

UNITED NATIONS
ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Work Session on Statistical Data Editing
(Paris, France, 28-30 April 2014)

INFORMATION NOTICE No.1*

The National Institute of Statistics and Economic Studies (INSEE), France will host the Work Session at the Maison de la mutualité, 24 rue Saint-Victor, 75005 Paris, France from 28 to 30 April 2014

I. Purpose of the work session

1. At its 2012 plenary session, the Conference of European Statisticians included the Work Session on Statistical Data Editing in its 2013-2014 meeting programme. The work session will be held from 28 to 30 April 2014 in Paris (France) at the kind invitation of the National Institute of Statistics and Economic Studies (INSEE) of France. The study programme of the work session focuses on methods and evaluation with particular emphasis on standardization and quality related issues. Such a programme will support streamlining initiatives in national statistical organizations.
2. The target audience of the work session includes statisticians dealing with editing and imputation of statistical data derived from surveys, censuses, administrative and external sources and spanning various subject-matter areas, both social and economic.

II. Agenda of the meeting

3. The programme of the work session will consist of the following substantive topics:
 - (i) Selective editing / macro editing;
 - (ii) New and emerging methods;
 - (iii) Getting the support of all people when implementing data editing;
 - (iv) Editing of Census and social data;
 - (v) International collaboration and processing tools.
4. Detailed explanatory notes on the nature and expected outcomes of topics (i) – (v) are provided in Section V of this Information Notice.

* The provisional agenda and timetable will be issued at a later date as WP.1. Information on the venue and local arrangements in Paris (hotel accommodation, transportation etc.) will be provided in Information Notice No. 2. See also Section VI 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 work session. 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.**

6. All participants attending the work session 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 France in the country in which the participant resides, with a reference to the UNECE Work Session on Statistical Data Editing. A letter to facilitate obtaining a visa can be requested from the UNECE secretariat.

7. Statistical offices and international organizations should inform the UNECE secretariat **before 16 December 2013** if their organization intends to participate at the work session and/or submit a contribution on the topics mentioned above. Please send a short abstract and indicate the topic to which the paper will refer as well as participant(s) name(s) and their contact data.

8. Participants should complete the registration form, which can be found on the last page, and return it to the UNECE secretariat **by 26 February 2014**.

IV. Documentation, methods of work and official languages

9. The working language of the work session is English. The documents should, therefore, be submitted in English only. The following deadlines and requirements apply:

- A short abstract of the paper should be submitted to the UNECE secretariat (diane.serikoff@unece.org) as soon as possible and **by 16 December 2013 at the latest**.
- The papers, not exceeding 10 pages, should be sent electronically to the UNECE secretariat **by 26 February 2014 at the latest**.
- Documents should be submitted in MS Word or PDF formats. Templates in Word and Latex will be made available on the web site or may be requested by e-mail (diane.serikoff@unece.org).

10. All participants are welcome to submit a paper on any topic of the agenda. For each topic, authors will have the opportunity to give a short presentation of their paper. The Session Organizers may ask a few of these speakers to give a longer presentation.

11. About 20 minutes will be allocated for long presentations while short presentations will be given about 5-10 minutes (time permitting) to simply highlight the main issues raised in their papers. Presenters may use PowerPoint or Adobe Acrobat full screen for their presentations. The UNECE secretariat cannot provide translation of the presentations.

- Presentations in PowerPoint or Adobe Acrobat should be sent to the UNECE secretariat by **14 April 2014 at the latest**. These will be installed on the conference room computer prior to the meeting.

12. All papers and presentations will be made available on the website of the UNECE secretariat at the following location: <http://www.unece.org/stats/documents/2014.04.sde.html>

13. Participants are encouraged to download the papers from the website and bring their own copies to the work session. Documents posted on the website before the work session will **not** be distributed in the conference room.

