

Written Report Pandas Challenge

Summarizes the analysis:

In my Py-City School Analysis, I first started off with setting up my dependencies, loading my files, reading it into Pandas data frame and finally merging both my student and school files. With that I got 39170 rows x 11 columns of data that I can utilize for the further requirements regarding module 4 challenge. The first data frame called district_summary included calculation for the total amount of schools (15), there altogether budget, amount of students, average math & reading scores, the percentage of students that passed math and reading, and average of overall passing/overall % passing. This helped create a larger but denser data frame called per_school_summary to showcase each high school's budget, budget per student, average math & reading scores, average % of passing math & reading and overall passing rates. Through this data frame I could then Sort the schools by % overall passing in descending order and display the top 5 rows. Then, to showcase lowest 5 rows Sort the schools by % overall passing in ascending order and display the bottom 5 rows. And saved the results in a Data Frame called "bottom schools". For the data frame to show math scores by grade I performed the necessary calculations to create a Data Frame that lists the average math score for students of each grade level (9th, 10th, 11th, 12th) at each school. And the same for data frame of reading scores by grade except I did the necessary calculations to show average reading scores for students for each grade. Lastly, for scores for school spending, school size and school type I created tables that broke down performance based on average spending ranges (per student), school size (small, medium or large) and school type (charter or district). This summary can be seen in the .pynb file showcasing all data frames and there calculations.

Draws two correct conclusions or comparisons from the calculations:

For District Summary: it can be concluded that out of 15 school (both charter and district) only 65.17% of students passed math & reading.

For School Summary: it can be concluded that Cabrera High School had the highest overall passing rate while Rodriguez High School had the lowest overall passing rate of students.

For Scores by School Type: it can be concluded that charter schools beat district schools in all metrics. Charter schools have higher averages in math and reading scores therefore they have higher passing rates of students as well.

For Scores by School Size: it can be concluded that the bigger the school the poorer the scores and passing rates of students as a whole. If a school is small or medium sized it will more likely have higher passing scores and averages.

For Scores by Spending: it can be concluded that schools with higher budgets do not yield better results. School with higher spending per student underperformed compared to schools with smaller budgets.