Phone Addiction and Human Behavior

J. Bou Karam, R. El Jammal, P. Abi Chacra

The analysis of the relation between phone dependency and human personality is a crucial subject nowadays, for technology plays a major role in everything anyone does. Several studies on the link between these two topics have been conducted in which papers were published, research was done, and in this case, visualizations are developed. The variety in all kinds of work regarding this topic has shown how serious, wide, and important the internet has become, to a point where it is as essential as food is to some. For this reason, in this paper, and in particular, this project, data, collected from an application that monitors phone usage, is presented in several forms to show the diversity in visualization, the growing addiction, and the extent of influence phones establish on their users to the point where every part of their life becomes affected. The proposed visualizations unravel trends, patterns, and the reality of the topic dealt with. As a result, this project helps users understand the extremity of their dependency and its negative consequences.

# Introduction

Technology is a huge part of people’s life nowadays. It offers several benefits to its users, starting from the basic services all the way to connecting the world to become one global worldwide network. However, technology has become a huge negative consequence in everyone’s life. In fact, phone addiction is a reality that almost all users suffer from. By definition, phone addiction is produced from excessive use of the internet or mobile applications with the users’ inability to focus on other tasks. For this reason, this project focuses on the effect of addiction on human behavior in terms of personality, productivity, and lifestyle.

Moreover, according to Sehar Shoukat, psychological problems gradually develop along with the increasing dependency like anxiety, depression, and insomnia. These disorders affect both the young and old by limiting their focus, wasting their time, and missing opportunities. The visualization presented throughout this project deals with several hypotheses that are to be proven or refuted like the negative relationship between smartphone usage leading to its addiction and the user’s wellbeing when it comes to his/her professional, academic, and personal life.

Furthermore, it sheds the light on how dependent users have become on their phones by calculating the number of times they unlock their devices, the amount of time spent on each category and, in turn, subcategory. Since technology is now an essential in most people’s lives, this project targets all kinds of audience, poor and rich, young and old, man and woman.

Although several people claim that technology offers more benefits to its users than liabilities, we, as well as Konok, Gigler, Bereczky, and Mikosi, strongly believe that negative underlying forces develop along with the growing addiction, like psychological, physiological, physical, and emotional.

# Existing work

Several studies have been conducted in order to assess the relation between phone addiction and productivity. One of the most related and remarkable is “Beyond Self-Report: Tools to Compare Estimated and Real-World Smartphone Use” by Sally Andrews, David Ellis, Heather Shaw, and Lukasz Piwek. It focuses on the actual use of mobile phones and the users’ estimation of how much they spend time on their devices. Furthermore, it approximates the user’s phone usage with respect to time based on his/her age. The data is presented through bar and line charts in order to clearly show the trends and patterns, compared the data, and find relevance among categories. The derived content in this paper will be used throughout the visualization in order to either accept or reject the hypothesis mentioned above, which is the negative effects phone dependency produces on its users. This work can be resorted to since the dataset used in this project has similar attributes, which boosts its credibility.

Second, “Optimal Screen and Study Time for Achievement of High Academic Performance in Adolescents” deals with the influence screen time has when it comes to students’ academic performance. It has been deduced that lack of attention, a decrease in academic performance, speech deficiencies, demonstration of violent behavior, and substance use can be all results of smartphone addiction. The channels and marks used in this study consist of bar charts with labels and colors to clearly show the difference among categories and tables in order to display the data with their attributes. This paper will be used to determine the hypothesis status in relation to the survey results gathered in the project.

Last but not least, “Cell phone addiction and psychological and physiological health in adolescents” and “Humans' attachment to their mobile phones and its relationship with interpersonal attachment style” are significant examples of the psychological disorders that might develop with the growing addiction.

All the papers mentioned above are remarkable pieces of work that are helpful in understanding the topic at hand, the obtained results from the papers, the results from the visualizations along with their datasets, and the final hypothesis decision, which is either approved or refuted.

