Data Production for SDG indicators in Ghana

Prepared by: Ghana Statistical Service

March 2017

Ghana Statistical Service
Head Office,
P. O. Box GP 1098,
Accra, Ghana.
Fax: +233-302-664304
E-mail: info@statsghana.gov.gh
Website: http://www.statsghana.gov.gh

Background

Ghana has been actively involved in the process to agree the Sustainable Development Goals (SDGs) since consultations began in 2012. As the successors to the Millennium Development Goals (MDGs), the SDGs, and the underpinning Agenda 2030, sought from the beginning to be inclusive and participatory. Ghana was one of the initial 20 nations selected for national consultations on the theme, "*The World We Want",* for the post-2015 development agenda and played numerous other roles at global level throughout the process leading to the adoption of the SDGs.

The SDGs, which came into effect on 1 January 2016, will shape global sustainable development policy through to 2030. In addition to adopting the 17 goals and 169 targets in March 2016, the UN Statistical Commission agreed on an indicator framework to track SDGs at the global and national level. In order to provide the information that will be required by the indicator framework, robust data production and tracking systems will need to be built in every country so that achievements at the national and local levels can be assessed and fed into the wider global framework. These national mechanisms will need to be robust, effective and integrated, and will help countries to optimize their resources and actions to achieve the goals.

In Ghana, an SDG Implementation Coordination Committee has been established to provide technical support to the implementation and monitoring of the agenda. This multi-stakeholder committee with cross-government representation as well as members from Civil Society Organisations (CSOs), the private sector, and academia, is chaired by the National Development Planning Commission (NDPC). Ghana Statistical Service (GSS) sits on the Committee and acts as the data champion for the SDGs at the national level.

Around the world, National Statistical Systems (NSSs) are being challenged to produce more and better quality data than ever before. The SDG indicators not only require data on areas of enquiry that NSSs have not traditionally produced, they also demand new levels of disaggregation to ensure that we *Leave No One Behind.* What is becoming clear moreover, is that the traditional NSS cannot produce all of the data required alone and new actors such as CSOs and private sector organisations must now be engaged with and thought of as part of a wider data ecosystem.

Since early 2016, GSS has been spearheading a number of activities to determine data availability for indicator production, uncover gaps in the data system, and identify current and future activities that could produce data for monitoring the SDGs. This has included consultations with a wide range of stakeholders as well as conducting audits on its own data production. In late 2016, GSS also published its second National Strategy for the Development of Statistics which was largely aligned with the SDGs and included those government agencies that will be responsible for producing the majority of SDG indicator data. These actions confirm GSS' commitment to embracing new and innovative methods of data production, analysis and dissemination as well as strengthening traditional data collection methods.

In November 2016, the SDG Implementation Committee in collaboration with UNICEF and UNDP held a forum to discuss data production and use for national development. Following from this forum, the GSS in collaboration with SDG Implementation Coordination Committee and supported by the Global

Partnership for Sustainable Development Data (GPSDD) and UNDP with advice from the UN Data Group, is planning a National Data Roadmap Forum in April 2017 focused specifically on SDG indicator data production and use. The Data Roadmap Process which this Forum will begin has 3 main aims:

1. Data producers will be clear about their responsibilities with regard to data production for the SDGs in the short, medium and long-term. The objective is also to engage with non-state actors in this regard as part of a broader ecosystem for data.
2. Data users will be facilitated in finding information and will have fora where they can also interact with the data producer community.
3. Ghana's National Statistical System (NSS) will have a clearer picture of resource availability for its activities and ways to manage coordination.

The Roadmap Forum will bring together key national data producers and users as well as international experts supported by the GPSDD. It will result in a National Data Roadmap, the implementation of which will be overseen by a multi-stakeholder advisory committee that has been established for that purpose and to oversee the organization of the forum itself.

The purpose of this document

The purpose of this document is to present the current situation in Ghana with regard to data production for SDG indicators as understood by Ghana Statistical Service based on a series of consultations with stakeholders, and to invite comments and feedback from stakeholders.

It is further intended to provide the basis for discussion at the upcoming National Data Roadmap Forum and to inspire future work.

The scope of the document

This document presents the current situation of data production for SDG indicators in Ghana, assesses key data gaps and presents some initial ideas about opportunities and approaches to filling data gaps. The data gaps and possible methods of addressing these presented here are by no means exhaustive and stakeholders are encouraged to provide their comments (see final section).

