PyCitySchools Analysis

By Lauren Cadogan

In PyCitySchools, two datasets were merged to compare math and reading scores across a school district. The district\_summary dataset shows the total school count, student count, budget and average grades. The groupby method was used to create the per\_school\_summary dataset which shows the same breakdown [as the district\_summary] but by school. Other methods used when analyzing the data were ‘mean’, ‘merge’, ‘unique’, ‘count’ and ‘sum’, among others.

The data shows the top 5 performing schools were all charter schools. A surprising finding was found in the spending\_summary\_dataframe. The highest overall passing percentage was by those that spent the least per student and worse scores were earned by the schools that spend the most per student. In the size\_summary dataframe, the was not a major difference in scores between small (<1000 students) and medium (1000-2000 students) sized schools but schools with 2000-5000 students had grades that were negatively impacted.