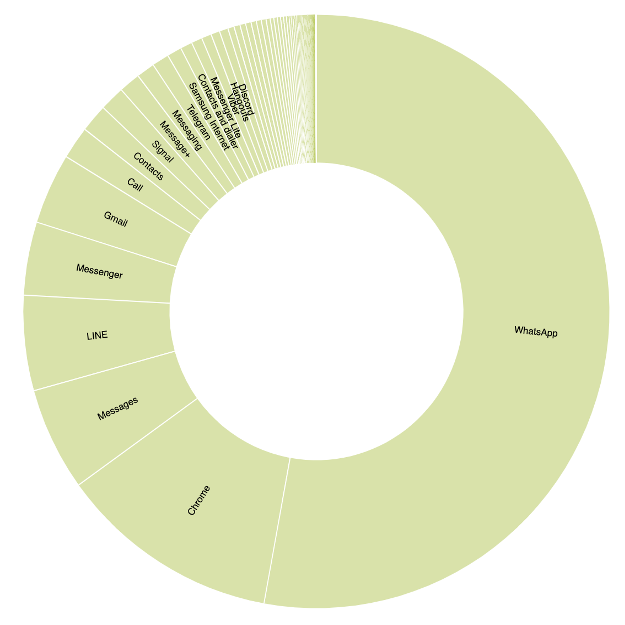
# Methodology

Chart, sunburst chart

Description automatically generated

Figure 1a Sunburst Chart with Application Types and Names.

The project focuses on different effects caused by phone addiction on users by measuring the time spent on applications, calculating application usage, computing the number of locks and unlocks recorded on Space, and analyzing other features, which will be further discussed in detail. Three datasets were used in developing the visualizations to reach our final hypothesis. The first dataset records the number of locks and unlocks done by all users at specific dates and times. The second one deals with the application types and names to show how often they are used through redundant data, and the last one is a survey done on more than four hundred people from different cultures, genders, and ages. For efficient and accurate use of data, the dataset files underwent preprocessing to extract the needed information in order not to overwhelm the visualization with unnecessary information. Before we go into detail, preprocessing was implemented in python using *pandas*, a library specific for data analysis and manipulation. After that, we exported the files into either csv or json format.



**Figure 1b** Sunburst Chart with Application Names

(after clicking on a type).

First, we wanted to get the most frequent used types of applications, and to do so, we created a dictionary in python that contains all application types as keys and their corresponding applications as values, which are also called their descendants. The applications were repeated among the users, so we got the count of each, i.e., how many times it is repeated. So, the result was a json file that contains the types as keys and the actual applications as values along with their count. This leads to proof on what type of apps the users spend their time on and what apps are most commonly used.

A sunburst containing the types of applications and their subcategories is shown. The inner part represents the category as a whole, communication for example, while the outer portions are its decedents, like WhatsApp. One channel to focus on is the size of every portion, which shows how often an application is used. Moreover, every category has its own color to differentiate between the different types. Some helpful tools to help better visualize the data is the tooltip which appears every time we hover over a piece, hence showing the name of the application or parent with the total number of times used. When clicking on a classification, the sunburst goes deeper into the attribute chosen and shows more applications that might not have been shown at an earlier stage. The sunburst is a great visualization in this case since it deals with a huge dataset containing ordinal data and shows how dependent users are on their phones by the size of every piece in the graph. It can be deduced that people are mostly addicted to applications belonging to the communication category, most specifically WhatsApp.

One of the data that was preprocessed was from the “*CLEANED Unlocks (cleaned R13-08-20).csv*” file where only the *User Key, Operations (Locks and Unlocks), and Date* columns were extracted and put in a json file where the main key was the type of operation. We continued by computing the total number of usages based on operation type, whether lock or unlock, for every date by looping over all values under the same key. Then, we found the total number of unique users performing the operation at a certain date as well for the goal of showing the user how significant the data is. Null values were removed since they do not help in counting the number of locks and unlocks in this case. The resulting dictionary was then saved as a json file (*locksUnlocksProcessing.json*).

Chart, scatter chart

Description automatically generated

Figure 2 Scatterplot showing the relation between the number of locks and unlocks.

