Investigating the Change in Smartphone Usage and Mental Health of STA304H5 Fall 2022 Students From Before To During COVID-19

By: Purumidha Sharma, Ruthba Tahiya, Saad Afridi, Osama Saleh, Marvin Roopchan, Mautasim Siddiqui, Abdur Imtiyas

STA304: Surveys, Sampling and Observational Data
Dr. Luai Al Labadi

November 29, 2022

Table of Contents

Introduction

In 2020, the world was faced with a pandemic that forced students to be indoors while participating in online learning. This caused students to change their lifestyle due to the restrictions that were put into place in Ontario. Due to the rapid change in student's lifestyles, we decided to conduct a study and use simple random sampling in lecture, tutorials and on Piazza to determine if there was a change in the average smartphone usage and mental health score for Fall 2022 STA304H5 students, before and during Covid-19. The results that can be obtained from this study are important as it will shed light on the possible effects of Covid-19 in students' lives and can be an important factor to consider when choosing a learning method for students.

Variables

Variable	Definition
Gender	Male, Female, or Non-Binary
Smartphone Usage before Covid-19	Measured in hours spent per day (0 - 8)
Smartphone Usage during Covid-19	Measured in hours spent per day (0 - 8)
Stress / Anxiety Level before Covid-19 (Inverted)	Indicates anxiety level from a range of 1 - 7, where 7 is very low stress/anxiety and 1 is very high
Stress / Anxiety Level during Covid-19 (Inverted)	Indicates anxiety level from a range of 1 - 7, where 7 is very low stress/anxiety and 1 is very high
Quality of Sleep before Covid-19	Indicates quality of sleep from a range of 1-7, where 7 is very good quality of sleep and 1 is very low
Quality of Sleep during Covid-19	Indicates quality of sleep from a range of 1-7, where 7 is very good quality of sleep and 1 is very low
Social Interactions Rating before Covid-19	Indicates quality of social interaction from a range of 1-7, where 7 is very good and 1 is very low
Social Interactions Rating during Covid-19	Indicates quality of social interaction from a range of 1-7, where 7 is very good and 1 is very low
Mental Health Score before Covid-19	Calculated as : Stress / Anxiety Level (inverted) + Social Interaction Level + Sleep Quality Score (all before Covid-19). Also, the mental health score is

	calculated per Student. Ranges from 3-21, where 3 indicates extremely poor mental health and 21 indicates extremely well mental health.
Mental Health Score during Covid-19	Calculated as : Stress / Anxiety Level (inverted) + Social Interaction Level + Sleep Quality Score (all during Covid-19). Also, the mental health score is calculated per Student. Ranges from 3-21, where 3 indicates extremely poor mental health and 21 indicates extremely well mental health.

Results

In this section, we will start by explaining and showing the necessary computations executed for our data analysis based on our survey data. Later, some of the findings and methodologies, including statistical tests used, visualizations, will be discussed. The main goal of this section is to describe the statistical significance of results obtained from our survey data analysis of the association among variables of interest related to our study's objective (as mentioned above).

Computations

In order to ask students to fill out our survey, we first needed to determine the number of participants to answer our survey in a manner that truly represents the student population of the STA304H5 Fall 2022 lecture. The number of participants required is called sample size, denoted by n for statistical analysis. We are performing simple random sampling (without replacement), equivalent to randomly selecting participants from our class. To calculate n, please refer to the following formula:

$$n = \frac{N\sigma^2}{(N-1)D+\sigma^2}, \text{ where } D = \frac{B^2}{4},$$

where σ^2 represents the variance, N represents the total number of Fall 2022 STA304 students, 2 and B represents the bound of error, which essentially tells you how far your estimate could be. In our scenario, we have N = 200 (students), since we don't know the σ^2 (generally unknown),

we calculate it through this formula: $\sigma^2 = \frac{Range\ (R)}{4} = \frac{8}{4} = 2$. We also want a bound of error of $B = 0.285$. Hence, we compute the following:

$$n = \frac{N\sigma^2}{(N-1)D+\sigma^2} = \frac{200(2)}{(200-1)(0.285^2/4)+(2)} = \frac{400}{(199)(0.0203)+(2)} \simeq 66$$

