Journal of Studies in Social Sciences and Humanities
http://www.jssshonline.com/
Volume 10, No. 1, 2024, 13-31
ISSN: 2413-9270

Statistical Analysis of Survey on Addiction to Social Media and Its Implications on Physical and Mental Health

Priyanka Aggarwal, Ph.D.
Department of Statistics, Hindu College, University of Delhi, Delhi, India

Samridhi Mehta, Ph.D.*
Department of Mathematics, Hindu College, University of Delhi, Delhi, India
Email: sifathisaid@hotmail.com

Neha Midha, Ph.D.
Department of Statistics, Hindu College, University of Delhi, Delhi, India

***Corresponding author*

Abstract
Iannotti et al. (2009), Lenhart et al. (2015) and Keles et al. (2019) studied the impact of social media particularly on the adolescents. Catering to a larger section of the population in terms of age and occupation, this study through the survey aims to understand the contribution of the factors leading to addiction to social media and their interplay. The impact of addiction to social media on physical and mental health of the respondent is also studied. The survey was conducted through a Google Form and 1056 responses were garnered. The data collected on 19 parameters was analyzed through various statistical tools. Through the survey, it was found that an increased use of social media has a negative effect on physical health and mental health. The findings obtained are data dependent.

Key words: Chi-square test of independence, Cronbach's Alpha, Z-test for difference of proportions, Logistic Regression, Discriminant Analysis

Introduction

Social media refers to websites and internet-based applications that help people communicate and share content. Social media platforms facilitate the exchange of information for entertainment, social interactions and are used by everyone ranging from celebrities, entrepreneurs and individuals. These platforms have also helped people get a place to express themselves, interact with people worldwide having shared interests and for various other activities such as education, art, music, reading among others. Considering the way, it has seeped into our lives, it is important that one pauses to see for any possible negative effect of it on our physical health, mental health and relationships. Understanding any possible implications of social media can help us deal with the associated issues (emotional, physical and general) in a better manner. Some of the possible perils of social media are stress, depression, low self-esteem, disturbed sleep, neglecting work, cyberbullying, violation of one's privacy etc.

To gain a deeper insight on addiction to social media and its implications on physical and mental health of the individual, this paper focuses on analyzing the results obtained from a survey, which was conducted

through a Google Form. Some contributing factors leading to possible addiction to social media were also examined through statistical techniques.

The review of recent studies conducted on social media and its implications is given in Section 2 of the paper. Framework of the questionnaire and the tools used to conduct and analyze the survey are underlined in Section 3 of the paper. Section 4 outlines the various socio-demographic aspects of the respondents of the survey. The study of characteristics contributing to the addiction to social media is deliberated upon in Section 5. Various statistical techniques used for analysis of the data collected through the survey have been discussed in the Section 6. This is followed by a discussion on findings obtained and references.

Literature Review

Iannotti et al. (2009) focused on the interrelationships of adolescent physical activity, screen-based sedentary behavior, and social and psychological health. They observed that the use of internet, when done in excess raises the risk of health problems. Lenhart et al. (2015) studied the role of social media on how the teens meet and interact with friends. They had identified the 13–17 age group as particularly heavy users of social media users, with 87% having access to a computer, and 58% to a tablet device. A report by the Royal Society for Public Health, & Young Health Movement (2017) suggested impaired sleep as a mechanism. Kim (2017) attempted to understand the impact of social media on adolescents' well-being. Primack and Escobar-Viera (2017) found that the number of social media accounts correlated with the level of anxiety, due to overwhelming demand. Marino et al. (2018) did a meta-analysis of 23 studies showing a correlation of problematic Facebook use and psychological distress in adolescent and young adults. Keles et al. (2019) studied the influence of social media on depression, anxiety and psychological distress in adolescents.

Materials and Methods

The survey consisting of 19 questions of qualitative and quantitative nature, was conducted using an online Google Form with no imposition of restrictions on age, gender, educational background, job profile etc. The participation in the survey was voluntary and the data was collected using snowball sampling. Through the questionnaire, the responses were collected on the number of hours spent on social media, impact of use of social media on one's personal life, physical and mental health, frequency of using social media, among others. The data obtained through the survey was first cleaned in order to remove duplicate/missing observations. 1056 responses to the survey were received where the age of the respondent ranges from 11 to 70. Various types of graphs such as histogram, pie-chart, bi-directional horizontal bar graph, cluster bar chart, and funnel chart were created to graphically represent the collected data. This aids in giving an overall quick and visual view to the information gathered and helps in comparative study and identifying the trends.

Several statistical tools were used to understand the interplay of factors and their effect. Methodologies in inferential statistics such as hypothesis testing and regression analysis have been used in the paper. Other statistical tests and techniques, namely, Chi-square test for independence of attributes applied by Pearson (1900), coefficient Alpha due to Cronbach (1951), Z-test for difference of proportions, Logistic Regression and Discriminant Analysis given by Verhulst (1838) and Fisher (1936) respectively, have also been applied to study the factors and draw relevant conclusions. The statistical analysis was carried out using the Statistical Package for Social Sciences (SPSS) Software version 23, Microsoft Excel and R Software.

Study of Socio-Demographic Characteristics

The respondents of the survey were from different professional backgrounds including students, academicians, businessmen, professionals in sectors such as accountancy, engineering, data analysts, law, banking, interior decoration, management, medicine, pilot, sports, software development. Among the 1056 respondents, 503 (47.63%) were female and 550 (52%) were male and 3 preferred not to say. The age and gender wise distribution of the respondents is shown in the histogram in Figure 1.

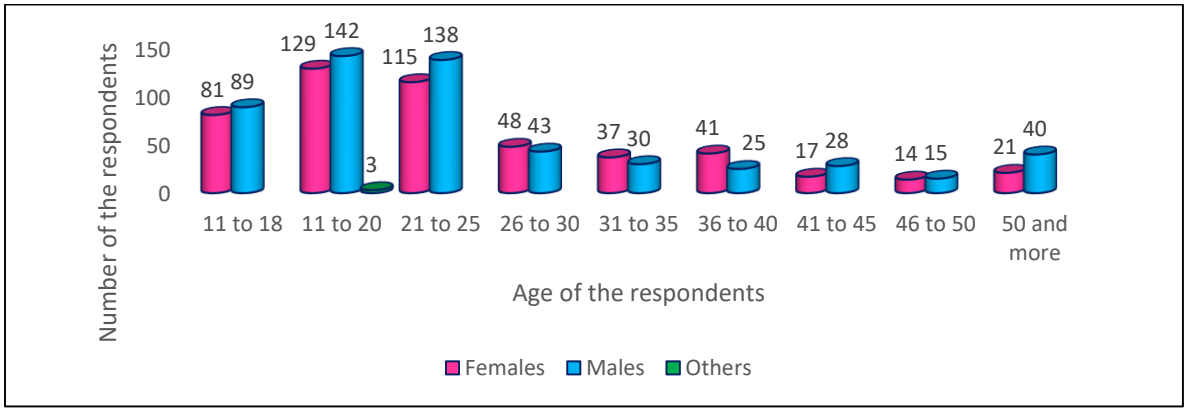


Figure 1: Age-wise and gender-wise distribution of the respondents

As per the data received on the relationship status of the respondents, it was found that 701 (66.38%) were single, 290 (27.46%) were married, 59 (5.59%) are unmarried but committed and remaining 6 included those who are either in a live-in relationship, divorced or widowed. The graphical representation of the same is provided in Figure 2.

