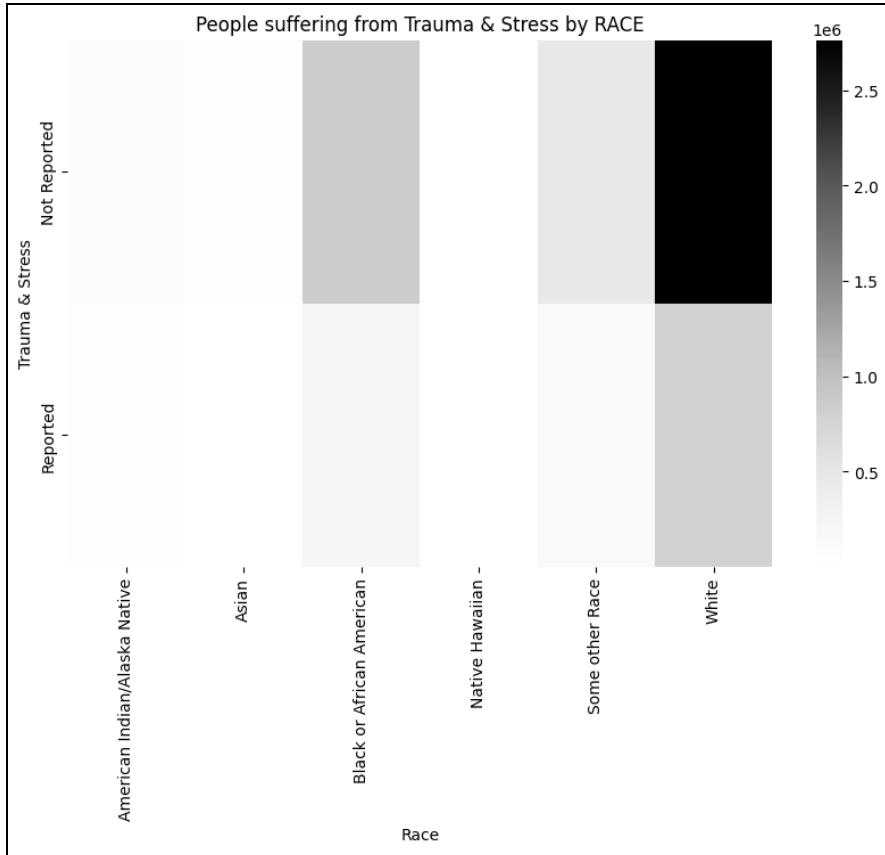


Fig. 13

Fig. 13 is a heatmap showing the distribution of trauma and stress related disorders among various racial groups. Again, the highest number of reported cases are for caucasians, followed by African Americans.

Reported Cases in the Asian Community for Various Mental Conditions:

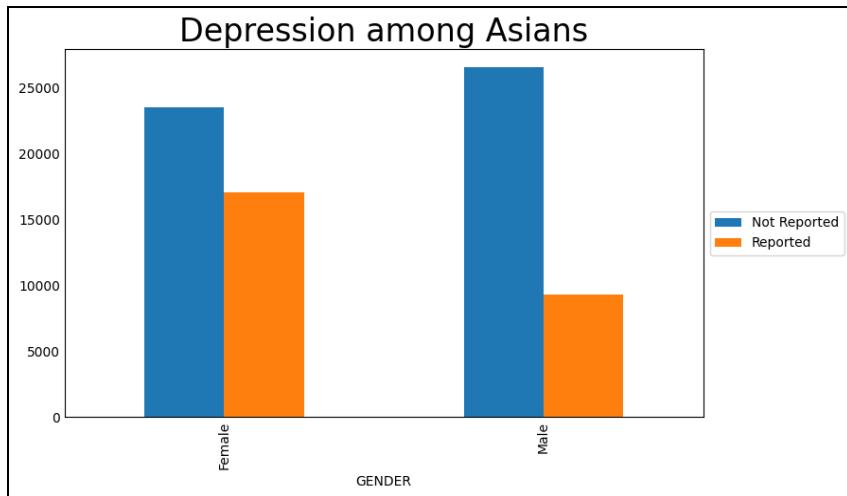


Fig. 14

Fig. 14 shows depression cases reported or not reported classified between female and male population. More not reported cases of depression can be seen amongst Asian men than female

