

Al Día Redatam Informa

*Software for processing and map data from censuses
and surveys for local and regional analysis*



LC/L.3429

Vol. 17

1

December 2011

25 years of REDATAM software development

The dawn of widespread computer use

Today's cellular telephone is one thousand times faster and more powerful than the computers of the 1970s. In those days, computers were enormous machines that weighed tons, with very little internal memory, which normally carried out processes according to instructions from punch cards. They stored large volumes of data on magnetic tape, which, because of its design, was read linearly.

Generally, population censuses contain millions of records and variables (microdata); they were managed and stored in these enormous machines, and they required a programmer to produce the tabulations needed for subsequent analysis. Because of the cost and time to read tapes with census microdata, it was almost mandatory to create, print and publish a series of documents with the official tabulations at the country level and for large administrative divisions. This limited census analysis to defined or limited topics, and it

A publication from
CELADE,
Population Division of
ECLAC



NACIONES UNIDAS

CEPAL

Table of contents

- 1. 25 years of REDATAM software development....1
- 2. REDATAM 7: speed, efficiency and ease of use..... 13
- 3. On-line census.....14
- 4. Products of the Redatam Family.....15
- 5. REDATAM 7: speed, efficiency and ease of use..... 14
- 6. Potential of census information and REDATAM in the twenty-first Century.....16
- 7. Interview with Jose Miguel Guzman, Chief of the Population and Development Branch, UNFPA..... 16
- 8. Interview with Byron Villacis, Director of INEC of Ecuador.....18
- 9. Interview with Edwin St Catherine, Director of Statistics at the Central Statistical Office of Saint Lucia 19
- 10. How to obtain Redatam+ SP..... 20

reduced the possibility of exploring new analysis lines. Given the difficulties in programming and producing special tabulations, and the fact that only National Statistical Offices (NSO) have census microdata, it was nearly impossible for researchers to obtain additional specific tabulations. There was virtual no feedback for data and analysis.

In this technological context, and thanks to the economic backing of the International Development Research Centre (IDRC) of Canada, Arthur Conning, head of Information and Data Processing of the Latin American Demographic Centre (CELADE) –the Population Division of the Economic Commission for Latin America and the Caribbean (ECLAC)– visited different national statistics and planning offices and other institutions of Latin America and the Caribbean, in pursuit of a project for using bibliographic descriptions to identify and locate existing official and specific census tabulations and publications. The inspiration for this idea was the success of DOCPAL (Population Documentation System for Latin America and the Caribbean), which created and used a bibliographic database of summaries and keywords in a mainframe computer.

Serendipity: Fortuitous discovery

Conducted in several countries in the region, the findings of this feasibility study were conclusive: The proposed plan DID NOT solve the real census information management needs of the region. NSO interests pointed in another direction, toward:

- Having the capacity to generate specific census tabulations which were not considered in the official findings published for each census, and
- Having census information available for disaggregated divisions outside of the country’s large administrative areas, and even for territories that do not necessarily fit into any official administrative arrangement (the conclusions of Box 1 provide more details).

The conclusion was clear: the original project had no future, but there was the unintentional and serendipitous discovery of a much more interesting and significant challenge for the region at that time.

Box 1

Main results of lifting NSO requirements (1983)

Trinidad and Tobago, Saint Lucia, Costa Rica, Plurinational State of Bolivia, Chile, Peru and the Brazilian State of Sao Paulo

- 1.- To consider the spatial characteristics and distribution of the local labour supply and of the population that would be positively or negatively affected by the implementation of a specific project.
- 2.- Summarized demographic data are rarely enough for planning a spatially localized project (for example, determining the number of maternity beds needed for a specific municipality).
- 3.- In the majority of developing countries, population censuses are the only source of available demographic information, with numbers of households and people, large enough to obtain useful pictures of any small geographic area of the country.
- 4.- To build areas of analytical interest by aggregating the census information of different, smaller areas.
- 5.- It is not very practical for statistical offices to publish volumes or keep tabulations of every possible scenario that could happen in the future for every small area in the whole of the country (many Latin American countries have more than 100,000 divisions of census enumeration).
- 6.- If existing tabulations cannot provide required information, the only recourse is to go back to the microdata and reprocess the information.
- 7.- The majority of statistical offices in Latin America and the Caribbean do not have the equipment for quickly reprocessing census microdata at a low cost. Other pressing duties, in addition, delay the response to users.
- 8.- Even when the majority of statistical offices considered in the study are able to produce census tabulations, there are often months-long delays or the cost is prohibitive for most users.
- 9.- The solution must eliminate dependency on programmers and large computers, which would be possible if users had direct access to low-cost microcomputers.

The response: inverted vectors, hierarchal organization and microcomputers

The origins of REDATAM date back to the early 1980s. By that point, CELADE had the relevant experience and could offer a technological response to the needs mentioned above. The era’s interest in studying fertility changes and their consequences on the people’s quality of life led to, among other initiatives, the World Fertility Survey (WFS). To make it easier to use these surveys, CELADE, toward the end of the 1970s, developed a software for mainframe computers, based on the “Rapid” program (Statistics Canada www.statcan.ca), combined with the survey processing capabilities of the SPSS program (Statistical Package for the Social Sciences), whose innovating feature was the ability to invert and then “read” questionnaire matrices and their variables.¹ Inverting variable files made it possible

¹ Most database information processing applications “read” records horizontally, i.e, they review the values of each entry in their entirety and “conserve” the value of the variable of interest. In the case of inverted vectors, the software only reads the variable in question, saving resources and time, speeding up the attainment of the process’s final result.

to read the data much faster. By only accessing variables of interest, inversion saves significant time in the process, and, with SPSS, it was possible to produce tabulations and statistics using these data. The creation of this computer program was an important step forward in census microdata processing for Latin American countries, and its development kept pace with other technological breakthroughs in computing.

In the early 1980s, the penetration and use of personal microcomputers in Latin American countries began to grow. Early on, they were much slower than mini and mainframe computers, but with the CELADE-developed strategy of reading the file variables in reverse already receiving recognition, it was possible to accelerate processing times, with programs obtaining a specific tabulation by only accessing required and chosen variables and entries. This processing logic, in addition to the advances made in personal computer technology, opened up a new horizon in the development of census microdata processing tools.

With this step in computer development, on top of the potential for censuses to produce information geographically disaggregated into small units –of people, of households, and of dwellings in a given country– the REDATAM software developing project was born. The name comes from the acronym for REtrieval of DATA for small Areas by Microcomputers.

Figure 1:
Logos used in REDATAM versions



The philosophy behind the development of this software was to offer a user-friendly tool that was accessible to all public policy and population researchers and decision makers.

Originally programmed for the DOS operating system, REDATAM used a simple programming language of indicators and tabulation generation, expanding access to a broader group of users, in addition to computer programmers. The first version emerged between 1985 and 1986, thanks to the resources of IDRC Canada, the support of United Nations Population Fund (UNFPA), and the Canadian International Development Agency (CIDA). The

first version, released in 1986, was called REDATAM DOS. Today, 25 years after the initial launch, we are on the verge of releasing REDATAM 7, a version which, by retaining the original logic and structures, takes all current computer and technological capabilities and converts them into a unique, reliable and easy-to-use program for processing and analysing large bases of microdata, extending beyond the borders of Latin America and the Caribbean and reaching other regions of the world.

