Analyzing the Contributing Factors of Mental Health in Tech

Qi Dong

Abstract—Mental health is an important concept in our life that enables people to cope with stresses, realize their abilities, and contribute to their community. In this report, the author analyzes the possible features that could relate to mental health issues based on a survey in tech industry. The data was also divided into subgroups to discover hidden relationships that was hidden by the larger population. At last, the author provides some suggestions to employers who are concerned with their employees' mental health and discusses the limitations of this project.

Index Terms—Visualization, Mental Health, Technology

1 Introduction

Mental illness is a growing concern in society, affecting the way individuals think and deal with stress. This is particularly relevant in the workplace, where public speaking and coordination skills are critical in measuring performance, and where mental health issues and excessive stress can affect these skills. A 2021 American Psychological Association survey [1] found that nearly 3 in 5 employees (59%) reported experiencing negative impacts of work-related stress, including lack of interest, motivation, or energy, and lack of effort at work.

As a student interested in working in the tech industry, I have analyzed the contributing factors to mental health issues in this field. While mental health issues can affect anyone regardless of age, gender, or occupation, certain factors may increase the risk in the tech industry. I explored questions such as whether older employees are more likely to experience mental health issues, whether the situation differs between male and female employees due to potential discrimination, and which state has the highest frequency of mental illness. By targeting these questions, employers can better understand the types of people who are at the highest risk for mental health issues and take appropriate action to address and prevent them.

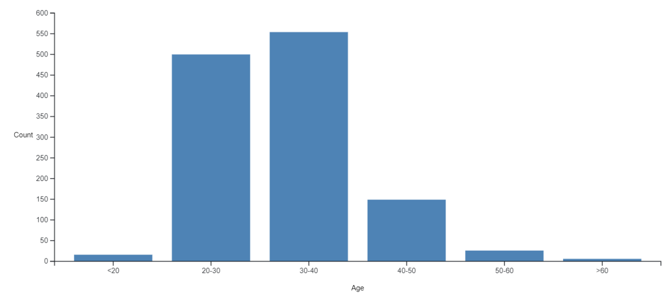

Fig. 1: Distribution over Age

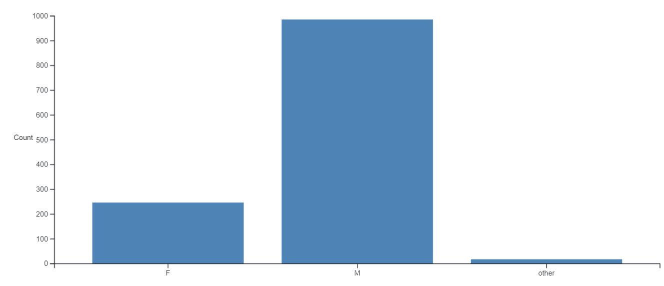

Fig. 2: Distribution over Gender

2 Data

The data used in this analysis was obtained from a survey [2] on Kaggle, consisting of 27 columns, each representing a question related to mental health. To facilitate the creation of visualizations, the author selected 10 features, including age, gender, work interference, remote work, family history, treatment, coworkers, supervisors, country, and state. Prior to analysis, the data was preprocessed by removing null values in these columns, filtering out extreme values, converting the numerical column into categorical bins, and reducing the 47 unique gender values to 3 categories: Female, Male, and Other.

3 Methodology

3.1 General Relationship

The author begins by examining the connection between mental illness and various factors, such as age, gender, work disruption, remote work, family history, and treatment. As all of the factors are categorical, and the count is quantitative, the author opted to use a bar chart to illustrate the relationship, as position is an effective channel that can be used to compare quantity.

Based on the age count depicted in Figure 1, it appears that employees aged 30-40 constitute the largest percentage, followed by those aged 20-30. However, a demographic study conducted in the Information Technology field [3] indicates that employees over the age of 40 comprise 56% of the population, with 30-40 year olds making up 30%, and 20-30 year olds representing 15%. While this research may not perfectly align with the survey data, it raises some concerns. Specifically, the fact that employees over the age of 40 only make up 15% of the population is significantly lower than the research findings. This discrepancy could be due to several reasons, such as employees over the age of 40 being less prone to mental illness or not participating in the survey.

