

## **1. Introduction**

There was a fundamental shift in the education, learning and the overall lifestyle of students due to the lockdown and online mode of learning. This impacted every one of us and especially the student call whose mode of learning went through a dramatic shift which affected their lifestyle. Students had to adapt to the new easy of attending their classes, self-learning, and align their overall social and personal lifestyle around the same. The approach of a more digital way of learning has a certain impact on their mental well-being and as a IB psychology student who got the opportunity to explore the Digital world of psychology, I felt it was interesting and important to explore the impact of online classes and activities on students, which can help develop an idea into the mental well-being of students in relation to the constant change that surrounds them and maybe conduct more research on it. (Pirdehghan et al.)

We will look at a dataset of a survey of the students response to the effects of this change and try to analyse how their online activities (which includes their online classes and online social activities) affected their overall lifestyle including their sleep, fitness, self-study, meal intake, body-weight and more. The data set was obtained from a data repository, which is a storage space for researchers to deposit data sets associated to their research.

## **2. Aim**

- The aim of this investigation is to explore the given data set by :-
- Identify the various behaviour patterns between various age groups for each attribute.
- Identify patterns and trends in each attribute.
- Find correlation and analyse positive relation and influences between variables.
- Impact of health issues on various attributes.
- Create a Logistic Regression Model for the dependant variable "time spend online" i.e. 'Time spent on Online Class' & 'Time spent on social media' to identify the pattern and predict trends towards the patter

### 3. Descriptive Analysis

#### 3.1. Structure of the data set

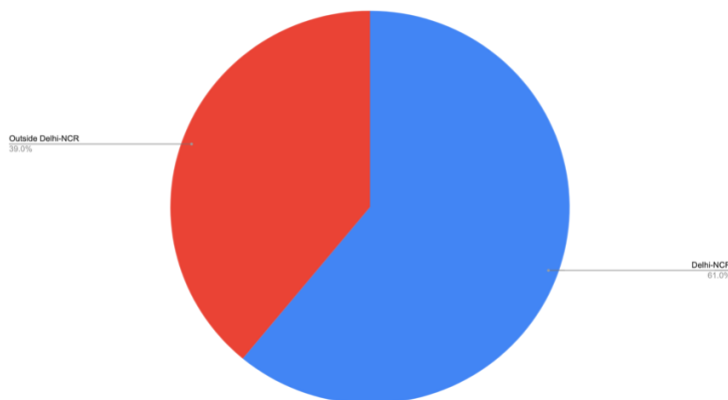
The data set we are analyzing contains 11 columns and 1183 rows. The variables that we are including in this investigation is Region of residence of the students, Age of the subject, Time spent on online Class, Medium of online class, Time spent on self-study, Time spent on Fitness, Time spent on sleep, Time spent on TV, Time spent on social media, preferred social media platform and number of meals per day. A section of the data set is shown in the appendix below.

### 4. Exploratory Analysis

We will now Understand the distribution of the qualitative and quantitative data to get a better view that helps form opinions and knowledge about the dataset.

- Which region do the students belong to?

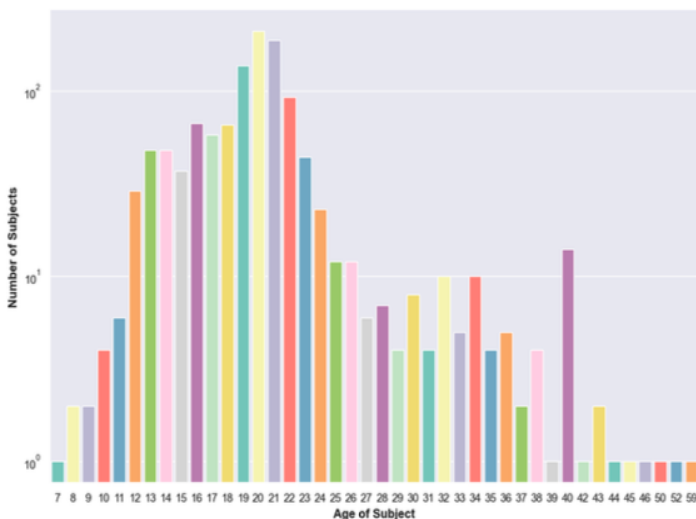
Distribution of Region of residence



From the pie chart we observe that the data set is broken into two categories, Delhi-NCR and outside Delhi-NCR implying the urban and rural breakdown respectively. There is an approximate of 60% to 40% split of urban to rural region of residence in the dataset.

Graph 1- Pie chart showing the Distribution of the region of residence based on the students in the dataset

- What is the age wise distribution of the students in the dataset?

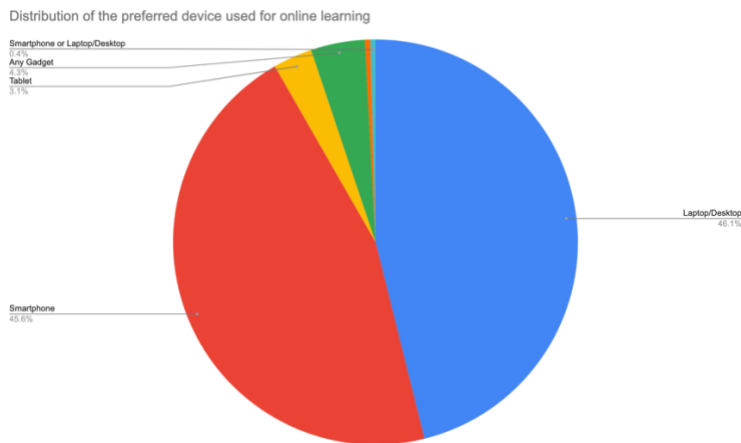


Graph 2 – Bar graph showing the distribution. age wise of students in the data set

From the bar graph we are clearly able to understand the ages of the students who gave the survey. The data here seems right skewed that yet follows a normal distribution. From further analysis, it is clear that the 16-25 age group is represent four times compared to the next age group of 5 to 15 years.

