Quality Assessment of Statistical Processes and Products at SORS

Seljak, Rudi
Statistical Office of the Republic of Slovenia
Vožarski pot 12
1000 Ljubljana, Slovenia
Rudi.seljak@gov.si

Introduction

Quality assessment and quality management activities are constantly declared as one of the fields with high priority in the strategic documents of the Statistical Office of the Republic of Slovenia (hereinafter SORS). Since 2003, when the more systematic approach in this area was introduced, a lot of effort has been put into this work and also a lot of concrete results have been achieved.

SORS is the national statistical authority but at the same time also part of the European Statistical System (hereinafter ESS), where a lot of work in the direction of common quality framework definition has also been done. Therefore, when defining our quality assessment and quality management concepts and principles, two main principles were followed:

- To take into account the work done inside the ESS and try to take into account as much as possible most of the results of this work in our development activities.
- To take into account the specific features of the statistical process in our office, the specific institutional environment in which the statistical production is carried out and the specific needs of our "domestic" users when setting up our own concepts and approaches.

In other words, when setting up the quality framework, we tried to equally consider the needs and demands of European as well as domestic users.

If the first years were more or less devoted to working in the theoretical field, in the recent years we have mostly focused on the implementation of the defined concepts into the reality of the everyday statistical production. The special attention in these recent years' activities has been devoted to the development and implementation of the quality indicators. Quality indicators are numerical values which should measure and follow the quality of the statistical processes and results. These indicators should be understandable in their definition, easy to calculate and sensibly interpretable for the users and producers. Our aim is to make these indicators a useful tool which would serve two main purposes:

- To enable quick and efficient insight into the statistical process, enable the producers to assess its quality and to timely detect possible errors or deficiencies.
- To provide the "core material" for the user-oriented quality reports. The indicators in these quality reports should show the degree of quality of the disseminated results through different dimensions. In the case that some of the results are assessed to be of lower quality, this should be clearly indicated and the limitations of the interpretability of these results should be stated in a transparent way.

In the first part of the paper we briefly describe our quality framework and its connections to the framework developed inside the ESS. We then move to the implementation part, trying to explain the work that has been done in order to make the theoretical principles useful in practice. The main part of the paper will be devoted to the quality indicators and their role in different parts of the statistical process.

Quality assessment and reporting at SORS

As mentioned, SORS started the preparatory work for systematic quality assessment and reporting in 2003. The basic document which served us as a starting point for the initial activities was the Eurostat's methodological document Standard Quality Report. The document was at that time also translated into the Slovene language. On the basis of this document a special internal working group prepared the general framework including the draft template for quality reports and the draft list of quality indicators. These two elements were later accomplished through the experiences gained in the pilot study, where the exhaustive quality report for the Retail Trade Survey was prepared.

The backbone of the system are six well known standard quality dimensions (relevance, accuracy, timeliness and punctuality, accessibility and clarity, comparability, coherence) with the seventh, supplementary dimension costs and burden. These dimensions are taken directly from the ESS framework, with some slight adjustments in the definition of the sub-dimensions. The structure of our quality reports hence always follows these dimensions and consequently defined sub-dimensions. Closely connected to the quality dimensions are the related quality indicators, which were also defined in the framework of the initial work of the working group. The list of quality indicators was also defined with the adjustment of the Eurostat's list, where some of the indicators were omitted and some new were added.

After the main theoretical elements of quality assessment were defined, we turned to the implementation work and started to prepare the quality reports for some of our surveys. The first quality reports, based on the rules and principles defined by the working group, were hence prepared in 2004 and 2005. These reports were still entirely producer-oriented, meaning that they were prepared for the internal purposes and were not publicly disseminated. In 2006 the first user-oriented and publicly disseminated reports were prepared. From then on, we have tried to gradually cover more and more statistical areas, preparing each year as many quality reports as feasible.

The preparation of the quality reports is now a common practice at SORS. In general, two types of quality reports are prepared. The exhaustive standard quality reports should be prepared every five years. In these reports all the quality dimensions should be thoroughly studied and reported. The proper presentation of the values of the standard quality indicators as well as all the textual explanations should be included in this report. Standard quality reports are usually 20-30 pages long. The annual quality reports are on the other hand much shorter, consisting mostly of the values of the quality indicators, but should be prepared each year. These reports are usually around 10 pages long. The standard quality reports are disseminated only in the Slovene language, while the annual quality reports are translated and disseminated also in English.

