TECHNICAL REPORT ON DUKE & DUCHESS COLLEGE ACADEMIC METRICS- 2024

Prepared by

Oludiran Ayoyinka Damilola

VEPH/20B/DA133

OUTLINE

1. INTRODUCTION

This report provides a detailed analysis of academic metrics at Duke & Duchess College for 2024. The analysis covers various performance indicators, including student performance, exam results, study habits,IT knowledge, school location, students ag, gender,parent level of education and involvement and socio-economic influences on education.

Objective

The main objective of this project is to improve students performance in external examination while the specific objectives are as follows:

a) To assess student performance based on key academic metrics.

b) To evaluate factors influencing exam results and learning outcomes.

c) To provide insights for academic improvement and policy recommendations.

Problem to be Addressed

Academic success is a primary goal for Duke & Duchess College, yet student performance in standardized examinations, such as SAT, varies significantly across different groups. While classroom learning is essential, other external factors may contribute to student outcomes. The institution seeks to move beyond traditional assumptions and empirically test various parameters that could influence academic performance.

This study investigates whether factors such as study habits, teacher quality, socioeconomic background, parental involvement, IT knowledge, and school type significantly impact student exam results. By analyzing these relationships, the college aims to determine which elements play a crucial role in shaping academic success.

Instead of drawing immediate conclusions, this project **is** exploratory in nature. It tests different academic and non-academic variables to identify the most impactful predictors of student performance. The insights from this study will inform targeted interventions**,** policy recommendations, and strategies to improve overall academic standards.

Key Questions Addressed:

1. To what extent do study hours per week influence SAT scores?

2. How does teacher quality impact student performance?

3. What role does socioeconomic background and parental involvement play in learning outcomes?

4. Does IT knowledge contribute to better academic results?

5. Are there significant differences in performance between students in urban vs. rural schools?

6. What age tends to perform the best in an external examination?

7. Does parent education influences the learning outcomes of students?

STORY OF DATA

Data Source:

The dataset was sourced from Kaggle, a well-known open data repository, and was collected through an online survey sent to students who recently gained admission to Duke & Duchess College. Since the responses come directly from students, the data reflects self-reported academic behaviours, study habits, and socio-demographic factors.

Data Collection

The data was gathered through an online questionnaire sent via email to admitted students. The survey captured academic, personal, and socio-economic information relevant to student performance. Students self-reported details such as study hours, parental involvement, school location, and IT knowledge. The collected responses were compiled into a structured dataset for analysis.

Data Structure

The dataset is organized in a tabular format, where:

- Each row represents an individual student's response, capturing various factors influencing their academic performance.

- Each column represents a specific variable, such as student ID**,** SAT score, age, gender, study habits, and socio-economic background.

Important Features & Their Significance

i. **SAT Score**

- The primary measure of academic performance in the dataset.

- Helps in assessing the impact of various factors (study habits, socioeconomic status, parental involvement, etc.) on student success.

ii **Study Hours Per Week**

- Captures the time commitment students put into studying.

- Helps determine whether more study hours lead to higher SAT scores.

iii **School Location (Urban vs. Rural)**

- Indicates whether a student attended an urban or rural high school.

- Urban schools may have better facilities and teachers, which could influence SAT scores.

Iv . **Parental Involvement (Low, Medium, High*)***

- Measures the level of parental support in a student's education.

- Higher involvement is often linked to better academic performance.

v   **Extra Tutorials (Yes/No)**

- Expected to reveal whether attending tutorials contributes to higher SAT scores and overall academic improvement.
- Provides insight into whether expanding tutorial sessions would benefit more students.

vi   **Socioeconomic Status (SES) (Low, Medium, High**)

- Determines access to quality education, tutoring, and study materials.

- High SES students may have advantages like private tutoring and SAT prep courses.

vii   **IT Knowledge (Low,High , Medium,)**

- Measures digital literacy and technology access on student performance (e.g., computers, online courses).

- Higher IT knowledge can provide better access to educational resources and enhance student academic performance.

viii   **Distance to School (Short, Moderate, Long)**

- Evaluates the influence of commuting distance on student performance.

- Supports recommendations for transportation assistance, flexible schedules, or remote learning options.

Ix   **Teacher Quality Rating** *(1-5 scale, self-reported)*

- Students rate the effectiveness of their teachers.

- Helps determine whether higher teacher quality correlates with higher SAT scores.

Data Biases

Since students self-reported their study hours, parental involvement, and teacher quality, their responses may be influenced by memory recall issues or exaggeration. They might overestimate study hours to appear more dedicated or rate teachers based on perception rather than objective criteria. Parental involvement is also subjective, as students may interpret it differently. To minimize bias, cross-validating with external records and using precise, structured questions can improve data accuracy.

