Assignment 1

Principles of Data Science

Name : Phanindra Chowdary Borra

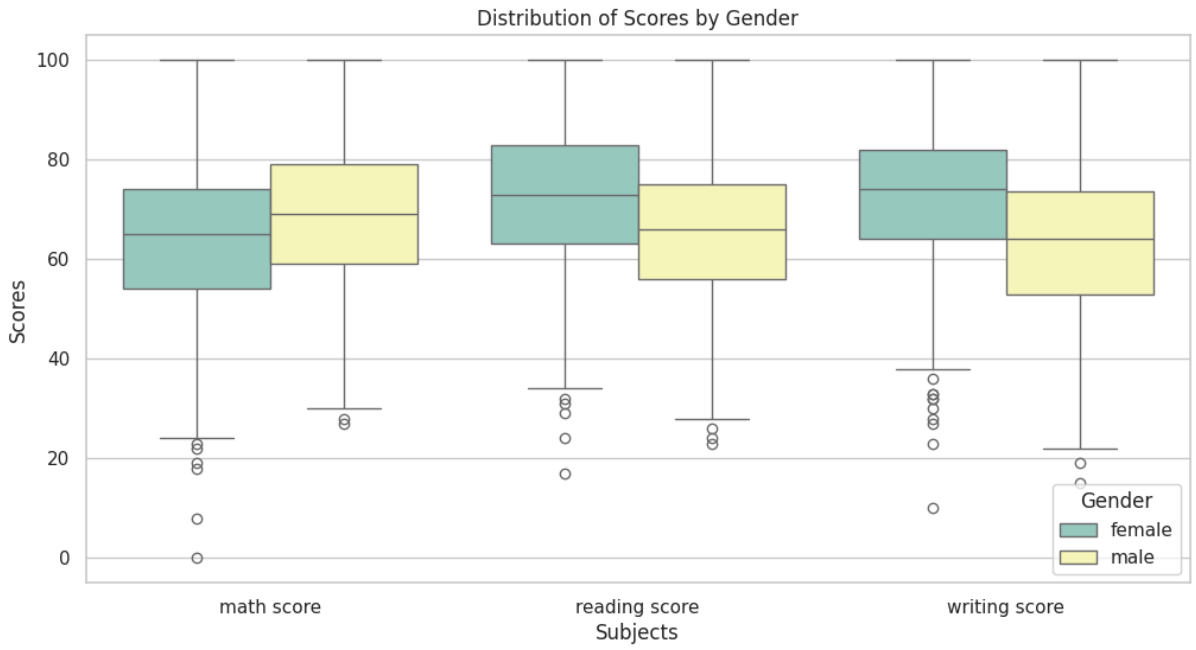
Student Id : 16352169

1. Distribution of Scores (Math, Reading, and Writing) by Gender

Visualization: ***Boxplot*** of math, reading, and writing scores by gender.

Purpose: This plot helps to visualize the distribution and central tendency of scores for each subject by gender.

Analysis: It becomes easier to compare how male and female students perform in each subject. We can assess whether there's a notable difference in median scores and identify the presence of outliers. For instance, if female students consistently have higher scores in reading and writing, this can point to gender-related performance differences.