To optimize use of census or other data, REDATAM databases are normally built with microdata, that is, with variables which refer

to individuals, households or other analytical elements; these variables can then be combined in tabulations to produce findings for any geographic area, as defined by the user. The data are organized in a hierarchy, making it possible to derive, for example, new household variables based on the number or characteristics of the individuals who live in each home. It is possible to select any series of subareas to process only the chosen data subset, accelerating, in turn, the aggregate calculation needed to generate various types of tabulations. Furthermore, any derived variable can be stored permanently in the database for future use. No prior

| | | |
|--|---|---|
| <p>Ageing and Development</p> | <p>Intergenerational Transfers, Population Aging, and Social Protection.</p> | <p>Indigenous and Afrodescendant Peoples</p> |
| <p>International Migration and Development</p> | <p>Censuses of 2010</p> | <p>Internal Migration & Spatial Distribution of the Population</p> |
| <p>Visit us at our web site: www.cepal.org/celade</p> | | |

programming experience is needed to quickly obtain tabulations and other statistical results, since it has a graphical user interface. The program contains resources for processing and creating maps connected to database geographical levels. This makes it possible for users to see on-screen a spatial analysis of the aggregated statistics that the system produces in any geographical area. The visualization functions, in addition to data manipulation and high-speed processing, provide fast access to data and increase their value. Its Web Server modules, moreover, make it possible for users to access online microdata processing, thus facilitating information dissemination.

25 years of REDATAM development

Since its beginnings, REDATAM has met users' needs for more functions and easier tabulation processing; it has done this by staying in step with the development of personal computer technology.

In general terms, REDATAM, during its evolution, has remained accessible software with a user-friendly interface, in terms of ease of use, cost and distribution (it started out as low priced software and now is totally free), essentially designed to work with large bases of microdata, such as housing and population censuses. According to this logic, and since its inception, census data confidentiality has been maintained and the information of

each person, household or dwelling protected, pursuant to statistical secrecy norms; the internal structure of REDATAM databases reflect these efforts. Over the course of its existence, more functions have been added, such as the option of obtaining new tabulations and the possibility of processing other sources of information, for example, farming censuses, household surveys, industry surveys, administrative records, foreign trade, youth surveys, and so forth.

Figure 2:
Menu screen of REDATAM Plus version, 1991.

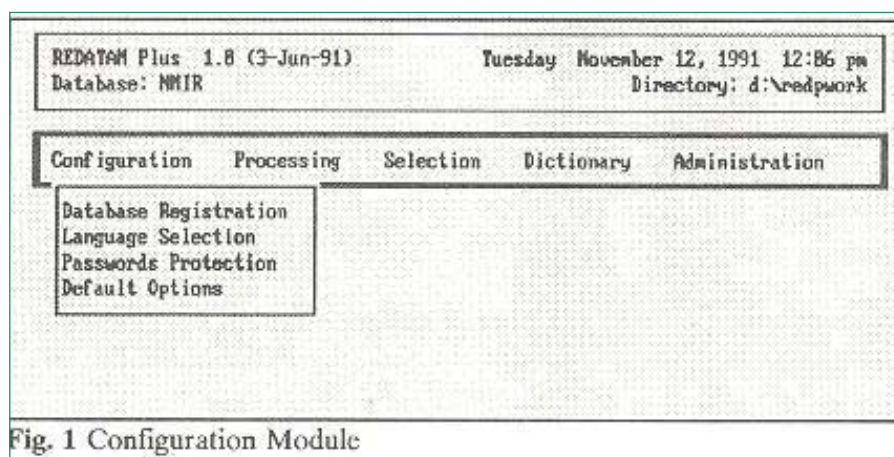


Fig. 1 Configuration Module

An important characteristic of the software's evolution is the concurrence of published REDATAM versions² and census rounds (every decade) and the emergence of new computer technology, thus we have:

| REDATAM Version | Generation | Publication | Operating system, environment, platform | Census round |
|-----------------|----------------|-------------|---|--------------|
| REDATAM DOS | 1st generation | 1985 | DOS | 1980 |
| REDATAM 3.1 | 1st generation | 1986 | DOS | |
| REDATAM Plus | 2nd generation | 1991 | DOS | 1990 |
| winREDATAM+ | 3rd generation | 1997 | Windows | |
| REDATAM+G4 | 4th generation | 2001 | Windows, WebServer | 2000 |
| REDATAM+SP | 4th generation | 2004 | Windows, WebServer | |
| REDATAM 7 | 5th generation | 2012 | Windows, WebServer, 64 bits, multilingual Unicode | 2010 |

REDATAM DOS and REDATAM 3.1 (1st generation)

This was the first version of the program to come out in 1985; in 1986, it was presented to the public as REDATAM 3.1. REDATAM DOS's main characteristics were:

- It allowed access to and processing of millions of records, such as households

² The REDATAM Family: <http://www.cepal.org/id.asp?id=7395>

and population censuses, disaggregated geographically and organized at the levels of dwellings, households and individuals, with data confidentiality ensured.

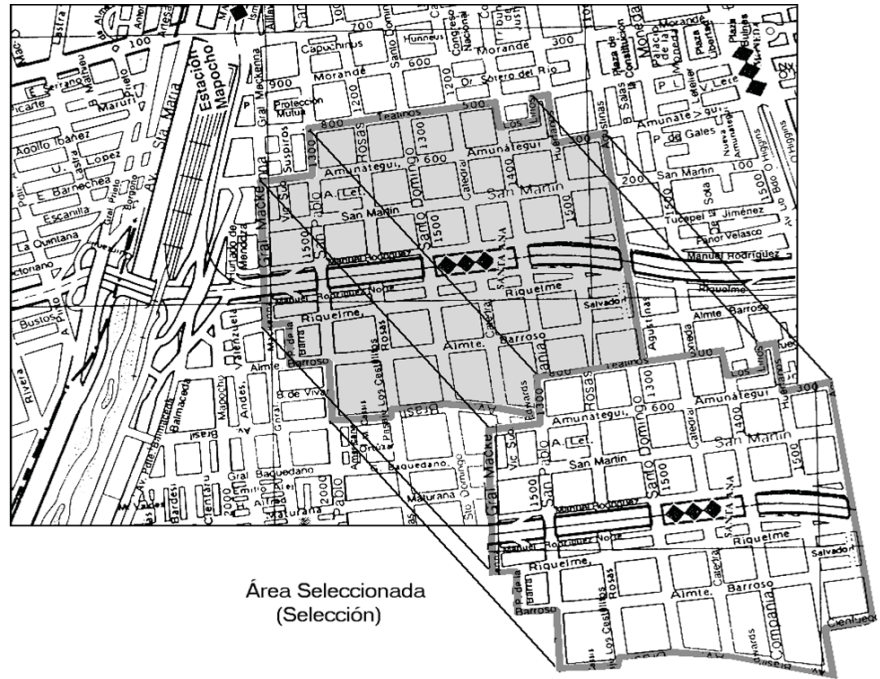
- The microdata internal organization is based on a hierarchal structure, making it possible to easily select geographical areas ad hoc, with even more specific geographical disaggregation, such as a census block.
- Since its inception, it promoted the democratization of population information and the widespread dissemination of census databases.
- The program's growing use and the identification of its capabilities in population information analysis created the conditions for devising and designing new ways to combine census microdata with other sources of external information, for example, data on educational equipment by district or municipality.

