

PART III:

SDG DATA SOURCES AND GAPS

Accurate, timely and comparable data for all SDG indicators is essential for a “robust, voluntary, effective, participatory, transparent and integrated”⁷ review framework for the 2030 Agenda. However, three years into the SDG era, such data are still far from comprehensive. This section addresses the following questions:

- How many global SDG indicators have sufficient data to assess progress of Goals and SDG Targets in the Asia-Pacific region?
- How does data coverage vary across subregions and the 17 Goals?
- To what extent is disaggregated data available to address the principle of “leaving no-one behind”?
- What are the primary sources of data for SDG indicators at national level?
- What are priorities for statistical investment to increase availability of SDG data?

7 United Nations (2015) “Transforming our world: the 2030 Agenda for Sustainable Development” (A/RES/70/1)

1. SDG DATA AVAILABILITY IN ASIA-PACIFIC REGION

To produce the present SDG progress report, data for Asia and Pacific countries were drawn from the Global SDG indicator database as of August 2018, and from online databases of designated custodian agencies in October 2018. The 244 SDG indicators (232 unique indicators) were classified based on the following criteria for the region:

- **Sufficient data:** Indicators with at least an underlying data series with two data points or more between 2000 and 2018 for at least half the countries in the region (or half of the countries in a subregion). To estimate a historical trend, two data points for at least half of the countries is deemed sufficient.
- **Insufficient data:** Indicators with an underlying data series with at least one data point (or two data points, but for less than half of the countries in a subregion). While indicators with such limited data availability may shed light on the current status of the region, no historical trend can be estimated.
- **No data:** Indicators with no data for any of the 58 countries of the region.

A large number of SDG indicators include disaggregation (women/men, urban/rural, age groups) or sub-indicators (e.g., SDG indicator 2.2.2 on malnutrition measures wasting and overweight children). As a result, SDG indicators can include two or more data series. In those cases, the series with the largest number of countries with data points is retained to measure data availability at the indicator level.

Sufficient data for only 36% of the global SDG indicators

Only 83 of the 232 Global SDG indicators have enough data to assess regional progress towards achieving the 2030 Agenda in Asia-Pacific. This number may appear low, nevertheless, this is a significant increase compared to last year. In 2017 there were sufficient data on only 63 SDG indicators⁸. Despite a clear improvement in data availability, it is still challenging to estimate SDG progress with insufficient or no data on almost two-thirds of the global SDG indicators.

Insufficient data for 30 per cent of Tier I indicators

To facilitate implementing the global indicator framework, the Inter-agency and Expert Group on SDG Indicators (IAEG-SDG) classified the 232 global SDG indicators into three tiers⁹:

- Tier I: Indicator is conceptually clear, has an internationally established methodology and standards, and countries regularly produce data (101 indicators);
- Tier II: Indicator is conceptually clear, has an internationally established methodology and standards, but countries do not regularly produce data (84 indicators);
- Tier III: Indicator has no internationally established methodology or standards, but methodology/standards are, or will be, developed or tested (41 indicators).

Most indicators with sufficient data in Asia-Pacific region (71 out of 79) are Tier I (or have at least one sub-indicator component classified as such for multiple-tiered indicators). The remaining 8 indicators are classified as Tiers II or III¹⁰.

However, the Asia-Pacific region has sufficient data for only 71 out of the 101 indicators classified as Tier I by the IAEG-SDG. Given availability of international standards and long tradition of data collection on these indicators, they are potential short-term priorities for statistical capacity building in the Asia-Pacific region.

Environmental indicators the most data-poor

The SDGs, in principle, integrate three dimensions of sustainable development – economic, social and environmental. However, most Goals can be predominantly associated to one of the three dimensions. As a result, agencies have used different ways to classify the 17 Goals and 169 SDG Targets¹¹. Though no perfect classification is possible, to analyse indicator gaps in Asia-Pacific, each Goal is classified under only one dimension of development based on the most prominent concentration of a Goal's objectives. The classification is as follows:

Economy	SDGs 8, 9	(29 indicators)
Social	SDGs 1-5; 10-11; 16	(128 indicators)
Environment	SDGs 6,7; 12-15	(62 indicators)
Not classified	SDG 17	(25 indicators)

8 The data availability analysis published in the 2017 Statistical Yearbook for Asia and the Pacific used similar criteria to define data availability, but divided indicators with insufficient data into two groups: "Status OK" for indicators having only one data point for more than half of the countries and "Status limited" for indicators having one data point for less than half of the countries. The two groups are merged into "Insufficient data" in this present data availability review.

