Mental Health Analysis in Work Place

Name:	Shweta Kulkarni
Project Title:	Mental Health Analysis in Work Place GitHub Link: Project Link
Dataset Source:	Dataset

1. Project Summary

This project explores mental health trends in the technology industry using a real-world dataset sourced from Kaggle. The dataset captures various attributes of individuals including age, gender, employment details, and their attitudes toward mental health support. Through extensive data preprocessing, including cleaning outliers and normalizing inconsistent gender entries, I prepared the data for insightful visualizations and statistical analysis. This analysis aimed to uncover age and gender-specific patterns, organizational support factors, and correlations between workplace policies and the likelihood of seeking mental health help.

The findings reveal that younger professionals, especially those aged 25–35, are more likely to seek mental health support, with notable gender differences in willingness to seek help. Moreover, organizational factors such as the presence of wellness programs, mental health benefits, and supportive leave policies show a positive correlation with employees seeking help. A heatmap of correlation values further highlights how structural support within companies plays a crucial role in mental health outcomes. This project not only uncovers hidden patterns but also emphasizes the importance of proactive mental health initiatives in the tech workplace.

2. Introduction

The objective of this study is to understand the factors that influence mental health support-seeking behavior among professionals in the tech industry. By examining demographic details and workplace-related parameters, I aim to uncover trends that reveal how age, gender, and organizational support impact an individual's decision to seek mental health help or treatment.

This analysis is particularly important due to the increasing mental health challenges faced by individuals in high-pressure, fast-paced environments like the tech sector. Long work hours, remote isolation, job insecurity, and burnout contribute to rising mental health concerns, yet stigma and lack of support often prevent professionals from seeking help. Identifying which workplace factors positively or negatively affect this behavior can inform better policies and promote a healthier work culture.

The dataset used in this project was sourced from Kaggle, titled "Mental Health in Tech Survey", which contains anonymized survey responses from tech employees across different countries. It includes features such as age, gender, company size, remote work status, and perceptions around mental health in the workplace.

3. Data Preprocessing

Before performing any analysis, I cleaned and prepared the dataset to ensure accuracy and consistency. First, I removed unrealistic age entries — specifically, respondents below 18 and above 80 years old — as these were likely errors or outliers that could skew the results.

Next, I standardized gender entries. Since participants had entered gender in various formats (e.g., "Male", "male", "M", "man"), I grouped all male-identifying responses under "M" and did the same for female and other gender identities. This helped simplify gender-based analysis and visualization.

Lastly, I handled missing values. For columns with a significant number of missing values that could not be reasonably filled (e.g., open-text fields like comments), I dropped them. For other missing fields where possible, I used logical imputation or removed only the affected rows to maintain dataset integrity without losing valuable information.

4. Exploratory Data Analysis

Before performing any analysis, I cleaned and prepared the dataset to ensure accuracy and consistency. First, I removed unrealistic age entries — specifically, respondents below 18 and above 80 years old — as these were likely errors or outliers that could skew the results.

Next, I standardized gender entries. Since participants had entered gender in various formats (e.g., "Male", "male", "M", "man"), I grouped all male-identifying responses under "M" and did the same for female and other gender identities. This helped simplify gender-based analysis and visualization.

Lastly, I handled missing values. For columns with a significant number of missing values that could not be reasonably filled (e.g., open-text fields like comments), I dropped them. For other missing fields where possible, I used logical imputation or removed only the affected rows to maintain dataset integrity without losing valuable information.

4.1. Age-wise Trends

- Number of people per age vs those who sought help.

