Quality Assurance for Record Databases Manual

Table of Contents

Introduction

The importance of record database quality lies in its being one of the most important sources for official statistics. Moreover, its use and proliferation and dependency represent a challenge for all statistical centers to assure their quality and place them within the general framework of the quality of statistical data, hence the need to focus on quality control in the production of statistical figures, where DSC issued in 2013 statistical data quality Charter for Dubai and developed it in 2016 to fit with the established quality frameworks and foundations worldwide.

DSC took it upon itself to find a systematic and clear framework for the quality of statistical data from different sources, as this is clearly evident in our system applicable in DSC, adopted from the global model Generic Statistical Business Process Management "GSBPM", so as to ensure the highest precision for statistical data in all its phases and to achieve the aspirations, the needs and desires of users and decision makers.

Given the importance of statistic numbers in building plans and developing programmers and policies in various areas of sustainable development, and the fact that DSC is the only authorized entity to produce, issue and publish statistical data in the Emirate from different sources, hence DSC has prepared this comprehensive guide to all record databases available in bodies, authorities and departments offering data to be a common frame of reference for the most important key principles for the quality of their record databases.

The fundamental principles of official statistics

In 1992, the United Nations Economic Commission affiliated with Europe adopted the fundamental principles of official statistics in the ECE region, and then in 1994 the United Nations Statistical Commission adopted these principles at the global level. They were also approved by the Economic and Social Council affiliated with the Commission for Basic Principles of Official Statistics in 2013. In January of 2014, it was adopted by the General Assembly. This recognition at the highest political level confirms that official statistics are reliable and objective information and critical for making decisions.

The first principle: proportionality, impartiality and equal access to official statistics

"Official statistics are an element that cannot be dispensed with in the information system of any democratic society, through the data it provides to the Government, the economy and the public on the economic, demographic, social and environmental situation. In order to achieve that, the official statistics authorities are required to gather statistical data that prove practically useful and make it available to citizens impartially in fulfillment of their right to seek and verify information."

The second principle: Professional standards and ethics

"To maintain confidence in official statistics, the statistical agencies are charged according to precise technical criteria including scientific principles and professional ethics; identify methods and procedures for statistical data collection, processing, storage and presentation."

The third principle: Accountability and transparency

"To facilitate the correct interpretation of the data, the statistical agencies display - according to scientific standards - information concerning statistical sources, methods and procedures that apply."

The fourth principle: Preventing improper use of official statistics

"The statistical agencies may comment on any erroneous interpretation and misuse of statistics."

The fifth Principle: Sources of official statistics

"Data for statistical purposes may be obtained from any source, whether from statistical surveys or administrative records. When statistical agencies choose sources, they should take into account the elements of quality, timeliness, costs and the burden on respondents. "

The sixth principle: confidentiality (Privacy)

"Data of individuals collected by statistical agencies for statistical compilation, whether concerning natural or legal persons, should be kept secret and must necessarily be used for statistical purposes."

The seventh principle: Legislation

"Make laws, procedures and measures public according to which the statistical systems work

The eighth principle: National coordination

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

The ninth principle: Use of international standards

"To enhance statistical agencies in each country to use international concepts, classifications and methods, which creates harmony and efficiency of statistical systems at all official levels."

The tenth principle: International cooperation

"Bilateral and multilateral cooperation in statistics contribute to the improvement of systems of official statistics in all countries."

Record databases, definition and concept

They are those data owned by an official entity, containing information collected for the purpose of applying regulatory, administrative and legal procedures.

There are many types of record databases, including:

Civil records: Civil records in developed countries that possess these records are considered a major source of data on individuals. These records are also known to be complete. The laws associated with these records allow obtaining detailed data including demographic, social and economic data. In addition, the data is updated continuously by the individuals in those States, where they provide the responsible civil registry with any changes regarding their address and their economic and professional activities ...etc. The U.A.E. owns a registry of civil status supervised by the Emirates Identity Authority-affiliated with the Ministry of Interior. This registry records all Emirati and non-Emirati individuals who have residence permit or who wish to stay in the State. It is linked to many relevant government agencies such as the General Directorate of Residency and Foreigners Affairs and vital events registration centers such as births and deaths. Civil registry suffers from some disadvantages that affect its accuracy and completeness, especially such defects of repetition and deletion, inclusion of the deceased in records of individuals in some cases. Efforts are being made to update this record and revise it currently. In addition, data types collected through it are being developed, by adding new variables related to home address, economic and educational aspects... etc., or in terms of activating binding laws concerning giving this information.

