PyCity School – Module 4 Challenge

Ryan James Submission

**City School Report**

In this report, as data analysts, we will scrutinise and investigate the relationships between two data sets within the local Government area, whereby, 15 schools with be analysed. Furthermore, with these two data sets, the information contained within two CSV datasets will be merged to form one uniform dataset for investigation. This merged data-frame was used to provide the following conclusions with regards to a school’s budget, number of students, and each students math and reading scores.

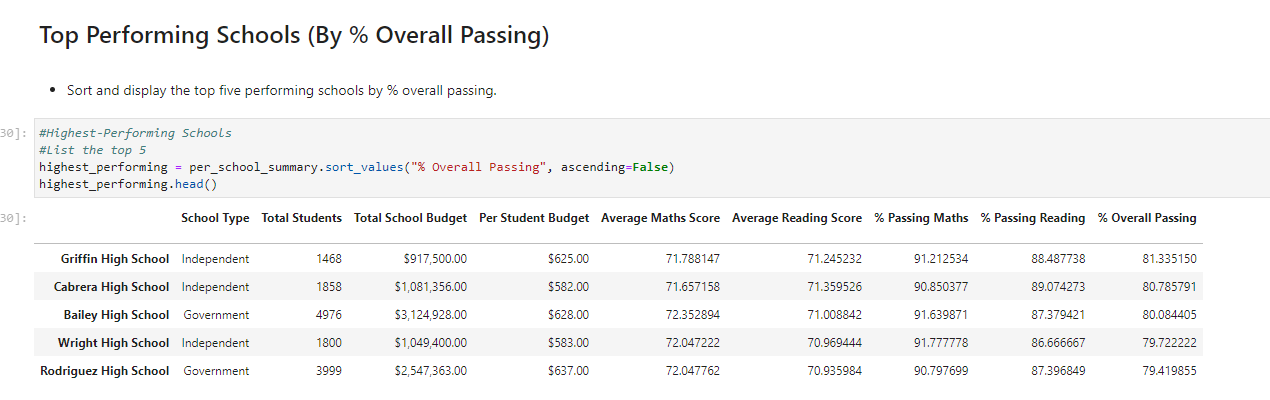
Throughout this report, both the Local Government Area (LGA) and School summaries were achieved – thus, the following conclusions were made when combining the two datasets:

* High-Performing Schools (top 5)
* Lowest-Performing Schools (bottom 5)
* Student Math Score for each individual year group
* Student Reading Score for each individual year group
* Scores by School Spending
* Scores by School Size
* Scores by School Type

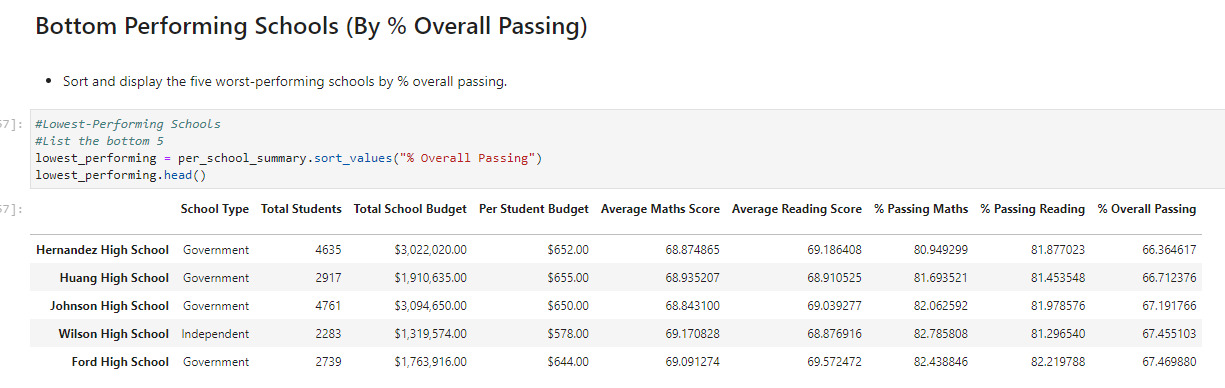
The effectiveness of this to draw more accurate conclusions to increase the validity of our findings. The sample size is small, however, due to high number data entries within a small sample size of schools, the validity of these results are accurate within this Government Area. However, these findings are localised and do not provide highly accurate assumptions – sample size is relatively small within the context of global assumptions.

**Findings Conclusions**

***Highest-Performing Schools (Overall Percentage who Passed)***

****

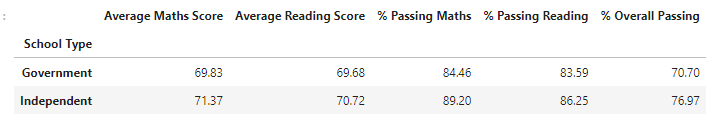
This dataset provide reflects the top 5 best performing schools, with regards to students who are passing math and reading. The pass rate across these schools is 50% - thus, these schools have high passing rates. Furthermore, all of these schools have signific budgets – these school have access to better resources, more skilled teachers, and provide students with more effective tools (technology & libraries) for students to achieve higher grades. On the contrary, when looking at the “***Lowest-Performing Schools (Overall Percentage who Passed):***

******

Despite these schools depicting a very close relationship to the “highest performing school” with regards to “Per Student Budget” and “Total School Budget”, these values do not correlate as to why students in these schools (Hernandez High School, Huang High School, Johnson High School, Wilson High School, Ford High School) are performing significantly lower than those school with similar budgets. These school may be “lower areas” of society, whereby, crime and less-fortunate families have their children attend these schools – thus, the high budgets for these schools, are incentives for children in these areas to build and develop an education.

**Lastly,**

In this Local Government Area (LGA), we can summaries that “independent” students outperformed “Government” students. The following dataset depicts this:



The overall passing rate for both maths and reading was 6.27% higher, which is statistically significant. This conclusion showcases independent students outperforming Government student. Furthermore, Independent schools attract a smaller subset of students (wealthy families, religious denomination, etc) – thus, these independent schools appear to have students who work harder.