

Contents

Introduction	2
Data.....	2
Associated Variables	2
Exploratory Analysis.....	3
Project Objective.....	8
Model Development	8
Model Interpretation	12
Relevant Prediction.....	13
Conclusion & Suggestion.....	13

Introduction

This project aims to investigate the relationship between the overall academic performance of high school students and their institutional setting. In this study, we analyze data from 372 high schools in New York City. The data consisted of the average SAT scores of 2014-2015 school's cohorts, along with various school and cohort's attributes, such as the school's borough and ethnicity proportion. Our intention is to identify variables that potentially affect the overall academic outcomes, and to quantify the extent that such variables have.

Data

Data	Source	Description
score.csv	https://www.kaggle.com/nycopendata/high-schools	Average SAT scores(Math, Reading, Writing), along with various attributes of 435 schools in NYC. The data pertained to 2014-2015 cohorts
demographics.csv	http://schools.nyc.gov/NR/ronlyres/46093164-D8AA-40DD-A400-8F80CEBC8DD5/0/DemographicSnapshot201112to201516Public_FINAL.xlsx	Contains information about the gender proportion of each school
survey_2014.csv	http://schools.nyc.gov/documents/misc/2014%20Public%20Data%20File%20SUPPRESSED.xlsx	2014 survey result collected from parents and teachers

Associated Variables

In this analysis, we combine data from the mentioned 3 files. We narrow down the associated features from the original sources to just 26 variables. We use R-script(mungdata.R) to perform the data munging and collect the processed result in processed_score.csv. Noted that the total number of observation that we analyze is reduced to 372 instances because of the missing SAT score in some of the data in score.csv. The explanation of each variable in processed_score.csv is shown in the following table.

Variable Name	Description	Type
DBN	School's unique identifier	Character
Borough	School's Borough. Comprised of 5 area: Staten Island, Queens, Manhattan, Brooklyn, Bronx	Character
City	City where the school is located	Character
Latitude	School's Latitude	Numeric
Longitude	School's Longitude	Numeric