The trend we observe here is that the concentration of the data is between the age group 16 and 25, which would constitute the major students from schools, high schools, and undergraduate studies.

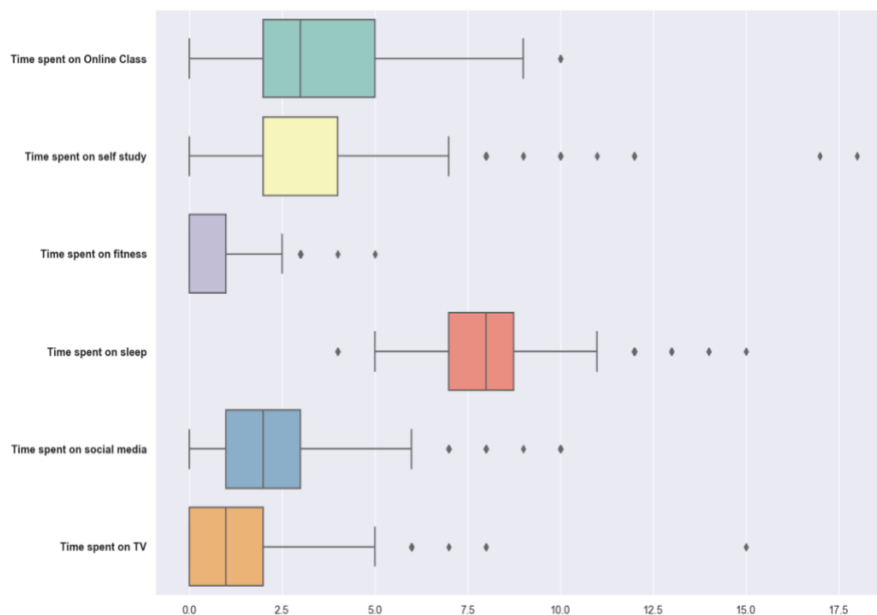
- What is the most preferred device used for online learning?



From the pie chart we observe that the data set is broken into five categories: Smartphone, Laptop/Desktop, Tablet, Any gadget and smartphone or Laptop/Desktop. The bulk of devices used for online learning seems to be smartphone and laptop/Desktop with a similar approximate of 46%.

Graph 3- Distribution of the preferred device used for online learning.

- How did the students spend their time?

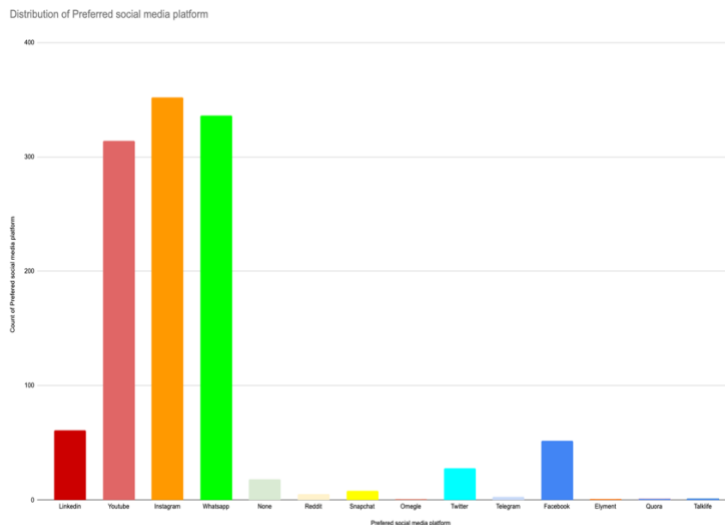


Graph 4 – Box and whisker plot showing the different ways students spend their time during Covid-19.

This graph shows six different ways in which students spend their time- on online classes, self-study, fitness, sleep, social media, and television. Seeing the graph, we understand on an average the students spend 2 to 5 hours daily with an average of approximately 3 hours in general. Students indulged in self-studies between 2 to 4 hours daily. Though we see quite a few outliers suggesting that there were students who have studied for as long as 8, 9, 10, 11, 12 and even 16 and 18 hours a day.

Moreover, there has been no initiative towards a fitness regime with very few and hardly any hours dedicated towards fitness. The average sleeping time is between 7 to 9 hours. Lastly the time spent online on social media activities is between 1 and 3 hours. There are some who have over indulged and spent 6 to 10 hours on social media.

- Where do they spend most of their time on social media?



This bar graph represents 13 different social media platforms that was popular during the covid-19 period. It seems that Instagram followed by WhatsApp and YouTube shows the majority in social media sites for entertainment and communicating with family and friends.

Graph 5 – Bar graph representing the preferred social media platform.

## 5. Bi-Variant Analysis (Correlation)

We now take the quantitative data from the data set and compare them to time spent online to see which has a highest correlation and try to develop an understanding.

- Time spent on online class in relation to Time spent on fitness.



Graph 6- Time spent on Online classes vs Fitness

This scatter plot is showing the correlation on how does the about of time spent online affect one's time on fitness. This scatter plot is clustered between 0 and 1 hour which denotes that the amount of time spent online does not affect their decision to spend time

on fitness. Though we see some of the students spend between 2 to 6 hours online they also spend 2 to 3 hours for their fitness training.

