A PROJECT REPORT

ON

ASSESSING THE INFLUENCE OF INTERNET USAGE PATTERNS ON ACADEMIC PERFORMANCE OF UNIVERSITY STUDENTS: A STATISTICAL STUDY UTILIZING PRIMARY DATA, MS EXCEL, SPSS AND R PROGRAMMING

SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENT

FOR THE AWARD OF THE DEGREE OF

Bachelor of Science

In

Statistics

BY

Mohammad Sarim	21STBSA201
Mohmmad Ariz	21STBSA228
Mohammad Zaid	21STBSA246
Navjot Singh	21STBSA249
Asif Rustam	21STBSA276
Zaemuddin	21STBSA304
Nasir Javed	21STBSA368
Md Zeeshan Ahmad	21STBSA447
Aamir Suhail	21STBSA456
Abdullah Khizr	21STBSA458
Faiz	21STBSA468
Mohammad Mustafa	21STBSA469
Mohammad Umair	21STBSA477

Under the supervision Of

Prof. Aquil Ahmed & Dr. Aijaz Ahmad Dar

Department Of Statistics and Operations Research

Aligarh Muslim University

2023-24

DEPARTMENT OF STATISTICS AND OPERATIONS RESEARCH

ALIGARH MUSLIM UNIVERSITY,

ALIGARH (202002) UP, INDIA

Prof. Aquil Ahmed

Dr. Aijaz Ahmad Dar

CERTIFICATE

This is to certify that the project titled "**ASSESSING THE INFLUENCE OF INTERNET USAGE PATTERNS ON ACADEMIC PERFORMANCE OF UNIVERSITY STUDENTS: A STATISTICAL STUDY UTILIZING PRIMARY DATA, MS EXCEL, SPSS AND R PROGRAMMING**" has been submitted by the following group of students of Bachelors of Science (Statistics), Final Year, 6th Semester, 2023-24, under our supervision for the course "**STB6S1-Project**".

Name	Roll No.
Mohammad Sarim	21STBSA201
Mohmmad Ariz	21STBSA228
Mohammad Zaid	21STBSA246
Navjot Singh	21STBSA249
Asif Rustam	21STBSA276
Zaemuddin	21STBSA304
Nasir Javed	21STBSA368
Md Zeeshan Ahmad	21STBSA447
Aamir Suhail	21STBSA456
Abdullah Khizr	21STBSA458
Faiz	21STBSA468
Mohammad Mustafa	21STBSA469
Mohammad Umair	21STBSA477

PROF. AQUIL AHMED **DR. AIJAZ AHMAD DAR**

ACKNOWLEDGEMENT

In the Name of **Almighty**, the Most Gracious, the Most Merciful.

All praise and thanks to **Almighty**, the Lord of the Worlds, for His blessings and guidance throughout the completion of this project. We thank Him for His infinite mercy and grace.

We would also like to express our deep gratitude to our supervisors, **Prof. Aquil Ahmed** (Professor) & **Dr. Aijaz Ahmad Dar** (Assistant Professor), Department of Statistics & Operations Research, Aligarh Muslim University, Aligarh for their constant guidance, valuable insights, and support throughout the project. Their expert advice and encouragement have been invaluable to our success.

We also extend our heartfelt thanks to the Chairman of the Department for providing us with all the necessary resources and facilities to carry out our project work.

Furthermore, we would like to thank our parents, siblings, and friends for their unwavering support and encouragement throughout this journey. Their love and encouragement were the driving forces behind our success.

We would express our sincere appreciation to all those who have contributed to this project and helped us in one way or another. May the **Almighty** reward them abundantly.

Mohammad Zaid

Md Zeeshan Ahmad

Nasir Javed

Aamir Suhail

Abdullah Khizr

Navjot Singh

Mohammad Sarim

Mohammad Mustafa

Asif Rustam

Mohammad Umair

Faiz

Zaemuddin

Mohmmad Ariz

CONTENTS

CHAPTER 1

INTRODUCTION

[1]

1.1. Internet

The internet has become an integral part of modern life, profoundly influencing how we communicate, access information, and perform everyday tasks. Since its inception, the internet has transformed from a niche tool for researchers and academics into a global network that connects billions of people around the world. Here, we explore the general usage of the internet, its benefits, and its potential impact on various aspects of life, including education.

General Usage of the Internet

i) The internet serves as a vast repository of information, offering users instant access to a wealth of knowledge and resources. It is used for a wide range of activities, including:

ii) Communication: Email, instant messaging, and social media platforms like Facebook, Twitter, and Instagram have revolutionized the way people stay in touch with friends, family, and colleagues, enabling real-time communication regardless of geographical barriers.

iii) Information Retrieval: Search engines like Google and Bing allow users to quickly find information on virtually any topic. Online encyclopedias, news websites, and digital libraries provide in-depth knowledge and up-to-date information on current events and scholarly topics.

iv) Entertainment: The internet offers endless entertainment options, from streaming services like Netflix and YouTube to online gaming platforms and music streaming services such as Spotify. These platforms provide a wide array of content, catering to diverse interests and preferences.

v) E-commerce: Online shopping has grown exponentially, with platforms like Amazon, eBay, and Alibaba allowing consumers to purchase goods and services

[^] The references of this chapter are provided on page 87 from the comfort of their homes. This has not only changed consumer behavior but also expanded the reach of businesses.

vi) Social Interaction: Social media and online communities enable users to share experiences, join interest-based groups, and participate in discussions. These platforms can foster a sense of community and provide support networks for individuals with shared interests.

vii) Education and Learning: The internet has significantly impacted education by providing access to online courses, educational videos, and academic resources. Platforms like Coursera, Khan Academy, and edX offer courses from top universities, making education more accessible and flexible.

The Internet in Education

- The educational landscape has been particularly transformed by the internet. Students and educators use online tools for a variety of purposes:

- Research and Study: Access to academic journals, e-books, and other resources allows students to conduct research and complete assignments more efficiently. Online databases and digital libraries provide a wealth of academic material.

- Online Learning Platforms: Virtual classrooms and e-learning platforms enable students to take courses from anywhere in the world. This flexibility supports lifelong learning and professional development.

- Collaboration Tools: Tools such as Google Docs, Slack, and Microsoft Teams facilitate collaboration among students and between students and teachers. These tools enable group projects, real-time feedback, and seamless communication.

- Educational Resources: Websites like Khan Academy, YouTube, and Wikipedia offer educational content that can supplement traditional learning. Students can find tutorials, lectures, and explanatory videos on a wide range of subjects.

- While the internet offers numerous benefits, it also presents challenges such as the potential for distraction, cyberbullying, and misinformation. Therefore, it is crucial for users, especially students, to develop good internet usage habits and critical thinking skills to navigate the digital world effectively.

In summary, the internet's general usage encompasses a broad spectrum of activities that have become essential to daily life. Its role in education is particularly significant, offering unprecedented access to information and learning opportunities. Understanding these usage patterns is key to assessing their impact on academic performance, which is the focus of this study.

1.2. Internet Usage

Students spending less time studying and more time on internet are two trends that all colleges and universities will have to confront. Lowering academic standards by rewarding minimum effort and achievement (expecting less) is certainly a short-term strategy, but one that will have negative long-term consequences. A more productive way to handle these concerns is to conduct empirical research to determine to what extent these trends will negatively impact the academic performance of students and use the findings from these studies to improve our academic programs. The influence that personal variables, such as motivation and ability, have on academic success is well documented, but there is a paucity of research investigating the influence that time students spend on various activities such as studying outside of class and working has on their academic success. One reason for a lack of research in this area may be the common belief among most students and academicians that more time spent studying outside of class positively influences academic performance and that more time spent working negatively influences academic performance. Another, more plausible reason for this lack of research may be the complex nature of these relationships when evaluated in the presence of other variables, such as student ability and motivation. For example, it is likely that time spent studying outside of class will have a differential impact on the academic performance of students who vary in ability. That is, the relationship that ability has with student performance will be stronger for those students who spend more time outside of class studying than for students who spend less time studying.

1.3. Types of Internet Usage

In today's digital age, internet usage among students can vary widely, and understanding these different types of usage is crucial for our study on the impact of internet usage on academic performance.

1.3.1. Social Media Engagement

Social media engagement is a prevalent form of internet usage among students, involving platforms like Facebook, Instagram, and Twitter. It serves as a primary means of social interaction and communication, but its extensive use can lead to distractions from academic responsibilities.

Excessive social media use may impact productivity and contribute to procrastination, potentially affecting students' study habits and overall academic performance. Understanding these dynamics is crucial for exploring how internet usage patterns, including social media engagement, influence students' study time and educational outcomes.

1.3.2. Entertainment and Media Consumption

Students also use the internet for entertainment purposes, such as streaming videos, music, or playing online games. These activities can be engaging and entertaining but may also consume a considerable amount of time that could otherwise be spent on academic pursuits.

1.3.3. Academic Research and Learning

On the other hand, internet usage for academic purposes is also prevalent. This includes conducting research online, accessing educational resources, participating in online courses, and communicating with professors or classmates for academic discussions. These activities can contribute positively to academic performance if managed efficiently.

1.3.4. Online Distractions

Apart from academic and entertainment purposes, students may also face online distractions such as excessive browsing, irrelevant content consumption, or falling into unproductive internet use patterns. These distractions can potentially impact their focus and study habits.

1.3.5. Balancing Internet Usage

The challenge lies in balancing these different types of internet usage. Our study aims to explore how the hours spent on different internet activities, including social media, entertainment, academic research, and distractions, correlate with the time students dedicate to studying outside of class and, ultimately, their academic performance.

By categorizing and understanding these various types of internet usage, we can delve deeper into the factors influencing students' online behaviour and its consequences on their academic outcomes. This knowledge will guide our regression analysis to determine the specific impacts of internet usage patterns on students' study time and academic performance.

1.4. Academic Performance and its importance:

Academic Performance and Its Measurement

Academic performance is a key indicator of a student's success and potential in the educational environment. It is commonly measured through various metrics, with Grade Point Average (GPA) being one of the most widely used. GPA represents an average of a student's grades over a specific period, typically a semester or an academic year, and provides a standardized measure of academic achievement.

Importance of Academic Performance

Academic Progress and Opportunities: A strong academic record can open doors to advanced study opportunities, scholarships, and academic recognitions. It often determines eligibility for honors programs, research positions, and other academic advancements.

Career Prospects: Many employers consider academic performance during the hiring process. A high GPA can enhance a student's resume, indicating diligence, competence, and a strong work ethic. It is particularly important for fields that require advanced technical knowledge and skills.

Personal Development: Achieving good grades fosters a sense of accomplishment and boosts self-confidence. It encourages the development of essential skills such as critical thinking, time management, and problem-solving, which are valuable beyond the academic context.

Graduate and Professional School Admissions: For students aiming to pursue further education, such as graduate or professional school, academic performance is a critical criterion. Admissions committees often use GPA as a key measure to evaluate applicants' readiness for advanced study.

Factors Influencing Academic Performance

Study Habits: Effective study routines, including consistent review, active participation in class, and utilization of academic resources, contribute significantly to academic success.

Time Management: Balancing study time with other activities is essential. Procrastination and poor time management can lead to stress and lower academic performance.

Internet Usage: As discussed in this study, the way students use the internet—whether for academic purposes or distractions—can greatly affect their academic outcomes.

External Support: Support from family, teachers, and peers can provide motivation and help students overcome academic challenges.

Health and Well-being: Physical and mental health are vital for academic success. Issues such as sleep deprivation, anxiety, and poor nutrition can negatively impact performance.

So, Academic performance, primarily measured through GPA, is a critical aspect of a student's educational journey. Understanding the factors that influence academic performance, including the role of internet usage, is essential for developing strategies to enhance student success. This section highlights the importance of GPA as a measure and the various factors that can affect it, setting the stage for the subsequent analysis in this study.

1.5. Academic Performance Metrics

Academic performance metrics are essential tools used to evaluate and quantify a student's learning and academic achievements. These metrics provide a standardized way to assess and compare the educational progress and performance of students across different subjects, courses, and institutions. Key performance metrics include the Grade Point Average (GPA), standardized test scores, class rankings, and qualitative assessments. Understanding these metrics is crucial for students, educators, and institutions to track progress, identify areas for improvement, and make informed decisions about academic policies and strategies.

Grade Point Average (GPA)

One of the most widely used metrics for measuring academic performance is the Grade Point Average (GPA). The GPA is a numerical representation of a student's average performance across all their courses over a specific period, typically a semester or an academic year.

How GPA is Calculated

Assigning Grade Values: Each letter grade earned in a course is converted into a numerical value. Commonly, an 'A' might correspond to a 4.0, a 'B' to a 3.0, a 'C' to a 2.0, a 'D' to a 1.0, and an 'F' to a 0.0. Some grading systems include plus and minus variations to provide a more nuanced assessment (e.g., B+ might be 3.3, and A- might be 3.7).

Credit Hours: Each course is assigned a certain number of credit hours based on its importance and the amount of work required. Core courses often have more credit hours than electives.

Calculating the GPA: The GPA is calculated by multiplying the numerical grade value by the credit hours for each course, summing these products, and then dividing by the total number of credit hours. The formula is as follows:

$$GPA = \frac{\sum(Grade\ value \times Credit\ Hours)}{\sum(Credit\ Hours)}$$

This calculation provides a weighted average that accurately reflects the student's overall academic performance.

Standardized Test Scores

Standardized tests are another crucial metric for assessing academic performance. These tests are designed to evaluate a student's knowledge and skills in specific subject areas and provide a benchmark for comparison across different student populations. Examples of standardized tests include the SAT, ACT, GRE, and various state-specific assessments.

Class Rankings

Class rankings compare a student's academic performance to their peers within the same institution. This metric is often used in high schools and colleges to identify the top-performing students. Rankings are usually based on GPA, but they can also consider other factors such as participation in honors or advanced placement courses.

Qualitative Assessments

Qualitative assessments provide a more comprehensive view of a student's academic performance beyond numerical scores. These assessments include teacher evaluations, peer reviews, and self-assessments. They consider aspects such as class participation, critical thinking, creativity, and leadership skills.

Attendance and Participation

Regular attendance and active participation in class are significant indicators of academic performance. Students who consistently attend classes and engage in discussions tend to perform better academically. Attendance records and participation grades are often used to monitor and encourage student engagement.

Assignments and Projects

Assignments and projects are practical tools for evaluating a student's understanding and application of course material. They require students to demonstrate their knowledge, conduct research, and develop problem-solving skills. The quality and timely submission of assignments and projects contribute to the overall assessment of academic performance.

Extracurricular Activities

While not a direct measure of academic performance, involvement in extracurricular activities can positively impact a student's academic success. Participation in sports, clubs, and other activities fosters skills such as teamwork, time management, and leadership, which can enhance overall academic performance.

Comprehensive Exams and Thesis Work

For graduate and postgraduate students, comprehensive exams and thesis work are critical performance metrics. These assessments evaluate a student's depth of knowledge in their field of study and their ability to conduct independent research.

Successful completion of these requirements is often necessary for graduation and professional advancement.

Peer and Self-Assessment

Peer and self-assessment practices involve students evaluating their own and each other’s work. These assessments help develop critical thinking and self-reflection skills. They also provide insights into a student's learning process and areas that may need improvement.

Importance of Academic Performance Metrics

i. Tracking Progress: They help students and educators track academic progress over time, identifying strengths and areas for improvement.

ii. Informed Decision-Making: Metrics provide valuable data for making informed decisions about curriculum development, teaching methods, and academic policies.

iii. Accountability: They hold students accountable for their learning and performance, encouraging them to strive for excellence.

iv. Benchmarking and Comparison: Metrics allow for benchmarking and comparison across different educational institutions, ensuring consistent standards of education.

v. Motivation and Recognition: High academic performance can motivate students to continue excelling and provide recognition through awards, scholarships, and honors.

vi. Career and Educational Opportunities: Strong academic performance opens doors to higher education and career opportunities, enhancing a student’s future prospects

Academic performance metrics encompass a range of quantitative and qualitative measures that together provide a comprehensive view of a student's academic achievements. From GPA and standardized test scores to class rankings and qualitative assessments, these metrics are essential for evaluating and enhancing educational outcomes. Understanding and effectively utilizing these metrics can lead to improved teaching strategies, better student engagement, and higher overall academic success.

1.6. Possible Factors Influencing Academic Performance

Academic performance is influenced by various factors that can affect how well students do in school. Understanding these factors is crucial for our study on how internet usage and study time relate to students' academic success.

One important factor is individual study habits. Students who have effective study routines, manage their time well, and stay organized tend to perform better academically. Our study will explore how the time students dedicate to studying outside of class interacts with their internet usage habits and impacts their grades.

Another factor is the level of motivation and interest in learning. Students who are motivated and interested in their studies often put in more effort, leading to better academic outcomes. We will consider how internet usage patterns might influence students' motivation levels and engagement in academic tasks.

Additionally, socioeconomic factors can play a role in academic performance. Factors such as access to resources, family support, and educational opportunities can impact students' abilities to excel in school. By examining how internet usage habits vary across different socioeconomic backgrounds, we can better understand the broader context of academic success.

This regression analysis will focus into these factors and their interactions with internet usage and study time, providing insights into how various aspects of students' lives contribute to their overall academic performance. This knowledge can inform strategies to support students in achieving better educational outcomes in the digital age.

1.7. Research Objective

The research objective can be comprehended and elucidated by breaking down its components

i. Characterize Internet Usage Patterns:

- Identify the types of online activities students engage in, including academic-related tasks (e.g., research, accessing educational resources) and non-academic activities (e.g., social networking, entertainment).
- Quantify the frequency, duration, and intensity of students' internet use for different purposes and across various devices (e.g., computers, smartphones, tablets).

ii. Examine Study Time Allocation:

- Analyse how students allocate their study time across different subjects and academic tasks.
- Investigate variations in study time based on personal factors (e.g., academic interests, learning preferences) and external influences (e.g., workload, extracurricular commitments).

Iii. Evaluate Academic Performance Metrics:

- Measure students' academic performance using objective metrics such as grades, exam scores, and overall academic achievement.
- Explore variations in academic outcomes based on students' study habits, time management strategies, and patterns of internet usage.

iv. Investigate Relationships and Correlations:

- Explore the relationships between internet usage, study time, and academic performance through statistical analysis (e.g., correlation coefficients, regression modeling).
- Examine potential moderating or mediating factors that influence these relationships (e.g., student motivation, technological proficiency).

v. Identify Factors Influencing Academic Success:

- Identify key factors that contribute to positive academic outcomes among students, considering the interplay between internet usage behaviors, study habits, and performance indicators.

- Determine whether specific patterns of internet use or study time allocation are associated with higher or lower academic achievement.

1.8. Significance of Study

The significance of this study Assessing the influence of internet usage patterns on academic performance of University students is multifaceted and crucial for several reasons like

i) Informing Educational Practices: Understanding how students utilize the internet and allocate their study time can inform educators and policymakers about effective teaching and learning strategies. By identifying patterns of internet usage that correlate with improved academic performance, educators can tailor instructional methods and curriculum delivery to leverage digital resources more effectively.

ii) Optimizing Student Time Management: Examining study time allocation provides insights into students' time management skills and habits. This knowledge can help students optimize their study schedules, prioritize tasks, and strike a balance between online activities and academic responsibilities, ultimately enhancing productivity and academic success.

iii) Addressing Digital Distractions: The research sheds light on the impact of non-academic internet use (e.g., social media, entertainment) on academic performance. By understanding how digital distractions affect students' focus and study habits, interventions can be developed to mitigate distractions and promote a more conducive learning environment.

iv) Promoting Digital Literacy and Responsibility: Analyzing internet usage patterns encourages discussions around digital literacy and responsible online behavior. Educators and parents can use research findings to educate students about using the internet as a valuable educational tool while emphasizing the importance of responsible online conduct.

v) Enhancing Student Support Services: Research outcomes can guide the development of support services for students, such as time management workshops, study skills training, and digital literacy programs. Institutions can tailor interventions based on empirical evidence to address specific challenges faced by students in managing their academic workload and online activities.

vi) Contributing to Policy Development: Findings from this research can contribute to the formulation of educational policies related to technology integration and digital learning. Policymakers can use evidence-based insights to advocate for resources that support effective internet use in educational settings and promote equitable access to digital tools for all students.

vii) Adapting to Evolving Educational Landscapes: As technology continues to shape education, studying internet usage and study habits among students provides timely information on adapting educational practices to evolving digital landscapes. This research helps educators stay abreast of trends in student behavior and preferences, facilitating continuous improvement in teaching methodologies.