**REDATAM-Plus
 (2nd generation)**

The following version, REDATAM-Plus, debuted in 1991. This version was rewritten with the basic functionality of its predecessor completely intact, but with several new and important innovations:

- Multisector databases: The capability of incorporating other databases into a unique branch REDATAM base, designed with the same hierarchal and geographical structure as the original; examples of additional databases include a different population census, education records, school records, agricultural censuses, youth surveys, health statistics, etc. With this, it became possible to incorporate various branches and

Figure 3:
REDATAM makes it possible to focus on small areas



- variables at any level of the base structure. A database diagram was included to help understand the multisector base structure.
- More microdata organizational levels became available (the previous version only allowed dwellings and persons).
- A statistical processor with new and better commands. There were new commands for creating variables at higher levels (for example, counting the number of female engineers in each district) to be incorporated into databases. Table output was improved.
- Improved interface with Geographical Information Systems (GIS) for creating thematic maps, using REDATAM outputs in dBase formats.
- It became possible to import IMPS data files and dictionaries, which in turn made it possible

to create databases directly in REDATAM.

- Multi-user databases: it became possible for more than one user to simultaneously work with and access databases, using a local network.
- REDATAM-Plus for Windows
 (winR+) (3rd generation)**

In 1997, when Microsoft's Windows Operating System became widely disseminated and used in personal computers, REDATAM Plus, for Windows 3.1 and Windows 95, was released. The version was completely rewritten in another programming language to take advantage of Windows' Graphical User Interface (GUI). It is more powerful, more open and faster than its predecessor, REDATAM-Plus for DOS. This new breakthrough

was framed in the “Applications of winR+SIG to help the decentralized local development” project, between CELADE and Canada’s University of Waterloo, with funding primarily from CIDA. This project led to the development of various tools related to the population’s distribution and spatial analysis, in support of decision-making. CELADE developed REDATAM Plus (winR+) for Windows and the ZonPlan module, for identifying target populations in small areas for social policy, while the University of Waterloo created three tools: AccessPlan, for analysing health-care resource accessibility and allocation; EduPlan, for the same purpose in the area of education; and TourPlan, for assessing the impact of tourism projects on the population and the environment. These three tools shared a common denominator: to obtain population information from areas that REDATAM had disaggregated.

The following are among some of winR+’s main differences with previous versions:

- Introducing the Windows environment improves the user interface by providing interactive windows, dynamic menus and the mouse.
- Processing speed increased significantly compared to previous versions.
- A new clearer, simpler language was developed to allow the user to process tabulations and generate indicators with just three commands (RUNDEF, DEFINE and TABLE).
- All winR+ users create their own “work space,” with a personal copy of their own data dictionary and documents, including their own processing programs, maps and geographic selections.
- Makes it possible to use external databases (dBase) and annex them to a REDATAM base.
- Data access on different hierarchal levels improves, which, in turn, makes it possible to derive results from two or more levels with greater ease than with other software.
- Creating a selection file becomes simpler, done in graphical form. One choice can be made up of any number and type of areas. Obtaining values of one or more variables for every subarea contained in the selection (called “area lists”) becomes possible.
- This version has Spanish and English interfaces, and it is possible to change from one to the other at any time.
- A cartographic deployment window is introduced in REDATAM itself, simplifying the creation of thematic maps.
- The ZonPlan module was developed –to identify target populations in small areas for public policy– allowing users to choose predefined socioeconomic and demographic indicators from a list, thus automatically generating results tables and maps.
- Microdata confidentiality and security are ensured. External databases are typically converted into REDATAM format, which compresses, encrypts and inverts the original data to combine information confidentiality with efficiency. This process provides certain security controls for sensitive data. There are three types of protection: first, during the creation of a REDATAM

Figure 4:
A winR+ working session using GUI in 1997.

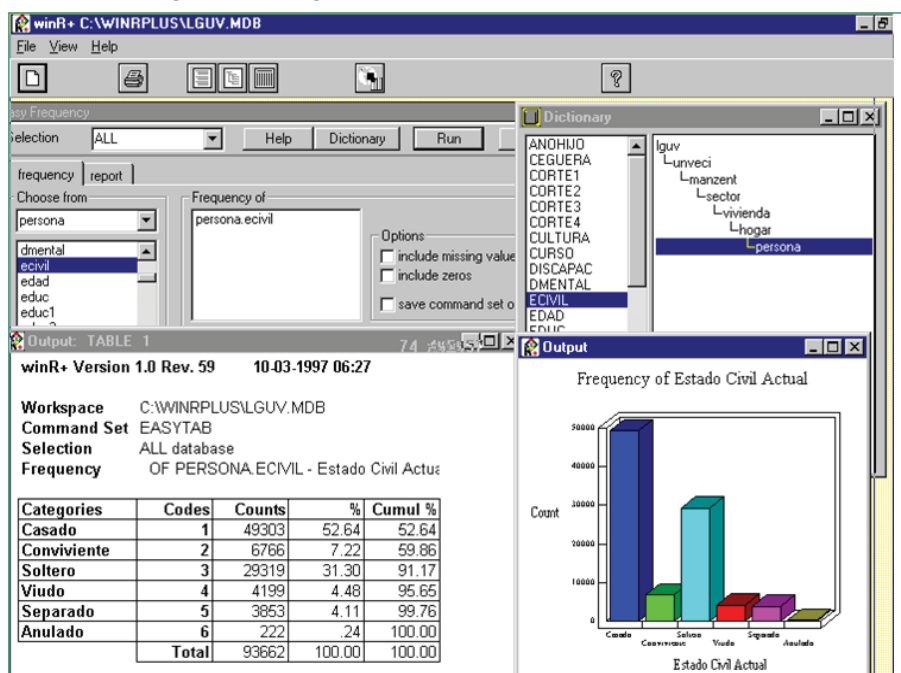


Figure 5:
 Thematic map generated with winR+ at the census block level
 (1992 Census of Chile, processed with winR+)



base, dwellings, households and persons are defined as “unselectable,” eliminating their identification, while the geographical areas that link them, such as blocks, districts, etc., are declared “selectable” with codes and names; second, there is a system of “key words” which reserve access for specific users; and third, database vector files are inverted and encrypted to prevent any direct access to them.

REDATAM+G4 (R+G4)

This fourth generation of

REDATAM, launched in 2001, makes use of the 32-bit Windows 95, 98, 2000 and NT4.0, which existed at the time. Its improvements over the previous version (winR+) include processing efficiency, speed and microdata storage.