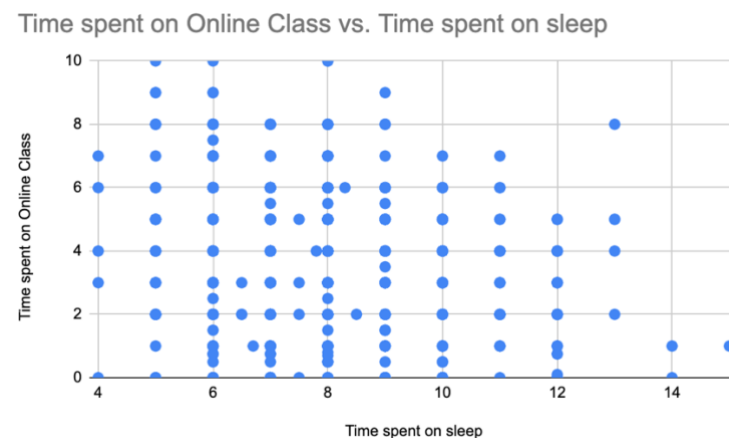
- Time spent on online classes in relation to time spent on self-study.



Graph 7- Time spent on Online classes vs Self-Study.

This scatter plot is showing the correlation on how does the about of time spent online affect one's time spent on self-study. There seems a linear relation here with most of the students who spend up to 2 hours online are investing between 1 to 4 hours on self-study. Many of the students who spend up to 8 hours online are investing 1 to 4 hours on self-study. Lasty. There are quite a few outliers where students are investing between 8 to 12 hours on self-study while they indulge in online class for 1 to 6 hours approximately.

- Time spent on online classes in relation to time spent on sleep.

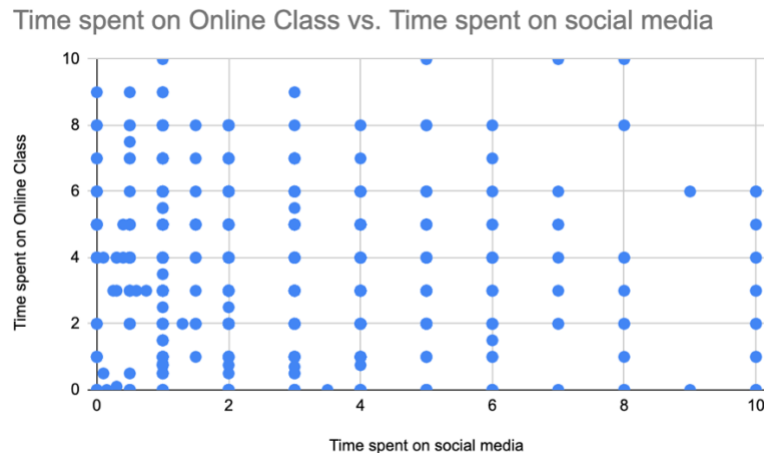


Graph 8- Time spent on Online classes vs Sleep.

This scatter plot is showing the correlation on how does the about of time spent online affect one's time spent sleeping the data here seems scattered all over place that is there does not seems to be a distinct pattern between the online class and sleeping patterns.

There are clusters seen for 1 to 2 hours of online class and 6 to 10 hours of sleep and 2 to 4 hours of online classes and 6 to 10 hours of sleep.

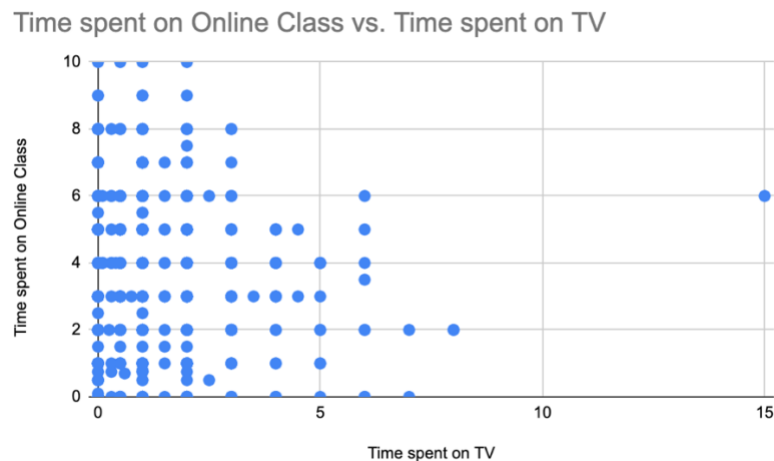
- Time spent on online classes in relation to time spent on social media.



Graph 9- Time spent on Online classes vs Social media.

This scatter plot is showing the correlation on how does the about of time spent online affect one's time spent on social media. As expected, there is a direct correlation between both these variables as both are online activity.

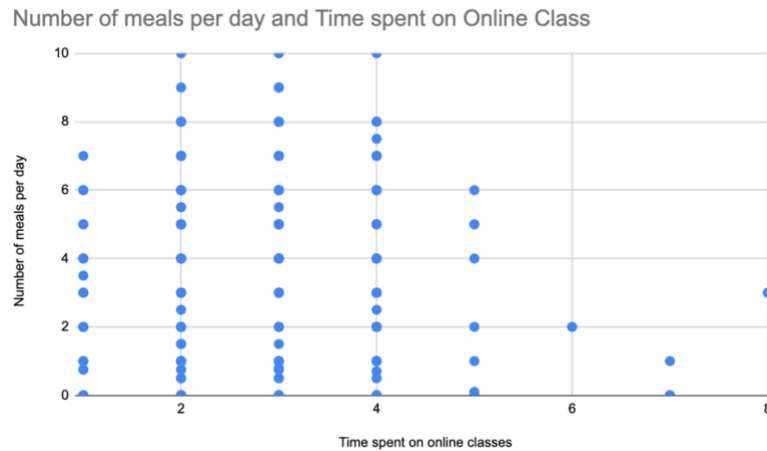
- Time spent on online classes in relation to time spent on television.



Graph 9- Time spent on Online classes vs TV.

