

**UNITED NATIONS**  
**ECONOMIC COMMISSION FOR EUROPE**

**CONFERENCE OF EUROPEAN STATISTICIANS**

**Work Session on Statistical Data Editing**  
(Budapest, Hungary, 14-16 September 2015)

**INFORMATION NOTICE No.1\***

Hungarian Central Statistical Office  
will host the Work Session at the Danubius Hotel Flamenco, Tas vezér utca 3-7, Budapest, Hungary  
from 14 to 16 September 2015

**I. Purpose of the work session**

1. At its 2014 plenary session, the Conference of European Statisticians included the Work Session on Statistical Data Editing in its 2014-2015 meeting programme. The work session will be held from 14 to 16 September 2015 in Budapest (Hungary) at the kind invitation of the Hungarian Central Statistical Office. 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 and macro-editing
  - (ii) Managing and supporting changes related to editing and imputation
  - (iii) Software tools and international collaboration
  - (iv) Evaluation and feedback
  - (v) Emerging methods and data revolution
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. Information on the venue and local arrangements in Budapest (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 Hungary 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 Hungarian Central Statistical Office (see also section VI of this Information Notice).

7. Statistical offices and international organizations should inform the UNECE secretariat **before 27 April 2015** 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 register **by 6 July 2015** by completing on-line registration form available here: <https://www2.unece.org/uncdb/app/ext/meeting-registration?id=jYodMv>

### 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 ([tetyana.kolomiyets@unece.org](mailto:tetyana.kolomiyets@unece.org)) as soon as possible and **by 27 April 2015 at the latest.**
- The papers, around 10 pages, should be sent electronically to the UNECE secretariat **by 6 July 2015 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 ([tetyana.kolomiyets@unece.org](mailto:tetyana.kolomiyets@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 **31 August 2015 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/2015.09.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 and macro-editing

**Session Organizers:** Pedro Revilla, National Statistical Institute of Spain ([previlla@ine.es](mailto:previlla@ine.es)) and Emmanuel Gros, INSEE, France ([emmanuel.gros@insee.fr](mailto:emmanuel.gros@insee.fr))

14. Selective editing (score function approaches, significance editing, optimization methods) and macro-editing (output editing, top down editing) are selection techniques to identify a limited amount of data with influential errors for which manual review is worthwhile. They could be used before or after data aggregation or to complement the existing micro-editing procedures, and are particularly appropriate in the context of business surveys, where usually a few of the largest units have a decisive influence on final results.

15. These procedures are increasingly used in the data editing process of national statistical institutes (NSIs), since they allow to save time and optimize manual review resources, while maintaining the quality of the statistical data products. All contributions to this topic that help to improve the state-of-the-art in selective and macro-editing are of course welcome. Contributions should especially focus on the following aspects:

- Unification of selective editing approach;
- Selective editing for classification variables;
- Quality assessment of selective editing;
- Complementarity between selective editing – throughout the data collection process – and macro-editing after the data collection process.

### Topic (ii): Managing and supporting changes related to editing and imputation

**Session Organizers:** Marco Di Zio, ISTAT, Italy ([dizio@istat.it](mailto:dizio@istat.it)) and Claude Poirier, Statistics Canada ([claudio.poirier@statcan.gc.ca](mailto:claudio.poirier@statcan.gc.ca))

16. New standardized procedures and methods have been studied in recent years aimed at improving the harmonization of the editing process and at improving efficiency in order to achieve a level of quality in the statistical products that is as high as possible with a level of cost of editing processes that is as low as possible. Although, in general, the process and methodological improvements are clearly analyzed and well established, their introduction in the statistical production process requires additional efforts to ensure changes are thoroughly and smoothly implemented and lasting benefits of change are achieved.

17. To realize this aim, efforts should focus on:

- involving and obtaining support and cooperation from top-level management and the editing staff as early, openly and fully as possible;
- understanding where an NSI wants to be and determining the initiatives to get there;
- planning development work in appropriate, achievable and measurable stages.

18. Getting the support of people requires making stakeholders aware of the importance of the new procedures/methodologies in terms of improvements in the production process as clearly demonstrated using quality indicators. It is of utmost importance that changes to the editing and imputation process are smoothly introduced in the production process to facilitate transition. It is also important to monitor the impact of changes to make valid conclusions on expectations that are met or explain those which are not. Changes would be easier to introduce, manage and accept if the data editing system is standardized and based on an industrialized statistical production process.

19. Contributions with concrete experiences discussing the following are welcome.

- Engaging top-level management and each of the levels below in supporting changes in culture and editing and imputation methods and processes;

- Training and education to support changes in culture and editing and imputation methods and processes;
- Monitoring and interpreting the efficiency of an editing and imputation process with respect to meeting users' needs and minimizing the use of resources.

Contributions may discuss the importance of a standardized data editing procedure for managing and supporting changes.

