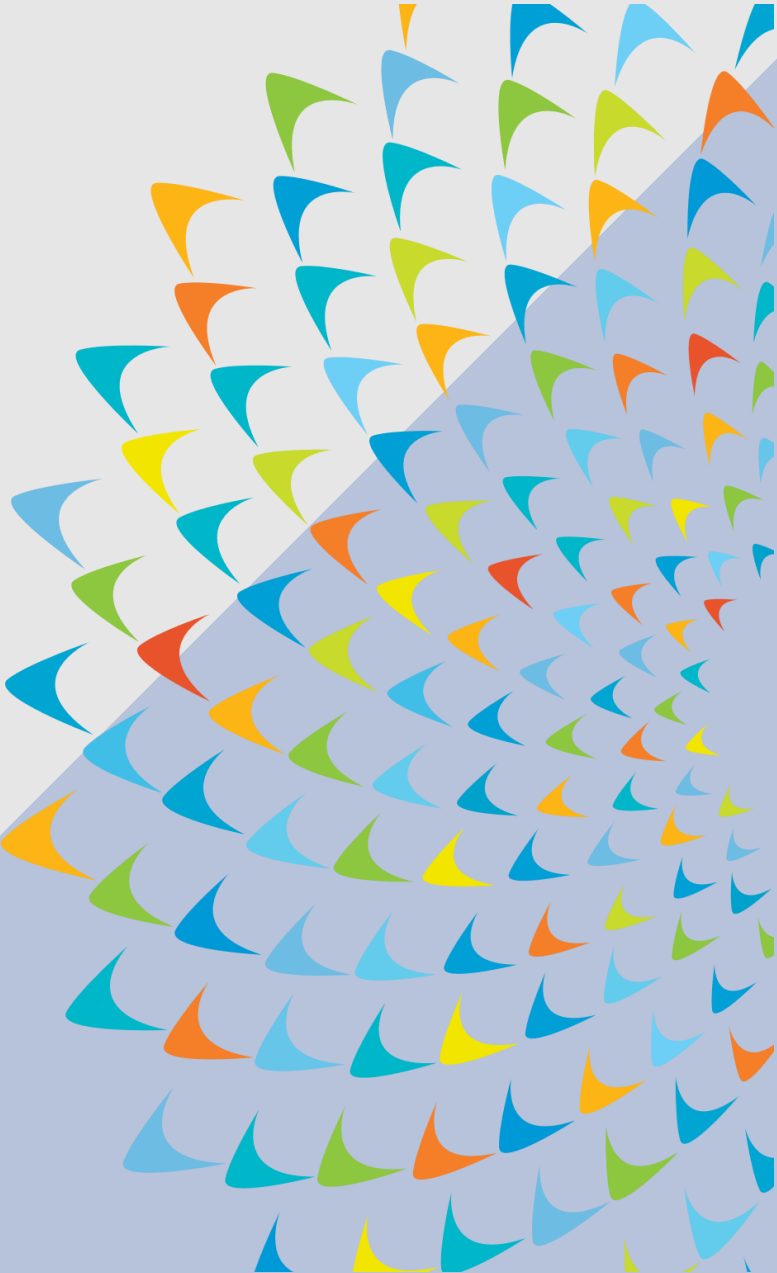# Data disaggregation:
## The *whys* and (some) *hows*

**PRACTICAL GUIDEBOOK ON DATA DISAGGREGATION FOR THE SUSTAINABLE DEVELOPMENT GOALS**

MARCH 2021

**CHAPTER 1** DATA DISAGGREGATION AND THE SUSTAINABLE DEVELOPMENT GOAL INDICATORS

**CHAPTER 2** INTEGRATING POLICY DEMANDS ON INCLUSIVE DEVELOPMENT WITH DATA

**CHAPTER 3** SOURCES OF DISAGGREGATED DATA: STRENGTHS, POTENTIALS AND LIMITATIONS

**CHAPTER 4** ANALYSIS OF DISAGGREGATED DATA IN AID OF LEAVE-NO-ONE- BEHIND

**CHAPTER 5** REPORTING, COMMUNICATING, AND IMPROVING USE OF DISAGGREGATED DATA

**CHAPTER 6** PRODUCING AND USING DISAGGREGATED DATA— FROM POTENTIAL TO POSSIBLE

ADB

- The whys-- LNOB and the SDG
- The whys-- the statistical concept of data disaggregation
- The hows-- the disaggregation dimensions and categories
- The hows– current standards

DATA DISAGGREGATION AND THE SUSTAINABLE DEVELOPMENT GOAL INDICATORS

## Developing a Strategy for Getting Disaggregated Data

ADB

- The whys-- LNOB and the SDG
- The whys-- the statistical concept of data disaggregation
- The hows-- the disaggregation dimensions and categories
- The hows– current standards

### Figure 1.1: Inclusive Data Charter Vision and Principles

**Principle One - All populations must be included in the data**
We can only achieve the "leave no one behind" goal by empowering the furthest behind. This means ensuring their voices are heard and their experiences are represented through data and analytics. We need to acknowledge all people, make them visible in the data to understand their lives, and inc-lude them in the development process.