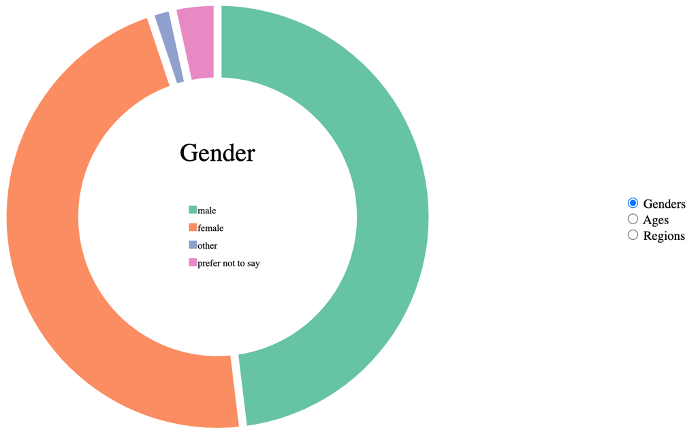


Figure 3a Donut Chart to show the genders participating in the survey.

A scatterplot is represented showing the top twenty locks and unlocks by date based on the total number of operations made. It begins by showing both the locks and unlocks simultaneously. In the figure below, the operations are almost overlapping since it happened that they are approximately equal. The blue colors represent the unlock, while the red are the locks. Clicking on a legend makes its appropriate data disappear; hence making it easier for the user to easily receive the data needed. Clicking on the same legend again makes the dots reappear When hovering over a circle, a tooltip appears that shows the type of operation, date, number of times performed the operation, and the total number of users. The reason for showing only a portion of the data is not to overwhelm the users with the huge dataset; therefore, only showing the significant data. This visualization was chosen to identify patterns in the data and the relationships between the two operations, which is proportional and almost the same.

After that, we go into the more important part of our project, which is the survey which either proves or refutes our hypothesis.

Before developing the visualization for showing the kind of participants in the survey, some preprocessing was made on the survey csv where only the *Gender, Age, and Eastern\_Western* columns were extracted. Then, we computed the total number of people answering the survey, which was 468, by finding the length of the data - len(data) – and then computing the total number of every category alone. In other words, within the genders, we have four categories: male, female, prefer not to say, and other where the total number of males and females 225 and 219 respectively. Then, we computed the probability of each grouping and stored them in their appropriate variables. Null values were considered as “Not Specified” since they are crucial in getting the exact statistics and number of people. As a result, we obtained a json file (*genderProcessed.json*) containing the category, sub-category, number of records and probability belonging to the latter, and the total number of people participating.

To make the user familiar with the figures, a donut chart is first visualized to introduce the participants, ranging from their gender, region (eastern/western), and age to show the distribution of data belonging to different users. The donut chart was also made interactive where each piece shows the percentage and total number of people belonging to a certain category. We find three radio buttons on the right of the chart where we can alternate between the three different labels. The legends in the middle of the donut chart, along with their colors, change according to the category they are representing. This donut chart is a user-friendly one that shows the necessary data to better understand the coming visualizations. It can be concluded that mostly males, people between the age of 20 and 29, and users of western culture are completing the survey.

Fourth was the extraction of the *emotion* *times check per hour* and *emotion time use phone per day* in order to show the distribution of emotions*.* These attributes were merged into a single json file for the purpose of presenting them in a single alternating pie chart, which will be discussed later. The probabilities of every emotion were also counted and saved in a json file where the null values were removed due to the lack of contribution in achieving this purpose.

A fourth figure produced is the pie chart that is made up of two json files showing the same structure of data but different attributes. In other words, the emotions, happy, somewhat happy, unhappy, and somewhat unhappy, are represented based on the derived percentages belonging to each column. To begin, the first pie chart shown is the “Emotion Times Check Phone Per Hour” where its corresponding columns were computed. Clicking on the second button, “Emotion Time Use Phone Per Day” updates the pie chart with transition time so it looks like it’s moving where the percentage of occurrence of every emotion is displayed when hovering over a piece. This pie chart is a clear way to show the difference among attributes in an efficient and organized manner where all necessary information is shown accordingly. As a result, it can be concluded that phone dependency leads to more unhappiness among users since both negative emotions dominate the positive ones.

