United Nations
Economic Commission for Africa

Manual for Statistical Development Indicators

United Nations
Economic Commission for Africa

Manual for Statistical Development Indicators

The designations used and the presentation of material in this publication do not imply the expression of any opinion whatsoever on the part of the United Nations Economic Commission for Africa (ECA) concerning the legal status of any country, territory, city or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries, or its economic system or degree of development. Designations such as "developed", "industrialized" and "developing" are intended for statistical convenience and do not necessarily express a judgment about the stage reached by a particular country or area in the development process.

Cover design, layout and printing: ECA Printing and Publishing Unit. ISO 14001-2015 certified.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Economic Affairs Officer from the Macroeconomic Policy Division of ECA, George Kararach. Gratitude is also expressed both to the senior staff assistant of the Statistical Development, Data Innovation and Outreach Section, Roman Legesse, for her immense contribution to the layout of the documents, for providing logistical and administrative support, and for organizing the two expert group meetings from the inception of this project; and to the intern with the African Centre for Statistics intern, Tissie Nadzanja, for her excellent taking of the minutes and report writing during the expert group meetings, and for her assistance in compiling the results for the volunteer pilot countries. Thanks also go to the ECA Publications and Conference Management Section for their invaluable work on this document. For those whose names were not mentioned, please kindly accept our apologies.

Oliver Chinganya

Director, African Centre for Statistics
Economic Commission for Africa of the United Nations
United Nations
Addis Ababa, Ethiopia
Tel.: +251 11 544 3595
Email: chinganya@un.org

Tinfissi Joseph Ilboudo

Chief, Statistical Development, Data Innovation and Outreach Section
African Centre for Statistics
United Nations
Addis Ababa, Ethiopia
Tel: +251 115 445 533
Email: ilboudo@un.org

PREFACE

To meet the increasing demand for high-quality official statistics in Africa, the Economic Commission for Africa developed the African Statistical Development Index, called ASDI and revamped it into the "Statistical Development Indicators" (StatDI), which was endorsed by the fifth joint session of the Statistical Commission for Africa-Committee of Directors General of National Statistical Offices (StatCom-Africa/CoDG), held in Abidjan, Côte D'Ivoire from 28 November to 2 December 2016. ECA presented the StatDI methodology during the Marketplace event organized by the Partnership in Statistics for Development in the 21st Century (Paris21) Annual Meeting in April 2017.

The StatDI is intended to assist countries by: (i) addressing the main challenges that arise in the absence of reliable data and information to track and report achievements in national development, the 2030 Agenda for Sustainable Development and the African Union's Agenda 2063; and (ii) monitoring the progress in statistical development. StatDI is built on the following four fundamental dimensions which covers most aspects of national statistical system, namely, A Functional institutional and organizational framework; Capacity for an efficient statistical system; Production of relevant statistics that meet international quality standards; and Good dissemination policy and effective use of the statistics produced for analysis and research.

The StatDI has unique features as a composite index, and the methodology is based on comprehensive regional strategic and quality frameworks, namely, the African Charter on Statistics, the Strategy for the Harmonization of Statistics in Africa, and the Reference Regional Strategic Framework for Statistical Capacity-Building in Africa. It is aligned with the National Strategy for the Development of Statistics. The StatDI methodology adheres to the United Nations Fundamental Principles on Official Statistics.

The StatDI makes it possible for countries to identify strengths and weaknesses in their national statistical systems, thereby providing mechanisms for addressing them to facilitate timely policy intervention. StatDI will prospectively contribute to enhancing tracking and reporting mechanisms quality of the data and statistics as well as the national statistical system to better inform that national development plan objectives, the Sustainable Development Goals and Agenda 2063.

To ensure ownership by ECA member States of the StatDI, and in response to StatCom-Africa/CoDG recommendation, ECA has developed a manual and the accompanying online application for implementing StatDI. The StatdI uses information that will be captured through an online questionnaire on the four dimensions. To validate the questionnaire, preliminary pilot results from 11 volunteer countries were presented during an Expert Group Meeting that was held in Addis Ababa from 7 to 9 December 2017. The meeting recommended that the manual and accompanying questionnaire be subjected to further extensive internal review.

The manual is intended to operationalize the StatDI, guide and inform policymakers, national statistical offices, academic and research institutions, and other stakeholders on its use and benefits. The manual guides member State on how to effectively collect, compile and assess each indicator in each dimension and subdimension, respectively, while also assessing the overall performance of national statistical development mechanisms. StatDI's manual will provide users with tools for monitoring progress in their respective national statistical development measurements, as well as in harmonizing comparisons among countries. The

ultimate objective of this manual is to provide a practical approach based on StatDI to address key challenges facing statistical information research in the region.

Due to the unique feature of the StatDI, we recommend that statistical agencies in other regions use it in measuring progress in statistical development

I would like to thank the volunteer countries, the consultant and all experts who contributed to producing the manual and the online application, which contributed to the various processes of development of the StatDI.

Oliver Chinganya
Director
African Centre for Statistics
Economic Commission for Africa

ABBREVIATIONS AND ACRONYMS

ASDI   African Statistical Development Index

GDDS   General Data Dissemination System

ICT   Information and communications technology

IMF   International Monetary Fund

OECD   Organization for Economic Cooperation and Development

SDDS   Special Data Dissemination Standard

StatDI   Statistical Development Indicators

ECE   Economic Commission for Europe

CHAPTER 1

INTRODUCTION

A. Objectives

1. The main objective of this manual is to assist and guide users regarding the usage of Statistical Development Indicators (StatDI) in a national context. It provides information on the framework of StatDI and on the methodology to compile the index.

2. More specifically, the manual aims to help users adopt an effective and efficient approach to measuring progress in national, subregional and African statistical development within the context of the implementation of the Sustainable Development Goals. The manual will contribute to promoting ownership and enhancing the sustainability of the statistical development process in Africa. Consequently, it may provide a model to be emulated by other developing countries outside the African continent. StatDI can serve as a tool to enable countries to capture, measure and assess progress in all relevant dimensions of statistical development in a national context. Its objectives include monitoring progress achieved in statistical development and supporting the production of high-quality data to meet the challenge posed by missing, timely and reliable data that are required for tracking and reporting progress in the area of national development, the 2030 Agenda for Sustainable Development and Agenda 2063 of the African Union.

3. The purpose of the questionnaire (see annex 2) is to collect information on indicators for all dimensions and subdimensions used to calculate StatDI. Only one questionnaire per country should be completed with the results managed by the agency in charge of coordinating statistical activities in that country, such as that country's national statistical office, the ministry responsible for supervising the production of statistics, the national statistics council or any other similar organization, in collaboration with other relevant stakeholders.

4. This manual has been prepared for the purpose of preparing the data entry portion of the survey regarding StatDI. Chapter D discusses directives for the codification of questions and variables within the questionnaire, which is comprised of four modules.

5. To a large extent, the success of the process depends on the quality of the data collected. The accuracy and precision of the data collected will positively influence analyses conducted after data collection. Therefore, it is both necessary and critical for all questions in the survey questionnaire to be fully understood by those who are responsible for completing it. It is also essential that users are able to assimilate the content of the manual before completing the questionnaire and putting the methodology into practice.

B. Background

1. Context

6. During the 1990s, the African continent witnessed the launch of a number of initiatives aimed at improving the social, economic and political conditions of its citizens. Since the adoption of the Abuja Treaty on 3 June 1991, African leaders have remained committed to the establishment of an African common market in order to increase economic self-sufficiency and promote self-sustaining development on the continent. Initiatives in that regard include Poverty Reduction Strategy Papers, the New Partnership for Africa's Development, the 2030 Agenda for Sustainable Development and Agenda 2063 of the African Union.

7. The monitoring of policy implementation with a view to achieving sustainable development in Africa has considerably increased the demand for harmonized and reliable statistical data in all areas, including peace and security, governance, multilateral surveillance, and the monitoring and evaluation of relevant programmes. As a result, African national statistical systems and subregional and continental organizations focusing on statistics

and statistical development have faced challenges, but have also had an opportunity to raise public awareness regarding the importance of statistics and development on the continent. Those stakeholders have been able to harness resources to build the capacities of African countries to meet increased demand for high quality and comparable statistics resulting from efforts related to development.

2. Regulatory frameworks

8. To overcome the shortcomings of past efforts aimed at enhancing statistical capacity in Africa, national stakeholders, including the directors of national statistical offices and external stakeholders were consulted on the statistical framework. The framework was endorsed at a meeting of the directors of national statistical offices in Africa, held in Addis Ababa from 6 to 8 February 2006 and at the Forum on African Statistical Development, held in the same city on 9 and 10 February 2006.[1]

9. The Reference Regional Strategic Framework for Statistical Capacity-Building in Africa is expected to create synergies, prevent the duplication of efforts and lead to sustainable capacity for statistical development in Africa.

10. To ensure effective monitoring of the Reference Regional Strategic Framework, a tool was needed to assess the progress made by African countries in developing national statistical systems[2] that supported development initiatives.

11. During the third meeting of StatCom-Africa, held in Cape Town, South Africa from 18 to 23 January 2012, ECA presented a report on ongoing efforts to develop ASDI, a composite index aimed at supporting the monitoring and evaluation of the implementation of Strategic During the meeting, it was suggested that the methodology of ASDI be re-examined by a committee.

12. In response to that suggestion, the African Centre for Statistics came up with a new measuring tool referred to as the "African Statistical Development Indicators" (ASDIs). ASDIs are based on the Fundamental Principles of Official Statistics (see annex 1a). All African statistical system[3] organizations and all professionals in the field of statistics at the national, regional and continental levels respect the principles established in General Assembly resolution 68/261 on Fundamental Principles of Official Statistics,[4] which were endorsed by the Statistical Commission on 3 April 1994, as well as a commitment to best practices.

13. ASDIs were developed on the basis of the African Charter for Statistics (see annex 1c) and on Strategic Framework (see annex 1b), while remaining consistent with the strategic vision of the Strategy for the Harmonization of Statistics in Africa, second generation (see annex 1d), the National Strategy for the Development of Statistics and other relevant regional and international recommendations. As a result, ASDIs are global in scope and address most of the fundamental principles of statistical development. The report on ASDIs was endorsed during the fifth meeting of StatCom-Africa, held in Abidjan, Côte d'Ivoire from 28 November to 2 December 2016.

14. ECA presented ASDIs during an event organized by Partnership in Statistics for Development in the 21st Century, in Paris, in April 2017 at which it

---

1 For additional information see www.uneca.org/sites/default/files/PublicationFiles/stat_rrsfdocument_final.pdf.
2 The national statistical system is the group of statistical organizations and units within a country that jointly collect, process and disseminate official statistics on behalf of the national government. See glossary for further details.
3 As defined by the African Charter, the African statistics system is a partnership comprised of national statistical systems (data providers, producers and users, statistics research and training institutes and statistics coordination bodies), statistics units in regional economic communities, regional statistics organizations, regional training centres, statistics units of continental organizations and coordination bodies at the continent level.
4 Further information on the Fundamental Principles is available at https://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx.

was noted that ASDIs was global in scope and, therefore, the methodology used to determine the indicators it contained need not be restricted to Africa. Consequently, ECA chose to rename ASDIs and refer to them henceforth as StatDI, in order to allow for broader applicability of that measuring tool. Applicability means that other regions outside Africa may use StatDI to measure the quality of their proprietary statistical systems.

C. Presentation of Statistical Development Indicators

1. Description of Statistical Development Indicators

15. The Statistical Development Indicators (StatDI) is a new tool developed by the African Centre for Statistics. The tool is intended to measure, monitor and report the strengths of national statistical system components and identify areas that require further focus and enhancement. Through an integrated and harmonized approach, StatDI provides a framework for statistical development, allowing for enhanced collaboration and partnerships among relevant stakeholders.

16. This innovative new tool will provide an opportunity to make appropriate decisions, including with regard to policy interventions. Accordingly, StatDI will highlight statistical activities that need improvement, particularly those pertaining to reporting to the achievement of the 2030 Agenda and Agenda 2063 of the African Union. Consequently, countries will be better positioned to report performance regarding national development plans, the 2030 Agenda and Agenda 2063.

17. The StatDI framework is focused on four core dimensions, namely:

    (a) Dimension I: Functional institutional and organizational framework;

    (b) Dimension II: Capacity for an efficient statistical system;

    (c) Dimension III: Production of relevant statistics that meet international quality standards;

    (d) Dimension IV: Effective dissemination policy and use of statistics produced for analysis and research.

Each dimension constitutes a module within the StatDI questionnaire.

Functional institutional and organizational framework:

    (a) This dimension focuses on the organization and coordination of a country's national statistical system. The aim is to provide an overall perspective on the operational and legal organization of a system, the status of stakeholders and how African and international standards may be applied appropriately.

    (b) This dimension addresses variables pertaining to the organization and management of a national statistical system, as well as collaboration among stakeholders and between the national statistical system and external stakeholders, which may include regional, subregional and international organizations.

Capacity for an efficient statistical system:

    (a) This dimension traces the financial, human, technical, material and infrastructural resources to ensure the efficient functioning of the national statistical system of a country.

    (b) The aim is to identify the contributions of partners who support a properly functioning system.

Production of relevant statistics that meet international quality standards:

    (a) This dimension focuses on data production.

    (b) This dimension also focuses on needs assessments, statistical production

programming, the comprehensiveness, periodicity and timeliness of data collected through censuses and surveys, data collection methods, quality assurances with regard to generated data, archiving, storage and data security issues, as well as data management.

Effective dissemination policy and use of statistics produced for analysis and research:

(a) This dimension focuses on compliance with international standards for data dissemination and includes:

    (i) Raising awareness of the activities of national statistical offices and systems in order to strengthen the statistical culture within a country and to underscore the reliability of official statistics;

    (ii) The formulation of a clear data dissemination strategy in a format that is transparent, easily understandable, practical and appropriate.

18. StatDI helps underscore the importance of decision-maker advocacy and commitment to statistical development. The StatDI approach also supports statistics generation by enhancing data collection, data compilation, and statistical analysis processes. Several components of the third StatDI dimension (Production of relevant statistics that meet international quality standards) highlight relevant aspects of data collection, compilation and analysis standards, and the importance of improving those standards in order to enhance statistical development.

19. The four dimensions of StatDI include 21 subdimensions, which contain 147 variables or indicators. However, only 140 indicators are relevant to StatDI calculations, with the remaining main seven indicators provide supplemental information (see table 1).

20. StatDI will be compiled annually. The first stage of operationalizing StatDI will be to conduct a harmonized survey in order to collect information on the aforementioned statistical development dimensions. Indicators will be computed for each country and analysis conducted to assess the statistical performance of each surveyed country.

Table 1: Dimensions and subdimensions of Statistical Development Indicators

Dimension	Subdimensions	Number of variables (Indicators)	Number of variables (for scoring)
Module I: Functional institutional and organizational framework	A. Legislative and regulatory framework for statistics	3	3
	B. Framework for organizing and coordinating national statistics	5	5
Module II: Capacity for an efficient statistical system	A. Funding	6	5
	B. Good governance of the statistical office	9	9
	C. Human capital	10	8
	D. Physical and material infrastructure of the national statistical office	4	4
	E. Statistical infrastructure	7	7
	F. ICT infrastructure	14	13
	G. Promotion of bilateral and multilateral cooperation	6	4
Module III: Production of relevant statistics that meet international quality standards	A. Proper evaluation of data requirements	5	5
	B. Statistical programming	2	2
	C. Comprehensiveness, periodicity, and timeliness of data collected through censuses and surveys	23	23
	D. Data collection	8	8
	E. Development and implementation of data quality assurance tools	7	6
	F. Archival, storage and security of data	5	5
	G. Data management	8	8
	H. Analytical work within the national statistical office	3	3
Module IV: Effective dissemination policy and use of statistics produced for analysis and research	A. Effective dissemination policy and use of statistics produced for analysis and research	2	2
	B. Data dissemination policy, plan or strategy	12	12
	C. Raising awareness of the activities of national statistical offices and systems in order to strengthen the culture of statistical data dissemination	6	6
	D. Promotion of the use and analysis of generated data	2	2
Total		**147**	**140**

Table 2 summarizes how StatDI maintains the strengths of ASDI and addresses dimensions that are often missing in national statistical systems.

Table 2: The value-addition of Statistical Development Indicators over the African Statistical Development Index

African Statistical Development Index		Statistical Development Indicators
Component	**Subcomponent**	
National statistical system operation: organization and coordination of the statistical system	• National Strategy for the Development of Statistics • Advocacy • Legal framework • The position of the national statistical office within the national statistical system • Coordination mechanism • Monitoring of activities	Dimension 1: An institutional and organizational framework compatible with the subdimensions • The existence of a functional framework organizing and coordinating the national statistical system • The existence of a specific legislative and regulatory framework for statistics that is updated regularly
Condition of statistical infrastructure	• User needs assessment • Data development • Harnessing ICT • Data analysis	Dimension 3: Production of relevant statistics that meet international quality standards • Proper evaluation of data requirements • Statistical programming • Comprehensiveness, periodicity, and timeliness of data collected • Data collection and validation • Quality assurance of the data produced • Archival, storage and security of data • Data management • Analytical work within the national statistical office
Data dissemination	• Data dissemination	Dimension 4: Effective dissemination policy and use of statistics produced for analysis and research • Communication regarding the activities of national statistical offices and systems to strengthen a statistical culture within society • The existence of a data dissemination strategy • Promotion of research and a thorough analysis of statistical data
Human capital development	• Human capital development	This is a subcomponent of Dimension 2
Funding	• Funding	Dimension 2: Capacity for an efficient statistical system • Statistics funding • Good governance of the national statistical office • Human capital • Physical and material infrastructure of the national statistical office • Statistical infrastructure • ICT infrastructure • Active promotion of bilateral and multilateral cooperation for statistical capacity-building

StatDI is based on high-quality information and strategic frameworks, while ASDI is based primarily on the Reference Regional Strategic Framework for Statistical Capacity-Building in Africa. As compared with ASDI, the key distinguishing features of StatDI are noted in box 1 (below).

Box 1: **From the African Statistical Development Index to Statistical Development Indicators**

A comparison of the structures of StatDI and ASDI, reveals the following:
- The subcomponents of the ASDI second component, with the exception of information and communications technology (ICT), which StatDI addresses within its second dimension, are only some of the subcomponents contained in the corresponding third dimension of StatDI.
- ASDI treats data dissemination as a dimension while StatDI considers it as a subcomponent within its fourth dimension.
- The fourth component of ASDI is devoted to human capital development and the fifth component of ASDI focuses exclusively on funding while StatDI considers other components in addition to these two subcomponents in its second dimension.

---

2. Data collection using an application

21. An application was developed to facilitate the management of the StatDI questionnaire and accompanies the manual.

22. The features of the application are as follows:

  (a) **Management of country list:** Administrators manage the list of participating countries.

  (b) **Management of users:** Administrators can create new users, reset passwords and link users to countries in order to update survey responses. Users will be required to change their password after the first successful login.

  (c) **Survey questions:** The application converts an existing question to the format required by the application, and uploads data to the system.

  (d) **Baseline and target:** Users may set a baseline and a target at the variable level for a selected country. Initially, targets for variables will be set by the administrator, using benchmarks for the scoring methodology as defined in this manual. Users may subsequently establish their own targets for their national statistical system. The baseline will be derived from the first survey operation of the StatDI questionnaire.

  (i) Based on the baselines and the results of the subsequent survey, each country will be able to extrapolate outcomes in order to design strategic reforms to facilitate progress towards achieving national or regional targets for the relevant dimension.

  (e) **Questionnaire response:** Survey questions will be visible to users and progress will be shown to the user as the responses are saved. It may take multiple sessions for a user to complete all responses, and in such cases the system will begin with the last question answered in the previous session. The responses will be saved in draft mode by default. Once all responses have been entered and reviewed, the user will confirm them. Scores will be calculated only after all responses have been confirmed. The user will not be able to modify responses after confirmation.

  (f) **View score and report:** The user will be able to view the score for the country for which he or she is responsible. The user may see score changes for any dimension/subdimension over multiple surveys. The administrator will be able to view the score for any country.

(g) Collect data offline: Users will be able to collect and save data locally if an Internet connection is not available. The user will be able to upload data to the server when Internet is available. This function is a separate standalone application run on Microsoft Excel.

(h) Miscellaneous: Passwords will be saved in encrypted form.

(i) User Acceptance Testing: The application must be deployed on the ECA web server for testing. Server details and accompanying privileges are required and users will need to test the system and provide feedback.

3. Methodology for calculation of Statistical Development Indicators

23. The calculation of StatDI is performed in three steps:

(a) Data processing and editing (processing, imputation), as explained in annex 4

(b) Normalization

(c) Weighting and aggregation

Normalization of Statistical Development Indicators

24. As was the case with ASDI, the scoring for most of the 140 variables in StatDI, has a value within a range of 0.1, which guarantees normalization. However, for the purposes of computing the weights using principal components analysis and factor analysis, the following known standardized formula may be used:

$$S_k = \frac{X_k - \overline{X}_k}{\sigma_k}$$

.... (1.0)

In which case rather than working with the original variables X_k (k=1,2,3,…Q) we standardize them to S_k (k=1,2,3,…Q) where S_k is defined by the above standardized formula, in which

$\overline{X}_k$ is the sample mean, and for instance is the mean for *T* countries with regard to the *Kth* variable (indicator);

σ_k is the standard deviation so that the StatDI builder will use S_k (k=1,2,3,…Q) instead of X_k (k=1,2,3,…Q); and Q denotes the number of variables

Computational structure of Statistical Development Indicators

25. Within a national statistical system, let $X_{iik} \in (0,1)$ be the score of the *kth* indicator (variable) for the *ith* subdimension within the *jth* dimension, $i = (1,2,3,…,S; j = 1,2,3,…M; k = 1,2,…,N)$, under the assumption that each variable has the same weight when explaining statistical development progress. To compute the weights of sub-dimensions and dimensions of statistical development, we consider Matrix I below:

Matrix I: Distribution of indicators among subdimensions within the dimensions of a system of statistical development

Dimensions											
Subdimension	D_1					....	D_j				
D_{ij}	D_{11}	X_{111}	X_{112}	...	X_{11n11}	....					
	D_{21}	X_{211}	X_{212}	...	X_{21n21}	....					
	...	...	...	...	...	....					
	$D_{s(1)1}$	$X_{s(1)11}$	$X_{s(1)12}$	...	$X_{s(1)1ns(1)1}$	....					
	...	...	....	....	....	.....					
	D_{1j}						X_{1j1}	X_{1j2}	...	X_{1jn1j}	....
	D_{2j}						X_{2j1}	X_{2j2}	...	X_{2jn2j}	....
	...						...	...	...	...	....
	$D_{s(j)j}$						$X_{s(j)j1}$	$X_{s(j)j2}$	...	$X_{1s(j)jns(j)j}$	....
	...						...	....	....	....	.....

Note: Matrix I displays the distribution of N indicators among S subdimensions within M dimensions of a statistical development system.

26. From the above matrix each subdimension D_{ij} has n_{ij} indicators (scored variables), so that

$$\sum_i \sum_j n_{ij} = N \ldots\ldots\ldots.\ldots (1.1)$$

Computation of weights for the subdimensions

27. Dimension D_j has M_j indicators, so we define the weight W_{ij} of sub-dimension D_{ij} to be:

$$W_{ij} = n_{ij}/M_j \ldots\ldots\ldots..\ldots (1.2)$$

28. The above suggests that the more indicators accommodated within a given subdimension, the greater the weighting associated with that subdimension. In addition, those weightings are normalized, thereby contributing to overall uniformity.

Computation of weights for the dimensions and the aggregation scheme

29. In a similar manner, we define the weight W_j for each D_j dimension as follows:

$$W_j = M_j/N \ldots\ldots\ldots..\ldots (1.3)$$

This again suggests a normalized weight within the M dimensions

(a) Final score for the subdimensions

30. To aggregate the score for subdimension D_{ij}, assuming that each variable has the same importance, or equal weight, in explaining progress in statistical development, then a simple weighted arithmetic or geometric mean can be used. For instance, let us denote this score by S_{ij}, in the arithmetic expression:

$$S_{ij} = \frac{\sum X_{s(j)n_{ij}}}{n_{ij}} \ldots\ldots\ldots.\ldots (1.4)$$

(b) Final score for each dimension

31. Since the weights are normalized, the straightforward final score Sj for dimension Dj may be calculated using either weighted arithmetic or geometric mean.

For example, the arithmetic aggregation would be:

$$S_j = \sum_i W_{ij} S_{ij} \quad \ldots\ldots\ldots\ldots (1.5.a)$$

and the geometric aggregation would be:

$$S_j = \prod_. S_{ij}^{W_{ij}} \quad \ldots\ldots\ldots\ldots (1.5.b)$$

Where **Wij** denotes the weight of the **ith** subdimension within the **jth** dimension.

(c)   Aggregate score for all dimensions
32.   Similarly, we can aggregate the scores of all the dimensions using either weighted geometric or arithmetic mean. If we denote the final score for measuring the overall performance in statistical development for a given national statistical system as **S,** then using the weighted normalized arithmetic expression we can write:

$$S = \sum_j W_j S_j \quad \ldots\ldots\ldots\ldots (1.6)$$

Example 1
Statistical Development Indicators calculation for a hypothetical national statistical system

33.   Variables are measured according to different scales and some variables are reported on an ordinal scale ranging from a minimum of zero to a maximum of one. Other variables are reported as percentages, including, for example, the proportion of the national statistical office budget financed by the government, or in absolute numbers, including the number of computers per person in the national statistical office.

➢   Let us consider the scoring of country Z. Table 3 summarizes the calculation of scores and weights for each subdimension and corresponding dimensions.

Table 3.A: Summary of the calculation of scores and weights for each subdimension and corresponding dimension for hypothetical country Z

C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
From StatDI questionnaires				From (1.2)	From the survey results. Compiler uses (1.4)	From (1.5)	From (1.3)
Dimension	Sub dimension	n_{ij}	M_j	$W_{ij}=n_{ij}/M_j$	S_{ij}	S_i	$W_j=M_j/N$
D_1	D_{11}	3	8	0.375	0.753	0.7136	0.057
	D_{21}	5		0.625	0.690		
D_2	D_{12}	5	50	0.100	0.752	0.6309	0.357
	D_{22}	9		0.180	0.685		
	D_{32}	7		0.160	0.317		
	D_{42}	4		0.080	0.400		
	D_{52}	7		0.140	0.971		
	D_{62}	17		0.260	0.617		
	D_{72}	3		0.080	0.667		
D_3	D_{13}	5	60	0.083	0.500	0.6618	0.429
	D_{23}	2		0.033	0.477		
	D_{33}	23		0.383	0.737		
	D_{43}	9		0.133	0.844		
	D_{53}	6		0.100	0.167		
	D_{63}	5		0.083	0.712		
	D_{73}	8		0.133	0.819		
	D_{83}	3		0.050	0.479		
D_4	D_{14}	2	22	0.091	0.756	0.6380	0.157
	D_{24}	12		0.545	0.634		
	D_{34}	6		0.273	0.548		
	D_{44}	2		0.091	0.814		
$S = \sum_j W_j S_j = 0.6500$							

Note: The summary of the calculation shown in table 3 is based on the assumption that each indicator is equally significant in explaining statistical development. That assumption will be reviewed as data become available.

Step 1: In our case, the total number of indicators is **N=140** (from 1.1).

Step 2: The score of the first subdimension D_{11} in dimension D_1 is obtained using (1.4) as:

$S_{11}=2.26/3=0.753$

Where 2.26 is the total score of the 3 indicators corresponding to D_{11}

Step 3: The score of the first dimension D_1 is obtained using (1.5) as:

$S_1= 0.375*0.753 + 0.625 *0.690 = 0.7136$

Step 4: The overall score for the whole statistical system is given using (1.6) as:

S= 0.7136*0.057) +(0.6309*0.357) +(0.6618*0.429) +(0.6380*0.157) = 0.6500

34. The scores will be computed within the national statistical office under ECA guidance. After a certain period of time and ideally by 2020, the weighting schemes may be changed to principal component analysis, as shown in annex 4, part A. As suggested in the StatDI report, the aggregation scheme that would be used at that time would be the geometric mean from equation 1.5.b as shown above. The actual calculation as illustrated in table 3 could serve as a baseline for pilot studies conducted by each country on the basis of this manual.

4. Dissemination of the outcomes of Statistical Development Indicators

35. The dissemination of StatDI outcomes would serve three purposes:

(a) **Awareness-raising**: it would be very useful for the general public to be made aware of the tremendous work that has been undertaken in the development of StatDI. The ultimate goals would be to initiate a "word of mouth" effect that would help national statistical offices and ECA in their work, in addition to building an identity and a strong profile for StatDI.

(b) **Understanding**: various groups and audiences will be targeted in the dissemination process. It would be useful if they were not merely aware of the work being done, but also had a deep understanding of the StatDI project and results.

(c) **Action**: this is one of the main purposes of the dissemination process. It would be ideal if StatDI initiated a change in statistical systems practices. That change should

result in the use of outputs, materials, and approaches adopted by the StatDI project.

36. The main targets of StatDI are firstly national statistical office and, secondly, all other public and private-sector users of the projects undertaken by national statistical offices. The StatDI project will use a number of tools during the dissemination process, including the following:

(a) **Project website**: the website will host a rich array of content types related to StatDI. The website may have both public and non-public parts and will serve as the entry point for all interested stakeholders. The project website will serve as a major communication and dissemination tool throughout the project. Moreover, the StatDI website will be linked to the websites of participating national statistical offices. The purpose of the website will be to increase awareness about the StatDI project and outcomes on a broad and international scale and will serve as a comprehensive and growing base of knowledge for the StatDI team. The website will be the primary means by which upcoming events will be announced and alerts will be made regarding outcomes, policy decisions, and newsletter publications. The website will provide downloadable materials and contain features for online registration/ invitations to project meetings and workshops. The website will also include an advanced search engine, a Frequently Asked Questions section and a StatDI blog. The StatDI website will be regularly updated and use made of social media to enhance the engagement of a wide and diverse range of stakeholders with the project. The website will aim to integrate rich content in the form of web design, press releases, summaries, and short presentation videos in order to ensure a

high level of impact and a global profile for the StatDI project. Measuring the impact of the dissemination tool will be done by means of Google analytics and website statistics, including those on deliverable/document downloads, comments received, and requests for information.

(b) Blog: the blog will be developed as part of the website, with the aim of serving as a virtual forum for engaging with stakeholders and end users, and as a forum for exchanging and discussing experiences, methodologies, and results related to StatDI and statistical development. The goal of the blog will be to promote active interactions between project partners and interested parties participating in project activities and collect feedback on the project. Data measured on the blog will include the number of comments posted, the number of participants, the variety of participating sectors and the backgrounds of stakeholders.

(c) Social media: in addition to the website, dedicated StatDI Twitter, LinkedIn, Facebook and YouTube pages will be established to enhance the project's interactions with a broad-based audience. Outcomes measured will include the number of followers, the number of posts and the geographical mapping of the online audience.

(d) Audiovisual materials: these will include presentations of the outputs of and lessons learned from the project. These materials will be developed to accompany the StatDI project results and other disseminated materials in order to deliver project outputs to the public in a clear, direct and visual manner. Summaries, lasting two or three minutes and presented in layman's terms will be publicly available on the StatDI website and social media channels. The measured outcomes will include the

number of video abstracts, the number of views, the number of downloads and the number of pickups.

(e) Open access and peer-reviewed publications and conference presentations: given the architecture of StatDI, it will be possible to identify areas of research related to statistical development. The identification of those areas of research may give rise to calls for papers or for prospective conferences on StatDI.

(f) Technical materials: these will consist of leaflets, fact sheets, policy briefs and one pager documents that will be produced throughout the project to facilitate the communication of results to policymakers and the general public. StatDI policy and business briefs and guidelines, including an overview of dimension innovations, plus best practices, policy recommendations, and assessment tools will all be delivered at key stages of the project in order to inform and support decision-making by governments, national statistical offices and other stakeholders. Measured outcomes will include the number of technical materials produced and distributed, the number of events/channels distributing those materials, and the number of downloads from the project web portal.

(g) StatDI newsletter: a newsletter will be published every six months so as to inform stakeholders about StatDI project outcomes, upgrades and events. The newsletter will be distributed to stakeholders via the StatDI mailing list and will also be available on the website. The StatDI newsletter will aim to alert users regarding key developments and results, as well as encourage participation in project activities. The measured outcomes will include the number of contacts to whom the newsletter is sent, the number of requests made to join the mailing list

after receipt from third parties, the number of web-based newsletter downloads and hard copy distribution numbers.

(h) Targeted stakeholder workshops and events: It is essential that the knowledge related to evidence and innovation acquired through the StatDI project be translated from scientific finding to policy, and thus to tools relevant to citizens. In that regard, the dissemination of methods and results in meetings and conferences will be performed through national, regional and international workshops open to all regions and participating countries. Those events will be organized in collaboration with the main partners of the StatDI project, namely the Mo Ibrahim Foundation, the World Bank, Partnership in Statistics for Development in the 21st Century, the African Development Bank and the International Monetary Fund (IMF). Strong involvement by multiple partners and networks in major international scientific programmes and conventions will ensure a direct and efficient transfer of knowledge at different levels of the policymaking process and will represent a point of reference for ECA and for national statistical systems in shaping future statistical development policies, while taking inputs from a broad audience into consideration. Reports and presentations will be tailored to the needs of relevant policy communities. The measured outcomes will include the number of event participants, the number of publications distributed at events, the representation of policy-makers and certain stakeholder groups, the number of project invitations for external events, and the inclusion of workshop recommendations in policymaking.

37. A dissemination timetable will be elaborated as part of StatDI project implementation.

CHAPTER 2

DIMENSIONS AND SUBDIMENSIONS OF STATISTICAL DEVELOPMENT INDICATORS

A. Guidelines for completing the Statistical Development Indicators questionnaire

1. Prerequisites

38. The purpose of the questionnaire is to collect information on indicators for all dimensions and subdimensions used to calculate StatDl. **Only one questionnaire per country should be completed** and managed by the agency in charge of coordinating statistical activities in the country in question, which may be the national statistical office, a ministry responsible for supervising the production of statistics, the national statistics council, or another authority. If necessary, this may be done in collaboration with other relevant stakeholders in the national statistical system. When collecting data from more than one source or department, the office in charge of data collection must share the contents of this manual together with a message to questionnaire respondents explaining the significance of StatDl in statistical development and its use as a tool that enables statisticians to explore the weaknesses and strengths of a given national statistical system. That message may result in higher questionnaire response rates.

2. Module I - Functional institutional and organizational framework

39. Functional institutional and organizational frameworks aim to provide an overview of the organization of the national statistical system of a given country as well as the operational and legal organization of that system, the status of stakeholders and of how they can appropriately apply African and international standards.

40. This dimension includes variables pertaining to the organization and management of the national statistical system and relies on collaboration among stakeholders and between the national statistical system and international regional, subregional stakeholders, including international organizations.

This dimension is comprised of the subcomponents mentioned below.