Current state of data production for the SDGs in Ghana

At present, the indicator framework for the SDGs is still being refined at the global level. In its most recent iteration (December 2016), there were 230 unique indicators agreed upon. These indicators are divided into 3 tiers. Tier 1 indicators are those which are conceptually clear, for which there is internationally agreed metadata (method of producing the data), and that countries around the world regularly produce. Tier 2 indicators are conceptually clear, have internationally agreed metadata but countries do not regularly produce. Tier 3 are those for which there is no internationally agreed metadata and which countries do not regularly produce, though some may be producing these indicators according to their own methodology.

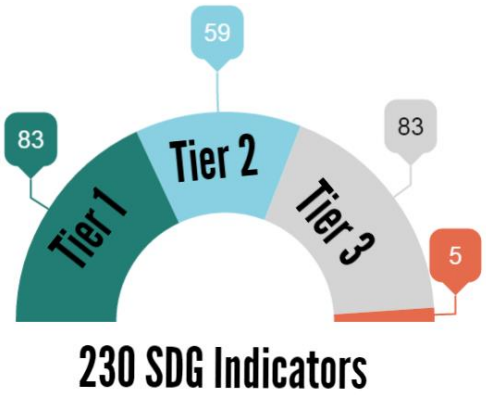


Figure 1: Breakdown of unique SDG indicators by Tier

Of these 230 Indicators, 83 have been designated Tier 1, 59 are Tier 2, 83 are Tier 3, and 5 are composite indicators, the components of which have multiple tiers (figure 1).

Some of the indicators are repeated more than once across the goals bringing the total to 241 with repetitions included. The breakdown across the goals (repetitions included) is as shown in Figure 2 below.

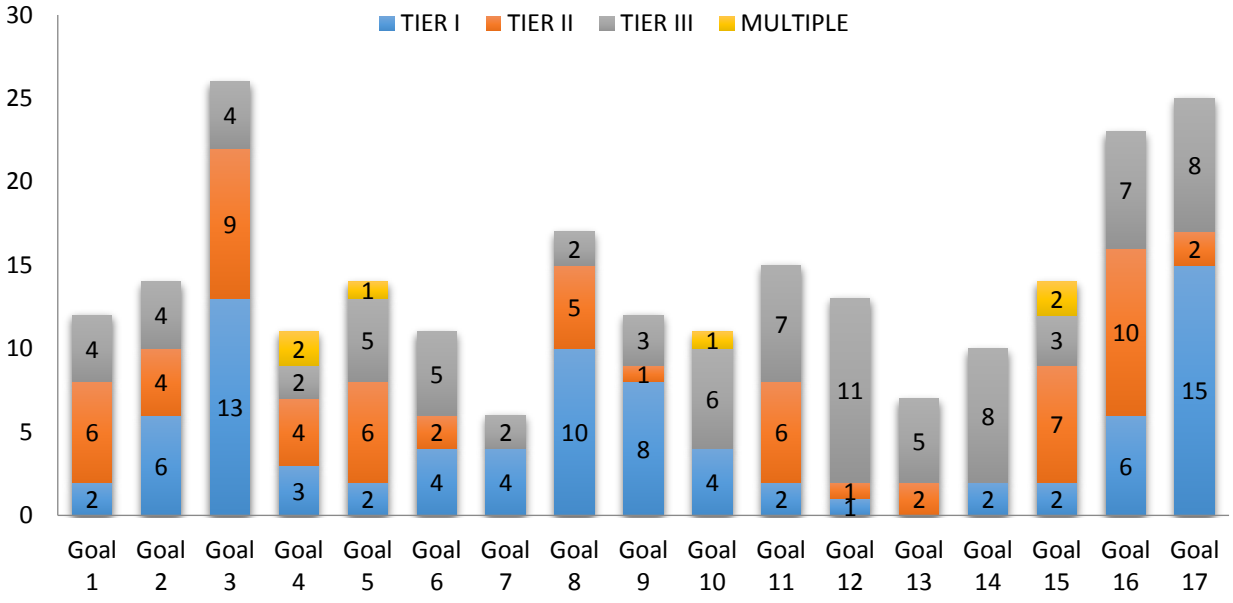
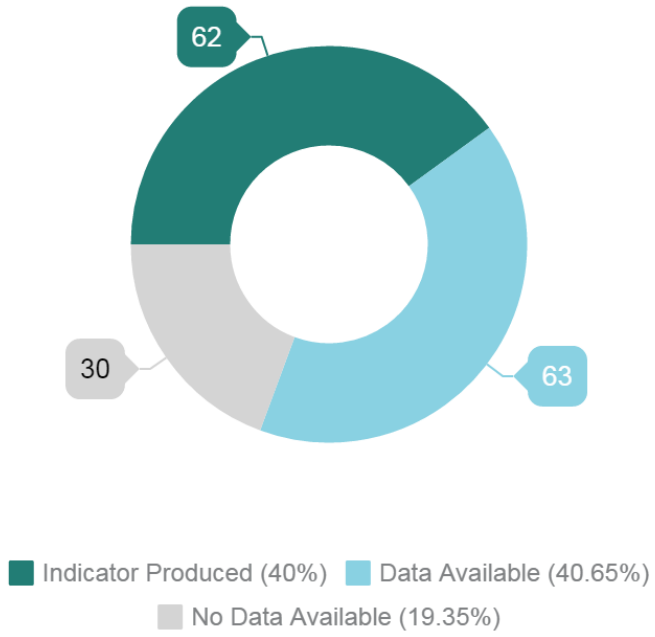


