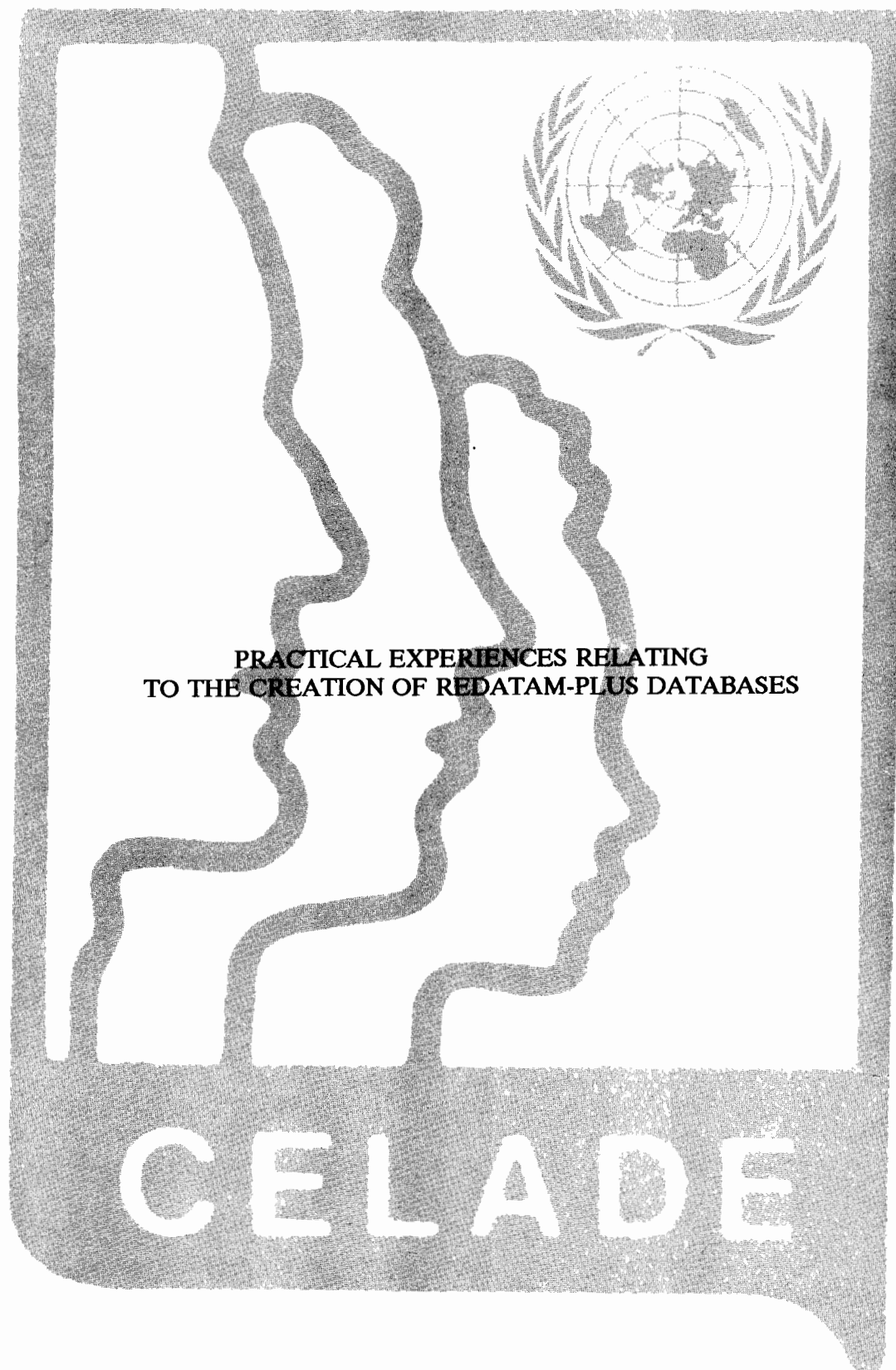


**NACIONES UNIDAS - CENTRO LATINOAMERICANO DE DEMOGRAFIA
UNITED NATIONS - LATIN AMERICAN DEMOGRAPHIC CENTRE**

Distr.
RESTRICTED

LC/DEM/R.158
Series A, No. 254
November 1991

ORIGINAL: ENGLISH



•

•

•

•

Distr.
RESTRICTED

LC/DEM/R.158
Series A, No. 254
November 1991

ORIGINAL: ENGLISH

C E L A D E

Latin American Demographic Centre

PRACTICAL EXPERIENCES RELATING
TO THE CREATION OF REDATAM-PLUS DATABASES

Document to be presented to the Interregional Workshop
on Population Databases and Related Topics
Harare, Zimbabwe, 9-13 December 1991

