

# **Background Information on Education Statistics in the UIS Database**

September 2020



## Objective

This document serves as a reader's guide to help data users understand and interpret the education data disseminated by the UIS. It presents the following information:

- The list of symbols accompanying data.
- The annual data releases cycle.
- The data sources used by the UIS and their updates, including population and economic data used to calculate education indicators.
- Specific country notes.
- Production process of learning assessment-based indicators.

### The following symbols are used:

.. Data not available

(a) Category not applicable

(n) Magnitude nil or negligible

(+) National estimation

(‡) UIS estimation (For regional averages: partial imputation due to incomplete country coverage ranging from 33% to 60% of population depending on the indicator.)

## Data releases

As of September 2018, the UIS has moved to one education data release in September of each year, with new national and regional data. National data are also updated in February, completing the UIS publication of educational data for the round of surveys conducted in the previous reference year.

This release includes national data and regional averages for the school or reference year ending in the previous year and includes data collected from administrative and household surveys, including educational attainment and literacy statistics (every two years), and learning assessments.

Stay informed of the latest data releases by signing up for the UIS email alert service at <http://uis.unesco.org>.

## Data sources

To produce internationally comparable education indicators, the UIS uses several main sources of data as listed below. These data sources differ in their coverage as well as the timing and way in which they are collected.

- **Administrative data** are based on information collected in the management of the education system. These are usually available from EMIS and used by Ministries of Education for



management and planning purposes, and are typically updated on an annual basis. EMIS should typically cover many types of educational paths and levels, including Early Childhood Education (ECE), higher education, and Technical and Vocational Education and Training (TVET). Most of the international monitoring of previous global development agendas (Education For All and the Millennium Development Goals) was based on administrative data produced by countries and compiled by international organizations on school EMIS, which has since been expanded. Although some administrative data, such as school feeding programmes and teachers' salaries, are usually available from non-EMIS sources in Education Ministries, these should ideally be linked to EMIS.

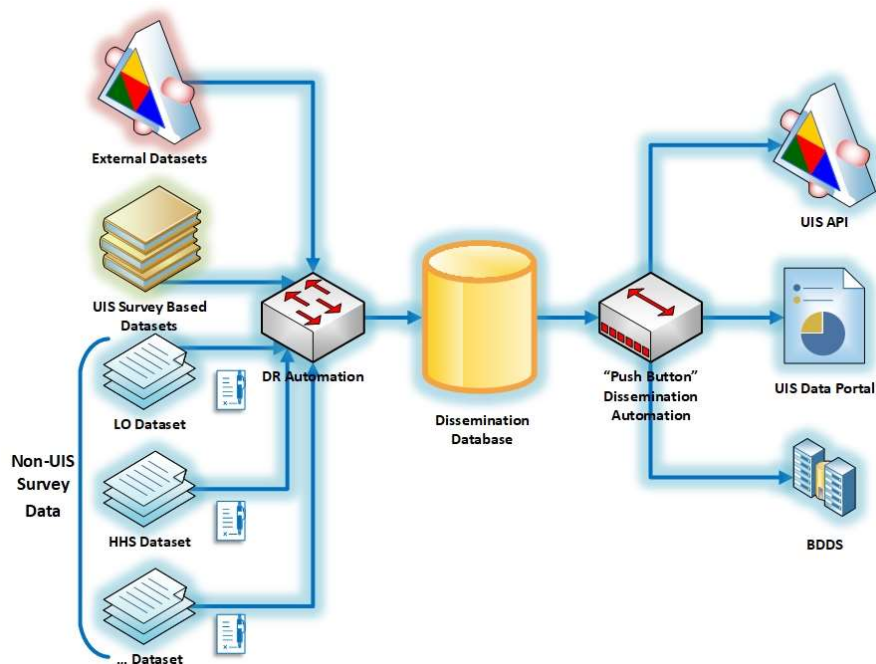
- **Household surveys** are an important source of data on access, participation and educational attainment. Surveys differ in terms of coverage, frequency, objectives and questionnaire design. In contrast to administrative data, they are collected less frequently, and by a variety of organizations and countries. Household survey data are usually available from non-EMIS sources outside of Ministries but should ideally be linked to EMIS.
- **Learning assessments** include national school-based assessments designed to measure specific learning outcomes at a particular age or grade that are considered relevant for national policymakers. They also encompass cross-national initiatives (either regional or global) that are based on a common, agreed-upon framework and follow similar procedures to yield comparable data on learning outcomes. Assessment data can also be collected from households. The data are usually available from non-EMIS sources outside the Ministries but should ideally be linked to EMIS.
- **Financial and expenditure data** contain information on different sources of income and expenditure, including government spending on education. This source commonly encompasses data on the construction and maintenance of schools, teacher salaries and household spending on education, including supplies, transport and other costs. Some administrative data are typically available from non-EMIS sources within Ministries but should ideally be linked to EMIS (e.g. school feeding programmes and teacher salaries), depending on how the country organizes the data.
- **Population censuses** are important as a source of population estimates (the number of persons by age and sex in a country). The UIS uses population estimates from the United Nations Population Division, which are based on a reliable methodology that is internationally accepted. Some of the administrative data are usually available from non-EMIS sources outside the Ministries and again should ideally be linked to EMIS.

