LOSING SCHOOLS TO THE ECONOMIC CHASM:

A MULTIPLE REGRESSION ANALYSIS OF FAMILY INCOME AND SCHOOL PERFORMANCE PROFILE SCORES

A Thesis

Submitted to McAnulty College and Graduate School of Liberal Arts

Duquesne University

In partial fulfillment of the requirements for

the Master of Science degree of Computational Mathematics

By

Lisa Over

May 2015

LOSING SCHOOLS TO THE ECONOMIC CHASM:

A MULTIPLE REGRESSION ANALYSIS OF FAMILY INCOME AND

SCHOOL PERFORMANCE PROFILE SCORES

By

Lisa Over

Approved March 13, 2015 of Defense

John C. Kern II, Ph.D.
Chair and Associate Professor of Statistics
Department of Mathematics and
Computer Science
(Committee Chair)

James Schreiber, Ph.D.
Professor of Education
Department of Educational Foundations
and Leadership
(Committee Member)

James Swindal, Ph.D.
Dean and Professor of Philosophy
McAnulty College and Graduate School
of Liberal Arts

Jeffrey Jackson, Ph.D.
Director of Graduate Studies
Professor of Computer Science
Department of Mathematics and
Computer Science

ABSTRACT

LOSING SCHOOLS TO THE ECONOMIC CHASM:

A MULTIPLE REGRESSION ANALYSIS OF FAMILY INCOME AND

SCHOOL PERFORMANCE PROFILE SCORES

By

Lisa Over

May 2015

Dissertation supervised by John C. Kern II, Ph.D.

The No Child Left Behind (NCLB) era of educational reform, which is now at the close of its 14th year, focuses on improving student learning by using research-based teaching techniques, making evidence-based decisions, hiring quality teachers, holding schools and school districts accountable for student performance, and allowing families to leave a school that is designated as "failing" to attend one with a better report card. The goal of NCLB is to make sure each and every child reaches certain learning outcomes before he or she graduates from high school. However, school performance measures show that fourteen years of NCLB funding and efforts have not reached every school and, consequently, not every child. Many schools are still designated as "failing" as a result of numerous low performing students.

The purpose of this study is to examine the amount of variance in the new Pennsylvania school performance data that is explained by IRS-based income values. Upon initial inspection

of school performance, it appears that scores decrease as income decreases. If this correlation is true and a statistically and practically significant amount of variance in school performance can be explained by income, then income is an underlying problem that cannot be resolved by NCLB efforts alone. The problem of low performing schools cannot be fixed until underlying issues are recognized and resolved. The goal of formally exposing income as a predictive variable in the school performance model is to bring a new awareness to the persistence of the economic chasm that separates school districts, schools, and students and their families.

DEDICATION

To my sister, Amy Netter, for her many years of dedicated service as a teacher of and mentor to at-risk, disadvantaged youth.

ACKNOWLEDGEMENT

Throughout my graduate education Dr. John Kern and Dr. Frank D’Amico have inspired and fostered my interest in statistics, and Dr. Jeffrey Jackson and Dr. Adam Drozdek have reinforced my interest in computer science. Dr. Kern approached my questions and incorrect thought patterns with patience and clarity. Dr. D’Amico’s contagious passion for statistics shone through his many accounts of his work in the hospital. Dr. Jackson and Dr. Drozdek willingly engaged my interests in web programming and database management systems through independent studies. Although the focus of this project was statistics, specifically multiple regression analysis, the data collection and management required knowledge of databases and the companion website required knowledge of web programming. The work of all four professors facilitated the integration of statistics and computer science into this project.

From the beginning of this project to its completion, Dr. James Schreiber, professor of Education from the department of Education and Leadership, has provided ideas, guidance, and support in many forms. His knowledge of both statistics and public education trends and policies has proven invaluable to the completion of this project.

Finally, my family has been a tremendous source of support and encouragement throughout this entire program. Thank you Joel, Nate, and Abbi for your love and commitment.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

Introduction/Theory

The School Performance Profile (SPP) score is a quantitative score assigned to each school in Pennsylvania to summarize the school's performance as indicated by a number of academic measures including students' standardized test scores. Much like grades given to students, the SPP score is measured on a 100-point scale. (PA School Performance, 2013c) "Appendix A: School Performance Profile Score" on page 44 provides a detailed explanation of the indicators used to calculate SPP scores. Approximately 90% of each SPP score is derived either directly from student standardized test scores or from some comparison or analysis of student standardized test scores.

The SPP scores vary from school to school, and this variability is explained by the differences in the indicators used to calculate the score. However, the purpose of this study is to explore another possible explanation for the variability in SPP scores: family income. SPP scores for schools in lower income communities appear to be lower than the SPP scores for schools in higher income areas.

Research Questions

The questions this study attempts to answer are "Does a correlation exist between family income and SPP scores?" and "If a correlation exists, how much variability in SPP scores can be explained by income?"

Method

Sampling

The sample of data chosen for this study includes 257 public schools from Allegheny County in Pennsylvania. Charter schools were excluded because they are independent from the public school system with some managed by private, for-profit businesses. A complete list of schools appears in "Appendix B: School and Estimated Income Data" on page 49. This sample can be regarded as a random sample representative of a larger population that includes counties with demographics, tax structures, and forms of government that are similar to Allegheny County.

A Microsoft Access® database was created to manage the sample data for this study. Four data files were downloaded and imported into separate tables in the database. Two school-related data files were retrieved from the PA School Performance website *http://www.paschoolperformance.org*. Two income-related data files were retrieved from the U.S. Census Bureau FactFinder website *http://factfinder2.census.gov*.

School data. The file "2012.2013.SPP.Scores.School.List.12.11.13.xlsx" contains the county, LEA name (Local Education Agency), school name, AUN (Administrative Unit Number), school number, and 2012-2013 SPP score for all schools in Pennsylvania (PA School Performance, 2013a). It was imported to have access to each school's SPP score as well as to the county in which each school resides. This study focuses on Allegheny County schools but further study is planned for all schools in Pennsylvania. Grouping the schools by county better organizes these endeavors.

The file "SPP.FF.2012.2013.txt" contains fast fact data for each school district and for each school in Pennsylvania. (PA School Performance, 2013b) This file contains multiple records for each school with each record containing six fields. Four fields specified the school: LEA Name, School Name, AUN, and School Number, and two fields specified the fact: Data Element and Display Value. Queries were created to retrieve facts of interest. Among the records for each school was the display value "School Zip Code," which is the physical zip code of the school building and necessary to connect school data to income data. A query retrieved the zip codes, which were then copied to a new table along with the corresponding school number.

Many of the school zip codes from the fast fact file (2013b) included the 4-digit zip code extension. This presented a problem for matching income to schools because the zip codes obtained from the mean income data file (U.S. Census Bureau, 2012a) and the median income data file (U.S. Census Bureau, 2012b) did not have the 4-digit extension. This extension had to be eliminated before school performance data could be matched to income data. A Visual Basic script was created to reference the zip code table, retrieve the school zip code, and copy only the first 5 digits to a new field in the table. The code for this script can be found in "Appendix C: VB Script to Extract 5-digit Zip Codes" on page 64.

Income data. The file "ACS_12_5YR_S1902_with_ann.csv" contains 2008-2012 American Community Survey 5-Year Estimates of mean income by zip code in the past 12 months and in 2012 inflation-adjusted dollars. (U.S. Census Bureau, 2012a) The file "ACS_12_5YR_S1903_with_ann.csv" contains 2008-2012 American Community Survey 5-Year Estimates of median income by zip code in the past 12 months and in 2012 inflation-adjusted dollars. (U.S. Census Bureau, 2012b)

According to the Census bureau document, "A Compass for Understanding and Using American Community Survey Data: What General Data Users Need to Know," the 5-year estimates are more reliable and precise than the 1-year and 3-year estimates. (U.S. Census Bureau, 2008) Although 5-year estimates are not as current as the other estimates, current data is not as important as precise data in this study. This study is concerned with the precise, relative difference in income for different schools, which is obtained using the 5-year estimates.

Connecting schools to income data. The 5-year estimated mean (U.S. Census Bureau, 2013a) and median (U.S. Census Bureau, 2013a) incomes were matched by zip code to each school's physical address. These income values were further estimated, hereafter referred to as the multiple-zip estimate, to account for students attending a school but living outside the zip code of the school's physical address. For practical reasons, this study assumes that each elementary school services students who reside in the same zip code as the school building. Obtaining an exact income for each school would require knowing not only which zip codes feed into a school but also the distribution of students from each zip code. This level of precision was deemed impractical to achieve because it would require obtaining lists of students and addresses for each school. Therefore, the multiple-zip estimation was performed for schools whose students came from different elementary and/or middle schools and included mostly middle schools and high schools from districts that serve students from more than one zip code. For each school that qualified for the multiple-zip estimation, a weighted average of the incomes from all schools that feed into the qualifying school was calculated and recorded in a new field.

Data

The data for this study includes SPP scores, mean income, and median income for 257 public schools in Allegheny County. There is no missing data. All 257 schools have the outcome variable, SPP score, and the two explanatory variables, mean and median incomes.

SPP Scores. The distribution of SPP scores, illustrated with a box plot and histogram in Figure 1 below, is skewed left with approximately 73% of the scores being at least 70. The mean score is 77.7 with standard deviation 13.7. There is one outlier noted as a dot outside the bounds of the box plot. An outlier is an observation whose value is below the first quartile minus 1.5 times the interquartile range or above the third quartile plus 1.5 times the interquartile range.

This low outlying score, 36.3, is for Wilkinsburg Senior High School of Wilkinsburg Borough School District.

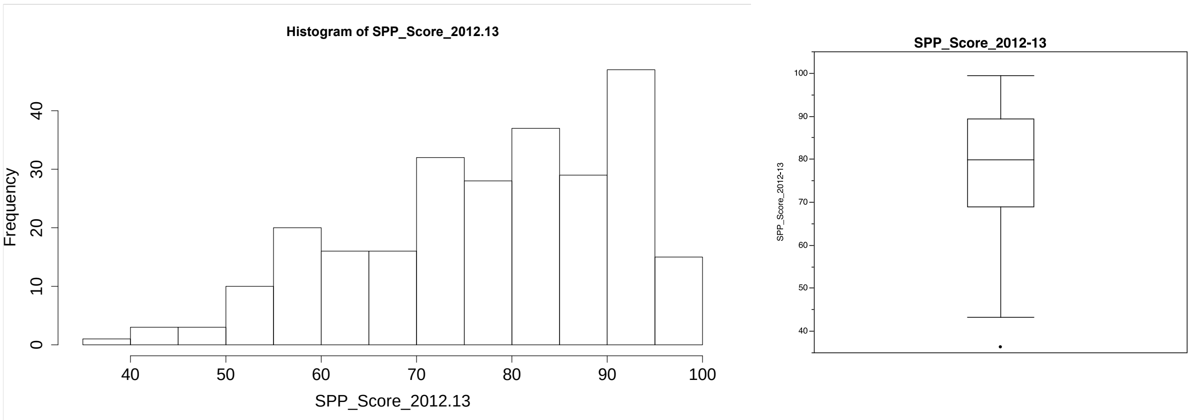


Figure 1 Distribution of SPP Scores for Allegheny County (JMP, 2012)

Mean income. The distribution of mean income after the multiple-zip estimation (defined in "Connecting schools to income data" on page 4), illustrated with a box plot and histogram in Figure 2 below, is skewed right with 86.4% of the data being below $100,000 and 53.7% of the data being below $70,000. The average mean income is $74,166 with standard deviation $25,669. There are 22 high outliers noted as dots outside the box plot. All 22 outliers, summarized in Table 1 below, have mean incomes between $127,000 and $153,000.

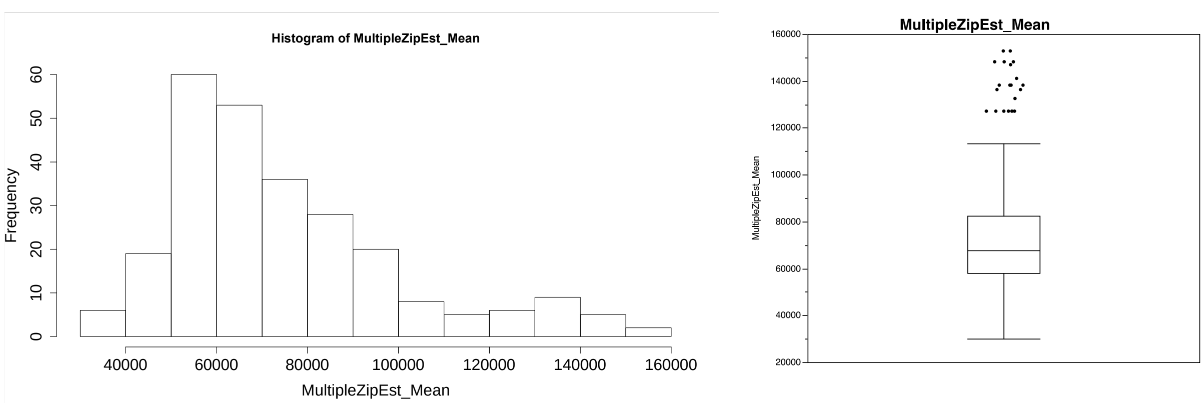


Figure 2 Distribution of Mean Income after Multiple-zip Estimation (JMP, 2012)

School District	School Name	Mean Income ($)
Fox Chapel Area SD	Dorseyville MS	136,271
Fox Chapel Area SD	Fairview El Sch	148,135
Fox Chapel Area SD	Fox Chapel Area HS	136,271
Fox Chapel Area SD	Hartwood El Sch	148,135
Fox Chapel Area SD	OHara El Sch	148,135
Hampton Township SD	Poff El Sch	127,080
North Allegheny SD	Bradford Woods El Sch	141,033
North Allegheny SD	Franklin El Sch	132,475
North Allegheny SD	Marshall El Sch	152,756
North Allegheny SD	Marshall MS	146,895
Pine-Richland SD	Eden Hall Upper El Sch	127,080
Pine-Richland SD	Hance El Sch	127,080
Pine-Richland SD	Pine-Richland HS	127,080
Pine-Richland SD	Pine-Richland MS	127,080
Pine-Richland SD	Richland El Sch	127,080
Pine-Richland SD	Wexford El Sch	152,756
Upper Saint Clair SD	Baker El Sch	138,167
Upper Saint Clair SD	Boyce MS	138,167
Upper Saint Clair SD	Eisenhower El Sch	138,167
Upper Saint Clair SD	Fort Couch MS	138,167
Upper Saint Clair SD	Streams El Sch	138,167
Upper Saint Clair SD	Upper Saint Clair HS	138,167

Table 1 List of Schools with Mean Incomes that Qualify as Outliers

Median income. The distribution of median income after the multiple-zip estimation (defined in “Connecting schools to income data” on page 4), illustrated with a box plot and histogram in Figure 3 below, is skewed right with 85.6% of the data being below $70,000. The average median income is $55,844 with standard deviation $18,275. There are 4 outliers noted as dots outside the box plot. All 4 high outliers, summarized in Table 2 below, have median incomes between $107,000 and $116,000.