•

•

•

•



**PRACTICAL EXPERIENCES RELATING
to the CREATION of REDATAM-PLUS DATABASES**

Ari Silva and Arthur Conning

**Interregional Workshop
on Population Databases and Related Topics
9-13 December 1991
Harare, Zimbabwe**

**UNITED NATIONS / ECLAC
Latin American Demographic Centre (CELADE)
Casilla 91, Santiago, Chile
Fax: 208-0252 Tel: 208-5051**

•

•

•

•

TABLE OF CONTENTS

| | |
|--|----|
| INTRODUCTION | 1 |
| NEED FOR INTEGRATED DATABASES | 2 |
| DATABASE CREATION: PROBLEMS AND CONCERNS | 6 |
| SOME APPLICATIONS | 9 |
| GIS INTERFACE | 10 |
| BIBLIOGRAPHY | 12 |

•

•

•

•

INTRODUCTION

This document discusses the creation and usage of multidisciplinary REDATAM-Plus databases, requirements for their creation, existing applications in Latin America and the Caribbean, and their connection to the emerging geographical information systems.

The Latin American Demographic Centre (CELADE), in 1983, carried out an examination of public and private sector requests for population information made to their respective national statistical offices (NSO) in various Latin American and Caribbean countries. The major finding was that there was a strong unmet demand for tailored small-area population and housing information and that the NSO were unable to answer most requests within a reasonable period, because they did not have the programmers nor the mainframe time available for reprocessing the census microdata nor the tradition of providing such services.

To help alleviate this problem, CELADE began work in 1985, on the REDATAM system (REtrieval of DATA for small Areas by Microcomputer) to provide the NSO with the technology that would enable them, without using scarce programmer skills, to supply tailored population and housing census tabulations rapidly and at low cost. To permit the efficient creation of statistics for any small areas of a country or region, the system was designed to store compressed microdata in such a fashion that processing is very rapid because it is limited only to records and variables belonging to the area of interest, while maintaining available all records and variables to meet other user needs. There are now census and other databases using REDATAM version 3.1 in over 25 Latin American and Caribbean countries.

To overcome some of the REDATAM version 3.1 limitations and to broaden the spectrum of the system's applications, CELADE developed an entirely new version, named REDATAM-Plus (or REDATAM+), with all the previous features, plus other powerful tools like the ability to store multiple branches and variables at all levels of the hierarchy, interfaces with Geographical Information Systems (GIS), camera-ready tabulations, and much more.

For more information on REDATAM-Plus,
please refer to the attached description.

REDATAM and REDATAM-Plus were developed with grants from the International Development Research Centre (IDRC) of Canada and with additional assistance from the United Nations Population Fund (UNFPA) and the Canadian International Development Agency (CIDA).

NEED FOR INTEGRATED DATABASES

Databases with information from different sources, like a population census and an economic survey, or from different instances in time of the same source, like a periodic survey, are described by such adjectives as: combined, integrated, multidisciplinary, multi-purpose, etc. The objective in merging information is to broaden the scope for analysis, by including other variables from different fields. For example, in addition to accessing socio-economic and demographic data collected in national housing and population censuses and specialized surveys (such as agricultural and industry-specific surveys), regional and local governments are now concerned with establishing and maintaining information capabilities on diverse issues pertaining to health care, education, land use, resource management, land ownership and tenure, and infrastructure such as roads, sewers, electricity and power supply (Hall, 1991).

Furthermore, in many countries of the Latin American and Caribbean region there is a growing trend toward institutional and territorial decentralization. The predominance of the central government and capital city is lessening and regional and municipal institutions, to varying degrees, are obtaining an increasing participation in the planning and implementation of local development projects and the management of their own affairs. Among the many implications is that sub-national authorities will require suitable technology and training for organizing and utilizing information from many fields.

Dekker (1991) conducted a survey on census processing methods, and one of his findings was that "more than half of the responding countries have built, or plan to build, population databases". He also notes that "Databases of macrodata are more common than those holding unit record data". This fact can be explained by the sometimes tremendous amount of storage needed for the often millions of records of microdata, and by the apparent lack of any existing software to handle it.

Need for microdata when producing tailored population statistics for small-areas

Pre-conceived standardized aggregate tabulations on variables, when available for administrative areas, may be adequate for some purposes, but they are unlikely to be available for individual city blocks as is often required, and information usually has to be specially "tailored" for detailed analyses of specific problems or projects. Researchers and planners may also desire to "interact with the data" to derive improved indicators and understanding. In these situations, or when suitable aggregate data is not available for all levels of census geography, it is imperative to be able to **reprocess** the census **microdata**. In the past this has usually meant working on mainframe or minicomputers, since census microdata files for a whole country or large region or city may involve many millions of cases. This places the processing of census data in the developing countries out of the reach of most sub-national institutions, since they cannot easily draw on the computing power, resources or programming skills required for mainframe census data processing.