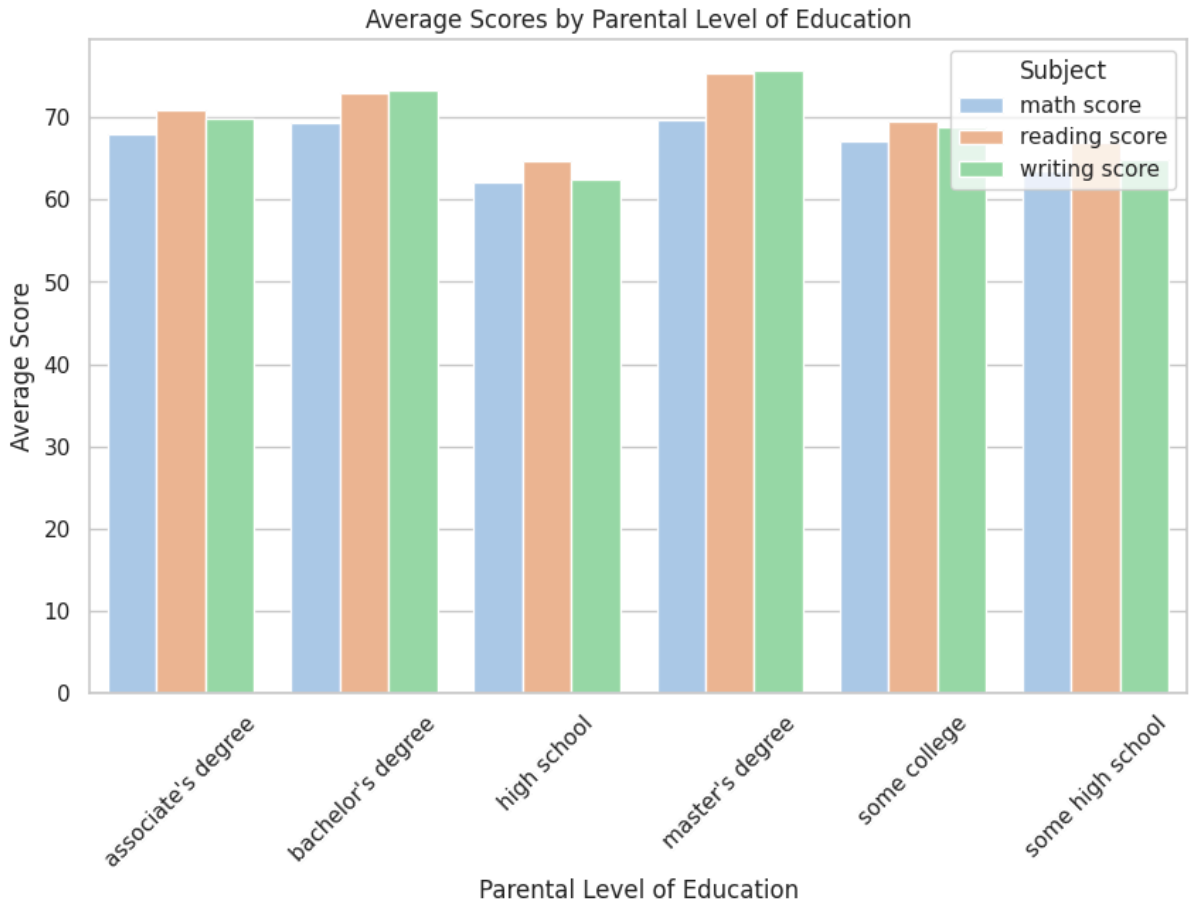


2. Average Scores by Parental Level of Education

Visualization: ***Bar chart*** showing average math, reading, and writing scores grouped by the parental level of education.

Purpose: This bar chart allows us to see the correlation between the parental education level and student performance across subjects.

Analysis: With this, it's easier to analyze whether a higher parental level of education tends to correspond to higher student scores. This could help in understanding the socio-economic or educational support impacts on student performance.

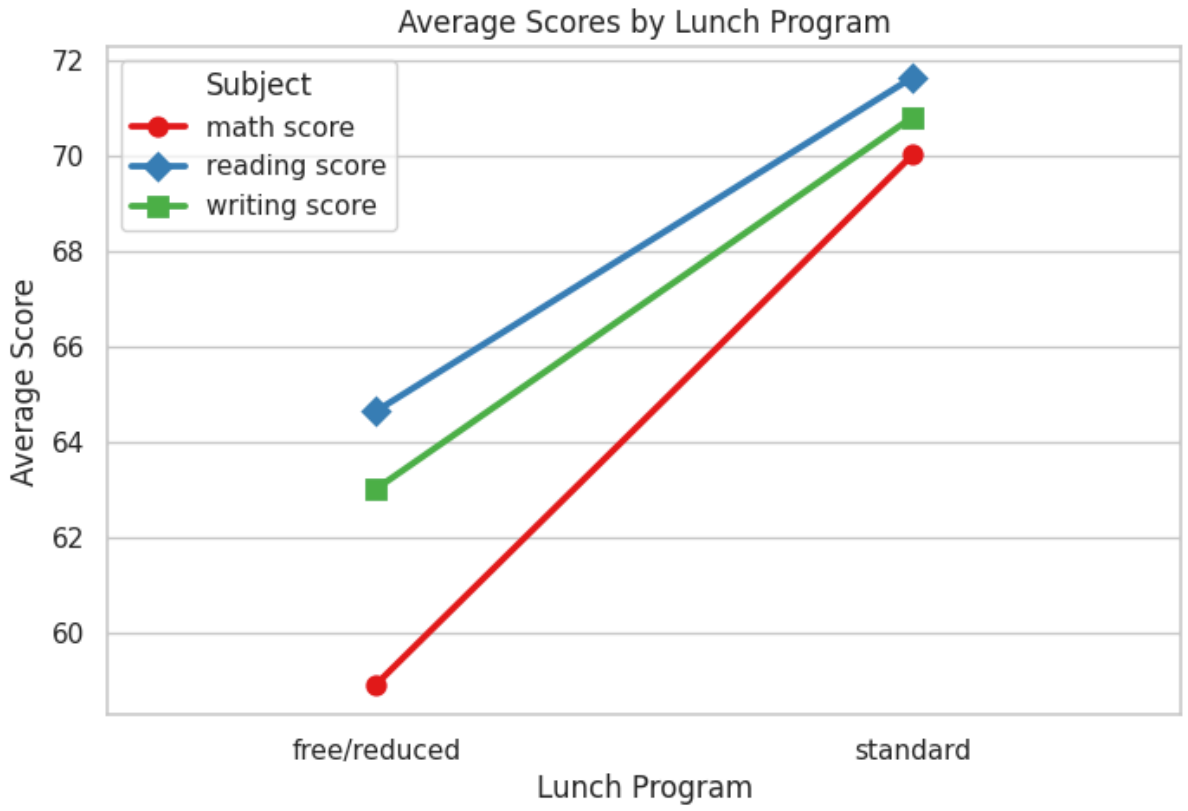


3. Point Plot Lunch Program Effect on Scores:

Visualization: ***Point Plot*** Lunch Program Effect on Scores.

