Student Performance Analysis Report

By Mehul Bhochiya

Content

Introduction

Numerous circumstances outside of the classroom have an impact on academic success. A student's learning outcomes are greatly influenced by factors like motivation, lifestyle, school environment, family history, and study habits. Understanding how these factors interact with academic achievement is essential for developing effective educational strategies and support systems.

This analysis is based on a student performance dataset containing information from 6,607 students and 20 attributes related to their academic behavior, personal habits, family and economic background, and learning environment. The central outcome variable in this dataset is Exam Score, which represents the final academic performance of students. The dataset combines both numerical and categorical variables, allowing for a comprehensive analysis of both quantitative trends and qualitative influences on student outcomes.

This analysis's main goal is to pinpoint and assess the critical elements that affect students' exam performance. Exam scores are specifically examined in relation to factors including hours studied, attendance, sleep patterns, motivation levels, parental participation, teacher quality, access to resources, and economic status. The report aims to identify the variables that are most strongly positively or negatively correlated with academic success by examining these correlations.

These insights can be used by educators to provide focused interventions for kids who are at risk of performing poorly. Parents are better able to comprehend how their children's academic progress is influenced by their involvement, motivation, and home environment.Additionally, the results provide a strong foundation for further applications such as student performance prediction models and early-warning systems.

Overall, this report seeks to transform raw educational data into actionable insights that support informed decision-making and contribute to improved student learning outcomes.

Problem Definition

Students' academic achievement is difficult to understand because there are so many interrelated factors that influence it, including study habits, family support, the school environment, lifestyle choices, and access to resources. These variables are often considered in isolation; hence, their real impact on test results and the need for targeted treatments are hard to decipher.

The core problem that this analysis addresses is a lack of measurable insights on how different academic, social, and environmental factors affect the students' exam performance. More precisely, the goal of this study is to extend beyond assumptions by providing empirical evidence of how various student-related attributes influence the target variable, Exam_Score.

To solve this overall problem, the review is organized around five specific sub-problems:

1. Study Habits Impact on Academic Performance:
Students display variation in study behaviors across hours studied, attendance, tutoring participation, and levels of motivation. The first problem concerns determining the extent to which study habits directly impact exam performance and which specific study-related behaviors yield higher scores.

2. Family Background and Student Success:
The family environment influences the students' academic career significantly. This problem probes parental involvement in the student's life, family income, and the education levels of parents; whether or not greater family support will necessarily lead to the best possible results in students' school performance.

3. School Environment versus Exam Score:
The quality of the school environment can greatly affect the effectiveness of learning. This problem aims at assessing how exam performance varies with regard to school type, teacher quality, peer influence, and access to learning resources all showing how institutional and social settings affect student achievement.

4. Lifestyle Analysis: Sleep, Physical Activity & Academic Scores:
Outside of the classroom, lifestyle choices that students make in their daily lives can impact their ability to focus and learn effectively. This problem explores the relationship between sleeping habits, exercise, access to the internet, and exam performance in an attempt to answer the question: are healthy behaviors associated with better academic performance?

5. Predicting Top Factors Affecting Student Exam Score:
Where there are lots of interrelated variables, determining which ones really make a difference becomes critical. The last problem involves the ranking of and identification of the best predictors of exam performance through analytical techniques such as pivot tables and

interactive slicers that help in differentiating high-impact factors from those with minimal consequences.

This analysis aims to provide educators and parents with a structured, data-driven view of student performance regarding the problems identified in these five areas. This report helps them focus on areas that hold the most potential for impact on improved academic outcomes. The problem definition provides the basis for the methodical investigation, insight creation, and visualization discussed in the subsequent sections of this report.

Data Preparation

Preparing the dataset was done in advance of any analysis or the construction of dashboards, in order to ensure that the results are accurate, consistent, and reliable. In fact, proper data preparation is a very critical step, as any analytical insight is as good as the quality of the underlying data. This phase focuses on the cleaning and validation of data to make the dataset suitable for meaningful analysis.

Data Cleaning

First, inconsistencies and irrelevant entries in the original dataset needed examination. In order to ensure that every student observation was unique and not overrepresented in the analysis, duplicate records were determined and eliminated by using Excel's Remove Duplicates feature.