**Population and housing information:
the census as a source of small-area data**

Population and housing statistics are among the important types of information required for specific development activities at the regional and local levels, as well as at the national level. The population statistics required may be relatively conventional measures, such as the total population and density in each area of interest, or may be non-standard as a complex measure of poverty based on both housing and population variables, or estimates of numbers of family members according to their occupational status, attendance by type of educational institution, and their household's characteristics, such as might be employed in an urban traffic model.

There are frequently various sources of population information in the developing countries including specialized surveys, vital statistics, etc. However, normally only the population and housing census can provide data on each small area comprising any given region, since the census is designed to be universal and the person and housing records have geographical codes for areas usually down to city blocks or smaller. The data in the census for very small areas also allows *ad hoc* larger zones to be constructed for specific purposes.

REDATAM-Plus and integrated databases

The information in a REDATAM+ database is hierarchically organized in terms of entities (provinces, counties, population, housing, schools, hospitals, etc.). Each entity can be viewed as a rectangular matrix where the rows are the elements, which are the individual provinces for the "province" entity and the individual person records for the "population" entity; and the columns are the variables (attributes). Entities relate to other entities forming a tree structure, where each entity can have only one superior entity and one or more inferior entities. A REDATAM+ database can have a single branch of information (corresponding to a single original file) such as shown in Fig. 1 for a census geographical hierarchy, or be very complex, limited only by the capacity of the computer storage, such as shown in Fig. 2, which is the demonstration database for New Miranda, that comes with the software. For more details, refer to CELADE (1991, Chapter 12) and Silva et al (1990).

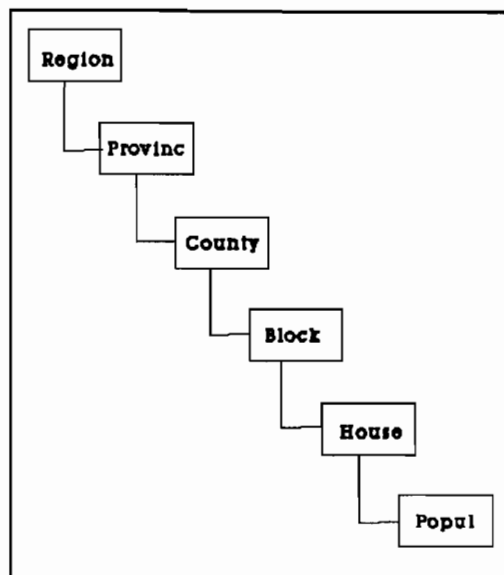


Fig. 1 A simple database structure

When combining different datasets in the same database, the way they will be placed in the database will depend upon the geographical hierarchies in the files. If they are the same, records in each dataset will be placed under the same branch. If not, new hierarchies (branches) can be defined.