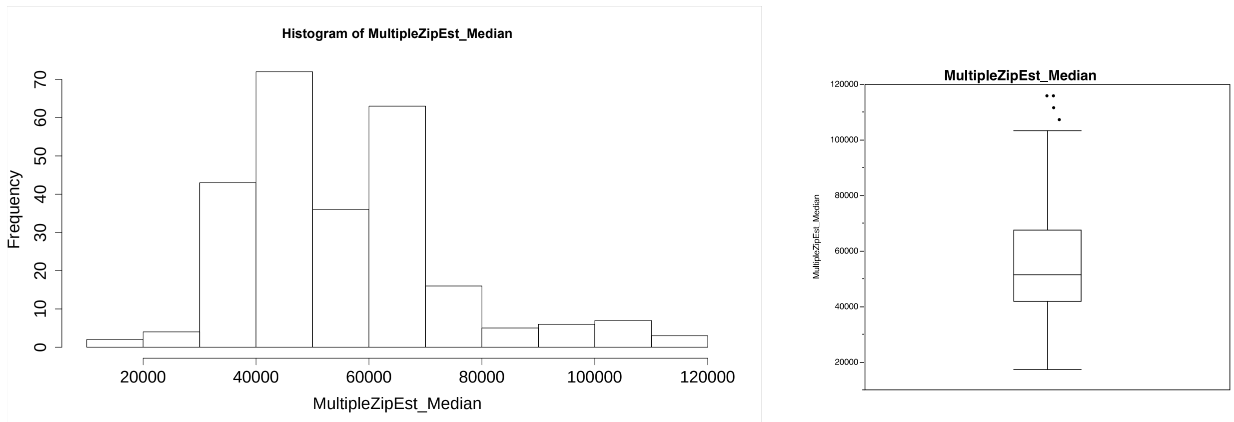


Figure 3 Distribution of Median Income After Multiple Zip Estimation (JMP, 2012)

School District	School Name	Median Income ($)
North Allegheny SD	Bradford Woods El Sch	107,188
North Allegheny SD	Marshall MS	111,488
North Allegheny SD	Marshall El Sch	115,787
Pine-Richland SD	Wexford El Sch	115,787

Table 2 List of Schools with Median Incomes that Qualify as Outliers

Income and SPP scores. Figure 4 and Figure 5 below illustrate the relationship SPP scores have with mean family income and median family income, respectively. According to Figure 4, as mean income increases, the mean SPP score increases. The mean SPP score is 57.15 when mean family income is 1.5 to 2.5 standard deviations below the average mean income, while the mean SPP score is 91.41 when mean family income is 2.5 to 3.5 standard deviations above the average mean income. Likewise, according to Figure 5, as median income increases, the mean SPP score increases. The mean SPP score is 57.48 when median family income is 1.5 to 2.5 standard deviations below the average median income, while the mean SPP score is 92.54 when median family income is 2.5 to 3.5 standard deviations above the average median income.

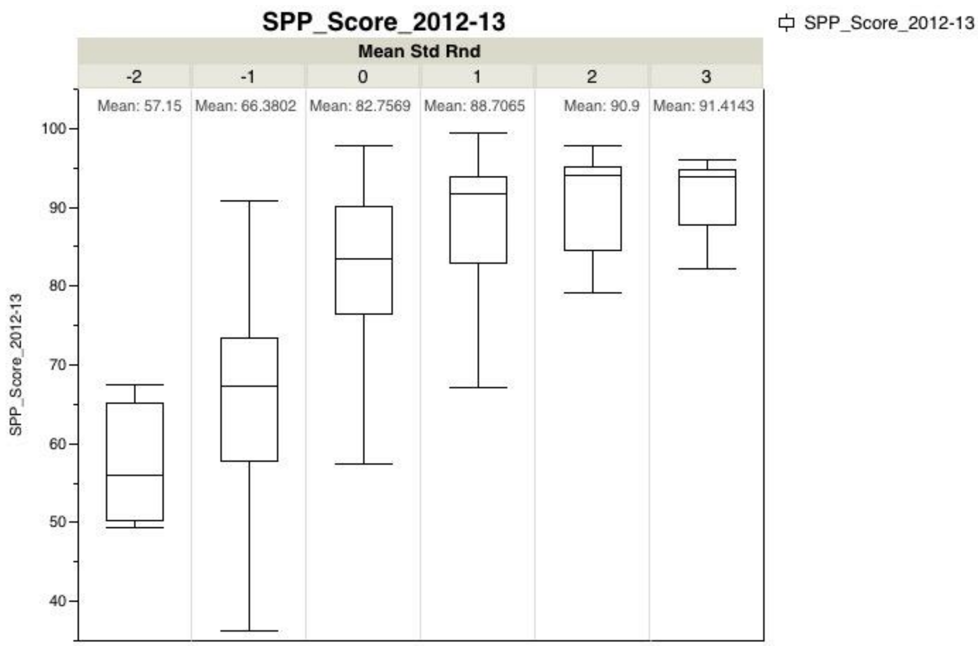


Figure 4 Box Plots Relating Mean Income to SPP Scores (JMP, 2012)

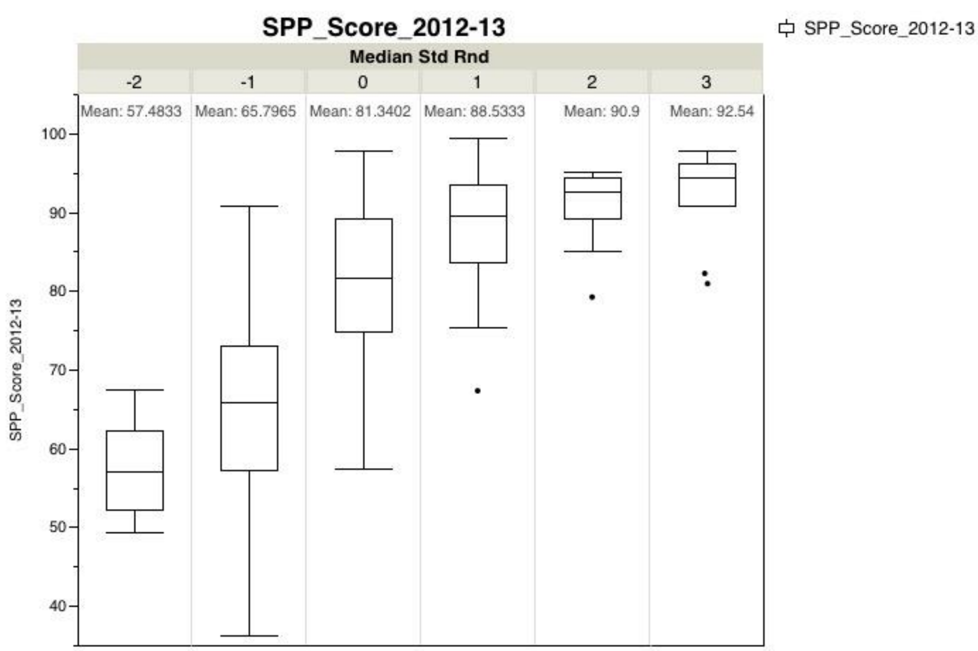


Figure 5 Box Plots Relating Median Income to SPP Scores (JMP, 2012)

Methodology

Multiple linear regression is the approach used to determine if income is correlated with school performance. Regression involves finding a linear equation that best describes the relationship between SPP scores and income and then evaluating that equation to determine how well it fits the data.

Exploring the relationships. From “Income and SPP scores” on page 7, a relationship between both mean and median incomes is apparent. Therefore, two simple linear models were explored involving each type of income. Figure 6 below illustrates the relationship between SPP scores and mean income. The equation of the regression line is

$$SPP = \beta_0 + \beta_1(mean\ income) + \text{error} \tag{1}$$

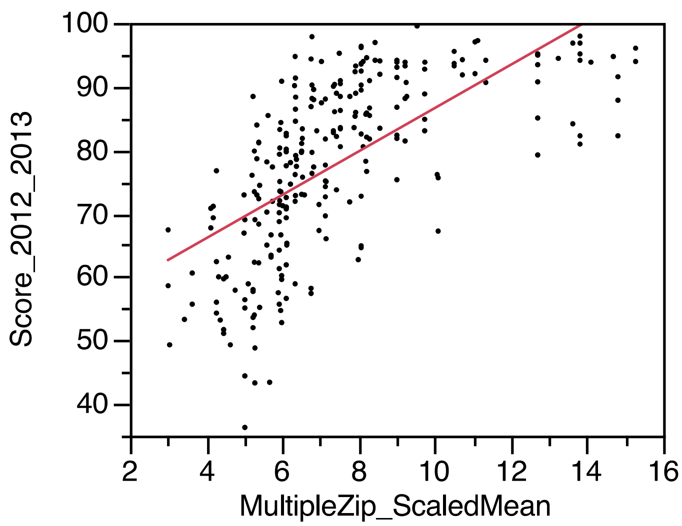


Figure 6 Regression Plot for Model with Mean Income (JMP, 2012)

Figure 7 below illustrates the relationship between SPP scores and median income. The equation of the regression line is

$$SPP = \beta_0 + \beta_1(median\ income) + \text{error} \tag{2}$$

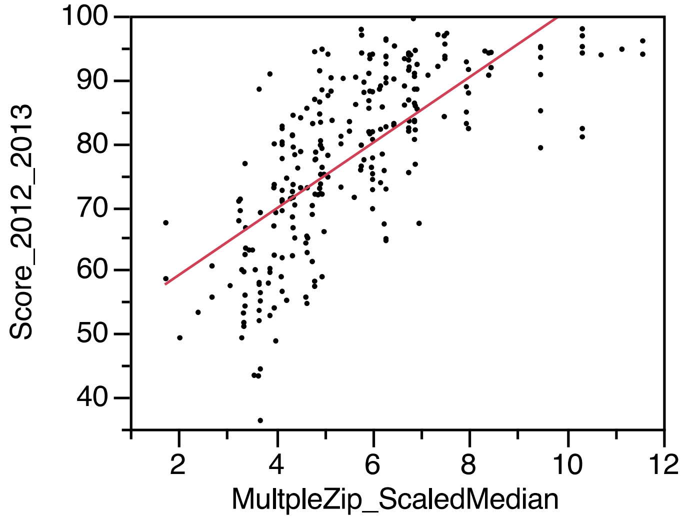


Figure 7 Regression Plot for Model with Median Income (JMP, 2012)

In equations (1) and (2) above, each β_j is a constant determined by the least squares method, which will be discussed in detail later when a model is chosen for further analysis. Both equations are statistically significant indicating that either income measure (mean income or median income) can be used to explain the variation in SPP scores. However, the idea of linear regression is to find a line that best describes the data. The points in both Figure 6 and Figure 7 above do not appear to follow their respective lines very closely; there is curvature to the data in both graphs. This usually indicates that a variable, perhaps an interaction, is missing from the equation.

The possible interaction between mean income and median income was explored next. Figure 8 illustrates this interaction between mean and median incomes. The mean SPP score is 60 when both mean and median incomes are lowest and 91.3 when mean and median incomes are highest. When mean income stays constant but median income drops, the mean score drops and vice versa.

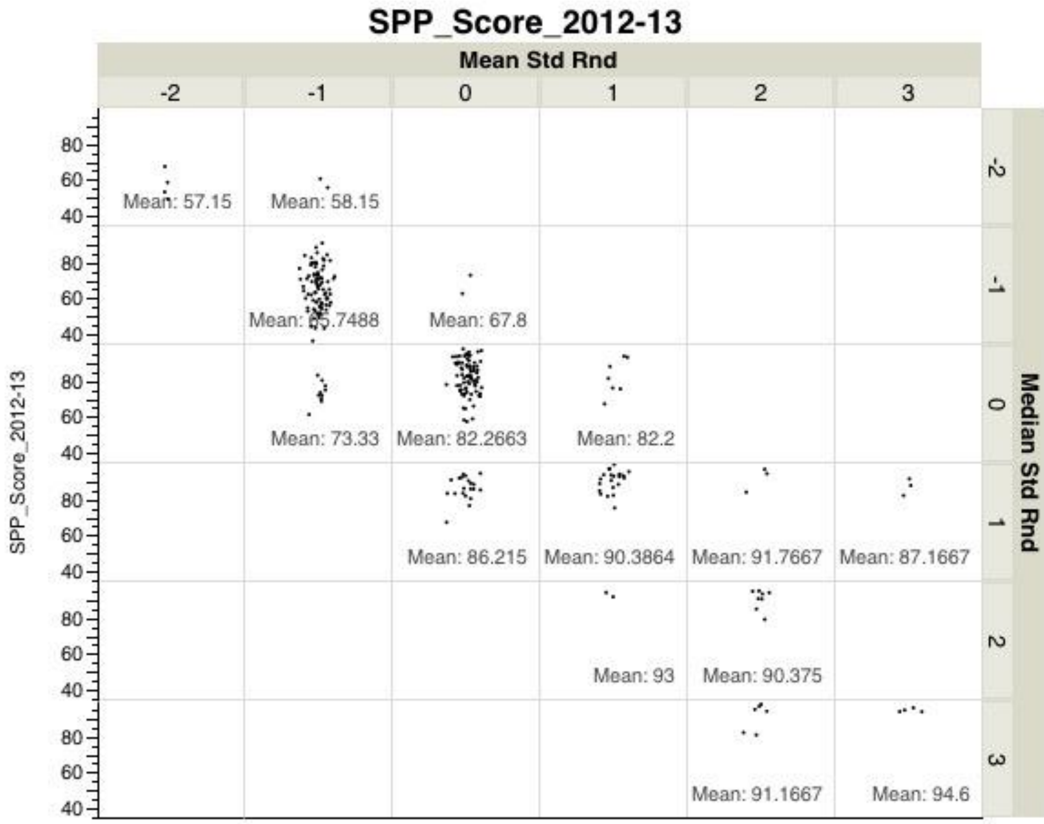


Figure 8 Interaction between Mean and Median Incomes (JMP, 2012)

The interaction between mean and median incomes is not only statistically significant, but also intuitively significant. In many cases, the mean income is much higher than the median, which indicates that a few wealthy families reside in a community that otherwise consists of middle- or low-income families. Including the interaction term, which has a negative coefficient, accounts for this phenomenon by reducing the SPP score estimate and subsequently preventing the predicted SPP score from overestimating the true value for schools with high mean income but low median income.

The linear model. The following multiple linear regression equation involving both mean and median incomes as well as the interaction between the two, i.e., the product of the mean and median incomes, was developed and is analyzed in the "Analysis" section on page 15.

Note that mean and median incomes were scaled by a factor of 10,000 to represent income in terms of $10,000s, and the mean and median incomes in the interaction term were centered on their respective means, 74,974.63 for mean income and 56,395.36 for median income, for easier interpretation of the coefficients. Centering on the mean involves subtracting the overall mean income from each value before multiplying the two.

$$SPP_i = \beta_0 + \beta_1(\mathrm{mean}_i) + \beta_2(\mathrm{median}_i) + \beta_3(mean_i)(\mathrm{median}_i) + \varepsilon_i \tag{3}$$

$$\widehat{SPP}_i = \hat{\beta}_0 + \hat{\beta}_1(\mathrm{mean}_i) + \hat{\beta}_2(\mathrm{median}_i) + \hat{\beta}_3(mean_i)(\mathrm{median}_i) \tag{4}$$

Equation (3) represents the true relationship between SPP scores and the model, where i=1 to 257 to represent the 257 Allegheny County schools, SPP_i is the actual SPP score for the i^{th} school, and ε_i is the difference between the actual SPP score and the predicted SPP score for the i^{th} school. Each β_j, j=0 to 3, represents a model parameter that is a fixed constant, where β_0 is the intercept of the regression equation and β_1, β_2, and β_3 are the coefficients of the three variables: mean income, median income, and the (mean)(median) interaction, respectively. Equation (4) represents the relationship between the predicted SPP scores and the model, where i=1 to 257 and $\widehat{SPP}_i$ is the estimated SPP score for the i^{th} school. Each $\hat{\beta}_j$, j=0 to 3, represent the estimated parameters.

