

Project Purpose and Overview

This project investigates the relationship between socioeconomic conditions and student academic performance in U.S. high schools. The primary focus is on average ACT or SAT scores at the school level, combined with community-level socioeconomic data such as household income, unemployment rate, adult educational attainment, and family structure. By linking these variables, the project seeks to identify whether and how the social and economic environment of a school district influences educational outcomes. Two main datasets are used. The EdGap dataset provides school-level ACT/SAT scores and socioeconomic indicators derived from the U.S. Census Bureau's American Community Survey and state departments of education. The NCES dataset (from the National Center for Education Statistics) supplements this with detailed school information such as school type, level, and location. Together, these sources enable a comprehensive analysis of both academic achievement and the contexts in which schools operate. The purpose of this project goes beyond simply measuring test scores. It aims to reveal patterns of inequality across schools and districts, offering insights into the broader social factors that shape educational opportunities. For instance, examining the relationship between test performance and access to free or reduced lunch programs may help highlight disparities tied to poverty. Similarly, comparing outcomes across different school types or regions can uncover systemic differences in access to educational resources. Finally, the project also emphasizes the ethical handling of educational data. Since the datasets come from public sources, issues of informed consent are minimal. However, attention is given to potential collection bias, data quality, and the importance of protecting sensitive information. A focus on fairness, transparency, and careful interpretation ensures that the results can be used responsibly to inform policy and practice. In short, the purpose of this project is to explore how community socioeconomic conditions connect to student academic performance, with the ultimate goal of contributing to a deeper understanding of educational inequality in the United States.