Then, any missing or blank values were taken care of by selecting all blank cells using Ctrl + G → Special → Blanks and deleting the rows that corresponded to those cells using Ctrl + –. Such incomplete records would misrepresent averages, correlations, and comparisons during analysis.

Data Validation

The cleaned dataset was then validated for logical correctness and usability. All numerical columns were verified to confirm that no negative values exist in fields where negative values are not logically possible, such as hours studied, attendance percentage, sleep hours, and exam scores. In addition, new categories were derived to support the analysis better and visualizations of the data. This included Range_Studied_Hours, which grouped study hours into more meaningful ranges; Attendance_Range, the categorization of attendance percentages; and Performance_Category, a classification based on exam score levels. These derived fields thus enabled proper segmentation, comparison, and interpretation of student performance patterns. Preparation of the data was done in a way that cleaned, validated, and structured the dataset to provide a very strong base for accurate analyses, reliable dashboards, and data-driven insights that follow later in this report.

Exploratory Data Analysis & Charts

EDA was performed with the main aim of scrutinizing the dataset systematically to uncover meaningful patterns, relationships, and trends that may influence student academic performance. The basis of this stage, therefore, has been thoroughly understanding how individual factors and their respective interactions relate to the target variable Exam_Score prior to finalizing dashboards and conclusions.

In this regard, a total of 20 pivot tables along with charts were built to study the relations among major academic, demographic, and environmental variables as part of the process for EDA. In most of the analyses, pivot tables were used to summarize large volumes of data, compute group-wise averages, and thus facilitated comparison across categories. Charts were thereafter derived from the pivot tables to identify visually the trends and relative performance.

Pivot-based analyses were performed as follows:

- Relationship Between Hours Studied and Exam Score
- Impact of Attendance Percentage on Exam Score
- Effect of Tutoring Sessions on Exam Score
- Influence of Sleeping Hours on Exam Score
- Impact of Motivation Level on Exam Score
- Relationship between Parental Education Level and Previous Scores
- Relationship between Parental Education Level and Exam Score
- Effect of Family Income on Exam Score
- Association between Parental Involvement and Motivation Level
- Combined Effect of Teacher Quality and Access to Resources on Exam Score
- Comparison of School Type (Public vs Private) and Exam Score
- Impact of Teacher Quality on Exam Score
- Relationship between Access to Resources and Previous Scores
- Effect of Access to Resources on Exam Score
- Influence of Peer Influence on Exam Score
- Analysis of Sleep Hours Across Exam Score Levels
- Impact of Internet Access on Exam Score
- Relationship between Physical Activity and Exam Score
- Effect of Learning Disabilities on Exam Score
- Comparison of Performance Category and Hours Studied

Detailed analyses were conducted to check the individual and combined effects of variables through these pivot tables and visualizations. The numerical variables were analyzed by looking at grouped averages and trends, while the categorical variables were examined in terms of frequency distribution and comparative bar and column charts. This helped in

finding dominant performance drivers, potential risk factors, and segments of students showing invariably better or worse academic outcomes.

The EDA phase was critical in confirming assumptions, locating significant relationships, and underlining aspects that needed further analysis in more depth. The insights developed in this stage informed the choice of KPIs, the structuring of the dashboard, and the ultimate interpretation of results. In summary, EDA laid a sound analytical framework for raw educational data to be turned into useful insights, as will be represented in further sections of this report.

Dashboards

To effectively communicate insights derived from the analysis, a set of five interactive dashboards was designed, each aligned with one of the defined problem statements. The dashboards were built using Excel pivot charts and slicers to enable dynamic filtering, comparative analysis, and deeper exploration of student performance patterns. This modular dashboard approach ensures clarity while allowing stakeholders to focus on specific dimensions of academic performance.

Each dashboard contains a carefully selected combination of charts and slicers, enabling users to analyze both high-level trends and detailed subgroup variations.

A consistent black, white, and blue color scheme was used across all dashboards to ensure visual clarity and professional presentation. According to color theory, blue is associated with trust, focus, and analytical thinking, making it well-suited for data visualization. Black and white provide strong contrast, reduce visual noise, and improve readability, especially when displaying multiple charts and filters simultaneously.