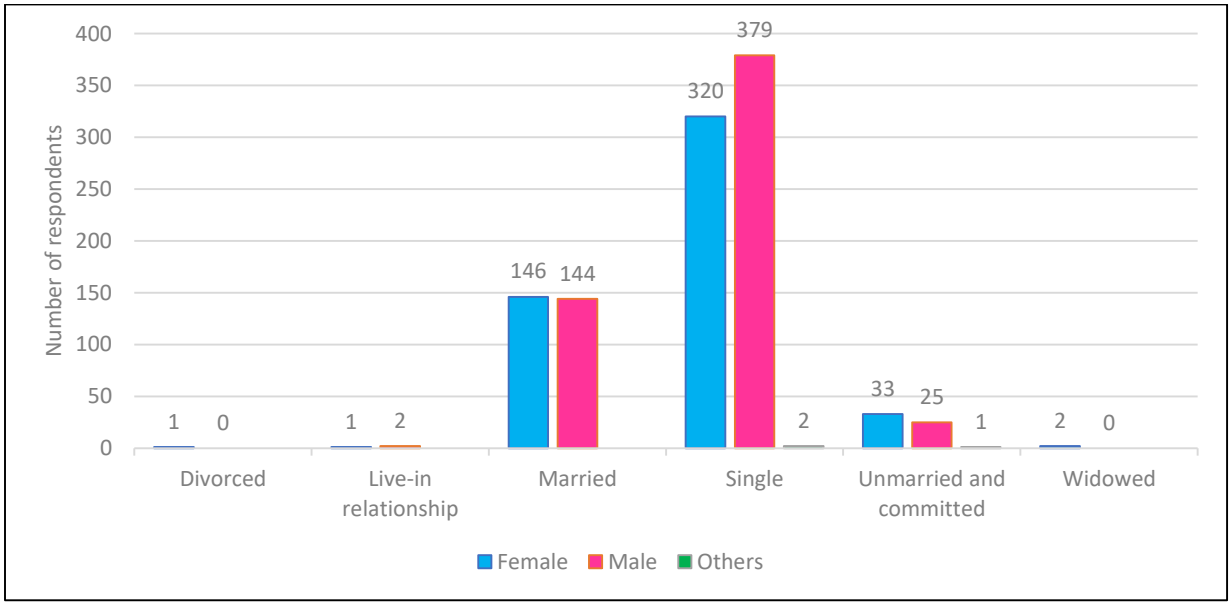


Figure 2*:* Relationship status of the respondents

Study of Attributes Contributing to the Addiction to Social Media

There are various characteristics that contribute to the addiction to social media of a respondent. Some of them are amount of time spent by an individual using social media, number of posts made by an individual in a day and the amount of information revealed through them. Figure 3 exhibits the gender wise distribution of the number of respondents and the corresponding hours spent by them per day on social media through a bi-directional horizontal bar graph.

It was observed that the maximum number of respondents in both male and female response categories devoted 1-2 hours per day on social media and the least number of respondents fall in the 'spend more than

5 hours on social media per day' category. Among the 503 females who undertook the survey, 127 spent 1-2 hours per day on social media and 36 spent more than 5 hours. On the other hand, out of 550 males, 154 spent 1-2 hours per day on social media and 48 spent more than 5 hours.

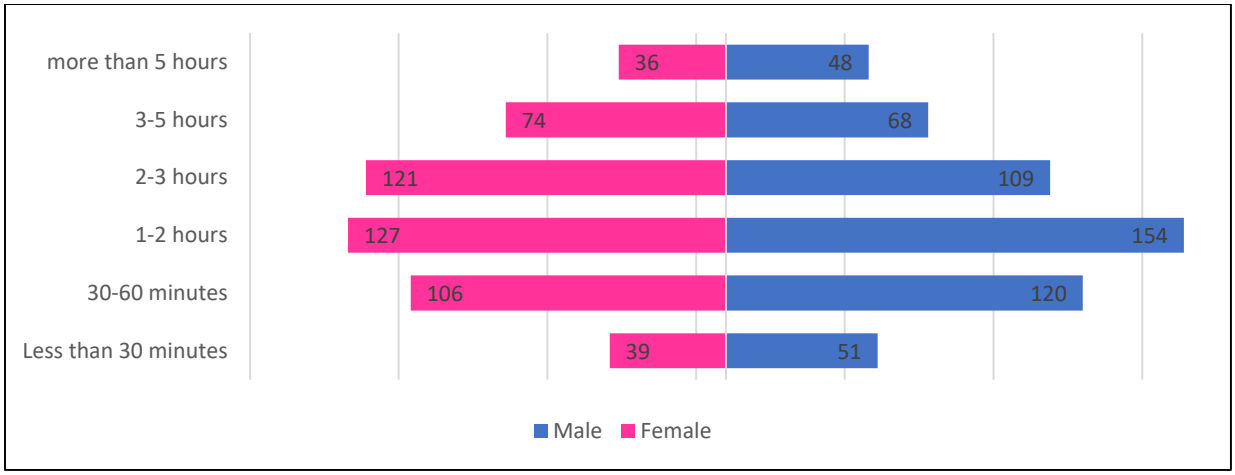
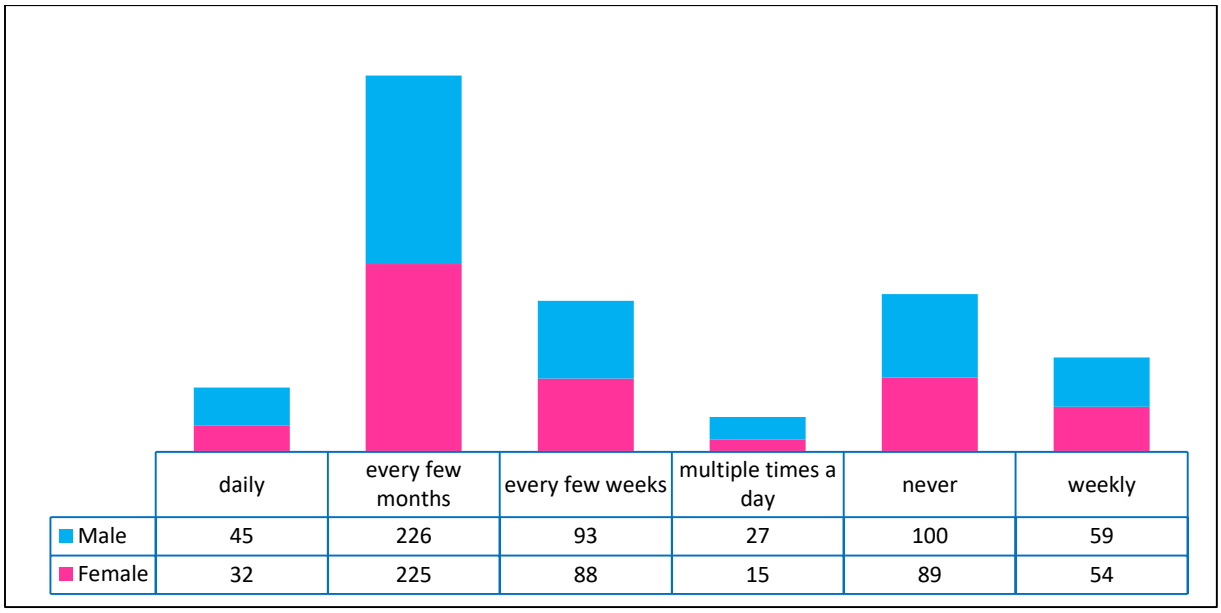


