

CONTENTS

- 1). Introduction
- 2). Methodology of Survey
- 3). Concept of Chi Square Test
- 4). Concept of Regression analysis
 - Simple linear regression
 - Multiple Linear regression
- 5). Data Analysis
 - Association between mental health and Gender using chi square test
 - Individual association between different lifestyle behaviors and Mental Health
 - Combined association between lifestyle behavior and mental health
- 6). Conclusion
- 7). Limitations
- 8). Field experience and difficulties
- 9). Reference
- 10). Questionnaire

Introduction

Mental health includes our emotional, psychological, and social well-being. It affects how we think, feel, and act. It also helps determine how we handle stress, relate to others, and make choices. Mental health is important at every stage of life, from childhood and adolescence through adulthood.

Over the course of the life, if a person experience mental health problems, his thinking, mood, and behavior could be affected.

Good mental health is more than just the absence of mental illness. It can be seen as a state of mental health that allows one to flourish and fully enjoy life.

Everyone experiences down times in life. The ability to cope with negative experiences varies greatly from one person to another and, in large part, determines whether people enjoy their lives.

Healthy lifestyles are important contributors to both physical and mental health. Getting high-quality sleep, engaging in physical activity, and eating well not only have advantages to physical health but also have advantages to mental health such as reduced risk of depression and anxiety and increased psychological well-being. Healthy lifestyles may be especially important for the mental health and well-being of young adults. Emerging adulthood is a time of both developmental and ecological changes, marked by increased responsibility, new roles, and changing life circumstances. This developmental period often coincides with a transition to work or university, with changing routines, academic demands, and living situations, which can disrupt health behaviors. Emerging adults also appear more vulnerable to poorer mental health, which could suggest a role for unhealthy lifestyles contributing to poorer emotional functioning.

One of the essential steps for research on any disease is to study how it affects the population. Scientists study the factors that make an individual vulnerable to an ailment and the number of people the disease affects. These are epidemiological studies and are essential to assess the burden of a particular disease on a region, whether it's a country, a state, or its districts. Based on these studies, governments can tailor their policies to help the people. These studies also serve as a goldmine of data for scientists, who study the cause and therapeutic strategies for a particular disease.

About one in seven persons in India suffered from mental disorders of varying severity in 2017, with depression and anxiety disorders being the most common mental disorders affecting 45.7 million and 44.9 million people each in the country, noted the findings of a first comprehensive study done by Indian Council of Medical Research (ICMR) on disease burden due to mental disorders in India.

The research has shown that there is a significant increase in the incidence of mental health disorders in India over the years. The ICMR findings have clearly noted that approximate 197 million persons, roughly one in seven Indians, suffered from any kind of mental disorders of the year 2017. These include depression, anxiety disorders, schizophrenia, bipolar disorders, idiopathic developmental intellectual disability, conduct disorders, and autism.

According to Prof Balram Bhargav, Director General, ICMR informed ANI "We found that depression is highest in older adults which has significant implications for the ageing population of India. Depression contributed 33.8 per cent of all mental disorder, followed by anxiety disorders (19.0 per cent), idiopathic developmental intellectual disability (10.8 per cent), and schizophrenia (9.8 per cent),"

Paying attention to the mental health of students is the need of the hour. According to a study published in the Asian Journal of Psychiatry, over 53% of Indian university students suffer from moderate to extremely severe depression.

Psychological distress is defined as a state of emotional suffering typically characterized by symptoms of depression and anxiety, which is recognized as a common mental health problem in the community. In general, mental health and related issues are not recognized in public, and with the global pandemic, these 'silent' and insidious issues can go unnoticed. Confinement to physical spaces, lack of mobility, panic buying, fear of contraction, loss of income, adaptation to the new normal and the growing ambiguity were some of the observed collective experiences, affecting the overall well-being during the lockdown. The COVID-19 outbreak has introduced additional stressors that further impinge the mental health status of the general population, making them vulnerable to psychological disorders. Studies have revealed that high distress and anxiety due to COVID-19 has resulted in poor psychological well-being, increase in suicidal tendencies,

exasperated pre-existing mental health conditions to name a few. It has also severely affected the family relationships and social dynamics, increase in cases of domestic violence, and abuse of alcohol.

The Covid-19 pandemic and lockdown haven't been kind to our student population. They are amongst the worst hit emotionally -- 58% of Indian college students experienced a significant increase in their stress levels and severe deterioration in their emotions of anger, anxiety, loneliness, hopelessness, and happiness.

