

OSMI Mental Health in Tech Survey

Project Report

Group 4

Authors:

Jay Panchal, Julie Bosworth, Nidhi Davawala, Shezan Rohinton Mirzan, Shirley Xu

ABSTRACT

This project aims to understand how mental health is viewed in the Tech industry, the spread of mental disorders among the employees and to gauge the system present to tackle these conditions. D3 based self-read interactive website depicting the survey results of 3 different years is implemented to gain a multi-faceted view of the data. These designs are beneficial to the current employees of the Tech industry, the companies themselves and the external organizations that work for the cause of mental well being in the IT industry. Insights gained through these visualizations serve as a feedback to the entire system and thus can be highly beneficial to all the key stakeholders.

1 Introduction

In the present times, the IT industry is responsible for providing employment to a vast population of the United States. Mental well being of the employees of such a large sector is of utmost importance. To understand the current scenario of mental health in the IT industry, we have referred to the OSMI Mental Health Surveys of Tech Industry. On visualizing these survey responses and drawing meaningful inferences, the stakeholders could gauge how the entire support system for mental health is performing and whether it can be improved further.

1.1 Data & problem selection

The Open Sourcing Mental Illness (OSMI) community runs large scale surveys in the IT/Tech industry to gain an understanding of how mental health disorders are viewed in the industry, what provisions and resources are made available by the employer and how effective the entire system is. Each of the surveys consists of about 1200 responses from the employees of the Tech industry that would help the users understand the current employees opinions about how mental health conditions affect the workplace environment.

We have used survey results for the years 2014, 2016 and 2019. This forms an extensive base for us to understand the current statistics of the Tech industry and find corresponding trends and correlations too. The data was initially pre-processed to make it adaptable for the visualizations. The visualizations are aimed to be interactive and provide multiple views of our survey data. Being in the field of Computer Science, this problem motivated us to understand how mental health is viewed in the IT industry and understand the trends changing over years.

1.2 Intended Users

Users Identification:

1. **Employees of the Tech Companies:** Present employees and staff of the IT Companies participating in the survey can understand the policies and resources provided by their employer.
2. **Mental Health Volunteers:** Unions and groups working for the cause of mental health could understand the present scenario in the IT industry and suggest changes to improve the conditions.
3. **Company's Human Resources Team:** The company could understand their employees' opinions, gauge and improve their existing system for supporting a healthy workplace environment.

Task Identification:

1. Check for correlation between the company's support for helping those with mental health conditions and workers' willingness to reach out for help.
2. Understand correlation between anonymity and consequences of discussing mental disorders as perceived by the employees.
3. Visualize the state-wise distribution of mental disorders.
4. Check for correlation between company size and employee's willingness to reach out for help.
5. Study the impact family history of mental health conditions has on employee's willingness to reach out for help
6. Check for correlation between company size and it's provisions for employee mental health care.
7. Identify parameters that help potential employees make a decision on which company to join.

1.3 Related Work

Some of the recent implementations show visualizations that relate to mental health. They tend to focus on mental health in general. The first one looks at mental on a global level to show how common different mental disorders are as well as what countries have the most cases of depression[1]. The second one focuses on the prevalence of mental disorders by different months[2]. However the third visualization is about how mental health and physical health are correlated[3]. They tend to use barcharts, stacked bar charts, and a geomap in the visualizations.

Although all of them focus on mental health, our visualizations are mostly based on mental health considerations in the Tech Industries primarily based in the United States. Apart from just providing the basic demographic view, our website enables users to gain deeper insights into the mental health state by providing views around employee mental state while discussing their mental health problems and factors affecting the work environment for people with mental health issues. The interactive website not only allows viewers to play around and filter different data settings, the site posits conclusions and key-findings related to each view which makes it easier for the viewer to get quick insights.

1.4 Data and task abstraction

Data Abstraction:

- **Field 1**
 - Name: mental_health_consequence
 - Meaning: If discussing a mental health disorder with the employer had negative consequences?
 - Type: Categorical
 - Cardinality/Range: 3
- **Field 2**
 - Name: physical_health_consequence
 - Meaning: Check if there are any negative consequences if employee discusses about mental health with their employer
 - Type: Categorical
 - Cardinality/Range: 2
- **Field 3**
 - Name: benefits
 - Meaning: Whether mental health benefits are included in the employer's health coverage plan.
 - Type: Categorical
 - Cardinality/Range: 2
- **Field 4**
 - Name: family_history
 - Meaning: Check if the employee had any family history of mental illness
 - Type: Categorical
 - Cardinality/Range: 2
- **Field 5**
 - Name: anonymity
 - Meaning: Check if the employee's anonymity is protected if they use employer mental health benefits
 - Type: Categorical
 - Cardinality/Range: 2
- **Field 6**
 - Name: state
 - Meaning: State to which the employee belongs to

