

## SC008, Data Quality Standard for Surveys

	<i>outlier, does not necessarily indicate incorrect data.</i>	
7.2.3	Data editing procedures should include cross-checks between related values, and ideally, should also include checks on summary totals for such related fields. <i>For example, this would include checking subtotals in Snap tables.</i>	6.2
7.2.4	For repeated surveys, editing procedures should compare values submitted previously for a specific school. Ranges are set for acceptable changes from one year to the next, and changes in values for individual schools are checked and verified against these ranges.	6.2
7.2.5	Editing procedures should include checks on the validity of values, such as checks against list tables for codes, values falling outside acceptable ranges, as well as required fields or sections that have not been completed.	6.2
7.2.6	Data cleaning procedures must be revised to take into account changes in the data collection instruments from one year to the next, which makes it impossible to successfully compare data from one year to the next. Such revisions must be documented for metadata purposes.	6.3
7.2.7	Methodologies used for editing should follow accepted standards, guidelines or good practices (national, international and peer-agreed).	7.3
7.2.8	Care should be taken if there are revisions to data ( <i>for instance by a province that submitted an initial dataset and then submitted a revised dataset</i> ). If any data is submitted after publication of the results and such data is included in the national dataset, this additional data need to be tagged as such in the dataset.	7.5

**2.3.2.8 STEP 8: Aggregation and analysis****Objectives to be reached during this step:**

When the collected data is analysed, it is important to use sound methodologies for the analysis. The focus of the analysis should be on providing answers to key questions and not just to show all the results. Of further importance is to document all the procedures and methods used. Documentation of the procedures provides evidence of the acceptability and quality of the obtained results, and also ensures the repeatability of the results.

It should be noted that an analysis could involve providing descriptive statistics, such as aggregated totals and counts, but it could also comprise an in-depth (technical) analysis of the collected data. The guidelines cover general aspects that relate to both these kinds of analysis, as well as aspects that relate primarily to technical analysis.

**QG8: Guidelines pertaining to actions carried out during this step:**

	<b>Action points: Aggregation and analysis</b>	<b>Related SASQAF indicator</b>
<b>8.1</b>	<b>General</b>	
8.1.1	The methods planned for the analysis must be motivated by the objectives of the study and by statistical considerations, rather than by the data itself or by political interference. Therefore, the choice of the level of analysis, as well as the relevant analysis methods, should preferably be finalised before data collection takes place.	8.5

## SC008, Data Quality Standard for Surveys

	<b>Action points: Aggregation and analysis</b>	<b>Related SASQAF indicator</b>
8.1.2	The analysis should, as far as possible, be aligned with the analysis plan that was drawn up during the planning and design steps, in order to ensure that the analysis provides answers that are in line with the objectives of the study.	
8.1.3	Electronic systems designed for data analysis must be designed so as to be compatible with the systems designed for data collection and capture. <i>If systems are developed by a third party (i.e. outside the Department of Basic Education), then this software product must link up with the Department of Basic Education capture and analysis tools and produce analysis outputs that are acceptable to the department.</i>	
8.1.4	The terms and conditions, including confidentiality, under which statistics are collected, processed and disseminated, and which was planned for the survey, should be adhered to during the analysis.	8.1
<b>8.2</b>	<b>Imputation of missing values</b>	
8.2.1	Imputation is the responsibility of national department. Provincial departments are not responsible for imputations.	
8.2.2	If imputation routines are used to fill in missing data, it is important to limit imputation to variables for which values are missing randomly, and not due to a systematic difference or bias. It is furthermore important to study and find an imputation method that is appropriate for a specific survey. Methods used must be reported in the metadata.	7.3
<b>8.3</b>	<b>Technical analysis</b>	
8.3.1	All assumptions underlying the analysis should be stated clearly, and appropriate methods for the analysis should be selected accordingly.	7.3
8.3.2	Assumptions, analysis methods and any limitations to the analysis should be documented for metadata purposes.	4.13
8.3.3	Measures of sampling errors for key variables must be calculated from the actual data that was collected, and these measures must be compared to the targets set for these values during the survey planning. As far as possible, results should be reported in accordance with the relevant standard errors or confidence intervals. This information must also be provided in the metadata that reports on data quality.	2.1, 4.13
8.3.4	Results should be reported in such a way that it supports appropriate conclusions from the analysis. Assumptions, limitations, data quality issues, response rates, potential biases and other important aspects that affect the analysis should be taken into account.	
8.3.5	If possible, measures of non-sampling errors must be calculated – such as errors picked up in the sample frame or during editing. It is also important to consider the effect of errors, such as deficiencies in sample frame or instruments, or other underlying biases to the analysis on the quality of the data and analysis. These errors and their impact should be documented for metadata purposes.	2.2
8.3.6	With regard to census surveys, it is important to report on non-response. The extent of non-response, possible reasons for the non-response and measures (if any) taken to address the non-response should be documented and reported.	2.2
8.3.7	Documentation must be provided to describe record-matching methods and techniques used (for instance matching of collected data to the MLI), and other analysis and editing methods used. It is important to record what was done, what errors were found and how these errors were rectified.	2.7

