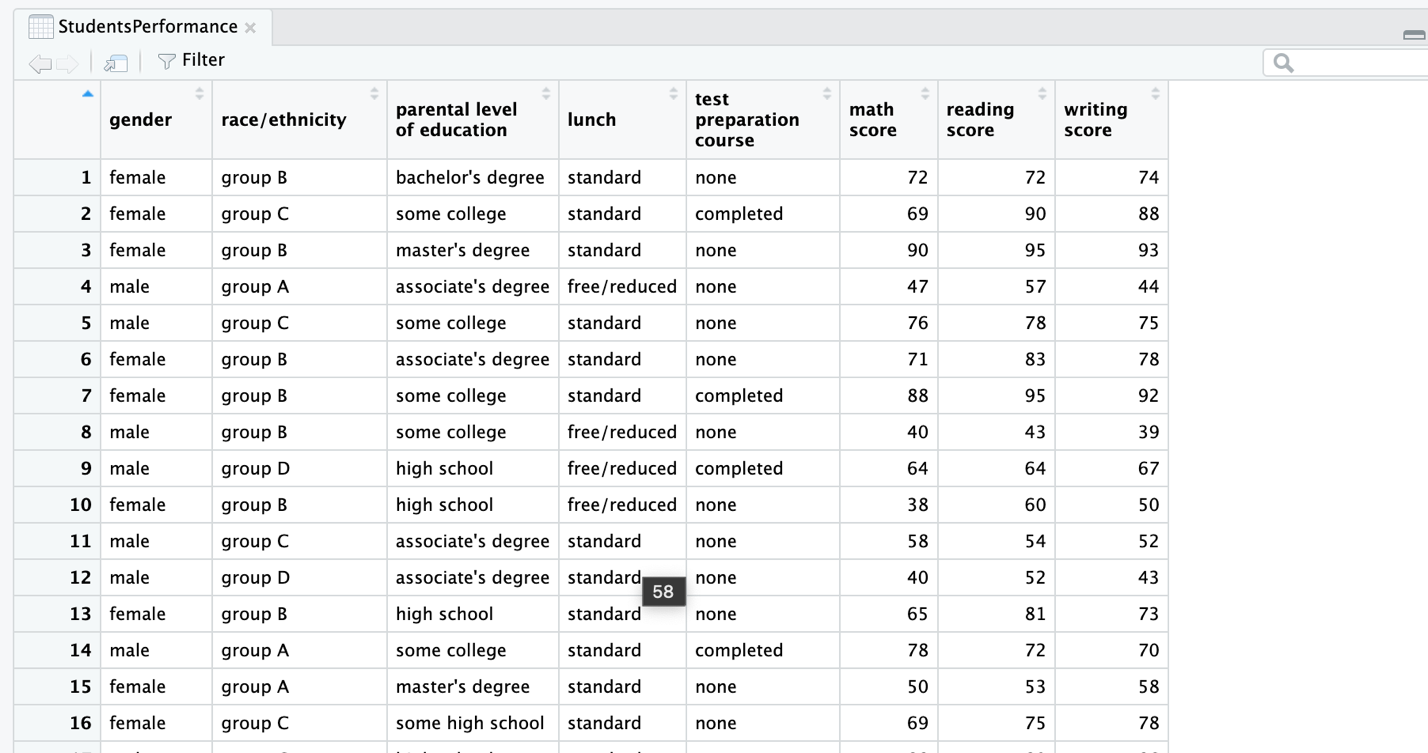
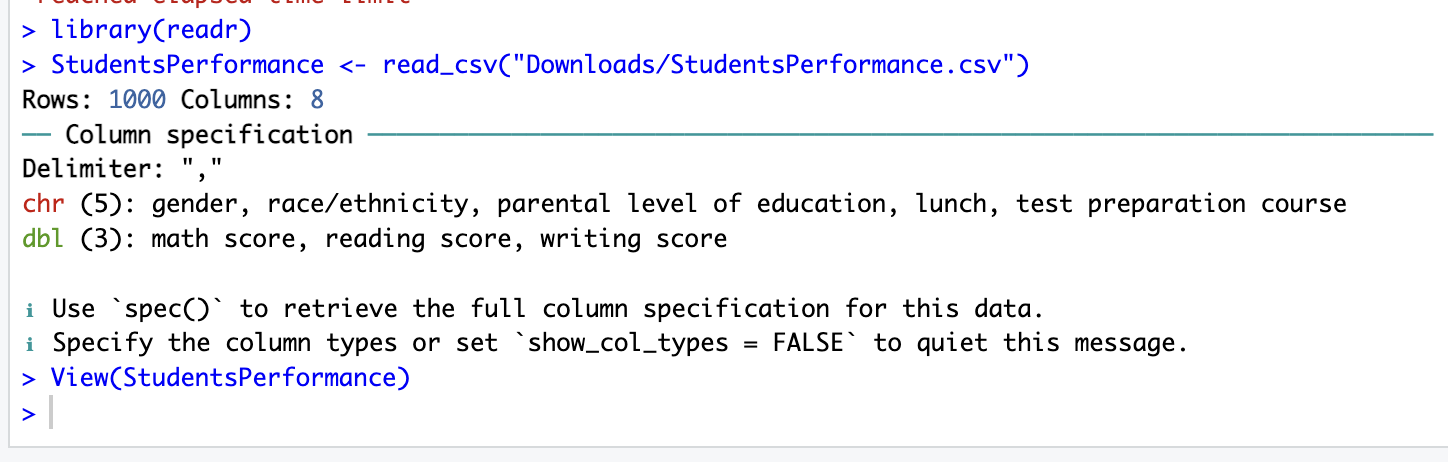
