2022

Analyzing Behavioral Adaptation to COVID-19 And Return To Pre-Pandemic Baselines in a Cohort of College Seniors

Vlado Vojdanovski
Dartmouth College, vlado.vojdanovski.22@dartmouth.edu

Analyzing Behavioral Adaptation to COVID-19 And Return To Pre-Pandemic Baselines in a Cohort of College Seniors

Undergraduate Computer Science Honors Thesis
Dartmouth College

Vlado Vojdanovski

June 1, 2022

Contents

Abstract

As the critical phase of the COVID-19 pandemic seems to be winding down, it is important to analyze the adjustment to COVID-19 and return to normalcy of various populations. In this study we focus on the behavioral adjustments exhibited by a cohort of N=114 college seniors. To infer COVID-19 adjustment we compare the 2021 year (second year of COVID-19) to the 2020 year (first year of COVID-19) and 2019 (pre-pandemic baseline year). We begin with a broad analysis between the second and first covid year, finding that the second year of COVID-19 shows significant returns to pre-pandemic baselines on multiple sensing features. Further, we run statistical comparisons between the terms of Fall 2020 (lockdown fall), Fall 2019 (pre-covid fall) and Fall 2021 (post-lockdown fall) and note statistically significant differences between Fall 2021 and Fall 2019 on four variables of interest. We find that activity variables surpass their pre-pandemic baseline, while smartphone usage variables still lag in their return. This suggests that disruptions to physical activity are easier to correct for, whereas smartphone and technology use display more permanent shifts once disrupted. We then use a multivariate forecasting method trained on Fall 2019 to forecast the entirety of Fall 2021, yielding an average Mean Absolure Relative Range Error of 12.15 indicating similarity between the terms. Finally, we perform a clustering analysis to understand whether there are any differences in how students react to the omicron and delta waves of COVID-19. One of our clusterings returns a cluster of students with a delayed return to baseline, while the other returns a few outlier students that exhibit dramatic shifts in behavior around the time the Omicron variant appears.

1 Introduction

In this study, we seek to examine and quantify the adaptation of students to the COVID-19 pandemic and to understand the dynamics of their return to baseline as measured by mobile sensing. We define adaption to broadly encompass all the shifts in behavior that occur during the COVID-19 pandemic. More precisely, we examine how key mobile sensing variables evolve over the second year of the COVID-19 pandemic. Return to baseline is defined as the narrowing of behavioral differences between the year 2021 and a pre-pandemic baseline represented with the year 2019. We find this to be a relevant topic for several reasons. Firstly, the disruptions caused by the COVID-19 pandemic were great both on the societal level as well as on the individual level. Understanding which behaviors return, or exceed, their respective baseline and which don't can tell us a lot about the malleability of human behavior. Further, analyzing the pace at which said variables return to pre-pandemic levels is interesting in-itself as it gives context to how quickly people can recover from pandemics or similarly shocking events. Finally, analyzing this data of college students allows us to make comparison with other groups of interest and highlight differences in future studies. Our study is driven by the following broad exploratory research questions:

- ***(Q1)*** *How does student behavior change during the second year of the COVID-19 pandemic?*
- ***(Q2)*** *Is the 2021 Fall term statistically similar to the last pre-pandemic fall?*
- ***(Q3)*** *Based on the results of* ***(Q2)*** *can we build a model to forecast the 2021 Fall?*
- ***(Q4)*** *Is there significant variance in how students react the Delta and Omicron waves?*

This study utilizes passively collected mobile sensing data as collected in the StudentLife dataset. As such, we are less reliant on using exclusively self-reports and are able to offer a fine-grained, detailed analysis of the second COVID-19 year and students' gradual return to pre-pandemic levels and patterns of activity. The smartphone app used in this work has been validated in a number of prior clinical studies [1, 2, 3]. As the study has been ongoing since 2017, we are able to directly compare and contrast the behavior of our entire cohort before, during, and after the COVID-19 pandemic. In addition, we collect Ecological Momentary Assessments from our cohort, enabling us to better understand and potentially connect the behavioral changes we see to various mental health features. Our contributions are as follows:

- To our knowledge this is the first and only study that analyzes the return to pre-pandemic baselines in any population after the disruption caused by the COVID-19 pandemic.
- This is also the first paper that shows the impact (or lack thereof) of the delta and omicron waves on college students, as evidenced by shifts in behavior.
- We find that the second year of COVID-19 is marked by a significant increase in physical activity, locations visited, and lower phone usage compared to the first year of COVID-19. We quantify this by comparing the 2021 and 2020 years in our longitudinal dataset.
- We show that pre-pandemic behavioral data can be used to train a model to forecast data nearly two years removed in time. Our model performs significantly better than a Naive baseline model, highlighting the potential for transfer learning of behaviors despite large longitudinal gaps.
- We perform clustering on the delta and omicron waves, identifying a cluster of students with a lagged return to their pre-pandemic baseline and a small number of students that exhibit significant decreases in activity during the omicron wave.

The structure of this paper is as follows. In Section 2 we cover the work that has already been done to shed light on this important topic. Following this, we detail our study design, ground truth and dataset in Section 3 and report on the analysis on behavioral change of students in Section 4. Given that our results seem to imply that aggregate 2021 behaviors become very close to their 2019 levels around the Fall Term of 2021, we perform statistical tests between the two time periods. After finding generally encouraging results, we build a deep-learning forecasting model to predict 2021 Fall data using the 2019 Fall as a training set. We then proceed to analyze the Delta and Omicron waves in Section 7, in an attempt to answer the question of whether all students react to the second and third COVID-19 wave the same way. Using cluster analysis, we find that the delta wave is experienced in similar ways among students, with the main difference being the pace of return to normalcy – with one cluster seemingly lagging behind by around a month. When clustering during the omicron wave, we find 3 outlier students who show significant decreases in their activity levels. We discuss our findings in Section 8, address the limitations of this study in Section 9 and present concluding remarks in Section 10.

2 Related Work

Our study comes in a long line of work aimed at using mobile sensing for assessment and analysis of human behavior. Such work has been used to evaluate mental health, personality, and even workplace performance. However, despite the wide usage of sensing data in studies, very few studies on the COVID-19 pandemic actively use smartphone data. Those that do are mostly concerned with contact tracing and tracking, and there has been a great amount of work done in that field [4, 5, 6, 7]. Most studies focused on the pandemic, though instrumental in advancing our understanding of the issues besetting the population, have only utilized self-reports [8, 9, 10]. This presents an issue for two reasons, firstly self-reports strongly rely on the participants memory, and secondly the absence of sensing data does not allow for a multi-modal comparison between self-reports and behavior.

However, there have been a few studies combining self-reports and sensing. Notably, researchers have analyzed the relationship between behavior and COVID news-coverage [11], behavior and COVID fatigue [12], and most recently, the links between mental health issues, COVID-19 concern and behavior [13]. A study by Sañudo et al. [14] also analyzes the initial reactions to the COVID-19 pandemic, noting a decrease in physical activity and increases in sleep duration. Other works have explored the relationships between having the COVID-19 disease and behavior, noting a 0.72 AUROC score in predicting COVID-19 concern [13]. The same study found significant decreases in walking duration, biking duration, number of unique locations visited as well as increases in sedentary time, sleep duration and a later sleep start time for students when comparing the 2020 year to the 2019 year. This strongly suggests that 2020 was still heavily impacted by the COVID-19 pandemic as the study did not find any behaviors that remained at their baseline levels. Work by Sun et al. [15] further validates many of the findings above.

Notably, all of these studies focus on the negative effects that the COVID-19 pandemic had on human behavior. That is, they are analyzing either the months immediately after COVID-19 or, as in Nepal et al. [13] the year following COVID-19. Our study is the first to provide insight into the behavior changes after the release of the vaccines and the advent of the delta and omicron waves.

