USING LABOUR FORCE SURVEYS TO PRODUCE EDUCATION INDICATORS

2024

2024 CONFERENCE ON
EDUCATION DATA
AND **STATISTICS**

Table of contents

1. Introduction

LFS and SDG4 indicators

Whether performed annually, quarterly, or continuously, household survey data is crucial for obtaining most of the timely and relevant SDG4 indicators. Furthermore, survey data may supplement, and sometimes even replace, administrative statistics on school attendance in nations, particularly for adult engagement in formal and non-formal education. To calculate SDG indicators 4.3.1 (Participation rate of youth and adults in formal and non-formal education and training in the previous 12 months, by sex), 4.3.3 (Gross enrolment ratio for tertiary education, by sex), and 4.4.3 (Youth/adult educational attainment rates by age group and level of education), the fifth meeting of the TCG, held in Mexico in November 2018, focused on using labour force surveys as the most comprehensive data source.

The majority of nations throughout the world undertake labor force surveys (LFS) in accordance with ILO guidelines and definitions. Given that, each household-based survey is designed to measure different social phenomena, the magnitude of measurement errors in estimates obtained for the same indicators might vary significantly from country to country and within different estimates over time within a country. In addition, the availability of population estimates impacts the precision of the estimates, such as, for example, participation rate of the youth/adult population in schooling. While the administrative data sources collected by national governments are coupled with population data from the United Nations Population Division, for the survey data, the population on which rely the data weighting procedure is the same as the population frame used in the sampling process. This does not mean that survey estimates are completely precise, on the contrary, there are measurement errors associated with them, but the magnitude of these errors is lesser than when combining administrative data and population estimates. This document's goal is to analyze how labour force survey data is used to calculate various SDG 4 indicators and to provide some insight into the benefits and drawbacks of this data source.

SDG 4.3.1 – participation rate of youth and adults in formal and non-formal education and training in the previous 12 months

One of the main sources for the calculation of SDG 4.3.1 is the household-based survey data compiled by the Department of Statistics of the International Labour Organization (ILO). ILO maintains a global database of national LFS-s or other relevant household-based surveys that cover labour market but contain valuable information on participation in education and training disaggregated by sex and age. While countries that conduct LFS following the internationally agreed concepts and definitions as regards labour market main indicators, there is not the same for the participation in education and training. For the calculation of SDG 4.3.1 there are three components to be in compliance with the UIS definition.

First, the distinction between "formal education and training" and "non-formal education and training". The information in some countries is collected only for formal education, and no other data source is available for non-formal education and training. Most LFS questionnaires restrict the scope of non-formal education and training to job-related courses and training. Another problem with LFS data is that it

excludes the category "guided-on-the-job training," which accounts for a significant portion of non-formal education and training. As a result, data comparability between countries is not feasible.

Second, the reference period required for the calculation of SDG 4.3.1 is the previous 12 months. Regarding the reference period for which the information is collected, there is a significant divergency between countries. The most common reference period in LFS questionnaire is "the last four weeks proceeding the date of interview". Participation in formal and non-formal education and training over the previous 12 months is thought to give a more complete measure of adult learning than one based on a 4-week reference period. For countries that conduct EU-LFS, starting from the reference year 2022, data on participation in education and training refer to the last 12 months. The periodicity of data collection with a "last 12 months" reference period is biennial. That means that EU-LFS will provide estimates in even years, 2022, 2024, etc.). For other countries, the information on education variables is collected for the current situation, which means that there is not clear if "currently" covers the last week, the last four weeks, the last 12 months, or the current academic year.

Table 1: LFS Reference period by SDG region

SDG region	Reference period				No of questionnaires
	4 weeks	*12 months*	*Currently*	*Not specified*	
Central and Southern Asia	0	1	4	1	**6**
Eastern and South-eastern Asia	0	0	5	5	**10**
Europe and Northern America	36	26	0	1	**38**
Latin America and the Caribbean	0	2	10	11	**23**
Northern Africa and Western Asia	5	2	3	3	**12**
Oceania	1	1	4	0	**6**
Sub-Saharan Africa	0	2	17	15	**34**
Total	42	34	43	36	**129**

Table 1 shows the different types of reference periods used by countries in their LFS questionnaires. There are analyzed the questionnaires of 129 countries (the latest version of available questionnaires used over the period 2018-2022). As shown in Table 1, the reference period used for the education variables differs from region to region. In countries of Latin America and Sub-Saharan Africa the reference period is either "currently" or not specified.