All charts include clear titles, labeled axes, and data labels where appropriate to ensure that insights are easily interpretable without additional explanation. This design approach enhances usability, reduces cognitive load, and enables stakeholders to quickly derive insights from the dashboards.

Overall, the dashboard design emphasizes interactivity, clarity, and analytical depth, allowing users to explore student performance from multiple perspectives and supporting data-driven decision-making.

1. Impact of Study Habits on Academic Performance
This dashboard focuses on understanding how core study-related behaviors influence exam outcomes.

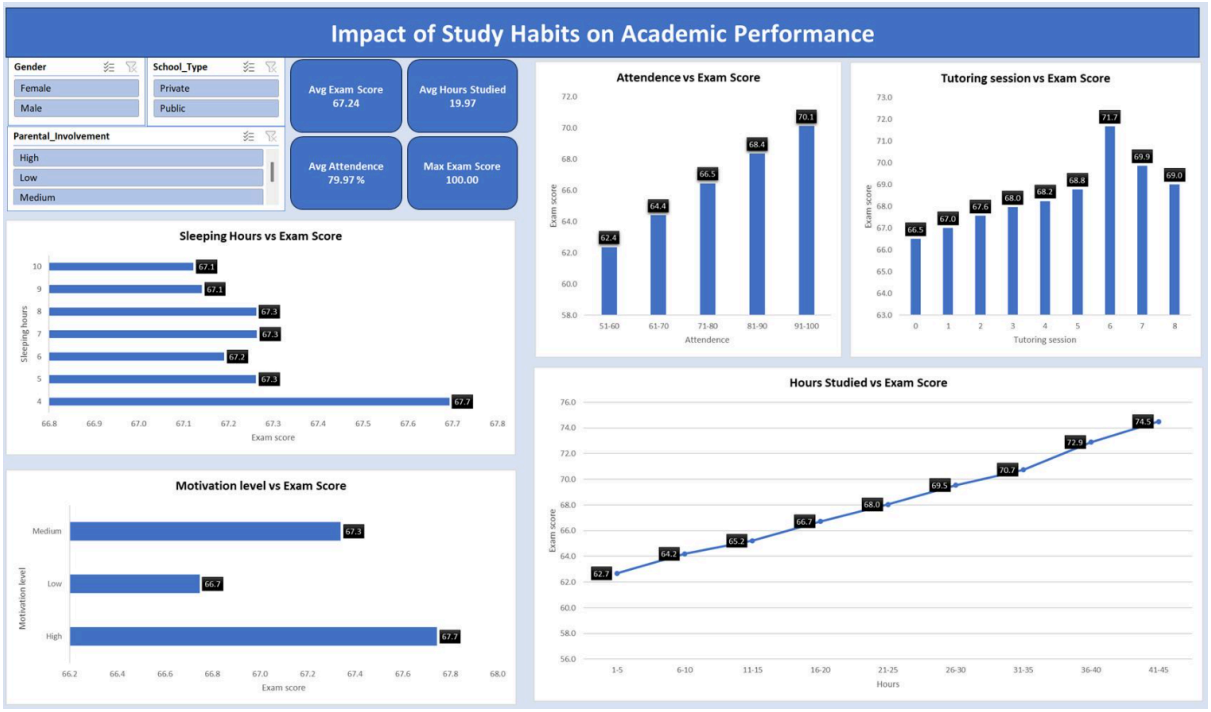
Charts (4):

- Hours Studied vs Exam Score
- Attendance vs Exam Score
- Tutoring Sessions vs Exam Score
- Sleeping Hours vs Exam Score

Slicers (3):

- Gender
- School Type
- Parental Involvement

These visualizations allow users to observe how changes in study behavior impact performance and how these effects differ across gender, school type, and levels of parental involvement.



2. Family Background Influence on Student Performance:
This dashboard examines the influence of family-related factors on academic outcomes.