3.   DATA SPLITTING AND PREPROCESSING

Data Cleaning

Before analyzing the academics performance data, I conducted a data quality check to ensure accuracy and reliability. This involved identifying missing values, duplicate entries, and inconsistencies using Microsoft Excel's built-in functions.

i)   Handling Missing Values

A prerequisite data quality check was conducted using Microsoft Excel's built-in functions to assess missing values. The check confirmed that there were no blank spaces in the dataset. This validation ensured data integrity and reliability before proceeding with the analysis. Missing values can distort statistical analysis and affect the accuracy of visualizations, leading to misleading conclusions. A dataset free of missing values ensures that calculations and comparisons remain consistent across all records. If missing values are present and not handled properly, they can introduce bias in model predictions or trend interpretations.

ii)   Removing Duplicates

After checking for missing values, a duplicate records assessment was conducted using Microsoft Excel's built-in functions. The test confirmed that no duplicate entries existed in the dataset, indicating that data cleaning had occurred before the analysis. Duplicate records can distort statistical analyses and misrepresent trends. Removing duplicates ensures that each observation is unique, thereby maintaining the accuracy of the findings. Since no duplicates were found, further data cleaning was unnecessary. This validation reinforced the dataset's integrity, ensuring that the

analysis would yield meaningful and reliable insights. By confirming the absence of duplicates, the study upholds a high standard of data quality, ultimately leading to more accurate and actionable conclusions.

iii)   **Data Transformation**

To enhance data clarity and consistency, the *JAMB Score* column was renamed to *SAT Score* to align with the analysis framework. The SAT scores were then categorized in intervals of 50 to facilitate comparative assessments across performance bands. Additionally, the *Distance to School* variable was grouped in increments of 5 units, allowing for a structured evaluation of its impact on academic performance. These transformations improved data interpretability, making it easier to identify trends and derive meaningful insights.

iv)   **Data Splitting**

In this study, the **SAT Score** serves as the **dependent variable**, representing students' academic performance. The goal is to analyze how various factors influence SAT scores and identify actionable insights for academic improvement.

The **independent variables** include:

- **Demographic Factors:** Age, Gender

- **Academic and Learning Support:** Study Hours, Extra Tutorial Attendance, Teacher's Quality

- **Environmental and Socioeconomic Influences:** Distance to School, Parental Involvement, Socioeconomic Status

- **Technical Proficiency:** IT Knowledge

By structuring the data in this way, the analysis aims to determine the extent to which these independent variables impact SAT scores, enabling the college to develop strategies that support student success.

v)   **Industry Context**

The data belongs to the education sector, specifically a college that aims to improve student academic performance. The focus is on identifying key factors influencing SAT scores and areas for improvement.

vi)   **Stakeholders**

The primary stakeholders are the school board of directors, who will use the insights to implement data-driven policies for academic improvement. Other beneficiaries may include faculty members, academic advisors, and policymakers looking to enhance student outcomes.

vii)   **Value to the Industry**

This analysis is crucial for the college as it provides data-driven insights into the factors influencing students' academic performance, particularly their SAT scores. By identifying key determinants, such as extra tutorial attendance, study hours, distance to school, and socioeconomic status, decision-makers can implement targeted interventions to enhance student learning outcomes.

4)   PRE-ANALYSIS

Identify Key Trend

An initial review of the dataset suggests potential trends in factors affecting student academic performance, particularly their SAT scores. Key areas of interest include identifying:

i)   **Impact of Extra Tutorials:** Students who attend extra tutorials may exhibit higher SAT scores compared to those who do not.

ii)   **Effect of Study Hours:** A possible trend indicates that students who dedicate more study hours tend to achieve better SAT performance.

iii)   **Influence of Distance to School:** Students with longer commuting distances may show lower SAT scores due to fatigue and reduced study time.

iv)   **Parental Involvement Correlation:** Higher parental involvement might be linked to improved academic performance.

v)   **Socioeconomic Status Disparities:** Students from higher socioeconomic backgrounds may have access to better resources, resulting in higher SAT scores.

vi)   **IT Knowledge Influence:** A trend suggesting that students with stronger IT knowledge perform better, possibly due to access to digital learning resources.

vii) **Teacher's Quality Contribution:** Schools or classes with higher-rated teachers may demonstrate better overall student performance.