- Type: Categorical
- Cardinality/Range: all the states in the USA
- **Field 7**
 - Name: no_employees
 - Meaning: Size group to which the company belongs to
 - Type: Categorical
 - Cardinality/Range: 6, the size groups into which company sizes are divided into
- **Field 8**
 - Name: leave
 - Meaning: Check if a sick leave would be available for mental health issue
 - Type: Categorical
 - Cardinality/Range: 3

For a full list of data abstraction, [click here](#)

Task Abstraction:

1. [**Which states employees suffer the most due to mental health conditions and which disorders are the major contributors?**](#)
Fields/attributes: state, condition
{Discover correlation, Explore distribution, Compare Values}
2. [**What are workers who do have a family history of mental health conditions more or less likely to do?**](#)
Fields/attributes: family_history, seek_help, mental_health_interview, phys_health_interview
{Discover correlation, Explore distribution, Compare Values}
3. [**How does anonymity affect whether or not people perceive there to be consequences?**](#)
Fields/attributes: anonymity, mental_health_consequence, physical_health_consequence
{Discover Correlation, Locate Attributes, Summarize Dependency}
4. [**Is there a correlation between the company's support for helping those with mental health conditions and workers' willingness to reach out for help?**](#)
Fields/attributes: negative_consequences_employer, willing_to_discuss
{Discover Correlation, Locate Attributes, Compare Values}
5. [**Are smaller or larger companies more likely to provide benefits to people with mental health conditions?**](#)
Fields/attributes: no_employees, benefits
{Discover Correlation, Locate Attributes, Compare similarity}
6. [**What features of a company should someone look at to predict whether they will be comfortable when asking for mental health help?**](#)
Fields/attributes: mental_health_consequence, phys_health_consequence, leave, anonymity, seek_help, wellness_program, care_options, benefits
{Derive dependency, Explore Distribution, Identify Features}

1.5 General Data Statistics

Following figures[1a and 1b] provide a brief description of the data. The data comes from three years, 2014, 2016 and 2019 with the 2019 data being a bit smaller. The average number of respondents being approximately ~1000. The data was mostly dominated by males and with most respondents falling in the range of 19-40 years. However, all these data is skewed across different age groups, genders, state-wise respondents etc.

The distribution of mental disorders is also skewed with predominantly people having Mood, Anxiety and ADH Disorders.

		2014	2016	2019
Gender Distribution	Male	993 (~78%)	1061 (~74%)	516 (~67%)
	Female	248 (~20%)	343 (~24%)	200 (~28%)
	Other	18 (~2%)	29 (~2%)	36 (~5%)
Age Groups Distribution	0-18	8	2	0
	19-30	572	541	121
	31-40	531	625	139
	40+	148	267	92
People with mental illness		-	40.1%	41.7%
Total Respondents		1259	1433	752

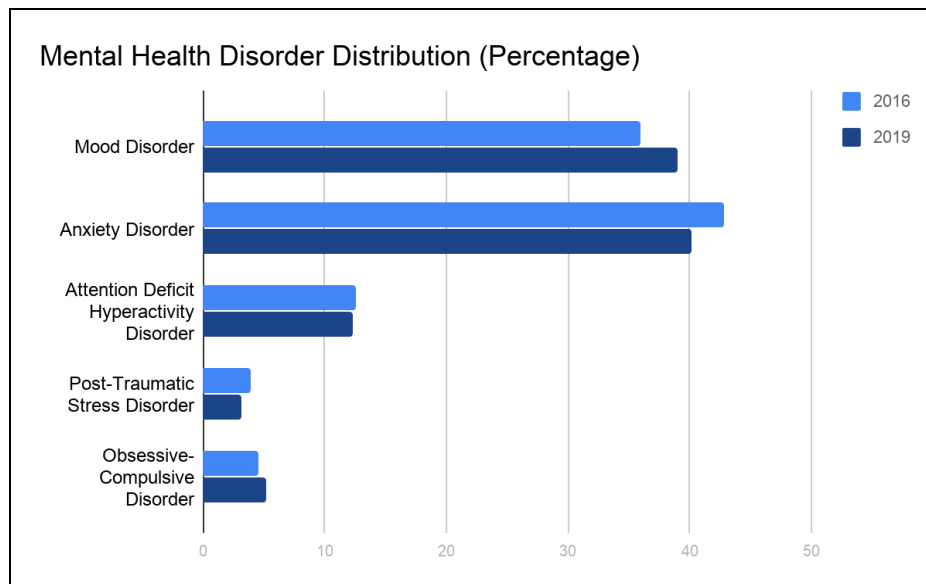


Figure 1a) General data statistics describing total respondents, their age and gender distribution
1b) Disorder-wise distribution of major mental health disorders in the survey

1.6 Design Overview/Solution

Our self-read interactive website will contain multiple interactive views that will help the users to visualize the survey responses over the years and also gain insights from the data. The overall design is built to get customizable results based on the filters like the category of diseases, age/gender of the respondents as well as the year of the survey. This will help the user get a dynamic view of the entire survey and thus draw meaningful inferences from it. The following sections describe the different views incorporated into our website. The views have been accompanied with the questions that it'll answer, a demo prototype and the interactivity that will be included in the final designs.