In summary, the significance of this study Assessing the influence of internet usage patterns on academic performance of University students lies in its potential to optimize educational practices, support student development, and inform policies that foster responsible and effective use of technology in education. This research contributes to creating a more student-centered and digitally inclusive learning environment that promotes academic excellence and holistic student success.

CHAPTER 2
RESEARCH METHODOLOGY

2.1. Objective

The main goal of this analysis is to study how the time students spend on the internet and studying affects their academic performance. With the internet becoming a big part of education, there are concerns about how it might distract students and affect their grades. This study plans to use regression analysis and data from different students to understand these relationships better.

We'll look at how internet use, study time, and grades are connected for students from different backgrounds. By using statistical methods, like looking at grades, test scores, and overall academic performance, we hope to learn more about what helps students succeed in school.

We also want to go beyond just finding connections and dive into why these connections exist. Factors like how much access students have to technology, their self-discipline, and their learning styles can all influence how internet use affects their study habits and grades. By considering these different factors, we aim to get a full picture of what affects student success in today's digital learning environment.

This research aims to share its findings with educators, policymakers, and others involved in education. By understanding how internet use and study habits relate to academic performance, we can work towards better strategies to help students do well in school in the digital age.

2.2. Target Population[2]

The target population for this study is university students, a demographic that offers a unique and valuable perspective on the relationship between internet usage and academic performance. University students, typically ranging from 18 to 25 years old, are in a critical phase of their academic and personal development. This age group is not only immersed in the academic demands of higher education but also heavily engaged with digital technology. The ubiquity of the internet in their daily lives for both academic and recreational purposes makes them an ideal group for this study. University students use the internet extensively for accessing educational resources, conducting research, participating in online discussions, and completing assignments. Simultaneously, they also use the internet for social media, entertainment, and communication, which can significantly impact their time management and study habits.

[2] The References of this chapter are provided on page 87

The academic environment of universities provides a structured setting where academic performance can be objectively measured through grades, GPAs, and other standardized assessments, allowing for precise quantification of academic outcomes. Moreover, the diversity within the university student population, encompassing various disciplines, year groups, socio-economic backgrounds, and cultural contexts, ensures a comprehensive and nuanced understanding of the impact of internet usage. This diversity is crucial as it allows the study to explore how different factors such as field of study, year of enrollment, and personal circumstances influence the relationship between internet usage and academic performance.

Furthermore, university students are at a stage where they are developing self-regulation and time management skills, making them particularly susceptible to the potential distractions posed by excessive internet usage. The insights gained from studying this population can be invaluable for developing targeted interventions and support systems to enhance academic performance and promote healthy internet habits. Understanding how university students allocate their time between online activities and academic responsibilities can inform policies and practices at educational institutions aimed at fostering a balanced and productive academic environment. By focusing on university students, this study aims to contribute to a deeper understanding of the complex dynamics between digital engagement and educational outcomes, providing actionable recommendations to support student success in the digital age.

To effectively target the university student population for this study, a structured approach encompassing various strategies will be employed. The target population will primarily consist of university students aged between 18 and 25, currently enrolled in academic programs. The geographic scope may encompass a single university, multiple universities within a region, or a broader national or international representation, depending on the study's objectives and resources. Stratified sampling methods will be utilized to ensure diversity within the sample, considering factors such as year of study, major field of study, and socio-economic background. Recruitment efforts will involve collaboration with university administrations to access student databases and official communication channels for recruitment purposes. Additionally, online surveys distributed through university email lists and social media platforms, along with in-person recruitment strategies such as setting up booths in high-traffic areas and classroom announcements, will be employed to maximize participant recruitment. Ethical considerations, including obtaining informed consent and ensuring confidentiality, will be strictly adhered to throughout the data collection process. By employing a comprehensive approach to target the university student population, this study aims to gather diverse and representative data to analyze the relationship between internet usage and academic performance effectively.

2.3. Sample Size

Statistical In this project, We aimed to assess the influence of internet usage patterns on the academic performance of university students. We collected data from 215 respondents. However, due to certain constraints, such as limited time and the occurrence of exams, we were unable to gather a larger sample size.

Constraints Faced

Time Limitations:

The data collection period was constrained by a tight schedule, leaving us with limited time to reach out to more respondents.

Academic Exams:

During the data collection phase, university exams were ongoing, making it challenging for students to participate in the survey. This significantly reduced our ability to gather more responses.

Despite these challenges, a sample size of 215 can still provide meaningful insights into the relationship between internet usage and academic performance. However, a larger sample size is generally preferable to increase the reliability and generalizability of the findings.

Ideal Sample Size and Its Calculation

The ideal sample size for a study depends on various factors, including the desired level of confidence, the margin of error, the population size, and the expected variability in the data. A commonly used formula to calculate the sample size for a given confidence level and margin of error is:

$$n = \frac{z^2 p(1-p)}{e^2}$$

Where:

n is the required sample size.

Z is the Z-value (the number of standard deviations from the mean) corresponding to the desired confidence level. For a 95% confidence level,

Z is 1.96.

p is the estimated proportion of the population that has the attribute of interest. If unknown,

p is often assumed to be 0.5 for maximum variability.

e is the margin of error (the desired precision of the estimate).

For example, if we want to calculate the sample size with a 95% confidence level and a 5% margin of error, assuming

p = 0.5

then,

n = 384.16

Therefore, the ideal sample size would be approximately 384 respondents. This would provide a more precise estimate of the population parameters with a 95% confidence level and a 5% margin of error.

While our study was conducted with a sample size of 215 due to practical constraints, ideally, a larger sample size of around 384 would be preferable to enhance the accuracy and generalizability of the results. Future studies should aim to collect data from a larger number of respondents to achieve more robust conclusions about the impact of internet usage on academic performance.

2.4. Data Set

For this project, we collected primary data from university students to assess the influence of internet usage patterns on their academic performance. The dataset comprises responses from 215 students and includes various variables: email address, age, gender, field of study, level of course, hours spent on the internet for non-academic purposes, perceived correlation between internet usage and academic performance, usage of specific online resources for studying, impact of internet usage on study effectiveness, ability to manage online time efficiently, hours spent studying outside of class per week, procrastination tendencies due to internet use, preferred study methods, overall study habits, current grade point average (GPA), the impact of internet usage on grades, strategies for balancing internet usage and study time, overall academic performance, importance of attendance, and educational board background. This diverse set of variables was chosen to comprehensively explore the multifaceted relationship between internet usage and academic performance, providing insights into how different aspects of internet use can affect students' educational outcomes. The inclusion of both continuous variables (such as GPA and hours spent studying) and categorical variables (such as gender and field of study) allows for a robust analysis using various statistical methods, enhancing the validity and depth of the study's findings.

2.4.1. General overview of variables

Age of Respondents: This variable represents the age of each participant in the study. Age can provide insights into the demographic composition of the sample population and may be relevant for analysing correlations with other variables such as academic performance or internet usage habits.

Gender of Respondent: This variable categorizes participants based on their gender identity, typically recorded as male or female. Gender may influence various aspects of academic performance and internet usage patterns, making it an important demographic variable to consider in the analysis.

Field of Study: This variable identifies the academic discipline or field in which each respondent is enrolled. Field of study can influence study habits, resource utilization, and academic performance, making it a key factor in understanding the relationship between internet usage and academic outcomes.

Level of Course: This variable indicates the educational level or stage of the course in which participants are currently enrolled, such as undergraduate or graduate level. The level of course may impact study workload, time management, and academic performance, influencing the variables under investigation.

Hours Spent on the Internet for Non-academic Purposes: This variable quantifies the amount of time each respondent typically spends on the internet for activities unrelated to academic purposes, such as social media or entertainment. It serves as a measure of internet usage behaviour outside of educational contexts.

Have you noticed any correlation between the amount of time spent on the internet and your academic performance?: This variable captures respondents' perceptions of any observed correlation between their internet usage habits and academic performance. It provides insights into their self-awareness regarding the impact of internet usage on academic outcomes.

Do you use any specific online resources or platforms for studying? : This variable identifies the online resources or platforms utilized by respondents for academic purposes, such as educational websites or online learning platforms. It offers insights into the diversity of digital learning tools used by students.

Do you believe there are specific times of the day when internet usage has a more significant impact on your study effectiveness?: This variable explores respondents' perceptions of the timing of internet usage and its influence on study effectiveness. It sheds light on the potential relationship between internet usage patterns and study habits.

How would you rate your ability to manage your online time efficiently?: This variable assesses respondents' self-perceived ability to manage their internet usage effectively. It provides insights into their level of digital self-regulation and time management skills.

How many hours per week do you typically dedicate to studying outside of class?: This variable quantifies the amount of time allocated by respondents for studying outside of formal class hours. It reflects their study workload and commitment to academic pursuits.

Do you ever find yourself putting off studying because you spend too much time on the internet?: This variable examines respondents' tendencies to procrastinate studying due to excessive internet usage. It highlights potential challenges in balancing internet activities and academic responsibilities.

What are your preferred study methods?: This variable identifies the study techniques or methods favored by respondents for learning and academic preparation. It offers insights into their learning preferences and strategies.

How would you rate your overall Study Habits?: This variable allows respondents to evaluate their overall study habits on a rating scale. It provides a self-assessment of their study behaviors and practices.

What is your current grade point average (GPA)?: This variable records respondents' current GPA, serving as a measure of their academic performance. GPA reflects their cumulative achievement across academic courses.

Have your grades ever been negatively affected by your internet usage?: This variable explores whether respondents perceive a negative impact of internet usage on their academic grades. It provides insights into their awareness of potential challenges posed by excessive internet use.

Do you believe your internet usage has a positive or negative impact on your academic performance?: This variable captures respondents' beliefs regarding the impact of internet usage on their academic performance. It reflects their attitudes and perceptions towards the relationship between digital technology and learning outcomes.

Are there any strategies you employ to balance your internet usage and study time effectively?: This variable identifies the strategies or techniques employed by respondents to manage and balance their internet usage and study time. It offers insights into adaptive coping mechanisms and self-regulatory behaviors.

How would you rate your overall Academic Performance?: This variable allows respondents to assess their overall academic performance on a rating scale. It provides a subjective evaluation of their achievements and success in academic endeavors.

Is there anything else you would like to share about your internet usage, study habits, or academic performance?: This open-ended variable allows respondents to provide additional comments or insights related to their internet usage, study habits, or academic performance. It offers a platform for qualitative feedback and reflections.

Do you think Attendance is important?: This variable explores respondents' beliefs regarding the importance of attendance in academic settings. It provides insights into their attitudes towards attendance requirements and its perceived impact on academic success.

From Which Board did you study?: This variable identifies the educational board or system from which respondents completed their previous education. It offers insights into their educational background and institutional affiliations.

2.5. Statistical Methods

Statistical methods are a collection of techniques used for analysing, interpreting, and presenting data. Statistical methods involve using mathematical techniques to analyse data. In a project examining the effect of internet usage and study time on students' academic performance, these methods can be crucial. They allow us to perform regression analysis, which helps in understanding and quantifying the relationship between these variables. By applying statistical methods, we can predict how changes in internet usage and study time may influence academic outcomes, thus providing valuable insights for improving educational strategies

2.5.1. Descriptive Statistics

Descriptive statistics play a crucial role in the initial analysis of our dataset, providing a summary of the main features of the data. For this project, descriptive statistics help us understand the general patterns and distributions of the variables we collected from the 215 university students. By summarizing data through measures such as mean, median, mode, standard deviation, and frequency distributions, we can gain insights into the central tendencies and variability of our dataset.

For example, examining the mean and standard deviation of students' GPA allows us to understand the average academic performance and how much individual GPAs deviate from the average. Similarly, frequency distributions for categorical variables such as gender, field of study, and level of course provide a clear picture of the demographic composition of our sample. This information is essential for identifying any potential biases in our sample and ensuring that it is representative of the broader student population.

Descriptive statistics are also important for checking data quality and preparing for further statistical analyses. They help identify outliers, missing values, and data entry errors that might affect the validity of our results. Furthermore, understanding the distribution of our variables guides the selection of appropriate statistical tests and models. For instance, if we find that a variable is not normally distributed, we may opt for non-parametric tests or transform the data to meet the assumptions of parametric tests.

Descriptive statistics provide a foundational understanding of our data, highlighting key characteristics and guiding subsequent analyses. They ensure that we have a clear and accurate picture of the dataset before diving into more complex inferential statistics and modelling, ultimately enhancing the reliability and interpretability of our research findings.

2.5.2. Graphical visualization

Graphical visualization is an attractive method of showcasing numerical data that help in analysing and representing quantitative data visually. A graph is a kind of a chart where data are plotted as variables across the coordinate. It became easy to analyse the extent of change of one variable based on the change of other variables. Graphical representation of data is done through different mediums such as lines, plots, diagrams, etc.

Definition of Graphical Representation of Data

A graphical representation is a visual representation of data statistics-based results using graphs, plots, and charts. This kind of representation is more effective in understanding and comparing data than seen in a tabular form. Graphical representation helps to qualify, sort, and present data in a method that is simple to understand for a larger audience. Graphs enable in studying the cause-and-effect relationship between two variables through both time series and frequency distribution. The data that is obtained from different surveying is infused into a graphical representation by the use of some symbols, such as lines on a line graph, bars on a bar chart, or slices of a pie chart. This visual representation helps in clarity, comparison, and understanding of numerical data.

Representation of Data

The word data is from the Latin word Datum, which means something given. The numerical figures collected through a survey are called data and can be represented in two forms - tabular form and visual form through graphs. Once the data is collected through constant observations, it is arranged, summarized, and classified to finally represented in the form of a graph. There are two kinds of data - quantitative and qualitative. Quantitative data is more structured, continuous, and discrete with statistical data whereas qualitative is unstructured where the data cannot be analysed.

Uses of Graphical Representation of Data

The main use of a graphical representation of data is understanding and identifying the trends and patterns of the data. It helps in analysing large quantities, comparing two or more data, making predictions, and building a firm decision. The visual display of data also helps in avoiding confusion and overlapping of any information. Graphs like line graphs and bar graphs, display two or more data clearly for easy comparison. This is important in communicating our findings to others and our understanding and analysis of the data.

Types of Graphical Representation of Data

Data is represented in different types of graphs such as plots, bar graph, histogram, scatter plot, correlation matrix. etc. They are as follows:

- Bar Graph

  A bar chart, also known as a bar graph, is a graphical representation of data using rectangular bars. The length or height of each bar is proportional to the value it represents, and the bars can be oriented either horizontally or vertically. The x-axis represents the categories being compared, while the y-axis represents the scale or value of the data being measured. Bar charts are commonly used to compare the values of different categories or groups, as well as to display trends over time. They are often used in business, finance, economics, and other fields to visualize data and make it easier to understand. Bar charts can be created using various software programs, including Microsoft Excel, Google Sheets, and other data visualization tools. They can be customized with different colours, fonts, labels, and other design elements to make the data more visually appealing and easier to interpret.

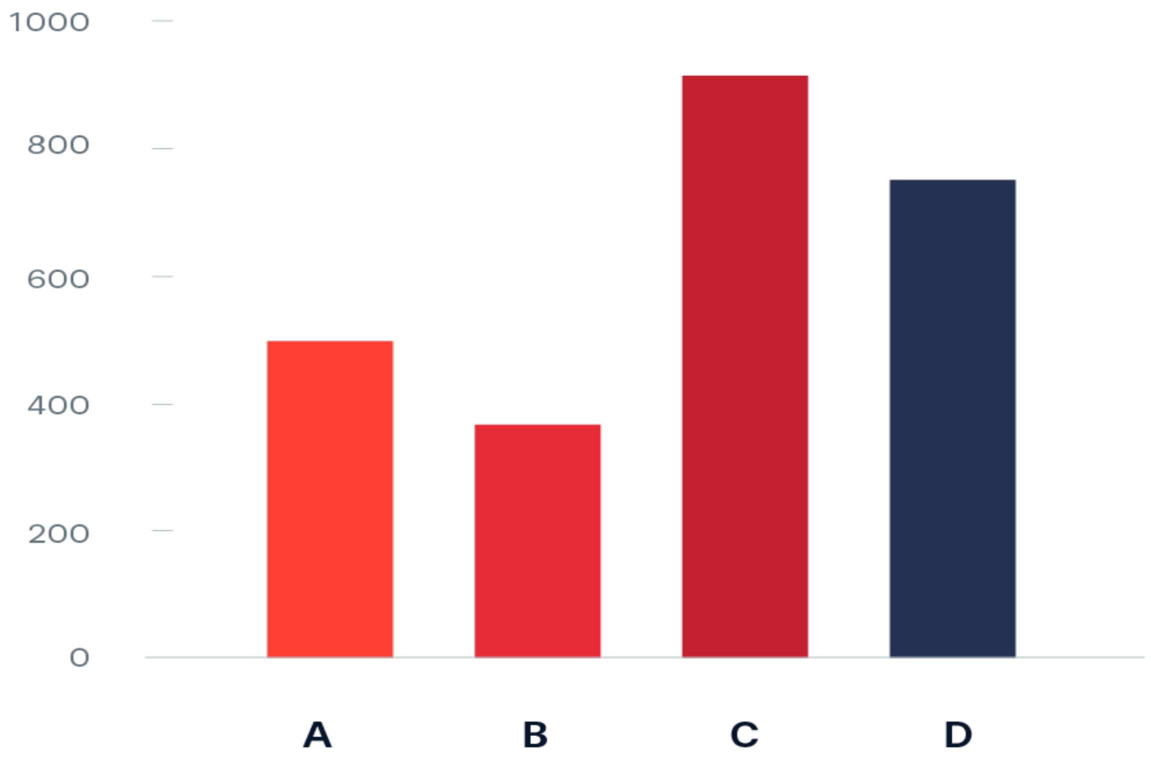


Fig.1: Bar Graph Visualisation

- Histogram

A histogram is a graphical representation of the distribution of numerical data. It consists of a series of rectangles or bins that are aligned next to each other along an axis. The length of each rectangle or bin corresponds to the frequency or proportion of data that falls within its range. The bins are typically evenly spaced, although they can be of varying widths if the data is not evenly distributed. Histograms are commonly used to visualize the shape of a dataset, including the range, central tendency, and variability of the data. They can help identify patterns and outliers in the data, as well as any potential gaps or overlaps. Histograms are also useful for checking the assumptions of statistical models, such as normality and symmetry. The x-axis of a histogram represents the range of values for the data, and the y-axis represents the frequency or proportion of data that falls within each bin. The bins can be adjusted to be narrower or wider, depending on the granularity of the data and the desired level of detail in the visualization. Histograms are easy to create and interpret and can provide valuable insights into the distribution of the data.

