COGS 108 Final Project Group 005

Jennifer Jiang, Hannah Ordonez, Daniela Garcia, Ghada Barhoush, Stephanie Li, Cynthia Delira

Background

The tech industry has been historically dominated by men and gender-based discrimination has proven to be a common problem in the workplace, even today. Studies have shown that both male and female employees are shown to be more likely to dismiss female applicants' performances and that women are not often found in managerial positions. Here are a few articles that we found relevant to these issues:

Feld, J., Ip, E., Leibbrandt, A., & Vecci, J. (2022, October 7). Identifying and overcoming gender barriers in tech: A field experiment on inaccurate statistical discrimination. https://www.cesifo.org/en/publications/2022/working-paper/identifying-and-overcoming-gender-barriers-tech-field-experiment

Hyrynsalmi, S. (May 2019) The Underrepresentation of Women in the Software Industry: Thoughts from Career-Changing Women https://www.researchgate.net/publication/335496057_The_Underrepresentation_of_Women_in_the_Software_Industry_Thoughts_from_Career-Changing_Women

McKinsey & Company.(1 March 2022). Repairing the Broken Rung on the Career Ladder for Women in Technical Roles. https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/repairing-the-broken-rung-on-the-career-ladder-for-women-in-technical-roles.

Given this information, we wanted to focus on the prevalence of mental health issues within workplaces in the tech industry, specifically when it comes to gender.

Our Question

How might the experiences of women and men differ in tech in terms of mental health issues/comfortability of discussing mental health issues?

Hypothesis

Female-identifying people in the tech industry are **more likely** to struggle with mental health issues and addressing those mental health gaps with their employer.

Women are often a minority in the tech workplace and as a result may feel less comfortable reaching out to coworkers or employers.

Additionally, they may experience gender discrimination or harassment from men in the workplace that can lead to a variety of mental health issues.

Data Cleaning

Data Cleaning

1. Removing columns

There were 27 total columns in our original dataset. We decided to drop some of the columns that we felt were not relevant to our analysis and the question we wanted to answer. These variables/columns are shown on the right:

- Timestamp
- State
- wellness_program
- seek_help
- anonymity
- phys_health_consequence
- coworkers
- supervisor
- mental_health_interview
- phys_health_interview
- mental_vs_physical
- comments

Data Cleaning

2. Handling outliers in data

Some of the questions allowed for open-answer responses, which led to outliers in certain columns. One of these columns was the “Age” column, where we had numbers as low as -1726 and as high as 99999999999.

Solution: we took the lower .05% and upper .05% of the 'Age' range, which removed all ages below 18 and above 60. This addressed the issue of the extreme outliers and only took out 1% of our observations

Data Cleaning

3. Addressing non-uniformity in data

Another column that allowed for open-answers was the “Gender” column. There were several variations of ‘male’ or ‘female’ that would cause problems later in data analysis.

Solution: We replaced all values that didn't match 'M' or 'F' exactly with the appropriate labeling of either 'M', 'F', or 'NB' (Non-Binary).

Data Cleaning

4. One-Hot Encoding Categorical Variables

We decided to adapt some of the columns with strictly Yes/No values to follow one-hot encoding. That is, we replaced 'Yes' values with a 1, and 'No' values with a 0 for ease of later calculations and analysis.

self_employed	family_history	treatment
NaN	No	Yes
NaN	No	No
NaN	No	No
NaN	Yes	Yes
NaN	No	No
...	...	...
No	No	Yes
No	Yes	Yes
No	Yes	Yes
No	No	No
No	Yes	Yes

→

Family History	Sought Treatment	Work Remotely
0	1	0
0	0	0
1	1	0
0	0	1
1	0	0
...	...	...
1	1	1
0	1	0
1	1	1
1	1	0
0	0	1

Data Cleaning

5. Reordering columns

We then re-ordered the remaining 15 columns to place columns with string data (non-numerical values) at the beginning, and then put our one-hot encoded columns at the end. This was done for ease of interpretation.

