Mental Health Disorder Analysis

Project Objective:

The first milestone of the Mental Health Disorders Analysis Project has established a solid foundation for analyzing mental health trends and coping strategies across different countries. This phase focused on data collection, preprocessing, and exploratory data analysis (EDA). The data collection process integrated datasets containing information on the prevalence of various mental health disorders, including schizophrenia, anxiety, bipolar, and eating disorders, while documenting dataset sources, dimensions, and variable descriptions for transparency. Data preprocessing addressed missing values through techniques like mean imputation, scaled numerical features using Min-Max scaling for better comparability, and simplified column names for improved readability. The exploratory data analysis utilized descriptive statistics and visualizations to uncover initial insights. Bar charts and grouped bar plots were created to analyze trends in disorder prevalence over time. A notable visualization was a grouped bar chart showing mental health disorder trends in Afghanistan from 2000 to 2002, revealing that anxiety disorders consistently had the highest prevalence (~25-30%), followed by schizophrenia (~10-12%), while bipolar and eating disorders showed lower rates. This analysis also identified "talking to friends/family" as a commonly used coping strategy for managing anxiety. The comprehensive approach to data handling and initial analysis has set the stage for deeper exploration and more advanced analytical techniques in the subsequent phases of the project, including the potential for predictive modeling of future mental health trends and the development of actionable recommendations for policymakers and healthcare providers.

Data Used:

The analysis utilized a subset of data from the "Mental Health" dataset available on Kaggle (https://www.kaggle.com/datasets/amirhoseinmousavian/mental-health). While the full dataset contains 50 files, this project focused on 14 specific files relevant to the research objectives. These selected files provide comprehensive information on various mental health disorders, their prevalence across different countries and years, and common coping strategies employed by individuals. The data encompasses key mental health indicators, allowing for a nuanced exploration of trends and patterns in mental health across diverse populations and time periods.

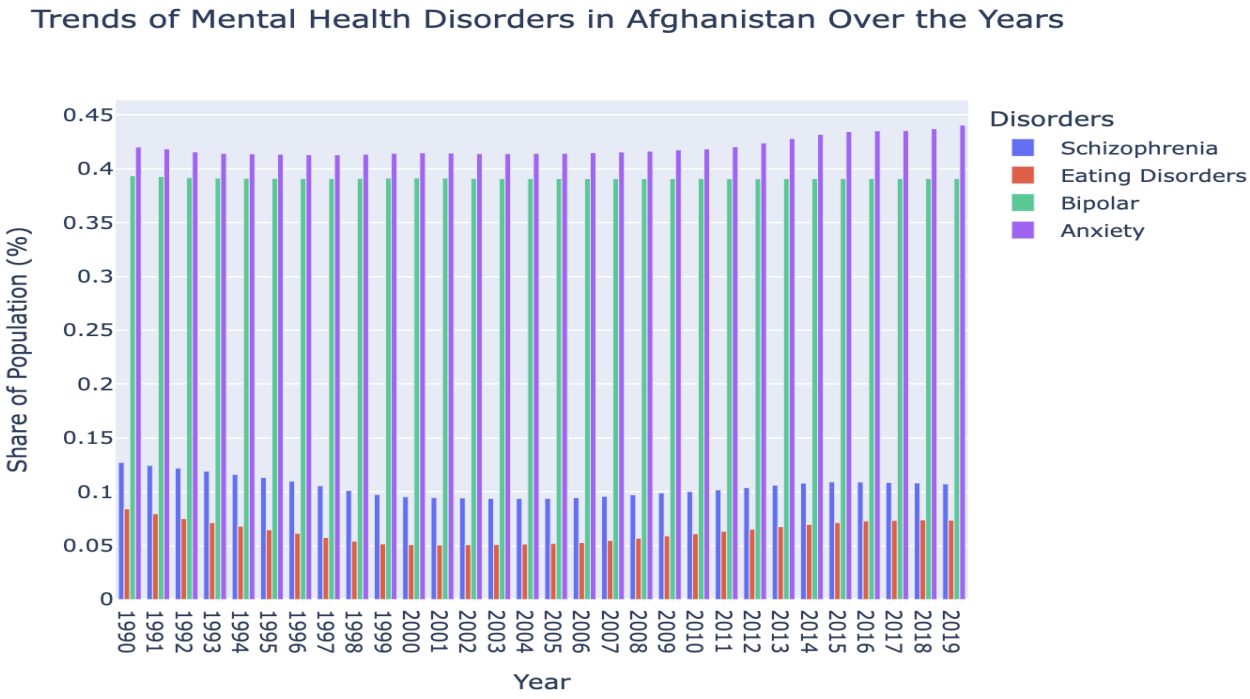
Tech Stack:

The project leveraged a robust set of Python libraries and tools for data manipulation, analysis, and visualization:

1. Pandas: For efficient data manipulation and analysis of structured data.
2. NumPy: To support large, multi-dimensional arrays and matrices, along with a collection of mathematical functions.
3. Matplotlib: For creating static, animated, and interactive visualizations in Python.
4. Seaborn: To produce statistical graphics based on matplotlib, enhancing the visual appeal of plots.
5. Sketch: For additional data visualization capabilities.
6. OS: To interact with the operating system, particularly for file and directory operations.
7. Plotly Express: To create interactive plots and charts with minimal code.
8. Plotly Graph Objects: For more customizable interactive visualizations.

This tech stack provided a comprehensive toolkit for data processing, statistical analysis, and the creation of both static and interactive visualizations, enabling a thorough exploration of the mental health dataset.

EDA Report and Key Insights:
The types of disorder trends over the 1999-2019 time period over Afghanistan. The same analysis can be done for different countries of the data.

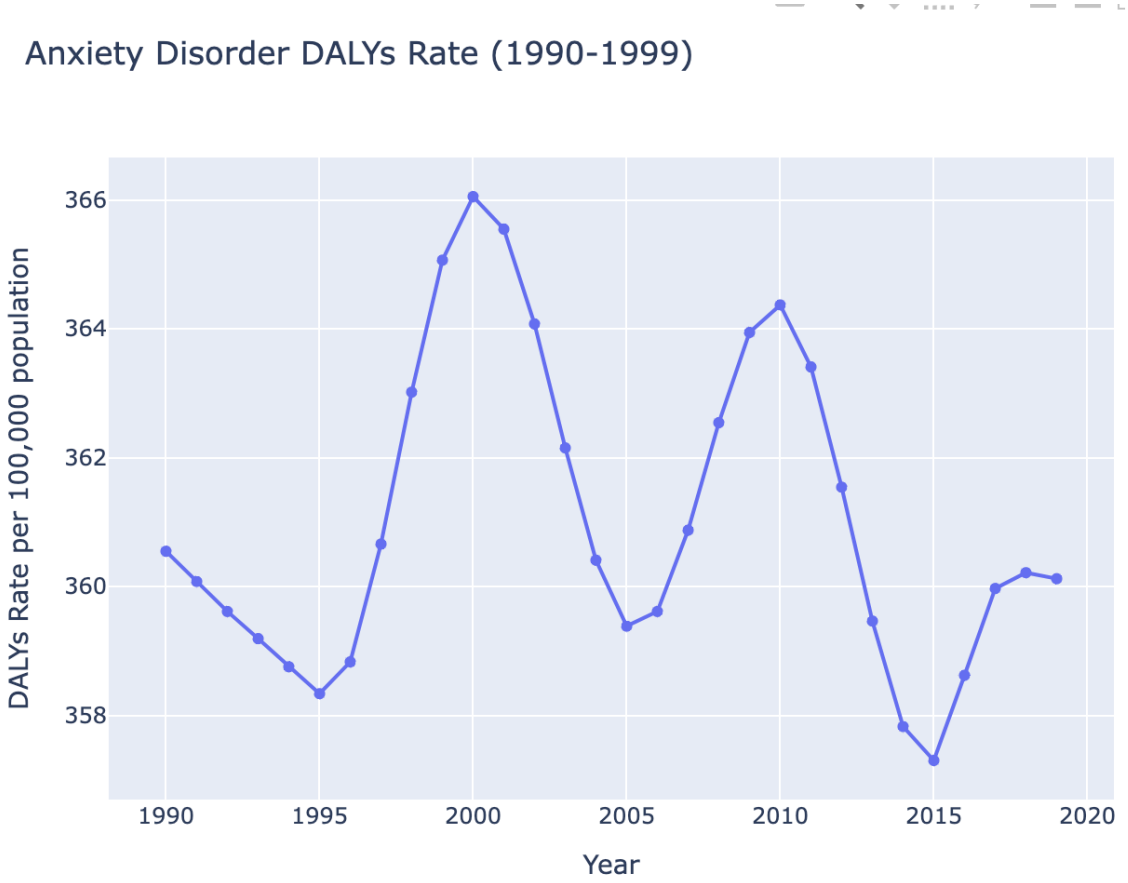


Observations from the Graph

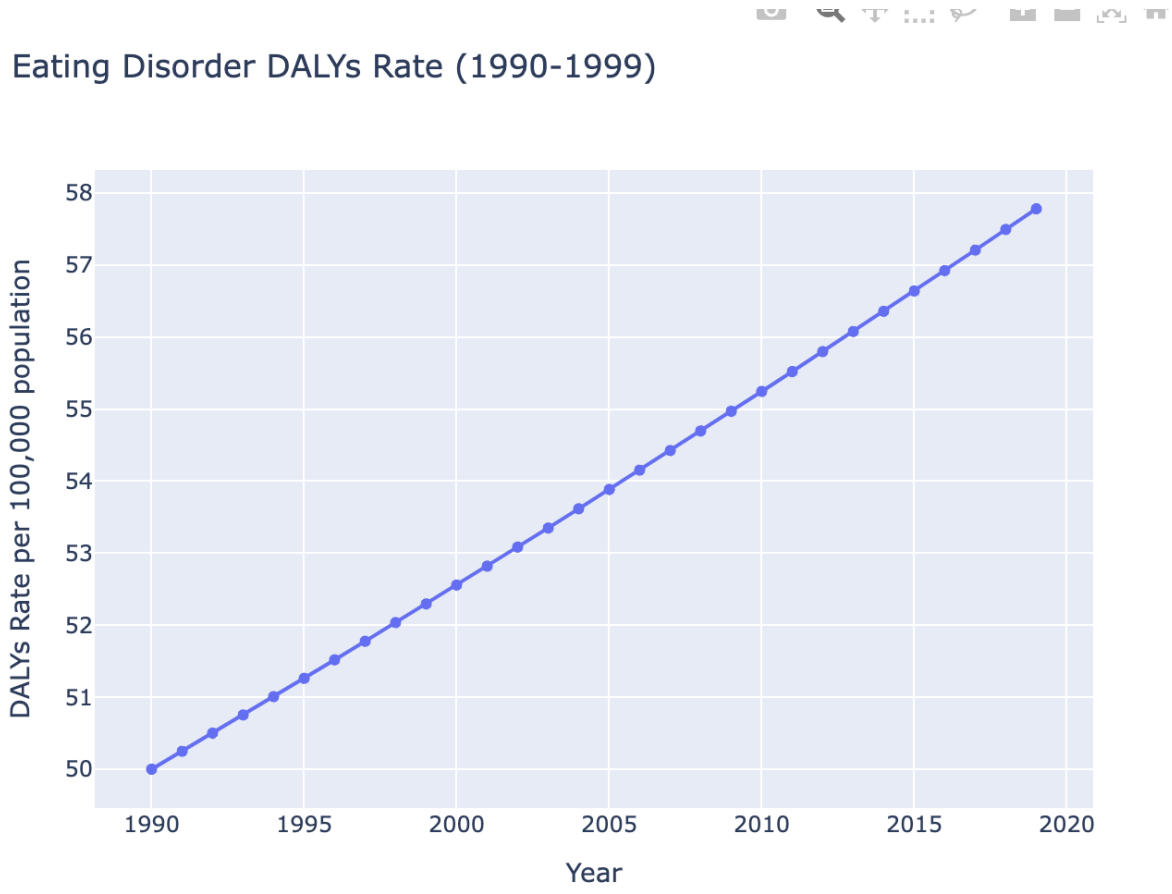
1. Anxiety Disorders Dominate:
    - Anxiety disorders consistently have the highest prevalence among the four mental health disorders in Afghanistan, with a share of the population nearing 0.4% across all years (1990–2019).

    - This indicates that anxiety is a significant mental health concern in the region.
2. Schizophrenia as the Second Most Prevalent:
    - Schizophrenia disorders maintain a steady prevalence of around 0.1%, making it the second most common disorder after anxiety.
3. Bipolar Disorders and Eating Disorders:
    - Bipolar disorders and eating disorders have much lower prevalence rates compared to anxiety and schizophrenia.
    - Bipolar disorders hover around 0.05%, while eating disorders remain below 0.03% throughout the observed years.
4. Stable Trends Over Time:
    - The prevalence rates for all four disorders remain relatively stable from 1990 to 2019, with no significant increases or decreases observed.
    - This suggests that mental health trends in Afghanistan have not undergone major shifts during this period.
5. Clear Visual Hierarchy:
    - The graph highlights a clear hierarchy of disorder prevalence: anxiety > schizophrenia > bipolar > eating disorders.
6. Implications for Mental Health Interventions:
    - The high prevalence of anxiety disorders suggests a need for targeted interventions and resources to address anxiety-related issues.
    - Schizophrenia also requires attention, while bipolar and eating disorders, though less prevalent, should not be overlooked.

