Module 4 Challenge – PyCitySchools

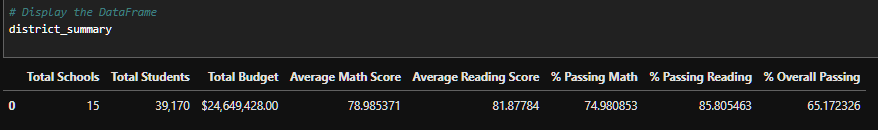
William Davis | [william.c.davis@nasa.gov](mailto:william.c.davis@nasa.gov)

----

**District Summary (20 points)**

* Calculate the total number of unique schools (2 points)
* Calculate the total number of students (2 points)
* Calculate the total budget (2 points)
* Calculate the average (mean) math score (2 points)
* Calculate the average (mean) reading score (2 points)
* Use the code provided to calculate the percentage of students who passed math (2 points)
* Calculate the percentage of students who passed reading (2 points)
* Use the code provided to calculate the percentage of students that passed both math and reading (2 points)
* Create a new DataFrame for the above calculations called district\_summary (4 points)

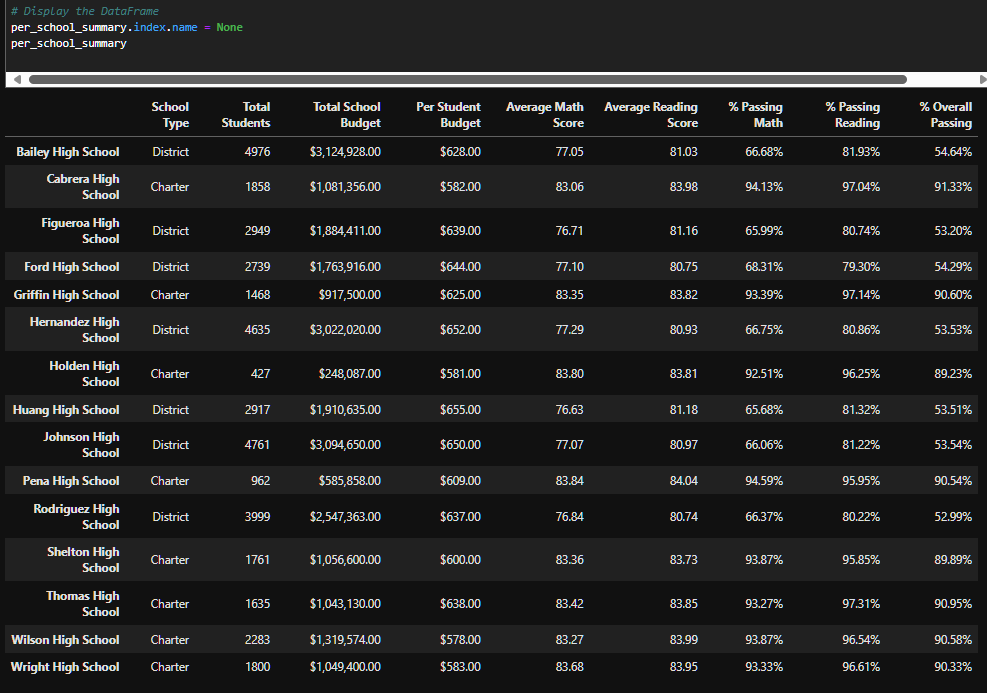
See attached file: PyCitySchools.ipynb



**School Summary (20 points)**

* Use the code provided to select the school type (2 points)
* Calculate the total student count (2 points)
* Use the code provided to calculate the per capita spending (2 points)
* Calculate the average test scores (2 points)
* Calculate the number of schools with math scores of 70 or higher (2 points)
* Calculate the number of schools with reading scores of 70 or higher (2 points)
* Use the provided code to calculate the schools that passed both math and reading with scores of 70 or higher (2 points)
* Use the provided code to calculate the passing rates (2 points)
* Create a new DataFrame for the above calculations called per\_school\_summary (4 points)