	Age	Gender	Country	Number Employees	Work Interfere	Offer Benefits	Knowledge of Care Options	How Easy to take Medical Leave	Mental Health Consequences	Self Employed	Family History	Sought Treatment	Work Remotely	Observed Consequences	Worl Interfer Ordina
0	37	F	United States	6-25	Often	Yes	Not sure	Somewhat easy	No	NaN	0	1	0	0	4.
2	32	M	Canada	6-25	Rarely	No	No	Somewhat difficult	No	NaN	0	0	0	0	2.
3	31	M	United Kingdom	26-100	Often	No	Yes	Somewhat difficult	Yes	NaN	1	1	0	1	4.
4	31	M	United States	100-500	Never	Yes	No	Don't know	No	NaN	0	0	1	0	1.
5	33	M	United States	6-25	Sometimes	Yes	Not sure	Don't know	No	NaN	1	0	0	0	3.
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	..
1240	29	M	United States	100-500	Sometimes	Yes	Yes	Don't know	Yes	0.0	1	1	1	0	3.
1242	26	M	United Kingdom	26-100	NaN	No	No	Somewhat easy	No	0.0	0	1	0	0	Na
1243	32	M	United States	26-100	Often	Yes	Yes	Somewhat difficult	No	0.0	1	1	1	0	4.
1244	34	M	United States	More than 1000	Sometimes	Yes	Yes	Somewhat difficult	Yes	0.0	1	1	0	0	3.
1245	46	F	United States	100-500	NaN	No	Yes	Don't know	Yes	0.0	0	0	1	0	Na

Data Analysis

Analyzing Levels of Mental Health Work Interference and Gender

Participants were asked the question, "if you have a mental health condition, do you feel that it interferes with your work?"

Participants were able to choose from the answers: "Never," "Rarely," "Sometimes", or "Often."

During data cleaning, we mapped the text answers to ordinal numbers, 1 to 4 respectively. Through data visualization, we were able to analyze the data in concern to number and proportion of employees per gender who felt that their mental health interfered with their work, as well as average levels of mental health interference as a whole.

1.

Analyzing Levels of Mental Health Work Interference and Gender

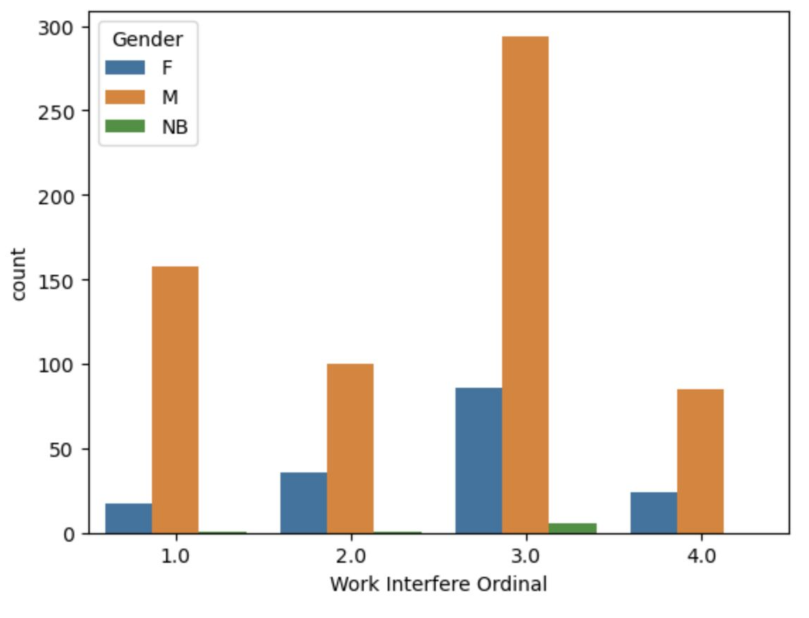


Figure 1:

This bar graph compares numbers of female, male, and non-binary employees by levels of mental health work interference.

1.