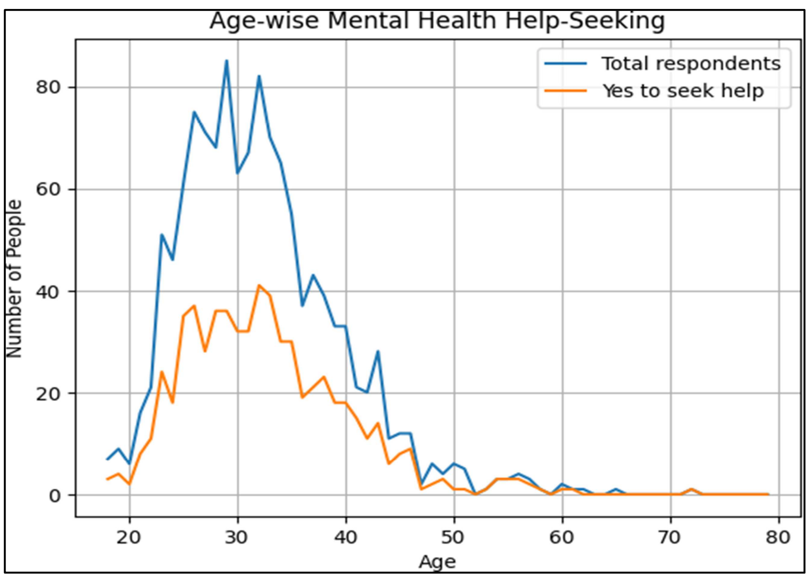

Figure 1: Age-wise Mental Health Help-Seeking

The graphs indicate that individuals between the ages of 25 to 35 form the largest segment of survey respondents, with a noticeable peak around age 30. This age group not only shows the highest participa- tion but also the highest absolute number of individuals reporting that they have sought mental health support. This suggests that people in this age range are both more engaged in mental health discussions and more proactive in seeking help when needed.

This trend could be attributed to various factors, such as increased mental health awareness, digital liter- acy, and the prevalence of workplace stress in early-to-mid career stages. It reflects a growing willingness among younger professionals to acknowledge and address mental health concerns.

Beyond the age of 40, both the total number of respondents and those seeking help decline significantly. This drop-off may indicate lower engagement from older age groups, possibly due to reduced survey participation, greater stigma, or generational differences in attitudes toward mental health. It emphasizes the need for targeted outreach and awareness campaigns for older demographics.

- Percentage of people seeking help by age

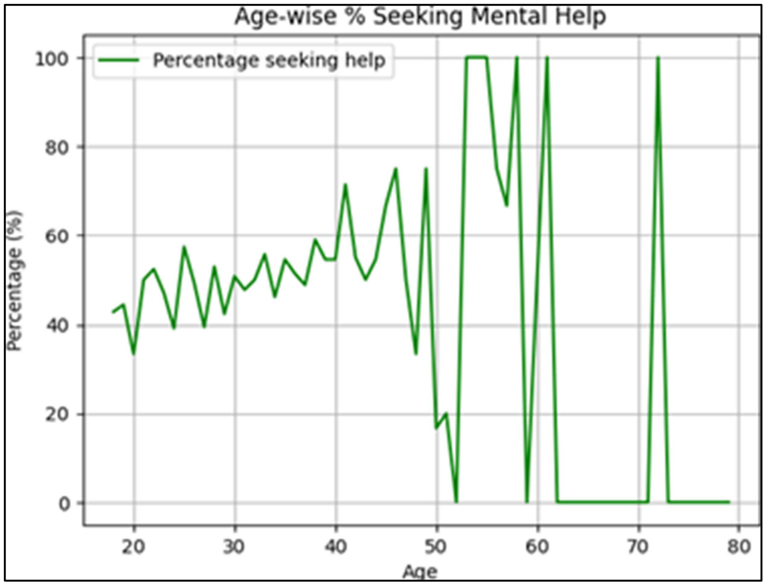

Figure 2: Age-wise % Seeking Mental Help

In the second chart, which visualizes the percentage of individuals seeking mental health help across different ages, a relatively stable trend can be observed from ages 18 to 45. Within this range, the help- seeking rate consistently falls between 40% and 60%, indicating that a significant portion of individuals in early to mid-adulthood are open to addressing mental health concerns.

Interestingly, the data shows a gradual increase in help-seeking behavior from the mid-30s to mid-40s. This upward trend may reflect a growing sense of self-awareness, greater emotional maturity, or increased exposure to mental health education and support systems within that age group. Individuals in their late 30s and early 40s are often in demanding stages of life — balancing careers, family responsibilities, and long-term goals — which may elevate the need for mental health support.

After the age of 45, however, the graph begins to fluctuate more abruptly. This volatility is likely due to a smaller number of survey respondents in these older age groups, making the percentages less statistically reliable. Consequently, while the apparent rise or fall in help-seeking beyond 45 could suggest behavioral shifts, it is more likely a result of limited data and should be interpreted with caution.

