REPUBLIC OF SOUTH AFRICA

STATISTICIAN-GENERAL

NSS
National Statistics System

QUALITY APPROVED

South African Statistical Quality Assessment Framework (SASQAF)

Statistics South Africa
Preferred supplier of quality statistics

Second edition

South African Statistical Quality Assessment Framework (SASQAF)

Second edition

Foreword

It is with great pleasure that I, on behalf of the Statistics Council, congratulate the Statistician-General, Statistics South Africa (Stats SA) and members of the National Statistics System team with the publication of the South African Statistical Quality Assessment Framework (SASQAF), which was gazetted on September 2009. This framework for the quality improvement of all official statistics will over time prove to be one of the major milestones in building the National Statistics System. It is particularly gratifying to see the extent to which different government departments embraced SASQAF as a guide to improve the quality of their information systems and surveys.

Our challenge remains to build trust in official statistics. Central to this challenge is the improvement of the quality of the data produced in Stats SA and all organs of state. The higher objective is present to our country with very high integrity statistics – capable of informing the development agenda in government and useful in planning and decision-making in the private sector and non-governmental organisations.

To achieve this objective of the delivery of data of high integrity we need to address the question of quality. SASQAF, with its framework of the eight quality dimensions, will play a vital role in achieving this objective. The eight dimensions of quality are:

- Relevance;
- Accuracy;
- Timeliness;
- Accessibility;
- Interpretability;
- Comparability and Coherence;
- Methodological soundness; and
- Integrity.

These dimensions of quality are in line with the requirements of the Statistics Act (No. 6 of 1999) which defines the purpose of official statistics as follows:

"(1)   The purpose of official statistics is to assist organs of state, businesses, other organisations or the public in –
(a)   planning;
(b)   decision-making or other actions;
(c)   monitoring or assessment of policies, decision-making or other actions.

(2)   Official statistics must protect the confidentiality of the identity of, and the information provided by, respondents and be –
(a)   relevant, accurate, reliable and timeous;
(b)   objective and comprehensive;
(c)   compiled, reported and documented in a scientific and transparent manner;
(d)   disseminated impartially;
(e)   accessible;
(f)   in accordance with appropriate national and international standards and classifications; and
(g)   sensitive to distribution by gender, disability, region and similar socio-economic features."

Also, the SAQAF as a framework is in line with the United Nations Fundamental Principles of Official Statistics, which means that there is broad coherence with international best practice.

The most important challenge now is to build the capacity within Stats SA and all organs of state to build a statistical system in our country that will deliver data that government and the broader South African society can trust. This statistical system will be a foundation for a development plan that will deliver a better life for all.

Howard Gabriels
Chairperson: SA Statistics Council

Preface

I am very pleased to write this preface for a second edition of the South African Statistical Quality Assessment Framework (SASQAF). SASQAF has lived up to its original intention and more. Through self-assessment it has also been used in the development of statistical quality in organs of state and other agencies. Most of all it has become the main driver in the implementation of the coordination of statistical production and use in many agencies.

SASQAF has proved to be a main enabler of the full implementation of the Statistics Act (No. 6 of 1999) which mandates the Statistician-General to develop standards for all organs of state and other agencies that produce statistics; to designate as official, statistics or class of statistics produced by any organ of state, including Stats SA; and to comment on any statistics produced in the public domain.

Why a second edition within a period of one-and-a-half years of the publication of the first edition? The second edition is a result of three developments – training in statistical quality of client organs of state and other agencies, a review of the National Survey of Research and Experimental Development (R&D survey) of the Department of Science and Technology (DST), and development of the SASQAF Operational Standards and Guidelines. Since the publication of the first edition of SASQAF and its appearance in the Government Gazette, demand for training on the use of the quality framework for statistical development and certification has outstripped existing resources, forcing development of a rollout strategy. During the training lack of clarity was noted in certain areas of SASQAF, albeit on a small scale. Secondly, a review of the R&D survey also helped identify some areas that needed some minor adjustments for clarity. SASQAF provides for periodic reviews of statistical series that have attained the status of official statistics to ensure that they maintain the good practice.

Assessment renders transparent the process of statistical production and fosters trust in the statistical system as the assessment results are published for access by the public. Thirdly, we noticed that SASQAF on its own was not quite user-friendly as it only identifies and describes what needs to be done but does not indicate how it should be done. To fill this gap we developed generic standards, benchmarks and guidelines with respect to each indicator of SASQAF to guide the development of standards, benchmarks and guidelines pertaining to specific series.

SASQAF is one of the statistical tools at the centre of the statistical system, enabling the implementation process of statistical coordination and thus of the South African National Statistics System (NSS). It is accordingly a driver of both the new strategic direction and strategic shifts in the work-programme of Stats SA and of the monitoring and evaluation aspects of the work-programmes of the rest of the organs of state and other agencies. In this respect it is enabling production of statistics in the country in four specific ways, among others.

First, it facilitates extension of the domain of statistical production to various sub-national levels of the administrative geography of the country, where availability of data is insufficient. The strategic shift and direction at Stats SA lie in a new emphasis on taking up the challenge of statistical leadership through the development and coordination of standards, a challenge which up to now the department has been relatively reluctant to confront. Second, the country-wide adoption of SASQAF is facilitating the linking into one system, the NSS, the various subsystems of statistical production that have largely occurred in isolation. The result will be an increase in the extension of a statistical base relevant to user needs. Third is the shift most directly associated with SASQAF, that of the move by Statistics South Africa from setting standards for its own products alone to setting standards for statistical products of all organs of state within a statistical system. The result is the filling in of the statistical quality gap.

Fourth is the shift from developing statistical capacity (skills and infrastructure) exclusively for itself to developing capacity for all organs of state producing statistics within a systems framework. Statistical production cannot be sustained without a well developed and maintained capacity. In this respect Statistics South Africa is to play a facilitation role as it would neither be feasible nor desirable for the department to provide human resources to each organ of state that produces statistics. This is because the many different organs of state have different mandates to execute, and Statistics South Africa would not be in a position to develop such capacity in its own ranks.

This is indeed a time of great hope in the history of statistical development in South Africa.

PJ Lehohla
Statistician-General

Contents

1.   Introduction

South Africa's first democratically elected government was voted into power with a strong mandate to transform society. Provision of services to the historically dispossessed, education, housing, poverty alleviation, job creation, economic development, and more-equitable distribution of wealth were all identified as high priorities.

However, the new government inherited a statistics void, at least as far as reliable information fit to be used in benchmarking and monitoring progress in service delivery was concerned. Equally, statistics gathered at the time had little value as a basis for informed decision-making, the development of policies, or the planning required for a massive programme of social transformation.

Institutionally, the new government needed to integrate and rationalise the production of national statistics by amalgamating the homeland statistical offices with the then Central Statistical Service (CSS), the statutory agency responsible for national statistics. Then the CSS itself had to be transformed both by reviewing and re-engineering the statistical series it produced, and by transforming its human capacity and statistical infrastructure.

The CSS was transformed into Statistics South Africa (Stats SA), a national government department deriving its new mandate and role from the Statistics Act (Act No.6 of 1999). Although Stats SA was the only institution tasked with producing official statistics, there were – and continue to be – many other producers of statistical information: market research companies, parastatal bodies, government departments, universities and research institutions, and the private sector. In this way, the decentralised and fragmented system of statistical production inherited in 1994 has endured into the present.

The existence of competing statistics which are mobilised in the interests of various contending parties is part of an ordinary and healthy democratic process. However, this does not mean that all statistics are of equal value or quality.

Statistics can and should be assessed in terms of their quality and fitness for required purpose, and this requires common standards and criteria as a basis for evaluation. This is one of the statutory mandates allocated to Stats SA under the Statistics Act.

Although there has been some progress in implementing this mandate, the current state of national statistics is still characterised by
- an information gap in terms of relevant statistics to meet the needs of users;
- a quality gap in terms of common standards, including concepts, definitions, classifications, methodologies and sampling frames; and
- a capacity gap in terms of both human resources and infrastructure.

Stats SA has used its statutory mandate for statistical leadership to develop various strategies aimed at closing these gaps. One of the most important of these involves implementation of a National Statistics System (NSS) to align the use and production of statistics, particularly those collected within the various institutions and organisations of national, provincial and local government, and other organs of state. Statistics are also produced by other institutions and organisations, in the private sector, research bodies and non-governmental organisations. Where these statistics are in the public domain, and where they might have an influence on the development of government policy, or on the measurement and monitoring of government programmes; they too can be evaluated within the framework of the NSS using the South African Statistical Quality Framework (SASQAF).

Alignment and evaluation of statistics requires a rational, transparent and sustainable framework for assessing the quality of those statistics. SASQAF has been developed for this purpose: it provides the framework and criteria used for evaluating and certifying statistics produced by government departments and other organs of state and, in some circumstances, by non-governmental institutions and organisations.

Within the NSS framework, SASQAF draws a distinction between 'official' and 'national' statistics. National statistics refer to those statistics used in the public domain but which the SG has not certified as being official. Official statistics are those statistics that have been certified by the SG as being official in terms of Section 14(7)(a) of the Statistics Act. Certification of statistics produced by organs of state involves a standard assessment procedure undertaken by a Data Quality Assessment Team (DQAT), established by the SG.[1]

For assessment of data for quality to begin, the submitting organ of state and the statistics under review need to comply with three initial criteria:

- The producing agency should be a member of the NSS;
- The statistics need to meet user needs beyond those specific and internal to the producing agency; and
- The statistics produced should be part of a sustainable series, not a once-off collection.

On meeting these initial criteria, assessment of the data begins. The DQAT is required to report on the statistics, classifying them as one of the following:

- quality statistics;
- acceptable statistics;
- questionable statistics; or
- poor statistics.

If the statistics being evaluated are not classified as *quality statistics,* the DQAT is required to make recommendations indicating areas for improvement which might lead to this status. Once the statistics are classified as being quality statistics, in line with the quality dimensions set out in SASQAF, the SG will formally designate the data as *official statistics*, which become subject to periodic reviews determined by the SG in consultation with the head of the producing agency or department.

This book is the culmination of a comprehensive process of learning and consultation, and has benefited from the input of a number of stakeholders. During the revision of this document, some improvements in the quality indicators and dimensions were deemed necessary. The definitions of some of the dimensions were expanded. Coherence was renamed "Comparability and Coherence" since there was no clear distinction made previously between the two concepts. Some indicators were dropped while new, but more, appropriate indicators were developed and others were moved under more suitable dimensions. Changes to these indicators can also be found in the Appendix F.

2.   Purpose of the framework

The main purpose of SASQAF is to provide a flexible structure for the assessment of statistical products. SASQAF can be used for:

- self-assessment by producers of statistics;
- reviews performed by a DQAT in the context of the NSS work;
- assessment by data users (e.g. financial market participants) based on the producing agency's quality declaration;
- assessment by international agencies (e.g. the International Monetary Fund) based on the quality declaration.

---

[1] See Annexure A

3.   Glossary or terms

administrative data: the set of units and data derived from an administrative source

administrative source: a data holding containing information collected and maintained for the purpose of implementing one or more administrative regulations

benchmark: a recognised standard, or a reference point, that forms the basis for assessment or comparison.

catalogue: an ordered list of statistical products available in the organisation

micro-data: observation data collected on an individual object or statistical unit.

data confidentiality: a property of *data*, usually resulting from legislative measures, which prevents it from unauthorised disclosure

data quality: fitness for use of statistical information, i.e. the degree to which a set of inherent characteristics in the statistical data fulfils user requirements; measured in terms of the prerequisites and eight dimensions of quality, namely: relevance, accuracy, timeliness, accessibility, interpretability, comparability and coherence, methodological soundness and integrity

electronic media: dissemination media that allow electronic exchange of data such that software, or a combination of individuals and software, can put the data in a compatible form at the receiving end

estimate: the particular value yielded by an estimator in a given set of circumstances

estimator: a rule or method of estimating a parameter of a population, usually expressed as a function of sample values

guidelines: directions or principles used in the development, maintenance and application of rules; they may or may not be necessarily mandatory, but are provided as an aid to interpretation and use of rules

misclassification: when a subject is falsely classified into a category in which the subject does not belong. It may result from misreporting by study subjects, from the use of less than optimal measurement devices, or from random error

quality indicator: an attribute of statistical information that is used to measure its quality

reference period: the period of time relevant for a particular *question*.

respondent burden: the effort, in terms of time and cost, required for respondents to provide satisfactory answers to a survey

scope: coverage or sphere of what is to be observed. It is the total membership or population of a defined set of people, objects or events

standard: a comprehensive set of guidelines for *survey*s and administrative sources collecting information on a particular topic, including definitions, statistical units, classifications, coding processes, questionnaire modules, and output categories

survey: a process which collects, examines, and reports on data concerning variables of interest for a *reference period*

statistical programme: a programme for producing statistics in a particular socio-economic sphere

4. Definition of data quality

Stats SA defines data quality in terms of 'fitness for use'. Data quality is further defined in terms of prerequisites and the eight dimensions of quality, namely, relevance, accuracy, timeliness, accessibility, interpretability, coherence, methodological soundness and integrity. Five of the eight SASQAF quality dimensions are also covered in the Data Quality Assessment Framework of the International Monetary Fund (IMF).

Prerequisites of quality refer to the institutional and organisational conditions that have an impact on data quality. These include the institutional and legal environment, and availability of human, financial and technological resources.

The *relevance* of statistical information reflects the degree to which it meets the real needs of clients. It is concerned with whether the available information sheds light on the issues of most importance to users.

The *accuracy* of statistical information is the degree to which the output correctly describes the phenomena it was designed to measure. It relates to the closeness between the estimated and the true (unknown) values. Accuracy is measured by means of two major sources of error, namely, sampling error and non-sampling error.

The *timeliness* of statistical information refers to the delay between the reference points to which the information pertains, and the date on which the information becomes available. It also considers the frequency and punctuality of release. The timeliness of information will influence its relevance.

The *accessibility* of statistical information refers to the ease with which it can be obtained from the agency. This includes the ease with which the existence of information can be ascertained, as well as the suitability of the form or medium through which the information can be accessed. The cost of the information may also be an aspect of accessibility for some users.

The *interpretability* of statistical information refers to the ease with which users can understand statistical information through provision of metadata. This information normally includes the underlying concepts, definitions and classifications used the methodology of data collection and processing, and indicators or measures of the accuracy of the statistical information.

The *coherence* of statistical information reflects the degree to which it can be successfully brought together with other statistical information within a broad analytical framework and over time. The use of standard concepts, classifications and target populations promotes coherence, as does the use of common methodology across surveys.

Methodological soundness refers to the application of international, national, or peer-agreed standards, guidelines, and practices to produce statistical outputs. Application of such standards fosters national and international comparability.

The *integrity* of statistical information refers to values and related practices that maintain users' confidence in the agency producing statistics and ultimately in the statistical product.

These dimensions of quality are overlapping and interrelated. Achieving an acceptable level of quality is the result of addressing, managing and balancing these elements of quality over time with careful attention to programme objectives, costs, respondent burden and other factors that may affect information quality or user expectations. Each dimension has to be adequately managed if information is to be fit for use. Failure to comply with any one dimension will impair the usefulness of the information.

5. Structure of the framework

SASQAF covers the various quality aspects of the entire statistical value chain (i.e. need, design, build, collection, processing, analysis and dissemination), and certifies national statistics on one of four levels. Level 4, certification (quality statistics), indicates optimal conditions for statistical production, while Level 1 (poor statistics) indicates the least favourable conditions. In outline, the four levels of certification are as follows:

Level Four: Quality Statistics – These are statistics that meet all the quality requirements as set out in SASQAF. They are designated as **quality statistics** to the extent that deductions can be made from them, and are 'fit for use' for the purpose for which they were designed. Level 4 applies to highly-developed statistical activities with respect to the corresponding indicator.

Level Three: Acceptable Statistics – These are statistics that meet most, but not all the quality requirements as stipulated in SASQAF. They are designated as **acceptable** to the extent that, despite their limitations, deductions can be made, and are 'fit for use' for the purpose for which they were designed. Level 3 refers to moderately well-developed activities with reference to a particular indicator.

