MGT 6203 Group Project Final Report

Spring 2023

Team 68
Members: Philip Murray, Jessica Wurst, Daniel Catlin, Ian Francis, Mardoqueu Mesquita

OBJECTIVE/PROBLEM

Project Title: Food Deserts and Educational Impact

Background Information:

In this project, we discuss Food deserts, which are residential areas within the United States that have limited access to affordable and nutrient-rich foods. The food atlas, displays large areas within the United States, referred as geographic areas that lack sufficient access to grocery stores. These areas tend to have higher rates of abandoned or vacant homes, lower levels of education, higher unemployment, and lower income. Their access to food can also be limited by distance to grocery stores and lack of public transportation. The U.S. Department of Agriculture estimates that more than 13.5 million people (about twice the population of Arizona) live in a food desert, with 82% of those people living in urban areas. Our project aimed to analyze and exploit potential relationships and causes. In our conclusions, we provide recommendations for areas of future research.

Problem Statement:

Food access and availability are closely related to wellness, which disproportionately affects low-income communities. There are several factors driving this; however, this analysis focuses on limited food availability due to insufficient physical access to food, i.e., food deserts. This project's purpose is to better understand the relationship between food access and academic success in urban areas to show the impact of aiding underserved communities.

Primary Research Question (RQ):

Does improvement in food access correlate with academic success in urban areas?

Supporting Research Questions:

1. What is the correlation between food access and academic success?
2. Has the correlation remained the same over time?
3. Can the correlation be isolated from other factors?

Business Justification:

A high school dropout will cost the economy over $270,000 more over their lifetime than their peers who graduated, driven by lower taxes, a reliance on government services, and an increased incarceration rate. [R1]

Literature Survey:

There is a significant amount of academic research that has been conducted in the field of food insecurity and academic growth of children. *Agriculture & Food Security*, a journal that addresses the challenge of global food security, found that household food insecurity has a strong contribution to students' poor school attendance [R2]. *The University of Alberta* ran various statistical analyses and came to similar conclusions that low household food security is associated with poor academic achievement, especially in reading and mathematics [R3]. As mentioned in the business justification, *The National Center for Education Statistics* quantified the significant impact on the economy that a high school dropout can have [R1].

These are only to name a few sources which are supporting evidence of how researchers are better understanding the issue of food insecurity and academic success. Moreover, there is a new understanding of the economic potential of food deserts. Research shows that States and local businesses have established programs to meet the food access needs of underserved communities. Those programs make available programmatic resources tied to their specific initiative, land, technical assistance and/or coordinated philanthropic resources [R4].

DATASET SOURCES AND DESCRIPTION

Data Sources:

- USDA Food Access Atlas Download
- 'EdFacts' - U.S. Department of Education
- County / Metro Area Lookup
- United States 2020 Census – School District Reference Map
- School District to Census Tract Lookup Table

Data Description:

- USDA Food Access Atlas contains 3 years of county-level food access data points (approximately ~220K rows w/ 147 variables)
- The EDFACTs data set is a breakdown of a school district's N-size population and the % proficient in the state reading level test. In the original dataset, demographic fields are available for RNO (Race National Origin), Gender, Disability, Limited English Proficiency, homeless, Migrant, Foster Care, and Military Connected student bodies. To limit the scope of this project, only the Full Student Body Full Grade groups were used along with the Economically Disadvantaged Student columns.
- The County / Metro area lookup table contains a breakdown of county name, state name, metropolitan area (Atlanta, Chicago, etc.), and state abbreviation
- The School District to Census Tract Lookup Table contains numeric Census Tract Identifiers that tie to the school district, allowing us to aggregate data to the school district level

Initial Variables:

The 147 variables in the Food Access Atlas can be potentially significant predictor variables and the dependent variable will be fourth grade or high school reading levels. We anticipate variables such as ‘low-income population count beyond 10 miles from a supermarket’ and ‘poverty rate’ will be correlated to educational success.

Initial Hypotheses:

Our initial hypothesis stated that high levels of food access in urban/metropolitan communities positively correlate with academic success; yet it would be difficult to isolate food access and low-income data's relative impact on academic success due to potential multicollinearity effects. The team’s approach to use various regression techniques and different success metrics was fundamental to allow us to draw reasonable conclusions on said hypothesis.

What business decisions will be impacted by the results of your analysis? What could be some benefits?

We would be able to identify which urban zones would benefit the most from subsidized Grocery store development. In fact, we found that the combination of Low-income-Low-Access between 1/2 and 10 miles, appeared as a significant factor through regression; thus, allowing us to identify areas with low reading rates and low-income and low-access to grocery store which could be potential candidates for assistance.

GENERAL MODELING APPROACH:

The team will join, cleanse, and aggregate the various data tables described above into a data set that will be used to run various regression and clustering models. The regression and clustering models will help to understand the food access predictor variables that are statistically significant with educational success, whether the relationship can be isolated from specific factors, and if there is empirical evidence supporting the expansion of subsidies and relief in certain food deserts. The team will use exploratory data analysis, variable selection methods, and modeling techniques learned in this course to drive these results.

The data set was filtered to only include counties which were identified as central to Metropolitan areas.

DATA MANIPULATION AND EXPLORATORY DATA ANALYSIS

Data Manipulation:

The team leveraged both SQL and R for data ingestion, manipulation, and aggregation. To do so, the team constructed an R file that imports both the DBI and RSQLite packages. The DBI package allowed the team to create a relational database system in the R file and the RSQLite package embeds a database engine interface paired with the DBI package. Using these tools, the team instantiated a local database, assigned file pathways to variables, and wrote to a database object to import multiple datasets for easy querying and manipulation.

The first group of datasets the team modified were the USDA Food Access Research Atlas across our three years of data. The Food Access Research data is broken down on the Census Tract level (geographical region in a county for census tracking) with associated counties and states. For each of the Census Tracts, there are approximately 147 variables related to the area's food access characteristics. The team leveraged both the county to metropolitan area lookup file and the school district to census tract lookup to create new variables in the data set that will allow the team to aggregate food access data at the census tract, school district, county, or metropolitan level.

The second group of datasets the team modified is an export of the U.S. Department of Education's data for Reading and Language Arts (RLA) test scores. The yearly data is broken down by school district, grade (4 – 12), the number of students taking the exams, and the percentage of students that passed their RLA exams. The team leveraged the school district to county lookup file, the county to metropolitan area lookup file, and the school district to census tract lookup file to create new variables in the data set that will allow the team to group educational success data at the census tract, school district, county, or metropolitan level.