The website is an easy 10 minute read with user-friendly interactions to enable users to gain different insights pertaining to the data. The website is completely modularised by keeping similar views near each other to avoid a lot of navigation by the user and to give the effect of a multi-view experience. The whole website is primarily divided into four modules:

1. **Introduction:** This part clearly introduces our problem and motivation along with shedding light on the OSMI Mental Health Dataset. The module also contains a sub-module which talks about the general data statistics derived from the surveys.
2. **Views:** All visualizations are categorized into 3 main views that would throw light on different aspects of the dataset and would help users navigate through related visualizations with ease. To give context to the user, each view is accompanied with a brief description of what the visualization does, its intended users and the key inferences deduced from the design.
 - 2.1. **Demographics View:**
 - 2.1.1. **View 1:** State wise distribution of disorder
 - 2.1.2. **View 2:** Illness Family History vs Willingness to discuss with peers
 - 2.2. **Employee mental state while discussing mental health:**
 - 2.2.1. **View 3:** Anonymity vs perceived consequences
 - 2.2.2. **View 4:** Willingness vs Company Benefits
 - 2.3. **Factors affecting work environment for people with mental health issues**
 - 2.3.1. **View 5:** Company Size vs Benefits provided
 - 2.3.2. **View 6:** Factors vs. Perception of Consequences
3. **Challenges and Issues:** Being the survey responses of multiple years, the dataset for each year had small variations which needed to be serialized. This section describes the different challenges faced and the pre-processing steps taken to make the datasets more usable.
4. **Conclusion:** This section summarizes all the key takeaways derived from each of the visualizations and provides suggestive measures to mitigate the detrimental effects of having mental health issues in the tech-industry.

1.7 Implementation Details

Tools:

1. **Data pre-processing:** Python libraries (Pandas, Numpy)
2. **Visualization Generation:** D3
3. **Vector Graphics:** SVG
4. **Website Creation:** Bootstrap/CSS

Data Pre-processing Steps:

1. **Missing Data:** The data was presented in '.csv' format and seemed to have consisted of many missing values. Wherever required, some of the data was extrapolated by either the mean/median value of that particular row. For some visualization, where the data couldn't be extrapolated, the rows were dropped keeping in mind that this wouldn't cause any low-data bias.
2. **Free-filled Data:** The data was collected without predefined possible field values for each column. This has led to many different field values for something that represents the same original value. For instance, in 2016 data, there are 70 unique genders. In such scenarios, the genders were grouped while respecting the participants' answers. The columns having custom responses were dropped because those are difficult to analyze without going through each response manually.
3. **Missing Columns:** Some important columns were missing altogether in a particular year's dataset. For instance, the type of disorder was completely missing in 2014's data. This has led us to completely ignore this data in particular visualization for trend analysis. In the 2019 data, there was no column for perception of negative consequences, so instead, although different but a similar column was used to infer the data by taking the inverse of the responses for whether participants would be willing to talk to employers about mental health issues.
4. **Data Skewness:** The Data isn't well distributed across all age groups, states and company sizes. This creates some kind of biases in the visualizations. To avoid the same, percentages are reported instead of absolute counts to compare data across categories. However, some biases are inevitable.

