Budget Educator

Aarti Patel, Gavin Chan, and Tony Huynh

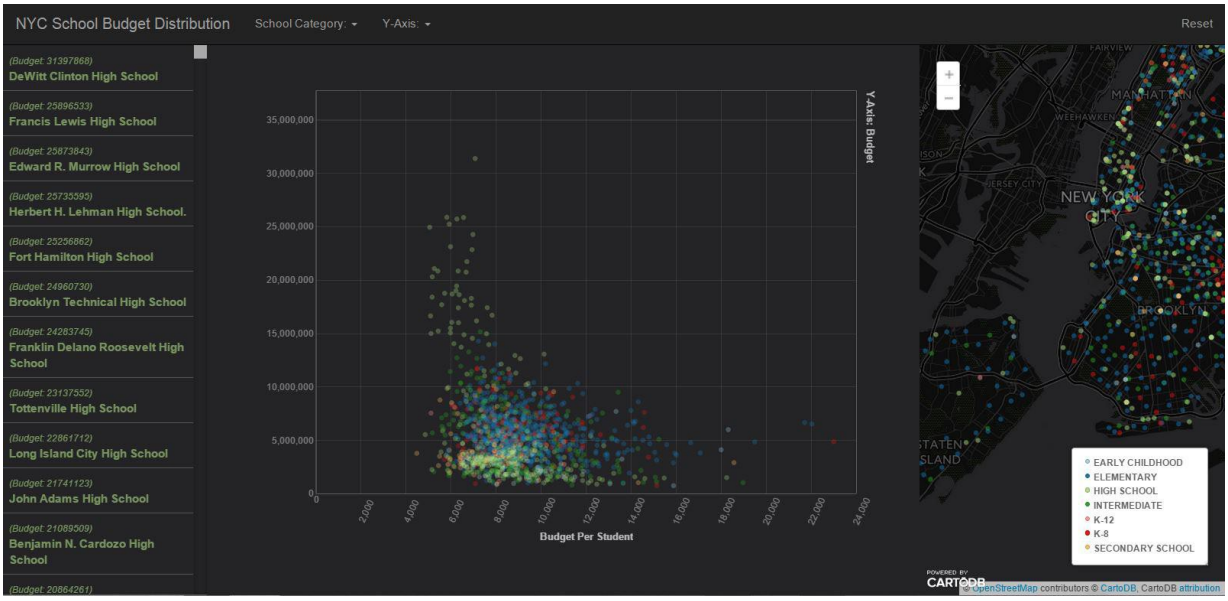

Fig. 1. The interface for Budget Educator as it is presented after loading the web page.

Abstract—Budget educator is an information visualization tool that displays the budget data from NYC public schools with respect to various academic measurements, such as ELA and Math exam grades. The tool is split into three views: a list, a scatterplot, and a map, each displaying school information depending on the selected y-axis value. This y-axis value can be selected from the drop-down menu on the navigation bar at the top of the webpage. Schools can also be filtered by their category, such as elementary, intermediate, etc.

Index Terms—Budget, data exploration, public schools, student performance

1 Introduction

In 2006, the court case "CFE vs State" reached a verdict to provide a minimum of roughly $5 billion dollars across 4 years for NYC School funding to allow students the opportunity to a "sound basic education" [1]. Over the past few years, however, funding decisions have been in flux as governors have been constantly trying to decide what variables are best suited to determine a school's monetary needs. The budget cuts in education continues with the recent proposal for a $1.3 billion cut in education aid by the New York State Legislature and the decline in money per child of students in New York City compared to outer-borough students [2], [3].

Those in the Department of Education and the organizations are responsible for budget allocation for NYC schools. Currently, they have determined how to allocate the budget based on a traditional and "Fair Student Funding" (FSF) methodology. This system uses student-body count, exam grades, poverty statistics, special education needs, school specialization, and various other factors to determine how much a given school should be funded. The FSF distribution would fluctuate depending on the budget amount available as a whole, thus may differ year to year.

With the release of agency data made available in 2013 by the New York City Open Data portal from the city government came further information on exam results from grades 3-8 and progress reports from 2006-2015. Similarly, with data on the each school's total granted budget from the New York City Department of Education, the individuals responsible for budget allocations can get an understand of whether or not a 'sound basic education' has been achieved and to what extent a school's budget influences the quality of education.

The visualizations arising from Budget Educator may become vital tools for understanding what factors can show an individual school's need. Those in charge of budgeting would need to rely less on intuition and experimentation for their budgets and be able to refer a high level visualization to gain an understanding of the current situation across the NYC school system. It may similarly reveal the adverse effects of poor funding distribution across the various schools and raise awareness of the ones that are hidden in the large collections of data.

The primary task is to be able to visualize the results of the "Fair Student Funding" (FSF) methodology, looking into student performance and school environment. In doing so, we can start to see what schools are being favored and if that preference has allowed all students to get a "sound, basic education". To begin, the project should focus on readily available data, such as exam test results in each grade. Over the course of building the project, we can integrate more factors about the classroom environment (such as average class sizes) to see what conditions are like in schools with higher/lower performing students and see what situations are potentially being ignored. With this data, we can generate and verify hypotheses regarding these factors and the effect of funding on them.

2 Related Work

After the verdict of 'CFE v. State of New York' in 2006, there had been several others who were interested in the results of budget distribution for NYC schools. Professors at NYU and Syracuse have taken interest in this topic and, in 2009, wrote a paper discussing the budgeting decisions for the districts in NY. They were curious if equity be improved in any way, so they needed to

understand the current budget methods at the time. Although it is not entirely similar to this proposed project, it gives us an insight on how that past budgeting data was analyzed and what kind of figures these authors presented.

For the most part, the analysis was primarily mathematical and used tables containing various calculations to express their findings. The traditional methods of explanation they used, however, could be improved upon by using visualizations to emphasize the trends they were trying to describe. Descriptions such as, "Funding does not respond 'crisply' to changes in characteristics of schools, even over a three year period (2001 – 2004)...there seems to be a good deal of 'inertia' in funding with sluggish responses to changes in school needs" [4] could be more easily described if there was a visual representation to clearly present that trend.

In addition to Schwartz's paper, Gecker from Princeton also presented an interest in the equity and fairness of the new fund distribution methodology by doing her research with more recent data in 2012 [5]. From a cursory reading, the approach is similar to our proposed project where she used collected data to see the trends and patterns of distributed funds across schools depending on their needs (families that receive public financial assistance) and their exam performance. The conclusion that was reached was that "spending per student was overall positively associated with student need", which would suggest that FSF was, in fact, being more fair in fund distribution. However, it was difficult to make the claim that it was an improvement from the previous method of funding since there was insufficient data prior to 2006.

