Policy Interventions for Student Success: A Look at Disparities at the School District of Beloit

My Giang Ngoc Le'25 and Yaksh Toyesh Ujoodha'25

DSDA310: Senior Seminar for Data Science and Data Analytic: Harnessing Data for the Common Good

Beloit College

10th November 2024

Executive Summary

Overview: The School District of Beloit is diverse with students coming from various racial, ethnic, and socio-economic backgrounds. The data shows an intersection between student race, parental involvement, level of English proficiency, and academic outcomes. Thus, we aim to conduct this analysis to identify the challenges that minority groups are facing and propose policies to enhance better academic performance of these students.

Key Findings:

1. Health disparities were not proven to be statistically significant enough to influence the academic outcomes of students in the sample.
2. Belonging to a minority group adversely affects academic performance, even being controlled for several factors such as parental involvement, old school environment, and English proficiency.
3. The level of English proficiency significantly impacts the academic outcomes for each student on Reading scores, while controlling for the same race, parental involvement, and old-school environment.
4. Robinson Elementary stands out with dramatically poor academic outcomes within the whole system.

Policy Recommendations

1. Develop tutoring and mentoring programs, especially for minority students
2. Use culturally responsive pedagogy for those who use English as a second language
3. Revitalize Robinson Elementary

Code: URL

Introduction

Context

> *"In the state of Wisconsin, the large differences in academic achievement among different student groups have persisted for at least two decades. Therefore, this study was conducted to investigate the role of school versus family factors in explaining students' academic performance at the School District of Beloit (SDB)" (From An examination of test scores of elementary school students at the School District of Beloit by Diep Phan, Linh Anh Le and Hoodish Domun)*

The authors of the above study stressed the importance of determining a student's academic performance in the long run. Consequently, they used the standardized RIT test scores which had the benefit of tracking and comparing students' learning progress over time. However, our approach consists of only looking at the scores of a specific semester Winter of the Academic Year 2018 - 2019.

Analyzing scores over time requires the use of more complex time-series statistical techniques which is beyond the scope of this project. Moreover, we also avoid introducing variability due to semesters or varying individual academic growth progression, making our model more straightforward and easy to interpret.

Research Questions

- Are there significant differences in academic performance between students of different racial/ethnic backgrounds?
- Do health disparities contribute to these academic differences?
- Is low language proficiency a barrier to academic success?

Scope: Our dataset consists of 215 observations based on the survey and data methodology carried out by Professor Diep Phan, Linh Anh Le, and Hoodish Domun for their previous research. While health and race disparities remain the focus of our work, we also employed several techniques (later discussed in this report) to understand what other factors could affect the academic performance of students at the School District of Beloit. One major decision that we implemented was to propose two different regression models for each subject namely Math and Reading. We believe that certain demographic variables could potentially have a more significant effect on one subject than the other.

Data Exploration and Cleaning

Data Overview

Our dataset is comprised of 215 observations with a total of 30 explanatory variables. 27 of these variables are of type categorical while 3 of them are of type numeric. The dataset was mainly divided into two groups, including student demographics, parental involvement, and academic information. Our key variables for this study would be WT18_19MathRIT, WT18_19ReadRIT, Majority_Group, and child_health_level_very_good_compare.

Data Cleaning and Data Transformation

Note: More details on our data cleaning process can be found with the code Jupyter Notebook shared along this report.

Below are the steps we carried out to clean our data:

1. Rename survey questions to meaningful names.
2. Re-encode 'A', 'B', C', 'D' … values to meaningful string values.
    a. Some questions had the option 'Other' and these were labeled as the string 'NULL'.
    b. Keep scores from WT18_19MathRIT and WT18_19ReadRIT.
3. Generate a quick EDA report.
4. Remove unnecessary columns:
    a. Columns such as 'Student ID', 'DOB', 'Gr', 'Age', 'Lunch Code', 'Q3_3_1', 'Q4_3_1', Q5_3_1' did not provide much interesting information and had little variation
    b. Columns such as 'Hispanic', 'Multi Race', and 'Fed Race' were dropped since 'L_Race' was chosen to explain the effect of race in our study. We believe that Local Race would be the best explanatory variable when looking at the effect of race on academic performance.
    c. Columns like 'LEP' were dropped since 'Eng_Prof' was instead used to understand the English Proficiency levels of students.
    d. Columns like 'bool_serious_health', 'str_mother_job', 'str_father_job', 'str_guardian_job', 'bool_legal_guardian', 'cat_guardian_edu_level', 'cat_guardian_work_hours' had imbalanced category representation and lots of missing values.

5. Creating new features
    a. To deal with the imbalanced representation of race, we created a variable called Majority_Group based on L_Race where 1 was assigned if the student was 'CAUCASSIAN' else 0 if not. This would allow us to understand the effect of belonging to a minority group.
    b. To deal with the imbalanced representation of child health level, we created a new variable called child_health_level_very_good_compare based on child_health_level where a value of 1 was assigned if the student was labeled as 'Very Good' else 0 if not. This would allow us to understand the effect of belonging to a health group of 'Very Good' compared to other groups such as 'Good' **as perceived by the parents.**
    c. To deal with the imbalance of English Proficiency, we created a new variable called Eng_Prof_Full based on Eng_Prof where a value of 1 was assigned if the student was labeled as '7-Fully English Proficient' else 0 if not, allowing us to control the level of English Proficiency and allow for better interpretation.
6. Removed columns 'L_Race', 'child_health_level' and 'Eng_Prof'
7. Handling Missing Values
    a. We found that 6 students did not fill out any information in the numeric columns such as test scores, num_adults_in_the_house and num_children_in_the_house. Thus, we decided to drop those rows leading to 215 observations rather than 221 observations.
    b. Missing values in categorical values were filled with the string 'NULL' instead. We believed that this strategy would allow us to keep as much of the data as possible while not hurting the statistical investigation of our study.
8. Perform one-hot encoding on all categorical variables
    a. We believe that each group is independently occurring and we did not feel that it was necessary to factor any inherent ordering in our categorical variables.

Analysis and Findings

Exploratory Data Analysis (EDA)

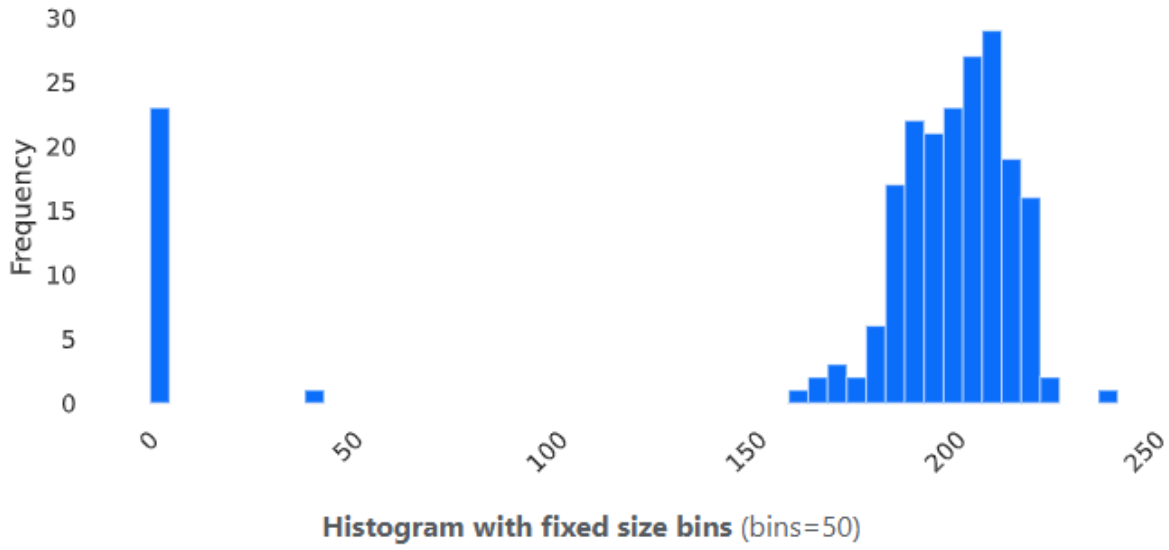
For Math Scores

Quantile statistics

Minimum	0
5-th percentile	0
Q1	189
median	202
Q3	210
95-th percentile	221
Maximum	243
Range	243
Interquartile range (IQR)	21

