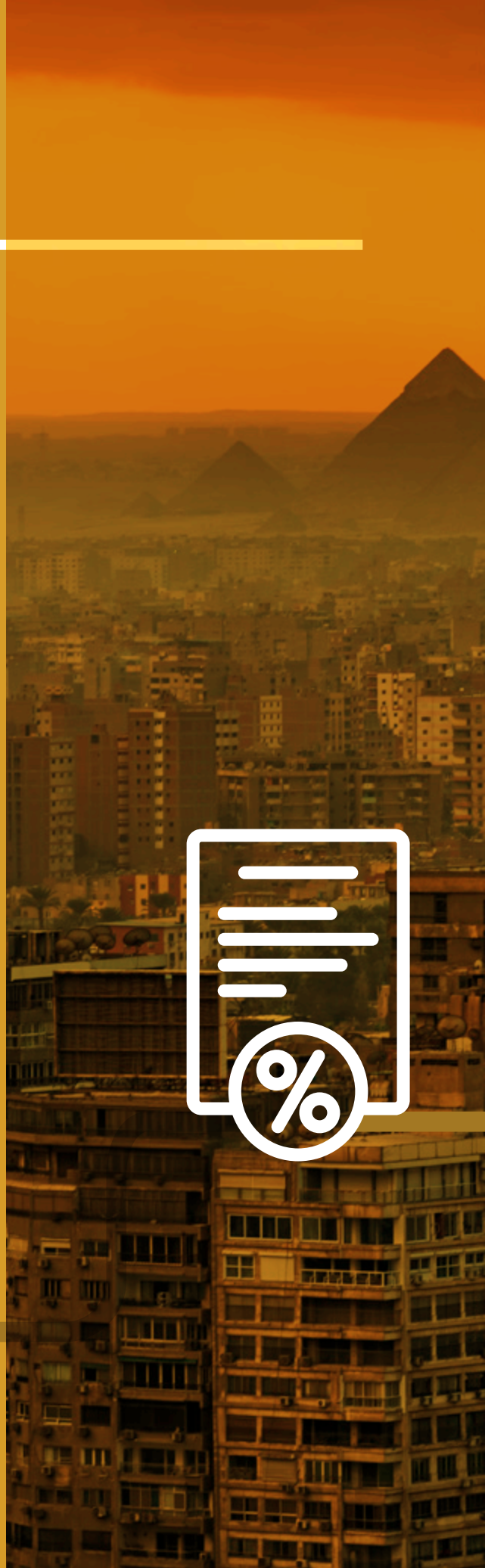
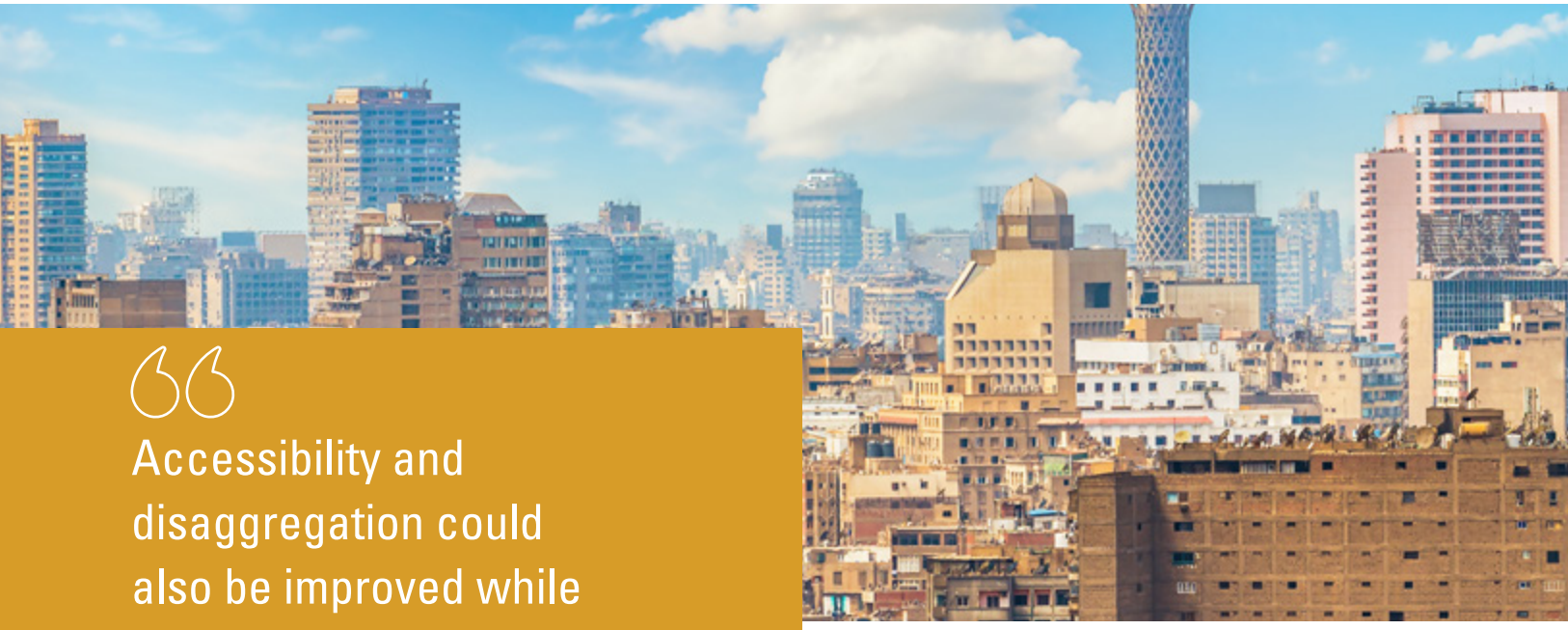
Data and data systems

