

User Research Paper

Research Methods

For our research, several methods were implemented to evaluate and understand a person's state of mind during an average college day. First and foremost, a preliminary survey was conducted to get a grasp on the demographics, gender, year in college, major, location of stay (on or off-campus), and initial expectations of the application after a brief description was given to the users. Following the initial survey was the subjective analysis of the state of minds of the users. The subjective analysis was conducted in person, and online through emails. One of the main components of the application is being able to discern emotions from the word choice of the users. At any given point in time, the application should be able to accurately give suggestions to the user depending on how he/she is feeling.

In an effort to make the results as accurate as possible, we have performed a manual analysis along with a *sentiment analysis* tool called VADER [6], which involved using a Python script to ensure that the manual scoring and the scoring done by the algorithm are within the same range. This tool is an algorithm that implements machine learning by using a predefined collection of words with an assigned score for each. Depending on the words used in the responses, the goal of the algorithm would be to predict the nature of the responses as either being *positive* or *negative* in nature. If the responses are deemed to be negative, the app would require further enhancements. To analyze the subjective responses given by the respondents, we created a WordCloud [1] to better categorize the results. WordCloud was our method of organizing words and terms that users frequently used to describe how their day was going.

During user research, we found that having the users rate their responses using a Likert scale proved very helpful because *sentiment analysis* relied heavily on the algorithm's ability to predict feedback for users depending on their responses. Having the users rate their responses simplified the process of predicting the feedback using the software tool, VADER [6]. Details about the questions asked during these evaluations can be found in the *Appendix*.

Findings

Of the 36 people surveyed, 20 people said that app would help them relieve stress to some extent. Additionally, of those 20 surveys, 17 people said their days were rated 3 or above on scale of 1 to 5, where 1 was awful and 5 was awesome. Coming to our hypothesis, 27 out of 36 people corroborated our initial hypothesis that social interaction alleviates stress as shown in Figure 1 below. The general trend as seen in the graph below is that people that were having a good day were usually the ones with an equally good social life. Of all the 36 respondents that took the survey, 55.6%, as shown in Figure 2 said they would find a mental health tracking app helpful. Granted that it won't substitute for a professional therapeutic session, the general consensus seemed to agree that having a way to track one's state of mind over a period of time would better help analyze their mental health. Also, based on the survey, we concluded that gender and ethnicity had no bearing on a participant's state of mind. As a result, we opted to exclude them as factors to consider in our user research.

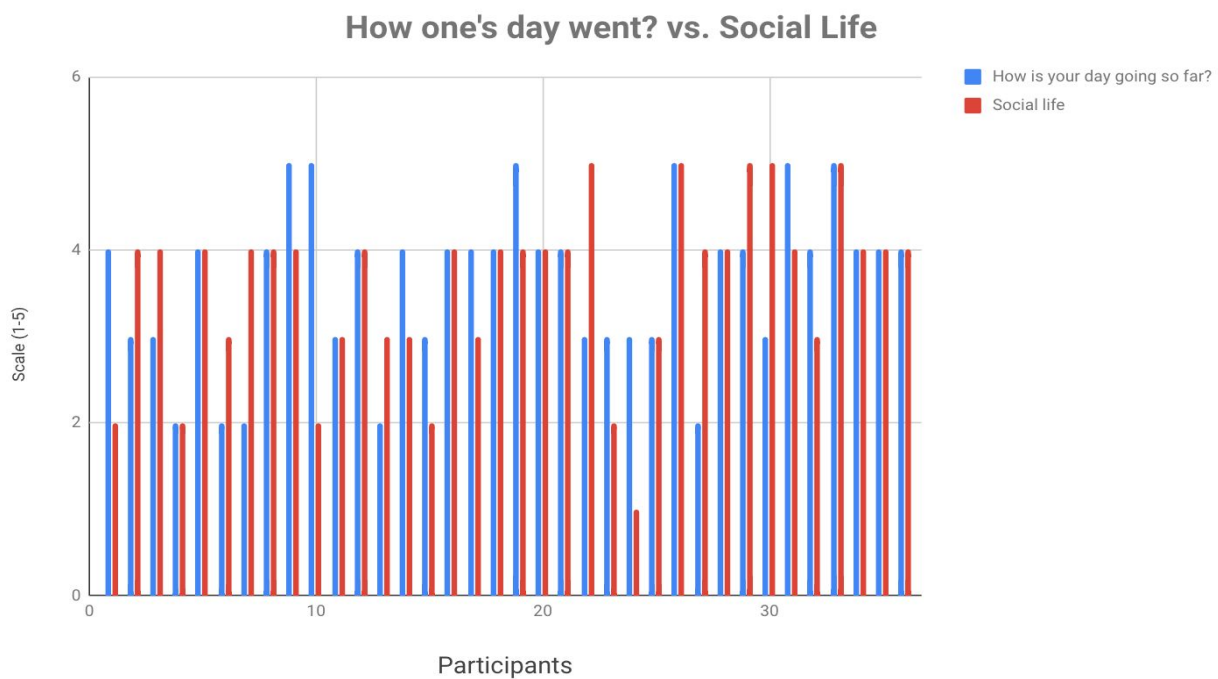


Figure 1: Relationship between how one's day is going and overall social life.