The parameters are determined by the method of least squares. Figure 9 below illustrates the fact that many reasonably good fitting lines could be drawn through the data.

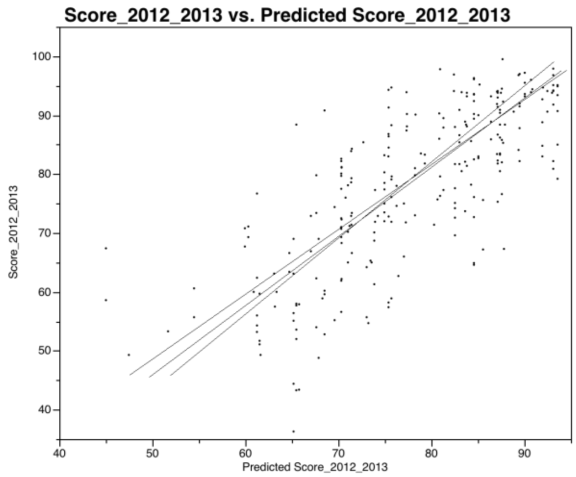
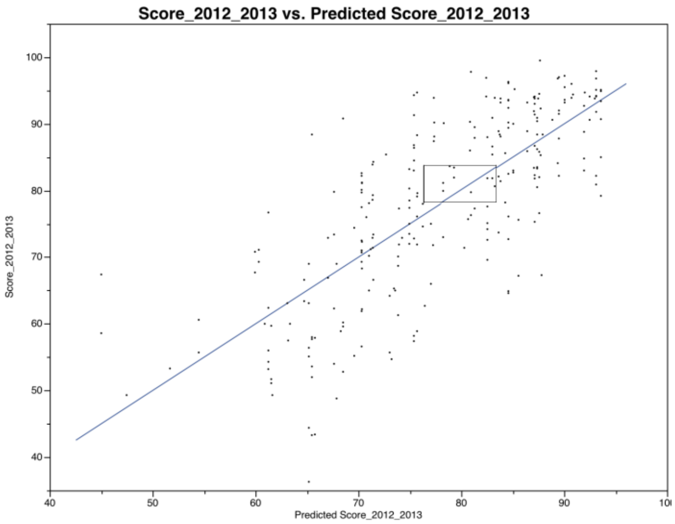
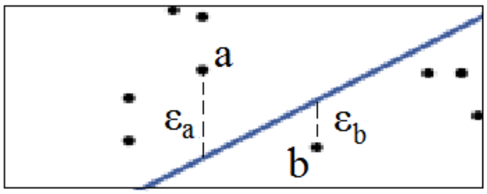


Figure 9 Actual SPP by Predicted SPP with Possible Regression Lines (JMP, 2012)

The least squares method of finding the best fitting line involves finding coefficients for each variable in the model that, together, minimize the amount of unexplained variance. Matrix algebra and Calculus are used by computer software to perform this task, and JMP was used for this study (JMP, 2012). The vertical distance from the point (x_i, y_i) to a line is called the residual error and is denoted with ε_i, where i=1 to 257. The expanded view in Figure 10 below shows point a, which has ε_a=2.5 units, and point b, which has ε_b=-1.3 units, represented by the dotted lines. Points above the line have a positive residual, and points below the line have a negative residual.





Point a: Harrison MS, Baldwin Whitehall SD
Actual SPP 81.8, Predicted SPP 79.3
Residual ε_a=2.5

Point b: Abraham Lincoln El Sch, Bethel Park SD
Actual SPP 79.6, Predicted SPP 80.9
Residual ε_b=-1.3

Figure 10 Actual SPP by Predicted SPP and Exploded View (JMP, 2012)

The residual error is used to calculate the variation in the data that is not explained by the model. To better interpret and compare the error, each residual is 1) squared to obtain uniformly positive errors and 2) summed to obtain an overall, positive measure of the error for each line. The resulting sum is called the sum of squared error and is what JMP uses to select the best fitting line. The line with the least sum of squared error, i.e., the line that explains the most variation in the data, is the best fitting line, which is also called the least squares line.

Equation (4) on page 12 can be rewritten with each parameter, $\hat{\beta}_j$, replaced by its estimated value from the least squares line:

$$\widehat{SPP}_i = 43.34 + 0.99(\text{mean}_i) + 5.34(\text{median}_i) - 0.63(\text{mean}_i)(\text{median}_i). \qquad (5)$$

Having found a least squares line is not proof that there is a significant relationship between SPP scores and the model. Before a significant relationship between SPP scores and income can be confirmed, the model must meet the conditions required by linear regression and be interpreted with respect to a series of statistical tests to determine how well the model describes the data. In addition, any SPP scores that are far away from the average predicted SPP score for a particular combination of mean and median incomes must be evaluated to verify that none of them influence the least squares line, i.e., none of them pull the line away from the overall trend of the data.

Analysis

Conditions for the linear model. The linear model in equation (5) above meets the conditions required for multiple linear regression, which include linearity, constant variance, independence, randomness, and normality, which are addressed in detail below.

Linearity. A positive linear relationship between SPP scores and the model is evident in Figure 11, a plot of the actual SPP scores against the predicted scores from the model. The upward trend of the data indicates that as income increases, on average, actual SPP scores increase.

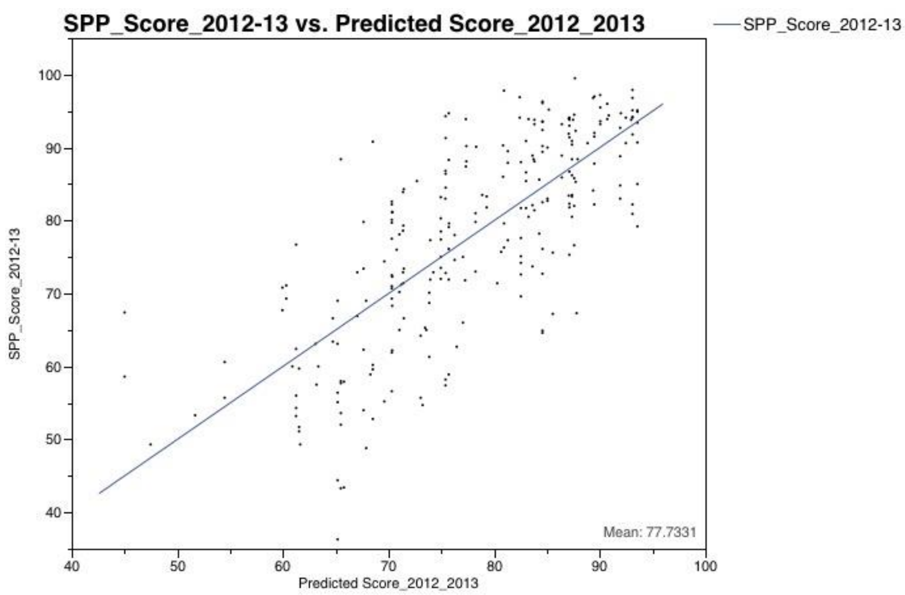


Figure 11 Plot of Actual SPP Scores by Predicted SPP Scores (JMP, 2012)

Constant Variance. The variability in the residuals is the same for each predicted SPP score. Figure 12, the plot of residuals against the fitted values, shows the plotted points to be randomly and roughly evenly spread above and below zero in a band of relatively constant width. This indicates that the variability in the residuals is constant.

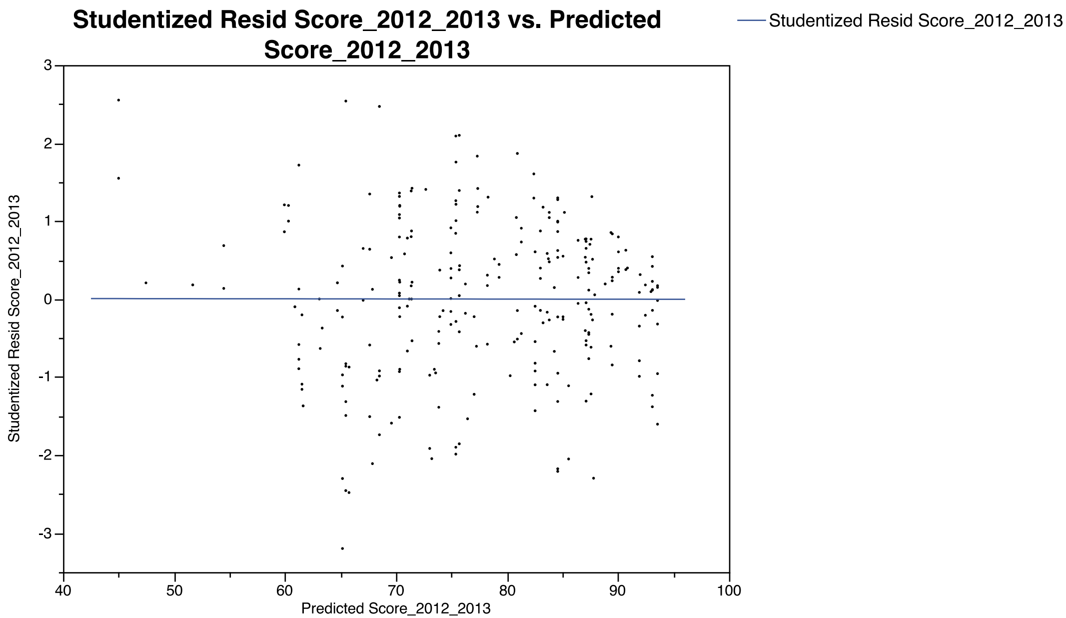


Figure 12 Plot of Residuals by Predicted SPP Scores (JMP, 2012)

Independence. Independence can be inferred from the context of the situation. The actual SPP scores, and therefore, the residuals from the model, are independent. There is no reason to believe that the SPP score of one school depends on that of another.

Randomness. The sample of 257 Allegheny County schools is representative of the population, which includes all counties with similar demographics, tax structures, and forms of government. For the purposes of this study, the sample can be regarded as a random sample.

Normality. In order to use standard distributions to construct and interpret confidence intervals and perform statistical tests, the random errors in the population must be normally distributed. If the sample residuals are independent, random, and normally distributed, the population residuals can be assumed to follow a normal distribution. The Normal Quantile Plot in Figure 13 is a plot of the observed residuals against the theoretical residuals, i.e., the residuals that are expected from a normally distributed sample of the same size. The linear pattern of this plot indicates that the residuals are roughly normally distributed.

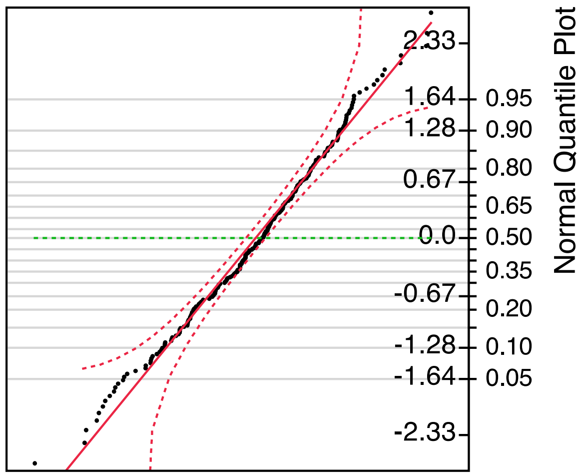


Figure 13 Normal Quantile Plot (JMP, 2012)

Evaluating parameter estimates. Table 3 below shows each parameter estimate from equation (5) on page 14 with its respective standard error, t ratio, and p-value (Prob>|t|). The t ratio is computed as the coefficient divided by its respective standard error. Under the null hypothesis for a specific coefficient, which says the coefficient is equal to zero, the ratio of the coefficient to its standard error follows a t-distribution. The t ratio is then compared to the t-distribution to obtain the p-value, i.e., the probability that a random sample would yield a ratio equal to or more extreme than the one obtained. The level of confidence for the tests outlined in Table 4 below is 95% with error rate $\alpha = 0.05$. If the p-value of a particular coefficient is less than 0.05, that coefficient is statistically significant and contributes to the model.

The intercept, 43.34, can be interpreted as the predicted score for a school if both mean and median incomes were $0.00. While the intercept is essential to the model in that it is used to calculate the predicted SPP scores, the previous interpretation has no practical significance.

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	($\hat{\beta}_0$) 43.338391	1.988584	21.79	<.0001*
MultipleZip_ScaledMean	($\hat{\beta}_1$) 0.9899209	0.666444	1.49	0.1387
MultipleZip_ScaledMedian	($\hat{\beta}_2$) 5.3362228	0.893452	5.97	<.0001*
(MultipleZip_ScaledMedian-5.58444)* (MultipleZip_ScaledMean-7.41662)	($\hat{\beta}_3$) -0.627071	0.091768	-6.83	<.0001*

Table 3 Parameter Estimates (JMP, 2012)

In Table 3 above, JMP uses an asterisk, *, in two ways 1) to indicate multiplication of the two incomes, (MultipleZip_ScaledMedian-5.58444)*(MultipleZip_ScaledMean-7.41662), for the interaction term and 2) to indicate that a p-value is statistically significant.

Table 4 below outlines the tests and results for each coefficient. The median income and the interaction between the mean and median incomes are statistically significant, i.e., each has a p-value less than 0.05. In other words, if the true coefficient for the median term is zero, i.e., the null hypothesis H_0: β_2 = 0 is true, the probability of obtaining the coefficient 5.34 is less than 0.0001. And, if the true coefficient for the interaction term is zero, i.e., the null hypothesis H_0: β_3 = 0 is true, the probability of obtaining the coefficient -0.63 is less than 0.0001. This is strong evidence that the median income and the interaction between the mean and median incomes contribute to the model, thus supporting the linear relationship between SPP scores and incomes. The coefficient for mean income was not statistically significant, however, because mean income is used in the interaction term, it remains in the model. Multicollinearity between mean and median incomes may explain why mean income is insignificant when median income and the interaction of mean and median incomes are added to the model. See “Multicollinearity” on page 24 for an explanation.

95% Confidence Level α = 0.05	Null and Alternative Hypotheses	p-value	Conclusion
Coefficient for Mean $\hat{\beta}_1$: 0.9899209	H_o: $\beta_1 = 0$ H_a: $\beta_1 \neq 0$	0.1387	There is *not* enough evidence to conclude that $\hat{\beta}_1$ is not zero.
Coefficient for Median $\hat{\beta}_2$: 5.3362228	H_o: $\beta_2 = 0$ H_a: $\beta_2 \neq 0$	<0.0001	There is enough evidence to conclude that $\hat{\beta}_2$is not zero.
Coefficient for (Mean)(Median) $\hat{\beta}_3$: -0.627071	H_o: $\beta_3 = 0$ H_a: $\beta_3 \neq 0$	<0.0001	There is enough evidence to conclude that $\hat{\beta}_3$is not zero.

Table 4 Summary of Hypothesis Tests for Estimated Coefficients $\boldsymbol{\beta_1}$, $\boldsymbol{\beta_2}$, and $\boldsymbol{\beta_3}$

Analysis of Variance. Analysis of variance (AOV) is a way to assess the overall fit of the model by comparing the portion of the variance explained by the model to an F distribution to test whether the portion explained by the model is significant. The AOV table, Table 5 below, summarizes the calculations and the test used to determine if the overall model is a good fit for the data. The variation is measured using the sum of squares explained in “The linear model” on page 11.

