Various Factors Impacting Mental Health Disclosure in the Tech Industry

Anika Washburn & Joy Fan

Abstract

In the United States, 1 in 5 adults experience mental illness, but in the tech industry in 2016, just over 50% of employees experienced mental illness [1, 2]. As potential future employees in the tech industry, we were interested in determining what factors impacted whether employees felt comfortable disclosing mental health issues with others, such as coworkers and supervisors. Discussing mental health decreases stigma surrounding the topic and encourages people to seek treatment, which is vital, as the longer the mental illness is not treated, the more severe it can become [3]. Through hypothesis testing, we determined which factors were correlated with the willingness of employees to disclose mental health issues. We found that the availability of resources does not impact employee disclosure, while perceived negative consequences and gender do. We conclude that it is most important to create an environment that allows employees to feel comfortable sharing any mental health issues, regardless of gender.

1. Introduction

Mental health includes our social, psychological, and emotional well-being which can shape how we think, feel, and act [4]. It has an impact on every aspect of our lives and is an integral part of our decision-making process. With 63% of Americans part of the US labor force, the workplace is vital in shaping the well-being and mental health of individuals [5]. Because 50% of employees experienced mental health illness in 2016, our project aims to specifically help tech companies improve their employees' mental health by determining which factors related to the workplace tend to impact employees' willingness to disclose any mental health issues [2]. With more employees sharing these issues, employers can take action to provide better resources and create an open and inclusive environment to all.

1.1 Data

The data was provided by Open Sourcing Mental Health (OSMI) [6]. OSMI is an organization whose purpose is "raising awareness, educating, and providing resources to support mental wellness in the tech and open source communities" [6]. One way in which the group performs research is using an annual survey called Mental Health in Tech to further understand employees' experiences regarding mental and physical health issues as well as learn about employer mental health offerings. All responses are voluntary and it targets employees at companies in the tech industry. For our analysis, we utilized the Mental Health in Tech survey responses from the years 2014 and 2019 [7, 8]. The abundance of data surrounding mental health for employees in the tech industry allowed us to more closely analyze a variety of factors that may impact whether an employee feels comfortable disclosing any mental or physical health issues to others.

1.2 Approach

Originally, our main interest was determining whether there is a correlation between the availability of resources and employees' comfort in sharing their mental health issues with coworkers and supervisors. After performing some tests, we decided to broaden our scope to look at other factors that could impact disclosure so we could provide a more detailed analysis and recommendation to companies. After cleaning the data, we then began to check if there was any correlation between the variables. All of the factors we were interested in were categorical variables, so we used chi-squared tests to determine if there was dependency between the factors and heatmaps to visualize the associations. For variables that were dependent on one another, we created bar charts to provide a more clear visual representation of the association. Additionally, we used hypothesis testing to compare and contrast the 2014 and 2019 datasets. This helped us understand whether there was a significant change in factors like mental health benefits, wellness programs, and employee disclosure from year to year.

1.3 Summary & Insights

Through the tests mentioned, we were able to draw conclusions and make suggestions about how tech companies should approach mental health in the workplace. We found that the availability of resources such as a wellness program or employer benefits do not have any impact on employee disclosure. However, the strong belief of negative consequences surrounding disclosing mental health prevents employees from sharing information. From 2014 to 2019, we see an increase in employer benefits, discussions, and knowledge of care options for mental health, but no increase in disclosure of mental health issues to supervisors or coworkers. We conclude that although there are more discussions concerning mental health, we may not be having the right discussions in these programs. Instead, we should focus on creating an environment where it is clear there are no negative consequences surrounding mental health.

Additionally, we found that females believe there are negative consequences for mental health disclosure more than males and share mental health issues in interviews less than males. This could be due to feelings of imposter syndrome and potential inequality towards women during the hiring process. Companies should ensure that men and women have equal opportunities in the workplace and continue to hire a diverse group of individuals.

2. Methods

The raw data came from OSMI's Mental Health in Tech Survey responses for the years 2014 and 2019 [7, 8]. There were many columns in the two datasets, but for our analysis we first analyzed the following factors and their relation to mental health disclosure to coworkers and supervisors:

- Employee knowledge of mental health benefits
- Discussion of mental health in wellness programs
- Perceived negative consequences of disclosing mental and physical health issues
- Gender

Our findings indicated that gender is an important factor in disclosure, so we decided to perform additional testing to check whether gender was correlated to any of the following factors:

- Perceived negative consequences of disclosing mental and physical health issues
- Willingness to disclose mental or physical health issues in interviews

Originally, there were 1259 rows x 27 columns for the 2014 survey results and 352 rows and 82 columns for the 2019 survey results [7, 8].