3 Study Methodology

3.1 Study Design

This study uses data from a longitudinal mobile study tracking N=220 College Students through their time in college. The dataset in question collects both mobile sensing data and self reports. We analyze N=114 of these students for the purposes of this paper. This is because the initial cohort of 106 people has already graduated college and thus their participation in this study has ended. All participants in this study were asked to install a continuous mobile sensing app on their iPhone or Android. They were also asked to keep the app running for the entire duration of their four years at the College, including academic breaks. As the goal of this study is to track student mental health, the participants were compensated $10 per a week for answering a set of Ecological Momentary Assessments regarding their mental health each week. A set of optional COVID-19 EMAs were introduced in the study after the start of COVID-19.

3.2 Demographics

Table 1 shows the demographics of the 114 students used in our analysis. The majority (66.1%, N=74) of our participants identify as females. In terms of race, 59.8% (N=67) are White, 27.7% (N=31) are Asians, 3.6% (N=4) are Black or African American, 1.8% (N=2) are American Indian/Alaska Native and 5.4% (N=6) belong to more than one race.

3.3 Ecological Momentary Assessments

This study utilizes self-reported data on stress, depression, anxiety and social level through EMAs. We use the 4-item Patient Health Questionnaire (PHQ-4) for our depression and anxiety measurements. We use the question "Are you feeling stressed now" with a 5 point Likert scale ranging from "Not at All" to "Extremely" for our stress measurement. We use three questions from the State Self-Esteem Scale to measure self-esteem levels with a question from the social, performance, and esteem categories [16]. Social levels ("Have you spend most of your time alone or with others today?") are measured on a 5-point Likert Scale ranging from "Almost Always Alone" to "Almost Always With Others". In addition, as noted above, we added optional COVID-19 EMAs to gauge the impact of the pandemic on student concern levels, social media usage, and

Table 1: Demographics of the participants. The table below lists the demographic composition of the students in our study.

Category	**Count**	**Percentage**
Sex		
Female	74	66.1%
Male	38	33.9%
Race		
White	67	59.8%
Asian	31	27.7%
Black or African American	4	3.6%
American Indian/Alaska Native	2	1.8%
More than one race	6	5.4%
Not reported	2	1.8%

behavioral trends. Psychologists in our research team developed these questions to minimal time to answer as to maximize response rate. Students were asked to respond to the COVID related EMAs once a week at a random time, but could also answer them at any time of their choosing, much like the EMAs noted previously. For this study, we mostly rely on the first question of our COVID EMAs "How concerned are you about COVID-19?" as we find it to be the broadest and best at capturing group-level dynamics. The use of this question is mostly during our clustering segment, where we cluster during periods of increasing COVID-19 concern to attempt to capture shifts in behavior. All COVID EMAs are detailed in Table 8.

3.4 Features

Some of the features we collect using our passive sensing app are listed below. Note that we collect these features hourly, as well as group them into epochs by summing several hours. Thus, epoch 1 covers the period between 9 am and 7 pm, epoch 2 covers the period between 7 pm and 1 am, and epoch 3 covers the period between 1 am and 9 am. Epoch 0 is, thus, used to denote summed data over the entire day. All the features used in this paper are described in Table 7.

Phone Usage. We record the number of phone unlocks that the participants make as well as the duration between phone unlocks and locks, inferring phone duration usage. As such we are able to have a proxy for screen time. Researchers find that phone usage is correlated with depressive symptoms and anxiety [2, 17].

Mobility. We sample GPS location of users every 10 minutes and use this information to derive the number of unique locations visited with the DBSCAN [18] algorithm, as well as distance distance travelled. Mobility features from mobile phones relate to anxiety and depression, based on several prior works [17, 19].

Physical Activity. We identify which physical activity the user is engaging in (ie walking, running, biking, etc) using APIs provided by the phone manufacturers.

Sleep. We derive sleep duration, sleep start, and sleep end based on the method described in [20, 21].

Semantic locations. We identify the home duration of participants by tracking the location where they spend their nights. We further use geo-fencing of areas on campus to determine various locations visited, including study areas, exercise areas, and others. Researchers have found associations between location types visited and mental health [22, 23].

4 Q1: How does student behavior change during the second year of the COVID-19 pandemic?

We begin our analysis by first exploring the change in behavior of the participants as a result of COVID-19. As we note in our methodology, the dataset we are working with is longitudinal and thus contains data from both before and during the COVID-19 pandemic.

We begin by comparing aggregate daily mobile sensing features between year one of COVID-19 (March 2020 - March 2021) and year two of COVID-19 (March 2021 - March 2022). To do so, we perform a Wilcoxon signed rank test between each day in the time period. A Benjamini-Hochberg False Discovery Rate correction was performed on the p-values. The results are shown in Figure 1.

As can be seen in Figure 1, all features that had significantly declined during the first year of COVID-19 as compared to our pre-pandemic baseline, have now significantly increased. Walking during all periods of the day sees strong increases (+29%) particularly in epoch 1 (+44%) and epoch 3 (+34%). We further note decreases in overall sedentary behavior (-4%) and in particular sedentary behavior in epoch 3(-6%). The later could potentially be a result of the reopening of bars and restaurants throughout 2021. Students also visited significantly more unique locations per day as compared to the pandemic period (+56%). Paradoxically, we also note that we see a decrease in running activity during the entire day (-21%) while also noting an increase in biking activity (+40%). Seeing as similar studies have found running to have suffered an initial decrease during the first year of COVID-19 [13], we expect this to be an indication that some may have dropped running as a hobby. Out of all the variables we tested, only the differences sleep duration and the unlock number during epoch 3 remained statistically insignificant. Since the sleep duration of the second COVID-19 year is elevated compared to the baseline year, it makes sense that the smartphone unlock number at night remains at a lower level as well.

Following this comparison between the first and second year of COVID-19, the question arises of how the second COVID-19 year compares to a pre-

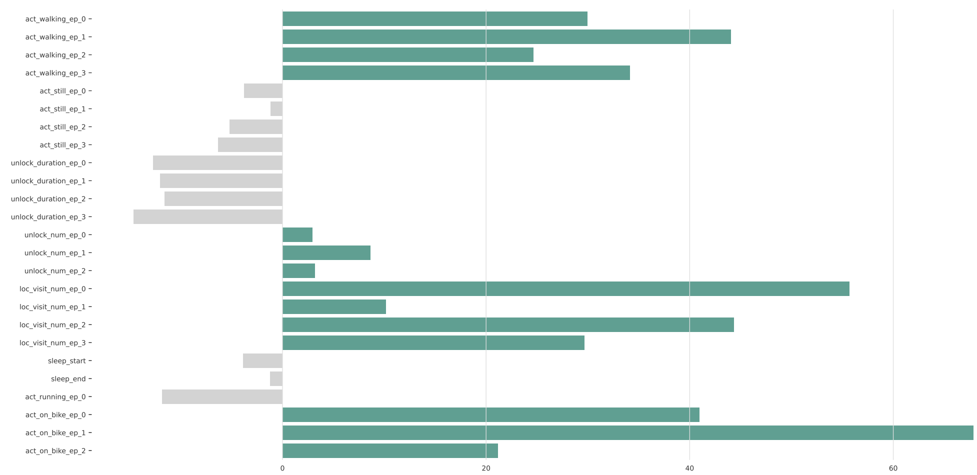

Figure 1: Features with significant differences between pre-COVID-19 baseline year vs first COVID-19 year. Green bars indicate a feature that has increased during the time period while gray bars indicate a features that has decreased. The value is derived by comparing the feature means, while the significance is calculated using a non parametric paired t-test. The x-axis is the percentage change and the y-axis is the feature name. All results presented in this figure meet the following criteria: significant with a p-value of less than 0.01 after correcting for multiple comparisons using the Benjamini/Hochberg FDR correction procedure.

pandemic baseline. To determine whether we are observing a return to a pre-pandemic level of activity, we plot the difference between the mean daily values for the 2021 and the 2019 years. Our results are shown in Figure 2.