41. The section of the questionnaire related to this subdimension focuses on political implications as well as the role of the national statistical office within the national statistical system, and relies on eight variables:

Legislative and regulatory frameworks for statistics

42. A country's statistics Act provides a framework for ensuring the professionalism of its national statistical office; ensuring that the functions of each of the components of the national statistical system are clearly established; providing for the coordination, harmonization, and effectiveness of the national statistical system; and empowering authorities to collect data. That Act should be in line with the Fundamental Principles of Official Statistics, as endorsed by the United Nations Statistical Commission and enshrined in the African Charter on Statistics. In addition, that legal framework establishes rules for compliance with statistical confidentiality to ensure that individual records are not accessed by unauthorized individuals or shared with political authorities or regulatory and tax authorities.

43. The following questions (1.A.1, 1.A.2 and 1.A.3) explore to what extent a given country respects international, regional or national rules for the national statistical system.

1.A.1.a. Specify the status of the statistics law/ act that regulates the statistical activities in the country

This question aims to confirm the existence of statistical acts or laws. More specifically, this question attempts to explore the status of such a law in terms of its existence and enforcement. In other words, whether it exists but has not yet entered into force; whether it has been drafted but not yet ratified; or drafted but not yet signed; or whether no such law or act has been formulated. Since this manual assumes

that most, if not all, countries have laws on statistics, the question is merely for informational purposes.

1A.1.b. Assuming the existence of a law/act that regulates statistical activities in the country concerned,

then subquestion 1.A.1.b aims to assess public access to those statistics. Four options are given, from which respondents choose one:

> ➢ Publicly available only on the website, in which case the score given is 0.35.

> ➢ Publicly available only as a hard copy version at the statistical office, in which case the score given is 0.15.

> ➢ Publicly available on both the website and as a hard copy version, in which case the score given is 0.5.

> ➢ Not publicly available, in which case the score given is zero.

Furthermore, the manual further explores the core provisions of the content of the law/Act with question 1.A.1.c.

Scoring scheme

Answer: Subquestions of question 1.A.1.c	Score
Does it explicitly make reference to the Fundamental Principles of Official Statistics?	
○ Yes	0.125
○ No	0.0
Is there language in the statistics law/Act stating that official statistics should be developed, produced and disseminated in line with the provisions of the African Charter on Statistics?	
○ Yes	0.125
○ No	0.0
Does the statistics law/act include provisions to address the following?	
○ Delineating the national statistical system	0.0625

○ Identifying producers of official statistics	0.0625
○ Neither of the above	0.0
Does the statistics law/act include provisions that specify the roles of the following official statistics producers within the national statistical system?	
○ National statistical office	0.0625
○ Other producers	0.0625
○ Neither of the above	0.0

According to the glossary of statistical terms of the Organization for Economic Cooperation and Development (OECD), a national statistical system is defined as, the ensemble of statistical organizations and units within a country that jointly collect, process and disseminate official statistics on behalf of the national Government. For the purposes of this manual, when an Act or law delineates the national statistical system, it means that it specifies and clearly describes the components or elements of the national statistical system.

I.A.2. What is the status of ratification of the African Charter on Statistics?

Scoring scheme

Answer: Status of ratification	Score
Not yet signed	0.25
Signed	0.5
Ratification instruments submitted to the African Union Commission	0.75
Ratified	1.0

I.A.3. Which of the following elements are clearly addressed by provisions in the statistics law/ act or policy, or other commitments that ensure statistical confidentiality? (You may select multiple options.)

Scoring scheme

Answer: elements addressed by provisions of the law/Act	Score
Protection of privacy	1/3
The law/Act makes a distinction among various types of "infractions", including, for example, carelessness, improper behaviour, use of confidential information for personal benefit	1/3
Penalties against persons who wilfully breach the statistical confidentiality sanctions exist on an administrative, penal and disciplinary level	1/3

Framework for organizing and coordinating the national statistical system

This subsection focuses on the coordination of the national statistical system, the prerequisites for organizing the national statistical office, whether the law/act specifies the status and functional responsibilities of the head of the national statistical office, and the scope of the work plan of the national statistical office.

I.B.1.a. Is there is a national body or institution responsible for the coordination of the national statistical system?

Score of 0.5 if such a body exists and zero if it does not exist.

Such bodies may include a national council that is responsible for the overall coordination of national statistics in the country. The existence of such a body implies that it ensures effective coordination of national statistics.

I.B.1.b. If yes, then how does that national body or institution coordinate the national statistical system?

This attempts to explore how coordination takes place in practice if a national body has been established. The respondent may tick more than one box.

Scoring scheme

Answer: Methods of coordination	Score
Regulations, agreements, memorandums of understanding or laws	0.10
Supervisory, advisory, or technical committees	0.10
Through the implementation of the annual or multi-annual statistics plan or National Strategy for the Development of Statistics	0.10
Joint data collection	0.10
Pre-approval of survey methodologies	0.10
None of the above	0.0
Others (to be specified)	0.10 if one or more of the above methods are not satisfied. Otherwise a score of 0 is assigned to provide a maximum total score for I.B.1.(b) of 0.5

I.B.2. Does the national statistical office satisfy the following conditions?

The conditions include whether scientific methods and procedures are established in full or partial accordance with professional considerations, and to what extent the national statistical office is autonomous with regard to managing its budget and establishing work programmes and its statistical release calendar.

Scoring scheme

Condition	Score
I.B.2.a. Scientific methods and procedures are always decided according to:	
Entirely professional considerations	1/4
Partial professional considerations	1/8
Absence of professional considerations	0
I.B.2.b. The budget is managed as follows:	
Autonomously	1/4
Semi-autonomously	1/8
Non-autonomously	0
I.B.2.c. Establishing the annual and multi-annual statistical work programmes is done:	
Autonomously	1/4
Semi-autonomously	1/8
Non-autonomously	0
I.B.2.d. Establishing the statistical release calendar is done:	
Autonomously	1/4
Semi-autonomously	1/8
Non-autonomously	0

I.B.3. Are the following elements regarding the status and functional responsibilities of the head of the national statistical office addressed in the law/Act?

This question is related to the provisions in the law on the conditions for appointment, dismissal and responsibilities of office of the head of the national statistical office.

Scoring scheme

The status and the functional responsibilities of the head of the national statistical office	Score
Appointment	1/4
Dismissal	1/4
Responsibilities of office	1/4
Duration of office	1/4

I.B.4. Level within the hierarchy of public service of the country of the national statistical office head or chief executive officer:

The head of the statistical office should be the most senior official in the national statistical system, and as such, the position should be granted a sufficiently high status in the government hierarchy in order to be effective. If that official is entrusted with an important level of decision-making power and autonomy, this will help to ensure that official statistics produced are of high quality.

The hierarchy levels for government positions include the following: Level 1 (Minister); Level 2 (Permanent Secretary); Level 3 (Director General); Level 4 (Director); Level 5 (below Director).

If the head of the national statistical office is at Level 1 or Level 2 the score given is 1, Level 3 would result in a score of 0.75, Level 4 a score of 0.5 and Level 5 a score of 0.

I.B.5. Does the scope of the national statistical office work plan cover the whole national statistical system?

This question is designed to determine whether or not the national statistical office work plan framework reflects and captures all relevant statistical activities conducted in order to produce official statistics within the national statistical system.

A score of 1 if it does and 0 if it does not.

3. Module II - Capacity for an efficient statistical system

45. The part of the questionnaire that covers capacity for an efficient statistical system traces the financial, human, technical, material and infrastructural resources needed to ensure the efficient functioning of the national statistical system of a country. The aim is to identify the capacities and

contributions of partners who support the effective functioning of the system.

Funding

46.   The funding subdimension examines the funding of statistical activities. The level of funding shows the involvement of the government in the development of statistics, and provides a better understanding of its importance in the development processes of the country. Variables included as part of this component are the proportion of the statistical budget financed by the government, which indicates the level of its commitment to provide adequate statistics and sustainable financial resources, as well as the contribution made by various stakeholders.

47.   The African Union Specialized Technical Committee on Finance, Monetary Affairs, Economic Planning and Integration has recommended dedicating at least 0.15 per cent of a country's national budget to statistics research. Questions in this area are designed to learn if the country is complying with that recommendation and what share of the national budget is allocated to the national statistical office.

II.A.1. What percentage of last year's total national budget was allocated for all statistical activities at the national statistical office?

Scoring is done as follows:

The proportion of the government budget allocated to the national statistical office (in terms of execution)	Score
If >= 0.15 per cent	1
If <0.15 per cent	Real allocation, as a percentage, divided by 0.15 per cent (in terms of execution)

The question aims to understand whether the national statistical office budget exceeds, meets or

is less than the 0.15 per cent of the total national budget, as recommended by Specialized Technical Committee on Finance, Monetary Affairs, Economic Planning and Integration. The percentage is calculated by dividing the allocated and executed national statistical office budget by the total national budget. For example, if X is the allocated and executed budget for the national statistical office and R is the total national budget, then the percentage is simply X divided by R.

Note: The national statistical office has been chosen instead of the national statistical system because it could be difficult for many countries to calculate allocations to sectoral statistical offices that operate separately from the national statistical office.

II.A.2. Complete the following table:

Provide the actual share (percentage) of the overall national statistical office budget by funding source (last five years, indicating both allocated and received funds).										
Funding source	Last year (Y-1)		Y-2		Y-3		Y-4		Y-5	
	A	R	A	R	A	R	A	R	A	R
Government										
National statistical office										
International partners										
Other sources										

This table shows the different sources of funding received by the national statistical office. The letter A denotes allocated funds, while the letter R denotes the funds actually received. For scoring purposes, a focus is placed on the proportion of the statistical budget financed by the government. Scoring is based on the proportion of funding of the national statistical office budget in terms of allocation during the last five years. The provided data attempts

International partners									
Other sources									

to assess the regularity of government provided funding as a part of the national statistical office budget over the previous five years.

To calculate the government percentage share in any given year, let us assume the national statistical office overall budget to be (X+Y+Z+W), where X is funded by the government, Y is funded by the national statistical office, Z is funded by international partners and W is funded by other sources, then the requested share is X divided by (X+Y+Z+W). The percentage shares of other funding sources may be calculated using the same formula but with the order of variables adjusted.

The score for each year will range from 0 if the government has a zero contribution rate to 0.2, if the total budget of the national statistical office is funded by the government. The score for any given year may be computed using the formula:

Score for a given year $=($ percentage share $)/5$

The final score for question II.A.2 is calculated as the sum of the five annual scores divided by five using the above formula, namely as a simple average of the five percentage shares.

II.A.3. Complete the following table:

Provide the actual share (percentage) of the overall national statistical system budget by funding source (last five years, indicating both allocated and received funds).									
Funding source	Last year (Y-1)		Y-2		Y-3		Y-4		Y-5
	A	R	A	R	A	R	A	R	A	R
Government									
National statistical office									

This question is for informational purposes only. It can serve as a proxy indicator if the information for question II.A.2 is missing or incomplete. The letter A denotes allocated funds and the letter R denotes received funds.

II.A.4. Provide the percentage share of the last population and housing census budget funded by the government's national budget.

The score is proportional to the share, ensuring that the score will be between 0 and 1, and will encourage countries to become self-reliant in terms of funding their population censuses.

II.A.5. Are there delays between the commitment and the release of funds by the government to the national statistical office?

In general, there are delays in this process. If the answer is yes and delays undermine the country's statistical activities, a score of 0 is assigned. If the answer is yes but the delays do not undermine the country's statistical activities, then a score of 0.8 is assigned. If there are no delays, then a score of 1.0 is assigned.

II.A.6.a. Does the country have an operational statistical development fund?

A score of 0.7 is assigned if such a fund exists while 0 is assigned if it does not.

Question II.A.6.b adds a bonus of 0.3 if the fund is sustainable in the short term, and if there has been a steady or increasing share of funding over the previous five years. Question II.A.6.c is for informational purposes only.

Good governance of the statistical office

48.   One of the lessons learned from past efforts to promote statistical development in African countries is that the creation of a plan is not all that is required. Stakeholders must also take action to ensure that that plan is implemented, and implementation must be closely monitored. There is a need to ensure the monitoring of statistical activities included in plans developed by countries so that progress, or a lack thereof, may be monitored. In addition, the national statistical office should regularly produce reports on statistical activities undertaken during a given period.

49.   This part of the questionnaire consists of a subdimension with variables related to the governance of the national statistical office and the overall national statistical system. The questions and indicators are as follows:

II.B.1. How are the senior managers of various units within the national statistical office appointed?

Senior managers comprise the top management of the national statistical office, and include general directors, and their deputies and assistants.

Scoring scheme

Answer	Score
Based on clear and transparent recruitment processes, such as a call for applications, for example	1
Appointment by the national statistical office governing body	0.5
Nominated by a higher authority, such as by the president, prime minister or a minister	0
Other	The score will depend on the answer. For example, a score may be 0.75 if the process is transparent and directed by the national statistical office, or the score may be 0.25 if the process is partly a choice by the national statistical office and partly an independent nomination.

II.B.2. Is there an annual statistical work plan?

A score of 1 is given if the answer is yes and the plan covers the entire national statistical system, a score of 0.75 is given if the plan covers only the national statistical office, and a score of 0 is assigned if no work plan has been developed.

50.   Coordination, collaboration, and partnership mechanisms are essential for the smooth functioning of the national statistical system. A coordination body should ensure the coherence and comparability of information collected through various sources in the system.

II.B.3. Is there a unit or team in charge of the coordination and implementation of statistical activities at the national statistical office?

A score of 1 is given if a unit or team exists and if not, a score of 0 is given.

II.B.4. Is there a framework for the evaluation of staff at the national statistical office?

Scoring scheme

Answer	Score
Yes, and this includes indicators for the promotion of staff	0.25
Yes, and this includes indicators for enhancing staff capacity	0.25
Yes, and this provides incentives and supports staff morale	0.25
Yes, and this matches staff with work duties	0.25
None of the above, but other frameworks have been developed	0.50
No framework has been developed	0.00

II.B.5. Annual report on the implementation of statistical programmes

Scoring scheme

Answer	Score
Only the national statistical office prepares the annual report on the implementation of statistical programmes	0.20
The national statistical office prepares the annual report on the implementation of statistical programmes in close coordination with the other producers of official statistics	0.40
The annual report on the implementation of statistical programmes includes the budgetary execution	0.20
The annual implementation report is published on the website	0.10
The annual implementation report is submitted to the statistical council or equivalent governing body for input	0.20
The opinion of the statistical council or the equivalent governing body on the annual implementation report is published	0.10
There is no annual report	0.00

51.   The following questions are related to the integrity of the financial accounts produced by the national statistical office, and the public availability of audited financial statements.

II.B.6.a. Are the financial statements of the national statistical office audited and certified on an annual basis by an external control body?

A score of 0.5 is given if the answer is yes and 0 if the answer is no.

II.B.6.b. Are the results of the auditing and certification of the financial statements made public?

A score of 0.5 is given if the answer is yes and 0 if the answer is no.

II.B.7.a. Is there a national quality assurance framework?

Scoring scheme

Answer	Yes, and this is publicly accessible online (1)	Yes, but it is not publicly accessible online (0.5)	No (0)
Score	0.50	0.25	0

II.B.7.b. Are the following topics covered by the national quality assurance framework?

Scoring scheme

Answer	Managing the statistical system	Managing the institutional environment	Managing statistical processes	Managing statistical outputs
Score	0.125	0.125	0.125	0.125

52.   This part of the questionnaire aims to ascertain the organization of the national statistical system with respect to statistical production.

II.B.8.   This question focuses on efforts to strengthen national statistical office statistical production. The relevant indicator is calculated

by looking at the proportion of ministries that have statistics units producing regular statistical outputs in coordination with the national statistical office relative to the total number of ministries that have statistics units.

Score = The figure obtained from II.B.8.b divided by the figure obtained from II.B.8.a.

In addition, based on the table, it is possible to calculate the proportion of ministries for which statistics are produced by the national statistical office. However, this proportion calculation is for informational purposes only.

II.B.9. Is there a user focused official statistics committee, or equivalent, or a consultation mechanism that ensures that the statistical information published is relevant to users?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

Human capital

53.   The human capital subdimension focuses on staff professionalism and training efforts in the national statistical system. The aim is to identify, maintain and develop staff skills for optimal performance.

54.   A human capital development plan is expected to provide the following:

   (a)   Training for skills growth in order to ensure updated scientific knowledge;
   (b)   Raising staff awareness of the benefits of standardization;
   (c)   An adequate implementation budget.

For more details, see the Fundamental Principles of Official Statistics: Implementation Guidelines.[5]

The following questions are related to the variables included in this subdimension:

11.C.1.a. Is there a human resources development plan in the country's national statistics system?

A score of 1.0 is given if a plan has been developed and this covers the whole national statistical system.

11.C.1.b. Does a plan exist only at the national statistical office?

A score of 0.7 is given if a plan exists only within the national statistical office.

Note: Question II.C.1.(a) and II.C.1.(b) are mutually exclusive.

11.C.2. Is there a human resources management plan at the national statistical office that addresses the following issues?

Scoring scheme

Answer	Appointment of staff	Promotion of staff	Mobility of staff	Dismissal of staff	There is no human resources management plan at the national statistical office
Score	0.25	0.25	0.25	0.25	0

II.C.3. What is the proportion of professional staff (staff with a qualification in statistics, demography or related field) dedicated to the

production and dissemination of statistics at the national statistical office?

The score is calculated as a proportion of the actual number of the statisticians, demographers and economists relative to the required number of staff in those roles.

Score = actual number (statisticians + demographers + economists) divided by the required number (statisticians + demographers + economists)

II.C.4.

This question is similar to the question posed in II.C.3, and the score for this indicator is the proportion of professionals plus non-professionals relative to the corresponding required number of both categories of staff within a national statistical office.

II.C.5. Indicate the number of national statistical office professional staff who left the institution during the previous year and their reasons for leaving:

Professional staff must have obtained at least a first degree. The focus of this question is to capture the percentage share of professional staff who resigned from the national statistical office during the previous year. A high turnover rate will undermine the sustainability of professional staff at a national statistical office.

Score = 1 minus the percentage share of those who resigned due to a lack of job satisfaction. Other data are for informational purposes only.

55.   Training is assessed as per the number of university students enrolled and specializing in statistics. Measurements include: the total number of students per 100,000 inhabitants; the provision of government scholarships for students in statistics at national or regional schools and universities; the existence of an initial and/or ongoing training

programme for national statistical office statisticians; the existence of "train the trainers" programmes; and the organization of ongoing training sessions for the national statistics system staff by the office of statistics or other training centres.

II.C.6. Provide the number of students enrolled in the field of statistics in the country's universities last year.

The answer to this question is for informational purposes only.

II.C.7. Provide the number of students enrolled in the field of statistics who received a scholarship last year.

The answer to this question is for informational purposes only.

II.C.8. Is there a recycling programme for national statistical office staff?

A score of 1 is given if a recycling programme exists and 0 if it does not.

II.C.9.a. Is there a training programme for human capacity-building and development in statistics and/or related fields, including ICT?

A score of 0.3 is given if the answer is yes, but only within the national statistical office, while a score of 0.5 is given if the answer is yes, and the training programme is designed for the entire national statistical system. A score of 0 is given if the answer is no, and no training programme exists.

II.C.9.b If a training programme exists, please comment on its implementation:

Scoring scheme

Answer	Score
Implementation is led by the national statistical office	0.2
Implementation is led by another body	0.1
Implementation runs smoothly and without financial and technical challenges	0.3
Implementation process faces certain financial and technical challenges	0.1
No programme is being implemented	0

Note: the first two options in II.C.9 (b) are mutually exclusive, as are the last three options.

II.C.10. Does the national statistical office conduct performance assessments of staff?

Employee performance is assessed frequently, and at least annually, and new performance targets are set on the basis of those assessments.

A score of 1 is given if such a practice exists and 0 if it does not.

Physical and material infrastructure of the national statistical office

56.   The following part of the questionnaire addresses the subdimension devoted to the physical and material infrastructure of the national statistical office.

II.D.1. Does the national statistical office have its own building?

A score of 1 is given if the building is owned by the national statistical office, while a score of 0.75 is given if it is owned by the government. If it is rented

a score of 0.5 is given and if the national statistical office has no building at all a score of 0 is awarded.

II.D.2. Please tick the circles that best describe the building that houses the national statistical office headquarters.

The scoring scheme assigns 1/15 for each characteristic met. The total scoring for this question is the sum of scores from II.D.2.a, II.D.2.b, and II.D.2.c.

II.D.3. Does the staff working space satisfy the following requirements? Multiple options may be selected.

A score of 1/5 is given for each chosen option and 0 if none are chosen. The final score is the sum of the chosen options. That is to say

Answer	Score
A desk is assigned to each professional staff member	1/5
The statistical agency provides a personal computer to each professional staff member	1/5
The statistical agency allocates filing cabinets to each professional staff member	1/5
All staff members have access to a means of communication	1/5
Access to a means of communication, such as the telephone and the Internet, is provided in the working spaces of each staff member	1/5
None of the above	0

II.D.4.a. Has fieldwork conducted by the national statistical office been delayed or negatively affected by insufficient transport infrastructure, which has prevented staff from carrying out all scheduled activities at a given time?

This question assesses the availability of transport infrastructure.

A score of 0.5 is given if the response is no and 0 if it is yes.

II.D.4.b. Are the vehicles assigned for fieldwork by the national statistical office, including those used to conduct censuses and surveys, in good condition and appropriate for fieldwork activities?

This question assesses the quality of transport infrastructure.

A score of 0.5 is given if the response is yes and 0 if it is no.

Statistical infrastructure

57. This subdimension addresses statistical infrastructure by posing the following questions:

II.E.1 Is there a unit/section, or an internal arrangement within the national statistical office, for the development of standards and classifications in the area of statistics?

A score of 1 is given if the answer is yes and 0 if the answer is no.

II.E.2a. Is there a unit/section in charge of a statistical register within the national statistical office?

A score of 0.5 is given if the answer is yes and 0 if the answer is no.

II.E.2b If yes, then scoring is as follows:

Answer	The statistical register has a database	The database is regularly updated	The statistical register is linked to censuses of the population and households	Guidelines on the use of the register have been developed and made available to the public, such as on the national statistical office website
Score	1/8	1/8	1/8	1/8

II.E.3. Is there a unit dedicated to statistical methodologies, including sampling and questionnaire design, within the national statistical office?

A score of 1 is given if the answer is yes and 0 if the answer is no.

II.E.4. Is there a unit in charge of the geographical information system within the national statistical office?

A score of 1 is given if such a system exists and 0 if it does not.

II.E.5.a. Does the national statistical office use licensed statistical software?

A score of 1 is given if the answer is yes and 0 if the answer is no.

II.E.5b. Please specify the type of licensed software.

This is for informational purposes only.

II.E.6. What electronic devices does the national statistical office use for data collection during censuses and national surveys? Choose multiple options where appropriate.

Scoring scheme

Answer	Lap-top	Tablets, smart-phones and other handheld devices.	Other, please specify	No de-vice used
Score	0.5	1 (for at least one device)	0	0

Note: The maximum score for *II.E.6* should not exceed 1.

II.E.7. What type of data collection is usually used in censuses and national surveys? Multiple options may be selected.

PAPI: Paper Assisted Personal Interview
CATI: Computer Assisted Telephone Interview
CAPI: Computer Assisted Personal Interview
CAWI: Computer Assisted Web Interview
MAPI: Mobile Assisted Personal Interview

Scoring scheme

Answer	PAPI	CATI	CAPI	CAWI	MAPI	Other, or none of the above
Score	0	0.5 for at least one			1.0	0

Note: The maximum score for II.E.7 should not exceed 1.0.

Information and communications technology infrastructure

58. This part of the questionnaire assesses within this subdimension the use of ICT by the national statistical office.

II.F.1. Is there an information technology (IT) department/unit/division at the national statistical office?

A score of 1 is given if one exists and 0 if one does not.

II.F.2.a. What is the total number of staff in the IT department?

This is for informational purposes only.

II.F.2.b. How many IT personnel work in the IT department?

IT personnel refers to those with a minimum of either a two-year post-secondary school diploma or in-house training in IT, or at least four years of work experience in the IT field.

This question is for informational purposes only and no score is given.

IIF.3. Does the head of the IT department/unit/division have post-graduate credentials in IT or a statistics-related field?

Post-graduate credentials mean at least a one-year diploma programme taken after completion of a first degree at a higher learning institution. In general, a postgraduate degree in IT requires at least five years at an institution of higher learning after completing secondary school education.

Scoring scheme

Answer	Score
Yes, IT skills	0.5
Yes, statistical skills	0.5
None of the above	0

II.F.4. How many professionals at the national statistical office have been allocated a personal computer?

Professional staff include those with a minimum of a first degree. This question only considers the number of functional personal computers and the total number of professionals as defined by II.C.3 and II.C.4.

The score is calculated as the ratio of the number of computers to the number of professional staff members. The ratio is calculated by dividing the number of functional personal computers allocated to the professional staff members by the total number of professional staff working in the national statistical office.

II.F.5. Is there a local area network (LAN) within the national statistical office?

A score of 1 is given if the answer is yes and 0 if the answer is no.

II.F.6. Does the national statistical office have its own servers?

A score of 1 is given if the answer is yes and 0 if the answer is no.

II.F.7. Is there a centralized databank66 that is managed by the national statistical office?

A score of 1 is given if a centralized databank exists and 0 if it does not.

II.F.8 Does the national statistical office use the following geospatial information system tools?

Geospatial information system tools include geospatial software, remote sensing software, Global

Positioning System equipment and software and a geospatial database.

Scoring scheme

Tool	Score
Geospatial software	0.2
Remote sensing software	0.2
Global Positioning System equipment	0.2
Global Positioning System software	0.2
Geospatial database	0.2
None of the above	0.0

II.F.9.a. Is there a national statistical office website that is regularly updated?

A score of 0.5 is given if the answer is yes and 0 if the answer is no.

II.F.9.b. If there is a website estimate the time lag between the release of statistical information by the national statistical office and the availability of this information on the website.
Scoring scheme

Time lag	Score
More than one month	0.0625
Within one month	0.125
Within one week	0.25
Within one day	0.5
Website is not updated	0

6 A centralized databank is a database that is located, stored and maintained in a single location.

II.F.10 Does the IT security policy at the national statistical office satisfy the following? Multiple options may be selected.

Scoring scheme

Dimension	Score
Availability (statistical agencies provide internal and external users with access to data as needed)	0.25
Integrity (statistical agencies ensure that survey and processing methods are effective and prevent human or technological manipulation of data)	0.25
Confidentiality (statistical agencies uphold provisions regarding confidentiality and data protection and ensure that data are only used for statistical purposes)	0.25
National statistical office staff are familiar with the IT security policy	0.25
There is no IT security policy at the national statistical office	0.0

II.F.11.a. What is the Internet bandwidth at the national statistical office?

This question is for informational purposes only.

II.F.11.b. What is the maximum Internet download speed for users?

This question assesses the reliability of Internet access.

Scoring scheme

Answer	Score
At least 10Mb/s	1.0
At least 5Mb/s, but less than 10Mb/s	0.5
At least 1Mb/s, but less than 5Mb/s	0.25
Less than 1 Mb/s	0.125
No Internet available	0

II.F.12. Is there an IT master plan?

A master plan setting forth the overall IT strategy, helps provide guidance on IT for an organization.

While an IT strategy focuses on how IT will help the business succeed, an IT master (or strategic) plan is a roadmap that can facilitate efforts by a business to implement the strategy.

A score of 1 is given if an IT plan exists and 0 is given if it does not.

II.F.13 Does the national statistical office maintain an integrated database from a range of data sources?

A score of 1 is given if one is maintained and 0 if it is not.

II.F.14. How many national statistical office professional staff use their own computer (and not one provided by the office) to undertake national statistical office-related work?

This question explores the extent to which the professional staff at a national statistical office use their own computers for work rather than the computers provided by the office.

The score awarded is 1 minus the proportion of professional staff who use their own computer divided by the total number of professionals.

Active promotion of bilateral and multilateral cooperation for statistical capacity-building

59.   This part of the questionnaire aims within this subdimension to gather information regarding efforts made to facilitate bilateral and multilateral cooperation.

II.G.1. Which of the following has your national statistical office undertaken as part of its engagement with other national statistical offices in the last five years?

Scoring scheme

Methods of engagement	Score
Official visits of heads of national statistical offices to other offices	0.25
Technical meetings to exchange knowledge and best practices	0.25
Publication of newsletters or a webpage providing information about current projects	0.25
Hosting of trainees from other statistical institutes	0.25
Other (please specify)	0.25
None of the above	0
Note: The total score should not exceed 1	

II.G.2. Does the website disseminate information on best practices and lessons learned that can be applied by other national statistical offices?

Note: This question is related to question *II.F.9*. Therefore, if the answer to *II.F.9* is no, then this question should be disabled in the system.

A score of 1 is given if the website disseminates that information and a score of 0 awarded if it does not.

II.G.3. In which of the following activities has the national statistical office participated in the past year?

Scoring scheme

Activity	Score
Meetings of the United Nations Statistical Commission	1/3
Meetings of the Statistical Commission for Africa	1/3
Meetings of the African Union Committee of Directors General of National Statistical Offices	1/3
None of the above	0

II.G.4.a. How many national statistical office professional staff have participated in training programmes carried out by global and regional statistical agencies in the past year?

A score of 0.5 is given if one or more professionals participated in a training programme and 0 if no professionals participated.

II.G.4.b. How many national statistical office professional staff have participated in a training programme abroad and received financial support from the government in order to attend that programme?

A score of 0.5 is given if one or more professionals received financial support from the government in order to participate in a training programme, and a score of 0 is given if no professionals did so.

II.G.5. How many national statistical office professional staff are members of professional bodies or networks?

This question is for informational purposes only.

II.G.6. Does the national statistical office encourage its staff to become chartered statisticians?

This question is for informational purposes only.

4. Module III - Production of relevant statistics that meet international quality standards

60.   The section on the production of relevant statistics that meet international quality standards focuses on the process by which data are produced by assessing needs, the programming of statistical production, as well as the comprehensiveness, periodicity and timeliness of data collected through censuses and surveys. In addition, data collection

methods, data management and quality assurances regarding data production, archival, storage and security are assessed.

Proper evaluation of data requirements

61.   This part of the questionnaire within this subdimension seeks to identify efforts made by the national statistical office to satisfy the demand for statistics through systematic evaluations of data needs, as required by the country's national development plan, the 2030 Agenda, and Agenda 2063 of the African Union, as well as the existence of mechanisms and surveys used to assess the specific needs of other stakeholders including relevant stakeholders in the private sector, and civil society.

III.A.1. Is there a systematic evaluation of data needs as per a national development plan or equivalent?

A full evaluation is when a statistical programme (for example the National Strategy for the Development of Statistics) is entirely based on the country's national development plan or equivalent. A **partial evaluation** is when the statistical programme takes into account the needs arising only from major sectors of the national development plan or its equivalent, and **no evaluation** is when there is no link between a statistical programme and the national development plan.

Scoring scheme

Status of evaluation	Score
Yes, a full evaluation is carried out	1.0
Yes, a partial evaluation is carried out	0.5
No evaluation is carried out	0
Other evaluation, specify type	0.25

III.A.2. Is there a systematic evaluation of the data requirements of initiatives to facilitate

achievement of the Sustainable Development Goals?

A **full evaluation** is when the statistical programme (for example the National Strategy for the Development of Statistics) contains all of the nationally prioritized Sustainable Development Goal indicators, while a **partial evaluation** is when some of the needs of the nationally prioritized Sustainable Development Goal indicators are taken into account. **No evaluation** is when there is no link between the statistical programme and the data requirements of Sustainable Development Goal initiatives.

Scoring scheme

Status of evaluation	Score
Yes, a full evaluation is carried out	1.0
Yes, a partial evaluation is carried out	0.5
No evaluation is carried out	0
Other evaluation, specify type	0.25

III.A.3. Is there a systematic evaluation of the data requirements of initiatives to facilitate implementation of Agenda 2063 of the African Union?

A **full evaluation** is when the statistical programme (for example the National Strategy for the Development of Statistics) contains all of the data needs for Agenda 2063, while a *partial evaluation* is when some of the target requirements of Agenda 2063 are taken into account in the statistical programme. **No evaluation** is when the statistical programme is not aligned with Agenda 2063.

Scoring scheme

Status of evaluation	Score
Yes, a full evaluation is carried out	1.0
Yes, a partial evaluation is carried out	0.5
No evaluation is carried out	0
Other evaluation, specify type	0.25

III.A.4. Is there a systematic evaluation of the data requirements of various ministries, departments and government agencies?

A **full evaluation** is when a national statistical office or other authorized body conducts a formal survey or assessment of the needs of all government ministries, departments and agencies while a **partial evaluation** is when a national statistical office or other authorized body conducts a formal survey or assessment of some of the needs of government ministries, departments and agencies, and **no evaluation** is when there is no direct link between the statistical programmes and the programmes of government ministries, departments and agencies.

Scoring scheme

Status of evaluation	Score
Yes, a full evaluation is carried out	1.0
Yes, a partial evaluation is carried out	0.5
No evaluation is carried out	0
Other evaluation, specify type	0.25

III.A.5.a. Is there a mechanism (such as a survey or a user-producer forum) to assess the data needs of other key data users including those in the private sector, civil society or the media?

A **full assessment** is when there a formal assessment of data needs for non-governmental users, including those in the private sector, is carried out on a regular basis, while a **partial assessment** is when formal assessments of the data needs of non-governmental users are carried out on a non-regular basis. **No assessment** is when there is no formal assessment of the requirements of non-governmental users.

Scoring scheme

Status of assessment	Score
Yes, a full assessment is carried out	1.0
Yes, a partial assessment is carried out	0.5
No assessment is carried out	0
Other evaluation, specify type	0.25

III.A.5.b. If yes, please explain the mechanism used.

This answer is for informational purposes only.

Statistical programming

62.   This subdimension examines statistical programming issues in the context of the National Strategy for the Development of Statistics or equivalent. It describes the process of developing a strategy and steps to ensure the strategy's completeness and consistency, formal adoption and funding by the government, as well as the funding of activities to collect key data within a specified timeframe, and the adoption of a multi-year programme of censuses and surveys.

III.B.1.a. Is there a multi-year strategic plan or strategy for the development of statistics (National Strategy for the Development of Statistics) in the country?

Scoring scheme

Activity	Score
Yes, it is being implemented	0.2
Yes, but it has expired	0.1
Yes, and it covers the whole national statistical system	0.2
Yes, but only for the national statistical office	0.1
No, there is no multi-year strategic plan or strategy	0

Note: the first four options are mutually exclusive.

III.B.1.b. If the answer to III.B.1.a is yes, check the following boxes as applicable. It is possible to make multiple choices.

Scoring scheme

Choice	Score
The strategy or programme design process was participatory and unrestricted and consultations involved multiple stakeholders, such as government ministries and departments, the private sector, civil society and members of academia	0.1
The strategy or programme has been officially adopted by the government	0.1
The strategy or programme is fully funded by the government	0.1
The strategy or programme is partially funded by the government	0.1
The strategy or programme is aligned with the national development plan or equivalent	0.1
The strategy or programme is aligned with the Sustainable Development Goals	0.1

III.B.2. Which of the following sectors have adopted a strategic plan for statistics? Multiple options may be selected.

A score of 0.1 is given for each chosen sector, and if no sector is chosen, a score of 0 is given. The total score should not exceed 1.0.

