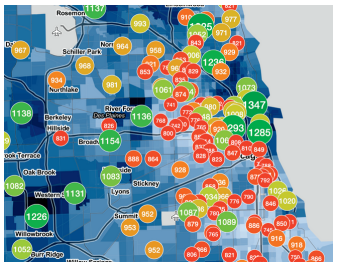
Public Transportation Access and Academic Performance in the Chicagoland Area

Sophie Maday, Lena Goodelle, GT Moy, Maya Marks-Strauss

ABSTRACT:

Does a lack of access to public transportation impact a student's ability to perform in formal education? Applying statistical data, we will be investigating the relationship between students' access to public transportation and their performance in the education system. Referencing SAT scores and graduation rates across Chicagoland High Schools, we will examine whether these metrics share a relationship with these students' ease of access to physical resources.

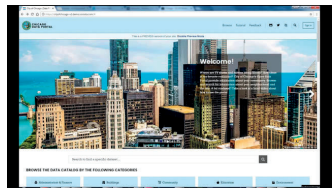


INTRODUCTION:

In this study, we will be delving into the impact of public transportation access on students' academic performance in Chicagoland High Schools. Through the use of statistical analysis, we aim to investigate the relationship between students' access to transportation and their educational outcomes. By analyzing data on SAT scores and graduation rates across high schools in Chicago, we hope to examine whether these metrics are influenced by students' physical access to transportation resources. The findings of our research provide valuable insights into the potential effects of transportation barriers on academic achievement, which can inform policy and decision-making in the education sector.

METHODS AND RESEARCH DESIGN:

In order to find a correlation between student performance in the Chicago high school system and transportation accessibility in the Chicagoland area, we chose 65 public high schools and began gathering data on student performance metrics, and identified CTA train stops in the area. We then calculated accessibility measures for each school based on their proximity to train stops and conducted statistical analysis to determine whether there is a significant relationship between transportation accessibility and student performance (measured by SAT scores and graduation rates). Finally, we interpreted the results as well as identified areas for future research and implications for education policy and practice.



Step 1: Acquire data sets via the Chicago Data Portal, including “CTA - System Information - List of 'L' Stops” and “Chicago Public Schools - School Profile Information SY1617” as well as average SAT score data through EdGap.com.

Step 2: Calculate distances between 65 public high schools in the Chicago area and the nearest CTA train stop.

- Import necessary libraries (pandas and geopy).
- Read the data from the two CSV files and store them as dataframes.
- Loop through each row in the high schools dataframe, and for each school, calculate the distance to the nearest public transportation location by looping through each row in the public transportation dataframe and comparing the distances (via “LONGITUDE” and “LATITUDE” variables).
- Add the calculated distances to the high school dataframe as a new column called "Distance".
- Sort the high schools dataframe by distance in descending order and print out the names of the high schools along with their distances in kilometers.

Step 3: Perform a multiple linear regression analysis using calculated distances.

- Import necessary libraries (pandas and statsmodels).
- Read the data from the CSV file and store it as a dataframe.
- Define the X and y variables for the regression. In this case, X is a dataframe with two predictor variables: "SAT_Score" and "Graduation_Rate", and y is a series containing the dependent variable "Distance_CTA_Train".
- Add a constant column to the X dataframe using the add_constant() function from statsmodels.
- Fit a linear regression model to the data using the OLS() function from statsmodels. This function takes the y variable as the first argument and the X variable as the second argument.
- Print out a summary of the model using the summary() method of the model object.

Step 4: Create a scatter plot of variables “Distance_CTA_Train” and “SAT_Score” and/or Graduation_Rate”.

- Import necessary libraries (pandas, matplotlib, and numpy).
- Read the data from the CSV file and store it as a dataframe.
- Extract the variables “Distance_CTA_Train” and “SAT_Score” or Graduation_Rate” for comparison.
- Create a scatter plot using the plot() function from matplotlib.
- The output of the script using the show() function will be a scatter plot with the distance to CTA Train Stop on the x-axis, SAT scores on the y-axis and each data point color-coded based on SAT scores.