## SC008, Data Quality Standard for Surveys

**2.3.2.9 STEP 9: Publication of results****Objectives to be reached during this step:**

Publication is done in order to make results available to users in an accessible way. In order to ensure proper use of the results, metadata, explaining the contents of the data and the methods used to obtain and analyse it, should be provided with the results.

**QG9: Guidelines pertaining to actions carried out during this step:**

	<b>Action points: Publication of results</b>	<b>Related SASQAF indicator</b>
<b>9.1</b>	<b>Before disseminating the data</b>	
9.1.1	Determine whether there is a data dissemination policy relevant to the survey. <i>For instance, this would be a publication policy that governs the contents, the form in which it is distributed, to whom it is distributed, etc<sup>4</sup>.</i>	4.10
9.1.2	Determine whether there are clearance procedures ensuring that the document has been approved by the unit producing it and whether there are guidelines on the issues to be addressed before publication could be approved. Ensure that these prescriptions are followed before publication of the results.	4.10
9.1.3	The Director-General: Basic Education will need to approve the contents before releasing it to the general public. Before information is referred to the DG: BE for approval, it must furthermore be approved by the relevant staff members internal to the department. These include the staff members responsible for conducting the survey, and also the EMIS Officer. Determine the requirements and processes in obtaining approval for release and ensure that these are correctly followed.	4.10
9.1.4	For qualitative data (e.g. <i>whole school evaluation</i> ), the publication should be in a format that would allow the data to be integrated and analysed.	4.10
9.1.5	Adding or removing results due to political pressure should be avoided.	8.5
9.1.6	Ministerial commentary when data is released should be identified as such, and should not be seen as part of the official statistics.	8.4
<b>9.2</b>	<b>Reporting on the quality of the results</b>	
9.2.1	Statistical measures of data quality should be reported with the results. It should also be indicated how the obtained accuracy differed from the planned levels, including the response rate that was actually achieved. It should be noted that it could be important to indicate if a particularly bad response was obtained with regard to certain aspects, even if the overall response rate was satisfactory. <i>This helps the user to understand what the results can and cannot be used for.</i>	2.2
9.2.2	The metadata collected throughout the survey should be collated. Issues recorded in the metadata that affect data quality or influence the use of the results should be published in conjunction with the results. <i>This could include objectives, methodology, data sources, accuracy, instruments, sampling plan, editing and imputation.</i>	4.13
9.2.3	There will be metadata that should be published with the data, metadata that may be released and metadata that will only be released if legally	4.13

<sup>4</sup> See document SC012, Standards for Publication of Education Statistics in this regard.

## SC008, Data Quality Standard for Surveys

	required to do so. Metadata for internal use carries a lot of detail, while metadata for publication carries less detail. Determine which portions of the recorded metadata should be released with the results. <i>A person who reads the report must understand how the study was conducted and what exceptions or unanticipated aspects were found in the data or the processes.</i>	
9.2.4	Concepts, definitions and classifications underlying the data should be published in order to clarify the data. <i>This should form part of the reference section in all reports.</i>	5.1
9.2.5	Information on the scope, sampling, data sources and statistical techniques used should be published. Differences relating to accepted standards, guidelines or good practices should be pointed out and motivated.	5.1, 5.2
9.2.6	Some information must also be published in terms of the processing of the data. For internal use, metadata should include the actual information one would need to process the data, such as the statistical code used to run the analysis, and this code must be available if requested.	4.13, 5.2
9.2.7	All the statistical information published (tables, etc.) must be accompanied by an explanation of the contents, and clarification of the key findings. <i>There must be sufficient information to guide the user in the interpretation of the information.</i>	5.3
9.2.8	If there are revisions of data, or if data collection continues after publication of the results, information pointing this out must be included with the published results. <i>This must be done to warn users about differences between the published results and the survey data disseminated.</i>	7.4, 7.5, 7.6

