

Distr.
RESTRICTED
LC/R. 610
October 1987
ENGLISH
ORIGINAL: SPANISH

E C L A C

Economic Commission for Latin America
and the Caribbean

Meeting of Directors of Statistics of the Americas

Santiago, 23-25 September 1987

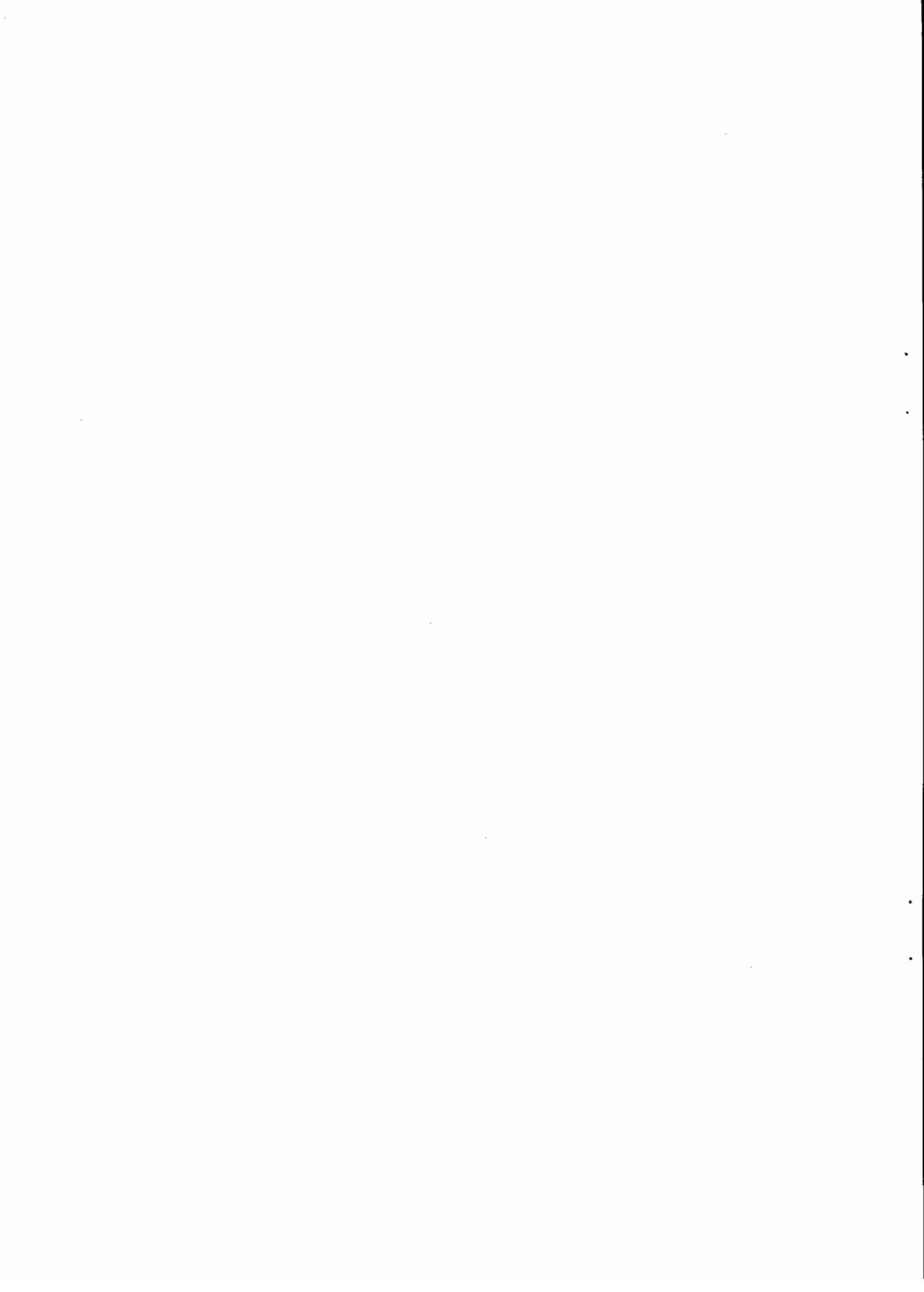
DISTRIBUTION OF STATISTICAL INFORMATION

This document was prepared by the Statistics and Projections
Division of ECLAC.

87-10-1462

CONTENTS

| | <u>Page</u> |
|---|-------------|
| Introduction | 1 |
| 1. Distribution channels..... | 2 |
| 1.1 The main technological options open at present | 2 |
| 1.1.1 Publications | 2 |
| 1.1.2 Microfiche | 3 |
| 1.1.3 Magnetic tape and disks | 3 |
| 1.1.4 Flexible or floppy disks | 4 |
| 1.1.5 WORM disks | 4 |
| 1.1.6 CD-ROM disks | 5 |
| 1.1.7 Supports which allow for direct access | 5 |
| 1.1.7.1 Data banks | 6 |
| 1.1.7.2 Telephone answering services | 8 |
| 1.1.8 Summary | 9 |
| 1.2 Multimedia distribution | 10 |
| 2. Institutional organization of the distribution of statistical information | 14 |
| 2.1 Institutions | 14 |
| 2.1.1 National agencies | 14 |
| 2.1.2 Regional and international agencies | 15 |
| 2.1.3 Private agencies | 15 |
| 2.2 Inter-agency relations | 16 |
| 2.2.1 Provision of data on a free or nearly free basis | 16 |
| 2.2.2 The marketing of tapes and other supports | 17 |
| 2.2.3 Redistribution | 17 |
| 2.2.4 Concessions | 18 |
| 2.3 Markets and prices | 18 |
| 2.3.1 Markets | 18 |
| 2.3.2 Prices | 20 |
| 3. Conclusions and points for discussion | 20 |
| 3.1 Considerations and questions | 20 |
| Notes | 22 |



Introduction

One of the ongoing duties of the ECLAC Secretariat is to concern itself with the collection, arrangement and interpretation of economic information and to disseminate the results of those activities. In this context, the promotion of the exchange of information among the countries of the region and between them and the rest of the world and of the methodologies used in analysing it is a continuous effort to which high priority has been given. Since the inception of the Commission, some of its tasks have been centred around this subject, in particular those carried out by the Statistics Division and the Economic Projections Centre.