Figure 1 describes the workflow outlined below. After creating our BitBucket repo, we performed data cleaning to filter and normalize both datasets. We then created contingency tables, and based on our results, we did additional testing including chi-squared tests and hypothesis testing. Finally, we created visualizations of our findings.

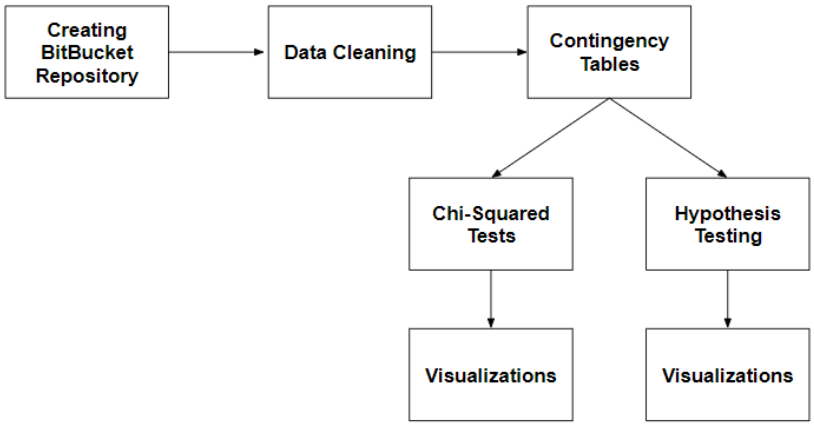

Figure 1: Project Workflow

2.1 Data Cleaning

Because we wanted to compare the 2014 and 2019 survey results, we cleaned each dataset by renaming columns and standardizing the results (i.e. True → Yes, False → No) to make them look as identical as possible. Both datasets had misspellings, different capitalization, and a variety of values for data within a specific column so we created standard values and categorized entries to these values. For example, the gender column contained Female, female, Femake, f, Woman, Cis Female, and more to indicate a female gender. We replaced these values with 'F'. To do this, we used NumPy library's `unique` function to see which specific values were in the column and the NumPy library's `replace` function to convert those values to standardized values. We only included data for individuals working in the United States.

2.1.1 Cleaning 2014 Dataset

The Kaggle Mental Health in Tech survey data from 2014 consists of 1260 rows and 27 columns [7]. This dataset contained attributes indicating the time the survey was taken and whether the

entry was from an individual in the tech industry, so the data was filtered to only include 2014 entries of those working at a tech company. After cleaning, our dataset contained 581 rows and 24 columns.

2.1.1 Cleaning 2019 Dataset

In the Mental Health in Tech survey results from 2019, there are 352 rows and 81 columns [8]. In order to make this dataset look more like the 2014 dataset, we omitted unnecessary columns. Additionally, since the column names in the 2019 dataset were the lengthy questions asked in the survey, we shortened them to mimic the 2014 dataset column names. After filtering and standardizing the data, our dataset contained 126 rows and 23 columns.

2.2 Testing

Because all of the survey data is categorical, we needed to use contingency tables to show the frequency of different variables and their associations. To do this, we used the Pandas library's `crosstab` function. Additionally, for visualizations purposes, we used sns library's `heatmap` function to differentiate between the frequencies of each association. We could then use these contingency table values to conduct chi-squared tests and hypothesis testing.

2.2.1 Chi-Squared Tests

Using the contingency tables we created for the frequency of different variables, we performed chi-squared tests to see if different variables are dependent on each other. We used SciPy.stats library's `chisquare2_contingency` function to find the test statistic, p-value, degrees of freedom, and expected value of the contingency tables. Then, using the test-statistic and comparing it against the critical value of $\propto = 0.05$, we determined whether the variables are independent or dependent. To determine the critical value, we used SciPy.stats library's `chi2.ppf(q, df)` function, which uses a significance level (1 - 0.05 = 0.95 in our case) and degrees of freedom (calculated in the original chi-squared test function). If the test-statistic was greater than the critical value, the variables are dependent, meaning there is association. Otherwise, the variables are independent, meaning there is no association. Additionally, we looked at the p-values in relation to $\propto = 0.05$ to ensure that the decision regarding the hypotheses was consistent.