Source		**DF**		**Sum of Squares**		**Mean Square**	**F Ratio**
Model	(df_{model})	3	(SSModel)	26935.714	(MSModel)	8978.57	109.0311
Error	(df_{error})	253	(SSE)	20834.235	(MSE)	82.35	**Prob > F**
C. Total	(df_{total})	256	(SSTotal)	47769.949			<.0001*

Table 5 Analysis of Variance (JMP, 2012)

Analysis of variance involves partitioning the variability in SPP scores into two components—the portion explained by the model and the portion not explained by the model.

$$SSTotal = SSModel + SSE \tag{6}$$

SSTotal. In equation (6), SSTotal, the total variation in the data, is calculated by obtaining the difference between each SPP score and the mean SPP score, *squaring* each difference, and *summing* those squared differences:

$$\sum_{i=1}^{257}(y_i - \bar{y})^2. \tag{7}$$

In equation (7), y_i is the SPP score for the i^{th} school and $\bar{y}$ is the mean SPP score.

SSModel. In equation (6), SSModel, the variation explained by the model, is calculated by obtaining the difference between each predicted SPP score and the mean SPP score, *squaring* each difference, and *summing* those squared differences:

$$\sum_{i=1}^{257}(\hat{y}_i - \bar{y})^2. \tag{8}$$

In equation (8), $\hat{y}_i$ is the predicted SPP score for the i^{th} school and $\bar{y}$ is the mean SPP score.

SSE. In equation (6), SSE, the variation not explained by the model, is calculated by obtaining the difference between each SPP score and the corresponding predicted SPP score obtained from equation (5), *squaring* each difference, and *summing* those squared differences:

$$\sum_{i=1}^{257}(y_i - \hat{y}_i)^2. \tag{9}$$

In equation (9), y_i is the SPP score for the i^{th} school and $\hat{y}_i$ is the predicted SPP score for the i^{th} school.

Degrees of freedom: df_{model}, df_{error}, df_{total}. The degrees of freedom refer to the number of observations that are free to vary. For example, in calculating the SSTotal in equation (7), the sample mean SPP score was subtracted from each SPP score. This sample mean SPP score is an estimate of the population mean SPP score, which is now a fixed value. The SPP scores for only 256 schools are free to vary because once 256 of the scores are plugged into equation (10) below, the 257^{th} score must be equal to a specific value that makes the equation true:

$$77.7 = \frac{x_1 + x_2 + x_3 + x_4 + \cdots + x_{256} + x_{257}}{257}. \quad (10)$$

In equation (10), 77.7 is the mean SPP score for the sample. The degrees of freedom for the SSTotal is n-1, where n is the sample size. For this model the total degrees of freedom, df_{total}, are n-1=256. The parameter estimates for the model place additional, similar constraints on the degrees of freedom. With three estimated parameters, β_1, β_2, and β_3, the degrees of freedom for the model, df_{model}, is 3. The degrees of freedom for the error, df_{error}=253, is the model degrees of freedom subtracted from the total degrees of freedom. Just as with the sum of squares components, the model degrees of freedom plus the error degrees of freedom equal the total degrees of freedom.

MSModel. The mean square for the model is the average sum of squares for the model, which is calculated by dividing SSModel by df_{model} as in equation (11) below.

$$MSModel = \frac{SSModel}{df_{model}} = \frac{26935.714}{3} = 8978.57 \quad (11)$$

MSE. The mean square error is the average sum of squares for the error, which is calculated by dividing SSE by df_{error} as in equation (12) below.

$$MSE = \frac{SSE}{df_{error}} = \frac{20834.235}{253} = 82.35 \quad (12)$$

F Ratio. The F ratio follows an F distribution when the null hypothesis, which says all coefficients are equal to zero, is true. The F ratio is calculated as the ratio of MSModel to MSE: $F = \frac{MSModel}{MSE} = 109.0311$. This value is compared to the F distribution using the df_{model}=3 and df_{error}=253 as the numerator and denominator degrees of freedom, respectively. The null and alternative hypotheses for this test follow:

H_o: $\beta_1 = \beta_2 = \beta_3 = 0$

H_a: at least one $\beta_j \neq 0$

The null hypothesis, H_o, says that all coefficients are equal to zero, which means there is no linear relationship between SPP scores and the model. The alternative hypothesis, H_a, says that at least one coefficient is not equal to zero; if true, the alternative hypothesis says that the portion of variability explained by the model is significant.

Prob>F. The Prob>F, or p-value, is the probability of obtaining a sample with this F ratio or a more extreme F ratio when the null hypothesis is true. If all coefficients are equal to zero, i.e., the null hypothesis is true, the probability of obtaining a sample with F Ratio≥109.0311 is less than 0.0001. In other words, the probability of obtaining a random sample that explains as much or more of the variability in SPP scores is less than 0.0001. With an error rate of $\alpha=0.05$, the AOV test supports rejecting the null hypothesis. There is enough evidence to conclude that at least one coefficient, β_1, β_2, or β_3, does not equal zero. This means that the model is effective in explaining the variability in SPP scores and further supports the inclusion of all three terms.

Summary of Fit. The coefficient of determination, R^2 (RSquare in Table 6 below) and the root mean square error (RMSE) each provide additional ways of assessing how well the model fits the data. Table 6 below summarizes these values.

RSquare	(R^2)	0.563863
RSquare Adj	(Adj R^2)	0.558692
Root Mean Square Error	(RMSE)	9.074621
Mean of Response	($\bar{y}$)	77.73307
Observations (or Sum Wgts)		257

Table 6 Summary of Fit (JMP, 2012)

The coefficient of determination, R^2, is a ratio of the model variation to the total variation:

$$R^2 = \frac{variability\ explained\ by\ the\ model}{total\ variability\ in\ SPP\ scores} = \frac{SSTotal - SSE}{SSTotal} = \frac{26935.714}{47769.949} = 0.563863 \qquad (13)$$

In equation (13), SSTotal is the total variation in SPP scores from equation (7) on page 20, and SSE is the variation in SPP scores not explained by the model from equation (9) on page 20.

The R^2 value in equation (13) is the unadjusted R^2 for the model. The coefficient of determination increases each time a predictor is added to the model. A model with too many predictors tends to model the random noise in the data rather than actual patterns, which is called "overfitting." Because the model represented by equation (5) on page 14 has three predictor variables: mean income, median income, and (mean×median) income, an adjustment is necessary to prevent this "overfitting" of the model. The Adjusted R^2 (RSquare Adj in Table 6 above) accounts for the number of predictors in the model and is the value used in interpreting the ratio of variability. When a new predictor enters a model, the adjusted R^2 only increases if the additional variability that is explained by the new term is more than what would be expected by chance. The Adjusted R^2 for the model represented by equation (5) is 0.558692, or ~56%. This means that the model explains approximately 56% of the variability in SPP scores. Because the model is a function of the mean and median incomes associated with each school, income explains approximately 56% of the variability in SPP scores.

The root mean square error (RMSE) is a measure of error for the model that can be compared directly to the standard deviation of the SPP scores.

$$RMSE = \sqrt{MSE}. \qquad (14)$$

Equation (14) can be rewritten in terms of SSE and degrees of freedom using equation (12) on page 21 in place of MSE.

$$RMSE = \sqrt{\frac{SSE}{df_{error}}} = \sqrt{\frac{20834.235}{253}} = \sqrt{82.35} = 9.074621. \tag{15}$$

The difference between the standard deviation of SPP scores, 13.7, and the RMSE is 4.63, which means the deviation in the predicted SPP scores is less than that of the actual SPP scores. This reduced error, together with an R^2 of 56%, indicates that the model is a good fit for the data.

Multicollinearity. One or more independent variables in a multiple regression model may be correlated with another independent variable, which in this study indicates that mean income may be correlated with median income. If mean and median incomes are highly correlated, the overall predictive power of the model is not affected. However, multicollinearity may explain why the coefficient for mean income becomes insignificant when median income and the interaction between mean and median income are in the model. The mean income explains some of the same variation already accounted for by the median income. One way to test for multicollinearity is to compute the variance inflation factor (VIF) for the two explanatory variables, mean and median incomes.

$$VIF = \frac{1}{1-R^2}. \tag{16}$$

In equation (16), the R^2 in the denominator is the coefficient of determination for a model that predicts one of the explanatory variables from the other(s). In this study, the R^2 for a model using median income to predict mean income is 0.88 resulting in a VIF=8.33. Because a VIF greater than 5 indicates multicollinearity, mean and median incomes are highly correlated. Despite this correlation, mean income remains in the model because 1) multicollinearity does not affect the power of the model to explain the variability in SPP scores and 2) an interaction term involving mean income is present in the model.

Model outliers and influential points. Model outliers are those schools with SPP scores that are far away from their predicted scores, and therefore, do not follow the linear trend of the other schools. The magnitude of the residuals for these schools is large; a point with a residual more than two standard deviations away from the residual mean of zero is questionable. This extreme distance may influence the least squares line by pulling the line toward the outlying points, which places the parameter estimates in question. To assess the magnitude of the residuals, JMP software (JMP, 2012) was used to calculate the studentized residuals, which are plotted against the predicted SPP scores in Figure 14 below. Fifteen schools have residuals more than two standard deviations from the residual mean of zero; ten of these schools have studentized residuals greater than two standard deviations below the residual mean of zero, and five of them have studentized residuals greater than two standard deviations above the residual mean of zero. These appear as the larger black dots in Figure 14.

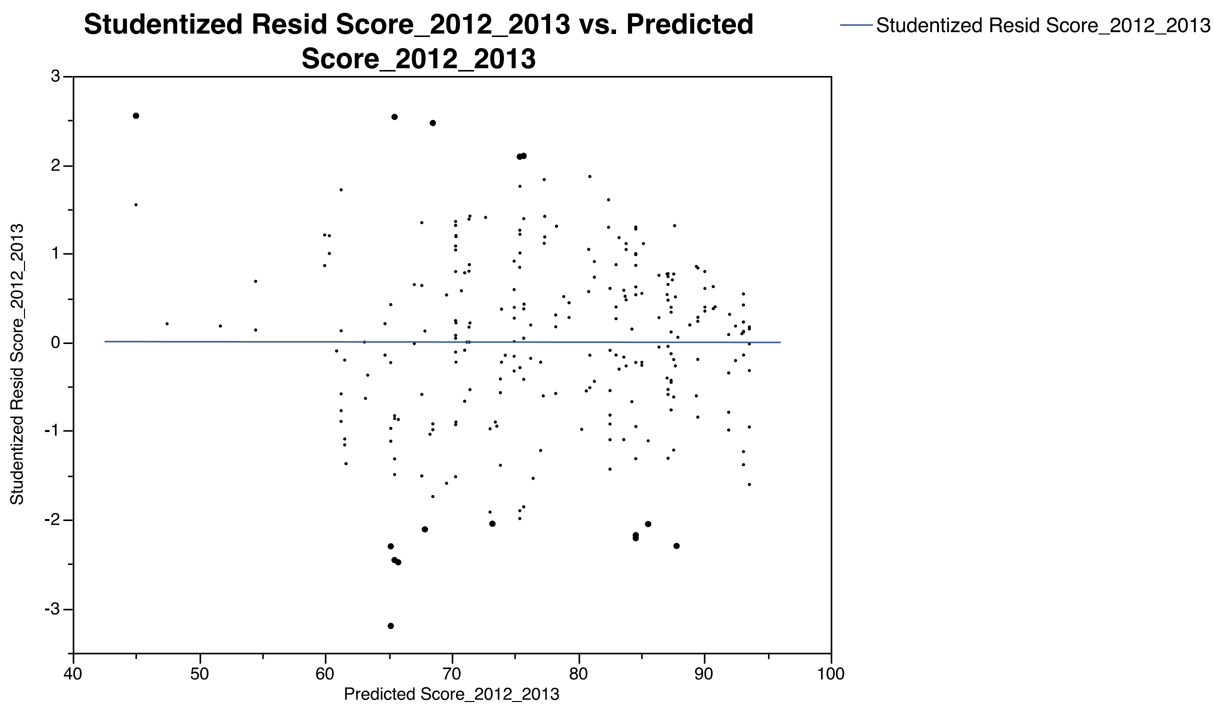


Figure 14 Studentized Residuals for the Model

Although the aforementioned fifteen schools have extreme residuals, the schools that are influential are those whose SPP score is far from the average predicted SPP score of 77.7, i.e., they appear far to the right or left on a graph where the predicted SPP score is the horizontal axis. The distance a school's SPP score is from the mean of the predicted values is called the leverage; and the farther away a school's SPP score is from this mean, the higher the leverage. Points with high leverage have more potential to influence the line.

To assess whether schools influence the line away from the general trend, both the studentized residuals and the leverage were considered. These two properties are represented by Cook's Distance, which was calculated for each point using JMP software (JMP, 2012). The Cook's Distance (D_i) for the i^{th} school is given by

$$D_i = \left(\frac{stdres_i^{\,2}}{k+1}\right)\left(\frac{h_i}{1-h_i}\right). \quad (17)$$

In equation (17), $stdres_i$ is the studentized residual for the i^{th} school; k is 3, the number of predictors; and h_i is the leverage for the i^{th} school. The outliers for the model are outlined in Table 7 below along with their studentized residuals and Cook's Distance (Cook's D Influence) from JMP (JMP, 2012).

School District	School Name	SPP Score	Predicted SPP Score	Studentized Residuals	Cook's Distance
Cornell	Cornell El	64.6	84.6	-2.209	0.0090
Cornell	Cornell HS	64.9	84.6	-2.176	0.0087
Keystone Oaks	Keystone Oaks MS	67.3	87.8	-2.296	0.0441
Montour	J W Burkett El Sch	90.8	68.5	2.470	0.0160
Mt Lebanon	Lincoln El Sch	94.7	75.7	2.099	0.0055
Pittsburgh	Academy at Westinghouse 6-12	43.4	65.8	-2.481	0.0204
Pittsburgh	Pittsburgh Brashear HS	54.7	73.2	-2.046	0.0054

School District	School Name	SPP Score	Predicted SPP Score	Studentized Residuals	Cook's Distance
Pittsburgh	Pittsburgh Minadeo K-5	67.2	85.6	-2.050	0.0293
Pittsburgh	Pittsburgh Perry HS	43.3	65.5	-2.456	0.0155
Pittsburgh	Pittsburgh Weil K-5	67.4	45	2.551	0.1113
Shaler Area	Reserve Primary Sch	88.4	65.5	2.539	0.0162
West Jefferson Hills	Gill Hall El Sch	94.3	75.4	2.090	0.0067
Wilkinsburg Borough	Wilkinsburg MS	44.4	65.2	-2.301	0.0134
Wilkinsburg Borough	Wilkinsburg SHS	36.3	65.2	-3.198	0.0259
Woodland Hills	Woodland Hills JHS	48.8	67.9	-2.110	0.0085

Table 7 Table of Outliers with Respect to the Model

A Cook's Distance greater than 0.5 indicates a moderately influential point, and a Cook's Distance greater than 1 indicates a very strong influential point. None of the outliers in Table 7 have Cook's Distances greater than 0.1113. Therefore, the outliers for the model do not influence the fit of the least squares line.

Conclusion

Results. The following model is the best model for describing the relationship between income data and School Performance Profile scores.

$$\widehat{SPP}_i = 43.34 + 0.99(\text{mean}_i) + 5.34(\text{median}_i) - 0.63(\text{mean}_i)(\text{median}_i) \tag{18}$$

In equation (18), mean and median incomes are each scaled by a factor of 10,000, and for the interaction term, the average of each type of income is subtracted from the respective value, i.e., the average mean family income is subtracted from the mean value and the average median family income is subtracted from the median value, for each school.

