The Demographics Behind SAT Performance

An Analysis of New York City Public Schools

Lian Benz and Zoe Tzaivragos
May 23, 2025
Data Science for Social Impact - Mr. Jarcho

Intro

Background/Context

Standardized tests have played a crucial role in the US college admissions process for many decades, with the SAT being a widely recognized metric for assessing academic readiness. Originally designed to provide a uniform measure of student ability across different high schools, the SAT has increasingly been scrutinized for what it may reveal about broader social inequalities. Critics have argued that standardized test scores often reflect more than just academic preparedness. They also mirror patterns of wealth, family background, school quality, and access to resources.

In recent years, questions about the fairness and effectiveness of the SAT have grown louder, especially with rising concerns about equity in higher education. Colleges and universities are now re-evaluating the weight placed on test scores in admissions decisions, especially in light of evidence suggesting that scores are strongly associated with race, income, and parental education levels. Further, as some institutions opt for test-optional policies, researchers and educators are examining whether the SAT serves as a valid predictor of college success or if it reinforces structural inequalities.

New York City, as one of the most demographically diverse and educationally complex urban centers in the US, offers an exceptional context for exploring these questions. With large disparities in school funding, neighborhood wealth, and student support systems, NYC public schools reflect many of the broader challenges facing American students. By focusing on SAT scores from NYC public high schools in 2012 and analyzing them in relation to census tract-level demographic data, this study aims to uncover localized patterns that speak to national concerns.

Literature Review

Standardized testing, such as the SAT, has long played a significant role in US college admissions. However, its effectiveness and fairness as a predictor of academic success has sparked an ongoing debate. Existing research explores how SAT scores correlate with academic outcomes such as college GPA and retention, but also examines how demographic variables intersect with these results. This literature review surveys the current body of work on the predictive validity of the SAT across diverse student populations and different institutional settings. It also identifies critical gaps in the research, especially with long-term outcomes beyond college graduation, and the influence of recent policy changes. By summarizing these findings, this review lays the groundwork for our own investigation on how SAT scores might reflect or reinforce broader structural inequalities in education and socioeconomic opportunity.

A growing body of research has examined the fairness and predictive validity of the SAT, raising concerns about whether it serves as an equitable measure of academic potential. Sackett et al.

(2012) found that while SAT scores correlate with college GPA, socioeconomic status plays an important role in mediating this relationship, suggesting that the SAT might reflect privilege more than ability.[1] Zwick and Himelfarb (2011) further emphasize this point by showing that socioeconomic factors such as high school environment influence how predictive SAT scores and GPAs are for college performance.[2] Their findings support the claim that structural inequalities, rather than individual skill alone, are central to understanding disparities in student performance on the SAT.[3]

Several studies have also explored how specific demographic variables, such as race, gender, and family background, relate to SAT outcomes. Fleming and Garcia (1998) focused on differences in SAT predictive validity for Black students at historically Black colleges versus predominantly white institutions, finding that test performance was not uniformly predictive across these contexts.[4] Young and Fisler (2000) examined gender differences in SAT performance, identifying a persistent gap that favors males, particularly on the math section, and highlighted the role of social factors in this.[5] Similarly, Daly and Miller (1975) investigated how writing apprehension, influenced by gender and self-perception, negatively affected SAT performance and willingness to take advanced courses.[6] Although their study is dated, it emphasizes how internalized academic anxiety can be an underacknowledged contributor to performance disparities.[7]

Beyond test scores themselves, some research has focused on how SAT-related policies affect educational access. Hurwitz et al. (2017) studied a College Board policy that allowed low-income students to send SAT scores to more colleges for free. Their findings show that even minor changes in testing policy can meaningfully improve college application behavior, enrollment, and graduation rates

---

[1] Sackett, P. R., Kuncel, N. R., Beatty, A. S., Rigdon, J. L., Shen, W., & Kiger, T. B. (2012). The Role of Socioeconomic Status in SAT-Grade Relationships and in College Admissions Decisions. Psychological Science, 23(9), 1000–1007. http://www.jstor.org/stable/23260359

[2] Zwick, R., & Himelfarb, I. (2011). The Effect of High School Socioeconomic Status on the Predictive Validity of SAT Scores and High School Grade-Point Average. *Journal of Educational Measurement*, *48*(2), 101–121. http://www.jstor.org/stable/23018085

[3] Ibid.

[4] Fleming, J., & Garcia, N. (1998). Are Standardized Tests Fair to African Americans?: Predictive Validity of the SAT in Black and White Institutions. *The Journal of Higher Education*, *69*(5), 471–495. https://doi.org/10.2307/2649106

[5] Young, J. W., & Fisler, J. L. (2000). Sex Differences on the Sat: An Analysis of Demographic and Educational Variables. *Research in Higher Education*, *41*(3), 401–416. http://www.jstor.org/stable/40196399

[6] Daly, J. A., & Miller, M. D. (1975). Further Studies on Writing Apprehension: SAT Scores, Success Expectations, Willingness to Take Advanced Courses and Sex Differences. *Research in the Teaching of English*, *9*(3), 250–256. http://www.jstor.org/stable/40170633

[7] Ibid.

among disadvantaged students.[8] This research relates to the broader point that SAT-related outcomes are not only shaped by student characteristics but also by institutional policies that can either widen or narrow gaps in access.[9]

Together, these studies provide essential context for our own analysis, which focuses on New York City public schools and how census tract demographics correlate with SAT scores. The findings from Marsh et al. (2008) add nuance by exploring how high school GPA and standardized tests together predict college performance, although they also question the consistency of these measures given variations in school grading systems and student support.[10] Overall, these sources affirm the importance of investigating the SAT not just as a test of academic ability but as a window into how structural inequalities in race, income, education, and family life shape education opportunity. These are insights that our study seeks to localize and expand upon using NYC-specific data.

Essential Question & Hypothesis

In our study, the specific question we work to address is: To what extent do demographic factors (such as household income, family structure, and racial composition) correlate with 2012 SAT scores in New York City public high schools?

We hypothesize that schools located in census tracts with lower median household incomes, higher percentages of single-parent households, and higher proportions of Black and Latino residents will, on average, report lower SAT scores. On the other hand, schools in areas with higher median household incomes, lower percentages of single-parent households, and larger proportions of white and Asian residents are expected to have higher average SAT scores. We expect these patterns to reflect structural inequalities rather than differences in academic potential.

Methodology

Each year, teenagers in the US and worldwide take the SAT, a factor that can dictate how competitive students are for college placement. In fact, the College Board reports that in the high

---

[8] Hurwitz, M., Mbekeani, P. P., Nipson, M. M., & Page, L. C. (2017). Surprising Ripple Effects: How Changing the SAT Score-Sending Policy for Low-Income Students Impacts College Access and Success. *Educational Evaluation and Policy Analysis*, *39*(1), 77–103. http://www.jstor.org/stable/44984571

[9] Ibid.

[10] Marsh, C. M., Vandehey, M. A., & Diekhoff, G. M. (2008). A Comparison of an Introductory Course to SAT/ACT Scores in Predicting Student Performance. *The Journal of General Education*, *57*(4), 244–255. http://www.jstor.org/stable/27798113

school class of 2024, over 1.97 million students took the SAT.[11] Despite the intentions of having the SAT be standardized across the board, countless researchers have noted that average SAT scores vary across demographic factors such as race, gender, and socioeconomic status. Our study aims to investigate the extent to which demographics impact SAT scores, as well as differences in how much of an impact different demographic aspects have.

Many academics have worked on tackling this question in the past, and we'll add to this growing body of research by specifically analyzing SAT scores from New York City public high schools in 2012. Through our research question—How do demographic factors of New York City census tracts such as income, racial composition, parental education, and family structure correlate with 2012 SAT scores from public high schools in those same census tracts?—we attempt to examine the relationship between said demographic variables and SAT scores but on a more localized scale. By focusing on a smaller group of students, we're building a jumping-off point for more generalizable research to follow. Additionally, relationships between SAT scores and these demographic factors that our study may reveal could potentially be analyzed in a deeper cultural context, given the specific focus on New York City.

In order to explore the relationship between SAT scores and demographic factors such as income, racial composition, parental education, and family structure, we'll observe and analyze various bivariate and multivariate regression models. Our specific visualizations, coded in R Studio, will include scatter plots, violin plots, density plots, and more.