Descriptive statistics

Standard deviation	64.636097
Coefficient of variation (CV)	0.35871866
Kurtosis	3.7798804
Mean	180.18605
Median Absolute Deviation (MAD)	10
Skewness	-2.3260332
Sum	38740
Variance	4177.825
Monotonicity	Not monotonic



Histogram with fixed size bins (bins=50)

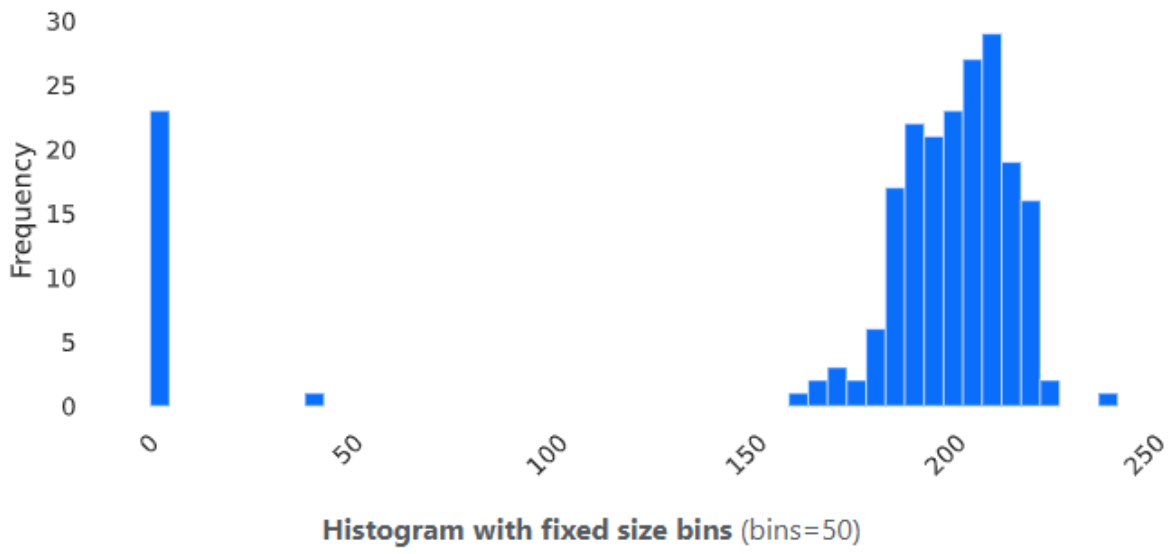
For Reading Scores

Quantile statistics

Minimum	0
5-th percentile	0
Q1	164
median	194
Q3	207.5
95-th percentile	218
Maximum	232
Range	232
Interquartile range (IQR)	43.5

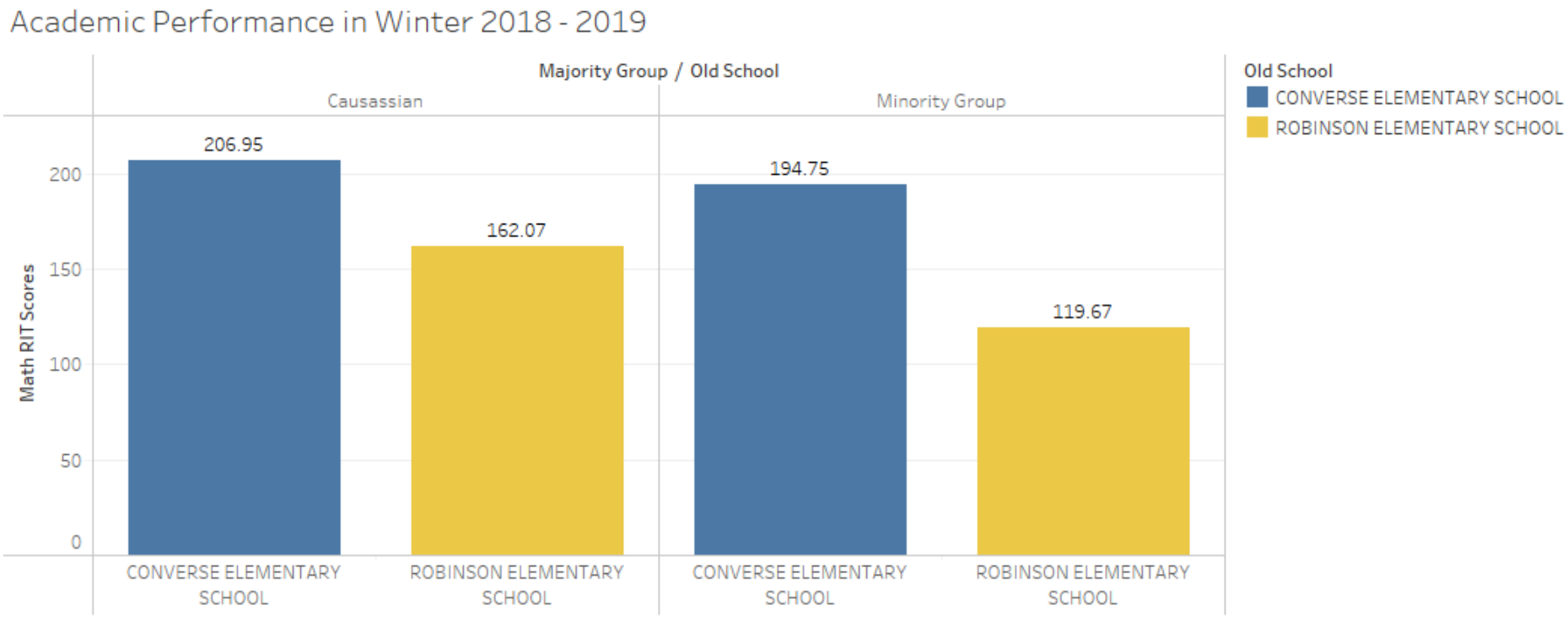
Descriptive statistics

Standard deviation	83.291627
Coefficient of variation (CV)	0.53760732
Kurtosis	-0.2381612
Mean	154.93023
Median Absolute Deviation (MAD)	17
Skewness	-1.275583
Sum	33310
Variance	6937.4951
Monotonicity	Not monotonic