2.2.2 Hypothesis Testing

Another part of our study was determining if there was a significant change between certain variables in 2014 and 2019. Our null hypothesis states that these attributes stayed the same from 2014 to 2019 and our alternative hypothesis says that there was an increase or decrease in these attributes between the years. We performed hypothesis tests by finding the difference of two proportions of the answers of "Yes" to the various variables we tested. We again calculated the test statistic and p-value. If the p-value was less than $\propto = 0.05$, we rejected the null hypothesis which indicated the attributes stayed the same. Otherwise, we failed to reject the null hypothesis.

3. Results and Discussion

We created heatmaps to visualize contingency table results for a variety of factors and their relation to employees' willingness to disclose mental health issues with their coworkers and supervisors.

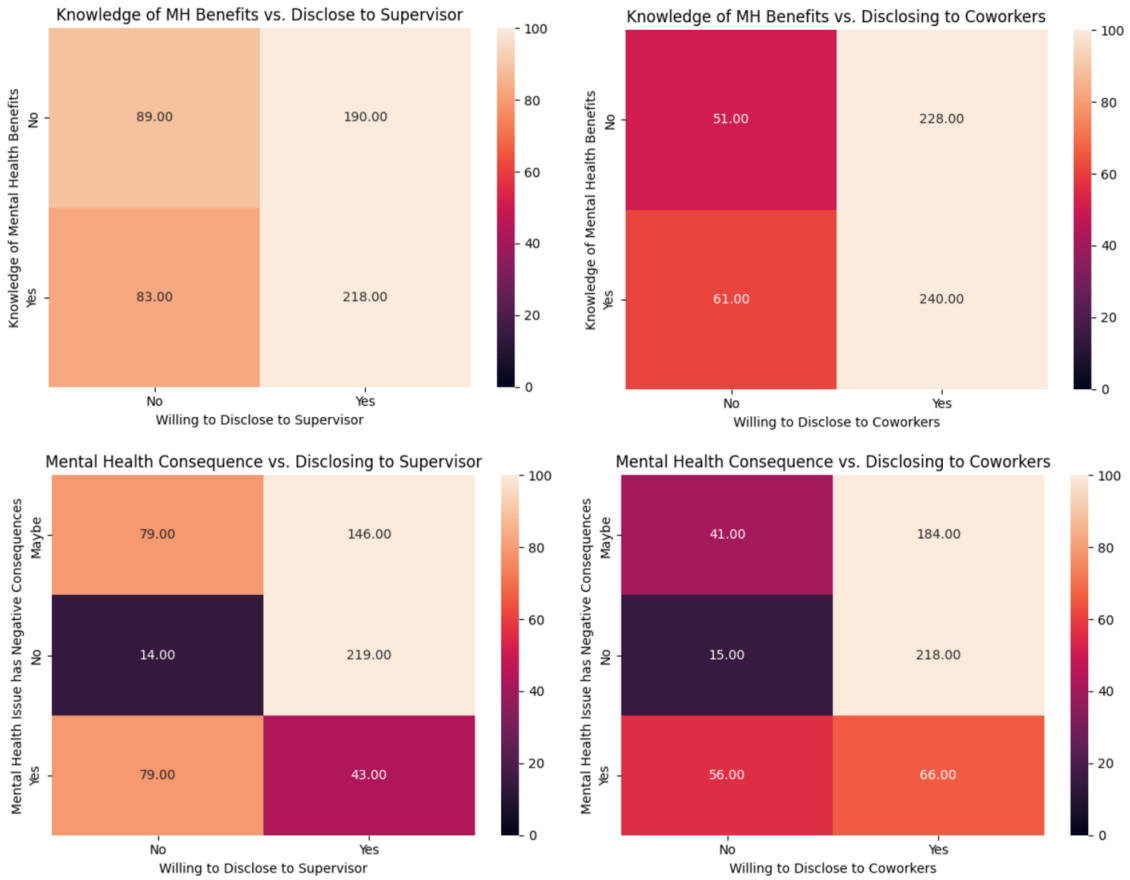

Figure 2: Heatmaps to visualize correlation between variables

Figure 2 shows these visualizations and through chi-squared tests, we were able to determine which factors impacted disclosure. The two upper graphs had 1 degree of freedom, $\propto = 0.05$, and a critical value ($\chi_{\propto}^{2}$) of 3.841. For knowledge of mental health benefits vs. disclosing to supervisor, the test statistic (χ^2) was 1.099 and the p-value was 0.2944. For knowledge of mental health benefits vs. disclosing to coworkers, $\chi^2 = 0.250$ and the p-value was 0.6169. In both cases, the test statistic is below the critical value and the p-value is greater than alpha, indicating independence between the variables. We created similar contingency tables that indicated no dependency between disclosure and mental health discussions via a wellness program as well as disclosure and employee knowledge surrounding their care options.

