SDG Push:
Methodological Note and User Guide of the SDG Push Diagnostic

The purpose of this guide is to provide an overview of the Sustainable Development Goals (SDG) Push Diagnostic. More specifically, it guides users on how to follow the three steps of the Diagnostic, together with the methodological background that was used to produce the Diagnostic. The three steps focus on SDG Trends, Current Priorities, Future Scenarios, and Interlinkages which can be used sequentially as part of the SDG Push process or as stand-alone investigations.

Overview

The SDG Push Diagnostic is a key element of the SDG Push: An Integrated Approach to Accelerate the SDGs, a process that combines analytical capabilities with qualitative methods to assess context, options and bottlenecks to determine effective policy choices, investments and pathways. The Diagnostic forms part of the Scoping Work, during which gaps, priorities, interlinkages and opportunities are identified.

The three steps of the Simulator focus on the following:

- **SDG Trends:** Visualizing data to easily identify SDG gaps, progress and interlinkages that allows for a holistic dialogue

- **Current Priorities:** Supporting faster real-time learning and identifying priorities using advanced text-analytics

- **Future Scenarios:** Mapping the projected impact of the SDG Push scenarios to provide an example of an ambitious set of actions that can accelerate SDG and national development goals.

SDG Trends

The SDG Trends section provides a straightforward visualization of current SDG progress for selected countries. The user can select the country of interest by clicking the 'Change Country' option, and the default country is South Africa.

The SDG Trends tab contains three sections: Current Trends, Target Overview, and SDG Overview. All these sections use four colour-coded, categorical scales to classify the progress of SDGs, targets, or indicators:

- Green for "On Track" —The country is on track to fulfill the SDG by 2030.

- Yellow for "For Review" – With current progress, the country will miss the SDG by 2030 by a small margin.

- Red for "Identified Gap" – With current progress, the country will miss the SDG by 2030 by a large margin.
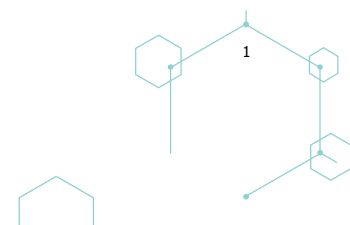
- Grey for "Gap Unidentified" – The country does not have enough data to identify the progress of the SDG.

Data source

All the time series data for all the indicators in all the SDGs are either fetched using UNStats API (https://unstats.un.org/sdgapi/swagger/#!/Goal/V1SdgGoalDataGet), or the data are provided by the government to the national country team.

The aim is to initiate a dialogue with the national country teams and governments, work with them to obtain the most up-to-date data and also define which time series they deem important to use for the simulator.

METHODOLOGY TO IDENTIFY THE CATEGORIES FOR EACH SUSTAINABLE DEVELOPMENT GOAL

A single **SDG** has multiple **targets;** each **target** has multiple **indicators**, and each **indicator** can have multiple time series**.** The data that are fetched from UNStats are for these time series. Some of the time series have very granular disaggregation, which might not be required in each case; therefore, key time series for each indicator are identified and filtered.

Once the key time series are identified and the data are available, the indicators are classified into five categories following the methodology given by UNStats (technical note link):

- Target achieved (if the country has already reached the target value)

- On track

- Fair progress but acceleration needed

- Limited or no progress

- Deterioration.

Please note that we are not able to classify all the indicators in these categories because for some, we might not have enough data for a country, or we might not have the target values for some indicators. This limits us to use only the time series that are mentioned in the technical note by UNStats. In some cases, after consulting with the governments, different UNDP country offices may provide target values for indicators for which we might not have a global target value.

All the time series within an indicator (since a single indicator can have multiple time series) are assigned scores of 1 to 4: 1, "*Target achieved*" or "*On track*"; 2, "*Fair progress but acceleration needed*"; 3, "*Limited or no progress*"; and 4, "*Deterioration*". All the time series for which the categories cannot be identified due to lack of sufficient data are ignored.

