HEALTH ANALYSIS AND INFORMATION SYSTEMS (AIS)
PAN AMERICAN HEALTH ORGANIZATION (PAHO)

Geographic Information System in Epidemiology and Public Health

SIGEPI

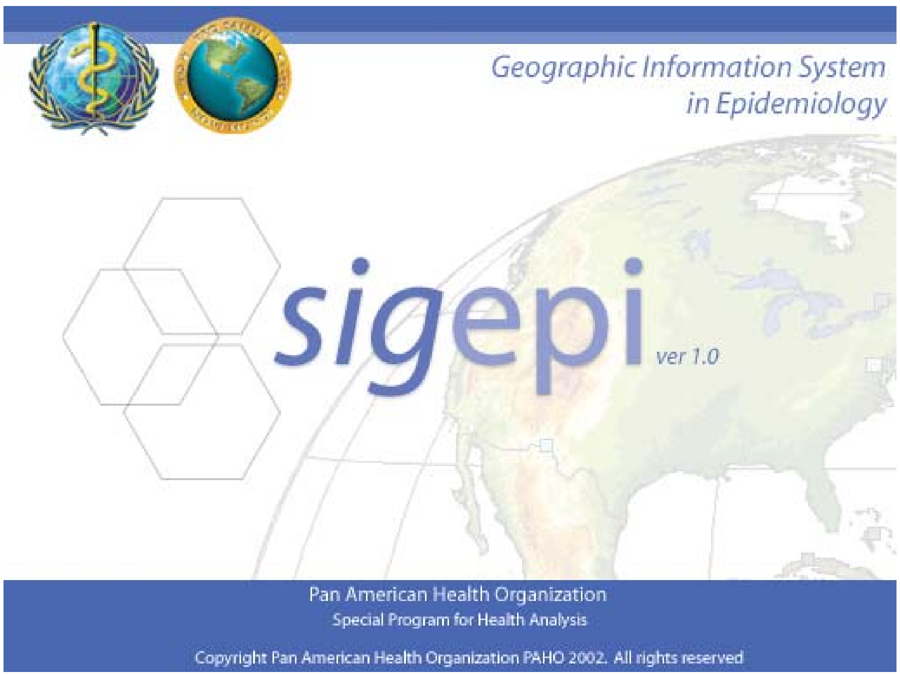
Version 1.26 March 2003

User's Manual

HEALTH ANALYSIS AND INFORMATION SYSTEM(AIS)
PAN AMERICAN HEALTH ORGANIZATION (PAHO)

GEOGRAPHIC INFORMATION SYSTEM IN EPIDEMIOLOGY AND PUBLIC HEALTH

User's Manual



Developed by:

Ramón Martínez Piedra, *Information and Technology Group, AIS/PAHO*
Manuel Vidaurre, *Informaiton and Technology Group, AIS/PAHO*
Katia Díaz, *Information and Technology Group*, AIS/*PAHO*
Dr. Enrique Loyola, *Analysis Group, AIS/PAHO*
Dr. Carlos Castillo-Salgado, Manager*, Área of Health Analyisis and Information Systems, AIS/PAHO*

Pan American Health Organization (PAHO).
525 23rd Street NW, Washington D.C., 20037
Telephone (202) 974 3327 • Fax (202) 974
E-mail: sha@paho.org

TABLE OF CONTENTS

INTRODUCTION AND INSTALLATION ... 9
SUMMARY ... 9
COMPONENTS OF SIGEPI ... 9
DISTRIBUTION ... 10
WHAT IS SIGEPI? ... 10
PRINCIPAL CHARACTERISTICS ... 10
FUNCTIONS OF SIGEPI ... 11
SIGEPI REQUIREMENTS AND INSTALLATION ... 15
Technical Requirements ... 15
Installation ... 15
Uninstalling SIGEpi ... 15
SIGEPI USER INTERFACE ... 16
STARTING TO WORK WITH SIGEPI ... 19
GETTING HELP IN SIGEPI ... 20
TYPES OF DATA IN SIGEPI ... 21
SPATIAL DATA ... 22
Vector Format ... 23
Raster Images ... 24
ATTRIBUTES TABLES ... 25
HOW TO OBTAIN DATA TO USE IN SIGEPI ... 26
ORGANIZING YOUR WORK IN A PROJECT ... 27
WHAT IS A PROJECT? ... 27
PROJECT WINDOW ... 28
WORKING WITH THE PROJECT ... 29
Creating a New Project ... 29
Opening an Existing Project ... 30
Rename, Copy, or Delete a Project ... 30
Saving the Project ... 31
Closing the Project ... 31
ACTIVATING A WINDOW ... 32
DELETING A WINDOW ... 32
PRINTING A WINDOW ... 32
CREATING AND MANAGING MAPS ... 33
WHAT IS A MAP? ... 33
Maps Window ... 33
WHAT IS A CARTOGRAPHIC LAYER? ... 34
CREATING AND WORKING WITH A MAP ... 35

TABLE OF CONTENTS

Adding Thematic Layers ... 37
Displaying Layers ... 37
Selecting and making active a Layer ... 38
Changing the position of the Layers ... 38
Defining the Properties of the Layers ... 38
Duplicating a Layer ... 39
Defining a Variable for ToolTip ... 40
Creating a Shapefile ... 41
Showing the Attributes Table of the Layer ... 41
Finding geographic units in the Layers ... 42
Select by Attributes ... 43
Select by Layer ... 45
Selecting all the geographic units of a Layer ... 47
Clearing a previous selection of features ... 47
Other Tools related to Layers ... 47
TOOLS FOR MAP MANAGEMENT ... 48
DEFINING THE PROPERTIES OF THE MAPS WINDOW ... 50
SPATIAL SELECTION PROPERTIES ... 51
THEMATIC MAPS ... 52
Creating a Thematic Map ... 52
Types of Thematic Maps ... 53
Thematic Map using Ranges ... 54
Thematic Map of Graduated Symbols ... 57
Thematic Map of Dot Density ... 58
Thematic Map using Bar Charts ... 61
Thematic Map using Pie Charts ... 63
Thematic Map using Single Values ... 65
Modifying a Thematic Map ... 67
Updating the Map Legend ... 67
Delete a Thematic Map ... 68
WORKING WITH LABELS ... 70
Saving a Label File ... 70
Showing a Label File ... 71
Clearing a Label File ... 71
DISPLAYING A MAP WINDOW ... 72
COPYING A MAP ... 72
WORKING WITH THE DATA TABLES AND MAP LAYERS ... 73
Adding Relationships of Data Tables to Layers ... 73
Converting a Relationship to a Shapefile ... 74
Removing Relationships ... 75
Updating the display of the Attributes Table ... 75
Copying the Attributes Table ... 75
Editing the values of an Attributes Table ... 75
DATABASE ... 77

TABLE OF CONTENTS

WHAT IS A DATABASE? ... 78
TABLES AND QUERIES ... 78
DATA SOURCES FOR TABLES ... 78
WORKING WITH THE DATABASE ... 80
Creating a new Database ... 80
Opening a Database ... 81
Closing the Database ... 82
MANAGING TABLES ... 82
Creating and Designing a Table ... 83
Importing and Linking External Tables ... 85
Displaying and Editing Tables ... 87
Creating an Index in a Table ... 88
Exporting Tables ... 89
Deleting a Table ... 90
MANAGING QUERIES ... 91
Creating and Designing a Query ... 91
Showing Queries ... 95
Deleting a Query ... 96
Exporting the Query to a Table ... 96

WORKING WITH GRAPHS ... 98

CREATING A GRAPH ... 99
EDITING A GRAPH ... 101
TYPES OF GRAPHS ... 102
Bar Chart ... 102
Stacked Bar Chart ... 102
Line Graph ... 103
Area Graph ... 104
Pie Chart ... 104
COPYING A GRAPH ... 104

ANALYTICAL PROCEDURES ... 106

DESCRIPTIVE STATISTICS ... 106
FREQUENCY DISTRIBUTION ... 108
CORRELATION ANALYSIS ... 110
SIMPLE AND MULTIPLE LINEAR REGRESSIONS ... 112
CALCULATION, STANDARDIZATION, AND SPATIAL SMOOTHING OF RATES ... 115
Calculation of Rates ... 117
Rate Standardization ... 120
DIRECT METHOD (DAR) ... 120
INDIRECT METHOD (IAR) ... 124
Smoothing of Rates ... 128
STANDARDIZED RATE RATIO SMOOTHER ... 130

TABLE OF CONTENTS

IDENTIFICATION OF PRIORITY AND CRITICAL AREAS 133
COMPOSITE INDEX IN HEALTH 137
SPATIAL ANALYSIS 140
PLOTTING POINTS FROM A TABLE 140
CREATING BUFFERS 142
EXPLORATORY SPATIAL DATA ANALYSIS 144
- *Outlier Map* *144*
- *Spatial Smoother* *147*
- *Spatial Lag Map* *150*
- *Global and local indexes of spatial autocorrelation* *152*

TIME-SPACE ASSOCIATION OF CASES (KNOX) 158
EXPOSURE-EFFECT ASSOCIATION 161

OBTAINING RESULTS 164

RESULTS WINDOW 164
DISPLAYING THE RESULTS SHEET 165
SELECTING AND COPYING RESULTS 166
CREATING AN HTML FILE 167
SELECTING ALL THE RESULTS 167
CLEARING AN EXISTING SELECTION 168

CREATING LAYOUTS 169

CREATING AND EDITING LAYOUTS 169
Adding Text to a Layout *171*
Adding a Map to a Layout *171*
Adding a Graph to a Layout *172*

TECHNICAL SUPPORT 174

CREDITS 175

GLOSSARY OF TERMS 177

BIBLIOGRAPHIC REFERENCES 180

Introduction and Installation

1

Summary

SIGEpi was created in response to the need for public health managers to have an adequate low-cost Geographic Information System to carry out epidemiological data analysis in health more efficiently.

SIGEpi offers simplified methods and procedures for public health analysis, including the functions of a Geographic Information System (**GIS**) as well as analytical, epidemiological, and statistical capacities. In this respect, it is the first software program to include spatial analysis methods and techniques focused towards health that are not available in frequently used commercial software programs.

SIGEpi was developed following several workshops and work with **PAHO** Collaborating Groups on the **"Application and Development of Geographic Information Systems in Epidemiology and Public Health"** Project, as well as collaboration with other health professionals and experts.

SIGEpi is intended for epidemiologists, health professionals, technical personnel, managers, administrators, and all personnel who make decisions and establish health policies at different levels: national, regional, community or local, who have limited access to commercial **GIS** and require simplified procedures and methods for data analysis.

Components of SIGEpi

SIGEpi is a **GIS** software program for epidemiology and public health made up of:

- **SIGEpi**: A component-based **Package of Executable Programs (System)** developed in Visual Basic that includes the use of tools and the components; Map Object, SQL, MS-Access, ActiveX control, HTMLBrowser, and HTMLHelp.

- **SIGEpi Help**: An on-line **Help System** that allows you to work with **SIGEpi**, plan and work with the System, design and work with a Project, and make use of the potential functions that it includes for data analysis.

- **SIGEpi Manual**: A **User's Manual** with sections for learning and training in the use of basic concepts and operation of **SIGEpi**, that allows you to develop your work and easily and expeditiously obtain results from data analysis and processing. This Manual supports the on-line help system.

