

Zelma Education:
State Assessment Data
Documentation



June 2024

Version 1.1

Zelma Education
Brown University

Emily Oster, Zelma Principal Investigator

Clare Halloran, Zelma Project Director

Suggested citation:

Zelma Education: State Assessment Data Documentation, Version 1.1. (2024). Brown University.
Retrieved [date] from zelma.ai.

Please send questions or comments about this technical guide to zelmadata@gmail.com. We value your feedback.

Contents

Overview	1
2.1 Document Publication.....	1
2.2 Data Publication & Citations.....	1
2.3 Data Sources	2
2.4 Assessment Inclusion Criteria	2
2.5 Changes in Assessments.....	2
2.6 School Years Included	3
2.7 Data Suppression & Missing Data.....	3
2.8 The Incomparability of State Assessment Data.....	3
2.9 Updates to Data & Documentation.....	4
File Format	5
2.0 State-Year-Data Level Variables.....	5
2.1 District and School Identifiers	6
2.2 Assessment Variables.....	8
2.3 Student Variables.....	10
2.4 Proficiency and Participation.....	14
2.5 Indicators of Assessment Changes ("Flags").....	18
2.6 NCES Descriptive Variables	18
Appendix A. Technical Guide Version History	23
Appendix B. Data Sources.....	24
Appendix C. State-Subject-Year Data Availability	29
Appendix D. New York City NCES District IDs	34
Appendix E. Proficiency Criteria by State-Year-Subject.....	35
Appendix F. Zelma Frequently Asked Questions (FAQs).....	38

Overview

Zelma Education is a comprehensive, interactive, AI-powered U.S. state assessment data repository that aims to make state assessment data more widely accessible and engaging for the general public. The repository includes publicly-available assessment data from all 50 states and D.C. for students in Grades 3-8. This technical guide provides an overview of the project's data sources and inclusion criteria, as well as documentation about the file format, variables, and data decisions to produce the Zelma data files.

2.1 Document Publication

This document corresponds to Zelma Version 1.1. See *Appendix A. Technical Guide Version History* for the Version History of this document.

2.2 Data Publication & Citations

Zelma Education published Version 1.1 on June 18, 2024. SEA data released or corrected after this date will be incorporated into subsequent versions. Data files used to produce information for the general public such as articles, reports, or presentations, should cite the data source as follows:

Citing Zelma data files

Raw Data Citation Format:

- Zelma Education (Version 1.1). 2024. *Zelma assessment data files*, [state name (years)]; [state name (years)]; [state name (years)]. Accessed at <https://www.zelma.ai/data> on Month DD, YYYY.

Raw Data Citation Examples:

- Zelma Education (Version 1.1). 2024. *Zelma assessment data files*, Minnesota (2018 - 2023); Wisconsin (2018 - 2023). Accessed at <https://www.zelma.ai/data> on June 18, 2024.
- Zelma Education (Version 1.1). 2024. *Zelma assessment data files*, All States (2018 - 2023). Accessed at <https://www.zelma.ai/data> on June 18, 2024.

Citing queries

Query Citation Format:

- Zelma Education (Version #). [year]. "[query title]." Accessed at [query link] on Month DD, YYYY.

Query Citation Example:

- Zelma Education (Version 1.1). 2024. "Math scores by race in New York over time." Accessed at <https://www.zelma.ai/q/JTS8BN29Y8> on June 18, 2024.

2.3 Data Sources

State assessment data are sourced from state education agencies (SEAs) and are combined with supplementary school and district information from U.S. Department of Education (ED) resources, including the National Center for Education Statistics (NCES) and *EDFacts*, as described below.

- a. **State Education Agencies (SEAs).** All proficiency data are sourced directly from State Education Agencies, from either their websites/data portals or via data request (no personal identifiable information is included). In some cases, SEA enrollment data are used as a proxy for student tested counts (see Table 16). A complete list of SEA sources by state can be found in *Appendix B. Data Sources*.
- b. **National Center for Education Statistics (NCES).** Zelma integrates information about district and school characteristics from the Common Core of Data, a U.S. public school database available from the National Center for Education Statistics (NCES).
- c. **EDFacts.** The U.S. Department of Education (ED) annually collects state proficiency data in ELA and math from SEAs through an initiative called *EDFacts*. In some cases, *EDFacts* counts of students tested are used as a proxy for student tested counts that are not available in SEA data files (see Table 16).

2.4 Assessment Inclusion Criteria

Zelma includes all available achievement data from annual statewide summative assessments that are used to report outcomes to the U.S. Department of Education under Title I, Part A of the Elementary and Secondary Education Act (ESEA). Zelma does not include other assessments administered by states, such as alternate assessments or English language proficiency assessments. Currently, Zelma only contains assessments administered in English.

The ESEA requires that state assessments be administered to students annually in Grades 3-8 and at least once in high school. Zelma currently limits data to Grades 3-8.

Several states include schools that are supported by the Bureau of Indian Education (BIE). Zelma data files do not currently include schools represented by BIE.

Assessment data availability varied by state, school year, and subject. See *Appendix C. State-Subject-Year Data Availability* for a complete list of subject availability by state, including the first and last year of available data.

2.5 Changes in Assessments

Changes in state assessments are noted in two ways for data users:

- a. in the data files, there are variables that indicate (or “flag”) if the subject-area assessment has changed either its name or how the state defines proficiency from the prior year (for ELA, math, science, and social studies), and
- b. below all figures produced by Zelma, there are details included as part of the “Notable Events” that explain if states have administered a new assessment in a given year, if there have been changes in proficiency cut scores, if an assessment has not been administered due to field testing, if participation rates were lower than a typical year, or if assessment data are not currently available.

2.6 School Years Included

Longitudinal data varies by state, with the first year of available data ranging from 1998 to 2018. A few notes:

- **No assessments in Spring 2020:** No states administered assessments in Spring 2020 due to the COVID-19 pandemic; therefore, no data are included as part of Zelma for this school year.
- **Missing years:** There may be periodic years of missing data for all subjects within a state or for a particular subject area; often this is due to field testing a new assessment. Some states received assessment waivers from ED in Spring 2021 due to the pandemic.

The full range of available years and subjects for each state is included in *Appendix C. State-Subject-Year Data Availability*.

2.7 Data Suppression & Missing Data

According to the Family Educational Rights and Privacy Act (FERPA), states are required to ensure that a student's individual identity cannot be determined when included as part of datasets with educational records. Therefore, states implement data rules that "suppress" or "mask" certain data fields that would otherwise reveal the outcomes for small groups of students; often, states suppress data when a cell size is less than 10 students, though this varies. Zelma did not request any personally identifiable information (PII) from states. Suppressed data are represented with an asterisk (*) in the Zelma data files. Information that was not provided in the original data files is represented with dashes (--) in the Zelma data files (see Table 1).

Table 1. Data File Symbols

Symbol	Description
*	Data suppressed
--	Data missing

Ranges: Some states suppress data by including a range of either students tested or the proficiency outcomes. Zelma data files retain these ranges. State that use "<" or ">" are converted to decimal ranges (e.g., a participation rate of ">.95" is represented as ".95-1").

2.8 The Incomparability of State Assessment Data

All states define proficiency according to their state's unique grade-level and subject-area learning standards. For this reason, results are *NOT comparable across states*, as states administer distinct assessments. These tests are often designed to assess student progress on state-specific standards and are reported based on state-specific definitions of proficiency.

2.9 Updates to Data & Documentation

- ★ *Changes from the prior data release will be highlighted for the user throughout this technical guide with this symbol.*

This technical guide documents the approach that Zelma Education has used to standardize the state assessment data files produced by SEAs, including data decisions, important information about each variable, and data sources.

UPDATED for Version 1.1:

- **File format.** All files include updated flag variables marking changes in assessments compared to the prior year (ELA and math flags are updated; science and social studies flags are new), and the removal of some variables (seasch, State_lead, Flag_CutScoreChange_read, Flag_CutScoreChange_oth).
- **Student counts.** All files now include student counts through Spring 2023, from raw state data files as available, or *EDFacts* if not. Level counts are derived across all files if tested counts and achievement level percents are available.

File Format

All Zelma files include an identical file format, including identical variables. However, because not all states report all data components, some variables will be *empty* for some states. In this section, we present variables included in Zelma Version 1.1.

2.0 State-Year-Data Level Variables

Table 2. State-Year-Data Level Values

Variable Name		Description
a	State	State Name
b	StateAbbrev	Two-letter state abbreviation
c	SchYear	School year for which the data were reported (e.g., 2021-22)
d	DataLevel	Level at which the data are reported
e	StateFips	The two-digit American National Standards Institute (ANSI) code for the state

a. State

Value labels reflect state names. All U.S. states are included as part of Zelma, as well as the District of Columbia.

b. StateAbbrev

Value labels reflect all two-letter state abbreviations for U.S. state and the District of Columbia (see Table 3).

c. SchYear

Value labels reflect the school year for the year the state administered the assessment; typically, these assessments are completed in the spring of each school year. For example, assessments completed in Spring 2023 are represented as part of the “2022-23” school year.

d. DataLevel

Value labels include **State**, **District**, and **School**. Data are included at each level to the extent that they were available from the SEA; that is, not all files will have all data levels available for all school years and subjects. Districts represent local education agencies (LEAs) as defined by NCES.

e. StateFips

Value labels reflect the two-digit American National Standards Institute (ANSI) code for state (see Table 3).

Table 3. State Abbreviations and FIPS Codes

State	StateAbbrev	FIPS	State	StateAbbrev	FIPS
Alabama	AL	01	Montana	MT	30
Alaska	AK	02	Nebraska	NE	31
Arizona	AZ	04	Nevada	NV	32
Arkansas	AR	05	New Hampshire	NH	33
California	CA	06	New Jersey	NJ	34
Colorado	CO	08	New Mexico	NM	35
Connecticut	CT	09	New York	NY	36
Delaware	DE	10	North Carolina	NC	37
District of Columbia	DC	11	North Dakota	ND	38
Florida	FL	12	Ohio	OH	39
Georgia	GA	13	Oklahoma	OK	40
Hawaii	HI	15	Oregon	OR	41
Idaho	ID	16	Pennsylvania	PA	42
Illinois	IL	17	Rhode Island	RI	44
Indiana	IN	18	South Carolina	SC	45
Iowa	IA	19	South Dakota	SD	46
Kansas	KS	20	Tennessee	TN	47
Kentucky	KY	21	Texas	TX	48
Louisiana	LA	22	Utah	UT	49
Maine	ME	23	Vermont	VT	50
Maryland	MD	24	Virginia	VA	51
Massachusetts	MA	25	Washington	WA	53
Michigan	MI	26	West Virginia	WV	54
Minnesota	MN	27	Wisconsin	WI	55
Mississippi	MS	28	Wyoming	WY	56
Missouri	MO	29			

2.1 District and School Identifiers

All Zelma data files include important information on district and school names and identifiers (see Table 4), as discussed below.