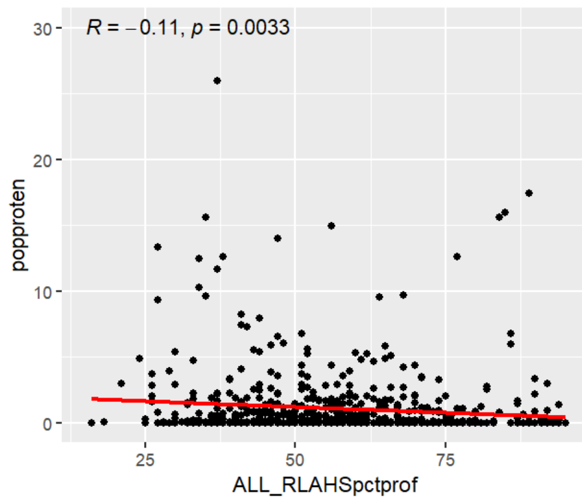
The final stage of data preparation was combining the two previously mentioned datasets and cleaning up any missing data or poorly formatted variables. The team combined the educational and food access datasets by merging on the census tract unique ID number. For additional data cleansing, any school districts that had null values for RLA exam scores or Census Tracts that had missing food access values were removed from the data set. For any educational or food access variables that had interval data, the team took the midpoint of the intervals and updated the dataset.

Exploratory Data Analysis:

The team started by choosing a handful of predictors to investigate and create some visuals to try and understand the initial relationship between our predictors and response. The initial charts identified that Poverty Rate and Low Income, Low Access data was slightly correlated with reading levels, while simply Low Access children by itself was likely not statistically significant [3].

This first investigation of variable distribution and correlation charts indicated that the metropolitan area was too large of an aggregation for these data points. The team hypothesized that this is due to many potentially significant predictors to explain the variability in educational success on more granular levels were suppressed because of the large amount of variation when joined at the metropolitan area. In addition, there are metropolitan areas that contain multiple counties and aggregating at that high a level reduced the diversity of our sample. Lastly, suburban areas may have been aggregated with urban areas. The team determined the next best aggregation level to examine would be the county level.

Now that the educational and food access data was aggregated to the county level, the team modeled the relationship between the State Reading Level pass rates and various food access related predictors using correlation tests [4]. The High School student reading proficiency did have a notable example of statistical significance with its correlation with the percent of the county's population that lived within 10 miles of a grocery store. Unfortunately, this model may be overfitted since ~41% of the 670 counties tested do not have a population that lives more than 10 miles away from a supermarket and an additional ~35% had less than 1% of their population living in such areas, and 19% of counties had less than 5% of its population living in a tract 10miles away from a supermarket. In summary, any model would be developed on ~5% of the counties in our data.

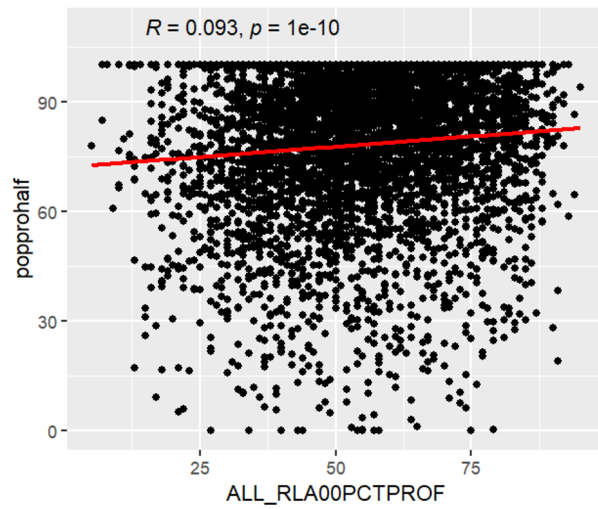


Correlation Test for HS Reading Proficiency and % of County > 10m from Grocery

Given the initial struggles of finding enough quality data, the team started searching for additional data or to change the scope of the problem itself. At this point we found a way to map new school district academic data with our census tract food access data. An extra level of granularity and additional data points gave the team hope that there were still quality conclusions to draw on the relationship between food access and educational success.

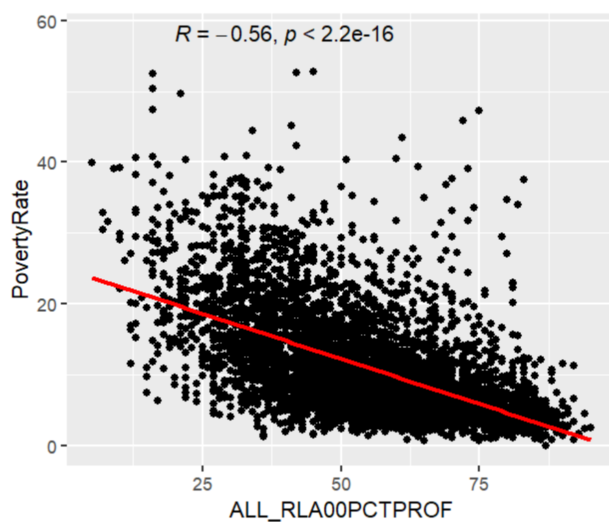
School District Level Data

Mapping the data to the school district provided us with a sample size of approximately four thousand as opposed to 671 data points for metropolitan counties. The team experimented with combining each of the grade level reading proficiency scores and testing across the same subset of predictors. Firstly, the correlation significance testing was much better than the initial data set, but the R-squared value remained very weak [5]. The team confirmed that the same conclusions were drawn from running the correlation tests on each of the individual grade's reading proficiency scores.



Correlation Test for the District Reading Proficiency and % of County > .5m from Grocery

As a comparison to the low access models, the team evaluated the school district poverty rate and all grade reading scores. Poverty Rate provided a much better linear fit than the previously examined predictors for distance from a supermarket indicating economics are a significant factor in the reading level scores.



Correlation Test of Poverty Rate and All Grade Reading Proficiency Scores

The visualizations and correlation tests indicated that the school district aggregation was the most appropriate level of specificity. The next section will go in more depth surrounding the modeling process the team took for school district data.