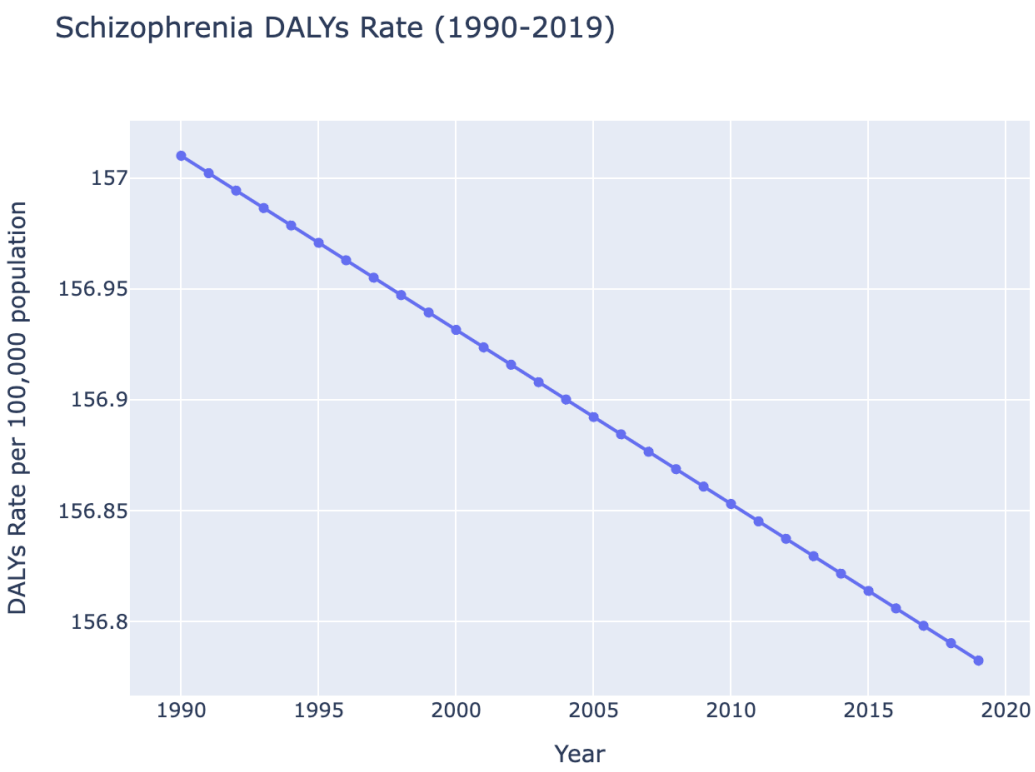
Dalys Rate of Disorders :



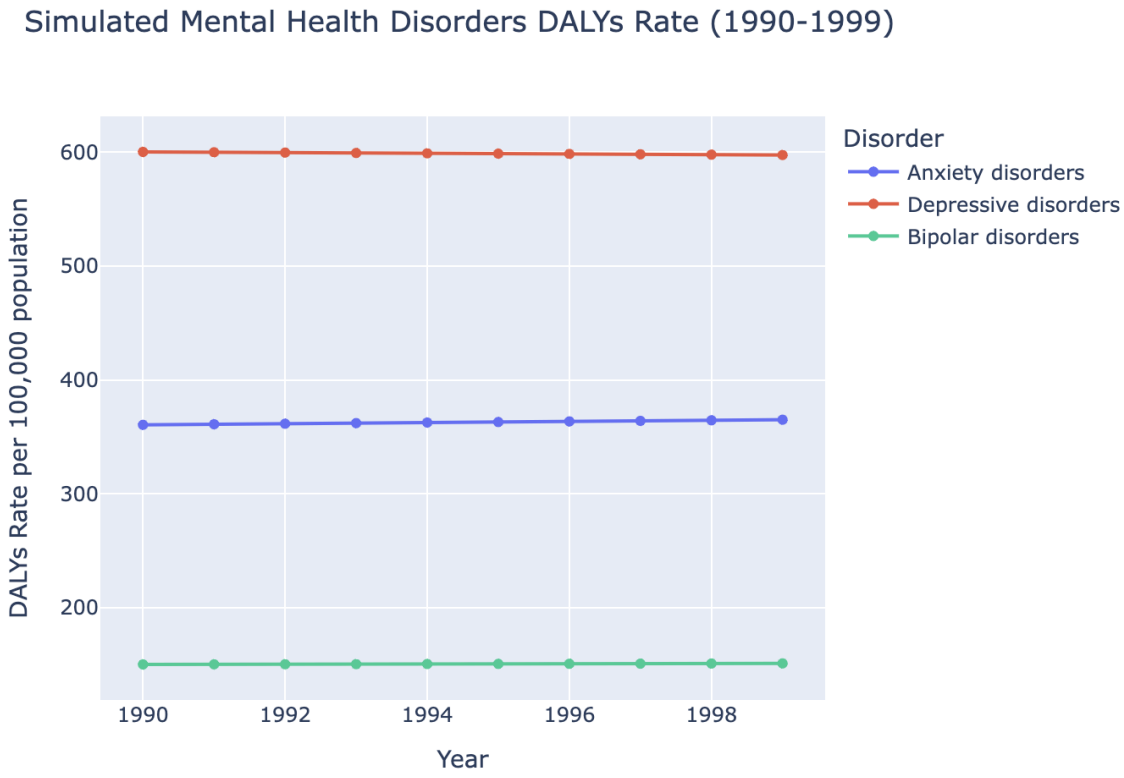
- Anxiety Disorder DALYs Rate (1990-2019): Fluctuates, starting around 360.5, peaking at 366 in the early 2000s, dipping to about 358 in the mid-2010s, and ending near 360, showing variability but relative stability.



