Education Inequality Analysis Communication

Sam Farias

What is the problem? .

This project aims to tackle the issue of unequal access to education in American high schools by concentrating on the typical performance of students in the ACT or SAT examinations, which are crucial components of college admission requirements. We ask the question does socioeconomic factors impact student testing scores?

One particular question we can explore is: which factors play the most significant role in determining ACT scores?

A specific question that we investigate later in this notebook is the relationship between ACT/SAT scores and socioeconomic factors in Washington state compared to Illinois state. Since these two states share similarities in terms of politics, median income, and the number of children who benefit from the free/reduced price lunch program, we are interested in exploring whether their test scores would be comparable or not

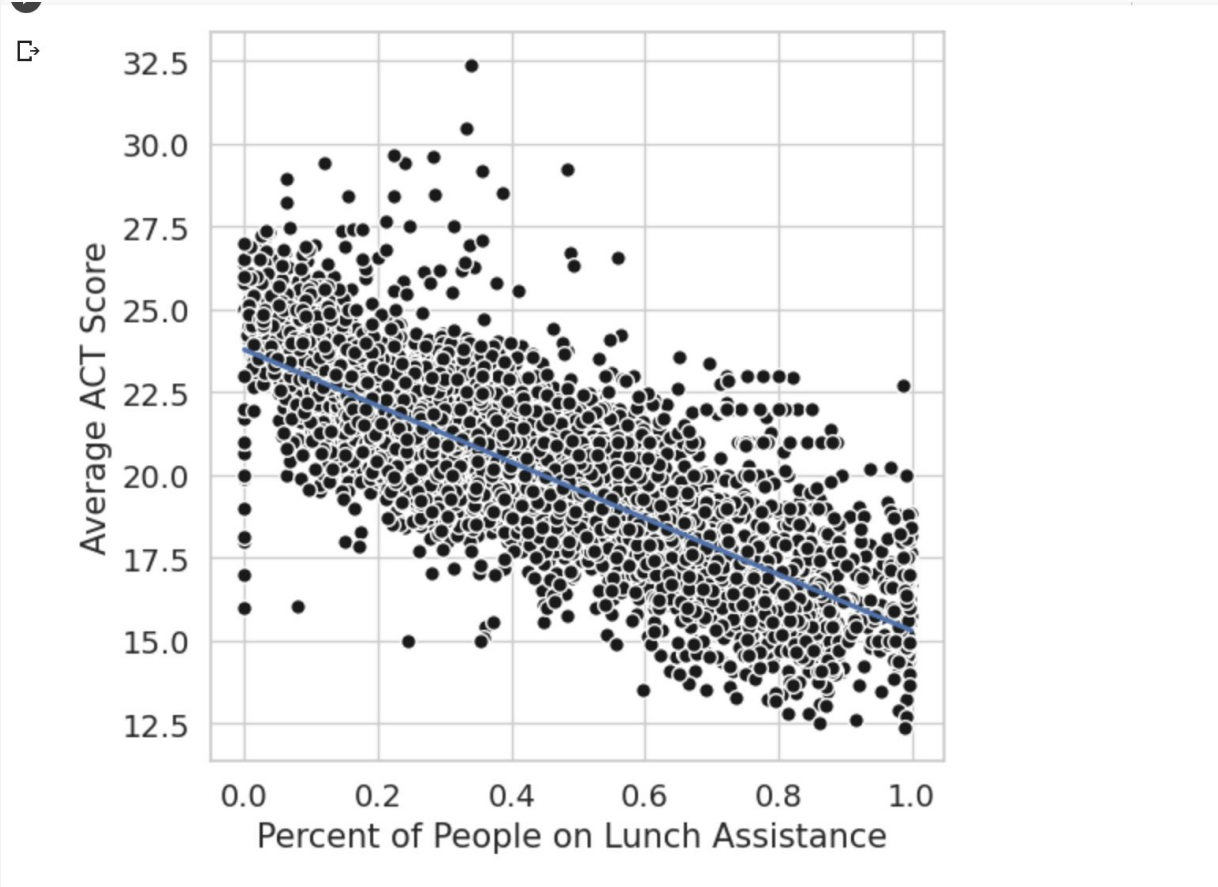
How did I get my data:

The primary source of information utilized is the EdGap data set, sourced from EdGap.org, which features data from 2016 encompassing the average ACT or SAT scores of schools, along with various socio-economic indicators of the school district. The secondary data set consists of fundamental details regarding each school, acquired from the National Center for Education Statistics.