<b>9.3</b>	<b>General issues</b>	
9.3.1	All activities should be monitored and controlled in order to ensure that publication deadlines are met.	3.2, 3.3, 3.5
9.3.2	When the results are published, it should be indicated whether this is a repeated survey and how often the survey has been repeated.	3.4
9.3.3	Ensure that the results will reach key users by employing the appropriate types of media/channels. Make sure that the correct level of confidentiality is maintained in the published results. <i>Do not release raw data or data that allows individuals to be identified.</i>	4.1, 4.4
9.3.4	Make sure that the release of data before publication adheres to the planned timelines and conditions.	8.2
9.3.5	Put systems in place to obtain feedback on the published results from key users. Furthermore, ensure that published results are provided to those who provided data for the survey.	1.5, 1.7

**2.3.2.10 STEP 10: Handling queries****Objectives to be reached during this step:**

The survey is only completed once all data has been finalised and is ready for release. Furthermore, once a study has been completed and results are published, there may be requests from users for the data attached to the survey. In clause 24 of the Education Information Policy, it is stated that the needs of researchers to analyse data pertaining to the education sector should be anticipated and that data should be prepared accordingly.

## SC008, Data Quality Standard for Surveys

It is important to disseminate data from the study according to the same principles as the release of the publication itself, and therefore many of the requirements for release of the data are similar to those for publication purposes. However, in releasing data it is also important to ensure that the correct context is maintained. The context provided in terms of the entire published report should also be clearly applied to sections of the report released in the form of data or as individual summary tables.

**QG10: Guidelines pertaining to actions carried out during this step:**

	<b>Action points: Handling queries</b>	<b>Related SASQAF indicator</b>
<b>10.1</b>	<b>General</b>	
<b>10.1.1</b>	All the requirements in terms of clearance before release, accompanying metadata, confidentiality and comments on data quality listed in terms of publication of the results, should also be followed for the release of data. <i>For queries and for reports published on the Internet, make sure that the information retains the correct context by providing appropriate accompanying metadata, since extracts from the report may be misinterpreted if read out of context.</i>	2.2, 4.10, 4.13, 5.1, 5.2, 5.3, 8.5
10.1.2	The types of media/channels used for providing requested data should be adequate and should preserve confidentiality. <i>Also plan for a way to provide the relevant metadata in conjunction with the data.</i>	4.4
10.1.3	Catalogue systems to identify available information from the survey should be made available to users, and this catalogue should be regularly updated.	4.12
10.1.4	Data must be made accessible to users by ensuring that the data could be provided in a variety of formats that will satisfy the requirements of all users.	4.5
10.1.5	Make sure that data from more than one instance of a repeat survey is provided in a comparable format.	6.2, 6.3
10.1.6	Preliminary and revised data should also be identified in the data.	7.5
10.1.7	Data that is released from a specific survey should contain references to that survey. The date on which the survey was completed/finalised, as well as the date on which the data was supplied, must be indicated on the dataset.	
10.1.8	Publication policies should address more than just the published hard copy and should include policies pertaining to electronic copies, Internet versions of reports, queries and all output provided to users.	4.10
10.1.9	Make sure that part of the process for providing data includes a check on the methods or codes used to extract data in response to queries. <i>Someone must double-check queries used to extract data.</i>	
10.1.10	Put in place systems to obtain feedback on the published results from key users.	1.5, 1.7
10.1.11	Put in place systems to log all the queries in order to keep all queries responded to, on file.	

## SC008, Data Quality Standard for Surveys

**APPENDIX A: SASQAF Data Quality Indicators**

Tables 2 to 10 that follow in this subsection provide the details on the SASQAF standard to be used to assess data quality, organised into the eight quality dimensions and the prerequisites for quality. Note that the numbering convention for the SASQAF indicators is to use the dimension as the first level of the number, so that the indicator numbered 3.2 refers to the second indicator under dimension 3. The table headings list the number of the dimension along with the name of the dimension.

Each table contains an indicator number, a description of the indicator, as well as the four possible assessment levels. When a certain statistic is measured in terms of an indicator, it is placed in one of the four levels, with level 4 indicating best and level 1 lowest quality.