Figure 2: SDG indicator tiers by goal (repeated indicators included)

Ghana Statistical Service
Head Office,
P. O. Box GP 1098,
Accra, Ghana.
Fax: +233-302-664304
E-mail: info@statsghana.gov.gh
Website: http://www.statsghana.gov.gh

Since the beginning of 2016, Ghana Statistical Service has led a process to identify the current state of data production for SDGs in the country. Of the 155 Tier 1, 2 & Multiple Tier indicators (including repetitions), Ghana currently produces 62 (Figure 3). Across the National Statistical System, there is some data available that with improvement could be used to produce a further 63 indicators, though it should be noted that this data comes with limitations, discussed further in the following sections. There are therefore, 30 Tier 1, 2 & Multiple Tier indicators for which Ghana currently holds no data.

Figure 3: SDG Indicator production/ data availability for Ghana



Broken down across the tiers, Ghana produces or holds data for 72 Tier 1 indicators, 48 tier 2 indicators, and 5 Multiple Tier indicators (including repetitions) (figure 4).

Of the Tier 1, 2, and Multiple Tier indicators that are already produced or for which there is data available, 33% come from censuses and surveys, 57% come from administrative data sources and 10% come from a combination of census/ survey and administrative data (Figure 5).

Figure 5: Indicator Production/ Data Availability across tiers in Ghana

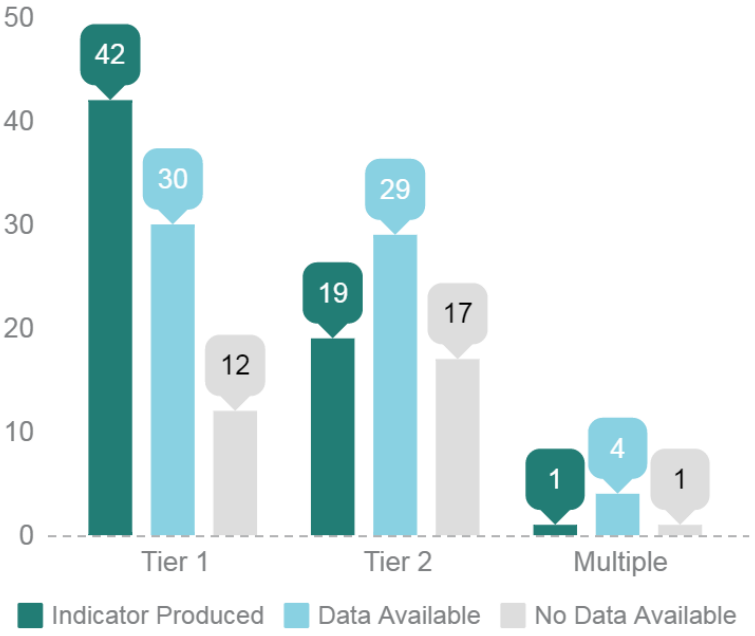
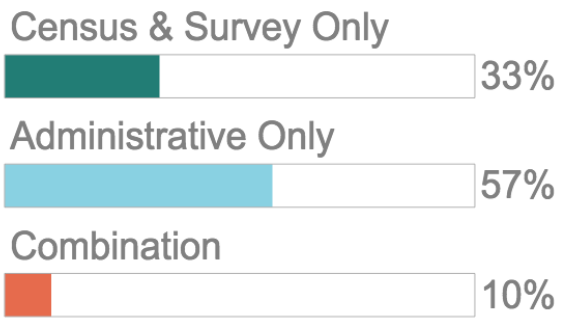


