Name: Chia Bing Xuan

Matriculation Number: A0259419R

Title: DSA4262 Assignment 1 – Visualising the Severity of Mental Health Conditions

1. Introduction

In this assignment, I have chosen to investigate the problem of poor mental health among today's population. To approach this task, we shall first analyse how the prevalence of mental health conditions has been trending in recent times, whilst discovering specific demographics that are more prone to these disorders. Subsequently, we will discuss the difference in lifestyle habits between the target groups which we have identified prior. Last but not least, we will consider the association of social media usage with one's mental health. This would allow us to make a more informed decision on how best to deploy our intervention in a targeted manner, thereby benefitting the mental well-being of the population to the fullest extent.

2. Plot 1

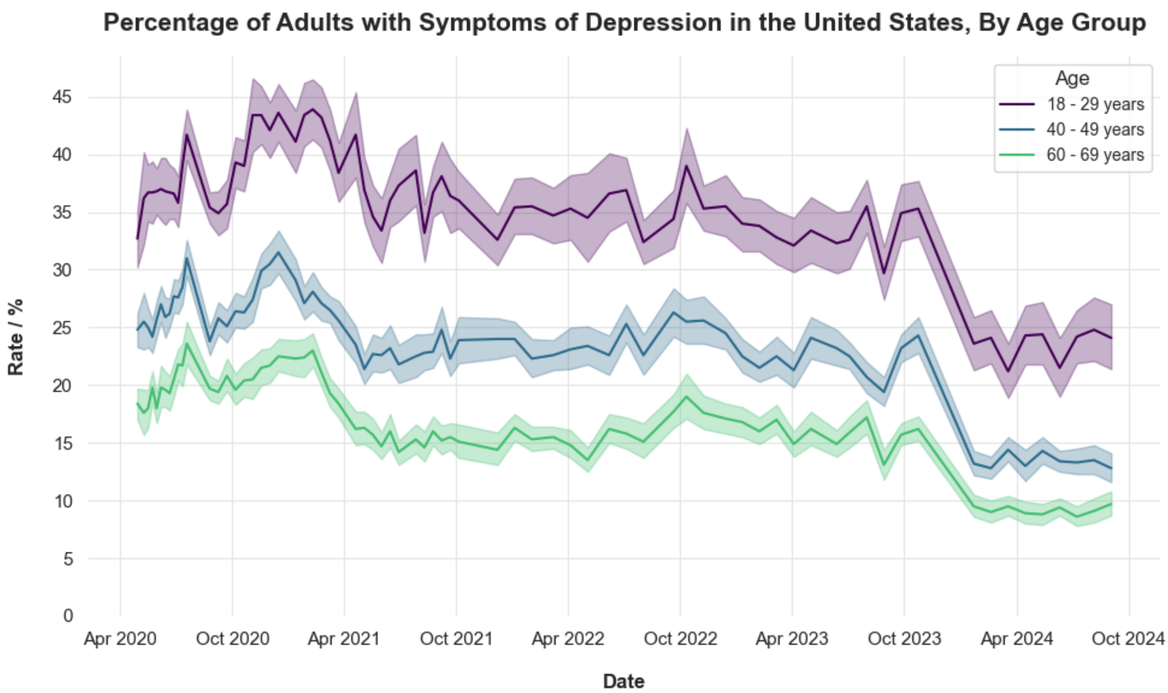

Fig. 1: Plot 1

To begin, I considered how the number of people with poor mental health has been evolving over time, across several age demographics. I found it important to do so, because I would then be better able to appreciate the current severity of the issue at hand, with respect to each of these groups of interest. In particular, given that depressive disorder is a prominent mental health condition, I looked into the proportion of adults suffering from symptoms of depression. To maximise the relevance of the visualisation and frame my work in a Singaporean context, I actively sourced for data from developed countries. Consequently, I opted to make use of a dataset from the United States Centers for Disease Control and Prevention (National Center for Health Statistics [NCHS], 2024). The CSV file consists of results from the Household Pulse Survey, which was conducted in regular time windows from 23 April

2020 to 16 September 2024. The data is able to describe the percentage of American adults who display depressive symptoms, across several age groups. To best illustrate how these rates changed with respect to the timeline of the survey, I opted to construct a multiple line graph as shown in Plot 1 (Fig. 1). The Viridis colour palette was used to ensure colour-blind friendliness. For each line, the confidence interval is represented by the faintly shaded region. Note that for each data point, the date was taken to be the midpoint of the corresponding time window.

