Analysis of Mental Health in the Tech Industry

Abstract

The alarming mental health disorder statistics among adults inspired us to investigate deeper into the issue and produce data-backed insights that are actionable and would help industries adopt the recommendations. We consider the mental health dataset for tech companies where people respond to questions on potential factors contributing to mental illness.

Initial exploratory data analysis produces trends and correlations, which are further tested by Poisson and logistic models. Firstly, we look to understand the role Geography plays on the mental health, perceived negative sentiments about the workplace, and an employer's coverage policy on providing mental health benefits. It is observed through a Poisson regression model that the expected number of mental illnesses is greater if individuals experience negative sentiment regarding mental health in the workplace as opposed to having no negative experiences. It is also observed using another Poisson regression model that an employer providing coverage is associated with a greater number of employees seeking treatment than if the employer does not provide coverage. This providing coverage decision is even greater if the employee is from the United States through a statistically significant interaction term coefficient between providing coverage and location in the United States. Out of all the countries included in the survey, the United States has a greater expected number of mental illnesses and greater expected number of individuals seeking treatment for mental health.

With so many features in the dataset, a worthwhile problem to look at is the predominant factors that puts a tech professional at risk of compromised mental health. We build a logistic regression model to understand the relationship between factors and predict the current mental illness status. The model indicates a strong association of factors like past mental illness, Age, Employer's take on mental illness, and medical coverage with a tech professional's current mental health. Finally, we build a predictive model using the Random Forest algorithm that helps us predict whether a person is at risk for mental illness based on several factors. The variable importance chart from Random Forest models helps us see the extremely strong dependence of past mental illness on the risk. We also see that the perception of suffering a negative consequence

either from their current employers or their past employers plays a significant role in putting them at risk of the illness.

We conclude by recommending a few crucial actions. It is necessary to have healthy discussion around mental health in the workplace. Promoting a positive attitude about discussing mental health is very important, and its absence is associated with a greater presence of mental illness. Employers are highly recommended to offer their employees coverage for mental health visits, especially if the company is in the United States and if their employee has had a history of mental illness in the past.

Introduction and Problem Statement

Mental health is an increasingly important concern in today's world. It is estimated that 16.2 million adults in the United States, or 6.7 percent of American adults, have had at least one major depressive episode each year. Depression tends to affect people in their prime working years and is one of the top 3 workplace problems.

The purpose of our project is to tackle this problem by evaluating data related to prevalence of mental health disorders in the IT workplace and generate a profile for adults most prone to depression. This will help understand the risk factors that contribute to mental health disorders. The insights can also help guide company policies regarding mental health resources to protect their employees in the workplace. The following list comprise the research questions proposed:

- Investigate the impact of geographical location on the amount or the probability of seeking treatment for mental illnesses.
- Determine the effect of a company's mental health coverage policy on the amount of employees seeking treatment.
- Investigate the effect of negative workplace sentiment on the number of people seeking treatment.
- Determine whether a person with a history of mental illness is more prone to facing another mental health issue.
- Determine if an employee's gender influences whether they feel that mental health concerns can be brought up in their workplace.
- Determine factors that are significantly determine the probability of people seeking treatment for mental illness.

The major challenge in the dataset is to handle categorical variables and the missing fields in the dataset in addition to a smaller number of data points. The report format starts off with a Data Collection and Pre-Processing section, followed by the methodology section, where 3 different approaches have been described. It is then followed by Analysis and Results, Conclusions and Lessons Learned. The appendix has the detailed results of our study.

Data Collection and Pre-Processing

The dataset has been obtained from “https://www.kaggle.com/osmi/mental-health-in-tech-2016". It has 1433 responses from a survey to measure attitudes towards mental health in the IT/tech workplace and examine the frequency of mental health disorders among tech workers. The dataset contains 63 columns, including categorical and text-based columns. Since it is typical of survey data to have a lot of missing values, data cleaning and feature engineering has been done. The data pre-processing and variables used vary between the research questions asked and models used to answer the research questions.

The data cleaning process related to the research question of analyzing the effect of negative sentiment involves engineering a new feature that combines information from multiple columns. Since there are many questions asked in the survey it could add complexity to treat every question related to negative sentiment separately, combining them reduces this complexity. The column is created by combining the 4 columns “Do you think discussing a mental health disorder with your employer would have negative consequences”, “Would you feel comfortable discussing a mental health disorder with your coworkers”, “Would you feel comfortable discussing a mental health disorder with you supervisor”, and “Do you feel that your employer takes mental health as seriously as physical health” into one column called ‘negativeSentimentWork’. If an individual answered any of the four questions with a response that affirms negative sentiment, then the new feature would be set to 1. If no negative sentiment is present in the questions, then the new feature is set to 0 for that individual. In creating this feature, it is required to exclude individuals who left these questions blank. This type of exclusion results in omitting 287 of the 1433 rows of data resulting in 1146 rows used for the workplace negative sentiment analysis.

