UNSD-DFID PROJECT ON SDG MONITORING

MODULE 5

Drafting Guidance

Module structure

- Guidance on drafting definitions
- Guidance on drafting section on methodology
- Guidance on drafting section on data sources
- Others

Drafting the metadata

- If you are using the E-handbook or other materials as a reference to draft elements of the metadata please ensure that they are adapted to your actual circumstances.

Definitions

Guidance on drafting definitions and key terms

- Provide descriptive information on the definition of the concepts associated with;
  - the indicator (i.e. births, disease, etc.)
  - any classifications used (i.e. industry, financial, environmental, rural/urban, occupations, age groups etc.)
- Often includes specific examples of what is and is not included in particular categories.

Example

Indicator 6.1.1; UNICEF

Definition and key terms

Proportion of population using safely managed drinking water services is defined as the proportion of population using an improved drinking water source which is accessible on premises, available when needed, and free of faecal (and priority chemical) contamination.

Key terms:

- 'Improved' drinking water sources include: piped water into dwelling, yard or plot; public taps or standpipes; boreholes or tubewells; protected dug wells; protected springs; packaged water (bottled, sachet); delivered water (tanker trucks, small cart) and rainwater.
- A water source is considered to be 'accessible on premises' if the point of collection is within the dwelling, yard, or plot.
- 'Available when needed': households report being able to access sufficient quantities of water when needed.
- Free from faecal and priority chemical contamination': *E. coli* or thermotolerant coliforms are the preferred indicator for microbiological quality, and arsenic and fluoride are the priority chemicals for global reporting.

Example

8.a.1; OECD

The Development Assistance Committee (DAC) defines *ODA* as those flows to countries and territories on the DAC list of ODA recipients and multilateral institutions which are:

(1) Provided by official agencies, including state and local governments, or by their executive agencies; and

(2) Each transaction of which:

a. is administered with the promotion of the economic development and welfare of developing countries as its main objective; and
b. is concessional in character and conveys a grant element of at least 25 percent (calculated at a rate of discount of 10%).

All donors refer to DAC donors, other bilateral providers of development cooperation and multilateral organizations.

Aid for Trade is captured in the OECD's Creditor Reporting System (CRS) as follows:

– Economic infrastructure (transport and storage (CRS codes 210xx), communications (CRS codes 220xx) and energy (CRS codes 230xx)),

– Trade policy and regulations and trade-related adjustment (CRS codes 331xx)

– The trade development policy marker, which identifies trade development/activities which have trade development as an explicit objective within the "building productive capacity" category which is defined as banking and financial services (CRS codes 240xx) , business and other services (CRS codes 250xx), agriculture, forestry, fishing (CRS codes 31xxxx), industry (CRS codes 321xx), mineral resources and mining (CRS codes 322xx), and tourism (CRS codes 332xx)),

See for reference: http://www.oecd.org/dac/financing-sustainable-development/development-finance-standards/dacandcrscodelists.htm

Tips for compilation

- Define key concepts and components of the indicator, including its further disaggregations (where necessary)
- Keep it succinct and brief as possible
- Make reference to international or national publications on sector standards & classifications where possible.
- Best to cite these rather than just providing a URL (as the URL could change), i.e. *Author, Year Published, Publication Title, Publisher [website]*

Methodology

Guidance on drafting methodology

- Provide descriptive information on any calculations made with the source data to produce the indicator

- Including formulas, adjustments and weighting particularly where mixed sources are used or where the calculation has changed over the time (i.e. discontinuities in the series).

- References to documentation related to various aspects of the data, such as detailed methodological documents or papers covering concepts, scope, classifications and statistical techniques.

Example

Indicator 8.1.1

Method of Computation and Other Methodological Considerations

Computation Method:

In order to calculate the annual growth rate of real GDP per capita the following steps need to be taken:

Step 1: Divide the real GDP for year *t* and *t+1* with the total population of the country in the respective years to get the real GDP per capita for the two consecutive years.

$$G_t = \frac{Real\ GDP\ year\ t}{Total\ population\ in\ year\ t}$$

$$G_{t+1} = \frac{Real\ GDP\ \ year\ (t+1)}{Total\ population\ in\ year\ (t+1)}$$

where,

G_t is the real GDP per capita for the year *t;* and

G_{t+1} is the real GDP per capita for the year *t+1*.