It also incorporated for the first time a module structure for REDATAM’s main processes:

- R+G4 Create for creating databases from other formats (this version accepts IMPS, ISSA, dBase and ASCII formats) in REDATAM format,
- R+G4 Process for statistical

information processing,

- R+G4 WebServer, the last member of the REDATAM family appeared in 2002 and made it possible to consult REDATAM databases over an Intranet or Internet network. Using the REDATAM WebServer engine lets all users process census microdatabases online.
- R+G4 xPlan for generating interfaces with predefined indicators and controlled database access (former ZonPlan).

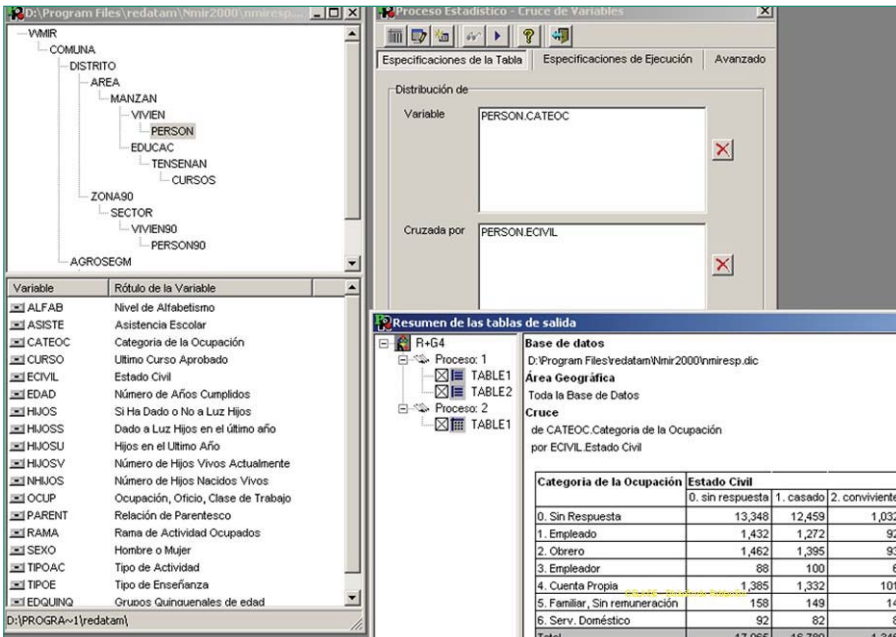
The R+G4 and Xplan applications offer access control for information contained in a R+G4 base, through a graphical interface, where users are not required to understand the internal mechanism or operations of the application; they only need to manage the parameters with a user-friendly and easy-to-use interface. These applications are available on CD separately or uploaded to the Internet to provide direct access to all online users.

In turn, there is improved visibility of database variable thematic maps and of area-specific generated indicators. This is a very important characteristic for planning, given that maps show straight away if the areas have a similar value or if the areas are different, according to the assigned variable.

It has an interactive assistant with hyperlinks similar to Windows.

In addition to Spanish and English, this version offers interfaces in French and Portuguese, and changing from one language to the other can be done at any time.

Figure 6:
REDATAM+G4 session with dictionary open, simple frequencies window and results table



REDATAM is now free and available online, allowing all users to acquire and update it at no cost.

REDATAM+SP

Finally, we have reached the current version, REDATAM+SP, officially launched in 2004. The most up-to-date version of REDATAM+SP is Revision V. The name was changed from REDATAM+G4 to REDATAM+SP in honour of Serge Poulard, a CELADE employee and one of the principal brains behind the new REDATAM versions, who passed away returning from a mission in 2004.

In this version, the source code was completely rewritten in C++ Builder and the processing speed improved considerably. There is the same hierarchical structure, which along with REDATAM's internal storage system (inverted

variable files, in which the variables associated with each entry are stored separately), makes it possible to process data at extremely high speed (to illustrate, a computer with a 900 MHz Pentium III processor can process nearly one million entries per second to obtain a frequency, and nearly 500,000 entries per second to generate a tabulation), allowing REDATAM to prepare the information of an entire country in a single process.

Thanks to the hierarchical structure, which organizes REDATAM databases into levels, the system supplies lists of names and codes for all selectable geographic areas. This version has improved tools for rapidly defining a specific geographic area "selection;" to improve its efficiency, the software reads and processes only the data contained in the selection. It is also

possible to define geographic areas and store them in a selected file using the thematic maps generated and deployed in the REDATAM maps window.

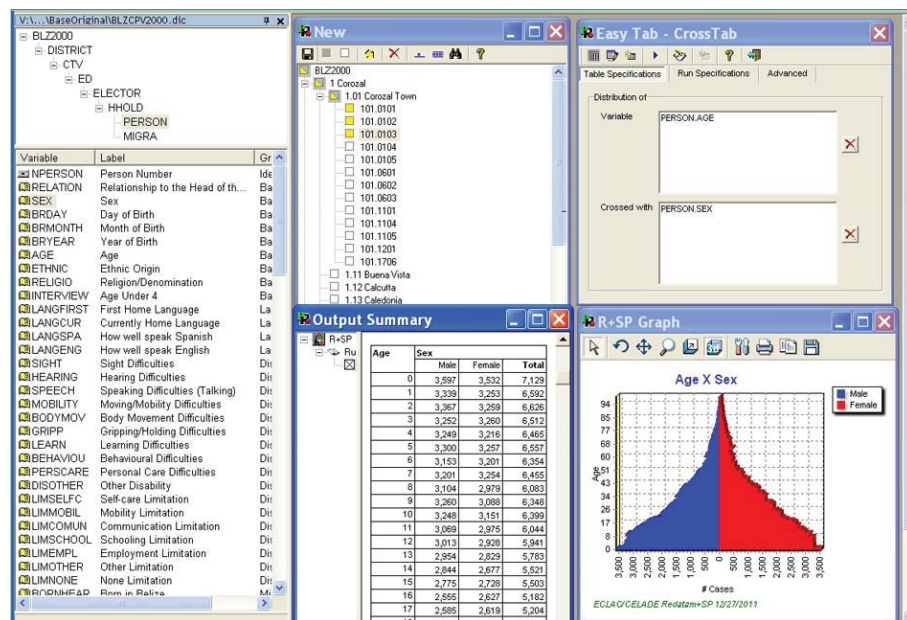
As stated above, the R+SP xPlan module offers programmers the possibility of creating applications tailored to the different users who will participate in reporting the data. R+SP xPlan application users do not need to understand the internal database characteristics; content parameters are chosen using a simple interface that can be used intuitively. In addition to presenting results in a tabular format, graphics and thematic maps can be generated automatically, with the possibility of modifying the result or exporting it into a specialized software format.

Chile's National Statistical Institute (INE) achieved a significant milestone in 1994 with the CD distribution, in REDATAM format, of municipal databases of the 1992 Population and Housing Census,³ as it did again for the 2002 census. This was emulated by other countries, including Costa Rica, Venezuela, Ecuador, the Dominican Republic, and later Peru, which, in addition to census microdata, distributed a wizard for producing tables and indicators, using the R+SP xPlan module, thereby making it possible to access and analyse census information at disaggregated levels.

With the advent of the Internet, REDATAM made use of this platform for uploading these R+Sp xPlan

³ CELADE-Population Division of ECLAC, REDATAM Report, Vol. 3, No 1/2, March 1994.