In answering the research question regarding geographical location, it is necessary to only control for countries that have enough observations to gain meaningful insights from them. Figure 1 displays a bar chart showing the observation count of the top 11 countries. It is seen that the observations are significantly greater in the United States than in other countries and there are many countries with a small number of observations. A country will be individually controlled for if their observation count is 40 or more. This cutoff results in choosing the United States, the United Kingdom, Germany, the Netherlands, and Canada as the locations to control. These 5 locations are controlled for by creating binary variables for each country with 1 indicating the individual is from that country and 0 if the individual is not from that country. Each observation in the sample has a country distinction so no rows are omitted.

For questions regarding coverage policy, only Yes and No responses need to be included. Figure 2 shows a bar graph displaying the number of observations for each possible value for if the employer provides mental health benefits as part of healthcare coverage variable. The variable can take on values Yes, No, Not Eligible for Coverage/NA, I don’t know, and omitted (individuals left question blank). Only selecting individuals that answered Yes or No results in omitting 689 rows leaving 744 rows for the observation analysis.

Some of the research questions involve asking information related to knowing the expected number of people seeking treatment or the expected number of people with a mental health disorder diagnosed by a medical professional. The survey data in its current form is not initially set up to handle this kind of question so a transformation is required in order to create count data. This transformation can be performed by summing how many people have gone to seek treatment over the respective variables of interest. An example of the result of this data aggregation can be seen in Tables 1 and 2 which show aggregation for analyzing negative sentiment with country designation and for analyzing coverage policy with country designation.

To identify factors that are significant in explaining mental illness, it is necessary to reduce multiple categories of survey responses in the dataset to closely related fewer categories. Input responses for gender like "Male", "Female", "Woman", "Non-Binary", "Transgender" are merged to three categories: "M", "F", and "Other”. We merge text-based responses like 'No, I don't think it would', 'No, it has not' to a broader category 'No'.

Mice package is used to impute missing data values to avoid dropping of valuable data points.

Proposed Methodology

Poisson regression:

Since the research questions regarding geographic location, negative sentiment, and coverage policy are based on analyzing a count response variable, a Poisson regression model is proposed for a descriptive analysis of each factor of interest's impact on the count of people having a mental disorder or the count of people seeking treatment for a mental disorder. Since coverage policies and workplace dynamics can be highly impacted by the culture or government of a surrounding area, geographic location and either negative sentiment or coverage policy will be analyzed together. Since coverage policy and negative sentiment variables results in differing exclusion of rows, they will not be analyzed together as excessive omission of observations for one category might alter the results of the analysis. This set up leads to two proposed Poisson models: one for analyzing negative sentiment and country designation, and the second for analyzing coverage policy and country designation. Some dispersion is expected in both models but having dispersion does not invalidate the coefficient estimates of the variables based on a Poisson regression model property. However, having dispersion can distort the confidence levels used in hypothesis tests of coefficients' significance. As a result, hypothesis testing of coefficients' significance needs to be handled in a robust way. Multiple methods for handling dispersion exist such as correcting the Poisson model for dispersion followed by using Z-tests to carry out the hypothesis testing. An alternative method is using an F-test to check for each coefficients' significance. Since the F-test can be more robust over the dispersion corrected Poisson model, it will be used to evaluate each coefficients' significance.

Logistic Regression:

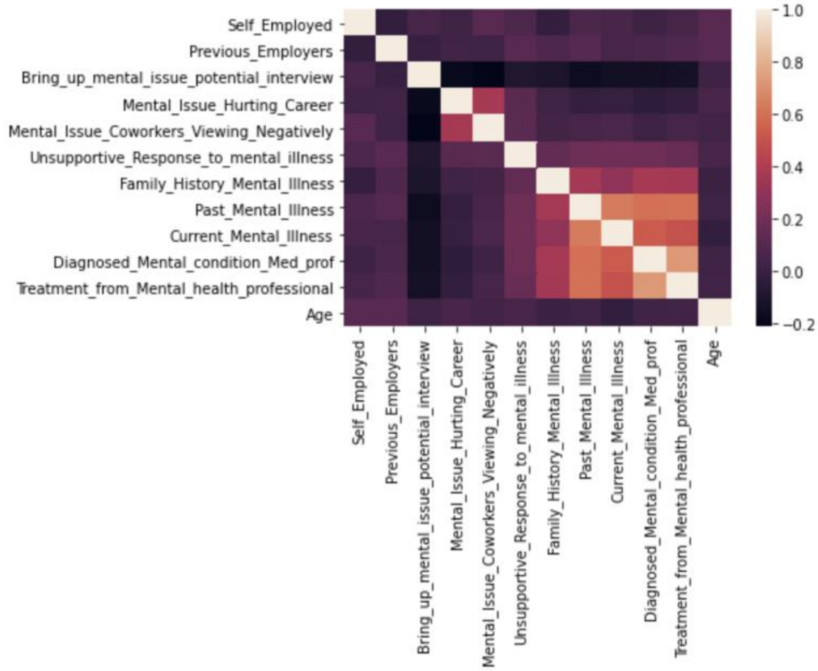
Logistic regression model helps us in classification type problems. It lets us understand the relationship between a dependent binary variable and other independent variables. We can clearly see which factors play a role in the predicting the dependent variable. Probability of current mental illness, and a tech professional's intent to seek remedy for a mental illness are two important responses in the survey that will help us answer our research questions on factors that lead to mental illness, history of illness causing recurrence, and employer's views on mental health well-being of their employees. Since these responses are categorical, a logistic regression model is best suited to predict the responses and identify significant factors contributing to them. We propose to build two logistic models. The first model predicts the probability of a tech professional's current mental state – if they have a mental illness or not. The second model predicts the probability of a person with mental illness seeking remedy for his/her health issues. We are specifically interested to see the impact of some factors like past mental illness, gender, age, employer's attitude towards mental health, medical coverage, etc on mental illness. Our aim is to identify factors that are significant / not-significant based on results from both the models to derive necessary conclusions. Hypothesis tests (p-value) of the coefficients is used to determine the significance of the factors. The coefficients of the factors will also be used to relate the correlation and level of impact on the outcome to make necessary conclusions.

