

Module 4: A Written Analysis of the Results

Caitlyn Castillo

U of A Bootcamp

Overview

The team was tasked with analyzing school data, specifically student standardized test scores and student allocated funding. The team must aggregate all data to visualize trends in data to show correlations between test scores and grade, school type, school budget, and school size. These visualizations will aid the school board in making decisions regarding future school decisions.

This specific task was to clean data, by removing inaccurate 9th grader data and replacing it with a Not a Number (NaN) value. The team first focused on Thomas High School (THS).

Results

1. How is the district summary affected?

District data was not affected by the removal of one data set, as the District has a large amount of data. Below is the District data prior to the data clean up:

	School Type	Total Students	Total School Budget	Per Student Budget	Average Math Score	Average Reading Score	% Passing Math	% Passing Reading	% Overall Passing
Bailey High School	District	4976	\$3,124,928.00	\$628.00	77.048432	81.033963	66.680064	81.933280	54.642283
Cabrera High School	Charter	1858	\$1,081,356.00	\$582.00	83.061895	83.975780	94.133477	97.039828	91.334769
Figueroa High School	District	2949	\$1,884,411.00	\$639.00	76.711767	81.158020	65.988471	80.739234	53.204476
Ford High School	District	2739	\$1,763,916.00	\$644.00	77.102592	80.746258	68.309602	79.299014	54.289887
Griffin High School	Charter	1468	\$917,500.00	\$625.00	83.351499	83.816757	93.392371	97.138965	90.599455
Hernandez High School	District	4635	\$3,022,020.00	\$652.00	77.289752	80.934412	66.752967	80.862999	53.527508
Holden High School	Charter	427	\$248,087.00	\$581.00	83.803279	83.814988	92.505855	96.252927	89.227166
Huang High School	District	2917	\$1,910,635.00	\$655.00	76.629414	81.182722	65.683922	81.316421	53.513884
Johnson High School	District	4761	\$3,094,650.00	\$650.00	77.072464	80.966394	66.057551	81.222432	53.539172
Pena High School	Charter	962	\$585,858.00	\$609.00	83.839917	84.044699	94.594595	95.945946	90.540541
Rodriguez High School	District	3999	\$2,547,363.00	\$637.00	76.842711	80.744686	66.366592	80.220055	52.988247
Shelton High School	Charter	1761	\$1,056,600.00	\$600.00	83.359455	83.725724	93.867121	95.854628	89.892107
Thomas High School	Charter	1635	\$1,043,130.00	\$638.00	83.418349	83.848930	93.272171	97.308869	90.948012
Wilson High School	Charter	2283	\$1,319,574.00	\$578.00	83.274201	83.989488	93.867718	96.539641	90.582567
Wright High School	Charter	1800	\$1,049,400.00	\$583.00	83.682222	83.955000	93.333333	96.611111	90.333333

As well as the data after the data clean up:

	School Type	Total Students	Total School Budget	Per Student Budget	Average Math Score	Average Reading Score	% Passing Math	% Passing Reading	% Overall Passing
Bailey High School	District	4976	\$3,124,928.00	\$628.00	77.048432	81.033963	66.680064	81.933280	54.642283
Cabrera High School	Charter	1858	\$1,081,356.00	\$582.00	83.061895	83.975780	94.133477	97.039828	91.334769
Figueroa High School	District	2949	\$1,884,411.00	\$639.00	76.711767	81.158020	65.988471	80.739234	53.204476
Ford High School	District	2739	\$1,763,916.00	\$644.00	77.102592	80.746258	68.309602	79.299014	54.289887
Griffin High School	Charter	1468	\$917,500.00	\$625.00	83.351499	83.816757	93.392371	97.138965	90.599455
Hernandez High School	District	4635	\$3,022,020.00	\$652.00	77.289752	80.934412	66.752967	80.862999	53.527508
Holden High School	Charter	427	\$248,087.00	\$581.00	83.803279	83.814988	92.505855	96.252927	89.227166
Huang High School	District	2917	\$1,910,635.00	\$655.00	76.629414	81.182722	65.683922	81.316421	53.513884
Johnson High School	District	4761	\$3,094,650.00	\$650.00	77.072464	80.966394	66.057551	81.222432	53.539172
Pena High School	Charter	962	\$585,858.00	\$609.00	83.839917	84.044699	94.594595	95.945946	90.540541
Rodriguez High School	District	3999	\$2,547,363.00	\$637.00	76.842711	80.744686	66.366592	80.220055	52.988247
Shelton High School	Charter	1761	\$1,056,600.00	\$600.00	83.359455	83.725724	93.867121	95.854628	89.892107
Thomas High School	Charter	1635	\$1,043,130.00	\$638.00	83.350937	83.896082	93.185690	97.018739	90.630324
Wilson High School	Charter	2283	\$1,319,574.00	\$578.00	83.274201	83.989488	93.867718	96.539641	90.582567
Wright High School	Charter	1800	\$1,049,400.00	\$583.00	83.682222	83.955000	93.333333	96.611111	90.333333