Analyzing Levels of Mental Health Work Interference and Gender

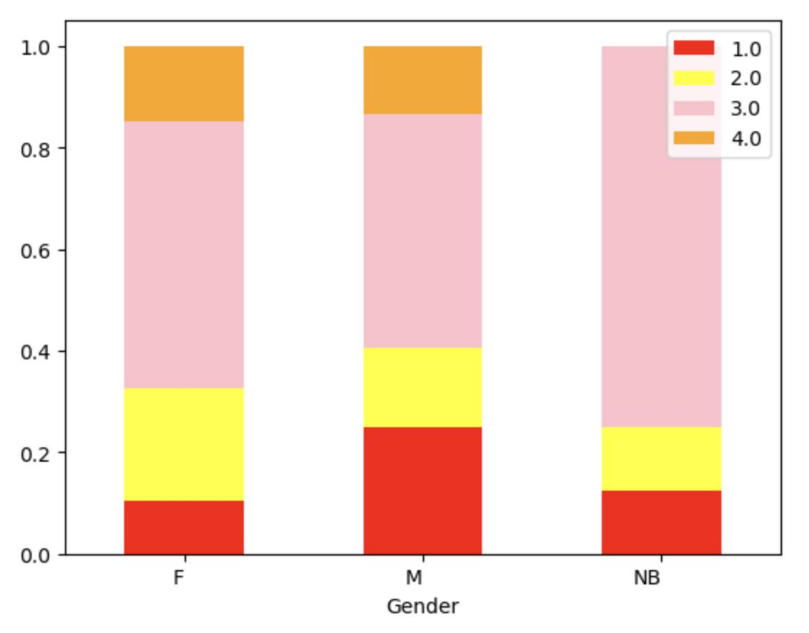


Figure 2:

This stacked bar plot compares proportions of perceived level of mental health and work interference among female, male, and non-binary employees.

1.

Analyzing Levels of Mental Health Work Interference and Gender

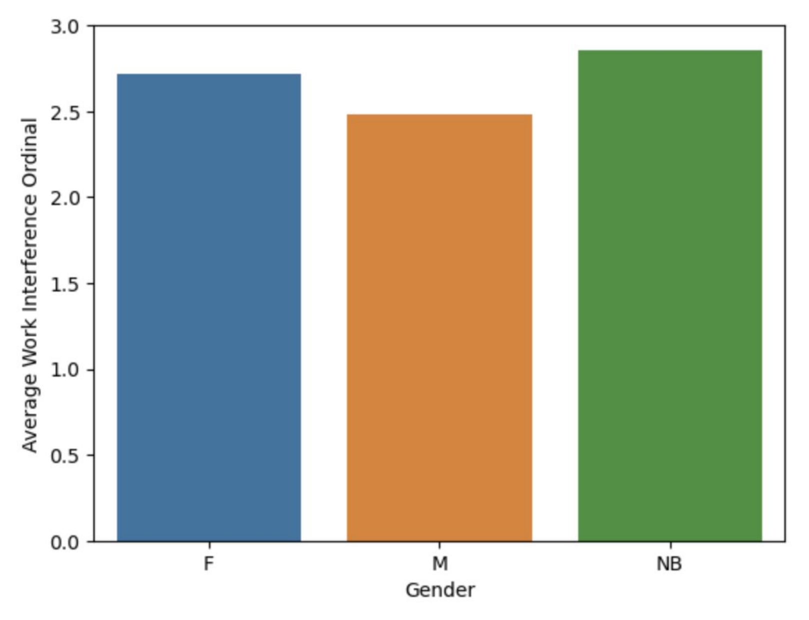


Figure 3:

This bar plot compares averages of perceived mental health work interference among female, male, and non-binary employees.

1.

Analyzing Levels of Mental Health Work Interference and Gender

Through comparing levels of mental health interfering with employees' work, we were able to observe that **female and non-binary employees** in the tech industry noticed a **slightly greater** level of work interference from their mental health compared to their male counterparts. This observation is imperative in searching for and addressing potential harms in work environments that may influence non-male employees differently than male employees.