In general, from May 2020 to July 2024, we see that the rates of depressive symptoms were decreasing across the three age groups shown. This can be attributed to the fact that the COVID-19 pandemic had been becoming increasingly more manageable in the United States, within this period of time. As the disease started to diminish in severity, more businesses and services were able to reopen, enabling the economy to recover steadily. The smaller degree of uncertainty – in terms of health and economy – meant that more people could become accustomed to a post-COVID way of life, in which they need not live in quarantine any longer. Hence, this might have reduced the levels of unhealthy stress and loneliness experienced by Americans, likely boosting their mental health. On top of this, there was an especially notable drop in rates at the end of 2023, around the Christmas festive season. This suggests that holiday periods could be another contributing factor towards improved mental health, given how people are able to spend more quality time with their loved ones, away from work and other sources of stress. We also observe that 18–29-year-olds are consistently more prone to developing symptoms of depression than 40–49-year-olds, who are in turn more at risk than 60–69-year-olds. With reference to the confidence intervals, we can conclude that this difference is significant. It is possible that younger adults experience more instability in their lives, stemming from events such as job searches, student loans and the formation of personal identities. This, coupled with their relative inexperience, could mean that they are less able to manage their emotions in an effective manner, leading to more harmful mental health outcomes. Ultimately, however, this hypothesis cannot be directly implied from the plot – more studies have to be conducted in order to verify that younger adults are indeed more vulnerable to depression, and for these specific reasons.

Despite the general trend observed, it may be inaccurate to conclude that poor mental health is becoming less serious of a problem. The chosen dataset is limited by the time period of the Household Pulse Survey, which only commenced in April 2020. We do not have data on pre-COVID depression rates. Hence, even though rates of depressive symptoms have declined since the initial COVID-19 outbreak, they may still be above pre-pandemic levels. Moreover, the latest rate for 18–29-year-olds remains concerningly high (around 1 in 4) and should be addressed. As for the older generation, it is entirely possible that their figures are simply under-reported, as a result of social stigma.

3. Plot 2

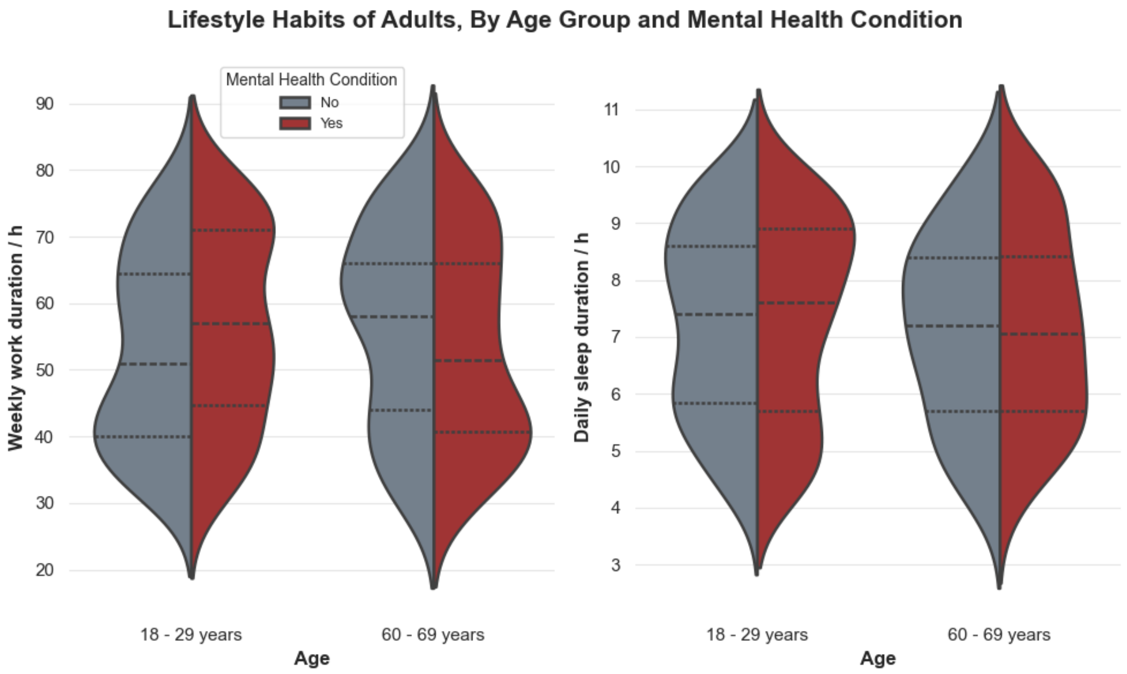

Fig. 2: Plot 2

From Plot 1, we know that a significant fraction of adults aged 18-29 suffers from symptoms of depression. Hence, I found it important to study this group in greater detail, so as to identify any key characteristics that may be linked to unsatisfactory mental health. In particular, I decided to compare the lifestyle habits of adults with and without mental health conditions, across two age brackets (18–29-year-olds and 60–69-year-olds). The 60–69-year-old age group was chosen in an effort to contextualise to Singapore's demographic composition. Owing to Singapore's rapidly ageing population, the growing number of senior citizens living in isolation is becoming an increasingly pertinent problem. If this is not managed well, Singaporean seniors may also be prone to suffering from elderly depression. Thus, it is worth contrasting the mental health profiles between these two subgroups. This would allow us to gain a better understanding of how poor mental health might potentially manifest, for these sections of the population. To visualise this, I made use of a dataset that contained information on 1000 individuals (Mohit, 2024). The CSV file describes the mental health status of these respondents, as well as several lifestyle factors – such as their weekly working hours and daily sleep duration. I designed a violin plot to compare and contrast the distributions of these variables, as shown in Plot 2 (Fig. 2). The quartiles of each distribution are marked with dotted lines.