For this study, we merged three separate datasets using R Studio. Our primary data set, which we found on Kaggle, was released by Jack Murray three years ago and titled "SAT scores, 5+ datasets, project - (intermediate)."[12] This data contains 2012 SAT scores for all New York City public high schools. Each school is identified by name and DBN (District Borough Number) and includes average scores for each section: critical reading, math, and writing. Our second dataset, from the Opportunity Atlas, which was developed by Raj Chetty and other professors from Harvard and Brown University, provides demographic data at the census tract level for New York City. Variables include median household income, percentage of those below the poverty line, percentage of single-parent households, and racial composition. Our third dataset, used to merge the prior two datasets, is the 2014-2015 NYC High School Directory, which provides the census tract number for each school. Using DBN as the common link, we matched the SAT score dataset with the directory, and then used census tract

---

[11] College Board. (2024, September 24). *SAT Participation Continues To Grow As The SAT Suite Successfully Completes Its Transition To Digital Testing*. https://newsroom.collegeboard.org/sat-participation-continues-grow-sat-suite-successfully-completes-its-transition-digital-testing

[12] Murray, J. (n.d.). *SAT scores: 5 datasets project (intermediate)* [Computer software]. Kaggle. https://www.kaggle.com/code/jackmurray94/sat-scores-5-datasets-project-intermediate.

numbers to merge in the demographic data from the Opportunity Atlas. With these three datasets, our population of interest is public high school students in New York City who chose to take the SAT, and our dataset is a comprehensive reflection of that population.

All data cleaning, merging, and variable creation was conducted in R Studio. SAT section scores were numeric and added together to create a total SAT score per school. Demographic percentages (ex. percent white, percent Black, percent single-parent households) and medians (ex. household income) were kept as continuous variables.

Our independent variables are the demographics of New York City, gathered from the Opportunity Atlas, such as income, racial composition, parental education, and family structure for each individual census tract. Our dataset measures these variables with median and percentage calculations from the Opportunity Atlas.

Our dependent variables are the average SAT scores from each school. Specifically, we're considering the average scores for each section: critical reading, math, and writing. Our dataset measures these variables by simply providing the average scores from each school, with the highest possible score for each section being 800.

The control variables in our study include the time period and the schools from which all the data is gathered. Having the time period as a control is relevant, given that the College Board changes the SAT from time to time, which can impact students' scores. Focusing on the same public schools in New York City is also important, given potential differences in students' ability to succeed on the SAT depending on their state or school district.

Overall, we predict that schools with a higher Black and Latino student enrollment will, on average, receive lower SAT scores in comparison to schools with predominantly white and Asian populations. We also predict that schools in census tracts with a higher percentage of single-parent households will generally receive lower SAT scores than schools in census tracts with fewer single-parent households. Finally, we predict that students at schools in census tracts with lower median household incomes or higher poor shares will receive lower SAT scores than students from schools in census tracts with the opposite.

Our predictions are not necessarily a reflection of the students' inherent capabilities, but instead their access to resources and support systems. For example, areas with high Black and Latino populations tend to have less access to quality educational resources and proper test preparation. In comparison, areas with a predominantly white and Asian population have more overall access to resources, specifically quality education, and therefore proper test preparation. Additionally, a student who is a part of a single-parent household may not have the same support system as one coming from a two-parent household. This lack of support can limit resources and lower mental health as well. Finally, students from low-income census tracts have limited access to educational resources, especially in

comparison to students in wealthier census tracts. Again, our predictions are primarily based on access to resources and not the students' capabilities.

Throughout our study, we will use various bivariate and multivariable regression models to analyze and identify patterns between demographics and SAT scores. The primary method of analysis for our study is a multivariate linear regression, which is appropriate for examining how several independent demographic variables simultaneously relate to a continuous dependent variable, which, in our case, is average SAT score. By using this method, we can isolate the unique contribution of each factor while holding the others constant. This will help us answer our research question of which demographics impact SAT scores and how much of an impact different demographic aspects have. We also use bivariate regression models and visualizations, such as scatter plots, violin plots, and density plots, to examine individual relationships and identify patterns or outliers prior to building our multivariate models. These preliminary visuals help with contextualizing findings from the regression and revealing non-linear trends or demographic thresholds. Our regression models are especially appropriate for this analysis because they allow us to quantify disparities in SAT performance across census tracts. They also allow us to sort out overlapping effects of different demographic variables, which is an important step given the often strong correlation between race, income, and family structure in many communities.

In general, there aren't any serious ethical problems related to our research regarding both data collection and the way the data is used. After all, even though this study gathers sensitive information on students (such as their SAT scores, family income, race, family circumstance, etc.), all data is provided in the form of school or census tract averages, with nothing singling out specific students. Additionally, all of the data is entirely anonymous, with no connections to the students.

However, when considering implications for those taking the SAT, it's important to remember that many researchers and professors argue that the SAT disadvantages students from lower socioeconomic backgrounds and students who are Black or Hispanic. Given how important the SAT can be in determining college admissions results and, therefore, potential opportunities that come with certain college experiences, being disadvantaged on the SAT can have negative long-term effects. With this, it's crucial to assess how ethical the SAT is when considering students of many different backgrounds and the potential for it to be an unfair test.

A limitation of our study is that we are determining the demographics of students based on the census tracts in which their school falls. However, not every student lives within the same census tract as their school and might even live far away, which allows for some potential discrepancies. Another possible limitation is that the SAT has changed drastically over the years, especially given the recent shift from paper to electronic test-taking. The data we're using does not account for these new changes, given that it's around a decade old, which could mean our results might not perfectly represent potential results drawn from more recent data.

To sum up, our study investigates how demographic factors such as income, racial composition, and family structure correlate with 2012 SAT scores in New York City public schools. By focusing on data on SAT scores and using census tracts to estimate school-level demographics, our research will provide a localized analysis of potential disparities in SAT performance. This approach allows us to contribute to a broader understanding of how systemic inequality can manifest in standardized testing outcomes.

Ultimately, this research is significant even beyond New York City since it'll highlight different structural factors that can shape academic opportunities. Additionally, it'll offer insight into the already ongoing debate around equity in college admissions. By understanding how demographics impact SAT scores, we can better inform educational policies and potential reforms.

Results

To examine how demographic factors influence SAT performance in New York City public high schools, we conducted many different statistical analyses using data merged from school-level SAT scores and census tract demographics. We began by analyzing correlations between individual SAT score sections to determine whether a total SAT score could effectively represent overall performance on the test. After confirming strong positive correlations between the three sections, we created a total score variable and used it to explore bivariate relationships with key demographic variables such as household income, poverty rate, family structure, and racial composition. These initial visualizations revealed consistent patterns, namely that schools in wealthier, majority white or Asian census tracts tended to have higher average SAT scores, while schools in poorer, majority Black or Hispanic tracts scored lower. To further test these patterns, we used a multivariate regression to evaluate the specific contribution of each demographic factor while controlling for the others. This section presents the findings from both our bivariate and multivariate analyses, with scatter plots, violin plots, density plots, and more.

The following table outlines all of the variables we included in our analysis. We classify each as a response or explanatory variable based on its role in the study. Definitions are included to clarify how each variable was measured or derived from our merged dataset. This classification helps structure our multivariate regression models and supports our interpretation of how different demographic factors relate to SAT scores.

