The Educational Opportunity Project
at Stanford University

Stanford Education Data Archive Technical Documentation

SEDA2022 2.0

April 2023

Erin M. Fahle, NWEA

Sean F. Reardon, Stanford University

Benjamin R. Shear, University of Colorado Boulder

Andrew D. Ho, Harvard University

Jim Saliba, Stanford University

Demetra Kalogrides, Stanford University

Suggested citation:

Fahle, E. M., Reardon, S. F., Shear, B. R., Ho, A. D, Saliba, J. & Kalogrides, D. (2023). Stanford Education Data Archive Technical Documentation: SEDA2022 2.0. Retrieved from http://purl.stanford.edu/db586ns4974.

Table of Contents

What is SEDA2022?

The Stanford Education Data Archive (SEDA) is created by the Educational Opportunity Project (EOP) at Stanford University (https:\\edopportunity.org). The EOP aims to generate and share data and research that can help scholars, policymakers, educators, and parents learn how to improve educational opportunities for all children. SEDA is the flagship data product of the EOP; it showcases how state accountability test data can be used to study educational opportunity in the U.S.

SEDA2022 is a special release of the Stanford Education Data Archive. It is part of a larger partnership with Harvard University's Center for Education Policy Research. This release is designed to provide insight into how school district average achievement in 2022, two years after the onset of the COVID-19 pandemic, compares to achievement in 2019, the year prior to the pandemic. The construction of SEDA2022 was supported by grants from the Bill and Melinda Gates Foundation and the Carnegie Corporation of New York. Some of the data used in constructing the SEDA files were provided by individual states and the National Center for Education Statistics (NCES). The findings and opinions expressed in our research and reported here are those of the authors alone; they do not represent the views of the U.S. Department of Education, NCES, individual states, or any of the aforementioned funding agencies.

Data Use Agreement

Prior to downloading the data, users must sign the data use agreement (https://edopportunity.org/get-the-data/).

Source Data

The data used to construct SEDA2022 test score estimates come from two primary sources: (1) the ED*Facts* data system; and (2) state-reported accountability data. We also use the National Assessment of Educational Progress (NAEP) data to link the state data to a common scale.

ED*Facts*. The ED*Facts* data system collects aggregated test score data from each state's standardized testing program as required by federal law. Specifically, each state is required to test every student in grades 3 through 8 in math and Reading Language Arts (RLA) each year. States have the flexibility to select or design a test that measures student achievement relative to the state's standards. Additionally, states set their own benchmarks or thresholds for the levels of performance, e.g., "proficient," in each grade and subject. States select 2 to 5 performance levels, where one or more levels represent "proficient" grade-level achievement. To ED*Facts*, states report the number of students in each school, subgroup, subject, grade, and year scoring at each of their defined performance levels; *no* *individual student-level data is reported*. ED*Facts* currently contains these school assessment outcomes for eleven consecutive school years from 2008-09 through 2018-19 in grades 3 to 8, and one grade in high school, in RLA and math.[1] The student subgroups include race/ethnicity, gender, and socioeconomic disadvantage, among others. The raw ED*Facts* data used in SEDA include no suppressed cells, nor do they have a minimum cell size for reporting. The data are reported by school, subject, grade, year, and subgroup and include schools in every state. For SEDA2022, we use the 2018-19 school-level proficiency data in grades 3-8 for all students, as well as the Black, Hispanic, White, economically disadvantaged (ECD), and not economically disadvantaged (not-ECD) subgroups. All state-subjects are represented in the 2018-19 data, except West Virginia RLA, which was removed by ED*Facts* due to data concerns (*State Assessments in Reading/Language Arts and Mathematics- School Year 2018-19* ED*Facts Data Documentation*, n.d.).

State-reported accountability data. Because ED*Facts* has not yet released 2021-22 school year data, we relied on state-reported data. Most states publicly report their school and/or district proficiency data as part of federal accountability. We collected this data in multiple ways. For most states, we use data scraped from state public websites. Two states provided data that was not released on their website to the EOP team for this purpose.