In the current wake of the pandemic, WHO states that India is the most depressed country in the world. One in six children and teenagers aged between 10-19 years of age suffer from depression. As we speak there is one suicide attempt every three seconds and one death by suicide every forty seconds by our youth. These statistics alone are alarming enough to take cognizance of the fact that mental health among students is going to be the next crisis.

Currently, our students are stressed, worried and anxious about the uncertainties of the future

- Most teenagers and adolescents have expressed feeling detached, lonely and extremely bored which creates a lot of frustration. Added to this, there is a sense of social detachment and a loss of social skills, group dynamics and relationships.
- There are many students who have turned towards the virtual world and virtual reality in creating friendships, interacting.
- This excessive online presence or screen time itself is creating mood swings, irritability, social withdrawal, changes in sleeping and eating patterns, difficulty with attention, focus and concentration, and further detachment from family or the real world. There is a significant loss of sports, play, movement, exercise, sleep, rest and outdoor activities. This leads to unhealthy diets, hygiene concerns, skin or hair related issues and emergence of multiple underlying medical concerns such as vitamin D3 deficiency, borderline diabetes, obesity, and increased lipid or cholesterol profiles.
- Families at home are also going through severe stress with conflicts, job stressors, financial problems, parental conflicts and even abusive relationships. These kinds

of stressors and abuse also greatly increase the risk for emotional and psychological setbacks in children. Parents working from home and students' online education make the whole work-life balance environment stressful at home.

As mental health professionals, we are also seeing a large surge of diagnosed mental health disorders such as depression, anxiety, panic attacks, OCD and substance abuse. Setbacks with learning disabilities and dyslexia or dysgraphia are also emerging.

According to Prof Balram Bhargav, Director General, ICMR informed ANI "We found that depression is highest in older adults which has significant implications for the ageing population of India. Depression contributed 33.8 per cent of all mental disorder, followed by anxiety disorders (19.0 per cent), idiopathic developmental intellectual disability (10.8 per cent), and schizophrenia (9.8 per cent),"

Paying attention to the mental health of students is the need of the hour.

According to a study published in the Asian Journal of Psychiatry, over 53% of Indian university students suffer from moderate to extremely severe depression.

The Covid-19 pandemic and lockdown haven't been kind to our student population. They are amongst the worst hit emotionally -- 58% of Indian college students experienced a significant increase in their stress levels and severe deterioration in their emotions of anger, anxiety, loneliness, hopelessness, and happiness.

Mental health is influenced not only by trait markers, general living conditions and major life events, but also an increasing evidence indicates, by simple everyday behaviors that can be altered by an individual.

The current study investigated how self-reported physical activity, mental activity dietary factors and social media usage individually and together predicted differences in mental health and well-being in young adults and whether there were any higher-order interactions among these behaviors in the prediction patterns.

Methodology of Survey

In any type of survey, pre-planning and systematic approach plays an important role to arrive at the best possible conclusion and successful completion in minimum time and cost. Also, its success depends upon resources, quality, timing and integrity of the surveyor who compiles the primary data. So, it is very important to manage all the available resources which make impact of the quality of the survey.

1). PLANNING OF THE SURVEY: A full proof planning is essential part of any statistical survey to complete it in a successful manner at a minimum cost, labour and time. Planning of the survey includes selection of topic, preparation of a short questionnaire and selection of sample covering almost all the area. From this questionnaire, I have selected those questions for analysis which seem to be relevant and in context to the topic and for I got appropriate number of responses.

2). OBJECTIVES OF THE SURVEY: The main aim of the current study is to evaluate the individual and combined associations between different aspects of everyday life and mental health of college student.

3). AREA OF THE SURVEY: To conduct any survey we should select a sample from a population. For performing the survey, we have collected data from the final year students of BA and BSc

4). SAMPLING METHOD: A simple random sampling method was used to collect the data. Total number of respondents in my survey was 113.

5).COLLECTION OF DATA: We were required to collect quantitative data for our project. The different methods for collecting quantitative data are observation method, interview method, questionnaire method experiments etc. For my survey I used online questionnaire method. Data was collected using an online questionnaire(Google form) in English language, during one-month period (march, 2022) The questionnaire was self-administered to seek details regarding the student's height, weight, frequency of exercise, frequency of hobby, time spent on social media and fast food consumption . Mental health status was determined by using General Health Questionnaire (GHQ-12).

6).DATA ANALYSIS: To find the association between gender and mental health we will perform chi square test. To examine the individual relationship between lifestyle behaviors and mental health several simple linear regression analyses were conducted. To assess the combined effect of lifestyle behavior and mental health a multiple linear regression was conducted. After collection of raw data we put them in a coded form so that the analysis becomes easy. To analyze the present data we will use MS Excel.

