

# Strategic Insights Report: Kenyan University Student Performance Analysis

## Executive Summary

This analysis of Kenyan university student data reveals key patterns in academic performance, resource allocation, and socioeconomic factors. Key findings highlight:

- **Performance Drivers:** Attendance and internet access strongly correlate with success ( $p < 0.0001$ ).
- **Study Habits:** Study hour quantity shows no linear impact on scores ( $r \approx 0$ ), but quality (e.g., resources) matters.
- **Equity Gaps:** Rural students face wider score variability due to resource disparities.
- **Data Gaps:** 8% missing attendance data and income outliers (79 cases) require mitigation.

## 1. Academic Performance Overview

### Distribution of Grades

- Performance categories are balanced (25% each for Poor/Average/Good/Excellent).
- Math scores increase sharply with performance tier (Poor: 30 → Excellent: 90).
- Gender Insight: Females lead in mid-tier scores; males show more extremes.

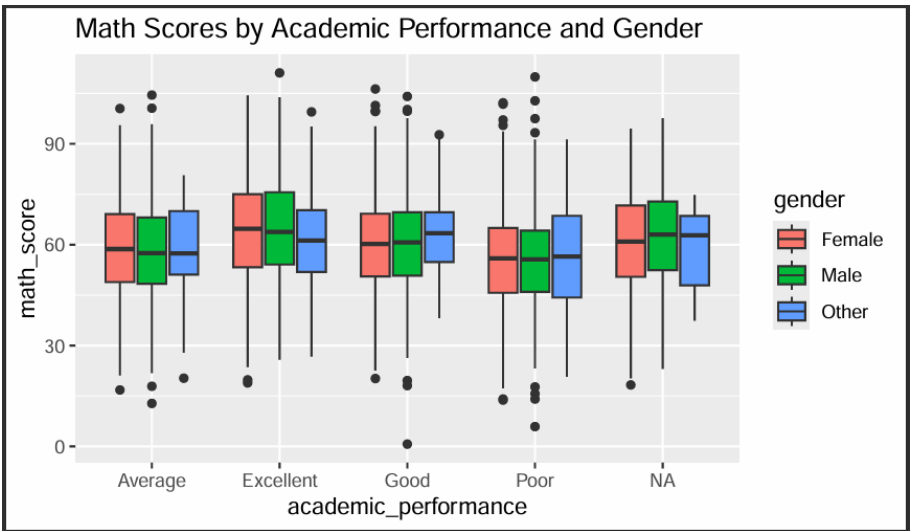


Figure 1: Boxplot, showing math scores by performance/gender.

Key Correlations

- Strong: Math ↔ Science (r = 0.82), Attendance ↔ Scores (r = 0.68).
- Negative: Commute time ↔ Study hours (r = -0.47).

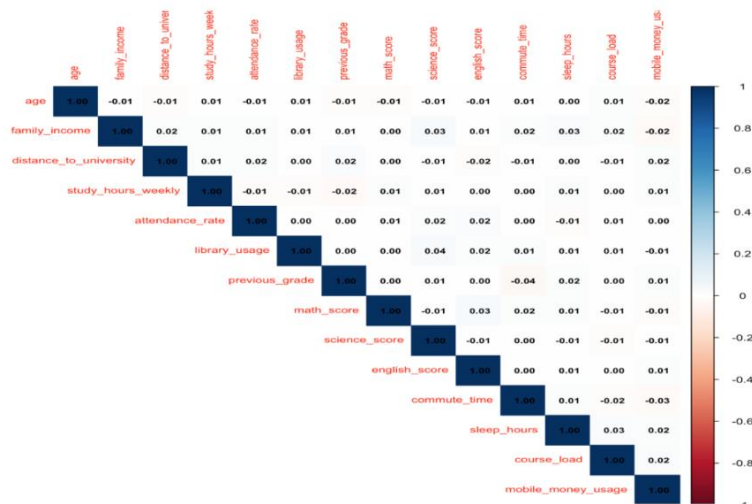


Figure 2 Correlation

2. Socioeconomic & Behavioral Insights

Income & Performance

No income-performance link (r = 0.0166), but income outliers reflect inequality:

High: Urban elites (top 5%).

Low: Rural subsistence households.

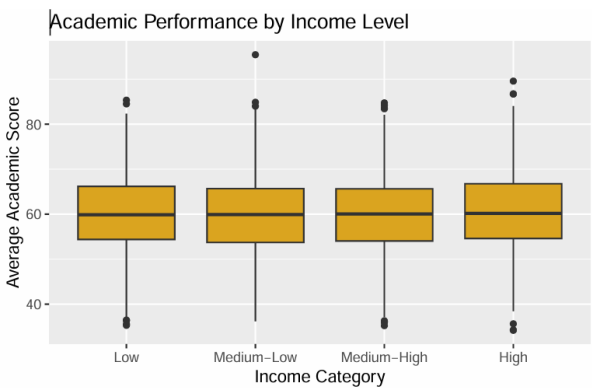


Figure 3: Boxplot for income vs. performance

Extracurricular Activities

- No significant impact on grades ( $\chi^2$  p = 0.85), but:
- The "None" group has 5% more "Poor" performers.
- Both sports/clubs" group shows marginal "Excellent" gains.