- This scatter plot is showing the correlation on how does the about of time spent online affect one's time spent on television. For most of the number of hours spent, we witness a concentration of around 0 to 2 hours of TV viewing pattern. This indicates online class does not have much influence on TV viewing habits.

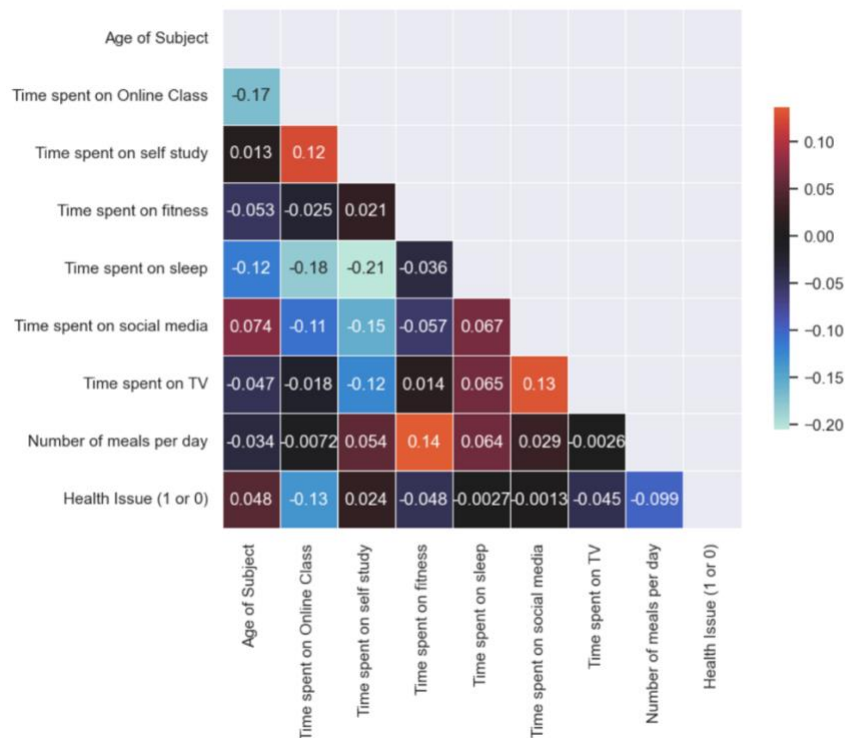
- Time spent on online classes in relation Number of meals per day.



Graph 10- Time spent on Online classes vs No. of meals per day

This scatter plot is showing the correlation on how does the about of time spent online affect one's time spent on television. Students attending online classes between 0 to 8 hours, are consuming around 4 meals per day. Additionally, few students between 0 to 6 hours were seen consuming 5 meals.

### Overall Correlation of all the variables



Graph 11 - Heat map for an overall correlation

Above is a heat graph; A heat graph is a type of data visualization that uses color to describe data. It displays data with each cell indicating a two-variable combination. Cells are colored according to the values they represent, with darker colors denoting greater values. ("What Is Heat Map (Heatmap)? - Definition from WhatIs.com")

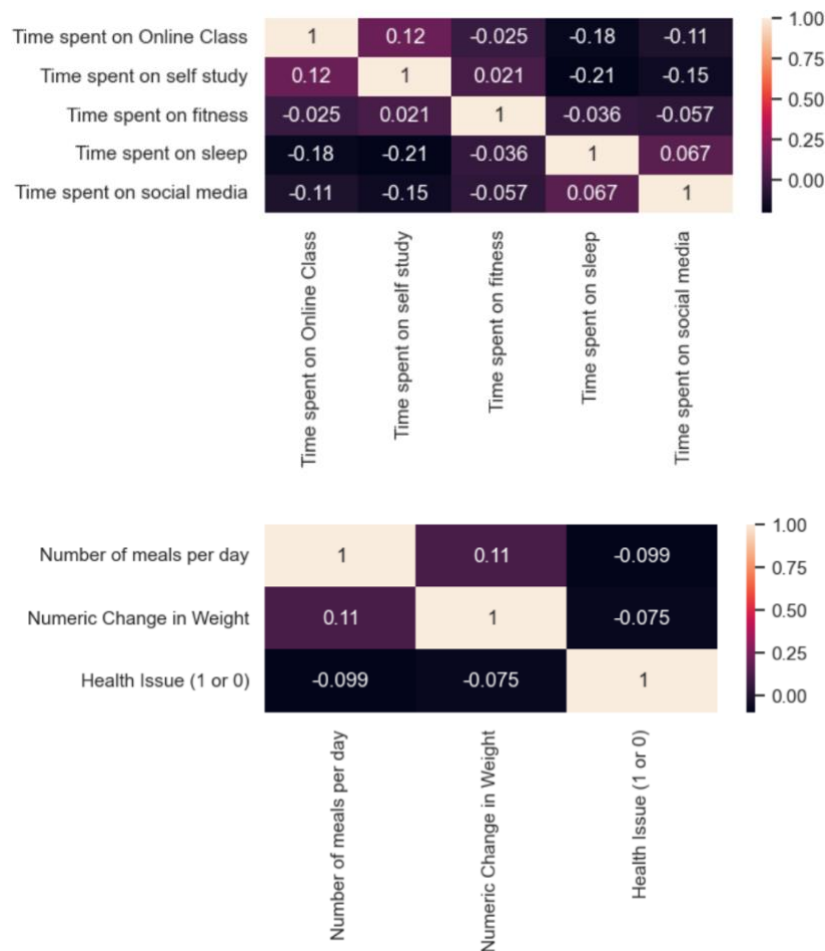
We do a correlation for the overall variable too to see as a whole which of the important factors causes high relation with each other. Knowing this concept, we find a strong correlation between self-study time and time spent on online classes (0.12), implying that the more time spent studying online positively affects time spent self-studying for greater academic accomplishments. We also detect positive correlation between the number of meals consumed each day and the amount of time spent exercising (0.14), meaning that the more one exercises, the greater their appetite. The last strong correlation was observed between time spent on television and time spent on social media (0.13), which basically implies that whether it's watching TV or using social media, a student's life is disrupted by the excessive use of technology, particularly in lockdown situations where technology and the internet are the only sources of entertainment.