Random Forest Model:

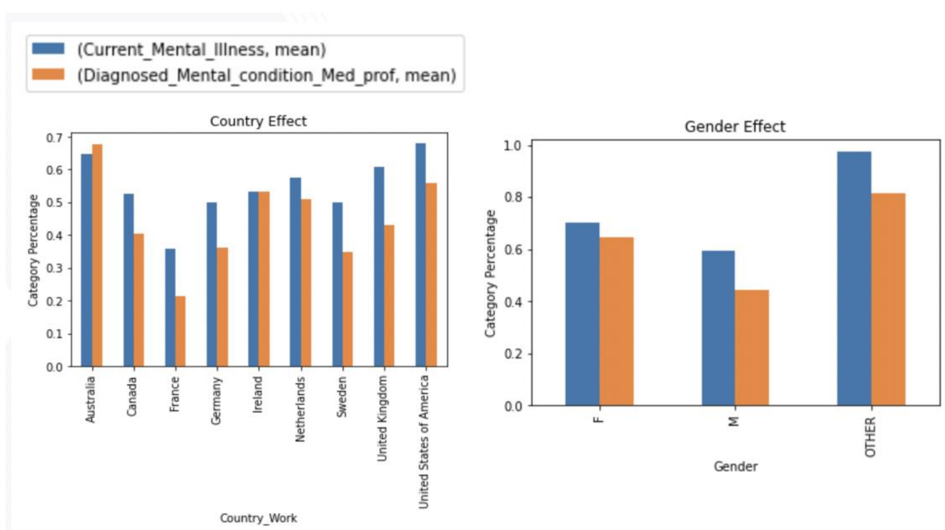
A random forest model is a collection of individual decision trees. As a result, the decision taken by the ensemble consisting of uncorrelated trees turn out to be better than any individual tree. Decision Trees and their extension Random Forests are robust and easy-to-interpret machine learning algorithms for both predicting the classification and regression tasks. In our project, we initially use all variables to see which factors have good predictive power. The model serves the dual purpose of understanding the importance of variables and predicting the risk of mental illness. The Variable Importance chart shows us a list of variables in decreasing order of their predictive power towards the built Random forest model. We then evaluate it using the Receiver Operating Characteristic (ROC) curve. It is a plot that can be used to determine the performance and robustness of a binary or multi-class classifier. The x-axis is the false positive rate (FPR) and the y-axis is the true positive rate (TPR). This ensures we are considering the accuracies of both classes while classifying.

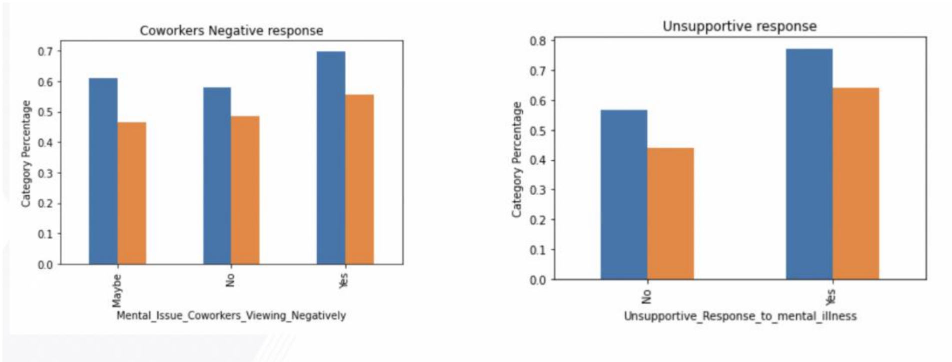
Analysis and Results

As a part of Exploratory Data Analysis (EDA), to observe the relation and dependency between the various columns of the data, we created a correlation matrix. In addition to that, we performed a rudimentary analysis on the dependency of "Current Mental Illness" and "Diagnosed Mental Condition" on other categorical columns which were not captured by the correlation matrix. From the correlation matrix, having a current illness is highly correlated with a past mental illness or family history of mental illnesses. It is a key observation and must be focused upon by tech companies to keep track of their employees and those that might be at a higher risk of facing a mental illness.