As we can see, the 2021 year begins with a lower level of walking activity than the 2019 baseline but is trending toward a return to it's pre-pandemic baseline levels. We see similar trends toward baseline in both the unlock duration (a noticeable decrease) and number of phone unlocks (a significant increase). By the time that the Fall Term of 2021 starts (the first fully in person term since the start of the pandemic), the student population exhibits behaviors near baseline. It is noteworthy, however, that despite the great rebounds in behavior, we still see a somewhat increased overall unlock duration and a somewhat lowered unlock number for the participants. At the same time, the features tracking activity seem to have nearly fully rebounded. This hints that physical activity may have faster dynamics and return to a baseline faster than other variables tracked. Finally, it is interesting to note that the number of locations visited also

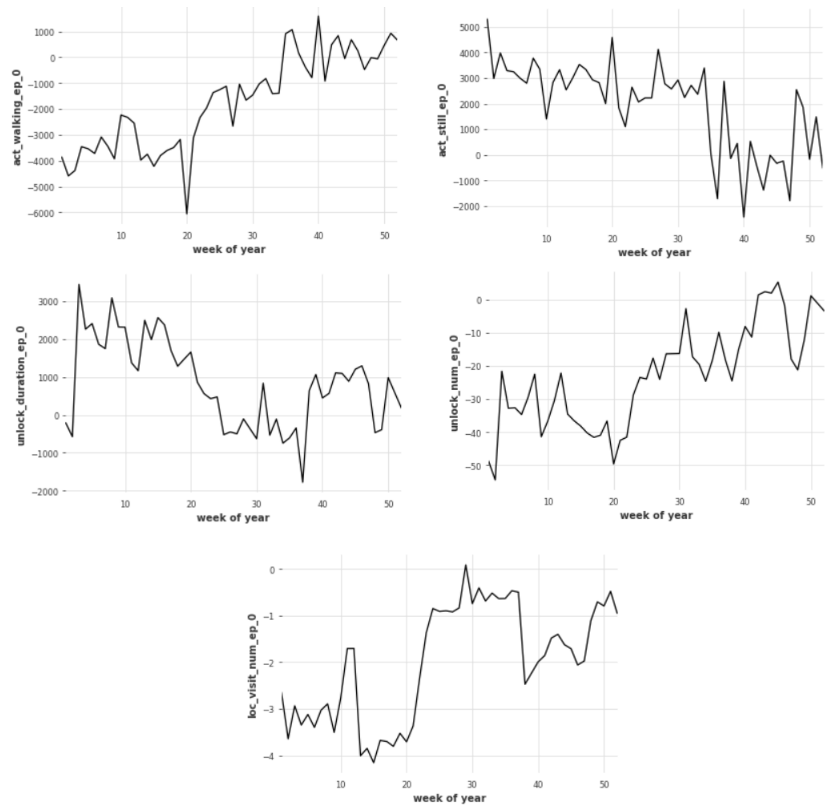

Figure 2: The figure above shows the difference between the daily mean aggregated values for four different features: walking, sedentary time, unlock duration, and unlock number between the years 2019 and 2021. Note: Weeks 29 and 51 of year 2019 suffered a software glitch on the unlock duration and number so that over 30 % of the respondents were incorrectly reported as zeroes. As such, these years are ommited from the graph and interpolated.

seems to be slightly lower through the year, but sees a further, drastic decrease once students return to campus for the fall term of 2021. This could indicate that despite being relative mild, the restrictions the college imposed that term had a noticeable effect on student behavior. For example, for the duration of fall 2021 the college fitness center had a strictly enforced mask mandate which could've yielded less frequent visits to the gym. In fact, 65 students spent at least one session at the gym (defined as exercising for 40 or more minutes) in the fall of 2019, while the respective number for 2021 was 31. This implies that casual gym users were less likely to utilize the athletics facilities than previous terms. We also saw decreases in time spent in other students' dorms as well as study and social spaces (excluding Greek spaces). Some of this is explained by the fact that the college was not hosting any social events that it had prior to the pandemic, potentially driving the social scene more-so toward the Greek houses. The college was, at the time, also not hosting any guest speakers, further reducing the number of unique locations that students could visit. As the mask mandate in the library was enforced significantly less stringently, at least some of the shift in behavior may be due a desire to spend more time outdoors, perhaps as an over-correction to the pandemic lockdowns. The behaviors may also be explained by students exploring the surrounding area more and eating outdoors.

Aside from effects of college policies, we may note that neither the Delta, nor the Omicron wave seem to have had any discernible population-wide effects on the behavioral trend as captured by mobile sensing. Also notable is that the inoculation campaign among the student body (approximately weeks 20-25 of 2021) seems to have had an impact on walking duration, unlock duration and unlock number but not the level of sedentary behavior. Tracking aggregate behavior like this may be too broad, so we attempt to more finely model the period when we note a relative return to baseline activity which seems to be around the time of the Fall 2021 Term at the College.

5 Q2: Is the 2021 Fall term statistically similar to the last pre-pandemic fall?

Prompted by the observation of a return to baseline above, we seek to answer the question of whether Fall of 2021, the first fully in-person term since the start of the pandemic, constitutes a return to baseline for the cohort on the key variables of physical activity and cellphone usage. As we noted above, some variables still display significant shifts (such as the exact locations the students visit) and note that the number of unique locations visited seems to be lowered by college policies. Thus, due to these observed difference as well as the narrowness of some of these variables, they are not included in this analysis where we attempt to provide measurements of broad behavioral patterns.

We aim to realize this objective by analyzing some of the main mobile sensing features that were affected by the COVID-19 pandemic and seem to have most reliably returned to pre-pandemic levels. We look at daily walking duration, daily sedentary time, phone unlock duration and number of phone unlocks per day. To better capture the trends in the data, we perform weekly aggregation per user and week of term. We present visual representations of the 2021 Fall term (post-lockdown Fall), the 2020 Fall Term (lockdown Fall), and 2019 Fall Term (pre-covid Fall) in Figure 3 . The lines represent the mean value for each term whereas the standard error is represented by the shaded area around the lines. Interestingly, the sedentary time and walking duration seem to be somewhat closer in value to baseline than the unlock duration or unlock number. To test for statistical significance, we used linear mixed effects models fit by log-likelihood as implemented in the lmer package [24]. We built separate models

