African Conference:
Transformative Agenda for Official Statistics

Co-organized by the African Development Bank (AfDB), Statistics Canada and the United Nations Statistics Division (UNSD)

With the support of the African Union (AU), the African Centre for Statistics (ACS) and the Statistical Office of the European Union (Eurostat)

Libreville, Gabon

November 2015

Session 2:
Innovation and Modernization through a Standard-based Statistical Business Architecture

Mr. António dos Reis Duarte, President of the National Statistical Institute of Cape Verde

(…) High-quality statistical information and data are essential for the proper planning and measurement of development outcomes. Africa should generate its own data to enable it to better monitor and track economic and social targets, including the goals and objectives of Agenda 2063. A data revolution in Africa would afford our continent the opportunity to interact with diverse data communities and to embrace a wide range of data sources, tools and innovative technologies, which would enable the continent to produce disaggregated data, including gender-disaggregated data, for decision-making, service delivery and citizen engagement[1].

Reliable statistics are essential for countries to progress. Besides being reliable, the challenge also lies with the appropriateness of the data; in other words, data must be available in a timely manner so that it should be possible to formulate more effective public policies or adjust them. We all know the cost of policy making based on incorrect or non-existent data. These costs are compounded in many African countries by the lack of financial and skilled human resources when confronted with the challenges to statistical systems. This excerpt from PARIS21 clearly elucidates the issue:

However, it should be borne in mind that the production of statistical data is not an end in itself, but a means for defining policies and measuring development outcomes. Hence, the need for accurate, reliable, and accessible statistical data that allowing it to assume its assigned role: to contribute to decision-making and to be one of the scientific pillars thereof.

The number of users of statistical data is growing and the quest for greater data diversity and details are an enormous constraint for national statistical systems, which must, consequently, be prepared to meet these new challenges. The Millennium Development Goals were a challenge for statistical systems to track the 61 indicators. The current proposal of the Sustainable Development Goals (SDGs), whose indicators are still in the design phase, currently include a total of 223 indicators, in various themes. The transformation Agenda includes topics such as education and health but also includes new themes such as governance. As concerns the African continent, Agenda 2063 represents another challenge in terms of monitoring even though the alignment process[2] between the

---

[1] Ministerial Declaration of the Eighth Joint Annual Meetings of the African Union Specialized Technical Committee on Finance, Monetary Affairs, Economic Planning and Integration and the Economic Commission for Africa Conference of African Ministers of Finance, Planning and Economic Development, Addis Ababa 30 and 31 March 2015.

[1] http://unstats.un.org/sdgs/meetings/iaeg-sdgs-meeting-02

[1] From 7 to 12 de Setembro decorreu o *First Technical Working Group Meeting on Developing a Measurement Framework within the Context of SHASA for Agenda 2063 for the First Ten-Year Implementation*, organizado por Statistics South Africa, e a Comissão da União Africana.

[2] De 7 a 12 de Setembro decorreu o *First Technical Working Group Meeting on Developing a Measurement*

Global Agenda and the regional Agenda is expected to allow for greater consistency in policies, but also, an articulated request for information. However, the SDGs and the 2063 Agenda must be seen as an opportunity for improving statistical systems.

Modernising the entire system is one of the possible strategies for adjusting the Statistics System to new times and challenges. This modernization must take place for issues related to instruments, data processing and should also include the data collection and dissemination process and at the same time be more flexible. Analysis of the institutional environment is also needed for the establishment of procedural rules and standards, given the massificacao of use of non-traditional data collection sources such as *Big Data* collection.

National Statistics System Law as Prerequisite for the Modernisation of Statistics

The basis for modernising official statistics hinges on legislation allowing for a national statistical system which is functional and open to change. As we know, statistical systems include national institutes of statistics, central banks, and sometimes, ministries or other types of institutions. To have an integrated system, dialogue between the various parts of the system is fundamental and should be institutionalized. Communication mechanisms should be established and be known by the parties. This is possible, only by enacting a clear and objective law that establishes the duties and responsibilities of the national statistics system as well as of all the component parts. This National Statistics System should:

- Foster a thematically encompassing statistical production. The plurality of integral institutions in the system allows for a greater variety of field of action;

- Promote the facilitated use of administrative registers. The legislation should include a requirement to provide administrative data for statistical use. However, just supplying data is not enough, because the law must provide for the validation of data models to populate IT platforms.

- Facilitate coordination of the system. To do this, it is necessary to set up an entity above the data-producing entities, which fosters dialogue between partners, producers, and also defines priority statistical operations. This entity must be independent and not answerable to Government authorities. The Law should provide

Framework within the Context of SHASA for Agenda 2063 for the First Ten-Year Implementation, organizado por Statistics South Africa, e a Comissão da União Africana.

for the existence of a National Statistical Council with the capacity of coordinating and regulating the production of official statistics.