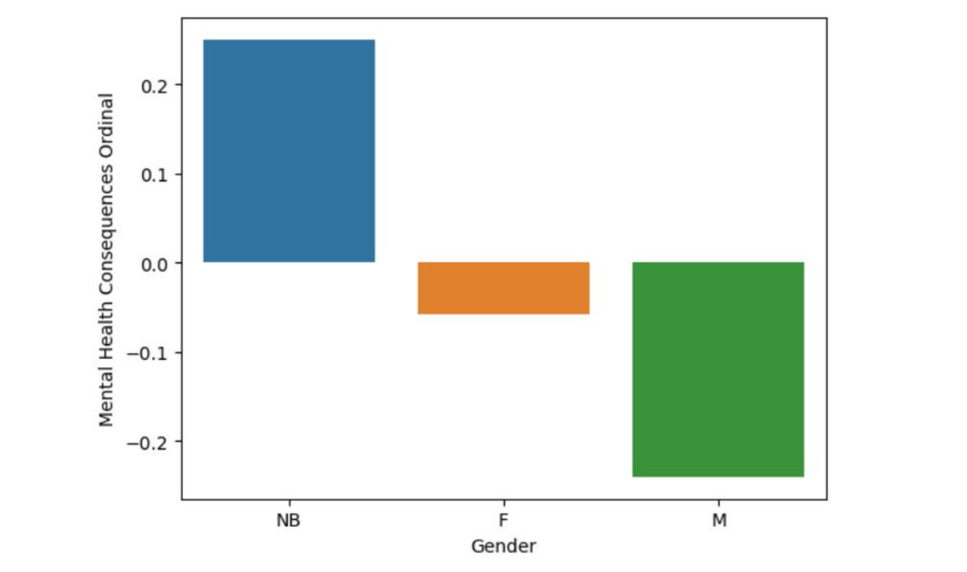
Analyzing Mental Health Consequences and Gender

Participants were asked the question, "Do you think that discussing a mental health issue with your employer would have negative consequences?"

This is an important question regarding the research that we are analyzing. Our research question asks whether or not there is a difference between how comfortable women feel in terms of discussing their mental health issues with their employers as compared to men. As part of our hypothesis and background research, we believe that women would have more trouble opening up about these issues due to their fear of not being taken seriously, or being judged as they can be already undermined due to their gender.

2.

Analyzing Mental Health Consequences and Gender

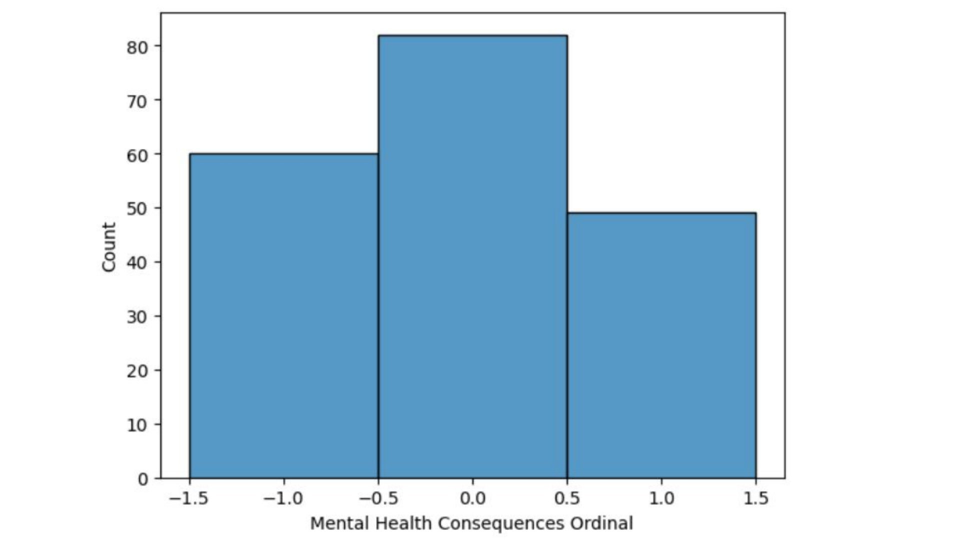


In this section, we will compare the variables of "Mental Health Consequences" with values of -1 for 'No', 0 for 'Maybe and 1 for 'Yes'. "Gender" has values of M for Male, F for Female, and NB for non-binary. Firstly, we found the mean for the responses from all the genders:

Non-binary employees have the highest average, followed by female employees. Remember that a value of ‘1’ means they do feel like they would face consequences if they share their issues.