Figure 3: Gender-wise distribution of the number of respondents and the corresponding hours spent by them per day on social media

The divided bar graph of Figure 4 shows gender wise distribution of how often respondents post on social media. It was observed that about 451 respondents (43%) prefer to post every few months on social media. Also, only 42 respondents (4%) prefer to post multiple times a day which is followed by 77 respondents (7%) who posted daily. In all categories, associated with the frequency of posting on social media, it was observed that men outnumber female respondents.



	daily	every few months	every few weeks	multiple times a day	never	weekly
Male	45	226	93	27	100	59
Female	32	225	88	15	89	54

Figure 4*:* Gender-wise distribution of how often respondents post on social media

From the survey, it was found that the majority of the respondents preferred to share very little personal information on social media platforms. This could also be due to the awareness among the respondents about the cyber-crimes and other related disadvantages of the same. Only 1% of the respondents revealed that they shared everything; 4% specified that they share a lot of personal information on social media. Approximately 18% shared no personal information at all and 52% favored to reveal very little personal information on social media as displayed in Figure 5.

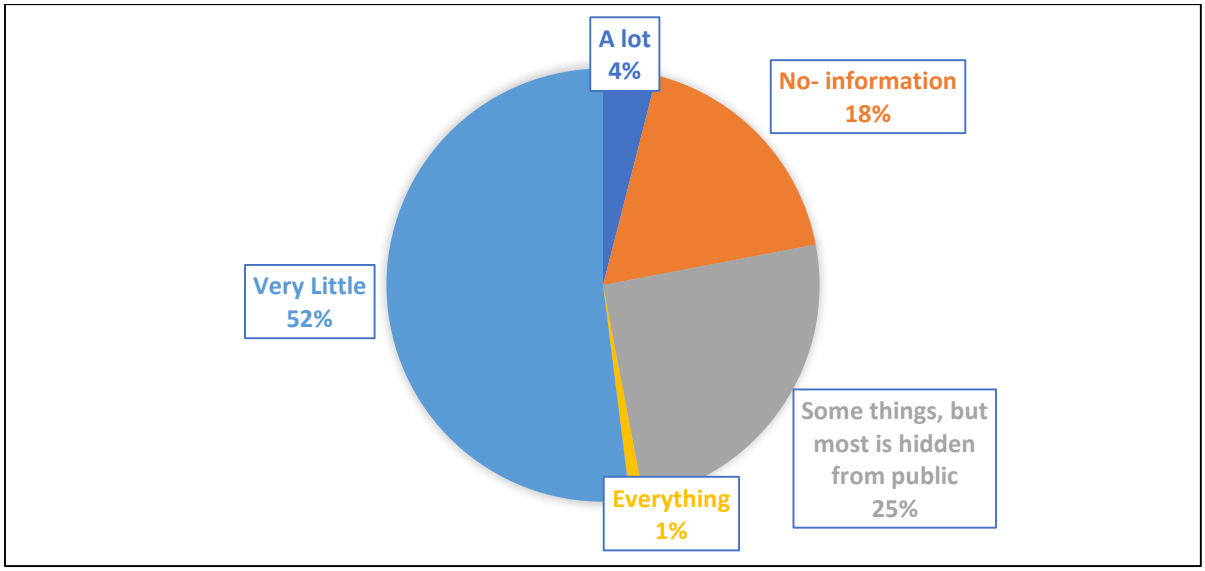
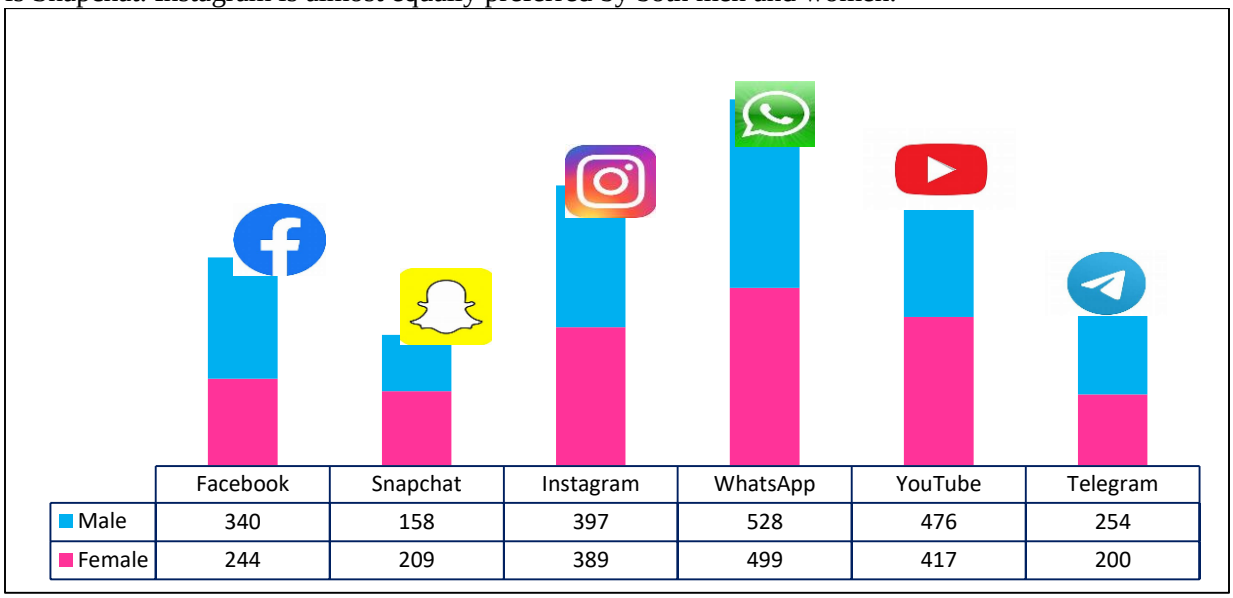


Figure 5*:* Percentage of respondents exhibiting the extent of personal information revealed by them on social media

The divided bar graph in Figure 6 shows the preferred social media platforms of the respondents of the survey. WhatsApp is the most popular platform followed by YouTube whereas the least preferred platform is Snapchat. Instagram is almost equally preferred by both men and women.



	Facebook	Snapchat	Instagram	WhatsApp	YouTube	Telegram
Male	340	158	397	528	476	254
Female	244	209	389	499	417	200

Figure 6: Preferred social media platforms by the respondents

It can be easily observed from the funnel chart drawn in Figure 7 that most of the respondents used social media for keeping in touch with friends and family. It is also a preferred medium to spend free time by generally browsing the web and bringing about awareness by watching news. However, among our respondents, dating was the least preferred choice as the use of social media.

