

Distr.  
RESTRICTED

LC/R. 516  
31 August 1986

ENGLISH  
ORIGINAL: SPANISH

---

E C L A C

Economic Commission for Latin America and the Caribbean

Ninth Inter-American Conference on Statistics organized  
by the Permanent Administrative Committee of the Inter-  
American Conference on Statistics.

Rio de Janeiro, 15-18 September 1986.



COMPUTING SYSTEMS FOR STATISTICAL TASKS  
Possibilities of interregional co-operation \*

\* This document was prepared by the Statistics and Quantitative Analysis  
Division of ECLAC.



## A. Objectives of the document

1. Electronic processing of censuses, surveys and other statistical data has made remarkable progress in recent years. This brief document, prepared at the request of COM-CIE, has the following objectives: a) to present a brief outline of a series of computing systems of proved efficiency currently used in developed countries for statistical tasks (paras. 2 to 12); b) to describe the tasks carried out by ECLAC in this field and the possibilities of regional co-operation which are being explored (paras. 13 to 18).

## B. Computing systems for statistical tasks

2. The available computing systems will be classified according to the step in the electronic processing of statistical data for which they were designed. Three essential steps can be distinguished: a) edition, which includes the validation and correction of data; b) management, which deals with data storage and retrieval; c) tabulation and analysis, where information is processed both for tabulation purposes and for computerized statistical analysis (analysis of variance, models, simulations, etc).

3. In the last few years, the systematization of methods and techniques, connected with another stage or aspect of electronic data processing and management has acquired great importance, namely, the need for data on data. This need has given rise to what are known as Meta-Data or Meta-Information Systems. These systems are being designed basically as methodologies that establish criteria to define, identify and encode available data with the final objective of informing the external users of the Statistics System about which data are available, where they are, how they are identified, how to gain access to them and where they are published. Data on data may, for example, be translated into DIRECTORIES which may be printed or on-line. Another simple form of these directories are the familiar footnotes in statistical tables or texts.

4. To tackle and solve the problems posed by statistical data processing the tendency has been to design and develop Generalized Systems for each of the steps already mentioned. They are generalized in that the user is provided with a single tool adjustable for each particular case, thus making it possible to concentrate efforts on the specific details of the case and not on the development of ad hoc systems/programmes. No less important is the fact that they provide a language and a technique capable of being easily shared by statistical experts and computing specialists, since these Systems use languages which are closer to the peculiarities of statistical programmes than to those typical of computing - which are generally only suitable for the initiated.

5. By way of indicating to what extent serious efforts have been made to arrive at the solutions described above, and by whom, some Systems for the most commonly used mainframes and the Institutions that develop them are listed below:

EDITING

- . CONCOR : Census Bureau of the United States
- . CANEDIT : Statistics Bureau of Canada
- . EDIT 78 : Statistics Bureau of Sweden
- . UNEDIT : United Nations Statistical Office
- . DIA : Statistics Bureau of Spain
- . AERO : Central Statistics Bureau of Hungary

MANAGEMENT

- . RAPID : Statistics Bureau of Canada
- . SAS : SAS Institute (private commercial) United States
- . SPSS : SPSS Inc. (private commercial) United States
- . ADABAS : Software AG (private commercial) United States
- . TOTAL : Cincom Systems Inc. (private commercial) United States

TABULATION AND ANALYSIS

- . CENTS : Census Bureau of the United States
- . TPL : Labour Statistics Bureau of the United States
- . TAB 68 : Statistics Bureau of Sweden
- . TAU : Office of Population Censuses and Surveys of Great Britain
- . XTALLY : United Nations Statistical Office
- . INTERTAB : Multinational European Project, SCP
- . SAS : SAS Institute
- . SPSS : SPSS Inc.

6. ECLAC, through its Statistics and Quantitative Analysis Division, launched in 1984 a permanent and serious effort to take an active part in the thorough analysis and use of these systems. This decision was taken because of the opportunity offered by an important European multinational project, SCP (Statistical Computing Project), with which ECLAC agreed on a co-operation plan that provides the necessary support for disseminating the results of the project among the Latin American countries.

7. The SCP was formally begun in 1980 as a result of meetings held in 1979 by several government statistics offices of European countries faced with the challenge of finding solutions for the electronic processing of statistical data in an optimum integral way. The conclusion drawn was that given that the complexity of tackling all the aspects of this task, which demanded a long time and considerable financial resources, the best solution was to jointly undertake and co-ordinate all the studies and development of techniques for their generalized application in this area.