Figure 7:
Online processing of 2002 Chilean Census.



applications on the Web and making it possible to directly process online all types of census microdata using its REDATAM WebServer engine.

NSOs took advantage of this development and used it as their platform for disseminating census information. In addition to publishing the results, they also made it possible to process online census microdata from the 2000 round and previous ones as well.

REDATAM is not only used to disseminate census information at the microdata level, but also in conjunction with information systems or special projects, which make it possible to integrate information from various data sources –particularly population and dwellings censuses– at the larger administrative and regional levels. With this, it is possible to

track and compare development indicators. Examples include: the Integrated System of Indicators for Development (SID, Panama); the Integrated System of Georeferenced Information on Youth (INJUMAP, Chile); the Integrated System of Indicators on Youth (SIJES, El Salvador); the international summit tracking systems, for events such as the 1994 Cairo Conference on Population and Development and the Madrid International Plan of Action on Ageing for Latin America; and the Caribbean and the Socio-Demographic System of Indicators for Indigenous Peoples (SISPPi).

With the REDATAM+SP version, new applications were developed: (1) for calculating income distribution and poverty indicators (such as the Gini, Atkinson, Theil and Sen indexes) by using census microdata

and household surveys; (2) for indirectly estimating early infant mortality (EIM), the traditional indicators used in this case are the probability of death before one and five years; and (3) for indirectly estimating fertility (EIF), using census data and vital statistics. The estimated indicators in this case include fertility rates by age group for mothers and global fertility rates.⁴

Evaluation, projections and challenges

Today 14 Latin American and 5 Caribbean countries have Internet applications with microdata for the census rounds of 2000 and before, for processing and analysing information online using REDATAM WebServer. Of the countries that have already completed their 2010 census round, Panama and Ecuador uploaded their applications to the Web, and the hope is for the region to move towards an integrated statistics system, with this tool being used in the dissemination of 2010 census results, not only in the region but in the entire world.

CELADE continues to be a leader in demographic data production and analysis. With regard to the increased technical capacity of Latin American countries to monitor population trends and their effects on public policy, every year there are on average six training courses

⁴ The indirect procedure for estimating the fertility and mortality indicators was developed by William Brass (1974). A complete description of the original methodology and its variants, including the steps to follow, can be found in United Nations Manual X (1983).

Figure 8:
 Online processing of 2008 Cambodian Census.



for using the different functions of REDATAM software. Furthermore, technical support is constantly provided to create and launch census databases, and develop online processing applications, not only for Latin America and the Caribbean, but for other regions in the world as well. African, Middle Eastern and Asian Pacific countries are also using REDATAM for data analysis inside statistics institutes and for information dissemination. For example, Cambodia put its 1998 and 2008 Population Censuses on the Web for direct processing – by incorporating the multi-language function, it will be possible to do this in various languages—including Arabic, Bahasa, Japanese, and other software interface languages—in its new version, REDATAM 7.

Main aspects of technical cooperation

The first technical support mission conducted by CELADE outside of the region was in Viet Nam in 1992, and its first remote mission was with Cambodia in August 2000, to generate the 1998 Population Census database in REDATAM format. By 1992, IDRC had included REDATAM in its book 100 technologies - From the South for the South (IDRC, 1992).

Over these 25 years, countless REDATAM training workshops and consultations have taken place around the world. As the map below shows, the number has kept rising.

Global overview: percentage of REDATAM penetration. Countries, 2010.

Figure 9:
 Main window of the Indicator Tracking System of the 1994 Cairo Conference on Population and Development.

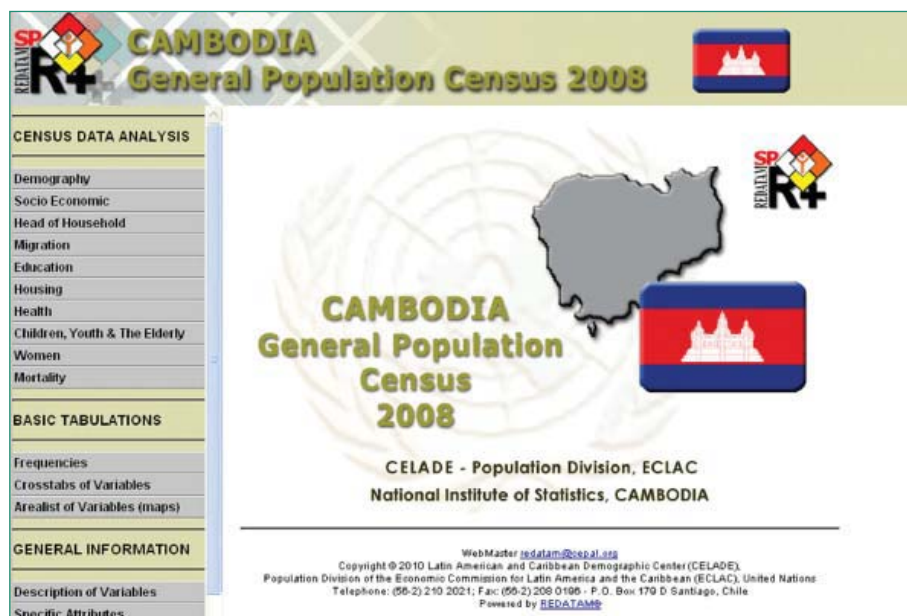


Figure 10:
Main window of the Indicator Tracking System of the 1994 Cairo Conference on Population and Development.



CELADE is not the only body promoting the use of REDATAM. UNFPA and its regional offices –UNFPA-Africa, UNFPA-Asia and the Pacific and UNFPA-LACRO– have provided REDATAM technical support and training to government organizations by promoting the “DfD” (Data for Development) concept and setting up REDATAM-based IMIS (Integrated Management Information Systems) through technical assistance, workshops and seminars. Especially in the last decade, REDATAM use has spread on its own in Latin America and the Caribbean and in other parts of the world. Clearly the explanation for this is the utility, accessibility, data safety and easy management that REDATAM provides. REDATAM has proven itself to be an ideal tool for sociodemographic data analysis and dissemination, and it has promoted cooperation among

specialists of different continents, who have built a community of users who intensify their collaboration through Facebook and Twitter.

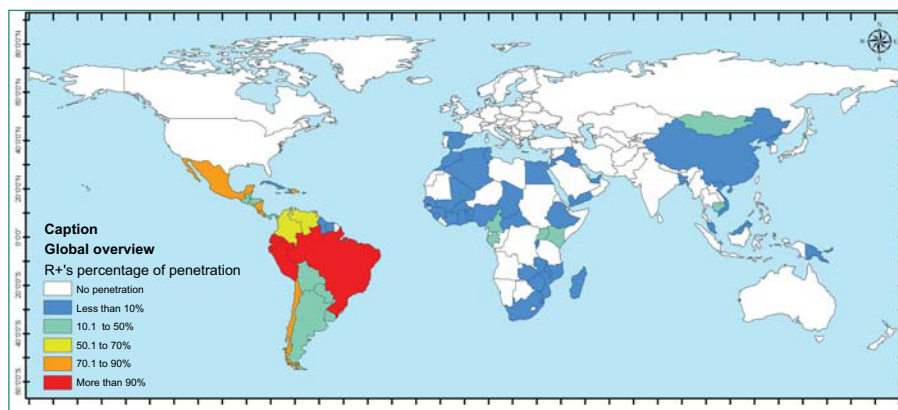
REDATAM has provided a strong response to countries’ need for census information and census data analysis. Nevertheless, the road has not been entirely smooth. Without a doubt, one of the biggest obstacles to the software’s widespread use has been the

availability of microdatabases. Countries disseminating census information by releasing databases is a recent development, especially when it comes to geographically disaggregated analysis to be used for decision-making at the local level. Certainly, the issues of urban planning, decentralization, domestic migration, housing deficits and targeting specific groups are enormously relevant to decision-making.