In recent years computers have been increasingly used in Latin America and the Caribbean, which has considerably changed the working methods of statistical agencies in the region.

To co-operate with governments and fulfil the tasks for which it is responsible, the Secretariat is carrying out a number of activities. A project financed by UNDP for transferring to the region computer packages designed by statistical agencies in Europe, the United States and Canada has been designed and recently put into operation.¹ Activities are also being carried out in close contact with some countries of the region in the use of microcomputers, especially within the framework of projects in which ECLAC is the executing agency.

Other activities related to the creation of data banks and to the exchange of data in machine-readable form are also being carried out.

Finally, the Statistics and Projections Division has concerned itself with the whole question of dissemination of information. This involves examining new methods of disseminating statistical information, primarily in the economic and social domain, and the changes which these new methods may bring about in the philosophy of gathering, storing and retrieving information.

This document was prepared for the meeting of Directors of Statistical Agencies in the Americas to be held at ECLAC from 23 to 25 September 1987. Its preparation has been assisted by Mr. Raimundo Beca, an expert in this field. The opinions contained in the document are not binding upon or necessarily shared by either him or the institution for which he works.

The document is divided into three parts. Part One contains a description of the distribution channels now available and presents a new approach which would make it possible to link distribution and production with the

alternative use of these potential channels in an approach known as "multimedia distribution".

In part Two consideration is given to the institutional organization of the distribution of statistical information. Possible roles of the public and private sectors and the relationship between them make up a subject to which little thought has so far been accorded in the region.

In view of the close relationship between institutional organization and payments policies, a rundown is given of the various forms of payments practised at present.

Finally, in part Three, some considerations are given and questions asked with the sole objective of co-operating in the debate to be held by the Directors of the statistical agencies of the countries of the region.

1. Distribution channels

The dissemination of national, regional or international statistics, a task which was traditionally carried out by publications, can now potentially be carried out by a variety of media or supports relating to the communication of information.

Although in practice a given distribution circuit favours the use of a particular medium (such as publications), there is no reason why a distributor of data must feel bound to any particular physical support. Moreover, sooner or later technological innovation will make it necessary for a distributor of data to select one or more channels of distribution.

1.1 The main technological options open at present

Today the most widely used channels of distribution include the following:

- publications
- microfiche
- magnetic tape and disks
- flexible or floppy disks
- WORM (write one-read multiple) disks
- CD-ROM (compact disk-read only memory) disks
- computerized data banks and on-line access to them
- telephone reply services

1.1.1 Publications

Publications are and for some time to come will certainly continue to be the means usually employed for the dissemination of statistics. It is estimated that no more than 10% of all statistics are distributed through the use of physical supports other than paper and, moreover, in those few cases in

which recourse is had to other means, the information concerned is also published.

The use of photocomposition and electronic publishing has not only significantly reduced reproduction costs but has also made it possible to align formats and the numbers of copies produced to the needs of any particular group of users.

Their greatest drawbacks continue to be the space they occupy in libraries and documentation centres, the difficulties of correcting something which has already been printed, their vulnerability to natural disaster and delays in making information available, caused by the usual delays in printing, distribution and accessibility.

1.1.2 Microfiche

In order to relieve some of the disadvantages of publications, recourse is commonly had to microfiche and microfilms. The use of "micro-tools" makes it possible to reduce the space required for filing considerably, and if they are handled appropriately, they may last over one hundred years in normal conditions of preservation. On the other hand, it is inconvenient to consult them, and in many statistical centres, microfiche are in fact used as an intermediate support from which copies are made on paper for purposes of consultation.

Initially microfiche and microfilms were no more than miniature copies of publications which, as the name suggests, were put on film. Now, however, it is possible to produce microfiche independently of publications and even prior to their production. Actually, Computer Output Microfilms (COM) make it possible to publish microfiche directly from any computer magnetic support. On the other hand, it is also possible to publish microfiche directly through photocomposition with no need to resort to paper. There can be no doubt that these "microtechniques" help to reduce considerably the cost of producing and reproducing publications.

It is no wonder, then, that the use of microfiche is now prevalent; at present they are commonly used in the statistical agencies of various developed countries.

1.1.3 Magnetic tape and disks

For more than 20 years, magnetic tapes have been the most widely used physical support for the exchange and distribution of information among statistical centres. Their comparatively low cost, their great storage capacity in a small amount of space, the fact that they are easily handled and their high degree of standardization have made magnetic tapes preferable to magnetic disks. However, the use of one or the other comes to exactly the same thing, and in some cases, no distinction is made between them.

Initially magnetic tapes were used to convey international trade statistics between customs authorities and national statistical centres or

between countries and regional or international centres. Today they constitute a medium which is used with enormous frequency and is the usual means on which leading national, regional and international statistical agencies disseminate or market their output.

1.1.4 Flexible or floppy disks

With the appearance of microcomputers and their rapid penetration into the market, it also began to be usual to distribute statistical information on flexible or floppy disks. Of course some statistical material stored on floppy disks was already circulating for use in word processors. Thus, for example, some technical journals and market studies included floppy disks containing statistical tables to facilitate insertion in reports and publications. Now microcomputers also make it possible to carry out some processing locally and even to accompany the data with the software used in applying them.

In theory floppy disks perform the same function as magnetic tapes and disks, but, on the one hand, they are easier to handle although, on the other hand, considerably less information can be recorded on them. They are usually available in formats of 3 1/2 inches and 5 1/4 inches. The fact that these formats are relatively standard has resulted in the rapid diffusion of this means of distributing information.