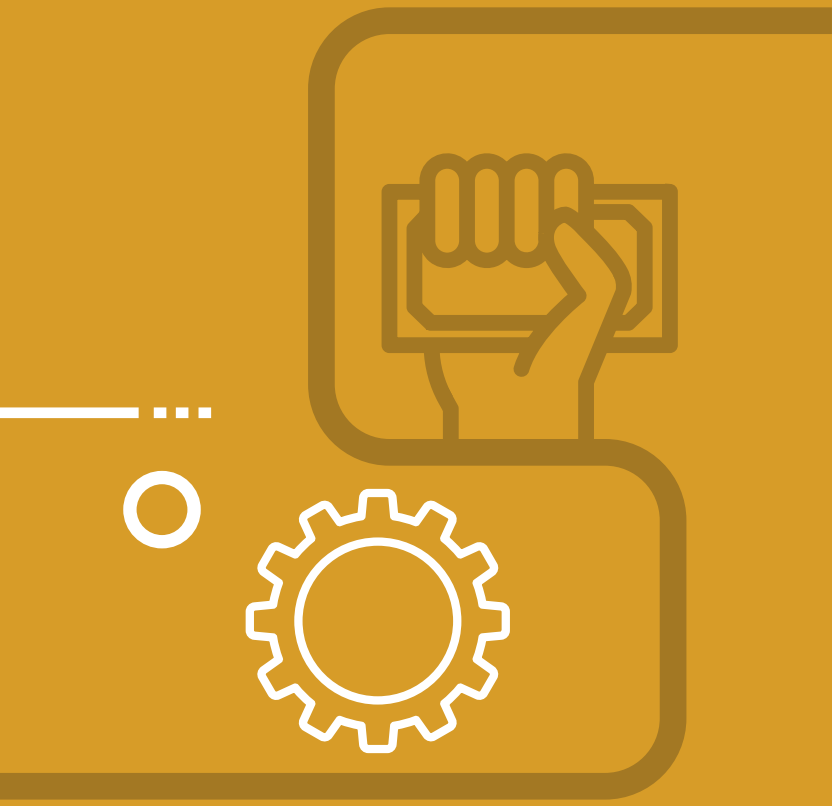
by Mazen Hassan and Engi Amin



02



> Accessibility and disaggregation could also be improved while maintaining confidentiality.



A. The role of data availability in policy selection

Data are produced at an unprecedented scale. According to Oracle, a computer technology corporation, approximately 90 per cent of the world's data were created in the last two years. Every two years, the volume of data across the world doubles in size. Data have become an essential part of running operations and making decisions in both the public and private spheres. This chapter focuses on the connection between data and sustainable development in Egypt by investigating three questions:

90%

of the world's data were created in the last two years.

> Data collected by various government institutions are scattered across different platforms, published in different formats and are not integrated in a central database.

1. To what extent are data on the SDGs available, accessible and usable in Egypt, and how could they help the Government to identify possible gaps and courses of action?

2. What possible bottlenecks need to be addressed to improve the speed and efficiency of the Government's data generating process?

3. Even if not fully updated or complete, could available data inform the Government on where best to mobilize resources in order to achieve the maximum possible impact for development?

The objective of this chapter is to encourage investing in efficient data systems to further support data-driven decisions related to development. Section B demonstrates the significance of generating reliable data for the current situation in Egypt. Section C assesses official public data, particularly on the SDGs, using various data quality parameters (primarily availability and timeliness). Section D presents a qualitative analysis of data generation bottlenecks that, if addressed, could make the process much smoother. Section E shows the result of a brief exercise demonstrating how data could drive development policy trade-offs. Section F concludes with policy recommendations.

B. The need for data-driven decisions

Given the country's ongoing and unprecedented level of public spending and reforms, the need for data-driven decisions has never been greater. Before exploring the three questions mentioned in the previous section, it is worth exploring the arguments in favour of investing in and making use of credible data systems.

The Egyptian Government is embarking on a significant public investment programme, likely

to be of unprecedented scale. Deciding on a course of action has therefore never been more consequential. The Government has spent over LE 4 trillion on public investments in the past six years, a figure that is unmatched by any comparable period in recent history. As a result, spending decisions have become extremely consequential. Given that spending levels might not continue to rise for long, the Government should make the most of current opportunities. It is therefore clear that public spending decisions must be data-driven. There are countless examples, including determining whether to continue subsidizing school meals, prioritize additional metro lines or increased capacity for ports, and prioritize schools over hospitals and vice versa. The price tag for each of these decisions has never been higher, given the large and justified appetite for public investments. The data guiding these decisions must therefore be reliable and complete.

In addition, Egypt is undergoing a transformational institutional reform. Parliament is establishing laws at an almost incomparable rate; nearly 900 laws were passed in the most recent legislative term. The Government is also expanding institutionally. New organizations are being created, new cities are being constructed and a new administrative capital is being erected. In this context, data are particularly relevant to continuously monitor performance and to ensure that expanding State institutions are fit for purpose. More importantly, data are needed to pre-emptively avert potentially harmful civil service practices, which far too often accompany government expansion in emerging economies.

Lastly, governments in general, and the Egyptian Government in particular, have perhaps never been in greater need of objective, neutral and impression-free foundations for decision-making. They are under rising pressure to justify decisions, a task that could be better accomplished with greater access to data-driven arguments. In Egypt, where a majority of young people aged 18 to 35 years are informed via social media, the public debate over government decisions must be continuously enriched with sound, data-informed arguments. The more data are made available, and the more credible and up to date they are, the more informed the public will be. There would also be less room for baseless or politicized views. Furthermore, decision-making processes inside the Government itself would benefit from impartial data to make better choices.

C. Assessment of multiple data quality parameters

This section presents the results of an evaluation of the availability, accessibility and completeness of official data on development indicators produced by the main statistical institutions in Egypt.[1]

1. The Central Agency for Public Mobilization and Statistics

The Central Agency for Public Mobilization and Statistics (CAPMAS) is the official statistical agency of Egypt responsible for collecting, storing, processing and publishing data in all major national and State fields at the national, governorate and subgovernorate levels.[2] CAPMAS is also the only official agency charged with conducting the national population and housing census and has recently begun conducting an economic census as well.

This section assesses the quality of data publicly available on the CAPMAS website and published in its reports and bulletins. To this end, the seven data quality dimensions adopted by OECD will be examined (i.e. accessibility, coherence, credibility, accuracy, interpretability, timeliness

and relevance).[3] Additionally, three other data dimensions that are particularly relevant to this exercise will be assessed: data completeness, and the frequency and level of data collection.[4]

Accessibility refers to the extent to which data are available and whether they can be retrieved quickly and easily. In general, primary data produced by CAPMAS are not publicly available.[5] The website does offer public access to select reports, such as the economic census and an annual statistical yearbook that includes descriptive tables and charts for demographic, social and economic data collected throughout the year. These data, however, are only available in Portable Document Format (PDF), which makes processing them for further analysis very time-consuming. Registered users can download additional publications, such as monthly, quarterly, biannual and annual statistical bulletins that contain data tables for several indicators (also available in PDF rather than spreadsheets). There are macro time-series data for a few indicators that can be downloaded as a spreadsheet but only with a maximum of two columns (year and value). This format makes it difficult to build data files with multiple variables, which are necessary to perform in-depth statistical analyses. The CAPMAS website does incorporate an interactive tool that contains data from the 1996 and 2006 censuses. It allows users to create queries and generate tables, charts and maps. Unfortunately, data from the 2017 census are not available. In conclusion, although some valuable data sets appear to be available, accessibility requires further improvement.

Completeness refers to the number of missing entries in the data. CAPMAS collects data at the micro and macro levels.[6] At the micro or individual level, data are collected primarily through surveys. Since primary data are not publicly available, the only way to assess the completeness of individual-level data is through the response rates and sampling procedures provided in the appendices or metadata sections of published reports. Based on these criteria, CAPMAS enjoys a respectable reputation for following sound sampling and statistical procedures when conducting surveys. At the macro level, the indicators, particularly economic indicators, are mostly complete as a result of the automated nature of data collection (e.g. exchange rates, stock market performance, financial accounts, etc.). Other national institutions are also responsible for monitoring national economic performance indicators, namely CBE, the Ministry of Finance, and the Ministry of Planning and Economic Development, in collaboration with CAPMAS under the Special Data Dissemination Standard framework.

