METHODS, GAPS, LIMITATIONS AND FURTHER RESEARCH

Methods

The SDGs are made up of 17 Goals that cover a wide range of topics, including ending poverty and hunger, protecting life on land and in water, creating decent jobs, building sustainable infrastructure, ending inequality, and promoting just institutions. They are measured by 169 targets and 232 unique indicators that are often repeated across Goals. This is because the Goals are meant to be achieved together and through each other.

To measure SDG achievement in the US, this report uses 103 indicators across 15 of the 17 Goals. Indicator values are transformed (normalized) into a 0-100 scale, where 100 represents achieving that particular indicator or Goal, and 0 represents no progress towards that Goal. Indicator scores are then averaged across each Goal to get a Goal score. Goal scores are averaged to get overall rankings. A full list of indicators can be found in the Annex.

To determine SDG progress, data was collected for each indicator going back as far as 2000, when available. For 92 of the 102 indicators, there is data for more than one year. SDG scores were calculated for each year for which there is data, starting in 2015 when the Goals began (or the closest year for which there is data) and continuing to 2020 (or the last year for which there is data). To measure growth rates, the linear average growth was calculated for the period closest to 2015-2020. These rates are then compared with the linear growth rate needed to reach a score of '100' by 2030, from the baseline values in 2015. For more detailed information, see the Full Methodology in the Annex.

How does this report compare to the 2018 edition?

This report includes additional information about how quickly and in what direction states are moving to achieve the SDGs. It sheds light on where states may be performing well right now, but getting worse, and where poor performance may mask improvement. In addition, this report includes 11 new indicators, many of which focus on measuring the 'Leave no one behind' agenda. 17 indicators were removed because recent data was not available, more precise measures were found, or alternative measures that included longitudinal data were substituted. The source, units, or definition of 13 indicators changed from those of previous report. More detailed information can be found in the Annex.

FULL METHODOLOGY

The Sustainable Development Report of the United States measures progress towards the internationally agreed Sustainable Development Goals. Using publicly available, recent data from reputable sources, this index presents an aggregate snapshot of development progress in US states.

The methodology below builds on the methodology built by SDSN and Bertelsmann Stiftung for the SDG Index and Dashboards Report. It has adapted those efforts and those from the version presented in the 2018 Sustainable Development Report of the United States. This section includes: 1) information on indicator and data selection, 2) rescaling and normalizing the data, 3) aggregating composite index and adding colors and, 4) tracking trends over time.

INDICATOR SELECTION CRITERIA

To determine quality, technically-sound, indicators for selection we used the following criteria:

1. **SDG and US state relevance:** Data is matched to the SDG targets, then matched to suggested indicators as closely as possible. From this list, indicators are selected that are most relevant to state contexts, for example: the index excludes international cooperation indicators. Finally, when possible, indicators should be relevant to a policy context and/or support communities and leaders in policy-making decisions. Alignment of each indicator to the SDG target or indicator is noted on the sources pages.

2. **Statistical quality:** Data must be from a reputable source that produces data in a replicable and reliable way. Preference is given to datasets that are updated routinely, so progress can be tracked to 2030, and to datasets that have disaggregated data available, to track progress for all groups.

3. **Timeliness:** Data must be published recently, with preference given to data covering years 2017 or later.

In 7 instances, data from earlier years was used because it was the most reliable source to cover an essential issue (see the source annex for more information on specific data sources and years covered).

4. **Coverage:** Datasets must provide data for at least 80% of states.

While all variables have more than 80% coverage, there are four variables that have missing values in their latest year available: Racial disparity in child poverty, Eviction rates, Non-carbon footprint, and Dam safety.

5. **Comparability:** Data was chosen that has a reasonable or scientifically determined threshold.

There are several indicators that the UN has recommended for monitoring purposes that aren't well

suited for comparison in an index because there is no consensus on 'best' level of achievement, and indeed 'best' levels may vary by location. This is the case, for example, with passenger and freight volumes (Indicator 9.1.2) or percent of employment in the manufacturing sector (Indicator 9.2.2) from Goal 9, neither of which have an optimal level of achievement at the state level.