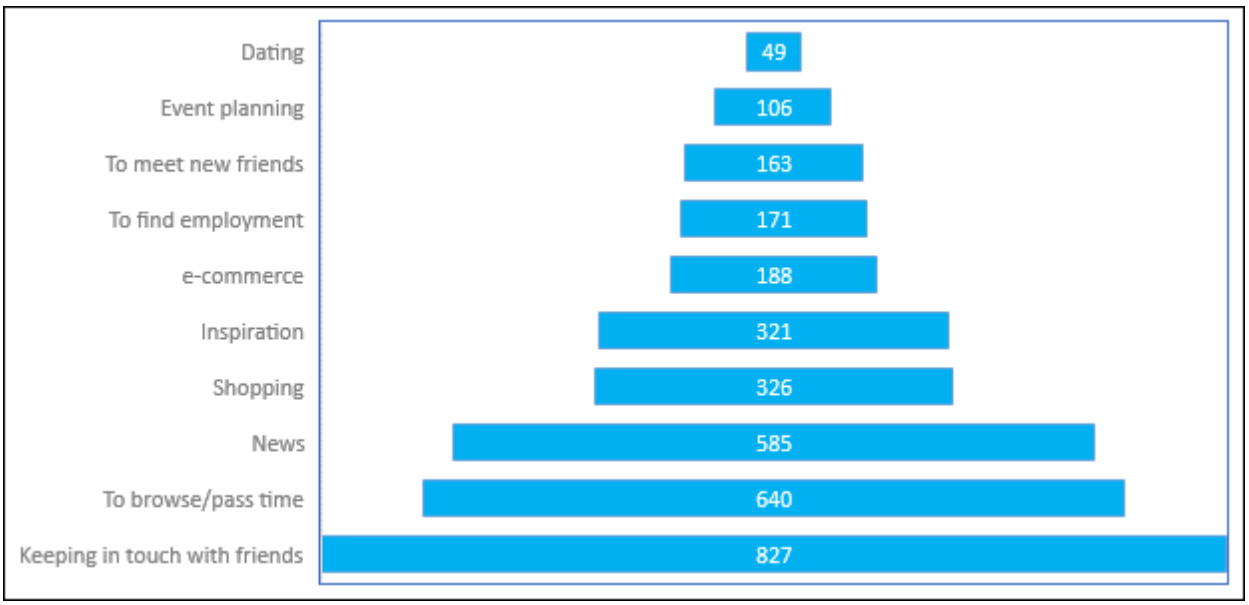


Figure 7: Various uses of social media

Data Analysis Using Various Statistical Techniques

Chi-Square Test for Independence of Attributes

Chi-square test for independence of attributes is conducted to ascertain the relationship between two categorical attributes. It has been used to examine the association between addiction to social media and mental health, studies, relationship status and tendency to postpone important activities.

Interdependence between 'feeling when without social media' and 'mental health'

In the survey, the respondents were asked to choose from one of the three options about how they felt when they were without social media for a long period of time. The options given to them and their corresponding number (indicated in brackets) of responses are:

*A*1: Anxious, Depressed, Lonely, like I am missing out on things (86);
*A*2: I feel like I might be missing out but it's not a big deal (587) and
*A*3: It doesn't affect me in any way (383).

Table 1

Crosstabulation Table

		How do you feel without social media for a long period of time?			
		*A*1	*A*2	*A*3	Total
Do you think use of social media platforms No	Count	20	293	243	556
	Percentage of total	1.9%	27.7%	23%	52.7%

has negatively affected your mental health?	Yes	Count	66	294	140	500
		Percentage of total	6.2%	27.8%	13.3%	47.3%
Total		Count	86	587	383	1056
		Percentage of total	8.1%	55.6%	36.3%	100%

Among the 86 respondents belonging to category $A1$, 66 respondents agreed that the use of social media has negatively affected their mental health while 20 disagreed. Also, 587 respondents fall in the category $A2$ in which 294 respondents felt a negative effect on their mental health due to use of social media whereas 140 respondents felt the same from the total of 383 respondents falling in $A3$ category. It was also observed that 500 respondents agreed that the use of social media platforms by them has negatively affected their mental health and remaining 556 reported against it. This is mentioned in Crosstabulation Table 1 and is graphically represented through cluster bar graph of Figure 8.

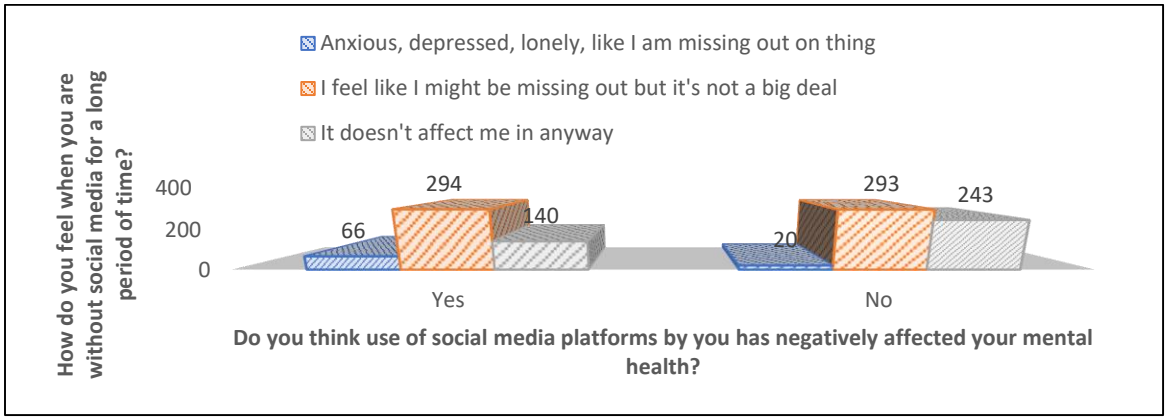


***Figure 8**:* Cluster Bar graph for Table 1

When the two attributes, viz., 'feeling when without social media' and 'mental health' are tested for independence using the Chi-square test in SPSS, the p-value was found to be 0.000 as given in Table 2. This suggested that they are dependent upon each other at 5% level of significance.

Table 2
Chi-Square tests

	Value	Degrees of Freedom	p-Value (two-tailed)
Pearson Chi-Square	49.476	2	0.000
Likelihood Ratio	51.012	2	0.000
Valid Cases	1056		

Interdependence between 'Studies' and 'Time Spent on Social Media'
Among the 1056 respondents of the survey, 611 were students. An analysis was done to understand the interdependence between the time spent on social media and its impact on their studies.

Table 3
Crosstabulation Table

			How many hours do you spend on social media per day?			
			Less than 1 hour	Between 1 hour and 3 hours	More than 3 hours	Total
Do you think use of social media platforms has negatively affected your studies?	No	Count	77	105	41	223
		Percentage of total	12.6%	17.2%	6.7%	36.5%
	Yes	Count	66	203	119	388
		Percentage of total	10.8%	33.2%	19.5%	63.5%
Total		Count	143	308	160	611
		Percentage of total	23.4%	50.4%	26.2%	100%

