

Egypt University of Informatics

Computer and Information Systems

Data Analysis Course

The Analysis of the Performance of Data Analysis Students

Submitted by: Ahmed Wael

19/3/2025

# Introduction

Understanding the relationship between study time and academic performance is crucial for improving students' learning outcomes. This study aims to analyse how the amount of time students dedicates to studying outside of class affects their GPA. By examining this relationship, we can determine whether more study hours lead to better academic performance and whether other factors influence student success.

# Research Question

Does the amount of time spent studying outside of class have a significant impact on students' GPA?

# Hypothesis

We hypothesize that students who study more hours outside of class will have a higher GPA compared to those who study less. However, we also consider that study methods and consistency might play a significant role in performance.

# Population of Interest:

Computer Science students across different academic years at Egypt University of Informatics.

# Sampling Method:

The survey was conducted using a **convenience sampling method**, where we shared the form link in CS student groups covering all academic levels. This method was chosen due to its accessibility and efficiency in collecting responses. However, it has limitations as it does not ensure a completely random or representative sample of the entire student population.

# Bias Identification:

To minimize bias, we designed clear and neutral questions to avoid leading responses. However, some biases may still exist:

* **Self-Reporting Bias:** Since students report their own GPA and study habits, there is a possibility of inaccurate responses.
* **Selection Bias:** The survey was shared online, meaning students who actively check their university groups were more likely to participate. This could exclude students with different study habits.
* **Limited Sample Representation:** Not all CS students participated, so results may not fully represent the entire department.

# Survey Questions:

1. **What is your current cumulative GPA?**
2. **How many hours do you study outside of class per day?**
3. **How frequently do you study before exams?**
4. **What study methods do you use (e.g., self-study, group study, online courses)?**
5. **How often do you take breaks while studying?**

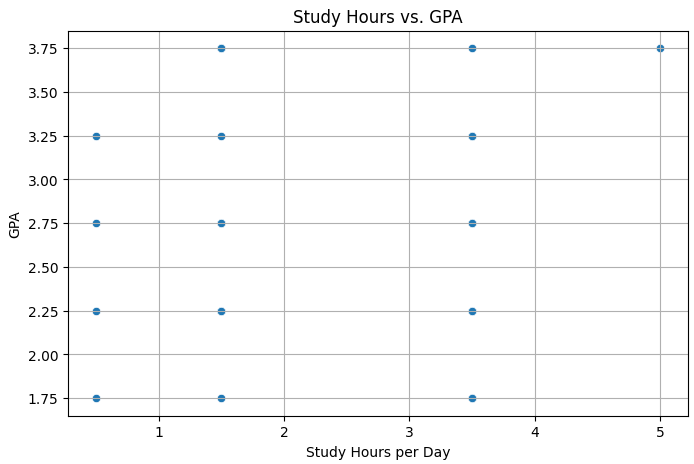
Online survey link: https://forms.gle/X8ymXR3s412j5L2M7

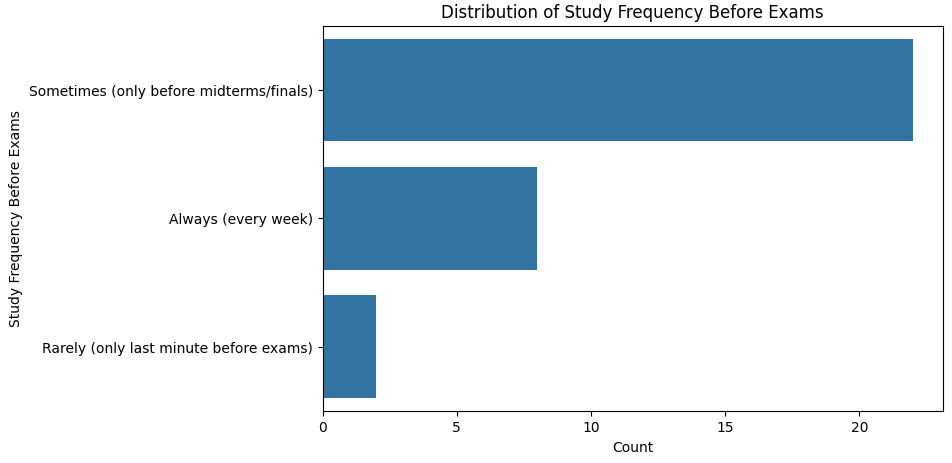
Number of samples collected: 32

# Analysis:

To analyse the collected data, we used statistical methods such as mean, median, and correlation analysis. The key focus was on identifying patterns between study hours and GPA

This one is wrong because when I collected the data, I should have collected the data as numbers not as a categorical (The GPA)

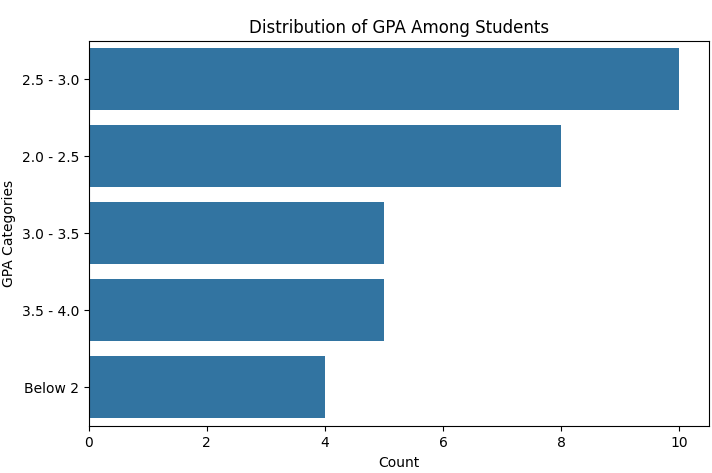




This bar chart illustrates the distribution of students based on how frequently they study before exams. The x-axis represents the number of students, while the y-axis lists different study frequency categories.

* The largest group consists of students who study **"Sometimes (only before midterms/finals)"**, indicating that most students tend to prepare closer to major exams rather than consistently throughout the semester.
* The second-largest group comprises students who study **"Always (every week)"**, suggesting a smaller but significant number maintain a regular study routine.
* The smallest group consists of students who study **"Rarely (only last minute before exams)"**, showing that very few students rely solely on last-minute cramming.

This visualization highlights that while some students maintain a consistent study schedule, the majority adopt a periodic or last-minute approach.



This bar chart represents the distribution of students across different GPA categories. The x-axis denotes the count of students, while the y-axis lists the GPA categories. The longest bar corresponds to the GPA range **2.5 - 3.0**, indicating that most students fall within this range. The second most common category is **2.0 - 2.5**, followed by **3.5 - 4.0** and **3.0 - 3.5**, which have similar counts. The smallest group consists of students with a GPA **below 2**. This visualization effectively shows the overall GPA distribution, highlighting the most and least common categories.

# Conclusion

The results suggest that students who study consistently tend to have higher GPAs, but excessive study hours do not always guarantee better performance. Study quality and effective learning techniques seem to matter more than just the number of hours studied.

# Any potential issues

1. **Self-Reported Data:** Some students may have misreported their GPA or study hours.
2. **Sampling Bias:** Convenience sampling may not represent all CS students equally.
3. **External Factors Not Considered:** The survey did not account for stress, sleep patterns, or external distractions that might affect GPA.
4. Wrong data collection: I collected all data as categorical only while I wanted some numerical values to collect some data as mean median and some graphs