6. **Repeated indicators:** Data should not repeat across Goals.

Within the SDGs official indicators, there are indicators that are repeated across multiple Goals. This promotes the idea that the SDGs are interconnected and interdisciplinary. However, in order to prevent double counting of indicators within the index calculations, indicators were not repeated across Goals. In cases where an indicator could reasonably fit within multiple SDGs, it was placed within the Goal with the target that was determined to most closely/directly match the language/intent of the indicator.

7. **Outcome indicators:** Whenever possible, data should measure outcomes.

In cases where outcome data was unavailable, process or output indicators were used to track policies or actions that have research-supported impact on outcomes.

Goals 14 and 17 are not included in this index due to issues of data availability, jurisdiction, and lack of state-level comparability.

Rescaling and normalizing the data

To rescale and normalize the data, the index followed the methodology developed by SDSN and Bertelsmann Stiftung, which is detailed below. Indicators were rescaled so they could be compared with one another. The choice of upper and lower bounds with which to rescale the data is a sensitive one and can introduce unintended effects into datasets if extreme values and outliers are not taken into account. (Note: in this section the term "upper bound" is used to refer to the target value, even if the indicator data is descending and the most progress is represented by a smaller number.) Lower bounds are particularly sensitive to outliers as they can impact the rankings of the data.[21] Detailed information about each indicator, it's bounds, and the rationale for those bounds can be found in Annex 3. To account for these considerations, this index used the following methodology for determining upper and lower bounds:

The upper bound for each indicator was determined using a five-step decision tree developed by SDSN and Bertelsmann Stiftung:[22]

1. **Use the absolute quantitative thresholds outlined in the SDGs and targets:** e.g. zero poverty, universal school completion, universal access to water and sanitation, full gender equality. Some SDG targets also propose relative changes (e.g. halve poverty).

2. **Where no explicit SDG target is available, set upper bound to universal access or zero deprivation for the following types of indicators:**

   a. Measures of poverty (e.g. working poor), consistent with the SDG ambition to "end poverty in all its forms everywhere" (Goal 1).
   b. Public service coverage (e.g. preschool access).
   c. Access to basic infrastructure (e.g. broadband access, road conditions, etc.).
   d. Leave no one behind (e.g. workplace discrimination), consistent with the SDG ambition to eliminate disparate treatment for all vulnerable groups including those identified by race, indigenous status, religion, gender, sexual orientation, disability, poverty, location, and age.

3. **Where science-based targets exist that must be achieved by 2030 or later, use these to set 100% upper bound:** target value of 1.7 tons of CO_2/capita by 2050 as outlined in the Deep Decarbonization Pathways report for the United States (e.g. Goal 13: Energy-related CO_2 emissions).

4. **Where even the best performing states lag significantly behind the international community, and the indicator matches one used in international contexts, use the average of the top 5 OECD performers or the top 5 Global Index performers.**

5. **For all other indicators use the average of the top 5 performers.**

The lower bound for each indicator was determined using a two-step decision tree:

1. Use science-based thresholds for lowest acceptable or safe performance.

2. Use the 2.5 percentile score of the available data to account for outliers.

For both the upper and lower bounds:

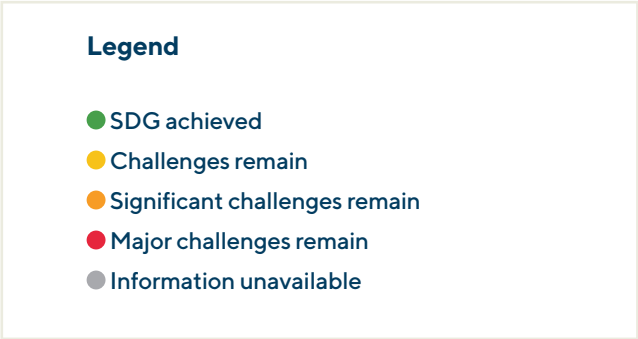
Each indicator distribution was censored, so that all values exceeding the target value scored 100, and values below the lower bound scored 0. In cases where the bounds were scientifically determined, the normalized score can be interpreted as percent of progress made towards achieving the SDGs, with 100% meaning achieving that indicator. In many cases, however, a score of zero is simply the lower benchmark of current progress of US states. In cases where the average of the top 5 is used to determine the score of '100', a '100' indicates only that this threshold level of achievement can be reasonably expected in the US context.

Calculating the index and assigning colors

Goal scores were created by taking the arithmetic average of the normalized indicator scores. Overall score was calculated by averaging the score for the 15 included SDGs.

Color scales were developed by creating interior thresholds that benchmark progress towards achieving the SDGs. The colors reflect the following scale:

Figure 25: Colors used in SDR



Source: SDSN

Green should not be interpreted as meeting the SDG indicator, but rather as an indication that the state is within range of achievement by 2030. Readers are cautioned to interpret this data in conjunction with data on rates of progress, as states could be slowing progress or moving away from achievement, or could be within range of achievement but not moving quickly enough to actually achieve the Goal by 2030.

Interior thresholds were developed, when available, by expert or scientifically determined levels. When this wasn't possible, interior thresholds were determined using summary statistics, such as using the mean (yellow/orange threshold) and the standard deviation (to set the yellow/green and orange/red thresholds) and then adjusted for clustering within the data. When the target value was more than two standard deviations away from the mean, colors were determined by evenly dividing distance to the target and adjusting for clustering. When there was just a three-point scale, three colors were used: red, yellow and green. The

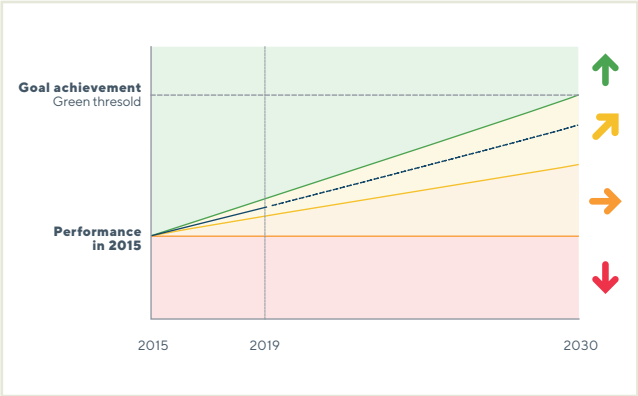
colors for Goal-level achievement were determined by mapping the indicator colors to a four-point scale (0-3), and then averaging the value across all indicators for a specific Goal.

Tracking trends over time

Historic data is used to estimate how fast a state has been progressing towards an SDG and determine whether —if extrapolated into the future—this pace will be sufficient to achieve the SDG by 2030. For each indicator, SDG achievement is defined by the Goal or achievement value (100 value) set for the SDG Dashboards. The difference in value between the target and the state value denotes the gap that must be closed to meet that goal. To estimate trends at the indicator level, we calculated the linear annual growth rates (annual percentage improvements) needed to achieve the target by 2030 (from 2015–2030), which we compared to the average annual growth rate over the most recent period (usually 2015–2019). Progress towards achievement on a particular indicator is described using a 4-arrow system (Figure 23). Figure 22 illustrates the methodology graphically.

Since projections are based on past growth rates over several years, a state may have observed a decline in performance over the past year (for instance due to the impact of COVID-19) but still be considered as being on track. This methodology emphasizes long-term structural changes over time since the adoption of the SDGs in 2015, with less emphasis on annual changes that may be cyclical or temporary.

Figure 26: Graphic representation of trends methodology



Source: SDSN

Figure 27: Four-arrow system for determining trends

Source: SDSN