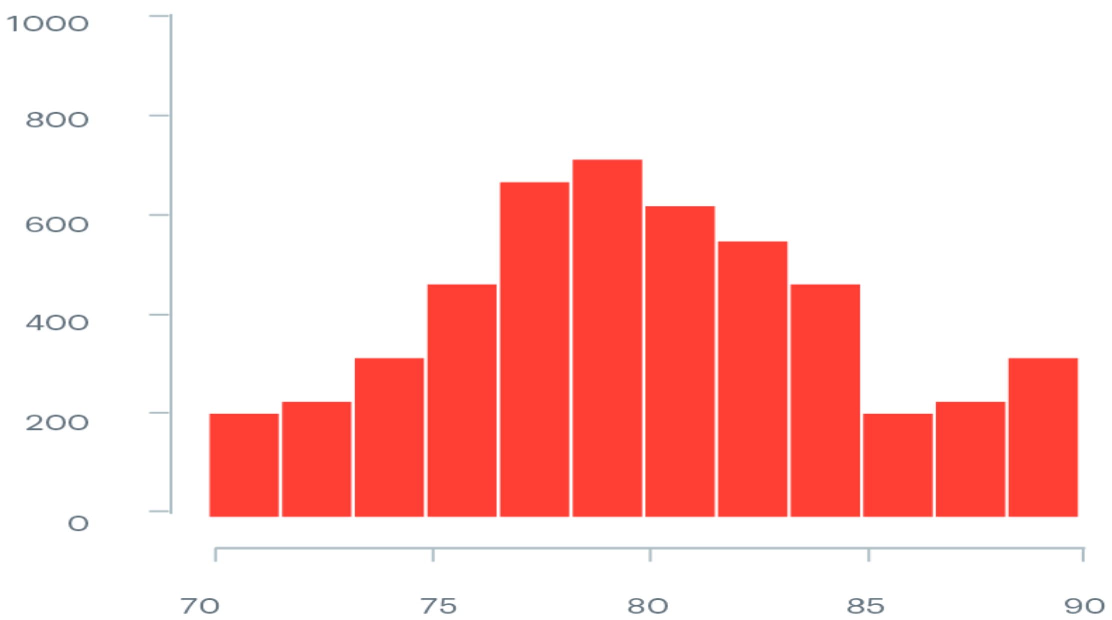
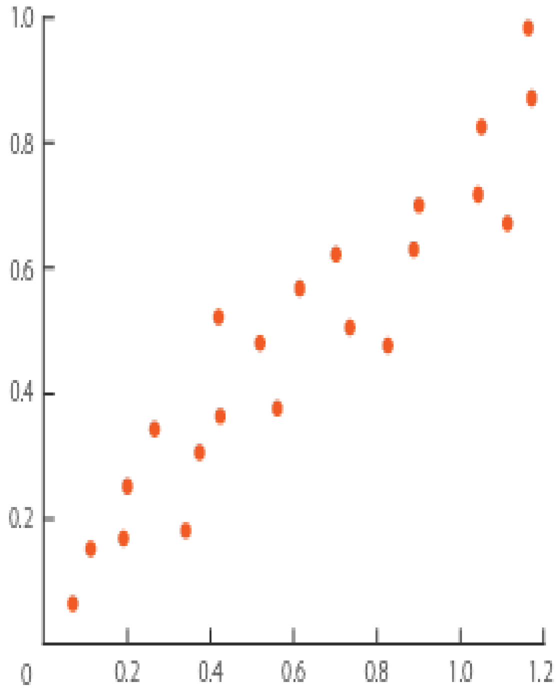


Fig.2: Histogram Visualisation

- Scatter Plot

  Scatter Plots are described as one of the most useful inventions in statistical graphs. Originally, the scatter plot was presented by an English Scientist, John Frederick W. Herschel, in the year 1833. Herschel used it in the study of the orbit of the double stars. He plotted the positional angle of the double star in relation to the year of measurement. The scatter plot was used to understand the fundamental relationship between the two measurements. Even though bar charts and line plots are frequently used, the scatter plot still dominates the scientific and business world. It is very easy for people to look at points on a scale and understand their relationship. A scatter plot (also called a scatter graph, scatter chart, scattergram, or scatter diagram) is a type of plot or mathematical diagram using Cartesian coordinates to display values for typically two variables in a set of data. A scatter plot can be used either when one continuous variable is under the control of the experimenter and the other depends on it or when both continuous variables are independent.



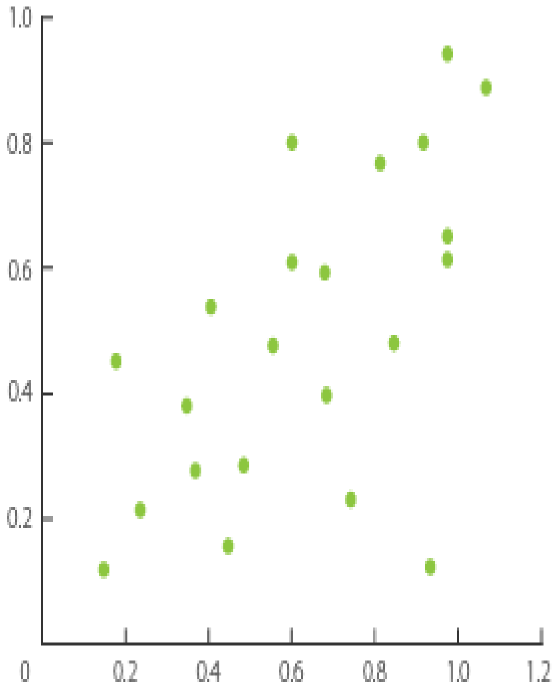


Fig.3: Scatter Plot Visualisation

- Correlation Matrix

The correlation matrix is a matrix that shows the correlation between variables. It gives the correlation between all the possible pairs of values in a matrix format. We can use a correlation matrix to summarize a large data set and to identify patterns and make a decision according to it. We can also see which variable is more correlated to which variable, and we can visualize our results. A correlation matrix involves a rows and columns table that shows the variables. Every cell in a matrix contains the correlation coefficient. The correlation matrix is in conjunction with other types of statistical analysis. It's very useful for regression techniques like simple linear regression, multiple linear regression and lasso regression models. In the regression technique, we have several independent variables, and based on that, we are predicting the dependent variables A correlation matrix is a statistical technique used to evaluate the relationship between two variables in a data set. The matrix is a table in which every cell contains a correlation coefficient, where 1 is considered a strong relationship between variables, 0 a neutral relationship and -1 a not strong relationship. It's most commonly used in building regression models.

2.5.3. Correlation and Association

In this project, "Assessing the Influence of Internet Usage Patterns on Academic Performance of University Students," examining correlation and association is critical to understanding how different variables relate to one another. These statistical concepts help us determine whether changes in one variable might be associated with changes in another, providing insights into potential relationships and dependencies within our data.

Correlation

Correlation measures the strength and direction of a linear relationship between two continuous variables. In our project, we can use correlation analysis to explore relationships such as the one between the number of hours spent on the internet for non-academic purposes and students' GPA. The correlation coefficient (r) ranges from -1 to 1, where values close to 1 or -1 indicate a strong relationship, and values close to 0 indicate a weak relationship. A positive correlation suggests that as one variable increases, the other tends to increase, while a negative correlation suggests an inverse relationship.

For example, calculating the Pearson correlation coefficient between internet usage hours and GPA can help us identify if excessive internet use is linked to lower academic performance. This information can guide us in understanding whether and how internet usage might impact students' grades.

Association

Association analysis, on the other hand, is used to explore relationships between categorical variables. In our project, we use Chi-Square tests to assess the association between categorical variables such as the field of study and the time spent studying, or the level of course and hours spent on the internet. The Chi-Square test helps us determine whether there is a significant association between these categorical variables.

We examined the association between "Time Spent Studying" and "Field of Study" using the Chi-Square test. The results indicated whether students from different fields of study spent significantly different amounts of time studying. Similarly, we assessed the relationship between "Hours Spent on the Internet" and "Field of Study" to see if students' internet usage patterns varied by their academic disciplines.

Importance of Correlation and Association

Identifying Relationships: These analyses help identify whether and how different factors are related, which is essential for understanding the dynamics between internet usage and academic performance.

Guiding Further Analysis: The insights gained from correlation and association analyses can guide further, more detailed investigations, such as regression analysis, to model these relationships and control for other variables.

Informing Interventions: By identifying significant associations and correlations, we can inform university policies and interventions aimed at improving academic performance. For example, if a strong negative correlation is found between non-academic internet use and GPA, strategies can be developed to help students manage their internet usage better.

Correlation and association analyses provide foundational insights into the relationships between variables in our study, helping us to understand the influence of internet usage patterns on academic performance. These analyses form the basis for more complex statistical modelling and ultimately contribute to the development of effective strategies to support student success.

2.5.4. Hypothesis Testing Procedure

In this project, "Assessing the Influence of Internet Usage Patterns on Academic Performance of University Students," hypothesis testing is a crucial statistical method used to determine whether there is enough evidence to support specific claims about the data. Hypothesis testing helps us make informed conclusions about the relationships between variables such as internet usage patterns and academic performance. Here, we outline the general procedures and specific tests used in this project.

General Hypothesis Testing Procedure

Formulate Hypotheses: Begin by stating the null hypothesis (H0) and the alternative hypothesis (H1).

Null Hypothesis (H0): There is no relationship between the variables.

For example, H0: There is no relationship between time spent studying and academic performance.

Alternative Hypothesis (H1): There is a relationship between the variables.

For example, H1: There is a relationship between time spent studying and academic performance.

Select the Appropriate Test: Choose the statistical test based on the type of data and the hypotheses.

For relationships between categorical variables, we use the Chi-Square test.

For comparing means between groups, we use the t-test.

For examining correlations between continuous variables, we use the Pearson correlation test.

Set the Significance Level: Typically, the significance level (alpha) is set at 0.05. This means there is a 5% risk of concluding that a difference exists when there is no actual difference.

Calculate the Test Statistic: Using statistical software like SPSS, R, or Excel, calculate the test statistic based on the data.

Make a Decision: Compare the test statistic to a critical value or use the p-value to determine whether to reject or fail to reject the null hypothesis.

If the p-value is less than the significance level ($p < 0.05$), reject the null hypothesis.

If the p-value is greater than the significance level ($p \geq 0.05$), fail to reject the null hypothesis.

Interpret the Results: Draw conclusions based on the hypothesis test results and relate them back to the research questions and objectives.

Specific Hypothesis Tests Used in Our Project

Chi-Square Test for Independence

Objective: To determine if there is an association between categorical variables such as the field of study and time spent studying.

Example:

H0: There is no association between time spent studying and field of study.

H1: There is an association between time spent studying and field of study.

Procedure: We used SPSS to perform the Chi-Square test, entering the categorical variables and interpreting the p-value to test our hypothesis.

Independent Samples t-Test:

Objective: To compare the means of GPA between different groups, such as undergraduate (UG) and postgraduate (PG) students.

Example:

H0: There is no difference in GPA between UG and PG students.

H1: There is a difference in GPA between UG and PG students.

Procedure: We conducted the t-test in SPSS, inputting the GPA data for UG and PG groups, and examined the test results to draw conclusions.

Pearson Correlation

Objective: To assess the strength and direction of the relationship between continuous variables such as hours spent on the internet and GPA.

Example:

H0: There is no correlation between hours spent on the internet and GPA.

H1: There is a correlation between hours spent on the internet and GPA.

Procedure: Using SPSS, we calculated the Pearson correlation coefficient and interpreted the p-value to understand the relationship.

Importance of Hypothesis Testing in Our Project

Validate Relationships: Confirm or refute suspected relationships between variables, such as internet usage and academic performance.

Support Data-Driven Decisions: Provide a statistical basis for making informed decisions and recommendations.

Enhance Academic Understanding: Contribute to the academic body of knowledge by rigorously testing theoretical assumptions with empirical data.

By following these hypothesis testing procedures, we ensure that our findings are robust, reliable, and statistically significant, thereby providing valuable insights into the influence of internet usage on academic performance among university students.

2.5.5. Regression Analysis

Regression is a statistical method used to establish a relationship between a dependent variable and one or more independent variables. In simpler terms, regression analysis helps to find the relationship between a dependent variable (outcome) and one or more independent variables (predictors). Regression analysis can be used to explore and understand the relationships between variables and to predict the future values of the dependent variable based on the values of the independent variable(s). Regression analysis is often used in many fields, including economics, finance, marketing, biology, psychology, and engineering. It is used to investigate the causal relationship between variables and make predictions based on historical data. There are many types of regression models, including linear regression, logistic regression, polynomial regression, and multiple regression. The choice of model depends on the nature of the data and the research question being addressed.

Linear regression is one of the most widely used regression models. It assumes a linear relationship between the dependent variable and the independent variable(s). The goal of linear regression is to find the best-fit line that explains the relationship between the variables. Multiple regression is used when there are two or more independent variables that affect the dependent variable. The goal of multiple regression is to find the best-fit line that explains the relationship between the variables. Regression analysis is a powerful tool for understanding and predicting the relationships between variables. A linear model specifies a linear relationship between a dependent variable and independent variables.

$$\boldsymbol{y} = \boldsymbol{\beta 0} + \boldsymbol{\beta 1 * X 1} + \ldots\ldots\ldots\ldots\ldots + \boldsymbol{\beta n * X n} + \in$$

where y is the dependent variable,

Xi are independent variables

βi are parameters of the model.

I. Assumptions of Regression Analysis

i. Linearity: The relationship between the dependent variable and the independent variable(s) should be linear.
ii. Independence: The observations should be independent of each other.
iii. Homoscedasticity: The variance of the errors should be constant across the range of the independent variable(s).
iv. Normality: The errors should be normally distributed.
v. No multicollinearity: The independent variables should not be highly correlated with each other.

.

CHAPTER 3

STATISTICAL ANALYSIS

3.1. Descriptive Statistics

Descriptive analysis is the initial step in statistical analysis, aimed at summarizing and describing the characteristics of the variables under study. In the context of our research on "A Regression Analysis on the Hours Spent on the Internet and Time Spent to Study on Academic Performance of Students," descriptive analysis provides a comprehensive overview of the variables involved, including hours spent on the internet, time spent studying, and academic performance.

3.1.1. Continuous Variables

I. Hours Spent on Internet

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Variance	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
Recoded NonAcdHours	215	1	5	3.23	1.047	1.095	-.108	.166
Valid N (listwise)	215							

Range: The range indicates the difference between the minimum and maximum values of the recoded non-academic hours spent on the internet. In this case, the range is 4, suggesting that the data varies from 1 to 5 (as "NA" has been excluded from the analysis).

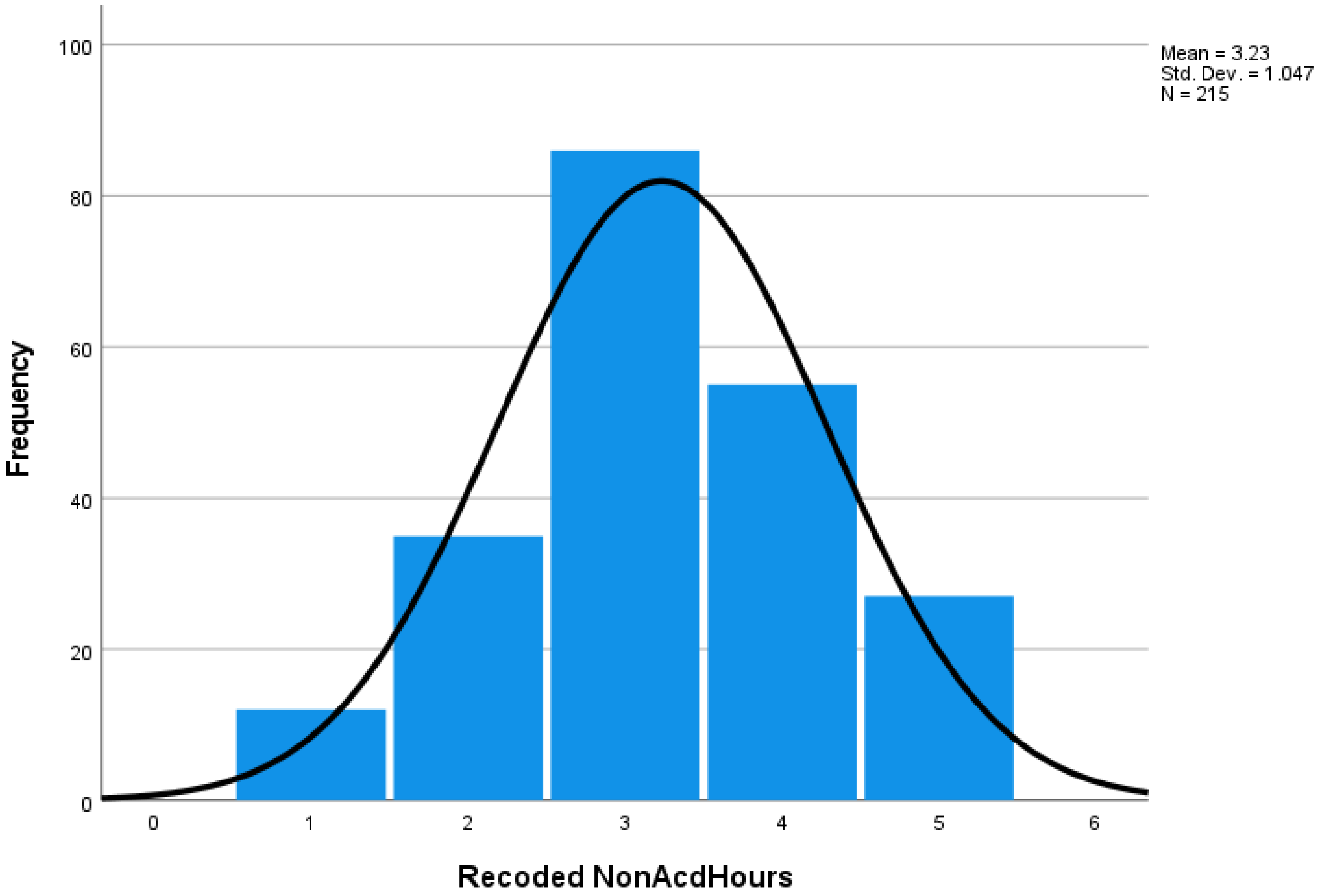
Minimum and Maximum: The minimum value represents the lowest recoded non-academic hours spent on the internet, which is 1 (corresponding to "Less than 1 hour"). The maximum value is 5 (corresponding to "More than 6 hours"), indicating the highest recoded hours.

Mean: The mean recoded value for non-academic hours spent on the internet is approximately 3.23. This suggests that, on average, respondents spend between 2 to 4 hours on the internet for non-academic purposes.

Standard Deviation and Variance: The standard deviation measures the dispersion or spread of the recoded non-academic hours around the mean. A standard deviation of approximately 1.05 indicates that the recoded hours vary moderately around the mean. The variance, which is the square of the standard deviation, provides a measure of the variability in the recoded hours.

Skewness: Skewness measures the symmetry of the distribution. A skewness value close to 0 (-0.108 in this case) suggests that the distribution is approximately symmetric.

Based on the recoded numerical values assigned to the original categories, it can be inferred that respondents, on average, spend around 2 to 4 hours on the internet for non-academic purposes. However, further analysis may be needed to explore the relationship between these recoded hours and other variables of interest in our study.



The distribution analysis graph visually represents the spread and shape of the data related to hours spent on the internet. This graphical representation provides insights into the distribution of these variables, highlighting any patterns, central tendencies, variability, and potential outliers present in the dataset. By examining the distribution of these variables, we gain a better understanding of how students allocate their time between internet usage.

The distribution analysis graph illustrates the frequency distribution of non-academic hours spent on the internet by students. The x-axis represents the categories of non-academic hours, while the y-axis indicates the frequency of students falling into each category.

The graph depicts a bell-shaped curve, indicative of a normal distribution pattern. In a normal distribution, data is symmetrically distributed around the mean, with the majority of observations clustered near the mean and fewer observations towards the tails.

The key statistics derived from the data are as follows:

Mean: 3.23 hours

Standard deviation: 1.047 hours

Sample size (N): 215 students

The normal distribution suggests that most students spend approximately 3.23 hours on the internet for non-academic purposes, with a standard deviation of 1.047 hours.

Given the approximate normal distribution pattern observed in the graph, it indicates that the distribution of non-academic internet usage among students follows a typical bell-shaped curve. This suggests that most students tend to cluster around the average internet usage time.