**Table 2 Prerequisites of quality (SASQAF dimension 0)**

Indicator number	Indicator description	Assessment Levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
0.1	The responsibility for producing statistics is clearly specified.	The responsibility for producing statistics is explicitly specified through a legal framework.	The responsibility for producing statistics is specified through a legal framework.	The responsibility for producing statistics is implied through a legal framework.	The responsibility for producing statistics is not specified.
0.2	Standards and policies are in place to promote consistency of methods and results.	All standards and policies are in place to promote consistency of methods and results, and are adhered to.	The majority of standards are in place to promote consistency of methods and results.	Some standards are in place to promote consistency of methods and results.	No standards are in place to promote consistency of methods and results.
0.3	Data sharing procedures and coordination among data-producing agencies are clearly specified and adhered to.	Data sharing procedures and coordination among data-producing agencies are explicitly specified through a legal framework.	Data sharing procedures and coordination among data-producing agencies are specified through a legal framework.	Data sharing procedures and coordination among data-producing agencies are implied through a legal framework.	Data sharing procedures and coordination among data-producing agencies are not specified.
		A data-sharing policy exists and is regularly updated and adhered to.	A data-sharing policy exists and for the most part is adhered to. It may not be up to date.	A data-sharing policy exists, but is rarely adhered to. It may not be up to date.	No data-sharing policy exists.
0.4	Measures are in place to ensure that individual data are kept confidential, and used for statistical purposes only.	Measures (e.g. policies, documented procedures) exist and are fully enforced so that individual data are always kept confidential.	Measures exist and are partially enforced so that individual data are always kept confidential.	Measures exist, but are not enforced to always keep individual data confidential.	There are no measures that ensure confidentiality.

## SC008, Data Quality Standard for Surveys

Indicator number	Indicator description	Assessment Levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
0.5	Resources are commensurate with the needs of statistical programmes (staff, facilities, computing resources, financing)	All resources are completely commensurate with statistical programmes.	Resources are partially commensurate with statistical programmes.	Resources are inadequately commensurate with statistical programmes.	Resources are not commensurate with statistical programmes.
0.6	Measures to ensure efficient use of the above resources in 0.5 are implemented.	Measures (e.g. project plans and sign-off documentation) to ensure efficient use of resources are systematically implemented.	Measures to ensure efficient use of resources are often implemented.	Measures to ensure efficient use of resources are seldom implemented.	Measures to ensure efficient use of resources are not implemented.
0.7	Processes are in place to focus on, monitor and check quality.	Processes are consistently in place to focus on, monitor and check quality.	Processes are to some extent in place to focus on, monitor and check quality.	Processes are seldom in place to focus on, monitor and check quality.	Processes are not in place to focus on, monitor and check quality.

Table 3 Relevance (SASQAF dimension 1)

Indicator number	Indicator description	Assessment Levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
1.1	Have both the internal and external users of the data been identified?	All users of the data have been identified with their most recent contact details.	All users of the data have been identified with some of the contact details not up to date.	Attempts have been made to create a user list.	No attempt has been made to create a user list.
1.2	Is there a process to identify user needs?	User needs are identified as a matter of course.	User needs are usually, but not always, identified.	User needs are identified on an ad hoc basis.	No attempt is made to identify user needs.
1.3	Are user needs and the usage of statistical information analysed?	User needs and the usage of statistical information are always analysed.	User needs and the usage of statistical information are often analysed.	User needs and the usage of statistical information are seldom analysed.	There is no effective interaction with users.
1.4	Changes made as a result of user needs assessments.	The results of the assessment are always built into the corporate processes and influence decisions on the design of the survey/series.	The results of the assessment are often built into the corporate processes and influence decisions on the design of the survey/series.	The results of the assessment are seldom built into the corporate processes and influence decisions on the design of the survey/series.	No action taken to incorporate the results of assessments.

## SC008, Data Quality Standard for Surveys

Indicator number	Indicator description	Assessment Levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
1.5	Is there a process to determine the satisfaction of users?	User satisfaction is measured and to a large extent has made an impact on the output.	User satisfaction is measured and to some extent has made an impact on the output.	User satisfaction is measured but has made no impact on the output.	User satisfaction is not measured.
1.6	To what extent are the primary data (e.g. administrative data and other data) appropriate for the statistical product produced?	The primary data are fully aligned to the statistical product released.	The primary data are mostly aligned to the statistical product released.	The primary data have limited relevance to the statistical product released.	The primary data are not at all relevant to the statistical product released.
1.7	Were special requests for estimates of statistical characteristics met?	All special requests were met.	Some special requests were met.	The majority of special requests were not considered.	No special requests were met.

Table 4 Accuracy (SASQAF dimension 2)

Indicator number	Indicator description	Assessment Levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
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)	Sampling errors are calculated for the main variables and are available for the other variables on request, and fall within acceptable standards.	Sampling errors are calculated and made available for the main variables, and fall within acceptable standards.	Sampling errors are calculated but not made available, and fall outside the acceptable standards.	No sampling errors are calculated.