Female and non-binary employees having values closer to 1 as compared to male employees implies a difference between the gender identities in how comfortable they feel with discussing their mental health and not fearing consequences.

Analyzing Mental Health Consequences and Gender

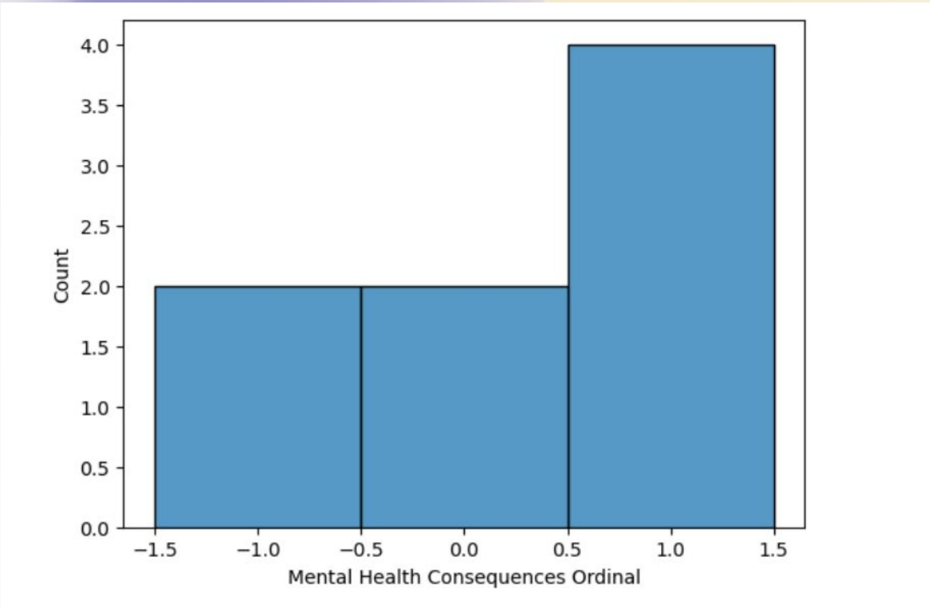


Second, we used a histogram to visually look at the responses from each separate gender:

The responses from female employees:

2.

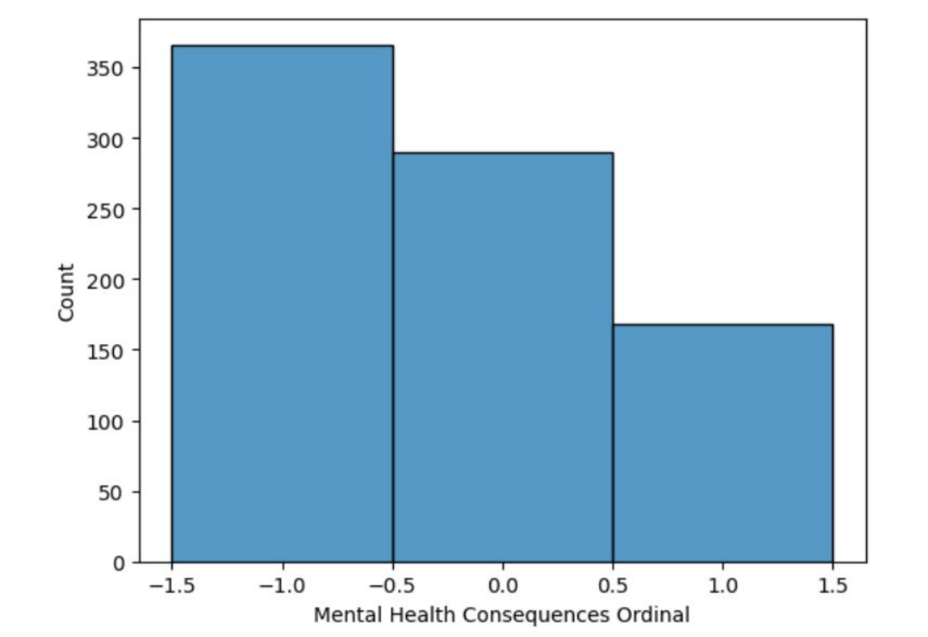
Analyzing Mental Health Consequences and Gender



