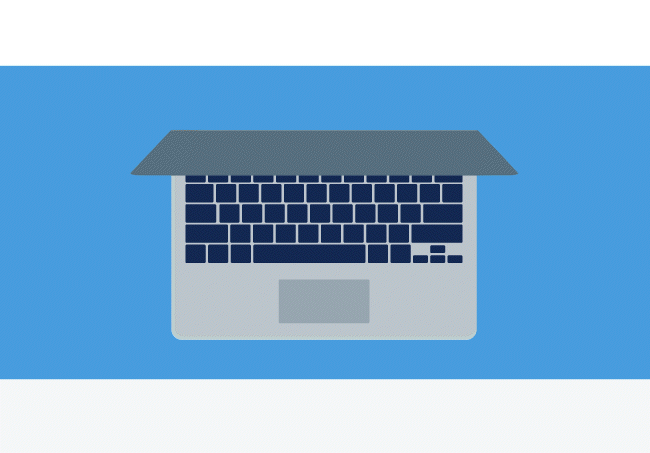
International Data Explorer Help Guide

NCES IDE Team

2024-01-11

Table of Contents

# International Data Explorer Help Guide



# 1 Overview of the International Data Explorer (IDE)

## 1.1 What is the IDE?

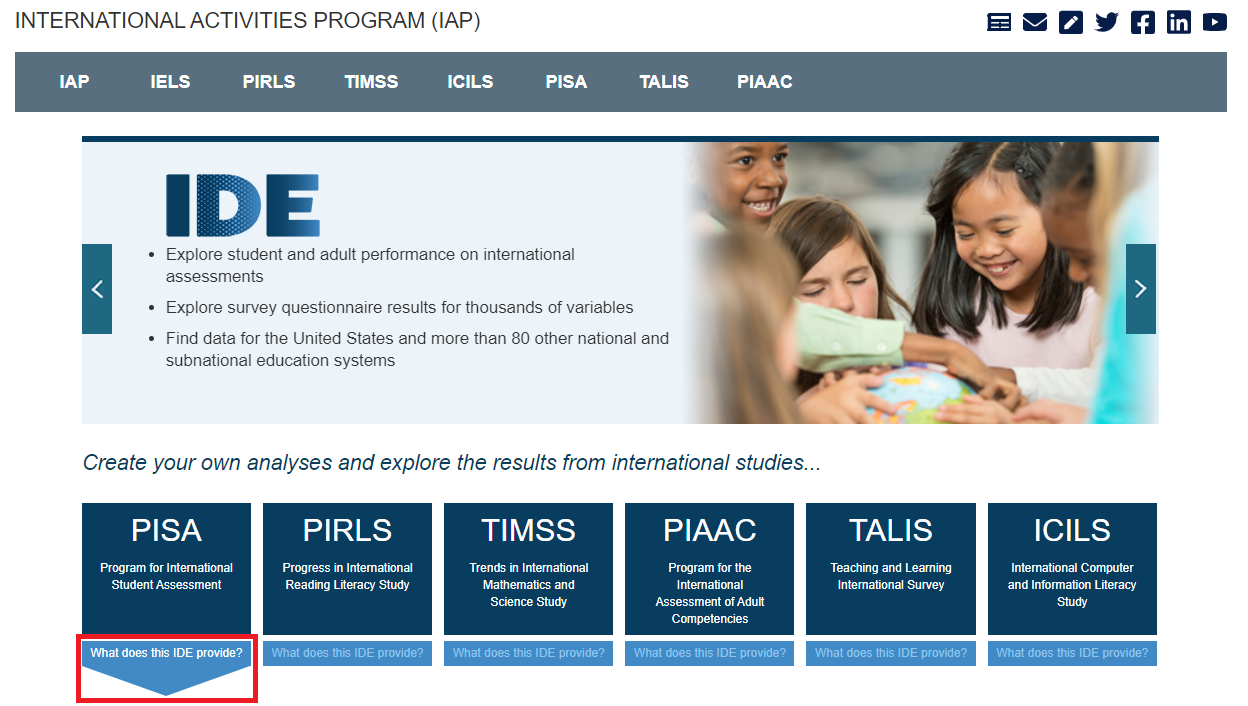
The National Center for Education Statistics (NCES) has made it easy to explore and analyze large-scale international education study data via various tools, including the international Data Explorer (IDE). The IDE is an interactive online tool with data from the following studies:

* [Program for International Student Assessment (PISA)](https://nces.ed.gov/surveys/pisa/)
* [Progress in International Reading Literacy Study (PIRLS)](https://nces.ed.gov/surveys/pirls/index.asp)
* [Trends in International Mathematics and Science Study (TIMSS)](https://nces.ed.gov/timss/)
* [Program for the International Assessment of Adult Competencies (PIAAC)](https://nces.ed.gov/surveys/piaac/index.asp)
* [Teaching and Learning International Survey (TALIS)](https://nces.ed.gov/surveys/talis/)
* [International Computer and Information Literacy Study (ICILS)](https://nces.ed.gov/surveys/icils/)

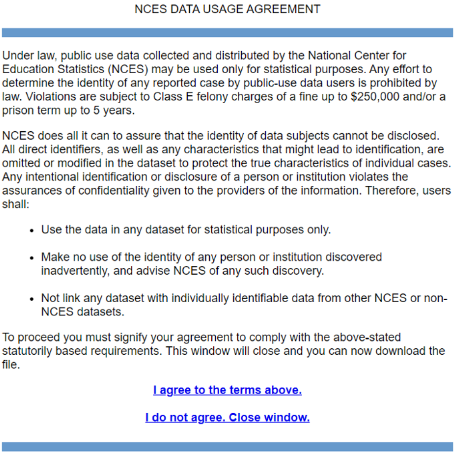
The IDE provides a wide set of functions, including:

* Explore student and adult performance on international assessments
* Explore survey questionnaire data for thousands of variables
* Find data for the United States and more than 80 foreign education systems
* Create tables, charts, and maps
* Calculate averages, percentages, standard deviations, percentiles, and performance/proficiency levels
* Run statistical tests, including gap analyses

## 1.2 How do I access the IDE?

1. Visit the IDE home page <https://nces.ed.gov/surveys/international/ide/>.
2. Select the assessment or survey of interest and start exploring. View “What does this IDE provide?” under each study to learn more 

## 1.3 NCES data usage agreement

In order to be able to enter the IDE, it is necessary to review and agree to the terms of the NCES Data Usage Agreement. You may see the following screen when visiting each of the IDE page from the IDE home page. After you click on the button agreeing to the terms, you will be taken to the Data Explorer. If you do not do this, you will be redirected to the IDE home page. 

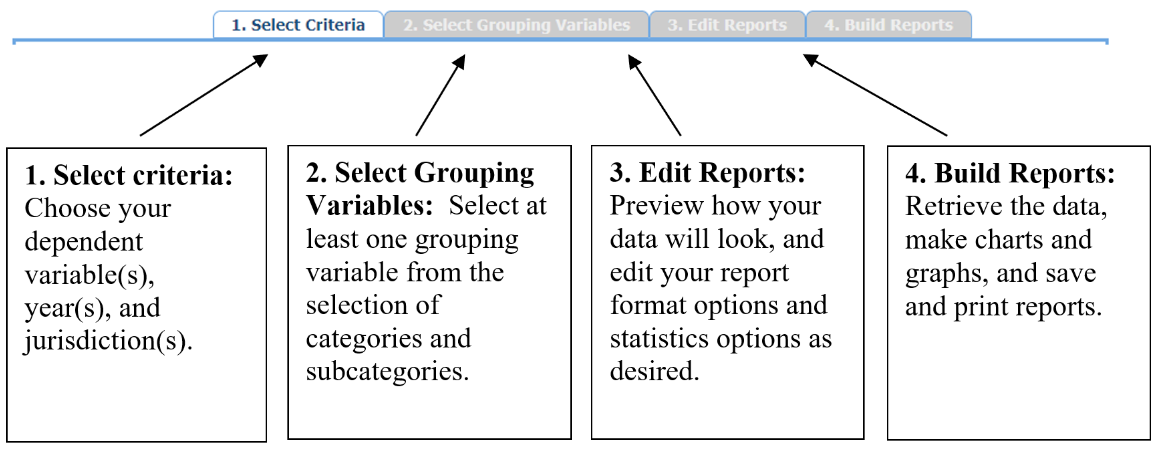
## 1.4 Computer requirements for the IDE

The IDE performs best under the following requirements.

* Screen resolution should be 1024 x 768 pixels or higher.
* Browsers: Google Chrome, Apple Safari, Internet Explorer (IE) version 10 or higher, FireFox 3.0 or higher.
* Enable JavaScript and pop-ups in your browser.
* The TIMSS IDE requires Flash version 9.0.115 or higher (download Adobe Flash Player at <http://get.adobe.com/flashplayer/>).
* Exports of files to Microsoft Office can be opened with Office 2003 or later.
* Exports of files to PDF can be read with Adobe Acrobat Reader.
* Screen reader software should be JAWS 8.0 or higher.

# 2 IDE Walkthrough

There are four general steps for exploring each IDE page (see exhibit 3). Each step is described in more detail in the following sub-sections.

Exhibit 3. What you will see in the IDE environment and what each step entails 

The following subsections covers the steps in details to explore data in the IDE. The IDE process is similar across all studies, so this guide primarily uses the screenshots from the PISA IDE by default. However, relevant screenshots from the other IDEs are included as needed to illustrate differences between studies. By following the instructions below, users can explore the data from any of IDEs.

## 2.1 Select Criteria

### 2.1.1 Overview

Your data query in the IDE begins on the Select Criteria screen (see exhibit 2). Choose one *Subject*, one *Grade*, and one or more *Measures*, *Years*, and *Jurisdictions* for the data you wish to view or compare. Use the *Reset* button, located in the upper-right portion of the screen (just below the *Help* button), to cancel your selections and begin again. Click on a red sideways-facing arrow (►) to open up a category and click on a red downward-facing arrow (▼) to close a category.

### 2.1.2 Choose Subject

Under *Subject*, you have the choice of *Mathematics and Science*, *TIMSS Advanced: Advanced Mathematics*, or *TIMSS Advanced: Physics*.

### 2.1.3 Choose Year

At the top of the *Measure* and (Jurisdiction) sections, you have the choice of selecting 2019, 2015, 2011, 2007, 2003, 1999, and/or 1995 by checking the appropriate box. To include data from all years, check the “All Years” box to the left of the individual years. Mathematics and science data are available for 2019, 2015, 2011, 2007, 2003, 1999, and 1995. In 1999, no data for grade 4 were collected in mathematics or science. Advanced mathematics and physics data are available for 2015 only.

### 2.1.4 Choose Measure

After choosing a subject, you can choose between the overall scale and/or any of the subject’s subscales. The overall scale and subscales can be used for trend analyses across years where applicable.

In addition, there are a number of continuous variables other than scale scores that you may choose as a measure of analysis. These variables fall under different categories, such as “Student and Family Characteristics” and “Teacher Background Characteristics, Formal Education, and Training” and include variables such as age, teaching experience, and class size.

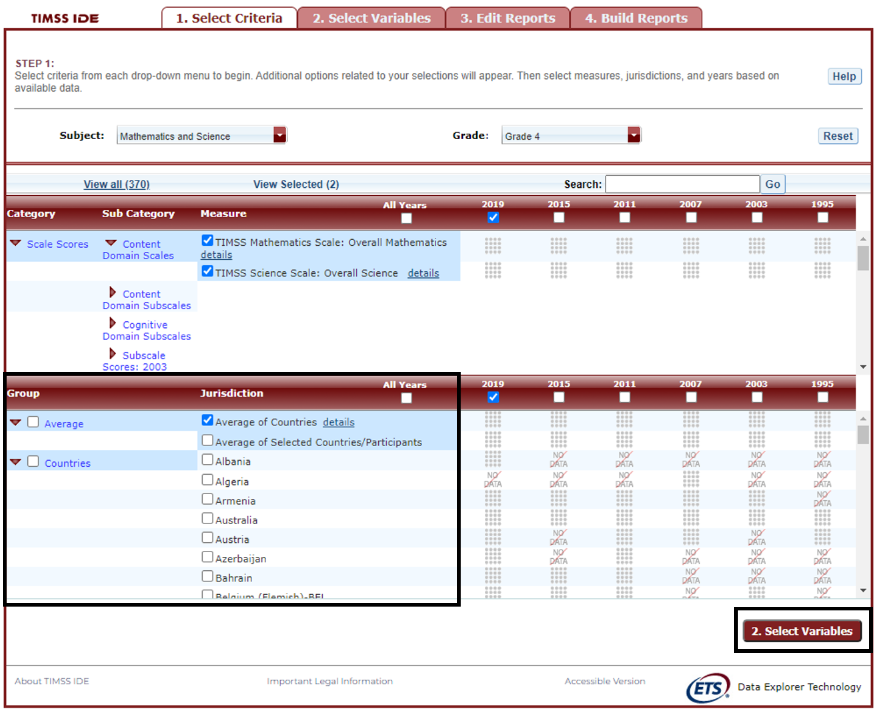
### 2.1.5 Choose Jurisdiction

With your\* Measure(s)\* and *Year(s)* selected, next choose at least one *Jurisdiction*.

Jurisdictions are found under the following groups: *Countries*, *U.S. Jurisdictions*, and *Benchmarking Participants*. There is also a group category called *Average*, with options to display the *Average of Countries* and the *Average of Selected Countries/Participants*. *Average of Countries* displays the average statistic for all available jurisdictions under the “*Countries*” group, except when “*All students*” is selected at step 2, in which case *Average of Countries* displays the TIMSS scale centerpoint of 500.