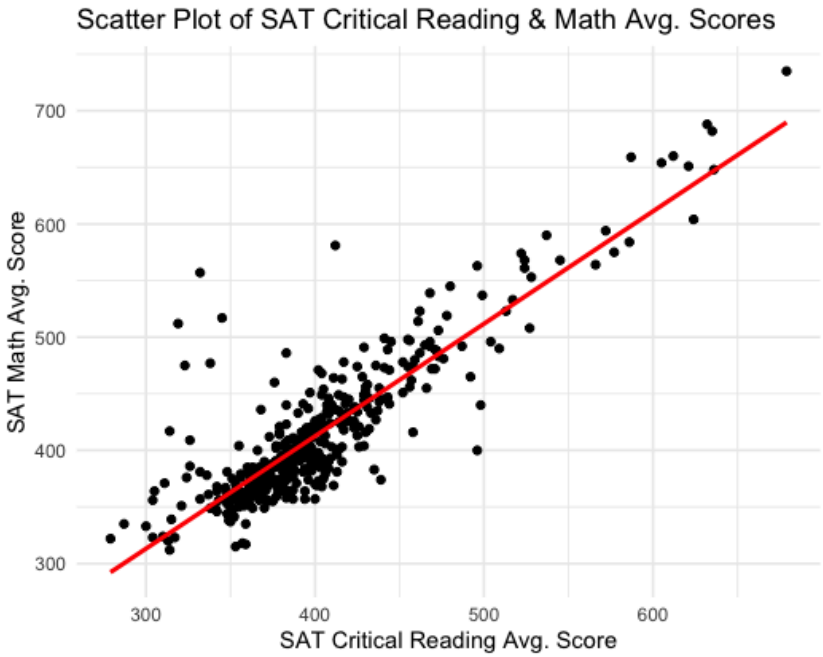
Variable Name	**Type**	**Definition**	**Role**
SAT Critical Reading	Quantitative	Average SAT critical reading section	Response

Score (2012)		score of students in a given school	(Dependent)
SAT Math Score (2012)	Quantitative	Average SAT math section score of students in a given school	Response (Dependent)
SAT Writing Score (2012)	Quantitative	Average SAT writing section score of students in a given school	Response (Dependent)
Total SAT Score (2012)	Quantitative	Average SAT score of students in a given school	Response (Dependent)
White Share (2010)	Proportion	Proportion of the census tract identifying as white	Explanatory (Independent)
Black Share (2010)	Proportion	Proportion of the census tract identifying as Black	Explanatory (Independent)
Hispanic Share (2010)	Proportion	Proportion of the census tract identifying as Hispanic	Explanatory (Independent)
Asian Share (2010)	Proportion	Proportion of the census tract identifying as Asian	Explanatory (Independent)
Non-White Share (2010)	Proportion	Proportion of the census tract identifying as non-white	Explanatory (Independent)
Median Household Income (2016)	Quantitative	Median household income of the census tract (in US dollars)	Explanatory (Independent)
Poor Share (2010)	Proportion	Proportion of the census tract living below the poverty line	Explanatory (Independent)
Single-Parent Household Share (2010)	Proportion	Proportion of households in the census tract with only one parent	Explanatory (Independent)
Majority white	Categorical	Census tracts with a white share greater than 50%	Explanatory (Independent)
Majority Black	Categorical	Census tracts with a Black share greater than 50%	Explanatory (Independent)
Majority Hispanic	Categorical	Census tracts with a Hispanic share greater than 50%	Explanatory (Independent)

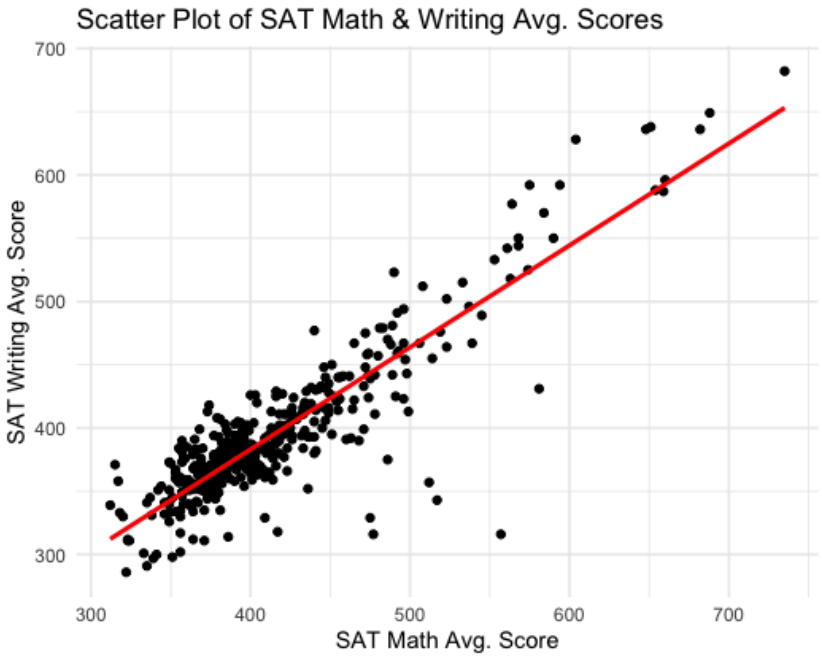
Majority Asian	Categorical	Census tracts with an Asian share greater than 50%	Explanatory (Independent)

Our initial dataset from Kaggle had different columns for scores from each section of the SAT (Critical Reading, Math, and Writing). The scores for each of the three columns ranged from zero to 800, with 2,400 being the highest possible total score. When analyzing our data, we started by exploring the relationship between the scores of the three sections, working to understand if high scores in one section were connected to high scores in the other two sections, and vice versa with low scores. We hypothesized that if a school had a higher average score for one section, it'd also have similarly high average scores for the other sections, with the same being true for a lower average score. If this hypothesis were correct, we'd create a new column for total SAT scores, combining the scores from each section, and use just that variable when comparing with the various demographic factors, given that there wouldn't be a notable difference between a school's scores on the three sections.

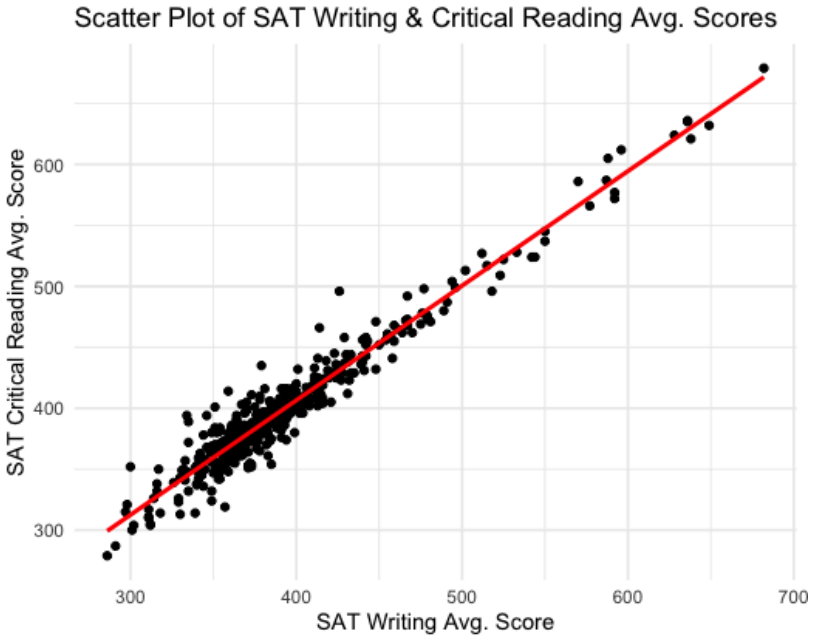
To explore the relationship between the scores of the three sections, we coded three scatter plots with lines of best fit and calculated correlation coefficients, which are shown below.



Correlation coefficient (SAT Critical Reading & Math Avg. Scores): 0.8724116



Correlation coefficient (SAT Math & Writing Avg. Scores): 0.8884561



Correlation coefficient (SAT Writing & Critical Reading Avg. Scores): 0.970342

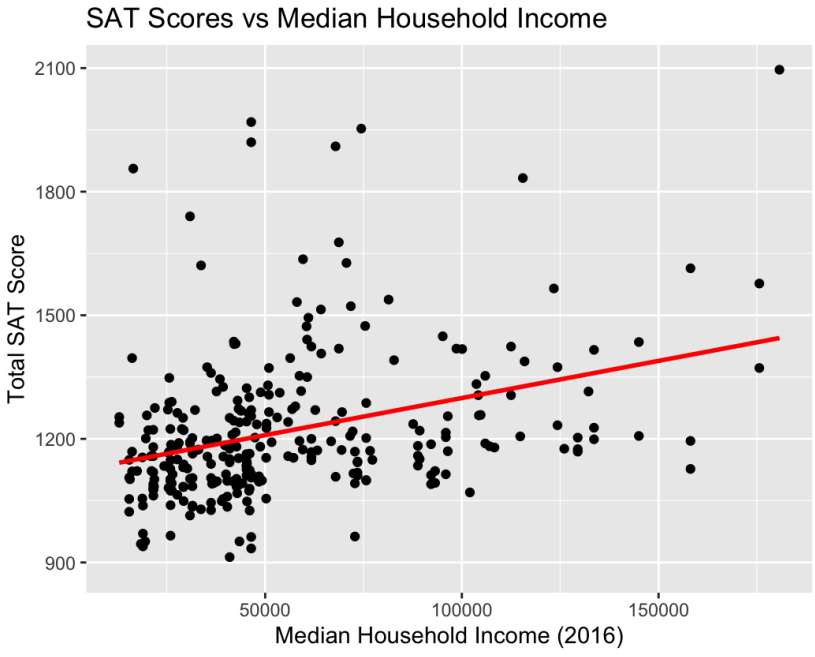
To determine whether a single total SAT score could effectively represent performance across all three sections (Critical Reading, Math, and Writing), we first examined the relationships between

the individual section scores. The correlation coefficients between each pair of SAT sections were all very high: 0.87 for Critical Reading and Math, 0.89 for Math and Writing, and an especially strong 0.97 for Writing and Critical Reading. These values, along with the accompanying visualizations, indicate strong positive linear relationships, suggesting that schools with higher average scores in one SAT section tend to also perform well in the other sections.