Coherence refers to the extent to which data are presented in the same format for comparability across different time points and data sets. CAPMAS closely follows international statistical standards in the collection and reporting of internationally defined indicators. There is coherence in terms of the units of measurement as well as consistent representation of data by CAPMAS, thus ensuring comparability across time and for country comparisons.

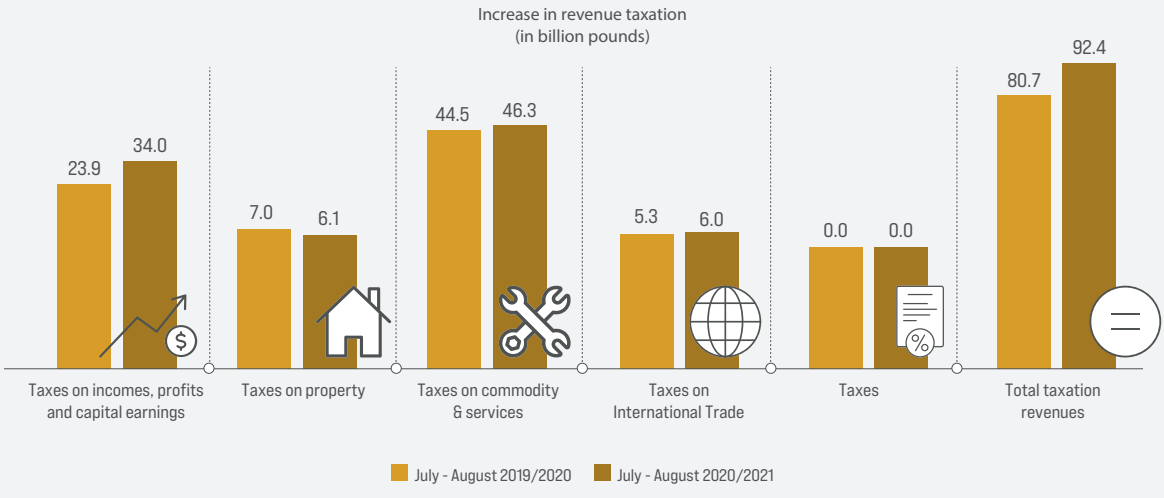
Credibility refers to the reputation of the data producer and user confidence. Accuracy refers to the degree to which data are correctly estimated and described. CAPMAS does enjoy an established track record as the main statistical body in Egypt, and the Agency follows international protocols for reporting statistics. Every report published by CAPMAS includes a methodology and/or a metadata section that documents all sampling, collection, statistical procedures and calculations to ensure both credibility and accuracy.

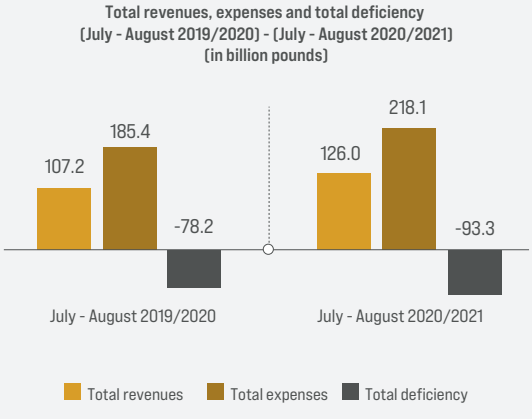
Interpretability refers to the ease with which users can understand the data produced and the extent to which data are reported in appropriate units, language, definition, etc. As mentioned, any data or statistics reported by CAPMAS are accompanied by full documentation on the way in which they were collected, produced and measured. All available figures and graphs contain defining information and variable description.

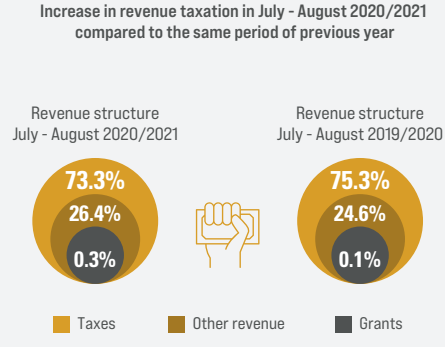
With regard to the frequency of data collection, this section focuses on data that must be collected from the field and are not generated automatically, particularly nationwide surveys tracing longitudinal, SDG-related data on the individual.[7] In Egypt, such surveys include the Household Income, Expenditure and Consumption Survey, the Demographic and Health Survey, and the Labour Force Survey. Since 2008/09, the Household Income, Expenditure and Consumption Survey has been conducted every two years instead of every five; however, the last available survey dates to 2017/18. As for the Demographic and Health Survey, its standard version is carried out by the Ministry of Health and Population rather than CAPMAS, and it is typically fielded every five years to form a database with longitudinal data. There are also interim surveys, fielded between the standard surveys, to provide updated data on key performance monitoring indicators. The latest standard survey, however, was conducted in 2014, and no interim surveys have been conducted to date, to the authors' best knowledge. Lastly, CAPMAS conducts a quarterly Labour Force Survey that includes longitudinal data on the size of the labour force, its characteristics and geographical distribution. This Survey also experiences some delays in the publication of its data, as will be discussed in the following paragraph. In conclusion, delays exist in fielding many of these longitudinal surveys, some of which were due in 2019 and 2020 but were not fielded until the first quarter of 2021. The COVID-19 pandemic was partially to blame and also exacerbated such delays. To address this problem, examples from developed countries provide innovative solutions to fielding censuses during the pandemic.[8] Although phone surveys present a number of constraints, such options could be explored in order to collect data on crucial SDG indicators quickly and to avoid any further delays. In fact, the urgency of the COVID-19 pandemic drove several United Nations bodies to conduct a number of phone surveys in collaboration with CAPMAS and relevant line ministries, with a view to assessing the pandemic's impact on vulnerable groups such as women and MSMEs in 2020. These rapid data collection efforts prove that data could be collected more quickly. The results of such crisis-driven surveys could also be linked to the SDG framework and indicators for monitoring purposes.

