



United Nations
Educational, Scientific and
Cultural Organization



UNESCO
INSTITUTE
FOR
STATISTICS



TECHNICAL
COOPERATION
GROUP



May 2021

SDG 4 Indicators Regional/Global Aggregation Methodology A Brief for the Technical Cooperation Group

This paper is co-authored by Silvia Montoya and Manos Antoninis.

Regional/global aggregates have so far been reported for just 5 of the 12 SDG 4 global indicators in the SDG indicator database (and the statistical annex of the UNSG report on progress towards the SDGs): 4.2.1, 4.2.2, 4.a.1, 4.b.1 and 4.c.1 (although indicators 4.a.1 and 4.c.1 are based on national definitions, which makes them non-comparable and not entirely suitable for regional/global aggregation). However, the 2021 UNSG report will feature for the first time regional/global aggregates for all SDG 4 indicators except 4.3.1. Aggregates are also not reported for indicator 1.a.2 on public education expenditure as share of total government expenditure.

This note has two purposes:

1. It summarizes the methods used to develop regional and global averages for the SDG 4 global indicators with respect to the following criteria: country groupings; general aggregation method; population or other weights used; handling of missing data; criteria to publicly display a regional or global estimate; estimate year; year range used; and modelling to generate yearly estimates (**Table 1**).
2. It identifies outstanding questions that need to be resolved, where possible proposing options for decisions that need to be made. These issues are grouped by type of indicator: learning assessment-based, survey-based, administrative-based, parity indices, and public expenditure (**Table 2**).

In March 2021, UNDESA assembled a task team in order to harmonize and improve existing methodologies, share experiences and best practices in calculating aggregates, provide guidance to other international and regional organizations, and make methodology used transparent. Currently, different methodologies are being used by different international agencies, for instance with respect to the use of population or income estimates, weights, country grouping classifications, and the handling of missing data. As part of the work in progress, the task team has agreed to start mapping custodian agencies' current practices and to do a stock-taking on how the different issues are addressed.

The seventh TCG meeting in October 2020 touched upon the issue of regional/global aggregates but only addressed two partial issues related to aggregates based on survey data:

- whether to publish ranges to reflect uncertainty due to sampling and imputation
- what population weight to use:
 - the size of the three cohorts roughly corresponding to the relevant education level; or
 - the size of the school-age population of the relevant education level

It therefore did not look comprehensively at the full set of issues related to regional/global aggregates for all indicators and the decisions that are needed to report on them.

The following issues can be identified by indicator:

Indicator 1.a.2 (public education expenditure as share of total government expenditure)

- There are several challenges hampering the estimation of aggregates for this indicator (as well as the corresponding indicator with total public expenditure in the denominator):
 - Data are available with a long lag.
 - Some countries have consistently not reported data, which makes it more difficult to impute.
 - With changes in GDP growth rates being correlated across countries, there is strong reason to report regional/global aggregates for the same year.

- It is unclear whether and how the data should be weighted. Arguments can be made both for:
 - using GDP weights; and
 - not using any weights, reporting medians, without imputation.
- However, random data gaps for individual countries in individual years need to be taken into account to estimate robust regional aggregate series.
- One approach for predicting missing values, used in the 2021 Education Finance Watch, is through regression analysis, using historical information on statistical relationships between fiscal, economic and education spending data, controlling for unobservable country and time-specific variations: public education spending as share of GDP was expressed as a function of total government expenditure as share of GDP and GDP per capita, with country and time fixed effects. However, UIS data were combined with data from the World Bank and IMF, which may not ensure comparability among sources and methodologies used for each of them.
- A more robust approach ultimately requires more and better data on both expenditure and GDP. Depending on how gaps are filled, there may or may not be a case for imputation to be pursued.

Indicator 4.1.1 (minimum proficiency level)

- There have been numerous challenges:
 - Data are not available for many low- and lower-middle-income countries.
 - Comparability is based on both the minimum proficiency level and alignment methods.
 - Categories other than sex have different definitions across surveys hindering the comparability of the disaggregation.
- Imputations have been carried out using a very simple model but there is work underway to improve the modelling strategy.