The notably high correlation between Critical Reading and Writing suggests that these two sections are especially closely connected, likely as a result of overlapping skill sets associated with language proficiency. While the math section is slightly less correlated with the other two sections, there is still a significant association. These results support our initial hypothesis and justify the decision to create a new column for total SAT score, given that it serves as a reliable measure of overall academic performance across all three sections. By using this total score, we focus our analysis and avoid being redundant, while also not sacrificing any meaningful variations in the data.

Next, we individually compared each school's average total SAT score with each demographic variable on the census tract level. The demographic variables we analyzed were median household income, poor share, single-parent share, and racial composition. Racial composition specifically included white share, Black share, Hispanic share, Asian share, and non-white share. We explored these relationships through scatter plots with best-fit lines and calculated correlation coefficients, which are shown below. For each of the scatter plots, Total SAT Score is the dependent variable on the y-axis, and the various demographic factors are the independent variables on the x-axis.

Total SAT Score vs. Median Household Income

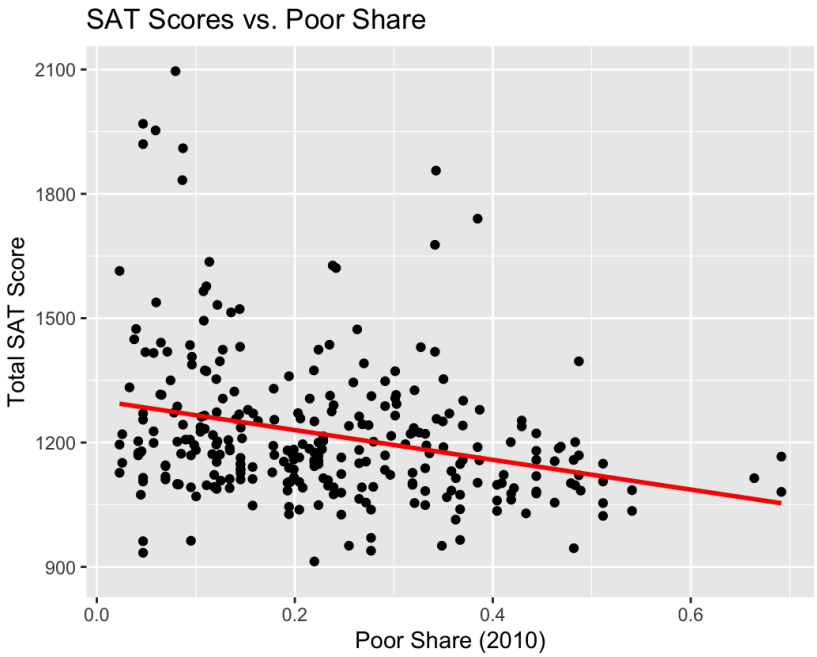


Hypothesis: As median household income increases in census tracts, average total SAT scores from schools in those census tracts will also increase. There's going to be a positive linear correlation between the two variables.

Correlation coefficient: 0.3408323

Analysis: This scatter plot shows a moderate positive relationship between median household income and total SAT scores (r ≈ 0.34). The upward trend in the best-fit line suggests that, on average, schools in wealthier census tracts report higher SAT scores. While this relationship isn't extremely strong, it supports the idea that income-related advantages, such as access to test prep, academic resources, and better-funded learning environments, positively influence standardized test scores.

Total SAT Score vs. Poor Share

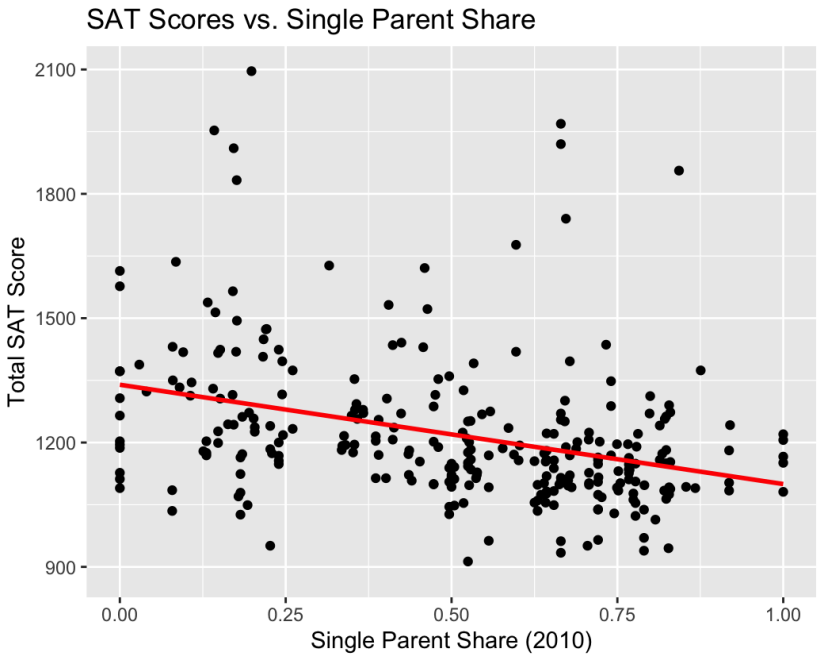


Hypothesis: As the poor share increases in census tracts, average total SAT scores from schools in those census tracts will decrease. There's going to be a negative linear correlation between the two variables.

Correlation coefficient: -0.2792302

Analysis: There is a moderate negative correlation (r ≈ -0.28) between the share of individuals living in poverty and SAT scores. This suggests that schools in poorer census tracts tend to have a lower average SAT performance. The inverse relationship aligns with existing research showing that economic hardship can stunt academic achievement through limited access to resources, elevated stress levels, and under-funded schools.

Total SAT Score vs. Single Parent Share

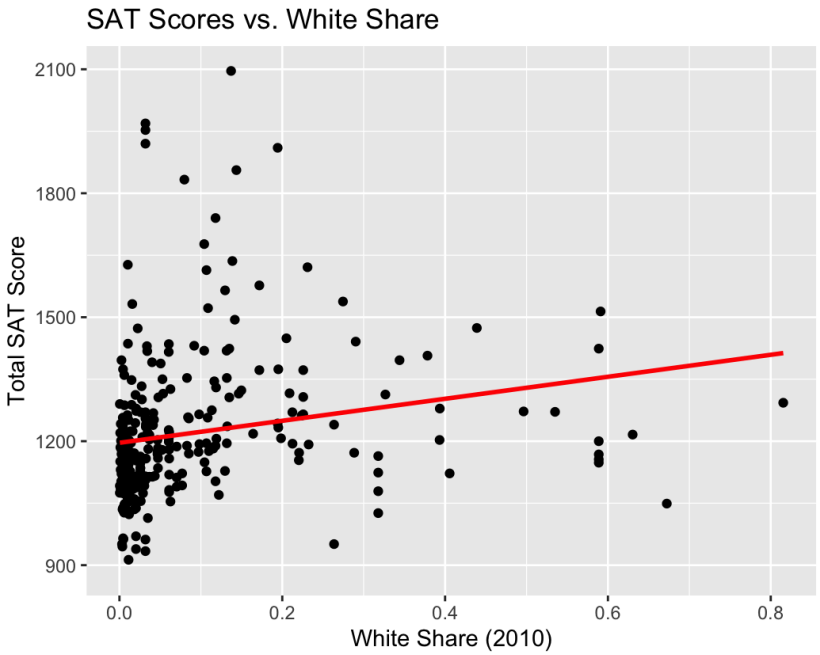


Hypothesis: As the single parent share increases in census tracts, average total SAT scores from schools in those census tracts will decrease. There's going to be a negative linear correlation between the two variables.

