Mental Health in Tech

# **Andrea Fox Summer 2021 https://github.com/anfox86/Mental-Health-in-Tech.git**

# Which Domain?

My main dataset is coming from Kaggle that was listed as a task. The following are references I have looked into and found may be useful in my analysis:

1. <https://nycdatascience.com/blog/student-works/mental-health-in-tech-industry/> this goes through analysis using the same dataset I plan to use. I’d like to see how someone else approaches it, but plan to do my own analysis and work.
2. <https://thriveglobal.com/stories/are-we-tackling-mental-health-in-tech-companies-the-right-way/> get more background on mental health in tech.
3. <https://thriveworks.com/blog/mental-health-tech-industry-employee-wellbeing/> more information on employee mental wellbeing. Talks through some of the consequences of not addressing the issue.
4. <https://blog.hubspot.com/service/survey-results?toc-variant-a=> walks through how to evaluate survey results. I’ve not worked with survey results much, so want to make sure I understand the methods before starting.
5. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4601897/> paper on understanding and evaluating survey research. Since this is opinion based data there may be some tips in here for evaluation.
6. <https://towardsdatascience.com/how-to-analyze-survey-data-in-python-c131764ea02e> analyzing data in python. Might have some tips on how to get started once I read in the survey results.
7. <https://www.abtassociates.com/insights/perspectives-blog/the-growing-importance-of-data-science-methods-to-survey-research> this talks specifically about machine learning and the importance of data science methods when applied to survey research. Again not sure how technical this project will be so using this as a learning opportunity.
8. <https://towardsdatascience.com/data-science-in-mental-health-ccd09ba2148a> talks through data science in the healthcare space. It’s extremely popular but has to be done with caution to not create bias and maintain a level of privacy.
9. <https://www.nature.com/articles/s41562-018-0470-9> paper on advancing mental health research through data science. This aligns nicely how I want to wrap up my project.
10. <https://www.mastersindatascience.org/resources/how-data-science-can-improve-mental-health-care/> another resource that talks through how data science can help elevate treatment for mental health which will helps in some of the wrap up of the project.

# Which Data?

The dataset I intend to use for this project came from Kaggle and is titled Mental Health in Tech Survey, which was completed in 2014.

<https://www.kaggle.com/osmi/mental-health-in-tech-survey>

This survey goes through a variety of questions after first establishing the person’s age, gender, country, and state (if applicable). It looks at people from both self employed or employed by tech companies then asks about whether treatment has been sought out before, does the employer have care options or a wellness program like EAP. It then gets into privacy related questions such as can you seek treatment anonymously or do you think you will be chastised for having a mental health condition. There are 27 columns in this set that I believe I will keep, but until I start dissecting it I won’t know for sure.

# Research Questions? Benefits? Why analyze these data?

There are two questions on the dataset explanation that I think are worth exploring which are:

1. How does the frequency of mental health illness and attitudes towards mental health vary by geographic location?

This could be fascinating. I’m friends with a few people that are European and they are typically less vocal on things like mental health. I don’t think it’s because they don’t take it seriously, it’s just not widely talked about.

1. What are the strongest predictors of mental health illness or certain attitudes towards mental health in the workplace?

So, for this one I wonder if the columns asking about negative consequences in relation to talking or seeking help for mental illness. That could point to a toxic environment that should be pointed out to an employer.

Some other questions that come to mind:

1. Does age or gender factor into the person being more willing to discuss a mental health issue with a coworker or supervisor?

Since the push for better mental health has been a more recent occurrence in the workplace, I would suspect younger adults would be more willing to seek help and know about available benefits.

1. Is there a trend between family history of having a mental health issue and the person having one themselves? In other words, is this a potential heredity issue?

Heredity is an odd thing. Some things are passed down while others aren’t, so I would like to see if survey respondents have that trend because someone else in their family did.

1. Because remote work has been a more normal thing now, I’m curious if the survey respondent has had to seek treatment more or less while working remote.

I know this survey was 2014, but I think because of Covid it makes me think whether employees have better mental health because they are home surrounded by familiar things or if it has caused issues. I know me personally it took awhile to adjust, so my mental health wasn’t the best during that adjustment period, while others on my team absolutely love it.

1. Is there a trend between the column mental vs physical and the columns of benefit, care options, wellness program, or seek help? For example, if the respondent believes that their employer doesn’t take mental health as seriously as physical health are they likely to know what options there are out there benefit wise?

So, this one is a bit personal for me. My stepfather never believed in mental health or mental illness, so he never took it seriously. If an employer doesn’t take mental health as seriously as it does physical health, is it likely the employee would know what options are available to them?

# What Method?

I plan to come to the project acting as a consultant that wants to know more about how mental health is addressed in the workplace, specifically in technology. My objectives is to get an idea of if employees know what resources are available to them, are they using said resources, and do they feel comfortable using those resources and benefits. I also plan to do some evaluation on location of employees since this could vary how mental health is accepted. Traditionally Europeans are more closed off and keep their problems internal versus Americans who are known to be more open about talking about things (speaking from my experience not as a knowledge expert here). I think there will be a lot of visual analysis to get ideas on trends and totals. That seems to be the way most of the articles I’ve looked at regarding survey data analysis have completed a large part of their analysis. I did find an article that talked through probability sampling methods and non-probability sampling methods so I may go that route and do some clustering.

# Potential Issues?

So far, the biggest challenge I think I may face is mismatched data. I included all of the columns and went down the list in Kaggle to get an idea of what all was included, and it shows whether there is valid, mismatched, or missing data. Some of the variables I want to use look like they have mismatched data, so not sure what that means exactly yet but could potentially cause problems in clean-up and analysis.

The other challenge I might face is keeping myself on track. This subject is fascinating, so I have to be careful I don’t go down a rabbit hole or two while I analyze.

# Concluding Remarks

I plan to approach this project as a consultant who is interested in finding out how employees in the tech industry feel about mental health and whether they believe their employers take mental health seriously by providing programs and benefits. My final report will be whether I feel like most survey respondents feel as if their mental health is being taken care of, and if I find otherwise provide analysis and suggestions on how this could be handled better. I wanted to approach this project as if I was doing something for a client because it makes questions and analysis easier than just approaching the project like I did in the first project.