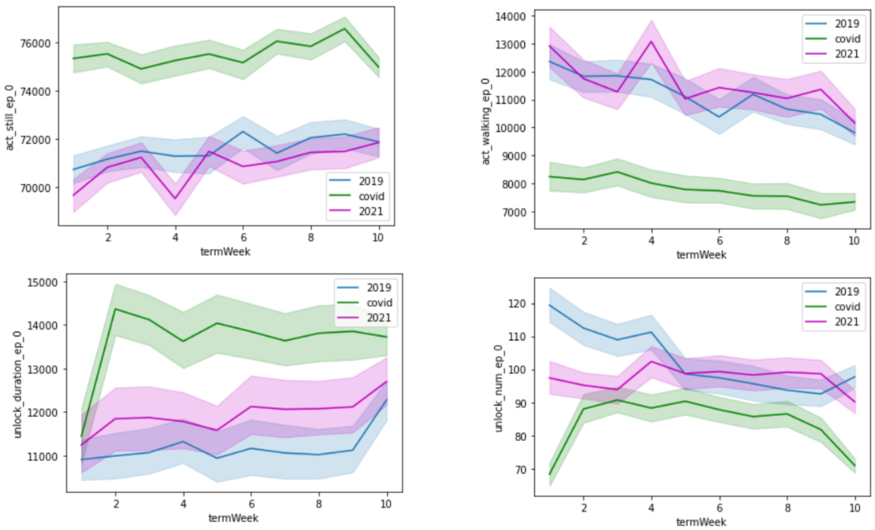

Figure 3: Figure representing plots for sedentary time, walking duration, phone unlock duration, and phone unlock number for each of the 2019,2020, and 2021 Fall terms. The x-axis represents the week of the term and the y-axis represents the value measured by the user's smartphone

to compare the post-lockdown fall and pre-covid fall on each feature. All the models included a binary factor, 'is_fall_2021' to label whether the term is is the fall of 2021. All models also included a linear 'termWeek' variable indicating the week of the term. Our first model included just the variables above. The next model included an additional interaction between the is_fall_2021 variable the the term of the week. Our third model added a quadratic term for the term of the week to the second model. Finally, the fourth model added an additional interaction between the fall indicator variable and the quadratic term of the week variable. For each variable, the model with the lowest deviance, tested with ANOVA to statistical significance, was selected. P values were calculated using the Satterthwaite method as implemented in lmerTest [25], accounting for the fact that we have multiple observations. The results for the models comparing the post-lockdown fall to the pre-covid fall are shown in Table 2.

We note in the table below that we find statistical differences between the two falls on nearly all the features of interest. Notably, some of these differences indicate the waning effects that COVID-19 has on the student population, while others may indicate lingering effects of the pandemic. In particular, we note that while walking duration and sedentary duration are different than during the pre-covid fall they show a positive improvement. That is, we see a higher walking duration in the post-lockdown fall, and a lower overall sedentary time. This is interesting in itself, as it could imply a desire to be more outdoorsy and active after a year in lockdown. On the opposite end, the unlock duration and unlock

Table 2: Table summarizing the results from the linear mixed effects model comparing Fall 2021 and Fall 2019

Variable	Dependent Variables							
	Walking Duration		Sedentary Duration		Unlock Num		Unlock Duration	
	parameter	p-value	parameter	p-value	parameter	p-value	parameter	p-value
is_fall_2021	3.0e-02	<0.001***	-1.7e-02	0.005 **	-4.8e-02	<0.001 ***	1.1e-02	0.002 **
termWeek	-7.7e-02	<0.001 ***	4.4e-02	<0.001 ***	-7.2e-02	<0.001 ***	1.2e-02	0.028 *
is_fall_2021*termWeek	NA	NA	NA	NA	7.3e-02	<0.001 ***	NA	NA
termWeek^2	NA	NA	NA	NA	3.2e-01	0.005 **	NA	NA
is_fall_2021*termWeek^2	NA	NA	NA	NA	-4.9e-01	0.004 **	NA	NA

number seem to still display patterns of behavior introduced by COVID-19. The unlock number feature is lower compared to the pre-covid term, and unlock duration is higher. These were patterns we noted when COVID-19 originally appeared, and it is interesting to see that they still remain. Overall, it seems that mobility features and smartphone unlock duration exhibit more similar dynamics over both periods, with smartphone unlock number lagging behind. This could imply that changes in social media and smartphone use may be harder to rectify than changes to physical behavior.

6 Q3: Based on the results of *(Q2)* can we build a model to forecast the 2021 Fall?

Given the complex relationship between the two time periods above, we ask whether a deep model can extract the necessary information to forecast key features of the post-lockdown fall by learning them on the pre-covid fall, nearly two years before the forecasting date. This question is of relevance due to the fact that statistical similarity tells us little on whether the behaviors can be learned and whether they are transferable. In fact, we note that the two period above are statistically different in a plethora of ways. However, we can also visually see that the two time periods still share some similar dynamics. Effectively, we aim to answer to questions. The first is whether these time periods are similar enough that one can forecast the other, and the second is whether human behavior can be forecast successfully despite a large longitudinal gap.

To fully utilize the seasonality of our data, as well as the fact that we are working with a longitudinal time series, we transform the three epochs of each feature we are forecasting into a single feature. That is, instead of having columns for the morning, afternoon, and evening epoch, we now have a single column containing all three sequentially. Motivated by the statistical analysis above, we choose to predict the users walking duration, sedentary time, unlock duration, and unlock number.

The specific task that we aimed to accomplish was a forecasting attempt on the entirety of Fall of 2021, using data from the 20 days prior to the first day of classes. Note that all data was scaled between 0 and 1 to aid convergence.

6.1 Model Selection

The period between September 16, and November 27, 2019 was split into a training and validation period by a split of 0.8:0.2. That is, 80 percent of the days were used for model training and 20 percent of the days were used for model validation. We then tested several models and tuned hyper-parameters accordingly. All calculations were done with the random seed set to 0. It should be noted that models were trained in a multivariate way, attempting to learn the relationship between all four variables of interest at once. However, no aggregation was done. That is, every user represented an individual time series and the model was not trained on the mean values of the dataset.
The model that performed best on the validation set was a BlockRNN model as implements in the Darts package [26]. This model is comprised of an RNN block that serves as an encoder for the input, and a FCN (Fully Convolutional Network) that produces the fixed size output as a forecast. The parameters that worked optimally were 20 days as the predictor input (that is 60 datapoints) producing 10 days (30 datapoints) as the predictor output. The RNN used was a vanilla RNN. The model was trained for 30 epochs, with the learning rate stepping down from 1×10^{-3} to 1×10^{-4} at epoch 20.
After model selection, the model was re-trained on the entire fall 2019 period with an additional 20 days before September 16, 2019. As such our training period ranges from August 27^{th}, to November 27^{th} of 2019. As a result, the forecasting task was to predict the behavior of the above noted variables in Fall 2021 based on the 20 days prior the beginning of that term.

6.2 Choosing a baseline

To choose a baseline to test against we tested several Naive models. These included an ARIMA model, an Exponential Smoothing Model, a NaiveMean model, a NaiveSeasonal model, and an Ensemble Model containing both the NaiveMean and the NaiveSeasonal models. Out of the above, the NaiveSeasonal with a K of 3 (that is, repeating the last 3 values for the length of the test set) performed best. The ARIMA and Exponential Smoothing models performed up to 4 times worse than the NaiveSeasonal model. This serves to show that it is reasonably difficult to perform better than a naive repetition in this dataset, even with a statistical state of the art such as ARIMA.

6.3 Model Performance

As the tests were ran on the individual time-series of the participants, we were able to run statistical tests on whether the differences in distributions are significant. The metric we used to evaluate the two models was MARRE (Mean

Average Relative Range Error) as define in the darts package [26]. The reason for this metric was that since our scaling was done on a zero to one scale, using the Mean Average Percentage Error would yield artificially high, inaccurate values. The results are shown in the table below. The model achieves

Table 3: Table summarizing the differences in performance between the deep model and the Naive Baseline

Predicted Variable	MARRE		p-value
	Baseline	BlockRNN	
walking duration	17.7	12.9	1.01e-5
sedentary duration	14.2	9.1	1.44e-9
unlock duration	18.0	13.8	7.96e-6
unlock number	15.9	12.8	0.002

a significant improvement over the best naive baseline, indicating that behavioral dynamics from the pre-pandemic fall are useful in inferring behavior in Fall 2021. This lends further credibility to the claim that Fall 2021 represents a return to baseline behavior for our cohort and also showcases that learned behavioral patterns can be useful in forecasting even across long time-periods. Even the unlock number, which had the strongest statistical differences between the terms, sees a significant improvement when utilizing the deep model.