**Figure 1** presents the UIS database and dissemination cycle, illustrating the centralized collection of data that originate from different sources. UIS data can be obtained from the following sources:

- [UIS Global Education Database](#), which houses internationally comparable data on all levels of education for more than 200 countries and territories.
- [Bulk data download service \(BDDS\)](#), which enables access to all UIS databases in comma-separated values (CSV) format.



**Figure 1. UIS database production and dissemination cycle**



Source: UNESCO Institute for Statistics.

## The UIS education survey

The UIS collects education statistics in aggregate form from official administrative sources at the national level. Collected information encompasses data on educational programmes, access, participation, progression, completion, literacy, educational attainment and human and financial resources. These statistics cover formal education in public (or state) and private institutions (early childhood education, primary and secondary schools, and colleges, universities and other tertiary education institutions), and special needs education (both in regular and special schools).

These data are gathered annually by the UIS and its partner agencies through the following three major surveys that can be downloaded from the UIS website at <http://uis.unesco.org/en/uis-questionnaires>.

i) UIS Survey of Formal Education

The UIS education questionnaires are sent to UNESCO Member States annually. The questionnaires are based on international standards, classifications and measures that are regularly reviewed and modified by the UIS in order to address emerging statistical issues and improve the quality of data.

ii) UOE Survey of Formal Education

UNESCO-UIS, the OECD and Eurostat (UOE) have jointly administered this annual data collection since 1993. The UOE questionnaire compiles data from high- and middle-



income countries that are generally members or partner countries of the OECD or Eurostat. The UOE survey gathers more detailed education statistics.

### iii) UIS Survey on Literacy and Educational Attainment

The UIS gathers attainment data through its biennial literacy and educational attainment survey, which is sent to all UNESCO Member States. The questionnaires collect internationally comparable data used to calculate literacy and attainment rates as well as mean years of schooling. The data are derived from national population censuses, household surveys and labour force surveys.

The UIS supplements literacy and educational attainment data provided by countries by calculating indicators based on information obtained from the United Nations Statistics Division and international survey programmes such as the Multiple Indicator Cluster Survey (MICS) and the Demographic and Health Surveys (DHS). The UIS also applies methodologies such as the Global Age-specific Literacy Projections Model (GALP) to produce literacy estimates for years with missing data and to produce regional and global averages.

In addition to the three surveys described above, the UIS compiles data for some Sustainable Development Goal 4 indicators directly from partner organizations, such as:

- UNICEF:
  - SDG indicator 4.2.1: Proportion of children aged 24-59 months who are developmentally on track in health, learning and psychosocial well-being, by sex
  - SDG indicator 4.2.3: Percentage of children under 5 years experiencing positive and stimulating home learning environments
- OECD:
  - SDG indicator 4.5.5: Percentage of total aid to education allocated to least developed countries
  - SDG indicator 4.b.1: Volume of official development assistance flows for scholarships by sector and type of study
- Global Coalition to Protect Education from Attack:
  - SDG indicator 4.a.3: Number of attacks on students, personnel and institutions

For the indicators listed below, some national values are calculated by the UIS from household survey and population census data. Observations derived from survey and census data are identified as such in the UIS database.