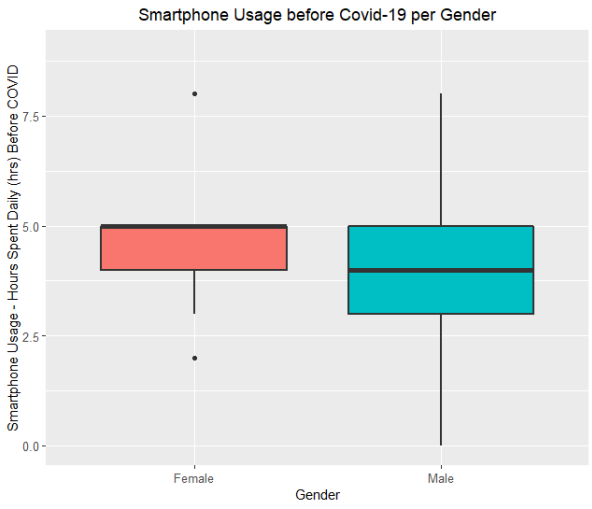
Therefore, we should sample 66 students. Fortunately, we were able to reach our target for a sample of 66 through mediums such as lecture/tutorials/Piazza.

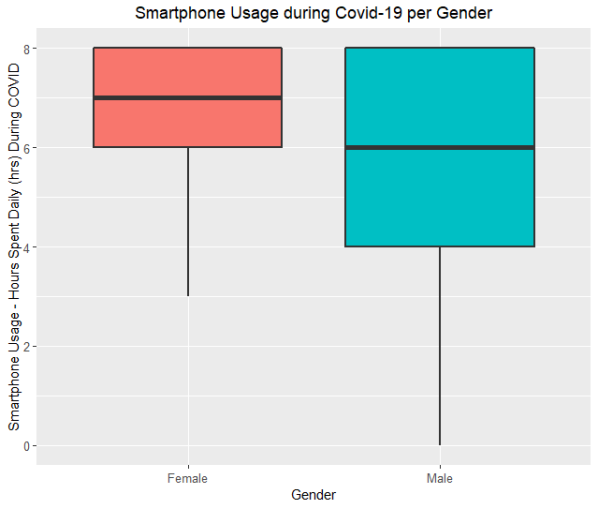
After collecting responses, we then computed the average smartphone usage before Covid-19 which was 4 hours, the average smartphone usage during Covid-19 which was 6 hours. Furthemore, the average mental health score before Covid-19 was 14 and the average mental health score after Covid-19 was 12.

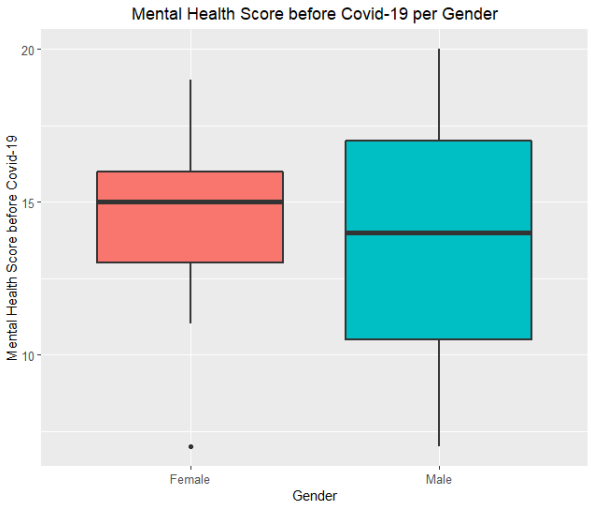
For the next section, we will discuss these results to see if there are any differences between computations that are irrelevant.

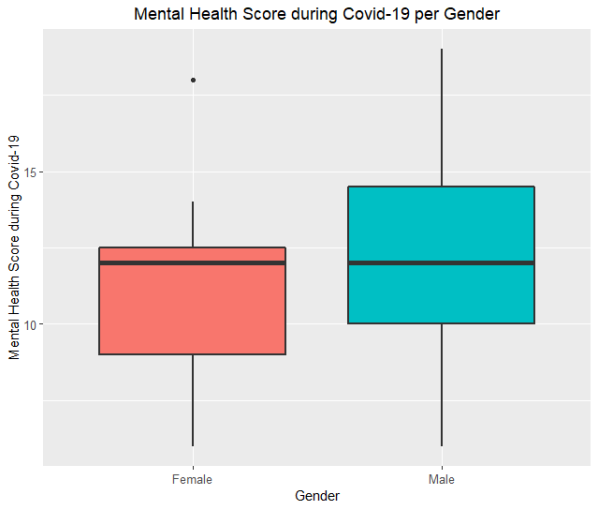
Discussion

To explore our outcomes more visually, we utilized the following plots. As seen below, we have two boxplots which portray smartphone usage (daily hours spent on smartphone) before Covid-19 and during Covid-19 and we, also, have two boxplots for the mental health scores of the students before Covid-19 and during Covid-19.









