

SURVEY ON USE OF INFORMATION AND COMMUNICATION
TECHNOLOGY (HOUSEHOLD SURVEY)

Statistics of Living Standard Sector

Last update: June 2019

TEMPLATE FOR DOCUMENTATION OF STATISTICAL PROCESSES GSBPM

1. Identify needs

1.1 Identify needs

Survey on use of information and communication technology (ICT) is a statistical survey conducted with households and provides an overview of the use of computers and the internet for personal purposes in our society. ICT statistics are based on:

- Law no. 04 / L-036 on Official Statistics of the Republic of Kosovo,
<http://ask.rks-gov.net/media/2023/ligji-per-statistikat-zyrtare-te-republikes-se-kosoves.pdf>

The data collected by the ICT also provide information on access to different electronic devices, the type of internet connection and the barriers using the Internet; information about computer usage, internet usage, e-commerce and e-skills.

1.2 Consultation and confirmation of needs

Each year, the working group reviews the set of questionnaires in order to reflect the requirements that may have been related to the contents of the questionnaires of ICT.

The questionnaires' reviews are as follows:

- Review of Household Questionnaire
- Review of the individual questionnaire

Based on the planning of the sampling frame, the financial resources, materials and needs are confirmed for:

- Enumerators / interviewers
- Supervisors
- Data entry operators

1.3 Check data availability

Data collection is done through:

- Completing the household questionnaire and
- Completing the questionnaire for an individual living in that household that is aged 16-74 years old.

1.4 Design draft project

At the end of the each year (December), the activity plan for each stage / phase of the observation / survey is prepared according to the GSBPM. This plan contains information about:

- Responsible persons;
- Responsible Sector;

TEMPLATE FOR DOCUMENTATION OF STATISTICAL PROCESSES GSBPM

- Responsible Department;
- The start and end date of the process;
- The risk of non-realization of the process;
- Estimation of costs and benefits as well as external constraints (e.g. projections for supporting staff).

2. Design

2.1 Design output

The ICT output is documented in the quality report published on the KAS website, referred to the link: http://ask.rks-gov.net/media/3834/esqrs_ict.pdf (Quality Report).

2.2 Description of variables

The ICT covers all variables based on EU.

2.3 Design of the data collection

In this process are located all the templates used in the process of data collection, such as:

- PAPI data collection method (questionnaires)
- Interviewers Manual
- Template of non-responses
- Authorization and ID Card for interviewers
- Distribution in accordance with regions and questionnaires and enumerators

If any of the documents will be updated or modified, the decision is taken with the approval of the working group established for this purpose.

2.4 Sampling frame and design

The summarized methodology has defined the methodological rules for determined the size, distribution, sampling, and procedures to assess the outcome.

The sample framework was based on data and mapping from the 2011 Census. For the census counting purposes, Kosovo was divided into enumeration areas (EA), which are relatively small operational segments set for counting registration. A total of 4626 EAs were designated for Kosovo and these were used as the primary sampling unit (PSU) selected in the first sampling phase for ICT of 750 enumeration area.

2.5 Design processing and analysis

At this stage, routines are defined for coding, editing, imputation, data estimation, and estimation and finalizing of datasets.

In the process of estimation and finalization of datasets, the following processes are followed:

TEMPLATE FOR DOCUMENTATION OF STATISTICAL PROCESSES GSBPM

- Check and (merge) data for preparing the final file;
- Consult the indicators that will be published in advance;
- Logic management and finding possible errors.

2.6 Design of production systems and workflow

At the end of each year (December) the activity plan and work charts are prepared for each survey phase according to GSBPM.

3. Build

3.1 Build data collection instruments

ICT data collection tools:

- ICT Household Questionnaire;
- Individual questionnaire for a person aged 16-74;

The IT Department reflects all changes to the data entry program, since the method of data collection is implemented by the PAPI method - with a paper questionnaire.

The application has defined the logical conditions and code of the data collection program in CsPro.