1.8 Final Product Sample

Demographic Views

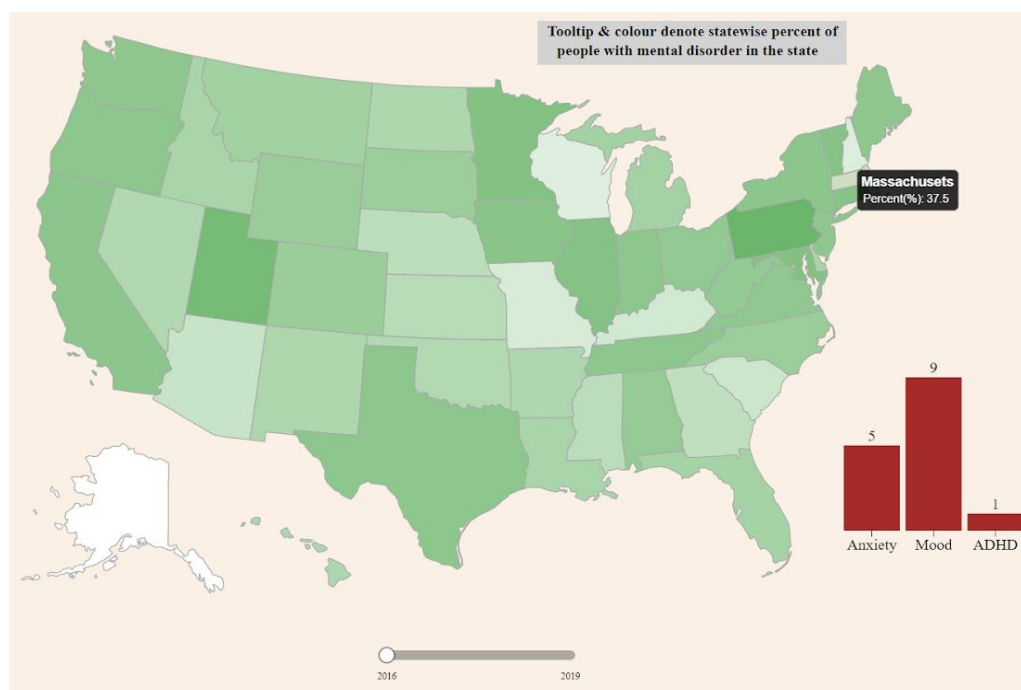
View 1: State wise distribution of disorders

This visualization shows the statewise demographic spread of the top 3 mental disorders faced by the employees of the IT industry. Mood, anxiety and attention deficit hyperactivity disorders are the most commonly faced illnesses among the employees and a bar chart shows the number of cases of these disorders for each state. The tooltip shows the percentage of people suffering from mental illness in that state and year-wise spread can be viewed using the timeline slider.

The current employees of the industry could use this design to check for trends in each state which might be useful for future job perspectives. The state mental health volunteers could use this design to gauge the situation of mental well being in their state and suggest required changes.

Fields/attribute: state, condition

Demo:



Justification: The choropleth map would be effective for this as the users could easily navigate over the state of their interest and check for the number of people affected with mental disorders. Moreover, a sequential colormap would help understand the intensity of diseases in each of the states.

Interactivity:

- **Timeline:** Data for different years could be seen by changing the slider
- **Barchart:** Hovering over each state shows the number of cases of key mental disorders
- **Tooltip:** Shows percentage of people suffering from mental illness in the hovered over state

Insights:

- Tech hubs like California, Washington and New York have the largest section of employees with mental disorders, which shows the overall stressful nature of the tech-industry
- Mood Disorder and Anxiety disorders are the most commonly faced mental disorders within all states over the years

Suggestion: Mental health volunteers working for this cause can address the issue of mood disorder, and suggest solutions to the companies by analyzing viable options to mitigate its effects

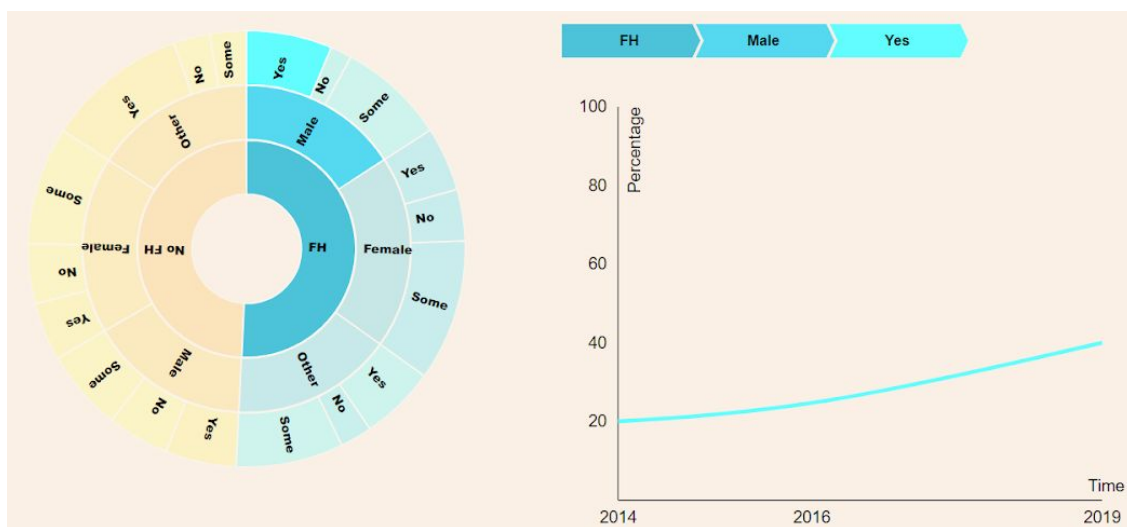
View 2: Impact of family history of mental health conditions on employee's willingness to reach out for help

This visualization shows how having a family history of mental illness affects the employee's willingness to discuss their mental health situation with their peers. Gender based division is also made to better understand the changing trends within the genders. Sunburst diagram is an efficient way to show the inherent hierarchy of this design, and each response in the outer circle is accompanied with the trends of that response changing over 3 years.