The concern for Budget Educator will then be whether the funding changes from 2007 have made any impact on the education performance of schools while keeping in mind that other prior investigations have shown that the distribution of that money have been relatively "fair". In addition, the previous documents have only used tables or line graphs to present the information and, while efficient, may only be useful to those who understand the mathematical data that is being presented. The hope for this project is that it can be understood and used by councils and other organizations involved in the distribution of funding who may not have an extensive background in data analysis.

3 Data Analysis and Abstraction

Our data sets all come from the NYC Open Data and the city's Department of Education site. The main datasets from NYC Open Data we will be using are Math Test Results from 2006-2012, English Language Arts (ELA) Test Results from 2006-2012, School Demographics and Accountability Snapshot from 2006-2012, Class Sizes per School from 2006-2011, and Progress Reports of All Schools from 2006-2011. To observe the amount of funding schools receive, we turn to the 'School Budget Overview' page of the NYC Department of Education's site. There, we can find the budgets across 2007 to 2015 identified by school name.

The Math Test Results from 2006-2012 dataset and English Language Arts (ELA) Test Results from 2006-2012 dataset, provide student exam results for the respective exams from 2006-2012 (grades 3-8), can be viewed citywide, which provides a top-level view of how students scored across the whole city (~50 rows of data), and can also be viewed by borough, which provides a deeper look of how students scored per borough (~250 rows of data). Each view, citywide and borough, which provides attributes of the grade, year, demographics/category, number of students tests, mean score, and a breakdown of the exam results based on the number of students to score in the respective range (level 1, level 2, level 3, and level 4), all of which will be useful to our project, can be further segmented into viewing exam results based off gender, race/ethnicity, and english proficiency status. The deeper we dive into the datasets, the larger it gets. For instance, if we look at exam results based off english proficiency status citywide, we get ~500 row of data, and if we look at exam results based off race/ethnicity by borough, we get about 1000 rows of data across a span of multiple years.

The School Demographics and Accountability Snapshot from 2006-2012 dataset, provide accounts of NYC public school student populations served by grade, special programs, ethnicity, and gender. Within this dataset, we see a breakdown of each school throughout 2006-2012 and can analyze its student population, taking in account number of students in each school, the number of students in each grade, and a breakdown of the student population according to race/ethnicity and gender. This dataset is quite large with over 10,000 rows of data, because it takes into account multiple years.

The Progress Reports of All Schools from 2006-2011 dataset, provides a general overview of each school, what type of school they are (high school, elementary/middle school) and the grade letter that they received for the corresponding years. This dataset has approximately 1700+ rows of data.

The budget data was extracted from the "Galaxy Table of Organization Budget" from the NYC Department of Education by using phantomjs and the web-scraping framework, pjscrape, by Nick Rabinowitz. Although the table on their page describes how the budget was allocated in more detail, because of the inconsistency of how that information is stored, the only information that could be reliably extracted was the grand total budget for each school.

As a whole, the data we have combined and processed has returned 1434 rows, each identified by a school's location code (DBN) and containing data from multiple attributes of interest. Due to lack of school grade data in more recent years, we made the design decision to focus this iteration on students graduating in the school year 2009 to 2010. The table below describes the number of entries we currently have as the input, including variables used in the table, in the scatterplot, and in the map. Since this dataset relies on the schools themselves entering the information, there have been several instances of incomplete entries, so we have attempted to maximize the number of available schools to present by allowing NULL values. The attributes on the following table marked with an asterisk (*) are ones used as the y-axis in the scatterplot. The x-axis in the scatterplot is fixed on the 'Budget Per Student' attribute. See Table 1 on the next page.

Because the main focus of Budget Educator is on the effects of each school's budget per student and the school's performance, we narrowed the scope of attributes specifically to fields most obvious to performance - exam scores, progress report grades, and number of students. In this final iteration, we excluded social influences that may or may not have affected budget allocation, such as racial demographics, gender demographics, and statistics on students with special needs. Therefore it should be noted that a school of special needs, such as one focused on students with disabilities, will have the same weight as a school without such a focus.

Similarly, there is no uniform reporting of how much each school spent on each student. Each school may have varying degrees of cost per student due to other external factors, such as renovation and teachers' salaries. Thus, there was a gross simplification on the 'Budget Per Student' as a transformation of the total school budget divided by the number of total number of students enrolled at the school. Because this does not take into account the external factors, this number is treated as the upper-bound of the 'Budget Per Student' attribute, or the ideal value should a school entirely invest in its students. Table 2 will describe the primary attributes received as input for Budget Educator.

Table 1. Summary of Data Variables by Row Count

Data Variable	**# of rows**	Percent of Cohort Who Took Regents*	362
DBN / Location Code	1434	Percent of Graduates Who Took Regents*	362
School Name	1434	Percent of Cohort Who Took Advanced Regents*	362
Total Budget	1434	Percent of Graduates Who Took Advanced Regents*	362
Total Enrollment*	1434	Percent of Cohort Who Took Regents without Advanced*	362
Budget Per Student	1434	Percent of Graduates Who Took Regents without Advanced*	362
School Type*	1434	ELA Mean Scale Score*	1053
School Category*	1434	Percent of Students Scoring Level 1 in ELA*	1053
Street Address	1434	Percent of Students Scoring Level 2 in ELA*	1053
Borough	1434	Percent of Students Scoring Level 3 in ELA*	1053
[Longitude, Latitude] / Coordinates	1434	Percent of Students Scoring Level 4 in ELA*	1053
Report Grade*	1434	Math Mean Scale Score*	1053
Number of Students in Cohort*	362	Percent of Students Scoring Level 1 in Math*	1053
Percent of Graduating Students in Cohort*	362	Percent of Students Scoring Level 2 in Math*	1053
Percent of Cohort Who Are Still Enrolled*	362	Percent of Students Scoring Level 3 in Math*	1053
Percent of Cohort Who Dropped Out*	362	Percent of Students Scoring Level 4 in Math*	1053

