

# Principles of Data Science – 5530

## Assignment 1

Nikhil Jagadeesh Sriram

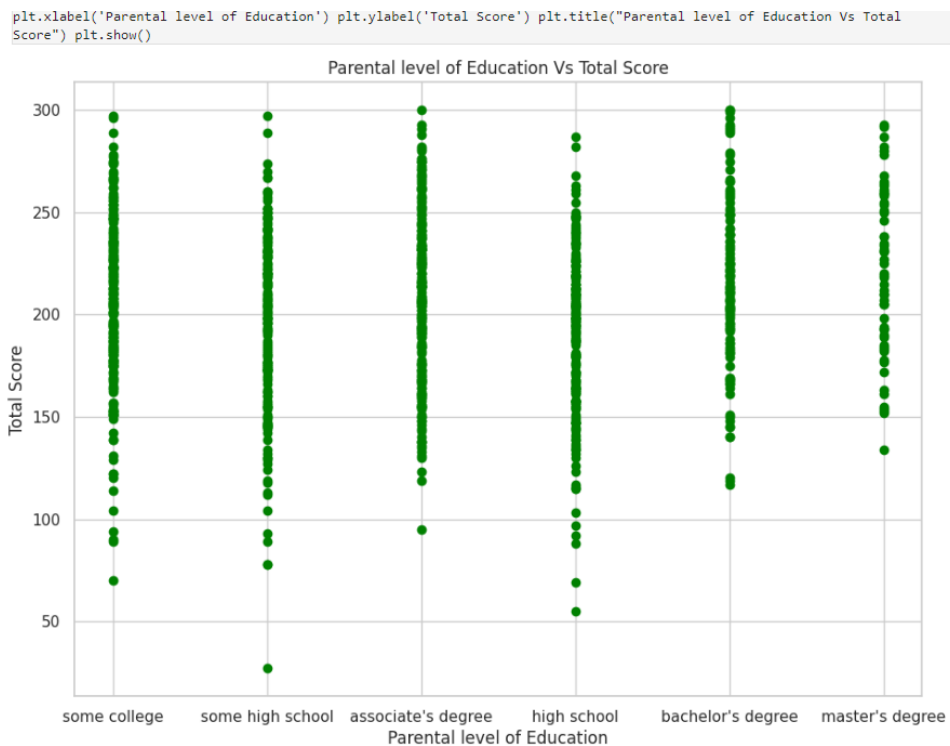
Student ID: 16352573

### Question2: Student Performance Dataset

#### Visualizations:

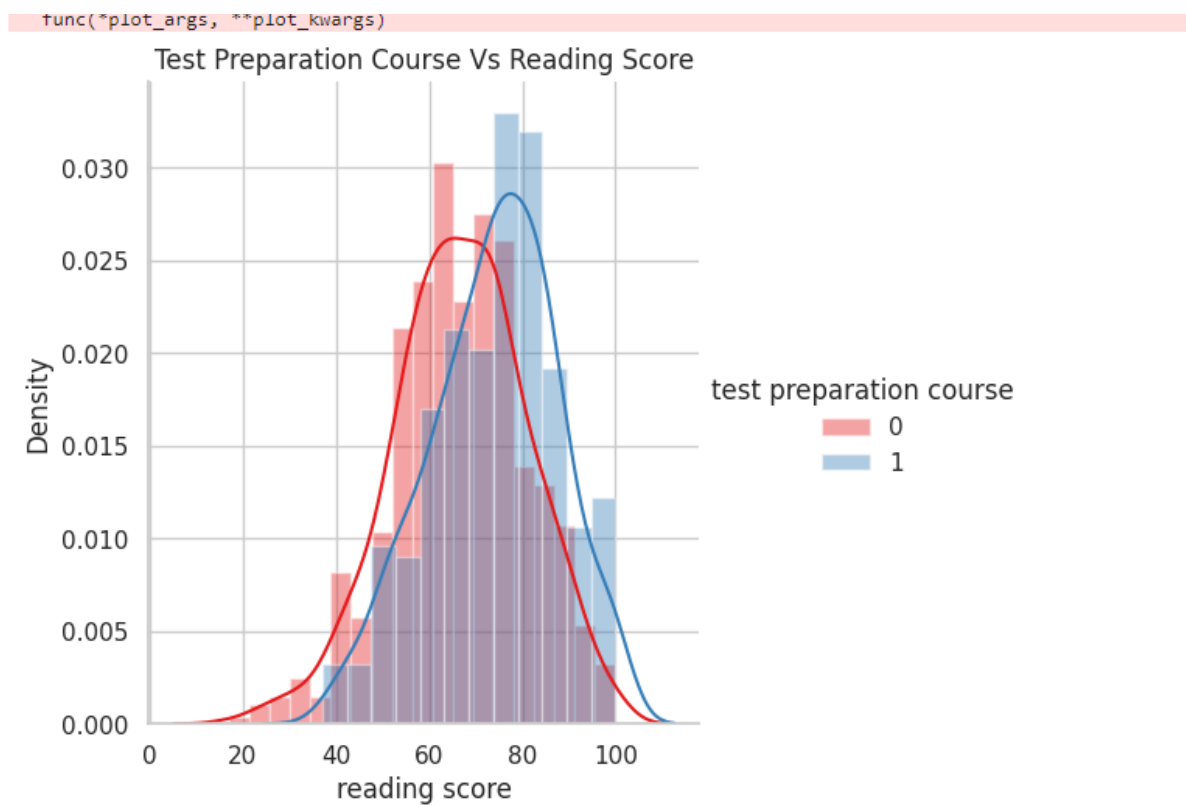
##### Plot1:

A scatter plot visually represents the relationship between total scores and the parental level of education. This graphical representation allows for an exploration of how total scores vary based on different levels of parental education.



## Plot2:

A faceted grid representation visualizes the reading scores concerning the selection of the test preparation course, with a focus on distinguishing between males and females. This graphical representation aims to provide a clear comparison of reading performance based on the test preparation course choice within each gender category.



### Plot3:

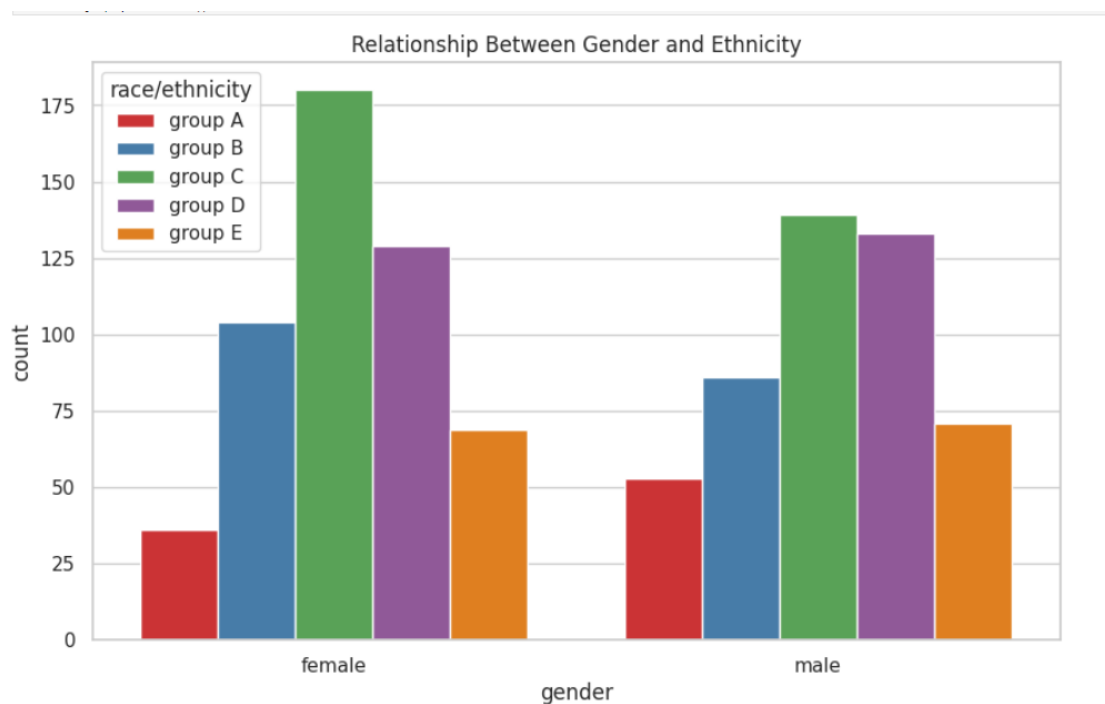
Examining the relationship between males and females based on their lunch behaviour, this analysis delves into whether individuals opt for a standard or free lunch. The goal is to understand potential differences in lunch preferences between genders.

```
[37]: <seaborn.axisgrid.FacetGrid at 0x79597170a000>
```



