



United Nations

UNECE Weekly

Economic Commission for Europe

Issue N° 213 – 26 February-16 March 2007

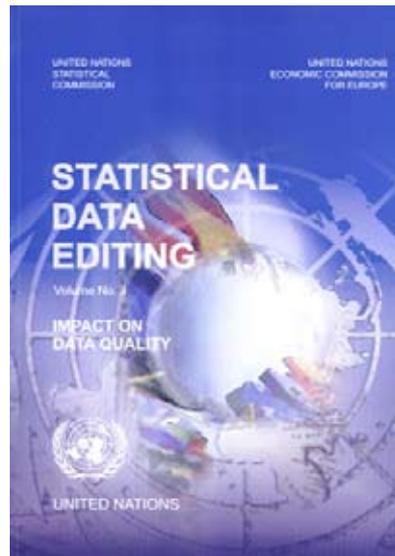
Statistical Data Editing - Data Quality

Why edit statistical data?

The data collection phase has a decisive role in determining the quality of the whole statistical process and the final output. There are errors and even missing values within the raw data sets. While such errors and omissions are intentional or not, statisticians need a more complete picture of the reality.

The role of editing comprises the following:

- Identifying problems in data collection and processing with a focus on getting accurate responses, or getting “true” values, from all of the respondents.
- Providing intelligence related to both survey design as well as data dissemination.
- Identifying means of improving data quality.



Impact on Data Quality

This volume examines how editing affects data quality. Data quality is understood in its broadly accepted meaning, that is, its fitness for use, or fitness for purpose. The publication addresses issues of accuracy, relevance, coherence and interpretability. Various stakeholders in the survey process are taken into account, such as the users, respondents or data producers themselves.

Chapter 1 provides examples of evaluation frameworks of the editing process: what to plan for, what to measure, what to retain, so we can learn from the process. It also illustrates the practical implementation with an example from the Italian National Statistical Institute.

Chapter 2 addresses specific issues of what to measure. Various quality indicators are proposed, both from a theoretical and practical point of view. This chapter examines how editing affects data quality in a quantitative way. The impact of imputation on the total survey error is taken into account. The chapter concludes by

Ukraine adheres to the CMR Convention

On 16 February 2007, Ukraine deposited with the United Nations Secretary-General an instrument of accession to the Convention on the Contract for the International Carriage of Goods by Road (CMR). This brings the total number of States Parties to the CMR to 52. The CMR will enter into force for Ukraine on 17 May 2007.

The CMR Convention fixes the conditions for the contract between the forwarder and the carrier for the international carriage of goods by road. In particular, it contains provisions on the consignment note of the goods carried and establishes the conditions for liability, for example in the case of loss of the goods or delays.

The UNECE Working Party on Road Transport is currently negotiating an additional Protocol to the CMR, which would allow the use of electronic consignment notes in international road transport. ❖

For more information contact José Capel Ferrer (jose.capel.ferrer@unece.org) or Marie-Noëlle Poirier (marie-noelle.poirier@unece.org).

UNECE experts discuss innovation and competitiveness policies

Why is innovation the main driver of competitiveness in the knowledge-based economy? How do the results of technological development and innovation diffuse in the modern economy and which are the most effective organizational models? What role can the public sector play in facilitating the generation and diffusion of innovation? What is the experience of different countries with policies facilitating technological development and innovation?

These are some of the issues addressed and discussed at the first meeting of the UNECE Team of Specialists on Innovation and Competitiveness Policies on 8-9 March in Geneva.

The TOS-ICP was established by decision of the UNECE Committee on Economic Cooperation and Integration at its first session in September 2006. The work of the Team will focus on the identification of good practices leading to policy recommendations related to innovation and competitiveness. These outcomes will

be broadly disseminated in the UNECE region, providing the basis for future capacity building and technical assistance activities in response to demands by member States.

The purpose of the first meeting of the TOS-ICP was to focus the Team's work on those aspects of innovation and competitiveness policies where multilateral cooperation could add the most value and where national experiences can be most easily taken on by other countries, and to adopt an implementation plan for its activities in 2007 and 2008.

A web-based virtual platform for information exchange developed by the UNECE secretariat was presented. This innovative tool will serve to facilitate the collaborative efforts of the TOS-ICP members and other associated experts between sessions as well as the dissemination of information and communications among them. ❖

For more information contact José Palacín (jose.palacin@unece.org).



cont.

suggesting what quality information should be provided to the data users and in what way.

Chapter 3 addresses the problem of improving data quality and the question:

What can statisticians learn from the editing process in order to be able to continuously improve the product? Detailed examples of how the data editing process should fit into the statistical infrastructure are presented. More globally, the chapter also addresses the problem of where and how editing should fit into the various modes of data collection, as it relates to electronic data reporting and the use of alternate data sources.

