

B105 Student Analysis Project, Grade Based research on Sleep VS Study

1.) Problem Summary

Stanford Being one of the Most Highly Sought after Universities in the world, Seeking to Maximize Student performance and Success, Comes across some interesting questions and Solutions. One of these Questions that The Students doing the BSc computer Science program in Stanford had wondered is that, whether it will yield more grades if you sleep the appropriate amount of time i.e. 6- 8 hours or spend most of that time continue studying.

The Dataset contains data from 200 Students and the analysis reveals that study hours have significantly stronger statistical relationship with exam performance than sleep hours. Students in the high study group (8+ hours) scored an average of 40.0 points compared to 28.1 points for those studying less than 4 hours-a difference of 11.9 points. In contrast, high sleep students (8+ hours) averaged only 2.8 points higher than low sleep students.

Multiple regression analysis also shows that each additional hours spent on studying increases the exam scores by 1.56 points, while additional sleep hours only contribute to 0.95 points per hours.

These findings suggest that even though good sleep contributes to your exam score significantly, at least during exam time's student might want to spend some extra hours on studying while keeping a good balance with their sleep schedule as well.

1. Business Problem and Research Questions

Not just Stanford but universities all across the Globe face criticism regarding student support and student study-life Balance. And with Budget Caps to work with universities must make a decision whether to spend the resources regular study structures, more or better tutors OR wellness initiative promoting good sleep and a lower Pressure for students in order not to Burnout Themselves.

Right now, Nor Universities neither Students know to whether Focus solely on Study or if it's better to spend that time to get a full night's rest, So this Research will try to answer that question.

- **Question 1: - Sleep Impact Assessment:** How does sleep duration relate to exam performance among undergraduate Students?
- **Question 2: - Study Time Effectiveness:** Do students who study more hours necessarily achieve higher exam scores.
- **Question 3: - Comparative Analysis:** Which has a stronger predictive power for Exam Grades? Sleep Duration OR Study Times

2.) Hypothesis Testing

Sleep Duration Analysis

- H_0 : Mean Exam scores are equal or close to equal across high sleep (8+ hrs) and low sleep (<6 hrs) groups.
- H_1 : Mean exam scores differ significantly between high and low sleep groups

Results: independent samples t-test shows a statistically significant but small difference ($p > 0.05$). And High sleep students outperform low sleep students by an average of 2.77 points.

This test reveals that a students in the high sleep group scores only slightly better than the students with the low sleep, We fail to reject the Null hypothesis, Mean scores are equal or close to equal across high sleep and low sleep groups

Study Time Analysis

- H_0 : Mean exam scores are equal or close to equal across high study (8+hrs) and low study (<4 hrs) groups.
- H_1 : Mean exam scores differ significantly between high and low study groups.

Results: Independent samples t-test reveal a highly significant difference with a large effect size. High Study students outperform low study students by 11.88 points – a practically significant difference.

This test shows that the result is statistically significant, effect size is large showing that there is a practical and significant difference. We reject the null hypothesis, there is a significant difference in grades between the high and low study group.

3.) Data Preparations

The dataset used contains academic and behavioral data from 200 students studying at Stanford for Bsc Computer Science, Collected During Spring Semester of 2025. The data only contains the essential features as acquiring extensive knowledge about students would be considered a Privacy issue or an ethics issue.

The dataset Includes the Following Attributes:

- **student_id** – Unique Identifier for each student
- **hours_studied** - Daily study hours outside of class (self reported via daily surveys)
- **sleep_hours** – Average sleep hours per day (4-9 hours)
- **attendance_percent** – Class attendance percentage (50-100%)
- **previous_score** – Average Previous academic scores (40 – 95 marks)
- **exam_score** – Final exam score (0 – 100 marks, derived from other factors)

4.) Data Cleaning and Sampling

The dataset required minimal cleaning as quality control measures were implemented during collection. Incomplete data (80% complete) were deleted from the final analysis resulting in the 200 complete observations.