ii) Potential Correlations

An initial analysis suggests a positive correlation between various factors and SAT **scores**, meaning that an increase in these factors is likely associated with improved academic performance. The observed relationships include:

a   **Student Age vs SAT Score**

Younger students might have lower SAT scores due to less academic experience, while older students might perform better with more years of study.

b   **Parental Involvement vs. SAT Score**

Students with high parental involvement **tend** to achieve higher SAT scores compared to those with medium or low involvement.

c   **Study Hours vs. SAT Score**

A higher number of study hours per **week** is linked to better SAT performance.

d   **Extra Tutorial Attendance vs. SAT Score**

Students who attend extra tutorials (Yes) typically have higher SAT scores than those who do not.

e   **Distance to School vs. SAT Score**

A negative correlation may exist, where students who travel longer distances tend to have lower SAT scores.

f.   **Socioeconomic Status vs. SAT Score**

Students from higher socioeconomic backgrounds tend to perform better on the SAT.

g.   **IT Knowledge vs. SAT Score**

Students with higher IT knowledge may demonstrate better academic performance.

h   **Teacher's Quality vs. SAT Score**

Higher-rated teachers contribute to better student performance.

iii)   **Initial Inisights**

Before conducting a detailed statistical analysis, a preliminary review of the dataset reveals promising insights and areas that warrant further investigation:

a. Does age have an impact on the academic performance of student?
b. Do students who attend extra tutorials perform significantly better on the SAT than those who do not?
c. How many study hours per week lead to the best academic performance?
d. Does a longer distance to school negatively affect SAT scores?
e. How does parental involvement impact student performance?
f. Do students from higher socioeconomic backgrounds perform better than those from lower-income families?
g. Does having IT knowledge give students an academic advantage?
h. How does teacher quality influence student SAT scores?

5. IN-ANALYSIS

i) Unconfirmed Insights

a)   Age and SAT Performance:

The fact that only three students aged 17-19 were top scorers suggests a potential link between younger students and high SAT scores. However, this correlation is not conclusive, as other variables such as study habits, prior knowledge, or access to resources might play a more significant role.

b)   Study Hours and JAMB Score:

While a strong positive relationship is noted between study hours (35-40 per week) and academic performance, this does not imply causation. Other factors that may influence outcomes include the quality of study sessions, teaching methods, or individual learning styles, may also contribute.

c)   IT Knowledge and Academic Performance:

The data suggests that students with higher digital literacy perform better. However, it remains uncertain whether IT knowledge directly improves academic performance or if students with strong IT skills also excel in other areas (e.g., critical thinking or resource accessibility).

d)   Extra Tutorials and SAT Scores:

A slight advantage (177 vs. 171) was noted for students attending extra tutorials, but the small difference indicates other factors may be at play. Further investigation is needed to determine whether the tutorials were effective or if high-achieving students were more likely to attend.

e)   Distance to School by SAT Score

A negative correlation suggests that students living closer to school tend to perform better (average SAT score of **184**) compared to those living farther away (average SAT score of **168**). This may be due to reduced travel fatigue, more study time, or better access to academic resources. However, this insight is not yet verified, as other factors such as socioeconomic status, parental involvement,

and school quality may also influence this trend. Further statistical testing is needed to confirm this relationship.

f)   Performance Analysis by Teacher Quality

Students who rated their teachers highly (**5/5**) had a significantly higher average SAT score (**217**) than those who rated their teachers poorly (**152**). This suggests that teacher effectiveness plays a major role in student performance, possibly due to better instruction, engagement, and support. However, this correlation has not been fully validated, as other variables such as student motivation, curriculum difficulty, and access to additional learning support may also contribute to these differences. Further investigation is required to establish causation.

g)   Parent Involvement by SAT Score

Higher parental involvement in learning is linked to better performance (average SAT score of **189**), while students with low parental support scored lower (**165**). This highlights the importance of a supportive home environment in academic success. However, it remains unclear whether parental involvement directly influences SAT scores or if other factors, such as socioeconomic status, school quality, or student self-discipline, play a more significant role. Additional analysis is needed to confirm this relationship.

h)   Socioeconomic Status by SAT Score

Students from high socioeconomic backgrounds performed better (average score of **191**) than those from lower socioeconomic backgrounds (**165**). This could be due to factors such as access to better schools, private tutoring, learning materials, and a conducive study environment. However, this insight is not yet confirmed, as personal determination, school support systems, and scholarship opportunities may mitigate the impact of socioeconomic status on performance. Further research is necessary to determine the extent of this effect.

**ii)	Recommendation**

**a)	Encourage Consistent Study Habits**

Promote structured study routines, especially targeting 35-40 study hours per week, as they correlate with higher scores.