Table 2. Detailed Description of Attributes

Attribute Name	Type	Range/Size	Description
DBN / Location Code	Identifier	N/A	DBN, or Location Code, is the unique identifier for schools on the New York City Department of Education's database.
School Name	Identifier	N/A	The school name, as the name suggests, is the name the school is referred to by.
Total Budget	Quantitative	$762,163 - $31,397,868	The total budget is the total amount of funding granted to the school by Department of Education.
Total Enrollment	Quantitative	48 - 4947	The total enrollment is the number of students enrolled in the school (from all grade levels).
Budget Per Student	Quantitative	$4,473.07 - $22,996.58	The budget per student is the calculation of the total budget divided by the total enrollment. This value represents the upper-bound amount of spending for each students, not taking into considering cost of renovation, teachers' salaries, and other external factors.
School Type	Categorical	Career Technical, General Academic, Transfer School, Special Education	The school type is what the school focuses on. For instance, career technical schools dedicates their studies to specific career fields and special education schools cater to meeting the educational requirements of students with special needs.
School Category	Categorical	Early Childhood, Secondary School, Elementary, Intermediate, High School, K-8, K-12	The school category specifies what grade level(s) the school takes in. For instance, intermediate schools take in students entering grades 6, 7, and 8, while elementary schools only take in K-5.
Street Address	Identifier	N/A	The street address, as the name suggests, is the street location of the school.
Borough	Categorical	Bronx, Brooklyn, Queens, Staten Island, Manhattan	The borough is the borough of where the school is located in, specific only to schools in NYC.

Coordinates	Identifier	N/A	The coordinates place provides the latitude and longitude of the school, and placement on a map.
Report Grade	Categorical	A, B, C, D, F, None	The report grade presents the letter grade as a result of a Department of Education School Quality report. It grades the overall school's practice, environment, and performance based on feedback by parents and students.
Number of Students in Cohort	Quantitative	0 - 1,134	The number of students in cohort is the number of students who are educated at the same period of time, often referring to the graduating year.
Percent of Graduating Students in Cohort	Quantitative	0% - 100%	The percent of graduating students in cohort is the percentage of students in the cohort (students educated at the same time) who are graduating.
Percent of Cohort Who Are Still Enrolled	Quantitative	0% - 100%	The percent of students in cohort still enrolled is the percentage of students in the cohort (students educated at the same time) who are not graduating and will be staying for an extended period of time.
Percent of Cohort Who Dropped Out	Quantitative	0% - 100%	The percent of students in cohort who dropped out is the percentage of students in the cohort (students educated at the same time) who have left or dropped out of school.
Percent of Cohort Who Took Regents	Quantitative	0% - 100%	The percentage of cohort who took regents is applicable only to those in grade levels 10-12 (high school) and refers to the percentage of students in the cohort who took the Regents exams (testing on core high school subjects).
Percent of Graduates Who Took Regents	Quantitative	0% - 100%	The percentage of graduates who took regents is applicable only to those in grade levels 10-12 (high school) and refers to the percentage of graduating students within the cohort who took the Regents exams (testing on core high school subjects).
Percent of Cohort Who Took Advanced Regents	Quantitative	0% - 100% (Data: 0% - 99%)	The percentage of cohort who took regents is applicable only to those in grade levels 10-12 (high school) and refers to the percentage of students in the cohort who took the advanced form of the Regents exams (testing on core high school subjects).
Percent of Graduates Who Took Advanced Regents	Quantitative	0% - 100%	The percentage of graduates who took regents is applicable only to those in grade levels 10-12 (high school) and refers to the percentage of graduating students within the cohort who took the advanced form of the Regents exams (testing on core high school subjects).
Percent of Cohort Who Took Regents without Advanced	Quantitative	0% - 100% (Data: 0% - 93%)	The percentage of cohort who took regents is applicable only to those in grade levels 10-12 (high school) and refers to the percentage of students in the cohort who took the standard form of the Regents exams (testing on core high school subjects).
Percent of Graduates Who Took Regents without Advanced	Quantitative	0% - 100%	The percentage of graduates who took regents is applicable only to those in grade levels 10-12 (high school) and refers to the percentage of graduating students within the cohort who took the standard form of the Regents exams (testing on core high school subjects).
ELA Mean Scale Score	Quantitative	1 - 4 (Data: 1.262 - 3.596)	The ELA Mean Scale Score refers to the average score of the students who took the ELA standardized test, applicable only to grades 3-8. The marks fall into quadrants of 1 (lowest) to 4 (highest).
Percent of Students Scoring Level 1 in ELA	Quantitative	0% - 100% (Data: 0% - 74%)	The percent of students scoring level 1 in ELA is the percent of students who took the ELA standardized test and scored a Level 1 on the exam.
Percent of Students Scoring Level 2 in ELA	Quantitative	0% - 100% (Data: 0% - 69%)	The percent of students scoring level 2 in ELA is the percent of students who took the ELA standardized test and scored a Level 2 on the exam.
Percent of Students Scoring Level 3 in ELA	Quantitative	0% - 100% (Data: 0% - 72%)	The percent of students scoring level 3 in ELA is the percent of students who took the ELA standardized test and scored a Level 3 on the exam.
Percent of Students Scoring Level 4 in ELA	Quantitative	0% - 100% (Data: 0% - 59%)	The percent of students scoring level 4 in ELA is the percent of students who took the ELA standardized test and scored a Level 4 on the exam.

Math Mean Scale Score	Quantitative	1 - 4 (Data: 1.676 - 3.920)	The Math Mean Scale Score refers to the average score of the students who took the Mathematics standardized test, applicable only to grades 3-8. The marks fall into quadrants of 1 (lowest) to 4 (highest).
Percent of Students Scoring Level 1 in Math	Quantitative	0% - 100% (Data: 0% - 38%)	The percent of students scoring level 1 in Math is the percent of students who took the Mathematics standardized test and scored a Level 1 on the exam.
Percent of Students Scoring Level 2 in Math	Quantitative	0% - 100% (Data: 0% - 67%)	The percent of students scoring level 2 in Math is the percent of students who took the Mathematics standardized test and scored a Level 2 on the exam.
Percent of Students Scoring Level 3 in Math	Quantitative	0% - 100% (Data: 0% - 61%)	The percent of students scoring level 3 in Math is the percent of students who took the Mathematics standardized test and scored a Level 3 on the exam.
Percent of Students Scoring Level 4 in Math	Quantitative	0% - 100% (Data: 0% - 92%)	The percent of students scoring level 4 in Math is the percent of students who took the Mathematics standardized test and scored a Level 4 on the exam.