Table 4. District and School Identifiers

Variable Name	Description
f DistName	District name
g SchName	School name
h NCESDistrictID	7-digit NCES district ID
i StateAssignedDistrictID	District ID as used by the state
j NCESSchoolID	12-digit NCES district ID
k StateAssignedSchoolID	School ID used by the state

★ **V1.1 Change:** Version 1.1 removes variables State_leadid and seasch.

f. DistName

District names reflect the name as spelled and reported publicly in SEA assessment data files, which may include variations over time (see Table 5).

Table 5. DistName Values

Data Level	DistName
State	DistName = "All Districts"
District	DistName will reflect the name as publicly reported in state assessment data files from the SEA. As such, there may be differences from year to year in how the state reports the district name (e.g., "Central Public Schools" one year may be reported as "Central PS" in another year). Note that district IDs would remain constant if the district did not change.
School	DistName will reflect the district that the given school is part of for the given school year, as defined by the National Center for Education Statistics.

g. SchName

School names reflect the name as spelled and reported publicly in SEA assessment data files, which may include variations over time (see Table 6). Not all states report school-level data; data at the school level typically have more data suppression than district- or state-level files.

Table 6. SchName Values

Data Level	SchName
State	SchName = "All Schools"
District	SchName = "All Schools"
School	SchName will reflect the name as publicly reported in state assessment data files from the SEA. As such, there may be differences from year to year in how the state reports the school name (e.g., "Franklin Elem." one year may be reported as "Franklin Elementary" in another year). Note that school IDs would remain constant if the school did not change.

h. NCESDistrictID

NCES assigns all U.S. local education agencies (LEAs) with a 7-digit identification number. These IDs are typically stable from year to year. All districts and schools in the assessment data files are matched to their NCES LEA ID to support the ability to merge data with other datasets. In some cases, newly-opened districts did not yet have publicly-available NCES IDs at the time of data release and are reported in the Zelma assessment files as "Missing/not reported." These IDs will be updated as available.

New York City Public Schools

In the case of New York City schools, there are multiple smaller local education agencies (LEAs) with unique IDs as part of the Common Core of Data. The Zelma data output includes these unique IDs. However, for visualization purposes, the data for these smaller LEAs are aggregated to represent the outcomes for NYC schools. See *Appendix D. New York City NCES District IDs* for a complete list of IDs as reported by NCES.

i. StateAssignedDistrictID

Values reflect the district identification numbers used by SEAs as part of their own data management systems and data files.

j. NCESSchoolID

NCES assigns U.S. schools a 12-digit identification number. The first 7 digits are consistent with the NCES district ID, and the subsequent 5 digits are unique to the school. All schools in the assessment data files are matched to their NCES school ID to support the ability to merge data with other datasets. In some cases, newly-opened schools did not yet have publicly-available NCES IDs at the time of data release and are reported in the assessment files as “Missing/not reported.” These IDs will be updated as available.

k. StateAssignedSchoolID

Values reflect the school identification numbers used by SEAs. These may or may not be unique for all schools in the file. For example, there may be distinct schools with a StateAssignedSchoolID of “1” that belong to different districts.

2.2 Assessment Variables

All Zelma data files include a set of variables to properly and easily identify student outcomes, including the assessment name, the assessment type, the assessed subject area, the assessed grade level, the student group category of analysis (e.g., Race/Ethnicity) and number of tested students in the respective category, and the student subgroup of analysis (e.g., Asian) and the number of tested students in the respective subgroup (see Table 7), as discussed below.

Table 7. Assessment Variables

Variable Name		Description
l	AssmtName	Name of state assessment
m	AssmtType	Assessment type
n	Subject	Assessment subject area
o	GradeLevel	Grade tested
p	StudentGroup	Student group category of analysis
	StudentSubGroup	Student subgroup within the larger StudentGroup category
q	StudentGroup_TotalTested	Number of students tested in the designated StudentGroup
	StudentSubGroup_TotalTested	Number of students tested in the designated StudentSubGroup

l. AssmtName

Value labels reflect the name of the assessment for the given state-year-subject corresponding to the student outcome data. A change in the assessment name from the prior year is indicated by the variable Flag_AssmtNameChange, as explained below. In most but not all cases, a change in the assessment name from the prior year indicates a change in the SEA’s definition of proficiency.

m. AssmtType

Value labels reflect the type of assessment for the given state-year-subject corresponding to the student outcome data. While the Zelma repository aims to include data for each state’s general state

standardized assessments, some states *aggregate* the regular and alternate assessment data. Alternate assessments are annual state assessments based on alternate achievement standards for students with significant cognitive disabilities. To reflect the difference in state assessments included in the dataset, this variable is included for additional clarity for the user (see Table 8).

Table 8. Assessment Types

AssmtType	Description
Reg	Regular/general assessment data only
Reg and alt	Aggregated data from the regular/general assessment AND the state's alternate assessment

n. Subject

All states include data for English language arts (ELA) and math. To the extent that additional subjects are publicly available, Zelma also integrates data for science, reading (if different from ELA), writing, social studies, and STEM (see Table 9). However, not all states have a required assessment for subject areas other than ELA and math. Subjects are included as reported by SEAs. States may assess in additional subject areas that are not included here (e.g., Spanish language arts).

Table 9. Subjects

Subject	Description
ela	English language arts
math	Mathematics
sci	Science
soc	Social studies
wri	Writing
read	Reading (only available for Arkansas, 2015-16 to 2022-23; and Georgia, 2010-11 to 2013-14)
eng	English (only available for Arkansas, 2015-16 to 2022-23)
stem	Science, technology, engineering and math (only available for Arkansas, 2015-16 to 2022-23)

Some states use the subject “ELA” to report student outcomes related to reading and writing, while other states use the term “reading” or “reading language arts.” For standardization purposes, all states have been recorded as “ela” in the data files *unless the state included both ELA and reading outcomes* in the same year, in which case with subject value labels were retained. This was the case for two states (see Table 10).

Table 10. States with both ELA and Reading Outcome Data

State	Year	Description
Arkansas	2015-16 to 2022-23	State includes ELA outcome data separately from reading outcome data. ELA outcome data is the average of three subjects that are included in the file: English, reading, and writing. Arkansas stopped reporting outcome data for <i>writing</i> after the 2016-17SY.
Georgia	2010-11 to 2013-14	State includes ELA outcome data separately from reading outcome data.

o. GradeLevel

Grade level value labels are limited to Grades 3-8, as well as an aggregated measure (“G38”), if available (see Table 11). *Not that for visualization purposes, zelma.ai relies on data disaggregated by grade level.*

Table 11. Grade Levels

GradeLevel	GradeLevel Value Labels
G38	Aggregated data for all of Grades 3-8
G03	Grade 3
G04	Grade 4
G05	Grade 5
G06	Grade 6
G07	Grade 7
G08	Grade 8
GZ [★]	Aggregated data for more than Grades 3-8. Disaggregated grade level data not available.

[★] **V1.1 Change:** Not all states disaggregate student outcomes by grade level. This is the case for Maine and New Mexico (see Table 12). Beginning in V1.1, these cases are indicated with GradeLevel value “GZ”.

Table 12. States without Disaggregated Grade-Level Data

State	Year	Description
Maine	2014-15 to 2022-23	State does not have disaggregated data available. GradeLevel values are presented as “GZ” as they include values outside of Grades 3-8.
New Mexico	2016-17	State does not have student <i>subgroup</i> data disaggregated by grade level; these values are presented as “GZ” as they include values outside of Grades 3-8. State does have data for “All Students” disaggregated by grade level.
	2017-18	
	2018-19	
	2020-21	State has student subgroup data available aggregated for Grades 3-8. GradeLevel values are presented as “G38” only. Data disaggregated by grade level are not available.
	2021-22	State does not have any student data disaggregated by grade level; these values are presented as “GZ” as they include values outside of Grades 3-8.
	2022-23	First year available with data disaggregated by grade and subgroup.

2.3 Student Variables

p. StudentGroup and StudentSubGroup

Value labels for **StudentGroup** reflect the broader student classification categories of which the **StudentSubGroup** classifications are a part (see Table 13). Each state uses its own criteria to define student subgroups. For this reason, there may be differences from one state to another related to how students are classified; they are not necessarily comparable. Zelma reflects each state’s own data classifications for student subgroups.

★ **V1.1 Change:** Ethnicity was removed as a StudentGroup and combined with RaceEth.

Table 13: StudentGroup and StudentSubGroup Value Labels

StudentGroup	StudentSubGroup Value Labels
All Students	<ul style="list-style-type: none"> All students
RaceEth	<ul style="list-style-type: none"> American Indian or Alaska Native Asian Black or African American Native Hawaiian or Pacific Islander Two or more White Hispanic or Latino Not Hispanic or Latino Filipino ★ Unknown
EL Status	<ul style="list-style-type: none"> English Learner English Proficient EL Exited ★ EL Monitored or Recently Exited ★ EL and Monitored or Recently Exited ★ Long-term EL (LTEL) ★ Ever EL ★
Economic Status	<ul style="list-style-type: none"> Economically Disadvantaged Not Economically Disadvantaged
Gender	<ul style="list-style-type: none"> Male Female Gender X ★ Unknown
Disability Status ★	<ul style="list-style-type: none"> SWD (students with disabilities) ★ Non-SWD (students without disabilities) ★
Migrant Status ★	<ul style="list-style-type: none"> Migrant ★ Non-Migrant ★
Homeless Enrolled Status ★	<ul style="list-style-type: none"> Homeless ★ Non-Homeless ★
Foster Care Status ★	<ul style="list-style-type: none"> Foster Care ★ Non-Foster Care ★
Military Connected Status ★	<ul style="list-style-type: none"> Military ★ Non-Military ★

Note: ★ Indicates a change or new addition from V1.0 to V1.1.

States have some flexibility in determining how they will report student outcome data by race/ethnicity (as granted by the ESEA). Therefore, reported race/ethnicity categories are not uniform, and some states may change student classification conventions over time, add new categories, remove categories, or add more nuance to student classifications.

The race/ethnicity student subgroup with the most variation was for the "Asian" and "Native Hawaiian or Pacific Islander" categories. For example, Hawaii was the only state to disaggregate results for "Native Hawaiian and "Pacific Islander", while several state aggregated results for "Asian" and "Native Hawaiian/Pacific Islander." See Table 14 for details on state differences in race/ethnicity reporting.

Table 14. Race/ Ethnicity Category Variations

State	Year	State Subgroup	Zelma Subgroup
Alaska	2016-17 to 2021-22	"Asian/Pacific Islander"	"Asian"
Hawaii	2015-16 to 2016-17	"Pacific Islander"	"Pacific Islander"
		"Asian/Pacific Islander"	"Asian"
Minnesota	2000-01 to 2012-13	"Asian/Pacific Islander"	"Asian"
	2018-19 to 2021-22	"American Indian or Alaska Native students" under <u>Federal</u> categories	"American Indian or Alaska Native"
	2022-23	"American Indian Students" under <u>State</u> categories)	"American Indian or Alaska Native"
Missouri	2014-15 to 2022-23	"Asian/Pacific Islander"	"Asian"
New Hampshire	2018-19 to 2022-23	"Asian+PI+Hawaiian"	"Asian"
New York	2005-06 to 2021-22	"Asian or Native Hawaiian/ Other Pacific Islander"	"Asian"
Ohio	2015-16 to 2022-23	"Asian or Native Hawaiian/ Other Pacific Islander"	"Asian"
Virginia	1997-98 to 2004-05	"Asian/Pacific Islander"	"Asian"

English Learner Data

Value labels for the **EL Status** Student Group reflect a wide range of classifications used by states (see Table 15). Each state uses its own criteria to define EL students and these additional subcategories. For this reason, there may be differences from one state to another related to how students are classified.; they are not necessarily comparable. Zelma reflects each state's own data classifications for student subgroups.