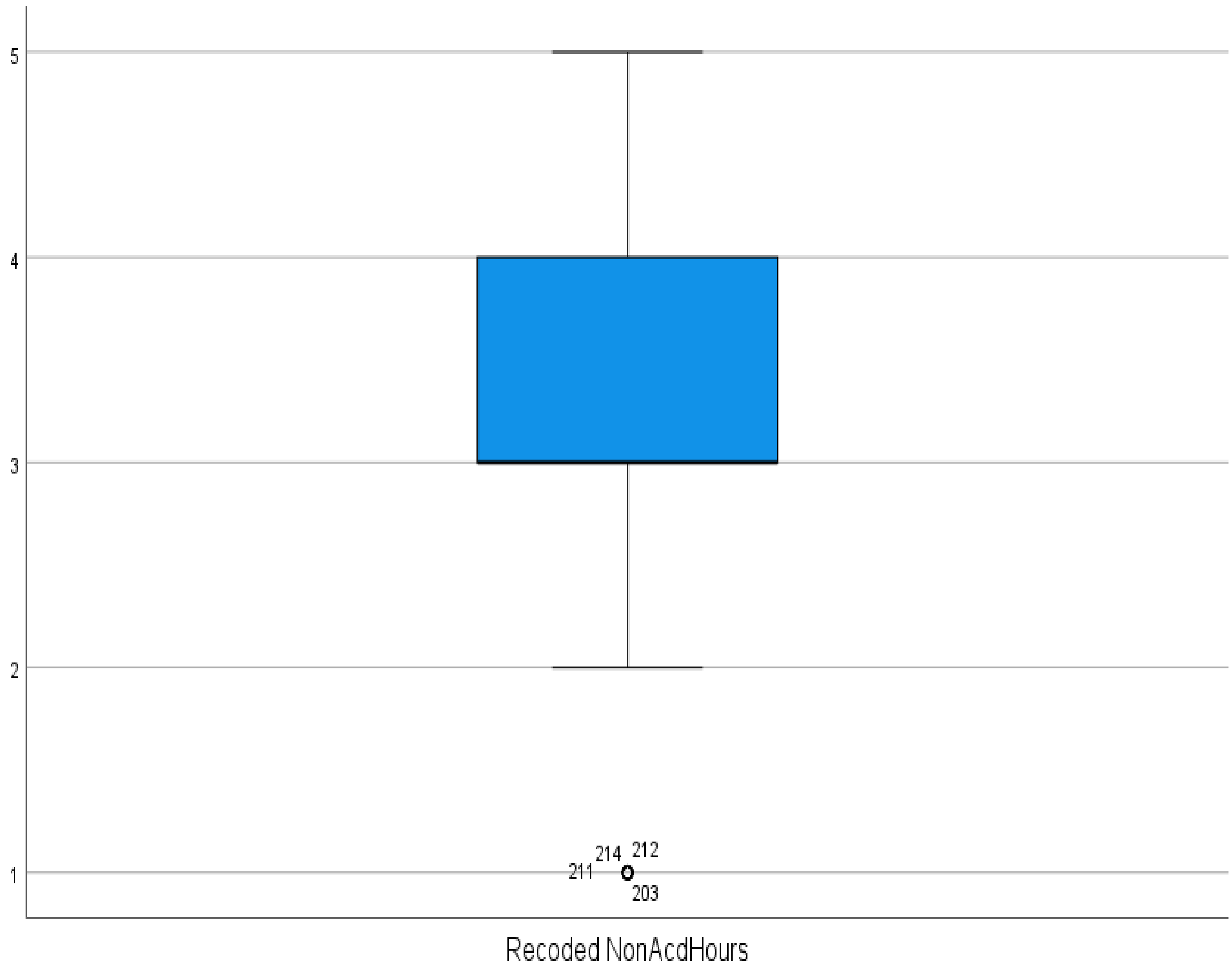
Considering the recoded categories:

Low Usage (0 & <1 hour): The combined frequency of students spending 0 hours (likely missing data) and less than 1 hour represents a smaller group. This indicates that there is a portion of students who either do not use the internet for non-academic purposes or use it very minimally.

Moderate Usage (1-4 hours): The frequencies for the 1-2 hours and 2-4 hours categories likely encompass a significant portion of students. This aligns with the central bulge of the normal distribution, indicating that this range is the most common internet usage among students.

High Usage (>6 hours): The frequency for "More than 6 hours" suggests a smaller group of students who spend a considerably higher amount of time on the internet compared to the majority. This group represents students who are heavy users of the internet for non-academic purposes.

Overall, the distribution analysis provides insights into the varying levels of non-academic internet usage among students, with the majority falling within the moderate usage range and smaller proportions in the low and high-usage categories.



Box Plot of Time Spent on Internet for non-Academic Purpose

II. Time Spent Studying

Descriptive Statistics

	N	Minimum	Maximum	Mean		Std. Deviation	Variance
	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic
OutsidecalssRecoded	208	0.00	8.00	4.9904	0.12192	1.75832	3.092
Valid N (listwise)	208						

Based on the descriptive statistics provided for the variable "Hours Spent Studying Outside of Class" with assigned numerical values, In SPSS it is done by Analyze > Descriptive Statistics > Descriptives. The descriptives are coming out to be:

Mean (Assigned Values): The average number of hours spent studying outside of class is approximately 4.99 hours per week.

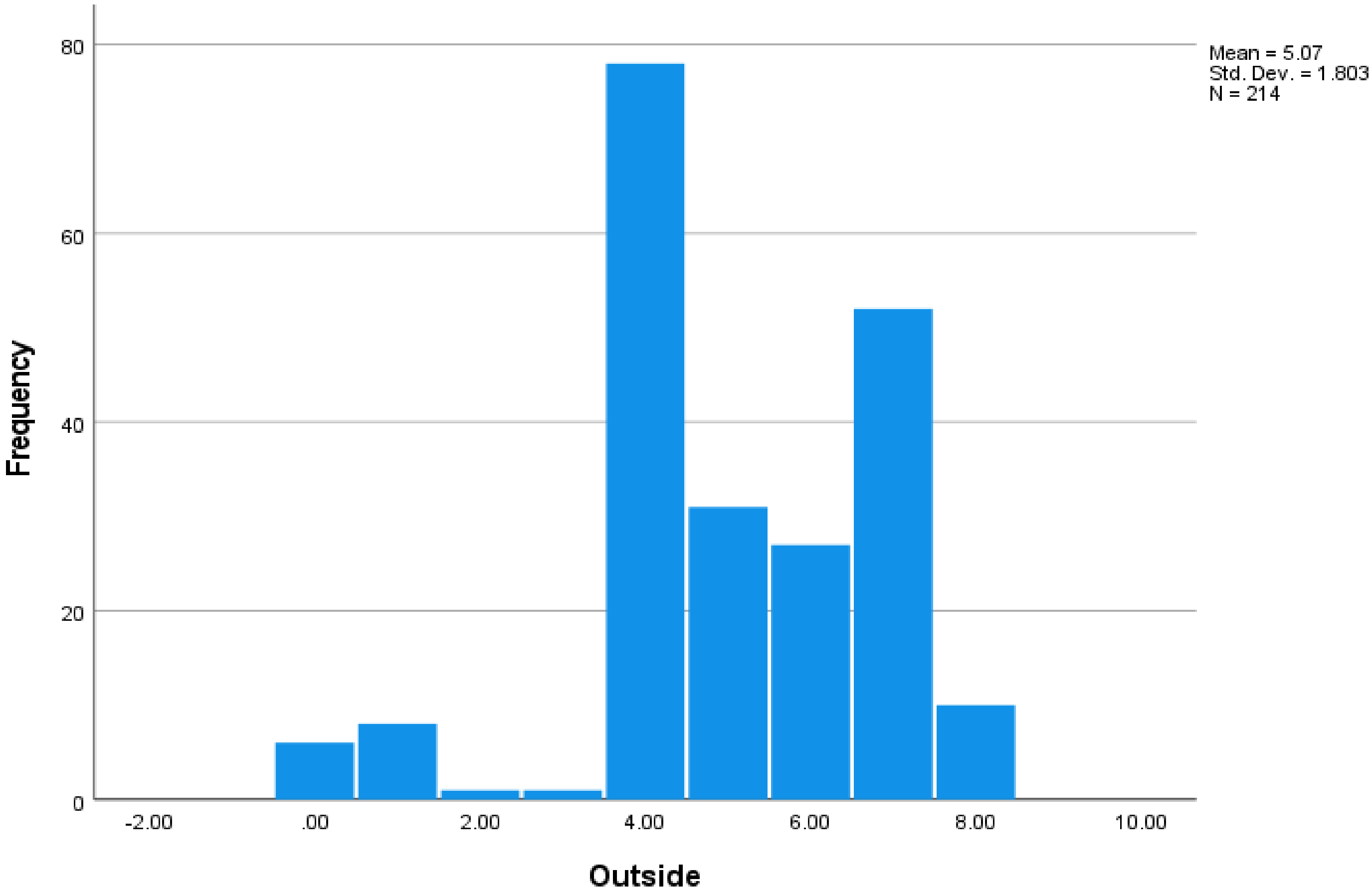
Minimum (Assigned Values): The minimum reported value (assigned as 0) indicates that some respondents reported spending no time studying outside of class.

Maximum (Assigned Values): The maximum reported value (assigned as 8) indicates that some respondents reported spending up to 8 hours studying outside of class per week.

Standard Deviation: The standard deviation measures the dispersion or variability of the data around the mean. In this case, the standard deviation is approximately 1.76, indicating that the study hours vary by about 1.76 hours from the mean.

Variance: The variance is another measure of the spread of the data. It is the square of the standard deviation and indicates how much the data values deviate from the mean squared. The variance is approximately 3.09.

These statistics, along with the assigned numerical values for different time ranges, provide insights into the distribution and variability of study hours among the respondents. The mean value indicates the average study time, while the standard deviation and variance quantify the spread of study hours around the mean. This information, along with the assigned values, can be valuable for understanding students' study habits and academic engagement outside of class.



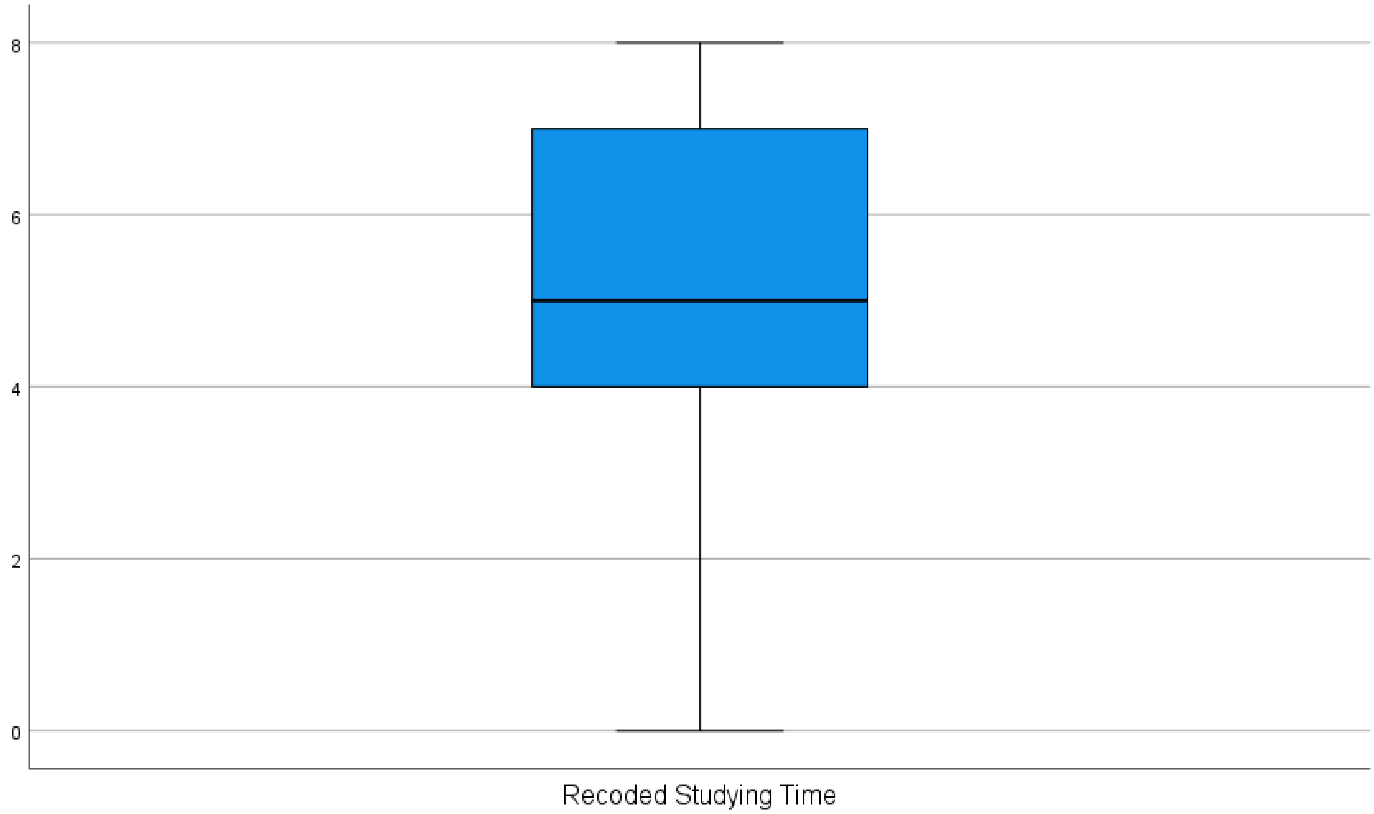
Before delving into the graph illustrating the distribution of time spent studying, it's essential to understand the pivotal role of study hours in academic performance. The allocation of time to study outside of class is a critical factor that influences a student's comprehension, retention of knowledge, and overall academic success. By visualizing the distribution of study hours, we can gain insights into how students manage their study time and identify any patterns or trends that may impact their performance.

The bar graph illustrates the distribution of time spent studying outside class by students. On the x-axis, the categories represent various intervals of hours spent studying, ranging from "1-2 hours" to "10-12 hours". The y-axis indicates the frequency, depicting the number of students falling within each study time interval.

Although the distribution exhibits slight asymmetry, it predominantly displays a bell-shaped curve, indicative of a normal distribution pattern. A normal distribution is characterized by symmetry around the mean, with the majority of data points concentrated near the mean and fewer data points as one moves away from the mean in both directions.

Key statistical measures derived from the data include a mean study time of 5.07 hours, with a standard deviation of 1.803 hours, based on a sample size of 214 students. Given the distribution's shape and statistical parameters, it is plausible to infer that the distribution of time spent studying outside class adheres to a normal distribution.

The intervals on the x-axis, ranging from "NA" to "10-12 hours", provide insights into the study habits of students across different time brackets.



Box Plot of Time Spent studying outside class

3.1.2. Categorical Variables

I. Gender

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
Recoded Gender	215	1	2	1.25	.435	1.156	.166
Valid N (listwise)	215						

Sample Size (N): The data includes responses from 215 participants.

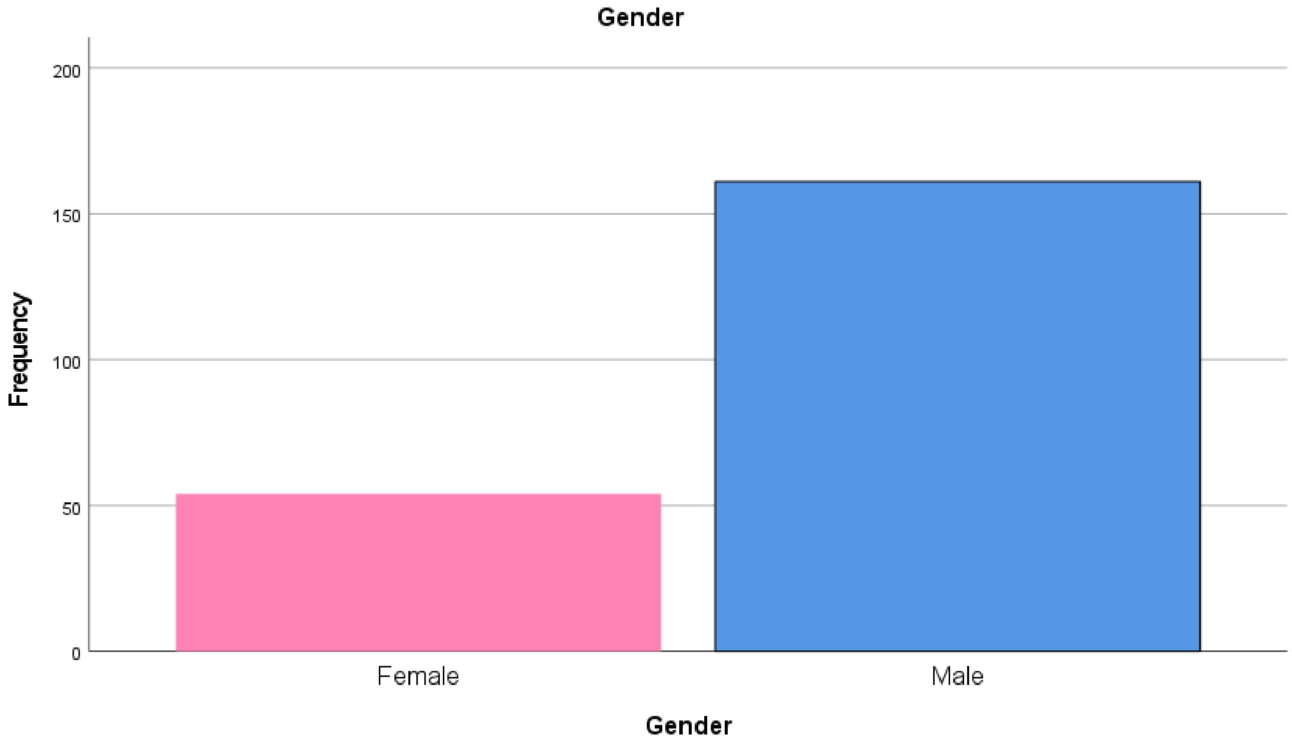
Minimum and Maximum Values: The values for the gender variable range from 1 to 2, where 1 represents "Male" and 2 represents "Female."

Mean (1.25): The average value of the gender variable is 1.25, indicating that the sample has more males than females. Given that 1 represents males and 2 represents females, a mean closer to 1 suggests a higher proportion of male respondents in the sample.

Standard Deviation (0.435): This value measures the amount of variation or dispersion of the gender values from the mean. A standard deviation of 0.435 indicates that the gender distribution is somewhat dispersed around the mean.

Skewness (1.156): The skewness value of 1.156 indicates that the distribution of the gender variable is positively skewed. This means there are more males (coded as 1) in the sample than females (coded as 2), and the distribution is not symmetrical. A positive skewness value confirms that the tail on the right side (representing females) is longer or fatter than the left side.

Standard Error of Skewness (0.166): The standard error of skewness is 0.166, which helps assess the significance of the skewness value. Given that the skewness value is considerably larger than twice the standard error (1.156 > 2 * 0.166), the positive skewness is statistically significant.



Bar Plot of Gender

II. Field of Study

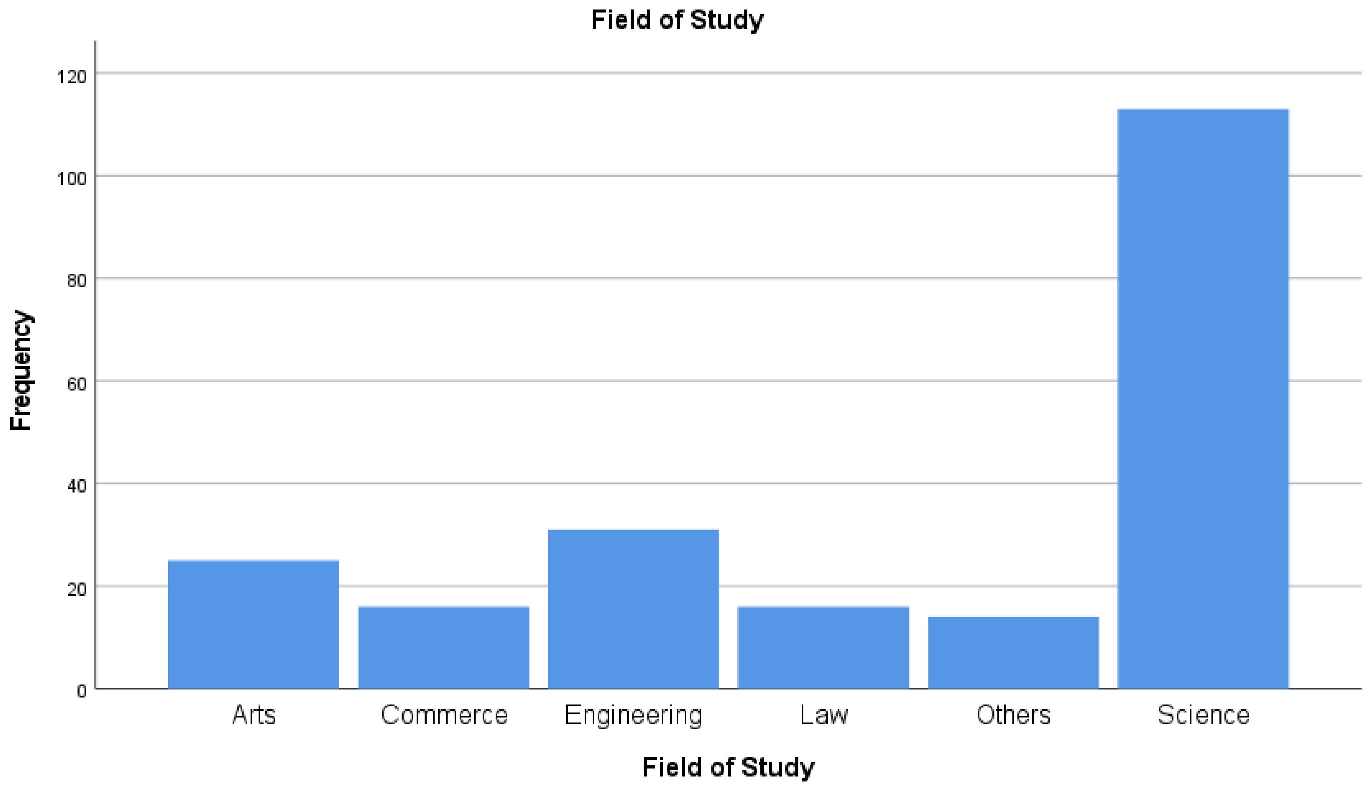
Field of Study

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Arts	25	11.6	11.6	11.6
	Commerce	16	7.4	7.4	19.1
	Engineering	31	14.4	14.4	33.5
	Law	16	7.4	7.4	40.9
	Others	14	6.5	6.5	47.4
	Science	113	52.6	52.6	100.0
	Total	215	100.0	100.0	

- Dominance of Science Students: The majority of respondents (52.6%) are from the Science field, indicating a strong representation of Science students in the sample. This could be reflective of the university's enrollment trends or the study's focus area.
- Diverse Representation: While Science dominates, there is still a varied representation across other fields such as Arts, Commerce, Engineering, Law, and Others. This diversity is beneficial for analyzing the influence of internet usage patterns on academic performance across different fields of study.