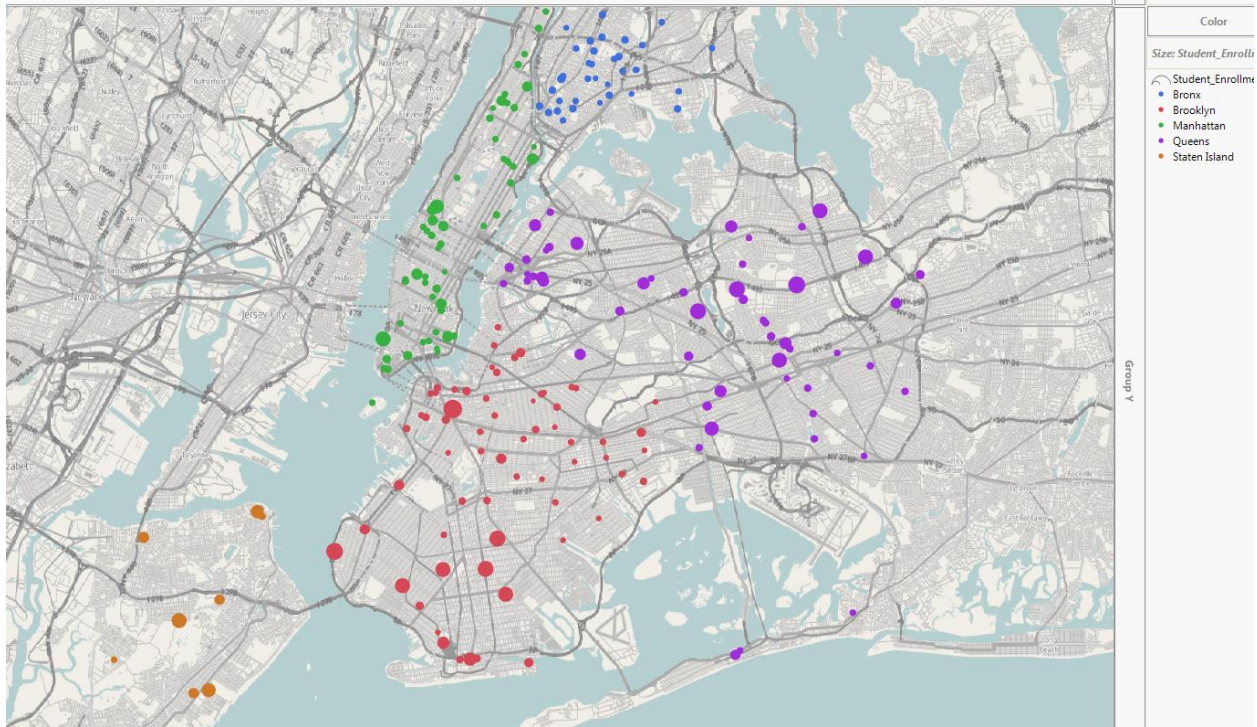
Start_Time	School's Opening hour	Numeric(e.g.: convert from 8:15 AM to 8.15)
End_Time	School's Ending hour	Numeric(e.g.: convert from 4:00 PM to 16.00)
Student_Enrollment	Number of school's enrollment	Numeric
Percent_White	%White students in 2014-2015 cohort	Numeric
Percent_Black	%Black students in 2014-2015 cohort	Numeric
Percent_Hispanic	%Hispanic students in 2014-2015 cohort	Numeric
Persent_Asian	%Asian students in 2014-2015 cohort	Numeric
*Average_SAT_Math	Average SAT Math score of 2014-2015 cohort	Numeric
**Average_SAT_Reading	Average SAT Reading score of 2014-2015 cohort	Numeric
**Averate_SAT_Writing	Average SAT Writing score of 2014-2015 cohort	Numeric
Female_Percent	%Female students in 2014-2015 cohort	Numeric
Male_Percent	%Male students in 2014-2015 cohort	Numeric
Disabilities_Percent	%Disability students in 2014-2015 cohort	Numeric
EngLearner_Percent	%English learner students in 2014-2015 cohort	Numeric
Poverty_Percent	%Poverty students in 2014-2015 cohort	Numeric
Parent_Response_Rate	Parent response rate on 2014 school's survey	Numeric
Teacher_Response_Rate	Teacher response rate on 2014 school's survey	Numeric
Instructional_Core_Satisfaction	%Response regarding instructional satisfaction	Numeric
Systems_for_Improvement_Satisfaction	%Response regarding system satisfaction	Numeric
School_Culture_Satisfaction	%Response regarding culture satisfaction	Numeric
Class_Hours	School's operating duration	Numeric(difference in hour: end_time – open_time)