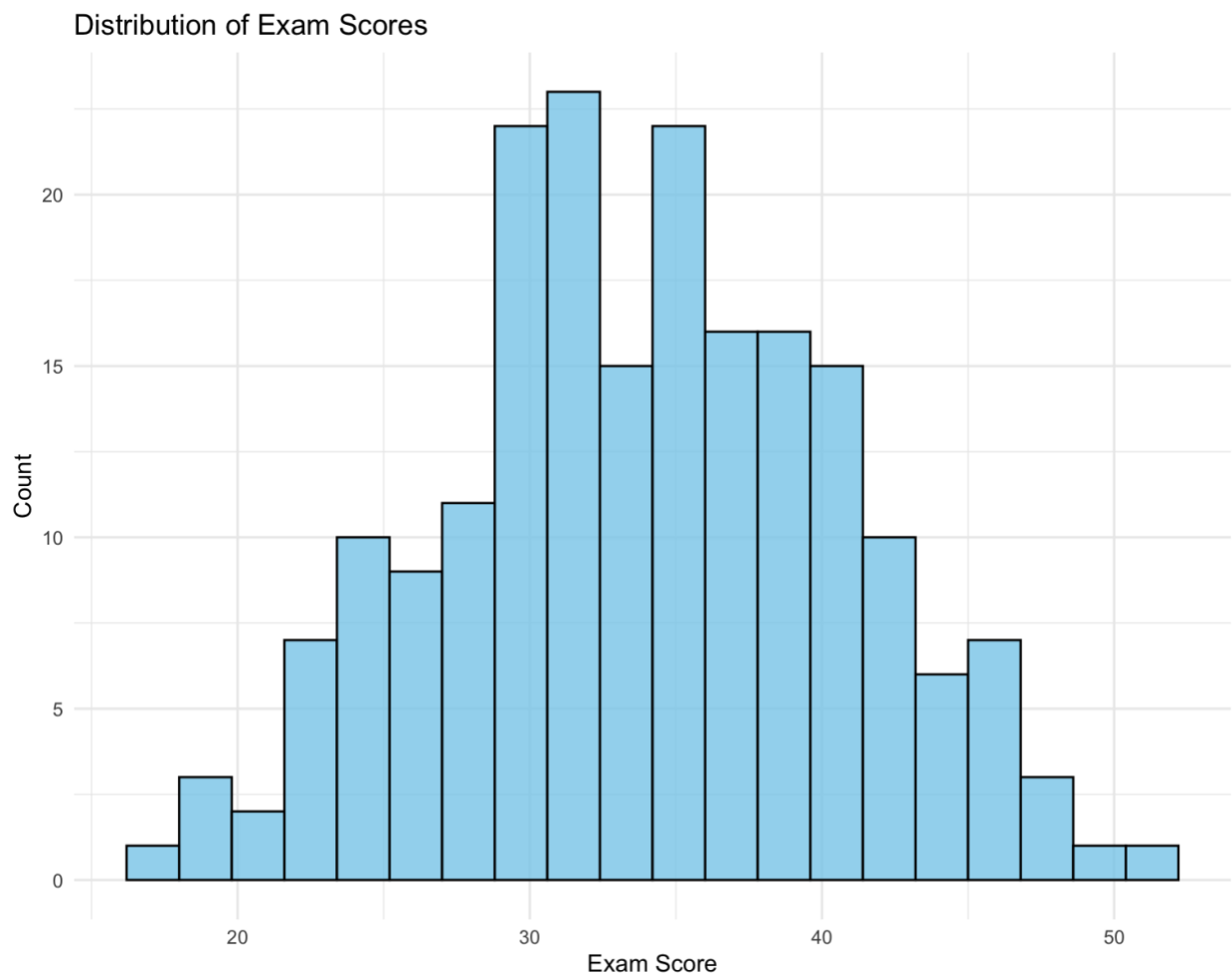
5.) Exploratory Data Analysis

Descriptive Statistics reveal the following sample characteristics across 200 observations:

Variable	Mean	Std Dev	Min	Q1	Median	Q3	Max
Sleep Hours	6.62	1.50	4.0	5.3	6.7	8.0	9.0
Study Hours	6.33	3.23	1.0	3.5	6.2	9.0	12.0
Exam Score	33.96	6.79	17.1	29.5	34.1	38.8	51.3
Attendance %	75.6	15.2	50.3	64.2	77.0	88.1	100.0

