Sishir Yerra Doddi

Data Visualization

Homework 1

Visualization of Student Performance

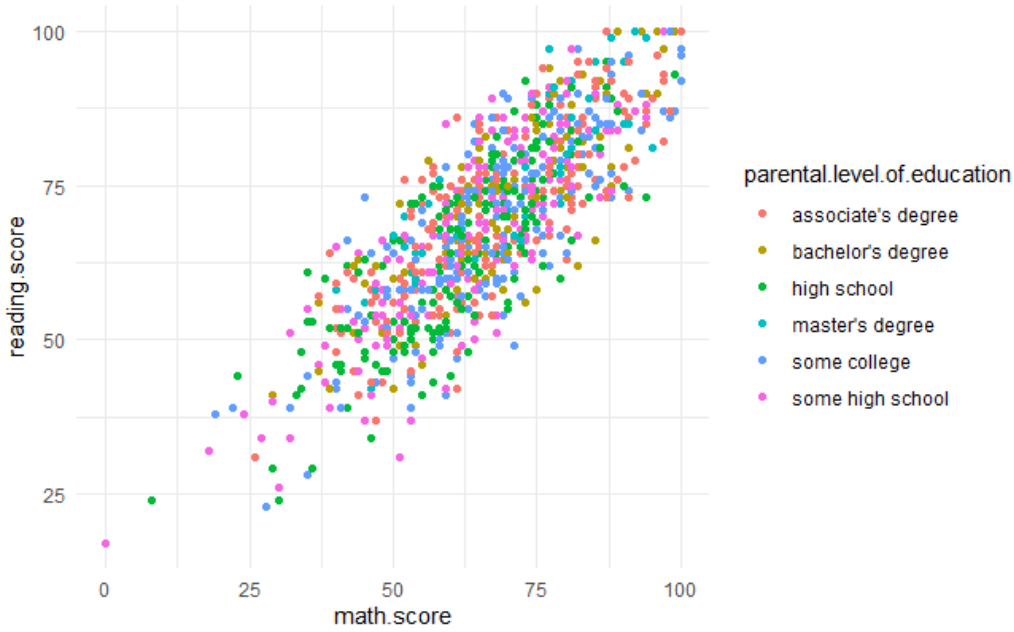
SUMMARY

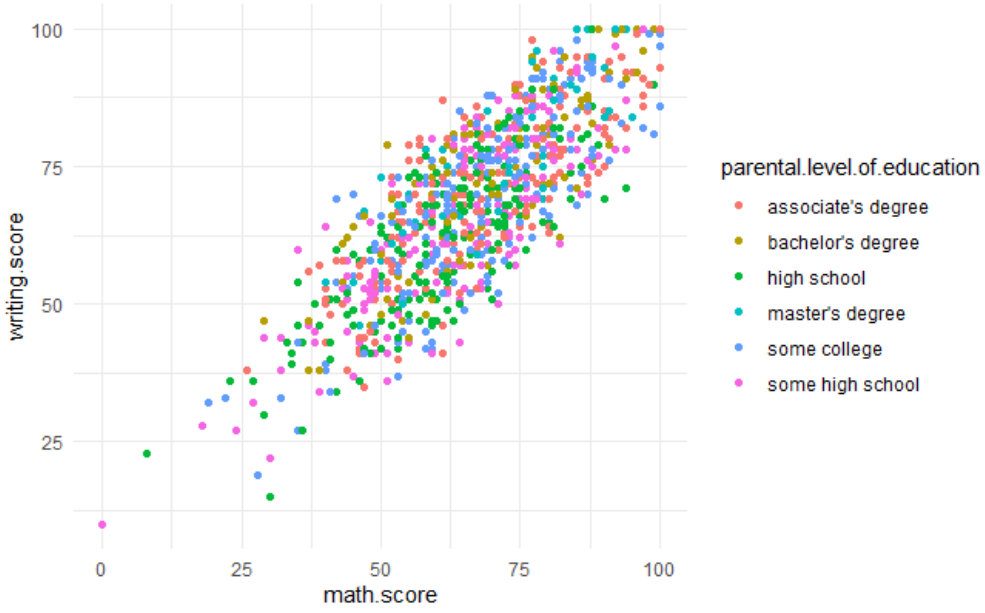
The topic I choose is about student performance. By looking at the data set one could interpret the student represent a students GRE or SAT scores. The dataset further separates the scores based on gender and level of education. The reason I choose this dataset is simply because the limited data felt easy to use and visualize on R. I initially generated plots on a different data sets, however it was hard to work with massive amount of data which could have been easy to work with prior experience in dplyr. Working with this dataset gave me better understanding of how I can use R for visualization.

