Principles of Data Science-5530

Assignment-2

Snigdha Setty Chandaluri

Student Id:16352688

Question 2: Student performance Dataset

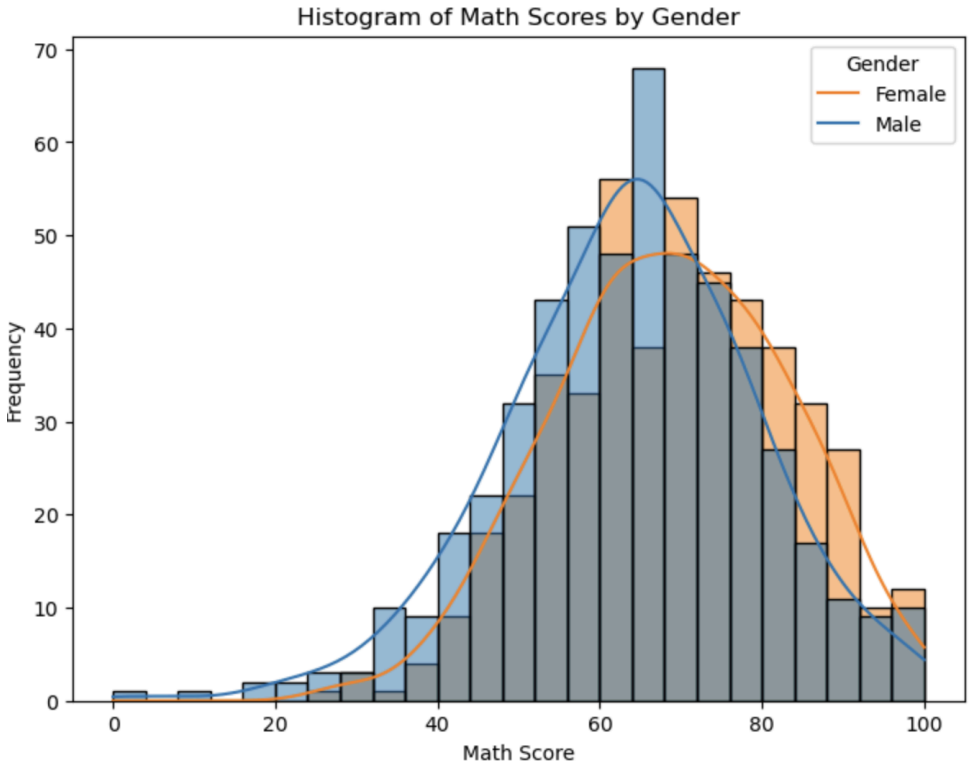
Visualizations:

Plot:1

The below image is the histogram of math scores by gender. The x-axis shows the math score, and the y-axis shows the frequency of that score. The blue bars represent the scores of female students, and the green bars represent the scores of male students.

The histogram shows that there are more females than males who scored between 20 and 40 on the math test. There are also more females than males who scored between 60 and 80 on the math test. However, there are more males than females who scored above 80 on the math test.

Overall, the histogram shows that the distribution of math scores is different for males and females.



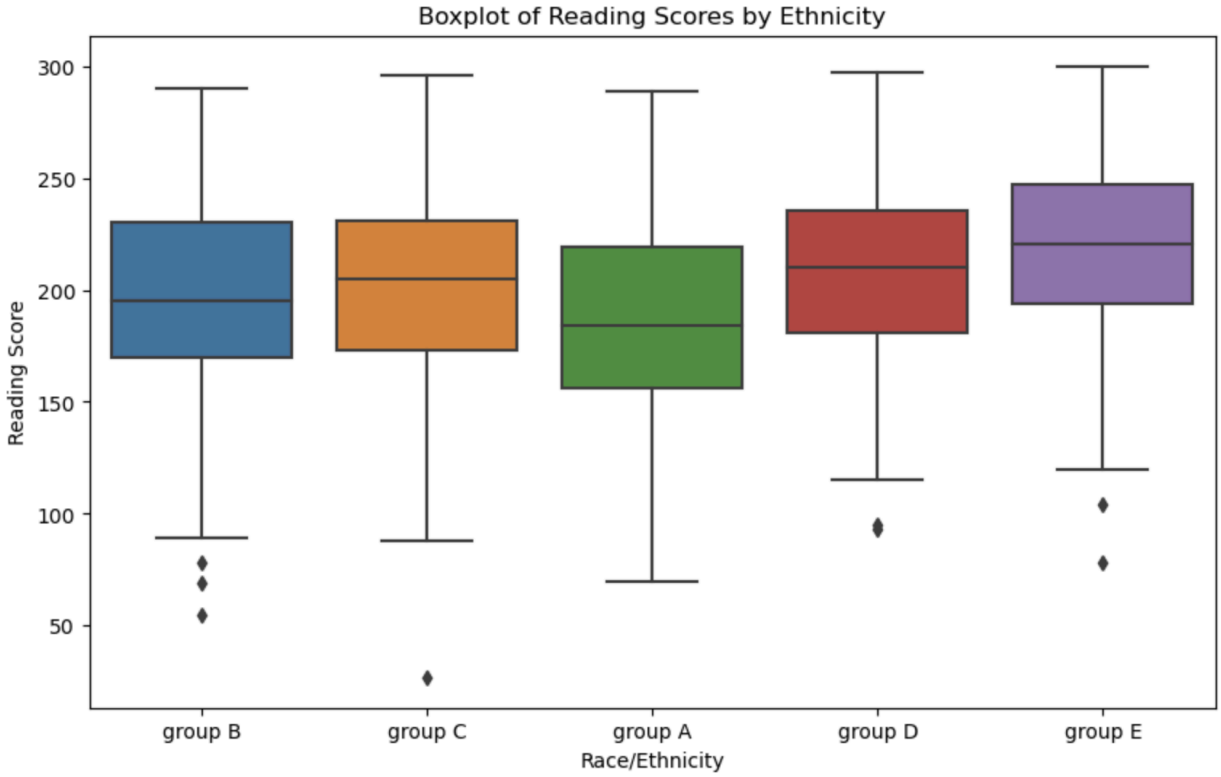
Plot 2:

Plot 2:

Imagine a bunch of boxes lined up next to each other. Each box represents the reading scores of a different group of people based on their ethnicity.

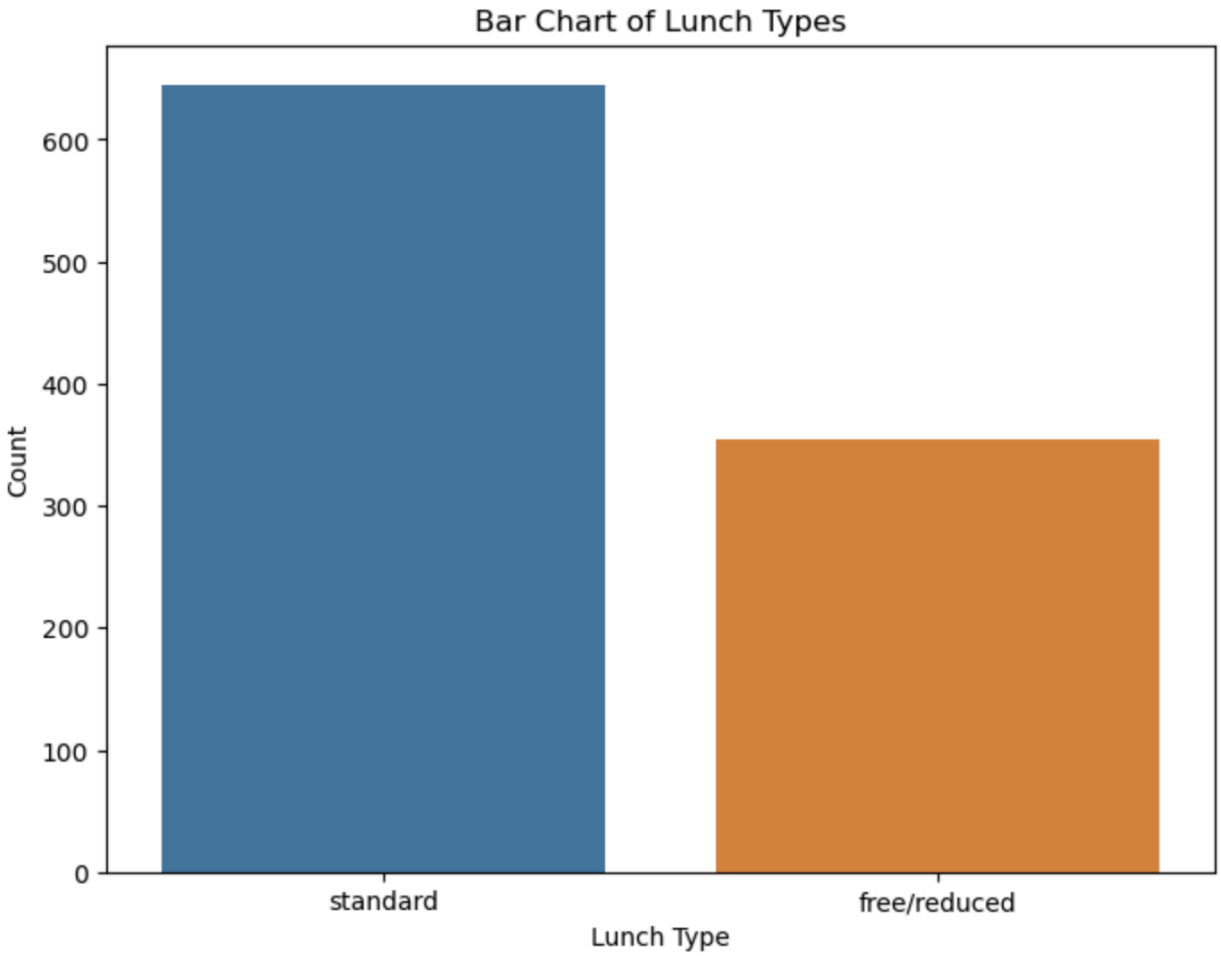
Inside each box, there's a bunch of dots. Most of the dots are clustered in the middle of the box, showing the typical scores for that group.
A line in the middle of the box shows the score that half the people scored higher than and half scored lower than.
The ends of the box (like little fences) show the range of scores that most people got.
Any dots way outside the box are outliers, scores much higher or lower than most people.
Remember, this picture doesn't tell the whole story. Scores might be bunched up in some groups and more spread out in others. There could be a few people with really high or low scores, but we don't know exactly how many.