Table 8 below summarizes the fit of the model in equation (18). Together, mean and median family incomes explain approximately 56% of the variation in SPP scores. The root mean square error of 9.075 is a reduction in error of about 4.625 units from 13.7, the standard deviation of SPP scores without the model.

RSquare	0.563863
RSquare Adj	0.558692
Root Mean Square Error	9.074621
Mean of Response	77.73307
Observations (or Sum Wgts)	257

Table 8 Summary of Fit (JMP, 2012)

Limitations. The limitations of this study involve certain assumptions that had to be made regarding the multiple-zip estimate for income data. For each elementary school, it was assumed that students attend the school whose physical zip matches their zip code of residence. From this assumption, a weighted average of income for all elementary schools that feed into middle and high schools was computed for each middle and high school. For perfectly accurate data, it would have been necessary to obtain a list of students for each school to determine the exact proportion of students from each zip code. A weighted average could then be calculated from knowing these exact proportions. It was neither realistic within the scope of this study nor possible with current security concerns to obtain such a list from any school.

Given the aforementioned constraints, the income data should be considered fairly accurate. In many cases, smaller children do attend their neighborhood school. Any deviation from this may be considered not to effect the estimation in a significant way.

Interpretation and Application. This study is consistent with findings from two recent studies involving the relationship between income and student achievement as measured by standardized test scores: The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit (Dahl and Lochner, 2010) and The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations. (Reardon, 2011) Because 90% of the SPP score is derived from student standardized test scores (see “Appendix A: School Performance Profile Score” on page 44), these two studies are not only relevant to the current study but are also necessary for interpreting the relationship between SPP scores and income and applying this new understanding in a way that will benefit students and schools.

According to the PA School Performance website, “The Pennsylvania School Performance Profile (SPP) is an *integral* part of the Educator Effectiveness System (teacher and principal evaluation).” (2013c, p. 1, emphasis added) SPP scores account for 15% of teacher and principal evaluations. This means that teachers who work in a failing school will find it challenging or difficult to earn greater than 85% on their individual effectiveness score. The results of this study require a second look at how teachers and administrators are evaluated. Upon initial inspection of the distribution of SPP scores, it is tempting to conclude that the existence of failing schools necessitates school choice through vouchers. But this is only true if failing schools happen to employ poor teachers and administrators. If there are non-school-related factors causing or contributing to school failure, one cannot conclude that vouchers are the solution. In exercising the right to school choice, factors outside of school will not change.

The fact that SPP scores are highly correlated with family income raises the question, “What factor(s) cause or contribute to a school receiving a ‘failing’ SPP score?” If the relationship between SPP scores and income is not simply because failing schools employ poor teachers and administrators, any use of that score to evaluate teachers and administrators is discriminatory. There are several possible non-school-related factors that contribute to student achievement directly, as measured by standardized test scores, and to SPP scores indirectly, because the SPP score is primarily derived from standardized test scores.

Dahl and Lochner (2010) found that an increase in income resulted in improved math and reading achievement. By studying the effects of the Earned Income Tax Credit (EITC) government program that provides extra income to low-income families, Dahl and Lochner were able to control for extraneous variables that might be correlated with the outcome, such as adverse or unstable living conditions, and estimated that “a $1,000 increase in family income raises math and reading test scores by about 6% of a standard deviation.” (p. 2) Reardon (2011) found many possible non-school-related factors that contribute to low achievement. Reardon explains that an increase in understanding and knowledge of the cognitive abilities and development of young children has led parents to nurture the cognitive development of their children from birth through the preschool years by exposing them to intellectually, socially, and emotionally stimulating activities. The extent to which parents nurture the cognitive development of their children is proportional to the resources secured by the parents, i.e., parents with higher education have higher incomes, and thus, more resources to invest in their children’s cognitive development. (p. 17-19) Reardon also states, “the income achievement gap is large when children enter kindergarten and does not appear to grow (or narrow) appreciably as children progress through school.” (p. 1) If the gap exists before children enter school and does not increase or decrease as children progress through school, it is reasonable and necessary to consider that factors outside of school, and beyond the control of teachers and administrators, contribute significantly to the income achievement gap.

The results and interpretation of this study do not support the cessation of all attempts to close the achievement gap between high- and low-income students. On the contrary, it serves to promote further, more effective work toward that end. Without an understanding of what is clearly not working to solve a problem, a new solution cannot be discovered and implemented. While school reformers continue to ignore the compelling evidence that factors outside of school play a significant role in student achievement, and as long as they continue to put more and more pressure on teachers and administrators to overcome these factors that are beyond their control, the system will fail. There are certainly ineffective teachers within the system; however, using student standardized test scores, by way of SPP scores, is an ineffective way to find them.

There is enough evidence from the data in this study to conclude that SPP scores, like student standardized test scores, are highly correlated with income. As long as there is reasonable doubt that this correlation is related to factors outside of school, SPP scores should not be used to evaluate teachers and administrators. Ninety-percent of each SPP score is derived from student standardized test scores—forty-percent directly from student standardized test scores, ten-percent from indicators of closing the achievement gap with student standardized test scores as indicators, and forty-percent from indicators of student growth with student standardized test scores as indicators. The Department of Education claims that the PVAAS system for measuring student growth is independent of student demographics. However, both Dahl and Lochner (2010) and Reardon (2011) have found that student achievement is dependent on factors outside of school, i.e., poor income and/or factors related to poor income contribute to poor achievement. If this is true, student growth will also be affected by factors outside of school and outside the control of schools, administrators, and teachers.

The studies by Dahl and Lochner (2010) and Reardon (2011) have established reasonable doubt that poor schools are the reason students fail. If poor income and/or factors related to poor income contribute to poor achievement, then using SPP scores to evaluate teachers is arbitrary and ineffective. There will be effective teachers in low-performing schools who are demoted, reprimanded, refused compensation, and/or professionally scarred. Likewise, there will be ineffective teachers in high-scoring schools who will receive the compensation and professional accolades that should be awarded only to effective teachers, such as the aforementioned effective teacher in the low-performing school. The act of judging an individual based on inaccurate and arbitrary data that is beyond their control is discrimination. Furthermore, valuable and scarce resources will be designated to provide training to improve the instruction of an already effective teacher. These wasted resources would better serve students if used in areas where improvements are actually needed.

In conclusion, this study supports the correlation between income and SPP scores and brings a necessary, new awareness to the persistence of the achievement gap. If school reformers recognize the correlation between income and SPP scores and recognize that income or income-related factors contribute significantly to student achievement, better decisions can be made regarding evaluation of teachers and administrators and the distribution of resources.

Annotated Bibliography

Carson Middle School Building Secretary. North Allegheny School District, personal communication November 24, 2014.

This personal communication improved estimates of mean and median incomes for communities served by North Allegheny School District. The district's schools were grouped by feeder pattern so the weighted mean and median incomes of elementary schools in each feeder pattern could be calculated for the respective middle school(s). The Carson MS feeder pattern includes McKnight, Hosack, and Peebles Elementary Schools; the Marshall MS feeder pattern includes Marshall and Bradford Woods Elementary Schools; and the Ingomar MS feeder pattern includes Franklin and Ingomar Elementary Schools.

Dahl, Gordon and Lochner, Lance. (2010, November). *The impact of family income on child achievement: Evidence from the Earned Income Tax Credit.* by Gordon Dahl and Lance Lochner. Retrieved from http://www.nber.org/papers/w14599.

Dahl and Lochner estimate the effect that family income has on academic achievement of children. They claim that their study was designed to measure the causal relationship between income and achievement. Income as a cause for academic achievement had not been addressed until this study because of the difficulty in separating the effects of income from the effects of many potentially extraneous factors such as adverse living conditions or relationships and changes in family circumstances or dynamics. This study takes advantage of the Earned Income Tax Credit (EITC) government program that provides extra income to low-income families. The EITC was expanded in the late 1980s and 1990s, which included as increase in both the maximum benefit amount and the

range of family income that qualified families to receive the EITC. As stated in the study publication, these expansions of the EITC provided an “exogenous source of income variation for American families that [Dahl and Lochner] use to identify the effects of family income on child achievement.” (p. 1) From longitudinal data on nearly 4,500 children matched to their mothers in the Children of the National Longitudinal Survey of Youth (NLSY), Dahl and Lochner had access to income and demographic data as well as to five repeated measures of math and reading test scores for each child. This longitudinal design allowed them to control for extraneous variables that cannot be observed and measured. The estimates from this study “suggest that current income has a significant effect on a child’s math and reading achievement — a $1,000 increase in family income raises math and reading test scores by about 6% of a standard deviation.” (p. 2) The significance of this study to the current study lies in the use of standardized achievement tests in the computation of SPP scores; a large portion of each SPP score is derived from student standardized test scores. Although the authors claim that poor income causes poor achievement, the current study addresses income as a contributing factor to poor achievement because causation is difficult to establish.

Duquesne City School District. (2009). *Duquesne City School District website.* Accessed November 24, 2014 from http://www.dukecitysd.org/content.aspx?id=47.

The Duquesne City School District website confirmed that Duquesne City students in grades 7-12 attend either East Allegheny or West Mifflin School District. Both the high school and middle school in each of these districts received a multiple-zip estimate that includes the income from 15110, the zip code of Duquesne City Elementary School.

Jefferson Middle School Building Secretary. Mt. Lebanon School District, personal communication November 24, 2014.

This personal communication improved estimates of mean and median incomes for communities served by Mt. Lebanon School District. The district's schools were grouped by feeder pattern so the weighted mean and median incomes of elementary schools in each feeder pattern could be calculated for the respective middle school(s). The Jefferson MS feeder pattern includes Jefferson, Lincoln, Hoover, and Markham Elementary Schools; and the Mellon MS feeder pattern includes Foster, Howe, and Washington Elementary Schools.

JMP Pro 10. (2012). *JMP®* (10.0.0 64-bit Edition) [computer software] Duquesne University Mathematics and Computer Science department Site ID 70123647: SAS Publishing.

JMP Pro 10 was used to create the linear model and to produce the supporting tables and figures.

O'Connell, John. Pittsburgh Public Schools, personal communication. October 3, 2014.

This personal communication improved estimates of the mean and median incomes for communities served by Pittsburgh Public Schools (PPS). The district's schools were grouped by feeder pattern so the weighted mean and median incomes of elementary

schools in each feeder pattern could be calculated for the respective middle school(s) and high school. Additionally, three schools in the district had incorrect grades listed after the school names. This communication allowed for this error to be corrected. Specifically, Pittsburgh Morrow K-5, Pittsburgh Weil K-8, and Pittsburgh Montessori K-8 were changed to Pittsburgh Morrow K-8, Pittsburgh Weil K-5, and Pittsburgh Montessori K-5, respectively. The following table summarizes the information gained from this communication:

School Name	Feeds into MS	Feeds into HS	Magnet
Pittsburgh CAPA 6-12			Magnet
Pittsburgh Miller K-5	Arsenal	Milliones	
Pittsburgh Milliones 6-12			
Pittsburgh Weil K-5	Arsenal	Milliones	
Acadamy at Westinghouse 6-12			
Pittsburgh Allderdice HS			
Pittsburgh Arsenal 6-8		Milliones	
Pittsburgh Arsenal K-5	Arsenal	Milliones or Westinghouse	
Pittsburgh Colfax K-8		Allderdice	
Pittsburgh Dilworth K-5			Magnet
Pittsburgh Faison K-5		Westinghouse	
Pittsburgh Fulton K-5	Arsenal	Milliones	Half Magnet
Pittsburgh Greenfield K-8		Allderdice	
Pittsburgh Liberty K-5			Magnet
Pittsburgh Lincoln K-5		Westinghouse	
Pittsburgh Linden K-5			Magnet
Pittsburgh Minadeo K-5	Sterrett	Allderdice	
Pittsburgh Montessori K-5			Magnet

School Name	Feeds into MS	Feeds into HS	Magnet
Pittsburgh Obama 6-12			Magnet
Pittsburgh Science and Technology Academ			Magnet
Pittsburgh Sterrett 6-8		Allderdice	
Pittsburgh Sunnyside K-8		Milliones	
Pittsburgh Woolslair K-5	Arsenal	Milliones	
Pittsburgh Allegheny 6-8			Magnet
Pittsburgh Allegheny K-5			Magnet
Pittsburgh King K-8		Perry	
Pittsburgh Manchester K-8		Perry	
Pittsburgh Morrow K-8		Perry	
Pittsburgh Perry HS			
Pittsburgh Schiller 6-8		Perry	Half Magnet
Pittsburgh Spring Hill K-5	Shiller	Perry	
Pittsburgh Arlington K-8		Carrick	
Pittsburgh Beechwood K-5	South Hills	Carrick	
Pittsburgh Brashear HS			
Pittsburgh Brookline K-8		Brashear or Carrick	
Pittsburgh Carmalt K-8			Magnet
Pittsburgh Carrick HS			
Pittsburgh Concord K-5	South Brook	Carrick	
Pittsburgh Grandview K-5	South Hills	Brashear	
Pittsburgh Mifflin K-8		Allderdice	
Pittsburgh Phillips K-5			Magnet
Pittsburgh Roosevelt K -5	South Brook	Carrick	
Pittsburgh South Brook 6-8		Brashear or Carrick	

School Name	Feeds into MS	Feeds into HS	Magnet
Pittsburgh South Hills 6-8		Brashear	
Pittsburgh West Liberty K-5	South Brook	Brashear or Carrick	
Pittsburgh Whittier K-5	South Hills	Brashear	
Pittsburgh Banksville K-5	South Hills	Brashear	
Pittsburgh Classical 6-8			Magnet
Pittsburgh Langley K-8		Brashear	
Pittsburgh Westwood K-5	South Hills	Brashear	

PA Department of Education. (2014). *Pennsylvania Value Added Assessment System (PVAAS).* Retrieved December 28, 2014 from http://www.portal.state.pa.us/portal/server.pt/community/state_assessment_system/20965/pennsylvania_value_added_assessment_system_(pvaas)/1426500.

The PA Department of Education website provides a detailed description of the new PA Value Added Assessment System, PVAAS, which measures growth of students by comparing current standardized test scores, such as from the PSSA, to prior test scores.

PA School Performance. (2013a). *2012-2013 SPP Scores – School List* [data file]. Retrieved August 6, 2014 from http://paschoolperformance.org/Downloads.

The file "2012.2013.SPP.Scores.School.List.12.11.13.xlsx" contains the county, LEA name, school name, AUN, school number, and 2012-2013 SPP score for all schools in Pennsylvania. The data in this file was used to select schools in Allegheny County and to retrieve the school data necessary for this study, i.e., school names and numbers with corresponding SPP scores.

PA School Performance. (2013b). *Data for SY 2012-2013* [data file]. Retrieved August 6, 2014 from http://paschoolperformance.org/Downloads.

The folder *SPP* contains two text (.txt) files in a pipe-delimited format. The file of interest to this study was "SPP.FF.2012.2013.txt," which contains fast fact data for each school. School zip codes were retrieved from this file for use in matching income data from the Census Bureau to schools. The file "SPP.APD.2012.2013.txt" contains academic performance data, including the SPP score for each school (stored as data element "Final Academic Score").

PA School Performance. (2013d). *Pennsylvania School Performance Profile: Mt. Lebanon SHS.* Retrieved October 28, 2014 from http://paschoolperformance.org/Profile/4377.

