

Impacts of Food Deserts, Racial Background, and Health on Academic Performance

MINK: mdlu02, ireyesp107, nhaseley, kkianaa

Hypothesis

Academic performance can influence a person's life by unlocking opportunities that would otherwise be difficult to access. Although K-12 education is provided to all children in the US, there are many factors outside of access to schools that impact academic performance. In this project, we decided to focus on the effects of food accessibility, racial background, and health conditions, such as sleep deprivation and mental health status, on the academic performance of students on their grade-level standardized tests. More specifically, we hypothesized that being located in a food desert and having poor health both had negative impacts on academic performance.

Data

We joined four datasets: [Food Access Research Atlas](#), [CDC PLACES: Census Tract Data](#), [National Center for Education Statistics \(NCES\) School District Geographic Relationship](#), and [Stanford Education Data Archive \(SEDA\)](#). The Food Access Research Atlas has information that was used to define areas as food deserts. The CDC PLACES dataset has information on mental health status and sleep. The NCES dataset included census tract information for local education agency identification numbers which helped join the SEDA dataset to the first two datasets. The SEDA dataset included the standardized test scores of students from grades 3-8. The datasets were not dirty but did have missing values. After removing these, we still had over 90,000 data points which was sufficient for our analysis.

Findings

There is a statistically significant difference in academic performance across race groups.

Using pairwise two-sample t-tests, we compared the mean education scores across racial groups. All pairwise t-tests resulted in statistically significant differences ($p < 0.05$) between all racial groups (Table 1). Negative t-statistics indicated that the "Race 1" group had a higher mean score and vice versa.

Race 1	Race 2	t-statistic	p-value	Significance
White	Asian	-41.321	0.0000e+00	Significant
White	Black	364.736	0.0000e+00	Significant

White	Hispanic	299.276	0.0000e+00	Significant
White	Native American	244.930	0.0000e+00	Significant
Asian	Black	364.126	0.0000e+00	Significant
Asian	Hispanic	305.830	0.0000e+00	Significant
Asian	Native American	259.130	0.0000e+00	Significant
Hispanic	Black	73.368	0.0000e+00	Significant
Hispanic	Native American	-31.568	3.4019e-218	Significant
Black	Native American	-98.411	0.0000e+00	Significant

There is a significant difference in average education scores between urban/rural food deserts and non-food desert communities.

We used a one-sample t-test to compare the mean education scores of students in food deserts against the standard known mean score of students in non-food desert communities. As shown in Table 2, the p-value was significant ($p = 0.0000e+00$) when comparing the scores of students in both urban/rural food deserts and those of students in non-food deserts with students in non-food deserts performing better.

Food Desert Type	Non-Food Desert Baseline Mean	t-statistic	p-value	Significance
Urban	5.371	-172.606	0.0000e+00	Significant
Rural	5.371	-47.711	0.0000e+00	Significant

There is a significant difference in sleep deprivation between urban/rural food deserts and non-food desert communities.

We used pairwise two-sample t-tests to compare the mean prevalence of sleep deprivation across urban/rural food desert and non-food desert areas. As shown in Table 3, the resulting p-values were significant ($p = 0.0000e+00$) and showed that non-food deserts had less sleep deprivation ($t\text{-statistic} = 426.850/100.032$).

Group 1	Group 2	t-statistic	p-value	Significance
Urban Food Desert	Non-Food Desert	426.850	0.0000e+00	Significant
Rural Food Desert	Non-Food Desert	100.032	0.0000e+00	Significant

There is a statistical difference in academic performance across racial groups when comparing communities with high versus low rates of sleep deprivation.

We used two-sample t-tests to compare the mean educational scores across racial groups in areas labeled as having high or low prevalence of sleep deprivation. The resulting p-values and t-statistics, displayed in Table 4, were significant and indicated that high rates of sleep deprivation are correlated with low academic performance ($p = 0.0000e+00$; t-statistic = -34.330).

Race 1	Race 2	t-statistic (high sleep deprivation prevalence)	t-statistic (low sleep deprivation prevalence)	Significance (p-value < 0.005)
White	Asian	-36.306	-22.686	Significant
White	Black	239.147	276.969	Significant
White	Hispanic	194.147	229.291	Significant
White	Native American	158.128	187.958	Significant
Asian	Black	253.055	261.920	Significant
Asian	Hispanic	212.198	220.333	Significant
Asian	Native American	179.544	186.981	Significant
Hispanic	Black	51.211	52.539	Significant
Hispanic	Native American	-23.701	-21.038	Significant
Black	Native American	-70.910	-68.372	Significant

Health factors are a stronger predictor of academic performance than food deserts.

We used XGBoost for a supervised prediction task with academic performance and health features as inputs and food desert/urban status as our target variable. Our trained models performed well considering the imbalance of our target variable. Examining XGBoost's internal feature importance scores revealed that sleep and mental health were the most impactful factors.