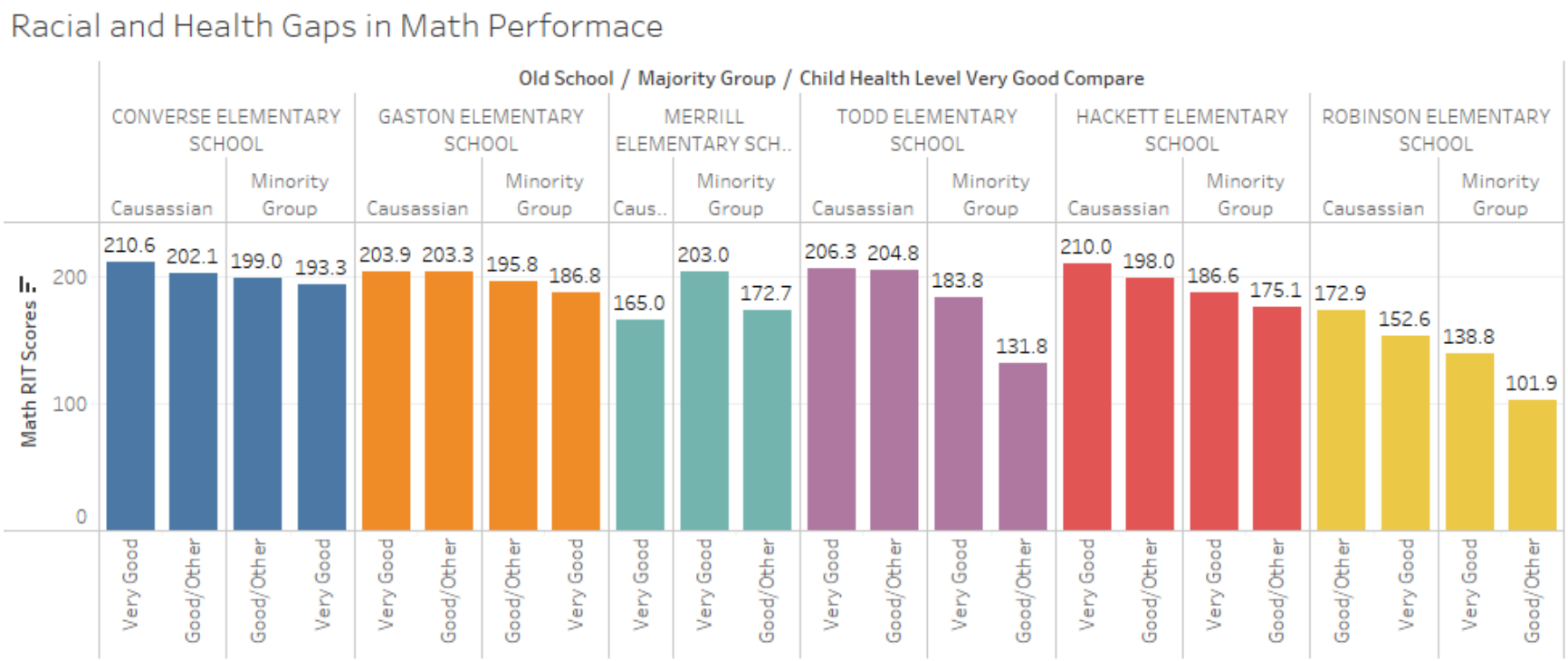


Histogram with fixed size bins (bins=50)

Interactions for Math Scores



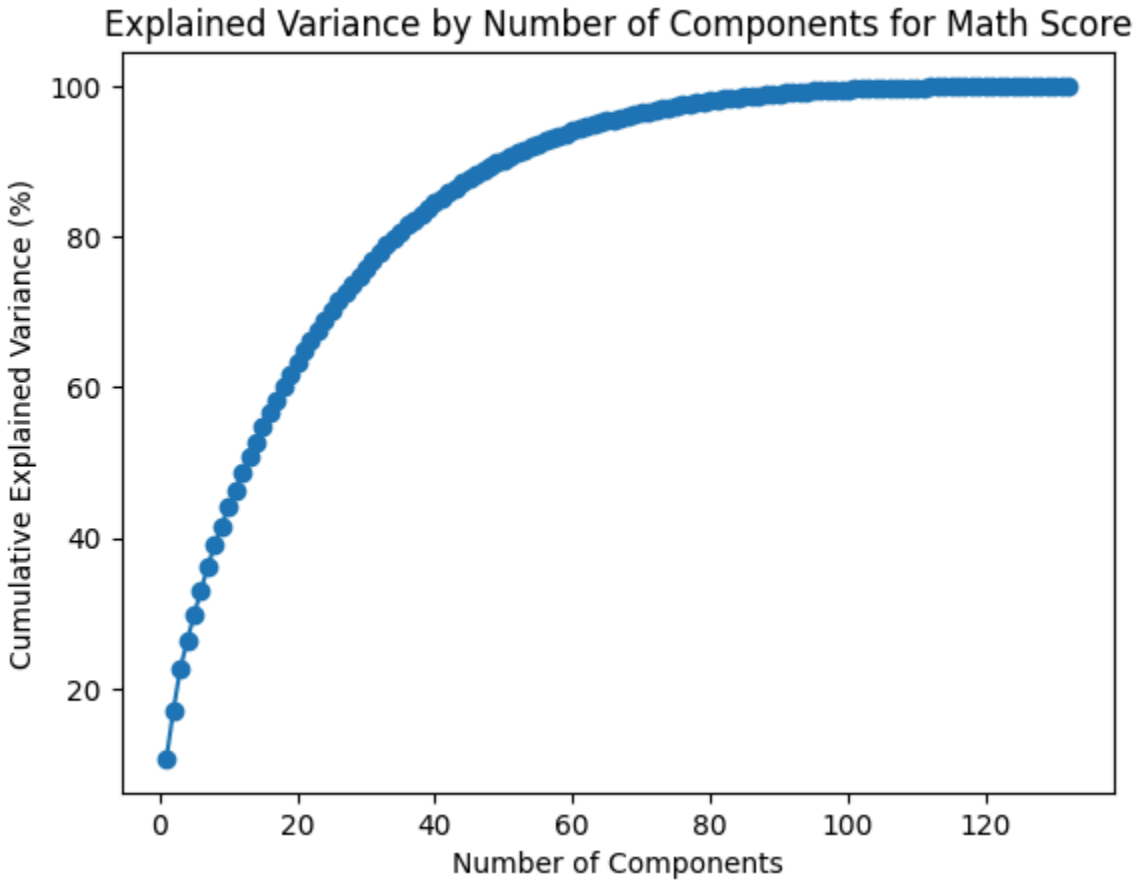
This figure shows how different schools as well as belonging to different racial groups have an impact on the average Math RIT Score.



It is interesting to see that the effect of health is only significant in Robinson Elementary School. However, there is an interesting trend across all schools where the Caucasian group does have higher average Math RIT Scores than the minority group.

Principal Component Analysis of Math Scores

To better understand the variation in math scores, we performed a PCA analysis on our data.



The graph suggests that 80% of the variance is explained by almost 40 explanatory variables. Finding the right explanatory variables out of more than 120 variables is quite a challenging task, therefore we decided to implement Lasso Regression to find out what would be the most significant features.