Figure 4: Source of data



Ghana Statistical Service
Head Office,
P. O. Box GP 1098,
Accra, Ghana.
Fax: +233-302-664304
E-mail: info@statsghana.gov.gh
Website: http://www.statsghana.gov.gh

Key Data Gaps

From the above analysis, it is clear that Ghana either produces or has some data for just over half of the SDG indicators. None of the Tier 3 indicators have been considered in this analysis and at present the National Statistical System has not prioritized the production of these indicators, which may be problematic for those goals which have a high proportion of tier 3 indicators (see Figure 2). Of the 155 Tier 1, 2, and multiple tier indicators (repetitions included), these can be broken down to identify key thematic data gaps (see Figure 6) and cross cutting data issues to be addressed for Ghana. The following discussion relates only to Tier 1, 2 and Multiple Tier indicators.

Coverage of thematic areas

Across the 17 Goals, there are a number of thematic areas on which Ghana performs better than others with regard to data production. Where the SDGs have adopted the legacy indicators of the MDGs for example, the Ghana NSS regularly produces data to report on these. However, for new areas, the capacity to produce data to meet reporting requirements is currently low.

Goal 1 (End poverty in all its forms everywhere) and **Goal 2 (End hunger, achieve food security and improved nutrition and promote sustainable agriculture)** are the bedrock for achieving the SDGs. Of the 17 indicators for these goals, Ghana currently produces 4 indicators, data is available for a further 10 and there is no data at all for 3 indicators. However, much of the data currently being utilized or identified as a possible source is drawn from surveys or would require strengthened and collaborative administrative data systems in order to produce the indicator on a regular basis.

Indicator 1.3.1 for example, *Proportion of population covered by social protection floors/systems, by sex, distinguishing children, unemployed persons, older persons, persons with disabilities, pregnant women, newborns, work-injury victims and the poor and the vulnerable,* would require coordination of inputs from institutions such as the Ministry of Gender, Children and Social Protection, Ghana Education Service, Ghana Health Service, National Health Insurance Authority, and the Ministry of Labour and Employment. Ensuring that these inputs can be utilized together, i.e. are interoperable, will be key to achieving success (see next section).

Goal 3 Ensure healthy lives and promote wellbeing for all, at all ages carries the largest number of indicators of any of the SDGs. Ghana currently produces 11 of the 22 indicators, with data available for a further 3 and no data for 8 indicators. Many of the indicators currently produced are from census or survey data and so can only give estimates intermittently. For example, indicator 3.7.1 *Proportion of women of reproductive age (aged 15-49 years) who have their need for family planning satisfied with modern methods,* is produced from Ghana Demographic and Health Survey or the Multiple Indicator Cluster Survey which are conducted every 5 years each. If SDG reporting is to be done on an annual basis this indicator will stay at the same level in the intervening years or between GDHS/ MICS rounds. To overcome this issue, Ghana will need to rely on estimates or find a way to produce this indicator on a more regular basis.