Chi-Square Test

A chi squared test is a data analysis on the basis of observations of a random set of variables. Usually it is a comparison of two statistical data sets. This test was introduced by **Karl Pearson** in 1900 for categorical data analysis and distribution. So it was mentioned as **Pearson's chi-squared test**.

The chi-square test is used to estimate how likely the observations that are made would be, by considering the assumption of the null hypothesis as true.

A hypothesis is a consideration that a given condition or statement might be true, which we can test afterwards. Chi-squared tests are usually created from a sum of squared falsities or errors over the sample variance.

Chi square test is applicable only for categorical data.

P-Value

P stands for probability here. To calculate the p-value, the chi-square test is used in statistics. The different values of p indicates the different hypothesis interpretation, are given below:

- $P \leq 0.05$; Hypothesis rejected
- $P > 0.05$; Hypothesis Accepted

Probability is all about chance or risk or uncertainty. It is the possibility of the outcome of the sample or the occurrence of an event.

Formula

The chi-squared test is done to check if there is any difference between the observed value and expected value. The formula for chi-square can be written as:

$$X^2 = \sum \frac{(\text{Observed value} - \text{Expected value})^2}{\text{Expected Value}}$$

The chi square test of independence is used to determine the association between the categorical variables. It is considered as a non-parametric test. It is mostly used to test statistical independence.

Regression Analysis

Regression analysis is a statistical technique for estimating the relationship among variables which have reason and result relation. Main focus of univariate regression is analyze the relation between a dependent variable and independent variable and formulates the linear relation equation between dependent and independent variable. Regression model with one dependent variable and more than one independent variables are called multiple linear regression.

Simple linear regression

The simple linear regression model can be expressed as

$$Y = \beta_0 + \beta_1 x + \varepsilon$$

Where the intercept β_0 and the slope β_1 are unknown constants and ε is a random error component.

Simple linear uses the following null and alternate hypotheses:

- Null Hypothesis, $H_0: \beta_1 = 0$
- Alternate Hypothesis, $H_1: \beta_1 \neq 0$

The null hypothesis states that the coefficient β_1 is equal to 0 that means there is no statistically significant relationship between the predictor variable, x, and the response variable, y. Whereas the alternative hypothesis states that there is a statistically significant relationship between x and y.

Multiple linear regression

If there is more than one regressor, it is called multiple linear regression. In general, the response variable Y may be related to k regressors, $x_1, x_2, x_3, \dots, x_k$ so that

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon$$

Multiple linear regression uses the following null and alternative hypotheses:

- Null hypothesis, $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$
- Alternative hypothesis, $H_1: \beta_1 = \beta_2 = \dots = \beta_k \neq 0$

The null hypothesis states that all coefficients in the model are equal to zero that is , none of the predictor variables have a statistically significant relationship with the response variable y.

We will use a model to determine whether engaging time in exercise & hobby, fast food consumption, time spent on social media and BMI are related to mental health of a person. The variables in analysis are as follows:

- Mental health : Dependent variable
- Exercise : Independent variable
- Hobby : Independent variable
- Fast food : Independent variable
- Social media : Independent variable
- BMI : Independent variable

The dependent variable mental health score of a person was assessed using GHQ-12 questionnaire, i.e. a self rated 12 item questionnaire developed by Goldberg used to screen the mental health of participants. Total score ranges from 0 to 36 with a score ≥ 20 are considered severe problems with psychological stress.

Assumptions of a linear model

Linear regression analysis is based on four fundamental assumptions:

1. The dependent and independent variables show a linear relationship between slope and the intercept.
2. The residuals are independent.
3. The residuals have a constant variance at every level of x.
4. The residuals of the model are normally distributed.

The various components of regression analysis:

- **Multiple R:** This is the correlation coefficient. It tells us the strength of the linear relationship
- **R-squared:** is the percentage of the response variable variation that is explained by a linear model. It is always between 0 and 100%. R-squared is a statistical measure of how close the data are to the fitted regression line. It is also known as the coefficient of determination, or the coefficient of multiple determination for multiple regression.
- **Adjusted R square:** The adjusted R-squared value helps us compare regression models with differing numbers of independent variables. It is generally used in multiple linear regression analysis
- **Standard error of the regression:** The standard error of the regression indicates the typical size of the residuals. This statistic shows how wrong the regression model is on average. Lower value signifies that the distances between the data points and the fitted values are smaller. Conveniently, this value uses the measurement units of the dependent variable.
- **Significance F:** This is the p-value for the F-test of overall significance. This test determines whether your model with all of its independent variables does a better job explaining the dependent variable's variability than a model with no independent variables. If this test result is statistically significant, it suggests you have a good model.
- **Coefficient:** It gives the least square estimate for the regression model.
- **P-Values:** The p-values for the coefficients indicate whether the dependent variable is statistically significant. When the p-value is less than

your significance level, you can reject the null hypothesis that the coefficient equals zero. For 95% confidence level, the p-value is 0.05.

