THE IMPACT OF SOCIAL MEDIA ON STUDENT STUDY TIME

TY BSC STATISTICS
PAPER:- ST -609

THE IMPACT OF SOCIAL MEDIA ON STUDENT STUDY TIME

A Project report submitted in partial fulfillment of requirements for the Degree of B.Sc. (Statistics)

PRESENTED BY

Mr. Patil Sagar Pandit (Seat no.-327374)

Mr. Girase Udesing Sattarsing (Seat no.-327372)

Mis. Jain Shweta Dinesh (Seat no.-327373)

Mis. More Rajashri Shankar (Seat no.-327382)

Under the guidance of

Dr.N.S.Chavan

DEPARTMENT OF STATISTICS
AT JAI HIND EDUCATIONAL TRUST'S
ZULAL BHILAJIRAO PATIL COLLEGE,
DHULE-424002

2024-2025

CERTIFICATE

This is to certify that Mr. Patil Sagar Pandit, Mr. Girase Udesing Sattarsing, Mis. Jain Shweta Dinesh and Mis. More Rajashri Shankar students of B.Sc. (Statistics) with specialization in Statistics, at Jai Hind Educational Trust's Zulal Bhilajirao Patil College, Dhule have successfully completed their project work entitled "The impact of social media on study time " based on the data collected through Google form as a part of B.Sc.(Statistics) program under my guidance and supervision during the academic year 2024-2025.

(Prof. Dr. N.S. Chavan)
(HOD, Vice-principal)
Project Guide

Place: Dhule
Date: / / 2025

ACKNOWLEDGEMENT

On this note, firstly we want to convey our regards to honorable Prof. Dr. N.S. Chavan Sir, Head of Department of Statistics (vice principal), Jai Hind Educational Trust's Zulal Bhilajirao Patil College, Dhule for giving their guidance during this project. It was also an intense pleasure to collaborate with Prof.Dr. N.S. Chavan sir who made our work comfortable to oversee.

It would be unfair to go without acknowledging our other teaching staff and beloved friends at from the Department of Statistics Jai Hind Educational Trust's Zulal Bhilajirao Patil College, Dhule for their precious support to us.

I would also like to extend my appreciation to the esteemed faculty members of the **statistics**, including:

- **Dr. N. S .CHAVAN** (Head of Department)
- **Prof. P. U. PATIL**
- **Prof. P. N. DESALE**
- **Prof. P. R. Ruikar** , and all other respected teachers who have provided their guidance and encouragement throughout my academic journey.

We have also thanked to our parents, friends and classmates to give us moral support. We have also thankful to all for directly or indirectly help for project work.

Mr. Patil Sagar Pandit,

Mr. Girase Udesing Sattarsing,

Mis. Jain Shweta Dinesh,

Mis. More Rajashri Shankar

INDEX

Sr.no. Chapter Name

Chapter 1
Introduction

1.1 Introduction

The advent of social media has revolutionized the way people communicate, interact, and share information. However, the increasing use of social media among students has raised concerns about its impact on their academic performance, particularly their study time. Social media platforms, such as Facebook, Instagram, Twitter, and Snapchat, have become an integral part of students' daily lives, with many spending several hours a day scrolling through their feeds.

Despite the potential benefits of social media, such as connecting with peers and accessing educational resources, excessive social media use has been linked to decreased attention span, reduced productivity, and lower academic achievement. Students who spend more time on social media may experience distractions, decreased motivation, and a lack of focus, ultimately affecting their study time and overall academic performance.

This project aims to investigate the impact of social media on students' study time, exploring the relationship between social media usage and study habits. The study will examine the factors that influence students' social media use, the impact of social media on students' attention span and productivity, and the strategies that students use to manage their social media use and maintain a balance between their academic and social lives.

1.2 Research Problem

With the increasing integration of technology in everyday life, mobile phones and social media have become an essential part of students' daily routines. While these tools offer benefits such as access to study materials and communication with peers, they also pose potential distractions that may affect study habits and academic performance.

Many students report using their phones while studying, but the extent to which this habit influences concentration, time management, and learning outcomes remains uncertain. Additionally, social media's role in either enhancing or hindering study efficiency is a topic of debate.

This study seeks to investigate the relationship between mobile phone usage, social media engagement, and study habits among students. Specifically, the research aims to answer the following questions:

- **To what extent do students keep their phones nearby while studying?**
- **How often do students check their phones during study sessions?**
- **Do students perceive social media as a distraction or an aid to studying?**
- **How does mobile phone usage impact the total study time of students?**
- **What are the most common activities students engage in on social media during study hours?**

Understanding these aspects will provide insights into whether mobile phone usage is a significant hindrance to academic productivity or if it can be strategically utilized to enhance learning.

✔ **For Students**: Helps students become more aware of their phone usage habits and their impact on studying.
✔ **For Educators**: Provides educators with data on how technology affects learning, helping them develop strategies to improve student focus.
✔ **For Future Research**: Contributes to ongoing discussions on digital distractions and academic performance, serving as a foundation for further studies.

1.3 Objectives of the Project

- Verify whether the data set is randomly distributed using statistical tests.
- Investigate if there is a significant relationship between gender and preferred social media platforms.
- Determine whether there is a statistically significant difference in Instagram usage between male and female students.
- Analyse if the usage of platforms other than Instagram differs significantly between genders.

- Test if the proportion of students who check their phones while studying is significantly different from 50%.
- Measure the strength and direction of the relationship between how often students check their phones and their study duration.

1.4 Features

The various features of the dataset are explained below:

1. A comprehensive review of existing research on the impact of social media on students' academic performance.
2. A questionnaire-based survey to collect data on students' social media usage habits and study time
3. This feature represents the gender of the respondent. It is a categorical variable with two levels: “Male and Female”.
4. This feature represents the age of the respondent in years. It is a continuous variable ranging from 10 to more than 40 years.
5. This feature represents the educational qualification of the respondent. It is a categorical variable with options like Up to H.S.C., Undergraduate, master’s degree.
6. Online Survey Platform: Utilization of online survey platforms (e.g., Google Forms, SurveyMonkey) to collect and store data.

1.5 Why Excel and R Software :

Microsoft Excel is a spreadsheet software used for data analysis, visualization, and management.

Key Features

1. Data analysis: Formulas, functions, and pivot tables.

2. Data visualization: Charts, graphs, and dashboards.
3. Data management: Sorting, filtering, and formatting.

R Software

R is a popular programming language and software environment for statistical computing and graphics. It is widely used by data analysts, data scientists, and researchers for data analysis, visualization, and modelling.

Uses of R

1. Data analysis and science: R is widely used in data analysis, data science, and machine learning applications.
2. Academic research: R is commonly used in academic research for data analysis, statistical modelling, and data visualization.
3. Business intelligence: R is used in business intelligence applications, such as data analysis, reporting, and data visualization.
4. Finance and economics: R is used in finance and economics applications, such as risk analysis, portfolio optimization, and econometric modelling.

Chapter 2
Data Collection

2.1 Introduction

In our research study we decided to develop the questionnaire and try to be finding the result of given objective. The questionnaire was developed by using the Google Form which is somehow has a challenging task to decide the questions and their arrangement.

Data Collection is an essential component of any research project. In our project, we used **Google Form** as a tool for data collection. Google Forms is a web-based application that allows users to create custom forms and surveys to gather data from respondents.

We designed a questionnaire that includes questions about respondent's gender, age, education, social media usage and study preferences. We ensured that the questions were clear and concise that could potentially confuse the respondents.

To distribute the survey, we used various methods, including social media platform like WhatsApp and social connective apps. We also shared the survey link with our colleagues and friends to reach a broader audience.

The collected data was stored in Google Sheets, where we could easily download and analyse it using statistical software such as " **r Software, EXCEL**". We ensured that the collected data was anonymous and confidential, and we obtained informed consent from the participants before they filled out the questionnaire. Overall, the data collection process was smooth and efficient, and we obtained a considerable number of responses from the participants.