**b)	Improve Teacher Effectiveness**

Invest in teacher training programs to enhance teaching quality.Encourage student-teacher engagement to foster a better learning environment.

**c)	Enhance Parental Involvement**

Schools should create programs that actively involve parents in their children's education. Educate parents on strategies to support academic performance at home.

**d)	Leverage IT Knowledge for Academic Success**

Integrate digital literacy courses into school curricula. Provide students with access to technology and resources to enhance learning.

**e)	Expand Access to Extra Tutorials**

Offer additional tutoring sessions to support students who may need extra academic assistance. Explore personalized tutoring strategies to maximize effectiveness.

**f)	Address Socioeconomic Barriers**

Provide scholarships and financial aid for students from lower socioeconomic backgrounds. Increase access to learning resources such as books, tutoring, and online courses.

**g)	Minimize the Impact of Distance to School**

Schools could implement flexible study schedules or virtual learning options to support students who live far away. Provide transportation support or boarding options to reduce travel fatigue.

iii)   Analysis Techniques Used in Excel

a)   **Pivot Tables:** Used to summarize SAT scores by factors such as study hours, distance to school, teacher quality, and socioeconomic status. This helped identify trends and correlations in student performance.

b)   **Pivot Charts:** Created to visually represent relationships, making it easier to identify patterns, such as the impact of study hours on SAT scores or the difference in performance based on parental involvement.

c)   **Conditional Formatting:** Applied color scales to differentiate high-performing students (e.g., darker blue for top scores) and low-performing students (e.g., lighter shades for lower scores)..

d)   **Statistical Functions (AVERAGE, COUNT, MAX, MIN):** Used to calculate the average SAT scores for different groups and MAX to determine the highest performing student.

6. POST ANALYSIS AND INSIGHT

i)   Key Findings

a)   **Student'Age and Performance:** Student age had little to no influence on academic performance, as younger (17-19) students, who were expected to perform lower, outperformed older ones.

b)   **Study Hours & Performance:** Students studying 35-40 hours per week scored higher, indicating structured study time improves SAT performance.

c)   **Distance to School:** Students living closer to school scored higher (184 vs. 168), likely due to reduced travel fatigue and better access to resources. Proximity to school positively impacts academic performance.

d)   **Teacher Quality:** Higher teacher ratings (5/5) correlate with better SAT scores (217 vs. 152), highlighting the impact of effective teaching.

e)   **Parental Involvement:** Students with strong parental support performed better (189 vs. 165), emphasizing the role of a supportive home environment.

f)   **IT Knowledge:** Higher IT literacy is linked to better academic performance.

g)   **Socioeconomic Status:** Students from higher-income backgrounds scored better (191 vs. 165), likely due to access to better resources.

h)   **Extra Tutorials:** Attending tutorials provided a small advantage (177 vs. 171), suggesting additional support helps but isn't a primary success factor.

ii)   Comparison with Initial Findings

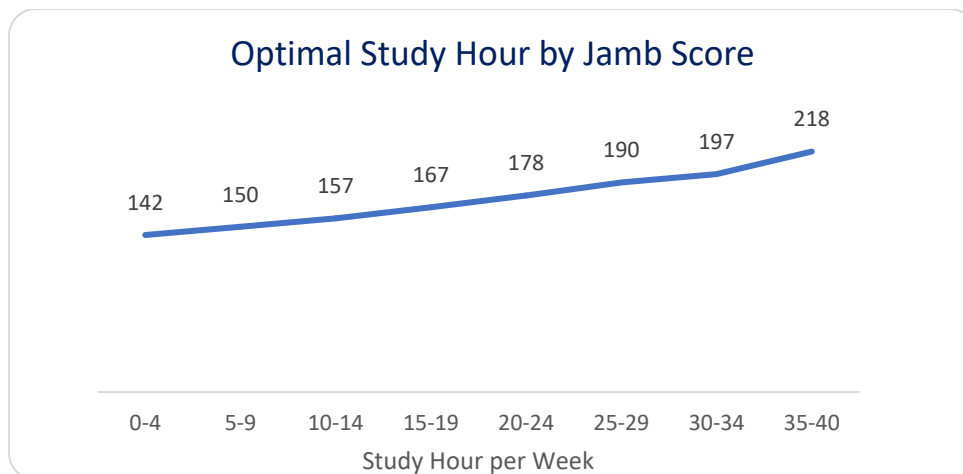The analysis largely supports the initial expectations, with a few nuanced differences:

a) **Student Age and Performance:** The hypothesis that age influence academic performance was not confirmed, while younger students achieved higher scores, age alone may not be a strong predictor of academic success. The correlation between age and SAT scores appears weak, suggesting that age is not a determining factor in academic success.