An average of the score of all the time series within an **indicator** is calculated (ignoring all the those for which the categories cannot be identified due to lack of sufficient data). This average value ranges from 1 to 4. Based on this average score, the **indicators** are classified into four categories (defined earlier) as follows:

- On Track (for scores of 1–1.5),

- For Review (for scores of 1.5–2.5)

- Identified Gap (for scores of 2.5–4)

- Gap Unidentified (for indicator with no time series with a category).
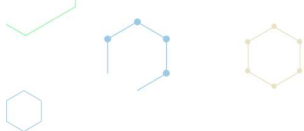
Using the above categorization, **indicators** within a **target** are assigned scores of 1 to 3 (1, "*On Track*"; 2, "*For Review*"; and 3, "*Identified Gaps*"). An average of the score of all the indicators within a target is calculated (ignoring all the indicators belonging to the category Gap Unidentified). This average value ranges from 1 to 3. Based on this average score, the **targets** are classified into four categories (defined earlier) as follows:

- On Track (for scores of 1 – 1.5)

- For Review (for scores of 1.5 – 2.5)

- Identified Gap (for scores of 2.5 – 3)

- Gap unidentified (for target with no indicators with a category).

Using the above categorization all the **targets** within an **SDG** are assigned scores of 1 to 3 (1 for "*On Track*", 2 for "*For Review*", and 3 for "*Identified Gaps*"). An average of the score of all the targets within an **SDG** is calculated (ignoring all the targets belonging to the category "Gap Unidentified"). This average value ranges from 1 to 3. Based on this average score, the **SDGs** are classified into four categories (defined earlier) as follows:

- On Track (for scores of 1 – 1.5)

- For Review (for scores of 1.5 – 2.5)

- Identified Gap (for scores of 2.5 – 3)

- Gap Unidentified (for target with no indicators with a category).

Using the above categorization all the **targets** within an **SDG** are assigned, each of the three sections and tabs are explained below.

SDG TRENDS

The section has a sub-header that provides a quick summary of how many SDGs out of the 17 SDGs fall under either the On Track, Identified Gap, or For Review categories. For example, for South Africa, it states, "2 are On Track, 8 are Identified Gaps, and 5 are For Review," as demonstrated in Figure 1.
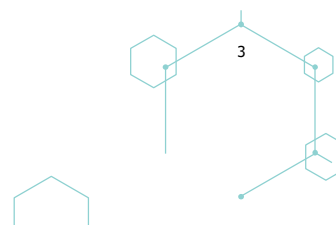
For **South Africa**, out of 17 SDG goals, **2 are On Track, 8 are Identified Gaps** and, **5 are For Review**

Figure 1. Summary of four categories of SDGs in South Africa. Available from: https://data.undp.org/diagnostic-simulator/current-sdg-gaps/ZAF

Followed by the sub-header is a colour-coded doughnut chart formed by the four categories that each SDG belongs to, as stated earlier. On the right of the chart is a list of the same SDG categories that form the chart.

The doughnut chart and the accompanying colour-coded, visual classification of a country's progress towards the 2030 Agenda for Sustainable Development is displayed per the country/region chosen by the user. Additionally, users can identify where the SDGs fall into each of the four classifications, including whether they are unidentified due to lack of data by looking at the legend to the right of the donut chart. If users wish to further explore the targets and indicators specifications for each SDG, they can click on the SDG to the right of the chart, and they will be directed to SDG Overview section of the page (more on this later).

SDG OVERVIEW

The *Target Overview* section is followed by the *SDG Overview* section, which provides a drop-down menu of all 17 SDG options that users can select from. For the given country/region earlier, the subsection below the drop-down is followed by a box that provides a quick summary of status of targets for that SDG. Then, it breaks down the selected SDG into its targets. In addition, it describes the target and colour-codes it to indicate to the user where the country stands in terms of achieving its 2030 Agenda for Sustainable Development. For example, for South Africa, when the user chooses SDG 1, it lists seven targets, 1.1, 1.2, 1.3, 1.4, 1.5, 1.a and 1.b, in three colours, with yellow indicating "For Review" for target 1.3, red indicating "Identified Gap" for target 1.1, and grey indicating missing data for the rest. Then, it lists each of these targets with the corresponding colour code and a clickable link with an option to expand further to explore all the indicators that comprise each of these targets. When the user clicks on the indicator, depending on data availability, it provides the user with a graphical representation of time series data for that indicator.