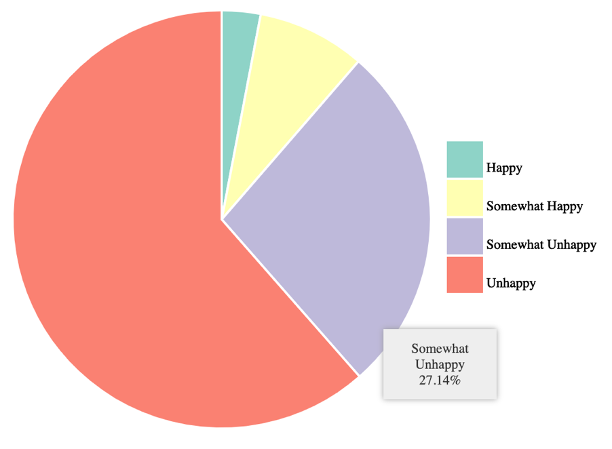


Figure 4 Pie Chart showing the difference among user emotions.

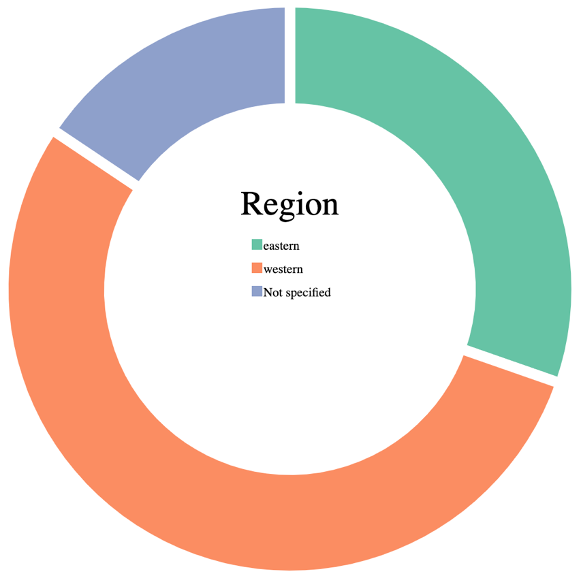


Figure 3c Donut Chart that shows the users’ regions.

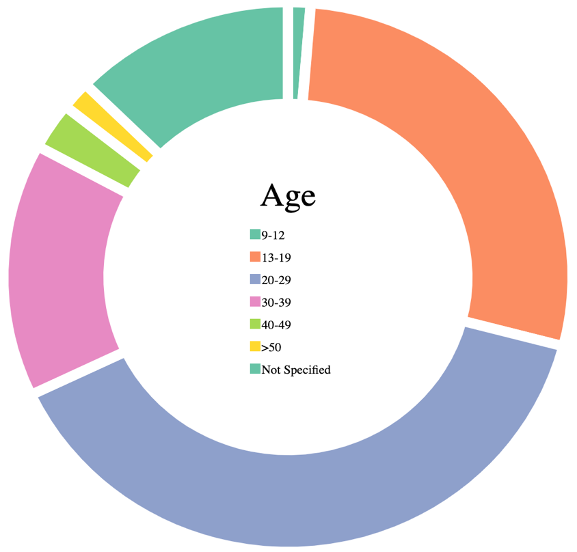


Figure 3b Donut Chart that shows the distribution of the ages.

The fifth preprocessing was made on two features from the “*CLEANED SSCoDA Survey Data in CSV Converted from SPSS*” dataset, which are the user’s culture and emotion towards the reported amount of time they spend on their mobile phones, be it on social media, gaming, or any other application. The null values were removed due to their irrelevance in this case. We then got the total number of users that are happy, unhappy, somewhat happy, and somewhat unhappy respectively, where we finally obtained a dictionary, like the proceeding example:

{

“emotion”: “Somewhat happy”,

“east\_count”: “13”,

“west\_count”: “19”

}

The fifth visualization is the grouped bar chart which shows a more detailed representation of the emotions according to the region the participant is in. in the chart below, each region is represented using a color where the corresponding values can be automatically derived or viewed on hover over a single bar. The reason for using this visualization is that grouped bar charts are most optimal for visualizing side-by-side categories, and since the features are categorical, we found this to be the best way to see the data clearly and efficiently. We wanted to prove whether the culture the user belongs to has any effect on his/her emotion regarding phone usage. According to the presented results, we can conclude that generally, eastern users are happier than western.