* - Dependent variable

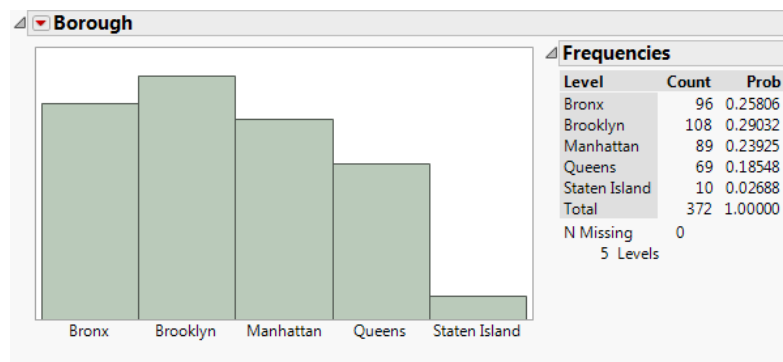
** - Not included in the analysis

Exploratory Analysis

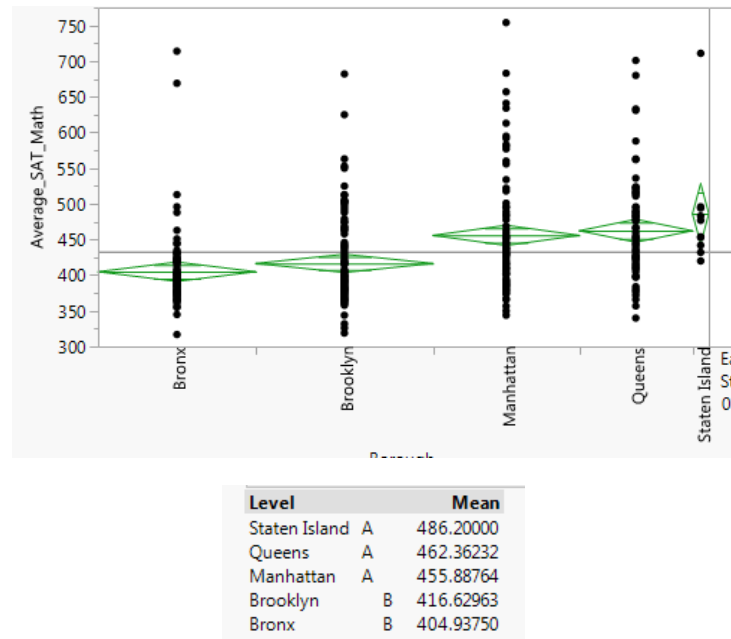
In this analysis, because of a limited time frame, we will quantify the academic performance solely based on the SAT-Math score outcome. The objective is to develop a model to predict the Average_SAT_Math using all other independent variables.



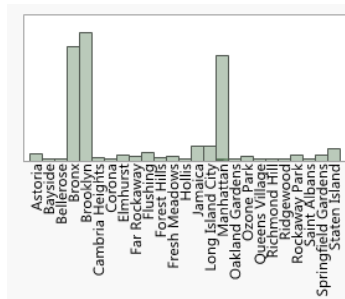
We indicate the location of each school in the above graph. The circle's size corresponds to the size of enrollment. The school's distribution in each borough can be summarized as follows:



We investigate the effect of spatial information (Borough) on SAT Math score by performing ANOVA. We find that this variable may be helpful in predicting the SAT score.



We also investigate the distribution of schools in each City but find that this variable is too fine-grained and decide to drop it as fear of running into overfitting.



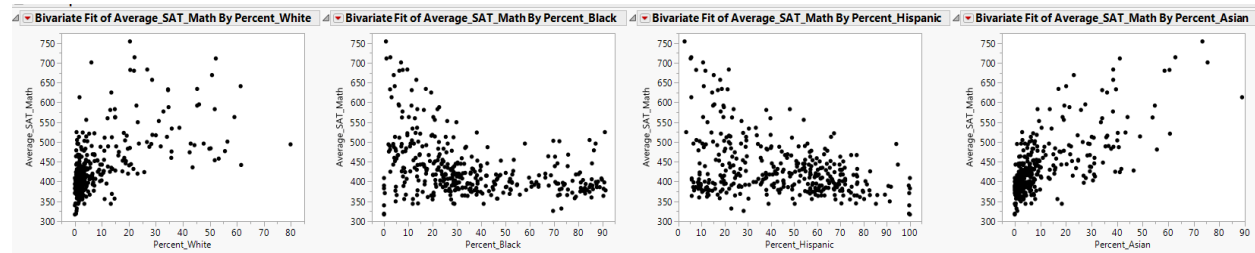
The statistics of other variables are shown in the following table.

Variable	Statistics
Start_Time	<div><div><div><div><div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div></div></div><div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div></div></div><div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div></div></div><div><div></div><div></div><div></div><div></div><div></div></div><div><div></div><div></div><div></div><div></div><div></div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> <div><div></div><div></div><div></div><div></div><div></div></div> </

End_Time	<div><div><div><div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></</div></div></div></div></div></div>
----------	---

Disabilities_Percent	<div><div><div><div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div></div></div></div>
----------------------	---

For preliminary analysis, the scatter plot between ethnicity proportion and SAT-Math score shows some predictive power and indicates that these variables should be included in the model.



Project Objective

Uncover the relationship between the academic setting and cohort's academic outcomes(as measured by the average SAT-Math scores). Linear Regression is chosen as our base model to fit the data on because of its simplicity and interpretability.

