REDATAM

USER'S
MANUAL

UNITED NATIONS
ECONOMIC COMMISSION FOR LATIN AMERICA AND THE CARIBBEAN
LATIN AMERICAN DEMOGRAPHIC CENTRE
RETRIEVAL OF CENSUS DATA FOR SMALL AREAS BY MICROCOMPUTER

(REDATAM)

VERSION 2.00 (31 March 1987)

USER’S MANUAL

A project financed by the International Development Research Center (IDRC) and with support from the United Nations Fund for Population Activities (UNFPA) and the Canadian International Development Agency (CIDA).

ECONOMIC COMMISSION FOR LATIN AMERICA AND THE CARIBBEAN (ECLAC)

LATIN AMERICAN DEMOGRAPHIC CENTRE (CELADE)
Casilla 91 - Santiago, Chile
Tel: 2283206 Cable: UNATIONS
# TABLE OF CONTENTS

## ACKNOWLEDGEMENTS

1. INTRODUCTION: Read this Section Please
   1.1. Capabilities and Requirements
   1.2. The Database
   1.3. How to Use this Manual
   1.4. How to Start REDATAM

2. OVERVIEW: What REDATAM Can Do
   2.1. The Data Dictionary
   2.2. Geographic Selection
   2.3. Statistical Processor
   2.4. Database Management Facilities

3. TUTORIAL
   3.1. Accessing REDATAM
   3.2. The Data Dictionary
   3.3. Tutorial Case 1
   3.4. Tutorial Case 2
   3.5. Tutorial Case 3
   3.6. Miranda Geography and Data

4. REFERENCE: Functions and Commands
   4.1. Statistical Processor Commands
   4.2. Geographic Selection
   4.3. Data Dictionary
   4.4. Database Management Functions
   4.5. Keyboard Usage
   4.6. REDATAM File Names

5. SPECIAL FEATURES: Advanced Users
   5.1. Statistics from Hierarchical Files
   5.2. Use of the WRITE Command
   5.3. Use of the REVIEW Command

APPENDIX

A. Equipment Required to Use REDATAM
B. System Installation and Testing
C. Utilization of a Laser Disk
D. Data Dictionary and Database Management Functions
   Data Dictionary
   Database Management Functions
E. Specifications and Limitations

INDEX
ACKNOWLEDGEMENTS

One of CELADE's priority concerns relates to the needs of the Latin American and Caribbean countries regarding numerical population data. In 1983, thanks to a grant from the International Development Research Centre (IDRC) of Canada, Mr. Arthur Conning, Chief of the CELADE Documentation and Data Processing Division, met with national authorities from statistical, planning and other agencies in Saint Lucia, Trinidad and Tobago, Costa Rica, Perú and Brazil in order to identify specific problems encountered by the countries in obtaining quantitative population data. Starting with a suggestion from Ms. Merle Alexander, then Chief Statistician of Saint Lucia, in each country it was clear that perhaps the major unmet need in this field was for rapidly available and low-cost population and housing information for small geographical areas. The trip was the basis for conceiving the REDATAM project which IDRC agreed to fund.

The technical design of the REDATAM software was carried out by Mr. Lawrence Finnegan, who was Chief of the CELADE Data Processing Unit until December 1986, when he resigned to return to his home country. Mr. Ari Silva, now Chief of the Unit, took over the responsibility for completing Version 2.00 and the operational tests. Mr. Sergio Somerville was responsible for the actual programming throughout the entire development period and worked in close collaboration with Mr. Claudio Meza. They had valuable assistance from Messrs. Sergio Acevedo, Carlos Olivares and from Mr. Jorge Balzo, who also generated the Chilean data base. The consultant, Mr. Guido Haymann, wrote the "User's Manual" in both English and Spanish, a task which he undertook with great dedication and to which he contributed many very useful suggestions.

Particular gratitude is due to the many persons who helped set up and carry out the two operational tests of REDATAM in an English-speaking Caribbean and a Latin American country, Saint Lucia and Chile, respectively. In Saint Lucia, Mr. Brian Boxhill, Chief Statistician of the Statistical Department played the principal role in fostering utilization of REDATAM. In the Instituto Nacional de Estadística de Chile, Mr. Alejandro Rojas, Chief of the División de Informática, and Mr. Eduardo Carrasco, Chief of the División de Estadísticas Demográficas y Sociales Contínuas, worked with their respective staffs to help create the database and to publicize and provide small area services.

During the visits to the Caribbean, valuable logistical and substantive support were provided by the Director of the Sub-regional Headquarters for the Caribbean of the Economic Commission for Latin America and the Caribbean (ECLAC), Mr. Clyde Applewhite, and his staff. Mr. Michael Hewitt, Chief of the ECLAC
Computer Centre in Santiago advised on equipment selection and loaned additional equipment to speed the work.

Special thanks must be expressed to IDRC, which once again made it possible for CELADE to open up a new direction of work in uncharted waters. IDRC also donated to the two statistical offices (and to CELADE) all the required microcomputer equipment for undertaking the operational tests and providing small area data services. Appreciation is also due to the United Nations Fund for Population Activities (UNFPA), CELADE's major donor, which covered the salaries of many of the persons who worked on REDATAM and to the Canadian International Development Agency (CIDA) which, through the CELADE-Canada Exchange Programme, supported specific components of the system development and database creation.

CELADE, April 1987
Part 1: INTRODUCTION

REDATAM is a user-friendly, interactive, microcomputer-based system that provides access to large geographically arranged data files, normally censuses of countries or large survey files. The principal purpose of REDATAM ("REtrieval of census DATA for small Areas by Microcomputer") is to organize and store all the data in these large files in the limited capacity of microcomputers in such a way that any tabulations or other statistics can be produced readily by the user for the smallest geographical area defined in the data, such as a city block, or for any grouping of such geographical units.

The system was developed by the United Nations Latin American Demographic Centre (CELADE) with a grant from the International Development Research Centre (IDRC) of Canada and support to CELADE from the United Nations Fund for Population Activities (UNFPA) and the Canadian International Development Agency (CIDA).

REDATAM has been designed to minimize your dependence on a manual, by including instructions and help facilities at every step of the program. However, you will achieve a faster and better understanding of the system if you familiarize yourself with its "philosophy". It is therefore recommended, that you at least read this introductory part and go through the TUTORIAL before setting out to use the system.
1.1. Capabilities and Requirements:

REDATAM can store the complete census, or other similar data for small and medium sized countries (such as Chile with a population of about 12 million inhabitants) or regions or cities within a country, and can provide interactive access to the data to produce tabulations and other statistics rapidly for specific geographical areas. While the rapidity of processing tabulations is fast in comparison with generalized statistical packages for similar microcomputers, the speed varies depending on the complexity of the tables (that is, the nature of the statistics, the number of variables used and the data transformations requested) and the total population of the area(s) of interest.

REDATAM, while efficient, is not intended normally to be employed to process a complete census of a medium-sized country, including tabulating all the data for all desired statistics, which would be done more efficiently on a larger computer and with a high-speed printer.

REDATAM has been written with the capability for basic data manipulation and production of statistics such as frequencies, cross tabulations and averages. However, REDATAM can also produce extract ("down-loaded") files of selected records and/or selected variables for the areas of interest, to permit further analysis by other statistical packages available for microcomputers, such as SPSS-PC and SL-MICRO.

REDATAM, which is written in the "C" language, operates on the IBM PC family of microcomputers. It has been tested on some 'IBM Compatible' computers, and should run on most of the fully compatible machines. System requirements are:

- 640KB of RAM Memory.
- 1 Floppy Disk.
- A Hard Disk unit of 20-100 MB (Size depending on the volume of the files to be loaded. A 20 MB disk should be able to store at least 200,000 census records or up to 1,000,000 records if data compression is used).
- A Monochrome or Color Graphics Monitor.
- A Printer with a carriage width of at least 80 characters.
- The IBM PC-DOS Operating System Version 2.0 or higher.

REDATAM also supports a "Write Once, Read Many Times" (WORM) laser-optical disk if very large storage capacity is required (presently up to 120 MB on a single removable disk).

See the APPENDIX for more details on system requirements and installation.
1.2. The Data Base:

The REDATAM system is designed to work with large data files that previously were not available on a microcomputer. Initial access to these files is gained by construction of a special database and transmission from a mainframe computer.

Since a database needs to be generated only once, most users will find the REDATAM Data Base of their national census (or other database) already created and ready to use.

The REDATAM Data Base may be viewed as a simple table or matrix, where each line corresponds to a census record and each column to a variable, or it can be a hierarchical type file, containing two record types such as for housing or household data and for each person within each house. In either case, the database is described by a Data Dictionary that contains a description of all units of geography, record types and variables. Normally, to take full advantage of REDATAM, the user must also have adequate maps of the geography of the country and the corresponding census codes so as to be able to correctly specify the areas of interest for tabulation.

The REDATAM system comes with a sample database, containing data about a fictitious country. This database is used for the TUTORIAL and for all examples and may be used for practice in the use of the system.

The construction of a REDATAM Data Base is a relatively complex undertaking, normally involving a mainframe computer and data transmission to the microcomputer. If a database is to be created by the users' institution, refer to the APPENDIX for requirements and instructions.
1.3. How to Use this Manual:

This manual is divided into 5 parts, appendices and an index.

If you need to install the system on your computer, read the Equipment, Installation and Testing Appendices (A and B). If you also need to install a Laser Disk, read Appendix C.

If you would like to obtain more detailed information about the system and its capabilities, read Part 2 (OVERVIEW), where all of the facilities available with the REDATAM system are explained. It is recommended that all new users read this part.

If you are a first-time user or are not very familiar with microcomputers, proceed to the TUTORIAL (Part 3), where you will be taken step by step through most of REDATAM's facilities by working out specific examples on your computer with a small database that comes with the REDATAM system.

Once you start using the system, the REFERENCE (Part 4) will be useful in explaining each of REDATAM's options and commands. It is organized in such a way that you can go directly to the specific feature you want to learn about.

When you are ready for more complex work, Part 5, SPECIAL FEATURES, will show you how to use the more advanced options of REDATAM.

APPENDIX E summarizes general specifications and limitations of the REDATAM system.

Finally, if you need to generate or modify your own database, refer to APPENDIX D.
1.4. How to Start REDATAM:

If REDATAM is already installed (if not, see APPENDIX B), you are ready to start using it. First, turn on your computer. You will be under the control of the Operating System (DOS). Either the A> or C> prompt should be the screen. If necessary, select the disk unit where your working directory resides by, for instance, entering:

C: [ENTER]

where [ENTER] represents the ENTER key (usually labeled with a bent arrow, or with Return in some computers). C may be replaced by the disk unit to use. Then, select your REDATAM working directory by, for instance, typing:

CD \WORK [ENTER]

where "WORK" is to be replaced by the actual name of your directory. Load REDATAM into memory by entering the name of the database you wish to use such as in:

MIRANDA [ENTER]

After a few seconds you should see REDATAM's Main Menu screen. You are ready to go.

REDATAM supports multiple languages. If you wish to change it, see the installation procedure in APPENDIX B.

If you are a new user, please follow the TUTORIAL given in Part 3 of this manual.

---

SPSS-PC is a registered trademark of SPSS Inc., 444 N. Michigan Ave., Chicago, Ill 60611.
SL-MICRO is a registered trademark of QSC, Box 23056, Lansing, MI 48909.
IBM is a registered trademark of International Business Machines Corporation, Armonk, NY.
Part 2: OVERVIEW: What REDATAM Can Do

This part provides a general introduction to the different aspects of REDATAM, its capabilities and usage.

REDATAM consists of 4 major modules:

- **The Data Dictionary**: The Data Dictionary contains the definition of all variables available in the REDATAM database.

- **Geographic Selection**: This feature is used to define the geographical areas that will be used for producing the statistics.

- **Statistical Processor**: This module has a series of facilities for data manipulation, record selection, the generation of statistics and the printing of the results. It also provides interfaces between REDATAM and other statistical packages if a more elaborate analysis is required.

- **Data Management Facilities**: These facilities assist in the process of generating and loading new REDATAM databases.

Each of these modules will be reviewed in the sections that follow. Details on each module are given in the REFERENCE (Part 4) and the APPENDIX.
2.1. The Data Dictionary:

The Data Dictionary provides a bridge between the physical data files and the user. It lets the user view the data in terms of variables and census records, thereby isolating him or her from the physical details of how the data is stored and manipulated.

A Data Dictionary contains a list of all available variables, and the values they can take. A description of each value is also provided. See section 3.6 for a sample Data Dictionary.

The Data Dictionary for a given data-set of interest is normally already defined, so most users will use it only to obtain a listing of the variables and their range of values in the database. However, options are available to modify an existing dictionary to suit specific reporting requirements or to create a new one.

Multiple databases, and therefore Data Dictionaries, are supported. The user may select which one to use.
2.2. Geographic Selection:

The Geographic Selection module is used to specify the geographic areas of the country, region, or city for which statistics are required.

The selection procedure uses the organization of the data files to provide access to small areas, such as city blocks, defined in the data. In order to do this, the data files that are loaded into the REDATAM system are rigorously checked for sequence and order within the geographic codes assigned to each area. The geography is assumed to have a hierarchical organization, that is, geographical areas have subdivisions, which in turn may also be subdivided. For example, provinces within regions, districts within provinces, and so on. The number and nature of geographical levels will vary with each country.

To access the geographic information of the data files, REDATAM uses an interactive full-screen-oriented approach to display the names and codes of the different levels of the geography, and allows the user to select the areas of interest.

Within each level of geography selected, the user is prompted to select component areas, continuing with this geographic selection process, as necessary, to the lowest level of geography available in the data file such as block or enumeration district. At any level, the user may specify that all of the unit is to be selected, eliminating further detailed selection at the subsequent lower levels.

REDATAM also allows the user to define groupings of the different geographical areas in ways other than permitted by the hierarchical organization of the country. Thus, for example, districts belonging to different provinces can be grouped together.

The geographical areas selected, are stored for subsequent use by the system. This avoids having to re-specify the geographic selection for multiple analyses of the same area(s) and allows REDATAM to use a more efficient system to access the data for the particular unit of geography. Existing geographical selections may also be modified.
2.3. **Statistical Processor:**

The Statistical Processor produces statistics for the geographical area of interest. This area must have been previously defined by using the Geographic Selection capabilities.

The data for a specific area may be further manipulated by criteria based on the variables (e.g., generate a tabulation using only females over 14 years old), by re-coding variables (e.g., make age groups from single years of age), by combining variables to derive new variables (e.g., create an index of housing quality), etc.

Basic statistical measures are provided, including frequencies, cross-tabulations and averages. The results of the tabulations can be viewed directly on the screen, they can be printed, or stored on a disk file for further analysis by statistical or spreadsheet programs, for preparation of graphics or for editing prior to printing.

The statistical command sequences may be saved on a file for later use with different geographical areas, they can be modified or incorporated in whole or in part into new command sequences.

As part of the Statistical Processor, REDATAM provides facilities to generate subsets of the data file ("down-loading"). Records are extracted from the REDATAM files and written to disk. Several formats are available, including some popular statistical packages for IBM microcomputers such as SPSS-FC and SL-MICRO, in addition to DOS ASCII output and REDATAM's own internal format.

The down-loading of a portion of a REDATAM database in its internal format, may be used by a city or region, for example, to obtain its data directly from the national REDATAM census database for use on its own microcomputer with REDATAM, normally without the need to utilize a large hard disk or laser disk for storing the entire census database.
2.4. Database Management Facilities:

These facilities allow the user to define and load a new REDATAM database. Most users will not need to concern themselves with this aspect, since the census files they use will have already been loaded by programmers.

A new REDATAM database has to be generated on a mainframe computer, compressed if very large, and transmitted to the microcomputer, where further processing is required. The Database Management Facilities take the transmitted files and convert them to a format usable by REDATAM.

Technical assistance for the creation of a REDATAM Database, or in certain cases, its actual creation, may be requested from CELADE.
Part 3: TUTORIAL

The purpose of this part is to provide an introduction to the capabilities and usage of the REDATAM system, without having to study the details of each option and command.

This is done in the form of a tutorial. The tutorial is based on a fictitious country called MIRANDA, whose data is supplied with the REDATAM System. Three different examples will be worked out, and you should follow them with your computer. Each of them is increasingly complex, but since the cases are independent of each other, you do not have to do them in order.

Imagine that you work for a regional planning agency of Miranda, concerned with the improvement of education in the rural areas of the country. Given limited resources, you want to determine what type and location of school will be most beneficial for the community. This means that you will need to obtain information on the distribution of population in rural areas, the present educational level, age, etc., and compare this with the country's requirements, priorities and resources in order to decide your course of action.

You have at your disposal a population and housing census of Miranda, which has been loaded into the REDATAM database. REDATAM can help you in organizing and presenting this information in many different ways to assist you in your work.

Before you start to work with REDATAM, you want to make sure that you have a clear idea of what you want to do, as detailed in the following checklist:

a) You are familiar with the geography of the country, and the geographical divisions of the REDATAM census file.

b) You know what information is available from the census, that is, what variables and their categories or range of values.

c) You have decided what geographic subdivisions make up the area of interest, and how you are going to group these subdivisions.

d) You know what tabulations or other statistical information you want to obtain for these areas, and how to obtain them using the REDATAM statistical processor.

In order to satisfy point a), a map of Miranda is included (Section 3.6) indicating the geographical subdivisions. Miranda is divided into two Provinces: FLORINDA and PERDITA. Each of them has one Town, the rest of the area being rural. Florinda's town is called VILCHES, while the town in Perdita is TUMACO. The towns
and rural areas of each province are divided into Enumeration Districts, or EDs, as shown in the map. Take some time to familiarize yourself with Miranda. A hierarchical diagram of Miranda's subdivisions is also included in section 3.6. You may find it useful to refer back to it for the examples.

For point b), we need to know what variables make up the census file. This can be found in the Data Dictionary. The Data Dictionary contains a list of all available variables and the values they can take. You may obtain a printout of the Data Dictionary directly from your microcomputer as described in 3.2.

Points c) and d) will be discussed prior to each tutorial case.

While each case is independent of the others, we suggest that you first go through tutorial case 1, where you will find more detailed help on the use of the screens and keyboard.
3.1 Accessing REDATAM:

This example assumes that you have a basic familiarity with the use of microcomputers, that is, you know how to turn it on, and are familiar with the keyboard. If this is not so, we suggest you take a look at KEYBOARD USAGE, Section 4.5.

We assume that at this point the REDATAM program and tutorial data files have been installed in your system. If this is not the case, refer to APPENDIX B or ask your REDATAM coordinator. After turning on your computer you will be under the control of the Operating System, also known as DOS. Either the A> or C> prompt should be on the screen. If the A> prompt is on the screen, type:

\[ C: \mediumbreak \quad [\text{ENTER}] \]

where [ENTER] represents the ENTER key (usually labeled with a bent arrow, or with Return in some computers), that is, do not actually type the word "ENTER". You should now see the C> prompt on your screen. Then, select your REDATAM working directory by typing:

\[ \text{CD \ mediumbreak } \text{directory} \quad [\text{ENTER}] \]

Where "directory" is to be replaced by the name of your directory. If you do not have one, see APPENDIX B or ask your REDATAM coordinator. Load REDATAM by entering:

\[ \text{MIRANDA} \quad [\text{ENTER}] \]

With this, REDATAM starts loading, and after a few seconds you should see the following Menu screen:

---

REDATAM System 2.0  Data Base: m80  Date: 27/06/1986  Time: 14:57:41

Statistical Processor
Geographic Selection
Data Dictionary
database Management
Exit

^V Move cursor. Select with ENTER or first letter. ESC to exit

---

Figure 3-1
This is the REDATAM Main Menu. It shows all the available system functions:

- Perform Statistical Processes: To define and print statistics on the selected geographic areas.

- Do a Geographic Selection: To define and group the Geographic Areas that are to be analyzed.

- Access and maintain a Data Dictionary: The Data Dictionary contains information about the variables available in the data files.

- Perform Database Management Functions: This option allows you to define and install new databases.

Some items on the screen will appear underlined and others highlighted. Adjust the contrast and brightness controls on your screen so that you see them properly.

At the top of the screen, REDATAM displays the currently active database. If it does not correspond to MI80 (as the Miranda database is named), as shown here, it will have to be changed in order to proceed with this tutorial. Refer to APPENDIX B or ask your REDATAM coordinator for assistance.

Here, and in all other menus, you can choose an option either by moving the cursor with the Up or Down Arrows until the desired option is highlighted (or blinking or underlined, depending on the type of monitor you have), or by directly typing the underlined or highlighted letter (usually the first one) of the description of that option. Try pressing the up and down arrows in order to see how the different options are highlighted.

You can exit this menu (and leave REDATAM) by selecting the Exit option or pressing the ESC (or ESCape) key. In general, when working within REDATAM, pressing ESC will return you to the previous menu, or cancel the current operation.

If you hear a beep, then you have pressed an invalid key for this particular menu. Always watch the screen with attention. The bottom line generally contains instructions.

Help or assistance related to the context or menu you are currently at, is often available, as shown on the bottom line. If so, just press ? (that is, hit the key with the "?" symbol while holding down the SHIFT key) if you need help or further information.
3.2. The Data Dictionary:

You may want to go through this section if you need to print a copy of the Miranda Data Dictionary.

Choose the Data Dictionary option from the Main Menu (see section 3.1. for how to access and use the Main Menu), that is, move the cursor until it is highlighted and then press ENTER, or press the letter D. The following screen will appear:

| REDATAM System | Data Base: mi80 | Date: 27/06/1986 | Time: 14:59:18 |

```
List Active Dictionary
Print Active Dictionary
Management Function
Exit to Main Menu
```

^v Move cursor. Select with ENTER or first letter. ESC previous level

Figure 3-2

This is the Data Dictionary Menu. From here we control all Data Dictionary functions. The only item you will be concerned with for this tutorial is the second option.

If you need to print a copy of the Data Dictionary for Miranda, first make sure that it is the active Data Dictionary (MI80). If not, see APPENDIX B or ask your REDATAM coordinator for assistance. Then, make sure that your printer is turned on and on-line. Finally, select the Print Active Dictionary option from the Data Dictionary Menu. The printer should start printing shortly. If it does not, check whether it is on-line and connected to the computer.

When done with the Data Dictionary, press ESCape in order to return to REDATAM's Main Menu.
3.3. Tutorial Case 1:

The purpose of this example is to illustrate the basic mechanics of using REDATAM. This is therefore a rather simple example. At the end of this case you will have used:

- The Geographic Selection option in order to select a geographic area using total and partial selections.
- The following Statistical Processor commands:
  BROWSE
  CLEAR
  CROSSTABS
  EXIT
  FREQUENCIES
  GEOGRAPHY
  LIST
  PROCESS
  ? (Help)

Since the scenario calls for a study of rural areas, let us choose all rural areas of Miranda for this example, in order to get an idea of the numbers we are dealing with. This takes care of checklist point c) from page 3-1.

For this geographic area, we want to obtain three frequency tables, each listing respectively the frequencies of occurrence of age, number of people attending school and highest educational level, and a cross-tabulation of age and highest educational level.

Look up the Data Dictionary listing: The first page will show some general information about it, and subsequent pages will list each of the available variables. For each variable there is a short name (as known by REDATAM), a description, and the range of values it can take. Then, if applicable, follows a list of these values, together with a description for each of them.

The names of the variables we require are AGE, ATTSCH and HIGHSCHT. With this, item d) of the checklist of page 3-1 is complied with.

Proceed then to select the geography. You have to start from the Main Menu, as shown in Fig. 3-1. If this is not what your screen is showing, you have to get there by following the instructions on your screen. From the Main Menu select the Geographic Selection option. You will be presented with the following screen:
This screen is used to load or create a Geographic Selection File. This file is used by REDATAM to store the selected areas, for which tabulations or other statistical operations can then be performed.

The number of files shown will vary depending on how many have been created on your directory. As the instructions say, you may select a file by moving the up and down arrows until the desired selection is highlighted. In this case we want to create a new file, so select the CREATE FILE option and press ENTER. Next to that option the following will appear:

CREATE FILE Name: __________

The cursor will be positioned at the beginning of the dashes that follow "Name:". At that point, REDATAM is expecting you to enter the name of the Geographic Selection File. The name can be any combination of up to eight alphanumeric characters, starting with a letter. This file will be later used by the Statistical Processor. For this example type:

CREATE FILE Name: casel____

"casel" is the name by which REDATAM will know our selection. If that name already exists since someone else has tried this tutorial, add your initials to the name. If you make a mistake you can reposition the cursor using the left and right arrows and then retype the name. Use the DEL key to erase the character at the cursor position. Press ENTER after completing the name. At the bottom of the screen you will see:

Label: ____________________________________________

REDATAM is requesting a short descriptive comment about the file that will help in later identifying it. Up to 60 characters of text, numbers or symbols may be entered. For this case type something like:
You can correct mistakes as described. Hit ENTER when you are finished typing. The file has been created and you will be presented with the following screen:

```
REDATAM V2.0 geography Selection
PROVINCE TOWN ED
Data Base: M180 Date: 25/3/1987
File: case1.sel

<table>
<thead>
<tr>
<th>Code</th>
<th>Area Name</th>
<th>Sel</th>
<th>Recode</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Florinda</td>
<td></td>
<td></td>
</tr>
<tr>
<td>02</td>
<td>Perdita</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

P Partial Selection  T Total Selection  ESC Abandon  ? Help

Figure 3-6

This screen displays the highest geographic level within the country, that is, Province, and shows all available selections, namely Florinda and Perdita. Also shown on the top of the screen is the name of the database in use and the Selection File. The names of all the Geographical Levels are also displayed, the current one being underlined.

The cursor will be positioned under the "Sel" column and can be moved by using the up and down arrow keys. The Sel column is used to select a given area. You can do this in three ways: no entry, Partial (P) or Total (T) selection. If we specify T, all subdivisions under the one selected will be included in our example. If a more detailed selection is desired, P is entered.

As stated, in this example we wish to select all of Miranda's rural areas. Since each province has both urban and rural areas, we cannot select any one in its totality, but need to move down to the town level of both Florinda and Perdita.

In consequence, we need to do a partial selection on both provinces. This is accomplished by entering a P on the "Sel" column in front of both the Florinda and Perdita provinces. Use the up and down arrows to move from one line to the next. When complete, the portion of the screen with the geography should look like:

```
<table>
<thead>
<tr>
<th>Code</th>
<th>Area Name</th>
<th>Sel</th>
<th>Recode</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Florinda</td>
<td>p</td>
<td></td>
</tr>
<tr>
<td>02</td>
<td>Perdita</td>
<td>p</td>
<td></td>
</tr>
</tbody>
</table>
```

If at this point you get somewhat confused about the different areas of the country, it may help to take a second look at Miranda's map and at the geographical hierarchy diagram provided in
If you are creating a new area, as in this first example, pressing:

HOME

automatically takes you to the next level of geography, Town.

Should you forget what to do at any time, looking at the bottom of the screen, you will see that help is available. It can be called up by pressing ?, the Help key (the SHIFT key pressed together with the key with the symbol "?"). A window pops up over your screen. It lists the options available together with instructions on how to use the help screen. Read them carefully.

You will find that pressing HOME takes you to the next incomplete level or that you can proceed to the next lower Geographic Level by pressing SHIFT-down arrow (hit the down arrow while holding down the Shift key). You must exit Help before you can proceed. Do this by hitting the ESC key. The window will disappear.

If you make a mistake at any point during the geographic selection, position yourself at the line requiring correction and reenter the correct letter (hit the space bar to eliminate that selection). You may cancel the whole selection and start from scratch by pressing ESC (you will be asked to confirm your request).

After pressing HOME or SHIFT-down arrow the Geographical Selection screen will change to:

<table>
<thead>
<tr>
<th>REDATAM System</th>
<th>Data Base: MIB0</th>
<th>Date: 16/12/1986</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROVINCE TOWN ED</td>
<td>File: case1.sel</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Code</th>
<th>Area Name</th>
<th>Sel</th>
<th>Recode</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Florinda</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01</td>
<td>Vilches</td>
<td></td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>Rural</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3-7

You will see the two subdivisions of Florinda: Vilches and the rural area surrounding it. Note that Florinda, the hierarchically higher geographic division is shown, with the towns indented. The screen heading also shows your present geographic level. The cursor will be under the Sel column as before and you can move between the two towns with the up and down arrows.

Since our geographical area of interest includes only rural areas, choose all of the Rural area (by entering a T under the Sel column) and leave blank the town of Vilches (leaving it
thereby out of our selection). After entering T, the cursor will move to the Recode column. Hit ENTER in order to return to the Sel column, since we are not using this option in this example. The screen should now appear with:

<table>
<thead>
<tr>
<th>Code</th>
<th>Area Name</th>
<th>Sel</th>
<th>Recode</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Florinda</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01</td>
<td>Wilches</td>
<td></td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>Rural</td>
<td>t</td>
<td></td>
</tr>
</tbody>
</table>

Next, we have to select the rural areas of the province of Perdita, that is, we want to move "horizontally" to the next province at the same level, that of town. See the geographic hierarchy at the end of this chapter, in order to clarify this concept.

Since we are creating a new area, pressing HOME takes us automatically to the remaining incomplete portions of the geography, which in this case is the town level of Perdita.

By recalling the help menu (with ?) you will find out that the left and right arrows also let you move horizontally across the same levels.

With either HOME or Right arrow, the next screen that appears should be:

<table>
<thead>
<tr>
<th>Code</th>
<th>Area Name</th>
<th>Sel</th>
<th>Recode</th>
</tr>
</thead>
<tbody>
<tr>
<td>02</td>
<td>Perdita</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01</td>
<td>Tumaco</td>
<td></td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>Rural</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3-8**

Proceed exactly as in the previous case, and select the totality of the Rural area of the province with a T. When done, hit the HOME key to ensure that there is no incomplete geographic selections remain. When you press HOME, REDATAM checks whether there are any pending nodes, that is, if some partial selections are not closed by Total selections at a lower level. If that is the case, you are taken to the pending level in order to complete the selection, otherwise, REDATAM moves to the highest geographical level.

Now that you are finished with your selection, hit

END
in order to save your selection in a "Selection File" and exit the Geographical Selection to be returned to the Main Menu (fig. 3-1). The selection information is saved in the file named CASE1.SEL, which is the name you gave it with the suffix .sel added by REDATAM.

Note that you may print your Geographic Selection through an option of the file selection menu (fig. 3-5).

You are now ready to define the statistics you want to do on the selected data. Choose then the Statistical Processor option from the Main Menu.

The Statistical Processor is a command driven system (as opposed to the menus we have seen up to now). As you enter it, you will be presented with a blank screen with the following heading:

```
REDATAM V2.0 Statistical Processor
Data Base: M180 Date: 18/3/1987
RECORD:
```

**Figure 3-9**

The Statistical Processor is now waiting for you to enter a command to tell it what to do. You will see the cursor blinking at the bottom of the screen; this is where commands are entered. Commands are entered by typing them and then hitting the ENTER key.

Now, tell the processor which geographic area to select, that is, the name of the Geographic Selection File, by typing the following command (of course, if you used a different name, use that one):

```
GEOGRAPHY case1 [ENTER]
```

where [ENTER] represents the ENTER key. Commands may be typed in upper or lower case. For the sake of clarity, commands are shown here in upper case and the arguments in lower case, although you can enter them in upper or lower case.

The command will be moved to the upper part of the screen after being checked for syntax. If you typed it correctly, no error message will appear under the command, indicating that REDATAM has found the command to be valid and correct, and that it is ready to accept another.

If a mistake is found, the screen will show an error message such as "no such command". The command just entered will also be re-displayed at the bottom of the screen. The right and left arrows, the HOME, and the END keys can be used to position the cursor at the point needing correction without reentering the whole sentence. The DEL key can be used to delete the character at the
cursor position. ESC erases the entire command and let you start anew. After correcting the command, press ENTER to reissue it.

Although the fully-written commands are shown in these examples, you need only enter as many letters as will uniquely distinguish it from all the other commands. For instance, REDATAM will recognize G, GE, GEO, etc., as the command GEOGRAPHY.

When the available space on the screen fills up, it will "scroll" moving up each line to make space at the bottom.

The set of statistical commands and their syntax is listed in the REFERENCE. The GEOGRAPHY command we just used, tells the Statistical Processor which Geographic Selection File to use. By specifying the CASE1 file, you have indicated the geographic area just defined.

At the beginning of this example, we decided to order frequency tables for age, school attendance and educational level. The REDATAM command to accomplish this is:

```
FREQUENCIES AGE ATTSCH HIGHSCHT
```

Remember to end this and all other commands with the ENTER key.

Recall that AGE, ATTSCH and HIGHSCHT are, according to the Data Dictionary, the names that REDATAM uses for the desired variables. No processing (other than syntax checking) takes place at this point. If you get an error message, correct the command and reissue it as explained above.

The FREQUENCIES command could have been entered in its abridged form, such as F. The variables however, cannot be abridged.

The next information we wanted, was a cross-tabulation between age and highest educational level. The command to use is:

```
CROSSTABS AGE BY HIGHSCHT
```

At processing time, this will result in a joint frequency distribution table of cases according to age and educational level. The values of AGE will make up the lines or rows, and those of HIGHSCHT the columns.

The Statistical Processor has a Help facility available, which can be accessed by pressing the "?" key while holding down the SHIFT key. When this is done, the screen is overlayed with a list of available commands. You can request a brief description of each command by following the instructions at the bottom of the screen. Those same instructions tell you that the ESCape key brings you back to the Statistical Processor to the point you left it.
If you need information about the variables available within the database, the BROWSE command lets you display on the screen the same information that you have in your Data Dictionary listing. The BROWSE command is menu-driven and instructions are displayed at the bottom of the screen. Again, the ESCape key returns you to the Statistical Processor.

Having now entered all the desired commands, we can display the information entered so far into the Statistical Processor. First clear the screen and then list the commands by entering:

```
CLEAR

LIST
```

Whereas the following should display:

```
LIST
GEOGRAPHY CASE1
FREQUENCIES AGE ATTSCH HIGHSCHT
CROSSTABS AGE BY HIGHSCHT
```

You are now ready to start the actual process and print the results. Make sure at this point that the printer is on line. Then enter the following command:

```
PROCESS OUTPUT = PRINTER
```

REDATAM should answer with

```
Tables require 4964 bytes
Writing output to PRINTER
Process started at 09:38:47
```

and start processing your request. OUTPUT = PRINTER is an option that tells REDATAM to direct the output to the printer. If not specified, the output is displayed on the screen only. You may want to do that if you do not have a printer available at the moment. Instructions on the screen will tell you how to proceed.

It may take a minute or so before REDATAM starts printing. On the upper right hand side of the screen, you should see the message:

```
RECORD  nnn
```

This is to tell you that REDATAM is working. The 'nnn' characters, which should change fairly rapidly, show the number of census records that have been processed in increments of 10.

You should get 3 frequency tables and a cross-tabulation, whose first few lines should look like this:
### Frequency for Age
Age of Respondent

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>CODE</th>
<th>ABSOLUTE</th>
<th>RELATIVE</th>
<th>ACCUMULATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>FREQU</td>
<td>FREQU</td>
<td>FREQU</td>
</tr>
<tr>
<td>0</td>
<td>98</td>
<td>2.4</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>103</td>
<td>2.5</td>
<td>4.9</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>126</td>
<td>3.0</td>
<td>7.9</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>142</td>
<td>3.4</td>
<td>11.3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>126</td>
<td>3.0</td>
<td>14.4</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>129</td>
<td>3.1</td>
<td>17.5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>121</td>
<td>2.9</td>
<td>20.4</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>143</td>
<td>3.5</td>
<td>23.9</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>123</td>
<td>3.0</td>
<td>26.8</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>138</td>
<td>3.3</td>
<td>30.2</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>121</td>
<td>2.9</td>
<td>33.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>etc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3-10**

### Frequency for ATTSCH
Attendance at School

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>CODE</th>
<th>ABSOLUTE</th>
<th>RELATIVE</th>
<th>ACCUMULATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fulltime Student</td>
<td>1</td>
<td>1252</td>
<td>30.3</td>
<td>30.3</td>
</tr>
<tr>
<td>Parttime Student</td>
<td>2</td>
<td>7</td>
<td>0.2</td>
<td>30.5</td>
</tr>
<tr>
<td>Not Applicable</td>
<td>3</td>
<td>2861</td>
<td>69.3</td>
<td>99.9</td>
</tr>
<tr>
<td>Not Stated</td>
<td>4</td>
<td>6</td>
<td>0.1</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td>4126</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

0 VALUES OUT OF RANGE

**Figure 3-11**
FREQUENCY FOR HIGHSCHT Highest Educational Level, Type School

GEOGRAPHY IS GLOBAL

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>CODE</th>
<th>FREQ</th>
<th>REL FREQ</th>
<th>ACC FREQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>1</td>
<td>1135</td>
<td>27.5</td>
<td>27.5</td>
</tr>
<tr>
<td>Nursery·Infant</td>
<td>2</td>
<td>470</td>
<td>11.4</td>
<td>38.9</td>
</tr>
<tr>
<td>Primary</td>
<td>3</td>
<td>2390</td>
<td>57.9</td>
<td>96.8</td>
</tr>
<tr>
<td>Second or Comp</td>
<td>4</td>
<td>94</td>
<td>2.3</td>
<td>99.1</td>
</tr>
<tr>
<td>Other Second</td>
<td>6</td>
<td>11</td>
<td>0.3</td>
<td>99.4</td>
</tr>
<tr>
<td>University</td>
<td>7</td>
<td>5</td>
<td>0.1</td>
<td>99.5</td>
</tr>
<tr>
<td>Other</td>
<td>8</td>
<td>15</td>
<td>0.4</td>
<td>99.9</td>
</tr>
<tr>
<td>Not Stated</td>
<td>9</td>
<td>6</td>
<td>0.1</td>
<td>100.0</td>
</tr>
</tbody>
</table>

TOTAL 4126 100.0

0 VALUES OUT OF RANGE

Figure 3-12


CROSSTAB OF AGE Age of Respondent
BY HIGHSCHT Highest Educational Level, Type School

GEOGRAPHY IS GLOBAL

<table>
<thead>
<tr>
<th>AGE</th>
<th>CODE</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>98</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>98</td>
</tr>
<tr>
<td>1</td>
<td>102</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>103</td>
</tr>
<tr>
<td>2</td>
<td>121</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>126</td>
</tr>
<tr>
<td>3</td>
<td>125</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>142</td>
</tr>
<tr>
<td>4</td>
<td>93</td>
<td>32</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>126</td>
</tr>
<tr>
<td>5</td>
<td>65</td>
<td>56</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>129</td>
</tr>
<tr>
<td>6</td>
<td>21</td>
<td>73</td>
<td>27</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>121</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>92</td>
<td>44</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>143</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>58</td>
<td>61</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>122</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>39</td>
<td>95</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>138</td>
</tr>
</tbody>
</table>

etc.

Figure 3-13

REDATAM first prints the commands you entered, and then the tables. The frequency tables (Figs. 3-10, 3-11 and 3-12) begin with a heading indicating the variable being counted. The "GEOGRAPHY IS GLOBAL" message indicates that the numbers that follow correspond to those for the total selected area (as specified
during the geographic selection that was stored in the CASE1 File). The following column headings are printed:

CATEGORY: Labels (as defined in the Data Dictionary or by the user) for the values of the variable in the CODE column.

CODE: Is the value that the variable takes.

ABSOLUTE FREQ: Is the total number of records whose value corresponds to that under the CODE column.

RELATIVE FREQ: Is the percentage from the total number of records that the ABSOLUTE FREQUENCY represents.

ACCUMULATIVE FREQ: Is a cumulative total of the RELATIVE FREQUENCY.

Totals are printed for the Absolute and Relative Frequency, and a final line appears indicating the number of values found to be out of the range specified in the Data Dictionary. This is useful in order to assess the validity of your data.

The cross-tabulation table (Fig. 3-13) starts with a heading indicating the two variables being crossed and also with a message indicating the geography included.

Then, a matrix is printed where AGE makes up the rows while the columns are made up by the values of HIGHSCHT, Highest Educational Level. It is possible that your printer was not wide enough to print all the columns. In this case, the part of the line that does not fit gets printed on the next line, and your matrix may look a little strange. You can fix that by using "compressed type" on your printer. See your PC specialist if you do not know how to do it.

Note that the column headings correspond to the descriptions of the values of HIGHSCHT truncated to the width (8 letters) of the column. Each cell contains the number of records (within the area selected) that have the values of AGE and HIGHSCHT of the corresponding row and column. Horizontal and vertical totals are printed as well as the number of records out of range.

Finish this session by typing:

EXIT or EX [ENTER]

You will be returned to REDATAM's Main Menu (Fig. 3-1). You can now proceed with Tutorial Case 2 or exit REDATAM by hitting the ESC key. This takes you back to the operating system, DOS.
3.4. Tutorial Case 2:

This example illustrates a more detailed geographical selection than that of Tutorial Case 1, and also shows the use of a few more statistical commands. At the end of this case you will have used:

- The Geographic Selection partial and total selection options.

- The following Statistical Processor commands:
  
  AREABREAK
  AVERAGE
  CLEAR
  CROSSTABS
  EXIT
  GEOGRAPHY
  IF
  LIST
  NEW
  PROCESS
  RECODE
  UNIVERSE
  VALUE LABEL
  VAR LABEL

In continuing with our study of the educational needs of rural Miranda, we will look in detail into a few selected districts. For those districts, we would like to study the following:

First, we would like to determine the number of people and their average age, as related to the highest type of school attended (not necessarily completed). We would like these results both for people currently in school and those who are not. We are only interested in people of ages between 6 and 30 years.

Next, we would like a table showing the number of children in primary school by age, along with the number of children who have dropped out or never attended primary school. We shall limit this study to children between the ages of 6 and 15.

The area of interest will be Enumeration Districts (EDs) 95, 96, 101 and 102 (see Miranda's map), which will have to be extracted from the total geography. This takes care of checklist point c) (page 3-1).

Take the Data Dictionary listing and identify the variables that will be used:

  AGE needs to be used.

  ATTSCH will be used to determine whether people are studying
or not.

TYPESCH will be used to determine the type of school currently or ever attended.

Normally, you would decide at this point which statistical commands to use in order to complete point d) of the checklist of page 3-1. However, since this is a tutorial, we will analyze the commands as we enter them.

As mentioned above, this example does not elaborate as much on each REDATAM feature as Tutorial Case 1 does. References are made to that example for figures and procedures. We suggest that if you have not already done it, you go through that example first in order to feel more comfortable with the REDATAM environment.

Enter REDATAM as described in section 3.1. If you are already in, go to the Main Menu (fig. 3-1) by following the indications on the screen. Select then the Geographic Selection. Figure 3-5 will be displayed.

Make sure that the CREATE FILE option is highlighted and press ENTER. Type a name as follows:

CREATE FILE Name: case2

If this file already exists, add your initials to the name. End your entry with ENTER and at the bottom of the screen enter a description like:

Label: Tutorial Case 2 for Miranda

Another ENTER ends the creation of the file and lets you select the geographical areas. The highest geographical level of Miranda will be shown (fig. 3-6).

Since the task defined calls for selected rural areas, we only want to select parts of any province, thus as indicated on the bottom of the screen, specify a Partial selection by entering a P on the "Sel" column in front of both the Florinda and Perdita provinces. The up and down arrows let you move from one line to the next. Notice that help is available for this screen. Press ? to see how to use the keyboard.

The completed screen should look like this:

<table>
<thead>
<tr>
<th>Code</th>
<th>Area Name</th>
<th>Sel</th>
<th>Recode</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Florinda</td>
<td>p</td>
<td></td>
</tr>
<tr>
<td>02</td>
<td>Perdita</td>
<td>p</td>
<td></td>
</tr>
</tbody>
</table>

Next, press HOME to move automatically to the next incomplete portion of the geography being selected. SHIFT-down arrow (the
down arrow while holding down the Shift key) also allows you to proceed to the next lower geographical level, that of town (fig. 3-7).

Again, since our geographical area of interest includes only a few rural areas, indicate a Partial selection on Florinda's rural area, by entering a P, as in the previous level. The completed screen should look like:

<table>
<thead>
<tr>
<th>Code</th>
<th>Area Name</th>
<th>Sel</th>
<th>Recode</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Florinda</td>
<td></td>
<td></td>
</tr>
<tr>
<td>07</td>
<td>Villages</td>
<td></td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>Rural</td>
<td></td>
<td>p</td>
</tr>
</tbody>
</table>

Another HOME, or SHIFT-down arrow, takes you to the lowest geographical level, that of ED. The screen (with footnotes omitted) should look like:

<table>
<thead>
<tr>
<th>REDATAM V2.0 Geography Selection</th>
<th>Data Base: M180 Date: 25/3/1987</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROVINCE TOWN ED</td>
<td>File: casel.sel</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>Code</td>
<td>Area Name</td>
</tr>
<tr>
<td>01 Florinda</td>
<td></td>
</tr>
<tr>
<td>99 Rural</td>
<td></td>
</tr>
<tr>
<td>089</td>
<td></td>
</tr>
<tr>
<td>090</td>
<td></td>
</tr>
<tr>
<td>091</td>
<td></td>
</tr>
<tr>
<td>092</td>
<td></td>
</tr>
<tr>
<td>093</td>
<td></td>
</tr>
<tr>
<td>094</td>
<td></td>
</tr>
<tr>
<td>095</td>
<td></td>
</tr>
<tr>
<td>096</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 3-14**

You can now see all the rural districts of Florinda. Now select (with a T) all the EDs of interest as indicated at the beginning, that is EDs 95 and 96. After you enter the T, the cursor will move to the Recode column. We will not be using it now, so hit ENTER to return to the Sel column. The screen with the completed items will look like:
Press HOME to continue automatically to the next portion of the incomplete geography. This allows us to select the rural areas of the province of Perdita on the screen depicted in fig. 3-7. In terms of the hierarchy, we have moved "horizontally" to the next province at the same level (it may be helpful to look at the geographical hierarchy diagram of section 3.6 at this point). Just as with Florinda, select the rural area with a P. The resulting screen should be:

Hit HOME in order to proceed to the ED level where you will see:

Proceed exactly as in the previous case, and select EDs 101 and 102 as Total selections. The completed screen is:
When done, hit the HOME key to verify that there are no pending selections and automatically return to the highest geographical level. Hitting END saves the selection and takes you to the Main Menu (fig. 3-1).

You have now defined your geographical area of interest and are ready to proceed to the Statistical Processor. Select that option.

The Statistical Processor is a command driven system (see Tutorial Case 1, section 3.3), and after the Statistical Processor screen appears, it is ready to accept your commands at the bottom line of the screen.

First, tell REDATAM which is the Geographic Selection File to be used by typing:

```
GEOGRAPHY case2
```

which is the file you just defined. Refer to Tutorial Case 1 on how to correct mistakes if REDATAM returns an error message. Commands may be entered in both upper or lower case. For the sake of clarity, commands are shown here in upper case and the arguments in lower case. Also, only as many letters as will uniquely identify a command need to be entered for REDATAM to recognize it.

Remember to end all commands with the ENTER key.

Since the first task calls for a table that is to include only people of ages 6 to 30, we need to restrict our "universe" to this group. This can be done with the UNIVERSE command and by recalling that the variable AGE contains the age of the respondent:

```
UNIVERSE age >= 6 AND age <= 30
```

The UNIVERSE command affects all other commands. It must be issued at the beginning and it may only be used once.

In order to have sufficient detail, we would like to request a set of tables for each ED, with a set of subtotal tables for each town and a set for the whole selected area. For this we issue the command:
AREABREAK town ed

The AREABREAK command provides subtotals at the requested levels for each of the output tables. Note that the geographical levels specified in AREABREAK must be in hierarchical order.

One of the variables that we plan to use, ATTSCH (Attendance at School), is not in the most convenient format, since it has four possible values, two corresponding to people who are in school and two to those who are not (see the Data Dictionary). It would be better if we could group these values into just two alternatives: Attending, or Not Attending. This can be accomplished with the RECODE command. Enter it as follows:

RECODE attsch TO school (1 THRU 2 = 1) (3 THRU 4 = 2)

The RECODE command creates a new variable, which we have named SCHOOL, and whose values correspond to groups of values of the ATTSCH variable. Thus, when ATTSCH is either 1 or 2, SCHOOL will have a value of 1. Likewise, values 3 and 4 correspond to a value of 2 for SCHOOL. In this way, SCHOOL will be 1 if the person is attending school, and 2 if he or she is not.

Note that REDATAM requires blank spaces between operandi.

SCHOOL may be used now like any other variable of the database. You can also add descriptions for it and its values with the following commands:

VAR LABEL school "Attending School"

VALUE LABEL school 1 "Attending" 2 "Not Attending"

With these commands, you have assigned descriptions to both the new variable and to its values. These descriptions will be printed on the output. Note that a maximum of 14 letters are printed when the labels are used as a row variable in a cross tabulation, but only the first 8 letters of each label appear when used as a column variable.

With the VALUE LABEL command we have also set the limits for the range of values that the new variable can take, that is 1 and 2. This is important in order to conserve memory.

Looking at the Data Dictionary, we can determine that TYPESCH shows the highest school level attended, whereas SCHOOL tells us whether the person is attending school or not. We can now directly request the first table with the AVERAGE command:

AVERAGE typesch BY school BY age

At processing time, this will result in a cross-tabulation of the
first two variables with the average of the third variable, AGE, in each cell of the table. A row will be printed for each value of TYPESCH and a column for each of the two possible values of SCHOOL. The total population count for each combination will be printed at the intersection, together with the corresponding average age.

For the next table requested, we need to specify that only children of ages between 6 and 15, and who have attended no more than primary school are to be selected. We can generate a logical expression for this requirement with the TYPESCH and AGE variables. We could then issue an IF command with the following clause:

\[
\text{IF typesch} \leq 3 \text{ AND } (\text{age} \geq 6 \text{ AND age} \leq 15)
\]

However, since we have used a UNIVERSE command, all people whose age is less than 6 have already been excluded, and thus we do not need to specify that condition again. So, just enter:

\[
\text{IF typesch} \leq 3 \text{ AND age} \leq 15
\]

Fewer logical conditions speed up the processing of the command. Unlike the UNIVERSE command, which affects all subsequent commands, this IF command remains in effect only until another IF is issued, setting a new condition.

Note that TYPESCH has a "Not Stated" response available. For a complete study, you might want to look at the number of cases for which this response was used in order to assess the data quality.

Also, we have implicitly assumed that children over 6 are not attending Nursery (code 2 of TYPESCH) and that there are no children of 14 or 15 who have already completed primary school. The latter cases may be excluded by also specifying that the highest educational level attained has to be less than primary (using the variable HIGHSCHT).

We may now request a cross-tabulation between AGE and SCHOOL. Enter the command:

\[
\text{CROSSTABS age BY school}
\]

The CROSSTABS command, like the AVERAGE command, also generates a joint frequency distribution, but does not compute averages. The first variable defines the rows and the second the columns. A row is produced for each value of AGE and a column for each value of SCHOOL. As AGE has up to 100 values and SCHOOL only two, the former must be the row variable in order to print on the paper or screen properly (the number of 9 character columns --8 characters for a label plus one space-- that fit on the paper depend on the pitch, usually 10, 12 or 16 characters per inch, and the size of the paper).
You can now see all the commands entered on a clean screen by typing:

CLEAR

LIST

CLEAR will clear the screen. The following should display:

LIST
GEOGRAPHY CASE2
UNIVERSE AGE >= 6 AND AGE <= 30
AREABREAK TOWN ED
RECODE ATTSCH TO SCHOOL (1 THRU 2=1) (3 THRU 4=2)
VAR LABEL SCHOOL "Attending School"
VALUE LABEL SCHOOL 1 "Attending" 2 "Not Attending"
AVERAGE TYPESCH BY SCHOOL BY AGE
IF TYPESCH <= 3 AND AGE <= 15
CROSSTABS AGE BY SCHOOL

You can now instruct REDATAM to start the actual processing and print the results. Make sure that the printer is on line:

PROCESS OUTPUT = PRINTER

REDATAM should answer with

Tables require 7232 bytes
Writing output to PRINTER
Process started at 15:40:39

and should start processing the request. It may take a couple of minutes before printing, but you will know that the computer is working since the number of census records processed, shown in the upper right corner, will change periodically.

Each output table, looking like figs. 3-16 and 3-17, will be printed for each geographic subdivision specified in the AREABREAK command, that is for each ED and for each Town, as well as for the entire area selected. The meaning of the cross-tabulation tables is explained in Tutorial Case 1 (section 3.3).

Note that in the AVERAGE table, the average age is listed under the total head count in each cell. In all other respects the table is the same as a CROSSTABS table. Sample output is listed on the next page.

There is also a way in which you can ask REDATAM to display the output on the screen and let you browse through it. In order to do this, do not add the OUTPUT = PRINTER clause of the PROCESS command. See more details on how to use this option in the REFERENCE part.
If you would like to work out some other statistics, you can clear all the commands entered by typing:

NEW

When done, you can exit the Statistical Processor and return to REDATAM's Main Menu (fig. 3-1) by entering:

EXIT or EX

You can now go to Tutorial Case 3 or exit REDATAM by hitting the ESC key. This takes you back to the PC's operating system.

<table>
<thead>
<tr>
<th>AVERAGE OF TYPESCH</th>
<th>Type of School Attended</th>
<th>By School</th>
<th>Attending School</th>
<th>By Age</th>
<th>Age of Respondent</th>
</tr>
</thead>
<tbody>
<tr>
<td>GEOGRAPHY IS ED, CODE SEQUENCE IS 1 99 95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CODE</th>
<th>Attendin Not Atte</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>None</td>
<td>1</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td>0.0</td>
<td>20.1</td>
</tr>
<tr>
<td>Nursery - Infant</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>7.4</td>
<td>0.0</td>
</tr>
<tr>
<td>Primary</td>
<td>3</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>9.7</td>
<td>0.0</td>
</tr>
<tr>
<td>Second or Comp.</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>15.8</td>
<td>0.0</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td>92</td>
</tr>
<tr>
<td>TOTAL AVERAGE</td>
<td></td>
<td>9.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

0 VALUES OUT OF RANGE

Figure 3-16
CROSSTAB OF AGE BY SCHOOL

AGE OF RESPONDENT

ATTENDING SCHOOL

GEOGRAPHY IS ED, CODE SEQUENCE IS 1 99 95

<table>
<thead>
<tr>
<th>CODE</th>
<th>Attend</th>
<th>Not Atte</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>14</td>
<td>3</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>13</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>11</td>
<td>14</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>12</td>
<td>7</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>13</td>
<td>10</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>14</td>
<td>4</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>15</td>
<td>2</td>
<td>8</td>
<td>10</td>
</tr>
</tbody>
</table>

TOTAL 88 22 110

0 VALUES OUT OF RANGE

Figure 3-17
3.5. **Tutorial Case 3:**

This example makes use of all the options available for the geographical selection, and shows the use of statistical processing commands that were not shown in the previous examples. At the end of this case you will have used:

- The Geographic Selection partial and total selection options, plus the "recode" option.

- The following Statistical Processor commands:
  AREABREAK
  CLEAR
  COMMENT
  COMPUTE
  CROSSTABS
  EXIT
  GEOGRAPHY
  IF ... INCLUDE
  LIST
  LOAD
  NEW
  OPTION
  PROCESS
  RECODE
  RUN NAME
  SAVE
  VAR LABEL
  VALUE LABEL

In this example we will study the feasibility of providing adult education, in existing schools, for people who live in the rural parts of Miranda and who are "functionally illiterate". By this we will mean people who have had 4 or less years of schooling, are not studying, and are 15 years of age or older. In order to determine the type of program required, we need a cross-tabulation between age and years of primary school. The latter has to be presented in two columns: No attendance at all, and 1 to 4 years. We also want to present age in 5 year groups. Separate tables for each sex are required.

The areas of interest will be Enumeration Districts (EDs) 91, 92, 93, 94, and 103 (see Miranda's map). Two schools are available in these EDs: One in ED 103, (school-1), and the other in ED 94 (school-2). While the school in ED 103 is large and can easily accommodate EDs 91 and 93, periodic flooding of the bridges makes access from ED 93 unreliable, and thus alternative arrangements need to be explored.

People from the different districts could be distributed in the two schools in two ways:
a) School-1: EDs 91, 93 and 103
   School-2: EDs 92 and 94

b) School-1: EDs 91 and 103
   School-2: EDs 92, 93, and 94

EDs will have to be grouped into these areas, and cases a) and b) will have to be studied separately. We have completed checkpoint c) of the checklist of page 3-1.

The following are the Data Dictionary variables that will be used:

AGE is needed for the cross tabulation.
SEX is needed for the different tables
ATTSCCH is required to select those who are not studying.
HIGHSCHY is required in order to determine the number of years studied.

At this point you would decide which statistical commands to use in order to complete point d) of the checklist (page 3-1). For this tutorial, we will analyze the commands as we enter them.

As mentioned, this example assumes certain familiarity with the REDATAM environment. We suggest that if you have not already done so, go through Tutorial Case 1 (section 3.1) first, in order to feel more comfortable with REDATAM.

Enter REDATAM (see section 3.1). If you are already inside REDATAM go to the Main Menu (fig. 3-1) as indicated on the screen. Choose the Geographic Selection option. Figure 3-5 will be displayed.

Move the cursor if needed until the CREATE FILE option is highlighted and press ENTER. Type the following name:

CREATE FILE     Name: case3

If this file already exists, add your initials to the name. End your entry with ENTER and at the bottom of the screen enter a description:

Label: Tutorial Case 3 for Miranda

Another ENTER takes you to the highest geographical level of Miranda (fig. 3-6).

The task defined calls for selected rural areas. Since the EDs we are interested in are both in Florinda and Perdita, specify a Partial selection by entering a P on the "Sel" column in front of
both provinces (see the instruction line at the bottom of the screen). The up and down arrows let you move from one line to the next. Notice that help is available for this screen. Press ? to summon it.

After you have entered a P for both provinces, press HOME to move automatically to the next pending selection or use SHIFT-down arrow (hit the down arrow while holding the Shift key down) in order to proceed to the next lower geographical level, that of town (fig. 3-7).

Since we will study rural areas only, specify a Partial selection on Florinda's rural area, by entering a P, as in the previous level. Another HOME or SHIFT-down arrow takes you to the lowest geographical level, that of ED. The screen (with footnotes omitted) will look like that of fig. 3-14.

Use Miranda's map and the geographical hierarchy diagram at the end of this chapter for reference to the different areas.

You can now see all the rural districts of Florinda. We will now group the EDs according to the school distribution a) stated above. For this, mark (with a T) all the EDs of interest, that is EDs 91, 92, 93, and 94. After you enter each T, the cursor will move to the Recode column. There, type a label to indicate to which of the schools that ED will be associated to. Use the labels S-1 and S-2 for schools 1 and 2 respectively. Hit ENTER to proceed to the next ED after entering the label. If you make a mistake, position yourself at the line requiring correction and reenter both the T in the Sel and label in the Recode column (hit the space bar to eliminate a selection). When you are done, the screen should look like this:

```
<table>
<thead>
<tr>
<th>Code</th>
<th>Area Name</th>
<th>Sel</th>
<th>Recode</th>
</tr>
</thead>
<tbody>
<tr>
<td>091</td>
<td>99 Rural</td>
<td>t</td>
<td>s•1_</td>
</tr>
<tr>
<td>092</td>
<td></td>
<td>t</td>
<td>s•2_</td>
</tr>
<tr>
<td>093</td>
<td></td>
<td>t</td>
<td>s•1_</td>
</tr>
<tr>
<td>094</td>
<td></td>
<td>t</td>
<td>s•2_</td>
</tr>
<tr>
<td>095</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>096</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

**Figure 3-18**

The label under the Recode column will be used to regroup the EDs when we generate the statistics.
We still need ED 103 from the province of Perdita. We hit HOME to move to the next unfinished node. You will see the screen depicted in fig. 3-7.

Just as with Florinda, select the rural area with a P and hit HOME or SHIFT-down arrow in order to proceed to the ED level where you will see the screen shown in fig. 3-15.

In this case, select only ED 103, which is to be associated to school 1 (S-1). Proceed as you did with Florinda. Enter a T, and as the cursor moves to the Recode column type s-1 and hit ENTER. The screen should look like:

```
REDATAM V2.0 Geography Selection
PROVINCE TOWN ED

Code  Area Name  Sel  Recode
02    Perdita
99    Rural
101
102
103  t  s-1
```

**Figure 3-19**

You can review your selections by using the left and right arrows in order to move "sideways" across the same level, and the SHIFT-up arrow and SHIFT-down arrow key combinations in order to move up or down a level. A message will appear if there are no more levels in the requested direction asking you to press ESC. You must press the HOME key to verify that there are no pending selections and to return to the highest geographical level. Move around your geographic selection and verify that it is correct.

When done, hit the END key to tell REDATAM that you wish to save your selections and return to the Main Menu (fig. 3-1). The selection information is saved in file CASE3.SEL, where REDATAM has added the suffix .SEL to your filename.

You have now defined your geographical area and groupings thereof, and are ready to proceed to the Statistical Processor. Select that option.

The Statistical Processor is a command driven system (see Tutorial Case 1, section 3.3), and after its screen appears, it is ready to accept your commands at the bottom line.