Indicator 4.1.2 (completion rate)

- Several challenges have hampered the estimation of aggregates for survey-based indicators, such as the completion rate:
 - Surveys take place every 3-5 years and their data are made available after 1-2 years.
 - Many countries have multiple surveys, with potentially inconsistent estimates.
 - Survey information has been used inefficiently: reported data reflect the year of the household survey, which is subject to various survey errors, while surveys contain much more information that could be used to establish long-term trends and attenuate the influence of these errors.
- The GEM Report has developed a [model](#) that addresses these issues and produces consistent country time series as well as regional/global aggregate estimates, including a short-term projection to the current year (although this would not be able to estimate the potential negative shock to completion rates caused by COVID-19). These have been used for the 2021 UNSG SDG Report.
- The estimates are weighted by the UNPD data for the closest cohort for each of the three levels of education: 10-14 for primary; 15-19 for lower secondary; and 20-24 for upper secondary completion. An alternative approach would have been the size of the school-age population, which gives a bigger weight to countries with longer duration for a particular level (e.g. to countries with 4-year relative to those with a 2-year upper secondary cycle).
- There is scope for imputing data for missing countries (e.g. estimating missing completion rates as a function of, for instance, enrolment data) but population coverage is in the range of 90% for all regions (with notable gaps for SIDS).

- The model has been extended to also estimate completion rates disaggregated by sex.
- However, as the model is based on retrospective data series, it does not so far allow for disaggregation by location and wealth, which are characteristics that change over time. Further methodological work is being carried out to address this challenge.

Indicator 4.2.1 (children developmentally on track)

- The main challenge in reporting regional/global aggregates for this indicator is its exclusive reliance on one source (MICS), which means low population coverage (31%). Under usual assumptions of minimum 50% population coverage, regional/global aggregates for this indicator can probably not be reported.
- The indicator's methodology has been thoroughly overhauled, which will break the series and means that population coverage will further decline in coming years.

Indicator 4.2.2 (early childhood attendance rate)

- The general methodology for the estimation of regional/global aggregates based on administrative data was presented as an information document at the seventh meeting of the TCG ([UIS, 2020](#)).
- The weight used is the population of children age one year before primary entry age.
- A sequence of imputations is made based on: (i) values for the nearest year; (ii) use of the unweighted group mean; and (iii) manual interventions.

Indicator 4.3.1 (adult education participation)

Indicator 4.4.1 (ICT skills)

- The main challenge in reporting regional/global aggregates for these two indicators is the relatively low population coverage. Currently, no regional aggregates are reported. However, under usual assumptions of minimum 50% population coverage, it should be possible to report some aggregates.

Indicator 4.5.1 (parity index)

- The index 'is calculated from the regional figures of its numerator and denominator' ([UIS, 2020](#)).
- However, an argument can also be made that the unit of measurement should not be the population (weighted mean) but the country (unweighted median). The latter choice may also be justified by the challenging issue of the population weights to be used for characteristics other than sex, such as location and wealth, let alone for characteristics such as migration or language.

Indicator 4.6.1 (literacy proficiency rate)

- The main challenge in reporting regional/global aggregates for these is the relatively low population coverage and the low frequency of data collection

Indicator 4.a.1 (school characteristics)

- The general methodology for the estimation of regional/global aggregates based on administrative data was presented as an information document at the seventh meeting of the TCG ([UIS, 2020](#)).

- The indicator reported is based **on national standards** and will need revision as learning assessments (which have standardized definitions for participating countries) have been approved by TCG and IAEG-SDG as a source for reporting.
- The weight applied is the number of schools at each level of education.
- A sequence of imputations is made based on: (i) values for the nearest year; (ii) use of the unweighted group mean; and (iii) manual interventions.

Indicator 4.b.1 (scholarships)

- This indicator has full coverage but the challenge is that more than one-third of the total volume of scholarships is not assigned to individual countries: therefore the global aggregate is not the sum of regional aggregates. There is no evidence basis for assigning this volume to countries and regions.

Indicator 4.c.1 (trained teachers)

- The general methodology for the estimation of regional/global aggregates based on administrative data was presented as an information document at the seventh meeting of the TCG ([UIS, 2020](#)).
- The indicator reported is based **on national standards**
- The weight applied is the number of teachers.
- Examples 6 and 7 in the methodology refer to imputations made for teachers.

