

8. POPULATION CENSUS DATA PROCESSING

Certain conditions, such as the sufficient number of rooms, tables, chairs, computers, heating equipment, stationery, etc. were necessary for the implementation of processing of the Population Census data. Given that, still in summer of 2009 the Population Census Department of NSS RA was provided with additional rooms for the employees, also for the storage of more than 12 000 portfolios, as well as for the implementation of Population Census data processing works.

120 metal boxes for the storage of portfolios (each with the capacity to store 100 portfolios), as well as desks, chairs, high quality computers, printers, copy machines and heaters have been purchased.



8.1 POPULATION CENSUS DATA CODING PROCESS

125 temporary employees (coders) were hired on a competitive basis to implement the Population Census coding process. During the competition the persons with higher education and students, as well as those who participated in the previous phases of the Population Census were privileged. Taking into account the abilities of person, in some cases, the persons with preliminary professional education were also hired. The table presented below is characterizing the sex and age composition and educational level of coders.

Sex, age	Total	Level of education		Students
		Higher	Secondary professional	
Total	125	100	4	21
19-29	49	27	1	21
30-39	25	25	0	0
40-49	27	26	1	0
50-59	20	19	1	0
60-69	4	3	1	0
Male	30	23	1	6
19-29	20	13	1	6
30-39	3	3	0	0
40-49	3	3	0	0
50-59	3	3	0	0
60-69	1	1	0	0
Female	95	77	3	15
19-29	29	14	0	15
30-39	22	22	0	0
40-49	24	23	1	0
50-59	17	16	1	0
60-69	3	2	1	0



The training of the coders was carried out from 30-31 November. On 1 December they were divided by turns, got information on the working rooms and places, after that they were provided with the necessary classifications and manuals. The coding works started on 3 December. A week later 30 persons were separated from the coders (this number could change each month), who were assigned to check the work done by coders.

The coders and controllers had a special enumeration sheets (diaries) where they took notes on the numbers of portfolios checked/coded during the day, the name/surname of coder/controller, the turn, month and date of receipt and handing over.

It should also be mentioned that, it was not prohibit to the controller from time to time to go back to coding. In this case, new controllers were chosen from coders.

The coding/checking process was implemented in two turn's working days within 2 working rooms foreseen for coding. The staff of Population Census Department worked also in two turns working regime. Their duties included:

- to get from the storage room the portfolios to be registered during the day by noting their number in the corresponding registration book and to arrange them in the special separated place in the corresponding working room,
- to follow that the coders would mark the coded portfolios with the certain color (yellow) and would arrange them in the place foreseen for the coded portfolios,
- to follow that the controllers would mark the checked portfolios with another color (blue) and would arrange them in the special place for taking back to the storage room,
- after the end of each turn the abovementioned diaries were collected from the controllers,
- to enter the information of checked portfolios from the diaries into the computer, to give back the diaries to the controllers the next day,
- to implement the control of coders' attendance and of ensuring proper working atmosphere in the rooms.

The control of the portfolios movement from the storage rooms to the working



rooms and vice versa was carried out by special documentations.

The entry of everyday information received from the controllers' diaries in the computer has gradually formed the database that allowed implementing additional control over the flow of portfolios. In addition, the mentioned database also allowed to be informed at any time at what stage of the process was the portfolio, who has coded (was coding) or has checked (was checking) it. Through this database it was also possible to control the entire coding process on time and to follow the defined terms. The problems requiring quick solutions have been occurring when implementing the coding process. One of them, as during the coding of 2001 RA Population Census data referred to the question 27 (the type of economic activity in brief). For this time also the significant part of enumerators has filled in this question by abbreviations (in contrary to the requirement of instruction to complete the Population Census questionnaire and to the department staff warnings). However in difference with the previous coding process, for this time the skilled enumerators have used the Internet. Several enumerators have been chosen to find out the responses to the mentioned questions and the searches mainly have been giving positive results.

The coding process lasted 6 months and was over on 1 June 2012.



8.2. POPULATION CENSUS QUESTIONNAIRE DATA ENTRY AND DATABASE CREATION OPERATIONS

110 operators of various ages and professions were hired for the Population Census questionnaire data entry. They were selected by testing results.

The data entry was carried out by two turns in the terms of 6 months from 1 February 2012 to 1 August 2012.

The data entry was implemented in two working rooms. The best two operators from each room were selected to control the data entry work.

The data entry was arranged via the network. The geographical codes of the questionnaires and the codes' created file have been sending via the network to the computer of the operator to enter the data of the questionnaire of the given portfolio.

The entered data were immediately accumulated in the relevant folders of the server. The copies of the folders also have been created for security purpose.

The controllers (verifiers) were also selected from operators, who have undergone the 70 % of portfolios to double entry. The most important fields of the questionnaire were selected for double entry.

After completing the entry works the files of appropriate folders were united creating the initial database. This database underwent to structural checking, i.e. it was checked if the entered data are consistent with the data of portfolios registration file.

After those works the information base was ready to go to the phase of logical control.



8.3. POPULATION CENSUS DATABASE AUTOMATED EDITING

In parallel with the RA 2011 Population Census data entry, the testing of automated editing program of database logical errors has been started. At that phase, when the data for some marz have being completed, the additional tabular method was used to edit errors. The tables were derived by crossing different variables. The appropriate corrections were made when finding out the inconsistencies between those variables. At that phase the Informational Technologies and Programming Division has been closely cooperating with the Data Processing and Classification Division.

When, after the completion of entry work the database was created and that database was cleaned of structural errors, the program of automated editing was used for that base, cleaning it of errors and the cleaned base was saved in the new file.

However, until then the editing program underwent some changes, since some errors had been found in the base that had not been taken into account in the program. The new commands for that data detection and correction were added to the program.

The program of automated editing is comprised of three parts. The first part corrects the data regarding 1 to 9 questions of the questionnaire, that are the serial number of the person, the category of residence, the duration of absence or temporary presence, location and cause, the relationship of a person with the person written the first, the mother's (father) serial number, gender, day, month, year of birth and age. The base cleaned up in the result of this program was considered to be input base for the errors correction's second program. The second part of the program refers to the 10 to 33 questions of the questionnaire. The third part of the program corrects the data on household. The base cleaned up from all the errors has been obtained as a result of the third program errors correction.