**Principle Two - All data should, wherever possible, be disaggregated in order to accurately describe all populations**
We recognize that data should be disaggregated by sex, age, geographic location, and disability status and, where possible, by income, race, ethnicity, migratory status, and other characteristics relevant in national contexts.

**Principle Three - Data should be drawn from all available sources**
We recognize the need to make high-quality, timely data from official and non-official sources accessible, and that these should include new data sources, where consistent with internationally accepted statistical standards.

**Principle Four - Those responsible for the collection of data and production of statistics must be accountable**
We will balance the principles of transparency - maximizing the availability of disaggregated data - confidentiality, and privacy to ensure personal data is not abused, misused, or putting anyone at risk of identification or discrimination, in accordance with national laws and the Fundamental Principles of Official Statistics.

**Principle Five - Human and technical capacity to collect, analyze, and use disaggregated data must be improved, including through adequate and sustainable financing**
We recognize that collecting and analyzing disaggregated data needs specific skills and these must be built. We recognize the need to finance data collection, analysis, and use appropriately and sustainably so that high-quality data can be collected and used by governments as well as by businesses, civil society, and citizens.

Source: Global Partnership for Sustainable Development Data. Inclusive Data Charter vision and principles. https://www.data4sdgs. org/sites/default/files/2018-08/IDC_onepager_Final.pdf.
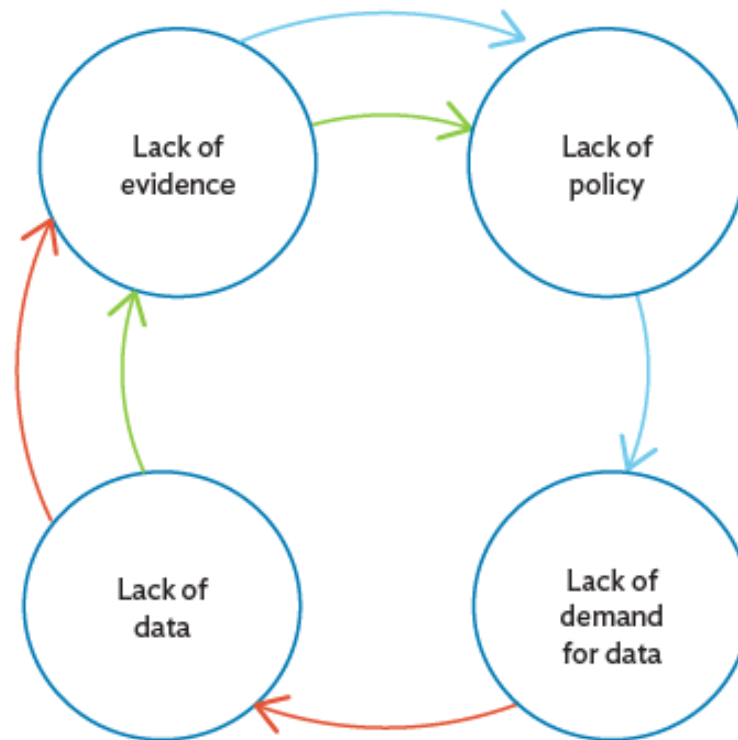
ADB

*Guiding principles*
- Policy-Data and Data-Policy Nexus
- HRBA to Data

*Policy priorities for 7 vulnerable populations (IAEG-SDGs)*

Figure 2.1: The Policy-Data and Data-Policy Nexus



Source: A. Bidarbakht-Nia. 2018. Connecting policymakers and data producers. *United Nations Economic and Social Commission for Asia and the Pacific blog.* 22 August. https://www.unescap.org/blog/connecting-policymakers-and-data-producers.

*Guiding principles*

- Policy-Data and Data-Policy Nexus
- HRBA to Data

*Policy priorities for 7 vulnerable populations (IAEG-SDGs)*

Figure 2.1: The Policy-Data and Data-Policy Nexus



Lack of evidence

Lack of policy
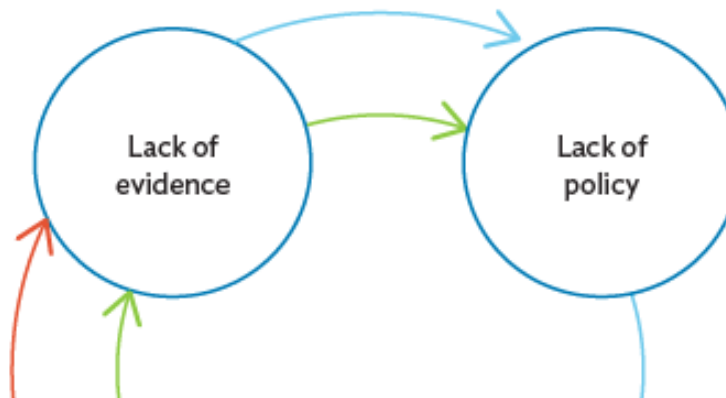
| Population Group | Policy Priorities |
| --- | --- |
| Poor | • Inclusive and pro-poor growth<br>• Social protection systems and floors<br>• Effective governance, including participation and use of available resources |
| Women and girls | • Poverty eradication<br>• Food insecurity and health<br>• Education<br>• Access to economic resources and decent work for all<br>• Gendered impacts of climate change |