Count of Do you think a mental health tracking app would help relieve stress to some extent?

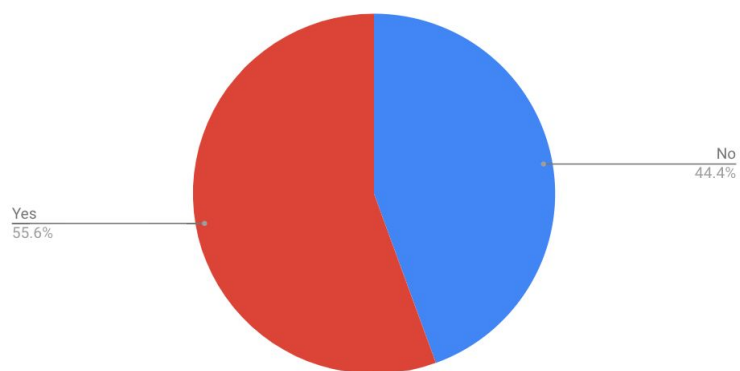
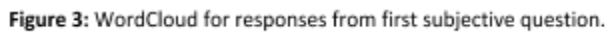


Figure 2: Responses about whether an app like ours would be useful.

To analyze our data, we used a variety of different methods. Since our user research produced results in the form of both qualitative and quantitative data, we used a combination of manual categorization/scoring and software tools to aid us in our analysis. The methods we used are detailed in the remainder of this section.

The first subjective question asked in the survey was “In a short response, explain a moment during the day when you felt stressed.” Figure 3 shows the output of WordCloud from the responses to this question. We can see that many of the largest words shown in the WordCloud seem to be related to school. “Class”, “lab”, “exam”, “physics”, “project”, etc. are all words that are particularly relevant among college students. The fact that these words showed up so frequently in responses when we asked participants what caused them stress that day strongly supports our hypothesis that college students struggle with stress which is a direct result of their status as a student. Many students also indicated that they felt crunched for time frequently, that they had not been getting enough sleep, and that WordCloud, one could not conclude that these responses where it became useful to manually review some of the was still at the root of these stressful situations. Only experiencing any stress whatsoever. Of the remaining school induced. This further strengthens our motivational health.



[illegible]

Figure 4: WordCloud for responses from second subjective question.

Similar to our methods of analysis for the surveys, we used a software tool called VADER [6] to aid us. VADER is a sentiment analysis tool that takes a string (sentence) as input, performs an algorithm on the string which calculates the overall emotion within (positive, negative, or neutral), and produces a score between -1 and 1 (-1=negative, 0=neutral, 1=positive). By comparing these scores with our manual analysis, we could gauge overall confidence in what we felt the participant meant. VADER's lexicon model is pre-trained with social media posts as well as online product reviews. Though VADER is relatively young compared to other similar tools, benchmarks have shown that its accuracy and reliability match or exceed even the most well established tools [7]. Additionally, Eric used this tool while working with a group of data scientists over the summer and can personally vouch for the effectiveness and accuracy of its output.

After manually analyzing responses from Interview 1, we had the impression that the overall emotion about our app was positive, and emotion about mental health was slightly negative. The VADER results show that both categories were slightly negative, though these numbers are both quite small and could almost be considered as having neutral emotion since they are so close to zero. Regardless, the VADER results did not match our manual assessment as closely as we had hoped. However, our impressions of Interview 2 were very positive overall, and VADER seemed to agree with us this time which suggests that we interpreted the responses correctly. If we were to conduct more interviews, we are confident that our consistency would increase and our analysis would be very accurate.