The general procedures for selecting one or more jurisdictions are as follows:

1. To open or close jurisdictions, click on the arrow. Jurisdictions in the group are open and can be selected when the red arrow points down (see exhibit 3).
2. Click the checkboxes next to the specific jurisdictions that you are interested in, or uncheck those jurisdictions that you wish to deselect. If you click the checkbox next to the group name (e.g., “Countries”), you will select all the jurisdictions within that group. If desired, uncheck the group name to deselect all.
3. If you want to close a group (e.g., close the list of countries in order to readily see the benchmarking participants), click the red arrow next to the group name. The closed group’s arrow points to the right. Be advised that closing the group will not deselect your choices.

Exhibit 3. Choosing jurisdictions 

To continue in the IDE, click the *Select Variables* button at the bottom right of the page or the tab at the top of the page to go to the next screen (see exhibit 3).

## 2.2 Select Variables

### 2.2.1 Overview

### 2.2.2 Search Using Category and Sub Category Lists

### 2.2.3 Search Function

## 2.3 Edit Reports

### 2.3.1 Overview

### 2.3.2 Preview Report

### 2.3.3 Edit Report

### 2.3.4 Create New Variables

### 2.3.5 Create New Report

### 2.3.6 Format Options

### 2.3.7 Statistics Options

### 2.3.8 Select Reports to Build

## 2.4 Build Reports

### 2.4.1 Overview

### 2.4.2 View Reports as Data Tables

### 2.4.3 Charts

### 2.4.4 Create Charts—Chart Options

### 2.4.5 Significance Tests

### 2.4.6 Gap Analysis

### 2.4.7 Regression Analysis

### 2.4.8 Export Reports

# 3 IDE definitions by study

This section provides an overview of the definitions in each IDE, including the kinds of criteria and variables that are used to form data queries, as well as the kinds of data that are available and the statistical methods used to assess them. Each of the subsections below covers the same topics and provides a detailed description of the definitions used in each IDE study.

* Criteria
  + Subject
  + Grade
  + Years
  + Measures
  + Jurisdictions
* Variables
  + Proficiency levels or benchmarks
  + Index Variables
* Statistics options
  + Averages
  + Percentages
  + Standard deviations
  + Percentiles
* Cross-tabulations
* Statistical notations and other notes

## 3.1 PISA IDE

### 3.1.1 Criteria

Each data query must include at least one selection from five criteria choices: language, subject, year(s), measure(s), and jurisdiction(s). Shown below is an outline of these selection criteria followed by a brief description.

1. Language:
   * English
   * Spanish
2. Subject:
   * Science literacy
   * Reading literacy
   * Mathematics literacy
   * Financial literacy
   * Problem solving
   * Collaborative problem solving
3. Year:
   * 2018 (data available for reading, reading subscales, mathematics, science, and financial literacy)
   * 2015 (data available for science, science subscales, reading, mathematics, financial literacy, and collaborative problem solving)
   * 2012 (data available for mathematics, mathematics subscales, reading, science, financial literacy, and problem solving)
   * 2009 (data available for reading, reading subscales, mathematics, and science)
   * 2006 (data available for reading, mathematics, science, and science subscales)
   * 2003 (data available for reading, mathematics, mathematics subscales, and science)
   * 2000 (data available for reading, reading subscales, mathematics, and science)
4. Measure:
   * Mathematics scale: Overall mathematics
   * Reading scale: Overall reading
   * Science scale: Overall science
   * Mathematics subscale: Employ
   * Mathematics subscale: Formulate
   * Mathematics subscale: Interpret
   * Mathematics subscale: Space and shape
   * Mathematics subscale: Change and relationships
   * Mathematics subscale: Quantity
   * Mathematics subscale: Uncertainty
   * Reading subscale: Locate information
   * Reading subscale: Understand
   * Reading subscale: Evaluate and reflect
   * Reading subscale: Access and retrieve
   * Reading subscale: Integrate and interpret
   * Reading subscale: Reflect and evaluate
   * Reading subscale: Continuous text
   * Reading subscale: Noncontinuous text
   * Science subscale: Identifying scientific issues
   * Science subscale: Explaining phenomena scientifically
   * Science subscale: Using scientific evidence
   * Science competency subscale: Evaluate and design scientific enquiry
   * Science competency subscale: Explain phenomena scientifically
   * Science competency subscale: Interpret data and evidence scientifically
   * Science knowledge subscale: Content Knowledge
   * Science knowledge subscale: Procedural and Epistemic Knowledge
   * Science system subscale: Earth and space
   * Science system subscale: Living systems
   * Science system subscale: Physical systems
   * Attitude scale: Interest in science
   * Attitude scale: Support for scientific inquiry
   * Financial literacy scale
   * Problem-solving scale
   * Collaborative problem-solving scale
5. Jurisdiction:
   * International average (OECD countries)
   * Average of the selected jurisdictions
   * OECD
   * Non-OECD
   * U.S. states

#### 3.1.1.1 Language

The PISA IDE currently provides the option to view all steps of the IDE and build reports in English or Spanish. The Help Guide currently is only offered in English.

#### 3.1.1.2 Subject

PISA assesses reading literacy, mathematics literacy, and science literacy at each administration. In addition, the IDE contains data from the administration of the 2012, 2015, and 2018 PISA financial literacy assessments, the 2012 PISA problem-solving assessments, and the 2015 PISA collaborative problem-solving assessment.

#### 3.1.1.3 Measures

The PISA IDE includes measures for each subject when selected, such as an overall scale, subscales (if applicable), and continuous variables.

Although each administration of PISA assesses mathematics, reading, and science, one of these subjects is assessed in depth in each administration. You can choose between the overall scale and/or any of the subject’s subscales as your measure. However, subscales are only available for a subject area in the years in which it was the major domain. The major subject area assessed in 2000 was reading literacy; in 2003, mathematics literacy; and in 2006, science literacy. The cycle fully repeated itself in 2009 and began again in 2018. Subscales are constituent parts of the major overall subject scale for an assessment and are specified by the PISA assessment frameworks. In the years when a subject area is a minor domain, only an overall scale is available, and it is based on a set of items of varying difficulty that represent the range of topics covered by the full assessment. Only overall scale scores are reported in the IDE for financial literacy, problem solving, and collaborative problem solving. Please see Section I. Background, for more information.

In 2015 and 2006, science was the major domain, and reading and mathematics were minor domains. Therefore, for these years, subscales are only available for science data; only single composite scales are available for PISA reading and mathematics data.

In 2012 and 2003, mathematics was the major domain, and reading and science were minor domains. Therefore, for these years, subscales are only available for mathematics data; only single composite scales are available for PISA reading and science data.

In 2018, 2009, and 2000, reading was the major domain, and mathematics and science were minor domains. Therefore, for these years, subscales are only available for reading data; only single composite scales are available for PISA mathematics and reading data.

In addition, there are continuous variables other than scale scores that you may choose as a measure of analysis. These variables fall under different categories, such as Student and Family Characteristics and School and Classroom Climate, and include variables such as student age in years, size of class, and an index of computer availability.

#### 3.1.1.4 Years

Currently, data availability in the IDE is dependent on the measure selected. If the measure chosen is an overall literacy scale, you can choose one or multiple years: 2018, 2015, 2012, 2009, 2006, 2003, and 2000. If the measure chosen is one of the science subscales, you can choose 2015 and/or 2006. If you choose between any of the mathematics subscales, you can choose 2012 and/or 2003. If you choose any of the reading subscales, you can choose 2018, 2009, and/or 2000. Subscales are not available for financial literacy, problem solving, or collaborative problem solving.

#### 3.1.1.5 Jurisdictions

All listed jurisdictions can be selected for any analyses, provided data are available for the selected year. In 2018, a total of 79 jurisdictions participated in the mathematics, reading, and science literacy PISA assessments: 37 Organization for Economic Cooperation and Development (OECD) countries and 42 non-OECD jurisdictions. The non-OECD jurisdictions include some subnational education systems, such as Hong Kong-China. Data are not available for some of these 79 jurisdictions for 2015, 2012, 2009, 2006, 2003, and/or 2000, either because they did not participate in that PISA cycle or because their data were suppressed due to reporting standards not being met (for example, PISA 2018 data for Vietnam were suppressed due to international reporting standards not being met, and PISA 2015 data for Argentina, Malaysia, and Kazakhstan were suppressed due to international reporting standards not being met).

Data are available for 73 jurisdictions (35 OECD and 38 non-OECD) in 2015, 65 jurisdictions (35 OECD and 30 non-OECD) in 2012, 65 jurisdictions (35 OECD and 30 non-OECD) in 2009, 57 jurisdictions (35 OECD and 22 non-OECD) in 2006, 41 jurisdictions (31 OECD and 10 non-OECD) in 2003, and 38 jurisdictions (29 OECD and 9 non-OECD) in 2000.

Also included in the IDE are the 5 U.S. states or territories that participated in PISA 2012 and PISA 2015. Data from the 43 jurisdictions that participated in the administration of the problem-solving assessment in 2012 are included in the IDE, as well as the 51 jurisdictions that participated in the 2015 collaborative problem-solving assessment. All jurisdictions that participated in the financial literacy assessment in at least one year (2012, 2015, 2018) are included in the IDE.

Jurisdictions for which data are not available for a selected year are identified by the icon representing “no data”. Note that the IDE contains a few U.S.-specific background variables (e.g., race/ethnicity) that, when selected, will not yield information for any other jurisdictions.

Jurisdictions listed in the IDE as OECD countries are those that are currently members of the OECD. In some cases, countries which are current members of the OECD were not members during a prior administration or release of PISA. For example, Latvia was an OECD country at the time of the 2015 PISA release, but not during earlier PISA cycles. The IDE recalculates OECD averages for previous PISA cycles based on the current count of 37 OECD countries as of the 2018 release. Please note that the recalculation of the OECD average based on the current count explains why OECD averages calculated by the IDE for earlier years (e.g., 2015 or 2012) do not match the OECD averages from OECD and NCES reports published in earlier years.

### 3.1.2 Variables

In the PISA IDE, questions from two types of questionnaires (student and school), as well as variables that are derived from background information, are organized into categories that have shared characteristics and can be selected as a group when examining and generating tables.

Content category and subcategory titles may overlap, but specific variables appear only once in a subcategory. Use Search in the Select Variables step to locate variables.

Note that some variables might be similar in content, but not comparable over the years, either due to differences in the question asked or differences in their response categories. For example, an index variable for students’ family structure is available for 2012, 2009, 2003, and 2000. Each index variable is based on students’ responses to the same question asking who usually lived at home with them. However, these three variables (STP5437 in 2012, FAMSTR09 in 2009 and FAMSTR00 in 2003 and 2000) are not comparable due to differences in response categories. In 2012, the response categories were “single-parent (natural or otherwise),” “two parents (natural or otherwise),” and “other”; in 2009, the categories were “single-parent family,” “two-parent family,” and “other”; and in 2003 and 2000, the categories were “single-parent family,” “two-parent family,” “mixed,” and “other.” The icon representing “no data”— —will help in identifying the year for which the variable has data available for analysis.

#### 3.1.2.1 Proficiency levels

Achievement results for PISA are reported using discrete proficiency levels for reading, mathematics, science, financial literacy, problem solving, and collaborative problem solving. Increasing levels represent the knowledge, skills, and capabilities needed to perform tasks of increasing complexity. Based on the statistics option chosen, IDE can report the average scores of students at each proficiency level or the percentage of students performing at each of the predefined levels for the chosen jurisdictions. Two statistics options, standard deviations and percentiles, will not generate reports as proficiency levels are not reportable using these statistical analyses. Proficiency levels for any subject should be analyzed with the scale of that same subject; for example, the reading literacy proficiency levels should be analyzed with the reading literacy scale.

Mathematics literacy: Administered in all cycles (2000, 2003, 2006, 2009, 2012, 2015, 2018). In 2000, an interim scale was used, and cut-off points between mathematics literacy proficiency levels were not established. Thus, proficiency levels cannot be analyzed in the IDE for 2000 mathematics literacy. From 2003 to 2018, mathematics literacy results were reported using 6 proficiency levels (level 1–level 6); the IDE shows 7 categories (below level 1, level 1, level 2, level 3, level 4, level 5, level 6).

Science literacy: Administered in all cycles (2000, 2003, 2006, 2009, 2012, 2015, 2018). Proficiency levels did not have strict definitions until 2006, when science literacy was the major domain and had a non-interim scale for the first time. Thus, proficiency levels cannot be analyzed in the IDE for 2000 and 2003 science literacy. In 2006, 2009, and 2012, science literacy results were reported using 6 proficiency levels. In 2015 and 2018, science literacy results were reported using 7 proficiency levels, as level 1 was broken into level 1b and level 1a. The cutpoint score for level 1a in 2015 and 2018 is the same as level 1 in 2006, 2009, and 2012; the cutpoint score for level 1b is set significantly lower. IDE programmers have retroactively calculated level 1b for 2006, 2009, and 2012 to allow for trend comparisons, so the IDE shows 8 categories for all years (below level 1b, level 1b, level 1a, level 2, level 3, level 4, level 5, level 6).

Reading literacy: Administered in all cycles (2000, 2003, 2006, 2009, 2012, 2015, 2018). In 2000, 2003, and 2006, 5 proficiency levels were used (level 1–level 5). Starting in 2009 and continuing in 2012 and 2015, reading literacy results were reported using 7 proficiency levels, with level 1 broken into level 1b and level 1a, followed by level 2 through 5 and a new top level (level 6). In 2018, a new lowest proficiency level (level 1c) was added, and the full list of 8 reading literacy proficiency levels became level 1c, level 1b, level 1a, level 2, level 3, level 4, level 5, and level 6. The cut point for level 1a from 2009 to 2018 is the same as for level 1 from 2000 to 2006. IDE programmers have retroactively calculated level 1c for pre-2018 years and level 1b and level 6 for pre-2009 years to allow for trend comparisons, so the IDE shows 9 categories for all years (below level 1c, level 1c, level 1b, level 1a, level 2, level 3, level 4, level 5, level 6).

