

Quality Data Sheets

Thorsten Tümmeler

Federal Statistical Office (Destatis), Wiesbaden, Germany; thorsten.tuemmler@destatis.de

Abstract

Measuring quality is one of the major challenges of quality management in statistics. The European standard structures for quality and metadata reporting increasingly require quantitative quality indicators to underpin qualitative assessment of quality of statistics provided in quality reports.

In a heavily decentralized national statistical system where most statistics are produced in collaboration by 15 independent statistical institutes it is difficult to acquire information from the statistical production process that is essential in order to calculate quality indicators like item response rates, imputation rates and other process indicators. In Germany, the field work is generally carried out by 14 independent state statistical offices that transmit aggregated results on state level to the Federal Statistical Office. The Federal Statistical Office provides methodology, publishes statistical results for Germany as well as relevant metadata like quality reports. So the statistical production process is segmented by two dimensions, geography (each state statistical office is responsible for the field work in its state) and process step (field work by the state statistical offices, methodology, publishing and quality reporting by the Federal Statistical Office).

Attaining the data that is necessary to generate quality indicators and other information on the statistical production process cannot be solved by organizational means alone. In order to keep the workload for producing such data at a minimum level, IT-systems need to fully support compilation, analysis and management of these data.

This paper describes the general idea and the goals of Quality Data Sheets as well as projected costs and benefits regarding the implementation. Finally, the vision for the extension of Quality Data Sheets for purposes of process and resources management will be illustrated.

Keywords: Implementing of quality reporting, quality indicators, decentralized statistical system.

1. Introduction: Why do we need Quality Data Sheets?

Federal statistics in Germany are basically produced conjointly by the 14 state statistical offices of the Länder (in the following: SSOs) and the Federal Statistical Office (in the following: Destatis). This working association is referred to as “the system of statistical offices” of the Federation and the states and the term regional decentralization describes the underlying principle. Accordingly, the SSOs are generally responsible for conducting statistical surveys prescribed by law. They are administratively and financially independent of the Federation and not subject to directions from Destatis or the federal ministries.

Usually the SSOs are responsible for collecting and processing data for their state. Therefore, the core processes of statistics production are performed at the SSOs for about two thirds of the federal statistics (Destatis produces the remaining third of federal statistics by itself). Among other things Destatis is in charge of compiling and disseminating federal results. Compiling federal results specifically includes the validation of the provided results and the aggregation to results on the federal level. For each statistical domain an expert meeting chaired by Destatis has the tasks of coordination and further development of the statistical production processes for the respective statistics. In this respect the situation in Germany is quite similar to the situation of the European Statistical System.

Destatis is also responsible for quality reporting to users and to Eurostat. However, obtaining all the information required to compute even basic quality indicators like the rates of over-coverage, unit non-response, item non-response and imputation as well as the proportion of common units between administrative data and survey data from 15 different independent statistical institutions is difficult. In practice, these indicators often cannot be computed because Destatis has no access to the information that is necessary to do so. Consequently, German quality reports often do not contain these quality indicators. And even more importantly, lack of this information seriously compromises systematic quality management.

Being aware of this situation the national Working Group “Quality of the Statistical Processes and Products” developed an instrument called Quality Data Sheets (QDS), to obtain the required information during the statistical production process directly from the tools that are used in the production process.

2. Objectives of Quality Data Sheets

The QDS are an instrument of quality management for all statistics working with micro data. They are used to collect comprehensive quality- and process-related information with each collection cycle of a survey for each statistical office. They refine this information and provide facts on process and data quality for each survey and each individual SSO as well as a summary on national level. QDS thus enable documentation, monitoring and controlling of important characteristics of the collection and processing phases of statistics production.

The QDS serve the following main objectives:

- The QDS provide the information necessary for the national quality reporting and for the quality reporting to Eurostat for all but especially for decentralized statistics. They enable the calculation of five quality indicators for accuracy, namely the rates of over-coverage, unit non-response, item non-response and imputation as well as the proportion of common units between administrative data and survey data.
- In addition to the quality indicators for accuracy the QDS provide information on the data collection and data validation processes that support proper interpretation of the indicators. This information enables fact-based discussion of the quality of the statistical production processes as well as the quality of the products in the respective expert meetings.
- The QDS provide important information about the data collection and data validation processes as well as on the accuracy of the results broken down by SSOs to the respective expert meeting. Thus, QDS enable the comparison of these indicators between SSOs and allow benchmarking of the underlying production processes. This transparency is a huge improvement since SSOs used to be very guarded about revealing any information on their internal processes which they deem to be of no concern to Destatis due to the principle of subsidiarity. That way QDS are expected to contribute to the identification and spread of good practices in the statistical offices.
- It will be possible to generate QDS at any point during the statistical production process. This allows monitoring of current process quality and current accuracy of the statistical results and thus enables better steering of quality. It also facilitates resource

management since staff can be shifted between different statistics depending on work progress that has or has not been achieved.

- The strategic goals of Destatis demand to ensure the high validity of data during the entire data production process (strategic goal Q2), the European Statistics Code of Practice and the Quality Assurance Framework of the European Statistical System require procedures to monitor the quality at different stages of the statistical production process (method 4.2.2). QDS are an instrument that will help to meet these requirements.