Level Two: Questionable Statistics – These are statistics that meet few of the quality requirements as stipulated in SASQAF. They are designated as **questionable** to the extent that very limited deductions can be made, and they are therefore not 'fit for use' for the purpose for which they were designed. Level 2 refers to activities that are developing but still have many deficiencies.

Level One: Poor Statistics – These are statistics that meet almost none of the quality requirements as stipulated in SASQAF. They are designated as **poor statistics** to the extent that no deductions can be made from them, and are not 'fit for use' for the purpose for which they were designed. Level 1 refers to activities that are underdeveloped.

6. Quality dimensions

The chapters which follow specify the quality dimensions of SASQAF in considerable detail, setting out descriptions and key components, the related indicators for each dimension, standard(s) that need to be adhered to in relation to a quality indicator, as well as providing benchmarks for the standards which determine the four different quality levels as prescribed.

The prerequisites and eight dimensions of quality
- Prerequisites of quality
- Relevance
- Accuracy
- Timeliness
- Accessibility
- Interpretability
- Comparability and Coherence
- Methodological soundness
- Integrity.

Chapter 1: Prerequisites of quality

1.1   Description

The prerequisites of quality refer to the institutional and organisational conditions that have an impact on data quality. It defines the minimum set of necessary conditions that have to be met in order to produce good quality statistics. It therefore serves as the foundation on which all other dimensions of data quality should be premised on.

1.2   Key components

- Legal and institutional environment (including Memoranda of Understanding (MoUs) or Service Level Agreements (SLAs)
- Privacy and confidentiality
- Commensurability of resources
- Quality as the cornerstone of statistical work

1.3   Quality indicators, standards and benchmarks

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
1.1 The responsibility for producing statistics is clearly specified.	1.1.1 A legal arrangement exists that explicitly mandates the production of statistics.	A law or legal arrangement exists that explicitly provides the mandate for the production of statistics.	A law or legal arrangement exists that implies that statistical production is part of its mandate.	No law or legal arrangement exists, but an informal agreement exists for statistical production.	No arrangement exists.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	**Acceptable Statistics Level 3**	**Questionable Statistics Level 2**	**Poor Statistics Level 1**
1.2 Standards and policies are in place to promote consistency of methods and results.	1.2.1 A set of policies must exist which covers all aspects of the statistical value chain.	Policies exist which cover all aspects of the statistical value chain and are adhered to.	Policies exist which cover all aspects of the statistical value chain but are not always adhered to.	Policies exist, but do not cover all aspects of the statistical value chain.	No policies exist.
	1.2.2 A set of standards related to appropriate policies must exist.	Standards exist which cover all aspects of the statistical value chain and are adhered to.	Standards exist which cover all aspects of the statistical value chain but are not adhered to; or existing standards do not cover all aspects of the statistical value chain but are all adhered to.	Standards exist, but do not cover all aspects of the statistical value chain.	No standards exist.

Indicator		Standards		Assessment Levels			
				Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
1.3	Data sharing and coordination among data-producing agencies are clearly specified.	1.3.1	A legal arrangement must exist which allows for the timely and efficient sharing of data between the collecting agency and the secondary user.	A legal arrangement exists and is adhered to.	A legal arrangement exists but is not adhered to.	An informal arrangement exists.	No arrangement exists.
		1.3.2	Regular contact must occur between the data-producing agencies and secondary users/agencies to resolve statistical issues.	Regular meetings are held to address all statistical issues to be resolved. Commitment to resolving them is achieved.	Regular meetings are held to address all statistical issues to be resolved. Commitment to resolving is not achieved.	Ad hoc meetings occur.	No contact.
1.4	Measures are in place to ensure that individual data are kept confidential, and used for statistical purposes only.	1.4.1	There must be a law or policy that ensures information collected are kept confidential and used for statistical or administrative purposes only.	A law/policy guaranteeing the confidentiality of respondents' information, that it is used for statistical purposes only exists and is adhered to at all times.	A law/policy guaranteeing the confidentiality of respondents' information, that it is used for statistical purposes only exists but is not adhered to at all times.	Although no law/policy exists, efforts are made to keep respondent information confidential and ensure they are used for statistical purposes only.	No law or policy exists.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	**Acceptable Statistics Level 3**	**Questionable Statistics Level 2**	**Poor Statistics Level 1**
1.5 Measures to oblige response are ensured through law.	1.5.1 There must be a law or other formal measures that inform respondents of their obligation to provide information; and any sanctions which may apply if they fail to do so.	There is a law that requires respondents to provide information. Respondents are informed of sanctions which may occur if they fail to respond, and sanctions are enforced.	There is a law that requires respondents to provide information. Respondents are informed of any sanctions which may occur if they fail to respond, though sanctions are not enforced.	A law exists, but there are no efforts made to inform respondents of the consequences of non-response.	No law or other formal measures exists.
1.6 Resources are commensurate with the needs of statistical programmes. • Staff • Facilities • Computing resources • Financing.	1.6.1 Collection programmes must be adequately staffed with skilled personnel.	The programme is adequately staffed with skilled personnel.	The staffing is inadequate with respect to the number of personnel, but their skill levels are high.	The staffing with respect to the number of personnel is adequate, but their skill levels are low.	The staffing is inadequate with respect to the number of personnel as well as the skills they possess.
	1.6.2 There must be a statistics unit or component responsible for compiling statistics.	A statistics unit with qualified statisticians exists.	No statistics unit exists, but there are qualified statisticians.	A statistics unit exists but there are no qualified statisticians.	No statistics unit exists. There are no statisticians.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	**Acceptable Statistics Level 3**	**Questionable Statistics Level 2**	**Poor Statistics Level 1**
	1.6.3 Facilities must have the infrastructure to manage the needs of statistical programmes.	Facilities are well equipped and have the necessary infrastructure for the production of statistics.	N/A	N/A	Facilities are ill-equipped and do not have the necessary infrastructure for the production of statistics.
	1.6.4 Computer hardware resources must be adequate in terms of ▪ data storage; ▪ data back up media; ▪ power supply (uninterrupted); ▪ memory; and ▪ other necessary equipment (notebooks, desktop, etc).	The hardware computing resources is sufficient in terms of data storage, data backup media, power supply, memory, and computers.	The hardware computing resources is sufficient in terms of data backup media, power supply, memory, and computers.	The hardware computing resources is sufficient in terms of power supply, memory, and computers.	The hardware resources are entirely insufficient.
	1.6.5 A disaster recovery and business continuity plan must exist.	A disaster recovery and business continuity plan exists.	N/A	N/A	A disaster recovery and business continuity plan does not exist.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	**Acceptable Statistics Level 3**	**Questionable Statistics Level 2**	**Poor Statistics Level 1**
	1.6.6 Computer software resources must be adequate in terms of • capturing systems; • editing systems; • coding systems; • statistical software; • up-to-date licences; • virus protection; and • appropriate access rights.	Computer software resources are adequate in terms of statistical & other specialised software, up-to-date licences, virus protection, and appropriate access rights.	Computer software resources are adequate in terms of appropriate access rights, virus protection, statistical & other specialised software.	Computer software resources are adequate in terms of appropriate access rights, statistical & other specialised software.	Computer software resources are entirely inadequate.
	1.6.7 Budgets must be adequate.	Budgets are sufficient for the needs of statistical programmes.	N/A	N/A	Budgets are entirely inadequate.
1.7 Measures to ensure efficient use of resources in 1.6 are implemented.	1.7.1 Staff of a statistical programme must be employed in positions that are aligned with their skills profile. Metrics: $a = \dfrac{\text{number of aligned staff}}{\text{total number of staff}} \times 100$	90% $\leq$ a. $\leq$ 100%.	75% $\leq$ a. $\leq$ 89%.	40% $\leq$ a. $\leq$ 74%.	0% $\leq$ a. $\leq$ 39%.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	**Acceptable Statistics Level 3**	**Questionable Statistics Level 2**	**Poor Statistics Level 1**
	1.7.2 Asset management policies that prevent the abuse of facilities (e.g. vehicles, telephones etc.) must be developed, adopted and implemented.	Asset management policies have been developed, adopted and monitored for compliance.	Asset management policies have been developed and adopted. No compliance monitoring exists.	Asset management policies have been developed but are not adopted by the organisation.	No asset management policies exist.
	1.7.3 Policies and procedures governing the use of ICT resources must exist, so as to maximise the return on investment.	All documented policies and procedures on the usage of IT resources have been implemented.	There are well documented policies and procedures on the usage of IT resources. Some have been implemented.	There are policies and procedures on the usage of IT resources. None have been implemented.	No policies or procedures on IT resource usage exist.
	1.7.4 Budgets must be reviewed and audited to ensure that financial resources are used in the best possible way.	Budgets are reviewed and audited to ensure that financial resources are used in the best possible way.	Financial audits take place, however budgets are not reviewed as per the policy.	Budgets are reviewed but there is no auditing procedure in place.	Budgets are neither reviewed nor audited.

Indicator		Standards		Assessment Levels			
				Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
1.8	Processes are in place to focus on, monitor and check quality.	1.8.1	The agency must have a quality management system in place.	The agency has a quality management system in place.	N/A.	N/A.	The agency has no quality management system in place.
		1.8.2	The data producing agency must have an independent data quality audit process.	The data producing agency has an independent data quality audit process.	The data producing agency has developed and implemented a data quality audit process, but it is not independent.	The data producing agency has developed, but not implemented, a data quality audit process.	The data producing agency has no audit process developed.
		1.8.3	Staff members in production areas must have a data quality management requirement as part of their performance agreements or job descriptions.	All staff members have a data quality management requirement as part of performance agreements or job descriptions, with clear sanctions for failure to comply.	All staff members have a data quality management requirement as part of performance agreements or job descriptions, without clear sanctions for failure to comply.	Staff members do not have a data quality management requirement as part of performance agreements or job descriptions.	There are no performance agreements or job descriptions.

Chapter 2: Relevance

2.1 Description

Relevance of statistical information reflects the degree to which the data meet the real needs of clients. It is concerned with whether the available information sheds light on the issues of most importance to users.

2.2 Key components

- Why do you need to conduct the survey or collect data?
- Who are the users of the statistics?
- What are their known needs?
- How well does the output meet these needs?
- Are user needs monitored and fed back into the design process?

2.3 Quality indicators, standards and benchmarks

Indicator		Standards		Assessment Levels			
				Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
2.1	Have both the internal and external users of the data been identified?	2.1.1	An up-to-date user database must exist.	An up-to-date user database exists.	A user database exists but is not up-to-date.	Users are known but not recorded in a database.	Users have not been identified.
2.2	Is there a process to identify user needs?	2.2.1	A process to identify user needs must exist.	A process to identify user needs exists.	N/A	N/A	A process to identify user needs does not exist.

Indicator		Standards		Assessment Levels			
				Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
2.3	Are user needs and the usage of statistical information analysed?	2.3.1	A report containing the findings of user needs and the usage of statistical information must be available.	A report containing the findings of user needs and the usage of statistical information is available.	User needs and the usage of statistical information are analysed, but a report is not available.	One of user needs or usage of statistical information is analysed, but a report is not available.	User needs and usage of statistical information are not analysed.
2.4	Changes are made as a result of user needs assessments.	2.4.1	The results of the user needs assessment must influence decisions on the statistical value chain of the survey or on administrative data collection systems, where feasible. Documented reasons for not implementing user needs must be provided as feedback to users.	The results of the user needs assessment influence decisions on the statistical value chain of the survey or administrative data collection systems, where feasible. Documented reasons for not implementing user needs are provided as feedback to users.	The results of the user needs assessment influence decisions on the statistical value chain of the survey or administrative data collection systems, where feasible. Documented reasons for not implementing user needs are not provided as feedback to users.	The results of the user needs assessment do not influence in any way decisions on the statistical value chain of the survey or on administrative data collection systems. Documented reasons for not implementing user needs are provided as feedback to users.	The results of the user needs assessment do not influence in any way decisions on the statistical value chain of the survey or on administrative data collection systems. Documented reasons for not implementing user needs are not provided as feedback to users.

Indicator		Standards		Assessment Levels			
				Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
2.5	Is there a process to determine the satisfaction of users with the statistical information?	2.5.1	A formal process must exist to determine the satisfaction of users with the statistical information.	Formal processes exist to determine the satisfaction of users with the statistical information. The results are incorporated into the statistical production process.	Formal processes exist to determine the satisfaction of users with the statistical information. The results are not incorporated into the statistical production process.	Processes exist to determine the satisfaction of users with the statistical information, but they are not formal.	There is no process at all.

Chapter 3: Accuracy

3.1 Description

The accuracy of statistical information is the degree to which the output correctly describes the phenomena it was designed to measure. Source data available provide an adequate basis to compile statistics.

3.2 Key components

- Assessment of sampling errors where sampling was used.
- Assessment of coverage of data collection in comparison to the target population.
- Assessment of response rates and estimates of the impact of imputation.
- Assessment of non-sampling errors and any other serious accuracy or consistency problems with the survey results or register based statistics.
- Data capture, data coding and data processing errors.
- Source data available provide an adequate basis to compile statistics (e.g. administrative records).
- Source data reasonably approximate the definitions, scope, classifications, valuation, and time of recording required.
- Source data are timely.

3.3 Quality indicators, standards and benchmarks

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
3.1 Measures of sampling errors for key variables are calculated. Amongst others these are: • standard error • coefficient of variation (CV) • confidence interval (CI) • mean square error (MSE) • design effect (DEFF).	3.1.1 Measures of sampling errors must be calculated for the main variables. They must be available for the other variables on request.	Measures of sampling errors are published for the main variables. Measures for other variables are available on request.	Measures of sampling errors are published for the main variables. Measures for other variables are not available on request	Measures of sampling errors are available on request for the main variables only.	Measures of sampling errors are not calculated.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
	3.1.2 Measures of sampling errors must fall within acceptable standards. At a minimum the following must be calculated: standard error, coefficient of variation, confidence interval, mean square error. The low accuracy of variables (if these exist), are explained. Metrics: $$SE = \sqrt{Var(\hat{\theta})}$$ $$CV = \frac{\sqrt{Var(\hat{\theta})}}{E(\hat{\theta})}$$ $$MSE(\hat{\theta}) = Var(\hat{\theta}) + B^2(\hat{\theta}),$$ Where $\hat{\theta}$ is an estimator of a parameter of interest	CV< or = 5% and $B^2(\hat{\theta}) = 0$	CV< or = 5% and $B^2(\hat{\theta}) > 0$	5%<CV< or = 30%	CV >30%
	3.1.3 Scientific Sampling techniques must be used. Metrics: a = Design effect $(\theta) = 1 + \delta(n-1)$	Scientific sampling techniques are used. 0<a≤1.	Scientific Sampling techniques are used. 1.0< a. <1.2.	Scientific Sampling techniques are used. 1.2 ≤ a. ≤1.5.	Scientific Sampling techniques are not used or a.>1.5.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
3.2 Measures of non-sampling errors are calculated, viz.: • Frame coverage errors (e.g. duplication in the frame/register used to conduct a survey, number of statistical units out of scope (i.e. number of ineligible units)	3.2.1 The extent of measures of non-sampling errors must be kept to an acceptable level. Metrics: $$a= \frac{\sum |\text{final weights - design weights}|}{\sum \text{design weights}}$$ $$b = \max \left[\frac{|\text{final weights - design weights}|}{\text{design weights}} \right]$$	a. < 15% and b. < 15%.	$15\% \leq$ a. <25%; or $15\% \leq$ b. < 25%.	a. $\geq$ 25% or b. $\geq$ 25%.	a. and b. are not calculated.
• Misclassification errors • Systematic errors to determine the extent of bias introduced for both administrative records and surveys • Measurement errors (e.g. questionnaire effects, data collection effects, interviewer effects and respondent effects) • Processing errors (e.g. data entry errors rates, coding errors, editing failure rates, imputation rates)	3.2.2 Delays between newly-registered administrative units and the corresponding statistical unit births must be known. Update procedures are enforced to correct for under-coverage.	Delays between newly-registered administrative units and the corresponding statistical unit births are known. Corrective actions for under-coverage are undertaken.	Delays between newly-registered administrative units and the corresponding statistical unit births are known. Corrective actions for under-coverage are not undertaken.	N/A	Delays between newly- registered administrative units and the corresponding statistical unit births are unknown.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
• Model assumption errors • Non-response errors (e.g. item non-response rates, unit non-response rates and overall response rates).	3.2.3 The measures of under-coverage fall within acceptable standards. where a= accepted standard x,y,z = agreed thresholds such that $z>y$	$x \geq a$	$a-y \leq x < a$	$a-z \leq x < a-y$	$x<a-z$ or measures of under-coverage is not calculated.
	3.2.4 Delays between de-registering of administrative units and the corresponding statistical unit deaths must be known. Update procedures are enforced to correct for over-coverage.	Delays between de-registering of administrative units and the corresponding statistical unit deaths are known. Corrective actions for over-coverage are undertaken.	Delays between de-registering of administrative units and the corresponding statistical unit deaths are known. Corrective actions for over-coverage are not undertaken.	N/A	Delays between de-registering of administrative units and the corresponding statistical unit deaths are unknown.
	3.2.5 The duplication rate must be at an acceptable level. where a= accepted standard x,y,z = agreed thresholds such that $z>y$	$x \leq a$	$a+y \geq x > a$	$a+z \geq x > a+y$	$x>a+z$ or duplication rate is not calculated.

