A logo with blue text

Description automatically generated

MSc Data Science and Engineering

Research Methods

AC52012: Quantitative Assessment

Submitted By

Reema Dutta: 2628084

Submitted to:

Professor Ramon Ruiz-Dolz

On

March 24th, 2024

Introduction: -

The New York Times recently published an insightful article titled “Epidemic of Anxiety, Depression, and Other Mental Health Issues: Whose Fault Is That?” This article delved into a comprehensive survey analyzing the Happiness Quotient across ten countries over the past 50 years or more. One surprising revelation was that people living in Chile were overall considered to have had a notably happy life. This finding contrasted with the conventional wisdom that Nordic countries are typically viewed as the happiest in the world, as suggested by the UN report. A particularly intriguing aspect highlighted by the survey was the observation that citizens of Chile exhibit a smaller Facebook user base compared to other countries, yet they rank high in other metrics. Despite Chile's GDP not being as high as that of other developed countries, factors such as its unique history, government policies, and cultural attributes were identified as potential influences on its current state. However, determining a causal relationship or correlation between the Happiness Quotient and social media consumption remains challenging, given the presence of other variables such as which specific social media platforms are popular in each country. Further research is warranted in this complex ecosystem to better understand the relationship between mental health issues and daily social media usage

Question: -

Inspired by the article, the topic of quantitative analysis is “What is the relationship between time spent on social media and mental health issues?”

Hypothesis: -

H0 : - There was an evidence of strong negative co-relationship between social media and mental health issues.

H1: - Some positive relationship between social media usage for a limited time and mental health of its users. This case study involves social media influencers or creative content creators who have made a living off these platforms.

Independent and Dependent variables: -

Independent variable – Hours spent on social media by one User for one day.

Dependent variable -- The mental health of a user, measured using a metric system of 1-10, where 10 indicates feeling happy and 1 indicates feeling least happy. This measurement is taken at the end of the day and is dependent on the time spent by the user on social media.

Expected Extraneous variable: -

In addition to the independent and dependent variables, several extraneous variables may influence the outcome of our study. These include demographic factors such as age, geographical location, socioeconomic status, and pre-existing health conditions. Furthermore, the specific social media platform being used by the individual may also introduce variability into the analysis, as different platforms may have varying effects on mental health.

How to control the extraneous variable: -

To reduce the potential impact of extraneous variables on our study, we have taken several measures. First, based on the findings of an online experiment, age has been deemed irrelevant and is not considered a variable that affects social media usage and mental health. Additionally, all participants in our study are selected from the same geographical location to control for regional differences in social media usage patterns. Furthermore, participants are required to have similar physical health statuses and financial backgrounds, and none have pre-existing mental health conditions. Finally, to ensure consistency in our analysis, all participants are using social media platforms with similar structures, designs, and algorithms.

Three statistical methods: -

Descriptive Statistics: This method allows us to summarize and analyse the data collected from our study participants. We calculate the mean, standard deviation, and mode of the happiness quotient to assess central tendency, variability, and distribution of mental health scores.

Pearson’s correlation coefficient: It is the measure of the linear relationship between happiness quotient and the hours spent online. Its range is between 1 and -1. For example, if the correlation coefficient is -1, it shows a perfect negative relationship between social media and hours spent on social media.

T test: T-test can be used to compare the means of happiness quotient of the two groups with different social media platforms. For each group there is a null hypothesis and an alternate hypothesis. If we know the mean and standard deviation of the happiness hypothesis , then we can calculate the p value from t score using t table

Calculate the p value of each set and compare it with the critical value of t distribution. And assess whether or not to reject the null hypothesis. We can also calculate the confidence interval . Confidence level is twice the significance level. For alpha = 0.1 t score is 1.68. If the mean lies between the confidence interval , we again fail to reject null hypothesis.

The p-value is the probability of obtaining a t value as extreme or more extreme than the calculated value if the hypothesis were true. Assuming Final result is we fail to reject the null hypothesis. Therefore there is statistically significant difference in happiness quotient between who spend more hours on social media and those spend less time on social media.

Conclusion:-

By employing these statistical methods, we aim to gain a comprehensive understanding of the relationship between time spent on social media and mental health issues, thereby contributing valuable insights to the field of mental health research.