7 Q4: Is there significant variance in how students react the Delta and Omicron waves?

The data in Figure 2 implies that the omicron and delta waves, albeit causing an increase in overall covid-related concern, do not impact the cohort's overall return to baseline on the cellphone usage and activity level variables. However, we are still left with the question of whether different student groups react differently to the waves. Thus, we set out to explore these potential difference through means of cluster analysis. In particular, we attempt to cluster based on the initial reaction and adjustment to the different waves. For the purposes of this analysis, we define the adjustment period of the delta wave to start at the local minima of the COVID-1 self-reported concern nearest to when a delta case was registered in the U.S, and end once the COVID-1 concern hits a peak. If two peaks were present, we chose the latter one to maximize series length. The omicron wave adjustment period is defined similarly, and both are visualized in Figure 4. Such a definition was most likely to capture any behavioral differences. To capture a reaction to COVID-19, changes in behavior could be as relevant as the behaviors themselves. This approach yielded series of length 42 days for the delta period, and 51 days for the omicron period.

Important to note is that, throughout the pandemic, the number of unique users reporting their COVID-19 concern is continually dropping. Seeing as one reason for lack of reporting could be considering COVID-19 to not be a relevant part of one's life, the COVID-19 concern levels for the delta and omicron waves may be artificially high. This point is further reinforced by the relatively stable answer rate to the PHQ-4 questionnaires. Where the number of unique people answering COVID-19 questionnaires drops from 50 to 33 (-32%) in the 2021 year, the number of people answering the COVID-19 EMA drops from 83 to 64 (-23%). The response number curves are visualized in Figure 5 Nevertheless, the concern trajectories still remain relevant.

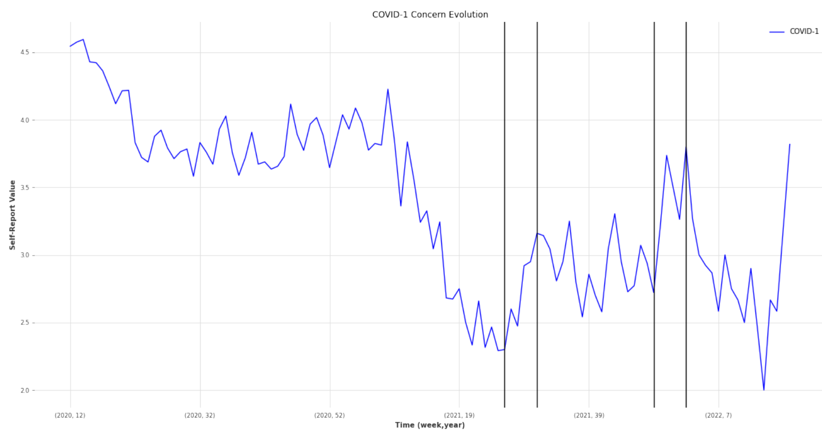

Figure 4: The evolution of the mean self-report value of the EMA "How concerned are you about COVID?" from March 2020 to February 2022. Denoted with black lines are what the author has defined to be the delta adjustment period (first pair of black lines) and the omicron adjustment period (second pair of black lines). The sudden spike of concern toward the end of the graph is likely due to the extremely small sample-size responding that week (9 people).

To perform the clustering we used the TimeSeriesKMeans algorithm as implemented in [27], and scaled the data before clustering. Since we used a euclidean distance metric, and are thus looking at the entire series and not averages, we decided to only keep the students that had data for the entire range we were analyzing. This left us with 58 students for the delta period and 52 students for the omicron wave. The features over which we were clustering were all three epochs of walking time, sedentary time, unlock duration, and unlock number. For this analysis, we kept the epochs as separate variables with the assumption that treating as separate series could yield more accurate results. We tried a number of clusters of 2,3,4 for both the delta and omicron periods. We get silhouette scores of (0.12 0.07, 0.08) and (0.4, 0.07, 0.05) respectively. For the delta period our clustering profile is summarized in Table 4.

We observe that the distribution of participants between the clusters is slightly skewed, with cluster one having 62% of the students. As we can see, cluster zero displays higher walking duration, lower sedentary duration and a higher smartphone unlock number. Interestingly, the cluster also has a higher average response to the "How much have you supported others?" COVID EMA

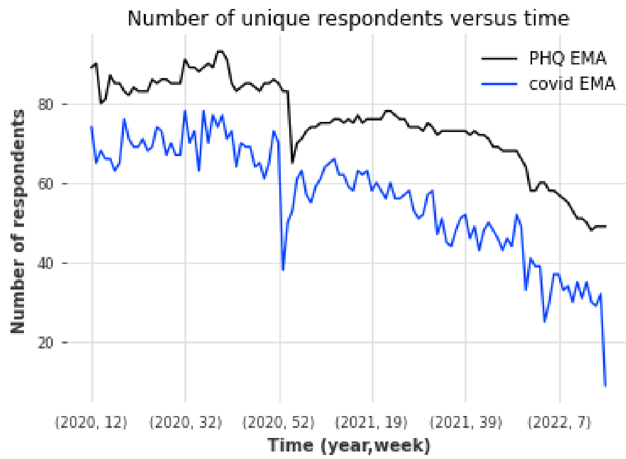

Figure 5: Number of participants answering specific EMAs for each week starting in March 2020

during this period. Such a result seems to be in line with the fact that support-giving has been associated with lower inflamatory markers as well as exhibits a calming effect on the nervous system [29, 30]. Further, when analyzing the clusters' multi-year behavioral trajectory, cluster one seems to have a slightly delayed return to baseline levels of activity as compared to cluster zero. This is particularly noticeable in the morning sedentary levels and morning walking duration, where cluster one does not exhibit a full return to normalcy even by the end of the 2021 calendar year. Another relevant fact about the discovered clusters is the observed increase in differences in behavior of the clusters. That is, initially, the clusters share some difference but are still somewhat close. When the COVID-19 pandemic hits, however, we see this difference significantly increase and then fail to fully return to previous levels. We quantify this observed difference using the Dynamic Time Warping Metric in Table 5. We choose to calculate a DTW (Dynamic Time Warping) distance to calculate the distance between the time series before the COVID-19 pandemic hits and then a second DTW distance after the onset of the pandemic. We choose the DTW metric here as we wish to allow for a more robust, non-linear definition of distance. A notable exception to this increase in difference are the variables representing unlock number. This seems to be due to a significant difference between the clusters in Fall of 2018, when the study was not fully enrolled. We also visualize the differences between the clusters in some key features in Figure 6.

Before proceeding to analyze the clustering during the omicron period we also summarize the gender distribution of the two clusters. Cluster 1, the cluster with a delayed return to normalcy, is gender imbalanced with 29 out of the 36 students being female, 7 being male. Cluster 0 has 12 male students and 10 female students. This implies that cluster 1 is over-represented in the number of

Feature	Mean Value Cluster Zero	Cluster One	Significance
act_walking_ep_0	1.64*10^4	9.6*10^3	<0.001
act_walking_ep_1	2.1*10^3	1.3*10^3	<0.001
act_walking_ep_2	8.0*10^3	5.0*10^3	<0.001
act_walking_ep_3	6.2*10^3	3.4*10^3	<0.001
act_still_ep_0	6.7*10^4	7.3*10^4	<0.001
act_still_ep_1	3.0*10^4	3.1*10^4	0.004
act_still_ep_2	2.2*10^4	2.5*10^4	0.04
act_still_ep_3	1.4*10^4	1.7*10^4	<0.001
unlock_num_ep_0	1.1*10^2	6.8*10	<0.001
unlock_num_ep_1	1.3*10	9.3	0.015
unlock_num_ep_2	5.9*10	3.7*10	<0.001
unlock_num_ep_3	4.0*10	2.3*10	<0.001
COVID-9	5.07	3.58	0.011

Table 4: Clustering results for the delta adjustment period. P-values were calculated by comparing the participant averages over the time period using the Mann-Whitney U non-parametric test. Al p-values are corrected using the Benjamini-Hochberg FDR Procedure. [28]

females, while cluster 0 is over-represented in the number of males. This may be an indication that women may be more likely to feel the impacts of COVID-19 for longer. This is in line with research indicating that females were more likely to show higher anxiety as a result of COVID-19 [8]. The racial make-ups of the two clusters seems to be similar, however we include .