9 Figures as of 31 December 2018

10 IAEG-SDG classifies SDG indicators 8.4.1 and 12.2.1 as Tier III due to unresolved methodological issues. Yet, the dataset published by UNEP is used in this report.

11 See (i) United Nations Development Programme (2017) "SDG Accelerator and Bottleneck Assessment"; (ii) Organisation for Economic Co-operation and Development (2016) "An SDG-based results framework for development co-operation"; and (iii) German Council for Sustainable Development (2015) "Sustainable Development Goals and Integration: Achieving a better balance between the economic, social and environmental dimensions".

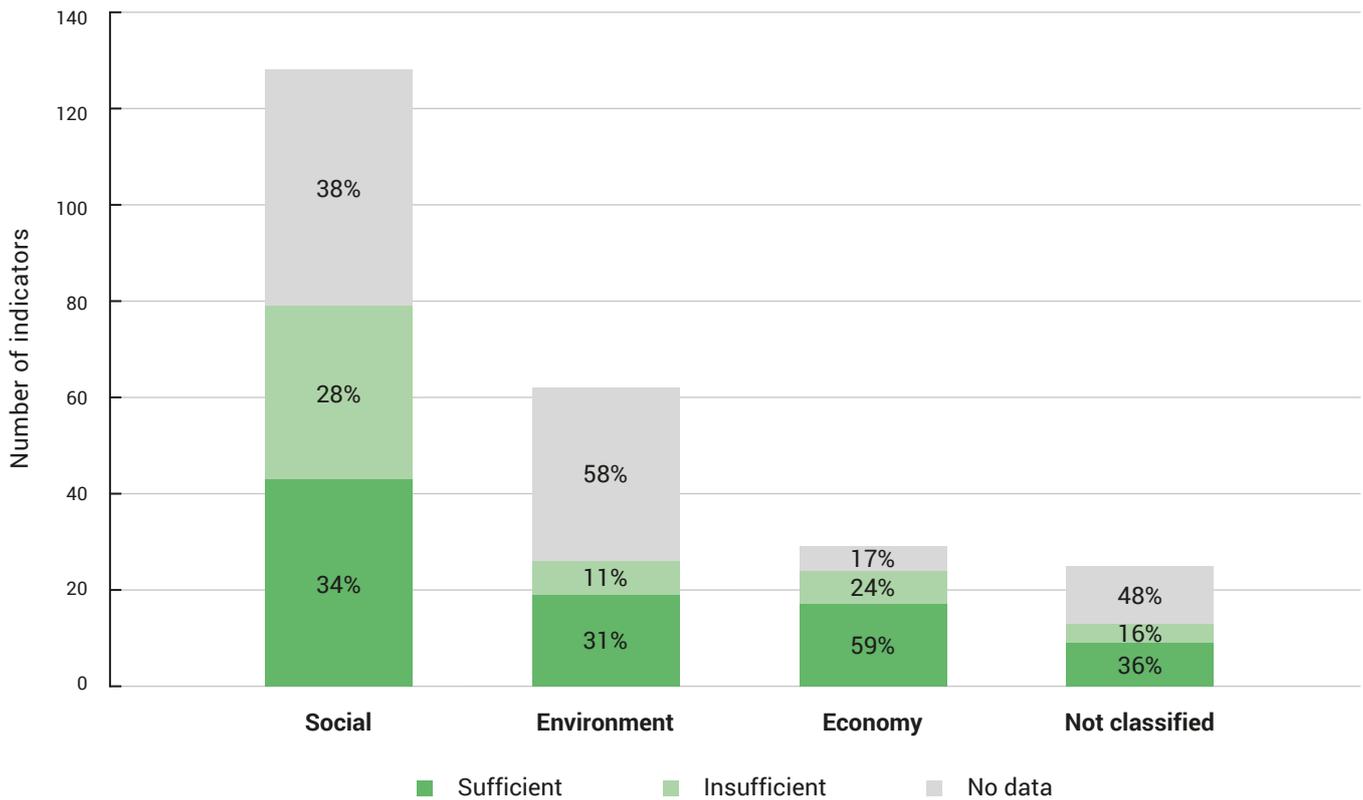


Figure 25– SDG data availability by development dimension

In general, whilst the social dimension has the most indicators, data availability is better for Goals with a strong economic dimension (Goals 8 and 9). Nearly 60 per cent of the SDG indicators under the economic dimension are estimated to have sufficient data for Asia and the Pacific region.

By contrast, only 34 per cent and 31 per cent of SDG indicators under the social and environmental dimensions, respectively, have sufficient data.

Better economic data availability might be explained by countries in the region prioritizing production of data for economy-related indicators and the existence of long-established, experienced national data compilation systems for such indicators (including national account systems, labour force surveys and establishment surveys).

Conversely, poor data availability under the environmental dimension, with nearly 60 per cent of environmental SDG indicators having no data for any country in the region, might be explained by the relative novelty of environmental measurement. While some data can be gathered through long-standing survey methods (e.g. agriculture), others like geo-spatial data come from newer technologies,

including remote sensors monitoring air and water quality. National and international statistics' compilation systems may not be as well equipped to collect and harmonize country data for such indicators.

Furthermore, methodological development for environment indicators may lag behind economic indicators. Nearly half the 41 SDG indicators classified as Tier III belong to the environmental dimension (against only two indicators in Tier III for the economic domain).

