Introduction

This panda’s challenge creates a scenario whereby a new Chief Data Scientist for a city’s school district has been tasked with making a strategic decision regarding future school budgets and priorities.

Information on students’ math and reading scores, as well as the schools they attended, was made available. The focus of this analysis is to aggregate the data to show case clear trends in school performance. Two data sets were used in this analysis, schools\_complete data set and students\_complete data. The two data sets were merged for the analysis. Pandas function pd. merge () was used to merge the two data sets.

Analysis procedure

I started this challenge by importing the two data into python. All command prompts were entered on jupyter notebook in vs code. Pandas were imported as pd, from pathlib I imported Path. These libraries were called to read the data set.

DataFrame was created by using pd. read\_csv() function, the two data sets were merged using pd.merge() function. Various functions and commands were used to determine the data type, perform statistical analysis, filter, slice and sort the data, and retrieve data from specific column of the DataFrame.

Report On Summary Results

summary statistics for combine data set.

A screenshot of a computer screen

Description automatically generated

It could be inferred from the above table that the average students reading scores were relatively better than the math score, the budget of a school depends on the school size. Schools with bigger sizes have higher budgets.

District Summary Report

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Total Schools | Total students | Total Budget | Average math score | Average reading score | % passing math | % passing reading | % overall passing |
| 15 | 39170 | 24649428 | 78.98 | 81.87 | 74.98 | 85.80 | 65.17 |

The percentage of students passing math is lower than the district average, whereas the percentage of students passing reading is higher than the district average. We can conclude that overall students tend to score higher or perform better in reading.

School Summary

A screenshot of a computer

Description automatically generated

School summary report shows that district schools have higher populations compared to charter schools. Because of the higher enrollment in district schools, they tend to have higher budgets. Although budgets for charter schools are lower than district schools the overall percentage passing for charter schools is higher than that of districts schools.

We can conclude that students in charter schools perform better than their counterparts in district schools.

All 5 top or higher performing schools are charter schools whereas the bottom five or lowest performing schools are district schools.

Size Summary

A screenshot of a black screen

Description automatically generated

Students in small and medium size schools perform better than their counterparts in large schools.

This conclusion could be inferred from the table above.

School Type Summary

A screenshot of a computer

Description automatically generated

Students in charter schools perform better than their counterparts in district schools. This assertion could be inferred from the above table.

Spending Summary

A screenshot of a computer

Description automatically generated

The overall students performance is negatively correlated with spending. This is because charter schools are relatively smaller than districts schools, hence, the small spending.

Two Correct Conclusion

* Performance of students in charter schools is relatively better than their counterparts in districts schools.
* School spending is dependent on school size but does not correlate positive with performance.