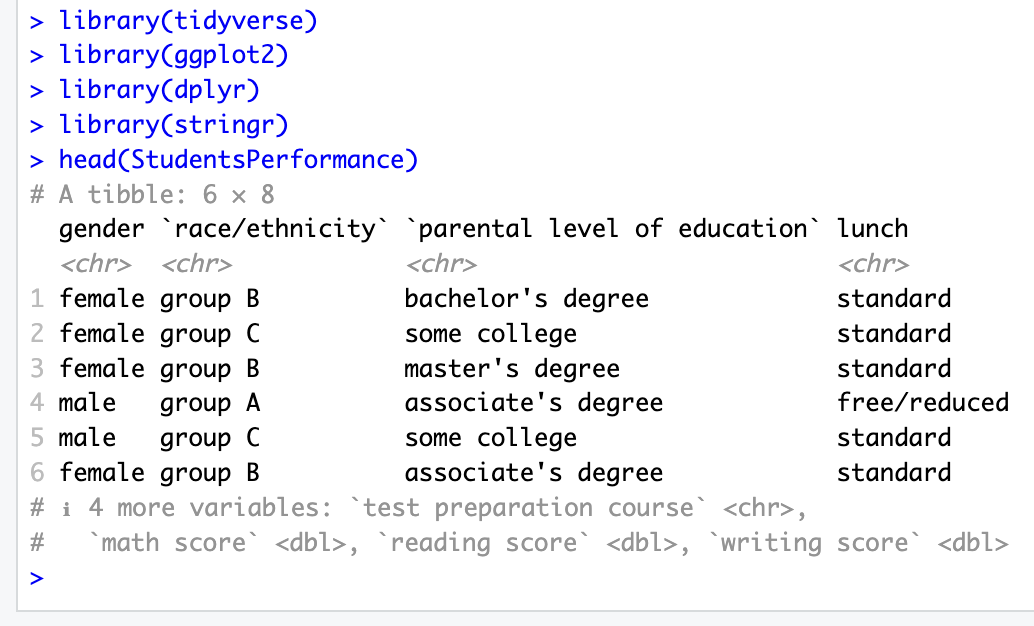
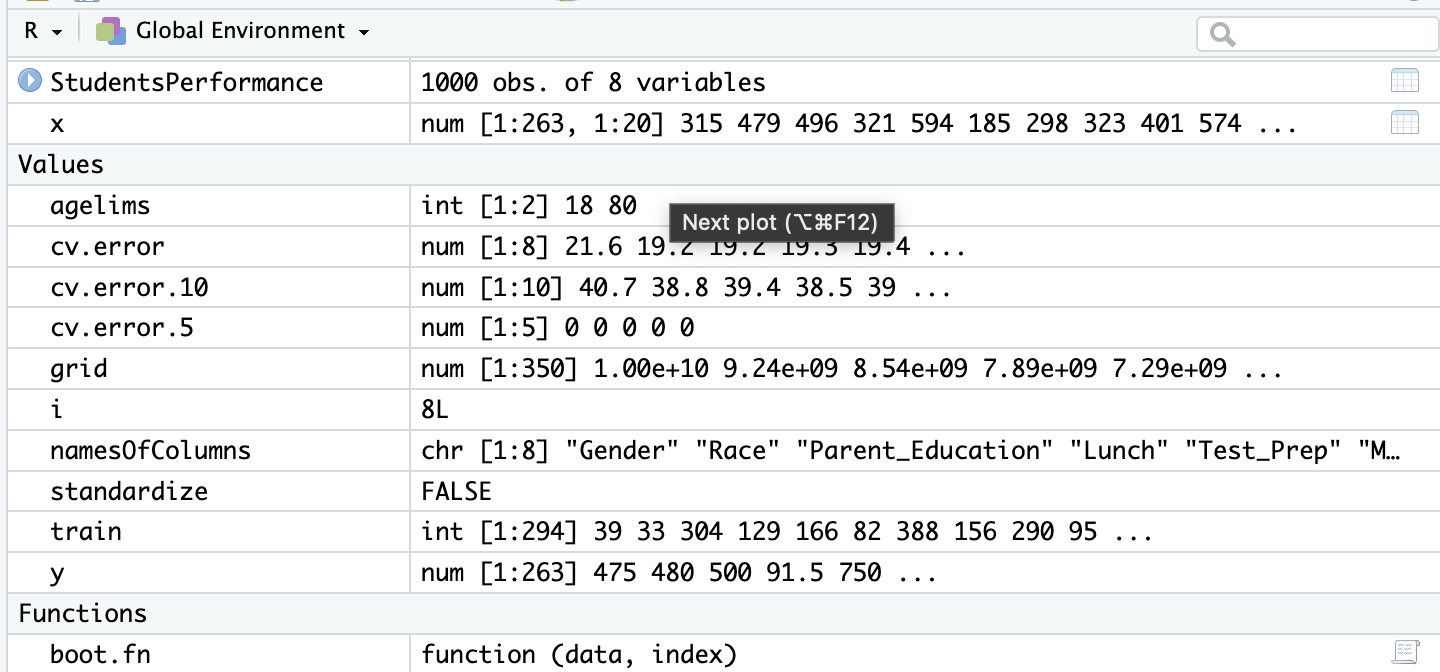
