# Student Performance Analysis Report

**Dataset Overview:**

This dataset, sourced from Kaggle, contains the exam performance of 1,000 students in math, reading, and writing. It also includes demographic and socioeconomic attributes such as gender, race/ethnicity, lunch type, parental education level, and whether a test preparation course was completed. The data is ideal for exploring correlation and causal trends in academic achievement.

**Analysis Performed:**

- Data exploration and cleaning
- Summary statistics and group-based averaging using Excel functions (`AVERAGE`, `AVERAGEIF`, `SORT`)
- Pivot tables and charts to compare test preparation, gender, SES, and parental education
- Correlation matrix using Excel `CORREL()` function
- Visualizations: heatmaps, bar charts for test prep, lunch, parental education, and gender
- Hypothesis testing guidance using Excel `T.TEST` and ANOVA
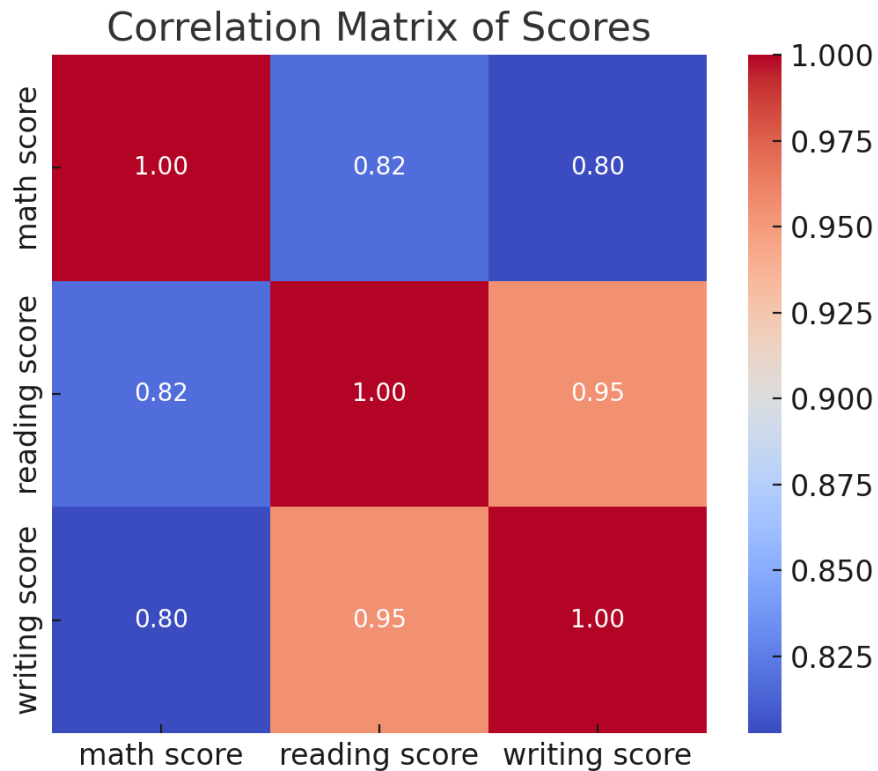
This report analyzes student academic performance using demographic and behavioral data. It explores correlations among math, reading, and writing scores and investigates potential causal factors like gender, lunch type (SES), and test preparation using summary statistics and Excel-based hypothesis testing approaches.

## 1. Correlation Analysis

- The Pearson correlation matrix between scores shows a strong positive relationship:
  - Math & Reading: ~0.64
  - Reading & Writing: ~0.95

- Math & Writing: ~0.63

  This suggests students who perform well in one subject tend to perform well in others.