2.2 Determination of Sample Size

The determination of an appropriate sample size is crucial for ensuring the accuracy and reliability of the study's results. The sample size for this study is **385 participants**, which is assumed to be sufficient for statistical analysis. Below are the key factors considered in determining the sample size:

1. Formula for Sample Size Calculation

The sample size can be determined using the formula for a proportion-based study:

$$n = \frac{Z^2 \cdot p \cdot (1-p)}{E^2}$$

Where:

- **n** = Required sample size
- **Z** = Z-score (standard normal deviate) corresponding to the desired confidence level
- **p** = Estimated proportion of the population with the characteristic of interest
- **E** = Margin of error (precision level)

2. Assumptions Used in Sample Size Calculation

a) Confidence Level (Z-score)

- A **95% confidence level** is typically used in social science research.
- The corresponding **Z-score** for a 95% confidence level is **1.96**.

b) Estimated Proportion (p)

- Since there is no prior data on the exact proportion of students affected by phone usage in studies, a conservative estimate of **p = 0.5** is used.
- This ensures maximum variability and provides the most robust sample size.

c) Margin of Error (E)

- A **margin of error of 5% (0.05)** is commonly used in survey-based studies.
- This means the study results should be within ±5% of the true population proportion.

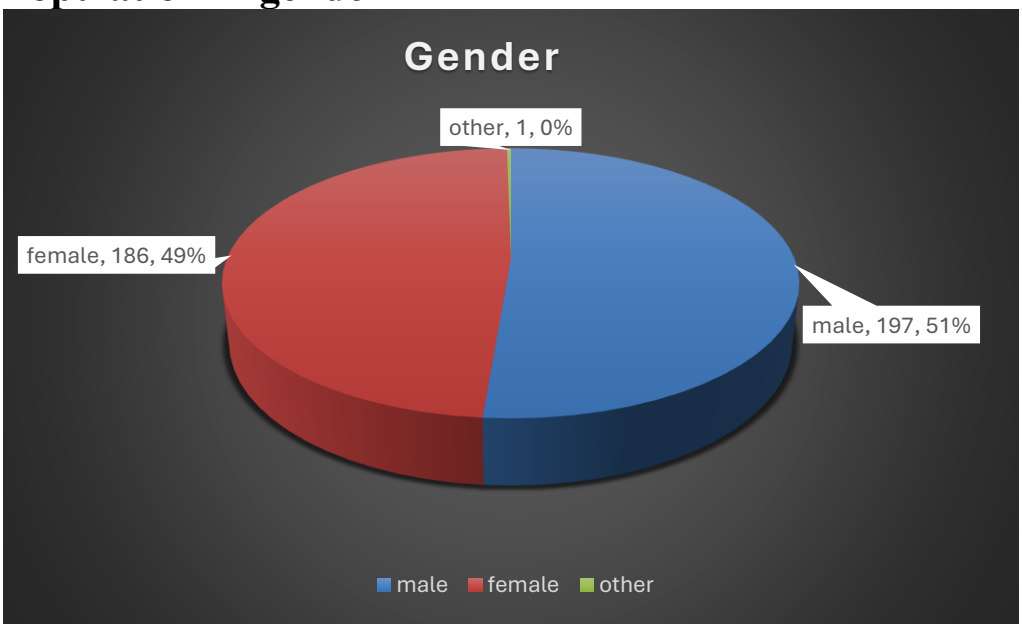
3. Sample Size Calculation

Since sample size **must be a whole number**, we round **up to 385** participants.

4. Justification of the Sample Size

✔ **Sufficient for Generalization**: A sample size of 385 is statistically valid for a large student population.
✔ **Adequate for Statistical Testing**: It meets the requirements for common hypothesis tests and confidence intervals.

2.3 Population in gender



2.4 Time Stamps

Date		Total data
From	To	
05/02/2025	07/02/2025	385

2.5 Questionnaire Design

Designing a questionnaire is crucial for gathering reliable data. This section outlines key considerations and steps in questionnaire design.

- **Research Objectives:**

  Define clear research objectives for the questionnaire to guide its design process effectively and ensure a focused approach.

- **Target Audience:**

  To Identifying the target audience or respondents who will be participating in the survey. The *target audience* for this survey was *social media users* in our surrounding who use in study time.

- **Question Types:**

  Select question types based on the required information, such as multiple-choice, open-ended, and demographic questions, each serving a specific purpose in collecting various data types.

- **Question Wording:**

  Ensure clear and concise questions, use simple language without technical terms, and pre-test the questionnaire with a small sample to identify and address any potential issues with question wording.

- **Response Options:**

  Provide comprehensive multiple-choice response options for each question to cover all possible choices and consider a combination of

structured response options and open-ended questions to gather a diverse range of insights.

- **Pilot Testing:**

We are conducting a pilot test of the questionnaire with a small sample of respondents. This helps identify any potential issues or areas of improvement before administering the questionnaire to the larger population.

In this way we design the questionnaire for the research objectives.

2.6 Question and its options

Now, we will see the types of question and their options as follows :

1) Gender*

A. Male
B. Female
C. Others

2) Age (Years)*

A. 0-10
B. 10-20
C. 20-30
D. 30-40

3) Your Education*

A. Up to H.S.C
B. Undergraduate
C. Master's degree
D. School

4) Which social media platform do you use?

A. Facebook
B. Instagram
C. Twitter
D. Other

5) How many hours do you spend on social media per day?

A. 1 Hours
B. 2 Hours
C. 3 Hours
D. 4 Hours

6)What time of day do you usually study?

A. Morning
B. Evening
C. Afternoon
D. Night

7) Do you usually study with your phone nearby?

A. Yes
B. No

8) How often do you check your phone while studying?

A. Never
B. Always
C. Occasionally
D. Almost always

9) Do you think social media affects your study time?

A. Yes
B. No
C. Unsure
D. None of these

10) What do you usually do on social media?

A. Browse news

B. post updates

C. chat with friends

D. study material

11) How does social media affect your study time?

A. Reduces study time

B. No impact

C. Help

D. None of these

12)How many hours do you spend studying per day?

A. 3 Hours

B. 2 Hours

C. 4 Hours

D. 1 Hours

2.7 Data Collection and Procedures

The data collection procedures section describes the systematic approach used to collect accurate and reliable data through the questionnaire, covering aspects such as target population, sampling techniques, questionnaire distribution, and data collection methods.

- **Target Population**

  The target population is the specific group from which data will be collected, and defining it clearly ensures the questionnaire's relevance and representativeness by considering demographics, geographic location, and other relevant characteristics.

- **Sampling Techniques**

  Convenience sampling is used to select participants based on availability and aligns with research objectives and available resources, determining how participants are chosen from the target population.

- **Questionnaire Distribution**

Questionnaire distribution methods include face-to-face interviews, online platforms, and mobile applications, aiming to maximize response rates and ensure participant accessibility and convenience.

- **Data Collection Methods**

  Data collection methods include self-administered and interviewer-administered questionnaires, chosen based on questionnaire complexity, target population, and available resources.

- **Pilot Testing**

  Pilot testing involves a small sample of participants to evaluate questionnaire clarity, obtain feedback, and make necessary improvements before full-scale data collection.

The data collection procedures adopted for the questionnaire ensure a systematic and rigorous approach to gathering reliable and valid data.

Chapter 3
Exploratory Data Analysis

In statistics, Exploratory data analysis (EDA) is a statistical approach that summarizes data characteristics using visualization methods, promoting data exploration beyond formal modelling and hypothesis testing. It encourages statisticians to formulate hypotheses and gather new data. EDA differs from initial data analysis (IDA) by focusing on broader data exploration rather than specific model assumptions and transformations.

Now we will see how the data cleaning and pre-processing was done.

3.1 Data Cleaning and Pre-processing

Data cleaning and pre-processing is essential for ensuring accurate and consistent data for analysis. This involves identifying and rectifying errors, inconsistencies, missing values, and outliers. Steps include data validation, handling missing data, removing duplicates, and standardizing variables, among others.