V. Explanatory notes to the agenda

Topic (i): Selective editing / macro editing

Session Organizers: Orietta Luzi, Istat, Italy (luzi@istat.it), Rudi Seljak, Statistical Office of Slovenia (rudi.seljak@gov.si) and Pedro Revilla, National Statistical Institute of Spain (previlla@ine.es)

14. It is well known that use of macro editing and selective editing procedures can significantly improve the efficiency of the data editing process. This fact has been proven by the many practical procedures that were implemented in the statistical process of national statistical institutes (NSIs) in recent years. This is especially true in the case of business surveys where usually a small amount of the largest units have a decisive influence on the final results and therefore the appropriate use of the combination of macro editing and selective editing techniques can significantly reduce the cost of the data editing process with no major impact on the quality of the final results. Potential contributions should focus on the following aspects:

- Practical implementation of local and global score functions
- Visualization approaches for a macro editing strategy
- Tools to support a macro editing strategy
- Integration of selective and macro editing techniques within the statistical process.

15. Contributions that present a new and innovative implementation strategy inside the statistical process are particularly warmly welcome.

Topic (ii): New and emerging methods

Session Organizers: Maria Garcia, United States Census Bureau (maria.m.garcia@census.gov) and Jeroen Pannekoek, Statistics Netherlands (jpnk@cbs.nl)

16. This topic covers new and emerging methods or other methods for improving and/or optimizing the process of data editing and imputation. In recent years there have been major contributions based on probability models, robust statistical methods, time series analysis, Bayesian networks, data visualization, and machine learning techniques. This topic aims to bring together contributions that include innovative ideas and applications related to various aspects of editing and imputation, for example:

- Defining what an edit failure is
- The detection and treatment of systematic errors
- The detection and treatment of outliers
- The use of graphical editing techniques
- Temporal editing and imputation
- Editing sub-annual data in a timely fashion
- Editing categorical variables
- Editing of data when using direct access to company information
- Editing of data acquired through electronic collections

17. Contributions can also report on innovative applications of existing methods not covered by the other topics. For instance, methods and tools for automatic editing, editing in the context of multiple sources and methods for analysing the efficiency of editing processes. Papers dealing with editing in the context of multiple sources or the impact on editing and imputation when incorporating data from external sources are particularly welcome.

18. Presentations could report on advancements in developing theory and techniques, empirical comparisons and evaluations of methods. They could also highlight the expected impact that new methods have on the statistical agency including how they contribute to standardizing concepts, terminology, methods, data structures and the quality of its data products.

Topic (iii): Getting the support of all people when implementing data editing

Session Organizers: Li-Chun Zhang, Statistics Norway (lcz@ssb.no) and Philippe Brion, INSEE, France (philippe.brion@insee.fr)

19. Editing procedures are one of the most consuming parts of the statistical process. New methods have been introduced in recent years in many statistical offices, aiming at improving efficiency and reducing time and costs of editing processes while ensuring the quality of the disseminated results, but it seems that there is still work to do to implement these methods in a more systematic way.
20. Different reasons may explain this situation and in particular:
- more support has to be obtained from top-level management;
 - support and cooperation have to be found from the editing staff;
 - the renewed data editing process has to be integrated within the data and metadata system of the statistical office, which requires resources across the organization.
21. To address the first two needs, paradata and metadata, in addition to methodological studies, it may be useful to convince different “non-methodologists” of the efficiency of the methods. In particular, they may show how a continuous improvement may be introduced from the process management point of view. They can also take into account the cost aspect.
22. The need for standardization and integration of the data and metadata system at the national statistical office raises broader technical, infrastructural and even organizational issues that seem to lie beyond the department responsible for data editing in a given statistical domain. Papers about concrete experiences on these topics are welcome.

Topic (iv): Editing of Census and social data

Session Organizers: Daniel Kilchmann, Swiss Federal Statistical Office (Daniel.Kilchmann@bfs.admin.ch) and Felipa Zabala, Statistics New Zealand (felipa.zabala@stats.govt.nz)