A challenge for this decade is to remain a leader in census data preparation, processing and utilization, as CELADE has done before, by tackling significant challenges associated with standardized methodology, updating content to reflect new issues –such as international migration and ageing– and ensuring the democratization and dissemination of information.



Figure 11
Global overview: percentage of REDATAM penetration. Countries, 2010.



REDATAM 7: speed, efficiency and ease of use

During these 25 years of the REDATAM Project, enormous jumps in computer technology have taken place (DOS operating system, Windows –16 bits, then 32 bits and currently 64 bits–), which CELADE, the Population Division of ECLAC, with the support and funding of various donors, has addressed by providing countries with a suitable and up-to-date tool. For this reason, we are presenting the new generation of REDATAM, called REDATAM 7, which incorporates new development technology -- based on C++, Delphi, Java and JavaScript -- and which, with the original program's logic and organizational structure intact, strives to improve processing speed and indicator programming, as well as make the user's interaction with the different REDATAM modules easier. Currently it is only available for the Windows platform.

REDATAM 7 improvements

A noteworthy development in these improvements has been the significant contributions of the network of users in Latin American countries and in other parts of the world, who have conveyed to us their expectations for and difficulties with using previous REDATAM generations; with that in mind, our primary improvements are:

- 1.The standardized use of the XML language: making it possible to synchronize documentation tasks, training and programming, as well as the interconnection with other, country-required computer tools, while shrinking the gap between development and user know-how.
- 2.A newly designed and programmed compiler: starting with the definition of a grammar or syntax that adjusts the language to new requirements, improving detection errors and deployment of error messages in REDATAM language.
- 3.Unicode Support: for generating dissemination applications in other languages used in the region –Quechua, Creole, Guarani– and of other Arab and Chinese regions, etc., in addition to the already existing languages: English, Spanish, Portuguese, French and Bahasa (Indonesia).
- 4.Presentation of user defined tabulations: access to each one of

On-line census

Countries that have made accessible their censuses for online processing using the Redatam Web Server (developed by CELADE-Population Division of ECLAC).



Argentina
Censo Nacional de Población, Hogares y Viviendas 2001
Instituto Nacional de Estadística y Censos (INDEC)



Belize
2000 Housing and Population Census
Tabulation System



Bolivia (Plurinational State of)
Censo de Población y Vivienda 2001
Instituto Nacional de Estadística (INE)



Chile
Censos Nacionales de Población y Vivienda 1992 y 2002
Instituto Nacional de Estadísticas (INE)



Colombia
Sistema de Consulta Información Censal 2005
Departamento Administrativo Nacional de Estadística (DANE)



Costa Rica
Censos de Población y Vivienda 1973, 1984 y 2000
Instituto Nacional de Estadística y Censos (INEC)



Ecuador
Censos de Población y de Vivienda 1990, 2001 y Galápagos 2006
Instituto Nacional de Estadística y Censos (INEC)



El Salvador, Republic of
VI Censo de Población y V de Vivienda 2007
DIGESTYC - Dirección General de Estadística y Censos



Honduras
Censos de Población y de Vivienda 2001
Instituto Nacional de Estadística



Mexico
II Censo de Población y Vivienda 2005
Instituto Nacional de Estadística (INEGI)



Nicaragua
Censo de Población y Vivienda 2005
Instituto Nacional de Información para el Desarrollo (INEC)



Panama
Censos Nacionales de Población y Vivienda 1990 y 2000
Dirección de Estadística y Censo. CELADE/CEPAL



Paraguay
Censo Nacional de Población y Viviendas 2002
Dirección General de Estadística Encuestas y Censos (DGEEC)



Peru
Sistema de Recuperación de Datos de los
Censos Nacionales 1981, 1993 y 2005
Instituto Nacional de Estadística e Informática (INEI)



Dominican Republic
Censos de Población y Vivienda 1981 y 2002
Oficina Nacional de Estadística (ONE)



Saint Lucia
2001: A Count 4 Everyone
Census Tabulation Plan (CARICOM Proposal)
The Statistical Department of Saint Lucia



Trinidad and Tobago
2000 Housing and Population Census
Census Tabulation Plan (CARICOM Proposal)
Central Statistical Office



Venezuela, (Bolivarian Republic of)
Sistema Integrado de Indicadores Sociales
para Venezuela (SISOV)

the elements involved in tabulation presentation, defining which ones to present and their position; the outputs are presented in XML for export to other applications.

5. Unlimited number of dimensions: currently the maximum number of variables that can be included in a tabulation are five plus an area break; with the newly designed REDATAM engine, tabulations may have unlimited dimensions.

Innovations included in REDATAM 7

We went back to the idea of working with a project rather than a dictionary; with the incorporation of the new RedAdmin module, we separated the shared statistical processing tasks from REDATAM databases' sporadic administrative tasks; we tried to make REDATAM programming more user-friendly; and we improved the presentation of xPlan/Web environments; enhancements include:

1. Like the WinR+ version, REDATAM 7 gets back to the Workspace concept, called Project, to facilitate the organization of all files (selections, programs, maps) and now documents (images of questionnaires, manuals, etc.) used in REDATAM database management; additionally, it is possible to preview the content of these files and documents without opening them.
2. The MULTITALLY command represents a big jump in information analysis by allowing users to obtain statistics from an uncategorized variable (for example: income, total arable area) in a single process: Cases, Sum, Maximum, Minimum, Average; conversely, one of these basic statistics can be used for several variables.
3. In addition to the existing tabulation commands (frequency, crosstabs, count, arealist), there are some new advanced, preprocessed calculations, such as GINI and NTIL.
4. One user requirement is alphanumeric variable support, for example, causes of death; with previous version, this

Look of updated news in facebook:
<http://www.facebook.com/redatam>


and Twitter:

<https://twitter.com/REDATAM>

requirement was only available for obtaining tabulations. Now, users can work with this type of variable as they do with all the other ones.

5. Just as there are easy ways for obtaining frequency, tabulation, count or area list, there is now the Process module, an assistant for generating easy indicators (EasyIndic, Fraction, QTS) that is useful for users who are unfamiliar with REDATAM language.
6. Rapid access to tabulation labels has been incorporated into the programming language, for example, “freq sex,” instead of “table tt as frequency of person. sex.” This concept also includes “freq person,” which obtains frequencies for all person labels and “freq migration,” whose results are all the variables of the “migration” group.
7. Innovations also include redesigned assistants for guiding REDATAM programming. Now a series of sequential steps are deployed with an automated assistant, which guides users to choose the best command for the tabulation/indicator they require.
8. The management and deployment of thematic maps has been completely redesigned, primarily to support other types of formats –previously only polygons were supported– and the incorporation of various sectors, labels, different forms of classification, colours, etc.