ADB

*Guiding principles*
- Policy-Data and Data-Policy Nexus
- HRBA to Data

*Policy priorities for 7 vulnerable populations (IAEG-SDGs)*



**A HUMAN RIGHTS-BASED APPROACH TO DATA**

LEAVING NO ONE BEHIND IN THE 2030 AGENDA FOR SUSTAINABLE DEVELOPMENT

PARTICIPATION · DISAGGREGATION · SELF-IDENTIFICATION · TRANSPARENCY · PRIVACY · ACCOUNTABILITY

| Population Group | |
|---|---|
| Poor | • Inc...<br>• So...<br>• Eff... |
| Women and girls | • Po...<br>• Fo...<br>• Ed...<br>• Ac...<br>• Ge... |

ble resources

ADB

*Guiding principles*
- Policy-Data and Data-Policy Nexus
- HRBA to Data

*Policy priorities for 7 vulnerable populations (IAEG-SDGs)*

*Tools: EPIC, ADAPT, StaTact*

Development Principles
(Social, Economic, Environmental)

Policy-Data void

Policy

Data

Policy waste

Not monitored
(Data gap)

Monitored
(Data used)

Data unused
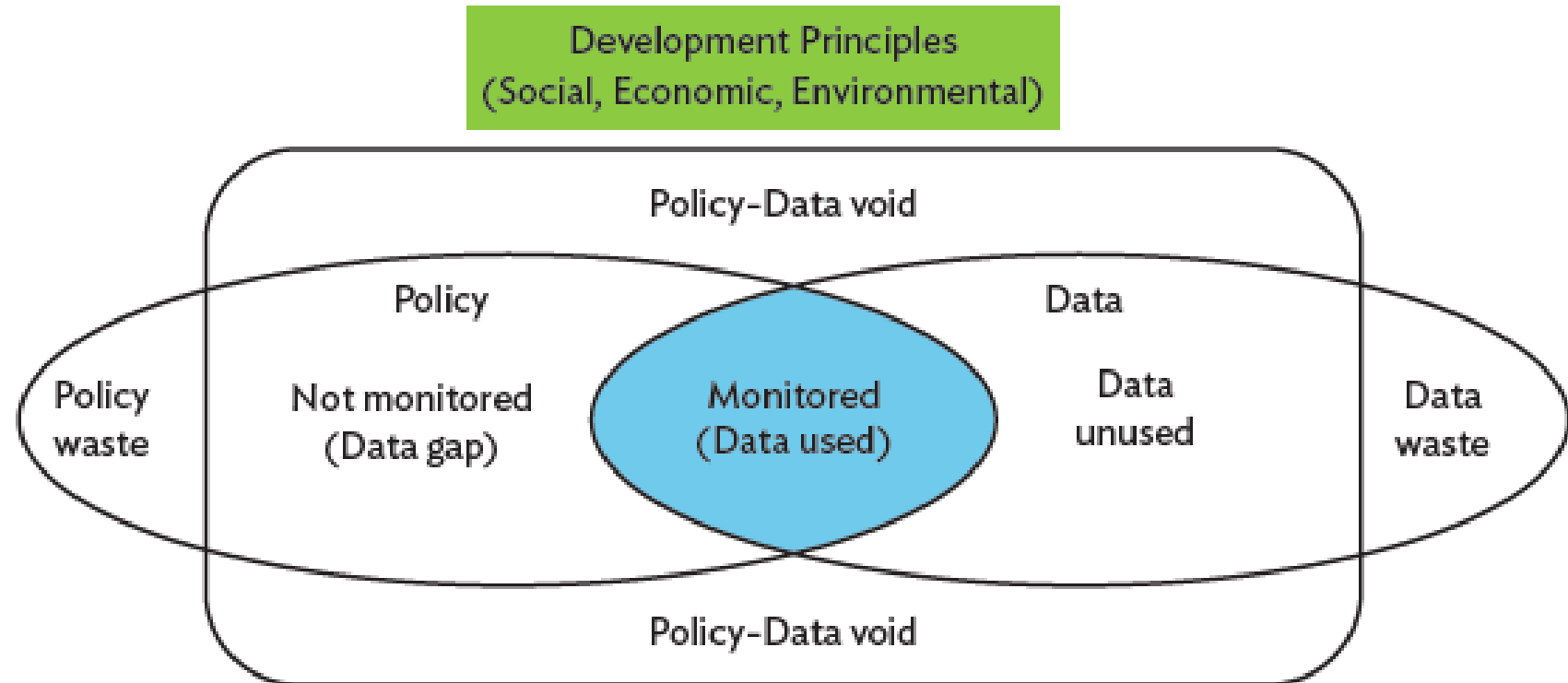
Data waste

Policy-Data void

8

Guiding principles
- Policy-Data and Data-Policy Nexus
- HRBA to Data

Policy priorities for 7 vulnerable populations (IAEG-SDGs)

