PROJECT REPORT ON THE TOPIC

"Online VS Offline education:- understanding and analyzing students's preferences"

By-

VIKASH KUMAR

CONTENTS

INTRODUCTION

The world of education has been significantly transformed by the COVID-19 pandemic, which forced schools and universities to rapidly switch from traditional classrooms to online platforms. This shift has not only changed where and how students learn but has also raised important questions about the effectiveness and future of online education. My project explores these changes, examining how students have adapted to online learning environments and their preferences for future educational modes.

This study is particularly relevant now as we navigate the potential lasting impacts of the pandemic on education. It will explore several key areas: how accessible and inclusive online learning is, how satisfied students are with their online learning experiences, and whether they would prefer the traditional classroom mode or the online learning platforms.

By investigating these aspects, my project titled ***"Online vs Traditional: Analyzing Future Educational Preferences "*** aims to provide insights that can help educators, policymakers, and platform developers improve online educational practices and make informed decisions about the future of learning. This study will help ensure that the educational strategies implemented post-pandemic are effective, equitable, and responsive to the needs of all students.

METHODOLOGY OF SURVEY

In any type of survey, preplanning and systematic approach play an important role in arriving at the best possible results and successful completion in minimum time and cost. Also, the success depends upon resources, quality, timing, and integrity of the surveyor who compiles the primary data. So, it is a very important task to manage all the available resources which make an impact on the quality of the survey.

(1) PLANNING OF THE SURVEY:

Full-proof planning is an essential part of any statistical survey to be completed successfully at an optimized cost, labor, and time. Planning of the survey includes the selection of topics and preparation of a short questionnaire covering all the required areas.

In this survey, it was ensured that the questionnaire was neither too long nor too short, so the respondent generally answered all the questions, and all the important topics were covered.

(2) OBJECTIVE OF THE SURVEY:-

A complete and well-defined objective should be mentioned at the very beginning stage. The objectives for this project are

- To understand the current usage patterns of online learning platforms among students.
- To identify the most popular online learning platforms among students and the reasons for their popularity.
- To analyze the time spent by students in traditional classrooms versus online learning platforms.
- To assess the effectiveness of online learning platforms in meeting the educational needs of students.

- To evaluate student satisfaction with the education provided by online learning platforms.
- To determine the factors that students consider most important when choosing online learning platforms.
- To investigate the association between various demographic factors (such as sex, residential area, field of study, income bracket, and level of education) and the preference for online or offline modes of education.
- To predict future trends in educational preferences among students based on current data.

The hypotheses to be tested in this study are as follows:-

- There is no association between the sex of the students and their preferred mode of education (online or offline).
- There is no association between the permanent residence of students (rural and urban) and their preferred mode of education.
- There is no association between the faculties in which students are currently enrolled and their preferred mode of education.
- There is no association between the family income of the students and their preferred mode of education.
- There is no association between the present class of students and their preferred mode of education.
- On average, students dedicate equal time to both types of education, online mode, and traditional classroom mode.
- The proportion of students which prefers online mode is the same as the ones who prefer traditional classroom mode.

(3) AREA OF SURVEY:

For performing any survey, a sample is selected from a population. I decided to collect my sample from the students of the various faculties from the different collages. I surveyed a total of 160 respondents.

(4) SAMPLING TECHNIQUES:

The sampling technique adopted for the study is a random sampling technique according to the convenience of the projector. A questionnaire was administered to different students to obtain data for analysis through Google Forms.

(5) SAMPLE SIZE:

The population being large the survey was carried out among 160 respondents. These samples can be considered adequate to represent the characteristics of the entire population.

(6) SAMPLE DESIGN:

Data has been presented with the help of a bar diagram, pie chart, etc. I have also used cross-tabulation and the **CHI-SQUARE test** to check the association between some attributes, **Z-TEST FOR THE SINGLE PROPORTION** and **T-TEST FOR THE TWO INDEPENDENT MEANS**.

(7) ANALYSIS AND REPORTING:

After the collection of raw data, we put them in coded form so that the analysis becomes easy. For this purpose, I used Python pandas software, and I also used MS Excel and MS Word to give a perfect frame to my whole analysis.

Pictorial representation of data:

Any statistical data can be represented by using a chart or diagram. A chart is a graphical representation of data, in which “the data is represented by symbols,

such as bars in a bar chart, slices in a pie chart ". Charts are often used in the understanding of large quantities of data and the relationships between parts of the data. Charts can usually be read more quickly than the raw data that they are produced from. We can represent raw data in

1. Bar chart or bar diagram

2. Pie chart

(1). Bar chart or Bar diagram:

A bar chart or bar graph is a chart with rectangular bars whose length is proportional to the values that they represent. The bars can be plotted vertically or horizontally. A vertical bar chart is sometimes called a column bar chart. A bar graph is a chart that uses either horizontal or vertical bars to show comparisons among categories. Some bar graphs present bars clustered groups of more than one (grouped bar graphs), and others show the bars divided into sub-parts to show cumulative effect (stacked bar graphs).

(2) Pie chart:

A pie chart is a circular chart divided into sectors illustrating numerical proportions. In a pie chart, the length of each sector (and consequently its central angle and area), is proportional to the quantity it represents, while it is named for its resemblance to a pie that has been sliced, there are variations in the way it can be presented. Pie charts are very widely used in the business world and the media. However, they have been criticized, and many experts recommend avoiding them, pointing out that project has shown it is difficult to compare different sections of a given pie chart or to compare data across different pie charts. Pie charts can be replaced in most cases by plots such as the bar chart.

CHI-SQUARE TEST FOR INDEPENDENCE OF ATTRIBUTES

The Chi-square test for independence is a statistical method used to determine if there is a significant association between two categorical variables. This test is commonly applied in projects to analyze data from contingency tables, and it helps determine whether the distributions of one variable differ across the levels of another variable.

Steps to Perform a Chi-square Test for Independence

1. Data Organization:

Arrange the data in a contingency table format where each cell represents the frequency count of cases for a specific combination of the categories of the two variables.

2. Assumptions:

The data in the cells should be frequencies or counts of cases.

A sufficient sample size is required. Typically, no more than 20% of the cells should have expected frequencies less than 5, and all individual expected frequencies should be 1 or greater.

3. Hypotheses:

***Null hypothesis (H_0)*:** There is no association between the two variables (they are independent).

Alternative hypothesis (H_1): There is an association between the two variables (they are not independent).

4. Calculate the Expected Frequencies:

For each cell in the table, calculate the expected frequency, which is given by:

$$Eij = \frac{\{(Row\ Total_i \times Column\ Total_j)\}}{Grand\ Total}$$

Where Eij is the expected frequency for cell i, j, Row $Total_i$ is the total of row i, Column $Total_j$ is the total of column j, and Grand Total is the total number of observations.

5. Chi-square Statistic:

Compute the chi-square statistic using the formula:

$$\chi 2 = \Sigma \frac{(Oij - Eij)2}{Eij}$$

Where O_{ij} is the observed frequency in cell i, j and E_{ij} is the expected frequency. The summation extends over all cells.

6. Determine the Degrees of Freedom:

Degrees of freedom for this test are calculated as:

(Number of Rows - 1) × (Number of Columns - 1)

7. Critical Value and P-value:

Compare the calculated chi-square statistic to a critical value from the chi-square distribution table at the desired significance level (commonly 0.05).

Alternatively, compute the p-value and compare it to the significance level to decide whether to reject the null hypothesis.

8. Conclusion:

If the chi-square statistic is greater than the critical value or if the p-value is less than the significance level, reject the null hypothesis, suggesting a significant association between the variables.

Otherwise, do not reject the null hypothesis, suggesting no significant association.

This test provides a way to objectively analyze the relationship between two categorical variables based on sample data.

Z-TEST FOR A SINGLE PROPORTION

The z-test for a single proportion is a statistical method used to compare an observed proportion to a theoretical one when the sample size is large. It is often used to determine whether the proportion of a certain characteristic in a sample is significantly different from the expected proportion.

Steps to Perform a Z-Test for a Single Proportion

1. Data Collection:
Gather a random sample from the population, and note the number of successes (i.e., occurrences of the characteristic of interest).