Indicator	Standards	Assessment Levels			
		Quality Statistics **Level 4**	**Acceptable Statistics** **Level 3**	**Questionable Statistics** **Level 2**	**Poor Statistics** **Level 1**
	3.2.6 The proportion of units which are out of scope must be at an acceptable level. where a= accepted standard x,y,z = agreed thresholds such that z>y	$x \leq a$	$a+y \geq x > a$	$a+z \geq x > a+y$	x>a+z or the proportion of units which are out of scope is not calculated.
	3.2.7 The proportion of units which are misclassified must be at an acceptable level. where a= accepted standard x,y,z = agreed thresholds such that z>y	$x \leq a$	$a+y \geq x > a$	$a+z \geq x > a+y$	x>a+z or the proportion of units which are misclassified is not calculated.
	3.2.8 Systematic errors must be identified and reported.	Systematic errors are both identified and reported.	Systematic errors are identified, but not reported.	N/A.	Systematic errors are not identified.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
	3.2.9 Every 10th statistical unit is independently double collected. The two outputs must be compared and corrective action must be taken. Records must be kept.	Every 10th statistical unit is independently double collected. The two outputs are compared and corrective action is taken. Records are available.	Fewer than one-in-ten statistical units are independently double collected. The two outputs are compared and corrective action is taken. Records are available.	Fewer than one-in-ten statistical units are independently double collected. The two outputs are compared, but no corrective action is taken. Records are available.	Double collection is not done.
	3.2.10 Data collection error rates calculated from fieldwork records must be at an acceptable level. where a= accepted standard x,y,z = agreed thresholds such that z>y	$x \le a$	$a+y \ge x > a$	$a+z \ge x > a+y$	x>a+z or data collection error rates is not calculated.
	3.2.11 The effects of data collection instruments must be determined and reported.	The effects of data collection instruments are reported.	N/A	N/A	The effects of data collection instruments are not determined.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	**Acceptable Statistics Level 3**	**Questionable Statistics Level 2**	**Poor Statistics Level 1**
	3.2.12 The effects of the data collection mode must be determined and reported.	The effects of data collection mode are reported.	N/A	N/A	The effects of data collection mode are not determined.
	3.2.13 The effects of the interviewers must be determined and reported.	The effects the interviewers are reported.	N/A	N/A	The effects of the interviewers are not determined.
	3.2.14 Respondent effects must be determined and reported.	Respondent effects are reported.	N/A	N/A	Respondent effects are not determined.
	3.2.15 Proxy responses must be separately categorised (flagged). Proxy response rate must be at an acceptable level. where a= accepted standard x,y,z = agreed thresholds such that z>y	$x \leq a$	$a+y \geq x > a$	$a+z \geq x > a+y$	x>a+z or Proxy response rate is not calculated.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
	3.2.16 Data entry error must average an acceptable accuracy rate where a= accepted standard x,y,z = agreed thresholds such that z>y	$x \leq a$	$a+y \geq x> a$	$a+z \geq x > a+y$	x>a+z or data entry error is not calculated.
	3.2.17 Coding error must average an acceptable accuracy rate. where a= accepted standard x,y,z = agreed thresholds such that z>y	$x \leq a$	$a+y \geq x> a$	$a+z \geq x > a+y$	x>a+z or coding error is not calculated.
	3.2.18 Editing rate must average an acceptable level. where a= accepted standard x,y,z = agreed thresholds such that z>y	$x \leq a$	$a+y \geq x> a$	$a+z \geq x > a+y$	x>a+z or editing rate is not calculated.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	**Acceptable Statistics Level 3**	**Questionable Statistics Level 2**	**Poor Statistics Level 1**
	3.2.19 Editing failure rate must average an acceptable level. where a= accepted standard x,y,z = agreed thresholds such that z>y	$x \leq a$	$a+y \geq x > a$	$a+z \geq x > a+y$	x>a+z or editing failure rate is not calculated.
	3.2.20 The imputation rate for item non-response must average an acceptable level. where a= accepted standard x,y,z = agreed thresholds such that z>y	$x \leq a$	$a+y \geq x > a$	$a+z \geq x > a+y$	x>a+z or imputation rate for item non-response is not calculated.
	3.2.21 The imputation rate for unit non-response must average an acceptable level. where a= accepted standard x,y,z = agreed thresholds such that z>y	$x \leq a$	$a+y \geq x > a$	$a+z \geq x > a+y$	x>a+z or imputation rate for unit non-response is not calculated.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	**Acceptable Statistics Level 3**	**Questionable Statistics Level 2**	**Poor Statistics Level 1**
	3.2.22 The model assumptions must be stated. All models used in the estimation of statistics must be described.	The model assumptions are stated. All models used in the estimation of statistics are described. The model assumptions are valid, and evidence is provided for this.	The model assumptions are stated. Not all models used in the estimation of statistics are described. The model assumptions are valid, and evidence is provided for this.	The model assumptions are not stated. All models used in the estimation of statistics are described. The model assumptions are not valid, or no evidence is provided for validity.	The model assumptions are not stated. All models used in the estimation of statistics are not described. The model assumptions are not valid, or no evidence is provided for this.
	3.2.23 Item non-response rate must be within acceptable levels where a= accepted standard x,y,z = agreed thresholds such that z>y	$x \leq a$	$a+y \geq x > a$	$a+z \geq x > a+y$	x>a+z or item non-response rate is not calculated.
	3.2.24 Unit non-response rate must be within acceptable levels. where a= accepted standard x,y,z = agreed thresholds such that z>y	$x \leq a$	$a+y \geq x > a$	$a+z \geq x > a+y$	x>a+z or unit non-response rate is not calculated.

Indicator	Standards		Assessment Levels			
			Quality Statistics Level 4	**Acceptable Statistics Level 3**	**Questionable Statistics Level 2**	**Poor Statistics Level 1**
3.3 To what extent is the primary data appropriate for the statistical product produced?	3.3.1	Source data must be consistent with the scope, definitions, and classifications of the statistical product produced.	Source data are consistent with the scope, definitions, and classifications of statistical product produced.	Source data is consistent with of scope and definitions.	Source data is consistent with the scope of the product.	Source data is consistent with only definitions and/or classifications of the statistical product.
3.4 Data from the primary source have been quality assessed.	3.4.1	Source data must be accompanied by a quality report.	Source data is accompanied by a quality report.	N/A	N/A	Source data does not have a quality report.
3.5 Register/frame maintenance procedures are adequate. • updates. • quality assurance. • data audit.	3.5.1	Maintenance procedures of register/frame must be documented and must be performed on a regular basis in line with what has been documented.	Maintenance procedures of register/frame are documented and performed on a regular basis in line with what has been documented.	Maintenance procedures of register/frame are documented and are performed on an ad hoc basis.	Maintenance procedures of register/frame are not documented but are performed on an ad hoc or regular basis.	Maintenance procedures of register/frame are not documented and are not performed on a regular basis.
	3.5.2	The impact of frame maintenance must be measured, monitored, analysed and reported on.	Maintenance impact is measured, monitored analysed and reported on.	Maintenance impact is measured and monitored, reported on but not analysed.	Maintenance impact is measured but not monitored or analysed nor reported on.	Maintenance impact is not measured, monitored or analysed.

Indicator	Standards		Assessment Levels			
			Quality Statistics Level 4	**Acceptable Statistics Level 3**	**Questionable Statistics Level 2**	**Poor Statistics Level 1**
	3.5.3	A survey for the assessment of quality must be conducted for every release, based on a sample drawn from the administrative records. A register improvement survey must be undertaken periodically to improve the quality of an identified deficiency. Feedback from both surveys must be regularly incorporated into update procedures.	A survey for the assessment of quality is conducted for every release, based on a sample drawn from the administrative records. A register improvement survey is undertaken periodically to improve the quality of an identified deficiency. Feedback from both surveys is regularly incorporated into update procedures.	A survey for the assessment of quality is conducted for every release, based on a sample drawn from the administrative records. A register improvement survey is undertaken periodically to improve the quality of an identified deficiency. Feedback from both surveys is not incorporated into update procedures.	A survey for the assessment of quality is conducted for every release, based on a sample drawn from the administrative records. A register improvement survey is not undertaken periodically to improve the quality of an identified deficiency. Feedback is not incorporated into update procedures.	A survey for the assessment of quality is not conducted for every release. A register improvement survey is not undertaken periodically to improve the quality of an identified deficiency.
3.6 Are data collection systems sufficiently open and flexible to cater for new developments?	3.6.1	The system must be flexible in all its components. It must be designed to allow for new developments (e.g., changes in definitions, classifications, etc).	The system is flexible in all its components. It is designed to allow for new developments. There is no need for an ad hoc collection system.	N/A	N/A	The system is inflexible.

Indicator	Standards		Assessment Levels			
			Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
3.7 Description of record-matching methods and techniques used on the administrative data sources. • match rate as a percentage of total records • measure of false negative matches (missed matches) • measure of false positive matches (mismatches)	3.7.1	Procedures and/or algorithms must be fully described. The description of the procedures and/or algorithms must be independently verified by replicating the output.	Procedures and/or algorithms are fully described. The description of the procedures and/or algorithms is independently verified by replicating the output.	Procedures and/or algorithms are not fully described. But, the algorithms are sufficiently described to allow for testing. The algorithms are independently verified by replicating the output.	Procedures and/or algorithms are fully described. The description of the procedures and/or algorithms is not verified by replicating the output.	Procedures and/or algorithms are not fully described.
	3.7.2	False negative matches (missed matches) created as a result of missed matches must be at an acceptable level. Metrics: $a = \dfrac{\text{number of missed matches}}{\text{total number of true and false matching pairs}} \times 100$	a. < 5%	5% $\leq$ a. <15%	a. $\geq$ 15%	No missed match rates calculated.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
	3.7.3 False positive matches (mismatches) created as a result of missed matches must be at an acceptable level. Metrics: $a = \dfrac{\text{number of missed matches}}{\text{total number of true and false matching pairs}} \times 100$	a. < 5%	5% $\leq$ a. <15%	a. $\geq$ 15%	No mismatch rates calculated.

Chapter 4: Timeliness

4.1   Description

Timeliness of statistical information refers to the delay between the reference point to which the information pertains and the date on which the information becomes available. Timeliness also addresses aspects of periodicity and punctuality of production activities within the statistical value chain.

4.2   Key components

- Statistics production time
- Timely receipt of administrative records.
- Periodicity of statistical release.
- Punctuality of statistical release.

4.3   Quality indicators, standards and benchmarks

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
4.1 Average time between the end of reference period and the date of the preliminary results.	4.1.1 The preliminary results must be released according to the prescribed standard. Metrics: $a = \dfrac{\text{date of preliminary results - end of reference period}}{\text{prescribed timeframe}}$ Where a. > 1 means the prescribed timeframe has been exceeded.	$0.9 < a. \leq 1$. Preliminary results are released in accordance with prescribed timeframes.	$1 < a. < 1.25$ times the prescribed lapse.	$a < 0.9$ or $a \geq 1.25$ times the prescribed lapse.	No preliminary results are released.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
4.2 Average time between the end of reference period and the date of the final results.	4.2.1 The final results must be released according to the prescribed standard. Metrics: $a = \dfrac{\text{date of final results - end of reference period}}{\text{prescribed timeframe}}$ a. > 1 means the prescribed timeframe has been exceeded.	0.9 < a. ≤ 1. Final results are released in accordance with prescribed timeframes.	1 < a. < 1.25 times the prescribed lapse.	a < 0.9 or 1.25 < a. < 1.5 times the prescribed lapse.	a. ≥ 1.5 times the prescribed timeframe.
4.3 Production activities within the statistical value chain are within the planned timelines, viz.: • data collection • data processing • data analysis • dissemination.	4.3.1 Project plan/schedule of key deadlines related to the statistical value chain must be compiled.	A schedule of key deadlines does exist.	N/A	N/A	A schedule of key deadlines does not exist.
	4.3.2 Updates to registers must occur within clearly specified timeframes.	Updates to registers occur within clearly specified timeframes	N/A	N/A	There is no time frame to update the register.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
	4.3.3 Data collection must follow the project plan/schedule Metrics $a = \frac{actual\ duration - planned\ duration}{planned\ duration} + 1$ a. > 1 implies that collections the phase exceeded planned duration. a. < 1 implies that collections phase completed within the planned duration.	a. ≤ 1 times the planned duration for data collection.	1< a. <1.25 times the planned duration for data collection.	1.25 ≤ a. <1.5 times the planned duration for data collection.	a. ≥1.5 times the planned duration for data collection.
	4.3.4 A protocol for the timely delivery of administrative data must exist and must be adhered to.	A protocol for the timely delivery of administrative data exist and it is adhered to.	A protocol for the timely delivery of administrative data exist and it is not adhered to.	N/A	A protocol for the timely delivery of administrative data does not exist.
	4.3.5 Data processing must follow the project plan/schedule. Metrics: $a = \frac{actual\ duration - planned\ duration}{planned\ duration} + 1$ a. > 1 implies that the processing phase exceeded the planned duration a. < 1 implies that the processing phase was completed within the planned duration.	a. ≤ 1 times the planned duration for data processing.	1< a. <1.25 times the planned duration for data processing.	1.25 ≤ a. <1.5 times the planned duration for data processing.	a. ≥1.5 times the planned duration for data processing.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
	4.3.6 Data analysis must follow the project plan/schedule. Metrics: $a = \frac{actual\ duration - planned\ duration}{planned\ duration} + 1$ a. > 1 implies that the analysis phase has exceeded the planned timeframes. a. < 1 implies that the analysis phase was completed within the planned timeframes.	a. ≤ 1 times the planned duration for data analysis.	1< a. <1.25 times the planned duration for data analysis.	1.25 ≤ a. <1.5 times the planned duration for data analysis.	a. ≥1.5 times the planned duration for data analysis.
	4.3.7 Dissemination must follow the project plan/schedule. Metrics: $a = \frac{actual\ duration - planned\ duration}{planned\ duration} + 1$ a. > 1 implies that the dissemination phase exceeded the planned timeframes. a. < 1 implies that the dissemination phase was completed within the planned timeframes.	a. ≤ 1 times the planned duration for dissemination.	1< a. <1.25 times the planned duration for dissemination.	1.25 ≤ a. <1.5 times the planned duration for dissemination.	a. ≥1.5 times the planned duration for dissemination.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
4.4 Periodicity of release.	4.4.1 The periodicity (e.g. monthly, quarterly, and annual) of release must conform to a data dissemination standard.	The periodicity of release conform to a data dissemination standard.	N/A	N/A	The periodicity of release does not conform to a data dissemination standard.

Chapter 5: Accessibility

5.1 Description

The accessibility of statistical information and metadata refers to the ease with which it can be obtained from the agency. This includes the ease with which the existence of information can be ascertained, as well as the suitability of the form or medium through which the information can be accessed. The cost of the information may also be an aspect of accessibility for some users.

5.2 Key components

- Catalogue systems are available in the organ of state or statistical agency
- Delivery systems to access information
- Information and metadata coverage is adequate
- Measure of catalogue and delivery systems performance
- Means of sharing data between stakeholders

5.3 Quality indicators, standards and benchmarks

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
5.1 Are statistical products (e.g. data, metadata) available to the public?	5.1.1 The statistical products must be disseminated to the public.	The statistical products are disseminated to the public.	N/A.	The statistical products are disseminated to selected users.	The statistical products are not disseminated.