SDG 4.4.3 Youth/adult educational attainment rates by age group and level of education (thematic indicator)

According to the UIS definition, the educational attainment of an individual is defined as the highest ISCED level completed by the individual. Educational attainment is measured with respect to the highest education programme successfully completed, which is typically certified by a recognised qualification. This indicator measures for each level of education the percentage of the population who completed at least that level (UIS metadata). Thus, it is a cumulative percentage. The LFS is the most common data

source for information on educational attainment, and the ILO database of national LFS-s or other relevant HH-surveys collects this information disaggregated by age and sex. In this database, the highest level of education completed is classified according to the ISCED-11 and ISCED-1997 (for a limited number of countries), and it provides the correspondence table between the aggregated levels of education. For countries that do not conduct LFS regularly, other household surveys can be used to calculate educational attainment rates since the related question is almost standardized for all household-based surveys.

SDG 4.3.3 Participation rate in technical and vocational programmes (15–24-year-old)

According to the UIS definition, the youth participation rate in technical and vocational programmes is the percentage of young people aged 15 to 24 who are enrolled in technical or vocational education, either it be formal or non-formal education, on a given date or during a specified period. LFS data are a potential data source for obtaining estimates for this indicator in cases when the information on school participation is collected by ISCED levels.

The methodological issues related to SDG indicator 4.3.1

The methodological issues related to SDG indicator 4.3.1 have been in focus of TCG since the fourth and fifth meetings in 2018. In the fourth TCG meeting, the development of methodology for Indicator 4.3.1 is presented. Then, the recommendations on definitions, methodology and formulation of questions for global indicator 4.3.1 (drafted by Manos Antoninis and Lotta Larson) were consulted by members of an working group (17 countries) established by the TCG Secretariat.

The agreed recommendations from working group 1 (TCG5/Ref/4: Indicator 4.3.1, 2018, pg. 4) regarding the methodological decision were as following:

1. Adopt an alternative classification of formal and non-formal education and training that can also be used for indicators 4.3.3 and 4.3.6.

Figure 1: Proposed classification for SDG Indicator 4.3.1
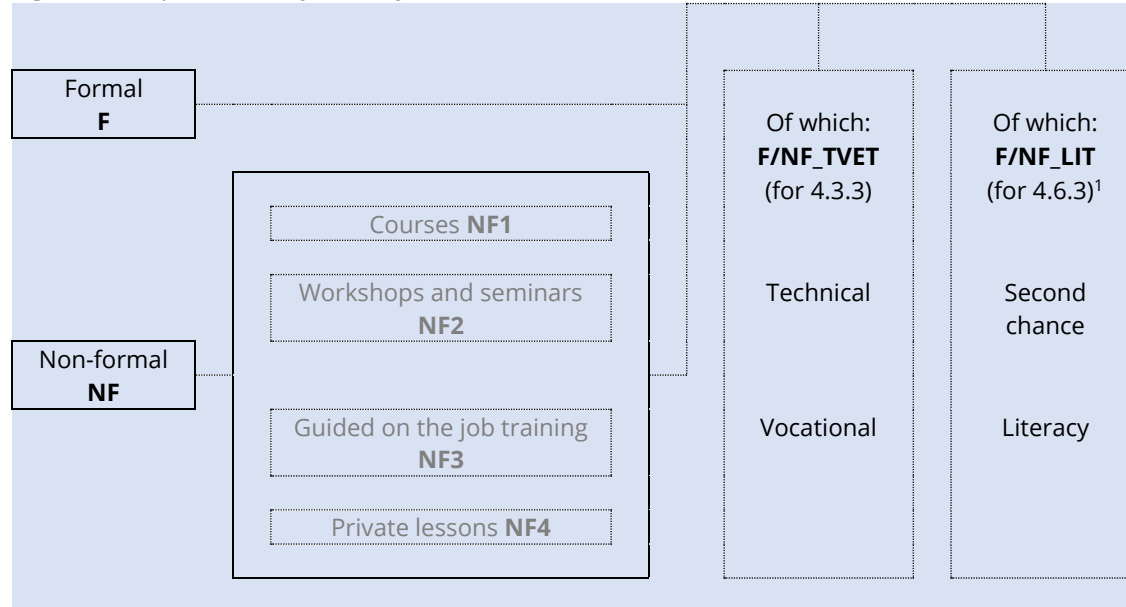
Source: UIS, TCG5/Ref/4: Indicator 4.3.1, 2018