Scoring scheme

Sector	Score
Education	0.1
Health	0.1
Agriculture	0.1
Water supply	0.1
Manufacturing	0.1
Employment/labour	0.1
Transport	0.1
Environment	0.1
Energy	0.1
Other, please specify	0.1
None of the above	0

Comprehensiveness, periodicity and timeliness of data collected through censuses and surveys

63. This part of the questionnaire within this subdimension assesses the completeness, periodicity and regularity of data collected through multiple surveys and/or censuses, including population and housing censuses conducted over the previous ten years, agriculture censuses conducted over the previous five years, livestock censuses conducted over the previous ten years, poverty surveys with specific income and expenditure modules conducted over the previous five years, surveys with modules focused on demography, health, education, gender, children, violence against women, and the environment conducted over the previous five years and surveys with employment and business modules conducted over the previous ten years.

III.C.1. When was the last population and housing census completed?

A score of 1 is given if the last census was completed in the previous ten years, while a score of 0 is given if this is not the case.

III.C.2. When were the three most recent agriculture censuses completed?

A score of 1 is given if the last census was completed in the previous five years, while a score of 0 is given if this is not the case.

III.C.3. When were the three most recent livestock censuses completed?

A score of 1 is given if the last census was completed in the previous ten years, while a score of 0 is given if this is not the case.

III.C.4. When were the three most recent surveys that included income and expenditure modules completed?

A score of 1 is given if the last census was completed in the previous five years, while a score of 0 is given if this is not the case.

III.C.5. When were the three most recent surveys that included a demography module completed?

A score of 1 is given if the last census was completed in the previous five years, while a score of 0 is given if this is not the case.

III.C.6. When were the three most recent surveys that included a health module completed?

A score of 1 is given if the last census was completed in the previous five years, while a score of 0 is given if this is not the case.

III.C.7. When were the three most recent surveys that included an education module completed?

A score of 1 is given if the last census was completed in the previous five years, while a score of 0 is given if this is not the case.

III.C.8. When were the three most recent surveys that included a governance module completed?

A score of 1 is given if the last census was completed in the previous five years, while a score of 0 is given if this is not the case.

III.C.9. When were the three most recent surveys that included a module on gender completed?

A score of 1 is given if the last census was completed in the previous five years, while a score of 0 is given if this is not the case.

III.C.10. When were the three most recent surveys that included a module on children completed?

A score of 1 is given if the last census was completed in the previous five years, while a score of 0 is given if this is not the case.

III.C.11. When were the three most recent surveys that included a module on violence against women completed?

A score of 1 is given if the last census was completed in the previous five years, while a score of 0 is given if this is not the case.

III.C.12. When were the three most recent surveys that included a module on the environment completed?

A score of 1 is given if the last census was completed in the previous five years, while a score of 0 is given if this is not the case.

III.C.13. How many surveys that included a module on employment have been completed in the last ten years?

A score of 1 is given if there have been one or more, and 0 if no surveys have been carried out.

III.C.14. How many surveys that included a module on the establishment of businesses have been completed in the last ten years?

A score of 1 is given if there have been one or more, and 0 if no surveys have been completed.

64.   This part of the questionnaire examines the registration of births and deaths in the country, the edition of the IMF Balance of Payments Manual used in the country, the periodicity of national accounts indicators, the frequency of industrial production indicators, the frequency of price statistics at the national level, the frequency of employment

indicators and the existence of a formal framework for the organization of the current statistical output intended to improve thematic and geographical coverage.

III.C.15. As a percentage of total births, what was the birth registration rate in the country last year, as per civil registration figures?

A score of 1 is given if the percentage share is 70 per cent, while a score of 0.75 is given if it is less than 70 per cent but greater than or equal to 50 per cent, and a score of 0 is given if the percentage is less than 50 per cent.

III.C.16. As a percentage of total deaths, what was the death registration rate in the country last year, as per civil registration figures?

A score of 1 is given if the percentage share is 70 per cent, while a score of 0.75 is given if it is less than 70 per cent but greater than or equal to 50 per cent, and a score of 0 is given if the percentage is less than 50 per cent.

III.C.17. Which edition of the IMF Balance of Payments Manual is used in the country?

This information sheds light on whether the country is using the most recent edition of the Manual. The sixth edition is currently the latest edition and incorporates all current innovations.

A score of 1 is given if the sixth edition is being used, a score of 0.5 if the fifth edition is being used and a score of 0.25 if any other edition is being used.

III.C.18.a. How long is the delay in the provision of statistics to the national statistical office so that it can make annual national accounts indicators available to the public?

Scoring scheme

Periodicity of delay	Score
Six months at most	0.5
More than six months, but equal to or less than nine months	0.25
More than nine months	0

III.C.18.b. How long is the delay in the provision of statistics to the national statistical office so that it can make quarterly national accounts indicators available to the public?

Scoring scheme

Periodicity of delay	Score
Two months at most	0.5
More than two months, but equal to or less than three months	0.25
More than three months	0

III.C.19. How long is the delay in the provision of statistics to the national statistical office so that it can make quarterly agriculture production indicators available to the public?

Scoring scheme

Periodicity of delay	Score
Three months at most	1
More than three months, but equal to or less than six months	0.5
More than six months	0

III.C.20. How often does the national statistical office publically release quarterly industrial production indicators?

Scoring scheme

Frequency	Score
Regularly	1
Not on a regular basis	0.5
No releases to the public	0

III.C.21. a. How often does the national statistical office release national consumer price statistics?

Scoring Scheme

Frequency	Score
Monthly	1/3
Less often than monthly	1/6
Not released to the public	0

III.C.21.b. How often does the national statistical office release national industrial price statistics?

Scoring scheme

Frequency	Score
Quarterly	1/3
Less often than quarterly	1/6
Not released to the public	0

III.C.21.c. How often does the national statistical office release national consumer service price statistics?

Scoring scheme

Frequency	Score
Quarterly	1/3
Less often than quarterly	1/6
Not released to the public	0

III.C.22. How often does the national statistical office release employment indicators?

Scoring scheme

Frequency	Score
◯ Quarterly	1
◯ Every six months	0.75
◯ Annually	0.5
◯ Every five years	0.25
◯ Less often than every five years	0.125
◯ Not released to the public	0

III.C.23 Is there a policy for updating the production of high-frequency indicators, including prices at the national or subnational levels?
A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

Data collection

65.　This part of the questionnaire within this subdimension assesses data collection modalities by ascertaining if there is a legal mandate authorizing data collection, assessing the quality of the overall system of survey design, and ascertaining if there is a systematic data review and revision mechanism during data collection, codification and entry. It also determines if the country has adopted standardized procedures for the imputation of missing data, the timeliness and transparency of revisions, the role of the national statistical office in the design of administrative data systems that make data more appropriate for statistical use, and whether an agreement has been concluded with administrative data holders to ensure that they use the data only for statistical purposes.

III.D.1. Is there a clear legal mandate authorizing the producers of official statistics to collect data in your country?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

III.D.2. Please choose the appropriate boxes in order to rate the quality of the overall system of survey design at the national statistical office.

Scoring scheme

Characteristics	Score, if the answer is yes
The target population is well defined	1/6
The sampling frame is well defined	1/6
The implementation parameters are well defined	1/6
The weighting of the sample data is well defined	1/6
There is a high quality report produced for each survey	1/6
Sampling errors are estimated for each survey	1/6

III.D.3. Does the national statistical office conduct systematic control review and correction of data during data collection, when necessary?

Scoring scheme

Answer	Score
Yes, for all data collection operations	1
Yes, for some data collection operations	0.5
Not at all	0

III.D.4. Does the national statistical office conduct systematic control, review, and correction of data during data codification and entry, when necessary?

Scoring scheme

Answer	Score
Yes, for all data entry operations	1
Yes, for some data entry operations	0.5
Not at all	0

III.D.5. Does the national statistical office adopt scientific methodologies in the imputation of missing data?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

III.D.6. Are there standardized procedures for data revisions?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

III.D.7. Is the national statistical office involved in the design or revision of administrative data or records in order to make the data more suitable for statistical use?

Scoring scheme

Answer	Scoring
Yes, for all ministries, departments and agencies	1
Yes, for some ministries, departments and agencies	0.5
No, not at all	0

III.D.8. Are there agreements with the producers of administrative records to facilitate data access and data sharing for statistical purposes?

Scoring scheme

Answer	Scoring
Yes, for all ministries, departments and agencies	1
Yes, for some ministries, departments and agencies	0.5
No such agreements exist	0

Development and implementation of quality assurance tools

66. This subdimension is concerned with assurances as to the quality of any data produced. It seeks to establish whether there is a clear data quality policy, and the number of user guides published by the national statistical office for national statistical system stakeholders on concepts, methods, classifications, and standards, which should be independently defined and in line with professional standards. The subdimension is also concerned with the organization of annual staff training sessions on concepts, methods, classifications, and standardizations in statistics.

III.E.1. Is there a clearly defined data quality policy?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

III.E.2. Is the data quality policy publicly available on the statistical office website?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

III.E.3 How many user guides, defined independently and in compliance with professional rules and ethics, are published by the

national statistical office for national statistical system stakeholders regarding:

➢ Concepts

➢ Methods

➢ Standards classification

Provide the number of user guides published by the national statistical office for each category.

The answers to question III.E.3. is for informational purposes only.

III.E.4.a. Does the national statistical office publish information for staff on its website on the following themes?

For each theme, concept and method discussed in questions III.E.4.a.i, III.E.4.a.ii, and III.E.4.a.iii, respectively, a score of 1/6 is given if the answer is yes, and a score of 0 is given if the answer is no.

Scoring scheme

Question	Score	
	Yes	No
III.E.4.a.i.	1/6	0
III.E.4.a.ii.	1/6	0
III.E.4.a.iii.	1/6	0

III.E.4.b. Does the national statistical office regularly assess the use of documents on the following themes?

For questions III.E.4.b.i., III.E.4.b. ii. and III.E.4.b.iii., a score of 1/6 is given if the answer is yes, and a score of 0 is given if the answer is no.

Scoring scheme

Question	Score	
	Yes	No
III.E.4.b.i.	1/6	0
III.E.4.b.ii.	1/6	0
III.E.4.b.iii.	1/6	0

67.   This part of the questionnaire evaluates the frequency of assessments, corrections and updates to the business register and household surveys/individuals frame and determines whether there is a quality certification system for all data generated by the national statistical system that complies with quality principles for sustainability, data sources, accuracy and reliability, continuity, coherence and comparability, timeliness, the integration of African specificities, and awareness raising among data providers on the importance of statistics.

III.E.5. How often does the national statistical office update the business register?

A score of 1 is given if the register is updated regularly, a score of 0.5 is given if it is updated often and a score of 0.25 is given if it is rarely updated and 0 if it is not updated.

Note: Regularly indicates that an update is done more than once per year, often indicates that an update is done once per year, and rarely indicates that updates are done less than once per year.

III.E.6. How often does the national statistical office update the household surveys/individuals frame?
A score of 1 is given if the household surveys/ individuals frame is updated regularly, a score of 0.5 is given if it is updated often and a score of 0.25 is given if it is rarely updated and 0 if it is not updated.

Note: Regularly indicates that an update is done at least once per year, often indicates that an update is done at least once every two to five years, and rarely indicates that an update is done less than once every five years.

III.E.7. Is there a quality certification system for all data generated by the national statistical system that ensures strict compliance with the quality principles enshrined in the Fundamental

Principles of Official Statistics and the African Charter on Statistics?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

Archival, storage and security of data

68. This subdimension examines archiving, storage and data security issues. Survey responders are asked to estimate the quality of archival and storage infrastructure for data, the quality of the hardware devices used to ensure the security and integrity of statistical databases, the quality of the technical devices used to ensure the security and integrity of statistical databases and the quality of the data anonymization system used to keep data secure.

III.F.1. Does the national statistical office have a data archiving and storage policy?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

III.F.2. Does the national statistical office have an IT security policy?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

III.F.3. Are national statistical office staff familiar with the IT security policy?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

III.F.4. Are data security procedures periodically examined?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

III.F.5. Does the national statistical office systematically anonymize microdata before distributing it to researchers and other users?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

Data management

69. This part of the questionnaire investigates data management. Inter alia, the results indicate whether data management is a dedicated activity within the national statistical office, if data managers are appropriately qualified, and whether data management is dependent on the type of software used or is carried out in line with national statistical office strategy. Relevant variables are derived from the proportion, as measured in percentage terms, of financial resources devoted to data management, and how georeferenced data are accessed by users. This subdimension also looks at the existence or otherwise of metadata for all stored data, who is allowed to extract data from the national statistical office server and whether there are clear guidelines on how the general public can access the data.

III.G.1. Does the national statistical office have a data management team/unit dedicated to data management activities?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

III.G.2. Has the data management team acquired appropriate skills in their fields of study?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

III.G.3. What guides the data management process?

A score of 1 is given if it is dependent on the national statistical office strategy, a score of 0.75 is given if it is dependent on software and a score of 0.25 if it is dependent on any other factor.

III.G.4. As a percentage of the total, what proportion of the national statistical office budget is devoted to data management?

This question ascertains the proportion of total financial resources that is devoted to data management.

Score: Percentage share devoted to data management.

III.G.5. Does the public have access to the georeferenced data?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

III.G.6. Are data systematically stored with metadata?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

III.G.7. Who is allowed to extract data from the national statistical office server?

This information provides an indication of level of security of the data.

Scoring scheme

Security status	Score
Only those with permission have access	1
Everybody	0

III.G.8. Are there clear guidelines on how the general public can access national statistical office data?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

Analytical work within the national statistical office

70. This subdimension deals with analytical work carried out within the national statistical office. It explores the status of current national statistical office publications with respect to statistical analysis and determines whether the staff of the national statistical office are able to use innovative data sources for the production of official statistics and whether there is a functional research unit within the statistical office.

III.H.1. Do current national statistical office publications contain analytical content??

Analytical content should include data sets that explore capture the characteristics of populations. Commonly, such analysis applies statistical modelling and uses statistical software to draw conclusions from the data set. Scenarios may also be run to provide projected results that can guide planning and future actions. For further information, see the paragraph on data analysis in the glossary.

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

III.H.2. Are the staff of the national statistical office able to use innovative data sources in the production of official statistics?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

III.H.3. Is there an operational statistics, economics or social research unit within the statistical office?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

5. Module IV - Effective dissemination policy and use of statistics produced for analysis and research

71. Effective dissemination policies and use of the statistics produced for analysis and research is dependent on adherence to international standards for data dissemination, effective communication regarding the activities of national statistical offices and systems with a view to strengthening a statistical culture within the country and fostering favourable perceptions of official statistics, and the dissemination of data in a clear, understandable, practical and appropriate format.

Adherence to international standards for data dissemination

72. This subdimension focuses on data standards initiatives. Responses indicate whether the country has subscribed to the Enhanced General Data Dissemination System and the Special Data Dissemination Standard of IMF. See annex 1(e) for more information.

IV.A.1. Does the country adhere to IMF data dissemination standards?

A score of 1/3 is given for each dissemination standard adhered to and a score of 0 is given if the country does not adhere to any of those standards.

IV.A.2. Does the country adhere to any other international standards for data dissemination?

That may include other recognized standards, such as the African Charter on Statistics.

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

Data dissemination policy, plan or strategy

73. This part of the questionnaire asks about the data dissemination strategies in place, including the existence of an operational plan for the distribution of data produced by the statistical office, the adoption of a publication schedule for data produced by the statistical office, deadlines and delays between the reference dates of publications and the actual dates of release, and, lastly, the nature of the statistical products released.

74. Questions ask respondents to indicate if the national statistical office has a user-friendly website, online databases, a billing policy and whether the national statistical office disseminates metadata upon request in response to user needs.

75. In addition, there are questions regarding public accessibility to data releases and the periodicity of publications.

76. The question asks respondents to indicate if the national statistical office has a policy of correcting results due to material errors in publications using standard statistical practices, or, in cases in which a serious error is made, whether the national statistical office suspends the publication, and whether the national statistical office clearly explains to users the reasons for corrections or suspensions.

77. Respondents are asked to specify if there are strict protocols that apply to external users' access to statistical microdata for research purposes, if the national statistical office regularly assesses the satisfaction and confidence of data users, if current national statistical office publications contain analytical content and whether national statistical

office staff are able to use innovative data sources in the production of official statistics.

IV.B.1.a. Does the national statistical office have a dissemination strategy and policy for its statistical products?

Scoring scheme

Answer	Dissemination policy only	Dissemination strategy only	Both	None
Score	0.25	0.25	0.5	0

IV.B.1.b. If the country has a dissemination strategy and/or dissemination policy in place, choose the appropriate options regarding public accessibility:

Scoring scheme

Answer	Accessible on the Internet only	Access as a hard copy only	Both	Neither on the Internet or as a hard copy
Score	0.35	0.15	0.5	0

IV.B.2.a. Is a release calendar for official statistics issued in advance?

A score of 0.5 is given if the answer is yes and a score of 0 is given if the answer is no.

IV.B.2.b. If yes, choose the appropriate boxes regarding the release calendar:

Scoring scheme

Answer	The calendar is only for the national statistical office	The calendar is for the entire national statistical system	The calendar is accessible on the Internet to the public
Score	0.0625	0.125	012.5

Answer	Punctuality is monitored on the basis of the calendar	If there is a divergence from the preannounced calendar date, the reasons for that divergence are explained to users and a new release date is announced
Score	0.125	0.125

Note: the first two options are mutually exclusive.

IV.B.3. The statistical products released satisfy the following:

Scoring scheme:

Answer	There is equal access for all users	Pre-release access privileges are restricted	Pre-release access privileges are publicly available
Score	0.25	0.15	0.35
Answer	In the case of statistical information leaks, pre-arrangements are revised so as to assure impartiality		None of the above
Score	0.25		0

IV.B.4. In the following table regarding the release of statistical products, list any key national statistical office publications released on a regular basis, and, for each of those publications, provide the announced reference dates and the actual dates of release last year.

Scoring scheme

Question	The announced date of release (A)	Actual date of release date (R)	Timeliness (time lag between the announced reference date and the actual date of release = R-A)	Score
IV.B.4a:	Publication by periodicity: Annual			
			One month or less	0.25
			More than 3 months but less than 6 months	0.125

				More than 6 months	0
IV.B.4.b:	Publication by periodicity: Half-yearly				
				Two months or less	0.25
				More than 45 days but less than 3 months	0.125
				More than 3 months	0
IV.B.4c:	Publication by periodicity: Quarterly				
				One month or less	0.25
				Less than 2 quarters	0.125
				More than 2 quarters	0
IV.B.4d:	Publication by periodicity: Monthly				
				Ten days or less	0.25
				More than 10 days but less than 20 days	0.125
				Twenty days or more	0

IV.B.5.a. Does the national statistical office website have the following features?

Scoring scheme

Feature	Score
Mobile compatibility	01.25
Accessible to all users (including persons with disabilities – please see manual)	0.125
Fast load times	01.25
Regularly updated (website is updated as soon as there is a new release or information to share)	0.125
None of the above	0.0625
No website has been created	0

IV.B.5.b. Do the online national statistical office statistical databases have the following features?

Scoring scheme

Feature	Score
Clarity, permitting the user to reuse the data	0.25
Data are machine readable	0.25
Neither of the above	0.0625
Databases are not available online	0.0625

No website has been created	0

IV.B.6. Does the statistical services policy provided to users by the national statistical office have the following features?

Scoring scheme

Feature	Score
The policy is derived from the country's statistical law/act or other regulatory frameworks	1/3
The policy is publicly available on the national statistical office website	1/3
The policy defines pricing for extra statistical services or statistical products, including publications	1/3
There is no statistical services policy	0

A fee should be charged for extra services in order to cover costs that are not covered in the national statistical office core budget but which are incurred when fulfilling user requests. In addition, fees charged can discourage the misuse and waste of the provided services.[7]

IV.B.7. Is metadata provided as part of dissemination practices?

Scoring scheme

Practice	Score
Metadata systematically accompany all releases	1.0
Metadata accompany some releases	0.5
Metadata can be provided upon request	0.25
Metadata are not provided	0

IV.B.8. The national statistical office engages with users to address their disaggregated data needs in the following way:

---

7 See, for example, the pricing policy of Statistics South Africa: http://adp.ihsn.org/sites/default/files/South%20Arica_SSA_pricing_policy.pdf.

Scoring scheme

Practice	Score
Provides disaggregated data for non-standard categories whenever possible	1
Provides disaggregated data for all feasible standard categories	0.75
Provides disaggregated data for some feasible standard categories	0.5
There is no mechanism for engaging with users on disaggregated data	0

IV.B.9. When errors are discovered in published statistical products, which of the following corrective measures are usually taken? Multiple options may be selected.

Scoring scheme

Corrective measures	Score
Errors are always corrected at the earliest opportunity	0.4
Errors are sometimes corrected at the earliest opportunity	0.2
Corrected data or information are publicized	0.2
Corrected data or information are not publicized	0.1
In the case of very serious errors, the dissemination of those statistics is suspended	0.2
In the case of suspension of the dissemination of those statistics, the reasons for the suspension are publicized	0.2
No corrective measures are taken	0

Note: the third and fourth options are mutually exclusive, but the fifth and sixth are not.

IV.B.10.a. Is there is a policy or procedure in place by which users may access microdata?

Scoring scheme

Answer	Score
Yes, but the policy or procedure is not publicly available on the national statistical office website or similar platform	0.2
Yes, the policy or procedure is publicly available on the national statistical office website or similar platform	0.4
There is no policy or procedure in place	0

Note: the first and second options are mutually exclusive*.*

IV.B.10.b. Which of the following practices are followed when microdata are provided to users?

Scoring scheme

Practice	Score
Microdata are anonymized before being provided to users	0.15
Only the data needed to address the request are provided	0.15
For data provided to external users, all individuals who will participate in the handling of microdata sign a privacy undertaking	0.15
The national statistical office ensures that the party receiving the microdata has the technical infrastructure and an organizational framework for the protection of confidential data, in compliance with the statistics law/act	0.15
None of the above	0
Microdata are not provided to external users outside the national statistical system	0
Microdata are not provided to internal users inside the national statistical system	0

Users may be internal to the national statistical system, including those involved in statistics production, or external, including researchers.[8]

Note: the first four options are not mutually exclusive and may occur together.

IV.B.11.a. Does the national statistical office conduct user satisfaction surveys?

A score of 0.4 is given if the answer is yes and a score of 0 is given if the answer is no

---

8 See article 24.4 of the Generic Law on Official Statistics, pp. 12-13, available at https://unstats.un.org/unsd/nationalaccount/workshops/2016/eecca/Generic_Law_Official_Statistics_EN.pdf.

IV.B.11.b. If the answer to IV.B.11.a is yes, which of the following practices are in place for communicating with users of official statistics?

Scoring scheme

Practice	Score
Effective and regular engagement with users of statistics to promote trust and enhance public perceptions of the value of statistics through activities such as statistic user/producer forums	0.1
Conduct user satisfaction surveys to investigate and document the needs and practices of users of official statistics, including in terms of how official statistics are used and the types of decisions they inform.	0.1
Involve users in the evaluation of statistics	0.1
Consult and inform users before making changes that affect statistics	0.1
Consult users about statistics usage	0.1
Draw up contracts and memoranda of understanding to foster exchanges among statistical institutes and user groups (e.g., memoranda of understanding with universities, researchers and the scientific community)	0.1
No formal consultations with users of official statistics take place	0

IV.B.12. Does the national statistical office take action based on the findings of user satisfaction surveys? Multiple options may be selected.

Scoring scheme

Action	Score
User needs are integrated into the work programme	0.2
The national statistical office publicly disseminates user experiences with the statistical services it provides	0.2
The national statistical office provides details, including in high quality reports, regarding gaps between measured statistical concepts and users' concepts	0.2
The national statistical office informs users regarding changes to data as a result of consultations with users	0.2
Other actions are taken (please specify)	0.2
None of the above	0
No action is taken	0

Note: the first five options are not mutually exclusive and may occur simultaneously.

Communication of the activities of national statistical offices and systems in order to strengthen the country's statistical culture and enhance the status of official statistics.

78.   This subdimension looks at communication activities conducted by the national statistical office and the activities of the national statistical system in order to enhance opinions of statistics among members of the public and decision-makers.

79.   Variables are related to the existence or otherwise of an internal unit within the national statistical office that is responsible for promoting awareness of statistical products, and a communication plan for activities of national statistical offices and systems. The questions aim to establish whether the national statistical office works closely with the media to organize training and communication programmes, whether sessions are organized during the year by the office, how the office works to publicize and/or strengthen public knowledge regarding the interpretation of statistical data, and whether the country celebrates and has representation at African Statistics Day and World Statistics Day.[9]

IV.C.1. Is there an internal unit/department within the national statistical office responsible for communication regarding statistical products?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

---

9 African Statistics Day is celebrated on 18 November and was adopted in May 1990 by the Conference of African Ministers of Finance, Planning, and Economic Development to be celebrated each year in order to "increase public awareness about the important role which statistics play in all aspects of economic, social and cultural life of citizens in African countries and the continent."

IV.C.2.a. Is there a communication plan for the national statistical office that includes national statistical system activities?

A score of 0.5 is given if the answer is yes and a score of 0 is given if the answer is no.

IV.C.2.b. If the answer to IV.C.2.a. is yes, how is communication with other departments within the national statistical system carried out? Multiple options may be selected.

Scoring scheme

Tool	Score
Research reports	0.1
Peer review papers	0.1
Press releases,	0.1
Policy briefs,	0.1
Other tools, please specify	0.1

IV.C.3. Does the national statistical office work closely with the media in their training and communication programmes?

A score of 1 is given if the answer is yes and a score of 0 is given if the answer is no.

IV.C.4. Choose the correct descriptions regarding any sessions organized during the past year by the national statistical office.

IV.C.4.a. Have any measures been taken to increase the number of users who can interpret statistical data?

Scoring scheme

Answer	Score
Yes, for the public	0.1
Yes, for the government	0.1
Yes, for the media	0.1
Yes, for the private sector	0.1
Yes, for civil society and NGOs within the country	0.1
No sessions have been organized during the past year	0

IV.C.4.b. Have any measures been taken to strengthen the capacity of users to interpret statistical data?

Scoring scheme

Answer	Score
Yes, for the public	0.1
Yes, for the government	0.1
Yes, for the media	0.1
Yes, for the private sector	0.1
Yes, for civil society and NGOs within the country	0.1
No sessions have been organized during the past year	0

IV.C.5 Did the country celebrate African Statistics Day and/or World Statistics Day last year?

A score of 0.5 for each event is given if the answer is yes and a score of 0 is given if the answer is no.

IV.C.6. If the answer to one of the questions above is yes, did a government official participate in the celebrations held to commemorate that event?

A score of 0.5 for each event is given if the answer is yes and a score of 0 is given if the answer is no.

Promoting research through the use and analysis of produced data

80. This subcomponent investigates collaboration between the national statistical office and various stakeholders in the area of data analysis and applied research.

IV.D.1. Which of the following practices are implemented by the national statistical office in order to enhance the use of produced statistics? Choose the appropriate option(s).

Scoring scheme

Practice	Score
Guidelines have been formulated to address the misuse and misinterpretation of data	0.25
Training is provided to users on the proper use of statistics	0.25
Open data copyright is offered (data may be used free of charge if the source is cited correctly and provided the rights of any third parties are respected)	0.25
Investigations of user practices are conducted (an official may be assigned to review data usage by the media)	0.25
None of the above	0

IV.D.2 Is there effective collaboration between the national statistical office and the following stakeholders to encourage data use and analysis?

Scoring scheme

Answer	Score	
	Yes	No
Academia	0.25	0
Research centres outside of universities	0.25	0
NGOs and civil society organizations	0.25	0
The private sector	0.25	0

CHAPTER 3

ANNEXES

Annex 1. Regulatory frameworks

Annex 1a. Fundamental principles of official statistics

These principles were adopted pursuant to the General Assembly resolution 68/261 of 29 January 2014:

The General Assembly, recalling recent resolutions of the General Assembly and the Economic and Social Council highlighting the fundamental importance of official statistics for the national and global development agenda, Bearing in mind the critical role of high-quality official statistical information in analysis and informed policy decision-making in support of sustainable development, peace and security, as well as for mutual knowledge and trade among the States and peoples of an increasingly connected world, demanding openness and transparency, Bearing in mind also that the essential trust of the public in the integrity of official statistical systems and confidence in statistics depend to a large extent on respect for the fundamental values and principles that are the basis of any society seeking to understand itself and respect the rights of its members, and in this context that professional independence and accountability of statistical agencies are crucial, Stressing that, in order to be effective, the fundamental values and principles that govern statistical work have to be guaranteed by legal and institutional frameworks and be respected at all political levels and by all stakeholders in national statistical systems, Endorses the Fundamental Principles of Official Statistics set out below, as adopted by the Statistical Commission in 1994 and reaffirmed in 2013, and endorsed by the Economic and Social Council in its resolution 2013/21 of 24 July 2013: **Fundamental Principles of Official Statistics**

Principle 1. Official statistics provide an indispensable element in the information system of a democratic society, serving the Government, the economy and the public with data about the economic, demographic, social and environmental situation. To this end, official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honor citizens' entitlement to public information.

Principle 2. To retain trust in official statistics, the statistical agencies need to decide according to strictly professional considerations, including scientific principles and professional ethics, on the methods and procedures for the collection, processing, storage and presentation of statistical data.

Principle 3. To facilitate a correct interpretation of the data, the statistical agencies are to present information according to scientific standards on the sources, methods, and procedures of the statistics.

Principle 4. The statistical agencies are entitled to comment on erroneous interpretation and misuse of statistics.

Principle 5. Data for statistical purposes may be drawn from all types of sources, be they statistical surveys or administrative records. Statistical agencies are to choose the source with regard to quality, timeliness, costs and the burden on respondents.

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 7. The laws, regulations, and measures under which the statistical systems operate are to be made public.

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 promote the consistency and efficiency of statistical systems at all official levels.

Principle 10. Bilateral and multilateral cooperation in statistics contributes to the improvement of systems of official statistics in all countries.

Annex 1b. Reference Regional Strategic Framework for Statistical Capacity-Building in Africa

The Reference Regional Strategic Framework sets out a coherent framework to build African Statistical capacity, in line with the Marrakech Action Plan for Statistics. The overall objective of the Framework is to improve development outcomes and good governance by strengthening national statistical systems in Africa.

Specific objectives of the Framework include: raising awareness of the role of statistical information in society, increasing users' satisfaction by enhancing the quality and usability of statistical information, promoting greater use of statistical information, and achieving synergy and cost-effectiveness in national statistical systems.

Specific strategies aimed at achieving the objectives include: invigorating statistical advocacy; mainstreaming statistics as a cross-cutting sector in the development process; updating the legal and regulatory framework; assessing and prioritizing user needs; undertaking data development; fostering coordination, collaboration and partnership; enhancing statistical infrastructure; harnessing ICT; developing human capital; improving data analysis; improving data dissemination; and improving the funding and sustainability of statistical activities.

The Strategic Framework provides a set of 79 recommendations in line with the aforementioned objectives and strategies. The recommendations are directed at different stakeholders, including national statistical offices, national statistical councils or boards, line ministries, subregional organizations, regional organizations, and development partners. The Framework provides guidance to countries on how to improve statistics and increase their use in policymaking and decision-making. Key stakeholders in the national statistical systems in Africa should work together and within this framework in order to realize those objectives.

Annex 1c. The African Charter on Statistics

The African Charter on Statistics is intended

➢ To serve as a policy framework for statistics development in Africa, especially the production, management, and dissemination of statistical data and information at national, regional and continental levels;

➢ To serve as an advocacy tool and instrument for statistics development in the continent;

➢ To ensure improved quality and comparability of the statistics required to monitor the economic and social integration process in the continent;

➢ To promote adherence to fundamental principles of production, storage, management, dissemination and use of statistical information in the African continent;

➢ To contribute to enhancing coordination of statistical activities and statistics institutions in Africa, including the coordination of partners' interventions at national, regional and continental levels;

➢ To build institutional capacity of Statistics authorities in Africa thus ensuring their autonomy in operations, while paying

particular attention to the adequacy of human, material and financial resources;

➢ To serve as a reference framework for the exercise of African statistician profession, professional code of ethics and best practices;

➢ To promote a culture of evidence-based policy formulation, monitoring, and evaluation;

➢ To contribute to improved and effective functioning of the African statistics system and experience sharing; and

➢ To ensure that there is no duplication in the implementation of statistics programmes.

Annex 1d. Strategy for the Harmonization of Statistics in Africa, second generation

In July 2009, the Assembly of Heads of State and Government of the African Union, meeting in Sirte, Libya, mandated the African Union Commission, in collaboration with ECA, the African Development Bank (AfDB) and members of the African statistical system, to develop SHaSA.

The updating of the Strategy for the Harmonization of Statistics in Africa was initiated in 2013 to support the implementation of development and integration programmes in Africa and to improve coordination and collaboration among national statistical institutes, regional and continental statistical organizations, and development partners. The Strategy for the Harmonization of Statistics in Africa, second generation, which covers the period 2017-2026, is the outcome of joint efforts exerted by three continental organizations, namely the African Union Commission, ECA, and AfDB, as well as member States and other stakeholders.

The vision of the African Statistical System is, "to generate timely, reliable, and harmonized statistical information, covering all aspects of political, economic, social, and cultural integration for Africa". The vision is anchored on four strategic pillars and will be achieved through the establishment of a strong and operational African statistical system and actualizations of the African statistical renaissance. The four themes are: to produce quality statistics for Africa; to coordinate the production of quality statistics for Africa; to build sustainable institutional capacity in the African statistical system; and to promote a culture of quality decision-making.

Annex 1e. International Monetary Fund Data Standards Initiative

The IMF Data Standards Initiative was launched to enhance member countries' data transparency and to promote the development of sound statistical systems. The need for data standards was highlighted by the financial crises of the mid 1990s, in which information deficiencies were seen to play a role. Under the Data Standards Initiative, the IMF established the Special Data Dissemination Standard (SDDS) in 1996 to provide guidance to countries that have or seek access to capital markets to disseminate key data, so that users and financial market participants in particular, have adequate information to assess the economic situations of individual countries.

In 1997, the IMF introduced the General Data Dissemination System, which was replaced in 2015 with the Enhanced General Data Dissemination System, (e-GDDS) to provide a framework for countries that aim to develop their statistical systems, and within which they can work toward disseminating comprehensive and reliable data that meet SDDS requirements.

In late 2012, the IMF established a third tier of data standards initiatives with the creation of SDDS Plus, which builds on SDDS and guides member countries

on the provision of economic and financial data to the public, which in turn supports domestic and international financial stability.

As part of the Data Standards Initiative, and in support of access by the public to information on countries' dissemination practices, the IMF established an electronic bulletin board, called the Dissemination Standards Bulletin Board (DSBB), which can be accessed on the IMF website. The DSBB posts information that e-GDDS countries make available to the IMF regarding statistical practices and the website offers direct links to the economic, financial, and sociodemographic data that countries disseminate on their national websites. The DSBB also displays information that SDDS subscribers provide to the IMF regarding their dissemination practices and offers direct links to economic and financial data disseminated in accordance with SDDS.