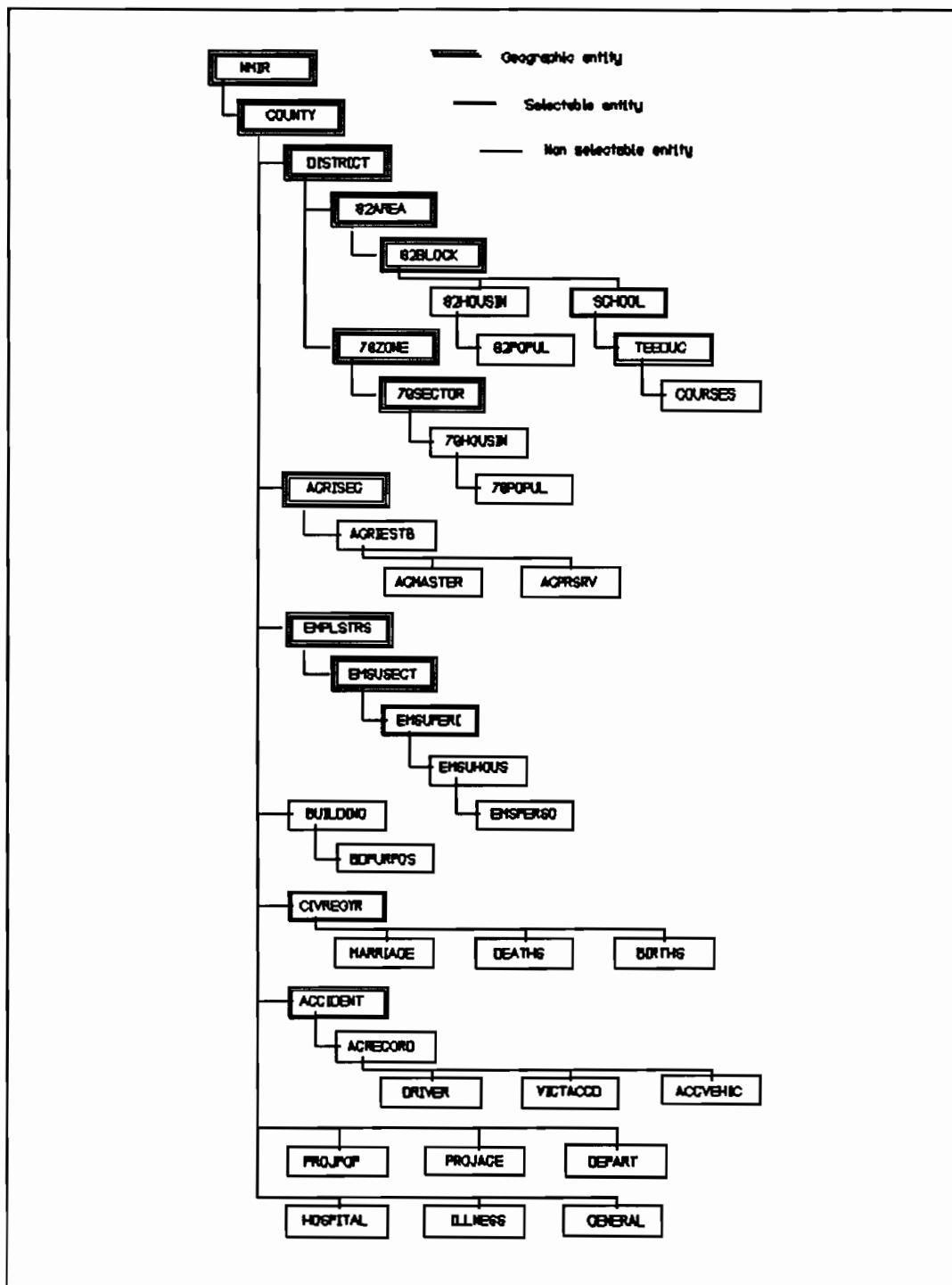


Fig. 2 A complex database structure

An example using an integrated database

Suppose we have three different datasets: a) population census data; b) marriages (from vital statistics); and c) housing construction records. Each can be manipulated to produce specific tabulations. The value added of a database integrating these three datasets under the same geographical hierarchy is that one can answer questions like: "How many houses should be built by region and economic status?". Using the census branch of the database we can calculate the initial deficit (from the number of different families in the same dwelling or household). The records of new marriages (meaning that they will need a home) are added to the initial count, and finally, the construction records are used to subtract what was recently built.

DATABASE CREATION: PROBLEMS AND CONCERNS

No matter which software is involved (REDATAM+ or other), the establishment of a database requires careful attention to the input data. Some aspects are very straightforward like the integrity and internal consistency of the original information (for example, in-range responses, consistency of two or more variables within the same record). The process to correct these errors may vary, depending upon the volume of the original information, from manual corrections in small surveys up to automatic imputations being applied by computer programs based on user-defined rules.

Another consideration that has to be taken into account is the definition of the database dictionary that will hold the metadata ("information about information"). The variables in the input file need to be defined, their categories and labels must be entered, and if possible, comments about each variable's behavior, imputation process, method of calculation (if it is a derived variable), etc. All this information in the dictionary needs to be precisely defined, in order to support a standardized usage of the database by different persons, with a minimum of outside documentation.

To be able to define the data dictionary, it is necessary to have some knowledge about the input file, at least the frequency distribution of each variable (to establish the categories), the number of cases of each record type (to control the creation process), and special characteristics like expansion factors.

If the database will have data from different sources, for example, different instances in time of the same survey, or a population and an economic census, the problems are much more complex, because other issues need to be taken into account. For "time series" surveys, a full standardization of codes is mandatory, to guarantee historical comparisons (for example, the codes for the variable "Roof material", must be the same and have the same meaning throughout the periods of the survey). Since a census can be viewed as a special case of a periodic survey, with a long interval (ten years on average), the historical comparison concern also applies, even if in weaker terms. Particularly important is geography, since many datasets use the same names for geographical areas, but the data can be compared only if the datasets refer to areas with the same boundaries.

Hence, the involvement of different sources of information in an integrated database requires solutions to complex problems such as: a) standardization of coding schemes; b) compatibility of criteria; c) universe of coverage; and d) geographical limits.

a) **Standardization of coding schemes:** Especially for the control variables (like the geographical ones that will provide the link between the various branches of information, or files), the codes for the variables must be the same (e.g., the identification codes for each department, county, zone, etc.), otherwise they will have to be converted to a common coding scheme before loading the database. These problems are frequent, particularly when the files are produced by different agencies.

b) **Compatibility of criteria:** For example, dwelling units and households might have different definitions in the census and in a survey, or conflicting definitions of unemployment for a person which is seeking his/her first job.

c) **Universe of coverage:** The scope and universe of different datasets may be different and not overlapping, such as a fertility survey that includes only women over 15 years of age, except the students living with their parents, a universe that is different from that used in the population census.

d) **Geographical limits:** Geographical mismatches are very common. This can occur not only in different periods of time, such as from one census to the next when the administratively lower divisions of a country might have undergone various degrees of restructuring during the intercensus period, but also in the same time period, such as census data according to political boundaries joined in a database with health data according to the areas of influence of the existing hospitals (that is, the country is subdivided in different ways).