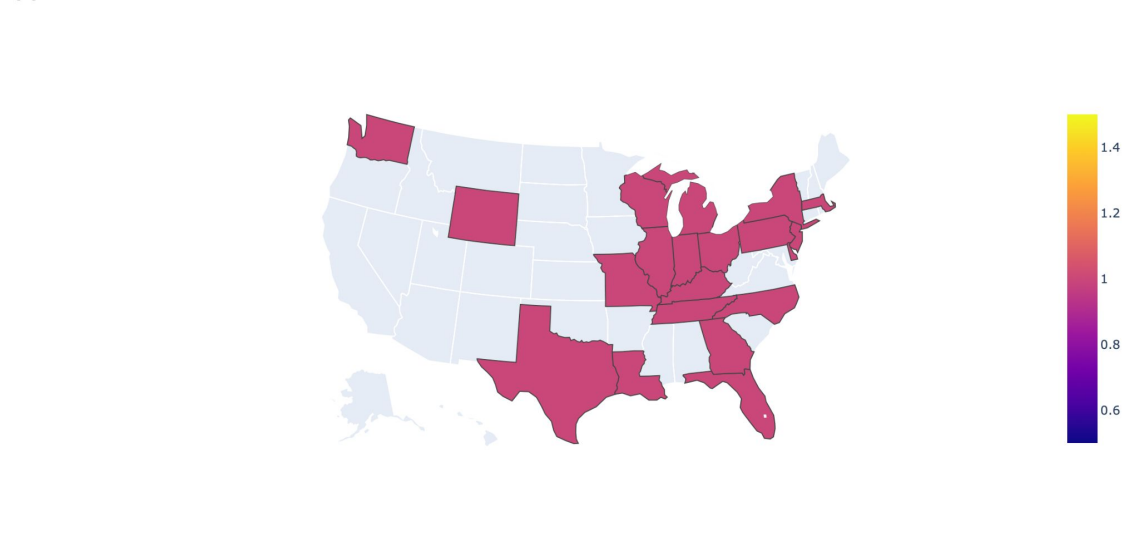
Steps I took:

To begin the data analysis, I imported the necessary libraries and loaded the data. After that, I inspected the contents of each data set and converted the data types of columns to the correct types. Then, I selected relevant subsets of the data and renamed the columns. To combine the data sets, I joined the EdGap and school information data frames while keeping relevant columns. To ensure data quality, I identified missing values in the data sets. I crafted a train test split for further analysis and did data imputation to handle missing values.

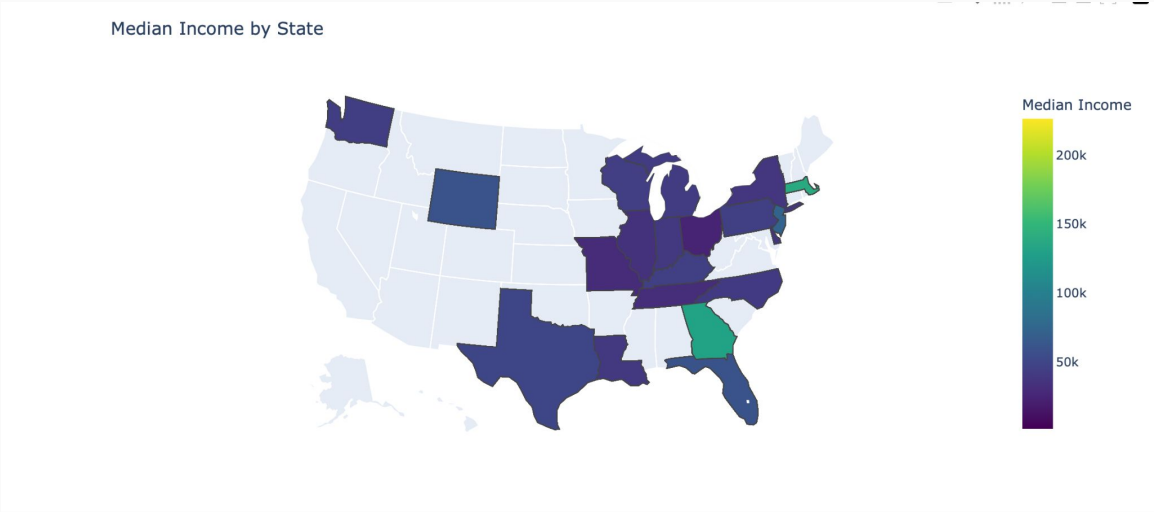
A scatter plot with a regression line for Percent of Students of Assisted Lunch and Average ACT Score.