I have downloaded the data in csv format from Kaggle with the help of google database. I have decided to work with gender, parental level of education and different subject scores. Upon understanding the data. I wanted to learn the relationship between the test score and the factors that are influencing the test scores. I decided to compare math scores with reading and writing scores to understand their correlation and also included the education level as a factor to understand if higher degree of education help achieve better scores or not.

PLOT 1

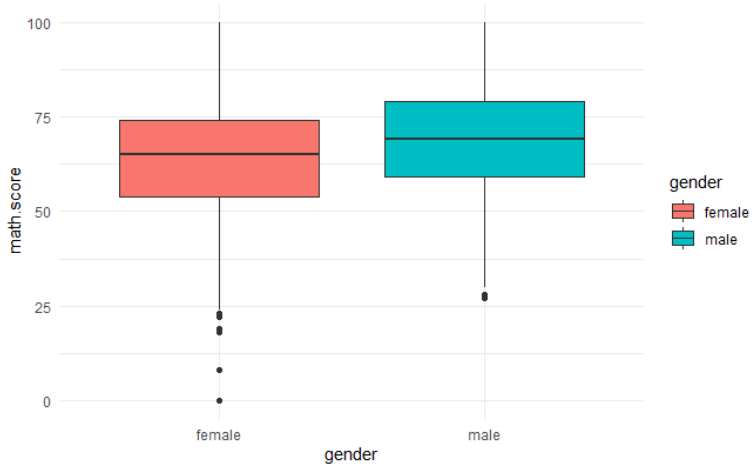
For the first plot I choose scatter plot to understand the relationship math scores had with writing and reading scores based on the level of education. I decided to choose math as the common variable to find out in which category the math scorers had a better performance in.