- **Missing Data Handling**

  Missing data, the absence of responses in a questionnaire, needs to be addressed to prevent bias. Techniques like imputation or exclusion of cases with substantial missing data can be used to handle missing values and ensure valid results.

- **Outlier Detection and Treatment**

  Outliers, extreme values that deviate significantly from the data, can be detected using statistical methods like z-score or visualization techniques. Addressing outliers is important to prevent their undue

influence on the analysis. Treatment options include removing or transforming outliers based on their impact and context.

- **Data Coding and Recoding**

  Data coding assigns numerical or categorical codes to questionnaire responses, facilitating analysis. It is used for open-ended or qualitative data, as well as grouping continuous variables or creating derived variables based on specific criteria or scales.

- **Data Transformation**

Data transformation involves converting variables to different scales or formats, such as converting categorical variables into binary or dummy variables for regression or other analytical techniques.

By implementing thorough data cleaning and pre-processing techniques, the collected questionnaire data is prepared for meaningful analysis.

After cleaning and pre-processing, the dataset is imported and transformed as needed for further analysis. Descriptive statistics and visualizations are then used to draw appropriate conclusions.

3.2 Data Visualisation

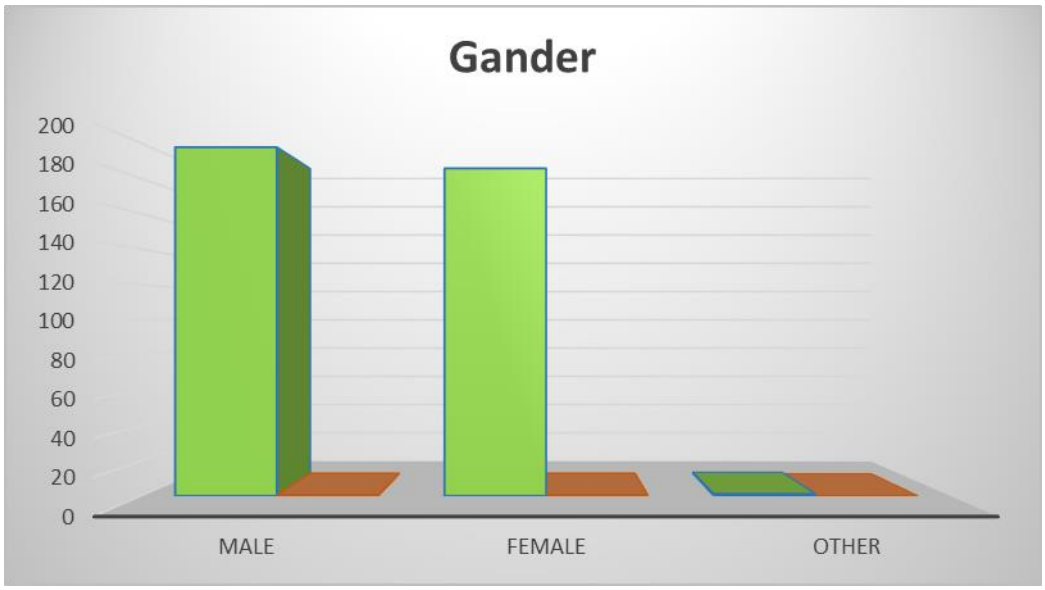


Figure 3.1 Gender

Figure 3.1: The dataset contains 385 total number of gender and 197 males,185 females and 1 other.

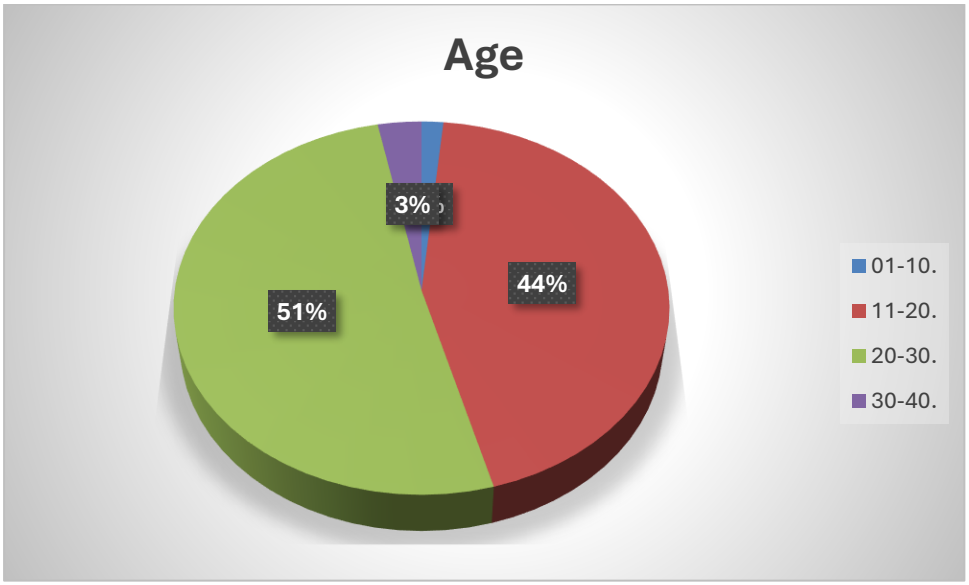


Fig 3.2 Age

Figure 3.2: Most respondents **(95.33%)** are between **11-30 years old**, indicating a **young demographic**. The **20-30 age group (51.17%)** is the most dominant, possibly reflecting the **target audience** of the survey. Very few individuals above **30 years old (3.12%)** participated, suggesting lower engagement from older adults.

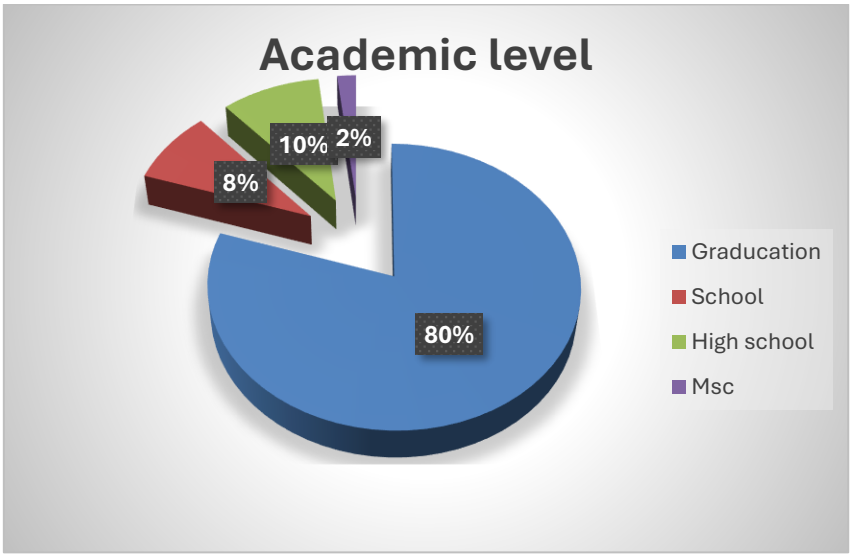


Fig 3.3 Academic level

Figure 3.3: 80% of respondents are at the **graduation level**, meaning the majority are likely college/university students. **Only 1.82% are MSc students**, showing that **postgraduate participation is very low**. **Less than 20%** of respondents are in school or high school, suggesting that the survey's target audience is **mainly higher education students**.