We can further analyze some relatively positive correlations, implying these variables do to a certain extent may affect each other but we cannot generalize it to a wider population, these include: Time spent on social media and age of subject (0.074), Health issues and age of subject (0.048), Number of meals and time spent on self-studies (0.054), Time spent on social media and Time spent on sleep (0.067), Time spent on TV and time spent on sleep (0.065) and Number of meals and time spent on sleep (0.064). An example

Lastly, factors that have a negative correlation are Health issues and number of meals (-0.099), Health issue and time spent on online class (-0.13), time spent on sleep and age of subject (-0.12), time spent on social media and time spent on online class (-0.11), time spent on TV and time spent on self-study (-0.12). This is because in this data set these had no effect on each other, but it can be different for a wider population and another data set.



## Sub setting correlation



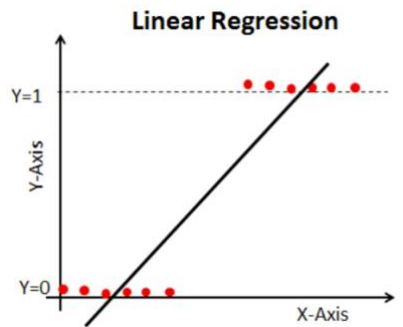
Graph 12 and 13 – Heat maps of subgroups

We proceed with a sub setting correlation as we can obtain more targeted and nuanced insights into the relationship between the two variables and identify any patterns or trends that may be specific to certain subgroups. From these two heat maps we see there is no strong correlation witnessed amongst the subset variables. We see relatively positive correlations between time spent self-studying and time spent on online class, time spent on sleep and time spent on social media and Number of meals per day and numeric change in weight.

## 6. Linear Regression Model

Linear Regression is a statistical technique used to model the relationship between a dependent variable and one or more independent variables by fitting a linear equation to the observed data. It is a method of finding the best linear relationship between the variables. We will be doing two different regression model summary one for the time spent online and one for the time spent on social media as we want to see the impact of online

classes and online activities on students. The goal of linear regression is to find a straight line that best fits the data.



General graph of a linear regression

### 6.1. Regression Model summary for Time spent online

OLS RESULT	
Dependent variable	Time spent on social media
Model	OLS
Method	Least squares
No.of observations	1131
Df residuals	1125
Df rmodel	5
Co variance type	Non Robust
R- squared	1
Adj R-squared	1
F statisitc	6.90E+31
Prob F statistic	0
Log Likelihood	36156
AIC	-7.23E+04
BIC	-7.23E+04

	coef	standard deviation	t	P> t	[0.025	0.975]
const	5.33E-15	6.03E-16	8.84	0	4.15E-15	6.51E-15
Time spent on Online Class	-2.06E-16	4.66E-17	-4.416	0	-2.97E-16	-1.14E-16
Time spent on self study	-2.08E-16	4.59E-17	-4.538	0	-2.98E-16	-1.18E-16
Time spent on fitness	-9.16E-16	1.31E-16	-6.988	0	-1.17E-15	-6.59E-16
Time spent on sleep	2.36E-16	6.12E-17	3.856	0	1.16E-16	3.56E-16
Time spent on social media	1	5.48E-17	1.82E+16	0	1	1

Omnibus	324.279
Prob(Omnibus)	0
Skew	1.435
Kurtosis	6.493
Durbin watson	1.122
Jarque-Bera (JB)	963.033
Prob(JB):	7.59E-210
Cond. No.	61.2

This observation presents the results of a multiple regression analysis with time spent on social media as the dependent variable and four other independent variables: time spent on online class, time spent on self-study, time spent on fitness, and time spent on sleep. The method used to create this model is Ordinary Least Squares (OLS), which is a standard method for estimating the parameters of a linear regression model.

This data represents the result of a linear regression analysis, where the dependent variable is "Time spent on social media". The key statistics are:

- R-squared: This statistic measures the proportion of variance in the dependent variable that is explained by the independent variables in the model. The R-squared value of 1.000 indicates that the model explains all the variance in the dependent variable.
- Adj. R-squared: This is the adjusted R-squared, which considers the number of independent variables in the model. In this case, the adjusted R-squared is 1.00, which is similar to R-squared.
- F-statistic: This statistic measures the overall significance of the model, by comparing the variance explained by the model to the variance not explained by the model. The F-statistic of 6.896e+31 and its associated p-value of 0.00 suggest that the model as a whole is statistically significant. - The F-statistic is a measure of the overall significance of the model, and it has a very high value of 6.896e+31. This value is so high that it suggests that the model is almost certainly significant and has explanatory power. The R-squared value of 1.000 indicates that the model explains all the variance in the dependent variable, and the F-statistic of 6.896e+31 and its associated p-value of 0.00 suggest that the model as a whole is statistically significant.
- Prob (F-statistic): This is the p-value associated with the F-statistic, which measures the probability of obtaining an F-statistic as extreme as the one observed, if the null hypothesis (that all coefficients are zero) were true.
- Log-Likelihood: This is a measure of how well the model fits the data, based on the likelihood of the observed data under the model. A lower log-likelihood indicates a better fit. There are 1131 observations in the dataset, and the log-likelihood is 36156.