Correlation coefficient: -0.3445438

Analysis: This scatter plot reveals the *strongest* negative correlation of all the demographic scatter plots ($r \approx -0.34$), indicating that higher shares of single-parent households in a census tract are associated with lower average SAT scores. This may reflect disparities in support systems, beneficial time spent with parents, and financial stability, all of which are factors that can significantly impact a student's academic preparation and performance.

Total SAT Score vs. White Share

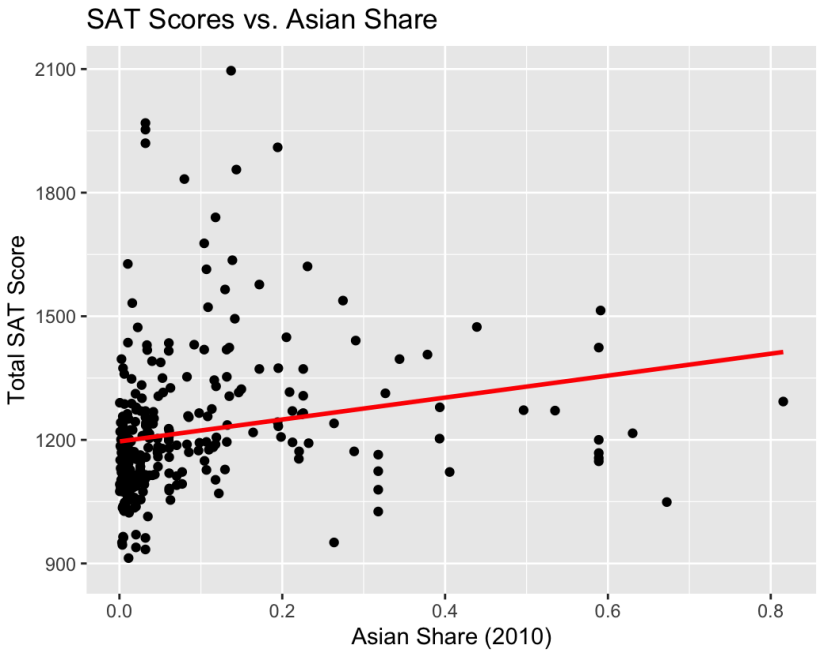


Hypothesis: As the percentage of white individuals increases in census tracts, average total SAT scores from schools in those census tracts will also increase. There's going to be a positive linear correlation between the two variables.

Correlation coefficient: 0.3663269

Analysis: The moderate positive correlation ($r \approx 0.37$) between the white population share and SAT score suggests that schools in predominantly white census tracts tend to perform better on the SAT. This may reflect broader systemic advantages commonly found in majority-white areas, such as higher levels of funding, greater access to test prep services, and stronger general education support.

Total SAT Score vs. Asian Share

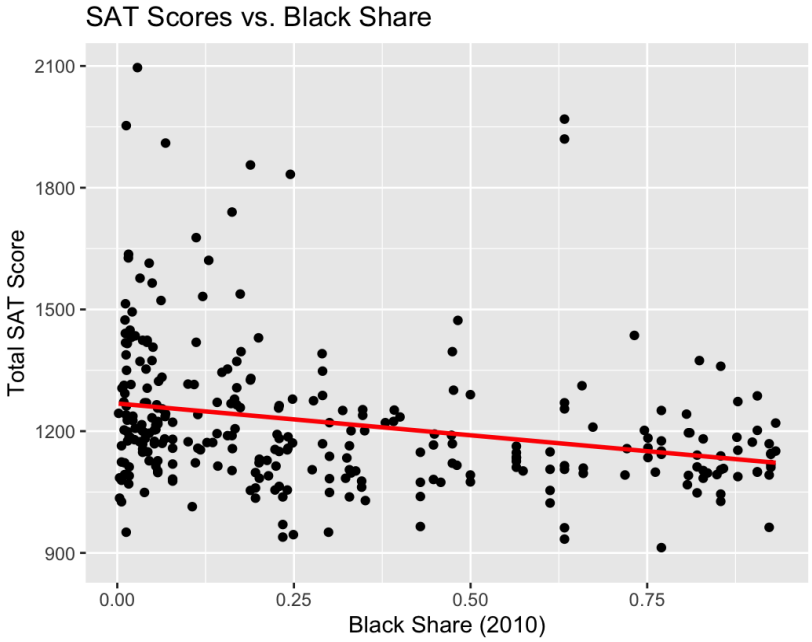


Hypothesis: As the percentage of Asian individuals increases in census tracts, average total SAT scores from schools in those census tracts will also increase. There's going to be a positive linear correlation between the two variables.

Correlation coefficient: 0.202618

Analysis: This scatter plot shows a weaker positive correlation (r ≈ 0.2), indicating that while schools in census tracts with more Asian residents tend to score slightly higher on the SAT, the relationship is less pronounced than with the white share. The relatively modest correlation could be due to greater variation in socioeconomic status across Asian communities in New York City, which could impact educational resources and test preparation strategies.

Total SAT Score vs. Black Share

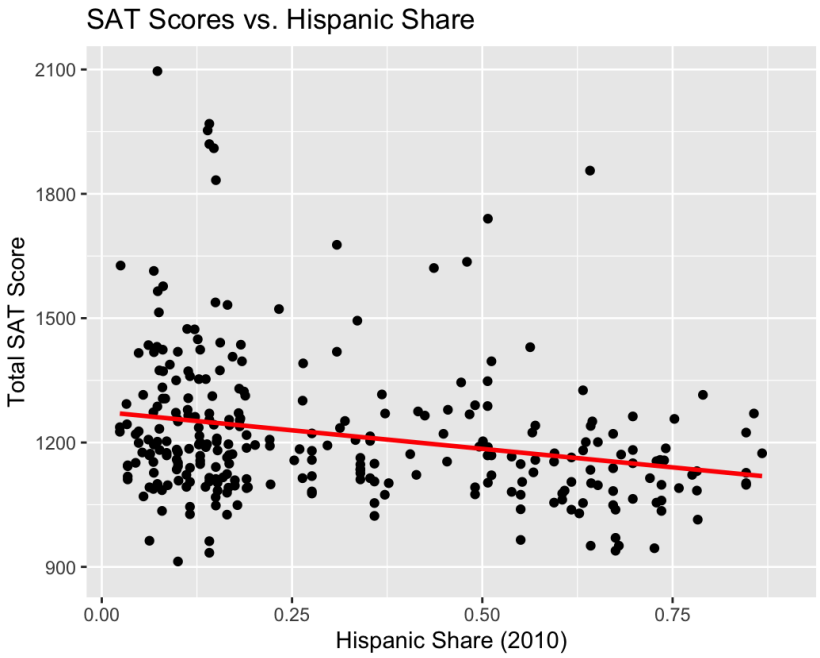


Hypothesis: As the percentage of Black individuals increases in census tracts, average total SAT scores from schools in those census tracts will decrease. There's going to be a negative linear correlation between the two variables.

Correlation coefficient: -0.2630636

Analysis: There is a moderate negative correlation (r ≈ -0.26), showing that schools located in census tracts with larger Black populations tend to have lower average SAT scores. This may reflect systemic disadvantages that disproportionately affect Black communities, such as underfunding in schools, residential segregation, and barriers to test prep access.