Distribution

SIGEpi is stand-alone software, with its own functions that does not depend on another commercial GIS. It is distributed by the Area of Health Analysis and Information Systems of the Pan American Health Organization (AIS/PAHO) to the public health institutions of the member countries of PAHO through inter-institutional agreements of technical cooperation, with specific commitments of use and application in solving of health and decision-making problems. The AIS Area will keep a record of the users, as well as the uses and applications that the health institutions give to SIGEpi, and it will offer the technical support and advisory services required through technical cooperation agreements.

SIGEpi is distributed on a Compact Disk that contains the Package of Programs, the User's Manual, and the On-Line Help System. The initial installation it is under a Demonstrative License of Use of 30 days, period in which the user should register the product. For the registry, the user should contact AIS/PAHO at sigepi_info@paho.org.

The Pan American Health Organization (PAHO) wishes to thank Environmental System Research Institute (ESRI) for their generous contribution that allows PAHO to distribute SIGEpi without charge for use of ESRI® GIS software components required in SIGEpi.

What is SIGEpi?

SIGEpi is a Geographic Information System (GIS) designed mainly for use in Public Health and Epidemiology. It is a computer program that offers a set of techniques, procedures, and methods for spatial analysis in Epidemiology integrated in a GIS environment. It offers a basic set of functions for the management and display of geographically referenced data.

The element that distinguishes **SIGEpi** from other currently available GIS is that it offers a group of analytical procedures specifically for data analysis and decision-making in public health and epidemiology. Moreover, these functions are provided in a simplified way and a user-friendly environment to facilitate their use at various levels of decision-making in public health.

Health Analysis and Information System Area of the Pan American Health Organization (AIS/PAHO) have developed **SIGEpi** as part of the Technical Cooperation in Application and Development of Geographic Information Systems in Epidemiology and Public Health in the Americas.

Public health professionals and technicians, as well as researchers, academicians, students, and anyone who needs to analyze and visualize their data spatially can use SIGEpi**.**

Principal Characteristics

SIGEpi has been designed to present data in four types of windows: Maps, Tables, Graphs, and Results. It can handle as many windows of the Maps, Tables, and Graphs types as needed during a work session. Whenever one of these windows is activated, its options menu and corresponding toolbar are shown; these elements are part of the SIGEpi User Interface.

A Project Window makes it possible for you to have control over the different elements and components of the System, such as the Database and its Tables, Maps, Graphs and Results Window.

Functions of SIGEpi

The basic functions of **SIGEpi** are described briefly for easier understanding of the System's potential and special features.

The functions of **SIGEpi** include:

a) the generic GIS functions for storage, manipulation, and processing of data and spatial attributes,

- *Manipulation of geographic data:*

- *Display geographic data on a Map*

  Different geographic data or Thematic Layers can be displayed to create a Map. For example, geographic boundaries of countries, administrative boundaries of states, departments or provinces, municipalities, health areas, access routes, cities, and geographic characteristics such as types of vegetation, crops, etc. Each of these, with its own characteristics, superimposed in an orderly manner one on top of another, can be used to create a Map made up of different Thematic Layers, with good quality for spatial data analysis.

  See Creating and Working with a Map...

- *Creation of Thematic Maps*

  Attributes of Layers can be presented utilizing different colors and symbols on the Map through Thematic Maps. Thematic Maps can be created in a very simple way to show data values on a Map. The types of Thematic Maps that can be created are: Ranges, Graduated Symbols, Dot Density, Bar Charts, Pie Charts, and Single Values.

  See Thematic Maps ...

- *Spatial selection of geographic features*

  Geographic features from Thematic Layers can be selected utilizing several spatial selection tools and strategies. Spatial elements selected from a Layer can also be used to select geographic features from another Thematic Layer of the Map. The selected elements can be converted into a new Thematic Layer if desired.

  See Tools for management of Maps ... and *Tools for management of Layers ...*

- *Seek and identify features in a layer by attribute values*

  The user can seek elements of a Map knowing some values for the attributes of the Layers. In addition, with only a click on a layer feature, the user can obtain the values of its attributes.

  See Select using Layer ... and *Working with Labels ...*

- *Manage of non-geographic data:*

▶ *Manage, process, and show non-geographic data*

Non-spatial (non-geographical) data can be handled by the System Database. The native format of the Database is the *Microsoft Access* standard (.mdb); however, other popular file formats, such as dBASE, can be imported. Non-spatial data can be created, updated, saved, and processed.
See Database ...

Available data files generated by other systems are easily handled by **SIGEpi**. They are shown in the Tables or Queries windows. Also, they may be linked with Thematic Layers and presented as a Thematic Map.
See Linking the Table with attributes of the Thematic Layer ...

▶ *Selection of geographic features by their attributes*

A layer can be queried to show geographic features that meet certain conditions. For example: countries with a population of more than 25,000,000 inhabitants, or departments with an infant mortality rate higher than 45 per 1,000 live births. As a result, geographic features that meet the condition are selected.
See Select using attributes ...

- *Visualization of data (Charts and Statistics results):*

▶ *Create Graphs based on attributes of Layers*

The values for attributes of the Thematic Layers and data Tables can be presented as graphs.
See Working with Graphs ...

▶ *Obtain the Result Sheet from a data analysis*

The results generated by applying analytical processes for data are presented in the Results Window. This window functions as a document viewer in HTML format to present the information. During a work session with **SIGEpi** there is a single Results Window.
See Obtaining Results ...

▶ *Make a Layout*

The Layout Window makes it possible for you to prepare a page for printing that can contain texts, maps, and graphs, as well as tables. It generates a print-out identical to the one designed for the layout sheet.
See Obtaining Layouts ...

b) Specific procedures for Epidemiological Analyses

- *Basic set of exploratory data procedures:*

▶ *Descriptive Statistics*

For each attribute from the Thematic Layers and data in the Database you can calculate a set of statistical measures of central trend and dispersion.
See Analytical Procedures / Descriptive Statistics ...

- *Frequency Distributions*
  You can calculate the Frequency Distribution of the values of each attribute of the Thematic Layers and associated data. Each variable's Frequency Distribution is presented graphically and is also shown in the Results Window.
  See Analytical Procedures / Frequency Distribution ...

- *Correlation Analysis*
  This procedure calculates the correlation matrix for a set of selected variables or attributes. It makes it possible for you to determine the correlation among selected variables and helps determine the important variables to include in the study analysis.
  See Analytical Procedures / Correlation Analysis ...

- *Regression Analysis*
  This procedure makes it possible to apply simple and multiple regression between independent variables and a dependent variable. Several regression models can be applied. The results are presented in the Results Window; in the case of a simple linear regression, a regression graph is generated.
  See Analytical Procedures / Simple linear and multiple regression ...

- *Analytical, statistical, and epidemiological procedures necessary for data analyses and decision-making in Public Health and Epidemiology.*

- *Calculation, Standardization, and Smoothing of Rates*
  This procedures permit calculation of rates, both crude and specific, by population groups or strata defined by the user. The calculated rates can be standardized (adjusted) according to a criterion (such as age group, etc.), applying direct or indirect methods. The calculated and standardized rates are added as new columns in a new layer and are displayed in Thematic Maps.
  See Analytical Procedures/ Calculation, Standardization, and Spatial Smoothing of Rates ...

- *Identification of Critical Areas*
  This function allows you to identify the areas, regions, cities, or geographic units that meet the conditions or criteria established by the user. This function is very useful for identifying the geographic areas in which the worst health indicators converge. A selection of the set of geographic units that satisfy the defined condition is obtained and can be converted to a new Thematic Layer.
  See Analytical Procedures / Identification of Priority or Critical Areas ...

- *Construction of a Composite Health Index*
  This function allows you to calculate a Composite Health Index (SENDES/PAHO) by geographic unit. The calculation is based on a set of health indicators selected by the user. A calculated indicator is obtained in a new Thematic Layer and presented in a Thematic Map.
  See Analytical Procedures / Composite Health Index Calculation ...

- *Time-Space Association of Cases (Knox method)*
  This procedure applies the Knox method to determine the association of cases in time and space. Using a different approach than other software programs that have implemented this

function, SIGEpi takes the necessary information for the application of this method from the Map where the cases are to be plotted. It is required for this application that the attributes of the cases contain a variable with the date of the report or appearance of the symptoms. SIGEpi offers the capacity to plot cases on the Map from a case record Table that contains two columns with location values (latitude and longitude) of the case.
See Analytical Procedures / Space-Time Association of Cases ...

- *Exposure-Effect Association*
  This procedure offers a set of quantitative methods of analytical epidemiology at the individual level, making it possible to show the statistical significance of a possible association between exposure factors and the effect being studied. It should be possible to geographically represent the exposure on the Map, as well as the population sample being studied. This offers techniques that can be applied in unmatched cohort and case-control studies.
  See Analytical Procedures/ Exposure-Effect Association ...

c) exploratory spatial data analysis.

- *BoxMap*
  This procedure allows you to obtain a Thematic Map from the indicator under study utilizing the BoxMap method for analyzing and spatially representing the distribution of the variable's values, identifying the outliers of the indicator that may occur in some of the areas or geographic units of a region.
  See Spatial Analysis / Outlier Map ...

- *Spatial Smoothing of Rates and Proportions*
  This applies a procedure for spatial smoothing of the indicators' values that are stored as attributes of the Thematic Layer displayed on the Map, making it possible to identify with greater clarity the trends in the variable values by reducing the differences that can be due to local characteristics.
  See Spatial Analysis / Spatial Smoother ...

- *Spatial Lag Map*
  This allows you to create a Thematic Map from the indicator under study by applying the Spatial Lag procedure and identify the relationship that exists between the variable and estimated value of its neighbors in each geographic unit.
  See Spatial Analysis / Spatial Lag Map..

- *Global and local indexes of spatial autocorrelation*
  Geographic patterns in the data can be evaluated using measures of spatial autocorrelation linked to tests of statistical significance. The latter evaluate whether the spatial distribution of a variable's values follows a certain distribution due to chance or not. Indexes of spatial autocorrelation make it possible to detect the existence of statistically significant spatial patterns in health events, or detect spatial concentration of geographic units with similar indicator values.
  See Spatial Analysis / Spatial Autocorrelation ...

SIGEpi Requirements and Installation

SIGEpi is a stand-alone system programs that runs on Microsoft Windows 95/98/Me/NT/2000 operating systems. All its functions are integrated into the System itself so that it does not require other programs.

Technical Requirements

Recommended technical characteristics

Processor and Memory	*SIGEpi requires a 486 or Pentium Processor, or above, with 128 MB of Memory (RAM).*
Space available on disk	*The installation of the system requires total space of approximately 40 MB.*
Space for data	*It is important that space be available on the disk for the data and geographic bases.*

Installation

SIGEpi is distributed on a compact disk. It can be installed in the Microsoft Windows 95/98/Me/NT/2000 operating systems.

For installation the following steps are recommended:

1. Place the compact disk in the Compact Disk or DVD drive. If installing from 3 ½" disks, place the first disk of the distribution package in the disk drive.
2. Click the Windows 95/98/NT **Start** button.
3. Select **Settings/Control Panel**.
4. The Control Panel window appears. Double ***click*** the **Add/Remove Programs** icon; the program **Setup.exe** is selected automatically.
5. Click the **Next>** or **OK** buttons and follow the steps of the Installation program.

SIGEpi creates its own folder, which can be defined in one of the installation steps. The program will automatically be included among the options in the Programs Menu of the Windows operating system.