In addition to that, other features were also extracted from the “CLEANED SSCoDA Survey Data in CSV Converted from SPSS” file. The columns are education of the user, employment type (full-time, part-time, unemployed…), age, number of hours of sleep, minutes spent on phone. For the preprocessing, we took the features needed and removed the null values. Then we saved it in a csv file to visualize.

One of the most complicated implemented visualizations was the pair plot where several features were extracted and represented, each according to its own specification. For the visualization, we wanted to get the correlation among the age, number of hours of sleep, minutes spent on phone, education of the user, and employment type because we initially thought that they might affect one another, so for example, if you sleep less, maybe it means that you spend a lot of time on your phone, or maybe that younger ages spend more time than the older generation, and so on. For this reason, the hypothesis here is the correlation and how one feature may influence others. Therefore, a pair plot is the best way to do that in order to see the actual relationship between the features. We notice that the features aren’t really correlated, and when we see dots placed horizontally or vertically, that is because the features are categorical. We can notice in the plot between the Age and MinsSmartphone features, that the younger ages tend to use their phones for longer period of time which makes sense.

One of the most interesting visualizations developed is derived from the main survey csv. After cleaning unnecessary headers for the purpose of showing the difference between the self-reported time spent and actual time spent, three columns where taken; the first one is the user index, the second one was the Q26\_Self\_Report\_Mins, and the last one was the Q26\_Actual\_Usage\_Mins. Finally, we then cleaned the data by dropping rows that had null values since we can’t correlate a data with an undefined data in this case.

The difference chart below represents the user indices on the x-axis and the actual and assumed time spent on the y-axis. When the self-reported minutes are higher than the actual usage mins, the difference between the two is filled in green and when the actual usage minutes are higher, the difference is filled in red. We can deduce that the majority of people think that they spend more time on their phones than what is actually recorded, but the actual data refuted this thought and showed that more than half of the users spend less time than what is assumed. In other words, about 40% of users spend more time than expected and the last bit of users don’t have a noticeable difference between self-reported and actual time.