3. Implementation of Quality Data Sheets

3.1. Generic approach

Quality indicators on accuracy as well as some basic indicators on process quality are relevant for all types of statistics that work with micro data. The Working Group “Quality of Statistical Processes and Products” decided to adopt a generic approach in order to better coordinate and harmonize the computation of this information. Key of this approach is a generic list of items that a typical QDS should comprise:

- Over-coverage broken down by causes: Number of units that were used for the collection, but for which was noted during the survey that they do not belong to the target population. Typical causes for over-coverage in business statistics are an outdated NACE code or the unit ceased to exist before the survey period. This indicator is a measure for the quality of the sampling frame.
- Number of units that changed the sampling stratum for sample surveys: This is an additional measure for the quality of the sampling frame and it indicates possible consequences on accuracy.
- Number of responses received after each wave of reminders: This indicator aims at measuring the effectiveness of the reminder mechanism.

- Number of respondents by data collection tool used: Examples for tools are paper questionnaire, online questionnaire or a web-based application that retrieves statistical data directly from the information systems of businesses (e.g. accounting software).
- Number of unit non-responses: This indicator aims at measuring the possible impact on accuracy of statistical results due to reduced sampling precision and bias.
- Item non-response: For key variables this is an indicator for a possible bias caused by non-response.
- Treatment of non-response: In order to better assess the impact of non-response, it is useful to know, how non-response was handled in the editing process. To do this, four cases should be distinguished:
 - Number of cases, where non-response was remedied by contacting and retrieving the required information from respondents.
 - Number of cases, where manual imputations were made.
 - Number of cases, where automatic imputations were made.
 - Number of cases, where missing data has not been substituted.
- Number of data sets containing errors and number of data sets containing warnings: This indicates the quality of the raw data provided by respondents. It also serves as an indicator for the documentation of time and effort invested.
- Number of errors and warnings: Since a data set can contain more than one error or warning, the quality of the raw data as well as the effort necessary to correct errors can also be indicated by this measure.
- Treatment of errors and warnings: As with the treatment of non-response the treatment of errors and warnings potentially influences accuracy of statistical results. In order to better gauge this influence and potential risks of model-inherent bias and over-editing as well as document time and effort invested, four cases of treatment should be distinguished:
 - Number of cases, where an error was corrected or a warning was checked by contacting the respondent and clarifying the issue by gaining additional information from him.

- Number of cases, where errors or warnings were corrected by manually imputing a value.
- Number of cases, where errors or warnings were corrected by an automated imputation procedure.
- Number of cases where warnings did not result in modification of the original value either because the original value proved correct or because the warning has been ignored.

Starting from this generic list, statisticians in the subject-matter departments should develop specific QDS that fulfill their individual information requirements by adding new items, by modifying items or by dropping items that are not relevant. These specific lists for statistics then need to be agreed in the respective expert meeting with the SSOs. The expert meetings then have to provide detailed specifications for the updates that are necessary in the respective processing tools in order to collect the required information and they have to trigger actual programming of the updates.

3.2. Technical implementation

All of the items in the generic list emerge at some point during the statistical production process and they change from one survey run to another. In order to document the quality of the data that are published, QDS need to be generated for every survey run.

Since additional manual work for the generation of QDS is beyond all questions and thus to be avoided, IT-systems need to be able to adequately register, store and utilize the required information.

