Written Report:

For this project, I have been tasked with assisting the school board and mayor in making strategic decisions regarding future school budgets and priorities. Specifically, I am required to analyze district-wide standardized test results and provide insights into school performance. The data sets that I have at my disposal include all student information as well as school information.

To begin my analysis, I merged the data together based on the school name column that each student attended. The resulting data set includes student information, student math and reading scores, school type, school name, and budget information for each school. I then calculated the count of schools, the number of students, and the total budget, which I stored in a variable called "total\_budget."

During this calculation, I noticed that there were duplicated values in both the budget and school name columns. To address this issue, I removed the duplicates in the total\_budget variable, so that they were only added to the total once.

Next, I calculated the average math and reading scores, as well as the overall passing rate for both subjects, for all districts. I created a summary table called "district\_summary" to display these calculations.

In the next step, I calculated how many students passed the math and reading tests per school, as well as the per-school budget and capita. To summarize all of these calculations, I created a summary table called "per\_school\_summary," which contains the same information as the district\_summary table, but also includes the school type.

Using these two tables, we can see the total number of schools, students, and budget, as well as the average math and reading scores and the percentage of students passing each subject. Additionally, we can see the overall passing percentage for each school. And that allowed me to compare each row of the per\_school\_summary to the district summary.

it seems that there may be a correlation between the per student budget and the overall performance of schools. Specifically, schools with lower per student budgets may have lower overall passing rates. This is an interesting finding and may be worth exploring further by creating visualizations to better understand the distribution of per student budget across schools and the relationship between this two.

Furthermore, I created two tables to identify the highest and lowest performing schools based on their overall passing percentage. This analysis provided some confirmation of my initial hypothesis that there may be a relationship between per student budget and overall school performance.

To further investigate this relationship, I created a summary table that grouped schools by their per student spending. To accomplish this, I created spending\_bins and labels to categorize schools into different spending ranges. This summary table allowed me to better understand how school performance varied across different levels of per student spending, and provided additional evidence to support my hypothesis.