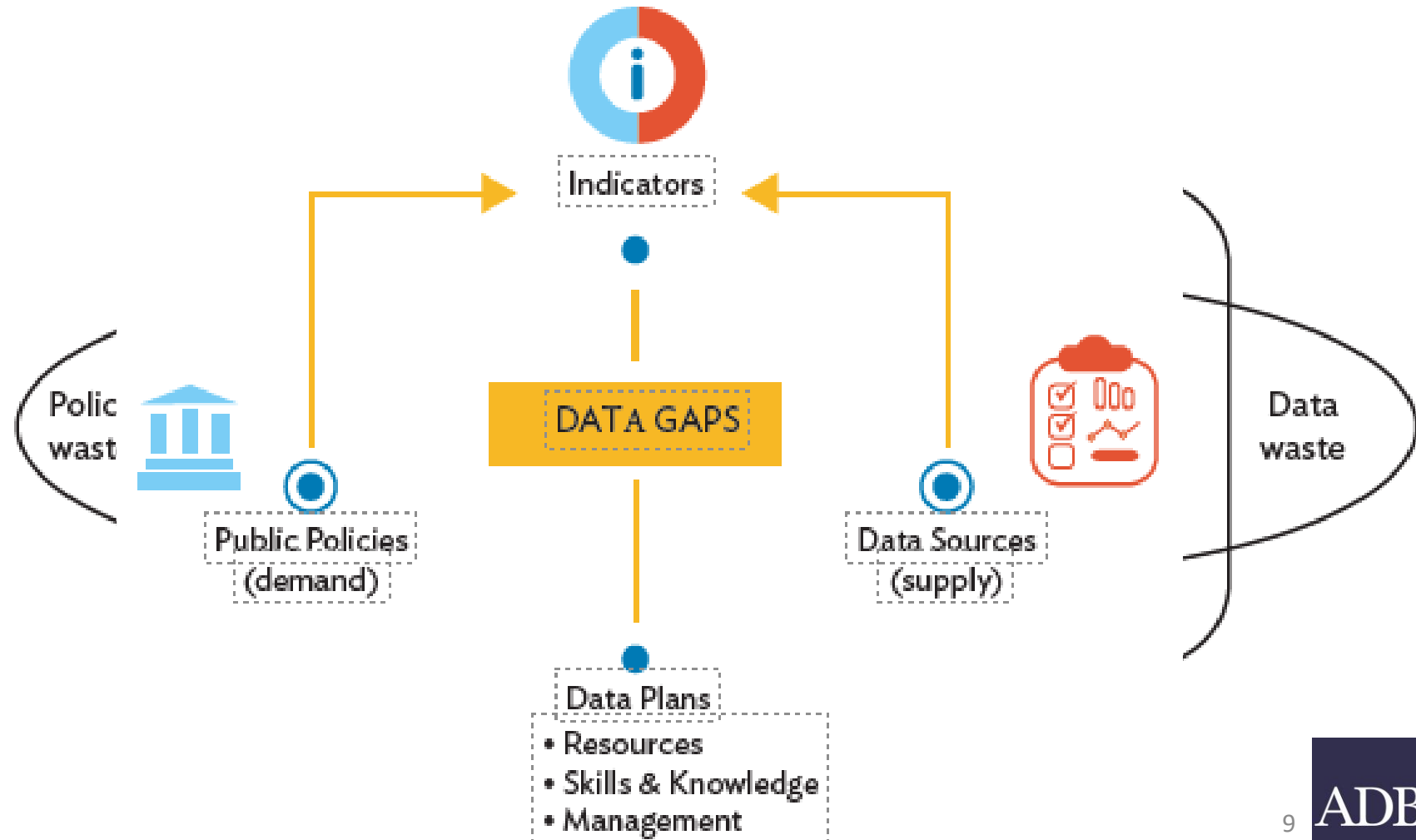Tools: EPIC, ADAPT, StaTact

CHAPTER 2

INTEGRATING POLICY DEMANDS ON INCLUSIVE DEVELOPMENT WITH DATA

Indicators

DATA GAPS

Polic wast

Public Policies (demand)

Data Sources (supply)

Data waste

Data Plans
- Resources
- Skills & Knowledge
- Management

ADB

## INTEGRATING POLICY DEMANDS ON INCLUSIVE DEVELOPMENT WITH DATA

*Guiding principles*

- Policy-Data and Data-Policy Nexus
- HRBA to Data

*Policy priorities for 7 vulnerable populations (IAEG-SDGs)*

*Tools: EPIC, ADAPT, StaTact*

- Census data
- Survey data
- Administrative reporting systems ()
- Small-area estimation
- Big data, geospatial data
- Data integration

