Analyzing the Effects of COVID-19 on Mental Health Through Tweets

Miranda Palmer#1

#College of Information Sciences and Technology, The Pennsylvania State University, United States of America

1mrp5636@psu.edu

***Abstract*—This document outlines research done to analyze the effects of Covid-19 on mental health. The researcher investigates the frequency of Tweets indicative of psychological stress as well as identifies any change in topics frequently associated with stress. Two datasets from before and after Covid-19 are labeled to identify, for each Tweet, whether or not it uses stressed language, the primary emotions conveyed, and the general topics included. Each dataset is analyzed separately using TweetNLP [1] so that the results can be compared to determine the change in mental health. First, the frequency of Tweets containing stressed language is calculated. Then, the emotions most frequently associated with stressed language are identified. Finally, for Tweets that are classified as one of these emotions, the most frequent topics are identified.**

After the emotions and topics representative of stress are identified, the two datasets can be compared to demonstrate in what ways society's general mental health has changed in response to Covid-19. The findings of this investigation can be further compared to that of related work which identifies stressors specifically related to Covid-19 [3], as well as another piece of related work that identifies the correlation between certain keywords and psychological stress [2].

Keywords—Covid-19, pandemic, mental health, psychological stress, Natural Language Processing, TweetNLP, Twitter, social media data

I. Introduction

Covid-19 created a series of unprecedented impacts to the way society and individuals function. One of those impacts was a widespread decline in mental health. As reported by the World Health Organization, "In the first year of the COVID-19 pandemic, global prevalence of anxiety and depression increased by a massive 25%", which demonstrates a measurable increase in mental illness since the start of the pandemic [4]. In order to identify possible causes for this decline in mental health, social media data could provide some unique insights into the general mental wellbeing of society.

In previous works of Natural Language Processing, authors suggest that social media data can be indicative of the raw mental state of users and has been utilized to investigate the types of emotions or related topics that users are experiencing [1]. Due to the more personal and seemingly anonymous nature of social media, user's language can be indicative of their personal emotions and used to make assumptions about the general public. In this research, Twitter data is used to determine the general mental wellbeing of users.

While there is no one metric to quantify mental health, stress has been proven to serve as a reliable indicator of the mental wellbeing of an individual. The authors of related works [2] suggest that stress can be indicative of a person's mental state and that it is more useful to investigate general psychological stress rather than acute mental illnesses like depression or anxiety. Their main justification to this claim is that stress serves as a precursor to acute mental health issues, so by identifying possible causes of stress and intervening to eliminate them, any subsequent mental health issues can be alleviated or avoided altogether. The research outlined in this document also identifies stress as a primary indicator of the mental wellbeing of a society to encompass all types of mental issues and make claims about the general public.

Using stress as a central topic, the first part of this research is conducted by utilizing two different datasets, one from 2018 and one from 2020, to add three additional variables to each Tweet. The datasets are each analyzed using strategies outlined in related work to label Tweets as containing stressed language or not [2]. Then, methods used in related work are used to identify the primary emotions conveyed in each Tweet [1]. The frequency that each emotion appears in stressed versus not stressed Tweets is meant to be representative of the relationship between that emotion and stress. As a final measure of stress, each Tweet is labeled with the topic of its contents by utilizing methods of related work [1]. Again, the frequency that each topic appears in stressed versus not stressed Tweets represents the relationship between that topic and stress.

The results of the first part of this research produce two datasets which are labeled in terms of the presence of stressed language, primary emotions of the Tweet, and primary topic or keyword of the Tweet. The second part of the research compares the two datasets in terms of their three additional variables to determine any changes between them as well as compares the results to findings from previous work. The frequency of stressed language will be compared between each dataset. The emotions that are most frequently related to stress are calculated and compared. Finally, the topics that are most frequently related to stress are also calculated and compared. These findings are evaluated by comparing the kinds of emotions identified in previous work as having a high correlation to stress [2].

Also, the findings are compared to previous work that identifies certain topics that are correlated to negative emotions about Covid-19. The results of the previous work are compared to those of this research to validate that the TweetNLP model is able to accurately identify emotions related to high stress as well as the topics indicative of high stress during Covid-19.