It is seen that USA and UK are the top 2 countries with highest percentage of tech employees reporting a current mental illness or having medically diagnosed with a mental illness. It is also seen employees who identify to non-binary gender face a much higher risk for facing mental illness. We can also see that negative sentiment in society, and more specifically in the workplace, appears to have higher percentages of mental illness.





It is crucial for tech companies, especially the ones in US and UK, to create a safe work environment, where employees are comfortable in discussing mental health with coworkers and their managers.

The analysis for the Poisson regression model on determining the relationship between country designation and negative workplace sentiment on the number of individuals diagnosed with a mental illness begins with fitting a full model with 6 variables including indicator variables for each of the top 5 countries and for negative workplace sentiment. The baseline case, represented by the intercept, captures the response of countries not comprising the top 5 countries with observations and observations that do not have negative sentiment in the workplace. Through an F-test it is observed that the coefficient for the UK has a high p-value at 0.646 so it is dropped from the model. A second model is fit without the UK indicator variable and each variable in the new model is determined significant with near zero p-values from the F-test. Since each coefficient is significant, this model is chosen as the final model to evaluate the impact of negative sentiment and country designation variables. The equation for the model is displayed in Equation 1 and a table showing the F-test values for each coefficient is displayed in table 3. Each of the positive coefficient values for the US and negative sentiment show that the expected number of mental illnesses for the US and negative sentiment are higher than the expected number of mental illnesses from the baseline response. On the other hand, the negative coefficient values for Canada, Netherlands, and Germany indicate that these countries have lower expected number of mental illnesses than the baseline response. As a result, it is observed that negative sentiment in the workplace is associated with higher counts of mental illnesses than if negative sentiment does not exist in the workplace. Also, the US is expected to have higher counts of mental illnesses over other countries included on the survey.

In determining the relationship between country designation and employer's coverage policy on the number of individuals seeking treatment for a mental illness, the full model is initially fit with 6 variables including indicator variables for each of the top 5 countries and an indicator variable for if the employer offers coverage. The baseline case captures the response of countries not comprising the top 5 countries with observations and observations that do not offer coverage for mental health benefits. It is found that the only variables with statistically significant coefficients are offering coverage and the US. This conclusion is determined through successive model fits with the coefficient with the least significant coefficient taken out through F-tests. The final model equation is shown in Equation 2 and the final p-values from the F-test is shown in Table 4. This analysis shows that the number of people seeking treatment in Canada, Germany, the Netherlands, and the United Kingdom are not significantly different than individuals in other countries other than the U.S. while coverage policy for mental health benefits is accounted for. On the other hand, the U.S. has significantly higher number of people seeking treatment than in other countries when offering coverage is accounted for. Offering coverage is seen to be associated with higher number of people seeking treatment than for individuals whose employer's do not provide coverage.

An interaction effect is tested between the US and coverage policy variables and it is shown to be statistically significant. The new model equation showing this interaction effect is displayed in Equation 3. The significance of this interaction term is that there is an added effect when individuals are from the United States and their employer provides mental health benefits as opposed to individuals from the United States without employer provided benefits. The coefficient value of this interaction term at 2.275 represents the amount to add to the coefficient of the United States variable at 1.18 to get the added effect on the expected number of people seeking treatment. Therefore, it is observed that offering coverage in the United States is highly associated with higher counts of individuals seeking treatment over individuals not having coverage in the United States.

Logistic Regression:

For the first logistic regression model, the response variable is "Current Mental Illness." This indicates the probability of a professional undergoing mental illness currently or not using "logit" link in the "glm" package in R. To understand the factors that significantly contribute to this, we initially regress all the predictors from the dataset (excluding country, since we are interested in the overall data and not-country-wise) against the response variable. We then perform stepwise selection to remove the insignificant predictors to identify the impact from the most significant factors. We consider a 95% confidence interval to estimate significance based on p-value (i.e factors with p-value < 0.05 are considered significant). One of the observations here is that "Gender" is not very influential in determining mental illness.

Regressing mental illness against the significant factors produced some interesting observations. The results of the model with the coefficients and its significance are in the Appendix (Table 6). Some interesting observations from this model is that "Past Mental Illness" is a significant factor in determining the current mental illness. Comparing the coefficients of the two categories of Past Mental Illness (Yes/No), we can see that a person with a past mental illness is more likely to have a mental health issue in the present (observe the coefficient of "No" is –2.135 and it increases to –0.527528 for a "Yes" category). "Age" is another factor that explains current mental illness. We can observe that "Age" is inversely correlated with a coefficient of –0.033, with a significant p-value. A person with a higher age is less likely to encounter a mental illness as compared to a younger adult. Another interesting observation is the significance of "Previous employers Weighing consequences of mental health." We see that the coefficients of employers weighing mental health seriously is lower than the employers where only some weigh mental health seriously. This indicates that employers that weigh mental concerns of their employees seriously reduce the chances of mental health issues among their employees. Refer Table 6: Logistic Model 1 Results in Appendix.