Based on the gender count displayed in Figure 2, we can observe that females and other genders comprise 26% of the population affected by mental illness, which is consistent with the research findings from the demographic study [3] that reports females making up 23% of the overall workforce.

Figure 3 indicates that the majority of respondents find mental illness to interfere with their work only **sometimes**, followed by those who report it **never**, **rarely**, and **often**. Figure 4 illustrates that most respondents do not have the option of remote work, and figure 5 shows that the majority do not have a family history of mental illness. Figure 6 indicates that around half of the respondents seek some form of treatment. As some factors, such as remote work, do not appear to have a clear relationship with mental illness even when broken down into subgroups, their corresponding images will not be displayed in the following sections to save space.

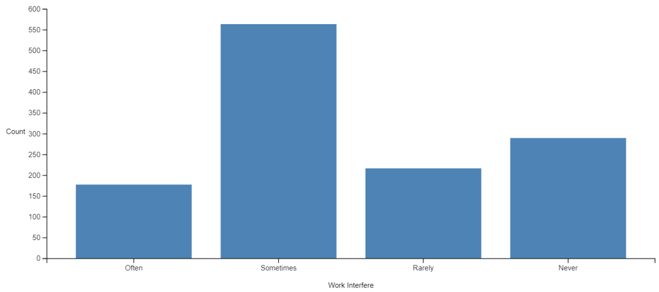

Fig. 3: Distribution over Work Interference

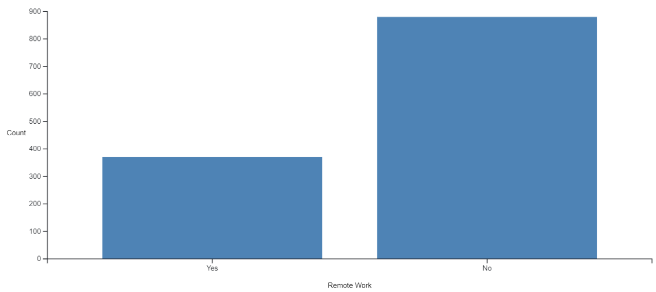

Fig. 4: Distribution over Remote Work

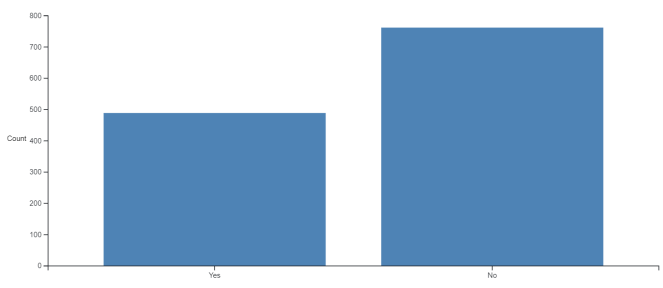

Fig. 5: Distribution over Family History

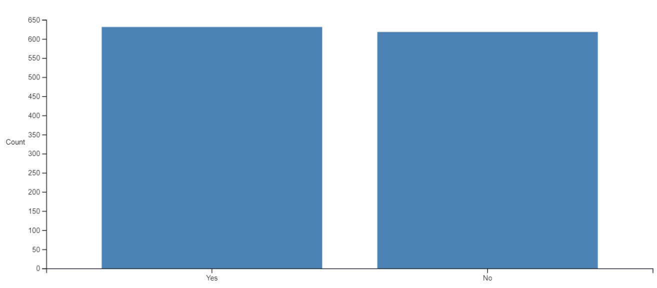

Fig. 6: Distribution over Treatment

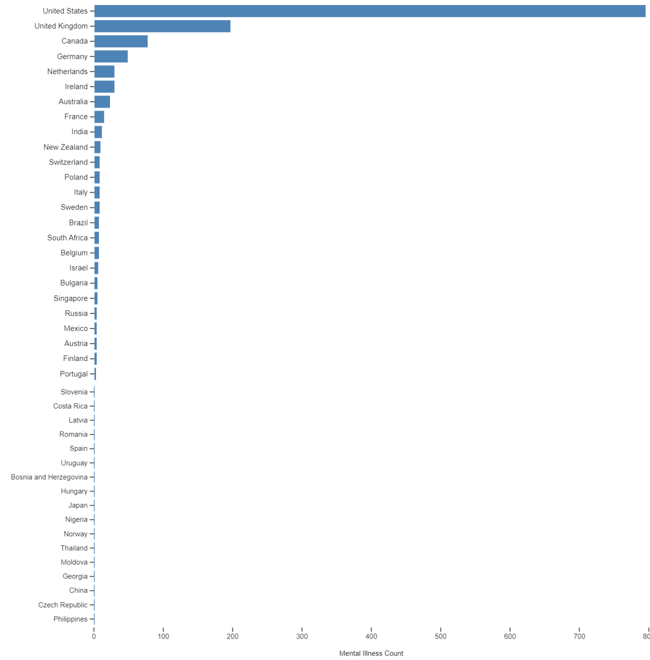

Fig. 7: Distribution over Country

3.2 Geographical Relationship

The author's next step is to examine the geographical information of the respondents, such as their countries and states. As both the country and state variables are categorical and the counts are quantitative, the author has chosen to use a bar chart to display the information.

Figure 7 indicates that the majority of respondents are from the United States, followed by the United Kingdom and Canada. Figure 8 further breaks down the United States respondents by state, with the highest counts observed in California, Washington, and New York. However, it is important to note that having a larger absolute number of respondents in these regions does **not** necessarily mean that people in these areas are more prone to mental illness. It simply reflects the fact that these regions are home to a greater number of tech companies.

3.3 Gender Relationship

The author has further segmented the data by gender and analyzed the relationship between various factors for each gender. Since all the factors are categorical, the author has used a grouped bar chart to present the information. The x-axis in this chart represents the three genders: female, male, and other. Each group of bars represents the distribution of the respective factor across the genders. As the distribution of the "other" gender is similar to that of the female gender, it will not be discussed separately in the chart analysis.

Figure 9 illustrates that the largest number of males falls within the age range of 30-40, while the largest number of females falls within the 20-30 age range. Figure 10 reveals that for both males and females, the majority reported that mental illness **sometimes** interferes with their work. However, the second largest group of females reported that mental illness **rarely** interferes with their work, while the second largest group of males reported that it **never** interferes with their work. In Figure 11, over half of the females reported having a family history of mental illness, which differs from the general distribution presented earlier, where most respondents reported not having a family history. This discrepancy arises because the count of males is significantly greater than the count of females, obscuring the relationship for females. Finally, Figure 12 demonstrates another difference between males and females, where most females seek some form of treatment, while most males do not.

The author investigates the reasons why males are less likely to seek treatment for mental health issues. To do so, she focuses on two