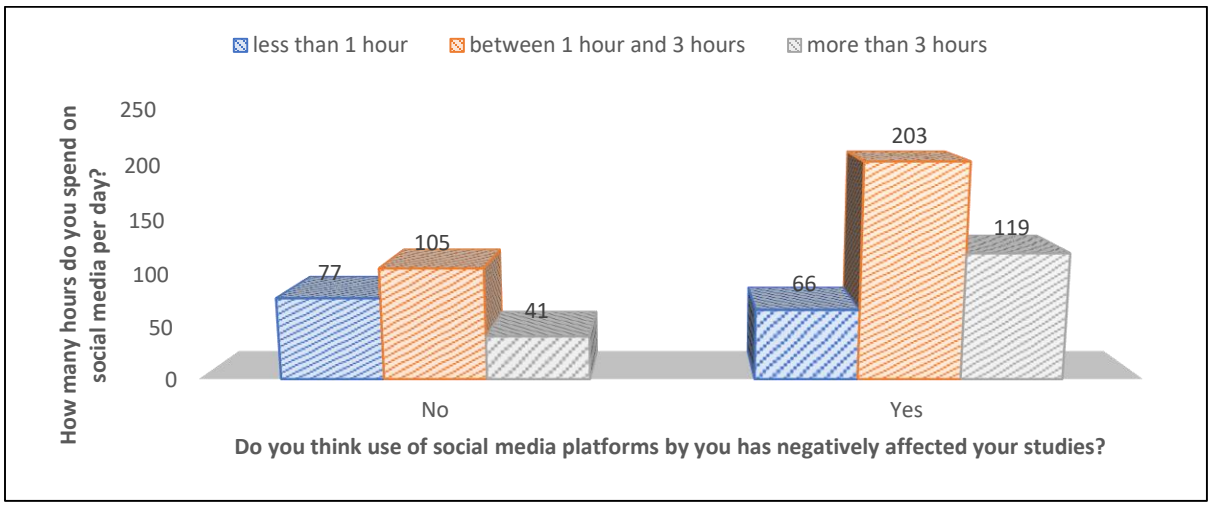


***Figure 9**:* Cluster Bar graph for Table 3

All the data received on these subjects was cross-tabulated and presented in Table 3 and graphically in Figure 9. In order to examine the dependence relationship between the amount of time spent on social media in a day and the negative effect on studies, the chi square test was applied at 5% level of significance and it was found that the p-value was 0.000, as shown in the last column of Table 4. This leads to the conclusion that the time spent by students on social media negatively affects their studies, which is quite expected.

Table 4
Chi-Square Tests

	Value	Degrees of Freedom	p-Value (two-tailed)
Pearson Chi-Square	27.500	2	0.000
Likelihood Ratio	27.157	2	0.000
Valid Cases	611		

Interdependence between 'Addiction to Social Media' and 'Relationship Status'

Chi-square test was conducted to analyze if the relationship status of the respondents bears any connection with the addiction to social media.

Table 5

Crosstabulation Table

			Relationship Status		Total
			Single	Not Single	
Do you consider yourself addicted to social media?	No	Count	426	243	669
		Percentage of total	40.3%	23%	63.4%
	Yes	Count	278	109	387
		Percentage of total	26.3%	10.3%	36.6%
	Total	Count	704	352	1056
		Percentage of total	66.7%	33.3%	100%

It was found that among the 704 respondents who are 'Single' (comprising of single, widowed or divorced categories), 278 confessed being addicted to social media and out of 352 respondents who are 'Not Single' (comprising of married, living in a live-in relationship and unmarried but committed categories), 109 respondents declared themselves addicted to social media. The information is cross tabulated in Table 5 and graphically represented using cluster bar graph in Figure 10.

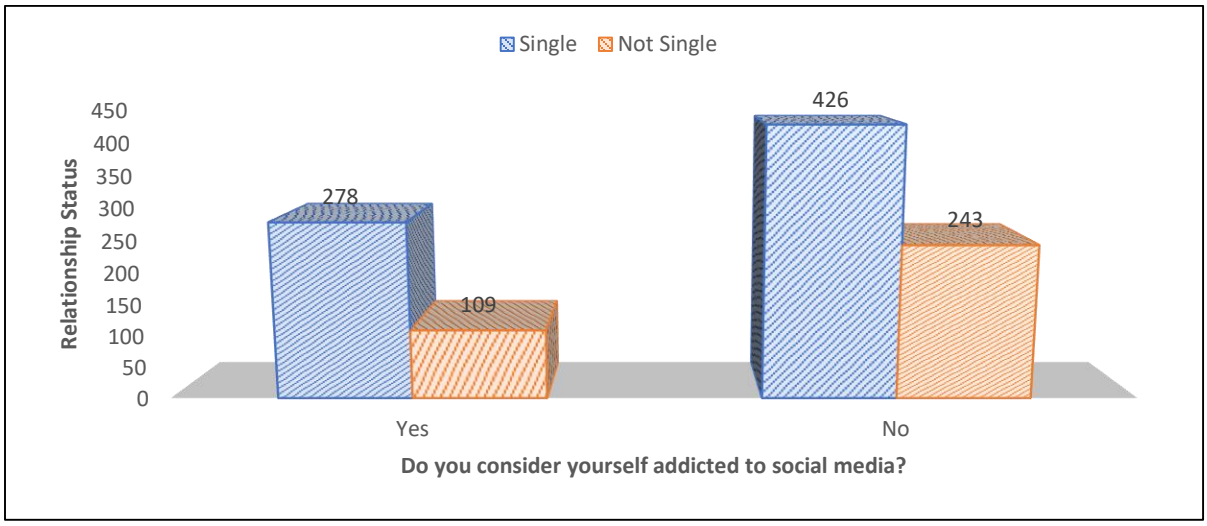


***Figure 10**:* Cluster Bar graph for Table 5

On conducting the chi-square test at 5% level of significance, the p-value in Table 6 suggested rejection of the hypothesis of independence of the two attributes. Thus, the relationship status bears a direct consequence on addiction to social media.

Table 6

Chi-Square Tests

	Value	Degrees of Freedom	p-Value (two-tailed)
Pearson Chi-Square	7.342	1	0.007
Likelihood Ratio	7.440	1	0.006
Valid Cases	1056		

Interdependence between 'Addiction to Social Media' and 'Postponement of Academic/Important Activity for Spending More Time on Social Networking Sites'

The survey enquired about how often the respondents postpone their academic/important activity for spending more time on social networking sites. The respondents were asked to respond on a 5-Point Likert Scale: Never, Rarely, Sometimes, Often, Always.

The crosstabulation between 'Addiction to Social Media' and 'Postponement of academic/important activity for spending more time on social networking sites' is shown in Table 7 which is further graphically represented in Figure 11.

Table 7
Crosstabulation Table

			Do you usually postpone your important academic activity for spending more time on social networking sites?					
			Never	Rarely	Sometimes	Often	Always	Total
Do you consider yourself addicted to social media?	No	Count	423	125	93	22	6	556
		Percentage of total	40.1%	11.8%	8.8%	2.1%	0.6%	52.7%
	Yes	Count	146	83	90	44	24	500
		Percentage of total	13.8%	7.9%	8.5%	4.2%	2.3%	47.3%
Total		Count	569	208	183	66	30	1056
		Percentage of total	53.9%	19.7%	17.3%	6.2%	2.8%	100%