Member countries' subscription to SDDS, participation in e-GDDS, or adherence to SDDS Plus is voluntary. Countries adhering to SDDS Plus must observe certain good dissemination practices, however, and the IMF monitors their performance in that regard. Countries participating in e-GDDS must also develop plans to improve the quality of their statistics. Empirical studies suggest that subscription to SDDS or participation in e-GDDS, may help reduce borrowing costs in international capital markets. A number of e-GDDS countries have graduated to become SDDS subscribers using the e-GDDS process as a stepping stone. In addition, enhanced data transparency and data quality in member countries has proven to be important to the IMF work in surveillance and crisis prevention.

Annex 2.   Statistical Development Indicators Questionnaire

African Centre for Statistics

STATISTICAL DEVELOPMENT INDICATORS

StatDI

Questionnaire

The main purpose of data collection in this questionnaire is to establish a baseline and subsequent data sets for all participating countries. Confidentiality, as per Principle 6 of the Fundamental Principles of Official Statistics, will be strictly adhered to during the pilot study.

The present questionnaire collects information to measure national and regional progress in statistical development. It is comprised of four modules that collect data to help calculate indicators in the areas addressed by the questionnaire. To ensure the provision of high-quality data, it is recommended that reference is made to the accompanying manual, which explains and defines the terms and concepts used in this questionnaire.

The African Centre for Statistics wishes to thank you for your time in completing this questionnaire.

IDENTIFICATION

Country: _____   Year: _____

Name of respondent: _____

Title: _____

Phone number: _____

Email address: _____

Module I: Functional institutional and organizational framework

Legislative and regulatory framework for statistics

I.A.1.a. Existence of a statistics law/act. Specify the status of the statistics law/act that regulates statistical activities in the country	◯ Exists and in force
	◯ Exists but not yet in force
	◯ Drafted but not yet ratified
	◯ Drafted but not yet signed
	◯ No law/act exists
I.A.1.b. Accessibility to the public of the statistics law/act (if it exists). Choose only one option	◯ Publically available only on the website
	◯ Publically available only as a hard copy from the statistical office
	◯ Publically available on both the website and as a hard copy
	◯ Not publically available
I.A.1c. If a law/act exists, does it satisfy the following conditions? Multiple options may be selected	Does it explicitly make reference to the Fundamental Principles of Official Statistics? ◯ Yes ◯ No
	Is there language in the statistics law/act stating that official statistics should be developed, produced and disseminated in line with the provisions of the African Charter on Statistics? ◯ Yes ◯ No
	Does the statistics act include provisions to address the following? ☐ Delineating the national statistics system ☐ Identifying producers of official statistics ☐ Neither of the above
	Does the statistics law/act include provisions that specify the roles of the following official statistics producers within the national statistical system? ☐ The national statistical office ☐ Other producers ☐ Neither of the above
I.A.2. What is the status of ratification of the African Charter on Statistics?	◯ Not yet signed
	◯ Signed
	◯ Ratification instruments submitted to the African Union Commission
	◯ Ratified
I.A.3. Which of the following elements are clearly addressed by provisions in the statistics law/act or policy, or other commitments that ensure statistical confidentiality? Multiple options may be selected	☐ Protection of privacy
	☐ The law/act makes a distinction among various types of "infractions", including, for example, carelessness, improper behaviour, use of confidential information for personal benefit
	Penalties against persons who wilfully breach the statistical confidentiality sanctions exist on an administrative, penal and disciplinary level

B. Framework for organizing and coordinating the national statistical system

I.B.1.a. Is there a national body or institution responsible for the coordination of the national statistical system?	◯ Yes, Please specify (for example, is the body a political or other body and what is its mandate?)
	◯ No if No, skip to question I.B.2.)

I.B.1.b. If the answer to I.B.1.a is yes, how does that national body or institution coordinate the national statistical system? Multiple options may be selected	☐ Regulations, agreements, memoranda of understanding or laws ☐ Supervisory, advisory, or technical committees ☐ Through the implementation of the annual or multi-annual statistics plan or National Strategy for the Development of Statistics ☐ Joint data collection ☐ Pre-approval of survey methodologies ☐ None of the above methods ☐ Others (please specify)
I.B.2. Does the national statistical office satisfy the following conditions?	*I.B.2.a.* Scientific methods and procedures are always decided according to: ○ Entirely professional considerations ○ Partial professional considerations ○ Absence of professional considerations
	I.B.2.b. The budget is managed as follows: ○ Autonomously ○ Semi-autonomously ○ Non-autonomously
	I.B.2.c. Establishing the annual and multi-annual statistical work pro-grammes is done: ○ Autonomously ○ Semi-autonomously ○ Non-autonomously
	I.B.2.d. Establishing the statistical release calendar is done: ○ Autonomously ○ Semi-autonomously ○ Non-autonomously
I.B.3. Are the following elements regarding the status and functional responsibilities of the head of the national statistical office addressed in the law/act?	☐ Appointment
	☐ Dismissal
	☐ Responsibilities of office
	☐ Duration of office
I.B.4. What is status of the national statistical office head or chief executive officer level in the hierarchy of public service of the country?	○ Minister (Level 1)
	○ Permanent Secretary or equivalent (Level 2)
	○ Director General or equivalent (Level 3)
	○ Director or equivalent (Level 4)
	○ Below Director (please specify) (Level 5)
I.B.5. Does the scope of the national statistical office work plan cover the whole national statistical system	○ Yes
	○ No
Module II: Capacity for an efficient statistical system	
A. Funding	
1. II.A.1. What percentage of last year's total budget was allocated for all statistical activities at the national statistical office?	In terms of execution: Share (per cent) _____ In terms of allocation: Share (per cent) _____
II.A.2. Please complete the following table: A denotes allocated and R denotes received	

Provide the actual share (percentage) of the overall national statistical office budget by funding source (last five years, indicating both allocated and received funds).

Funding source	Last year (Y-1)		Y-2		Y-3		Y-4		Y-5	
	A	R	A	R	A	R	A	R	A	R
Government										
National statistical office										
International partners										
Other sources										

II.A.3. Complete the following table: A denotes allocated and R denotes received

Provide the actual share (per cent) of the overall national statistical system budget by funding source (last five years, indicating both allocated and received funds).

Funding sources	Last year (Y-1)		Y-2		Y-3		Y-4		Y-5	
	A	R	A	R	A	R	A	R	A	R
Government										
National statistical office										
International partners										
Other sources										

II.A.4. Provide the percentage share of the last population and housing census budget that was funded by the government's national budget.	Share: __ per cent
II.A.5. Are there delays between the commitment and the release of funds by the government to the national statistical office?	◯ Yes, delays exist and jeopardize statistical activities
	◯ Yes, delays exist but do not jeopardize statistical activities
	◯ No, there are no delays
II.A.6.a. Does the country have an operational statistical development fund?	◯ Yes
	◯ No (Skip to II.A.6.c)

II.A.6.b. If the answer to II.A.6.a is yes, complete the following table:

Provide the percentage share and the total amount of the national fund devoted to statistical development					
	Last year Y-1	Y-2	Y-3	Y-4	Y-5
Total amount of the fund (Local currency)					
Share of national budget (%)					

II.A.6.c. In the table below, provide information regarding the main partners of the national statistical office and the support received from those partners last year

Partners	Type of partners (national public, national private, international, etc.)	Type of support given (training, deputation of experts, equipment)	Value in United States dollars ($)	Percentage of national statistical office budget

B.   Good governance of the statistical office

II.B.1. How are the senior managers of various units within the national statistical office appointed?	○ Based on a clear and transparent recruitment process, including but not limited to, a call for applications
	○ Appointment by the national statistical office governing body
	○ Nominated by a higher authority, such as by the president, prime minister or a minister
	○ Other (please explain) _____
II.B.2. Is there an annual statistical work plan?	○ Yes, covering the entire national statistical system
	○ Yes, but covering only the national statistical office
	No work plan has been developed
II.B.3. Is there a unit or team in charge of the coordination and implementation of statistical activities at the national statistical office?	○ Yes
	○ No
II.B.4. Is there a framework for the evaluation of staff at the national statistical office? Multiple options may be selected.	☐ Yes, and this includes indicators for the promotion of staff
	☐ Yes, this includes indicators for enhancing staff capacity
	☐ Yes, and this provides incentives and supports staff morale
	☐ Yes, and this matches staff with work duties
	☐ None of the above, but other frameworks have been developed
	☐ No framework has been developed

II.B.5 How is the annual report on the implementation of statistical programmes prepared? Multiple options may be selected.	☐ Only the national statistical office prepares the annual report on the implementation of statistical programmes
	☐ The national statistical office prepares the annual report on the implementation of statistical programmes in close coordination with the other producers of official statistics
	☐ The annual report on the implementation of statistical programmes includes the budgetary execution
	☐ The annual implementation report is published on the website
	☐ The annual implementation report is submitted to the statistical council or equivalent governing body for input
	☐ The opinion of the statistical council or equivalent governing body on the annual implementation report is published
	☐ There is no annual report
II.B.6.a. Are the financial statements of the national statistical office audited and certified on an annual basis by an external control body?	○ Yes
	○ No (go to question II.B.7.a)
II.B.6.b. Are the results of auditing and certification of the financial statements made public?	○ Yes
	○ No
II.B.7.a. Is there a national quality assurance framework?	○ Yes, and this is publicly accessible online
	○ Yes, but it is not publicly accessible online
	○ No (go to question II.B.9)
II.B.7.b. Are the following topics covered in the national quality assurance framework? Multiple options may be selected	☐ Managing the statistical system
	☐ Managing the institutional environment
	☐ Managing statistical processes
	☐ Managing statistical outputs
II.B.8. Kindly provide the following information:	
II.B.8.a. Number of ministries with statistics units	Number:
II.B.8.b. Number of ministries with statistics units that regularly produce statistics in coordination with the national statistical office	Number:
II.B.8.c. Number of ministries for which the national statistical office directly produces statistics	Number:
II.B.9. Is there a user-focused official statistics committee, or equivalent, or a consultation mechanism that ensures that the statistical information published is relevant to users?	○ Yes
	○ No
C. Human capital	
II.C.1 Is there a human resources development plan at each of the following levels?	
II.C.1.a. At the country's national statistical system?	○ Yes ○ No
II.C.1.b. Only at the national statistical office?	○ Yes ○ No

II.C.2. Is there a human resources management plan at the national statistical office level that addresses the following issues? Multiple options may be selected.	☐ Appointment of staff
	☐ Promotion of staff
	☐ Mobility of staff
	☐ Dismissal of staff
	☐ There is no human resources management plan at the national statistical office

Kindly complete the following table regarding staff profiles at the national statistical office

Professional categories	Actual number (previous calendar year)		Required number
	Total	With an advanced degree	
II.C.3.a. Statisticians			
II.C.3.b. Demographers			
II.C.3.c. Economists			
II.C.3.d. GIS experts			
II.C.3.e. ITC experts			
II.C.4.a. Others – professional positions			
II.C.4.b. Others – non-professional positions			
Total			

II.C.5. Indicate the number of national statistical office professional staff who left the institution during the previous year and their reasons for leaving:	
	Number:
Training	
Seconded	
Retired	
Resigned	
II.C.6. Provide the number of students enrolled in the field of statistics in the country's universities last year.	Total number _____ Total per 100,000 inhabitants _____
II.C.7. Provide the number of students enrolled in the field of statistics who received a scholarship last year.	Total number _____
II.C.8. Is there a recycling programme for national statistical office staff?	◯ Yes ◯ No
II.C.9.a. Is there a training programme for human capacity-building and development in statistics and/or related fields, including ICT?	◯ Yes, but only for the national statistical office
	◯ Yes, for the entire national statistical system
	◯ No (go to question II.C.10)
II.C.9.b. If a training programme exists, please comment on its implementation. Multiple options may be selected.	☐ Implementation is led by the national statistical office
	☐ Implementation is led by another body
	☐ Implementation process runs smoothly and without financial and technical challenges
	☐ Implementation process faces certain financial and technical challenges
	☐ No programme is being implemented
II.C.10. Does the national statistical office conduct performance assessments of staff?	◯ Yes ◯ No

D. Physical and material infrastructure of the national statistical office

II.D.1. What is the ownership status of the national statistical office operations building?	○ No specific building exists
	○ It is rented
	○ It is owned by the government
	○ It is owned by the national statistical office

II.D.2.. Please tick the boxes that best describe the building that houses the national statistical office headquarters:

II.D.2.a. External appearance:

II.D.2.a.i. No external building flaws	○ Yes ○ No
II.D.2.a.ii. Clean, with decent surroundings such as landscaping	○ Yes ○ No
II.D.2a.iii. Signposts have been installed	○ Yes ○ No
II.D.2.a.iv. External security is provided, including security guards and security checks	○ Yes ○ No
II.D.2.a.v. Parking is available	○ Yes ○ No

II.D.2.b. Internal appearance

II.D.2.b.i. No internal building flaws	○ Yes ○ No
II.D.2.b.ii Clean	○ Yes ○ No
II.D.2.b.iii. Internal security is provided	○ Yes ○ No
II.D.2.b.iv. There is an information desk	○ Yes ○ No
II.D.2.b.v. There is adequate office space, i.e.,15 square metres per employee	○ Yes ○ No

II.D.2.c. Furniture

II.D.2.c.i. There is a meeting room	○ Yes ○ No
II.D.2.c.ii. Office meets minimum furniture requirements, including a chair and a table for each employee	○ Yes ○ No
II.D.2.c.iii. The furniture is in good condition	○ Yes ○ No
II.D.2.c.iv. There is an area equipped with cupboards and shelves that can be used to store and archive questionnaires or for other purposes	○ Yes ○ No
II.D.2.c.v. Printing and copying machines hare provided	○ Yes ○ No

II.D.3. Do staff working areas satisfy the following requirements Multiple options may be selected	☐ A desk is assigned to each professional staff member
	☐ The statistical agency provides a personal computer to each professional staff member
	☐ The statistical agency allocates filing cabinets to each professional staff member
	☐ All staff members have access to a means of communication
	☐ Access to a means of communication, such as the telephone and the Internet, is provided in the working space of each staff member
	☐ None of the above
II.D.4.a. Has fieldwork conducted by the national statistical office been delayed or negatively affected by insufficient transport infrastructure, which has prevented staff from carrying out all scheduled activities at a given time?	○ Yes ○ No
II.D.4.b Are the vehicles assigned for fieldwork by the national statistical office, including those used to conduct census and surveys, in good condition and appropriate for fieldwork activities?	○ Yes ○ No

E. Statistical infrastructure	
II.E.1. Is there a unit/section, or an internal arrangement within the national statistical office, for the development of standards and classifications in the area of statistics?	○ Yes ○ No
II.E.2.a. Is there a unit/section in charge of a statistical register within the national statistical office?	○ Yes ○ No, go to II.E.3
II.E.2.b. If the answer to II.E.2.a is yes, choose the appropriate boxes.	☐ The statistical register has a database ☐ The database is regularly updated ☐ The register is linked to censuses of the population and households ☐ Guidelines on the use of the register have been developed and made available to the public, such as on the national statistical office website
II.E.3. Is there a unit dedicated to statistical methodologies, including sampling and questionnaire design, within the national statistical office?	○ Yes ○ No
II.E.4. Is there a unit in charge of the geographical information system within the national statistical office?	○ Yes ○ No
II.E.5.a. Does the national statistical office use licensed statistical software?	○ Yes ○ No (go to II.E.6)
II.E.5.b. Please specify the type of licensed software	☐ Statistical analysis system ☐ Statistical Package for the Social Sciences ☐ STATA ☐ S Plus (S+) ☐ E views ☐ Other (please specify) _____
II.E.6. What electronic devices does the national statistical office use for data collection during censuses and national surveys? Multiple options may be selected.	☐ Laptop ☐ Tablets ☐ Smart phones or other handheld devices ☐ Other devices (please specify) _____
II.E.7. What type of data collection is usually used in censuses and national surveys? Multiple options may be selected	☐ PAPI (Paper Assisted Personal Interview) ☐ CATI (Computer Assisted Telephone Interview) ☐ CAPI (Computer Assisted Personal interview) ☐ CAWI (Computer Assisted Web interview) ☐ MAPI (Mobile Assisted Personal Interview) ☐ None of the above

F. ICT Infrastructure	
II.F.1. Is there an IT department/unit/division at the national statistical office?	○ Yes ○ No (go to question II.F.4)
II.F.2.a. What is the total number of staff in the IT department? II.F.2.b. How many IT personnel work in the IT department?	Number _____ Number _____
II.F.3. Does the Head of the IT department/unit/division have post-graduate credentials in IT or a statistics-related field? Multiple options may be selected.	☐ Yes, IT skills ☐ Yes, statistical skills ☐ None of the above
II.F.4. How many professionals at the national statistical office have been allocated a personal computer?	Number: _____

II.F.5. Is there a local area network (LAN) within the national statistical office?	○ Yes
	○ No
II.F.6. Does the national statistical office have its own servers?	○ Yes
	○ No
II.F.7. Is there a centralized databank that is managed by the national statistical office?	○ Yes
	○ No
II.F.8. Does the national statistical office use the following geospatial information system tools? Multiple options may be selected.	☐ Geospatial software
	☐ Remote sensing software
	☐ Global Positioning System equipment
	☐ Global Positioning System software
	☐ Geospatial database
	☐ None of the above
II.F.9.a. Is there a national statistical office website that is regularly updated?	○ Yes
	○ No (go to question II.F.10)
II.F.9.b. If the answer to question II.F.9 is yes, estimate the time lag between the release of statistical information by the national statistical office and the availability of this information on the website?	○ More than one month
	○ Within one month
	○ Within one week
	○ Within one day
	○ Website is not updated
II.F.10. Does the IT security policy at the national statistical office satisfy the following? Multiple options may be selected.	☐ Availability (statistical agencies provide internal and external users with access to data as needed)
	☐ Integrity (statistical agencies ensure that survey and processing methods are effective and prevent human or technological manipulation of data)
	☐ Confidentiality (statistical agencies uphold provisions regarding confidentiality and data protection and ensure that data are only used for statistical purposes)
	☐ national statistical office staff are familiar with the IT security policy
	☐ There is no IT security policy at the national statistical office
II.F.11.a. What is the Internet bandwidth at the national statistical office?	/_____/ Mbps
II.F.11.b. What is the maximum Internet download speed for users?	○ At least 10Mb/s
	○ At least 5Mb/s, but less than 10Mb/s
	○ At least 1Mb/s, but less than 5Mb/s
	○ Less than 1 Mb/s
	○ No Internet available
II.F.12. Is there an IT master plan? Please provide details	○ Yes ○ No
II.F.13. Does the national statistical office maintain an integrated database from a range of data sources?	○ Yes ○ No
II.F.14. How many national statistical office professional staff use their own computer (and not one provided by the national statistical office), to undertake national statistical office-related work?	Number _____
G. Active promotion of bilateral and multilateral cooperation for statistical capacity-building	

II.G.1. Which of the following has your national statistical office undertaken as part of its engagement with other national statistical offices in the last five years? Multiple options may be selected.	☐ Official visits of heads of national statistical offices to other national statistical offices (from your national statistical office to another or vice versa)
	☐ Technical meetings to exchange knowledge and best practices
	☐ Publication of newsletters or webpage providing information about current projects
	☐ Hosting of trainees from other statistical institutes
	☐ Other (please specify)
	☐ None of the above
II.G.2 Does the website disseminate information on best practices and lessons learned that can be applied by other national statistical offices?	◯ Yes ◯ No
II.G.3. In which of the following activities has the national statistical office participated in the past year	☐ Meetings of the United Nations Statistical Commission
	☐ Meetings of the Statistical Commission for Africa
	☐ Meetings of the African Union Committee of Directors General of National Statistical Offices
	☐ None of the above
II.G.4.a. How many national statistical office professional staff have participated in training programmes carried out by global and regional statistical agencies in the past year?	Number _____
II.G.4.b. How many national statistical office professional staff have participated in a training programme abroad and received financial support from the government in order to attend that programme?	Number _____
II.G.5. How many national statistical office professional staff are members of professional bodies or networks?	Number _____
II.G.6. Does the national statistical office encourage its staff to become chartered statisticians?	◯ Yes
	◯ No

Module III: Production of relevant statistics that meet international quality standards

A. Evaluation of data requirements

III.A.1. Is there a systematic evaluation of data needs as per a national development plan or equivalent?	◯ Yes, a full evaluation is carried out
	◯ Yes, a partial evaluation is carried out
	◯ No evaluation is carried out
	◯ Other evaluation (please specify)
III.A.2. Is there a systematic evaluation of the data requirements of initiatives to facilitate achievement of the Sustainable Development Goals?	◯ Yes, a full evaluation is carried out
	◯ Yes, a partial evaluation is carried out
	◯ No evaluation is carried out
	◯ Other evaluation (please specify)
III.A.3. Is there a systematic evaluation of the data requirements of initiatives to facilitate implementation of Agenda 2063 of the African Union?	◯ Yes, a full evaluation is carried out
	◯ Yes, a partial evaluation is carried out
	◯ No evaluation is carried out
	◯ Other evaluation (please specify)

III.A.4. Is there a systematic evaluation of the data requirements of various ministries, departments and government agencies?	◯ Yes, a full evaluation is carried out
	◯ Yes, a partial evaluation is carried out
	◯ No evaluation is carried out
	◯ Other evaluation (please specify)
III.A.5.a. Is there a mechanism in place (such as a survey or user-producer forum) to assess the data needs of other key data users, including those in the private sector, civil society or the media?	◯ Yes, a full assessment is carried out
	◯ Yes, a partial assessment is carried out
	◯ No assessment is carried out
	◯ Other assessment (please specify)
III.A.5.b If yes, explain the mechanism used.	………………………………………..

B. Statistical programming	
III.B.1.a. Is there a multi-year strategic plan or National Strategy for the Development of Statistics s in the country?	◯ Yes, it is being implemented
	◯ Yes, but it has expired
	◯ Yes, and it covers the whole national statistical system
	◯ Yes, but only for the national statistical office
	◯ No, there is no multi-year strategic plan or strategy (go to question III.B.2)
III.B.1.b. If the answer to III.B.1.a. is yes, check the following boxes as applicable? Multiple options may be selected.	☐ The strategy or programme design process was participatory and unrestricted and consultations involved multiple stakeholders, such as government ministries and departments, the private sector, civil society and members of academia
	☐ The strategy or programme was officially adopted by the government.
	☐ The strategy or programme is fully funded by the government
	☐ The strategy or programme is partially funded by the government
	☐ The strategy or programme is aligned with the national development plan or equivalent
	☐ The strategy or programme is aligned with the Sustainable Development Goals
III.B.2. Which of the following sectors have adopted a strategic plan for statistics? Multiple options may be selected.	☐ Education
	☐ Health
	☐ Agriculture
	☐ Water supply
	☐ Manufacturing
	☐ Employment/Labour
	☐ Transport
	☐ Environment
	☐ Energy
	☐ Other (please specify)
	☐ None of the above
C. Comprehensiveness, periodicity, and timeliness of data collected through censuses and surveys	

III.C.1. When was the last population and housing census completed?	I__I__I__I__I
III.C.2. When were the three most recent agriculture censuses completed?	I__I__I__I__I I__I__I__I__I I__I__I__I__I
III.C.3. When were the three most recent livestock censuses completed?	I__I__I__I__I I__I__I__I__I I__I__I__I__I
III.C.4. When were three most recent surveys that included income and expenditure modules completed?	I__I__I__I__I I__I__I__I__I I__I__I__I__I
III.C.5. When were the three most recent surveys that included a demography module completed?	I__I__I__I__I I__I__I__I__I I__I__I__I__I
III.C.6. When were the three most recent surveys that included a health module completed?	I__I__I__I__I I__I__I__I__I I__I__I__I__I
III.C.7. When were the three most recent surveys that included an education module completed?	I__I__I__I__I I__I__I__I__I I__I__I__I__I
III.C.8. When were the three most recent surveys that included a governance module completed?	I__I__I__I__I I__I__I__I__I I__I__I__I__I
III.C.9. When were the three most recent surveys that included a module on gender completed?	I__I__I__I__I I__I__I__I__I I__I__I__I__I
III.C.10. When were the three most recent surveys that included a module on children completed?	I__I__I__I__I I__I__I__I__I I__I__I__I__I
III.C.11. When were the three most recent surveys that included a module on violence against women completed?	I__I__I__I__I I__I__I__I__I I__I__I__I__I
III.C.12. When were the three most recent surveys that included a module on the environment completed?	I__I__I__I__I I__I__I__I__I I__I__I__I__I
III.C.13. How many surveys that included a module on employment have been completed in the last ten years?	Number-----------
III.C.14. How many surveys that included a module on the establishment of businesses have been completed in the last 10 years?	Number-----------
III.C.15. As a percentage of total births, what was the birth registration rate in the country last year, as per civil registration figures?	----- %
Indicate the number of births registered as a percentage of all births in the country.	

III.C.16. As a percentage of total deaths, what was the death registration rate in the country last year, as per civil registration figures?	----- %
Indicate the number of deaths registered as a percentage of all deaths in the country.	
III.C.17. Which edition of the IMF Balance of Payments Manual is used in the country?	○ Sixth edition
	○ Fifth edition
	○ Other edition
III.C.18.a. How long is the delay in the provision of statistics to the national statistical office so that it can make annual national accounts indicators available to the public?	○ At most six months
	○ More than six months, but equal to or less than nine months
	○ More than nine months
III.C.18.b. How long is the delay in the provision of statistics to the national statistical office so that it can make quarterly national accounts indicators available to the public?	○ At most two months
	○ More than two months but equal to or less than three months
	○ More than three months
III.C.19. How long is the delay in the provision of statistics to the national statistical office so that it can make quarterly agriculture production indicators available to the public?	○ At most three months
	○ More than three months, but equal or less than six months
	○ More than six months
III.C.20. How often does the national statistical office publically release quarterly industrial production indicators?	○ Regularly
	○ Not on a regular basis
	○ Not released to the public
III.C.21.a. How often does the national statistical office release national consumer price statistics?	○ Monthly
	○ Less often than monthly
	○ Not released to the public
III.C.21.b. How often does the national statistical office release national industrial price statistics?	○ Quarterly
	○ Less often than quarterly
	○ Not released to the public
III.C.21.c. How often does the national statistical office release national consumer service price statistics?	○ Quarterly
	○ Less often than quarterly
	○ Not released to the public
III.C.22. How often does the national statistical office release employment indicators?	○ Quarterly
	○ Every six months
	○ Annually
	○ Every five years
	○ Less often than every five years
	○ Not released to the public
III.C.23. Is there a policy for updating the production of high-frequency indicators, including prices at the national or subnational levels?	○ Yes
	○ No
D. Data collection	
III.D.1. Is there a clear legal mandate authorizing the producers of official statistics to collect data in your country?	○ Yes
	○ No
III.D.2. Please choose the appropriate circles in order to rate the quality of the overall system of survey design at the national statistical office.	
III.D.2.a. The target population is well defined	○ Yes ○ No
III.D.2.b. The sampling frame is well defined	○ Yes ○ No

III.D.2.c. The implementation parameters are well defined	○ Yes ○ No
III.D.2.d. The weighting of the sample data is well defined	○ Yes ○ No
III.D.2.e. There is a high quality report produced for each survey	○ Yes ○ No
III.D.2.f. Sampling errors are estimated for each survey	○ Yes ○ No
III.D.3. Does the national statistical office conduct systematic control, review and correction of data during data collection, when necessary?	○ Yes, for all data collection operations ○ Yes, for some data collection operations ○ Not at all
III.D.4. Does the national statistical office conduct systematic control, review and correction of data during data codification and entry, when necessary?	○ Yes, for all data entry operations ○ Yes, for some data entry operations ○ Not at all
III.D.5. Does the national statistical office adopt scientific methodologies in the imputation of missing data?	○ Yes ○ No
III.D.6. Are there standardized procedures for data revisions?	○ Yes ○ No
III.D.7. Is the national statistical office involved in the design or revision of administrative data or records in order to make the data more suitable for statistical use?	○ Yes, for all ministries, departments and agencies ○ Yes, for some ministries, departments and agencies ○ Not at all
III.D.8. Are there agreements with the producers of administrative records to facilitate data access and data sharing for statistical purposes?	○ Yes, for all ministries, departments and agencies ○ Yes, for some ministries, departments and agencies ○ Not at all
E. Development and implementation of data quality assurance tools	
III.E.1. Is there a clearly defined data quality policy?	○ Yes ○ No (go to question III.E.3)
III.E.2. Is the data quality policy publicly available on the statistical office website?	○ Yes ○ No
III.E.3. How many user guides, defined independently and in compliance with professional rules and ethics, are published by the national statistical office for national statistical system stakeholders regarding:	
Concepts and definitions	Number: _____
Methods	Number: _____
Standards classification	Number: _____
III.E.4.a. Does the national statistical office publish information for staff on its website on the following themes?	
III.E.4.a.i. Concepts and definitions	○ Yes ○ No
III.E.4.a. ii. Methods	○ Yes ○ No
III.E.4.a.iii. Classifications and standards	○ Yes ○ No
III.E.4.b. Does the national statistical office regularly assess the use of documents on the following?	
III.E.4.b.i. Concepts and definitions	○ Yes ○ No
III.E.4.b. ii. Methods	○ Yes ○ No
III.E.4.b.iii. Classifications and standards	○ Yes ○ No

III.E.5. How often does the national statistical office update the business register?	○ Rarely
	○ Often
	○ Regularly
	○ There is no business register

Rarely = update is done less often than once per year
Often = update is done once per year
Regularly = update is done more than once per year

III.E.6. How often does the national statistical office update the household surveys/individuals frame?	○ Rarely
	○ Often
	○ Regularly
	○ The frame is not updated

| III.E.7. Is there a quality certification system for all data generated by the national statistical system that ensures strict compliance with the quality principles enshrined in the Fundamental Principles of Official Statistics and the African Charter on Statistics? | ○ Yes |
| | ○ No |

F. Archival, storage and security of data

III.F.1. Does the national statistical office have a data archiving and storage policy?	○ Yes ○ No
III.F.2. Does the national statistical office have an IT security policy?	○ Yes ○ No (go to question III.F.5)
III.F.3. Are national statistical office staff familiar with the IT security policy?	○ Yes ○ No
III.F.4. Are data security procedures periodically examined?	○ Yes ○ No
III.F.5. Does the national statistical office systematically anonymize microdata before providing it to researchers and other users?	○ Yes ○ No

G. Data management

III.G.1. Does the national statistical office have a data management team/unit dedicated to data management activities?	○ Yes
	○ No
III.G.2. Has the data management team acquired appropriate skills in their fields of study?	○ Yes
	○ No
III.G.3. What guides the data management process?	☐ The software used
	☐ The national statistical office strategy
	☐ Other factors (please provide details)_____
III.G.4 As a percentage of the total, what proportion of the national statistical office budget is devoted to data management?	_____%
III.G.5. Does the public have access to the geo-referenced data?	○ Yes
	○ No
III.G.6. Are data systematically stored with metadata?	○ Yes
	○ No
III.G.7 Who is allowed to extract data from the national statistical office server?	○ Only those with permission have access
	○ Everybody

III.G.8. Are there clear guidelines on how the general public can access national statistical office data?	○ Yes
	○ No
H. Analytical work within the national statistical office	
III.H.1. Do current national statistical office publications contain analytical content?	○ Yes
	○ No
III.H.2. Is the staff of the national statistical office able to use innovative data sources in the production of official statistics?	○ Yes
	○ No
III.H.3. Is there an operational statistics, economics or social research unit within the statistical office?	○ Yes
	○ No

Module IV: Effective dissemination policy and use of statistics produced for analysis and research

A. Adherence to international standards for data dissemination

IV.A.1. Does the country adhere to the following IMF data dissemination standards?	☐ SDDS
	☐ SDDS-Plus
	☐ e-GDDS
	☐ None of the above
IV.A.2. Does the country adhere to any other international standards for data dissemination?	○ Yes ○ No

B. Data dissemination policy, plan or strategy

IV.B.1.a. Does the national statistical office have a dissemination strategy and policy for its statistical products?	○ Dissemination strategy only
	○ Dissemination policy only
	○ Both of the above
	○ Neither of the above (go to question IV.B.2.)
IV.B.1.b. If the country has a dissemination strategy and/or dissemination policy in place, choose the appropriate options regarding public accessibility:	○ Accessible on the Internet only
	○ Accessible in hard copy only
	○ Both
	○ Neither of the above
IV.B.2.a. Is a release calendar for official statistics issued in advance?	○ Yes
	○ No (go to question IV.3)
IV.B.2.b. If the answer to IV.B.a. is yes, choose the appropriate options regarding the release calendar:	☐ The calendar is only for the national statistical office
	☐ The calendar is for the entire national statistical system
	☐ Accessible to the public on the Internet
	☐ Punctuality is monitored on the basis of the calendar
	☐ If there is a divergence from the preannounced calendar date, the reasons for that divergence are explained to users and a new release date is announced

IV.B.3 The statistical products released satisfy the following: Multiple options may be selected.	☐ There is equal access for all users
	☐ Pre-release access privileges are restricted
	☐ Pre-release privileges are publically available
	☐ In the case of statistical information leaks, pre-arrangements are revised so as to assure impartiality
	☐ None of the above

IV.B.4. In the following table regarding the release of statistical products, list any key national statistical office publications released on a regular basis, and, for each of those publications, provide the announced reference dates and the actual dates of release last year

Question	Publication by periodicity	Announced release date	Actual release date	Comment
IV.B.4.a.	Annually			
IV.B.4.b.	Half-yearly			
IV.B.4.c.	Quarterly			
IV.B.4.d.	Monthly			

IV.B.5.a. Does the national statistical office website have the following features? Multiple options may be selected.	☐ Mobile compatibility
	☐ Accessible to all users (including persons with disabilities, please refer to the manual)
	☐ Fast load times
	☐ Regularly updated (website is updated as soon as there is a new release or information to share)
	☐ None of the above
	☐ No website has been created
IV.B.5.b. Do the national statistical office online statistical databases have the following features?101 Multiple options may be selected.	☐ Clarity, permitting the user to reuse the data
	☐ Data are machine readable
	☐ Neither of the above
	☐ Databases are not available online
	☐ No website has been created

---

10 For further information, see www.socialmediatoday.com/content/web-design-11-characteristics-user-friendly-website.

IV.B.6. Does the statistical services policy provided to users by the national statistical office have the following features: 112 Multiple options may be selected.	☐ The policy is derived from the statistical law/act or other regulatory frameworks
	☐ The policy is publicly available on the national statistical office website
	☐ The policy defines pricing for extra statistical services or statistical products, including publications
	☐ There is no statistical services policy
IV.B.7. Is metadata provided as part of dissemination practices? Choose only one option.	◯ Metadata systematically accompany all releases
	◯ Metadata accompany some releases
	◯ Metadata can be provided upon request
	◯ Metadata are not provided
IV.B.8. The national statistical office engages with users to address their disaggregated data needs in the following way: Choose only one option.	◯ Provides disaggregated data for non-standard categories whenever possible
	◯ Provides disaggregated data for all feasible standard categories
	◯ Provides disaggregated data for some feasible standard categories
	◯ There is no mechanism for engaging with users with regard to disaggregated data
IV.B.9. When errors are discovered in published statistical products, which of the following corrective measures are usually taken? Multiple options may be selected.	☐ Errors are always corrected at the earliest opportunity
	☐ Errors are sometimes corrected at the earliest opportunity
	☐ Corrected data or information are publicized
	☐ Corrected data or information are not publicized
	☐ In the case of very serious errors, the dissemination of those statistics is suspended
	☐ In the case of suspension of the dissemination of those statistics, the reasons for the suspension are publicized
	☐ No corrective measures are taken
IV.B.10.a. Is there is a policy or procedure in place by which users may access microdata? Choose only one option.	◯ Yes, but the policy or procedure is not publicly available on the national statistical office website or similar platform
	◯ Yes, the policy or procedure is publicly available on the national statistical office website or similar platform
	◯ There is no policy or procedure in place

---

11 See Fundamental Principles of Official Statistics Guidelines, United Nations General Assembly, 29 January 2014, p. 40.