The two lower graphs in Figure 2 had 2 degrees of freedom, $\propto = 0.05$, and $\chi_{\propto}^{2} = 5.991$. For mental health consequence vs. disclosing to supervisor, $\chi^2 = 137.711$ and the p-value was 1.248 x

10^{-30}. For mental health consequence vs. disclosing to coworkers, $\chi^2 = 80.314$ and the p-value was 3.631×10^{-18}. These values indicate that there is dependency between the variables because the test statistic is above the critical value and the p-value is less than alpha. Thus, we found that employee's knowledge of mental health benefits do not have an impact on their willingness to disclose to coworkers or supervisors, but when employees believe there is a negative consequence associated with sharing mental health issues, they are less likely to disclose this information to a supervisor or coworker.

Figure 3 shows the difference in 2014 and 2019 of the proportions of the availability of mental health benefits, mental health discussions, and knowledge of options. The blue bars represent 2014 and the orange bars represent 2019 data. We wanted to see if there was an increase in mental health benefits, mental health discussions, and knowledge of options in 2019 from 2014. To see if this was true, we conducted hypothesis testing.

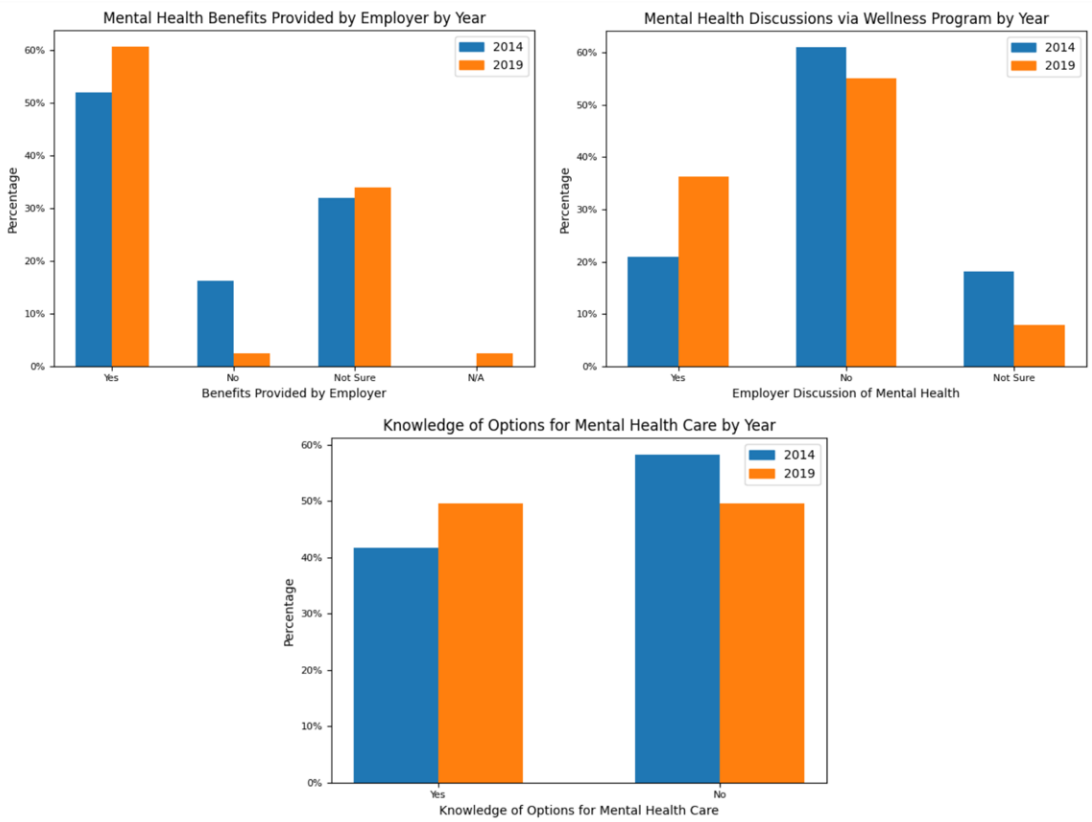

Figure 3: Difference in mental health benefits, mental health discussions, and employee knowledge of options for mental health care from 2014 to 2019