Financial literacy: Administered in 2012, 2015, and 2018. In all 3 years, financial literacy results were reported using 5 proficiency levels (level 1–level 5); the IDE shows 6 categories (below level 1, level 1, level 2, level 3, level 4, level 5). Problem solving: Administered in 2012. Problem-solving results were reported using 6 proficiency levels (level 1–level 6); the IDE shows 7 categories (below level 1, level 1, level 2, level 3, level 4, level 5, level 6).

Collaborative problem solving: Administered in 2015. Collaborative problem solving results were reported using 4 proficiency levels (level 1–level 4); the IDE shows 5 categories (below level 1, level 1, level 2, level 3, level 4).

As noted above, the IDE also provides available data for students performing below proficiency level 1 for mathematics literacy, financial literacy, problem solving, and collaborative problem solving; below level 1b for science; and below level 1c for reading literacy. Patterns of responses for students in the proficiency levels below each subject’s lowest level (e.g., below level 1 for mathematics literacy, below level 1c for reading literacy, etc.) suggest that these students are unable to answer at least half of the items from those levels correctly; for this reason, the cognitive capabilities of students scoring below these levels are unclear and not defined by OECD. Proficiency at and below these low levels is sometimes combined in reports and referred to as below level 2 (e.g., for reading literacy, below level 2 refers to levels 1a, 1b, lc, and below level 1c.) Descriptions that characterize typical student performance at each proficiency level are shown in the following tables for reading, mathematics, and science literacy, as well as financial literacy, problem solving, and collaborative problem solving.

For more information on benchmarks, please visit <https://nces.ed.gov/surveys/pisa/2018technotes-6.asp>.

#### 3.1.2.2 Index Variables

In addition to scale scores representing performance in various subjects, PISA uses indices derived from the student, parent, teacher, and school questionnaires to contextualize PISA results or to estimate trends that account for demographic changes over time.

Information on indices for each year of administration can be found in the chapters referenced in the summary table below. The PISA technical reports can be found on the OECD PISA publications page (<http://www.oecd.org/pisa/publications/>).

| Year of PISA administration | PISA technical report | Links |
| --- | --- | --- |
| 2018 | Chapter 16 | [Link](http://www.oecd.org/pisa/data/33688233.pdf) |
| 2015 | Chapter 16 | [Link](http://www.oecd.org/pisa/data/pisa2003technicalreport.htm) |
| 2012 | Chapter 16 | [Link](http://www.oecd.org/pisa/data/42025182.pdf) |
| 2009 | Chapter 16 | [Link](http://www.oecd.org/pisa/data/pisa2009technicalreport.htm) |
| 2006 | Chapter 16 | [Link](http://www.oecd.org/pisa/data/pisa2012technicalreport.htm) |
| 2003 | Chapter 17 | [Link](http://www.oecd.org/pisa/data/2015-technical-report/) |
| 2000 | Chapter 17 | [Link](http://www.oecd.org/pisa/data/pisa2018technicalreport/) |

### 3.1.3 Statistics Options

The IDE reports PISA data with several statistics options:

• Averages • Percentages • Standard deviations • Percentiles

#### 3.1.3.1 Averages

This statistic provides the average value for a selected continuous variable or overall score for the combined literacy scale (for example, science literacy) or score for one of the subscales corresponding to the subject chosen (for example, the science competency subscale: interpret data and evidence scientifically).

For the PISA assessment, student performance is reported on scales that range from 0 to 1,000. PISA scales are produced using item response theory (IRT) to estimate average scores for mathematics, reading, science, financial literacy, and problem solving for each jurisdiction. IRT identifies patterns of response and uses statistical models to predict the probability of answering an item correctly as a function of the students’ proficiency in answering other questions. That is, student responses to the assessment questions are analyzed to determine the percentage of students responding correctly to each multiple-choice question and the percentage of students achieving each of the score categories for constructed-response questions.

#### 3.1.3.2 Percentages

This statistic shows the percentage of students as a row percentage. For example, if a categorical variable is selected and the jurisdictions are listed in the table stub, the percentage data for the response categories will sum to 100 percent in each jurisdiction. By default, the percentage distributions do not include missing data, although there is an option to include them.

#### 3.1.3.3 Standard deviations

The standard deviation is a measure of how widely or narrowly dispersed scores are for a particular dataset. Under general normality assumptions, 95 percent of the scores are within two standard deviations of the mean. For example, if the average score of a dataset is 500 and the standard deviation is 100, it means that 95 percent of the scores in this dataset fall between 300 and 700. The standard deviation is the square root of the variance.

#### 3.1.3.4 Percentiles

This statistic shows the threshold score (or cut point) for the following:

• 10th percentile – the bottom 10 percent of students • 25th percentile – the bottom quarter of students • 50th percentile – the median (half the students scored below the cut point and half scored above it) • 75th percentile – the top quarter of students • 90th percentile – the top 10 percent of students

### 3.1.4 Cross-tabulations

Cross-tabulation is a method of combining separate variables into a single table. Normally, each variable has its own table. If you have selected two or three variables (not counting *All students*) and when you go to the *Edit Reports* step, you will automatically get a list with one table for each variable (including one for *All students*); at the end of that list you will get one cross-tabulation for the two or three variables selected.

If you have chosen four or more variables (not counting *All students*), you will get tables for each variable, but you won’t get the cross-tabulation. Be advised that if you go back to add another variable without subtracting one to keep the total under four, you will lose any edits you might have made to the cross-tabulation.

### 3.1.5 Statistical Notations and Other Notes

Statistical notations and other notes are found at the end of a data table, as applicable to that table:

* — Not available.
* † Not applicable. (For instance, the standard error for the statistic cannot be reported because the statistic does not meet reporting standards.)
* # The statistic rounds to zero.
* ‡ Reporting standards not met. (For instance, the sample size is insufficient to permit a reliable estimate.)
* NOTE: A general note pertains to any special characteristics of the data in the table.
* SOURCE: Source information is listed for all PISA data and should be cited when data are used in a publication or presentation.

#### 3.1.5.1 Calculation of OECD averages

The IDE generates the OECD average for the selected measures and variables if “International Average (OECD Countries)” is clicked under “Jurisdiction.”

Jurisdictions listed in the IDE as OECD countries are those that are currently members of the OECD. In some cases, countries which are current members of the OECD were not members during a prior administration or release of PISA. For example, Latvia was an OECD country at the time of the 2015 PISA release, but not during earlier PISA cycles. The IDE recalculates OECD averages for previous PISA cycles based on the current count of 35 OECD countries as of the 2015 release. Please note that the recalculation of the OECD average based on the current count explains why OECD averages calculated by the IDE for earlier years (e.g., 2012 or 2009) do not match the OECD averages from OECD and NCES reports published in earlier years.

Furthermore, there are certain OECD countries that are excluded from the OECD averages both in the IDE and published OECD reports due to issues listed below:

• Four current OECD countries (Estonia, the Slovak Republic, Slovenia, and Turkey) did not participate in 2000 and 2003. • Data for the Netherlands and the United Kingdom were suppressed in 2000 due to international reporting standards not being met.  
• The reading literacy scores are not reported in the 2006 cycle for the United States due to a printing error in the test booklets.[[1]](#footnote-102) • The OECD average for the optional financial literacy assessment is calculated based on the average scores of the 14 participating countries in 2012. • The OECD average for the optional problem-solving assessment is calculated based on the 28 participating countries in 2012. • Data for Vietnam were suppressed in 2018 due to international reporting standards not being met. • The reading literacy scores for Spain were not reported in 2018 due to sub-optimal response behaviors from students.

Please note that OECD averages are affected by data suppression rules (discussed on the next page). This means that in some cases the OECD average generated by the IDE when a variable is chosen may not match the PISA 2018 OECD and NCES reports for that variable. This occurs when an OECD country’s data is suppressed in either the IDE or the OECD or NCES reports, but not both. If a country’s data is suppressed in the IDE, it will not be included in the calculation of the average score. For example, the OECD excluded Spain’s reading data from its first report presenting the results of the PISA 2018 survey (OECD, PISA 2018 Results (Volume I): What Students Know and Can Do, available at <http://www.pisa.oecd.org>) because of a concern over sub-optimal response behaviors from students. NCES also excluded these data from its 2018 report. After further investigation, the OECD decided to release all available PISA 2018 data for Spain, but this change was not reflected in the NCES report.

#### 3.1.5.2 Statistical Comparisons

Comparisons of achievement across years are made using independent t-tests with a linking error taken into account. Comparisons between jurisdictions are also treated as independent. As of December 2016, all comparisons within a jurisdiction, within the same year, are made using dependent t-tests. Prior to this, only male-female comparisons within a jurisdiction were treated as dependent. Because of this change, the results of statistical significance testing may differ slightly from the results obtained using earlier versions of the PISA IDE. The alpha level for all t-tests is .05.

#### 3.1.5.3 Data Suppression

Data suppression may be handled slightly differently in the PISA IDE and the OECD PISA International Reports. For the IDE, the Rule of 62 is applied to suppress data to avoid reporting results for groups about which little of interest could be said due to lack of power. The Rule of 62 is borrowed from the IDE’s counterpart, the National Assessment of Educational Progress (NAEP) Data Explorer (NDE). This rule states that statistics for a group are suppressed if they are based on less than 62 cases. These statistics are means, standard errors, standard deviations, and a set of percentiles. The rule serves to assure a minimum power requirement to detect moderate differences at nominal significance level (.05). The minimum power is 0.80 and the moderate effect size is 0.5 standard deviation units. A design effect of 2 is assumed to derive an appropriate complex sample standard deviation.

## 3.2 PIRLS IDE

### 3.2.1 Criteria

Each data query must include at least one selection from three criteria choices: measure(s), year(s), and jurisdiction(s). Shown below is an outline of these selection criteria followed by a brief description.

1. Subject
   * PIRLS
   * ePIRLS
2. Measure
   * Scale scores
     1. PIRLS
        + PIRLS Reading Scale: Combined Reading
        + PIRLS Reading Scale: Acquire and Use Information
        + PIRLS Reading Scale: Literary Experience
        + PIRLS Reading Scale: Interpreting, Integrating, and Evaluating
        + PIRLS Reading Scale: Retrieving and Straightforward Inferencing
     2. ePIRLS
        + ePIRLS Reading Scale: Online Informational Reading
        + ePIRLS Reading Scale: Online Interpreting, Integrating, and Evaluating
        + ePIRLS Reading Scale: Online Retrieving and Straightforward Inferencing
   * Student and Family Characteristics
   * Student Perception of Reading
   * Student Perception of School
   * Student Characteristics (Teacher)
   * English Language and Reading Instruction (Teacher)
   * Class Resources (Teacher)
   * Teacher Characteristics
   * School Characteristics
   * School Instruction Time
   * Curriculum (School)
   * Reading Instruction (School)
   * School Resources
   * School Climate and Behavior Problems
   * Teacher Collaboration
   * Principal Characteristics
3. Jurisdiction
   * Average of Countries
   * Average for Selected Countries/Participants
   * Countries
   * Benchmarking Participants
   * Off-Grade Participants
4. Years
   * 2001
   * 2006
   * 2011
   * 2016
   * All Years

#### 3.2.1.1 Subject

PIRLS is a study of the reading literacy, and ePIRLS is a study of online informational reading. Both are subjects that can be selected in the IDE.

#### 3.2.1.2 Measures

PIRLS focuses on overall reading literacy, but within this broad category, four subscales are available: two focusing on the purposes of reading (literary experience and acquire and use information) and two focusing on the processes used for reading (interpreting, integrating, and evaluating and retrieving and straightforward inferencing). The 2001 and 2006 reading subscales have been rescaled to allow for comparisons to 2011 and later years. Subscales are constituent parts of the composite subject scale for an assessment and are specified by the assessment framework. The weighted average of these is the basis for the reading composite scale, as described in the PIRLS framework. Subscales are based on fewer observations than the combined scale and, as a result, may have larger standard errors. ePIRLS which focuses on online informational reading does not include subscales focusing on the purposes of reading, since the entire assessment focuses on reading to acquire and use information. ePIRLS does include two subscales focusing on the processes used for reading (interpreting, integrating, and evaluating and retrieving and straightforward inferencing). Similar to PIRLS, ePIRLS also includes a composite online reading scale. In addition, there are a number of dependent (or continuous) variables, other than scale scores, that you may choose as a measure. These variables fall under different categories, such as Student and Family Characteristics and School Characteristics.

#### 3.2.1.3 Years

Currently 2001, 2006, 2011, and 2016 PIRLS data, 2016 ePIRLS data are available through the IDE. Each year can be selected separately or all years can be selected together, by selecting All Years.

#### 3.2.1.4 Jurisdictions

In 2001, there were 35 countries and subnational education systems that participated in PIRLS. Two benchmarking jurisdictions also participated, the Canadian provinces of Ontario and Quebec. In addition, Sweden assessed a smaller sample of 3rd-graders. In 2006, there were 45 countries and subnational education systems that participated in PIRLS, and 5 benchmarking jurisdictions that participated. In addition, Norway and Iceland assessed a smaller sample of 5th-graders. In 2011, there were 57 countries and subnational education systems that participated in PIRLS, and 9 benchmarking jurisdictions that participated. The total of 57 includes 4 education systems that only gave the 4th-grade assessment to 5th- or 6th-graders. In 2016, there were 50 countries and subnational education systems that participated in PIRLS, and 11 benchmarking jurisdictions that participated. Denmark administered the 4th-grade assessment to both 3rd- and 4th-graders. South Africa administered the 4th-grade assessment to 5th-graders who spoke English, Afrikaans, and Zulu. In 2016, Norway chose to assess fifth and ninth grades to obtain better comparisons with Sweden and Finland, but also collected benchmark data at fourth and eighth grades to maintain trend with previous PIRLS cycles. At the 4th grade, five education systems participated in PIRLS Literacy (Egypt, Iran, Kuwait, Morocco, and South Africa), and two of these education systems completed both PIRLS and PIRLS Literacy (Iran and Morocco). Because Iran and Morocco participated both in PIRLS and PIRLS Literacy, their data reported is based on the average of both assessments. There were 14 countries and subnational education systems that participated in ePIRLS, and 2 benchmarking jurisdictions that participated. All listed jurisdictions can be selected for any analyses. However, the IDE contains a few U.S.-specific background variables (e.g., race/ethnicity) that, when selected, will not yield information for any other jurisdictions.