Model Development

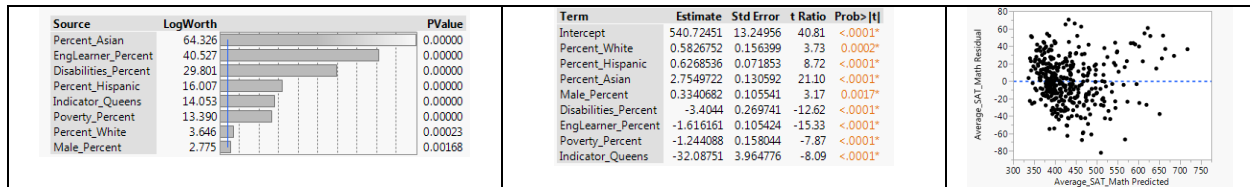
The initial model building consisted of all independent variables (exclude DBN, City, Latitudes, Longitudes, and Female_Percent (as this variable is reflected in Male_Percent)). We obtain the following model:

Effect Summary			Parameter Estimates					Residual Plot	
Source	LogWorth	PValue	Term	Sum of Squares	df	Mean Square	F Ratio	Prob > F	VIF
Intercept	10.125	0.0000	Intercept	444.7095	1	444.7095	4.27	<.0001	1.572814
Borough_Percent	20.385	0.0000	Borough(Bronx)	145.1210	1	145.1210	4.14	<.0001	1.572814
Disabilities_Percent	12.205	0.0000	Borough(Bronx)	155.1210	1	155.1210	1.07	0.2852	1.493237
Poverty_Percent	10.841	0.0000	Borough(Manhattan)	8.349345	1	8.349345	2.42	0.0107	1.493237
Borough	3.210	0.0006	Borough(Queens)	21.77940	1	21.77940	5.94	<.0001	1.392460
Male_Percent	3.179	0.0006	Start_Time	6.024818	1	6.024818	0.61	0.5406	0.2076285
School_Culture_Satisfaction	2.895	0.0027	End_Time	1.00047	1	1.00047	0.00	0.9480	0.9449051
Percent_White	1.621	0.02395	Student_Enrollment	0.00468	1	0.00468	1.83	0.0728	2.133068
Parent_Response_Rate	1.139	0.07284	Percent_White	1.018128	1	1.018128	1.12	0.3445	85.50208
System_for_Improvement_Satisfaction	0.909	0.11779	Percent_Black	0.122947	1	0.122947	0.38	0.5375	29.40118
Instructional_Core_Satisfaction	0.763	0.26654	Percent_Hispanic	0.00205	1	0.00205	0.01	0.9133	173.3143
Percent_Hispanic	0.578	0.31098	Male_Percent	0.00203	1	0.00203	0.29	0.5837	98.20267
Class_Hours	0.441	0.38827	Disabilities_Percent	0.00475	1	0.00475	0.01	0.9133	2.938046
End_Time	0.371	0.42604	Engagement_Percent	0.00475	1	0.00475	0.01	0.9133	2.938046
Teacher_Response_Rate	0.270	0.53765	Poverty_Percent	1.229843	1	1.229843	7.48	<.0001	1.733064
Start_Time	0.267	0.54063	Parent_Response_Rate	0.175259	1	0.175259	2.27	0.0308	1.368086
Percent_Black	0.240	0.57529	Teacher_Response_Rate	0.00889	1	0.00889	0.62	0.5376	1.661082
			Instructional_Core_Satisfaction	0.00889	1	0.00889	0.17	0.6812	1.000000
			System_for_Improvement_Satisfaction	0.00889	1	0.00889	0.17	0.6812	1.000000
			School_Culture_Satisfaction	0.00889	1	0.00889	0.17	0.6812	1.000000
			Client_Hours	0.00889	1	0.00889	0.17	0.6812	1.000000