1.1.5 WORM disks

WORM disks also known as draw (direct read after write) disks make it possible, as their name indicates, to record information without its being erased later. Technically speaking, this is done by a laser ray which burns a sensitive surface, while another less powerful laser ray is used for reading. Erroneous information may, however, be corrected by destroying the area recorded and recording the corrected information elsewhere on the disk.

The advantages of WORM disks consist in the small storage space they require, the large volume of information which can be recorded on them (between 200 million and one billion characters) and their long useful life (over 10 years). On the other hand, their reading and writing devices are relatively expensive.

Their most widely used formats are 12 inches and 5 1/2 inches. The introduction of WORM laser disks as peripheral parts of microcomputers ensures that this type of support will be widely accepted on the market. A 5-inch format compatible with compact disks also exists, but this product has so far not been widely accepted.

For the distribution of statistical information, data may be placed on WORM disks either locally from magnetic tapes, data banks (down-loading, for example), floppy disks, direct entry, scanners, etc.; or the information may be recorded centrally, in which case the user has only to add updated material.

1.1.6 CD-ROM disks

With the same technology as that used in the compact disk, whose acceptance and rapid penetration in the world music reproduction market is one of the greatest industrial success stories in recent years, this support has begun to be used for recording information for microcomputers. As its name implies, it consists in a read only memory (ROM) which is punched out on the basis of an identical master copy. As in the compact disk used for music, it can only read the information recorded.

The CD-ROM belongs to a family made up of derivatives of the audio compact disk - CD-V (VIDEO), CD-I (interactive), CD-VI (video-interactive), etc., all of which usually have a 5-inch diameter. Like its relatives, it makes use of all the economies of scale associated with a technology which has already sold 25 million reading devices and 250 million disks. Thus, the marginal cost of an ROM disk, or any other disk in the family, is about US\$2 for a disk of which over 1 000 copies have been made.

CD-ROM disks obviously offer the same advantages as WORM disks; i.e., they require only a small amount of storage space, they make it possible to store enormous volumes of information (600 million characters, equivalent to some 200 000 pages) and have a long useful lifespan (over 10 years). However, the ROM reading devices are significantly more economical than the WORM devices (a ROM reader costs under US\$1 million, while WORM reader-recorders cost US\$3 500 or more).

The CD-ROM technology makes it possible to integrate textual information, data and software in the same support. Thus, for purposes of statistical information, for example, a CD-ROM can store statistical tables, documentation on those tables, statistics, observations and meta-data, application programmes, instructions and examples pertaining to use, etc.

Of course, other supports, such as WORM disks, in particular, and magnetic supports in general, also allow for such integration. However, the superiority of ROM disks resides in the high degree of standardization achieved by this technology which ensures that any reader can be connected to any compatible PC and that any ROM disk may be read by any reader (specification of the High Sierra Group).2/

1.1.7 Supports which allow for direct access

The supports described so far have in common that they are physical media on which information is recorded or printed and conveyed to a device or appropriate place where the user has access to it. There are two additional supports which have different characteristics; they are kept in the hands of the producer or distributor, and the user may have access to the data they contain in order to study it and possibly to process it; access is achieved through various means such as telecommunications or locally by means of computer terminals connected to the central device which supports the information. One such support is the data bank; for the other, less well-known, one, we shall use the term "Telephone reply service".

1.1.7.1 Data banks

The administration of information by means of computerized data banks has brought about a notable change in the management of large volumes of data thanks to the computer techniques developed for this purpose.

Before proceeding further, it would seem useful to explain what we mean by a "data bank"; from the onset, let us say that it is something more than a voluminous collection of data stored on a single support with a more or less complex structure, to which access can be had through methods somewhat more sophisticated than the sequential method.

Without being too academic about it, we may say that a data bank is really a COMPUTER PHILOSOPHY, in which data constitute the hub of the doctrine. Less esoterically it may be said that the organization of the administration of the information is no longer viewed functionally and becomes a step in which data are given priority in and of themselves as an entity with their own value and importance, independently of the application in which they are used.

In this functional approach to organization, the data are arranged in independent computer files oriented to specific applications; for example, the data or variables of a survey will comprise a separate file, independent of other files pertaining to similar surveys or of files prepared for processing in accordance with the plan of analysis established for the particular survey concerned; naturally, the format, structure of the file and support used will be selected on the basis of the software available.

In contrast, the approach promoted by the data bank philosophy will be to identify and classify each variable so that it may be incorporated into an integrated information system; for example, let us take the variable (question) "branch of employment"; if this approach is taken, this variable will be given an identification code by which it will be recognized and set apart from other variables within the bank; this key or code will be designed to identify this variable as a branch of employment in a given survey, corresponding to the person interviewed in a certain household belonging to a given sampling unit; a number or title will of course be associated with it, and the different values or codes it may be given in relation to some standardized classification, such as, for example, the ISIC, Rev. 2, will be documented, and a description will be made of the techniques used to obtain those data from the person interviewed and their relationship with other variables will be defined; once all these specifications have been met, the information obtained may begin its own life in a collection of data (voluminous or otherwise), and picking the variable in question out from the rest will not be a laborious task.

It follows that once such specifications have been developed for each of the variables in a given survey and for each of the variables processed in a computer centre, the final result will be a data bank whose contents are well-defined, well-identified and well-integrated.

This collection of specifications constitutes what is known as a DICTIONARY OR CATALOGUE OF DATA contained in a data bank; all information

identifying and describing the data available is now known as meta-information or meta-data; at present research is being conducted in this field to establish the methodologies and standards needed to describe statistical data so that it can be disseminated in a way which is universally manageable; a study is also being made of the most effective techniques by which a user can employ such a DICTIONARY easily and rapidly.

