

Low Income Student Analysis for Illinois Public Schools

Qumars Jadali

Department of Information Technology: University of The Cumberland

DSRT-834-B02: Advanced Statistical Applications

Dr. Kona Jones

12/7/2025

Low Income Student Analysis for Illinois Public Schools

Introduction

The purpose of this analysis is to see whether there is a difference between the percent of students who come from low-income households, that enroll in Illinois high schools versus Illinois middle schools in 2025. Analyzing how low-income student enrollment differs between middle schools and high schools shows how demographics can change as students progress through the Illinois educational system and can help give educational administrators and policymakers more insight on how big this hypothesized change is, and probe further research into discovering the influencing factors of this change. The participant's and research setting in this study include students who go to Illinois public high schools and middle schools, however no personal identifiable information is associated with the students, only the percentage of students who come from low-income households that enroll in high schools and middle schools are included (Illinois State Board of Education, 2025). The data utilized for this analysis is from the 2025 Illinois Report Card Public Data Set (Illinois State Board of Education, 2025). This data set includes a wide array of information that ranges from student demographics, teacher demographics, school types, as well as key performance metrics that the schools are measured with (Illinois State Board of Education, 2025).

Research Question & Variables

The research question for this analysis asks if there is a difference in the percent of students who come from low-income households that enrolled in Illinois public middle schools versus Illinois public high schools. One of variables that is used for this analysis includes the field named 'School Type' which is a categorical variable, and the categories that make up this variable are 'High School' and 'Middle/Junior High School'. The second variable used in this assignment is '% Student Enrollment – Low Income', which is a numeric variable, and refers to the percentage of students that come from a low-income household that enrolled in that particular year, and ranges from zero to 100. The null and alternative hypotheses are as follows:

H₀: There is no significant difference in the percent of students who come from low-income households that enrolled in Illinois public middle schools versus Illinois public high schools.

H_a: There is a significant difference in the percent of students who come from low-income households that enrolled in Illinois public middle schools versus Illinois public high schools.

Statistical Test

The statistical test chosen for this analysis is an independent samples t-test. The reason why an independent samples t-test is the most appropriate test to use for this analysis is that I am looking to see if there is a difference between two independent groups that contain numeric data that is not paired (Spatz et al., 2023).

Results

An independent samples t-test was performed to identify whether there was a difference in the percentage of students who come from low-income households that enrolled in Illinois public high schools versus middle schools. A Levene test found that the assumption of homogeneity of variance has been met ($p = .393$). A Shapiro-Wilk test showed that the two groups did not have data that was normally distributed ($p < .001$). Due to the Shapiro-Wilk value, the assumption of normality could not be met that is required for an independent samples t-test. A Mann-Whitney U test was performed to see if there was a difference in the percentage of students who come from low-income households that enrolled in Illinois public high schools versus middle schools. The percentage of students who come from low-income households that enrolled in Illinois public high schools ($Mdn = 46.80$) did not differ significantly from those who enrolled in Illinois public middle schools ($Mdn = 47.35$), $U = 216317$, $p = .083$. The null hypothesis was not able to be rejected. The rank-biserial correlation was $-.056$, which indicates a large negative effect size between the two groups.

Conclusion & Summary

This research analyzed if there was a difference between the percentage of students that come from low-income households that enrolled in Illinois public high schools versus middle schools. Even though the Levene test showed that the variances are not significantly different, since the Shapiro-Wilk

test showed that the data is not normally distributed, a non-parametric test was still needed to correctly analyze the data. The results of the analysis did not provide significant results, even after adjusting the statistical test to a more appropriate non-parametric test due to the violation of the assumption of normality. This study could be replicated over time to see if the results still continue to be insignificant, and if later on the results do end up having a significant difference, researchers could look into the driving factors that influence the differences in percentages of students that come from low-income households that enrolled in Illinois public high schools versus middle schools.

References

- Illinois State Board of Education. (2025). *2025 Report Card Public Data Set*. [Data set]. Illinois State Board of Education. <https://www.isbe.net/pages/illinois-state-report-card-data.aspx>
- Spatz, C., Peszka, J. J., Kennedy, L., & Sestir, M. (2023). *Exploring statistics: Tales of distributions*. Outcrop Publishers.

Appendix

Independent Samples T-Test

Independent Samples T-Test

	U	df	p	Rank-Biserial Correlation	SE Rank-Biserial Correlation
% Student Enrollment – Low Income	216317		.083	–0.056	0.032

Note. For the Mann-Whitney test, effect size is given by the rank biserial correlation.

Note. Mann-Whitney U test.

Assumption Checks

Test of Normality (Shapiro-Wilk)

Residuals	W	p
% Student Enrollment – Low Income	0.976	< .001

Note. Significant results suggest a deviation from normality.

Test of Equality of Variances (Levene's)

	F	df ₁	df ₂	p
% Student Enrollment – Low Income	0.731	1	1282	.393

Descriptives

Group Descriptives

	Group	N	Mean	SD	SE	Coefficient of variation	Mean Rank	Sum Rank
% Student Enrollment – Low Income	High School	692	50.57	24.35	0.926	0.481	659.1	456095
	Middle/Junior High School	592	47.23	23.71	0.974	0.502	623.1	368876

Descriptive Statistics ▼

	% Student Enrollment – Low Income	
	High School	Middle/Junior High School
Median	46.80	47.35