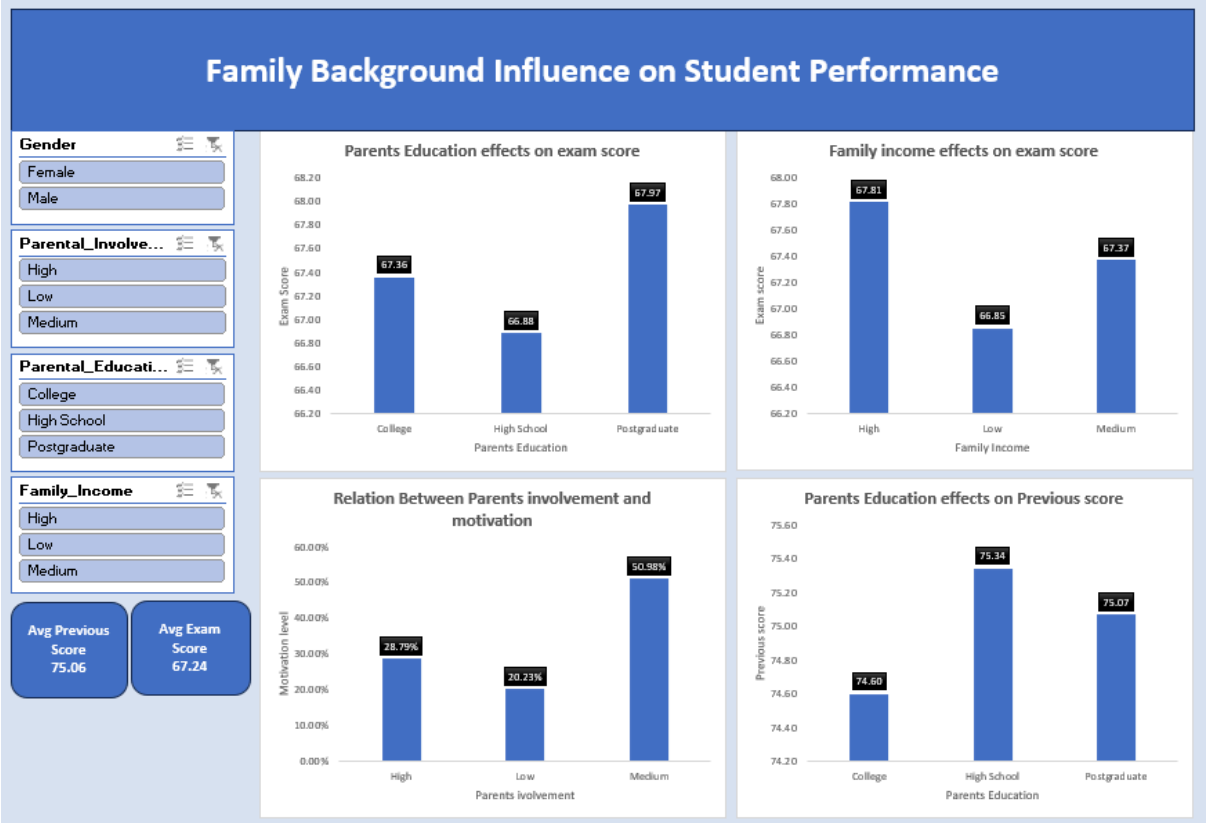
Charts (4):

- Motivation Level vs Exam Score
- Parental Education Level vs Previous Scores
- Parental Education Level vs Exam Score
- Family Income vs Exam Score

Slicers (4):

- Gender
- Parental Involvement
- Parental Education Level
- Family Income

The dashboard enables detailed comparison of academic performance across different socio-economic and educational family backgrounds.



3. School Environment & Exam Score:

This dashboard analyzes how institutional and social environments contribute to student performance.