Current Priorities

The second tab, Current Priorities, provides an in-depth analysis of text documents with respect to the SDGs using natural language processing technologies. Accordingly, it serves a two-fold purpose: First, it provides

users with results of the analysis of the latest Voluntary National Report (VNR) submitted by the selected country. The VNRs are drawn from the website of the High-Level Political Forum on Sustainable Development. While most VNRs are written in English, some countries submit their reports in a different official United Nations language such as Arabic, Chinese, French, Russian or Spanish. In such cases, the analysis is performed on a machine-translated English version of the original document. For VNRs that are available in multiple languages, the English version is always used for the analysis.

Second, the Current Priorities tab enables users to go beyond VNRs and analyze a report, policy document or any other text document of interest by uploading an arbitrary PDF file. Note that this must be a searchable PDF and not a scanned copy of a document. It should also be noted that, unlike VNRs, in the October 2022 version, these documents must be in English since non-English texts will not be translated.

The analysis of VNRs and user-provided documents is performed using a **custom-built machine learning model for SDG classification.** The model analyses the document by linking its constituent pieces to SDGs and then aggregates the statistics before presenting the results to the user. The underlying assumption of the analysis is that the amount of text linked to each SDG defines how important each SDG is. In this context, the "importance" of an SDG is referred to as "salience". Thus, the most salient SDG is the one to which most of the text pieces, e.g., paragraphs, could be linked.

There are two main user interface (UI) elements that present the results of the analysis. The Relative Salience bar chart depicts the salience of all 17 SDGs. The most salient SDG, by definition, obtains a relative salience of 1.0, and other SDGs can have any value from 0.0 (i.e., this SDG is not covered in a text at all), to 1.0 (i.e., this SDG is the most prominent/salient in the text). This measure of relative salience is used to categorize SDGs into High-, Medium- and Low-priority tiers based on thresholds. The thresholds were determined such that, on average, approximately 33.3 percent of SDGs would fall into the Low-priority group across all VNRs.

Relative Salience

Relative Salience is a measure of the amount of text content linked to each SDG as compared to the Goal which is the most salient in the text. Relative Salience can help to understand which of the SDGs covered in the document[receive most attention and which ones are only briefly treated.
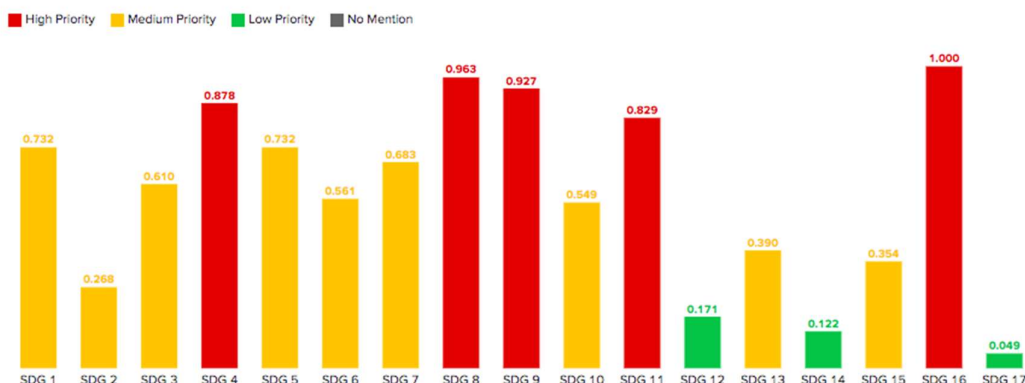
Below the chart, there is a *SDGs by Priorities* section, which shows all SDGs grouped by priority. By clicking on any of the SDGs, users can also see key topics that are discussed in the document in relation to the selected SDGs.