It can be seen that there is a small shift in Thomas High School's scores and percentages, however this is a small deviation in the larger data set. Overall, the district's scores would not be affected enough to warrant different funding or support in the future.

2. How is the school summary affected?

On a school level, the data changes are more impactful.. Below is a closer look at Thomas High School's scores prior to the data clean up:

83.418349 83.848930 93.272171 97.308869 90.948012

As well as after the clean up:

83.350937 83.896082 93.185690 97.018739 90.630324

The first number is the average math score for the school, this average only decreased by less than .1. The second number is the average reading score for the school, this average actually increased by .05, which would help the school in funding decisions and teacher performance pay. The third and fourth numbers are the percentages for passing math and reading, these percentages only decreased by less than a full 1% each. The final number is the overall passing percentage, which was only decreased by .3% after the data was cleaned.

3. How does replacing the ninth graders' math and reading scores affect Thomas High School's performance relative to the other schools?

These two tables are arranged by highest and lowest performing schools after the data was cleaned. The top table is the 5 highest, and the bottom is the 5 lowest scoring schools.

It can be seen that even after the Thomas High School lost some data due to inaccuracies in the 9th grade, they still remained the second highest scoring school in the district. When the original data for Thomas High School is substituted, they are still 2nd place. The team used this to determine that the scores relative to other schools was not impacted.

	School Type	Total Students	Total School Budget	Per Student Budget	Average Math Score	Average Reading Score	% Passing Math	% Passing Reading	% Overall Passing
Cabrera High School	Charter	1858	\$1,081,356.00	\$582.00	83.061895	83.975780	94.133477	97.039828	91.334769
Thomas High School	Charter	1635	\$1,043,130.00	\$638.00	83.350937	83.896082	93.185690	97.018739	90.630324
Griffin High School	Charter	1468	\$917,500.00	\$625.00	83.351499	83.816757	93.392371	97.138965	90.599455
Wilson High School	Charter	2283	\$1,319,574.00	\$578.00	83.274201	83.989488	93.867718	96.539641	90.582567
Pena High School	Charter	962	\$585,858.00	\$609.00	83.839917	84.044699	94.594595	95.945946	90.540541

	School Type	Total Students	Total School Budget	Per Student Budget	Average Math Score	Average Reading Score	% Passing Math	% Passing Reading	% Overall Passing
Rodriguez High School	District	3999	\$2,547,363.00	\$637.00	76.842711	80.744686	66.366592	80.220055	52.988247
Figueroa High School	District	2949	\$1,884,411.00	\$639.00	76.711767	81.158020	65.988471	80.739234	53.204476
Huang High School	District	2917	\$1,910,635.00	\$655.00	76.629414	81.182722	65.683922	81.316421	53.513884
Hernandez High School	District	4635	\$3,022,020.00	\$652.00	77.289752	80.934412	66.752967	80.862999	53.527508
Johnson High School	District	4761	\$3,094,650.00	\$650.00	77.072464	80.966394	66.057551	81.222432	53.539172

4. How does replacing the ninth-grade scores affect the following:

- Math and reading scores by grade

9th grade was impacted as it contained inaccurate data, making the group as a whole not countable. The team had to replace 9th grade at Thomas High School with a NaN score. The other grades were not impacted, scores and percentages were not changed.

