Introduction

For my project, I will examine data to see how income levels affect SAT scores. By doing this, I will be able to see which schools average higher scores on the tests.

Every year, thousands of students take the SAT's in order to be accepted to most colleges. The scores on these tests vary from student to student. In most cases, students who attend private schools come from homes where the parents have a larger income level. But, for the research for this project I will look at public schools. By looking at the scores of public school while also looking at their income level, I may be able to find a correlation.

As these tests are the same for students, whether they are in public school, private school, or any other form of schooling. These tests are required by a majority of colleges for admission.

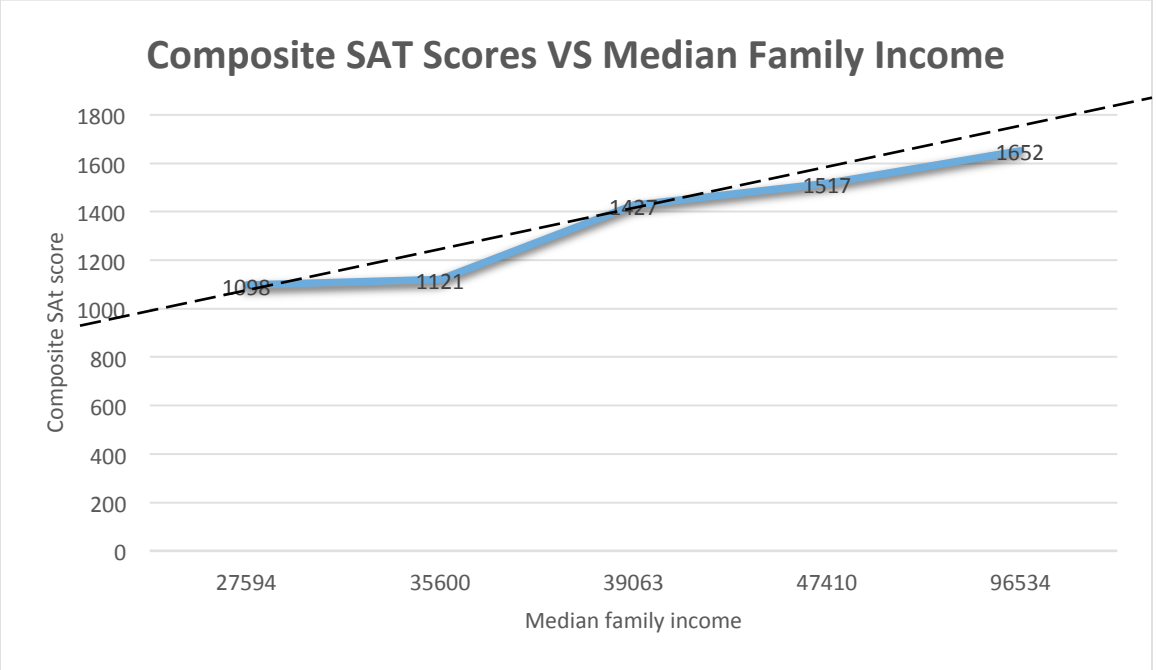
Methodology

The population of this study is made of public school students in Pennsylvania. The school districts included are Burrell School District, Northeast High school, Perry Traditional Academy, Pine Richland, and Aliquippa. The sample will include all students who took the tests in the school districts in 2014. This is a representative sample of the population because it includes all members who took it, not just the high scorers or low scorers. The observational unit is the student, and the data to be used to compare the scores is the quantitative variable of income (dollars).

By choosing 5 school districts at with varying scores from low to high, I was able to then go and see the median household income of that area to see if the amount was high or low. Without knowing if the income level would be high or low, I could not control the data. The data was as follows:

	Composite Score	Median Family income
Burrell	1517	47410
Pine Richland	1652	96534
North East	1427	39063
Perry	1121	35600
Aliquippa	1098	27594

From this data, correlation is shown below



Results:

After calculating the correlation between these numbers, I found the correlation to be r=0.7731. This correlation shows that there is moderate to strong relation between these numbers. Although it is not very strong, it still shows a pattern between the data points. The line of

regression can be calculated as $\hat{y}$=992.606+0.007x. By using this line, we can assume that if we set x to $0, we would find that $\hat{y}$ would equal 992.606 which would mean that families with no income at all will average a score of 992 on the SAT. Continuing with the regression line, if we set x to $100000, $\hat{y}$ would be 1693. Designated on the chart with the black dash is the regression line. By looking at this line, we can also see the relation between these two and see how closely it follows to the plotted points.

Conclusion:

By looking at all of this data, the relationship between median family income and the composite SAT scores of Pennsylvania school districts is seen. With a correlation of r=0.7731 and by looking at the regression line, we can conclude that the two pieces of data are related. The line of regression is similar to the actual points and the correlation is moderate to very strong which proves this conclusion. The higher the median income of a school district, the higher the SAT scores seem to be on average. This can help to understand the importance of money in education as well as what the ability to afford tutors may alter. Lower income schools need more funding in order to get their students to the same level as other schools as they cannot alter the amounts of money the parents in the districts make.

Future study:

If I were to do this project again in the future, I would like to do 2 different states to compare those scores as well to see if the regression line is any different. I would also use 10 schools with each data set instead of 5 to get a better reading. The schools I chose may have just been coincidences and although I believe my information is correct, I would not know without doing more samples. I know that because of my time constraints I may have not been able to get

accurate enough results, but I believe that I did find very good information for the time that I did have. As a future teacher, this project really interested me and I would like to look into it further in the future.