DICTIONARIES are the means available to the data-management system to ensure that users are familiar with a bank's contents and characteristics, and they ordinarily provide additional information relating to the policies and techniques of access to the bank.

Storing all the information on computers, once it has been classified in this way, structured rationally or hierarchically or in a network on supports of a sequential type or direct-access supports is a problem which is solved with the techniques available in this connection, which, for obvious reasons, we will not attempt even to hint at here; in the same way that the storage techniques are highly developed, the techniques used for retrieving data have also been elaborated on the basis of sound principles.

In addition to mentioning logical organization and to outlining the physical organization of its computing facilities, we should say something about systems of access to the bank --access which it will be possible to obtain locally by terminals connected directly to the computer used, remotely through the use of telecommunication systems or by means of deferred procedures using batch processing, or through any combination of these three approaches. When this is done, we shall have described a data bank system combining characteristics of integration, non-redundancy and data independence, compatibility, reliability, safety and, most important of all, versatility.

Although the development of computing systems in a processing centre for managing all the tasks relating to the operation of a data bank is a challenge which can be accepted, it will call for enormous effort in terms of the acquisition of highly specialized personnel, the logical and physical design of the base and (no less important) the purchase of computer equipment.

For a number of years sophisticated systems have existed for the management of data banks which enable them to be installed efficiently and within reasonable lapses of time; in computer jargon, these systems are known as Data Base Management Systems (DBMS) and have been designed and developed both privately and in government, non-profit making institutions. Certainly the most important decision to be made in connection with the installation of a data bank relates to choosing the DBMS whose characteristics make it possible to satisfy the requirements of the processing centre concerned.

Very frequently the computing system which manages data banks provides for various data processing systems which can be generally applied and have a high processing capacity, the latter being an aspect which satisfies one of the most frequently made demands on the part of users. When DBMSs are used in conjunction with those systems, access to data is obtained and they are subsequently processed in a way which is very efficient, very direct and easy for the user.

Looking back over our description of various dissemination supports, we see that the data bank, as here described, is another technical support which may be used for dissemination.

The characteristics of data banks provide the user with new ways of gaining access to information, which are totally different from those already available to him such as, for example, the traditional publications approach. A bank means that it is no longer necessary to be tied to a rigid format, that it is possible to select only the information needed; that if on-line access to data is available, it will be possible to obtain usefully updated information immediately and also to use the bank's processing capacities to make tabulations and statistical analyses, to design econometric models, to recover data in printed form and even to copy selected information onto some magnetic support (an operation known as "down-loading").

In this way, the user plays an active role in the dissemination process, which was not true in the case of earlier supports.

It may be stated that for all these reasons, the data bank is called upon to play a mediating role between a user with a problem and the data and systems used in processing them.

Finally, it is worth repeating that in respect of the dissemination of information, the data bank provides a new support of great versatility and effectiveness.

1.1.7.2 Telephone answering services

Telephone answering services may be adapted for the dissemination of statistical information. In some cases, the message which the user receives is standard; in other cases, he has several options to choose from; for example, a dial or a keyboard may be used in selecting the data required (for example, dialing the number 12 may connect the user with the price index or with the stock quotations of a given exchange).

The responder operates either through pre-recorded messages or directly on the basis of a data bank or a computerized file which is read by a voice synthesizer. Often the voice synthesizer provides only the variable datum (e.g., the value of the quotation), and the rest is pre-recorded. In France, the introduction of this procedure has met with great success. Thus, 280 000 of the 400 000 requests addressed to INSEE concerning statistical indexes in 1985 were dealt with by telephone answering services.

It should be added that some systems for distributing information by telephone answering services also offer the user the possibility of recording personal messages concerning, for example, his comments concerning the information received, or asking additional questions.

1.1.8 Summary

The synoptic table which follows summarizes the advantages and disadvantages of each of the supports discussed and evaluates the comparative frequency of their use at present.

| | Advantages | Disadvantages | Comparative frequency of use |
|-------------------|--|---|------------------------------|
| Publications | - Familiarity - Ease of production | - Storage - Updating - Selection - Vulnerability | + + + + + |
| Microfiche | - Mobility - Storage - Reproduction - Durability | - Updating - Consultation | + + + |
| Magnetic tape | - Selection - Volume of data | - Equipment - Vulnerability | + + + + |
| Floppy disks | - Mobility - Selection | - Volume of data - Equipment - Vulnerability | + |
| WORM disks | - Selection - Volume of data - Storage - Durability - Mobility | - Equipment | + |
| ROM compact disks | - Selection - Volume of data - Storage - Reproduction - Durability - Integration of text, data and programmes | - Updating | - + - |

| | | | |
|------------------------------|-----------------------------|-------------------------|-------|
| Data banks | - Selection | - Cost of operation | |
| | - Volume of data | - Equipment | |
| | - Versatility | - Telecommunications | + + |
| | - Timing of reply | | |
| | - Updating | | |
| | - Processing of information | | |
| <hr/> | | | |
| Telephone answering services | - Timing of reply | - Volume of information | + + + |
| | - Updating | - Telecommunications | |
| <hr/> | | | |

1.2 Multimedia distribution

The information contained in the foregoing section concerning the various media available for the dissemination of statistical information gives rise immediately to the question whether it is desirable or necessary to adopt a policy of multimedia distribution in the field of statistics.

In reply, consideration will be given to the available supports mentioned above from the point of view of both the production and the use of data.

In many of the centres producing statistics in the region, publications, and to a much lesser extent, microfiche are regarded as indispensable and are virtually the only dissemination media used. Publications, in particular, have so far been unhesitatingly regarded as the natural medium of access to information; once the publishing process has been completed, distribution organized and the mechanisms of user access established, the information they contain is immediately available to be used. With the advantages and disadvantages indicated in the synoptic table, publications will, at least in the near future, remain the indispensable medium for the ordinary user and for the presentation of specialized information designed to meet specific requirements; it should be noted that the techniques needed to produce publications are firmly entrenched in the production centres and that, on the other hand, users will certainly for a long time demand this medium because of their familiarity with it and the ease of consultation which publications have offered since they first appeared.