Notably, not all states reported usable data. To be used in our estimation process, the data must: (a) include *at least three* proficiency categories; and (2) be disaggregated by school or school district, subgroup, subject, grade, and year. The list of states that met these criteria can be found in **Table 1**.[2,3]

The usable data that states report also varied in content. There are three patterns of data reporting:

(1) Number of students scoring in each proficiency category. These are the necessary student-level data and can be used "as is."
(2) Total number of students tested and the percent scoring in each proficiency category. From this data, we can derive the counts in each category (with some rounding error).

---

[1] Data is also available for the 2020-21 school year; however, due to limited testing in that year, it is not comparable to other years.
[2] A number of states only reported two categories (proficient and not proficient): Alaska, Delaware, Hawaii, Iowa, and New Mexico. These were excluded from SEDA2022. Additionally, New York data was excluded from SEDA2022 due to low test participation rates in the 2018-19 school year and Colorado was excluded from SEDA2022 due to low test participation rates in the 2019-22 school year. While DC provided usable data, we were unable to produce reliable estimates and it was therefore excluded from SEDA2022.
[3] Most states with 3+ proficiency categories reported data in the required disaggregated format. For one state (Pennsylvania) we used 2022 data disaggregated by schools (not districts).

(3) Only the percent scoring in each proficiency category. We estimate the total number of test takers (in each school district, subgroup, subject, grade, and year) using the 2022 Common Core of Data (CCD) and the 2019 ED*Facts* data.[4] We can then derive an estimated count scoring in each category. In the downloadable datafiles, we flag all cells where we estimated counts.

Additionally, not all states reported data for all subgroups. States also used different suppression rules to protect student privacy. Some states suppress entire rows of data that do not meet reporting thresholds typically based on sample size, while others use partial suppression (suppressing some cells of data within a row). The extent and type of suppression affected our methodology for cutscore estimation, and sometimes prohibited our ability to produce estimates for individual districts. **Table 1** also includes information on the subgroups reported in the data, whether counts were estimated, and whether the state used partial suppression.

NAEP. Because different states use different tests and proficiency thresholds, the test score estimates derived from the above data sources are not readily comparable across states, grade, or years. Therefore, we also draw on the National Assessment of Educational Progress (NAEP) 2019 and 2022 national and state assessment data in 4th and 8th grade math and reading to link the estimates to a scale that is comparable among states and over grades and years.

Construction

Construction of the SEDA2022 test score estimates occurs in a series of steps. These steps are largely similar to those described in the SEDA 4.1 Technical Documentation. Here we provide an overview of the process and highlight where the SEDA2022 process deviates from the SEDA 4.1 process.

Step 1: Creating the 2019 School-to-District Crosswalk

This step links each public school to an administrative school district (e.g., NCES leaid) in 2019. A school-to-leaid crosswalk is not needed in 2022 as the proficiency data are already reported by administrative districts.[5,6]

---

[4] Specifically, we estimate the number of test takers in a district-subgroup-subject-grade as the number of enrolled students in the district-subgroup-subject-grade from the CCD 2022 multiplied by the participation rate for that district-subgroup-subject-grade from the 2019 ED*Facts* data. If the number of enrolled students is missing for a given district-subgroup-grade in 2021-2022, then imputed values are used.
[5] For the one state where we used 2022 school-level data (Pennsylvania), we use the NCES leaids reported in the CCD data to aggregate the schools to districts.
[6] While many states report data by school in 2022, there is too much suppression to reliably construct district estimates from the school level data. Because of this, it is possible that the set of schools assigned to a given district in 2019 is not identical to that in 2022. We acknowledge this limitation.

This crosswalk process deviates from that used for SEDA 4.1, which links schools to geographic school districts. Administrative districts differ from the geographic districts in two ways. First, for geographic school districts in SEDA 4.1, we "reassign" charter schools, magnet schools, and schools operated by secondary districts to the district in which they are physically located (regardless of the entity that operates the schools). Second, we exclude schools classified as "Special Education" from geographic districts and combine them into statewide special education districts. For SEDA2022 administrative districts, we do not reassign schools; charter, magnet, secondary, and special education schools are attached to the traditional public or charter district that operates them. For more information on geographic districts, we refer to you to the SEDA 4.1 Technical Documentation.