Indicator		Standards		Assessment Levels			
				Quality Statistics Level 4	**Acceptable Statistics Level 3**	**Questionable Statistics Level 2**	**Poor Statistics Level 1**
5.2	Rules governing the restricted availability of administrative records are well described and documented.	5.2.1	A policy document having clear rules governing the restricted availability of administrative records must exist.	A policy document having clear rules governing the restricted availability of administrative records exists and is adhered to.	A policy document having clear rules governing the restricted availability of administrative records exists and is not adhered to.	A policy document having vague rules governing the restricted availability of administrative records exists.	There is no policy document at all.
5.3	Types of media and/or channels used for sharing data amongst stakeholders are adequate and preserve confidentiality.	5.3.1	Data must be accessible through a variety of channels with mechanisms that ensure confidentiality.	Data are accessible through a variety of channels with mechanisms that ensure confidentiality.	Data are accessible through a variety of channels though loopholes exist that may compromise confidentiality.	Limited channels exist for stakeholders to access data and no mechanisms exist to ensure confidentiality.	At most one channel exists for stakeholders to access data.
5.4	Data is accessible in a format beyond the producing agency.	5.4.1	The data must be available in a variety of file formats.	It is available in a variety of file formats.	N/A	N/A	It is available in one file format only.
5.5	Statistics are released on a pre-announced schedule.	5.5.1	Statistics must be released on a pre-announced schedule.	Statistics are released on a pre-announced schedule which is made available to users annually and is adhered to.	A pre-announced schedule exists and is made available to users annually but not adhered to.	A pre-announced schedule exists which is made available to users in an ad hoc manner but is not adhered to.	No pre-announced schedule exists.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
5.6 Statistical products are made available to all users at the same time.	5.6.1 Statistical release must be made available to all users at the same time.	Statistical release has an embargo date and time and is made available to all users at the same time.	Statistical release does not have an embargo date and time, but are made available to all users at the same time.	Statistical release has an embargo date and time, but is not made available to all users at the same time.	Statistical release does not have an embargo date and time, and are not made available to all users at the same time.
5.7 Statistics/administrative records not routinely disseminated are made available upon request.	5.7.1 Statistics/administrative records not routinely disseminated must be made available, and the terms and conditions on which they are made available must be publicised.	Statistics/administrative records not routinely disseminated are made available, and the terms and conditions are publicised.	Statistics/administrative records not routinely disseminated are made available, but the terms and conditions are not publicised.	N/A	Statistics/administrative records not routinely disseminated are not available to users.
	5.7.2 Special requests must be considered and be met.	80-100% of special requests are met.	60 – 79% of special requests are met.	40-59% of special requests are met.	0-39% of special requests are met.
5.8 User support services exist and are widely publicised.	5.8.1 User support services must exist and be widely publicised.	User support services exist and are widely publicised.	User support services exist, but are not widely publicised.	N/A	User support services do not exist.

Indicator		Standards	Assessment Levels			
			Quality Statistics Level 4	**Acceptable Statistics Level 3**	**Questionable Statistics Level 2**	**Poor Statistics Level 1**
5.9	Does a data dissemination policy exist, and is it accessible?	5.9.1 A data dissemination policy must be accessible.	A data dissemination policy exists, freely accessible and is adhered to.	A data dissemination policy exists, freely accessible but is not adhered to.	A data dissemination policy exists but it is not freely accessible.	No dissemination policy is in place.
5.10	Does the pricing policy governing dissemination exist, and is it accessible?	5.10.1 A pricing policy must be accessible.	A pricing policy exists, freely accessible and is adhered to.	A pricing policy exists, freely accessible but is not adhered to.	A pricing policy exists but it is not freely accessible.	No pricing policy is in place.
5.11	Catalogues of publications and other services are available to users of statistics.	5.11.1 Catalogues of publications and other services must be accessible to users of statistics.	Catalogues of publications and other services exist and freely accessible to users of statistics.	N/A.	Catalogues of publications and other services exist, but is not freely accessible to users of statistics.	No catalogue is in place.
5.12	Metadata are readily accessible to users.	5.12.1 Minimum metadata required for interpreting the product must be accessible.	Minimum metadata required for interpreting the product is readily accessible to users.	Minimum metadata required for interpreting the product is available, but not readily accessible to users.	Incomplete metadata is available.	No metadata is available at all.

Chapter 6: Interpretability

6.1 Description

Interpretability of statistical information is refers to the ease with which users understand statistical information through the provision of metadata.

6.2 Key components

- Concepts and definitions, and classifications that underlie the data;
- Metadata on the methodology used to collect and compile the data;
- Key findings, giving the summary of the results;
- Presentation of statistics in a meaningful way.

6.3 Quality indicators, standards and benchmarks

Indicator		Standards		Assessment Levels			
				Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
6.1	Documented metadata (definitional, operational, methodological, system and dataset) are sufficient to understand data.	6.1.1	Metadata must be documented according to the accepted standards, guidelines or good practices.	A complete set of metadata is documented according to the standard(s).	Sufficient metadata is not documented according to the standard(s).	Incomplete set of metadata is documented according to the standard(s).	No metadata available.
6.2	Statistics are presented in a clear and understandable manner.	6.2.1	The presentation of the statistics must be according a standard.	The presentation of the statistics is according the standard.	N/A	Standard exist, but is not adhered to.	There is no standard.
6.3	Statistical releases contain a summary of the key findings.	6.3.1	Statistical releases must contain a summary of key findings.	Statistical releases contain a summary of key findings.	N/A	N/A	Statistical releases are published without summary of key findings

Chapter 7: Comparability and Coherence

7.1   Description

Comparability of statistical information is the ability to compare statistics on the same characteristic between different points in time, geographical areas or statistical domains. The coherence of statistical information reflects the degree to which it can be successfully brought together with other similar statistical information from different sources within a broad analytic framework and over time. It is the extent to which differences between two sets of statistics are attributable to differences between the estimates and the true value of the statistics.

7.2   Key components

- The use of common concepts and definitions within and between series.
- The use of common variables and classifications within and between statistical series.
- The use of common methodology and systems for data collection and processing within series.
- The use of common methodology for various processing steps of a survey such as editing and imputations within series.

7.3   Quality indicators, standards and benchmarks

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	**Acceptable Statistics Level 3**	**Questionable Statistics Level 2**	**Poor Statistics Level 1**
7.1 Data within series and administrative systems are based on common concepts and definitions, classifications, and methodology, and departures from this are identified in the metadata.	7.1.1 All data (including source data, related frame data, and related survey data) within the same series must use the same concepts and definitions. Departures from common concepts and definitions must be identified in the metadata and archived.	All statistical products within the same series use the same concepts and definitions.	At least one statistical product within the same series uses different concepts and definitions. Departures from common concepts and definitions are identified in the metadata and archived.	At least one statistical product within the same series uses different concepts and definitions. Departures from common concepts and definitions are identified in the metadata, but not archived.	At least one statistical product within the same series uses different concepts and definitions. Departures from common concepts and definitions are not identified in the metadata and are not archived.
	7.1.2 All data (including source data, related frame data, and related survey data) within the same series must use the same classifications. Departures from common classifications must be identified in the metadata and archived.	All statistical products within the same series use the same classifications.	At least one statistical product within the same series uses different classifications. Departures from common classifications are identified in the metadata and archived.	At least one statistical product within the same series uses different classifications. Departures from common classifications are identified in the metadata, but not archived.	At least one statistical product within the same series uses different classifications. Departures from common classifications are not identified in the metadata and are not archived.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
	7.1.3 All data (including source data, related frame data, and related survey data) within the same series must use the same methodology. Departures from common methodology must be identified in the metadata and archived.	All statistical products within the same series use the same methodology.	At least one statistical product within the same series uses different methodology. Departures from common methodology are identified in the metadata and archived.	At least one statistical product within the same series uses different methodology. Departures from common methodology are identified in the metadata, but not archived.	At least one statistical product within the same series uses different methodology. Departures from common methodology are not identified in the metadata and are not archived.
7.2 Statistics are consistent or reconcilable over time.	7.2.1 Statistics must be consistent over time.	Statistics are consistent over time.	N/A	N/A	Statistics are not consistent over time.
	7.2.2 The statistics must follow an expected trend established over time. Any inconsistencies in the key variables must be reconciled.	The statistics follow an expected trend established over time. Inconsistencies in the key variables are reconciled.	N/A	N/A	The statistics do not follow an expected trend established over time.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
7.3 Data across comparable series, or source data are based on common frames, identifiers, concepts and definitions, and classifications, and departures from these are identified in the metadata.	7.3.1 Data across comparable series or source data must be based common identifiers. Departures from common identifiers must be identified in the metadata and archived.	Data across comparable series or source data are based on common identifiers.	Data across comparable series or source data are not based on common identifiers. Usage of different identifiers is identified in the metadata, and is archived.	Data across comparable series or source data are not based on common identifiers. Usage of different identifiers is identified in the metadata, but is not archived.	Data across comparable series or source data are not based on common identifiers. Usage of different identifiers is not identified in the metadata and is not archived.
	7.3.2 Data across comparable series or source data must be based on common identifiers. Departures from common identifiers must be identified in the metadata and archived.	Data across comparable series or source data are based on common identifiers.	Data across comparable series or source data are not based on common identifiers. Usage of different identifiers is identified in the metadata, and is archived.	Data across comparable series or source data are not based on common identifiers. Usage of different identifiers is identified in the metadata, but is not archived.	Data across comparable series or source data are not based on common identifiers. Usage of different identifiers is not identified in the metadata and is not archived.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	**Acceptable Statistics Level 3**	**Questionable Statistics Level 2**	**Poor Statistics Level 1**
	7.3.3 Data across comparable series or source data must use the same concepts and definitions. Departures from common concepts and definitions must be identified in the metadata and archived.	Data across comparable series or source data are based on common concepts and definitions.	Data across comparable series or source data are not based on common concepts and definitions. Usage of different concepts and definitions is identified in the metadata, and is archived.	Data across comparable series or source data are not based on common on concepts and definitions. Usage of different concepts and definitions is identified in the metadata, but is not archived.	Data across comparable series or source data are not based on the same concepts and definitions. Usage of different concepts and definitions is not identified in the metadata and is not archived.
	7.3.4 Data across comparable series or source data must use the same classifications. Departures from common classifications must be identified in the metadata and archived.	Data across comparable series or source data are based on common classifications.	Data across comparable series or source data are not based on common classifications. Usage of different classifications is identified in the metadata, and is archived.	Data across comparable series or source data are not based on common classifications. Usage of different classifications is identified in the metadata, but is not archived.	Data across comparable series or source data are not based on common classifications. Usage of different classifications is not identified in the metadata and is not archived.

Indicator		Standards		Assessment Levels			
				Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
7.4	Statistics are checked for consistency with those obtained through other data sources.	7.4.1	Statistics must be checked for consistency with a comparable dataset. Inconsistencies must be reconciled.	The data producer demonstrates that the statistics are consistent. Any inconsistencies in the key variables are reconciled.	The data producer does not demonstrate that the statistics are consistent. Any inconsistencies in the key variables are reconciled.	The data producer demonstrates that the statistics are consistent. Any inconsistencies in the key variables are not reconciled.	The data producer does not demonstrate that the statistics are consistent. Any inconsistencies in the key variables are not reconciled.
7.5	A common set of identifiers (for the purpose of record matching) exist and have been agreed upon by data producers.	7.5.1	A common identifier must be agreed upon by the data producers.	The identifier is the same across all datasets and is agreed upon by the data producers.	The identifier is the same across all datasets but is not agreed upon by the data producers.	There is a common identifier across any two datasets. However, there is not one identifier common to all datasets.	There is no common identifier across all datasets to be matched.
		7.5.2	The common identifier must be unique in every dataset. Rules and practices must be agreed upon to ensure uniqueness.	The common identifier is unique in every dataset. Rules and practices to ensure uniqueness are agreed upon.	The common identifier is unique in every dataset. Rules and practices to ensure uniqueness exist independently of each other.	The common identifier is unique in every dataset. Rules and practices to ensure uniqueness do not exist.	The common identifier is not unique in every dataset.

Chapter 8: Methodological soundness

8.1   Description

It refers to the application of international, national, or peer-agreed standards, guidelines, and practices to produce statistical outputs. Application of such standards fosters national and international comparability.

8.2   Key components

- International norms and standards on methods.
- Data compilation methods employ acceptable procedures.
- Other statistical procedures employ sound statistical techniques.
- Transparent revision policy and studies of revisions are done and made public.

8.2   Quality indicators, standards and benchmarks

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
8.1 Concepts, definitions, and classifications used follow accepted standards, guidelines or good practices (national, international, peer-agreed).	8.1.1 The concepts and definitions must satisfy accepted standards, guidelines or good practice in line with national, international, peer-agreed norms; and must be documented. Deviations from the standard must be formally approved, and be fully documented.	The concepts and definitions satisfy accepted standards, guidelines or good practice in line with national, international, peer-agreed norms; and are documented. Deviations from the standard are formally approved, and fully documented.	The concepts and definitions satisfy accepted standards, guidelines or good practice in line with national, international, peer-agreed norms; and are documented. Deviations from the standard are not approved, and fully documented.	The concepts and definitions are documented, but do not satisfy accepted standards, guidelines or good practice.	No documented concepts and definitions exist.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
	8.1.2 The classifications must satisfy accepted standards, guidelines or good practice in line with national, international, peer-agreed norms; and must be documented. Deviations from the standard must be formally approved, and be fully documented.	Classifications satisfy accepted standards, guidelines or good practice in line with national, international, peer-agreed norms; and are documented. Deviations from the standard are formally approved, and fully documented.	Classifications satisfy accepted standards, guidelines or good practice in line with national, international, peer-agreed norms; and are documented. Deviations from the standard are not approved, and fully documented.	Classifications are documented, but do not satisfy accepted standards, guidelines or good practice.	No documented classifications exist.
8.2 The scope of the study is consistent with accepted standards, guidelines or good practices.	8.2.1 The scope of the study must be appropriate for the intended topic. The scope of the study must be consistent with accepted standards, guidelines or good practices in line with the survey constraints.	The scope of the study is appropriate for the intended topic and is consistent with accepted standards, guidelines or good practices in line with the survey constraints.	N/A	N/A	The scope of the study is inappropriate for the intended topic.

Indicator	Standards		Assessment Levels			
			Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
8.3 Methodologies used follow accepted standards, guidelines or good practices (national, international, peer-agreed), viz.: • questionnaire design • sampling methods • sampling frame design • frame maintenance • piloting • data collection methods • data editing and imputation methods • data analytical methods • revision procedures	8.3.1	The designing of the questionnaire must follow accepted standard, sets of guidelines or good practices.	The designing of the questionnaire follows accepted standard, sets of guidelines or good practices.	N/A	N/A	The designing of the questionnaire does not follow accepted standard, sets of guidelines or good practices.
	8.3.2	The sampling methods must follow accepted standard, sets of guidelines or good practices.	The sampling methods follow accepted standard, sets of guidelines or good practices.	N/A	N/A	The sampling methods do not follow accepted standard, sets of guidelines or good practices.
	8.3.3	The frame maintenance methods must follow accepted standard, sets of guidelines or good practices.	The frame maintenance methods follow accepted standard, sets of guidelines or good practices.	N/A	N/A	The frame maintenance methods do not follow accepted standard, sets of guidelines or good practices.
	8.3.4	The piloting methods must follow accepted standard, sets of guidelines or good practices.	The piloting methods follow accepted standard, sets of guidelines or good practices.	N/A	N/A	The piloting methods do not follow accepted standard, sets of guidelines or good practices.

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
	8.3.5 Data collection methods must follow accepted standard, sets of guidelines or good practices.	Data collection methods follow accepted standard, sets of guidelines or good practices.	N/A	N/A	Data collection methods do not follow accepted standard, sets of guidelines or good practices.
	8.3.6 Editing and imputation methods must follow accepted standard, sets of guidelines or good practices.	Editing and imputation methods follow accepted standard, sets of guidelines or good practices.	N/A	N/A	Editing and imputation methods do not follow accepted standard, sets of guidelines or good practices.
	8.3.7 The methods of analysis used must follow accepted standards, sets of guidelines or good practices.	The methods of analysis used follow accepted standards, sets of guidelines or good practices.	N/A	N/A	The methods of analysis used do not follow accepted standards, sets of guidelines or good practices.
	8.3.8 Revision methods used must follow accepted standards, sets of guidelines or good practices.	Revision methods used follow accepted standards, sets of guidelines or good practices.	N/A	N/A	Revision methods used do not follow accepted standards, sets of guidelines or good practices.