For younger adults, we see that there is a clear difference in the weekly work duration, between individuals suffering from poor mental health and those who do not. The distribution of healthy 18–29-year-olds are slightly right-skewed, with approximately half of them working less than 50 hours a week. On the other hand, the corresponding quartiles of unhealthy 18–29-year-olds are all higher – around 25% of these people work more than 70 hours per week, which is excessively long. This shows that younger adults with poor mental health tend to have longer working hours. As for 60–69-year-olds, we see that healthy individuals have a more left-skewed distribution as opposed to those with mental health conditions. Hence, 60–69-year-olds with poor mental health tend to work for a shorter duration of time. One could possibly hypothesise that this group of people is mainly made up of retired seniors who struggle to keep themselves occupied with other activities. For these individuals, the consequent loss of social interaction and a sense of purpose may be a contributor towards mental health issues like geriatric depression. However, such a causal mechanism is not a conclusion that can be drawn directly from the plot, given that correlation is not equivalent to causation – more evidence is needed to verify these claims.

For 18–29-year-olds, the sleep distribution of unhealthy individuals are slightly more left-skewed, compared to those who are healthy. Furthermore, around 25% of 18–29-year-olds with mental health conditions sleep more than 9 hours a day, which is a sign of oversleeping. This suggests that younger adults with poor mental health generally have a higher tendency to oversleep. Looking at 60–69-year-olds, there is no obvious disparity in sleep duration between healthy and unhealthy seniors – the corresponding quartiles do not differ by a notable amount.

Thus, it appears that 18–29-year-olds with poor mental health tend to work more excessively and are more likely to oversleep, while 60–69-year-olds with mental health conditions tend to work for shorter durations. While we can draw these insights from the graph, we cannot conclude that these distinctions are indeed noteworthy. Moving forward, tests should be conducted in order to confirm that the differences in distributions are statistically significant. In addition, a limitation of this dataset is that it is imbalanced with respect to age group. In the final data frame that was used to create the violin plot, 63% of data points belonged to adults aged 18-29, while only 37% corresponded to 60–69-year-olds. To achieve a more insightful visualisation, the next step would be to obtain more samples from the latter age group – this would allow the final dataset to be balanced and more representative of these seniors.

