

**UNITED NATIONS STATISTICAL COMMISSION and  
ECONOMIC COMMISSION FOR EUROPE  
CONFERENCE OF EUROPEAN STATISTICIANS**

**Work Session on Statistical Data Editing**  
(Ottawa, Canada, 16-18 May 2005)

**INFORMATION NOTICE No.1**

At the invitation of the Government of Canada,  
the Work Session will be held in Ottawa from 16 to 18 May 2005.

**I. PURPOSE OF THE MEETING**

1. At its 2004 plenary session, the Conference of European Statisticians included the Work Session on Statistical Data Editing in its 2005-2006 meeting programme. At the kind invitation of Statistics Canada, the meeting will be held from 16 to 18 May 2005 in Ottawa (Canada). The study programme of the meeting focuses on methods and evaluation, with a particular emphasis on quality related issues. The publication "Statistical Data Editing – Volume 3" concerning the impact of editing on data quality will also be discussed at the meeting.

2. The target audience of the meeting includes statisticians dealing with editing and imputation of statistical data derived from surveys, censuses, and administrative sources and spanning various subject-matter areas, both social and economic.

**II. AGENDA OF THE MEETING**

3. The programme of the meeting will consist of the following substantive topics:

- (i) Editing administrative data and combined data sources;
- (ii) Implementing editing strategies and links to other parts of processing;
- (iii) Electronic data reporting – editing nearer source and multimode collections;
- (iv) New and emerging methods, including automation through machine learning, imputation, evaluation of methods;
- (v) Quality indicators and quality reporting.

4. Detailed explanatory notes on the nature and expected outcomes of topics (i) – (v) are provided in Section V of this Information Notice.

**III. PARTICIPATION AND ACCREDITATION**

5. Representatives of all Member States of the United Nations and of interested intergovernmental organizations are welcome to participate in the meeting. Participants representing non-governmental organizations in a consultative status with the United Nations Economic and Social Council may also attend. All participants must be accredited by the competent authorities of their country or international organization. All participants attending the meeting are requested to have a valid passport and, if required, a visa. Applications for visas should be made as soon as possible to the Embassy of Canada in the country in which the participants reside, with a reference to the UNECE Work Session on statistical data editing. The statistical offices and international organizations should inform the UNECE secretariat **before 15 December 2004** whether their organization intends to participate at the meeting and/or submit a contribution on the topics mentioned above. Please indicate the topic to which the paper will refer as well as *participant(s) name(s) and their contact data*.

6. Participants should complete the registration form, which can be found on the last page, and return it to both the UNECE Secretariat and Statistics Canada **by 1 April 2005**.

#### IV. DOCUMENTATION, METHODS OF WORK AND OFFICIAL LANGUAGES

7. The working language of the work session is English. The documents may be submitted in any of the official languages of the UNECE (English, French, Russian), but no translation and interpretation facilities will be available. The following deadlines and requirements apply:

- The intention to submit a paper should be communicated to the UNECE Secretariat **by 15 December 2004**.
- The papers, not exceeding 10 pages, should be sent electronically to the UNECE Secretariat **by 15 February 2005 at the latest**.
- Documents should be submitted in MS Word and/or PDF formats. Templates in Word containing the header, document number and format will be sent to authors or may be requested by e-mail (diane.serikoff@unece.org). Alternatively, templates in LaTeX can also be provided.

8. For each substantive item on the agenda the discussants will invite 2-3 speakers to write an **invited paper** and to present them at the meeting. All other participants are welcome to submit a **supporting paper** on any topic of the agenda. The supporting papers will be distributed to the delegates. They will be summarized by the discussants and the issues raised in them will also be discussed.

9. Authors of invited papers will have about 20 minutes per paper for the presentation. Authors of supporting papers may be given (time permitting) about 5 minutes to highlight the main issues raised in their papers. Presenters may use PowerPoint presentations, Adobe Acrobat full screen presentations or A4/letter transparencies. UNECE secretariat cannot provide translation of the presentations.

- Presentations in PowerPoint or Adobe Acrobat should be sent to the UNECE secretariat **by 2 May 2005 at the latest**, so that these can be installed on the conference room computer prior to the meeting.