The primary motivation for this research is to combine and improve upon methods conducted in three documents of related work [1], [2], [3], for the purpose of investigating the possible causes for an increase in mental health issues after Covid-19. These causes can be identified at an individual and a societal level to reveal several possible points of action to alleviate stress before it becomes a chronic issue.

II. Literature Review

This research utilizes three main documents as related work. Each of these documents provides a unique insight into the topic of stress as it is related to general Tweets and those about Covid-19 specifically.

[1] *TweetNLP: Cutting Edge Natural Language Processing for Social Media*

As mentioned in the previous section, this document introduces the TweetNLP model in which the authors improved upon current methods for Natural Language Processing of social media data and created a tool that can be used to study individual and group communication through social media. Their main approach to achieve this goal involved using transformer-based language models that are trained using only raw, unstructured social media data. They evaluated the TweetNLP model against general purpose NLP models as well as Twitter specific NLP models to understand how their model compared with current technologies.

As they discuss in their literature review, most NLP libraries and technologies are not specialized in social media data. The most recent approach for addressing this problem is the development of a Twitter-specific python library that provides models for tokenization and lemmatization. However, they discuss the fact that these libraries are only useful if the model is able to process unclean data. In order to target the weaknesses in many current NLP methods, they specifically trained their model on raw and relatively messy Twitter data. The language models used in TweetNLP rely on RoBERTa-base or XLM-R-base architecture, which are pre-trained models that are recognized as the most efficient models for the purpose of social media specific data. They utilized these models and continued to train with their own Twitter data to build upon the training that was done previously.

After creating the TweetNLP model, they evaluated its performance through a series of Twitter-based tasks and compared the results with four general purpose models as well as three Twitter specific models. They found that the TweetNLP model performed slightly better on some datasets than others but was able to perform about the same as or better than current models. While the TweetNLP model did not consistently outperform current methods, their hypothesis that general purpose NLP models would not perform as well as Twitter specific models was true. They conclude by stating that the use of Twitter specific models and libraries can improve the performance of sentiment analysis in Tweets, and that their TweetNLP model is intended to be utilized for research in human communication through social media.

The two TweetNLP tasks that are implemented in my research are used to create new variables in each of the datasets. The first task that will be utilized is emotion recognition, and it will be used to determine the emotions present in every Tweet of each dataset. While the two datasets already contain labels for their emotions, it may provide more comparable results to label each dataset by use of the same emotion recognition model. Following the same logic as the emotion recognition task, the Tweets of both datasets will also be classified by the topic classification task. A limitation of this work is understanding how language changes over time, specifically in an environment like social media where slang terms and phrases are used very frequently.

[2] *Understanding and Measuring Psychological Stress Using Social Media*

The second document that was utilized as a form of related work investigates the use of language around stress in social media and how it is indicative of mental wellbeing. As mentioned in the previous section, the authors of this document emphasize the importance of identifying causes of general psychological stress before investigating acute mental illnesses like depression. The authors conducted a study to determine the stress level of their participants as well as collect their social media data. They analyzed this data to determine how stress is talked about on social media. They then altered user-level models to work on community-level data and validated the predicted stress of a participant by factoring in health and socioeconomic factors. Finally, they validate the predicted stress of a participant by considering health and socioeconomic factors.

In their literature review, they discuss the lack of language models to predict psychological stress, which is their main goal in their research. They also discuss the fact that there is a lack of validation against region-level ground truth for socioeconomic stress factors.

The first step of their research included a study in which they conducted a survey and asked participants to answer questions from the Cohen Stress scale, as well as provide personal information such as their location and age. The participants were also asked to give access to their Twitter and Facebook data in order to correlate the results of the survey with their language use on social media. They then used normalized frequency distributions to represent the language of each user and county, calculating a mean stress score for each user by detecting expressions of stress at the sentence level. The second step of their research involved making stress predictions on a set of geo-located Tweets and validating these predictions against county-aggregated stress labels in order to examine the correlation between socioeconomic factors and stress.