- **Confidence Interval:** The confidence interval for a coefficient indicates the range of values that the actual population parameter is likely to fall.

Correlation Coefficient

The correlation coefficient is a measure that determines the degree to which the movement of two different variables is associated. The most common correlation coefficient is generated by the Pearson product moment correlation is used to measure the linear relationship between two variables. The possible range of values of correlation coefficient is -1.0 to 1.0. A correlation of -1.0 indicates a perfect negative correlation whereas a correlation of 1.0 indicates a perfect positive correlation. If the correlation coefficient is greater than zero, it is a positive relationship. Conversely, if the value is less than zero, it is a negative relationship. A value of zero indicates that there is no relationship between the two variables.

The correlation coefficient between two variables X and Y is given by

$$\text{Correlation} = \rho = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{(\sum(X_i - \bar{X})^2)(\sum(Y_i - \bar{Y})^2)}}$$

Where, \bar{X} , is the mean of the variable X and \bar{Y} is the mean of the variable Y.

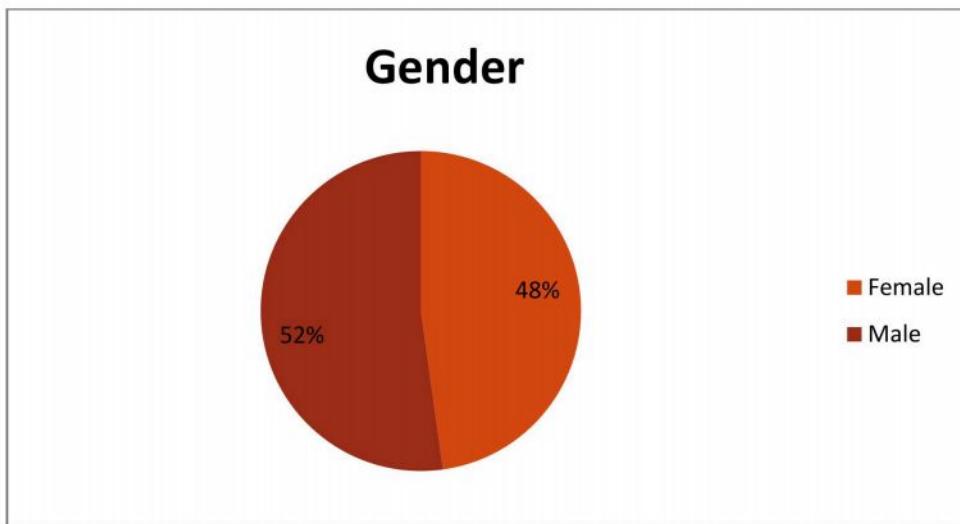
DATA ANALYSIS

Gender wise distribution of respondents:

	No. of respondents	Percentage
Female	54	47.79
Male	59	52.21
Total	113	100

Out of the total 113 respondents, **47.79%** were male and **52.21%** were females.

Graphical Representation of the data



Distribution of Various lifestyle behaviors

1). Physical activity

No. of times exercised in a week	Number of respondents
0	16
1	13
2	18
3	20
4	14
5	10
6	11
7	11
Total	113

2). Hobby

No. of days in a week engaged in Hobby	Number of respondents
0	12
1	9
2	15
3	13
4	10
5	19
6	15
7	20
Total	113

3). Fast food

Fast food consumption	No. of respondents
Never	19
Once a month	24
2 to 4 times a month	34
2 to 3 times a week	19
4(or more) times a week	17
Total	113

4). Social Media

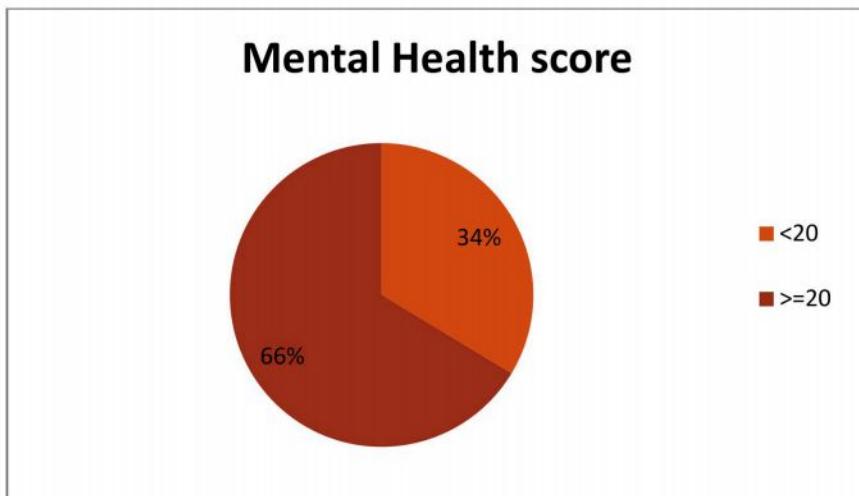
Time spent on social media (in hrs)	Number of respondents
0	10
1	34
2	21
3	20
4	12
5	7
6	9
Total	113