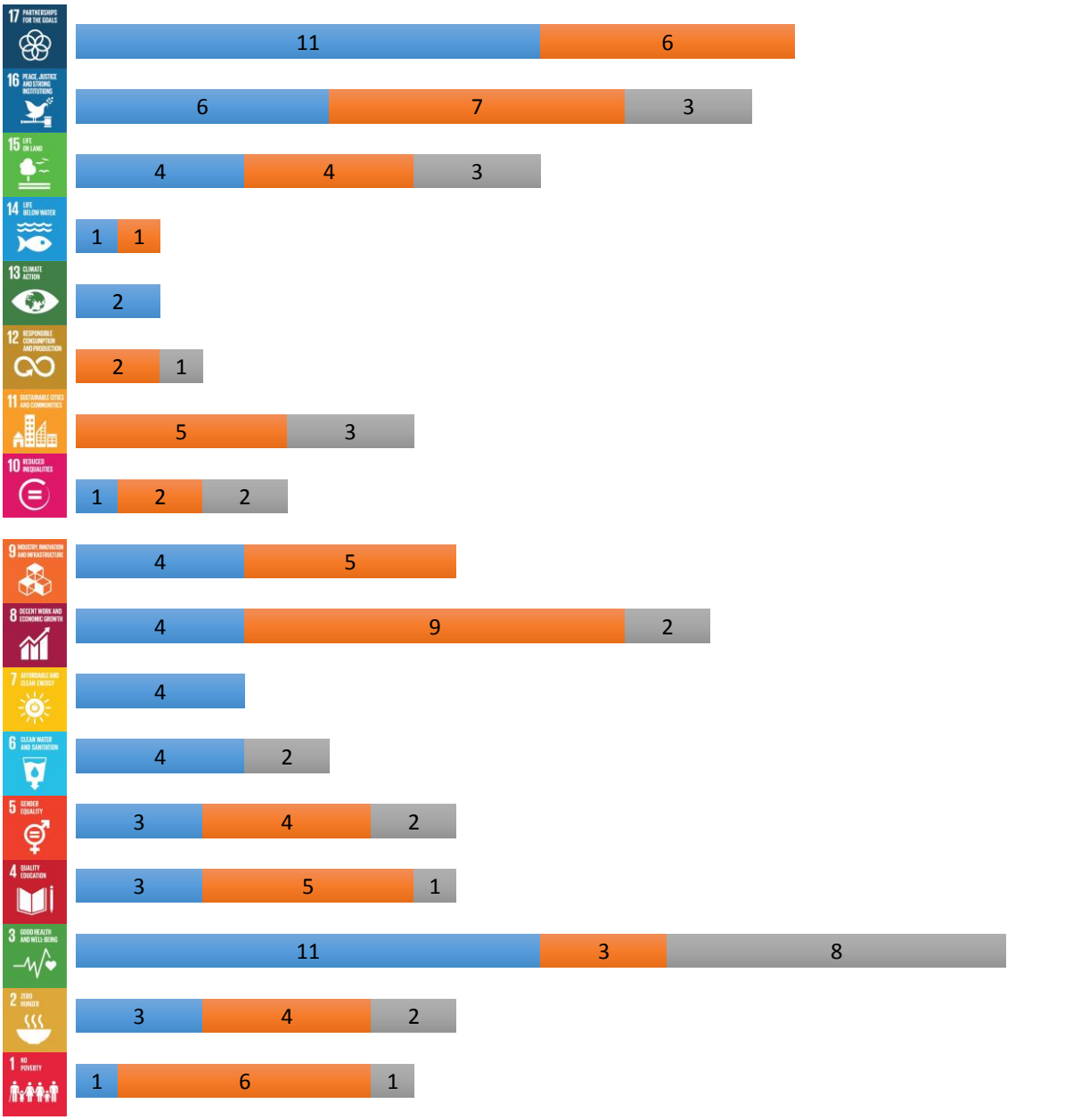


Figure 6: Current state of data production in Ghana for Tier 1 & 2 & Multiple Tier indicators

Ghana Statistical Service
Head Office,
P. O. Box GP 1098,
Accra, Ghana.
Fax: +233-302-664304
E-mail: info@statsghana.gov.gh
Website: http://www.statsghana.gov.gh

Gaps are apparent for other goals such as education and gender for which less than half of the indicators are currently produced. New areas such as decent work (Goal 8), industry and innovation (Goal 9), and reducing inequality (Goal 10) have more data available than is currently being used and mechanisms will need to be put in place to transform this data to regular indicator production. For others such as sustainable cities (Goal 11) and responsible consumption (Goal 12), Ghana currently does not produce any of the SDG indicators (Tier I&II). Goals 13-15 will require substantial amounts of new and improved environmental data, while for Goal 16 (Peace, Justice and Governance), Ghana produces 6 of the 15 indicators. Goal 17 relating to partnerships for SDGs, has 17 indicators associated with it, of which Ghana currently produces 11.

Apart from key thematic data gaps, there are a number of issues that must be addressed across all areas in order to produce the type of quality, nuanced, and timely data required by the SDGs.

The cross-cutting issue of data disaggregation

Underpinned by the commitment to *Leave No One Behind,* the Inter Agency Expert Group on the Sustainable Development Goals has stated that "indicators should be disaggregated, where relevant, by income, sex, age, race, ethnicity, migratory status, disability, and geographic location, or other characteristics, in accordance with the Fundamental Principles of Statistics".

Data disaggregation is critical because aggregate level data tend to hide lower level information. For example, national poverty estimates can mask the differing levels of poverty between different parts of the country or population sub groups. Similarly, the lack of gender disaggregated data means that it is impossible to measure the different experience of men and women in society. Without disaggregated data then, it is not possible to build effective policies and interventions that will have the positive outcomes intended for all of the groups they impact.

While disaggregated data is extremely important however, it can also prove difficult to produce. Many countries, including Ghana, rely heavily on census and survey data and have relatively weak administrative data systems. Administrative data is that which is collected primarily by government departments for the purpose of registration or record keeping usually during the delivery of a service, for example healthcare or education services. Well-designed administrative data systems can provide very detailed data on socio-economic characteristics and at a faster pace than censuses and surveys which are typically conducted at 5 or 10 year intervals. In Ghana, the wider administrative data system has received some attention in recent years (see next section for more details).