Step 2:The annual growth rate of real GDP per capita is then calculated as follows:

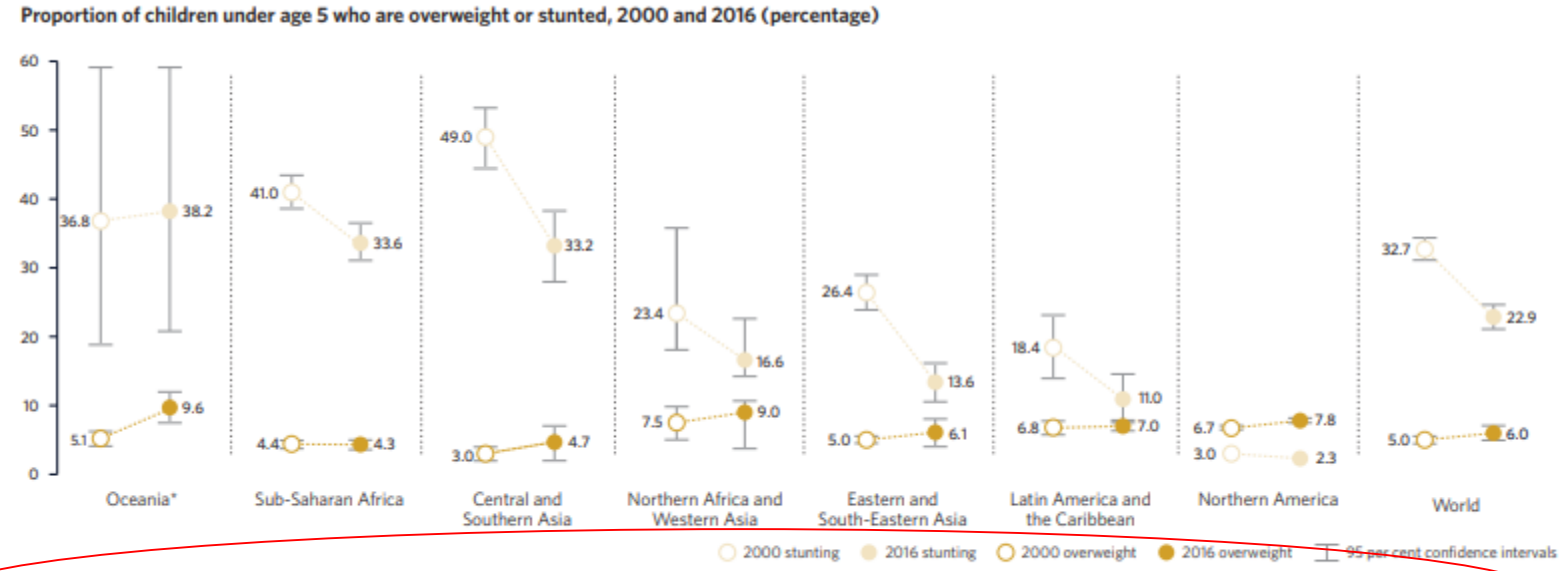
$$= \frac{[G_{t+1} - G_t]}{G_t} \times 100$$

Comments and limitations:

Although countries or areas calculate GDP using the common principles and recommendations in the United Nations System of National Accounts (SNA), there are still issues with the international comparability of GDP estimates due to countries potentially using different versions of the SNA (e.g. 1968, 1993, or 2008) or different degrees of coverage of informal and non-observed economic activities in the GDP estimates.

Example

Indicator 2.2.1, 2.2.2



Note: Central and Southern Asia 2016 overweight data had consecutive low population coverage. Northern America regional average for stunting is based only on United States data; hence, confidence intervals are not available.

Example

Indicator 8.4.2

Method of Computation and Other Methodological Considerations

Computation Method:

It is calculated as direct imports (IM) of material plus domestic extraction (DE) of materials minus direct exports (EX) of materials measured in metric tonnes.

DMC= IM + DE- EX

UN Environment is publishing a global manual which should be released by July 2018. This global manual is based largely on the EUROSTAT Economy Wide Material Flow Accounting compilation guide 2013. MFA accounting is also part of the central framework of the System of integrated Environmental-Economic Accounts (SEEA). More information on data sources and compilation is provided in the below section.

DMC should be divided by the GDP and the population of the country to get standardized measures such as DMC per capita and DMC per GDP.

Comments and limitations:

DMC cannot be disaggregated to economic sectors due to the fact that there is no way to link the consumption of materials to a particular economic sector.

Example

United States, Indicator 6.1.1

Description of actual indicator available	Describes the percent of US population that receives drinking water from community water systems in compliance with drinking water standards.
Method of computation	The EPA Office of Ground Water and Dr inking Water (headquarters) calculates this measure using data reported in the Safe Drinking Water Information System - Federal (SDWIS - FED) and provides the results to EPA regions. This measure includes federally - regulated contaminants of the following viola tion types: Maximum Contaminant Level, Maximum Residual Disinfection Limit, and Treatment Technique violations. It includes any violations from currently open and closed community water systems (CWSs) that overlap any part of the most recent four quarters.
Comments and limitations	EPA classifies Public Water Systems (PWS) into 3 types according to the number of people they serve, the source of their water, and whether they serve the same customers year-round or on an occasional basis.