Since it was not feasible to produce the quality reports for all the different surveys at once, we gradually supplemented the list of surveys with disseminated reports. We prepared the standard procedure, shortly described as follows:

- Each year a list of 10-20 new surveys which are designated to prepare the standard quality report is determined.
- The standard quality reports for these surveys are prepared by the survey manager and other persons involved in the survey process. Before the dissemination, the reports are edited by the quality coordinators from the General Methodology Unit.
- On the basis of the standard quality report, shorter annual report for the same year is prepared and disseminated.
- From this year on the annual report should be prepared (updated) each year.
- By the current plan, all the standard quality reports should be updated every five years after the first year of the preparation.

The initially defined concepts of quality assessment and quality reporting were mainly targeting the classical surveys, the surveys where the data are still gathered primarily for statistical purposes, using some kind of survey questionnaires. Since the use of the administrative sources has a long history and has especially increased in the last decade, we are now more and more frequently faced with the surveys which completely or at least partly use administrative data as the direct data source. In these cases the need for the adjustment of some of the approaches is more and more clearly pointed out. Therefore, in 2009 a slight revision of our main concepts was carried out, aiming at adjusting the quality reporting templates and adjusting the list of quality indicators to fit better to the wider definition of the statistical survey. One very important task, which is left for the near future, is to prepare the quality assessment framework for the registers and large administrative databases, which are used by several surveys.

Quality indicators

The exhaustive quality information could always be seen as to consist of two parts: the qualitative and the quantitative one. While the qualitative part textually explains and describes the different aspects of the quality dimensions, the quantitative part provides the numerical values which should (also properly explained) show the degree of the attained quality in a more exact, mathematical way. These indicators should, of course, be defined in an understandable and "easy to calculate" way, if they should serve our purposes well. It is also important that their definitions are harmonized when we use them in different countries and in different statistical domains. Otherwise their usage is very locally limited and their explanatory power is much lower.

The modern quality assessment approaches consider the statistical product as a result of a complex statistical process and when assessing the quality of the statistical product it is of great importance to monitor and assess different parts of the process as well as the final statistical product and also the institutional environment where the statistical production is carried out. Hence, also the quantitative information that is provided by the quality indicators, should take into account all these different aspects. Especially when the statistical process is in question, the use of properly defined, properly calculated and properly interpreted quality indicators can provide a great benefit.

SORS has defined the first list of standard quality indicators for the purposes of quality reports, but at the same time they were also considered to become a tool for statistical production monitoring (see Seljak, Zaletel 2004). The whole list of our indicators could hence be conceptually divided into two parts: predominantly product-oriented and predominantly process-oriented indicators. We will now firstly present the whole list of indicators, then explain which of them are considered as process-oriented and finally describe how they are used in our statistical process.

Whole list of SORS's standard quality indicators

Here we present the complete list of standard indicators, which were defined through the last revision in 2009 when we added some indicators designed for the purposes of the surveys which use administrative data as a direct data source. These indicators are: Coherence of the reference dates (R2); Bias due to non-coverage (A1_1); Rate of unsuccessful linkage of the data source (A3_1); Rate of coherence of the data sources (A7). The "extended" form of the notation (e.g. A1_1) means that this indicator should replace the primarily defined indicator (e.g. A1) when the administrative data source is used.

In our model the quantitative quality assessment can refer to three different objects: survey as a whole variable, statistics (statistical result). Therefore, all the indicators in the table presented below are grouped into three groups with accordance to the related object of assessment, using notations Sv (survey), Vr (variable) and St (statistics).

Table 1: List of standard quality indicators used at SORS

Quality Component	Notation	Indicator	Object
Relevance	R1	Rate of unavailable statistics	Sv
	R2	Coherence of the reference dates	Sv
Accuracy	A1	Sampling error	St
	A1_1	Bias due to non-coverage	St
	A2	Unit non-response rate	Sv
	A3	Item non-response rate	Vr
	A3_1	Rate of unsuccessful linkage of the data source	Vr
	A4	Overcoverage rate	Sv
	A5	Imputation rate	Vr
	A6	Editing rate	Vr
	A7	Rate of coherence of the data sources	Vr
Timeliness and punctuality	T1	Timeliness of the first release	Sv
	T2	Timelines of the release of the final results	Sv
	T3	Punctuality of the first release	Sv
Accessibility and clarity	AC1	Number of means used for dissemination	Sv
Comparability	C1	Length of comparable time series	St
Coherence	CH2	Coherence between first and final results	St
	CH3	Coherence with comparable data from other sources	St