Table 1. Summary of methodologies currently used to report regional and global aggregates for SDG 4 indicators

| | 1.a.2 | 4.1.1 | 4.1.2 | 4.2.1 | 4.2.2 | 4.3.1 or 4.4.1 | 4.5.1 | 4.6.1 | 4.a.1* | 4.b.1 | 4.c.1* |
|--|-----------------|--|--|--|---|----------------------------------|---|--|---|---|---|
| STATUS | | | | | | | | | | | |
| SDG database | Not reported | Not reported | Not reported | Reported (World, SSA) | Reported | Not reported | Not reported (except 4.c.1) | Not reported | Reported | Reported (LAC, SSA) | Reported |
| UIS database | Not reported | Not reported | Not reported | Not reported | Reported | Not reported | Reported for 4.2.2 | Not reported | Reported | Not reported | Reported |
| 2021 SDG Report | Not reported | Reported | Reported | Reported (2020 only, 31% cover) | Reported | Not reported | Reported for 4.1.1, 4.2.2 and 4.c.1 | Reported (2020 only) | Reported | Reported | Reported |
| 2020 GEM Report | Reported | Not reported | Reported | Reported for some regions | Reported | Reported for some regions | Reported | Reported for some regions | Reported | Reported | Reported |
| CHARACTERISTICS | | | | | | | | | | | |
| Regional groupings | — | SDG, M49, World Bank | SDG, M49, World Bank | SDG, M49, World Bank | SDG, M49, World Bank | — | SDG, M49, World Bank | SDG, M49, World Bank | SDG, M49, World Bank | SDG, M49, World Bank | SDG, M49, World Bank |
| General method | — | Population-weighted administrative data | Population-weighted modelled estimates | Population-weighted survey data | Population-weighted administrative data | — | Population-weighted administrative data (applied to underlying indicators, not the index) | Population-weighted | Population-weighted administrative data | Sum | Population-weighted administrative data |
| Population or other weights | — | 4.1.1a and b: primary enrolment; 4.1.1c: lower secondary enrolment. For disaggregation: enrolment by sex. | Cohort size 10-14 primary 15-19 lower s 20-24 upper s | (...) | Population of age one year before primary entry | — | Same weight used as for the underlying indicators | Population aged 15 to 64 available in 5-y increments; weighted average of preceding and proceeding years used for in-between years | Total number of schools by level | No weights | Total number of teachers by level |
| Handling of missing data and implied assumption | — | Data for latest year with non-missing values for indicator and weight were used. Countries without either were excluded. No imputation was used. | Country has regional value | (...) | Sequence of imputations: nearest year, auxiliary data, unweighted group mean, manual. | — | Imputation is applied to underlying indicators, not to parity indices | Data for latest year with non-missing values for indicator and weight were used. Countries without either were excluded. No imputation was used. | Sequence of imputations: nearest year, auxiliary data, unweighted group mean, manual. | No missing data but large share of total volume not assignable to countries | Sequence of imputations: nearest year, auxiliary data, unweighted group mean, manual. |

| | 1.a.2 | 4.1.1 | 4.1.2 | 4.2.1 | 4.2.2 | 4.3.1 or 4.4.1 | 4.5.1 | 4.6.1 | 4.a.1* | 4.b.1 | 4.c.1* |
|---|-------|--|---|-------|---|----------------|------------------------------|--|---|-----------------------------|---|
| Criteria to display regional or global estimates: population coverage | — | No criteria were applied due to low data coverage, particularly in low and middle income countries | 80% | (...) | Data for ≥60% of population: publishable Data for <60% and ≥33% of population: published as UIS estimate. Data for <33% of population: not published. | — | Same as underlying indicator | No criteria were applied due to low data coverage, particularly in low and middle income countries | Data for ≥60% of population: publishable Data for <60% and ≥33% of population: published as UIS estimate. Data for <33% of population: not published. | No missing data | Data for ≥60% of population: publishable Data for <60% and ≥33% of population: published as UIS estimate. Data for <33% of population: not published. |
| Estimate year | — | Latest year for each country | Current year (modelled) | (...) | 2 years before current year | — | Same as underlying indicator | Latest year for each country | 2 years before current year | 2 years before current year | 2 years before current year |
| Years represented / Year ranges used to calculate regional estimates with survey data (no modelling) | — | Most recent country data in last X years | Most recent country data in last 5 years modelled and projected to current year | (...) | Most recent country data in last 5 years | — | Same as underlying indicator | Latest years are represented | Any country data used if average distance from estimate year is ≤4 years | Country data two years back | Any country data used if average distance from estimate year is ≤4 years |
| Modelling: Short description including covariates | — | No | ABC model | No | No | — | No | No | No | No | No |
| Other methods used | — | No | No | No | No | — | No | No | No | No | No |