**INTRODUCTION TO SMALL AREA ESTIMATION TECHNIQUES**

A Practical Guide for National Statistics Offices

MAY 2020

**MAPPING POVERTY THROUGH DATA INTEGRATION AND ARTIFICIAL INTELLIGENCE**

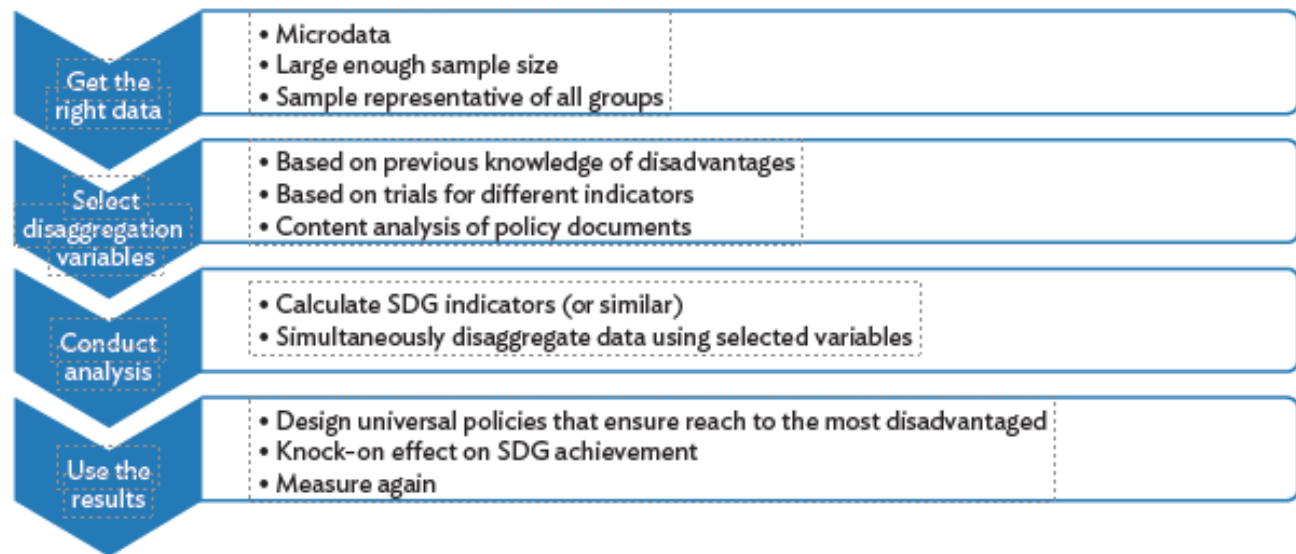*A Special Supplement of the Key Indicators for Asia and the Pacific 2020*

ADB

*The whys and hows*

- Multi-level disaggregation analysis of SDG indicators
- Equity assessments
- Telling stories through data visualization

Figure 4.6: Framework for Leave No One Behind Analysis

**Get the right data**
- Microdata
- Large enough sample size
- Sample representative of all groups

**Select disaggregation variables**
- Based on previous knowledge of disadvantages
- Based on trials for different indicators
- Content analysis of policy documents

**Conduct analysis**
- Calculate SDG indicators (or similar)
- Simultaneously disaggregate data using selected variables

**Use the results**
- Design universal policies that ensure reach to the most disadvantaged
- Knock-on effect on SDG achievement
- Measure again

SDG = Sustainable Development Goal.