Magnetic tape is a specialized medium which has been used to a great extent for many years to disseminate statistical data; it is the medium which has made it possible to provide access to large quantities of information in a computerizable format so that the producer can provide the user with recently updated information of broader coverage in terms of time and subject matter --assets which it is considerably more difficult for publications to offer owing to their nature.

It should be specified that the users of this medium tend to be specialized centres which are well endowed with computer equipment; it should also be noted that the producer-user exchange associated with this medium is preeminently an inter-agency exchange.

The incorporation of micro-computers ("micros") in statistical agencies of the region, in international and regional bodies and, to no lesser extent, in private professional activities makes floppy disks a universal means of dissemination.

The use of floppy disks for dissemination is already an ongoing reality to which special attention must be paid in view of the fact that the technical means of using them are in the hands of a very broad range of users. The standardization of formats and programmes for data processing creates virtually total compatibility between producers and users; moreover, owing to the physical aspects of the connections between micros and major equipment, incompatibility between techniques and systems relating to the major equipment are no longer a barrier to the exchange of computerizable data since micros perform the role of entry and exit facilities for highly compatible data.

The at present limited capacity of floppy disks is totally overcome by the fact that it is possible to reach a much greater variety of types of user; as in the case of magnetic tapes, exchange takes place among agencies, but it is not limited to them since, as has already been mentioned, the equipment needed to process floppy disks is to be found in a wide variety of places, including at the level of individual professionals; for the latter reason, in particular, the data concerned is disseminated to a spectrum of users who cannot be reached in any other way, so that a new approach to dissemination is brought into being.

The media analysed so far show us that a multimedia approach to dissemination is useful and necessary. In actual fact, the variety of user types, the technical requirements of many of those users for data which is directly computerizable and the more specialized demand in terms of timeliness and updating call for an approach of this nature.

The characteristics of the equipment at present available to the statistical agencies of the region and the technical know-how on which they can draw make it possible, on the basis of the supports mentioned above, to adopt a multimedia dissemination policy without making an enormous effort in terms of either staff or development and even less so in terms of computer equipment; a decision to provide census samples, surveys, foreign trade data, economic and social data series, etc., recorded on magnetic supports such as those referred to above call for very little organizational effort and for the development of computer procedures which present no technical challenge which cannot be met with the resources now to hand; such a decision will allow the institutes of the region to fulfil their duties of providing information for users who require this type of support in a timely and efficient manner; although it may be obvious, it should be emphasized that meeting the requirements of this type of user at official and private level amounts to placing information in the hands of the people best placed to perform the activities most important to the development of the region.

The decision to meet the requirements of multimedia dissemination with the supports described above can be re-examined in the light of the possibility of adopting data bank methods and systems.

If the management and processing of information were organized around data banks, all the activities involved, and those relating to dissemination in particular, would doubtless have to be changed substantially.

The dissemination of statistical data in a way which makes use of the resources referred to offers an advantage which is so far an impossibility --local or remote on-line access for the user to all the information he is permitted to have. The functional characteristics of data banks as described in section 1.1.7.1 make it possible for the user to be directly responsible for deciding what data he needs, in what format and on what medium it should be recorded; moreover, since it is now possible, by organizing data banks, to provide the user with the capacity of processing data at the same time as he is given access to it, he will not only be able to obtain the data recorded in the bank but will also be able to manipulate it by means of a computer in order to adapt it his requirements; for example, he will be able to use the resources associated with these systems to obtain "final products" such as statistical tables prepared through the use of generalized tabulation systems and printed in the camera-ready style obtainable in LASER printers, for example, which will undoubtedly be incorporated in the equipment used in a data bank.

It can also be stated that once the data bank system has been selected, the diffusion of information by the means described above becomes a simple operation of copying the data found in the bank; the copying process can be performed as an activity assigned to the processing centre itself in order to satisfy standardized dissemination policies or may be carried out directly by the user (down-loading), depending on his own needs.

It should be borne very much in mind that the use of data banks to impart great versatility and efficiency to all the activities carried out by a processing centre requires, in exchange, a considerable investment in organization, highly qualified staff, equipment and systems. If access to the data bank is of the remote, on-line variety, equipment and facilities for the exchange of data by telecommunications will be required along with its corresponding system of programmes for the management of communications, and it will certainly be necessary to apply to public or private enterprises with a view to leasing lines of communication associated with networks at national, regional and international level.

In order to understand all this better, imagine that a multipurpose household survey is processed in a data bank. A survey of this kind will cover various subjects, such as housing, demography, income, occupation, health, etc.

Once this survey has been incorporated in a data bank, the data or variables it contains will be available to any user who is qualified to use any of the media described above.

From the user's point of view, for reasons relating to the structuring of data banks and systems of access to them, the survey will be reduced to only those variables needed by him for his purposes. The user may, for example, obtain calculations relating to combinations of various variables; he may use the variables to make simple or sophisticated statistical analyses, and from

the point of view of dissemination, he may copy data from the survey, using any of the supports described above. Although no previous mention may have been made of the matter, the user may of course take possession of the data, provided his connection is remote, for transmission along the lines of communication he is using for his connection to the bank; naturally, the device which the user is using in his connection must have the capacity to store, in some kind of electronic memory, the data which reach him. This kind of operation is very simple when the connecting device is a micro-computer and slightly more difficult when the user requests that the data be transmitted between the computer supporting the base and his own computer which supports the connecting device --usually a video terminal.

The user, as may be seen from the above, has at his disposal numerous possibilities from which he can, in each case, choose the most appropriate for his objectives.

If we now consider the survey from the point of view of the producer, its organization within a data bank also offers many advantages.