2. Assumptions:
The sample size should be large enough for the Central Limit Theorem to apply (typically $n \times p \geq 5$ and $n \times (1-p) \geq 5$ where n is the sample size and p is the proportion of successes).

The sample is randomly selected and independent.

3. Hypotheses:
Null hypothesis (H_0): The observed proportion p equals the hypothesized proportion p_0 (e.g., $p = p_0$).
Alternative hypothesis (H_1): The observed proportion p does not equal the hypothesized proportion p_0

(e.g., $p \neq p_0$).

4. Test Statistic:

Calculate the z-statistic using the formula:

$$z = \frac{p - p0}{\sqrt{\frac{p0(1-p0)}{n}}}$$

Where p is the sample proportion, p_0 is the null hypothesis proportion, and n is the sample size.

5. Decision Rule:

Determine the critical z-value from the standard normal distribution corresponding to the desired level of significance (commonly $\alpha = 0.05$).

For a two-tailed test, use $\pm z_{\alpha/2}$;

For a one-tailed test, use z_α or $-z_\alpha$ depending on the direction of the alternative hypothesis.

6. Conclusion:

If the absolute value of the z-statistic is greater than the critical z-value, reject the null hypothesis.

Otherwise, do not reject the null hypothesis.

7. P-Value:

Calculate the p-value based on the z-statistic from the standard normal distribution to provide the probability of observing such an extreme test statistic under the null hypothesis.

T-TEST for testing the equality of two means

The t-test for two independent means is a statistical method used to compare the means of two independent groups. It is often used to determine whether there is a significant difference between the means of two groups.

Steps to Perform a T-Test for Two Independent Means

1. **Data Collection:** Gather random samples from the two populations, and note the means and standard deviations of each sample.
2. **Assumptions:**
    - The samples are randomly selected and independent.
    - The populations from which the samples are drawn are normally distributed.
    - The populations have the same variance (this assumption can be relaxed with the use of a Welch's t-test).
3. **Hypotheses:**
    - **Null hypothesis (H0):** The means of the two populations are equal (e.g., μ1 = μ2).
    - **Alternative hypothesis (H1):** The means of the two populations are not equal (e.g., μ1 ≠ μ2).
4. **Test Statistic:** Calculate the t-statistic using the formula:

$$t = \frac{\overline{x_1} - \overline{x_2} - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Where $\overline{x_1}$ and $\overline{x_2}$ are the sample means, μ_1 and μ_2 are the population means (usually 0 under the null hypothesis), s_1^2 and s_2^2 are the sample variances, and n_1 and n_2 are the sample sizes.

5. **Decision Rule:** Determine the critical t-value from the t-distribution corresponding to the desired level of significance (commonly α = 0.05).
    - For a two-tailed test, use ±tα/2;
    - For a one-tailed test, use tα or -tα depending on the direction of the alternative hypothesis.

6. **Conclusion:** If the absolute value of the t-statistic is greater than the critical t-value, reject the null hypothesis. Otherwise, do not reject the null hypothesis.
7. **P-Value:** Calculate the p-value based on the t-statistic from the t-distribution to provide the probability of observing such an extreme test statistic under the null hypothesis.

DATA ANALYSIS

&

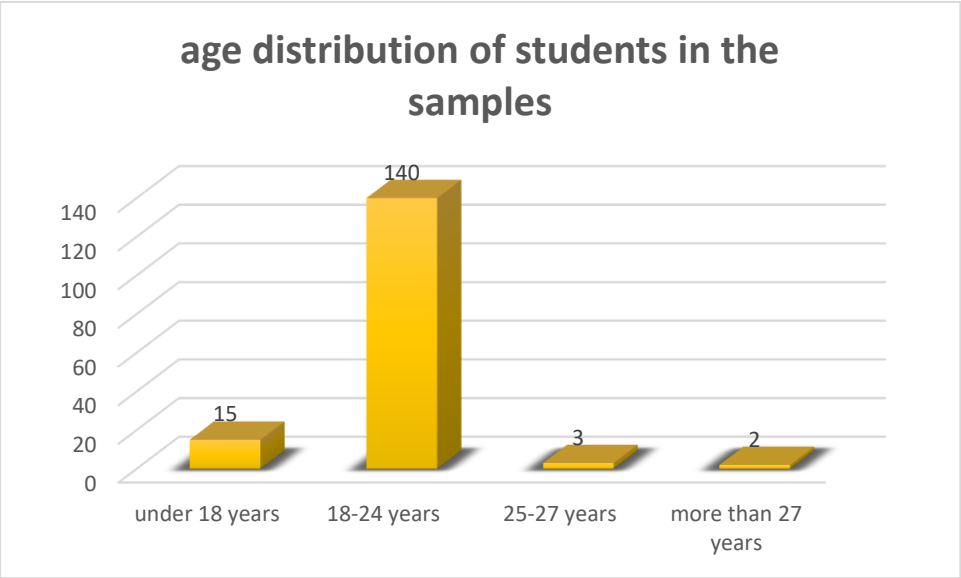
INTERPRETATION

>TABLES

>GRAPHS & CHART

Table-1
Age distribution of students

age of students	
under 18 years	15
18-24 years	140
25-27 years	3
more than 27 years	2



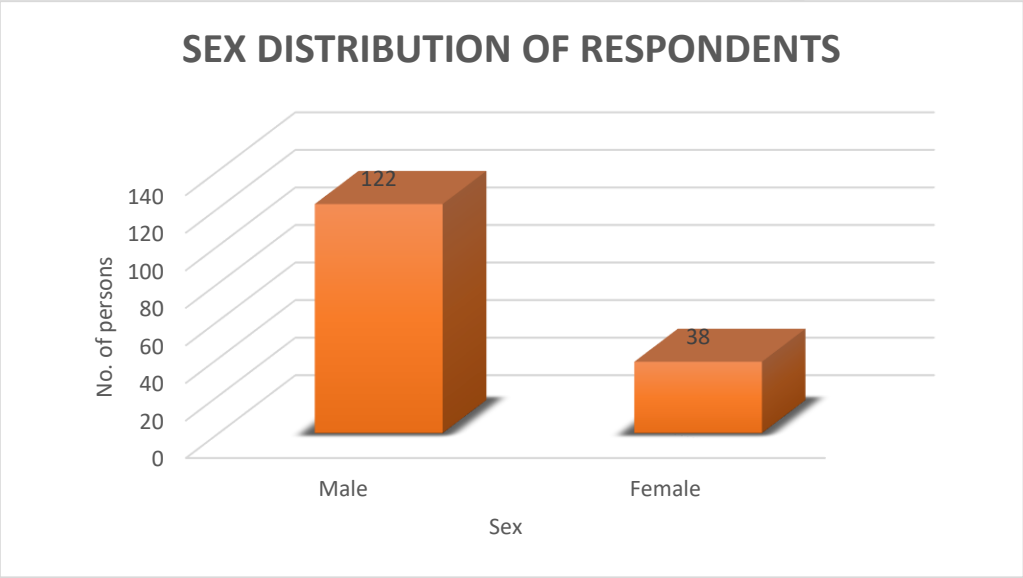
From the above table and bar graph, it is evident that a vast **majority** of respondents fall under the age category of 18-24 years.
This implies that very few people are outside this age group.

This is primarily because most of the students from graduation and Post-graduation are under this age category, thus we find their numbers exceptionally higher.

Sex distribution of students

Table-2

Sex of students	Number	percentage
Male	122	76.25
Female	38	23.75



From this table and graph, we find that **122** out of a total of **160** respondents are Male while the other **38** are female.

