**Analysis Summary:**

For this exercise, we are simulating being a Chief Data Scientist for a city’s school district. We have been tasked with assisting the school board and mayor as they prepare to make strategic decisions regarding future school budgets and priorities. In order to do so, we’ve been given access to every student’s math and reading scores as well as various information for all schools within the district. Utilizing two different CSV files containing the above data in conjunction with Python and Pandas, a thorough analysis was conducted. Dataset/dataframe manipulation and visualization allowed us to present cohesive information and actionable insights for the key decision makers. Analysis includes visualizations with key metrics for a district summary, a per school summary, both highest & lowest performing schools, math & reading scores by grade, scores categorized by school spending (per student), scores categorized by school size, and scores categorized by school type (district vs. charter).

**Conclusions:**

* When taking a look specifically within our “Spending Ranges (Per Student)” binning analysis, the data seems to show an inverse relationship between the amount of per student funding/spending and respective math/reading scores and math/reading passing percentages. In other words, the <$585/student range exhibited the highest metrics (across average math score, average reading score, % passing math, % passing reading, % overall passing) while each category/bin thereafter shows a decrease. Initially, while this data shows on average that the more funding each student receives within the district, the lower their aggregate academic achievement scores, there are a number of other contributing factors. For example, the data also shows that schools with a higher overall budget and hence a higher per student budget, are also schools that tend to have the highest number of enrolled students (size). As such, while each school’s size contributes to their relative funding and associated scoring aggregates, factors associated with having a larger school population inherently could come as a detriment to scores. This is evidenced in the below point.
* While there is a very miniscule aggregate scoring difference between small (<1000) and medium (1000-2000) schools within the district, there is a considerable decrease in scoring aggregates when comparing small and medium sized schools with large (2000-5000) schools. There is an approximately six, two, twenty-four, fourteen, and thirty-two average score/percentage point decrease across average math score, average reading score, % passing math, % passing reading and % overall passing respectively.
* Overall, the data conveys that the “School Type” variable (categorization between charter and district) has the most significant impact on scoring aggregates for schools within the district. Not only are the top five schools within the district (according to our % overall passing metric) all categorized as charter schools, the bottom five schools (according to our % overall passing metric) within the district also happen to be categorized as district schools. This relationship is further evidenced by the final visualization within our analysis that shows the scoring aggregate discrepancy between charter and district schools. Charter schools are exhibiting roughly six points better regarding average math score, three points better in terms of average reading score, twenty-seven percentage points more students passing math, sixteen percentage points more students passing reading and thirty-seven percentage points more students passing both math and reading.
* Individual school reading and math scores, seem to have very little variance depending on a student’s grade. This small degree of variance conveys that there is less of an impact on reading/math scores due to grade and perhaps there exists more worthwhile areas to focus on for improvement.
* Although our analysis highlights key factors and their ability to influence aggregate academic achievement metrics, the actionable insights derived will have to be carefully “unpacked” and given proper time/resource to truly understand how budgeting/priority decisions will ultimately affect the city’s school district. From a data professional’s perspective, size, type, per student funding seem to represent the most influential variables. As such, perhaps rezoning/building more schools within the city’s district may promote more favorable outcomes. This could allow each student to garner more resource availability, a lower student to teacher ratio, etc. When it comes to the type of school, key decision makers should pay special attention to the results dichotomy between charter & district schools and determine if the community might be better served with a model that emphasizes one over the other. Focusing on each school’s characteristics (variables within our analysis) could prove to be the main “pulls and levers” that better the individual school and in turn the district’s metrics a whole.