Through their evaluation of the survey results against social media data, they found that the language of stress is often very self-focused and shows a perceived lack of control or resources. Stressed users tended to post about exhaustion, losing control, focus, and physical pain as opposed to non-stressed users who post about travel and family. They also determined that the most frequent words positively correlated with stress were first person phrases like "I" or "me" while non-stressed users more frequently used phrases that indicate community such as "we" and "our". This reiterates previous research on language use in mental health and demonstrates that the same logic can be applied to social media data. The addition of social media language is able to outperform models that are only trained on sociodemographic variables. Also, the model is able to accurately predict correlations to known sociodemographic stressors identified in related research. This means that social media language can accurately determine the stress of its user and, through the investigation of sociodemographic factors, can be used to determine the cause of a particular user's stress. However, they caution that it cannot always be assumed that social media is representative of the emotional state of real-life users without the addition of sociodemographic data. That being said, they hope that their research is able to help design a social media based intervention at the user level that is able to "enable a low stress lifestyle" and develop a better understanding of regional variations in stress.

To utilize the methods of this document within my research, the Tweets are going to be evaluated using the same lexicon of stressed language in order to designate a stress score for each Tweet. This is for the purpose determining the level of stressed language that each Tweet contains in order to correlate high-stress Tweets with other factors of the Tweet. The only limitation to this method is that it assumes that social media data is representative of the stress level of real life users and that their social media habits reflect their mental state.

[3] *What Are We Depressed About When We Talk About Covid-19: Mental Health Analysis on Tweets using Natural Language Processing*

The final document of related work that is utilized in my research investigates Covid-19 related Tweets and their correlation to negative emotions. They seek to investigate how and why the public feels negatively about Covid-19. This is the document that proposes that the utility of social media data is due to its anonymous nature. The authors reason that because users post their feelings and emotions on social media, it can be useful to analyze language to estimate mental health conditions, especially after an event like Covid-19. The authors intended to create a dataset with Covid-19 related keywords to handle the issue of the lack of a labeled Covid-19 dataset. Then, they utilized multilingual models to do single and multi-level classification to analyze trends in certain emotions.

The authors begin their research by tackling three main limitations with existing research on Covid-19 Twitter data. The first limitation is the fact that there is a lack of available labeled Covid-19 dataset. In order to solve this issue, they also address the problem of limited API availability by manually creating their own labeled dataset. The third limitation that they attempt to remedy is the fact that most datasets and NLP models are only trained on English Tweets, so they extend their research to include multilingual Tweets. After considering the possible limitations of their work and taking steps to solve them, the authors consider several different models for the purpose of predicting emotions and classifying Tweets. They decided to use the BERT model because of its multilingual capabilities and the fact that it is considered to perform relatively better than most other NLP models.

They follow three main steps in their research. The first being that they use Twitter API to collect Tweets with certain keywords from March to April 2020. The keywords include Covid-19 specific topics such as "lockdown", "mask", "sick", etc. After collecting the Tweets, they use a pre-trained multilingual BERT model to classify and label the emotion of each Tweet. They then used the classifications to further identify the keywords that are most frequently correlated with specific emotions. Finally, they investigate their findings by creating a time series graph to visually analyze the distribution of emotions for each keyword. Each of these steps are executed with the intention of determining the changes in certain emotions during the early months of the lockdown.

They found that negative emotions like fear, anger, and sadness were most frequent among Tweets that contained the word "lockdown" and had a noticeable spike around the time that mandatory lockdowns were being announced. Alternatively, the word "mask" had a generally positive emotions tied to it such as optimism and joy. This research helped to identify how and why the public feels negatively about Covid-19 and provides ways to gather real-time data about the emotions present among a community.

This research is utilized in this document as it provides a set of topics and emotions that are correlated with poor mental health as well as Covid-19. These findings will be used to validate the emotions labels of the second dataset and to provide topics to investigate within the second dataset as reasons for negative emotions surrounding Covid-19. For example, based on the research in this document, it is useful to investigate terms like "lockdown" and "mask" as they may have a strong correlation to Covid-19 related Tweets.

These three documents together provide unique methods for analyzing the level of stress in Tweets and identifying the emotions and topics related to Tweets that are indicative of psychological stress.