LINEAR REGRESSION APPROACH

Variable Selection Methods / Parameter Optimization:

Prior to beginning modeling, it was hypothesized that the further away from a supermarket, the children's population, the lower the reading score would be. This would mean we would expect to see a negative correlation between the reading level proficiency and the proportion of children in a low access area. To identify the best variables for the model, we utilized forward, backward, and stepwise variable selection methods [10].

After analyzing the results we determined the best variables to build our models on are median family income (MedianFamilyIncome), poverty rate (PovertyRate), proportion of residents living in an urban tract (Urban), proportion of residents living who are considered low income living in a low access area .5 - 19 miles from a supermarket (lilahalf10), and the proportion of kids who live in a low access area at least .5 miles from a supermarket (lakidshalf). The p-values for the model indicated that all the factors were significant at a greater than 99.9% confidence interval, so no features were dropped from the final model and the R-squared was determined to be strongest when all factors were present.

The two strongest factors were median family income and poverty rate. By itself, median family income had an r squared of .39+ vs. .44+ found in the final model indicating a strong correlation between economic factors and reading level proficiency.

From the variables selected, the best performing model for the entire student body is seen below.

```
Residuals:
    Min      1Q  Median      3Q     Max 
-52.339  -8.120  -0.358   7.741  46.004 

Coefficients:
                     Estimate Std. Error t value Pr(>|t|)    
(Intercept)         3.426e+01  1.731e+00  19.791  < 2e-16 ***
PovertyRate        -3.244e-01  4.821e-02  -6.728 2.02e-11 ***
MedianFamilyIncome  2.160e-04  9.542e-06  22.633  < 2e-16 ***
lilahalf10         -8.398e+00  1.146e+00  -7.326 2.94e-13 ***
Urban               2.915e+00  7.082e-01   4.116 3.94e-05 ***
lakidshalfp         5.914e+00  1.249e+00   4.734 2.29e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 12.35 on 3370 degrees of freedom
Multiple R-squared:  0.4375,    Adjusted R-squared:  0.4367 
F-statistic: 524.2 on 5 and 3370 DF,  p-value: < 2.2e-16
```

Linear Regression Model Summary

Considering that the US student population is greater than 18 million, having an R-squared of 43.67% gives the team relatively high confidence that the model has been tuned well. When considering the number of factors that likely have a measurable effect on a student's reading level proficiency, such as the economic factors we identified above, being able to identify low access areas to food as an important factor is an important step towards testing our hypothesis.

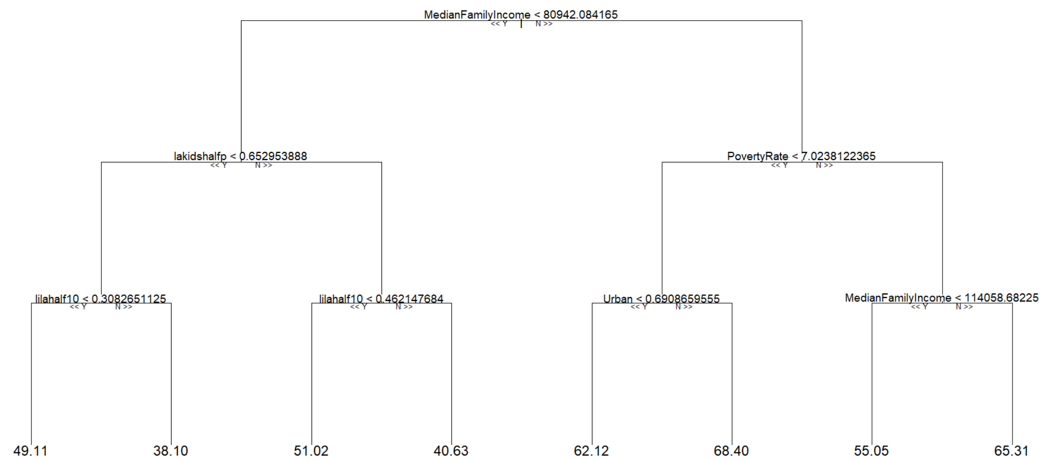
While the R-squared value is relatively high, one area where the model could be improved is the RMSE, which is approximately 12. Considering the possible range of values for reading level proficiency is 0-100, an RSME of 12 could be interpreted as a 12% error rate. One notable issue with the Reading Level Proficiency scores is that a majority of them were originally a range of values. Further analysis would need to be done on the original range of values to determine how far outside the original range of scores the RMSE is, but this is one area that could be improved on.

When the model was run against the individual class student bodies (4^{th}, 8^{th}, and HS) the adjusted R-squared dipped into the 30s for all 3 groups while the P value remained relatively high. Additionally, Urban was no longer a significant factor for either HS and 8^{th} grade groups and the proportion of kids who live in a low access area at least .5 miles from a supermarket (lakidshalfp) was no longer significant for HS. This could potentially indicate a diminishing effect of these variables on reading level proficiency as students grow older, but such an impact would need to be studied in future research.

Categorical Model Overview:

We additionally sought to test a clustering model, however after running several iterations, it was determined that the performance of any clustering model would be heavily affected by our proportion of school districts that either had 100% or 0% of its population the selected distance from the supermarket. As an example of this issue, our highest correlated distance variable was used with our student population reading proficiency response variable with the recommended number of clusters from our elbow chart analysis [6]

Due to the clustering model proving unsuccessful, the team pivoted and built a random forest model with the same variable set as the linear regression model we built. After cross validating to determine the optimal number of nodes for our data [7], we selected a forest model with 8 nodes and produced the output below.



Random Forest

After running this model, a similar RMSE to the linear model above (~12.9) with an R-squared of ~.459 were the results of the model indicating very similar performance, which is what we should expect considering we are using the same variables. What stands out in the model is the separation of the economic factors to the low access factors. Once the first split occurs, low access factors are used for the remaining four nodes. This could possibly indicate an increased effect of low access within lower economic populations. As expected, the two factors MedianFamilyIncome and PovertyRate remain the highest impact [8].

PROJECT CONCLUSIONS

Poverty Rate, percentage of population in an urban area, and the percentage of population in low-income census tracts with low accessibility at ½ to 10 miles, were the most significantly significant predictors of reading proficiency across all models.

Low food access alone did not always indicate lower reading scores. Percent or population with Low access at ½ mile and 20 miles were both positively correlated with reading rates. However, the percentage of population with low access at 1 mile and 10 miles was negatively correlated with reading proficiency rates. While these factors were over a 0.05 significance rate, they are not as significant as the factors mentioned above, highlighting the importance of economic and environmental factors in understanding the impact of low food access.