IV.B.10.b. Which of the following practices are followed when microdata are provided to users?123 Multiple options may be selected.	☐ Microdata are anonymized before being provided to users
	☐ Only the data needed to address the request are provided
	☐ For data provided to external users, all individuals who will participate in the handling of microdata sign a privacy undertaking
	☐ The national statistical office ensures that the party receiving the microdata has the technical infrastructure and an organizational framework for the protection of confidential data, in compliance with the statistics law/act
	☐ None of the above
	☐ Microdata are not provided to external users outside the national statistical system
	☐ Microdata are not provided to internal users inside the national statistical system
IV.B.11.a. Does the national statistical office conduct user satisfaction surveys? Choose only one option.	◯ Yes
	◯ No (go to question IV.C.1)
IV.B.11.b. If the answer to question IV.B.11.a. is yes, which of the following practices are in place for communicating with users of official statistics? Multiple options may be selected.	☐ Effective and regular engagement with users of statistics to promote trust and enhance public perceptions of the value of statistics through activities such as statistic user/producer forums
	☐ Conduct user satisfaction surveys to investigate and document the needs and practices of users of official statistics, including in terms of how official statistics are used and the types of decisions they inform
	☐ Involve users in the evaluation of statistics
	☐ Consult and inform users before making changes that affect statistics
	☐ Consult users about statistics usage
	☐ Draw up contracts and memoranda of understanding to foster exchanges among statistical institutes and user groups (e.g. memoranda of understanding with universities, researchers and the scientific community)
	☐ No formal consultations with users of official statistics take place
IV.B.12 Does the national statistical office take action based on the findings of user satisfaction surveys? Multiple options may be selected.	☐ User needs are integrated into the work programme
	☐ The national statistical office publicly disseminates user experiences with the statistical services it provides
	☐ The national statistical office provides details, including in high quality reports, regarding gaps between measured statistical concepts and users' concepts
	☐ The national statistical office informs users regarding changes to data as a result of consultations with users
	☐ Other actions are taken (please specify)
	☐ None of the above
	☐ No action is taken

---

12 Manual users can be internal to the national statistical system (for official statistics production) or external (researchers). With regard to privacy, see article 24.4, Generic Law on Official Statistics, pp. 12-13. Available at www.unece.org/fileadmin/DAM/stats/publications/2016/ECECESSTAT20163_E.pdf.

C. Communication of the activities of national statistical offices and systems in order to strengthen the country's statistical culture and enhance the status of official statistics	
IV.C.1. Is there an internal unit/department within the national statistical office responsible for communication regarding statistical products?	○ Yes ○ No
IV.C.2.a. Is there a communication plan for the national statistical office that includes national statistical system activities?	○ Yes ○ No (go to question IV.C.3)
IV.C.2.b. If the answer to IV.C.2.a. is yes, how is communication with other departments in the national statistical system carried out? Multiple options may be selected.	☐ Research reports ☐ Peer review papers ☐ Press releases ☐ Policy briefs ☐ Other tool (please specify)
IV.C.3. Does the national statistical office work closely with the media in their training and communication programmes?	○ Yes ○ No
IV.C.4 Please comment regarding the sessions organized during the past year by the national statistical office to enhance the following aspects of users' capacity:	
IV.C.4.a. Have any measures been taken to increase the number of users who can interpret statistical data? Multiple options may be selected.	☐ Yes, for the public ☐ Yes, for the government ☐ Yes, for the media ☐ Yes, for private sector ☐ Yes, for civil society and NGOs within the country ☐ No sessions have been organized during the past year
IV.C.4.b. Have any measures been taken to strengthen the capacity of users to interpret statistical data? Multiple options may be selected.	☐ Yes, for the public ☐ Yes, for the government ☐ Yes, for the media ☐ Yes, for private sector ☐ Yes, for civil society and NGOs within the country ☐ No sessions have been organized during the past year
IV.C.5. Did the country celebrate:	
IV.C.5.a. World Statistics Day last year?	○ Yes ○ No
IV.C.5.b. African Statistics Day last year?	○ Yes ○ No. (If the answer to both IV.C.5.a. and IV.C.5.b. is no, go to submodule D)
IV.C.6. If the answer to one of the questions above is yes, did a government official participate in the celebrations held to commemorate that event?	
IV.C.6.a. World Statistics Day?	○ Yes ○ No
IV.C.6.b. African Statistics Day?	○ Yes ○ No
D. Promoting research through the use and analysis of produced data	

IV.D.1. Which of the following practices are implemented by the national statistical office in order to enhance the use of produced statistics? Multiple options may be selected.	☐ Guidelines have been formulated to address the misuse and mis-interpretation of data
	☐ Training is provided to users on the proper use of statistics
	☐ Open data copyright is offered (data may be used free of charge if the source is cited correctly and provided the rights of any third parties are respected)
	☐ Investigations of user practices are conducted (an official may be assigned to review data usage by the media)
	☐ None of the above
IV.D.2 Is there effective collaboration between the national statistical office and the following stakeholders to encourage data use and analysis?	
IV.D.2.a. Academia	○ Yes ○ No
IV.D.2.b. Research centres outside universities	○ Yes ○ No
IV.D.2.c. NGOs and civil society organizations	○ Yes ○ No
IV.D.2.d. The private sector	○ Yes ○ No
Thank you for your time and cooperation.	

Annex 3. Scoring of variables, subdimensions and dimensions of Statistical Development Indicators in the piloting phase

The following is a summary of the scoring scheme used with the StatDI questionnaire.

	Proposed scoring scheme for Statistical Development Indicators			
Module I: Functional institutional and organizational framework	**Answer (score)**			
A. Legislative and regulatory framework for statistics				
Q (I.A.1.a.)	For information			
Q (I.A.1.b.)	Publically available only on the website (0.35)	Publically available only as a hard copy from the statistical office (0.15)	Publically available on both the website and as a hard copy, (0.5)	Not publically available (0)
Q (I.A.1.c.)	Explicitly makes reference to the Fundamental Principles of Official Statistics Yes (0.125), No (0)		Produced and disseminated official statistics in line with the provisions of the African Charter on Statistics Yes (0.125), No (0)	
	The law/act delineates the national statistical system (0.065)	The law/act identifies all producers of official statistics (0.065)	The law/act both delineates the national statistical system and identifies all producers of official statistics (0.125)	The law/act neither delineates the national statistical system nor identifies all producers of official statistics (0)
	The law/act includes provisions on the role of the national statistical office (0.065)	The law/act includes provisions on the role of other producers of statistics (0.065)	The law/act includes provisions on both the role of the national statistical office and the role of other producers of statistics (0.125)	The law/act neither includes provisions on the role of the national statistical office nor on the role of other producers of statistics (0)
Q (I.A.2.)	Not yet signed (0.25)	Signed (0.5)	Ratification instruments sent to the African Union Commission (0.75)	Ratified (1.0)
Q (I.A.3.)	Protection of privacy (1/3)	The law/act makes a distinction among various types of infractions (1/3)	The law/act includes provisions regarding penalties against persons who wilfully breach statistical confidentiality (1/3)	

B. Framework for organizing and coordinating the national statistical system	
	Answer (score)
Q (I.B.1.a.)	Existence of a national body, Yes (0.5), No (0)

Q (I.B.1.b.)	Method of coordination				
	Regulations, agreements, memoranda of understanding or laws (0.1)	Supervisory, advisory or technical committees (0.1)	Through the implementation of the annual or multi-annual statistics plan of the NSDS (0.1)	Joint data collection (0.1)	Pre-approval of survey methodologies (0.1)
	None of the above methods (0)		Others (0.1) or (0) if the above five methods are satisfied.		

	I.B.2.a Scientific methods and procedures are always decided according to:		
Q (I.B.2.)	Fully professional considerations (1/4)	Partial professional consideration (1/8)	Absence of professional considerations (0)
	I.B.2.b. Managing the budget		
	Autonomously, (1/4)	Semi-autonomously, (1/8)	Non-autonomously (0)
	I.B.2.c. Establishing the annual and multi-annual statistical work programmes		
	Autonomously (1/4)	Semi-autonomously (1/8)	Non-autonomously (0)
	I.B.2.d. Establishing the statistical release calendar		
	Autonomously (1/4)	Semi-autonomously (1/8)	Non-autonomously (0)

Q (I.B.3.)	Appointment (1/4)	Dismissal (1/4)	Responsibilities of office (1/4)	Duration of office (1/4)	
Q (I.B.4.)	Level 1 (1.0)	Level 2 (1.0)	Level 3 (0.75)	Level 4 (0.5)	Level 5 (0)
Q (I.B.5.)	Yes (1.0)	No (0)			

	Proposed scoring scheme for Statistical Development Indicators					
Module II: Capacity for an efficient statistical system	**Answer (score)**					
A. Funding						
Q (II.A.1.)	If the proportion is greater than or equal to 0.15%, the score is 1.0	Otherwise the score is the real allocation (in per cent) divided by 0.15 per cent				
Q (II.A.2.)	For each year the score is the share (%) divided by 5	The final score is obtained by adding the five annual scores together				
Q (II.A.3.)	For informational purposes only					
Q (II.A.4.)	% share					
Q (II.A.5.)	The score is 0 if the answer is: yes, delays exist and jeopardize statistical activities The score is 0.8 f the answer is: yes, delays exist but do not jeopardize statistical activates The score is 1.0 if the answer is: no, there are no delays					
Q (II.A.6.a.)	Yes (0.7), No (0)					
Q (II.A.6.b.)	Steady or increasing share (0.3), other (0)					
Q (II.A.6.c.)	For informational purposes only					
B. Good governance of the statistical office						
Q (II.B.1.)	Based on a clear and transparent recruitment processes (1.0)	Appointment by the national statistical office (0.5)	Nominated by a higher authority (0)			
	Others: score may be 0.75 if the process falls between the first and second categories, or 0.25 if the process falls between the second and third categories.					
Q (II.B.2.)	Yes, covering the entire national statistical system (1.0)	Yes, but covering only the national statistical office (0.75)	No work plan has been developed (0)			
Q (II.B.3.)	Yes (1.0), No (0)					
Q (II.B.4.)	Yes, and this includes guidelines for the promotion of staff (0.25)	Yes, and this includes guidelines for enhancing the capacities of staff (0.25)	Yes, and this provides incentives and supports staff morale (0.25)	Yes, and this matches staff with work duties (0.25)	None of the above, but other frameworks have been developed (0.50)	No framework has been developed (0)

Q (II.B.5.)	Only the national statistical office prepares the annual report (0.2)	national statistical office prepares the annual report in close coordination with other producers of official statistics (0.4)	The annual report includes the budgetary execution (0.2)	The annual report is published on the website (0.1)
	The annual report is to the statistical council or equivalent governing body for input (0.2)		The opinion of the statistical council or equivalent governing body on the annual report is published (0.1)	
	There is no annual report (0)			
Q (II.B.6.a.)	Yes (0.5), No (0)			
Q (II.B.6.b.)	Yes (0.5), No (0)			
Q (II.B.7.a.)	Yes, and this is publicly accessible on-line (0.5)	Yes, but it is not publicly accessible online (0.25)	No (0)	
Q (II.B.7.b.)	Managing the statistical system (0.125)	Managing the institutional environment (0.125)	Managing statistical processes (0.125)	Managing statistical outputs (0.125)
Q (II.B.8.)	Score = (Number of ministries with statistics units that regularly produce statistics in coordination with the national statistical office divided by the number of ministries with statistics units			
Q (II.B.9.)	Yes (1.0), No (0)			
C. Human capital				
Q (II.C.1.a.)	Yes (1.0), No (0)			
Q (II.C.1.b.)	Yes (0.7), No (0)			
Q (II.C.2.)	Appointment of staff (0.25)	Promotion of staff (0.25)	Mobility of staff (0.25)	
	Dismissal of staff (0.25)	There is no human resources management plan at the national statistical office (0)		
Q (II.C.3.)	Score = The actual number of statisticians + demographers + economists divided by the corresponding required totals			
Q (II.C.4.)	Score = The actual number of other professionals + non-professionals divided by the corresponding required totals			
Q (II.C.5.)	The score = (1 minus the percentage share of those who resigned)			
Q (II.C.6.)	For informational purposes only			
Q (II.C.7.)	For informational purposes only			
Q (II.C.8.)	Yes (1.0), No (0)			
Q (II.C.9.a.)	Yes, but only for the national statistical office (0.3)	Yes, for the entire national statistical system (0.5)	No (0)	

	Implementation process				
Q (II.C.9.b.)	is led by the national statistical office (0.2)	is led by another body (0.1)	runs smoothly without financial and technical challenges (0.3)	faces certain financial and technical challenges (0.2)	No programme is being implemented (0.1)
Q (II.C.10.)	Yes (1.0), No (0)				
D. Physical and material infrastructure of the national statistical office					
Q (II.D.1.)	No specific building exists (0)	It is rented (0.5)	It is owned by the government (0.75)	It is owned by the national statistical office (1.0)	
Q (II.D.2.a.)	For each characteristic: Yes (1/15), No (0)				
Q (II.D.2.b.)	For each characteristic: Yes (1/15), No (0)	The final score of Q (II.D.2.) is obtained by adding the scores of Q (II.D.2.a), Q (II.D.2.b) and Q II.D.2.c			
Q (II.D.2.c.)	For each characteristic: Yes (1/15), No (0)				
Q (II.D.3.)	A desk is assigned to each professional staff member (1/5)	The statistical agency provides a personal computer to each professional staff (1/5)			
	The statistical agency allocates filing cabinets to each professional staff member (1/5)	Each staff member has access to a means of communication (1/5)			
	Access to a means of communication, such as the telephone and the Internet is provided in the working space of each staff member (1/5)	None of the above (0)			
Q (II.D.4.a.)	Yes (0), No (0.5)				
Q (II.D.4. b.)	Yes (0.5), No (0)				
E. Statistical infrastructure					
Q (II.E.1.)	Yes (1.0), No (0)				
Q (II.E.2.a.)	Yes (0.5), No (0)				
Q (II.E.2.b.)	The statistical register has a database (1/8)	The database is regularly updated, (1/8)			
	The register is linked to censuses (1/8)				
	Guidance on the use of the register has been made available to the public (1/8)				
Q (II.E.3.)	Yes (1.0), No (0)				
Q (II.E.4.)	Yes (1.0), No (0)				
Q (II.E.5.a.)	Yes (1.0), No (0)				
Q (II.E.5.b.)	For informational purposes only				

Q (II.E.6.)	If a laptop is provided, the score is 0.5. If at least one tablet, smartphones or another handheld device is provided, the score is (1.0). If at least one other handheld device is provided, the score is (1.0) If no handheld device is provided, the score is (0). Total score should not exceed a maximum of (1.0)		
Q (II.E.7.)	PAPI (1.0) If at least one or more of CATI, CAPI, CAWI, is used, the score is (0.5), MAPI or others, the score is (0). Total score should not exceed a maximum of (1.0)		
F. ICT Infrastructure			
Q (II.F.1.)	Yes (1.0), No (0)		
Q (II.F.2.a.) and Q (II.F.2.b.)	For informational purposes only		
Q (II.F.3.)	Yes, IT skills (0.5) Yes, statistical skills (0.5) Neither of the above (0)		
Q (II.F.4.)	The score is the ratio obtained by dividing the number of professionals in the national statistical office to whom a functional computer has been allocated by the total number of professionals employed last year.		
Q (II.F.5.)	Yes (1.0), No (0)		
Q (II.F.6.)	Yes (1.0), No (0)		
Q (II.F.7.)	Yes (1.0), No (0)		
Q (II.F.8.)	Geospatial software (0.2)	Remote sensing software (0.2)	
	Global Positioning System equipment (0.2)	Global Positioning System software (0.2)	
	Geospatial database (0.2)	None of the above (0.2)	
Q (II.F.9.a.)	Yes (0.5), No (0)		
Q (II.F.9.b.)	More than one month (0.0625)	Within one month (0.125)	Within one week (0.25)
	Within one day (0.5)	Website is not updated (0)	
Q (II.F.10.)	Availability (0.25)	Integrity (0.25)	Confidentiality (0.25)
	National statistical office staff are familiar with the IT-security policy (0.25)	There is no IT security policy at the national statistical office (0)	
Q (II.F.11.a.)	For informational purposes only		
Q (II.F.11.b.)	At least 10Mb/s (1.0)	At least 5Mb/s, but less than 10Mb/s (0.5)	
	At least 1Mb/s, but less than 5Mb/s (0.25)		
	Less than 1 Mb/s, (0.125)	No internet available (0)	
Q (II.F.12.)	Yes (1.0), No (0)		
Q (II.F.13.)	Yes (1.0), No (0)		
Q (II.F.14.)	Score = 1 minus (number of those who own their own computers divided by the total number of staff)		
G. Active promotion of bilateral and multilateral cooperation for statistical capacity-building			

Q (II.G.1.)	Official visits (0.25)	Technical meetings (0.25)	Total score should not exceed (1.0)
	Publication of newsletters or a web-page (0.25)	Hosting of trainees (0.25)	
	Other (0.25)	None of the above (0)	
Q (II.G.2.)	Yes (1.0), No (0)		
Q (II.G.3.)	The United Nations Statistical Commission (1/3)	Meetings of the Statistical Commission for Africa (1/3)	
	Meetings of the African Union Committee of Directors General of National Statistical Offices (1/3)		None (0)
Q (II.G.4.a.)	If one or more staff members have participated, the score is (0.5) If no staff members have participated, the score is (0)		
Q (II.G.4.b.)	If one or more staff members have participated, the score is (0.5) If no staff members have participated, the score is (0)		
Q (II.G.5.)	For informational purposes only		
Q (II.G.6.)	For informational purposes only		

	Proposed scoring scheme for Statistical Development Indicators		
Module III: Production of relevant statistics that meet international quality standards	**Answer (Score)**		
A. Evaluation of data requirements			
Q (III.A.1.)	Yes, full evaluation, (1.0)	Yes, partial evaluation (0.5)	No evaluation (0)
	Other evaluation (0.25)		
Q (III.A.2.)	Yes, full evaluation, (1.0)	Yes, partial evaluation (0.5)	No evaluation (0)
	Other evaluation (0.25)		
Q (III.A.3.)	Yes, full evaluation, (1.0)	Yes, partial evaluation (0.5)	No evaluation (0)
	Other evaluation (0.25)		
Q (III.A.4.)	Yes, full evaluation, (1.0)	Yes, partial evaluation (0.5)	No evaluation (0)
	Other evaluation (0.25)		
Q (III.A.5.a.)	Yes, full assessment, (1.0)	Yes, partial assessment (0.5)	No assessment (0)
	Other assessment, (0.25)		
Q (III.A.5.b.)	For informational purposes only		
B. Statistical programming			
Q (III.B.1.a.)	Yes, it is being implemented (0.2)	Yes, but it has expired (0.1)	
	Yes, and it covers the entire national statistical system (0.2)	Yes, but only for the national statistical office (0.1)	No (0)

Q (III.B.1.b.)	The strategy was participatory (0.1)		The strategy was officially adopted (0.1)
	The strategy is fully funded by the government (0.1)		The strategy is partially funded by the government (0.1)
	The strategy or programme is aligned with the national development plan or equivalent (0.1)		The strategy or programme is aligned with the Sustainable Development Goals (0.1)

Q (III.B.2.)	Education (0.1)	Health (0.1)	Agriculture (0.1)
	Water supply (0.1)	Manufacturing (0.1)	Employment/Labour (0.1)
	Transport (0.1)	Environment (0.1)	Energy (0.1)
	Others (0.1) for each sector		None of the above (0)

C. Comprehensiveness, periodicity, and timeliness of data collected through censuses and surveys

Q (III.C.1.)	Score (1.0) if the latest census was completed within the last 10 years, and (0) if otherwise
Q (III.C.2.)	Score (1.0) if the latest census was completed within the last 5 years, and (0) if otherwise
Q (III.C.3.)	Score (1.0) if the latest census was completed within the last 5 years, and (0) if otherwise
Q (III.C.4.)	Score (1.0) if the latest survey was completed within the last 5 years, and (0) if otherwise
Q (III.C.5.)	Score (1.0) if the latest survey was completed within the last 5 years, and (0) if otherwise
Q (III.C.6.)	Score (1.0) if the latest survey was completed within the last 5 years, and (0) if otherwise
Q (III.C.7.)	Score (1.0) if the latest survey was completed within the last 5 years, and (0) if otherwise
Q (III.C.8.)	Score (1.0) if the latest survey was completed within the last 5 years, and (0) if otherwise
Q (III.C.9.)	Score (1.0) if the latest survey was completed within the last 5 years, and (0) if otherwise
Q (III.C.10.)	Score (1.0) if the latest survey was completed within the last 5 years, and (0) if otherwise
Q (III.C.11.)	Score (1.0) if the latest survey was completed within the last 5 years, and (0) if otherwise
Q (III.C.12.)	Score (1.0) if the latest survey was completed within the last 5 years, and (0) if otherwise
Q (III.C.13.)	Score (1.0) if one or more has been completed, and (0) if no survey has been conducted
Q (III.C.14.)	Score (1.0) if one or more has been completed, and (0) if no survey has been conducted
Q (III.C.15.)	Score is (1.0) if the percentage is ≥70%; (0.75) if the percentage is <70% and ≥ 50%; and (0) if otherwise
Q (III.C.16.)	Score is (1.0) if the percentage is ≥70%; (0.75) if the percentage is <70% and ≥ 50%; and (0) if otherwise
Q (III.C.17.)	Sixth edition (1.0), fifth edition (0.5), and any other edition (0.25).

Q (III.C.18.a.)	At most 6 months (0.5)	More than 6 months but equal to or less than 9 months (0.25)	More than 9 months (0)
Q (III.C.18.b.)	At most 2 months (0.5)	More than 2 months but equal to or less than 3 months (0.25)	More than 3 months (0)
Q (III.C.19.)	At most 3 months (1.0)	More than 3 months but equal to or less than 6 months (0.5)	More than 6 months (0)
Q (III.C.20.)	Regularly (1.0)	Not on a regular basis (0.5)	Not released to the public (0)
Q (III.C.21.a.)	Monthly (1/3)	Less often than monthly (1/6)	Not released to the public (0)
Q (III.C.21.b.)	Quarterly (1/3)	Less often than quarterly (1/6)	Not released to the public (0)
Q (III.C.21.c.)	Quarterly (1/3)	Less often than quarterly (1/6)	Not released to the public (0)

Q (III.C.22.)	Quarterly (1.0)	Every 6 months (0.75)	Annually (0.5)	Not released to the public (0)
	Every 5 years (0.25)	Less often than every 5 years (0.125)		

Q (III.C.23.)	Yes (1.0), No (0)				
D. Data collection					
Q (III.D.1.)	Yes (1.0), No (0)				
Q (III.D.2.)	Yes (1/6), No (0) (for each of the six characteristics)				
Q (III.D.3.)	Yes, for all data collection operations, (1.0)	Yes, for some data collection operations (0.5)	Not at all (0)		
Q (III.D.4.)	Yes, for all data collection operations (1.0)	Yes, for some data collection operations (0.5)	Not at all (0)		
Q (III.D.5.)	Yes (1.0), No (0)				
Q (III.D.6.)	Yes (1.0), No (0)				
Q (III.D.7.)	Yes, for all ministries, departments and agencies (1.0)	Yes, for some ministries, departments and agencies (0.5)	Not at all (0)		
Q (III.D.8.)	Yes, for all ministries, departments and agencies (1.0)	Yes, for some ministries, departments and agencies (0.5)	Not at all (0)		
E. Development and implementation of quality assurance tools					
Q (III.E.1.)	Yes (1.0), No (0)				
Q (III.E.2.)	Yes (1.0), No (0)				
Q (III.E.3.)	For informational purposes only				
Q (III.E.4.a.)	Yes (1/6) and No (0) (for each of the three categories)				
Q (III.E.4.b.)	Yes (1/6) and No (0) (for each of the three categories)				
Q (III.E.5.)	Regularly (1.0)	Often (0.5)	Rarely (0.25)	Not updated (0)	
Q (III.E.6.	Regularly (1.0)	Often (0.5)	Rarely (0.25)	Not updated (0)	
Q (III.E.7.)	Yes (1.0), No (0)				
F. Archival, storage and security of data					
Q (III.F.1.)	Yes (1.0), No (0)				
Q (III.F.2.)	Yes (1.0), No (0)				
Q (III.F3.)	Yes (1.0), No (0)				
Q (III.F.4.)	Yes (1.0), No (0)				
Q (III.F.5.)	Yes (1.0), No (0)				
G. Data management and accessibility					
Q (III.G.1.)	Yes (1.0), No (0)				
Q (III.G.2.)	Yes (1.0), No (0)				
Q (III.G.3.)	National statistical office Strategy (1.0)	Software used (0.75)	Other factors (0.25)		
Q (III.G.4.)	Proportion (%)				
Q (III.G.5.)	Yes (1.0), No (0)				
Q (III.G.6.)	Yes (1.0), No (0)				
Q (III.G.7.)	Only those with permission (1.0), Everyone (0)				
Q (III.G.8.)	Yes (1.0), No (0)				

H. Analytical work within the national statistical office				
Q (III.H.1.)	Yes (1.0), No (0)			
Q (III.H.2.)	Yes (1.0), No (0)			
Q (III.H.3.)	Yes (1.0), No (0)			

	Proposed scoring scheme for Statistical Development Indicators			
Module IV: Effective dissemination policy and use of statistics produced for analysis and research	**Answer (score)**			
A. Subscription to international standards for data dissemination				
Q (IV.A.1.)	SDDS (1/3)	SDDS-Plus (1/3)	e-GDDS (1/3)	None (0)	
Q (IV.A.2.)	Yes (1.0), No (0)			
B. Data dissemination policy, plan or strategy				
Q (IV.B.1.a.)	Dissemination strategy only (0.25)	Dissemination policy only (0.25)	Both (0.5)	Neither (0)
Q (IV.B.1.b.)	Accessible on the Internet only (0.35)	Accessible in hard copy only (0.15)	Both of the above (0.5)	Neither of the above (0)
Q (IV.B.2.a.)	Yes (0.5), No (0)			
Q (IV.B.2.b.)	The calendar is only for the national statistical office (0.0625)	The calendar is for the entire national statistical system (0.125)	Accessible to the public on the Internet (0.125)	
	Punctuality is monitored on the basis of the calendar (0.125)	If there is a divergence from the pre-announced calendar date, the reasons for that divergence are explained to users and a new release date is announced (0.125)		
Q (IV.B.3.)	Equal access provided to all users (0.25)	Pre-release access privileges are restricted (0.15)	Pre-release access privileges are publicly available (0.35)	
	In the case of statistical information leaks, pre-arrangements are revised so as to assure impartiality (0.25)	None of the above (0)		
Q (IV.B.4.a.)	Annual publications			
	Released monthly or less (0.250)	Released more than every 3 months and less than six months (0.125)	Released more than every six months (0)	
Q (IV.B.4.b)	Half yearly publications			
	Released every two months or less (0.250)	More than 45 days and less than three months (0.125)	Released more than every three months (0)	
Q (IV.B.4.c)	Quarterly publications			

	Released every one month or less (0.250)	Released less than every two quarters (0.125)	Released less often than more than every two quarters (0)
Q (IV.B.4.d)	Monthly publications		
	Released every 10 days or less (0.250)	More than 10 days and less than 20 days (0.125)	Released every 20 days or more (0)
Q (IV.B.5.a.)	Mobile compatibility (0.125)	Accessible to all users (including people with disabilities) (0.125)	
	Fast load times (0.125)	Regularly updated (the website is updated as soon as there is a new release or information to share) (0.125)	
	None of the above (0.0625)	No website has been created (0)	
Q (IV.B.5.b.)	Clarity, permitting the user to reuse the data (0.25)	Data are machine-readable (0.25)	Neither of the above (0.0625)
	Databases are not available online (0.0625)	No website has been created (0)	
Q (IV.B.6.)	The policy is derived from the statistical law/act or other regulatory frameworks, (1/3)		The policy is publicly available on the national statistical office website (1/3)
	The policy defines pricing for extra statistical services or statistical products, including publications (1/3)		There is no statistical services policy (0)
Q (IV.B.7.)	Metadata systematically accompany all releases (1.0)	Metadata accompany some releases (0.5)	
	Metadata can be provided upon request (0.25)	Metadata are not provided (0)	
Q (IV.B.8.)	Provides disaggregated data for non-standard categories whenever possible (1.0)	Provides disaggregated data provided for all feasible standard categories, (0.75)	
	Provides disaggregated data provided for some of feasible standard categories, (0.5)	There is no mechanism for engaging with users with regard to disaggregated data (0)	
Q (IV.B.9.)	Errors are always corrected at the earliest opportunity (0.4)	Errors are sometimes corrected at the earliest opportunity (0.2)	
	Corrected data or information are publicized (0.2)	Corrected data or information are not publicized (0.1)	
	In the case of very serious errors, the dissemination of those statistics is suspended (0.2)		
	In the case of suspension of the dissemination of those statistics, the reasons for the suspension are publicized (0.2)		
	No corrective measures are taken (0)		
Q (IV.B.10.a.)	Yes, but the policy or procedure is not publicly available on the national statistical office website or similar platform (0.2)		
	Yes, the policy or procedure is publicly available on the national statistical office website or similar platform (0.4)	There is no policy or procedure in place (0)	

Q (IV.B.10.b.)	Microdata are anonymized before being provided to users (0.15)	Only the data needed to address the request are provided (0.15)	
	For data granted to external users, all persons who will handle the microdata sign a contract to guarantee privacy (0.15)		
	The national statistical office ensures that the party receiving the microdata has the technical infrastructure and an organizational framework for the protection of confidential data, in compliance with the statistics law/act (0.15)		
	None of the above (0)	Microdata are not provided to external users outside the national statistical system (0)	Microdata are not provided to internal users inside the national statistical system (0)
Q (IV.B.11.a.)	Yes (0.4), No (0)		
Q (IV.B.11.b.)	Effective and regular engagement with users of statistics to promote trust and enhance public perceptions of the value of statistics through activities such as statistics user/producer forums (0.1)		
	Conduct user satisfaction surveys to investigate and document the needs of users of official statistics, including in terms of how official statistics are used and the types of decision they inform (0.1)		
	Involve users in the evaluation of statistics (0.1)	Consult and inform users before making changes that affect statistics (0.1)	Consult users about statistics usage (0.1)
	Draw up contracts and memoranda of understanding to foster exchanges among statistical institutes and user groups (0.1)		
	No formal consultations with users of official statistics take place (0)		
Q (IV.B.12.)	User needs are integrated into the work programme (0.2)		
	The national statistical office publicly disseminates user experiences with the statistical services it provides (0.2)		
	The national statistical office provides details, including in high quality reports, regarding gaps between measured statistical concepts and users' concepts (0.2)		
	The national statistical office informs users regarding changes to data as a result of consultations with users (0.2)		
	Other actions are taken (0.2)	None of the above (0)	No action is taken (0)
C. Communication of the activities of national statistical offices and systems in order to strengthen the country's statistical culture and enhance the status of official statistics			
Q (IV.C.1.)	Yes (1.0), No (0)		
Q (IV.C.2.a.)	Yes (0.5), No (0)		
Q (IV.C.2.b.)	Research reports (0.1)	Peer review papers (0.1)	Press releases (0.1)
	Policy briefs (0.1)	Other tool (0.1)	
Q (IV.C.3.)	Yes (1.0), No (0)		
Q (IV.C.4.a.)	Yes, for the public (0.1)	Yes, for the government (0.1)	Yes, for the media (0.1)
	Yes, for the private sector (0.1)	Yes, for civil society and NGOs within the country (0.1)	No sessions have been organized during the past year (0)

Q (IV.C.4.b.)	Yes, for the public (0.1)		Yes, for the government (0.1)		Yes, for the media (0.1)
	Yes, for the private sector (0.1)	Yes, for civil society and NGOs within the country (0.1)	No sessions have been organized during last year (0)		
Q (IV.C.5.a.)	Yes (0.5), No (0)				
Q (IV.C.5b.)	Yes (0.5), No (0)				
Q (IV.C.6.a.)	Yes (0.5), No (0)				
Q (IV.C.6.b.)	Yes (0.5), No (0)				
D. Promoting research through the use and analysis of produced data					
Q (IV.D.1.)	Guidelines have been formulated to address the misuse and misinterpretation of data (0.25)				
	Training is provided to users on the proper use of statistics (0.25)	Open data copyright is offered (0.25)		Investigations of user practices are conducted (0.25)	None of the above (0)
Q (IV.D.2.a.)	Yes (0.25), No (0)				
Q (IV.D.2.b.)	Yes (0.25), No (0)				
Q (IV.D.2.c.)	Yes (0.25), No (0)				
Q (IV.D.2.d)	Yes (0.25), No (0)				

Annex 4. Statistical methods

1.   *Data processing and editing*

Processing

Data processing is the collection of data to produce meaningful results. Results are achieved by doing calculations on a given set of data to extract required information in an appropriate form such as through diagrams, reports, tables, manipulating input data manually or with an application programme.

Imputation of missing data

Imputation of missing data is quite appealing, since, as in any survey, it will not be possible to obtain a 100 per cent response rate from all countries in an StatDI survey. Thus, as recommended by the OECD, consideration should be given to different approaches for imputing missing values. In addition, extreme values should be examined as they can become unintended benchmarks. Providing a complete dataset through either single or multiple imputations requires:

> ➤   estimating missing values;

> ➤   measuring the reliability of each imputed value so as to assess the impact of the imputation on the composite indicator results;

> ➤   discussing the presence of outliers in the dataset.

Various types of missingness[13]

There are three main patterns of missing values: (a) Missing completely at random; (b) Missing at random; and (c) Missing not at random.

(a) Missing completely at random

For example, missing values in variable, whether or not a country or its national statistical office subscribed to the General Data Dissemination System (GDDS) will follow the "missing completely at random" type if (a) on average the value of the national statistical offices do not report their answer assumed to be the same as those of national statistical offices that reported their answer;; and if (b) each of the other variables in the data set would have to be the same, on average, for the national statistical offices that did not report their answer to this variable and the national statistical offices that did report their answers, in this case. (Yes or NO).

(b) Missing at random

Missingness does not depend on the variable in question, but is conditional on other variables in the dataset. For example, the missing values in rating the quality of the building housing the national statistical office would be "missing at random" if the probability of missing data on this rating variable depended on whether or not the national statistical office had its own building but, within each category of owned-building, the probability of the missing rating is unrelated to the value of rating. Another example of a "missing at random" pattern may be one missing by design, e.g. if survey question 1 is answered yes, then survey question 2 is not to be answered.