4 Task Analysis and Questions

As described in the introduction, the decisions of budget allocation are based on a series of variables, such as student-body count, exam performance, and so on. This application aims to reveal whether that methodology is being followed by providing an overview of budget allocations in NYC public schools based on these attributes of interest. In particular, the questions that need to be answered are:

1. How does school funding play a role in overall student performance?
   - It is important to get a sense of how money is being used to aid student learning and whether schools with more resources are showing higher exam and graduation results. There may be instances where schools are being given high budgets, but not achieving high performance. On the flip side, it will be interesting to see schools that are receiving below average funding, but outputting above average exam and graduation results.
2. What correlations do we see with schools with higher budgets and student exam results?
   - Although there may not be a causal relation between higher budgets and student exam performance,, it would be beneficial to observe how the two factors interact with one another. A high correlation would suggest that poor performing schools will continue to fall behind since they are receiving less resources to work with for their students. However, a low correlation would suggest that there little relation between how schools are using those resources to improve their students' exam results.
3. How does the school budget affect the graduation/dropout rates of the school?
   - By observing any trends between graduation/dropout rates of schools based on budget, it is possible to get a sense of how well a school's budget might be used to help students that are falling behind or in risk of dropping out. If there does not seem to be any notion of a trend, then it is possible to consider that schools are not investing a large portion of these resources to helping students in need.
4. How are school budgets and performances affected by the location of the school?
   - Using location as an additional parameter of consideration, the user can see how budget allocation may vary depending on the type of area the school resides. Lower income areas may receive more funding than areas that are well off because of the demographic of students that might require more resources during school hours.

5 Visualization and Interaction Design

5.1 Overview

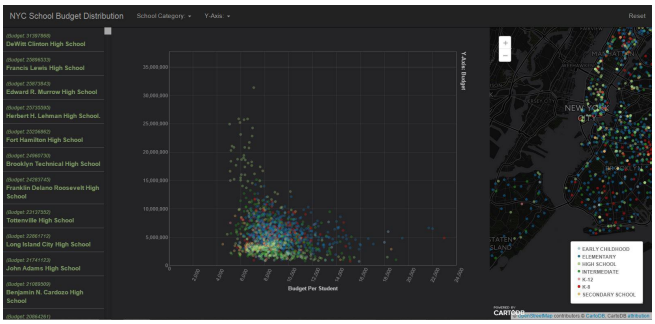

Fig. 2. The interface is broken into four components. The navigation bar on top, the list of schools on the left, the scatterplot in the middle, and the map on the right.

Budget Educator focuses on getting an understanding of the current situation across the NYC school system and, thus, features a filtering navigation bar, a list of schools, a scatter plot, and a map. On the navigation bar are two drop-downs that allow the user to filter and limit the data to specific school categories and attributes related to a school's academic performance. On the sidebar is a list of school names currently present in the scatterplot. The scatterplot measures a variety of academic performance factors against an estimated budget per student attribute, allowing the user flexibility to choose which attribute they want to observe and form conclusions based on the visualization presented. The map plots the schools present in the scatterplot on the map of New York City, allowing the user to observe whether there are any geographical correlation in the data. Users also have the option to reset the tool to its default state using the 'Reset' button in the navigation bar.

5.2 Navigation Bar

The navigation bar consists of two drop-downs and a reset button. The drop-downs function as filters to allow those selected to be reflected in the table, the scatterplot, and the map.

The drop-down on the left, titled 'School Categories,' allows the user to filter out schools that are not selected. If there are no selections, the default would display all schools from all categories. Selections of the filter are toggle-able, so multiple school categories can be selected and displayed at once. An initial implementation without multiple selection revealed that, because schools of different categories are not mutually exclusive, it would be difficult to compare different categories of schools on the same map for a common attribute. The first two school categories selected will be

displayed in the navigation bar for ease of remembrance, while those selected will be shaded grey in the drop-down. For instance, if a user only wants to view elementary and intermediate schools, they would select 'Elementary' and 'Intermediate' from the drop-down. This would then be reflected as shaded cells in the drop-down, displayed in the navigation bar, and update the school list, the scatterplot, and the map to include only elementary schools and intermediate schools.

The 'Y-Axis' drop-down enables the user to switch between different attributes that can be used as the y-axis of the scatterplot. Currently, it is defaulted to the first attribute on the drop-down, the 'Total Enrollment' of a school. Only one attribute can be selected at a time and will be used against the fixed x-axis, the 'Budget Per Student.' This gives the user flexibility to choose and visualize what attributes influence the estimated budget per student allocation for schools.

The reset button, located at the far right of the navigation bar, allows the user to reset the current selection back to its default state - that is, no selection. There will be no filter on the schools by categories, so all schools will be shown. All schools currently highlighted in the scatterplot or selected in the school list will be unhighlighted and deselected. The y-axis selection, however, will remain unchanged because a user is more likely to click this button only to unfilter and deselect for a clear slate of the interface.

5.3 School List

The school list is a scrollable list of schools ordered by descending y-value. Each cell consists of a school's name in bold and their value given the current y-axis selection, colored based on their school category in the legend. In its deselected state, the cell is seen in 50% opacity but will revert to a full opacity on a black background upon selection. Hovering over an unselected cell will similarly display in full opacity to increase readability, but will not change the background color. The shift in opacity and background color provides a sharper contrast, increasing readability and indicator toward user's focus. A school can be selected and deselected by clicking their cell on the school list. Once selected via the school list, the selected schools will display in full opacity in the scatterplot and visible on the map.

5.4 Scatterplot

To the right of the school list is the scatterplot, which presents each school as a point, colored depending on that school's category. The x-axis will always represent the budget-per-student value for each school whereas the y-axis will change depending on the currently selected drop-down item found on the navigation bar. When a user clicks and drags on the scatterplot, they will be able to highlight points of interest that persist through any y-axis changes. This allows a user to easily see how particular schools stand compared to other schools in various academic measurements. The selection is also reflected on the school list and the map. This design was chosen for its intuitive presentation and interaction as opposed to multiple graphs and views that may flood the page with information. A single academic measurement is presented at one time to allow the user to focus on that particular topic and highlight the ones that may require further investigation, which is enabled through the y-axis changes.

5.5 Map

The rightmost view is a map that is zoomed to fit the New York City area. This view is populated with points representing the location of each school and is colored by school category with the legend described on the bottom right corner. Each point can be clicked for further information about the school's name. Selections made on the navigation bar's school category drop-down, the school list, and/or the scatterplot will be reflected on the map. The map aims to assist users in understanding how schools of similar behavior may cluster in certain locations or even show how widespread a particular issue may be in the area. Without the map, the user loses a sense of where these problems are occurring and lessens the impact of patterns that might have been found in the scatterplot.