For the mental health benefits, we can see in Figure 3, for 2014 the proportion that have benefits is 51.9% and for 2019 the proportion that have benefits is 61.1%. When conducting the hypothesis test, we found the test statistic to be $z = -1.8797$ and a p-value of 0.03005. For the mental health discussions via a wellness program, Figure 3 shows that in 2014 the proportion of

companies with those discussions is 20.9% and for 2019 the proportion that have benefits is 36.5%. When conducting the hypothesis test, we found the test statistic to be z = -3.7458 and a p-value of 0.00009. For the knowledge of care options bar chart in Figure 3, in 2014 the proportion of companies providing care options is 41.7% and for 2019 the proportion that have benefits is 50%. When conducting the hypothesis test, we found the test statistic to be z = -1.6998 and a p-value of 0.04457.

Due to the fact that the p-value for all of these tests is less than the significance level $\propto = 0.05$, we can reject the null hypothesis and conclude that there is an increase in the availability of resources in terms of mental health benefits, mental health discussions via a wellness program, and knowledge of care options in 2019 compared to 2014. Tech companies are taking initiatives to provide more resources towards mental health, but this does not necessarily mean employees are more willing to disclose mental health issues.

Figure 4 shows the difference in 2014 and 2019 of the proportions of employee's willingness to disclose their mental health issues to coworkers vs. supervisors. To see if there was an increase in disclosure, we performed hypothesis tests again.

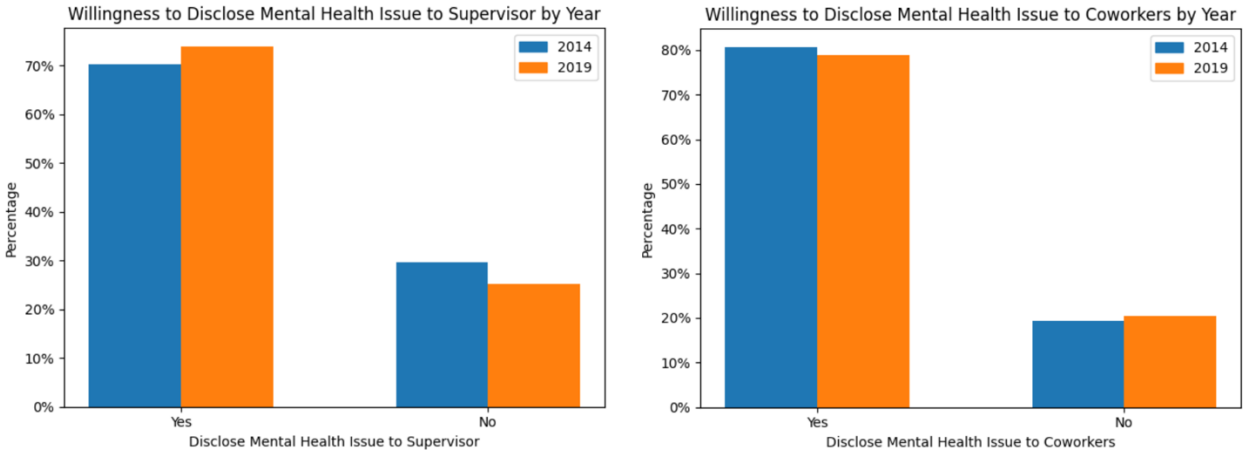

Figure 4: Difference in mental health disclosure to supervisors and coworkers in 2014 and 2019

For the willingness to disclose to supervisors graph in Figure 3, we see that in 2014 the proportion is 70.3% and in 2019 the proportion is 74.6%. The results of the hypothesis test showed that the test statistic z = -0.9558 and p-value = 0.16853. For the willingness to disclose to coworkers graph in Figure 3, we see that in 2014 the proportion is 80.58% and in 2019 the proportion is 79.4%. When conducting the hypothesis test, the results show that the test statistic z = -0.3398 and the p-value = 0.36393.