The clustering for the omicron period yielded different results. The clusters were quite unequally balanced, with cluster containing 3 students and cluster zero the other 49. The clusters had significant differences in both mobile sensing features as well as EMAs. Cluster one had a significantly higher walking duration over all three epochs as well as a significantly lower sedentary duration over the epochs. However, cluster one also exhibited a significantly higher cellphone unlock duration and cellphone unlock number. Cluster one was also statistically more anxious, depressed, and stressed. We thus suspect that the higher level of walking duration may be a protective mechanism that the students in cluster one have developed to guard off their significantly heightened anxiety and depression levels. Unfortunately, only a single person from cluster one filled out any COVID-19 EMAs making a statistical comparison of COVID-19 concern difficult. The COVID-19 EMAs were, however, higher for cluster one than

Table 5: Table showcasing the distance between the two clusters identified over the delta period for each time series, before and after COVID-19

Feature	**Pre-COVID DTW Distance**	**COVID DTW Distance**
act_walking_ep_0	0.17	0.2
act_walking_ep_1	0.05	0.09
act_walking_ep_2	0.21	0.24
act_walking_ep_3	0.34	0.41
act_still_ep_0	0.21	0.24
act_still_ep_1	0.06	0.10
act_still_ep_2	0.28	0.39
act_still_ep_3	0.37	0.46
unlock_duration_ep_0	0.14	0.23
unlock_duration_ep_1	0.05	0.07
unlock_duration_ep_2	0.19	0.34
unlock_duration_ep_3	0.24	0.35
unlock_num_ep_0	0.37	0.22
unlock_num_ep_1	0.09	0.06
unlock_num_ep_2	0.27	0.22
unlock_num_ep_3	0.29	0.17

cluster zero. Effectively, our clustering algorithm served as an outlier detection algorithm, flagging three students who may have had a significant reaction to the news of the omicron wave. The most notable behavior change that we note is that while cluster zero shows no significant changes in behavior, cluster one shows a decrease in walking duration of over 30% and an over 15% increase in sedentary time. The stress levels of members of cluster one also show some increase. There do not seem to be any relevant changes to the unlock duration or unlock number for either cluster. These changes in behavior were not present during the 2019 year and coincide with increasing news coverage of the omicron wave. These findings are visualized in Figure 7 and summarized in Table 6

8 Discussion

In this paper, we examine the behavior of a group of college seniors (N=114) through the second year of COVID-19. We cover both the omicron and the delta waves as well as the first fully in-person term for the cohort. While sev-

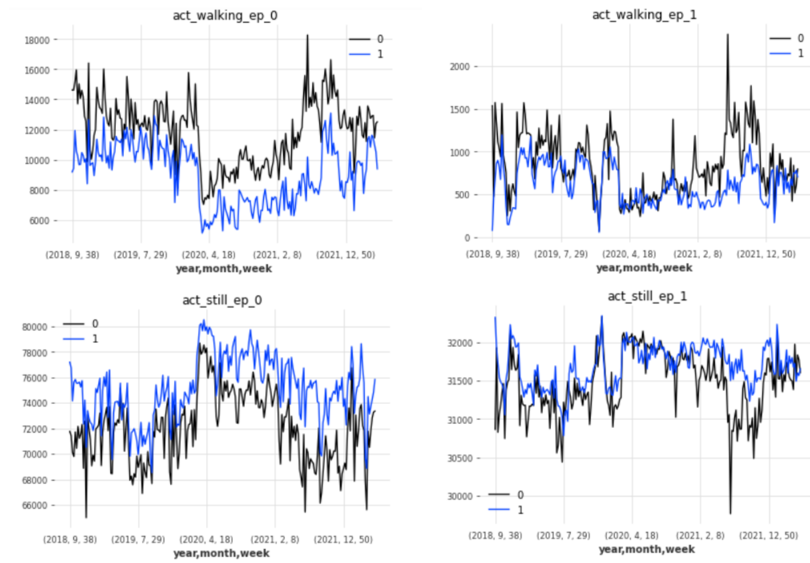

Figure 6: Behavior of the mean of each cluster over the entire study. P-values were calculated by comparing the participant averages over the time period using the Mann-Whitney U non-parametric test. Al p-values are corrected using the Benjamini-Hochberg FDR Procedure.

eral studies have analyzed the effects of COVID-19, all of them have focused on the effects of the initial strain of the virus on human behavior. This is the first paper that analyzes the aftermath of the COVID-19 pandemic. We use a longitudinal dataset spanning several years, allowing us to directly compare the second year of COVID-19 to a 2019 baseline, noting the behavioral changes that occur. We find that students' behavior shows dramatic shifts, and significant returns, and in some cases improvements, to pre-pandemic baseline levels. We find significant increases in walking duration, reductions in sedentary time, drastic increases in number of locations visited and changes to sleeping behavior offsetting the shifts that occurred during the first year of COVID-19. We also analyze cohort aggregate values finding a significant shift in behavior immediately after the vaccination efforts among the college cohort (April and May) with an improvement in walking duration and sedentary time compared to a pre-pandemic baseline. We also note a lag in returning to baseline in the unlock duration, and unlock number variables. We also find that the number of unique locations visited is a variable particularly affected by college policies.
In addition to the above analyses, we perform clustering on the omicron and delta waves to discover whether there are any differences in how students react to the waves. Our clustering during the rise of concern in the delta period leads us to discovering two clusters of students exhibiting significant differences in behavioral trends. One cluster exhibits a faster return to baseline in physical

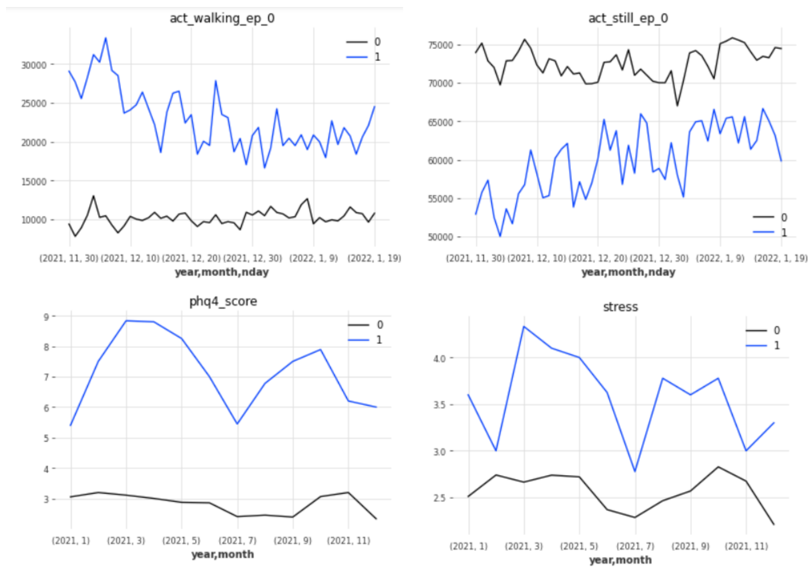

Figure 7: Behavior of the means of the two omicron clusters during the omicron adjustment period.

activity, namely walking duration and sedentary time while the other has partial exhibits delays in this return. We also observe that behavioral differences between the clusters become exacerbated after COVID-19 pandemic hits, which presents a highly interesting result. Our clustering during the omicron period produces three students with the most drastic shift in behavior, differing in all of clustering variables as well as self-reported anxiety and overall PHQ-4. Unfortunately, the number of COVID-19 concern EMAs is too low for this time period to establish whether significant differences exist between the groups.

