



Economic and Social Council

Distr.: General
1 September 2014

Original: Russian/English

Economic Commission for Europe

Conference of European Statisticians

Group of Experts on Population and Housing Censuses organized jointly with Eurostat

Sixteenth Meeting

Geneva, 23-26 September 2014

Item 5 of the provisional agenda

Operational aspects of censuses

Publication on the Internet of non-personal micro databases of All-Russia Population Censuses of 2002 and 2010

Note by the Federal State Statistics Service (Russian Federation)

1. In 2012 the Federal State Statistics Service provided a full and unlimited access to the micro database of the All-Russia population census of 2002 for Internet users. The database contains the information concerning all private households in Russia. Since 2013 the micro database of All-Russia Population Census of 2010 on the population of private households has become available. It contains the data of the All-Russia Population Census of 2010, as well as the data of the All-Russia Population Census 2002 as to comparable indicators. To gain access to these micro databases the software package SuperSTAR (its developer - Space-Time Research) is applied.

SuperSTAR is the addition to the standard method of publication of the results of population censuses and surveys in tabular form in the format of pdf, Word and Excel. SuperSTAR functionality allows its users:

- to work out a discretionary table layout regarding any measures of the program of the All-Russia Population Census of 2010;
- to form the results of the census form concerning any territorial units, including human settlements;
- to generate data tables of the All-Russia Population Census of 2010 in comparison with the results of the All-Russia Population Census of 2002;
- to form inquiries with overlapping features of various objects (e.g. households and dwellings, individuals and households, etc.);
- to provide the data in the form of tables and graphs;

- to export the generated table in MS Excel, xml or csv formats; and diagrams in png or pdf.

The advantages of software system SuperSTAR are:

- The system fast response;
- The possibility to use a single hardware and software system for dealing with several databases, such as population census, agricultural census, sample survey of household budgets, population survey of employment problems, a sample survey of reproductive plans of population, etc .;
- An access to software modules is performed either directly or remotely through web access;
- A software package is designed to be used by specialists of statistical authorities and external users via the Internet;
- An intuitive graphical interface does not require programming skills in constructing database inquiries;
- While dealing with SuperSTAR on the Internet, there is no need to install additional software, since it is necessary to have only a standard web browser - Internet Explorer;
- SuperSTAR allows us to download information from databases of any format.

However, for dealing with the software package it is necessary to know the programs and methodologies of statistical surveys. Otherwise, it will not be able to form an inquiry containing the correct information.

To train users to work with databases the official website of Rosstat provides detailed methodological explanations of the All-Russia Population Censuses of 2002 and 2010, as well as a video showing an example how to construct a database inquiry.

In addition, the Federal State Statistics Service conducted workshops with users of statistical information how to apply the results of the All-Russia Population Census of 2010. Everyone could participate in the seminar. During the seminar the operation of the program complex SuperSTAR was also demonstrated.

If users have any questions as to the search of information, or how to deal with databases, Rosstat consults on the phone and by e-mail.

The formation of tables with the results of the census occurs rapidly within a few minutes, despite the large volume and complex structure of a micro database (the information concerning more than 300 million people with 250 attributes in the form of complex tables as to the premises, households, family units and persons). Such rapid work of the program SuperSTAR is achieved due to a special storage format.

While providing an access to a micro database, it is paramount to ensure the confidentiality of personal information. Despite the fact that the primary information in the database is impersonal, it is still possible to identify a particular respondent by unique features of a surveyed group of persons. Such characteristic may be the date of birth or the nationality of inhabitants that is non-relevant to the location, etc. First of all, the risk of disclosure of personal data is high for residents of settlements with the small amount of population.

For the protection of personal data in conjunction with the Federal State Statistics Service Space-Time Research has developed an algorithm to protect personal data.

It is an algorithm of distortion (perturbation) of data that changes figures slightly in the table. This makes impossible to obtain information about a specific respondent or household. The use of this algorithm provides the representation of data in the absence of

access restrictions as to census results by aggregating the values of indicators or territorial units.

In the tables with modifications to hide confidential data, there is a warning message concerning introduced distortions.