The use of administrative districts (rather than geographic districts) is preferred for SEDA2022 for two reasons. First, one of the aims of SEDA 2022 is to help school districts understand their learning recovery needs. Administrative districts have authority to set policy for their schools, as such it is most useful for the estimates to reflect only the schools under their operation. Second, to construct geographic school districts, we need data for individual charter schools. While many states report such data in 2022, data for many schools is suppressed due to the small numbers of students taking assessments. Because of this we cannot reliably construct geographic school district estimates for the 2022 school year.

Step 2: Data Cleaning

This step removes data not used in the estimation process and prepares data for estimation.

State-subject-grade-year removals. There are two primary reasons why a state-subject-grade was removed from the 2019 or 2022 data prior to estimation:

1. Test participation in a state-subject-grade-year was too low. We use a threshold of 94%[7] state-wide participation and remove any state-subject-grade-year where the 2019 participation rate falls below this threshold. Notably, the auxiliary participation rate data used to determine state-level removals is not available in 2022, so the 2019 state participation was used to remove state-subject-grade-year in both years.[8]
2. More than 5% of students took a test that is not the primary grade-level state assessment used for accountability. This occurs for two reasons. First, in some states students may take

[7] We use 94% (rather than 95% as done in SEDA 4.1) to preserve as much sample as possible. This affects a single state in our data, Oregon, where the participation rate in three subject-grades is between 94% and 95% in 2019. By including these cells with slightly lower participation, we can produce more estimates that meet our reliability thresholds.
[8] The exception to this is Colorado, which was removed entirely from SEDA2022 due to low state-reported test participation in 2022.

an end-of-course assessment. This is common in 8th grade math, when a subset of students takes the Algebra I test in place of the 8th grade math assessment. Second, a Spanish-language version of an assessment may be used in some subjects and grades. Spanish- and English-language versions of the assessments are not always equivalent, particularly when assessing reading and language skills. Because our estimation methodology relies on the fact that all students took a common test (within a state-subject-grade-year), we remove cases where fewer than 95% of students took the primary state accountability assessment for a given state-year-subject-grade. We identified these cases using two types of information:

a. State-reported information on the number of students taking each type of assessment. If more than 5% of students were reported to take the non-primary test in a subject-grade-year, it was removed.
b. State-level data on the number of students tested by grade-subject. If the reported number tested in a grade-subject cell was substantially lower than other grade subjects (suggesting a lower testing rate), it was removed.

A list of the states-subject-grades excluded from data construction are shown in **Table 2**.

Other removals. In addition to these state-level removals, we also remove all virtual schools in 2019. The 2022 data may or may not include virtual schools that are reported as part of administrative districts.

Data cleaning. For the 2019 data, we combine performance data for regular and alternate assessments. For 2022, we use data as reported, which may or may not include alternative assessments.

Step 3: Estimating and Linking Cutscores

In this step, we use Heteroskedastic Ordered Probit (HETOP) models or the inverse cumulative standard normal distribution function to estimate the state-grade-subject-year cutscores. For all state-subject-grades in 2019, we use the HETOP model to estimate the cutscores. In 2022, we use the HETOP model when the state-subject-grade data meets the following requirements: (1) counts are not estimated; and (2) the district-level data represents at least 95% of the students in the state, subject, and grade (e.g., there is limited suppression). When either of these requirements is not met, we calculate the cutscores from the state proficiency count data using the inverse cumulative standard normal distribution function.

We then link the estimated cutscores to the NAEP scale and standardize the NAEP-linked cutscores relative to a reference cohort of students. The resulting cutscores are in the Cohort

Standardized (CS) scale and are comparable across states and years under the linking assumptions reviewed in the 4.1 documentation.

Additional detail on the HETOP estimation and NAEP linking process can be found in the SEDA 4.1 Technical Documentation, Reardon, Shear, Castellano and Ho (2017), and Reardon, Kalogrides, and Ho (2019).