III. Contributions and Novelty

In my own work, I combine the insights found in the three related works mentioned above to provide an in-depth analysis of the effects of Covid-19 on mental health. To perform this analysis, I utilize two different datasets, one which was used in the first paper [1] as a training set for emotion recognition and is from the year 2018. The second dataset is for sentiment analysis from Kaggle and spans from 2020-2022. I split each dataset into their own respective training and testing sets. The training sets are labeled and evaluated separately for the analysis portion of my research to determine the differences between Tweets before and after Covid-19. The testing sets are combined into one large set. Given that the TweetNLP [1] library was developed on data that was collected before Covid-19, I wanted to evaluate the accuracy of the models with Post-Covid data to see if the models still perform with similar results. While my contributions were not created with the intention of improving the accuracy of the models, the purpose of this testing set is to evaluate the accuracy of the results that I gathered in order to supplement my research and understand any anomalies that may occur.

The first stage of my contributions involves the use of the TweetNLP [1] natural language processing to individually label each Tweet based on the presence of irony or offensive language, as well as label it's emotion and most relevant topics. While my analysis does not specifically focus on the presence of offensive language or irony, I include these classifications as a means of determining how the language differed between the two datasets. The emotion classification labels Tweets as either containing joy, optimism, anger, and sadness. Finally, the topic classification includes 19 different categories of topics which include, arts and culture, business and entrepreneurs, celebrity and pop culture, diaries and daily life, family, fashion and style, film tv and video, fitness and health, food and dining, gaming, learning and educational, music, news and social concern, other hobbies, relationships, science and technology, sports, travel and adventure, and youth and student life. For each Tweet, I create an ordered list of the three topics that the Tweet was predicted to be most related to. This list is stored within a separate column of the new labeled datasets. I perform this process separately on the pre- and post- Covid datasets.

Through this process of labeling each Tweet in the datasets, I found that the platform I was using to conduct this research was not capable of handling large volumes of data. I had initially decided to utilize Jupyter Notebook because I would be able to divide my code into separate chunks and explain each step of my process in text boxes above the code. Unfortunately, while I did intend on using a large amount of data to perform this investigation, due to time constraints I needed to reduce the number of Tweets in each dataset. To do this, I took a random sample of 10,000 Tweets from each dataset and performed my analysis on only those Tweets. I will discuss the impacts of this decision Tweets in the Future Work section.

The second stage of my contributions includes the use of methods developed in the second related work in order to determine the stress level of each Tweet. In order to do this, I utilize the lexicon of stressed language [2] that was developed as a result of extensive research into how psychological stress is represented in social media posts. Using the stress lexicon, I was able to identify a stress score for each word that was included in the lexicon. From there, I aggregated each word's score to create a total stress score for a given Tweet. This was the recommended procedure for utilizing the stress lexicon which was discussed in its documentation. This process was also performed separately for each of the datasets, and the stress score is represented as a numerical value in a new column of the labeled datasets.

After labeling each Tweet for a series of classifications as well as obtaining a stress score, I investigate the correlations of these results within each dataset as well as compare those results between datasets. In the following section, I describe how I perform my analysis to investigate the types of emotions and topics that are most frequently associated with stressed language.

IV. Findings and Analysis

In order to understand the state of mental health Pre- and Post-Covid, I conduct an analysis of each dataset separately. First, after obtaining the stress scores, I wanted to understand the distribution of stress scores and determine what would be considered a high versus low stress score as well as investigate how the stress score changed between datasets. I found that the average stress score of the Pre-Covid data was about 196 and that of the Post-Covid data was about 328, which indicates that the average stress score was slightly higher in the Post-Covid dataset. From here, I wanted to establish a threshold for a "high stress" Tweet. To do this, I found that the maximum stress score of the Pre-Covid data was about 4025 and the maximum stress score of the Post-Covid data was about 2258. Based on this information, I chose a stress score of 800 to represent high stress. This score is significantly higher than the average and would encompass enough of the high stress Tweets to be able to perform a meaningful analysis.