The company representatives could refer to this visualization to gain a better understanding of the reservations faced by the employees to open up about their problems. The mental health volunteers could also use the results to analyze the effects of family history and suggest ways for employees to open up about their problems.

Fields/attributes: family_history, seek_help, mental_health_interview, phys_health_interview

Demo:



Justification: There exists an inherent hierarchy in this view. With the sunburst diagram, the user could easily navigate through different levels of hierarchy. Responses are as organized via the levels of: Presence/Absence of family history of disorders, the gender of the responder within that category and at the final level it answers the question of whether the employee would feel comfortable discussing their mental disorders with their peers. Sunburst is one of the most optimal ways to show this hierarchy and it provides ease of access to the user.

Interactivity: Hovering over responses generates trends of that response over 3 years.

Insights:

- Over the years, the tendency to open up about discussing mental illness has increased.
- People with family history of mental illness are more willing to reach out for help
- Employees without a history of illness show more reservations in opening up about it
- Men were found to be more willing to discuss their mental health
- Over years, females are now more willing to discuss their mental health

Suggestion: Companies can organize counselling sessions, focused especially on the employees without family history of illness so that they feel comfortable discussing their issues and finding the solutions.

Employee mental state while discussing mental health

View 3: Correlation between anonymity and perceived consequences i.e. if protection of the identity of the person with a condition affects how he/she thinks about the consequences of discussing mental and physical health.

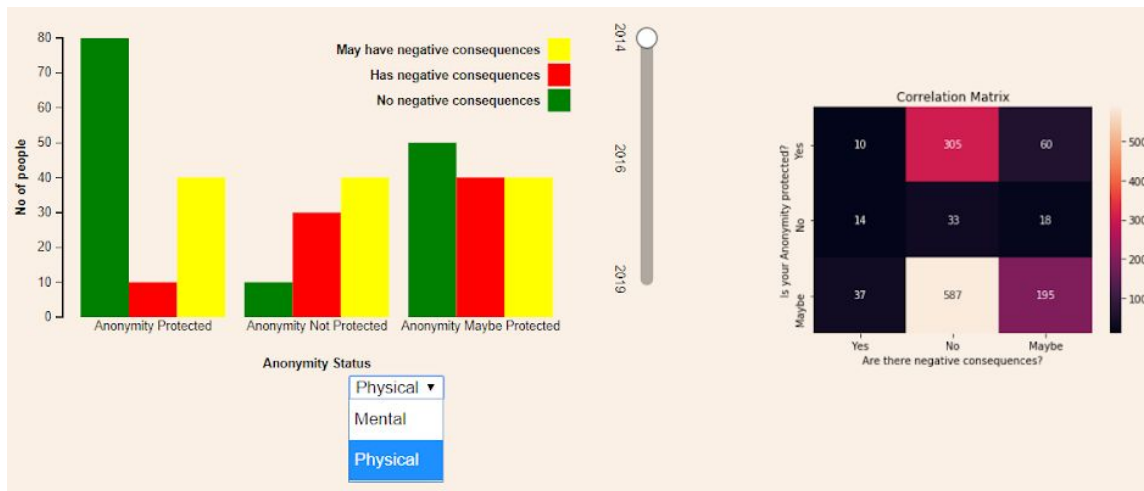
Protection of anonymity plays an important role in the way employees perceive there to be negative consequences attached to revealing their mental health issues. This correlation can be understood from the visualization given below. The bar chart shows the employee's perceived consequences against the status of anonymity protection in their respective companies. The heat map associated with each bar gives a better understanding about these correlations. The drop down helps the user navigate between the perceived consequences of the both mental and physical disorders.

This visualization can be used by the company representative/HR teams to modify or formulate their anonymity policies that best suit their employees comforts in opening up about their mental illnesses.

Fields/attributes: anonymity, mental_health_consequence, physical_health_consequence

Justification: We want to observe if protecting anonymity affects how people think of the consequences of discussing mental and physical health. This bar chart would show clear comparison between the responses of the employees and help to see if there exists any correlation. The plot also concisely shows the results of both mental and physical health disorders. Adjacent bars also make it easy for comparisons between as well as within categories. Heatmap adds to the user readability and makes it easy to understand the correlation.