A set of components that you may already have installed on your computer is included in the **SIGEpi** installation package. If the component on your computer is more recent than the one which **SIGEpi** installs, a warning will appear, so that you may decide whether to keep the component that is already installed or copy the version included in **SIGEpi;** in this case we recommend that you keep the most recent version, in order to avoid affecting the operation of other programs that are already installed.

Uninstalling SIGEpi

SIGEpi *can be uninstalled with* Add/Remove Programs, *which appears in the Windows Control Panel. We recommend that you always uninstall it using this way. Deleting the contents of the*

***SIGEpi** installation folder should be avoided to prevent components copied to other folders on the hard disk from remaining.*

To uninstall SIGEpi, take the following steps:

1. Click the Windows 95/98/NT **Start** button.
2. Select the **Settings/Control Panel** option.
3. The Control Panel Window appears. Double ***click*** the **Add/Remove Programs** icon.
4. Select the option **SIGEpi** from the list of installed programs.
5. Click the **Remove** button and follow the steps for the uninstall process.

The System and all its components will automatically be uninstalled. A group of components installed by **SIGEpi** can be shared by other programs installed on your computer; therefore, we recommend that you take care during the uninstall process, where you are alerted to decide whether to uninstall a component shared by other applications.

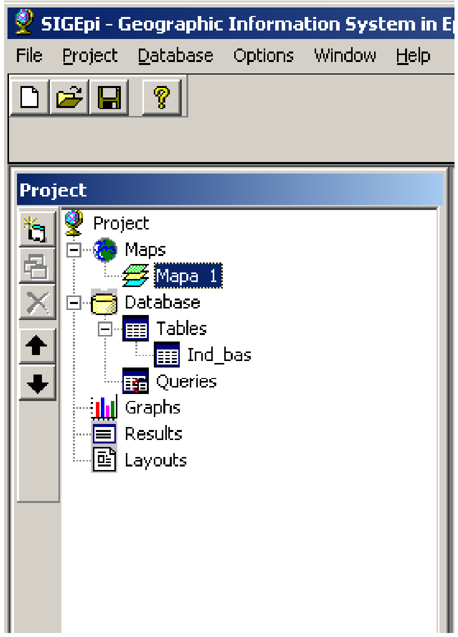
SIGEpi User Interface

SIGEpi is an application with a User Interface of multiple types of documents. In **SIGEpi** your data are displayed in the form of *Maps, Tables, Graphs, Results, and Layouts*. Each one of these forms for displaying data has its own characteristics and they are organized in a Project.

The System allows you to work with a single Project open at a time. The Project makes it possible to maintain all the components and elements of a work session together and organized. The Project can be saved in a file on disk and opened at a subsequent work session with **SIGEpi,** allowing you to keep all the components created during a work session without the need to repeat tasks that you have already carried out.

Project Window

The Project Window appears in the left part of the application's work area when a new Project is created or an existing Project is opened; it contains a list in a tree format where the different components of the Project are organized. The components are grouped under Maps, Tables, Queries, Graphs, Results, and Layouts. Each group contains elements of its type that have been created during a work session, distinguished by their names; each of these components can be modified. See Project



Project Window

Maps, Tables, Graphs, Results, and Layouts Windows

Each type of component is shown in its own window, for example:

- geographic data, geographic databases, and raster images are shown in the Maps Window,
- attributes of geographic data, Database Tables, and Database Queries are shown in the Tables Window,
- different types of graphs are shown in the Graphs Window,
- results of analytical and statistical processes are shown in the Results Window,
- presentation of data taken from the different types of windows that you want to display are shown in the Layouts Window.

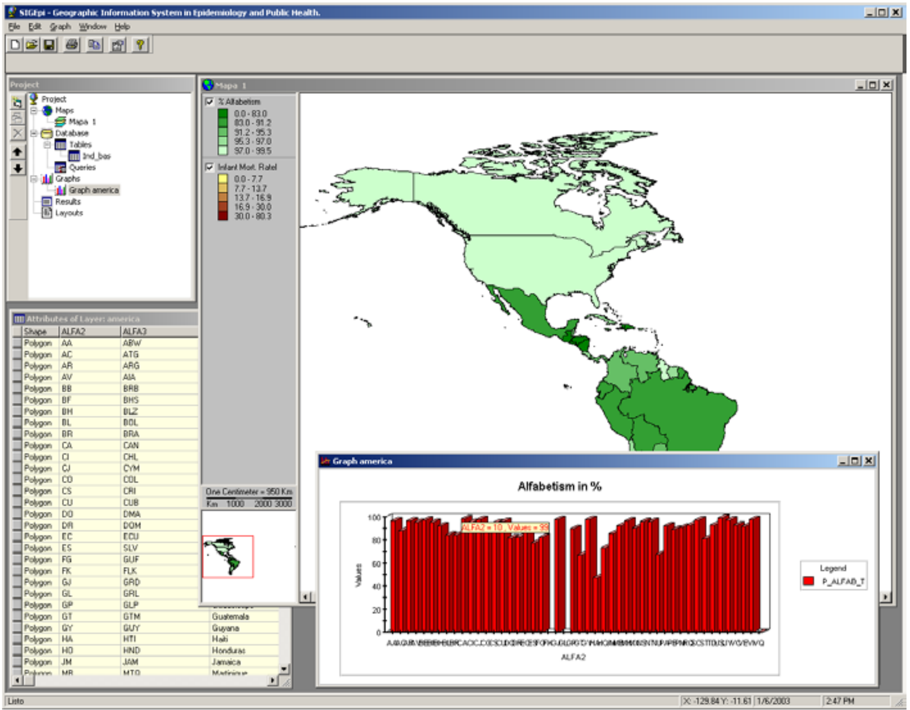
There can be as many open windows as needed, however, only one window is active at a time.

All the open windows will appear listed on the **Window Menu** and are activated by selecting them from the windows list or by ***clicking*** the title bar of the window that you want to activate.

The **active** window is always in front of the other windows, which you can see fully and work with, and for which you can modify the characteristics. When you activate a window, it moves to the front of all the other windows. Each window has its own functions, actions, and commands.

For example, when the Maps window is active you will see the buttons, tools, and menus that correspond to working with Maps.

NOTE: SIGEpi allows opening as many windows as it is needed, depending on the amount of memory (RAM) available in the computer. In beta testing stage of SIGEpi, it has been detected different non-wished effects when memory is overflow due to projects with many windows or map window with many large cartographic layers.



SIGEpi's Interface with several types of windows in the Project open at one time

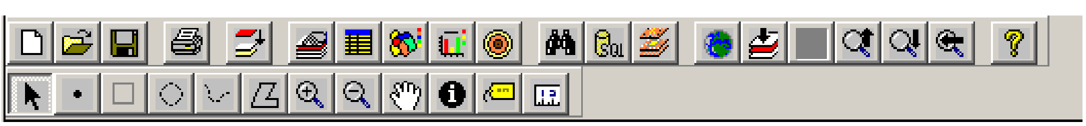
Menu Bar

The Menu Bar is always shown on the top line of the main application window. This bar shows the menus and options for the active window. By activating a different window than the one that is currently active will change the options in the Menu Bar.

Menu Bar of the Maps Window

Toolbar

The Toolbar appears under the Menu Bar. It is made up of two lines of buttons that simplify access to the menu options. The tool buttons are also related to the window types. As with the Menu Bar, a different toolbar appears when another type of window is activated.



Toolbar of the Maps Window

Starting to work with SIGEpi

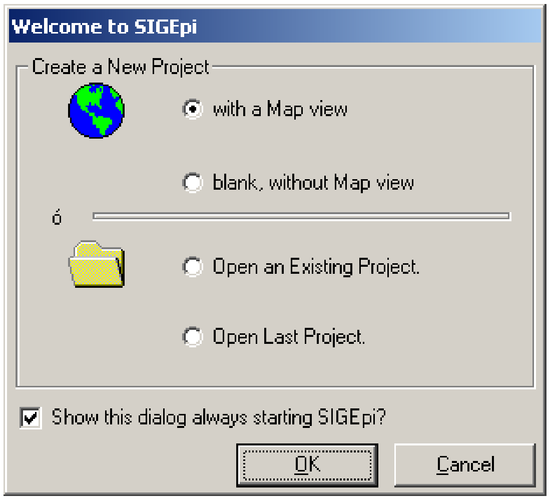
To begin to work with **SIGEpi** make sure the program is installed satisfactorily.
The process of starting to run **SIGEpi** is similar to that for any software program that runs on the Windows Operating System platform.

- **To run SIGEpi, take the following steps:**

1. ***Click*** the Windows 95/98/NT **Start** button.
2. Select the option **Programs/SIGEpi/SIGEpi**

A splash identification window is displayed followed by the principal **SIGEpi** window or Project Window.

To help you begin working with **SIGEpi**, a welcome window is displayed asking you to select the characteristics with which to begin work with **SIGEpi**.



Welcome Window

- **In this window you can select:**

1. Create a New Project
   a.) with a Map window
   b.) blank, without a Map window
2. Open an existing Project
3. Open the last Project.

If you select **Open the last Project**, this has the same effect as choosing **Open an Existing Project**, except that in this case the last Project with which you were working is selected, since **SIGEpi** saves the name and location of the Project with which you have been working whenever you: save the Project, exit the System, or close the current Project and create or open another Project.

Getting Help in SIGEpi

SIGEpi has an on-line Help System that allows you to obtain information on each step that you find yourself.

Each button, tool, or menu of options will tell you what it can do; just moving the cursor to the desired object shows you briefly its function (status bar).

To get Help on a window or Dialogue Box:

- Click on Help button in each dialog and you will get information about the process you are doing..

To display the contents of SIGEpi Help:

1. From the Help Menu on the Menu Bar, select Help ***Contents***.
2. ***Click*** the Content or Topic you want to display.

To search in Help according to the Topic Index:

1. From the **Help** Menu on the Menu Bar, select Help ***Contents***.
2. ***Click*** the ***Index***, and select the Topic for which you want to display the content.

To search in Help for one word in particular:

1. From the **Help** Menu on the Menu Bar, select Help ***Contents***.
2. ***Click Search***, and type the keyword for which you want to display the content.

To get Help on a phrase, subject, or word highlighted in the text

- ***Click*** the highlighted phrase, subject, or word.

Chapter 2

Types of data in SIGEpi

SIGEpi organizes and structures work with different types of data

SIGEpi allows you to work with: **Spatial Data**, geographically referenced data (Maps, Layers, Images), and

Non-spatial Data, non-geographic data (Databases, Tables, Graphs).

Data that describe the Earth's surface or the location of a part of it are called geographic data; this includes cartographic elements represented by graphics object, raster images, and their attributes.

Today, more than 80% of information can be represented geographically, using area codes, street addresses, locations, cities, states, and countries. **SIGEpi** allows working with geographic data and linking them the health data to bring the capability of visualizing and analyzing health variables in relation to the geographic space. This capability makes possible the application of statistics and epidemiological methods to the data.

Spatial Data

Spatial data is the core of a Geographic Information System. Spatial data is geographic data that describe the form and location of an object or feature on the Earth's surface, linking it with the attributes that characterize and describe it.

The formats most commonly used to represent spatial data are: vector and raster.

Data Attributes (non-spatial)

Data attributes, or non-spatial data, are those that we are already accustomed to working with; we usually have our data in a Table format such as dBASE, EXCEL, etc. These data are already available in existing information systems and databases that we already use.