b) **Study Hours and Performance** – The hypothesis that more study hours lead to better performance was confirmed, especially for students studying 35-40 hours per week, where a strong positive correlation was observed.

c) **Extra Tutorials** – While initial expectations suggested a noticeable boost in SAT scores for students attending extra tutorials, the analysis revealed only a slight improvement (177 vs. 171), indicating that other factors may also contribute to academic success.

d) **IT Knowledge** – The assumption that IT literacy improves performance was validated, with students proficient in IT achieving better results, suggesting that digital literacy enhances learning outcomes. The influence of IT knowledge on performance was more pronounced than expected, reinforcing the growing role of digital skills in education.

e) **Distance to School** – The analysis confirmed that students living farther from school performed worse on average, supporting the initial assumption that travel fatigue and reduced study time negatively impact SAT scores.

f) **Teacher Quality** – A significant gap in SAT scores based on teacher ratings (217 vs. 152) reinforced the assumption that teacher effectiveness plays a crucial role in academic performance.

g) **Parental Involvement** – The data supported the initial hypothesis, showing a direct link between high parental involvement and higher SAT scores (189 vs. 165), emphasizing the role of family support in academic achievement.

h) **Socioeconomic Status** – The trend of higher-performing students coming from wealthier backgrounds was confirmed, highlighting the advantage of better educational resources and support systems.

7.   DATA VISUALIZATION & CHARTS

Performance Analysis by Student Age

| Age	Score
17	367
18	367
19	367
20	366
21	355
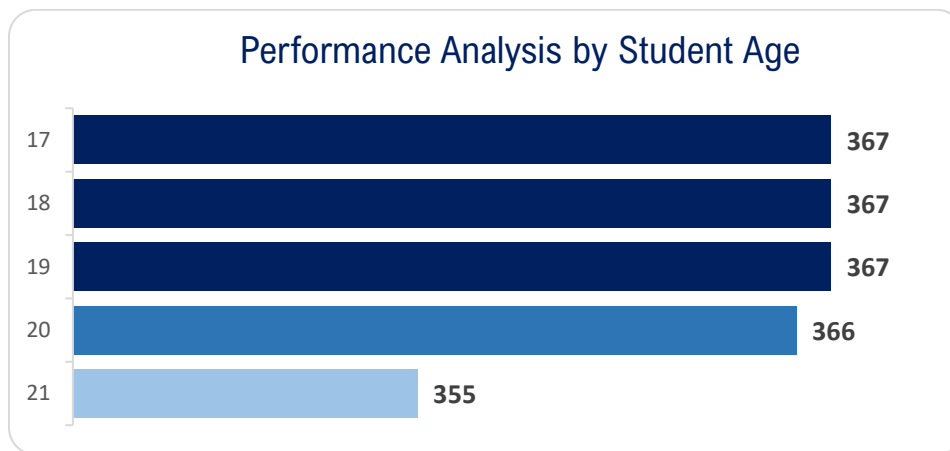
The chart highlights the top five students in the SAT examination along with their ages. It aims to identify the highest score and examine whether age influences performance. Contrary to my initial assumption that younger students might struggle due to limited experience, the results showed that age has little to no effect, as three of the top performers were between 17-19 years old.

Optimal Study Hour by Jamb Score

142   150   157   167   178   190   197   218

0-4   5-9   10-14   15-19   20-24   25-29   30-34   35-40

Study Hour per Week

The analysis reveals a positive correlation between study hours per week and SAT performance. Students who studied more than 25 hours generally performed better, with the optimal study range for excellent results identified between 35-40 hours per week, yielding an average SAT score of 218. In contrast, students who studied between 0-4 hours had a significantly lower average score of 142.

IT Knowledge by SAT Score
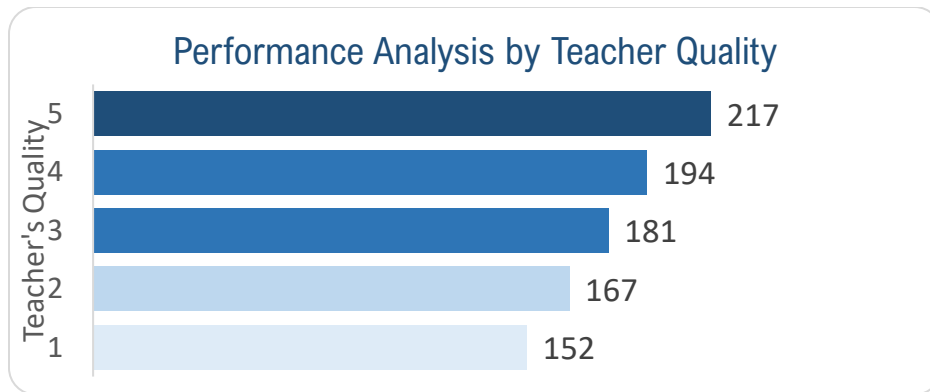
187   173   167