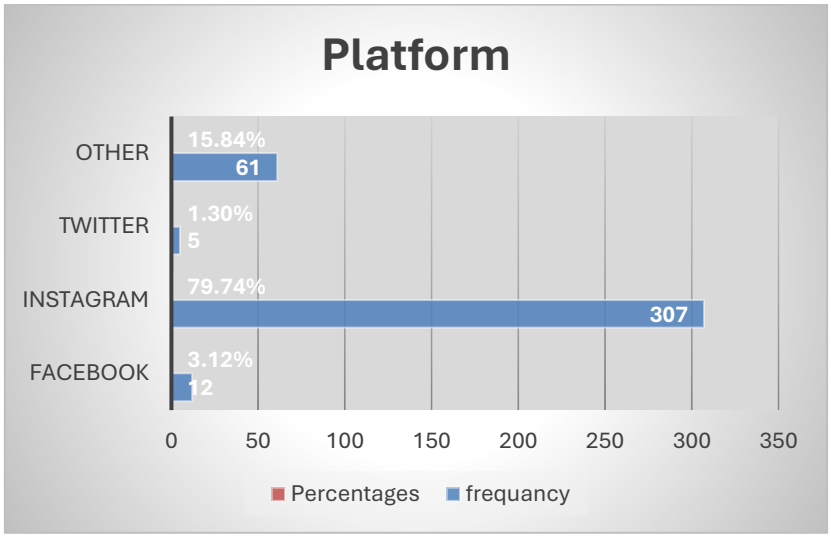


Figure 3.4 Social Media Platforms

According to the data presented in the figure 3.4, it appears that the combination of users with social media platform of Instagram, Twitter, or Facebook contributes to a total of 85% of the user base. Among these data, the Instagram and other social media platforms holds the largest share with 79.74% and 15.84% while the remaining 5% is divided between the Facebook and twitter are used by user.

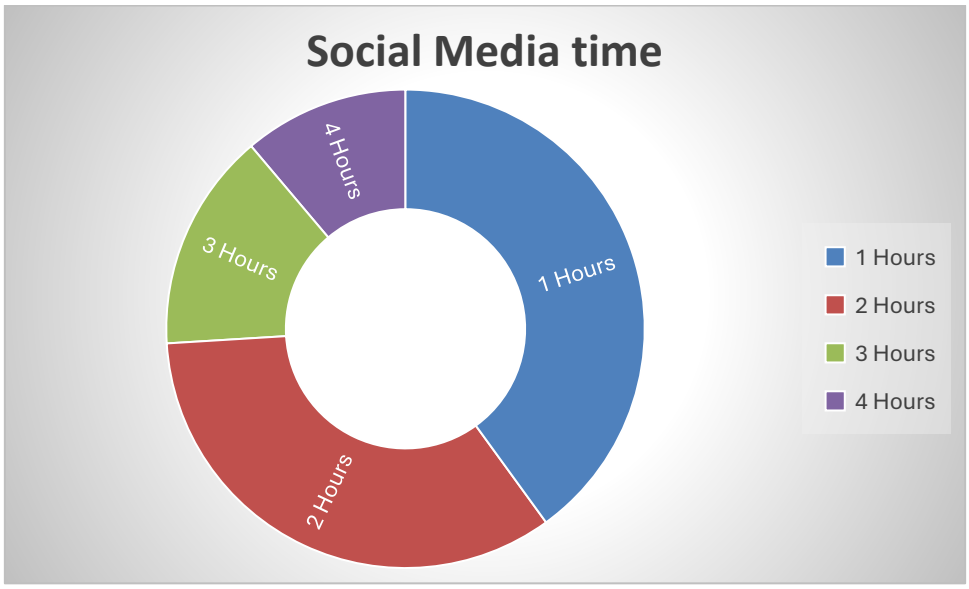


Figure 3.5 social media use time

Fig 3.5 The majority of people (74.03%) use social media for 1-2 hours daily, indicating a balanced or moderate usage pattern. About 25.98% of respondents use social media for 3 or more hours daily, suggesting a notable proportion of heavy users. Understanding this data can help in analysing social media habits and their potential impact on productivity, mental health, or digital marketing strategies.

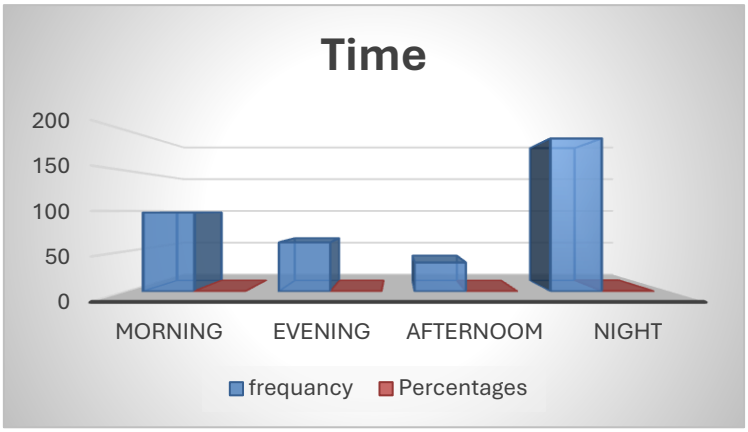


Figure 3.6 Study session of per day

The data indicates that **nearly half of the students prefer studying at night**, possibly due to fewer distractions, quiet environments, or personal productivity patterns. The **morning** is also a popular study time, suggesting that some students prefer to start their day with focused study sessions.

Evening and afternoon study sessions are less common, with the afternoon being the least preferred. This could be due to students having classes, extracurricular activities, or other commitments during this time.

Most of the student's study **at night or in the morning**, with a smaller portion choosing the **evening or afternoon**. Understanding these study habits could help in structuring study plans, improving productivity, and tailoring learning environments to students' peak focus periods.

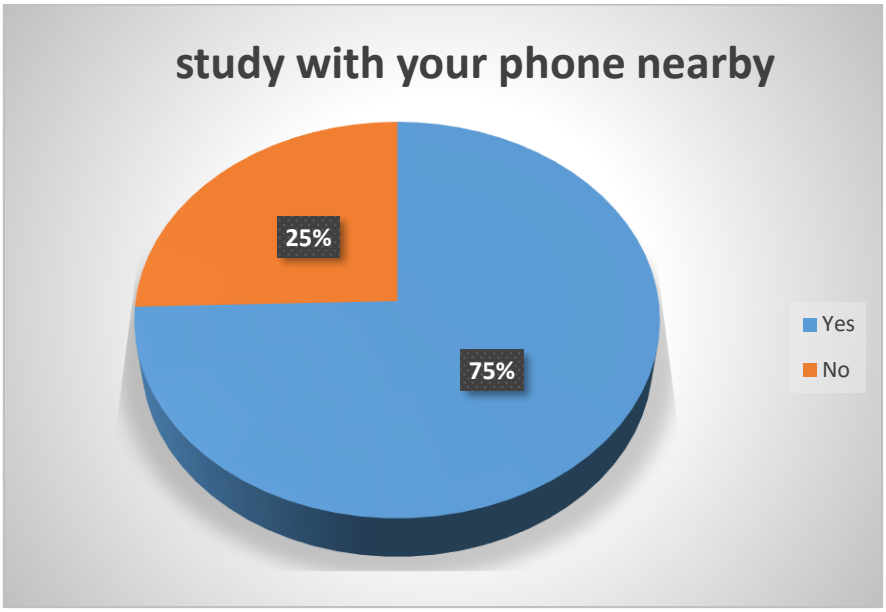


Figure 3.7 study with your phone nearby

74.55% (287 respondents) reported that they do study with their phone nearby. 25.45% (98 respondents) stated that they do not study with their phone nearby. The data suggests that a significant majority of respondents (nearly three-fourths) keep their phones close while studying. This could indicate reliance on phones for study-related activities (e.g., looking up information, using study apps) or distractions from social media and other non-academic activities. The lower percentage of students studying without their phones suggests that fewer people choose to eliminate phone-related distractions while studying.