School performance data for Mt. Lebanon SHS illustrate the SPP score dependence on standardized test scores. At this link, selecting the "Academic Performance" tab and the "View Calculation" button reveals the break-down of standardized test scores that make up the SPP score for a school. This data is necessary for understanding SPP scores and how they are determined.

PA School Performance. (2013c). *Pennsylvania School Performance Profile Frequently Asked Questions.* Retrieved October 25, 2014 from http://paschoolperformance.org/FAQ.

This document explains how the PA School Performance Profile (SPP) is calculated and used. The SPP score is a quantitative score assigned to each school in Pennsylvania to summarize the school's performance as indicated by a number of academic measures. Much like grades given to students, the SPP score is measured on a 100-point scale. SPP scores are a significant part of Pennsylvania's Educator Effectiveness System, i.e., the system that evaluates teacher and principal performance.

Reardon, Sean F. (July 2011). *The widening academic achievement gap between the rich and the poor: New evidence and possible explanations.* Retrieved from http://cepa.stanford.edu/content/widening-academic-achievement-gap-between-rich-and-poor-new-evidence-and-possible.

Sean Reardon is a professor of education and sociology at Stanford University where he has studied the correlation between income and student achievement as measured by standardized test scores. In this document, chapter 5 from the book *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, 2011, New York: Russell Sage Foundation Press, Reardon compares the achievement of students in the 90th percentile of the income distribution with those in the 10th percentile. The average difference in achievement, called the income achievement gap, has grown to be twice as large as the gap between African American and Caucasian students, a race-related gap that has been targeted for improvement for the past forty years. (p. 1) Reardon explores four possible causes for the increase of the income achievement gap: 1) the difference in income between families in the 90th and 10th percentiles of the income distribution has increased;

2) high-income families invest more in the cognitive development of their children compared to low-income families; 3) the positive correlation between income and other socioeconomic factors that benefit children has increased enabling high-income families to secure related resources; and 4) segregation by income, i.e., rich and poor live in vastly different geographic locations, has created similar differences in the quality of schools and availability of resources because wealthy families can obtain more resources for their schools. (p. 13) In summary, Reardon concludes that 1) rising income inequality does not explain the rising income achievement gap, rather, the dollar simply "buys more achievement" than it once did so high-income families can buy more opportunities for the cognitive growth of their children (p. 17); 2) parents with high incomes invest more in the cognitive development of their children by deliberately focusing their children's activities around "intellectual and socio-emotional development" (p. 19); 3) the structure of the family has become divided into two extremes with low-income mothers more likely to be young, unemployed, and single/divorced and high-income mothers are more likely to be older, employed, and married. "This polarization in family structure implies a corresponding polarization in key resources (income, parental time) available for children, which may have important implications for the distribution of children's academic achievement" (p. 21); and 4) there is not enough evidence to determine whether an increase in income segregation has created a difference in the quality of schools and thus, caused the increase in the income achievement gap. (p. 24) The significance of this study to the current study lies in the use of standardized achievement tests in the computation of SPP scores; a large portion of each SPP score is derived from student standardized test scores.

U.S. Census Bureau. (2012a). *Mean income in the past 12 months (in 2012 inflation-adjusted dollars) 2008-2012 American Community Survey 5-year estimates* [data file]. Retrieved September 5, 2014 from http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_12_5YR_S1903&prodType=table.

The folder *ACS_12_5YR_S1902* contains various files relating to the American Community Survey (ACS) 5-year estimates of mean income in the past 12 months. The file containing the data of interest to this study is "ACS_12_5YR_S1902_with_ann.csv," a comma separated file with numerous fields of mean income and demographic data by zip code. The income data used in this study was retrieved from the column "Mean income (dollars); Estimate; All households," and corresponding zip codes were retrieved from the column "Id2."

U.S. Census Bureau. (2012b). *Median income in the past 12 months (in 2012 inflation-adjusted dollars) 2008-2012 American Community Survey 5-year estimates* [data file]. Retrieved September 5, 2014 from http://factfinder2.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_12_5YR_S1902&prodType=table.

The folder *ACS_12_5YR_S1903* contains various files relating to the American Community Survey (ACS) 5-year estimates of median income in the past 12 months. The file containing the data of interest to this study is "ACS_12_5YR_S1903_with_ann.csv," a comma separated file with numerous fields of median income and demographic data by zip code. The income data used in this study was retrieved from the column "Median income (dollars); Estimate; Households," and corresponding zip codes were retrieved from the column "Id2."

U.S. Census Bureau. (2008). A compass for understanding and using American Community Survey data: What general data users need to know. Retrieved September 5, 2014 from http://www.census.gov/acs/www/Downloads/handbooks/ACSGeneralHandbook.pdf.

This document was referenced for its information on how to select and use American Community Survey (ACS) data. "Appendix 1: Understanding and Using ACS Single-Year and Multiyear Estimates" starting on page A-1 provides details about each type of estimate and guidelines for choosing which type to use. The section "Differences Between Single-Year and Multi- year ACS Estimates" states that multi-year estimates are more reliable than single while single-year estimates are more current. The sample sizes in the multiyear estimates are larger and accounts for more precise estimates. Because this study is interested in the relative differences between incomes, precise data is more important than current data; therefore, the 5-year income estimates were used. (p. A-2, A-3)

Zipmap. (2014). Scott Township zip code boundaries. Accessed September 10, 2014 from http://www.zipmap.net/Pennsylvania/Allegheny_County/Scott_Township.htm.

This website confirmed that Chartiers Valley School District encompasses five zip codes. Four zip codes were used in the multiple-zip estimate: 15243, 15220, 15106, and 15017. The portion of zip code 15216 within Scott Township was not significant enough to include it.

Appendix A: School Performance Profile Score

According to the PA School Performance website, SPP scores are weighted heavily on standardized test scores. Below is the break down of how SPP scoring was determined for each school in the 2012-2013 school year with specific illustrations showing how the SPP score was determined for Mt. Lebanon SHS. Figure 15 below is a copy of the key for symbols used to communicate visually a score for each indicator before it is entered into the calculation. (PA School Performance, 2013d)

Figure 15 Key for Visual Interpretation of SPP Score Indicators and SPP Scores (PA School Performance, 2013d)

Forty-percent of each school's SPP score is derived from indicators of academic performance, which include

- Percent proficient or advanced on various PSSA tests and/or Keystone exams
- Percent competent or advanced on industry certification exams, if applicable
- Percent meeting college readiness benchmarks measured by the SAT and ACT exams, if applicable

Figure 16 below breaks down academic achievement indicators for Mt. Lebanon SHS. Thirty-percent of the SPP score is based on various PSSA and Keystone exams. Each exam—Mathematics/Algebra I, Reading/Literature, Science/Biology, and Writing—accounts for 7.5% of the score. The Writing PSSA score is not applicable, but the website is unclear about how the 7.5% for writing is distributed to reach the 40% total for indicators of academic achievement. Industry Standards-Based Competency Assessments and SAT/ACT College Ready Benchmark account for 5% each. The red circle shows the percent break down for each indicator.

Data Element	Maximum Measure	Performance Measure	x	Factor Value	=	Earned Points	Possible Points
Indicators of Academic Achievement							
Mathematics/Algebra I - Percent Proficient or Advanced on PSSA/Keystone	100.00	91.96	x	7.50%	=	6.89	7.50
Reading/Literature - Percent Proficient or Advanced on PSSA/Keystone	100.00	94.81	x	7.50%	=	7.11	7.50
Science/Biology - Percent Proficient or Advanced on PSSA/Keystone	100.00	82.50	x	7.50%	=	6.18	7.50
Writing - Percent Proficient or Advanced on PSSA	100.00	NA	x	7.50%	=	NA	NA
Industry Standards-Based Competency Assessments - Percent Competent or Advanced	100.00	IS	x	5.00%	=	IS	0.00
Grade 3 Reading - Percent Proficient or Advanced on PSSA	100.00	NA	x	0.00%	=	NA	NA
SAT/ACT College Ready Benchmark	100.00	100.00	x	5.00%	=	5.00	5.00

Figure 16 Indicators of Academic Achievement (PA School Performance, 2013d)

Ten-percent of each school's SPP score is derived from indicators of closing the achievement gap—five-percent for all students and five-percent for historically underperforming students, i.e., economically disadvantaged, English Language Learners, students with disabilities. Figure 17 below shows that achievement gap indicators include PSSA tests or Keystone exams for each subject—Mathematics/Algebra I, Reading/Literature, Science/Biology, and Writing. The red circle shows the percent break down for each test.

Measurement involves comparing the percent proficient or advanced on each test to the goal of 100% proficiency. This measurement is not used for 2012-2013 SPP scores; it will be used starting 2013-2014. The website is unclear how the 10% from these indicators will be distributed to reach 100% total for all indicators.

Indicators of Closing the Achievement Gap - All Students *							
Mathematics/Algebra I - Percent of Required Gap Closure Met	100.00	NA	x	1.25%	=	NA	0.00
Reading/Literature - Percent of Required Gap Closure Met	100.00	NA	x	1.25%	=	NA	0.00
Science/Biology - Percent of Required Gap Closure Met	100.00	NA	x	1.25%	=	NA	0.00
Writing - Percent of Required Gap Closure Met	100.00	NA	x	1.25%	=	NA	0.00
Indicators of Closing the Achievement Gap - Historically Underperforming Students *							
Mathematics/Algebra I - Percent of Required Gap Closure Met	100.00	NA	x	1.25%	=	NA	0.00
Reading/Literature - Percent of Required Gap Closure Met	100.00	NA	x	1.25%	=	NA	0.00
Science/Biology - Percent of Required Gap Closure Met	100.00	NA	x	1.25%	=	NA	0.00
Writing - Percent of Required Gap Closure Met	100.00	NA	x	1.25%	=	NA	0.00

Figure 17 Indicators of Closing the Achievement Gap (PA School Performance, 2013d)

Forty-percent of each school's SPP score is derived from indicators of academic growth from the Pennsylvania Value-Added Assessment System (PVAAS), which follows groups of students over time. According to the Department of Education website (PA Department of Education, 2014), growth is a measure of the current achievement on a quality assessment, such as the PSSA, compared to all prior results from the same quality assessment. (PA Department of Education, 2014, "What is value-added assessment?") Figure 18 illustrates the use of standardized test scores for each subject—Mathematics/Algebra I, Reading/Literature, Science/Biology, and Writing. The red circle shows the percent break down for each test.

The website states that PVAAS is a statistical analysis that is independent of student demographics. (PA Department of Education, 2014, "What is the difference between achievement and growth data?, Growth") Benefits of PVAAS include the ability to track and use data for the following data-driven practices

- Monitor student achievement to ensure all students have opportunities to grow academically.
- Evaluate the impact of educational practices, curricula, instructional methods, and professional development on students' academic performance with the goal of improving upon such practices.
- Identify student programs and professional development for teachers that help students perform at higher levels and focus resources on such programs.

(PA Department of Education, 2014, "What are the benefits of PVAAS?")

Indicators of Academic Growth/PVAAS								
Mathematics/Algebra I - Meeting Annual Academic Growth Expectations	100.00		100.00	x	10.00%	=	10.00	10.00
Reading/Literature - Meeting Annual Academic Growth Expectations	100.00		100.00	x	10.00%	=	10.00	10.00
Science/Biology - Meeting Annual Academic Growth Expectations	100.00		100.00	x	10.00%	=	10.00	10.00
Writing - Meeting Annual Academic Growth Expectations	100.00		NA	x	10.00%	=	NA	0.00

Figure 18 Indicators of Academic Growth for Mt. Lebanon SHS (PA School Performance, 2013d)

Ten-percent of each school's SPP score is derived from other indicators such as attendance rates, graduation rates, grade promotion rates, advanced course credits, International Baccalaureate participation, etc. The large red circle shows the percent break down for each indicator. The red circle in the lower right-hand corner shows the SPP score, 96.33, before extra credit is incorporated into the score.

Other Academic Indicators								
Cohort Graduation Rate	100.00		97.46	x	2.50%	=	2.44	2.50
Promotion Rate	100.00		NA	x	0.00%	=	NA	0.00
Attendance Rate	100.00		95.95	x	2.50%	=	2.40	2.50
Advanced Placement, International Baccalaureate Diploma, or College Credit	100.00		100.00	x	2.50%	=	2.50	2.50
PSAT/Plan Participation	100.00		100.00	x	2.50%	=	2.50	2.50
* Indicators of closing the achievement gap will be included in 2013 - 2014						Total Points	65.02	67.50
						Calculated Score = Total Earned Points/Possible Points =	96.33	

Figure 19 Other Academic Indicators (PA School Performance, 2013d)

Each school can earn extra credit for students scoring advanced on PSSA or Keystone exams, advanced on Industry Standards-Based Competency Assessments, and/or passing an advanced placement course exams. Figure 20 shows the break down for Mt. Lebanon SHS. The large red circle shows the percent break down for each indicator. Figure 20 also illustrates that Mt. Lebanon's SPP score increased from 96.33 to 99.5 because of extra credit, as indicated by the red circle in the lower right-hand corner.

Extra Credit for Advanced Achievement							
Percent PSSA/Keystone Advanced - Mathematics/Algebra I	100.00	57.68	x	1.00%	=	0.57	NA
Percent PSSA/Keystone Advanced - Reading/Literature	100.00	35.14	x	1.00%	=	0.35	NA
Percent PSSA/Keystone Advanced - Science/Biology	100.00	29.79	x	1.00%	=	0.29	NA
Percent PSSA Advanced - Writing	100.00	NA	x	1.00%	=	NA	NA
Percent Advanced - Industry Standards-Based Competency Assessments	100.00	IS	x	1.00%	=	IS	NA
Percent 3 or Higher on an Advanced Placement Exam	100.00	100.00	x	2.00%	=	2.00	NA
Final Score = Calculated Score + Extra Credit for Advanced Achievement =						99.5	

Figure 20 Extra Credit for Advanced Achievement (PA School Performance, 2013d)

Appendix B: School and Estimated Income Data

List of Schools by District	Multiple-zip Estimate Notes	Zip code	Census Mean ($)	Multiple-zip Mean ($)	Census Median ($)	Multiple-zip Median ($)	Score 2012-13
Allegheny Valley SD							
Acmetonia Primary Sch	All schools estimated as average income for zip codes 15024 and 15144.	15024	72,906	65,128	54,872	53,551	73
Colfax Upper El Sch		15144	57,350	65,128	52,230	53,551	81
Springdale JSHS		15144	57,350	65,128	52,230	53,551	79.8
Avonworth SD							
Avonworth El Sch		15237	90,121	90,121	67,527	67,527	75.3
Avonworth HS		15237	90,121	90,121	67,527	67,527	82.3
Avonworth MS		15237	90,121	90,121	67,527	67,527	81.8
Baldwin-Whitehall SD							
Baldwin SHS	Weighted average income for zip codes 15236 (x2) and 15227 (x1)	15236	68,272	65,302	58,291	55,308	83.3
Harrison MS	Weighted average income for zip codes 15236 (x2) and 15227 (x1)	15236	68,272	65,302	58,291	55,308	81.8
McAnnulty El Sch		15236	68,272	68,272	58,291	58,291	87.9
Paynter El Sch		15227	59,362	59,362	49,342	49,342	83.2
Whitehall El Sch		15236	68,272	68,272	58,291	58,291	89.5
Bethel Park SD							
Abraham Lincoln El Sch		15234	67,820	67,820	57,730	57,730	79.6
Benjamin Franklin El Sch		15102	82,166	82,166	68,875	68,875	76.6
Bethel Memorial El Sch		15102	82,166	82,166	68,875	68,875	82
Bethel Park HS	Weighted average income for zip codes 15234 (x1) and 15102 (x4)	15102	82,166	79,297	68,875	66,646	88.9
George Washington El Sch		15102	82,166	82,166	68,875	68,875	94.5
Independence MS	Weighted average income for zip codes 15234 (x1) and 15102 (x4)	15102	82,166	79,297	68,875	66,646	85.9