- No. Observations: This is the number of observations used in the analysis, which is 1131.
- AIC and BIC: These are information criteria that measure the quality of the model, based on how well it fits the data and how many parameters it has. Lower values indicate a better model. The AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) are both negative, which is normal for these measures.
- Df Residuals: This is the degrees of freedom of the residuals, which is the difference between the number of observations and the number of parameters in the model. In this case, the Df Residuals is 1125.
- Df Model: This is the degrees of freedom of the model, which is the number of parameters estimated in the model. In this case, the Df Model is 5.
- Covariance Type: This indicates the type of covariance matrix used in the analysis. In this case, The covariance type is non robust, meaning that the model assumes that the errors are normally distributed with constant variance. The coefficients for each independent variable represent the effect of that variable on time spent on social media while holding other variables constant.
- ("Akaike Information Criterion - an Overview | ScienceDirect Topics"), (CFI Team), ("F-Statistic and T-Statistic - MATLAB & Simulink"), (kenton), ("Omnibus Test"), (Stephanie), (Wikipedia Contributors, "Jarque-Bera Test"), (Wikipedia Contributors, "Ordinary Least Squares")

Notably, the coefficient for time spent on social media is 1.000, indicating that a unit increase in time spent on social media leads to a unit increase in time spent on social media, which is not surprising. The coefficients for the other variables are all small in magnitude and are either negative or positive, suggesting that they have weak or negligible effects on time spent on social media. However, it's important to note that some of these variables (e.g., time spent on online class and time spent on self-study) have statistically significant coefficients with p-values of 0.000, indicating that they do have some effect on time spent on social media.

The Omnibus test, Jarque-Bera test, and Durbin-Watson test are statistical tests that help check assumptions of the regression model, and in this case: the Omnibus and Jarque-Bera tests show that the residuals are not normally distributed and the Durbin-Watson test indicates that there may be some autocorrelation present. The coefficient for the constant is also statistically significant with a p-value of 0.000.

Overall, the results suggest that while time spent on social media is strongly influenced by itself, it may also be influenced by other factors such as **time spent on online classes and self-study**.

This observation presents the results of a multiple regression analysis **with time spent on social media as the dependent variable and four other independent variables: time spent on online class, time spent on self-study, time spent on fitness, and time spent on sleep**.

## 6.2. Regression Model summary for Time spent on social media.

OLS RESULT	
Dependent variable	Time spent on social media
Model	OLS
Method	Least squares
No.of observations	1131
Df residuals	1125
Df rmodel	5
Co variance type	Non Robust
R- squared	1
Adj R-squared	1
F statisitc	6.90E+31
Prob F statistic	0
Log Likelihood	36156
AIC	-7.23E+04
BIC	-7.23E+04

	coef	standard deviation	t	P> t	[0.025	0.975]
const	5.33E-15	6.03E-16	8.84	0	4.15E-15	6.51E-15
Time spent on Online Class	-2.06E-16	4.66E-17	-4.416	0	-2.97E-16	-1.14E-16
Time spent on self study	-2.08E-16	4.59E-17	-4.538	0	-2.98E-16	-1.18E-16
Time spent on fitness	-9.16E-16	1.31E-16	-6.988	0	-1.17E-15	-6.59E-16
Time spent on sleep	2.36E-16	6.12E-17	3.856	0	1.16E-16	3.56E-16
Time spent on social media	1	5.48E-17	1.82E+16	0	1	1

Omnibus	324.279
Prob(Omnibus)	0
Skew	1.435
Kurtosis	6.493
Durbin watson	1.122
Jarque-Bera (JB)	963.033
Prob(JB):	7.59E-210
Cond. No.	61.2

This observation presents the results of a multiple regression analysis with **time spent on social media as the dependent variable and four other independent variables: time spent on online class, time spent on self-study, time spent on fitness, and time spent on sleep.**

The method used to create this model is Ordinary Least Squares (OLS) as we did similarly in the previous regression model. This data represents the result of a linear regression analysis, where the dependent variable is "Time spent on social media". The key statistics are:

- R-squared: This statistic measures the proportion of variance in the dependent variable that is explained by the independent variables in the model. The R-squared value of 1.000 indicates that the model explains all the variance in the dependent variable, and the F-statistic of 6.896e+31 and its associated p-value of 0.00 suggest that the model as a whole is statistically significant.
- Adj. R-squared: This is the adjusted R-squared, which takes into account the number of independent variables in the model. In this case, the adjusted R-squared is 0.041, which is slightly lower than the R-squared.
- F-statistic: This statistic measures the overall significance of the model, by comparing the variance explained by the model to the variance not explained by the model. The F-statistic is a measure of the overall significance of the model, and it has a very high value of 6.896e+31. This value is so high that it suggests that the model is almost certainly significant and has explanatory power.
- The R-squared value of 1.000 indicates that the model explains all the variance in the dependent variable, and the F-statistic of 6.896e+31 and its associated p-value of 0.00 suggest that the model as a whole is statistically significant.
- Prob (F-statistic): This is the p-value associated with the F-statistic, which measures the probability of obtaining an F-statistic as extreme as the one observed, if the null hypothesis (that all coefficients are zero) were true.
- Log-Likelihood: This is a measure of how well the model fits the data, based on the likelihood of the observed data under the model. A lower log-likelihood indicates a better fit. There are 1131 observations in the dataset, and the log-likelihood is 36156.
- No. Observations: This is the number of observations used in the analysis, which is 1131.
- AIC and BIC: These are information criteria that measure the quality of the model, based on how well it fits the data and how many parameters it has. Lower values indicate a better model. The AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) are both negative, which is normal for these measures.
- Df Residuals: This is the degrees of freedom of the residuals, which is the difference between the number of observations and the number of parameters in the model. In this case, the Df Residuals is 1125.
- Df Model: This is the degrees of freedom of the model, which is the number of parameters estimated in the model. In this case, the Df Model is 5.