REDATAM+ databases: special concerns

Besides all the considerations mentioned above, the creation of a REDATAM+ database involves additional concerns. The very first one is the sorting of the input records according to the geographical variables. The way REDATAM+ implements the data structure and the relationship between hierarchical levels (Silva, 1990) requires this sorting to insure that records corresponding to the same areas will be together. Embedded in this sorting process is also another decision to be considered, which is the definition of the geographical area levels that will define the database hierarchical levels of information. That is, which are the subdivisions of the country that will be used for data storing and querying within REDATAM+.

Another decision that may have to be made refers to redundant information in the database. By redundant information we mean data that can be derived from other variables in the database, for example, the income per capita of a municipality (if the database stores the income for each person), or the age distribution of the persons given by quinquennial age groups (if there is a variable with the single year age of the persons). As these variables can always be recalculated, they waste storage space. However, it may be advantageous to have them stored directly in the database,

either because they may be difficult to calculate, or because they are needed very frequently. There is a trade-off between space and processing time.

If the database to be created will contain information from a single source of data, and if the original file is already stored in a microcomputer, the one-time activity to create this database is very simple, involving the definition of the data dictionary and the execution of an automatic loading feature. However, in the case of large files (like a census) that do not fit on the microcomputer hard disk, this task is more complex and time consuming, and it may require a skilled programmer/analyst or external assistance. An alternative that is often used is to partition the original file into logical pieces (like departments or regions), and then generate each part, finally appending them to make a single database, or generating them separately to produce an individual database for each region.

The paragraphs that follow apply only to integrated (or multidisciplinary) general purpose databases, where the need to integrate information from various sources imposes other complexities that must be dealt with.

Needless to say that the very first question to be addressed is which information to enter into the database. Some relationship must exist between the data in the files in order for them to be considered as participating in the integrated database, otherwise it is worthless to associate them, since no one will ever use them as such. This is not as simple as it might look, primarily because even if you cannot think of any application using the combined data, just the fact that they can be "seen" together in the same database can trigger somebody else's imagination.

The crucial point when integrating different files in the same REDATAM+ database, if they will share the same geographical subdivisions, is to establish the universality of these subdivisions. For example, when integrating the data from both the population and the economic establishment censuses down to the block level, even if they use the same coding scheme, there will be city blocks with no information on households (because there are none) and vice-versa. This means that if one uses the population census as the basis for the hierarchical structure of the REDATAM+ database, some blocks will not appear in this structure, which will inhibit the creation of the other branch of information (the economic data) for these blocks. In order to combine both datasets, it is necessary first to merge the geographical structure from both files, creating an output that will contain the universe of blocks in the country.

Last, but not least, it is important to note that REDATAM+ is not meant to deal with volatile information (files that change very often). You can append information to an existing database (such as from another year of the survey), but there is no facilities in the REDATAM+ system to update (or delete) specific records and/or variables in the database, such as in general database packages (dBase, Paradox, etc.). On the other hand, REDATAM+ is much more efficient than these for storing very large masses of "frozen" information that do not change in time (for example, the 1980 census will not change, even though the population in the country might have increased), and for accessing and processing selected areas very rapidly.

SOME APPLICATIONS

Adoption and use of REDATAM has become fairly widespread and databases now exist for one, and in some countries two, censuses in eleven countries in the Caribbean and at least thirteen countries in South and Central America. Although REDATAM was designed primarily to hold census data, its processing speed and ease of use have led it to be applied to smaller surveys and other data processing. As REDATAM+ has not been officially released yet, most of the examples given below refer to REDATAM version 3.1.

Costa Rica

The 1984 population census database was downloaded into six parts, corresponding to the Ministry of Planning's administrative regions. These sub-databases were installed in each regional office to attend to local consultations.

A vital statistics database was made with the records of the deaths occurring in different years.

The Ministry of Planning built a database with information about the "Road System" to identify road sections that need maintenance. The elements of this database are the road segments, hierarchically stored according to the country's administrative divisions, with data such as length, type of pavement, other classifications like responsibility for maintenance (federal, municipal, etc.), number of lanes, etc.

Chile

An integrated REDATAM database at two instances in time of a periodic household survey has been used to build social-demographic indicators (Bravo, 1990). Use of the data is facilitated because the survey maintains a uniform coding scheme through time, it is easier to handle conceptually the data. The most obvious advantage of such a database is the easy and simultaneous access to all the information in the survey, corresponding to different moments in time. The interesting feature here is that time is used to define a hierarchical level.

Uruguay

Using the REDATAM 3.1 version, the National Statistical Office (NSO) generated both 1975 and 1984 population census databases, and the 1988 economic census. These databases are being accessed by the users through a Local Area Network (LAN).