High   Medium   Low

.
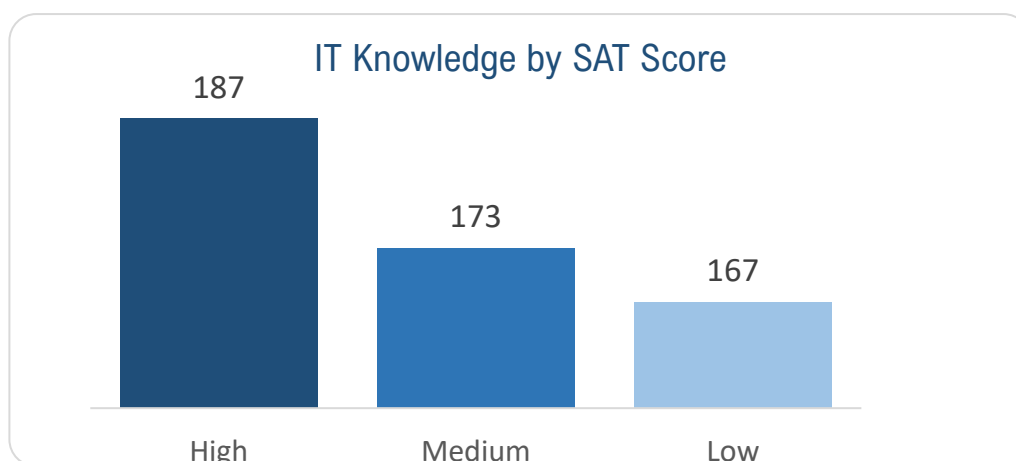
Students with high IT knowledge performed better with an average SAT Score of 187 while those with low knowledge had an average of 167. The influence of IT Knowledge on SAT Score was high implying policies should be geared towards improving student learning status in IT.

Extra Tutorials by SAT Score

171

177

- Yes
- No
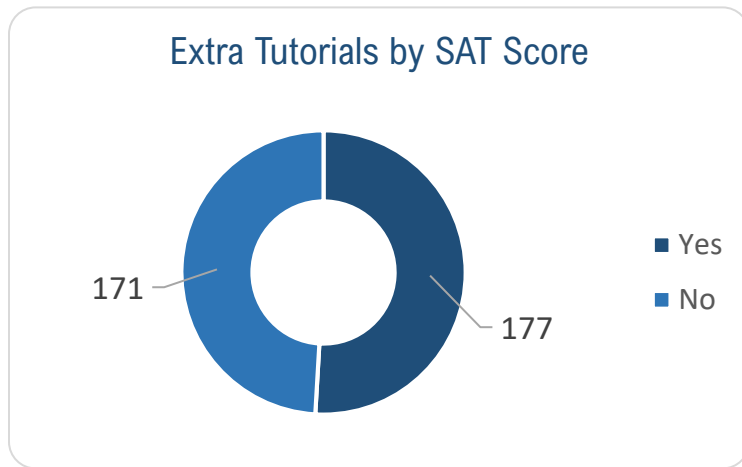
Attending extra tutorials had only a slight improvement (177 vs. 171), indicating that other factors may also contribute to academic success.