Since the producer manages the information and in that capacity has access to all the variables in the survey, he can easily draw up publication schedules in respect of topics or groups of topics, set publication priorities, establish the procedures by which he will meet the requirements of multimedia dissemination in accordance with his policies in this connection and will also be able to satisfy any given user's requirements with regard to a particular support.

Although all the activities described above may be carried out without the intermediary of a data bank, the versatility and integrity provided by data banks mean that the producer can attain all these objectives surely, rapidly and efficiently.

In short, incorporating a data bank as the backbone of the activities of a statistical data processing centre leads to qualitative and quantitative changes in all the centre's activities, from the stage at which the collection of data is organized and preparations are made for incorporating them into the bank up to the processing and dissemination stages in all their aspects.

A virtually unlimited field of possibilities opens up to the user, with virtually no effort required on his part. Viewed from the perspective of the producer-user relationship, this technical resource synthesizes all the possibilities for multimedia dissemination and combines them with a processing capacity, which is very difficult to obtain with any other kind of strategy.

The rest of the supports referred to, and the WORM and CD-ROM devices in particular, are in the process of gathering general acceptance by the market. Given their characteristics, we may be certain that in the medium term they will acquire importance in the dissemination of statistical information which needs little or no periodic updating, such as census data, survey data, foreign trade data, production statistics, historical demographic series, etc.

It is important to pay particular attention to the development of market activities in this area since, in view of the evolution of the prices of these

devices and their increasing capacity for storage, it may be necessary to reconsider established policies which may prove obsolete by comparison with policies which might result from widespread access to these devices.

In short, from the producer's point of view, in order to re-examine the resources which are being used for disseminating information and consequently for establishing multimedia dissemination policies, it may be necessary to act more promptly and efficiently in the task of giving users access to the large volume of data which are usually somewhat neglected when very limited channels of dissemination are available; by analysing the technical profiles of users, it will be possible to decide the stages in which the supports utilized will be diversified, the final goal always being the establishment of a data bank, which international experience shows is bound to be a highly desirable solution, considering the progress made in the field of electronic data processing.

The results obtained with multimedia dissemination policies are decisive for users if they are to reach goals of efficiency and timeliness in their computing activities. On the other hand, choosing a support on which information can be received enormously simplifies the complex and slow task of collecting data by other means while at the same time optimizing the situation as regards expenditure on equipment.

Finally, it should be specified that multimedia dissemination enables the producer and the user to interact with the maximum degree of efficiency, allowing both of them to invest their resources virtually exclusively in their own respective tasks.

2. Institutional organization of the distribution of statistical information

2.1 Institutions

2.1.1 National agencies

The functions, organization and location of the national statistical agencies of the countries of Latin America and the Caribbean differ appreciably. In general, they are producers of statistics. In addition to this primary function, there are other functions relating to the processing and distribution of information. In many countries they compile and process information produced by sectoral offices, ministries or specialized agencies. On occasion, part of the distribution function has been delegated to private or public bodies engaged in the dissemination of information.

2.1.2 Regional and international agencies

Just as national offices operate at country level, regional and international agencies operate at the level of a geographical region or at world level.^{3/}

In general these agencies do not gather primary information but compile and process data provided by their member countries and ensure the comparability of those data.

Computer-to-computer communication of statistics among national agencies and international organizations is usually very infrequent. Some relatively ambitious projects now exist in the European Community.

The main task to which these organizations must address themselves is the harmonization of data and occasionally the verification of information. Actually, in spite of the vast effort made, problems of nomenclature or of special treatment and interpretation by one country or another will always persist. New interpretations, adjustments and commitments are often necessary.

Appropriate documentation of these meta-data is a sine qua non for the effective utilization of the statistical data compiled by regional and international agencies.

The large majority of regional and international agencies distribute statistical information directly by means of publications. Some of these organizations, such as, for example, the International Monetary Fund, the Statistical Office of the United Nations Secretariat, ILO, FAO, WHO, UNIDO, the World Bank, OECD, EEC, etc., also distribute data in the form of microfiche and magnetic tape and, in some cases, even permit on-line access. In some cases, magnetic tape and on-line access are strictly reserved to country members. In other cases, information available through such means is also sold or made available to third parties.

A special situation arises when information is sold or made available to a redistributor. In such cases, the most frequent approach at present is to charge redistributors a higher price than that charged to final users (on the order of four times higher).

2.1.3 Private agencies

Commercial data banks, private agencies which prepare diagnoses and economic forecasts and other private organizations which sell statistical information and process data helpful in making use of that information are in general more often identified as distributors than as producers of information. What happens most commonly is that private agencies purchase information from numerous sources and try to add as much value as possible to it. Value added in this way consists in (ranging from the simplest to the most complex): the selection of pertinent series; harmonization, listing and adjustment of series; documentation; data processing; integration into models; forecasts; advisory services; etc.

However, some private agencies also produce statistics. These include organizations such as the New York Stock Exchange, which, to all intents and purposes, produces only primary information which is later marketed and distributed by third parties.

Two basic differences in behaviour may be observed between private agencies and their public counterparts at national, regional or international level.

The first of these differences is that official or public institutes charge hardly anything for producing information, the cost usually being covered by the budget of the organization in question. The prices charged in fact roughly approximate the cost of the physical supports plus the cost of data processing specifically associated with the information distributed. The production of private agencies is, on the other hand, invoiced on the basis of its real cost.

The second difference is that, commercial considerations notwithstanding, private agencies have no problem in making value judgements concerning the quality of data or even in making estimates in this connection. Official agencies are, on the other hand, subject to many restrictions and much pressure, particularly with regard to the following: statistical secrecy, commercial secrets, national and international policy considerations, etc. Thus, for example, official bodies, for political reasons cannot freely divulge evaluations of the comparative risks involved in paying the external debt. Another dimension of this same lack of symmetry between official and private agencies may be observed in the limitations placed on the former, or which they place on themselves, in respect of processing data for a given user.