Source: United Nations Entity for Gender Equality and the Empowerment of Women (UN Women). 2020. *Module 7: Multilevel Disaggregation Analysis to Monitor the SDGs from a Leave No One Behind Perspective—Training Syllabus.* https://data.unwomen.org/sites/default/files/documents/Asia-Pacific-Training-Curriculum/Module7/Module7_Syllabus_LNOB.pdf.
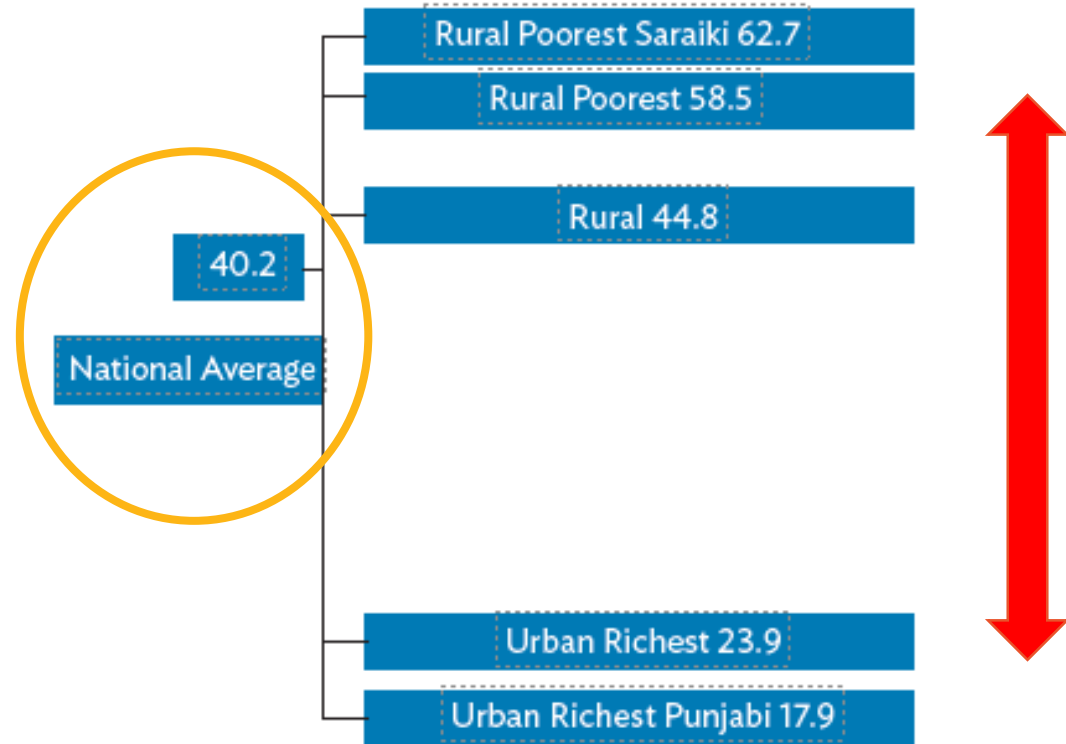
ADB

*The whys and hows*
- Multi-level disaggregation analysis of SDG indicators
- Equity assessments
- Telling stories through data visualization



ANALYSIS OF DISAGGREGATED DATA IN AID OF LEAVE-NO-ONE- BEHIND

CHAPTER 4

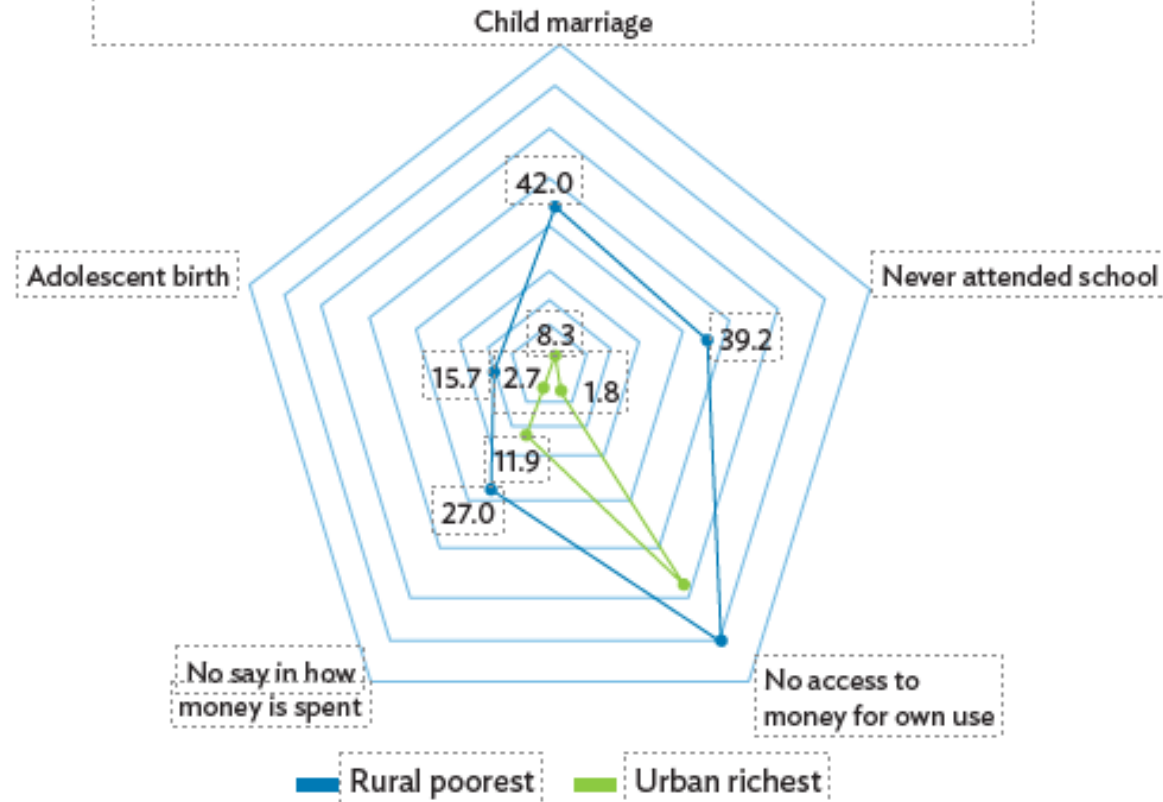Figure 4.4: Proportion of Women Aged 18–49 Who Were Married or in a Union before Age 18 in Pakistan, 2012-2013