Distribution of mental health score of respondents

Mental Health Score	No. of respondents	Percentage
>=20	38	33.63
<20	75	66.37
Total	113	100

We already know that students scoring above 20 in mental health questionnaire are suffering with some psychological distress, so in our survey out of 113 respondents, 38 respondents have some psychological distress.

Graphical representation of data



Gender-wise distribution of mental health distress

	Suffering	Not Suffering	Total
Male	24	35	59
Female	14	40	54
Total	38	75	113

Here suffering indicates the students who have mental health score >20 and are likely to have some mental health issue.

Chi Square Test for association between Gender and mental health distress

The null and alternate hypotheses for this test will be

H_0 : Gender has no effect on mental health.

H_1 : Gender has effect on mental health.

2x2 contingency table

		Suffering	Not suffering	Total
Male	Observed	24	35	59
	Expected	19.8	39.2	59.0
Female	Observed	14	40	54
	Expected	18.2	35.8	54.0
Total	Observed	38	75	113
	Expected	38.0	75.0	113.0

Chi Square Values

	Value
Chi Square Calculated	2.8
Chi Square Tabulated	3.841

Level of significance: 5%

Degree of Freedom: 1

From the contingency table, we got the calculated chi square value as 2.8. The tabulated chi square value at 5% level of significance for 1 degree of freedom is 3.841.

Since, the calculated chi square value is less than the tabulated chi square value, we will not reject the null hypothesis.

Therefore, at 95% confidence level, we can conclude that there is no association between gender and mental health distress that means gender has no effect on mental health.

Individual Association of Different Lifestyle Behaviors and Mental Health

To assess the individual effect of each lifestyle behaviors on mental health we will perform several simple linear regressions.

1). Relationship between exercise and Mental Health

Dependent Variable: Mental health (MH)

Independent Variable: Exercise

Hypotheses:

Null Hypothesis, H_0 : There is no statistically significant relation between exercise and mental health of a person

Alternate Hypothesis, H_1 : There is statistically significant relation between exercise and mental health of a person

Regression Statistics				
Multiple R	R Square	Adjusted R Square	Standard Error	Observations
0.746402822	0.557117173	0.553127237	4.943162813	113

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	24.60208531	0.818888636	30.04326	3.92E-55	22.9794028	26.22476776
Exercise	-2.514066035	0.212758215	-11.8165	2.35E-21	-2.93566063	-2.092471441

Correlation Coefficient = -0.74640

Interpretation:

- The correlation coefficient for the two variables Exercise and mental health is -0.75. The negative sign indicates that on increasing the frequency of engaging yourself in exercise in a week the mental health score can be reduced.
- Roughly 56% of the variation in the mental health score can be explained by the number of times a person exercises in a week.
- The observed value fall an average of 4.9432 units from the regression line.
- The p-value 0.000 is less than 0.05 hence we will reject the null hypothesis which indicates that there is a statistically significant association between exercise and mental health score.

The estimated regression equation is :

$$\text{MH Score} = 24.6021 - 2.514 \text{Exercise}$$

This tells us that for additional time one exercises, the mental health score is expected to decrease by 2.514, on average. The expected mental health score for a person who exercises 0 times a week is 24.60.

2). Relationship between hobby and mental health

Dependent Variable: Mental Health

Independent Variable: Hobby

Hypotheses:

Null Hypothesis, H_0 : There is no statistically significant relation between hobby and mental health of a person

Alternate Hypothesis, H_1 : There is statistically significant relation between hobby and mental health of a person

Regression Statistics				
Multiple R	R Square	Adjusted R Square	Standard Error	Observations
0.8214626	0.674800867	0.67187114	4.235797	113

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	26.88931521	0.784289524	34.28493	6.94E-61	25.33519321	28.44343721
Hobby	-2.615107492	0.172311764	-15.1766	7.69E-29	-2.95655475	-2.273660239

Correlation Coefficient = -0.8214626

Interpretation:

- The correlation coefficient between hobby and mental health score is -0.82. On increasing the frequency of engaging yourself in hobby in a week the mental health score will be decreased.
- Approx 67% of the variation in the mental health score can be explained by the number of times a person engages himself in a week.
- The observed value fall an average of 4.236 units from the regression line.
- The p-value 0.000 is less than 0.05 hence we can reject the null hypothesis which indicates that there is a statistically significant association between hobby and mental health score.