Demo



Interactivity:

- **Timeline:** Data for different years could be seen by changing the slider
- **Dropdown:** Choose between the type of disorder to view their perceived consequences
- **Heatmap:** All bar charts accompanied with heatmaps for better visualization

Insights:

- Protection of anonymity plays a big role in making employees comfortable to discuss their illness.
- Employees with anonymity protected/mostly protected are confident that there won't/might not be negative consequences.
- Discussing mental illness would have more negative consequences than discussing physical issues

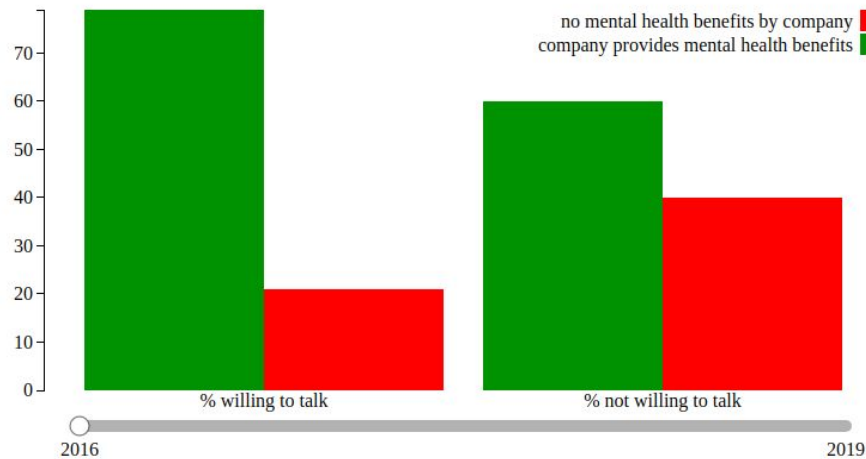
Suggestion: Companies could formulate policies that maintain employee anonymity to encourage a healthy workplace environment open to discussing mental and physical illness.

View 4: Is there a correlation between the company's support for helping those with mental health conditions and workers' willingness to reach out for help?

Employees suffering from mental health disorders might shy away from discussing it if they feel it is not professional to talk about it at the workplace. So it is required that their employers provide a formal stage for them to discuss their mental health disorders in order to function up to their potential. Companies often wish to provide benefits to hire and retain their employees and hence might be interested in knowing if they can help their employees tackle any problems they face at the workplace including their health. This is done through health care packages which cover employee wellbeing. However, these packages often ignore mental health disorders, which in fact can be a real struggle for the ones affected by it. Studying the data we have can provide us insights about whether including mental health disorder related benefits can help employees feel comfortable at their workplace. Such insights can also help employees decide about their next workplace if they wish to move to their next jobs.

Fields/attributes: benefits, willing_to_discuss

Demo:



Justification: The bar chart would be a suitable chart for visualizing if there is any correlation in employees being provided mental health benefits and them being willing to talk about it with their supervisors. Side-by-side presence of the bars in this chart allows easy comparison of these numbers across the categories as well as in the sub-categories.

Interactivity:

1. **Timeline:** Data for different years could be seen by changing the slider
2. **Barchart:** Hovering over each bar shows the percentage of employees
3. **Legend:** Hovering on legend helps highlight the corresponding category

Key Insights:

- Employees tend to be willing to discuss their mental health issues more often when their companies provide mental health support in the health care package.
- Increase in the percentage of companies offering mental health support from year 2016 to 2019. This shows that more companies are becoming aware about providing mental health care benefits.

Suggestion: The visualization enables us to understand the fact that employees becoming more aware and discussing it under professional scenarios needs companies helping them through mental health benefits and setting up a stage for them to discuss their mental health issues.

Factors affecting work environment for people with mental health issues

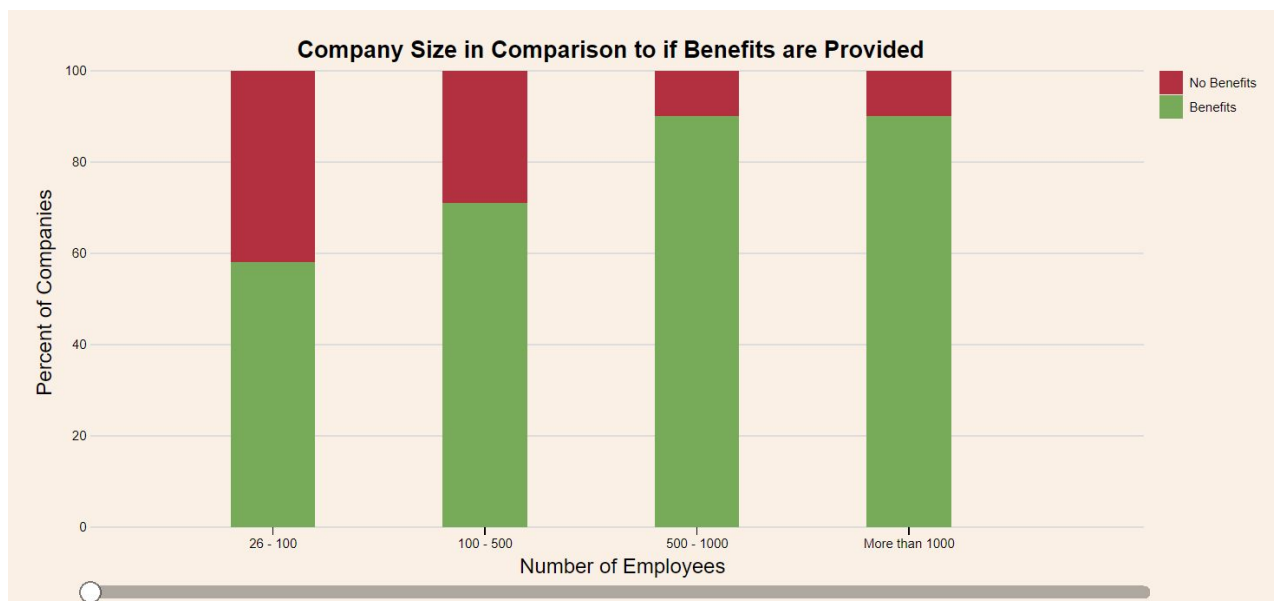
View 5: Are smaller or larger companies more likely to provide benefits to people with mental health conditions?