High Priorities (5)
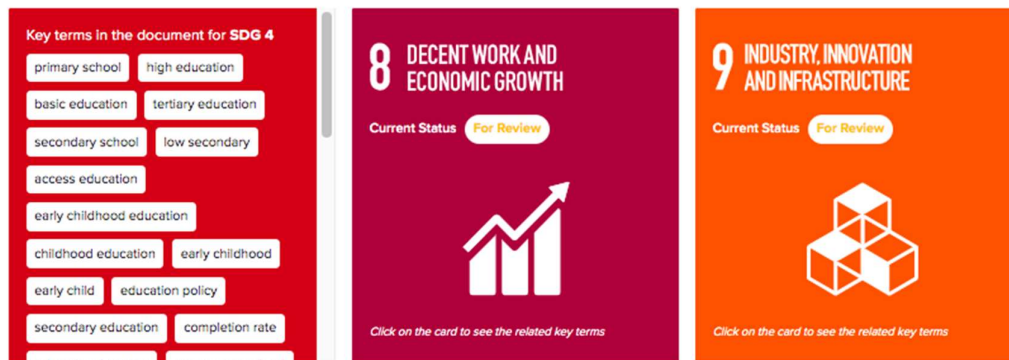
It is also possible to compare two documents by uploading two files. Here, the users are presented with a table that lists each SDG and its priority tier in both documents side by side. This may be useful to compare VNRs from different years for the same country, two VNRs from two different countries, or two completely different policy documents.

Future Scenarios

The final tab of the main page is labelled "Future Scenarios". The results in this section build on UNDP's SDG Push scenario, which features an ambitious yet feasible set of SDG accelerators across governance, social protection, green recovery and digitalization – the diagnostic supports governments to mainstream and disseminate accelerators to recover better from the COVID-19 pandemic.

It has two parts: the first part describes Future Scenarios, which is based on the International Futures (IFs) model used to forecast the development path and the effects of combined set of interventions in the areas of Governance, Social Protection, Green Economy, and Digital Disruption across different and predefined set of scenarios. IFs is an open-source integrated assessment modeling tool that simulates interactions within and across 186 countries and 12 core systems: agriculture, demographics, economics, education, energy, environment, finance, governance, health, infrastructure, international politics and technology. These systems are dynamically connected so that changes in one system may lead to changes across all others, which can illuminate spillover effects as well as tradeoffs and synergies with different policy choices.

The figure below shows how the models within the IFs correspond to the 5 Ps of the SDGs and associated SDGs.
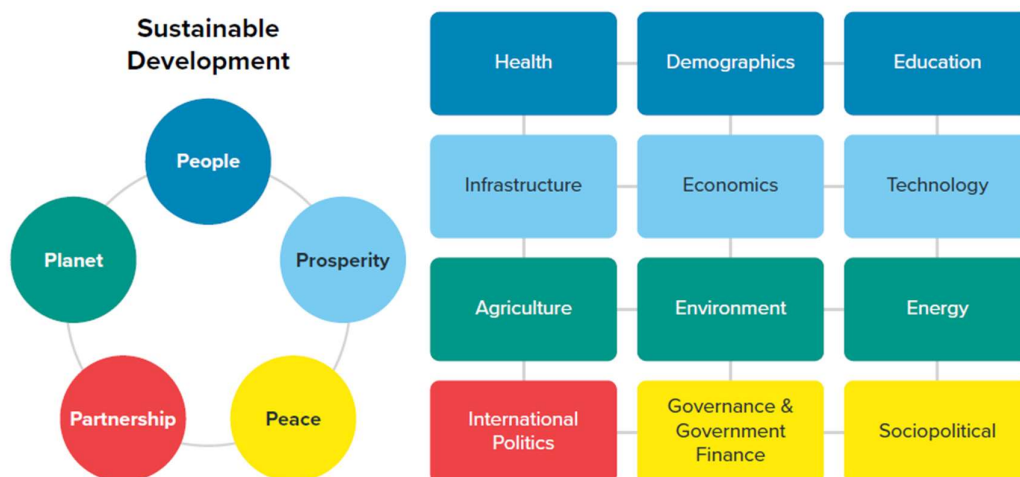
The second part illustrates the impact of Future Scenario investments in five SDGs and its targets and compares it with the baseline scenario where such investments are not undertaken. The time series graphs show progress until 2050 on a yearly basis for every target related to specific SDG tab shown in the subheading. Each of these graphs provides two scenarios for each country in terms of what the gaps would look like with and without the SDG Push scenarios, colour-coded in green and red, respectively.