8. In October 1980, with the co-ordination of the Director of the Statistical Division of the Economic Commission for Europe, the project was formally presented in a document submitted to the United Nations Development Project (UNDP) in which the objectives, organization and work-plans for the next four years were set forth. The UNDP agreed to give financial support to the project, in view of the importance and relevance of its objectives and the success to be expected given the expertise of the countries concerned. The Economic Commission for Europe was designated as the Executing Agency, and in order to carry out the work, a project Co-ordinator based in Geneva and a technical consultant to the Co-ordinator, based in Bratislava, Czechoslovakia, were appointed.

9. To deal with the tasks of analysis and development, working groups were formed for each of the topics mentioned above.

For data edition, Hungary was the leader country, with the participation of Spain, Sweden, Yugoslavia, France, Ireland, Bulgaria and Greece.

The leader with respect to data management was Sweden, with the participation of Canada, Italy, Hungary and the German Democratic Republic.

In the case of tabulation and statistical analysis the leader country was Denmark and the participant countries were England, Czechoslovakia, Finland, German Democratic Republic, Poland, Romania and Turkey.

The leader in the case of META-DATA was Czechoslovakia, with Canada, the Netherlands, Poland and Bulgaria as participants.

These groups worked from 1981 until the end of 1984 under the original terms of the SCP project.

The editing group used AERO as the basis for its work. The management group used RAPID, while the tabulation and analysis group directed its efforts to the design and development of a new system to collect all the existing experience in this area and to overcome all the drawbacks detected. The META-DATA group had as its aim the design of a methodology for the management and supply to users of data on statistical data.

10. The results of the work carried out by the groups, by the end of 1984, were as follows:

EDITING: A new version of AERO was developed, known as AERO Ver.2, which incorporated the experiences and suggestions of the countries. The necessary documentation was prepared for its installation, utilization and definition of its characteristics. Yugoslavia designed and incorporated functions interactive with Version 2, thus producing Version 2.5. It is expected to improve the latter version in terms of programming techniques and functional characteristics, to produce Version 2.6, which will be the version finally used as the basis for maintenance and distribution.

MANAGEMENT: Since RAPID is already a fully operating and successful system, the group developed new functions (BASE OPERATORS) to make it a real relational Data Base Management System (DBMS); interactions with analysis packages such as SAS, SPSS and others are live. An interesting instruction manual for the use of RAPID which provides information on its characteristics, philosophy and functions was prepared. The complete necessary background documentation for the installation and use of both RAPID and of the BASE OPERATORS is now available.

TABULATION AND ANALYSIS: The group designed, programmed and installed, the INTERTAB system and all the necessary background documentation on its installation, definition of its characteristics and utilization was prepared. The installed system corresponded to Version 1.0 and is still being tested -although already live- in the countries forming this group, with a view to improving its execution efficiency and correcting some problems in the printing module.

META-DATA: This group produced a Manual for users of META-DATA Systems in which a model is defined for the creation, management and utilization of this type of information. Basically, a methodology has been established to solve this problem. The aspects of computerizing a system of this nature have been presented as elements to be considered in this phase, but the design and installation of a System satisfying the requirements of this model remains a task for the future.

To give information on the goals reached and disseminate the availability of these Systems, a seminar-workshop was organized in November 1984 in Budapest in which international organizations and national statistical bureaux participated. On behalf of Latin America only Cuba and Argentina attended. The results obtained by the project were considered successful and the Executing Agency was requested to seek means to carry on with the tasks of improvement and new developments.

11. All the Systems produced under the SCP project are available to developing countries free of charge; this is the direct consequence of one of the objectives which UNDP established as a requirement for giving financial support to the project. Both the Economic Commission for Europe and the countries participating in the project were pleased to accept this condition since it is consistent with policies of shared technological transfer to developing countries. It was decided to give the Regional Economic Commissions responsibility for dissemination of, and technical assistance in, the use of the System by the countries of their respective Regions.

12. As a consequence of the positive results of the SCP project, the member countries have decided to go on with the development work. As the financial support from UNDP has come to an end, the European countries involved and the Economic Commission for Europe held co-ordination meetings in 1985 in order to continue the work in an orderly manner. At the 33rd Plenary Session of the Conference of European Statisticians (Geneva, June 1985) it was decided to incorporate the SCP project as one of the regular activities of the Conference.

With this decision the SCP project has become a significant element within European statistical activities, so that it is possible to guarantee the continuity of its work and the continued official support of the governments of the member countries.