Table 15. EL Status Value Labels

EL Status	Value Labels
English Learner	Students who meet the state's criteria for being classified as an "English learner" (Note: all states use their own process and procedures for English learner classification and re-classification as "English proficient").
English Proficient	Students who meet the state's criteria for being classified as "English proficient" (Note: all states use their own process and procedures for English learner classification and re-classification as "English proficient").
EL Exited★	Students who had previously been classified as an "English learner" (EL) but who have ever "exited" or "ended" their EL status by attaining English proficiency as measured by the state's EL assessment. These students are considered "English proficient". (Note: all states use their own process and procedures for English learner classification and re-classification as "English proficient").

EL Status	Value Labels
EL Monit or Recently Ex ★	EL Monitored or Recently Exited: Students who had previously been classified as an "English learner" (EL) but who had "exited" or "ended" their EL status by attaining English proficiency in the past 1-4 years as measured by the state's EL assessment. This is different from the "EL Exited" category, which includes ALL EL students who have ever exited the EL status. These students are considered "English proficient". (Note: all states use their own process and procedures for English learner classification and re-classification as "English proficient").
EL and Monit or Recently Ex ★	EL *AND* Monitored or Recently Exited: Students who are *currently* English learners, as well as students who have "exited" or "ended" their EL status by attaining English proficiency in the past 1-4 years as measured by the state's EL assessment. Some students in this category are therefore "English learners" while others are "English proficient". (Note: all states use their own process and procedures for English learner classification and re-classification as "English proficient").
LTEL ★	Long-term English learners: Students who the state has classified as English learners for an extended period of time, as determined by the state. Typically, LTELs are students who have been classified as ELs for a minimum of 6 years. These students are considered "English learners". Definitions vary by state (and not all states define LTELs).
Ever EL ★	Students who have ever been classified as an English learner. Some students in this category are therefore "English learners" while others are "English proficient."

Note: ★ Indicates a change or new addition from V1.0 to V1.1.

q. StudentGroup_TotalTested and StudentSubGroup_TotalTested

Values for **StudentGroup_TotalTested** and **StudentSubGroup_TotalTested** reflect the counts of students tested as reported by the SEA. Where the state did not report counts of students tested, data are applied from the applicable reporting year as available from *EDFacts* (which is currently only available through Spring 2022 assessments), from the prior year's assessment data, or enrollment data. However, *EDFacts* only reports outcomes and tested counts for ELA and math; as a result, many gaps still exist for science and other subjects. In addition, *EDFacts* does not report outcomes for the Native Hawaiian/Pacific Islander student subgroup, as these are aggregated into counts with the "Asian" student subgroup. See

Table 16 for details on student count values for states that did not provide this information in their raw data.

There may be cases where the summation of **StudentSubGroup_TotalTested** does not equal **StudentGroup_TotalTested**. This may be due to the fact that not all students are necessarily accounted for in the student subgroups.

Note: Visualizations on Zelma.ai are dependent on counts of students tested. Where possible, other sources are utilized to best represent the data. However, not all missing states-years have been matched with other data sources.

Table 16. Data Sources for Files with Missing Student Tested Counts for ELA and Math

State	Year	StudentGroup_TotalTested	Zelma Data Substitute
Alabama	2022-23	State removed tested variable as available in prior years.	State's 2021-22SY counts applied
Arizona	2009-10 to 2015-16	No counts	<i>EDFacts</i>
Arkansas	2018-19 to 2022-23	No counts	<i>EDFacts</i> [2021 applied to 2021-22SY and 2022-23SY]

State	Year	StudentGroup_TotalTested	Zelma Data Substitute
Connecticut	2020-21	No counts	<i>EDFacts</i>
Illinois	2014-15 to 2021-22	No counts	<i>EDFacts</i> [2021 ela also applied to math due to missing data]
	2022-23	No counts	<i>EDFacts</i> 2022 applied
Indiana	2004-05 to 2012-13	No counts	No current substitute applied
Kansas	2014-15 to 2021-22	No counts	<i>EDFacts</i>
	2022-23	No counts	<i>EDFacts</i> 2022 applied
Kentucky	2021-22 to 2022-23	State removed tested variable as available in prior years.	<i>EDFacts</i> 2022 applied
Mississippi	2013-14	Partial counts	<i>EDFacts</i>
Nebraska	2015-16 to 2016-17	No counts	<i>EDFacts</i>
New Mexico	2016-17 to 2018-19	No counts	<i>EDFacts</i>
North Dakota	2014-15 to 2021-22	No counts	<i>EDFacts</i>
	2022-23	No counts	<i>EDFacts</i> 2022 applied
Utah	2013-14 to 2020-21	No counts	<i>EDFacts</i>
	2021-22 to 2022-23	No counts	Utah DoE enrollment data due to insufficient <i>EDFacts</i> data
Virginia	1997-98 to 2004-05	No counts	No current substitute applied
West Virginia	2014-15 to 2021-22	No counts	<i>EDFacts</i>
	2022-23	No counts	<i>EDFacts</i> 2022 applied

2.4 Proficiency and Participation

All Zelma data files include important information on data by proficiency level (also known as “achievement levels” or “performance levels”, the criteria used by the SEA to define proficiency for the given year-subject, and the total count and percent of students considered to be at “proficient or above” on the state assessment (see Table 17). Additional variables include the average scale score and the participation rate, if reported. Details are provided below.

Table 17. Proficiency and Participation Variables

Variable Name	Description	
r	Lev1_count	Count of students performing at Level 1
	Lev1_percent	Percent of students performing at Level 1
	Lev2_count	Count of students performing at Level 2
	Lev2_percent	Percent of students performing at Level 2
	Lev3_count	Count of students performing at Level 3
	Lev3_percent	Percent of students performing at Level 3
	Lev4_count	Count of students performing at Level 4
	Lev4_percent	Percent of students performing at Level 4
	Lev5_count	Count of students performing at Level 5
	Lev5_percent	Percent of students performing at Level 5
	Lev6_count*	Count of students performing at Level 6 (only in Ohio file)
	Lev6_percent*	Percent of students performing at Level 6 (only in Ohio file)
s	ProficiencyCriteria	Achievement levels included in state's definition of proficiency
t	ProficientOrAbove_count	Count of students attaining proficiency of above, as defined by the state
u	ProficientOrAbove_percent	Percent of students attaining proficiency of above, as defined by the state
v	AvgScaleScore	Average scale score for the applicable student subgroup.
w	ParticipationRate	Participation rate for the applicable student subgroup

Note: * Indicates a change or new addition from V1.0 to V1.1.

r. Level counts and percents

States vary in the number of achievement levels used to describe student outcomes. Most commonly, state-subject assessments use four achievement levels, such as:

- *Level 1 – Below basic*
- *Level 2 – Basic*
- *Level 3 – Proficient*
- *Level 4 – Advanced*

However, there are some state-subjects that have either fewer than four or more than four achievement levels. For standardization purposes, each state's *lowest achievement level* has been mapped onto the Zelma category **Lev1_count** and **Lev1_percent**, with subsequent levels progressing up through Level 6, as available. There is one case where the *lowest* achievement level was not mapped onto Level 1 due to the variable naming conventions used by the state – Maryland – as presented in Table 18.

In the event that state data provide a *range* of values for each achievement level (e.g., 0.20 – 0.29 percent of students), Zelma retains these ranges in the dataset.

Table 18. Level 1 Exception

State	Years/ Subject	Description	Zelma Var Mapping
Maryland	2017-18 to 2018-19 (science)	• [no Level 1 variable]	Not mapped
		• Level 2 Pct	Lev2_percent
		• Level 3 Pct	Lev3_percent
		• Level 4 Pct	Lev4_percent
		• Level 5 Pct	Lev5_percent

★ **V1.1 Change:** North Carolina’s reported performance levels range from Level 2 to Level 5; for Version 1.0, Zelma files mapped these categories onto Levels 2 to 5 to maintain consistency with the state. However, for Version 1.1, these categories have been re-mapped onto Levels 1 to 4 to reflect greater consistency across all state files (see Table 19).

Table 19. North Carolina Performance Level Re-Mapping in Zelma V1.1

Years	Version	Description	Zelma Var Mapping
2021-22 to 2022-23 (ELA, math, science)	Version 1.0	• [No Level 1 variable]	Not mapped
		• pct_notprof	Lev2_percent
		• pct_l3	Lev3_percent
		• pct_l4	Lev4_percent
		• pct_l5	Lev5_percent
	Version 1.1	• pct_notprof	Lev1_percent
		• pct_l3	Lev2_percent
		• pct_l4	Lev3_percent
		• pct_l5	Lev4_percent

★ **V1.1 Change:** For Version 1.0, Zelma did not have data disaggregated by achievement level for Ohio. This has been incorporated for Version 1.1; as a result, data are now available for all of Ohio’s six achievement levels. These levels have been retained in the Zelma dataset; the Ohio datafiles are thus the only Zelma files that include variables Lev6_count and Lev6_percent.

States with Re-Scaled Proficiency Level Data

The majority of states report outcomes by proficiency level (or “achievement level”) based on the count and percent of students that fall in each of the designated performance categories, typically between Level 1 and Level 4. However, the states of **Kansas**, **Washington**, and **Wisconsin** report proficiency level outcomes in a way that *includes* students for whom there is no test score. To standardize our reporting of proficiency level outcome data across states, Zelma files have re-scaled the proficiency level percentages to remove the students for whom there is not a valid test score. In other words, the percentage of students achieving at each of Levels 1-4 will now add to 100%, compared to raw SEA data files which would have a total of less than 100% due to the additional reporting category of students without a valid test score.

No Achievement Level Data

Some states reported only a single measure for percent of students reaching proficiency or above; these states, therefore, do not have achievement level data. See Table 20 for states that do not provide achievement level data for select years and subjects.