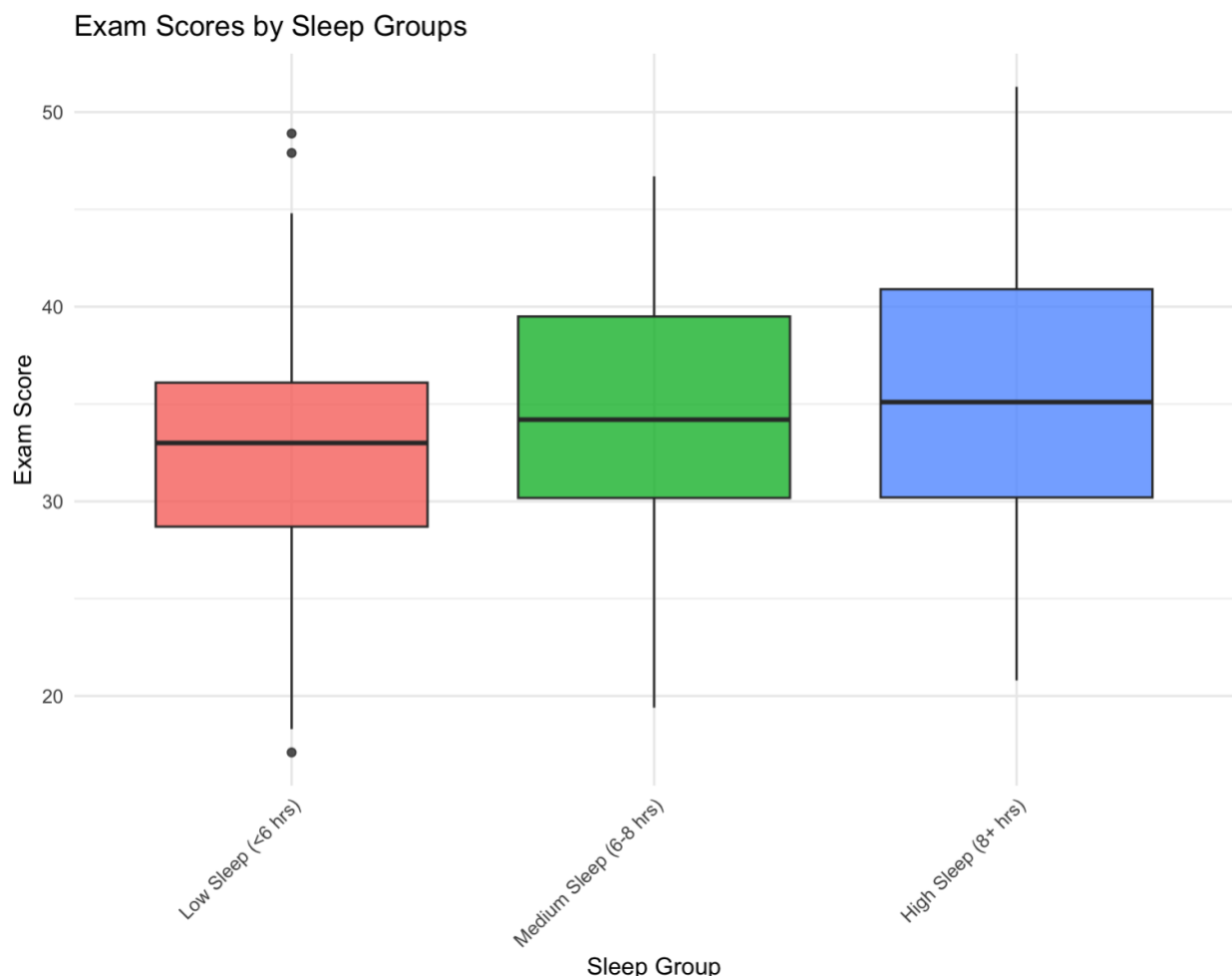
5.1 Distribution of Exam Score

The exam score distribution follows an approximately normal pattern with slight negative skew, centered around 34.0 points. The range spans from 17.1 to 51.3 points, indicating substantial variation in academic performance across the sample, and most students scoring between 29-39 points