Total SAT Score vs. Hispanic Share

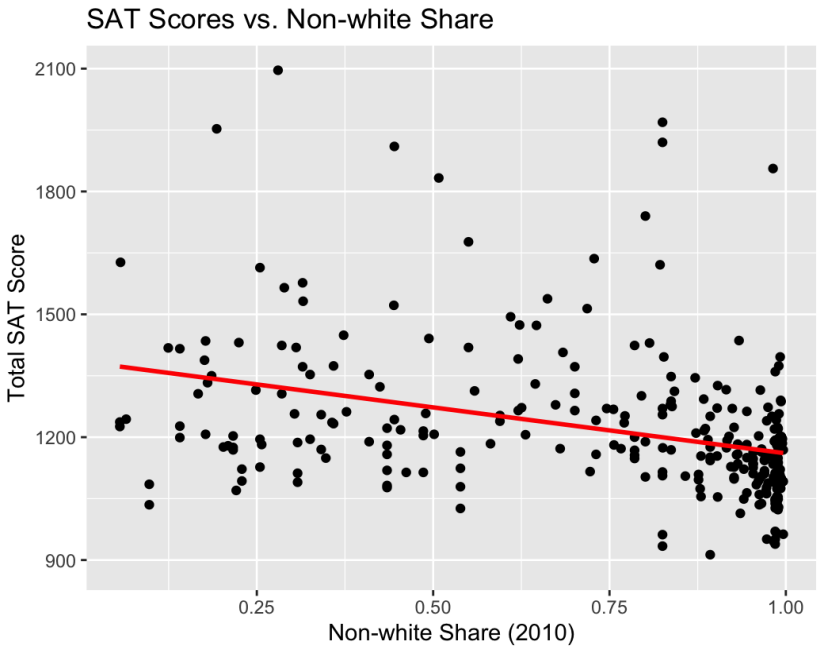


Hypothesis: As the percentage of Hispanic individuals increases in census tracts, average total SAT scores from schools in those census tracts will decrease. There's going to be a negative linear correlation between the two variables.

Correlation coefficient: -0.2415261

Analysis: The negative correlation ($r \approx -0.24$) between Hispanic share and SAT score follows similar trends to those seen in the Black share analysis. Schools in census tracts with larger Hispanic populations tend to have lower average SAT scores. This pattern likely reflects barriers that can affect Hispanic students' access to high-quality instruction and resources, such as linguistic obstacles, immigration-related stressors, and educational inequalities.

Total SAT Score vs. Non-White Share

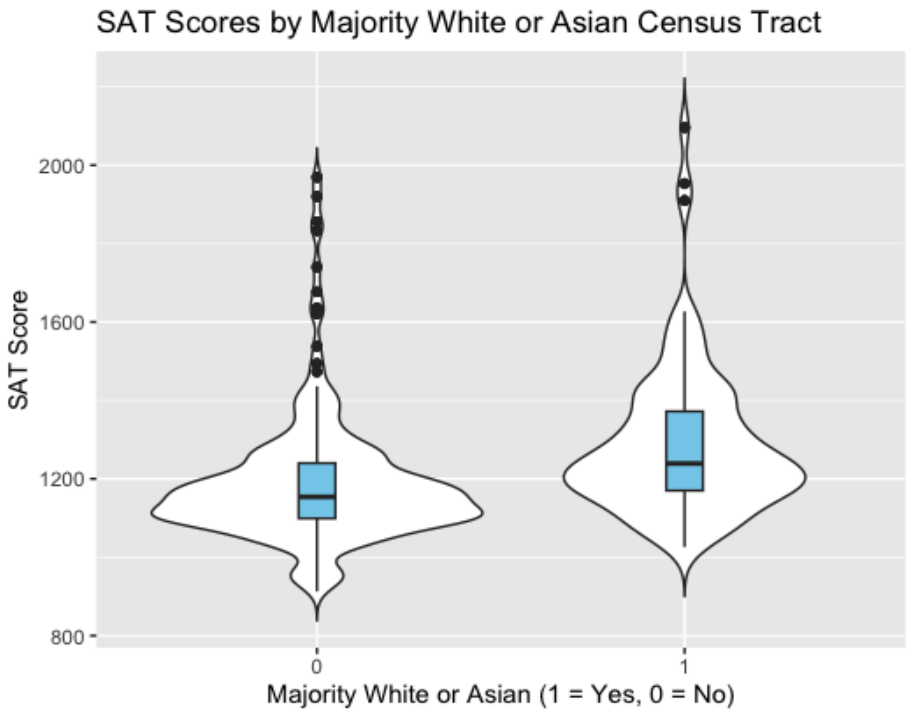


Hypothesis: As the percentage of non-white individuals increases in census tracts, average total SAT scores from schools in those census tracts will decrease. There's going to be a negative linear correlation between the two variables.

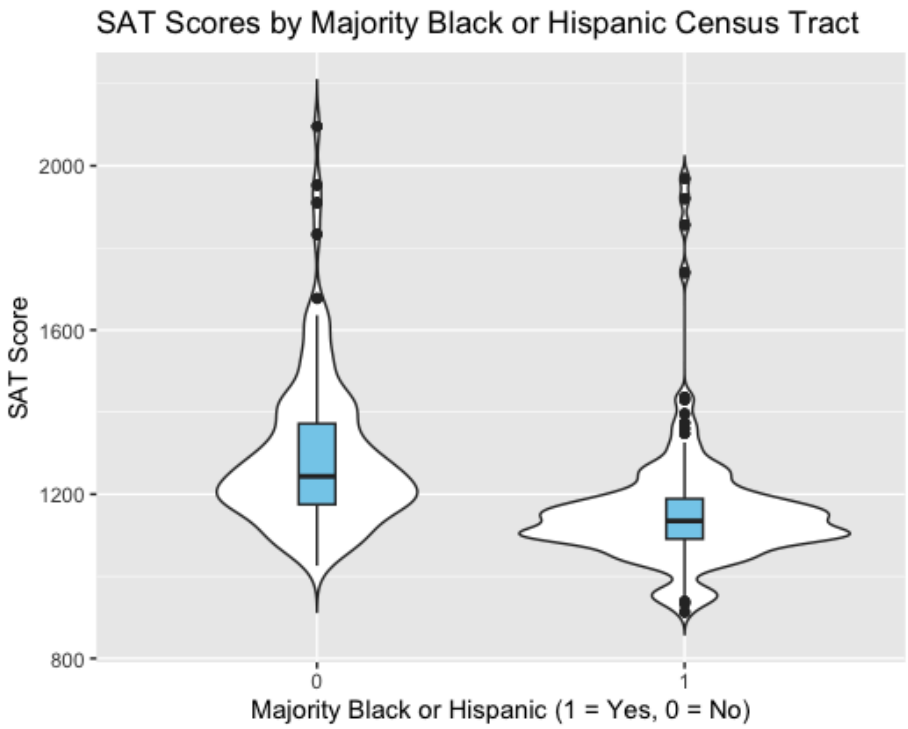
Correlation coefficient: -0.3606265

Analysis: This relationship ($r \approx -0.36$) shows a moderate-to-strong negative correlation between non-white population share and SAT scores. As the non-white share in census tracts increases, SAT scores from those tracts tend to decrease. This scatter plot speaks to the broader pattern of racial inequality in academic opportunity, likely tied to structural racism, school segregation, and unequal funding across schools.

When considering racial composition's impact on SAT performance, our analysis thus far has shown that census tracts with a higher white and Asian share tended to have higher average SAT scores. In comparison, census tracts with a higher Black and Hispanic share tended to have lower average SAT scores. With this, we used R Studio to filter for schools with a majority population of each race. Then, we combined schools with a majority white population and a majority Asian population. We also combined schools with a majority Black population and a majority Hispanic population. After doing so, 101 schools had a majority white or Asian population, and 167 had a majority Black or Hispanic population. With the new categorizations of majority white or Asian schools and majority Black or Hispanic schools, we coded violin plots, shown below, to clearly visualize differences in SAT outcomes based on race. We hypothesized that these plots would show higher average SAT scores for schools in majority white or Asian census tracts and lower average SAT scores for schools in majority Black or Hispanic census tracts.