Following the stress score threshold, I created two new datasets that only included the Tweets from the Pre- and Post-Covid datasets with high stress scores. The resulting datasets also demonstrate that frequency of high stress Tweets increases between the datasets. There were 318 Tweets in the Pre-Covid dataset that had a stress score over 800, while there were 723 of these Tweets found in the Post-Covid dataset. This confirms the findings that I mentioned earlier comparing the average stress score between datasets.

I continue my analysis using the new datasets that only contain high stress Tweet to determine the number of Tweets in each dataset that contained offensive language or irony. As I mentioned earlier, the purpose of this step of my analysis was to gain an understanding of how the language might have changed over time. I found that in the Pre-Covid dataset, about 10% of tweets contained offensive language and in the Post-Covid dataset, offensive language was present in about 9% of Tweets. This demonstrates that the use of offensive language did not vary significantly between the two datasets. In addition, about 53% of the Tweets in the Pre-Covid dataset were labeled as ironic and about 48% of the Tweets in the Post-Covid dataset were labeled as ironic. This demonstrates a moderate change in the frequency of irony in Tweets. This slight difference could account for some abnormalities in my results as the use of irony did appear to change over time.

Next, I determined the emotions that were most frequently associated with the high stress Tweets in the new datasets. Interestingly, my results showed that from the Pre-Covid dataset, joy was the emotion that was most frequently associated with high stress Tweets. About 66% of high stress Tweets from the Pre-Covid dataset were labeled as joy while only 17% were labeled as anger, the second most frequent emotion. Compared to the Post-Covid dataset, which had anger as the leading emotion, about 65% of Tweets were labeled with anger. Also, joy was the least frequent emotion among this set of Tweets, with optimism being the second, found in 14% of Tweets. While the results for the Post-Covid data aligned with my expectations, the overwhelming presence of joy within high stress Tweets before Covid-19 was unexpected. Based on these results, I decided to investigate the topics that were most frequently associated with anger and joy.

For the Pre-Covid data, the results showed that the topic most frequently associated with joy was celebrity and pop culture, followed by news and social concern, and finally music. About 22% of Tweets labeled "joy" included celebrity and pop culture as part of their most relevant topics. While only 16% and 14% of these Tweets contained news and social concern or music in their list of relevant topics, respectively.

As for the emotion of anger, the Pre-Covid data showed that the topics of news and social concern, celebrity and pop culture, and sports were the most frequent topics associated with anger before Covid-19. About 28% of Tweets labeled "anger" included news and social concern in their list of relevant topics. In comparison to celebrity and pop culture or sports which were found in 22% and 13% of Tweets respectively.

For the Post-Covid data, the results showed that the topic most frequently associated with joy was news and social concern, followed by celebrity and pop culture, and diaries and daily life. News and social concern accounted for about 27% of Tweets, while celebrity and pop culture or news and social concern were present in about 14% and 13% of Tweets respectively.

After Covid-19, the topics most frequently associated with anger were news and social concern, diaries and daily life, and celebrity and pop culture. Tweets about news and social concern accounted for about 33% of the data, while diaries and daily life and celebrity and pop culture was present in 27% and 22% of Tweets respectively.

After determining the topics most frequently associated with joy and anger, I analyze these results in comparison to the related works I identified in the Literature Review section as part of my Conclusion.

V. Conclusions

To evaluate the accuracy of my results and further investigate my findings, I compare my results to the research conducted in the second [2] and third [3] documents of related work. To aid in my analysis, I created two charts that summarize the topics most frequently associated with joy or anger for each of the datasets.

Pre-Covid		
	Topic	Frequency
Joy	celebrity and pop culture	22%
	news and social concern	16%
	music	14%
Anger	news and social concern	28%
	celebrity and pop culture	22%
	sports	13%

Table 1: Pre-Covid Emotion-Topic Correlations

Post-Covid		
	Topic	Frequency
Joy	news and social concern	27%
	celebrity and pop culture	14%
	diaries and daily life	13%
Anger	news and social concern	33%
	diaries and daily life	27%
	celebrity and pop culture	22%

Table 2: Post-Covid Emotion-Topic Correlations

Based on these findings, my initial observations are that the topic of news and social concern tends to be very frequent among both datasets. In the Post-Covid data, it accounted for the most Tweets regardless of the emotion the Tweet was conveying. Similarly, celebrity and pop culture seemed to be very frequent within both of the datasets. Due to this, it can be inferred that the frequency of such topics does not necessarily have to do with the emotions that the Tweet is conveying but that they were the most frequent topics of all of the Tweets with high stress scores.