4.2. Gender-wise Trends

- Total number of males vs females vs help-seeking count

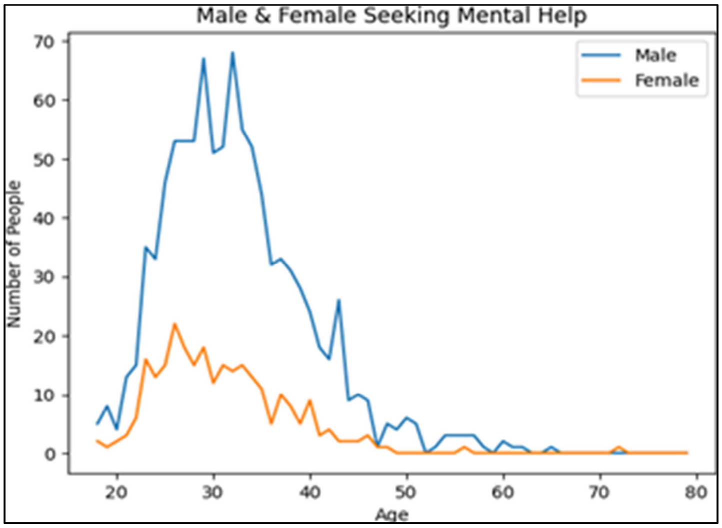

Figure 3: Male Female Seeking Mental Help

This graph presents the raw count of male and female respondents who reported seeking mental health help across different age groups. The highest number of help-seeking individuals is observed in the age range of 25 to 35, with a clear peak around age 30. Males significantly outnumber females in absolute numbers, which could reflect the overall male dominance in the tech workforce or dataset sample.

After age 40, there is a noticeable decline in the number of people from both genders reporting mental health help-seeking behavior, suggesting either reduced engagement or possible underrepresentation of older individuals in the survey data.

- Total number of males vs females vs help-seeking count

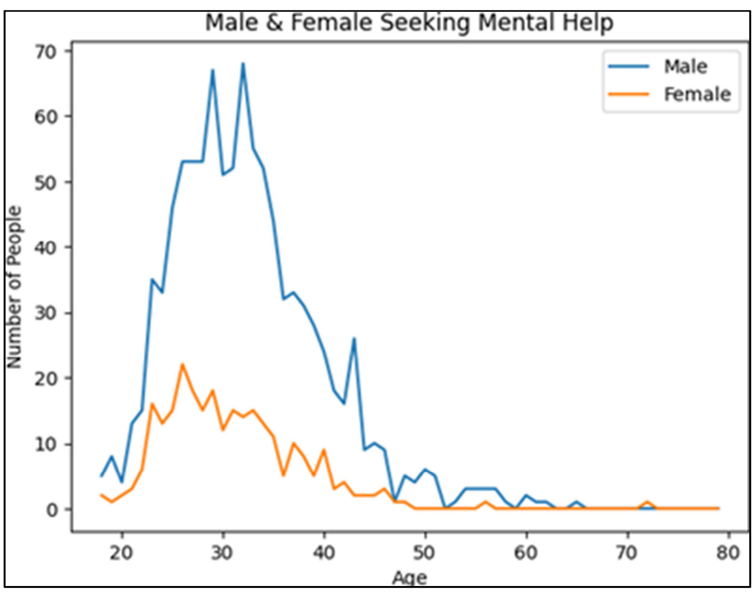

Figure 4: Percentage of Male Female Seeking Mental Help

This graph displays the percentage of males and females within each age group who reported seeking mental health support. Contrary to common assumptions, the trend shows that male respondents exhibit a higher percentage of help-seeking behavior than females across most age ranges, especially between the ages of 25 and 40. The percentage for males peaks sharply around the early 30s, reaching over 30%, while females peak below 15% in the same age range. After the age of 40, both lines decline significantly, which may be due to reduced participation from older individuals, resulting in more erratic and less reliable data. This unexpected finding challenges conventional beliefs about gendered mental health behavior and could indicate that, at least in the tech industry, men might be more open to seeking help than previously thought — or possibly that female respondents are underrepresented or less willing to report.