(c) Missing not at random

Missingness can be fully accounted by the variables themselves. For example, high-income households are less likely to report their income.

There are three general methods for normalizing missing data: case deletion, single imputation, and multiple imputations.

---

13   Donald B. Rubin, Multiple Imputation for Nonresponse in Surveys, New York: Wiley, 1987.

Case deletion: Also known as complete case analysis, this method simply omits the missing records from the analysis. However, this approach ignores possible systematic differences between complete and incomplete samples and produces unbiased estimates only if deleted records are a random subsample of the original sample. Furthermore, standard errors will generally be larger in a reduced sample as less information is used.

Single imputation: As the name of this method suggests, a single value will be computed to replace the missing values. There are two types of single imputation defined as follows:

Implicit modelling focuses on an algorithm, with implicit underlying assumptions that need to be verified in terms of whether they are reasonable and fit for purpose. Implicit modelling includes hot-deck imputation, substitution, and cold deck imputation. For the definitions of those techniques within single imputation, refer to the glossary.

Explicit modelling. In this case, single imputation will generate a predictive distribution based on a formal statistical model where the assumptions are explicitly stated. Examples of explicit modelling of single imputation include unconditional imputation, regression imputation and expectation maximization. The following box defines and describes these techniques.

Box 2: Handling of missing data

Explicit modeling of single imputation

➢ **Unconditional mean/median/mode imputation**

The sample mean (median, mode) of the recorded values for the given individual indicator replaces the missing values.

Let D_{ij} (i=1,2,3,…,S;j=1,2,3,…M) denote the **(ith)** subdimension within the **(jth)** dimension in StatDI,

Let X_k (k=1,2,3,…,Q) denote the random variable associated with the **kth** indicator, and let x_{kc} denote the value of this indicator for country **c; (c=1,2,3,…,T)** Let t_k be the number of reported/recorded or non-missing values (number of countries) on $X_{k'}$ so that there are $-t_k$ missing values in that subdimension. The unconditional mean is then given by:

$$\overline{x_k} = \frac{1}{t_k} \sum_{c=1}^{t_k} x_{kc} \quad \text{………………. (A.1)}$$

Similarly, the median and the mode of the distribution can be calculated on the available sample and substitute missing values. However, by "filling in" blank spaces with the sample mean, the imputed value becomes a biased estimator of the population mean, even in the case of "missing completely at random" mechanisms, and the sample variance underestimates true variance, thus underestimating the uncertainty in the composite due to the imputation.

➢ **Regression imputation**

Missing values are substituted by the predicted values obtained from regression. The dependent variable of the regression is the individual indicator hosting the missing value, and the regressor(s) is (are) the individual indicator(s), showing a strong relationship with the dependent variable, i.e. usually a high degree of correlation.

Suppose a set of **H<Q** fully observed individual indicators $(X_1, X_2, \ldots X_H)$. Let a specific indicator X_k be reported by/recorded for only R country so that it is missing for **T-R** countries. Regression imputation computes the regression of X_k on $(X_1, X_2, \ldots X_H)$ using the **R** complete observations, and imputes the missing values as a prediction from the regression:

$$\widehat{x_{ik}} = \widehat{\beta}_0 + \sum_{j=1}^{H} \widehat{\beta}_j x_{ij} + \varepsilon_i \quad \ldots\ldots\ldots (i = 1,2,3, \ldots\ldots.. T - R) \ldots\ldots\ldots (A.2)$$

Where ε_i is a random variable $N(0, \widehat{\sigma^2})$ and $\widehat{\sigma^2}$ is the residual variance from the regression of X_k on $(X_1, X_2, \ldots X_H)$, using the **R** complete observations.

The strategy to define the 'best' regression is a two-step procedure. First, all different subsets of predictors are adopted in a multiple regression manner. Then the best subset(s) is determined using the following criteria: the value of the explanation rate R^2; the value of the residual mean square RMS;

➢ **Expectation Maximization (EM) imputation**

This model focuses on the interdependence between model parameters and the missing values. The missing values are substituted by estimates obtained through an iterative process. First, the missing values are predicted based on initial estimates of the model parameter values. These predictions are then used to update the parameter values, and the process is repeated. The sequence of parameters converges to maximum-likelihood estimates, and the time to convergence depends on the proportion of missing data and the flatness of the likelihood function. If simplicity is its main appeal, an important limitation of the single imputation method is its systematic underestimation of the variance of the estimates (with some exceptions for the EM method, where the bias depends on the algorithm used to estimate the variance). Therefore, this method does not fully assess the implications of imputation or the robustness of the composite index derived from the imputed data set.

Note: Although an advantage of single imputation is its simplicity, which may address some missing values, its main limitation is its systematic underestimation of the variance of the estimates. There are some exceptions to this in the EM method, where the bias depends on the algorithm used to estimate the variance. Therefore, single imputation does not fully assess the implications of imputation or the robustness of the composite index derived from the imputed data set.

Multiple Imputation: Generally, multiple imputation procedure is built on three steps:

(a) Deciding to impute

At the imputation stage, the main question that arises is whether to impute data or to perform analysis on cases with complete data only. This question is important to answer since imputation processes are very demanding in terms of implementation time. Another issue to solve at this stage is the theoretical foundation of the approach adopted.

It is also important to consider the power issue. The decision to impute data may be guided by the fact that one wishes to avoid losing data observations. The approach adopted thus leads to use of the information available in those observations that contain missing values, and this can lead to less confidence in the overall results.

Thus, the decision to impute should be made after running tests to identify the specific pattern of missing values affecting the data set.

(b) Creating imputation models

Once imputation is chosen as a course of action, it becomes important to create models to support the process.

In theory, an imputation model estimates the distribution of all the variables it contains. But there are some approaches, including multiple imputation using chained equations for example, that break down the problem into a series of estimates of one variable in the model. An issue with this type of technique is that a series of models that includes the distributions of individual variables do not necessarily add up to one consistent outcome for the joint distribution. In the process of selecting imputation models, there should be a focus on the choice of variables and methods, be they continuous non-normal variables, transformations, bounded variables, or the inclusion of non-linear terms and interaction terms. Ideally, better results are obtained when real data are gathered rather than relying on imputation methods.

Basically, if imputation is done N times, it creates N number of "complete" datasets. The parameters of interest are estimated on each data set, together with their standard errors. Average estimates are combined using the aforementioned N sets and the "between-and-within" imputation variance is calculated.

Several types of imputation method may be used in multiple imputations. For example, regression imputation could be used repeatedly by drawing N values of the regression parameters using the variance matrix of estimated coefficients. However, one of the most commonly-used methods is the Markov Chain Monte Carlo (MCMC) method. MCMC is a sequence of random variables in which the distribution of an actual element depends on the value of the previous one. The method assumes that data are drawn from a multivariate normal distribution and requires "missing at random" or "missing completely at random" assumptions.

The main question a modeller must answer when creating imputations is which method a model should be used to fill in missing data. There is no definitive answer to this question, but a number of guidelines have been formulated. The choice depends primarily on the dataset available. Unfortunately, data expressed on a continuous scale or ordinal data, as in the case of data collected for StatDI, do not allow for methods such as MCMC. Generally speaking, with regard to single imputation, the amount of missing data, as compared to the scope of the dataset,[14] as well as the identity of the country and the indicator for which the data are missing, should be noted. Therefore, as recommended by OECD, we do not suggest that a specific method should be used, but rather that the method should be in line with the characteristics of any missing information.[15] For multiple imputations, regression imputation could, however, be used repeatedly.

(c) Imputing

This process consists of implementing the imputation models as defined previously. This is when an assessment will be made on the treatment of outliers.

The detailed concept of multiple imputations

The use of multiple imputations for missing data is often used to handle missing data in multivariate analysis. The pioneer of multiple imputations for missing data is Donald B. Rubin.[16]

The procedure for conducting multiple imputations for missing data as described by Rubin is as follows:

➤ The **first** step of multiple imputations for missing data is to impute the missing values by using an appropriate model from any available information that may be useful. Potential models include linear regression, logistic regression, and non-linear models.

➤ The **second** step is to repeat the first step several times (generally between three and five times).

➤ The **third** step is to perform the desired analysis on each dataset by using standard and complete data methods. This is done when the main purpose of the imputation is to run an estimated regression model.

➤ The **fourth** step is to average the values of the parameter estimates across the missing value samples in order to obtain a single value estimate. Alternatively, one may also simply calculate an average of the imputed missing values.

➤ The **fifth** step is to calculate the standard errors, firstly by averaging the squared standard errors of the missing value estimates, then by calculating the variance of the missing value parameters across the samples, and finally by combining the two quantities in multiple imputations for missing data.

Example 2: Simple illustration of multiple imputations

Let us suppose we only have the two variables X and Y in the dataset. Let us now consider that variable X has no missing value and Y has some missing values.

---

14   A small amount of missing data in a large dataset is unlikely to require sophisticated imputation methods such as MCMC.
15   For further information, see www.oecd.org/std/42495745.pdf.
16   Donald B. Rubin, Multiple Imputation for Nonresponse in Surveys, New York: Wiley, 1987.

Therefore $Y = (Y_M, Y_O)$

(M=missing, O=observed).

Table 4.A.1: Simulated data to illustrate multiple imputation

Unit	Data	
	Y	X
1	1.1	3.4
2	1.5	3.9
3	2.3	2.6
4	3.6	1.9
5	0.8	2.2
6	3.6	3.3
7	3.8	1.7
8	?	0.8
9	?	2.0
10	?	3.2

In the table above, we notice that the last three observations for Y are missing. These can be imputed by means of multiple imputations.

The first step in multiple imputations is to use the data from units where both *Y* and *X* are observed, (Y_O, X) — in order to identify the relationship between *Y* and *X.*

Table 5.A.2: Data from units where both Y and X are observed

Unit	Data	
	Y_O	X
1	1.1	3.4
2	1.5	3.9
3	2.3	2.6
4	3.6	1.9
5	0.8	2.2
6	3.6	3.3
7	3.8	1.7

Given that both variables are continuous, we will run a simple linear regression, with Y being the dependent variable and X being the explaining variable.

Once the regression is performed, if x represents the vector of **X** values from individuals with missing **Y** values, then we use the previous linear regression results to complete the missing values of **Y**. This process is then performed four times, with the results shown in table **A.3**.

As described in the Rubin process, once the four complete datasets have been imputed, the imputed values may be combined into a single value using the average.

Thus, for unit 8, in which the value was previously missing, the imputed value would be the mean

Table 6.A.3: Imputed datasets

Unit	Data		Imputation 1		Imputation 2		Imputation 3		Imputation 4	
	Y	X	Y	X	Y	X	Y	X	Y	X
1	1.1	3.4	1.1	3.4	1.1	3.4	1.1	3.4	1.1	3.4
2	1.5	3.9	1.5	3.9	1.5	3.9	1.5	3.9	1.5	3.9
3	2.3	2.6	2.3	2.6	2.3	2.6	2.3	2.6	2.3	2.6
4	3.6	1.9	3.6	1.9	3.6	1.9	3.6	1.9	3.6	1.9
5	0.8	2.2	0.8	2.2	0.8	2.2	0.8	2.2	0.8	2.2
6	3.6	3.3	3.6	3.3	3.6	3.3	3.6	3.3	3.6	3.3
7	3.8	1.7	3.8	1.7	3.8	1.7	3.8	1.7	3.8	1.7
8	?	0.8	0.2	0.8	0.8	0.8	0.3	0.8	2.3	0.8
9	?	2.0	1.7	2.0	2.4	2.0	1.8	2.0	3.5	2.0
10	?	3.2	2.7	3.2	2.5	3.2	1.0	3.2	1.7	3.2

of all values from the imputed dataset. The imputed value of unit 8 for variable Y would be (0.2+0.8+0.3+2.3)/4=0.9. The value of 0.9 will be accompanied by its standard value given that it would be useful for the purposes of uncertainty analysis.

There are various software programmes available that can be used to perform multiple imputations.

2. Weighting under the principal components analysis/factor analysis framework [178]

The below section describes how to use principal components analysis and factor analysis to compute weights and indicators.

As a starting point, some basic background statistics and mathematics required to understand the whole process are outlined below. We then describe the two techniques and present a step-by-by example to illustrate how the mechanism works.

Basic Statistical Background

In this subsection, we explain the mathematics behind the process to compute component weights using principal components analysis.

Let X and Y be two sets of data. Note that X_i represents the ith number in the data set X.

Mean

The mean of the dataset or variable X is given by:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Unfortunately, the mean does not reveal much about the data, except an approximate middle point.

For instance, the two following sets have the same mean, but are clearly not identical.

X= [0 8 12 20] and Y= [8 9 11 12]

The difference between those two sets is in the way they spread. The spread of the data is represented by the standard deviation.

The concepts of statistical variation and its basic measures

Standard deviation

The standard deviation is the average distance from the mean of the data set to a point. The way to calculate it is to compute the squares of the distance from each data point to the mean of the set, add them together, and divide the figure so obtained by n if you are working on the whole set of data or by n-1 if you are working on a sample of the complete data:

$$s = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{(n-1)}}$$

Example 3: Computation and interpretation of indicator variation using standard deviation

Case A: To illustrate how to measure the variation of a set of variables, assume X has four observations: X= [0, 8, 12, 20]. To calculate the standard deviation, we follow the steps arranged in column X. (See table A.4).

---

17 For further information, please see Handbook on Constructing Composite Indicators: Methodology and User Guide, Paris: OECD, 2008. Available at www.oecd.org/std/42495745.pdf.

Table 7.A.4: Computation of standard deviation – Case A

X	$(X_i - \overline{X})$	$(X_i - \overline{X})^2$
0	-10	100
8	-2	4
12	2	4
20	10	100
Total		208
Divided by (n-1)		69.333
Square Root		8.3266

Case B: Consider another set of Y with four observations: Y = [8, 9, 11, 12]. Following the same steps as in column X, we observe that 0 as standard deviation is drastically smaller when compared with the standard deviation shown in table 7(a).

Table 8.A.5: Computation of standard deviation - Case B

Y	$(Y_i - \overline{Y})$	$(Y_i - \overline{Y})^2$
8	-2	4
9	-1	1
11	1	1
12	2	4
Total		10
Divided by (n-1)		3.333
Square Root		1.8257

As expected, the first set has a much larger standard deviation due to the fact that the data are much more spread out from the mean.

Case C: Consider set Z, say:

Z= [10, 10, 10, 10]

In this case, Z has a mean of 10, but a standard deviation of 0 since all the values are the same and none of them deviate from the mean.

Variance

Variance is another measure of the spread of data in a dataset. In fact, it is almost identical to standard deviation. The formula is:

$$s^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{(n-1)}$$

The variance is the squared standard deviation. Both measurements are measures of the spread of the data. Standard deviation is the most common measure, but variance is also used.

Covariance

Variance and standard deviation are purely one dimensional. For example, data sets could be number of statisticians as a proportion of the total number of staff at African national statistical offices, or the number of partnership agreements concluded by a national statistical office with regional or global statistical agencies.

However, many data sets have more than one dimension (variable), and the aim of the statistical analysis of these data sets is usually to see if there is any relationship between the various dimensions. For example, we might have as our data set both the number of statisticians as a proportion of the total number of staff at African national statistical offices and the number of partnership agreements concluded by a national statistical office. We could then perform statistical analysis to see if the number of partnerships affects the composition of the staff.

Unfortunately, standard deviation and variance only operate in one dimension, so that you may only calculate standard deviation for each dimension of the dataset independently from other dimensions. However, it is useful to have a similar measure to find out how much the dimensions vary from the mean with respect to one another.

Covariance, however, is a measure that allows this type of analysis. Covariance is always measured between two dimensions. If you calculate the covariance between one dimension and itself, you get the variance.

Covariance is obtained as follows:

$$\text{cov}(X,Y) = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{(n-1)}$$

What does covariance tell us? Its exact value is not as important as its sign (positive or negative). A positive value indicates that both dimensions increase together. If the value is negative, then as one dimension increases, the other decreases.

Correlation coefficient

A correlation coefficient is a normalized form of the covariance coefficient. It is derived from the covariance of those variables and their respective standard deviations and is computed as follows:

$$corr(X,Y) = \frac{\text{cov}(X,Y)}{s_X s_Y}$$

Covariance Matrix

Recall that covariance is always measured between two dimensions. If we have a dataset with more than two dimensions, then more than one covariance measurement can be calculated.

A useful way to get all the possible covariance values between different dimensions (variables) is to calculate them all and put them in a matrix.

Correlation Matrix

Let X_{nxp} be the following matrix of data:

$$X = \begin{pmatrix} x_{11} & \dots & \dots & x_{1p} \\ x_{21} & \dots & \dots & x_{2p} \\ \dots & \dots & \dots & \dots \\ x_{n1} & \dots & \dots & x_{np} \end{pmatrix}$$

Where X_{ij} is the value of the row i and column j.

The correlation matrix R is defined as follows:

$$R = \begin{pmatrix} 1 & r_{21} & \dots & r_{1p} \\ r_{21} & 1 & \dots & r_{2p} \\ \dots & \dots & \dots & \dots \\ r_{p1} & r_{p2} & \dots & 1 \end{pmatrix}$$

Where $r_{ij} = corr(X_i, X_j)$ is the correlation coefficient of variables x_i and x_j as defined above.

Basic matrix algebra

Background is provided here for the matrix algebra required in principal components analysis. We will focus on eigenvectors and eigenvalues within a given matrix. We assume the reader has a basic knowledge of matrices.

For a given matrix A, an eigenvector v is a vector that verifies:

$$Av = \lambda v$$

Where λ is a scalar called the eigenvalue of the matrix A.

Principal components analysis) and factor analysis

General definition

Principal components analysis is a method that allows linear transformations of a large number of correlated variables in order to obtain a relatively limited number of uncorrelated components. This approach facilitates analysis by grouping

data into smaller sets and addressing challenges related to multicollinearity between variables.

Given a set of observations described by exclusively numerical variables $\{X_1, X_2, ..., X_p\}$, the objective of principle components analysis is to describe this same set of data with new and reduced variables. Those new variables will be linear combinations of the original variables, and will be called Principal Components.

In general, reducing the number of variables used to describe a dataset results in a loss of information. Principal components analysis minimizes that loss of information.

Principal components analysis can therefore be seen as a technique for reducing dimensionality.

Although the objective is generally to use only a few Principal Components, principal components analysis initially constructs p components equal to the number of original variables. The analyst decides how many

components to retain only at a later point in the process.

"Retaining k Main Components" means "replace the original observations with their orthogonal projections in the k-dimensional subspace defined by the first k Principal Components."

As per OECD guidelines, combining principal components analysis with factor analysis can provide a multiple weighting method for multivariate analysis; factor analysis is a simple way to endogenously determine the weights to be used in the summation of a composite index, whereas the principal components analysis/ factor analysis method utilizes the top Principal Components that account for most of the variance of the variables considered in factor analysis.[18]

Description of the principal components analysis/factor analysis technique

Within a national statistical system, let $X_{ijk} \in (0,1)$ be the score of kth the indicator (variable) for the ith

Matrix II: Arrangement of the scores of the variables for the application of the principal components analysis/factor analysis technique

Subdimension	Dimensions					Dj				
	Dj				....					
D_{ij}	D_{11}	X_{111}	X_{112}	...	X_{11n11}	....				
	D_{21}	X_{211}	X_{212}	...	X_{21n21}	....				
	...	...	...	...	...	....				
	$D_{s(1)1}$	$X_{s(1)11}$	$X_{s(1)12}$	...	$X_{s(1)1ns(1)1}$	....				
	...	...	....	....	....	.....				
	D_{1j}						$X_{1/1}$	$X_{1/2}$	...	X_{1jn1j}	....
	D_{2j}						$X_{2/1}$	$X_{2/2}$	...	X_{2jn2j}	....
	...						...	...	...	...	....
	$D_{s(j)j}$						$X_{s(i)i1}$	$X_{s(i)i2}$	...	$X_{1s(i)jns(j)}$	....
	...						...	....	....	....	.....

---

18 For further information, see http://publications.jrc.ec.europa.eu/repository/bitstream/JRC31473/EUR%2021682%20EN.pdf.

subdimension within the *jth* dimension, $i = (1,2,3, \dots, S; j = 1,2,3, \dots M; k = 1,2, \dots, N)$, under the assumption that each variable has the same importance in explaining the progress in statistical development (having equal weights). To compute the weights of subdimensions and dimensions of statistical development, recall Matrix I, or consider Matrix II below:

From Matrix II above, let us focus on the D_{ij} the *ith* subdimension within the *jth* dimension.

This subdimension has n_{ij} variables, namely:

$$| X_{ij1} \mid X_{ij2} \mid \dots \mid X_{ijk} \mid \dots \mid X_{ijnij} |$$

Based on the notation of Matrix II, let **Q = nij** ; and $X_{ijk} = X_k$ **(k=1,2,3,...Q)** that is to say, drop the subscripts for the subdimension and dimension and focus only on the variables within a given subdimension. The row data above can now be rewritten as:

$$| X_1 \mid X_2 \mid \dots \mid X_k \mid \dots \mid X_Q |$$

Principal components analysis

Let us now find linear combinations of these **Q** variables X_k **= (k=1,2,3,...Q)** to produce **Q**

Principal components Z_k **= (k=1,2,3,...Q)** such that we have a system of **Q** linear equations

$$Z_k = a_{k1}X_1 + a_{k2}X_2 + \dots + a_{kQ}X_Q; \quad (k = 1,2,3, \dots Q)$$

$\dots\dots$ (1)

Equation (1) above tells us that we have **Q** Principal Components combining the **Q** original variables (or indicators), X_k **= (k=1,2,3,...Q)**. The weights a_{kj} applied to the variables X_j in the *Kth* expression of Equation (1) are called components or factor loadings, (*j*=1,2,3,...Q) These weights

are chosen so that the principal components Z_k = **(k=1,2,3,...Q)** satisfy the following conditions:

(a) They are uncorrelated (orthogonal);

(b) The first principal component accounts for the maximum possible proportion of the variance of the set of the variables X, the second principal component accounts for the maximum of the remaining variance, and so on until the last of the principal components absorbs all the remaining variance not accounted for by the preceding components, and

    (i) each of the Q expressions of Equation (1) above satisfy the identity

$$a_{k1}^2 + a_{k2}^2 + \dots + a_{kQ}^2 = 1 \dots\dots (2)$$

Next, let us define the sample covariance matrix (CM), of *T* observations, (*T* may be a set of time series data for a given country or number of countries, etc.), so that:

$$CM = \begin{pmatrix} m_{11} & m_{12} & m_{..} & m_{QQ} \\ m_{21} & m_{22} & \cdots & m_{2Q} \\ \vdots & \vdots & \ddots & \vdots \\ m_{Q1} & m_{Q2} & \cdots & m_{QQ} \end{pmatrix}$$

$\dots\dots\dots\dots\dots\dots\dots\dots$ (3)

Where the diagonal element m_{kk} is the variance of , while is the covariance of variables $X_{k'}$ and $X_{l'}$ (l=1,2,3,...Q) Principal components analysis requires finding the eigenvalues $\lambda_k, k = 1,2,3, \dots, Q$ of the sample covariance matrix (CM) given by Equation (3) above.

> **Box 3: The eigenvalues of a matrix**
>
> Definition (1): The eigenvalues of the covariance matrix by definition are the variances of the principal components

The eigenvalues of the matrix CM can be found by solving the characteristic equation:

$$\det (CM - \lambda I) = |CM - \lambda I| = 0$$
.... (4)

Where *I* the identity matrix with the same order as *CM*, and *t* $\det (CM - \lambda I)$ is the determinant of matrix

$(CM - \lambda I)$. This characteristic equation defines a polynomial in the variable λ. The solution λ of this characteristic equation are called eigenvalues.

An important property of eigenvalues is that they add up to the sum of the diagonal elements of CM. That is:

Box 4: The relation between eigenvalues and the variances of the original variables

The sum of the variances of the principal components is equal to the sum of the variances of the original variables

So that:
$$\lambda_1 + \lambda_2 + \cdots + \lambda_Q = m_{11} + m_{22} + \cdots + m_{QQ}$$

.... (5)

Standardization of the original variables

Next, to prevent one variable from having an undue influence on the principal components, it is common to standardize the variables to have zero means and unit variances at the start of the

analysis. So that rather than to start the process with the original variables X_k = (k=1,2,3,...Q) we first standardize them to S_k = (k=1,2,3,...Q) where S_k is defined by the standardized formula:

$$S_k = \frac{X_k - \overline{X}_k}{\sigma_k}$$

....................................(6)

Where X_k is the sample mean, calculated, for example, for *T* countries with regard to the *kth* variable (indicator) and where σ_k is the standard deviation. Working with S_k = (k=1,2,3,...Q) instead of X_k = (k=1,2,3,...Q) will reduce the covariance matrix given by Equation (3) into a correlation matrix, since $\sigma_k^2 = 1$ for all (k=1,2,3,...Q)

So,

$$m_{11} = m_{22} = \cdots = m_{QQ} = 1$$
.... (7)

And of course, in this case, m_{ki} is the correlation between variables S_k and S_l

(l=1,2,3,...Q)

Arrangement of the principal components

Next, we arrange or list the principal components in a descending order according to the value of their eigenvalues (variances) as shown below:

Table 9.A.6: Arrangement of the principal components according to their eigenvalues

Pc	Eigenvalue (variance of principal components)	% (variance)	Accumulative (%)
1	λ_1	$\lambda_1/\Sigma\lambda_k = V_1$	V_1
2	λ_2	$\lambda_1/\Sigma\lambda_k = V_2$	$V_1 + V_2$
...	....	....	....
Q	λ_Q	$\lambda_Q/\Sigma\lambda_k = V_Q$	$V_1+V_2+.......+V_Q=100\%=1$
---	$\Sigma\lambda_k = Q$		

Note from the last row that the sum of the variances is equal to the number of principal components since we are working with standardized variables $S_k = (k=1,2,3,...Q)$ so that from Equations (5) and (7) the sum of

$X_{kk} = (k=1,2,3,...Q)$ is equal to Q.

Let us consider the following set of eight eigenvalues (3.3, 1.7, 1, 0.9, 0.5, 0.3, 0.2, 0.1).

Component loadings:

In the case of an uncorrelated X, $r(Z_k, X_j)$ then, reduces the weights *akj* applied to the variables *Xj* in the *kth* expression of Equation (1), which are also called components or factor loadings

High and moderate loadings (>0.50) indicate how the individual indicators are related to the principal components. To illustrate this, we can consider the eight indicators shown and arrange the correlation coefficients between the principal components Z and the variables X as shown in table A.9.

Each of the individual indicators, except x_1 and x_2 is entirely accounted for by one principal component alone.

The high and moderate loadings are all found in the first four principal components.

An undesirable property of these components is that two individual indicators x_1 and x_2 are related strongly to two principal components.

The four first components account for 88 per cent of the total variance.

Table 11.A.8: Component loadings for Q individual indicators

X (indicators)	Principal component (Zk)				
	1	2	3	....	Q
	Component loading				
X1	r(Z1, X1)	r(Z2, X1)	r(Z3, X1)	....	r(ZQ, X1)
X2	r(Z1, X2)	r(Z2, X2)	r(Z3, X3)	....	r(ZQ, X2)
X3	r(Z1, X3)	r(Z2, X3)	r(Z3, X3)	....	r(ZQ, X3)
....	....	....	....	....	....
XQ	r(Z3, XQ)	r(Z3, XQ)	r(Z3, XQ)	....	r(ZQ, XQ)

Table 12.A.9: Component loadings for eight individual indicators

X (indicators)	Principal components							
	1	2	3	4	5	6	7	8
	Component loadings							
X_1	-0.11	-0.75	0.13	0.60	-0.10	-0.12	-0.17	0.05
X_2	-0.56	-0.48	0.22	-0.54	0.27	-0.17	-0.04	0.10
X_3	0.92		0.02	-0.10	0.04	0.11	-0.27	-0.13
X_4	0.35	-0.85	0.01	-0.13	0.11	0.35	0.06	-0.08
X_5	-0.76	-0.39	-0.16	-0.16	-0.41	-0.16	0.16	-0.09
X_6	-0.91	0.13	0.01	0.07	-0.19	0.30	0.04	0.16
X_7	-0.74	0.11	0.37	0.39	0.33	-0.02	0.20	-0.07
X_8	-0.36	-0.12	-0.87	0.15	0.26	-0.03	0.02	0.02
% variance explained by components	0.42	0.22	0.12	0.11	0.06	0.04	0.02	0.01
Cumulative (%)	0.42	0.64	0.76	0.88	0.93	0.97	0.99	1.00

Source: Handbook on Constructing Composite Indicators: Methodology and User Guide, Paris: OECD, 2008.

Note: Extraction method: principal components analysis loadings greater than 0.5 (as absolute values) are highlighted, and n=23 countries. In addition, the percentage of variance explained by each component assumes that variables are standardized.

Note: In the case of uncorrelated X, then this percentage variance is given by the squared weights a^2_{kj}.

Component scores

Box X: computation of standard deviation

> **Box 8:** The relationship between indicator scores, component loading and principal component scores
>
> Definition (4): The component scores of each principal component are the scores of each case or observation. In our example, these include each country or different points of time for a given country.

The component score for a given case for a principal component is calculated by taking the case's standardized value for each variable, multiplying it by the corresponding loading of the variable for the given principal component, and summing those products. Table A.10 illustrates the component scores

for an example country (referred to as Country A) by using three variables (indicators).

Factor analysis

Factor analysis aims to describe a set of Q variables X_1, X_2,…X_Q or briefly X_k (k=1,2,3,…Q) in terms of a smaller number of unobservable m factors (also called latent factors) and to highlight the relationship between these variables. In other words, factor analysis is a statistical technique used to describe variability among observed, correlated variables in terms of a potentially lower number of unobserved or indirectly measured variables called factors.

Factor analysis assumes that the data are based on the underlying factors and that the data variance can be decomposed into that accounted for by common and unique factors. The model is a system of Q equations given by:

$$X_k = \alpha_{k1} F_1 + \alpha_{k2} F_2 + \cdots + \alpha_{km} F_m + e_k ; \quad (k = 1,2,3,…Q)$$

Table 13.A.10: Component score for Country A

Indicator/Variables (xs)	Standardized value of Xs	First principal component (1)	Second principal component (2)
X_1	S_1	$r(Z_1, X_1)$	$r(Z_2, X_1)$
X_2	S_2	$r(Z_1, X_2)$	$r(Z_2, X_2)$
X_3	S_3	$r(Z_1, X_3)$	$r(Z_2, X_3)$
The CS for Country A for PC (1) and PC (2)		$\sum_{k=1}^{3} S_k \, r(Z_1, X_k)$	$\sum_{k=1}^{3} S_k \, r(Z_2, X_k)$

and $m < Q$ ……….. (8)

Where the original Q variables X1, X2, …XQ are standardized with zero mean and unit variance;

$a_{k1}, a_{k2}, \ldots a_{km}$ are m the factor loadings related to the variable $X_1, X_2, \ldots X_Q$;

$F_1, F_2, \ldots F_m$ are uncorrelated common factors, each with zero mean and unit variance; and

e_k $(k=1,2,3,…Q)$ are the Q random errors assumed to be independently and identically distributed with zero mean Since the factor loadings, $\alpha_{kj}, (k=1,2,3,…,Q; j=1,2,3,…m)$ can be interpreted as standardized regression coefficients, one could also say that the variable X_k has a correlation of a_{jk} with Factor F_j.

To construct the system of the Q equations shown by Equation (8), the most common approach is to use principal components analysis to extract the first m principal components and to consider them as factors, neglecting those remaining. Principal components factor analysis is most preferred in the development of composite indicators. This approach has the virtue of simplicity and allows for the construction of weights representing the information content of individual indicators.

It should be noted, however, that different extraction methods supply different values for the factors and thus for the weights, influencing the score of the composite and the corresponding country ranking.

Adopting principal components analysis in order to define and select factors, will ascertain the following:

(a) Percentage of the variance in a given variable explained by a given factor

> **Box 9: The connection between the share of the variance of a given indicator as explained by a factor and the factor loading**
>
> Definition (5): The squared factor loading a^2_{kj} $(k=1,2,3,…,Q; j=1,2,3,…m)$ is the percentage of variance in that variable X_k explained by the Factor F_j

Note:

(i) To ascertain the percentage of the variance in all the variables accounted for by each factor, add the sum of the squared factor loadings for that factor (column) and divide by the number of variables.

(ii) The number of variables equals the sum of their variances as the variance of a standardized variable is 1. This is the same as dividing the factor's eigenvalue by the number of variables.

(b) The variance as explained by a given factor

> **Box 10: The variance as explained by a given factor captured from shares of variances from all indicators in the study**
>
> Definition (6): The variance explained by a given factor is the sum of the squared factor loading $\sum_{(k=1)}^{Q}\alpha^2_{kj'}$ (k=1,2,3,…,Q;j=1,2,3,…m) of all the variables X_k that explained by the Factor F_j

(c) Factor scores

> **Box 11: The relation between the scores of indicators, the factor loading and the factor scores**
>
> Definition (7): The factor scores for a case are the standardized scores of each variable (row) for each factor (column)

With component scores, to compute the factor score for a given case and factor, one takes the case's standardized score for each variable, multiplies it by the corresponding loadings of the variable for the given factor and sums those products. Computing factor scores allows for an examination of factor outliers. Factor scores may also be used as variables in subsequent modelling.

Communality

> **Box 12: The relationship between the variation of indicators and the associated factors**
>
> Definition (8): The sum of the squared factor loadings for all factors for a given variable (row) is the variance in that variable and is accounted for by all the factors. This is called communality.

Communality measures the percentage of the variance in a given variable as explained by all factors jointly, and may be interpreted as the reliability of the indicator.

Stopping rules

Standard practice is to choose factors that:

Table 14.A.11: Eigenvalues of eight individual indicators and the corresponding per cent variance and cumulative variance

A	B	C	D
Principal component	Eigenvalue (variance of principal components)	Percentage (variance)	Cumulative (per cent)
1	3.3	41.25	41.25
2	$\lambda_2=1.7$	21.25	62.5
3	$\lambda_3=1.0$	12.5	75
4	$\lambda_4=0.9$	11.25	86.25
5	$\lambda_5=0.5$	6.25	92.5
6	$\lambda_6=0.3$	3.75	96.25
7	$\lambda_7=0.2$	2.5	98.75
8	$\lambda_8=0.1$	1.25	100%=1
-	$\sum\lambda_k=8$		

Source: Handbook on Constructing Composite Indicators: Methodology and User Guide, Paris: OECD, 2008.

(a) Have associated eigenvalues larger than one;

(b) Contribute individually to the explanation of overall variance by more than 10 per cent;

(c) Cumulatively contribute to the explanation of the overall variance by more than 60 per cent.

The first four factors are those with eigenvalues close to unity, as shown in Table A.7. Individually, each factor explains more than 10 per cent of the total variance, and together they count for approximately 86.25 per cent of the variance.

After choosing the number of factors to be retained, it is standard practice to perform a rotation so as to enhance the interpretability of the results.

The sum of eigenvalues is not affected by rotation. Changing the axes will alter the eigenvalues of particular factors and the factor loadings.

Various rotational strategies have been proposed. The most common rotation method is the varimax rotation, but the goal of all methods is to obtain a clear pattern of loadings. However, different rotations imply different loadings and thus varying significances for principal components. This is sometimes cited as a drawback to the method.