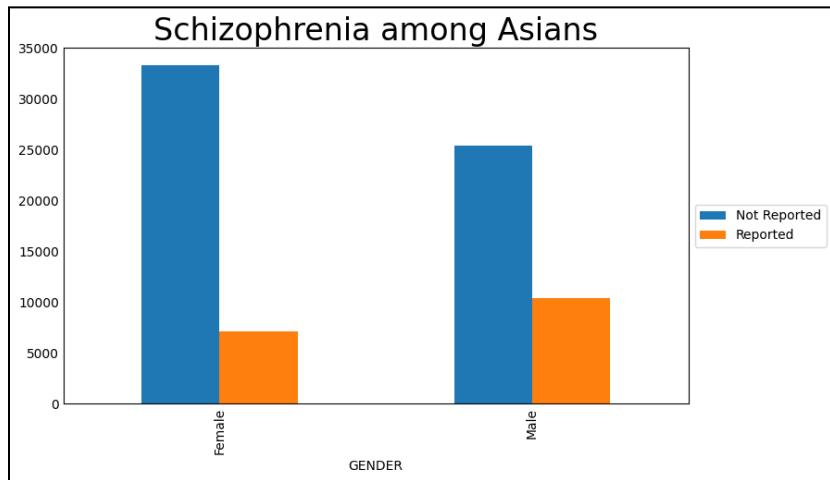


Fig. 15

Fig. 15 shows the variation in schizophrenia among Asian men and women. It is apparent that it is more prevalent among males, and this percentage is higher even when compared to schizophrenia prevalence among the entire population (all racial groups).

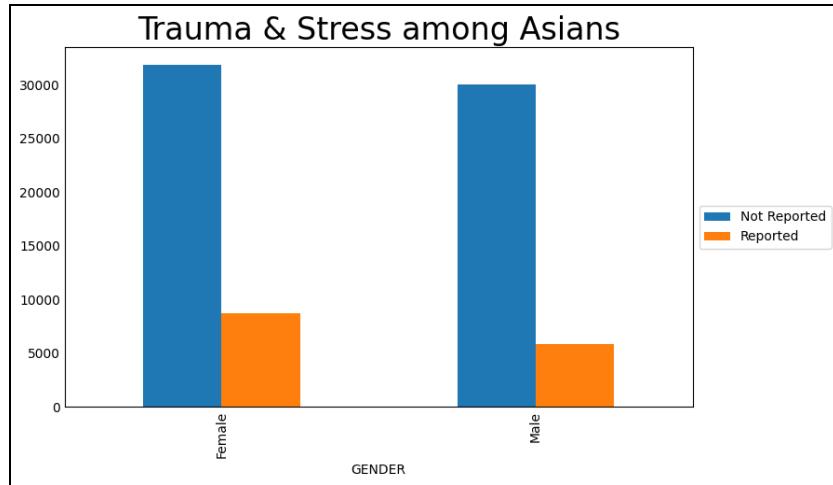


Fig. 16

Fig. 16 shows the occurrence of trauma and stress related disorders among men and women from the Asian community. The higher levels of stress and trauma may be attributed to the treatment of women in some Asian communities.

