**PyCity School Analysis**

Two sets of data are provided:

1. School Data, information on schools in details including type, budgets, size etc.
2. Student Data, information on students in details including their grades and math/reading scores etc.

Initial step was to merge the data based on school names to get wholistic view of the data set. Based on the data provided, I was able to gather “District Summary” data frame to look at the data overall. Based on this table, it was confirmed that there was total of 15 schools, around 39,000 students with budget of $24.6M. Out of these students, the average Math Score is 79, and average Reading Score is 82. Their % passing is 75 in Math, 86 in Reading and 65 overall.

Utilizing the schools’ name, “School Summary” table is created based on all the information that was provided. School Type, Total Students and Total Budget, Per Student Budget and as well as the Math, Reading and Overall scores are calculated and placed in the table. By looking at this table, we were able to present the Highest 5 and Bottom 5 Performing Schools based on their % Overall passing.

Further analysis is done based on Reading and Math scores by Grades. The data is sorted by grade level 9th, 10th, 11th, and 12th and Math/Reading Scores are grouped by school name. Then the two data set merged into a data frame to see the passing scores with all the grades based on school names.

Next analysis is done to see if there is an impact on spendings with Reading and Math scores. I binned the spending into buckets and added another column to label the school names based on the student per budget data. Table, Scores by School Spending, of Math, Reading and Overall scores is created based on budget spending per student.

Similar analysis done to see if the school size has an impact on Reading and Math scores. A bin list is created to label the size of schools and the column is added into the data frame to label the school sizes. Table, Score by School Type, of Math, Reading and Overall scores is created based on the school size.

**Conclusion:**

1. Comparing the school types, Chater type school has higher scores than District type school over all. Especially, the % Passing Math, Reading and Overall, we can see significant differences.
2. Comparing the school sizes, there are not difference btw Small (<1000) and Medium (1000-2000) size schools, but when it comes to much larger schools, we can see that average Math/Reading scores drops, especially % Overall Passing significantly lower against the other two variables.
3. One interesting thing we can conclude is the spendings per student. As the spendings are increasing, the % Overall Passing is gradually decreasing. Based on the Scores by School Spending table, higher spending on a student does not mean that they will have higher Math and Reading scores. Spending does not impact student’s academic success.