List of Schools by District	Multiple-zip Estimate Notes	Zip code	Census Mean ($)	Multiple-zip Mean ($)	Census Median ($)	Multiple-zip Median ($)	Score 2012-13
Neil Armstrong 5-6 MS	Weighted average income for zip codes 15234 (x1) and 15102 (x4)	15102	82,166	79,297	68,875	66,646	93.2
William Penn El Sch		15102	82,166	82,166	68,875	68,875	85.8
Brentwood Borough SD							
Brentwood MS		15227	59,362	59,362	49,342	49,342	75
Brentwood SHS		15227	59,362	59,362	49,342	49,342	72
Elroy Avenue El Sch		15227	59,362	59,362	49,342	49,342	73.5
Moore Sch		15227	59,362	59,362	49,342	49,342	80.3
Carlynton SD							
Carlynton JSHS	Average income for zip codes 15106 and 15205	15106	60,138	65,835	43,135	46,601	72.9
Carnegie El Sch		15106	60,138	60,138	43,135	43,135	71.2
Crafton El Sch		15205	71,531	71,531	50,067	50,067	75
Chartiers Valley SD							
Chartiers Valley HS	Average income for zip codes 15243, 15220, 15106, and 15017 (Zipmap, 2014)	15017	83,239	82,940	59,092	59,316	86.6
Chartiers Valley Intrmd School	Average income for zip codes 15243, 15220, 15106, and 15017 (Zipmap, 2014)	15220	77,770	82,940	61,501	59,316	81.7
Chartiers Valley MS	Average income for zip codes 15243, 15220, 15106, and 15017 (Zipmap, 2014)	15017	83,239	82,940	59,092	59,316	90.9
Chartiers Valley Primary Sch	Average income for zip codes 15243, 15220, 15106, and 15017 (Zipmap, 2014)	15017	83,239	82,940	59,092	59,316	85.4

List of Schools by District	Multiple-zip Estimate Notes	Zip code	Census Mean ($)	Multiple-zip Mean ($)	Census Median ($)	Multiple-zip Median ($)	Score 2012-13
Clairton City SD							
Clairton El Sch		15025	67,593	67,593	48,120	48,120	57.4
Clairton MS/HS		15025	67,593	67,593	48,120	48,120	58.2
Cornell SD							
Cornell El Sch		15108	80,729	80,729	62,833	62,833	64.6
Cornell HS		15108	80,729	80,729	62,833	62,833	64.9
Deer Lakes SD							
Curtisville Pri Ctr	Average income for zip codes 15084 and 15024	15084	54,068	63,487	43,426	49,149	84.5
Deer Lakes HS	Average income for zip codes 15084 and 15024	15024	72,906	63,487	54,872	49,149	72.8
Deer Lakes MS	Average income for zip codes 15084 and 15024	15024	72,906	63,487	54,872	49,149	86.4
East Union Intrmd Sch	Average income for zip codes 15084 and 15024	15024	72,906	63,487	54,872	49,149	91.3
Duquesne City SD							
Duquesne El Sch		15110	30,456	30,456	20,330	20,330	49.3
East Allegheny SD							
East Allegheny HS	Average income for zip codes 15137 and 15110 (Duquesne City SD, 2009)	15137	52,089	41,273	44,646	32,488	67.7
Green Valley Primary Sch		15137	52,089	52,089	44,646	44,646	76
Logan MS	Average income for zip codes 15137 and 15110 (Duquesne City SD, 2009)	15137	52,089	41,273	44,646	32,488	70.8

List of Schools by District	Multiple-zip Estimate Notes	Zip code	Census Mean ($)	Multiple-zip Mean ($)	Census Median ($)	Multiple-zip Median ($)	Score 2012-13
Elizabeth Forward SD							
Central El Sch		15037	65,356	65,356	57,624	57,624	75.7
Elizabeth Forward MS	Weighted average income for zip codes 15037 (x1), 15063 (x1), and 15135 (x2)	15037	65,356	62,186	57,624	50,849	78
Elizabeth Forward SHS	Weighted average income for zip codes 15037 (x1), 15063 (x1), and 15135 (x2)	15037	65,356	62,186	57,624	50,849	74.6
Greenock El Sch		15135	63,550	63,550	49,631	49,631	88.3
Mt Vernon El Sch		15135	63,550	63,550	49,631	49,631	79.1
William Penn El Sch		15063	56,288	56,288	46,508	46,508	85.4
Fox Chapel Area SD							
Dorseyville MS	Weighted average income for zip codes 15215 (x1), 15238 (x3)	15238	148,135	136,271	79,857	74,856	84.1
Fairview El Sch		15238	148,135	148,135	79,857	79,857	91.5
Fox Chapel Area HS	Weighted average income for zip codes 15215 (x1), 15238 (x3)	15238	148,135	136,271	79,857	74,856	96.8
Hartwood El Sch		15238	148,135	148,135	79,857	79,857	82.2
Kerr El Sch		15215	100,680	100,680	59,853	59,853	76.1
OHara El Sch		15238	148,135	148,135	79,857	79,857	87.8
Gateway SD							
Dr Cleveland Steward Jr El Sch		15146	71,331	71,331	60,073	60,073	69.6
Evergreen El Sch		15146	71,331	71,331	60,073	60,073	74.2
Gateway MS		15146	71,331	71,331	60,073	60,073	88
Gateway SHS		15146	71,331	71,331	60,073	60,073	81.7
Moss Side MS		15146	71,331	71,331	60,073	60,073	77.6
Ramsey El Sch		15146	71,331	71,331	60,073	60,073	72.6
University Park El Sch		15146	71,331	71,331	60,073	60,073	75.1

List of Schools by District	Multiple-zip Estimate Notes	Zip code	Census Mean ($)	Multiple-zip Mean ($)	Census Median ($)	Multiple-zip Median ($)	Score 2012-13
Hampton Township SD							
Central El Sch		15101	97,402	97,402	79,421	79,421	84.8
Hampton HS	Weighted average income for zip codes 15044 (x1) and 15101 (x2)	15101	97,402	107,295	79,421	84,520	91.8
Hampton MS	Weighted average income for zip codes 15044 (x1) and 15101 (x2)	15101	97,402	107,295	79,421	84,520	94.2
Poff El Sch		15044	127,080	127,080	94,717	94,717	85
Wyland El Sch		15101	97,402	97,402	79,421	79,421	83
Highlands SD							
Fairmount Primary Center	Average income for zip codes 15065, 15084, and 15014	15014	52,575	53,848	41,481	43,544	81.1
Fawn Primary Center	Average income for zip codes 15065, 15084, and 15014	15065	54,900	53,848	45,724	43,544	81.2
Grandview Upper El Sch	Average income for zip codes 15065, 15084, and 15014	15084	54,068	53,848	43,426	43,544	72.3
Highlands MS	Average income for zip codes 15065, 15084, and 15014	15065	54,900	53,848	45,724	43,544	68.3
Highlands SHS	Average income for zip codes 15065, 15084, and 15014	15065	54,900	53,848	45,724	43,544	62.2
Keystone Oaks SD							
Dormont El Sch		15216	63,401	63,401	49,661	49,661	76.1
Fred L Aiken El Sch		15220	77,770	77,770	61,501	61,501	88.4
Keystone Oaks HS	Average income for zip codes 15216, 15234, and 15220	15216	63,401	69,664	49,661	56,297	71.4
Keystone Oaks MS	Average income for zip codes 15216, 15234, and 15220	15216	63,401	69,664	49,661	69,664	67.3

List of Schools by District	Multiple-zip Estimate Notes	Zip code	Census Mean ($)	Multiple-zip Mean ($)	Census Median ($)	Multiple-zip Median ($)	Score 2012-13
Myrtle Ave Sch		15234	67,820	67,820	57,730	57,730	76.3
McKeesport Area SD							
Centennial El Sch		15132	36,382	36,382	26,955	26,955	60.6
Founders Hall Sch	Weighted average income for zip codes 15131 (x1) and 15132 (x2)	15132	36,382	44,657	26,955	33,533	59.7
George Washington Sch		15132	36,382	36,382	26,955	26,955	55.7
McClure Intrmd Sch	Weighted average income for zip codes 15131 (x1) and 15132 (x2)	15131	61,208	44,657	46,689	33,533	51.1
McKeesport Area SHS	Weighted average income for zip codes 15131 (x1) and 15132 (x2)	15132	36,382	44,657	26,955	33,533	51.7
White Oak El Sch		15131	61,208	61,208	46,689	46,689	65
Montour SD							
David E Williams MS	Average income for zip codes 15108 and 15136	15108	80,729	70,300	62,833	50,868	93.9
Forest Grove El Sch		15108	80,729	80,729	62,833	62,833	92.4
J W Burkett El Sch		15136	59,871	59,871	38,903	38,903	90.8
Montour HS	Average income for zip codes 15108 and 15136	15136	59,871	70,300	38,903	50,868	87.4
Moon Area SD							
Allard El Sch		15108	80,729	80,729	62,833	62,833	96.1
Bon Meade El Sch		15108	80,729	80,729	62,833	62,833	89.4
Hyde El Sch		15108	80,729	80,729	62,833	62,833	72.7
J H Brooks Sch		15108	80,729	80,729	62,833	62,833	90.2
McCormick Elem Sch		15108	80,729	80,729	62,833	62,833	93.5
Moon Area Lower MS		15108	80,729	80,729	62,833	62,833	82.5
Moon Area Upper MS		15108	80,729	80,729	62,833	62,833	93.6
Moon SHS		15108	80,729	80,729	62,833	62,833	96.3

List of Schools by District	Multiple-zip Estimate Notes	Zip code	Census Mean ($)	Multiple-zip Mean ($)	Census Median ($)	Multiple-zip Median ($)	Score 2012-13
			Mt Lebanon SD				
Foster El Sch		15234	67,820	67,820	57,730	57,730	97.8
Hoover El Sch		15243	110,612	110,612	73,537	73,537	97
Howe El Sch		15228	105,144	105,144	74,980	74,980	95.5
Jefferson El Sch		15243	110,612	110,612	73,537	73,537	92
Jefferson MS	Weighted average income for zip codes 15228 (x1), 15216 (x1), and 15243 (x2)	15243	110,612	97,442	73,537	67,929	93.8
Lincoln El Sch		15216	63,401	63,401	49,661	49,661	94.7
Markham El Sch		15228	105,144	105,144	74,980	74,980	93.2
Mellon MS	Weighted average income for zip codes 15234 (x1) and 15228 (x2)	15228	105,144	92,703	74,980	69,230	88.4
Mt Lebanon SHS	Weighted average income for zip codes 15228 (x3), 15216 (x1), 15234 (x1), and 15243 (x2)	15228	105,144	95,411	74,980	68,486	99.5
Washington El Sch		15228	105,144	105,144	74,980	74,980	93.6
			North Allegheny SD				
Bradford Woods El Sch		15015	141,033	141,033	107,188	107,188	93.8
Carson MS	Weighted average income for zip codes 15101 (x1) and 15237 (x2)	15237	90,121	92,548	67,527	71,492	90.6
Franklin El Sch		15143	132,475	132,475	83,232	83,232	94.4
Hosack El Sch		15101	97,402	97,402	79,421	79,421	88.8
Ingomar El Sch		15237	90,121	90,121	67,527	67,527	86.7
Ingomar MS	Average income for zip codes 15143 and 15237	15237	90,121	111,298	67,527	75,380	97.2
Marshall El Sch		15090	152,756	152,756	115,787	115,787	93.9
Marshall MS	Average income for zip codes 15090 and 15015	15090	152,756	146,895	115,787	111,488	94.7
McKnight El Sch		15237	90,121	90,121	67,527	67,527	93

List of Schools by District	Multiple-zip Estimate Notes	Zip code	Census Mean ($)	Multiple-zip Mean ($)	Census Median ($)	Multiple-zip Median ($)	Score 2012-13
North Allegheny IHS	Weighted average income for zip codes 15101 (x1), 15237 (x3), 15090 (x1), 15143 (x1), and 15015 (x1)	15237	90,121	113,433	67,527	84,030	90.6
North Allegheny SHS	Weighted average income for zip codes 15101 (x1), 15237 (x3), 15090 (x1), 15143 (x1), and 15015 (x1)	15090	152,756	113,433	115,787	84,030	94.1
Peebles El Sch		15237	90,121	90,121	67,527	67,527	94.1
			North Hills SD				
Highcliff El Sch		15229	73,745	73,745	56,596	56,596	90.3
McIntyre El Sch		15237	90,121	90,121	67,527	67,527	91.4
North Hills JHS	Average income for zip codes 15229 and 15237	15229	73,745	81,933	56,596	62,062	78.2
North Hills SHS	Average income for zip codes 15229 and 15237	15229	73,745	81,933	56,596	62,062	85.6
Ross El Sch		15237	90,121	90,121	67,527	67,527	93.8
West View El Sch		15229	73,745	73,745	56,596	56,596	86
Avalon El Sch		15202	59,671	59,671	43,670	43,670	73.4
Bellevue El Sch		15202	59,671	59,671	43,670	43,670	84.3
Northgate MSHS		15202	59,671	59,671	43,670	43,670	66.6
			Penn Hills SD				
Forbes El Sch		15147	59,289	59,289	43,703	43,703	71.4
Linton MS	Weighted average income for zip codes 15235 (x2) and 15147 (x1)	15235	59,228	59,248	47,639	46,327	55.7
Penn Hebron El Academy		15235	59,228	59,228	47,639	47,639	68.7
Penn Hills SHS	Weighted average income for zip codes 15235 (x2) and 15147 (x1)	15235	59,228	59,248	47,639	46,327	64.2
Washington El Sch		15235	59,228	59,228	47,639	47,639	70.1

List of Schools by District	Multiple-zip Estimate Notes	Zip code	Census Mean ($)	Multiple-zip Mean ($)	Census Median ($)	Multiple-zip Median ($)	Score 2012-13
Pine-Richland SD							
Eden Hall Upper El Sch		15044	127,080	127,080	94,717	94,717	79.2
Hance El Sch		15044	127,080	127,080	94,717	94,717	93.4
Pine-Richland HS		15044	127,080	127,080	94,717	94,717	95.1
Pine-Richland MS		15044	127,080	127,080	94,717	94,717	94.9
Richland El Sch		15044	127,080	127,080	94,717	94,717	90.7
Wexford El Sch		15090	152,756	152,756	115,787	115,787	96
Pittsburgh SD							
Acadamy at Westinghouse 6-12	Average income for zip codes 15208, 15201, and 15206	15208	58,989	56,715	30,704	35,648	43.4
Pittsburgh Allderdice HS	Weighted average income for zip codes 15217 (x2), 15207 (x1), and 15208 (x1)	15217	100,939	77,716	62,462	48,837	71.8
Pittsburgh Allegheny 6-8	Average income for all Pittsburgh zip codes	15212	52,357	61,066	36,685	41,429	56.6
Pittsburgh Allegheny K-5	Average income for all Pittsburgh zip codes	15212	52,357	61,066	36,685	41,429	70.7
Pittsburgh Arlington K-8		15210	42,761	42,761	33,772	33,772	54.3
Pittsburgh Arsenal 6-8	Weighted average income for zip codes 15219 (x2), 15201 (x1), 15206 (x1), and 15224 (x1)	15201	54,077	43,317	42,348	33,099	60
Pittsburgh Arsenal K-5		15201	54,077	54,077	42,348	42,348	55.2
Pittsburgh Banksville K-5		15216	63,401	63,401	49,661	49,661	71.9
Pittsburgh Beechwood K-5		15216	63,401	63,401	49,661	49,661	58.9
Pittsburgh Brashear HS	Weighted average income for zip codes 15226 (x2), 15210 (x1), 15211 (x1), 15205 (x1), and 15216 (x1)	15216	63,401	59,610	49,661	46,534	54.7
Pittsburgh Brookline K-8		15226	57,447	57,447	48,193	48,193	71.9
Pittsburgh CAPA 6-12	Average income for all Pittsburgh zip codes	15222	102,653	61,066	58,319	41,429	82.2