Administrative records: Administrative records possessed by most ministries and governmental and non-governmental institutions are major sources of statistical data in various sectors and areas that are used to calculate the indicators for these data types. DSC coordinates with all local and federal governmental and non-governmental organizations with regard to obtaining statistical data published by the Center in one of its most important publications; namely, the Annual Statistical Yearbook. Statistical data that DSC obtains from administrative registries of government institutions cover many aspects. DSC concerned departments coordinate with various governmental institutions to provide data of those institutions, where DSC connects and communicates with institutions using all available means of communication such as telephone, fax, e-mail and follow up with them online through Dubai Statistics Smart System through smart electronic window "Statistical Sources System" for obtaining all statistical data or to inquire about some data. After obtaining the data from different sources, the concerned authority approve, process and prepare it in a manner appropriate to the statistical approach and related standards, as data are validated, coded, numbered and the output designed on Dubai Smart Statistical System through "Interactive Statistics System" and producing them in the form of statistical classifications including the available variables in a way consistent with the criteria used in DSC via the website or through interactive graphics and diagrams through the system of statistical indicators "idashboard". The cycle of obtaining data from Government sources is different, as DSC is provided on a monthly basis with some data, such as arrivals and departures, marriage and divorce, while other data as the education and health data is obtained on a yearly basis.

Special administrative records: Private organizations contribute to supplying the statistical data obtained by DSC, where these organizations provide DSC concerned departments with some statistical data concerning certain statistical topics such as associations of public benefit and related information falls under this data information on sports clubs in Dubai.

General guidelines for record databases

Record databases of various kinds in Dubai must abide by some general guidelines or key features, as follows:

1. **Universal coverage:** record databases should include all the relevant facts, events or actions in the Emirate of Dubai and in every demographic group in a specific time period.

2. **Continuity:** the principle of continuity is critical in collecting and compiling data records, where data is needed to clarify short-term fluctuations, including seasonal movements, as well as long-term movements. Continuity is most easily achieved when the record database is fully established, where periodical reporting; monthly, quarterly or yearly becomes a recurring activity in the system. When using the supplementary additions to the record database by sample surveys, and in order to obtain estimates of some modifiers, special efforts may be needed to ensure the availability of data on a frequent and regular basis.

3. **Confidentiality:** Personal information in partial data to the record database and any associated statistical reports must be protected and consistent with the intended use of those documents and specific statistics for administrative purposes. Also, procedural facts-based statistical reports, documentaries, whether derived from the system registry or obtained by any other means, such as sample surveys, must be utilized in the widest possible use to suit certain methods concerning providing confidentiality to individuals whose data contribute to these statistics.

   It must be noted that law number 28 for 2015 concerning DSC confirms in article 9 confidentiality of statistical data as follows: "Personal data and information obtained as a result of any count, census or survey are considered confidential and may not be communicated or reviewed by any entity or individual except through DSC, as per a written consent from the person concerned or at the request of the competent judicial authorities, nor may they be published for non-statistical purposes. "

4. **Use:** the principle of confidentiality is based on the individual right to expect not to use the information given in confidence to the registrar or interviewer except only for permissible statistical or administrative purposes. The authority entitled to collect procedural facts-based statistical data and expected to maintain their confidentiality are required to perform their role. Data communicated by individuals is expected to be complete and accurate, regardless of the sensitivity of the information.

5. **Regular publication**: the aim of collecting data should be as follows:

- Provide summarized statistics for total procedural facts-based statistical data periodically, every month or every quarter, within the timeline that provides information for many purposes fast enough, the most important may be health interventions and programs to estimate the population and for administrative use or for any other needs.

- Produce detailed annual classifications for each type of procedural facts-based statistical data comprehensively classified by demographic, social and economic characteristics.

- Using them as a major/supporting/complementary input to issue official statistics.