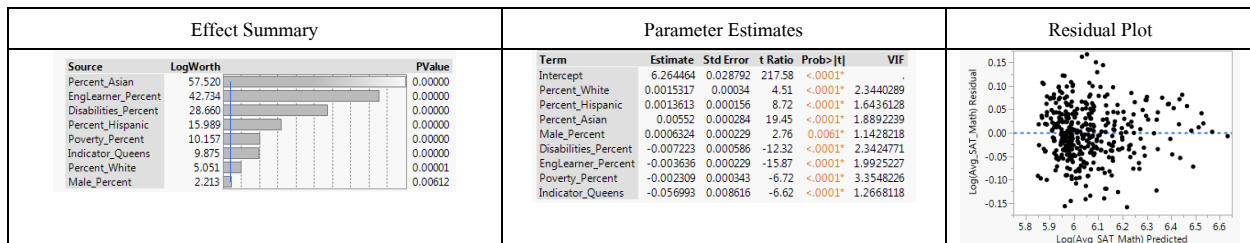
Although the base model shows a strong predictive power (RSquare = 0.87 and RMSE = 26), it has many undesirable properties; the model contains many variables that are not statistically significant, some independent variables are highly correlated (as shown by VIF), and the Residual Plot shows Heteroscedasticity problem (verified by Park-Test). To attenuate these effects, we perform a series of model development, which can be summarized as follows:

1. Manually create a dummy variable based on Borough. As opposed to the one generated by JMP, this will allow us to remove an individual borough that we found not significant. Bronx is treated at the base level since it has the lowest Average SAT-Math score means.
2. Re-fit the model. Iteratively remove variables with high VIF and P-Value that exceeds 0.01 significant threshold.

Effect Summary	Parameter Estimates	Residual Plot
----------------	---------------------	---------------

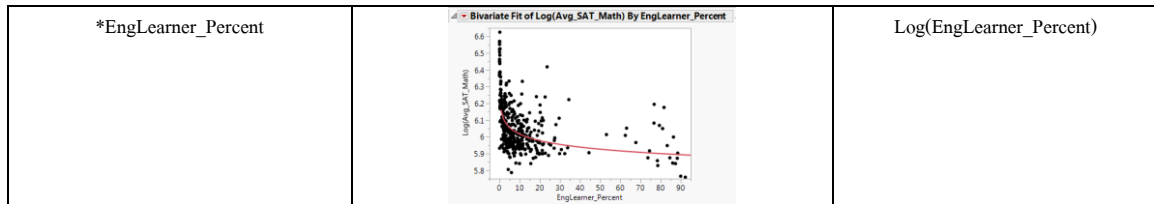


3. As Heteroscedasticity is still presented, we apply Log transformation to Average SAT-Math score and re-fit the model. Drop any unnecessary variable as stated in 2)



4. We try to eliminate Heteroscedasticity by plotting every independent variable against Log(Avg SAT Math) and transform them appropriately if we think that leads to a more linear relationship.

Variable	Plot of variable vs Log(Avg SAT Math) and transformation fit	Transformation Taken
*Percent_White		Log(Percent_White + 0.1)
*Percent_Black		Log(Percent_Black + 0.1)
*Percent_Hispanic		Log(Percent_Hispanic + 0.1)
*Percent_Asian		Log(Percent_Asian + 0.1)

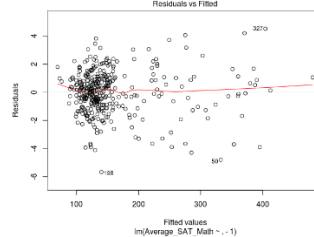


* Add 0.1 to the original value before applying the Log transformation because some instances have 0 value

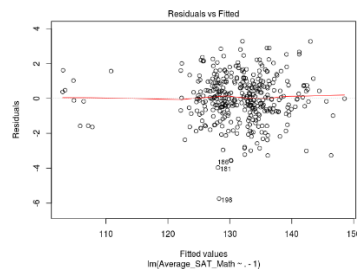
5. Re-fit the model using the transformed variables. Drop any unnecessary variable as stated in 2)