Second, we used a histogram to visually look at the responses from each separate gender:

The responses from non-binary employees:

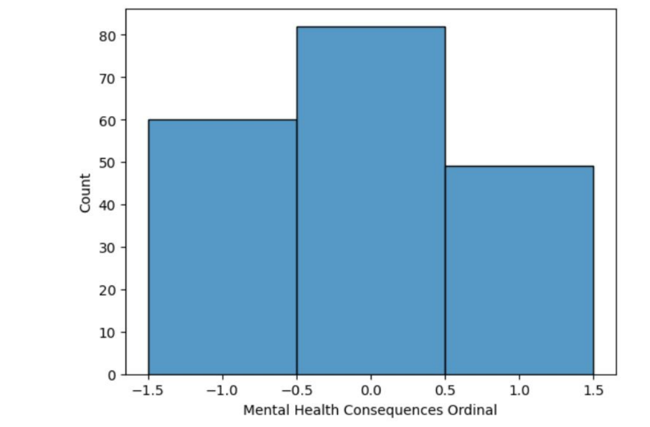
Analyzing Mental Health Consequences and Gender



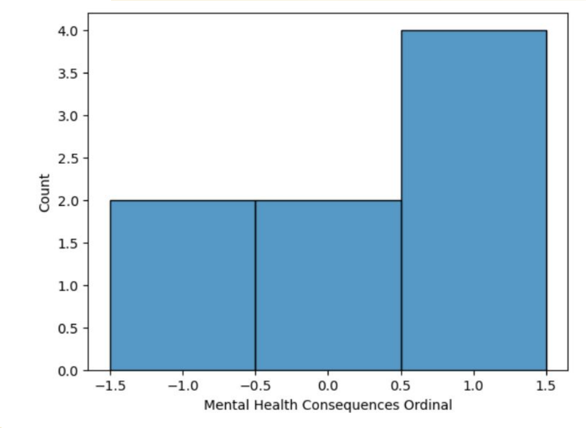
Second, we used a histogram to visually look at the responses from each separate gender:

The responses from male employees:

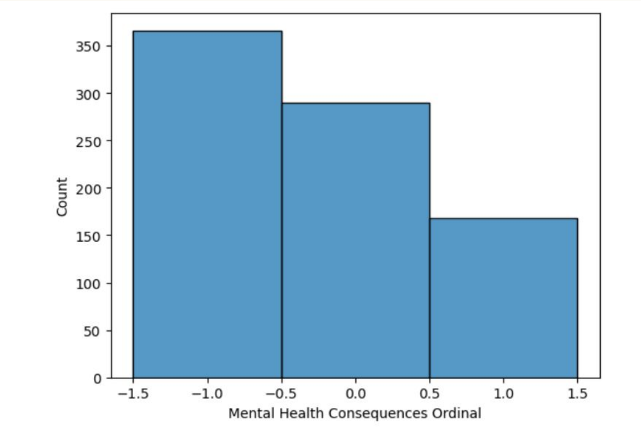
Analyzing Mental Health Consequences and Gender



Female employees



Non-binary employees



Male employees

As you can see, male employees face the least discomfort with sharing mental health issues compared to female and non-binary employees.

Analyzing Participants Who Have Sought Treatment and Gender

Participants were asked the question, "Have you sought treatment for a mental health condition?"

Participants answered this question with a ‘yes’ or ‘no’ response.

These responses were recorded with a specific ordinal number with using 1 for ‘yes’ and 0 for ‘no’. This was done with the intention to facilitate data collection.

3.