6. **The availability of resources:** when planning the detailed classification program, it is important to ensure the availability of resources to complete it on a consistent basis regularly according to a timetable to ensure using analysis of the interrelationship between demographic, social and economic factors in the planning, operation and evaluation of all statistical programmer input effectively for the preparation and evaluation of social and economic plans. Moreover, as far as possible, those statistics should be comparable within the Emirate of Dubai and in various data sources at the national level to allow for a proper analysis.

For the purposes of using record databases statistically, the following conditions must be met:

1. The record database contains "descriptive data" about all variables contained in the record database.

2. Levels of confidentiality and security of data must be specified with the record database provider to avoid any conflict of interest.

3. The reference period of the record database must be consistent with the reference period of the statistical output.

4. The record database must cover its relevant specific community.

5. The record database must be free from redundancy or deficiency.

6. The missing values for record database variables as statistically at their minimum must be acceptable and can be compensated without a radical change to the consistency and integrity of data.

7. All record database variables must be classified according to specific categories set by the entity, with a recommendation that it be according to international standard classifications and manuals.

8. Maintain periodicity, consistency, structuralism and definitions of the key variables contained in the record database.

Record database quality management procedures

Where the main procedures for managing the quality of record databases are present in three major axes between DSC "data producer" and between all entities, departments and authorities "data provider", are as follows:

1. **Networking**: it may include the following:
   - Description of data collection procedures.
   - Regular dialogue with the data providers.
   - Document the quality guidelines for each set of record databases.
   - Describe the effect of record databases on official statistics in terms of the effect of errors and bias.

2. **Management**: it includes the following:
   - Participatory relationship building with partners, data providers.
   - Preparation of a guide for all statistical requirements for record databases.
   - Verification tools associated with data quality.

3. **Verification**: and includes the following:
   - Apply statistical quality assurance standards of data providers.
   - Provide the audit results of record databases.
   - Define drawbacks and where there is lack of clarity and impartiality in record databases.
   - Determine the impact of distortions in record databases on statistical targets and performance indicators.

Document Classification: General
تصنيف المستند: خارجي
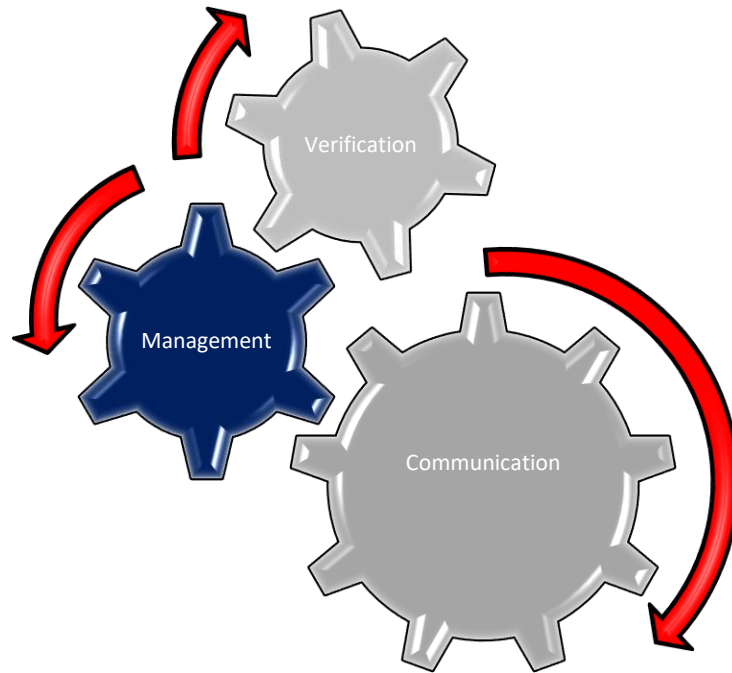
Statistical quality dimensions for record databases

Eight dimensions of statistical quality of record databases in Dubai have been identified, so as to ensure providing the institutional environment governing record database quality assurance, as well as to promote the importance of quality assurance in all data providers alike. These dimensions are:

1. Institutional Environment

Law number 28 for 2015 concerning DSC is the basis for management of the statistical system in Dubai, where it guarantees moral and legal personality to manage statistical work by DSC in Dubai. Quality and environment policy which we in DSC systematically follow, ensure the quality of our statistical services to meet the needs of partners and customers to ensure the preservation of the environment, by working according to the standards of institutional excellence and quality and environmental management systems ISO 90012015 and ISO 14001:2015 and system of security and confidentiality of information and customer service management system 10004:2012 and customer complaints management system 10002:2014 and health and occupational safety specifications 18001:2007.