4.3. Family History Trends

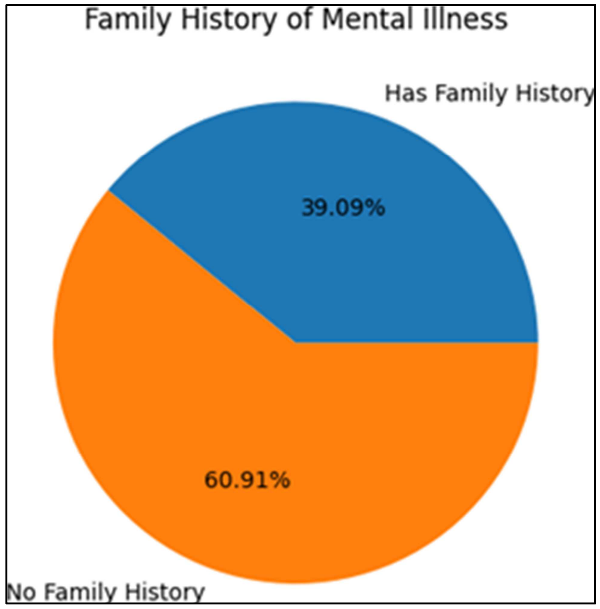

Figure 5: Family History of Mental Illness

The first chart shows that approximately 39.09% of the respondents reported having a family history of mental illness, while 60.91% did not. This indicates that a significant portion of individuals in the tech industry are aware of mental health issues within their families, suggesting a potential genetic or environmental predisposition to mental health concerns among this workforce.

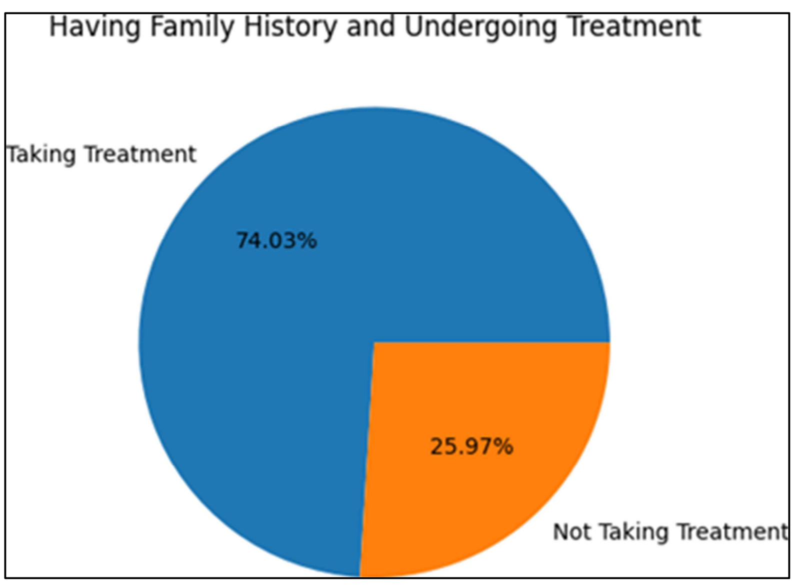

Figure 6: Having Family History and Undergoing Treatment

The second chart reveals that among individuals who do have a family history, 74.03% are undergoing treatment, while only 25.97% are not. This suggests that having a family history may positively influence one's likelihood of seeking help or treatment. Awareness, past exposure, and possibly encouragement from family members familiar with mental illness could be contributing factors behind the higher treatment rate in this group.