Feature Selection - Results of Lasso Regression for Math Scores

Our analysis revealed that 'Old_School_ROBINSON ELEMENTARY SCHOOL' and 'Majority_Group_1' were two significant features that could explain the variation in Math Scores

Regression Model for Math Scores

Our proposed regression model is thus as follows:

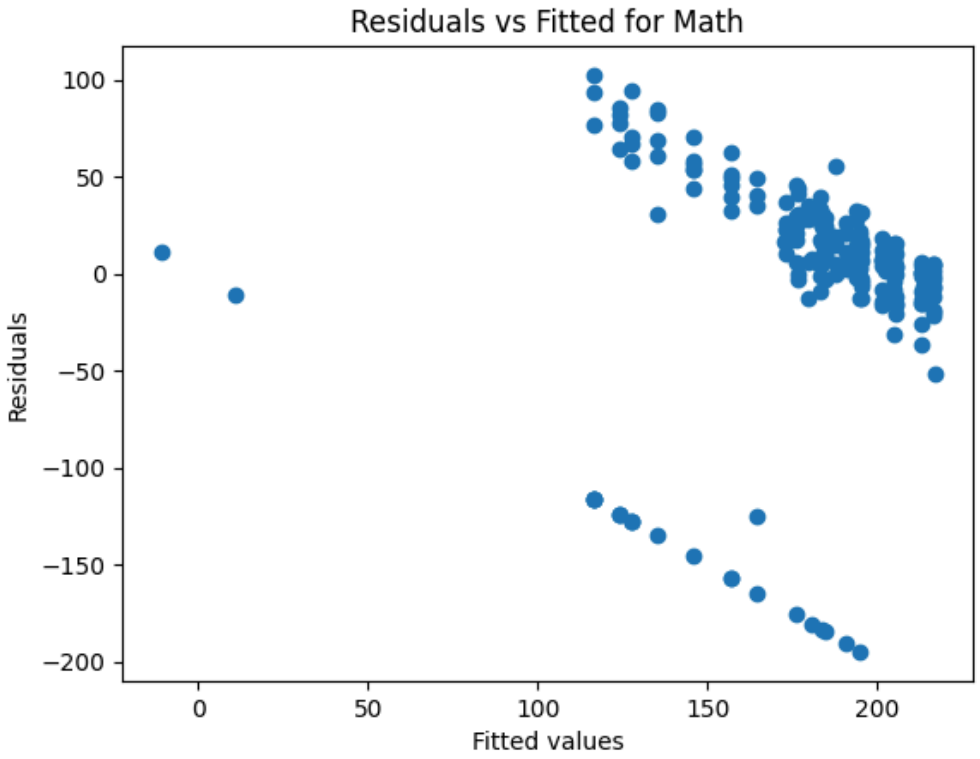
$$WT18\_19MathRIT = \alpha + \beta1 \cdot Old\_School\_GASTON\ ELEMENTARY\ SCHOOL + \beta2 \cdot Old\_School\_HACKETT\ ELEMENTARY\ SCHOOL + \beta3 \cdot Old\_School\_MERRILL\ ELEMENTARY\ SCHOOL + \beta4 \cdot Old\_School\_ROBINSON\ ELEMENTARY\ SCHOOL + \beta5 \cdot Old\_School\_TODD\ ELEMENTARY\ SCHOOL + \beta6 \cdot Old\_School\_NULL + \beta7 \cdot child\_health\_level\_very\_good\_compare\_1 + \beta8 \cdot Majority\_Group\_1 + \beta9 \cdot Eng\_Prof\_Full\_1$$

Checking for Heteroskedasticity for Math Scores

We conducted both White's test and Breusch-Pagan's test to verify if our model suffered from heteroskedasticity. Consequently, we applied robust standard errors in our analysis.

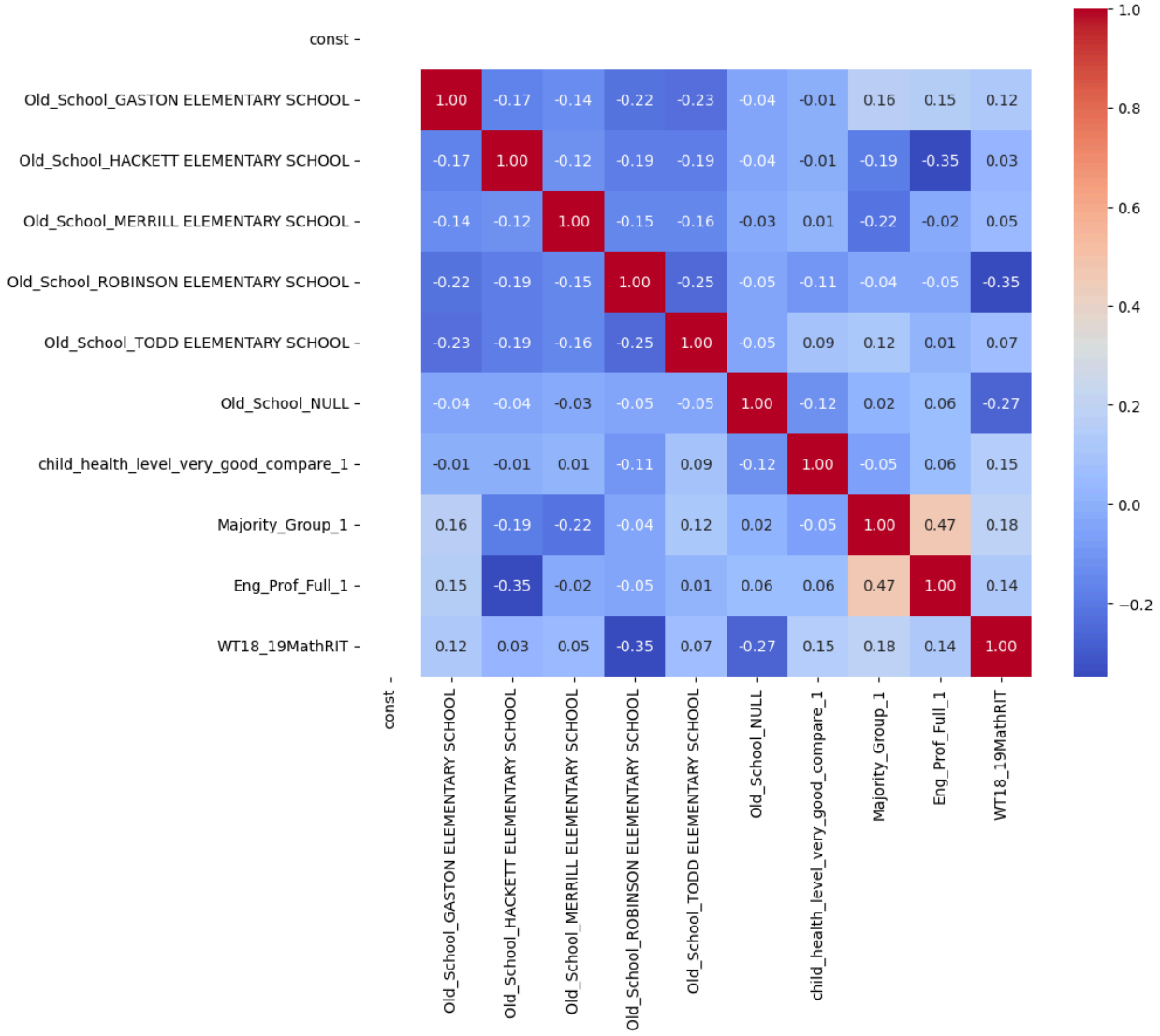
The results are as below:

- White's Test p-value for Math: 0.0034034630700949375
- Breusch-Pagan p-value for Math: 7.14597750315618e-06



Checking for Multicollinearity for Math Scores

We also verified for multicollinearity in our model to ensure that it is robust and not prone to statistical errors. Below is a correlation heat map of our regressors and dependent variable:



All correlations were below the absolute value of 0.4 and hence, we concluded that the multicollinearity effect was not strong.