So, while the picture shows some differences in reading scores between different groups, it's important to remember that it's just a snapshot, and there's more to the story than what you see at first glance.



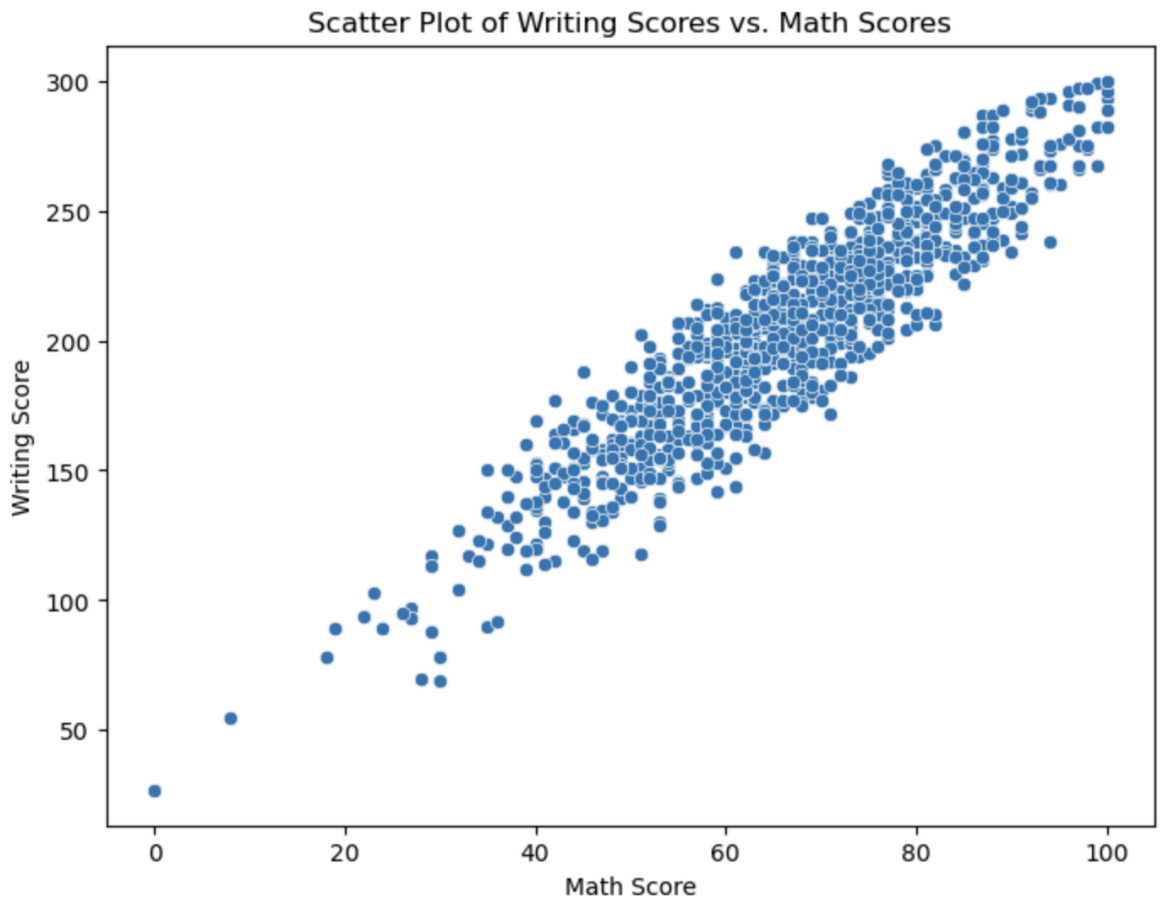
Plot 3:

The image you sent me is a bar chart that shows the number of different lunch types offered. The x-axis shows the lunch type, and the y-axis shows the number of lunches of that type. For example, there are more standard lunches offered than free/reduced lunches.



