Background
Like many other American cities, Chicago experiences widespread racial and income inequalities. Although the overall population is evenly racially divided, 18 neighborhoods in Chicago, with most in the South Side, are over 90% black. Recent studies have also found that Chicago’s white households are wealthier than the national average, but black and white students have substantially less wealth than the national average. A central theme in the literature is that economic reform in Chicago in the 1980s–1990s led to a drain in investment in lower-income, non-white neighborhoods, including social infrastructure such as schools and public transportation. In order to study the result of this trend, I will examine the association between high school racial and socioeconomic student demographics and respective student attendance and graduation outcomes across 49 neighborhood high schools in Chicago. I will also study the role of Chicago Transit Authority (CTA) Rail Stations in bridging the gap in educational outcomes across high schools with differing racial and socioeconomic demographics, and map locations for new CTA Rail Stations to optimally do so.

Methodology
For this analysis, I used High School Profile and Progress Report Excel data, and Attendance Zone shapefile data from Chicago Public Schools, available on the Chicago Data Portal. All data was imported into ArcMap as dbase tables or shapefiles, and a raster calculator was used to generate low-income and white student percentage fields in the Progress Report dbase table. Then, an attribute join was done with the High School Attendance boundaries shapefile using School ID as the unique identifier. Figures 1 and 2 were created then generating symbology based upon Low-income and White Student Proportion attributes, using a 5-class, Natural Jenks classification. Figure 3 was generated in Microsoft Excel. The mean Student Attendance and Graduation rates were obtained using statistics of these fields, after school zones were selected by income and racial categories using the Select by Attributes option in the Attribute Table.

Results
From Figure 3, and visually comparing Figures 1 and 2 with Figures 5 and 6, we observe that high schools with a higher proportion of low-income and non-white students experience lesser student attendance and 4-year graduation rates. The gap in student graduation is larger than the gap in student attendance. From Figures 5 & 6, we can conclude that the presence of 2 or more CTA Rail stations is associated with greater student attendance and graduation rates, even when compared among schools with similar racial or socioeconomic profiles. Finally, Figure 9 shows us that attendance zones where new CTA Rail Stations are most likely to bridge the educational gap are largely in the southern and western parts of Chicago. In particular, these consist of zones where no CTA Rail stations currently exist.

Conclusion
From my analysis, I can infer that there is a link between racial/socioeconomic profiles and educational outcomes in Chicago’s high schools. I can also support the hypothesis that the lack of equitable investment in CTA infrastructure across Chicago’s neighborhoods has contributed to gaps in educational outcomes. My findings are consistent with previous studies that shed light on disparities across neighborhoods with racial and socioeconomic differences in Chicago. Although past literature has advocated for construction of new CTA Rail Stations, this is the first analysis that maps the intersection of inequities in both transportation and educational infrastructure across Chicago’s neighborhoods.

This analysis was limited by a small sample size and little was done to control for confounding when measuring educational gaps between schools of different racial and socioeconomic profiles. For example, the gap in graduation rates was larger than that of attendance rates, likely due to time-variant socioeconomic factors. In addition, there was a very small count of majority-white high schools, making it difficult to create reliable conclusions about educational outcomes in these schools. Besides these limitations, the model is useful for policymakers to identify which neighborhoods in Chicago would most benefit from new rail stations. These findings may then be leveraged towards building policy that incentivizes CTA to invest in new stations in selected areas. In the future, a similar model should be created that ranks need for new CTA Rail stations where a CTA Rail Line already exists. This would be less costly for CTA to implement, and more likely to be applied. Finally, a cost-effective analysis of new CTA rail stations would be a great policy resource and should be built by comparing marginal utility of new rail stations with incremental costs of rail station construction and maintenance.

Acknowledgments