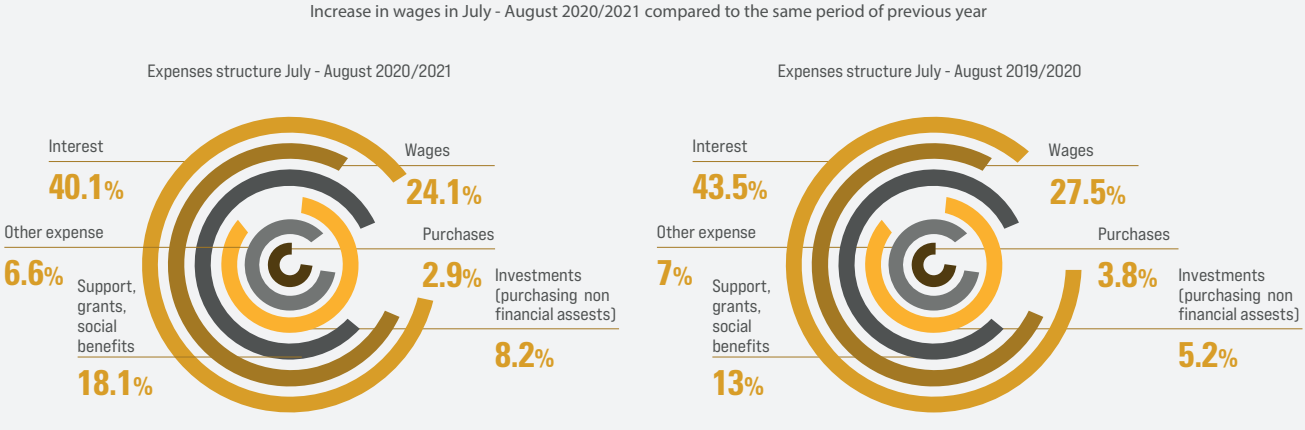
Timeliness refers to the extent to which data are up to date and whether lags exist between the observation date and the publication date. CAPMAS reports are published based on the frequency of data collection and measurement: monthly, quarterly, biannual, annual and special reports. Since data can only be accessed through reports and bulletins, time lags exist between the date on which data are measured and their publication on the website. For example, monthly bulletins provide information on indicators that are primarily economic, such as average consumer prices for most food commodities, consumer and producer price indices and foreign trade. Although a separate report is published for each of these indices every month, there is usually a lag of two to three months, sometimes more, before the bulletin is published. Figure 20 shows some statistics from the Informatics Monthly Statistical Bulletin for November 2020, published by CAPMAS, which was made available on the website at the beginning of December 2020. Most of the statistics are from August 2020 or earlier; very few indicators are actually from November 2020. Although improvements have been made regarding data timeliness compared to previous years, it could be improved even further. Quarterly bulletins, such as the Labour Force Survey, are also published at the end of the following quarter. An examination of these reports also reveals several lags in the reported data for biannual and annual bulletins. Most of the indicators reported (e.g. those related to information and communications technology (ICT), utilities and services, health care, security and safety) have a lag of one year. As such, the yearly bulletins published at the end of 2020 contain figures from 2019.[9] Timeliness therefore is indeed a significant drawback to the data made available through the CAPMAS website. Significant improvements could be made if data are collected and made available more quickly.

Figure 20. Excerpt from the CAPMAS informatics monthly statistical bulletin of November 2020









Source: CAPMAS, Egypt, 2020. Informatics Monthly Statistical Bulletin (November), pp. 21-22.

The level of data, or disaggregation, reflects the level at which indicators are broken down by subcategories such as gender, income, age, employment sectors, geographic location and others. Disaggregated data are a core requirement for measuring and keeping track of sustainable development indicators. They reveal inequalities between different sub-categories that are essential for policymaking. Demographic indicators reported in CAPMAS' bulletins and reports are usually sub-grouped by gender, age groups, urban/rural areas, and sometimes governorates. However, there is more room for reporting statistics using further disaggregation, especially on a geographic and administrative level. Egypt is divided into three levels of administrative hierarchies (governorates; cities/districts (marakez) and neighbourhoods/ villages). Each level at the same hierarchy is heterogeneous, making the task of measuring development indicators indispensable.

Relevance assesses the extent to which data produced are relevant to the intended task. Based on the analysis conducted, data produced on a majority of the SDGs and their respective indicators speak to the designated indicator. Although published data might be outdated or incomplete, they correctly track relevant phenomena or conditions.

In conclusion, at least three dimensions could be improved to realize greater potential from CAPMAS data in monitoring performance on the SDGs and in guiding policies. First, data must be made available more quickly to avoid diminishing their value. Second, accessibility and disaggregation could also be improved while maintaining confidentiality. Interactive platforms could be integrated to offer instant graphs, multiple selections, trend analysis, breakdowns by quantile and other features.[10] Disaggregation requires gathering more data at subnational levels, which might mean higher costs. Nevertheless, given the high scores on data reliability and consistency at CAPMAS, it is unfortunate that the full potential of data cannot be realized because of delays in publication or the use of formats that complicate data processing.

Lastly, even though CAPMAS is the country's official data hub, it serves as the primary source of data for fewer than half of the SDG indicators, covering only 8 of the 17 Goals in the 2021 voluntary national review (VNR). For the remaining Goals, various ministries are responsible for compiling data. These include the Ministries of Education and Technical Education, Health and Population, Water Resources and Irrigation, Electricity and Renewable Energy, and Environment. Since their primary function is to provide public services, compiling SDG data and making them available to the public or to CAPMAS is not a priority. It is therefore essential to improve the ability of these ministries to collect relevant data and share them more efficiently with CAPMAS or a wider audience.

It should be noted that the National Strategy for the Development of Statistics was launched by CAPMAS in 2019/20. It aims to achieve better coordination between data producers and data users and make available the data required for development decisions.[11] Such a strategy is an essential first step. As it is still in the early stages, its progress and funding requirements must be closely monitored in order to achieve these objectives and carry out the suggested improvements.

2. Other public institutions providing data related to the Sustainable Development Goals

Other national institutions collaborate with CAPMAS to produce and disseminate economic and financial statistics and data.[12] This section focuses on four of these institutions: CBE, which provides data for the financial, banking and external sectors; the Ministry of Finance, which provides public finance statistics; the Ministry of Planning and Economic Development, which provides national accounts and production index statistics based on data from other ministries as well; and the Egyptian Cabinet's Information and Decision Support Center.

The CBE website has an online portal that publishes data on a monthly basis for economic indicators regarding the performance of the financial and banking sectors in Egypt. Statistics on inflation rates, exchange rates, liquidity, interbank rates and volumes, economic indices and stocks are all publicly available and downloadable in the form of spreadsheets. The statistics are complete, consistent and provided in a timely manner every month. Given the macro nature of economic and monetary data and statistics, they are expected to be of higher quality than demographic and social microdata. It is also important to mention that the design and user engagement of the CBE website is much better than that of CAPMAS.

The Ministry of Finance has recently updated its website and data portal to include a monthly financial bulletin that reports statistics on general economic and financial outlooks, sector indicators, domestic prices, government debt and fiscal and monetary sectors. The bulletin also has a comparative analysis section that compares the performance of Egypt with other comparable countries in key economic indicators. Statistics on fiscal indicators (i.e. tax/non-tax revenues and expenditures breakdown) can be downloaded in a spreadsheet format. The website has undergone significant improvements to incorporate interactive components, adopted a user-friendly interface, and made important fiscal data and performance indicators available online. Nevertheless, a timeliness issue remains. There is a lag of approximately two to three months for the monthly financial bulletin.[13] Moreover, some indicators and data included in these bulletins have a lag of one to two months (table 2). It should be noted that these data lags are much shorter than those found on other governmental websites, hence data are more up to date.