Data availability for the social dimension is mixed. For example, good health and well-being (Goal 3) has the highest data availability among the 17 Goals, with a large number of indicators carried over from the Millennium Development Goals (MDG) framework that rely on well-established data production and dissemination mechanisms. But this dimension also includes new measurement areas such as quality of education (Goal 4), sustainable cities and communities (Goal 11) and peace, justice and strong institutions (Goal 16) with the percentage of data availability ranging from less than 20 per cent for Goals 11 and 16 to about 35 per cent for Goal 4.

Sufficient data on less than a quarter of SDG indicators in Pacific subregion

SDG data availability is uneven across the five ESCAP subregions. The Pacific has the lowest share of SDG indicators with sufficient data (22 per cent), while the highest share occurs in South and South-West Asia with 40 per cent of SDG indicators with sufficient data. Apart from the fact that not all SDG indicators are equally relevant to all subregions, most variability can be explained by diversity in statistical capacity, level of demand for official statistics and investment in statistical development.

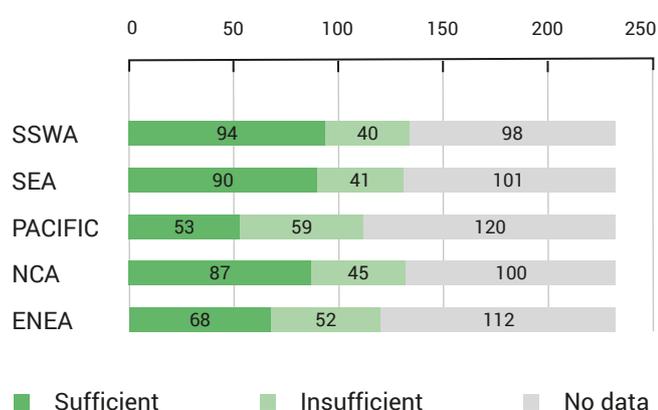


Figure 26 – SDG data availability by subregion

While surveys remain a major source for SDG indicators, administrative data provide the best coverage

Global SDG data originate from countries' national statistical systems that compile data from different sources:

- **Surveys** are the most common data compilation procedures used in many domains. They include sample surveys, such as household income and expenditure surveys, labour force surveys, agricultural and enterprise surveys, and censuses as total enumeration of the target population.
- **Administrative data:** The primary source of information to generate statistical information is an administrative register, such as: tax data (business/profits tax, property taxes, import/export duties),

social security data, health/education records, registration systems for persons/businesses/property/vehicles, as well as private businesses with data holdings (credit agencies, business analysts, utility companies, telephone directories, retailers with store cards).¹²

- **Digital data sources:** Digital data sources include satellite imagery whose primary purpose might be research, but from which statistical information can be derived, or indicators from remote sensors that monitor air or water quality.
- **Others:** Indicators for which the underlying compilation instrument is not clearly determined, for example, country compliance with international agreements that are measured by a yes/no metric versus a quantitative one.