Process quality indicators

As mentioned, some of the above listed indicators are especially important when the monitoring of the statistical process is in question. Here we provide the list of these indicators. We also briefly explain for which part of the process(es) the indicator should (at least indirectly) refer. The indicators which are by our model considered as process indicators are:

- **Sampling error**. The indicator refers to the sampling as well as to the estimation procedure. In the case that non-response adjustments have been taken into account when the sampling error was estimated, it also (indirectly) reflects the quality of the data collection.
- **Unit and item non-response rate.** Directly indicates the quality of data collection part of the process. It also (at least indirectly) tells something about (at least some parts of) the survey design (e.g. selection of the data collection mode). In the case that administrative data are used, it can also indicate the (ir)relevance of the used data sources.
- **Editing and imputation rate.** It also mostly refers to the data collection part of the process. In the case that the automated editing system is used, it can also refer to the data processing part. Namely, the large number of corrected values can be in these cases the consequence of too strictly defined logical check.
- **Overcoverage rate.** It mostly refers to the sampling frame construction. High value of this indicator can also indicate the possible deficiencies of the main data sources (registers) which were used for

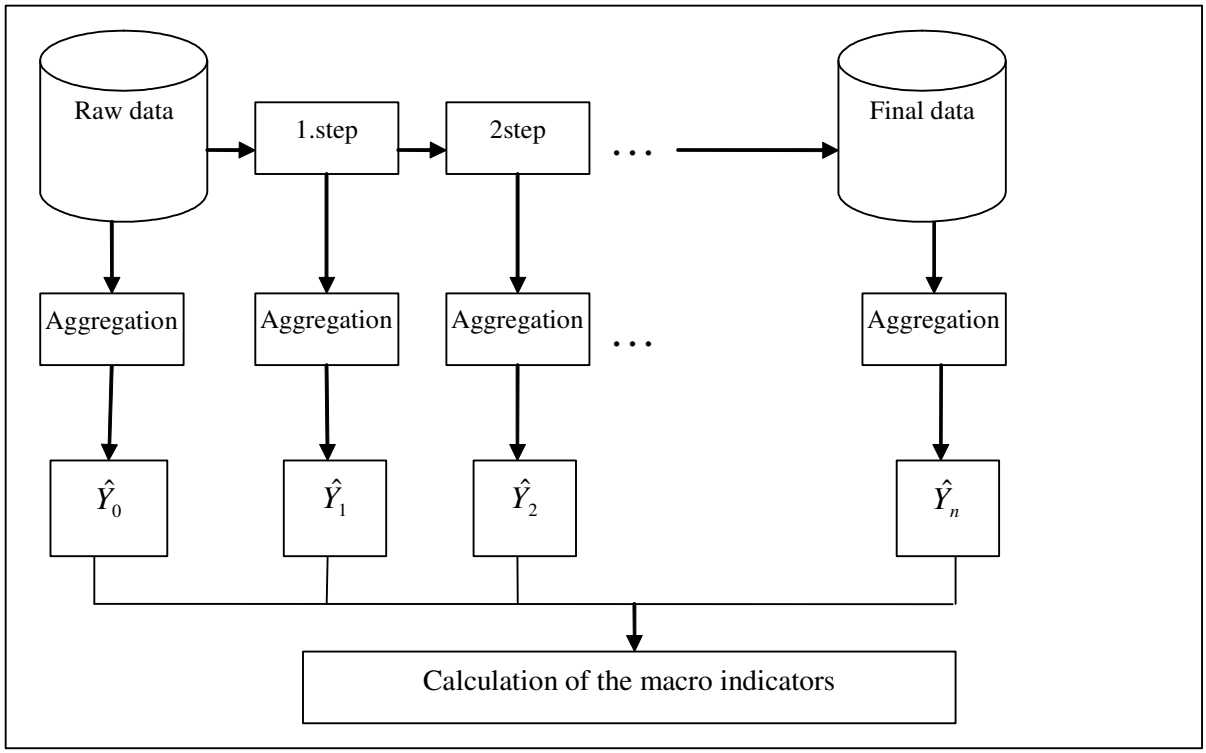
the purposes of the sampling frame construction.