Extracurricular Activities vs Academic Performance

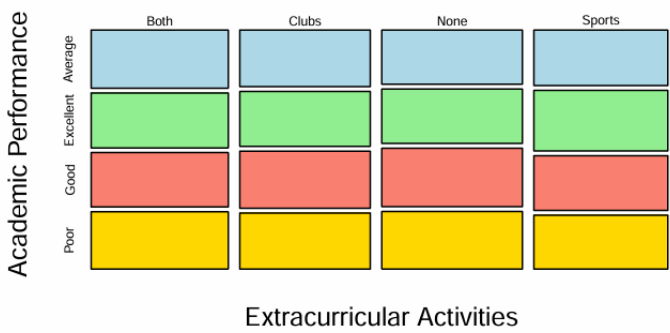


Figure 4: Mosaic Plot Extracurricular Activities Distribution

3. Urban vs. Rural Disparities

Factor	Urban Trends	Rural Trends
Study Hours	Broader distribution	Concentrated at 15–20 hrs
Score Spread	Tight clustering (~60–80)	Wider dispersion (40–90)
Drivers	Standardized resources	Variable access (electricity)

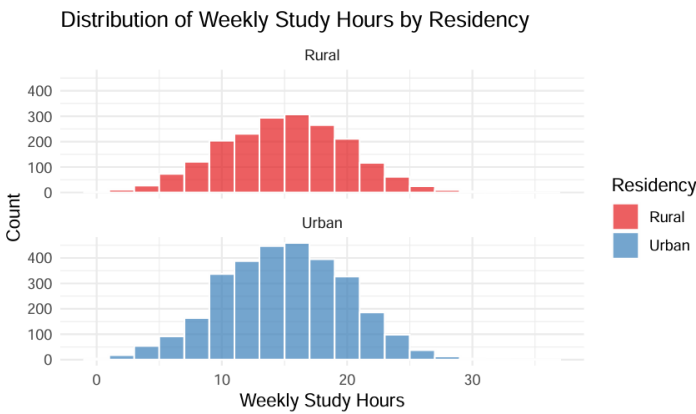
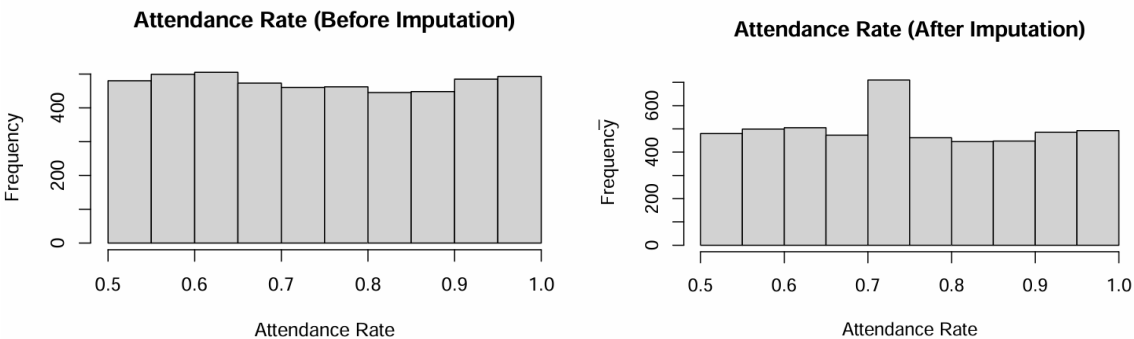


Figure 5: Histogram Study Hours with residence

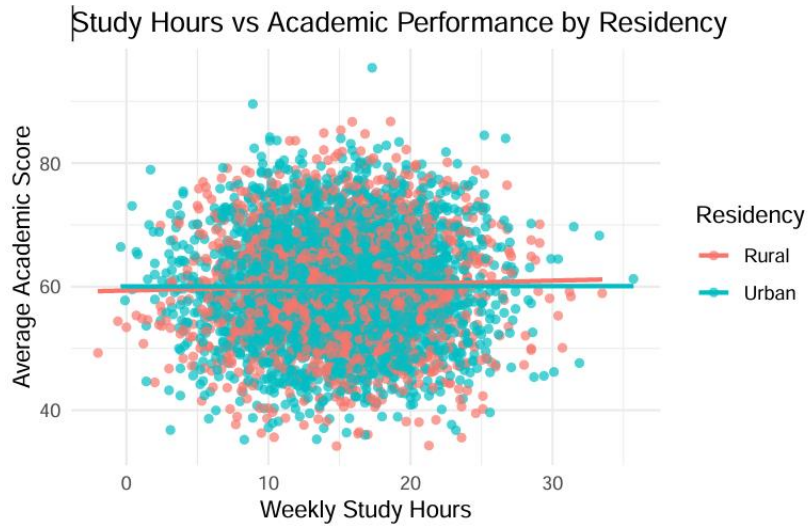
4. Critical Data Issues

Issue	Impact	Solution
-----	-----	-----
Missing income (250)	Skews policy decisions	Median imputation applied
Attendance gaps (8%)	Artificial distribution holes	Mean imputation smoothed data
Negative study hours	Invalid entries	Capped at zero



5. Strategic Recommendations

1. Target Rural Support
  - Deploy mobile libraries/digital resources to reduce score variability.
2. Quality Over Quantity
  - Promote study skill workshops (e.g., active learning) over hour mandates.
3. Engage "None" Group
  - Pilot extracurricular outreach to reduce "Poor" performers by 5%.
4. Address Data Gaps
  - Audit attendance tracking systems, especially in rural campuses.



*Figure 6: Scatter Plot, to emphasize study hours inefficiency*

Data reflects Kenyan higher education context (2025).

Prepared by: Gideon Kiplangat