There are various formats for managing and storing data. **SIGEpi** uses Microsoft Access (.mdb) as its native format for managing non-spatial data. It can also handle the following formats:

- *dBASE III*
- *dBASE IV*
- *FoxPro*
- *Excel*

Spatial Data

Spatial data store geometric or spatial locations of geographic characteristics of the Earth's different surfaces or areas, together with information on attributes or characteristic data that they represent. These are the graphic objects on the Map, such as political limits, territories, roads, river networks, location of health units, etc. Spatial data may contain an explicit geographic reference such as latitude and longitude, or an implicit reference such as an address or zip code.

Data on geometric locations are frequently stored in a raster or vector data structure, and they correspond to the attributes or data that they represent, which are stored in Tables of geographically related or georeferenced data, along with the characteristics that describe them. This is known as a georeferenced data structure.

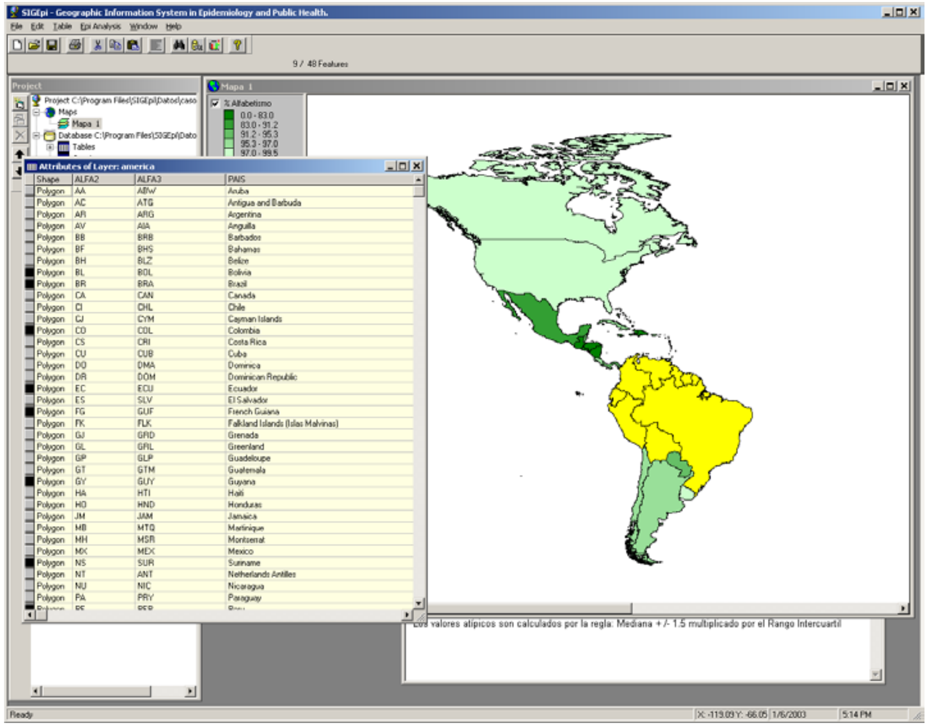
Below are some of the characteristics of the spatial and cartographic data

- In spatial data there is an explicit relationship between the geometric or geographic information and their attributes or characteristics; that is, both are always available and directly related when working with them. For example, if a particular area of a Maps Window is selected, **SIGEpi** automatically highlights or shows the record of the associated Table that contains the attributes of the area, if the Table has been opened in a Tables Window.

- Spatial data are georeferenced in order to find and determine locations or areas of the Earth's surface. To ensure that these locations are correctly specified, spatial data always use a specific system of coordinates, together with the size and projection of the Map. When spatial data is displayed on your screen, a particular scale for representation is used, just as when a map is shown on paper.

- The spatial data structure is the main characteristic of the cartographic data. It is designed with the capacity to manage and analyze the relation among geographic objects.

- Spatial data are organized in layers. For each geographic theme there is a layer containing the graphic representation and attributes of each geographic object. For example, country boundaries, roads, land elevations and rivers are each stored separately in layers that can be organized in a specific order in a map.



Maps window with a layer and its attribute table.

Vector Format

A vector data structure is used for spatial data representation. Common examples of vector data include digital maps and digital cartographic databases.

The vector format consists of a chain of coordinates and uses three types of graphic elements to represent the geographic objects of a Map: points, lines, and areas.

A point is represented by a pair of Cartesian X,Y coordinates, usually referenced geographically, for example: the reference of the real location on the surface of the Earth, of a city on a country-level Map, a case of disease in a city represented on a Map.

A line is represented by a set of pairs of X,Y coordinates linked sequentially to form segments of a line (topology), for example: it can represent a river, a street, a route between states.

An area is represented by a set of pairs of X,Y coordinates, linked in the form of lines; however, in this case the last pair of coordinates joins the first pair, creating a surface enclosed by a polygon. Examples of areas include the representation of states or countries on a Map or the soil types of a region. In the representation of lines and areas, each pair of coordinates forms a node, linked by lines that constitute the vector object. At present, files of coordinates obtained from Global Positioning System (GPS) receptors are being used more widely.

The vector format is commonly used to describe discrete entities, such as a City, State, or Municipality of a Country; it is less used to describe continuous entities (as is the case of the *raster* format), such as soil types, accesses to health care units, etc.

Spatial data file formats (Vector type) supported by SIGEpi

ESRI Shapefiles

(*.shp)
ESRI *Shapefiles* is a simple non-topological format for storing data in geometric locations and their attributes that describe the geographic characteristics of a given area.
The *shapefile* format defines the geometry and attributes of geographically referenced characteristics in approximately 5 files (.shp, .shx, .dbf, .sbh and .sbx, .ain and .aih)

ESRI Coverages

(*.adf, *.tat)
ESRI *Coverages* are topological data structures that represent geographic characteristics. The *coverage* format is designed for spatial analysis and applications that manage large volumes of geographic data. Defining the spatial topological characteristics in a coverage format can optimize data storage, reducing redundancy of coordinates and facilitating several spatial operations, such as an *overlay and adjacency* of geographic units or areas.
The ARC/INFO *coverage* format is one of the widely available formats for digital mapping and GIS applications.

CAD Drawings

(*.dwg, *.dxf)
CAD *drawings* is a data format obtained utilizing Computer Assisted Design (CAD) techniques; it is a standard of the CAD industry for the exchange and transfer of data between CAD and GIS applications. You can work on CAD *drawing* files to represent Layers in a Map just as with a *Shapefile format*.

VPF Data

(*.pft, *.lat, * aft, *. tft)
Working with VPF data is similar to working with the other data formats; **SIGEpi** allows you to work with data in the Vector Product Format (VPF), obtained from ArcView. VPF data can be added to a Map as Thematic Layers, Symbols, Relations, etc.

In any of these cases, you can work with these files in the Maps Window using the Work with Maps options.

Raster Images

An image is a graphic representation or description of an object that is obtained through an optical or electronic source. Common examples of an image data type are data obtained through remote sensing, such as satellite images, scanned data, and aerial photographs.

An image is a form of *raster* data where every grid-cell or pixel has a value depending on how the image was captured and what it represents. For example, if the image is a satellite image from a remote sensor, each pixel represents the light energy reflected from a part of the Earth's surface. However, if the image is a scanned document, each pixel represents a brilliance value associated with a particular part of the document.

Raster data are commonly, but not exclusively, used to store information on geographic characteristics that vary continuously on a surface or area. An image is a form of *raster* data in which every grid-cell or pixel stores a value obtained through an optical or electronic data source.

Raster data's resolution depends on the data source where they are obtained. The size of each part or area is adjusted, as when a zoom adjustment is made on a displayed Map, so that you see only the image of the part or area selected.

Raster Image Formats for spatial data that SIGEpi supports

GRID Data

(.hdr, *.adf)*
GRID data is an ESRI format that supports 32-bit integer and 32-bit floating point *raster grids*. GRIDs are specially designed to represent geographic phenomena that can vary continuously in space, and for spatial modeling and analysis of curves and surfaces, as in hydrology.

Standard Image Formats

(.bmp, *.dib, *.tif, *.jpg, *.jpeg, *.jff, *.bil, *.bip, *.bsq, *.gis, *.lan, *.rlc, *.sid, *.sun, *.rs, *.ras. *.svf, *.img, *.gif)*
The *standard* image formats display an image that shows the graphic representation or description of an object obtained through some of the most commonly available electronic or optical devices. Example: images obtained through scanners, photographs, etc.

SIGEpi is able to load and show on a map window all the raster image formats mentioned above, but it does not allow processing the attributes of each cell or pixel.

To add a raster image to a Map window, see Adding Layers to a Map.

To change the form in which the image is shown, modify the Properties of the Layers.

Attributes Tables

Attributes Tables describe and/or define characteristics of the object or feature they represent. Information stored in the Tables of a Database describes particular characteristics of the geographic unit.

Data attributes are the variables that characterize or are related to spatial data, such as population of a territory, number reported of cases of a disease, incidence or mortality rate of a disease, etc. There are different data sources useful in public health: census records, epidemiological surveillance records, medical statistics, health surveys, etc.

A Table is a set of elements that has a horizontal dimension (rows) and a vertical dimension (columns) integrated in a Relational Database System. A Table has a specific number of columns but can have N number of rows. A Table is also called a relation. The rows stored in a Table are equivalent to the records in any kind of data file.

The spatial data formats that we use in SIGEpi are similar to those in other GIS. The native format is ESRI Shapefile, however, it can manage other standard formats of cartographic and attributes data. See Tables and Queries.

Adding your data Table to a Map

In **SIGEpi** existing data tables can be joined to the attributes of layers. It is important to use a key field in both tables to relate them. This capability allows you to visualize spatial characteristics based on the data values in external tables. This also allows you to identify areas based on their characteristics.
See Working with the Attributes Table.

How to obtain data to use in SIGEpi

Data that comes with SIGEpi

The set of data that comes with SIGEpi has the purpose of being used to show the functions and analytical procedures implemented in SIGEpi, through Case Studies. This set of data is of public domain.
There is a large amount of sources of cartographic data; a lot of them are available through Intenet.
Some International Organizations are developing different projects and efforts to offer standardized cartographic data to facilitate the use of GIS Technology.

Some links to websites in Internet where it is possible to have access to cartographic data:

Geographic Network of Environmental System Research Institute (ESRI):	http://www.geographicnetwork.com/free.cfm
ESRI ArcData Online, GIS Data Online:	http://www.esri.com/data/online/browse.html
ESRI Packaged Geographic Data Sets – Global:	http://www.esri.com/data/online/datapacks/index.html
Center for Disease Control and Prevention of Atlanta (CDC), Shapefiles that are being distributed by EpiInfo/EpiMap 2000:	http://www.cdc.gov/epiinfo/EIshape.htm

National Sources of cartographic data

It is recommended to contact the national institutions in charge of census and/or cartographic data of each country with the intention of identifying national sources of cartographic data in digital formats.

Using data of other formats

Through this User's Manual you can identify the cartographic data formats that can handle SIGEpi. However, there are other cartographic formats very popular as MapInfo (*.map) that SIGEpi can not read directly, in this case it is recommended to use the MapInfo function Universal Translator to convert the MapInfo format to ESRI ShapeFile format.

Chapter 3

Organizing your work in a Project

The Project is SIGEpi's work area that organizes and supports your data

A **Project** in **SIGEpi** is a special type of file containing all the data and information related to a workspace or session using SIGEpi. It offers the ability of saving and recovering the status of a work session. The content of this file is generated by the System itself using XML language.