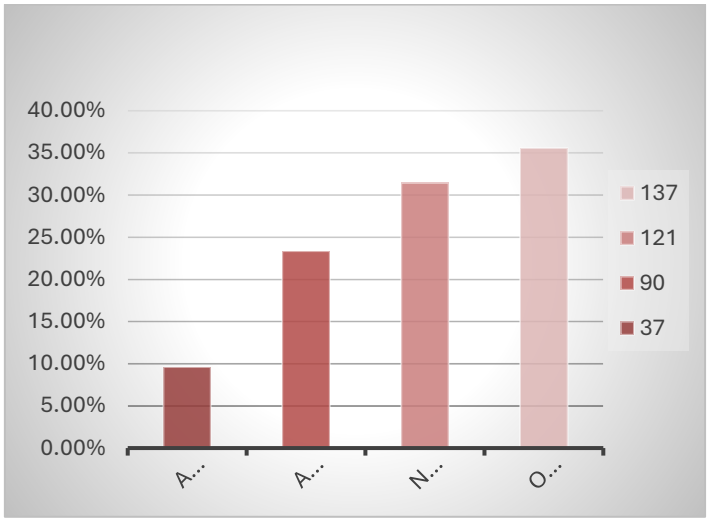


Figure 3.8 check your phone while studying

Most students (around 65%) check their phones at least occasionally while studying, with **23.38% always using their phone** and **9.61% checking it frequently**. However, **31.43% manage to avoid using their phones completely**. The highest percentage (35.58%) falls under the **occasional** category, indicating that most students try to balance their phone use while studying.

These results suggest that while phones are a common distraction, many students attempt to limit their usage. The relatively high percentage of students who never check their phones indicates that some individuals prioritize focus and concentration.

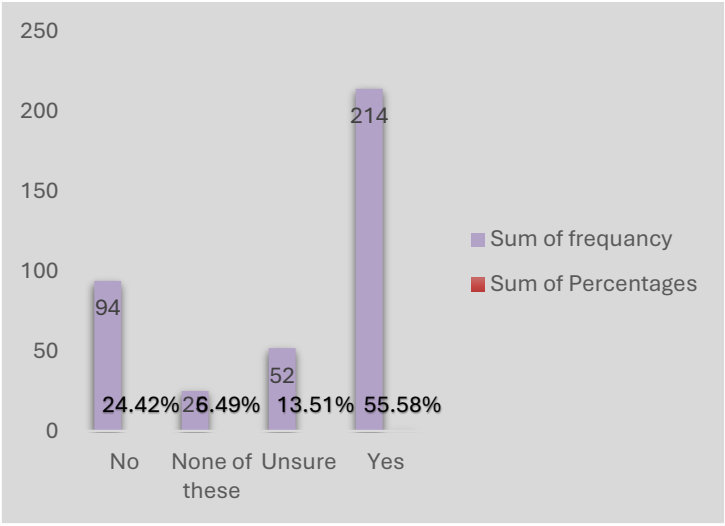


Figure 3.9 check your phone while studying

A **majority (55.58%) acknowledge that social media affects their study time**, indicating that distractions from social platforms could be a significant issue for students. However, **24.42% believe it does not affect them**, suggesting that some students may have strong self-control or structured study habits.

The **13.51% who are unsure** might occasionally experience distractions but not consistently enough to confirm an impact. The **6.49% choosing "None of these"** might have different interpretations of the question or believe other factors influence their study time more than social media.

This data suggests that social media plays a major role in students' study habits, often acting as a distraction. However, a significant portion of students either manage to control its influence or do not perceive it as a problem.

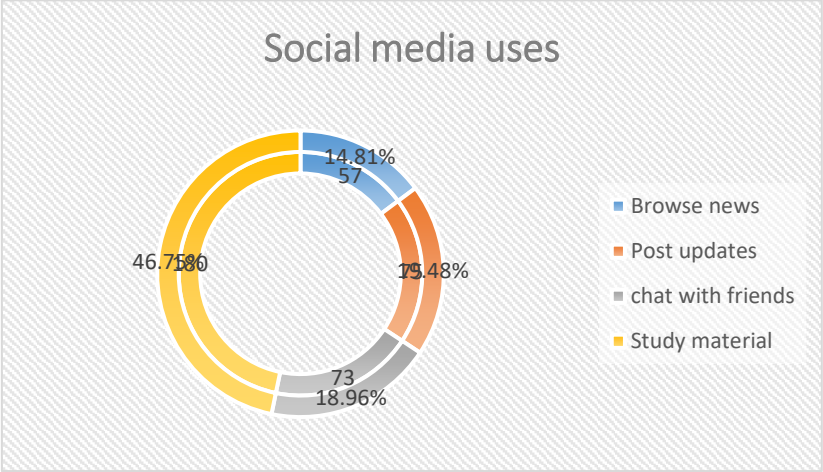


Figure 3.10 social media use

The data suggests that while social media is often associated with distractions, nearly **half of the respondents (46.75%) use it for educational purposes**. This indicates that platforms may serve as a valuable resource for learning, group discussions, and accessing study materials.

However, a **significant portion (over 50%) use social media for non-academic activities**, such as posting updates, chatting, or browsing news. This highlights the dual nature of social media—while it can support learning, it also presents potential distractions.

Social media plays a mixed role in students' lives, with many leveraging it for educational benefits while others engage in social interactions and news consumption. Encouraging effective use of social media for learning while minimizing distractions could enhance productivity.

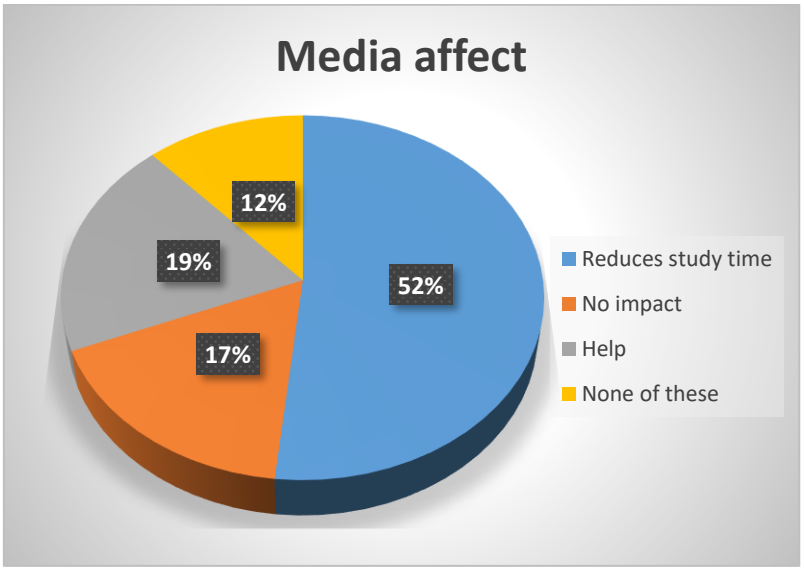


Figure 3.11 Media affect your study time

A **majority (51.69%) feel that media negatively impacts their study time**, likely due to distractions from entertainment, social media, and other non-educational content. However, **19.48% find media beneficial for their studies**, possibly using it for online resources, study videos, or digital learning tools.

The **17.14% who see no impact** may have strong self-discipline or structured study habits that prevent media distractions. The **11.69% selecting "None of these"** suggests that some respondents might see media's impact as situational or dependent on content type.

The data suggests that media can both hinder and aid learning, depending on how it is used. While many students struggle with media-related distractions, some effectively incorporate it into their study routine. Encouraging responsible media consumption and using it for educational purposes could help students maximize its benefits.