Indicator		Standards		Assessment Levels			
				Quality Statistics Level 4	**Acceptable Statistics Level 3**	**Questionable Statistics Level 2**	**Poor Statistics Level 1**
8.4	Are revisions schedule followed? Are they regular and transparent?	8.4.1	A revisions schedule must exist for surveys which conduct revisions. The revisions schedule must be publicly available and accessible. The revisions must take place as per the schedule.	A revision schedule is publicly available, accessible and adhered to.	The revisions schedule is available and adhered to, but not publicly accessible.	A revisions schedule is publicly available and accessible but is not adhered to.	There is no revisions schedule.
8.5	Preliminary and revised data are identified in the metadata.	8.5.1	Preliminary and revised data must be identified in the metadata. Metadata must contain an explanation of the changes.	Preliminary and revised data are identified in the metadata. Metadata contains an explanation of the changes.	Preliminary and revised data are identified in the metadata. Metadata does not contain an explanation of the change.	N/A	No revisions are identified in the metadata.
8.6	Studies of revisions and their findings are made public.	8.6.1	Regular studies of revisions or upcoming revisions must be done and their findings must be made public.	Regular studies of revisions or upcoming revisions are done and their findings are made public.	Regular studies of revisions or upcoming revisions are done and their findings are not made public.	Studies of revisions are done on an ad hoc basis.	No revision studies are done.

Chapter 9: Integrity

9.1 Description

The integrity of statistical information refers to values and related practices that maintain users' confidence in the agency producing statistics and ultimately in the statistical product. This includes, among others, the need for the statistical system to be based on the United Nations (UN) principles of official statistics and includes principles of objectivity in collection, compilation and dissemination of data to ensure unbiased statistics which are not subject to confidentiality breaches or premature releases.

9.2 Key components

- Professionalism and ethical standards which guide policies and practices.
- Assurances that statistics are produced on an impartial basis.
- Ethical standards are guided by policies and procedures.

9.3 Quality indicators, standards and benchmarks

Indicator	Standards	Assessment Levels			
		Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
9.1 The terms and conditions, including confidentiality, under which statistics are collected, processed and disseminated are available to the public and follow the UN principles of official statistics.	9.1.1 A terms and conditions document must be available and accessible to the public.	A terms and conditions document is accessible to the public. The terms and conditions follow the UN principles of official statistics.	A terms and conditions document is not accessible to the public. The terms and conditions follow the UN principles of official statistics.	A terms and conditions document is accessible to the public. The terms and conditions do not follow the UN principles of official statistics.	A terms and conditions document does not exist.

Indicator		Standards		Assessment Levels			
				Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
9.2	Describe the conditions under which policy-makers, specifically government, may have access to data before release. Are the conditions published?	9.2.1	A data dissemination policy detailing the conditions under which policy-makers have access to the data must be available.	Policy-makers have controlled access to the data. The conditions, along with reasons for their access, are for published and adhered to.	Policy-makers have controlled access to the data. The conditions, along with reasons for their access, are published and not adhered to.	Policy-makers have uncontrolled access to the data. The conditions, along with reasons for their access, are not published.	There are no conditions preventing policy makers getting access to the data prior to the release.
9.3	Advance notice is given of major changes in methodology and source data.	9.3.1	Advance notice of at least 6 months must be given of major changes in methodology and source data.	Advance notice of at least 6 months is given of major changes in methodology and source data.	Advance notice of between 3 months to 6 months is given of major changes in methodology and source data.	Advance notice of less than 3 months is given of major changes in methodology and source data.	No advance notice is given of major changes in methodology and source data.
9.4	Government commentary, when data are released, should be identified as such, and not be seen as part of the official statistics.	9.4.1	Government commentary, when data are released, must be identified as such, and not be seen as part of the official statistics.	Government commentary, when data are released, is clearly identified as such, and not seen as part of the official statistics.	N/A	N/A	Government commentary is regarded as part of the official statistics.

Indicator		Standards		Assessment Levels			
				Quality Statistics Level 4	Acceptable Statistics Level 3	Questionable Statistics Level 2	Poor Statistics Level 1
9.5	Choice of source data, techniques and dissemination decisions are informed solely by statistical considerations.	9.5.1	The choice of source data, techniques and dissemination decisions must be informed solely by statistical considerations.	The choice of source data, techniques and dissemination decisions are solely informed by statistical considerations.	N/A	N/A	The choice of source data, techniques and dissemination decisions are not solely informed by statistical considerations. Results are either suppressed or/and manipulated.
9.6	Ethical guidelines for staff behaviour are in place and are well known to the staff.	9.6.1	A professional code of conduct must be in place providing ethical guidelines for staff behaviour.	A professional code of conduct is in place providing ethical guidelines for staff behaviour. The code of conduct is well known and enforced.	A professional code of conduct is in place providing ethical guidelines for staff behaviour. The code of conduct is well known, but is not enforced.	A professional code of conduct is in place providing ethical guidelines for staff behaviour. The code of conduct is not well known and not enforced.	A professional code of conduct is not in place.

7. References

Brackstone, G. (1999). Managing data quality in a statistical agency. Survey Methodology, 25, 139-149

Eurostat, (2004) FAO Statistical Data Quality Framework: A multi-layered approach to monitoring and assessment

IMF (DQAF), Source: http://dsbb.imf.org/vgn/images/pdfs/dqrs_Genframework.pdf 13 February 2008

UNECE (2009), Generic Statistical Business Process Model, Joint UNECE/Eurostat/OECD Work Session on Statistical Metadata (METIS) Version 3.1 – December 2008

Statistics South Africa, (2008), South African Statistical Quality Assessment Framework (SASQAF), edition 1, Statistics South Africa, 2008

OECD (2003/1). Quality Framework and Guidelines for OECD statistical activities, Version 2003/1, September 2003

Statistics Canada (2000). Policy on Informing Users of Data Quality and Methodology. Statistics Canada

Statistics Canada (2003). Statistics Canada's Quality Guidelines, Fourth Edition Catalogue no. 12-539-XIE, October 2003

Statistics Netherlands, (2004), Conceptual metadata and process metadata: Key elements to improve the quality of the statistical system

Statistics Act, Act No 6 of 1999, Republic of South Africa, 1999

Sundgren, B.(2000) The Swedish Statistical metadata system. Eurostat conference, 2000

Svante Öberg, Director General (2005) Quality Issues in the European Statistical System, Statistics Sweden

World Bank, Development Data Group and UNESCO: Institute for Statistics: A Framework for Assessing the Quality of Education Statistics

Annexure A: Protocol specifying the procedure for the Statistician-General to designate statistics as official statistics

Introduction

Purpose of this protocol

The purpose of this protocol is to specify the requirements that must be met, and the procedures to be followed, before statistics can be certified as official statistics.

Mandate

Section 14(7)(a) of the Statistics Act (No. 6 of 1999) empowers to the Statistician-General to 'designate as official statistics any statistics or class of statistics produced from statistical collections by Statistics South Africa; or other organs of state, after consultation with the head of the organ of state concerned'. In certain circumstances, statistics which are not produced by an organ of state can be evaluated and certified in the same way as statistics produced by an organ of state.

Definitions

Organ of state:
'(a)   any department of state or administration in the national, provincial or local sphere of government; or

(b)   any other functionary or institution:
  (i)   exercising a power or performing a duty in terms of the Constitution or a provincial constitution; or
  (ii)   exercising a public power or performing a public duty in terms of any legislation, but does not include a court or a judicial officer'.

Official statistics:

Statistics produced by an organ of state which the Statistician-General has designated as official in terms of Section 14(7)(a) of the Statistics Act (No. 6 of 1999).

National statistics:

Statistics produced by an organ of state that are within the public domain, and have not been designated as official statistics.

Member of National Statistics Systems (NSS):

"An organ of state or other organisation that produces, supplies or uses statistics, and has signed a memorandum of understanding with the Statistician-General committing to adhere to common statistical quality criteria, standards, and procedures as set down by the Statistician-General in terms of Sections 7(2) (e) and 14 (6) of the Statistics Act."

Scope of the protocol

The protocol covers the role of the Statistician-General, the function of the National Statistics System (NSS), the general principles that guide evaluation of statistics, and the procedure for evaluating, and designating statistics as official by the Statistician-General.

Statistical principles

The Statistics Act empowers the Statistician-General to coordinate statistics and develop standards for both Statistics South Africa and other organs of state that produce statistics. Section 3(2) of the Act specifies that official statistics must be

'(a)   relevant, accurate, reliable and timeous;

(b)   objective and comprehensive;

(c)   compiled, reported and documented in a scientific and transparent manner;

(d)   disseminated impartially;

(e)   accessible;

(f)   in accordance with appropriate national and international standards and classifications; and

(g)   sensitive to distribution by gender, disability, region and similar socio-economic features.'

For statistics to be certified as official, they need to be aligned with the above principles. The principles that underpin the certification process, and provide the framework for the conduct of the parties involved in the certification process, are presented below[2].

Principle 1: The aim of the National Statistics System (NSS) is to provide a framework for effective and comprehensive coordination of statistical output, quality, and standards.

Principle 2: It is a long-term goal of the NSS that only official statistics will be used to inform government policies, programmes and projects, including the Government-wide Monitoring and Evaluation (GWM&E) system.

Principle 3: It is a long-term goal of the NSS that all national statistical collections that meet the relevant criteria should be accorded official status.

Principle 4: National statistics qualify as official statistics when they are relevant beyond the organ or agency that collected them; when their production is sustainable; when they meet quality criteria and standards specified by the Statistician-General; and when they are accessible as a public good.

Principle 5: Assessment of statistical quality is guided by the South African Statistical Quality Assessment Framework (SASQAF).

Principle 6: Statistics declared as official will be reviewed at regular intervals, to be determined jointly by the Statistician-General and the head of the relevant organ or agency, in order to ensure that they remain relevant and of specified quality.

Principle 7: Official and national statistical series and other statistical products in the public domain may be evaluated for effectiveness, efficiency and comparative benchmarking at periodic intervals, determined jointly by the Statistician-General and the head of the producing organ or agency. Evaluations may incorporate peer reviews.

Principle 8: Collection, processing and analysis of data should be governed exclusively by scientific principles in accordance with international or peer-agreed best practice, within the parameters of available resources.

Principle 9: Statistical processes, procedures and methodology should be fully documented to enable users to assess fitness for purpose.

Principle 10: Custody of data designated as official statistics will normally be with Statistics South Africa, unless the Statistician-General and the head of the producing organ of state or agency agree otherwise. Ownership of data remains with the originating organ of state or agency.

---

[2]   Some of the principles also appear in the South African Statistical Quality Assessment Framework (SASQAF). However, SASQAF provides an operational framework and more detailed criteria for assessment, while the principles are wider in scope and provide an 'environmental' framework.

Procedure for designating statistics from organs of state as official statistics

1. The Statistician-General will publish and regularly update SASQAF as a framework within producing agencies who may apply to have data designated as official statistics. The Statistician-General, in consultation with the head of the producing organ of state or agency, determines the elements or outputs of the producing organ of state or agency to be designated as official statistics. These could include a survey, a register, a dataset, indicators, a data table, etc.

2. An organ of state or agency will apply, through the division responsible for the NSS at Statistics South Africa, to the Statistician-General to have their statistics designated as official statistics.

3. Applications will be referred to a Data Quality Assessment Team (DQAT) constituted by the Statistician-General, drawn from
   - Statistics South Africa
   - Applicant (organ of state)
   - Subject-matter expert(s) (recommended by the organ of state and/or the Statistician General)
   - Statistics Council member (observer status).

4. Appointed DQAT members will sign Terms of References for the review and an Oath of Confidentiality document.

5. For assessment to begin, the submitting organ of state and the statistics under review need to comply with three initial criteria:
   - The producing agency must be a member of the NSS.
   - The statistics are used to meet user needs beyond those specific and internal to the producing agency.
   - The statistics produced should be part of a sustainable series, not a once-off collection.

6. DQAT will assess the quality of the product(s) in terms of SASQAF requirements, assigning a SASQAF quality level to the product. The assessment process is as follows:

   a. the applicant will identify all the SASQAF indicators that are relevant to the product under evaluation, and motivate why the remaining indicators are not relevant;

   Note: The selection is based on the requirements of the product or the properties of the data. In principle, indicators that provide useful information to users should be selected. Not all indicators are relevant for all products.

   b. once DQAT and the applicant reach agreement on which indicators are relevant, and on the standard for each indicator, they will sign an agreement to this effect;

   c. the applicant will then be asked to produce a quality declaration for their product, for all the agreed indicators; and

   d. DQAT will assess these quality statements against the relevant standards, and based on the results, assign one of the four quality levels (quality, acceptable, questionable or poor), and will identify areas of improvement in the quality statements.

7. DQAT will recommend the overall SASQAF level of the product.

8. If the product submitted for evaluation is not classified as *quality statistics* in terms of the SASQAF levels of evaluation, DQAT will advise the applicant on areas of improvement.

9. If the product satisfies the requirements o*f quality statistic*s set out in SASQAF, the Statistician General will designate the product as *official statistics.*

10. Once the product has been designated as official statistics, it will be published with the Statistician-General's official seal of approval (the Official Statistics Mark), and stored in the NSS archive for public access.

11. The Statistician General will issue a notice in the government gazette to the effect that a product has been designated as official statistics

12. The product then becomes subject to periodic reviews, determined by the Statistician General in consultation with the head of the producing agency or department.

13. The Statistician General will publish the results of the assessment or review for access by the public.

Annexure B: Fundamental principles of official statistics relative to Statistics South Africa

In its endeavour to fulfill the purpose of providing users with quality information, Statistics South Africa (Stats SA) has adopted the following principles developed by the Economics and Social Council Statistical Commission of the United Nations:

Principle 1: Relevance, impartiality and equal access
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 to honour citizens' entitlement to public information.

Principle 2: Professional standards and ethics
To retain trust in official statistics, Stats SA will 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: Accountability and transparency
To facilitate the correct interpretation of data, Stats SA will present information according to scientific standards on the sources, methods and procedures of statistics.

Principle 4: Prevention of misuse
Stats SA is entitled to comment on erroneous interpretation and misuse of statistics.

Principle 5: Cost-effectiveness
Data for statistical purposes may be drawn from all types of sources, be they statistical surveys or administrative records. Stats SA will choose the source with regard to quality, timeliness, costs and the burden on respondents.

Principle 6: Confidentiality
Individual data collected by Stats SA for statistical compilation, whether they refer to natural or legal persons, will be strictly confidential and used exclusively for statistical purposes.

Principle 7: Legislation
The laws, regulations and measures under which the statistical systems operate will be made public.