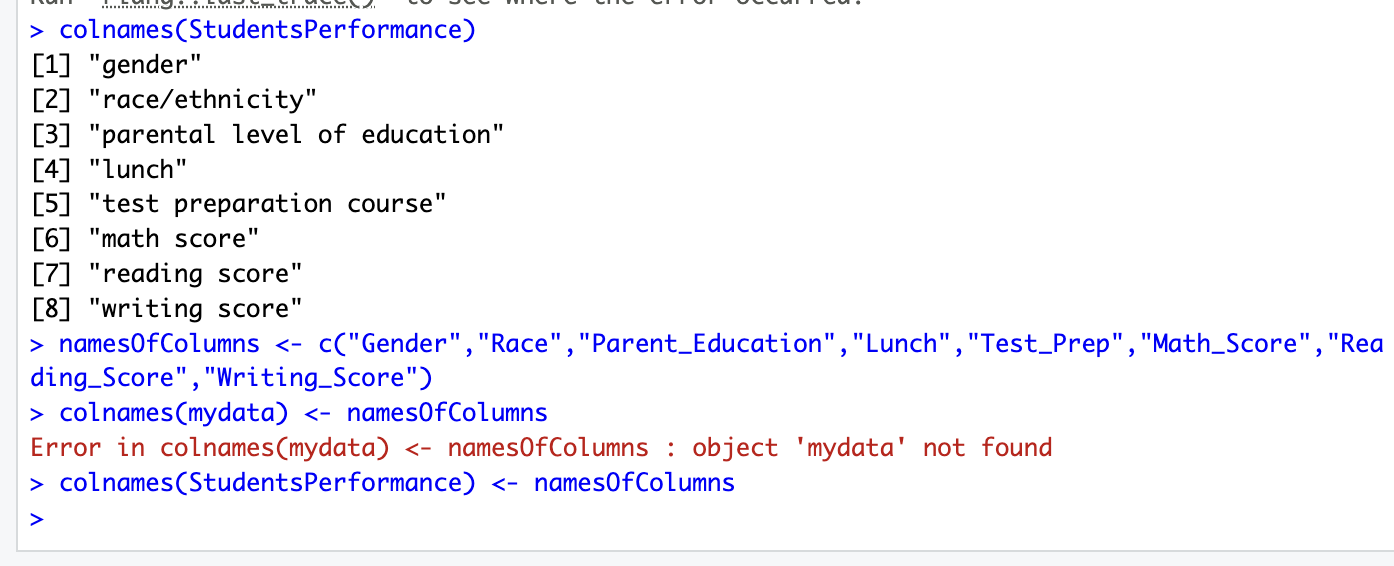
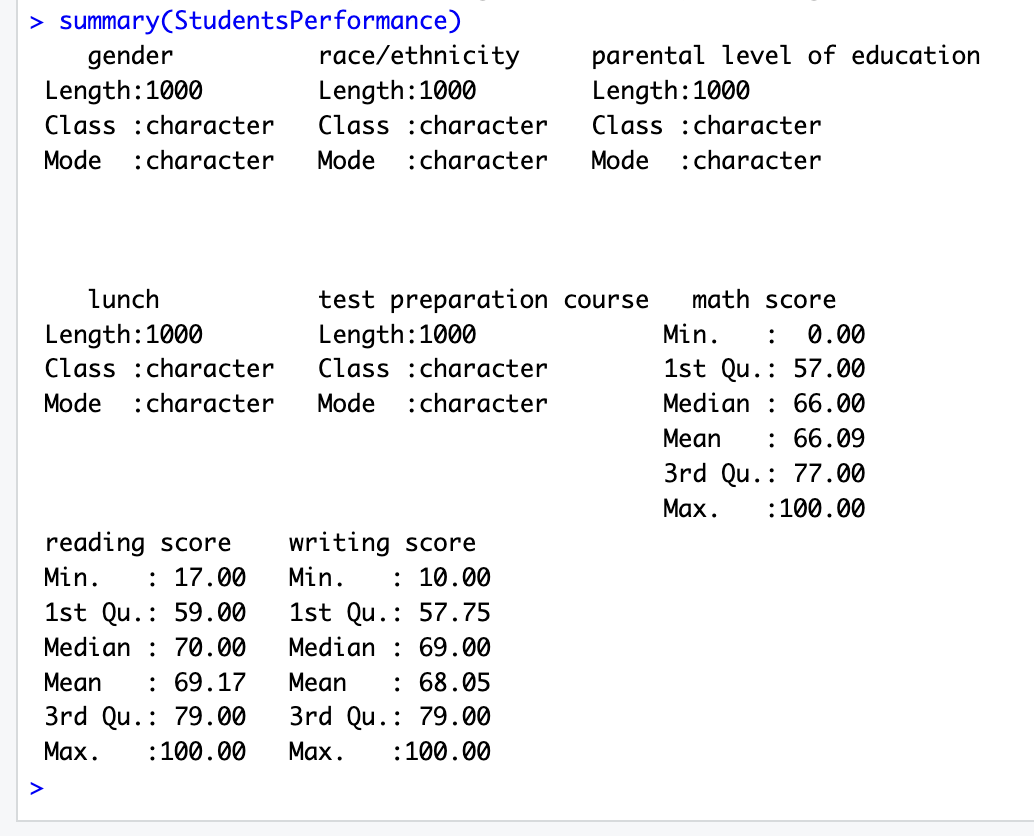
PDS Assignment-1