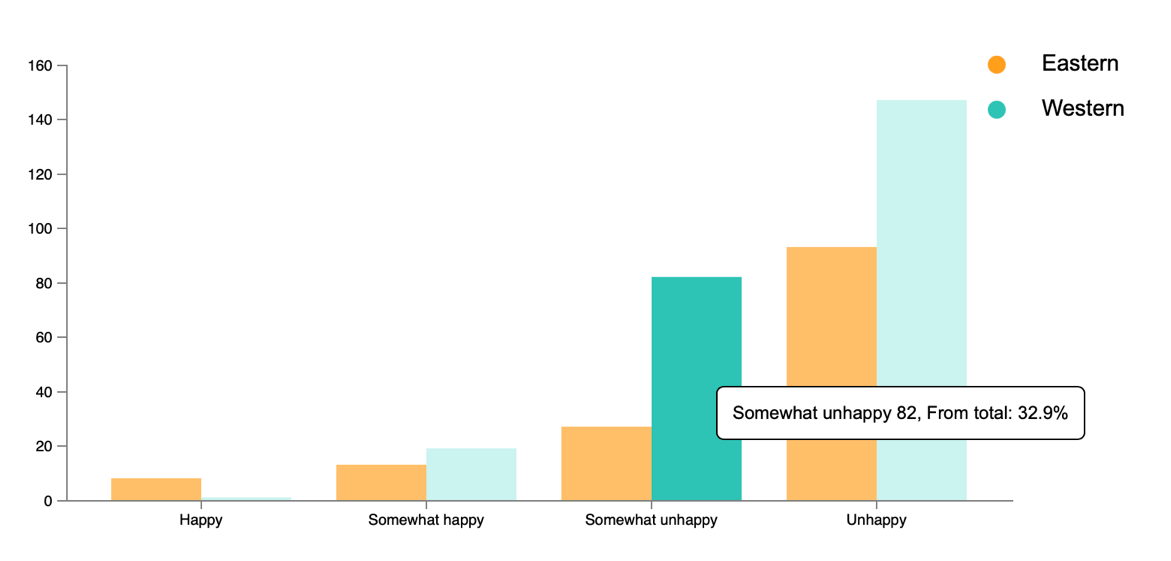
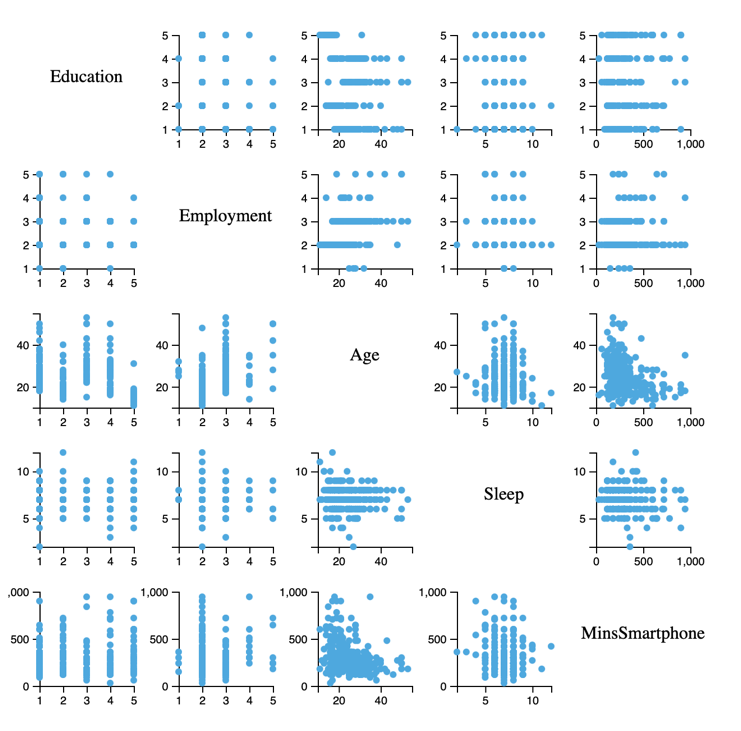


Figure 5 Grouped Bar Graph showing the distribution of emotions among cultures.



**Figure 6** Pair Plot that shows correlation among attributes.

Last but not least, the scatterplot below uses a csv created from the main csv containing three main attributes; the emotions, total time spent on the phone, and time spent on every specific category. Five scatterplots were created to represent every category (gaming, messaging, etc.…) One important note to keep in mind is that null values were removed in this case due to the chart’s showing only emotions, so null values are irrelevant. All five plots have the emotion and total time spent in common where the latter is represented on the y-axis while the time spent on a certain application type is measured on the x-axis. Those visualizations help us see on which time range there is more concentration in usage and in terms of emotions where a filter was added as well by hovering over the legends which filters out all other legends. Buttons on top of the plot can help filtering the plot by emotions in an organized manner as well. For example, we can deduce that almost everyone is unhappy with high phone usage and dependency is high since the minimum usage is around an hour a day based on the figures below. Therefore, we conclude that higher phone usage leads to lower happiness regardless the application type.