Step 4: Selecting and Preparing Data for Mean Estimation

This step selects data for district-subgroup-subject-grade-year cases that will be used in estimation. For 2019, the same rules are used in SEDA2022 and SEDA 4.1. These are as follows:

1. The participation rate is less than 95%. In these cases, the population of tested students on which the mean and standard deviation estimates are based may not be representative of the population of students in that school.
2. Incomplete data reported by student demographic subgroups. There are a small number of cases where the total number of test scores reported by race or gender is less than 95% of the total reported test scores for all students. We are concerned about subgroup data quality in these cases.
3. More than 40% of students take alternate assessments. Measurement error may affect district-subgroup-subject-grade-year cases where students take alternate assessments. These assessments typically differ from the regular assessment and generally have different proficiency thresholds or meanings. The threshold for exclusion is 40%.
4. Students scored only in the top or only in the bottom proficiency category. We cannot obtain maximum likelihood estimates of unique means for these cases and therefore remove them prior to estimation.
5. District-subgroup-subject-grade-year cells that do not meet the minimum statistical estimation requirements. When all cells for a district are insufficient (e.g., have all observations in a single middle category; have all observations in only 2 adjacent categories; have only 2 proficiency categories (one cut score); or have all observations in only the top and bottom categories) or small (have fewer than 100 test scores) we do not have sufficient information to produce mean and standard deviation estimates.

For 2022, the information to make exclusions based on participation, representation, and alternate assessments is not available. Thus, we only make removals based on 4 and 5, above.

Step 5: Estimating Means

This step uses the pooled HETOP model to estimate district-subgroup-subject-grade-year means and standard deviations, along with their standard errors, based on the cutscores from Step 3 and the data prepared in Step 4. Unlike SEDA 4.1, the SEDA2022 means are estimated separately by year and subject such that information is only pooled over grades. So, we estimate two pooled HETOP models in 2019 (one for math and one for RLA) and two pooled HETOP models in 2022. The resulting estimates are on the CS scale, and are comparable across states, within subjects and grades. For more details on the pooled HETOP model, see the SEDA 4.1 Technical Documentation and Shear and Reardon (2021).

Step 6: Creating Reporting Scales

This step creates the three scales reported in SEDA 2022: the Year Standardized (YS), Grade Year Standardized (GYS), and NAEP Point (NP) scales.

To create the YS scale, we standardize the estimates to the 2019 national average in each grade and subject. In this scale, each unit is equivalent to a 2019 national standard deviation in the same subject and grade.

To create the GYS scale, we first approximate the average amount student test scores grow in a grade on NAEP using the 4^{th} and 8^{th} grade estimates by subject in 2019. We calculate the amount the test scores changed between 4th and 8th grade as the average score in 8th grade in 2019 minus the average score in 4th grade in 2019. Then, to get an estimate of per-grade differences, we divide that value by 4. We scale the data using these parameters, such that in the GYS scale each unit is interpretable as 1 grade level.

To put estimates on the NAEP scale, we interpolate NAEP means in each grade and subject and use those to scale such that each unit is interpretable as 1 NAEP point.

Step 7: Pooling Mean Estimates

We use a different HLM pooling model in SEDA2022 than in SEDA 4.1. Separately for each subject, state and scale, we fit the model shown in Equation (1):

$$\begin{aligned}\widehat{mn}_{dgy}^{scale} = \beta_{0d} + \beta_{1d}\left(grade_{dgy} - 5.5\right) + \beta_{2d}\left(D2022_{dgy}\right) \\ + \beta_3\left(\left[grade_{dgy} - 5.5\right] \times D2022_{dgy}\right) + e_{dgy}\end{aligned}$$

$$\begin{aligned}\beta_{0d} &= \gamma_{00} + v_{0d} \\ \beta_{1d} &= \gamma_{10} + v_{1d} \\ \beta_{2d} &= \gamma_{20} + v_{2d}\end{aligned} \tag{1}$$

$$e_{dgy} \sim N\left(0, \widehat{\omega}_{dgy}^2\right);\ \epsilon_{dgy} \sim N(0, \sigma^2); \begin{bmatrix} v_{0d} \\ v_{1d} \\ v_{2d} \end{bmatrix} \sim MVN(0, \boldsymbol{\tau}^2)$$

$grade$ is the grade-level and $D2022$ is a dummy variable equal to one if the year is 2022. In this model, β_{0d} is the average score on the selected scale for district d in 2019, adjusting for linear trend across grades 3-8 in 2019. β_{2d} is the change in average score between 2019 and 2022 for district d, adjusting for linear trend across grades. This model is estimated using the HLM software (Raudenbush & Congdon, 2021), and produces both ordinary least squares ("OL") and shrunken Empirical Bayes ("EB") estimates of the changes. For more detail on OL and EB estimates, see the SEDA 4.1 Technical Documentation.