Principle 8: National coordination
Stats SA will promote coordination among statistical producers within South Africa in order to advance consistency and efficiency in the statistical system.
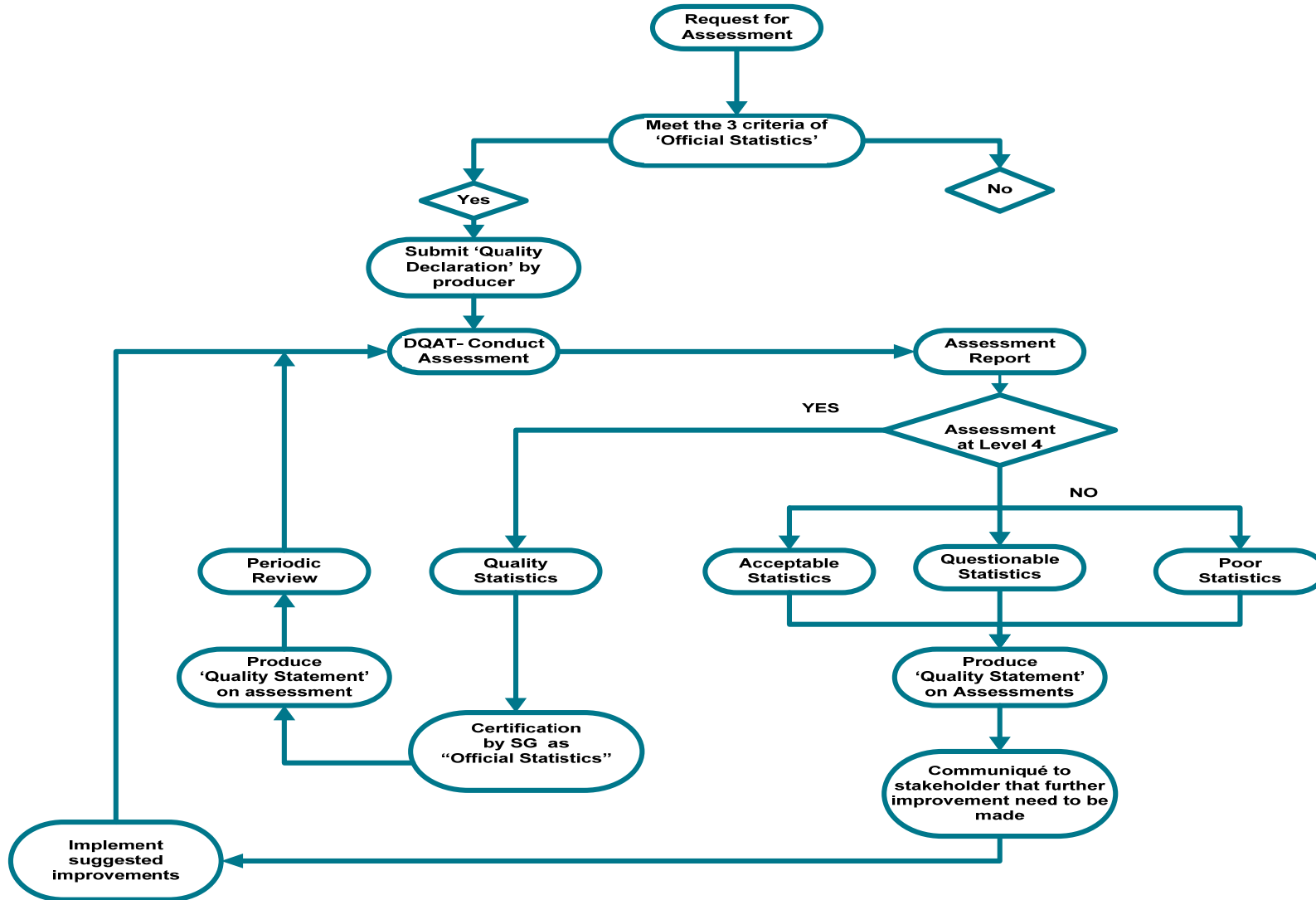
Principle 9: International standards
Stats SA will use international concepts, classifications and methods, where possible, to promote the consistency and efficiency of statistical systems between countries.

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

Annexure C: The process of assessing statistics and designating them as official

```
                                    Request for
                                    Assessment
                                        │
                                        ▼
                                 Meet the 3 criteria of
                     ┌───────────  'Official Statistics'  ───────────┐
                     │                                               │
                     ▼                                               ▼
                   Yes                                              No
                     │
                     ▼
             Submit 'Quality
             Declaration' by
                producer
                     │
      ┌──────────────┤
      │              ▼
      │          DQAT– Conduct ───────────────────────────────► Assessment
      │          Assessment                                      Report
      │              ▲                                              │
      │              │                                              ▼
      │              │                                         Assessment
      │              │                       YES                 at Level 4
      │              │          ┌─────────────────────────────────┤
      │              │          │                                  │ NO
      │              │          │                    ┌─────────────┼─────────────┐
      │              │          ▼                    ▼             ▼             ▼
      │          Periodic    Quality            Acceptable   Questionable      Poor
      │          Review      Statistics         Statistics   Statistics      Statistics
      │              ▲          │                    │             │             │
      │              │          │                    └─────────────┼─────────────┘
      │          Produce        │                                  ▼
      │          'Quality       │                             Produce
      │          Statement'     │                             'Quality Statement'
      │          on assessment  │                             on Assessments
      │              ▲          ▼                                  │
      │              │     Certification                           ▼
      │              │     by SG as                           Communiqué to
      │              └──── "Official Statistics"               stakeholder that further
      │                                                        improvement need to be
      │                                                        made
      │                                                            │
    Implement ◄─────────────────────────────────────────────────┘
    suggested
    improvements
```
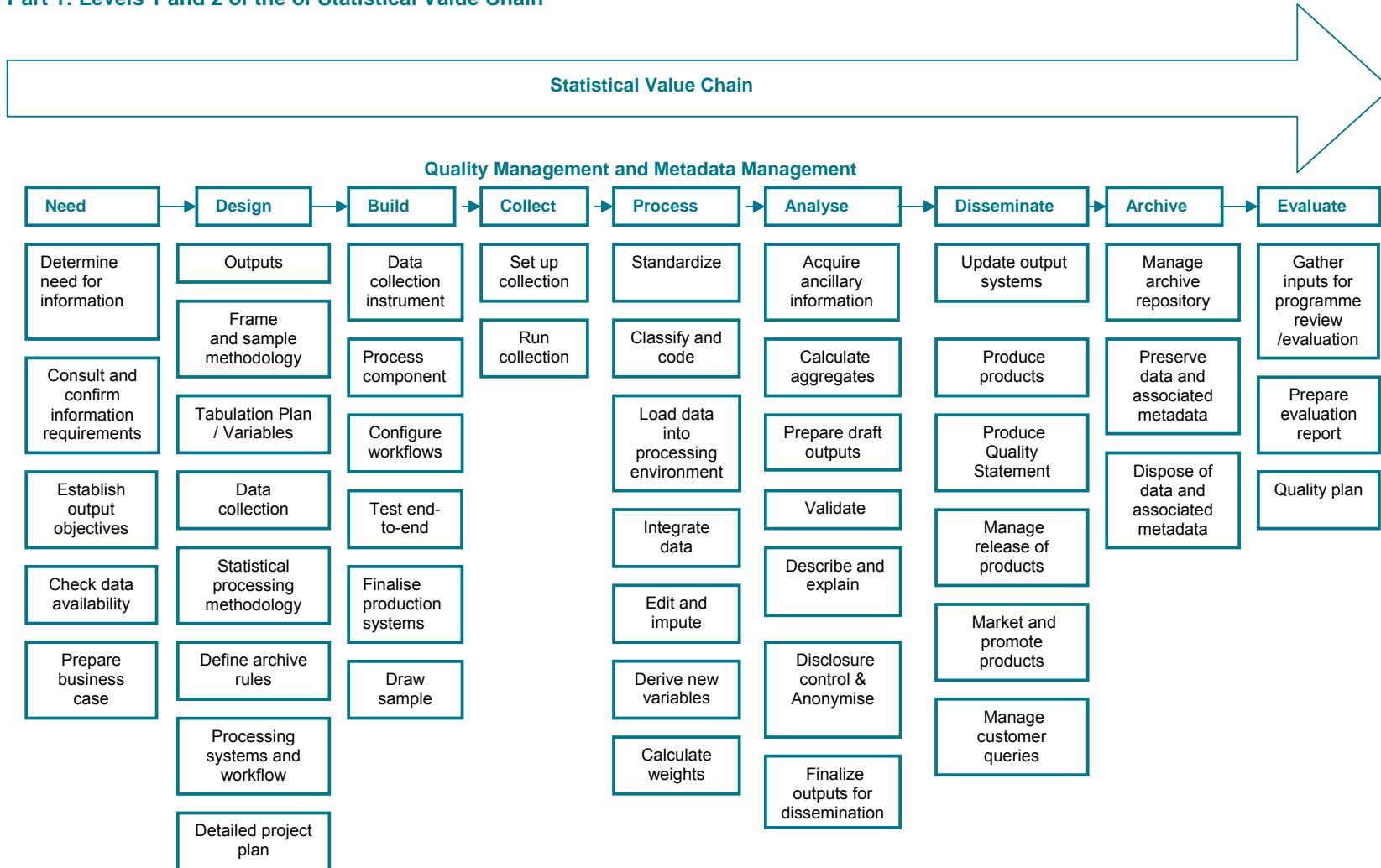
Annexure D: Statistical Value Chain (SVC)

Statistics South Africa has adapted the Statistical Value Chain as developed by the Joint UNECE/Eurostat/OECD Work Session on Statistical Metadata (METIS) in 2008. The statistical process involves a range of statistical operations, which are enabled by various support functions.

The Statistical Value Chain can be divided into four levels:

- Level 0: the statistical business process;
- Level 1: the nine phases of the statistical business process;
- Level 2: the sub-processes within each phase;
- Level 3: a description of those sub-processes.

A diagram showing the phases (Level 1) and sub-processes (Level 2) is included as Part 1. The sub-processes are described in detail in Part 2. The SVC also contains two over-arching processes that apply throughout the nine phases, and across statistical business processes, which are Quality Management and Metadata Management.

Part 1: Levels 1 and 2 of the of Statistical Value Chain

Statistical Value Chain

Quality Management and Metadata Management

Need	Design	Build	Collect	Process	Analyse	Disseminate	Archive	Evaluate
Determine need for information	Outputs	Data collection instrument	Set up collection	Standardize	Acquire ancillary information	Update output systems	Manage archive repository	Gather inputs for programme review /evaluation
Consult and confirm information requirements	Frame and sample methodology	Process component	Run collection	Classify and code	Calculate aggregates	Produce products	Preserve data and associated metadata	Prepare evaluation report
Establish output objectives	Tabulation Plan / Variables	Configure workflows		Load data into processing environment	Prepare draft outputs	Produce Quality Statement	Dispose of data and associated metadata	Quality plan
Check data availability	Data collection	Test end-to-end		Integrate data	Validate	Manage release of products		
Prepare business case	Statistical processing methodology	Finalise production systems		Edit and impute	Describe and explain	Market and promote products		
	Define archive rules	Draw sample		Derive new variables	Disclosure control & Anonymise	Manage customer queries		
	Processing systems and workflow			Calculate weights	Finalize outputs for dissemination			
	Detailed project plan							

Part 2: Levels 2 and 3 of the Statistical Value Chain

This part of annexure considers each phase in turn, identifying the various sub-processes within that phase, and describing their contents. It therefore covers Levels 2 and 3 of the SVC.

1. Need

This is the first phase of the SVC, which involves all the necessary planning for a new survey. This phase is triggered when a need for new statistics is identified, or feedback about current statistics initiates a review. It determines whether there is a presently unmet demand, externally and/or internally, for the identified statistics, and whether the statistical organisation can produce them. The following are activities that need to be carried out during the Need phase:

1.1 Determine need for information
- Initial investigation and identification of what statistics are needed and what is needed of the statistics.
- Consideration of practice amongst other statistical organisations, & methods used by those organisations.

1.2 Consult and confirm need
- Consulting with the stakeholders and confirming in detail the need for the statistics. Statistical organisations should know what it is expected to deliver, when, how, and, perhaps most importantly, why.
- Determining whether previously identified needs have changed. This detailed understanding of user needs is the critical part of this sub-process.

1.3 Establish output objectives
- Identifies the statistical outputs that are required to meet the user needs identified in sub-process 1.2 (Consult and confirm need).
- Agreeing the suitability of the proposed outputs and their quality measures with users.

1.4 Check data availability
- Checks whether current data sources could meet user requirements, and the conditions under which they would be available, including any restrictions on their use.
- Research into potential administrative data sources and their methodologies, to determine whether they would be suitable for use for statistical purposes.
- Prepare a strategy for filling any remaining gaps in the data requirement

1.5 Prepare business case
- Documents the findings of the other sub-processes in this phase in the form of a business case to get approval to implement the new or modified statistical business process.
- Such a business case would typically include:
  - A description of the "As-Is" business process (if it already exists), with information on how the current statistics are produced highlighting any inefficiencies and issues to be addressed
  - The proposed "To-Be" solution, detailing how the statistical business process will be developed to produce the new or revised statistics;
  - An assessment of costs and benefits, as well as any external constraints.

2. Design

The phase describes the development and design activities, and any associated practical research work needed to define the statistical outputs, concepts, methodologies, collection instruments and operational processes. For statistical outputs produced on a regular basis, this phase usually occurs for the first iteration, and whenever improvement actions are identified in Phase 9 (Evaluate) of a previous iteration.

2.1 Outputs
- Contains the detailed design of the statistical outputs to be produced, including the related development work and preparation of the systems and tools used in the Disseminate phase.
- Outputs should be designed, wherever possible, to follow existing standards, so inputs to this process may include metadata from similar or previous collections, international standards, and information about practices in other statistical organisations from sub-process Determine need for information.

2.2 Frame and sample methodology
- Identifies and specifies the population of interest, defines a sampling frame (and, where necessary, the register from which it is derived), and determines the most appropriate sampling criteria and methodology (which could include complete enumeration). Common sources are administrative and statistical registers, censuses and sample surveys.
- Describes how these sources can be combined if needed.
- Analysis of whether the frame covers the target population should be performed.
- A sampling plan should be made: The actual sample is created in sub-process 3.6 (Draw sample), using the methodology specified in this sub-process.

2.3 Tabulation plan/Variables
- Defines the variables to be collected via the data collection instrument, as well as any other variables that will be derived from them in sub-process 5.6 (Derive new variables), and any classifications that will be used. It is expected that existing national and international standards will be followed wherever possible. This may need to run in parallel with sub-process 2.4 (Data collection), as the definition of the variables to be collected, and the choice of data collection instrument may be inter-dependent to some degree.
- Preparation of metadata descriptions of collected and derived variables and classifications is a necessary precondition for subsequent phases.

2.4 Data collection
- Determines the most appropriate data collection method(s) and instrument(s). The actual activities in this sub-process vary according to the type of collection instruments required, which can include computer assisted interviewing, paper questionnaires, administrative data interfaces and data integration techniques.
- Design of questions and response templates (in conjunction with the variables and classifications designed in sub-process 2.3 (Tabulation plan/Variables)).
- Design of any formal agreements relating to data supply, such as memoranda of understanding, and confirmation of the legal basis for the data collection.
- This sub-process is enabled by tools such as question libraries (to facilitate the reuse of questions and related attributes), questionnaire tools (to enable the quick and easy compilation of questions into formats suitable for cognitive testing) and agreement templates (to help standardise terms and conditions).

- Design of process-specific provider management systems.

2.5   Statistical processing methodology

- Designs the statistical processing methodology to be applied during Phase 5 (Process), and Phase 6 (Analyse).
- This can include developing and testing routines for coding, editing, imputing, estimating integrating, verifying and finalising data sets.

2.6   Define archive rules

- This sub-process is where the archiving rules for the statistical data and metadata resulting from a statistical business process are determined. The requirement to archive intermediate outputs such as the sample file, the raw data from the collect phase, and the results of the various stages of the process and the analyse phases should also be considered. The archive rules for a specific statistical business process may be fully or partly dependent on the more general archiving policy of the statistical organisation, or, for national organisations, on standards applied across the government sector. The rules should include consideration of the medium and location of
the archive, as well as the requirement for keeping duplicate copies. They should also consider the conditions (if any) under which data and metadata should be disposed of. (Note – this sub-process is logically strongly linked to Phase 2 – Design, at least for the first iteration of a statistical business process).

2.7   Processing systems and workflow

- Determines the workflow from data collection to archiving, taking an overview of all the processes required within the whole statistical production process, and ensuring that they fit together efficiently with no gaps or redundancies. Various systems and databases are needed throughout the process.

- A general principle is to reuse processes and technology across many statistical business processes, so existing systems and databases should be examined first, to determine whether they are fit for purpose for this specific process, then, if any gaps are identified, new solutions should be designed.

2.8   Detailed project plan

- Develop a project plan giving details on activities to be carried out, start date and duration of each activity; and human resources allocated to each activity.

3.   Build

This phase builds and tests the production systems to the point where they are ready for use in the "live" environment. For statistical outputs produced on a regular basis, this phase usually occurs for the first iteration, and following a review or a change in methodology, rather than for every iteration. It is broken down into five sub-processes, which are generally sequential, from left to right, but can also occur in parallel, and can be iterative.