- Covariance Type: This indicates the type of covariance matrix used in the analysis. In this case, The covariance type is non robust, meaning that the model assumes that the errors are normally distributed with constant variance.

The coefficients for each independent variable represent the effect of that variable on time spent on social media while holding other variables constant. Notably, the coefficient for time spent on social media is 1.000, indicating that a unit increase in time spent on social media leads to a unit increase in time spent on social media, which is not surprising. The coefficients for the other variables are all small in magnitude and are either negative or positive, suggesting that they have weak or negligible effects on time spent on social media. However, it's important to note that some of these variables (e.g., time spent on online class and time spent on self-study) have statistically significant coefficients with p-values of 0.000, indicating that they do have some effect on time spent on social media.

The Omnibus test, Jarque-Bera test, and Durbin-Watson test are statistical tests that help check assumptions of the regression model, and in this case: the Omnibus and Jarque-Bera tests show that the residuals are not normally distributed and the Durbin-Watson test indicates that there may be some autocorrelation present. The coefficient for the constant is also statistically significant with a p-value of 0.000.

## **Conclusion**

Having explored the dataset and analysed it with various techniques, some important pointer with regarding to this sample dataset and its variables can be concluded as follows:

- Time spent on social media may be influenced by factors such as time spent on online classes and self-study.
- Observation on the activities of age group 16-25 which forms most of the observations shows an average of 1 to 2 hours and 2 to 4 hours of online classes, which follows a pattern of 7 to 9 hours of sleep which seems normal. There is no significant variance in sleep patterns observed between infected and healthy students.
- There is a pattern observed between online classes and self-study. Most students who spend up to 8 hours on online classes are investing 1 to 4 hours in self-studies.
- There is no influence on following a fitness regime and the activity with fitness is as low as 0 to 1 hour in most observations.
- The normal meal pattern seems 4 times a day for students attending up to 8 hours of online class.
- Not major pattern of weight loss is observed due to the infection.
- Infected students also are attending online class and indulging in self-studies, even though the time may be lesser.

The Linear Regression Analysis tells us that the data has a high statistical significance which means that the relationship between the independent variable(s) and the dependent variable is unlikely to have occurred by chance.

It also suggests that our dataset has the ability of the model to accurately predict the dependent variable based on the independent variables, hence can predict the dependent variable accurately. This also suggest that the independent variable(s) are strongly related to the dependent variable and the model makes accurate predictions.

This particular model fits this data and helps us to come out with some dependable inferences with regards to the effect of lock down and its effect on student's activities.

## Appendix

ID	Region of residence	Age of Subject	Time spent on Online Class	Medium for online class	Time spent on self study	Time spent on fitness	Time spent on sleep	Time spent on social media	Preferred social media platform	Time spent on TV	Number of meals per day	Health issue during lockdown
R1	Delhi-NCR	21		2 Laptop/Desktop		4	0	7	3 LinkedIn		1	4 NO
R2	Delhi-NCR	21		0 Smartphone		0	2	10	3 Youtube		0	3 NO
R3	Delhi-NCR	20		7 Laptop/Desktop		5	0	6	2 LinkedIn		0	3 NO
R4	Delhi-NCR	20		3 Smartphone		2	1	6	5 Instagram		0	3 NO
R5	Delhi-NCR	21		3 Laptop/Desktop		3	1	6	3 Instagram		1	4 NO
R6	Delhi-NCR	21		0 Smartphone		6	0	5	1 Youtube		0	1 YES
R7	Delhi-NCR	19		2 Smartphone		2	1	5	4 Instagram		0	3 NO
R8	Outside Delhi-NCR	19		2 Tablet		1	1	10	5 Instagram		0	3 YES
R9	Delhi-NCR	21		3 Laptop/Desktop		4	1	8	2 Whatsapp		1	3 NO
R10	Outside Delhi-NCR	20		0 Laptop/Desktop		1	0.5	8	5 Instagram		3	3 YES
R11	Delhi-NCR	21		3 Laptop/Desktop		3	1	8	3 Instagram		1	4 NO
R12	Delhi-NCR	21		1 Laptop/Desktop		0	1	7	3 Instagram		1	2 YES
R13	Delhi-NCR	21		3 Laptop/Desktop		0	0	6	3 Instagram		0	3 YES
R14	Outside Delhi-NCR	22		1 Laptop/Desktop		2	0	7	0 None		0.5	3 NO
R15	Delhi-NCR	20		5 Laptop/Desktop		1	0	8	3 Instagram		0	3 NO
R16	Delhi-NCR	22		3 Smartphone		3	1	6	2 Instagram		1	3 NO
R17	Outside Delhi-NCR	20		0 Smartphone		0	0	8	2 Reddit		1	3 NO
R18	Delhi-NCR	20		1 Smartphone		2	1	6	1 Youtube		0	3 NO