3.2 Build and test IT support for data processing

It includes the activities on building and testing IT support needed for different activities of data processing and analysis, described in sub-process 2.5 “Design processing and analysis“. It includes building and testing informatics solutions for data integration, for manual and automatic data modification after editing, for the calculation of weights, aggregates of tables, etc.

3.3 Build dissemination components

The materials for publication include information about:

- access to information and communication technology,
- use of computers,
- use of the Internet,
- the use of the E-Government,
- use of purchases or trade,
- technological or computer skills.

ICT data is not made available at the micro level as a result of maintaining confidentiality.

Aggregated data is the only type of data provided to external users.

Although micro data is not published, users do not have direct access to micro data, but researchers and institutions may require access by Law.

TEMPLATE FOR DOCUMENTATION OF STATISTICAL PROCESSES GSBPM

3.4 Configure workflow

Materials and design of the work plan for HBS are set out in point 2.6.

3.5 Test statistical process

This sub-process includes conducting test or pilot survey, usually on a smaller sample, to test the instruments of data collection, processing and analysis, etc. in order to check the expected statistical processes in terms of its organization and expected outputs. Pilot surveys may be repeated more than once (especially in the case of the censuses) until the correct statistical process is ensured.

3.6 Finalization of the statistical output system

It includes producing technical documentation for computer support for the collection (developed in sub-process 3.1), processing (developed in sub-process 3.2) and dissemination (developed in sub-process 3.3) of data.
Produce of the documentation related to the components of the process including technical documentation and user manuals. One of the materials needed at this stage is the Supervisors Manual which will be used in the training of the enumerators.
The Manual provides information's and instructions needed to conduct the interview process.

4. Collect

4.1 Create a frame and select sample

The sample frame is based on data and mapping from the Census of 2011. For counting and enumeration purposes, Kosovo has been divided into Enumeration Areas (EAs), which are relatively small operational segments set for counting during the Census.

In total there are 4626 EAs in the entire territory of Kosovo and these were used as primary sampling unit (PSU). The sampling of the Household Budget Survey is stratified in 7 regions, (urban / rural) and the sample is selected randomly selected.

- The sampling selection for ICT is done in two phases: In the first phase of the sample, 750 EAs in 7 regions were selected systematically (PPS - probability proportional to size) within each stratum (region, urban / rural).
- 5 households were selected in the second phase for a total sampling size of 3750 households per year.

4.2 Set up data collection

Building a data collection process involves the preparation for data collection which includes process planning, preparation for their monitoring, development and testing of the program, and training of enumerators / controllers and data entry operators involved in the process.

TEMPLATE FOR DOCUMENTATION OF STATISTICAL PROCESSES GSBPM

During the testing process of the questionnaire, the Sector pays attention how to fill in or complete questionnaires or special sections where attention or care shall be taken. For this reason, the attached material has been prepared “INSTRUCTIONS ON HOW TO COMPLETE QUESTIONNAIRES”.

4.3 Data collection

Data collection is performed by KAS throughout the year from 1 January to 31 December. The households, that has been selected to be surveyed, is firstly explained and presented to the purpose of the survey, the questions for which the household will provide information, and ensures that the individual data collected by the ICT will be handled by maintaining confidentiality, therefore the data will be used exclusively for statistical purposes and will be published only as aggregated data.

The enumerator should help households during the 14 days to complete a diary for food, beverages and goods, through visits and phone calls. The enumerator must provide his or her telephone number to the household so that the household can contact him / her if necessary. In the case of refusal, the enumerator should insist on finding a reliable and convincing way to conduct the interview. If after three times of persistence the interview results with rejection, the enumerator should indicate the reason for the refusal.

In case the apartment (house) is not occupied (is empty) or the household is not found, the enumerator should try two more times to contact the household at the address given and after these three visits failed then the enumerator shall report and notify the person responsible for the survey or the survey manager.

4.4 Finalise data collection

The logical control of the questionnaire is done by the Living Standard Sector.

Data entry continues in the data entry center. The data entry was realized through the CsPro program. The process of data entry also integrates logical conditions for verifying the data entered, as: logical corrections and verification of coherence between the different parts of the questionnaire.