Through this analysis, we concluded that low food access in low-income areas impacts academic performance. To maximize the impact of government subsidized grocery store construction, low-income low access areas should be targeted.

When filtering data for reading rates at 20% or Lower, Urban Areas, and Low-income Low Access rates of greater than 75%, the following school districts are returned. Many counties return to multiple districts and many of the metropolitan areas have multiple districts.

State	AreaName	County	SchoolDistrict	POP2010	Pct Low Income / Low Access	Percent Proficent in
Alabama	Birmingham-Hoover. AL	Jefferson County	Midfield City School District	7,130	1	19
Alabama	Birmingham-Hoover. AL	Jefferson County	Tarrant City School District	4,989	1	19
Arizona	Phoenix-Mesa-Glendale. AZ	Maricopa County	Isaac Elementary District	31,406	1	19
Arizona	Phoenix-Mesa-Glendale. AZ	Maricopa County	Murphy Elementary District	12,269	1	16
Colorado	Denver-Aurora-Broomfield. CO	Arapahoe County	Sheridan School District 2	9,232	1	19
Illinois	Chicago-Joliet-Naperville. IL-IN-WI	Cook County	Dolton School District 148	11,109	1	10
Illinois	Chicago-Joliet-Naperville. IL-IN-WI	Cook County	Dolton School District 149	22,576	1	13
Illinois	Chicago-Joliet-Naperville. IL-IN-WI	Cook County	Hazel Crest School District 152-5	2,902	1	13
Illinois	Chicago-Joliet-Naperville. IL-IN-WI	Cook County	Hoover-Schrum Memorial School District 157	5,873	1	18
Illinois	Chicago-Joliet-Naperville. IL-IN-WI	Cook County	Thornton Township High School District 205	64,870	0.76	10
Illinois	Chicago-Joliet-Naperville. IL-IN-WI	Cook County	West Harvey-Dixmoor Public School District 147	3,537	1	9
Illinois	Chicago-Joliet-Naperville. IL-IN-WI	Kane County	Aurora East Unit School District 131	83,848	0.93	18
Illinois	Chicago-Joliet-Naperville. IL-IN-WI	Lake County	Zion Elementary School District 6	14,808	0.77	16
Illinois	Chicago-Joliet-Naperville. IL-IN-WI	Will County	Bloom Township High School District 206	3,144	1	17
Illinois	Decatur. IL	Macon County	Decatur School District 61	50,237	0.82	12
Illinois	Peoria. IL	Tazewell County	Creve Coeur School District 76	5,509	1	12
Illinois	St. Louis. MO-IL	Madison County	Madison Community Unit School District 12	3,061	1	8
Illinois	St. Louis. MO-IL	Madison County	Venice Community Unit School District 3	1,890	1	10
Illinois	St. Louis. MO-IL	St. Clair County	Cahokia Community Unit School District 187	19,222	1	5
Illinois	St. Louis. MO-IL	St. Clair County	East St. Louis School District 189	42,297	0.88	13
Indiana	Chicago-Joliet-Naperville. IL-IN-WI	Lake County	Gary Community School Corporation	75,443	0.84	17
Kansas	Kansas City. MO-KS	Wyandotte County	Kansas City Unified School District 500	117,284	0.79	19
Michigan	Battle Creek. MI	Calhoun County	Battle Creek Public Schools	45,037	0.94	19
Michigan	Detroit-Warren-Livonia. MI	Macomb County	East Detroit Public Schools	38,812	0.83	20
Michigan	Detroit-Warren-Livonia. MI	Wayne County	River Rouge School District	7,903	0.76	16
Michigan	Detroit-Warren-Livonia. MI	Wayne County	Westwood Community Schools	23,131	1	17
Michigan	Flint. MI	Genesee County	Beecher Community School District	10,079	1	10
Michigan	Flint. MI	Genesee County	Flint City School District	102,395	0.98	13
Michigan	Flint. MI	Genesee County	Westwood Heights Schools	2,841	1	16
Missouri	St. Louis. MO-IL	St. Louis County	Normandy Schools Collaborative	38,592	0.96	15
Missouri	St. Louis. MO-IL	St. Louis County	Riverview Gardens School District	38,321	1	15
New Jersey	Atlantic City-Hammonton. NJ	Atlantic County	Egg Harbor City School District	4,243	1	15
New Jersey	Philadelphia-Camden-Wilmington. PA-NJ-DE-MD	Camden County	Lindenwold Borough School District	17,613	0.87	19
Oklahoma	Oklahoma City. OK	Oklahoma County	Crutcho Public School	4,098	1	7
Oklahoma	Oklahoma City. OK	Oklahoma County	Millwood Public Schools	3,177	1	11
Oklahoma	Oklahoma City. OK	Oklahoma County	Western Heights Public Schools	22,087	0.76	13
Pennsylvania	Pittsburgh. PA	Allegheny County	Duquesne City School District	5,565	1	17
Rhode Island	Providence-New Bedford-Fall River. RI-MA	Providence County	Woonsocket School District	41,186	0.77	16
Wisconsin	Janesville. WI	Rock County	Beloit School District	37,262	0.85	18

Unfortunately, the team was not able to investigate whether the relationship between food access and educational success has changed over time. The team would dig deeper into this secondary analysis question if there was additional time.

Appendix:

1. USDA Food Access Data Screenshot

a.

CensusTract	State	County	Urban	Pop2010	OHU2010	GroupQuartersFlag	NUMGQTRS
01001020100	Alabama	Autauga County	1	1912	693	0	0
01001020200	Alabama	Autauga County	1	2170	743	0	181
01001020300	Alabama	Autauga County	1	3373	1256	0	0

