# PyCitySchools Challenge

## Overview of Project

###Original Project

Maria, the chief data scientist for a city school district, has requested my assistance in preparing an analysis of school testing proficiency in the school district. I assisted Maria in analyzing data on student funding and students’ standardized test scores. The project used Anaconda and Jupyter Notebook software. It also used the Pandas library, a preferred tool for data analysis. I created and activated a development environment for the project. I also created and cloned a GitHub repository for the project. Maria walked me through a series of exercises working with the data, Anaconda, Jupyter Notebook, and the Pandas library. This included opening and inspecting csv files containing data on the school district; cleaning the data; creating and merging dataframes; filtering the data for specific analyses; and developing reports for the school board.

### Revision of the Project

The school board subsequently notified Maria and her supervisor that the student\_complete.csv file showed evidence of academic dishonesty, specifically, reading and math grades for Thomas High School ninth graders appeared to have been altered. Although the school board did not know the full extent of the academic dishonesty, they wanted to uphold state-testing standards and turned to Maria for help. Maria asked me to replace the math and reading scores for Thomas High School with NaNs while keeping the rest of the data intact. Once I replaced the math and reading scores, Maria requested that I repeat the school district analysis that I did earlier and describe how these changes affected the overall analysis. </p>

## Challenges

Throughout working on Maria’s series of data exercises and the challenge, I found the code did not compile, showing errors. Sometimes finding the errors was simple, such as just running the entire code from the beginning rather than just the cell. But more times than not, finding the error was very difficult, usually requiring review of existing code from a prior student, XGUILXR and his “PyCitySchools\_Challenge.ipynb” file. I found that most of my code that did not compile had minor errors such as a missing bracket or parentheses. However, in some cases, it was more, typically where I did not break the code down enough. The screenshot below is one example of this in which I attempted to do grades 10-12 of Thomas High School together in one line of code and the solution was breaking it down by grade, so multiple lines of code.

![THS\_Overall\_Passing\_Percentage\_code.png](https://github.com/Robertfnicholson/School\_District\_Analysis/blob/031e541bfdf64fb0d53cff4431ce491359cf6f29/THS\_Overall\_Passing\_Percentage\_code.png) </p>

## Results:

PyCitySchools Challenge provided the following key results:

* The district summary changed slightly with the removal of the Thomas High School 9th grader test results. The average math score for the district decreased slightly from 79.0% to 78.9%. Since the report rounded values, the summary does not indicate other changes. See the original and revised District Summary below.

![District\_summary\_original.png](<https://github.com/Robertfnicholson/School_District_Analysis/blob/031e541bfdf64fb0d53cff4431ce491359cf6f29/District_summary_original.png>)

![District\_summary\_revised.png](https://github.com/Robertfnicholson/School\_District\_Analysis/blob/031e541bfdf64fb0d53cff4431ce491359cf6f29/District\_summary\_revised.png)

* The school summary changed for Thomas High School, with the average math score decreasing from 83.418% to 83.351% rounding to the nearing thousandth. However, Thomas High School’s average reading scores increased slightly to 83.896% from 83.845%, again rounding to the nearing thousandth. See the original and revised Per School Summary below.

![Per\_school\_summary\_original.png](<https://github.com/Robertfnicholson/School_District_Analysis/blob/031e541bfdf64fb0d53cff4431ce491359cf6f29/Per_school_summary_original.png>)

![Per\_school\_summary\_revised.png](https://github.com/Robertfnicholson/School\_District\_Analysis/blob/031e541bfdf64fb0d53cff4431ce491359cf6f29/Per\_school\_summary\_revised.png)

* Replacing the ninth graders’ math and reading scores at Thomas High School did not impact Thomas High School’s performance relative to the other schools. The schools passing percentages remained unchanged when rounding the values to whole numbers.
* Replacing the Thomas High School ninth-grade scores had the following affect, after rounding:
  + No change in the scores by school spending.
  + No change in the scores by school size.
  + No change in the scores by school type.
  + This is understandable given that the change from the removal of a relatively small set of grades is limited when using an average measure and percentage across the district. </p>

## Key Reports

Maria and I generated the following key reports for the school board as part of the project:

* Per School Summary

![Per\_school\_summary\_revised.png](https://github.com/Robertfnicholson/School\_District\_Analysis/blob/031e541bfdf64fb0d53cff4431ce491359cf6f29/Per\_school\_summary\_revised.png)

* Size Summary

![Size\_summary\_revised.png(https://github.com/Robertfnicholson/School\_District\_Analysis/blob/031e541bfdf64fb0d53cff4431ce491359cf6f29/Size\_summary\_revised.png)

* Spending summary

![Spending\_summary\_revised.png](https://github.com/Robertfnicholson/School\_District\_Analysis/blob/031e541bfdf64fb0d53cff4431ce491359cf6f29/Spending\_summary\_revised.png)

* Type Summary

![Type\_summary\_revised.png(https://github.com/Robertfnicholson/School\_District\_Analysis/blob/031e541bfdf64fb0d53cff4431ce491359cf6f29/Type\_summary\_revised.png) </p>

## Summary

Four changes in the updated school district analysis after reading and math scores for the ninth grade at Thomas High School were replaced with NaNs were the following:

* The number of students with test scores in the district decreased 461 from 39,170 to 38,709.
* The school summary changed for Thomas High School, with average math score decreasing from 83.418% to 83.351% and average reading score increasing from 83.845% to 83.896%.
* The number of Thomas High School students with test scores decreased 461 from 1,635 to 1,174 after eliminating the 9th grader scores.
* The overall passing percentage for Thomas High School went from 90.948% to 90.630%, rounding to the nearest thousandth.</p>