Table 20. States-Years Not Reporting Student Outcomes by Achievement Level

State	Year(s)	Subject(s)
Alaska	2016-17 to 2021-22	ELA, math, science (<i>available beginning 2022-23</i>)
Delaware	2014-15 to 2022-23	ELA, math
	2018-19 to 2022-23	Science
	2018-19 to 2021-22	Social studies
Illinois	2015-16 to 2022-23	Science
Indiana	2005-06 to 2012-13	ELA, math, science, social studies
Iowa	2002-03 to 2013-14	ELA, math
Maine	2014-15	ELA, math, science
New Mexico	2014-15 to 2015-16	ELA, math, science
	2020-21 to 2022-23	ELA, math, science
Oregon	2020-21	ELA, math, science
Virginia	1997-98 to 2001-02	ELA, math, science, social studies, writing

★ **V1.1 Change:** For Version 1.0, Zelma did not have data disaggregated by achievement level for Hawaii, Indiana (2014 to 2023), or Ohio. These states now have achievement level data incorporated into Version 1.1.

s. ProficiencyCriteria

Values include the achievement levels that the state uses for a given year-subject to determine proficiency (e.g., “Levels 3-4”). These are the achievement levels that correspond to the ProficientOrAbove_count and ProficientOrAbove_percent variables, even if achievement level data are not available for a given state. See Appendix E for a full list of proficiency criteria by each state, year, and subject.

t. ProficientOrAbove_count

Values represent the count of students achieving proficiency or above according to the state’s designated proficiency criteria. Values are represented as a whole number or as a range.

u. ProficientOrAbove_percent

Values represent the percentage of students achieving proficiency or above according to the state’s designated proficiency criteria. Values are represented as a decimal or as a decimal range.

v. AvgScaleScore

Value labels for AvgScaleScore reflect the mean scale score for the given student subgroup, if available.

w. ParticipationRate

Values are expressed as a decimal or a decimal range (e.g., If the state reports the participation rate as <25%, it is represented in Zelma as 0-.25). If the state does not provide participation rate data, it

is represented as "--" for all observations. Participation rates are sourced from the original data or derived if the state reports "percent not tested."

2.5 Indicators of Assessment Changes ("Flags")

All Zelma data files include indicators of assessment changes ("flags") from the prior year (see Table 21).

Table 21. Flags for Assessment Changes

Variable Name		Description	Value Labels
x	Flag_AssmtNameChange	Indicator if the state-subject assessment name has changed from the <i>prior year's</i> administration. Typically, but not always, this is accompanied by a change in cut scores.	<ul style="list-style-type: none"> • Y = Yes • N= No change • Not applicable = Subject-area assessment not administered for the relevant school year
y	Flag_CutScoreChange_ELA	Indicator if the state's ELA cut scores have changed from <i>the prior year's</i> administration or not, affecting comparability to prior years. A change represents a new baseline in the ELA assessment.	
z	Flag_CutScoreChange_math	Indicator if the state's math cut scores have changed from the <i>prior year's</i> administration, affecting comparability to prior years. A change represents a new baseline in the math assessment.	
aa	Flag_CutScoreChange_sci ★	Indicator if the state's science cut scores have changed from the <i>prior year's</i> administration, affecting comparability to prior years. A change represents a new baseline in the science assessment.	
bb	Flag_CutScoreChange_soc ★	Indicator if the state's social studies cut scores have changed from the <i>prior year's</i> administration, affecting comparability to prior years. A change represents a new baseline in the social studies assessment.	

Note: ★ Indicates a change or new addition from V1.0 to V1.1.

★ **V1.1 Change:** Version 1.1 removes Flag_CutScoreChange_read and Flag_CutScoreChange_oth.

2.6 NCES Descriptive Variables

All Zelma data files include a set of descriptive variables from NCES, including an indicator if the school or district is part of a charter district, the reported school level, a virtual school indicator, the county name, and the county code (see Table 22), as discussed below.

★ **V1.1 Change:** Version 1.1 adds the DistLocale variable.

Table 22. Descriptive Variables

Variable Name		Description
cc	DistType	District type
dd	DistCharter	Charter indicator for the district
ee	DistLocale ⁺	An indication of school's location relative to a populous area. The locales assigned to school districts are based on the locale code of their schools, weighted by the size of the schools' membership.
ff	SchType	School type
gg	SchLevel	School level
hh	SchVirtual	Virtual school indicator
ii	CountyName	County in which the district is located.
jj	CountyCode	County code in which the district is located, also referred to as the county-level FIPS code

Note: ⁺ Indicates a change or new addition from V1.0 to V1.1.

cc. DistType

Value labels for the type of district come from NCES, which classifies districts into a variety of district types (see Table 23).

Table 23. DistType Values

DistType	Description
Regular local school district	Locally governed agency responsible for providing free public elementary or secondary education; includes independent school districts and those that are a dependent segment of a local government such as a city or county.
Component district	Regular local school district that shares its superintendent and administrative services with other school districts participating in the supervisory union.
Local school district that is a component of a supervisory union	Regular local school district that shares its superintendent and administrative services with other school districts participating in the supervisory union.
Regional education service agency	Agency providing specialized education services to a variety of local education agencies, or a county superintendent serving the same purposes.
State-operated agency	Agency that is charged, at least in part, with providing elementary and/or secondary instruction or support services. Includes the State Education Agency if this agency operates schools. Examples include elementary/secondary schools operated by the state for the deaf or blind; and programs operated by state correctional facilities.
Federal-operated agency	A federal agency that is charged, at least in part, with providing elementary or secondary instruction or support services.
Charter agency	All schools associated with the agency are charter schools.
Supervisory union	An education agency that performs administrative services for more than one school district, providing a common superintendent for participating districts.
Specialized public school district	A specialized public school district is a school district that operates one or more schools that are designed for a specific educational need or purpose.
Other education agency	Agency providing elementary or secondary instruction or support services that does not fall within the definitions of other agency types.
Missing/not reported	School type not available.

dd. DistLocale

Value labels come from NCES, which classifies districts based on their proximity to populous areas. The value labels used until the 2005-06SY are referred to as the “metro-centric locale” (see codebook for these values), while value labels beginning in the 2006-07SY are referred to as the “urban-centric locale.” The locales assigned to school districts are based on the locale code of their schools, weighted by the size of the schools' membership. The value labels and their definitions are presented in Table 24.

Table 24. DistLocale Values

DistType	Description
City, Large	Territory inside an urbanized area and inside a principal city with population of 250,000 or more.
City, Midsize	Territory inside an urbanized area and inside a principal city with population less than 250,000 and greater than or equal to 100,000.
City, Small	Territory inside an urbanized area and inside a principal city with population less than 100,000.
Suburb, Large	Territory outside a principal city and inside an urbanized area with population of 250,000 or more.
Suburb, Midsize	Territory outside a principal city and inside an urbanized area with population less than 250,000 and greater than or equal to 100,000.
Suburb, Small	Territory outside a principal city and inside an urbanized area with population less than 100,000.
Town, Fringe	Territory inside an urban cluster that is less than or equal to 10 miles from an urbanized area.
Town, Distant	Territory inside an urban cluster that is more than 10 miles and less than or equal to 35 miles from an urbanized area.
Town, Remote	Territory inside an urban cluster that is more than 35 miles of an urbanized area.
Rural, Fringe	Census-defined rural territory that is less than or equal to 5 miles from an urbanized area, as well as rural territory that is less than or equal to 2.5 miles from an urban cluster.
Rural, Distant	Census-defined rural territory that is more than 5 miles but less than or equal to 25 miles from an urbanized area, as well as rural territory that is more than 2.5 miles but less than or equal to 10 miles from an urban cluster.
Rural, Remote	Census-defined rural territory that is more than 25 miles from an urbanized area and is also more than 10 miles from an urban cluster.

ee. DistCharter

Value labels for the district charter indicator come from NCES, which indicate if a district is classified as a charter (Yes) or not (No). Observations missing this information are reported as “Missing/not reported.” This variable will only have values for district- and school-level data.

a. SchType

Value labels for the type of school come from NCES, which classifies districts according to distinct characteristics (see Table 25).

Table 25. SchType Values

SchType	Description
Regular school	A public elementary/secondary school that does not focus primarily on vocational, special, or alternative education.
Special education school	Public elementary/secondary school that focuses primarily on the following: hard of hearing, deaf, speech-impaired, health-impaired, orthopedically impaired, mentally retarded, seriously emotionally disturbed, multi-handicapped, visually handicapped, deaf and blind, and adapts curriculum, materials or instruction for students served.
Vocational school	Public elementary/secondary school that focuses primarily on vocational education, provides education in one or more semi-skilled technical operations.
Other	A public elementary/secondary school that addresses the needs of students which typically cannot be met in a regular school, provides nontraditional education, serves as an adjunct to a regular school, and falls outside of the categories of regular, special education, vocational education.
Reportable program	A program within a school that may be self-contained, but does not have its own principal, and is not a school according to the NCES definition of a school.
Missing/not reported	School type not available.

b. SchLevel

Value labels for the school level indicator come from NCES, which classifies all schools according to their primary level of instruction (see Table 26). This variable will only have values for school-level data.

Table 26. SchLevel Values

SchType	Low Grade	High Grade
Primary	Prekindergarten through 03	Up to 08
Middle	Grade 04 through 07	04 through 09
High	Grade 07 through 12	12 only
Ungraded	Ungraded	Ungraded
Not applicable	Not applicable	
Other	Any combination of grade levels not included above.	
Missing/not reported	School type not available.	

c. SchVirtual

Value labels for the virtual school indicator come from NCES, which classifies all schools by whether or not they operate as virtual schools. The first year with available data is the 2015-16 school year, so years prior to this will have "Missing/not reported" for all school values. This variable will only have values for school-level data.

d. CountyName

Values reflect the name of the county in which the district is located, available from NCES. This variable will only have values for district- and school-level data. In some cases where a school or district falls on the border of two states, the school/district's county may differ from the state with the school/district assessment data.

e. CountyCode

Values reflect the county code in which the district is located, also referred to as the county-level FIPS code, available from NCES. This variable will only have values for district- and school-level data.

Appendix A. Technical Guide Version History

Table A1. Technical Guide Version History

Version	Date Released	Description
Version 1.0	Dec 6, 2023	Initial data launch
Version 1.1	June 18, 2024	<ul style="list-style-type: none"> • Added remaining 2023 assessment data [now includes all 50 states] • Added new student subgroup proficiency data, including outcomes by disability status, migrant status, homeless enrolled status, foster care status, and military connected status. • Added new subgroups within existing student groups of race/ethnicity, gender, and English learner status. • Added new grade level x subgroup data for Arkansas, Florida, Hawaii, Indiana, Iowa, Nevada, Oklahoma. • Added new outcomes by achievement level for Hawaii, Ohio. • Added achievement level counts for all states based on achievement level percents. • Added new flags for changes in science and social studies assessments compared to prior year.