### 3.2.2 Variables

In the PIRLS IDE, questions from three types of questionnaires (student, teacher, and school) as well as variables that are derived from background information are organized into categories that have shared characteristics and can be selected as a group when designing and generating tables. Content category and subcategory titles may overlap, but specific variables appear only once in a subcategory. Use “Search” in the Select Variables step to locate variables.

#### 3.2.2.1 Achievement Levels

In addition to average scale scores, achievement results for PIRLS and ePIRLS are reported using achievement levels. The achievement levels are international benchmarks based on collective judgments about what students should know and be able to do relative to the body of content reflected in each subject-area assessment. The overall reading literacy scale is divided into international benchmarks. International benchmarks for the reading levels are as follows:

• Below low—below 400 • At low—between 400 and 474 • At intermediate—between 475 and 549 • At high—between 550 and 624 • At advanced—at or above 625

For more information on benchmarks, please visit <https://nces.ed.gov/surveys/pirls/pirls2016/technotes_intlbenchmarks.asp>.

### 3.2.3 Statistics Options

The IDE reports PIRLS data with several statistics options:

• Averages • Percentages • Percentiles • Standard deviations

#### 3.2.3.1 Averages

This statistic provides the average value for a selected continuous variable or the average scale score for the combined reading scale or one of the reading subscales. For the PIRLS assessment, student performance is reported on scales that range from 0 to 1,000, with the scale centerpoint fixed at 500 and a standard deviation of 100. PIRLS scales are produced using item response theory (IRT) to estimate average scores for reading literacy for each jurisdiction. IRT identifies patterns of response and uses statistical models to predict the probability of answering an item correctly as a function of students’ proficiency in answering other questions. That is, student responses to the assessment questions are analyzed to determine the percentage of students responding correctly to each multiple-choice question and the percentage of students achieving in each of the score categories for constructed-response questions.

#### 3.2.3.2 Percentages

This statistic shows the percentage of students as a row percentage. For example, if the first column lists countries, then each country will display its own percentage distribution across its row. If the table cell for Black female students in the United States is 9 percent, then Black females constituted 9 percent of U.S. fourth-graders. By default, percentage distributions do not include missing data, though there is an option to include the missing data.

#### 3.2.3.3 Standard deviations

The standard deviation is a measure of how widely or narrowly dispersed scores are for a particular dataset. Under general normality assumptions, 95 percent of the scores are within two standard deviations of the mean. For example, if the average score of a dataset is 500 and the standard deviation is 100, it means that 95 percent of the scores in this dataset fall between 300 and 700. The standard deviation is the square root of the variance. In the IDE, you may obtain standard deviations as one of your two choices for “Statistics Options” in the Edit Reports step.

#### 3.2.3.4 Percentiles

This statistic shows the threshold (or cutpoint) score for the following: • 10th percentile—the bottom 10 percent of students • 25th percentile—the bottom quarter of students • 50th percentile—the bottom half of students (half the students scored at or below the cutpoint and half scored above it) • 75th percentile—the top quarter of students • 90th percentile—the top 10 percent of students

### 3.2.4 Cross-tabulations

Cross-tabulation is a method of combining separate variables into a single table. Normally, each variable has its own table. If you have selected two or three variables (not counting All Students), when you go to the Edit Reports step, you will automatically get one table for each variable (including one for All Students); at the end of that list, you will get one cross-tabulation for the two or three variables selected. If you have chosen four or more variables (not counting All Students), you will get tables for each variable, but you won’t get the cross-tabulation. Be advised that if you go back to add another variable without subtracting one to keep the total under four, you will lose any edits you might have made to the cross-tabulation.

### 3.2.5 Statistical Notations and Other Notes

Statistical notations and other notes are found at the end of a data table, as applicable to that table:

• — Not available. • † Not applicable. (For instance, the standard error for the statistic cannot be reported because the statistic does not meet reporting standards.) • # The statistic rounds to zero. • ‡ Reporting standards not met. (For instance, the sample size is insufficient to permit a reliable estimate.) • NOTE: A general note pertains to any special characteristics of the data in the table. • SOURCE: Source information is listed for all PIRLS data and should be cited when data are used in a publication or presentation.

The general note (NOTE) warns users of jurisdiction-specific changes in population coverage, participation rates, or sampling procedures which deviated from international standards. Data from these jurisdictions have issues that interfere with proper trend analysis: Azerbaijan, Israel, Kuwait, Lithuania, Morocco, Poland, Qatar, and South Africa. Please be aware of these concerns for the following jurisdictions (years in parentheses): Alberta-CAN (11), Austria (16), Azerbaijan (11), Belgium (Flemish) (06), Belgium (French) (16, 11), Bulgaria (06), Canada (16, 11), Croatia (11), Denmark (16, 11, 06), England (11, 01), Florida-USA (11), Georgia (16, 11, 06), Greece (01), Hong Kong (16, 11), Israel (16, 11, 06, 01), Latvia (16), Lithuania (11, 01), Malta (16), Morocco (11, 01), the Netherlands (16, 11, 06, 01), Northern Ireland (11), Norway (11, 06), Oman (11), Ontario-CAN (11), Portugal (16), Qatar (11), Russian Federation (06, 01), Scotland (06, 01), Singapore (16, 11), the United States (16, 11, 06, 01), Quebec-CAN(16), and Madrid-ESP(16). Exclusion rates for Azerbaijan and Georgia for 2011 are slightly underestimated as some conflict zones were not covered and no official statistics were available. The TIMSS & PIRLS International Study Center has reservations about the reliability of the average 2011 achievement scores of Morocco and Oman because the percentage of students with achievement too low for estimation exceeds 15 percent. Results for Canada (Ontario) in 2006 may differ slightly from the IEA PIRLS 2006 International Report because the 2006 data shown include public and private schools, whereas the IEA report excluded private schools from trend analysis for Canada (Ontario) in 2006 to match its 2001 sample.

#### 3.2.5.1 Linking Teacher Data

Results shown in the PIRLS IDE may differ slightly from those in the International Association for the Evaluation of Educational Achievement (IEA) PIRLS International Reports because of a slightly different procedure used in linking teacher data to the students. In the IEA report, when a student has more than one teacher, the student’s weight is distributed equally amongst the teachers, and all the teacher data are used in the analysis. For the same case, the IDE randomly selects one of the teachers for the student, and the entire weight for the student is assigned to this teacher.

#### 3.2.5.2 Statistical Comparisons

The alpha level to establish significance for all comparisons is .05. All comparisons within a jurisdiction, within the same year, are made using dependent samples t-tests. Comparisons between jurisdictions, and comparisons between years, even for the same jurisdiction, are made using independent samples t-tests. The PIRLS IDE also uses independent samples t-tests, between a country and a subnational entity that is participating as a benchmarking entity (for instance, in order to compare scores between the United States and Florida, since they each are an independent sample).

#### 3.2.5.3 Data Suppression

Finally, data suppression may be handled slightly differently in the PIRLS IDE and the IEA PIRLS International Reports. For the IDE, the Rule of 62 is applied to suppress data to avoid reporting results for groups about which little of interest could be said due to lack of power. The Rule of 62 is borrowed from the IDE’s counterpart, the National Assessment of Educational Progress (NAEP) Data Explorer (NDE). This rule states that statistics for a group are suppressed if they are based on less than 62 cases. Statistics are: means, standard errors, standard deviations, a set of percentiles, and a set of achievement-level percentages. The rule serves to assure a minimum power requirement to detect moderate differences at nominal significance level (0.05). The minimum power is 0.80 and the moderate effect size is 0.5 standard deviation units. A design effect of 2 is assumed to derive an appropriate complex sample standard deviation.

## 3.3 TIMSS IDE

### 3.3.1 Criteria

Each data query must include at least one selection from four criteria choices: subject, grade, measure(s), and jurisdiction(s). Shown below is an outline of these selection criteria followed by a brief description.

1. Subject:
   * Mathematics and Science
   * TIMSS Advanced: Advanced Mathematics
   * TIMSS Advanced: Physics
2. Grade:
   * Grade 4
   * Grade 8
   * End of High School
3. Years
   * 2019
   * 2015
   * 2011
   * 2007
   * 2003
   * 1999
   * 1995
4. Measure:
   * TIMSS scale scores
     + Mathematics: Grade 4
       - Overall scale
       - Subscales
     + Mathematics: Grade 8
       - Overall scale
       - Subscales
     + Science: Grade 4
       - Overall scale
       - Subscales
     + Science: Grade 8
       - Overall scale
       - Subscales
     + TIMSS Advanced: Advanced Mathematics: End of High School
       - Overall scale
       - Subscales
     + TIMSS Advanced: Physics: End of High School
       - Overall scale
       - Subscales
   * Student and Family Characteristics
   * Student Computer Use
   * Student Activities Outside of School
   * Student Perception/Valuing of Mathematics/Science
   * Teacher Background Characteristics, Formal Education, and Training
   * Teacher Perception of Mathematics/Science Teaching/Learning
   * Teacher Preparation and Collaboration
   * Teacher Activities Outside of School (when Mathematics and Science is selected)
   * Classroom Characteristics
   * Classroom Instruction
   * Role of Homework (Teacher)
   * School Characteristics
   * School Resources
   * Home Involvement (School)
   * School Climate and Safety
5. Jurisdiction:
   * Average of Countries
   * Average of the Selected Countries/Participants
   * Countries
   * U.S. Jurisdiction (when Mathematics and Science is selected)
   * Benchmarking Participants (when Mathematics and Science is selected)

#### 3.3.1.1 Subject & Grade

TIMSS is a study of mathematics and science, and those are the subjects that can be selected. The Mathematics and Science subject can only be selected with either Grade 4 or Grade 8 options. TIMSS Advanced is a study of advanced mathematics and physics, and those are the subjects that can be selected. The TIMSS Advanced: Advanced Mathematics and the TIMSS Advanced: Physics options can only be selected with the End of High School grade option.

#### 3.3.1.2 Measures

TIMSS focuses on overall mathematics and science knowledge, but within these broad categories a variety of subscales are available each year, including the environmental awareness subscale introduced in TIMSS 2019. Subscales are constituent parts of the composite subject scale for an assessment, and are specified by the assessment framework for that year. The weighted average of these is the basis for the mathematics and science composite scales, as described in the TIMSS and TIMSS Advanced frameworks. Subscales are based on fewer observations than the composite scales and, as a result, may have larger standard errors. In addition, there are a number of continuous variables other than scale and subscale scores that you may choose as a measure of analysis. These variables fall under different categories, such as Student and Family Characteristics and School Characteristics, and include variables such as age, teaching experience, and class size.

#### 3.3.1.3 Jurisdictions & Years

Note that some country counts overlap because some countries participated at both the fourth- and eighth-grade levels. Also, benchmarking participants are currently available in the IDE for 2019, 2015, 2011, 2007, and 2003. So they are only listed below for those years. In 2019, a total of 64 education systems participated in TIMSS at the 4th grade, while 46 systems participated at the 8th grade. Most of these education systems are member countries of the International Association for the Evaluation of Educational Achievement (IEA), the group that sponsors TIMSS internationally; a small number at each grade are nonmember subnational entities that joined TIMSS 2019 as “benchmarking participants”. In 2015, there were 49 countries and subnational education systems, as well as 6 benchmarking participants that participated in TIMSS at the fourth-grade level. At the eighth-grade level, 38 countries and subnational education systems participated along with 6 benchmarking participants. Nine countries participated in TIMSS Advanced at the end of high school. Also, for TIMSS 2015, countries where students were expected to find the TIMSS assessments too difficult for their fourth- or eighth-grade students were given the option to assess students at a higher grade. Accordingly, one country (South Africa) administered the fourth grade assessment to fifth grade students and two countries (Bostwana and South Africa) administered the eighth grade assessment to ninth grade students. All off-grade participants (i.e., countries that tested students at grades other than four and eight) have the tested grade in parentheses within the IDE system. For example, South Africa’s label is “South Africa (5)” and “South Africa (9)”. In 2015, Norway chose to assess fifth and ninth grades to obtain better comparisons with Sweden and Finland, but also collected benchmark data at fourth and eighth grades to maintain trend with previous TIMSS cycles. In 2019, Norway continued assessing at the fifth and ninth grade level and did not test grades four and eight. Additionally, for TIMSS 2015, 7 countries and 1 benchmarking education system participated in the Numeracy assessment (newly developed TIMSS Numeracy assessment, a less difficult version of the fourth grade mathematics assessment), including Bahrain, Indonesia, Iran, Kuwait, Jordan, Morocco, and South Africa as well as Buenos Aires. Each of these participants gave the fourth-grade assessments in mathematics and science as well as the Numeracy assessment, except for Jordan and South Africa, which each participated exclusively in Numeracy. In 2011, there were 52 countries and subnational education systems, as well as 7 benchmarking participantsthat participated in TIMSS at the fourth-grade level. At the eighth-grade level, 45 countries and subnational education systems participated along with 14 benchmarking participants. Also, for TIMSS 2011, countries where students were expected to find the TIMSS assessments too difficult for their fourth- or eighth-grade students were given the option to assess students at a higher grade. Accordingly, three countries administered the fourth grade assessment to their sixth grade students and the eighth grade assessment to their ninth grade students. In 2007, there were 37 countries and subnational education systems, as well as 7 benchmarking participantsthat participated in TIMSS at the fourth-grade level. At the eighth-grade level, 50 countries and subnational education systems participated along with 7 benchmarking participants. In 2003, there were 25 countries and subnational education systems, as well as 3 benchmarking participantsthat participated in TIMSS at the fourth-grade level. At the eighth-grade level, 48 countries and subnational education systems participated along with 4 benchmarking participants. In 1999, there were 38 countries and subnational education systems that participated in TIMSS at the eighth-grade level. Fourth-grade students were not assessed in TIMSS 1999. In 1995, there were 29 countries and subnational education systems that participated in TIMSS at the fourth-grade level. At the eighth-grade level, 46 countries and subnational education systems participated.