Looking at the figures above, we can notice that the average smartphone usage before Covid-19 and during Covid-19 changes. For example, for females, the median line for smartphone usage before Covid-19 points at 5 hours, but during Covid-19, it rises to 7. Moreover, we see a clear decrease in the mental health score from before Covid-19 and during Covid-19 for both genders. We do notice the evident changes with the plots; however, the plots are not sufficient to conclude our initial claims. To establish a better conclusion we used a type of statistical test, known as hypothesis testing. Hypothesis testing allows us to use data from samples to draw conclusions about the population in interest.

We, specifically, used a paired-sign test because it allows us to efficiently use our before and after sample data. We used the test to test out whether the average smartphone usage before Covid-19 and during Covid-19 differ and also, to check whether the mental health scores of the students before Covid-19 and during Covid-19 are different. Alongside our hypothesis tests, we calculated Cronbach's alpha, a coefficient that allows us to determine whether or not our questionnaire has internal consistency and if it is reliable. Cronbach's alpha is measured on a scale from 0 to 1.0 where a high interrelatedness amongst the questionnaire is depicted when values are from 0.7 to 1.0.

After our hypothesis tests and calculated Cronbach's alpha, we made the following observations:

1. We are 95% confident that the average smartphone usage before Covid-19 and during Covid-19 differ.

2. We are 95% confident that the mental health scores of the students before Covid-19 and during Covid-19 differ.
3. Our Cronbach's alpha depicts that, unfortunately, our questionnaire was inconsistent and unreliable since alpha was below 0.7.

Conclusion

The objective of this observational study was to examine whether Covid-19 had an impact on STA304H5 students' smartphone usage and their mental health. Our findings suggest that the average STA304H5 students' smartphone usage and the average mental health score differed from before Covid-19 to during Covid-19.

However, there were several limitations to this observational study that may have impacted the overall results. Majority of the participants were chosen during lectures and office hours, so we may have missed students that never attended in-person, indicating that our study might be using an inaccurate representation of the population as the students were conveniently chosen. Moreover, using Cronbach's alpha, we observed that there was unreliability and inconsistency with the questionnaire . If we had provided more of an incentive to answer the questionnaire, we may have gotten more responses, which results in a more comprehensive data set . Better yet, we could have improved the length of our survey with more mental health-related questions so that the reliability and consistency of our data were higher.

Appendix

R-Codes

Libraries

```
library(ggplot2)
library(ltm) # for cronbach.alpha
library(BSDA)
library(psych)
```