Regression Results for Math Scores

OLS Regression Results

Dep. Variable:	WT18_19MathRIT	R-squared:	0.248
Model:	OLS	Adj. R-squared:	0.215
Method:	Least Squares	F-statistic:	66.01
Date:	Sun, 10 Nov 2024	Prob (F-statistic):	9.98e-56
Time:	16:21:03	Log-Likelihood:	-1170.3
No. Observations:	215	AIC:	2361.
Df Residuals:	205	BIC:	2394.
Df Model:	9		
Covariance Type:	HC1		

	coef	std err	z	P>|z|	[0.025	0.975]
const	176.4838	11.775	14.988	0.000	153.405	199.563
Old_School_GASTON ELEMENTARY SCHOOL	-3.9875	8.072	-0.494	0.621	-19.809	11.834
Old_School_HACKETT ELEMENTARY SCHOOL	-3.1381	13.681	-0.229	0.819	-29.953	23.676
Old_School_MERRILL ELEMENTARY SCHOOL	0.1513	13.032	0.012	0.991	-25.391	25.693
Old_School_ROBINSON ELEMENTARY SCHOOL	-59.9221	15.470	-3.874	0.000	-90.242	-29.602
Old_School_TODD ELEMENTARY SCHOOL	-11.7325	8.441	-1.390	0.165	-28.276	4.811
Old_School_NULL	-194.6822	11.121	-17.506	0.000	-216.479	-172.886
child_health_level_very_good_compare_1	11.1926	8.834	1.267	0.205	-6.122	28.507
Majority_Group_1	21.8101	9.032	2.415	0.016	4.109	39.512
Eng_Prof_Full_1	7.2934	12.507	0.583	0.560	-17.220	31.807

Omnibus:	81.509	Durbin-Watson:	1.631
Prob(Omnibus):	0.000	Jarque-Bera (JB):	202.820
Skew:	-1.742	Prob(JB):	9.08e-45
Kurtosis:	6.241	Cond. No.	16.4

Notes:
[1] Standard Errors are heteroscedasticity robust (HC1)

With only 9 explanatory variables we were able to explain almost 25% of the variation in Math Scores, with an R-squared of 0.248. The combined effect of all our explanatory variables is **extremely strong** with an F-Statistic of 66.01 with an associated p-value of 9.98 with 10 to the power of -56. Within the 5% significance level, we observe that Majority_Group_1 and Robinson Elementary School are significant with p-values of 0.016 and 0.000 respectively.

First Finding: While controlling for different schools, races, health levels, and English proficiencies, we found that a student who studied at Robinson Elementary School has on average a Math RIT Score that **is less** than other student groups **by 60 points.**

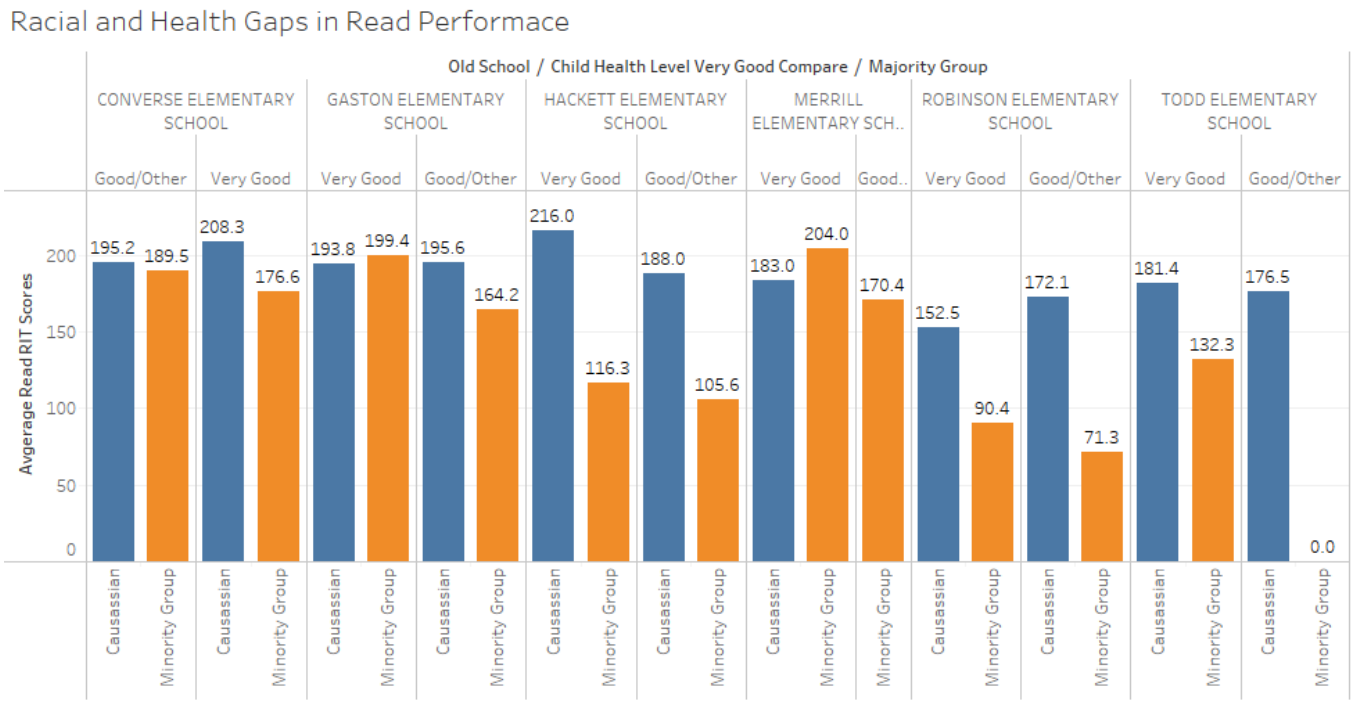
Note: The base group here is Converse Elementary School

Second Finding: While controlling for different schools, races, health levels, and English proficiencies, we found that a student who is Caucasian has on average a Math RIT Score that **is more** than the minority student groups **by 23 points.**

