**City School Math & Reading Test Results Analysis:**

**In Python using Jupyter Notebook**

**Overview of City School Test Result Project:**

After completing an initial analysis of the district high school standardized tests score for Maria to report to the school board. It was determined that there were questions as to the validity of the ninth-grade scores for Thomas High School. Maria asked us to help her by replacing the scores in question with “NaN” values and then completed the analysis again to get updated results. She also asked us to do a brief write-up of explaining how these changes to the data effected the final analysis results.

**Process of Data Modifications:**

**Remove Thomas High Schools 9th grade test scores:**

Using the “loc” function to find the ninth grade Math and Reading test scores for Thomas High Schools and used the Nympy replace function to change them to Nan:

***Code:*** *student\_data\_df.loc[(student\_data\_df["grade"]== "9th") & (student\_data\_df["school\_name"]=="Thomas High School") , "reading\_score"] = np.nan*

Next, we used the same “loc” function to “count” the number of ninth grade students at Thomas High school (which came to 461).

***Code:*** *ThomasHS\_grade\_nine = student\_data\_df.loc[(student\_data\_df["grade"]== "9th")*

*&(student\_data\_df["school\_name"]=="Thomas High School")].count()['Student ID']*

With this information we recalculated the district averages after reducing the “total student count” by the 461 Thomas high school ninth graders. This provided us with the new district averages. Since 461 is just over 1% of the total number of students there was minor impact to the overall averages.

To get the school details we had to recalculate the averages and percentage for Thomas High School excluding the ninth-grade scores that were now Nan values. We again used the “loc” and “count” the passing grades for both Math and Reading for each grade (10-12) individually. Using code like this:

***Code:*** *ThomasHS\_passing\_math\_df = student\_data\_df.loc[(student\_data\_df["math\_score"]>=70) &(student\_data\_df["school\_name"]=="Thomas High School") &*

*((student\_data\_df["grade"]== "10th")|*

*(student\_data\_df["grade"]== "11th")|*

*(student\_data\_df["grade"]== "12th"))].count()['Student ID'*]

Once we had the count of passing students for each grade, we recalculated the average for passing students using the new student counts that excluded the Thomas High School nineth-graders to get the updated percentages.

Once again using the “loc” function the percentages “replaced” the original percentages.

***Code:*** *per\_school\_summary\_df.loc[["Thomas High School"],["% Passing Math"]] = ThomasHS\_passing\_math\_percent*

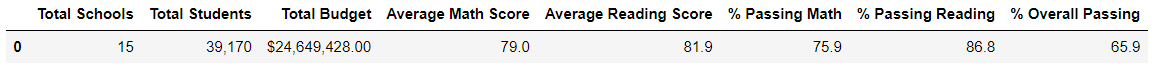
Since the number of ninth grades at Thomas High School was almost 40% of their total student count it did have significant impact on that schools’ results. See result below:

**Results and Analysis of After Data Modifications:**

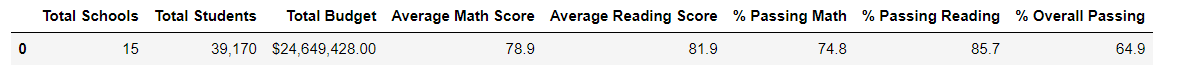
After making the modification to remove the ninth-grade scores and complete the analysis again with the new data set. It had the following results:

**Total Results: *(Only slight change to total percentage)***

(Original Analysis)



(Updated Analysis)

****

**Example of School Results: *(Modifications only impacted Thomas High School as at school level)***

(Original Analysis) **Graphical user interface, text, application

Description automatically generated**

(Updated Analysis)Graphical user interface, text, application, email

Description automatically generated

**School Average Test Scores by Grade: *(See changes to Thomas High School 9th Grade scores only)***

(Original Analysis)

**Table

Description automatically generatedTable

Description automatically generated**

(Updated Analysis)

Table

Description automatically generated Table

Description automatically generated

**Overall Results by per Student Spend: *(No change to a tenth of %)***

(Original Analysis)

**Table

Description automatically generated with medium confidence**

(Updated Analysis)**Graphical user interface

Description automatically generated**

**Overall Results by size: *(No change to a tenth of %)***

(Original Analysis)

**Graphical user interface

Description automatically generated**

(Updated Analysis)

**Table

Description automatically generated**

**Overall Results by Type: *(No change to a tenth of %)***

(Original Analysis)

**Table

Description automatically generated with low confidence**

(Updated Analysis)

**Graphical user interface, application

Description automatically generated**

**Summary**

In conclusion the modification to change all the Thomas High School nineth grade scores to “Nan” had only small effect on the results of the analysis. It did lower the total average scores and percentages slightly. Average math score down a tenth of a point, and average reading less than a tenth of a point. The “Overall Passing Percentages” did go down, by about 1.5%. There was an even small impact to the Thomas High School results with their passing percentages only dropping by less than a percentage point.

A picture containing graphical user interface

Description automatically generated