Since the p-value is greater than the significance level $\propto = 0.05$ for both tests, we cannot reject the null hypothesis and conclude that there is not enough evidence to suggest that there was an

increase in employees disclosing to their supervisors and coworkers in the years 2014 to 2019. From these hypothesis tests, it proves that even though there was a significant increase in the availability of resources, the willingness to disclose mental health issues to coworkers and supervisors has not increased. Therefore, there may be other factors to consider when looking at what can impact an employee's willingness to disclose their mental health issues.

Because we found that the willingness to disclose mental health issues to coworkers and supervisors was dependent on perceived negative consequences of sharing, we decided to look into gender and if it impacted any of these factors. The top bar chart in Figure 5 shows that regardless of gender, more employees believe there is a negative consequence associated with mental health issues than physical health issues. The remaining two bar charts indicate that, in general, more females believe there are negative consequences for sharing mental and physical health issues than males. To verify that the difference is significant, we again used hypothesis testing and found this to be true. Females may feel less comfortable sharing information surrounding their mental and physical health in the workplace than men because they generally believe there are negative consequences. This could be due to feelings of imposter syndrome or believing they are at a disadvantage as a female and sharing this information would further disadvantage themselves.

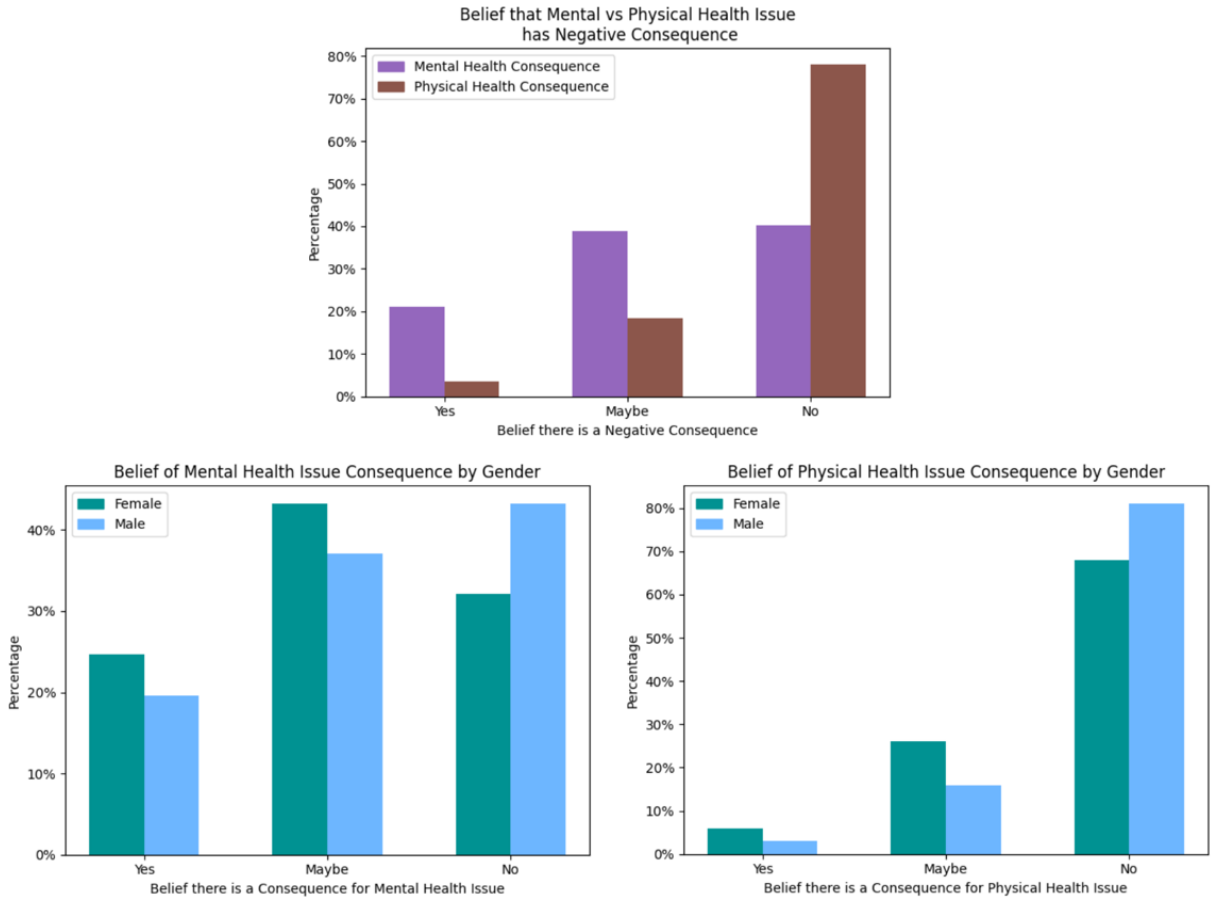

Figure 5: Belief that mental or physical health issue has negative consequence based on gender

We also performed hypothesis testing to determine whether gender had an impact on the willingness to share mental or physical health issues in an interview. The top graph of Figure 6 shows that regardless of gender, more employees are comfortable sharing physical health issues in interviews than mental health issues. This is likely due to the stigma associated with mental health. The remaining two graphs in Figure 6 show that males are more likely to share mental or physical health issues in interviews than females. This could be due to inequality towards females during the hiring process and the belief that sharing this information could hurt a female's chances of receiving a job offer.

