# Exploratory Data Analysis Report

## Objective

This report summarizes an exploratory data analysis (EDA) conducted on a simulated dataset of student performance. The purpose of this analysis is to understand how various factors—such as study habits, demographics, and family income—affect academic outcomes.

## Dataset Overview

The dataset contains information about 500 students, including their demographic details, study habits, and academic performance. The following attributes were analyzed:

1. `student\_id`: A unique identifier for each student.  
2. `age`: Student age, ranging from 15 to 22 years.  
3. `study\_hours`: Weekly study hours, ranging from 0 to 40.  
4. `previous\_gpa`: GPA from the previous semester (2.0 to 4.0).  
5. `extracurricular\_activities`: Number of activities the student participates in (0 to 5).  
6. `final\_grade`: The final grade in a course (0 to 100).  
7. `gender`: Male or Female.  
8. `family\_income`: Annual family income, ranging from $20,000 to $200,000.

## Key Findings

### 1. Factors Influencing Academic Performance

The analysis revealed a few important trends:  
- Students who study more tend to achieve higher grades, as indicated by a moderate positive correlation between `study\_hours` and `final\_grade`.  
- A strong correlation between `previous\_gpa` and `final\_grade` suggests that consistent academic performance plays a key role in success.  
- Family income, surprisingly, shows little to no impact on grades.

### 2. Academic Performance by Gender

When comparing grades between males and females:  
- Both genders perform similarly on average, with almost identical median grades.  
- However, male students exhibit slightly more variation in their performance, as shown by the broader spread of their grades.

### 3. How Students Spend Their Study Time

The distribution of weekly study hours reveals that:  
- Most students dedicate 10 to 25 hours per week to studying.  
- A small number of students either study very little or invest over 30 hours per week, representing the outliers in the dataset.

### 4. Balancing Activities and Academics

The analysis shows an interesting relationship between extracurricular involvement and grades:  
- Students who participate in 1 to 3 activities tend to perform the best academically.  
- Too much involvement (4-5 activities) appears to negatively affect grades, likely due to time constraints.

### 5. Family Income and Performance

Grouping students by income quartiles showed only minor differences in average grades. This suggests that while family income may provide access to resources, it does not significantly influence overall academic performance.

## Conclusion

This analysis highlights the factors that most influence student performance. Study hours and previous GPA are strong predictors of academic success, while gender and family income have limited impact. Additionally, moderate involvement in extracurricular activities is linked to better grades. These insights can help educators and students focus on strategies that promote academic excellence.