

Topic IV: Progress in the implementation of SDC methods and techniques in central and eastern Europe

RESPONDENTS' ATTITUDES TO STATISTICAL CONFIDENTIALITY

Contributed paper

Submitted by the National Statistical Committee of Kyrgyzstan¹

1. The Kyrgyz Republic's Law on State Statistics gives the National Statistical Committee responsibility for collecting, accumulating, processing and storing statistical information, presenting it to the appropriate users and ensuring its completeness, integrity, reliability, freshness and adequacy. Under the said Law, the statistical authorities have the right to obtain reliable statistical and accounting information from all economic operators at prescribed times, free of charge and in full. In this way, the economic operators fulfil their obligation to provide information, which they do at their own expense.
2. The exacting demands placed on statistics derive from the need to produce official statistics in accordance with information requirements under market conditions and in accordance with international exchanges of information.
3. State confidence in statistics was built up over years of laborious efforts by several generations of statisticians, when the consumers of official statistical information were a clearly defined group. Having built up this reserve of trust, we are exploring new ways of preserving and increasing it. The information providers and users are made up of practically all tiers of society.
4. The independent status of the Kyrgyz Republic's statistical body allows it to exert maximum influence on the output of accurate data and actively conduct research. Data collection takes the form of both full-scale and sample surveys on territorial divisions, in accordance with registers compiled on the basis of the Single State Register of Statistical Units (ÅGRSÅ) and state classifications.
5. Primary statistical information is collected from economic operators by district and urban statistical committees, which perform checks, make the necessary changes and forward the information to the regional statistical committees on the primary statistical report forms.
6. At regional level, incoming information is checked and corrected and put into electronic form with a view to compiling summary information for publication and dissemination to various users.
7. As regards the completeness and content of the information we collect, it is adequate in terms of volume and substance:
 - firstly, the information is complex, reflecting both the economic situation of economic operators and social characteristics: the numbers and turnover of personnel, changes in average monthly production, educational attainment, standards of living, employment and unemployment, etc.;
 - secondly, the information comprises authentic, quantitative, estimative and qualitative data.

¹ Prepared by Anatoly Kim.

8. We produce the following two types of information from the chronological point of view:

- dynamic, reflecting changes in one or other indicator;
- static, i.ä. data relating to the particular time (date) in which the survey or inquiry is carried out.

9. In functional terms, the structural subdivisions of state statistics are as follows:

- planning, methodology and checks on the implementation of software used for statistical operations;
- organisation of statistical surveys on the basis of comprehensive and sample surveys, counts and censuses;
- the collection and preliminary checking of statistical inputs in the activities of market operators;
- the preparation, processing, compilation and accumulation of itemised data sets, tables required by law, regional information sources, databases and data banks, the protection and presentation of information sources in connection with user inquiries;
- analytical activity, including the compilation of aggregate informational and analytical material, the preparation and release of official statistical publications, and cooperation with users.

10. Recently, there has been considerable growth in the demand for statistical data (both micro- and aggregate data) and, at the same time, a decrease in willingness to provide information. As in many other countries, we are faced with the problem of non-response, complaints about the administrative burden associated with participation in statistical surveys and scepticism about data confidentiality. The National Statistical Committee treats its information providers with understanding. This is reflected in the fact that duplication is avoided in the creation of statistical accounting forms, the time taken to interview respondents is being reduced, and, of course, a guarantee is provided concerning the strict secrecy in which the information will be treated.

11. Under these circumstances, the confidentiality of the information provided is guaranteed by the Law on Statistical Confidentiality, which lays down not only the rights of state statistical bodies, but also their responsibilities: "Officials of state statistical bodies shall be liable for prosecution in accordance with the law of the Kyrgyz Republic, for false or unpunctual provision of statistical information, and for disclosing state or commercial secrets".

12. Confidentiality is ensured by the fact that all users may only access data within the limits established for each authorised employee.

13. In cases where a respondent and consumer are one and the same person, his or her attitude toward confidentiality is ambiguous. As an information provider, an operator may be a keen advocate of confidentiality, yet, in his capacity as a user of statistical information, he often requires information which is classed as confidential. He requires detailed statistical information, obtained via questionnaires, and wants to know about enterprises (numbers, average wages, etc.) and individuals, i.e. confidential information. The problem lies deeper: the respondent now knows what the user wants as soon as he himself becomes one, and so is very careful when answering the questions. The duality of the situation makes the respondent less trusting towards statistics.