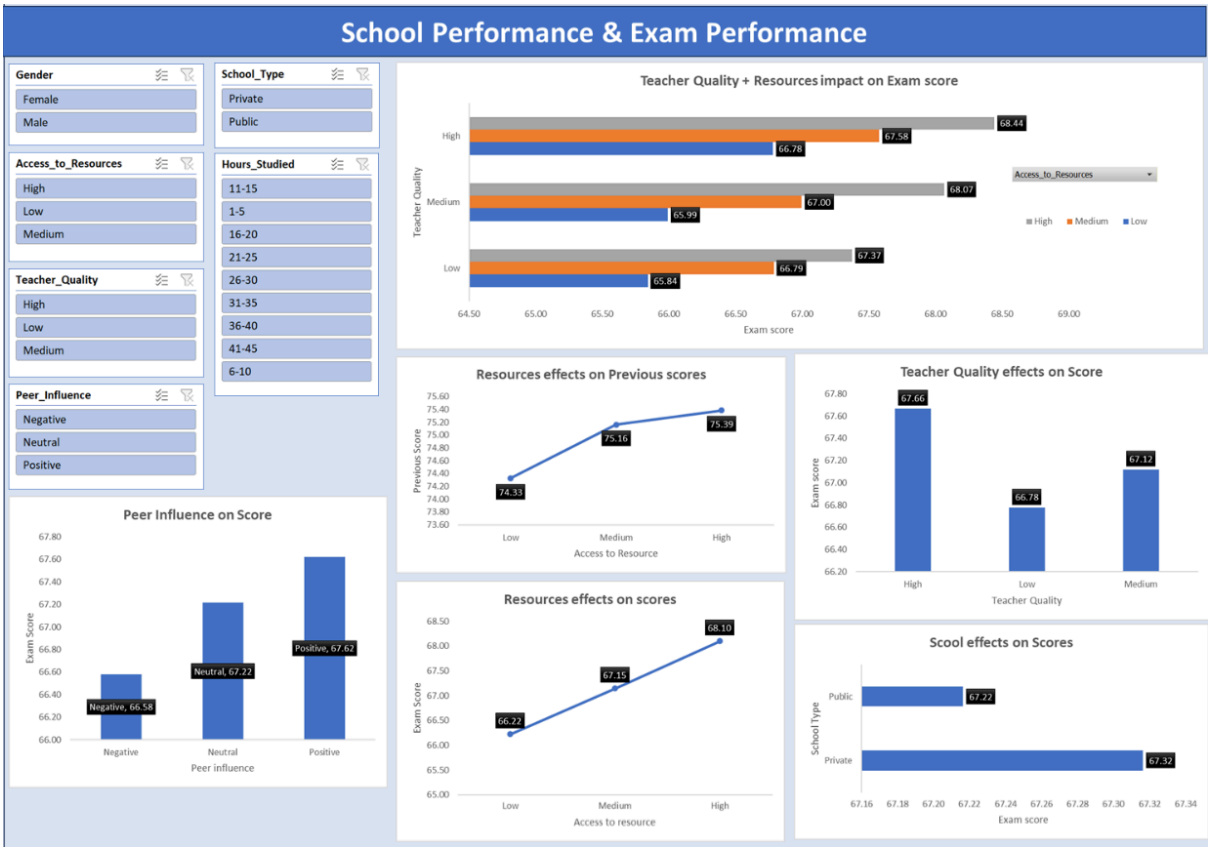
Charts (6):

- Parental Involvement vs Motivation Level
- Teacher Quality and Access to Resources vs Exam Score
- School Type vs Exam Score
- Teacher Quality vs Exam Score
- Access to Resources vs Previous Scores
- Access to Resources vs Exam Score
- Peer Influence vs Exam Score

Slicers (6):

- Gender
- School Type
- Hours Studied
- Access to Resources
- Teacher Quality
- Peer Influence

These visuals allow users to evaluate both individual and combined environmental effects on academic outcomes.



4. Lifestyle & Daily Habits impacts on Scores:

This dashboard focuses on non-academic lifestyle factors that may influence learning effectiveness.

