Data and Methodology Codebook

January 2025

Data Sources

Geographic unit boundaries: Geographic boundaries for census units (i.e., census tracts, counties, and state legislative districts) are obtained from the U.S. Census Bureau's TIGER/Line shapefile repository. School district boundary data come from the Education Demographic and Geographic Estimates Program (EDGE), Composite School District Boundaries Files.

School district revenues: Revenues from federal, state, and local sources come from the U.S. Census Bureau, Annual Survey of School System Finances, also known to researchers by the survey number F-33. Revenues include total and per-pupil revenues from federal, state, and local sources, including a disaggregated category from property taxes, plus specific totals for federal IDEA and Title I, III, and VII allocations. The following adjustments were made for some or all school districts:

- We exclude revenue for capital outlay and debt service programs from state revenues, because they can contribute to large fluctuations in district revenues from year to year. Similarly, we exclude money generated from the sale of property from local revenues.

- We subtract from state and local revenues the total amount of money sent to separate charter local education agencies (LEAs)—an expenditure category included in the F-33 survey—divided proportionally across the local, state, and federal revenue categories based on the percent of each district's revenues that come from these sources. This adjustment addresses the fact that, in just under 2,000 districts, revenues received by local school districts include funds that are transferred to charter schools that are operated by charter LEAs even though the students educated by these schools are not counted in enrollment totals. This artificially inflates the per-pupil revenues in these school districts.

- In the state of Arkansas, some revenues that are collected locally are categorized as state revenues. Before analysis, we subtracted the values of these collections from state revenues and added them to local revenues. The misattribution of revenue for each district is described in the F-33 documentation as C24, Census state, NCES local revenue. See the Annual Survey of School System Finances: Public Elementary-Secondary Education Finance Data Technical Documentation for state-specific notes in relation to education finance data.

School district enrollments and racial composition: School district enrollment characteristics come from the U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD). This includes data reporting graduation rates and the number of English language learners, students with disabilities, and students experiencing homelessness. In some cases, where counts are very small, data are suppressed. We recode these values as null. In conjunction with the above finance data, these enrollment statistics are used to calculate per-pupil revenue figures.

School district school-age poverty rates: School-district-level data on poverty rates among school-age children come from the U.S. Census Bureau's Small Area Income and Poverty Estimates (SAIPE).

Community indicators: Socioeconomic indicators (i.e., median household income, unemployment rates) and social indicators (i.e., educational attainment and information on internet accessibility), demographic indicators (population by racial group), select housing indicators (i.e., home value, rent indicators, and vacancy and owner occupancy rates, but not student homelessness rates or measures of housing insecurity) are derived from the U.S. Census Bureau's American Community Survey five-year estimates. School district-level data are provided by EDGE, drawing upon data from the U.S. Department of Education's National Center for Education Statistics, and U.S. Census Bureau.

Community and public health indicators: Census tract-level data recording health and health-linked outcomes for adult populations are recorded by the PLACES initiative of the Centers for Disease Control (CDC) beginning in 2019. The Social Determinants of Health included in this dataset are available only for 2022 and include information about access to health insurance and doctor's visits; incidence of chronic health issues, including obesity and asthma; access to reliable transportation; measures of food and housing insecurity; and utility shut-off notices.

Air quality: Environmental data recording the median air quality index at the census tract level is obtained from the [Environmental Protection Agency's air data repository](#).

Spatial Estimation

The fundamental challenge of making all the data available in a single dashboard is that the data sources are reported at different geographic levels. Our goal is to produce estimates for any given geography (e.g., a school district, a county, a state legislative district, or a census tract) using consistent methodologies. The key steps in this spatial estimation approach are as follows:

1. **Geographic standardization:** We begin by collecting and standardizing the spatial boundary files for the various geographies. These boundary files are obtained from authoritative federal sources and are projected into WGS84, a common coordinate system. To ensure accurate spatial calculations, we preserve the highest common level of precision found in the collected shapefiles.

2. **Overlay and intersection:** We overlay the source geography onto the target geography and compute the area intersection. This process identifies source-target pairs and their resulting polygon geometries. We remove "hallucinated" intersections smaller than 1.0 m2 that result from spatial imprecision.

3. **Calculation of overlap weights:** For each intersection between a source polygon (representing the geographic unit at which the requested data are reported) and a target polygon (representing the geographic unit for which we wish to estimate the values), we calculate an overlap fraction:

$$Overlap = \frac{Intersection\ Area}{Area\ of\ Source\ Polygon}$$

We assume a uniform spatial distribution of the variable of interest within the source polygon. This assumption is central to basic spatial interpolations and implies that the attribute is evenly distributed over the polygon's area. For variables not inherently uniform, this assumption has the potential to introduce uncertainty. In the absence of more granular data, it is a widely accepted methodological approach. However, users are advised to interpret results as estimates rather than exact values.

4. **Variable allocation:** After computing the overlap fraction, we allocate the source polygon's attribute value proportionally:

$$Allocated\ Value = Overlap * Source\ Polygon\ Value$$

5. **Reaggregation within the target geography:** Because each target geography may overlap multiple source polygons, we sum all allocated values from all intersecting source polygons:

$$Final\ Target\ Value = \sum_{All\ Source\ Intersections} Allocated\ Values$$

For count variables (e.g., number of students, total revenue), we directly sum these allocated counts to produce a total. For rates or proportions (e.g., poverty rates), we aggregate the underlying counts (e.g., number of children in poverty) and then rederive the value with the newly aggregated total counts (e.g., total school-age population).