Step 8: Suppressing and Flagging Data for Release

Our goal is to ensure that the data we release is useful for various education stakeholders. We take caution to not report data that is unreliable and to flag estimates that require additional information for interpretation.

Data suppression. We suppress estimates for three reasons:

(1) The estimates that do not reflect at least 20 unique students in the 2019 school year. We do not suppress based on 2022 sample size because all 2022 source data was public.

(2) The estimates are too imprecise to be useful. The precision threshold we use is 0.33 grade levels. For each district and subgroup, we keep both math and reading estimates when one or both have a change estimate (β_{2d}) with an OL standard error less than 0.33 grade levels.

(3) The district does not have a corresponding geographic boundary in the 2019 school district shape files. These tend to be specialized administrative districts, like charter school and virtual school districts, and tend to not serve stable populations over time.

Data flags. We flag districts in the data where there were large changes in the overall enrollment and/or in the racial composition using grade 3-8 CCD enrollment data by district-subgroup from fall 2019 and 2022. We identified districts as having large changes in overall enrollment if the proportional change in grade 3-8 CCD enrollment was greater than .20 (20%), calculated as:

$$\frac{|g38enroll_{2022} - g38enroll_{2019}|}{\min(g38enroll_{2022}, g38enroll_{2019})} > .20$$

We then identified districts as having large changes in racial composition if the percentage of students of any given racial group changed by more than 5 percentage points, calculated as:

$$\max(|\%asn_{2022} - \%asn_{2019}|, |\%blk_{2022} - \%blk_{2019}|, |\%hsp_{2022} - \%hsp_{2019}|, |\%hpi_{2022} - \%hpi_{2019}|, |\%mtr_{2022} - \%mtr_{2019}|, |\%nam_{2022} - \%nam_{2019}|, |\%wht_{2022} - \%wht_{2019}|) > .05$$

The final flag shown on the website and reported in the data is a combination of these two, such that we flag any places with large changes in the overall enrollment and/or in the racial composition per the above measures. We opted to flag these places to indicate that changes in achievement should be interpreted with caution given the shift in student population.

Data Quality and Validation Checks

Inferences regarding changes in the estimated test scores from 2019 to 2022 hinge on the comparability of the data over time. There are a few potential threats to these inferences:

Data discrepancies. Ostensibly, because state reported proficiency data are the source data for ED*Facts*, the data reported in the two sources should be equivalent (save for differences in data suppression). However, as part of the ED*Facts* data collection process, state proficiency data are vetted and cleaned. As such, it is possible that the 2022 state data are of lower or different quality than the 2019 ED*Facts* data. Additionally, there may be differences in how the data are reported, for example, whether alternate assessments are reported in the count data for each district.

To better understand any potential discrepancies in the data, we cleaned the 2019 state reported data using the same rules (described above) as the 2022 state reported data. We then compared the count data and estimates produced from both sources. Differences in estimated means are associated with differences in the underlying count data at the state and/or district levels (i.e., the two sources reported different proportions of students scoring at the same proficiency level). On average, these differences in the estimated means were small (.07 grade levels). Only in a small percentage of cases (6.6%) did estimates differ by more than .2 grade levels. More than three-quarters of cases with differences larger than .2 grade levels are in district-subject-grades with fewer than 100 test-takers; only 3% are in cases with more than 500. Thus, while there are differences in the data, they would impact the estimates we report in limited ways for relatively few test-takers.

Notably, as a result of this analysis, we determined that the data reported in ED*Facts* for Arkansas RLA is not comparable to the data reported by the state in either Reading or English. As such RLA data for Arkansas was removed from website and data files.

NAEP Linking. We replicated a subset of the analyses from Reardon, Kalogrides, and Ho (2021) to demonstrate that the NAEP linking continues to perform well.