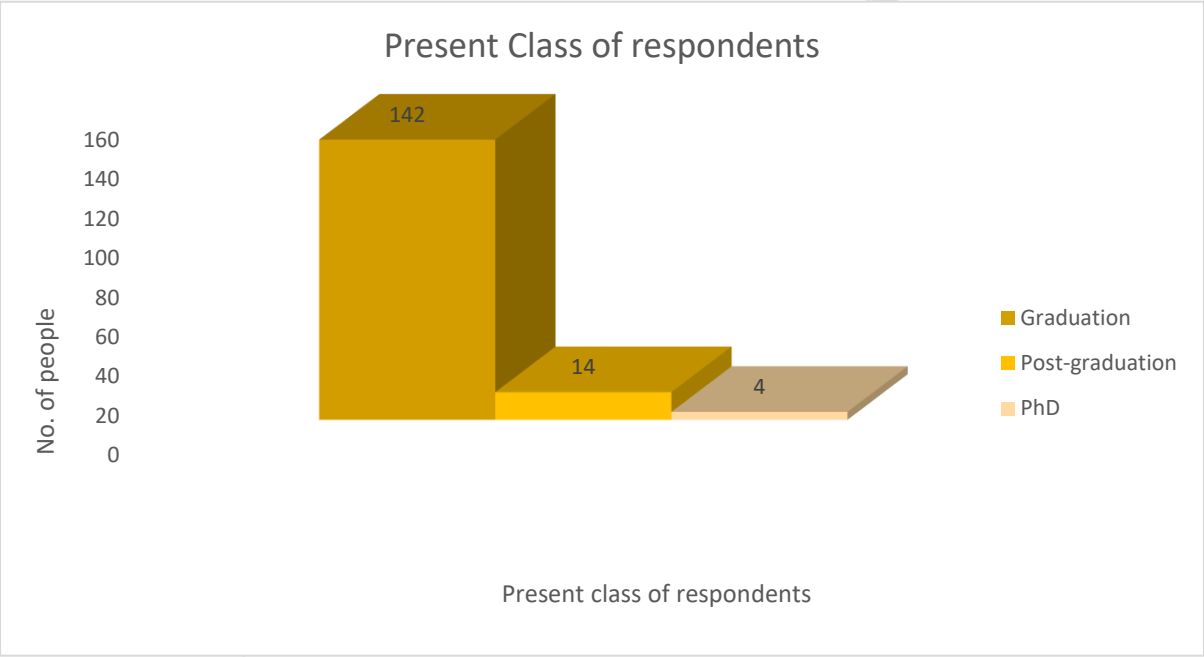
That implies around **76.25%** of students are male and **23.75%** are female.

This heterogeneity is bound to happen since the total student population follows a similar distribution (by general observation).

Present class of students

Table – 3

Current Class	No. of students	%
Graduation	142	88.75
Post-graduation	14	8.75
PhD	4	2.5



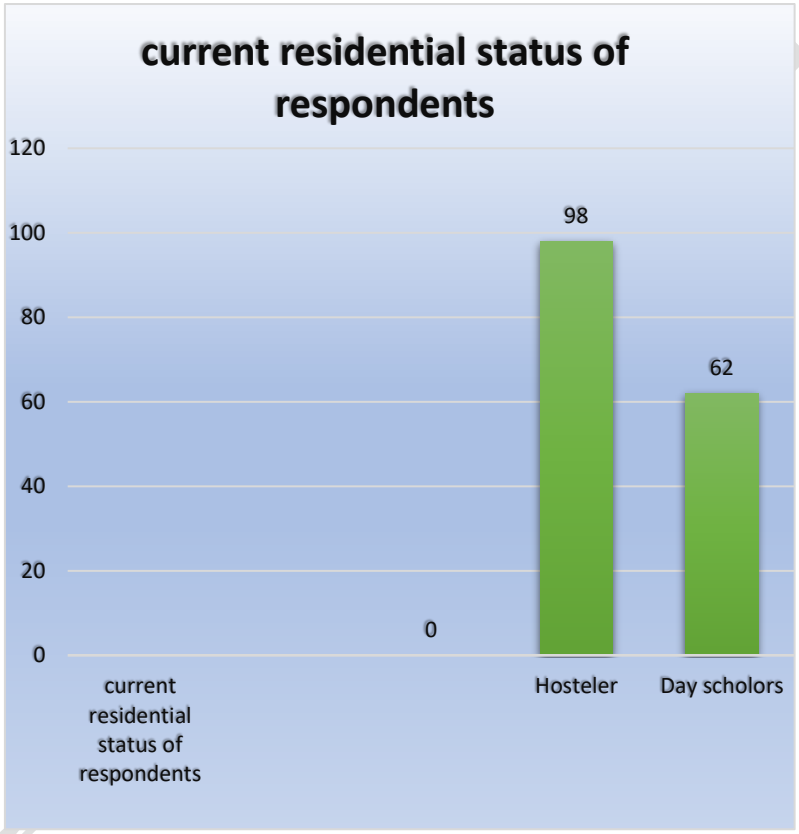
From the data obtained through the samples, we find that most of the respondents were currently pursuing graduation.

- **142** out of **160** (**88.75%**) students are pursuing graduation.
- **14** students (**8.75%**) are currently in Masters courses.
- **4** respondents (**2.5%**) are currently doing PhD.

Current Residential status of students

Table -4

current residential status of respondents		
	no. of students	percentage
Hostelers	98	61.25
Day scholars	62	38.75



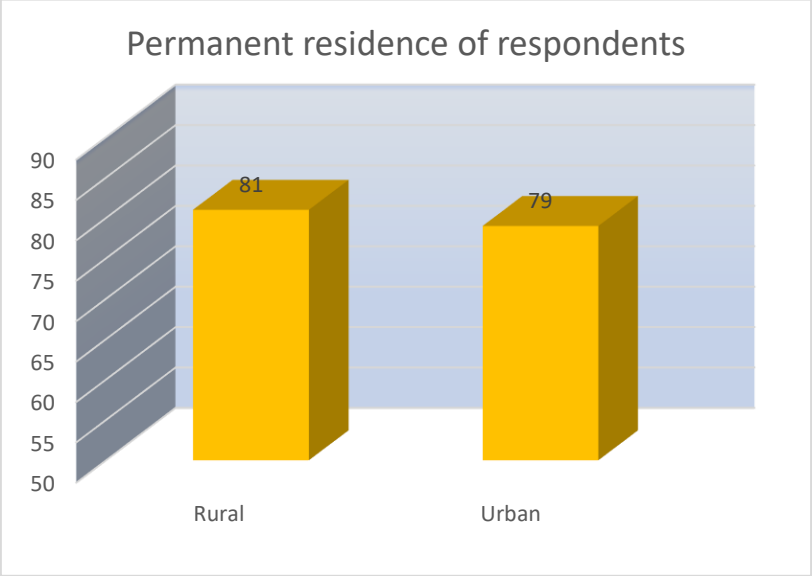
According to the samples obtained,

- **98** out of **160** respondents are hostelers. i.e. **61.25%** of students are hostlers.
- **62** students are day scholars. (**38.75%**)

Permanent residence of students

Table-5

permanent residential status of respondents		
	No. of students	**percentage**
Rural	81	50.625
Urban	79	49.375



According to the samples obtained, we find that

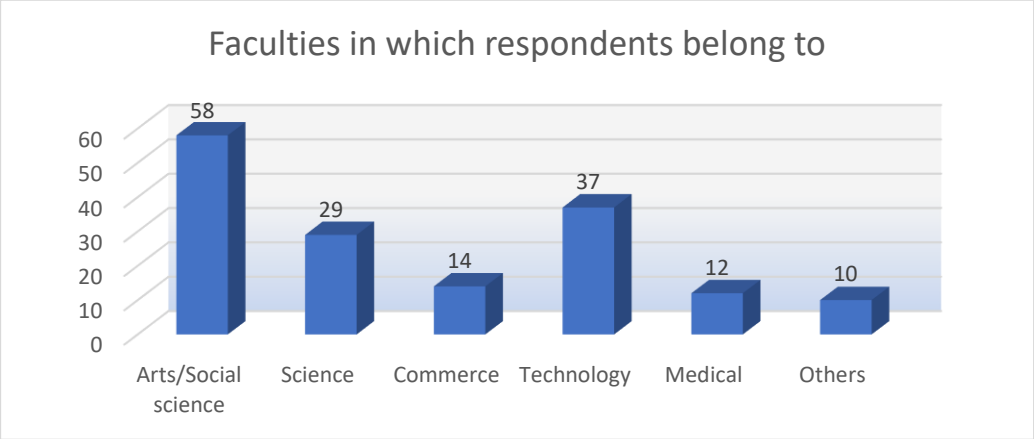
- **81** out of **160** respondents belong to Rural areas. That means **50.625%** of students are from rural backgrounds.
- **79** out of the total students are from urban areas. (**49.375%**)

This implies that there are an equal number of students from both Rural and Urban backgrounds.