Analyzing Participants Who Have Sought Treatment and Gender

Figure 1

This figure counts the number of yes or no responses on whether or not they sought treatment depending on gender.

```
In [46]: sns.countplot(data = sought_treatment, x = "Sought Treatment", hue = "Gender")

Out[46]: <AxesSubplot:xlabel='Sought Treatment', ylabel='count'>
```

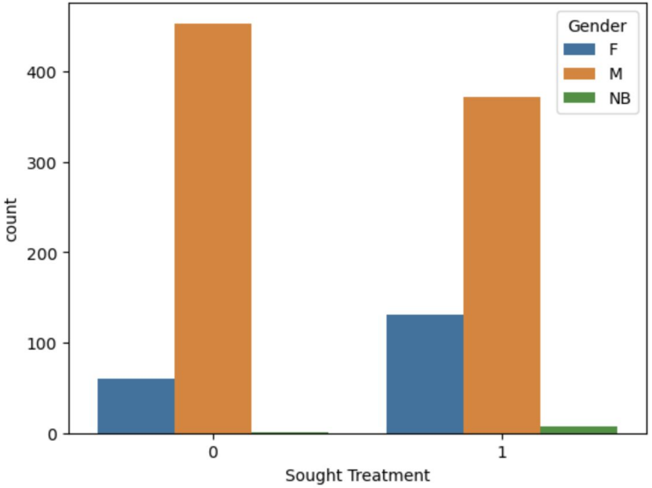


Figure 1:
A count box was plotted that demonstrated how men are more likely to seek treatment but this graph does not take into account the proportions. Since it does not take into account the proportions the number of men who sought treatment is much higher compared to the female and nonbinary people.

Analyzing Participants Who Have Sought Treatment and Gender

Figure 2

This figure also comapres the yes or no responses to the same survey question but takes into account the proportion of genders.

In [49]:
```
plot_proportion = sns.barplot(data=treatment, x=treatment['Sought Treatment'], y=treatment['prop'], hue=treatment['Gender'])
plot_proportion
```

Out[49]: `<AxesSubplot:xlabel='Sought Treatment', ylabel='prop'>`

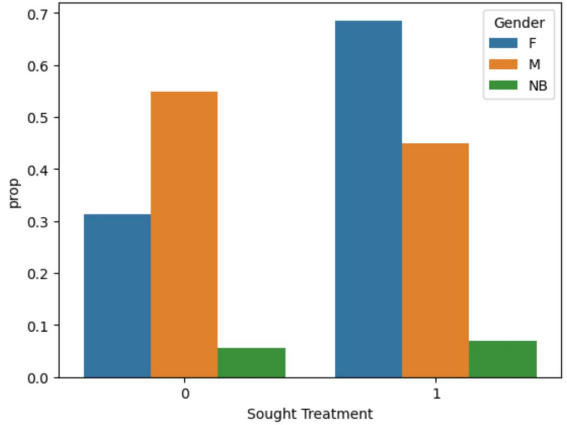


Figure 2:

Taking into consideration of the proportions for genders, figure 2 demonstrates how female and nonbinary identifying people often seek more mental health treatment despite the number of female/nonbinary employees being so low compared to the male employee.

3.

Analyzing Participants Who Have Sought Treatment and Gender

The graphs demonstrated a clear understanding that participants' who identified as female or nonbinary were more likely to seek mental health treatment compared to the male participants. We can say that women and nonbinary people are more likely in seeking treatment for mental health issues compared to male workers.

Conclusion

Through all three topics of observation, we are able to conclude that males and non-males within the tech industry have significant differences in experience regarding mental health in the workplace. To summarize what we found from our data analysis:

	Male Employees	Female and Non-binary Employees
Seeking mental health treatment	Less likely	More likely
Experiencing work interference from their mental health	Less likely	More likely
Comfortability discussing mental health issues with their employers	More likely	Less likely

Conclusion

Correlation vs. Causation:

- There is a clear positive correlation between greater mental health struggles and non-male gender identification, although it is unclear what the causes of such outcomes are.
- Our question focused primarily on the outcomes (employees’ attitudes towards mental health in the workplace), and not the causes for these outcomes
- Possible causes to look into in the future: whether employees felt discriminated against in the workplace due to their gender identity/expression, whether employees felt that their peers treated them differently based on their gender, etc.

Overall, understanding and observing patterns of mental health across different employees within the tech field is an imperative first step to addressing any and all prejudice; as such, this field of research is crucial in supporting employees' mental wellness.