2. Refrain from setting a minimum duration for non-formal programs.
3. Adopt the widely used intervals of 15-24 years (youth) and 25-64 years (adults).
4. Adopt a single household/labour force survey as data source for the global indicator at the country level.
5. Consider alternative sources if the adoption of a single survey results in low country coverage for the indicator.
6. Adopt a standard formulation for the relevant questions to be included in surveys.

What is the status of play five years after the acceptance of these recommendations?

R1. While the Working Group on Indicator Development agreed upon the proposed classification of FE and NFE, the implementation of this classification is not straightforward. Countries have designed their household questionnaires according to their national needs, where the variables on education serve mostly for describing the profile of the interest groups (for example, the educational profile of unemployed persons / of poor, etc.). In the analysis of 129 different questionnaires for 129 countries (of which 112 are LFS questionnaires and 17 are HIES questionnaires), retrieved by ILO webpage on data

---

1 The 9th meeting of TCG approved to discontinue the reporting on SDG indicator 4.6.3 (the participation rate of illiterate youth/adults in literacy programs) since it has not been reported for the last 11 years and the coverage rate for this indicator in the UIS database in the period from 2010 to 2022 is 0%.

catalogue, only 76 countries have questions that collect information on NFE. European countries have a better methodological coverage of Indicator 4.3.1. They collect information for both FE and NFE (see Table 1). Starting from the reference year 2022, the EU-LFS data will collect biennially FE and NFE data based on a 12-months reference period. Half of Sub-Saharan African countries collect information for the NFE, but only four of them make the distinction between FE and NFE. For Latin America and the Caribbean countries, only nine countries have questions on NFE, mostly formulated for the training related to any occupation.

Table 2: Number of LFS and HIES questionnaires with NFE questions

SDG Region	Number of questionnaires /countries	Has information on NFE	FE distinguished from NFE	LFS	HIES/HS
Central and Southern Asia	6	2	1	5	1
Eastern and South-Eastern Asia	10	2	2	10	
Latin America and the Caribbean	23	9	8	20	3
Northern America and Europe	38	38	38	38	
Oceania	6	1	1	4	2
Sub-Saharan Africa	34	17	4	23	11
Western Asia and Northern Africa	12	7	5	12	
Total	**129**	**76**	**59**	**112**	**17**

Source: Elaborated by the author based on countries questionnaires in ILO data catalogue:
https://www.ilo.org/surveyLib/index.php/catalog/?page=1&ps=15

As mentioned in the previous section, for European countries, the Adult Education Survey instead of LFS is preferred to obtain the estimates for Indicator 4.3.1. There are two issues to be further elaborated. First, while EU-LFS and AES have harmonized the methodology on measuring FE and NFE, the estimates differ significantly; even for the FE (one should expect less bias for FE). That brings in question which survey to rely on. Second, a more detailed analysis is needed to explore the bias created by these two data sources. On the one hand, the exploitation of LFS data for the calculation of SDG 4.3.1 ensures a better global geographical coverage, but the participation in the non-formal education and training seems to be underestimated. On the other hand, while AES covers better the participation in non-formal education and training, the likelihood of overestimating it is quite high. Therefore, given the fact that EU-LFS and AES are two surveys with fully harmonized methodology and can serve as a model for other countries, a working group could help elaborate the factors that cause the estimates bias of each source.

R2. For the second agreed recommendation regarding Indicator 4.3.1 there is no elaboration at this stage.

R3. Regarding the adaption of age intervals for youth and adults, UIS has progressed a lot. The estimated for Indicator 4.3.1 are disaggregated sex and by following age groups: 15 to 24 (youth), 25-54 years old, 55-64 years old, 15-64 years old, and 15 years and above.