Faculties to which students belong to :

Table -6

Faculties to which respondents belong to		
	Number of students	**percentage**
Arts/Social science	58	36.25%
Science	29	18.13%
Commerce	14	8.75%
Technology	37	23.13%
Medical	12	7.50%
Others	10	6.25%



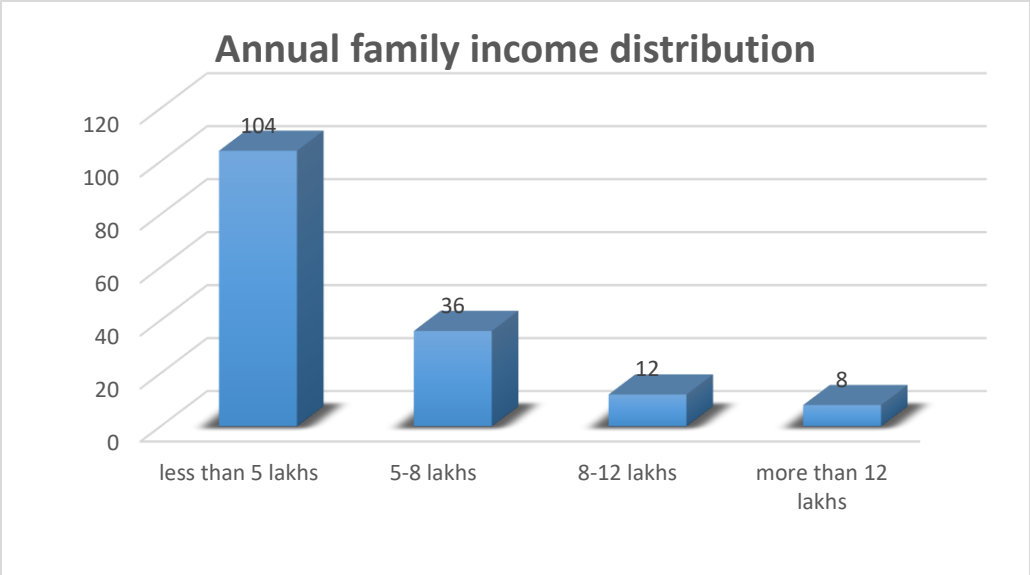
We find that

- **58** out of the total students belong to the faculty of Arts/Social Sciences **(36.25%).**
- **29** students belong to the Institute of Science **(18.13%).**
- **14** students belong to the faculty of commerce **(8.75%).**
- **37** students belong to the technical background. (**23.13%**).
- **12** students are from Medical background. (**7.50%**).
- The remaining **10** students are from other faculties (**6.25%**).

Annual Family Income distribution

Table -7

Annual family income distribution		
	Number of students	**percentage**
less than 5 lakhs	104	65.00%
5-8 lakhs	36	22.50%
8-12 lakhs	12	7.50%
more than 12 lakhs	8	5.00%



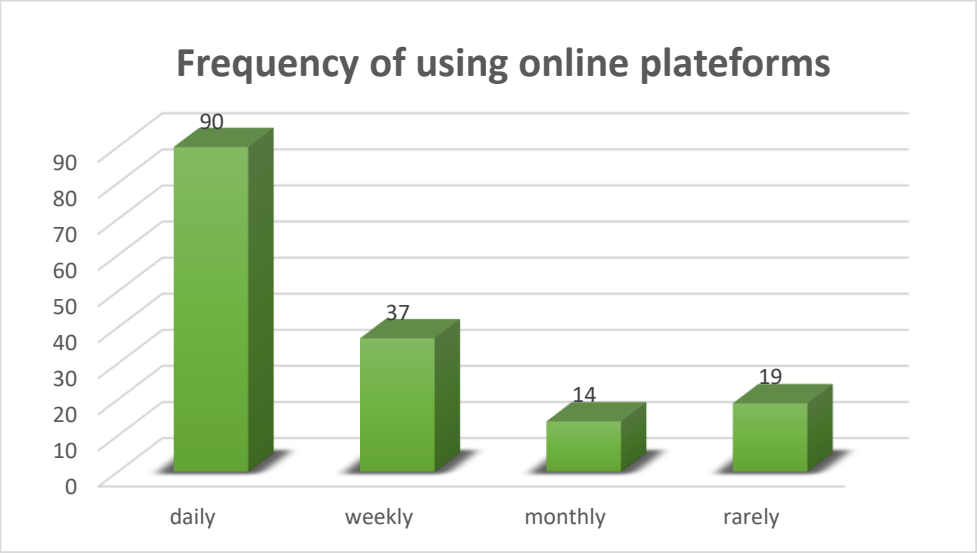
We find that

- 104 out of 160 students have an annual family income of less than 5 Lakhs. This means 65% of students do not fall in any of the income tax brackets. They have low to moderate-income backgrounds.
- 36 students fall under the bracket of 5-8 lakhs (22.50%).
- 12 students are under the 8-12 lakhs category (7.50%).
- 8 students are under more than 12 lakhs category (5%).

Frequency of using online learning platforms

Table – 8

frequency of using online platforms	number of students	percentage
daily	90	56.25%
weekly	37	23.13%
monthly	14	8.75%
rarely	19	11.88%



We find that

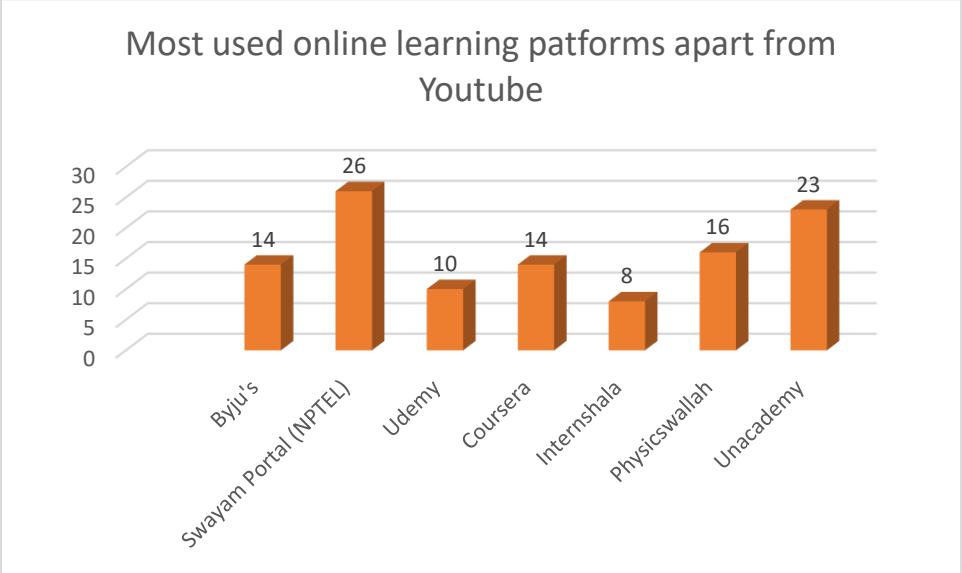
- 90 out of 160 respondents (56.25%) use online platforms daily.
- 37 students (23.13%) use these platforms weekly.
- 14 students (8.75%) use these online learning platforms monthly.
- 19 students (11.88%) use these platforms rarely.

We can interpret that almost 80% of students use online platforms regularly and consistently in one way or another. This data is expected since a recent surge has been observed among students toward using these platforms. Also, the availability of Internet and communication devices at considerably low prices has supported this movement.