Table 2. Credit provided by Banks

	Jun - 16	Jun - 17	Jun - 18	Jun - 19	Jun - 20	Dec - 20	Jan - 21	Feb - 21	Mar - 21
Total lending	**717,999**	**925,660**	**1,426,457**	**1,629,664**	**1,854,326**	**2,493,370**	**2,532,469**	**2,573,022**	**2,648,549**
	(22.1)	(30.4)	(51.3)	(14.2)	(13.8)	(30.5)	(24.0)	(22.9)	(22.9)
To Government[2/]	**66,421**	**172,047**	**354,723**	**452,917**	**542,446**	**853,888**	**861,833**	**881,110**	**932.049**
	(62.8)	(173.1)	(104.2)	(27.7)	(19.8)	(60.2)	(47.6)	(37.7)	(41.4)
In local currency	10,855	100,473	142,710	216,549	235,838	560,590	587,139	607,138	658,937
In foreign currency	55,566	71,574	212,013	236,368	306,608	275,298	274,694	273,972	273,112
The non-government	**651,578**	**753,613**	**1,071,734**	**1,176,747**	**1,311,880**	**1,657,481**	**1,670,636**	**1,691,912**	**1,716,500**
	(19.1)	(16.5)	(39.4)	(9.8)	(11.5)	(19.4)	(14.6)	(16.3)	(14.8)
In local currency	468,502	556,967	724,503	849,775	1,038,221	1,406,849	1,420,939	1,441,445	1,469,754
In foreign currency	183,076	196,646	347,231	326,972	273,659	250,633	249,697	250,467	246,746
Memorandum items (per cent)									
Credit to private sector[3/] / Total credit	82.0	72.0	64.7	62.4	62.0	60.1	59.8	60.2	59.4
Non government loans/Deposits[4/]	43.8	43.6	42.6	38.8	38.6	38.3	38.3	38.2	37.8
Government loans/Deposits	26.3	49.0	67.6	84.9	88.8	96.2	93.6	93.2	95.0
Foreign currency denominated credit to total credit	33.2	29.0	39.2	34.6	31.3	21.1	20.7	20.4	19.6
Government foreign currency denominated credit to total government credit	83.7	41.6	59.8	52.2	56.5	32.9	31.9	31.1	29.3
Denominated credit to total non-government credit	28.1	26.1	32.4	27.8	20.9	15.1	14.9	14.8	14.4

Source: Ministry of Finance, Egypt, 2021. The Financial Monthly, vol. 16, No. 7 (May), p. 50.

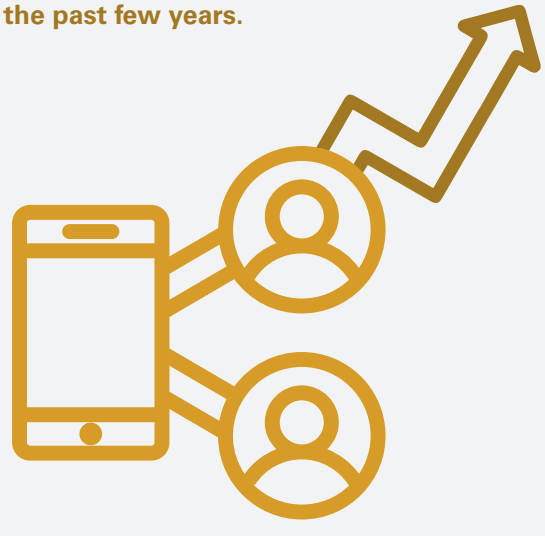
The Ministry of Planning and Economic Development is responsible for establishing sustainable development plans and implementing the country's strategic vision in coordination with other ministries, stakeholders and development partners. It is the national institution in charge of developing the statistical capabilities to measure the impact, performance and project completion rates to achieve the SDGs and the Egypt Vision 2030. The Ministry's website has a National Accounts interactive data portal that contains statistics on the gross domestic product (GDP) and public investments calculated at current and fixed prices. It can be queried for various economic activities, time frequencies and governorates. The Ministry collaborates with CAPMAS, CBE, and the Ministry of Finance to produce and publish these statistics. According to the website of the Ministry of Planning and Economic Development, Egypt meets the requirements of the IMF Special Standard for Data Dissemination in the dissemination of economic and financial data to the public with regard to the timing of publication and the availability of economic and financial data (i.e. on a quarterly basis with a maximum lag of three months).

The Information and Decision Support Center, founded in 1985, is considered the think tank of the Egyptian Cabinet and provides information and expertise to support policy and decision-making. One of the Center's recent strategic goals is to provide information and knowledge not only to support policymaking but also to support the transformation to an information-based society. The Center conducts public opinion polls and publishes the results in periodic reports. In 2021, it created a dedicated portal and mobile application that provide a number of national, African and global data points, performance indicators and statistics. In addition to providing data in the form of downloadable reports and Excel tables, there are also interactive dashboards and maps. The mobile application increases accessibility for a larger sector of the public, given that the number of mobile users in Egypt has grown exponentially in the past few years. These data are not generated by the Center itself but are predominantly collected and provided by other sources.

Other ministries that provide SDG-related data include the Ministry of Communications and Information Technology, which publishes a monthly report on ICT indicators. In 2020, the Ministry of International Cooperation established a mechanism to analyse the contribution of development cooperation to the SDGs by mapping all ongoing projects in its portfolio to align them with the relevant SDGs. Within this mapping exercise, two distinct approaches were used: (a) single SDG mapping, which identified the primary SDG towards which each project contributed; (b) multiple SDG mapping, which considered the multidimensional linkages that projects might have to different SDGs.[14]

Until 2017, the Ministry of Environment published a yearly State of the Environment Report, comprising data related to environmental and climate change indicators. The Ministry of Trade and Industry also has an interactive dashboard (in Arabic) with indicators related to economic development. The Ministry of Electricity and Renewable Energy publishes yearly reports on indicators related to energy, the most recent of

The number of mobile users in Egypt has grown exponentially in the past few years.



which was for 2018/19. Although the Ministry of Health and Population is primarily responsible for the Demographic and Health Survey and other surveys from which all data related to health indicators are derived, it does not have a data portal. Instead, health data are made available through international institutions such as the Demographic and Health Surveys Program or UNDP. It is therefore clear that data collected by various government institutions are scattered across different platforms, published in different formats and are not integrated in a central database (such as CAPMAS). While initiatives by ministries or CBE to disclose information and publish data are to be commended, the data collection process would be accelerated by assigning the responsibility of SDG indicators to one entity and ensuring coordination among government bodies.

3. The need to complete and update data on Sustainable Development Goals

The United Nations classifies SDG indicators into three tiers based on the following criteria: whether the indicator is conceptually clear, whether it has an internationally established methodology and whether data are regularly produced by a majority of countries.[15] An indicator that meets the three criteria is placed in Tier 1. An indicator with a clear concept and established methodology that lacks regular data belongs to Tier 2. An indicator that does not meet the three criteria is placed in Tier 3. At the time of writing this chapter, three major national sources had reported on the availability of SDG data and statistics in Egypt.