The fundamental advantage of the Project is to keep a set of Maps, Tables, Graphs, Results, and Layouts available to work with, avoiding the need to open and display each one every time it is needed.

The Project offers easy access to all your data organized in the different window types (interfaces), allowing you to manipulate and analyze it.

SIGEpi automatically creates a new Project when a work session begins and displays the Project Window; however, you can start a new project or open an existing one at any time.

It is recommended that you save the Project whenever you have made changes during a work session in order to save time and effort in future sessions.

Starting a Project begins by opening the Maps files.

What is a Project?

A Project is a type of file in which your work in **SIGEpi** is managed and subsequently stored or saved.

> **Project:** A Project contains all the views, windows, tables, graphs, maps, layouts, and results that have been created and with which you are working during an application with **SIGEpi.**

All of the resources you use and have available in the work area are a component of the Project.

For example:
If you are using **SIGEpi** to find possible areas or localities in which certain characteristics are found together, such as a low infant mortality rate, good educational level of mothers, etc. You can keep all the Windows with Maps, Thematic Layers, Tables, Graphs, and the Result sheet from analytical processes applied to your data and layouts in this Project. This way, you can continue analyzing and handling each kind of data appropriately. These can be stored when you conclude the Project work session. The next time you need to work with these data, opening the saved Project will make all the components available just as you left them. Project files have the extension .sml

What type of data does a Project contain?

A Project file does not contain the actual data that you use in **SIGEpi,** such as spatial data stored in Shapefiles and tables. A Project stores references to these data sources, so the same data can be used in several Projects without duplicating it.

What does the Project Window do?

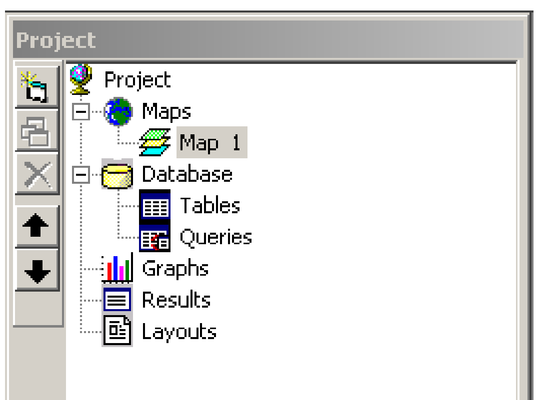
When you open a Project, all the Project components are listed or displayed in the Project Window. From the Project Window you can create new Project components, open or rename existing components, or remove or delete Project components. When the Project Window is active, its menu of options and buttons is available, allowing you to carry out operations and manage all its components.

How do you personalize the Project?

Each Project component in **SIGEpi** can be personalized by selecting it from the corresponding Menu Bar of the active window or double ***clicking*** the appropriate button, to define or modify the Properties of the component. The characteristics that can be modified depend on the type of component.

Project Window

Each Project in **SIGEpi** has a Project Window that is used to manage all the components, such as the **Maps, Tables, Graphs, Results, and Layouts** Windows.



Project Window

The **Project Window** has a list of buttons on the left side of the work area that represent the Project components; you can use these buttons to select the component you want to activate or display on the screen.

For example:

To see what Maps windows the Project contains, double ***click*** the **Maps** button and you are shown the list of all the Maps windows; to open one of them, double ***click*** the **name** of the Maps

Window you want to select. If you want to add new Layers to your Map, select the **Add Layers** option from the **Layers** menu on the Menu Bar, or ***click*** the **Add Layers** button that is available when the Maps Window is active. A dialog box will appear asking which Layer you want to add from the available geographic bases. You can choose to display or not to display the selected layers by ***clicking*** the **Map Legend** in each Layer selected; a check mark appears for each Layer displayed on the Map.

Every button in the **Project Window** represents a way of displaying and showing your data; each button represents a document or a user interface for handling and managing your information. The Project Window contains interfaces for every type of file that **SIGEpi** supports for the analysis of your data (Maps, Tables, Graphs, Results, and Layouts).

What does it mean when we say that several data analyses can be organized into a **Project** in **SIGEpi**? It means that you can work on several subjects of interest in health, organized into different Windows of **Maps, Tables, Graphs, Results, and Layouts**.

For example:

- An "Analysis of the leading causes of a cholera epidemic in several countries of the region" are represented in a Maps Window, which would correspond with the data in the associated Tables and their graphic representation, analytical procedures used, etc.,

- In another Maps Window you are analyzing the "Distribution of the incidence of tuberculosis in a region", linked to tables that reflect the characteristics and health patterns of the areas, epidemiological and/or statistical procedures applied, etc.,

- This way, there are as many components of each type as are needed for an efficient analysis with several alternative data sets, organized into a Project to support your work.

WARNING: Even though it is possible to have so many components as it is required during a work session with SIGEpi, it is suggested to be cautious mainly with the quantity of windows of maps opened simultaneously, there should be taken into account the size of the cartographic files that are being used and the availability of memory of the computer, in order to avoid effects unwanted during its use.

Working with the Project

Creating a New Project

When a Project is created in **SIGEpi**, a file is created that contains all the **Maps, Tables, Graphs, Results, and Layouts** components that are used to create or modify the Project.

- **To create a Project follow the steps:**

1. From the **File** menu, choose the option **New Project**
2. Or, from the Toolbar under the Menu Bar, click the **New Project** button

SIGEpi close all windows and components of the application, asking the user to save the current project. Afterwards the new empty project is created with no name. SIGEpi allows you to set a name to the project when the project is saved. The name of the Project can be changed by clicking the Properties option of the Project menu from the menu Bar.

When a work session has ended, the Project should be saved. To do this, select the option **Save Project** from the **File** menu or simply click the **Save Project** button. **SIGEpi** displays a dialog box that allows you to enter or change the Project name. If a Project name that already exists is selected, you are asked whether you want to replace it.

Opening an Existing Project

When you open a Project in **SIGEpi,** a file is opened that contains references to the geographic data, **Maps, Tables, Graphs, Results, and Layouts** that make up the **Project**.

To open an existing Project:

1. From the **File** menu, choose the option **Open Project**
2. Or, from the Toolbar under the Menu Bar, click the **Open Project** button

SIGEpi displays the **Open Project** dialog box, where you select or specify the name and location of the Project file you want to open.

As a result, all the commands stored in the Project are carried out, rebuilding all the components created during a previous work session.

Rename, Copy, or Delete a Project

When working with a Project file, you may want to assign a different name to your Project, make copies, or delete Projects that you no longer need.

With the Project Window active, select **Save Project As** from the **File** menu to save your Project with a new name.

You can also rename or delete a Project file using the commands and standard editing functions of the operating system.

To copy or rename a Project:

1. From the **File** menu, select the option **Save Project As**; you are shown a dialog box that allows you to assign another name to your Project and save it; this way, you have two Project files with different names and identical content.

To rename or delete a Project:

1. **Close** the Project; in a multi-user environment ensure that all users have closed the Project.
2. Use the commands or actions corresponding to your Operating System.

Saving the Project

In **SIGEpi** the work you have carried out can be saved; this will cause all the references to the components of the Project to be stored in a file.
When a component is activated, its window will display it in the same state as it was left in the last time you worked with it.

The Project can be saved at any time during the work session. When a Project is saved, all the work done up to that point is saved.

If the Project has not been saved, or if you have made changes after the last save, **SIGEpi** will ask you to save the Project when you open a new Project or close the existing one.

To save a Project:

1. From the **File** menu, choose the option **Save Project**
2. Or, from the Toolbar under the Menu Bar, choose the **Save Project** button

To save a existing Project with a different name:

1. Activate the Project Window
2. From the **File** menu, choose the option **Save Project As**...
3. In the **dialog box** that appears, select the folder and write the file name.

Closing the Project

During a work session with **SIGEpi** always a project exist. Remember that the visual appearance of the Project is the Project Window.

There is any menu option to close de Project. The ways the user can close a Project are: creating a new project, opening an existing project or exiting **SIGEpi**.

In any of those alternatives, **SIGEpi** will ask if you want to save the current Project

When a Project in **SIGEpi** is closed, all the windows associated with each component of the Project are closed. This closes all displayed windows, including the Project window. **SIGEpi** will set up a new Project where you can begin working.

If the **Exit** option of the **File** menu is selected, you are also asked if you want to save the Project; in case you click on yes, it is saved, and you exit the **SIGEpi** System. If you do not want to save the Project, the application still closes without saving your work.

Activating a Window

During a work session, windows are created for each type of Project component; they are organized in a list structured as a tree and displayed on the Project window. At any time while working with **SIGEpi** you can activate a window.

To activate a window:

1. Click the **name** of the Window, if it is open,
2. Or, select the **name** of the Window you want, from the **Window** menu on the Menu Bar,
3. Or, **double *click*** the **name** of the Window in the tree structure in the Project Window.

Deleting a Window

You may want to delete some of the Windows or components (Maps, Tables, Graphs, Results and Layouts) of the Project Window.

To delete a Project component:

1. Select the component you want to delete, ***clicking*** its window to activate it, or selecting it from the list of Project components displayed on the left side of your Project Window.
2. ***Click*** the **Delete Window** button on the Toolbar that appears to the left of the list of the Project's tree structure; note that depending on the component selected, certain actions are either activated or not.

Printing a Window

Each Window or component of the Project Window (Maps, Tables, Graphs, Results and Layouts) can be printed individually if you want.

- **To print a Window or Project component:**

1. Activate the Window that you want to print
2. From the **File** menu, choose the **Print** option
3. Or, click the **Print** button from the Toolbar
4. Follow the directions in the **dialog box** for printing, and define the characteristics or properties of the print-out that you want.

Chapter 4

Creating and Managing Maps

SIGEpi facilitates the management of your data and allows you to represent your data spatially

The Map is the principal cartographic product used in a GIS.

Maps can be classified by their content into **General Maps** and **Thematic Maps**

- ***General Maps***
  These are maps that represent the lines and borders of geographic information.
  For example: administrative borders of municipalities, regions, elevation, watersheds, communication routes, cities, etc. They are prepared with different techniques and by manual or computerized procedures ranging from on-site methods to remote methods with satellite images and Global Positioning Systems (GPS).

- ***Thematic Maps***
  These are maps that emphasize a particular aspect or theme, either natural or social.
  These maps represent information or data on a Map in the form of colors, shading, etc, displaying some characteristic of interest. For example: representing the incidence of a disease in an area.

What is a Map?

A Map is the basic data structure that can be used in **SIGEpi** for representation, analysis, and obtaining results from information supported by cartographic bases and tables of statistical, epidemiological, and health data. Maps show the distribution, location, size, and relationships of different natural and social phenomena in a region or area on the earth's surface.