Most used online learning platforms

Table-9

Platforms	No. of students	percentage
Youtube	100	62.50%
Byju's	14	8.75%
Swayam Portal (NPTEL)	26	16.25%
Udemy	10	6.25%
Coursera	14	8.75%
Internshala	8	5.00%
Physicswallah	16	10.00%
Unacademy	23	14.38%

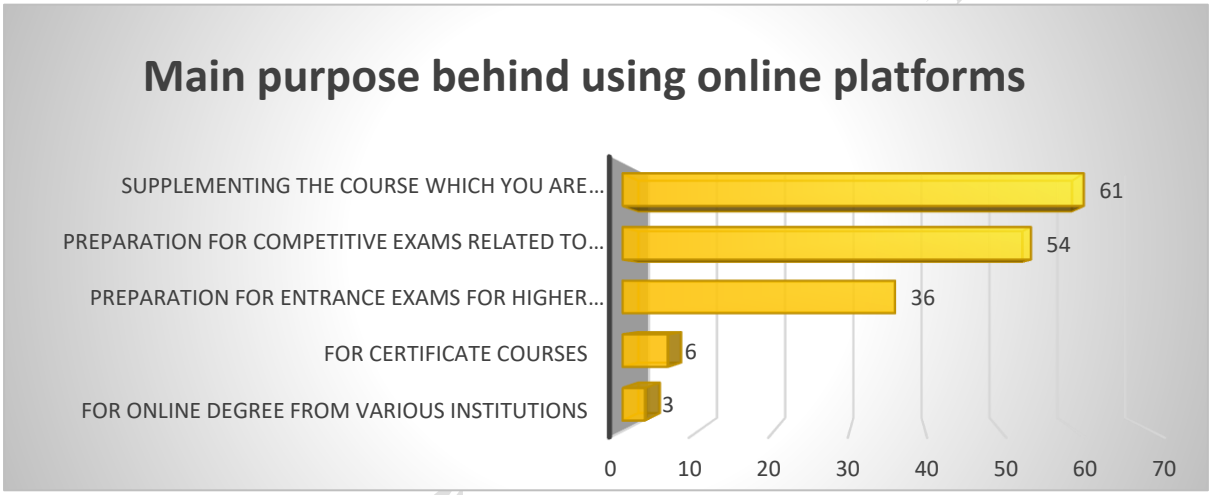


- Apart from YouTube, Swayam Portal, unacademy, Physicswallah, Udemy, Coursera, etc are some of the platforms that are mostly used by students for online learning.

The main purpose for using online learning platforms

Table-10

the main purpose behind using online platforms	number of students	percentage
for online degree from various institutions	3	1.88%
for certificate courses	6	3.75%
preparation for entrance exams for higher education	36	22.50%
preparation for competitive exams related to Jobs	54	33.75%
Supplementing the course that you are currently enrolled in	61	38.13%



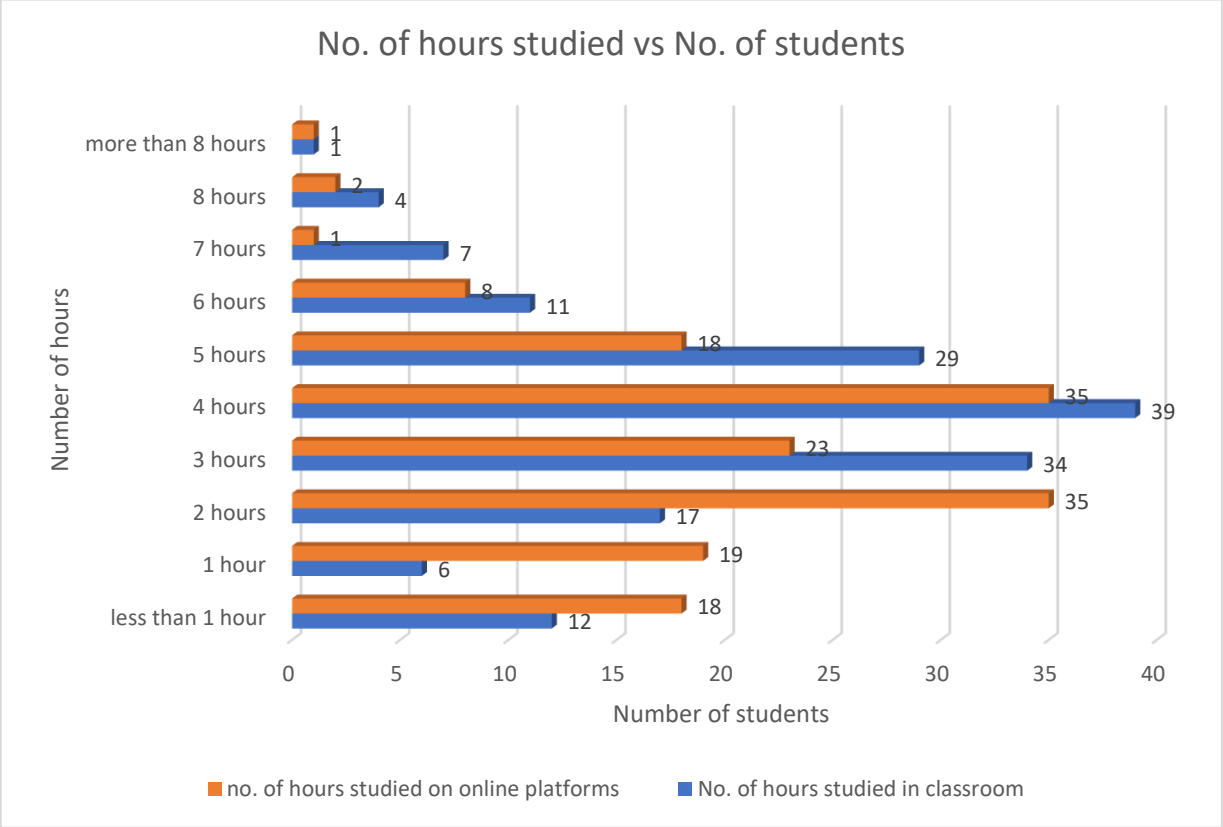
From the above data and the graph, we observe that

- 61 out of 160 people stated that they use online learning platforms to supplement their current course. This is about 38.13% of all students.
- 54 students (33.75%) students use these platforms for the preparation of job-related competitive exams.
- 36 students (22.50%) students use these platforms for the preparation of competitive exams for higher education.
- 6 students (3.75%) students use these platforms for the certificate courses.
- 3 students (1.88%) students use these online learning platforms for getting online degrees.

Data on average time dedicated to online study and traditional classroom study by students

Table-11

No. of hours studied	No. of students (in the classroom)	No. of students (Online learning platforms)
less than 1 hour	12	18
1 hour	6	19
2 hours	17	35
3 hours	34	23
4 hours	39	35
5 hours	29	18
6 hours	11	8
7 hours	7	1
8 hours	4	2
more than 8 hours	1	1



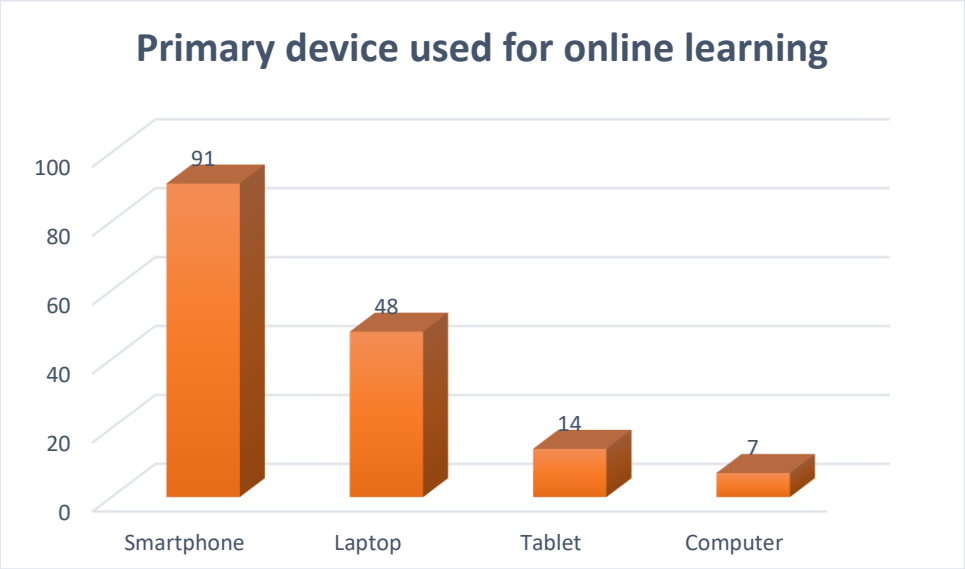
Observing this table and graph, we find that

- Majority of the students **(56.25%)** spend 2 hours to 5 hours daily in the traditional classroom.
- On the other hand, **81.5%** of students spend from 1 hour to 5 hours on online learning platforms. **58.125%** of students spend from 3 to 5 hours on these platforms.
- The shape of the amount of average time spent in traditional classrooms is somewhat bell-shaped with little skewness.
- However, the data on time spent on online learning platforms is skewed to the right.
- It indicates that a very small proportion of students use online learning platforms for more than 5 hours. More students prefer the traditional classroom method when the number of study hours exceeds 5.