The VNR was published by the Ministry of Planning and Economic Development with the support of UNDP in 2018 and 2021.[16] Only 35.7 per cent of indicators in the 2018 report are classified as Tier 1. Some of these indicators were further refined by the United Nations Statistics Division (UNSD) in 2020 in order to measure goals with higher specificity,[17] which further reduced the percentage of Tier 1 indicators. Irrespective of the tier classification, the 2018 report states that data exist on 106 of the 244 indicators, or 43.5 per cent. That number increased to 47.5 per cent in the 2021 VNR.[18]

According to the National Statistical Report for Monitoring SDGs in Egypt, published by CAPMAS in 2019, data are available for 116 indicators (47.5 per cent).

The website of the Egypt SDG Observatory, operated by CAPMAS, is an official national digital platform for SDG indicators and data. As at 14 October 2021, data on the website were mostly outdated, not sufficiently disaggregated and missing many key indicators that were present in both the 2018 VNR and the 2019 CAPMAS National Statistical Report. Only 53 of the 247 indicators are available (or 21.5 per cent). Table 3 displays the extent to which data on the SDGs (as defined by the 2020 global indicator framework) are available for Egypt through the official SDG Observatory website.

It is worth mentioning that an independent Egyptian Women's Observatory provides data on indicators related to women's empowerment and gender equality (Goal 5), as well as some updated statistics related to other indicators. It is published by the National Council for Women in cooperation with Baseera, the Egyptian Center for Public Opinion Research, an independent and non-partisan organization established in 2012.

Data are available for 116 indicators of 224

47.5%

in 2021 VNR

Table 3. Tallying available Sustainable Development Goal indicators on the SDG Observatory

SDG	Number of targets	Total number of indicators	Available	Missing	Date of most recent indicator	Indicators with at least one level of disaggregation
1. No poverty	7	13	4 (30.8%)	9 (69.2%)	2016	2
2. Zero hunger	8	14	5 (35.7%)	9 (64.3%)	2016	2
3. Good health and well-being	13	28	9 (32.1%)	19 (67.9%)	2016	5
4. Quality education	10	12	4 (33.3%)	8 (66.7%)	2017	3
5. Gender equality	9	14	7 (50%)	7 (50%)	2017	5
6. Clean water and sanitation	8	11	1 (9.1%)	10 (90.9%)	2015	0
7. Affordable and clean energy	5	6	1 (16.7%)	5 (83.3%)	2016	0
8. Decent work and economic growth	12	16	6 (37.5%)	10 (62.5%)	2016	4
9. Industry, innovation and infrastructure	8	12	4 (33.3%)	8 (66.7%)	2016	0
10. Reduced inequalities	10	14	1 (7.1%)	13 (92.9%)	-	0
11. Sustainable cities and communities	10	14	1 (7.1%)	13 (92.9%)	2015	1
12. Responsible consumption and production	11	13	1 (7.7%)	12 (92.3%)	-	0
13. Climate action	5	8	0 (0%)	8 (100%)		0
14. Life below water	10	10	2 (20%)	8 (80%)	2015	0
15. Life on land	12	14	1 (7.1%)	13 (92.9%)	2015	0
16. Peace, justice and strong institutions	12	24	2 (8.3%)	22 (91.7%)	2015	1
17. Partnerships for the Goals	19	24	4 (16.7%)	20 (83.3%)	2017	1
Total	169	247	53 (21.5%)	194 (78.5%)		24

Source: Author, with information from the Egypt SDG Observatory dashboard.
Note: Produced in October 2020 and updated in October 2021.

The discrepancy in the percentages on data availability reported in the three national resources highlights a problem with data integration. SDG data are scattered across various repositories, making it more difficult to assess data availability and quality. A mechanism for integrating existing and new sources of data is therefore essential in order to build an inclusive national database that can guide policies more effectively.

Additionally, delving deeper into data gaps per indicator in these reports[19] shows that some SDGs have a particularly small number of indicators for which data are available (Goals 11 to 15, which are mainly related to the environment). Goal 11 on sustainable cities and communities, for example, has only three data points in the CAPMAS report and one data point in the Observatory (out of 14). Goal 12 on responsible consumption and production has only one data point in the CAPMAS report and one in the Observatory (out of 13). Goal 13 on climate action has three data points in the CAPMAS report and none in the Observatory (out of 8). Goal 14 on life below water has three data points in the CAPMAS report and two in the Observatory (out of 10). Goal 15 on life on land has five data points in the CAPMAS report and one in the Observatory (out of 14).

D. Effectiveness of data collection mechanisms

This section shares some of the insights gained from attempts to obtain development data on Egypt. The authors also conducted interviews[20] to pinpoint whether there were structural problems or bottlenecks in generating official development data that impacted their potential to guide decision-making. Two groups of obstacles were identified.

1. Problems at the source

There seems to be a lack of what could be called "trained data officers" in various ministries and government departments, particularly at the local level where disaggregated data are supposed to be collected. While the Government has recently embarked on an ambitious project to create digital transformation units in all government offices, it is still unclear how much training or authority the responsible officers will have. Although CAPMAS and the Ministry of Planning and Economic Development organized workshops in 2019 to train data officers in collecting and sending SDG data, the high turnover rate in these positions sometimes hampers the execution of this process. For example, in some cases, CAPMAS had to resend the methodologies agreed upon to the individual ministries because the officers who had been trained had moved to other departments. Consequently, there is a need for continuous rather than one-time capacity-building training for these officers,[21] as well as measures to decrease turnover. For example, officers who receive training could be required to remain in their positions for one or two years. The National Strategy for the Development of Statistics, supported by the World Bank, does include a component on capacity-building, and its effective implementation is crucial.

The digital transformation initiative began in 2019, so it might still be too early to evaluate its progress. Nevertheless, based on interviews conducted by the authors, some institutions are still applying a slow, low-tech process for gathering and sharing data. One such example is sending a PDF, which either has to be retyped as a Microsoft Word document or filled out manually, both of which cripple data collection and the application of relatability checks. Another example involves saving the forms on compact disks (CDs), which must be sent back and forth by courier, and many new computers no longer have CD drives. Although this observation certainly does not apply to all government departments (CAPMAS, for example, conducted the latest census using electronic rather than paper forms), such institutional variation slows down the overall process. Lastly, some ministries also tend to wait to collect as much data as possible before sending their SDG data reports to CAPMAS, which means that data collected earlier become outdated while other data are being completed. Real-time sharing of data therefore is an option that could be utilized more often.

Stronger networking among CAPMAS, the Ministry of Planning and Economic Development and other ministries could also speed up the data collection process and help to identify bottlenecks more quickly. In many cases, CAPMAS cannot directly contact data collection officers in the ministries because requests must be sent to the minister's office. Again, such bureaucratic hurdles slow down the process.