Chart, line chart

Description automatically generated

Figure 7 Difference Chart showing the variation in actual time spent vs. the assumed time

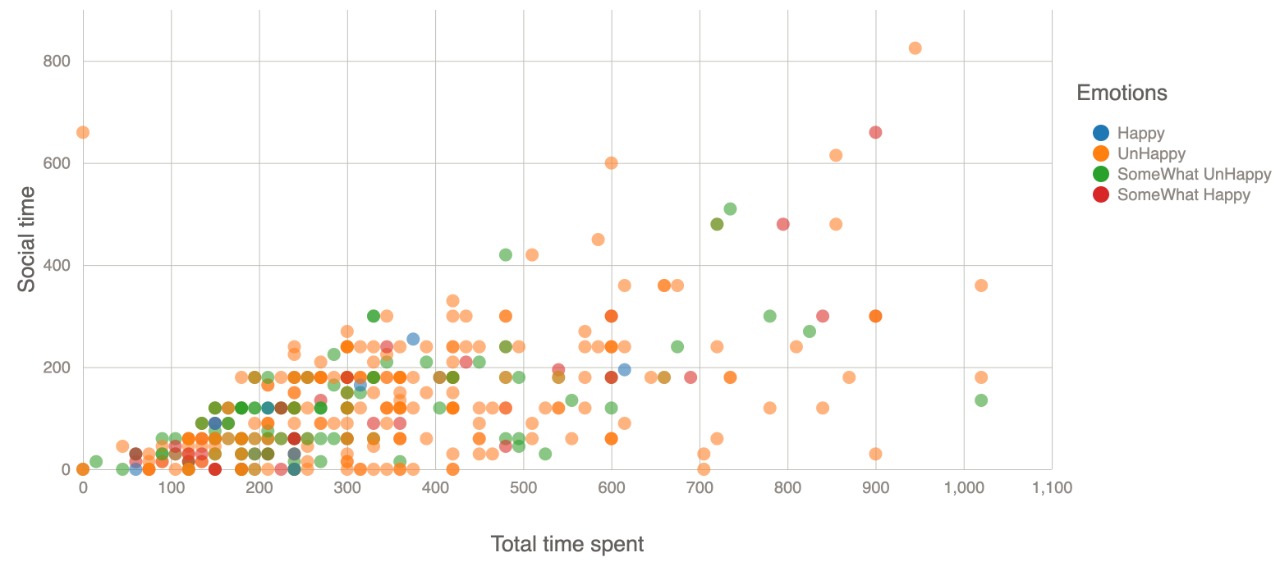


Figure 8a Scatterplot showing the emotions corresponding to social applications.

# Discussion

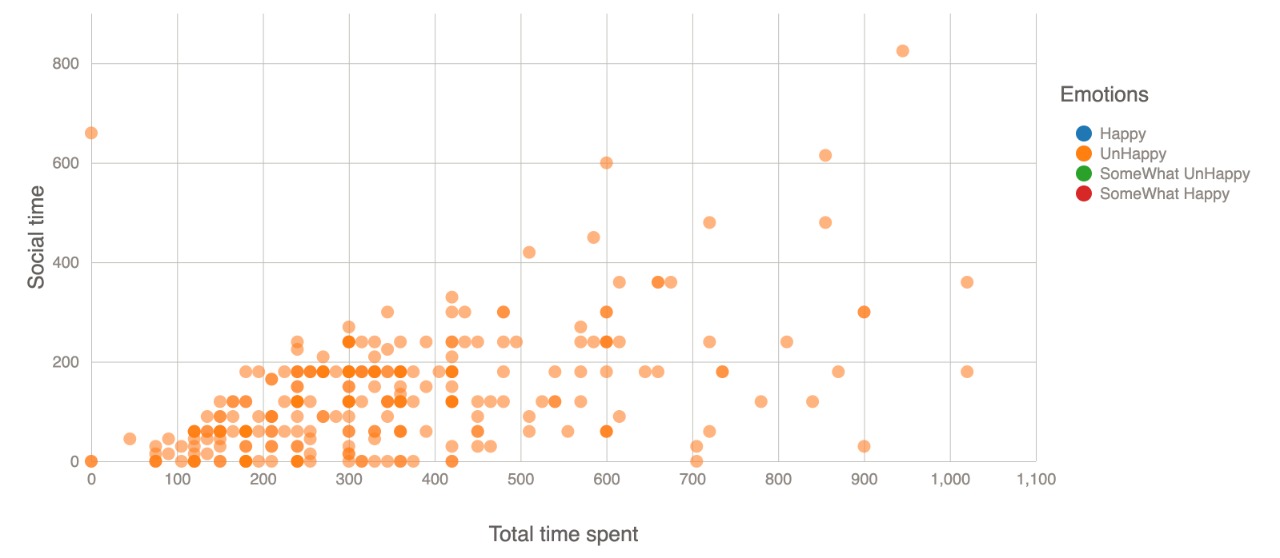


Figure 8b Scatterplot showing unhappiness by hovering over the corresponding legend.