Censuses and surveys themselves have a number of issues that mean that they cannot provide all of the data required for SDG indicators, including: cost, coverage, and timeliness. Firstly, censuses and surveys are very expensive to conduct. They require large numbers of field staff, as well as resources such as vehicles and more recently electronic tablets for data collection. Secondly, coverage while very good during the census, is less inclusive for household surveys. For a typical household survey conducted in Ghana, households will be selected to ensure that the sample is representative at national and regional levels. This means that there must be enough households selected from each region to ensure that accurate regional level estimates can be achieved. GSS does not currently conduct household surveys that

are representative at the district or community level as the cost to implement the sample size required would be prohibitive. Household surveys often exclude certain sections of the population, such as institutionalized individuals and homeless households. The third issue is the speed with which the data can be generated. The process of preparing for a survey, data collection, data cleaning, data analysis and report writing can sometimes take years which means that by the time the data is released many things would have changed and the policies derived therefrom may not yield the desired results. In addition, if SDG reporting is to be done on an annual basis and surveys are conducted 5 years apart then many indicators estimates will remain stagnant until the next iteration of the survey.

While censuses and surveys present challenges, they will remain central to SDG indicator data production and act as sources of validation for other data sources. New sources of data and technology have begun to be used to provide levels of disaggregation elsewhere and may warrant further investigation and investment.

Data disaggregated by location

In recent years, Ghana Statistical Service has made significant improvements in its ability to capture and produce data with location markers. Census enumeration area maps are digitized using a Geographic Information System (GIS), which when combined with other data can give a substantial amount of information about the situation of people in different locations. For example, using a technique called Small Area Estimation, GSS was able to combine the 2010 Population and Housing Census Data with GLSS 6, including the use of location data taken from both the census and the survey to produce poverty maps for all 216 districts in the country. This has been very helpful for policy makers at both national and district levels to direct resources and tackle the issues in their areas.

Given that MMDAs are the local authority responsible for ensuring policies and programmes impact communities within their jurisdiction, statistical information at this level is extremely useful to help District Assemblies make informed decisions. In the recent past however, censuses (population and industrial) are the only time that GSS produced data at this level. Given that censuses are conducted infrequently, this means that district level data is only updated once a decade. As was discussed earlier, household surveys that are conducted more frequently cannot produce district-level estimates because the sample sizes are not large enough to provide such estimates.

Location data will be central to achieving the SDGs and many advances have been made in this area with new technology. This is discussed further in the next section.

Data disaggregated by sex and gender sensitive data

Gender data are vital to ensure that we can capture the different experiences and realities of women and men, girls and boys. During the last national forum on data and in the first National Strategy for the Development of Statistics, there was a clear call for Ghana to produce both sex disaggregated data and gender sensitive data.

Sex disaggregated data is that which merely produces separate estimates for men and women. This is extremely important data and must be prioritized throughout the data production process across each of the SDGs. This means that when we produce data for health, education, relating to the environment or

the economy along with all other thematic areas, we must produce data that can provide separate counts for men and women. While censuses and surveys are generally quite good at producing sex-disaggregated data, the administrative data system will need to be strengthened in this area.

Gender-sensitive data production seeks to go further than simply disaggregating by sex and tries to uncover or reflect underlying gender relations. For example, gender sensitive data might focus on areas that impact men and women differently such as domestic violence or labour force participation. For example, in nearly every country, women spend more time than men on unpaid labour. Household surveys often struggle to capture this and other traditional data collection methods can also fail to capture women's economic contribution because they are often employed in the informal sector. This often leads to a distorted view of the roles that men and women actually play in society.

Thus, while there is a standalone goal relating to gender, Goal 5: "Achieve gender equality and empower all women and girls", this is not the only time that we need to think about gender. We must think about each of the indicators we produce from a gender perspective. Moreover, while many of the Goal 5 indicators are Tier III, this does not mean that no gender data can be produced. Data 2X, a global organisation responsible for improving gender data, has identified 16 "Ready to Measure" indicators across the SDGs that most countries will be able to collect.