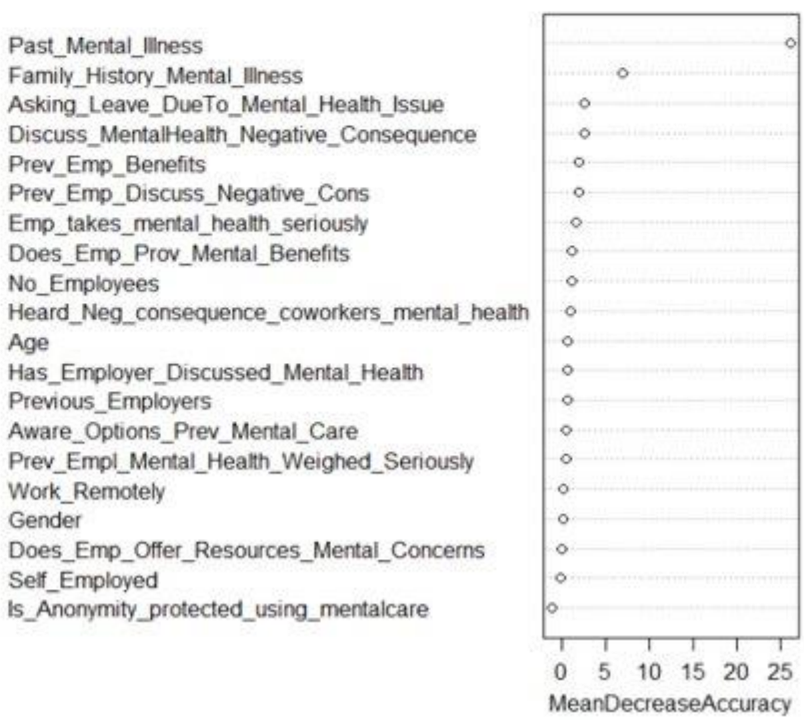
The second logistic regression model is built regressing "Employees seeking treatment for mental illness" against the available dataset. Again, stepwise selection is performed regression is performed to identify the most significant predictors and the results are analyzed. "Past Mental Illness" turns out to be a significant predictor in this case, again contributing to people seeking treatment for illness. Comparing their coefficients, people with a history of mental illness are more likely to seek treatment as compared to people without a prior history of illness (the odds increase from a factor of –1.027 to 0.811). Other observations from this model are that Employers that discuss mental health with their employees increase their chances of seeking mental health treatment. "Mental Health Coverage and it's awareness" seems to be significant and positively influences an employee in seeking treatment for mental health issues. Refer Table 7: Logistic

Model 2 Results in the Appendix. Overall, these factors can be observed most essential in mitigating risks leading to depression.

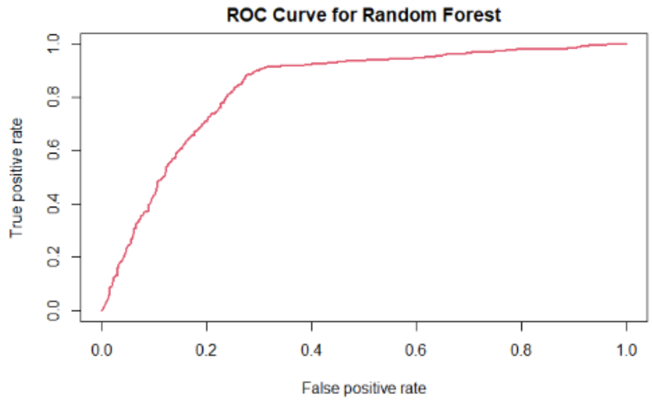
Random Forest:

Here, we seek to identify factors that increase the risk of Mental illness. In this case our dependent variable is positive when either the respondent is currently having a mental illness (reported by them) or are diagnosed for a mental illness by a mental health professional. The model essentially is an ensemble of several decision trees and handles both regression and classification. An advantage of using Random forest here is we get the list of variables in order of their importance based on their splits in the member trees.

<u>Model 1:</u>



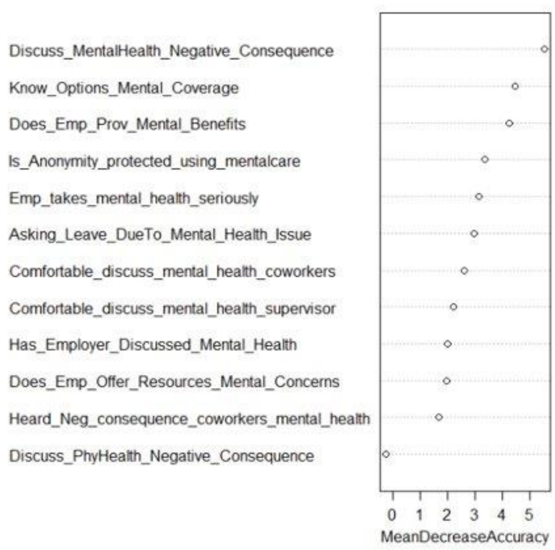
The chart above shows the variable importance of the predictors and we can conclude that 'Past Mental Illness' seems to be a very important factor – meaning, this variable highly predicts the risk of an individual. This is in agreement to the logistic regression results. The predictive model with the variables listed has an accuracy of ~87%.