Distance to School by SAT Score

Distance in Km

| 0-4 | 184 |
4-8	176
8-12	174
12-16	171
16-20	168
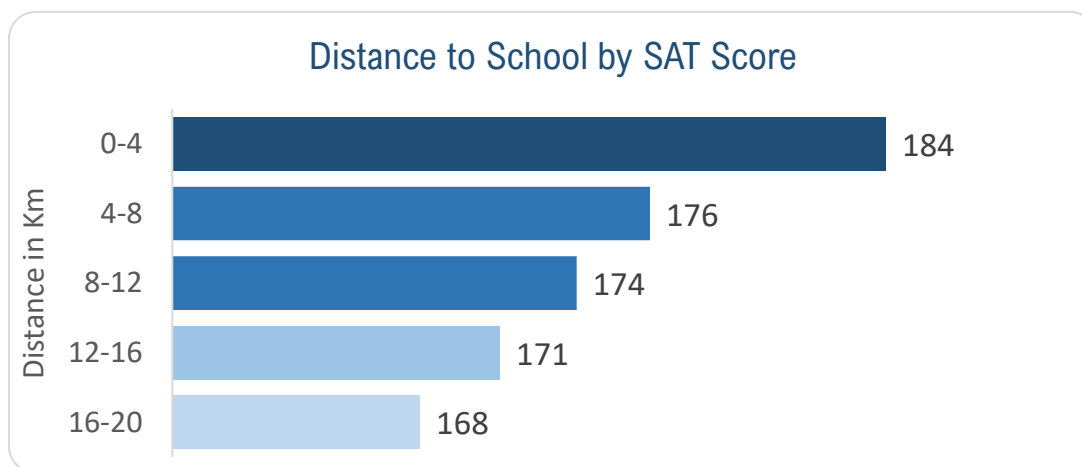
Students with low distance (0-4) to school performed better with an average of 184 against those who have a longer distance (16-20km) to cover.

Performance Analysis by Teacher Quality

| Teacher's Quality	Average Score
5	217
4	194
3	181
2	167
1	152

Teacher's quality had a strong effect on the performance of students in an examination as those who rated the quality of their teacher's teaching 5/5 had an average score of 217 while those who rated their teacher's teaching 1/5 had an average score of 152.

Parent Involvement by SAT Score

	Average Score
High	189
Medium	172
Low	167
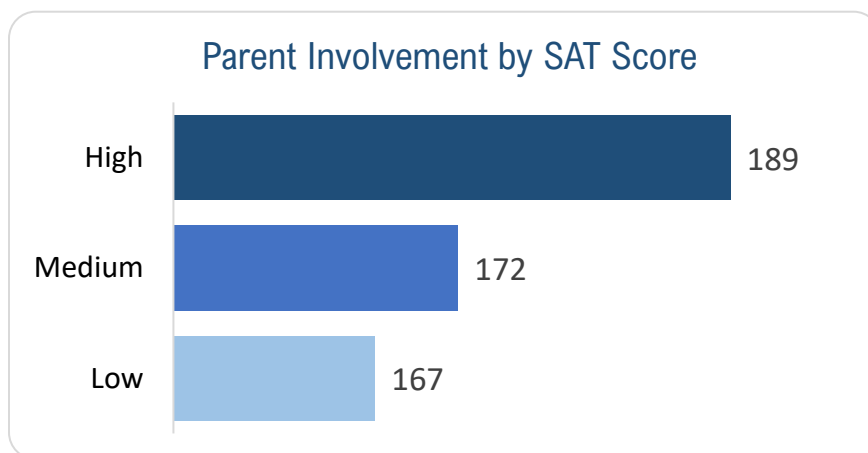
The chart above revealed that parents have a crucial role to play in determining the academic performance of their ward. The students who had high parental involvement in their learning performed better with an average score of 189, which is higher than the two other categories (Medium and low).