This visualization focuses on how company size plays an impact on if mental health benefits are provided. In particular, the data focuses on four group sizes: 26 - 100, 100 – 500, 500 – 1000, and more than 1000. From the data we gathered if there were any or no benefits provided. The visualization could help people looking for work in tech realize how in recent years all companies do a good job at providing benefits, but large companies are reliable when it comes to providing benefits.

Fields/attributes: no_employees, benefits

Intended Users: The intended users are people who are interested in going into tech who would like benefits for their mental health

Demo:



Justification: The stacked bar chart would be effective for this question because it enables for a user to quickly see how many companies are providing benefits based on the size, while keeping the chart from looking too cluttered. The color scheme also makes it so that a user will quickly be able to figure out what is “no” and what is “yes” because no is commonly correlated with red and yes is commonly correlated with green.

Interactivity: When a section of a bar in the bar chart is hovered over, it displays what percent of the bar the section makes up.

Key Inferences/Insights:

- Larger companies are more willing to provide mental health benefits on average up until 2019
- Over time more companies have begun to provide benefits regardless of size, hinting that companies are realizing that mental health is important

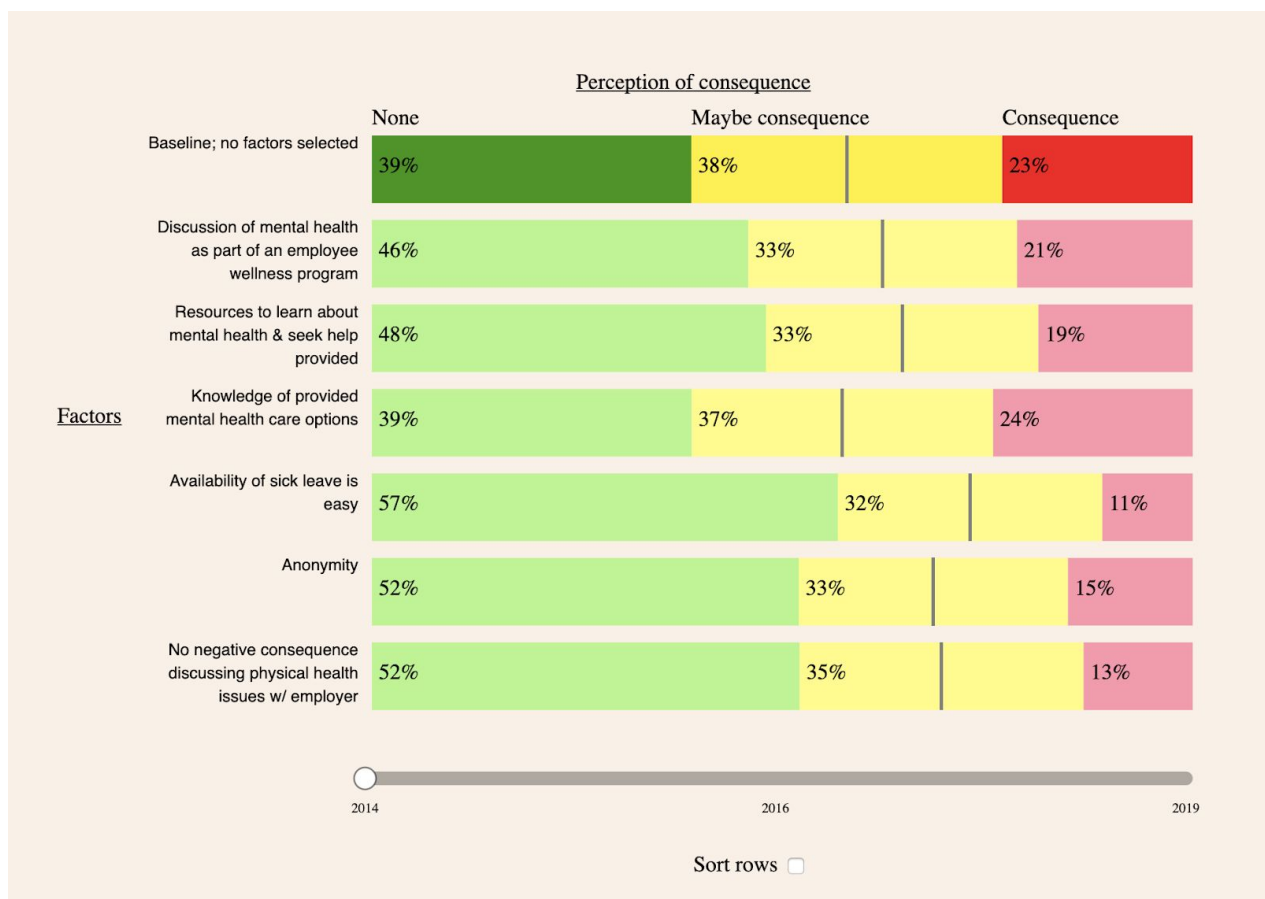
Suggestions: Smaller companies should provide equal amounts of support as larger companies and attempt to catch up to the larger companies. Even for the larger companies, about 10% of them still do not provide the required support. Large companies tend to have enough resources to support mental wellbeing programs and thus should consider implementing them in their structure.