14. The task facing us is to construct a client-server information system based on a local network with a central database. We will select one of the popular multi-user data-base operating systems (SQL Server) and software allowing the high-speed processing of applications (Delphi, Visual Basic), thus creating a system which is a combination of a database and appropriate applications allowing this particular task to be carried out.

15. The underlying principles of the information system used by the National Statistical Committee are based on modern technological platforms, information system topologies and architectures, guaranteeing a high level of security, reliability, integrity and security for information.

16. The applications software is written in the Clipper programming language. There are currently about 150 different sets of such software which are used for processing statistical accounts. Results are presented in the form of summary tables for each branch of the economy. Primary information on economic operators is input in coded form, based on the ÅGRSÅ and state classifications. This ensures a certain level of confidentiality regarding information about an economic operator.

17. Although programs written in Clipper can be used with any computer, even hopelessly obsolete ones, they still meet the requirements of the system used by the National Statistical Committee. The technical shortcomings of this system are the embedded graphics and the narrow range of data processing operations, and insufficient possibilities for guaranteeing data security. However, the transition to the creation of software products with the application of modern technology, using data accumulated previously, and ensuring the security and protection of information, is still under way.
18. The software is gradually being converted into a language which supports work with an MS SQL Server. Creating a system of statistical indicators makes it possible to organise appropriate access to databases. MS SQL Server uses a C-2 information protection device installed in Windows NT.
19. Used with SQL, the various utilities make it possible to set the functions of user applications, linked to databases. Using audit criteria, the administrator can apply filters, depending on the need to supervise specific users, applications and computers.
20. The system should be as easy to use as possible without harming the key objective of security.
21. Consideration should be given to the type of processing aids to be used to build an information system based on the above-mentioned operating system, and to how the applications will handle the data.
22. Moreover, it is helpful to use various software packages for statistical analysis intended for solving practical problems in a specific area, for example, as part of the analysis of time series. Automated data analysis systems (statistical software packages) are thought of as high-performance software products, but are perhaps most widely used in practical work and research in the most varied types of areas.
23. The degree of "intelligence" of a software package depends, first of all, on the organisation of an operating environment for the statistical software, where the user has sufficiently high-level statistical assistance throughout the entire statistical analysis process, i.e. when identifying the nature of the data to be analysed, choosing appropriate models and methods, integrating them in the technical network and interpreting results, etc.
24. It is a question of a developed system of interactive statistical support, covering the various stages of performing the task:
- guiding users to existing literature sources in the application of statistical methods, helping them with the terminology and concepts which are used and pointing to solutions to similar problems;
 - assisting with the formulation of the tasks, and detailed preliminary analysis of primary data, with the accent on drawing users' attention to the identification of the type of data and their peculiarities;
 - selecting an appropriate model and technical processing module network;
 - describing the most common statistical "traps", and how to avoid them;
 - helping with the interpretation of intermediate and final results of statistical analysis;
 - proposing the thrust of further research.
25. Every user has rights, based on his needs and the role of the National Statistical Committee. The network users bear some of the responsibility for protecting the information.
26. The level of protection against unauthorised access to software and databases is based on hardware, software and system-related methods and organisational wherewithal:
- the introduction of multi-stage password protection and assigned levels of access to information using standard software, operating systems and internal means of controlling access, thus ensuring a high degree of confidentiality;
 - the use of standard, approved software and applications for protecting information from unauthorised access, software "intrusion" and viruses;
 - fixed allocation of work stations (so that a given user can only gain access from a particular work station);
 - restrictions on user applications (user operations are recorded);
 - limitations on access to applications.

27. User access to the system is controlled and managed by the system administrator, who assigns the system users specific user names and passwords, levels of authorisation of access to clearly defined functions, items and data, and these reserves functions for specific users.

28. It should also be borne in mind that there are a large number of highly varied programs for storing and processing confidential information and ensuring computer security in a Microsoft Windows environment. The correct use of these programs makes it possible to increase the security of information in several ways. In fact, thanks to modern encryption algorithms, it is possible to provide effective protection, while keeping the information on one's own computer. This requires no more than a well-defined security policy and appropriate software.