In the second paper [2], the authors conclude that topics that have to do with more personal aspects such as diaries and daily life or news and social concern would be most indicative of high stress. This is consistent with my findings, and the increased presence of diaries and daily life among high stress Post-Covid Tweets indicates that people were more frequently tweeting about their personal life after Covid-19. Additionally, topics like music and sports were most frequently recorded before Covid-19, indicating that the general public was more concerned with issues of news, social concern, and their daily life after Covid-19 rather than their usual hobbies or interests. This could be considered a potential side-effect of the increased stress caused by Covid-19 and is a possible reason for the increase in mental health issues recorded after the pandemic. As is mentioned in the second related paper [2], a loss of interest in hobbies or usual activities is common among people experiencing mental health problems such as depression. Similarly, an increase in concern over personal problems represented by "diaries and daily life" or social issues indicates a strong focus on the self and could also point to an increase in concerns about losing control over their social situation.

Comparing this to the findings of the third related paper [3], the authors found that among Covid-19 related Tweets, people were most often feeling negative emotions about the lockdown. On the other hand, people were most often feeling positive emotions about the use of masks. This is a possible explanation for the fact that the Post-Covid dataset included relatively more Tweets that talked about issues of news and social concern or diaries and daily life, but that it was not significantly different between emotions of joy and anger.

To conclude, my findings were consistent with research conducted in both the second [2] and third [3] related works that I mentioned in my Literature Review section. The results of my analysis demonstrate that there was a noticeable increase in the average stress score of Tweets before and after Covid-19, as well as an increase in the frequency of high stress Tweets. This affirms the fact that there has been a documented increase in mental health issues since the start of the pandemic and indicates that people have been using social media to express this. Based on the content of these Tweets, public concern has shifted away from hobbies and interests and towards societal and personal issues, which could potentially explain the increase in stress.

In the following section I will discuss the utility of these findings and how I could have conducted my research differently to account for some complications I encountered.

VI. Future Work

In order to improve the work that I have already done as well as further my research in this topic, I identify three main steps that I will take in my future work. These initial steps are meant to tackle some of the pitfalls of my own research while also creating a additional information to be analyzed.

The first step that I would take in my future work would be to implement my code on a platform that is capable of supporting large datasets. In my research, I decided to use the same platform that was utilized by my first parent paper because it was already split up into modular chunks of code that could easily be manipulated. However, I found that this platform, Jupyter Notebook, was incapable of processing the larger datasets that I gathered for my Post-Covid data. Due to this, I had to select a smaller random sample of Tweets in order to label each Tweet. While my analysis shows that the results I gathered were consistent with research done in my second and third parent papers, there was a lot of room for outliers and other confounding factors that could skew my results. My results and the insights that I gained from them would be more meaningful with a larger amount of data. Specifically, there would be less variation between the results of different datasets and the conclusions I make could be more accurately generalized to a population. However, there could be some potential complications that could arise as a result of this approach. First, if I were to use a different platform, it would be harder to explain each step of my process which could reduce the reproducibility of my research. I specifically chose this platform because of how it can be easily divided into sections and each block of code can be accompanied by a text box. Depending on the platform that I choose, I may not be able to easily section off my code. Second, the use of a different platform would mean that I would have to transform the code from my parent paper onto a different platform and essentially start over. While I think that this could create some complications and would be very time consuming, I think that this is a necessary step to improving the accuracy and meaningfulness of my research. After taking this step to include larger amounts of data, I intend to investigate two new aspects of the data.