Applying the above classification, administrative registers are the largest data source for the global SDG indicators – 91 of the 232 global SDG indicators can be sourced from administrative registers. Surveys are the next most frequent data source (seventy-one indicators can be obtained primarily from surveys) and then digital sources (11 indicators can be classified as sourced from digital sources). The remaining 59 SDG indicators could not be classified for varied reasons: some SDG indicators do not need to be compiled at country level (many indicators are for “number of countries”); international agencies estimate certain indicators; or indicators lack official metadata.

Nearly half the SDG indicators from administrative sources have sufficient data in the Asia-Pacific region, while only 32 per cent of the SDG indicators coming from surveys have sufficient data.

Global SDG indicators sourced from surveys show the highest rate of indicators with insufficient data (39 per cent). Surveys are typically implemented only every few years due to the high financial cost, resulting in time series with fewer data points.

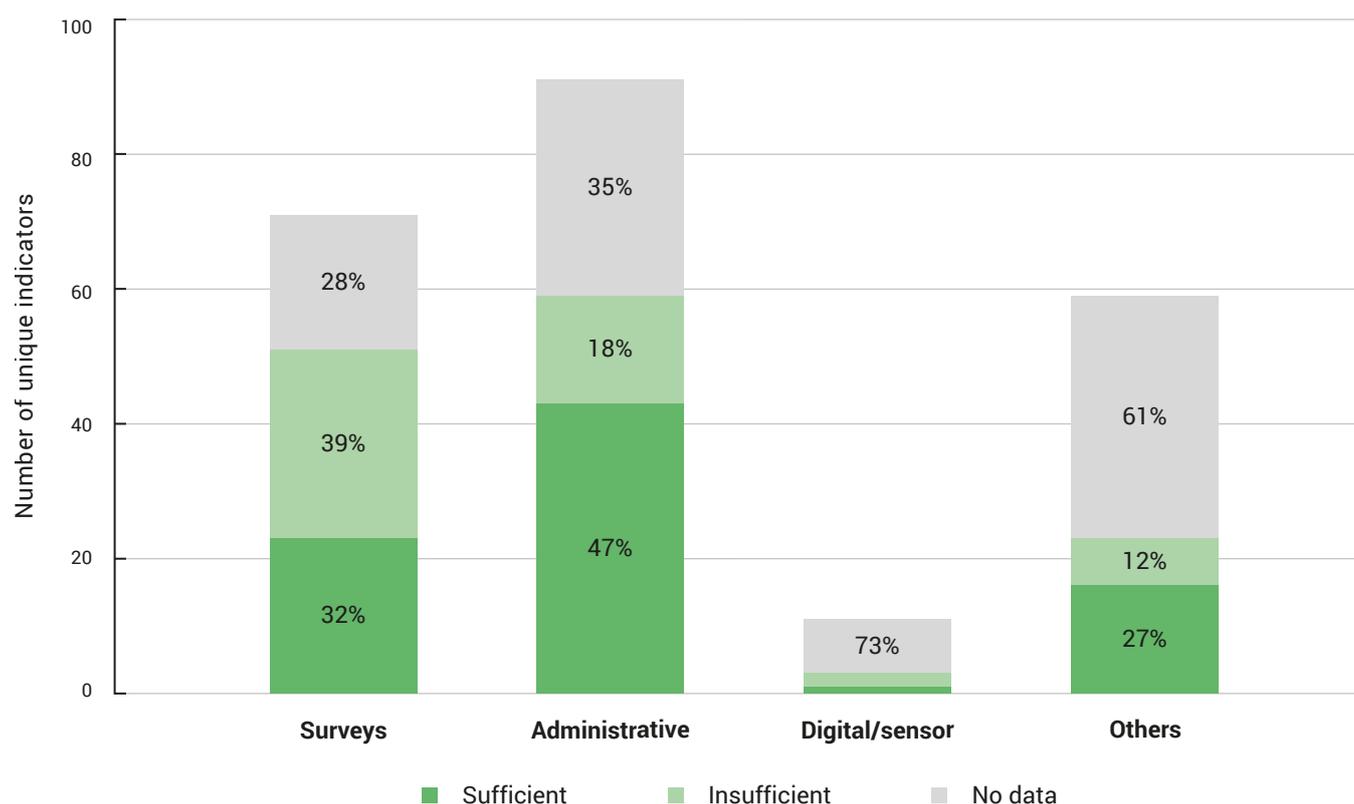


Figure 27 – Availability by type of national data source

Data for indicators sourced from administrative sources tend to be available more rapidly, increasing their usefulness in decision making. Examining the global SDG indicator dataset, the average time lag, measured as the average number of years between the year for the latest available data year and the current year (2018) across all countries in the region, is nearly four years for indicators sourced from surveys and only 2.6 years for indicators sourced from administrative data. While the time required by the international statistical system to compile and disseminate data is assumed to be the same in both cases, the availability of data at the national level appears to be significantly faster for administrative sources than surveys.

Indicators sourced from digital/sensor sources are too few (only 12 indicators identified) to draw conclusions. Only one indicator out of 12 has sufficient data, and nine indicators did not have any data. As the statistical methodology for this type of compilation is still relatively new and under development, eight of 12 indicators in this category are classified as Tiers II or III.

Study limitations

To ensure proper use of data availability estimates published in this part, notable limitations are:

- **Time lag.** There is often a substantial time lag between data being compiled, processed and published at national and international levels.
- This assessment uses only the best series under each indicator to estimate data availability at the indicator level. Consequently, availability does not cover all disaggregated series.