9 Limitations

This work's main limitation is the narrowly defined nature of the population being examined. That is, we only posses data from a small number of college seniors at a small North-Eastern university. As such, we lack the ability to compare the adjustments between cohorts of different age-groups, which would yield a stronger analysis. Nevertheless, our work still presents an informative picture of the cohort that we are studying, allowing for comparisons to be made by different studies tracking different cohorts.

Another downside of this study is the lack of access to behavioral data through smartwatches or other, more precise, tools aimed at capturing behavioral shifts. We are working with a number mostly-iPhone users which limits the scope of the data we are able to collect. Finally, as the cellphone manufacturers do not provide regular updates on the performance on the algorithms aimed at inferring

	Mean Value		Significance
Feature	Cluster Zero	Cluster One	
act_walking_ep_0	10232.563621	22939.869281	<0.001
act_walking_ep_1	1062.808928	2575.091503	0.016
act_walking_ep_2	5963.671673	12307.424837	<0.001
act_walking_ep_3	3206.083020	8057.352941	<0.001
act_still_ep_0	72572.942845	59869.431373	<0.001
act_still_ep_1	31032.756362	29516.006536	0.02
act_still_ep_2	23948.897789	17841.450980	0.0012
act_still_ep_3	17591.288694	12511.973856	<0.001
unlock_duration_ep_0	11775.359735	17962.529209	0.039
unlock_duration_ep_2	5979.536208	9636.287111	0.02
unlock_num_ep_0	75.471840	178.568627	<0.001
unlock_num_ep_1	8.486859	20.640523	0.021
unlock_num_ep_2	41.342511	92.098039	<0.001
unlock_num_ep_3	25.990822	66.287582	0.02
phq_anx	1.244444	3.142857	0.021
phq4_score	2.437037	5.571429	0.021
stress	2.351852	3.357143	0.047
phq4-1	0.718519	1.500000	0.049
phq4-2	0.525926	1.642857	0.02
phq4-4	0.540741	1.428571	0.024

Table 6: Clustering summary results for the omicron adjustment period. P-values were calculated by comparing the participant averages over the time period using the Mann-Whitney U non-parametric test. Al p-values are corrected using the Benjamini-Hochberg FDR Procedure.

activities, we are also at a disadvantage with regard to the ability to quantify the errors we are getting from the measurements themselves. Nevertheless, given that these algorithms are tuned on large datasets, we can infer a somewhat high accuracy.

10 Conclusion

This paper analyzed the evolution of student COVID-19 concern over the second year of the pandemic as well as the behavioral trends that occurred in that time period. We analyze these trends broadly, and then proceed to make a narrower statistical comparison between the fall terms of 2021 and 2019 to determine whether we note a significant shift to baseline activity. We then use a deep learning forecasting model to test whether fall 2019 can be used to forecast the fall of 2021. Finally, we perform clustering on both the omicron and delta rise-of-concern periods and find clusters of students that have differing reactions during those periods.

We believe that this study opens the door many future avenues of research. We note the possibility of analyzing the differences in behavior between different colleges, that is colleges that loosened restrictions later or sooner than the college in question. This would elaborate on the connection between the local policies, college policies, and student behavior. Then, there is the opportunity to compare the behaviors of this dataset to the behaviors of people in more advanced age, as well as potentially clinical populations. We note that students with higher PHQ-4 scores seems to have a stronger shift in behavior during the omicron wave. Exploring this further, within a diagnosed population, could yield further knowledge of stressors response between different populations to COVID-19. Even a study of an older, non-clinical population, would add further context to this study.

11 Appendix

Table 7: Table describing the most frequently used features in the paper

Feature	Description
act_walking_ep_0	Total walking duration over entire day
act_walking_ep_1	Total walking duration between 9am and 7pm
act_walking_ep_2	Total walking duration between 7pm and 12am
act_walking_ep_3	Total walking duration between 12an and 9am
act_still_ep_0	Total sedentary duration over entire day
act_still_ep_1	Total sedentary duration between 9am and 7pm
act_still_ep_2	Total sedentary duration between 7pm and 12am
act_still_ep_3	Total sedentary duration between 12an and 9am
unlock_duration_ep_0	Total unlock duration over entire day
unlock_duration_ep_1	Total unlock duration between 9am and 7pm
unlock_duration_ep_2	Total unlock duration between 7pm and 12am
unlock_duration_ep_3	Total unlock duration between 12an and 9am
unlock_num_ep_0	Total number of phone unlocks over entire day
unlock_num_ep_1	Total number of phone unlocks between 9am and 7pm
unlock_num_ep_2	Total number of phone unlocks between 7pm and 12am
unlock_num_ep_3	Total number of phone unlocks between 12am and 9am
loc_visit_num_ep_0	Total unique locations visited the entire day
loc_visit_num_ep_1	Total unique locations visited between 9am and 7pm
loc_visit_num_ep_2	Total unique locations visited between 7pm and 12am
loc_visit_num_ep_3	Total unique locations visited between 12am and 9am
act_on_bike_ep_0	Total biking duration over the entire day
act_on_bike_ep_1	Total biking duration between 9am and 7pm
act_on_bike_ep_2	Total biking duration between 7pm and 12am
act_on_bike_ep_3	Total biking duration between 12am and 9am
act_running_ep_0	Total running duration over the entire day
act_running_ep_1	Total running duration between 9am and 7pm
act_running_ep_2	Total running duration between 7pm and 12am
act_running_ep_3	Total running duration between 12am and 9am

References

[1] Dror Ben-Zeev, Rachel Brian, Rui Wang, Weichen Wang, Andrew T Campbell, Min SH Aung, Michael Merrill, Vincent WS Tseng, Tanzeem Choudhury, Marta Hauser, et al. Crosscheck: Integrating self-report, behavioral sensing, and smartphone use to identify digital indicators of psychotic relapse. *Psychiatric rehabilitation journal*, 40(3):266, 2017.

Table 8: COVID EMAs

Question
COVID
1. How concerned are you about COVID-19?
2. How much has the COVID-19 situation impacted your day to day activities in the last week?
3. How much have you changed your behaviors in response to the COVID-19 situation in the last week?
4. How concerned are you for yourself regarding COVID-19?
5. How concerned are you for your classmates regarding COVID-19?
6. How concerned are you for your family regarding COVID-19?
7. How concerned are you about obtaining food, supplies, etc.?
8. How supported do you feel?
9. How much have you supported others?
10. Is your social media usage:
1 (much less than normal) - 7 (much more than normal)

[2] Rui Wang, Weichen Wang, Alex daSilva, Jeremy F Huckins, William M Kelley, Todd F Heatherton, and Andrew T Campbell. Tracking depression dynamics in college students using mobile phone and wearable sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(1):1–26, 2018.

[3] Rui Wang, Weichen Wang, Min SH Aung, Dror Ben-Zeev, Rachel Brian, Andrew T Campbell, Tanzeem Choudhury, Marta Hauser, John Kane, Emily A Scherer, et al. Predicting symptom trajectories of schizophrenia using mobile sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3):1–24, 2017.

[4] Tania Martin, Georgios Karopoulos, José L. Hernández-Ramos, Georgios Kambourakis, and Igor Nai Fovino. Demystifying COVID-19 digital contact tracing: A survey on frameworks and mobile apps. *Wireless Communications and Mobile Computing*, 2020:1–29, October 2020.

[5] Enrique Hernandez-Orallo, Pietro Manzoni, Carlos Tavares Calafate, and Juan-Carlos Cano. Evaluating how smartphone contact tracing technology can reduce the spread of infectious diseases: The case of COVID-19. *IEEE Access*, 8:99083–99097, 2020.