### 3.3.2 Variables

In the TIMSS IDE, questions from three types of questionnaires (student, teacher, and school) as well as variables that are derived from background information are organized into categories that have shared characteristics and can be selected as a group when examining and generating tables. Content category and subcategory titles may overlap, but specific variables appear only once in a subcategory. Use *Search* in the *Select Variables* step to locate variables.

#### 3.3.2.1 Benchmarks

In addition to average scale scores, achievement results for TIMSS and TIMSS Advanced are reported using benchmarks. The benchmarks are internationally set levels based on collective judgments about what students should know and be able to do relative to the body of content reflected in each subject-area assessment. Using score cutpoints, the average scale scores are divided into four international benchmarks for TIMSS (low, intermediate, high, and advanced) and three international benchmarks for TIMSS Advanced (intermediate, high, and advanced). TIMSS benchmark data for grades 4 and 8 are presented in a discrete format. This “discrete” format presents the percentage of students performing at each international benchmark: at low, at intermediate, at high, and at advanced, with an additional category created for those students scoring below the low benchmark (below low). (Note that there is simply too little information to know what students scoring below the low benchmark can actually do.) TIMSS Advanced benchmark data are presented in a discrete format. This “discrete” format presents the percentage of students performing at each international benchmark: at intermediate, at high, and at advanced, with an additional category created for those students scoring below the intermediate benchmark (below intermediate). Please note that the TIMSS assessment is not designed to assess students scoring below the intermediate benchmark. For more information on benchmarks, please visit <https://nces.ed.gov/timss/technotes.asp#_Toc94791995>.

#### 3.3.2.2 Index Variables

In addition to scale scores representing performance in various subjects, TIMSS and TIMSS Advanced use indices derived from the student, teacher, and school questionnaires to contextualize results or estimate trends that account for demographic changes over time. Information on indices for each year of administration can be found in the chapters referenced in the summary table below.

| Study Name | Study Year | Links |
| --- | --- | --- |
| TIMSS | 2019 | [Link](https://timss2019.org/international-database/downloads/T19_UG_Supp3-derived-context-variables.pdf) |
| TIMSS | 2015 | [Link](http://timssandpirls.bc.edu/timss2015/international-database/downloads/T15_UG_Supplement3.pdf) |
| TIMSS | 2011 | [Link](https://timssandpirls.bc.edu/timss2011/downloads/T11_UG_Supplement3.pdf) |
| TIMSS | 2007 | [Link](https://timssandpirls.bc.edu/TIMSS2007/PDF/T07_UserGuide_Supp3.zip) |
| TIMSS | 2003 | [Link](https://timssandpirls.bc.edu/timss2003i/PDF/t03_ug_s3.pdf) |
| TIMSS | 1999 | [Link](https://timssandpirls.bc.edu/timss1999i/data/bm2_supplement3.pdf) |
| TIMSS | 1995 | [Link](https://timssandpirls.bc.edu/timss1995i/database/UG1_Sup4.pdf) |
| TIMSS Advanced | 2015 | [Link](https://timssandpirls.bc.edu/timss2015/advanced-international-database/downloads/TA15_UG_Supplement3.pdf) |
| TIMSS Advanced | 1995 | [Link](https://timssandpirls.bc.edu/timss1995i/database/UG3_Sup3.pdf) |

### 3.3.3 Statistics Options

The IDE reports TIMSS data with several statistics options:

• Averages • Percentages • Standard deviations • Percentiles

#### 3.3.3.1 Averages

This statistic provides the average value for a selected continuous variable or overall score for the combined scale (e.g., TIMSS Mathematics Scale: Overall Mathematics) or score for one of the subscales corresponding to the subject chosen (e.g., TIMSS Mathematics Scale: Algebra). For the TIMSS and TIMSS Advanced assessment, student performance is reported on scales that range from 0 to 1,000, with the TIMSS scale centerpoint fixed at 500 and a standard deviation of 100. Scale scores can show the standard error and are often accompanied by data showing percentages and standard deviations.

TIMSS scales are produced using item response theory (IRT) to estimate average scores for mathematics, science, advanced mathematics, and physics for each jurisdiction. IRT identifies patterns of response and uses statistical models to predict the probability of answering an item correctly as a function of students’ proficiency in answering other questions. That is, student responses to the assessment questions are analyzed to determine the percentage of students responding correctly to each multiple-choice question and the percentage of students achieving in each of the score categories for constructed-response questions.

The TIMSS achievement scale was established in 1995 based on the combined achievement distribution of all countries that participated in TIMSS 1995. To provide a point of reference for country comparisons, the scale centerpoint of 500 was located at the mean of the combined achievement distribution. The units of the scale were chosen so that 100 scale score points corresponded to the standard deviation of the distribution. In the IDE, Average of Countries shows the TIMSS scale centerpoint when “All students” is selected at Step 2 as the independent variable.

#### 3.3.3.2 Percentages

This statistic shows the percentage of students as a row percentage. For example, if the first column lists countries, then each country will display its own percentage distribution across its row. By default, percentage distributions do not include missing data, although there is an option to include them.

#### 3.3.3.3 Standard deviations

The standard deviation is a measure of how widely or narrowly dispersed scores are for a particular dataset. Under general normality assumptions, 95 percent of the scores are within two standard deviations of the mean. For example, if the average score of a dataset is 500 and the standard deviation is 100, it means that 95 percent of the scores in this dataset fall between 300 and 700. The standard deviation is the square root of the variance.

#### 3.3.3.4 Percentiles

This statistic shows the threshold (or cutpoint) score for the following:

• 10th percentile—the bottom 10 percent of students • 25th percentile—the bottom quarter of students • 50th percentile—the median (half the students scored below the cutpoint and half scored above it) • 75th percentile—the top quarter of students • 90th percentile—the top 10 percent of students

### 3.3.4 Cross-tabulations

Cross-tabulation is a method of combining separate variables into a single table. Normally, each variable has its own table. If you have selected two or three variables (not counting *All students*) and when you go to the *Edit Reports* step, you will automatically get one table for each variable (including one for *All students*); at the end of that list, you will get one cross-tabulation for the two or three variables selected. If you have chosen four or more variables (not counting *All students*), you will get tables for each variable, but you won’t get the cross-tabulation. Be advised that if you go back to add another variable without subtracting one to keep the total under four, you will lose any edits you might have made to the cross-tabulation.

### 3.3.5 Statistical Notations and Other Notes

Statistical notations and other notes are found at the end of a data table, as applicable to that table:

• — Not available. • † Not applicable. (For instance, the standard error for the statistic cannot be reported because the statistic does not meet reporting standards.) • # The statistic rounds to zero. • ‡ Reporting standards not met. (For instance, the sample size is insufficient to permit a reliable estimate.) • NOTE: A general note pertains to any special characteristics of the data in the table. Population coverage, participation rates, sampling procedures, reliability standards, and trend comparability issues are addressed here. See details below. • SOURCE: Source information is listed for all TIMSS and TIMSS Advanced data and should be cited when data are used in a publication or presentation.

Population coverage, participation rates, sampling procedures, or reliability standards deviated from international standards in the following jurisdictions:

*Grade 4*:

| Region/Country | Years Participated |
| --- | --- |
| Abu Dhabi-UAE | 15 |
| Alberta-CAN | 11, 07 |
| Australia | 03, 95 |
| Austria | 95 |
| Azerbaijan | 11 |
| Bahrain | 15 |
| Belgium (Flemish)-BEL | 19, 15 |
| British Columbia-CAN | 07 |
| Canada | 19, 15 |
| Croatia | 11 |
| Denmark | 19, 15, 11, 07 |
| Dubai-UAE | 19, 07 |
| England-GBR | 19, 03, 95 |
| Florida-USA | 15, 11 |
| Georgia | 19, 15, 11, 07 |
| Hong Kong-CHN | 19, 15, 11, 03 |
| Hungary | 95 |
| Israel | 95 |
| Italy | 15 |
| Kazakhstan | 19, 11, 07 |
| Kosovo | 19 |
| Kuwait | 15, 11, 95 |
| Latvia | 19, 07, 95 |
| Lithuania | 19, 15, 11, 07, 03 |
| Massachusetts-USA | 07 |
| Minnesota-USA | 07 |
| Mongolia | 07 |
| Morocco | 19, 15 |
| Netherlands | 19, 15, 11, 07, 03, 95 |
| New Zealand | 19 |
| North Carolina-USA | 11 |
| Northern Ireland-GBR | 19, 15, 11 |
| Norway | 19, 11 |
| Ontario-CAN | 19, 07 |
| Pakistan | 19 |
| Philippines | 19 |
| Portugal | 19, 15 |
| Qatar | 11 |
| Quebec-CAN | 15, 07 |
| Russian Federation | 19 |
| Saudi Arabia | 19, 15 |
| Scotland-GBR | 07, 03, 95 |
| Serbia | 19, 15, 11 |
| Singapore | 19, 15, 11 |
| Slovak Republic | 19 |
| Slovenia | 95 |
| South Africa | 19 |
| Spain | 15 |
| Sweden | 15 |
| Thailand | 95 |
| Turkey | 19 |
| United States | 19, 15, 11, 07, 03 |

*Grade 8*:

| Region/Country | Years Participated |
| --- | --- |
| Abu Dhabi-UAE | 19 |
| Alabama-US | 11 |
| Alberta-CAN | 11 |
| Australia | 95 |
| Austria | 95 |
| Belgium (Flemish)-BEL | 99, 95 |
| Belgium (French)-BEL | 95 |
| British Columbia-CAN | 07 |
| Buenos Aires-ARG | 15 |
| California-USA | 11 |
| Canada | 15 |
| Chile | 19, 15 |
| Colombia | 95 |
| Colorado-USA | 11 |
| Connecticut-USA | 11 |
| Denmark | 95 |
| Dubai-UAE | 19, 07 |
| Egypt | 19, 15 |
| England-GBR | 11, 07, 03, 99, 95 |
| Florida-USA | 15, 11 |
| Gauteng-ZAF | 19 |
| Georgia | 19, 15, 11, 07 |
| Greece | 95 |
| Honduras-Grade 9 | 11 |
| Hong Kong-CHN | 19, 07, 03, 99 |
| Indiana-USA | 11 |
| Indonesia | 03 |
| Iran, Islamic Rep. of | 15 |
| Israel | 19, 15, 11, 07, 03, 99, 95 |
| Italy | 15 |
| Jordan | 19, 15 |
| Kazakhstan | 19 |
| Kuwait | 19, 15, 95 |
| Latvia | 99, 95 |
| Lithuania | 15, 11, 07, 03, 99, 95 |
| Macedonia | 03 |
| Massachusetts-USA | 11, 07 |
| Minnesota-USA | 11, 07 |
| Mongolia | 07 |
| Morocco | 19, 15, 03, 07 |
| Netherlands | 03, 99, 95 |
| New Zealand | 19, 15 |
| North Carolina-USA | 11 |
| Norway | 19 |
| Oman | 19, 15 |
| Ontario-CAN | 11, 07 |
| Qatar | 19, 15 |
| Quebec-CAN | 19, 15, 07 |
| Romania | 95 |
| Russian Federation | 19, 11 |
| Saudi Arabia | 19, 15 |
| Scotland-GBR | 07, 03, 95 |
| Serbia | 07, 03 |
| Singapore | 19, 15, 11 |
| Slovak Republic | 19 |
| Slovenia | 95 |
| South Africa | 19 |
| Sweden | 19 |
| Switzerland | 95 |
| Thailand | 95 |
| United States | 19, 15, 11, 07, 03, 95 |
| Western Cape-ZAF | 19 |

*TIMSS Advanced*:

| Region/Country | Years Participated |
| --- | --- |
| Lebanon | 15 |
| Portugal | 15 |
| United States | 15 |

Jurisdictions with a number after their name (for instance, Norway (5) and Norway (9)) have participated with a grade different than most other jurisdictions. The number in parentheses indicate the grade.

TIMSS Advanced assesses the advanced mathematics and physics knowledge and skills of students in their final year of secondary school who were taking or had taken courses in advanced mathematics and physics; the percentage of the age cohort enrolled in these courses and considered eligible for the TIMSS Advanced study varied across participating jurisdictions (ranging from 2% to 34% in 2015, and was 11% in the United States for advanced mathematics and 5% for physics).

In TIMSS Advanced 2015, the Russian Federation participated with two populations of students for Advanced Mathematics —results for students in intensive courses (6 or more hours per week) are reported separately from the results for other students from the Russian Federation taking courses that involve 4.5 hours per week.

In 2015, Armenia tested the same cohort of students as other countries, but later in the assessment year.

Data from these jurisdictions have issues that interfere with proper trend analysis: Armenia, Australia, Botswana, Canada, Finland, Indonesia, Israel, Italy, Kazakhstan, Kuwait, Latvia, Morocco, Norway, Philippines, Poland, Qatar, Saudi Arabia, Slovenia, South Africa, Syrian Arab Republic, Thailand, Turkey, and Yemen. For more details on trends with 2019 data, see Appendix A in the IEA [TIMSS 2019 International Reports](https://timss2019.org/reports/), which lists all countries with previous years of data not comparable for measuring trends to 2019, primarily due to countries improving translations or increasing population coverage.