- **Non-applicability.** In practice, certain indicators may not apply to certain countries. For example, marine preservation in landlocked countries. This aspect has not been considered in this analysis as ESCAP is not in a position to identify relevance of indicators to countries.
- **Statistical quality.** Dimensions of quality (such as relevance, accuracy, reliability, timeliness, punctuality, clarity, coherence, comparability, and methodological soundness) are beyond the scope of this review.

2. SUMMARY AND CONCLUSIONS

Lack of data hinders SDG progress assessment.

Despite a significant increase in availability of SDG indicators in the Asia-Pacific region, lack of data remains one of the biggest challenges for the region to measure progress towards the 2030 Agenda for Sustainable Development; less than 40 per cent of the global SDG indicators are available for regional SDG progress assessment.

A two-pronged approach needed for statistical support.

Tier I SDG indicators are one area of short-term action that will deliver gains in the near future. On the other hand, environmental indicators (mostly Tiers II and III) and some of the social indicators such as quality education, gender equality and strong institutions, require long-term planning and investment for methodological development, capacity building and data production. Data in the economic domain is generally more plentiful than in the social and environmental domains.

There is a wide gap in data availability across subregions.

The Pacific, where 75 per cent of indicators lack sufficient data, faces the biggest challenge to assess SDG progress.

The importance of administrative sources for SDG data at the national level is evident.

Survey data remain a major source of SDG indicators at the national level, yet 68 per cent of indicators sourced from surveys lack data at the regional level. Administrative sources are a major primary source for SDG indicators at national level. Administrative data sources present several advantages over surveys: produced at a lower cost, more rapidly and at a higher frequency. Administrative data alleviate the increasing difficulty faced by statistical offices in obtaining data from survey respondents (response burden) and decrease the long-term cost of producing official statistics.

The region needs to enhance use of alternative data sources for SDG indicators.

The use of alternative sources of data (such as remote sensing) for SDGs is minimal in the region. Statistical systems need better access to methods and tools that enable them to harness the power of new and alternative sources to complement surveys and administrative data.