2. The lack of a legal framework for data availability

Egypt has made various attempts to produce its own Freedom of Information Act. According to interviews with individuals who have been involved in such attempts over the years, there have been at least five draft laws. Some were

drawn up exclusively by the Government and some exclusively by civil society institutions, which sometimes included the private sector. There was also one joint draft put forward by the Government and civil society. Although attempts began as early as 2008, no such law has been passed and no bill appears to be in the pipeline. Based on interview data, the authors posit that the two main obstacles are a lack of ownership and concerns about the misuse of data. Both must be addressed to allow Egypt to begin producing better data at a faster pace in order to guide development decisions.

With regard to the lack of ownership, previous drafts of the bill have been introduced by different ministries. For example, the Ministry of Justice introduced the bill to address legal procedures, and the Ministry of Communications and Information Technology introduced it to address technical issues surrounding the publication of data. Other ministries have been on board at different times as well. While their participation is essential, given the bill's complexity, no single ministry has taken ownership of the process in order to push it forward. Admittedly, this is a common problem when regulating new and multifaceted areas. Nevertheless, the issue must be solved for the bill to be taken up again. Assigning responsibility to one institution would be a first good step.

Members of Parliament have repeatedly expressed concerns about the misuse of data to fuel the production of "fake news", which also explains the lack of legislative enthusiasm for the bill. While this is a reasonable concern, it must be emphasized that a lack of data would not solve the problem of misuse of information. Misinformation and disinformation have become a fact of life in a world where billions of inhabitants are connected to social media every day.

Two further dimensions stress the need to produce such a law. First, it remains an unfulfilled constitutional requirement. According to article 68 of the Constitution, on access to information and official documents: "Information, data, statistics and official documents are owned by the people. Disclosure thereof from various sources is a right guaranteed by the State to all citizens. The State shall provide and make them available to citizens with transparency. The law shall organize rules for obtaining such, rules of availability and confidentiality, rules for depositing and preserving such, and lodging complaints against refusals to grant access thereto".[22] Although the Constitution does not set a time frame for such a law to be promulgated, its value to development in Egypt underlines its urgency. Second, Egypt has achieved several successes in fighting corruption, and each periodic anti-corruption government strategy underscores the need for information to be free and publicly available.

E. How data could guide development decisions

This section explores the potential for data to guide development decisions with an exercise that demonstrates that even incomplete data could inform decisions concerning where to direct resources to achieve maximum impact. It uses principal component analysis for each Goal to determine the relative weight of its individual indicators[23] in driving change. This can be used as a first step to guide policy decisions when choosing among several courses of action.

The exercise began by examining SDG data for Egypt from 2000 to 2018.[24] Data from the UNSD Global SDG Indicators Database were used, as there are insufficient historical data points in official national sources.[25] Indicators missing over 60 per cent of the data points were excluded. Missing

values for indicators with less than 40 per cent of the data points missing were imputed.[26] Redundant indicators were dropped (Spearman's rank-order correlation > 0.9), and then the principal component analysis with rescaling was conducted.[27] Results are shown in table 4. Detailed information on the indicators used, the principal components and the eigenvalues (i.e. loadings) are presented in annex 3 and are available upon request.

With respect to Goal 1 (no poverty), policies lifting people from below the poverty line appear to be the most important, particularly with a focus on access to basic services among the poor (e.g. education, health care, water and sanitation, etc.). The Egyptian Government recently announced a large-scale programme to spend approximately $32 billion on the poorest Egyptian villages. Given the results of this exercise, the data appear to justify this programme.

With regard to Goals 2 and 3 (zero hunger and good health and well-being), a particular focus on children, in terms of nutrition and mortality rates, is most capable of making progress. As for Goal 8 (decent work and economic growth), providing aid for exports and reducing unemployment should receive particular attention. The Egyptian Government has announced an ambitious goal to increase exports from $26.1 billion to $100 billion. Concrete policies in this regard are likely to provide a significant boost to growth.

Table 4. Most influential indicators per Sustainable Development Goal

SDG	The two most influential indicators in improving the relevant Goal	Number of indicators included
Goal 1: No poverty	- Proportion of population living below the national poverty line (percentage) - Proportion of population with access to basic drinking water services (percentage)	6
Goal 2: Zero hunger	- Prevalence of malnutrition among children (wasting and overweight) - Agriculture orientation index for government expenditures	7
Goal 3: Good health and well-being	- Infant mortality rate - Neonatal mortality rate	10
Goal 4: Quality education	- Gender parity index for participation rate in organized learning - Total official flows for scholarships	4
Goal 5: Gender equality	N/A[a]	2
Goal 8: Decent work and economic growth	- Total official flows (disbursement) for Aid for Trade - Unemployment rate - Number of commercial bank branches per 100,000 adults	7

Source: Authors' calculations.
[a] Not enough data to perform analysis.

F. Conclusion and policy recommendations

This chapter set out to examine the extent to which national data on SDGs are available, accessible and usable in Egypt. The analysis points to a need for SDG data to be produced more quickly and with more detailed disaggregation in order to better identify both geographical and social gaps in development. A mechanism for integrating existing and new sources of SDG data, which are currently scattered across different repositories, is essential to building an inclusive national database. Moreover, some data bottlenecks could be addressed to improve the efficiency of the Government's data generating process. These mainly pertain to training data officers and improving networking among the various data gathering institutions.

There are two additional arguments in favour of data-informed development decisions. First, the coming years will almost certainly be even more data-driven, as both individuals and Governments are doing more online than ever before. The COVID-19 pandemic has only accelerated this trend, which provides an opportunity for the Egyptian Government to collect more data in real time. This trend also creates the potential for big data analysis, particularly as the Egyptian Government is launching a number of e-portals to provide specific services. While some government projects already use big data, mainly to detect fraud and suspected corruption, there do not appear to be any initiatives to monitor development, assess the impact of development projects or compensate for the lack of data on many SDG indicators. The use of big data would first require investing in data specialists, which would be a good starting point to unlock this potential.

Second, in the age of misinformation, continuous publication of data is the best way to counter "fake data". In today's world, individuals are and will continue to be hungry for data every day, if not every hour. Whether governments approve or not, data about government performance and indicators will continue to pour from unauthorized sources. Sustainable development indicators are prime candidates for potential fake data. Indicators on poverty, the spread of diseases, pollution and social spending are but a few examples. Consequently, disclosing data about such indicators in a timely, transparent and accessible manner is becoming essential for a country in which a majority of the population obtains information from social media outlets. By investing in a culture of trust in government data, whether they are positive or negative, the space for fake data shrinks. If accurate official data are easily accessible to all (the public, researchers, decision makers, media and intelligentsia), users would develop a collective resistance to fake data.

In light of the analysis presented in the chapter, the authors would like to propose the following policy recommendations in the area of strengthening data and data systems:

> Stronger networking among CAPMAS, the Ministry of Planning and Economic Development and other ministries could also speed up the data collection process and help to identify bottlenecks more quickly.