2. National Center for Education Statistics

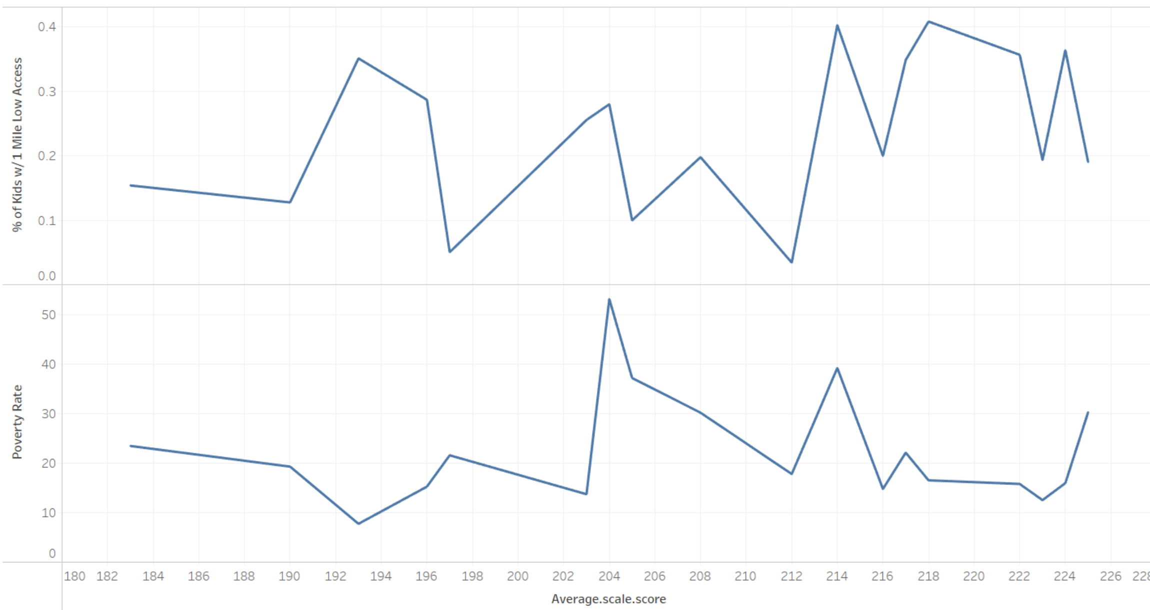
a.

National Center for Education Statistics

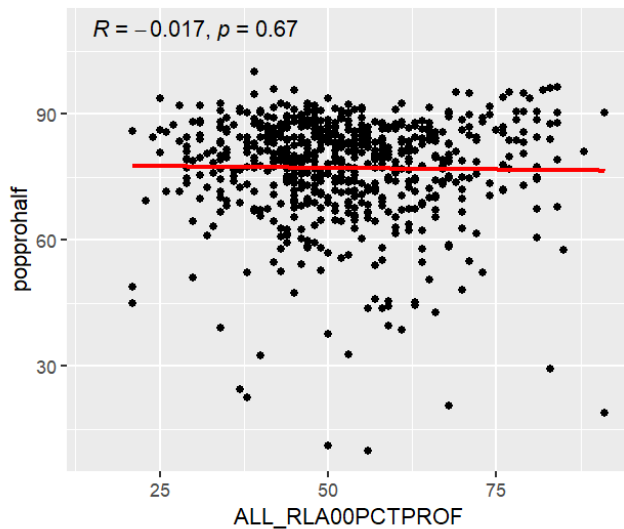
2022 NAEP Mathematics Assessment at Grade 4: Trend Results for Percentage Distribution, Average Score, Percentile Score

Jurisdiction	Statistics	Variable	Category	2022	Significance comparison results	Prior assessment year	2022	Prior assessment year	Difference between 2022 and prior years
NP	PERCENT	GENDER	MALE	2022		2019	51	51	-0.1
NP	PERCENT	GENDER	FEMALE	2022		2019	49	49	0.1
NP	PERCENT	GENDER	MALE	2022		2017	51	51	-0.2
NP	PERCENT	GENDER	FEMALE	2022		2017	49	49	0.2
NP	PERCENT	GENDER	MALE	2022		2015	51	51	0.0
NP	PERCENT	GENDER	FEMALE	2022		2015	49	49	0.0
NP	PERCENT	GENDER	MALE	2022		2013	51	51	-0.2
NP	PERCENT	GENDER	FEMALE	2022		2013	49	49	0.2

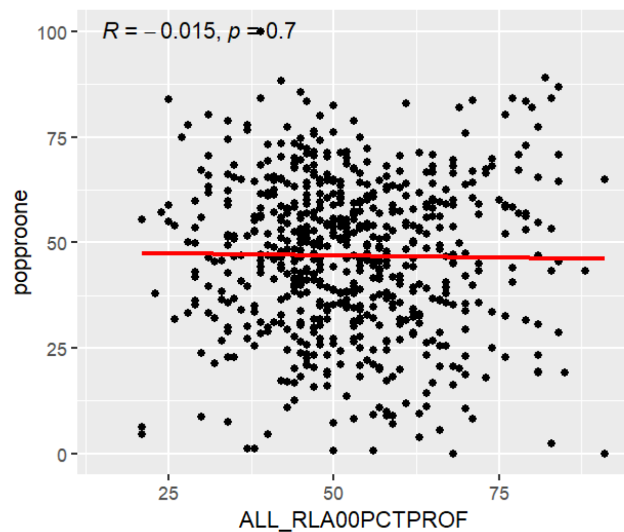
3. Variable Analysis on Poverty Rate, % of Kids w/1 Mile of Grocery, and 4th Grade Reading Success