Socioeconmic Status by SAT Score

165 191 174
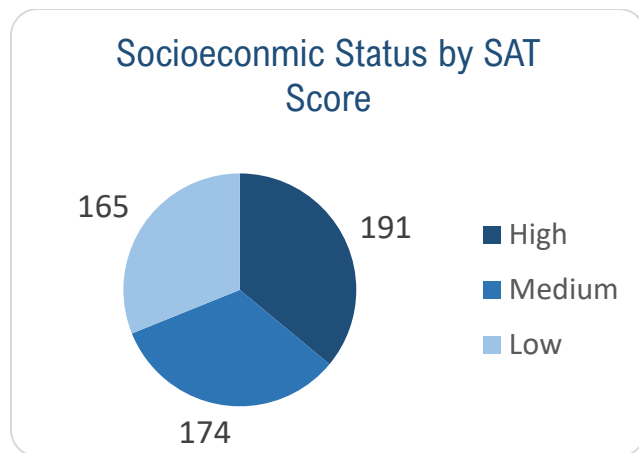
- High
- Medium
- Low

Beyond parent involvement, this chart revealed that the socioeconomic status can also influence the student performance, as student from wealthier home had the highest average of 191 when compared from those Medium and Low Socioeconomic status respectively.
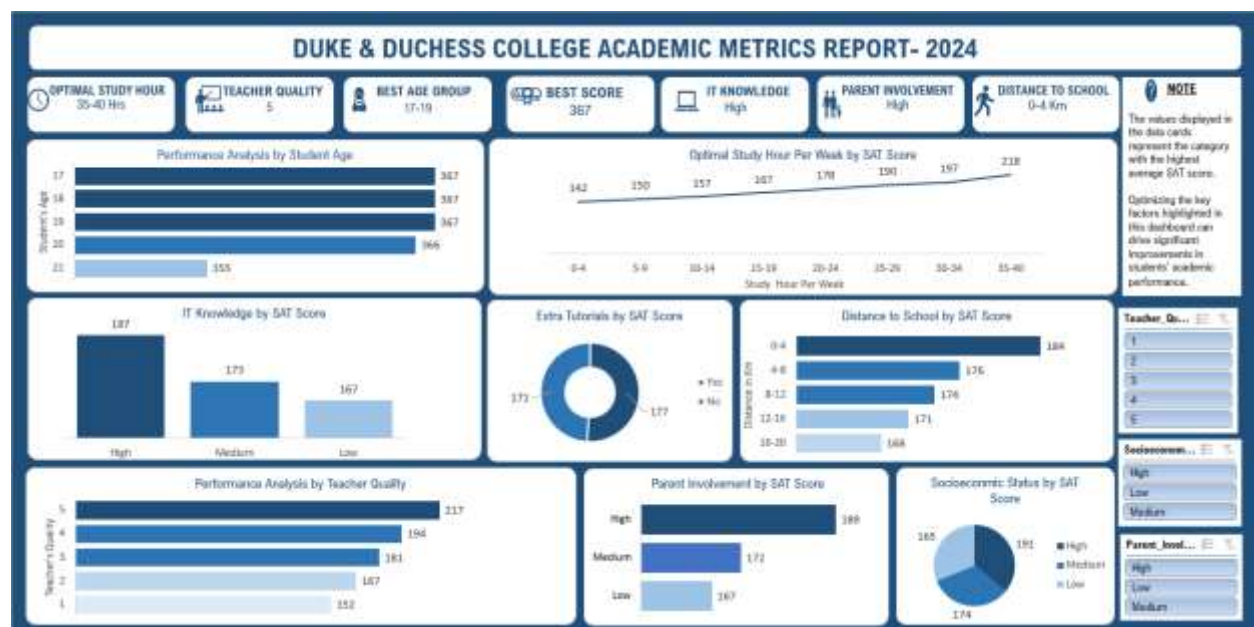
8. RECOMMENDATION AND OBSERVATION

i.   **Encourage Adequate Study Hours**

Students should aim to study for at least 25 hours per week, with an optimal range of 35-40 hours for improved SAT performance.Schools should provide structured study plans and time management strategies.

ii. **Improve Teacher Quality**

High teacher quality positively impacts academic performance across various factors. Schools should invest in teacher training and development programs.

iii. **Leverage IT Knowledge for Better Performance**

Encouraging digital literacy and IT knowledge acquisition can boost students' academic performance. Schools should integrate IT skills into their curriculum.

iv. **Promote Extra Tutorials**

Extra tutorials significantly improve SAT scores. Schools and parents should encourage students to participate in structured tutorial sessions.

v. **Support Socioeconomic and Parental Involvement**

Higher socioeconomic status and parental involvement correlate with better SAT performance. Schools should engage parents and provide additional resources to students from lower socioeconomic backgrounds.

vi **Address Socioeconomic and Parental Involvement Disparities**

The impact of socioeconomic status and parental involvement is more pronounced among students living far from school. Schools should implement strategies to engage parents and provide academic support to students from low-income backgrounds.

vii **Minimize the Impact of Distance on Academic Performance**

Schools and policymakers should explore transportation support for students who live far from school to reduce the negative impact of long commutes on performance.

CONCLUSION

The analysis reveals that factors such as study hours, IT knowledge, extra tutorials, and teacher quality significantly impact students' SAT performance. Students who study between 35 and 40 hours per week tend to achieve the highest SAT scores, especially when supported by high-quality teachers, strong socioeconomic backgrounds, and parental involvement. IT knowledge plays a crucial role in academic success, as students with higher digital literacy consistently outperform those with lower IT skills. Attendance in extra tutorials proves beneficial, as students who engage in additional learning sessions achieve higher SAT scores, particularly when combined with strong teacher support. Distance to school negatively affects academic performance, but this gap narrows with high teacher quality, highlighting the importance of investing in skilled educators. Overall, improving teacher quality, promoting parental involvement, supporting students from lower socioeconomic backgrounds, and encouraging structured study habits can significantly enhance SAT performance.

Reference

Dataset used was retrieved from:https://www.kaggle.com/datasets/idowuadamo/students-performance-in-2024-jamb(Accessed:12th March,2025)