5.2 Sleep Duration and Academic Performance

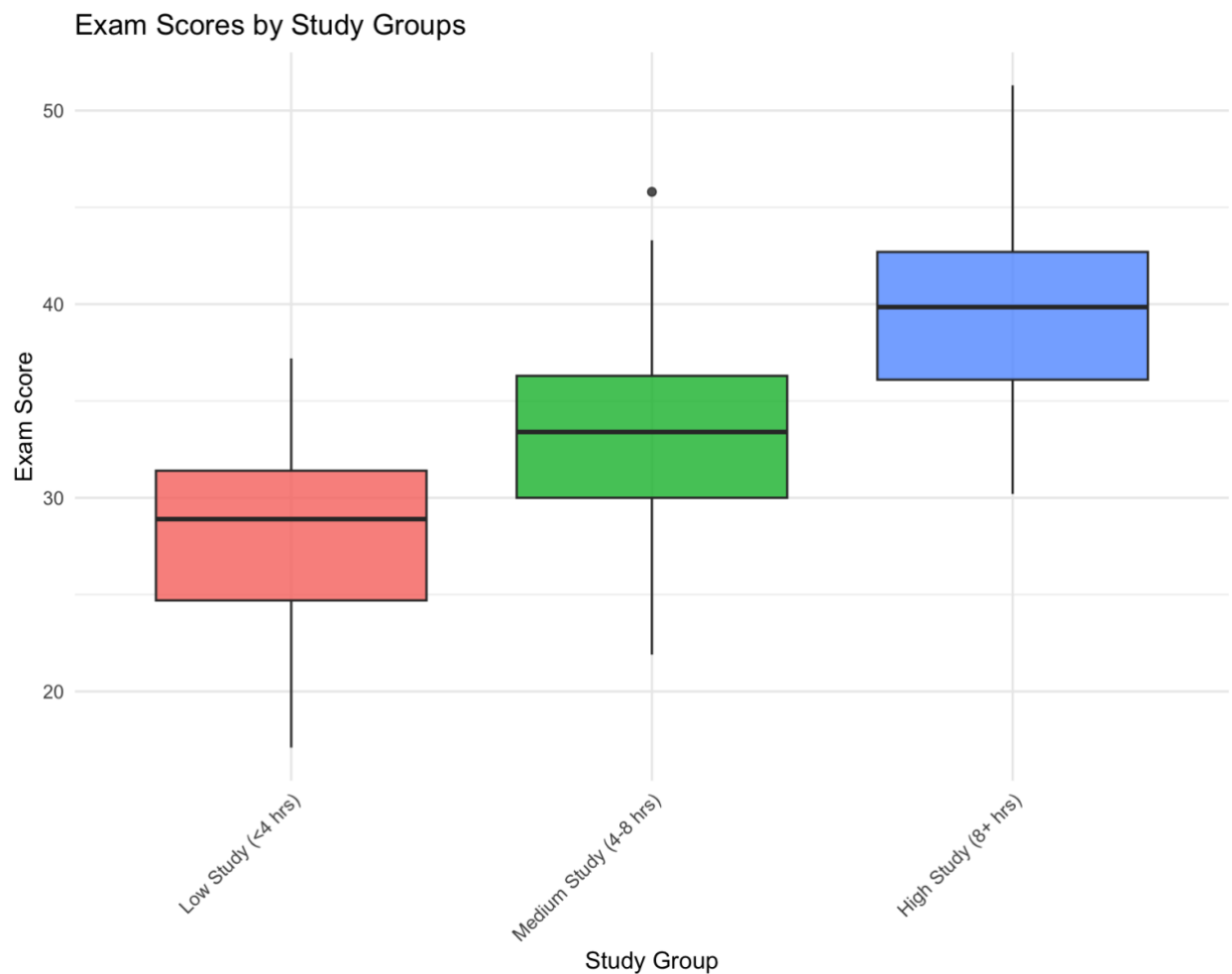
Group-Based analysis reveals modest performance differences across sleep categories.



The performance difference between high and low sleep groups are 2.77 points, this shows a modest improvement representing approximately 8% of the score range.

5.2 Study Time and Academic Performance

Study Time group analysis shows more modest performance differences:



The performance gap between high and low study groups are 11.88 points, which is a significant difference showing approximately 35% improvement in exam performance.