Data disaggregated by other population sub-groups

Data must also be disaggregated across some population sub-groups such as income group, by age, ethnicity, migratory status, or disability status. The difficulty with producing estimates for these groups is similar to that for producing estimates at district-level, one of sample size and methodology. In addition however, there is an issue with coverage of surveys and the groups that are often left out of the sampling frame are the homeless population, those in prisons or institutional care, those who may not be legally documented, or are nomadic and without a fixed location.

Similar to the discussion above, census data can only produce data once a decade and would likely not cover topics specifically geared towards these sub-populations. While there have been efforts made to give estimates for persons with disability, for example in GLSS 7, the sample size is not sufficient enough to provide regional level estimates. New ways will need to be explored to produce data for these population sub-groups.

The challenge of producing timely, relevant, quality, and interoperable data

As can be seen from the forgoing discussion, one of the key challenges apart from disaggregating data is producing timely and relevant estimates. Because censuses and surveys can only be conducted years apart, the data we draw from them remains stagnant until the next iteration of the survey. For the SDGs this presents a significant challenge as progress will need to be tracked annually. Thus, for many of the indicators which we have indicated above that Ghana currently produces, where these are from survey/ censuses, these are only produced periodically. For example, for indicator 1.2.1, *Proportion of population living below the national poverty line, by sex and age* which is produced using GLSS data is available at

intervals of every 3-5 years and therefore progress on this indicator can only be measured each time the survey is conducted.

For other indicators that are produced from administrative data, the rate of production of data is much faster and cheaper. However the administrative data that is currently produced cannot always be used to produce statistics or indicators due to inherent issues with the data collected. Also, statistical capacity with the Ministries, Departments and Agencies that collect these data has been low although there has been ongoing efforts to strengthen capacity.

Data produced must also be relevant to data users. SDG goals, targets and indicators must be adapted to and prioritized according to Ghana's context. However, given that context evolves with time, the optimum way to ensure that data remains relevant is for producers and users to engage in an ongoing dialogue in which users advise and producers have the flexibility to adapt. The relative lack of Open Data in Ghana has also meant that data users cannot easily access the data available, and many data producers do not know which data is held by other producers leading to duplication and sometimes conflicting estimates.

Ensuring that data produced is of a high enough quality to produce accurate estimates for SDG indicators is also crucial. Good quality data requires attention at each stage of the production process, from question phrasing, to enumerator selection and training, to data cleaning and analysis. In addition, data must be harmonized across the National Statistical System so that actors are working from agreed metadata to avoid conflicting estimates. Further to harmonizing current data production, making data interoperable would also help not only to monitor SDG progress but also help implementing agencies to achieve the SDGs. For example, ensuring that classifications used for education data at Ministry of Education matches exactly the classifications used by Ministry of Finance in its budget statements would be very powerful in allowing for the tracking of spending. Open Data would also be helpful here, as it would ensure that data is made available in formats that can easily be integrated. For example, files available in ".csv" format can easily be integrated, while files in PDF, excel, Word, or other formats require a large amount of effort to integrate.

Opportunities and approaches to filling data gaps

While there are many challenges that the data system must face in order to produce the data required, the SDGs also mark a significant opportunity for National Statistical Systems (NSSs) around the world. From the beginning of the process to agree Agenda 2030, there has been focus on the need to make them data driven. This means that there is now more attention than ever before being paid to the production of data which must be harnessed to build the robust and sustainable data ecosystems of the future.

Moreover, since the agreement of the MDGs, the world has undergone a "data revolution". Major technological advances have meant that there are now more methods of collecting and producing data as well as new types of data than ever before.

To capitalize on the high-level goodwill towards data and take advantage of the data revolution, Ghana must focus its data production efforts on 3 main areas to begin filling its data gaps.

Strengthening and rationalising Censuses and Surveys

Censuses and surveys will remain central to data production over the lifetime of the SDGs. Ghana is fortunate to have a history of regularly conducting household surveys, the largest being Ghana Living Standards Survey, Ghana Demographic and Health Survey, and the Multiple Indicator Cluster Survey. It also has a strong history of conducting the census at regular 10-year intervals and is due to conduct its next in 2020.

Improving coordination and harmonization across these data collection activities will ensure that data for certain indicators can be produced at regular intervals. For example, indicator *1.2.2 Proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions,* requires information on non-monetary aspects of poverty. If a standard set of questions on this can be agreed and implemented across all surveys, the same indicator could be produced every time one of these surveys is conducted. In addition, timing the surveys will be helpful to spread data production to illustrate progress over time.

There could also be further exploitation of methods such as Small Area Estimation which combines censuses and surveys to allow for district level data production. As was discussed earlier, this was used to produce poverty maps by combining the 2010 Population and Housing Census and GLSS 6 data.