Appendix B. Data Sources

Table B1. Data Sources

State	Data Source(s)
Alabama	<p>Alabama State Department of Education. <i>Report Card: Student Participation & Proficiency, SY 2014-15 – 2022-23.</i> https://reportcard.alsde.edu/SupportingData_Proficiency.aspx</p> <p>Alabama State Department of Education. <i>Reports & Data (School Performance: Proficiency).</i> https://www.alabamaachieves.org/reports-data/school-performance/</p>
Alaska	<p>Alaska Department of Education & Early Development (DEED). <i>Proficiency data received via data request, 2016-17SY – 2021-22SY</i> (Received July 26, 2023).</p> <p>Alaska Department of Education & Early Development (DEED). <i>Proficiency data received via data request, 2022-23SY</i> (Received April 24, 2024).</p>
Arizona	<p>Arizona Department of Education. <i>State Assessment Results (ela and math), State Science Assessment Results (science).</i> https://www.azed.gov/accountability-research/data</p>
Arkansas	<p>Arkansas Department of Education. <i>Assessment: Test Scores.</i> https://dese.ade.arkansas.gov/Offices/learning-services/assessment-test-scores/2023</p> <p>Arkansas Department of Education. <i>ADE Data Center: My School Info (2016-2023 ACT Aspire).</i> https://myschoolinfo.arkansas.gov/</p>
California	<p>California Department of Education. <i>English Language Arts/Literacy and Mathematics: Smarter Balanced Summative Assessments, SY 2014-15 – 2022-23</i> (Research Files). https://caaspp-elpac.ets.org/caaspp/ResearchFileListSB</p> <p>California Department of Education. <i>STAR Data Files, SY 2002-03 – 2012-13</i> (Research Files). https://caaspp-elpac.ets.org/caaspp/</p>
Colorado	<p>Colorado Department of Education. <i>CMAS – Mathematics, English Language Arts, Science and Social Studies Data and Results.</i> https://www.cde.state.co.us/assessment/cmas-dataandresults</p>
Connecticut	<p>Connecticut State Department of Education – EdSight. <i>Smarter Balanced Achievement/Participation.</i> https://public-edsight.ct.gov/performance/smarter-balanced-achievement-participation</p> <p>Connecticut State Department of Education – EdSight. <i>Next Generation Science Standards (NGSS) Assessment.</i> https://public-edsight.ct.gov/performance/smarter-balanced-achievement-participation</p>
Delaware	<p>Delaware Department of Education – Delaware Open Data. <i>Student Assessment Performance, SY 2014-15 – 2022-23.</i> https://data.delaware.gov/Education/Student-Assessment-Performance/ms6b-mt82/about_data</p> <p>Delaware Department of Education. <i>2017 DCAS Science State Summary.</i> https://www.doe.k12.de.us/Page/3470</p> <p>Delaware Department of Education. <i>Assessments: DCAS State Summary Reports, SY 2014-15 – 2015-16.</i> https://www.doe.k12.de.us/Page/2416</p>
District of Columbia	<p>DC Office of the State Superintendent of Instruction (OSSE). <i>The Partnership for Assessment Readiness for College and Careers (PARCC) (ELA and Mathematics Assessment Results).</i> https://osse.dc.gov/parcc</p> <p>DC Office of the State Superintendent of Instruction (OSSE). <i>DC Science Assessment (Assessment Results).</i> https://osse.dc.gov/science</p>

Table B1. Data Sources (continued)

State	Data Source(s)
Florida	<p>Florida Department of Education. <i>Know Your Data PK-12 Advanced Reports, SY 2017-18 – 2022-23</i> (Assessments). https://edudata.fldoe.org/AdvancedReports.html</p> <p>Florida Department of Education. <i>K-12 Student Assessment: Results, SY 2014-15 – 2016-17</i> (Results). https://www.fldoe.org/accountability/assessments/k-12-student-assessment/results/</p> <p>Florida Department of Education. <i>FCAT 2.0 Historical, SY 2010-11 – 2013-14</i> (Retrofitted Statewide Assessment Scores). https://www.fldoe.org/accountability/assessments/k-12-student-assessment/archive/fcat-2-0/retrofitted-statewide-assessment-score/</p> <p>Florida Department of Education. <i>FCAT Historical, SY 1997-98 – 2009-10</i> (Scores & Reports). https://www.fldoe.org/accountability/assessments/k-12-student-assessment/archive/fcat/scores-reports/index.shtml</p>
Georgia	<p>Governor’s Office of Student Achievement. <i>Georgia Milestones End-of-Grade (EOG) Assessments (by grade)</i> (Downloadable Data). https://gosa.georgia.gov/dashboards-data-report-card/downloadable-data</p>
Hawaii	<p>Hawai’i State Department of Education. <i>Proficiency data received via data request, 2014-15SY – 2022-23SY</i> (Received May 29, 2024).</p>
Idaho	<p>Idaho Department of Education. <i>Proficiency data received via data request, SY 2015-16 – 2022-23.</i> (Received November 27, 2023).</p>
Illinois	<p>Illinois State Board of Education. <i>Data & Accountability: Report Card Data Library.</i> (Report Card Public Data Sets). https://www.isbe.net/Pages/Illinois-State-Report-Card-Data.aspx</p> <p>Illinois State Board of Education. <i>Assessment: Illinois Science Assessment (ISA).</i> (Illinois Science Assessment Results). https://www.isbe.net/Pages/Illinois-Science-Assessment.aspx</p>
Indiana	<p>Indiana Department of Education. <i>Data Reports Archive, SY 2013-14 – 2021-22</i> (ILEARN, ISTEP+). https://www.in.gov/doe/it/data-center-and-reports/data-reports-archive/#ILEARN</p> <p>Indiana Department of Education. <i>Data Center & Reports, SY 2022-23</i> (ILEARN Assessment Results). https://www.in.gov/doe/it/data-center-and-reports/</p> <p>Indiana Department of Education. <i>Proficiency data by grade and subgroup received via data request, SY 2013-14 – 2022-23.</i> (Received April 3-June 14, 2023).</p>
Iowa	<p>Iowa Department of Education. <i>Proficiency data received via data request, SY 2014-15 – 2022-23</i> (Received December 1, 2023).</p> <p>Iowa Department of Education. <i>PK-12 Education Statistics, SY 2002-03 – 2013-14</i> (Student Performance: Assessments – Proficiency Rate). https://educate.iowa.gov/pk-12/data/education-statistics#Student_Performance</p>
Kansas	<p>Kansas State Department of Education (KSDE). <i>Performance Indicators</i> (Longitudinal Performance Level Reports). https://ksreportcard.ksde.org/assessment_results.aspx</p> <p>Kansas State Department of Education (KSDE). <i>Performance Indicators</i> (Participation Summary Report). https://ksreportcard.ksde.org/part_details.aspx</p>

Table B1. Data Sources (continued)

State	Data Source(s)
Kentucky	Kentucky Department of Education. <i>School Report Card: Data Sets, SY 2020-21 – 2022-23</i> (Academic Performance: State Assessments). https://www.kyschoolreportcard.com/datasets
	Kentucky Department of Education. <i>School Report Card Datasets, SY 2012-13 – 2018-19</i> (Assessments/Accountability). https://openhouse.education.ky.gov/Home/SRCData
Louisiana	Louisiana Department of Education. <i>Proficiency data received via data request, SY 2014-15 – 2022-23</i> (Received September 13, 2023; December 12, 2023).
Maine	Maine Department of Education. <i>ESSA Dashboard, SY 2015-16 – 2022-23</i> (Academic Performance on the Assessments by Student Population). https://www.maine.gov/doe/dashboard
	Maine Department of Education. <i>Maine Assessment Legacy Data, SY 2014-15</i> (Academic Performance on the Assessments by Student Population). https://www.maine.gov/doe/data-reporting/reporting/legacy-assessment-data
Maryland	Maryland State Department of Education. <i>Data Downloads: Public Use Data for Download</i> (MCAP, PARCC, MSA data). https://reportcard.msde.maryland.gov/Graphs/#/DataDownloads
Massachusetts	Massachusetts Department of Elementary and Secondary Education (DESE). <i>Statewide Reports</i> (Assessment and Accountability). https://profiles.doe.mass.edu/statereport/#Assessment%20and%20Accountability
	Massachusetts Department of Elementary and Secondary Education (DESE). <i>Next Generation MCAS Achievement Results</i> (E2C Hub MCAS Data Explorer). https://educationtocareer.data.mass.gov/Assessment-and-Accountability/Next-Generation-MCAS-Achievement-Results/i9w6-niyt/about_data
Michigan	Michigan’s Center for Educational Performance and Information (CEPI). <i>K-12 Data Files, SY 2014-15 – 2022-23.</i> https://www.mischooldata.org/k-12-data-files/ (accessed September 11, 2023)
Minnesota	Minnesota Department of Education. <i>Data Center, Data Reports and Analytics</i> (Assessment Files). https://public.education.mn.gov/MDEAnalytics/DataTopic.jsp?TOPICID=1
Mississippi	Mississippi Department of Education. <i>Student Assessment</i> (Assessment Results). https://www.mdek12.org/OPR/Reporting/Assessment
	Mississippi Department of Education. <i>Files received via data request, SY 2015-16 – 2018-19</i> (Received March 5, 2024).
Missouri	Missouri Department of Elementary & Secondary Education. <i>Missouri Comprehensive Data System</i> (Students: Missouri Assessment Program (MAP) Data). https://apps.dese.mo.gov/MCDS/home.aspx?categoryid=2&view=2
Montana	Montana Office of Public Instruction. <i>Interactive Dashboards: Student Achievement & Performance.</i> https://gems.opi.mt.gov/student-data
	Montana Office of Public Instruction. <i>Proficiency data received via data request, SY 2015-16 – 2022-23</i> (Received January 6, 2024).
Nebraska	Nebraska Department of Education. <i>Data Downloads</i> (NSCAS, NeSA Assessment Data). https://nep.education.ne.gov/Links

Table B1. Data Sources (continued)

State	Data Source(s)
Nevada	Nevada Department of Education. <i>Data Interaction for Nevada Accountability Portal</i> (Group Summary Report). http://nevadareportcard.nv.gov/di/main/assessment
New Hampshire	New Hampshire Department of Education. <i>Bureau of Education Statistics: Assessment Data</i> (Disaggregated Data File Regular Denominator). https://www.education.nh.gov/who-we-are/division-of-educator-and-analytic-resources/bureau-of-education-statistics/assessment-data
New Jersey	New Jersey Department of Education. <i>Assessment</i> (Statewide Assessment Reports). https://www.nj.gov/education/assessment/results/reports/
New Mexico	New Mexico Public Education Department (NM PED). <i>Achievement Data</i> (New Mexico Accountability Data by Year). https://webnew.ped.state.nm.us/bureaus/accountability/achievement-data/ New Mexico Public Education Department (NM PED). <i>Proficiency data received via data request, SY 2016-17 – 2022-23</i> (Received December 6, 2023).
New York	New York State Department of Education (NYSDE). <i>Downloads</i> (Report Card Database). https://data.nysed.gov/downloads.php
North Carolina	North Carolina Department of Public Instruction. <i>Accountability Data Sets and Reports</i> (Disaggregated Performance Data). https://www.dpi.nc.gov/districts-schools/testing-and-school-accountability/school-accountability-and-reporting/accountability-data-sets-and-reports
North Dakota	North Dakota State Government. <i>Download Insights Data</i> (Assessment Performance; Assessment Participation). https://insights.nd.gov/Data
Ohio	Ohio Department of Education and Workforce. <i>Proficiency data received via data request, SY 2015-16 – 2022-23</i> (Received May 19, 2024).
Oklahoma	Oklahoma State Department of Education. <i>State Testing Resources</i> (State Assessments Summary Reports). https://sde.ok.gov/state-testing-resources Oklahoma State Department of Education. <i>Proficiency data received via data request, SY 2016-17 – 2022-23</i> (Received April 25, 2024).
Oregon	Oregon Department of Education. <i>Assessment Group Reports</i> (English Language Arts, Mathematics, Science). https://www.oregon.gov/ode/educator-resources/assessment/Pages/Assessment-Group-Reports.aspx
Pennsylvania	Pennsylvania Department of Education. <i>Files received via data request, SY 2014-15 – 2021-22</i> (Received September 25, 2023). Pennsylvania Department of Education. <i>Data and Reporting, Assessments, PSSA Results, SY 2022-23.</i> https://www.education.pa.gov/DataAndReporting/Assessments/Pages/PSSA-Results.aspx
Rhode Island	Rhode Island Department of Education. <i>Assessment Data Portal.</i> https://www3.ride.ri.gov/ADP
South Carolina	South Carolina Department of Education. <i>Data, Test Scores, State Assessments, SC READY, SY 2015-16 – 2022-23.</i> https://ed.sc.gov/data/test-scores/state-assessments/sc-ready/ South Carolina Department of Education. <i>Data, Test Scores, State Assessments, SC PASS, SY 2015-16 – 2022-23.</i> https://ed.sc.gov/data/test-scores/state-assessments/scpalmetto-assessment-of-state-standards-pass/
South Dakota	South Dakota Department of Education. <i>Proficiency data received via data request, SY 2014-15 – 2022-23</i> (Received October 16, 2023).