Note: * national values are based on national standards.

Explanation of each methodological characteristic for regional/global aggregates used in **Table 1**:

| | |
|--|---|
| Regional groupings | <p><i>SDG/M49</i>: World; Sub-Saharan Africa; Northern Africa and Western Asia (Northern Africa; Western Asia); Central and Southern Asia (Central Asia; Southern Asia); Eastern and South-Eastern Asia (Eastern Asia; South-Eastern Asia); Latin America and the Caribbean; Oceania (Australia and New Zealand; Oceania exc. Australia and New Zealand); Europe and Northern America (Europe; Northern America)</p> <p>Landlocked Developing Countries Least Developed Countries Small Island Developing States</p> <p><i>World Bank</i>: low-, lower-middle-, upper-middle- and high-income countries</p> |
| General method | <ul style="list-style-type: none"> • Population-weighted survey data • Population-weighted administrative data • Population-weighted modelled estimates using survey and/or administrative data as input data (often with a few covariates) to generate yearly estimates for countries • Modelled regional estimates (currently used for some malnutrition estimates but trying to retire this method) • Use of weights other than population, or no weights • Other |
| Population or other weights | <ul style="list-style-type: none"> • Population data source: WPP for the most part? • Population groups: usually as relevant to the indicator (e.g. U5, Births, 15-49) |
| Handling of missing data and implied assumptions | <p><i>Missing country level data</i> (= no data available for a country)</p> <ul style="list-style-type: none"> • Assume country value is zero • Assume country has regional value, i.e. assumed to have the weighted average of all countries with data in the region • Estimate country value, e.g. use other data source to generate estimate, use other variables to estimate • <i>Missing data for reporting year</i> • Assume latest value available is current; when out of range: include; treat as missing; extrapolate • Estimate country-year value: use available data from earlier years to estimate more recent value based on observed country trends or indicator (global/regional) trends • <i>Any other adjustments or imputations</i> |
| Criteria to display estimates: population coverage | <p>Minimum percentage of population covered for estimates to be presented</p> <p><i>Exceptions</i>: Certain big countries lacking data may be excluded from the region's total population (e.g. Brazil, China, India, Russian Federation)</p> |
| Estimate year | <p>Aggregates reported with a X-year lag, e.g. estimates published in 2020 are for 2019</p> |
| Years represented / Year ranges used (no modelling) | <p>Aggregates reflect most recent data available for a country over a range of years, e.g. data represent a six year period: for 2019, survey data between 2014-19</p> <p>Exceptions applied for big countries where the most recent data available is used to generate regional aggregates even if the latest data point is outside the year range</p> |
| Modelling: short description including covariates | <p>In general, most models try to fit country data. For indicators that are modelled, it is important to understand the model methods, including covariates that may influence the estimate for countries with very sparse or no data.</p> |