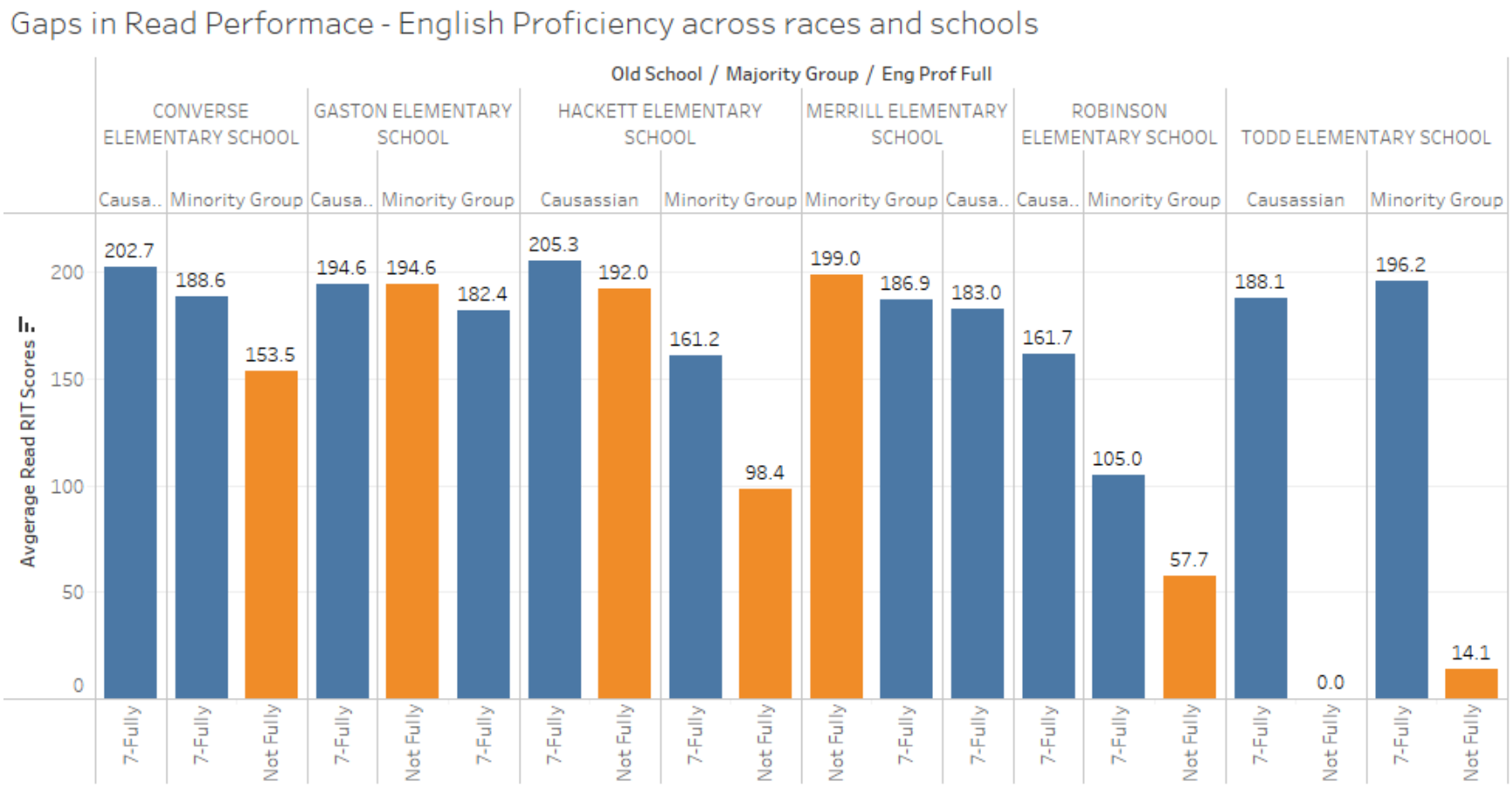
Note: The base group here is the Minority group

Third Finding: Health level is not found to be significant at the 5% level of significance, contradicting our hypothesis.

Interactions for Read Scores



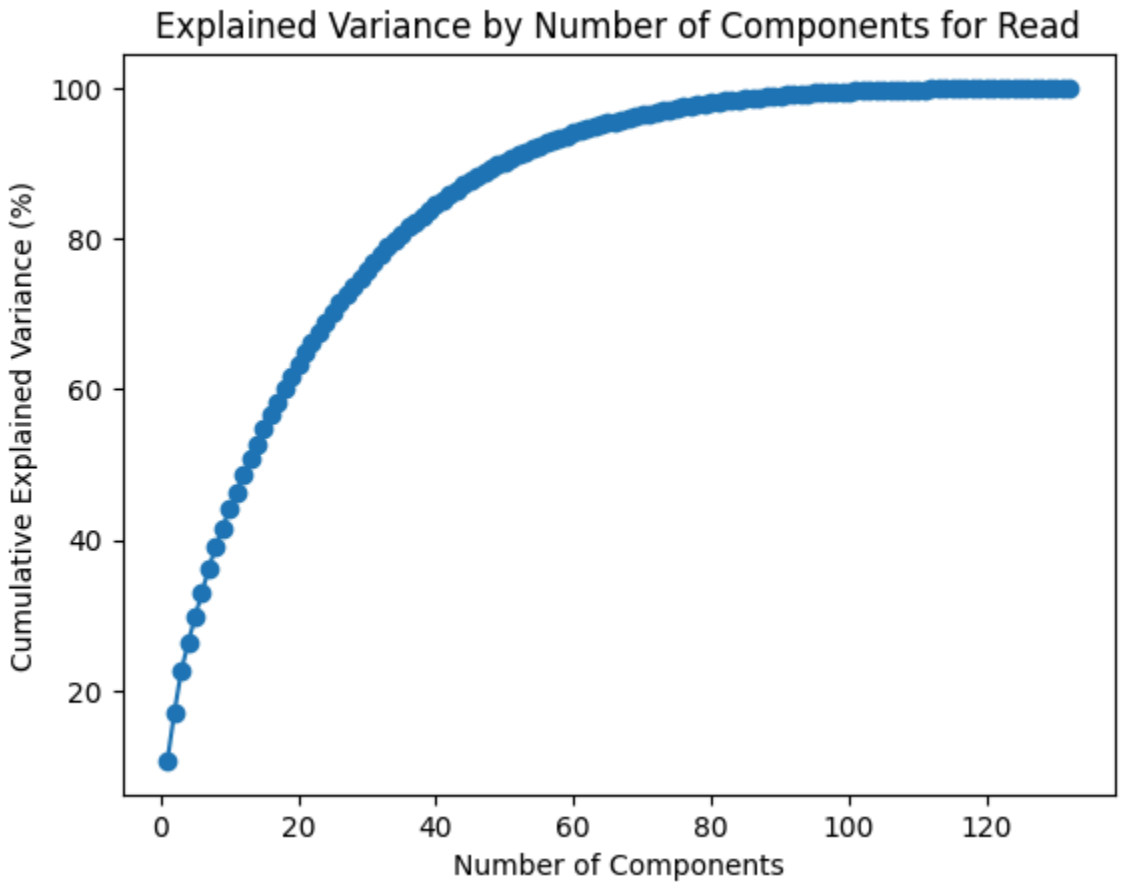
Similarly, we see that belonging to a minority group has a significant effect in decreasing the average Read RIT Scores. From the above graph, we see that both Robinson Elementary School and Todd Elementary School are behaving poorly across the Read subject. However, in terms of health level, we can only see a weak slight trend where students belonging to the 'Very Good' health level tend to perform better than those who belong to the other group.



This time, we are again controlling for racial groups, and across schools, we clearly see the trend that within a single group, a higher English proficiency score will lead to a higher average Read RIT Test Score.

Principal Component Analysis of Read Scores

To better understand the variation in math scores, we decided to perform a PCA analysis on our data.



As we can see from the above graph, 80% of the variance is explained by almost 30 explanatory variables. Finding the right explanatory variables out of more than 120 variables is quite a challenging task, therefore we decided to use Lasso Regression to find out what would be the most significant features.

Feature Selection - Results of Lasso Regression for Math Scores

Our analysis revealed that the following features were significant for Read Scores:

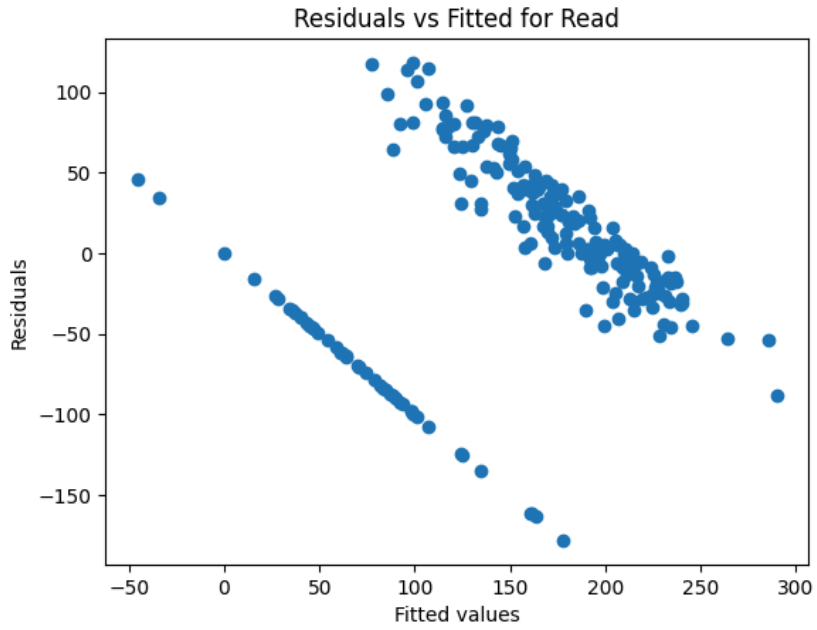
- num_adults_in_house
- num_children_in_house
- Old_School_MERRILL ELEMENTARY SCHOOL
- Old_School_ROBINSON ELEMENTARY SCHOOL