DSC also gives great importance to statistical coordination at various levels, where it formed the Statistical Coordination Committee in Dubai that has many local and official actors that provide data in Dubai. DSC also adopted our system according to the **Generic Statistical Business Process Model "GSBPM"** for governance of all projects, processes and statistical outputs from different sources.

2. Relevance

The importance of statistical output is connected to the rate of demand for the data in any area related to reality and society. Meeting those statistical outputs resulting from the record database of users' needs and the amount of interest in them is directly reflected on how important those outputs for them are.

3. Ensure Data Confidentiality & Security

Confidentiality and individual data protection must be provided and secured by all data providers by including in the rules and legislation of all data providers the related items and materials. In addition, partnership agreements and memoranda of understanding between DSC and the various data providers must emphasize the confidentiality and protection of

data exchanged between the two parties. Moreover, Article 9 in law number 28 for 2015 concerning DSC came to emphasize the confidentiality of statistical data, as it says "personal data and information obtained as a result of any count, census or survey is considered confidential and may not be communicated or reviewed by anyone or any entity except via DSC, against a written permission from the person concerned or at the request of the competent judicial authorities, nor may they be published for non-statistical purposes." This is in addition to the information security policy that the Centre has prepared and adopted and its commitment to provide customers, stakeholders, partners and employees with a safe environment for information processing.

4. Accuracy

It means how accurate record databases in the field or the subject are, the objective of collecting, measuring and clarifying their characteristics, to avoid errors in assessing them, which is usually biased due to sampling errors, generalization, representation and others. Therefore, the statistical output is of great precision and high reliability.

5. Consistency

It is the process of data consistency internally between various record database variables and their comparability, whether over time, with the other emirates, or other data from different sources.

6. Timeliness

Record databases lose their credibility and relevance if they exceed the assigned date to be adopted in measuring a particular area; good timing is the ability to publish statistical data and information within the specified period.

7. Accessibility & Availability

It is the process of publishing record databases in a clear and understandable way, in a way that is appropriate and convenient, and making them accessible. That is why the ability of users to easily access record databases is an important factor in determining the quality of such data.

8. Metadata

Metadata describe the data in general to give the user information about the used method, reference, chronology and the legal reference that the organization that produces the statistical work uses and the most important information on concepts, measurement basis and method of data availability and periodicity and providing all definitions of the variables contained in the record database.

Evaluation and quality assurance for record databases

Evaluating quality for record databases is contained in the statistical data quality Charter of the Emirate of Dubai and the practices adopted and recognized in the United Nations specialized agencies in the field. Where all statistical quality assurance levels of record databases have been defined through the adoption of practice field's model concerning quality assurance of record databases that are used as one of the most important sources of official statistics. In addition, a subjective evaluation of the record database has been prepared, including all data quality dimensions, which have been previously addressed, with a view to ensuring quality in all steps involved in each stage of the processes performed by the source.

Practice field forms for record database quality assurance

Areas of practice for quality assurance of record databases which are used to produce official statistics

Verification procedures and documentation by the data producers "	Principles, criteria and verification tools used by the record database providers"	Communicating with partners' record database providers"	Operational context and data collection	Level of quality assurance/guarantee
Quality assurance activities: Not taken into account, or not verified				Level 1: no guarantee of quality
There is no description of the methods of verification and documentation of quality assurance	No description of the procedures and criteria of the data producer concerning quality assurance	No communication	The data provider does not check or examine the operational context for data collection	The Dubai Statistics Center (producer of official statistics) did not verify, audit and apply quality assurance procedures for record databases at all