2.)

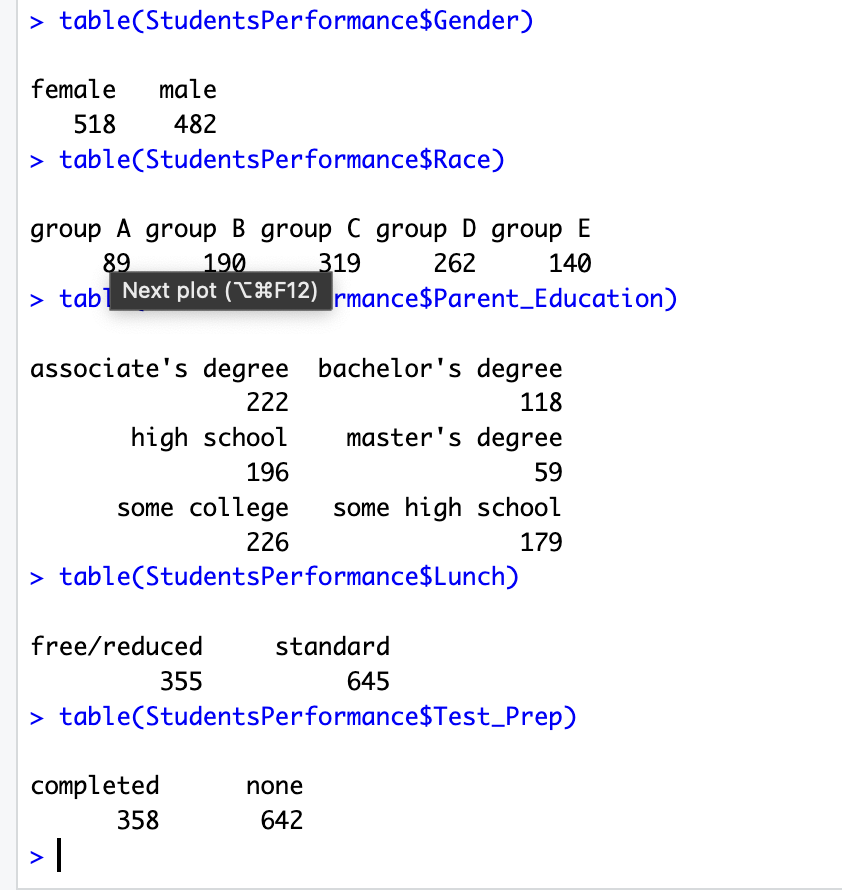
The given dataset is loaded into Rstudio, this dataset contains 8 variables and 1000 observations





