Joseph Moravitz

GT Data Analytics Boot Camp

Homework 4 – Pandas

For this homework we were given a choice of two assignments. I had opted to complete the assignment having to do with analyzing data about test scores in a school district. In the assignment we were tasked with reading in data from csv files and manipulating it using pandas dataframes to try to glean some information about trends about test scores based on various factors.

Without having looked at the data, one might have expected that the more money a school spends per student, the more equipment, resources, higher quality teachers and such a student attending that school would have access to. It would seem logical to expect that students in those situations would perform better on average. Surprisingly, this was not what the data suggests. In fact, what we actually see from the data is the exact opposite, that schools where spending is less actually had higher average test scores.

Looking into this a bit more, if we were to have sorted the schools by the per-student budget we find that the schools with smaller per-student budget ratios were in fact all charter schools and district schools consistently have higher per-student budgets. This suggests to me that our initial expectation that given all other factors being equal, a higher amount of spending per student should still likely increase performance. Our observations were skewed by not considering the difference in school type in our earlier attempt.

Indeed, when looking at the top performing schools and the bottom performing schools, the top performing schools are dominated by charter schools while the bottom performing schools are dominated by district schools. This begs the question, “What is it about charter schools that make them so successful?” It certainly isn’t that they have a larger budget available. I expect that this may be a matter of values. The types of families who would send their children to charter schools are often more affluent and such families and children would have school performance as a higher priority than those who would send their children to district schools. Not only is school performance a higher value on the family side of things, but it is certainly an important value from the school side as charter schools which underperform are often closed.

Charter schools also consistently have a smaller student population which may further explain differences in performance. Schools with very large populations may in turn have higher student-teacher ratios and less one-on-one time spent with students.

All being told, our data set is simply too small to be able to make particularly strong observations beyond the observation that charter schools tend to outperform district schools within the area we got our data from. I would be interested to look at data from several more areas to try to see what the trends are within district schools alone to see how funding and school sizes affect test score averages without charter schools getting in the way.