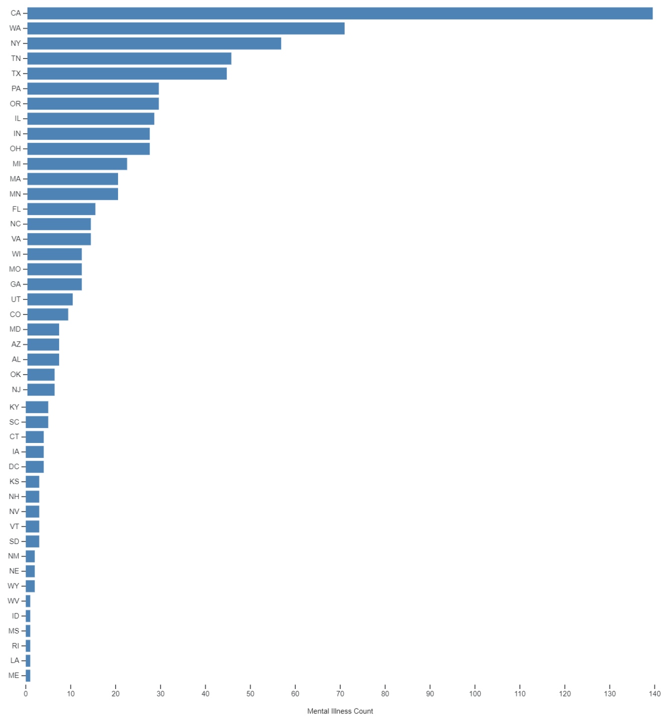

Fig. 8: Distribution over States

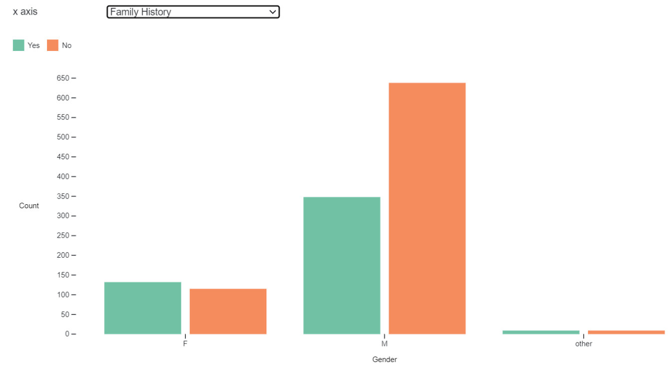

Fig. 11: Distribution of Family History over Gender

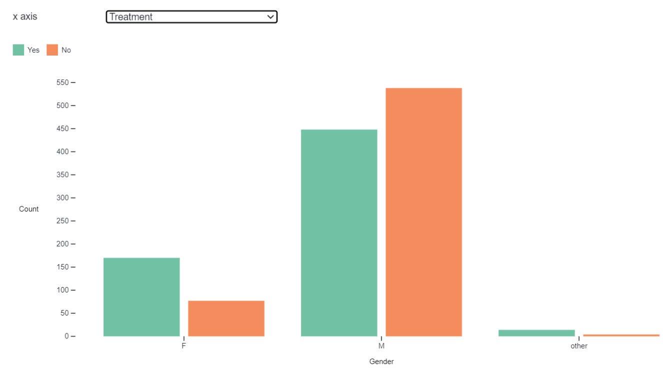

Fig. 12: Distribution of Treatment over Gender

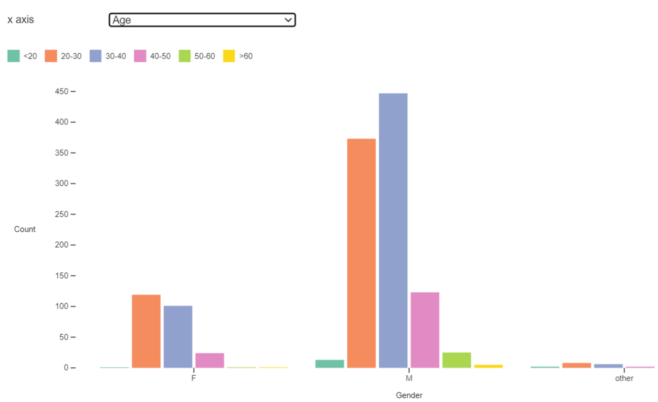

Fig. 9: Distribution of Age over Gender

features from the survey: whether or not males would discuss a mental health issue with their coworkers or supervisors. The same grouped bar chart used earlier is utilized again since the only categories that changed are the factors other than gender.

Figure 13 reveals that the distribution of males is comparable to that of females concerning their willingness to discuss mental health issues with coworkers. Most males are open to discussing their mental health problems with some of their coworkers. Figure 14 illustrates that most males responded affirmatively when asked if they would talk to their supervisors, even more so than females. Therefore, it appears that males are comfortable discussing their mental illness, but most of them do not seek treatment. The author will explore this further in the Treatment Relationship section.

