Wash U Data Boot Camp
Ben Rosensweig
Module 4 Homework
May 2024

Written Report

Analysis Summary

We created nine data frames out of two data files that summarized several different metrics among fifteen schools in a hypothetical school district. Among others, these metrics included school types, budgets, average scores, and passing rates. We used a number of different methods within pandas on the Jupyter Notebook to achieve this analysis. Besides the usual setup of importing, establishing a file path, and reading, we merged two sets of data into a single data frame. We utilized several different methods such as .unique(), .mean(), and conditionals. We created a new data frame using .set_index() based on each of the schools. We also used the method .groupby() to calculate average test scores. And while not particularly useful in pandas, we utilized the .map function for formatting. Of particular note, we created two sets of bins with the pd.cut() method to see test scores by school spending and school size.

Conclusions

There are two conclusions we can draw from this data. One of the most obvious and simple seems to be the difference between the two school types, Charter and District. The top five highest performing schools are all charter schools, whereas the lowest are all district schools.

The other conclusion is that there does not seem to be a strong correlation between the spending per student and test scores and passing rates. For instance, in the highest per capita range of $645 to $680 per student, the overall passing rate was 53.53%. Whereas for the lowest spending range, less than $585 per student, the overall passing rate was 90.37%. At first this didn't seem to make sense to me, and I wondered whether the generative data didn't match real world data. One would think that the more money a school spent per student, the higher the passing rates would be. However, according to Alice H. Johnson et al in the School Choice Demonstration Project from the University of Arkansas, "In the nine cities studied by the researchers, public schools received an average of $29,168 per pupil in FY 2020. Charter schools received around 70 percent of that at only $20,230 per student. But despite this gap, students in charter schools performed several points higher on a national standardized test."[1]

[1] Johnson, Alison H., et al. "Still a Good Investment: Charter School Productivity in Nine Cities." School Choice Demonstration Project, Department of Education Reform, University of Arkansas, November 2023.