6 Findings and Insights

6.1 Finding 1: "Highest Budget & Highest Budget Per Student versus Exam Performance"

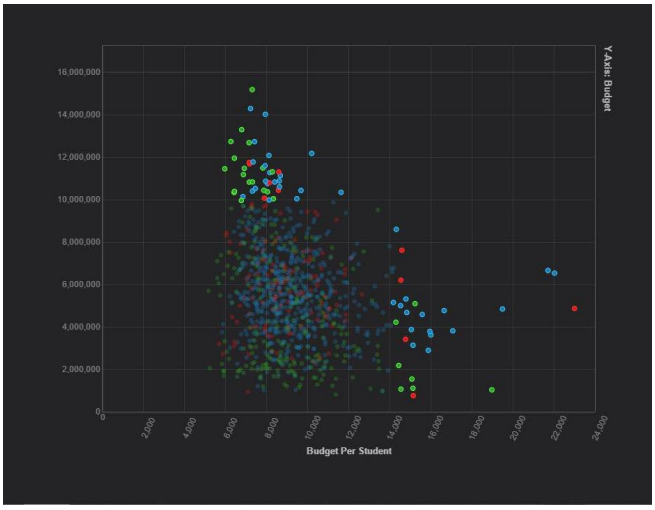

Fig. 3. Highlighted plots of the schools receiving the highest budget and the highest budget per student

When talking about budget, it is an immediate concern whether or not the money being spent has any noticeable impact. In this case, we want to observe whether or not the schools with the highest budget and the highest budget per student are outperforming the other schools. The schools with the highest budget are of interest because they are likely to be the ones with the most resources to spend on their students. Likewise, the schools with the highest budget per student may not necessarily have a lot of money to work with, but they have less students to spend on, so they might be expected to use that amount to its full extent. In Fig. 3, we only observe Elementary, Intermediate, and K-8 schools because those are the grade levels that have data for the Math and ELA exams we will use to measure exam performance. The figure above shows two distinct groups in the scatterplot: the top-left highlighted points that show the schools with the highest budget and the bottom-right highlighted points that show the schools with the highest budget per student. Note that the schools with the highest budgets do not exceed $12,000 per student and will remain on the left side of the scatterplot while the schools with the highest budget per student will remain to the right of the $14,000 tick mark on the x-axis.

After switching the y-axis to ELA Mean Scale Score in Fig. 4, the two groups are all seen to be below the 3.0 mark and predominantly performing the same as the rest of the schools in NYC. This may seem unintuitive at first, but it becomes more clear when you realize that the schools with the highest budgets also are the schools with vastly more students, which may make it hard for the schools to maintain a high performance for each person. In addition, the schools with the highest budget per student are schools that have special needs, such as 'The 47 American Sign Language & English Lower School' which needs to have specially trained staff to be able to communicate in sign language.

This average performance is also seen in the Percent of Students Scoring Level 4 in ELA, Math Mean Scale Scores, and the Percent of Students Scoring Level 4 in Math, pictured below in figures 5, 6, and 7 respectively. Note that Level 4 in ELA and Math indicates that the student is proficient in the subject.

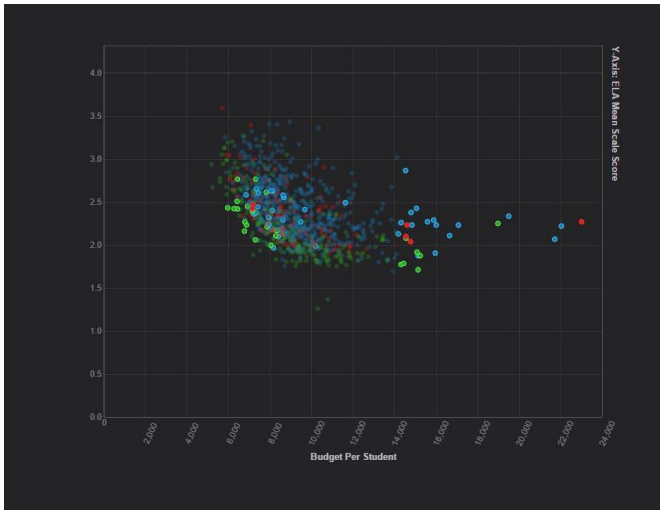

Fig. 4. Previous highlights on budget are now shown in relation to ELA Mean Scale Score

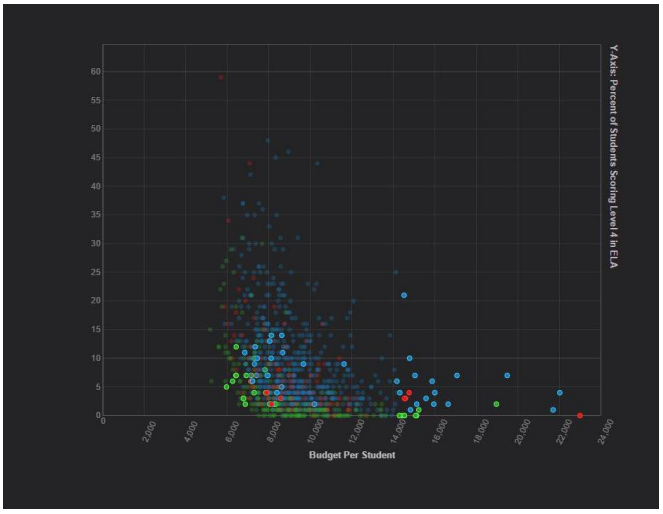

Fig. 5. Previous highlights on budget are now shown in relation to Percent of Students Scoring Level 4 in ELA

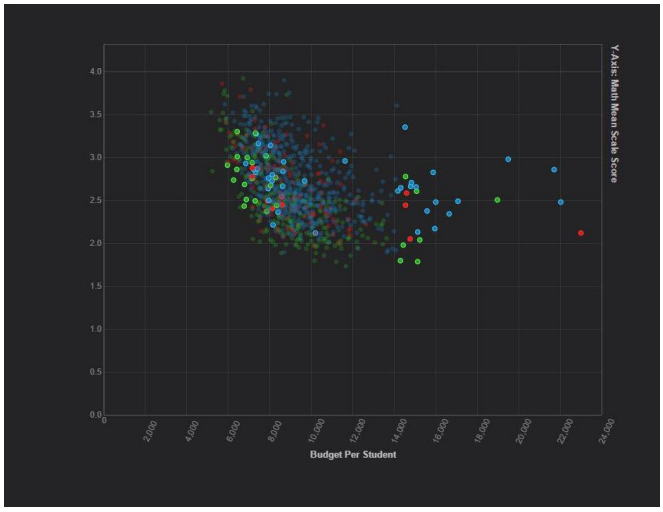

Fig. 6. Previous highlights on budget are now shown in relation to Math Mean Scale Score