- Scores by school spending

Even though Thomas High School lost 9th grade scores, the student count did not change between the original and data cleaned tables. While there was a small shift in how much the school spent per student, approximately \$2, this is based on the assumption of less data. However the original count of students remained the same as the school size does not differ. There was not a significant change for the District to see or to make different decisions in the future.

When compared to other school's spending reports, it can be seen that higher spending does not always equal higher scores. Thomas High School spent \$638 per student and was second highest in scores, while schools like Johnson High School and Hernandez High School spend \$650 and above, and scored in the lowest 5 schools.

- Scores by school size

Thomas High School did not have a change to the school size, as the student count was not changed in the cleaned data set from the original data set.

The school scores did change as we invalidated 9th grade, however this did not change the fact that these students are still enrolled and count for District decisions such as funding or teacher allocations.

When compared to other school's enrollment reports, it can be seen that enrollment numbers hold some impact on scores. Thomas High School as 1635 student enrolled and scored second in the district. Schools with triple the students,

like Johnson High School with 4761 students enrolled scored the lowest of all schools.

- Scores by school type

Thomas High School did not have a change to the school type, as the federal distinction of Charter School was not impacted by the cleaning of data.

When compared to other school's reported types, it can be seen that school type has a great impact on scores. All top five scoring schools are Charter schools, Thomas High School included. District controlled schools are all lowest scoring.

References:

Top 5 highest performing:

	School Type	Total Students	Total School Budget	Per Student Budget	Average Math Score	Average Reading Score	% Passing Math	% Passing Reading	% Overall Passing
Cabrera High School	Charter	1858	\$1,081,356.00	\$582.00	83.061895	83.975780	94.133477	97.039828	91.334769
Thomas High School	Charter	1635	\$1,043,130.00	\$638.00	83.350937	83.896082	93.185690	97.018739	90.630324
Griffin High School	Charter	1468	\$917,500.00	\$625.00	83.351499	83.816757	93.392371	97.138965	90.599455
Wilson High School	Charter	2283	\$1,319,574.00	\$578.00	83.274201	83.989488	93.867718	96.539641	90.582567
Pena High School	Charter	962	\$585,858.00	\$609.00	83.839917	84.044699	94.594595	95.945946	90.540541

Bottom 5 lowest performing:

	School Type	Total Students	Total School Budget	Per Student Budget	Average Math Score	Average Reading Score	% Passing Math	% Passing Reading	% Overall Passing
Rodriguez High School	District	3999	\$2,547,363.00	\$637.00	76.842711	80.744686	66.366592	80.220055	52.988247
Figueroa High School	District	2949	\$1,884,411.00	\$639.00	76.711767	81.158020	65.988471	80.739234	53.204476
Huang High School	District	2917	\$1,910,635.00	\$655.00	76.629414	81.182722	65.683922	81.316421	53.513884
Hernandez High School	District	4635	\$3,022,020.00	\$652.00	77.289752	80.934412	66.752967	80.862999	53.527508
Johnson High School	District	4761	\$3,094,650.00	\$650.00	77.072464	80.966394	66.057551	81.222432	53.539172

Summary

Four changes that occurred due to the cleaning of 9th grade data were to the school spending per student, reading percentages, math percentages, and overall passing percentages.

In the section of school spending per student, there was a small shift. While the math code was being created, there was an assumption of less data due to invalidating the 9th grade scores. This led to a small change, approximately \$2. However the overall student enrollment numbers did not change, which could lead to mistakes in future assessments.

In the section of percentages, the three major changes were to the reading percentage, math percentages, and overall passing percentages. Each percentage was changed by approximately .1% with the invalidation of 9th grade.

At a District level, the change of only .1% along with the reasoning, there was no major change for the District to be concerned about. Although, the District and Board may be concerned as to why a whole grade level at Thomas High School was deemed invaled. The team and school team should be prepared for questions as to why this data was invaled.