Effect Summary			Parameter Estimates						Residual Plot
Source	LogWorth	PValue	Term	Estimate	Std Error	t Ratio	Prob> t	VIF	
Log(Percent_Asian_adj)	26.680	0.00000	Intercept	6.1194874	0.06625	92.37	<.0001*	.	
Log(EngLearner_Percent_adj)	21.873	0.00000	Indicator_Queens	-0.040766	0.009668	-4.22	<.0001*	1.3138748	
Student_Enrollment	12.281	0.00000	Indicator_Staten_Island	-0.058806	0.022263	-2.64	0.0086*	1.206477	
Log(Percent_Black_adj)	7.769	0.00000	Student_Enrollment	4.0076e-5	5.348e-6	7.49	<.0001*	1.6645134	
School_Culture_Satisfaction	7.332	0.00000	Disabilities_Percent	-0.001553	0.00059	-2.63	0.0086*	1.9542378	
System_for_Improvement_Satisfaction	6.633	0.00000	Poverty_Percent	-0.001524	0.000327	-4.66	<.0001*	2.5098501	
Poverty_Percent	5.344	0.00000	Systems_for_Improvement_Satisfaction	-0.005927	0.001124	-5.27	<.0001*	4.7956064	
Indicator_Queens	4.504	0.00003	School_Culture_Satisfaction	0.0075669	0.001355	5.58	<.0001*	5.8679113	
Indicator_Staten_Island	2.065	0.00862	Log(Percent_Black_adj)	-0.022268	0.003859	-5.77	<.0001*	1.8124027	
Disabilities_Percent	2.054	0.00884	Log(EngLearner_Percent_adj)	-0.037004	0.003532	-10.48	<.0001*	2.5515644	
			Log(Percent_Asian_adj)	0.0377099	0.003195	11.80	<.0001*	1.9398177	

6. We have lessened the effect of Heteroscedasticity but the issue is still presented. We turn to Weighted Least Squares Regression approach by assuming the assumption of non-constant error variance. We switch to using R to conduct the analysis at this point (as performing the analysis in JMP can be quite tedious). The analysis code can be found in analysis.R . We use the transformed data collected from step 4). As the observations come from aggregated result, we firstly try to weight the error variance by the enrollment size (that is, multiply every variable by $\sqrt{\text{Student_Enrollment}}$) and fit the regression model with no intercept. The residual plot indicates that Heteroscedasticity is still presented.



7. Now, we try another approach where error variance is weighted by model's residuals; firstly, fit the regression model using transformed variables in step 4) and then use the mean square residual of each borough as a weight for WLS. The residual plot indicate that Heteroscedasticity issue is now fixed.

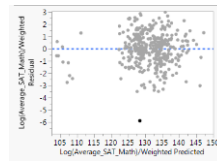


8. We import the transformed data back to JMP(weighted_score.csv) and drop any unnecessary variable as stated in 2).

We obtain the following model:

Effect Summary			Parameter Estimates					Residual Plot
Source	LogWorth	PValue	Term	Estimate	Std Error	t Ratio	Prob> t	
1/Weighted	253.427	0.00000	Student_Enrollment/Weighted	0.0000395	5.377e-6	7.34	<.0001*	
Log(Percent_Asian_adj)/Weighted	26.231	0.00000	Log(Percent_Black_adj)/Weighted	-0.021331	0.003884	-5.62	<.0001*	
Log(EngLearner_Percent_adj)/Weighted	21.764	0.00000	Log(Percent_Asian_adj)/Weighted	0.0368968	0.003159	11.68	<.0001*	
Student_Enrollment/Weighted	11.858	0.00000	Disabilities_Percent/Weighted	-0.00174	0.000577	-3.02	0.0027*	
School_Culture_Satisfaction/Weighted	7.881	0.00000	Log(EngLearner_Percent_adj)/Weighted	-0.036992	0.003542	-10.44	<.0001*	
Systems_for_Improvement_Satisfaction/Weighted	7.493	0.00000	Poverty_Percent/Weighted	-0.001402	0.000324	-4.33	<.0001*	
Log(Percent_Black_adj)/Weighted	7.407	0.00000	Systems_for_Improvement_Satisfaction/Weighted	-0.006309	0.001116	-5.65	<.0001*	
Poverty_Percent/Weighted	4.719	0.00002	School_Culture_Satisfaction/Weighted	0.0078705	0.001353	5.82	<.0001*	
Queens/Weighted	3.817	0.00015	1/Weighted	6.1170387	0.06605	92.61	<.0001*	
Disabilities_Percent/Weighted	2.561	0.00274	Queens/Weighted	-0.037877	0.009895	-3.83	0.0002*	

9. We drop the data point that poses too much influence on the model. Its Cook's Distance is 0.244, far exceeds the recommended threshold Cook's Distance ($4/372 = 0.01$).