23. This topic focuses on the methodological advances of editing and imputation (E&I) techniques applied to census or social data. Most of the variables of Census or social data are categorical variables which are mostly factual but may also be attitudinal. All E&I aspects applied especially to this kind of data may be discussed in this topic.
24. The availability of registers and administrative data has enabled countries to gather census data without surveying the whole population. Other innovations to assist in collecting census data include the use of a Geographic Information System and multi-mode data collection methods. This session is also interested in E&I methodological work related to these aspects.
25. The use of external data sources is a fast-developing field of investigation with the aim of reducing the response burden and accelerating the production process. To process the integration of multiple sources of data, the editing strategy plays a central role in the quality of statistical outputs produced from these sources. Contributions of E&I developments and applications in social surveys with respect to external data sources are warmly welcomed in this topic.
26. E&I of categorical variables or E&I in the presence of a two-stage sample design, which is quite common in social surveys, as well as considerations about the consistency over time of the E&I results of such data may be discussed.
27. Papers discussing the results of applications of E&I for census or social data should describe the methodologies used, with particular attention to the evaluation of the experiences (benefits/risks/open problems) derived during the E&I processing.

Topic (v): International collaboration and processing tools

Session Organizers: Claude Poirier, Statistics Canada (claudio.poirier@statcan.gc.ca), Gergely Horváth, Hungarian Central Statistical Office (gergely.horvath@ksh.hu) and Pedro Revilla, National Statistical Institute of Spain (previlla@ine.es)

28. The development of stronger international collaboration to improve the efficiency of statistical production can help national statistical offices when facing resource constraints. A variety of directions can be followed. Up to now, concepts and methods were generally more portable between organizations than were systems and tools. This explains the current appetite from the statistical community for interoperable “plug & play” tools. A working group has been created by the High-Level Group for the Modernization of Statistical Production and Services (HLG) to propose an architecture that would enable such an interoperability. As they mention in their draft Common Statistical Production Architecture (CSPA), there are two major threats to statistical production: the rigid processes and the aging environment. In reaction to that, organizations began to share and borrow tools – or parts thereof – with each other. Independent and interchangeable modules constitute a viable solution to this vision, as is the open source approach.

29. Papers focusing on portable solutions, on collaboration initiatives targeting software development or reuse of tools, or on the following aspects are welcome:

- Governing data editing open source tools
- Competing goals of local programs against international initiatives
- Modernization using Enterprise Architecture
- Managing strengths, weaknesses, opportunities and threats
- International collaboration to solve practical problems and to share knowledge and solutions
- Identifying national needs, schedules and expectations
- Moving national successes into an international perspective.

VI. Local arrangements

A. Meeting venue

30. The meeting will be held at:

Maison de la mutualité
24 rue Saint-Victor
75005 – PARIS
France

Tel. : +33-1 8392 2400

E-mail : contact@maisondelamutualite.com

B. Travel and Accommodation

31. Participants and/or their offices are requested to make their own travel arrangements and hotel reservations. INSEE will provide a list of recommended hotels in Information Notice No.2.

32. The organizers regret not being able to assist participants in their travel arrangements (air tickets, trains, local transportation, etc.). Neither INSEE nor the UNECE secretariat can take any responsibility for financial arrangements with respect to travel or accommodation.

VII. Further information

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

INSEE:

Mr. Philippe Brion
Head, Business Registers, Infrastructures and
Structural Statistics
INSEE
18, Bd. A. Pinard
75675 PARIS CEDEX 14

Tel: +33-1 4117 6537
e-mail: philippe.brion@insee.fr

UNECE:

Mr. Steven Vale
Head of Statistical Information and Methodology Unit
Statistical Division
United Nations Economic Commission for Europe
Palais des Nations
1211-GENEVA 10
Switzerland

Tel.: +41 22 917-3285
email: steven.vale@unece.org

UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS
Work Session on Statistical Data Editing
 (Paris, France, 28-30 April 2014)

REGISTRATION FORM

This form should be completed by each participant and returned
by 26 February 2014 to:

Ms. Diane Serikoff
 Tel: +4122 917-2241
 E-mail: diane.serikoff@unece.org

1. Surname (Family name):	Mr. <input type="checkbox"/> Ms. <input type="checkbox"/>	2. First name:
3. Representing (country or organization):		
4. Official function (in home country):		
5. Mailing address:		
6. Tel. No.:		
7. E-mail address:		
8. Please indicate topic ((i)-(v)) and title, if you intend to present a paper:		
9. Special requirements:		

Date _____

Signature _____