Technology is considered one of the major life necessities since it makes the world one global community. In other words, everything is reachable and possible to acquire with just a click of a button. This leads to more and more people spending most of their time on their phones instead of being productive either by studying, working, or completing some chores. The project is an important topic nowadays because the internet comes first to many people. For this reason, this project studies the effect of phone usage on the user’s emotions and personality in order to either prove the significance of this subject in our daily lives. The visualizations mentioned above provide a simple yet effective technique that guides users throughout the whole procedure step by step to finally reach the last conclusion; how serious technology influences our mood. This topic is a relevant one nowadays since phones may lead to some psychological issues, like depression (which can be attained through constant unhappiness), anxiety, and other terms. The dataset, although huge, covers only a small part of phone and internet users, which makes it somewhat not credible since it is evaluated on a small number of people. The more people who use the application Space that records their activity, the more our hypothesis can be proven to be accurate although its being accepted with the results at hand.

# Conclusion

In conclusion, all visualizations represented in this project lead to our final hypothesis: how phone dependency affects one’s mood and behavior. Some figures, like the scatterplot that shows the locks and unlocks, prove how addicted people have now become to social media, where a user may just check his phone for at least a thousand times a day. This proves the growing dependency produced by smart phones. Moreover, the sunburst shows where this dependency mostly resides, which is in the communication category since it makes up the biggest portion. Most importantly, the visualizations belonging to the survey prove that the more people use their phones, the more unhappy they are even if certain categories like the employment and sleep are not very related. The scatterplots showing a certain category’s time usage with respect to the total time show that almost everyone is unhappy no matter which application they use, and that there is a huge dependency where almost everyone spends their time on a certain category for more than two hours per day. As a result, our hypothesis was proven to be correct since increasing phone dependency leads to an increase in unhappiness.

**Authors’ Contributions**

* Joy Bou Karam
  + Sunburst: showing frequency in application usage.
  + Scatterplot: representing the pattern of locks and unlocks to show the high dependency.
  + Donut Chart: introducing the audience to the type of users contributing to the survey.
  + Pie Charts: showing distribution of emotions in two categories (Phone Usage and Phone Check)
* Rim El Jammal
  + Grouped Bar Chart: showing side-by-side categories (Eastern and Western) with emotions.
  + Pair Plot: showing correlation among features related to the survey.
* Pierre Abi Chacra
  + Scatterplots(5): showing the influence of time usage in a certain category on emotions.
  + Difference chart: showing the fluctuation and difference between the actual time spent reported by the application and the self-reported time spent.

However, each member contributed in the development of a visualization implemented by another team member for better team work and communication.

# References

1. Shoukat, S. (2019, February 04). Cell phone addiction and psychological and physiological health in adolescents. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6449671/>

2. Konok, V., Gigler, D., Bereczky, B. M., & Miklósi, Á. (2016, March 28). Humans' attachment to their mobile phones and its relationship with interpersonal attachment style. Retrieved from <https://www.sciencedirect.com/science/article/abs/pii/S0747563216302333>

3. Mehmet, K., Mehmet, C., Ahmet, A., Aslınur, K., Betül, K., Büşra, B., . . . Mustafa, T. (2014, June 11). (PDF) Optimal Screen and Study

Time for Achievement of High Academic Performance in Adolescents. Retrieved October 29, 2020, from <https://www.researchgate.net/publication/273662507_Optimal_Screen_and_Study_Time_for_Achievement_of_High_Academic_Performance_in_Adolescents>

4. Andrews, S., Ellis, D. A., Shaw, H., & Piwek, L. (2015). Beyond self-report: tools to compare estimated and real-world smartphone use. PloS one, 10(10), e0139004, from <https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0139004&type=printable>