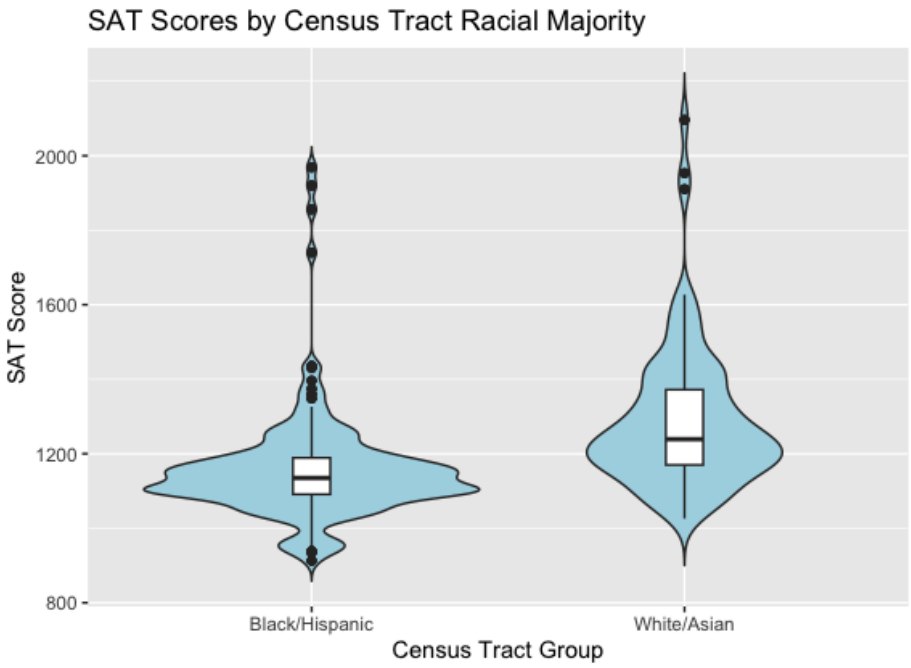


The first violin plot illustrates SAT score distributions for schools in census tracts with a majority white or Asian population (coded as 1) versus those that are not (coded as 0). Schools in majority white or Asian areas show a broader distribution of SAT scores with a higher median around 1300-1350, and several schools reaching scores above 1600, even approaching 2000. The overall range for these schools spans roughly from 1000 to over 2000, indicating both high performance and variability. In contrast, schools in non-majority white or Asian tracts have a lower median near 1150-1200 and a tighter spread, with fewer outliers on the upper end and most scores concentrated between 1000 and 1400.



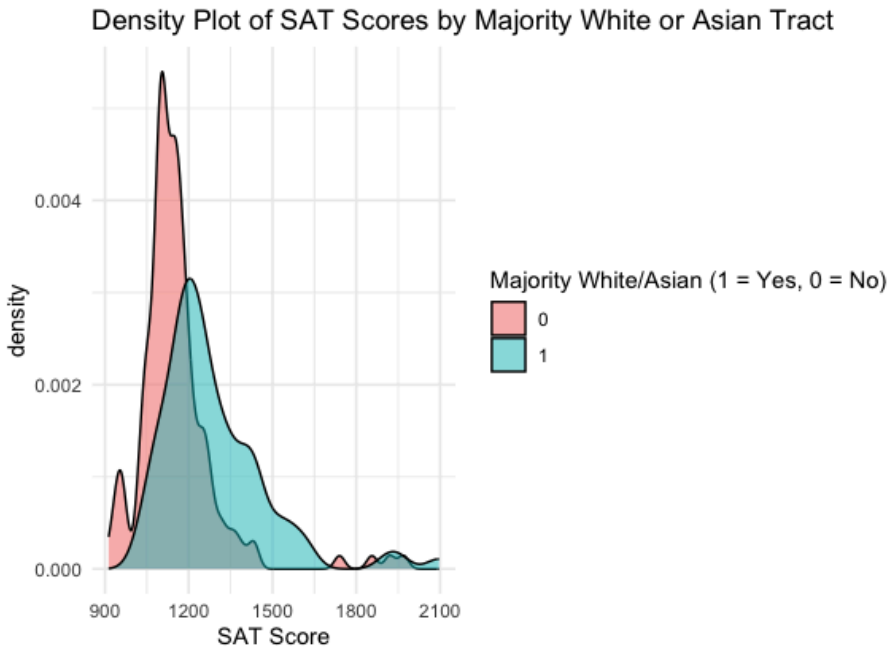
The second violin plot displays SAT score distributions for schools in census tracts with a majority Black or Hispanic population (coded as 1) compared to those that are not (coded as 0). Schools in majority Black or Hispanic areas show a more compressed distribution, with a median around 1100-1150 and most scores falling between 950 and 1300. Very few schools exceed 1400, and high-end outliers are sparse. In contrast, schools in non-majority Black or Hispanic tracts exhibit greater variability and a higher central tendency, with a median closer to 1250, and a long tail stretching beyond 2000, suggesting more schools achieving high scores.

Taken together, these two violin plots vividly illustrate the correlation between racial demographics and SAT performance at a school level. Schools in majority white or Asian census tracts not only tend to have higher average SAT scores but also display more variability on the upper end, suggesting that students in these areas may have greater access to advanced academic resources and test preparation. On the other hand, schools in majority Black or Hispanic census tracts show a more compressed distribution, with lower average scores and fewer outliers on the high end. To display these differences even more clearly, we created a third violin plot showing SAT scores for schools in majority Black or Hispanic census tracts on the left side and SAT scores for schools in majority white or Asian census tracts on the right side. This third plot fully supported the same conclusions we arrived at from the first two violin plots.

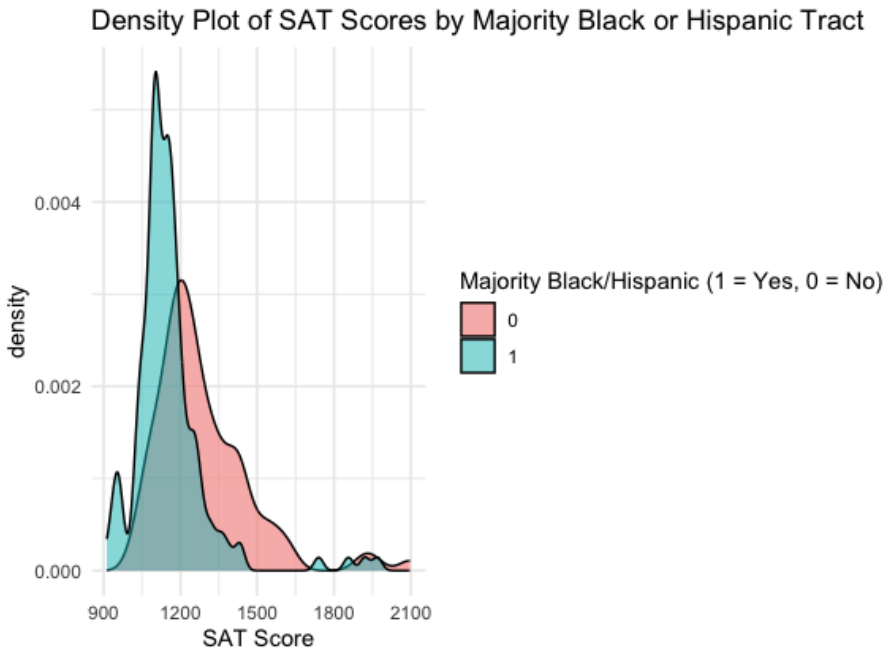


These patterns support the hypothesis that racial composition, which is intertwined with systemic factors such as socioeconomic disparities, school funding, and neighborhood segregation, has a substantial impact on education outcomes like SAT scores. The contrasting distributions across the plots highlight and strengthen the conclusion that majority white or Asian environments are associated with relative advantage, while majority Black or Hispanic environments are associated with disadvantage in terms of SAT performance. These visualizations help display the structural inequalities embedded in educational opportunities and outcomes.

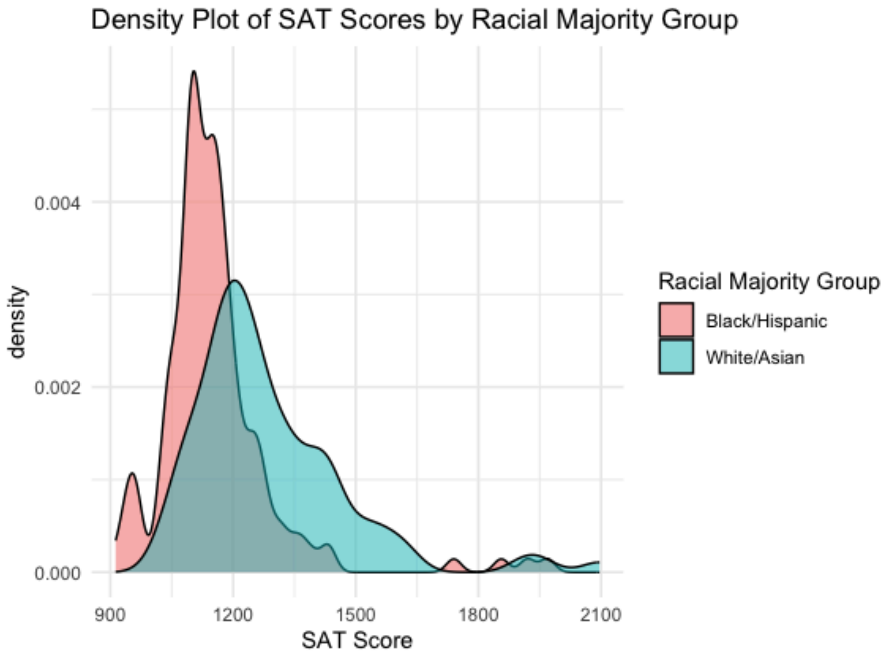
To further illustrate the impact of racial composition on school-wide SAT outcomes, we also coded three density plots in R Studio. As with the violin plots, we hypothesized that these density plots would show higher average SAT scores for schools in majority white or Asian census tracts and lower average SAT scores for schools in majority Black or Hispanic census tracts.



