

CONFERENCE OF EUROPEAN STATISTICIANS

UNECE Work Session on Statistical Data Editing

(27 – 29 May 2002, Helsinki, Finland)

TOPIC (IV): IMPACT OF NEW TECHNOLOGIES ON STATISTICAL DATA EDITING

Issues for discussion

Discussant: William E. Winkler (US Bureau of the Census)

Invited Papers:

- The performance of the Canadian Census Edit and Imputation System (CANCEIS) (Michael Bankier, Statistics Canada)
- Towards a new system for edit and imputation of the 2001 Italian Population Census data: A comparison with the Canadian Nearest-neighbour Imputation Methodology (Antonia Manzari and Alessandra Reale ISTAT, Italy)
- UniEdit: Standardized processing of structural business statistics in the Netherlands (Arjan De Jong, Statistics Netherlands)

From the invitation letter: Presentations of new methods and techniques should be accompanied with an evaluation of operational feasibility and the organizational issues associated with their implementation. The contributions could also describe on which kind of data and under which circumstances the new method performs best. It would be desirable to distinguish between those methods in the stage of theoretical development, those in the stage of test or pilot projects, and those that are implemented in practice.

This is a broad topic. Many ideas are acceptable. The importance is to describe the ideas in a manner that relates them to other existing ideas in statistical data editing and imputation.

1. **Is the method being described a new concept, a novel extension of existing ideas, or an application of concepts from some other field?** How does the method relate to other concepts in statistical data editing and imputation? Are the ideas primarily theoretical or are they practical in the sense of having been implemented?
2. **How practical are the ideas?** Have they been implemented in a research system or in a production system? To what types of data (discrete, continuous, combination of discrete and continuous) are the ideas applicable?
3. **Do the concepts rely on expert knowledge?** Some methods require the use of experts to develop and train other individuals in their use. Are the methods being described those that can be easily understood by potential appliers and users of the technology? Will the potential users need extensive training?
4. **Do the concepts rely on training data?** How were the training data created? How representative is the training data of representative situations in which statistical data editing and imputation are applied?
5. **How was testing performed?** What quantification of the performance of the new methods has been done?