They also designed an interface from MAPINFO (a GIS package) to REDATAM+, that makes use of the mapping display facilities of the GIS to select the areas of interest by drawing a polygon line around the blocks or zones of interest which can then be analyzed using the REDATAM+ Statistical Processor (Greising, 1991).

Venezuela

The NSO plans to create a REDATAM+ database on CD-ROM disks with the 1990 population and housing census, for data dissemination and decentralization.

GIS INTERFACE

Dekker (1991) stated that "Another recent development is the popularity of geographic information systems (GIS), which link various items of information on the common denominator of geographic location. Such systems usually are based on vector maps of the areas covered. Data pertaining to certain nodes, line segments or polygons on the map may be retrieved from an associated database, and displayed overlaid on the map. Such information can have any nature, provided it is attributable to a geographic structure. Statistical indicators are frequently used as information components of such geographic information systems. Several statistical offices have published geographically-oriented population databases on CD-ROM. Such databases may contain results of one or several censuses and surveys".

On the same line of thought, ESRI (1990) notes that "Geography (and the data describing it) is part of our everyday world; almost every decision we make is constrained, influenced or dictated by some fact of geography. Fire trucks sent to fires by the fastest available routes, central government grants are often awarded to local governments based on population, and diseases are studied by identifying areas of prevalence and rate of spread".

As many regional and local projects have aspects involving the spatial distribution of population, resources, infrastructure, etc., geographic information systems (GIS) can be employed as "information integrating machines" (Dangermond, 1988), taking advantage of the underlying geography to "relate" data from many different sources and fields, while also facilitating spatial display and analysis of the data. In the developed countries, both the public and private sectors increasingly recognize the benefits of using GIS (Ibid; Tomlinson, 1987), but this is only incipient in most Latin American and Caribbean countries. Until recently, GIS software was very expensive and required large computers, placing GIS out of reach and out of mind for most national as well as sub-national institutions.

Another common denominator and perhaps often the most important one, as suggested above, is population data; much done nowadays can be defined in terms of population. Things are constructed by people to be used by other people: industrial, commercial and agricultural activities, services (power, gas, water, phone, etc.), education, health, etc.

The first instance of an application of REDATAM+ with a GIS involves an ongoing pilot project being developed by the Ministry of the Interior of Chile to focalize social investments for the poor by detecting pockets of poverty within municipalities and then identifying the consequent social investments required in each depending on the specific nature of the poverty in each pocket detected. This work utilizes a REDATAM+ databases containing 1982 population and housing data for each of the municipalities of interest (downloaded from a larger census database), and data is

being added to each database from two specialized surveys. Using a basic needs approach to poverty, indicators have been calculated along with other information for each city block (the study will later consider rural areas) and these block-level aggregate indicators have been exported from REDATAM+ to the pcARC/INFO geographic information system to identify the boundaries of the poverty pockets through spatial display and analysis. The analysis then returns to REDATAM+ for more detailed examination of the needs of each of the identified areas, which will be compared with the existing social infrastructure to determine what and where the new social investments are desirable. See Gonzalez et al, (1991) and Silva, (1991).

The needs in the developing countries, particularly of regional and local authorities in the public and private sectors, are not often for software as such, but rather for solutions to specific recurring problems. Consequently, a major strategy for stimulating the integration of population in development planning and fostering increased use of the 1990-round of census data, as well as for increasing the dissemination of REDATAM+, will be through applications of REDATAM+ multi-disciplinary databases with different GIS packages to specific problems such as the detection of localized poverty and the analysis of the encroachment of urban growth on agricultural land. See Hall, et al, (1991) for further details on this strategy.

BIBLIOGRAPHY

- Bravo, Jorge, 1990, **Towards a System of Social-Demographic Indicators Based on Periodic Household Surveys: A REDATAM Application**, CELADE, Santiago, Series A-204 LC/DEM/R.104. [original in Spanish].
- CELADE, 1991, **REDATAM-Plus User's Manual**, CELADE, Santiago (November 1991). Series A-201 LC/DEM/G.90 [also in Spanish].
- Dangermond, Jack, 1988, **Trends in GIS and Comments**, Computer Environment and Urban Systems, Vol. 11, pp.137-159.
- Dekker, Arij, 1991, **New or Uncommon Computer Methods in Population Census Data Processing**, 48th I.S.I. Session, Cairo, 9 to 17 September 1991.
- ESRI, 1990, **Understanding GIS, the ARCINFO Method**, Redlands, California.
- Gonzalez, Juan Carlos, et. al., 1991, **Census Data Processing with REDATAM-Plus for GIS: An Application for the Detection of Poverty Pockets in Chilean Municipalities**, presented in the Third Latin American Conference on Geographical Information Systems, Viña del Mar, Chile, 21 to 25 October 1991.
- Greising, Edgardo, 1991, **Utilization of MAPINFO as a Graphic Interface for REDATAM-Plus**, presented in the Third Latin American Conference on Geographical Information Systems, Viña del Mar, Chile, 21 to 25 October 1991.
- Hall, Brent, Conning A., Silva A., Poulard S., 1991, **Development of Demonstration Projects Integrating Socio-Economic And Demographic Data with PC-Based GIS in Latin America and the Caribbean**, presented in the Third Latin American Conference on Geographical Information Systems, Viña del Mar, Chile, 21 to 25 October 1991.
- Silva, Ari, Conning A., et. al., 1990, **The REDATAM-Plus Model: Database Architecture and Philosophy**, June 1990, Revised version of a paper presented in the "Seminar on Use of Databases", Cuernavaca, Mexico, 27 November to 1 December 1989.
- Silva, Ivan, 1991, **Methodology for Detecting Intra-municipality Pockets of Poverty and Opportunities for Investment**, 1st progress report, 8 November 1991.
- Tomlinson, R.F., 1987, **Current and Potential Uses of Geographic Information Systems: The North American Experience**. Paper based on a report published within the report: **Committee of Enquiry Chaired by Lord Chorley**, 1987, Handling of Geographic Information (London: HMSO), Appendix 6.