Comparison to percent proficient data. As a face validity check of the data, for each state-subject we correlated our estimated changes in means (pooled over grades) with changes in the probit transformed reported percent proficient (pooled over grades). Correlations ranged from .56 to .97 in math and .65 to .98 in RLA. These suggest that our estimates tell a story that is largely consistent with that from the reported percent proficient.

Version Changes

There were several improvements made to the data cleaning and estimation processes between SEDA 2022 v1.0 (called "beta") and the current version, SEDA2022 v2.0. A summary of the major changes from SEDA2022 1.0 to SEDA2022 2.0 are listed below. We:

- Added 11 new states and removed one state (Colorado); Version 1.0 had 29 states, Version 2.0 has 40 states.
- Refined the estimation of tested counts from CCD when not reported by the state.
- Changed the state-subject-grade-year inclusion rules.
- Changed when the two cutscore estimation methods (HETOP vs. inverse normal) were used for 2022 data.
- Made other small improvements to estimation.

Definitions

Administrative school district: Administrative school districts operate sets of public and charter schools. The schools operated by each school district are identified using the National Center for Education Statistics (NCES) school and district identifiers. Most commonly, administrative school districts operate local public schools within a given physical boundary; these are what we refer to as "traditional public school districts." There are also specialized administrative districts that do not have a physical boundary, like charter school and virtual school districts.

Subgroup: Subgroups are defined per state accountability reporting requirements; our data include the following subgroups: all, Black, Hispanic, White, poor (economically disadvantaged), and non-poor (not disadvantaged).

Frequently Asked Questions

Why aren't results available for my state or district?

There are several reasons why we may not show data for a particular district.

1. We cannot construct an estimate for the district in either 2019 or 2022:

    a. Sufficient data for estimation were not reported by the state or district in either 2019 or 2022.
    b. The district changed its leaid between 2019 and 2022. This may occur as a result of a district merger, split, or takeover that occurred during this three-year span.
    c. Fewer than 95% of students in the state or district participated in testing in the subject in 2019.
    d. More than 40% of students in the district took alternative assessments rather than the regular tests in 2019.
2. The data is suppressed:
    a. The district is too small and/or has too few grades of data available to allow for the construction of reliable estimates.
    b. The district does not have a geographic boundary; such districts include charter districts and/or specialized local education agencies.

Why did the results for my district change between releases?

A number of improvements were made in our data inclusions/exclusion rules, as well as in our estimation methods that enabled us to get better estimates of the average achievement in 2019 and 2022, and the change in achievement been those two years. The improvements are detailed above in the section on *Version Changes*.

How have you accounted for population changes from 2019 to 2022 in the estimates?

We have not made any adjustment to the estimates to account for changes in the population between 2019 and 2022. On the website and in the downloadable files, we provide a flag that indicates if there was a large change in the total enrollment and/or racial composition of the district during the time period. The calculation of this flag is described above under the section *Step 8: Suppressing and Flagging Data for Release*.

References

Raudenbush, S.W., & Congdon, R.T. (2021). *HLM 8: Hierarchical linear and nonlinear modeling*. Chapel Hill, NC: Scientific Software International, Inc.

Reardon, S. F., Shear, B. R., Castellano, K. E., & Ho, A. D. (2017). Using heteroskedastic ordered probit models to recover moments of continuous test score distributions from coarsened data. *Journal of Educational and Behavioral Statistics*, 42(1), 3–45. https://doi.org/10.3102/1076998616666279

Reardon, S. F., Kalogrides, D., & Ho, A. D. (2019). Validation Methods for Aggregate-Level Test Scale Linking: A Case Study Mapping School District Test Score Distributions to a Common Scale. *Journal of Educational and Behavioral Statistics*, 107699861987408. http://doi.org/10.3102/1076998619874089

Shear, B. R., & Reardon, S. F. (2021). Using Pooled Heteroskedastic Ordered Probit Models to Improve Small-Sample Estimates of Latent Test Score Distributions. *Journal of Educational and Behavioral Statistics*, *46*(1), 3–33. https://doi.org/10.3102/1076998620922919

*State Assessments in Reading/Language Arts and Mathematics- School Year 2018-19 ED*Facts *Data Documentation*, U.S. Department of Education, Washington, DC: EDFacts. Retrieved 04/24/2023 from http://www.ed.gov/edfacts.