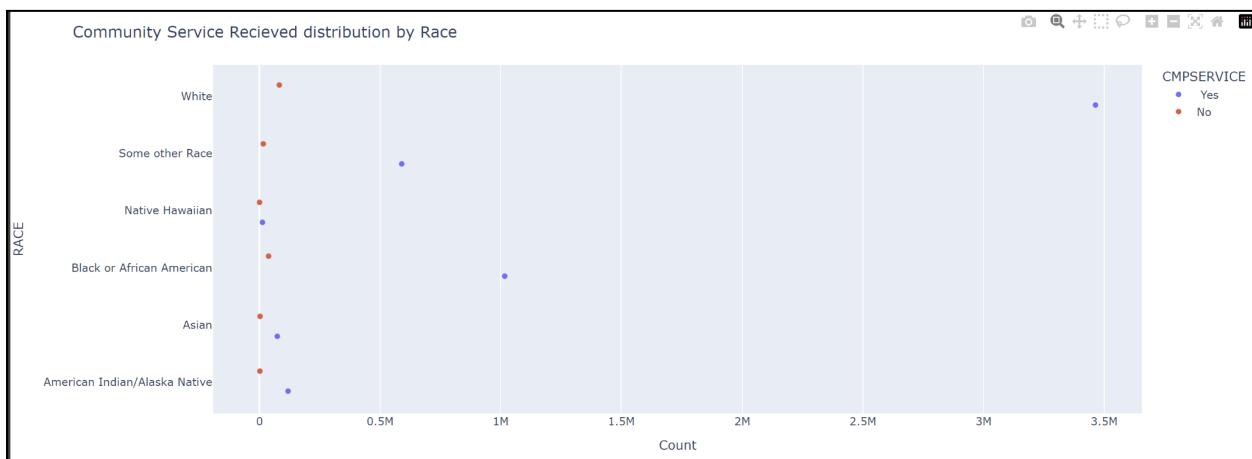


Fig. 17

People diagnosed with mental conditions are ready to receive community services if offered. The figures above determine that Asian people are reserved, and not everyone is ready to accept the condition they are going through.

Deliverable 2: Sentiment Analysis of Stories

Our main goal behind sentiment analysis was to know the overall emotional tone behind the stories. These stories are in text format collected through interviews or in dialogue from plays and dramas, primarily focusing on the Asian population in the United States. Our sponsor wants to spread awareness and learn more about how others perceive people with mental conditions by showcasing the “words” they have faced and analyzing the sentiment behind their stories.

What is Sentiment Analysis?

Sentiment Analysis is a powerful tool that uses Machine learning and Natural Language processing techniques to identify and extract information from social media and textual data. Text pre-processing, feature extraction, and classification are techniques used for analyzing and classifying the sentiment of a given text. Emotional tones are expressed in negative, neutral, and positive classes.

Tools Used in Analysis

Programming Language: Python

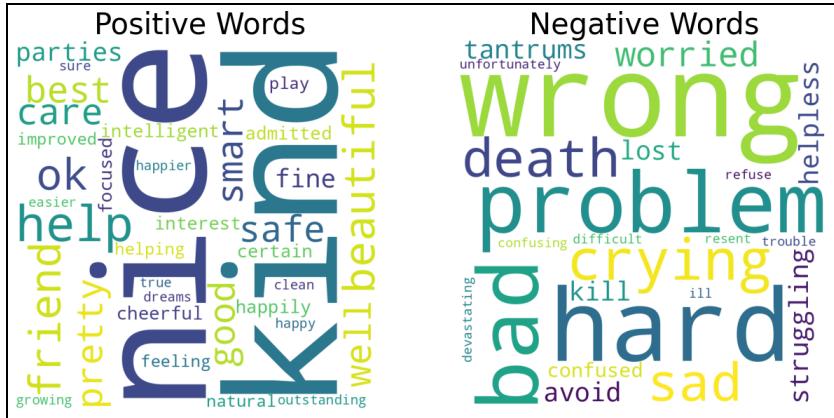
Libraries: NLTK, WordCloud, Matplotlib

Module: nltk.sentiment.vader

VADER is a pre-trained sentiment analysis model created to analyze text data, known for its ability to handle noisy, informal, and emotive language. The lexicon contains sentiment scores for many words and emoticons. The polarity scores method determines the sentiment score for each word. Positive words contain words with a compound score greater than 0.1 and negative words with less than -0.1. The lists are then used for generating word clouds for positive and negative words.