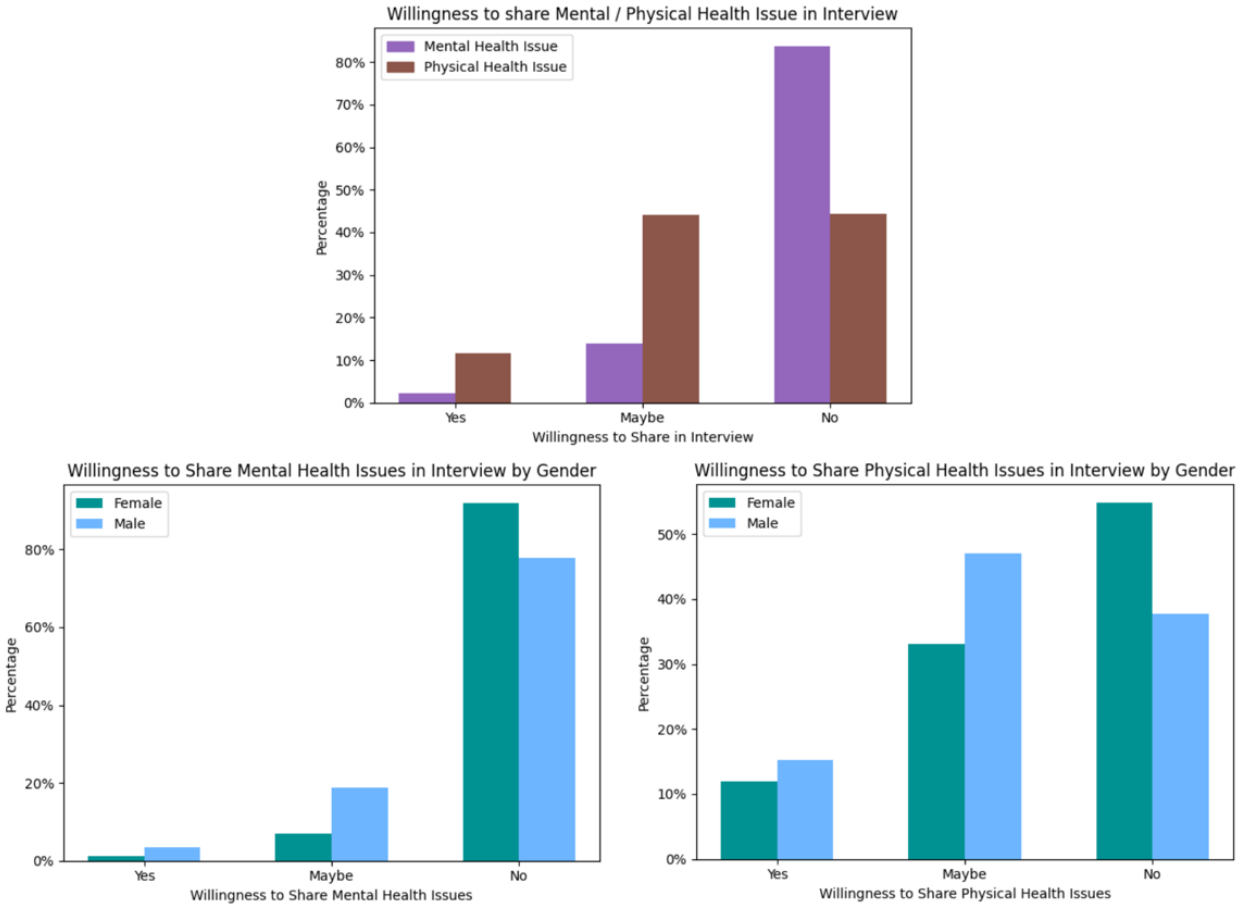

Figure 6: Belief that mental or physical health issue has negative consequence based on gender

3.1 Limitations

In terms of our project data and analysis, there are a few limitations worth noting. Due to the scope of our project, we were only able to use data from two years. Ideally, we would have cleaned data from more surveys and performed our analysis over these years. Additionally, we wanted to use a more current dataset, but the most recent one with enough data was from 2019. Even this dataset did not have nearly as much data as the 2014 dataset, but we accepted the tradeoff so we would have data from a more current year. This could have skewed results especially when comparing and contrasting findings between years. Next, the survey itself is

voluntary, so we may not have had an accurate representation of employees in the tech industry. For example, those who may feel particularly passionate about mental health in the tech industry may be more inclined to fill out the survey rather than those who are not passionate about the subject. Finally, our dataset had a large gender discrepancy with 519 males, 175 females, and 12 individuals identifying as non-binary. This could have impacted our results and conclusions specifically surrounding gender.

4. Conclusion

In the end, we found that the perceived negative consequences associated with mental health had the most impact on the willingness of employees to disclose mental health issues to coworkers or supervisors. The increase in the availability and knowledge of mental health resources from 2014 to 2019 did not result in more employees sharing this information. While the accessibility of mental health resources is important for employers to offer, we found that minimizing the perceived negative consequences surrounding mental health issues will encourage more disclosure. Another factor that impacts an employee's willingness to disclose is gender. More females believe that there is a perceived negative consequence for disclosing their mental