3.1   Data collection instrument

- Describes the activities to build the collection instruments to be used during the Phase 4 (Collect). The collection instrument is generated or built based on the design specifications created during Phase 2 (Design). A collection may use one or more collection modes to receive the data, e.g. personal or telephone interviews; paper, electronic or web questionnaires. Collection instruments may also be data extraction routines used to gather data from existing statistical or administrative data sets.

- Preparing and testing the contents and functioning of that instrument (e.g. testing the questions in a questionnaire). It is recommended to consider the direct connection of collection instruments to the statistical metadata system, so that metadata can be more easily captured in the collection phase. Connection of metadata and data at the point of capture can save work in later phases.

3.2 Process components

- Describes the activities to build new and enhance existing software components needed for the business process, as designed in Phase 2 (Design). Components may include dashboard functions and features, data repositories, transformation tools, workflow framework components, provider and metadata management tools.

3.3 Configure workflows

- Configures the workflow, systems and transformations used within the statistical business processes, from data collection, right through to archiving the final statistical outputs. It ensures that the workflow specified in sub-process 2.7 (Processing system and workflow) works in practice.

3.4 Test end-to-end

- Describes the activities to manage a field test or pilot of the statistical business process. Typically it includes a small-scale data collection, to test collection instruments, followed by processing and analysis of the collected data, to ensure the statistical business process performs as expected.

Following the pilot, it may be necessary to go back to a previous step and make adjustments to instruments, systems or components. For a major statistical business process, e.g. a population census, there may be several iterations until the process is working satisfactorily.

3.5 Finalise production systems

- Include activities to put the process, including workflow systems, modified and newly-built components into production ready for use by business areas. The activities include:
  - producing documentation about the process components, including technical documentation and user manuals
  - training the business users on how to operate the process
  - moving the process components into the production environment, and ensuring they work as expected in that environment.

3.6 Draw sample

- Establishes the frame and selects the sample for this iteration of the collection, as specified in sub-process 2.2 (Frame and sample methodology).
- includes the coordination of samples between instances of the same statistical business process (for example to manage overlap or rotation), and between different processes using a common frame or register (for example to manage overlap or to spread response burden). Quality assurance, approval and maintenance of the frame and the selected sample are also undertaken in this sub-process, though maintenance of underlying registers, from which frames for several statistical business processes are drawn, is treated as a separate business process.
- The sampling aspect of this sub-process is not usually relevant for processes based entirely on the use of pre-existing data sources (e.g. administrative data) as such processes generally create frames from the available data and then follow a census approach.

4. Collect

This phase collects all necessary data, using different collection modes (including extractions from administrative and statistical registers and databases), and loads them into the appropriate data environment. For statistical outputs produced regularly, this phase occurs in each iteration.

4.1 Set up collection

- Ensures that the people, processes and technology are ready to collect data, in all modes as designed. It takes place over a period of time, as it includes the strategy, planning and training activities in preparation for the specific instance of the statistical business process. Where the process is repeated regularly, some (or all) of these activities may not be explicitly required for each iteration. For one-off and new processes, these activities can be lengthy.
- This sub-process includes:
  - preparing a collection strategy
  - training collection staff
  - ensuring collection resources are available e.g. laptops
  - configuring collection systems to request and receive the data;
  - ensuring the security of data to be collected;
  - preparing collection instruments (e.g. printing questionnaires, pre-filling them with existing data, loading questionnaires and data onto interviewers' computers etc.).

4.2 Run collection

- This is where the collection is implemented, with the different collection instruments being used to collect the data. It includes the initial contact with providers and any subsequent follow-up or reminder actions. It records when and how providers were contacted, and whether they have responded.

- This includes the management of the providers involved in the current collection, ensuring that the relationship between the statistical organization and data providers remains positive, and recording and responding to comments, queries and complaints.
- For administrative data, this process is brief: the provider is either contacted to send the data, or sends it as scheduled. When the collection meets its targets (usually based on response rates) the collection is closed and a report on collection is produced.

4.3 Load data into processing environment

- Initial data validation, as well as loading the collected data and metadata into a suitable electronic environment for further processing in Phase 5 (Process). It may include automatic data take-on, for example using optical character recognition tools to extract data from paper questionnaires, or converting the formats of data files received from other organisations. In cases where there is a physical data collection instrument, such as a paper questionnaire, which is not needed for further processing, this sub-process manages the archiving of that material in conformance with the principles established in phase 8 (Archive).

5. Process

This phase describes the cleaning of data records and their preparation for analysis. It is made up of sub-processes that check, clean, and transform the collected data, and may be repeated several times. For statistical outputs produced regularly, this phase occurs in each iteration. The sub-processes in this phase can apply to data from both statistical and non-statistical sources (with the possible exception of sub-process 5.7 (Calculate weights), which is usually specific to survey data).

5.1 Standardise and anonymise

- This is where statistical units are derived or standardised, and where data are anonymised.
- Depending on the type of source data, this sub-process may not always be needed.
- Standardisation includes converting administrative or collection units into the statistical units required for further processing.
- Anonymisation strips data of identifiers such as name and address, to help to protect confidentiality. Standardisation and anonymisation may take place before or after sub-process 5.2 (Integrate data), depending on the requirements for units and identifiers in that sub-process.

5.2 Integrate data

- Integrates one or more data sources. The input data can be from a mixture of external or internal data sources, and a variety of collection modes. The result is a harmonised data set. Data integration typically includes:
    - matching / record linkage routines, with the aim of linking data from different sources referring to the same unit;
    - prioritising when two or more sources contain data for the same variable (with potentially different values).
- Data integration may take place at any point in this phase, before or after any of the other subprocesses. There may also be several instances of data integration in any statistical business process.

5.3 Classify and code

- Classifies and codes the input data. For example automatic (or clerical) coding routines may assign numeric codes to text responses according to a pre-determined classification scheme.

5.4 Edit and impute

- This applies to collected micro-data, and looks at each record to try to identify (and where necessary correct) missing data, errors and discrepancies. It can also be referred to as input data validation. It may be run iteratively, validating data against predefined edit rules, usually in a set order. It may apply automatic edits, or raise alerts for manual inspection and correction of the data. Where data are missing or unreliable, estimates are imputed, often using a rule-based approach. Specific steps include:
    - the identification of potential errors and gaps;
    - the selection of data to include or exclude from editing and imputation routines;
    - editing and imputation using one or more pre-defined methods e.g. "hot-deck" or "cold-deck"
    - imputation;
    - writing the edited / imputed data back to the data set, and flagging them as edited or imputed;
    - the production of metadata on the editing and imputation process;
- Editing and imputation can apply to unit records both from surveys and administrative sources, before and after integration.

5.5 Derive new variables

- This sub-process creates variables that are not explicitly provided in the collection and are needed to deliver the required outputs. It derives these new variables by applying arithmetic formulae to one or more of the variables that are already present in the dataset. It may need to be iterative, as some derived variables may themselves be based on other derived variables. It is therefore important to ensure that variables are derived in the correct order.

5.6   **Calculate weights**

- This sub process creates weights for unit data records according to the methodology created in sub-process 2.5: Statistical processing methodology. These weights can be used to "gross-up" sample survey results to make them representative of the target population, or to adjust for non-response in total enumerations.

6.   Analyse

In this phase, statistics are produced, examined in detail, interpreted, and made ready for dissemination. This phase includes the sub-processes and activities that enable statistical analysts to understand the statistics produced. For statistical outputs produced regularly, this phase occurs in every iteration. The Analyse phase and sub-processes are generic for all statistical outputs, regardless of how the data were sourced.

6.1   **Acquire ancillary information**

- This sub-process includes many ongoing activities involved with the gathering of intelligence, with the cumulative effect of building up a body of knowledge about a specific statistical domain. This knowledge is then applied to the current collection, in the current environment, to allow informed analyses. Acquiring a high level of domain intelligence will allow a statistical analyst to understand the data better, and to identify where results might differ from expected values. This allows better explanations of these results in sub-process 6.5 (Describe and explain).

6.2   **Calculate aggregates**

- This sub process creates aggregate data and population totals from micro-data. It includes summing data for records sharing certain characteristics, determining measures of average and dispersion, and applying weights from sub-process 5.6 to sample survey data to derive population totals.

6.3   **Prepare draft outputs**

- This sub-process is where domain intelligence is applied to the data collected to produce statistical outputs. It includes the production of additional measurements such as indices or seasonally adjusted series, as well as the recording of quality characteristics.

6.4   **Validate**

- This sub-process is where statisticians verify the quality of the outputs produced, in accordance with a general quality framework. Verification activities can include:
  - checking that the population coverage and response rates are as required;
  - comparing the statistics with previous cycles (if applicable);
  - confronting the statistics against other relevant data (both internal and external);
  - investigating inconsistencies in the statistics;
  - performing macro editing;

  - Verifying the statistics against expectations and domain intelligence.

6.5   **Describe and explain**

- This sub-process is where the in-depth understanding of the outputs is gained by statisticians. They use that understanding to interpret and explain the statistics produced for this cycle by assessing how well the statistics reflect their initial expectations, viewing the statistics from all perspectives using different tools and media, and carrying out in-depth statistical analyses.

6.6 Disclosure control and anonymise

- This sub-process ensures that the data (and metadata) to be disseminated do not breach the appropriate rules on confidentiality. This may include checks for primary and secondary disclosure, as well as the application of data suppression or perturbation techniques.

6.7 Finalize outputs for dissemination

- This sub-process ensures the statistics and associated information are fit for purpose and reach the required quality level, and are thus ready for dissemination. It includes:
    - completing consistency checks;
    - determining the level of release, and applying caveats;
    - collating supporting information, including interpretation, briefings, measures of uncertainty and
    - any other necessary metadata;
    - producing the supporting internal documents;
    - pre-release discussion with appropriate internal subject matter experts;
    - approving the statistical content for release.

7. Disseminate

This phase manages the release of the statistical products to customers. For statistical outputs produced regularly, this phase occurs in each iteration. It is made up of five sub-processes, which are generally sequential, from left to right, but can also occur in parallel, and can be iterative.

7.1 Update output systems

- This sub-process manages the update of systems where data and metadata are stored for dissemination purposes, including:
    - formatting data and metadata ready to be put into output databases;
    - loading data and metadata into output databases;
    - ensuring data are linked to the relevant metadata.

Note: formatting, loading and linking of metadata should preferably mostly take place in earlier phases, but this sub-process includes a check that all of the necessary metadata are in place and ready for dissemination.

7.2 Produce products

- This sub-process produces the products, as previously designed, to meet user needs. The products can take many forms including printed publications, press releases and web sites. Typical steps include:
    - preparing the product components (text, tables, charts etc.);
    - assembling the components into products;
    - editing the products and checking that they meet publication standards.

7.3 Produce 'quality statement'

This sub-process produces a quality report. This is a metadata required to declare the quality of a statistical product. This type of metadata should be documented according to the standard template for quality declaration.

7.4 Manage release of products

- This sub-process ensures that all elements for the release are in place including managing the timing of the release. It includes briefings for specific groups such as the press or ministers, as well as the arrangements for any pre-release embargoes. It also includes the provision of products to subscribers.

7.5 Market and promote products

- Whilst marketing in general can be considered to be an over-arching process, this sub-process concerns the active promotion and marketing of the statistical products produced in a specific statistical business process, to help them reach the widest possible audience. It includes the use of customer relationship management tools, to better target potential users of the products, as well as the use of tools including web sites, to facilitate the process of communicating statistical information to users.

7.6 Manage customer queries

- This sub-process ensures that customer queries are recorded, and that responses are provided within agreed deadlines. These queries should be regularly reviewed to provide an input to the over-arching quality management process, as they can indicate new or changing user needs.

8. Archive

This phase manages the archiving and disposal of statistical data and metadata. Given the reduced costs of data storage, it is possible that the archiving strategy adopted by a statistical organisation does not include provision for disposal, so the final sub-process may not be relevant for all statistical business processes. In other cases, disposal may be limited to intermediate files from previous iterations, rather than disseminated data.

8.1 Manage archive repository

- This sub-process concerns the management of one or more archive repositories. These may be databases, or may be physical locations where copies of data or metadata are stored. It includes:

  - maintaining catalogues of data and metadata archives, with sufficient information to ensure that individual data or metadata sets can be easily retrieved;
  - testing retrieval processes;
  - periodic checking of the integrity of archived data and metadata;
  - upgrading software-specific archive formats when software changes.

- This sub-process may cover a specific statistical business process or a group of processes, depending on the degree of standardisation within the organisation. Ultimately it may even be considered to be an over-arching process if organisation-wide standards are put in place.

8.2 Preserve data and associated metadata

- This sub-process is where the data and metadata from a specific statistical business process are archived. It includes:
  - identifying data and metadata for archiving in line with the rules defined in 8.1;
  - formatting those data and metadata for the repository;
  - loading or transferring data and metadata to the repository;
  - cataloguing the archived data and metadata;
  - Verifying that the data and metadata have been successfully archived.

8.3 Dispose of data and associated metadata

- This sub-process is where the data and metadata from a specific statistical business process are disposed of. It includes;
  - identifying data and metadata for disposal, in line with the rules defined in 8.1;
  - disposal of those data and metadata;
  - recording that those data and metadata have been disposed of.

9.   Evaluate

This phase manages the evaluation of a specific instance of a statistical business process. It logically takes place at the end of the instance of the process, but relies on inputs gathered throughout the different phases. For statistical outputs produced regularly, evaluation should, at least in theory occur for each iteration, determining whether future iterations should take place, and if so, whether any improvements should be implemented. However, in some cases, particularly for regular and well established statistical business processes, evaluation may not be formally carried out for each iteration. In such cases, this phase can be seen as providing the decision as to whether the next iteration should start from Phase 1 (Specify needs) or from some later phase (often Phase 4 (Collect)).

9.1   Gather inputs for programme review/ evaluation

- Evaluation material can be produced in any other phase or sub-process. It may take many forms, including feedback from users, process metadata, and system metrics and staff suggestions. Reports of progress against an action plan agreed during a previous iteration may also form an input to evaluations of subsequent iterations.
- This sub-process gathers all of these inputs, and makes them available for the person or team producing the evaluation.

9.2   Prepare evaluation report

- This sub-process analyses the evaluation inputs and synthesises them into an evaluation report. The resulting report should note any quality issues specific to this iteration of the statistical business process, and should make recommendations for changes if appropriate. These recommendations can cover changes to any phase or sub-process for future iterations of the process, or can suggest that the process is not repeated.

9.3   Quality plan

- This sub-process brings together the necessary decision-making power to form and agree an action plan based on the evaluation report. It should also include consideration of a mechanism for monitoring the impact of those actions, which may, in turn, provide an input to evaluations of future iterations of the process.

Quality management

This process is present throughout the model. It is closely linked to Phase 9 (Evaluate), which has the specific role of evaluating individual instances of a statistical business process. The overarching quality management process, however, has both a deeper and broader scope. As well as evaluating iterations of a process, it is also necessary to evaluate separate phases and sub-processes, ideally each time they are applied, but at least according to an agreed schedule. These evaluations can apply within a specific process, or across several processes that use common components.

Metadata management

Good metadata management is essential for the efficient operation of statistical business processes. Metadata are present in every phase, either created or carried forward from a previous phase. The key challenge is to ensure that they are captured as early as possible, and stored and transferred from phase to phase alongside the data they refer to. A metadata management strategy and system(s) are therefore vital to the operation of this model.