Primary device used for online learning

Table-12

primary device for online learning	Number of students	percentage
Smartphone	91	56.88%
Laptop	48	30.00%
Tablet	14	8.75%
Computer	7	4.38%



We observe that

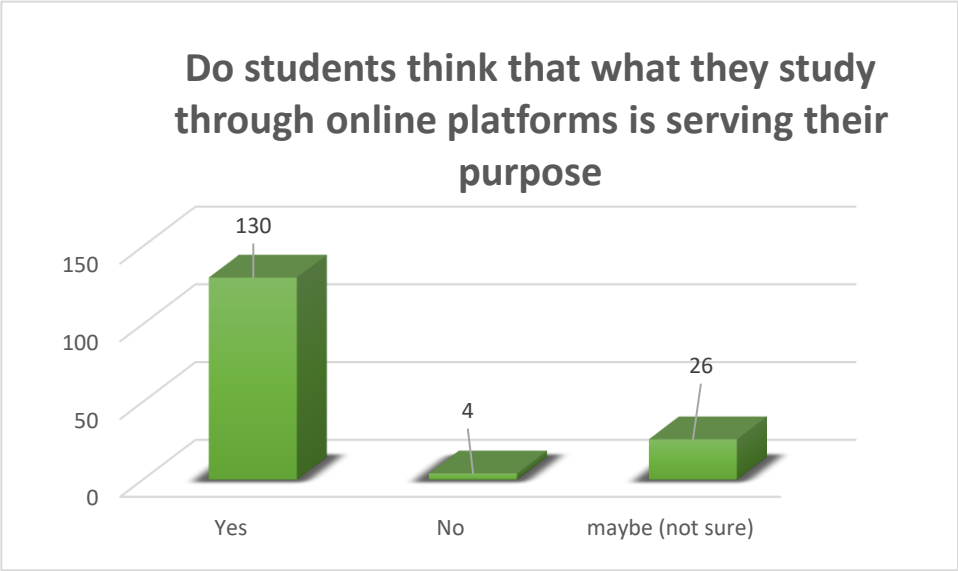
- Majority of students (91 out of 160 or 56.88%) use smartphones as their primary device for learning from online platforms.
- 48 students (30%) use laptops as their primary device.
- 14 students use Tablet (8.75%).
- 7 students use computers (4.38%).

It is quite evident that almost half of the total students do not have access to big-screen devices like laptops and tablets.

do students feel what they study online is serving their educational purpose?

Table-13

do students feel what they study online is serving their purpose?	No. of students
Yes	130
No	4
maybe (not sure)	26

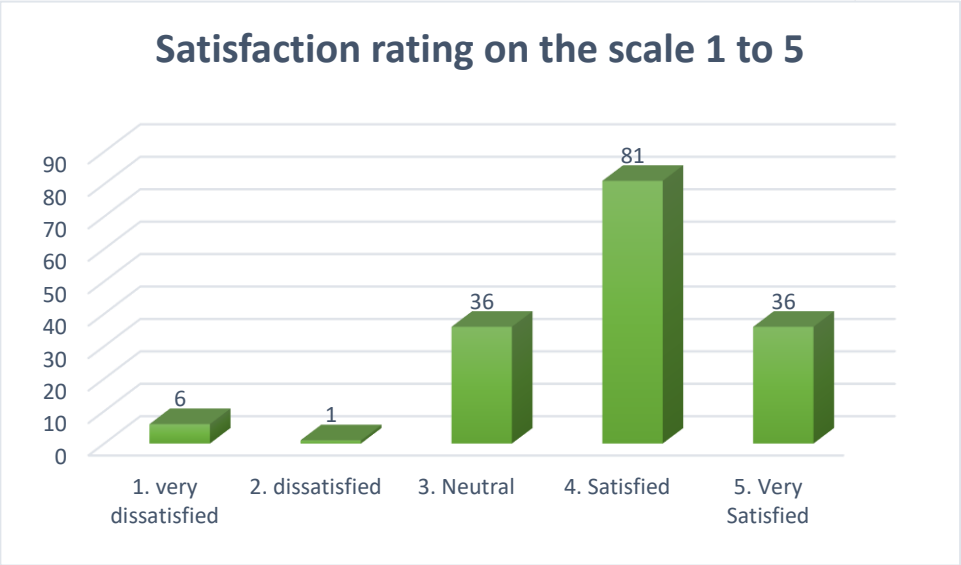


- A vast majority of students feel that what they study on these online platforms is serving their educational needs.
- However, a considerable portion of students (26 out of 160 or **16.25%**) feel that they are not completely sure whether these platforms are serving their requirements or not.

Student's satisfaction with online learning platforms

Table-14

satisfaction ratings	Number of students	percentage
1. very dissatisfied	6	3.75%
2. dissatisfied	1	0.63%
3. Neutral	36	22.50%
4. Satisfied	81	50.63%
5. Very Satisfied	36	22.50%



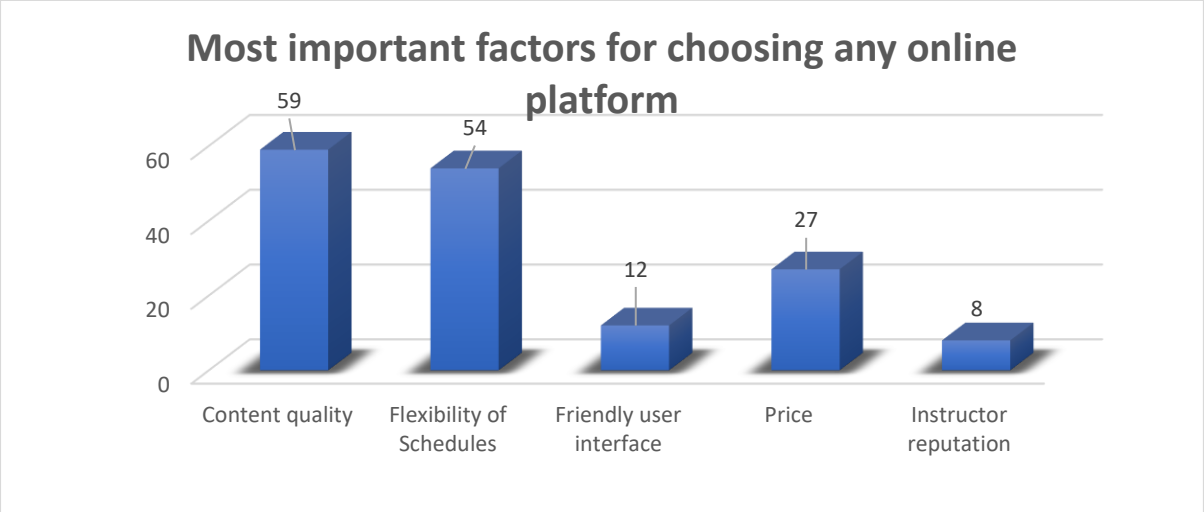
From this data, we observe that

- 72.50% of the total students feel that they are satisfied with the education from these platforms. 22.50% of them are very satisfied.
- The average trend of this data is that on average, students are satisfied with the education and learnings provided by these platforms.
- There are very few people who are not satisfied with these platforms.
- A considerable portion of students remain undecided on this question (**22.50%**).