List of Schools by District	Multiple-zip Estimate Notes	Zip code	Census Mean ($)	Multiple-zip Mean ($)	Census Median ($)	Multiple-zip Median ($)	Score 2012-13
Pittsburgh Carmalt K-8	Average income for all Pittsburgh zip codes	15226	57,447	61,066	48,193	41,429	80.1
Pittsburgh Carrick HS	Weighted average income for zip codes 15210 (x3), 15226 (x2), and 15216 (x1)	15210	42,761	51,096	33,772	41,227	58.9
Pittsburgh Classical 6-8	Average income for all Pittsburgh zip codes	15220	77,770	61,066	61,501	41,429	69.3
Pittsburgh Colfax K-8		15217	100,939	100,939	62,462	62,462	75.6
Pittsburgh Concord K-5		15210	42,761	42,761	33,772	33,772	62.4
Pittsburgh Dilworth K-5	Average income for all Pittsburgh zip codes	15206	57,080	61,066	33,891	41,429	71
Pittsburgh Faison K-5		15208	58,989	58,989	30,704	30,704	57.5
Pittsburgh Fulton K-5		15206	57,080	57,080	33,891	33,891	66.6
Pittsburgh Grandview K-5		15210	42,761	42,761	33,772	33,772	56
Pittsburgh Greenfield K-8		15207	49,996	49,996	39,721	39,721	72.9
Pittsburgh King K-8		15212	52,357	52,357	36,685	36,685	53.6
Pittsburgh Langley K-8		15204	43,758	43,758	33,435	33,435	53.2
Pittsburgh Liberty K-5	Average income for all Pittsburgh zip codes	15232	84,906	61,066	48,336	41,429	82.6
Pittsburgh Lincoln K-5		15206	57,080	57,080	33,891	33,891	63.4
Pittsburgh Linden K-5	Average income for all Pittsburgh zip codes	15208	58,989	61,066	30,704	41,429	77.5
Pittsburgh Manchester K-8		15233	57,088	57,088	34,588	34,588	63.1
Pittsburgh Mifflin K-8		15207	49,996	49,996	39,721	39,721	66.9
Pittsburgh Miller K-5		15219	30,074	30,074	17,441	17,441	58.6
Pittsburgh Milliones 6-12 (University Prep)	Weighted average income for zip codes 15219 (x2), 15201 (x3), 15206 (x1), and 15224 (x1)	15219	30,074	46,392	17,441	33,099	49.3
Pittsburgh Minadeo K-5		15217	100,939	100,939	62,462	62,462	67.2
Pittsburgh Montessori K-5	Average income for all Pittsburgh zip codes	15206	57,080	61,066	33,891	41,429	61.9
Pittsburgh Morrow K-8		15212	52,357	52,357	36,685	36,685	52

List of Schools by District	Multiple-zip Estimate Notes	Zip code	Census Mean ($)	Multiple-zip Mean ($)	Census Median ($)	Multiple-zip Median ($)	Score 2012-13
Pittsburgh Obama 6-12	Average income for all Pittsburgh zip codes	15206	57,080	61,066	33,891	41,429	71
Pittsburgh Perry HS	Weighted average income for zip codes 15212 (x4), 15214 (x1), and 15233 (x1)	15214	50,364	52,813	37,815	36,524	43.3
Pittsburgh Phillips K-5	Average income for all Pittsburgh zip codes	15203	69,143	61,066	43,360	41,429	79.7
Pittsburgh Roosevelt K -5		15210	42,761	42,761	33,772	33,772	76.7
Pittsburgh Schiller 6-8		15212	52,357	52,357	36,685	36,685	57.7
Pittsburgh Science and Technology Academ	Average income for all Pittsburgh zip codes	15213	47,704	61,066	22,508	41,429	72.5
Pittsburgh South Brook 6-8	Weighted average income for zip codes 15210 (x2) and 15226 (x1)	15226	57,447	47,656	48,193	38,579	57.9
Pittsburgh South Hills 6-8	Average income for zip codes 15210 (x1), 15211 (x1), 15205 (x1), and 15216 (x2)	15216	63,401	61,233	49,661	46,496	65.3
Pittsburgh Spring Hill K-5		15212	52,357	52,357	36,685	36,685	58
Pittsburgh Sterrett 6-8	Average income for zip codes 15217 and 15208	15208	58,989	79,964	30,704	46,583	62.7
Pittsburgh Sunnyside K-8		15201	54,077	54,077	42,348	42,348	74.4
Pittsburgh Weil K-5		15219	30,074	30,074	17,441	17,441	67.4
Pittsburgh West Liberty K-5		15226	57,447	57,447	48,193	48,193	77.3
Pittsburgh Westwood K-5		15205	71,531	71,531	50,067	50,067	66
Pittsburgh Whittier K-5		15211	65,073	65,073	49,319	49,319	79.6
Pittsburgh Woolslair K-5		15224	45,282	45,282	35,877	35,877	60
Plum Borough SD							
Center El Sch		15239	75,273	75,273	68,716	68,716	88.4
Holiday Park El Sch		15239	75,273	75,273	68,716	68,716	83.5
Oblock JHS		15239	75,273	75,273	68,716	68,716	86.2
Pivik El Sch		15239	75,273	75,273	68,716	68,716	90.4
Plum SHS		15239	75,273	75,273	68,716	68,716	80.5
Regency Park El Sch		15239	75,273	75,273	68,716	68,716	83.3

List of Schools by District	Multiple-zip Estimate Notes	Zip code	Census Mean ($)	Multiple-zip Mean ($)	Census Median ($)	Multiple-zip Median ($)	Score 2012-13
Stevenson El Sch		15239	75,273	75,273	68,716	68,716	90.9
Quaker Valley SD							
Edgeworth El Sch	Average income for zip codes 15143 and 15056	15143	132,475	92,275	83,232	59,503	81.4
Osborne El Sch	Average income for zip codes 15143 and 15056	15143	132,475	92,275	83,232	59,503	88.1
Quaker Valley HS	Average income for zip codes 15143 and 15056	15056	52,074	92,275	35,774	59,503	93.2
Quaker Valley MS	Average income for zip codes 15143 and 15056	15143	132,475	92,275	83,232	59,503	93.8
Riverview SD							
Riverview HS	Average income for zip codes 15139 and 15147	15139	78,855	69,072	51,724	47,714	83
Tenth Street El Sch		15139	78,855	78,855	51,724	51,724	83.5
Verner El Sch		15147	59,289	59,289	43,703	43,703	79.3
Shaler Area SD							
Burchfield Primary Sch		15101	97,402	97,402	79,421	79,421	92.7
Jeffery Primary Sch		15116	79,138	79,138	69,248	69,248	85.3
Marzolf Primary Sch		15209	63,244	63,244	54,004	54,004	90.1
Reserve Primary Sch		15212	52,357	52,357	36,685	36,685	88.4
Rogers Primary Sch		15116	79,138	79,138	69,248	69,248	92.3
Shaler Area El Sch	Weighted average income for zip codes 15101 (x1), 15116 (x2), 15209 (x1), 15212 (x1)	15116	79,138	74,256	69,248	61,721	73.7
Shaler Area HS	Weighted average income for zip codes 15101 (x1), 15116 (x2), 15209 (x1), 15212 (x1)	15209	63,244	74,256	54,004	61,721	88.9
Shaler Area MS	Weighted average income for zip codes 15101 (x1), 15116 (x2), 15209 (x1), 15212 (x1)	15116	79,138	74,256	69,248	61,721	82.1

List of Schools by District	Multiple-zip Estimate Notes	Zip code	Census Mean ($)	Multiple-zip Mean ($)	Census Median ($)	Multiple-zip Median ($)	Score 2012-13
South Allegheny SD							
South Allegheny Early Childhood Ctr		15133	55,979	55,979	44,026	44,026	70.2
South Allegheny El Sch		15133	55,979	55,979	44,026	44,026	78.1
South Allegheny MS/HS		15133	55,979	55,979	44,026	44,026	65
South Fayette Township SD							
South Fayette MS		15057	85,655	85,655	67,624	67,624	91.9
South Fayette Twp El Sch		15057	85,655	85,655	67,624	67,624	94
South Fayette Twp HS		15057	85,655	85,655	67,624	67,624	83.4
South Park SD							
South Park Elem Center		15129	73,444	73,444	64,464	64,464	82.7
South Park MS		15129	73,444	73,444	64,464	64,464	83
South Park SHS		15129	73,444	73,444	64,464	64,464	90
Steel Valley SD							
Barrett El Sch		15120	52,705	52,705	39,787	39,787	54
Park El Sch		15120	52,705	52,705	39,787	39,787	79.8
Steel Valley MS		15120	52,705	52,705	39,787	39,787	73.4
Steel Valley SHS		15120	52,705	52,705	39,787	39,787	62.3
Sto-Rox SD							
Sto-Rox El Sch		15136	59,871	59,871	38,903	38,903	60.2
Sto-Rox HS		15136	59,871	59,871	38,903	38,903	52.8
Sto-Rox MS		15136	59,871	59,871	38,903	38,903	59.6
Upper Saint Clair SD							
Baker El Sch		15241	138,167	138,167	103,326	103,326	82.2
Boyce MS		15241	138,167	138,167	103,326	103,326	95.1
Eisenhower El Sch		15241	138,167	138,167	103,326	103,326	97.9
Fort Couch MS		15241	138,167	138,167	103,326	103,326	96.8
Streams El Sch		15241	138,167	138,167	103,326	103,326	80.9
Upper Saint Clair HS		15241	138,167	138,167	103,326	103,326	94.1

List of Schools by District	Multiple-zip Estimate Notes	Zip code	Census Mean ($)	Multiple-zip Mean ($)	Census Median ($)	Multiple-zip Median ($)	Score 2012-13
West Allegheny SD							
Donaldson Elem Sch		15071	84,397	84,397	57,815	57,815	96.9
McKee El Sch		15071	84,397	84,397	57,815	57,815	94.1
West Allegheny MS	Weighted average income for zip codes 15126 (x1) and 15071 (x2)	15126	74,994	81,263	64,575	60,068	80.5
West Allegheny SHS	Weighted average income for zip codes 15126 (x1) and 15071 (x2)	15126	74,994	81,263	64,575	60,068	93.9
Wilson El Sch		15126	74,994	74,994	64,575	64,575	95.2
West Jefferson Hills SD							
Gill Hall El Sch		15025	67,593	67,593	48,120	48,120	94.3
Jefferson El Sch		15025	67,593	67,593	48,120	48,120	86.8
McClellan El Sch		15236	68,272	68,272	58,291	58,291	77.3
Pleasant Hills MS	Weighted average income for zip codes 15025 (x2) and 15236 (x1)	15236	68,272	67,819	58,291	51,510	90.2
Thomas Jefferson HS	Weighted average income for zip codes 15025 (x2) and 15236 (x1)	15025	67,593	67,819	48,120	51,510	88.1
West Mifflin Area SD							
Clara Barton El Sch		15122	53,352	53,352	45,228	45,228	83.9
Homeville El Sch		15122	53,352	53,352	45,228	45,228	72.9
New Emerson El Sch		15122	53,352	53,352	45,228	45,228	78.6
West Mifflin Area HS	Average income for zip codes 15122 and 15110 (Duquesne City SD, 2009)	15122	53,352	41,904	45,228	32,779	71.1
West Mifflin Area MS	Average income for zip codes 15122 and 15110 (Duquesne City SD, 2009)	15122	53,352	41,904	45,228	32,779	69.3
Wilkinsburg Borough SD							
Kelly El Sch		15221	50,222	50,222	36,930	36,930	56.4
Turner El Sch		15221	50,222	50,222	36,930	36,930	55.1
Wilkinsburg MS		15221	50,222	50,222	36,930	36,930	44.4

List of Schools by District	Multiple-zip Estimate Notes	Zip code	Census Mean ($)	Multiple-zip Mean ($)	Census Median ($)	Multiple-zip Median ($)	Score 2012-13
Wilkinsburg SHS		15221	50,222	50,222	36,930	36,930	36.3
Woodland Hills SD							
Dickson El Sch		15218	63,733	63,733	48,340	48,340	77.4
Edgewood El Sch		15218	63,733	63,733	48,340	48,340	78.5
Fairless El Sch		15104	34,366	34,366	24,106	24,106	53.3
Shaffer El Sch		15221	50,222	50,222	36,930	36,930	69
Wilkins El Sch		15235	59,228	59,228	47,639	47,639	61.3
Woodland Hills Academy		15145	45,893	45,893	35,319	35,319	63.1
Woodland Hills JHS	Weighted average income for zip codes 15218 (x2), 15104 (x1), 15221 (x1), 15235 (x1), and 15145 (x1)	15218	63,733	52,863	48,340	40,112	48.8
Woodland Hills SHS	Weighted average income for zip codes 15218 (x2), 15104 (x1), 15221 (x1), 15235 (x1), and 15145 (x1)	15221	50,222	52,863	36,930	40,112	69

(U.S. Census Bureau, 2012a) and (U.S. Census Bureau, 2012b)

Appendix C: VB Script to Extract 5-digit Zip Codes

```
‘The following script deletes the 4-digit extension from the school zip codes so school
‘data can be linked to income data, which uses only the 5-digit code
Private Sub RecordsetOpenTable()


'create variable db to represent the database needed for this script
Dim db As Database
'set db to equal the current database where this script is running
Set db = CurrentDb


'create varaible rs to represent the records that will be queried
Dim rs As DAO.Recordset
'open table SCHOOL_ZIP_CODES from the current database db; use SQL to set rs equal to all
records in the table
Set rs = db.OpenRecordset("SELECT * FROM SCHOOL_ZIP_CODES")


'if there are records in rs, perform action on record
If rs.RecordCount <> 0 Then


'loop through all records in rs
 'perform action until end of file (EOF) has been reached
 Do While Not (rs.EOF = True)
```

```
'create working variables to perform action on record
Dim str As String
Dim zip5d As String

'store zip code value from field "School Zip" in variable str
str = rs.Fields("School Zip")
'store the first 5 characters in variable zip5d
zip5d = Left$(str, 5)

'copy the zip5d to new field "ZipShort" in table SCHOOL_ZIP_CODES
With SCHOOL_ZIP_CODES
rs.Edit
rs("ZipShort") = zip5d
rs.Update
End With

'move to the next record
rs.MoveNext

Loop
End If

'close the record set and clear rs and db variables
```

```
rs.Close

Set rs = Nothing

Set db = Nothing

End Sub
```