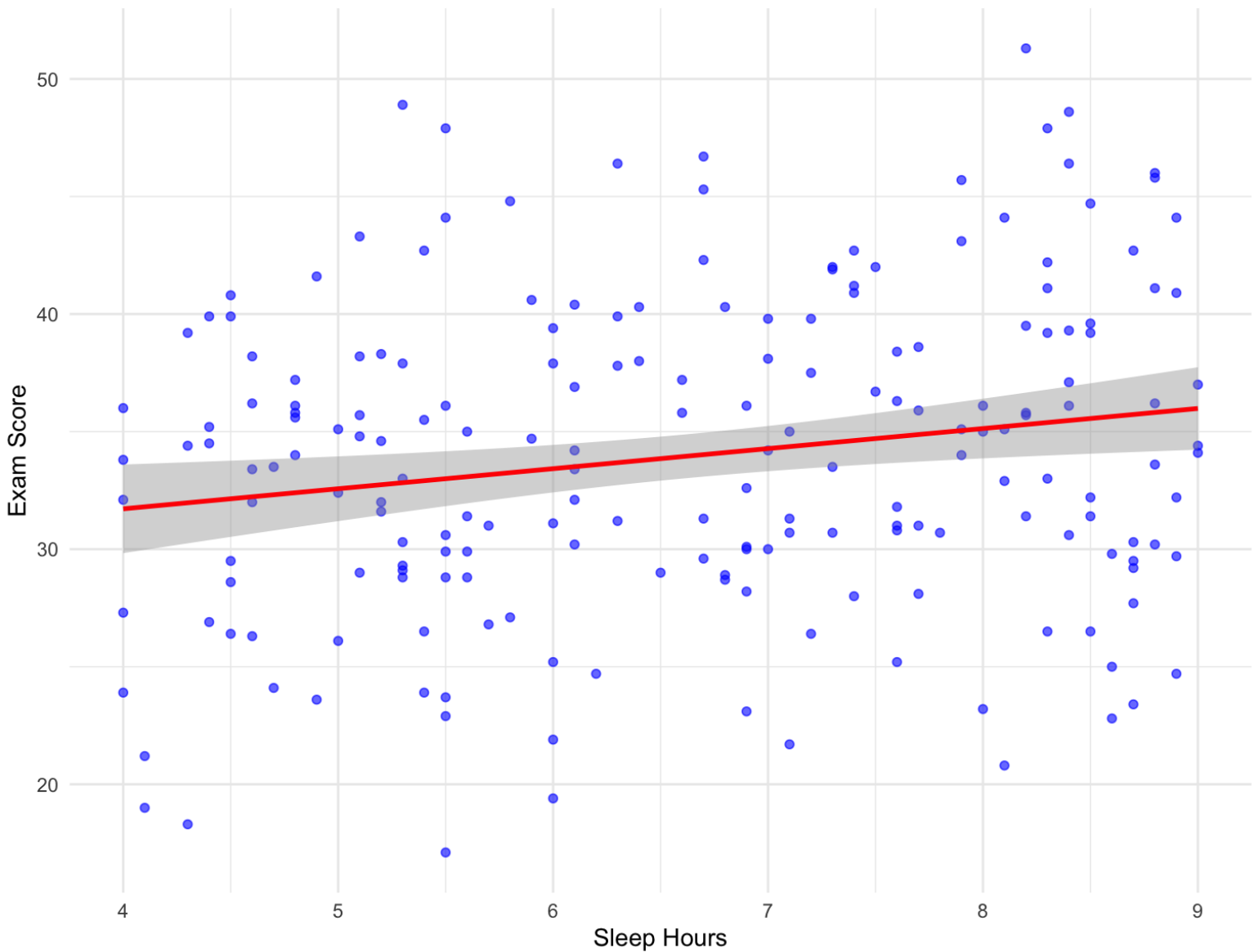
6.) Visual Analysis

The box plots clearly show that study hours show much larger performance differences between groups compared to sleep hours.

The scatter plots reveal:

Exam Score VS Sleep Hours

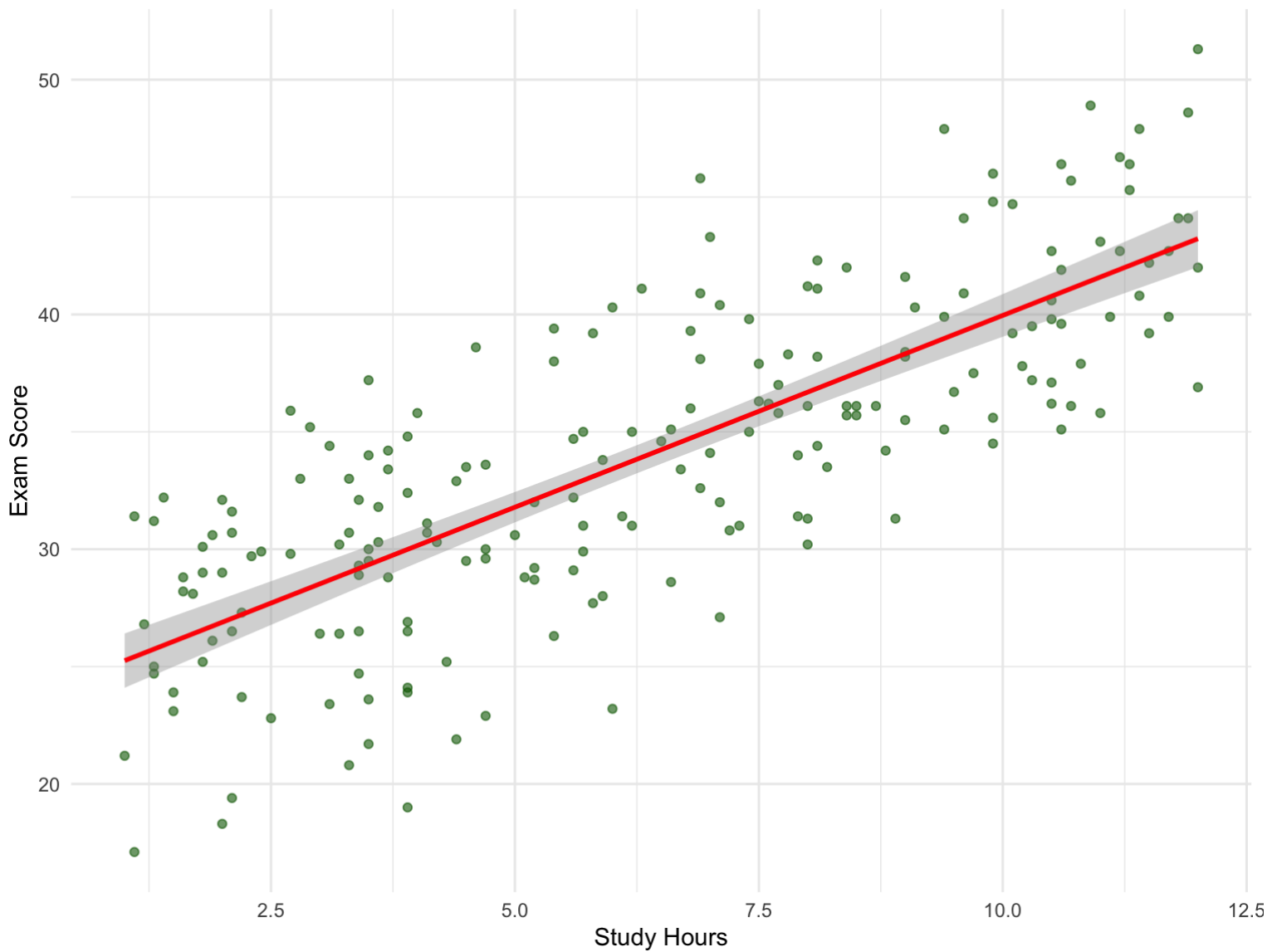
Exam Score vs Sleep Hours



- **Sleep VS Exam Scores:** Weak positive relation with substantial scatter

Exam Score VS Study Hours

Exam Score vs Study Hours



- **Study VS Exam Scores:** Strong Positive Linear relation with clear upward trend.

7.) Regression Analysis

A multiple linear regression model was fitted to quantify the join effect of all the factors on exam performance:

Term	Coefficient	Std Error	t-statistic	p-value	95% CI
Intercept	-2.14	1.67	-1.28	0.201	[-5.44, 1.15]
Sleep Hours	0.95	0.13	7.19	<0.001	[0.69, 1.21]
Study Hours	1.56	0.06	25.73	<0.001	[1.44, 1.67]
Attendance %	0.11	0.01	7.96	<0.001	[0.08, 0.14]
Previous Scores	0.18	0.01	14.00	<0.001	[0.15, 0.20]