Table 1. An example of a table formed without the application of algorithm of personal data protection (conventional data)

	All population	0 - 9	10 - 19	20 - 29	30 - 39	40 - 49	50 - 59	60 - 69	70 and over	The age is not specified
Bronnitsy	191	6	12	67	38	48	10	6	4	-
Dzerzhinsky	81	4	10	18	22	10	13	-	1	3
Dolgoprudny	796	23	47	256	260	141	31	6	6	26
Dubna	71	3	10	9	6	17	9	5	7	5
Zheleznodorozhny	86	4	12	16	13	16	8	4	3	10
Zhukovski	3712	35	105	985	1053	884	304	105	162	79
Ivanteevka	242	5	6	91	91	29	6	2	-	12
Klimovsk	93	7	2	15	37	23	2	4	2	1
Kolomna	115	10	6	35	32	12	8	6	-	6
Krasnoarmeysk	19	-	1	8	6	4	-	-	-	-

Table 2. An example of a table formed with the application of the algorithm of personal data protection (conventional data)

	All population	0 - 9	10 - 19	20 - 29	30 - 39	40 - 49	50 - 59	60 - 69	70 and over	The age is not specified
Bronnitsy	188	4	12	67	36	47	9	9	4	-
Dzerzhinsky	81	5	9	15	20	10	14	-	-	-
Dolgoprudny	797	21	47	260	257	139	33	5	7	30
Dubna	71	3	10	8	7	15	11	7	5	3
Zheleznodorozhny	84	9	13	15	15	13	6	5	4	10
Zhukovski	3713	36	108	988	1052	886	300	104	160	81
Ivanteevka	242	4	4	91	90	27	3	-	-	8
Klimovsk	92	9	6	13	38	22	3	3	-	-
Kolomna	112	11	9	32	37	13	9	7	-	7
Krasnoarmeysk	20	-	3	8	5	6	-	-	-	-

Any Internet user can get an access to the micro databases of the All-Russia Population Censuses of 2002 and 2010.

Micro database contains indicators as to all aspects of the population census program. In addition, some indicators characterizing households and family units were added.

During the operation of micro data bases of the All-Russia Population Census of 2002, the Federal State Statistics Service got some proposals from users to add to the database certain additional estimates. In connection with it, the indicators of family units that characterize his wife (the mother) and her husband (the father) were added to the database

of the All-Russia Population Census of 2010. This allows users to work out tables that are required to study married couples. An example can be the generation of the data characterizing the ethnic homogeneity of couples, educational level, economic activity, sources of livelihood, the place of birth, nationality, languages spoken and age composition of spouses. This function has become very popular among demographers as it was the first time when they managed to analyze the main indicators characterizing spouses. An example of table «A Husband's and Wife's Ages" is shown in Table 3.

Table 3 "The Age of a husband and the age of a wife in the Belgorod Region"

Territorial unit	Belgorod Region												
	18	19	20	21	22	23	24	25	26	27	28	29	30
Number of complete years of a wife													
Number of complete years of a husband													
18	24	37	4	8	7	6	6	0	4	0	7	0	0
19	84	91	66	42	25	20	24	8	10	17	3	0	0
20	150	233	253	153	95	63	50	27	18	21	9	13	7
21	191	246	414	501	287	209	156	93	76	52	23	30	27
22	139	371	611	737	938	561	383	224	118	84	99	43	34
23	153	340	739	1039	1228	1388	832	443	325	258	171	114	123
24	154	364	624	973	1411	1690	1716	955	595	418	214	139	117
25	86	234	476	928	1390	1824	1870	1977	1219	744	361	272	211
26	71	149	386	735	1131	1682	2282	2224	2047	1233	803	483	358
27	49	119	332	660	1065	1692	1949	2228	2397	2386	1309	789	580
28	29	101	200	395	745	1194	1758	1999	2417	2556	2102	1253	836
29	29	74	134	270	529	870	1258	1646	2164	2390	2528	2192	1265
30	14	51	114	267	421	664	1037	1557	1898	2482	2642	2475	2500

Currently Rosstat continues to improve the Internet users' operation with databases. In 2013, apart from the database of the All-Russia Population Census of 2010, the access to databases of sample survey of reproductive plans of population sample survey of behavioral factors influencing population health status. Furthermore, in 2014 it is planned to implement authorized access to the micro database. This will allow us to expand the functionality of the software for registered users.