For example, for the default country South Africa, users can see that SGD 1 tab produces three graphs, one for poverty set at under $1.90 per day by number of people and the second by percent of population, while the third one show female poverty headcount by number of people. The top line in each of these graphs is shown as a red line is a scenario when the SDG Push scenario is not implemented; the bottom line is when the SDG Push scenario is implemented and shown as a green line. Therefore, by having these graph lines right on top of each other, users can easily identify the gaps with and without the push scenario. See an example below.
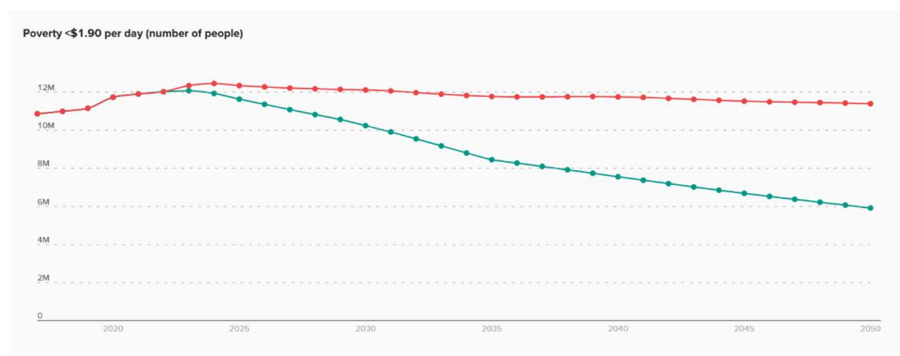

Figure 1. Projected number of people in extreme poverty in South Africa under Baseline and SDG Push scenario
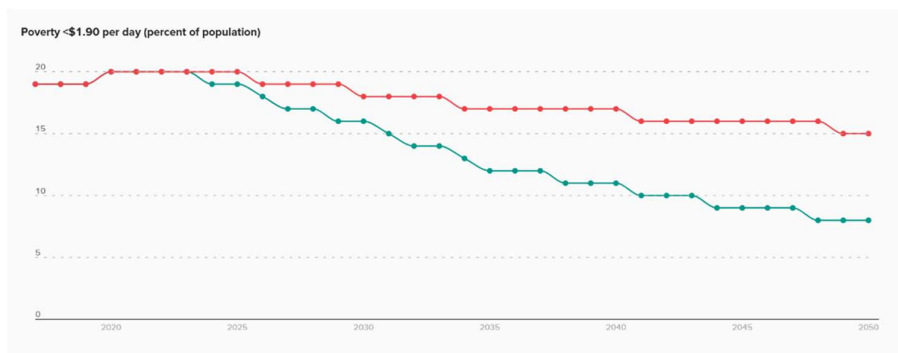
Figure 2. Projected percent of people in extreme poverty in South Africa under Baseline and SDG Push scenario
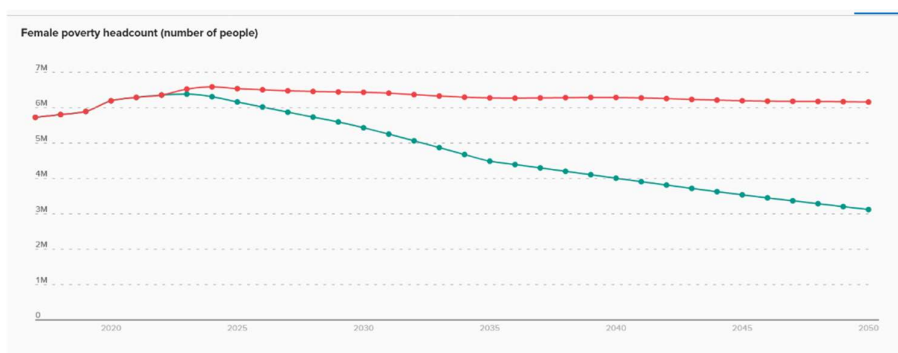

Figure 3. Projected female poverty numbers in extreme poverty in South Africa under Baseline and SDG Push scenario

SDG Interlinkages: Synergies and Tradeoffs

The final section covers the SDG Interlinkages and opportunities for accelerations. The SDGs do not exist in silos, understanding how the goals are interconnected, both positively and negatively, is essential to understanding the mechanisms for achieving the targets. Identifying the most interlinked and strongest SDGs is important for acceleration. Thus, important to interrogate the interlinkages at the national level. Examine how the 169 targets are interconnected and how this relates to the SDGs that are not on track to reach the 2030 goals.

