Applied Data Science Capstone
Final Project

Description of the Problem & Discussion of the Background

What is the problem?

American students fall behind their global peers in math performance, and it does not appear to be improving. The average math SAT score of college-bound high schools seniors remained stagnant between 1972 and 2011, despite the US' real GDP per capita doubling during that time.(1) The failure to improve throughout the years has compromised the US' position in the global marketplace.

Mathematical capability impacts earning potential. College graduates with math, engineering, or physical science majors earn 19% more than other fields.(1) American job opportunities in the field are robust. McKinsey's Global Institute estimates 2 million workers with advanced analytical and technical skills will soon be needed. However, if trends continue, a large fraction of these jobs will be outsourced to professionals abroad.(2)

Although there is a demand for it, US public schools have failed to position students to obtain these high paying jobs. US public schools have especially let down low-income students, who have a much greater need for such an opportunity. Since 1960, the gap in standardized test scores between affluent and low-income students has grown by 40%.(3)

What is the solution?

Learning outside of the classroom can have substantial impact on math achievement, particularly for low-income and minority students. After-school participation narrows the math achievement gab between high, middle, and low income students. Income differences in math achievement were eliminated all together for students who had consistent after school engagement across K-5. (5)

In order to improve the US' position in the global marketplace and increase earning potential for struggling American families, investments in after school math programs should be made in low income areas and directed toward students K-5.

Resources:

1. https://www.educationnext.org/solving-america%E2%80%99s-math-problem/
2. https://www.educationnext.org/are-u-s-students-ready-to-compete/
3. https://www.nytimes.com/2012/02/10/education/education-gap-grows-between-rich-and-poor-studies-show.html
4. 2015 DreamBox Learning, Inc, Tim Hudson, White Paper

Description of the Data & How It Will Be Used to Solve the Problem

For this project, I am going to determine the best location for an after school math program in New York city.

The data I will used will come from NYC Department of Education Info Hub (infohub.nyced.org). The data includes Math Test Results summarized by school name and economic status. I will also incorporate location data obtained from the National Center for Education Statistics (https://nces.ed.gov/ccd/elsi/expressTables.aspx).

I will identify a school that is centrally located to other low-income, under performing K-5 students. The idea being that both the students within the school chosen and the surrounding schools can attend this math-focused after school program.

Methodology

Exploratory data analysis was performed to determine if NYC school scores differed based on Economic (Dis)Advantage and location. Pandas was imported into python to perform these functions. Datasets were merged to create one dataframe to encompass the desired information (School Name, Address, Econ Status, Mean Score Scale, Grade, Year).

Geopandas was used to plot schools of interest to determine a centrally located area. Bronx schools categorized as Econ Disadv were plotted on a map of the Bronx.

Notebook with Code can be found at:
https://gist.github.com/4d53609755be4657f80398fe88ba8483

Results

The data confirms that public schools in NYC fits the literature's statement that test scores are impacted by economical advantage. Economically Disadvantage schools score almost 20 points less on the standardized math test than Not Economically Disadvantaged students across NYC.

Category	Mean Scale Score
Econ Disadv	345.932087
Not Econ Disadv	364.751642

To narrow down a location for an After-School Math Program, test scores were examined by borough. The Bronx scored at least 12 points lower than all other boroughs.

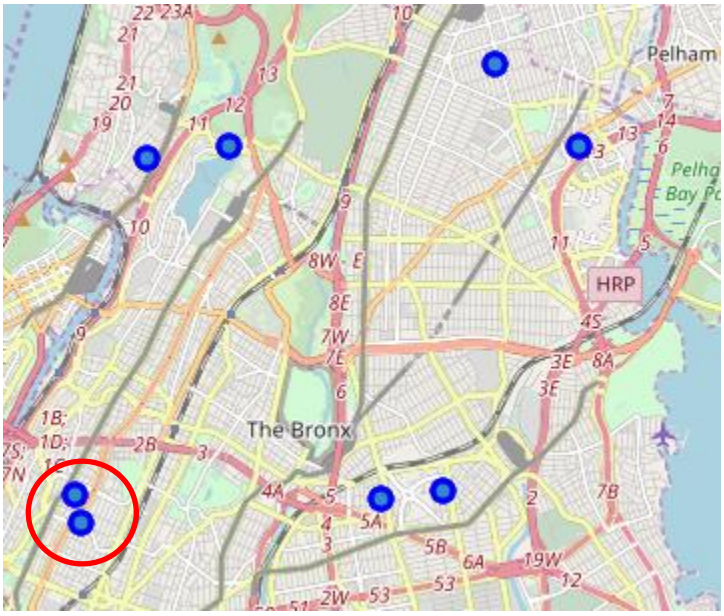
Borough	Mean Scale Score
BRONX	339.773573
BROOKLYN	353.677677
MANHATTAN	358.508799
QUEENS	359.232340
STATEN ISLAND	356.260452

It was confirmed that Economically Disadvantaged schools score less on standardized math teset than Economically Advantaged schools within the Bronx. Economically Disadvantaged schools scored less than Not Economically Disadvantaged schools by 12 points.

Category	Mean Scale Score
Econ Disadv	338.445680
Not Econ Disadv	350.498622

Therefore, the After-School Math Program will be at a school in the Bronx that is categorized as Economically Disadvantaged.

Two schools are in close proximity to each other, Lucero Elementary School and Sheridan Academy for Young Leaders.



Discussion

As the literature predicts, students in Economically Disadvantaged schools have the greatest need for an After-School Math Program. Students at Economically Disadvantaged schools located in the Bronx are in the most need within New York City.

By mapping Economically Disadvantaged Schools in the Bronx, the potential to reach more students than one school was found. Lucero Elementary School and Sheridan Academy for Young Leaders are very close to each other, which may allow for the transfer of students from one school to another after the main school day is over.

Sheridan Academy for Young Leaders has lower test scores between the two schools. Therefore, we will place the After-School Program there. Since it is nearby, it can also be a resource to students at Lucero Elementary School who wish to attend as well.

Conclusion

An After-School Math Program will be placed at Sheridan Academy for Young Leaders in the Bronx, with hope that students from Lucero Elementary School will attend as well giving the greatest impact to the community.