REDATAM Plus

REtrieval of census DATa for small Areas by Microcomputer
Software, sample database and manuals in English and Spanish

What is REDATAM-Plus

REDATAM-Plus (or REDATAM+) is a user-friendly interactive software system designed to make it convenient for users to obtain small-area information required for decision making, modelling and related purposes from large census datasets and similarly structured data, organized in a multidisciplinary hierarchical database. REDATAM+ stores compressed microdata of one or more entire population and housing censuses that may involve millions of cases, or the data of a region or city, on an IBM compatible micro-computer; the database can also include aggregated data at any level of geography. REDATAM+ produces tabulations, usually within minutes, for any variables and for any small geographical areas specified by the user, down to city blocks or groupings of such areas.

Planners, researchers and other substantive users can work with REDATAM+ without assistance from programmers, once the database of interest has been created. So that time is not lost processing the large amount of data that are always readily available in a REDATAM+ database, the system processes only the data in the areas selected by the user and only the variables of interest.

While the first databases created using the earlier REDATAM system were oriented to population and housing censuses, the system now supports any kind of hierarchically structured information from various censuses, surveys and other sources forming a real multipurpose database.

What does REDATAM+ do?

The following outputs can be generated rapidly from a large REDATAM+ population and housing census or other database for any geographical areas defined by the user:

- Cross-tabulations with up to four variables.
- Averages with up to four classification variables.

- Marginal frequencies.
- Export of the data for the user-defined areas as a DOS flat file or with complete dictionary parameters for direct input into SPSS-PC, LOTUS 123, etc.
- Export of self-contained REDATAM+ sub-databases for areas defined by the user, permitting work on smaller machines or decentralization of data to sub-regional or municipal offices.
- Export of database variables or results of internal processes to be manipulated by GIS packages.

For any output:

- Results can be obtained for sub-areas as well as for the entire user-defined area. The selection can be based on quantitative criteria (such as every district with more than 10,000 households) as well as on names and codes of the geographical entities.
- Cases can be weighted.
- Variables can be recoded.
- New variables can be derived.
- Subpopulations can be selected.
- Processing can be hierarchical, for example, within each household (or block, or county, etc.), new aggregated variables can be derived, such as the number of males attending secondary school by household income category.
- Command sets and geographical area definitions, can be saved for later use or edited to introduce changes.
- The number of cases can be limited or may be a sample.
- The database can be protected by passwords.
- Processing can be executed interactively or in batch.
- Network operation to allow multiple use of a database.
- Camera-ready output as an option.
- Import of Integrated Microcomputer Processing System (IMPS) dictionaries to permit the direct creation of a REDATAM+ database after data entry, editing and processing, for example, a census with IMPS.

Extensive help is available on-line. There are also extensive management features to facilitate the construction and editing of REDATAM+ databases and the addition of new variables or geographical areas to existing databases. A REDATAM+ database is compressed to around 25% of the storage needed for the input data.

Inputs required

The census and/or other datasets of interest must first be converted into a REDATAM+ database. While users can obtain results from an existing database without assistance, a programmer normally is required to make a large census or a multidisciplinary database. The one-time creation of a REDATAM+ database requires a very clean set of microdata, codebook information on each variable and the geographical hierarchy and associated names. Use of the database may require maps to locate the geographic subdivisions of interest.

For demonstration purposes, the REDATAM+ software package comes with a ready-made multidisciplinary database for the hypothetical country *New Miranda*, that includes, among other types of data, two population censuses with around 50,000 persons, an agricultural survey, aggregate data on the number of students in schools in each area, projections, health statistics, traffic accidents and many more.