Codebook

School District Revenues

Variable	Description
School Level	Whether a school district is intended to serve only elementary or secondary grades, or whether it is a unified district meant to serve all grade levels
Total Enrollment	Total number of enrolled students
Total Revenue Per Pupil ($)	Total education revenue from all federal, state, and local sources, divided by total number of enrolled students
Federal Revenue Per Pupil ($)	Education revenue provided through federal programs, divided by total enrollment
State Revenue Per Pupil ($)	Education revenue provided by state government, divided by total enrollment
Local Revenue Per Pupil ($)	Education revenue raised from local sources, including property taxes and other local income streams, divided by total enrollment
Local Property Tax Revenue ($)	Education revenue raised from property taxes specifically
Federal Title I Allocation ($)	Federal funding for the education of children from low-income families, allocated under Title I of the Elementary and Secondary Education Act
Federal Title III Allocation ($)	Federal funding for the education of English learners, allocated under Title III of the Elementary and Secondary Education Act
Federal Title VII Allocation ($)	Federal Impact Aid funding for districts that include federally owned land and/or enroll federally connected children, allocated under Title VII of the Elementary and Secondary Education Act
Federal IDEA Allocation ($)	Federal funding for the education of students with disabilities, allocated under the Individuals with Disabilities Education Act

School District School-Aged Poverty Rate

Variable	Description
Poverty Rate (%)	Percent of school-age children (those aged 5–17) residing in the geographic unit who live in households that fall below the federal poverty line

School District Enrollments and Racial Composition

Variable	Description
White (% Students)	White students as a percent of total enrolled students
Black (% Students)	Black students as a percent of total enrolled students
Latino (% Students)	Latino students as a percent of total enrolled students
Asian (% Students)	Asian students as a percent of total enrolled students
Native (% Students)	Native students as a percent of total enrolled students
Pacific Islander (% Students)	Pacific Islander students as a percent of total enrolled students
Other Race (% Students)	Students of other races as a percent of total enrolled students
Two or More Races (% Students)	Students of two or more races as a percent of total enrolled students
Students of Color (%)	All non-white students as a percent of total enrolled students
English Learners (%)	Percent of enrolled students designated as English language learners (ELLs)
Students with Disabilities (%)	Percent of students with an Individualized Education Program (IEP) or receiving disability services
4-Yr. High School Graduation Rate (%)	Percent of enrolled 9th grade students receiving high school diploma or equivalency within 4 years
Students Experiencing Homelessness (%)	Percent of enrolled students identified as experiencing homelessness

ACS Community Indicators

Variable	Description
Median Household Income ($)	Median income of all households in the geographic unit
Median Value, Owner-Occupied Homes ($)	Median market value of homes owned by occupants
Median Monthly Residential Rent ($)	Median total amount of gross rent paid before deducting expenses including utilities, maintenance, and property taxes
Median Rent as % of Income	Percent of median income attributed to gross rent as defined above
Total Population	Total number of people residing within a given geographic unit
White, including W. Hispanic (% Population)	White residents as a percent of total population
Black (% Population)	Black residents as a percent of total population
Native (% Population)	Native residents as a percent of total population
Asian (% Population)	Asian residents as a percent of total population
Pacific Islander (% Population)	Pacific Islander residents as a percent of total population
Other Race (% Population)	Residents of other races as a percent of total population
Two or More Races (% Population)	Residents of two or more races as a percent of total population
Housing Unit Vacancy Rate (%)	Percent of residential housing units that are vacant
Percent Homes, Owner-Occupied	Percent of occupied housing units that are owner-occupied
Mobile Internet Only (% Having)	Percent of households relying solely on mobile internet access
Broadband Internet (% Having)	Percent of households with broadband internet access
Other Forms Internet Access (% Having)	Percent of households with other forms of internet access besides mobile and broadband, including dial-up or satellite
High-School Diploma Attained (% Adults ≥25 Yrs)	Percent of adults aged 25 or older with at least a high school diploma or equivalent
College Degree Attained (% Adults ≥25 Yrs)	Percent of adults aged 25 or older with an associate or bachelor's degree
Unemployment Rate (%)	Percent of the civilian labor force that is unemployed
English as Primary Language (% Households)	Percent of households where English is the primary language spoken

CDC PLACES Community and Public Health Indicators

Variable	Description
Lacking Health Insurance (% Adults)	Percent of adults (aged 18+) without health insurance
Obesity Rate (% Adults)	Percent of adults (aged 18+) classified as obese
Asthma Rate (% Adults)	Percent of adults (aged 18+) with asthma
Doctor Checkup in Past 12 Months (% Adults)	Percent of adults who have had a medical checkup in the past year
SNAP Receipt in Past 12 Months (% Adults)	Percent of adults who received Supplemental Nutrition Assistance Program (SNAP) benefits in the past year (2022 data only)
Unreliable Transportation in Past 12 Months (% Adults)	Percent of adults reporting unreliable access to transportation (2022 data only)
Food Insecurity in Past 12 Months (% Adults)	Percent of adults reporting insufficient food access in the past year (2022 data only)
Housing Insecurity in Past 12 Months (% Adults)	Percent of adults reporting difficulty securing stable housing in the past year (2022 data only)
Utility Shut-Off Notice in Past 12 Months (% Adults)	Percent of adults receiving a utility shut-off notice in the past year (2022 data only)

EPA Air Quality Index

Variable	Description
Median Air Quality Index	Median value of daily U.S. Air Quality Index. This value aggregates values recording ground-level ozone, particulate matter, carbon monoxide, sulfur dioxide, and nitrogen dioxide. It runs on a scale of 0–500, with high values representing more hazardous air quality.