Word Cloud

Ashley



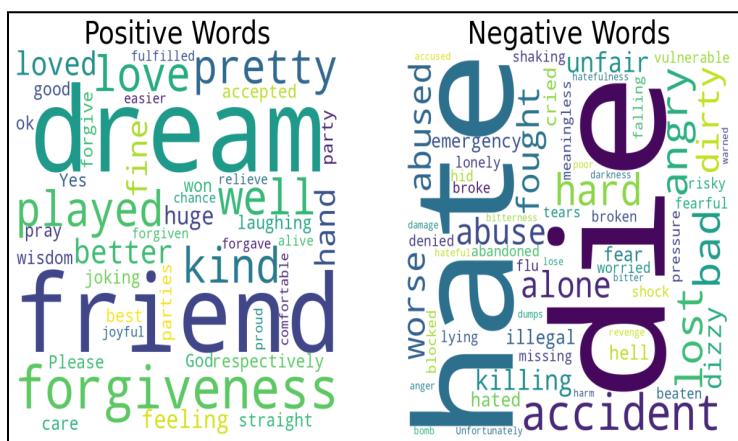
Polarity Scores

```
↳ {'neg': 0.079, 'neu': 0.807, 'pos': 0.114, 'compound': 0.9965}
```

Emotional Tone

The emotional tone of the text is positive.

Camille



Polarity Scores

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

Emotional Tone

The emotional tone of the text is negative.

Next Steps for the Sponsor

The sponsor can utilize the insights derived from data visualization and sentiment analysis to inform decision-making processes. The project opens avenues for ongoing collaboration and improvement in mental health services. The project can help the UIC Department of Disability and human development in several ways:

1. Enhanced Understanding: Through data visualization, the project will transform complex SAMHSA data into visual representations that are easy to comprehend. This will enable the department to gain a deeper understanding of the prevalence, patterns, and trends related to mental health issues in the target population.
2. Evidence-Based Decision-Making: By performing sentiment analysis on the textual data, the project will uncover the emotions, attitudes, and experiences expressed by individuals with disabilities and mental health issues. This information can guide decision-making processes, allowing the department to tailor their services and interventions to better address the specific needs of the community.
3. Improved Services and Support: The insights derived from the data visualizations and sentiment analysis can lead to actionable recommendations for enhancing mental health services and support. The UIC Department of Disability can utilize these recommendations to develop targeted interventions, allocate resources effectively, and provide better assistance to individuals with disabilities and mental health challenges.
4. Collaboration and Partnerships: The project's outcomes can serve as a valuable resource for collaborating with other stakeholders, such as researchers, community organizations, and policymakers. The data visualizations and analysis can foster dialogue, knowledge sharing, and collaborative efforts towards improving mental health outcomes and promoting inclusivity.

Conclusion:

Throughout the project, the team learned valuable lessons and overcome various challenges. The project's impact on promoting awareness and addressing the needs of individuals with disabilities and mental health issues was significant. The incorporation of business analytics principles and techniques proved valuable. The team expresses gratitude for the opportunity to work with the UIC Department of Disability and for the experiences gained from the capstone project. By analyzing SAMHSA data and presenting it through data visualizations and sentiment analysis, the project contributes to understanding mental health issues and supports evidence-based decision-making. The outcomes will facilitate informed discussions, promote awareness, and aid in the development of effective strategies and interventions to better serve individuals with disabilities and mental health issues. In conclusion, the project now focuses on SAMHSA Data Visualization and sentiment analysis, utilizing business analytics techniques to gain insights from the data. This revised approach will enable the UIC Department of Disability and other stakeholders to make informed decisions and improve mental health services based on data-driven evidence.