Blog 5

Group 1

## Blog5:

We are currently finalizing a comprehensive dataset that compiles district-level information spanning 18 years (2006–2023), with a particular focus on academic outcomes such as SAT performance, graduation rates, enrollment demographics, low-income percentages, and teacher-to-student ratios. To ensure consistency across this panel, we standardized district names and aligned year columns before merging graduation data with enrollment figures for each year. Additionally, SAT performance data was incorporated by matching district and year, with SAT scores appended only where districts already existed in the core panel. Empty cells were replaced with the placeholder “nil” to maintain clarity during downstream processing.

Our cleaned dataset, titled final\_merged, includes detailed insights into how various factors—such as income levels, school resources, and demographic shifts—may influence student achievement across districts. We also created a filtered version of this dataset by excluding charter schools to focus exclusively on traditional public schools. One key column, “Low Income %,” reflects the proportion of students from low-income households in each district, which we anticipate will be useful in identifying patterns between socioeconomic background and academic success. Moving forward, this dataset enables us to explore broader questions about the relationship between school funding, equity in education, and academic performance over time, positioning it as a strong foundation for longitudinal and policy-oriented analysis.