## SC008, Data Quality Standard for Surveys

Indicator number	Indicator description	Assessment Levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
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 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 (questionnaire effects, data collection mode effects, interviewer effects, respondent effects), processing errors (data entry error rate, coding errors, editing failure rates, imputation rates), model assumption errors, non-response errors (overall response rate, item response rate, unit non-response, weighted and unweighted response rates)	Non-sampling errors are extensively described and analysed, and the measures fall within acceptable standards.	Non-sampling errors are described and analysed, and the measures are not far off from acceptable standards.	Non-sampling errors are described and analysed, and the measures are far off from acceptable standards.	Non-sampling errors are not described.

## SC008, Data Quality Standard for Surveys

Indicator number	Indicator description	Assessment Levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
2.3	Data from the primary source have been quality assessed: Coverage, timeliness, coherence)	Quality declaration is attached and shows that data comply with acceptable standards.	Quality declaration is attached and shows that the deficiencies in the data do not invalidate use of the data.	Quality declaration is attached and shows that data deviate significantly from acceptable standards.	Quality declaration is not attached.
2.4	Does an agreement for relevant deadlines for transfer of data from the primary source exist and are they adhered to?	Measures (agreements, documented procedures) exist to ensure that agreed deadlines are adhered to.	Measures exist to ensure adherence to agreed deadlines but there are minor discrepancies regarding adherence.	Deadlines for reporting exist with no follow-up procedures to ensure the timely receipt of data.	No deadlines for reporting and no procedures to ensure timely receipt of data exist.
2.5	Register / frame maintenance procedures are adequate.	Maintenance and update procedures of register/frame are adequate, thoroughly documented and performed on a regular basis.	Maintenance and update procedures are adequate and performed on a regular basis, but are not thoroughly documented.	Maintenance and update procedures are inadequate and are performed on an ad hoc basis. Some documentation exists.	No maintenance and update procedures exist.
	Updates	Updates are typically live and are registered on the occurrence of the event.	Updates are typically after the event, but occur at regular intervals.	Updates are typically after the event, but occur on an ad hoc basis.	No maintenance and update procedures exist.
	Quality assurance	A regular follow up survey is conducted based on a sample drawn from the administrative records and matches the frequency of the release.	A follow up survey is conducted but is inadequate given the frequency of the release.	The follow up survey is conducted on an ad hoc basis	No follow up survey is conducted.
	Data audit	An analysis of alternate data source/s is conducted to determine the cause, extent and type of errors in the administrative record system / frame and matches the frequency of the release	An analysis of alternate data sources is done on a regular basis but is inadequate given the frequency of the release.	An analysis of alternate data sources is done on an ad hoc basis	No analysis of alternate data sources is done.

## SC008, Data Quality Standard for Surveys

Indicator number	Indicator description	Assessment Levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
2.6	Are data collection systems sufficiently open and flexible to cater for new developments (e.g., changes in definitions, classifications, etc)?	Data collection programmes are sufficiently robust, with changes causing minimal impact on systems.	Data collection programmes are sufficiently robust, with changes causing significant impact on systems.	Although the data collection programmes are weak, the changes would result in significant system changes (not a major overhaul).	Data collection programmes are weak, with changes requiring an overhaul of the entire system.
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)	Manual and electronic techniques used for matching records are thoroughly documented.	Although incomplete, a high degree of documentation exists on manual and electronic record-matching techniques used.	Some documentation exists on manual and electronic record-matching techniques used.	Manual and electronic techniques used for matching records are not documented at all.

Table 5 Timeliness (SASQAF dimension 3)

Indicator number	Indicator description	Assessment Levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
3.1	Average time between the end of reference period and the date of the first results.	Preliminary results are released within the recommended timeframes as specified in the relevant standards and good practices.	Preliminary results released approach the relevant standards and good practices.	Preliminary results released lag behind relevant standards and good practices.	Preliminary results released lag far behind the relevant standards and good practices.
3.2	Average time between the end of reference period and the date of the final results.	Final results are released within the recommended timeframes as specified in the relevant standards and good practices.	Final results released approach the relevant standards and good practices.	Final results released lag behind relevant standards and good practices.	Final results released lag far behind the relevant standards and good practices.