- Balanced Sample: Although there is a dominant field, the sample includes sufficient numbers of students from other fields to allow for meaningful comparisons and insights into how different fields may influence or relate to academic performance and internet usage.

This distribution provides a comprehensive overview of the study's participants, which is crucial for understanding how internet usage patterns might impact academic performance across various academic disciplines.



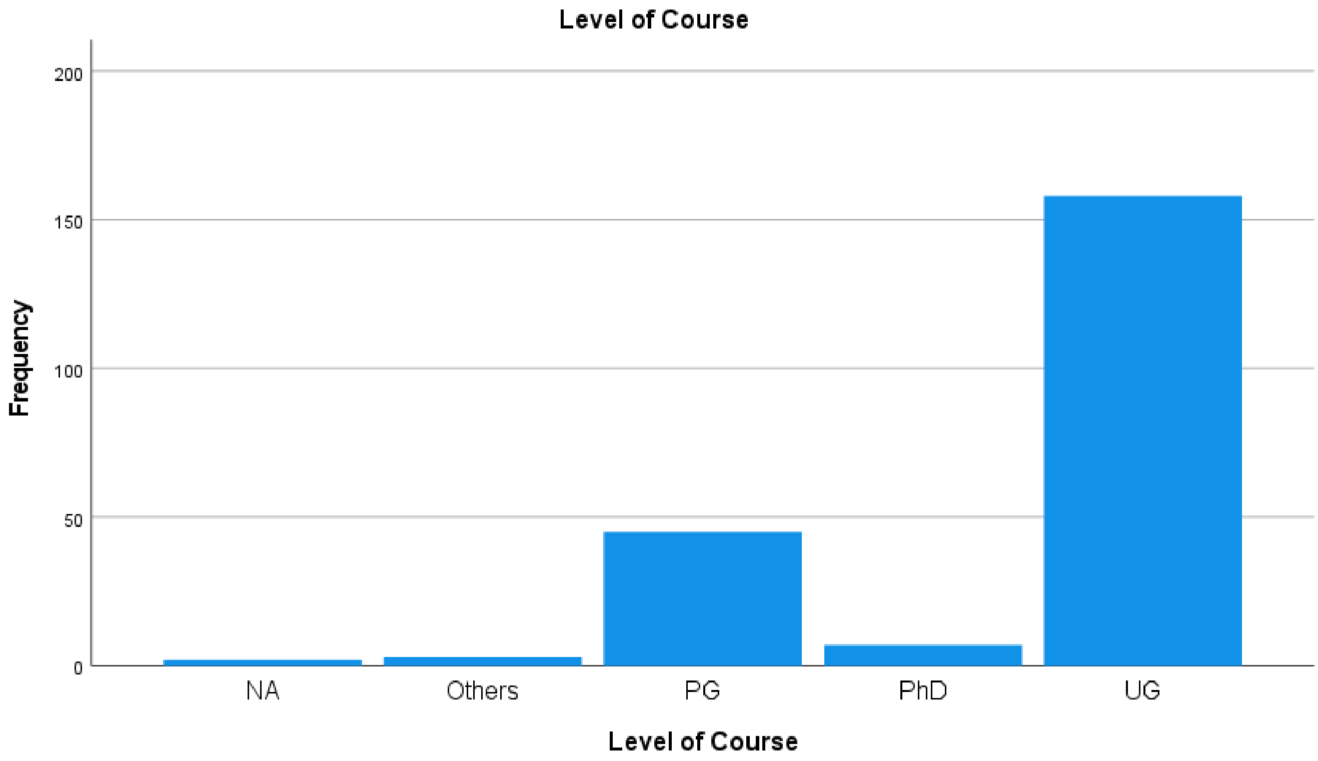
Bar Plot of Field of study

III. Level of Course

Level of Course

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	NA	2	.9	.9	.9
	Others	3	1.4	1.4	2.3
	PG	45	20.9	20.9	23.3
	PhD	7	3.3	3.3	26.5
	UG	158	73.5	73.5	100.0
	Total	215	100.0	100.0	

Frequency table of Level of Course



Bar Plot of Level of Course

IV. Internet Impact on Academic Performance of students

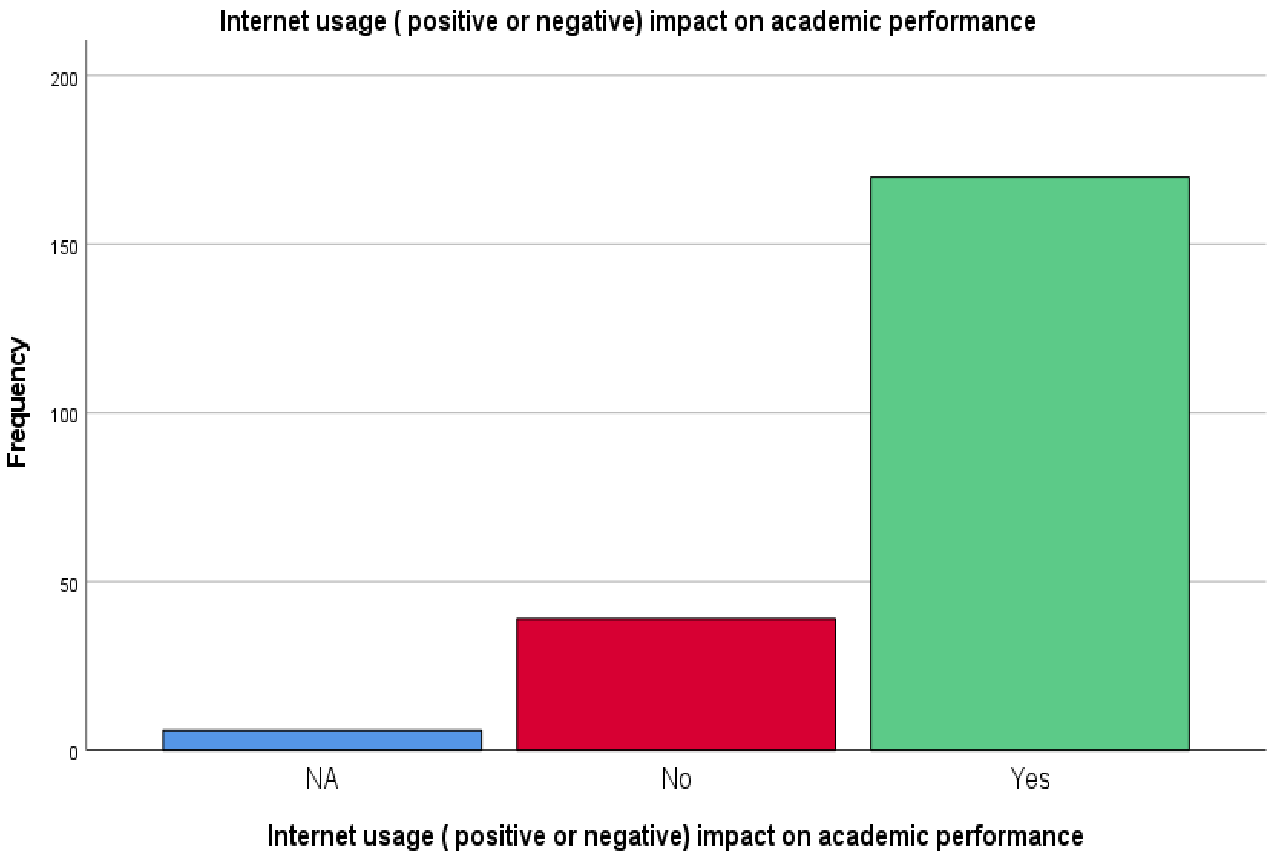
Do you believe your internet usage has a positive or negative impact on your academic performance?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	NA	6	2.8	2.8	2.8
	No	39	18.1	18.1	20.9
	Yes	170	79.1	79.1	100.0
	Total	215	100.0	100.0	

Predominant Positive Perception: A significant majority of students (79.1%) believe that their internet usage has a positive impact on their academic performance. This suggests that most students perceive the internet as a beneficial tool for their studies, possibly due to access to educational resources, online learning platforms, and other supportive tools.

Minority Negative Perception: Only 18.1% of students feel that their internet usage has a negative impact on their academic performance. This indicates that while there is a recognition of potential downsides, such as distraction or procrastination, it is not the dominant perception among the students surveyed.

Low Non-Applicability: A very small fraction of respondents (2.8%) chose "NA," indicating that the majority of students have formed a clear opinion on how internet usage affects their academic performance.



Bar plot of Internet Impact on Academic Performnace

3.2. Correlation Analysis between Hours Spent on the Internet and Time Spent Studying

Correlations

		Time Spent on Internet	Time Spent Studying
Time Spent on Internet	Pearson Correlation	1	-.153*
	Sig. (2-tailed)		.025
	N	215	214
Time Spent Studying	Pearson Correlation	-.153*	1
	Sig. (2-tailed)	.025	
	N	214	214

*. Correlation is significant at the 0.05 level (2-tailed).

Studying outside class reveals a Pearson correlation coefficient of -0.153*. The negative sign indicates an inverse relationship between the two variables, implying that as the hours spent on the internet increase, the time spent studying outside class tends to decrease, and vice versa.

The correlation coefficient of -0.153* is statistically significant at the 0.05 level (2-tailed), suggesting that this relationship is unlikely to have occurred by chance. This finding implies that there is a modest, yet statistically significant, negative correlation between hours spent on the internet and time spent studying outside class among the study participants.

In practical terms, this correlation suggests that students who spend more time on the internet for non-academic purposes tend to allocate less time to studying outside class, and conversely, those who spend less time on the internet allocate more time to studying. Understanding this relationship can help educators and policymakers develop interventions to promote effective time management and balance between internet usage and academic responsibilities among students.

3.3. Comparing Groups

In this project, we compare different groups of university students to understand the influence of internet usage patterns on academic performance. The comparisons are made using proportion tests and t-tests to examine the relationships between categorical and continuous variables.

3.3.1 Proportion Test

Descriptive analysis is the initial step in statistical analysis, aimed at summarizing and describing the characteristics of the variables under study. In the context of our research on "A Regression Analysis on the Hours Spent on the Internet and Time Spent to Study on Academic Performance of Students," descriptive analysis provides a comprehensive overview of the variables involved, including hours spent on the internet, time spent studying, and academic performance.

I. Hours Spent on the Internet v/s Academic Performance

Null Hypothesis (H0): There is no significant association between hours spent on the internet and academic performance. This means that the amount of time spent on the internet does not affect how well students perform academically.

Alternative Hypothesis (H1): There is a significant association between hours spent on the internet and academic performance. This means that the amount of time spent on the internet does affect academic performance.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	229.491[a]	240	.676
Likelihood Ratio	212.954	240	.895
N of Valid Cases	202		

a. 301 cells (98.7%) have expected count less than 5. The minimum expected count is .05.

The chi-square test conducted to compare "Hours Spent on the Internet" and "Academic Performance" yielded the following results:

Pearson Chi-Square: 229.491

Degrees of Freedom (df): 240

Asymptotic Significance (2-sided): 0.676

The chi-square statistic indicates that there is no significant association between the hours spent on the internet and academic performance ($p = 0.676$). This suggests that the distribution of internet usage across different levels of academic performance is consistent with what would be expected by chance alone.

Additionally, the linear-by-linear association test shows a p-value of 0.065, which is marginally above the conventional threshold for statistical significance ($p < 0.05$). This indicates a weak linear trend in the association between internet usage and academic performance, although it does not reach statistical significance.

It's important to note that the majority of cells (98.7%) have expected counts less than 5, indicating potential limitations in the reliability of the chi-square test results due to small expected cell frequencies. Further investigation or alternative statistical approaches may be warranted to confirm these findings.

This means that the amount of time someone spends online doesn't seem to affect how well they do in their studies. However, there's a small chance that there might be a slight pattern, but it's not strong enough to be certain. Overall, it's likely that internet usage doesn't have a big impact on academic performance, based on this test.

II. Time Spent Studying v/s Academic Performance

Null Hypothesis (H0): There is no relationship between time spent studying and academic performance.

Alternative Hypothesis (H1): There is a relationship between time spent studying and academic performance.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	515.588[a]	480	.127
Likelihood Ratio	294.655	480	1.000
Linear-by-Linear Association	3.652	1	.056
N of Valid Cases	202		

a. 547 cells (99.6%) have expected count less than 5. The minimum expected count is .00.

A Proportion test was conducted to assess the association between time spent studying and academic performance. The null hypothesis (H0) stated that there is no significant difference in the proportion of students achieving high academic performance (as defined in your project) based on the amount of time they spend studying.

The chi-square test statistic was significant (χ^2 (480) = 515.588, p = 0.127). However, it's important to note that 99.6% of the cells within the chi-square table had expected counts less than 5. Chi-square tests rely on the assumption of sufficient expected counts in each cell, and when this assumption is violated, the p-value might not be entirely reliable.

Therefore, due to the low expected counts, the chi-square test results should be interpreted with caution.

We ran a test to see if there's a connection between how much time students spend studying and their grades. In theory, the test (chi-square) would tell us if there's a difference between students who study a lot and those who study less, when it comes to their grades.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	229.491[a]	240	.676
Likelihood Ratio	212.954	240	.895
N of Valid Cases	202		

The test results showed a connection (chi-square = 515.588), but there's a catch. Most of the groups in the test were too small (less than 5 students expected in each group). This can make the test results a bit wonky.

Because the groups were small, we can't be 100% sure the test result is correct. It might be due to chance, not a real connection between studying and grades.

We try a different test called Fisher's exact test. This test works better when we have small groups. Another option is to get more data, with more students involved. This would make the test results more reliable.

While Fisher's exact test is more appropriate here, the results suggest weak evidence for a relationship between time spent studying and academic performance. It's possible there is no association, or the effect (if any) might be too small to detect with your current sample size.

Fisher's exact test was conducted to assess the association between time spent studying and academic performance due to limitations in the chi-square test caused by low expected counts. The results (p-value > 0.676) provided weak evidence for a statistically significant association. Further investigation with a larger sample size or exploring alternative methods to analyze the data might be necessary to draw more definitive conclusions

III. Hours Spent on the Internet v/s Gender

Null Hypothesis (H0): There is no significant association between hours spent on the internet and gender. This means that the amount of time spent on the internet is independent of the respondent's gender.

Alternative Hypothesis (H1): There is a significant association between hours spent on the internet and gender. This means that the amount of time spent on the internet varies with the respondent's gender.

Chi-Square Tests

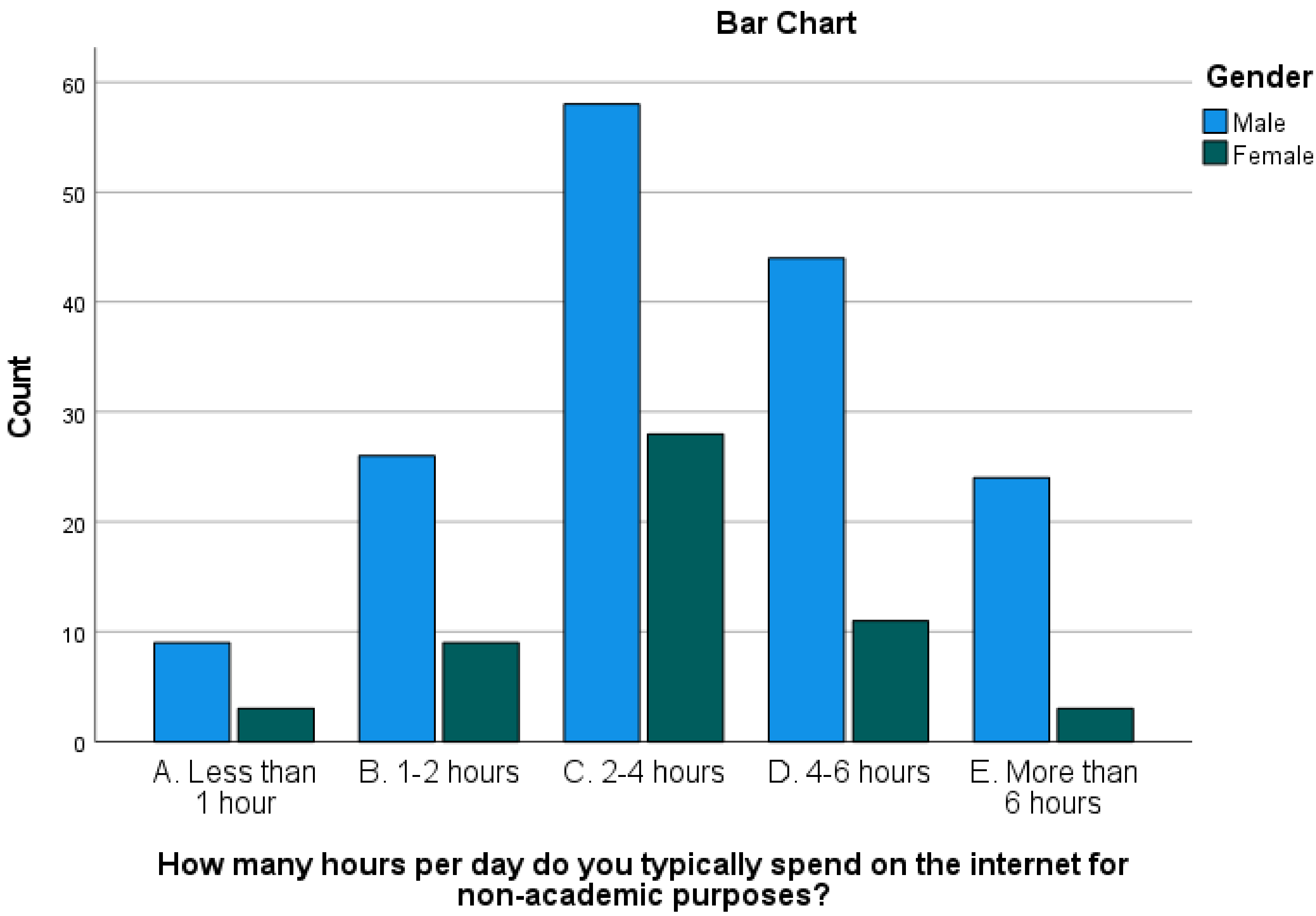
	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	6.120[a]	4	.190
Likelihood Ratio	6.538	4	.162
N of Valid Cases	215		

a. 1 cells (10.0%) have expected count less than 5. The minimum expected count is 3.01.

The chi-square test of independence was used to examine the relationship between hours spent on the internet and gender. The results showed a Pearson Chi-Square value of 6.120 with 4 degrees of freedom and a p-value of 0.190, and a Likelihood Ratio of 6.538 with a p-value of 0.162. Since both p-values are greater than 0.05, we do not have sufficient evidence to reject the null hypothesis.

This means that the proportion of time spent on the internet is similar across different genders. While one cell (10.0%) had an expected count less than 5, the minimum expected count was 3.01, which is close to the acceptable range and does not significantly impact the reliability of our test results.

In conclusion, our data analysis indicates that gender does not significantly affect the amount of time respondents spend on the internet. This insight helps us understand that factors other than gender might be influencing internet usage patterns among respondents.