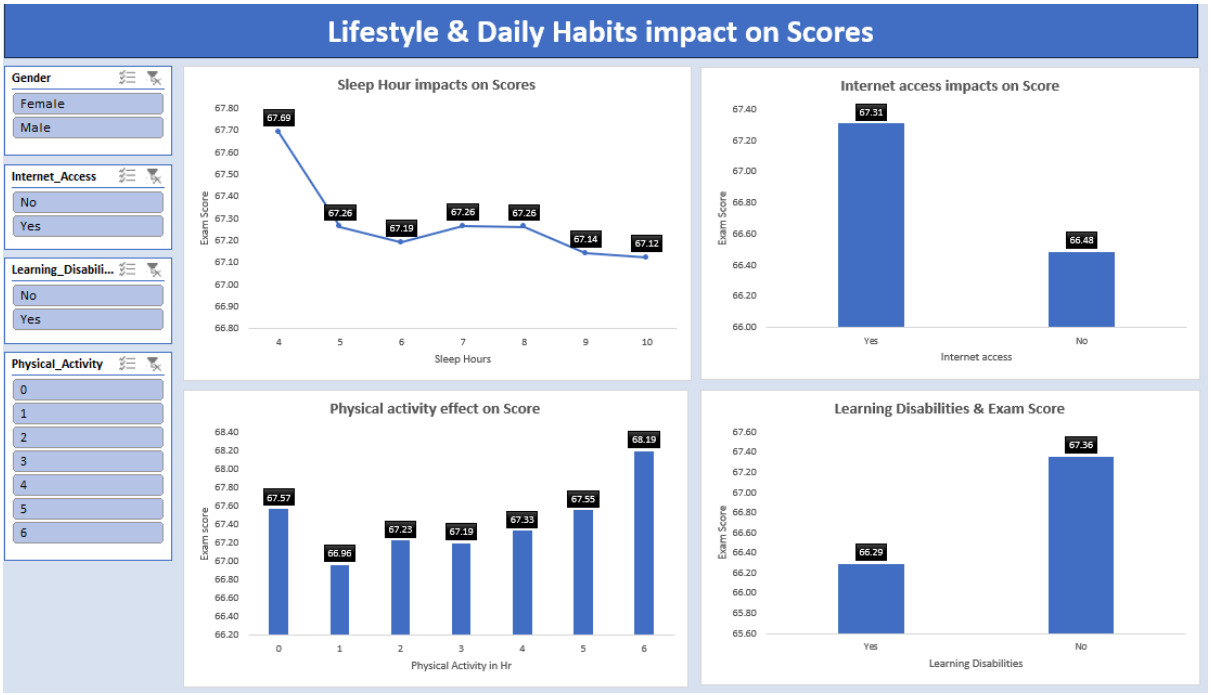
Charts (4):

- Sleep Hours vs Exam Score
- Internet Access vs Exam Score
- Physical Activity vs Exam Score
- Learning Disabilities vs Exam Score

Slicers (4):

- Gender
- Internet Access
- Learning Disabilities
- Physical Activity

The dashboard helps identify whether daily habits and health-related factors have measurable effects on academic performance.



5. Multi-Factor Analysis:

This dashboard integrates multiple variables to identify patterns associated with high and low academic performance.

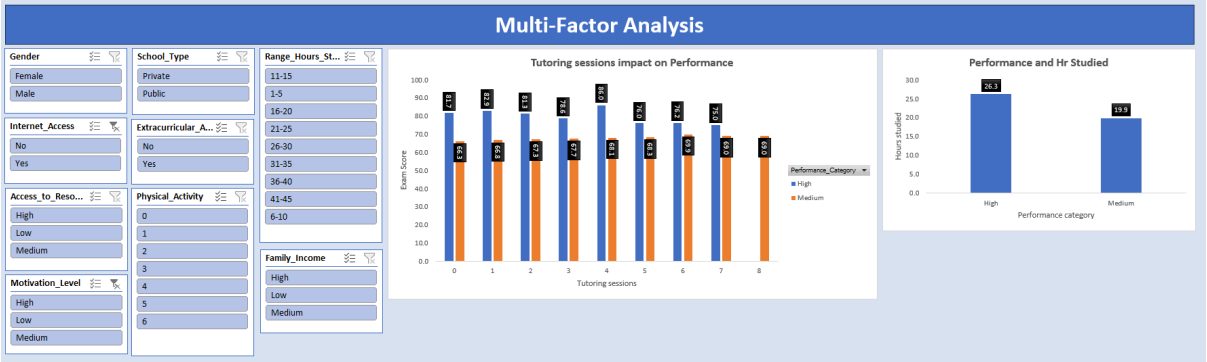
Charts (2):

- Performance Category vs Hours Studied
- Performance Category, Exam Score, and Tutoring Sessions

Slicers (9):

- Gender
- School Type
- Hours Studied
- Access to Resources
- Internet Access
- Motivation Level
- Extracurricular Activities
- Physical Activity
- Family Income

This dashboard enables advanced segmentation and comparative analysis, supporting the identification of key performance drivers and high-risk student groups.