- Allow for autonomy. The law must guarantee financial autonomy (adequately resourcing the national statistical systems to implement the action plans which are assessed by the National Statistics Commission responsible for setting national priorities in terms of production of official statistics), property autonomy and technical autonomy. The latter is fundamental because it gives to systems the freedom to implement methodologies according to good practices in terms of technical recommendations, without political interference**.**

Big Data

As already mentioned, national laws relating to national statistics systems are fundamental, and are one of the key issues concerning Big Data. National institutes of statistics should abide by the fundamental principles of official statistics, adopted by the United Nations Statistical Commission in 1994 and subsequently adopted by the United Nations General Assembly in 2014, and in particular the following principles:

- Principle 6. Individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes.

- Principle 8. Coordination among statistical agencies within countries is essential to achieve consistency and efficiency in the statistical system.

- Principle 9. The use by statistical agencies in each country of international concepts, classifications and methods promotes the consistency and efficiency of statistical systems at all official levels.

Thus, it is first of all necessary that Big Data producers be included in the national statistical systems, but this implies compliance with the approved legislation on statistics, that is, statistical coordination as well as other producing entities should be respected. Secondly, statistics should comply with the concepts, international classifications and methods, underpinned by the quality of information and comparability. Producers of statistics through Big Data will also have to comply with established standards and those that will be established. Thirdly, the confidentiality of information is a basic and

fundamental principle that must be respected by any producer. Thus, the statistical data producer must have rights and obligations, in the same way as other producers.

There is, however, an important factor to consider. Big Data refers to management, processing and disclosure of information from alternative sources. Information is not solely statistical. Simply counting *cases* is not statistics. There are statistical and mathematical rules to be respected. For starters, in the analysis, producers of information (which must not be included in national statistics systems) should be distinguished from the producers of statistics (which should be part of the system).

Innovation in Administrative Records

Statistics derived from administrative records are extremely important for countries since their production is fast and requires fewer resources; however, their structure and compilation are complex.

African countries are at different stages in terms of use of administrative records and within the same country, various statistics may be at different stages. This diversity requires an initial inventory of these statistics, firstly as concerns the collection mechanisms which can be diverse (transmission of databases to the national statistical system according to the schedule established, partial access to the database, full access to database). Concerning, for example, data of entry and exit from the country, Cape Verde's National Institute of Statistics has partial access to the database of the Directorate of Foreign Nationals and Borders-National Police, relating to the entry and exit of each passenger, except the name and passport number. Access has been granted allowing the INE a monthly download of the database.

There are a set of preconditions for the use of administrative records. At the national level, it should be recognized that information banks are not necessarily databases for statistical use. With the development of computerization, it is urgent that information lend itself to statistical processing. Using the example of birth records, if an institution decides to digitize birth records and keep them as images, their use for statistical purposes is almost zero. If on the other hand, it decides to create a data entry model, where information is entered, it is possible to use them, but nevertheless, it may not be sufficient. To make it possible to maximize the use of these databases, it is necessary that the latter be databases as well and that information processing should be facilitated; data input models should be

discussed with the national statistical system. Thus, the SEN Law should create the obligation of dialogue or validation of the data entry models of all SEN stakeholders.

Nevertheless, the constraint of information sharing between institutions persists. For instance, for vital statistics and civil status registration, a common platform is needed between health facilities and civil status registry; in fact, national institutes of statistics should be considered part of the communication between these two primary producers of information.

The various outlines imply that regional and continental harmonization of statistics is a delicate and slow process. For countries where the systems are already established, change can be difficult because of the classifications used; databases and transmission mechanisms are already defined and a change can imply a disruption of the current system. In the case of Cape Verde, the adoption of a new computerised customs system (ASYCUDA World) has led to the temporary suspension of production of external trade statistics. These experiences are useful for systems which have not yet been set up.

And regional and continental harmonization must necessarily consider national diagnostic systems in such a manner as to conduct comparative analysis between countries in order to identify the most efficient mechanisms for each type of statistic.

Innovation in Statistics Collection Operations

The Continent is banking on mobile technology-assisted collection; some countries have used it in Censuses (Cape Verde - 2010, Senegal - 2014, Cote d'Ivoire - 2014), and many other countries in surveys.

Mobile mechanisms have been used in two areas: data collection for surveys and/or censuses and digital mapping.

For data collection, the Personal Digital Assistant (PDA) was used, but with recent innovations, a transition is being effected using tablets to make up for the flaws of PDA, e.g. small screen and poor luminosity making it difficult to fill out the questionnaires. The biggest advantage lies in the fact that the tablet is equipped with a more powerful processor reducing the transition time between questions and modules such as the consistency checking process.