From Summary function we can know information about each variable such as Datatype, NA values. I have renamed the variables for cleaner look.

Using the table function I categorized each variable to see how many observations are in each variable

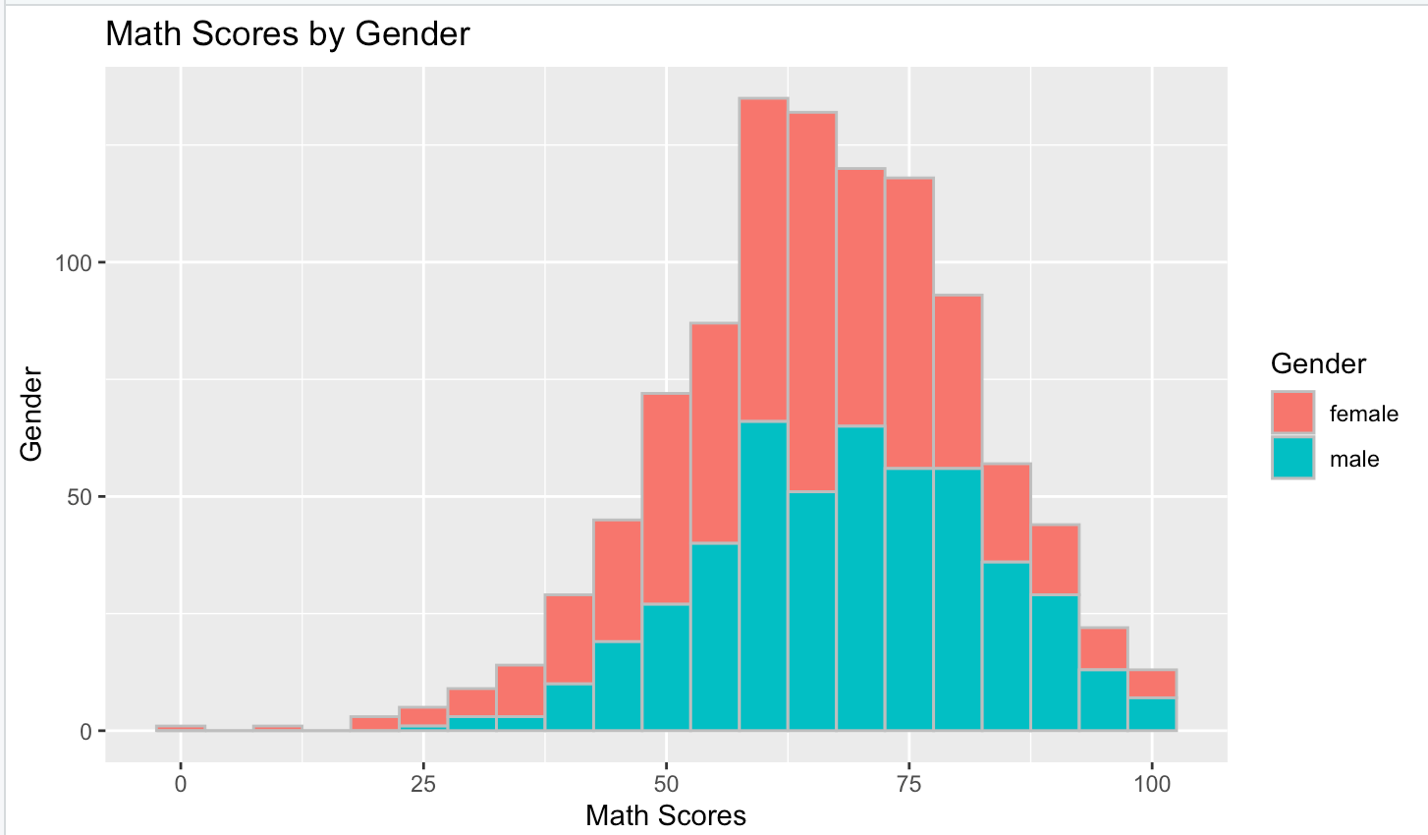
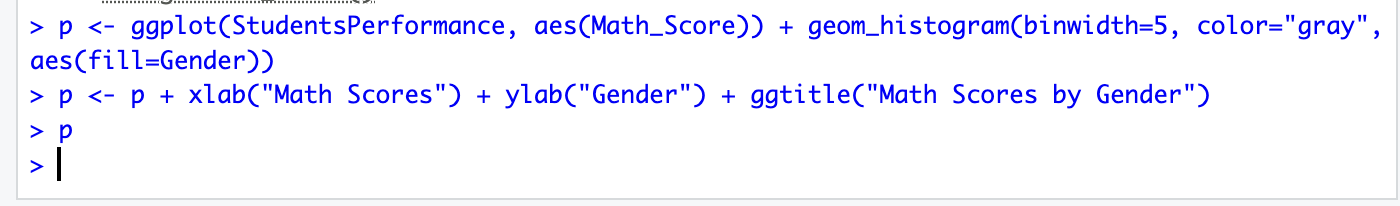


Data Visualization

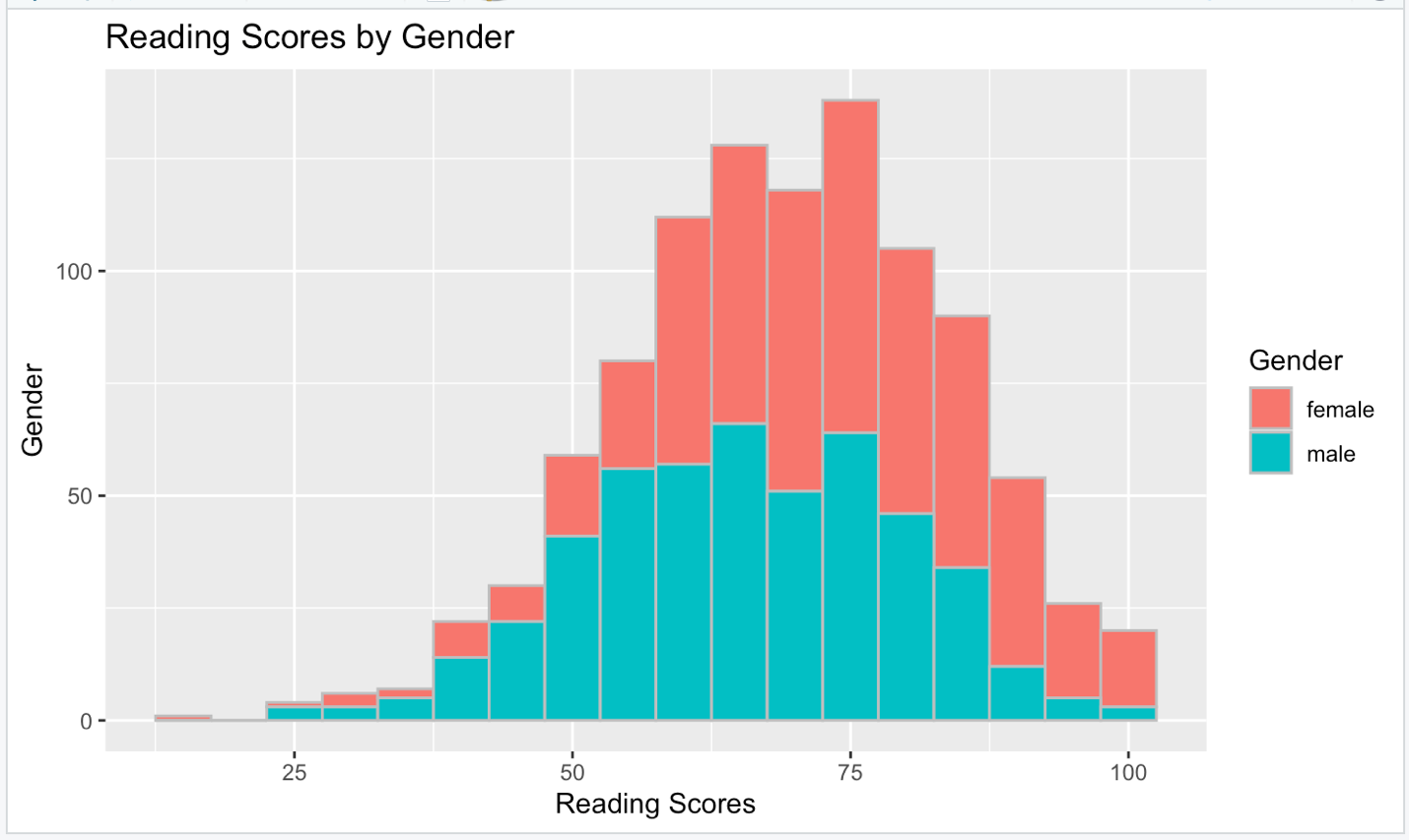
The visualization is used to find the visual patterns which helps us understand more about the data provided.