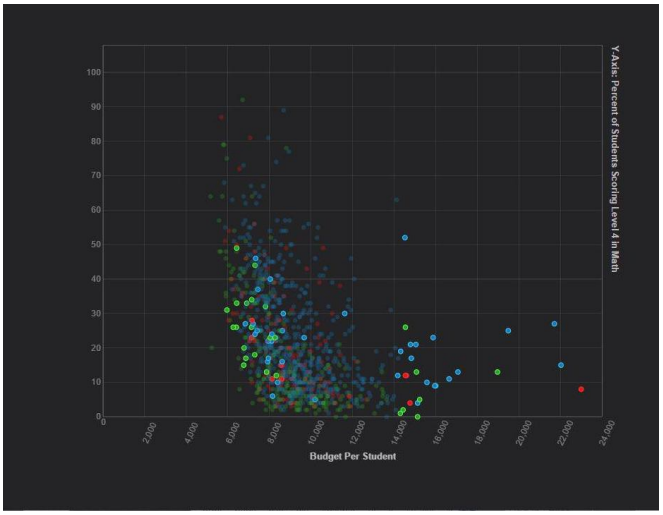

Fig. 7. Previous highlights on budget are now shown in relation to Percent of Students Scoring Level 4 in Math

6.2 Finding 2: "Top Performing Schools Do Not Have High Budgets"

From the previous finding, it was possible to observe that budgets do not seem to have a forceful impact on school performance, so it begs the question: what kind of budgets do high performance schools have? Fig. 8 and 9 show the process of highlighting the schools of interest, namely the schools that perform higher than a 3.0 on the ELA exams and 3.5 on the Math exams.

At an initial glance, it is possible to see that these school are not receiving an above-average budget per student. Apart from the outlier on the far right, these schools lie fairly closely to all the other schools on the x-axis. Even more so, there does not seem to be any indication of a correlation between increasing budget per student and higher performance. To emphasize this point, Fig. 10 shows that these same high-performing schools receive similar amounts as other schools.

Turning our attention towards the map (Fig. 11), notice that these highest performing schools are located in middle to lower Manhattan, downtown Brooklyn, and Queens. These areas are now locations of interest that can be further investigated to examine the demographic of the student body and this may reveal why the schools are performing better. However, although these schools are not receiving more budget than the majority, maybe they are receiv-

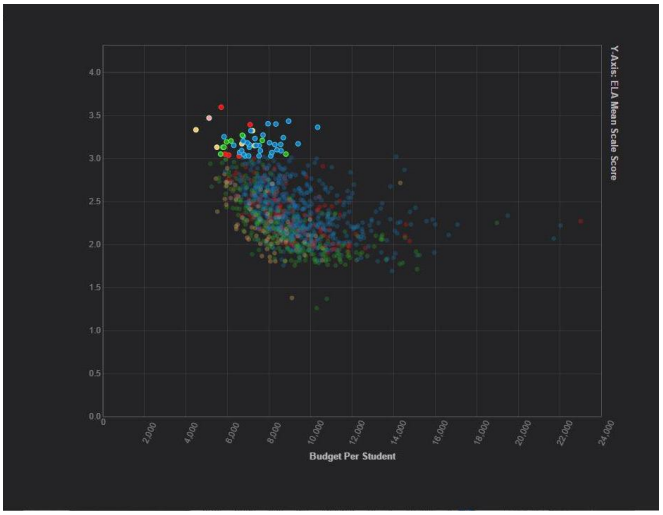

Fig. 8. Highest performing schools on ELA exams

-ing more aid than schools with poor performance.

Upon investigating that possibility, it is visible that this is not

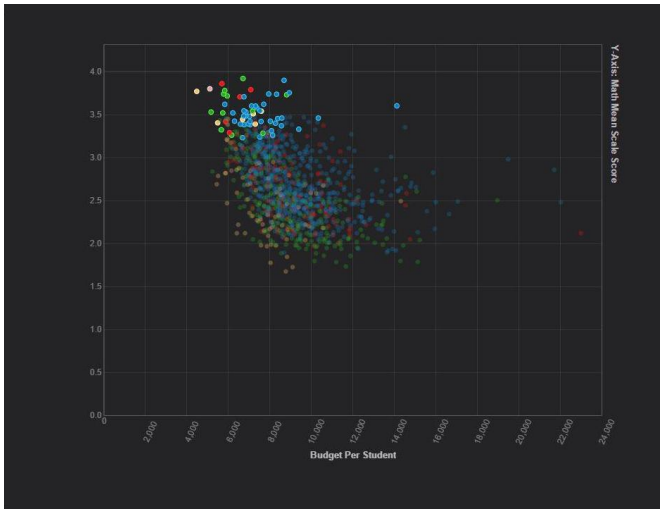

Fig. 9. Previous highlight of schools along with the schools with the best performance on Math exams

