In this project I was asked to analyze district-wide standardized test results. I had access to the student's math and reading scores, as well as various demographic details about the schools those students attend.

I broke the district's data down into key metrics in a DataFrame to see high-level snapshots of the district's schools including:
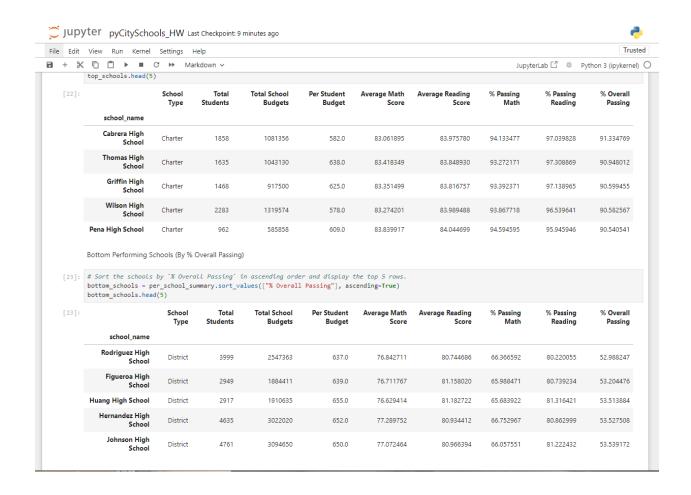
- Total number of unique schools

- Total students

- Total budget

- Average math score

- Average reading score

- % passing math (the percentage of students who passed math)

- % passing reading (the percentage of students who passed reading)

- % overall passing (the percentage of students who passed math AND reading)

I similarly did calculations and created a DataFrame with key metrics about each school, including:

- School name

- School type

- Total students

- Total school budget

- Per student budget

- Average math score

- Average reading score

- % passing math (the percentage of students who passed math)

- % passing reading (the percentage of students who passed reading)

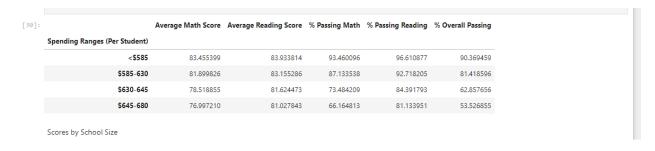- % overall passing (the percentage of students who passed math AND reading)

I then built more detailed DataFrames to find trends and analyze the data to find observable trends.

Such as:

```
top_schools.head(5)
```

[22]:

	School Type	Total Students	Total School Budgets	Per Student Budget	Average Math Score	Average Reading Score	% Passing Math	% Passing Reading	% Overall Passing
school_name									
Cabrera High School	Charter	1858	1081356	582.0	83.061895	83.975780	94.133477	97.039828	91.334769
Thomas High School	Charter	1635	1043130	638.0	83.418349	83.848930	93.272171	97.308869	90.948012
Griffin High School	Charter	1468	917500	625.0	83.351499	83.816757	93.392371	97.138965	90.599455
Wilson High School	Charter	2283	1319574	578.0	83.274201	83.989488	93.867718	96.539641	90.582567
Pena High School	Charter	962	585858	609.0	83.839917	84.044699	94.594595	95.945946	90.540541

Bottom Performing Schools (By % Overall Passing)

[23]:
```
# Sort the schools by `% Overall Passing` in ascending order and display the top 5 rows.
bottom_schools = per_school_summary.sort_values(["% Overall Passing"], ascending=True)
bottom_schools.head(5)
```

[23]:

	School Type	Total Students	Total School Budgets	Per Student Budget	Average Math Score	Average Reading Score	% Passing Math	% Passing Reading	% Overall Passing
school_name									
Rodriguez High School	District	3999	2547363	637.0	76.842711	80.744686	66.366592	80.220055	52.988247
Figueroa High School	District	2949	1884411	639.0	76.711767	81.158020	65.988471	80.739234	53.204476
Huang High School	District	2917	1910635	655.0	76.629414	81.182722	65.683922	81.316421	53.513884
Hernandez High School	District	4635	3022020	652.0	77.289752	80.934412	66.752967	80.862999	53.527508
Johnson High School	District	4761	3094650	650.0	77.072464	80.966394	66.057551	81.222432	53.539172

Where we can see that the top five performing schools in the district were all Charter Schools, while the lowest performing schools were all District schools.

We can also see:

[30]:

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

Scores by School Size

That spending more money per student did not equal higher test results. In fact, we see that the lowest spending schools, had the highest test results.