Process quality indicators at macro level

In addition to the above listed standard quality indicators that are through the process calculated at the micro level, we usually also calculate some indicators which are calculated at the level of the already aggregated data and therefore we call them *the macro quality indicators*. The main goal of these indicators is to show us how the particular parts of the statistical process influence the statistical results. Especially in the cases when the statistical process is fully automated, survey managers can through these indicators get at least indirect indication what is going on in the "black box" of the automated process.

The calculation procedure of these indicators can be quite generally presented. Let $\hat{Y}_0$, $\hat{Y}_1,\ldots,\hat{Y}_n$ be the estimated values of the target population parameter after the consecutive steps of the statistical process (e.g. $\hat{Y}_0$ - raw data estimate; $\hat{Y}_1$ - estimate after manual editing). The macro indicators are the differences between these "intermediate" and the final aggregate ($\hat{Y}_n$). In some case (e.g. totals) it is more appropriate to calculate the relative difference ($\left|\hat{Y}_k - \hat{Y}_n\right| / \hat{Y}_n$), while in others (e.g. indices) the absolute difference ($\left|\hat{Y}_k - \hat{Y}_n\right|$) should be calculated.

Figure 1: Graphical presentation of the macro indicators calculation



Automation of the process indicators calculation

If we want to make the process quality indicators an efficient tool for monitoring the statistical process, they should be at disposal quickly and in a usable form. Quick availability of the process indicators is of special importance in the case of short-term surveys, where timeliness is the key aspect. Since the results should be produced and disseminated as quickly as possible, also the process quality indicators should be calculated as soon as possible, so that the survey manager can discover some possible deficiencies before the data dissemination.

At SORS many surveys faced methodological and technical revisions in the recent years. One of the elements of these revisions was the automation of the calculation of the process indicators listed above. The goal is that these indicators would be at disposal at the same time as the statistical results. To enable such an automated system, the following two conditions should be fulfilled during the technical revision:

- The statistical process should never overwrite the original or any other version of the data. In other words, the changes should always be performed by creating a new data version.
- All the changes performed during the statistical process should be flagged in a transparent and unambiguous way. For each data change, the process metadata about what, why and how was changed should be created and stored.

The traceability of data changes in the statistical process is assured by the usage of the special metadata variables called the statuses of the variables. For each variable V, which could be subject of a change during the statistical process, the status variable V_s must be defined. These variables could be stored in the same table as the "original" variables or in a separate table. The values of the status variables are determined by the standard code list, consisting of 4-digit codes, which should cover all the possible steps in the process and all the possible methods that could be used in these steps. The main rule is that if the value of the variable is changed during the process, also the status variable must be changed respectively.

Example 1: If for instance the data for a certain variable are obtained by using a postal questionnaire, the initial status for this variable is "11.11" (*Original value obtained by the postal questionnaire)* for all the responding units. If later in the data editing phase the value was checked and corrected through the telephone re-contact, the status changes to "12.12" (*Value corrected on the basis of the telephone re-contact)*

Example 2: If for a certain variable the value is missing, the initial value of the status is "00.00" (*Item non-response*). If later in the process the missing value was imputed by using the hot-deck method, the status gets the value 41.15 (*Value imputed with the hot-deck method*). If in the stage of the final edits this value was corrected again, the status gets the value 42.15 (*Corrected imputed value*).

In the values of the status variables we record the information what has happened with the certain data item during the process. With this information it is then quite easy to calculate the values of the process quality indicators as well as the values of the macro quality indicators. The indicators are automatically calculated by using standard SAS macros then exported into the several EXCEL spreadsheets where they can be reached by the users. In the case of monthly surveys, calculation of the indicators is always done for the current as well as for several (usually at least 11) previous months. This enables the survey manager not only to see the indicators for the results that should be published, but also to compare them to the indicators for the results of the previous months. Indicators which could be seen in the form of EXCEL table could usually also be graphically presented by EXCEL chart.