R4 and R5. The adoption of a single household/labour force survey as data source for the global indicator 4.3.1 is realized by using the ILO database. The use of LFS and other household-based surveys has affected positively the rate of coverage for indicator 4.3.1. Table 2 presents the rate of coverage for this indicator. There is a significant improvement for countries that belong to Central and Southern Asia (an increase of the rate of coverage from 14 % in 2029 to 64 % in 2023), and for countries in Sub-Saharan Africa as well (See Table 3).
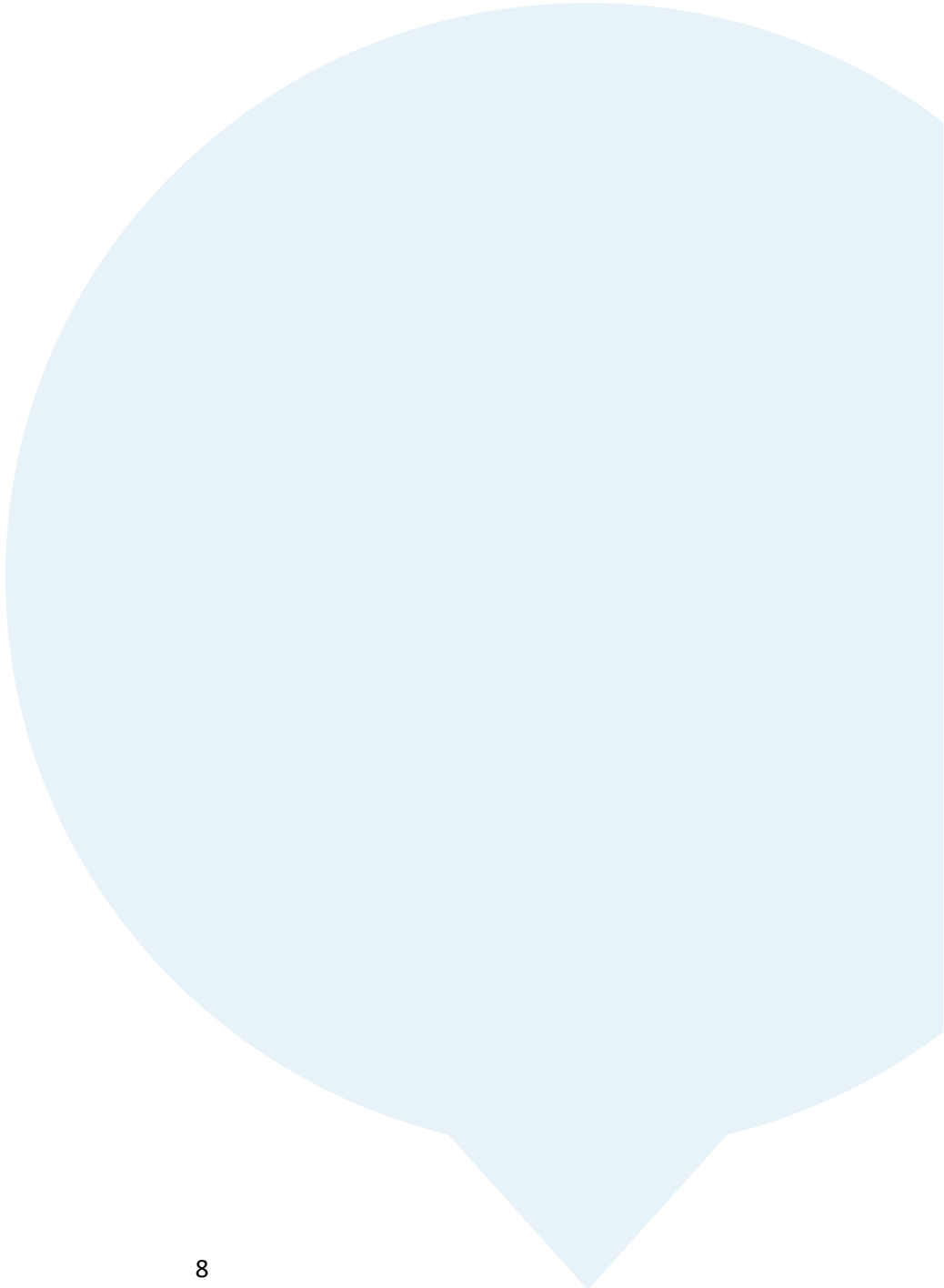
Table 3: Rate of coverage SDG Indicator 4.3.1

SDG regions	2019 (February)	2023 (based on reported estimates for the period 2018-2022)
Central and Southern Asia	14%	64%
Eastern and South-Eastern Asia	61%	56%
Latin America and the Caribbean	35%	62%
Northern America and Europe	76%	87%
Oceania	32%	30%
Sub-Saharan Africa	33%	75%
Western Asia and Northern Africa	41%	54%
World	**45%**	**66%**

Source: UIS Database, February 2019 (TCG6-WD-3-UIS-data-coverage-report) and September 2023 release.