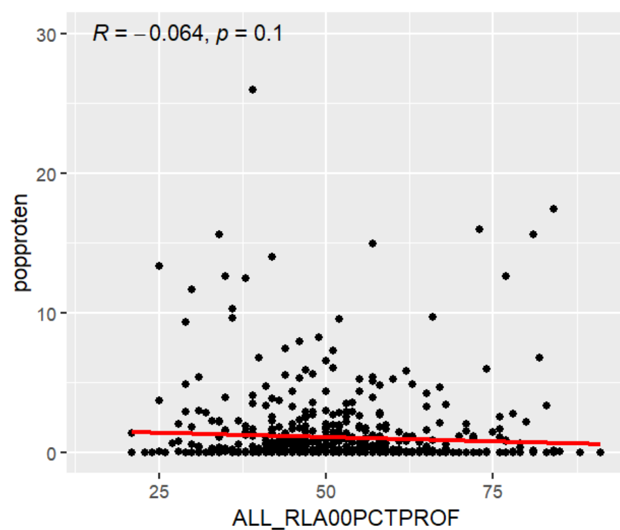
4. County Level Correlation Tests


R = −0.017, p = 0.67
popprohalf
ALL_RLA00PCTPROF
90
60
30
25
50
75

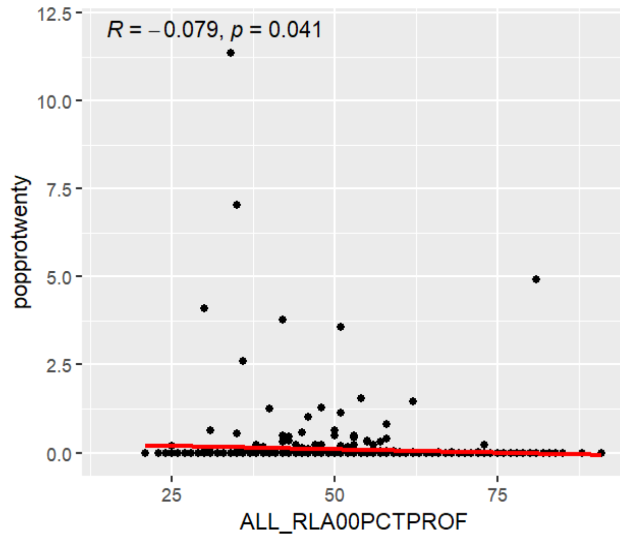
a.


R = −0.015, p = 0.7
popproone
ALL_RLA00PCTPROF
100
75
50
25
0
25
50
75

b.


R = −0.064, p = 0.1
popproten
ALL_RLA00PCTPROF
30
20
10
0
25
50
75

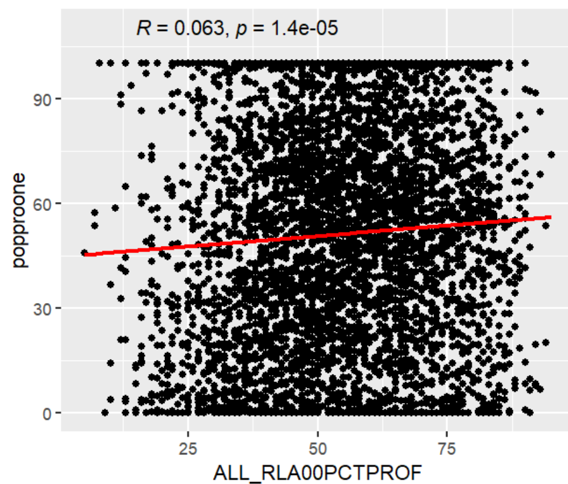
c.



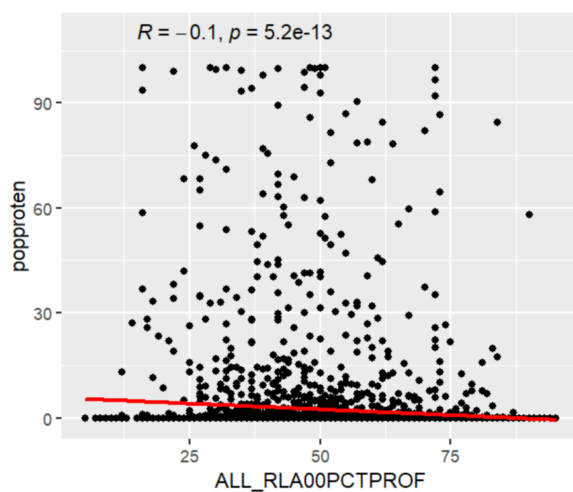
d.

5. School District Correlation Tests

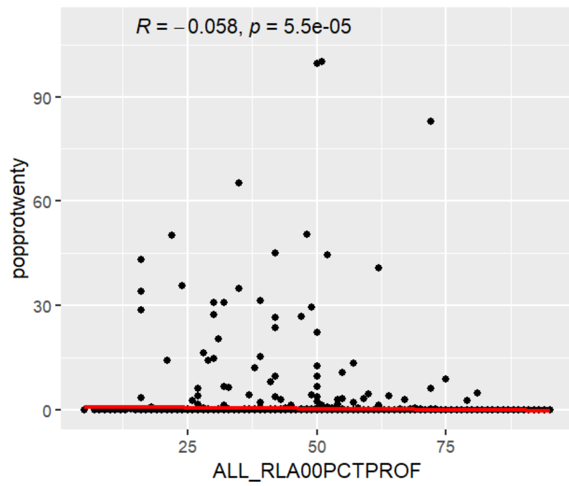
a.



b.



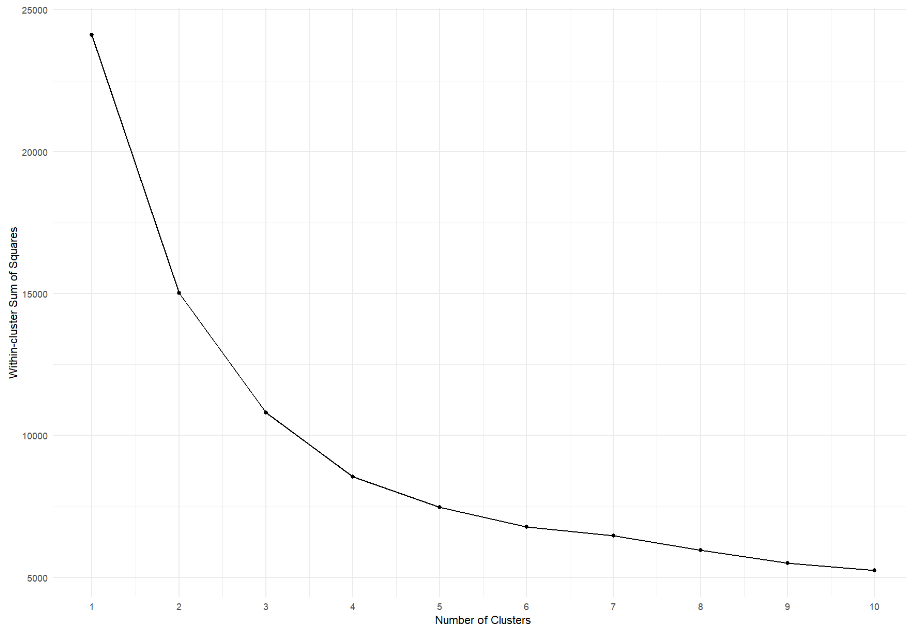
c.



d.