The most important factors while choosing online platforms

Table-15

most important factor for choosing any online platform	Number of students	percentage
Content quality	59	36.88%
Flexibility of Schedules	54	33.75%
Friendly user interface	12	7.50%
Price	27	16.88%
Instructor reputation	8	5.00%

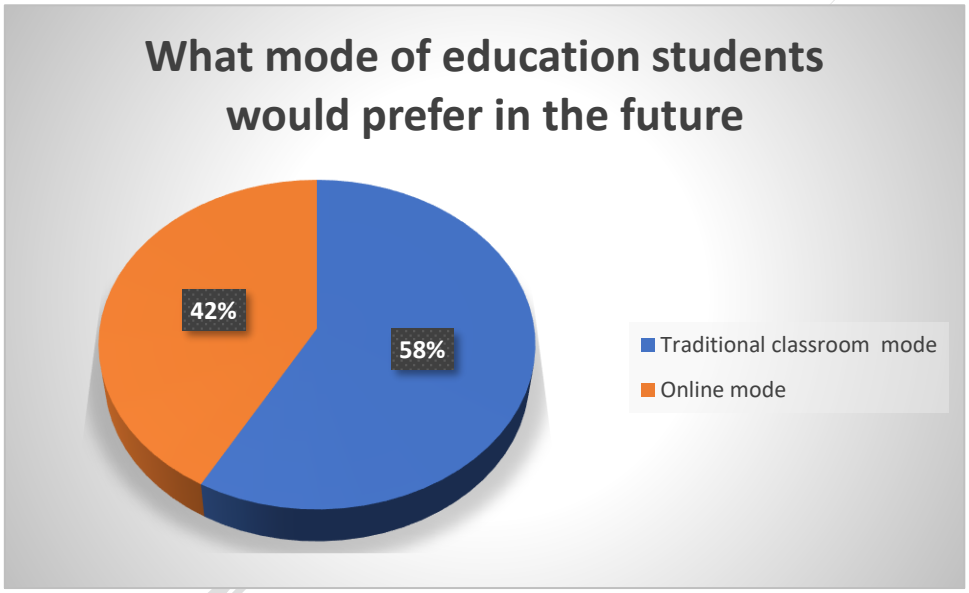


- 36.88% of students preferred Content Quality as the most important factor when choosing online platforms. On the other hand, 33.75% of students feel that flexibility of schedule is the most important one.
- A considerable portion of students (16.88%) felt that Price is the most important factor.

Preference between Online mode of education or Offline (traditional) mode

Table-16

what would students prefer	No. of students	percentage
Traditional classroom mode	93	58.13%
Online mode	67	41.88%



- Almost 58% of the students would prefer the Traditional Classroom method in the future.
- Other 42% feel the future is for technological advancement and they would prefer online mode of education in the future.

Testing the hypotheses

(i) There is no association between the sex of the students and their preferred mode of education (online or offline).

Null Hypothesis (H_0): There is no association between gender and preferred mode of education. In other words, the distribution of males and females is independent of whether they prefer online or offline modes of education.

Alternative Hypothesis (H_1): There is an association between gender and the preferred mode of education. In other words, the distribution of males and females is dependent on whether they prefer offline mode or online mode of education.

We use the chi-square test for the independence of attributes to test this hypothesis.

Contingency table:

	Male	**Female**	**Total**
Online	47	20	67
Offline	75	18	93
Total	122	38	160

Expected frequency table:

	Male	**Female**
Online	51.09	15.91
Offline	70.91	22.09

$\chi^2 = \Sigma \frac{(Oij-Eij)2}{Eij}$ = 2.3691.

Degrees of freedom = (Number of Rows - 1) $\times$ (Number of Columns - 1) = 1

The critical value of the chi-square distribution with 1 degree of freedom is **3.84.** Tabulated value = **3.84** and p-value = **0.123.**

Now, since the p-value is more than the level of significance (**0.05**) and hence we find no reason to reject the null hypothesis based on the samples collected.

Hence, we may accept the Null Hypothesis that there is no association between sex and preferred mode of education.

Male students are equally likely to prefer offline or online modes of education as female students.

(ii) There is no association between the permanent residence of students (rural and urban) and their preferred mode of education.

Null Hypothesis (H_0): There is no association between the residential area (rural or urban) and the preferred mode of education (online or offline). In other words, the residential area does not affect the mode of education preference.

Alternative Hypothesis (H_1): There is an association between the area (rural or urban) and the mode of preferred education (online or offline). This suggests that the preference for online or offline education may depend on whether the residential area is rural or urban.

We use the chi-square test for the independence of attributes to test this hypothesis.

Contingency table:

	Rural	**Urban**	**Total**
Online	21	46	67
Offline	60	33	93
Total	81	79	**160**

Expected frequency table:

	Rural	**Urban**
Online	33.92	33.08
Offline	47.08	45.92

$\chi^2 = \Sigma \frac{(Oij-Eij)2}{Eij}$ = **17.1447**

Degrees of freedom = (Number of Rows - 1) ✕ (Number of Columns - 1) = 1

The p-value is **0.000035.**

Since the p-value is less than the level of significance, we have enough evidence To reject the Null Hypothesis.

Hence, the null hypothesis that There is no association between the residential area (rural or urban) and the preferred mode of education (online or offline) is rejected.

This suggests that the preference for online or offline education may depend on whether the residential area is rural or urban.

(iii) There is no association between the faculties in which students are currently enrolled and their preferred mode of education.

Null Hypothesis (H_0): There is no association between the field of study and the preferred mode of education. In other words, the distribution of students across fields of study is independent of whether they are online or offline.

Alternative Hypothesis (H_1): There is an association between the field of study and the preferred mode of education. In other words, the distribution of students across fields of study is associated with whether or not they prefer online or offline modes of education.

We use the chi-square test for the independence of attributes to test this hypothesis.

Observed frequencies:-

	Arts/social science	**Technology**	**Science**	**Commerce**	**Others**	**Total**
Online	16	19	13	5	14	67
Offline	42	18	16	9	8	93
Total	58	37	29	15	22	**160**

Expected frequencies:-

	Arts/social science	Technology	Science	Commerce	Others
Online	24.29	15.49	12.14	5.86	9.21
Offline	33.71	21.51	16.86	8.14	12.79

$\chi^2 = \Sigma \frac{(Oij-Eij)2}{Eij}$ = **10.8328**

The degree of freedom for this test is (5-1) ×(2-1) = 4

The critical value of the chi-square distribution with 4 degrees of freedom is **9.487729037.**

Since the calculated value of the test statistic is greater than the critical value, we have enough evidence based on the sample to reject the Null Hypothesis.

We can conclude that the distribution of students across fields of study is associated with whether or not they prefer online or offline modes of education.

After a careful examination of the table, we find that the students from the faculty of social science and Arts prefer the offline mode of education more than the online mode of education.

(iv) There is no association between the family income of the students and their preferred mode of education.

We use the chi-square test for the independence of attributes to test this hypothesis.

Null Hypothesis (H_0): There is no association between income bracket and preferred mode of education. In other words, the distribution of individuals across income brackets is not associated with whether they prefer offline or online modes of education.

Alternative Hypothesis (H_1): There is an association between the income bracket and the preferred mode of education. In other words, the distribution of individuals across income brackets is associated with their preferred mode of education.

Observed frequencies:-

	Less than 5 lakhs	**5-8 lakhs**	**8-18 lakhs**	**More than 12 lakhs**	**Total**
Online	41	17	6	3	67
offline	63	19	6	5	93
Total	104	36	12	8	160

Expected frequencies:-

	Less than 5 lakhs	**5-8 lakhs**	**8-18 lakhs**	**More than 12 lakhs**
Online	43.55	15.08	5.03	3.35
Offline	60.45	20.92	6.98	4.65

Since there are two cells less than 5 in the expected frequencies table, we pool those two values and accordingly reduce the degree of freedom by one.