Annexure E: Mapping quality indicators to activities in the Statistical Value Chain (SVC)

Activities of the statistical value chain		Quality dimensions and indicators	
Phases	**Sub-processes**	**Quality dimension**	**Quality indicator**
Need	Determine need for information	Prerequisites of quality	1.1 The responsibility for producing statistics is clearly specified.
Need	Determine need for information	Prerequisites of quality	1.2 Standards and policies are in place to promote consistency of methods and results.
Need	Establish output objectives	Prerequisites of quality	1.3 Data sharing and coordination among data-producing agencies are clearly specified.
Need	Establish output objectives	Prerequisites of quality	1.4 Measures are in place to ensure that individual data are kept confidential, and used for statistical purposes only.
Need	Determine need for information	Relevance	2.1 Have both the internal and external users of the data been identified?
Need	Determine need for information	Relevance	2.2 Is there a process to identify user needs?
Need	Establish output objectives	Relevance	2.3 Are user needs and the usage of statistical information analysed?
Need	Establish output objectives	Relevance	2.4 Changes are made as a result of user needs assessments.
Need	Check data availability	Relevance	2.5 To what extent is the primary data appropriate for the statistical product produced?
Need	Prepare business case	Timeliness	4.4 Periodicity of release.
Need	Check data availability	Accessibility	5.1 Legal arrangements are in place to allow access to administrative records via manual, automated or electronic processes.
Need	Frame and sample methodology, Data collection	Methodological soundness	8.1 The scope of the study is consistent with accepted standards, guidelines or good practices.
Need	Check data availability	Integrity	9.5 Choice of source data, techniques and dissemination decisions are informed solely by statistical considerations.
Design	Detailed project plan	Prerequisites of quality	Resources are commensurate with the needs of statistical programmes. • Staff • Facilities

Activities of the statistical value chain		Quality dimensions and indicators	
Phases	**Sub-processes**	**Quality dimension**	**Quality indicator**
			• Computing resources • Financing
Design	Detailed project plan	Prerequisites of quality	1.7 Measures to ensure efficient use of resources in 1.6 are implemented.
Design	Frame and sample methodology	Accuracy	3.5 Register/frame maintenance procedures are adequate. • updates. • quality assurance ▪ data audit.
Design	Data collection	Accuracy	3.6 Are data collection systems sufficiently open and flexible to cater
Design	Data collection	Accuracy	3.7 Description of record-matching methods and techniques used on the administrative data sources. • match rate as a percentage of total records • measure of false negative matches (missed matches) • measure of false positive matches (mismatches)
Design	Outputs, Frame and sample methodology, Tabulation plan/variables, Data collection Statistical processing methodology, Define archive rules,	Comparability and coherence	7.3 Data across comparable series, or source data are based on common frames, identifiers, concepts and definitions, and classifications, and departures from these are identified in the metadata.
Design	Tabulation plan/variables; Identify concepts	Methodological soundness	8.1 Concepts, definitions, and classifications used follow accepted standards, guidelines or good practices (national, international, peer-agreed).
Design	Statistical processing methodology	Methodological soundness	8.2 Methodologies used follow accepted standards, guidelines or good practices (national, international, peer-agreed), viz.: • questionnaire design • sampling • sampling frame design • frame maintenance

Activities of the statistical value chain		Quality dimensions and indicators	
Phases	**Sub-processes**	**Quality dimension**	**Quality indicator**
			• piloting • data collection • editing and imputation of data • analysis of data • revision data
Design	Tabulation plan/variables Identify concepts	Methodological soundness	8.1 Concepts, definitions, and classifications used follow accepted standards, guidelines or good practices (national, international, peer-agreed)
Design	Statistical processing methodology	Methodological soundness	8.3 Methodologies used follow accepted standards, guidelines or good practices (national, international, peer-agreed), viz.: • questionnaire design • sampling • sampling frame design • frame maintenance • piloting
Process	Integrate data	Comparability and coherence	7.4 A common set of identifiers (for the purpose of record matching) exist and have been agreed upon by data producers.
Analyse	Validate	Accuracy	3.1 Measures of sampling errors for key variables are calculated. Amongst others these are: • standard error • coefficient of variation (CV) • confidence interval (CI) • mean square error (MSE) • design effect (DEFF)

Activities of the statistical value chain		Quality dimensions and indicators	
Phases	**Sub-processes**	**Quality dimension**	**Quality indicator**
Analyse	Validate	Accuracy	3.2 Measures of non-sampling errors are calculated, viz.: • Frame coverage errors • Systematic errors • Measurement errors • Processing errors • Model assumption errors • Non-response errors
Analyse	Define archive rules, Processing systems and workflow Validate	Comparability and coherence	7.2 Statistics are consistent or reconcilable over time
Analyse	Validate	Comparability and coherence	7.3 Statistics are checked for consistency with those obtained through other data sources.
Disseminate	Produce quality statement	Accuracy	3.3 Data from the primary source have been quality assessed..
Disseminate	Manage release of products	Timeliness	4.1 Average time between the end of reference period and the date of the preliminary results.
Disseminate	Manage release of products	Timeliness	4.2 Average time between the end of reference period and the date of the final results.
Disseminate	Manage release of products	Accessibility	5.1 Are statistical products (e.g. data, metadata) available to the public?
Disseminate	Manage release of products	Accessibility	5.2 Rules governing the restricted availability of administrative records are well described and documented.
Disseminate	Produce dissemination products	Accessibility	5.3 Types of media and/or channels used for sharing data amongst stakeholders are adequate and preserve confidentiality.
Disseminate	Produce dissemination products	Accessibility	5.4 Data is accessible in a format beyond the producing agency.
Disseminate	Manage release of products	Accessibility	5.5 Statistics are released on a pre-announced schedule.
Disseminate	Manage release of products	Accessibility	5.6 Statistical products are made available to all users at the same time.

Activities of the statistical value chain		Quality dimensions and indicators	
Phases	**Sub-processes**	**Quality dimension**	**Quality indicator**
Disseminate	Manage customer queries	Accessibility	5.7 Statistics/administrative records not routinely disseminated are made available upon request.
Disseminate	Market and promote products	Accessibility	5.8 User support services exist and are widely publicised.
Disseminate	Manage release of products	Accessibility	5.9 Does a data dissemination policy exist, and is it accessible?
Disseminate	Manage release of products	Accessibility	5.10 Does the pricing policy governing dissemination exist, and is it accessible?
Disseminate	Market and promote products	Accessibility	5.11 Catalogues of publications and other services are available to users of statistics.
Disseminate	Manage release of products	Accessibility	5.12 Metadata are readily accessible to users.
Disseminate	Produce dissemination products	Interpretability	6.1 Documented metadata (definitional, operational, methodological, system and dataset) are sufficient to understand data.
Disseminate	Produce dissemination products	Interpretability	6.2 Statistics are presented in a clear and understandable manner. Statistical releases contain a summary of the key findings.
Disseminate	Produce dissemination products	Interpretability	6.3 Statistical releases contain a summary of the key findings.
Disseminate	Processing systems and workflow	Methodological soundness	8.4 Are revisions schedule followed? Are they regular and transparent?
Disseminate	Manage release of products	Methodological soundness	8.5 Preliminary and revised data are identified in the metadata.
Disseminate	Manage release of products	Methodological soundness	8.6 Studies of revisions and their findings are made public.
Disseminate	Manage release of products	Integrity	9.1 The terms and conditions, including confidentiality, under which statistics are collected, processed and disseminated are available to the public and follow the UN principles of official statistics.
Disseminate	Produce quality statement	Integrity	9.2 Describe the conditions under which policy-makers, specifically government, may have access to data before release. Are the conditions published?
Disseminate	Produce quality statement	Integrity	9.3 Advance notice is given of major changes in methodology and source data.

Activities of the statistical value chain		Quality dimensions and indicators	
Phases	**Sub-processes**	**Quality dimension**	**Quality indicator**
	Check data availability		
Disseminate	Outputs, Frame and sample methodology, Tabulation plan, Data collection Statistical processing methodology, Define archive rules	Integrity	9.4 Government commentary, when data are released, should be identified as such, and not be seen as part of the official statistics.
Evaluation	Gather inputs for programme review/evaluation	Prerequisites of quality	1.5 Measures to oblige response are ensured through law.
Evaluation	Gather inputs for programme review/evaluation	Relevance	2.6 Is there a process to determine the satisfaction of users with the statistical information?
All phases	All phases	Timeliness	4.3 Production activities within the statistical value chain are within the planned timelines, viz.: • data collection • data processing • data analysis • dissemination
All phase	All phases	Integrity	9.5 Ethical guidelines for staff behaviour are in place and are well known to the staff.

Annexure F: Differences between SASQAF Edition 1 and SASQAF Edition 2

The following table outlines the adjustments and additions made to the quality indicators in the first edition of SASQAF:

Old Indicator number	Indicator (SASQAF Edition 1) – September 2008	New Indicator number	Indicator (SASQAF operational standards and guidelines)
0.3	Data sharing procedures and coordination among data-producing agencies are clearly specified and adhered to.	1.3	Data sharing and coordination among data-producing agencies are clearly specified
	New indicator	1.5	Measures to oblige response are ensured through law.
0.6	Measures to ensure efficient use of above resources in 0.5 are implemented.	1.7	Measures to ensure efficient use of resources in 0.6 are implemented.
1.4	Changes made as a result of user needs assessments.	2.4	Changes are made as a result of user needs assessments.
1.5	Is there a process to determine the satisfaction of users?	2.5	Is there a process to determine the satisfaction of users with the statistical information?
1.6	To what extents are the primary data (e.g. administrative data and other data) appropriate for the statistical product produced?	3.3	Moved to accuracy
1.7	Were special requests for estimates of statistical characteristics met?		Deleted
2.1	Measures of sampling errors for key variables are calculated. Amongst others these are: • Standard error • Coefficient of variation (CV) • Confidence interval (CI) • Mean square error (MSE).	3.1	Measures of sampling errors for key variables are calculated. Amongst others these are: • standard error • coefficient of variation (CV) • confidence interval (CI) • mean square error (MSE) • design effect (DEFF)

Old Indicator number	Indicator (SASQAF Edition 1) – September 2008	New Indicator number	Indicator (SASQAF operational standards and guidelines)

2.2	Measures of non-sampling errors are calculated, viz.:
	• frame coverage errors
	  ○ under-coverage errors
	  ○ over-coverage errors
	• duplication in the frame/register used to conduct a survey
	• the number of statistical units out of scope(i.e. number of ineligible units)
	• misclassification errors
	• systematic errors to determine the extent of bias introduced for both administrative records and surveys
	• measurement errors
	  ○ questionnaire effects
	  ○ data collection mode effects
	  ○ interviewer effects
	  ○ respondent effects
	• processing errors
	  ○ data entry errors rates
	  ○ coding errors
	  ○ editing failure rates
	  ○ imputation rates
	• model assumption errors
	• non-response errors
	  ○ overall response rate
	  ○ item response rates
	  ○ unit non-response (e.g. weighted and un-weighted response rates).
		3.2	Measures of non-sampling errors are calculated, viz.:
			• Frame coverage errors (e.g. duplication in the frame/register used to conduct a survey, number of statistical units out of scope (i.e. number of ineligible units)
			
			• Misclassification errors
			• Systematic errors to determine the extent of bias introduced for both administrative records and surveys
			• Measurement errors (e.g. questionnaire effects, data collection effects, interviewer effects and respondent effects)
			• Processing errors (e.g. data entry errors rates, coding errors, editing failure rates, imputation rates)
			• Model assumption errors
			• Non-response errors (e.g. item non-response rates, unit non-response rates and overall response rates).

1.6	Was moved from Relevance	3.3	To what extent is the primary data appropriate for the statistical product produced?

Old Indicator number	Indicator (SASQAF Edition 1) – September 2008	New Indicator number	Indicator (SASQAF operational standards and guidelines)
2.3	Data from the primary source have been quality assessed • accuracy. • timeliness. • coherence.	3.4	Data from the primary source have been quality assessed.
2.4	Does an agreement for relevant deadlines for transfer of data from the primary source exist and are they adhered to?		Deleted
2.6	Are data collection systems sufficiently open and flexible to cater for new developments (e.g., changes in definitions, classifications, etc.)?	3.6	Are data collection systems sufficiently open and flexible to cater for new developments?
2.7	Description of record-matching methods and techniques used on the administrative data sources. • match rate as a percentage of total records • measure of false negative matches (same unit but match was missed) • measure of false positive matches (record matched but relate to separate entities)	3.7	Description of record-matching methods and techniques used on the administrative data sources. • match rate as a percentage of total records • measure of false negative matches (missed matches) • measure of false positive matches (mismatches)
3.1	Average time between the end of reference period and the date of the first results.	4.1	Average time between the end of reference period and the date of the preliminary results.
3.4	Report on the frequency of release.	4.4	Periodicity of release
3.5	Punctuality of time schedule for publication.		Deleted
4.1	Are data and information available to the public?	5.1	Are statistical products (e.g. data, metadata) available to the public?
4.3	Legal arrangements are in place to access administrative records via manual/automated/electronic systems.		Deleted
4.4	Types of media/channels used for sharing data amongst stakeholders are adequate and preserve confidentiality.	5.3	Types of media and/or channels used for sharing data amongst stakeholders are adequate and preserve confidentiality
4.7	Statistics are made available to all users at the same time.	5.6	Statistical products are made available to all users at the same time.
4.9	User support services are widely publicised.	5.8	User support services exist and are widely publicised.

Old Indicator number	Indicator (SASQAF Edition 1) – September 2008	New Indicator number	Indicator (SASQAF operational standards and guidelines)
4.10	Does a data dissemination policy exist, and is it maintained and accessible?	5.9	Does a data dissemination policy exist, and is it accessible?
4.11	Does the pricing policy governing dissemination exist, and is it available to users?	5.10	Does the pricing policy governing dissemination exist, and is it accessible?
4.12	Catalogue systems (for survey, administrative records and other services) to identify information are available to users and are updated regularly.	5.11	Catalogues of publications and other services are available to users of statistics.
4.13	Metadata (a full range of information on underlying concepts, definitions, classifications, methodology, data sources, accuracy, etc.) are documented, available and readily accessible to users	5.12	Metadata are readily accessible to users.
5.1	Availability of concepts and definitions, classifications underlying the data (survey and administrative records). Differences from accepted standards, guidelines or good practices are annotated.	6.1	Documented metadata (definitional, operational, methodological, system and dataset) are sufficient to understand data.
5.2	Documents on scope, basis of recording, data sources and statistical techniques (methodology) used are available. Differences from accepted standards, guidelines or good practices are annotated.		5.1 and 5.2 merged
	New indicator	6.2	Statistics are presented in a clear and understandable manner
5.3	All the statistical releases produced are accompanied by primary messages clarifying the key findings.	6.3	Statistical releases contain a summary of the key findings.
6.1	Data within series and administrative systems are based on common frameworks, such as concepts, definitions, classifications, and methodologies, and departures from these are identified in the metadata.	7.1	Data within series and administrative systems are based on common concepts and definitions, classifications, and methodology, and departures from these are identified in the metadata.
6.2	Statistics are consistent and reconcilable over time.	7.2	Statistics are consistent or reconcilable over time.
6.3	Data across comparable series, or source data are based on common frames, common identifiers, concepts, definitions, and classifications, and departures from these are identified in the metadata.	7.3	Data across comparable series, or source data are based on common frames, identifiers, concepts and definitions, and classifications, and departures from these are identified in the metadata.

Old Indicator number	Indicator (SASQAF Edition 1) – September 2008	New Indicator number	Indicator (SASQAF operational standards and guidelines)
6.4	Statistics are checked for consistency with those obtained through other data sources (identify comparable datasets and incomparable ones).	7.4	Statistics are checked for consistency with those obtained through other data sources.
7.3	Methodologies used follow accepted standards, guidelines or good practices (national, international, peer-agreed), viz.: • Questionnaire design • Sampling methods • Sample frame design • Frame maintenance • Piloting • Standard collection methods • Standard editing and imputation methods • Standard analytical methods.	8.3	Methodologies used follow accepted standards, guidelines or good practices (national, international, peer-agreed), viz.: • questionnaire design • sampling methods • sampling frame design • frame maintenance • piloting • standard collection methods • standard editing and imputation methods • standard analytical methods • revision methods
7.4	Revisions schedule followed (explain the extent to which it is regular and transparent).	8.4	Are revisions schedule followed? Are they regular and transparent?
8.3	Advance notice is given of major changes in methodology, source data and statistical techniques.	9.3	Advance notice is given of major changes in methodology and source data.
8.4	Ministerial commentary, when data are released, should be identified as such, and not be seen as part of the official statistics.	9.4	Government commentary, when data are released, should be identified as such, and not be seen as part of the official statistics
8.5	Choice of source data, techniques and dissemination decisions are informed solely by statistical considerations (without political interference).	9.5	Choice of source data, techniques and dissemination decisions are informed solely by statistical considerations.
8.6	Ethical guidelines for staff behaviour are in place and are well known to the staff (professional code of conduct).	8.6	Ethical guidelines for staff behaviour are in place and are well known to the staff