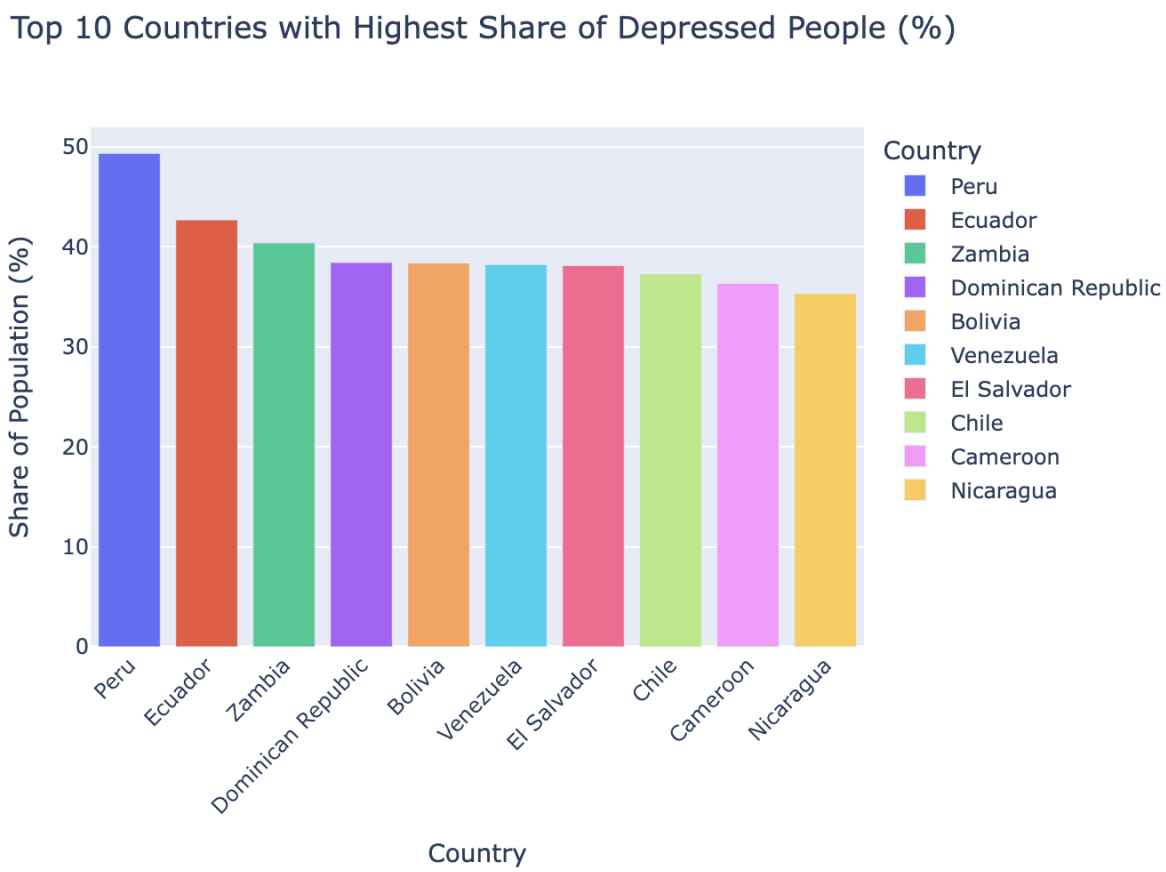
- Anxiety Disorder DALYs Rate (1990-2019): Fluctuates, starting around 360.5, peaking at 366 in the early 2000s, dipping to about 358 in the mid-2010s, and ending near 360, showing variability but relative stability.