Rural Poorest Saraiki 62.7
Rural Poorest 58.5
Rural 44.8
40.2
National Average
Urban Richest 23.9
Urban Richest Punjabi 17.9

ADB

*The whys and hows*

- Data flows in sharing and reporting SDG progress
- Dissemination tools
- Improving communication
- Use cases



**CHAPTER 5**

**REPORTING, COMMUNICATING, AND IMPROVING USE OF DISAGGREGATED DATA**

Figure 5.1: Spider Chart Visualization of Multidimensional Inequalities

Inequalities between poorest rural and richest urban Indian women, various indicators, percentage, 2015–2016

- Child marriage
- Never attended school
- No access to money for own use
- No say in how money is spent
- Adolescent birth

42.0
39.2
8.3
15.7  2.7  1.8
11.9
27.0

Rural poorest   Urban richest

ADB

*The whys and hows*
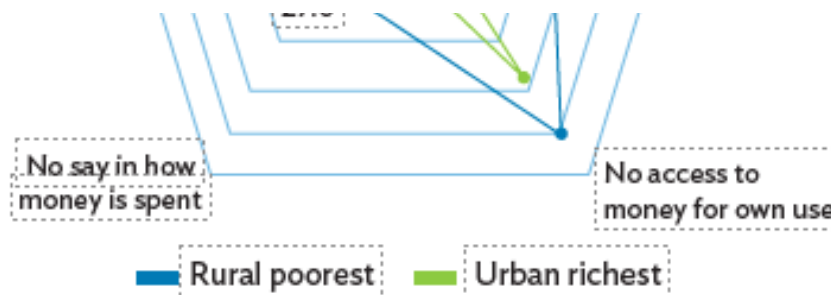- Data flows in sharing and reporting SDG progress
- Dissemination tools
- Improving communication
- Use cases



CHAPTER 5

REPORTING, COMMUNICATING, AND IMPROVING USE OF DISAGGREGATED DATA

Figure 5.1: Spider Chart Visualization of Multidimensional Inequalities

Inequalities between poorest rural and richest urban Indian women, various indicators, percentage, 2015-2016

Child marriage

No say in how money is spent

No access to money for own use

— Rural poorest   — Urban richest

**Understanding Usability of SDG National Reporting Platforms**

OPEN DATA WATCH

ADB

- Why required disaggregated data are often not available
- Ongoing capacity development initiatives that may help NSSs to address these issues

PRODUCING AND USING DISAGGREGATED DATA— FROM POTENTIAL TO POSSIBLE

Figure 6.2: Definition of Data Governance

"Data governance provides clear rules about how data should be treated throughout the data value chain to stimulate innovation, efficiency, sustainability, growth in use, and the value of data, while building trust and safeguarding against misuse of data."

GOOD DATA GOVERNANCE

INSTITUTIONS    LEGAL SYSTEMS    ICT

ADB

- Why required disaggregated data are often not available
- *How?* Ongoing capacity development initiatives that may help NSSs to address these issues



CHAPTER 6

PRODUCING AND USING DISAGGREGATED DATA— FROM POTENTIAL TO POSSIBLE

Figure 6.3: Big Data-Related Competencies

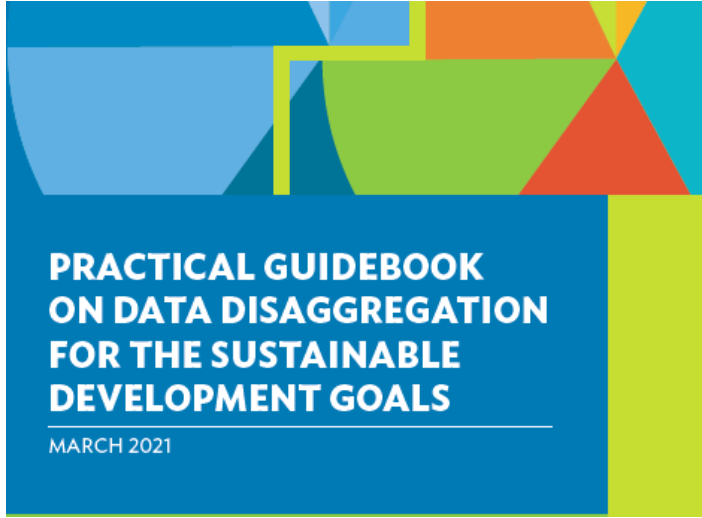| | Data acqusition | Data processing | Data analysis | Data visualization |
|---|---|---|---|---|
| **Core competencies** | Ethics and privacy | Ethics and privacy | Ethics and privacy | Ethics and privacy |
| | Data management | Data management | Mathematics | Statistics |
| | Machine Learning | Mathematics | Statistics | Programming |
| | Programming | Programming | Programming | Data visualization |
| | | Machine learning | Machine learning | |
| **Generic skills** | Product understanding | Curiosity | Curiosity | Product understanding |
| | Critical thinking | Business acumen | Adaptability | Business acumen |
| | Business acumen | Critical thinking | Critical thinking | Storytelling |
| | Curiosity | Communication | Communication | Tommunication |
| | Team player | Team player | Team player | Team player |
| | Agile project management | Agile project management | Agile project management | Agile project management |

ADB

- Why required disaggregated data are often not available
- *How?* Ongoing capacity development initiatives that may help NSSs to address these issues



**CHAPTER 6**

**PRODUCING AND USING DISAGGREGATED DATA— FROM POTENTIAL TO POSSIBLE**

Figure 6.3: Big Data-Related Competencies

Figure 6.6: What Is Citizen-Generated Data?