See the [IEA TIMSS 2015 International Reports](http://timssandpirls.bc.edu/timss2015/), [IEA TIMSS 2011 International Reports](http://timssandpirls.bc.edu/timss2011/index.html), the [IEA TIMSS 2007 International Reports](http://timss.bc.edu/TIMSS2007/intl_reports.html), and the [IEA TIMSS 2003 International Reports](http://timss.bc.edu/timss2015-advanced/frameworks.html) for further information on specific trend issues in previous years.

Because of national-level changes in the starting age/date of school, 1999 data for Australia and Slovenia cannot be compared to 2003 data. Because of changes in the population tested, 1995 data for Israel, Italy, New Zealand, and South Africa and 1999 data for Morocco cannot be used for trend analyses. Because only Latvian-speaking schools were included in 1995 and 1999 data for Latvia, 1995 and 1999 data cannot be compared to 2003, 2007, and 2011 data. Data for Kuwait, Indonesia, Saudi Arabia, Morocco, and Turkey cannot be used for trend analyses because comparable data across years are not available.

The Syrian Arab Republic participated in TIMSS 2003 at the 8th grade and Yemen participated in TIMSS 2003 at the 4th grade, but because the characteristics of their sample are not completely known, they were shown in an appendix in the TIMSS 2003 International Report and their 2003 data are excluded from the IDE.

South Africa and Bulgaria participated in TIMSS 1995 at the 8th grade, but due to problems with their background data, their 1995 data are excluded from the IDE.

#### 3.3.5.1 Linking teacher data

Results shown in the TIMSS IDE may differ slightly from those in the International Association for the Evaluation of Educational Achievement (IEA) TIMSS International Reports because of a slightly different procedure used in linking teacher data to the students. For Grade 4 and Grade 8, some students (mostly for Grade 8) may be assigned more than one science or mathematics teacher. Each teacher is asked to complete the teacher questionnaire, and the IEA TIMSS International Reports present results that are based on averaged data for these teachers. For the TIMSS IDE, if a student has more than one teacher for each subject, a student is linked to data from a single teacher for mathematics and science. The teacher is chosen randomly from the group of teachers (mathematics or science) who answered the questionnaire for each student.

#### 3.3.5.2 Statistical Comparisons

The alpha level to establish significance for all comparisons is .05. All comparisons within a jurisdiction, within the same year, are made using dependent samples t-tests. Comparisons between jurisdictions, and comparisons between years, even for the same jurisdiction, are made using independent samples t-tests. The TIMSS IDE also uses independent samples t-tests, between a country and a subnational entity that is participating as a benchmarking entity (for instance, in order to compare scores between the United States and Massachusetts or Minnesota, since they each are an independent sample).

#### 3.3.5.3 Data Suppression

Data suppression may be handled slightly differently in the TIMSS IDE and the IEA TIMSS International Reports. For the IDE, the Rule of 62 is applied to suppress data to avoid reporting results for groups about which little of interest could be said due to lack of power. The Rule of 62 is borrowed from the IDE’s counterpart, the National Assessment of Educational Progress (NAEP) Data Explorer (NDE). This rule states that statistics for a group are suppressed if they are based on less than 62 cases. Statistics are: means, standard errors, standard deviations and a set of percentiles. The rule serves to assure a minimum power requirement to detect moderate differences at a nominal significance level (0.05). The minimum power is 0.80 and the moderate effect size is 0.5 standard deviation units. A design effect of 2 is assumed to derive an appropriate complex sample standard deviation.

For information on creating and interpreting the TIMSS 2019 context questionnaire scales, see [Methods and Procedures: TIMSS 2019 Technical Report](https://timssandpirls.bc.edu/timss2019/methods/chapter-16.html).

## 3.4 PIAAC IDE

### 3.4.1 Criteria

Each data query must include at least one selection from four criteria choices: display, years/studies, measure, and jurisdiction. Shown below is an outline of these selection criteria followed by a brief description.

1. Display
   * U.S. Adults, 16–74 (Household and Prison) (data available for U.S. PIAAC 2017 and 2012/2014)
   * Young Adults, 16–34
   * Adults, 16–65
2. Years/Studies:
   * PIAAC 2017 (data available for literacy, numeracy and problem solving in technology-rich environments)
   * PIAAC 2012/2014 (data available for literacy, numeracy and problem solving in technology-rich environments)
   * ALL 2003–2008 (data available for literacy and numeracy)
   * IALS 1994–1998 (data available for literacy)
3. Measure:
   * PIAAC Literacy: Overall scale
     + PIAAC Reading Components scale
   * PIAAC Numeracy: Overall scale
   * PIAAC Problem solving in technology-rich environments: Overall scale
   * Other continuous variables from the background questionnaire, including international variables, derived variables, and U.S. national adaptations and additions to the International background questionnaire.
4. Jurisdiction:
   * Within U.S. Adults, 16–74 (Household and Prison) Display:
     + U.S. Household (16–74 years old)
     + U.S. Prison (16–74 years old)
     + U.S. Household (16–65 years old)
   * Within Young Adults, 16–34 or Adults, 16–65 Display
     + Average of All Jurisdictions
     + Average of the Selected Jurisdictions
     + OECD National Entities
     + OECD Sub-National Entities
     + Partners

#### 3.4.1.1 Display

The PIAAC IDE contains three different adult sample populations which can be selected for analysis from the *Display* drop down menu.

* *U.S. Adults, 16–74 (Household and Prison)*: This display contains U.S.-only comparable data from the PIAAC, including the 2017 U.S. Household Data (for ages 16–74, and 16–65), combined 2012 and 2014 U.S. Household Data (for ages 16–74, and 16–65), and Prison Data (for ages 16–74).
* *Young Adults, 16–34*: This display contains internationally comparable data from the 3 international rounds of PIAAC (2012-2017 for all countries, except the U.S. which combined 2012-2014 data only) Household Data, ages 16–34. This display does not include the 2017 U.S. Household Data.
* *Adults, 16–65*: This display contains internationally comparable data from the 3 international rounds of PIAAC (2012–2017 for all countries, except the U.S. which combined 2012–2014 data only) Household Data, ages 16–65. This display does not include the 2017 U.S. Household Data.

#### 3.4.1.2 Measures

You can choose the overall scale, which is each subject’s default measure in the PIAAC IDE or there are also a number of continuous variables other than scale scores that you may choose as a measure of analysis. These variables are continuous variables from the international and U.S. national background questionnaire (such as earnings or hours of work per week) and derived variables from PIAAC, ALL, and IALS. Derived variables from PIAAC include indices of literacy, numeracy, and computer use at work and at home and imputed years of formal education, among others. A fourth domain, called Reading Components, measures literacy at the very low end of the spectrum, in areas such as sentence completion, passage comprehension, and vocabulary. This domain was given to respondents who decided not to take the computer-based assessment or who did not pass a set of core information and computer technology tasks and a set of core literacy/numeracy tasks. The adults in the sample population that did not answer the assessment will be displayed along with those that did answer the assessment if you select *Percentage across full sample* under the *Population* category.

#### 3.4.1.3 Years/Studies

Currently, data availability in the PIAAC IDE is dependent on the *Display* and *Measure* selected in step 1, *Select Criteria*. If the *Display* chosen is *U.S. Adults, 16–74 (Household and Prison)* you can choose one or more years and studies between *PIAAC 2017* and *PIAAC 2012/14*. If the *Display* chosen is *Adults, 16–65* or *Young Adults, 16–34* you can choose one or more years and studies between *PIAAC 2012-2017*, *ALL 2003–2008*, and *IALS 1994–1998*.

#### 3.4.1.4 Jurisdictions

All listed jurisdictions can be selected for any analyses, provided data are available for the selected years/studies range. When PIAAC was first administered in 2012, a total of 24 jurisdictions participated, including the United States. Nine additional jurisdictions administered PIAAC in 2014 and five additional jurisdictions administered PIAAC in 2017. Data for these jurisdictions, with the exception of three, are available within the Adults, age 16–65 and Adults, age 16–34 displays. Data for three jurisdictions, Australia, Jakarta (Indonesia), and Russian Federation, are not available: Australia’s data are suppressed in the PIAAC IDE because of national restrictions on the use of their data; Jakarta’s data are suppressed because their data file is not publicly available; and Russian Federation’s data are suppressed in the PIAAC IDE because the data do not represent the entire resident population aged 16–65 years in Russia. Jurisdictions include some subnational entities, such as England/Northern Ireland. Data are not available for some of the 33 PIAAC-participating jurisdictions for ALL 2003–2008 or IALS 1994–1998, either because they did not participate in that assessment or because their data were suppressed due to reporting standards not being met (see Table 2). Data are available for 6 jurisdictions in ALL 2003–2008, and 15 jurisdictions in IALS 1994–1998. Jurisdictions for which data are not available for a selected year are identified by the icon representing “no data”— .

Table 2. PIAAC IDE jurisdictions with available data by year/study

| Jurisdiction Groups | PIAAC 2012–2017 | ALL 2003–2008 | IALS 1994–1998 |
| --- | --- | --- | --- |
| OECD National Entities | 27 | 7 | 16 |
| OECD Sub-National Entities | 4[[2]](#footnote-165) | 0 | 4[[3]](#footnote-166) |
| Partners | 6 | 0 | 0 |
| Total Jurisdictions[[4]](#footnote-167) | 37 | 7 | 20 |

*NOTE: In the U.S. Adults 16-74 (Household and Prison) Display, data for the U.S. Household (16–74 years old), U.S. Prison (16–74 years old), and U.S. Household (16–65) are the only selections available in the Jurisdiction menu. However, users may choose to work with the U.S. prison sample, or one of the U.S. household samples, or select more than one for analysis as jurisdictions.*

### 3.4.2 Variables

PIAAC requires in-person interviews to complete the background questionnaire before administering the direct assessments (i.e., literacy, numeracy, reading components, and/or problem solving in technology-rich environments (PS-TRE)). In the PIAAC IDE, measures are derived from two instruments: the computer-based assessment (CBA), given to respondents who were comfortable taking the assessment on a computer, and the paper-based assessment (PBA), given to respondents that were not familiar with computers or chose not to take the assessment on a computer. Variables derived from the background questionnaire were administered to each participating adult. Variables are organized into categories that have shared characteristics and can be selected as a group (category) when examining and generating tables. Content category and subcategory titles may overlap, but specific variables appear only once in a subcategory. Use *Search* in the *Select Variables* step to locate variables. Note that some variables might be similar in content, but not comparable over the years, either due to differences in the question asked or differences in their response categories. The icon representing “no data”— —will help in identifying the year for which the variable has data available for analysis. Except for the estimates for *All Adults*, the variables that can be compared across years are located under a special category called *Trend Variables*, sub-category *Trends to IALS and ALL*. Note that common variables such as age and gender, among others, can appear in other categories and sub-categories with the “no data” icon, but have data when selected under the *Trends Variables* category.

#### 3.4.2.1 Proficiency levels

Proficiency levels are available in the Proficiency Levels sub-category in the Major Reporting Groups category. Achievement results for PIAAC are reported using achievement levels for literacy, numeracy, and problem solving in technology-rich environments (PS-TRE). Increasing levels represent the knowledge, skills, and capabilities needed to perform tasks of increasing complexity. As a result, the findings are reported in terms of percentages of the adult population at each of the predefined levels. Based on the statistics option chosen, the IDE can report the average scores of adults at each proficiency level or the percentage of the adults performing at each of the predefined levels for the chosen jurisdictions. The statistics options to choose standard deviations and percentiles will not generate reports as proficiency levels are not reportable using these statistical analyses. The IDE can report percentage distributions of variables among the adults at each proficiency level (for example, the percentage distribution of adult population that are employed, unemployed, and out of the labor force [employment status] within each proficiency level). To conduct this type of analysis, you can select the relevant Overall scale or Percentage across full sample at step 1, Select Criteria. At step 2, Select Criteria, you can select the relevant proficiency levels in addition to other variable(s) of interest. At the Edit Reports step, you can select the Edit command for the cross-tabulated report to change the table layout and move the proficiency levels variable to the row and one (or both) of the other selected variables to the column (this step may not be necessary depending on the order of the selected variables) and to select Percentages as the Statistic. Results for combined proficiency levels (for example, the combined level 4/5 proficiency level used in most reporting of PIAAC literacy and numeracy results) can be produced by creating a new variable within the Edit Report page. (For further information, see section 3.D. Create Variables.) You can then proceed to the Build Reports step. Literacy and numeracy results in PIAAC 2012–2017 and ALL 2003–2008, and literacy results in IALS 1994–1998 were reported using six achievement levels: below level 1, level 1, level 2, level 3, level 4, and level 5. Literacy related non-response is also available. The number of achievement levels in problem solving in technology-rich environments (PS-TRE) differs from the number in literacy and numeracy for PIAAC 2012–2017 where four achievement levels were used: below level 1, level 1, level 2, and level 3. Four other levels are also available in problem solving in technology-rich environments (PS-TRE) achievement levels: no computer experience, failed ICT core, refused CBA, and literacy related non-response (which are further explained in the Description of PIAAC problem solving in technology-rich environments (PS-TRE) achievement levels table below.) For more information on proficiency levels, please visit <https://nces.ed.gov/surveys/piaac/measure.asp>.