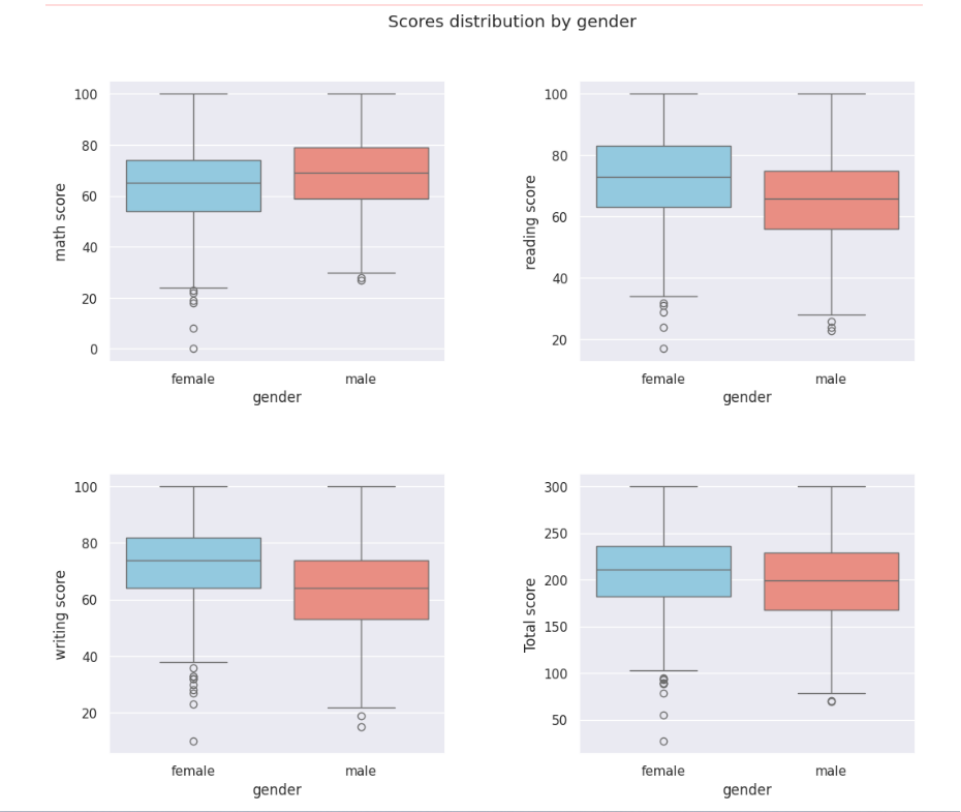
### Plot 4:

Exploring the correlation between gender and ethnic groups, this analysis investigates the distribution of males and females within various ethnic categories. It aims to uncover insights into gender representation across different ethnicities.



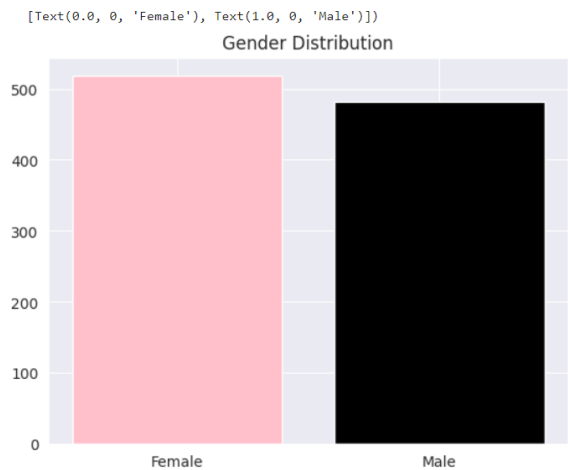
**Plot5:**

A series of four box plots examines gender disparities across different domains, including math, reading, writing, and total scores. These visualizations provide insights into score distributions and potential variations between genders.



**Plot6:**

The horizontal axis represents the categories 'Female' and 'Male,' while the vertical axis displays the respective counts. The visualization employs distinct pink and black bars to distinguish between genders, offering a distinct representation of the distribution based on gender.



## Plot7:

A paired scatter plot illustrates the relationship between gender and four distinct categories: reading score, writing score, math score, and total score. This visualization provides a comprehensive view of how these scores vary between male and female individuals.