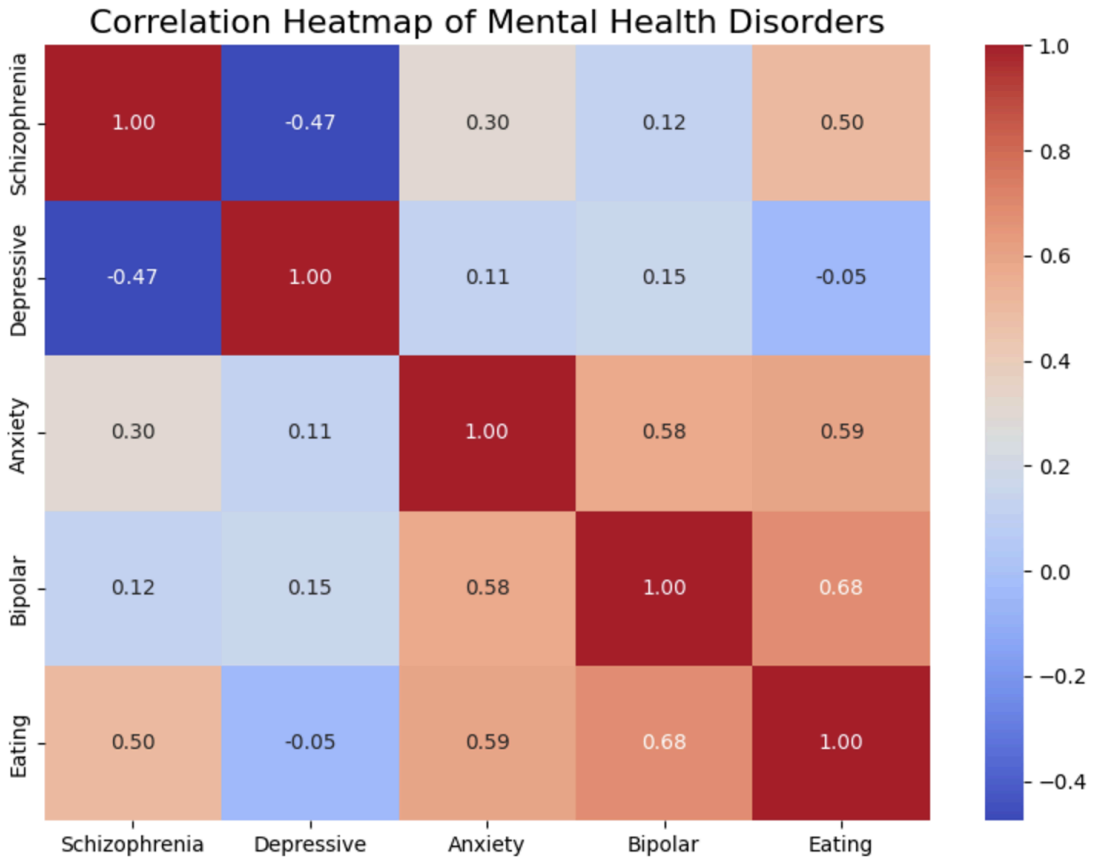
The graph illustrates the DALYs (Disability-Adjusted Life Years) rate for schizophrenia from 1990 to 2019. It shows a steady and gradual decline over the years, with a slight reduction from approximately 157 per 100,000 population in 1990 to around 156.8 in 2019. The overall change is minimal, indicating that the burden of schizophrenia on population health has remained relatively stable over the three decades. This trend suggests that while there may have been some improvements in managing schizophrenia, they are not significant enough to drastically reduce its impact.