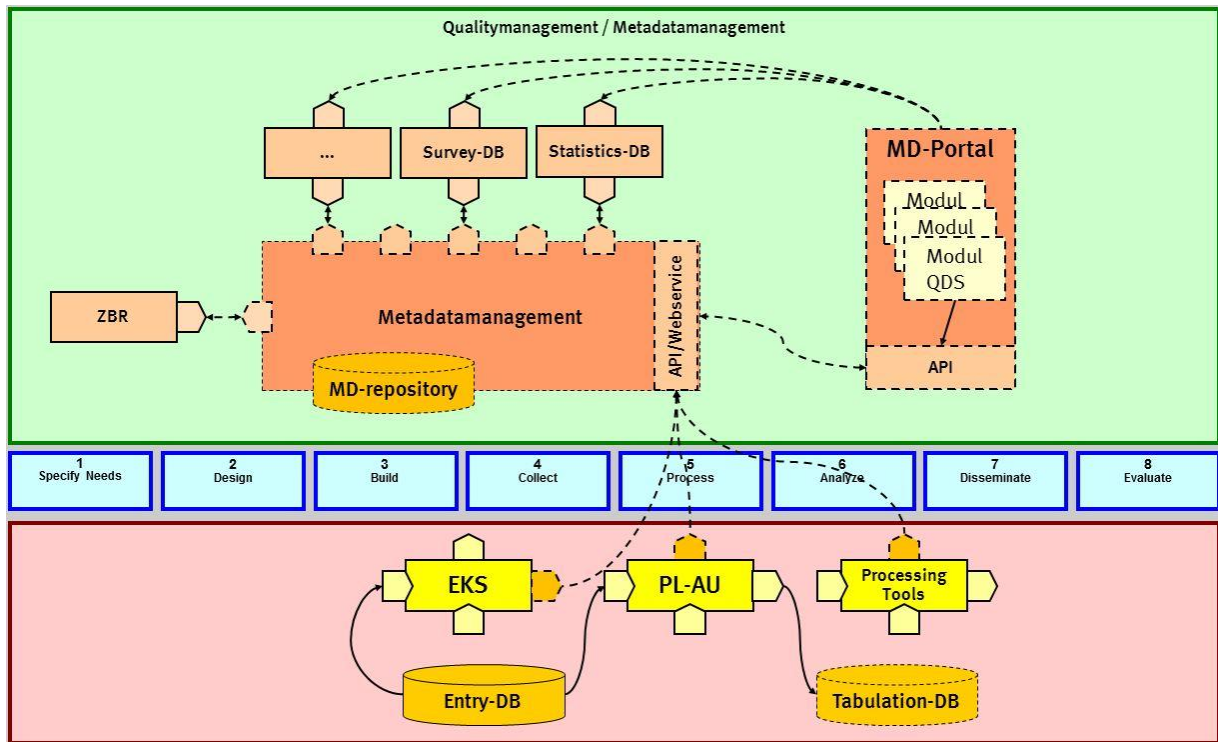


Fig. 1: Technical implementation

Naturally, all standardized IT-tools available for the statistical production process need to be updated so that they support the generation of QDS. The most important ones are input control systems (EKS) and the data editing executive environment (PL-AU) which cover the core of the production process. Unfortunately, statistical production in most cases is still based on specific processing tools rather than standardized tools. That means that many different tools need to be updated, thus significantly increasing the required investment.

The idea is that in all those IT-tools the survey managers of the statistical offices of the states can “press a button” to generate the data that is required for QDS at least at the end of each survey run. The IT-tools pass these data on to a metadata management system (MDMS, green area in the illustration above) using a standardized webservice interface. The MDMS receives the data deliveries from different IT-tools, augments the data using metadata from other systems like the survey-database and the statistics-database and generates a QDS for the respective statistic. Once every survey manager for each of the 14 SSOs has initiated the generation of QDS, the MDMS automatically computes a summary for Germany. This

summary then provides the quality indicators required for quality reporting on national level and to Eurostat. Using the metadata portal the survey managers can look at the data of their state and compare them to that of the other states as well as to the summary for Germany. From the metadata portal survey managers can also export the data for further analysis required for discussion on quality of the production processes as well as accuracy of the statistical product in the respective expert meetings.

3.3. Implementation strategy

Currently IT-tools do not support the generation of all the data required for QDS, so basically, every standardized IT-tool and every specific processing tool used in the statistical production process needs to be updated. The necessary investment for implementing QDS for all 330 statistics and in all SSOs and Destatis is estimated at about 1 000 programming months. As this is too much to realize in the short term, an incremental approach was adopted. In a first step, 30 statistics that are expected to need the information, that the QDS provide, the most, e.g. because of European requirements, ought to implement QDS immediately. That means the respective expert meetings need to be briefed, they have to develop specific QDS that contain the information that is needed for their purposes, the expert meetings have to specify the update required in the processing tools and to commission the actual programming. It is expected that QDS will be implemented for most of those 30 statistics by the end of 2017.

For all other statistics the QDS should be implemented at a convenient opportunity, i.e. when processing software is to be updated for other reasons.

In 2020 an evaluation will take place in order to determine the general implementation status. At that point all statistics that have not yet implemented QDS will have to present a working plan for the implementation.

4. Outlook

QDS are a very promising and important instrument that will support and improve quality assurance by measuring indicators for processing quality as well as accuracy of results. It will be a key instrument in helping statisticians to improve overall quality of statistics by making transparent existing differences in processing quality and accuracy of results for the different SSOs for the same statistic. These differences will have to be scrutinized by the respective expert meetings, their effect on statistical results will have to be discussed and measures will have to be taken if necessary.

From the experience of briefing 18 expert meetings, however, statisticians tend to regard QDS as a chore they have to fulfill because of a decision of the heads of the statistical offices. As with quality management in general, it is difficult to convincingly demonstrate the benefit for statisticians themselves. Especially statisticians from the SSOs are wary because for them increased transparency may mean that they have to explain and justify or even change their practices, which so far have been obscure for the other SSOs and especially Destatis. As a consequence the individual expert meetings tend to shorten the generic list of items in order to save on costs for updating their software or to avoid too much transparency.

However, once QDS are implemented for a range of statistics the expectation is that statisticians will experience the benefit of QDS and start using QDS to its full potential, i.e. as a powerful instrument supporting quality management, process management and even resource management. A key feature in this respect is that QDS can be generated at any point in the statistical production process and as frequently as needed. This allows for e.g. daily monitoring of working progress as well as of the quality level that is realized. For a survey manager that is responsible for multiple surveys this provides an overview on work progress as well as quality levels reached and enables a better control of resources deployment. Prerequisite is an adequate specific QDS that covers exactly the information that a survey manager needs in order to better manage the surveys in his area of responsibility.