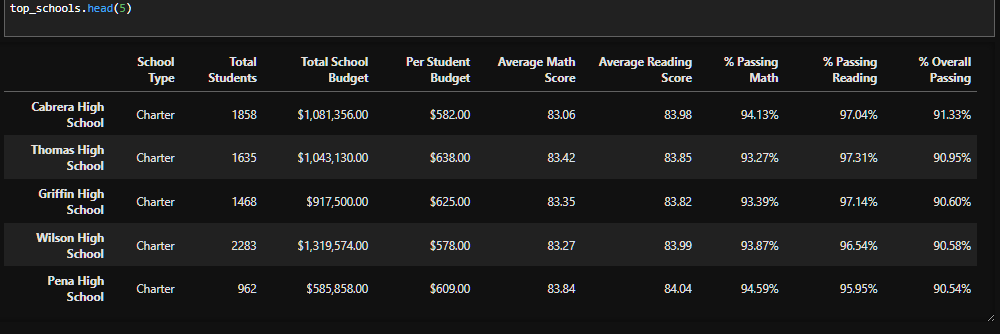
See attached file: PyCitySchools.ipynb



**Highest-Performing Schools by Percentage of Overall Passing (5 points)**

* Sort the schools by % Overall Passing in descending order (2 points)
* Save the results to a DataFrame called top\_schools (2 points)
* Display the first 5 rows (1 point)

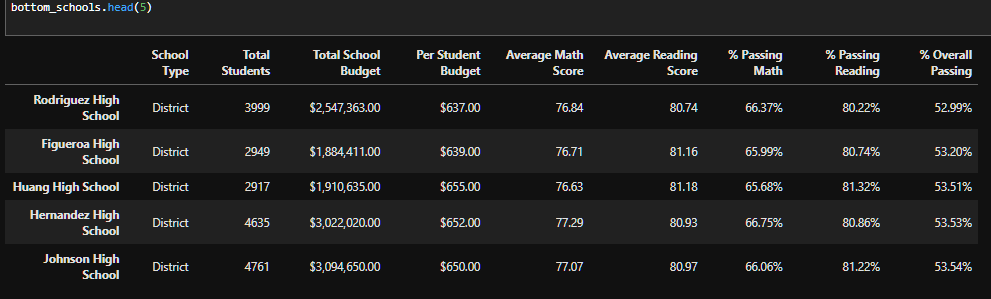
See attached file: PyCitySchools.ipynb



**Lowest-Performing Schools by Percentage of Overall Passing (5 points)**

* Sort the schools by % Overall Passing in ascending order (2 points)
* Save the results to a DataFrame called bottom\_schools (2 points)
* Display the first 5 rows (1 point)

See attached file: PyCitySchools.ipynb



**Math Scores by Grade (10 points)**

* Use the code provided to separate the data by grade (1 points)
* Group by "school\_name" and take the mean of each (4 points)
* Use the code to select only the math\_score (1 points)
* Combine each of the scores above into single DataFrame called math\_scores\_by\_grade (4 points)

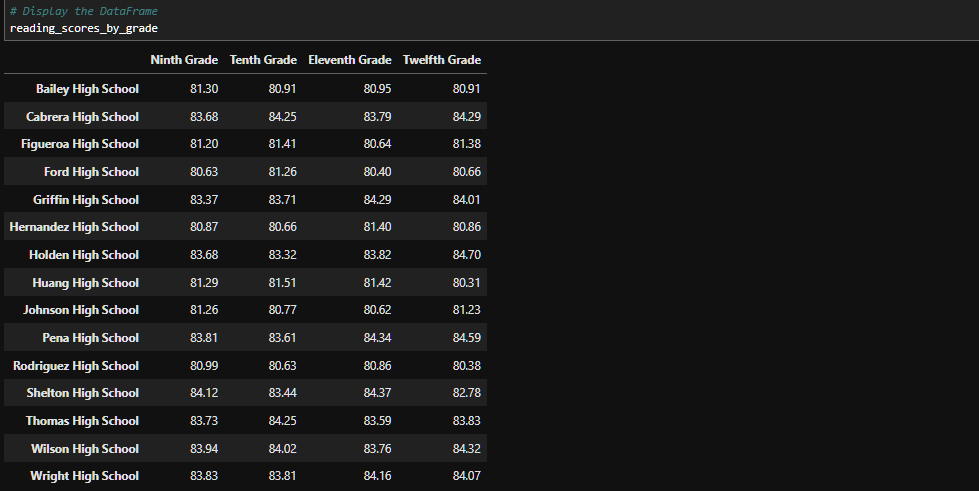
See attached file: PyCitySchools.ipynb



**Reading Scores by Grade (10 points)**

* Use the code provided to separate the data by grade (1 points)
* Group by "school\_name" and take the mean of each (4 points)
* Use the code to select only the reading\_score (1 points)
* Combine each of the scores above into single DataFrame called reading\_scores\_by\_grade (4 points)