- Old_School_TODD ELEMENTARY SCHOOL
- parent_relationship_Mother
- num_child_change_school_1.0
- cat_father_work_hours_31 - 40 hours
- cat_parent_real_aspire_child_edu_level_Bachelor Degree
- cat_freq_communication_school_staff_Sometimes
- cat_freq_read_with_child_3-4 times per week
- cat_freq_help_homework_3-4 times per week
- cat_freq_regulate_homework_Rule strict en
- Majority_Group_1
- Eng_Prof_Full_1
- child_health_level_very_good_compare_1

Checking for Heteroskedasticity for Math Scores

We used White's test and Breusch-Pagan's test to verify the heteroskedasticity of our model. Here are the results:

- White's Test p-value for reading: 0.6695573113445809
- Breusch-Pagan p-value for Read: 0.15721323355530223



Consequently, we did not apply robust standard errors in our analysis.

Checking for Multicollinearity for Math Scores

We also verified for multicollinearity in our model to ensure that it is robust and not prone to statistical errors. Due to the high dimensionality of our model and its explanatory variables, we tried to filter out any variables with a correlation of more than 0.7 (in absolute value).

We only found that these 2 variables were highly correlated: 'cat_freq_read_with_child_NULL' and ' cat_freq_help_homework_NULL'. We decided to leave them in our model. For the rest of our features, we concluded that the multicollinearity was not strong enough.

Regression Results for Read Scores: The result below demonstrates the regression model that we use

OLS Regression Results

Dep. Variable:	WT18_19ReadRIT	R-squared:	0.536
Model:	OLS	Adj. R-squared:	0.367
Method:	Least Squares	F-statistic:	3.180
Date:	Sun, 10 Nov 2024	Prob (F-statistic):	6.79e-09
Time:	16:21:07	Log-Likelihood:	-1172.9
No. Observations:	215	AIC:	2462.
Df Residuals:	157	BIC:	2657.
Df Model:	57		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	58.2332	91.904	0.634	0.527	-123.295	239.762
num_adults_in_house	-10.4862	5.866	-1.788	0.076	-22.072	1.100
num_children_in_house	-1.6138	4.587	-0.352	0.725	-10.673	7.445
Old_School_GASTON ELEMENTARY SCHOOL	-8.6983	16.585	-0.524	0.601	-41.456	24.059
Old_School_HACKETT ELEMENTARY SCHOOL	-11.4140	19.625	-0.582	0.562	-50.177	27.349
Old_School_MERRILL ELEMENTARY SCHOOL	10.1845	21.520	0.473	0.637	-32.322	52.691
Old_School_ROBINSON ELEMENTARY SCHOOL	-71.8691	17.903	-4.014	0.000	-107.230	-36.508
Old_School_TODD ELEMENTARY SCHOOL	-50.1094	16.923	-2.961	0.004	-83.535	-16.684
Old_School_NULL	-218.2802	58.623	-3.723	0.000	-334.072	-102.489
parent_relationship_Legal Guardian	7.0044	51.973	0.135	0.893	-95.653	109.662
parent_relationship_Mother	-34.2905	13.591	-2.523	0.013	-61.134	-7.447
parent_relationship_NULL	-34.1069	32.505	-1.049	0.296	-98.311	30.097
num_child_change_school_1.0	20.8647	12.098	1.725	0.087	-3.031	44.760
num_child_change_school_2.0	-7.0286	18.006	-0.390	0.697	-42.594	28.537
num_child_change_school_3.0	17.6563	18.950	0.932	0.353	-19.773	55.085
num_child_change_school_4.0	16.1751	26.529	0.610	0.543	-36.224	68.574
num_child_change_school_5.0	10.1024	39.474	0.256	0.798	-67.867	88.072

num_child_change_school_4.0	16.1751	26.529	0.610	0.543	-36.224	
num_child_change_school_5.0	10.1024	39.474	0.256	0.798	-67.867	88.072
num_child_change_school_6.0	75.1112	76.335	0.984	0.327	-75.664	225.887
num_child_change_school_NULL	69.1653	78.476	0.881	0.379	-85.841	224.171
cat_father_work_hours_31 - 40 hours	12.8341	35.848	0.358	0.721	-57.972	83.640
cat_father_work_hours_41 - 50 hours	17.5579	35.161	0.499	0.618	-51.891	87.007
cat_father_work_hours_51 - 60 hours	31.9812	40.831	0.783	0.435	-48.667	112.629
cat_father_work_hours_More than 60 hours	10.3344	42.825	0.241	0.810	-74.253	94.922
cat_father_work_hours_NULL	33.1297	37.946	0.873	0.384	-41.821	108.080
cat_father_work_hours_The child mother does not work	-9.2436	36.974	-0.250	0.803	-82.275	63.788
cat_father_work_hours_Less than 20 hours	38.5539	43.734	0.882	0.379	-47.830	124.938
cat_parent_real_aspire_child_edu_level_Bachelor Degree	-13.0855	25.266	-0.518	0.605	-62.991	36.820
cat_parent_real_aspire_child_edu_level_High School diploma or GED	-2.9926	33.936	-0.088	0.930	-70.022	64.037
cat_parent_real_aspire_child_edu_level_I don't know	2.9934	32.384	0.092	0.926	-60.971	66.957
cat_parent_real_aspire_child_edu_level_Less than High School	34.7545	78.707	0.442	0.659	-120.706	190.215
cat_parent_real_aspire_child_edu_level_Master Degree, PhD, JD, MD or other advanced degree	-16.8186	26.367	-0.638	0.524	-68.898	35.260
cat_parent_real_aspire_child_edu_level_NULL	13.7054	51.098	0.268	0.789	-87.224	114.634
cat_parent_real_aspire_child_edu_level_Some College	-61.6388	28.302	-2.178	0.031	-117.540	-5.738
cat_freq_communication_school_staff_Never	29.9363	77.935	0.384	0.701	-124.001	183.873
cat_freq_communication_school_staff_Often	46.0746	69.719	0.661	0.510	-91.633	183.782
cat_freq_communication_school_staff_Rarely	34.2698	70.302	0.487	0.627	-104.589	173.129
cat_freq_communication_school_staff_Very Regularly	46.0108	71.255	0.646	0.519	-94.732	186.753
cat_freq_communication_school_staff_Sometimes	51.7462	69.067	0.749	0.455	-84.675	188.167
cat_freq_read_with_child_3-4 times per week	-34.6858	16.038	-2.163	0.032	-66.363	-3.008
cat_freq_read_with_child_Almost every day (5 - 6 times per weeek)	-19.4347	19.346	-1.005	0.317	-57.647	18.778
cat_freq_read_with_child_Everyday	-27.4730	20.445	-1.344	0.181	-67.855	12.909
cat_freq_read_with_child_G	-29.1991	70.807	-0.412	0.681	-169.057	110.658
cat_freq_read_with_child_NULL	18.7484	91.594	0.205	0.838	-162.167	199.664
cat_freq_read_with_child_Never	-20.0769	37.209	-0.540	0.590	-93.571	53.417
cat_freq_read_with_child_Rarely	-13.9868	15.414	-0.907	0.366	-44.433	16.460
cat_freq_help_homework_3-4 times per week	11.8369	16.148	0.733	0.465	-20.058	43.732
cat_freq_help_homework_Almost every day (5 - 6 times per weeek)	0.8214	16.387	0.050	0.960	-31.546	33.189
cat_freq_help_homework_Everyday	12.0235	18.495	0.650	0.517	-24.508	48.555
cat_freq_help_homework_NULL	42.5062	106.007	0.401	0.689	-166.878	251.891

