# Student Score Analysis Report

## Overview

This report summarizes the analysis of student performance based on demographic and socio-economic factors using the dataset "Expanded\_data\_with\_more\_features.csv." The primary focus was on understanding how variables like parent education, gender, ethnic group, and marital status influence students' math, reading, and writing scores.

## 1. Data Summary

Dataset Details:  
- Total Entries: 30,641  
- Features: 15, including scores for Math, Reading, and Writing.  
- Missing Data: Several columns contained missing values, notably:  
 - EthnicGroup: 1,840 missing values  
 - ParentEduc: 1,845 missing values  
 - TestPrep: 1,830 missing values  
  
Key Statistics:  
- Math Score: Mean = 66.56, Min = 0, Max = 100  
- Reading Score: Mean = 69.38, Min = 10, Max = 100  
- Writing Score: Mean = 68.42, Min = 4, Max = 100

## 2. Gender Distribution

Findings:  
- The dataset contained more female students than male students.  
- Visualization: A bar chart confirmed the higher representation of females in the dataset.

## 3. Parent's Education and Student Scores

Analysis:  
Grouped the data by ParentEduc and calculated mean scores for Math, Reading, and Writing.  
  
| Parent Education Level | Math Score | Reading Score | Writing Score |  
|------------------------------|------------|---------------|---------------|  
| Associate's Degree | 68.37 | 71.12 | 70.30 |  
| Bachelor's Degree | 70.47 | 73.06 | 73.33 |  
| Master's Degree | 72.34 | 75.83 | 76.36 |  
| High School | 64.43 | 67.21 | 65.42 |  
| Some College | 66.39 | 69.18 | 68.50 |  
| Some High School | 62.58 | 65.51 | 63.63 |  
  
Conclusion:  
Higher levels of parental education positively correlate with better student scores across all subjects.  
Visualization: A heatmap highlighted this trend.

## 4. Parent's Marital Status and Student Scores

Analysis:  
Grouped data by ParentMaritalStatus to observe average scores:  
  
| Parent Marital Status | Math Score | Reading Score | Writing Score |  
|-----------------------|------------|---------------|---------------|  
| Divorced | 66.69 | 69.65 | 68.80 |  
| Married | 66.66 | 69.39 | 68.42 |  
| Single | 66.16 | 69.16 | 68.17 |  
| Widowed | 67.37 | 69.65 | 68.56 |  
  
Conclusion:  
Parental marital status has minimal to negligible impact on student performance.  
Visualization: A heatmap confirmed the lack of strong correlation.

## 5. Ethnic Group Distribution

Distribution:  
- Groups: A, B, C, D, E  
- Group C had the highest representation, followed by Groups B and A.  
  
Visualization:  
- A pie chart and bar chart displayed the distribution of ethnic groups in the dataset.

## 6. Weekly Study Hours

Findings:  
- Weekly study hours were categorized as < 5, 5 - 10, etc.  
- Standardized the format of this column for consistency.

## 7. Boxplot Analysis of Scores

Visualized the distribution of scores for Math, Reading, and Writing using boxplots.  
Observations:  
- Scores were well-distributed, with some outliers in all three categories.

## Conclusions

1. Parent education has a significant positive impact on student scores.  
2. Gender distribution is skewed towards females in the dataset.  
3. Parental marital status and ethnic group show minimal influence on performance.  
4. The dataset highlights the importance of socio-economic factors in shaping academic outcomes.

## Recommendations

- Encourage parental involvement in education to improve student performance.  
- Conduct further analysis to explore other potential factors influencing scores.  
- Address missing data issues to improve analysis accuracy.