First, tell REDATAM which is the Geographic Selection File to be used by typing:

```
GEOGRAPHY case3
```
which is the file just defined; the file suffix .SEL is not included in the name. Refer to Tutorial Case 1 on how to correct mistakes if REDATAM returns an error message. Commands can be entered in both upper or lower case. For the sake of clarity, commands are shown here in upper case and the arguments in lower case. Also only as many letters as will uniquely identify a command need to be entered for REDATAM to recognize it.

We next need to tell REDATAM that all data is to be grouped according to the Recode labels defined in the geographical selection. For this, use the following command:

AREABREAK RECODE

Remember to end all commands with the ENTER key.

This groups together all EDs with the same "recode" label. Tables are generated for each label in addition to a total for the whole selected geography.

The table requested involves only two columns: One for people who have never attended school and another for those that have attended 1 to 4 years. There is no variable with those values in the Data Dictionary, but the HIGHSCHY variable contains the information we require. Thus, we create a new variable that we will name ATTEND by using the RECODE command as follows:

RECODE h highs chy TO attend (1 = 1) (2 THRU 3 = 2) (4 THRU HIGHEST = 3)

The command should fit in just one line on the screen.

The new variable, ATTEND, has 3 possible values, each one corresponding to a range of values of HIGHSCHY. Spaces between the "(" , ")", "-" and "=" symbols should be respected. Also note the use of HIGHEST. It means that all values higher than 4 are to be assigned to 3. A LOWEST clause could also be used. Note that the values of HIGHSCHY (to the left of the = sign), must be in ascending order.

ATTEND will then have a value of 1 for people who have never attended school, 2 for those who have studied 1 to 4 years and 3 for the rest.

We also want to present the age in 5 year groups. We could do that by using the RECODE command again, but it is easier to use the COMPUTE command as follows:

COMPUTE age5 = (age / 5) + 1

That is, the new variable AGE5, is defined as the age divided by 5 plus 1. Since tables are printed with truncated values of the
variables, AGE5 will have a value of 1 for ages 1 through 4, a value of 2 for ages 5 to 9, and so forth. As in the RECODE command, spaces between operandi may not be omitted.

The new variables can be used as if they had been added to the Data Dictionary. We can also provide a description for them and for each of the values they take. These descriptions will appear in the tables generated. It is also desirable to tell REDATAM the range of values the new variables can take, so that no more than the necessary memory space is reserved. The VAR LABEL command provides descriptions for the variables, while the VALUE LABEL command provides descriptions for its values as well as setting their range:

VAR LABEL attend "Years of School"
VALUE LABEL attend 1 "None" 2 "1-4 Yrs" 3 "5 or More"

VAR LABEL age5 "Age Groups"
VALUE LABEL age5 4 "15 - 19" 5 "20 - 24" 6 "25 - 29",
7 "30 - 34" 8 "35 - 39" 9 "40 - 44" 10 "45 - 49"

As you type the last command, you will note that the command and arguments are longer than what fits in one line on the screen. You can either break line with a comma (,) as shown above, which tells REDATAM that the command continues on the next line, or you may repeat the command, such as in

VALUE LABEL age5 4 "15 - 19" 5 "20 - 24" 6 "25 - 29"
VALUE LABEL age5 7 "30 - 34" 8 "35 - 39" 9 "40 - 44"
VALUE LABEL age5 10 "45 - 49"

Note the spaces between numbers and symbols. They must be respected.

With the VAR LABEL command you have assigned descriptions to the ATTEND and AGE5 variables, and with the VALUE LABEL you have defined labels for its values. Note that we only assigned descriptions for age ranges starting at 15 years, since this is the minimum age that we are interested in. Also note that labels for several variables could have been assigned with one command. Separate commands are used here for the sake of clarity.

When large groups of commands are entered and saved for later use (as we shall see presently), it is sometimes useful to enter some comments about what is done in order to assist in understanding them later. For this you can use the COMMENT command as in:

COMMENT AGE and HIGHSCHY are recoded into AGE5 and ATTEND

The COMMENT (or COMM, its abbreviation) command is only saved,
and has no effect on the data processed by REDATAM.

Continuing with our example, in this case we are interested in "people who have had 4 or less years of schooling and are 15 years of age or older". You can filter-out all cases that do not match this description. Enter:

\[
\text{IF attend} < 3 \text{ AND age} \geq 15 \text{ INCLUDE}
\]

The IF ... INCLUDE command excludes all records that do not comply with the logical condition specified. You can verify that according to the values we assigned to the ATTEND variable, the logical condition corresponds to the selection statement.

Unlike the simple IF command, IF ... INCLUDE affects all commands that follow it and its effect cannot be reversed by further IFs. It is similar to the UNIVERSE command, except that several IF ... INCLUDEs may be issued, successively restricting the record set. Also, the UNIVERSE command has to be issued at the beginning, which would have been impossible here since the ATTEND variable had to be defined before. On the other hand, UNIVERSE processes much faster, and should be used whenever possible, when record selection is required.

We are now ready to generate the cross-tabulation requested. Type the following command:

\[
\text{CROSSTABS age5 BY sex BY attend}
\]

Remember that the output is not generated until the PROCESS command is used. You will notice that a 3-way cross tabulation has been requested here. A matrix with the joint frequency distributions of AGE5, SEX and ATTEND will be printed. SEX will make up the rows, and the values of ATTEND the columns. A matrix will be printed for each value of AGE5.

We can also request that along with the population counts, relative percentages be printed. For instance, if we want to know the percent of all people that each age group represents, both in relation to the total of each ATTEND category and in relation to the total selected population, we enter the following command:

\[
\text{OPTION PERCENT COLUMN TOTAL}
\]

This adds column-wise and total percentages to the tallies at each row - column intersection. Horizontal or row totals could also be requested.

You can now list all the commands entered, but first clear the screen for clarity:

\[
\text{CLEAR}
\]
LIST

The screen should show:

GEOGRAPHY CASE3
AREABREAK RECODE
RECODE HIGHSCHY TO ATTEND (1 = 1) (2 THRU 3 = 2) (4 THRU
HIGHEST = 3)
COMPUTE AGES5 = (AGE / 5) + 1
VAR LABEL ATTEND "Years of School"
VALUE LABEL ATTEND 1 "None" 2 "1-4 Yrs" 3 "5 or More"
VAR LABEL AGES5 "Age Groups"
VALUE LABEL AGES5 4 "15 - 19" 5 "20 - 24" 6 "25 - 29",
7 "30 - 34", 8 "35 - 39" 9 "40 - 44" 10 "45 - 49"
COMMENT AGE AND HIGHSCHY ARE RECODED INTO AGES5 AND ATTEND
IF ATTEND < 3 AND AGE >= 15 INCLUDE
CROSSTABS AGES5 BY SEX BY ATTEND
OPTION PERCENT COLUMN TOTAL

Recall that the task calls for two different groupings of the EDs (different school assignments). This means that after generating the tables for this case we will have to go back to the Geographic Selection, modify it, and return to the Statistical Processor. In order to avoid reentering all the commands, they can be "saved" on the disk to recall them at a later moment. For this use the SAVE command:

SAVE case3

The name following the SAVE command tells REDATAM under which name should the command set be saved. We have given it, for convenience, the same name as that of the Geographic Selection File. This will not cause any problem to REDATAM. It can distinguish the files from each other since selection files are given automatically the suffix .SEL and command files the suffix .ECF. If the file already exists, REDATAM will so inform you and ask you to confirm (y for "yes" and n for "no") before overwriting it.

You can now request the tables to be printed. However, since we will generate the same set of tables with a different geography later again, it is useful to be able to distinguish the "runs". We can label them as follows:

RUN NAME "Adult Education Study, School distribution A"

This header will be printed at the top of each output page, making it easy to identify it later. Now, making sure that your printer is on-line, enter:

PROCESS OUTPUT = PRINTER

REDAATAM should answer with
Tables require 1920 bytes
Writing output to PRINTER
Process started at 16:42:54

It may take a couple of minutes before REDATAM starts printing. The output consists of a cross tabulation for each school area (that is labels S-1 and S-2) and for the entire selected area. The meaning of the cross-tabulation table is explained in Tutorial Case 1 (section 3.3, fig. 3-13). The tables should look like fig. 3-20, on the next page.

Note that the requested percentages are printed under the tallies of each cell. A table is printed for each re-coded geography plus a total table for the combined subdivisions.

Also note that no column is printed for value 3 of ATTEND. This means that no cases were found. REDATAM does not print rows or columns that only have zeroes.


CROSSTAB OF AGE5 Age Groups
BY SEX Sex of Respondent
BY ATTEND Years of School

GEOGRAPHY IS RECODED by e-1

<table>
<thead>
<tr>
<th>CODE</th>
<th>None</th>
<th>1-4 yrs</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>4</td>
<td>1</td>
<td>3.7%</td>
</tr>
<tr>
<td>Column Percent</td>
<td>1.9%</td>
<td>3.7%</td>
<td>2.5%</td>
</tr>
<tr>
<td>Total Percent</td>
<td>1.3%</td>
<td>1.3%</td>
<td>2.5%</td>
</tr>
<tr>
<td>Female</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Column Percent</td>
<td>0.0%</td>
<td>3.7%</td>
<td>1.3%</td>
</tr>
<tr>
<td>Total Percent</td>
<td>0.0%</td>
<td>1.3%</td>
<td>1.3%</td>
</tr>
<tr>
<td>20-24</td>
<td>5</td>
<td>1</td>
<td>3.7%</td>
</tr>
<tr>
<td>Male</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Column Percent</td>
<td>1.9%</td>
<td>3.7%</td>
<td>2.5%</td>
</tr>
<tr>
<td>Total Percent</td>
<td>1.3%</td>
<td>1.3%</td>
<td>2.5%</td>
</tr>
<tr>
<td>Female</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Column Percent</td>
<td>3.8%</td>
<td>7.4%</td>
<td>5.0%</td>
</tr>
<tr>
<td>Total Percent</td>
<td>2.5%</td>
<td>2.5%</td>
<td>5.0%</td>
</tr>
<tr>
<td>25-29</td>
<td>6</td>
<td>1</td>
<td>3.7%</td>
</tr>
<tr>
<td>Male</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Column Percent</td>
<td>1.9%</td>
<td>3.7%</td>
<td>2.5%</td>
</tr>
<tr>
<td>Total Percent</td>
<td>1.3%</td>
<td>1.3%</td>
<td>2.5%</td>
</tr>
<tr>
<td>etc.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 3-20

When done, exit the Statistical Processor and return to REDATAM's Main Menu (fig. 3-1) by entering:
EXIT

For the second part of this example, we will repeat the study changing the grouping of the EDs according to the alternative b) for school assignment proposed at the beginning of this section. By typing EXIT you were returned to the Main Menu. Return to Geographic Selection by selecting that option. Figure 3-5 will be displayed.

Move the cursor until the file CASE3, which you previously defined, is highlighted and press ENTER. REDATAM will select CASE3 and take you to the highest geographical level of Miranda (fig. 3-6). The selections you already made will be shown (that is a P for both Florinda and Perdita).

For this case, you have to reassign the EDs to different labels. The HOME key will not move you to pending levels since they are all complete. Therefore, in this situation, in which a geographic selection is being modified, you must use SHIFT-up and down arrows, and the right and left arrows as indicated on the help screen.

For case b), ED 93 needs to be reassigned to S-2. Thus reenter the T under the Sel column of the corresponding line in order to move the cursor to the Recode column and change the label so that the screen looks as follows.

```
REDATEM V2.0 Geography Selection
PROVINCE TOWN ED

Data Base: M180 Date: 25/3/1987
File: case1.sel

Code Area Name Sel Recode
01 Florinda 089 t s·1
99 Rural 090 t s·2
091 t s·2
092
093 t
094
095
096
```

**Figure 3-21**

Since there are no more changes to be made, hit HOME to verify that there are, in fact, no pending selections and then END to save the information and return to the Geographic Selection Menu (Fig. 3-4). The previous Geographic Selection File has been replaced with the changes you just made. If we wanted to keep it for later use, we could have created a new selection file under a different name, or have used the DOS COPY command in order to
start with a copy of the original file.

From the Main Menu (fig. 3-1) select the Statistical Processor option. Then enter:

LOAD case3

This restores the whole command set just as it was when you entered the SAVE command earlier in this tutorial. REDATAM rechecks the command syntax.

Since the geography file now contains the new distribution, if you now issue the PROCESS command, you will obtain the same tables as for case a), but for the new geography. Do so, but also provide a new RUN NAME for proper identification of the output:

RUN NAME "Adult Education Study, School distribution B"

PROCESS OUTPUT = PRINTER

Make sure that the printer is on-line. Compare the two outputs.

You have now repeated the whole process with but a minimum number of keystrokes, by using your existing case. More cases could now be studied, further changing the geography file or the command set. For the latter, there is a Statistical Processor Editor called by the REVIEW command (or REV for short), that allows you to edit the commands entered. You can change, delete or insert commands, or extract a group of them to include them into another file. When you are ready for this, look at SPECIAL FEATURES (Part 5) to see a demonstration.

You can also review the output on the screen before printing it, which is accomplished issuing the PROCESS command without any arguments. The output can also be sent to a disk file that you can later edit with a word processor or enter into a spreadsheet like Lotus 123. To do this, replace the PRINTER clause in the PROCESS command with the name of the file you want to create.

There is also a WRITE command that lets you directly extract records from the REDATAM census file into the file you specify. A variety of output formats are available that let you process the selected records by other packages like SL-MICRO, SPSS-PC, etc. See the Reference (Part 4) and SPECIAL FEATURES (Part 5) to see how to use it.

You can end your processing by typing:

EXIT

or if you want, you can generate new statistics of your own. Use the NEW command to erase the present command set, typing:
This tutorial has given you with a quick look at some of the potentials of REDATAM. Additional commands and options beyond those presented in this tutorial are available. If there is something special you need to do and require additional help, take a look at the REFERENCE (Part 4). You might just find what you are looking for. Also, SPECIAL FEATURES (Part 5) contains examples and tutorials for the more advanced features of REDATAM.

Happy REDATAMing....
3.6. **Miranda Geography and Data**:

Miranda is an island. It is divided into two Provinces: Florinda and Perdita. Each of the provinces has a Town and a large rural area. The towns are Vilches and Tumaco and belong to the provinces of Florinda and Perdita respectively. The towns and the rural areas are divided into Enumeration Districts, EDs for short. EDs are numbered 84 through 103.

You will find a map of Miranda on the next page.

Miranda's subdivisions form a hierarchical structure, that may be represented by the following diagram:

```
Level 0:  
  Country

  ------------  
     MIRANDA    
  ------------

Level 1:  
  Province

  ------------  
   FLORINDA    
   PERDITA    
  ------------

Level 2:  
  Town

  ------------  
   VILCHES    
   RURAL    
   TUMACO    
   RURAL    
  ------------

Level 3:  
  ED

  84  85 | 89  90 | 97  98 | 101 102
  86  87 | 91  92 | 99 100 | 103   
  88     | 93  94 |        |       
        | 95  96 |
```
Part 4: REFERENCE: Functions and commands

This section provides detailed descriptions for the use of the REDATAM system. Each function and command is described separately to facilitate quick consultation and look-up.

It is assumed in this section that you have a good idea of the purpose and philosophy of REDATAM, and that you have some familiarity in the use of your microcomputer. If this is not the case, see the previous sections of this manual. It is also assumed that REDATAM has been installed on the microcomputer. If this is not so, see APPENDIX B.

REDATAM is started directly from DOS (the Disk Operating System of the PC). The C> prompt should be on the screen in order to load. Each user should have a work directory where his or her selection and command files are stored. This directory is selected with:

```
CD \directory [ENTER]
```

Where "directory" is to be replaced by the name of the user's directory. [ENTER] represents the ENTER key. A special batch file is required for each database to be used. Refer to APPENDIX B for instructions on how to prepare this file. Enter:

```
batch_file [ENTER]
```

Where "batch_file" is to be replaced by the actual name of the file, which usually corresponds to the name of the database to be used. REDATAM should start loading. In a few seconds, the following screen appears:

```
| REDATAM System 2.0 | Database: M180 | Date: 02/07/1986 | Time: 00:56:58 |
```

```
| Statistical Processor |
| Geographic Selection |
| Data dictionary |
| database management |
| Exit |

\'v Moves the cursor. Select with ENTER or first letter. ESC to end
This is the REDATAM MAIN MENU. It controls access to all system functions. Some general system data is displayed at the top of the screen, and instructions at the bottom.

REDATAM supports multiple languages. See APPENDIX B in order to change the language in which screens are to appear.

Options:

Statistical Processor:
This option loads the Statistical Command Processor. This is where the actual statistics are generated. They are described in Reference Section 4.1.

Geographic Selection:
This option accesses the Geographic Selection module, where the geographic areas for which statistics will be performed are defined. The option is described in Reference Section 4.2.

Data Dictionary:
This option accesses the Data Dictionary Processor to perform inquiries on, select and maintain Data Dictionaries. It is described in Reference Section 4.3.

dataBase management:
This option is used to create and load new REDATAM Databases. It is described in section 4.4.

Exit:
This option exits REDATAM and returns to the PC's operating system.

Usage:
An option is selected by either moving the up and down arrows until it is highlighted, and then pressing ENTER, or by directly pressing the letter that appears underlined and in capitals for the desired option.

The ESC key or the Exit option can be used to exit REDATAM and return to the operating system.
4.1. **Statistical Processor**:

This section contains an alphabetical list of all the commands available with the Statistical Processor.

It is in the Statistical Processor that the actual statistics are generated, after the geographical areas of interest have been defined with the Geographic Selection Processor.

The Statistical Processor offers a series of statistical functions and facilities to select records, print and format output, interface with other computer packages, etc.

The Statistical Processor is a command driven system. The user is presented with a blank screen, where commands are entered at the bottom and are terminated with ENTER. Commands are checked for syntax and consistency as they are entered and moved to the upper part of the screen. As a general rule, no further messages are displayed if the command is accepted.

If an error is found, a message is shown and the command is redisplayed at the bottom of the screen, where it must be edited or erased by using the following keys:

- The **left** and **right arrow** keys can be used to position the cursor within the command line.

- The **HOME** key moves the cursor to the left-most position on the command line.

- The **END** key moves the cursor to the right end of the command just entered.

- The **ESC** key erases the entire command line.

- The **DEL** key erases the character at the current cursor position.

- The **Backspace** key erases the character to the left of the cursor and moves the remaining characters one space to the left.

Note that while editing the command, REDATAM is in the "insert" mode, that is, new characters are inserted in front of the cursor position.

After correcting the command, pressing ENTER instructs REDATAM to check it again.

Help is available in the Statistical Processor by pressing the ? key in conjunction with the **SHIFT** key. When this is done, the following screen appears:
Following the instructions on the bottom line, the up and down arrows and left and right arrows are used to highlight the command for which information is required. Pressing ENTER shows a brief description, syntax and example of the command. The up and down arrows are then used to see the previous or next pages, while PgUp and PgDn are used to respectively see the previous or next command. ESC is used to choose another command, or if none is selected, to exit Help and return to the Statistical Processor.

Only as many letters as are necessary to uniquely identify the command need to be typed. For instance, if B is entered, REDATAM will assume that BROWSE was entered, since no other command starts with that letter. On the other hand, for the COMPUTE command, at least COMP needs to be entered in order to distinguish it from the COMMENT command. For the sake of clarity, commands are shown in their full syntax in this section.

It is important to note that the geographical variables (for example, TOWN, PROVINCE or ED) are available for table generation or data manipulation only up to the lowest level specified in the Geographic Selection. For instance, if the total selections have been specified at the TOWN level, the PROVINCE and TOWN variables are available but not the ED variable. Also note that no descriptions are provided for the values of these variables in tabular output unless specifically entered in the Data Dictionary. The geographic variables are identified by a "record type" of 0.

This section contains a short alphabetical list of commands with their syntax, followed by a detailed description of each command, including syntax, comments and examples. Examples are based on the dictionary included at the end of the TUTORIAL (Section 3.6).

The following notation convention is used:

a) If a command has more than one syntactical form, each alternative is shown on a separate line. For greater clarity, if a command has too many arguments to fit in one line, several lines are shown.
b) Required arguments are spelled in CAPITAL letters. They are to be typed as they appear.

c) If a command contains required arguments supplied by the user, these are indicated between the < and > signs in lower case letters. For instance, in the command:

LOAD <file name>

the expression <file name> is to be replaced by the actual file name, as in:

LOAD MYFILE

or

load myfile

As shown, commands and arguments can be entered in either upper case, lower case or mixed mode. For clarity, lower case is used here to show user selected input.

d) When a command includes optional items, these are enclosed in square brackets, like:

COMMENT [comments]

Note that commands may continue on more than one line, if they are longer than the 80 character screen width. A command may be broken at any point with a comma (,), which instructs REDATAM that the command continues on the next line. Virtually any number of continuation lines may be used.
LIST OF COMMANDS

Each command is shown in all its alternative forms and syntaxes. Minimum abbreviations are shown below in parenthesis. The symbols "var" and "val" are often used for "variable" and "value" respectively.

AREABREAK <geographic subdivision 1> [geographic subdivision 2] [...]

(AR)

AREABREAK RECODE

(AR RECODE)

AVERAGE [var 1] BY [var 2] BY <var 3> BY <var 4> BY <var 5>
    (panel 1) (panel 2) (row) (column) (average)

(AV)

BROWSE

(B)

CLEAR

(CL)

COMMENT [comments]

(COMM)

COMPUTE <variable> = <algebraic expression>

(COMP)

CROSSTABS [var 1] BY [var 2] BY <var 3> BY <var 4>
    (panel 1) (panel 2) (row) (column)

(CR)

DIRECTORY

(D)

ENDHP

(EN)

EXIT

(EX)

FREQUENCIES <variable 1> [variable 2] [variable 3] [...]

(F)

GEOGRAPHY <file name>

(G)

IF <condition>

(I)
IF <condition> COUNT <new variable> (I)

IF <condition> EXCLUDE (I)

IF <condition> INCLUDE (I)

IF <condition> QUALIFY (I)

IF <condition> SUM <new variable> = <existing variable> (I)

LIST (LI)

LOAD <file name> (LO)

NEW [Y] (N)

OPTION [FILENAME = <file name>] [DOCUMENTATION = [DOS] [PRINT] [REDATAM] [SPSS] [SLMICRO]] (O)

OPTION COUNT <level number> (O COUNT)

OPTION PERCENT [ROW] [COLUMN] [TOTAL] (O PERCENT)

PRINT (PRI)

PROCESS [OUTPUT = PRINTER]
PROCESS [OUTPUT = <file name>] (PRO)

RECODE <existing variable> TO <new variable> <recode list>
RECODE <existing user variable> <recode list>

where <recode list> may be any of the following:

(<existing val 1> [THRU <existing val 2>] = <new val 1>)
[(<existing val 3> [THRU <existing val 4>] = <new val 2>)]
[...] ...

(<existing val 1> [- <existing val 2>] = <new val 1>)
[(<existing val 3> [- <existing val 4>] = <new val 2>)]
[...] ...
( existing val 1> [, existing val 2> ] = <new val 1>)
[( existing val 3> [, existing val 4> ] = <new val 2>)
[...] ... 
(REC)

(REV)

RUN NAME "[text]"
(RU NAME)

SAVE <file name>
(S)

UNIVERSE <condition>
(U)

VALUE LABEL <variable> <value 1> "<description>"
[value 2] "[description]"
[...] ...
(VAL LABEL)

VAR LABEL <variable 1> "<description>"
[variable 2] "[description]"
[...] ...
(VAR LABEL)

VIEW
(VI)

WEIGHT <algebraic expression>
(WE)

WRITE [record #] <variable 1> [(Fn.m)] [variable 2] [(Fn.m)]
[...] ...
WRITE <variable 1> [variable 2] [...] ...
[/] [variable 3] [variable 4] [...] ...
(WR)
Usage: The AREABREAK command is used to request tables for smaller areas within the overall selected geography, for example, tables for each block within a local planning area. If not used, each table requested is presented once, with an overall total table, for the area defined in the Geographic Selection File.

Syntax: AREABREAK <geographic subdivision 1>
          [geographic subdivision 2] [...] ...

(AR) -- Minimum abbreviation

Notes: The geographic subdivisions specified, are used for control breaks with separate tables for each level indicated. A table at the global level, that is, for all the records included in the Geographic Selection File (modified by the UNIVERSE and/or IF commands) is also provided.

The "geographic subdivision" names must be valid Data Dictionary geographic variables and must have been included in the Geographic Selection File as either Partial or Total selections.

Example: Making reference to the TUTORIAL example (Section 3.6), if a table is requested, and the following command is issued:

AREABREAK province town

REDATAM will generate totals and subtotals for the following combinations and in the following order:

- Table for the town of Vilches in Florinda
- Table for the rural part of Florinda
- Table for all selected areas of Florinda (global level)
- Table for the town of Tumaco in Perdita
- Table for the rural part of Perdita
- Table for all selected areas of Perdita (global level)
- Table for all selected areas of Miranda (global level)

Note that even if the whole country of Miranda is selected for this example, each of the Provinces, Towns and rural areas has to be explicitly selected (with a "T" or "P" as appropriate) during the Geographic Selection if the AREABREAK command is to provide subtotals for each of the subdivisions. For example, in order to obtain the above tables for the whole country, a geographic selection that includes at least the fol-
lowing detail is required:

REDATAM V2.0 Geography Selection
PROVINCE TOWN ED

<table>
<thead>
<tr>
<th>Code</th>
<th>Area Name</th>
<th>Sel</th>
<th>Recode</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Florinda</td>
<td>p</td>
<td></td>
</tr>
<tr>
<td>02</td>
<td>Perdita</td>
<td>p</td>
<td></td>
</tr>
</tbody>
</table>

REDATAM V2.0 Geography Selection
PROVINCE TOWN ED

<table>
<thead>
<tr>
<th>Code</th>
<th>Area Name</th>
<th>Sel</th>
<th>Recode</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Florinda</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01</td>
<td>Vilches</td>
<td>t</td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>Rural</td>
<td>t</td>
<td></td>
</tr>
</tbody>
</table>

REDATAM V2.0 Geography Selection
PROVINCE TOWN ED

<table>
<thead>
<tr>
<th>Code</th>
<th>Area Name</th>
<th>Sel</th>
<th>Recode</th>
</tr>
</thead>
<tbody>
<tr>
<td>02</td>
<td>Perdita</td>
<td></td>
<td></td>
</tr>
<tr>
<td>01</td>
<td>Tumaco</td>
<td>t</td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>Rural</td>
<td>t</td>
<td></td>
</tr>
</tbody>
</table>

Should an AREABREAK on EDs be also desired, then each of them has to be selected individually.

If two or more tables are requested for the specified area, each one is printed for each of the specific areas of the above list.

See Also: AREABREAK RECODE
AREABREAK RECODE

Usage: The AREABREAK RECODE command is used to request tables for selected areas within the overall geography, grouped according to the recode labels specified in the geographic selection.

Syntax: AREABREAK RECODE

(AR RECODE) -- Minimum abbreviation

Notes: By using this command, records are grouped by the labels specified for the Recode field in the Geographic Selection. Subtotal tables are printed for each label, in addition to a total table for the entire area.

Example: If the following geographic selection is specified:

```
REDATAM V2.0 Geography Selection
PROVINCE TOWN ED
Code    Area Name   Sel  Recode
01      Florinda    p
02      Perdita     p
```

```
REDATAM V2.0 Geography Selection
PROVINCE TOWN ED
Code    Area Name   Sel  Recode
01      Florinda
01      Viches    t    urban
99      Aural     t    rural
```

```
REDATAM V2.0 Geography Selection
PROVINCE TOWN ED
Code    Area Name   Sel  Recode
02      Perdita
01      Tumaco    t    urban
99      Rural     t    rural
```

When the command AREABREAK RECODE is issued, each table requested is printed for the combined sub-totals of all areas labeled with "urban" and all areas labeled with "rural", in addition to a table for all the selected geography.

See Also: AREABREAK
The AVERAGE command is similar to the CROSSTABS command, in that it instructs REDATAM to generate a joint frequency distribution table of cases according to the classification variables, but in addition, it calculates the average of a third dependent variable.

**Syntax:**

