EDA REPORT

1. Objective

The objective of this Exploratory Data Analysis (EDA) was to understand the structure, patterns, and relationships within the student dataset and identify key factors influencing academic performance (`test_score`). The analysis aims to extract meaningful insights that can support data-driven academic interventions and predictive modeling.

2. Dataset Overview

The dataset contains student-level academic and demographic information, including:

- Age
- Gender
- Course Stream
- Attendance Rate (%)
- Study Hours per Week
- Prior GPA (10-point scale)
- Internet Access
- Scholarship Status
- Test Score (Target Variable)

Initial data cleaning steps were completed before analysis, including:

- Missing value imputation
- Outlier checks
- Business rule enforcement

3. Univariate Analysis Summary

Numeric Variables

1. **Study Hours per Week** shows moderate variability across students.
2. **Attendance Rate** is generally concentrated in medium to high ranges.
3. **Prior GPA** follows a fairly balanced distribution.
4. **Test Score** shows variation, allowing meaningful modeling.

Categorical Variables

1. Gender distribution appears balanced.
2. Calculating frequency of Scholarship.

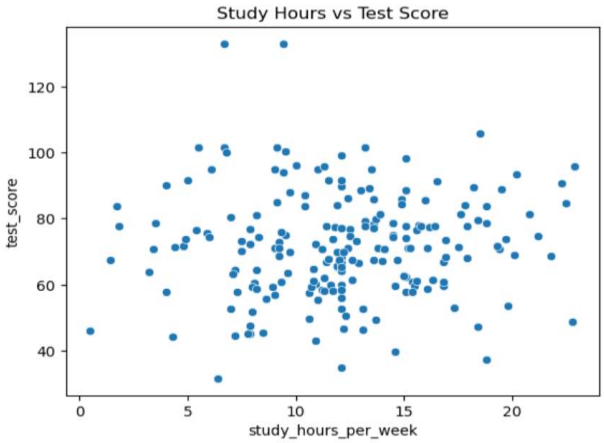
3. Majority of students have internet access.

4. Bivariate Analysis & Key Insights

1□. Study Hours vs Test Score

A positive relationship was observed between study hours and test score. Students who study more hours per week tend to perform better academically.

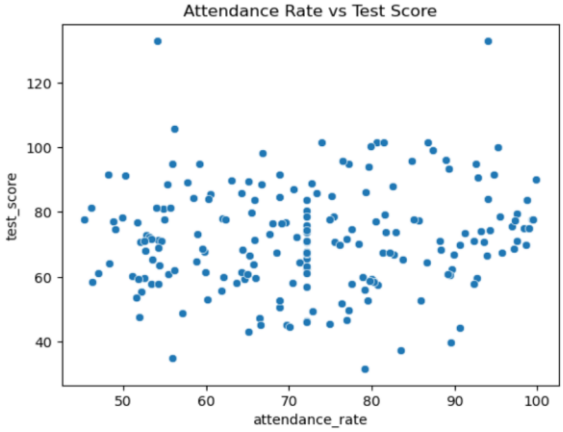
Actionable Insight:
Encourage structured study plans and time management workshops to improve performance.



2□. Attendance Rate vs Test Score

Attendance rate shows a clear positive association with test performance. Students with higher attendance consistently achieve higher scores.

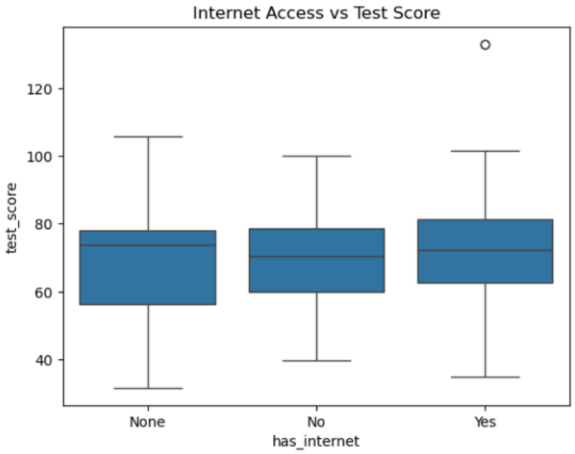
Actionable Insight:
Implement stricter attendance monitoring and early-warning systems for low-attendance student.



3□. Internet Access

Students with internet access demonstrate higher average test scores compared to those without internet access.