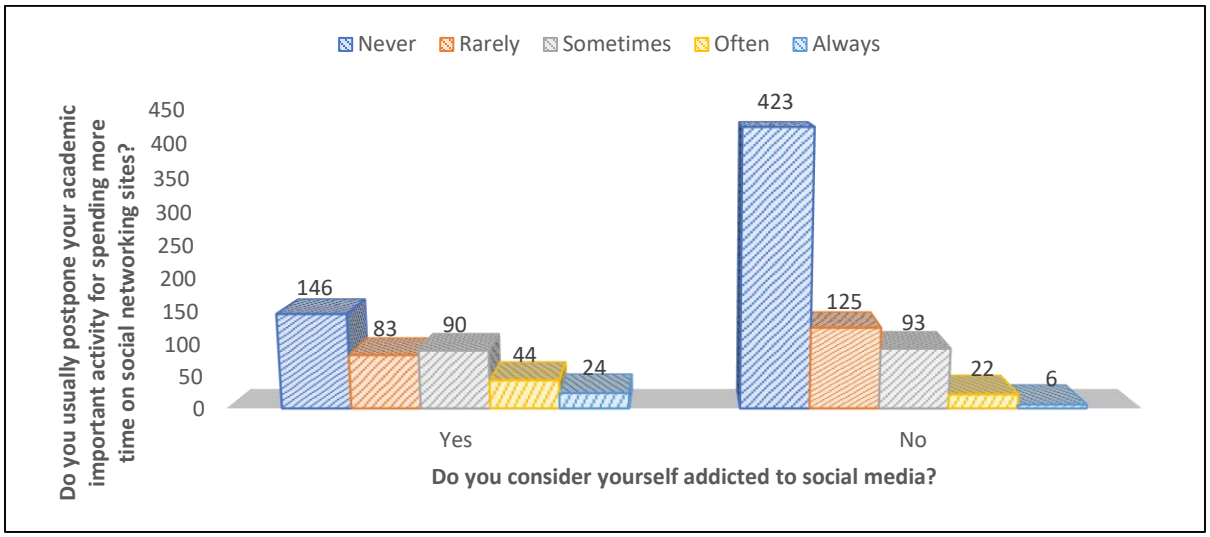


Figure 11: Cluster Bar graph of Crosstabulation Table 7

To study the relationship between the above two stated attributes, a Chi-square test was conducted. The p-value obtained in Table 8 indicated that there is a dependence between addiction to social media and postponement of important academic activities for spending more time on social networking sites.

Table 8
Chi-Square Tests

	Value	Degrees of Freedom	p-Value (two-tailed)
Pearson Chi-Square	92.825	4	0.000
Likelihood Ratio	92.156	4	0.000
Valid Cases	1056		

Cronbach's Alpha

Cronbach's Alpha measures reliability or internal consistency. It tests to see if multiple-question Likert scale surveys are reliable. These questions measure latent variables which are difficult to measure in real life. This test tells us how closely related a set of test items are as a group.

Through the survey, six questions were asked to measure the physical health of the respondents and seven questions were asked to measure their mental health. Each question was a 5-point Likert item: 'Never', 'Rarely', 'Sometimes', 'Often', and 'Always'. In order to understand whether all of these questions reliably measure the same latent variable (Physical health/ Mental health), a Cronbach's Alpha has been used on the responses obtained through survey.

Physical Health

The six contributing factors under study with regard to physical health of the respondents were
A: effect of social media on sleep,
B: effect of social media on interest in daily activities
C: effect of social media on daily exercise routine
D: frequency of checking social media before going to bed
E: frequency of checking social media immediately after getting up
F: sedentary lifestyle.

It was found that Cronbach's Alpha is 0.808 which indicates a high level of internal consistency for our scale with this specific sample. The value of 0.808 indicates that the above-mentioned contributing factors, A, B, C, D, E and F bear a direct relationship on physical health of the respondents.

Table 9
Inter-item Correlation Matrix

	A	B	C	D	E	F
A	1.000	0.608	0.532	0.354	0.305	0.408
B	0.608	1.000	0.657	0.296	0.307	0.439
C	0.532	0.657	1.000	0.255	0.293	0.427
D	0.354	0.296	0.255	1.000	0.590	0.364
E	0.305	0.307	0.293	0.590	1.000	0.403
F	0.408	0.439	0.427	0.364	0.403	1.000

Table 10
Factor's Statistics Table

Factor	Scale Mean, if factor deleted	Scale Variance, if factor deleted	Corrected Factor-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha, if factor deleted
A	14.46	19.826	0.608	0.436	0.770
B	14.56	19.513	0.644	0.538	0.762
C	14.66	18.882	0.591	0.477	0.773
D	13.12	20.793	0.507	0.389	0.791

E	13.72	19.599	0.509	0.396	0.793
F	14.46	19.856	0.559	0.314	0.780

The Table 9 of inter-item correlation matrix indicates the correlation between various contributing factors mentioned above. The final column of Table 10 shows the value of Cronbach's Alpha if that particular item was deleted from the scale. It is observed that the removal of any of the six contributing factors would not lead to a higher value of Cronbach's Alpha. Thus, all the contributing factors play a vital role and have a relevant relationship with the physical health of the respondents.

Mental Health

The contributing factors, $A, B\ C, D, E, F$ and G, under study with regard to mental health of the respondents were considered to be 'effect on sleep', 'effect on interest in daily activities', 'effect on mood due to lack of access to internet', 'feeling of jealousy due to seemingly extravagant lives of others on social media', 'judgmental of self by the number of likes one gets on social media', 'feeling sad when others get more likes than self' and 'postponement of academic/important activities for spending more time on social networking sites' respectively.

It was found that the actual value of Cronbach's Alpha is 0.766 which indicates a moderately high level of internal consistency for our scale with this specific sample. Table 11 shows the correlation between various contributing factors mentioned above. The final column of the Table 12 indicates that the removal of any of the seven contributing factors would not lead to a higher Cronbach's Alpha value. Thus, all the contributing factors play an important role and have a pertinent relationship with the mental health of the respondents.

Table 11

Inter-item Correlation Matrix

	A	B	C	D	E	F	G
A	1.000	0.358	0.217	0.135	0.245	0.279	0.259
B	0.358	1.000	0.477	0.414	0.311	0.261	0.296
C	0.217	0.477	1.000	0.627	0.287	0.197	0.232
D	0.135	0.414	0.627	1.000	0.270	0.186	0.200
E	0.245	0.311	0.287	0.270	1.000	0.425	0.433
F	0.279	0.261	0.197	0.186	0.425	1.000	0.608
G	0.259	0.296	0.232	0.200	0.433	0.608	1.000

Table 12

Factor's Statistics Table

Factor	Scale Mean, if factor deleted	Scale Variance, if factor deleted	Corrected Factor-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha, if factor deleted
A	12.04	18.731	0.376	0.177	0.759
B	12.90	17.071	0.531	0.339	0.727
C	13.35	18.277	0.498	0.458	0.735
D	13.60	19.354	0.451	0.418	0.745
E	13.18	17.624	0.506	0.279	0.732
F	12.50	17.255	0.511	0.413	0.731
G	12.59	17.091	0.532	0.420	0.727

Z-Test for Difference of Proportions

Z-test for difference of proportions is a statistical hypothesis test used to determine whether two proportions, based on specific characteristics, significantly differ from each other. In the survey, it has been conducted to check the equality between male and female respondents on being addicted to social media. The positive responses to the dichotomous survey questions have been provided in Figure 12. The Z-test for proportions showed that both, male as well as female respondents, equally consider themselves addicted to social media. Both the categories (male and female) of respondents agreed that they plan events/get-togethers/trips so that they have something nice to post on social media for that day. Female respondents were more in favor of application of filters to the photographs before posting them on social media, spending a lot of time selecting the correct caption for the picture, and considering social media sites as a form of escape from daily problems than male respondents. On the other hand, the male respondents were pro exhibiting different personalities on social media pages and agreed to social media negatively affecting their relationships with friends and family than the female counterparts.