Maps Window

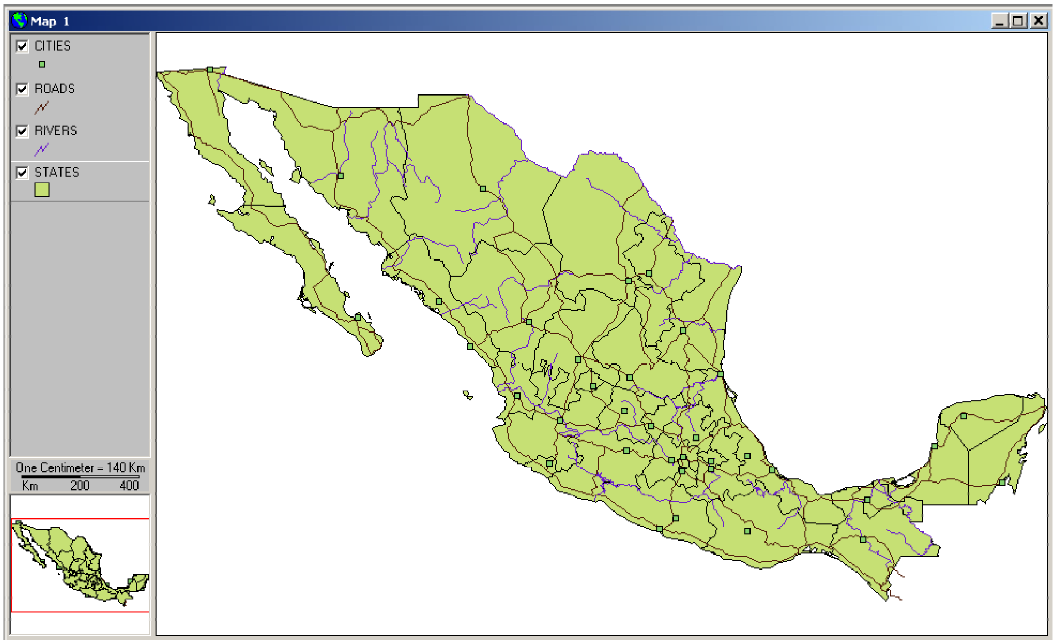
In **SIGEpi** all work with Maps is done in the Maps Window.

When displaying a Map, you can interactively show and manage geographic data on the right side of the Maps Window; that is, you select the geographic data to be used and how to display them.

The **Maps Window** also contains:

1. A **Legend** (located on the upper left)
2. The **Scale** (located on the center left), and
3. A **Localizer** (located on the lower left), which facilitates identification of the area presented on the Map.

You can select whether to display these components or not, choosing the options **Show Legend**, **Show Scale** and **Show Localizer** from **Map** menu.



Maps Window

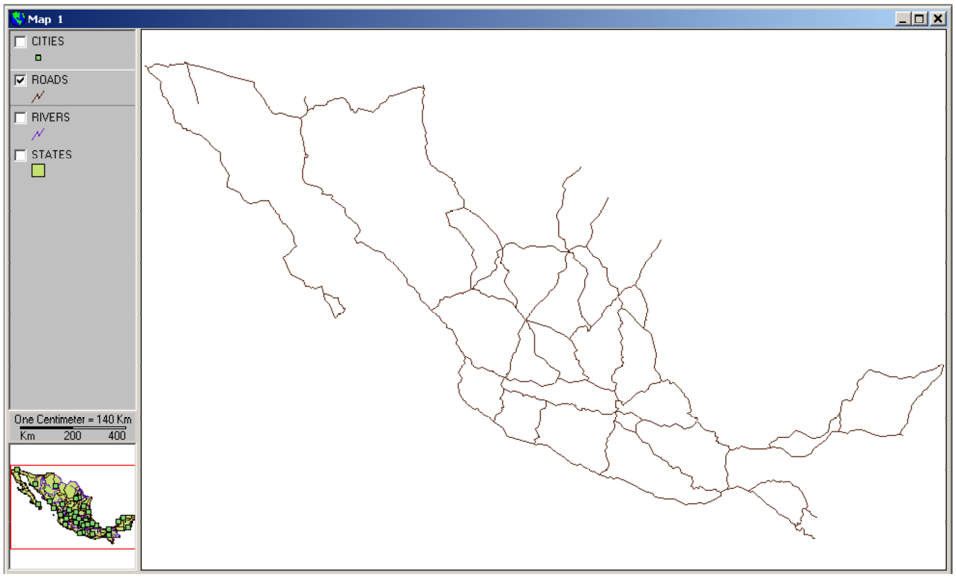
A Maps Window shows the geographic data that is used and managed for data analysis; however, it does not contain the data file itself, rather it references the data sources. This makes the Maps Window dynamic, because it reflects the current state of the data source; that is, if the data source is updated, its representation in the Project is updated automatically. For example: you can have a Map in your Project that shows the census of a city classified by population, and another Map of the same city that also shows groups not included in this census.

In a GIS, a Map is usually represented as a collection of cartographic layers.

What is a Cartographic Layer?

A Cartographic Layer represents a distinctive characteristic from the geographic data source.

For example: a Map displaying a country could have one layer for states, one for cities, one for communication routes, one for rivers, etc. The left side of the Maps Window lists the Layers contained in the Map, usually called Legend or Table of Content; from there you can choose whether to display each layer, organize them, or create Thematic Maps.



Maps Window displaying only the cartographic layer of main roads.

See Creating and Working with a Map.

Creating and Working with a Map

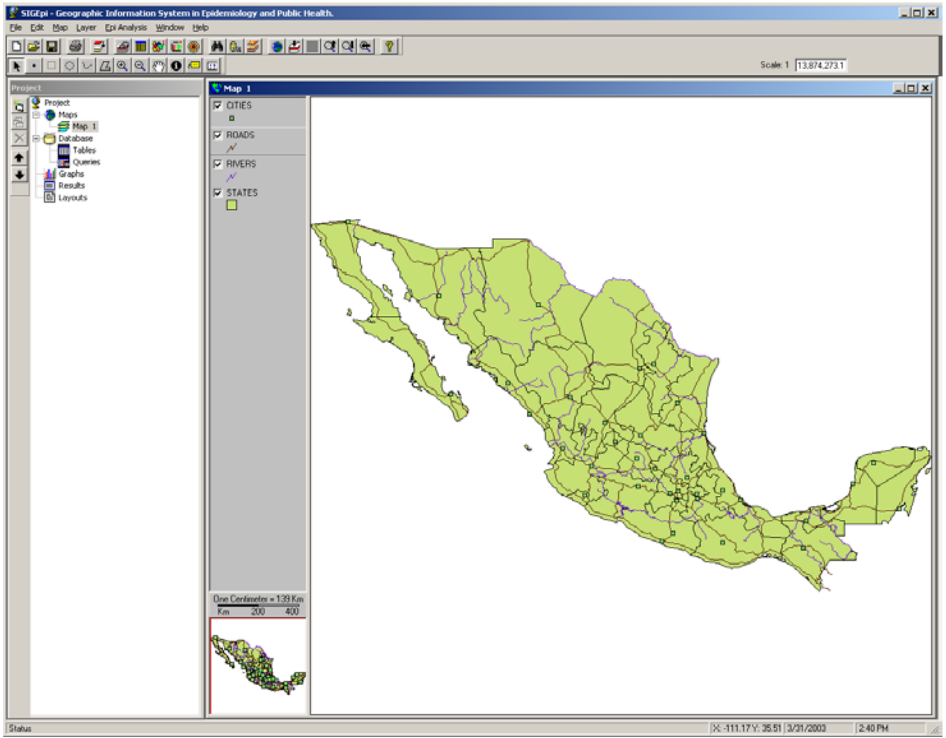
All work with **Maps** is done in the **Maps Window;** once the Project is opened, you can open a Maps Window by taking the following steps:

- **Open a Maps Window:**

1. The list of Project components appears on the left side of the Project Window; place your cursor on the **Map** component and ***click*** the **New Window** button of the Toolbar to the left of the Project components structure,
2. Similarly, once the **Map** component menu is selected, you can select the option **New Map** from the **Project** menu shown on the Menu Bar.

In either case, a new Maps Window is created and displayed, in which you can work to represent spatial data. See Adding Thematic Layers.

Note that when the **Maps Window** opens and activates, a **new menu** associated with the functions of this window is displayed in the **Menu Bar**. If you switch to the **Project Window, another menu** in the **Menu Bar** is activated. This menu and toolbar dynamism functions throughout the Project and for all the components of the Project during work with **SIGEpi**.



Maps Window and associated Menu Bars and Toolbars.

When a new Map window is created, it is named Map #, where # is a consecutive number for the Map window. Choosing the **Properties** option of the **Map** menu, you can change the Map's name, unit of measurement, and background color; to define the **Selection** properties, **Legend** characteristics**,** and **Scale.**

Show a Maps Window:

Once the Maps Window is created, you can select the Maps Window you want to display from the components structure of the Project Window, and:

1. **Double *click*** your selection
2. Or, ***click*** the **Show Window** button of the Toolbar to the left of the Project components structure
3. Or, select the option **Show Map** from the **Project** Menu of the Menu Bar.

In each case the selected Maps Window shown, activated, and placed in front of the other Windows of the Project.

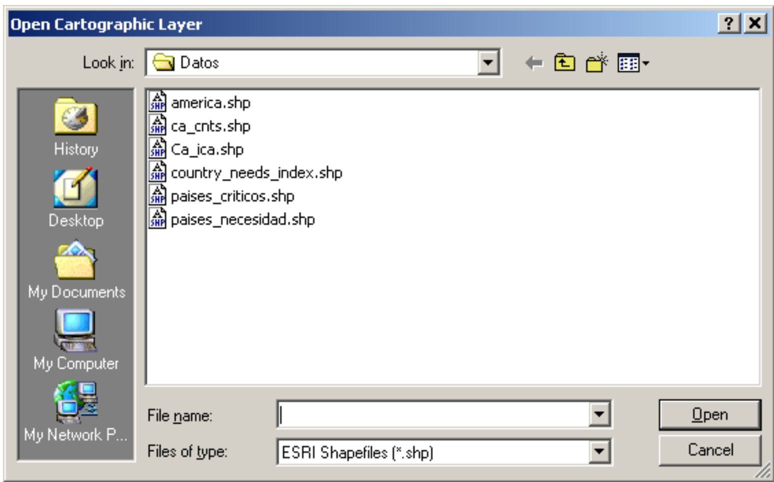
Adding Thematic Layers

Once the Maps Window is created, the first thing to begin displaying your spatial data is select and add the Thematic Layers of interest, depending on your cartographic bases.

- **Add a Thematic Layer to the Maps Window:**

1. Activate the Maps Window
2. From the **Layer** menu, choose the option **Add Layer...**,
3. Or, from the Toolbar under the Menu Bar, click the **Add Layer** button.

The **Open Layers** dialog box is displayed, which allows you to select the **location** and **name** of the spatial data files that you want to add to the Map. You can also select different spatial data formats that **SIGEpi** supports.



Dialog box. Open Cartographic Layers.

You can select one or more files at the same time. Just use the **Shift** and **Ctrl** keys following the procedures used in your Operating System. As a result, the selected **Cartographic Layers** are shown in the Maps Window. The Layers initially appear on the Map **Legend** and are initially visible. To learn how to hide and show a layer, see next topic Showing Cartographic Layers.

Displaying Layers

1. To make visible the **Cartographic Layer(s) click** the rectangle that appears to the left of the Layers' name in the **Legend** in the Maps Window.

This causes the layer to be displayed and a check mark ☑ to be shown.

You can display or hide each layer by ***clicking*** on the check box.

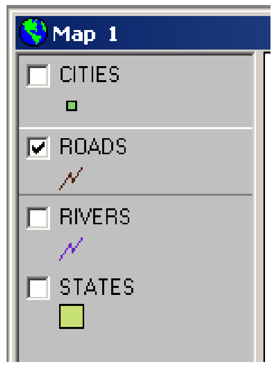
Only the layers that have a check mark ☑ in the Map Legend are displayed.