Despite the improvement of the rate of coverage, the methodological nuances remain persistently. Using LFS data from the ILO database is beneficial in terms of having a better geographical coverage. Although the definition of indicator 4.3.1 does not strictly require the distinction between FE and NFE, the ILO data miss the information on what is included in the participation in learning activities. Therefore, it is necessaire to have an update of LFS inventory for identifying countries with underestimated indicator 4.3.1 caused by the lack of questions on NFE. As comparability across countries is of a great importance, the LFS (HS) inventory will provide more metadata related to the indicator.

For countries that do not conduct LFS on regular basis, other household surveys fill the gap. For example, as shown in table 1, the Household Income and Expenditure Survey data are a good source for countries in Sub-Saharan Africa. That has increased the rate of coverage from 33% to 75%.

R6. A questionnaire should be delivered to all countries aimed to consult and assess their plans for changing the formulation of questions.

2. Methodological challenges

Using LFS data for estimating indicators 4.3.1, 4.4.3 and most likely 4.3.3 fills the gap in data reporting. It significantly improves the geographical coverage since almost all nations carry out labour force or other household-based surveys. However, regarding methodology, heterogeneity exists at a large scale across countries, making data comparability and accuracy less feasible. Below is a list of challenges associated with the production of SDG4 indicators.

- Because participation in education and training is not the primary focus of the LFS, respondents may underestimate their participation in education and training, and the interviewer may require

more time to insist on reformulating the related questions to properly collect education information.

- For the indicator 4.3.1, data comparability across countries is not always feasible due to methodological heterogeneity regarding to the reference period, and the distinction between formal and non-formal education and training. In the case of European countries, this indicator is estimated based on Adult Education Survey (AES), conducted every five years. AES definitions and concepts are fully aligned with the UIS definition for the indicator 4.3.1 and the reference period is last 12 months. The AES estimates on adult participation in education and training are higher than LFS ones.
- The set of variables available in the LFS differs from country to country. While in general, for some countries, such as those in EU-LFS, there is a dedicated module to education and training, for others, the number of questions is limited and placed in the household roster. Therefore, the information collected can most probably be used only for the calculation of SDG 4.3.1 and 4.4.3.
- Lack of adequate countries' LFS metadata (some countries have limited information that covers only the main definitions of employment and unemployment, and there are not given details about education data and classification).
- The translation and wording of questions are other aspects that might create estimate bias, affecting both data accuracy and comparability.

3. Advantages of using LFS data for the calculation of some SDG 4 indicators

Despite the methodological challenges related to the use of LFS data for the calculation of indicators 4.3.1, 4.3.3, and 4.4.3, labour force surveys provide and important source of information on education and training (participation and educational attendance).

- LFS data allows the disaggregation by age group and sex. Other variables included in the LFS questionnaire for most countries are disability, degree of urbanization (urban or rural area), migration, and type of school provider (public or private).
- Countries coverage is higher when using LFS data since labour surveys are conducted in more than 60% of countries (according to the ILO LFS data catalogue).
- Data collection periodicity: Most countries conduct LFS regularly (on a monthly, quarterly, continuous, or annual basis). Therefore, it offers the advantage of having timely estimates and creating time series data for monitoring the trends and developments regarding the participation in education and training for different populations of interest.
- Flexibility to adapt to international standards: LFS is aligned with statistically agreed international standards such as ISCED-11.
- Equity measure in education – LFS data collected regularly offer the possibility to measure the education progress of the most marginalized populations.