We hope users welcome REDATAM 7 and that it meets their expectations, because CELADE strives for continued progress in its development, taking steps forward to deliver a unique, easy-to-use, efficient and versatile tool to the user community.

REDATAM online (list of countries which put their census data online). 

Products of the Redatam Family

Redatam+SP (R+SP) is the latest version of the fourth generation of Redatam. Like the previous version, Redatam+SP has the following modules that allow to create, administrate, process and disseminate large census databases and/or other sources arranged hierarchically:

R+SP Process module allows you to process and analyze census or surveys data through the use of a programming language or programming wizard so you don't require any skills in programming. The user can obtain tables, graphs and maps of the variables, statistics and generated indicators. This module also includes tools for managing the database wither to expand, to concatenate, or to download subsets of the database.

R+SP Create module allows you to create hierarchical databases in a Redatam Proprietary format. It reads files in CSPro, ASCII, xBase, SPSS or IMPS format. The ASCII files should have the metadata (dictionary) defined in CsPro or CHILLAN format (DDF file).

R+SP xPlan module was developed to build data dissemination applications on CD ROM, it is intended for end users to access and explore without any knowledge of the Redatam the information, either microdata or aggregated indicators. Tables and statistics can be obtained easily.

R+SP VisualDDF extension, facilitates the creation of a data dictionary in CHILLAN format (ddf files) when the original data is in ASCII format and there is no dictionary or metadata attached to the base for the creation of a Redatam database.

R+SP WebServer is the Engine that allows the processing of Redatam databases through the Intranet or the Internet. It allows the user to send and process a request online of any Redatam database very fast and with a high degree of efficiency.

R+SP GINI module allows the calculation of inequality and poverty indicators. This module offers the calculation of several inequality indicators such as Gini, Theil, Atkinson and other indicators using income variables from a Redatam database. The module is built under the Redatam xPlan applications architecture (INL files) and includes mathematical computations for the generation of the indicators.

The Process, Create, XPlan and VisualDDF Redatam modules are installed automatically when you setup the Redatam+SP program. However, the Redatam Web Server installation must be done separately. This module requires previously installing any internet server in your computer (Apache or ISS for Windows OS).

“Potential of census information and REDATAM in the twenty-first Century”

In this special edition of the REDATAM Newsletter, commemorating the 25 years of its existence, we talked with three distinguished experts on population issues who are familiar with and have used REDATAM for many years to analyse information provided by census and other data sources in their studies and research. We spoke with Jose Miguel Guzman, chief of the Population and Development Branch of the Technical Division of UNFPA, INEC Ecuador Director Byron Villacis and Edwin St. Catherine, all of whom discussed in depth census information’s potential in the twenty-first century and the region’s use of a software like REDATAM.

Jose Miguel Guzman,
Chief of the Population and Development Branch, UNFPA Technical Division.

What is the potential use of 2010 census information?

The 2010 Census Round was done with more regional awareness about the value of data as instruments for planning and showing the true conditions of countries. Local planning has become more relevant in the region and there is an increasing and more diverse demand for information. The censuses of 2010 offer a unique opportunity for seeing demographic processes in the interior of countries, but they also will also provide a comparative picture which will help identify shared patterns. This round’s censuses also offer the possibility of analysing previously not-so-relevant issues, at least not from the perspective of census data use. The topics of gender, youth, ageing and environment have been added to urbanization, migration, demographic dynamics and others.

What are the main reasons for infrequent census information use?

Certainly, for many countries the use of census data is limited. Often this is due to data accessibility problems, poor analysis capabilities and the statistical world’s isolation from the world of real problems, which at times is seen in the disconnect between statistical producers and the ministries which should be the consumers of such information. But, at the same time it bears pointing out the significant progress in the region, greater even than in many other regions, especially when it comes to data accessibility.

How can the dissemination of census information be improved?

Four recommendations:

1. By putting dissemination and communication as integral parts of the Census Plan and not mere complementary activities.
2. By using all available media to reach the largest number of users. This implies combining old methods –such as three-page printouts, publications and other forms– with more innovative methods, such as Web-based applications, social media, Tablet (iPad) applications, interactive databases (REDATAM databases, for example) and others.
3. Think about communication content. For communication to be effective, the message must motivate the receiver, and the

way to do that is make sure that the message concretely reflects people’s real circumstances. Sometimes a message might be interesting from the producer’s point of view, but if it is not put in context and if it has no relation to people’s day-to-day lives, then it has very little media value.

4. Reach more users by broadening the traditional base. Reach schools, trade unions, the private sectors and other segments with original strategies. Essentially, the idea is to convert users into the main advocates of census taking and its importance.

What is your opinion of REDATAM?

REDATAM is unique for its type of software and it is free! I do believe there are many areas where further improvements and development are needed, but what is important is that it offers a broad platform for doing everything, or almost everything, in terms of census microdata management.

What do you believe to be its strongest attributes?

Its greatest attribute is that it has been globalized. Now there are countries all over the world that use it, but this still must be strengthened and extended.

What does it still need?


Better marketing and more partners outside of Latin America. Today people are talking about “cloud computing” and that is exactly what REDATAM has started

to do by putting data online and allowing anyone to process them from anywhere in the world.

How important is the Internet in the 2010 census?

Statistics institutes have the unique opportunity to broaden the use of census data through extensive Internet use. Many are doing this, but there are challenges. Users need more help interacting with data, they need to be more motivated, and they need tools, for example, for quickly and easily converting a database into a map or a figure that they can immediately cut and paste into a MS PowerPoint type presentation.

What support is UNFPA providing in this process?

Globally, UNFPA has contributed nearly US\$ 200 million to countries’ census processes. In Latin America, most of this funding goes toward technical assistance and national capacity building. The regional UNFPA office, based in Panama, coordinates this support. From the Technical Division, we are supporting three initiatives at the global level: (1) Manual preparation, for in-depth analysis of relevant topics (gender, youth, environment and others). (2) Support of REDATAM development and improvements, and of its expansion to other regions of the world. (3) Support of successful shared experiences on census analysis. 

What is the potential use of 2010 census information?

For the first time Ecuador has enough statistical information, especially for population, dwellings and economic censuses. We have more than 5 million pieces of data available to our users, which can be used in decision-making, as well as in the definition of public and private policy.

We must not forget that statistics are one of the main tools of decision-making; our goal has been to generate up-to-date and timely information, a goal we meet by delivering census findings in record time and by generating more than 40 surveys for the country's use.

What are the main reasons for infrequent census information use?

One of the main causes is that people are unaware of all the information that can be obtained in census findings, that and the lack of accessibility for the average user. Which is why our fundamental mission is to disseminate as much information as possible and make it accessible. On our web page, census databases have already been released, as well as user-friendly programs for using them

and a direct link with our users at inec@inec.gob.ec and with our social networks.