Table B1. Data Sources (continued)

State	Data Source(s)
Tennessee	Tennessee Department of Education. <i>Federal Programs and Oversight: Data Downloads & Requests</i> (State Assessments: Assessment Files). https://www.tn.gov/education/districts/federal-programs-and-oversight/data/data-downloads.html
Texas	Texas Education Agency. <i>Student Assessment Results</i> (STAAR Aggregate Data). https://tea.texas.gov/student-assessment/testing/student-assessment-results/staar-aggregate-data
Utah	Utah State Board of Education (USBE). <i>Reports</i> (Readiness Improvement Success Empowerment (RISE)/Utah Aspire Plus). https://www.schools.utah.gov/datastatistics/reports
Vermont	Vermont Agency of Education (AOE). <i>Vermont Education Dashboard: Assessment.</i> https://education.vermont.gov/data-and-reporting/vermont-education-dashboard/vermont-education-dashboard-assessment
Virginia	Virginia Department of Education. <i>Test Results Build-A-Table, SY 2005-06 – 2022-23</i> (SOL Test Results). https://p1pe.doe.virginia.gov/apex_captcha/home.do?apexTypeId=306 Virginia Department of Education. <i>SOL Test Pass Rates & Other Results: Archived Reports, SY 1997-98 – 2004-05</i> (1998-2005 Miscellaneous Reports). https://www.doe.virginia.gov/data-policy-funding/data-reports/statistics-reports/archived-reports Virginia Department of Education. <i>Participation rate data received via data request, SY 2015-16 – 2022-23.</i> (Received December 1, 2023).
Washington	Washington Office of Superintendent of Public Instruction. <i>Data Portal</i> (Report Card Assessment Data). https://ospi.k12.wa.us/data-reporting/data-portal
West Virginia	West Virginia Department of Education. <i>State Assessment Results.</i> https://zoomwv.k12.wv.us/Dashboard/dashboard/7301
Wisconsin	Wisconsin Department of Public Instruction. <i>WISEdash Data Files by Topic</i> (Forward data). https://dpi.wi.gov/wisedash/download-files/type?field_wisedash_upload_type_value=Forward
Wyoming	Wyoming Department of Education (WDE). <i>Assessment Reports: Grades 3-10 WY-TOPP/WY-ALT Performance Results</i> (Aggregated and Disaggregated Results). https://edu.wyoming.gov/data/assessment-reports/.

Appendix C. State-Subject-Year Data Availability

Longitudinal data varies by state, with the first year of available data ranging from 1998 to 2018. All Spring 2023 assessment data released as of June 18, 2024 are included as part of the Zelma Version 1.1 dataset. The data range in Table C1 presents the earliest possibly year available for the state-subject and the latest possible year available. A few notes:

- **No assessments in Spring 2020:** No states administered assessments in Spring 2020 due to the COVID-19 pandemic; therefore, no data are included as part of Zelma for this school year.
- **Missing years:** There may be periodic years of missing data for all subjects within a state or for a particular subject area. Often this is due to field testing a new assessment. Some states received assessment waivers from ED in Spring 2021 due to the pandemic.

Table C1. State-Subject-Year Data Availability

State	Subject	Data Start	Data End	Missing Years
Alabama	ELA	2015	2023	
Alabama	Math	2015	2023	
Alabama	Science	2015	2023	
Alaska	ELA	2017	2023	
Alaska	Math	2017	2023	
Alaska	Science	2017	2023	2021
Arizona	ELA	2010	2023	
Arizona	Math	2010	2023	
Arizona	Science	2010	2023	2021
Arkansas	ELA	2009	2023	
Arkansas	Math	2009	2023	
Arkansas	Reading	2016	2023	
Arkansas	Science	2009	2023	
Arkansas	STEM	2016	2023	
Arkansas	Writing	2016	2017	
California	ELA	2010	2023	2014
California	Math	2010	2023	2014
California	Science	2010	2013	
California	Social Studies	2010	2013	
Colorado	ELA	2015	2023	
Colorado	Math	2015	2023	
Colorado	Science	2015	2023	2022
Colorado	Social Studies	2015	2015	
Connecticut	ELA	2015	2023	
Connecticut	Math	2015	2023	
Connecticut	Science	2019	2023	
Delaware	ELA	2015	2023	
Delaware	Math	2015	2023	
Delaware	Science	2015	2023	2018

Table C1. State-Subject-Year Data Availability (continued)

State	Subject	Data Start	Data End	Missing Years
Delaware	Social Studies	2015	2022	2017, 2018
District of Columbia	ELA	2015	2023	2021
District of Columbia	Math	2015	2023	2021
District of Columbia	Science	2019	2023	2021
Florida	ELA	2015	2023	
Florida	Math	2015	2023	
Florida	Science	2015	2023	
Georgia	ELA	2011	2023	
Georgia	Math	2011	2023	
Georgia	Reading	2011	2014	
Georgia	Science	2011	2023	
Georgia	Social Studies	2011	2023	
Hawaii	ELA	2015	2023	
Hawaii	Math	2015	2023	
Hawaii	Science	2015	2023	
Idaho	ELA	2016	2023	
Idaho	Math	2016	2023	
Idaho	Science	2016	2023	2021
Illinois	ELA	2015	2023	
Illinois	Math	2015	2023	
Illinois	Science	2016	2023	
Indiana	ELA	2005	2023	
Indiana	Math	2005	2023	
Indiana	Science	2011	2023	
Indiana	Social Studies	2011	2023	
Iowa	ELA	2003	2023	
Iowa	Math	2003	2023	
Iowa	Science	2019	2023	
Kansas	ELA	2015	2023	
Kansas	Math	2015	2023	
Kansas	Science	2019	2023	
Kentucky	ELA	2012	2023	
Kentucky	Math	2012	2023	
Kentucky	Science	2012	2023	2015, 2016, 2017
Kentucky	Social studies	2012	2023	2021
Kentucky	Writing	2012	2023	
Louisiana	ELA	2015	2023	
Louisiana	Math	2015	2023	
Louisiana	Science	2015	2023	2018
Louisiana	Social Studies	2015	2023	2016

Table C1. State-Subject-Year Data Availability (continued)

State	Subject	Data Start	Data End	Missing Years
Maine	ELA	2015	2023	
Maine	Math	2015	2023	
Maine	Science	2015	2023	2021
Maryland	ELA	2015	2023	
Maryland	Math	2015	2023	
Maryland	Science	2015	2023	2017
Massachusetts	ELA	2010	2023	
Massachusetts	Math	2010	2023	
Massachusetts	Science	2010	2023	2015, 2016
Michigan	ELA	2015	2023	
Michigan	Math	2015	2023	
Michigan	Science	2015	2023	2018, 2019
Michigan	Social Studies	2015	2023	
Minnesota	ELA	1998	2023	
Minnesota	Math	1998	2023	
Minnesota	Science	2008	2023	
Minnesota	Writing	1998	2005	
Mississippi	ELA	2014	2023	
Mississippi	Math	2014	2023	
Mississippi	Science	2014	2023	
Missouri	ELA	2010	2023	
Missouri	Math	2010	2023	
Missouri	Science	2010	2023	2018
Montana	ELA	2016	2023	
Montana	Math	2016	2023	
Nebraska	ELA	2016	2023	
Nebraska	Math	2016	2023	
Nebraska	Science	2016	2023	2021
Nebraska	Writing	2016	2016	
Nevada	ELA	2016	2023	
Nevada	Math	2016	2023	
Nevada	Science	2017	2023	
New Hampshire	ELA	2009	2023	
New Hampshire	Math	2009	2023	
New Hampshire	Science	2009	2023	2014
New Jersey	ELA	2015	2023	2021
New Jersey	Math	2015	2023	2021
New Jersey	Science	2019	2022	2021
New Mexico	ELA	2017	2023	
New Mexico	Math	2017	2023	

Table C1. State-Subject-Year Data Availability (continued)

State	Subject	Data Start	Data End	Missing Years
New Mexico	Science	2017	2023	2021
New York	ELA	2006	2023	
New York	Math	2006	2023	
New York	Science	2006	2023	
New York	Social Studies	2006	2010	
North Carolina	ELA	2014	2023	
North Carolina	Math	2014	2023	
North Carolina	Science	2014	2023	
North Dakota	ELA	2015	2023	
North Dakota	Math	2015	2023	
North Dakota	Science	2015	2023	
Ohio	ELA	2016	2023	
Ohio	Math	2016	2023	
Ohio	Science	2016	2023	
Ohio	Social Studies	2016	2017	
Oklahoma	ELA	2017	2023	
Oklahoma	Math	2017	2023	
Oklahoma	Science	2017	2023	
Oregon	ELA	2015	2023	
Oregon	Math	2015	2023	
Oregon	Science	2015	2023	
Pennsylvania	ELA	2015	2023	
Pennsylvania	Math	2015	2023	
Pennsylvania	Science	2015	2023	
Rhode Island	ELA	2018	2023	
Rhode Island	Math	2018	2023	
Rhode Island	Science	2019	2023	
South Carolina	ELA	2016	2023	
South Carolina	Math	2016	2023	
South Carolina	Science	2016	2023	
South Carolina	Social Studies	2016	2019	
South Dakota	ELA	2003	2023	2014
South Dakota	Math	2003	2023	2014
South Dakota	Science	2007	2023	
Tennessee	ELA	2010	2023	2016
Tennessee	Math	2010	2023	2016
Tennessee	Science	2010	2023	2016, 2019
Tennessee	Social Studies	2013	2023	2015, 2016, 2017
Texas	ELA	2012	2023	
Texas	Math	2012	2023	