cat_freq_help_homework_Almost every day (5 - 6 times per weeek)	0.8214	16.387	0.050	0.960	-31.546	33.189
cat_freq_help_homework_Everyday	12.0235	18.495	0.650	0.517	-24.508	48.555
cat_freq_help_homework_NULL	42.5062	106.007	0.401	0.689	-166.878	251.891
cat_freq_help_homework_Never	-39.9955	43.798	-0.913	0.363	-126.504	46.513
cat_freq_help_homework_Rarely	24.3377	21.467	1.134	0.259	-18.063	66.738
cat_freq_regulate_homework_No rule	79.4344	38.995	2.037	0.043	2.411	156.457
cat_freq_regulate_homework_Rule mod. enf	64.3460	37.295	1.725	0.086	-9.319	138.011
cat_freq_regulate_homework_Rule no enf	75.6696	45.134	1.677	0.096	-13.478	164.817
cat_freq_regulate_homework_Rule strict enf	80.5712	36.037	2.236	0.027	9.391	151.751
child_health_level_very_good_compare_1	9.8979	10.894	0.909	0.365	-11.619	31.415
Majority_Group_1	27.5678	12.875	2.141	0.034	2.137	52.999
Eng_Prof_Full_1	60.6572	13.681	4.434	0.000	33.635	87.680

Omnibus:	8.577	Durbin-Watson:	1.711
Prob(Omnibus):	0.014	Jarque-Bera (JB):	8.443
Skew:	-0.462	Prob(JB):	0.0147
Kurtosis:	3.297	Cond. No.	181.

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

With only 57 explanatory variables we were able to explain almost 50% of the variation in Read Scores, with an R-squared of 0.532. The combined effect of all our explanatory variables is **very strong** with a F-Statistic of 3.180 with an associated p-value of 6.79 with 10 to the power of -9.

Within the 5% significance level, we observe the following variables with their respective p-values:

Variable	p-value
Old_School_ROBINSON ELEMENTARY SCHOOL	0.000
Old_School_TODD ELEMENTARY SCHOOL	0.004
Old_School_NULL	0.000
parent_relationship_Mother	0.013
cat_parent_real_aspire_child_edu_level_Some College	0.031
cat_freq_read_with_child_3-4 times per week	0.032
cat_freq_regulate_homework_No rule	0.043
cat_freq_regulate_homework_Rule strict en	0.027
Majority_Group_1	0.034
Eng_Prof_Full_1	0.000

First Finding: While controlling for different schools, races, health levels, English proficiencies, and several parental involvement variables, we found that a student who studied at Robinson Elementary School has on average a Read RIT Score that **is less** than other student groups **by 71 points.** Also, a student who studied at Todd Elementary School has on average a Read RIT Score that **is less** than other student groups by **50 points.**

Note: The base group here is Converse Elementary School

Second Finding: While controlling for different schools, races, health levels, English proficiencies, and several parental involvement variables, we found that a Caucasian student has on average a Read RIT Score that **is more** than the minority student groups **by 27.6 points.**

Note: The base group here is the Minority group

Third Finding: While controlling for different schools, races, health levels, English proficiencies, and several parental involvement variables, we found that a student who is English proficient has on average a Read RIT Score that **is more** than the student groups with lesser English Proficiencies **by 60 points.**

Note: The base group here is the Not Fully English Proficient group (less than 7)

Policy Recommendations

Recommendations:

4. Develop tutoring and mentoring programs, especially for minority students
5. Use culturally responsive pedagogy for those who use English as a second language
6. Revitalize Robinson Elementary

Justifications: Currently, the average score for minority groups is lower than their Caucasian peers; this indicates that students in these ethnic groups need targeted tutoring and mentoring programs to improve their academic performance. Firstly, we recommend the School District of Beloit develop a small-group tutoring and one-on-one mentoring services specifically for students from minority groups to bridge the academic performance gap in Math and Reading. Moreover, we found a correlation between level of English proficiency and the academic outcomes of these students, those who are not fully proficient, have worse academic performance. We recommend that schools within the system all integrate a culturally responsive pedagogy, which means teachers will be train to have specific teaching approaches to those who use English as a second language, and the tailored tutoring programs that we mentioned above.

Robinson Elementary is a notable case within the system. Students from this school face several challenges that hinder their ability and potential in both Math and Reading. These kids consistently struggle compared to other peers within the system. Even after controlling for variables including race, health, and level of English proficiency, the scores underperformed on average. The score disparities are also huge, with students from minority backgrounds scoring significantly lower than their Caucasian peers, suggesting that there's something that needs to be done about the equity issues within the school environment. Families are also reported to lack the capacity or knowledge to engage effectively in this school. All of these factors suggest that Robinson Elementary is facing distinct challenges, perhaps from the curriculum, teacher training,

or even the misallocation of resources. Thus, it needs a comprehensive school revitalization program.

The Wallace Foundation shows that schools that invest in teacher training and infrastructure would result in high student achievement (Grissom, 2021). They emphasize that well-resourced schools and effective leadership will directly lead to student outcomes. More resources should be available for this school when they initiate their tutoring programs, teacher training with a focus on cultural responsiveness, infrastructure building, and more health and wellness initiatives. Although Robinson Elementary is underperforming, we would not recommend it be shut down immediately. The sudden close-down of one school in the system would lead to multiple issues, including overcrowding for other nearby schools, and the confusion and stress for students and families currently enrolling in this school, potentially leading to even lower performance of students when they transfer to other schools (Kirshner et al., 2010). The sudden shutdown indicates the failure to address the root cause of the underperformance and denial of the community. Even if the school is underperforming after implementing the set of recommendations, we suggest turning Robinson Elementary into an after-school institution to offer the Beloit community parent education courses and other additional services that can support the overall School District of Beloit.

Limitations:

1. Missing data points, potential response bias in self-reporting data, and limitations on causal inference due to the observational nature of the data
2. The representation of each category is not uniform
3. The limited time prevents us from looking at the time-series data to provide more in-depth analysis and recommendations for the problems
4. The health information of the students should be more specific rather than just subjective health level as per the baseline; for example, weight, whole health index, healthy eating index, etc.

Works Cited

Kirshner, B., Gaertner, M., & Pozzoboni, K. (2010). Tracing Transitions: The Effect of High School Closure on Displaced Students. *Educational Evaluation and Policy Analysis*, *32*(3), 407–429. http://www.jstor.org/stable/40963085

Grissom, J. A., Egalite, A. J., & Lindsay, C. A. (2021). *How Principals Affect Students and Schools: A Systematic Synthesis of Two Decades of Research*. https://doi.org/10.59656/el-sb1065.001