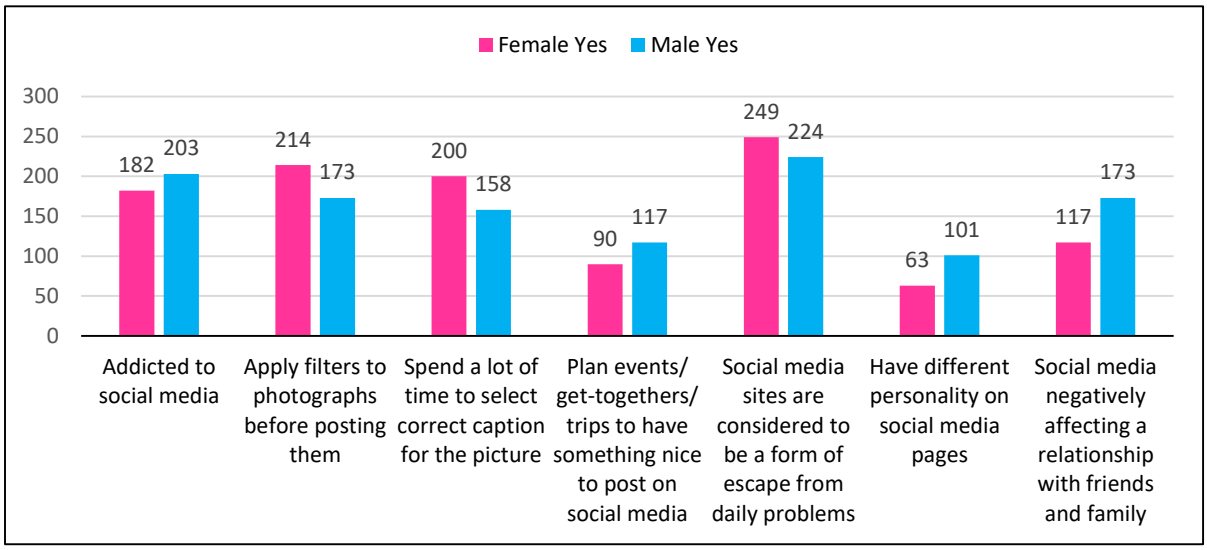


***Figure 12**:* Gender-wise break-up of the number of respondents reporting 'Yes' to dichotomous survey questions

Logistic Regression Analysis

Logistic Regression predicts a binary outcome based on a set of independent variables. Peng et al. (2002) reviews the analysis for logistic regression. In this section, logistic regression has been used to check if addiction to social media can be predicted on the basis of the following five variables:

X_1: number of social media accounts,
X_2: duration of sleep,
X_3: time spent on phone at a stretch,
X_4: hours spent on social media,
X_5: feeling of an individual when he/she stays away from social media.

The dependent variable is taken to be 'addiction to social media' which has been measured on a dichotomous scale - 'yes' and 'no'. Among the above listed five independent variables, the first three variables X_1, X_2 and X_3 are continuous variables and the last two variables X_4 and X_5 are categorical. The variable X_4 (hours spent on social media) is measured on a trichotomous scale (less than 1 hour, 1 to 3 hours or more than 3 hours) and the variable X_5 (feeling of an individual when he/she stays away from social media) is also measured on a trichotomous scale (Anxious/Depressed, I feel like I might be missing out but it's not a big deal or it doesn't affect me in any way). The Box-Tidwell test was conducted using

SPSS to confirm the existence of a linear relationship between the continuous independent variables and the logit transformation of the dependent variable.

Table 13
Classification Table

Observed response	Predicted response		
	Yes	No	Percentage Correct
Yes	237	150	61.2
No	93	576	86.1
Overall Percentage			77.0

Table 14
Variables in the Equation

Variables	B (Coefficients)	Standard Error	Wald	Degrees of freedom	p-value	Exp(B)
X_1	-.484	.060	64.022	1	.000	.616
X_2	.369	.070	27.872	1	.000	1.446
X_3	.066	.018	12.809	1	.000	1.068
X_4			57.832	2	.000	
$X_4(1)$	1.898	.252	56.870	1	.000	6.671
$X_4(2)$	1.034	.201	26.503	1	.000	2.813
X_5			47.606	2	.000	
$X_5(1)$	-2.461	.386	40.667	1	.000	.085
$X_5(2)$	-.921	.191	23.231	1	.000	.398
Constant	-.560	.622	.809	1	.369	.571

Tables 13 and 14 exhibit the output obtained after conducting Binomial Logistic Regression analysis using SPSS. From the Classification Table 13, it is observed that the percentage of cases that can be correctly classified as "Not addicted to social media" are found to be 77%. Sensitivity is observed at 61.2% which is the percentage of the cases having the observed characteristic "Addicted to Social Media" and are correctly predicted by the model. Specificity is detected to be 86.1% which is the percentage of cases that do not have the observed characteristic "Addiction to Social Media" and are also correctly predicted. The positive predicted value, that is the percentage of correctly predicted cases having the observed characteristic "Addicted to Social Media" compared to the total number of cases predicted to have the observed characteristic, is 71.8%. The negative predicted value, which is the percentage of correctly predicted cases not possessing the observed characteristic "Addicted to Social Media" compared to the total number of cases predicted not to have the observed characteristic, comes out to be 79.3%.

The Table 14 shows the contribution of each independent variable to the model and its statistical significance using the Wald's Test. From this table, it can be seen that X_1(p=0.000), X_2 (p=0.000), X_3 (p=0.000), X_4 (p=0.000) and X_5 (p=0.000) added significantly to the model. Using the second column of Table 16 on coefficients, the fitted Binary Logistic Regression Model can be written as

$$P = \frac{1}{1 + e^{-(-0.560 - 0.484X_1 + 0.369X_2 + 0.066X_3 + I_4 + I_5)}}$$

where

$$I_4 = \begin{cases} 0, & X_4 = \text{less than 1 hour} \\ 1.898, & X_4 = 1 \text{ to 3 hours} \\ 1.034, & X_4 = \text{more than 3 hours} \end{cases}$$

and

$$I_5 = \begin{cases} 0, & X_5 = \text{Anxious/Depressed} \\ -2.461, & X_5 = \text{ I feel like I might be missing out but it's not a big deal} \\ -0.921, & X_5 = \text{it doesn't affect me in any way} \end{cases}$$

As an illustration, consider a respondent having values of the variables as $X_1 = 7$, $X_2 = 4$, $X_3 = 4$, $X_4 =$ "more than 3 hours" and $X_5 =$ "Anxious/Depressed", P evaluates to 0.1494 which is significant at 5% level. Therefore, the particular respondent can be classified as addicted to social media.

Discriminant Analysis

Discriminant analysis is a multivariate technique used to ascertain the relationship between the categorical dependent variable and continuous independent variables. In the survey, this technique has been used to classify whether a teenage respondent is addicted to social media depending upon a set of continuous numeric independent variables. All the tables associated with the analysis have been generated through SPSS software.

Among the 1056 respondents of the survey, 327 were teenagers between the age-group 13 to 19. The independent variables include five continuous numeric variables, viz., number of social media accounts that one has (X_1), number of hours one sleeps in a day (X_2), amount of time (in hours) spent on phone at a stretch (X_3), age of the respondent (X_4) and amount of time (in hours) spent on social media (X_5). Addiction to social media (Y) is the dependent categorical variable having two groups 'Yes' or 'No'.