10. Invited and supporting papers will be made available on the website of the UNECE Secretariat at the following location: **<http://www.unece.org/stats/documents/2005.05.sde.htm>** Delegates are encouraged to download the papers from the website and bring their own copies to the meeting. Documents posted on the website before the meeting will **not** be distributed in the conference room.

#### V. EXPLANATORY NOTES TO THE AGENDA

##### (i) Editing administrative data and combined data sources

Discussants: Natalie Shlomo (N.Shlomo@soton.ac.uk) and Heather Wagstaff, United Kingdom (Heather.Wagstaff@ons.gsi.gov.uk)

11. Administrative data are collected for other purposes than statistical and incorporating them in the statistical processes poses unique challenges and opportunities for National Statistical Offices to improve the quality of their statistical data as well as reduce the costs and response burden. The session covers aspects relating to edit and imputation procedures when administrative data are used for enhancing and correcting survey data and for developing direct statistical databases such as registers, frames and censuses.

12. Under this topic, contributions are requested suggesting strategies, new methodologies and best practices when incorporating administrative data in the edit and imputation procedures for processing statistical data, and in particular, the impact on the quality of the statistical data.

13. The session will focus on these main areas of research:

- Effective methodologies and strategies for editing and imputing administrative data directly in order to produce reliable sources of information for use in statistical processes;

- Edit and imputation procedures and best practices for editing and imputing survey data when administrative data are available;
- Combining multiple sources of administrative and survey data for constructing high quality statistical databases using effective edit and imputation strategies which will ensure correct coverage, consistent and clean records.

14. Papers describing experiences in both demographic and business areas are requested and should include general issues and methodologies beyond specific applications with an emphasis on the impact and assessment of the quality of the statistical data.

**(ii) Implementing editing strategies and links to other parts of processing**

Discussants: Orietta Luzi, Italy (luzi@istat.it) and Carsten Kuchler, Germany (carsten.kuchler@destatis.de)

15. Designing and implementing a data editing strategy requires the management of generally complex data flows and the harmonization of editing with the other parts of the survey process. The focus of this topic is on the integration of single editing techniques with respect to comprehensive data editing strategies, particularly as reflected in the (re-)design of editing processes. Contributions should consider their impact on data quality, the organization of work and application flows or management issues. However, presentations of particular activities should relate to the underlying objectives pursued by the implemented data editing strategy. In order to share experiences, reports of ongoing or already completed implementation projects are especially welcome.

16. Of particular interest is the impact of editing strategies on related stages of the survey process. The work session especially encourages contributions that relate data editing to the following three processes:

- survey design (adapting sample design, improving questionnaire design, harmonizing and optimizing editing at the data capture stage, efficient use of auxiliary information and external data sources in data editing processes, monitoring and optimizing the efficiency of survey processes by exploiting information on errors and error sources from data editing, etc.);
- advanced estimation methods (modeling background knowledge from the editing process, assessing the impact of editing on the statistical properties of estimators, etc.);
- release of micro data (supporting disclosure elimination with information from data editing, providing background knowledge to users, multiple imputation, etc.).

**(iii) Electronic data reporting – editing nearer source and multimode collections**

Discussants: Pedro Revilla, Spain (previlla@ine.es); and Paula Weir, United States (pweir@eia.doe.gov)

17. Electronic data reporting (EDR) offers the possibility of using built-in edits in electronic questionnaires previously not possible in paper questionnaires or other modes of data collection. This topic covers all issues relating to editing as it pertains to EDR. This includes strategies and methodologies with respect to implementing editing at the point of data collection and its relationship to some other modes of collection and to editing that occurs in the post-collection processing. Topics of interest include:

- the impact of EDR on the editing strategy;
- the optimization of the effectiveness of both editing at data capture and at post-collection survey processing;
- performance measures and indicators of editing at data capture and post-collection processing as it affects the overall survey quality;
- challenges and issues such as security and confidentiality, respondents' burden, response rates, timeliness, and incentives;
- the use of focus groups, usability or cognitive testing of data providers with respect to editing at data collection.