## Section of the data set

Dep. Variable:	Time spent on social media		R-squared:	1.000			
Model:	OLS		Adj. R-squared:	1.000			
Method:	Least Squares		F-statistic:	6.896e+31			
Date:	Sun, 26 Feb 2023	Prob (F-statistic):	0.00				
Time:	18:27:26	Log-Likelihood:	36156.				
No. Observations:	1131		AIC:	-7.230e+04			
Df Residuals:	1125		BIC:	-7.227e+04			
Df Model:	5						
Covariance Type:	nonrobust						
	coef	std err	t	P> t	[0.025	0.975]	
	const	5.329e-15	6.03e-16	8.840	0.000	4.15e-15	6.51e-15
Time spent on Online Class	-2.056e-16	4.66e-17	-4.416	0.000	-2.97e-16	-1.14e-16	
Time spent on self study	-2.082e-16	4.59e-17	-4.538	0.000	-2.98e-16	-1.18e-16	
Time spent on fitness	-9.159e-16	1.31e-16	-6.988	0.000	-1.17e-15	-6.59e-16	
Time spent on sleep	2.359e-16	6.12e-17	3.856	0.000	1.16e-16	3.56e-16	
Time spent on social media	1.0000	5.48e-17	1.82e+16	0.000	1.000	1.000	
Omnibus:	324.279	Durbin-Watson:	1.122				
Prob(Omnibus):	0.000	Jarque-Bera (JB):	963.033				
Skew:	1.435	Prob(JB):	7.59e-210				
Kurtosis:	6.493	Cond. No.	61.2				

### OLS Regression Results

Dep. Variable:	Time spent on social media	R-squared:	1.000				
Model:	OLS	Adj. R-squared:	1.000				
Method:	Least Squares	F-statistic:	6.896e+31				
Date:	Sun, 26 Feb 2023	Prob (F-statistic):	0.00				
Time:	18:13:49	Log-Likelihood:	36156.				
No. Observations:	1131	AIC:	-7.230e+04				
Df Residuals:	1125	BIC:	-7.227e+04				
Df Model:	5						
Covariance Type:	nonrobust						
	coef	std err	t	P> t	[0.025	0.975]	
	const	5.329e-15	6.03e-16	8.840	0.000	4.15e-15	6.51e-15
Time spent on Online Class	-2.056e-16	4.66e-17	-4.416	0.000	-2.97e-16	-1.14e-16	
Time spent on self study	-2.082e-16	4.59e-17	-4.538	0.000	-2.98e-16	-1.18e-16	
Time spent on fitness	-9.159e-16	1.31e-16	-6.988	0.000	-1.17e-15	-6.59e-16	
Time spent on sleep	2.359e-16	6.12e-17	3.856	0.000	1.16e-16	3.56e-16	
Time spent on social media	1.0000	5.48e-17	1.82e+16	0.000	1.000	1.000	
Omnibus:	324.279	Durbin-Watson:	1.122				
Prob(Omnibus):	0.000	Jarque-Bera (JB):	963.033				
Skew:	1.435	Prob(JB):	7.59e-210				
Kurtosis:	6.493	Cond. No.	61.2				

## Image of the stat results results



## **Citations**

1. “Akaike Information Criterion - an Overview | ScienceDirect Topics.”  
*W*[www.sciencedirect.com](http://www.sciencedirect.com), [www.sciencedirect.com/topics/pharmacology-toxicology-and-pharmaceutical-science/akaike-information-criterion#:~:text=The%20Akaike%20information%20criterion%20\(AIC\)%20and%20the%20Bayesian%20information%20criterion](http://www.sciencedirect.com/topics/pharmacology-toxicology-and-pharmaceutical-science/akaike-information-criterion#:~:text=The%20Akaike%20information%20criterion%20(AIC)%20and%20the%20Bayesian%20information%20criterion).
2. CFI Team. “R-Squared.” *Corporate Finance Institute*, 4 Nov. 2022,  
[corporatefinanceinstitute.com/resources/data-science/r-squared/](http://corporatefinanceinstitute.com/resources/data-science/r-squared/).
3. “F-Statistic and T-Statistic - MATLAB & Simulink.” *W*[www.mathworks.com](http://www.mathworks.com),  
[www.mathworks.com/help/stats/f-statistic-and-t-statistic.html](http://www.mathworks.com/help/stats/f-statistic-and-t-statistic.html).
4. kenton, will. “Understanding the Durbin Watson Statistic.” *Investopedia*, 2019,  
[www.investopedia.com/terms/d/durbin-watson-statistic.asp](http://www.investopedia.com/terms/d/durbin-watson-statistic.asp).
5. “Omnibus Test.” *Wikipedia*, 15 June 2020, [en.wikipedia.org/wiki/Omnibus\\_test](https://en.wikipedia.org/wiki/Omnibus_test).
6. Pirdehghan, Azar, et al. “Social Media Use and Sleep Disturbance among Adolescents: A Cross-Sectional Study.” *Iranian Journal of Psychiatry*, vol. 16, no. 2, 17 Mar. 2021,  
[www.ncbi.nlm.nih.gov/pmc/articles/PMC8233562/](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC8233562/),  
<https://doi.org/10.18502/ijps.v16i2.5814>.
7. Stephanie. “Log Likelihood Function.” *Statistics How To*, 20 Apr. 2021,  
[www.statisticshowto.com/log-likelihood-function/](http://www.statisticshowto.com/log-likelihood-function/).
8. “VisibleBreadcrumbs.” *Mathworks.com*, 2019, [www.mathworks.com/help/stats/what-is-linear-regression.html](http://www.mathworks.com/help/stats/what-is-linear-regression.html).

9. “What Is Heat Map (Heatmap)? - Definition from WhatIs.com.”

*SearchBusinessAnalytics*, [www.techtarget.com/searchbusinessanalytics/definition/heat-map](http://www.techtarget.com/searchbusinessanalytics/definition/heat-map).

10. Wikipedia Contributors. “Jarque–Bera Test.” *Wikipedia*, Wikimedia Foundation, 23 Aug. 2019, [en.wikipedia.org/wiki/Jarque%E2%80%93Bera\\_test](https://en.wikipedia.org/wiki/Jarque%E2%80%93Bera_test).

11. “Ordinary Least Squares.” *Wikipedia*, Wikimedia Foundation, 28 Nov. 2019, [en.wikipedia.org/wiki/Ordinary\\_least\\_squares](https://en.wikipedia.org/wiki/Ordinary_least_squares).