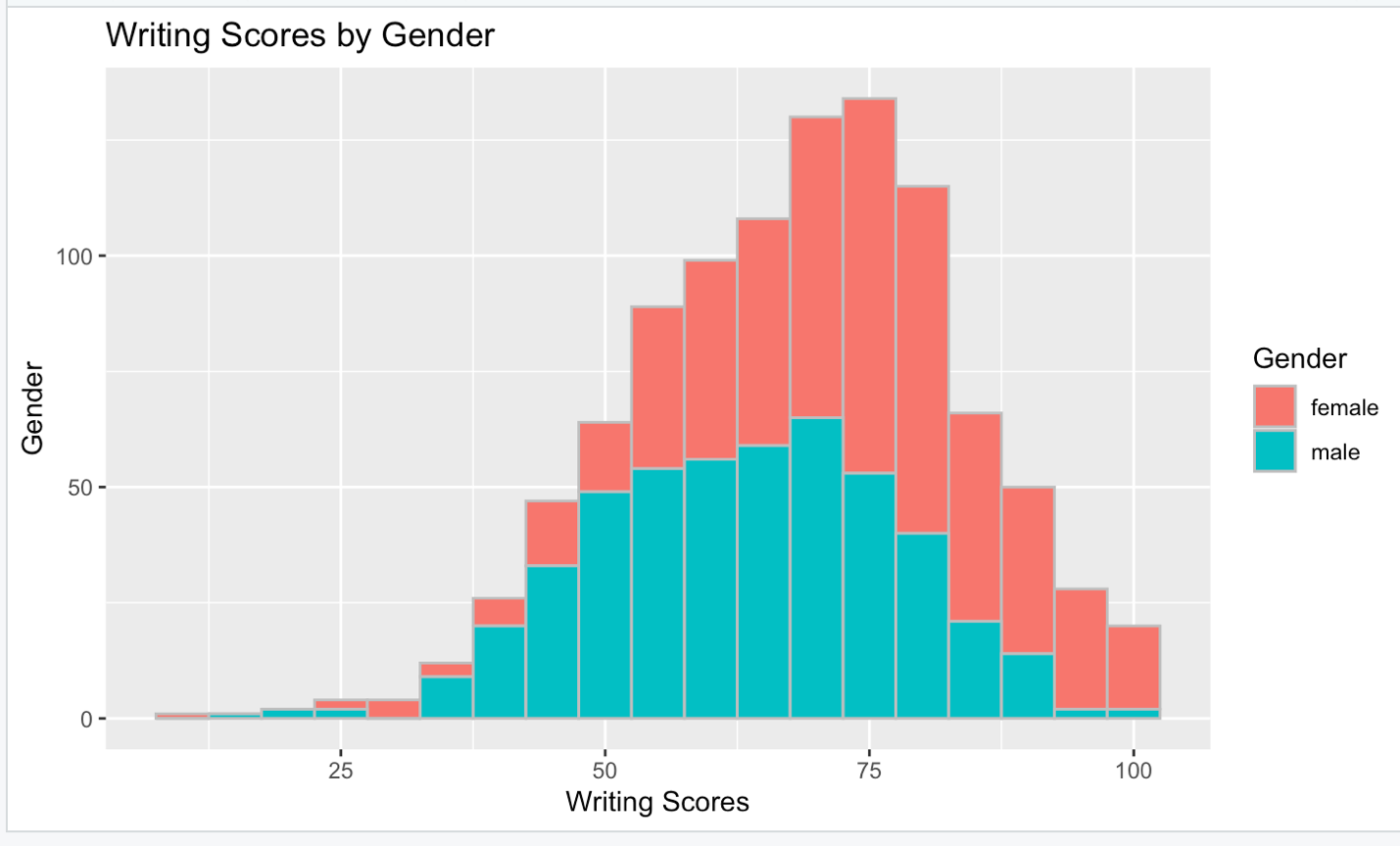
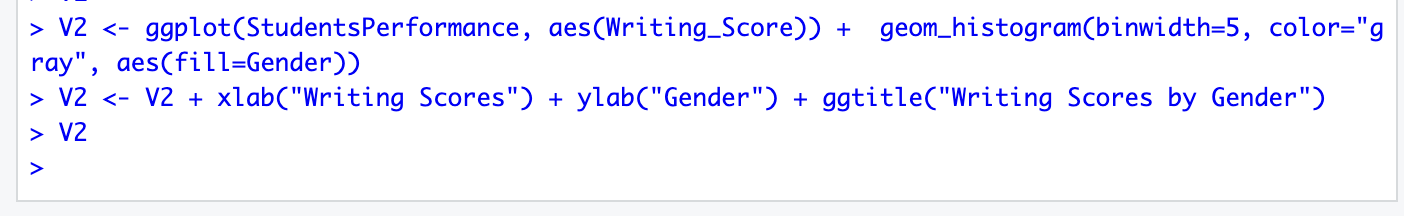
The next three Bar Graphs are plotted based on the gender on and the maths, reading and writing scores