Table 2. Summary of issues to address on regional/global aggregates for SDG 4 indicators

| Issues | 4.1.1 and learning assessment data-based indicators | 4.1.2 and survey data-based indicators | 4.2.2 and administrative data-based indicators (e.g. 4.1.4) | Parity indices | 1.a.2 and finance indicators |
|--|---|---|---|--|---|
| Regional groupings | | | | | |
| General method | | | | | |
| Population weights | Choose between: <ul style="list-style-type: none"> school age population i.e. UIS data enrolment | Choose between: <ul style="list-style-type: none"> cohort size 10-14 year old for primary 15-19 year old for lower sec 20-24 year old for upper sec (=unbiased by level duration) school age population i.e. UIS data (=consistent with OOS) | | Choose between: <ul style="list-style-type: none"> median of countries i.e. average over individual countries' parity indices mean of populations i.e. aggregate populations and divide the two groups | Choose between: <ul style="list-style-type: none"> unweighted average (mean or median) of country-level indicators weighted average of country-level indicators by GDP/total government expenditure |
| Handling of missing data and implied assumption | Choose between: <ul style="list-style-type: none"> assume country has regional value impute missing values based on other information | Choose between: <ul style="list-style-type: none"> assume country has regional value impute missing values based on other information | | | Choose between: <ul style="list-style-type: none"> Use regression model expressing the indicator as a function of total government expenditure as share of GDP and GDP per capita, with country and time fixed effects to impute missing values Sequential imputation (UIS): (i) use indicator values for country from nearest year (ii) use unweighted regional mean of national indicator values (iii) manual estimation No imputation |

| Issues | 4.1.1 and learning assessment data-based indicators | 4.1.2 and survey data-based indicators | 4.2.2 and administrative data-based indicators (e.g. 4.1.4) | Parity indices | 1.a.2 and finance indicators |
|---|---|--|---|---|--|
| Criteria to display regional or global estimates: population coverage | | | | Choose between: <ul style="list-style-type: none"> • 50% of countries • 50% of population | Choose between: <ul style="list-style-type: none"> • UIS methodology: <ol style="list-style-type: none"> i. Publishable national data for $\geq 60\%$ of GDP / total public expenditure in PPP\$ in region: publish without qualifier. ii. Publishable national data for $< 60\%$ and $\geq 33\%$ of GDP / total public expenditure in PPP\$ in the region: publish as UIS estimate. iii. Publishable national data for $< 33\%$ of GDP / total public expenditure in PPP\$ in region: not published. • Publishable national data for at least 50% of GRP / total public expenditure in region: publish without qualifier. |
| Estimate year | | | | | |
| Years represented / Year ranges used to calculate regional estimates with survey data (no modelling) | | | | | |
| Modelling: Short description including covariates | | | | | |
| Other issues | | | | | |
| Confidence intervals | Choose between: <ul style="list-style-type: none"> • confidence intervals (+how) • no confidence intervals | | | | |
| Disaggregation | Choose between: <ul style="list-style-type: none"> • aggregates for urban/rural and bottom/top quintile (+what weights) • no aggregates for urban/rural and bottom/top quintile | | | | |

ANNEX 1 -Proposed methodology for regional averages for SDG indicator 1.a.2: Decision points for the TCG

This paper is co-authored by Saïd Ould Ahmedou Voffal and Friedrich Huebler, UNESCO Institute for Statistics.

The UIS is proposing a methodology to calculate regional averages for the education component of SDG indicator 1.a.2, “proportion of total government spending on essential services.” The proposed calculation method, described in the Annex 1.2, is based on established methodology that is also used for other SDG indicators. It includes imputation methods for countries with missing data based on auxiliary information, as well as built-in reliability checks to assess whether a value is publishable.

The regional proportion of general government expenditure on education (all levels of education combined) as a percentage of total general government expenditure in a region is calculated as the sum of education expenditure in all countries in a region, divided by the sum of total government expenditure in a region. For the calculation, amounts expressed in local currency units are converted to international (US) dollars using a purchasing power parity conversion factor.

Decision points

1. Is the proposed methodology technically sound? Yes/No
2. Which of the following parameters should be used as population weight:
 - a. Mean of country-level indicator values weighted by total government expenditure
 - b. Simple unweighted average of country-level indicators values
3. Does the TCG approve the following sequential process to handle missing data: Yes/No
 1. Use indicator values for country from nearest year
 2. Use unweighted group mean of country-level indicator values across countries in region
 3. Manual estimation of national indicator values
4. Which of the following criteria on the population coverage does the TCG recommend for display of regional or global estimates:

Option 1 (current UIS methodology): Regional values are published based on the following criteria:

 - a. Publishable national data for $\geq 60\%$ of total general government expenditure in PPP\$ in the region: indicator published without qualifier.
 - b. Publishable national data for $< 60\%$ and $\geq 33\%$ of total general government expenditure in PPP\$ in the region: indicator published as UIS estimate.
 - c. Publishable national data for $< 33\%$ of total general government expenditure in PPP\$ in the region: indicator not published.