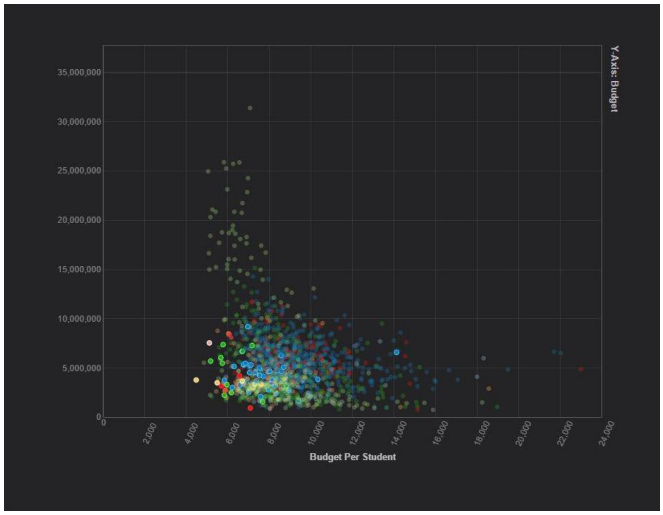

Fig. 10. Budget amounts of highest performing schools

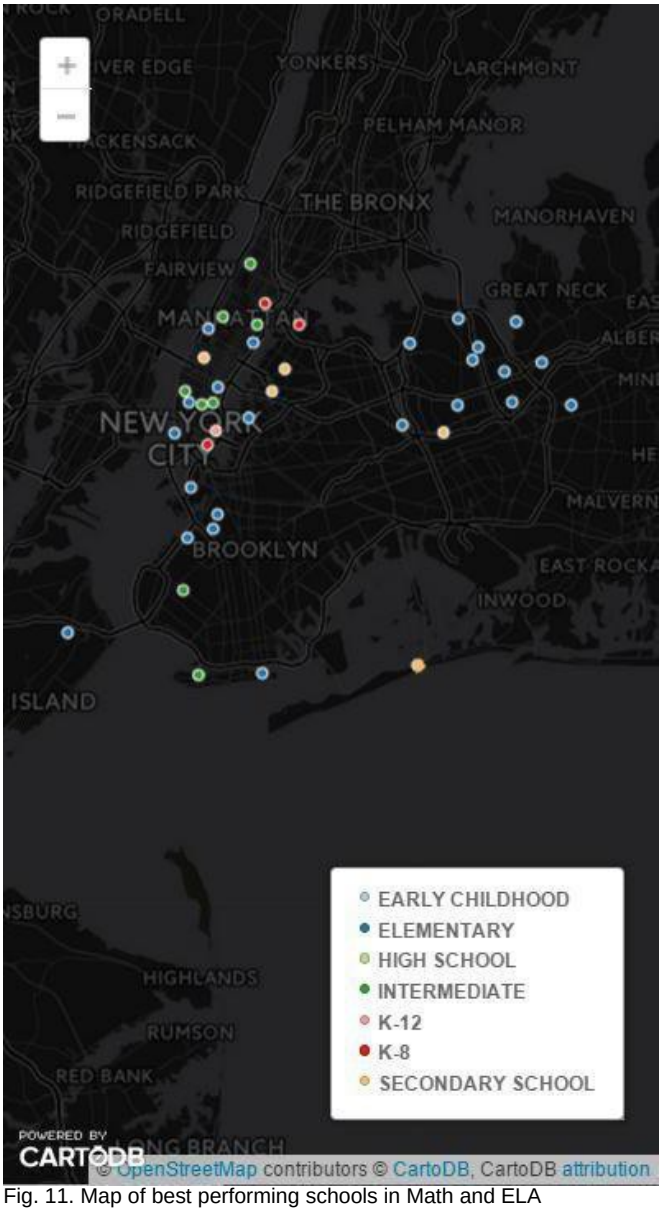

Fig. 11. Map of best performing schools in Math and ELA

the case and, if anything, poor performing schools are actually receiving a larger amount per student (Fig. 12, 13, and 14).

It is also possible to see that the schools with the lowest scores on exams cluster in different areas in NYC. In Fig. 15, there is a heavy concentration of low performing schools in upper Manhattan, Bronx, and the middle of Brooklyn. What might be interesting, is an investigation of how higher performing schools are using their budget and apply those practices to these lower performing schools. This can be made possible since the list of schools on the left side can inform the user of what schools to research.