3.4 Age Relationship

The author also segmented the data into different age groups to uncover potential hidden patterns. As all the factors analyzed are categorical,

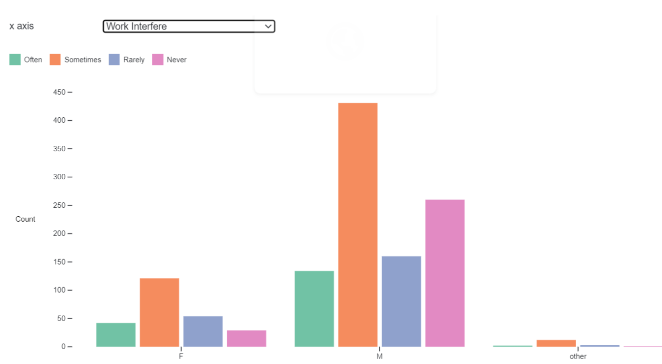

Fig. 10: Distribution of Work Interference over Gender

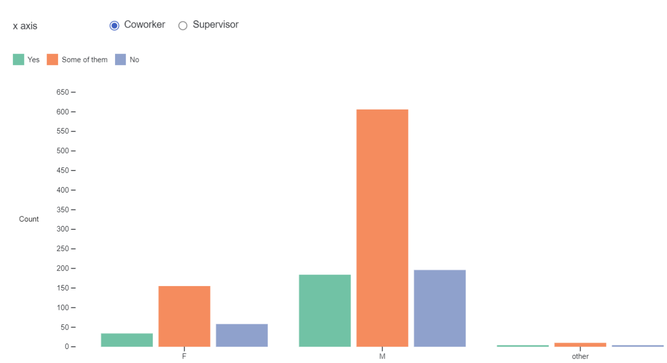

Fig. 13: Distribution of Coworkers over Gender

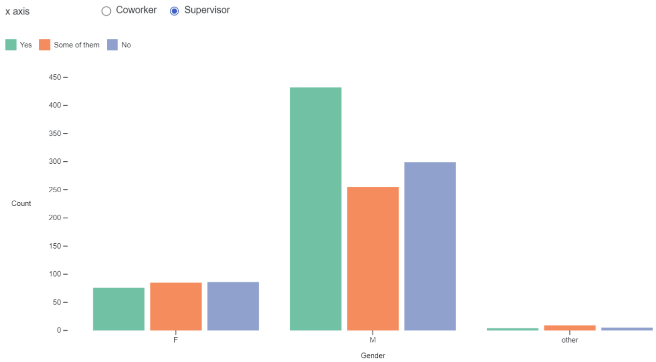

Fig. 14: Distribution of Supervisors over Gender

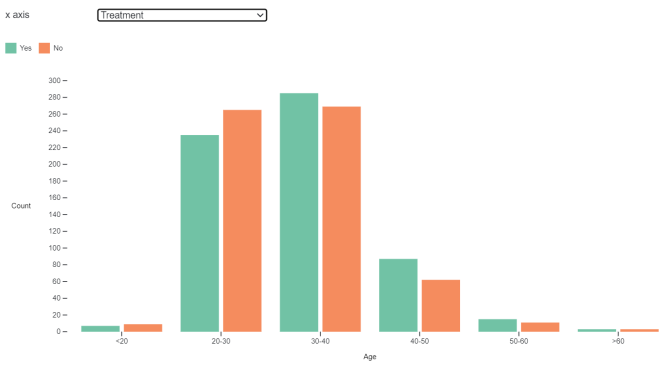

Fig. 15: Distribution of Treatment over Age

a grouped bar chart was employed once again. However, since the distribution for gender, work interference, remote work, and family history did not vary significantly in terms of age, the author opted not to include those plots in order to save space.

However, we can observe different patterns in the treatment seeking behavior of different age groups in Figure 15. We notice that over half of the employees under the age of 30 do not seek treatment, while over half of the employees aged 30 or older seek some form of treatment.