Table C1. State-Subject-Year Data Availability (continued)

State	Subject	Data Start	Data End	Missing Years
Texas	Science	2012	2023	
Texas	Social Studies	2012	2023	
Texas	Writing	2012	2021	
Utah	ELA	2014	2023	
Utah	Math	2014	2023	
Utah	Science	2014	2023	
Vermont	ELA	2016	2023	
Vermont	Math	2016	2023	
Vermont	Science	2019	2023	
Virginia	Math	1998	2023	
Virginia	ELA	1998	2023	
Virginia	Science	1998	2023	
Virginia	Social Studies	1998	2014	
Virginia	Writing	1998	2023	
Washington	ELA	2015	2023	
Washington	Math	2015	2023	
Washington	Science	2018	2023	
West Virginia	ELA	2015	2023	
West Virginia	Math	2015	2023	
West Virginia	Science	2019	2023	
Wisconsin	ELA	2016	2023	
Wisconsin	Math	2016	2023	
Wisconsin	Science	2016	2023	
Wisconsin	Social Studies	2016	2023	
Wyoming	ELA	2014	2023	
Wyoming	Math	2014	2023	
Wyoming	Science	2014	2023	
Wyoming	Writing	2014	2014	

Appendix D. New York City NCES District IDs

Table D1. New York City School District's Subordinate School Districts

Subordinate District Name	NCESDistrictID
New York City Geographic District #1	3600076
New York City Geographic District #2	3600077
New York City Geographic District #3	3600078
New York City Geographic District #4	3600079
New York City Geographic District #5	3600081
New York City Geographic District #6	3600083
New York City Geographic District #7	3600084
New York City Geographic District #8	3600085
New York City Geographic District #9	3600086
New York City Geographic District #10	3600087
New York City Geographic District #11	3600088
New York City Geographic District #12	3600090
New York City Geographic District #13	3600091
New York City Geographic District #14	3600119
New York City Geographic District #15	3600092
New York City Geographic District #16	3600094
New York City Geographic District #17	3600095
New York City Geographic District #18	3600096
New York City Geographic District #19	3600120
New York City Geographic District #20	3600151
New York City Geographic District #21	3600152
New York City Geographic District #22	3600153
New York City Geographic District #23	3600121
New York City Geographic District #24	3600098
New York City Geographic District #25	3600122
New York City Geographic District #26	3600099
New York City Geographic District #27	3600123
New York City Geographic District #28	3600100
New York City Geographic District #29	3600101
New York City Geographic District #30	3600102
New York City Geographic District #31	3600103
New York City Geographic District #32	3600097

Appendix E. Proficiency Criteria by State-Year-Subject

Table E1. Proficiency Criteria by State-Year-Subject

State	Year(s)	Subject	Proficiency Criteria
Alabama	2014-15 to 2022-23	ELA, math, science	Levels 3-4
Alaska	2016-17 to 2022-23	ELA, math, science	Levels 3-4
Arizona	2009-10 to 2022-23	ELA, math, science	Levels 3-4
Arkansas	2008-09 to 2013-14	ELA, math, science	Levels 3-4
	2014-15	ELA, math	Levels 4-5
	2014-15	Science	Levels 3-4
	2015-16 to 2022-23	ELA, math, science, reading, STEM, English	Levels 3-4
California	2009-10 to 2012-13	ELA, math, science, social studies	Levels 4-5
	2014-15 to 2022-23	ELA, math	Levels 3-4
Colorado	2014-15	Social studies	Levels 3-4
	2014-15 to 2022-23	Science	Levels 3-4
	2014-15 to 2022-23	ELA, math	Levels 4-5
Connecticut	2014-15 to 2022-23	ELA, math, science	Levels 3-4
Delaware	2014-15 to 2022-23	ELA, math, science	Levels 3-4
District of Columbia	2014-15 to 2022-23	ELA, math	Levels 4-5
	2018-19 to 2022-23	Science	Levels 3-4
Florida	2014-15 to 2022-23	ELA, math, science	Levels 3-5
Georgia	2010-11 to 2013-14	ELA, math, science, reading, social studies	Levels 2-3
	2014-15 to 2022-23	ELA, math, science, social studies	Levels 3-4
Hawaii	2014-15 to 2022-23	ELA, math, science	Levels 3-4
Illinois	2014-15 to 2022-23	ELA, math	Levels 4-5
	2015-16 to 2018-19	Science	Level 2
	2020-21 to 2022-23	Science	Levels 3-4
Indiana	2004-05 to 2017-18	ELA, math	Levels 2-3
	2010-11 to 2017-18	Science, social studies	Levels 2-3
	2018-19 to 2022-23	ELA, math, science, social studies	Levels 3-4
Iowa	2002-03 to 2022-23	ELA, math	Levels 2-3
	2018-19 to 2022-23	Science	Levels 2-3
Kansas	2014-15 to 2022-23	ELA, math	Levels 3-4
	2018-19 to 2022-23	Science	Levels 3-4
Kentucky	2011-12 to 2022-23	ELA, math, science, social studies	Levels 3-4
Louisiana	2013-14 to 2022-23	ELA, math, science, social studies	Levels 4-5
Maine	2014-15 to 2018-19	ELA, math, science	Levels 3-4
	2020-21 to 2021-22	ELA, math	Levels 2-3
	2021-22	Science	Levels 3-4
	2022-23	ELA, math, science	Levels 3-4

Table E1. Proficiency Criteria by State-Year-Subject (continued)

State	Year(s)	Subject	Proficiency Criteria
Maryland	2014-15 to 2015-16	Science	Levels 2-3
	2017-18 to 2018-19	Science	Levels 4-5
	2020-21	Science	Levels 2-3
	2014-15 to 2018-19	ELA, math	Levels 4-5
	2020-21	ELA, math	Levels 2-3
	2021-22 to 2022-23	ELA, math, science	Levels 3-4
Massachusetts ^a	2009-10 to 2014-15	ELA, math, science	Levels 3-4
	2014-15 to 2015-16	ELA, math	Levels 4-5
	2016-17 to 2022-23	ELA, math, science	Levels 3-4
Michigan	2014-15 to 2022-23	ELA, math, science	Levels 3-4
Minnesota	1997-98 to 2004-05	ELA, math	Levels 3-5
	2005-06 to 2022-23	ELA, math, science	Levels 3-4
Mississippi	2013-14	ELA, math, science	Levels 3-4
	2014-15 to 2022-23	ELA, math	Levels 4-5
	2014-15 to 2017-18	Science	Levels 3-4
	2018-19 to 2022-23	Science	Levels 4-5
Missouri	2009-10 to 2022-23	ELA, math, science	Levels 3-4
Montana	2015-16 to 2022-23	ELA, math, science	Levels 3-4
Nebraska	2015-16 to 2022-23	ELA, math, science	Levels 2-3
Nevada	2015-16	ELA, math	Levels 3-4
	2016-17 to 2022-23	ELA, math, science	Levels 3-4
New Hampshire	2008-09 to 2022-23	ELA, math, science	Levels 3-4
New Jersey	2014-15 to 2022-23	ELA, math	Levels 4-5
	2018-19 to 2022-23	Science	Levels 3-4
New Mexico	2016-17 to 2018-19	ELA, math	Levels 4-5
	2020-21 to 2022-23	ELA, math	Levels 3-4
	2016-17 to 2022-23	Science	Levels 3-4
New York	2005-06 to 2009-10	ELA, math, science, social studies	Levels 3-4
	2010-11 to 2022-23	ELA, math, science	Levels 3-4
North Carolina	2013-14 to 2017-18	ELA, math, science	Levels 3-5
	2019	ELA, science	Levels 3-5
	2019	Math	Levels 2-4
	2020-21 to 2022-23	ELA, math, science	Levels 2-4
North Dakota	2014-15 to 2022-23	ELA, math, science	Levels 3-4
Ohio	2015-16 to 2016-17	ELA, math, science, social studies	Levels 3-6
	2017-18 to 2022-23	ELA, math, science	Levels 3-6
Oklahoma	2016-17 to 2022-23	ELA, math, science	Levels 3-4

Table E1. Proficiency Criteria by State-Year-Subject (continued)

State	Year(s)	Subject	Proficiency Criteria
Oregon	2014-15 to 2017-18	Science	Levels 4-5
	2019-22 to 2022-23	Science	Levels 3-4
	2014-15 to 2022-23	ELA, math	Levels 3-4
Pennsylvania	2014-15 to 2022-23	ELA, math, science	Levels 3-4
Rhode Island	2018-19 to 2022-23	ELA, math, science	Levels 3-4
South Carolina	2015-16 to 2022-23	ELA, math	Levels 3-4
	2015-16	Science	Levels 2-3
	2016-17 to 2022-23	Science	Levels 3-4
	2015-16 to 2018-19	Social studies	Levels 2-3
South Dakota	2002-03 to 2022-23	ELA, math, science	Levels 3-4
Tennessee	2009-10 to 2022-23	ELA, math, science	Levels 3-4
	2012-13 to 2022-23	Social studies	Levels 3-4
Texas ^b	2011-12 to 2015-16	ELA, math, science, social studies, writing	Levels 2-3
	2016-17 to 2022-23	ELA, math, science, social studies, writing [through 2021]	Levels 2-4
Utah	2013-14 to 2022-23	ELA, math, science	Levels 3-4
Vermont	2015-16 to 2021-22	ELA, math	Levels 3-4
	2018-19 to 2022-23	Science	Levels 3-4
Virginia	1997-98 to 2022-23	ELA, math, science, writing	Levels 2-3
	1997-98 to 2013-14	Social studies	Levels 2-3
Washington	2014-15 to 2022-23	ELA, math	Levels 3-4
	2018-19 to 2022-23	Science	Levels 3-4
West Virginia	2014-15 to 2022-23	ELA, math	Levels 3-4
	2018-19 to 2022-23	Science	Levels 3-4
Wisconsin	2015-16 to 2022-23	ELA, math, science	Levels 3-4
Wyoming	2013-14 to 2022-23	ELA, math, science	Levels 3-4

- a. In MA during SY 2014-15, some used an assessment with 4 proficiency levels (Legacy MCAS) and some districts began an assessment with 5 proficiency levels (PARCC).
- b. TX has 4 proficiency levels, but includes Levels 2-4 as “passing”, including the Level 2 category of “approaching proficiency.” This varies from other states.

Appendix F. Zelma Frequently Asked Questions (FAQs)

Background

1. What is Zelma?

Zelma Education is a comprehensive, interactive, AI-powered U.S. state assessment data repository that aims to make state assessment data more widely accessible and engaging for the general public. The repository includes all publicly-available assessment data from all 50 states and D.C. for students in Grades 3-8.

Zelma is an AI-powered research assistant that can help answer your assessment-related questions!

2. Who is Zelma for?

Everyone. Zelma is intended to make understanding these data accessible to anyone who can ask questions, which is everyone. Our target audience includes educators and education leaders, journalists and policymakers, parents, researchers and students. Visuals help to answer questions that individuals may have about particular states, districts, or schools, and data files are available for anyone who wants to use the data itself for further exploration.