Option 2: Regional values are published if publishable national data represent at least 50% of total government expenditure in the region.

Option 3: Regional values are published if publishable national data represent at least 50% of total GDP in the region.
5. Does the TCG propose any changes to the methodology? Yes/No
6. Does the TCG approve implementation of the methodology by the UIS? Yes/No

ANNEX 1.2 - SDG indicator 1.a.2: proportion of total government spending on essential services (education): Proposed methodology for regional averages

The UIS is the custodian agency for the education component of SDG indicator 1.a.2, “Proportion of total government spending on essential services (education, health and social protection).” This document proposes a methodology for regional averages for the education component of this indicator, based on established UIS methodology for calculation of regional aggregates.

Indicator definition

The indicator is defined as total general (local, regional and central) government expenditure on education (current, capital, and transfers) regardless of the source of funding, expressed as a percentage of total general government expenditure on all sectors (including health, education, social services, etc.).

Calculation of national indicator values

Total general government expenditure for all levels of education combined is divided by total general government expenditure (all sectors), and the result is expressed as a percentage.

$$PXE_{c,t} = \frac{XE_{c,t}}{TX_{c,t}}$$

where

$PXE_{c,t}$ = total general government expenditure on education (all levels of education combined) as a percentage of total government expenditure, in country c , in financial year t

$XE_{c,t}$ = total general government expenditure on education (all levels of education combined) in country c , in financial year t

$TX_{c,t}$ = total general government expenditure in country c , in financial year t

Calculation of regional aggregates

The regional proportion of general government expenditure on education (all levels of education combined) as a percentage of total general government expenditure in a region is calculated as the sum of education expenditure in all countries in a region, divided by the sum of total government expenditure in a region. For the calculation, amounts expressed in local currency units are converted to international (US) dollars using a purchasing power parity conversion factor.

$$PXE_{r,t} = \frac{XE_{r,t}}{TX_{r,t}} = \frac{\sum_{c,r} \frac{XE_{c,t}}{PAN_{US,PPP,c,t}}}{\sum_{c,r} \frac{TX_{c,t}}{PAN_{US,PPP,c,t}}}$$

where:

$PXE_{r,t}$ = total general government expenditure on education (all levels of education combined) as a percentage of total general government expenditure in region r , in financial year t

$XE_{r,t}$ = total general government expenditure on education (all levels of education combined) in region r , in financial year t

$TX_{r,t}$ = total general government expenditure in region r , in financial year t

$XE_{c,t}$ = total general government expenditure on education (all levels of education combined) in country c of region r , in financial year t

$TX_{c,t}$ = total general government expenditure in country c of region r , in financial year t

$PAN_{US,PPP,c,t}$ = purchasing power parity conversion factor (for GDP), measured in local currency units per international (US) dollar, in country c of region r , in financial year t

Weight

The regional average is weighted by total general government expenditure (expressed in PPP\$) in each country in a region.

Imputation of missing data

The sequence of imputations for national values is as follows:

1. value of $XE_{c,t}$, $TX_{c,t}$ and $PAN_{US,PPP,c,t}$ from nearest year (if available),
2. unweighted group mean of $PXE_{c,t}$ across all countries in a region with data,
3. manual estimation.

Criteria for publication by UIS

Regional values for indicator 1.a.2 are published by the UIS based on the following criteria:

1. Publishable data for $\geq 60\%$ of total general government expenditure in PPP\$ in the region: indicator published without qualifier.
2. Data for $< 60\%$ and $\geq 33\%$ of total general government expenditure in PPP\$ in the region: indicator published as UIS estimate.
3. Data for $< 33\%$ of total general government expenditure in PPP\$ in the region: indicator not published.

Range of years with data considered during calculation

Any national data are used, as long as the average distance from the reference year is ≤ 4 years.