The first density plot shows SAT score distributions for schools in census tracts where the majority population is white or Asian (coded as 1) compared to those where it is not (coded as 0). The blue curve (majority white or Asian) is distinctly shifted rightward, peaking around 1250-1300. It also has a broader spread and maintains visible density even past 1500, with some schools reaching toward 1800-2000. This indicates greater variation and higher potential performance. Meanwhile, the pink curve (non-majority white or Asian tracts) peaks closer to 1100-1150, with density falling sharply beyond 1300, reflecting a more compressed score range. These findings reinforce the hypothesis that schools in majority white or Asian tracts tend to achieve higher average SAT scores with wider score distributions. At the same time, those in other areas are more concentrated in lower score ranges.



The second density plot illustrates the distribution of SAT scores for schools in census tracts where the majority population is Black or Hispanic (coded as 1) compared to those where it is not (coded as 0). The blue curve, representing majority Black or Hispanic tracts, peaks sharply around 1050-1100, indicating that most schools in these areas tend to have SAT averages within this lower score range. The distribution is narrow, showing limited variation and a rapid decline in density beyond 1250, with very few scores reaching above 1400. In contrast, the pink curve (non-majority Black or Hispanic tracts) is flatter and broader, peaking slightly higher around 1150-1200 and spreading more evenly across scores ranging from 1000 to over 1600. This suggests that while there are also lower scores in non-majority Black or Hispanic areas, these schools are more likely to reach mid and high range SAT scores compared to those in majority Black or Hispanic areas.



This third density plot overlays the SAT score distributions of schools in majority white or Asian tracts with those in majority Black or Hispanic tracts, providing a direct comparison between the two racial majority groupings. The pink curve (Black/Hispanic) peaks steeply between 1000 and 1100, with scores rarely exceeding 1300, indicating that most schools in these communities are clustered on the lower end of the SAT scale. In contrast, the blue curve (white/Asian) shifts notably to the right, peaking around 1250-1300 and tapering off more gradually, with density extending well beyond 1500 and up to 2000. The rightward skew suggests higher performance and greater score variability in white/Asian-majority tracts.

Taken together, the three density plots clearly illustrate a consistent pattern: schools located in majority white or Asian census tracts tend to have higher SAT scores and a broader range of performance, while schools in majority Black or Hispanic tracts are more tightly clustered at lower score levels. The first two plots show this trend separately for each racial grouping, and the third plot combines these distributions, making the disparity more visually pronounced. These findings strongly support the hypothesis that the racial composition of a school's surrounding census tract correlated with school-wide SAT outcomes. This likely reflects broader structural inequalities in resources, opportunity, and access.

Beyond the various bivariate regressions, we also ran a multivariate regression to assess how census tract-level demographic and socioeconomic characteristics correlate with school-wide SAT performance in New York City. The full results of the multivariate regression from R Studio are attached below.

```
Residuals:
   Min     1Q Median     3Q    Max
-282.1  -84.2  -37.7   57.0  752.9

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)              1.26e+03   1.33e+02    9.53   <2e-16 ***
med_hhinc2016            5.07e-04   4.86e-04    1.04    0.297
poor_share2010          -1.88e+02   1.08e+02   -1.74    0.083 .
singleparent_share2010 -1.79e+01   6.65e+01   -0.27    0.789
share_white2000          5.69e+01   1.12e+02    0.51    0.613
share_asian2010          6.51e+01   1.46e+02    0.45    0.657
share_black2010         -9.45e+01   1.26e+02   -0.75    0.452
share_hisp2010          -4.41e+01   1.31e+02   -0.34    0.737
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 165 on 275 degrees of freedom
  (86 observations deleted due to missingness)
Multiple R-squared:  0.181,     Adjusted R-squared:  0.16
F-statistic: 8.66 on 7 and 275 DF,  p-value: 1.3e-09
```

The dependent variable was the average total SAT score per school, while independent variables included median household income (2016), share of residents below the poverty line (2010), share of single-parent households (2010), and racial composition variables (share of white residents in 2000, share of Asian residents in 2010, share of Black residents in 2010, share of Hispanic residesnts in 2010). Although the model is statistically significant overall ($F(7, 275) = 8.66$, $p < 0.001$), it accounts for only a small amount of the variance in SAT scores ($R^2 = 0.181$; Adjusted $R^2 = 0.16$). This suggests that other unmeasured school-level or student-level factors play a substantial role in driving SAT outcomes. Below is a table that specifically breaks down each predictor in the regression model.

Variable	**Estimate**	**p-value**	**Interpretation**
Intercept	1260	<2e-16	Baseline SAT score when all predictors are 0
med_hhinc2016	0.000507	0.297	Not significant | $10,000 increase in median income → +5.07 SAT points
poor_share2010	-188	0.083	Marginally significant | 10% increase in poor share → -18.8 SAT points
singleparent_share2010	-17.9	0.789	Not significant
share_white2000	56.9	0.613	Not significant

share_asian2010	65.1	0.657	Not significant
share_black2010	-94.5	0.452	Not significant
share_hisp2010	-44.1	0.737	Not significant

Among the predictors, only the poverty rate (poor_share2010) approaches statistical significance ($p = 0.083$), with a negative coefficient of -188. This implies that a 10 percentage point increase in the neighborhood poverty rate is associated with an estimated 18.8 point decrease in the school's average SAT score, with everything else held constant. Other variables did not reach statistical significance ($p > 0.29$ for all), though some have large effect sizes in the expected direction. For example, the coefficient for share_black2010 is -94.5, suggesting a potentially negative relationship between the percentage of Black residents and SAT scores. However, the p-value (0.452) indicates low statistical confidence in this estimate.

Overall, this multivariate regression revealed complex interactions among the predictors. Specifically, most variables were not statistically significant when modeled together, which suggests overlapping effects and potential confounding. This means that certain demographic characteristics, such as income, race, and family structure, may be correlated with one another, making it challenging to isolate their individual contributions to SAT performance. For example, poverty rates may be intertwined with racial composition or household structure, which leads to shared variance that dilutes the unique statistical impact of any single variable. As a result, while some variables may show strong associations in isolation, those effects weaken or disappear once the model accounts for other related factors, highlighting the complexity of the social and structural influences on educational outcomes.

Conclusion

Our study worked to examine the extent to which demographic variables, such as income, family structure, and race, correlate with SAT scores in New York City public schools. Across many visualizations and regression models, we observed consistent patterns, namely that schools in census tracts with higher proportions of white or Asian residents, lower poverty rates, and fewer single-parent households tended to have higher average SAT scores. Although many of these variables weren't statistically significant in our multivariate regression, the direction and size of their coefficient were still aligned with our initial hypothesis. Notably, poverty rate was the only variable approaching significance, reinforcing the idea that economic conditions may drive disparities in academic outcomes.

These findings suggest that SAT scores are intertwined with structural inequalities that influence educational opportunity, instead of serving purely as a measure of academic potential.

While our results support our initial hypothesis, this study still has some limitations. For example, demographic characteristics were inferred based on school location rather than where students lived. This may have resulted in some misalignment, especially in a city like New York, where people are highly mobile. Additionally, our data reflects conditions from over a decade ago, and the SAT has undergone many changes since then, including becoming digitized. Despite these limitations, the broader implications of our findings clearly show that standardized tests like the SAT risk perpetuating social inequality when used as a crucial component in college admissions. As debates around test-optional policies continue, our study contributes to a growing body of evidence questioning the SAT's fairness and validity. If educational equity is a national goal, policymakers and institutions must carefully reconsider whether standardized testing should continue to play a key role in determining students' academic futures.