The estimated regression equation is:

$$\text{MH Score} = 26.889 - 2.6151\text{Hobby}$$

This tells us that for additional time you engage yourself in your hobby the mental health score is expected to decrease by 2.615 on average. The coefficient for the

intercept tells us that the expected mental health score for a person who engages 0 times a week in hobby is 26.89.

3). Relationship between time spent on Social Media and mental health

Dependent Variable: Mental Health

Independent Variable: Social Media

Hypotheses:

Null Hypothesis, H_0 : There is no statistically significant relation between social media usage and mental health of a person

Alternate Hypothesis, H_1 : There is statistically significant relation between social media usage and mental health of a person

Regression statistics				
Multiple R	R Squared	Adjusted R Square	Standard Error	Observation
0.9039936	0.8172044	0.8155576	3.1757278	113

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	7.220962716	0.517614661	13.95046	3.65E-26	6.195274743	8.24665068
Social Media	3.897550231	0.174963575	22.27635	9.13E-43	3.550848237	4.2442522

Correlation Coefficient: 0.90399

Interpretation:

- The correlation coefficient between social media and Mental health is 0.90. It suggests that on increasing the time spent on social media the mental health score is expected to increase.
- Approx 82% of our values fit in the regression analysis model. That is 82% of the dependent variables (i.e. mental health) are explained by the independent variables (i.e. social media).
- The observed values fall an average of 3.176 units from the regression line.
- The p-value 0.000 is less than 0.05, we will reject the null hypothesis which indicates that there is a statistically significant association between social media and mental health score.

The estimated regression equation is:

$$\text{MH Score} = 7.221 + 3.897 \text{Social media}$$

This tells us that for an additional hour spent on social media the mental health score is expected to increase by 3.897 on average. The coefficient for the intercept tells us that the expected mental health score for a person who spends 0 hours on social media is expected to have a mental health score 7.221.

4). Relationship between fast food consumption and mental health

Dependent Variable: Mental Health

Independent Variable: Fast food

Hypotheses:

Null Hypothesis, H_0 : There is no statistically significant relation between fast food consumption and mental health of a person

Alternate Hypothesis, H_1 : There is statistically significant relation between fast food consumption and mental health of a person

Regression Statistics				
Multiple R	R Squared	Adjusted R square	Standard error	Observations
0.8854426	0.7840086	0.7820627	3.4520627	113

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	6.887790974	0.584266902	11.78877	2.72E-21	5.73002714	8.045555
fast food	5.076864608	0.252925053	20.0726	9.8E-39	4.575676746	5.578052

Correlation Coefficient: 0.8854426

Interpretation:

- The correlation coefficient between fast food and Mental health is 0.90. It suggest that on increasing the number of times fast food consumed in a week the mental health score will increase.
- Approx 78% of our values fit in the regression analysis model. That is 82% of the dependent variables (i.e. mental health) are explained by the independent variables (i.e. fast food consumption).
- The observed value fall an average of 3.452 units from the regression line.

- The p-value 0.000 is less than 0.05, hence we will reject the null hypothesis which indicates that there is a statistically significant association between fast food and mental health score.

The estimated regression equation is :

$$\text{MH Score} = 6.888 + 5.0769 * \text{fast food}$$

This tells us that for an additional time fast food consumed in a week the mental health score is expected to increase by 5.0769 on average. The coefficient for the intercept tells us that the expected mental health score for a person who do not eat fast food in a week is expected to have a mental health score 6.888.

5). Relationship between BMI and mental health

Dependent Variable: Mental Health

Independent Variable: BMI

The BMI of a person is calculated using the formula:

$$\text{BMI} = \frac{\text{Weight(KG)}}{\text{Height(m)}^2}$$

The average BMI of the students was 21.55 kg/m²

Hypotheses:

Null Hypothesis, H_0 : There is no statistically significant relation between BMI and mental health of a person

Alternate Hypothesis, H_1 : There is statistically significant relation between BMI and mental health of a person

Regression Statistic				
Multiple R	R Squared	Adjusted R	Standard Error	observation
0.1198944	0.014374	0.00549516	7.3742256	113

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	22.87787616	4.953704785	4.618337	1.05E-05	13.06177974	32.69397
BMI	-0.289525499	0.227552885	-1.27234	0.20591	-0.74043671	0.161386

Observation: Since, the p-value is 0.20591 that is greater than 0.05 we will accept the null hypothesis and hence conclude that there is no effect of a person's BMI on his mental health score. So we will not include this variable in our model further.