**(iv) New and emerging methods, including automation through machine learning, imputation, evaluation of methods**

(Discussants: Ton de Waal, Netherlands (twal@cbs.nl) and Maria Garcia, United States (maria.m.garcia@census.gov))

18. This topic covers new methods of data editing and imputation. The contributions could report on new methods and software, implementations in different situations, or organizational issues related to the development or implementation of new editing and imputation methods. Topics of interest include:

- the detection and treatment of systematic errors;
- the detection and treatment of outliers;
- development of new software applications;
- impact of new systems on organization efficiency;
- the use of graphical editing techniques;
- temporal editing and imputation;
- selective editing techniques;
- imputation under edit restrictions;
- outlier-robust imputation;
- variance estimation in the presence of imputation; and
- machine learning methods for editing and imputation.

19. Presentations of new methods should be linked with an evaluation of the operational feasibility of applying the ideas in a variety of data or organizational situations. Empirical comparison studies and evaluations are particularly welcome. These studies and evaluations could focus on aspects such as the processing time, the requirements for storage capacities, and expected impact of edit systems on data quality.

**(v) Quality indicators and quality reporting**

(Discussants: Leopold Granquist, Sweden (leopold.granquist@scb.se); and Svein Nordbotten, Sweden (svein@nordbotten.com))

20. A similar subject was discussed in the Work Session on Statistical Data Editing in Madrid, October, 2003: "Development and use of data editing quality indicators". The Work Session concluded that editing indicators should focus on problems in the collection process to obtain ideas on how to improve the incoming data quality in future surveys. Therefore, papers that elaborate on the aims and properties of quality indicators described below will be welcome.

21. The indicators should be developed with the clear aim in mind of what they are supposed to measure and why. Questions that need to be answered are, e.g. for whom and for what purpose the quality indicators are needed, which indicators to choose and why, how to collect the data for them, and how to use the collected information for improving the survey process. It was agreed that the aims of the editing quality indicators are:

- (i) to obtain knowledge of measurement errors or severe problems for respondents to provide answers, in order to have a basis for improving the incoming data quality of the survey, and
- (ii) to have a basis on which to take measures to increase the efficiency of the data editing process.

22. In addition, the indicators should be useful not only for producers but also for the users of data. It is necessary to produce indicators that can inform the users about data quality in general and to have a strategy on how to communicate this information to users.

**VI. FURTHER INFORMATION**

23. For further information, please contact the following organizers:

**UNECE:** Mr. Juraj Riecan  
Head, Dissemination and Statistical Information Systems  
Statistical Division, United Nations Economic Commission for Europe  
Palais des Nations  
1211-GENEVA 10, Switzerland  
Tel.: +41 22 917 33 28 Fax: +41 22 917-0040  
e-mail: juraj.riecan@unece.org

**Statistics Canada:** Mr. Claude Poirier  
Chief, Business Survey Methods Division  
Statistics Canada  
120 Parkdale Ave.  
OTTAWA K1A 0T6, Canada  
Tel: +1-613 951-1491 Fax: +1-613 951-5711  
e-mail: poircla@statcan.ca

**UNITED NATIONS STATISTICAL COMMISSION and  
ECONOMIC COMMISSION FOR EUROPE  
CONFERENCE OF EUROPEAN STATISTICIANS**

**Work Session on Statistical Data Editing**  
(Ottawa, Canada 16-18 May 2005)

**REGISTRATION FORM**

This form should be completed in duplicate by each participant  
and returned **by 1 April 2005**

**1 copy to:** Ms. Diane Serikoff  
Tel: +4122 917 2241  
Fax: +4122 917 0040  
E-mail: diane.serikoff@unece.org

**1 copy to:** Mr. Claude Poirier  
Tel: +1-613 951-1491  
Fax: +1-613 951-5711  
E-mail: poircla@statcan.ca

|   |   |                |
|---|---|----------------|
| 1. Surname (Family name):   | Mr. ?<br>Ms. ?                              | 2. First name: |
| 3. Representing (country or organization):                                      |   |                |
| 4. Official function (in home country):   |   |                |
| 5. Mailing address:   |   |                |
| 6. Tel. No.:  | 7. Fax No.:                                 |                |
| 8. E-mail address:  |   |                |
| 9. Please indicate topic ((i)-(v)) and title, if you intend to present a paper: |   |                |
| 10. Date and time of arrival in Ottawa:   | 11. Date and time of departure from Ottawa: |                |
| 12. Special requirements:   |   |                |

Date \_\_\_\_\_

Signature \_\_\_\_\_