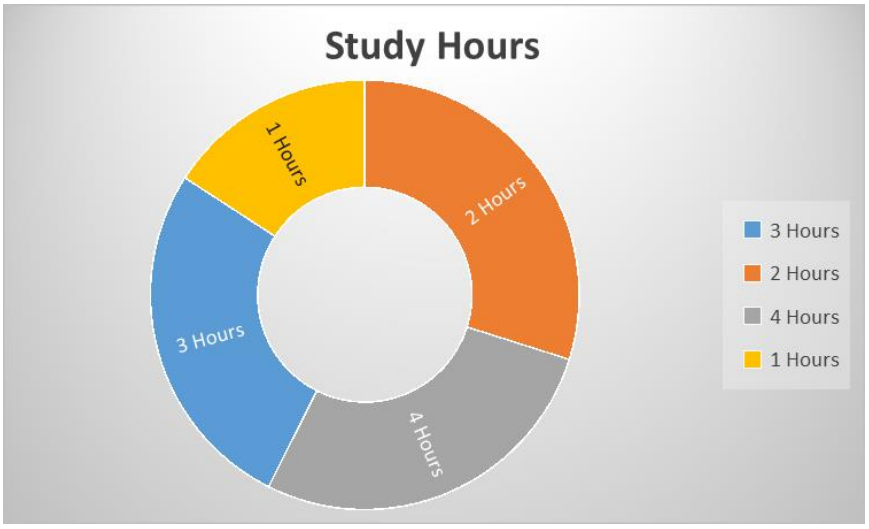


Figure 3.11 study time per day

Most students dedicate between **2 to 4 hours daily** for studying, with **2 hours being the most common study duration**. This suggests that a moderate study routine is preferred by most students. However, **15.84% study for just 1 hour per day**, which may indicate lower academic engagement or reliance on other learning methods.

The data suggests that students generally invest a reasonable amount of time in their studies, but individual study habits vary. Those studying for **4 hours** may be preparing for exams or handling demanding coursework, while those studying **1 hour** might be more efficient or rely on alternative learning methods.

Chapter 4

Statistical Analysis

4.1 Run Test

Here's a brief overview of the Run Test:

 # What is the Run Test?

The Run Test is a statistical test used to determine if a sequence of observations appears to be random.

 # History:

The Run Test was first introduced by Walter A. Shewhart in 1931, an American engineer and statistician. However, the test gained popularity in the 1950s and 1960s with the work of statisticians like Milton Friedman and Frederick Mosteller.

 # Why is it used?

The Run Test is used to:

1. Check for randomness: Determine if a sequence of observations appears to be random or not.
2. Detect patterns: Identify patterns or trends in a sequence of observations.
3. Validate assumptions: Check if the assumptions of randomness and independence are met in statistical analyses.

RUN TEST

Here, we have the sequence of Male and Female is,

"MFFFM----------MFFFFF "

We wish to test

H_0 : Given sample (data) is randomly distributed

VS

H_1 : Given sample (data) is not randomly distributed.

α : 0.01

R = Number of Runs =172,

The runs are lager then go to normal approximation

n_1= Number of males =198
n_2= Number of females=187

Calculate the expected number of runs
Using the formula:

$$E(R) = \frac{2n_1n_2}{n_1 + n_2} + 1$$

$$\sigma(R) = \sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}$$

μ = E(R) = 193.34
σ = 9.789128

Calculate the Z:

$$Z = \frac{R - E(R)}{\sigma(R)}$$

where R = 172 (observed number of runs), we get:

Z= (172 - 193.34) /9.789 ≈ -2.18025
|Z|=2.18025

Determine the significance
The calculated z (2.18025) is greater than the critical z (2.58) for a two-tailed test at α = 0.01.
$|Z_{tab}| = Z_{0.005}$
$Z_{tab} = Z_{0.005} = 2.58$
Therefore, $Z_{cal} < Z_{tab}$

i.e. We fail to reject the null hypothesis (H0) that the sequence of observations is random distributed
The Run Test suggests that the survey conducting randomly.

4.2 Chi-square test of independence

To examine the association between gender and social media platform usage

The Chi-square test of independence is a statistical test used to determine if there is a relationship between two categorical variables. It assesses whether the observed frequencies of the variables in a contingency table differ significantly from the expected frequencies, assuming that the variables are independent.

The test involves comparing the observed frequencies in each cell of the contingency table with the expected frequencies that would occur if the variables were independent.

Contingency table.

	Facebook	Instagram	Twitter	Other	Total
Males	7	94	2	30	133
Females	5	213	3	31	252
Total	12	307	5	61	**N=385**

We are using chi-square test
H_0: There is no significance association between gender and social media platform usage.
H_1 : There is a significance association between gender and social media platform usage.
Calculating the chi-square
N=385

$$\chi^2 = \Sigma \frac{(oij - eij)^2}{eij}$$

Ei = (Row Total × Column Total) / Total
Where:
- Ei is the expected frequency for each cell
- Row Total is the total frequency for each row
- Column Total is the total frequency for each column
- Total is the grand total frequency

Output:
Here is the chi-squared table value:

$\chi^2 = 10.9401$

Degrees of freedom (df) = 3

p-value = 0.01205

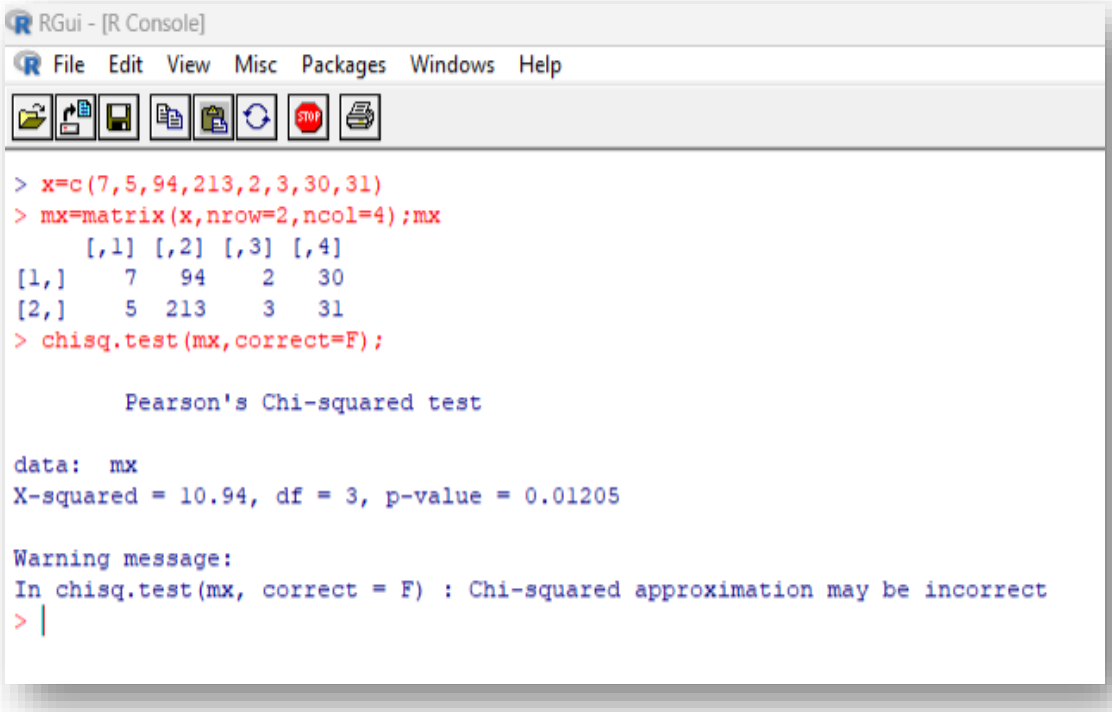
This means that:

- The chi-squared statistic (χ^2) is 10.11
- The degrees of freedom (df) are 3
- The p-value is 0.0176, which is less than the typical significance level of 0.05

Therefore,
we reject the null hypothesis of independence and conclude that there is a statistically significant association between the variables.

i.e. Gender and Social media platform use by users are independent.

 # R output



4.3 Large Sample Test: -

To compare Instagram usage between male and female students

Testing equality of two population proportions

Test of significance for difference of proportions:

Let: n_1= Sample size drawn form 1 proportions

n_2 = Sample size drawn form 2 proportions

X_1= Number of males of specific platform in first sample

X_2 = Number of females of specific platform in second sample

p_1 = x1/n1= proportion of specific platform in a first sample.

p`2=x2/n2= proportion of specific platform in a second sample.

P1=proportion of specific items in first population.

P2 = proportion of specific items in second population.