Data Processing

```
# load the data
dataset <- read.csv('c://FALL22/STA304/project/dataset.csv')
```

```
# High Stress/Anxiety => Bad
# High Q of sleep | High social interactions => Good

# Reverse Stress/Anxiety scores
dataset$Inversed.Stress.and.Anxiety.Level.Overall.Before.COVID.19..scale.fr
om.1.7. <-
dataset$Stress.and.Anxiety.Level.Overall.Before.COVID.19..scale.from.1.7.
dataset$Inversed.Stress.and.Anxiety.Level.Overall.During.COVID.19..scale.fr
om.1.7. <-
dataset$Stress.and.Anxiety.Level.Overall.During.COVID.19..scale.from.1.7.

dataset$Inversed.Stress.and.Anxiety.Level.Overall.Before.COVID.19..scale.fr
om.1.7.[dataset$Stress.and.Anxiety.Level.Overall.Before.COVID.19..scale.fro
m.1.7. == 7] <- 1
dataset$Inversed.Stress.and.Anxiety.Level.Overall.Before.COVID.19..scale.fr
om.1.7.[dataset$Stress.and.Anxiety.Level.Overall.Before.COVID.19..scale.fro
m.1.7. == 6] <- 2
dataset$Inversed.Stress.and.Anxiety.Level.Overall.Before.COVID.19..scale.fr
om.1.7.[dataset$Stress.and.Anxiety.Level.Overall.Before.COVID.19..scale.fro
m.1.7. == 5] <- 3
dataset$Inversed.Stress.and.Anxiety.Level.Overall.Before.COVID.19..scale.fr
```

```
om.1.7.[dataset$Stress.and.Anxiety.Level.Overall.Before.COVID.19..scale.fro
m.1.7. == 4] <- 4
dataset$Inversed.Stress.and.Anxiety.Level.Overall.Before.COVID.19..scale.fr
om.1.7.[dataset$Stress.and.Anxiety.Level.Overall.Before.COVID.19..scale.fro
m.1.7. == 3] <- 5
dataset$Inversed.Stress.and.Anxiety.Level.Overall.Before.COVID.19..scale.fr
om.1.7.[dataset$Stress.and.Anxiety.Level.Overall.Before.COVID.19..scale.fro
m.1.7. == 2] <- 6
dataset$Inversed.Stress.and.Anxiety.Level.Overall.Before.COVID.19..scale.fr
om.1.7.[dataset$Stress.and.Anxiety.Level.Overall.Before.COVID.19..scale.fro
m.1.7. == 1] <- 7

dataset$Stress.and.Anxiety.Level.Overall.Before.COVID.19..scale.from.1.7.
dataset$Inversed.Stress.and.Anxiety.Level.Overall.Before.COVID.19..scale.fr
om.1.7.

dataset$Inversed.Stress.and.Anxiety.Level.Overall.During.COVID.19..scale.fr
om.1.7.[dataset$Stress.and.Anxiety.Level.Overall.During.COVID.19..scale.fro
m.1.7. == 7] <- 1
dataset$Inversed.Stress.and.Anxiety.Level.Overall.During.COVID.19..scale.fr
om.1.7.[dataset$Stress.and.Anxiety.Level.Overall.During.COVID.19..scale.fro
m.1.7. == 6] <- 2
dataset$Inversed.Stress.and.Anxiety.Level.Overall.During.COVID.19..scale.fr
om.1.7.[dataset$Stress.and.Anxiety.Level.Overall.During.COVID.19..scale.fro
m.1.7. == 5] <- 3
dataset$Inversed.Stress.and.Anxiety.Level.Overall.During.COVID.19..scale.fr
om.1.7.[dataset$Stress.and.Anxiety.Level.Overall.During.COVID.19..scale.fro
m.1.7. == 4] <- 4
dataset$Inversed.Stress.and.Anxiety.Level.Overall.During.COVID.19..scale.fr
om.1.7.[dataset$Stress.and.Anxiety.Level.Overall.During.COVID.19..scale.fro
m.1.7. == 3] <- 5
dataset$Inversed.Stress.and.Anxiety.Level.Overall.During.COVID.19..scale.fr
om.1.7.[dataset$Stress.and.Anxiety.Level.Overall.During.COVID.19..scale.fro
m.1.7. == 2] <- 6
dataset$Inversed.Stress.and.Anxiety.Level.Overall.During.COVID.19..scale.fr
om.1.7.[dataset$Stress.and.Anxiety.Level.Overall.During.COVID.19..scale.fro
m.1.7. == 1] <- 7

dataset$Stress.and.Anxiety.Level.Overall.During.COVID.19..scale.from.1.7.
dataset$Inversed.Stress.and.Anxiety.Level.Overall.During.COVID.19..scale.fr
om.1.7.
```

```
# Create mental health score column
dataset$mental.health.score.before.COVID.19 <-
dataset$Inversed.Stress.and.Anxiety.Level.Overall.Before.COVID.19..scale.fr
om.1.7. + dataset$Quality.Of.Sleep.Before.COVID.19..scale.from.1.7. +
dataset$Rating.Social.Interactions.Before.COVID.19..scale.from.1.7.

dataset$mental.health.score.during.COVID.19 <-
dataset$Inversed.Stress.and.Anxiety.Level.Overall.During.COVID.19..scale.fr
om.1.7. + dataset$Quality.Of.Sleep.During.COVID.19..scale.from.1.7. +
dataset$Rating.Social.Interactions.During.COVID.19..scale.from.1.7.
```

Assumptions

```
# Normality test on smartphone diff
smartphone_differences =
abs(dataset$Hours.Spent.Daily.On.Smartphone.Before.COVID.19..0.8.hours. -
dataset$Hours.Spent.Daily.On.Smartphone.During.COVID.19..0.8.hours.)
shapiro.test(smartphone_differences)
```