- SDG indicator 4.1.2: Completion rate (primary education, lower secondary education, upper secondary education)
- SDG indicator 4.1.4: Out-of-school rate (primary education, lower secondary education, upper secondary education)
- SDG indicator 4.2.2: Participation rate in organized learning (one year before the official primary entry age), by sex



- SDG indicator 4.3.1: Participation rate of youth and adults in formal and non-formal education and training in the previous 12 months, by sex
- SDG indicator 4.5.1: Parity indices (female/male, rural/urban, bottom/top wealth quintile and others such as disability status, indigenous peoples and conflict-affected, as data become available) for all education indicators on this list that can be disaggregated
- SDG indicator 4.5.4: Education expenditure per student by level of education and source of funding

## Population estimates

The main source of population estimates is: United Nations, Department of Economic and Social Affairs, Population Division (2019). World Population Prospects: The 2019 Revision, (<https://population.un.org/wpp/Download/Standard/Population/>). Where UN Population Division (UNPD) estimates are not available or where population issues have arisen, national data or UIS estimates are used. Please see the last section of this document for further details.

## Summary of population data sources

The UN Population Division is the standard source of population data at the UIS. For 14 countries, population data from a different source are used to generate more precise indicator values. The exceptions, summarized in the table below, can be classified as follows:

- No UNPD population estimates are available. This is the case for some small countries (Liechtenstein, Monaco).
- Population data by UNPD represent a geographically larger area than education data. Without adjustment, this would lead to an underestimation of population-based indicators. Countries in this group include Azerbaijan, Cyprus, Republic of Moldova, and Serbia.
- Analysis by the UIS revealed important inconsistencies between education data and UNPD population estimates by single year of age that affected the accuracy of population-based indicators. This applies to the remaining countries listed in the table below.

The UIS is currently collaborating with the UN Population Division and the World Bank on guidelines for use of UN or non-UN population data for calculation of SDG indicators, as part of a task team established by the Committee for the Coordination of Statistical Activities (CCSA). The UIS will apply these guidelines, which are expected to be completed after the September 2020 data release, in the future to choose the source of population data for calculation of national indicators.



Source of population data	Number of countries	Criteria used
United Nations Population Division (UNPD): World Population Prospect 2019, released in June 2019	196	UNPD is the main source of population data used by the UIS for all countries-with the exception of the countries below.
Eurostat population database	10	<p><b>Countries for which education data do not cover certain areas but UNPD data do:</b></p> <ul style="list-style-type: none"> <li>- Azerbaijan (excluding Nagorno-Karabakh)</li> <li>- Republic of Moldova (excluding Transnistria)</li> <li>- Serbia (excluding Kosovo)</li> <li>- Cyprus (excluding areas not under the control of the national government)</li> </ul> <p><b>Small European countries for which UNPD did not provide data:</b></p> <ul style="list-style-type: none"> <li>- Liechtenstein</li> <li>- Monaco</li> </ul>
		<p><b>Countries with important inconsistencies between education data and UNPD data that make indicator values implausible.</b> Following recurrent objections from countries, it was decided to use national data instead of UNPD data:</p> <ul style="list-style-type: none"> <li>- France</li> <li>- Ireland</li> <li>- Latvia</li> <li>- Lithuania</li> </ul>
National data collected through UOE data collection	2	<p><b>Countries with important inconsistencies between education data and UNPD data that make indicator values implausible.</b> Following recurrent objections from countries, it was decided to use national data instead of UNPD data:</p> <ul style="list-style-type: none"> <li>- Brazil</li> <li>- United States of America</li> </ul>
National statistical offices submitting directly data to the UIS	2	<p><b>Countries with important inconsistencies between education data and UNPD data that make indicator values implausible.</b> Following recurrent objections from countries, it was decided to use national data instead of UNPD data:</p> <ul style="list-style-type: none"> <li>- Dominican Republic</li> <li>- Singapore (2016 - 2018)</li> </ul>

NB: Data for countries with a population under 90,000 inhabitants provided by the UN Population Division (UNPD)



should be used with caution because the annually interpolated population by age and sex based on 5-year periods and 5-year age groups can only provide crude approximations, especially since key demographic determinants of population change are often unavailable for these small countries.

## Economic statistics

Data on economic indicators such as Gross domestic product (GDP) and Purchasing power parity (PPP) are World Bank estimates as of July 2020. For countries where GDP estimates are not published by the World Bank, data are obtained from the December 2019 release by the United Nations Statistics Division (UNSD). This data release uses the Power Purchase Parties (PPP) conversion factor, GDP (LCU per international \$), of the most recent World Bank International Comparison Program of 2017. Data on total general government expenditure (all sectors) come from the October 2019 release of the International Monetary Fund (IMF)'s World Economic Outlook database.