View 6: What features of a company should someone look at to predict whether they will be comfortable when asking for mental health help?

This visualization shows the effects of factors on perception of negative consequences discussing mental health

Fields/attributes: mental_health_consequence, phys_health_consequence, leave, anonymity, seek_help, wellness_program, care_options, benefits

Demo:



Justification: The stacked bar chart breaks down each subset of responses, filtered by each factor, by the percentage of replies that are yes, maybe, or no. The black line drawn in the middle of each bar represents either the “yes” score or the “no” score plus half the “maybe” score, which gives the center of the data; the further right it is, the higher the number of “yes” responses, and thus the more the factor corresponds with a response of “yes” in regards to whether an individual is willing to talk about mental health. This allows various factors to be compared to determine

Interactivity:

1. **Timeline:** Data for different years could be seen by changing the slider
2. **Checkbox:** Allows for data to be sorted by impact of factors
3. **Focusing:** When a category is highlighted, the other categories are faded out for legibility

Key Inferences/Insights:

- The most impactful feature on perception of consequence was the availability of sick leave being easy, as it tended to have the 1st or 2nd highest correlation with a "none" response out of all factors.
- While resources and knowledge of mental health care options decreased perceptions of consequences, they were not as impactful as direct company policies, such as sick leave and anonymity.
- There is a positive correlation between no perception of negative consequences with mental and with physical health, suggesting that workplaces that care about physical health also tend to care more about mental health too.

Suggestions: Companies can make their employees feel more comfortable talking about mental health without fearing consequences if they add policies that help protect workers, such as anonymity and availability of sick leave, rather than just educational resources about mental health

Conclusion

Mental wellbeing is a pressing issue, especially in a high-pressure setting of the IT industry in the USA. Using the OSMI surveys, the trends and practices of the tech industry for the past 3 years were revealed. These results form a solid base for analyzing the entire system and adopting them while finding solutions for inducing a healthy and welcoming workplace environment. Opening up about the mental conditions is the first step towards mitigating the effects of the illness and thus all the stakeholders must act in a way to create an industry supportive of these conditions. Following are some suggestions based on the results of studying the visualization that could be adopted:

1. Mental Health communities and volunteers could address issues of mood and anxiety disorders and come up with solutions
2. Tech companies could do their bit by organize counselling session for their employees encouraging them to open up their mental conditions
3. Considering the mass employee opinions, the companies could add policies to protect workers such as anonymity and sick leave
4. Smaller companies should provide more mental health benefits and packages to their employees
5. The larger companies still not working towards the cause of mental well being should incorporate policies to support their employees
6. In general, all the stakeholder need to set up a stage for employees to talk about mental health issues

REFERENCES

1. <http://www.humanosphere.org/global-health/2013/06/visualize-mental-illness/>
2. <https://www.einfochips.com/blog/how-data-analytics-visualization-can-aid-in-mental-healthcare/>
3. <https://janezhang.ca/posts/documenting-mental-health-data/>
4. B. Saket, A. Endert and Ç. Demiralp, "Task-Based Effectiveness of Basic Visualizations," in *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 7, pp. 2505-2512, 1 July 2019.
5. Visualization Analysis and Design Tamara Munzner 2014
6. <https://github.com/d3/d3/wiki/Tutorials>
7. <https://www.kaggle.com/osmi/mental-health-in-tech-survey>