```
AVERAGE [var1] BY [var2] BY <var3> BY <var4> BY <var5>
(panel 1) (panel 2) (row) (column) (average)
```

(AV) -- Minimum abbreviation

**Notes:**

At least three variables have to be provided. A line is printed for each value of the row variable, and a column for each value of column variable. The tallies for the occurrences of each combination are printed at the intersection, together with the average of variable 5. If additional variables are provided, a row-column panel or matrix is printed for each combination of their values. Up to four classification variables may be provided.

Horizontal and vertical totals are also printed.

The variables specified must be in the Data Dictionary, or have been previously defined by the user, and if they are geographical variables, they must have been explicitly selected in the Geographic Selection, as indicated at the beginning of this section.

Lines or columns with only zeroes are not printed. For row variables, 14 letters of the value labels are printed, while for column variables only the first 8 characters are printed. Note that a column has a total of 9 characters since one is a blank to separate the columns.

When REDATAM tries to print more columns than may fit within the width of a page, the line "wraps around", making output somewhat difficult to understand. The following limitations need to be observed in order to avoid this:

Up to 7 columns fit on a regular sheet of paper (9.5 inches wide) using normal print type. If compressed type or wide (15 inch) paper is used, 12 columns can fit. 22 columns fit on wide paper when compressed type is used.
Example: AVERAGE highscht BY occstat BY age

would generate the following table:

AVERAGE OF HIGHSCHT Highest Educational Level, Type School
BY OCCSTAT Occupational Status
BY AGE Age of Respondent

GEOGRAPHY IS GLOBAL

| Code | Governor Private | Private Unpaid W With Pai W/o Paid Did No W Not Stat TOTAL |
|------|------------------|------------------------------------------|------------------|
|      | Gover 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| None | 640     | 7 | 133 | 5 | 3 | 16 | 111 | 210 | 10 | 1135 |
| Nursery - Infant | 287 | 3 | 12 | 4 | 2 | 3 | 27 | 30 | 2 | 470 |
| Primary | 851 | 89 | 377 | 39 | 21 | 39 | 235 | 716 | 23 | 2390 |
| Second or Comp | 64 | 8 | 6 | 0 | 1 | 1 | 2 | 10 | 2 | 94 |
| Other Second | 6 | 1 | 2 | 1 | 0 | 1 | 0 | 1 | 5 | 11 |
| University | 7 | 0 | 1 | 2 | 0 | 0 | 0 | 2 | 0 | 6 |
| Other | 8 | 1 | 5 | 2 | 0 | 0 | 0 | 6 | 1 | 15 |
| Not Stated | 9 | 0 | 0 | 35 | 0 | 0 | 0 | 63 | 0 | 33.7 |
| TOTAL | 1944 | 115 | 534 | 48 | 28 | 59 | 378 | 979 | 41 | 4126 |
| TOTAL AVERAGE | 7.5 | 29.5 | 35.7 | 32.9 | 29.2 | 48.7 | 44.1 | 36.3 | 38.2 | 23.3 |

6 VALUES OUT OF RANGE

The values and descriptions (truncated to the column width) of the first two variables are displayed in row-column format. The population counts for each combination are displayed at the intersections, together with the average value of the third variable (AGE in this case). The number of values out of range is indicated at the end.

If either HIGHSCHT or OCCSTAT had non-integer values, they would be truncated, as occurs in all other REDATAM tables.

See Also: OPTION COUNT, CROSSTABS, FREQUENCIES
Usage: This command interactively displays a list of names and descriptions of the variables of the current Data Dictionary.

Syntax: BROWSE

(B) -- Minimum abbreviation

Notes: This command is similar to the DATA DICTIONARY BROWSE FUNCTION and the BROWSE option within the Statistical Processor Editor (REVIEW). It allows a quick perusal of the available variables, thus obviating the need to have a Data Dictionary listing to look up variable descriptions, their exact spelling, labels, etc.

For details on its use, see the Keyboard Usage (Section 4.5) and the DATA DICTIONARY BROWSE FUNCTION, APPENDIX D.

While in BROWSE, pressing ESC returns to the Statistical Processor.
Usage: The CLEAR command clears the screen.

Syntax: CLEAR

(CL) -- Minimum abbreviation

Notes: CLEAR has no other effect on the commands already entered. The commands in the current set can be re-displayed on the screen with LIST. To actually erase the command set from memory, use NEW.

See Also: LIST, NEW
Usage: The COMMENT command provides the ability to enter descriptive text in the command set. This is ignored at processing time.

Syntax: COMMENT [comments]

(COMM) -- Minimum abbreviation

or

* [comments]

Notes: This command can be useful to document long sequences of commands that are stored for later use.

Example: COMM This command set is used for the Adult Education
        COMM Feasibility Study

* This command set is used for the Adult Education
* Feasibility Study

This may also be entered as:

COMM This command set is used for the Adult Education, Feasibility Study

See Also: RUN NAME
Usage: The COMPUTE command performs arithmetic operations on variables.

Syntax: COMPUTE <new variable> = <algebraic expression>

(COMP) -- Minimum abbreviation

Notes: This command computes the algebraic expression and stores it in the specified variable. The variable must be new. It may not be a Data Dictionary variable or have been created by a previous COMPUTE or other command. The variable may then be used as any other variable of the data file. The four basic arithmetic operations (+, -, *, /) plus exponentiation (^) and parenthesis may be used. Expressions are evaluated according to the rules of algebra.

The variables used in the algebraic expression must be in the Data Dictionary, or have been previously defined by the user, and if they are geographical variables, they must have been explicitly selected in the Geographic Selection, as indicated at the beginning of this section.

Blank spaces between arithmetical operandi and variables are required.

The new variable name can be up to eight characters in length and must start with a letter.

For maximum speed, the results of a COMPUTE are stored as integers. Hence, the fractional parts that may result from the arithmetic operations are truncated, that is, 3.8 would be stored as 3. If the decimal is needed for later use, the value should be multiplied, for example by 10. in the COMPUTE expression. Note that negative values resulting from a COMPUTE are stored and can be used in other commands, but they are not included in an output table.

If the range of the new variable is not defined via the VALUE LABEL command, REDATAM reserves space for the range 0 to 99.

In hierarchical files, the variables defined by COMPUTE, are considered population variables if at least one population variable is involved in the calculation. They are considered housing variables otherwise.
Example: COMPUTE lbirths = livebrth / (livebrth + stilbrth)

creates a variable called LBIRTHS that is the result of dividing the number of live births by the sum of live and still births. By using the VAR LABEL and VALUE LABEL commands, descriptions for the new variable and its values may be provided, as well as specifying the value range.

See Also: VAR LABEL, VALUE LABEL, RECODE
Usage: CROSSTABS generates a joint frequency distribution table with two, three or four variables. The result is shown as a matrix of population counts.

Syntax: CROSSTABS [var 1] BY [var 2] BY <var 3> BY <var 4>
        (panel 1) (panel 2) (row) (column)

        (CR) -- Minimum abbreviation

Notes: At least two variables have to be provided. A line is printed for each value of the row variable, and a column for each value of the column variable. The tallies for the occurrences of each combination are printed at the intersection.

If three or four variables are used, a table of variable 3 by variable 4 is printed for each combination of values of variables 1 and/or 2.

The variables specified must be in the Data Dictionary, or have been previously defined by the user, and if they are geographical variables, they must have been explicitly selected in the Geographic Selection, as indicated at the beginning of this section.

Horizontal and vertical totals are computed. With the use of the OPTION PERCENT command, row, column and/or total percentages may be requested. The OPTION PERCENT command must immediately follow the CROSSTABS command to which it refers.

Lines or columns with only zeroes are not printed. For row variables, 14 letters of the value labels are printed, while for column variables only the first 8 characters are printed. Note that a column has a total of 9 characters since one is a blank to separate the columns.

When REDATAM tries to print more columns than may fit within the width of a page, the line "wraps around", making output somewhat difficult to understand. The following limitations need to be observed in order to avoid this:

Up to 7 columns may fit on a regular sheet of paper (9.5 inches wide) using normal print type. If compressed type or wide (15 inch) paper is used, 12 columns fit. 22 columns can fit on wide paper when compressed type is used.
Example: CROSSTABS sex BY highscht BY occstat

would generate the following table:

<table>
<thead>
<tr>
<th></th>
<th>Governme</th>
<th>Private</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>8</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>Primary</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Second or Comp</td>
<td>2</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Multi High</td>
<td>7</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Other Second</td>
<td>5</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>University</td>
<td>13</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>67</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td>Not Stated</td>
<td>31</td>
<td>60</td>
<td>3</td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>1</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>Primary</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Second or Comp</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Multi High</td>
<td>6</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Other Second</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>University</td>
<td>11</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Other</td>
<td>36</td>
<td>26</td>
<td>3</td>
</tr>
<tr>
<td>Not Stated</td>
<td>19</td>
<td>40</td>
<td>6</td>
</tr>
<tr>
<td>TOTAL</td>
<td>212</td>
<td>193</td>
<td>22</td>
</tr>
</tbody>
</table>

0 VALUES OUT OF RANGE

The values and descriptions (truncated to the column width) of the variables are displayed in row-column format. The population counts for each combination are displayed at the intersections. One panel is generated for each value of SEX (two in this case). The number of values out of range is indicated at the end.

If either HIGHSCHT or OCCSTAT had non-integer values, they are truncated for this table. Negative values are considered out of range.

See Also: OPTION COUNT, OPTION PERCENT, AVERAGE, FREQUENCIES
Usage: The DIRECTORY command displays on the screen the Geographic Selection Files and Statistical Processor Command Files available in the user's directory.

Syntax: DIRECTORY

(D) -- Minimum abbreviation

Notes: Geographic Selection Files (extension SEL) are displayed first, followed by Statistical Processor Command Files (extension ECF).

Only the files in the default directory are shown (the directory from which REDATAM was started).

Example:

Selection Files
TEST.SEL   Test REDATAM Installation
CASE1.SEL   Sample case 1 for Miranda
CASE2.SEL   Sample case 2 for Miranda
CASE3.SEL   Sample case 3 for Miranda

Command Files
TEST.ECF   RUN NAME "Test REDATAM Installation"
CASE3.ECF   GEOGRAPHY CASE3

The label entered at creation time is shown for selection files. For command files, the first line is displayed. The file is easier to identify if the first line is a RUN NAME or COMMENT command.

See Also: RUN NAME, COMMENT
Usage: ENDHP is used with hierarchical files in order to mark the end of the hierarchical processing within the household.

Syntax: ENDHP

(EN) -- Minimum abbreviation

Notes: The ENDHP command is used in conjunction with the IF ... QUALIFY, IF ... COUNT and IF ... SUM commands. It must follow the last of these commands and precede any table generating commands (such as CROSSTABS, AVERAGE, or FREQUENCIES).

When an ENDHP is found, REDATAM retrieves the next record. If it belongs to the same household, it repeats all the commands that precede ENDHP. If the next record belongs to another household, REDATAM continues with the statements following ENDHP, processing all "qualified" records, until the end of the command set. It then re-initializes the COUNT and SUM variables and repeats the process for all population records of the next household unit.

The "qualified" population records that are available for tabulations or other statistical processes after the ENDHP are either:

a) The Head of household and only the Head, if no IF ... QUALIFY is used.

b) The population records that were accepted by an IF ... QUALIFY statement.

Note that a) is valid if, as should always be the case, the first record of each household is the Head.

Other selection commands, such as UNIVERSE or IF, may be used before or after the ENDHP command to further specify the population to be processed.

Example: A typical use of ENDHP in a hierarchical file would be:

GEOGRAPHY ...
UNIVERSE ...
IF ... / IF ... INCLUDE / IF ... EXCLUDE
RECODE ...
COMPUTE ...
IF ... COUNT ...
IF ... SUM ...
IF ... QUALIFY
ENDHP
IF ... / IF ... INCLUDE / IF ... EXCLUDE
RECODE ...
COMPUTE ...
FREQUENCIES ... / CROSSTABS ... / AVERAGE ...

All commands are processed once for each population record until ENDHP is encountered. The commands that follow, are executed once all population records of a particular household have been read and only for the records that have been "qualified" (or only the Head of household if no IF ... QUALIFY was used).

See SPECIAL FEATURES (Part 5) for more details about processing of hierarchical files.

See Also: IF ... QUALIFY, IF ... COUNT, IF ... SUM
Usage: This command exits the Statistical Processor and returns control to REDATAM's main menu.

Syntax: EXIT

(EX) -- Minimum abbreviation

Notes: The whole command set is lost, unless it was previously saved.

See Also: SAVE, NEW
Usage: The FREQUENCIES command creates one or more tables with the value distribution of the specified variables.

Syntax: FREQUENCIES <variable 1> [variable 2] [variable 3] ....

(F) -- Minimum abbreviation

Notes: For each of the specified variables, a table is generated, with total population counts for each of the values the variable takes. A line is printed for each value, showing the Absolute, Relative and Accumulative frequencies.

The variables specified must be in the Data Dictionary, or have been defined by the user, and if they are geographical variables, they must have been explicitly selected in the Geographic Selection, as indicated at the beginning of this section.

A table is generated for each variable.

Example: FREQUENCIES sex attsch

would produce the following two tables:

FREQUENCY FOR SEX Sex of Respondent
GEOGRAPHY IS GLOBAL

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>CODE</th>
<th>FREQUENCY</th>
<th>RELATIVE</th>
<th>ACCUMULATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FREQ</td>
<td>FREQ</td>
<td>FREQ</td>
<td></td>
</tr>
</tbody>
</table>

Male
1 453 57.0 57.0

Female
2 342 43.0 100.0

TOTAL 795 100.0

0 VALUES OUT OF RANGE
FREQUENCY FOR ATTSCH Attendance at School
GEOGRAPHY IS GLOBAL

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>CODE</th>
<th>FREQ</th>
<th>RELATIVE</th>
<th>ACCUMULATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fulltime Student</td>
<td>1</td>
<td>116</td>
<td>15.9</td>
<td>15.9</td>
</tr>
<tr>
<td>Parttime Student</td>
<td>2</td>
<td>537</td>
<td>73.7</td>
<td>89.6</td>
</tr>
<tr>
<td>Not Applicable</td>
<td>3</td>
<td>48</td>
<td>6.6</td>
<td>96.2</td>
</tr>
<tr>
<td>Not Stated</td>
<td>4</td>
<td>28</td>
<td>3.8</td>
<td>100.0</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td>729</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

0 VALUES OUT OF RANGE

The absolute and relative frequencies are totaled. The number of values out of range (as specified in the Data Dictionary) is also indicated.

See Also: OPTION COUNT, CROSSTABS, AVERAGE
Usage: The GEOGRAPHY command is used to indicate the geographic area of interest, by specifying the Geographic Selection File to be used for the run.

Syntax: GEOGRAPHY <file name>

(G) -- Minimum abbreviation

Notes: This command indicates to REDATAM which Geographic Selection File is to be used. The file must exist, having been previously created through the geographic selection process.

There can be only one active GEOGRAPHY command. If a second one is issued, it supersedes the first.

If the command is not issued, the last Geographic Selection File used in the geographic selection process is used.

<file name> is the name of an existing selection file, residing in the user's directory.

Example: GEOGRAPHY case4

uses the geographical area defined in file CASE4.SEL for all requested statistics. The suffix .SEL must not be written.
Usage: The IF command is used as a temporary filter to select records within a geography file.

Syntax: IF <condition>

(I) -- Minimum abbreviation

Notes: <condition> is an algebraic/logical expression whose result yields a true or false for a record. The commands following the IF are executed for all records for which the condition is true, until the next IF (or the end of the command set) is encountered.

The <condition> can include parenthesis, the basic four arithmetic operators and the logical operators, or their equivalents as listed below:

<, <=, =, >=, >, <, &&, ||
LT, LE, EQ, GE, GT, NE, AND, OR

The variables used in the condition must be in the Data Dictionary, or have been defined by the user, and if they are geographical variables, they must have been explicitly selected in the Geographic Selection, as indicated at the beginning of this section.

Example: IF sex = 1 AND (age <= 15 OR highschool < 3)

would consider only people who are male (value of SEX is 1) and who are either 15 or younger or who have not completed primary school (HIGHSCHOOL is less than 3).

An IF applies to all commands that follow until the next IF or until the end of the command set. When one or more IFs are used to select certain cases, for which statistical commands are entered, then another IF must normally be introduced later, before table producing commands, to eliminate the effect of the previous IF. For example:

COMMENT Males assigned Weight=3 and females Weight=4.
IF sex = 1
WEIGHT 3
IF sex = 2
WEIGHT 4
IF sex = 1 or sex = 2
CROSSTABS income BY sex

Leaving out the last IF would mean that the table would
be tabulated only for females (SEX = 2), since "IF sex = 2" continues to apply until changed.

Blank spaces between all arithmetical or logical operands are required.

See Also: IF ... INCLUDE, IF ... EXCLUDE, UNIVERSE
IF ... COUNT

Usage: The IF ... COUNT command is used in conjunction with hierarchical file processing, in order to count the occurrences of specific population records within a household. If used, it must be placed before the ENDHP command.

Syntax: IF <condition> COUNT <new variable>

(I) -- Minimum abbreviation

Notes: <condition> is an algebraic/logical expression whose result yields a true or false for a record. All population records, within each household, that meet the condition, are counted and the result stored in the indicated user variable. This variable can then be used like any other user-defined variable. Its value is reset to zero when a new housing record is retrieved.

The <condition> can include parenthesis, the basic four arithmetic operations and the logical operators, or their equivalents as listed below:

<, <=, =, >=, >, <, &&, ||
LT, LE, EQ, GE, GT, NE, AND, OR

The variables used in the condition must be in the Data Dictionary, or have been defined by the user, and if they are geographical variables, they must have been explicitly selected in the Geographic Selection, as indicated at the beginning of this section.

The name of the new variable may be up to eight characters in length and must start with a letter. It is treated as a housing variable.

If the range of the new variable is not defined via the VALUE LABEL command, REDATAM reserves space for the range 0 to 99.

Example: A typical use would be:

GEOGRAPHY ...
UNIVERSE ...
IF sex > 0 COUNT ntot
VAR LABEL ntot "Number of People in Household"
RECODE ntot (10 THRU HIGHEST = 10)
VALUE LABEL ntot 10 "10 or more"
IF age >= 15 QUALIFY
ENDHP
CROSSTABS ntot BY sex

The total number of people in each household is counted (SEX may be either 1 or 2, so the condition specified includes everyone). The cross tabulation will show the number of males and females who are 15 years or older, crossed with the total number of people in the household in which they reside.

Blank spaces between all arithmetical or logical operands are required.

See Also: IF ... SUM, IF ... QUALIFY, ENDHP, VAR LABEL, VALUE LABEL
Usage: The IF ... EXCLUDE command is used to permanently exclude a subset of records from processing within a geographic selection.

Syntax: IF <condition> EXCLUDE

(I) -- Minimum abbreviation

Notes: <condition> is an algebraic/logical expression whose result yields a true or false for a record. All records that meet the condition are discarded from further processing. Further IFs would add to the condition, which cannot be reversed.

The <condition> can include parenthesis, the basic four arithmetic operators and the logical operators, or their equivalents as listed below:

<, <=, =, >=, >, <>, &&, ||
LT, LE, EQ, GE, GT, NE, AND, OR

The variables used in the condition must be in the Data Dictionary, or have been defined by the user, and if they are geographical variables, they must have been explicitly selected in the Geographic Selection, as indicated at the beginning of this section.

Example: IF sex = 1 EXCLUDE

discards all people who are male (value of SEX is 1) for the remainder of the run. These records are no longer available, regardless of further IFs.

Blank spaces between all arithmetical or logical operands are required.

See Also: IF, IF ... INCLUDE, UNIVERSE
Usage: The IF ... INCLUDE command is used as a permanent filter to select a subset of records for processing within a geography file.

Syntax: IF <condition> INCLUDE

(I) -- Minimum abbreviation

Notes: <condition> is an algebraic/logical expression whose result yields a true or false for a record. Only records that meet the condition are processed and the rest are discarded from further processing. Further IFs would add to the condition, which cannot be reversed.

The <condition> can include parenthesis, the basic four arithmetic operators and the logical operators, or their equivalents as listed below:

<, <=, =, =>, >, <>, &&, ||
LT, LE, EQ, GE, GT, NE, AND, OR

The variables used in the condition must be in the Data Dictionary, or have been defined by the user, and if they are geographical variables, they must have been explicitly selected in the Geographic Selection, as indicated at the beginning of this section.

Note that when IF ... INCLUDE is used immediately after the GEOGRAPHY command, it has the same effect as the UNIVERSE command. However, the latter significantly speeds up the processing. The advantage of IF...INCLUDE is that it may be used with user-defined variables as well as several times within a given command set.

Example: IF sex = 1 INCLUDE

considers only people who are male (value of SEX is 1). All other records are discarded for the remainder of the run, so that a later IF such as

IF sex = 2

would select zero records, since all females were already excluded for the run.

Blank spaces between all arithmetical or logical operands are required.

See Also: IF, IF ... EXCLUDE, UNIVERSE
IF ... QUALIFY

Usage: The IF ... QUALIFY command is used in conjunction with hierarchical file processing in order to select specific population records within a household. If used, it must be placed before the ENDHP command.

Syntax: IF <condition> QUALIFY

(I) -- Minimum abbreviation

Notes: <condition> is an algebraic/logical expression whose result yields a true or a false for a record. All population records that meet the condition are selected and "qualified" for processing following the ENDHP command.

The <condition> can include parenthesis, the basic four arithmetic operators and the logical operators, or their equivalents as listed below:

<, <=, =, >=, >, <>, &&, ||
LT, LE, EQ, GE, GT, NE, AND, OR

The variables used in the condition must be in the Data Dictionary, or have been defined by the user, and if they are geographical variables, they must have been explicitly selected in the Geographic Selection, as indicated at the beginning of this section.

If an ENDHP command exists, but no IF ... QUALIFY is present, only the Head of Household or first record is "qualified". Multiple IF ... QUALIFY commands may be used.

An ordinary IF command prior to an IF ... QUALIFY does not alter the population "qualified". However, a UNIVERSE, IF ... INCLUDE, or IF ... EXCLUDE does.

Example: A typical use would be:

GEOGRAPHY ...
UNIVERSE ...
IF sex > 0 COUNT ntot
VAR LABEL ntot "Number of People in Household"
RECODE ntot (10 THRU HIGHEST = 10)
VALUE LABEL ntot 10 "10 or more"
IF age >= 15 QUALIFY
ENDHP
CROSSTABS ntot BY sex

For each household, all population records with AGE
greater or equal to 15 are "qualified". The cross tabulation shows the number of males and females who are 15 years or older, crossed with the total number of people in the household in which they reside.

Blank spaces between all arithmetical or logical operands are required.

See Also: IF ... COUNT, IF ... SUM, ENDHP
Usage: The IF ... SUM command is used in conjunction with hierarchical file processing, in order to add the value of a variable for specific population records. It must be used before the ENDHP command.

Syntax: IF <condition> SUM <new variable> = <existing variable>

(I) -- Minimum abbreviation

Notes: <condition> is an algebraic/logical expression whose result yields a true or false for a record. For all population records within each housing unit that meet the condition, the value of the indicated variable is summed to give a new variable. This new variable can then be used like any other user-defined variable. The value of <new variable> is reset to zero for each new housing record.

The <condition> can include parenthesis, the basic four arithmetic operators and the logical operators, or their equivalents as listed below:

<, <=, =, >, >>, &&, ||
LT, LE, EQ, GE, GT, NE, AND, OR

The variables used in the condition must be in the Data Dictionary, or have been defined by the user, and if they are geographical variables, they must have been explicitly selected in the Geographic Selection, as indicated at the beginning of this section.

The name of the new variable may be up to eight characters in length and must start with a letter. It is treated as a housing variable.

If the range of the new variable is not defined via the VALUE LABEL command, REDATAM reserves space for the range 0 to 99.

Example: A typical use would be:

GEOGRAPHY ...
UNIVERSE ...
IF sex > 0 COUNT ntot
VAR LABEL ntot "Number of People in Household"
RECODE ntot (10 THRU HIGHEST = 10)
VALUE LABEL ntot 10 "10 or more"
IF sex > 0 SUM tincome = income
IF age >= 15 QUALIFY
ENDHP
COMPUTE avinc = tincome / ntot / 1000
VAR LABEL avinc "Average Income: Thousands"
RECODE avinc (10 THRU HIGHEST = 10)
VALUE LABEL avinc 0 "0 - 999" 1 "1.000 - 1.999",
2 "2.000 - 2.999" 3 "3.000 - 3.999" 4 "4.000 - 4.999",
5 "5.000 - 5.999" 6 "6.000 - 6.999" 7 "7.000 - 7.999",
8 "8.000 - 8.999" 9 "9.000 - 9.000" 10 "10.000 OR MORE"
CROSSTABS avinc BY ntot

The total number of people in each household is counted and their incomes summed (SEX may be either 1 or 2, so the conditions specified include everyone). The cross tabulation will show the average income per person in relation to the number of people in the household.

Blank spaces between all arithmetical or logical operandi are required.

See Also: IF ... COUNT, IF ... QUALIFY, ENDHP, VAR LABEL, VALUE LABEL
Usage: This command displays on the screen all the commands in the current set.

Syntax: LIST

(LI) -- Minimum abbreviation

Notes: All commands entered (or if a NEW command was issued, those entered since then) are listed. Unlike the REVIEW command, LIST does not allow editing of the commands. Use the PRINT command to print the command set.

See Also: PRINT, REVIEW, NEW
Usage: The LOAD command retrieves from disk storage a previously SAVED command set, appending it to the current command set.

Syntax: LOAD <file name>
       (LO) -- Minimum abbreviation

Notes: When LOAD is invoked, the specified file is located on the disk, and if found, all the statements it contains are loaded, and appended after the last command of the current set. However, if LOAD is invoked from within the REVIEW command, the current command set is replaced.

The command file must be in the user's directory.

The commands are checked for syntax as they are loaded. They can be further modified with the REVIEW command.

See SPECIAL FEATURES (Part 5) for a tutorial on the use of editing commands.

Example: LOAD myfile

Loads the commands previously saved in file MYFILE.

See Also: SAVE, REVIEW, NEW
Usage: The NEW command eliminates the current command set from the computer's memory.

Syntax: NEW [Y]

(N) -- Minimum abbreviation

Notes: After invoking this command the Statistical Processor is restarted from scratch. Its effect is the same as EXITing the Statistical Processor and reentering it.

A new command set is started.

On entering NEW, REDATAM displays:

Confirm if you want to clear memory (Y/N):

If confirmed, all commands, including the current geography selection, are erased.

If the "Y" parameter is used, the memory is cleared without confirmation.

To clear the command set from the screen without erasing it from memory, use CLEAR.

Note that NEW has no effect on command sets previously saved on disk.

See Also: SAVE, EXIT, CLEAR
OPTION

**Usage:**
The OPTION command is used to define processing options for the WRITE command.

**Syntax:**
```
OPTION [FILENAME = <file name>]
[DOCUMENTATION = [DOS] [PRINT] [REDATAM] [SPSS] [SLMICRO]]
```

(O) -- Minimum abbreviation

**Notes:**
The OPTION command, used after a WRITE command, specifies the path and file name of the file being written to, as well as the output file format.

If OPTION FILENAME is not used after a WRITE, the file is named $xxxxxWRT.TMP$, where $xxxx$ corresponds to the name of the database. The file is always written to the user's directory. (See APPENDIX B). $<file name>$ must be four characters in length. No suffix specification is allowed.

The DOCUMENTATION option supports several format types, including REDATAM, SPSS-PC and SL-MICRO files. A DOS file format is assumed if this option is not used.

When a DOS file format is requested, the data is saved as a sequential ASCII file that can be seen with TYPE or similar commands under DOS.

The PRINT option generates a DOS file, but sends it to the printer.

For the SPSS and SLMICRO options, a command file is also generated. The user can later modify it to add his or her own commands and execute directly with the data obtained from REDATAM. The command file name is the name specified in the OPTION command plus the '.CMD' extension.

The REDATAM option generates a Data Dictionary and all necessary files to proceed to create a new REDATAM database with the Database Management Functions.

See the WRITE command description, and SPECIAL FEATURES (Part 5) for more details on the use of this command.

**Example:**
When used after a WRITE command to generate an SPSS-PC file:

```
WRITE water sex
```
OPTION DOCUMENTATION = SPSS FILENAME = spss

produces a file named SPSS. with the requested data, and another file named SPSS.CMD containing the following:

TITLE CASE1.SEL, mi80, Tue Jan 06 1987.
DATA LIST FILE = 'SPSS' FIXED
/WATER 1-1 SEX 2-2.
VARIABLE LABELS WATER 'Water Supply'/
SEX 'Sex of Respondent'.
VALUE LABELS WATER
   1 'Pub. Piped Dwell'
   2 'Pub. Piped Yard'
   3 'Prv. Piped Dwell'
   4 'Prv. Catchment'
   5 'Pub. Stand Pipe'
   6 'Pub. Tank'
   7 'Other'
   8 'Not Stated'/
SEX
   1 'Male'
   2 'Female'.

* Insert your SPSS commands here.
FINISH.

When used after a WRITE command to generate an SL-MICRO file:

WRITE water sex
OPTION DOCUMENTATION = SLMICRO FILENAME = slmi

Writes a file named SLMI. with the requested data, and another file named SLMI.CMD containing the following:

RUN NAME CASE1.SEL, mi80, Tue Jan 06. 1987
VARIABLE LIST WATER,SEX
INPUT FORMAT FIXED(F1.0,F1.0)
DATA FILE SLMI
VAR LABELS WATER,Water Supply/
SEX,Sex of Respondent/
VALUE LABELS WATER
   (1) Pub. Piped Dwell
   (2) Pub. Piped Yard
   (3) Prv. Piped Dwell
   (4) Prv. Catchment
   (5) Pub. Stand Pipe
   (6) Pub. Tank
   (7) Other
   (8) Not Stated /
SEX
   (1) Male
   (2) Female
The user can then modify the file and add the other commands required to run SPSS-PC or SL-MICRO. The TITLE (in SPSS-PC) or RUN NAME (in SL-MICRO) commands contain the name of the Geographic Selection File and the REDATAM database, the date and time.

See Also: WRITE
OPTION COUNT

Usage: The OPTION COUNT command is used to alter the level at which the cells of a table are tallied, that is, at the household or population level. It must immediately follow a CROSSTABS, FREQUENCIES or AVERAGE command.

Syntax: OPTION COUNT <level number>

where <level number> is: 1 for households
2 for persons

(O COUNT) -- Minimum abbreviation

Notes: OPTION COUNT may be used immediately after a table generating command to indicate the record type that is to be tabulated. Normally, REDATAM applies the following rules:

a) If only housing variables are used in the table: REDATAM counts at the household level.

b) If at least one population variable is used in the table: REDATAM counts at the population level.

OPTION COUNT permits to override these defaults by specifying that records are to be counted at either the housing level (level 1) or population level (level 2).

Example: FREQUENCIES water
OPTION COUNT 2

counts the number of people, rather than households that have each type of water facilities.

FREQUENCIES sex
OPTION COUNT 1

classifies the number of households by the sex of the first population record, probably the head. If OPTION COUNT 1 had not been used, the total number of males and females within the selected area would have been shown.

See SPECIAL FEATURES (Part 5) for details on the use of OPTION COUNT and other aspects of hierarchical processing.

The CROSSTABS command may be immediately followed by an OPTION command with COUNT and PERCENT options.

See Also: AVERAGE, CROSSTABS, FREQUENCIES, OPTION PERCENT
Usage: The OPTION PERCENT command is used to request percentages for the CROSSTABS command.

Syntax: OPTION PERCENT [ROW] [COLUMN] [TOTAL]

(O PERCENT) -- Minimum abbreviation

Notes: When used after a CROSSTABS command, OPTION PERCENT can include row-wise, column-wise and/or total percentages of the individual cell values. Any combination of the three options can be included.

Example: CROSSTABS sex BY occstat
OPTION PERCENT ROW COLUMN TOTAL

would generate the following table:

<table>
<thead>
<tr>
<th>SEX OF RESPONDENT</th>
<th>OCCSTAT</th>
<th>GEOGRAPHY IS GLOBAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex of Respondent</td>
<td>Occupational Status</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CODE</td>
<td>Governme</td>
<td>Private</td>
</tr>
<tr>
<td>1</td>
<td>336</td>
<td>109</td>
</tr>
<tr>
<td>Row Percent</td>
<td>74.2%</td>
<td>24.1%</td>
</tr>
<tr>
<td>Column Percent</td>
<td>58.0%</td>
<td>56.2%</td>
</tr>
<tr>
<td>Total Percent</td>
<td>42.3%</td>
<td>13.7%</td>
</tr>
<tr>
<td>Female</td>
<td>243</td>
<td>85</td>
</tr>
<tr>
<td>Row Percent</td>
<td>71.1%</td>
<td>24.9%</td>
</tr>
<tr>
<td>Column Percent</td>
<td>42.0%</td>
<td>43.8%</td>
</tr>
<tr>
<td>Total Percent</td>
<td>30.6%</td>
<td>10.7%</td>
</tr>
</tbody>
</table>

The percentages requested (all in this case) are printed together with the cross-tabulation figures for each cell, and are totaled as well.

The CROSSTABS command may be immediately followed by OPTION command with COUNT and PERCENT options.

See Also: CROSSTABS, OPTION COUNT
PRINT

Usage: The PRINT command is used to list the current command set on the printer.

Syntax: PRINT

(PRI) -- Minimum abbreviation

Notes: This command is the same as the LIST command, except that the commands are printed rather than listed on the screen.

See Also: LIST
The PROCESS command initiates the execution of the commands in the current set.

**Syntax:**

`PROCESS [OUTPUT = PRINTER,[NOFF]]`

`PROCESS [OUTPUT = <file name>,[NOFF]]`

(Pro) -- Minimum abbreviation

**Notes:**

PROCESS processes all current commands, generates the requested output, and returns control to the Statistical Processor.

The options are used to route the output to the printer, or the file whose name is specified.

If no options are supplied, the results and input commands are shown on the screen, where the user can browse through them. Instructions are displayed.

When the output is sent to the screen, it is saved in a file named `REDOUTPU.TMP`, in the user's directory. The file is available for browsing through the use of the VIEW command until another PROCESS command overwrites it. `REDOUTPU.TMP` may also be loaded into a resident editor, like Sidekick, for editing or printing without leaving REDATAM. See the VIEW command for further details on this option.

As the data file is read, the number of records processed, is displayed on the upper right hand corner of the screen in 10 household record increments.

PROCESS can be interrupted by pressing the ESC key.

The NOFF option will omit all form feed characters from the output listing.

**Example:**

`PROCESS OUTPUT = PRINTER`

starts the actual processing of the previously entered commands and sends the output to the printer.

**See Also:** VIEW
RECODE

Usage: This command permits the creation of temporary variables as grouped ranges of existing ones. These are then treated like any other variable.

Syntax: RECODE <existing variable> TO <new var> <recode list>

RECODE <existing user variable> <recode list>

where <recode list> may take any of the following forms:

- \((<\text{exist. val 1}> [\text{THRU} <\text{exist. val 2}>] = <\text{new val 1}>)\)
- \([(\text{<exist. val 3}> [\text{THRU} <\text{exist. val 4}>] = <\text{new val 2}>)]\)
- \([...]\) ...

- \((<\text{exist. val 1}> [- <\text{exist. val 2}>] = <\text{new val 1}>)\)
- \([(\text{<exist. val 3}> [- <\text{exist. val 4}>] = <\text{new val 2}>)]\)
- \([...]\) ...

- \((<\text{exist. val 1}> [, <\text{exist. val 2}>] = <\text{new val 1}>)\)
- \([(\text{<exist. val 3}> [, <\text{exist. val 4}>] = <\text{new val 2}>)]\)
- \([...]\) ...

(REC) -- Minimum abbreviation

Notes: By using RECODE, a new variable may be defined (first syntax form), its values corresponding to groups of values of another. The re-coded variable can then be used as if it were another variable of the data file. The second form is only permitted with variables that were previously created by the user.

The values of the existing variable must be in ascending order. Therefore to recode 3, 4 and 6 into 2, and 5 into 1, you must use \((3 - 4 = 2) \ (5 = 1) \ (6 = 2)\).

In both variations of the RECODE command, values that are not explicitly specified in the command remain unchanged.

The variables specified in the recode list must be in the Data Dictionary, or have been defined by the user, and if they are geographical variables, they must have been explicitly selected in the Geographic Selection, as indicated at the beginning of this section.

The new variable name can be up to eight characters in length and must start with a letter. It must not have been previously created.
By using the VAR LABEL and VALUE LABEL commands, descriptions for the new variable and its values can be provided, as well as a specification of their range (if no range is specified for new variables, REDATAM assumes 0 to 99).

Blank spaces between operandi and variables are required.

In hierarchical files, variables are assumed to be of the same type (housing or population) as that of the existing variable.

Example: RECODE age TO agegroup (LOWEST THRU 4 = 1) (5 THRU 9 =, 2) (10 THRU 14 = 3) (15 THRU 19 = 4) (20 THRU HIGHEST, = 5)

creates a variable called AGEGROUP that represents age groups. Note the use of "LOWEST" and "HIGHEST"; they can be used to respectively include all values below or above a given value.

Note also that the command has been broken down into continuation lines with the use of a comma (,).

The command combination:

COMPUTE agegroup = (age / 5) + 1
RECODE agegroup (5 THRU HIGHEST = 5)

has the same effect as the previous example. The COMPUTE reassigned ages in groups of five (fractional parts are truncated). With the RECODE, all values higher than 5, that is belonging to ages 20 and up are reassigned to code 5.

See Also: VAR LABEL, VALUE LABEL, COMPUTE
Usage: The REVIEW command allows the user to modify, delete and insert new commands in the current set.

Syntax: REVIEW

(REV) -- Minimum abbreviation

Notes: This command acts as an internal editor that can be used to make changes on a line by line basis. Each command is rechecked for syntax and context after leaving the editor.

External editors, such as Sidekick, may also be used, but of course, they will not check for syntax, which may have to be corrected when LOADING the file. To use an external editor, see SPECIAL FEATURES (Part 5), for the internal format of the command set.

The following keys may be used while editing:
The arrow keys move up, down or along the commands.

F1 - Inserts a blank line before the current command, where a new command can be entered.
F2 - Deletes the current line.
F3 - Collects a block of commands to be saved. Pressing it once activates collect mode. The cursor can then be moved down until the desired lines are highlighted. Pressing F3 again returns to edit mode while the lines remain highlighted. When F5 is pressed, that block of commands is saved. Pressing F3 twice returns all lines to normal.
F4 - Includes a file of commands immediately above the line where the cursor is currently at. REDATAM asks for the name of the command file to include.
F5 - Saves the entire set of commands. If a group of commands has been collected (with F3), only the commands in the highlighted block are saved. REDATAM asks for a file name to save to.
F6 - Loads an existing command file, replacing the current one.
F7 - Invokes BROWSE, with information on Data Dictionary variables.
F8 - Calls the Help facility, with information on each command.
F9 - Quits REVIEW, ignoring any changes made.
F10 - Exits REVIEW, returns to the Statistical Processor, and checks the commands for syntax.

See SPECIAL FEATURES (Part 5) for a tutorial on the use of the REVIEW editing commands.

See Also: BROWSE, LIST, LOAD, SAVE
RUN NAME

Usage: RUN NAME can be used to specify a label or title for the output.

Syntax: RUN NAME "[text]"

(RU) -- Minimum abbreviation

Notes: The "text" specified is printed at the top of each page of output along with the date and time of the process.

There can be only one active RUN NAME command. If a second one is issued, it supersedes the first.

It is recommended that this command be used to assist in later identifying printed output pages. Also, if it is placed on the first line of the command set, it helps in identifying its contents when using the DIRECTORY command.

Example: RUN NAME "Case 5 of Adult Education. Geography 1."

See Also: COMMENT, DIRECTORY
SAVE instructs the system to write to disk all the commands in use.

Syntax:
SAVE <file name>

(S) -- Minimum abbreviation

Notes:
The entire active command set is saved, that is, all commands entered since entering the Statistical Processor or the last NEW command.

If the file already exists on disk, REDATAM answers with:

File exists: Override (Y/N):

The user is asked to confirm, with a "y" for yes or a "n" for no, that the existing file is to be overwritten. If "n" is answered, the command is canceled.

The entire set of commands is lost if the user does not SAVE before EXITing the Statistical Processor.

See SPECIAL FEATURES (Part 5) for a tutorial on the use of this and other editing commands.

File names must be valid DOS (Disk Operating System) file names. REDATAM adds the suffix .ECF, so there is no danger that Command Files be confused with Geographic Selection Files, which carry the .SEL suffix.

Files are saved in the current user directory.

Example: SAVE myfile

saves the current command set in file MYFILE.ECF.

See Also: LOAD, REVIEW, EXIT, NEW
UNIVERSE

Usage: The UNIVERSE command filters out from the totality of records in the selected area, those records that do not meet a specified condition. This command affects all commands in the set.

Syntax: UNIVERSE <condition>

(U) -- Minimum abbreviation

Notes: <condition> is an algebraic/logical expression whose result yields a true or false. The UNIVERSE command uses the same syntax and has the same effect as the IF ... INCLUDE command, but unlike it, it must be issued before any other commands except GEOGRAPHY and AREABREAK. It can therefore not be used with user created variables. If geographical variables are used, they must have been explicitly selected in the Geographic Selection, as indicated at the beginning of this section.

If IF commands are issued in addition to the UNIVERSE command, their effect is added to that of the UNIVERSE command.

There can be only one UNIVERSE command for each command set, and it must be used right after the GEOGRAPHY and AREABREAK commands.

The UNIVERSE command executes significantly faster than the IF ... INCLUDE command, so it should be used whenever possible, when record selection is required.

The <condition> can include parenthesis, the basic four arithmetic operators and the logical operators, or their equivalents listed below:

<, <=, =, >=, >, <>, &&, ||
LT, LE, EQ, GE, GT, NE, AND, OR

Blank spaces between arithmetical and logical operands are required.

Example: UNIVERSE sex = 1 AND age >= 25

applies all the commands in the current set, only to males (SEX = 1) whose age is greater or equal to 25.

See Also: IF, IF ... INCLUDE, IF ... EXCLUDE
Usage: VALUE LABEL is used to specify descriptive labels for the values of user-created variables. It is also used to define the range of values these variables can take.

Syntax: VALUE LABEL <variable> <value 1> "<description>"
        [value 2] "[description]"
        [...] ...

(VAL LABEL) -- Minimum abbreviation

Notes: VALUE LABEL associates descriptions to the values of variables created with the COMPUTE, RECODE, IF ... COUNT and IF ... SUM commands. These descriptions are used in the output tables of the AVERAGE, CROSSTABS and FREQUENCIES commands.

As the existence of value labels does not affect the data printed in a table, it is not necessary to label any or all values of a variable. However its use is recommended both for ease of reading the output tables, and to define the range of values, as explained below.

When VALUE LABEL is used, REDATAM reserves space for all values between the lowest and highest value specified in the command. Any range of values between 0 and 32656 may be assigned. If VALUE LABEL is not used, the variable has the default range of (0, 99). Only the minimum and maximum values need to be specified with VALUE LABEL in order to define a range.

The range of each new variable should be properly specified, since it is used to calculate the space reserved for each table. If the default values are used, REDATAM may run out of memory during execution.

Note that value labels longer than one line, can be entered either by repeating the command, or by indicating continuation lines with a comma.

Blank spaces between operandi are required as shown.

Example: VALUE LABEL agegr 1 "4 or less" 2 "5 to 9" 3 "10 to 14"
          VALUE LABEL agegr 4 "15 to 24" 5 "25 or more"

or

VALUE LABEL agegr 1 "4 or less" 2 "5 to 9" 3 "10 to, 14" 4 "15 to 24" 5 "25 or more"

See Also: VAR LABEL, COMPUTE, RECODE, IF ... COUNT, IF ... SUM
VAR LABEL

Usage: VAR LABEL is used to provide descriptive labels for user-created variables.

Syntax: VAR LABEL <variable 1> "<description>"
          [variable 2] "[description]"
          [...] ...

(VAR LABEL) -- Minimum abbreviation

Notes: VAR LABEL associates descriptions to the variables created with the COMPUTE, RECODE, IF ... COUNT and IF ...
 ... SUM commands. These descriptions are used in printing the results of the AVERAGE, CROSSTABS and FREQUENCIES commands.

Blank spaces between operandi are required as shown.

Example: VAR LABEL agegroup "Age Group of Respondent"
         VAR LABEL average "Average Number of Children"

associates descriptions to the AGEGROUP and AVERAGE variables.

See Also: VALUE LABEL, COMPUTE, RECODE, IF ... COUNT, IF ... SUM
Usage: This command allows the user to re-display the last output sent to the screen.

Syntax: VIEW

(VI) -- Minimum abbreviation

Notes: VIEW shows on the screen the output of the last PROCESS command sent to the screen and saved in the file REDOUTPU.TMP. The user can scroll through the output.

Instructions are shown on the bottom line. The following keys may be used:

The Up and Down arrows scroll the output one line up or down respectively.

The left and right arrows move the output one screen column to the left or right respectively.

Home moves to the top of the output.

End moves to the bottom of the output.

PgUp moves up one half screen of output.

PgDn moves down one half screen of output.

Ctrl-Left arrow pans the screen one output column to the left.

Ctrl-Right arrow pans the screen one output column to the right.

Esc exits VIEW.

The REDOUTPU.TMP file may be edited by an external editor, such as SIDEKICK, in order to make changes for printing, for example.
Example: VIEW
invoked after a PROCESS command may bring up the following screen:

REDATAM System Output  Data base: M180  Date: 22/12/1986
REDATAM System Mon Dec 22 15:26:15 1986
1> GEOGRAPHY CASE1
2> IF SEX = 1
3> FREQUENCIES ATTACH HIGHSCHT

REDATAM Statistics Mon Dec 22 15:27:48 1986
FREQUENCY FOR ATTACH Attendance at School
GEOGRAPHY IS GLOBAL

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>CODE</th>
<th>FREQ</th>
<th>RELATIVE</th>
<th>ACCUMULATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fulltime Student</td>
<td>1</td>
<td>14</td>
<td>15.4</td>
<td>15.4</td>
</tr>
<tr>
<td>Parttime Student</td>
<td>2</td>
<td>71</td>
<td>78.0</td>
<td>93.4</td>
</tr>
<tr>
<td>Not Applicable</td>
<td>3</td>
<td>2</td>
<td>2.2</td>
<td>95.6</td>
</tr>
<tr>
<td>Not Stated</td>
<td>4</td>
<td>4</td>
<td>4.4</td>
<td>100.0</td>
</tr>
</tbody>
</table>

<^[move scr Home top End bot PgUp up PgDn down Ctrl< left Ctrl> right ESC quit

See Also: PROCESS
Usage: WEIGHT sets the tabulation weight factor to any value derived by a computation.

Syntax: WEIGHT <algebraic expression>

(WE) -- Minimum abbreviation

Notes: The factor computed from the algebraic expression or literal value, is applied to each record as it is processed. Variable weights may be assigned by including variables as parts of the expression. The four basic arithmetic operators (+, -, *, /) plus exponentiation (^) and parenthesis may be used. Expressions are evaluated according to the rules of algebra.

Integer results are used. Decimals resulting from a calculation are discarded.

This command would normally be used if a sample census record represented a number of people.

The WEIGHT function affects only the AVERAGE, CROSSTABS and FREQUENCIES commands to produce weighted tallies. It has no effect on COMPUTE or other commands.

The factor set by the WEIGHT command remains in effect until specifically changed by another WEIGHT command. Care must thus be taken to reset the weight to its default value (1) if needed.

The default value, that is, if WEIGHT is not used, is 1.

The variables used in the algebraic expression must be in the Data Dictionary, or have been defined by the user, and if they are geographical variables, they must have been explicitly selected in the Geographic Selection, as indicated at the beginning of this section.

Blank spaces between arithmetical operands are required.

Example: WEIGHT numpeopl

multiplies each record by the number of people declared in variable numpeopl.

See Also: COMPUTE
Usage: The WRITE command is used to create a subset of the REDATAM data file for further analysis inside or outside the system.

Syntax: WRITE [record #] <variable 1> [((Fn.m))]
[variable 2] [((Fn.m))]
[...]

WRITE <variable 1> [variable 2] [...] ...
[/] [variable 3] [variable 4] [...] ...

(WR) -- Minimum abbreviation

Notes: A file is written whose fields are the specified variables, which may belong to the database or be defined by the user. All records selected by the Geographic Selection, and record selection commands are written to the file.

The WRITE command is used in conjunction with the OPTION command. The first form of WRITE is used when OPTION DOCUMENTATION specifies either DOS, PRINT, SPSS or SILMICRO formats. The second form is used with OPTION DOCUMENTATION = REDATAM.

[record #] is an optional parameter used to specify the record type to which this variable belongs. A WRITE has to be issued for each record type if more that one is used. It must not be used when generating a REDATAM database.

[[(Fn.m))] can be used to specify the format in which the variable is to be written. "n" indicates the number of characters to be reserved for the variable, including the decimal point. "m" indicates the number of decimal places. If this parameter is not used, the Data Dictionary format is used, and if the variable was created by the user, 5 places are reserved (no decimals). It cannot be used when generating a REDATAM database.

[/] may be used in hierarchical files when generating a REDATAM database, in order to separate housing from population variables.

WRITE commands may be issued in sequence, that is, several WRITE commands can be used within the command set (with different record selections and/or record types), writing data to different files. The OPTION command must follow each WRITE command. However, only
one WRITE command is allowed for OPTION DOCUMENTATION = REDATAM.

See SPECIAL FEATURES (Part 5) for more information on the use of this command when generating a REDATAM database.

Example: WRITE 1 tenure
OPTION FILENAME = tst1 DOCUMENTATION = DOS
WRITE 2 age (F2.0) agepct (F4.1)
OPTION FILENAME = tst1 DOCUMENTATION = DOS

creates a DOS file named TST1 in the work directory. Two types of records are written. The first record type contains the TENURE variable in the format specified in the Data Dictionary. The second record type contains the AGE variable, written as two characters with no decimals, and the user created AGEPCT variable, of length 4 with one decimal.

See the OPTION command for more examples on the use of the WRITE command, and Part 5 for an explanation on its use in order to generate subset REDATAM databases.

See Also: OPTION
4.2. **Geographic Selection**

This section contains a description of the options and menus available with the Geographic Selection Processor.

REDATAM applies its statistics to one or more specific geographical areas selected by the user. Countries are assumed to have a hierarchical geographic division, that is, they have major divisions like provinces, which are in turn subdivided into smaller areas, like departments, and so on.

In order to use the data effectively, the analyst must be able to select and group different areas of the country in order to study them individually. The Geographic Selection Processor allows the user to indicate the areas of the country he or she wants selected, and optionally specify groupings thereof. REDATAM stores all the information in a Geographic Selection File, which is used by the Statistical Processor, when statistics are actually generated.

The Geographic Selection Processor is a menu driven system, that is, the user is presented with screens offering a series of choices. All screens are self explanatory, and further information and help is often available by pressing the ? key.

In this section, each screen and option is explained. Examples of screens, based on the sample geography included at the end of the TUTORIAL (Section 3.6), are shown. A list of available options with a short description is also provided.
Geographic Selection Option List

LOAD SELECTION FILE
Selects, creates and updates Geographic Selection Files.

GEOGRAPHY DEFINITION MENUS
Series of menus used to select and group geographical areas.
LOAD SELECTION FILE

The following screen is shown (the file names, may of course be different):

REDATAM V2.0 Load or Create File               Data Base: M150 date: 25/03/1987

Name       Date     Time     Contents

CREATE FILE

CASE1.SEL  22/12/86  11:39:13  Sample case 1 for Miranda
CASE2.SEL  22/12/86  11:36:10  Sample case 2 for Miranda
CASE3.SEL  16/12/86  16:52:21  Sample case 3 for Miranda
GEOG15.SEL 16/12/86  09:28:00  STUDY NO 15
DEMO.SEL   16/12/86  09:27:17  DEMONTRATION

CUR Up & CUR Down ENTER Sel File Del Del File ALT-P Print Sel ESC Abandon

A list of available Geographic Selection Files is shown, together with the date and time of creation or last change and a short description of each file. Existing files can be selected and new ones created. They can also be erased or printed. Instructions are shown at the bottom of the screen.

Options:

CREATE FILE:
This option is used to create a new Geographic Selection File.

Select File:
This option selects and loads an existing Geographic Selection File, and goes to the GEOGRAPHY DEFINITION MENUS.

Delete File (DEL):
This option deletes the file highlighted by the cursor.

Print File (Alt-P):
This option sends to the printer a copy of the currently selected geography file. It is presented as copies of the geography definition menus for the selected areas.

Usage:
A file is selected by using the up and down arrows until it is highlighted. ENTER loads the file, or any of the other options shown on the screen may be used. The screen scrolls when the cursor reaches the bottom and more files exist.
If an existing file is selected, it is loaded, and REDATAM goes to the GEOGRAPHY DEFINITION MENUS.

If the CREATE FILE option is selected, the following prompt appears on the screen:

CREATE FILE      Name: __________

The cursor is positioned at the beginning of the dotted line that follows "Name:". Any valid DOS file name can be entered. REDATAM adds a .SEL suffix to it. ENTER finishes the entry, and writes the following line at the bottom of the screen:

Label: ___________________________________________________________________

A short text description (up to 60 characters) may be entered. This description is shown next to the file to assist in identifying it. ENTER completes the creation process, loads the file and goes to the GEOGRAPHY DEFINITION MENUS.

When typing both the file name and the label, the right and left arrows, the DEL and Backspace keys can be used to edit the text before ENTER is pressed.

Pressing ESC cancels the file selection and returns to the MAIN MENU.

Selection files are saved in the current user directory.
These are a series of screens, each one showing a particular geographic level. Selections can then be made. Facilities for "navigating" between the different screens are provided. When first entering from the LOAD SELECTION FILE, the highest geographical level is shown.

REDATAM V2.0 Geography Selection
PROVINCE TOWN ED

File: CASE1.SEL

P Partial Selection T Total Selection ESC Abandon ? Help

The heading of the screen shows the name of the active database, the date, the current Geographic Level (shown underlined) and the name of the Geographic Selection File. The bottom of the screen shows instructions. A Help facility, accessible with the ? key is available.

All geographic subdivisions available at that level are shown, two in this case.

The hierarchical geographic division assumed for this example is shown in the following diagram:
Level 0:
  Country
    --------------
    MIRANDA

Level 1:
  Province
    ---------------
    FLORINDA    PERDITA

Level 2:
  Town
    ---------------
    VILCHES      RURAL    TUMACO    RURAL

Level 3:
  ED
    --- 84 | 85 | 89 | 90 | 97 | 98 | 101 | 102 ---
    --- 86 | 87 | 91 | 92 | 99 | 100 | 103 ---
    --- 88 |   | 93 | 94 |   |   |   |   ---
    ---   |   | 95 | 96 |   |   |   |   ---

If the geographic level is further down the hierarchy, all preceding levels are also shown. Lower levels are shown indented, as in the following screen (footnotes omitted):

REDATAM V2.0 Geography Selection
PROVINCE TOWN ED
Code Area Name Sel Recode
01 Florinda
  99 Rural
    089
    090
    091
    092
    093
    094
    095
    096

Options: Select a Geographic Area:
Areas can be selected in their totality or partially. If a partial selection is specified, the selection process has to continue to the next lower geographical level.

Re-code a Geographic Area:
Geographic areas may be re-coded, so that they can be grouped in ways other than the hierarchical structure of the geographic subdivisions. For instance, adjacent districts may be "joined" with another province to define a new geographical unit.

Move Between Geographical Levels:
The user can move "up and down" and "sideways" between the different geographical levels.

Save Geographic Selections:
When selections are done, they are saved in the current Geographic Selection File, with the .SEL suffix. REDATAM checks for consistency and completeness of the selection before saving.

Usage:
When entering a GEOGRAPHY DEFINITION MENU, the cursor is positioned at the first geographic area under the Sel column. The Sel column is used to indicate the type of selection for the corresponding geographical area. Three choices are available:

The column can be left blank: This means that this particular area is excluded from the geographic selection.

T can be entered: This indicates a Total selection of the area, meaning that all the sub-areas it contains are to be included in the Geographic Selection.

P can be entered: This indicates a Partial selection, meaning that just some of the sub-areas are to be selected, and that further specification is required at a lower level.

The cursor can be moved along the Sel column with the up and down arrows.

A T or a P should be entered into every area that in whole or in part, respectively, needs to be included in the Geographic Selection. They can be entered either in upper or lower case.

When a T is entered, the cursor automatically skips to the Recode column. The Recode column is used if the
geographic areas need to be grouped in a way other than that afforded by the hierarchical structure of the file. The Recode option is used in conjunction with the AREABREAK RECODE command of the Statistical Processor. By entering a label under that column, all areas with the same label are grouped in a tabulation or other statistical process, if the AREABREAK RECODE command is issued.

A Recode label may be up to six characters in length (no differentiation is made for upper and lower case). The entry is completed with ENTER. If no label is desired, it is left blank. A completed selection screen (footnotes omitted), may look like this:

```
REDATAM V2.0 Geography Selection
Data Base: M180 Date: 25/3/1987
PROVINCE TOWN ED
File: CASE1.SEL

<table>
<thead>
<tr>
<th>Code</th>
<th>Area Name</th>
<th>Sel</th>
<th>Recode</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Florinda</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>02</td>
<td>Perdita</td>
<td>T</td>
<td>AREA-1</td>
</tr>
</tbody>
</table>
```

This screen specifies a Partial selection on the province of Florinda and a Total selection on Perdita. Furthermore, Perdita has been Recoded with the "AREA-1" label. In order to complete the selection process, areas from the lower levels of Florinda (town and possibly ED) have to be selected, since only a partial selection has been requested here. All subdivisions of Florinda that are Recoded with "AREA-1", will be grouped together with all of the province of Perdita.

By accessing the help facility, information on the use of the keyboard is displayed. When pressing the ? key (the SHIFT key together with the ? key), the following "window" is superimposed on the current screen:

```
<table>
<thead>
<tr>
<th>DESCRIPTION</th>
<th>COMMAND</th>
<th>DESCRIPTION</th>
<th>COMMAND</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partial Selection</td>
<td>p</td>
<td>Next Selection</td>
<td>v</td>
</tr>
<tr>
<td>Total Selection</td>
<td>t</td>
<td>Previous Selection</td>
<td>^</td>
</tr>
<tr>
<td>Next Upper Level</td>
<td>SHIFT+</td>
<td>Enter Selection</td>
<td>HOME</td>
</tr>
<tr>
<td>Next Lower Level</td>
<td>SHIFT-</td>
<td>End Selection</td>
<td>END</td>
</tr>
<tr>
<td>Previous Node</td>
<td>&lt;</td>
<td>Cancel Selection</td>
<td>ESC</td>
</tr>
<tr>
<td>Next Node</td>
<td>&gt;</td>
<td>Help</td>
<td>?</td>
</tr>
</tbody>
</table>