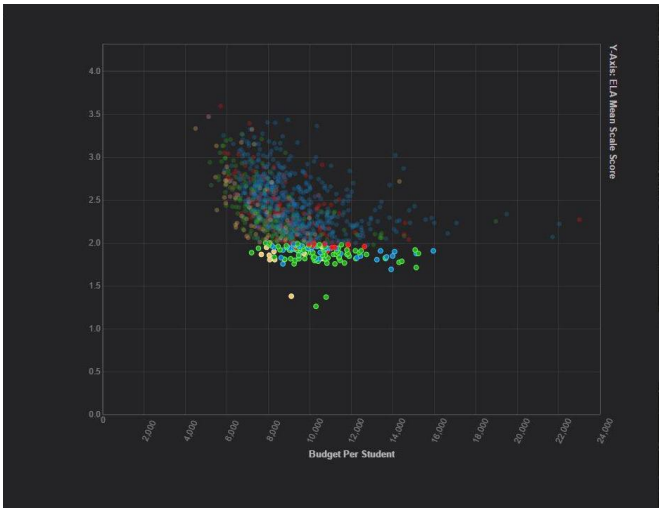

Fig. 12. Lowest performing schools on ELA

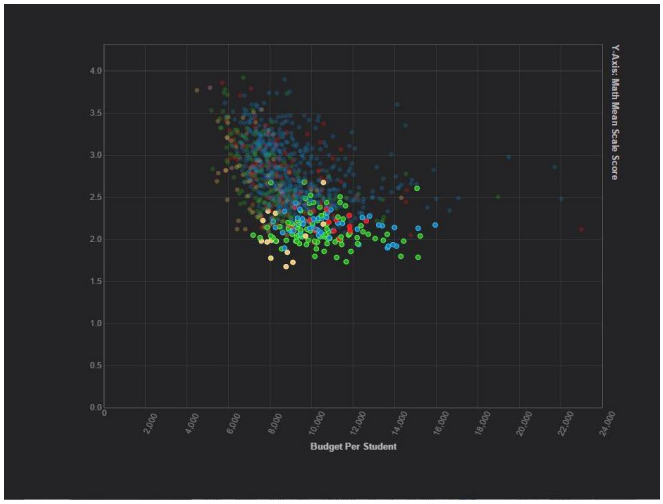

Fig. 13. Lowest performing schools on ELA and Math

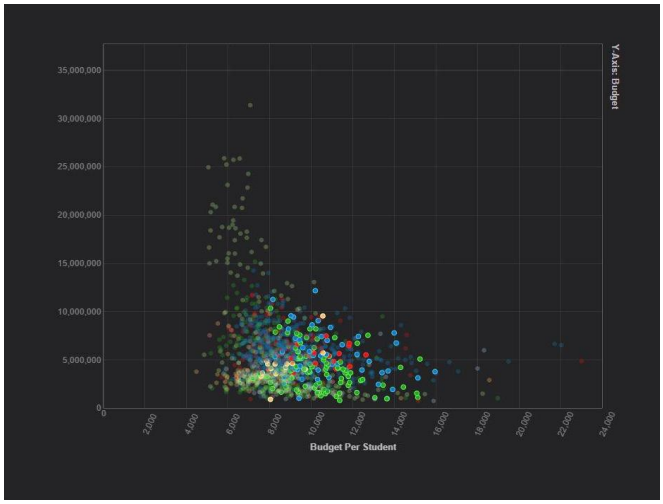

Fig. 14. Budget per student for lowest performing schools in ELA and Math

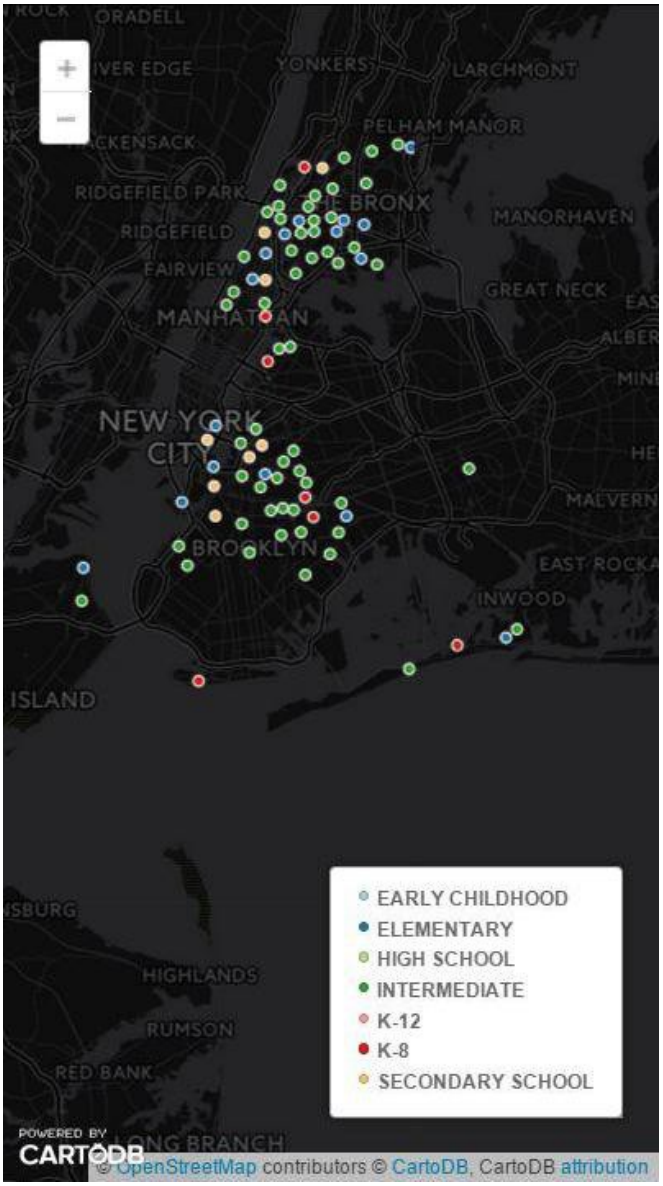

Fig. 15. Map of lowest performing schools in ELA and Math

7 Limitations and Future Work

As of now, Budget Educator provides visualization for the data on students supposedly graduating in the school year 2009-2010. Because only one school year's sample was used for this iteration, we are unable to see progression on how the budget allocation provided by the Department of Education changed across the years. Instead, we are only able to observe the effects of such budget provided in the school year 2009-2010 and its relationship to other performance factors within the same chronological scope. Similarly, there was an abundance of data related to a school's demographics and geographical location that can be used to branch towards other studies. Instead, Budget Educator focuses on factors directly associated with graduation rates and exam performances, thus only ranking a school by its academic achievements.

Should there be more time to develop the project, the first and foremost action is to address the social influences we excluded from this iteration, namely racial demographics, gender demographics, and statistics on students with special needs. With this expansion, we hope to observe whether these demographic influences play a role in not only the funding of the school, but the quality of education provided. Similarly, because this iteration only features students graduating in the school year 2009-2010, an additional feature we may implement would be to expand into years prior to 2010. The reason for this decision in the first place was due to lack of school grade data after 2010, thus the school year 2009-2010 was chosen to represent the most 'recent' sample. An expansion to years prior to 2010 would provide insight as to how the funding changed over the years, particularly after 2007 when school-level spending became more 'equitable' [5].

8 Conclusion and Lessons Learned

Initially, the project seemed fairly straightforward and the team had various ideas of how to approach the design of the visualization. Our main intention was to have the map inform the user of how budget varied in different locations and if that budget played a role in the students' performances. However, data collection was more difficult that we initially thought since the school locations were not immediately available to us in our initial data sets. In addition, the pre-processing needed to combine the various datasets together was an experience that led us all to reaffirm the power of SQL. We then

spent quite a bit of time using Tableau to produce prototypes of potential designs for the visualization to arrive at the final list/scatterplot/map interaction. Through this experience, our team gained a sense of how valuable quick prototypes are for learning how to display information in an understandable way and how some designs just cannot convey a message.

References

[1] Rebell, Michael. “CFE v State of New York: Past, Present and Future”. (Summer 2011) <http://schoolfunding.info/wp-content/uploads/2011/07/CFE-Past-Present-and-Future.pdf>

[2] DeMause, Neil, and Elizabeth Green. "The Campaign for Fiscal Equity Lawsuit Was the Best Hope for City Schools. It Failed." The Campaign for Fiscal Equity Lawsuit Was the Best Hope for City Schools. It Failed. Village Voice, 21 Jan. 2009. Web. 09 Mar. 2015. <http://www.villagevoice.com/2009-01-21/news/the-campaign-for-fiscal-equity-lawsuit-was-the-best-hope-for-city-schools-it-failed/>.

[3] Rebell, Michael. "Rebell: Cuomo's School Cuts Are Unconstitutional." NY Daily News. N.p., 17 Mar. 2011. Web. 09 Mar. 2015. <http://www.nydailynews.com/opinion/cuomo-school-cuts-unconstitutional-lead-lawyer-campaign-fiscal-equity-case-article-1.124492>

[4] Schwartz, Amy. Rubenstein, Ross. Stiefel, Leanna. “Why Do Some Schools Get More and Others Less? An Examination of School-Level Funding in New York City” (November 18, 2009) <https://steinhardt.nyu.edu/scmsAdmin/uploads/003/060/4-Why%20Do%20Some%20Schools%20Get%20More%20and%20Others%20Less.pdf>

[5] Gecker, Alexandra “‘Fair Student Funding:’ Did School-Level Spending Become More Equitable in New York City Public Schools After 2007?” (2012) <http://www.princeton.edu/economics/undergraduate/senior-independent-work/sampletheses/Sample_three.pdf>

Links

Github: https://github.com/NYU-CS6313-Projects/Budget-Educator
Demo: http://nyu-cs6313-projects.github.io/Budget-Educator/
Video uploaded: https://vimeo.com/128095858