Considering the verification of the following quality assurance activities:				Level 2: the fundamental guarantee of quality
The Dubai Statistics Center (producer of official statistics) established regular verification and examination checks on data sent by data providers. In addition, a description of the quality assurance of DSC concerning record databases has been published, as well as the definition of the general approach to quality and the findings achieved. It also identified strengths and limitations of the record databases and explained the degree of potential risk to data quality resulting from using record databases.	The data provider notifies DSC of the statistics examination procedures and provide a summary of these procedures and the periodic review of the record databases. Such as internal audit or external audit and describe the implications of these revisions on the data and the requirements of official statistics.	Communicating with data providers and determine the arrangements for the provision of data, including identifying cyclical needs, timing and coordination of data supply, and method of dispatch and communication about previous errors in data and consultation about the data users' views and quality improvement areas.	The Dubai Statistics Centre uses a brief presentation of the record database collection process, and it explained the main stages outlined in the context of GSBPM operations. It defined actions taken to reduce risks that affect the quality of the data and summarized the implications of accuracy and quality of data. DSC explained any changes in data collection or the context of the processes associated with them.	The Dubai Statistics Center (producer of official statistics) reviewed and published a summary about record databases quality assurance arrangements.

				Level 3: enhanced quality assurance
The Dubai Statistics Center (producer of official statistics) provides a thorough description of the main principles of quality assurance, quality indicators and examination methods of data used by it in record databases and details of the general approach to the measurement of quality indicators. It identifies strengths, limitations and restrictions concerning record databases and explains the degree of potential risks on quality posed by record databases.	The Dubai Statistics Centre (producer of official statistics) provides a description of the main principles of the quality assurance, quality indicators, examining data used by data providers and describing the role of scrutinizing of record databases during the collection process and practical preparation of record databases. Describe the implications of official statistics as regards quality issues identified by record database providers.	DSC (producer of official statistics) agreed to and documented the following: * record database requirements for statistical purposes. * Define the legal basis for data supply process by providers. * Define the procedures concerning data transfer process. * Define the data protection arrangements. *Define the data transfer methods by providers. • Create an effective way to communicate with contacts (for example with data collector of providers, IT systems and operational policy owners) and discussion of the ongoing statistical requirements in the system of data collection and quality of data provided. • Measuring needs, opinions, and satisfaction of users of statistics and resolving any issues regarding the quality of the record databases.	The Dubai Statistics Centre (producer of official statistics) provided users with a thorough description of the practical context and record database collection, arrangements and procedures, for example: process chart detailing the processes of data collection, explanations of the categories, selection and summarization of the potential sources of bias and error in the logs and databases, defining and describing the safeguards taken to reduce risks to the quality of the data. It also gave a detailed description of the effects resulting from the accuracy and quality of the data, including the effect of any changes in the arrangements for collecting record databases.	The Dubai Statistics Center (producer of official statistics) evaluated record database quality assurance procedures and published a thorough description of them.

Considering the verification of the following quality assurance activities:				Level 4: comprehensive quality guarantee
The Dubai Statistics Center (producer of official statistics) provides full explanation of its procedures to verify record database quality assurance (including validity, logic, coherence and consistency) and indicators to measure record database quality and made comparisons with data from other sources. It also defines the effects that may arise from the distortion in record databases on the statistical performance indicators. It identified strengths and opportunities for improvement in record databases, limitations and impediments relating to using them in official statistics.	The Dubai Statistics Center (producer of official statistics) describes the data quality criteria used by data providers, indicators and tools for verification and review data from the providers, such as: verification of data accuracy, completeness and coverage. DSC also identified and documented verification reports and audited record databases and described their uses to issue official statistics. It clarified its sustainability, adequacy and conformity to specifications and for the production of official statistics.	The Dubai Statistics Center (producer of official statistics) built and maintained a solid relationship with record database providers through signing an MOU outlining roles and responsibilities and clarifying the framework and legal reference for providing data. It also defined the style and method of providing data and defined data protection and security conditions. It explained the various revision levels for data, content and specifications, in addition to holding and arranging regular meetings between data personnel of the two parties to follow up changes in record databases and hold meetings to follow up the issues concerning quality.	The Dubai Statistics Center (producer of official statistics) provided data users with a detailed description of record database system and the context of their operations (explaining why data are gathered, by whom and how, and determine the differences between the data collected and the data saved and define the concepts of data terminology, lost data and redeemed data, if any. It also defined the most prominent performance indicators and their targets, as well as potential error sources in record databases and quality assurance procedures used to minimize risks that could affect record databases and their quality.	The Dubai Statistics Center (producer of official statistics) verified record database quality assurance arrangements and determined the results of the independent audit process and published detailed documentation about quality assurance and audits.