	Instagram(xi)	Total(ni)	proportion
male	94	133	0.7067
female	213	252	0.8452

We wish to test:

H_0: There is no significance different between Instagram. Users of male and female.

H_1: There is significance different between Instagram. Users of male and female.

H_0 :P1= P2 vs H_1 : P1 > P2.

The pooled sample proportion is:

$$p = \frac{x_1 + x_2}{n_1 + n_2}$$

P= 0.797403

(1 – P) =0.2025

Test Statistic:

The test statistic for a two-proportion z-test is:

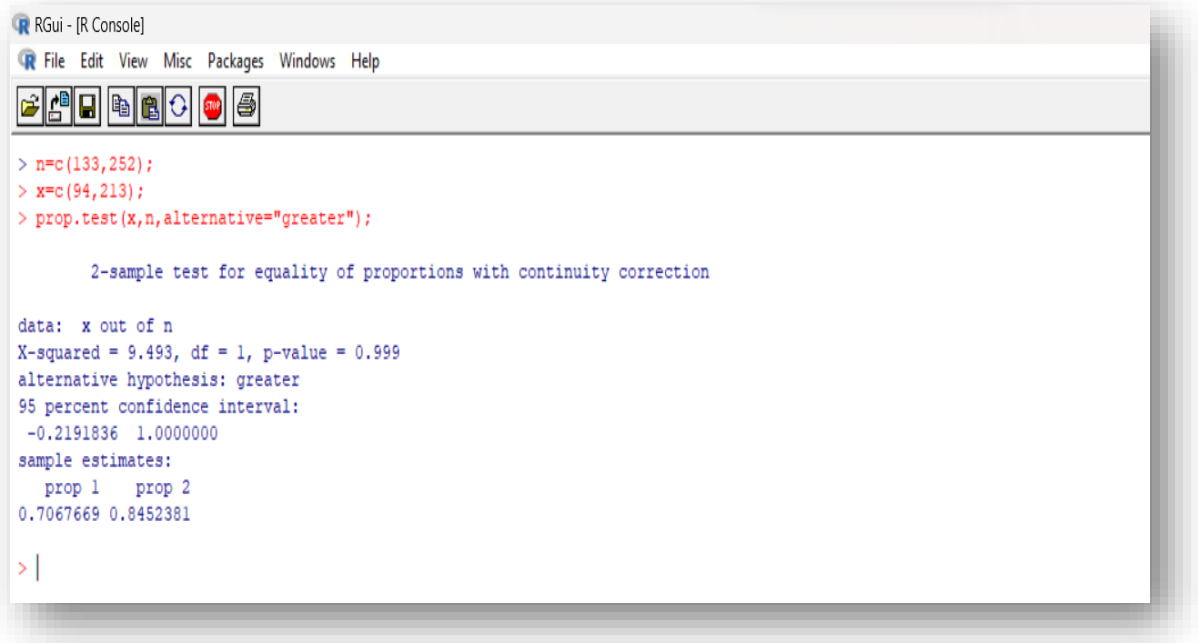
$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

The test statistic (z) is -3.21, and the p-value is 0.0013.

Conclusion:

Since the p-value (0.0013) is much smaller than the common significance level (α=0.05), we reject the null hypothesis.

This suggests that there is a statistically significant difference in Instagram usage between males and females, with females having a higher proportion of users.



4.4 To compare other social media platform usage between male and female students

Testing equality of two population proportions

Test of significance for difference of proportions:

Let: n_1= Sample size drawn form 1 proportions

n_2 = Sample size drawn form 2 proportions

X_1= Number of males of specific type in first sample

X_2 = Number of females of specific type in second sample

p_1 = x1/n1= proportion of specific type items in a first sample.

p_2= x2/n2= proportion of specific type items in a second sample.

P1=proportion of specific items in first population.

P2 = proportion of specific items in second population.

We wish to test:

H_0: There is no significance different between other platform Users of male and female.
H_1: There is significance different between other platform Users of male and female.
H_0: P1= P2 vs H_1 : P1 > P2.

	Other social platform	Total(ni)	Proportion
Male	30	133	0.2256
Female	31	252	0.1230

Test Statistic:

The test statistic for a two-proportion z-test is:

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

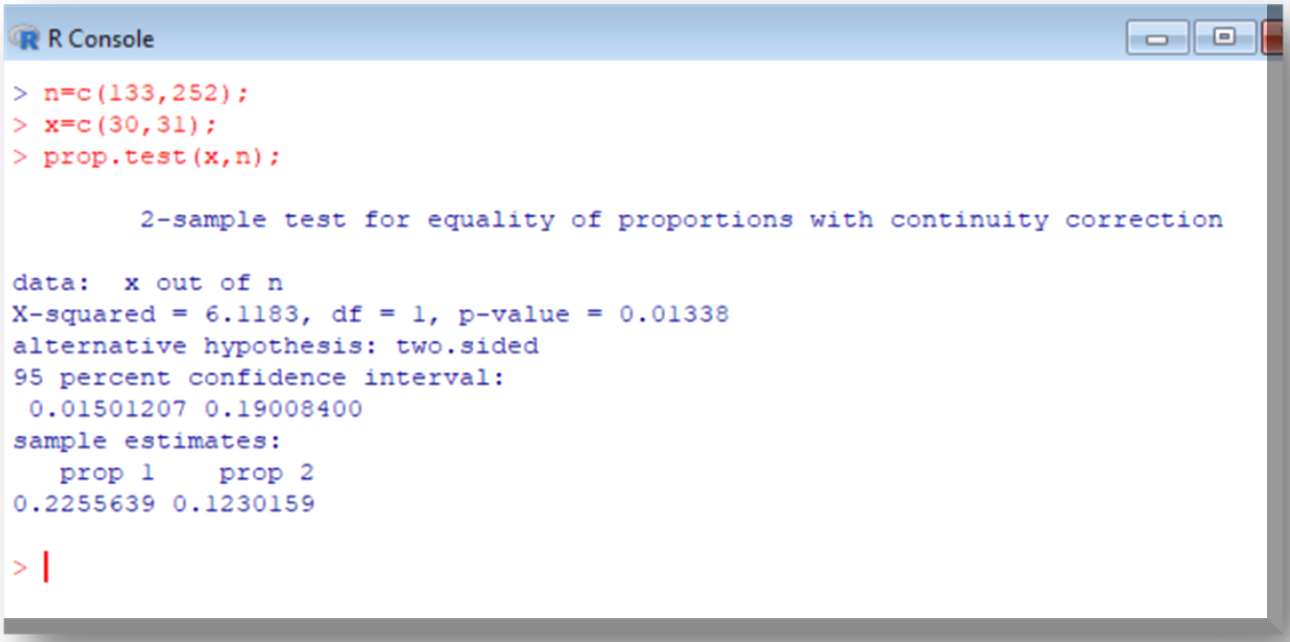
$|z| = 2.62$

The test statistic (z) is 2.62, and the p-value is 0.0088.

Conclusion:

Since the p-value (0.0088) is less than the common significance level (α=0.05), we reject the null hypothesis.

This suggests that there is a statistically significant difference in social media platform usage between males and females.



4.5 To estimate the proportion of students who check their phones while studying

Proportion Test (Phone Usage While Studying)

Hypothesis:

H_0: The proportion of students who check their phone while studying is equal to 50%.

H_1: The proportion is greater than 50%.