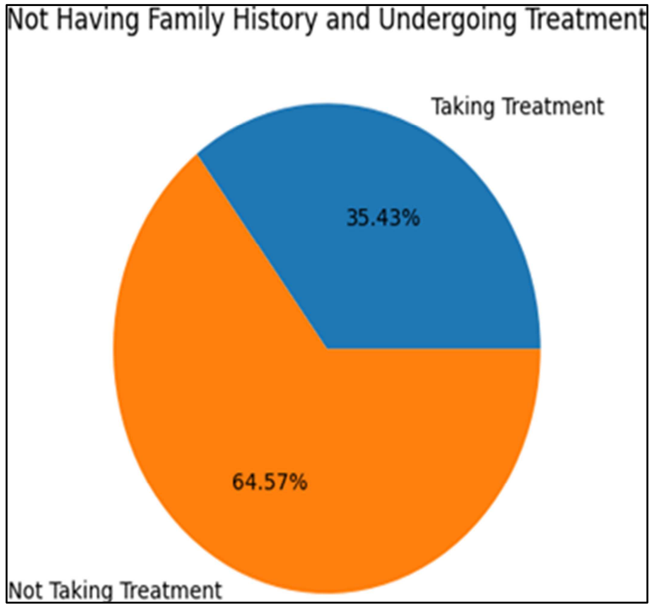

Figure 7: Not Having Family History and Undergoing Treatment

The third chart shows that among individuals without any known family history, only 35.43% are taking treatment, while a larger share, 64.57%, are not. This highlights a potential lack of perceived need or awareness among this group. Without a familial reference, these individuals may be more likely to dismiss symptoms, underestimate their severity, or be less open to seeking professional help. Overall, these charts strongly suggest that family history plays a vital role in mental health awareness and proactive treatment- seeking behavior. Individuals with a known history are more than twice as likely to undergo treatment compared to those without. This emphasizes the importance of mental health education and destigma- tization, particularly for those who may not have prior exposure to mental illness in their immediate circles.

4.4. Correlation Heatmap

- This heatmap visualizes the pairwise Pearson correlation coefficients between five variables related to workplace mental health support and behavior: treatment, benefits, care options, wellness program, and seek help. The values range from -1 (perfect negative correlation) to +1 (perfect positive correlation), with 0 indicating no correlation.

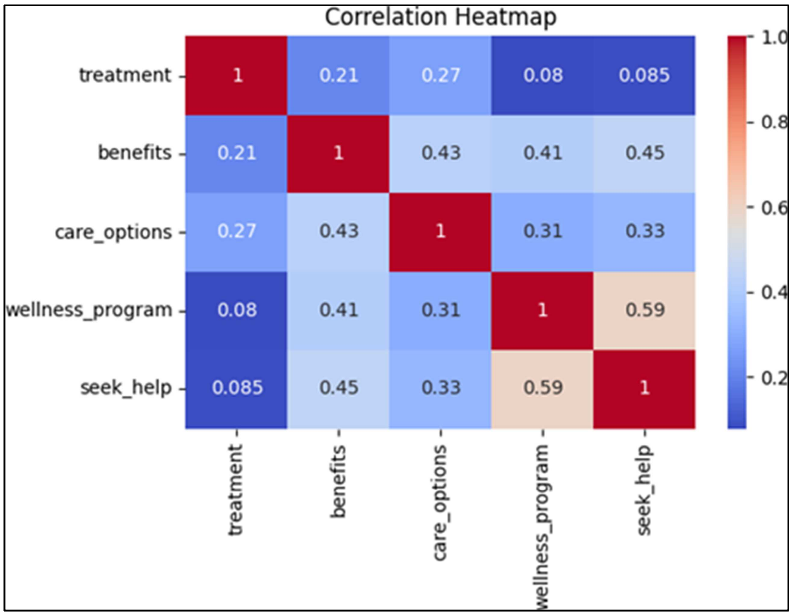

Figure 8: Correlation Heatmap

The heatmap illustrates the relationships between five key variables related to workplace mental health: treatment, benefits, care options, wellness programs, and seek help. Among these, the strongest correlation is observed between seek help and wellness program with a value of 0.59. This indicates that the presence of wellness programs in the workplace is positively associated with employees' likelihood of seeking mental health support. It suggests that when organizations actively promote wellness initiatives, it can significantly influence employees to come forward and seek help for their mental well-being.

Additionally, moderate correlations are seen between benefits and other variables. For example, benefits correlates with seek help (0.45), care options (0.43), and wellness program (0.41). This means that employees who are aware of or receive mental health-related benefits are also more likely to know about available care options, participate in wellness programs, and ultimately seek mental health support. These moderate correlations reflect the importance of accessible and well-communicated mental health benefits in driving awareness and encouraging action.

There is also a mild correlation between care options and seek help (0.33), suggesting that simply having options for care available slightly increases the chances of individuals seeking support. However, the relatively low strength of this correlation indicates that awareness of care options alone may not be sufficient to significantly boost help-seeking behavior. It likely needs to be accompanied by cultural shifts and proactive encouragement from employers.

On the other hand, the variable treatment shows weak correlations with all other factors — such as seek help (0.085), wellness program (0.08), and benefits (0.21). This is particularly revealing, as it implies that being in treatment for mental health issues is not strongly linked to the presence of workplace resources or encouragement. This disconnect points toward deeper challenges such as personal stigma, fear of professional consequences, or cultural norms, which may deter individuals from pursuing actual treatment even when support structures exist.