## SC008, Data Quality Standard for Surveys

Indicator number	Indicator description	Assessment Levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
3.3	Production activities within the statistical value chain are within the planned timelines, viz.: data collection, data processing, data analysis, dissemination	All elements within the statistical value chain are within the planned timelines.	Some elements within the statistical value chain are within the planned timelines.	Few elements within the statistical value chain are within the planned timelines.	All elements within the statistical value chain are not within the planned timelines.
3.4	Report on the frequency of release.	The standards and guidelines for the frequency of release exist and are adhered to.	The standards and guidelines for the frequency of release exist, but only some are adhered to.	The standards and guidelines for the frequency of release exist, but are not adhered to.	No standards and guidelines exist for the frequency of release.
3.5	Punctuality of time schedule for publication.	Statistical outputs are released are always within the relevant standards and good practices, e.g. see GDDS and SDDS as a standard.	Statistical outputs released are most of the time within the relevant standards and good practices.	Statistical outputs lag behind the relevant standards and good practices.	Statistical outputs lag far behind the relevant standards and good practices.

Table 6 Accessibility (SASQAF dimension 4)

Indicator number	Indicator description	Assessment Levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
4.1	Are data and information available to the public?	All statistics disseminated are available from a publicly accessible medium.	Most of the statistics disseminated are available from a publicly accessible medium.	Few statistics disseminated are available from a publicly accessible medium.	Statistics disseminated are not available from a publicly accessible medium.
4.2	Rules governing the restricted availability of administrative records are well described and documented.	All rules governing the restricted availability of administrative are well described and documented.	Some of the rules are defined and documented	Some of the rules are defined and documented	No rules are defined or documented
4.3	Legal arrangements are in place to access administrative	Only those with whom legal arrangements are in place are able to access	Administrative records are made accessible to those with whom legal	Administrative records are made accessible to those without any legal	Administrative records are made accessible without any legal arrangements in

## SC008, Data Quality Standard for Surveys

Indicator number	Indicator description	Assessment Levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
	records via manual/automated/electronic systems	administrative data via manual/automated/electronic systems	arrangements are not officially in place, but are pending via manual/automated/electronic systems	arrangements in place, but discussions have been entered into.	place.
4.4	Types of media/channels used for sharing data amongst stakeholders are adequate and preserve 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.	No channels exist for stakeholders to access data.
4.5	Data is accessible in a format beyond the producing agency.	Data is accessible in a variety of formats that satisfies the requirements of all users.	Data is accessible in a variety of formats that satisfies the requirements of most users.	Data is accessible in a variety of formats that satisfies the requirements of some users.	Data is accessible in a format that only meets the needs of the producing agency.
4.6	Statistics are released on a pre-announced schedule.	Statistics are always released according to an advance release calendar.	Statistics are most of the time released according to an advance release calendar.	Statistics are sometimes released according to an advance release calendar.	There are no advance release calendars.
4.7	Statistics are made available to all users at the same time.	Statistics are always made available to all users at the same time.	Statistics are often made available to all users at the same time.	Statistics are seldom available to all users at the same time.	Statistics are never released simultaneously to all interested parties.
4.8	Statistics/administrative records not routinely disseminated are made available upon request.	Statistics not routinely disseminated are always available on request; or Administrative records not routinely shared are always available on request (where a legal framework is in place)	Statistics not routinely disseminated are usually available on request; or Administrative records not routinely shared are usually available on request (where a legal framework is in place).	Statistics not routinely disseminated are occasionally available on request; or Administrative records not routinely shared are occasionally available on request (where a legal framework is in place).	Statistics/administrative records not routinely disseminated are not available on request.
4.9	User support services are widely publicised.	User support services are well known and widely utilized.	User support services are well known and utilized by some users.	User support services are known but they are not used.	User support services do not exist.
4.10	Does a data dissemination policy exist, and is it maintained and accessible?	A data dissemination policy exists, and is available and up to date.	A data dissemination policy exists but is outdated.	A data dissemination policy is under development.	No data dissemination policy exists.
4.11	Does the pricing	Pricing policy	Pricing policy	Pricing policy is	Pricing policy

## SC008, Data Quality Standard for Surveys

Indicator number	Indicator description	Assessment Levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
	policy governing dissemination exist, and is it available to users?	exists, and is available and up to date.	exists but is outdated.	under development.	does not exist.
4.12	Catalogue systems (for survey, administrative records and other services) to identify information are available to users and are updated regularly.	Catalogue systems to identify information are available and updated regularly.	Catalogue systems to identify information are partially available and updated regularly.	Catalogue systems are not readily available and are not updated regularly.	Information is not catalogued.
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	Metadata are always documented, available, and readily accessible.	Metadata are available and accessible to some users	Metadata are available but not readily accessible.	Metadata is not documented

Table 7 Interpretability (SASQAF dimension 5)

Indicator number	Indicator description	Assessment Levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
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.	Concepts, definitions and classifications underlying the data are available, and any deviations from acceptable standards are annotated.	Some deviations from acceptable standards are annotated.	Few deviations from acceptable standards are annotated.	Deviations from acceptable standards are not annotated.