Salient strengths

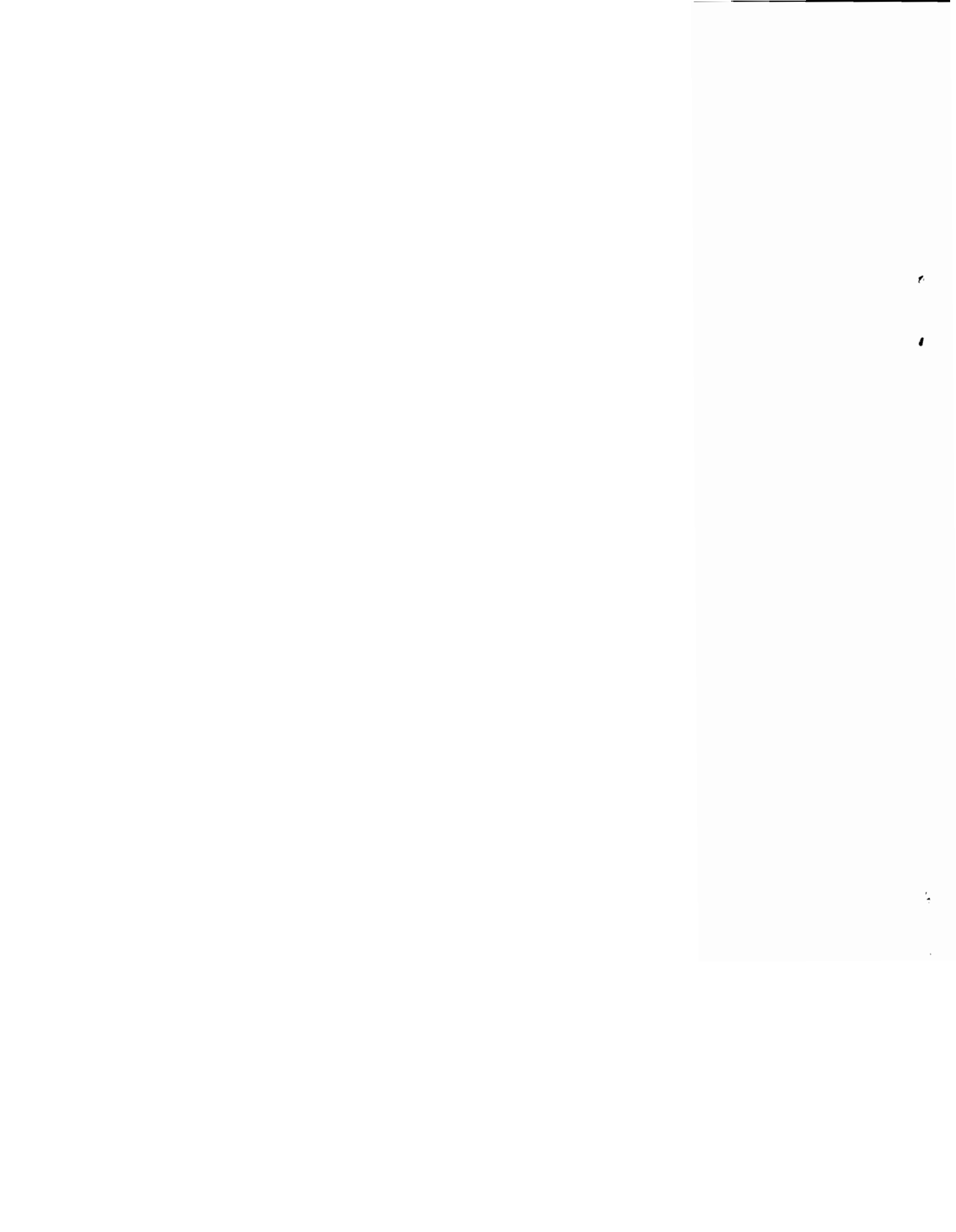
REDATAM+ makes it practical for a user with a microcomputer to have readily available the microdata of an entire hierarchically structured population and housing census (or other types of information), to define any geographic areas of interest and to produce rapidly any tabulations for the selected areas, which may have a few hundred records to hundreds of thousands or more.

REDATAM+ reduces significantly the amount of disk storage needed for the database compared to conventional methods. On average, a REDATAM+ database occupies 25% of the space required for the original dataset.

The REDATAM+ internal dictionary for the variables provides all the advantages of a meta-database (uniformity of documentation, less programming needed, etc).

Multidisciplinary multi-level database capabilities. Various hierarchical levels of microdata are permitted along with aggregate variables (generated from the microdata or imported from other sources) describing the geographical entities. Thus data can be from many different sources. To help the user visualize a database, REDATAM+ shows a hierarchical tree diagram.

Geographic information system (GIS) interfaces, including export of REDATAM outputs to pcARC/INFO for cartographic display and spatial analysis. Interfaces to other spatial display and GIS packages



0
1
2

3
4
5

07101:RLD
**** ARCHIVO DE DOCUMENTOS (CARMEN VERA) ****
BIBLIOTECA
CEPAL

CLI