Rotation methods

Rotation serves to make output more understandable by seeking "simple structure." Simple structure is a pattern of loadings where each item or variable loads strongly on only one of the factors and more weakly

Table 15.A.12: Rotated factor loadings for four individual indicators (method 1)

X (indicators)	Factors				Communality	Remarks
	1	2	3	4		
	Factor loadings					
X_1	0.07	0.97	0.06	0.06	$(0.07)^2 + (0.97)^2 + (0.06)^2$ $(0.06)^2 = 0.95$	By Definition (8)
X_2	0.13	0.07	-0.07	0.93	0.89	
X_3	0.79	-0.21	0.21	0.42	0.89	
X_4	-0.64	0.56	-0.04	0.36	0.86	
X_5	0.37	0.17	0.38	0.68	0.77	
X_6	0.82	-0.04	0.25	0.35	0.85	
X_7	0.88	0.23	-0.09	0.09	0.85	
X_8	0.08	0.04	0.96	0.04	0.93	
Explained variance	2.64	1.39	1.19	1.76		For Factor 1 by definition (6) is : $(0.07)^2 + (0.13)^2 + (0.79)^2 + ... + (0.08)^2 = 2.64$
Percentage of the variance in all the variables accounted for by each factor	33.00	17.38	14.88	22.00		Explained variance divided by the number of variables (See Note (i) under Definition (5))
Cumulative (per cent)	33	50	65	87		

Source: Handbook on Constructing Composite Indicators: Methodology and User Guide, Paris: OECD, 2008.

Table 16.A.13: Rotated factor loadings for four individual indicators (method 2)

X (indicators)	Factors				Communality	Remarks
	1	2	3	4		
	Factor loadings					
X_1	0.01	0.11	0.88	0.13	$(0.01)^2 + (0.11)^2 + (0.88)^2 \ ) \ (0.13)^2 = 0.80$	By Definition (8)
X_2	0.96	0.14	0.09	0.18	0.99	
X_3	0.31	0.56	-0.29	0.60	0.86	
X_4	0.29	-0.45	0.58	-0.14	0.65	
X_5	0.41	0.13	0.18	0.73	0.75	
X_6	0.13	0.57	-0.13	0.73	0.89	
X_7	0.14	0.95	0.10	0.14	0.95	
X_8	-0.01	0.03	0.03	0.39	0.15	
Explained variance	1.31	1.80	1.27	1.67		For Factor 1 by Definition (6) is: $(0.01)^2 + (0.96)^2 + (0.31)^2 \ ) + ... + (-0.01)^2 = 1.314$
Percentage of the variance in all the variables accounted for by each factor	16	23	16	21		Explained variance divided by the number of variables (See Note (i) under Definition (5))
Cumulative (per cent)	16	39	55	76		

Source: Handbook on Constructing Composite Indicators: Methodology and User Guide, Paris: OECD, 2008.

on the other factors. Rotations can be orthogonal or oblique, thus allowing the factors to correlate.

Varimax rotation

A varimax rotation yields results that attempt to make it as easy as possible to identify each variable with a single factor. Geometrically, varimax rotation is an orthogonal rotation of the factor axes designed to maximize the variance of the squared loadings of a factor (column) on all of the variables (rows) in a factor matrix. This has the effect of differentiating the original variables by an extracted factor. Each factor will tend to have either large or small loadings of each variable. However, the orthogonality (i.e. independence) of factors, is often an unrealistic assumption. Oblique rotations include orthogonal rotation, and for that reason, oblique rotations are a preferred method.

In oblique rotations, one calculates both a pattern matrix and a structure matrix. The structure matrix is simply the factor loading matrix in orthogonal rotation, which represents the variance in a measured variable and is explained by a factor on both a unique and common contributions basis. In contrast, the pattern matrix contains coefficients that represent unique contributions. The more factors involved, the lower the pattern coefficients, since more common contributions to the variance will be explained. For oblique rotation, the calculation examines both the structure and pattern coefficients when attributing a label to a factor. See table A.12.

Note: Estimation method: Principle Components, varimax rotation. Positive loadings greater than 0.5 are highlighted.

Factor 1 is dominated by X_3, X_6 and X_7 (high positive coefficients (loadings) >0.5;

Factor 2 is dominated by , X_1, and X_4

Factor 3 is dominated by X_8 so that X_8 is exclusively loaded in factor 3

Factor 4 is dominated by X_2, and X_5

The four first factors count for 87% of the total variance (cumulative percentage variance corresponding to factor 4)

A different method of extracting factors that examines non-correlation of specific factors would have yielded different results. Table A.13 presents the rotated factor loadings of the four factors using the "principal factors maximum likelihood" extraction method.

Principal components analysis and factor analysis aim to account for the highest possible variation in the indicator set using the smallest possible number of factors. Therefore, the composite no longer depends upon the dimensionality of the dataset, but rather is based on the statistical dimensions of the data.

Principal components analysis is often used to extract factors. In factor analysis, only a subset of principal components is retained, usually those that account for the largest amount of the variance.

Steps in factor analysis

The first step in factor analysis is to check the correlation structure of the data. If correlation between the indicators is weak, then it is unlikely that they share common factors.

The second step is to identify a certain number of latent factors representing the data. Each factor depends on a set of coefficients (loadings), and each coefficient measures the correlation between the individual indicator and the latent factor.

The third step is to examine the rotation of factors (see table A.14). The rotation is used to minimize

Table 17.A.14: Factor loadings of eight indicators based on principal components

	Factor loading				Squared factor loading (scaled to unity sum)			
	Factor1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2	Factor 3	Factor 4
X_1	0.07	0.97	0.06	0.06	0.00	0.68	0.00	0.00
X_2	0.13	0.07	-0.07	0.93	0.01	0.00	0.00	0.49
X_3	0.79	-0.21	0.21	0.42	0.24	0.03	0.04	0.10
X_4	-0.64	0.56	-0.04	0.36	0.15	0.23	0.00	0.07
X_5	0.37	0.17	0.38	0.68	0.05	0.02	0.12	0.26
X_6	0.82	-0.04	0.25	0.35	0.25	0.00	0.05	0.07
X_7	0.88	0.23	-0.09	0.09	0.29	0.04	0.01	0.00
X_8	0.08	0.04	0.96	0.04	0.00	0.00	0.77	0.00
Explained variance	2.64	1.39	1.19	1.76				
Expl./ ToT = W_f	0.38	0.20	0.17	0.25				

Source: Handbook on Constructing Composite Indicators: Methodology and User Guide, Paris: OECD, 2008.

the number of individual indicators that highly load the same factor. The idea behind transforming the factorial axes is to obtain a simpler structure of the factors. Ideally this is a structure in which each indicator is loaded exclusively on one of the retained factors. Rotation is a standard step in factor analysis since it changes the factor loadings, and hence the interpretation of the factors, while leaving unchanged the analytical solutions obtained ex-ante and ex-post the rotation.

Note: Expl./ToT is the share of the variance explained by the factor from the total variance of the four factors which in this case is 6.98. This share is calculated by taking the variance explained by the factor and dividing that figure by the total variance. For the purpose of obtaining final weights for the indicators, this share is also known as the weight of the respective factor (W_f).

To scale to unity sum for a given factor is obtained by dividing each squared factor loading by the total squared factor loadings in that factor.

The fourth and last step involves the construction of the weights from the matrix of squared factor loadings after rotation, given that the square of factor loadings represents the proportion of the total unit variance of the indicator, which is explained by the factor. One possible approach is to group the individual indicators with the highest factor loadings into intermediate composite indicators. The corresponding squared loadings of intermediate composite indicators are also known as domain weights (see table A.15).

Next, as shown in the below table, we aggregate the intermediate composite indicators using the share of variance explained by each factor.

Finally, we scale these weights such that their sum is unity, as shown below.

Example 4: Estimation of weights for Statistical Development Indicators using the principal components analysis/factor analysis technique in a hypothetical situation

We can use the principal components analysis and factor analysis technique to calculate weights for the first dimension (module) of StatDI, which has two subdimensions D_{11} and D_{21} respectively.

We have three indicators in D_{11} and five indicators in D_{21} so for D_{21} we can set $Q=5$ for this subdimension. Accordingly, for country XYZ, say we have

$$(X_1, X_2, X_3, X_4, X_5) = (1.0, 0.5, 1.0, 0.25, 1.0) \ldots\ldots \quad (A.1)$$

Table 18.A.15: Domain weights of eight indicators based on principal components

Indicators (variables)	Factor 1	Factor 2	Factor 3	Factor 4
	Intermediate composite indicators- OR-Domain weights (W_D)			
X_3, X_6, X_7	(0.24, 0.25.0.29)			
X_1, X_4		(0.68, 0.23)		
X_8			(0.77)	
X_2, X_5				(0.49,0.26)

Source: Handbook on Constructing Composite Indicators: Methodology and User Guide, OECD, 2008.

Table 19.A.16: Aggregating the intermediate composite indicators using the factor shares of variance

Indicators (variables)	Factor1	Factor2	Factor3	Factor4
	Weight Score = (domain weights) multiplied by (weight of the respective factor) = $(W_D)(W_f)$			
X_3, X_6, X_7	0.38(0.24, 0.25, 0.29) = (0.09,0.10,0.11)			
X_1, X_4		0.20(0.68, 0.23)= (0.14, 0.05)		
X_8			0.17(0.77) = (0.13)	
X_2, X_5				0.25(0.49, 0.26) (0.12, 0.07)

Source: Handbook on Constructing Composite Indicators: Methodology and User Guide, Paris: OECD, 2008.

Assume we received data from 20 countries, then we will have twenty expressions in the form of Equation (A.1) which will represent our sample.

The Principal components Z_k (k=1,2,3,…5) will reduce to:

$$Z_k = a_{k1} X_1 + a_{k2} X_2 + … + a_{k5} X_5; \ (k=1,2,3,…5) … \text{(A.2)}$$

Next, we standardize Xs using equation (6) above: $S_k = (X_k - \overline{X}_k)/\sigma_k)$ for each of the five indicators so that we are working with $(S_1, S_2, S_3, S_4, S_5)$ rather than $(X_1, X_2, X_3, X_4, X_5)$.

Solving the system of the seven equations given by Equation (A.2), however after replacing X_j by S_j so we can obtain the weights a_{kj} or components or factor loadings, (j,k=1,2,3,…5) for the principal components Z_k. At this stage, each of the seven indicators (variables) is checked.

$$a^2_{k1}2 + a^2_{k2} + … + a^2_{k5} = 1; (k=1,2,3,…5) ……… \text{(A.3)}$$

Next, we must calculate the covariance matrix, which, in this case, will be the correlation matrix (CoM), in order to calculate the eigenvalues of the Principal components. The matrix will therefore be:

Table 20.A.17: Scaled weights using principal component and maximum likelihood methods

	Principal component method		Maximum likelihood method
Indicators	A Unscaled weights (Weight score)	B Scaled weights (unscaled) divided by (sum of weights)	
X_1	0.14	0.17	0.19
X_2	0.12	0.15	0.20
X_3	0.09	0.11	0.07
X_4	0.05	0.06	0.07
X_5	0.07	0.09	0.15
X_6	0.10	0.12	0.11
X_7	0.11	0.14	0.19
X_8	0.13	0.16	0.02
	Sum of weights = 0.81		

Source: Handbook on Constructing Composite Indicators: Methodology and User Guide, OECD, 2008.

$$CoM = \begin{pmatrix} 1 & m_{12} & m_{..} & m_{55} \\ m_{21} & 1 & \cdots & m_{25} \\ \vdots & \vdots & \ddots & \vdots \\ m_{51} & m_{52} & \cdots & 1 \end{pmatrix}$$

……………… ……………… (A.4)

Where the diagonal elements denote the variance of the variables, while the off-diagonal elements represent the correlation between the variables.

Using Equation (A.4), we can compute the eigenvalues λ_j (j=1,2,3,…5) , which represent the corresponding variances of the Principal Components, and solve the following characteristic equation:

det(CM-λI)=|(CM-&λI)|=0 ……………. (A.5a)

Where, as defined before, I is the identity matrix with the order five as the CoM order; and **det(CM-λI)** is the determinant of matrix **(CoM- λI)** . This characteristic equation defines a polynomial in the variable **λ**. The factorized version of this polynomial will be:

$(R_1- \lambda_1)$ $(R_2- \lambda_2)$ $(R_3- \lambda_3)$… $(R_5- \lambda_5)$=0… (A.5b)

The solutions $\lambda=(\lambda_1,\lambda_2,\lambda_3,…,\lambda_5)$ of this characteristic equation are called eigenvalues,

which are the variances of the seven principal components.

At this stage in the calculation, check that:

$\lambda_1+\lambda_2+…+\lambda_5=m_{11}+m_{22}+…+m_{55} = 5$ …… (A.6)

Since,

$m_{11}=m_{22}=…=m_{55}=1$ …………….. (A.7)

Now we can arrange these five eigenvalues, according to their values and starting from the highest value such that:

$R_1 \geq R_2 \geq R_3 \geq … \geq R_5$, which are numerical values as shown in table A.18 below.

Next, let us examine the correlation matrix between with (S_1,S_2,S_3,S_4,S_5), and (Z_1,Z_2,Z_3,S_4,Z_5), which provides the component loadings for of seven individual indicators.

A software will report this matrix in a version as shown in the table A.19, a matrix of 5X5 elements.

Next, let us examine the correlation matrix between (S_1,S_2,S_3,S_4,S_5), and (Z_1,Z_2,Z_3,S_4,Z_5), which provides the component loadings for

Table 21.A.18: **Eigenvalues of five individual indicators and a functional framework for organizing and coordinating a national statistical office**

A	B	C	D
PC	Eigenvalue (variance of PC)	% (variance)	Cumulative (%)
1	$\lambda_1=R_1$	$V_1=R_1/5$	V_1
2	$\lambda_2=R_2$	$V_2=R_2/5$	V_1+V_2
3	$\lambda_3=R_3$	$V_3=R_3/5$	$V_1+V_2+V_3$
4	$\lambda_4=R_4$	$V_4=R_4/5$	$V_1+V_2+V_3+ V_4$
5	$\lambda_5=R_5$	$V_5=R_5/5$	$V_1+V_2+V_3+ V_4+V_5=100$
-	$\Sigma\lambda_k =5$		

seven individual indicators. Software can be

A	B	C	D
Principal component	Eigenvalue (variance of principal components)	Percentage (variance)	Cumulative (per cent)
1	$\lambda_1 = R_1$	$V_1 = R_1/5$	V_1
2	$\lambda_2 = R_2$	$V_2 = R_2/5$	$V_1 + V_2$
3	$\lambda_3 = R_3$	$V_3 = R_3/5$	$V_1 + V_2 + V_3$
4	$\lambda_4 = R_4$	$V_4 = R_4/5$	$V_1 + V_2 + V_3 + V_4$
5	$\lambda_5 = R_5$	$V_5 = R_5/5$	$V_1 + V_2 + V_3 + V_4 + V_5 = 100$
-	$\Sigma\lambda_k = 5$		

used to report this matrix, as shown in table A.19, as a matrix of 5X5 elements.

In the case of an uncorrelated **X**, then $r(Z_k, x_j)$ reduces to the weights a_{kj} applied to the variables X_j in the *kth* expression of Equation (1). These are also referred to as components or factor loadings. Certain software[19] will also report the squared component loading $r^2(Z_k, x_j)$ which, as defined earlier, is the percentage of variance in that variable X_j explained by the principal component z_k.

The next step involves an examination of the results of factor analysis as provided by the software that eventually will provide weights for the seven indicators of D_{21} (A functional framework organizing and coordinating the national statistical office). Using principal components analysis/factor analysis, let us assume the first two principal components shown in table A.1 satisfy the OECD criteria for factor selection. In this case, statistical software can solve the following system of five equations:

$$X_k = \alpha_{k1}F_1 + \alpha_{k2}F_2 + e_k \quad (k=1,2,3,\ldots,5) \ldots\ldots\ldots\ldots (A.8)$$

With the weights α_{kj} (k=1,2,3,..,5;j=1,2) representing the loading factor such that $\widehat{\alpha_{kj}}$ is the estimate for α_{kj} (k=1,2,3,..,5;j=1,2) as given by the statistical software.

Table 22.A.19: Component loadings for five individual indicators

X (indicators)	Principal component (Z_k)				
	Z_1	Z_2	Z_3	Z_4	Z_5
	Component loadings				
X_1	$r(Z_1, X_1)$	$r(Z_2, X_1)$	$r(Z_3, X_1)$	$r(Z_4, X_1)$	$r(Z_5, X_1)$
X_2	$r(Z_1, X_2)$	$r(Z_2, X_2)$	$r(Z_3, X_2)$	$r(Z_4, X_2)$	$r(Z_5, X_2)$
X_3	$r(Z_1, X_3)$	$r(Z_2, X_3)$	$r(Z_3, X_3)$	$r(Z_4, X_3)$	$r(Z_5, X_3)$
X_4	$r(Z_1, X_4)$	$r(Z_2, X_4)$	$r(Z_3, X_4)$	$r(Z_4, X_4)$	$r(Z_5, X_4)$
X_5	$r(Z_1, X_5)$	$r(Z_3, X_5)$	$r(Z_3, X_5)$	$r(Z_4, X_5)$	$r(Z_5, X_5)$

19 Potential software for such analysis includes BMDP, JMP (statistical software), Statistical Package for the Social Sciences and Stata.

Table 23.A.20: Rotated factor loadings of seven indicators based on principal components

Variables (Indicators)	Factor loading		Squared factor loading (scaled to unity sum)	
	Factor1	Factor2	Factor1	Factor2
X_1	$\widehat{\alpha_{11}}$	$\widehat{\alpha_{12}}$	$\widehat{\alpha_{11}^2}/A$	$\widehat{\alpha_{12}^2}/B$
X_2	$\widehat{\alpha_{21}}$	$\widehat{\alpha_{22}}$	$\widehat{\alpha_{21}^2}/A$	$\widehat{\alpha_{22}^2}/B$
X_3	$\widehat{\alpha_{31}}$	$\widehat{\alpha_{32}}$	$\widehat{\alpha_{31}^2}/A$	$\widehat{\alpha_{32}^2}/B$
X_4	$\widehat{\alpha_{41}}$	$\widehat{\alpha_{42}}$	$\widehat{\alpha_{41}^2}/A$	$\widehat{\alpha_{42}^2}/B$
X_5	$\widehat{\alpha_{51}}$	$\widehat{\alpha_{52}}$	$\widehat{\alpha_{51}^2}/A$	$\widehat{\alpha_{52}^2}/B$
Explained variance			$A= \sum_{k=1}^{5} \widehat{\alpha_{k1}^2}$	$B= \sum_{k=1}^{5} \widehat{\alpha_{k2}^2}$
Factor weight (*Wf*)			$W_1 = A/(A+B)$	$W_2 = B/(A+B)$

Next, to enhance the interpretability of the results, let us use Varimax rotation with the aim of identifying each variable as a single factor.

Let us depict the preliminary results of the software as shown in table A.20.

Note: Principal component extraction method with varimax normalized rotation.

From table A.20 reading the factor loading corresponding to each variable (row), assume that

$$\widehat{\alpha_{12}} > \widehat{\alpha_{11}} \text{ for } X_1; \ \widehat{\alpha_{22}} > \widehat{\alpha_{21}} \text{ for }$$

$$X_2; \ \widehat{\alpha_{31}} > \widehat{\alpha_{32}} \text{ for } X_3;$$

$$.... ; \widehat{\alpha_{51}} > \widehat{\alpha_{52}} \text{ for } X_5;$$

table A.20 also shows the percentage of variance in the variable X_k explained by the Factor F_j which is simply obtained by squaring $\widehat{\alpha_{kj}}$ to obtain $\widehat{\alpha_{kj}^2}$ $(k = 1,2, ...,5; j = 1,2)$, (see Definition 5) as well as the share of variance explained by a given factor

which is simply the sum of the squared factor loading, that is $A = \sum_{k=1}^{Q} \alpha_{k1}^2$, $(k = 1,2,3, ...,5)$ of all the variables X_k that are explained by the Factor F_1 and $B = \sum_{k=1}^{Q} \alpha_{k2}^2$, $(k = 1,2,3, ...,5)$, of all the variables X_k that are explained by Factor F_2 (See Definition 6). We use these shares to scale to unity the percentage of variance in variable X_k explained by a given factor as well as in defining the factor weight as displayed in table A.20.

Next, by reviewing the definition of the five variables (indicators) of the subdimension D_{11} from the original questionnaire, then table A.20 also tells us that Factor 1 captures most of the organizational status of the national statistical office, while Factor 2 reflects the coordination of the national statistical offices within that subdimension.

The next step is to define the weight score using the following expression:

Table 24.A.21: Scaled weights using the principal component for the first subdimension in Statistical Development Indicators

Variables in D_{11}	Weight Score	Scaled weights
X_1	$\omega_1 = W_2(\widehat{\alpha_{12}^2}/B)$	$W_1 = \omega_1/C$
X_2	$\omega_2 = W_2\widehat{\alpha_{22}^2}/B$	$W_2 = \omega_2/C$
X_3	$\omega_3 = W_1(\widehat{\alpha_{31}^2}/A)$	$W_3 = \omega_3/C$
X_4	$\omega_4 = W_2(\widehat{\alpha_{42}^2}/B)$	$W_4 = \omega_4/C$
X_5	$\omega_5 = W_1(\widehat{\alpha_{51}^2}/A)$	$W_5 = \omega_5/C$
	C= sum of the weight scores	

Weight Score = (domain weights) multiplied by (weight of the respective factor)

$$= (W_D)(W_f) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (A.9)$$

Table A.20 suggests we have two domain weights, $D_1 = (\widehat{\alpha_{31}^2}, \widehat{\alpha_{51}^2})/A$ and $D_2 = (\widehat{\alpha_{12}^2}, \widehat{\alpha_{22}^2}, \widehat{\alpha_{42}^2})/B$. Accordingly, the weight scores can be defined as shown in table A.21 by collecting the data from table A.20. Those scores may turn out to be unscaled weights. Table A.21 also shows the scaled weights, also known as the final resulting weights, for the variables under study for the first subdimension of StatDI. Those scores are simply obtained by dividing each of the unscaled weight by their total sum, say C.

Where: $(W_1, W_2 \dots W_3)$ are the desired weights under principal components analysis/factor analysis for the second subdimension D_{21} of the first dimension D_1 of StatDI.

In the same manner, the principal components analysis/factor analysis weighting procedure can be applied to aggregate the dimensions in order to calculate the global StatDI score.

Any national statistical office can therefore calculate StatDI on the basis of the equations provided in A.2 and A.8 and the correlation matrix provided in A.4, even if their statistical software has no option for principal components analysis/factor analysis. That is one of the value adds of this manual, since some national statistical offices may not have principal components analysis/factor analysis options for weights estimation built into their statistical infrastructure.

3. Uncertainty and sensitivity analysis

Sensitivity analysis can be used to assess the robustness of composite indicators. Uncertainty analysis and sensitivity analysis are a powerful combination of techniques that help provide useful insights during the process of composite indicator building, including by contributing to the quality definition of indicators and an assessment of country ranking reliability. A combination of uncertainty and sensitivity analysis can help to gauge the robustness of a composite indicator, and increase its transparency. Uncertainty analysis focuses on how uncertainty in input factors propagates the structure of the composite indicator and affects composite indicator values. Sensitivity analysis studies how much each individual source of uncertainty contributes to the output variance. In the field of

building composite indicators, uncertainty analysis is more often adopted than sensitivity analysis[20] and the two types of analysis are almost always used separately. Synergistic use of uncertainty analysis and sensitivity analysis has proven to be more powerful).[21]

The above is a leading question that could be addressed when conducting an indicator robustness check.

To answer this question, one may use the multi-modelling approach.

A multi-modelling approach would explore various combinations of the three main assumptions needed to build an index:

(i) A weight is given to each subcomponent/ variable

In terms of a weighting scheme, a weighted approach based on principal components analysis/ factor analysis could be developed. In practice, the approach is separate from the unique frameworks that may be used for weighting. In addition to a weighting scheme, two more alternatives could be tested: an equal weight approach and a composite indicator calculation.

(ii) The aggregation technique

Regarding the aggregation technique, a weighted geometric rule was used to build StatDI:

$$ASDIs = \prod_{i=1}^{n} SC_i^{w_i} \quad (1)$$

Within the multi-modelling process, other aggregation techniques may be examined. For example, a linear aggregation scheme may be tested in addition to the baseline scenario aggregation scheme. The baseline scenario is the original model used to compute the indicator.

(iii) The number of subcomponents included in the final measure

For the purposes of our multi-modelling approach, StatDI may be built either by keeping all subcomponents/variables or by excluding one at a time. That statistical procedure allows the robustness of inference to be tested, but should not be considered as a modification of the general framework that has been retained. For example, the dimension entitled, "Capacity for an efficient statistical system" has seven subcomponents, i.e. funding, good governance of the statistical office, human capital, physical and material infrastructure of the national statistical office, statistical infrastructure, ICT infrastructure, and the promotion of bilateral and multilateral cooperation.

The original indicator for that dimension has been built on all seven subcomponents. In this phase, one may remove either one of the variables, such as "the proportion of the national statistical office budget financed by the government", and replace it instead with, "the percentage of last year's total national budget allocated for all statistical activities at the national statistical office" within the funding subcomponent, or remove the entire funding subcomponent and rebuild the indicator. In so doing, one may obtain a new value for the indicator

---

20 Jamison, D. and M. Sandbu, World Health Organization ranking of health system performance. 2001, Science, 293, pp. 1595–1596; and Freudenberg, M, "Composite indicators of country performance: a critical assessment" 2003 report DSTI/ IND (2003)5. Organization for Economic Cooperation and Development: Paris.

21 Tarantola, S., J. Jesinghaus and M. Puolamaa, "Global sensitivity analysis: a quality assurance tool in environmental policy modelling". In Sensitivity Analysis (eds A. Saltelli, K. Chan and M. Scott), 2000, pp. 385–397, New York: Wiley.

and a different rank of the country being considered using this indicator.

By varying and combining the three assumptions, one may create hundreds of model scenarios for the new StatDI that may be compared to the original scenario. It is also possible to run the multi-modelling framework by incorporating other assumptions. For example, it is possible to introduce an assumption in the imputation technique. In this way, one may impute the missing values using techniques that are different from those in the suggested multiple imputation framework. One can also vary the normalization scheme.

The normalization scheme

The assumptions of aggregation, weighting, and indicator inclusion are the common assumptions that researchers use to run the multi-modelling analysis.

As a complement to uncertainty analysis, one may perform a sensitivity analysis, using the various scenarios derived from multi-modelling design. For example, for each country a calculation may be done regarding the absolute rank shift between the original index rank and the rank provided by the scenarios, and the Spearman rank correlation

coefficient and percentiles may be used to report shifts for all countries.

> **Box 14:** **Spearman rank correlation coefficient, percentile and absolute difference**
>
> Definition of the Spearman rank correlation coefficient: The Spearman rank correlation is the correlation between the rankings of two variables.
> Definition of a percentile: A percentile or centile is a measure used to indicate the value below which a given percentage of observations within a group falls.
> Based on that definition, the 50th percentile is the value below which 50 per cent of all observations in a data set may be found, and the 90th percentile is the value below which 90 per cent of all observations in a data set may be found.
> Definition of absolute difference: the absolute difference between two variables is the absolute value of the difference between those two variables.

To compute the Spearman rank correlation coefficient for two variables, one may use the following formula:

$$r_s = 1 - \frac{6\sum_{k=1}^{n} d^2_k}{n^3 - n}$$

Where:

r_s = Spearman's rank value

Table 25.A.22: Building the Spearman rank correlation coefficient

Input the values for variables x and y and the ranks in the table

Calculate the difference (d) between Rank x and Rank y

Square the difference

| n | Variable x | | Variable y | | Difference d (Rx-Ry) | Difference d2 |
	x	Rank x (Rx)	Y	Rank y (Ry)		
1	4	3	6	3	0	0
2	3	4	5	4	0	0
3	5	2	3	5	-3	9
4	7	1	7	2	-1	1
5	2	5	8	1	4	16
Total						$\sum_{k=1}^{n} d^2_k$ =26

d = difference in rank of the values of each matched pair

n = number of individuals in the sample

As an example, consider the following process in building the Spearman rank correlation coefficient with a sample containing five observations:

n=5, thus the Spearman rank correlation for our example is

$$r_s = 1 - \frac{6 * 26}{5^3 - 5} = -0.3$$

The Spearman rank correlation expresses the strength of the linkage between variable ranks in a single value between -1 and +1.

A positive correlation coefficient indicates a positive relationship between the two variables, while a negative correlation coefficient indicates a negative relationship. A correlation coefficient of 0 indicates no linear relationship between the two variables.

Example 5: Rating ten countries in three scenarios, a short illustration

Assume that we received data from ten countries and that we computed the indicator for the dimension "Capacity for an efficient statistical system".

Assume that, by varying and combining the three assumptions of aggregation, weighting, and indicator inclusion, we calculate 150 simulated scenarios, including the baseline.

After computing the indicator with the baseline and other scenarios for countries being considered, one may rank countries from first to tenth for each scenario.

Matrix III, shown below, shows the frequencies of a country's indicator rank as calculated in all scenarios. A frequency matrix such as this one synthesizes the rankings while highlighting any uncertainty.

Interpreting Matrix III

The results for Country 1 indicate that 100 per cent of scenarios rank Country 1 first or second. The results for Country 2 say that 97 per cent of scenarios rank Country 2 first or second, while three per cent place rank it third or fourth. This indicates that it is beyond doubt that Country 1 and Country 2 are the top two countries in those scenarios. As one may notice for Country 5, the different scenarios give widely varying

Matrix III: Rating a group of countries in various scenarios (ranking)

	1-2	3-4	5-6	7-8	9-10
Country 1	100	0	0	0	0
Country 2	97	3	0	0	0
Country 3	4	85	9	2	0
Country 4	0	92	3	3	2
Country 5	0	20	40	18	22
Country 6	0	14	66	5	15
Country 7	0	9	30	49	12
Country 8	0	3	15	75	7
Country 9	0	0	3	35	62
Country 10	0	0	15	13	72

Table 26.A.23: Impact of the assumptions on the indicator rankings

Scenario	Weighting	Aggregation	Excluded subcomponent	50th percentile	90th percentile	Spearman rank coefficient
1	Original (Principal components analysis/ factor analysis)	Original (Geometric)	Funding	4	11	0.872
2	Equal Weighting	Arithmetic	Good governance of the statistical office	3	11	0.892
3	Equal Weighting	Original (Geometric)	ICT infrastructure	3	10	0.921

results, with 50 per cent of the scenarios ranking Country 5 fifth or sixth. Thus, the results for Country 5 should be interpreted with caution.

Since the illustration is based on imaginary data, a detailed interpretation cannot be given. However, it should be noted that a detailed analysis can be done once real scenarios have been considered.

Once the uncertainty analysis is performed, we may assess the impact of each scenario on ranking. For this purpose, for each of the ten imaginary countries, we can compute the absolute difference between the original rank (derived from the baseline scenario) and a different scenario. We can then summarize the differences across all ten countries by using the 50th and 90th percentiles and the Spearman rank correlation coefficient.

Let us consider three alternative scenarios. The following table summarizes the differences between ten countries, with different scenarios calculated as described in the previous paragraph.

To obtain the 50th percentile for a second scenario, the process was as follows:

   (i)   We compute the rank of all countries according to the baseline scenario.
   (ii) We compute the rank of all countries according to the second scenario.

   (iii) For each country, we compute the absolute difference between the original rank and the rank derived from the second scenario.
   (iv) We compute the 50th percentile for the distribution of absolute differences.

Interpreting table A.23

The second line of the table tells us that scenario 1 uses equal weighting as a weighting scheme, an arithmetic aggregation rule, and does not include the subdimension "Funding" in the computation of the "Capacity for an efficient statistical system" indicator. Since the 50th percentile is 4, it therefore indicates that 50 per cent of the countries in the sample, which is five in this case, experience a shift of more than four positions as compared to the baseline scenario. Since the 90th percentile is 11, it therefore indicates that 10 per cent of the countries in the sample, which is one country in this case, experience a shift of more than 11 positions as compared to the baseline scenario. An examination of the previous table shows that scenario 1 has the highest impact given that the 50th and the 90th percentiles are the highest of all scenarios. For all of the imaginary scenarios, the Spearman rank coefficients with the baseline are high, positive and close to 1. These results indicate that those scenarios tend to give results that are consistent with the original model.

Annex 5. Glossary and references

1. Regulatory and instructional bodies and instruments

- **Abuja Treaty:** Adopted in Abuja on 3 June 1991 by what was then known as the Organization of African Unity and is now known as the African Union. The main objectives of the Treaty are to promote economic, social and cultural development and the integration of African economies; to establish, on a continental scale, a framework for the development, mobilization and utilization of human and material resources in Africa; and to promote cooperation in all fields of human endeavour. The Treaty entered into force on 12 May 1994. The text of the Treaty is available at: https://au.int/sites/default/files/treaties/7775-treaty-0016_-_treaty_establishing_the_african_economic_community_e.pdf.

- **African Charter on Statistics:** Adopted in Addis Ababa on 4 February 2009 by what was then known as the Organization of African Unity and is now known as the African Union. The Charter lists the following 10 main objectives: to serve as a policy framework for statistics development in Africa; to serve as advocacy tool; to ensure quality statistics for monitoring economic and social integration in Africa; to promote adherence to fundamental principles in the production, storage, management, dissemination and use of statistical information; to enhance coordination of statistical activities; to build the institutional capacity of statistical authorities in Africa; to serve as a reference framework for the exercise of the African statistician profession; to promote a culture of evidence-based policy making, monitoring and evaluation; to help enhance the functioning of the African statistical system; and to ensure that there is no duplication of efforts in the implementation of statistics programmes. The text of the Charter is available at https://paris21.org/sites/default/files/AU-English_African_Charter-web.pdf.

- **African Statistics Day:** This day is celebrated on 18th November. This Day was adopted in May 1990 by the Sixteenth Meeting of African Ministers responsible for Finance, Planning, and Economic Development to be celebrated each year in order to increase public awareness of the important role played by statistics in all aspects of the economic, social and cultural lives of citizens in African countries. Further information on the Day is available at: https://ecastats.uneca.org/acsweb/Home/AfricanStatisticsDay.aspx.

- **Agenda 2063:** This is a strategic framework for the socioeconomic transformation of the continent until 2063. The Agenda builds on and seeks to accelerate the implementation of past and existing continental initiatives for growth and sustainable development. Some of the past and current initiatives it builds on include: the Lagos Plan of Action, the Abuja Treaty, the Minimum Integration Programme, the Programme for Infrastructural Development in Africa, the Comprehensive Africa Agricultural Development Programme, the New Partnership for Africa's Development, as well as national and regional plans and programmes. The Agenda is available at www.un.org/en/africa/osaa/pdf/au/agenda2063.pdf.