Purpose: Displays average scores for students by lunch program using points connected by lines.

Analysis: The point plot makes it easy to compare average scores between different lunch programs for each subject at a glance, highlighting differences in performance without clutter.

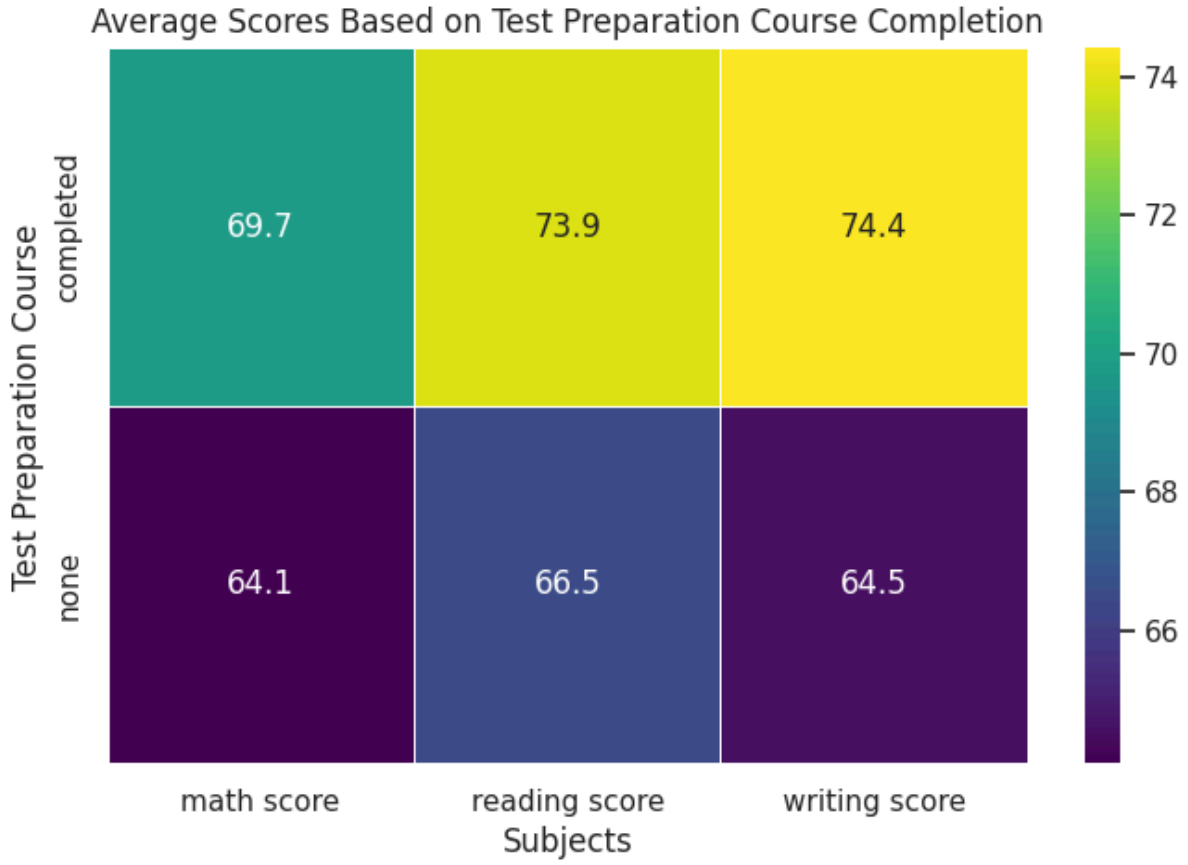


4. Heatmap for Test Preparation Course Impact on Scores:

Visualization: ***Heatmap*** for Test Preparation Course Impact on Scores:.

Purpose: Represents average scores for different subjects based on test preparation course completion in a grid format.

Analysis:The heatmap visually emphasizes the differences in average scores across subjects, making it easy to spot areas of strength or weakness associated with test preparation.



5. Scatter Plot of Math vs. Reading Scores, Colored by Race/Ethnicity

Visualization: Scatter plot with math scores on the x-axis and reading scores on the y-axis, with points colored by the race/ethnicity group.

Purpose: This scatter plot shows the relationship between math and reading scores, and how this relationship varies across different racial/ethnic groups.

Analysis: The scatter plot helps in understanding if students who score higher in math also tend to score higher in reading, and whether this trend differs by race/ethnicity. Clustering of points by certain race/ethnic groups can reveal group-specific patterns in academic strengths or weaknesses.