Analysis Report

Introduction:

This report analyzes the student performance dataset to gain insights into factors that may influence academic achievement. The dataset includes information on student demographics, parental education, test preparation courses, and scores in math, reading, and writing.

Key Findings:

1. Distribution of Math Scores:

- The distribution of math scores is slightly skewed to the left, indicating that a larger portion of students tend to score moderately well in math, with a few students achieving exceptionally high scores.

- The histogram demonstrates the frequency of math scores within specific ranges.

2. Gender Differences in Math Performance:

- The boxplot comparing math scores by gender reveals that males tend to have slightly higher median math scores than females.

- However, there is significant overlap in scores between the genders, suggesting that gender alone may not be a major determinant of math performance.

3. Correlation between variables:

- Correlation Matrix: The correlation matrix highlights the strength and direction of the linear relationship between different numerical variables.

- The heatmap illustrates the correlation between the various variables in the data. It shows strong positive correlation between math, reading and writing scores.

4. Parental Level of Education and Math Score

- The Scatter plot shows how the parental level of education influences the math score. Students from higher parental level of education tend to score higher.

- The color coding also indicates the impact of lunch type.

5. Test Preparation Course and Math Score

- The scatter plot shows the correlation between the test preparation course and the math score. Students who completed the preparation course tend to score higher than those who didn't.

6. Gender and Reading Score

- The Bar chart shows the average reading score by gender. There is a marginal difference between average reading scores for males and females.

7. Race/Ethnicity and Writing Score

- The bar chart shows the average writing score by race/ethnicity. It indicates the difference in average writing score across different race/ethnicity groups.

Conclusion:

The student performance dataset provides valuable information on the diverse factors that influence academic performance. While gender and race/ethnicity do demonstrate some correlation with scores, other variables like parental education and test preparation seem to exert a stronger influence.