- **Fundamental Principles of Official Statistics:** To ensure that national statistical systems produce appropriate and reliable data that adhere to certain professional and scientific standards, the Conference of European Statisticians developed and adopted ten Fundamental Principles of Official Statistics in 1991. Statisticians in other parts of the world soon realized that the principles were of much wider, global significance. The United Nations Statistical Commission then adopted the very same set of principles – with a revised preamble – as the United Nations Fundamental Principles of Official Statistics in April 1994. At its forty-second session, held in 2011, the Statistical Commission acknowledged that the Principles were still

relevant and that no revision of the Principles themselves was necessary. The Commission recommended, however, that the Friends of the Chair of the Statistical Commission revise and update the preamble of the Fundamental Principles in order to take into account new developments. At its forty-fourth session, held in 2013, the Statistical Commission adopted the revised preamble. In 2018 the Friends of the Chair of the Statistical Commission revisited the Principles and conducted a global survey, which was to be presented to the Statistical Commission in March 2019. The Principles are available at https://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx.

- **Committee of Directors General of National Statistical Offices:** The Committee was established in 2006 by African Union Conference of Ministers of Economy and Finance It meets every year under the aegis of the African Union Commission to discuss statistical development issues in Africa, including progress made in the implementation of the African Charter on Statistics and the Strategy for the Harmonization of Statistics in Africa. Members of Committee include Directors General of National Statistics Offices. Further information is available at https://au.int/en/newsevents/20151126/9th-committee-directors-generals-national-statistics-office-codgs.

- **Comprehensive Africa Agricultural Development Programme:** The Programme is Africa's policy framework for agricultural transformation, wealth creation, food security and nutrition, economic growth and prosperity for all. Further information is available at www.un.org/en/africa/osaa/peace/caadp.shtml.

- **General Data Dissemination System:** The System provides members with a basic framework for a broad national statistical development strategy. It covers a set of statistics needed by all countries in their policymaking and analysis. Further information is available at www.imf.org/en/Publications/Manuals-Guides/Issues/2016/12/31/The-General-Data-Dissemination-System-Guide-for-Participants-and-Users-40365.

- **Lagos Plan of Action for the Economic Development of Africa:** The Plan underscores the need for "breakdowns by sex in order to identify areas for action relevant to women, particularly in the sub-theme areas of employment, health and education. Lack of such sex breakdowns has in the past hindered the identification of special needs of women." To enhance statistics in the continent, the plan states that "As a matter of urgency, Member States must strengthen their statistical infrastructures as bases for effective policymaking and planning" The text of the Plan of Action is available at https://web.archive.org/web/20070106003042/http://uneca.org/itca/ariportal/docs/lagos_plan.PDF.

- **Marrakech Action Plan for Statistics:** Together with the Partnership in Statistics for Development in the 21st Century, and the Trust Fund for Statistical Capacity Building, the Marrakech Action Plan for Statistics aims to develop statistical capacity in developing countries. Further information is available at www.worldbank.org/en/data/statistical-capacity-building/marrakech-action-plan-for-statistics.

- **Minimum Integration Programme:** The Minimum Integration Programme, which was elaborated in close collaboration with the regional economic communities, was adopted by Ministers of Integration in 2009. This Programme includes different activities to speed up and bring to a successful conclusion the process of regional and continental integration. The Programme will be implemented by the regional

economic communities, Member States and the African Union Commission, in collaboration with Africa's development partners. Further information is available at https://au.int/en/ea/ric/mip.

- **National statistical council or board:** A body constituted by statistic authorities to address the challenges faced by statistical agencies in the country in relation to the production of data, coordination of the national statistical system, and the collection of objective data so as to restore public trust in the figures released by the government. It may be a decision-making body chaired by the prime minister, a minister or another senior figure in government, or a consultative body to a national statistical institute.

- **National statistical office:** The main government department responsible for the collection, processing and dissemination of official statistics in the country.

- **National statistical systems:** the ensemble of statistical organizations and units within a country that jointly collect, process and disseminate official statistics on behalf of the national Government.

- **New Partnership for Africa's Development (NEPAD):** A programme of the African Union established with the aim of eradicating poverty, fostering growth and sustainable development in African countries, strengthening Africa's status in the global economy and political landscape and empowering women. Further information is available at www.nepad.org/.

- **National Strategy for the Development of Statistics:** This is expected to provide a country with a strategy for developing statistical capacity across the entire national statistical system. The Strategy will provide a vision for where the national statistical system should be in five to ten years and will set milestones for getting there. It will present a comprehensive and unified framework for the continual assessment of evolving user needs and priorities for statistics and for building the capacity needed to meet these needs in a more coordinated, synergistic and efficient manner. It will also provide a framework for mobilizing, harnessing, and leveraging resources (both national and international) and a basis for the effective and results-oriented strategic management of the national statistical system. Further information is available at www.paris21.org/NSDS.

- **Partnership in Statistics for Development in the 21st Century, Statistical Capacity-Building Indicators:** These 34 qualitative and quantitative indicators address resources, inputs and statistical products and look at the statistical environment, the statistical process, and the relevance of a country's statistical products to users' needs. Further information is available at https://paris21.org/sites/default/files/scbi-final-en.pdf.

- **Programme for Infrastructural Development in Africa:** Designed to develop a vision and strategic framework for the development of regional and continental infrastructure, including in the areas of energy, transport, ICT and transboundary water resources). The overall goal of the Programme is to promote socioeconomic development and poverty reduction in Africa through improved access to integrated regional and continental infrastructure networks and services. Further information is available at www.afdb.org/en/topics-and-sectors/initiatives-partnerships/programme-for-infrastructure-development-in-africa-pida.

- **The African Statistics System:** A partnership composed of national statistical systems, including data providers, producers and users, statistics research and training institutes and statistics coordination bodies, as well as statistics units in the statistics organizations of regional economic communities, regional training centres, and the statistics units of continental organizations and coordination bodies at the continental level.

- Qualification criteria of the OECD, Giuseppe Nicoletti, Stefano Scarpetta and Olivier Boylaud (2000): This paper presents a database of indicators for most OECD countries. The paper illustrates a methodology on how to derive the weights of the different components for aggregating these detailed indicators into summary indicators. the paper is available at www.oecd.org/eco/outlook/1880867.pdf.

- **Reference Regional Strategic Framework for Statistical Capacity-Building in Africa**: The Framework is based on information obtained from various sources, including reports, earlier assessments of country statistical capacity, and an assessment of and consultation with more than 20 countries and 40 subregional, regional, and international organizations and donors' agencies undertaken in 2005. Further information is available at www.uneca.org/publications/reference-regional-strategic-framework-statistical-capacity-building-africa.

- **Strategy for the Harmonization of Statistics in Africa (second generation):** Information on the Strategy is available at www.africa.undp.org/content/rba/en/home/library/reports/strategy_for_the_harmonization_of_statistics_in_africa_shasa_governance_peace_and_security_gps_data.html.

- **Sustainable Development Goals:** The Sustainable Development Goals are the blueprint to achieve a better and more sustainable future for all. They address the global challenges we face, including those related to poverty, inequality, climate, environmental degradation, prosperity, and peace and justice. The aim is to achieve the 17 Goals by 2030. Further information is available at www.un.org/sustainabledevelopment/sustainable-development-goals/.

- **Statistical Commission for Africa-Committee of Directors General of National Statistical Offices (StatCom-Africa/CoDG):** To efficiently serve its member States in different areas, ECA has established a number of subsidiary bodies as part of its intergovernmental machinery. In April 2007, the Conference of Ministers endorsed the creation of the new subsidiary body in charge of statistics, namely the Statistical Commission for Africa (StatCom-Africa). Further information is available at www.uneca.org/statcomdodgggim-africa.

- **Special Data Dissemination Standard (SDDS):** Established by the IMF in 1996, SDDS guides countries that have, or might seek, access to international capital markets on how best to make their economic and financial data available to the public. Further information is available at http://dsbb.imf.org/Pages/SDDS/Home.aspx.

- **World Bank Statistical Capacity Indicator Dashboard:** This provides an overview of the statistical capacities of over 140 developing countries. It aims to assess the capacity of national statistical systems

and to monitor progress in strengthening statistical capacity over time. Further information is available at http://datatopics.worldbank.org/statisticalcapacity/.

2. Definitions of terms

A

- **Accuracy:** See "Data accuracy".

- **Administrative data:** "Data collected by, or on behalf of, national and local authorities other than a producer of official Statistics, for administrative purposes in conformity with legal bases other than statistical legislation" (Economic Commission for Europe (ECE), Generic Law on Official Statistics, available at www. unece.org/fileadmin/DAM/stats/publications/2016/ECECESSTAT20163_E.pdf). Administrative data also refer to "The set of units and data derived from an administrative source. (OECD Glossary of Statistical Terms, available at www.oecd- ilibrary.org).

- **Administrative data providers:** National and local authorities that provide producers of official statistics with data collected primarily for administrative purposes.

- **Administrative source:** "The organizational unit responsible for implementing an administrative regulation (or group of regulations), for which the corresponding register of units and the transactions are viewed as a source of statistical data." (OECD Glossary of Statistical Terms).

- **Analytical work (inside the national statistical office):** By analytical work within the national statistical office, this manual focuses on the following: (a) Exploring whether or not the publications of the national statistical office have an analytical content in the statistical sense; (b) Ascertaining the capacity of national statistical office staff to use innovative data sources for the production of official statistics, and; (c) Ascertaining whether there is an operational statistics, economics or social research unit.

- **Annual statistical programme:** An operational translation of the multi-year statistical programme within the national statistical system for updating the list of producers of official statistics and providing a legal basis for:

  (a) All official statistics to be released;
  (b) All statistical surveys to be carried out by producers of official statistics;
  (c) All transmissions of administrative data or data from other existing sources to producers of official statistics;
  (d) Main development activities in the area of official statistics;
  (e) Statistical registers to be maintained and developed.

- **Anonymized data:** "Data containing only anonymized records". An anonymized record is "a record from which direct identifiers have been removed" (OECD Glossary of Statistical Terms).

- **Autonomous institute/office/agency:** A body is described as autonomous when it is completely self-governing (see also "Semi-autonomous institute/office/agency" and "Non-autonomous institute/office/agency").

B

- **Baseline data:** Data available at the start of the period covered by the national implemented project or strategy, relating directly to outputs and outcomes targeted by that project or strategy, which serves as a basis for comparison with subsequently acquired data.

- **Benchmark:** A "reference point or standard against which performance or achievements can be assessed" (OECD, Glossary of Key Terms in Evaluation and Results Based Management, available at www.oecd.org/dac/evaluation/2754804.pdf.

- **Big machine:** A computer specifically designed with a large memory that can store and manipulate large amounts of data in a single place at one time.

- **Business register (foreign direct investment):** "A register of enterprises or establishments involved in foreign direct investment that is maintained by countries to assist in the compilation of their direct investment data" (OECD Glossary of Statistical Terms).

C

- **Centralized databank:** A database that is located, stored and maintained in a single location.

- **Code:** "A language-independent set of letters, numbers or symbols that represent a concept whose meaning is described in a natural language" (OECD Glossary of Statistical Terms).

- **Cold deck imputation:** Replacing the missing value with a value from an external source, e.g. from a previous realization of the same survey.

- **Consistency:** "Logical and numerical coherence" (OECD Glossary of Statistical Terms).

- **Coordination mechanisms:** These include interagency committees, resolutions on policies and standards, designations of statistical activities, and statistical development programmes. See, for example, Candido J. Astrologo, Jr., Statistical Coordination Mechanisms Towards Measuring the Information economy: The Philippine Experience, available at https://unstats.un.org/unsd/economic_stat/ICTKorea/Documents/Astrologo_Philippines.pd.

- **Correlation coefficient:** "A measure of the degree to which two variables tend to move together. The coefficient has a value between plus and minus 1, which indicates the strength and direction of association" (OECD Glossary of Statistical Terms).

- **Correlation matrix:** The correlation matrix of a data table is a matrix of the correlation coefficient of variables in the table. The correlation coefficient indicates the strength of the linear connection between two continuous variables. This coefficient takes values between -1 and 1. The closer the coefficient to a value of one, the more the linear link between the variables X and Y and the stronger the positive correlation: that is to say that a high value of X will be associated with a high value of Y. The closer the coefficient is to a value of -1, the more the negative correlation between the variables X and Y: that is to say a high value of X will be associated with a low value of Y. If the coefficient takes a value close to 0, this indicates that there is no linear connection between the variables X and Y. The relation between X and Y can be of another order but, in any case, it is not linear. However, this does not mean that the variables are independent.

D

- **Data:** "Characteristics or information, usually numerical, that are collected through observation. Context: Data is the physical representation of information in a manner suitable for communication, interpretation, or processing by human beings or by automatic means." (OECD Glossary of Statistical Terms).

- **Data accuracy:** "Closeness of computations or estimates to the exact or true values that the statistics were intended to measure. (…) The accuracy of statistical information is the degree to which the information correctly describes the phenomena it was designed to measure. It is usually characterized in terms of error in statistical estimates and is traditionally decomposed into bias (systematic error) and variance (random error) components" (OECD Glossary of Statistical Terms). Ensuring accuracy "requires proper concern for consistency across geographical areas and across time, as well as statistical measures of errors in the data". United Nations Statistics Division, Handbook of Statistical Organization, Third edition, available at https://unstats.un.org/unsd/publication/SeriesF/SeriesF_88E.pdf.

- **Data analysis:** "The process of transforming raw data into usable information, often presented in the form of a published analytical article, in order to add value to the statistical output." (OECD Glossary of Statistical Terms).

- **Data Archival:** "Data archiving is the process of retaining data for long term storage. The data might not be in use, however, it can be brought into use and can be stored for future purposes". (Techopedia, available at www.techopedia.com/definition/14677/data-archiving).

- **Database:** "A logical collection of information that is interrelated and that is managed and stored as a unit, for example in the same computer file. The terms database and dataset are often used interchangeably" (OECD Glossary of Statistical Terms).

- **Data precision:** "The property of the set of measurements of being very reproducible or of an estimate of having a small random error of estimation. Context: Precision is to be contrasted with accuracy, which is the property of being close to some target or true value. Precision is a quality associated with a class of measurements and refers to the way in which repeated observations conform to themselves; and in a somewhat narrower sense refers to the dispersion of the observations, or some measure of it, whether

or not the mean value around which the dispersion is measured approximates the "true" value." (OECD Glossary of Statistical Terms).

- **Data anonymization process/system**: See "Anonymized data".

- **Data collection:** "The process of gathering data". Context: Data collection encompasses such concepts as:

  (a) The type(s) of interview used for data collection (e.g. personal or by telephone, paper and pencil, facsimile, computer-aided personal or telephone interview (CAPI/CATI), or mailed questionnaires);
  (b) The duration of the field work (specify the dates); the period used for data collection;
  (c) Whether a permanent survey organization exists or personnel for each survey round is recruited, etc.

Data may be observed, measured or collected by means of questioning, as in survey or census response." (OECD Glossary of Statistical Terms).

- **Data collection tools:** "The methodologies used to identify information sources and collect information during an evaluation". Examples are informal and formal surveys, direct and participatory observation, community interviews, focus groups, expert opinion, case studies, literature search." (OECD Glossary of Statistical Terms).

- **Data confidentiality:** "A property of data, usually resulting from legislative measures, which prevents it from unauthorized disclosure". (OECD Glossary of Statistical Terms).

- **Data development:** Synonymous to statistical development, this term is used when designing and implementing a National Strategy for the Development of Statistics. In this manual, data development refers to the concept as expressed in the Reference Regional Strategic Framework for Statistical Capacity-Building in Africa, available at www.uneca.org/sites/default/files/PublicationFiles/stat_rrsfdocument_final.pdf.

- **Data Dissemination:** "The release to users of information obtained through a statistical activity." (OECD Glossary of Statistical Terms).

- **Data editing:** "The activity aimed at detecting and correcting errors (logical inconsistencies) in data. Context: Editing techniques refers to a range of procedures and processes used for detecting and handling errors in data. Examples of different techniques include the different approaches to editing such as micro-editing/ macro-editing, input/output editing, or to the various tools available for editing such as graphical editing, interactive editing, etc. Edit types refer to the actual nature of edits applied to data during input or output processing. These include:

  (a) Validation edits - to check the validity of basic identification of classificatory items in unit data;
  (b) Logical edits - ensure that two or more data items do not have contradictory values;
  (c) Consistency edits - check to ensure that precise and correct arithmetic relationships exist between two or more data items;

(d) Consistency edits - check to ensure that precise and correct arithmetic relationships exist between two or more data items;

(e) Range edits - identify whether or not a data item value falls inside a determined acceptable range;

(f) Variance edits - involve looking for suspiciously high variances at the output edit stage;

Edit types may also refer to whether these edits are fatal or query type, i.e. whether they detect errors with certainty or point to suspicious data items. Micro-editing and macro-editing may be distinguished in order to calculate the rate of edits." (OECD Glossary of Statistical Terms).

- **Data processing:** In general, data processing refers to the collection of items of data to produce meaningful information and manipulating that data to extract required information in an appropriate form.

- **Data standardization:** "The process of reaching agreement on common data definitions, formats, representation and structures of all data layers and elements." (OECD Glossary of Statistical Terms).

- **Data storage:** "A general term for archiving data in electromagnetic or other forms for use by a computer or device. Different types of data storage play different roles in a computing environment. In addition to forms of hard data storage, there are now new options for remote data storage, such as cloud computing, that can revolutionize the ways that users access data". (Techopedia, available at www.techopedia.com/definition/23342/data-storage).

- **Data validation:** "An activity aimed at verifying whether the value of a data item comes from the given (finite or infinite) set of acceptable values." "Validation" may be defined as: "A continuous monitoring of the process of compilation and of the results of this process. In Statistical Data and Metadata Exchange (SDMX), "Validation" describes methods and processes for routinely assessing source data – including censuses, sample surveys, and administrative records – and how the results of the assessments are monitored and made available to guide statistical processes." (OECD Glossary of Statistical Terms).

- **Dissemination policy:** Established by the national statistical office, the dissemination policy should:

  (a) Adopt transparent procedures and apply to the entire national statistical system;

  (b) Ensure that unified terminology is used in the dissemination of all official statistics.

- **Data dissemination strategy:** In the context of this manual, the strategy should: uphold the right of the public to information; adhere to international standards on impartiality, timeliness, quality and comparability; adhere to international standards on data dissemination; facilitate accessibility by the public to statistical data, including microdata; ensure compliance with the release calendar for official statistics and facilitate the publication and updating of online statistical databases; identify users' needs and employ appropriate ITC to meet those needs; take corrective measures in case of published errors; and conduct user satisfaction surveys and engage with users with a view to enhancing the strategy.

- **Data security:** "The measures taken to prevent unauthorized access or use of data. (OECD Glossary of Statistical Terms).

E

- **Efficiency:** "Achieving maximum output from a given level of resources used to carry out an activity." (OECD Glossary of Statistical Terms).

- **Eigenvalue:** in the principal components analysis method, one step is to transform the correlation matrix through the linear algebra method of diagonalization. The outputs of that process are called eigenvalues, which represent the inertia (variance) of a factor.

- **Evaluation:** "The systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation, and results. The aim is to determine the relevance and fulfilment of objectives, development efficiency, effectiveness, impact, and sustainability." (OECD Glossary of Statistical Terms).

- **Evaluation of data requirements:** This manual focuses on conducting systematic evaluations of data needs arising from the following: (a) The national development plan or equivalent; (b) The Sustainable Development Goals; (c) Agenda 2063 of the African Union; (d) The data needs of different government departments; (f) The needs of other users (private sector, civil society, etc.).

F

- **Factor:** In the principal components analysis method (as outlined in this manual) a Factor, although not directly observable, can be defined as a principal component that, in turn, represents the linear combination of the original variables.

- **Frequency, periodicity and timeliness on the release of indicators:** In this manual the frequency at which an indicator is released is simply the rate at which the release happens or is repeated. Periodicity refers to the pattern of release (which may be monthly, quarterly or annual). The terms frequency and periodicity are used interchangeably in this manual The periodicity defines the reference period (for example, if the periodicity is quarterly then within 12 consecutive months there are only 4 reference periods). Knowing the reference period and the actual release date, we can evaluate the timeliness as the difference between the actual release date and the reference period. For example, if a reference period concludes on 15 January but the actual release date is 15 March, the timeliness is 2 months or 60 days.

G

- **Geographical information system**: "A geographical information system can be seen as a system of hardware, software and procedures designed to support the capture, management, manipulation, analysis, modelling and display of spatially referenced data." (OECD Glossary of Statistical Terms).

- **Geo-referenced data:** Data with geographic location information included, such as latitude and longitude.

- **Good governance:** "Good governance is characterised by participation, transparency, accountability, rule of law, effectiveness, equity, etc." (OECD Glossary of Statistical Terms, available at https://stats.oecd.org/glossary/detail.asp?ID=7237).

H

- **Hot-deck imputation:** Filling in blank cells with individual data, drawn from "similar" responding units. For example, missing values for individual income may be replaced with the income of another respondent with similar characteristics, e.g. age, sex, race, place of residence, family relationships, job, etc.

I

- **Imputation of missing data (Data imputation):** "The substitution of estimated values for missing or inconsistent data items (fields). The substituted values are intended to create a data record that does not fail edits." (OECD Glossary of Statistical Terms).

- **Index of economic well-being:** An index consisting of consumption, wealth, equality, and economic security. The final aggregation of the index was compiled using observed weighting schemes including equal weight, expert weighting, factor analysis, and regression analysis. The index outlined both advantages and shortcomings of each weighting scheme. For more details, see: An assessment of weighting methodologies for composite indicators: the case of the index of economic well-being, Andrew Sharpe and Brendon Andrews, CSLS Research Report No. 2012-10, available at www.csls.ca/reports/csls2012-10.pdf.

- **Indicator:** A quantitative or qualitative factor or variable that provides a simple and reliable means to measure achievement, to reflect the changes connected to an intervention, or to help assess the performance of a development actor." (OECD, Glossary of Key Terms in Evaluation and Results Based Management).

- **International standards:** Should be determined only by statistical considerations and consider best practice. Without common standards and metadata, comparability of data produced by different agencies would be impossible. This applies equally within a country, and between countries (see Fundamental Principles of Statistics: Implementation Guidelines, available at https://unstats.un.org/unsd/statcom/50th-session/documents/BG-Item3b-FPOS-Implementation-guidelines-E.pdf). See also "Quality assurance framework".

- **IT master plan:** Guides an organization in formulating its overall IT strategy. While an IT strategy focuses on how IT will help the business succeed, an IT master plan is a roadmap to help the business implement that strategy.

- **IT security policy:** According to the Fundamental Principles of Statistics: Implementation Guidelines, IT security policy should cover the following three dimensions: (a) Availability: statistical agencies provide internal and external users access to data to the required extent: (b) Integrity: statistical agencies secure adequate survey methods and processing methods and guarantee that data are not falsified by human

or technical misbehaviour; (c) Confidentiality: statistical agencies assure that provisions regarding confidentiality and data protection are guaranteed and that data are only used for statistical purposes. Furthermore, the IT-security policy should be widely known to the staff of the statistical authority.

L

- **Legal framework (Legislative and regulatory framework for statistics):** A law or formal provision in force stipulates that statistical agencies are professionally independent and impartial, develop, produce and disseminate statistics in accordance with professional standards, and treat all users equitably. An example of a legal framework is the African Charter on Statistics. The laws and regulations governing official statistical activities within a country should be clear, comprehensive and up-to-date. Statistical laws and regulations should be coherent with other laws and regulations governing the activities of the national statistical office and the national statistical system.

- **Local Area Network (LAN):** "A network connecting computers in a small area, typically an office or building". (IT Career Glossary, available at www.itcareerfinder.com/brain-food/it-career-glossary.html).

M

- **Metadata** "Data and other documentation that describe statistical data and statistical processes in a standardized way by providing information on data sources, methods, definitions, classifications and data quality." (ECE, Generic Law on Official Statistics).

- **Monitoring:** "A continuing function that uses a systematic collection of data on specified indicators to provide management and the main stakeholders of an ongoing development intervention with indications of the extent of progress and achievement of objectives and progress in the use of allocated funds." (OECD, Glossary of Key Terms in Evaluation and Results Based Management).

- **Multi-year programme:** Establishes guidelines for the strategic development of the national statistical system in with a view to addressing existing and emerging user needs in the context of available resources.

N

- **National statistical coordinating body:** The highest policymaking and coordinating body on statistical matters.

- **National Statistical Office:** "The main producer of official statistics (…) and responsible for coordinating all activities for the development, production, and dissemination of official statistics within the National Statistical System". (ECE, Generic Law on Official Statistics).

- **National statistical system:** Consisting of producers of official statistics within a country, the national statistical system. comprises of the following two components:

(a) The national statistical office, which is the leading authority of the national statistical system;

(b) Other Producers of Official Statistics, consisting of organizational entities of national authorities that develop, produce and disseminate official statistics (ECE, Generic Law on Official Statistics). Alternatively, the national statistical system is "the ensemble of statistical organizations and units within a country that jointly collect, process and disseminate official statistics on behalf of the national government." (OECD Glossary of Statistical Terms).

Non-autonomous institute/office/agency: A body is described as non-autonomous when has no self-governing mechanism.

O

- **Official statistics:** Statistics describing on a representative basis the economic, demographic, social and environmental phenomena of a given country. In this manual, official statistics shall be developed, produced and disseminated in compliance with the provisions and principles of the African Charter on Statistics and the Fundamental Principles of Official Statistics (see ECE, Generic Law on Official Statistics). Alternatively, they may be defined as "statistics disseminated by the national statistical system, excepting those that are explicitly stated not to be official." (OECD Glossary of Statistical Terms).

- **Other producers of official statistics:** These producers are "responsible for the development, production and dissemination of official statistics, as assigned in the annual statistical programme. The functions shall be carried out in full conformity with statistical legislation and standards…" (ECE, Generic Law on Official Statistics).

- **Operational statistical development fund:** The functioning funds allocated for the studies and statistical capacity-building of the national statistical system.

P

- **Performance appraisal:** "Assessment against a set of predetermined criteria of the efficiency and effectiveness with which an individual fulfils an agreed set of tasks. Such appraisals are frequently used in assessing whether managers should qualify for pay increases or promotion. (OECD Glossary of Statistical Terms).

- **Periodicity:** "The frequency of compilation of the data." (OECD Glossary of Statistical Terms).

- **Poverty Reduction Strategy Paper (PRSP):** "The PRSP is a comprehensive, nationally owned strategy that is prepared by the borrowing country and endorsed in their respective areas of responsibility by the Boards of the IMF and World Bank. Funds are provided at an annual interest rate of 0.5 percent. They are repayable over 10 years, including a grace period of 5 1/2 years." (OECD Glossary of Statistical Terms).

- **Pricing regulation (policy):** "Price regulation refers to the policy of setting prices by a government agency, legal statute or regulatory authority. Under this policy, minimum and/or maximum prices may be set." (OECD Glossary of Statistical Terms).

- **Principal Component Analysis:**[22][23] is a descriptive statistics technique leading to approximate, but in a certain sense optimal, graphic representations of the contents of a data table: these provide a simultaneous overview of the links between variables and similarities between individuals. It is also a tool that can be used to reduce the dimensionality of a set of continuous variables, so that they can be used as an intermediary of calculation for further analyses.

- **Principal components analysis/factor analysis:** This analysis may be used to estimate composite indicators, assuming that the variables have different weights. Further details are provided in the OECD Handbook on Constructing Composite Indicators, Methodology and User Guide, available at www.oecd.org/std/42495745.pdf.

- **Physical and material infrastructure:** The building housing the national statistical office headquarters, the working spaces for staff and the vehicles used to conduct field activities.

Q

- **Quality assurance framework:** In the context of this manual, the framework should cover the following:

  (a) the statistical system
  (b) the institutional environment
  (c) statistical processes
  (d) statistical outputs

The framework should also be accessible to the public;

According to Holupka (2013), the quality assurance framework can be viewed as a "toolkit" or an "umbrella", and a single place to record, reference and organize the full range of quality concepts, policies, tools, and practices".[23][24]OECD defines the IMF data quality assurance framework (DQAF) as an assessment methodology that aims to provide structure and a common language for the assessment of data quality. The DQAF facilitates a comprehensive view of data quality, one that recognizes interrelations, including trade-offs, among elements of quality and allows emphases to vary across data categories and uses/users. It facilitates dialogue with national statistical agencies and country authorities, as well as a more homogenous approach to assessing data quality by Fund staff. (OECD Glossary of Statistical Terms). In 2012, the United Nations Statistical Commission endorsed a generic national quality assurance framework template to provide the general structure within which countries can develop and enhance their national assurance frameworks.

- **Quality assurance:** "A planned and systematic pattern of all the actions necessary to provide adequate confidence that a product will conform to established requirements." (OECD Glossary of Statistical Terms).

---

22 For further details please see the following: www.stat.ncsu.edu/people/bloomfield/courses/ST810J/slides/pca.pdf; www.cse.psu.edu/~rtc12/CSE586Spring2010/lectures/pcaLectureShort.pdf; www.cs.princeton.edu/picasso/mats/PCA-Tutorial-Intuition_jp.pdf and www.utdallas.edu/~herve/abdi-awPCA2010.pdf.
23 Mary Jane Holupka, National quality assurance frameworks, United Nations Statistics Division (2013). Available at www.unsiap.or.jp/programmes/sqm_materials/sqm6/4_Overview_NQAFs.pdf.

- **Quality assessment and certification processes**: "Guarantee the official nature of statistics produced in various parts of the national statistical system." (Recommendation of the OECD Council on Good Statistical Practice, available at www.oecd.org/statistics/good-practice-toolkit/.

- **Quality of official statistics:** "The Producers of Official Statistics shall be committed to continually assessing and improving the quality of official statistics in terms of relevance, accuracy, reliability, timeliness, punctuality, transparency, clarity, coherence and comparability." (ECE, Generic Law on Official Statistics).

- **Quality policy:** A quality policy ensures that the producers of official statistics systematically assess the quality of official statistics. The quality policy is publicly available through guidelines, frameworks, reports, etc. and staff members receive appropriate training for their application. (Recommendation of the OECD Council on Good Statistical Practice).

R

- **Regulatory framework**: see "Legal framework".

- **Register:** "A set of files (paper, electronic, or a combination) containing the assigned data elements and the associated information." (OECD Glossary of Statistical Terms).

- **Release:** "The dissemination activity by which official statistics, including revised statistics, become publicly known for the first time." (ECE, Generic Law on Official Statistics).

- **Release calendar:** "A statement on the schedule of release of data in terms of periodicity and timeliness." (OECD Glossary of Statistical Terms).

- **Reliability:** "The closeness of the initial estimated value(s) to the subsequent estimated values." (OECD Glossary of Statistical Terms).

- **Rotation:** rotation is a method[24][15] used to further analyse initial principal component analysis results with the goal of making the pattern of loadings clearer, or more pronounced. This process is designed to reveal a simple structure.

S

- **Statistical Analysis System:** Software that captures, stores, modifies, analyses and presents data. This software can be used in the analysis of cross-sectional surveys or time series data.

- **Statistical Package for the Social Sciences:** This statistical software in numerous fields of statistical analysis, including factor analysis, cluster analysis and geospatial analysis. This software can also be used in the analysis of cross-sectional surveys or time series data.

---

24 For further information, see: http://hosted.jalt.org/test/PDF/Brown31.pdf.

- **STATA:** This software can manipulate data by exploring, visualizing and modelling data sets. This software can be used in the analysis of cross-sectional surveys or time series data.

- **Semi-autonomous institute/office/agency:** A body that has a degree of, but not complete, self-government.

- **Standards:** "Defining and establishing uniform specifications and characteristics for products and/or services." (OECD Glossary of Statistical Terms).

- **Standards classification:** "Standard classifications are those that follow prescribed rules and are generally recommended and accepted. They aim to ensure that information is classified consistently regardless of the collection, source, point of time, etc." (OECD Glossary of Statistical Terms).

- **Statistical Methodologies:** "Theory and methods of data collection, processing, and analysis." (OECD Glossary of Statistical Terms).

- **Standardized procedures (or data standardization):** "The process of reaching agreement on common data definitions, formats, representation and structures of all data layers and elements. (OECD Glossary of Statistical Terms).

- **Statistical analysis:** Is the process by which an inference is made for a population on the basis of a sample.

- **Statistical infrastructure:** A unit/section or an internal arrangement within the national statistical system that is responsible for the development of standard classification, a statistical register and a geographical information system. The term also refers to the statistical software and electronic devices used by the national statistical office to collect data during censuses and national surveys.

- **Statistical data:** "Data from a survey or administrative source used to produce statistics." (OECD Glossary of Statistical Terms).

- **Statistical law/Act:** A set of legislative measures or the legal mandate to the statistical authorities empowering them to collect data, including, for example, from households, the general public, and administrative records with a view to producing African statistics. More precisely it is the domestic law governing the operation of the country's statistical systems. The law shall provide provisions for transparency in statistical production and on the accessibility of data by users. The law/Act must also be made available to the public. For further information, see the African Charter on Statistics.

- **Statistical programming:** Establishes the various features of the National Strategy for the Development of Statistics or similar multi-year strategy, determines whether the National Strategy for the Development of Statistics or multi-year strategy has been formally adopted, and is to be fully or partially funded, by the government, and ensures the alignment of the National Strategy for the Development of Statistics or strategy with the country's national development plan.

- **Statistical register:** "Lists of statistical units and their characteristics, including identifiers, that are necessary for statistical production." (ECE, Generic Law on Official Statistics).

- **Statistical standard**: "A statistical standard provides a comprehensive set of guidelines for surveys and administrative sources collecting information on a particular topic. Components of a standard include (a) definition(s); (b) statistical units; (c) classification(s); (d) coding process(es); (e) questionnaire module(s)." (OECD Glossary of Statistical Terms).

- **Substitution:** Replacing non-responding units with unselected units in the sample. For example, if a household cannot be contacted, then a previously non-selected household in the same housing block is selected.

- **Sustainability:** "The continuation of benefits from a development intervention after major development assistance has been completed". (OECD, Glossary of Key Terms in Evaluation and Results Based Management).

T

- **Timeliness of data collection:** To ensure timeliness "requires concern for issuing data as frequently as needed to reflect important changes in what is being studied, as well as disseminating data as soon as practicable after they are collected". (United Nations Statistics Division, Handbook of Statistical Organization, Third edition).

U

- **Uncertainty and sensitivity analysis:** Uncertainty analysis refers to a technique that shows how the uncertainty in the input factors or variables aggregated by the composite indicator or index affects the performance of the constructed composite indicator or the index. Sensitivity analysis refers to the study of how the uncertainty in the composite indicator or index can be divided and allocated to different sources of uncertainty in its inputs. For further information, see Andrea Saltelli, "Sensitivity Analysis for Importance Assessment", Risk Analysis, vol. 22 (3) (2002).

- **Unified framework:** A unified framework for the National Strategy for the Development of Statistics is one that addresses user needs and priorities for statistics, including the required capacity building to meet these needs in a coordinated, synergistic and efficient manner. The framework also sets forth how national and international resources should be mobilized and establishes a basis for the effective and results-oriented strategic management of the national statistical system.

- **User needs (for statistics):** "The data and metadata requirements of persons or organizations to meet a particular use or set of uses. Such needs may be specified in terms of the quality dimensions promulgated by international organizations or national agencies." (OECD Glossary of Statistical Terms).

- **User Satisfaction Survey:** "A statistical survey aiming to assess the satisfaction of users of statistics". (OECD Glossary of Statistical Terms).