Tips for drafting

- Keep brief and succinct as possible.
- Use notation and formula only where necessary
- Make reference to international or national documents on detailed methodology adopted where possible.
- Best to cite these rather than only providing a URL (as the URL could change), i.e. *Author, Year Published, Publication Title, Publisher [website]*

Data sources

Guidance on drafting section on data sources

- Provide descriptive information on the source.

- For example with a census/ survey source the following should be described;

  - the sample frame used & coverage
  - type of interview conducted
  - dates/ duration of fieldwork
  - sample size & coverage
  - response rate
  - history of the source (including breaks in series)
  - details of denominator (if from a different source)

Guidance for admin data

Important that data is methodologically transparent to users and accompanied by:

- Name of the source agency. If more than one specify whether data are from all agencies, or only from the main agency or ministry;
- Description of the purposes data were originally compiled and collected by the agency;
- An outline of the strengths and weaknesses of the data in terms of the statistical application.

Guidance for admin data

- Impact of issues relating to coverage, differences in concepts (i.e. non-standard classifications) and the use of unit concepts different from statistical ones.

- Description of processing or transformation (if any) undertaken by the statistical agency.

- Descriptions of the reliability of the data, including adherence to international norms and standards and caveats / limitations on the statistical use(s) of the data, e.g. for social indicator generation.

Example

Metadata from Cambodia

A summary of the data source was provided

Indicator name	Proportion of married women using modern birth spacing methods
Contact point in responsible ministry/agency	Ms. Khol Kemarary Chief of Statistics Office Department of Planning and Health Information Ministry of Health [redacted] Website: http://moh.gov.kh Mr. Youk Dararith Vice Chief of Statistics Office Department of Planning and Health Information Ministry of Health [redacted] Website: http://moh.gov.kh
Definition	It refers to the use of modern methods of contraception among married women of reproductive age in a country
Method of computation	Number of married women of reproductive age using modern methods divided by total number of married women of reproductive age.
Process of obtaining data	Cambodia Demographic and Health Survey (CDHS), is a household survey covering many areas related to the demographic and health situation (especially women's health) within the population. The survey has been conducted jointly by the National Institute of Statistics (NIS) and Ministry of Health (MoH) in 2000, 2005 and 2010. The CDHS is a sample survey and the 2000, 2005 and 2010 surveys collected data from a representative sample of 10,000 plus households. All women age 15-49 in these households and all men aged 15-49 in a sub-sample of one-half of the households were eligible to be individually interviewed. The questionnaire is conducted in three parts for household, women and men in the household supplemented by blood collection for HIV and hemoglobin testing.
Treatment of missing values	The weights for each survey observation are determined by the sampling design, design weights, and adjusted for non-response and other imperfections such as under coverage or, adjusted to improve the precision of estimates.
Comments and limitations	While expanded in content the 2010 survey is a successor to the 2005 and 2000 surveys and directly comparable.
Data availability	Every 5 years
Level of estimates	Estimates are available at the national, urban, rural and provincial level.
Expected time of release	October

Example

Indicator 1.1.1 global metadata

Data Sources and Collection Method

The indicator is produced globally using micro-level data on household income or consumption expenditures from nationally representative household surveys, which is reported to the World Bank's development research group and/or the ILO (for working poverty). Only nationally representative surveys that contain sufficient information to produce a comprehensive consumption or income aggregate (including consumption or income from own production) and allow for the construction of a correctly weighted distribution of per capita consumption or income are used.

Examples of surveys include household income and expenditure surveys (HIES), living standards measurement surveys (LSMS) with employment modules, or labour force surveys (LFS) that collect information on household income. Such surveys also offer the benefit of allowing the employment status and income (or consumption expenditure) variables to be derived from the same sampled households ideally for the same long observation period.

Comments and limitations /discrepancies/disaggregation

Guidance on drafting comments and limitations

- Describes comments and limitations of the indicators including issues such as:
  - Comparability with international data
  - Disaggregation dimensions, if possible
  - presence of wide confidence intervals

- Provides further details on additional non-official indicators commonly used together with the official SDG indicator.