## SC008, Data Quality Standard for Surveys

Indicator number	Indicator description	Assessment Levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
		All concepts used in administrative records are well defined and documented.	The vast majority of the concepts, definitions and classifications used in administrative records are well defined and documented.	Some of the concepts, definitions and classifications used in administrative records are well defined and documented.	None of the concepts, definitions and classifications used in administrative records are defined or documented.
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.	Adequate documentation on scope, basis of recording, data sources, and statistical techniques used is available and deviations from accepted standards, guidelines or good practices are annotated. The accepted standard is the metadata template.	Partial documentation on scope, basis of recording, data sources, and statistical techniques used is available and deviations from accepted standards, guidelines or good practices are annotated.	Inadequate documentation on scope, basis of recording, data sources, and statistical techniques used is available, and deviations from accepted standards, guidelines or good practices are annotated.	Scope, basis of recording, data sources, and statistical techniques used are not documented.
5.3	All the statistical releases produced are accompanied by primary messages clarifying the key findings.	Primary messages clarifying all key findings on each statistical release are available in detail.	Primary messages clarifying some key findings on each statistical release are available in detail.	Primary messages clarifying a few key findings on each statistical release are available but not in detail.	No primary messages clarifying key findings on each statistical release.

Table 8 Coherence (SASQAF dimension 6)

Indicator number	Indicator description	Assessment Levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
6.1	Data within series and administrative systems are based on common frameworks, such as concepts, definitions, classifications, and methodologies, and departures	All data within series are based on common frameworks, concepts, definitions, classifications, and methodologies and departures from this are identified in the metadata.	Most of the data within series are based on common frameworks, concepts, definitions, classifications, and methodologies and departures from this are identified in the metadata.	Limited data within series are based on common frameworks, concepts, definitions, classifications, and methodologies and departures from this are identified in the metadata.	Data within series are not based on common frameworks, concepts, definitions, classifications, and methodologies.

## SC008, Data Quality Standard for Surveys

Indicator number	Indicator description	Assessment Levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
	from these are identified in the metadata.				
6.2	Statistics are consistent and reconcilable over time.	Statistics are always consistent and reconcilable over time.	Statistics are sometimes consistent and reconcilable over time.	Statistics are seldom consistent and reconcilable over time.	Statistics are neither consistent nor 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.	All data across comparable series, or primary source data are based on common frames, common identifiers, concepts, definitions, and classifications, and any differences are identified and can be allowed for in the interpretation.	Most data across comparable series, or primary source data are based on common frames, common identifiers, concepts, definitions, and classifications, and any differences are identified and can be allowed for in the interpretation.	Limited data across comparable series, or primary source data are based on common frames, common identifiers, concepts, definitions, and classifications, and any differences are identified and can be allowed for in the interpretation.	No data across comparable series or primary source data are based on common frames, common identifiers, concepts, definitions, and classifications.
6.4	Statistics are checked for consistency with those obtained through other data sources (identify comparable datasets and incomparable ones).	Statistics are always checked for consistency with those obtained through other data sources.	Statistics are sometimes checked for consistency with those obtained through other data sources.	Statistics are rarely checked for consistency with those obtained through other data sources.	Statistics are not checked for consistency with those obtained through other data sources.
6.5	A common set of identifiers (for the purpose of record matching) exist and have been agreed upon by the data producers.	A common set of identifiers (for the purpose of record matching) exist and have been agreed upon by the data producers.	Some identifiers exist, facilitating record matching, but have not been agreed upon.	Some identifiers exist, but is insufficient for accurate record matching	No common identifiers exist