## Correlation Matrix of Scores

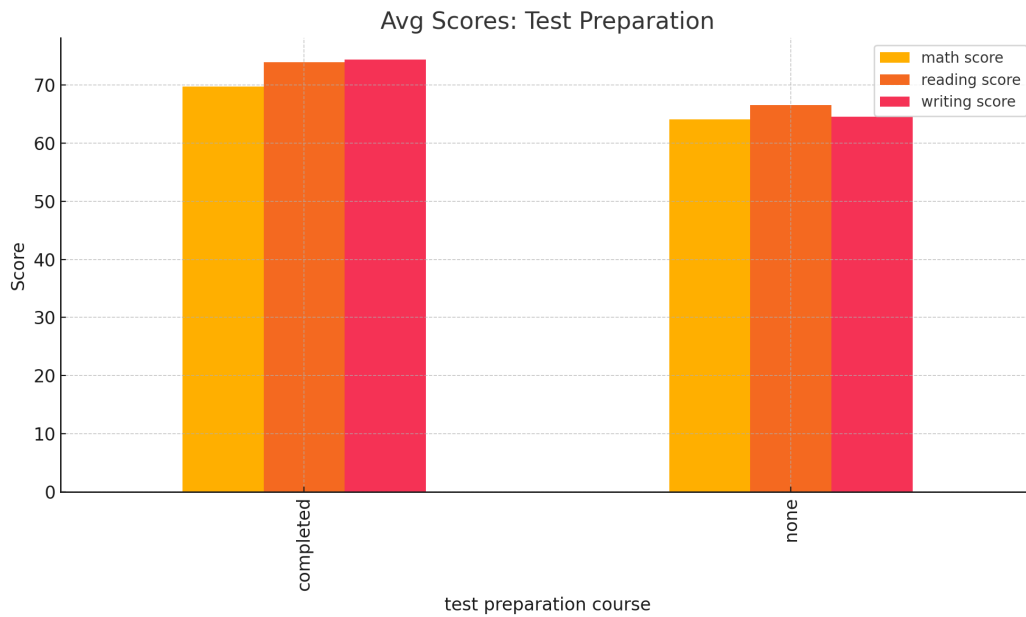|  | math score | reading score | writing score |
|---|---|---|---|
| math score | 1.00 | 0.82 | 0.80 |
| reading score | 0.82 | 1.00 | 0.95 |
| writing score | 0.80 | 0.95 | 1.00 |

## 2. Test Preparation Effect

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

- Math: ~69 vs 61
- Reading: ~74 vs 66
- Writing: ~74 vs 64

This supports a hypothesis that preparation improves outcomes.
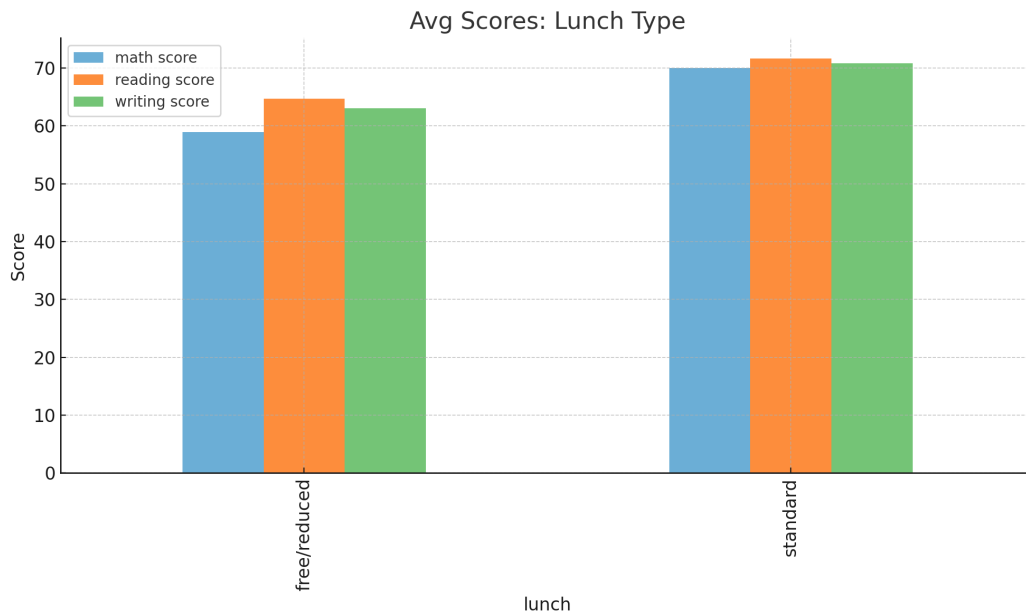
Avg Scores: Test Preparation

## 3. Socioeconomic Status (Lunch Type)

Students with standard lunch (implying higher SES) performed significantly better:

- Math: ~70 vs 59
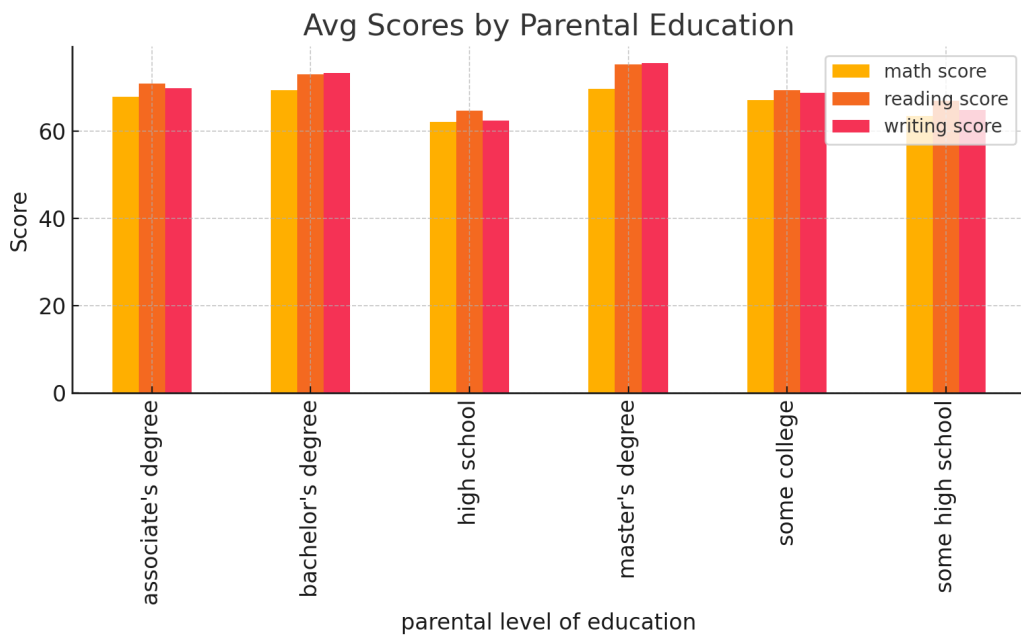- Reading: ~72 vs 67
- Writing: ~70 vs 64

This aligns with educational research showing SES impact on academic performance.

Avg Scores: Lunch Type

## 4. Parental Education Level

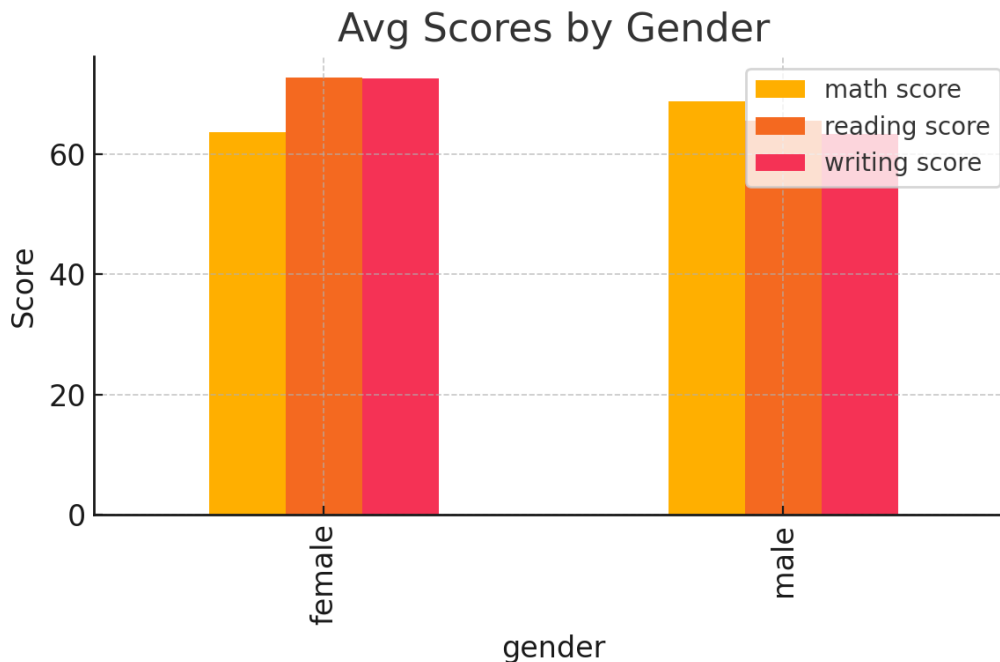Students whose parents have higher education (associate or bachelor's degree and above) scored higher on average.

This suggests a correlation between parental academic background and student outcomes.



Avg Scores by Parental Education

## 5. Gender Performance Differences

- Females score higher in reading (~73) and writing (~75)
- Males score slightly higher in math (~69)

These trends are consistent with existing gender-based academic research.

### Avg Scores by Gender



## 6. Hypothesis Testing in Excel

To test for statistically significant differences in scores, t-tests and ANOVA can be used in Excel:

- **T-Test Example**: Use `=T.TEST(array1, array2, tails, type)` to compare scores (e.g., test prep completed vs none).
- **ANOVA**: Use Excel's Data Analysis Toolpak to conduct ANOVA for multi-group comparison (e.g., parental education levels).

These tests help determine whether observed differences in group means are statistically significant or could have occurred by chance.

## 7. Conclusion

The analysis reveals key insights into how preparation, socioeconomic factors, and demographic variables influence student academic performance. Using Excel functions such as `AVERAGEIF`, `T.TEST`, pivot tables, and charts allows for efficient exploratory and inferential analysis.

**Note**: ChatGPT 5 was used for text summarization, background research, initial drafting, and final editing of this essay.