PyCitySchools – Observations

My observations are going to be based on the limited knowledge that I have of the actual public school system. While I understand this to be data that is illustrative of how to use the Pandas library, given that it’s required that we provide a statement of at least two observable trends, I will speak on the data as if it were actual data from a real school district. With that said here is what I can see based on the data.

The first observable trend is that on average the charter schools tend to have higher scores in both math and reading. Sometimes there are big disparities between student scores in district schools and charter schools. This success disparity is visible in the part where we segment the top and bottom performing schools. For example, Cabrera High has a pretty remarkable 97% passing in reading, 94% passing in math whereas Rodriguez High School did not do so well with 80% in reading and a dismal 66% passing rate in math. The differences are probably present for all the reasons a person would expect like district schools having students from a lower socioeconomic class, higher student to teacher ratio, budget that is completely dependent on state and federal funding, etc. Whereas the charter schools may charge tuition, or they may require a certain level of academic success to qualify for admission, finally and this looks to be backed up by the data, a lower student to teacher ratio.

This brings up another observable trend based on the data. That observation is that the larger district schools tend to have lower levels of academic success. That can be seen in the ‘scores by size’ section of the notebook. The smaller the school the higher the scores, this is especially apparent in the percentage of students passing math where small schools of less than 1000 students are passing at a rate of 93% whereas the largest schools are sitting at a much lower 69%. At the same time, when the budget per student numbers are factored in, the trend is that the district schools tend to have a higher budget per student. Something interesting to note is just how much bigger the district schools are in comparison to the charter schools. The charter schools having and average enrollment of a pretty modest 1524 students whereas the district schools are much larger with an average of 3853 students enrolled. So, between the higher cost per student, with more students, and likely the special needs and at-risk students at the district schools, it seems that the cumulative effect seems to be a lower degree of academic success.

Based on the two observations laid up in above, I believe that there is a glaring limitation of this data set. That limitation is that it paints, what I would consider to be likely in accurate picture of the school system in PyCity. At first glance, one could conclude that charter schools are superior schools to district schools given the better academic numbers. Likely, that is not the case. As a supporter of public schools, I realize there may be some bias in this analysis, but I believe that the story this data tells is, at best incomplete. While this data set does show raw numbers like, enrollment at various schools, averages and percentages of student achievement, and budgets, this data leaves out quite a bit of factors that could play into why these numbers are the way they are. For example, how much of the budget goes to payroll, maintenance, and other expenses related to running a school? Something else that should be considered is, how does funding work in the state the PyCity is in? Are charter school students and district students funded at the same level? As mentioned before about student to teacher ratio, this data mentions nothing about how many teachers are employed at each school, that could be important. What about special needs students, at-risk students, TAG students? These are groups of students that could skew the results. Ultimately, the main takeaway that I took from this data is that a smaller school will usually yield better test scores.