4. Proposal for a better use of LFS data

The UIS efforts to produce timely, reliable, comparable, and accurate data on SDG 4 indicators are challenged not only by the lack of relevant data sources but also by the methodological peculiarities of the same data source, such as LFS. Labour force or other household-based surveys collect information on the participation of youth and adults in learning activities. Although LFS primarily intends to collect labour-market information, it may be utilised to get at least two SDG indicators: 4.3.1 and 4.4.3. However, regarding compliance with UIS concepts and definitions, SDG indicator 4.3.1 is more complex since it requires measuring the participation in education and training (formal and non-formal) in the previous 12 months. The complexity is related to the LFS methodological heterogeneity regarding the reference period used by countries and the collection of data for non-formal education and training.

A harmonised implementation of the concept of "reference period" is needed. The 12-month reference period significantly impacts measuring participation in non-formal education and training. It generates a more comprehensive measure of participation in education and training since a longer reference period allows for capturing more education and training activities and is less influenced by seasonal effects.

Another challenge is harmonising questions in the LFS that cover participation in formal and non-formal education and training. Most countries have formulated LFS questions consistent with the concepts in the classification of learning activities. Education variables are harmonised in European countries on the Adult Education Survey and EU-LFS. It would be recommended that the wording of questions be the same. For example, "regular" education can be changed to "formal" education. The difficulty is in understanding the concept of non-formal education and training. Splitting the general question into as many questions as types of non-formal education and training can improve the coverage of NFE. A few short questions help respondents better understand the NFE concept. In the fifth TCG meeting, a set of questions is proposed after consultation with different countries (UIS TCG5-Ref4, 2018).

The UIS definition of SDG Indicator 4.3.1 does not require calculating the participation disaggregated by formal and non-formal education and training, but it implies including them in the calculation. However, the comparability of the indicator is mainly affected by the fact that not all countries collect information on non-formal learning activities. Given the lack of such information, there are countries where this indicator is underestimated and covers only formal education. A "relaxed" definition of indicator 4.3.1 would be a solution in such circumstances. Thus, the indicator can be disaggregated by type of education (formal and non-formal). Calculating three indicators instead of one might be cumbersome, but it allows both data comparability and accuracy. The proposed indicators are:

- Participation rate of youth/adult in formal education and training in the previous 12 months
- Participation rate of youth/adult in non-formal education and training in the previous 12 months
- Participation rate of youth/adults in formal and non-formal education and training in the previous 12 months (the existing indicator).

However, if the data is not provided by such disaggregation, then detailed metadata would help understand the accuracy of this indicator. It would be recommended to make an explanatory note in the tables and the related metadata, indicating whether the SDG Indicator 4.3.1 includes only formal education or formal and non-formal education and training.

EU countries collect information on formal and non-formal education and training using two main tools, EU-LFS (conducted continuously) and Adults Education Survey (conducted every five years). The aim of collecting it is to estimate the participation of adults in lifelong learning. Although standardization of the methodology on the measurement of education indicators, there still exist great discrepancies in estimates of SDG indicator 4.3.1. Consequently, these differences pose another challenge: Which data source is more reliable, and which one should be used to report this indicator? While the AES complies with the definition of UIS for the types of education as well as for the reference period, the EU-LFS uses the 4-week reference period and excludes "guided on-the-job training". EU countries are reporting the SGD indicator 4.3.1 based on AES. However, since for the majority of countries, the estimates are calculated based on LFS data, it would be more reasonable to use EU-LFS data as well for EU countries. Another solution is to keep estimates from both sources, but in order to avoid confusion among users, the estimates should always be accompanied with detailed metadata.

References

2018. "TCG5-Ref4: SDG Indicator 4.3.1." Consultation on Definitions, Methodology and Formulation of Questions". November 2018, Mexico.
https://tcg.uis.unesco.org/wp-content/uploads/sites/4/2019/01/TCG5-REF-4-indicator-4.3.1.pdf

2018. "TCG4: Development of SDG Global Indicator 4.3.1." Report TCG4/12 for January 2018 meeting of Technical Cooperation Group on the Indicators for SDG 4–Education 2030 (TCG). Montreal: UIS.
http://tcg.uis.unesco.org/files/resources/meetings/4th/TCG412%20Development%20of%20Indicator%204.3.1.pdf

2019. UIS Report on SDG 4 indicators data coverage, TCG6/WD/3. https://tcg.uis.unesco.org/wp-content/uploads/sites/4/2019/08/TCG6-WD-3-UIS-data-coverage-report.pdf

ILO Survey Catalogue: https://www.ilo.org/surveyLib/index.php/catalog/?page=1&ps=15