10. We refit the model using the remaining 371 observations:

Effect Summary			Parameter Estimates					Residual Plot
Source	LogWorth	PValue	Term	Estimate	Std Error	t Ratio	Prob> t	
1/Weighted	256.992	0.00000	Student_Enrollment/Weighted	3.9167e-5	5.224e-6	7.49	<.0001*	
Log(Percent_Asian_adj)/Weighted	27.713	0.00000	Log(Percent_Black_adj)/Weighted	-0.024356	0.003778	-6.45	<.0001*	
Log(EngLearner_Percent_adj)/Weighted	26.018	0.00000	Log(Percent_Asian_adj)/Weighted	0.037114	0.003073	12.08	<.0001*	
Student_Enrollment/Weighted	12.268	0.00000	Disabilities_Percent/Weighted	-0.001578	0.000562	-2.81	0.0053*	
Log(Percent_Black_adj)/Weighted	9.436	0.00000	Log(EngLearner_Percent_adj)/Weighted	-0.041791	0.003595	-11.62	<.0001*	
School_Culture_Satisfaction/Weighted	7.509	0.00000	Poverty_Percent/Weighted	-0.001045	0.000324	-3.23	0.0014*	
Systems_for_Improvement_Satisfaction/Weighted	7.210	0.00000	Systems_for_Improvement_Satisfaction/Weighted	-0.006013	0.001087	-5.53	<.0001*	
Queens/Weighted	3.228	0.00059	School_Culture_Satisfaction/Weighted	0.0074627	0.001319	5.66	<.0001*	
Poverty_Percent/Weighted	2.864	0.00137	1/Weighted	6.112735	0.064244	95.14	<.0001*	
Disabilities_Percent/Weighted	2.277	0.00528	Queens/Weighted	-0.033334	0.009669	-3.47	0.0009*	

As the highest Cook's Distance of data used to fit this model is 0.09 and not too severe, no further data are dropped.

11. We then perform Park-Test to verify for Heteroscedasticity. The P-Value for \hat{Y} when regressed on r^2 is 0.55 and when regressed on $\log(r^2)$ is 0.90. This indicates that the Heteroscedasticity is now fixed.

The model in step 8) is our selected model. The model equation is

$$\frac{\text{Log(AvgSatMath)}}{\text{Weight}} = 6.1 * \frac{1}{\text{Weighted}} + 0.000039 * \frac{\text{StudentEnrollment}}{\text{Weighted}} - 0.02 * \frac{\text{Log(PercentBlack)}}{\text{Weighted}} + 0.03 * \frac{\text{Log(PercentAsian)}}{\text{Weighted}} - 0.001 * \frac{\text{DisabilityPercent}}{\text{Weighted}} - 0.04 * \frac{\text{Log(EngLearnerPercent)}}{\text{Weighted}} - 0.001 * \frac{\text{PovertyPercent}}{\text{Weighted}} - 0.006 * \frac{\text{SystemForImprovementSatisfaction}}{\text{Weighted}} + 0.007 * \frac{\text{SchoolCultureSatisfaction}}{\text{Weighted}} - 0.03 * \frac{\text{Queens}}{\text{Weighted}}$$

Where Weighted is the means of Residual Square per group and has the following values:

Group	Weighted
Bronx	0.0463926118015564
Brooklyn	0.0447919246176774
Manhattan	0.0446127098917998