Presentation of quality indicators for the Retail Trade Monthly Survey

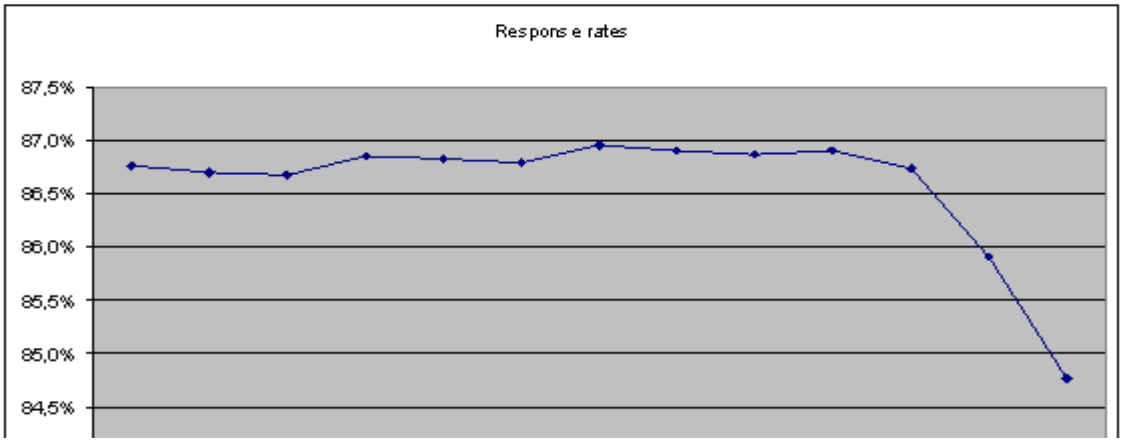
The calculation and presentation of the process and macro quality indicators will be presented for the concrete example of the Retail Trade Survey (hereinafter RTS). At SORS this survey is carried out monthly, by using a combination of data obtained from the classical survey (postal questionnaire) and the VAT data which are monthly provided by the Tax Authorities. Field data are collected for the smaller part of the largest units, which represent only 5% of the number of units, but more than 50% of the monthly turnover. The data for all the remaining 95% of units are obtained from the administrative source. In both cases of data collection we are facing a certain degree of non-response. Since in the case of the "field" survey we are dealing only with the large units, the non-response rate is quite low (approx. 3%). In the case of "administrative data units", the non-response is mostly due to the fact that some units are not obliged to report monthly but quarterly. Therefore, in some activity groups where there are a lot of small units the non-response rate on the monthly basis can be quite significant.

The non-response rate could be observed separately for the field and administrative data units or together for the whole set of observation units. Here we present how the values of the response rates for the whole set of observation units is presented. The values of the indicators are always calculated and graphically presented for the previous 13 months