Tables & Figures

Table 1. *2022 source data overview*

State	Subgroups reported	Estimated counts	Partial suppression
AL	All, Black, ECD, Hispanic, White students	Y	N
AR	All students	N	N
AZ	All, Black, Hispanic, White students	N	N
CA	All, Black, ECD, Hispanic, Not-ECD, White students	N	N
CT	All, Black, ECD, Hispanic, Not-ECD, White students	N	Y
FL	All students	N	N
GA	All students	N	N
ID	All, Black, ECD, Hispanic, Not-ECD, White students	Y	Y
IL	All, Black, ECD, Hispanic, Not-ECD, White students	Y	N
IN	All students	N	N
KS	All, Black, ECD, Hispanic, White students	Y	N
KY	All, Black, ECD, Hispanic, Not-ECD, White students	Y	N
LA	All, Black, ECD, Hispanic, Not-ECD, White students	Y	N
MA	All, Black, ECD, Hispanic, Not-ECD, White students	N	N
MD	All, Black, ECD, Hispanic, Not-ECD, White students	N	Y
MI	All, Black, ECD, Hispanic, Not-ECD, White students	N	Y
MN	All, Black, ECD, Hispanic, Not-ECD, White students	N	N
MO	All, Black, ECD, Hispanic, Not-ECD, White students	N	Y
MS	All students	N	N
NC	All, Black, ECD, Hispanic, Not-ECD, White students	N	Y
ND	All, ECD, Not-ECD	N	N
NE	All, Black, ECD, Hispanic, White students	Y	N
NH	All, Hispanic, White	Y	Y
NJ	All, Black, ECD, Hispanic, Not-ECD, White students	In some cells	N
NV	All students	N	Y
OH	All, Black, ECD, Hispanic, Not-ECD, White students	N	N
OK	All students	N	Y
OR	All, Black, Hispanic, White students	N	N
PA	All students	N	N
RI	All students	N	N
SC	All, Black, ECD, Hispanic, Not-ECD, White students	N	N
SD	All, Black, ECD, Hispanic, Not-ECD, White students	N	N
TN	All, Black, ECD, Hispanic, Not-ECD, White students	N	Y

TX	All, Black, ECD, Hispanic, Not-ECD, White students	N	N
UT	All students	Y	Y
VA	All students	Y	N
WA	All, Black, ECD, Hispanic, Not-ECD, White students	In some cells	N
WI	All, Black, ECD, Hispanic, Not-ECD, White students	N	N
WV	All, Black, White students	Y	N
WY	All, ECD, Hispanic, Not-ECD, White students	Y	N

Table 2. *Data cleaning, state-subject-grade removals*

State	Subject	Grade	Reason removed in 2019 (if blank, not removed)	Reason removed in 2022 (if blank, not removed)
AZ	Math	8	Primary Testing Rate < 95%	
FL	Math	7	Primary Testing Rate < 95%	Primary Testing Rate < 95%
FL	Math	8	Primary Testing Rate < 95%	Primary Testing Rate < 95%
OH	Math	7	Primary Testing Rate < 95%	Primary Testing Rate < 95%
OH	Math	8	Primary Testing Rate < 95%	Primary Testing Rate < 95%
TN	Math	8	Primary Testing Rate < 95%	Primary Testing Rate < 95%
TX	Math	7	Primary Testing Rate < 95%	Primary Testing Rate < 95%
TX	Math	8	Primary Testing Rate < 95%	Primary Testing Rate < 95%
TX	RLA	3	Primary Testing Rate < 95%	Primary Testing Rate < 95%
TX	RLA	4	Primary Testing Rate < 95%	Primary Testing Rate < 95%
VA	Math	7	Primary Testing Rate < 95%	Primary Testing Rate < 95%
VA	Math	8	Primary Testing Rate < 95%	Primary Testing Rate < 95%
NJ	Math	8	Primary Testing Rate < 95%	Primary Testing Rate < 95%
MO	Math	8	Primary Testing Rate < 95%	Primary Testing Rate < 95%
MD	Math	7	Primary Testing Rate < 95%	Primary Testing Rate < 95%
MD	Math	8	Primary Testing Rate < 95%	Primary Testing Rate < 95%