The Interlinkages colour-coded figure allows users to select any target for which a visual display of the target and its synergies with other targets (represented by green lines) are presented in default mode. Furthermore, users can click on the tradeoff tab to obtain a visual presentation of the target and its tradeoff (instead of synergies), represented by red lines.

The synergies and trade-offs are global, which means they are the same for all countries as mapped by the European Commission. However, the traffic light colouring of the targets is customised based on each countries data on the SDGs. See section "Methodology to Identify the Categories for each SDGs" above. To learn more about the methodology behind the Synergies and Trade-offs see details here.

Targets are greyed out because of lack of data, however, the information on interlinkages still applies.
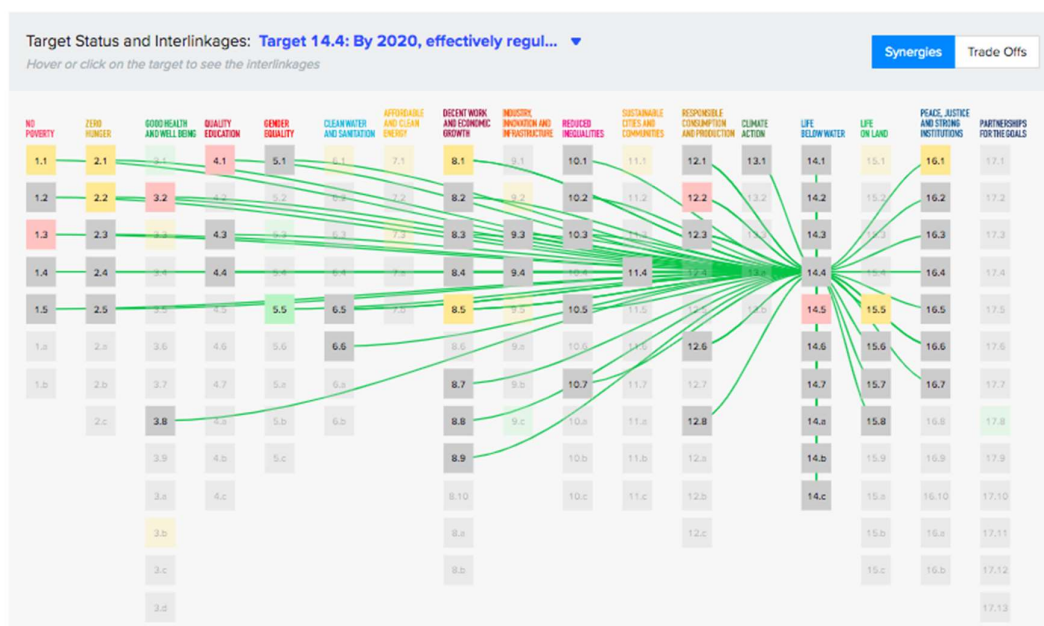
For example, for Indonesia, SDG 11 (Sustainable Cities and Communities) is identified as "Identified Gap". Through the visualisation the user can explore the target 11.1 (marked as "Identified Gap") to see how this target is interlinked both positively and negatively with other targets. Within Synergies, three other "Identified Gap" targets would benefit from improvements in 11.1 (3.3, 5.5, 8.5). Conversely, under trade-offs, 11.1 has counterproductive interlinkages with three other targets (2.2, 5.5, 14.5). This shows that two targets (11.1 and 5.5) may have synergies and trade-offs simultaneously.

Important to note is that multiple targets might be greyed out due to insufficient data, however, these might be important targets to consider based on the country-context.