Conclusion: From the above results we can say that a person's mental health is affected by his various lifestyle behaviors such as number of times a person exercise in a week, number of times he engages himself in a hobby, number of hours spent on social media and number of times a person consumes fast food in a week. And a person's BMI has no effect on his mental health. Next we will calculate the combined effect of the lifestyle behaviors which has statistically significant affect on a person's mental health.

Combined association between lifestyle behavior and mental health

Dependent Variables: Exercise, Hobby, Social Media, Fast food

Independent Variable: Mental Health

Regression Statistics				
Multiple R	R squared	Adjusted R square	Standard Error	Observation
0.96177997	0.92502071	0.92224370	2.06196246	113

ANOVA					
	df	SS	MS	F	Significance F
Regression	4	5664.94146	1416.235	333.0995	8.99464E-60
Residual	108	459.1824334	4.251689		
Total	112	6124.123894			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	13.29277369	1.0184834	13.05154	6.11E-24	11.27396296	15.31158442
exercise	-0.480305919	0.125917692	-3.81444	0.000228	-0.72989662	-0.230715216
Hobby	-0.701652781	0.132412726	-5.29898	6.21E-07	-0.96411777	-0.439187795
fast food	1.921558913	0.271347824	7.081534	1.52E-10	1.383700467	2.459417359
Social Media	1.625347948	0.216059947	7.522671	1.69E-11	1.197079664	2.053616232

Interpretation:

- 92.5% of the variation is mental health can be explained by number of times exercised, number of times engaged yourself in hobby, number of times one consumes fast food in a week, time spent on social media.
- The observed values fall from an average of 2.062 units from the regression line.
- SS is the ANOVA table is the sum of squares. We consider that smaller the Residual SS compared with the total SS, the better the model fits in the data. Here the residual SS is comparatively than the total SS, we can say that our model fits in the data better.
- The overall F statistic for the regression model is 333.0995.
- The p-value associated with the overall F statistic is 0.00 (<0.05)that indicates that the explanatory variables **exercise**, **hobby**, **fast food**, **social media** taken combined have a statistically significant association with mental health score.

Estimated regression equation:

$$\text{MH Score} = 13.293 - 0.48 * \text{Exercise} - 0.70 * \text{Hobby} + 1.92 * \text{fastfood} + 1.62 * \text{social media}$$

From the above equation we interpret that keeping other variables constant if exercise is increased by one unit mental health decreases 0.48 times. Similarly if hobby is increased by one unit, keeping other variables constant mental health

score decreases 0.70 times. If eating fast food is increased by one unit the mental health score increases 1.92 times. If time spent on social media increases by 1 hour the mental health score increases 1.62 times. If a person does not exercise or engages himself in hobby at all, and does not consume fast food and does not spend time on social media, his expected mental health score will be 13.293.

CONCLUSION

After analyzing all the different aspects of lifestyle behavior, I have concluded the points about the impact of lifestyle behaviors on mental health of a student

- ◆ Engaging oneself in a physical activity (exercise) one can reduce their mental health score 2.514 times. In other words in order to maintain a good mental health one should exercise regularly.
- ◆ Students must engage themselves in hobby or some stress relieving activity. As, our analysis has shown that engaging oneself in some kind of stress releasing activity one can reduce their mental health score 2.615 times and hence improving their mental health.
- ◆ Students must try to reduce their time spent on social media. We have found that for an additional hour spent on social media, the person's mental health score increases 3.897 times.
- ◆ Students must include healthy food in their diets. Our analysis has show that for an additional time fast food consumed in a week, one's mental health score increases 5.08 times.
- ◆ There is no statistically significant relation between a person's BMI and his mental health score.
- ◆ From the multiple linear regression we have found that mental health score gets affected significantly if a person simultaneously does exercise, engages himself in hobby, consumes fast food, and spends time on social media.
- ◆ As seen from the simple linear regression as well as multiple linear regression consumption of fast food has a greater impact on mental health than any other variable. Hence to maintain a

good mental health one should have a balanced diet with less fast food consumption.

- ◆ The number of healthy lifestyle choices predicts a lower mental health score. Healthy lifestyle behaviors seem to have an additive effect on mental health.
- ◆ From the chi square testing we found out that the mental health score is independent of the gender of a person.