In conclusion, while workplace programs, benefits, and awareness efforts do play a role in encouraging help-seeking behavior, they do not necessarily translate into individuals entering treatment. To address this, organizations must move beyond basic provisions and foster a work culture where mental health is openly discussed, support is normalized, and seeking treatment is free from fear or judgment. Only then can the gap between awareness and action be effectively bridged.

5. Predictive Modeling: Estimating Treatment Behavior

In this project, a logistic regression model was employed to predict whether individuals have sought mental health treatment based on various demographic and workplace-related features. The dataset included categorical variables such as gender, family history of mental illness, availability of benefits, care options, wellness programs, and the target variable treatment. These variables were encoded numerically to make the data suitable for model training. For example, gender was encoded as 0 for Male, 1 for Female, and 2 for Other, while binary responses such as "Yes" and "No" were encoded as 1 and 0 respectively. Multi-category responses like "Don't know" or "Not sure" were also assigned unique numerical values.

After preprocessing, a subset of relevant features — Age, Gender, family history, benefits, care options, and wellness program — was selected as inputs (X), and the target variable was treatment (y). The dataset was split into training and testing sets in an 80:20 ratio using train test split. A logistic regression model was then trained on the training set and used to make predictions on the test set.

The model achieved an accuracy of 68.92%, indicating it correctly predicted treatment-seeking behavior in nearly 69 out of 100 cases. The classification report further illustrated the model's performance across the two classes — ”Yes” and ”No”. For the ”No” class (those who did not seek treatment), the model had a precision of 0.63 and a recall of 0.75, resulting in an F1-score of 0.69. This indicates that while the model occasionally misclassified actual ”Yes” cases as ”No”, it was fairly reliable at correctly identifying ”No” instances. For the ”Yes” class (those who did seek treatment), the model had a higher precision of 0.76, meaning it was better at making accurate positive predictions. However, the recall was slightly lower at 0.64, suggesting that some actual treatment-seeking individuals were not detected by the model.

Overall, the model's F1-score was balanced across both classes at 0.69, showing reasonably consistent performance. However, there is room for improvement, especially in boosting recall for the ”Yes” class. Potential improvements include incorporating additional features, applying scaling techniques, tuning the decision threshold, or using more advanced classifiers like Random Forest or XGBoost. Despite being a baseline model, logistic regression has demonstrated its value in identifying patterns related to mental health treatment behavior in this dataset.

6. Conclusions

- This project provided a comprehensive analysis of mental health trends within the technology industry by utilizing a real-world dataset from the Kaggle "Mental Health in Tech Survey." Through careful data preprocessing and exploratory data analysis, several important insights emerged:
- Age-related Trends: Younger professionals, particularly those aged 25 to 35, are both the largest respondent group and the most likely to seek mental health support. This highlights a growing awareness and proactive attitude toward mental health in early and mid-career stages. Conversely, participation and help-seeking decline in older age groups, underscoring the need for targeted mental health outreach for these demographics.
- Gender Differences: Contrary to common expectations, male respondents in the tech industry exhibited a higher percentage of mental health help-seeking behavior compared to females, especially between ages 25 and 40. This finding challenges traditional assumptions and may reflect either industry-specific cultural factors or sampling biases. Further investigation is needed to understand gender dynamics in mental health support within tech workplaces.
- Family History Influence: Individuals with a family history of mental illness are significantly more likely to be undergoing treatment, suggesting that prior exposure to mental health issues fosters greater awareness and willingness to seek help. In contrast, those without such history tend to underutilize treatment options, indicating a critical

need for education and destigmatization efforts among this group.

- Workplace Support and Correlations: The analysis identified positive correlations between organizational support measures — such as wellness programs, mental health benefits, and care options — and the likelihood of employees seeking mental health help. Wellness programs, in particular, showed the strongest association, emphasizing their role in encouraging help-seeking behavior. However, the weak correlation between actual treatment and these workplace factors reveals a persistent gap between awareness/support and treatment uptake, likely due to stigma and cultural barriers.
- Overall, the findings underscore the critical importance of proactive and culturally sensitive mental health initiatives in the technology sector. While structural support like wellness programs and benefits can increase awareness and willingness to seek help, organizations must also foster an open and stigma-free environment to ensure that employees feel safe to pursue treatment. Such holistic approaches are essential for improving mental health outcomes and cultivating a healthier, more supportive workplace culture in the fast-paced tech industry.