5. Process

5.1 Data integration

Data integration is done after the data entry process is completed.

5.2 Classify and code

At the same time, during the data entry process, data is coded according to certain activities.

TEMPLATE FOR DOCUMENTATION OF STATISTICAL PROCESSES GSBPM

5.3 Review and validate

To review, evaluate the ICT data, the below steps are followed:

1. Data quality analysis

- Logical control of survey data;
- Comparison of time series data;
- Comparison of data with other data sources available at KAS.

2. Handling of non-responses

Non-responses are considered:

- No contact;
- Full refusal;
- Partial refusals (for variables or specific indicators).

5.4 Edit and Impute

No imputation

5.5 Derive new variables and units

The material provides information about the syntax for “Deriving new variables” in SPSS.

5.6 Calculate weights

In order that estimates from the ICT data to be representative of the population, sampling weights should be applied. The basic weight for each household in the sampling equals the inversion of probability of its selection (calculated by multiplying the probabilities at each sampling stage). The probabilities of the selection are based on a stratified two-phase sampling design. In the first phase is selected a sample of the Enumeration Area (EA) with proportional probability proportional to size (PPS) within each stratum (region, urban / rural) and in the second phase is selected a sample of 8 households in each EA of the sample. Based on this sampling design, the probabilities of selection for households in each EA in the sample may be as follows:

$$p_{hi} = \frac{n_h \times M_{hi}}{M_h} \times \frac{m_{hi}}{M_{hi}},$$

n_h = number of selected EAs sample in strata h for ICT

M_{hi} = the total number of households in the sampling frame of the EA sample in the layer h

M_h = the total number of households in the sampling frame for the strata h (that is the cumulative mass of the stratum size)

m_{hi} = number of sampling households selected in EA sample in stratum h.

The basic sampling weight is calculated as the inverse of this probability of selection. Based on the previous expression of probability, the weight can be calculated as follows:

TEMPLATE FOR DOCUMENTATION OF STATISTICAL PROCESSES GSBPM

$$W_{hi} = \frac{M_h}{n_h \cdot m_{hi}},$$

where:

W_{hi} = the basic sample weight of households in the EA sample in the strata h

It is important to match the weights for households in the sample to take into account the household's non-response in each EA of the sample. Once the weights are calculated at the level of EA of the sample, it is useful to adapt the weights at this level. The final weight (W'_{hi}) for the sample households in sample in the strata h can be expressed as follows:

$$W'_{hi} = W_{hi} \cdot \frac{n_h}{n'_h},$$

where:

n'_h = number of sample EA with interviews completed in strata h for ICT.

From 2013 to 2018, final weights are estimated by using the adaptation factor.

For ICT of the year 2018, weight are calculated as follows:

$$A_{HBS} = \frac{H_{2016}}{H_{2017}},$$

where

H_{2017} = the total number of households in Kosovo since last year (2017)

$H_{2018} = \sum_h \sum_i W'_{hi} \cdot n_{hi}$, = weighted estimates of households in Kosovo from ICT data of 2018 adjusted for non-response;

= the number of households in the EA sample in layer h in ICT data of 2018.

The final weights of ICT 2018 are calculated by multiplying the weights, adjusted for non-response, with this adaptation factor for households, as follows:

$$W''_{hi} = W'_{hi} \cdot A_{ICT},$$

where:

W''_{hi} = weighted final weight for households sample in EA sample in layer h.

5.7 Calculate Aggregates data

Aggregated ICT data contains information about all the variables that need to be considered for the ICT survey to make the country-level estimation.

5.8 Finalise data files

Preparing the final files follows the same procedure as explained in step 5.7. The final files of each questionnaire are stored on the server.

TEMPLATE FOR DOCUMENTATION OF STATISTICAL PROCESSES GSBPM

6. Analysis

6.1 Preparing the preliminary outputs

After preparing the final files, the process for preparing the results of the ICT according to the format set for publication begins, which contains information about:

- access to information and communication technology,
- use of computers,
- use of the Internet,
- the use of the E-Government,
- use of purchases or trade,
- technological or computer skills.