Plot 4:
The scatter plot showing the relationship between writing scores and math scores. Each dot in the plot represents a student's scores on both writing and math tests.

Higher up on the y-axis means a higher writing score.
Further to the right on the x-axis means a higher math score.
The dots are spread out across the graph, showing that there's no single pattern to how well students do in writing compared to math. Some students score high in both, some score low in both, and many score somewhere in between.

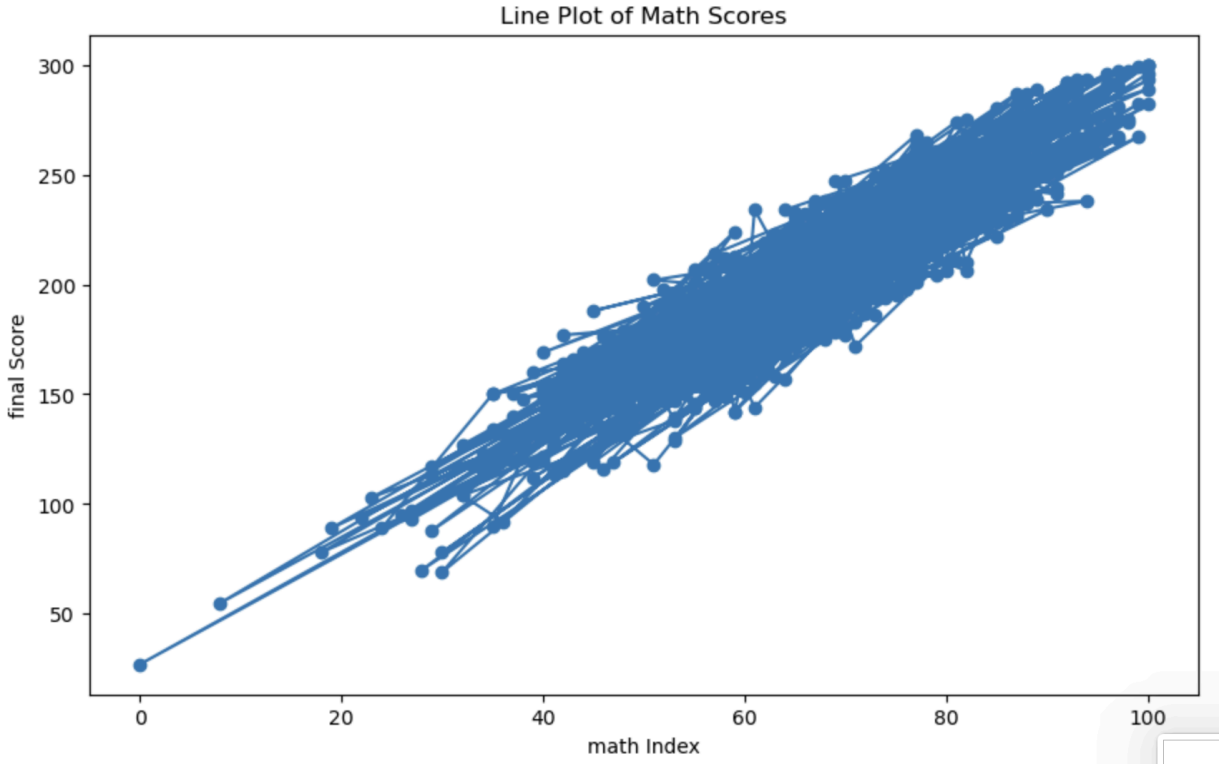


Plot 5:

The line plot of math scores against math index. It appears to show a positive correlation between the two variables. This means that as the math index increases, the math scores also tend to increase.

Here's a more detailed breakdown of the image:

The x-axis shows the math index, which seems to represent some kind of indicator of how challenging the math problems are.
The y-axis shows the math scores, which could be, for example, the percentage of correct answers.
Each dot on the line represents the average score for a certain math index level.
The line slants upwards from left to right, which indicates a positive correlation. This means that students tend to score higher on the math tests when they are given problems from a higher math index (presumably more challenging problems).


Line Plot of Math Scores
final Score
math Index
300
250
200
150
100
50
0
20
40
60
80
100