Summary:

In this assignment, the data should be analyzed in a way to help in making decisions regarding the future school budgets and priorities. Two datasets were given one about 15 unique schools and one about all the students in them, which in some cases they were analyzed individually and in other cases a merge of both datasets was analyzed.

The data is analyzed in Jupyter Notebook, using Pandas. So many different DataFrames formatting is used to achieve the results.

Analysis:

The comparison of data is between 15 unique schools with 39,170 students. The total budget for them is $24,649,428.00. Another considerable point is that in overall, students tend to pass the reading more than math, since the passing reading percentage is higher than the passing math percentage.

Results show that the highest performing schools are charter type with overall passing percentage of about 91.33. The highest performing school is Cabrera High School. This school has 1858 students and the total budget for them is $1,081,356.00.

On the other hand, the bottom performing Schools are district type with overall passing percentage of about 52.99. The bottom performing school is Rodriguez High School. This school has 3999 students and the total budget for them is $2,547,363.00.

According to the results, it can be concluded that the number of students in a school matters as the higher number of students result in a poor school performance. However, having a high budget does not increase the chance of a better performance.

Referring to the scores by school spending and scores by school size sections, the same conclusion can be achieved, that the higher number of students results in a poorer school performance, but a high budget does not result in a higher school performance.

Additionally, based on the results the school type charter has a much higher school performance compared to school type district.