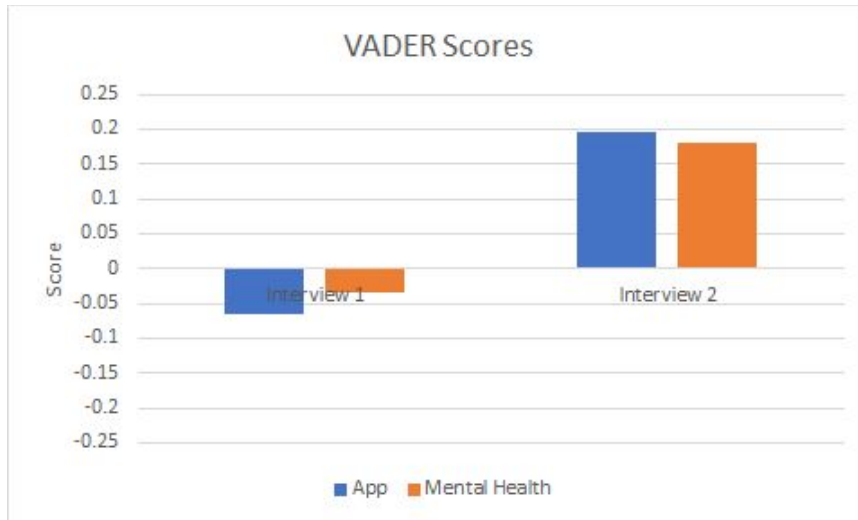


Figure 5: Visual representation of VADER results.

```
(project-env) root@LAPTOP-NLCC5G3D:/mnt/c/Users/Eric/repos/fall2018/4140/team_project# python3 analysis.py
Interview 1
app_avg = -0.0651375
mental_health_avg = -0.03387500000000001

Interview 2
app_avg = 0.19657499999999997
mental_health_avg = 0.1815875
```

Figure 6: VADER raw output.

Based on our detailed analysis, we believe that our original goal is still strong and that we can continue our research with little or no deviations from the course we were on prior to our user research. Gathering data about our target users has been invaluable to our study, and will certainly lead to an end product of higher quality than one created without consideration of its users.

References

- [1] A. Mueller. 2018. *WordCloud for Python*. Retrieved October 3, 2018 from https://amueller.github.io/word_cloud/
- [2] N. Alharbi et al. 2017. *Molecular Visualization of Computational Biology Data: A Survey of Surveys*. Eurographics Conference on Visualization (EuroVis) 2017.
- [3] A. Smith et al. 2017. *Evaluating Visual Representations for Topic Understanding and Their Effects on Manually Generated Topic Labels*. Transactions of the Association for Computational Linguistics, vol. 5.
- [4] F. Baumann and D. Roller. 2017. *Overview of German Additive Manufacturing Companies*. MDPI Data 2017.
- [5] J. Blasco and E. Quaglia. 2018. *InfoSec Cinema: Using Films for Information Security Teaching*. Information Security Group, Royal Holloway, University of London.
- [6] C.J. Hutto and Eric Gilbert. 2014. *VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text*. Association for the Advancement of Artificial Intelligence.
- [7] F. Ribeiro et al. 2016. *SentiBench - a benchmark comparison of state-of-the-practice sentiment analysis methods*. EPJ Data Science Journal.

Appendix

Basic Demographic Questions (Questionnaire Part 1)

1. What is your name?
2. How old are you?
3. Gender:
4. Ethnicity:
5. What is your major?
6. Do you think a mental health tracking app would help relieve stress to some extent?

Subjective & Likert-Scale Questions (Questionnaire Part 2)

1. On a scale of 1 to 5 (1 being awful, and 5 being terrific), how is your day going so far?
2. In a short response, explain a moment during the day when you felt stressed.
3. What did you, or are you doing to overcome this stress?
4. How many hours a day do you work on school-related, or work-related things on average?
5. How many hours do you spend on doing activities that you find entertaining/relaxing on average?
6. How would you rate your social life on a scale of 1 to 5 (Assume 1 means you have very little to no social interaction, and 5 means you have a good balance of work and social life)?
7. How often do you hangout with your close group of friend(s) in a week on a scale of 1 to 5 (1 being none, and 5 being five, or more times a week)?

Interview

1. What is your gender, age, and race?
2. What is your major?
3. Are you on or off campus?
4. Do you have any experience participating in research studies?
5. How helpful do you think an app for monitoring mental health would be to college students?
6. Do you know people that might use this if it became available?
7. What's your opinion on Clemson's mental health resources? Do they do enough or too little?
8. How important do you think that mental health is overall?
9. Do you think an app could provide the necessary attention or help that is needed for people struggling with mental health?
10. Do you think an app should have limits on how invasive or personal it is with one of its users with regards to mental health? Are there any topics that are off limits? Which ones?
11. Should the app be able to make decisions without the user knowing (for example, alert a professional if they show signs of depression or more serious symptoms)?
12. Do you think an app like this that helps college students handle mental health could be widely used?
13. What's the difference to you between someone being happy or sad (mood) and their mental health being out of whack (them being depressed)?
14. Do you think students struggling with mental health would be brave enough to use this app if it was available to them?
15. Would you ever use an app if you started struggling with mental health (and were too afraid to seek professional help)?
16. How passionate are you about mental health and providing resources to people who need help? (scale of 1 to 5)
17. Do you think society doesn't place enough value on mental health?