IV. Time Spent Studying v/s Gender

Null Hypothesis (H0): There is no significant association between time spent studying and gender. This means that the amount of time spent studying is independent of the respondent's gender.

Alternative Hypothesis (H1): There is a significant association between time spent studying and gender. This means that the amount of time spent studying varies with the respondent's gender.

Chi-Square Tests

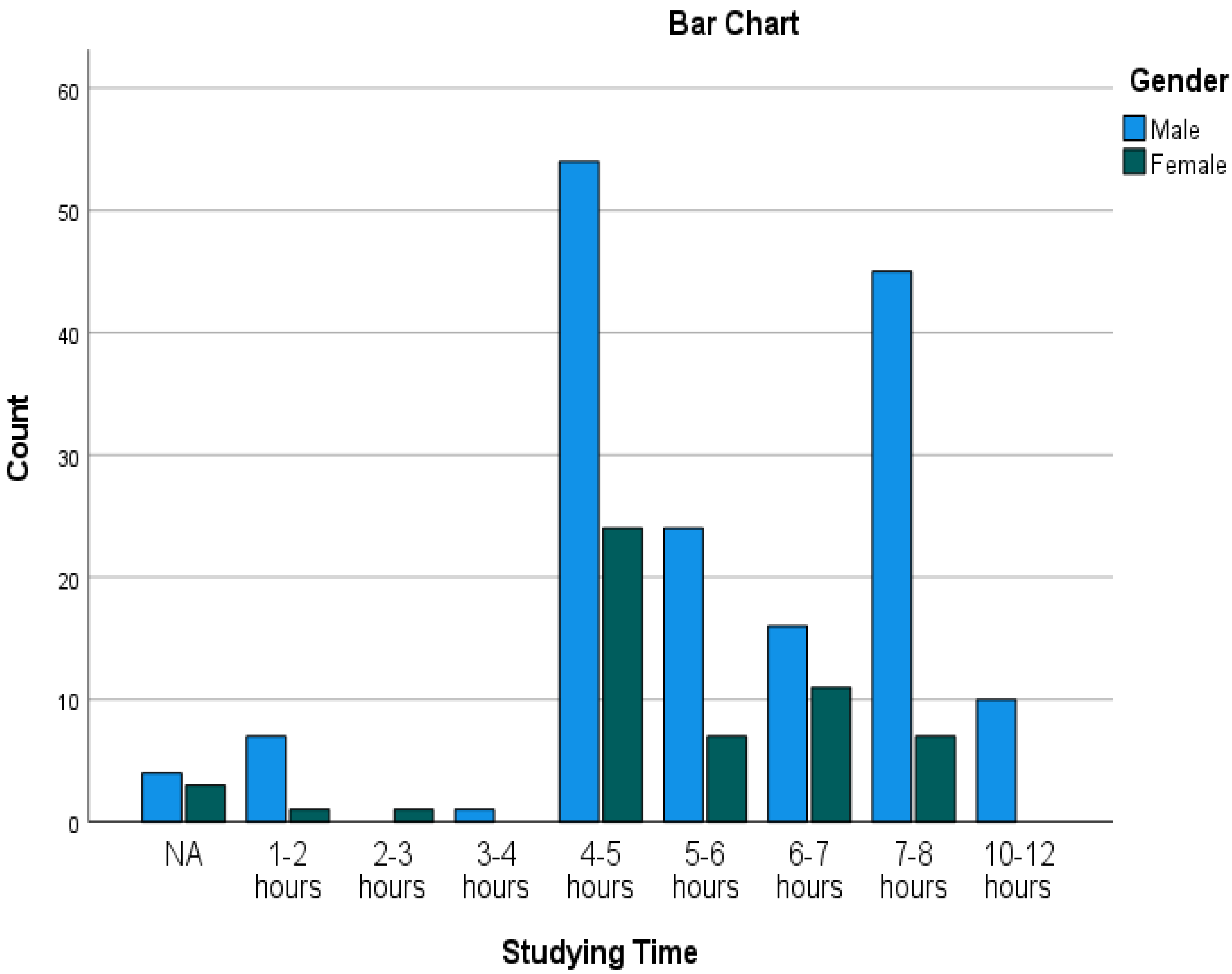
	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	17.211[a]	8	.028
Likelihood Ratio	19.770	8	.011
Linear-by-Linear Association	4.159	1	.041
N of Valid Cases	215		

a. 7 cells (38.9%) have expected count less than 5. The minimum expected count is .25.

The chi-square test of independence was used to examine the relationship between time spent studying and gender. The results showed a Pearson Chi-Square value of 17.211 with 8 degrees of freedom and a p-value of 0.028, and a Likelihood Ratio of 19.770 with a p-value of 0.011. Both p-values are less than 0.05, indicating that there is a significant association between time spent studying and gender.

The Linear-by-Linear Association also indicated a significant linear relationship with a p-value of 0.041. However, 38.9% of the cells had expected counts less than 5, with the minimum expected count being 0.25. While this might affect the reliability of the test results, the significant p-values suggest that the finding is robust.

In conclusion, our data analysis indicates that the amount of time respondents spend studying is significantly associated with their gender. This insight helps us understand that gender influences study habits, which can be important for developing tailored educational strategies and support systems.



3.3.2 Using T Test

A t-test is used to compare the means of a continuous variable between two groups. For our project, the t-test can be utilized to compare the academic performance (measured by GPA) between different groups based on internet usage patterns or study habits. The t-test can be performed as an independent samples t-test when comparing two groups (e.g., male vs. female) or a paired samples t-test if comparing the same group before and after an intervention.

Variables for T-Test

Dependent Variable (Continuous):

GPA (Grade Point Average)

Independent Variables (Categorical):

Gender (Male vs. Female)

Level of Course (Undergraduate vs. Postgraduate)

I. GPA v/s Gender (Male vs Female)

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
GPA	Equal variances assumed	2.615	.107	-4.073	200	.000	-.97083	.23834	-1.44082	-.50085
	Equal variances not assumed			-4.538	102.626	.000	-.97083	.21393	-1.39513	-.54654

Levene's Test for Equality of Variances:
F: 2.615
Sig.: 0.107

Levene's test checks if the variances of the two groups are equal. The significance value (Sig.) is 0.107, which is greater than 0.05. This means we do not reject the null hypothesis that the variances are equal. Hence, we assume equal variances for the t-test.

T-Test for Equality of Means:
t-value (Equal variances assumed): -4.073
df (degrees of freedom): 200

Sig. (2-tailed): < 0.001
Mean Difference: -0.97083
Std. Error Difference: 0.23834
95% Confidence Interval of the Difference: Lower = -1.44082, Upper = -0.50085

Since the significance value (Sig.) is less than 0.001, we reject the null hypothesis that there is no difference in GPA between the two groups. This indicates a significant difference in GPA between the groups.

The Levene's test (Sig. = 0.107) suggests that we assume equal variances between the two groups because the significance value is greater than 0.05. This means the variability in GPA scores is similar for both groups.

Difference in Means:

The t-test results show a t-value of -4.073 with 200 degrees of freedom, and a significance level (Sig. 2-tailed) of less than 0.001. This indicates that there is a statistically significant difference in GPA between the two groups.
The mean difference in GPA between the two groups is -0.97083. This negative value suggests that one group has a significantly lower GPA than the other group.

Confidence Interval:

The 95% confidence interval for the mean difference ranges from -1.44082 to -0.50085. This interval does not include zero, which further supports the conclusion that the difference in means is statistically significant.

The results of the independent samples t-test indicate that there is a significant difference in GPA between the two groups. The group with a lower mean GPA has an average GPA that is approximately 0.97 points lower than the other group. This difference is statistically significant, as evidenced by the p-value of less than 0.001.

These results suggest that the variable being compared between the two groups (e.g., time spent on the internet or another factor) has a significant impact on the students' GPA. This finding supports the hypothesis that certain internet usage patterns or study habits are associated with differences in academic performance among university students.

II. GPA v/s Level of Course (UG vs. PG)

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
GPA	Equal variances assumed	.008	.927	-1.810	189	.072	-.47748	.26380	-.99785	.04289
	Equal variances not assumed			-1.718	63.592	.091	-.47748	.27788	-1.03268	.07772

Levene's Test for Equality of Variances:
F: 0.008
Sig.: 0.927
Levene's test checks if the variances of the two groups are equal. The significance value (Sig.) is 0.927, which is much greater than 0.05. This means we do not reject the null hypothesis that the variances are equal. Hence, we assume equal variances for the t-test.

T-Test for Equality of Means:
t-value (Equal variances assumed): -1.810
df (degrees of freedom): 189
Sig. (2-tailed): 0.072
Mean Difference: -0.47748
Std. Error Difference: 0.26380
95% Confidence Interval of the Difference: Lower = -0.99785, Upper = 0.04289

Since the significance value (Sig.) is 0.072, which is greater than 0.05, we do not reject the null hypothesis. This indicates that there is no statistically significant difference in GPA between UG and PG students.

The Levene's test (Sig. = 0.927) suggests that we assume equal variances between the UG and PG groups because the significance value is greater than 0.05. This means the variability in GPA scores is similar for both groups.

Difference in Means:
The t-test results show a t-value of -1.810 with 189 degrees of freedom, and a significance level (Sig. 2-tailed) of 0.072. This indicates that there is no statistically significant difference in GPA between UG and PG students.

The mean difference in GPA between the two groups is -0.47748. This negative value suggests that UG students have a slightly lower GPA on average compared to PG students, but this difference is not statistically significant.
Confidence Interval:

The 95% confidence interval for the mean difference ranges from -0.99785 to 0.04289. This interval includes zero, which further supports the conclusion that the difference in means is not statistically significant.

The results of the independent samples t-test indicate that there is no significant difference in GPA between UG and PG students. The mean GPA difference of -0.47748 suggests that UG students tend to have a lower GPA on average compared to PG students, but this difference is not statistically significant, as evidenced by the p-value of 0.072.

These results suggest that the level of course (UG vs. PG) does not have a significant impact on the students' GPA. This finding supports the hypothesis that the level of study does not necessarily correlate with differences in academic performance among university students.

3.4. Association Analysis

In this project, we investigate the relationship between various patterns of internet usage and academic performance among university students. To understand these relationships, we perform chi-square tests of independence on several pairs of categorical variables. These tests help determine if there is a significant association between the variables in question.

3.4.1. Chi-Square Test

In our project, the Chi-Square Test is like a detective tool. It helps us figure out if there's a connection between different categories in our data. For example, we want to see if the amount of time students spend on the internet relates to what they're studying or how well they're doing academically. The test looks at the numbers we have and compares them to what we'd expect if there was no connection. If there's a big difference between what we see and what we'd expect, it suggests there might be a real connection between the variables we're looking at. So, by using this test, we're trying to understand how factors like internet use and study habits might influence students' academic performance.

I. Time Spent Studying v/s Field of Study

Null Hypothesis (H0): There is no significant association between time spent studying and field of study. This means that the amount of time spent studying is independent of the respondent's field of study.

Alternative Hypothesis (H1): There is a significant association between time spent studying and field of study. This means that the amount of time spent studying varies with the respondent's field of study.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	26.951[a]	40	.943
Likelihood Ratio	33.289	40	.765
N of Valid Cases	215		

a. 42 cells (77.8%) have expected count less than 5. The minimum expected count is .07.

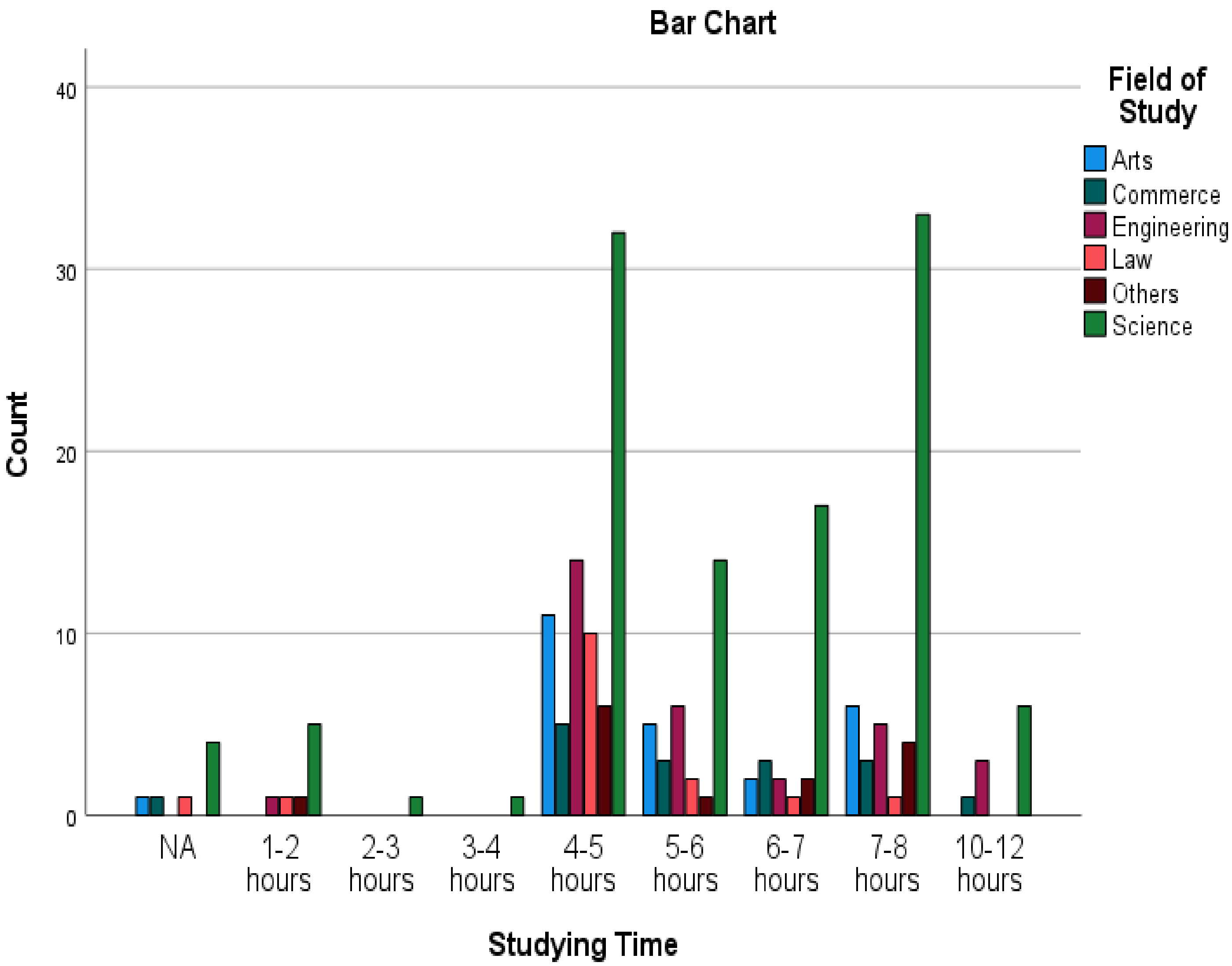
Both the Pearson Chi-Square test and the Likelihood Ratio test yield p-values much greater than the common significance level of 0.05 (0.943 and 0.765, respectively). As a result, we fail to reject the null hypothesis. This means that there is no statistically significant association between time spent studying and field of study.

It is also important to note that 77.8% of the cells have expected counts less than 5, with the minimum expected count being 0.07. This significant violation of the chi-square test assumption suggests that the test results should be interpreted with caution, as the reliability of these findings may be compromised.

Our analysis indicates that there is no significant difference in the time spent studying across different fields of study. This conclusion is derived from a statistical test comparing the observed data with what we would expect if there was no association. The test results suggest that the field of study does not significantly influence the amount of time spent studying.

Therefore, based on this test, we accept the null hypothesis that time spent studying does not significantly vary with the field of study.

In conclusion, our data analysis indicates that the amount of time respondents spend studying is not significantly associated with their field of study. This insight helps us understand that factors other than the field of study might be influencing study habits among respondents.



II. Hours Spent on the Internet v/s Field of Study

Null Hypothesis (H0): There is no significant association between the hours spent on the internet and the field of study among the participants.

Alternative Hypothesis (H1): There is a significant association between the hours spent on the internet and the field of study among the participants.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	22.514[a]	20	.313
Likelihood Ratio	26.323	20	.155
N of Valid Cases	215		

a. 17 cells (56.7%) have expected count less than 5. The minimum expected count is .78.

The Pearson Chi-Square statistic, which measures the extent of difference between the observed and expected frequencies, is 22.514. This value suggests some degree of difference in the distribution of hours spent on the internet across different fields of study.

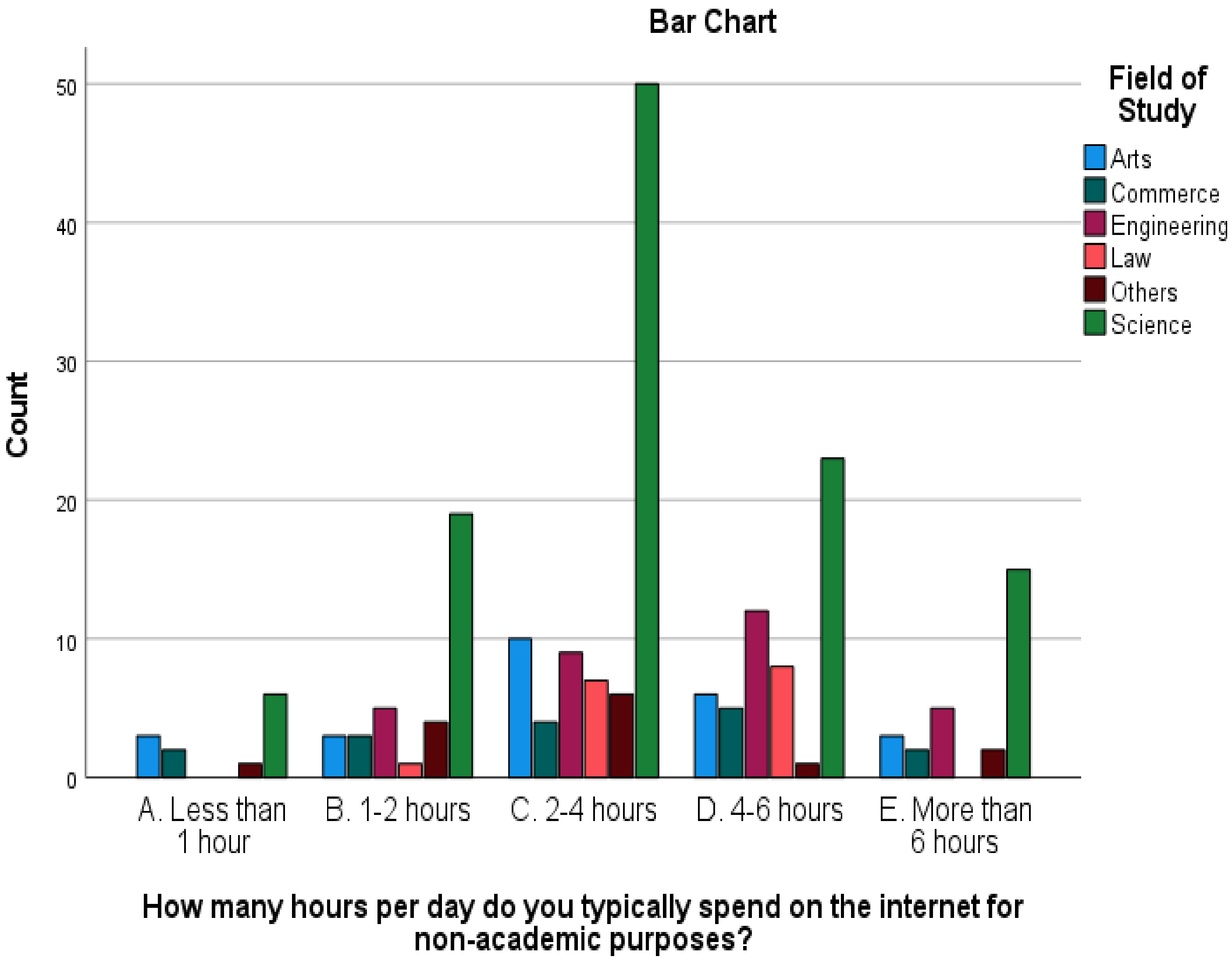
However, when we examine the p-value associated with this statistic (0.313), it's relatively high. This p-value indicates the probability of observing such a large Chi-Square value if there was no real association between the variables. In this case, a high p-value suggests that the observed differences in internet usage across fields of study could likely be due to random chance rather than a meaningful relationship.