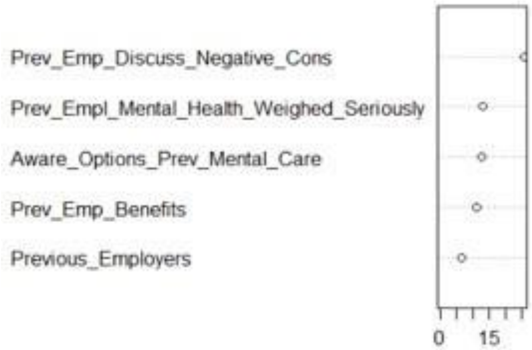
	Class 0	Class 1	Class Error
Class 0	337	113	25.11%
Class 1	74	909	7.52%

Sometimes the importance of other variables gets masked when one factor dominates in terms of the predictive power. To enable us to get more insights, the predictors are categorized into 3 categories and we build individual Random Forest models for each category that lets us have a closer look within the category.

<u>Model 2a: Variables pertaining to current employers:</u>

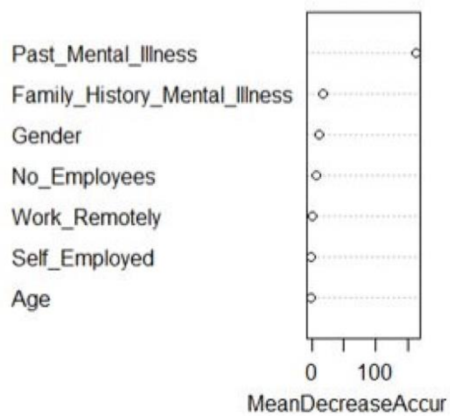


<u>Model 2b: Variables pertaining to previous employers:</u>



In both cases of previous and current employers, employees who perceive that a discussion around mental health could causes negative consequence on Mental Health are more likely to be affected.

<u>Model 2c: Variables pertaining to personal characteristics:</u>



Employees with past mental illnesses are more susceptible to have a mental illness. Overall, the top subsequent factors are variables belonging to a mix of different categories and hence highlights the importance of all categories.

Conclusions

Employers' discussion of mental health in the workplace and promoting a positive attitude about discussing mental health is very important, and its absence is associated with a greater presence of mental illnesses. Employers are highly recommended to offer their employees coverage for mental health visits, especially if the company is in the United States and if their employee has had a history of mental illness in the past.

In moving forward in the analysis, it is recommended to address some of the biases in the data. It is observed that most of the survey questions are answered by individuals working in the United States. Therefore, results and analyses could be swayed to the effect of individuals residing in the United States. It is recommended to obtain more observations of individuals in other countries to account for this bias and create a more balanced dataset. In addition, the survey can only be used in observing mental health in work environments in the tech industry. The findings from the analysis cannot be extrapolated to other lines of industry and work. It is recommended to carry out designed experiments in other professions to observe certain factors response on mental illnesses and compare the results to the analysis performed in the tech industry.

Lessons Learned

While some hypothesis may be logical in nature, it always helps to have a data backed result to be even more certain. Data backed insights help in decision making. It especially helps to concentrate our efforts and allocate resources on to the most important aspects maximizing the impact. Interpretation of categorical variable significance was robustly examined and learned in the analysis, especially for the Poisson regression model fits. There were many times were the p-values would constantly change depending upon what variables were controlled for leading to confusing results. It was learned through the investigation that it is important to connect the meaning of a categorical variable’s coefficient as the effect of that qualitative variable in comparison to the mean response of uncontrolled variable coefficients (the baseline or intercept). If the baseline changes it may change the significance of a variable. Also, the effect of the number of observations is seen in the data. If a certain controlled variable does not have enough observations, then the effect is difficult to interpret. Using a threshold number of observations to determine the effect of a variable is important in gauging the significance of that variable.

Appendix

Link to GitHub repo with all code used for the analysis:

https://github.gatech.edu/mganesan7/ISYE_6414_Project/tree/master/Code

Equations

$$log(\hat{u}) = \beta_{baseline} + X_{negative\ sentiment}\beta_{negative\ sentiment} + X_{USA}\beta_{USA} + X_{Canada}\beta_{Canada} + X_{Netherlands}\beta_{Netherlands} + X_{Germany}\beta_{Germany}$$

$$log(\hat{u}) = 3.223 + 0.223X_{negative\ sentiment} + 1.947X_{USA} - 0.815X_{Canada} - 1.144X_{Netherlands} - 1.326X_{Germany}$$

Equation 1: Poisson regression model equation for estimating of the count of mental illnesses based on country designation and negative sentiment in the workplace

$$log(\hat{u}) = \beta_{baseline} + \beta_{offers\ coverage}X_{offers\ coverage} + \beta_{USA}X_{USA}$$

$$log(\hat{u}) = 1.622 + 1.219X_{offers\ coverage} + 2.743X_{USA}$$

Equation 2: Poisson regression model equation for estimating the count of individuals seeking treatment based on country designation and employer's coverage policy to include mental health benefits.

$$log(\hat{u}) = \beta_{baseline} + \beta_{USA}X_{USA} + \beta_{offers\ coverage}X_{offers\ coverage} + \beta_{offers\ coverage,USA}X_{offers\ coverage}X_{USA}$$

$$log(\hat{u}) = 2.534 + 1.18X_{USA} - 0.272X_{offers\ coverage} + 2.275X_{offers\ coverage}X_{USA}$$