Original degree of freedom = (4-1)×(2-1) = 3

The expected vs observed table is

observed	expected	((O-E)/E)^2
41	43.55	0.003428499
63	60.45	0.001779458
17	15.07	0.016401672
19	20.92	0.008423238
6	5.03	0.0371884
6	6.98	0.019712482
8	8	0
Total	160	0.08693375

The reduced degree of freedom is (3-1)=2
The critical value of the chi-square distribution at the level of significance 0.05 and degree of freedom 2 is **5.991464547.**

Also,
$\chi^2 = \Sigma \frac{(Oij-Eij)2}{Eij}$ = 0.0869.

It can easily be observed that the calculated value of the test statistic is less than the critical value.

Hence we find no reason to reject the null hypothesis.

We conclude that the distribution of individuals across income brackets is not associated with whether they prefer offline or online modes of education.

(v) There is no association between the present class of students and their preferred mode of education.

We use the chi-square test for the independence of attributes to test this hypothesis.

Null Hypothesis (H_0):

There is no significant association between the level of education (Graduation or higher) and the preferred mode of education (online or offline).

Alternative Hypothesis (H_1):

There is a significant association between the level of education (Graduation or higher) and the preferred mode of education (online or offline).

Observed frequencies:-

	Graduation	**Higher than graduation**	**Total**
Online	59	8	67
Offline	83	10	93
Total	142	18	**160**

Expected frequencies:-

	Graduation	**Higher than graduation**
Online	59.46	7.538
Offline	82.54	10.462

The degree of freedom is (2-1)×(2-1) =1

The value of test statistic is

$\chi^2 = \Sigma \frac{(Oij-Eij)2}{Eij}$ = 0.055

The critical value of chi-square distribution with 1 degree of freedom and 0.05 level of significance is 3.84.

Since the calculated value of the test statistic is less than the critical value, we have no reason to reject the null hypothesis based on the samples available.

Hence, we conclude that there is no significant association between the level of education (Graduation or higher) and the preferred mode of education (online or offline).

(vi) On average, students dedicate equal time to both types of education, online mode, and traditional classroom mode.

We use the t-test for the testing of this hypothesis.

H_0: $\mu_1 = \mu_2$

Vs

H_1: $\mu_1 < \mu_2$

Here, The Null hypothesis states that there is no difference between the average time studied in the classroom and online platforms.

The alternative hypothesis states that students spend more time in traditional classrooms than on online learning platforms.

The data is too large to present here. So instead, I have used a data analysis tool in MS Excel. The summary is as follows.

t-Test: Two-Sample Assuming Equal Variances		
	Variable 2	*Variable 1*
Mean	3.7375	2.975
Variance	3.210534591	3.103144654
Observations	160	160
Pooled Variance	3.156839623	
Hypothesized Mean Difference	0	
df	318	
t Stat	3.838473813	
P(T<=t) one-tail	7.46825E-05	
t Critical one-tail	1.649659429	
P(T<=t) two-tail	0.000149365	
t Critical two-tail	1.967451948	

From this analysis, it is evident that the p-value of this test is 7.46825E-0E. This is very small and it is less than the level of significance 0.05.

Hence, the null hypothesis is rejected. We can conclude that students spend more time in traditional offline classes than the online learning platforms.

(vii) The proportion of students which prefers online mode is the same as the ones who prefer traditional classroom mode.

The Null Hypothesis: (H_0: p_1 =0.5)

VS

Alternative Hypothesis (H_1: p_1 >0.5)

The Null Hypothesis states that the proportion of students who prefer the traditional method of education over the online one is 0.5. On the other hand, the alternative hypothesis states that this proportion is different.

From the samples, given that p=0.58

Now,

The z-test statistic is $z = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$.

The calculated value of z is **2.023857703.**

The critical value of z at 0.05 level of significance and for the one-tailed test is 1.64485.

Based on these results, we have enough evidence suggesting that the null hypothesis should be rejected. Hence, we conclude that the proportion of students who want traditional classroom education in the future is significantly

more than 0.5. In other words, more than half of the students would still want the traditional classroom method of education over online learning.

CONCLUSION

After a thorough and careful examination of the samples, and testing the hypotheses, the following results are obtained.

- Most students use online platforms regularly and consistently in one way or another.
- Apart from YouTube, Swayam Portal, unacademy, Physicswallah, Udemy, Coursera, etc are some of the platforms that are mostly used by students for online learning.
- Most of the students use online learning platforms to supplement the current traditional course, prepare for the entrance exams, and prepare for job-related competitive exams.
- On average, a student spends **3.7375** hours per day in the traditional classroom. On the other hand, he/she spends **2.975** hours per day on online learning platforms. Through the testing of the hypothesis, it has been concluded that students, to this day, spend more time in regular classrooms than the online learning platforms.
- Majority of students use smartphones as their primary device for learning from online platforms.
- A vast majority of students feel that what they study on these online platforms is serving their educational needs.
- 72.50% of the total students feel that they are satisfied with the education from these platforms.
- 36.88% of students preferred Content Quality as the most important factor when choosing online platforms. On the other hand, 33.75% of students feel that flexibility of schedule is the most important one.
- There is no association between the preferred mode of education and the sex of the students. Male students are as likely to prefer any mode of education as female students.
- The preference for online or offline mode education may depend on whether the residential area of the student is rural or urban. From this sample, it has been observed that Rural area students prefer the traditional classroom method over the online learning platforms.

- The distribution of students across fields of study is associated with whether or not they prefer online or offline modes of education. After a careful examination of the table, we find that the students from the faculty of social science and Arts prefer the offline mode of education more than the online mode of education.

- The distribution of individuals across income brackets is not associated with whether they prefer offline or online modes of education. A person belonging to any income group is equally likely to prefer the traditional classroom method or the online learning method.

- students from various levels of education are equally likely to prefer the traditional classroom method or online mode of education in the future.

- More than half of the students would still want the traditional classroom method of education over online learning.

Suggestions

- **Enhance Online Learning Platforms:** As most students use online platforms regularly, efforts should be made to enhance the quality and variety of content available on these platforms.
- **Balance Between Traditional and Online Learning:** Given that students still spend more time in traditional classrooms, it's important to strike a balance between online and offline modes of education. Hybrid models of learning, which combine the best aspects of both modes, could be explored.
- **Device Accessibility:** Since the majority of students use smartphones for online learning, ensuring that educational content is mobile-friendly and accessible on various devices is crucial.
- **Address the Rural-Urban Divide:** Policymakers should address the observed preference for traditional learning in rural areas. This could involve improving internet connectivity and access to digital devices in rural areas.
- **Consider Field-Specific Preferences:** The preference for online or offline education varies across different fields of study. Policymakers should take this into account when designing educational policies and consider field-specific strategies.
- **Income and Education Level:** Since the preference for online or offline education does not seem to be associated with income level or education level, efforts should be made to ensure that quality education is accessible to all, regardless of these factors.
- **Future Preferences:** More than half of the students still prefer traditional classroom education. Policymakers should ensure that traditional methods of education are preserved and enhanced, even as online learning continues to grow.

CHALLENGES FACED

Encountering challenges is a common part of any new endeavor, and my experience with conducting an online survey via Google Forms was no exception. Despite facing various difficulties, I approached them with a positive attitude and learned valuable lessons along the way.

One of the challenges I encountered was ensuring a sufficient response rate to our survey. Despite my efforts to promote the survey through various channels, including social media and email, I found it challenging to reach our target audience and motivate them to participate.

Additionally, I faced technical issues with the Google Forms platform, such as formatting errors and difficulties in data analysis. These obstacles required patience and problem-solving skills to overcome, but they ultimately enriched my learning experience.

Despite the challenges, my project became more engaging and rewarding as I navigated through these obstacles. I had the opportunity to interact with diverse individuals and gained insights into the practical application of theoretical concepts in real-life situations.

Working with real-life data and deriving meaningful interpretations from it was a novel experience for me. I am grateful to my supervisor for his invaluable guidance and support throughout every stage of my project, which greatly contributed to my growth and development.