Regardless of the choice of mobile devices, there are common advantages: consistency check; real-time collection monitoring, elimination of the entry stage; this precludes the procurement and transportation of large amounts of paper; consistency check; autoskips; obligation to respond to nuclear matters; elimination of storage costs; time savings in the processing of results. In addition, there are financial gains, as the equipment can be used in several other statistical operations. In terms of analysis, focused analyses can be made, in reduced spaces, thanks to information georeferencing. It is possible to perform geostatistical analysis by associating geographic information with alphanumeric information. The challenge of transmitting information via the Internet should be noted; in fact, coverage in our Continent is gradually improving but there are still areas that are not covered. Concerning this specific constraint, online and offline software have been created. In many cases, collection is done in offline mode and work is done online, with the sole intention of setting up geographical housing contact information and thus be able to gather information, locate housing or synchronize information. Currently, NIS are more and more inclined to use these devices.

Digital mapping is the second potential for mobile devices and new technologies. Currently, it is a question of geo-referencing information through the assignment of geographic codes to each building. This enables the mapping of information gathered through questionnaires, but other uses are possible, for example, in agriculture. Second Carlos Lopes "*The success of large-scale agricultural investments on the continent will largely depend on the extent to which key land governance and policy reforms are undertaken by African countries. At the core of these reforms is an inclusive land tenure system. It would not only buttress the land rights of local communities, women and youth but also increase economic development, promote gender equality and inspire younger women and men to take up agricultural activities. An inclusive land tenure system would also improve efficiency of land use and climate change mitigation and contribute to Africa's economic transformation."*

To monitor progress, especially in the relationship between the characteristics of the population such as sex, age and agricultural production, comparison of of two elements of information is necessary, collection through questionnaires relating to families, with collection in the field, in relation to agricultural lands. It is possible to geo-reference farmland. The general agriculture census is ongoing in Cape Verde.

Again, the software can be adapted to a more effective response to information needs and policy priorities.

The advantages are obvious, but there are constraints. The procurement of PDA or tablets is an investment with a high initial cost, mainly if it is intended for censuses. It was observed that from 1995-2004, 18 countries on the African Continent did not conduct a census; from 2005 to 2014, 9 countries did not carry out a census and it is hoped that at the next round of censuses (2015-2024) the number of countries will decrease. In this set-up, the procurement of PDA or tablets for censuses on the Continent would not be possible. To do this, the likely alternative would be the procurement of equipment by a pan-African organization which would then hand them over or lease them to countries for the conduct of censuses and if necessary, for specific surveys. For the next round of censuses, the countries that have already indicated the date of their next census are 1 in 2015, 1 in 2016, 3 in 2017, 3 in 2018, 2 in 2019, 3 in 2020, 3 in 2021 and 1 in 2022. Thus, with proper management and programming, it will be possible support countries requiring such equipment.

A second factor, which is just as important, relates to technical qualification. IT consultation, especially in software development is expensive and knowledge is not reproducible.

In the information technology sector, progress is fast and just as quickly outdated. In the case of Cape Verde, technicians are trained in the use of Windows media software, but recently, these technicians needed training in Android technology in order to accompany the transition from PDA to the Tablet. Without a team that can be trained continuously in every innovation, it becomes necessary to call on expertise from new consultants; also recommended is the establishment of knowledge centres in software development, which would be the result of collaboration between international organizations and some national institutes of statistics. These centres would lend technical assistance to countries in need thereof.

Innovation in Dissemination

Dissemination needs to be effective; it is necessary that national statistical systems be closer to users in a dynamic, attractive and coherent manner.

The first challenge of the national institutes of statistics has been the creation of their own *websites*, and this challenge is now a reality. In 2015, the national institutes of statistics

of 50 African countries had their *website*[3]. However, modernization requires further progress in dissemination.

Dissemination software exist and are used by the national statistical systems. There is a Data Portal[4] that brings together data from 51 African countries[5], and also has national versions as is the case of Cape Verde, Nigeria and Rwanda. DevInfo[6] is another platform that facilitates the dissemination of socioeconomic statistics and has been the subject of improvements since 2002 and is currently in its 7th version. CensusInfo[7] is an IT byproduct which is a product of dissemination of Census indicators. In the dissemination of micro-data, there is the Microdata Cataloging Tool ((*NADA*)[8]). In the area of geostatistics, ArcGis[9] is an important dissemination tool. Cape Verde used ArcGIS for the dissemination of results of the 2010 General Population and Housing Census. In this regard, it was not limited only to dissemination. ArcGIS was used in all phases of the operation: pre-collection, collection, post-collection and dissemination[10].