Queens	0.0484199559569425
Staten Island	0.0584236025734739

The equation can be simplified to

$$\begin{aligned} \text{Log}(\text{AvgSatMath}) = & 6.1 * 0.000039 * \text{StudentEnrollment} - 0.02 * \text{Log}(\text{PercentBlack}) + 0.03 * \text{Log}(\text{PercentAsian}) - 0.001 \\ & * \text{DisabilityPercent} - 0.04 * \text{Log}(\text{EngLearnerPercent}) - 0.001 * \text{PovertyPercent} - 0.006 \\ & * \text{SystemForImprovementSatisfaction} + 0.007 * \text{SchoolCultureSatisfaction} - 0.03 * \text{Queens} \end{aligned}$$

Model Interpretation

The F-Statistics and P-Value of the model's coefficients are all less than 0.01 significant threshold, thus, indicate that they are all statistically significant. We also use the obtained model to predict the value of $\text{Log}(\text{Average_SAT_Math})$, using the original 372 instances, and transform the value back to the original space, Average_SAT_Math , to assess the model's performance. The model's performance is: R-Squared = 0.84 and RMSE = 28.25.

The model's interpretation is as follows:

- For every 100 persons increase in student enrollment, we expect a 0.39% increase in Average_SAT_Math score
- For a 1% increase in the ratio of Black students, we expect a 0.02% decrease in Average_SAT_Math score
- For a 1% increase in the ratio of Asian students, we expect a 0.03% increase in Average_SAT_Math score
- As the percentage of Disability students increased by 1, we expect a 0.1% decrease in Average_SAT_Math score
- For a 1% increase in the ratio of English learner students, we expect a 0.04% decrease in Average_SAT_Math score
- As the percentage of Poverty students increased by 1, we expect a 0.1% decrease in Average_SAT_Math score
- As the percentage of responders who satisfied with the system improvement increased by 1, we expect a 0.6% decrease in Average_SAT_Math score
- As the percentage of responders who satisfied with the school culture increased by 1, we expect a 0.7% increase in Average_SAT_Math score
- On average, the Average_SAT_Math score of schools in Queens borough is 3% lower than other schools

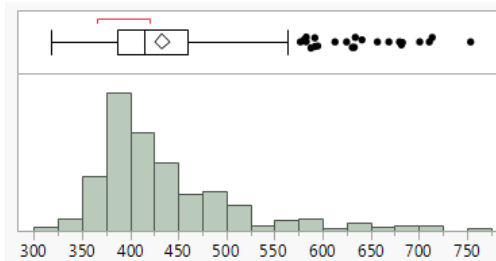
Model Observation:

- We presume that larger class size may deteriorate academic performance. The model instead tells us that this is not the case. Other variables, such as student-to-teacher ratio, should be included to make more accurate prediction.
- Increase in the ratio of Black students deteriorates academic performance, vice versa for Asian students. White and Hispanic students have no effect.

- It is understandable that the ratios of disability students, English learner students, and poverty students are negatively correlated with the academic performance.
- The satisfaction of system improvement is negatively correlated with the Average_SAT_Math score. This seems counter-intuitive and required further investigation.
- The means Average_SAT_Math of schools in Queens is the second highest among all 5 boroughs. Yet, the model suggests that the score of schools in this area is around 3% lower than other schools.

Relevant Prediction

We demonstrate the model prediction by using data from the school “New Explorations into Science, Technology and Math High School”(DBN: 01M539). The predicted Average_SAT_Math of the school is 646.8 while the actual score is 657. From the histogram plot of Average_SAT_Math, the score that the students obtained is quite on a high side. The prediction is fairly closed and agrees with the model’s interpretation since the school comprised of a high percentage of Asian students (much higher than the average), low percentage of Black students, nearly none English learners, etc.



Conclusion & Suggestion

With $R\text{-Squared} = 0.84$ and $RMSE = 28.25$, we are content with the model’s performance. However, highly accurate model is not our main goal here as we opt for simpler and interpretable model. For highest predictive performance reason, readers should also consider cross-validation technique to prevent overfitting, as well as other non-linear modellings.

We also quantify the effect of each variable based on just one aspect of academic outcome, the Average_SAT_Math score. It is also interesting to investigate Average_SAT_Reading and Average_SAT_Writing scores as well. It may turn out that each variable has a different impact on a different section of SAT score.