

Work Session on Statistical Data Editing
(Prague, Czech Republic, 14 to 17 October)

Item 3 of the provisional agenda

NATIONAL REPORT

Submitted by the Lithuanian Department of Statistics ¹

¹ Prepared by Vygandas Norkus.

I. BRIEF DESCRIPTION OF PROJECTS

1. Lithuanian Department of Statistics (LDoS) has two projects currently under development:

- a) Statistical Information System (SIS);
- b) Population and Housing Census project.

2. The goal of the **statistical information system** in LDoS is to introduce unified and coherent forms, methods and tools for data capture, storage, retrieval, processing and dissemination. The project includes the development of:

- a) register of statistical indicators;
- b) integrated metadata information system;
- c) database of statistical indicators;
- d) computerized work places for users with free access to databases.

Significant attention in this project is also paid to data security, integration of statistical registers and integration into the global information system of government institutions. The project was begun in 1995. Implementation of the register of indicators was started in 1996.

3. **The population and housing census** is planned to be carried out in 1999. The project consists of the following general stages:

- a) preparation of preprinted questionnaires using administrative sources;
- b) automatic coding;
- c) data entry with OCR (Optical Character Recognition) technology
- d) data editing;
- e) data dissemination using GIS (Geographical Information System).

4. The project was begun in February 1997. The pilot population census is planned to be carried out in October 1997. Questionnaire design, preprinting of questionnaires, data editing and automatic coding are implemented, taking into account the use of OCR technology. Generalized software packages will be used for data processing.

5. LDoS does not use any generalized software for data editing. For each survey a separate tailor-made data editing and tabulation software is developed. The main reasons for this costly and time-consuming activity are differences in file formats and structures. Furthermore, as the data checking is incorporated into the data entry stage, the use of different software for data entry makes it impossible to introduce standard software for data editing.

II. DATA EDITING METHODS AND TECHNIQUES USED

6. The SIS project plans to overcome these obstacles and to integrate the statistical processing on the same type of computer system, using standard software and metadata base. The project also covers the development of a generalized software package for data editing.

III. HARDWARE AND SOFTWARE USED

7. Servers

Name	CPU	RAM	HDD
ICL Super Server K524S	4*75 MHz SuperSparc	256 MB	40 GB RAID
ICL Super Server H824S	4*150 MHz SuperSparc	512 MB	40 GB RAID
Novell NetWare Server	2*100 MHz Pentium Pro	128 MB	16 GB RAID
Novell NetWare Server	2*166 MHz Pentium Pro	128 MB	16 GB RAID

8. Personal computers

	CPU	RAM	HDD
	75-100 MHz Pentium	8-32 MB	500-1000 MB
GB	166-200 MHz Pentium Pro	16-64 MB	820 MB - 4

9. Server software

Server name	Operating system	DBMS	Other software
Application servers	Unix NX (SVR 4.2)	Oracle 7.3.2 server	
File servers	Novell NetWare		
Internet server	Linux		Mail, FTP, Names server, Firewall

10. Software used on personal computers

Operating system	Data entry and editing	Data processing and analysis	Publication
Windows 95, Windows NT	Developer/2000, MS Office, FoxPro for Windows	Developer/2000, MS Office, Discoverer/2000, FoxPro for Windows, SAS	MS Office, Developer/2000

IV. PERSONNEL AND FINANCIAL IMPLICATIONS

11. The SIS development is a costly and labour consuming process covering the implementation of modern information technologies, methods and tools of statistical surveying and analysis. The project development group consists of IT specialists and subject-matter area specialists from each LDoS section. Two foreign companies are also involved. The contracts with "World Systems Ltd" (from Luxembourg) and "Partnership UK Ltd" (from United Kingdom) were signed for 24 months. The Steering Committee responsible for coordination of project development has been formed.

12. The SIS project is financed from PHARE funds. 2.25 MECU was invested during 1992-1996. Tenders for: (1) Computer and Software Equipment; (2) Office Equipment; (3) Technical Assistance and Training; (4) Classifications adoption have been organized. The Population and Housing project is financed from the state budget and PHARE funds.

V. PLANS FOR FUTURE DEVELOPMENT

13. The main task for the near future is the further development of the two projects mentioned above. The strategic plan of development for the next two years has already been prepared. It includes: (1) the implementation of a register of indicators and knowledge base; (2) introducing international classifications in statistical data coding and processing; (3) transferring statistical data into Oracle environment; (4) implementation of optical reading of statistical questionnaires; (5) development of generalized software packages for data editing, processing and dissemination.