Food and Agriculture Organization of the United Nations

SUSTAINABLE DEVELOPMENT GOALS

Guidelines on data disaggregation for SDG Indicators using survey data

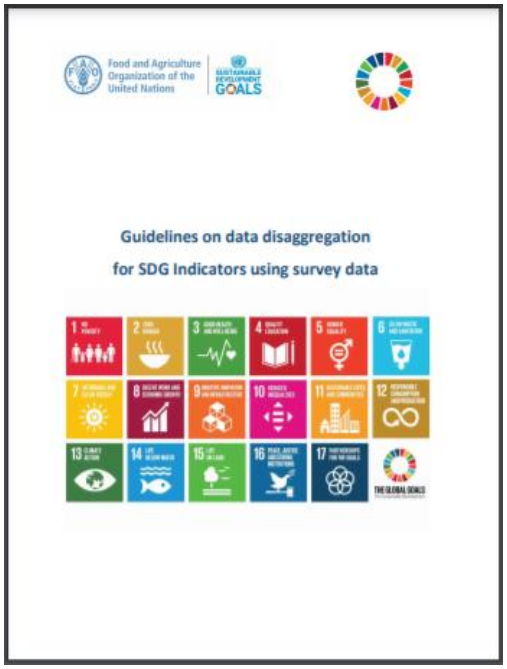
Open Virtual IAEG-SDG Meeting – Leave no one behind

Data Disaggregation for SDGs

March 25, 2021

Clara Aida Khalil– FAO, Office of the Chief Statistician

Guidelines on data disaggregation for SDG Indicators



Main objectives of the Guidelines:

- Offer methodological and practical guidance for the production of direct and indirect disaggregated estimates of SDG Indicators based on survey data.

- Provide tools to assess estimates accuracy and present strategies for data integration, including small area estimation (SAE) methods

Relevance of the guidelines

- Approximatively **30%** of the Global SDG Indicators are based on survey data

- **7 out of 21** SDG Indicators under FAO custodianship can be computed using data from household and agricultural surveys.

ISSUE ADDRESSED: The use of traditional sampling techniques imposes limitations on the production of disaggregated data and reliable estimates for small sub-populations.

Innovative techniques that could address some of these issues are far from being mainstreamed in National Statistical Offices.

Data disaggregation with sample surveys

Main idea of the guidelines:

Direct estimates of an indicator for a given sub-population: based only on sample information from the sub-population itself. **Two main issues:**

- Sampling size often not large enough to guarantee reliable estimates for small domains;
- Possibility of having non sampled sub-domains.

These issues can be addressed:

- **At design stage:** adopting sampling designs that guarantee an observed set of sampling units for every sub-population for which disaggregated data must be produced.
- **At the analysis stage:** producing indirect estimates, coping with the little information available for "small areas" by borrowing strength from other sources of data.

Addressing data disaggregation at the design stage

The guidelines illustrate **alternative sampling strategies** for direct domain sampling estimation:

- Most common domain estimators are discussed, introducing the context of their usability.
- Model-assisted and model-based estimation approaches are illustrated.
- Mainstreamed and innovative approaches to address data disaggregation at sampling design stage are discussed (with their pros and cons):
  - **Oversampling**
  - **Deeper stratification**
  - **Multiphase sampling with screening of respondents**
  - **Marginal stratification designs**
  - **Indirect sampling**
- Suitable software packages are suggested

Assessing estimates accuracy

Tools to assess the accuracy of direct estimates are provided:

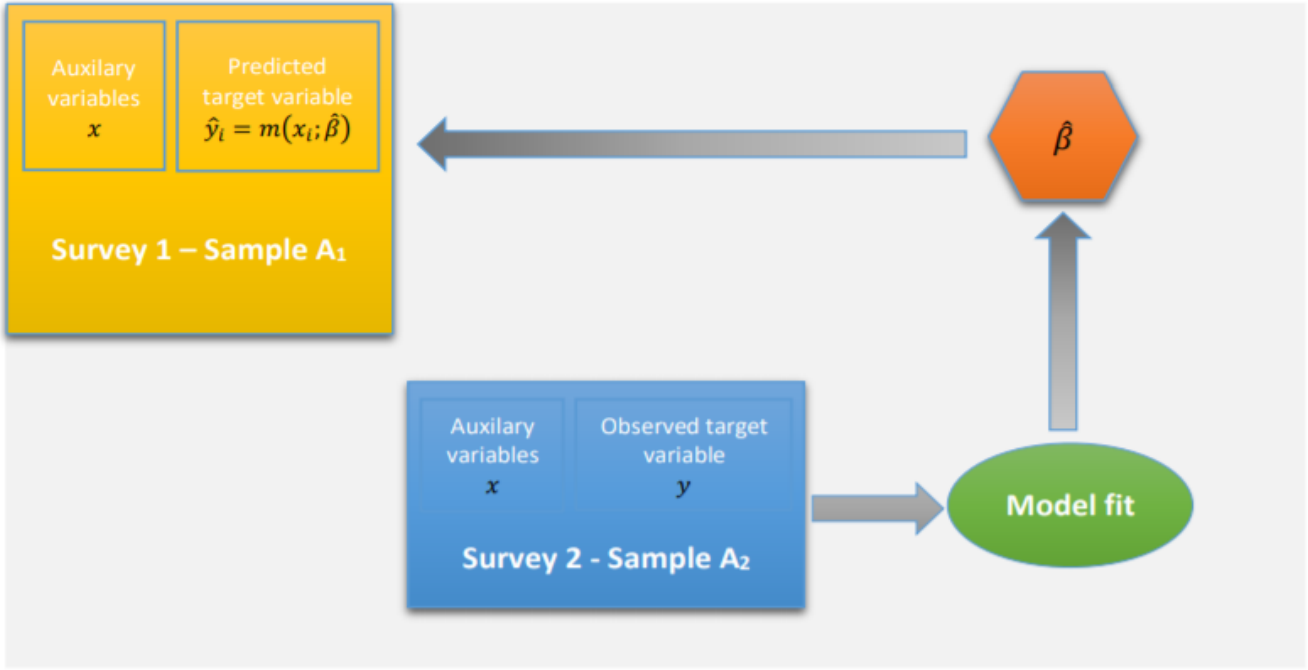
- **Sampling variance estimation:** to be used when indicators computation is based on the inferential properties of repeated sampling schemes.
- **Model Variance:** suitable when estimation relies on models using auxiliary variables.
- **Global Variance:** when model-based approaches are used jointly with inference based on the sampling design.

The publication emphasizes the importance of **estimating** and **disseminating** accuracy measures:

- Enables users to assess the fitness for use of an estimate.
- Build public trust in data and their use.

Addressing disaggregation at the analysis stage

Possible approach for indirect estimation: model-assisted approach based on the pioneering work of Kim and Rao (2012) considering the integrated use of: 1) a small survey collecting information on the variable of interest; 2) a more extensive survey or census not measuring the variable of interest but gathering relevant auxiliary variables.



Source: FAO, 2020.

A practical application based on SDG Indicator 2.1.2

The approach has been adopted to produce disaggregated estimates of **SDG Indicator 2.1.2** on the Prevalence of Moderate and Severe Food Insecurity based on the Food Insecurity Experience Scale (FIES).

Two data sources:

- Malawi's Fourth Integrated Household Survey (IHS4) 2016-17
- Malawi FIES survey module collected through the Gallup World Poll – 2016

Objective: Estimate Indicator 2.1.2 by sex, age class, and income quintile.

The guidelines present results along with their accuracy measures

An introduction to Small Area Estimation

The guidelines also provide an **introduction to small area estimation methods** (SAE) by:

- Presenting the process flow for SAE implementation;
- Providing an overview of main **unit-level** and **area-level** approaches;
- Indicating main references on the topic;
- Giving tools to assess the quality of small area estimates.

Way forward

Starting from this work, the FAO will:

- Develop **case studies** on additional indicators under its custodianship (e.g. 2.3.1, 2.3.2, 5.a.1);
- Use **small area estimation** to produce disaggregated estimates of SDG Indicators;
- Develop methodologies and guidelines to integrate survey data with additional data sources: **census data**, **administrative data**, **geo-spatial information**.

Food and Agriculture Organization of the United Nations

SUSTAINABLE DEVELOPMENT GOALS

Thank you!

Access the Guidelines here:
http://www.fao.org/3/cb3253en/CB3253EN.pdf