Selecting and making active a Layer

- **To select a Layer:**

1. Activate the Maps Window,
2. ***Click*** the Layer name in the map legend.

The selected Layer will appear with a highlighted rectangle in the Map Legend.



Map Legend. Selection of Layers.

As you can see in the example above, the ROADS has been selected and it is shown as highlighted among the Layers added to the Map Legend.

Changing the position of the Layers

The Layers added to the Map are placed top to bottom on the map, so the layer added last appears at the top. However, in order to have a clear representation of the map, you should change the positions of the layers.

It is important to take the geographic and geometric characteristics of each layer into account to define their position on the Map. For example, if we want to visualize in a map the boundaries of first administrative and political division of a country and the main cities of the country, the layer of cities should be at the top. This way, the overlapping of those layers on the map shows the desired result. Otherwise, we cannot see the cities on the map due to the administrative boundaries will cover the cities. The polygons (administrative boundary) hide the points (cities).

- **To change the order of Layers:**

1. Move the cursor to the **Legend** in the Maps Window
2. ***Click*** the layer you want to move and drag the cursor to the desired position.

This operation can be carried out as often as needed until you reach the order that best displays the Layers on the Map.

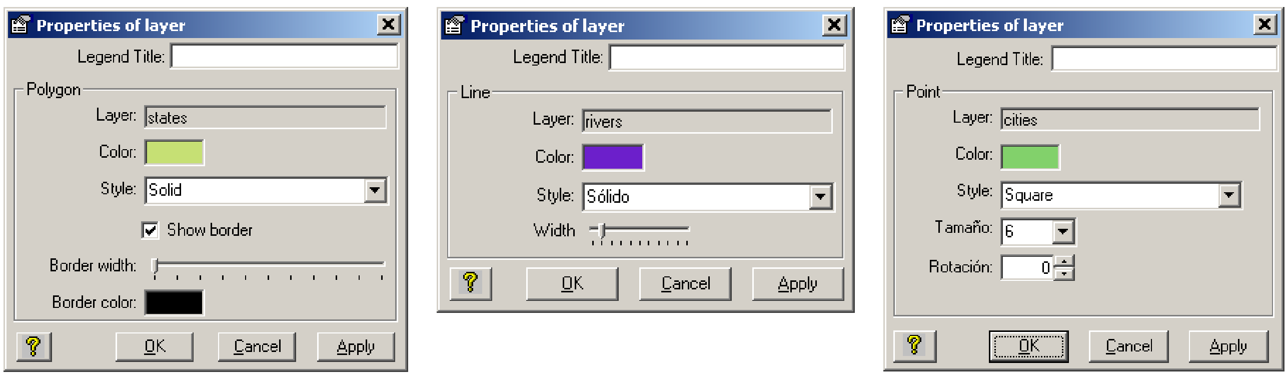
Defining the Properties of the Layers

To obtain a good-quality Map the properties of each layer in the map should be set according to the theme they are representing. Among layer properties there are: fill color, fill pattern, line thickness, types of symbols, etc.

To modify the characteristics or Properties of a Layer:

1. Select the **Thematic Layer** that you want to modify the **Properties** by ***clicking*** the name of the Layer in the **Legend**,
2. Choose the option **Properties**...from the **Layer** menu on the Menu Bar,
3. Or, choose the option **Properties**...from the menu displayed by **right *clicking*** the name of the selected Layer,
4. Or**, *click*** the **Layer Properties** button.

Dialogue Box is displayed, which allows you to modify the Properties of the layer. Each layer has it own dialogue box of properties, depending on the type of features it has: points, lines, or polygons.



Dialogue boxes. Layer Properties. Layers of polygons, lines, and points.

In addition, if the layer is a *raster* image, its Dialogue Box will include special characteristics for this type of layer.

Duplicating a Layer

You can duplicate any Layer in a Map. This way you can create more than one Thematic Maps for the same cartographic file, conducting data analysis with various parameters. You can change the Legend displayed for each Layer.

To duplicate a Layer:

1. Select the Layer that you want to duplicate,
2. Select the option **Duplicate Layer**, from the **Layer** menu on the Menu Bar or from the menu displayed by **right *clicking*** the Layer,
3. In case the duplicated layer is not visible, ***click*** on the check mark box that appears to the left of the layer's name in the **Legend**.

You can turn the display of a Layer on or off by successive ***clicks*** on this rectangle.

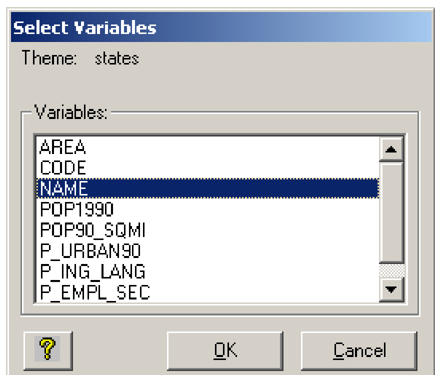
Defining a Variable for ToolTip

Selecting a **Variable for ToolTip** allows you to display the content of a variable or attribute as a label when you select a point in the active Thematic Layer of the Maps Window.

To select a Variable ToolTip:

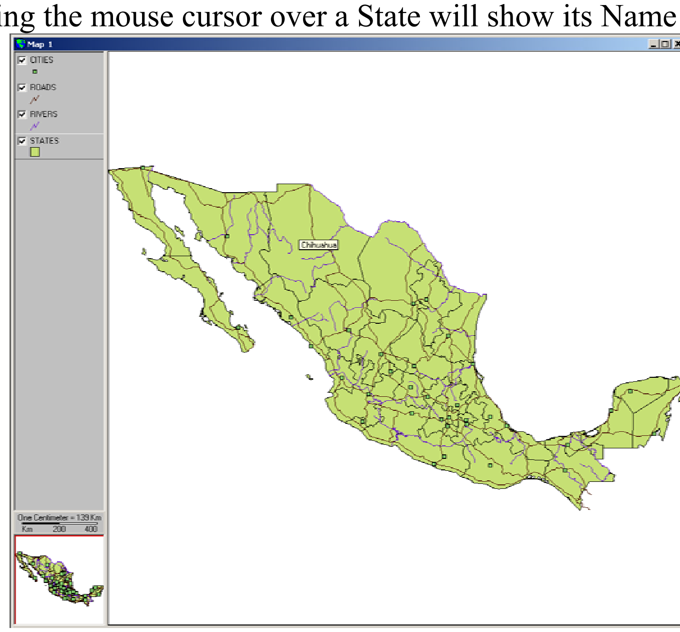
1. Activate the Maps Window,
2. Select the Layer of interest,
3. *Click* on the option **ToolTip Variable**... from the **Layer** menu on the Menu Bar or from the menu displayed by **right *clicking*** the selected Layer.

You will see the **Dialogue Box** that allows you to select a variable or attribute that it will be used as a Tooltip when the mouse cursor moves over a feature on the map.



Dialogue box. Select variable.

In this case the variable **NAME** was selected from the layer **STATES** that is the State boundaries in Mexico. Now moving the mouse cursor over a State will show its Name as a Tooltip.



Example. Maps Window and Attributes Table associated with the Active Layer.

The State's name, where the cursor is located, is displayed. In this case is Chiguagua.

Creating a Shapefile

SIGEpi has the capacity to create new cartographic files (Shapefiles) from existing layers.

For example:
You are working with a layer of municipality boundaries, and the municipalities that show the greatest health problems according to some health indicators are selected. We want to save this spatial selection as a new layer that can later be used in other work sessions.

Converting a Layer to a Shapefile:

1. Select the desired Layer in the active Maps Window,
2. Select the option **Convert to Shapefile ...,** from the **Layer** menu, or from the menu displayed by **right *clicking*** the selected Layer,
3. Write a name for the new Shapefile in the displayed dialog box.

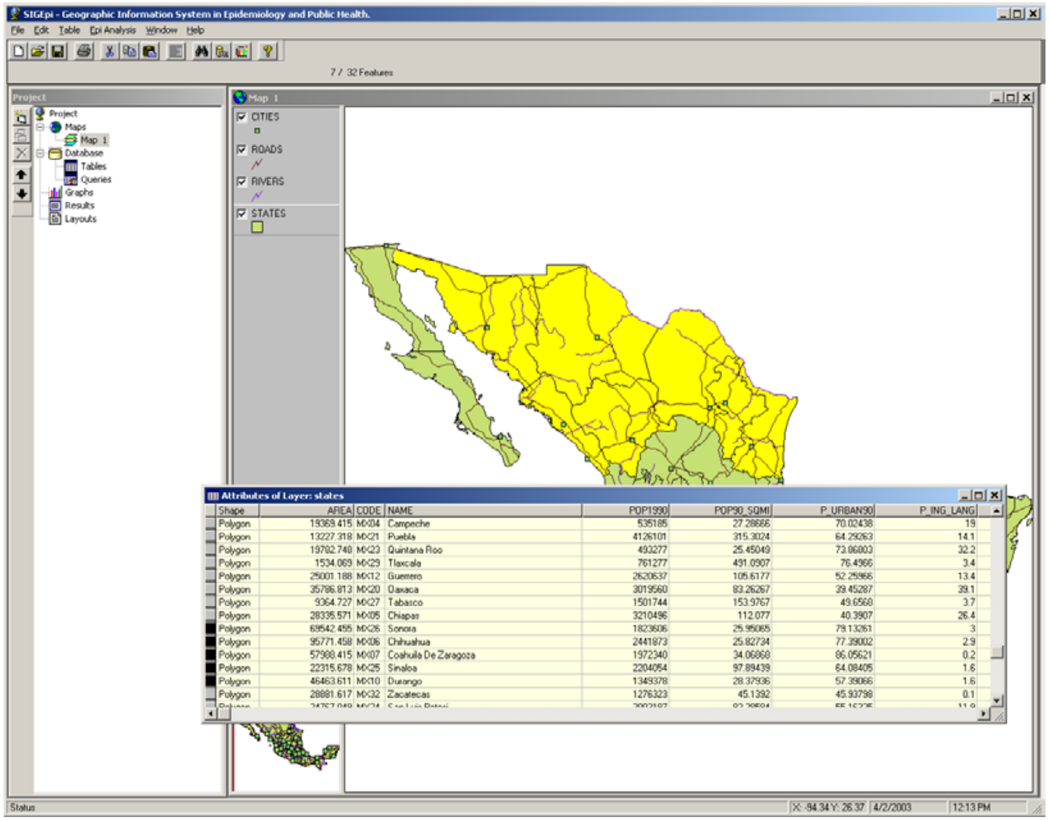
Once the Shapefile is saved, **SIGEpi** asks if you want to add this new Layer to your Map; this gives you the ability to continue to work with the current Map and the option to display the selection as a new Layer in the Map.

Showing the Attributes Table of the Layer

Each Thematic Layer displayed in the Maps Window has the property of being a spatial datum; as such, it has an associated Attributes Table that contains information that defines or characterizes the Layer.

To show the Attributes Table of a Layer:

1. Activate the Maps Window,
2. Select the desired layer,
3. Select the option **Open Attributes Table**, from the **Layer** menu, or from the menu displayed by **right *clicking*** on the selected Layer,
4. Or, ***click*** the **Attributes Table** button that is found on the Toolbar displayed under the Menu Bar.



Maps Window. Attributes table of the layer States.

A selection of a States is shown on the Map and this selection is shown in the attributes table too.