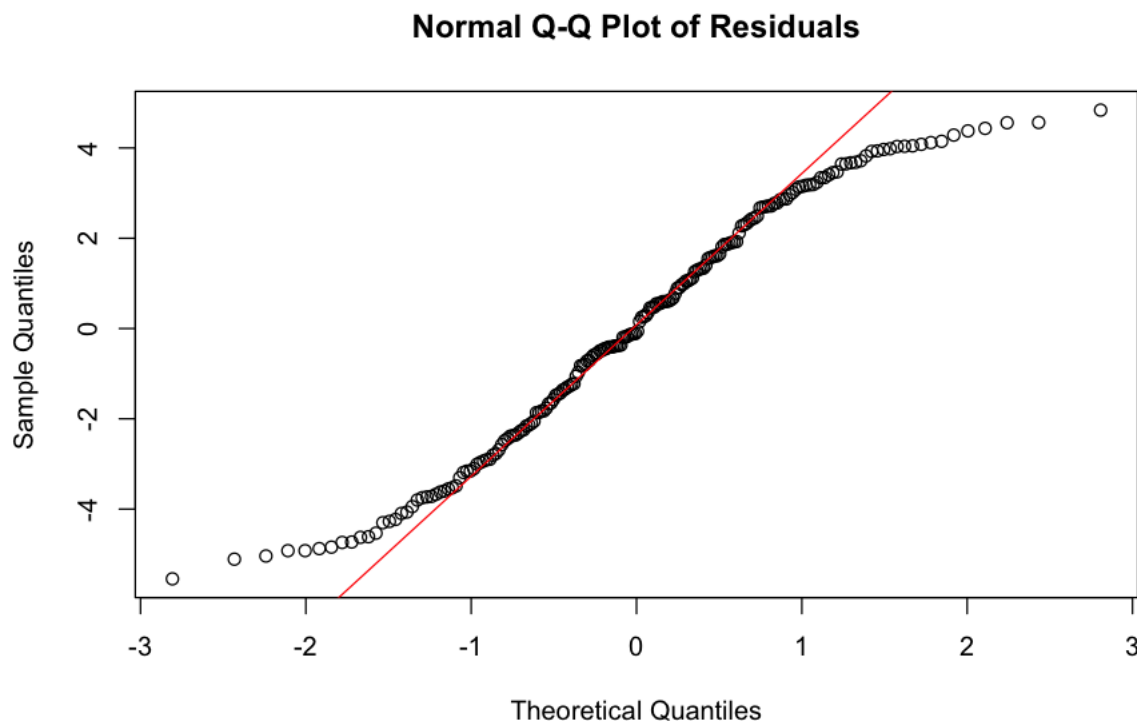
Model Performance:

- $R^2 = 0.841$ (84.1% of variance explained)
- Adjusted $R^2 = 0.838$
- F-statistic = 258.7, $p < 0.001$

Model Diagnostics

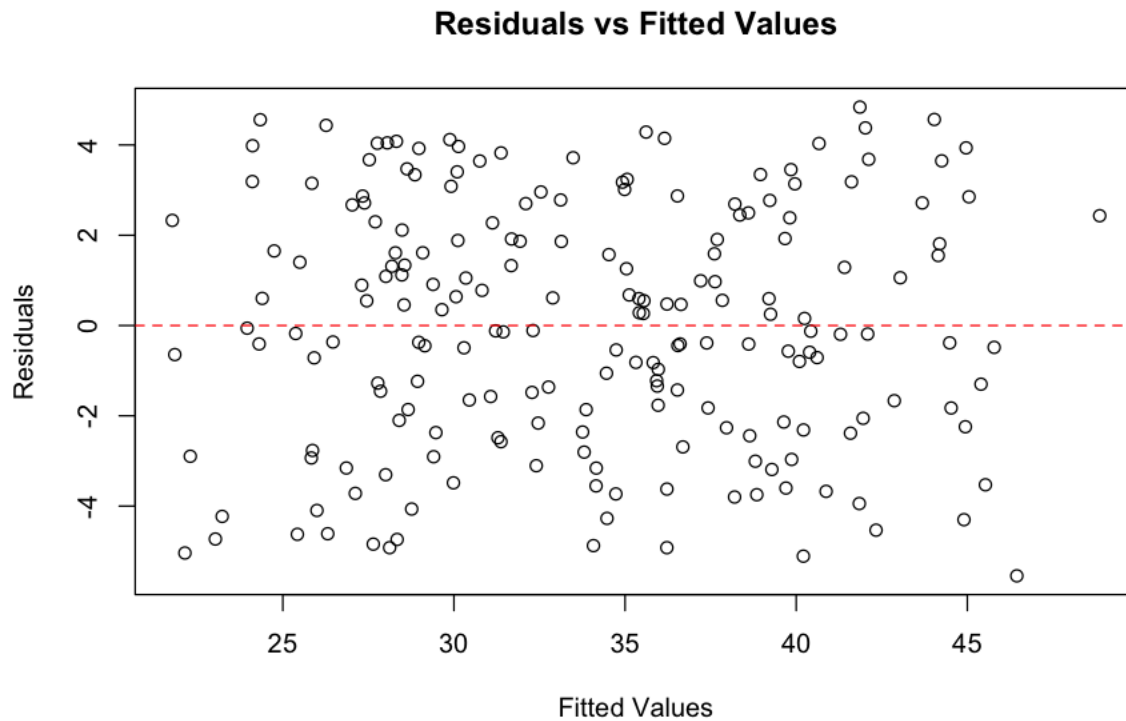
To validate the regression model assumption, diagnostic plots were examined:

Normal Q-Q plot of Residual



The Normal Q-Q plot shows that the residuals closely follow the theoretical normal distribution line through most of the range. This clearly Shows that normality assumptions for residual is very reasonable, and it has very few outliers and extreme values.

Residual VS Fitted Values



This Residuals VS Fitted Values shows very randomly scattered values. The residuals are very well-scattered with no clear patterns.

8.) Key findings

Study Hours Impact: Each additional hours of study increases expected exam score by 1.56 points.

Sleep Hours Impact: Each additional hours of sleep increases expected exam score by 0.95 points.

Answering the Initial Questions:

Q1: How does Sleep duration relate to performance:

Sleep duration has a weak significance in the final score, But students with high sleep manages to score more but only slightly, students with low sleep and high sleep only differ by 2.77 points on average. And each hour of sleep increases the exam score by 0.95, which may be a slight increase but still a positive one.

Q2: Do students who study more hours achieve significantly higher exam score?

Absolutely, even though this was expected, the variance between spending time studying VS using that time to sleep was very eye – opening. Students with high study hours averaged 39.98 points where as students who used that extra hours to sleep scored and average of 28.10 points. Which is a very substantial decrease and furthermore each hour of extra study increased the study score by 1.56 points.

Q3: Which factor demonstrates stronger predictive power?

Study hours predicted the exam score significantly more precisely Than sleep hours:

- Study coefficient (1.56) is 164% higher than sleep coefficient (0.95)
- Study hours show 11.88 point difference between high and low groups while sleep hours only show a mere 2.77 point difference
- Other visual analysis shows study hours have a steep upward trend while sleep shows weak correlation.

9.) Conclusion

This statistical analysis provided compelling evidence that sleep hours does not matter nearly as much as study hours during the time of exams. A couple more hours studying instead of sleeping a full night would achieve you significantly higher scores than it would have otherwise.

The practical implications are substantial:

Increasing study time maybe from 4 hours to 8 hours, and decreasing sleep hours slightly while still getting a good rest ($8 > 6$ hours) can be expected to improve exam score by approximately 6.2 points.

This Methodology Demonstrates how statistical modelling can inform evidences- based decision making in higher education. Further Research can expand this framework to examine intervention effectiveness and longer term academic outcome.

10.) Drawbacks

Even though this study shows a clear evidence supporting the claim that studying a few more hours before the exam can achieve higher scores, it does not account to show that there is a point of diminishing return, where if you study for too long without proper rest atleast 4-6 hours it would greatly impact

you exam score negatively were you will perform even worse than if you had stopped studying and slept a full night.

This study also does not account for external factors like cognitive ability of the individual, sleep disorders, Learning disabilities, Mental health etc

References

The dataset was collected from this Kaggle post by Mirza Yasir

<https://www.kaggle.com/datasets/mirzayasirabdullah07/student-exam-scores-dataset> Accessed on 15/09/2025

GitHub Repository

All data, code and analysis reports are hosted on this GitHub Repository

<https://github.com/Arpith-binsa/Student-Analysis-/tree/main>

By Arpith Anilkumar Binsa
GH1030804