### **Topic (iii): Software tools and international collaboration**

**Session Organizers:** Rudi Seljak, Statistical Office of Slovenia ([rudi.seljak@gov.si](mailto:rudi.seljak@gov.si)) and Claude Poirier, Statistics Canada ([claudio.poirier@statcan.gc.ca](mailto:claudio.poirier@statcan.gc.ca))

20. Usage of suitable software tools is inevitable when designing and implementing modern data editing strategies. Development of such tools can be a complex, long-term and very costly job. For smaller organisations, such development often exceeds their capability and budget constraints. The principle of sharing seems to be a natural solution in such cases. Although the statistical community has shared knowledge and established concepts, methods and best practices, there is still a lack of successful collaboration in the field of exchange of operational ready-to-use software tools. To some extent, the same is also true in the field of standardisation and harmonisation of data editing strategies at the global architecture level. If the development of the GSBPM and GSIM models have certainly been a success story on the concept level, their implementation is still an ongoing activity and a big challenge for the future developments.

21. Papers focusing on the following issues are especially welcome under this topic:

- Open source tools for data editing;
- Sharable data editing tools developed for national purposes;
- Building bridges between local and global solutions;
- International initiatives and related activities;
- Implementation of the Generic Statistical Information Model (GSIM);
- Modernisation and harmonisation by using Enterprise Architecture;
- Applications of the Common Statistical Production Architecture (CSPA).

### **Topic (iv): Evaluation and feedback**

**Session Organizers:** Daniel Kilchmann, Swiss Federal Statistical Office ([Daniel.Kilchmann@bfs.admin.ch](mailto:Daniel.Kilchmann@bfs.admin.ch)) and Sander Scholtus, Statistics Netherlands ([s.scholtus@cbs.nl](mailto:s.scholtus@cbs.nl))

22. During the last decades editing and imputation (E&I) strategies have been developed and integrated in most statistical agencies for most of their statistical processes with the aim of reducing costs and of managing the resources efficiently by concentrating expensive treatments to critical units.

23. More generally, the development and redesigning of data preparation processes were often driven by the following main objectives of E&I (EDIMBUS, 2007):

1. identify error sources in order to provide information for future improvements of the survey process;
2. provide information about the quality of the incoming/outgoing data;
3. identify and treat the most significant errors;
4. when needed, provide complete and consistent (coherent) individual data.

24. The main focus of this topic is on indicators – including paradata and metadata of any kind – and techniques which allow statistical agencies to evaluate whether and to which degree the objectives above are met. Of special interest in this topic are:

- methods for using the information from E&I to improve the survey process;
- methods for estimating the uncertainty due to E&I;

- different quality indicators for different groups of users (e.g., data editors, statistical analysts, process managers, methodologists).

25. Contributions discussing the methods and their implementation as well as the results from their application are warmly welcomed. Results covering the whole data preparation process as well as results covering individual steps of that process for any survey may be discussed.

### **Topic (v): Emerging methods and data revolution**

**Session Organizers:** Jeroen Pannekoek, Statistics Netherlands ([jpnk@cbs.nl](mailto:jpnk@cbs.nl)) and Li Chun Zhang, Statistics Norway ([lcz@ssb.no](mailto:lcz@ssb.no))

26. This topic covers new and emerging 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 including the ones mentioned above.

27. More and more often, sample surveys are replaced by other sources, or used in combination with them. Therefore, of special interest are methods for the detection and correction of errors and imputation in these alternative or new data sources, including register and Big Data, either on their own or in combination with survey data.

28. Presentations could report on advancements in developing theory and techniques, empirical comparisons and evaluations of methods. They also could 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.

## **VI. Local arrangements**

### **A. Meeting venue**

29. The meeting will be held at:

#### **DANUBIUS HOTEL FLAMENCO**

Ravel Conference Room

H-1113 Budapest, Hungary

Tas vezér utca 3-7.

[www.danubiushotels.com/flamenco](http://www.danubiushotels.com/flamenco)

### **B. Travel and Accommodation**

30. Participants and/or their offices are requested to make their own travel arrangements and hotel reservations. Information about the special room rates for the Danubius Hotel Flamenco will become available in the Information Notice No. 2.

31. A letter to facilitate obtaining a visa can be requested from the Hungarian Central Statistical Office. Please contact **Ms. Eszter Békés** ([eszter.bekes@ksh.hu](mailto:eszter.bekes@ksh.hu)) in case you need a special invitation letter to obtain a visa. Please note that, since the meeting is organised by the UN, the visas should be issued free of charge.

## VII. Further information

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

### HCSO:

Ms. Agnes Andics  
Methodologist  
Hungarian Central Statistical Office  
Methodology Department  
1024 Budapest, Keleti Károly u. 5-7.  
Hungary

Tel: +36-1 345-6643  
e-mail: [agnes.andics@ksh.hu](mailto:agnes.andics@ksh.hu)

Ms. Tamara Pál  
Head of the  
International Relations Section  
1024 Budapest, Keleti Károly u. 5-7.  
Hungary

Tel: +36-1 345-6018  
e-mail: [tamara.pal@ksh.hu](mailto:tamara.pal@ksh.hu)

### UNECE:

Ms. Tetyana Kolomiyets  
Statistical Information and Methodology Unit  
Statistical Division  
United Nations Economic Commission for Europe  
Palais des Nations  
1211-GENEVA 10  
Switzerland

Tel.: +41 22 917-4150  
email: [tetyana.kolomiyets@unece.org](mailto:tetyana.kolomiyets@unece.org)