### 3.4.3 Statistics Options

The IDE reports PIAAC data with several statistics options:

• Averages • Percentages • Standard deviations • Percentiles

#### 3.4.3.1 Averages

This statistic provides the average value for a selected continuous variable or the average scale score. For the PIAAC assessment, adult performance is reported on scales that range from 0 to 500. PIAAC scales are produced using item response theory (IRT) to estimate average scores for literacy, numeracy, and problem solving in technology-rich environments (PS-TRE) for each jurisdiction. IRT identifies patterns of response and uses statistical models to predict the probability of answering an item correctly as a function of the adults’ achievement in answering other questions. That is, all participants’ responses to the assessment questions are compiled and analyzed to determine the percentage of adults responding correctly to each multiple-choice question and the percentage of adults achieving each of the score categories for constructed-response questions.

#### 3.4.3.2 Percentages

This statistic shows the percentage of adults as a row percentage. For example, if the first column lists jurisdictions, then each jurisdiction will display its own percentage distribution across its row. By default, percentage distributions do not include missing data, although there is an option to include them. The adults in the sample population that did not answer the assessment will be displayed along with those that did answer the assessment if you select Percentage across full sample under the Population category.

#### 3.4.3.3 Standard deviations

The standard deviation is a measure of how widely or narrowly dispersed scores are for a particular variable. Under general normality assumptions, 95 percent of the scores are within two standard deviations of the mean. For example, if the average value of a variable is 500 and the standard deviation is 100, it means that 95 percent of the values in this variable fall between 300 and 700. The standard deviation is the square root of the variance.

#### 3.4.3.4 Percentiles

This statistic shows the threshold score (or cut point) for the following:

• 10th percentile – the bottom 10 percent of students • 25th percentile – the bottom quarter of students • 50th percentile – the median (half the students scored below the cut point and half scored above it) • 75th percentile – the top quarter of students • 90th percentile – the top 10 percent of students

### 3.4.4 Cross-tabulations

Cross-tabulation is a method of combining separate variables into a single table. Normally, each variable has its own table. If you have selected two or three variables (not counting All Adults), when you go to the Edit Reports step, you will automatically get a list with one table for each variable (including one for All Adults); at the end of that list, you will get one cross-tabulation for the two or three variables selected. If you have chosen four or more variables (not counting All Adults) you will get tables for each variable, but you will not get the cross-tabulation. Be advised that if you go back to add another variable without removing one variable (to keep the total under four) you will lose any edits you might have made to the cross-tabulation.

### 3.4.5 Statistical Notations and Other Notes

Statistical notations and other notes are found at the end of a data table, as applicable to that table:

* — Not available.
* † Not applicable. (Data were not collected or not reported.)
* # The statistic rounds to zero.
* ‡ Reporting standards not met. (Did not meet reporting standard.)
* NOTE: A general note pertains to any special characteristics of the data in the table.
* SOURCE: Source information is listed for all PIAAC data and should be cited when data are used in a publication or presentation.

#### 3.4.5.1 Calculation of OECD averages

The IDE generates the average of all jurisdictions included in the IDE for the selected measures and variables if Average of All Jurisdictions is chosen under Jurisdiction. This average generated by the IDE is based on 27 OECD national and 2 sub-national entities [Flanders (Belgium), England and Northern Ireland (UK)] and 6 partner jurisdictions in PIAAC 2012–2017, 6 OECD national or sub-national entities in ALL 2003–2008, and 18 OECD national or sub-national entities in IALS 1994–1998. Please note that there might be differences between the averages generated by the IDE and the OECD averages for literacy, numeracy and problem solving in technology-rich environments (PS-TRE) published in the PIAAC 2012/2014 OECD and NCES reports. Furthermore, the Average of All Jurisdictions generated by the IDE might differ from previously published results in OECD and NCES reports using PIAAC 2012–2017, ALL 2003–2008, and IALS 1994–1998 data. These differences might be due to the jurisdiction composition of the averages.

#### 3.4.5.2 Linking error

PIAAC 2012–2017, ALL 2003–2008, and IALS 1994–1998 are linked assessments. That is, the sets of items used to assess literacy and numeracy in these years and studies include a subset of common items, referred to as trend items. To establish common reporting metrics for PIAAC, the difficulty of the link items, measured on different occasions, is compared. The comparison of the item difficulties on the different occasions is used to determine a score transformation that allows the reporting of the data on a common scale. As each item provides slightly different information about the link transformation, it follows that the chosen sample of link items will influence the estimated transformation. The consequence is an uncertainty in the transformation due to the sampling of link items, just as there is an uncertainty in jurisdiction means due to the sampling of adults. The uncertainty that results from the link-item sampling is referred to as linking error, and this error must be taken into account when making certain comparisons using the PIAAC 2012–2017, ALL 2003–2008, and IALS 1994–1998 data. As with sampling errors, the likely range of magnitude for the errors is represented as a standard error. Significance tests for scores across years within the IDE take into account the linking errors applicable to each subject.

## 3.5 TALIS IDE

### 3.5.1 Criteria

Each data query must include at least one selection from four criteria choices: subject, education level, year(s), measure(s), and jurisdiction(s). Shown below is an outline of these selection criteria followed by a brief description.

1. Subject:
   * School
   * Teacher
2. Education Level:
   * ISCED 2 (Lower Secondary, default)
   * ISCED 1 (Primary)
   * ISCED 3 (Upper Secondary)
3. Year(s):
   * TALIS 2018 (data available for U.S.)
   * TALIS 2013 (data available for U.S.)
   * TALIS 2008 (data not available for U.S.)
4. Measure(s):
   * Full population estimate
   * Continuous variables from the school and teacher questionnaires, including international variables, derived variables, combined item scales, and U.S. national adaptations and additions to the international questionnaires.
5. Jurisdiction(s):
   * Average of All Jurisdictions
   * Average of Selected Jurisdictions
   * OECD National Entities
   * OECD Sub-National Entities
   * Partners

#### 3.5.1.1 Subject

Only one subject (either school level or teacher level) can be selected at a time in the IDE. Selecting the School option in the drop-down list provides school information that is an attribute of schools (thus estimates are reported, for example, as the “percentage of schools”), while selecting the Teacher option provides teacher or school information that is an attribute of teachers (thus estimates are reported, for example, in terms of the “percentage of teachers”).

#### 3.5.1.2 Education Level

Only one education level (ISCED 1, 2, or 3) can be selected at a time in the IDE. For more information on data availability by ISCED level, please see Jurisdiction(s) below.

#### 3.5.1.3 Measures

You can choose the full population estimate, which is the default measure at each education level in the TALIS IDE, or there are a number of continuous variables that you may choose as a measure of analysis. These continuous variables are from the international and U.S. national teacher and school questionnaires.

#### 3.5.1.4 Years

The TALIS IDE includes data from 2018, 2013, and 2008, the three years in which TALIS was administered to teachers and principals. Some of the variables included across years may differ; for example, the “culture of sharing success” variable under the School Climate and Safety (reported by Principal) subcategory has data available in TALIS 2013, but not in TALIS 2018 or 2008.

When a certain variable is not available for a corresponding year in the TALIS IDE, it will be noted with the symbol “ ”. The participating countries and subnational education systems across years also vary (more information provided under Jurisdiction(s) below).

To select both years of TALIS for analysis, check the box for “All Years.”

#### 3.5.1.5 Jurisdictions

All listed jurisdictions can be selected for any analysis, provided data are available for the selected year of TALIS. Please note the following inclusions and exclusions of TALIS participating country and subnational education system data in the OECD TALIS international reports and the NCES TALIS IDE:

• The Netherlands participated in TALIS 2008 but did not meet the sampling standards. Their data are not included in the TALIS IDE or in the OECD TALIS international report. • Cyprus’s data for TALIS 2013 were included in the OECD TALIS international report but were not made publicly available for use in the data files provided on the OECD website. Cyprus’s data for TALIS 2013 are not included in the TALIS IDE. • Iceland’s data for TALIS 2013 and 2018 were included in the OECD TALIS international report but were not made publicly available for use in the data files provided on the OECD website. Iceland’s data for TALIS 2013, but not for TALIS 2018, are included in the TALIS IDE. • In the OECD TALIS 2013 international report, all ISCED 2 estimates for the United States are shown separately from those for the other participating education systems. This is because the United States did not achieve an acceptable level of response based on the international response rate standards established for TALIS 2013. (To read more about the U.S. response rate, the steps taken to determine the level of bias in the estimates, and caveats about the U.S. data, see <https://nces.ed.gov/surveys/talis/talis2013/index.asp>.) However, in the TALIS IDE, report outputs do not show U.S. estimates separately from the estimates of all other jurisdictions.

### 3.5.2 Variables

In the TALIS IDE, variables are derived from two types of questionnaires: the school questionnaire (answered by school principals) and the teacher questionnaire (answered by teachers). TALIS gives teachers and school principals the opportunity to provide their perspectives on the state of education in their own countries in six reporting areas: (1) Learning environment, (2) Appraisal and feedback, (3) Teaching practices and classroom environment, (4) Development and support, (5) School leadership, and (6) Self-efficacy and job satisfaction. Variables are organized into categories (and subcategories) that have shared characteristics and can be selected as a group when examining and generating tables. Note that variable titles in the TALIS IDE may overlap or be repeated under categories or subcategories, but specific variables appear only once. Some variables might be similar in title and content, but not comparable over the years, either due to differences in the question asked or differences in their response categories. Use *Search* in the *Select Variables* step to locate and select variables in the TALIS IDE.

#### 3.5.2.1 Index Variables

TALIS uses indices derived from the teacher and school questionnaires to contextualize TALIS results and to estimate trends that account for demographic changes over time. Information on the indices available for each year of administration can be found in the chapters referenced in the summary table below.

| Year of TALIS administration | TALIS report chapter | TALIS technical report links |
| --- | --- | --- |
| 2018 | Chapter 11 | [Link](https://www.oecd.org/education/talis/TALIS_2018_Technical_Report.pdf) |
| 2013 | Chapter 10 | [Link](https://www.oecd.org/education/school/TALIS-technical-report-2013.pdf) |
| 2008 | Chapter 11 | [Link](https://www.oecd.org/education/school/44978960.pdf) |

### 3.5.3 Statistics Options

The IDE reports TALIS data with several statistics options:

• Averages • Percentages • Standard deviations • Percentiles

#### 3.5.3.1 Averages

For the TALIS assessment, teacher and principal averages for continuous variables are in the same units as the variables themselves (e.g., average age of teachers). By default, the standard errors of the averages are shown in parentheses. Note that averages will only display in the TALIS IDE if you have selected a continuous variable as a measure.

#### 3.5.3.2 Percentages

Percentages are the default statistic used for analysis in the TALIS IDE. This statistic shows the percentage of teachers or schools as a row percentage. For example, if the first column lists countries, then each country will display its own percentage distribution of teachers or principals across its row. By default, percentage distributions do not include missing data, although there is an option to include them.

#### 3.5.3.3 Standard deviations

The standard deviation is a measure of how widely or narrowly dispersed scores are. Under general normality assumptions, 95 percent of the scores are within two standard deviations of the mean. Thus, if the average score is 35 and the standard deviation is 5, it means that 95 percent of the scores fall between 30 and 40. The standard deviation is the square root of the variance.

#### 3.5.3.4 Percentiles

This statistic shows the threshold (or cut-point) for the following:

• 10th percentile – Bottom 10 percent of teachers or schools • 25th percentile – Bottom quarter of teachers or schools • 50th percentile – Median (i.e., half the teachers or schools reported values below the cut-point and half reported values above it) • 75th percentile – Top quarter of teachers and schools • 90th percentile – Top 10 percent of teachers and schools

### 3.5.4 Cross-tabulations

Cross-tabulation is a method of combining separate variables into a single table. Normally, each variable has its own table. If you have selected two or three variables (not counting *All cases*) and when you go to the *Edit Reports* step, you will automatically get a list with one table for each variable (including one for *All cases*); at the end of that list you will get one cross-tabulation for the two or three variables selected.

If you have chosen four or more variables (not counting *All cases*), you will get tables for each variable, but you won’t get the cross-tabulation. Be advised that if you go back to add another variable without subtracting one to keep the total under four, you will lose any edits you might have made to the cross-tabulation.

### 3.5.5 Statistical Notations and Other Notes

Statistical notations and other notes are found at the end of a data table, as applicable to that table:

* — Not available.
* † Not applicable. (For instance, the standard error for the statistic cannot be reported because the statistic does not meet reporting standards.)
* # The statistic rounds to zero.
* ‡ Reporting standards not met. (For instance, the sample size is insufficient to permit a reliable estimate.)
* NOTE: A general note pertains to any special characteristics of the data in the table.
* SOURCE: Source information is listed for all TALIS data and should be cited when data are used in a publication or presentation.

#### 3.5.5.1 Calculation of Average of All Jurisdictions in the TALIS IDE

For each ISCED level, the *Average of All Jurisdictions* option under *Jurisdiction* within step 1, *Select Criteria*, includes all jurisdictions with data for the year of interest.

#### 3.5.5.2 Statistical Comparisons

In the TALIS IDE, most comparisons are independent with an alpha level of .05, and dependent t tests are performed only for basic gender comparisons by country (with no additional variables included in the analysis). In contrast, reports published by the OECD employ a dependent testing methodology for all gender comparisons by country (i.e., even when additional variables besides gender and country are included in the analysis). Because of this difference, the statistical significance of gender differences by country may vary slightly between published reports and the IDE.

#### 3.5.5.3 Data Suppression

Data suppression may be handled slightly differently in the TALIS IDE and the OECD TALIS international reports. For the IDE, the Rule of 62 is applied to suppress data to avoid reporting results for groups about which little of interest could be said due to lack of statistical power from the data. The Rule of 62 is borrowed from the IDE’s counterpart, the National Assessment of Educational Progress (NAEP) Data Explorer. This rule states that statistics for a group—i.e., means, standard errors, standard deviations, and a set of percentiles—are suppressed if they are based on less than 62 cases. The rule serves to assure a minimum power requirement to detect moderate differences at a nominal significance level (.05). The minimum power is 0.80 and the moderate effect size is 0.5 standard deviation units. A design effect of 2 is assumed to derive an appropriate complex sample standard deviation.

