

# Principles of Data Science – 5530

## Assignment 1

Naga Sai Radhakrishna Siram

Student ID: 16356525

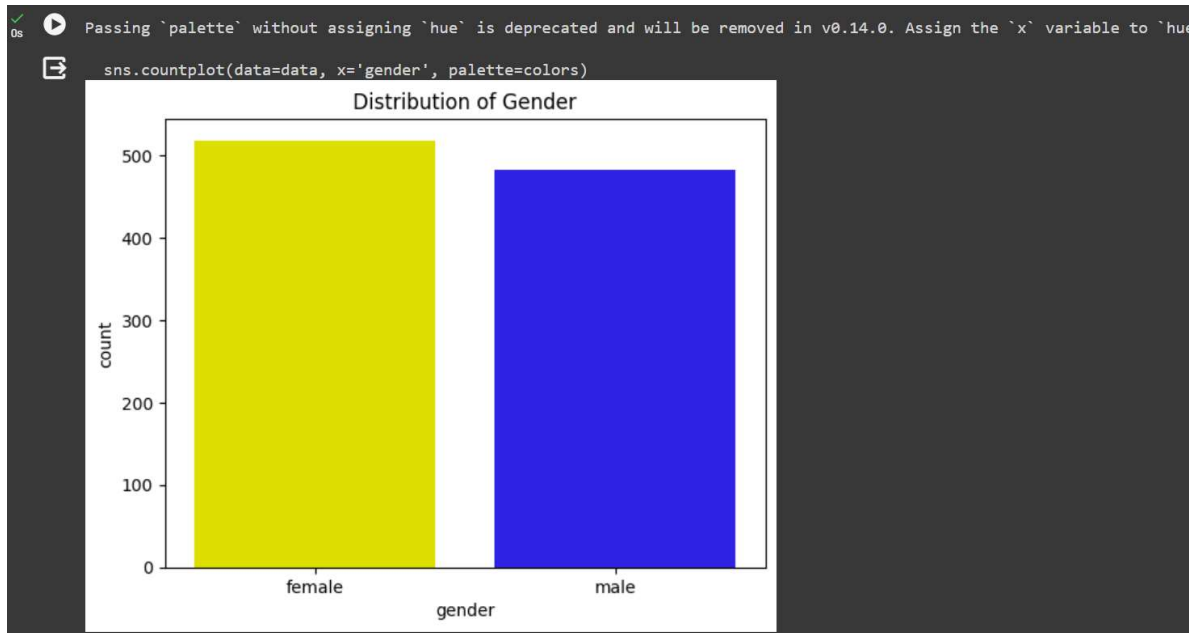
### Question2: Student Performance Dataset

#### Visualizations:

##### Plot 1:

The following bar graph illustrates the distribution of responses based on gender, showcasing the count of male and female respondents.

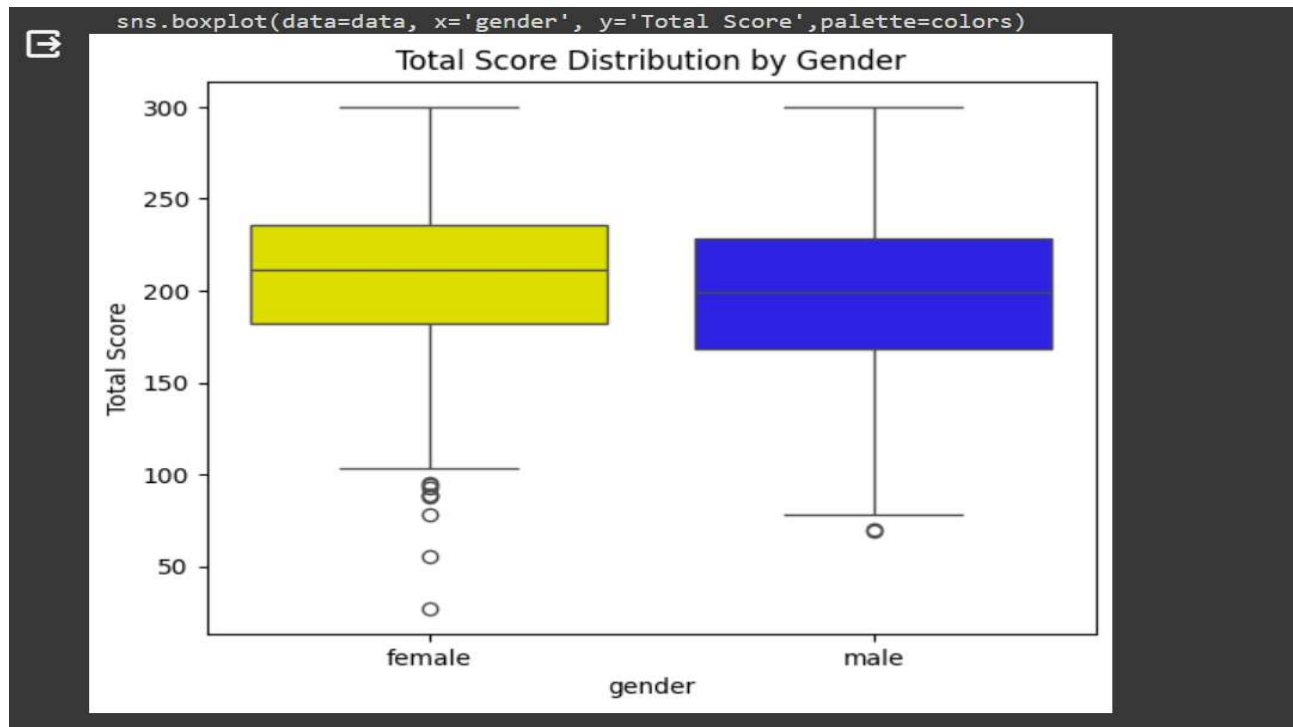
X-axis: Gender (Male, Female) Y-axis: Frequency of Responses