Table 15

Canonical Discriminant Function's Unstandardized Coefficients

Variable	Unstandardized Coefficients
X_1: Number of social media accounts one has	0.063
X_2: Number of hours one sleeps in a day	0.041
X_3: Amount of time (in hours) spent on phone at a stretch	0.185
X_4: Age of the respondent	0.071
X_5: Amount of time (in hours) spent on social media	0.399
Constant	-3.342

From Table 15, the discriminant model can be written as

$$Z = -3.342 + 0.063\,X_1 + 0.041\,X_2 + 0.185\,X_3 + 0.071\,X_4 + 0.399\,X_5$$

Table 16

Canonical Discriminant Function's Standardized Coefficients

Variable	Standardized Coefficients
X_1: Number of social media accounts one has	0.249
X_2: Number of hours one sleeps in a day	-0.192
X_3: Amount of time (in hours) spent on phone at a stretch	0.519
X_4: Age of the respondent	0.055
X_5: Amount of time (in hours) spent on social media	0.781

From Table 16, the value of coefficients for X_5 is the highest at 0.781, thereby highlighting the amount of time (in hours) spent on social media to be the best indicator.

It can be observed that from the Tables 17 and 18, the eigen value is 1.002 (>1) and Wilks' Lambda has a significant p-value at 5% level of significance. These two indicate that the two groups differ significantly.

Table 17
Eigenvalues

Function	Eigenvalue	Percentage of variance	Cumulative percentage	Canonical Correlation
First Canonical discriminant function	1.002	100	100	0.773

Table 18
Wilks' Lambda

Test of Function	Wilks' Lambda	Chi-square	Degrees of freedom	p-value
First Canonical discriminant function	0.403	293.169	5	0.000

Using the Table 19 for 'Classification results', it is observed that 45 out of 327 cases have been misclassified by the given discriminant function. This yields an accuracy of 86.2% of original grouped cases that have been correctly classified.

Table 19
Classification Results

Addiction to social media		Predicted Group Membership		Total
		Yes	No	
Original Count	Yes	98	29	127
	No	16	184	200
Percentage	Yes	77.2	22.8	100
	No	8	92	100

Table 20
Unstandardized Canonical Discriminant Functions at Group Centroids

Addiction to Social Media	Function Coefficient
Yes	1.252
No	-0.795

The discriminant criteria can be calculated using

$$\frac{('Yes'Coefficient \times Count\ of\ 'Yes') + ('No'Coefficient \times Count\ of\ 'No')}{Total\ count\ of\ 'Yes'\ and\ 'No'}$$

Using the values of count of 'Yes' and 'No' and their coefficients from Tables 19 and 20 respectively, the Discriminant Criteria is obtained as 0.00001223. This value of Discriminant Criteria is used to predict the classification based on the values of the independent variables. If for a particular set of values of independent variables, the discriminant model evaluates Z to be greater than the value of discriminant criteria, then the respondent will be considered as addicted to social media whereas if Z is less than the value of discriminant criteria, then the respondent will be considered as not addicted to social media.
As an illustration, when the set of values considered for X_1, X_2, X_3, X_4 and X_5 are 10, 5, 3, 15, 6 respectively, then the value of Z is obtained as 1.507, indicating that the person is addicted to social media. Similarly, another set of values considered for X_1, X_2, X_3, X_4 and X_5 as 5, 8, 0.5, 17, 2, yields a negative value for Z, that is -0.6015, implying that the individual with these set of responses is not addicted to social media.

Conclusion

Social media undeniably plays a crucial role in our lives. The increased use of social media often crosses the thin line between using social media for our benefit and addiction to social media. In this paper, the characteristics that contribute to addiction to social media were considered. Based on the responses gathered, it was observed that most men and women spend one to two hours per day and post once every few months on social media revealing very less personal information. Most people use social media for being connected with friends and family. Upon using Chi-square tests, it was found that the feeling that one experiences when one is without social media negatively affects mental health. From a student's perspective, the amount of time that he/she spends on social media negatively affects their studies. The relationship status of an individual is also dependent on addiction to social media. The dependence is also observed through Chi-square tests between postponement of important academic activities for spending more time on social networking sites.

Using Cronbach's Alpha, it was observed that the effect of social media on sleep, interest in daily activities, daily exercise routine, frequency of checking social media before going to bed, and immediately after getting up, sedentary lifestyle are six factors that contribute in affecting physical health of the respondent. Further, with regard to mental health of the respondents, the seven contributing factors are effect on sleep, effect on interest in daily activities, effect on mood due to lack of access to internet, feeling of jealousy due to seemingly extravagant lives of others on social media, judgmental of self by the number of likes one gets on social media, feeling sad when others get more likes than self and postponement of academic/important activities for spending more time on social networking sites. Also, both, male as well as female respondents, equally consider themselves addicted to social media.

Logistic regression and discriminant analysis reveal the same patterns in a set of data. The functional form for both of the techniques is the same and differ only in the method of estimation of their coefficient. Both the statistical techniques can be used to assess if a respondent falls in the category of addiction to social media or not. The data collected from respondents classifies the majority of them as addicted to social media thereby negatively affecting their physical and mental health.

References

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334.

Fisher, R.A. (1936). The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics*, **7**(2), 179-188.

Iannotti, R. J., Janssen, I., Haug, E., Kololo, H., Annaheim, B., Borraccino, A. & HBSC Physical Activity Focus Group. (2009). Interrelationships of adolescent physical activity, screen-based sedentary behaviour, and social and psychological health. *International Journal of Public Health*, 54(2), 191-8.

Keles, B., McCrae, Niall & Grealish, A. (2019). A systematic review: the influence of social media on depression, anxiety and psychological distress in adolescents. *International Journal of Adolescence and Youth*, 25(1), 79-93.

Kim, H. H. (2017). The impact of online social networking on adolescent psychological well-being (WB): A population-level analysis of Korean school-aged children. *International Journal of Adolescence and Youth*, 22(3), 364–376.

Lenhart, A., Smith, A., Anderson, M., Duggan, M., Perrin, A. (2015). Teens, Technology and Friendships. Pew Research Center. Retrieved from https://www.pewresearch.org/internet/2015/08/06/teens-technology-and-friendships/

Marino, C., Gini, G., Vieno, A., & Spada, M. M. (2018). The associations between problematic Facebook use, psychological distress and well-being among adolescents and young adults: A systematic review and meta-analysis. *Journal of Affective Disorders*, 226, 274–281. Elsevier B.V

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine*. Series 5. 50(302), 157–175.

Peng, J., Lee, K., & Ingersoll, G. (2002). An Introduction to Logistic Regression Analysis and Reporting. *Journal of Educational Research*, 96, 3-14.

Primack, B. A., & Escobar-Viera, C. G. (2017). Social media as it interfaces with psychosocial development and mental illness in transitional age youth. *Child and Adolescent Psychiatric Clinics of North America*, 26(2), 217–233.

Royal Society of Public Health and Youth Health Movement (2017) #StatusOfMind: Social Media and Young People's Mental Health and Wellbeing. *Royal Society for Public Health*, London. Retrieved from https://www.rsph.org.uk/our-work/campaigns/status-of-mind.html

Verhulst, P. F. (1838). Notice sur la loi que la population suit dans son accroissement. Correspondance mathematique et Physique, *publiee par A. Quetelet*, 10, 113-120.