C. Possibilities for interregional co-operation

13. ECLAC decided to take charge of the SCP project in the region through its Statistics and Quantitative Analysis Division. In order to carry out this task, the following activities have been envisaged:

a) To install all the SCP Systems in the ECLAC Computing Centre in order to carry out exhaustive tests on operation, transportability and ease of use. For this task we shall receive technical assistance from European experts who took part in their development; we shall evaluate the quality of the documentation on the installation and use of the Systems; we shall identify the problems to be solved in installing these Systems in the countries of the region, considering the characteristics of the hardware used, and will design the necessary solutions for each case. In the shortest possible time we shall prepare versions of the technical manuals which are suitable for use in the region in the light of the language, hardware and level of professional manpower typical of our countries.

b) Two workshop-seminars are envisaged (for the 1987-88 period), one in the Centre or the Southern Cone and another one in Central America, to present the packages and as a means of providing immediate general access to their use.

c) There may be agreements with the different countries to send an expert for in situ installation and staff training.

d) There may be agreements with the different countries to send their technical staff to Santiago to be trained in the use of the Systems.

e) Technical assistance to solve problems that may arise in the use of Systems will be provided through telephone consultations, correspondence or any other suitable means.

14. The region has already developed some expertise in the use of Generalized Systems, particularly, CONCOR, CENTS, TPL, SAS and SPSS. As we are aware that this store of expertise and know-how should not be wasted, the SCP Systems will be introduced using a comparative training method when such an approach is deemed desirable by the countries in which the work is carried out; ECLAC will co-operate with the countries in the use of these Systems; furthermore, it will engage in permanent work on the evaluation of existing and new Systems in order to analyze the feasibility of using them, in the light of criteria of advisability and real gains in efficiency and technical quality relative to the level reached with the Systems being used.

15. In the last few years the appearance of more and more powerful micro-computers, with better Generalized Systems and at prices which are markedly lower than mainframe computers, has opened up a whole new perspective as regards electronic data processing. The characteristics of this type of hardware in terms of easy use, price, ease of installation with respect to environmental requirements and adaptability to different levels of application have made them become a totally valid alternative to mainframe computers in the multiple usual applications of offices that make use of electronic statistical data processing.

In fact, for the first time and at long last, a tool has been made accessible to people who are not computing experts and who are thus enabled to undertake most of the computing tasks by themselves. On the other hand, the expert is given the same possibilities that the larger hardware offered, and possibly new facilities also.

16. In view of this, and since it is safe to assume that desk-top computers will have ever-increasing capability to deal with the type of process that we are concerned with, work is in progress to adapt the available Systems to the characteristics and potentialities of this type of hardware.

There are already different versions of CONCOR, CENTS, SPSS, SAS and other new alternatives for desk-tops, and some excellent data entry systems are also available.

The SCP project has decided to undertake work in line with this trend also, and is preparing a meeting to organize work that will allow the adaptation of the SCP Systems to this type of hardware.

The availability of all these Systems in their maxi and micro versions opens up a vast field of possible solutions to problems of the same nature but of a different scale, by means of the use of a single methodology and the same computing tool. The training periods to carry out the same task will thus be considerably reduced.

17. ECLAC is fully convinced that the right and most feasible way to deal with most of the demand for computing processes in the countries of the Region is by means of the use of micro-computers and Generalized Systems either already adapted or in process of adaptation.

To provide incentives and assistance to the countries of the Region for the adoption of these tools as soon as concrete results are available, ECLAC envisages the preparation of a project similar to the one outlined for the SCP, although of a more limited nature because of funding. An exchange of know-how and expertise with larger countries will be aimed for, concentrating direct assistance on countries which are smaller and have fewer resources and for which these solutions appear to be the most advisable, given the scale of the habitual applications.



18. ECLAC considers that suitable conditions exist at present for carrying out an important interregional co-operation effort. Considering that there are particularly favourable conditions for gaining access to efficient systems at practically no direct cost and, also, that the hardware available in most of the countries makes it feasible to use these systems, ECLAC is submitting a project to UNDP aimed at transferring these technological advances to the Region. It is expected, also, that by means of the activities described above a fruitful stage of regional co-operation will be launched, which will naturally enjoy the backing of the institution.

1. The first part of the paper discusses the importance of the research and the need for a new approach to the study of the history of the world.

2.

3.

4.

18. ECLAC considers that suitable conditions exist at present for carrying out an important interregional co-operation effort. Considering that there are particularly favourable conditions for gaining access to efficient systems at practically no direct cost and, also, that the hardware available in most of the countries makes it feasible to use these systems, ECLAC is submitting a project to UNDP aimed at transferring these technological advances to the Region. It is expected, also, that by means of the activities described above a fruitful stage of regional co-operation will be launched, which will naturally enjoy the backing of the institution.