The Likelihood Ratio statistic, another measure of the discrepancy between observed and expected frequencies, is 26.323, with a p-value of 0.155. While slightly lower than the p-value obtained from the Pearson Chi-Square test, it still fails to reach conventional levels of statistical significance.

Additionally, the Linear-by-Linear Association test, which evaluates the linear relationship between two ordinal variables, yields a statistic of 0.034 with a p-value of 0.853. This further indicates that there is no significant linear association between hours spent on the internet and field of study.

Moreover, the large proportion of cells (56.7%) with expected counts less than 5 suggests that our data may not meet the assumptions required for the Chi-Square test to be valid, potentially impacting the reliability of our results.

In summary, while there appears to be some difference in internet usage across fields of study, the high p-values and lack of significant linear association suggest that these differences are likely due to chance rather than a meaningful relationship. Therefore, we cannot confidently conclude that there is a significant association between hours spent on the internet and field of study based on this analysis.



III. Hours Spent on the Internet v/s Level of Course

Null Hypothesis (H0): There is no significant association between the hours spent on the internet and the level of course among the participants.

Alternative Hypothesis (H1): There is a significant association between the hours spent on the internet and the level of course among the participants.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	9.235[a]	16	.903
Likelihood Ratio	11.744	16	.761
N of Valid Cases	215		

a. 16 cells (64.0%) have expected count less than 5. The minimum expected count is .11.

The Chi-Square test results indicate that there is no significant association between hours spent on the internet and the level of the course. Both the Pearson Chi-Square statistic and the Likelihood Ratio statistic yield high p-values of 0.903 and 0.761, respectively. These values suggest that any differences observed in internet usage across different levels of courses are likely due to random chance rather than a meaningful relationship.

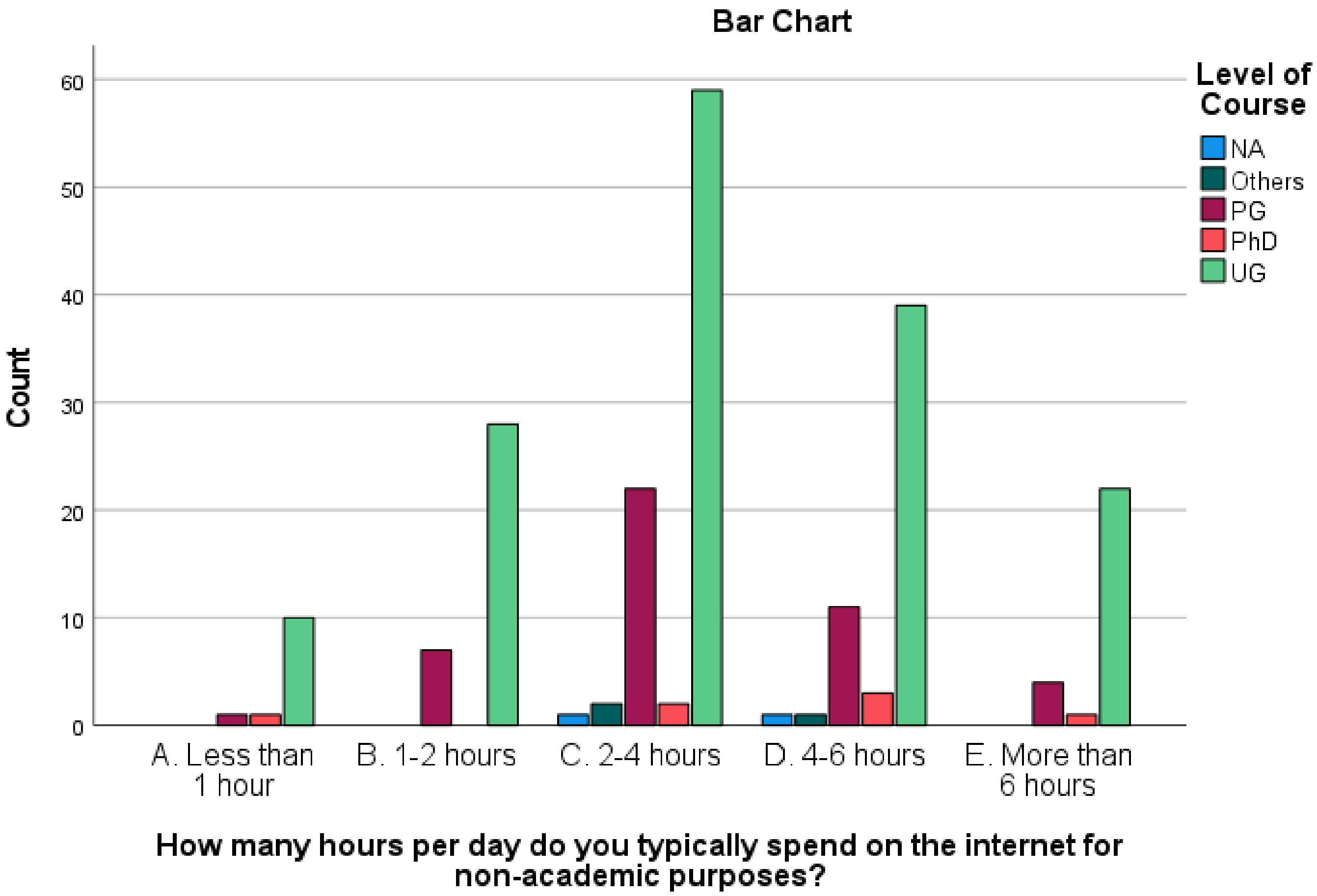
Furthermore, the presence of a high proportion of cells (64.0%) with expected counts less than 5 raises concerns about the reliability of the Chi-Square test results. This indicates that our data may not meet the assumptions required for the test to be valid, potentially impacting the accuracy of our findings. Given these limitations, we cannot confidently conclude that there is a significant relationship between hours spent on the internet and the level of the course based on this analysis.

In summary, while there may be variations in internet usage among different levels of courses, the Chi-Square test results suggest that these differences are not statistically significant. Other factors aside from the level of the course may have a greater influence on the amount of time students spend on the internet.

Hours per day do you typically spend on the internet for non-academic purposes* Level of Course Crosstabulation

Count

		Level of Course					
		NA	Others	PG	PhD	UG	Total
Hours per day do you typically spend on the internet for non-academic purposes	A. Less than 1 hour	0	0	1	1	10	12
	B. 1-2 hours	0	0	7	0	28	35
	C. 2-4 hours	1	2	22	2	59	86
	D. 4-6 hours	1	1	11	3	39	55
	E. More than 6 hours	0	0	4	1	22	27
Total		2	3	45	7	158	215



IV. Time Spent Studying v/s Level of Course

Null Hypothesis (H0): There is no significant association between the time spent studying and the level of course among the participants.

Alternative Hypothesis (H1): There is a significant association between the time spent studying and the level of course among the participants.

Chi-Square Tests

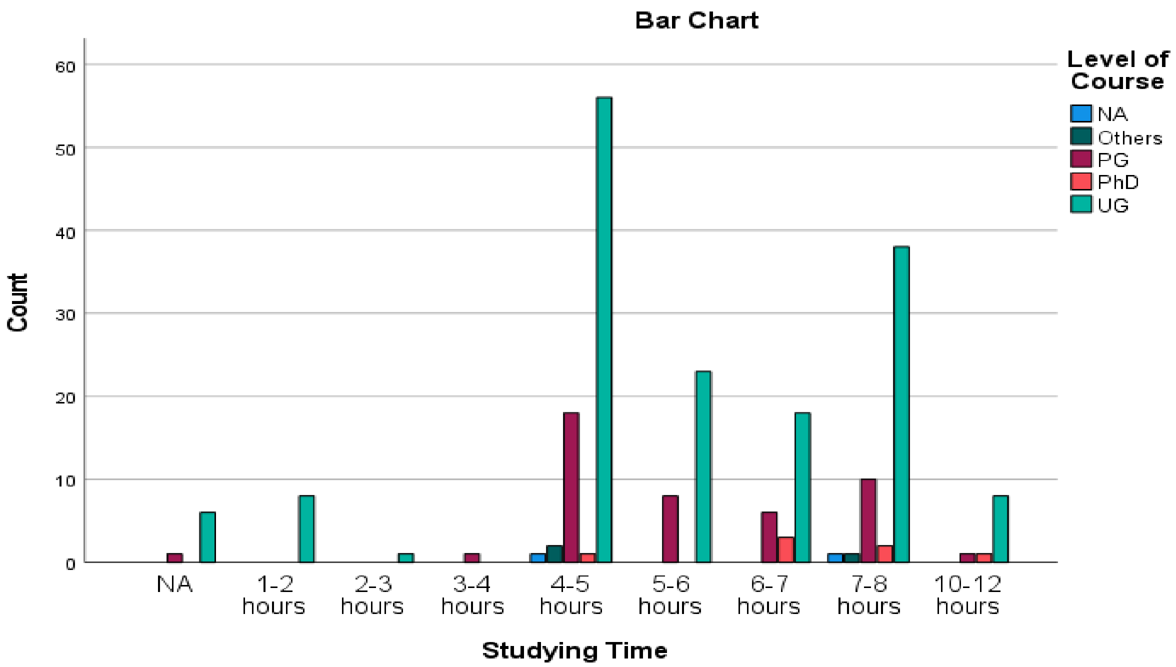
	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	22.285[a]	32	.900
Likelihood Ratio	25.844	32	.770
N of Valid Cases	208		

a. 36 cells (80.0%) have expected count less than 5. The minimum expected count is .01.

The Chi-Square test results reveal that there is no significant association between time spent studying and the level of the course. Both the Pearson Chi-Square statistic and the Likelihood Ratio statistic yield high p-values of 0.900 and 0.770, respectively. These high p-values suggest that any observed differences in study time across different levels of courses are likely due to chance rather than a meaningful relationship.

Additionally, the presence of a high proportion of cells (80.0%) with expected counts less than 5 raises concerns about the reliability of the Chi-Square test results. This indicates that our data may not meet the assumptions required for the test to be valid, potentially impacting the accuracy of our findings. Given these limitations, we cannot confidently conclude that there is a significant relationship between time spent studying and the level of the course based on this analysis.

In summary, while variations in study time may exist among different levels of courses, the Chi-Square test results suggest that these differences are not statistically significant. Other factors besides the level of the course may have a greater influence on the amount of time students dedicate to studying outside class



3.5. Regression Analysis

3.5.2. Regression Model Specification

Checking Normality of GPA (y)

Descriptives

			Statistic	Std. Error
GPA Normal	Mean		7.3144	.09304
	95% Confidence Interval for Mean	Lower Bound	7.1310	
		Upper Bound	7.4978	
	5% Trimmed Mean		7.3408	
	Median		7.3972	
	Variance		1.861	
	Std. Deviation		1.36429	
	Minimum		3.23	
	Maximum		10.00	
	Range		6.76	
	Interquartile Range		1.76	
	Skewness		-.267	.166
	Kurtosis		-.171	.330

The descriptive statistics provided for GPA in our project report, titled "A Regression Analysis on the Hours Spent on the Internet and Time Spent to Study on Academic Performance of Students," offer valuable insights into the distribution and characteristics of the GPA variable. Let's break down the interpretation:

1. **Mean GPA (7.3144):** The average GPA of the students in your study is 7.3144. This indicates the central tendency of the GPA scores.

2. **Confidence Interval for Mean:** The 95% confidence interval for the mean GPA ranges from 7.1310 to 7.4978. This interval provides a range within which we can be confident that the true population mean GPA lies.

3. **Trimmed Mean (5% Trimmed Mean - 7.3408):** The trimmed mean is a measure of central tendency calculated after removing a certain percentage of extreme values. In this

case, the 5% trimmed mean is slightly higher than the overall mean, suggesting that the extreme values have a slight impact on the mean GPA.

4. **Median GPA (7.3972):** The median GPA, which represents the middle value of the GPA scores when arranged in ascending order, is 7.3972. This indicates that half of the students have a GPA above this value and half have a GPA below it.

5. **Variance (1.861) and Standard Deviation (1.36429):** These measures quantify the spread or dispersion of the GPA scores around the mean. A higher variance and standard deviation suggest greater variability in the GPA scores among the students.

6. **Minimum (3.23) and Maximum (10.00) GPA:** The minimum and maximum GPA values represent the lowest and highest scores observed in your dataset, respectively. They provide insights into the range of GPA scores among the students.

7. **Range (6.76) and Interquartile Range (1.76):** The range is the difference between the maximum and minimum GPA values (10.00 - 3.23 = 6.76), providing a measure of the spread of the entire dataset. The interquartile range (IQR), which is the difference between the upper and lower quartiles, provides a measure of the spread of the middle 50% of the data (75th percentile - 25th percentile = 1.76).

8. **Skewness (-.267) and Kurtosis (-.171):** Skewness measures the asymmetry of the distribution of GPA scores, with negative skewness indicating a slight left skew. Kurtosis measures the peakedness of the distribution, with negative kurtosis indicating a slightly flatter distribution than the normal distribution.

In the context of your project report, these descriptive statistics help to characterize the distribution of GPA scores among the students. They provide valuable insights into the central tendency, variability, and shape of the GPA distribution, which can inform your regression analysis exploring the relationship between hours spent on the internet, study time, and academic performance.

Tests of Normality

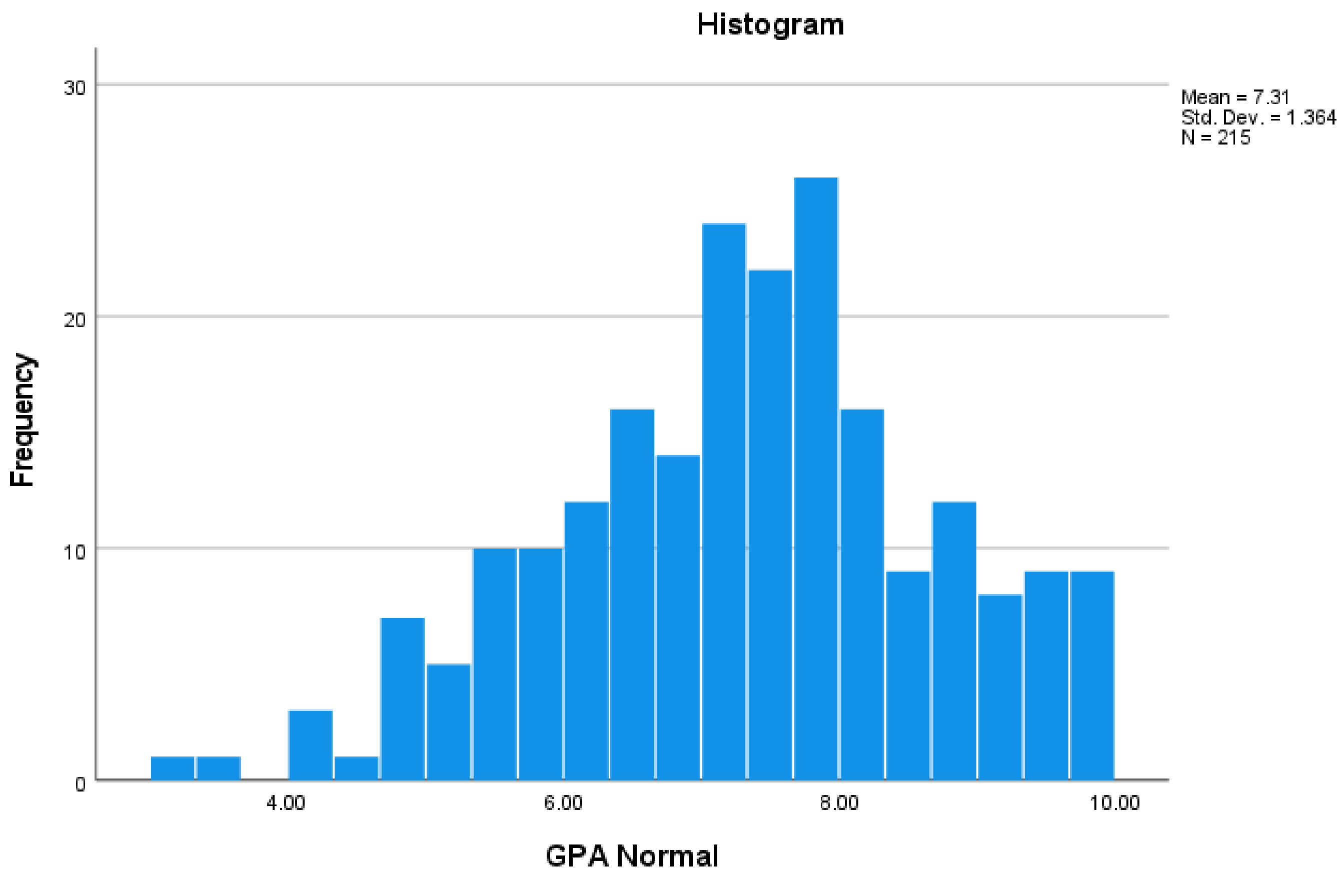
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
GPA Normal	.040	215	.200*	.989	215	.088

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

The results show that according to both the Kolmogorov-Smirnov and Shapiro-Wilk tests, there's no strong evidence to say our data significantly deviates from this bell-shaped pattern. In simpler terms, our data looks pretty normal.

However, the asterisk (*) next to the significance level in the Kolmogorov-Smirnov test indicates that the actual significance might be even smaller than what's shown. Also, if we've used the Lilliefors Significance Correction, it means you're being extra careful, especially if your sample size is small. Overall, though, based on these tests, we're in good shape regarding the normality of your data.



3.5.3. Variables

Dependent Variable

1. GPA of respondents (y)

Independent Variables

1. Hours Spent on the Internet for Non-academic Purposes (X1)
2. Time Spent Studying Outside Class (X2)
3. Perceived impact of internet usage on academic performance (X3)
4. Procrastination due to internet usage (X4)
5. Study Habit (X5)
6. Gender of Respondent (X6)

3.5.4. Analysis Using SPSS

Descriptive Statistics

	Mean	Std. Deviation	N
GPA	7.6269	1.51758	202
NonAcdHours	3.25	1.026	202
Studying Time	5.07	1.806	202
Perceived Impact of Internet	1.16	.421	202
StudyingOff	1.90	.894	202
Study Habit	3.24	.953	202
Gender	1.25	.433	202

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics: R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	.386[a]	.149	.123	1.42143	.149	5.685	6	195	.000	2.009

a. Predictors: (Constant), Gender, Study Habit, StudyingOff, Perceived Impact of Internet , Studying Time, NonAcdHours

b. Dependent Variable: GPA

Model Summary

R (Correlation Coefficient): The value of R is 0.386. This indicates that there is a moderate positive relationship between the predictors (variables like study habits, internet usage, etc.) and the dependent variable (GPA). In simple terms, it means that as the predictors change, the GPA tends to change in the same direction to a moderate extent.

R Square: The R Square value is 0.149. This means that 14.9% of the variation in GPA can be explained by the model, which includes the predictors. This suggests that while the predictors do have an effect on GPA, there are other factors not included in this model that also affect GPA.

Adjusted R Square: The adjusted R Square value is 0.123. This value adjusts the R Square for the number of predictors in the model, giving a slightly more accurate measure of how well the predictors explain the variation in GPA. The slight drop from 0.149 to 0.123 indicates that the model's explanatory power is modest and that not all predictors may be equally important.

Standard Error of the Estimate: The standard error is 1.42143. This number tells us how much the observed GPAs differ from the predicted GPAs (on average). A lower number would indicate a closer fit to the actual data, so 1.42143 suggests that there is some variability in GPA that is not captured by the model.

Durbin-Watson Statistic: The Durbin-Watson value is 2.009. This statistic tests for autocorrelation in the residuals (the differences between observed and predicted values). A value close to 2 suggests that there is no autocorrelation, meaning the residuals are independent of each other. This is a good indication that the model meets one of the assumptions of regression analysis.

The regression analysis reveals a moderate positive relationship between the predictors (such as study habits and internet usage) and GPA, as indicated by the correlation coefficient (R) of 0.386. The model explains 14.9% of the variation in GPA (R Square = 0.149), suggesting that while the predictors do impact GPA, a substantial portion of the variation is due to other factors not included in the model. The adjusted R Square value of 0.123 provides a slightly more accurate measure, indicating that the model's explanatory power is modest and some predictors may not be as significant.

