Exploratory Data Analysis Summary - Student Performance Dataset

1. Overview

This dataset contains exam scores and demographic details for high school students.

It includes scores in math, reading, and writing, along with gender, parental education, lunch type, and test preparation status.

The purpose of this analysis is to identify academic performance trends and factors that influence student outcomes.

2. Key Findings

Female students perform significantly better in reading and writing.

Male students slightly outperform in math.

Reading and writing scores are strongly positively correlated.

Students who completed the test preparation course scored higher across all subjects.

Group E students consistently achieved the highest average scores.

Students with standard lunch performed better than those on free/reduced lunch.

3. Visual Support

Countplot: Gender, parental education, lunch, test prep course

Histogram: Distribution of math, reading, and writing scores

Boxplots: Scores by gender and test prep status

Correlation Heatmap: Relationship between all three scores

Bar Charts: Average scores by gender and ethnicity

Pairplot: Multivariate score relationships

4. Educational Implications

Schools should invest in structured test preparation programs.

Special attention can be given to underperforming ethnic groups or lunch-based categories.

Targeted reading and writing support for male students may help bridge subject gaps.

High-performing patterns in Group E can be studied and replicated for other groups.