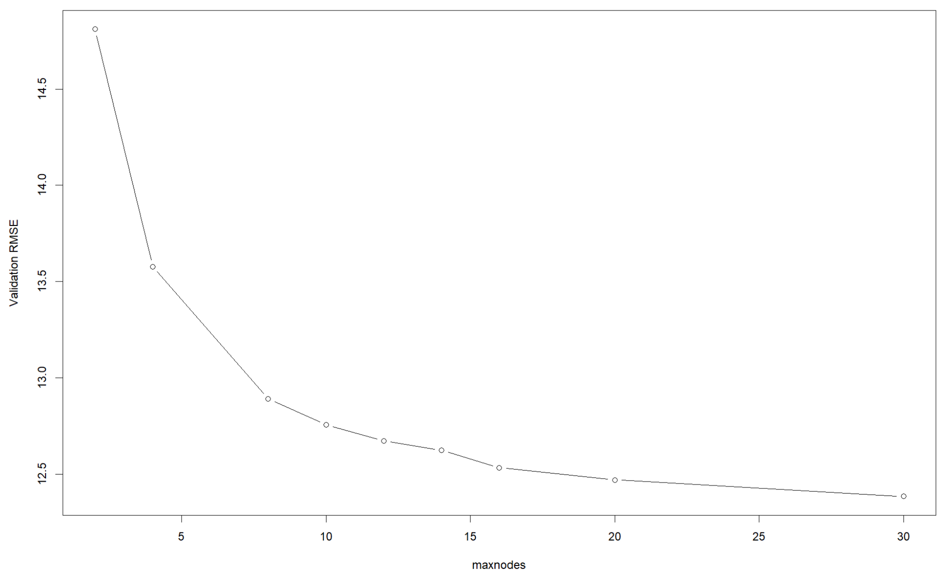
6. Clustering Analysis Output

a.



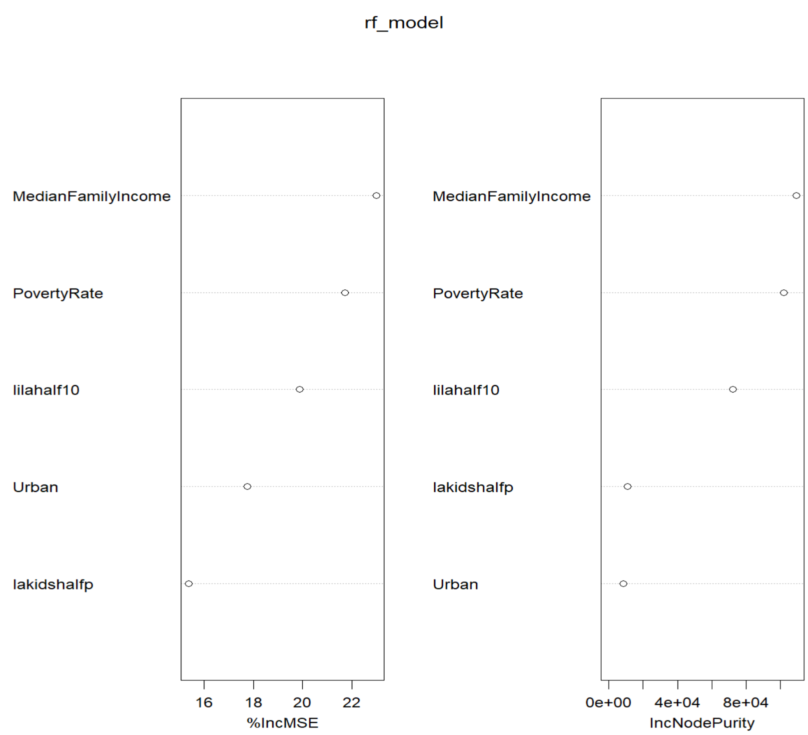
b.

7. Cross Validation for Random Forest



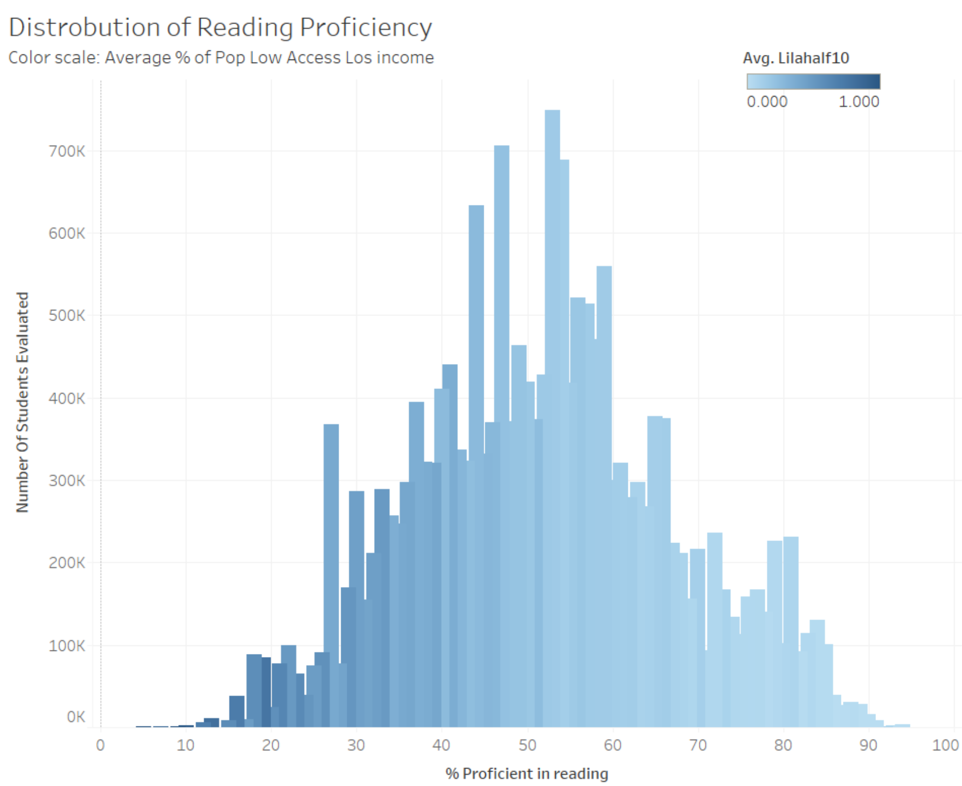
a.

8. Variable Impact on Random Forest Model



a.

9. Distribution of Reading Proficiency and Students Evaluated



a.