4. Plot 3

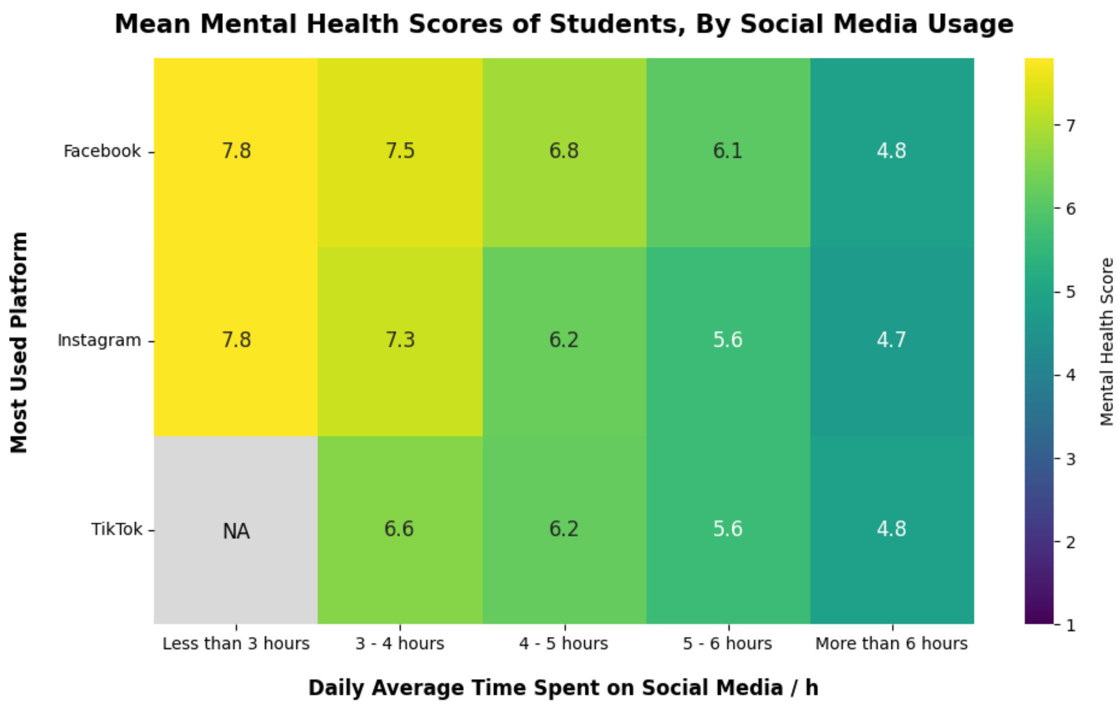

Fig. 3: Plot 3

From Plot 1, the reported proportion of 18–29-year-olds with symptoms of depression is higher than that of 60–69-year-olds. Hence, as a good first step, it would be wise to plan our intervention with these young adults as our target audience. From Plot 2, we now have a clearer picture of the behaviour of individuals with poor mental health, within this age bracket. Such learning points should be used to inform the message that we can bring across to these people, via our intervention efforts. Furthermore, our outreach should be maximised for 18–29-year-olds who are more vulnerable to suffering from worsening mental health. On this note, I decided to investigate whether there is a definite correlation between social media usage and mental health. Should this be the case, deploying our intervention on social media platforms would allow us to more effectively connect with at-risk young adults. Therefore, I opted to make use of a "Social Media Addiction" dataset, which contains the self-reported social media preferences and mental health scores of students aged 16-25 – similar to our demographic of choice (Shamim, 2025). To highlight intensity more effectively, I crafted a heat map that describes the severity of mental health issues, with respect to social media usage. This can be seen in Plot 3 (Fig. 3), in which the Viridis colour palette was used to ensure colour-blind friendliness. Note that for a given respondent, the duration of social media usage pertains to all platforms which the individual uses – not just his or her most commonly used platform.

Regardless of the most used platform, we see that the average score decreases as the duration spent on social media increases. Therefore, there is a negative correlation between social media usage and mental health – respondents who spend more time on social media tend to have poorer mental health. When Facebook is the most used platform, we also see that the mental health scores are consistently higher than (or equal to) those corresponding to Instagram and TikTok, for each time interval. In particular, at 3-4 hours, the mental health status for TikTok (6.6) is significantly worse than Instagram (7.3) and Facebook (7.5). As a result, respondents who spent most of their screen time on TikTok and Instagram tend to have poorer mental health. We can zero in on these platforms for our solution.

From this, we can structure our intervention as a social media campaign on Instagram and TikTok, for Singaporeans aged 18-29. A possible approach would be to collaborate with local mental wellness organisations – such as the Singapore Association for Mental Health – to produce creative Instagram Reels or short TikTok videos. In these videos, we can first mention relatable stress-inducing events that Singaporeans in this age bracket face – examples include academic difficulties in university, challenging job searches and early career troubles – in order to capture their attention. Subsequently, the later segments of these clips can shift towards raising mental health awareness and breaking any associated social stigma, whilst educating viewers about effective strategies to manage negative emotions. If we have access to a business account, we can also purchase advertisements on these platforms, in which users can be directed to the relevant support systems and helplines, should they require any such assistance. This would allow us to get in touch with as many of these young adults as possible, making it easier for the necessary help to be provided for those at risk, whilst also encouraging them to actively seek support.

However, a caveat of this visualisation is that correlation does not equate to causation. We cannot say that an increase in social media usage directly leads to the deterioration of an individual's mental health. More research has to be conducted to make sure that this causal link exists – this would make us more confident of our intervention's effectiveness. This dataset is also limited by the fact that it does not completely align with our desired age bracket of 18-29 years' old. Additional studies should aim to collect data that is exclusively within this range. Moreover, the dataset does not contain any respondents who not only has TikTok as their most used platform, but also uses social media for less than 3 hours a day. More data should be collected to account for this group of people.

5. Appendix

Please see below for a list of relevant links and references.

- GitHub repository: https://github.com/chiabingxuan/DSA4262

- Mohit, B. (2024). *Mental Health Dataset* [Data set]. Kaggle. https://www.kaggle.com/datasets/bhadramohit/mental-health-dataset/data
- National Center for Health Statistics (2024). *Indicators of Anxiety or Depression Based on Reported Frequency of Symptoms During Last 7 Days* [Data set]. Centers for Disease Control and Prevention. https://catalog.data.gov/dataset/indicators-of-anxiety-or-depression-based-on-reported-frequency-of-symptoms-during-last-7-
- Shamim, A. (2025). *Students' Social Media Addiction* [Data set]. Kaggle. https://www.kaggle.com/datasets/adilshamim8/social-media-addiction-vs-relationships