Chapter 4 considers quality in its full breadth. It reviews what has been done, and what was planned but not done, in evaluating quality. The authors argue that editing should be considered in a wider perspective by providing information to other survey processes in a systematic way. The future challenges are spelled out in this chapter:

The aim of this publication is to assist National Statistical Offices in assessing the impact of the data editing process on data quality, that is, in assessing how well the process is working. We hope that this publication will help the statistical community of the UNECE region to improve the quality of official statistics.*

For more information contact the UNECE Statistical Division (support.stat@unece.org). Available from the UN Publications and Sales Section, Palais des Nations, Geneva (unpubli@unog.ch).

How are statistical data edited?

One may say that the easiest solution would be to call back the respondents and discuss with them questions related to their responses. However, this is not always possible. The respondents are usually not happy to be called back, and consider data reporting as an administrative burden. Moreover, imagine a survey with several thousand responses, or a census with millions of responses. Statistical offices don't have sufficient manpower to verify data manually and call back all respondents. Most often imputation of missing values and editing of data is based on complex models and mathematical methods.

How much does it cost?

Assisted by the incredible computing power of modern technology, there is a significant number of highly qualified specialists involved in editing and imputation of statistical data. Depending on concrete cases, the data editing phase may be time-consuming and delay the release of final outputs. All this makes editing an expensive part of the statistical survey process.

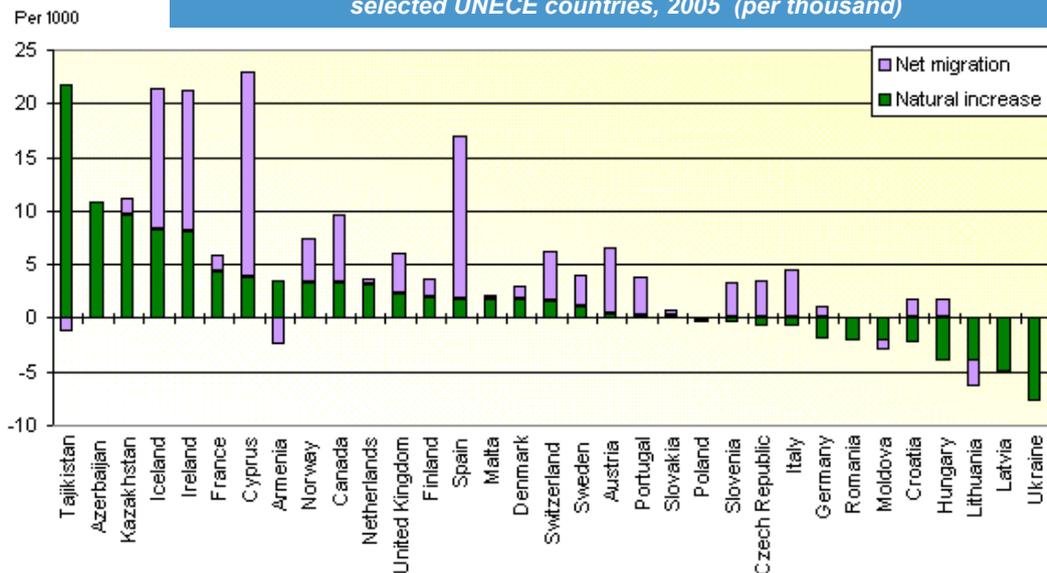
This raises an important question: Is data editing worth the time and resources needed? The answer is not simple. It is necessary to compare the quality before and after editing. The impact of editing and imputation on quality will depend on the nature of the data, the method of data collection, data editing methods used and several other aspects.

How can the costs of data editing and imputation be managed?

Comparing the quality of data before and after editing is very important for managers, who can decide on whether and "how much" to edit. It is a typical cost-benefit consideration, that is currently in the focus of managers and experts. One of the ways to decrease costs is to edit only selected records. Selective editing is cheaper, and it still improves the quality of statistical aggregates included in the output. Another approach is to use less severe criteria for identifying suspicious values, so called outliers. This might decrease the editing workload without compromising the desirable quality level.

Facts and figures

Contribution of natural increase and net migration to population change in selected UNECE countries, 2005 (per thousand)



Source: UNECE Statistical Database, compiled from national and international official sources.

Note: Data are arranged in order of the contribution of natural increase to the population growth.

Over the last two decades many UNECE countries experienced declines in fertility rates, which affected the pace of the overall population growth. In 2005, the population in one fifth of the UNECE countries could not be maintained through the natural balance (the difference between births and deaths). Apart from a few exceptions, in the rest of the UNECE countries the contribution of the natural increase of population was very low and was often exceeded by the net migration.

For example, in 2005, an influx of immigrants outweighing by far the number of emigrants greatly raised the overall population increase in Cyprus, Spain, Ireland and Iceland. On the other hand, in several UNECE countries the balance of migration was negative, which in some cases – Lithuania, Moldova, Romania and Poland – further aggravated the overall population decline.