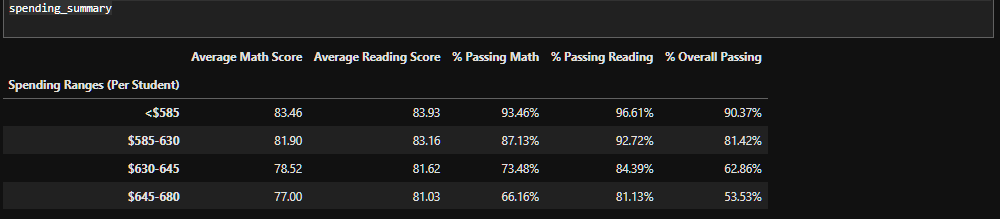
See attached file: PyCitySchools.ipynb



**Scores by School Spending (5 points)**

* Use pd.cut with the provided code to bin the data by the spending ranges (2 points)
* Use the code provided to calculate the averages (1 points)
* Create the spending\_summary DataFrame using the binned and averaged spending data (2 points)

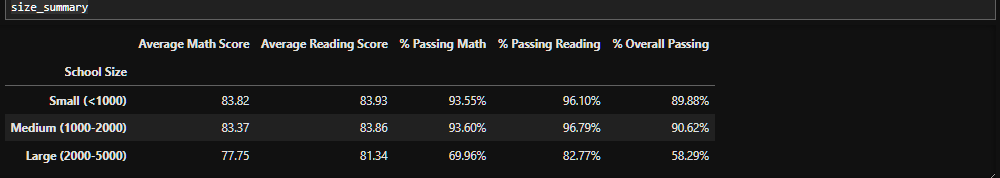
See attached file: PyCitySchools.ipynb



**Scores by School Size (5 points)**

* Use pd.cut with the provided code to bin the data by the school sizes (2 points)
* Use the code provided to calculate the averages (1 points)
* Create the size\_summary DataFrame using the binned and averaged size data (2 points)

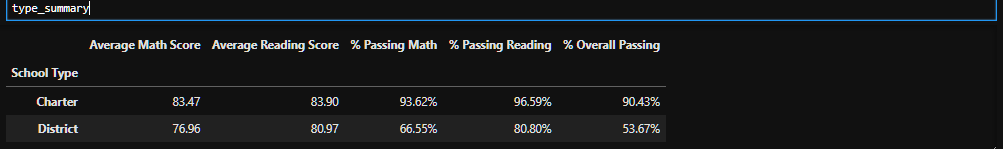
See attached file: PyCitySchools.ipynb



**Scores by School Type (5 points)**

* Group the per\_school\_summary DataFrame by "School Type" and average the results (2 points)
* Use the code provided to select the new column data (1 point)
* Create a new DataFrame called type\_summary that uses the new column data (2 points)

See attached file: PyCitySchools.ipynb



**Written Report (15 points)**

To receive all points, the written report presents a cohesive written analysis that:

* Summarizes the analysis (5 points)
* Draws two correct conclusions or comparisons from the calculations (10 points)

Within the PyCitySchools District there are a total of 15 High Schools serving a combined enrollment of 39,170 students. The operating cost of these 15 schools is approximately $24.65M. Standardized test scores within the district show approximately 3 in 4 students pass math (75.0%) and 5 in 6 students pass reading (85.8%); only 2 out of 3 students pass both math and reading (65.2%).

The PyCitySchools District includes 7 District schools and 8 Charter Schools. The District schools tend to have significantly larger enrollments, with the smallest being Ford High School with 2739 students and the largest being Bailey High School with 4976 students. This is in stark contrast to the Charter schools where the largest campus is Wilson High School with 2283 students and the smallest campus is Holden High School with just 427 students. At all schools, the average math score fluctuates between 76 – 84 whereas the reading score fluctuates within a smaller range of 80-84.

The five top performing schools are all Charter schools with an overall passing rate of 90.5% or better for all enrolled students. The five bottom performing schools are all District schools with an overall passing rate of 53.5% or lower for all enrolled students.

Math scores at each high school are generally flat across all 4 grades and do not show improvement or decline. Reading scores at each high school show the same trend and are also generally flat across all 4 grades and do not show any marked improvement or decline.

In comparing the amount of money spent per student to test score performance, there appears to be an inverse relationship to testing performance outcomes. The more that is spent, the fewer students who pass math, reading, or overall. Therefore, increasing spending does not correlate to improved test scores. Similarly, comparing the size of each school to test score performance, there again appears to be an inverse relationship. The larger the school, the fewer number of students who pass math, reading, or overall. The best test scores and highest pass rates occur at the smaller schools, which in PyCitySchools District tends to be Charter schools.