Equation 3: Poisson regression model equation for estimating the count of individuals seeking treatment based on country designation and employer's coverage policy to include mental health benefits with an interaction term

Figures

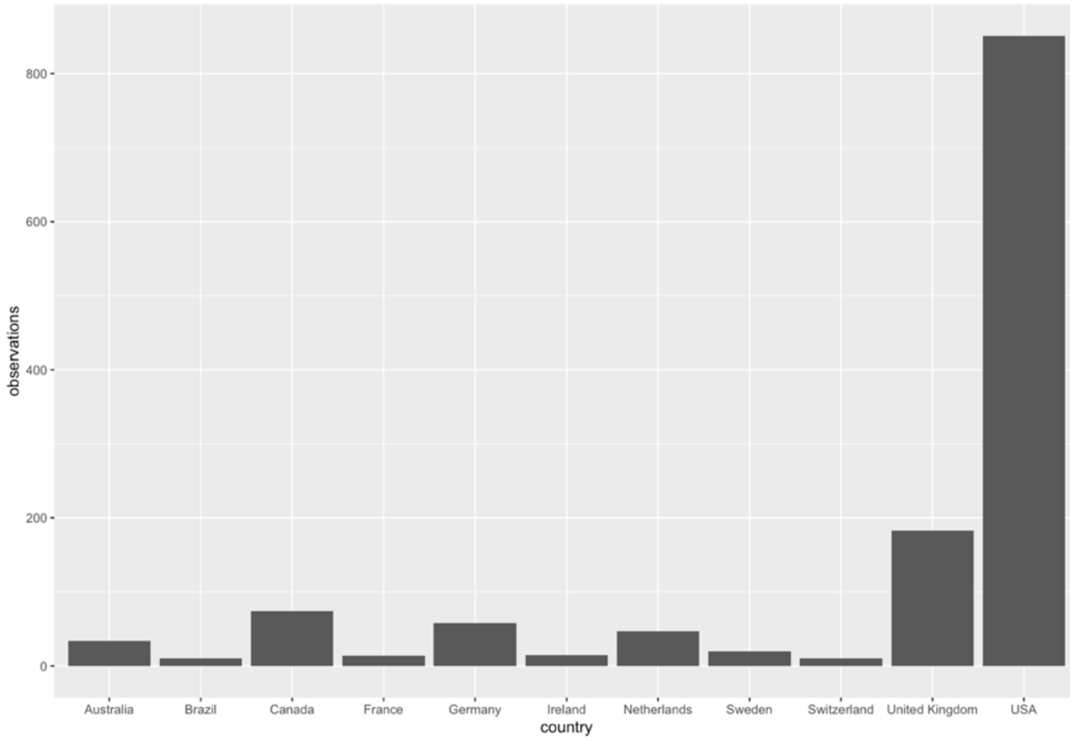


Figure 1: shows a bar plot of the number of observations in the survey for each of the top 11 countries with observations

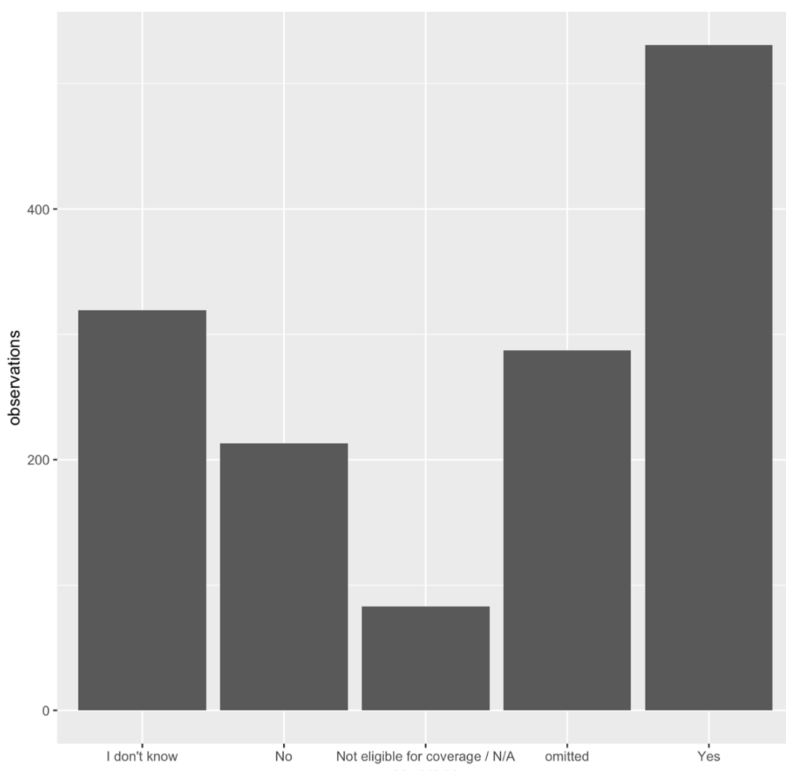


Figure 2: shows a bar plot of the number of observations in the survey for each value for the variable indicating whether the employer provides mental health benefits in their coverage policy.