Building effective administrative data system

As is clear from the above sections, the administrative data system is going to be key to producing SDG indicator data. Not only is it far cheaper to collect than survey data, it is also more timely and frequent and can produce far more detailed data than is possible from censuses and surveys.

Work has been ongoing in this area within the National Statistical System in Ghana. For example, a new Civil Registration and Vital Statistics Strategy began to be implemented in 2016. This 5 year plan aims to ensure that systems are in place to adequately record major life events, such as birth, marriage, divorce and death. In addition, GSS has been working with a number of line ministries to review the data collection templates (in the coming weeks) in line with the SDG indicators to ensure MDAs produce the required data for indicator computation. Further cooperation and harmonization will be required between GSS and MDAs to ensure that the required administrative data is produced.

To enable the administrative data system to flourish, there will be the need for investment into administrative data infrastructure. This has already begun in certain thematic areas such as education and health which have their own Management Information Systems (MIS). For the NSS to function effectively however, these MIS will need to be interoperable and use technology that is future proof. A suggestion for how this might be done was outlined in the National Strategy for the Development of Statistics using a cloud-based system.

Data production and use at MMDA level will also need to be strengthened and adequately resourced. As part of a recent restructuring in Ghana Statistical Service, some statistical officers were posted to district

assemblies. In this respect it has been programmed for the establishment of district statistics department in all MMDAs to facilitate the collection and collation, analysis and reporting local level statistics for planning, monitoring and evaluation.

Exploring new sources and types of data

New sources of and types of data have the potential to play a key role in the production of disaggregated data. Used in conjunction with traditional data sources for validation, new types of data can add different dimensions of information that traditional data sources does not currently produce.

Those that have shown the most promise in pilot projects elsewhere and in Ghana in producing SDG indicator data are:

- Geospatial (in particular, satellite data): Geospatial data, GIS data or geodata has explicit geographic positioning information included within it, such as a road network from a GIS, or a geo-referenced satellite image. Geospatial data has been shown to be useful for the production of environmental indicators and can be produced and analysed in real-time.

- Big Data: Big Data is an umbrella term referring to the large amounts of digital data continually generated by the global population. A large share of this output is "data exhaust," or records generated as a by-product of everyday interactions with digital products or services. Mobile phone data, for example, can tell us about migratory patterns, spending patterns, and where there are gender markers, can even help us to produce gender data.

  The benefits of Big Data for SDG indicator production are:

  - Big Data is continuously produced and if real time analysis is adopted can offer more timely estimates that censuses/ surveys
  - Big Data is passively produced which means that there is no direct cost associated with the production of the data
  - Big Data can add information about new topics and levels of disaggregation.

Conclusion

The SDGs and their accompanying indicators present both an opportunity and a challenge for Ghana's data ecosystem. With increased international and national attention given to the importance of data, data producers and users have the opportunity to harness the momentum to make lasting improvements but must first acknowledge the myriad challenges that must be overcome.

Ghana currently produces a quarter of SDG indicators and holds some data for another quarter (for indicators classified as Tier 1 & 2 & Multiple). Apart from key thematic data gaps, most notably for those new themes which the traditional NSS does not currently have the capacity to produce, there are a

number of cross-cutting issues that must be addressed. Data disaggregation, alongside producing timely, relevant, quality and interoperable data are issues that must be considered across the SDGs.

Strengthening censuses and surveys, administrative data systems, and exploring the use of new data would all help towards producing more and better indicators. What is also clear however, is that many of the indicators require cross-governmental and multi-stakeholder participation. Further discussion of how to build new data communities, mechanisms, and systems is therefore encouraged at the upcoming Data Roadmap Forum to identify innovative ways of working to enable SDG data production and use.

Call for comments

This short paper is intended to continue the conversation with stakeholders about data for the SDGs. It has outlined Ghana's current status in relation to data production, highlighted some of the key data gaps, and offered some initial ideas on how to address these, from the viewpoint of Ghana Statistical Service.

This paper aims to be a basis for further discussion and we would welcome interaction with all stakeholders.

If you have comments or feedback, please submit them to: Omar Seidu, SDG Co-ordinator at GSS, omar.seidu@statsghana.gov.gh or Eleanor Carey, Eleanor.carey@statsghana.gov.gh.

Ghana Statistical Service
Head Office,
P. O. Box GP 1098,
Accra, Ghana.
Fax: +233-302-664304
E-mail: info@statsghana.gov.gh
Website: http://www.statsghana.gov.gh