2.2 Inter-agency relations

The relationships between institutions of different kinds, whether private or commercial, in practice take various forms, including the following:

- free or nearly free provision of data on magnetic tapes or supports
- the marketing of data on magnetic tapes or other supports
- redistribution contracts
- concessions.

2.2.1 Provision of data on a free or nearly free basis

This is practised among countries and regional and international organizations. The relationship between associates and chambers of commerce or professional organizations is also frequently of this kind.

Physically, the most common practice is the exchange of magnetic tapes. However, on-line consultation is also practised in some cases as, for example, among the countries members of the European Community. Direct, computer-to-computer links, on the other hand, are, as noted above, extremely infrequent.

In general, this kind of relationship is symmetrical. Thus, if the relationship between a purveyor of raw information and the compiler of such information is based on an arrangement whereby the information is free or nearly free, the same is true in the opposite direction.

If payment is made, and the information is provided on a nearly free basis, no attempt is made to pay for more than the material costs of the magnetic tapes and other supports or special processing devices used. For budgetary reasons, instead of invoicing, other forms of compensation are practised. For example, the cost may be included with other costs charged to associated bodies or a tape with data may be exchanged for several virgin tapes, etc.

2.2.2 The marketing of tapes and other supports

Statistical agencies, aware of the economic value of the information they collect and process, increasingly engage in the practice of marketing those data, usually on magnetic tape.

The commercial practices engaged in by official bodies, at national and international level, differ substantially from those of private agencies, as has been seen above.

On the one hand, official bodies do not usually charge for the real value of the information they distribute nor for the costs of producing that information. Although there is a growing tendency to charge third parties higher prices than the nearly free provision of information made to other official bodies, this is still a far cry from charging the final users an opportunity cost or the market value of the information, when there is a market for it. It may, however, be observed that some official bodies are beginning to charge redistributors higher prices than those charged to final users (charges four times as high are common). This is at present the case in such countries as the United States, the United Kingdom, France and other OECD countries.

On the other hand, official bodies are constrained by their bylaws or impose limits on themselves in the matter of adding significant value to information. As will be seen below, it is for value added (evaluations, special processing, estimates, advisory services, etc.) that the market is best disposed to pay.

2.2.3 Redistribution

In general, there is a wide margin of possibilities for finding a suitable and balanced solution which takes into account the interests both of the redistributors of data and of the official agencies which provide them with those data.

The cost of producing basic statistics would be prohibitive and impracticable for a redistributor. Even when, as mentioned above, the sales price charged to a redistributor is four times as high as that charged to a

final user, if a redistributor has a large clientele, it is to his best advantage to purchase data. By contrast, however, when no official body collects reliable information on topics deemed to be essential, redistributors are forced to produce it themselves or to purchase it in the market at prices considerably higher than those charged by public institutions.

Official bodies may view a redistributor as representing an opportunity to reach final users which it would be difficult to identify and service directly. Naturally, it is the value added to the information which creates a product which is attractive to potential users.

The tendency nowadays is towards the establishment of contractual relationships between producers and redistributors. In addition to putting a price on the provision of data, the contracts involved in these relationships specify the obligations of the parties to them. In connection with such obligations, provision is made for such matters as frequency of delivery of data by the producer and the formats and supports used for those data and, on the part of the redistributor, for the frequency with which the information must be updated and for the commercial ethics involved.

2.2.4 Concessions

A concession is a special form of redistribution contract, in which, as in public service concessions, the beneficiary receives exceptional privileges such as exclusive rights in respect of certain brands or priority access if not absolute access, to data. The counterpart of these exceptional privileges is in general expressed in terms of free access to the services rendered by the beneficiary or preferential prices for certain users, the right to technical quality and frequency of updating of the services received and also the obligation to continue in certain activities although they may be uneconomical.

What usually happens is that the concessionary or concessionaries are selected on the basis of a public proposal and that the concession is for a certain duration of time (5 to 10 years). In France there are various examples of such concessionary contracts. In a number of European countries it is at present worthwhile establishing this approach to redistribution.

2.3 Markets and prices

2.3.1 Markets

The decisive factors in the success of any approach to the distribution of information are clearly the value added to the raw information and the ad hoc nature of the services rendered.

Where statistical and economic information is concerned, the value added consists in the selection of pertinent series and data; the documentation of such data; the evaluation of the data in terms of relevance, scope, limitation

and coverage; the software used; the forecast models; the forecast and advisory services in the use of the system and in economic strategy.

As for the ad hoc nature of the services rendered, it resides first in the capacity of the system to understand the questions formulated by one user in particular and reciprocally to formulate the replies in a language understandable by the user; secondly, it consists in the capacity to perform data processing in the manner specified and in such a way that the data supplied to the user can be used.

It is now estimated that the world market for computerized information amounts to about US\$4 billion a year. Of this, 60% is in the United States and about 80% of the world market is supplied by United States firms. The financial, commercial and demographic information sectors are those which enjoy the greatest success on the market. Since statistical agencies participate directly or indirectly in the commercial and demographic information sectors of the market, it is worth pausing to consider them in some detail.

The commercial information sector supplies data on the productive activities, finances and solvency of enterprises. The main users of this type of service are purchasing and marketing departments of enterprises which can have recourse to data banks for replies to questions such as who manufactures what?, who buys what? or what are the commercial risks in respect of any given client? In general, the information on the basis of which this type of data bank is constructed include corporation or trade registers which are frequently managed by national statistical agencies. Of course it is difficult for them to bestow a solvency or commercial risk rating on enterprises since rating or scoring procedures always have something subjective about them. In mixed economies this gives rise to a field of activity in which it is possible to design approaches to co-operation among public producers and private redistributors. The line between the public collection of information and the private sector is in many respects arbitrary, varying from one country to another, and is frequently a source of controversy.