Tables

Table 1: shows the aggregated form of the count of mental illnesses based on country designation and negative sentiment in the workplace

negativeSentimentWork	isInUS	isInUK	isInCA	isInDE	isInNL	Y_illness
0	0	0	0	0	0	22
0	0	0	0	0	1	6
0	0	0	0	1	0	8
0	0	0	1	0	0	10
0	0	1	0	0	0	26
0	1	0	0	0	0	180
1	0	0	0	0	0	36
1	0	0	0	0	1	9
1	0	0	0	1	0	10
1	0	0	1	0	0	15
1	0	1	0	0	0	29
1	1	0	0	0	0	216

Table 2: shows the aggregated form of the count of individuals seeking treatment based on country designation and employer's offering mental health benefits

offersCoverage	isInUS	isInUK	isInCA	isInDE	isInNL	Y_treatment
0	0	0	0	0	0	19
0	0	0	0	0	1	6
0	0	0	0	1	0	8
0	0	0	1	0	0	5
0	0	1	0	0	0	25
0	1	0	0	0	0	41
1	0	0	0	0	0	12
1	0	0	0	0	1	5
1	0	0	0	1	0	4
1	0	0	1	0	0	13

1	0	1	0	0	0	14
1	1	0	0	0	0	304

Table 3: shows the p-value for an F-test for each variable in the country designation and negative sentiment Poisson regression model. The null hypothesis is the variable's coefficient is zero and the alternative hypothesis is the variable's coefficient is non-zero. A small p-value indicates that it is not likely that the null hypothesis is true and therefore the coefficient is most likely significant

coefficient	F-test p value
negative sentiment	0.0027097
USA	0.0000000
Canada	0.0003137
Netherlands	0.0000778
Germany	0.0000442

Table 4: shows the p-value for an F-test for each variable in the country designation and coverage policy Poisson regression model. The null hypothesis is the variable's coefficient is zero and the alternative hypothesis is the variable's coefficient is non-zero. A small p-value indicates that it is not likely that the null hypothesis is true and therefore the coefficient is most likely significant.

coefficient	F-test p value
offers coverage	0.0099110
USA	0.0000343

Table 5: shows the p-value for an F-test for the interaction variable in the country designation and coverage policy Poisson regression model. The null hypothesis is the variable's coefficient is zero and the alternative hypothesis is the variable's coefficient is non-zero. A small p-value indicates that it is not likely that the null hypothesis is true and therefore the coefficient is most likely significant.

coefficient	F-test p value
offers coverage * USA	0.0021178

Table 6: Logistic Model1 Results

Coefficients	**Estimate**	**z-value**	**p-value**	**Significance**
Know_Options_Mental_CoverageYes	0.455901	2.278	0.022728	*
Discuss_MentalHealth_Negative_ConsequenceNo	0.371816	2.001	0.045422	*

Discuss_MentalHealth_Negative_ConsequenceYes	0.761857	3.18	0.00147	**
Past_Mental_IllnessNo	-2.135299	-9.094	< 2e-16	***
Past_Mental_IllnessYes	-0.537582	-2.004	0.04507	*
Diagnosed_Mental_condition_Med_profYes	0.681761	2.82	0.004806	**
Treatment_Interfere_workOften	2.79665	2.49	0.012767	*
Treatment_Interfere_workRarely	0.159804	0.715	0.47447	
NonTreatment_Interfere_workOften	2.564408	9.242	< 2e-16	***
NonTreatment_Interfere_workRarely	2.332153	9.809	< 2e-16	***
Age	-0.032973	-3.366	0.000761	***
Prev_Empl_Mental_Health_Weighed_SeriouslySome	-0.530082	-2.957	0.00311	**
Prev_Empl_Mental_Health_Weighed_SeriouslyYes	-0.738152	-1.515	0.129871	

Table 7: Logistic Model2 Results

Coefficients	**Estimate**	**z-value**	**p-value**	**Significance**
Has_Employer_Discussed_Mental_HealthYes	0.5377	2.493	0.01266	*
Asking_Leave_DueTo_Mental_Health_IssueEasy	-0.8256	-3.712	0.000206	***
Asking_Leave_DueTo_Mental_Health_IssueOther	-0.5159	-2.2	0.027802	*
Aware_Options_Prev_Mental_CareYes	0.7511	4.108	3.99E-05	***
Prev_Empl_Mental_Health_Weighed_SeriouslySome	0.5095	2.637	0.00836	**
Prev_Empl_Mental_Health_Weighed_SeriouslyYes	-0.8715	-1.571	0.116203	
Past_Mental_IllnessNo	-1.0271	-4.578	4.69E-06	***
Past_Mental_IllnessYes	0.8112	3.404	0.000664	***
Diagnosed_Mental_condition_Med_profYes	2.575	11.045	< 2e-16	***
Treatment_Interfere_workOften	1.0399	2.198	0.027948	*
Treatment_Interfere_workRarely	1.1665	6.048	1.47E-09	***