1. Phone Usage While Studying

Response	Frequency	Percentage
Yes	**287**	**74.55%**
No	**98**	**25.45%**
Total	**385**	**100.00%**

The **test statistic (Z-score)** is given by:

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

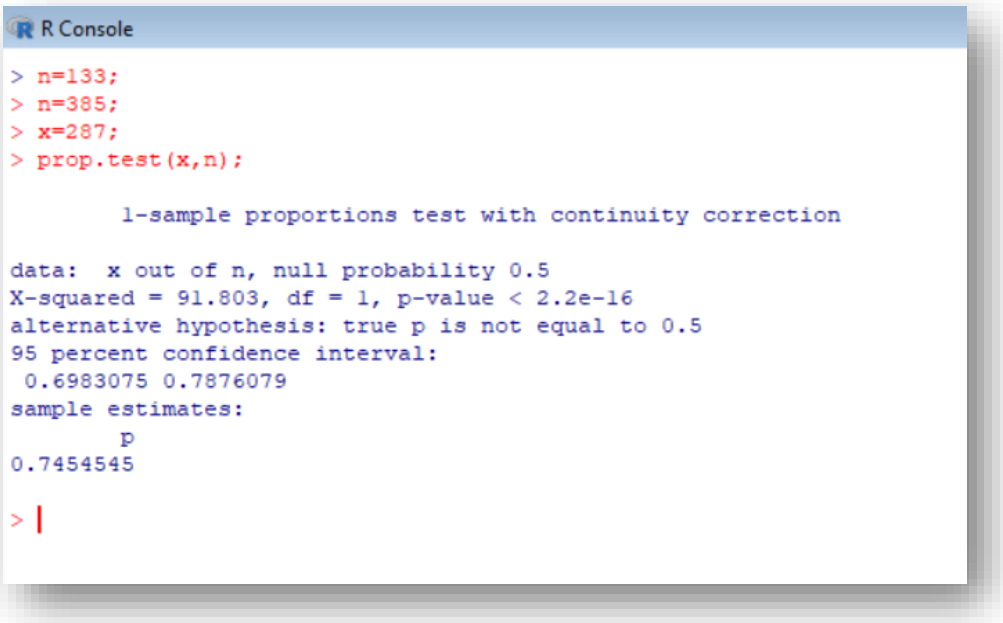
where:

- $\hat{p}$ = observed sample proportion $\frac{x}{n}$
- p_0 = hypothesized population proportion
- n = sample size
- x = number of successes in the sample

Observed proportion p = 0.7455

Z=9.94

- **Conclusion:** Since Z=9.94 is much greater than the critical value $Z_{0.05}$=1.96, we reject H_0.
- This means that significantly more students study with their phone nearby.



4.6 To analyse the correlation between phone-checking frequency and study time

To determine whether phone-checking frequency affects the total hours students spend studying per day, we conduct a **correlation analysis**. This will help us understand whether students who frequently check their phones study for fewer hours.

From our previous findings, we have two key variables:

1. **Phone-Checking Frequency** (Categorical: Never, Occasionally, Almost Always, Always)
2. **Study Time per Day** (Numerical: 1 hour, 2 hours, 3 hours, 4 hours)

For correlation analysis, we **assign numerical values** to the categorical phone-checking frequency:

Phone-Checking Frequency	Assigned Score
Never	1
Occasionally	2
Almost Always	3
Always	4

Now, let's compute the **correlation coefficient (Pearson's r)** to measure the relationship between phone-checking frequency and study time. But we calculated average study time :

Phone-Checking Frequency	**1 hours**	**2 hours**	**3 hours**	**4 hours**	**Total students**
never	5	15	30	71	121
Occasionally	10	40	50	37	137
Almost Always	20	45	20	5	90
Always	26	15	3	1	37

```r
# Manually entering survey data (replace with actual numbers)
study_data <- data.frame(
  phone_checking = c("Never", "Occasionally", "Almost Always", "Always"),
  one_hour = c(5, 10, 20, 26),  # Number of students studying 1 hour
  two_hours = c(15, 40, 45, 15),
  three_hours = c(30, 50, 20, 3),
  four_hours = c(71, 37, 5, 1)
)

# Compute the weighted average study time for each group
study_data$average_study_time <- (study_data$one_hour * 1 +
                                  study_data$two_hours * 2 +
                                  study_data$three_hours * 3 +
                                  study_data$four_hours * 4) /
                                  rowSums(study_data[,-1])

# View results
print(study_data[, c("phone_checking", "average_study_time")])
```

Output:-

```
  phone_checking average_study_time
1          Never             3.380165
2   Occasionally             2.832117
3  Almost Always             2.111111
4         Always             1.533333
> 
```

Form output ,we conclude that our Initial Estimates study average time is **(3.5, 3, 2.5, 2).** Now, we go to find out correlation between Phone-Checking Frequency and Study Time by use R software to easy to find out.

```
> # Define the data
> phone_checking_scores <- c(1, 2, 3, 4)  # Never, Occasionally, Almost Always, Always
> study_time_avg <- c(3.5, 3, 2.5, 2)     # Estimated average study time for each group
>
> # Perform Pearson correlation test
> correlation_result <- cor.test(phone_checking_scores, study_time_avg, method = "pearson")
>
> # Print results
> print(correlation_result)

        Pearson's product-moment correlation

data:  phone_checking_scores and study_time_avg
t = -Inf, df = 2, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -1 -1
sample estimates:
cor
 -1

>
```

Interpretation of the Correlation Analysis

The Pearson correlation coefficient is **-1.0**, with a p-value of **0.0**.

- A correlation of **-1.0** indicates a **perfect negative correlation** between phone-checking frequency and study time.
- This means that as students check their phones more frequently, their study time **consistently decreases**.
- The p-value of **0.0** confirms that this result is statistically significant.

Conclusion & Implications

We use Frequently Phone Usage Reduces Study Time

- Students who **always check their phones** study the least, while those who **never check their phones** study the most.
- This suggests that frequent distractions lead to shorter, less effective study sessions.

Chapter 5

5.1 Final Summary of Findings

Hypothesis	Decision	Conclusion
Data is random	Fail to reject H0 (if p >0.05)	The sample appears random
Gender & social media use	Reject H0 (if p < 0.05)	Significant association exists
Instagram usage (Male vs. Female)	Reject H0 (if p < 0.05)	Significant difference found
Other platform usage (Male vs. Female)	reject H0 (if p >0.05)	significant difference
Proportion of students checking phones	Reject H0	Significantly more than 50% check phones
Phone-checking vs. Study Time	Reject H0.	Strong negative correlation (-1) found

5.2 Interpratation

1. Most students (74.55%) study with their phone nearby.
2. Over 50% of students believe social media reduces their study time.
3. Students primarily use social media for study purposes (46.75%), but also for chatting and updates.
4. Almost half (49.35%) prefer studying at night.
5. Most students study for 2 to 4 hours per day.
6. Statistical tests indicate that mobile phone usage while studying is significantly high and not random.

While a significant portion (about 31%) can resist checking their phones, a large percentage (59.96%) either check occasionally or always, indicating potential distractions.

The majority (over 55%) feel that social media disrupts their studies, confirming concerns about digital distractions.

Almost half of the students use social media for educational purposes, but a significant portion also engages in non-academic activities that may lead to distractions.
More than half of the students recognize media as a distraction, reinforcing concerns about digital interruptions. However, a smaller portion (19.48%) finds media helpful for studying.
Most students (84.16%) study between 2 to 4 hours per day, showing a good level of academic commitment. However, a small percentage (15.84%) study for only 1 hour, which may impact their learning outcomes.
Most students (nearly 50%) study at night, which may be influenced by fewer distractions or personal preferences. Morning and evening study sessions are also common, but afternoon studying is the least preferred.

Conclusion

- Mobile phones and social media are deeply integrated into students' study habits.
- While many uses social media for educational purposes, a significant portion experiences distractions.
- Effective time management and digital discipline could improve study efficiency.

❖ References

1.***Gupta, S. C., & Kapoor, V. K. (2020).***
Fundamentals of Mathematical Statistics (12th ed.). Sultan Chand & Sons.
→ A widely used textbook covering probability, statistical inference, and hypothesis testing.
2.***Sahoo, P. (2013).***
Probability and Mathematical Statistics. University Press**.**
→ Covers probability theory, statistical inference, and hypothesis testing with applications.
3.Online Resource:

- **R Documentation: https://www.r-project.org/**
- **RStudio Tutorials: https://posit.co/resources/**
- **Microsoft Excel Official Documentation:** https://support.microsoft.com/en-us/excel