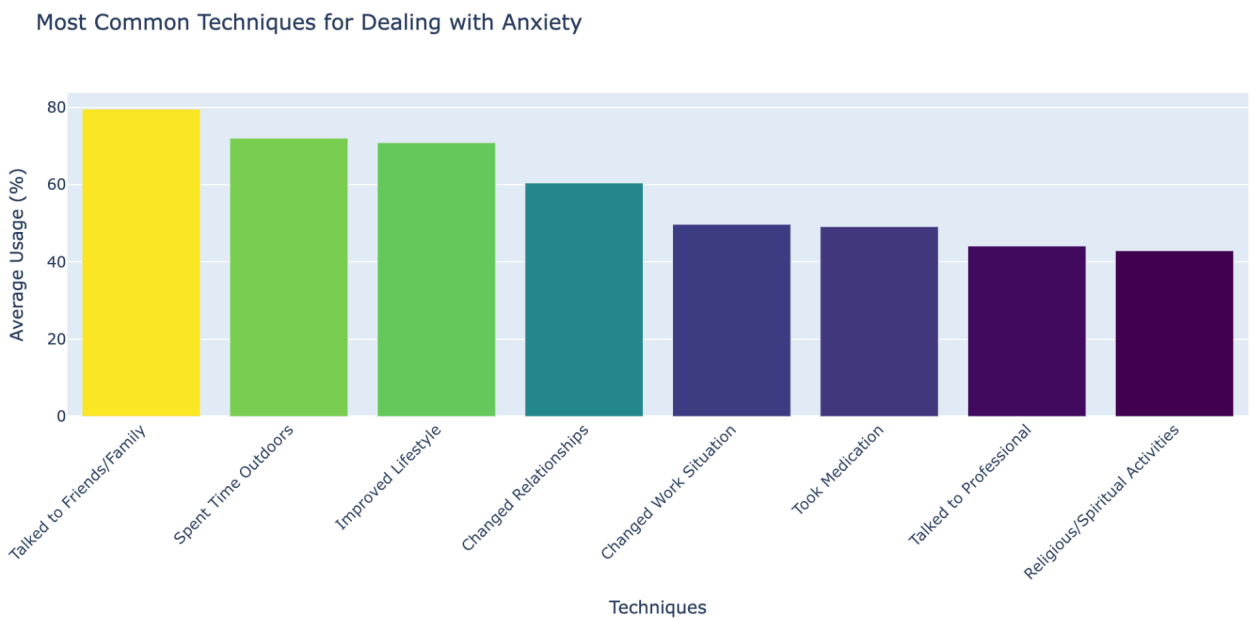
The graph compares the DALYs rates for anxiety, depressive, and bipolar disorders over the decade from 1990 to 1999. The rates remain constant throughout the period, with depressive disorders having the highest burden at around 600 per 100,000 population, followed by anxiety disorders at approximately 360 per 100,000 population, and bipolar disorders at about 150 per 100,000 population. The lack of variation indicates no significant changes in the prevalence or management of these disorders during this time. Depressive disorders stand out as having the greatest impact on public health among the three, highlighting the need for targeted interventions. Anxiety disorders also represent a substantial burden, while bipolar disorders have a comparatively lower impact. Both graphs emphasize the importance of sustained efforts to address mental health challenges globally.



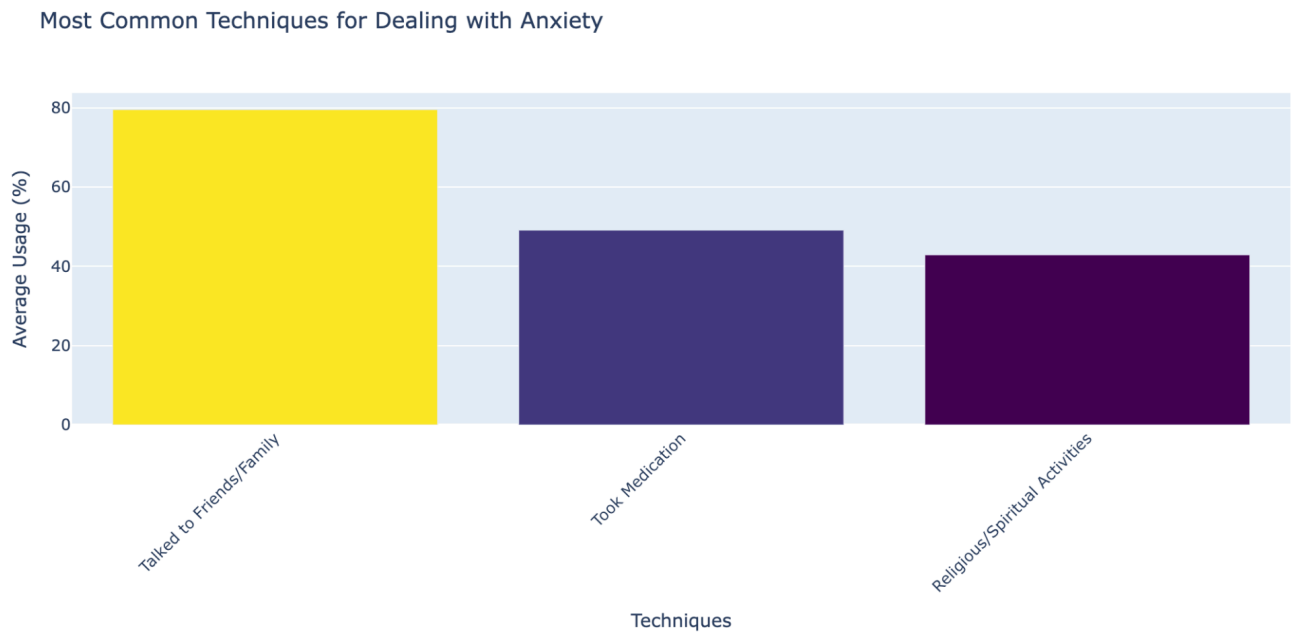
The chart highlights a notable prevalence of depression in Latin American countries such as Peru, Ecuador, Bolivia, and Venezuela, suggesting potential regional or cultural factors contributing to these high rates. African countries like Zambia and Cameroon also appear in the top 10, indicating that depression is a global concern across diverse regions. This visualization underscores the urgent need for targeted mental health interventions and resources in these countries to address the widespread impact of depression.



