Welcome to the Toolkit on Using Small Area Estimation for SDGs!

In committing to the realization of the 2030 Agenda for Sustainable Development, Member States recognized that the dignity of the individuals is fundamental and that the Agenda’s Goals and targets should be met for all nations and people and for all segments of society. Ensuring that these commitments are translated into effective action requires a precise understanding of the target populations and progress made in addressing their particular priorities.

To properly measure this, statistics need to be presented for different population groups and geographical areas. The Sustainable Development Goal (SDG) indicator framework has included an overarching principle of data disaggregation: SDG indicators should be disaggregated, where relevant, by income, sex, age, race, ethnicity, migratory status, disability and geographic location, or other characteristics, in accordance with the Fundamental Principles of Official Statistics.

As sound statistical methods are vital to overcome this challenge, Small Area Estimation (SAE) constitutes an important topic in the way forward. It covers a variety of methods used to produce survey based estimates for geographical areas or domains of study in which the sample sizes are too small, or even absent, to provide valid estimates. In order to obtain reliable estimates, additional datasets are generally brought to bear upon the process through a modelling procedure.

To enable national statistical offices to estimate disaggregated indicators, guidelines are needed to support the process. The idea of writing guidelines on how to use statistical methods and, in particular SAE, to produce disaggregated statistical indicators is not new. Some focus on methodological aspects, others provide methodology in a specific program language or focus on a specific topic such as poverty mapping. Usability of SAE for official statistics has also been carried out over the past 10 years. So how do these guidelines differ from the existing work?

The SAE4SDG Toolkit in Wiki is a space to provide information on methods to produce disaggregated data through small area estimation. It aims to complement and use the existing methodological work and case studies to encourage and enable national statistical offices to employ SAE for the monitoring of the SDGs. The Toolkit will be an evolving project/document that will incorporate newly available methods, case studies and practical examples in future versions. The Toolkit also focuses on key steps to help countries in moving from SAE experiment to official data production. Finally, the Toolkit aims to be a space for partners to document and include references for their work on small area estimation.

The Toolkit is produced by United Nations Statistics Division, under the guidance of the Inter-Agency and Expert Group on SDG Indicators (IAEG-SDGs) and the Inter-Secretariat Working Group on Household Surveys. Production of Toolkit is also supported by experts from national statistical offices, regional and international organisations, and academia. Contribution of experts to the Toolkit is covered in the Acknowledgement.
What to expect

The SAE4SDG Toolkit targets practitioners and technical staff in National Statistical Offices and other institutions within the National Statistical System that are interested in using SAE for the monitoring of the SDGs. While the Toolkit provides information on SAE methods and the process around building an estimation procedure, it also offers discussions around elements that help countries make the transition from SAE experimentation to production. This is to respond to the challenges that the use in official statistics is still rather limited even though the SAE methods have been around for a long time.

An outline of the Toolkit and brief description of each section are provided below:

- **Why is SAE important for SDG data disaggregation**: outlines challenges and limitations of producing data for SDG indicators solely based on survey data and explains why SAE is important.
- **Producing SAE**: explains each SAE production step, including assess data needs, availability of input data, SAE models, specification, analysis and adaptation and evaluation and benchmarking. The SAE production framework is adapted from the one proposed by Tzavidis et al. (2018). For each step of the SAE production, three SDG indicators are used as examples with available Synthetic dataset and R code. Users are encouraged to use them for hands-on exercises.
- **Communicating SAE methods and results**: provides guidance and examples on how to communicate SAE methods and results with different users.
- **From SAE experiment to production**: outlines challenges and enabling environment for countries to move from SAE experiment to official production, based on lessons learnt and experiences of national statistical offices.
- **SAE practices**: incorporates national examples and case studies through two angles: (a) documenting the lessons learnt and challenges of countries in using SAE for official data production; and (b) illustrating SAE practices for indicators under different SDG goals.
- **Additional reference materials are available in Software packages, key SAE readings and major national, regional and international SAE projects.**
- **FAQ** is a place to provide responses to commonly asked questions.

What not to expect

There is a wide range of small area estimation methods and the research is growing. As most other guidelines, this Toolkit covers the most common models and provides idea of extensions. Further literature is presented for readers who may be interested in more methodological aspects.

The guidelines do not have any research goal, instead existing literature and software is cited. For special issues in applications, the collaboration with experts may be recommendable.

The same applies to the presented case studies. There is a wide range of case studies available. However, the focus should be on case studies that are actually used as official statistics or at least as experimental statistics. Thus, the collaboration with national statistical offices is needed to find out to which case studies this applies.

What is next?

The Wiki platform will continue to be updated to incorporate:

1. **SAE case studies**: more national case studies will be incorporated. All readers are welcome to submit case studies (sdgindicators@un.org) and examples to be incorporated into the Wiki site.
2. **Introduction and link to latest methodological development in the area of SAE, including the use of non-traditional data sources.**
- **Haoyi Chen**
  - Page Acknowledgement
    updated May 25, 2022 (view change)
  - Page Mexico
    updated May 25, 2022 (view change)
  - Page SAE for countries and partners
    updated May 23, 2022 (view change)
  - Page Training materials
    updated May 13, 2022 (view change)
  - Page Software packages
    updated Apr 28, 2022 (view change)
  - Page SAE by SDGs
    updated Apr 23, 2022 (view change)

- **Aida Khalil (FAO)**
  - Page FAO
    updated Apr 19, 2022 (view change)
  - Page References
    updated Apr 19, 2022 (view change)

- **Haoyi Chen**
  - Page Indonesia
    updated Apr 08, 2022 (view change)
  - Page Italy
    updated Apr 08, 2022 (view change)
  - Page Enabling environment to foster the use of SAE for official data production
    updated Apr 08, 2022 (view change)
  - Page SAE4SDG
    updated Apr 08, 2022 (view change)

- **Ann-Kristin Kreutzmann**
  - Page Evaluation and Benchmarking
    updated Mar 24, 2022 (view change)

- **Haoyi Chen**
  - Page US Census Bureau
    updated Mar 01, 2022 (view change)
  - Page Why is SAE important for SDG data disaggregation
    updated Mar 01, 2022 (view change)
  - PNG File mapPlotMotivation2_red.png
    attached Mar 01, 2022
  - PNG File mapPlotMotivation1_red.png
    attached Mar 01, 2022