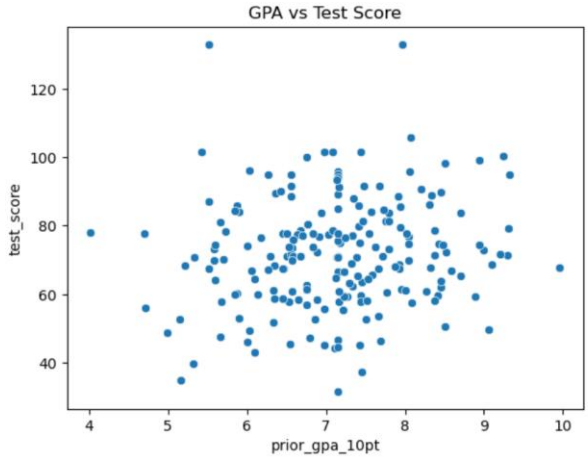
Actionable Insight:
Provide digital support programs or institutional internet facilities for students lacking access.



4□. Prior GPA as Performance Indicator

Prior GPA shows strong positive correlation with current test scores. Students who previously performed well tend to maintain good performance.

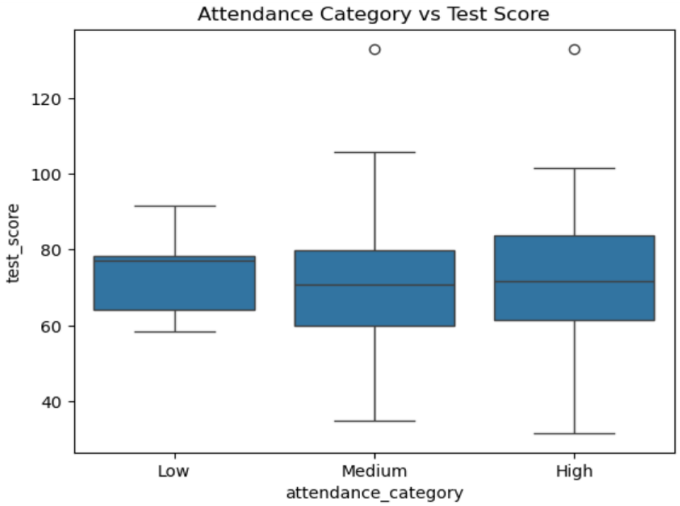
Actionable Insight:
Identify students with low prior GPA and provide early academic mentoring or tutoring.



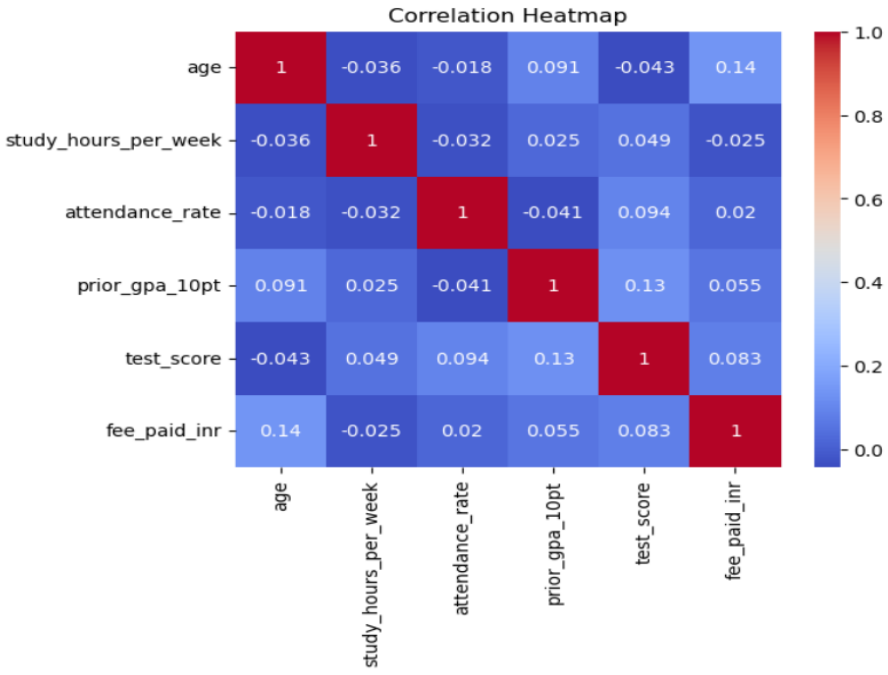
5□. Attendance Category Performance

Students categorized under "High Attendance" outperform those in Medium and Low categories.

Actionable Insight:
Develop intervention programs specifically targeting low-attendance students.



5. Correlation Analysis

The correlation heatmap indicates:

- Strong positive correlation between **prior GPA and test score**
- Moderate positive correlation between **study hours and test score**
- Positive relationship between **attendance rate and performance**