The heatmap illustrates the correlation between various mental health disorders, with correlation values ranging from -1 (perfect negative correlation) to 1 (perfect positive correlation). Schizophrenia and depressive disorders exhibit a moderately negative correlation (-0.47), indicating distinct patterns or factors influencing these conditions. Eating disorders show strong positive correlations with both bipolar disorders (0.68) and anxiety disorders (0.59), suggesting possible overlaps in their prevalence or shared contributing factors. Similarly, bipolar disorders and anxiety disorders have a moderate positive correlation (0.58), highlighting potential comorbidities between these conditions. On the other hand, depressive disorders display weak correlations with most other disorders, such as anxiety (0.11) and bipolar (0.15), indicating limited overlaps in their prevalence patterns. The diagonal values of 1.00 represent the perfect self-correlation of each disorder with itself. Overall, the heatmap reveals significant associations between certain mental health disorders while also highlighting distinct patterns for others, providing insights into potential shared risk factors and unique characteristics of these conditions.



This bar chart, titled "Most Common Techniques for Dealing with Anxiety", illustrates the average usage percentage of various coping strategies. The most widely used technique is "Talked to Friends/Family", with an average usage of nearly 80%, indicating the importance of social support in managing anxiety. The second most common method is "Spent Time Outdoors", with an average usage of around 70%, highlighting the therapeutic benefits of nature and outdoor activities. "Improved Lifestyle" and "Changed Relationships" follow closely, both with usage rates exceeding 60%, reflecting the significance of adopting healthier habits and adjusting personal dynamics.



The bar chart titled "Most Common Techniques for Dealing with Anxiety" highlights the average usage percentages of three coping strategies. The most commonly used technique is "Talked to Friends/Family", with an average usage of approximately 80%, emphasizing the importance of social support in managing anxiety. "Took Medication" is the second most utilized strategy, with an average usage of around 40%, indicating that medical intervention is a significant but less common approach. Lastly, "Religious/Spiritual Activities" is the least used technique, with an average usage of about 30%, suggesting that spiritual practices are less frequently relied upon for anxiety management. The chart underscores that individuals primarily depend on social connections for dealing with anxiety, while medical and spiritual approaches are secondary choices.

Project Timeline:

Data Collection:

- Timeline: February 5, 2025 - February 11, 2025 (7 days)
- Tasks:
    - Identified and acquired the dataset from a reliable source (Kaggle).
    - Verified dataset accessibility and ensured compliance with licensing or usage restrictions.
    - Documented the dataset source, dimensions, and variable descriptions.

Data Preprocessing:

- Timeline: February 12, 2025 - February 18, 2025 (7 days)
- Tasks:
    - Handled missing data using techniques such as mean imputation.

    - Addressed outliers using statistical methods like interquartile range (IQR).
    - Normalized features using Min-Max scaling for comparability.
    - Renamed long column names into simpler terms for better readability.

Exploratory Data Analysis (EDA):

- Timeline: February 19, 2025 - February 23, 2025 (5 days)
- Tasks:
    - Conducted descriptive statistics to summarize key variables.
    - Created visualizations such as bar charts and heatmaps to identify patterns and trends.
    - Analyzed correlations between mental health disorders and identified potential issues such as multicollinearity or skewed distributions.