LIMITATIONS

Chi Square test

Chi square, like any analysis has its limitations. One of the limitations is that all participants measured must be independent, meaning that an individual cannot fit in more than one category. If a participant can fit into two categories a chi-square analysis is not appropriate. When calculating the number of expected individuals for each class, there should be greater than 5 for each class for the most appropriate use of chi-square. Another consideration one must make is that the chi-square statistic is sensitive to sample size. Most recommend that chi-square not be used if the sample size is less than 50.

Linear Regression Analysis

1. Main limitation of Linear Regression is the **assumption of linearity** between the dependent variable and the independent variables. In the real world, the data is rarely linearly separable. It assumes that there is a straight-line relationship between the dependent and independent variables which is incorrect many times.
2. **Prone to noise and over fitting:** If the number of observations are lesser than the number of features, Linear Regression should not be used, otherwise it may lead to over fit because it starts considering noise in this scenario while building the model.
3. **Prone to outliers:** Linear regression is very sensitive to outliers (anomalies). So, outliers should be analyzed and removed before applying Linear Regression to the dataset.

4. Prone to multicollinearity: Before applying Linear regression, multicollinearity should be removed (using dimensionality reduction techniques) because it assumes that there is no relationship among independent variables.

In summary, Linear Regression is great tool to analyze the relationships among the variables but it over-simplifies real world problems by assuming linear relationship among the variables.

Field experiences and difficulties

This was the first project of my life and I have learnt many things during my project work. The success of this type of survey totally depends upon the quality of questions and honesty of the surveyor who collects the basic data. Also, the accuracy of the data collected and co-operation of the respondents depend on the ability of the investigator in expressing and explaining the objective and aim of the survey to the respondents.

This project was based on the real life applications of many concepts which we have learnt in these three years. This project provided me practical approach of the theory of Statistics which we have learnt so far and it was full of exciting experience. I have faced many difficulties in data collection. The main difficulty was lack of responses due to many reasons.

Keeping all the things in consideration, I have tried my best to get reliable data and information. I came across a variety of responses during the project work.

REFERENCE

- 1).** S.C. Gupta and V.K. Kapoor: Fundamental of mathematical statistics.
- 2).** Goon A.M., Gupta M.K. and Dasgupta B. (1980): An Outline of Statistical Theory, Vol. 2, The World Press Publishers Pvt. Ltd. Calcutta
- 3).** Rohtagi V.K. (1967): An Introduction to Probability Theory and Mathematical Statistics, John Wiley and Sons.
- 4)**.S.C. Gupta and V.K. Kapoor: Fundamentals of applied statistics.
- 5).** www.wikipedia.org
- 6).** www.Google.com
- 7).** www.Statisticsbyjim.com
- 8).** www.investopedia.com
- 9).** www.Statology.org

QUESTIONNAIRE



Impact of healthy lifestyle on mental health of college students

* Required

1. Email *

2. Name *

3. Gender *

Mark only one oval.

Female

Male

Other: _____

4. What is your height (in feet) *

5. What is your weight (in kg) *

6. Do you exercise regularly? If yes, with what frequency have you exercised in the last month? *

Mark only one oval.

Never
 Once a week
 Twice a week
 Thrice a week
 4 times a week
 Other: _____

7. Do you have a hobby/stress releasing activity? If yes, with what frequency have you kept yourself engaged with it in the last month? *

Mark only one oval.

Never
 Once a week
 Twice a week
 Thrice a week
 4 times a week
 Other: _____

8. On an average how many hours of sleep do you get in a day?

Mark only one oval.

- Less than 3
- 4 to 6 hrs
- 6 to 8 hrs
- 8 to 10 hrs
- Other: _____

9. How often do you eat fastfood? *

Mark only one oval.

- Never
- Once a month
- 2 to 4 times a month
- 2 to 3 times a week
- 4(or more) times a week
- Other: _____

10. Your average time spent on Social Media in a day(in hrs) *

Mark only one oval.

- 1
- 2
- 3
- 4
- Other: _____

11. Please read each statement and mark the appropriate option which indicates how much the statement applied to you over the past week. There are no right or wrong answers. (Do not spend too much time on any statement). Have you recently *

Mark only one oval per row.

	Less than usual	No more than usual	Rather more than usual	Much more than usual
Been able to concentrate on what you're doing?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Lost much sleep over worrying about something?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Felt you were playing an important part in things happening around/to you?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Felt capable of making decisions about important things?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Felt constantly under strain (something causing anxiety or worry)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Felt you couldn't overcome your difficulties?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Been able to enjoy your normal day-to-day activities?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Been able to deal with your problems?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Been feeling unhappy and depressed?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Been losing confidence in yourself?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Been thinking of yourself as a worthless person?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Been feeling reasonably happy, all things considered	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>