- Sources of reasons for discrepancy between data and metadata used for national and global monitoring

Example

Indicator 8.5.2

Comments and limitations	The questions noted above were added to the CPS in June 2008 to identify persons with a disability in the civilian noninstitutional population age 16 and older. The collection of these data is sponsored by the Department of Labor's Office of Disability Employment Policy. Statistics based on the CPS are subject to both sampling and nonsampling error. When a sample, rather than the entire population, is surveyed, there is a chance that the sample estimates may differ from the true population values they represent. The component of this difference that occurs because samples differ by chance is known as sampling error, and its variability is measured by the standard error of the estimate. There is about a 90-percent chance, or level of confidence, that an estimate based on a sample will differ by no more than 1.6 standard errors from the true population value because of sampling error. BLS analyses are generally conducted at the 90-percent level of confidence. The CPS data also are affected by nonsampling error. Nonsampling error can occur for many reasons, including the failure to sample a segment of the population, inability to obtain information for all respondents in the sample, inability or unwillingness of respondents to provide correct information, and errors made in the collection or processing of the data. Additional information

Example

Indicator 1.1.1

Comments and limitations:

In making international comparisons of poverty estimates, there are conceptual and practical problems to address. Potential problems include the following:

- Internationally comparable lines are useful for producing global aggregates of poverty. However, such a universal line is generally not suitable for the analysis of poverty within a country. For that purpose, a country-specific poverty line needs to be constructed that reflects the country's economic and social circumstances, and adjusted for different locations such as rural and urban areas.
- The reliability of poverty estimates using the international poverty line is significantly influenced by the underlying PPP data, national consumer price indices and their production timelines. Therefore, comparison across contries may not be accurate in terms of needs deprivation.
- Differences in the relative importance of consumption of non-market goods may affect poverty rate estimates. The local market value of all consumption in kind (including own production) should be included in total consumption expenditure. Similarly, imputed profit from the production of non-market goods should be included in income.
- This indicator measures poverty based on household per capita income/consumption, ignoring intra-household inequality in the distribution of resources, and does not take into account other dimensions of poverty such as vulnerability, people's feeling about relative deprivation and lack of voice and power of the poor.

Example

Indicator 2.2.2

Disaggregation

Disaggregated country data are available in a majority of household surveys, and typically include sex, age groups, household wealth, mothers' education, residence. UNICEF's expanded databases include disaggregated data.

Common pitfalls

Poor quality data are unfortunately all too common. Accurate estimates of stunting, overweight and wasting rely on accurate measurement of height and weight as well as child's age. Surveys with field personnel who are not well trained or well supervised may yield poor quality data, and so the global household survey programmes such as MICS and DHS not only provide detailed guidelines on training and fieldwork implementation but also run specific data quality checks on the collected data in order to assess data quality.

Example

Disaggregation

Disaggregation by place of residence (urban/rural) and socioeconomic status (wealth) is possible for nearly all countries for basic services and may be possible for safely managed services in future. Disaggregation by other stratifiers of inequality (subnational region, gender, education, disadvantaged groups, etc.) is possible in some countries but these are generally not available from administrative sources. Wherever possible sanitation services will also be disaggregated by the JMP by service level (including no services, basic, and safely managed services) following the sanitation ladder.

Discrepancies with national estimates: JMP estimates are based on national sources of data approved as official statistics. Differences between global and national figures arise due to differences in indicator definitions and methods used in calculating national coverage estimates. In some cases national estimates are based on the most recent data point rather than from regression on all data points as done by the JMP. In some cases national estimates draw on administrative records of infrastructure coverage rather than the nationally representative surveys and censuses used by the JMP which collect information directly from households.

Example

Explaining the difference between global and national metadata

Global Metadata

This table provides information on metadata for SDG indicators as defined by the United Nations Statistical Commission. Complete global metadata documentation on all indicators in Goal 8, unless otherwise noted, is provided by the UN Statistics Division.

See UN metadata for 8.6.1

SDG Indicator Name	Proportion of youth (aged 15-24 years) not in education, employment or training
SDG Target Addressed	By 2020, substantially reduce the proportion of youth not in employment, education or training.
Definition of SDG Indicator	The NEET is defined as the percentage of youth (15-24 years old) who are not in employment and not in education or training.
UN Designated Tier	1
UN Custodial Agency	ILO

U.S. Metadata

This table provides metadata for the actual indicator available from U.S. statistics closest to the corresponding global SDG indicator. Please note that even when the global SDG indicator is fully available from U.S. statistics, this table should be consulted for information on national methodology and other U.S.-specific metadata information.

Method of computation for global SDG indicator	
Graph Title	Percent of US population 16 to 24 years who are not enrolled in school and are either unemployed or not in the labor force
Actual indicator available	Percent of population 16 to 24 years who are not enrolled in school and are either unemployed or not in the labor force
Description of actual indicator available	Percentage of civilian noninstitutional population 16 to 24 years who are not enrolled in school and are either unemployed or not in labor force

Summary

To recap

- Be aware of the target audience
- Use clear and simple language
- Keep sentences and paragraphs short
- Avoid technical terms, jargon and acronyms
- Use a standard glossary of terms
- Develop style guide for how data / metadata are presented
- Ask colleagues to review the data and metadata – do they make sense?

Thank you for your attention!