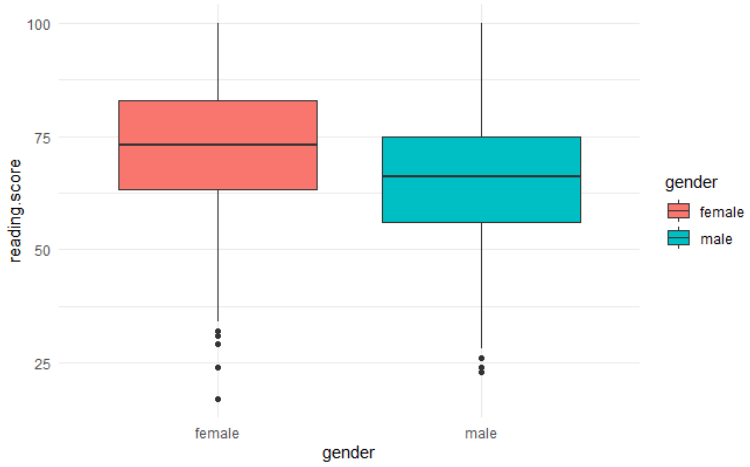


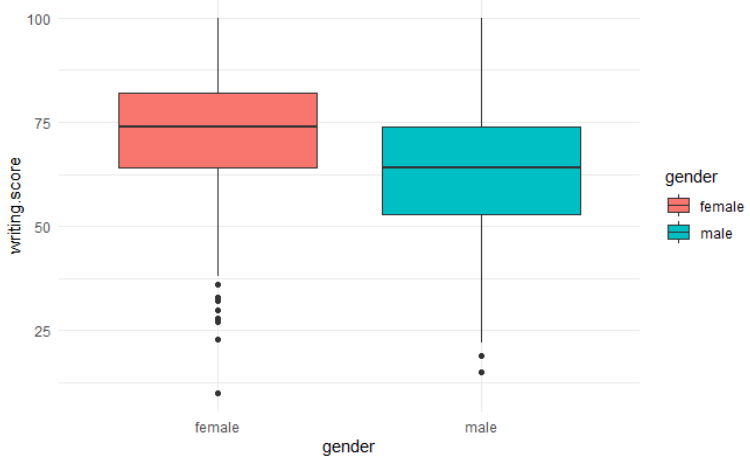


PLOT 2

For the second plot, I decided to work with bar plots to find out how dispersed the data is and to find any potential outliers. I have chosen gender as a common variable to learn who has better performance and find any sign of skewness among the different courses.

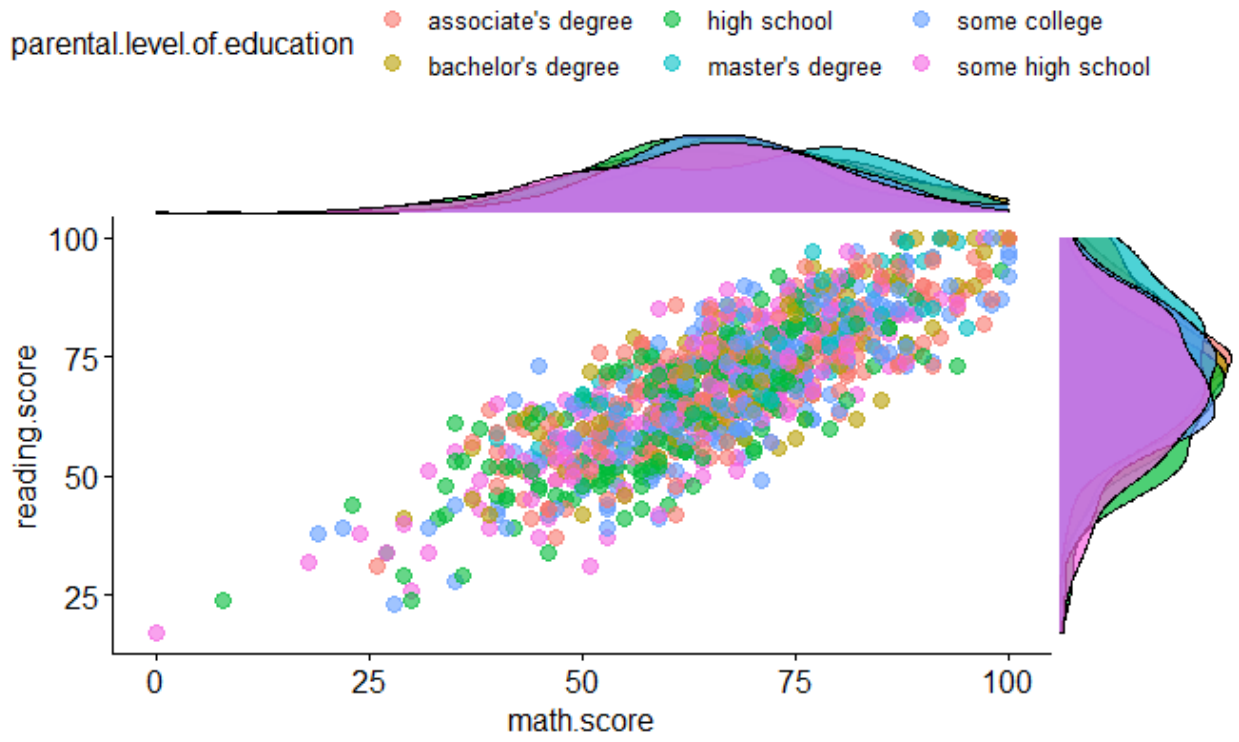






PLOT 3

For the third plot, I deiced to use a scattered histogram plot to learn about the density of each score based on the level of education.



Reference

[Students Performance in Exams | Kaggle]