1. Training – and decreasing the turnover of – data officers at the different line ministries while achieving greater networking between them and CAPMAS. The purpose of that training would be to increase human capacities in managing the data collection and data streaming processes. Such training could be supported by several UN bodies.

2. Ensuring greater integration of SDGs indicators into the periodic surveys published by CAPMAS. These surveys have the advantage of being tailored for disaggregated data and are conducted frequently. Linking them to the SGD framework would ensure that more SDG-related data are generated regularly. (short term)

3. Creating some binding benchmarks regarding the updating frequency and level of disaggregation of SDG data. These benchmarks should then be met by government agencies. Obviously, before the issuing of such benchmarks, the relevant resources (human and technological) have to be made available to enable meeting the benchmarks. (medium term)

4 Unifying data the depository for SDG data (possibly at CAPMAS), to make sure SDG data are integrated in one national and all-inclusive source that can be used more effectively to guide policies. (medium term)

5 Develop initiatives to use Big Data to monitor development, assess the impact of developmental projects or compensate for the unavailability of data on many SDG indicators. Certainly, Big Data is an area that first requires investing in data specialists. This would be a good starting point to unlock such potential. (short term)

6 Disclose data about SDG indicators, in a timely, transparent, and accessible manner and invest in a culture that supports trust in government data – whether positive or negative – the space for 'fake data' shrinks. This should be associated with adopting legislative reforms that contribute to strengthened governance, particularly laws on freedom of information. (short term)

The Egyptian Government announced a large-scale programme

$32 BILLION

The poorest Egyptian villages

The Egyptian Government has announced an ambitious goal to increase exports

$26.1 BILLION → **$100** BILLION

How data could guide development decisions

- No poverty
- Zero hunger
- Good health and well-being
- Quality education
- Gender equality
- Decent work and economic growth

Endnotes

1. Additional data on SDG indicators for Egypt are available on international data portals managed by the United Nations and its specialized agencies (e.g. ILO, UNCTAD, FAO, etc.). However, this data assessment section is only concerned with national Egyptian sources that are used to provide national policy recommendations.
2. These include public finance and national accounts, macroeconomic indicators, domestic and foreign investments, domestic and foreign trade, income and standards of living, poverty, subsidies, manufacturing, markets and prices, the labour force, transport, mining, information and communications technology, utilities and services, health care, security and safety, education and research, entrepreneurship, agriculture, water resources and irrigation, food security, demographics, youth, women, childhood and maternity, sports, social insurance, the environment, and press and media.
3. OECD, 2017.
4. Pipino and others, 2002.
5. For example, researchers or academics who wish to acquire raw data must submit a formal request for approval and pay a fee corresponding to the amount of data requested.
6. On the individual or micro level, CAPMAS conducts national surveys to collect data on population demographics, health care, education and access to services and technology, among others. On the macro level, CAPMAS collaborates with other public institutions such as the CBE, the Ministry of Finance and the Ministry of Planning and Economic Development to produce macroeconomic data (i.e. data on public finance, national accounts, trade, and domestic and foreign investments).
7. The authors differentiate between longitudinal and panel data in that longitudinal data arise from collecting observations from different units (i.e. individuals) over time (i.e. equivalent to time series), while panel data arise from repeated observations of the same unit (i.e. individuals) over time.
8. The Economist, 2021.
9. While the most recent biannual bulletin available on the CAPMAS website (accessed in December 2020) was dated December 2020, the second most recent issue dated back to 2017, leaving a significant time lag. Moreover, the latest annual bulletin available on the CAPMAS website was published in October 2020 but contained statistics on marriage and divorce from 2019.
10. While writing this chapter, the authors were informed that there were plans to upgrade the CAPMAS website substantially. This would certainly be a step in the right direction.
11. For more details, see the 2020 CAPMAS Quality Assurance Framework for Statistical Data.
12. Attempts were made to preserve the distinction between data (raw numbers and/or information) and statistics (data compiled and treated for later use) wherever possible when describing the information provided by each institution. Nevertheless, it should be noted that sometimes the lines of demarcation are not very straightforward.
13. For example, when accessed on 21 August 2021, the latest financial report available on the website was for May 2021.
14. More information on the mapping methodology can be found on the Ministry's website: https://www.moic.gov.eg/en-US/Sectors/Index?na=110.
15. UNSD, 2020.
16. Egypt published its third VNR in July 2021, highlighting challenges caused by the COVID-19 pandemic. Annex 2 has been updated to include the indicators available in that report.
17. Annual refinements of indicators are included in the global indicator framework as they occur.
18. For a list of available indicators and their sources, see annexes 1 and 2 in the VNRs.
19. Annex 2 includes details on the available indicators for the three national sources. For indicators with available data, the annex provides the year of the latest figure available and the levels of disaggregation (if they exist). Note that annex 1 of the 2018 VNR of Egypt provides the aggregate availability of data according to CAPMAS in 2018, while annex 2 reports the actual data points for some of the indicators, which are fewer than those reported in annex 1. The authors included indicators available for the 2021 VNR as well.
20. The authors conducted interviews with a senior official at CAPMAS responsible for managing SDG-related data that flow in from various government ministries. An interview was also conducted with a former advisor at the Ministry of Planning and Economic Development.
21. For example, Egypt may benefit from support from UNCTAD to improve the generation of statistics on trade in services. For more information, see the project "Strengthening statistics on internationally trade-in-services" for countries of the West African Economic and Monetary Union, as well as the project "Train for trade".
22. Constitute, 2021, p. 24.
23. Lafortune and others, 2018. Principal component analysis is a technique for dimensionality reduction and feature selection. The analysis is conducted using SDG data to elicit the most important indicators (i.e. assign them weights) to determine which are the main drivers of performance. Given that there are 17 Goals comprising 247 indicators, it is reasonable to assume that not all indicators have the same importance or significance in driving progress. Principal component analysis can help decision makers to set priorities in targeting demanding issues that hinder progress.
24. The authors acknowledge that indicators have changed during this period and include only consistent indicators and proxies.
25. Data downloaded from the Global SDG Indicators Database on 13 January 2021.
26. Imputation was done using the regularized iterative principal component analysis algorithm.
27. Eigenvectors (i.e. the main components that carry change and in which the most variance or information lies) are calculated and sorted to determine the principal components. The first principal component holds the most variance, then the second, then the third, etc. After determining the eigenvectors, eigenvalues (i.e. loadings) attached to each eigenvector are listed to give the amount of variance carried in each principal component. These loadings determine the weight or importance of each indicator in the principal component. In order to determine the most influential indicators that govern change for each Goal, loadings of the first principal component are investigated. Indicators with the highest loadings (i.e. eigenvalues) in the first component are therefore the most influential in determining the change in performance for the respective Goal. Knowing the most influential indicators can help decision makers to pinpoint the main indicators that drive change in SDG performance and act accordingly.