##### Plot 2:

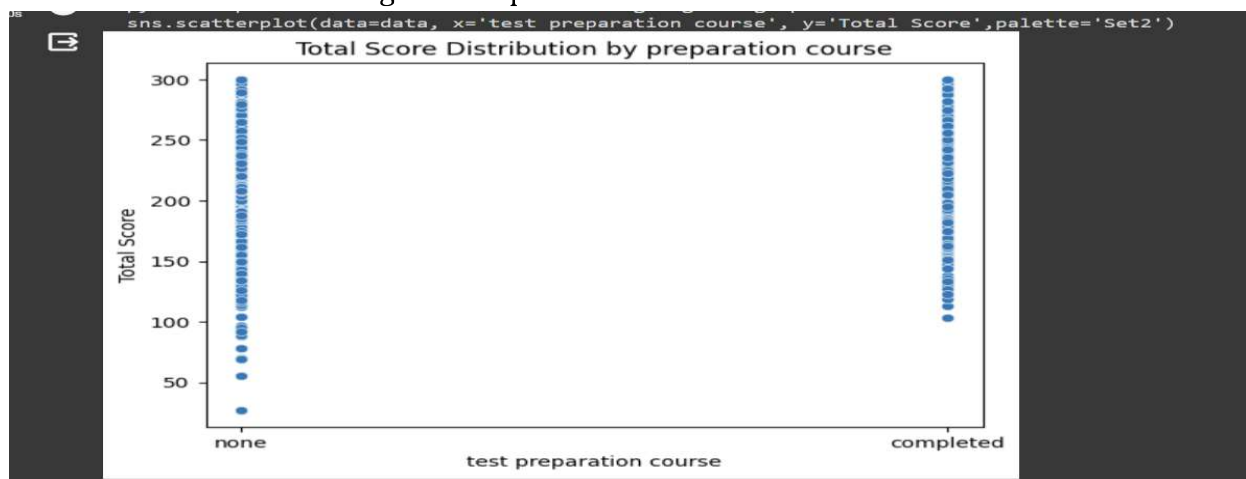
Below graph comparing gender and total score, the X-axis denotes the gender categories (e.g., male and female), while the Y-axis represents the total score achieved. This

visualization allows for a quick comparison of the overall performance between genders and highlights any disparities or trends in the total scores achieved



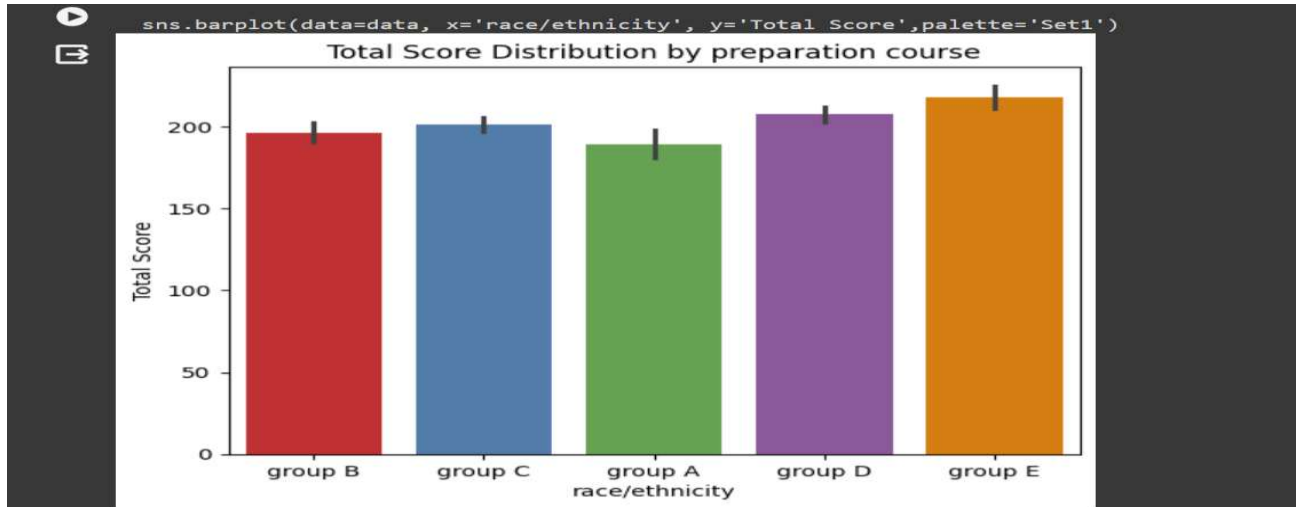
### Plot 3:

On the X-axis, the presence or absence of a test preparation course is represented, with one end indicating no course taken and the other indicating the course was completed. The Y-axis reflects the total score achieved on the test. The plot allows for visualizing any correlation between test preparation and performance, indicating whether those who took the course tend to score higher compared to those who didn't.



## Plot 4:

The X-axis represents the various racial or ethnic categories, while the Y-axis displays the total score obtained by individuals belonging to each group. Each bar on the plot corresponds to a specific racial or ethnic category, with its height indicating the average or total score achieved by individuals within that category. The length of each bar serves as a visual representation of the score distribution, providing insights into any disparities or patterns that may exist across different racial or ethnic group.



## Plot 5:

A scatter pair plot matrix was constructed to explore the relationships between math, reading, and writing scores, as well as their combined total score. The plot revealed strong positive correlations between math, reading, and writing scores, with each score positively influencing the total score. Additionally, a linear relationship was observed between each individual score and the total score, indicating a consistent impact on overall performance.