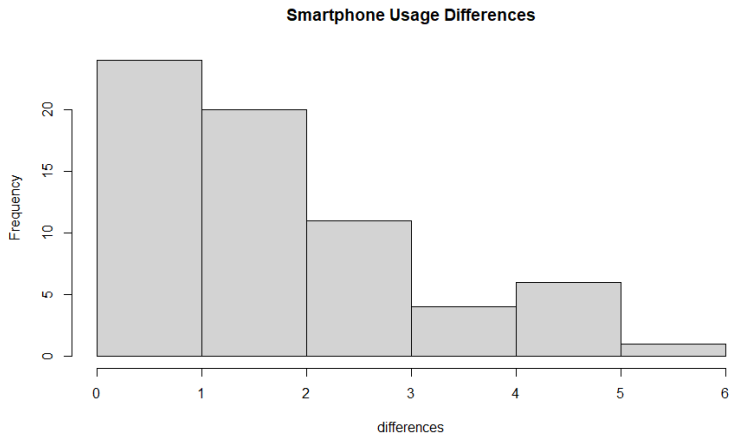
```
        Shapiro-Wilk normality test

data:  smartphone_differences
W = 0.89872, p-value = 5.601e-05
```

```
# Normality test on  mental health diff
mental_differences = abs(dataset$mental.health.score.before.COVID.19 -
dataset$mental.health.score.during.COVID.19)
shapiro.test(mental_differences)
```

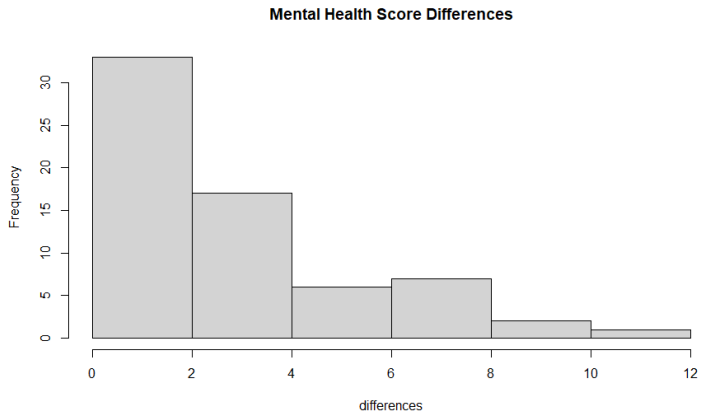
```
        Shapiro-Wilk normality test

data:  mental_differences
W = 0.88544, p-value = 1.814e-05
```

```
# Graphs for differences distribution
smartphone_before=dataset$Hours.Spent.Daily.On.Smartphone.Before.COVID.19..
0.8.hours.
smartphone_during=dataset$Hours.Spent.Daily.On.Smartphone.During.COVID.19..
0.8.hours.
```

```
differences = abs(smartphone_before-smartphone_during)
hist(differences, breaks="Scott", main="Smartphone Usage Differences")
```




```
differences = abs(dataset$mental.health.score.before.COVID.19 -
dataset$mental.health.score.during.COVID.19)
hist(differences, breaks="Scott", main="Mental Health Score Differences")
```



Paired-Sign Test for Hypothesis

```
# Paired sign test for pairs diff means (two-sided, !=)
SIGN.test(x =
dataset$Hours.Spent.Daily.On.Smartphone.Before.COVID.19..0.8.hours.,
          y =
dataset$Hours.Spent.Daily.On.Smartphone.During.COVID.19..0.8.hours.,
          alternative = "two.sided",
```

```
          conf.level = 0.95)

SIGN.test(x = dataset$mental.health.score.before.COVID.19,
          y = dataset$mental.health.score.during.COVID.19,
          alternative = "two.sided",
          conf.level = 0.95)
```

```
        Dependent-samples Sign-Test

data:  dataset$Hours.Spent.Daily.On.Smartphone.Before.COVID.19..0.8.hours.  and
 dataset$Hours.Spent.Daily.On.Smartphone.During.COVID.19..0.8.hours.
S = 1, p-value = 3.482e-13
```

```
        Dependent-samples Sign-Test

data:  dataset$mental.health.score.before.COVID.19 and dataset$mental.health.sc
ore.during.COVID.19
S = 42, p-value = 0.0008618
```