6.2 Data Validation

Regular logical checks are carried out with information from other KAS sources. In the output estimation, the internal consistency of the data is checked before the output is finalized. The links between the variables and the coherence in their series are also checked.

6.3 Interpretation and explanation of output

Interpretation and explanation of the output is a process that is carried out simultaneously with the preparation of the preliminary output. Along with the press release, you can also prepare the set of tables for website, including the PX-Web format, which can be found on the KAS website.

6.4 Apply disclosure control

In producing tables and graphs for publication, staff takes care not to directly or indirectly identify statistical units. For identity protection, the tables and graphs are checked for the total of at least three units and the unit's share should not exceed 85 percent of the total. When the data produced can identify the statistical unit they are grouped into categories in order to maintain confidentiality.

6.5 Finalizing of output

Output finalization refers to final data and text control from the Sector which is responsible together with the Director of Department, which analyse the data consistency, time series estimates, and trend. At the same time, the responsible staff ensures that the written text and tables that will be published on the KAS official website are finalized in three language versions: Albanian, Serbian and English.

TEMPLATE FOR DOCUMENTATION OF STATISTICAL PROCESSES GSBPM

7. Publication / release

7.1 Updating the output system

Along with the publication, tables for the on-line Pc-Axis database are also prepared. This sub-process includes the following processes:

- Formatting data to be put into the database.
- Uploading data to the database.
- Receive confirmation for uploaded data.

7.2 Producing the products for dissemination

All ICT products are passed to the Dissemination Sector to carry out the process of editing and checking if these products meet the set standards of publication.

7.3 Managing of dissemination products

Includes informing specific groups such as the media or ministries, as well as addressing possible prohibitions before publishing and delivering products to subscribers.

7.4 Promote the dissemination products

The promotion of ICT publication products is accomplished by posting the most important findings on Facebook. Facebook posts are based on the KAS Publications Calendar, which contains release dates for publications and other important events.

7.5 Manage user support

Applications are deposited in the Department for Communication and Coordination, which is supplemented by No. of the Protocol, the data of receipt of the request, the content of the file, the person assigned to follow the request and the deadlines for completing the application.

8. Evaluation

8.1 Gather evaluation inputs

To assess the quality of the survey and to determine the proposals for its improvement, the following should be considered:

- remarks or comments of the enumerators regarding the questionnaires;
- questions or requests of internal or external data users;
- results of the analysis of the data collected through quality indicators.

8.2 Conduct evaluation

The assessment of ICT is in accordance with the SIMS (Integrated Metadata Structure version 2.0.) template for quality reporting. SIMS template in Albanian, Serbian and English versions as well as an explanatory material of each concept:

TEMPLATE FOR DOCUMENTATION OF STATISTICAL PROCESSES GSBPM

- SIMS Format ENG
- SIMS Format ALB
- SIMS Format SER

(SIMS) is the quality report that summarizes both ESMS and ESQRS reports, orientated to producers and users. This version is suggested by EUROSTAT since it reduces the time of completion by producers of the statistics, the concepts are only filled in one document and ESMS and ESQRS are then produced.

ESMS is recommended by the ESS Code of Practice and Regulation 223/2009 of EUROSTAT as a user-oriented quality report structure as this model contains a basic level of quality information that is structured and at the same time easy to understand.

ESQRS is a quality reporting framework that was developed by EUROSTAT in 2010, oriented to producers of statistics. These reports are usually not published, but are used by statistical offices to improve process in statistical activity.

8.3 Produce and adopt action plan for process improvement

Based on the assessment of the survey results, as well as the quality report, the Sector plans a change, such as: improving the overall process. The changes / improvements that can be applied in the following years are:

- Improvement and modification of the questionnaire
- Improvement and modification of logical controls in the data entry program
- Changing and modifying the data process (refers to logical control rules of the imputation and editing process).