Figure 2: Table presentation of the RTS response rates

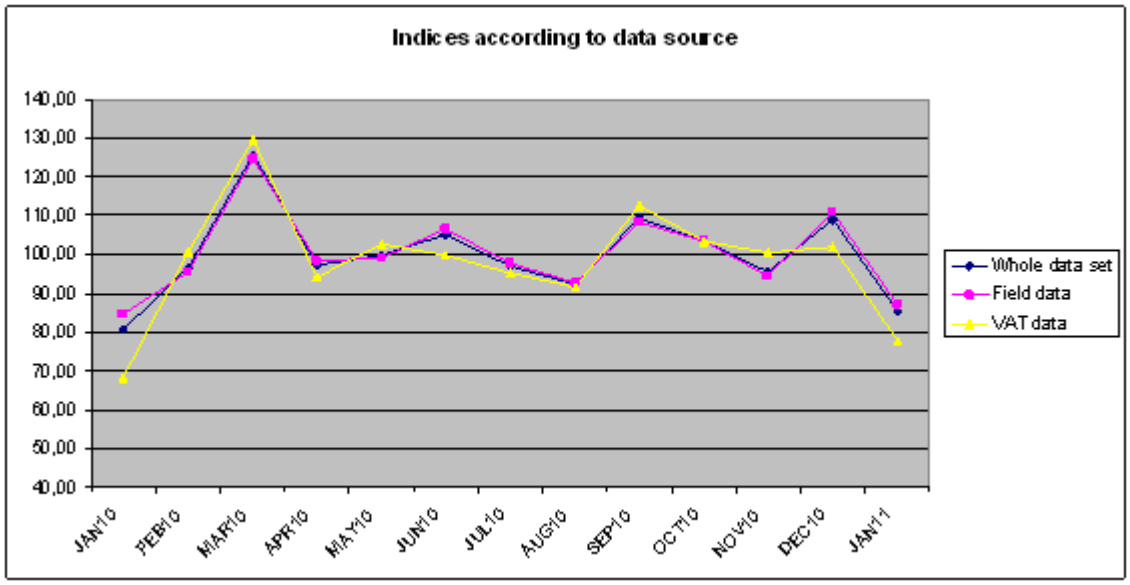
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		JAN10	FEB10	MAR10	APR10	MAY10	JUN10	JUL10	AUG10	SEP10	OCT10	NOV10	DEC10	JAN11
2	Activity group													
3	Total	86,8%	86,7%	86,7%	86,9%	86,8%	86,8%	87,0%	86,9%	86,9%	86,9%	86,7%	85,9%	84,8%
4	AG 01	94,3%	93,9%	93,9%	94,1%	93,9%	94,1%	93,9%	94,1%	93,7%	93,5%	93,2%	92,2%	91,8%
5	AG 02	60,3%	60,3%	59,7%	58,4%	58,4%	57,9%	57,3%	57,3%	57,3%	57,3%	57,3%	57,3%	62,2%
6	AG 03	91,5%	91,5%	91,4%	91,1%	91,1%	91,1%	91,6%	91,6%	91,3%	90,6%	90,5%	90,4%	88,1%
7	AG 04	88,4%	88,4%	88,4%	89,6%	89,6%	89,5%	88,9%	88,3%	88,8%	88,8%	88,1%	88,1%	83,1%
8	AG 05	82,9%	82,9%	82,9%	82,3%	82,3%	81,6%	81,0%	81,5%	81,4%	81,3%	81,2%	79,9%	84,4%
9	AG 06	97,6%	97,6%	97,0%	96,3%	96,3%	96,3%	96,3%	96,3%	96,3%	96,3%	96,3%	95,1%	97,0%
10	AG 07	78,2%	78,2%	78,5%	78,3%	78,3%	78,3%	78,9%	78,8%	79,2%	79,5%	79,8%	78,7%	78,3%
11	AG 08	94,8%	94,8%	94,8%	94,8%	94,8%	94,8%	94,7%	94,0%	94,0%	94,0%	93,2%	92,0%	92,5%
12														

Figure 3: Graphical presentation of the RTS response rates



For the RTS subject matter specialists it is also very important to see the difference between the statistical results coming out of the field data and the ones coming out of the administrative data. Therefore, one of our macro quality indicators is based on the differences between indices estimated on the basis of the field data, indices estimated on the basis of the administrative data and indices estimated on the basis of the whole set of data. These differences are also presented in the form of a table as well as graphically. Here we only show the graphical presentation.

Figure 4: Differences between indices calculated from different data sources



Conclusions

Quality assessment and quality reporting is one of the most important activities in the practice of SORS and in the recent years a lot of effort has been made in order to make these activities regular and systematic part of the statistical process. Two main goals should be fulfilled by the regular quality assessment. The first one is to provide to the users exhaustive information on different aspects of data quality, including their eventual deficiencies and limitations. The second goal is to enable the data producers the insight into the different phases of the automated statistical process, enable them easy and quick detection of the eventual critical points. This goal should be mostly fulfilled by the calculation of the process and macro quality indicators. Their actual effectiveness is achieved only if their calculation is fully automated and if they are at disposal at the same time as the statistical results.

A lot of work on the modernization of the quick and efficient quality monitoring has already been done, but there is still plenty of work to be done. The most important tasks for the future are further introduction of the automated system in most of the surveys and construction of the central database of the quality indicators where all the most important quality information would be stored centrally.

REFERENCES

[1] Lyberg L. et al.: Survey Measurement and Survey Quality, Wiley, 1997.

[1] Seljak R., Flander Oblak A.: Quality Assessment of the Register-based Slovenian Census 2011; Paper presented at the Joint UNECE/Eurostat Meeting on Population and Housing Censuses, Geneva, 13-15 May, 2008

[1] Seljak R., Ostrež T.: Quality Reporting at SORS – Experiences and Future Perspectives. Paper presented at the European Conference on Quality and Methodology in Official Statistics, Helsinki, 4-6 May, 2010.

[1] Wallgren A., Wallgren B.: Register-based Statistics; Administrative Data for Statistical Purposes: John Wiley & sons, 2007.

[2] Zaletel M., Seljak R.: Measurement of Data Quality in the Case of Short Term Statistics. Paper presented at the European Conference on Quality and Methodology in Official Statistics, Mainz, Germany, 24-26 May, 2004.