These are only a few examples of existing statistical dissemination platforms, easy to develop and use. The real advantage of the variety of existing platforms is the opportunity given to national statistical systems to adopt those that best fit their needs. But the danger comes from the simultaneous use of identical platforms, causing duplication of effort with the maintenance and updating of each of the platforms, as well as data splintering in several platforms; although each platform, individually, aims to achieve ease of use for users, the simultaneous use of platforms is difficult.

However, although these platforms receive users, however they do not interact with them. To do this, social networks have an important role to play. To expand statistical literature, but also make statistical data available to users who do not use the national statistical system, social networks complement dissemination through the media. Social networks play a double role; on the one hand, highlighting the mechanism that underlies statistical

---

3 http://www.uneca.org/es-blog/governance-and-large-scale-agricultural-investments-africa
3 Fonte: Divisão de Estatística das Nações Unidas, dados actualizados a 18 de Setembro de 2015.
3 Fonte: Divisão de Estatística das Nações Unidas, dados actualizados a 28 de Setembro de 2015.
4 http://www.afdb.org/en/knowledge/statistics/data-portal/
5 South Africa, Algeria, Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Cape Verde, Chad, Comoros, Côte d'Ivoire, Djibouti, Egypt, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea-Bissau, Guinea -Conakry, Equatorial Guinea, Lesotho, Liberia, Libya, Madagascar, Malawi, Mali, Mauritania, Mauritius, Morocco, Mozambique, Namibia, Niger, Nigeria, Kenya, Democratic Republic of Congo, Central African Republic, Republic of Congo, Rwanda, Sao Tomé and Principe, Senegal, Seychelles, Sierra Leone, Somalia, Sudan, Swaziland, Tanzania, Togo, Tunisia, Uganda, Zambia, Zimbabwe.
6 http://www.unicef.org/statistics/index_24300.html
7 http://www.censusinfo.net/
8 http://www.ihsn.org/home/software/nada
9 http://www.esri.com/industries/federal/statistics
10 https://www.esri.com/industries/www.esri.com/~/media/Files/Pdfs/industries/federal/pdf/G69563_CapeVerdeCenusFlier_147252_V5_REVISED_PRINT1

production (visits, training, partnerships, among others) and, on the other hand, disclosing statistical data themselves. Worthy of mention here is the case of Tunisia whose Facebook page has a variety of information but which requires constant updating, representing a constraint for the NIS (the last update of the page dates back to 27 April 2015). INECV has its platforms, updated website and Facebook page, but it faces the same problem concerning the Mobile INE (mobile version of the website); this is a human resources constraint.

Thus, investment in new technologies for statistics must be supported through an assessment of needs and technologies available for dissemination, and must take into account the human resources available for the implementation, maintenance and update of statistical information.

New Coordination and Cooperation Models

The African continent has several regional, organizations including the Arab Maghreb Union (AMU), the Economic Community of West African States (ECOWAS), the Central African Economic and Monetary Community (CEMAC), the Southern African Development Community (SADC), Common Market for Eastern and Southern Africa (COMESA) and the Intergovernmental Authority on Development (IGAD). Each regional body strives to meet the specificities of the region, and this influences statistical production that tends to be harmonized within the region and between different regions. The establishment a unified standard is needed for the continent, hence the Strategy for the Harmonization of Statistics in Africa (SHaSA). The strategy is structured in thrusts and works well with significant benefits; however, SHaSA would be more effective if it were part of an institute with a continental dimension and context. This African Union (AU) project needs an African Institute of Statistics, a technically strong institution with a capacity to coordinate the continent's statistical production in terms of nomenclature, Metadata, and further technical assistance capacity to countries, and an adequate knowledge of all national statistical systems in such a way as to facilitate South-South cooperation. To this end, the diagnosis of statistical production must be one of the core performance thrusts of the African Institute of Statistics, or of a similar body.

Conclusion

Modernization and innovation in statistics must be seen as an integrated process between national institutions, without neglecting the needs at regional, continental or global levels.

An investment is needed in human and financial resources and should be analysed and carefully planned so that innovation does not suffer regress or stagnation.

It is necessary to consider setbacks that will occur and which are unavoidable. However, through experience-sharing, South-South cooperation, it is possible to leverage the experience of the most advanced national statistical systems and which have already gained extensive experience in managing the constraints related to ICT use and have probably already found innovative solutions.

One last point, of utmost importance, concerns the creation of knowledge centres in the development of statistics software and programmes intended for administrative records and surveys. In addition to these centres, the creation of a technically strong continental Institute close to national statistical systems is fundamental.