Vhuhwavho Todani
INST414
Sprint 1
QSSR track

Executive Summary

How does technology/social media usage affect individuals' mental health? (QSSR track)

Access to technology and the online world has become increasingly widespread and is already a norm in some regions of the world. It's become fairly normal to gain access to things such as the internet and social media at very young ages, although most would acknowledge the obvious positives aspects such as being able to access information easier or connect with more people all over the world, it could be useful to identify potential negatives. A major question is how much does the screentime/social media/tech use affect mental health, especially in youth and young adults? Many different types of people could benefit from research in this area, from individuals themselves, parents ,teens etc. It could inspire people to reevaluate or think about how often they make use of their devices and whether or not it has good or bad effects on them. The QSSR track is ideal for this topic as mental health is a notable social science issue which can impact a lot of different areas and aspects of a community.

Background & Motivation

The reason this issue is interesting is because there has been a global increase in the access to technology and the online world in the past two decades. An interesting aspect of this is that younger people born into the current world essentially grow up with high internet access, social media and entertainment etc. it's increasingly common for children to have access to a smartphone by their early teens. Since this technology can come with a lot of positives and negatives, it's important to have some idea on how it influences us and how we interact with the world or life in general. I already have some knowledge on this topic as I have learned in the past how there have been studied associations between social media usage and negative mental health as well as read ups that focus on how remote communication has influenced things such as how often people decide to go out. I have also occasionally come across articles that stated that many

people spend hours on social media platforms such as tiktok daily. Studying this topic with data can be valuable as the information obtained could be used by individuals to reflect or potentially be used when building policies or programs that may center around the impacts of social media.

Data Source & Availability

https://www.kaggle.com/datasets/souvikahmed071/social-media-and-mental-health

This dataset with 482 records focuses on social media usage and mental health. The data was collected through making use of surveys. This dataset contains information such as the age of an individual, whether or not they use social media, time spent on social media, how often individuals report feeling down, whether or not they face issues sleeping along with other information which might be useful for finding trends in. I've accessed it and the necessary data is already in it.

https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/1WWCA5

This is the secondary dataset with few columns explores the amount of youth who needed mental health treatment in relation to the hours of screentime reported. It is collected from the Harvard dataverse and based on data from the National Survey of Children's Health that was released in 2020. This data is relevant as it allows us to spot potential trends in technology usage by screen time and chances of needing mental health treatment from ages 0-17. Any correlations found here could be valuable for research. It has around 41000 records. I've accessed this dataset already and the necessary information was fully available.

Approach overview

Using these datasets, I want to see if there are any correlations that would indicate a link between social media usage/screentime and an individual's reported mental health level. A notable positive/negative correlation could indicate that the amount of time a person spends engaged in social media or general screentime on technology has a notable role in their mental health. The primary dataset featured a very small minority who claimed they didn't use social media, I would also be interested in if those individuals had a notable difference in response in areas such as distraction or mood. The dataset also features metrics such as age and gender. It

would be valuable to see if groups such as men and women had any differences in reported metrics (e.g. were women impacted more by hours spent on social media than men?). Relationship status is also featured which could be a notable aspect in this research (e.g. do people who are in a relationship/married have the same trend in reports as those who listed themselves as single?)
During this project I would make use of visuals/statistics to properly display any trends that were found. For example a standard bar chart could be convenient when comparing male/female reports in the dataset. Statistical analysis methods will also be part of the research in addition to visualizations especially for identifying differences in responses from different groups. These will be carried out through the use of python, r or pandas

My expectation out of these sets is that there will likely be a correlation between increased social media/screentime and mental health whereas as social media/screentime increases, individuals' mental health may also show increased negative trends. This would fall inline with what I've previously been exposed to overtime. with the topic of technology usage/screentime/social media etc. A significant correlation in both the primary and secondary dataset would be enough to convince me that there is a notable impact that social media use/screentime has on individuals mental health however if the trends are weak or minor it may contradict my expectations.

I think my current datasets are functional and clean enough for the purpose of the paper and both of them are fully accessible so it's unlikely those will be issues moving forward. It could be an issue if the original question is unanswerable, for that case since I believe my topic has room for flexibility, I could consider finding data that may be adjacent to technology usage if available (for example internet browsing/gaming?)

CCDS installation + Github

https://github.com/Vtodani/INST414prjct