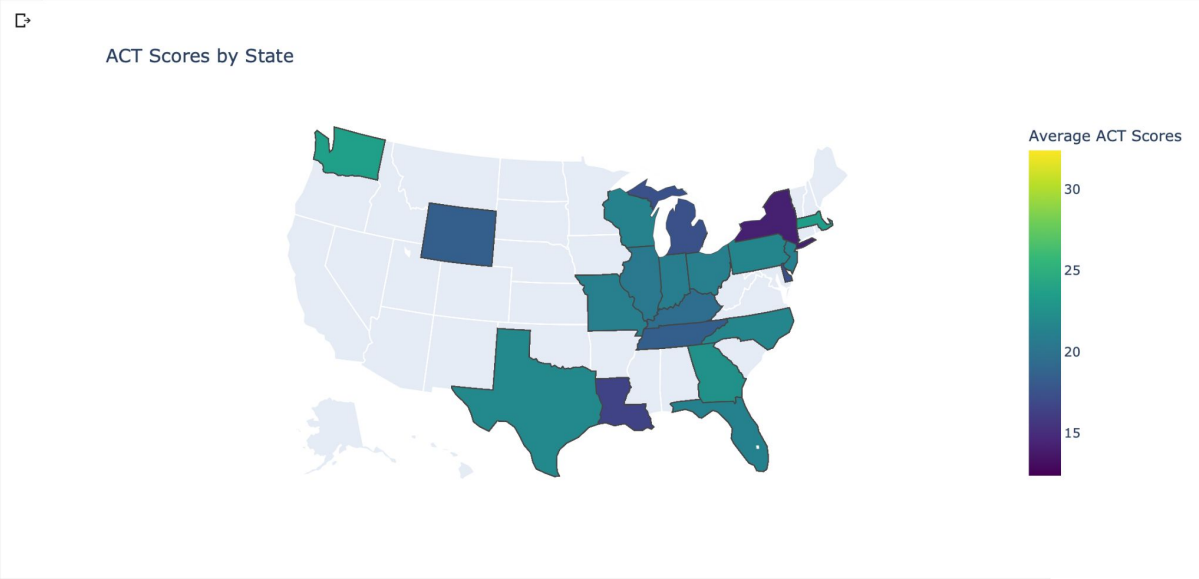
Data being pulled:



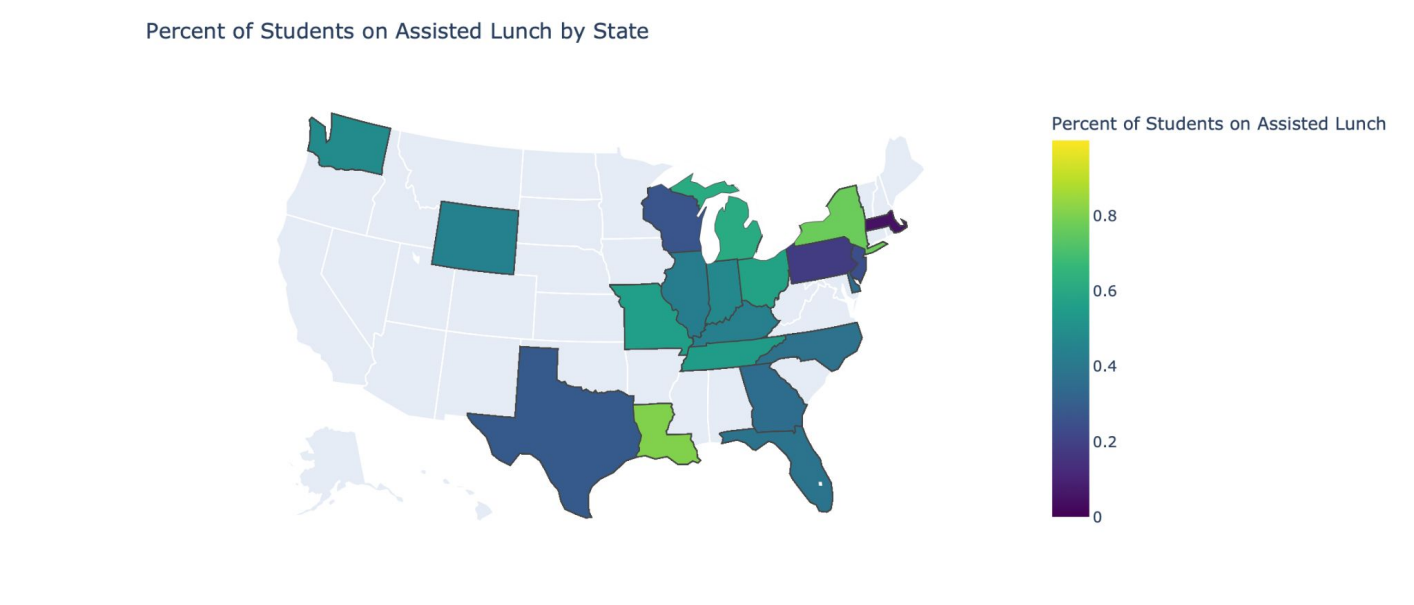
Median Income by State:

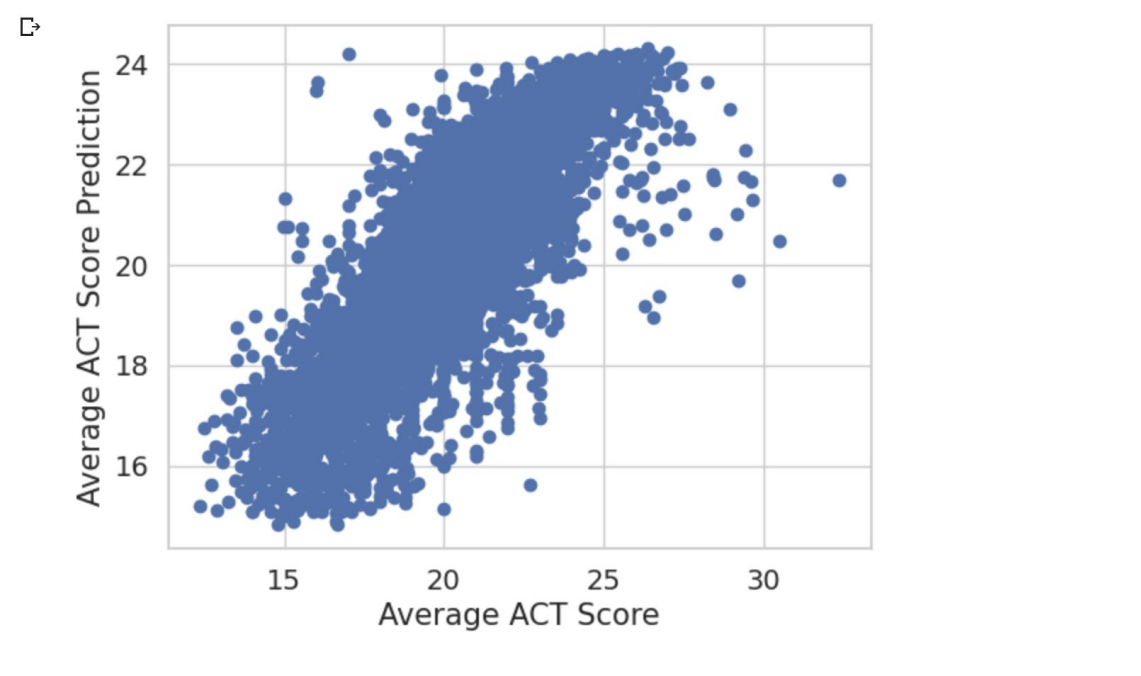


ACT Scores by State:

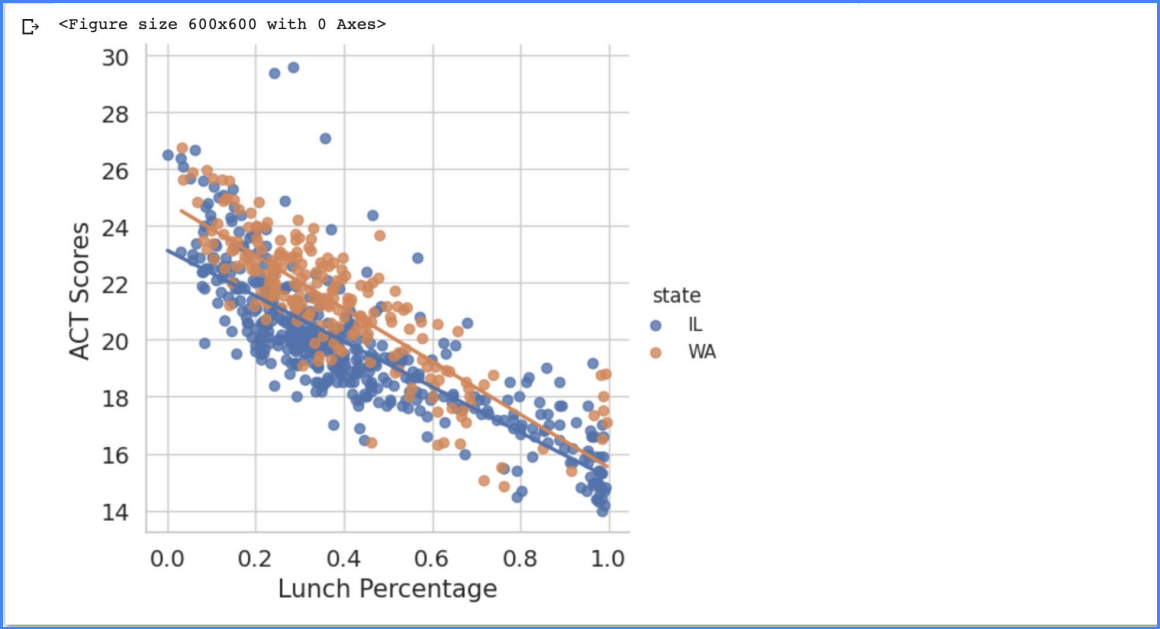


Students on Assisted Lunch by State





Predictive model designed to estimate ACT scores.



A scatter plot with a regression line for Per of Assisted Lunch and Average ACT Score. Washington vs Illinois.

A scatter plot with a regression line for Per of Assisted Lunch and Average ACT Score.

In conclusion, our project provides conclusive evidence of a robust relationship between ACT/SAT scores and socioeconomic factors. This is evident from the high accuracy (R2) of our predictions, the statistically significant coefficients associated with these factors, and the measure of accuracy expressed in ACT units. These findings highlight the crucial role of socioeconomic factors in shaping test scores and emphasize the significance of addressing educational inequality through targeted interventions.

In conducting this research project, it is important to acknowledge certain limitations that may have impacted our findings. Firstly, we must recognize that we may have missed important predictors, such as whether schools require ACT/SAT test-taking or not. Additionally, the availability of data varied across states, with some states having limited or missing overall data. Furthermore, it is worth noting that the SAT underwent changes in 2016, which could have influenced the results. Another limitation is the reliance on indirect measurements of certain conditions for students. Lastly, quality control issues, such as SAT/ACT conversions, may have introduced uncertainties into the analysis. Despite these limitations, our research provides valuable insights into the relationship between ACT/SAT scores and socioeconomic factors.