Graphs

```
# mental health score before covid boxplot per gender
g_box_mental_bfr<- ggplot(dataset, aes(x=Gender,
y=mental.health.score.before.COVID.19, fill=Gender)) +
  geom_boxplot(lwd=1) +
  ylab("Mental Health Score before Covid-19") +
  xlab("Gender") +
  theme(legend.position = "none") +
  theme(axis.text.x = element_text(size = 9.5)) +
  theme(axis.text.y = element_text(size = 9.5)) +
  theme(plot.title = element_text(hjust = 0.5)) +
  ggtitle("Mental Health Score before Covid-19 per Gender")

g_box_mental_bfr
```

```
# mental health score during covid per gender
g_box_mental_dur <- ggplot(dataset, aes(x=Gender,
y=mental.health.score.during.COVID.19, fill=Gender)) +
  geom_boxplot(lwd=1) +
  theme(legend.position = "none") +
  labs(color = "Gender") +
  ylab("Mental Health Score during Covid-19") +
```

```
  theme(plot.title = element_text(hjust = 0.5)) +
  ggtitle("Mental Health Score during Covid-19 per Gender") +
  xlab("Gender") +
  theme(axis.text.x = element_text(size = 9.5)) +
  theme(axis.text.y = element_text(size = 9.5))

g_box_mental_dur
```

```
# smartphone hours before covid per gender boxplot
g_box_smart_bfr <- ggplot(dataset, aes(x= Gender, y =
Hours.Spent.Daily.On.Smartphone.Before.COVID.19..0.8.hours., fill=Gender))
+
  geom_boxplot(lwd=1) + ylim(c(0,9)) +
  theme(legend.position = "none") +
  labs(color = "Gender") + ylab("Smartphone Usage - Hours Spent Daily (hrs)
Before COVID") +
  xlab("Gender") +
  theme(axis.text.x = element_text(size = 9.5)) +
  theme(axis.text.y = element_text(size = 9.5)) +
  theme(plot.title = element_text(hjust = 0.5)) +
  ggtitle("Smartphone Usage before Covid-19 per Gender")

g_box_smart_bfr
```

```
# smartphone hours during covid per gender boxplot
g_box_smart_dur <- ggplot(dataset, aes(x= Gender, y =
Hours.Spent.Daily.On.Smartphone.During.COVID.19..0.8.hours., fill =
Gender)) + geom_boxplot(lwd=1) + scale_fill_discrete("Gender",
labels=c('Female', 'Male')) + labs(color = "Gender") + ylab("Smartphone
Usage - Hours Spent Daily (hrs) During COVID") +
  xlab("Gender") + theme(axis.text.x = element_text(size = 9.5)) +
  theme(axis.text.y = element_text(size = 9.5)) +
  theme(plot.title = element_text(hjust = 0.5)) +
  theme(legend.position = "none") +
  ggtitle("Smartphone Usage during Covid-19 per Gender")

g_box_smart_dur
```

Advanced Methodologies

```
# Getting cronbach alpha of multi q's for mental health before covid
mental_health_qs_before_covid <- data.frame(

stress_before_covid=dataset$Inversed.Stress.and.Anxiety.Level.Overall.Before.COVID.19..scale.from.1.7.,

q_of_sleep_before_covid=dataset$Quality.Of.Sleep.Before.COVID.19..scale.from.1.7.,

social_before_covid=dataset$Rating.Social.Interactions.Before.COVID.19..scale.from.1.7.

)

cronbach.alpha(mental_health_qs_before_covid)
```

```
Cronbach's alpha for the 'mental_health_qs_before_covid' data-set

Items: 3
Sample units: 66
alpha: 0.522
```

```
# Getting cronbach alpha of multi q's for mental health during covid
mental_health_qs_during_covid <- data.frame(

stress_before_covid=dataset$Inversed.Stress.and.Anxiety.Level.Overall.During.COVID.19..scale.from.1.7.,

q_of_sleep_before_covid=dataset$Quality.Of.Sleep.During.COVID.19..scale.from.1.7.,

social_before_covid=dataset$Rating.Social.Interactions.During.COVID.19..scale.from.1.7.

)

cronbach.alpha(mental_health_qs_during_covid)
```

```
Cronbach's alpha for the 'mental_health_qs_during_covid' data-set

Items: 3
Sample units: 66
alpha: 0.2
```

Questionnaire

Link: https://forms.gle/1aEvTcEPfH37eQvi8

Technical Report

Link: https://docs.google.com/document/d/1GG-gENA0Z4xMx-uTp_KdUJlG1sniHx719ZF6UiOgIB8/edit?usp=sharing