### A. Technical notes on education data

Education data and indicators are based on the revised International Standard Classification of Education ([ISCED 2011](#)). ISCED 2011 was implemented in the 2014 education data collection and covers the entire formal education system, including both formal initial education programmes and formal adult education programmes. It should be noted that for the years before 2014, the UIS Survey of Formal Education did not cover formal adult education programmes. Data for countries involved in the UOE data collection did not cover formal adult education programmes for the years 2004-2012. Therefore, these limitations should be taken into consideration when comparing time series data.

SDG indicator 4.3.3: Participation rate in technical and vocational programmes (15- to 24-year-olds), by sex. Prior to 2018, data do not include enrolment in technical and vocational short-cycle tertiary programmes (ISCED 5) because the UIS only collected data up to ISCED level 4, although the indicator covers by definition all technical and vocational programmes. As of 2018, the UIS started collecting and incorporated enrolment at ISCED 5 in the calculation of this indicator. On average, participation rates are approximately 1.6% higher with inclusion of ISCED 5 programmes. Caution is required when comparing this indicator over time.

The education regional averages are based on both publishable data and on non-publishable estimated or imputed data. They are calculated based on data as of September 2020.

The literacy regional averages are based on the September 2020 release.

There are cases where an indicator should not exceed a maximum theoretical value (for example the adjusted net enrolment rate), but inconsistencies between demographic and school data may have resulted in the indicator exceeding the theoretical limit. In these cases, "capping" has been applied, while maintaining the same gender ratio. For more details, please refer to the definition of capping in the [UIS Glossary](#).

For enrolment, the percentage of females (% F) is included to provide information on the proportion of girls enrolled with respect to the total enrolment. For gender parity in enrolment, a more relevant measure is the gender parity index (GPI).





Starting with the September 2020 data release, the UIS has replaced all parity indices by adjusted parity indices. Adjusted parity indices are limited to a range of 0 to 2, are symmetrical around 1, and are therefore easier to interpret. Unadjusted parity indices, which were disseminated previously, have no upper bound and are not symmetrical around 1.

## B. Education finance

Expenditure on early childhood education or from international sources – both of which are often comparatively small – have been treated as negligible in cases where data were in fact missing. In these cases, the totals may be underestimated.

## C. Country notes

**Azerbaijan:** Education data do not cover Nagorno-Karabakh, whereas the 2019 Revision of the World Population Prospects data do. The population data used for the calculation of indicators were provided by national authorities and exclude Nagorno-Karabakh.

**Bolivia, Plurinational State of:** Population-based indicators should be interpreted with caution due to concerns regarding the coherence between national enrolment data and population estimates from the 2019 Revision of the World Population Prospects.

The rate of out-of-school adolescents of lower secondary (ISCED 2) and upper secondary (ISCED 3) age does not include data on students enrolled in tertiary education (ISCED 5 to 8). This means the number of out-of-school youth could be overestimated.

**Brazil:** Due to concerns regarding the coherence between national enrolment data and the population estimates from the 2019 Revision of the World Population Prospects, population data were provided by the *Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira* (INEP).

A break in the time series of pre-primary and primary indicators occurs in 2013 due to the following changes in the national education system:

- The duration of pre-primary education changed from 3 years to 2 years.
- The entry age to primary education changed from 7 years to 6 years and the duration from 4 years to 5 years.

In April 2019, there was a change in the methodology used to calculate the Brazilian Population Projections and for this reason the population data for years 2012 to 2018 have been revised by the country. Those changes will be implemented for the February 2021 data release.

**China:** Two Special Administrative Regions – Hong Kong and Macao – are reported separately from data for China.

**Cyprus:** Education data for Cyprus do not cover areas that are not under control of the national government, whereas the 2019 Revision of the World Population Prospects data do. The population



data used for the calculation of indicators were provided by national authorities and only cover the population living in government-controlled areas.

**Dominican Republic:** The National Statistical Office has revised the population estimates and projections that are not yet reflected into the Revision of the World Population Prospects.

**Ecuador:** The source of the education data used is the *Archivo Maestro de Instituciones Educativas* (AMIE). Indicators were calculated with a capping factor and with national population estimates. Therefore, indicators for Ecuador are not comparable with the data officially published by the country.

**Estonia:** The country has an integrated pre-primary education programme. Pupil's age is used as a proxy to disaggregate data for early childhood educational programmes between ISCED level 01 and ISCED level 02: enrolment of 2 years old children and younger is mapped as ISCED level 01 and enrolment of 3-years old children and older is mapped as ISCED level 02.