The second step of my future work would be to extract keywords from each Tweet along with labeling their topics. While the topic classification model provided by my first parent paper was central to my research, I believe that more meaningful insights could be gained by extracting specific keywords rather than categorizing each Tweet into a general topic. The results of this categorization would be easier to interpret because it would focus on specific keywords rather than a general topic category. This would also make it easier to compare the results that I gather to that of my second parent paper in which they discuss the certain keywords that are often indicative of high stress. This comparison would provide a better indication of the accuracy of my results and how they align with previous research. However, a potential problem of this approach is that many of the extracted keywords may not be useful. A large portion of Tweets will result in keywords that are not frequently found in other Tweets. There would be a very wide range of keywords to analyze, and it would be difficult to group Tweets by keyword in order to understand their correlation to stress. That being said, only a certain portion of Tweets would be useful in this approach, which would reduce the number of Tweets that are being analyzed. While there could be some potential complications that result from this approach, I think that it is a worthwhile step to consider in the interest of producing more interpretable results.

Finally, the third step that I would take to begin my future work has to do with extracting information about the user rather than just analyzing the content of the Tweets. In my second parent paper, the authors were able to gather a variety of personal and demographic information about users and could analyze that data along with the Tweets. I would like to utilize this approach in my own research because I think that the results would be easier to understand with the addition of user data. For example, one area that I would like to investigate is the age groups of users and how that affects their overall stress level or the topics that they are Tweeting about. This could provide some interesting insights into the way that each age group is experiencing the effects of Covid-19 and how that differs between groups. Similarly, I think that the addition of location data could be very useful in analyzing how certain populations were affected and create more meaningful results about the stress of certain regions. This would generate some additional areas of analysis that could strengthen my research and create a clearer picture of the state of mental health. However, this approach is easier said than done. While Twitter API does allow access to user data, Twitter accounts often do not include information like age or location. In the second parent paper, the authors conducted a study and asked participants to provide that information in addition to their social media data. I would have to take an approach similar to this because of a lack of available datasets that include that kind of personal information. Also, this approach would mean that I would not be able to research on the general population but rather on a subset that is willing to participate in a study. While there are a few issues that would prevent me from gathering this kind of information, it would take my research a step further and create more meaningful conclusions about a population. A step that I think is necessary to fully understand the complexities of mental health and the lasting effects of Covid-19.

With those three initial steps in mind, I seek to continue my research into the area of mental health because it has been and will continue to be a widespread issue that could affect anyone, regardless of age, location, gender, or economic status. The effects that Covid-19 had on mental health was unprecedented and reveals a genuine threat to the general public. It is now more important than ever that we continue to research the possible causes for this shift and begin to make progress in solving this issue.

REFERENCES

[1] Camacho-Collados, J., Rezaee, K., Riahi, T., Ushio, A., Loureiro, D., Antypas, D., Boisson, J., Espinosa-Anke, L., Liu, F., Martínez-Cámara, E., Medina, G., Buhrmann, T., Neves, L., & Barbieri, F. (2022, June 29). *TWEETNLP: Cutting-edge natural language processing for social media*. arXiv.org. Retrieved October 23, 2022, from https://arxiv.org/abs/2206.14774v1

[2] Guntuku, S. C., Ungar, L. H., Eichstaedt, J. C., Jaidka, K., & Buffone, A. (2019, April 4). *Understanding and measuring psychological stress using social media*. Papers With Code. Retrieved October 23, 2022, from https://paperswithcode.com/paper/understanding-and-measuring-psychological

[3] Li, I., Li, Y., Li, T., Alvarez-Napagao, S., Garcia-Gasulla, D., & Suzumura, T. (2020, June 8). *What are we depressed about when we talk about covid19: Mental health analysis on Tweets using natural language processing*. Papers With Code. Retrieved October 23, 2022, from https://paperswithcode.com/paper/what-are-we-depressed-about-when-we-talk

[4] World Health Organization. "Covid-19 Pandemic Triggers 25% Increase in Prevalence of Anxiety and Depression Worldwide." *World Health Organization*, World Health Organization, 2 Mar. 2022, https://www.who.int/news/item/02-03-2022-covid-19-pandemic-triggers-25-increase-in-prevalence-of-anxiety-and-depression-worldwide#:~:text=Wake%2Dup%20call%20to%20all,mental%20health%20services%20and%20support&text=In%20the%20first%20year%20of,Health%20Organization%20(WHO)%20today.