The purpose of this assignment was to perform exploratory data analysis using Pandas on data collected from 15 high schools, some charter and some district, or public schools. The data consisted of all students by name from each school, individual math and reading scores, the number of students at each school, and the budget of each school. Pandas provided a means of combining data from more than one source to create one merged dataframe, calculating aggregate statistics, and filtering the data based on various characteristics using indexing. By looking at the mean scores in math and reading among all the scores, we see that the overall average reading score is higher than the overall average math score. One may be interested in which factors are associated, or maybe, influence the difference in scores for math and reading.

If we filter the data based on whether the school is a charter school versus whether the school is a district school, it is apparent that charter schools in this dataset have higher marks in average math score, average reading score, percentage of students passing math, percentage of students passing reading, and percentage of students passing. The possible reasons that account for this discrepancy in scores between charter scores could be funding, socioeconomic status related factors, such as the need for students to hold part-time jobs to earn money, lack of emphasis from parents on education, tutoring paid for by parents, or other reasons.

However, further inspection of the data shows a lack of support for the idea that better funded schools have more students with better math and reading grades. It does not seem that the budget per student at charter schools is overwhelmingly greater than the budget per student at district schools, based on the appearance of the dataframe below.

school_name	School Type	Total Students	Total School Budget	Per Student Budget	Average Math Score	Average Reading Score	% Passing Math	% Passing Reading	% Overall Passing	Spending Ranges (Per Student)
Bailey High School	District	4976	$3,124,928.00	628.0	77.048432	81.033963	66.680064	81.933280	54.642283	$585-630
Cabrera High School	Charter	1858	$1,081,356.00	582.0	83.061895	83.975780	94.133477	97.039828	91.334769	<$585
Figueroa High School	District	2949	$1,884,411.00	639.0	76.711767	81.158020	65.988471	80.739234	53.204476	$630-645
Ford High School	District	2739	$1,763,916.00	644.0	77.102592	80.746258	68.309602	79.299014	54.289887	$630-645
Griffin High School	Charter	1468	$917,500.00	625.0	83.351499	83.816757	93.392371	97.138965	90.599455	$585-630
Hernandez High School	District	4635	$3,022,020.00	652.0	77.289752	80.934412	66.752967	80.862999	53.527508	$645-680
Holden High School	Charter	427	$248,087.00	581.0	83.803279	83.814988	92.505855	96.252927	89.227166	<$585
Huang High School	District	2917	$1,910,635.00	655.0	76.629414	81.182722	65.683922	81.316421	53.513884	$645-680
Johnson High School	District	4761	$3,094,650.00	650.0	77.072464	80.966394	66.057551	81.222432	53.539172	$645-680
Pena High School	Charter	962	$585,858.00	609.0	83.839917	84.044699	94.594595	95.945946	90.540541	$585-630
Rodriguez High School	District	3999	$2,547,363.00	637.0	76.842711	80.744686	66.366592	80.220055	52.988247	$630-645
Shelton High School	Charter	1761	$1,056,600.00	600.0	83.359455	83.725724	93.867121	95.854628	89.892107	$585-630
Thomas High School	Charter	1635	$1,043,130.00	638.0	83.418349	83.848930	93.272171	97.308869	90.948012	$630-645
Wilson High School	Charter	2283	$1,319,574.00	578.0	83.274201	83.989488	93.867718	96.539641	90.582567	<$585
Wright High School	Charter	1800	$1,049,400.00	583.0	83.682222	83.955000	93.333333	96.611111	90.333333	<$585

In addition, filtering the data based on budget range shows that the average marks and percentage of students passing for higher budget buckets are lower than those for lower budget buckets. Despite this seemingly negative correlation between school budget and student scores or student pass rates, one must be mindful not necessarily to attribute lower student scores or pass rates to a higher school budget. Perhaps, despite some schools having higher budgets per student, there are lower teachers to student ratios, so the schools with seemingly higher budgets have less academic attention given to each student. Or perhaps, there are other factors that are more powerful than school budget that influence students' academic performance, such as the ones mentioned above.

Spending Ranges (Per Student)	Average Math Score	Average Reading Score	% Passing Math	% Passing Reading	% Overall Passing
<$585	83.455399	83.933814	93.460096	96.610877	90.369459
$585-630	81.899826	83.155286	87.133538	92.718205	81.418596
$630-645	78.518855	81.624473	73.484209	84.391793	62.857656
$645-680	76.997210	81.027843	66.164813	81.133951	53.526855

Through this data, one can obtain a glimpse into the math and reading scores of a set of charter and district/public schools. From this data, it is apparent that charter schools, on average have higher math scores, reading scores, percentage passing math, percentage passing reading, and percentage passing both. The data also suggests that budget is not a factor influencing scores and pass rates. Whereas one could speculate endless reasons for these differences in academic performance, one reason slightly supported by the data is school size, and potentially consequently teachers to student ratios. One rudimentary solution is to allocate more budget funds from other areas to hiring more instructors to raise the teachers to student ratios. Another preliminary solution is to have regular 1-on-1 teacher-student meetings throughout the school year. For other potential causes, investigation into other data is needed, and experimentation with the factors underlying those reasons would be pure guesswork.