## SC008, Data Quality Standard for Surveys

**Table 9 Methodological soundness (SASQAF dimension 7)**

Indicator number	Indicator description	Assessment Levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
7.1	Concepts, definitions, and classifications used follow accepted standards, guidelines or good practices (national, international, peer-agreed).	All concepts, definitions, and classifications follow accepted standards, guidelines or good practices (national, international, peer-agreed).	Most concepts, definitions, and classifications follow accepted standards, guidelines or good practices (national, international, peer-agreed).	Few concepts, definitions, and classifications follow accepted standards, guidelines or good practices (national, international, peer-agreed).	Concepts, definitions, and classifications do not follow any standards, guidelines or good practices (national, international, peer-agreed).
7.2	The scope of the study is consistent with accepted standards, guidelines or good practices.	The scope of the study is completely consistent with accepted standards, guidelines or good practices.	The scope of the study is partially consistent with accepted standards, guidelines or good practices.	The scope of the study is inadequately consistent with accepted standards, guidelines or good practices.	The scope of the study is inconsistent with accepted standards, guidelines or good practices.
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	Methodologies used in all processes always follow accepted standards, guidelines or good practices.	Methodologies used in all processes sometimes follow accepted standards, guidelines or good practices.	Methodologies used in all processes seldom follow accepted standards, guidelines or good practices.	Non-standard methods used.
7.4	Revisions schedule followed (explain the extent to which it is regular and transparent).	Revisions schedule is always followed.	Revisions schedule is sometimes followed.	Revisions schedule is seldom followed.	No revisions schedule.

## SC008, Data Quality Standard for Surveys

Indicator number	Indicator description	Assessment Levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
7.5	Preliminary and revised data are identified in the metadata.	Preliminary and revised data are always identified and explained in metadata.	Preliminary and revised data are sometimes identified and explained in metadata.	Preliminary and revised data are seldom identified and explained in metadata.	Preliminary and revised data are not identified and explained in metadata.
7.6	Studies of revisions and their findings are made public.	Studies of revisions and findings are always made public.	Studies of revisions and findings are sometimes made public.	Studies of revisions and findings are seldom made public.	Studies of revisions and findings are never made public.

Table 10 Integrity (SASQAF dimension 8)

Indicator number	Indicator description	Assessment levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
8.1	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.	The terms and conditions, including confidentiality, under which statistics are collected, processed and disseminated, are available to the public and completely follow the UN principles of official statistics.	The terms and conditions, including confidentiality, under which statistics are collected, processed and disseminated, are available to the public and by and large follow the UN principles of official statistics.	The terms and conditions, including confidentiality, under which statistics are collected, processed and disseminated are available to the public and to some extent follow the UN principles of official statistics.	The terms and conditions, including confidentiality, under which statistics are collected, processed and disseminated are not available to the public and/or do not follow the UN principles of official statistics.
8.2	Describe the conditions under which policy-makers, specifically government, may have access to data before release. Are the conditions published?	Policy-makers always get the statistics at the same time as everyone else and this is publicly stated.	Policy-makers in exceptional cases get the statistics before everyone else and this is publicly stated.	Policy-makers often get the statistics before everyone else and this is not publicly stated.	Policy-makers routinely get the statistics before everyone else and this is not publicly stated.
8.3	Advance notice is given of major changes in methodology, source data and statistical techniques.	Advance notice of major changes in methodology, source data and statistical techniques is always given.	Advance notice of major changes in methodology, source data and statistical techniques is sometimes given.	Advance notice of major changes in methodology, source data and statistical techniques is seldom given.	Advance notice of major changes in methodology, source data and statistical techniques is never given.

## SC008, Data Quality Standard for Surveys

Indicator number	Indicator description	Assessment levels			
		Quality Statistics	Acceptable Statistics	Questionable Statistics	Poor Statistics
		Level 4	Level 3	Level 2	Level 1
8.4	Ministerial commentary, when data are released, should be identified as such, and not be seen as part of the official statistics.	Ministerial commentary, when data are released, is always identified as such, and is not seen as part of the official statistics.	Ministerial commentary, when data are released, is sometimes confused to some extent with the official statistics.	Ministerial commentary, when data are released, is often confused with the official statistics.	There is no clear distinction between Ministerial commentary, when data are released, and official statistics.
8.5	Choice of source data, techniques and dissemination decisions are informed solely by statistical considerations (without political interference).	Source data, techniques and dissemination decisions are informed solely by statistical considerations without any political interference.	Source data, techniques and dissemination decisions are informed by statistical considerations as well as limited political interference.	Source data, techniques and dissemination decisions are informed by statistical considerations with political interference.	Source data, techniques and dissemination decisions are informed solely by political interference.
8.6	Ethical guidelines for staff behaviour are in place and are well known to the staff (professional code of conduct).	Ethical guidelines for staff behaviour are in place, are well known to the staff and are adhered to.	Ethical guidelines for staff behaviour are in place, are known to the staff and are adhered to.	Ethical guidelines for staff behaviour are in place, are not well known to the staff and to some extent are adhered to.	Staff does not know ethical guidelines for staff behaviour.