How can the dissemination of census information be improved?

In INEC Ecuador we are pursuing several activities.

1. Information utilization workshops
2. Using REDATAM workshops
3. Xplan, a new tool for calculating variables
4. The Si Emprende application for the enterprising, a tool for combining population and economic census information that produces, as a result, business-specific supply and demand figures.
5. Custom-made user orders at nec@inec.gob.ec, and
6. Participation in social networks.

What tools are used most for processing census information?

REDATAM, SPSS and Customer Service Unit (inec@inec.gob.ec)

What is your opinion of REDATAM?

That it is an accessible didactic tool which makes it possible to quickly obtain personalized information, according to the user's needs.

***Byron Villacis,
Director of INEC of Ecuador***

What do you believe to be its strongest attributes?


1. That it is free software.
2. Some operations are possible: averages, frequencies, percentages.
3. Results can be exported into Excel.
4. Data from different geographical levels can be hierarchically combined to create aggregated variables.

What does it still need?

There could be improvements with the number of characters for observation, since this is somewhat limited, as well as with its capacity for processing large quantities of data.

How important is the Internet in the 2010 census?

Today, its importance is fundamental; the Internet democratizes and globalizes information. In INEC Ecuador, we have developed technical tools such as Si Emprende or Xplan. We have, in addition, some 7,945 followers on Twitter and 3,889 on Facebook. We are also active on Flickr and YouTube.

The Internet is our main tool for disseminating statistics. 

**Edwin St Catherine,
Director of Statistics at the Central Statistical Office of Saint Lucia**

How can the dissemination of census information be improved?

I think that the dissemination of census data in the now and in future will be very much dependent on its ubiquitous availability. Redatam is a central part of our strategy to ensure that we make this concept of ubiquity a reality. We do expect that as redatam has in the past evolved to leverage improvement in technology so it will in the future thus ensuring the availability of census and survey data on socio-economic conditions.

Redatam has evolved from desktop processing software with a special ability to output summary tables of information faster than standard database statistical processing software.


This occurred before the advent of the internet and the availability of high capacity storage media like CDroms and DVD disk. When these high capacity standalone storage devices became available redatam XPlan was developed to take advantage of this possibility and now in the age of the internet redatam has been developed further with the redatam webserver to allow for the dissemination of census data through a user driven interface online. The St Lucia Statistical Office has kept up with these developments with redatam and this

has allowed us to benefit in a cost effective way from the evolution of these technology overtime at no software cost to us.

We are now in the process of designing our website to allow a user driven interface for access to census 2010 data and we are once again benefitting in this process from the substantial improvements that redatam is continuing to make in this area. The redatam webserver is maturing into a product which will allow non-programmers with two weeks of training to have the ability to build a web interface to allow users' access to the variables in a census or other socio-economic database to build their own tables online.

We are deploying this cost effective, customized technology to allow us to improve the dissemination of census information. Once our redatam webserver site for the dissemination of Census 2010 data is completed we will have an official press launch to ensure that all our users are aware of its availability and are capable of using it to address policy formulation, monitoring and evaluation needs at the national level and for international requirements. It will also serve as a very useful tool to all our users from students to more sophisticated users.

Redatam must as it has in the past continue to evolve and consequently we expect that given the new and exciting developments in mobile technologies which are taking place there will be further developments to put the census data in the hands of our users. This can take the form of allowing users to query census data through online text messaging or through the development of windows/android or ipad/iphone based mobile applications specifically designed to allow data from the census to be requested by the individual in any location where he/she is able to use a telephone.

While redatam can provide the tools and the platform, we must however recognize that ultimately statistical offices must ensure by promoting the awareness of census data, that, it is used in all commonly available means of communication to guide policy actions and improve the socio-economic wellbeing of our population. Our people must think "census data" whenever they face important decisions for which socio-economic datasets can provide answers and redatam can continue to be key to making this a reality. 

HOW TO OBTAIN REDATAM

The Redatam+SP (R+SP) software is available in four languages: Spanish, English, Portuguese and French. To install the software, free of charge, please download it from:

www.cepal.org/redatam

In this page you will find three files to download which corresponds to the latest versions of Redatam+SP V.5, Redatam+SP WebServer V.5 and Redatam+SP GiniPlus V.1

- [Download completo de R+SP V5Rev06 \(35 MB\)](#)
R+SP V5 Rev06 Full
- [Descargar R+SP WebServer V5Rev06 \(11.6 MB\)](#)
R+SP WebServer V5Rev06 Esp
- [Download R+SP GINIPlus V1Rev00 \(7.3 MB\)](#)
R+SP GINIPlus V1 Rev00

You must download the .exe file to a temporary directory on your hard disk to start the installation. Then extract the installation files in the temporary directory and run the .exe file (double-click on the file).

MINIMUM RECOMMENDED CONFIGURATION

Redatam+SP was developed based on 32-bit technology and runs on any the Microsoft Windows platform: Windows 9 x, NT, 2000 or above, including Windows Vista.

Regarding the minimum hardware requirements recommended these are:

Pentium 133 Mhz or higher. 32 Mb RAM memory or higher. SVGA color monitor. Hard disk with at least 30 Mb free space and more for the storage of your own databases. Mouse. Color printer (optional if you want to print maps)

R+SP Webserver requires Windows Operating System and to have installed previously any Windows Internet Server such as Apache or IIS.

HOW TO INSTALL Redatam+SP

For old users of Redatam or in the case you need to use Redatam databases from earlier versions (Redatam G4 and winR v1.2): Redatam+SP dictionary and bases can coexist in the same directory with previous versions. Dictionaries of Redatam G4 (.dic files) are fully compatible with Redatam+SP, and don't required to be imported. In any case, it is recommended to uninstall previous versions. Redatam+SP allow importing databases from any previous versions.

Follow the installation instructions on the screen. The installation process displays a series of screens that allow the user to do a custom installation.

It is recommended that the Redatam+SP Webserver installation should be carried out by the web master or administrator because it requires previously the installation of an internet server such as Apache or IIS.

EXAMPLE DATABASE (NEW MIRANDA)

An example database (New Miranda) has been included in the setup so that users can begin to learn and work immediately with Redatam+SP Process module, without having to wait for the creation or acquisition of its own Redatam database.

This base is automatically loaded when you choose the typical installation option under the "NMIR" folder, which is a subdirectory of the Redatam system folder ("C:\Program Files\Redatam" by default). We recommend using "Redatam". It is possible to omit the installation of the New Miranda database by choosing the option custom during the setup. This demo database is implemented in four languages: Spanish, English, Portuguese and French.

WHAT IS NEW in Redatam SP?

You can keep your system updated by downloading and installing the latest version of Redatam+SP. We will be developing and uploading to the web site regularly new updates (identified by the revision number, Rev xx); given that they are cumulative, if you haven't installed the previous version you only need to install the latest update. ftware es por descar



Latinamerican and Caribbean
Demographic Center - CELADE Population
Division of ECLAC

Casilla 179-D, Santiago, Chile
Tel: (56-2) 210-2726 - Fax: (56-2) 208-0196
email: redatam@cepal.org