Press <ESC>|
```
No work may be done while the Help window is active. Pressing ESC clears it from the screen in order to resume work on the current selection.

The Help window shows how to move about the geographic levels:

SHIFT - up arrow (pressing the up arrow while holding down either of the SHIFT keys), displays the next hierarchically higher level.

SHIFT - down arrow displays the next hierarchically lower level.

left arrow moves to the next geographic area "to the left" at the same level.

right arrow moves to the next geographic area "to the right" at the same level.

Refer to the hierarchical diagram of Miranda shown above for clarification of these concepts.

If any of the requested moves is not possible (because there are no more geographic areas in the direction requested) a message indicating this appears on the screen.

HOME checks if there are any "pending" selections, that is, it follows all partial selections down to the lower levels to verify that they have been ended with a total selection. If pending selections are found, the incomplete screen is displayed for the user to finish. If no pending selections are found, the user is returned to the highest geographical level. Given this characteristic of the HOME key, it can be used instead of the arrow keys when defining a new geography to successively move to the different levels requiring completion. The geographic selection must always be completed with HOME in order to verify that all pending selections have been completed, before pressing END.

END terminates the selection process, saves the selection information in the Geographic Selection File and returns to the GEOGRAPHIC SELECTION MENU.

ESC cancels all selections after asking for confirmation from the user (y for "yes" and n for "no"). REDATAM returns to the MAIN MENU. The existing Geographic Selection File remains unchanged.
4.3. **Data Dictionary**:

The Data Dictionary is an integral part of the REDATAM system. It is a computer readable copy of the data file documentation, containing information on what variables are available in the data file, describing what those variables mean, what categories are available within each of the variables and providing other technical information about the variables that the REDATAM system needs to access and prepare tabulations.

For most users of the REDATAM system, the Data Dictionary describing the file or files that will be used will have already been prepared by the computer technical support staff within the user's organization, and will be available as part of the REDATAM database that is in use.

The main interest that most users will have in the Data Dictionary, is to review its contents and print a hard-copy for their individual use.

The TUTORIAL shows how to print a copy of the Data Dictionary. APPENDIX D contains a complete reference to all Data Dictionary functions, including the use of the browse function and the creation and modification of Data Dictionaries.
4.4. Database Management Functions:

The *Database* creation and loading will be done by programmers. This is because some parts of the load process are done on a large mainframe computer. The data is then transmitted via a communications link to the microcomputer. Once the data is in the microcomputer, the REDATAM system provides the necessary facilities to transform the transmitted files into a REDATAM Data Base. The processes involved are described in APPENDIX D.
4.5. Keyboard Usage:

REDATAM has been designed to make optimum use of the PC's keyboard. Many often used operations can be accomplished with just a few, sometimes one, keystrokes. For this, special keys are used.

This section provides a reference to the use of the keyboard in REDATAM. While every effort has been made to keep the function of each key as consistent and as intuitive as possible, its use may vary somewhat between each of the environments of REDATAM in order to accommodate different requirements. These differences are explained in this section, which is divided into the following parts:

GENERAL KEYBOARD USE AND CONVENTIONS
   Provides some general comments on the use of the keyboard in all environments, and explains naming conventions.

KEYBOARD USAGE IN MENU SCREENS
   Explains how to use the keyboard when choosing options from a menu screen.

CHARACTER AND SCREEN EDITING KEYS
   Explains all the keys that can be used when text needs to be entered or modified.

KEYBOARD USE WITH THE HELP FACILITY
   Shows how the keyboard is used when help (?) is invoked.

KEYBOARD USE WITH THE STATISTICAL PROCESSOR
   Presents the special keys used when under control of the Statistical Command Processor.

KEYBOARD USE WITH THE GEOGRAPHIC SELECTION
   Presents the special keys used in connection with the Geographic Selection Menus.

KEYBOARD USE WITH THE DATA DICTIONARY
   Presents the special keys used in connection with the Data Dictionary Processor.

KEYBOARD USE WITH THE DATABASE MANAGEMENT FUNCTIONS
   Presents the special keys used in connection with the Database Management Function option.
Most keys have always the same use as in other PC applications.

Letter Keys always display the letter they are marked with when struck, as in any typewriter.

Number Keys and Special Symbol Keys also display the number or symbol they are labeled with, except for the ? key, which is sometimes used to summon help as explained below. It is important to note, that your PC has two sets of number keys: One across the top, and another on a numerical pad on the right side. Most keyboard types use the numerical pad also for special cursor control, the number function being activated by depressing the Num Lock Key once. If this is the case, it is suggested that you do not use the numerical pad to enter numbers, since this will require constant shifting between the two functions.

The Space Bar always produces a blank space.

The SHIFT Key and the Caps Lock Key are used as in most typewriters. REDATAM does not discriminate between upper and lower case letters, so either, or any combination can be used.

The ? Key is often used to invoke a special help facility that provides information about general system usage. It is important to note that the combination SHIFT key and ? must be used to access help.

There are several special keys that are used for which some naming conventions are explained here:

The Up Arrow Key is the key that corresponds to the number 8 of the numeric keypad and also has an arrow pointing upwards.

The Down Arrow Key is the key that corresponds to the number 2 of the numeric keypad and also has an arrow pointing downward.

The Left Arrow Key is the key that corresponds to the number 4 of the numeric keypad and also has an arrow pointing to the left.

The Right Arrow Key is the key that corresponds to the number 6 of the numeric keypad and also has an arrow pointing to the right.

The HOME Key is the key that corresponds to the number 7 of the numeric keypad and is also labeled with "Home".

The END Key is the key that corresponds to the number 1 of the numeric keypad and is also labeled with "End".
The **PgUp** Key (pronounced "page up") is the key that corresponds to the number 9 of the numeric keypad and is also labeled with "Pg Up".

The **PgDn** Key (pronounced "page down") is the key that corresponds to the number 3 of the numeric keypad and is also labeled with "Pg Dn".

The **INS** Key (pronounced "insert") is the key that corresponds to the number 0 of the numeric keypad and is also labeled with "Ins". It is not used in REDATAM.

The **DEL** Key (pronounced "delete") is the key that corresponds to the decimal point of the numeric keypad and is also labeled with "Del".

All these keys are active only if the **Num Lock** key is "turned off". Some keyboards have a light indicating the status. If your's does not, some experimentation may be required. The computer always starts with Num Lock off.

The **ESC** Key (pronounced "escape"), is usually either over the numeric keypad or on the top left side of the keyboard. It is labeled "Esc".

The **CTRL** Key (pronounced "control"), is usually on the left side of the keyboard and labeled "Ctrl".

The **ALT** Key (pronounced "alternate" or "alt") is usually on the lower left side of the keyboard and labeled "Alt".

The **BACKSPACE** Key is usually to the right and over the numeric keypad and has a left pointing arrow ( <-- ).

The **ENTER** Key is on the right side of the keyboard, to the left of the numeric keypad and is labeled with "Enter", "Return" or with a bent arrow.

The **Fn** Keys (or function keys), where n is a number from 1 to 10 (1 to 12 in some keyboards), correspond to the special function keypad, usually to the left of the keyboard, and sometimes over the top, with keys labeled F1 through F10 (or F12).

Some keys are used in combination with others. One key is pressed first and held down while the other is pressed. This is shown by separating the keys with a dash. For instance, **SHIFT-Right Arrow** means holding down the **SHIFT** key, while pressing the **Right Arrow** key, just as when typing in upper case. Keys that are usually used in combination, are the **SHIFT**, **ALT** and **CONTROL** keys.

If an invalid key is pressed, REDATAM either ignores it or responds with a beep.
REDATAM 2.00 User's Manual

REFERENCE ------------------ 4.5 Keyboard: USE IN MENU SCREENS

---------------------------------
KEYBOARD USAGE IN MENU SCREENS
---------------------------------

REDATAM makes ample use of Menu Screens, with several options to select from. The following keys can then be used:

The Up and Down Arrow Keys can be used to highlight the desired option. The option can then be selected by pressing ENTER.

An option may also be selected by pressing the underlined or highlighted letter of the description of that option. It usually corresponds to the first letter of the description.

The ESC key returns the user to the previous menu, or to the operating system if located at the Main Menu.
Text needs to be entered often. This is done with the normal typewriter keys. However, corrections can be made by using the following keys:

The **Left and Right Arrow Keys** can be used to position the cursor within the text being entered.

The **DEL Key** erases the character at the current cursor position.

The **BACKSPACE Key** erases the character to the left of the cursor and moves the remaining characters one space to the left.

In addition, if text is entered in a full screen mode, that is, the cursor can be moved between different lines, the latter is accomplished with the **Up and Down Arrow Keys**.

Characters entered are inserted in front of the current cursor position, displacing the rest of the line to the right. The **INS Key** has no effect.
REDATAM 2.00 User's Manual

REFERENCE --------- 4.5 Keyboard: USE WITH THE HELP FACILITY

KEYBOARD USE WITH THE HELP FACILITY

REDATAM often offers on-line, context-sensitive help to the user, explaining the use of the different keyboard options. Help, when available, can be requested by pressing the ? key.

When ? is pressed (SHIFT key and the ? key), a "window" appears over the current screen with information on the present subject.

The Up and Down Arrow Keys can be used in order to switch between different help screens. The bottom of each screen informs if more screens are available.

The ESC Key exits the help facility and restores the screen to its previous state.

No work can be done while a help screen is on. Only the above mentioned keys are accepted.

When Help (?) is invoked from the Statistical Processor, the whole screen is replaced by a list of available commands. The user can select the command for which information is desired.

The Up and Down Arrow Keys and Left and Right Arrows Keys are used to highlight the command for which information is required.

As a command is selected, a brief description is shown. The PgUp and PgDn Keys are used to see the previous or next pages.

The ESC Key is used to exit Help and return to the Statistical Processor.
The Statistical Processor is a "command driven" system. This means that the user types in commands on the screen, REDATAM checks them for syntax and context and either accepts them or rejects them. If the command is rejected, it is re-displayed for the user to correct it. Commands are entered or corrected in one or more lines. They can be edited as described in CHARACTER AND SCREEN EDITING KEYS. In addition, the following keys can be used:

ENTER, tells REDATAM that the command has been entered or corrected and that it should be checked for syntax.

HOME, moves the cursor to the left-most position on the command line.

END, moves the cursor to the right end of the command line.

ESC, erases the entire command line.

The last line may be used to enter new commands.

When the REVIEW command is entered, the following keys are used:

The Up and Down, Right and Left Arrows are used to move up, down or along the command set.

F1 - Inserts a blank line before the current command, where a new command can be entered.

F2 - Deletes the current line.

F3 - Collects a block of commands to be saved. Pressing it once activates collect mode. The cursor can then be moved down until the desired lines are highlighted. Pressing F3 again returns to edit mode while the lines remain highlighted. When F5 is pressed, that block of commands is saved. F3 is pressed twice to "deselect" a block.

F4 - Includes a file of commands immediately above the line where the cursor is currently at. REDATAM asks for the name of the command file to include.

F5 - Saves the entire set of commands. If a group of commands has been collected (with F3), only the commands in the highlighted block are saved. REDATAM asks for a file name to save to.

F6 - Loads an existing command file, replacing the current one.

F7 - Invokes BROWSE, with information on Data Dictionary variables.
4.5 Keyboard: USE WITH THE STAT PROCESSOR

F8 - Calls the Help facility, with information on each command.
F9 - Quits REVIEW, ignoring any changes made.
F10 - Exits REVIEW, returns to the Statistical Processor, and checks the commands for syntax.

When the output of a process is shown on the screen or the VIEW command is entered, the following keys are used:

The Up and Down arrows scroll the output one line up or down respectively.

The Left and Right arrows move the output one screen column to the left or right respectively.

HOME moves to the top of the output.

END moves to the bottom of the output.

PgUp moves up one half screen of output.

PgDn moves down one half screen of output.

CTRL-Left arrow pans the screen one output column to the left.

CTRL-Right arrow pans the screen one output column to the right.

ESC exits VIEW.

Keyboard usage with the BROWSE command in the Statistical Processor, itself, or when called from within REVIEW, is the same as when used with the Data Dictionary:

The Up and Down and Left and Right Arrows are used to move the cursor between the different variables, the selected one being highlighted. The same keys can be used to scroll the screen when positioned at the last line.

The PgUp and PgDn Keys are used to scroll the Code / Description window when looking up a particular variable.

ESC, ends BROWSE and returns to the previous screen.

See KEYBOARD USE WITH THE HELP FACILITY for information on keyboard usage when Help (?) is called from within the Statistical Processor.
When performing geographic selections, the following keys are used:

P, performs a partial selection.

T, performs a total selection.

SPACE, excludes a subdivision from the selection.

SHIFT-Up Arrow, displays the next hierarchically higher level.

SHIFT-Down Arrow, displays the next hierarchically lower level.

Left Arrow, moves to the next geographic area "to the left" at the same level.

Right Arrow, moves to the next geographic area "to the right" at the same level.

HOME, moves to the next pending selection. If all selections are complete, returns to the highest geographical level. HOME should always be used before finishing.

END, ends the selection process, saves the information in a selection file and returns to the MAIN MENU.

ESC, cancels all selections and returns to the MAIN MENU.

If any of the requested moves is not possible (because there are no more geographic areas in the direction requested); a message appears on the screen.
When under control of the Data Dictionary Processor, some special keys are used in addition to those already defined.

When defining a new Data Dictionary or updating an existing one, the following keys are used in order to move between the different fields and screens:

HOME, positions the cursor at the beginning of the field.
END, positions the cursor at the end of the field.
CTRL-A, backs up the cursor to the previous word.
CTRL-F, moves the cursor to the next word.
ENTER, moves the cursor to the next item.
PgDn, moves to the next screen level.
PgUp, backs up to the previous screen level.
Down Arrow, moves the cursor to the next field.
Up Arrow, backs up the cursor to the previous field.
Alt-S, saves the Dictionary and continues work.
Alt-D, saves the Dictionary and ends the session.
Alt-A, abandons the session without saving any entries.
SHIFT-Right Arrow, moves to the next variable.
SHIFT-Left Arrow, backs up to the previous variable.
SHIFT-INS, inserts variables and codes.
SHIFT-DEL, deletes variables and codes.

When in the Data Dictionary Browse function (or invoking the BROWSE command from within the Statistical Processor or REVIEW), the following keys are used:

The Up and Down and Left and Right Arrows are used to move the cursor between the different variables, the selected one being highlighted. The same arrows can be used to scroll the screen when positioned at the last line.

The PgUp and PgDn Keys are used to scroll the Code / Des-
cripition window when looking up a particular variable.

ESC, ends BROWSE and returns to the previous screen.
The Database Management Functions are "menu driven" and "fill-in-the-blanks" type. In addition to the general keys used for typing and editing entries, the following keys are used:

HOME, moves the cursor to the beginning of the field.
END, moves the cursor to the end of the field.
ENTER, moves the cursor to the next field.
Down Arrow, moves the cursor to the next field.
Up Arrow, backs up the cursor to the previous field.
PgDn, starts a process or goes to the next step or screen.
ALT-A, abandons the session, no work done.
4.6. REDATAM File Names:

REDAATAM uses many files, which can be classified into several different categories, depending on their extensions (the DOS three letter suffix). These are:

**Program Files:**

*.en1  English text files for the English version of the system.

*.en2  Spanish text files for the Spanish version of the system.

*.exe  Executable program files.

*.hl1  Help text files in English.

*.hl2  Help text files in Spanish.

*.pa1  REDATAM screen image files in English.

*.pa2  REDATAM screen image files in Spanish.

*.sc1  REDATAM menu libraries in English.

*.sc2  REDATAM menu libraries in Spanish.

**Database Files:**

*.bin  Transposed REDATAM data files. One for each variable in the Data Dictionary.


*.dic  Data Dictionary text files.

*.geo  Contains the geographic structure of the database.

*.inx  Geographic index files (geographic codes and names).

*.nom  The geographic names file. Contains the names of the geographic areas.

*.red  Pointers to link household and population data files.

**User Files:**

*.cmd  SL-MICRO or SPSS-PC command format file documentation (generated by the WRITE command).
**.ecf**  User-created Statistical Processor Command Files (created by the SAVE command).

**.prn**  Print format output files (generated by the PROCESS command).

**.sel**  User-created Geographic Selection Files.

**.tmp**  Temporary system work files (may be erased). REDOUTPU.TMP is the file that contains the last REDATAM output sent to the screen.
Part 5: SPECIAL FEATURES: Advanced Users

This part covers in detail the more advanced features in REDATAM. It is a complement to the TUTORIAL and the REFERENCE sections and the user is expected to be familiar with the REDATAM environment.

The following subjects are covered in the sections that follow:

- Section 5.1 explains the "philosophy" of hierarchical file processing and shows how to use special commands for that purpose.

- Section 5.2 shows how to create subset REDATAM databases, using the WRITE command.

- Section 5.3 shows how to use the features of REVIEW, the statistical command editor, in order to reduce the work required when building and modifying command sets.
5.1. **Statistics from Hierarchical Files:**

One of the most useful aspects of REDATAM is that it can handle files with two record types, hierarchically structured. It is often the case that a census includes data about each dwelling or household and separate information for each person residing or belonging to it. Through the use of some special commands, REDATAM gives the user control over the statistics that may be obtained from the hierarchical structure.

In order to take advantage of this capability, the user has to be familiar with some definitions and rules. These are presented here.

1. **DEFINITIONS**

1.1. **Data Structure Definitions:**

   a) **HIERARCHICAL STRUCTURE:** A hierarchical structure is composed of two types of records: Household or Housing records, which from here on will be referred to as H-records, and Population or Individual Records, referred to as P-records.

   H-records, which contain information about households or dwellings, can have one or more P-records, containing personal information, with one record for each person in the household. REDATAM databases will usually have this structure if the required data is available.

   Note that even though in some databases the H-record represents a house or dwelling, it normally corresponds to a household. There may, therefore, be more than one household record within the house. As there are only two levels available in REDATAM at present, in this situation, houses per se, may be distinguished by selecting only the first household. There is usually a special variable that contains the household number. In the Miranda database, this variable is HHNO.

   b) **REPRESENTATION OF THE STRUCTURE:** In order to assist in understanding the relationship between H-records and P-records, the structure of a hierarchical file may be visualized as shown on the next page.

   Note that the diagram shows how the data structure may be conceived by the user. The actual physical representation is different, among other reasons because the records are not organized by person, but by variable.
Every Household must have a "Head of Household". REDATAM census databases should always be made with the record of the Head as the first person record in the Household.

1.2. Output Definitions:

a) TABLE "TALLY LEVEL": The frequency in each cell of a table is the result of counting H-records or P-records.

- If H-records are counted, it is called an H-tally.

- If P-records are counted, it is called a P-tally.

TABLES are produced by the following commands:

AVERAGE
CROSSTABS
FREQUENCIES

1.3. Input Definitions:

a) VARIABLES: There are three types of variables available in REDATAM:
- **Geographic Variables**, referred to as G-variables. Identified by a record type of 0 in the Data Dictionary.

- **Household Variables**, referred to as H-variables. Identified by a record type of 1 in the Data Dictionary.

- **Person Variables**, referred to as P-variables. Identified by a record type of 2 in the Data Dictionary.

Output tables may have just one type of variable, or they may be mixed. Geographical variables must be explicitly referenced in the Geographic Selection before they can be used for statistics.

b) **OPTION COUNT**: This command, placed immediately after a Table command may be used to control the tally level:

   OPTION COUNT 1  =>  H-tally
   OPTION COUNT 2  =>  P-tally

1.4. **Hierarchical Processing Definitions**:

   a) **NORMAL PROCESSING**: When each record is processed only once without separate processing within each household. The ENDHP command is not present.

   b) **HIERARCHICAL PROCESSING**: When there is processing "within household". Special commands are available for this purpose:

```
IF <condition> QUALIFY
IF <condition> COUNT <user variable>
IF <condition> SUM <user variable> = <existing variable>
```

These commands must always be followed by the ENDHP (END Hierarchical Processing) command to indicate that "within household" processing has ended and that normal processing continues.

The user variables created by IF ... COUNT and IF ... SUM are H-variables.

2. **RULES**

2.1. **General**:

   a) **OPTION COUNT** always produces the tally requested, P or H. For example, the following command set would give an H-tally without the OPTION COUNT 2 command. Its presence forces a P-tally.
GEOGRAPHY casel
CROSSTABS rooms BY toilfac
OPTION COUNT 2

b) If no OPTION COUNT is present, the variables in the TABLE determine the tally level.

- If the Table contains only H-variables and/or G-variables, an H-tally is produced for that table.
- If the Table contains at least one P-variable, a P-tally is produced for that table.

For example:

GEOGRAPHY casel
CROSSTABS tenure BY highsch

Produces a P-tally, since HIGHSCHT is a P-variable, even though TENURE is not.

c) In an H-tally table, the result is usually "indeterminate" when there are one or more P-variables, since each Household is counted only once. For example:

CROSSTABS tenure BY sex
OPTION COUNT 1

REDATAM does not know in which cell to put the H-tally for a house that is, for instance, "owned" (TENURE = 1) and has both a male and a female living in it. As a default, the first record of the household, which should correspond to the Head of Household, is used (if the table were explicitly limited to "Head of Household", for example, the indeterminacy would be eliminated since there is only one head per household).

Note that the variables created by a RECODE have the same level (H or P) as that of the variable being recoded. The variables created by a COMPUTE are P-variables if at least one of the variables that define them is a P-variable. Otherwise they are H-variables.

2.2. Within Household Processing:

a) IF ... COUNT and IF ... SUM may be used to respectively obtain the number of specified persons within a household and the sum of a variable for all specified persons of the household. These commands must be followed by an ENDHP command to signal the termination of "within household" processing.

If no other instruction is given (see next paragraph), on
finishing the "within household" processing, REDATAM automatically makes the household record, the newly created household variables and the record of the Head or first record of the household available for further processing. For example:

```
GEOGRAPHY ed
COMMENT count no. of persons, male or female, in H
IF sex > 0 COUNT ntot
VAR LABEL ntot "Household size"
VALUE LABEL ntot 1 "1 person" 20 "20 persons"
ENDHP
FREQUENCY ntot, sex
```

would give frequency tables of household size and the sex distribution of Heads (since only the record of the Head is available after the ENDHP).

b) **IF ... QUALIFY** must be used if more than the record of the Head is required after the ENDHP. **IF ... QUALIFY** is used to designate the P-records within each household that are desired for later processing. Hence, in the above command set, adding the command:

```
IF sex > 0 QUALIFY
```

before the ENDHP, makes all the P-records available for table production. Of course, a more restrictive expression could be used if not all persons were of interest in the FREQUENCY table. The **IF ... QUALIFY** expression does not affect **IF ... COUNT** or **IF ... SUM** or vice versa.

**IF ... QUALIFY**, with the obligatory ENDHP, are not normally required unless **IF ... COUNT** and/or **IF ... SUM** are used, since the same filtering can be done by an ordinary IF.

Note that an ordinary IF before an ENDHP continues to apply for the post-HP processing until a new IF command is given (see the IF commands in the REFERENCE Section).

**RECOMMENDATION:** Since hierarchical processing often involves a complex set of commands that may not give the expected output or, even when the commands are correct, the results may not be substantively meaningful, it is strongly suggested that the user take an area with a small population (perhaps around 500) to permit applying the command set and checking the results with other tabulations to ensure that they are correct.

As an example, a useful application of these commands is the ability to select the most outstanding case of some characteristic within the household. Consider the following command se-
COMMENT This sequence selects the person with the highest income in each household
COMMENT Establish income for first person
IF indno = 1
COMPUTE maxinc = income
COMMENT Determine if person's income is equal or greater, to previous maximum. If so qualify that record, and update household maximum
IF income >= maxinc QUALIFY
IF income >= maxinc
COMPUTE maxinc = income
COMMENT Go to next person in household
ENDHP
COMMENT Select only those whose income is equal to the, final maximum
IF income = maxinc INCLUDE
....
5.2. **Use of the WRITE Command:**

The WRITE command is a powerful utility that allows the user to extract files with selected data from the REDATAM database. Files can be generated as DOS character files, as well as with SPSS-PC or SL-MICRO data input file formats, including automatic generation of their associated command files. These cases are covered in the REFERENCE section (Part 4) under the WRITE and OPTION commands.

Another option permits the generation of a complete REDATAM database that contains a subset of the original files, as well as other user-defined variables. This option is useful if a lot of work needs to be done on a part of the database, such as on a particular geographic area, or if just a few variables are required. In this way, storage and processing time requirements may be reduced significantly. This section explains how to use this option.

The format of the WRITE command when generating a subset REDATAM database is the following:

```plaintext
WRITE <variable 1> [variable 2] [ ... ]... 
[/] [variable 3] [variable 4] [ ... ]... 
OPTION DOCUMENTATION = REDATAM FILENAME = <data base name>
```

The OPTION command must follow the WRITE command in the form specified. The `<data base name>` must be four characters in length and must start with a letter. No directory specification is allowed. All files are created in the user's work directory.

The WRITE command generates the following:

a) A file with the **Geographic Structure**, named `dbname.GEO`, where "dbname" is the name of the database supplied with the OPTION command.

b) The **Data Dictionary** and control files, containing the definition of all the variables that will be contained in the database. The files are named `dbname.DIC` and `dbname.CTL` respectively.

c) The **Data File**, named `dbname`.

After these files have been created, the AUTOLOAD FUNCTION of the Database Management Functions menu must be run in order to generate the transposed data and index files, thus completing the creation of the REDATAM database.

The user must make sure that there is sufficient disk space to complete the process. Note that a flat data file is created before the REDATAM database files are written. Therefore, during
the creation process, space equal to approximately twice the final database size is required.

If the size of any file exceeds the DOS 32 megabyte limit, the database has to be generated in parts, and must be consolidated with the APPEND VARIABLES FUNCTION, following the AUTOLOAD process.

Different types of REDATAM databases may be created, depending on how the WRITE command is used. There are several options:

I. WRITE <geographic variable>

If just one geographic variable is specified, all selected records are written, including all the variables of the current database. Any geographic variable may be entered. For instance:

```
GEOGRAPHY case1
UNIVERSE age > 5 AND age < 40
WRITE town
OPTION DOCUMENTATION = REDATAM FILENAME = rurl
```

would generate a database of all rural areas (as specified by the CASE1 Geographic Selection File) containing people whose ages fluctuate between 5 and 40. The database would be named RURL and it would contain all the variables of the original database with the same hierarchical structure. Any geographic variable could be specified instead of TOWN in the WRITE command, with identical results.

The following message would be printed after the process is completed:

* WRITE CASES to file 'RURL'. Number of cases are 2537.
* Record size is 81 bytes. Total space occupied 410594 bytes.

The following files would be created:

<table>
<thead>
<tr>
<th>File</th>
<th>Size</th>
<th>Date</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>RURL</td>
<td>229246</td>
<td>5-02-87</td>
<td>19:10</td>
</tr>
<tr>
<td>RURL</td>
<td>DIC 10734</td>
<td>5-02-87</td>
<td>18:33</td>
</tr>
<tr>
<td>RURL</td>
<td>CTL 861</td>
<td>5-02-87</td>
<td>18:33</td>
</tr>
<tr>
<td>RURL</td>
<td>GEO 114</td>
<td>5-02-87</td>
<td>18:33</td>
</tr>
</tbody>
</table>

II. WRITE <variable 1> [variable 2] [...] ...

In this case a non-hierarchical database is generated, including all the specified variables plus all the geographic variables and the record identification variable. For instance:
GEOGRAPHY CASE1
UNIVERSE AGE > 5 AND AGE < 40
WRITE WATER LIGHT AGE HIGHSCHT
OPTION DOCUMENTATION = REDATAM FILENAME = rur2

would generate a database of all rural areas (as specified by the CASE1 Geographic Selection File) containing people whose ages fluctuate between 5 and 40. The database would contain the variables WATER, LIGHT, AGE and HIGHSCHT in addition to PROVINCE, TOWN, ED and IDENT. It would be named RUR2. Only one record type would be written, each one containing both the housing and population data. The IDENT variable contains the record type and is created automatically.

The following message would be printed after the process is completed:

* WRITE CASES to file 'RUR2'. Number of cases are 2537.
* Record size is 15 bytes. Total space occupied 38055 bytes.

The following files would be created:

<table>
<thead>
<tr>
<th>RUR2</th>
<th>43129</th>
<th>5-02-87</th>
<th>19:41</th>
</tr>
</thead>
<tbody>
<tr>
<td>RUR2</td>
<td>DIC</td>
<td>2308</td>
<td>5-02-87</td>
</tr>
<tr>
<td>RUR2</td>
<td>CTL</td>
<td>126</td>
<td>5-02-87</td>
</tr>
<tr>
<td>RUR2</td>
<td>GEO</td>
<td>114</td>
<td>5-02-87</td>
</tr>
</tbody>
</table>

III. WRITE <var 1> [var 2] [...] ... / <var 3> [var 4] [...] ...

In this case a hierarchical database is generated, including all the specified variables plus all the geographic variables and record identification variables. The variables before the slash (/) must be housing variables and will make up the housing records, while those that follow it are written into population records. For instance:

GEOGRAPHY CASE1
UNIVERSE AGE > 5 AND AGE < 40
WRITE WATER LIGHT / AGE HIGHSCHT
OPTION DOCUMENTATION = REDATAM FILENAME = rur3

would generate a hierarchical database of all rural areas (as specified by the CASE1 Geographic Selection File) containing people whose ages fluctuate between 5 and 40. The database would contain the variables WATER, LIGHT, AGE and HIGHSCHT in addition to PROVINCE, TOWN, ED and IDENT. It would be named RUR3. The WATER and LIGHT variables would be written into the housing record, while AGE and HIGHSCHT would make up the population record.

The following message would be printed after the process is completed:
The following files would be created:

<table>
<thead>
<tr>
<th>File</th>
<th>Size</th>
<th>Date</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>RUR3</td>
<td>47019</td>
<td>5-02-87</td>
<td>21:07</td>
</tr>
<tr>
<td>RUR3</td>
<td>2328</td>
<td>5-02-87</td>
<td>20:56</td>
</tr>
<tr>
<td>RUR3</td>
<td>126</td>
<td>5-02-87</td>
<td>20:56</td>
</tr>
<tr>
<td>RUR3</td>
<td>114</td>
<td>5-02-87</td>
<td>20:56</td>
</tr>
</tbody>
</table>

Once the WRITE command, in any of the above forms, has been processed, the data file has to be converted into REDATAM format, which consists in one transposed data file for each variable and index files for the geographic variables. This is accomplished through the AUTOLOAD FUNCTION of the Database Management Functions. APPENDIX D, explains how to accomplish that process.

One of the problems that may be encountered when generating large databases, is that the data file may exceed the 32 MB file size limit imposed by DOS. This will usually be the case if the database requires a laser disk. The following procedure has to be employed in order to circumvent this problem:

A. Create several different databases (with several WRITE commands), each one containing only as many variables as are possible without exceeding the 32 MB limit. Note that REDATAM can process only one WRITE with OPTION DOCUMENTATION = REDATAM per run, so several processes have to be made each time changing the variables in the WRITE command.

B. If a single record or non-hierarchical database is desired, variables may be specified in any order in the WRITE command. However, if a hierarchical database is desired, the first WRITE must include two record types. The first record must include housing variables only, while the second must include only population variables. For instance:

For a single record type database:

```plaintext
WRITE sex  water  hignscht  light
OPTION DOCUMENTATION = REDATAM FILENAME = SET1
```

For a hierarchical database:

```plaintext
WRITE  water  light  /  sex  hignscht
OPTION DOCUMENTATION = REDATAM FILENAME = SET1
```

The name of the file in the first WRITE will be the final name of the REDATAM database.

C. Subsequent WRITEx must also include two record types for hierarchical databases, although housing and population
variables may be placed in any of the two records. For single record databases all WRITEs should include just one record type. For instance:

For a single record type database:

WRITE typescht toilfac age dwelling
OPTION DOCUMENTATION = REDATAM FILENAME = SET2

For a hierarchical database:

WRITE typescht toilfac / age dwelling
OPTION DOCUMENTATION = REDATAM FILENAME = SET2

Variables already specified in previous WRITEs may be repeated, but this will slow down the building of the database.

This process is repeated until all desired variables have been included in some WRITE command. It is very important that each WRITE include exactly the same records, that is, the same Geographic Selection and Statistical Selection Statements.

D. The AUTOLOAD FUNCTION of the Database Administration functions is run for each database, as described in APPENDIX D.

E. The different databases are now appended into one REDATAM database through the use of the APPEND VARIABLES FUNCTION, as described in APPENDIX D.
5.3. Use of the REVIEW Command:

Experienced REDATAM users will soon find that they spend a significant portion of their time preparing and modifying statistical command sets. Furthermore, a lot of command sets share common procedures, for which the user will have already developed proven and efficient techniques. As a result, the user often reenters the same command sequences over and over again.

This section provides a short tutorial on the use of the REVIEW command, which will help you reduce the number of keystrokes required to arrive at the desired result. You may follow this example directly on your computer.

The Geography Selection Files created in the TUTORIAL (Part 3) are used here. If they are not available, use any other, since it will not affect this exercise.

Imagine that you are doing a study that requires you to do cross tabulations with the age of people. Since you want to keep your tables reasonably short and you do not need single years of age, you group people's age into 5-year groups. Load the Statistical Processor and enter the following set of commands:

```plaintext
RUN NAME "Test Case 1"
GEOGRAPHY case2
AREABREAK town
COMPUTE age5 = age / 5
RECODE age5 (10 THRU HIGHEST = 10)
VAR LABEL age5 "Age Groups in 5 Year Increments"
VALUE LABEL age5 0 "Less than 5" 1 "5 - 9" 2 "10 - 14",
3 "15 - 19" 4 "20 - 24" 5 "25 - 29" 6 "30 - 34" 7 "35 - 39",
8 "40 - 44" 9 "45 - 49" 10 "50 or More"
IF mainact = 1
CROSSTABS age5 BY highschool
```

AGE5 will group all ages 0 through 50 into 5-year increments. Labels have been provided for the variable and all its values. Save this command set on disk by entering:

```
SAVE test1
```

If the file already exists, you will be asked if you want to overwrite it. If you are sure nobody needs it, proceed, otherwise, choose another name.

You may actually PROCESS this command set, but it is not really necessary for this example. Imagine that you did and now realize that you do not need to include people whose age is less than 15 in the tabulation, and on the other hand you want to group people up to age 60. Without REVIEW, you would have to reenter the entire command set. Instead type:
REVIEW

You will see the following screen:

```
REDATAM V2.0 Editor       1180e de Datos:       Fecha: 25/3/1987
RUN NAME "TEST CASE 1"
GEOGRAPHY CASE2
AREA BREAK TOWN
COMPUTE AGES = AGE / 5
RECODE AGES (10 THRU HIGHEST = 10)
VALUE LABEL AGES 0 "LESS THAN 5" 1 "5 - 9" 2 "10 - 14"
3 "15 - 19" 4 "20 - 24" 5 "25 - 29" 6 "30 - 34" 7 "35 - 39"
8 "40 - 44" 9 "45 - 49" 10 "50 OR MORE"
IF MAINACT = 1
CROSSTABS AGES BY HIGHSCHT
```

When REVIEW is invoked, the commands of the current set are displayed, but unlike in the standard command entry mode, you may, through the use of the up and down arrow keys, move the cursor to any command that needs to be changed.

Now, move the cursor, which should now be to the left of the first line, down to the RECODE command. Since we want to group up to age 60, we need to change both 10s in the RECODE command to 12s. Use the right and left arrow keys to move along the command line until you are over the 0 of the first 10. Then enter a 2. The 2 should be inserted between the 1 and the 0. Next move the cursor to the 0 and delete it by hitting the DEL key. Repeat this with the next 10. The command should now look like this:

```
RECODE AGES (12 THRU HIGHEST = 12)
```

Next, we need to change the VALUE LABEL command in order to suppress tabulation of values less than 3 and include values 11 and 12. Move down to the VALUE LABEL command and move to the right until you are over the 0 following AGES. Delete all characters except for the comma, by repeatedly pressing the DEL key. When you are done, the line should look like this:

```
VALUE LABEL AGES,
```

The next line does not need to be changed, so move down to the third line of the VALUE LABEL command and delete the "OR MORE" phrase. Then add the following:

```
- 54" 11 "55 - 59",
```
The line should end up looking like this:

\texttt{8 "40 - 48" 9 "45 - 49" 10 "50 - 54" 11 "55 - 59",}

We still need to add another range, so move down one line and press the F1 key once. A blank line should be inserted above the cursor. Note that this function is documented at the bottom line of the screen. Enter the following into the new line:

\texttt{12 "60 or more"}

All the required changes are now done without the need to retype the whole command set. We may now save it by pressing F5. REVIEW will ask for a file name. Enter the same name we used before, that is TEST1. You will be asked to confirm that you want the file overwritten. The commands are saved to disk, and you return to REVIEW.

There are a few other keys that you may use to edit your commands:

- \texttt{END} moves the cursor to the right end of the command.
- \texttt{HOME} moves to cursor to the left-most column of the screen.
- \texttt{BACKSPACE} erases the character to the left of the cursor.
- \texttt{ESCAPE} erases the whole command, leaving the line blank.

While in REVIEW, commands are entered in the usual way, that is, only the minimum abbreviation of the command needs to be entered (the editor puts in the full name when the cursor moves to another line or ENTER is pressed). The commands, however, are not checked for syntax at this time.

In addition to the F1 and F5 keys, you may use F2 in order to delete a whole line. F4 is the equivalent within REVIEW of the BROWSE command, providing information about the database variables. F8 invokes the help facility (same as ? within the Statistical Processor).

You can return to the Statistical Processor with either F9 or F10. F9 discards all the changes and reinstates the command set as it was when REVIEW was invoked. F10 updates the command set and rechecks it for syntax, just as if you had LOADed it. Hit F10 now.

You may now perform a PROCESS with the modified command set. Let us assume that you will need to build several command sets using the same recodification for age. You may prepare a "canned" recodification that can be included into each command set as needed.

Enter REVIEW again and move the cursor to the COMPUTE command. Then hit F3 (collect). The cursor will change from a rectangle to a line. Move the cursor down, and as you do, the lines should be
highlighted. Continue until the VALUE LABEL command and its continuation lines are highlighted. Then hit F3 again. The cursor should return to its normal form and you may now continue to work as usual. The lines that have been "collected" remain highlighted. As a matter of fact, you may even insert new lines within those collected, which will also be highlighted.

Now press the save key, F5 (save), once more. This time however, only the collected, that is highlighted, lines will be saved. When asked for a file name, you could enter something like AGE5, so that you may easily identify the file that contains the age recodification in 5 year groups. Remember though, that when one or more lines are highlighted, only those are saved. Otherwise, the whole command set is saved.

Note that the collected commands remain highlighted. Press F3 twice in order to de-select them.

Exit REVIEW with either F9 or F10 and then clear all commands from memory by entering:

```
NEW
```

Now enter the following command set:

```
RUN NAME "Test Case 2"
GEOGRAPHY case3
AREA BREAK RECODE
COMPUTE age5 = 0
IF occstat = 1 OR occstat = 2
CROSSTABS age5 BY highscht BY occstat
```

This is a new command set that involves the new age re-codification. Note that we have included a dummy COMPUTE in order to define AGE5 to REDATAM. Otherwise, the CROSSTABS command would not have been accepted. Next, enter REVIEW and move the cursor to the IF command. Then press F4. REDATAM will ask for the name of the file to "include". Enter AGE5, or whatever name you gave the collected commands. The whole file will be inserted above the line of the cursor. Now move to the dummy COMPUTE and erase it with F2, since the correct codification has been included. The new command set may be saved and/or processed.

F6 allows you to replace the current commands with a new command set from disk. Note that unlike INCLUDE (F4), that inserts a file, or the LOAD command, which appends the file to the existing commands, F6 replaces the current set with that of the file. Remember that all changes may be cancelled by exiting REVIEW with F3.

With this example, you have been shown the use of the REVIEW command to edit or create command sets and its capabilities to collect or merge different files. A significant amount of time
and effort can be saved if it is properly used, not to mention reduced data entry errors.

You may also use external editors, such as Sidekick, in order to work on your command files. Command files are saved as simple ASCII text files. You have to make sure that you respect REDATAM's command entry standards, such as 80 character lines and syntax. No extraneous characters should be added. Commands will be checked for syntax when they are LOADED.
APPENDIX A

Equipment Required to Use REDATAM

REDATAM operates on the IBM PC family of microcomputers, that is, the IBM-PC-XT or the IBM-PC-AT. It should also run on most of the fully "IBM compatible" machines, and as of this date it has been run on the following equipment:

- EPSON Equity I
- SANYO PC
- CORONA PC
- COMPAQ Portable
- GREAT WALL 0520C-H

The following minimum hardware and system software configuration is required:

- 640KB of RAM Memory.
- 1 Floppy Disk.
- A Hard Disk unit of 20-100 MB (Size depending on the volume of the files to be loaded. A 20 MB disk should be able to store at least 200,000 census records or up to 1,000,000 records if data compression is used).
- A Monochrome or Graphics Monitor.
- A Printer with a paper width of at least 80 characters.
- The IBM PC-DOS Operating System Version 2.0 or higher.

REDATAM also supports a "Write Once, Read Many Times" (WORM) laser-optical disk if very large storage capacity is required (presently up to 115 MB on a single removable disk). At present, REDATAM can work with an Information Storage Inc. (ISI) single side Superstore 2000 Model 525WC drive.

REDATAM is written in the C language, and has not shown any compatibility problems with popular memory-resident programs, such as Sidkick.

EPSON is a registered trademark of Epson America, Inc.
SANYO is a registered trademark of Sanyo Corporation
CORONA is a registered trademark of Corona Data Systems
COMPAQ is a registered trademark of Compaq Computer Co.
SIDEKICK is a registered trademark of Borland International
GREAT WALL is a Chinese People's Republic microcomputer
ISI is a registered trademark of Information Storage Inc.
APPENDIX B

System Installation and Testing

This appendix describes the steps required to install and test REDATAM on your computer. It is assumed that you are familiar with the use of microcomputers and their operating system.

First, make sure that your hardware configuration corresponds to at least the minimum indicated in APPENDIX A, and that your hard disk has enough space to load the system and the database.

Check your CONFIG.SYS file, on the root directory. It must contain the following parameters:

```
BUFFERS=20
FILES=20
```

If the parameters are not included in the file, or their values are smaller, you must add or change them. If the file does not exist, create it on your root directory.

A path to the REDATAM directory is also needed. Check the PATH command on your AUTOEXEC.BAT file and include a path to C:\REDATAM. If no AUTOEXEC.BAT file is present, you must create it in the root directory of the disk from which the system boots, and include the following command:

```
PATH C:\REDATAM
```

REDATAM provides an automated loading procedure. This procedure creates a directory for the programs, called REDATAM, and another for the example database included with the system, called MIRANDA.

To start the loading procedure, position yourself on the C: disk, inside a work directory. Enter the following:

```
C:
CD \work
[ENTER]
[ENTER]
```

where [ENTER] represents the ENTER or RETURN key, and "work" is to be replaced by the name of any working directory.

Insert Disk 1 of the REDATAM distribution diskettes into drive A:. Check the content of the README file on the disk for changes that could not be included in this manual. Enter:

```
A:TYPE README
[ENTER]
```

in order to display the file on the screen, or:
A: COPY README LPT1: [ENTER]

in order to send it to the printer. Unless instructed otherwise by the INSTALL file, start the installation process by entering:

A: INSTALL ENG [ENTER]

The files will then be copied onto the hard disk. The program will request the remaining diskettes as they are needed.

If you replace the ENG parameter with ESP, the install procedure will be run in Spanish.

When all files have been copied, a file named MIRANDA.BAT will have been placed in the current user directory. This batch file contains the following statements:

```
SET DBASE=C:\MIRANDA\nSET PROG=C:\REDATAM\nSET LANG=ENG
SET DBNAME=M180
REDATAM
```

This batch file is used to start REDATAM and load the example database named M180. The following environment variables are used:

- **DBASE**: Identifies the directory where the database to be used resides.
- **PROG**: Identifies the directory where the REDATAM program resides.
- **LANG**: Identifies the language that is to be loaded. Languages currently supported are English (ENG) and Spanish (ESP).
- **DBNAME**: Identifies the name of the database to use.

Similar files and directories may be created in order to access other databases.

The TEST.ECF and TEST.SEL files are also copied to the user directory. These are used to test the installation of the system.

From your working directory start REDATAM by entering:

```
MIRANDA [ENTER]
```

You should see the following screen:
Select the **Statistical Processor** option, and after it has loaded you should see a blank screen with the following heading:

```
REDAFAM V2.0 Statistical Processor
```

at the bottom of the screen enter:

```
LOAD TEST
```

You may enter it in either upper or lower case. REDATAM should display the following on the screen:

```
LOAD TEST
Loading...
RUN NAME "Test REDATAM Installation"
GEOGRAPHY TEST
File "TEST.SEL" selected
FREQUENCIES ATTSCH
COMPUTE GROUP = (AGE / 5) + 1
RECODE GROUP TO GROUP (9 THRU HIGHEST = 9)
VAR LABEL GROUP "5 Year Age Groups"
VALUE LABEL GROUP 1 "0 thru 4 years" 9 "40 years and +"
CROSSTABS GROUP BY ATTSCH
```

Turn on your printer at this point, and then enter:

```
PROCESS OUTPUT = PRINTER
```

Should you make a mistake on any of the above commands, REDATAM will respond with the appropriate error message, and re-display
the command, which you may correct by positioning the cursor with the left and right arrow keys, erasing a letter with the DEL key or erasing the whole command with the ESC key. Once corrected, reissue it with another ENTER.

REDATAM should start processing and display the following message:

PROCESS
Tables require 288 bytes
Output mode is VIEW
Process started at 14:06:04

After a minute or so, the following output should appear on your printer:

REDATAM System V2.0      Tue Mar 31 11:03:27 1987
1> RUN NAME "Test REDATAM Installation"
2> GEOGRAPHY TEST
3> FREQUENCIES ATTSCB
4> COMPUTE GROUP = (AGE / 5) + 1
5> RECODE GROUP TO GROUP (9 THRU HIGHEST = 9)
6> VAR LABEL GROUP "5 Year Age Groups"
7> VALUE LABEL GROUP 1 "0 thru 4 years" 9 "40 years and +"
8> CROSSTABS GROUP BY ATTSCB

"Test REDATAM Installation" Sun Feb 01 14:21:17 1987

FREQUENCY FOR ATTSCB Attendance at School
GEOGRAPHY IS GLOBAL

<table>
<thead>
<tr>
<th>CATEGORY</th>
<th>CODE</th>
<th>ABSOLUTE</th>
<th>RELATIVE</th>
<th>ACCUMULATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>FREQ</td>
<td>FREQ</td>
<td>FREQ</td>
</tr>
<tr>
<td>Fulltime Student</td>
<td>1</td>
<td>279</td>
<td>32.4</td>
<td>32.4</td>
</tr>
<tr>
<td>Parttime Student</td>
<td>2</td>
<td>1</td>
<td>0.1</td>
<td>32.5</td>
</tr>
<tr>
<td>Not Applicable</td>
<td>3</td>
<td>581</td>
<td>67.5</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td><strong>861</strong></td>
<td><strong>100.0</strong></td>
<td></td>
</tr>
</tbody>
</table>

0 VALUES OUT OF RANGE
"Test REDATAM Installation"  Sun Feb 01 14:21:17 1987

CROSSTAB OF GROUP  5 Year Age Groups
BY ATTSCH  Attendance at School

GEOGRAPHY IS GLOBAL

<table>
<thead>
<tr>
<th>CODE</th>
<th>Fulltime</th>
<th>Partime</th>
<th>Not Appl</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 thru 4 years</td>
<td>1 13</td>
<td>0 112</td>
<td>125</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>116</td>
<td>0 4</td>
<td>120</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>141</td>
<td>0 6</td>
<td>147</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>1 77</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0 67</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0 51</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0 35</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0 35</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>40 years and +</td>
<td>9</td>
<td>0 194</td>
<td>194</td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>279</td>
<td>1 581</td>
<td>861</td>
<td></td>
</tr>
</tbody>
</table>

0 VALUES OUT OF RANGE

Check whether these figures match those of your output.

If everything has proceeded as shown here, then REDATAM is properly installed in your system. Exit the Statistical Processor by typing:

EXIT

[ENTER]

And you should be returned to the main menu:
Select the **Exit** option to return to the operating system. The installation is complete.

In order to install your national REDATAM census files, you should follow the instructions that come with it. If your system has a laser-optical disk cartridge, see APPENDIX C.
APPENDIX C

Utilization of a Laser Disk

REDATAM has the built-in capability to support a laser optical disk if very large databases are required. Should this be your case, you will be supplied with a cartridge with the REDATAM database already loaded on it.

At present, REDATAM can work with an Information Storage Inc. (ISI) single side Superstore 2000 Model 525WC drive. In order to use it, turn on you laser disk unit (switch on the back), and insert the cartridge in the correct position into the slot on the front. You can later manually eject it by pressing the button in the front of the unit.

The laser disk is recognized as logical drive D:. Therefore, you need to specify drive D: for the DBASE environment variable in the batch file used to load REDATAM and access the database. See APPENDIX B for more information.
APPENDIX D

Data Dictionary and Database Management Functions

D.1. Data Dictionary:

The Data Dictionary is an integral part of the REDATAM system. It is a computer readable copy of the data file documentation, containing information on what variables are available in the data file, describing what those variables mean, what categories are available within each of the variables and providing other technical information about the variables that the REDATAM system needs in order to access and prepare tabulations.

For most users of the REDATAM system, the Data Dictionary describing the file or files that will be used by REDATAM will have already been prepared by the computer technical support staff within the user's organization, and will be available as part of the REDATAM data base that is in use. The main interest that most users will have in the Data Dictionary is to print a hard-copy for their individual use. It is also possible to review the variables in the dictionary directly from its main menu, however, this will be more conveniently done using the BROWSE command of the Statistical Processor.

In order to do this and a host of other functions, REDATAM has to be activated and the Data Dictionary menus be selected. The Data Dictionary Processor is a menu driven system, that is, the user is presented with screens offering a series of choices. All screens are self-explanatory, and some further information and help is available by pressing the ? key.

In this section, each screen is shown and all its options explained. Sample screens, based on the example included at the end of the TUTORIAL (Section 3.6), are shown.
DATA DICTIONARY PROCESSOR OPTIONS

DATA DICTIONARY MENU
Controls access to all Data Dictionary options and functions.

DATA DICTIONARY BROWSE FUNCTION
Used to interactively browse through the Data Dictionary, displaying information on variables as selected.

PRINT DATA DICTIONARY OPTION
Produces a hard-copy of the selected Data Dictionary on the printer.

DATA DICTIONARY MANAGEMENT FUNCTIONS
Controls access to functions that permit the creation and maintenance of Data Dictionaries.

CREATE DATA DICTIONARY
Allows the creation of a new Data Dictionary.

UPDATE DATA DICTIONARY
Used to modify the active Data Dictionary.
The active database and identification data is displayed at the top of the screen. Instructions are displayed at the bottom.

Options:

**List Active Data Dictionary:**
This option lets the user browse through the active Data Dictionary, providing information about any variable. The DATA DICTIONARY BROWSE FUNCTION is loaded.

**Print Active Data Dictionary:**
Produces a hard copy printout of the Data Dictionary and all its variables. See the PRINT DATA DICTIONARY OPTION.

**Management Functions:**
This group of functions allows the creation, maintenance and erasure of Data Dictionaries. The ADMINISTRATIVE DATA DICTIONARY FUNCTIONS menu is shown.

**Exit to Main Menu:**
Returns to the REDATAM MAIN MENU.

**Usage:**
An option is selected by either moving the cursor with the up and down arrows until it is highlighted, and then pressing ENTER, or by directly pressing the underlined letter of the desired option. ESC may also be used to return to the MAIN MENU.
DATA DICTIONARY BROWSE FUNCTION

A series of screens are displayed, that allow the user to interactively browse through the Data Dictionary. This same function is available from within the Statistical Processor:

REDATAM V2.0 Identification

Data Base: M180     Date: 25/3/1987

General Description

Author: CELADE
Original file identification: Miranda 80
Creation date: 29/12/1986
Update date: 00/00/0000
Update author: 
Record length: 84
Record types: 2
Record identification variable: IDENT
Initial position and length: 8 - 7

Press any key. <ESC> to abandon

The first screen provides general information about the Data Dictionary. For specific information on the meaning of each of the fields, consult the CREATE DATA DICTIONARY function.

Options: The user can move to the next screen or return to the DATA DICTIONARY MENU.

Usage: As indicated at the bottom of the screen, pressing ESC returns to the DATA DICTIONARY MENU. Any other key continues to the next screen.

REDATAM V2.0 Record Types

Data Base: M180     Date: 25/3/1987

Record Type Description:

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Housing</td>
</tr>
<tr>
<td>2</td>
<td>Population</td>
</tr>
</tbody>
</table>

Press any key. <ESC> to abandon

This screen lists all the record types in the database. In this instance, the census file is hierarchical and includes both indi-
individual and housing records, each with different types of variables. More information on how to use this type of file is provided in SPECIAL FEATURES (section 5.1).

Options: The user can move to the next screen or return to the DATA DICTIONARY MENU.

Usage: As indicated at the bottom of the screen, pressing ESC returns to the DATA DICTIONARY MENU. Any other key continues to the next screen.

This is the Miranda 1980 Census of Population database.

This database is used for the Tutorial and all examples of the REDATAM manual. It is a hierarchical database, containing housing as well as population records. The complete database contains about 1600 household records and 7000 population records.

Press any key. <ESC> to abandon

This screen provides a textual description of the Data Dictionary, as entered by the person that created it.

Options: The user can move to the next screen or return to the DATA DICTIONARY MENU.

Usage: As indicated at the bottom of the screen, pressing ESC returns to the DATA DICTIONARY MENU. Any other key continues to the next screen.
### REDATAM V2.0 Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PARISH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOWN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ED</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IDENT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHNO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DWELLING</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TENURE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WATER</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOILSHR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOILFAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DWELLYR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WALLMAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROOMS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIGHT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COOKFUEL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TFAC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>INDOM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RELAY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MARSTAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>USURES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>URESF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIRTHPLA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Variable Description

<table>
<thead>
<tr>
<th>Description</th>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>(7) Type of Tenure</td>
<td>1</td>
<td>Owned</td>
</tr>
<tr>
<td>Data type</td>
<td>2</td>
<td>Leased</td>
</tr>
<tr>
<td>Compression type</td>
<td>3</td>
<td>Private Rented</td>
</tr>
<tr>
<td>Original length/compress</td>
<td>4</td>
<td>Rent Free</td>
</tr>
<tr>
<td>Initial position</td>
<td>5</td>
<td>Squatted</td>
</tr>
<tr>
<td>Decimals</td>
<td>6</td>
<td>H. P.</td>
</tr>
<tr>
<td>Minimum/maximum range</td>
<td>7</td>
<td>Govt Rented</td>
</tr>
<tr>
<td>Out of range</td>
<td>8</td>
<td>Other</td>
</tr>
<tr>
<td>Record Type</td>
<td></td>
<td>More...</td>
</tr>
</tbody>
</table>

This is an interactive screen, where the characteristics of any variable in the Data Dictionary may be selectively displayed.

Variables are shown in the upper part of the screen, while the lower part is used to list their descriptions. Instructions are shown on the bottom of the screen.

**Options:** The user can choose the variable to display, examine its characteristics, or return to the DATA DICTIONARY MENU.

**Usage:** As indicated at the bottom of the screen, the left, right, up and down arrows are used to position the cursor over the desired variable, which will be highlighted. If there are more variables than can fit in the box, the screen scrolls as the cursor reaches the box boundaries.

As variables are selected, their characteristics are displayed in the bottom half of the screen. On the lower right part of the screen the values that the variable can take and their associated descriptions are shown.

In the case shown above, the TENURE variable has been selected. The PgUp and PgDn keys may be used to scroll the "Code Description" window, if all possible value
descriptions do not fit in the box. This is the case in the sample screen, as indicated by the "more.." message at the end.

ESC ends the browse and return to the DATA DICTIONARY MENU.
PRINT DATA DICTIONARY OPTION

This option prints a hard-copy of the active Data Dictionary. The format is similar to that of the DATA DICTIONARY BROWSE FUNCTION screens, that is, it includes basic data, a list of all record types, documentation text and a description of each variable and all the values it can take.

The printer must be on-line when this option is selected. The system will halt until printing is complete.
The following menu is displayed when this option is chosen:

```
REDATAM System 2.0     Date: 01/22/1987  Time: 15:19:18

Create Dictionary
Update Dictionary
Exit to Dictionary Menu

^V Move cursor. Select with ENTER or first letter. ESC previous level
```

The active database and some other identification data is displayed at the top of the screen. Instructions are displayed at the bottom.

Options:

Create Dictionary:
This option is used to define a new Data Dictionary. The CREATE DATA DICTIONARY option is loaded.

Update Dictionary:
This option is used to update the currently active Data Dictionary. The UPDATE DATA DICTIONARY screens are shown.

Exit to Dictionary Menu:
Returns to the DATA DICTIONARY MENU.

Usage:
An option is selected by either moving the cursor with the up and down arrows until it is highlighted, and then pressing ENTER, or by directly pressing the underlined letter of the desired option. ESC or the Exit option are used to return to the DATA DICTIONARY MENU.
CREATE DATA DICTIONARY

A series of screens are displayed, that allow the user to interactively create a new Data Dictionary. This is a necessary step prior to creating a new REDATAM database from a census file. However, this step is not necessary if the database is created with the WRITE command, since the Data Dictionary will be created automatically.

As the Data Dictionary provides the information necessary to build the REDATAM database from the original census, some of the input required refers to the structure of the census file that has been down-loaded onto the microcomputer.

When this option is selected, REDATAM requests the name of the Data Dictionary to be created:

NEW DICTIONARY:   

The Data Dictionary will be created in the current user directory. The name must be four characters long. The following screen is displayed next:

```
REDATAM V2.0 $identification  Data Base: XYZ  Date: 25/3/1987

Name of author
Original file identification
Creation date  17/3/1987
Last update  17/3/1987
Update authors name
Record length
Record types
Record identification variable
Initial position and length

Record Codes  Description
```

General information about the new Data Dictionary is entered here.

Options: The user can enter and correct the data for a new Data Dictionary, move to subsequent screens, save the data entered, cancel all entries, and return to the DATA DICTIONARY MENU.

Usage: The pertinent data has to be entered as requested. As
shown on the bottom line, the up and down arrows position the cursor at the desired field.

The following information has to be provided:

- The name of the author of the Data Dictionary.

- The name that identifies the original mainframe file.

- Creation and last update dates are defaulted to the system date, but may be changed.

- The name of the person responsible for file update may be provided, although this field is usually left blank at creation time.

- The record length, in bytes, of the down-loaded census file from which the REDATAM database will be generated.

- The number of different record types in the database (such as population and housing records in a hierarchical database).

- The name of the variable that identifies the record type in the census file.

- The position and length, in bytes, of the record identification variable.

When the upper portion of the screen is filled, the codes and descriptions that identify each record type are entered. REDATAM requests as many entries as specified in the "Number of record types" field. Housing records must be identified by a 1 and population records by a 2.

When done, pressing PgDn brings up the next screen. Pressing ALT-S saves to disk all the work done so far. ALT-D saves and exits the session. ALT-A cancels all entries and ends the session. These, and additional features may be displayed by invoking help (?), which opens the following two windows:
<table>
<thead>
<tr>
<th>DESCRIPTION</th>
<th>COMMAND</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positions Cursor at Beginning of Field</td>
<td>Home</td>
</tr>
<tr>
<td>Positions Cursor at End of Field</td>
<td>End</td>
</tr>
<tr>
<td>Moves Cursor One Space to the Right</td>
<td>-&gt;</td>
</tr>
<tr>
<td>Backs Up Cursor one Space to the Left</td>
<td>&lt;</td>
</tr>
<tr>
<td>Erases Character Under Cursor</td>
<td>Del</td>
</tr>
<tr>
<td>Erases Character to the Left of Cursor</td>
<td>BACKSPACE</td>
</tr>
<tr>
<td>Backs Up Cursor to Previous Word</td>
<td>Ctrl-A</td>
</tr>
<tr>
<td>Moves Cursor to Next Word</td>
<td>Ctrl-F</td>
</tr>
</tbody>
</table>

More...

^v Move Pages. ESC Abandons

<table>
<thead>
<tr>
<th>DESCRIPTION</th>
<th>COMMAND</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moves Cursor to the Next Item</td>
<td>ENTER</td>
</tr>
<tr>
<td>Moves to the Next Screen Level</td>
<td>PgDn</td>
</tr>
<tr>
<td>Moves Cursor to the Next Item</td>
<td>v</td>
</tr>
<tr>
<td>Backs Up Cursor to the Previous Item</td>
<td>^</td>
</tr>
<tr>
<td>Saves the Dictionary and Continues Work</td>
<td>Alt S</td>
</tr>
<tr>
<td>Saves the Dictionary and Ends the Session</td>
<td>Alt D</td>
</tr>
<tr>
<td>Abandons the Session. No Save.</td>
<td>Alt A</td>
</tr>
<tr>
<td>Help</td>
<td>?</td>
</tr>
</tbody>
</table>

End...

^v Move Pages. ESC Abandons

The next screen, accessed by pressing PgDn displays the following:

REDATAM V2.0 Documentation  Data Base: XYZK  Date: 25/3/1987

Documentation

^ Cursor Up v Cursor Down PgDn Next Level PgUp Previous Level ? Help

Options: The user can enter a text describing the new Data Dictionary, move to subsequent or previous screens, save the data entered, cancel all entries, and return to the DATA DICTIONARY MENU.
Usage: This screen allows the user to enter descriptive text about the Data Dictionary. The text is not processed in any way. It is only displayed or printed when information about the Data Dictionary is requested.

Text is entered on the dotted line. When the line is full or ENTER is pressed, a new line appears. As much information as fits on the screen may be entered. The right and left and up and down arrows can be used to position the cursor where desired.

Pressing PgDn brings up the next screen while PgUp returns to the previous screen. Pressing ALT-S saves to disk all the work done so far. ALT-D saves and exits the session. ALT-A cancels all entries and ends the session. The same features (and help windows) available for the previous screen can be used here.

The next screen, accessed by pressing PgDn displays the following:

```
REDATAM v2.0 Variables               Data base: X:ZX       Date: 25/3/1987

Variable number: 1                      Data type: c/b
Name:                                      Compression type:
Description:                              Compress length:
Data type (c/b):                          Original length:
Original length:                          Compress length:
Initial position:                         Decimals:
Minimum range:                            Minimum range:
Maximum range:                            Maximum range:
Out of range:                             Record type:

Code Description
```

Options: The user can define new variables, its characteristics and the values they can take, move between variables, insert or delete them, move to previous screens, save the data entered, cancel all entries, and return to the DATA DICTIONARY MENU.

Usage: This screen is used to define all the variables that make up the database. The values they can take, and descriptions thereof are also entered here.

All fields have to be filled in with the information as requested. The up and down arrows are used to position the cursor at the desired field. REDATAM automatically provides a sequential number for each new variable.
The following information has to be provided:

- Name of the variable. Up to 8 alphanumeric characters, starting with a letter.

- A verbal description of the variable. This will be printed with all output.

- The data type, indicating whether the variable is stored in character (C) or binary (B) form. The default value is C. It is ignored for geographical variables.

- The compression type field is not used at present. All compressed variables are represented in binary form.

- The length of the variable before compression if stored in binary form. Otherwise, it is ignored.

- The length of the compressed variable. It is ignored if the variable is not compressed. It is calculated automatically.

- The starting byte of the variable within the record of the down-loaded census file.

- The number of decimals field is not used in the present version. All data is assumed to be in integer form.

- The minimum value the variable may take.

- The maximum value the variable may take.

- The value that should be assumed for values outside the specified range. Calculations are done with this value, even though the number of values out of range are noted with each output table. It is not used at present.

- The type of record to which the variable belongs, as identified by the record code (population = 2, housing = 1 and geographical variables = 0).

Once the upper part of the screen is filled, the user can proceed to enter all the possible values and descriptions for the variable under the "code" and "description" columns. Any number of values may be entered (within the range specified). The screen "scrolls" if the bottom is reached.

Pressing Shift-right arrow or Shift-left arrow respec-
Ins and Shift-Del can be used to respectively insert or delete a variable. All remaining variables are renumbered. With the use of these keys, all variables can be created and corrected.

PgUp returns to the previous screen. ALT-S saves to disk all the work done so far. ALT-D saves and exits the session. ALT-A cancels all entries and ends the session. These and other features are available through the help windows (?). These are:

<table>
<thead>
<tr>
<th>DESCRIPTION</th>
<th>COMMAND</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positions Cursor at Beginning of Field</td>
<td>Home</td>
</tr>
<tr>
<td>Positions Cursor at End of Field</td>
<td>End</td>
</tr>
<tr>
<td>Moves Cursor to the Right</td>
<td>-&gt;</td>
</tr>
<tr>
<td>Backs Up Cursor to the Left</td>
<td>&lt;-</td>
</tr>
<tr>
<td>Erases Character Under Cursor</td>
<td>Del</td>
</tr>
<tr>
<td>Erases Character to the Left of Cursor</td>
<td>BACKSPACE</td>
</tr>
<tr>
<td>Backs Up Cursor to Previous Word</td>
<td>Ctrl-A</td>
</tr>
<tr>
<td>Moves Cursor to Next Word</td>
<td>Ctrl-F</td>
</tr>
<tr>
<td></td>
<td>More...</td>
</tr>
<tr>
<td></td>
<td>^ v Moves Pages. ESC Abandons</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DESCRIPTION</th>
<th>COMMAND</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moves Cursor to the Next Item</td>
<td>ENTER</td>
</tr>
<tr>
<td>Backs Up to the Previous Screen</td>
<td>PgUp</td>
</tr>
<tr>
<td>Moves Cursor to the Next Item</td>
<td>v</td>
</tr>
<tr>
<td>Backs Up Cursor to the Previous Item</td>
<td>^</td>
</tr>
<tr>
<td>Saves the Dictionary and Continues Work</td>
<td>Alt S</td>
</tr>
<tr>
<td>Saves the Dictionary and Ends the Session</td>
<td>Alt D</td>
</tr>
<tr>
<td>Abandons the Session, No Save.</td>
<td>Alt A</td>
</tr>
<tr>
<td></td>
<td>More...</td>
</tr>
<tr>
<td></td>
<td>^ v Moves Pages. ESC Abandons</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DESCRIPTION</th>
<th>COMMAND</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inserts variables and Codes</td>
<td>SHIFT Ins</td>
</tr>
<tr>
<td>Erases Variables and Codes</td>
<td>SHIFT Del</td>
</tr>
<tr>
<td>Moves to the Next Variable</td>
<td>SHIFT -&gt;</td>
</tr>
<tr>
<td>Backs Up to the Previous Variable</td>
<td>SHIFT &lt;</td>
</tr>
<tr>
<td>Help</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td>End</td>
</tr>
<tr>
<td></td>
<td>^ v Moves Pages. ESC Abandons</td>
</tr>
</tbody>
</table>
UPDATE DATA DICTIONARY

A series of screens are displayed that allow the user to interactively modify a Data Dictionary. The use of this option is similar to the CREATE DATA DICTIONARY option.

When the update function is selected the user is asked to enter the name of the Data Dictionary to update:

DATA DICTIONARY:

The Data Dictionary must reside in the directory of the currently active Data Dictionary, as specified by the DBASE environment variable (see APPENDIX B). The name must be four characters long.

IMPORTANT: Do not update the MI80 Data Dictionary, as the results may be unpredictable. If it does become damaged, reinstall it.

The screen clears and the following message appears:

Reading 58 variables

REDATAM starts loading the information of the selected Data Dictionary. When done, the following message appears:

Reading 58 variables 202 codes loaded. Press a key

The numbers, will of course change from case to case. Pressing any key brings up the following screen:

<table>
<thead>
<tr>
<th>REDATAM V2.0 Identification</th>
<th>Data Base: MI80</th>
<th>Date: 25/3/1987</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name of author</td>
<td>CELADE</td>
<td></td>
</tr>
<tr>
<td>Original file identification</td>
<td>Miranda 80</td>
<td></td>
</tr>
<tr>
<td>Creation date</td>
<td>29/12/1986</td>
<td></td>
</tr>
<tr>
<td>Last update</td>
<td>1/2/1987</td>
<td></td>
</tr>
<tr>
<td>Update author name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Record length</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td>Record types</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Record identification variable</td>
<td>IDENT</td>
<td></td>
</tr>
<tr>
<td>Initial position and length</td>
<td>8 1</td>
<td></td>
</tr>
</tbody>
</table>

Record Codes Description

1 Housing
2 Population
General information about the Data Dictionary is shown here.

Options:  The user can modify any of the existing information, move to subsequent screens, save the data entered, cancel all entries, and return to the DATA DICTIONARY MENU.

Usage: Any of the data fields may be modified. As shown on the bottom line, the up and down arrows position the cursor at the desired field.

See the CREATE DATA DICTIONARY function for an explanation of the meaning and content of these and all other data fields.

When done, pressing PgDn brings up the next screen. Pressing ALT-S saves to disk all the work done so far. ALT-D saves and exits the session. ALT-A cancels all entries and ends the session, leaving the original Data Dictionary unchanged. These, and additional features can be displayed by invoking help (?), which opens the following two windows:

**Description**

<table>
<thead>
<tr>
<th>Description</th>
<th>Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positions Cursor at Beginning of Field</td>
<td>Home</td>
</tr>
<tr>
<td>Positions Cursor at End of Field</td>
<td>End</td>
</tr>
<tr>
<td>Moves Cursor One Space to the Right</td>
<td>-&gt;</td>
</tr>
<tr>
<td>Backs Up Cursor One Space to the Left</td>
<td>&lt;-</td>
</tr>
<tr>
<td>Erases Character Under Cursor</td>
<td>Del</td>
</tr>
<tr>
<td>Erases Character to the Left of Cursor</td>
<td>BACKSPACE</td>
</tr>
<tr>
<td>Backs Up Cursor to Previous Word</td>
<td>Ctrl-A</td>
</tr>
<tr>
<td>Moves Cursor to Next Word</td>
<td>Ctrl-F</td>
</tr>
</tbody>
</table>

`More...`

```
*V Moves Pages. ESC Abandons
```

**Description**

<table>
<thead>
<tr>
<th>Description</th>
<th>Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moves Cursor to the Next Item</td>
<td>ENTER</td>
</tr>
<tr>
<td>Moves to the Next Screen level</td>
<td>PgDn</td>
</tr>
<tr>
<td>Moves Cursor to the Next Item</td>
<td>V</td>
</tr>
<tr>
<td>Backs Up Cursor to the Previous Item</td>
<td>^</td>
</tr>
<tr>
<td>Saves the Dictionary and Continues work</td>
<td>Alt S</td>
</tr>
<tr>
<td>Saves the Dictionary and Ends the Session</td>
<td>Alt D</td>
</tr>
<tr>
<td>Abandons the Session. No Save.</td>
<td>Alt A</td>
</tr>
<tr>
<td>Help</td>
<td>?</td>
</tr>
</tbody>
</table>

`End...`

```
*V Moves Pages. ESC Abandons
```
The next screen, accessed by pressing PgDn displays the following:

REDATAM V2.0 Documentation  Database: M180  Date: 25/3/1987

Dictionary Documentation:

This is the Miranda 1980 Census of Population database.

This database is used for the Tutorial and all examples of the REDATAM manual. It is a hierarchical database, containing housing as well as population records. The complete database contains about 1600 household records and 7000 population records.

Options: The user can modify the Data Dictionary textual description, move to subsequent or previous screens, save the data entered, cancel all entries, and return to the DATA DICTIONARY MENU.

Usage: This screen allows the user to modify descriptive text about the Data Dictionary. The text is not processed in any way. It is only displayed or printed when information about the Data Dictionary is requested.

Existing text may be modified or deleted. New text can be inserted or added. A dotted line, on which text can be entered, will appear when new lines are added. Up to a full screen of information may be entered. The right and left and up and down arrows can be used to position the cursor where desired. Further editing keys (and other features) can be viewed by invoking the help facility (?). The same windows as for the previous screen appear.

Pressing PgDn brings up the next screen while PgUp returns to the previous screen. Pressing ALT-S saves to disk all the work done so far. ALT-D saves and exits the session. ALT-A cancels all entries and ends the session.

The next screen, accessed by pressing PgDn displays the following:
REDATAM V2.0 Variables

Variable number: 6
Name: DWELLING
Description: Type of Dwelling
Data type (c/b): Compression type:
Original length: 1 Compress length: 0
Initial position: 12 Decimals: 0
Minimum range: 0 Maximum range: 9
Out of range: 0 Record type: 1

Code Description
1 Separate House
2 Flat\Apartment
3 Range\Barracks
4 Out Room
5 P\.[comm. Bld
6 Other Private
7 Group Dwelling

"v Move Cursor Shift< Previous Var Shift> Next Var Alt-A Abandon ? Help

The actual field content displayed varies depending on the variable selected.

Options: The user can modify existing variables and define new ones, move between variables, move to previous screens, save the data entered, cancel all entries, and return to the DATA DICTIONARY MENU.

Usage: This screen is used to modify the Data Dictionary variables. The values they can take, and descriptions thereof are also updated here.

See the CREATE DATA DICTIONARY function for an explanation of the meaning and content of these and all other data fields.

The up and down arrows are used to position the cursor at the desired field. The values and descriptions of the variable can also be modified. The screen "scrolls" if the bottom is reached.

Pressing Shift-right arrow or Shift-left arrow respectively shows the next or previous variable. With the use of these keys all variables can be corrected.
PgUp returns to the previous screen. ALT-S saves to disk all the work done so far. ALT-D saves and exits the session. ALT-A cancels all entries and ends the session. These and other features are available through the help windows (?). These are:
<table>
<thead>
<tr>
<th>Description</th>
<th>Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positions Cursor at Beginning of Field</td>
<td>Home</td>
</tr>
<tr>
<td>Positions Cursor at End of Field</td>
<td>End</td>
</tr>
<tr>
<td>Moves Cursor One Space to the Right</td>
<td>(\rightarrow)</td>
</tr>
<tr>
<td>Backs Up Cursor One Space to the Left</td>
<td>(&lt;)</td>
</tr>
<tr>
<td>Erases Character Under Cursor</td>
<td>Del</td>
</tr>
<tr>
<td>Erases Character to the Left of Cursor</td>
<td>BACKSPACE</td>
</tr>
<tr>
<td>Backs Up Cursor to Previous Word</td>
<td>Ctrl-A</td>
</tr>
<tr>
<td>Moves Cursor to Next Word</td>
<td>Ctrl-F</td>
</tr>
</tbody>
</table>

More...

\[^v\] Moves Pages, **ESC Abandons**

<table>
<thead>
<tr>
<th>Description</th>
<th>Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moves Cursor to the Next Item</td>
<td>ENTER</td>
</tr>
<tr>
<td>Backs Up to the Previous Screen Level</td>
<td>PgUp</td>
</tr>
<tr>
<td>Moves Cursor to the Next Item</td>
<td>^</td>
</tr>
<tr>
<td>Backs Up Cursor to the Previous Item</td>
<td>^</td>
</tr>
<tr>
<td>Saves the Dictionary and Continues Work</td>
<td>Alt S</td>
</tr>
<tr>
<td>Saves the Dictionary and Ends the Session</td>
<td>Alt D</td>
</tr>
<tr>
<td>Abandons the Session, No Save.</td>
<td>Alt A</td>
</tr>
</tbody>
</table>

More...

\[^v\] Moves Pages, **ESC Abandons**

<table>
<thead>
<tr>
<th>Description</th>
<th>Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moves to the Next Variable</td>
<td>SHIFT-&gt;</td>
</tr>
<tr>
<td>Backs Up to the Previous Variable</td>
<td>SHIFT&lt;-</td>
</tr>
<tr>
<td>Help</td>
<td>?</td>
</tr>
</tbody>
</table>

End...

\[^v\] Moves Pages, **ESC Abandons**
D.2. Database Management Functions:

The loading of the census data into the microcomputer will be done by national programmers. This is because some parts of the preparation and loading process are done on a large mainframe computer and will vary from census to census. The data is then transmitted via a communications link to the microcomputer. Once the data is in the microcomputer, REDATAM provides the necessary facilities to transform the transmitted files into a REDATAM Database. The mainframe file preparation is described in a separate reference document on the technical aspects of the REDATAM system.

The census data must be loaded into the microcomputer in the form of a single flat file, in ASCII form. Its records must be sorted according to the hierarchic geographic structure of country, with the records of the lowest subdivisions grouped together. Schematically, each record may have the following structure:

```
| Major Geo. | 2nd Geo. | ... | nth Geo. | Record Identification | Data ... |
| Division   | Division | ... | Division | Variable               |
```

All the above information must be provided, even though it may be in any position within the record.

Each record must contain the codes of all the geographic subdivisions to which it belongs. In addition, there must be a "Record Identification Variable", which is used to distinguish housing from population records in a hierarchical file. If more than one record type exists, the housing record must be immediately followed by all the population records that are associated to it.

The data must be supplied in character form. If necessary, REDATAM transforms it to the format indicated in the Data Dictionary. The file must be named "dbnm.", where dbnm is the four character name by which the REDATAM database will be known. This file may be down-loaded from another computer, created through the WRITE command (see section 5.2), or it could be generated directly on the microcomputer.

In addition, before generating the REDATAM database, the corresponding Data Dictionary has to be created, as well as the Geographic Names File.

The Data Dictionary is created automatically if the WRITE command is used to generate the "dbnm." file. Otherwise, section D.1 of this Appendix explains how to create or modify a Data Dictionary.

The Geographic Names File is named "dbnm.NOM" and includes the actual names of all the geographic subdivisions of the country.
It is a character ASCII file, where each record has the following structure:

11c1c2c3...cnlabel

11 is a 2 character numeric code identifying the geographic level to which the label belongs. For Miranda it is 01 for Province, 02 for Town and 03 for ED.

c1, c2, ... cn identify the geographic subdivision described by the label. In the case of Miranda, c1 is the code of the Province, c2 the code of the Town and c3 the code of the ED. The codes are numeric with the length specified in the Data Dictionary.

label is the actual name of the geographic subdivision.

For example, the Miranda Geographic Names File is named MI80.NOM and contains the following:

010100Florinda
020101Vilches
020199Rural
010200Perdita
020201Tumaco
020299Rural

Note that EDs are not named, so codes for them are not provided (only "11", "c1" and "c2" are required). Also note that 00 is used for the Town code of the Provinces.

The length of the names is limited by the space available on geographic selection screen.

In summary, the following files are required before proceeding to build the database:

- "dbnm."
- "dbnm.DIC"
- "dbnm.CTL"
- "dbnm.NOM"

The sorted census data.
The Data Dictionary.
A Data Dictionary control file that is built automatically when creating the dictionary.
The Geographic Names File.

These files must all reside in the directory specified by the DBASE environment variable (see APPENDIX B). The Data Management Functions are then used in order to generate the REDATAM database. The Data Management Functions are loaded from REDATAM's MAIN MENU. They are a menu-driven procedure, that is, the user is presented with screens offering a series of choices.

Very large databases (usually residing on a laser disk), may have to be generated in parts due to the 32 MB file size limit for DOS. In this case, the above files must be created several times
under different names, each one containing a different group of variables, but exactly the same records. Each file is converted into a separate REDATAM database with the processes that follow. The databases are then merged together as explained below. The user must assess whether this is necessary and if so, the number of variables to include in each database.
DATA MANAGEMENT OPTIONS

DATA MANAGEMENT MENU
Controls access to all Data Management Functions.

GEOGRAPHIC STRUCTURE GENERATION
Generates the Geography File that REDATAM uses for its processes. This step is not necessary if the data file was generated through the WRITE command.

AUTOLOAD
Creates the REDATAM database by generating the transposed data and index files.

APPEND VARIABLES
Appends new variables to an existing database from another database.
DATA MANAGEMENT MENU

When the Data Management Functions are chosen, the following menu will appear:

```
| REDATAM System 2.0       | Database: M180 Date: 02/02/1987 Time: 09:28:18 |
```

- Geographic Structure Generation
- Autoload
- Append Variables
- Exit

'v' Move cursor. Select with ENTER or first letter. ESC previous level

The active database and some other identification data is displayed at the top of the screen. Instructions are displayed at the bottom.

Options:

**Geographic Structure Generation:**
This option creates the Geography File from the active Data Dictionary. The GEOGRAPHIC STRUCTURE GENERATION option is loaded.

**Autoload:**
This option creates the transposed data and index files. The AUTOLOAD option is accessed.

**Append Variables:**
This option appends variables from another database. The APPEND VARIABLES option is used.

**Exit:**
This option returns to the MAIN MENU.

Usage:
An option is selected by either moving the cursor with the up and down arrows until it is highlighted, and then pressing ENTER, or by directly pressing the underlined letter of the desired option. ESC may also be used to return to the MAIN MENU.
The first step in loading the database (after building the appropriate Data Dictionary and loading the data file), is to create the Geographic Structure. The following screen appears:

```
REDATAM V2.0 Index Generation  Data Base: M180 Date: 25/03/1987

Level  Variable Name  Name Length

1

^ Cursor Up    v Cursor Down  PgDn Start Proc   Alt A Abandon
```

Some basic data is displayed at the top of the screen, including the active Data Dictionary. Instructions are displayed on the bottom line. Note that this process is not necessary when the database has been generated with the WRITE command.

**Options:**

The names of all the geographic variables are entered. The Geography File is then built and control is returned to the DATA MANAGEMENT MENU.

**Usage:**

The database being created must be active before proceeding. See APPENDIX B if this is not so. The name of the variables that identify each geographic level must be entered in hierarchical sequence. The up and down arrows are used in order to move from line to line. The following information is entered on each line:

- The name of the geographic variable.
- The length of the names or labels of the geographic area, as entered in the "dbnm.NOM" file.

The following files must be present before this process may be executed:

- "dbnm.DIC" The Data Dictionary.

When all geographic variables have been entered, the PgDn key initiates the process. The "dbnm.GEO" file is generated. Continue with the AUTOLOAD function in order to complete the process.

Pressing ALT-A cancels the process after asking the user to confirm his or her intention.
AUTOLOAD

This process creates the transposed data and index files for each of the variables. After this process is complete, the database is ready for use by REDATAM. The following screen is displayed:

REDATAM V2.0 Auto Load

Data Base M180 Date: 25/03/1987

Data Base Name

<- Cursor Left  -> Cursor Right  PgDn Start Proc Alt A  Abandon

Some basic data is displayed at the top of the screen. Instructions are displayed on the bottom line.

Options: The name of the database to be processed is entered. When the process is complete, control is returned to the DATA MANAGEMENT MENU.

Usage: The only information required is the name of the database to be generated. The following files must be present before processing begins:

- "dbnm."
  - The sorted census data.
- "dbnm.DIC"
  - The Data Dictionary.
- "dbnm.CTL"
  - The Data Dictionary control file that is built automatically when creating the dictionary.
- "dbnm.NOM"
  - The Geographic Names File.
- "dbnm.GEO"
  - The Geography File, generated by the previous process.

Note that if the "dbnm.NOM" file is not present, the ".NOM" file of the currently active database is used. If not available, an error message is displayed.

The process is started with PgDn. If the database to be generated is also the active one, a warning message is displayed, asking the user to confirm his or her intention to proceed.

Pressing ALT-A cancels the process.

At the end of this process, the following files will have been created:

- "dbnmiii.BIN" The transposed data files, one for
each variable. iii corresponds to the number of the variable in the Data Dictionary, right justified and padded with zeroes to the left.

- "dbnmjjjj.INX" The index file, one for each geographic variable. jjjj corresponds to the number of the variable in the Data Dictionary, right justified and padded with zeroes to the left.

- "dbnmPUNT.RED" A pointer file, used to link housing and population records.

The REDATAM database is now created and ready to be used. The "dbnm." file may be erased in order to save space. Make sure that you have a backup.

If the database is being built by parts, the AUTOLOAD process must be repeated for each sub-database. Then, the APPEND VARIABLES process must be executed.
This option appends new variables to an existing REDATAM database. Its main use is to consolidate several databases that were created separately in order to circumvent the 32 MB DOS file size limit. With this process, one REDATAM database is appended to another.

The Data Dictionary must have been created and the GEOGRAPHIC STRUCTURE GENERATION and AUTOLOAD processes must have been done for both databases, unless they were created with the WRITE command. In that case, only the AUTOLOAD process needs to be executed. Both databases must contain exactly the same records and must have the same hierarchical structure. The process is repeated for each database to be appended.

The following screen is displayed when this option is selected:

```
REDATAM V2.0 Append

Append to     Append from

Var Name    Variable Type    Variable Description

^ Previous Field    v Next Field    PgDn Start Proc    Alt A Abandon
```

Some basic data is displayed at the top of the screen. Instructions are displayed at the bottom.

**Options:** The names of the databases to merge is entered. When the process is complete, control is returned to the DATA MANAGEMENT MENU.

**Usage:** The names of the source and target databases must be provided. None of them may be active. The process is then started with PgDn.

If neither database is hierarchical, all non-geographic variables of the source database will be appended to the target database. If a variable already exists, it is ignored. Processing time increases however, so it is recommended not to repeat variables in sub-databases.

If one of the databases is hierarchical, the other one must be too for the process to work correctly. In this case, the name of each variable to be appended from the
source database is shown on the screen. The user may modify the record type of the variable, reassigning it to either a housing or population record. In this way the structure of the temporary hierarchical files created by the WRITE command may be modified to match the actual file structure.

After the process is complete, the target Data Dictionary will have been modified to include the new variables, and the transposed data files renamed accordingly. The source database and all its associated files are erased.
APPENDIX E

Specifications and Limitations

The REDATAM system has a few capacity limitations that may require some consideration if very large tabulations or complex command files are used. These are:

- Maximum number of continuation lines for a command: 20
- Maximum number of operands in an arithmetic expression: 64
- Maximum number of lines per command set: 128
- Maximum number of variables per command set: 128
- Maximum number of AVERAGEs per command set: 32
- Maximum number of COMPUTEs per command set: 64
- Maximum number of CROSSTABS per command set: 32
- Maximum number of FREQUENCIES per command set: 32
- Maximum number of RECODEs per command set: 64
- Maximum number of VALUE LABELs per command set: 128
- Maximum number of VAR LABELs per command set: 128
- Maximum CROSSTABS table size: 64K bytes

Also, due to DOS limitations, the combined size of all open files may not exceed 32 megabytes. Since each variable has a separate file (geographic variables have index files as well), this may limit the number of variables that may be employed in any one command set in large countries.

The following words are reserved, and may not be employed as variable names:

- AREABREAK
- AVERAGE
- BROWSE
- BY
- CLEAR
- COMMENT
- COMPUTE
- CROSSTABS
- DIRECTORY
- DRIVE
- ENDPHP
- EXIT
- FREQUENCIES
- GEOGRAPHY
- IF
- LIST
- LOAD
- NEW
- OPTION
- PRINT
PROCESS
RECODE
REVIEW
RUN NAME
SAVE
UNIVERSE
VALUE LABEL
VAR LABEL
VIEW
WEIGHT
WRITE
INDEX

, 4-5
.ECF suffix 4-53
.SEL suffix 4-65
Access REDATAM 3-3
AREABREAK 3-23, 4-9
AREABREAK RECODE 3-32, 4-11
Arithmetic expressions 4-17, 4-28, 4-30, 4-32, 4-33, 4-34, 4-36, 4-54, E-1
ASCII file 4-41
AUTOEXEC.BAT B-1
AVERAGE 3-23, 4-12, 4-59, 5-3, E-1
AVERAGE, Table 3-26, 4-13
Batch file 4-1, B-2, C-1
Beep 3-4
BROWSE 3-13, 4-14, 4-50, 5-15, D-4
CELDAE 1-1, 2-5
Census file B-6, D-21
Checklist 3-1
CIDA 1-1
CLEAR 3-13, 4-15
Collect 4-50, 5-15
Columns, Printing 4-12, 4-19
Comma 4-5, E-1
Command driven 4-3
Command editing 4-50, 5-13
Command files 4-39, 4-50, 4-53
Command list 4-6
Command set 4-40, E-1
Command syntax notation 4-4
Commands 3-11
Commands, Abbreviation 3-12, 4-4, 4-6
Commands, Clear 4-40
Commands, Collect 4-50, 5-15
Commands, Insert 4-50, 5-16
Commands, Length 4-5
Commands, Multiple lines 4-5
Commands, Save 4-50, 4-53, 5-15
Commands, Selection 4-28, 4-30, 4-32, 4-33, 4-34, 4-36, 4-54
COMMENT 3-33, 4-16
COMPUTE 3-32, 4-17, 4-55, 4-56, E-1
Computer system requirements A-1, 1-2
CONFIG.SYS B-1
CROSSTABS 3-13, 3-24, 3-34, 4-19, 4-45, 4-59, 5-3, E-1
CROSSTABS, Table 3-16, 3-27, 3-36, 4-20
Data Base, REDATAM 1-3
Data compression D-14
Data Dictionary 2-2, 3-2, 3-5, 4-71, 4-82, D-1, D-21
Data Dictionary, Browse D-4
Data Dictionary, Create D-10
INDEX

Data Dictionary, Erase D-20
Data Dictionary, Print 3-5, D-8
Data Dictionary, Update D-16
Database management 2-5, 4-72, 4-84, D-21
Database, Active 3-4
Database, Append Variables D-29
Database, Autoload D-27
Database, Geographic Structure Generation D-26
Database, Miranda 3-4
DBASE B-2, C-1, D-22
DBNAME B-2
DIRECTORY 4-21
Directory, Database B-1, D-22
Directory, Programs B-1
Directory, User 4-1, 4-39, 4-47, 4-53, B-1
Directory, Work 3-3, 4-1, 4-53, B-1
Down-loading 2-4, 4-41, 4-60, 5-8
ED 3-40
ENDHP 4-22, 5-4
Enumeration Districts 3-40
Error message 3-12, 4-3
ESC 3-4
Examples 3-1
EXIT 3-17, 4-24
Expressions, Arithmetic/Logical 4-17, 4-28, 4-30, 4-32, 4-33,
        4-34, 4-36, 4-54
Extract files 4-41, 4-60, 5-8
File, Census B-6, D-21
Files, Commands 4-39, 4-50, 4-53
Files, Directory 4-20
Files, Down-load 4-41, 4-60, 5-8
Files, Geographic Selection 4-62
Files, Hierarchical 5-2
Files, REDATAM 4-85, D-22, D-27, D-28
Florinda 3-40
FREQUENCIES 3-13, 4-25, 4-59, 5-3, E-1
FREQUENCIES, Table 3-15, 3-17, 4-25
Geographic variables 5-4
Geographic Selection 2-3, 3-8, 3-10, 3-20, 4-27, 4-62, 4-81
Geographic Selection File 3-7, 4-62, 4-64
Geographic Selection Processor 4-62
Geographic Selection, Definition 4-66
Geographic Selection, Partial 4-68, 3-9
Geographic Selection, Print 4-64
Geographic Selection, Recode 4-11, 4-68, 3-30, 3-37
Geographic Selection, Total 3-9, 4-68
Geographic variables 5-4, D-24
Geographic variables, Statistics with 4-4
GEOGRAPHY 3-12, 4-27
H-records 5-2
H-tally 5-3
Household variables 5-4
Head of Household 5-3
Help, Data Dictionary D-12, D-15, D-17, D-20
Help, Geographic Selection 4-69
Help, Statistical Processor 4-3, 4-50, 5-15
Help, Use 4-78
Hierarchical files 5-2, 5-3, 5-10, D-11, D-29
Hierarchical processing 4-22, 4-30, 4-34, 4-36, 4-44, 5-2
Hierarchical structure of Miranda 3-40
Hierarchical structure, Geography 4-67
HIGHEST 4-49
Household Records 5-2
Household Variables 5-4
Housing Records 5-2
IBM A-1
IBM compatibles, Equipment used A-1
IDRC 1-1
IF 3-24, 4-28
IF ... COUNT 4-30, 4-55, 4-56, 5-4
IF ... EXCLUDE 4-32
IF ... INCLUDE 3-34, 4-33, 4-54
IF ... QUALIFY 4-34, 5-4
IF ... SUM 4-36, 4-55, 4-56, 5-4
Include 4-50
Indeterminate H-tally table 5-5
Individual Records 5-2
Insert key 4-75
Insert mode 4-3
Installation, REDATAM B-1
Interface to Other Packages 2-4
Introduction 1-1, 3-1
Keyboard use 4-73
Keyboard use, Data Dictionary 4-82
Keyboard use, Data Management 4-84
Keyboard use, Editing 4-77
Keyboard use, General 4-74
Keyboard use, Geographic Selection 4-70, 4-81
Keyboard use, Help facility 4-78
Keyboard use, Menu screens 4-76
Keyboard use, Statistical Processor 4-3, 4-79
LANG B-2
Laser disk C-1
LIST 3-13, 4-38
List of commands 4-6
LOAD 3-38, 4-39, 5-16
Logical expressions 4-28, 4-30, 4-32, 4-33, 4-34, 4-36, 4-54
LOWEST 4-49
Manual, How to Use it 1-4
Memory, Clear 4-40
Menu, Data Dictionary 3-5, D-3
Menu, Data Dictionary Management D-9
Menu, Data Management D-25
Menu, Main 3-4, 4-2
MI80 3-4
Miranda 3-1
Miranda database 3-4
Miranda Geography and Data 3-40
Miranda Hierarchical Structure 3-40
Miranda, Map 3-41
NEW 3-39, 4-40
Notation, Command syntax 4-4
OPTION 4-41, 4-60
OPTION COUNT 4-44, 5-4
OPTION PERCENT 3-34, 4-45
Output, File 4-47, 4-60, 5-8
Output, Printer 4-47, 4-52
Output, Screen 4-47, 4-57
P-records 5-2
P-tally 5-3
Person variables 5-4
Perdita 3-40
Person Variables 5-4
Population Records 5-2
PRINT 4-46
Printer 3-14, 4-47
Printer, Width 4-12, 4-19
PROCESS 3-14, 3-25, 4-47, 4-57
PROG B-2
Province 3-40
RECODE 3-23, 3-32, 4-48, 4-55, 4-56, E-1
Recode (Geographic Selection) 4-68, 3-30
RECORD 3-14
Records, Population/Household D-11, D-14, 4-22, 4-44, 5-2
Records, Qualify 4-22, 4-34
REDA Datam 1-1, E-1
REDA Datam Database 1-3, D-21
REDA Datam, Capabilities 1-2, 2-1
REDA Datam, Files 4-85
REDA Datam, Installation B-1
REDA Datam, Languages B-2
REDA Datam, Load 1-5, 3-3, 4-1
REDA Datam, Modules 2-1
REDA Datam, Requirements 1-2, A-1, C-1
REDA Datam, Specifications and Limitations E-1
REDA Datam, Testing B-1
REDAOUTU.TMP 4-47, 4-57
REFERENCE 4-1
Replace 4-50
Requirements, Computer systems 1-2, A-1
Reserved words E-1
REVIEW 4-50, 5-13
Rows, Printing 4-12, 4-19
RUN NAME 3-35, 4-52
SAVE 3-35, 4-53, 5-13, 5-16
Screen, Clear 4-15
SL-MICRO 2-4, 4-41
SPECIAL FEATURES 5-1
SPSS-PC 2-4, 4-41
Statistical Processor 2-4, 3-11, 4-3, 4-79
Syntax, Command notation 4-4
Tables, AVERAGE 3-26, 4-13
Tables, by geographic area 4-9, 4-11
Tables, Crosstabs 3-16, 3-27, 3-36, 4-20, E-1
Tables, Frequencies 3-15, 3-17, 4-25
Tables, Percent 4-19, 4-45
Tables, Width 3-17, 4-12, 4-19
Tally level 5-3
Town 3-40
Tumaco 3-40
TUTORIAL 3-1
Tutorial Case 1 3-6
Tutorial Case 2 3-18
Tutorial Case 3 3-28
UNFPA 1-1
UNIVERSE 3-22, 4-33, 4-54
VALUE LABEL 3-23, 3-33, 4-55, E-1
VAR LABEL 3-23, 3-33, 4-56, E-1
Variable names 4-17, 4-30, D-14, E-1
Variables, Append 5-11, D-22, D-29
Variables, Browse 4-14
Variables, Environment B-2
Variables, Geographic 4-12, 4-17, 4-19, 4-25, 4-28, 4-30, 4-32, 4-48, 4-54, 5-4, 5-9, D-26
Variables, Integer/Fractional 4-17, 4-20
Variables, Labels 4-49, 4-55, 4-56
Variables, Population/Household 4-17, 4-49, D-14, 5-4
Variables, Range 4-17, 4-30, 4-55, D-14
Variables, User defined 4-17, 4-30, 4-36, 4-48, 4-55, 4-56
VIEW 4-57
Vilches 3-40
WEIGHT 4-59
WRITE 4-60, 5-8, D-21
Zero columns/rows 4-12, 4-19