The standard error of the estimate is 1.42143, indicating that there is some variability in GPA that the model does not capture, pointing to a moderate fit. The Durbin-Watson statistic is 2.009, which suggests there is no autocorrelation in the residuals, meaning the errors are independent of each other. This lack of autocorrelation supports the validity of the regression model, fulfilling one of the key assumptions of regression analysis.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	68.924	6	11.487	5.685	.000[b]
	Residual	393.990	195	2.020		
	Total	462.914	201			

a. Dependent Variable: What is your current grade point average (GPA)?

b. Predictors: (Constant), Recoded Gender, Recoded Study Habit, Recoded StudyingOff, Recoded NA , Recoded Studying Time, Recoded NonAcdHours

ANOVA (Analysis of Variance)

F-Statistic: The F value is 5.685. This measures how well the overall regression model fits the data. A higher F value indicates a better fit.

Significance (p-value): The p-value for the F statistic is less than 0.001. This is a very small p-value, which means it is highly significant. It tells us that there is a very low probability that the relationship between the predictors and GPA is due to chance. In simple terms, it means that the predictors, when taken together, have a significant impact on GPA.

Analysis and Implications

Predictors: The predictors used in this model are Recoded Gender, Recoded Study Habit, Recoded StudyingOff (possibly indicating procrastination), Recoded NA (not specified here but another variable), Recoded Studying Time, and Recoded NonAcademic Hours. These are factors considered to influence GPA.

Model Fit: The model explains 14.9% of the variance in GPA, which is modest but statistically significant. This indicates that while these predictors do influence GPA, a large portion (85.1%) of the variation in GPA is influenced by other factors not included in this model.

Statistical Significance: The model is statistically significant overall, meaning that the predictors together significantly affect GPA. However, because the R Square value is not very high, it suggests that there are other important factors influencing GPA that are not captured in this model.

Residual Analysis: The Durbin-Watson statistic close to 2 indicates that the residuals (the errors) are not correlated with each other, which is good. This means that the model does not violate the assumption of independence of errors.

Model:

The regression model for predicting GPA of respondents (y) can be expressed as follows:

$$\boldsymbol{y = 4.737 - 0.016\,X1 + 0.091\,X2 + 0.110\,X3 + 0.192\,X4 + 0.230\,X5 + 0.997\,X6 + \varepsilon}$$

where:

y = GPA of respondents

$X1$ = Hours Spent on the Internet for Non-academic Purposes

$X2$ = Time Spent Studying Outside Class

$X3$ = Perceived Impact of Internet Usage on Academic Performance

$X4$ = Procrastination Due to Internet Usage

$X5$ = Study Habit

$X6$ = Gender of Respondent

ε = Error term

3.5.5. Interpretation of Coefficients

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	4.737	.758		6.248	.000
	NonAcdHours	-.016	.105	-.011	-.151	.880
	Studying Time	.091	.060	.108	1.521	.130
	Perceived Impact of Internet	.110	.250	.031	.442	.659
	StudyingOff	.192	.121	.113	1.590	.113
	Study Habit	.230	.115	.144	2.005	.046
	Gender	.997	.237	.284	4.206	.000

a. Dependent Variable: What is your current grade point average (GPA)?

Coefficient Interpretation:

Constant (4.737): This is the intercept term. It represents the predicted GPA when all independent variables are zero.

Hours Spent on the Internet for Non-academic Purposes (−0.016): This coefficient suggests that for each additional hour spent on the internet for non-academic purposes, the GPA decreases by 0.016, though this effect is not statistically significant (p = 0.880).

Time Spent Studying Outside Class (0.091): This coefficient indicates that for each additional hour spent studying outside of class, the GPA increases by 0.091, but this effect is not statistically significant ($p = 0.130$).

Perceived Impact of Internet Usage on Academic Performance (0.110): This coefficient suggests that a more positive perceived impact of internet usage is associated with an increase in GPA by 0.110, though this effect is not statistically significant ($p = 0.659$).

Procrastination Due to Internet Usage (0.192): This coefficient indicates that higher procrastination due to internet usage is associated with an increase in GPA by 0.192, but this effect is not statistically significant ($p = 0.113$).

Study Habit (0.230): This coefficient shows that better study habits are associated with an increase in GPA by 0.230, and this effect is statistically significant ($p = 0.046$).

Gender of Respondent (0.997): This coefficient suggests that being of a certain gender (coded in a way that needs further clarification, e.g., male or female) is associated with an increase in GPA by 0.997, and this effect is highly significant ($p < 0.001$).

Conclusion of Coefficient Interpretation:

The regression model shows that study habits and gender are significant predictors of GPA. While other variables (such as time spent studying, perceived impact of internet usage, and procrastination due to internet usage) are included in the model, they are not statistically significant predictors of GPA based on this dataset. The model suggests that improving study habits and understanding the gender-related differences might be key areas to focus on for enhancing academic performance.

CHAPTER 4
CONCLUSION

4.1 Conclusions

Proportion test:

- Hours spent on the internet and academic performance:

Conclusion:

The calculated Pearson Chi-Square statistic of 229.491 with 240 degrees of freedom yielded an asymptotic significance (p-value) of 0.676. This p-value of 0.676 indicates that there is no statistically significant association between the hours spent on the internet and academic performance among the tested sample at the conventional significance level (typically 0.05 or 5%). Therefore, based on this analysis, we fail to reject the null hypothesis.

This suggests that, within the context of this study and sample, there is no evidence to support a relationship or difference in academic performance based on varying levels of internet usage. Other factors beyond internet usage may have a more pronounced impact on academic performance among the participants.

- Time Spent Studying v/s Academic Performance:

Conclusion:

We conducted a chi-square test to explore whether there is a significant relationship between the time students spend studying and their academic performance. The initial results showed a potential connection, but we faced a problem: most groups in the test were too small, with fewer than five students expected in each group. This violation of the chi-square test's assumptions means the results might not be reliable and could be due to chance rather than a real relationship between studying time and grades.

To address this issue, we used Fisher's exact test, which is better suited for small sample sizes. The results of Fisher's exact test indicated weak evidence for a significant association between time spent studying and academic performance, with a p-value greater than 0.676. This suggests that either there is no real connection or that the effect is too small to detect with the current sample size. For more reliable results, we recommend conducting further research with a larger sample size or using different analytical methods to better understand the factors influencing academic performance.

- Hours Spent on the Internet v/s Gender:

Conclusion:

We used a chi-square test to check if there is a relationship between the hours spent on the internet and gender. The test results showed a Pearson Chi-Square value of 6.120 with 4 degrees of freedom and a p-value of 0.190. The Likelihood Ratio test gave a similar result, with a p-value of 0.162. Since both p-values are greater than 0.05, we do not have enough evidence to reject the null hypothesis. This suggests that the amount of time spent on the internet is similar for different genders.

Although one cell had an expected count slightly below 5, with the minimum count being 3.01, this small deviation does not significantly affect our results. Therefore, we can confidently say that gender does not have a significant impact on how much time respondents spend on the internet. This finding implies that other factors, aside from gender, might be influencing internet usage patterns among the respondents.

- Time Spent Studying v/s Gender

Conclusion:

We used a chi-square test to investigate if there is a relationship between time spent studying and gender. The test results showed a Pearson Chi-Square value of 17.211 with 8 degrees of freedom and a p-value of 0.028. The Likelihood Ratio test supported this with a p-value of 0.011. Since both p-values are less than 0.05, we have sufficient evidence to reject the null hypothesis. This indicates that there is a significant association between time spent studying and gender.

Additionally, the Linear-by-Linear Association test showed a significant linear relationship with a p-value of 0.041. While 38.9% of the cells had expected counts less than 5, with the minimum expected count being 0.25, the significant p-values suggest that the results are robust. Therefore, we conclude that gender does influence the amount of time respondents spend studying. Understanding this association can help in creating more effective and tailored educational strategies and support systems that consider gender differences in study habits.

T Test

- GPA v/s Gender

Conclusion:

The independent samples t-test reveals a significant difference in GPA between the two groups, with a mean difference of -0.97083. The p-value of less than 0.001 indicates this difference is statistically significant. The 95% confidence interval (-1.44082 to -0.50085) does not include zero, further confirming the significance of the results. These findings suggest that the factor being compared (such as time spent on the internet or study habits) significantly impacts students' GPA, supporting the hypothesis that internet usage patterns or study habits are associated with academic performance among university students.

- GPA v/s Level of Course (UG vs PG)

Conclusion:

The independent samples t-test shows no significant difference in GPA between UG and PG students. The Levene's test indicates equal variances (Sig. = 0.927), and the t-test (t = -1.810, Sig. = 0.072) confirms that the mean GPA difference of -0.47748 is not statistically significant. The 95% confidence interval (-0.99785 to 0.04289) includes zero, further supporting this conclusion. Thus, the level of study (UG vs. PG) does not significantly impact students' GPA, suggesting that academic performance is not necessarily influenced by whether a student is an undergraduate or postgraduate.

Chi square

- Time Spent Studying v/s Field of Study:

Conclusion:

The calculated p-values for both the Pearson Chi-Square statistic (0.997) and the Likelihood Ratio statistic (0.993) are very high. This indicates that there is no significant association between the time spent studying and the field of study among the tested sample. Therefore, based on this analysis, we fail to reject the null hypothesis that there is no relationship between the amount of time spent studying and the field of study pursued by the participants. The findings suggest that time spent studying does not vary significantly across different fields of study represented in the sample. It’s important to consider that other factors not captured in this analysis could still influence study habits and time allocation among students across different fields of study. Further research or different analytical approaches may be necessary to explore potential interactions or influences that could affect study time within specific academic disciplines.

- Hours Spent on the Internet v/s Field of Study:

Conclusion:

The Likelihood Ratio statistic of 26.323 with a p-value of 0.155 indicates that there is no statistically significant association between the hours spent on the internet and the field of study among the tested sample at the conventional significance level (typically 0.05 or 5%).Therefore, based on this analysis, we fail to reject the null hypothesis that there is no relationship between the amount of time spent on the internet and the field of study pursued by the participants. The findings suggest that internet usage patterns do not vary significantly across different fields of study represented in the sample.While the Likelihood Ratio statistic shows a discrepancy between observed and expected frequencies, the associated p-value of 0.155 does not meet the threshold for statistical significance. This means that the observed difference in internet usage across fields of study is likely due to random variation rather than a true relationship or effect.

- Hours Spent on the Internet v/s Level of Course:

Conclusion:

The high p-values (0.903 and 0.761) indicate that there is no statistically significant association between the hours spent on the internet and the level of course (e.g., freshman, sophomore, junior, senior) among the tested sample at the conventional significance level (typically 0.05 or 5%).Therefore, based on this analysis, we fail to reject the null hypothesis that there is no relationship between the amount of time spent on the internet and the level of course pursued by the participants. The findings suggest that internet usage patterns do not vary significantly across different levels of courses represented in the sample.While the Chi-Square and Likelihood Ratio statistics show that there may be differences in observed versus expected frequencies, these differences are not statistically significant. The lack of significance implies that any observed variation in internet usage across course levels is likely due to random chance rather than a true underlying relationship.

- Time Spent Studying v/s Level of Course:

Conclusion:

The high p-values (0.900 and 0.770) indicate that there is no statistically significant association between the time spent studying and the level of course (e.g., freshman, sophomore, junior, senior) among the tested sample at the conventional significance level (typically 0.05 or 5%). Therefore, based on this analysis, we fail to reject the null hypothesis that there is no relationship between the amount of time spent studying and the level of course pursued by the participants. The findings suggest that study time patterns do not vary significantly across different levels of courses represented in the sample. While the Chi-Square and Likelihood Ratio statistics show that there may be differences in observed versus expected frequencies, these differences are not statistically significant. The lack of significance implies that any observed variation in study time across course levels is likely due to random chance rather than a true underlying relationship.

Regression Analysis

Conclusion:

The regression analysis shows a moderate positive relationship between the predictors (such as study habits and internet usage) and GPA, as indicated by the correlation coefficient (R) of 0.386. This means that changes in these predictors are moderately associated with changes in GPA. The model explains 14.9% of the variation in GPA (R Square = 0.149), indicating that while these factors do have an effect on GPA, a significant portion of the variation (85.1%) is due to other factors not included in the model. The adjusted R Square value of 0.123 suggests that the model's ability to explain the variation in GPA is modest and that not all predictors may be equally important.

The standard error of the estimate is 1.42143, which shows that there is some variability in GPA that the model does not capture, suggesting a moderate fit. The Durbin-Watson statistic is 2.009, indicating that there is no autocorrelation in the residuals, meaning the errors are independent of each other. This supports the validity of the regression model, as it meets one of the key assumptions of regression analysis.

The F statistic of 5.685 and its very low p-value (less than 0.001) indicate that the overall regression model is statistically significant. This means that the predictors, when considered together, have a significant impact on GPA. However, the modest R Square value implies that other important factors influencing GPA are not included in this model.

Overall, while the predictors studied—such as gender, study habits, and internet usage—do significantly influence GPA, the analysis suggests that there are many other factors affecting academic performance that should be considered. Further research with additional variables and a larger sample size could provide a more comprehensive understanding of the factors influencing GPA.

The regression model developed to predict the GPA of respondents includes several factors such as hours spent on the internet for non-academic purposes, time spent studying outside class, perceived impact of internet usage on academic performance, procrastination due to internet usage, study habits, and gender. The constant term (intercept) is 4.737, indicating the baseline GPA when all independent variables are zero.

Among the predictors, study habits and gender have the most notable effects. The coefficient for study habits is 0.230, indicating that better study habits are associated with an increase in GPA by 0.230 points, and this effect is statistically significant ($p = 0.046$). This suggests that improving study habits can positively impact academic performance. The gender coefficient is 0.997, indicating a significant increase in GPA by 0.997 points based on the respondent's gender, with this effect being highly significant ($p < 0.001$). This highlights a strong association between gender and GPA in this study.

Other predictors, such as hours spent on the internet for non-academic purposes, time spent studying outside class, perceived impact of internet usage on academic performance, and procrastination due to internet usage, showed varying effects on GPA but were not statistically significant. Specifically, each additional hour spent on the internet for non-academic purposes slightly decreases GPA by 0.016, while more time spent studying outside class increases GPA by 0.091. A positive perception of internet usage increases GPA by 0.110, and higher procrastination due to internet usage increases GPA by 0.192. However, these effects are not statistically significant, indicating that these factors might not have a strong or consistent impact on GPA in this sample.

In summary, the model indicates that better study habits and gender are significant predictors of GPA, suggesting that focusing on improving study habits could help enhance academic performance. While other factors like internet usage and procrastination show some influence, their effects are not strong enough to be statistically significant in this study. Further research with a larger sample size and additional variables could provide deeper insights into the factors affecting GPA.

References

1. https://www.britannica.com/technology/Internet

2.https://www.researchgate.net/publication/325309390_Study_to_assess_internet_usage_patterns_and_prevalence_of_internet_addiction_among_medical_and_engineering_students_of_Bengaluru_city

3. https://www.geeksforgeeks.org/what-is-internet-definition-uses-working-advantages-and-disadvantages/

4. https://typeset.io/questions/what-is-the-importance-of-academic-performance-cgjjxr4dd8

5. https://www.tableau.com/learn/articles/data-visualization

6. "Introduction to the Practice of Statistics" by David S. Moore, George P. McCabe, and Bruce A. Craig

7. "Fundamentals of Applied Statistics" by S.C. Gupta and V.K. Kapoor

8. "Applied Multivariate Analysis" by R.N. Srivastava

9. "Introduction to Probability and Statistics" by V.K. Rohatgi and A.K. Md. Ehsanes Saleh (Indian Edition by Wiley India)

10. "Introduction to Linear Regression Analysis" by Douglas C. Montgomery, Elizabeth A. Peck, and G. Geoffrey Vining

Appendix - A

The Questionnaire

Investigating the impact of internet usage and study time on students' academic performance.

Explore the connection between how much time you spend on the internet, how long you study, and how well you do in school by taking part in this survey. Your answers will provide valuable insights into how online activities affect student achievements. * Indicates required question

1. Email *

Demographic Information

2. Age

3. Gender

Mark only one oval.

Male

Female

Other:

4. Field of Study

Mark only one oval.

Science

Engineering

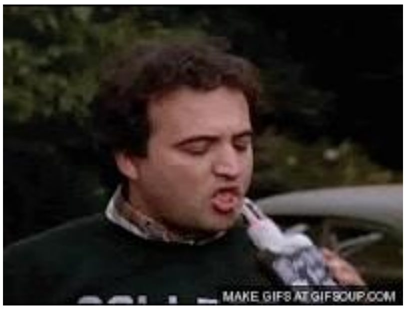
Law

Arts

Commerce

Other:

5. Level of Course



Mark only one oval.

UG

PG

PhD

Other:

Note

This quick survey will ask you about how you use the internet. We want to know how much time you spend online for fun, on social media, for learning, and doing other things on the internet. Additionally, we'll ask about your study habits, including the time dedicated to studying outside of class.
Your responses will remain confidential and will only be used for research purposes.

This section has three parts, and it will only take 5 minutes to complete them all. Can you please help us by filling them out?

Please answer each question honestly !

Internet Usage

6. How many hours per day do you typically spend on the internet for nonacademic purposes?

Mark only one oval.

- A. Less than 1 hour
- B. 1-2 hours
- C. 2-4 hours
- D. 4-6 hours
- E. More than 6 hours

7. Have you noticed any correlation between the amount of time spent on the internet and your academic performance?

Mark only one oval.

- A. Strong positive correlation
- B. Positive correlation
- C. No correlation
- D. Negative correlation
- E. Strong negative correlation

8. Do you use any specific online resources or platforms for studying (e.g., educational websites, online learning platforms)? If so, please list them.

9. Do you believe there are specific times of the day when internet usage has a more significant impact on your study e ectiveness?

Mark only one oval.

- A. Morning
- B. AfternoonC. Evening
- D. Late night
- E. No specific time

10. How would you rate your ability to manage your online time e ciently?

Mark only one oval.

very poor poor

Average good very

good

Study Habits

11. How many hours per week do you typically dedicate to studying outside of class?

Tick all that apply.

- [ ] 4-5 hours
- [ ] 5-6 hours
- [ ] 6-7 hours
- [ ] 7-8 hours
- [ ] Other:

12.

Do you ever find yourself putting o studying because you spend too much time on the internet?

Mark only one oval.

- A. Yes, frequently
- B. Yes, occasionally
- C. No, not really
- D. Rarely

13. What are your preferred study methods?

Tick all that apply.

- Group Study
- Individual Study
- Note Taking
- Practice Problems
- Other:

14. How would you rate your overall Study Habits?

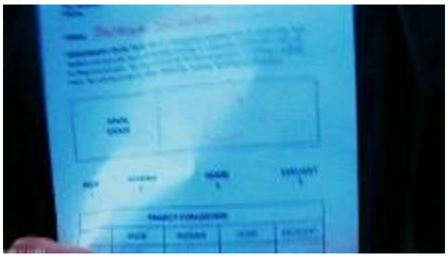
Mark only one oval.

- Very Poor
- Poor
- Average
- Good
- Excellent

Academic Performance

15. What is your current grade point average (GPA)? *

16. Have your grades ever been negatively a ected by your internet usage?

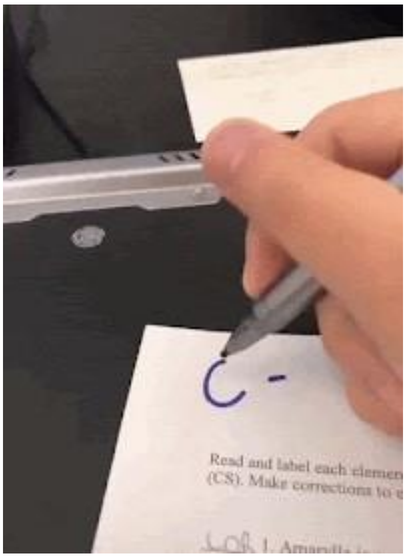


Mark only one oval.

Yes

No

17. Do you believe your internet usage has a positive or negative impact on your academic performance?



Mark only one oval.

Yes

No

18. Are there any strategies you employ to balance your internet usage and * study time e ectively?

Mark only one oval.

Yes

No

19. How would you rate your overall Academic Performance?

Mark only one oval.

- Very poor
- Poor
- Average
- Good
- Excellent

20. Is there anything else you would like to share about your internet * usage, study habits, or academic performance?

This content is neither created nor endorsed by Google.