10. Regression Model Variable Selection Output

Significant Variables, Forward Select

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.565557e+01	2.009509e+00	17.743419	2.994144e-68
NUMGQTRS	2.881220e-04	1.383776e-04	2.082144	3.738288e-02
MedianFamilyIncome	2.015816e-04	8.081931e-06	24.942256	3.985617e-129
lahisphalf	4.624561e-04	2.323541e-04	1.990307	4.661446e-02
TractKids	8.686456e-04	3.328593e-04	2.609648	9.092051e-03
TractHispanic	-2.276041e-04	8.472223e-05	-2.686474	7.246240e-03
lakidshalfp	2.561378e+01	1.106350e+01	2.315162	2.064660e-02
lakids1p	2.090437e+01	7.967182e+00	2.623810	8.723102e-03
lilahalf10	-1.128033e+01	1.414738e+00	-7.973436	1.919049e-15
lila110	5.964130e+00	1.630737e+00	3.657322	2.576314e-04
popproone	-2.519192e-01	8.200638e-02	-3.071946	2.138752e-03
popproten	-8.261138e-02	2.766823e-02	-2.985784	2.842984e-03

Significant Variables, Backward Select

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.565557e+01	2.009509e+00	17.743419	2.994144e-68
NUMGQTRS	2.881220e-04	1.383776e-04	2.082144	3.738288e-02
MedianFamilyIncome	2.015816e-04	8.081931e-06	24.942256	3.985617e-129
lahisphalf	4.624561e-04	2.323541e-04	1.990307	4.661446e-02
TractKids	8.686456e-04	3.328593e-04	2.609648	9.092051e-03
TractHispanic	-2.276041e-04	8.472223e-05	-2.686474	7.246240e-03
lakidshalfp	2.561378e+01	1.106350e+01	2.315162	2.064660e-02
lakids1p	2.090437e+01	7.967182e+00	2.623810	8.723102e-03
lilahalf10	-1.128033e+01	1.414738e+00	-7.973436	1.919049e-15
lila110	5.964130e+00	1.630737e+00	3.657322	2.576314e-04
popproone	-2.519192e-01	8.200638e-02	-3.071946	2.138752e-03
popproten	-8.261138e-02	2.766823e-02	-2.985784	2.842984e-03

Significant Variables Stepwise

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.603065e+01	1.922845e+00	18.738199	1.176315e-75
POP2010	-4.071536e-04	7.811604e-05	-5.212165	1.945147e-07
NUMGQTRS	3.539283e-04	1.135544e-04	3.116818	1.838998e-03
PovertyRate	-2.069207e-01	9.928676e-02	-2.084071	3.720668e-02
MedianFamilyIncome	2.023646e-04	7.947073e-06	25.464045	3.101894e-134
lalowihalf	2.492904e-04	1.146143e-04	2.175038	2.967617e-02
lakidshalf	-1.161190e-03	2.798647e-04	-4.149111	3.395711e-05
lablackhalf	6.922657e-04	1.465302e-04	4.724388	2.374310e-06
laasianhalf	5.250900e-04	2.515108e-04	2.087743	3.687387e-02
laomultirhalf	4.702878e-04	1.919696e-04	2.449804	1.432906e-02
lahisphalf	3.138639e-04	7.285465e-05	4.308083	1.679824e-05
lasnaphalf	-2.331275e-03	3.137080e-04	-7.431354	1.266600e-13
lapop1	1.118560e-02	3.548019e-03	3.152633	1.628071e-03
lawhite1	-1.090460e-02	3.546472e-03	-3.074773	2.118528e-03
lablack1	-1.137634e-02	3.537507e-03	-3.215920	1.308936e-03
laasian1	-1.100904e-02	3.648619e-03	-3.017317	2.563674e-03
laaian1	-1.163304e-02	3.574135e-03	-3.254786	1.142686e-03
laomultir1	-1.261586e-02	3.634353e-03	-3.471280	5.225688e-04
lanhopi10	-8.882867e-01	1.857774e-01	-4.781458	1.792593e-06
lahunv10	4.680307e-02	2.054308e-02	2.278289	2.275334e-02
lapop20	3.486721e-02	1.700688e-02	2.050183	4.040104e-02
lalowi20	2.112837e-02	9.542484e-03	2.214137	2.686641e-02
lawhite20	-4.168317e-02	1.725102e-02	-2.416273	1.571754e-02
laaian20	-4.840713e-02	1.768237e-02	-2.737593	6.211987e-03
TractLOWI	1.389420e-04	6.282686e-05	2.211506	2.704798e-02
TractKids	8.542516e-04	2.003622e-04	4.263537	2.050990e-05

```
TractWhite          4.101197e-04 7.141507e-05  5.742762  9.891006e-09
TractAsian          2.459812e-04 1.008436e-04  2.439235  1.475455e-02
TractNHOPI         -2.201636e-03 9.427347e-04 -2.335372  1.956519e-02
TractHispanic      -1.767861e-04 3.855552e-05 -4.585234  4.650624e-06
lakidshalfp         2.646745e+01 1.067437e+01  2.479534  1.318964e-02
lakids1p            2.000826e+01 7.204117e+00  2.777336  5.502022e-03
lilahalf10         -1.114481e+01 1.382691e+00 -8.060236  9.541967e-16
lila110             5.941037e+00 1.588259e+00  3.740597  1.857358e-04
popproone          -2.395508e-01 7.464981e-02 -3.208994  1.340811e-03
popproten          -7.118581e-02 2.340708e-02 -3.041208  2.369093e-03
```

References

1. Trends in High School Dropout and Completion Rates in the United States. National Center for Education Statistics. (2017). Retrieved March 8, 2023, from https://nces.ed.gov/programs/dropout/intro.asp#ref4
2. Tamiru, D., Belachew, T. The association of food insecurity and school absenteeism: systematic review. Agric & Food Secur **6**, 5 (2017). https://doi.org/10.1186/s40066-016-0083-3
3. The association between food insecurity and academic achievement in Canadian school-aged children Erin L Faught, Patty L Williams, Noreen D Willows, Mark Asbridge and Paul J Veugelers. https://tfss.ca/wp-content/uploads/2018/05/association_between_food_insecurity_and_academic_achievement_in_canadian_schoolaged_children.pdf
4. PolicyMap. 2018. How Opportunity Zones Can Expand Access to Fresh Food in America's Food Deserts [online]. Philadelphia, PA, and Los Angeles, CA. [Cited 12 April 2023]. https://www.policymap.com/2018/09/how-opportunity-zones-can-expand-access-to-fresh-food-in-americas-food-deserts/