3.5 Treatment Relationship

Finally, the author explores the relationship between seeking treatment and other related factors. The same grouped bar chart is utilized for consistency since all factors remain categorical.

In Figure 16, we can observe that most individuals who seek treatment find mental illness to **sometimes** interfere with their work, and some find it to **often** or **rarely** interfere with their work. Conversely, individuals who do not seek treatment most commonly reported that

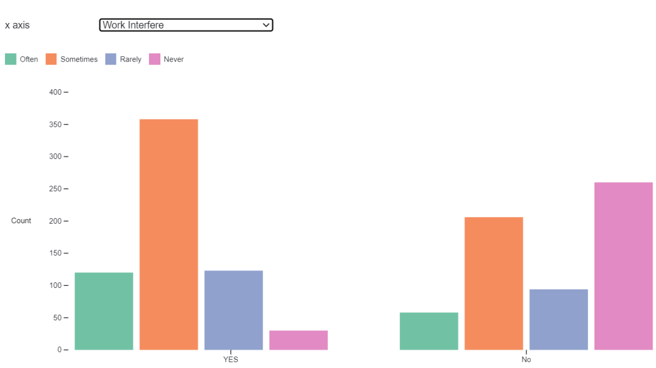

Fig. 16: Distribution of Work Interference over Treatment

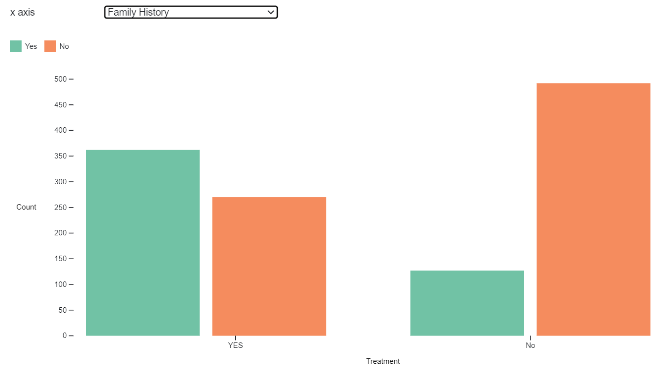

Fig. 17: Distribution of Family History over Treatment

mental illness **never** interferes with their work. This pattern is understandable because those who do not experience interference in their work due to mental illness are less likely to feel motivated to seek treatment. Figure 17 indicates that individuals who seek treatment more often have a family history of mental illness, while those who do not seek treatment are more likely to **not** have a family history of mental illness. This explains why most males do not seek treatment as many of them do not have a family history of mental illness and some do not perceive it to interfere with their work.

4 Discussion

Overall, it is important to pay attention to individuals who do not seek treatment, particularly males under the age of 30 with no family history, and who do not perceive mental illness as interfering with their work. It is essential to understand that not perceiving interference with work does not necessarily mean that mental illness will not affect other aspects of their lives. Therefore, it is crucial to always prioritize mental health and seek help when necessary.

The author has chosen to utilize bar charts for almost all of the visualizations since they are effective in conveying the necessary information and comparing quantitative values. She believes that using complex graphs simply for the sake of their complexity and visual appeal is unnecessary.

5 Conclusion

In this project, the author used data visualization to explore various factors that may be related to mental illness, such as age, gender, work interference, family history, treatment, and geographical information. To uncover hidden relationships and potential subgroups, the author divided the data into different gender, age, and treatment subgroups. However, the project has limitations since all charts are based on absolute numbers instead of normalized percentages. For instance, having more people in California does not necessarily mean people from California are more likely to have mental illness. It only indicates that California is home to many tech companies. Therefore, it is important to exercise caution and not draw conclusions about the positive association of certain factors with mental illness based solely on their absolute count.

References

[1] A. Abramson. Burnout and stress are everywhere, 2022. 1

[2] L. Open Sourcing Mental Illness. Mental health in tech survey, Nov 2016. 1

[3] ZIPPIA. Information technology professional demographics and statistics [2023]: Number of information technology professionals in the us, Sep 2022. 1