[6] Katherine Klise, Walt Beyeler, Patrick Finley, and Monear Makvandi. Analysis of mobility data to build contact networks for COVID-19. *PLOS ONE*, 16(4):e0249726, April 2021.

[7] Miklos Szocska, Peter Pollner, Istvan Schiszler, Tamas Joo, Tamas Palicz, Martin McKee, Aron Asztalos, Laszlo Bencze, Mor Kapronczay, Peter Petrecz, Benedek Toth, Adam Szabo, Attila Weninger, Krisztian Ader,

Peter Bacskai, Peter Karaszi, Gyozo Terplan, Gabor Tuboly, Adam Sohonyai, Jozsef Szoke, Adam Toth, and Peter Gaal. Countrywide population movement monitoring using mobile devices generated (big) data during the COVID-19 crisis. *Scientific Reports*, 11(1), March 2021.

[8] Jeremy A Bigalke, Ian M Greenlund, and Jason R Carter. Sex differences in self-report anxiety and sleep quality during covid-19 stay-at-home orders. *Biology of sex Differences*, 11(1):1–11, 2020.

[9] Cornelia Herbert, Alia El Bolock, and Slim Abdennadher. How do you feel during the covid-19 pandemic? a survey using psychological and linguistic self-report measures, and machine learning to investigate mental health, subjective experience, personality, and behaviour during the covid-19 pandemic among university students. *BMC psychology*, 9(1):1–23, 2021.

[10] Elizabeth V Goldfarb. Participant stress in the covid-19 era and beyond. *Nature Reviews Neuroscience*, 21(12):663–664, 2020.

[11] Jeremy F Huckins, Alex W DaSilva, Weichen Wang, Elin Hedlund, Courtney Rogers, Subigya K Nepal, Jialing Wu, Mikio Obuchi, Eilis I Murphy, Meghan L Meyer, Dylan D Wagner, Paul E Holtzheimer, and Andrew T Campbell. Mental health and behavior during the early phases of the COVID-19 pandemic: A longitudinal mobile smartphone and ecological momentary assessment study in college students (preprint). *Journal of Medical Internet Research*, May 2020.

[12] Dante L Mack, Alex W DaSilva, Courtney Rogers, Elin Hedlund, Eilis I Murphy, Vlado Vojdanovski, Jane Plomp, Weichen Wang, Subigya K Nepal, Paul E Holtzheimer, Dylan D Wagner, Nicholas C Jacobson, Meghan L Meyer, Andrew T Campbell, and Jeremy F Huckins. Mental health and behavior of college students during the covid-19 pandemic: Longitudinal mobile smartphone and ecological momentary assessment study, part ii. *J Med Internet Res*, 23(6):e28892, Jun 2021.

[13] Subigya Nepal, Weichen Wang, Vlado Vojdanovski, Jeremy F Huckins, Alex daSilva, Meghan Meyer, and Andrew Campbell. Covid student study: A year in the life of college students during the covid-19 pandemic through the lens of mobile phone sensing. In *CHI Conference on Human Factors in Computing Systems*, pages 1–19, 2022.

[14] Borja Sañudo, Curtis Fennell, and Antonio J. Sánchez-Oliver. Objectively-assessed physical activity, sedentary behavior, smartphone use, and sleep patterns pre- and during-COVID-19 quarantine in young adults from spain. *Sustainability*, 12(15):5890, July 2020.

[15] Shaoxiong Sun, Amos A Folarin, Yatharth Ranjan, Zulqarnain Rashid, Pauline Conde, Callum Stewart, Nicholas Cummins, Faith Matcham, Gloria Dalla Costa, Sara Simblett, Letizia Leocani, Femke Lamers, Per Soelberg Sørensen, Mathias Buron, Ana Zabalza, Ana Isabel Guerrero Pérez,

Brenda WJH Penninx, Sara Siddi, Josep Maria Haro, Inez Myin-Germeys, Aki Rintala, Til Wykes, Vaibhav A Narayan, Giancarlo Comi, Matthew Hotopf, and Richard JB Dobson and. Using smartphones and wearable devices to monitor behavioral changes during COVID-19. *Journal of Medical Internet Research*, 22(9):e19992, September 2020.

[16] Mark R. Leary and Roy F. Baumeister. The nature and function of self-esteem: Sociometer theory. In *Advances in Experimental Social Psychology Volume 32*, pages 1–62. Elsevier, 2000.

[17] Sohrab Saeb, Mi Zhang, Christopher J Karr, Stephen M Schueller, Marya E Corden, Konrad P Kording, and David C Mohr. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: An exploratory study. *Journal of Medical Internet Research*, 17(7):e175, July 2015.

[18] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231, 1996.

[19] Luca Canzian and Mirco Musolesi. Trajectories of depression: Unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '15, page 1293–1304, New York, NY, USA, 2015. Association for Computing Machinery.

[20] Zhenyu Chen, Mu Lin, Fanglin Chen, Nicholas D. Lane, Giuseppe Cardone, Rui Wang, Tianxing Li, Yiqiang Chen, Tanzeem Choudhury, and Andrew T. Campbell. Unobtrusive sleep monitoring using smartphones. In *2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops*, pages 145–152, 2013.

[21] Rui Wang, Fanglin Chen, Zhenyu Chen, Tianxing Li, Gabriella Harari, Stefanie Tignor, Xia Zhou, Dror Ben-Zeev, and Andrew T Campbell. Studentlife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*, pages 3–14, 2014.

[22] Yu Huang, Haoyi Xiong, Kevin Leach, Yuyan Zhang, Philip Chow, Karl Fua, Bethany A. Teachman, and Laura E. Barnes. Assessing social anxiety using gps trajectories and point-of-interest data. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, September 2016.

[23] Mehdi Boukhechba, Philip Chow, Karl Fua, Bethany A Teachman, and Laura E Barnes. Predicting social anxiety from global positioning system traces of college students: Feasibility study. *JMIR Mental Health*, 5(3):e10101, July 2018.

[24] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*, 2014.

[25] Alexandra Kuznetsova, Per B Brockhoff, and Rune HB Christensen. lmertest package: tests in linear mixed effects models. *Journal of statistical software*, 82:1–26, 2017.

[26] Julien Herzen, Francesco Lässig, Samuele Giuliano Piazzetta, Thomas Neuer, Léo Tafti, Guillaume Raille, Tomas Van Pottelbergh, Marek Pasieka, Andrzej Skrodzki, Nicolas Huguenin, Maxime Dumonal, Jan Kościsz, Dennis Bader, Frédérick Gusset, Mounir Benheddi, Camila Williamson, Michal Kosinski, Matej Petrik, and Gaël Grosch. Darts: User-friendly modern machine learning for time series, 2021.

[27] Romain Tavenard, Johann Faouzi, Gilles Vandewiele, Felix Divo, Guillaume Androz, Chester Holtz, Marie Payne, Roman Yurchak, Marc Rußwurm, Kushal Kolar, and Eli Woods. Tslearn, a machine learning toolkit for time series data. *Journal of Machine Learning Research*, 21(118):1–6, 2020.

[28] Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1):289–300, 1995.

[29] Mona Moieni, Michael R Irwin, Kate E Byrne Haltom, Ivana Jevtic, Meghan L Meyer, Elizabeth C Breen, Steven W Cole, and Naomi I Eisenberger. Exploring the role of gratitude and support-giving on inflammatory outcomes. *Emotion*, 19(6):939, 2019.

[30] Tristen K Inagaki and Naomi I Eisenberger. Giving support to others reduces sympathetic nervous system-related responses to stress. *Psychophysiology*, 53(4):427–435, 2016.