#### 3.5.5.4 ISCED

The International Standard Classification of Education (ISCED) is an internationally comparable method for describing levels of education across countries, created by the United Nations Educational, Scientific and Cultural Organization (UNESCO). TALIS used the ISCED classification system for all administration cycles. ISCED levels are defined as follows:

• Level 0 – The initial stage of organized instruction, designed primarily to introduce very young children to a school-type environment. ISCED level 0 programs can either be center or school based. Preschool and kindergarten programs in the United States fall into the level 0 category. • Level 1 – Consists of primary education, which usually lasts 4 to 6 years. ISCED level 1 typically begins between ages 5 and 7, and is the stage where students begin to study basic subjects, such as reading, writing, and mathematics. In the United States, elementary school (grades 1 through 6) is classified as level 1. • Level 2 – Also known as lower secondary education. Students continue to learn the basic subjects taught at level 1, but this level is typically more subject specific than level 1 and may be taught by specialized teachers. ISCED level 2 usually lasts 2 to 6 years and begins around the age of 11. Middle school and junior high (grades 7 through 9) in the United States are classified as level 2. • Level 3 – Also known as upper secondary education, student coursework at this level is generally subject specific and often taught by specialized teachers. Students often enter upper secondary education at the age of 15 or 16 and attend anywhere from 2 to 5 years. ISCED level 3 can prepare students for university, further schooling, or the labor force. Senior high school (grades 10 through 12) is considered level 3 in the United States. • Level 4 – Consists primarily of vocational education, and courses are taken after the completion of secondary school, although the content is not more advanced than the content of secondary school courses. ISCED level 4 programs in the United States are often in the form of 1-year certificate programs. • Level 5 – Divided into levels 5A and 5B, this level focuses on tertiary education. ISCED level 5A refers to academic higher education below the doctoral level. Level 5A programs are intended to provide sufficient qualifications to gain entry into advanced research programs and professions with high skill requirements. In the United States, bachelor’s, master’s, and first-professional degree programs are classified as ISCED level 5A. ISCED level 5B refers to vocational higher education. Level 5B programs provide a higher level of career and technical education and are designed to prepare students for the labor market. In the United States, associate’s degree programs are classified as level 5B. • Level 6 – Refers to the doctoral level of academic higher education. Level 6 programs usually require the completion of a research thesis or dissertation.

## 3.6 ICILS IDE

### 3.6.1 Criteria

Each data query must include at least one selection from five criteria choices: subject, year(s), measure(s), and jurisdiction(s). Shown below is an outline of these selection criteria followed by a brief description.

1. Display:
   * Student
   * Teacher
2. Measure(s):
   * Computer and Information Literacy (CIL): Overall
   * Computational Thinking (CT): Overall
3. Year(s):
   * 2018 (data available for CIL and CT)
   * 2013 (data available for CIL)
4. Jurisdiction(s):
   * Average of Countries
   * Average of the Selected Jurisdictions
   * Countries
   * Benchmarking participants

#### 3.6.1.1 Display

ICILS participating countries selected a nationally representative sample of 8th-grade students and teachers. As a result, when using the ICILS IDE, you have the option to run either a student- or teacher-level analysis. Selecting the Student option in the display drop-down list provides student or school information that is an attribute of students (thus estimates are reported, for example, as the “percentage of students”), while selecting the Teacher option provides teacher or school information that is an attribute of teachers (thus estimates are reported, for example, in terms of the “percentage of teachers”).

#### 3.6.1.2 Measures

The ICILS IDE includes measures for each display when selected, including overall scales (Student display only) and continuous variables. There are continuous variables other than scale scores that you may choose as a measure of analysis. These variables fall under different categories, such as Population and Student demographics, and include variables such as student age in years, ratio of school size and teachers, and an index of computer experience in years. Some of the continuous variables may have missing values if there are cases that did not answer the assessment or the questionnaire. In both the Student and Teacher displays, selecting the variable “Percentage across full sample” under the Population category will allow you to calculate percentage statistics based on the full sample.

#### 3.6.1.3 Years

Data from ICILS 2013 and 2018 are available in the IDE. Currently, data availability in the IDE is dependent on the measure selected. For example, if the measure chosen is CIL: Overall scale, you can choose one or both years: 2018 and 2013. If the measure chosen is CT: Overall scale, you can choose 2018.

#### 3.6.1.4 Jurisdictions

All listed jurisdictions can be selected for any analyses, provided data are available for the selected year. In 2018, a total of 12 countries and 2 benchmarking participants took part in the ICILS CIL assessment. Of those participants, 8 countries and 1 benchmarking participant opted into the ICILS CT assessment. In 2013, some 18 countries and 3 benchmarking participants took part in the ICILS CIL assessment. Jurisdictions for which data are not available for a selected year are identified by the icon representing “no data”— . Note that the IDE contains a few U.S.-specific background variables (e.g., race/ethnicity, or NATRACE-COLLAPSED) that, when selected, will not yield information for any non-U.S. jurisdictions.

### 3.6.2 Variables

In the ICILS IDE, variables from the student, teacher, principal, and ICT coordinator questionnaires are organized into categories that have shared characteristics. Content category and subcategory titles may overlap, but specific variables appear only once in a subcategory. Use Search in the Select Variables step to locate variables. Note that some variables might be similar in content, but not comparable over the years, either due to differences in the question asked or differences in their response categories. For example, a student background questionnaire variable, “In what country were you and your parents born/Mother or [female guardian]”, which is only available in 2013, is similar but not comparable to the variable “In what country were you and your parents born/[Parent or guardian 1]”, which is only available in 2018. The icon representing “no data”— —will help in identifying the year for which the variable has data available for analysis.

#### 3.6.2.1 Proficiency levels

Achievement results for ICILS are reported using discrete proficiency levels for CIL and CT. Higher levels represent the knowledge, skills, and capabilities needed to perform tasks of increasing complexity. Based on the statistics option chosen, IDE can report the average scores of students at each proficiency level or the percentage of students performing at each of the predefined levels for the chosen jurisdictions. Two statistics options, standard deviations and percentiles, will not generate reports as proficiency levels are not reportable using these statistical analyses. Proficiency levels for any subject should be analyzed with the scale of that same subject; for example, the CIL proficiency levels should be analyzed with the overall CIL scale. Computer and Information Literacy: Administered in 2013 and 2018. In both years, CIL results were reported using four proficiency levels (levels 1–level 4); the IDE shows five categories (below level 1, level 1, level 2, level 3, level 4). Computational Thinking: Administered in 2018. CT results were reported using three proficiency levels (lower region, middle region, upper region); the IDE shows these three categories. Descriptions that characterize typical student performance at each proficiency level are shown in the following tables for CIL and CT. For more information about the development of the proficiency levels, please see the [ICILS 2018 Technical Report](<https://www.iea.nl/sites/default/files/2020-05/ICILS> 2018 Technical Report-FINAL\_0.pdf.

#### 3.6.2.2 Index Variables

In addition to scale scores representing performance in various subjects, ICILS uses indices derived from the student, teacher, principal, and ICT coordinator questionnaires to contextualize ICILS results or to estimate trends that account for demographic changes over time. Information on indices for each year of administration can be found in the IEA publication chapters referenced in the summary table below.

| Year of ICILS administration | ICILS IEA User Guide chapters | Links |
| --- | --- | --- |
| 2018 | Appendix C | [Link](https://www.iea.nl/publications/user-guides/icils-2018-user-guide-international-database) |
| 2013 | Appendix 3 | [Link](https://www.iea.nl/publications/user-guides/icils-2013-user-guide-international-database) |

### 3.6.3 Statistics Options

The IDE reports ICILS data with several statistics options:

• Averages • Percentages • Standard deviations • Percentiles

#### 3.6.3.1 Averages

This statistic provides the average value for a selected continuous variable or overall scale. For the ICILS assessment, student performance is reported on scales that range from 100 to 700. ICILS scales are produced using item response theory (IRT) to estimate average scores for CIL and CT for each jurisdiction. IRT identifies patterns of response and uses statistical models to predict the probability of answering an item correctly as a function of the student’s proficiency in answering other questions. That is, student responses to the assessment questions are analyzed to determine the percentage of students responding correctly to each multiple-choice question and the percentage of students achieving each of the score categories for constructed-response questions.

#### 3.6.3.2 Percentages

This statistic shows the percentage of students as a row percentage. For example, if a categorical variable is selected and the jurisdictions are listed in the table stub, the percentage data for the response categories will sum to 100 percent in each jurisdiction. By default, percentage distributions do not include missing data, although there is an option to include them.

#### 3.6.3.3 Standard deviations

The standard deviation is a measure of how widely or narrowly dispersed scores are for a particular dataset. Under general normality assumptions, 95 percent of the scores are within two standard deviations of the mean. For example, if the average score of a dataset is 500 and the standard deviation is 100, it means that 95 percent of the scores in this dataset fall between 300 and 700. The standard deviation is the square root of the variance.

#### 3.6.3.4 Percentiles

This statistic shows the threshold score (or cut point) for the following:

• 10th percentile – the bottom 10 percent of students • 25th percentile – the bottom quarter of students • 50th percentile – the median (half the students scored below the cut point and half scored above it) • 75th percentile – the top quarter of students • 90th percentile – the top 10 percent of students

### 3.6.4 Cross-tabulations

Cross-tabulation is a method of combining separate variables into a single table. Normally, each variable has its own table. If you have selected two or three variables (not counting *All Cases*) and when you go to the *Edit Reports* step, you will automatically get a list with one table for each variable (including one for *All Cases*); at the end of that list you will get one cross-tabulation for the two or three variables selected.

If you have chosen four or more variables (not counting *All Cases*), you will get tables for each variable, but you won’t get the cross-tabulation. Be advised that if you go back to add another variable without subtracting one to keep the total under four, you will lose any edits you might have made to the cross-tabulation.

### 3.6.5 Statistical Notations and Other Notes

Statistical notations and other notes are found at the end of a data table, as applicable to that table:

* — Not available.
* † Not applicable. (For instance, the standard error for the statistic cannot be reported because the statistic does not meet reporting standards.)
* # The statistic rounds to zero.
* ‡ Reporting standards not met. (For instance, the sample size is insufficient to permit a reliable estimate.)
* NOTE: A general note pertains to any special characteristics of the data in the table.
* SOURCE: Source information is listed for all ICILS data and should be cited when data are used in a publication or presentation.

#### 3.6.5.1 Statistical Comparisons

Comparisons of achievement scores across years are made using independent t tests with a linking error taken into account. Comparisons between jurisdictions are also treated as independent. All comparisons within a jurisdiction, within the same year, are made using dependent t tests. The alpha level for all t tests is .05.

#### 3.6.5.2 Data Suppression

Data suppression may be handled slightly differently in the ICILS IDE and the reports from IEA or NCES. For the IDE, the Rule of 62 is applied to suppress data to avoid reporting results for groups about which little of interest could be said due to lack of power. The Rule of 62 is borrowed from the IDE’s counterpart, the National Assessment of Educational Progress (NAEP) Data Explorer (NDE). This rule states that statistics for a group are suppressed if they are based on less than 62 cases. These statistics are means, standard errors, standard deviations, and a set of percentiles. The rule serves to assure a minimum power requirement to detect moderate differences at nominal significance level (.05). The minimum power is 0.80 and the moderate effect size is 0.5 standard deviation units. A design effect of 2 is assumed to derive an appropriate complex sample standard deviation. In addition, the IDE does not support the calculation of the coefficient of variation.

# 4 Resources

Here are some resources that can help you use the IDE tool and learn more about the international studies supported by the IDE:

* Tutorial videos: The IEA-ETS Research Institute has created a series of tutorial videos that provide a comprehensive walkthrough of the IDE tool. These videos use TIMSS, PIRLS, and PIAAC as examples to illustrate how to use the tool. Watching these videos can be a great way to learn how to use the IDE tool effectively and efficiently.
* IDE Brochure: The IDE Brochure provides a general overview of how to use the IDE tool. It explains the main features of the tool and provides step-by-step instructions on how to access and use data from the international studies supported by the IDE. The brochure is a great resource for those who are new to the IDE tool and want to learn more about how it works.
* Distance Learning modules: The IDE Distance Learning modules provide in-depth information about the international studies supported by the IDE. These modules cover a wide range of topics, including study design, sampling procedures, and data collection methods. They are a great resource for those who want to learn more about the studies and how to use the IDE tool to analyze the data.
* IAP Publications and Products: NCES publishes a wide range of publications and products related to the international studies supported by the IDE. These publications include research reports, technical manuals, and user guides. They are a great resource for those who want to learn more about the studies and the data that is available through the IDE tool.

We welcome your suggestions on how to improve the IDE tool. If you have any feedback or suggestions, please send an email to [NCESinternational@ed.gov](mailto:NCESinternational@ed.gov). We are committed to continuously improving the IDE tool and ensuring that it meets the needs of our users.

1. While the Netherlands’ 2000 data were suppressed for the OECD release of the PISA 2000 results, the United Kingdom’s 2000 data were suppressed retroactively by the OECD after the release of the PISA 2000 results. [↑](#footnote-ref-102)
2. Includes England/Northern Ireland both combined and individually. [↑](#footnote-ref-165)
3. Includes England/Northern Ireland both combined and individually. [↑](#footnote-ref-166)
4. The count of countries and subnational education systems which have data in the PIAAC IDE is different from those listed as participating in the OECD PIAAC International Report. This is due to the omission of Australia, Jakarta (Indonesia), and Russian Federation data from the PIAAC IDE and the inclusion of England and Northern Ireland as individual entities in the PIAAC IDE. [↑](#footnote-ref-167)