What is known as demographic information is to a large extent obtained from population censuses and household surveys. This sector provides data on the demographic structure (the age structure and socioeconomic parameters) and consumer patterns for the smallest geographical area possible (the commune, the neighbourhood or the area covered by a zip-code, for example). The primary users of this kind of service are trading and publicity firms which strive to focus their commercial promotion campaigns on the basis of the socioeconomic characteristics and consumption patterns of the geographical areas and population groups at which they are aiming. The high production cost of census or household data has caused a proliferation of private agencies which conduct surveys. These agencies, whose work is to a large extent related to political surveys, considerably reduce production costs and the cost of demographic information by relying on pulls or panels, which are small permanent samples designed with specific objectives in mind. When these pulls or panels are not related to political surveys, they usually take the form of family budget censuses carried out by means of purchase or credit cards distributed to the households participating in the sample, surveys on the distribution of common goals which are usually carried out by examining the results shown on cash

registers or conducted by survey workers who observe the movements in shops or at counters or performed through audiomats which automatically monitor the tuning of radios and televisions.

This illustrates the decisive role played by the value added to information and the personalization of processing procedures.

2.3.2 Prices

The orders of magnitude of the prices now charged in the market are as follows:

- On-line access to a data bank costs between US\$100 and US\$200 an hour. In some cases an annual subscription fee of about US\$500 is collected.

- Commercial information is usually invoiced by unit; one unit of commercial data cost approximately US\$20. Frequently users must purchase a carnet containing between 20 and 50 tickets (one coupon = one information unit).

- Subscription to a CD-ROM-supported data bank costs between US\$1 000 and US\$3 000 a year. A subscription usually gives the holder the right to one new disk every three months, the new disk frequently cancelling and replacing all previous disks (releases). The CD-ROM reader, which when purchased in large amounts, costs about US\$500, in some cases also comes with the subscription.

- An official institution invoices a third party, not a redistributor, between US\$500 and US\$1 000 for a recorded magnetic tape.

- It is customary to invoice redistributors four times the price charged to third parties.

- The exchange of virgin tapes for tapes with data (a common practice) is carried out on the basis of three tapes for one.

3. Conclusions and points for discussion

3.1 Considerations and questions

Statistical bureaus in Latin America and the Caribbean are now incorporating new computer equipment and establishing connections with telecommunications systems which provide them with many possibilities for the dissemination of statistics. On the other hand, it is becoming apparent that it is possible and necessary to establish links between production and distribution procedures and the use of information.

This document contains a description of the various supports which can be used for disseminating information. First of all, the participants in the meeting can exchange the experience they have acquired. In this connection it seems worthwhile evaluating the suitability of different supports by studying

factors such as their acceptability or the demand for them on the part of different users, the degree to which they complement or replace publications, their cost and the facilities or difficulties associated with their use.

One of the most important of the supports listed is the computerized data bank to which the user has on-line access or access obtained through batch processing. This of course amounts to a qualitative change for the user who may select the information most useful to him and appreciably reduce the time gap between possession of facts and availability of information. On the other hand, the handling of this facility makes it necessary to design policies of a different nature on subjects as diverse as the determination of the data to which access will be provided to the setting of prices. In spite of the fact that experience in this field is hard to come by in the region, there can be no doubt that many directors are studying the problem and will want to discuss the matter.

The discussion concerning data banks bears on questions relating to organization and incorporation of techniques which must be carefully considered. The idea of a multimedia approach is introduced in the document, and the advisability of its being adopted in statistical agencies is suggested. In our opinion this is a subject which provides many opportunities for regional co-operation. It calls for the incorporation of new techniques, some of which are being developed in the OECD countries. At the same time, it opens the way to the definition of policies for exchanging information and for co-ordination or joint negotiation in acquiring access to or purchasing systems which we will all be needing shortly.

User interests play a very important role in the multimedia approach and in fact have always been a focal point for producers of statistics. At present, it seems possible to respond to a greater number of requests thanks to the new media available. It is interesting to note that one of the greatest verified demands in the developed countries is related to simultaneous access to information and to the software which can transform it, thereby adding "value" to it. Detailed and documented dictionaries or data catalogues which give users and redistributors the access they need to information and enable them to develop economic applications or models for the preparation of projections constitutes the kind of inputs which are in heavy demand from users. ECLAC is particularly concerned with this subject, and its project on statistics for the development of Latin America and the Caribbean, which has received a great deal of support from the directors of statistical agencies, bears on it. It would be desirable for participants also to propose new areas for joint action in this connection which might be carried out in the region.

Finally, question arises concerning the institutional framework in which the tasks of producing and distributing statistics are carried out. It is worth asking to what extent it is useful to draw a distinction between these tasks and to encourage separate agencies to specialize in them. This question is closely associated to another one which relates to the distribution and price policies which might govern the contractual links between producers and distributors. Information relating to these subjects is contained in the document, and we trust that viewed in the light of the experience of Latin America and the Caribbean, it will make it possible to outline different

strategies adapted to the special conditions and political decisions characterizing the different countries of the region.

NOTE: For a full list of regional and international statistical agencies see document E/CN.3/1987/2 entitled SPECIAL ISSUES: COMMUNICATIONS BETWEEN NATIONAL AND INTERNATIONAL STATISTICAL AGENCIES AND THE DEVELOPMENT OF DATA BASES.

Notes

1/ The contents of this project are described in document LC/R.516 entitled "Computing systems for statistical tasks" which was issued in August 1986.

2/ The High Sierra Group comprises the leading companies in the sector, such as Phillips, Sony and Microsoft, and has proposed standards which have been widely accepted in the industry.

3/ Data bank management software is a set of integrated programmes which make it possible to manage the storage, organization, retrieval and processing of data; on the market these are known as Data Base Management System (DBMS) packages.