Figure 1

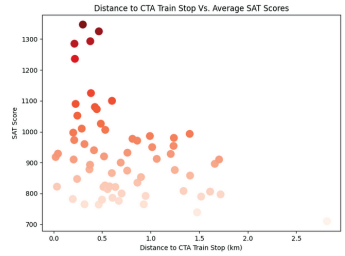
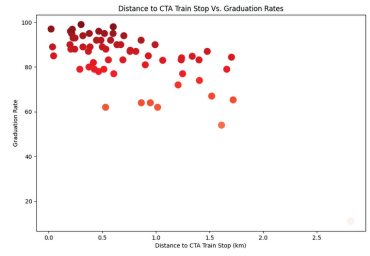


Figure 2



RESULTS:

The model explains 48.7% of the variance in the dependent variable, Distance_CTA_Train. This means that the independent variables included in the model (SAT_Score and Graduation_Rate) account for almost half of the variation in the distance between a public high school and the nearest CTA train station.

The negative coefficient for the variable Graduation_Rate (-0.0257) means that, on average, a one-unit increase in the Graduation_Rate is associated with a 0.0257 km decrease in Distance_CTA_Train, holding all other variables constant. This suggests that higher graduation rates are associated with shorter distances between a student's school and the nearest CTA train station.

The coefficient for SAT_Score is not significant at the 0.05 level, which means that there is no evidence of a significant linear relationship between SAT_Score and Distance_CTA_Train.

The intercept (represented by the constant of 3.1077) is the expected value of Distance_CTA_Train when all other independent variables are equal to zero. This means that if a student had a SAT_Score of zero and a Graduation_Rate of zero, the expected distance between their school and the nearest train station would be 3.1077 km.

It's important to keep in mind that our regression model is based on a specific sample of 65 observations and may not necessarily generalize to the population as a whole.

CONCLUSIONS:

Our finding that higher graduation rates are associated with shorter distances between a student's school and the nearest CTA train station suggests that public high schools with higher graduation rates tend to be located in areas that are more accessible to public transportation.

Policymakers could use this finding that higher graduation rates are associated with shorter distances to CTA train stations to target transportation infrastructure investments in areas where public schools have lower graduation rates. By improving public transportation access to schools in these areas, policymakers could help to improve student attendance, which could, in turn, improve graduation rates.

Some examples include increasing the number of CTA train stations, extending bus routes, or improving the frequency and reliability of public transportation in these areas. Policymakers can also work with public transportation providers to offer discounted or free transportation passes to students in order to alleviate the financial burden on families who cannot afford transportation costs. Policymakers can explore alternative transportation options such as carpooling, biking, or walking to school. They can work with community organizations, schools, and parents to promote these options and provide the necessary infrastructure such as bike lanes and pedestrian-friendly routes.

Second, the fact that there is no significant linear relationship between SAT scores and distance to the nearest train station suggests that access to transportation alone may not be sufficient enough to show improvements in academic performance. Policymakers could use this information to develop more comprehensive strategies that address both academic and transportation needs in underperforming schools.

Policymakers could also use our finding that the regression model is based on a specific sample of 65 observations to highlight the need for further research in this area. By investing in more comprehensive data collection and analysis, policymakers could gain a better understanding of the factors that are influenced by transportation access for students and develop more effective policies to address these issues.

WORKS CITED:

Chicago Public Schools. “Chicago Public Schools - School Profile Information SY1617.” Cityofchicago.org, 15 Sept. 2016, data.cityofchicago.org/Education/Chicago-Public-Schools-School-Profile-Information-/8i6r-et8s. Accessed 14 Apr. 2023.

Chicago Transit Authority. “CTA - System Information - List of ‘L’ Stops.” Cityofchicago.org, 4 Aug. 2011, data.cityofchicago.org/Transportation/CTA-System-Information-List-of-L-Stops/8pix-ypme. Accessed 14 Apr. 2023.

“EdGap.org | SAT/ACT College Readiness Map.” Edgap.org, 2016, www.edgap.org/#11/41.8340/-87.7320. Accessed 14 Apr. 2023.