3. Where do the data come from?

All data are from three primary sources. First, all proficiency data are sourced directly from State Education Agencies, from either their websites/data portals or via data request (no personal identifiable information is included). Second, Zelma integrates data from the National Center for Education Statistics (NCES) to provide information about district and school characteristics. Third, our data files have added “flags” to note for researchers if there have been changes in the subject-area assessment for a given year that might affect comparability over time.

4. Who created Zelma?

The Zelma research team is based at Brown University, led by Emily Oster and Clare Halloran. The website was created in collaboration with Novy.AI to develop the site’s design and AI features.

5. How can I create a Zelma.ai account?

Users may create their own Zelma account through our sign in page [here](#). Creating an account will allow users to save all queries.

Summative Assessments

6. What assessments does Zelma include?

Zelma includes data from annual statewide summative assessments. Zelma does not include other assessments administered by states, such as alternate assessments or English language proficiency assessments. Currently, Zelma only contains assessments administered in English.

7. My state adopted a new assessment recently. Can I still see prior achievement data?

Yes. Zelma includes all available achievement data for all available school years, regardless of changes. Changes in state assessments are noted in two ways for data users: 1) in the data files, there are variables that note if the subject-area assessment has changed either its name or how the state defines proficiency, and 2) all visuals produced by Zelma include detailed notes that indicate when states changed or administered new assessments in order to better inform interpretation.

8. How do states define proficiency?

All states define proficiency according to their state's unique grade-level and subject-area learning standards. For this reason, results are *NOT comparable across states*, as states administer distinct assessments.

9. How does Zelma account for changes in assessment administration?

Changes in state assessments that affect comparability over time are described below the figures as part of the "Notable Events." The Notable Events will explain if states have administered a new assessment in a given year, if there have been changes in proficiency cut scores, if an assessment has not been administered due to field testing, if participation rates were lower than a typical year, or if assessment data are not currently available. In the raw data files, variables indicate if there has been a change in the assessment name (for a given subject and year) or changes in the cut score criteria.

10. How is Zelma different from other sites that have state assessment data?

Zelma sets itself apart from other state assessment data sources for the following reasons:

- a. **Zelma provides users with the most up-to-date data publicly available.** Zelma provides updated assessment data promptly upon the public release from State Education Agencies for the most recent spring assessments. This means that individuals can quickly dive into the data to explore trends across districts, achievement gaps by student subgroups, and to explore trends across schools and districts by grade or student subgroup.
- b. **Zelma is highly user-friendly.** Zelma allows individuals to use natural language to ask questions and explore areas of interest, enabled by AI.
- c. **All data are easily accessible.** All state data are available to download via the [Codebook & Downloads](#) page. Individuals may select the state and year of interest and easily export as a .csv file.
- d. **Zelma offers broader subject-area coverage.** Existing data sources prioritize data for reading/ELA and math. However, State Education Agencies often provide academic achievement data across other subjects as well, such as science and social studies. We believe that integrating these subject areas captures other important areas of student learning.

Data Availability

11. How can I get notified about data updates?

We encourage all data users to subscribe to receive data updates via e-mail; you may do so at the bottom of the Codebook & Downloads page [here](#). Upon downloading Zelma data files or the full dataset, you will also have the opportunity to subscribe to receive information about new data updates as well.

12. What states are included as part of Zelma?

Data are available for all 50 U.S. states and the District of Columbia. However, states vary in the years and data components that are publicly available.

13. What subjects does Zelma include?

All states include data for English language arts (ELA) and math. To the extent that additional subjects are publicly available, Zelma also integrates data for science, reading (if different from ELA), writing, social studies, and STEM. However, not all states have publicly posted assessment data for subject areas other than ELA and math.

14. What school years are available as part of Zelma?

In Version 1.1, released June 2024, Spring 2023 assessment data are included for all 50 states and DC. Longitudinal data varies by state, with the first year of available data ranging from 1998 to 2018. The most recent year with available data is Spring 2023. No states administered assessments in Spring 2020 due to the COVID-19 pandemic; therefore, no data are included as part of Zelma for this school year. Users may view the years of availability for each state, by subject, via the [Codebook & Downloads](#) page.

15. What grades are available?

Zelma currently limits data to Grades 3-8.

16. What student groups are available in the data?

Currently, to the extent that states have published or shared student subgroup information, data are available by the following student groups:

- race/ethnicity
- gender
- English language status
- economic status
- disability status
- migrant status
- homeless enrolled status
- foster care status
- military connected status

17. How does Zelma define the student subgroups?

Each state uses its own criteria to define student subgroups such as race/ethnicity, gender, English learner status, and economic status. Zelma reflects each state's own data classifications for student subgroups. For this reason, results are NOT comparable across states, as states administer distinct assessments and may have unique state criteria for student classifications.

18. What variables are included in the data?

All state datasets include an identical set of variables, even if a state does not have data available for a particular variable. Specific NCES variables include: district and school NCES IDs, district and school state-assigned IDs, district and school types, indicators for virtual school or charter district, county names and IDs, and school-level information. Please see [Codebook & Downloads](#) for a complete list and description of all available variables.

19. Does Zelma include student-level information?

No, Zelma does not include any student-level data or "personal identifiable information" (PII). Aggregated data are only available at the state, district, or school level, depending on the state.

20. What kinds of schools are included in the datasets?

Zelma includes data from all U.S. public school districts (including charter schools) with available data from State Education Agencies for students in Grades 3-8 on the general summative assessments.

21. Will more data be added to Zelma?

Yes. Zelma will incorporate additional data: a) as it is released by State Education Agencies each year, b) as additional data is received via data request for missing components, and c) as NCES information is updated for new districts and schools. *The next scheduled data release for Zelma 2.0 is December 2024, which will include available assessment data from Spring 2024.* If there are data components you are interested in seeing in the future, please send your feedback to zelmadata@gmail.com.

Using Zelma.ai

22. What can I do with Zelma?

Zelma is a tool to ask questions about state test score data in the United States. The best way to use Zelma is to be specific about the following components of interest:

- State, district(s) or school(s) of interest
- Subject(s) of interest
- Time frame of interest (data over time? Or for a particular year?)
- Student groups of interest

For example, users can ask:

- *"Please show me math scores in Minnesota over time."*
- *"How do math test scores differ by race/ethnicity in New Jersey?"*

Users can also ask how districts within a state compare to others, or about which districts are the top performing in a certain subject:

- *"How do ELA scores compare in Brookline MA versus the rest of the state?"*
- *"What were the top 5 performing school districts in math in Mississippi in 2023?"*

The more detail that users can give Zelma, the better. And if she doesn't quite get it, try again! We're always improving.

23. Are proficiency data comparable across states?

No, the results are *NOT comparable across states*, as states administer distinct assessments. These tests are often designed to assess student progress on state-specific standards, and are reported based on state-specific definitions of proficiency. Rather, for each state, Zelma can provide insight into trends and outcomes across districts and schools.

24. How do I cite the data that I use?

Data files used to produce information for the general public such as articles, reports, or presentations, should cite the data source as follows (currently, the data posted to our site represent Version 1.1):

Citing Zelma data files

Raw Data Citation Format:

- Zelma Education (Version 1.1). 2024. *Zelma assessment data files*, [state name (years)]; [state name (years)]; [state name (years)]. Accessed at <https://www.zelma.ai/data> on Month DD, YYYY.

Raw Data Citation Example:

- Zelma Education (Version 1.1). 2024. *Zelma assessment data files*, Minnesota (2018 - 2023); Wisconsin (2018 - 2023). Accessed at <https://www.zelma.ai/data> on June 18, 2024.
- Zelma Education (Version 1.1). 2024. *Zelma assessment data files*, All States (2018 - 2023). Accessed at <https://www.zelma.ai/data> on June 18, 2024.

Citing queries

- **Query Citation Format:** Zelma Education (Version #). [year]. "[query title]." Accessed at [query link] on Month DD, YYYY.
- **Query Citation Examples:** Zelma Education (Version 1.1). 2024. "Math scores by race in New York over time." Accessed at <https://www.zelma.ai/q/JTS8BN29Y8> on June 18, 2024.

More information about our Terms of Service can be found [here](#).

25. What datasets are available to download?

All state data are available to download via the [Codebook & Downloads](#) page. Individuals may select the state and year of interest and easily export as a .csv file.

26. How can I merge the Zelma datasets with other federal or state datasets?

All data files include federal district and state identifiers (NCESDistrictID and NCESSchoolID), as well as state-assigned identifiers (StateAssignedDistrictID and StateAssignedSchoolID). These variables can be used to appropriately merge student assessment data with external datasets.

27. How do I share my queries?

There is a "Share" button available for every query in the upper-right corner. This button will allow you to share the query via X or for users to copy the query link.

28. What if I do not see anything appear for my query?

There are a few possible reasons that you may not see a visual appear for your query:

- **Zelma may need some more information.** Try asking the question in a different way - can the query be more specific (e.g., are you interested in a particular year or subject)?
- **The data may not be provided from states in a way that allows for visualizations.** In some cases, data are not available from states in a way that Zelma can use for visualization purposes. For example, Zelma relies on data disaggregated by grade level; states that only provide aggregated data for all of Grades 3-8 cannot be included here. In addition, Zelma relies on data related to *the number of students tested across the state, district, and schools*. When states omit the number of students tested, Zelma is not able to use proficiency data

for visualization purposes. However, proficiency rates may still be available in state files via the [Codebook & Downloads](#) page.

- **The data file may not be available.** There are several reasons that states may not have available assessment data for a given subject in a given year: the state may have only administered a field test in a given year, the state may have received a federal waiver from testing in 2021 due to the pandemic, the state may not post all assessment data online, or the state may have stopped administering an assessment.

You may always contact us at zelmadata@gmail.com and we are happy to look into any questions you have about your queries.

AI-powered Zelma

29. What kind of output can Zelma produce?

Zelma can produce line graphs, bar graphs, pie charts and tables.

30. How does Zelma work?

All state assessment data files have been re-formatted by Zelma's team using a standardized set of variable names. Zelma draws from these files to respond to user queries; in other words, Zelma cannot use any external data. To be fully transparent with how Zelma draws from the data files, each query also displays the associated SQL code to explain the commands used to generate the response. Zelma is always learning - and improving!

31. How does Zelma handle data suppression?

In order to protect personal identifiable information (PII), states have regulations in place to limit the amount of information shared publicly. For example, many states omit proficiency data if a given subgroup of students includes 10 or fewer students. In the raw data files, these cases of data suppression are designated with an asterisk (*), while missing data is designated as --. Zelma omits these cases from all visualizations.

In addition, states may present a data range rather than a specific number for a given variable. In the raw data files, the data ranges are preserved. For the purposes of the data visualizations, Zelma replaces the data range with the mid-point. For example, a state may indicate that 20-30 students completed the ELA assessment in a given school; for visualization purposes Zelma replaces this as 25.

Contacting Zelma

32. How do I contact Zelma?

All questions, comments, or corrections can be sent to our team at zelmadata@gmail.com. Thank you for your feedback!