Key Insights & Interpretation

1. Students studying 10 or more hours per week achieved a higher average exam score (67.39) compared to students studying less than 10 hours (63.81), indicating a moderate positive association (~5.6% improvement) between study time and academic performance.
2. Attendance percentage shows the strongest positive relationship with exam performance. Students with attendance between 90-100% achieved the highest average scores (70.13), while lower attendance groups showed progressively reduced performance.
3. Sleep duration shows no meaningful impact on exam scores. Average scores remain nearly constant across all sleep-hour groups, with a near-zero correlation (-0.016), indicating no clear performance advantage from increased sleep in this dataset.
4. Students attending tutoring sessions demonstrate slightly higher exam scores than those with no tutoring. The positive correlation (~0.154) suggests tutoring contributes to performance improvement, although the effect size is relatively small.
5. Motivation level exhibits a clear ordered relationship with academic performance. Students with high motivation consistently outperform those with medium and low motivation, highlighting motivation as a meaningful contributor to exam success.
6. High parental involvement is associated with better academic outcomes. Students receiving stronger parental support show higher average exam scores, reflecting a small to moderate positive influence on performance.
7. Family income has a positive relationship with exam scores, where students from higher-income households tend to perform better. However, its influence is weaker than core academic factors such as attendance and study hours.
8. Students whose parents have higher educational qualifications achieve higher exam scores, with performance improving progressively from high school to postgraduate parental education levels.
9. Previous academic performance is positively influenced by parental education, as students with more highly educated parents demonstrate higher previous scores, reinforcing the long-term academic impact of parental background.
10. Students with high parental involvement exhibit a higher proportion of high motivation levels, indicating a strong link between parental support and student motivation.
11. Students enrolled in private schools show a slight advantage in average exam scores compared to public school students, although the difference is relatively small.
12. Teacher quality has a noticeable positive impact on exam performance. Students taught by high-quality teachers consistently achieve higher average scores, though this effect remains smaller than attendance and study-time influences.
13. Peer influence plays an important role in academic outcomes. Students exposed to positive peer environments achieve higher exam scores than those with neutral or negative peer influence.

14. Better access to learning resources is associated with higher exam scores, showing a small to moderate positive effect on academic performance.
15. Combined analysis reveals that students experiencing both high teacher quality and high access to resources achieve some of the highest average exam scores (e.g., 68.44), demonstrating the importance of supportive learning environments.
16. Re-analysis of sleep patterns confirms that sleep hours alone do not significantly affect exam scores, with averages remaining stable across groups.
17. Physical activity levels show minimal variation in exam performance, suggesting that physical activity does not have a measurable impact on academic scores in this dataset.
18. Students with internet access perform slightly better on average than those without access, indicating that internet availability likely supports learning through improved access to educational resources.
19. Students with learning disabilities score modestly lower on average compared to students without learning disabilities, highlighting the need for additional academic support mechanisms.
20. High-performing students (Exam Score ≥ 75) exhibit distinct characteristics: higher study hours, better attendance, stronger previous academic performance, increased tutoring participation, higher motivation levels, greater parental involvement, and better access to learning resources.
21. Factors associated with reduced academic performance include greater distance from school and the presence of learning disabilities, as indicated by the strongest negative correlations with exam scores.
22. The strongest predictors of exam performance are attendance, hours studied, previous scores, access to learning resources, and parental involvement, emphasizing that consistent academic engagement and supportive environments are key drivers of student success.

Conclusion

This analysis used a structured, data-driven approach to draw inferences about the most influential academic, social, and environmental factors that affect student performance on exams. By applying comprehensive exploratory analysis, pivot tables, and interactive dashboards, the most powerful drivers of academic success were found to be attendance, study hours, motivation, parental involvement, and access to resources. Although lifestyle factors like sleep and physical activity did not show much direct influence, supportive learning environments and consistent engagement in academics proved vital contributors to higher performance. Overall, the findings in this paper show how educational data can be transformed into actionable insights to support targeted interventions, informed decision-making, and improved student outcomes.