health issues and overall, this can lead to less disclosure of their mental health to their coworkers, supervisors, and during interviews. While companies are leading initiatives to create a more diverse and equitable work culture, especially towards gender, there needs to be more affirmations for these minority groups that there will not be repercussions for disclosing this sensitive information.

This project could be expanded upon by looking at data from a more recent year and analyzing how the COVID-19 virus impacted the mental health for employees in the tech industry. Additional data with new factors that may impact an employee’s willingness to disclose could be helpful in building on our results and in making more suggestions to tech companies.

References

[1] “Mental Health By the Numbers,” *National Alliance on Mental Illness*, Jun. 2022. https://www.nami.org/mhstats (accessed Dec. 11, 2022).

[2] “OSMI Mental Health in Tech Survey 2016,” *osmi.typeform.com*. https://osmi.typeform.com/report/Ao6BTw/U76z (accessed Dec. 11, 2022).

[3] “Why do we need to talk about mental health? | Mass.gov,” *www.mass.gov*. https://www.mass.gov/info-details/why-do-we-need-to-talk-about-mental-health#:~:text=Mental%20health%20is%20just%20as (accessed Dec. 11, 2022).

[4] “Mental Health,” *Centers for Disease Control and Prevention*, 2019. https://www.cdc.gov/mentalhealth/index.htm (accessed Dec. 14, 2022).

[5] “Mental Health in the Workplace,” *Centers for Disease Control and Prevention*, Apr. 10, 2019. https://www.cdc.gov/workplacehealthpromotion/tools-resources/workplace-health/mental-health/index.html (accessed Dec. 13, 2022).

[6] “Open Sourcing Mental Illness,” *osmihelp.org*, 2015. https://osmihelp.org/ (accessed Dec. 11, 2022).

[7] “Mental Health in Tech Survey,” *Kaggle*, 2016. https://www.kaggle.com/datasets/osmi/mental-health-in-tech-survey (accessed Dec. 11, 2022).

[8] “OSMI Mental Health In Tech Survey 2019,” *Kaggle*, 2019. https://www.kaggle.com/datasets/osmihelp/osmi-mental-health-in-tech-survey-2019 (accessed Dec. 11, 2022).