Out of the three variables Maths score, Reading score and Writing score, Females have scored more in the Reading and Writing but males scored more in math on average.

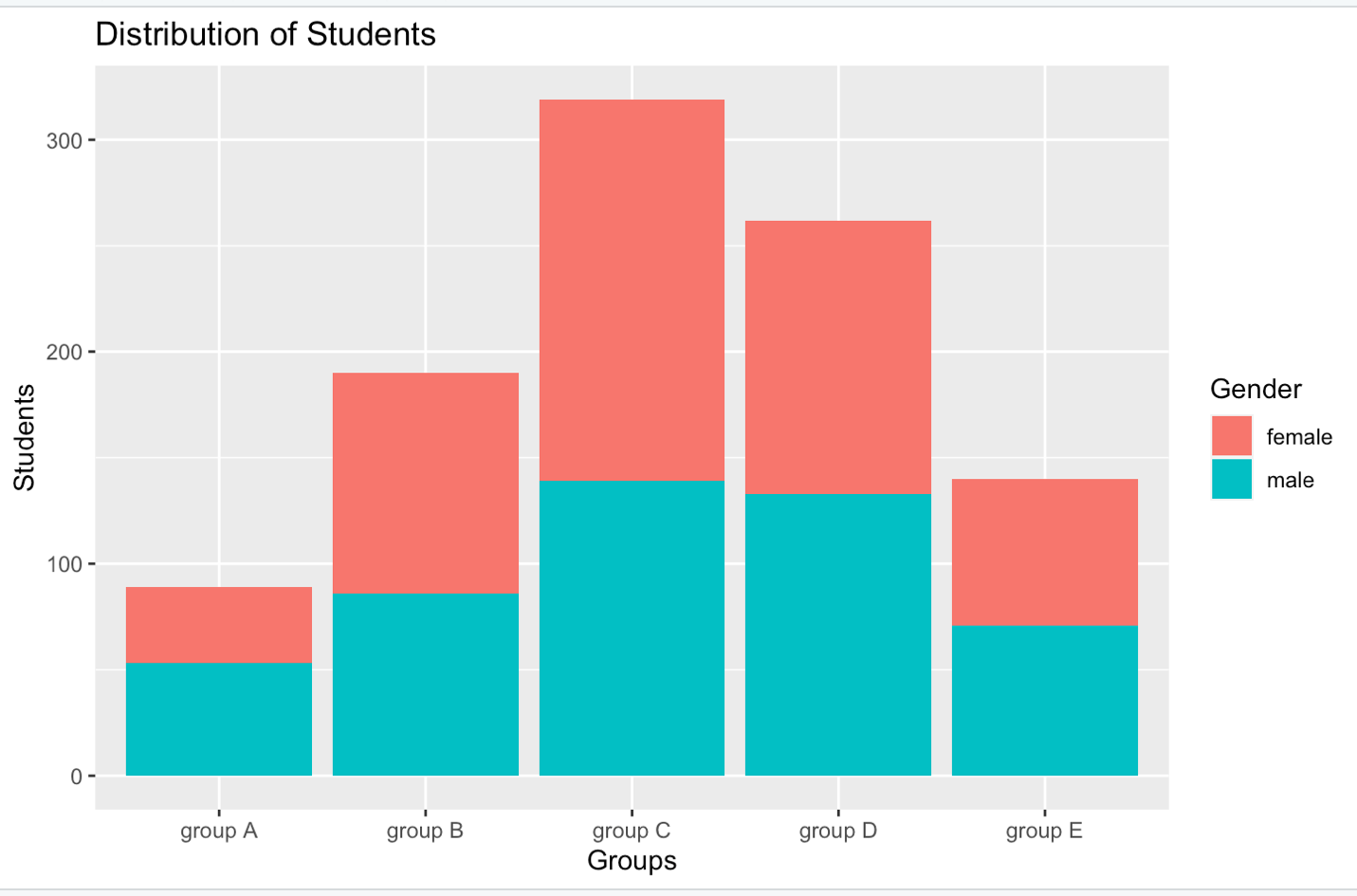


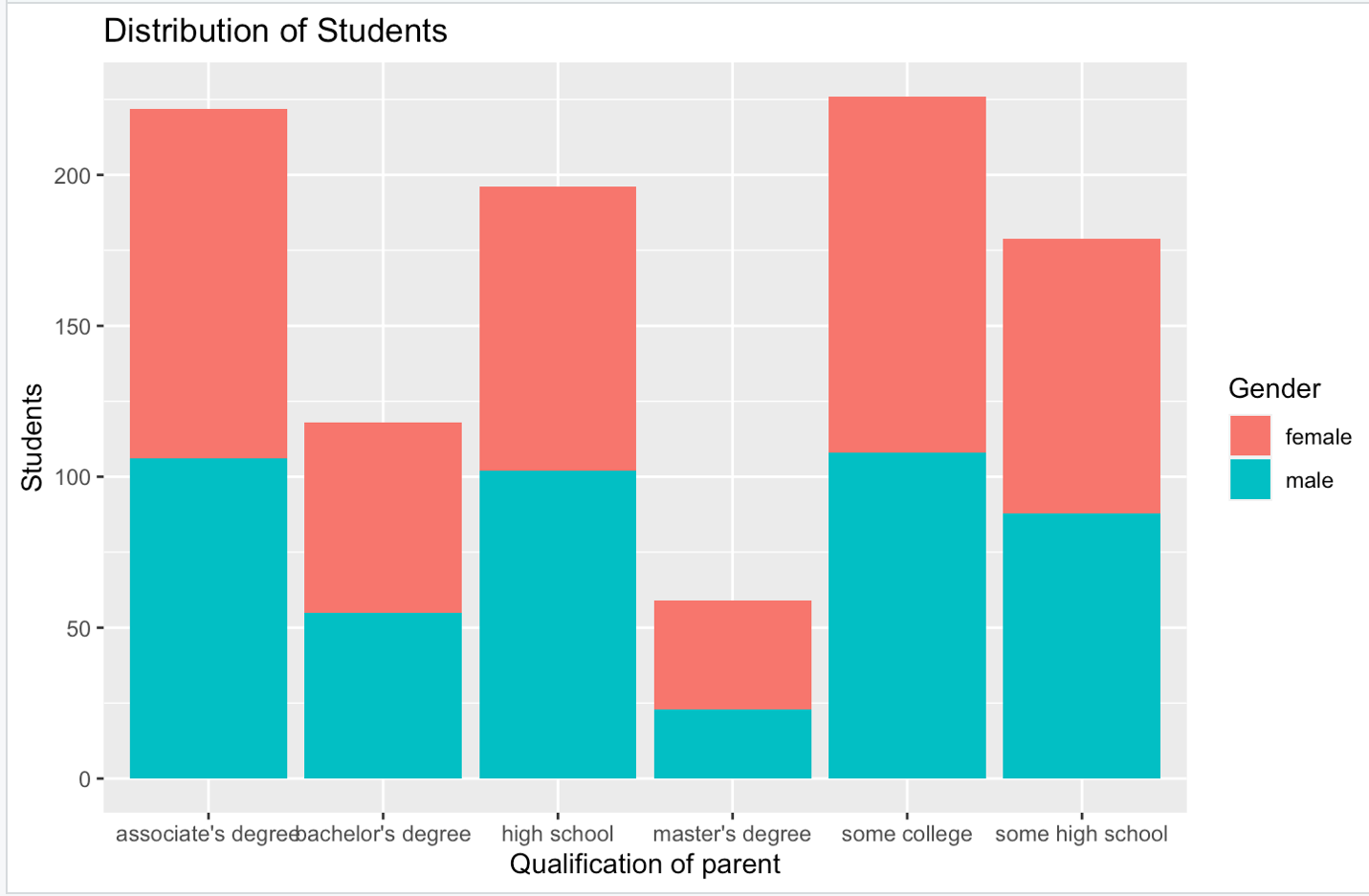
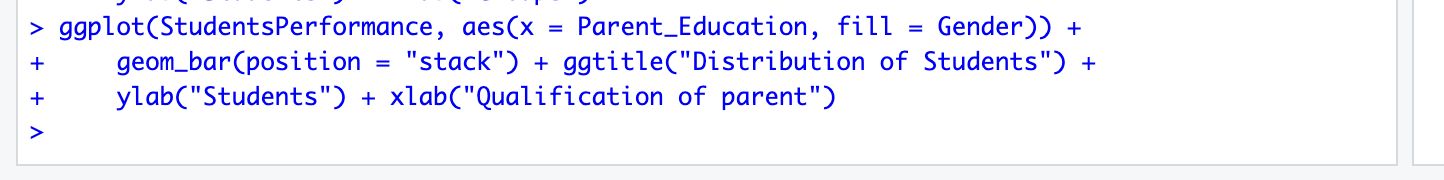
A screenshot of a computer

Description automatically generated



A screenshot of a computer

Description automatically generated

This Graph shows the ratio of Female to male in each Race and also the group C has more students followed by group D, B and E; group A has the least students.

From this histogram between students and parent qualification it can be observed that parents with master’s degree is very low and parents with associate and some college degree has slight variation. The ratio of female to male in the parent\_Education can be observed.