**PRACTICAL GUIDEBOOK ON DATA DISAGGREGATION FOR THE SUSTAINABLE DEVELOPMENT GOALS**

MARCH 2021
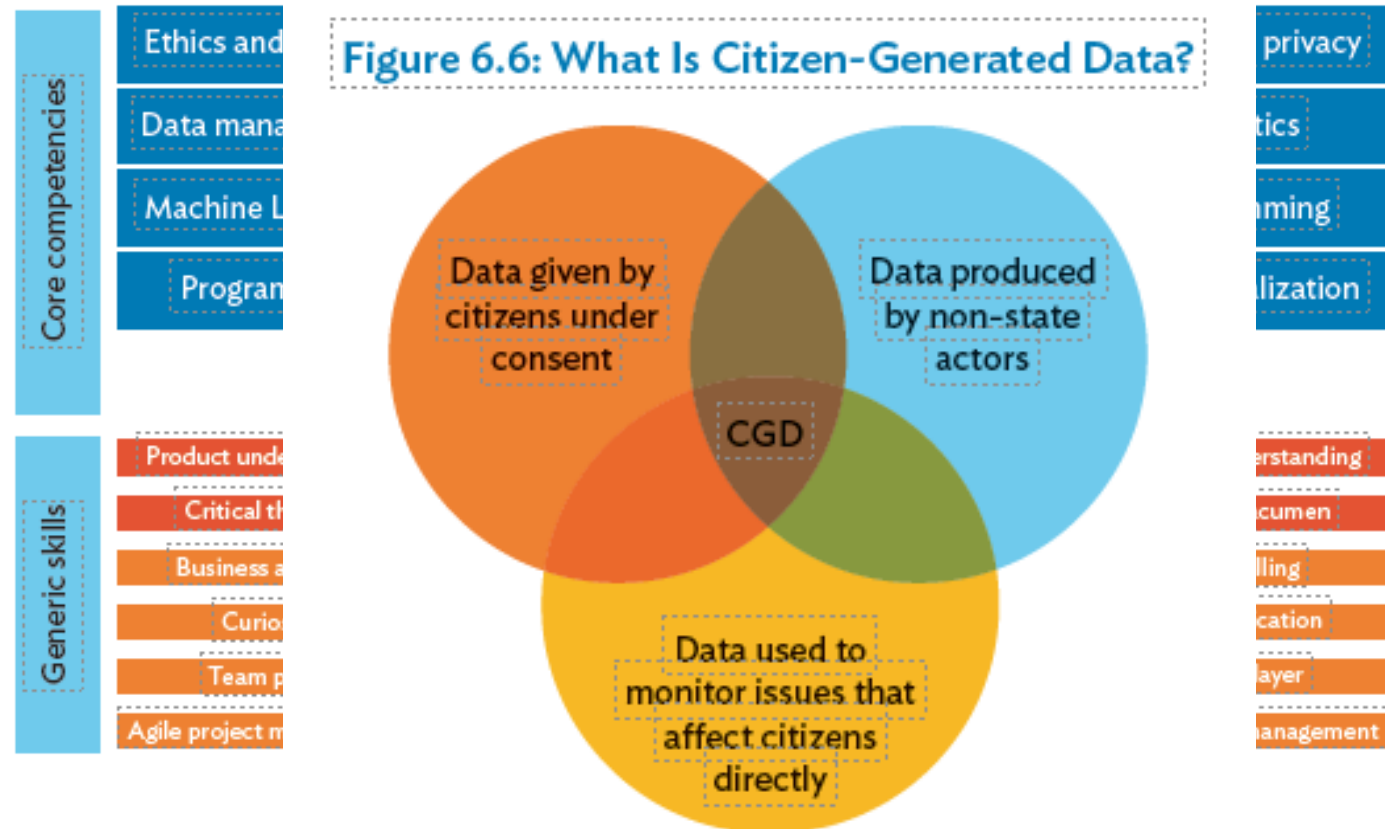
Elaine Tan (*Overall Supervision*)

Margarita F Guerrero (*Lead Author*)

Arturo Martinez Jr

Jose Ramon Albert

Yongyi Min (UNSD)

Heather Page (UNSD)

Mildred Addawe

Marymell Martillan

Joseph Bulan

Ron Lester Durante

IAEG-SDGs

UN Women

UN ESCAP

WHO

PARIS21