Self-assessment form for record databases

Self-assessment for record databases	
Introductory data	
	Project name:
	Name of Entity:
	Entity's contact person:
	DSC contact person:
Attach "metadata card"	**Purpose and use:**

0 I don't know	0 No	1 Yes	1. Institutional Environment
			1.1 Is there a law that supports the source collection procedures, saving and owning its record databases?
			1.2 Is there an agreement or memorandum of understanding between the source and DSC to exchange data and information amongst themselves?
			1.3 Is there an organizational unit in the source concerned with statistics for coordination regarding joint statistical work?
			1.4 Does the source have a policy concerning security and confidentiality of information?
0 I don't know	0 No	1 Yes	2. Relevance
			2.1 Can the availability of data from this source be useful as a substitute for survey data collection?
			2.2 Can the availability of data from this source be useful for supplementary checks the accuracy of official statistics?
			2.3 Are source data important for official statistics?
			2.4 Is there a high effect of using this source data on the respondent's burden?
0 I don't know	0 No	1 Yes	3. Ensure Data Confidentiality & Security
			3.1 Does the source data include a disclaimer on the protection and confidentiality of data? Is it applied?
			3.2 Does this disclaimer include data protection law? Is it applied?

0 I don't know	0 No	1 Yes	3.3 Is the process of transferring data from source to DSC safe?
			3.4 Are there special and specific requirements defined by the source for the data transfer process?
			3.5 Does gaining access to data necessitates that DSC buys equipment or special software to transfer data?
0 I don't know	**0 No**	**1 Yes**	**4. Accuracy**
			4.1 Is DSC informed of the data collection method and how to save it?
			4.2 Is DSC notified of any changes to the data collection and saving process when they occur?
			4.3 Is DSC notified of plans to change the databases and their definitions?
			4.4 Does the source allow DSC to ask questions and inquiries about checking data accuracy?
			4.5 Are there a documented verification procedures supported by the source to ensure accuracy?
0 I don't know	**0 No**	**1 Yes**	**5. Consistency**
			5.1 Are definitions in the source consistent with the definitions used by DSC?
			5.2 Are existing variables checked for consistency against the previous database variables?
			5.3 Is the time assignment schedule of the source databases consistent with the time assignment schedule for the purpose it is used for in DSC?
			5.4 Is the consistency of the source databases being verified across the length of time available?

0 I don't know	0 No	1 Yes	Timeliness
			6.1 Are the data delivery dates documented with the source?
			6.2 Are the data periodicals regular, "monthly, quarterly, yearly?
			6.3 Is data delivery always on time?
			6.4 Is DSC advised early enough of delaying the delivery of data?
0 I don't know	**0 No**	**1 Yes**	**7. Accessbility & Availability**
			7.1 Does the source have any reservations concerning making the databases available to users?
			7.2 Does the source publish its data to users?
			7.3 Does the source supply data to users against payment?
			7.4 Does the source provide interactive information window for users to interact with databases?
0 I don't know	**0 No**	**1 Yes**	**8. Metadata**
			8.1 Does the source provide structural metadata about database variables?
			8.2 Does the source provide referential metadata that describes a database?
			8.3 Does the source provide metadata about its target category or group?
			8.4 Are categories used by the source sufficiently clear and specific?
			8.5 Does the source have a specific system for metadata?

☐ GSS Risk Assessment Tool v1.0, Office for National Statistics- ONS (UK), 2014

☐ Data Quality Assessment Tool for Administrative Data, U.S. Census Bureau, 2013

☐ Guidelines for Measuring Statistical Output Quality v4.1, Office for National Statistics- ONS (UK), 2013

☐ Quality Assurance and Audit Arrangements for Administrative Data, UK Statistics Authority, 2013

☐ Quality, Methods and Harmonization Tool, Office for National Statistics- ONS (UK), 2011

☐ Checklist for the Quality evaluation of Administrative Data Sources, Central Bureau for Statistics (CBS), *Statistics Netherlands, 2009*