**France:** There is a change in methodology in 2006 and 2013 that limits the comparability of time series of SDG indicator 4.3.3 on participation rate in technical and vocational programmes (15 to 24 year olds).

**Germany:** Starting from the year 2009, there is no concept of repeaters in grades 1-2 of primary education. This explains a decrease in 2009 in the time series "Percentage of repeaters in primary education".

**Mexico:** A break in the time series of pre-primary indicators occurs in 2014 due to the following changes in the national education system:

- The entry age to pre-primary education changed from 4 years to 3 years.
- The duration of pre-primary education changed from 2 to 3 years.

**Palestine:** Education data do not cover East Jerusalem, whereas the 2019 Revision of the World Population Prospects data do. Population-based indicators may be underestimated and therefore should be interpreted with caution.

**Republic of Moldova:** Education data do not cover Transnistria, whereas the 2019 Revision of the World Population Prospects data do. The population data used to calculate indicators were provided by national authorities and exclude Transnistria.

**Russian Federation:** Starting from 2015, information provided by the Russian Federation includes statistical data for the Autonomous Republic of Crimea and the city of Sevastopol, Ukraine, temporarily occupied by the Russian Federation. The population data used for the calculation of indicators were provided by the UN Population Division and do not cover the Autonomous Republic of Crimea and the city of Sevastopol.

**Serbia:** Education data do not cover Kosovo, whereas the 2019 Revision of the World Population Prospects do. The population data used for the calculation of indicators were provided by Eurostat and exclude Kosovo.



**Singapore:** Education data starting in 2016 refer to residents (citizens and permanent residents) with local addresses and who were away from Singapore for a cumulative period of less than 6 months in the past 12 months prior to the reference date (end-June each year). The data should only be used for the purpose of computing education indicators and should not be compared against other population data published by the Singapore Government.

**Ukraine:** Education data do not cover Crimea and the part of Donbass that is not under government control from 2015 onwards, whereas the 2019 Revision of the World Population Prospects do. Population data excluding these areas are not available for the moment. Consequently, the UIS is not able to produce indicators from 2015 onwards.

**United Republic of Tanzania:** Education data do not cover Zanzibar, whereas the 2019 Revision of the World Population Prospects data do. The population of Zanzibar is approximately 3 per cent of the total population of the United Republic of Tanzania. Population-based indicators are underestimated and should therefore be interpreted with caution.

## Production process of learning assessment-based indicators

For more information on the data production cycle of SDG 4 Indicators, please refer to: [http://tcg.uis.unesco.org/wp-content/uploads/sites/4/2020/08/CN\\_LOS\\_DataProduction-.pdf](http://tcg.uis.unesco.org/wp-content/uploads/sites/4/2020/08/CN_LOS_DataProduction-.pdf) and for metadata information on the following SDG Indicators, use the associated link:

- 4.1.1: <http://tcg.uis.unesco.org/wp-content/uploads/sites/4/2020/08/metadata-4.1.1.pdf>
- 4.1.6: <http://tcg.uis.unesco.org/wp-content/uploads/sites/4/2020/08/Metadata-4.1.6.pdf>
- 4.2.1: <https://unstats.un.org/sdgs/metadata/files/Metadata-04-02-01.pdf>
- 4.2.3: <http://tcg.uis.unesco.org/wp-content/uploads/sites/4/2020/08/Metadata-4.2.3.pdf>
- 4.4.1: <http://tcg.uis.unesco.org/wp-content/uploads/sites/4/2020/08/Metadata-4.4.1.pdf>
- 4.5.2: <http://tcg.uis.unesco.org/wp-content/uploads/sites/4/2020/08/Metadata-4.5.2.pdf>
- 4.6.1: <http://tcg.uis.unesco.org/wp-content/uploads/sites/4/2020/08/Metadata-4.6.1.pdf>
- 4.7.4: <http://tcg.uis.unesco.org/wp-content/uploads/sites/4/2020/08/Metadata-4.7.4.pdf>
- 4.7.5: <http://tcg.uis.unesco.org/wp-content/uploads/sites/4/2020/08/Metadata-4.7.5.pdf>
- 4.a.2: <http://tcg.uis.unesco.org/wp-content/uploads/sites/4/2020/08/Metadata-4.a.2.pdf>