A dynamic relationship exists between the Layers and Attributes Tables; this means that selections made on the Map window are reflected in the Attributes Table window and vice versa. Under certain circumstances, i.e. the number of geographic units and the number of attributes are very large, the attributes table will not display. This will depend on the amount of memory available.

Finding geographic units in the Layers

Working with Maps window you may need **to find** specific geographic units based on the **values** of layer attributes. Example, locate all the geographic units that meet a characteristic.

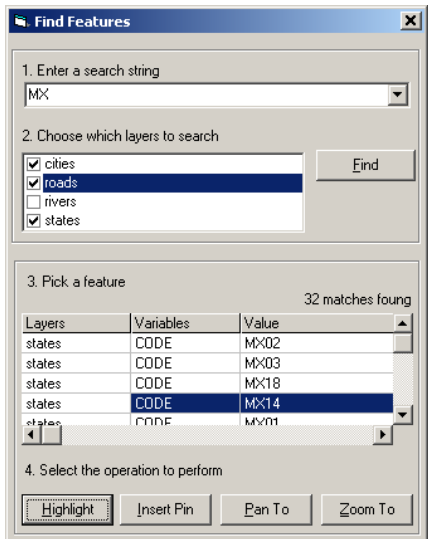
To find geographic units on a Map:

1. Activate the Maps Window (Does not require a Thematic Layer to be selected),

2. Select the option **Find...,**
   a.) from the **Layer** menu,
   b.) from the menu displayed by **right *clicking*** on any layer name of the Legend,
   c.) from the **Table** menu when is active the Attributes Table Window.
3. Or, ***click*** the **Find** button on the Toolbar displayed under the Menu Bar.
4. In the displayed **dialog box**, type the value (content of the attribute) you want to find

5. **Select** the Layer(s) you want included in the selection by ***clicking*** the rectangle that appears to the left of the name of each Layer.
6. ***Click*** the **Find** button; as a result, an edit box displays a report of geographic units that satisfy the condition of the completed search,
7. **Select** the geographic units of interest by clicking them in the corresponding edit box,
8. Choose the way you want the selection shown on the Map, by ***clicking*** the buttons **Highlight, Insert Pin, Center on, Expand to**. This helps you identify the geographic unit selected on the Map.

You can select other geographic units to find their geographic locations, and continue displaying them on the Map, through the actions described above.



Dialogue box. Find Geographic Units on the Map.

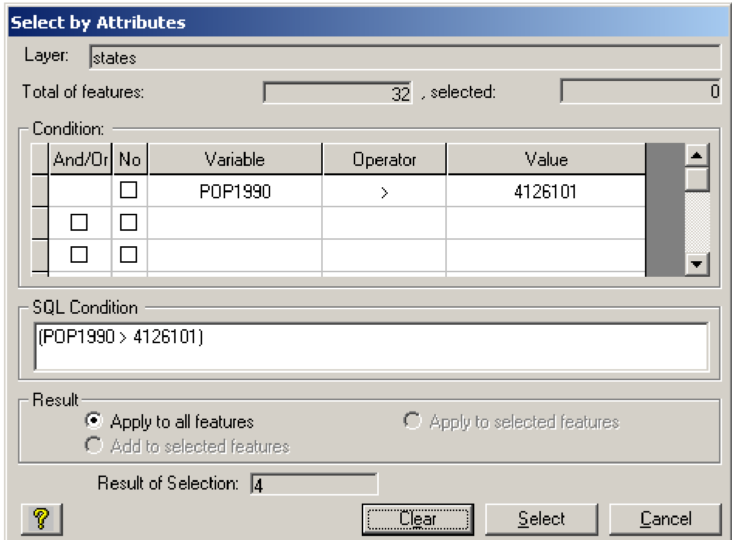
Select by Attributes

A usual way to select geographic features in a layer is using attribute values.

To select geographic features from a layer in a Map based on values of attributes:

1. Activate the Maps Window,
2. Select a Layer, and select the option **Select by Attributes...,**
   a.) from the **Layer** menu,
   b.) from the menu displayed by **right *clicking*** on the layer name in the Legend,
   c.) from the **Table** menu when the attribute table window is active.
3. Or, ***click*** the **Select by Attributes** button from the Toolbar located under the Menu Bar,

The dialog Select by Attributes is displayed, in this case:

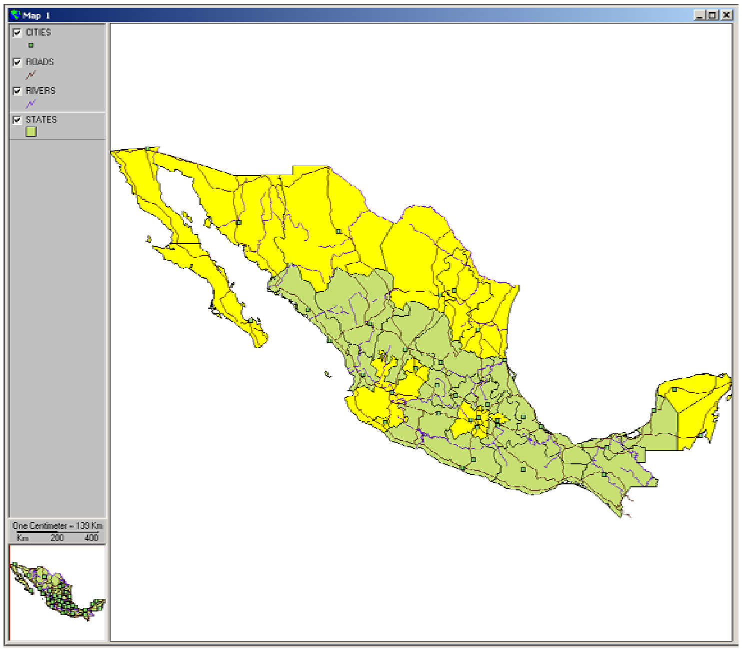


Dialogue box. Select by Attributes.

1. The **name** of the selected layer, number of features and selected features are displayed at the top of the dialogue.
2. In the **Condition** edit box, define the **conditions** you want. To do this, ***click*** the grid-cells of each column shown to select the **Variable**, **Operator**, and **Value** for building the condition expression. The conditional expression could be simple or complex. In case of creating complex condition use de logical operators AND, OR. The logical operator NOT is used to negate a expression or part of a expression.
3. The **conditional expression** is editable using the edition box "SQL Condition". Ensure to write a valid conditional expression if you decide to edit it.
4. ***Click*** the scope of the selection clicking on one given alternative in **Result** frame.
5. ***Click*** the **Clear** button if you want to delete the defined condition,
6. ***Click*** the **Select** button if you want to select the geographic units that meet the defined condition.

Example:

A Maps window shows a selection of States that meet a condition. The States that are selected meet the condition that the Area is more than 3,000 square miles and urban population is 60,000 inhabitants and over.



Maps Window. States selected by a condition.

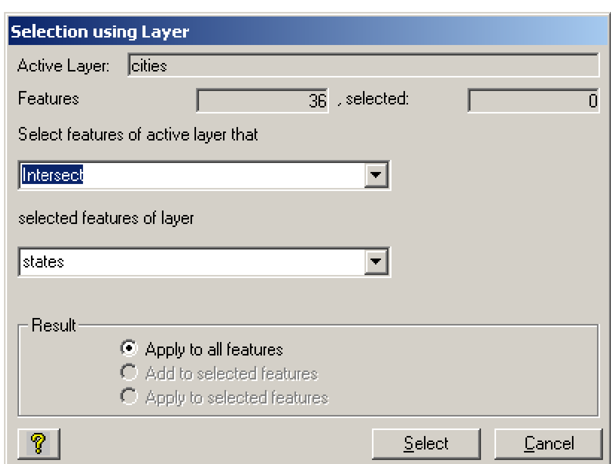
Select by Layer

In many cases it is necessary to select geographic units in a **Layer** based on a spatial **relationship** with geographic units of another **Layer**.

Selecting geographic units for a Layer based on another Layer:

1. Activate the Maps Window, and Select a layer,
2. Select the option **Select by Layer...,** from the **Layer** menu, or from the menu displayed by **right *clicking*** on the layer name in Legend,
3. Or, ***click*** the **Select by Layer** button from the Toolbar,

The dialogue **Select by layer** is displayed:



Dialogue Selection using Layer.

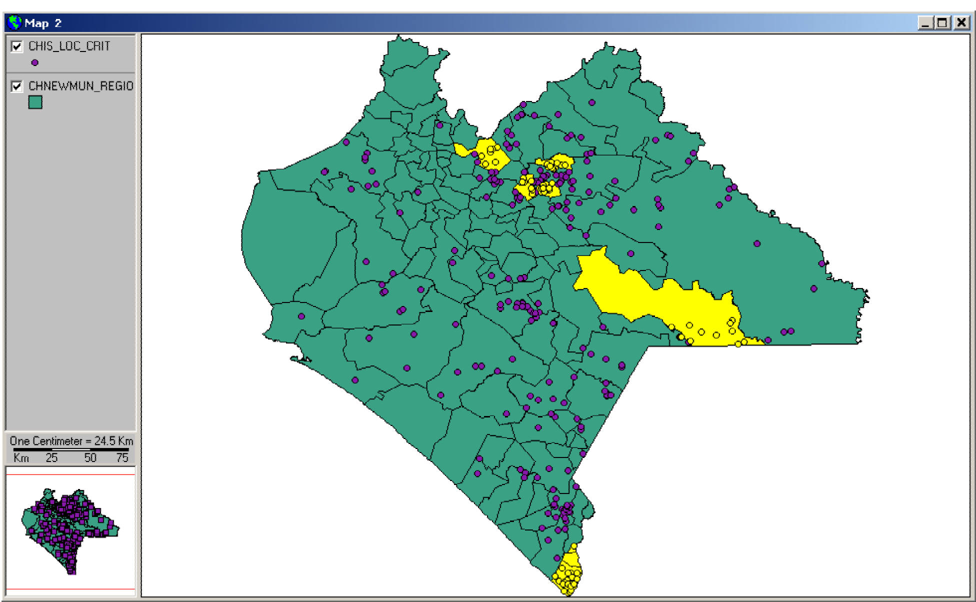
This dialog presents its control in natural language way to allow a better understanding.

4. The **name** of the selected or active layer is displayed. It shows total of features and selected features in the layer.
5. Choose the selection method
6. Choose the **Layer** which selected features will be used for selection, in the list "selected feature of layer"
7. ***Click*** the range you want to apply to the selection's **result**,
8. ***Click*** the Select button to choose the geographic units that meet the defined condition.

Example:
A team of epidemiologist needs to know the cities and localities that are in areas of higher risk of malaria. They have two cartographic layers in a map; one of Municipality boundaries with the indicator Annual Parasite Index (API), and other layer of cities and localities where the mayor percent of population is concentrated. The team has defined that an API higher that 4.0 is consider of high risk. As the API is measured at the municipality level, first, it is needed to select all the municipalities with API > 4.0. The team has done this step using the function "Select using attributes". The selected municipalities appear in yellow in the map

So, to select the localities that overlay the municipalities of higher risk of malaria, they choose the option Selection by Layer, on the Layer menu. Using the dialogue Select by Layer, they decided to select the localities that **intersect** with the selected municipalities from the municipality layer.

As a result, the localities that intersect the selected municipalities are now selected too. Small yellow circles in the map represent them.



Maps Window. Selection by layer.

Selecting all the geographic units of a Layer

To select all the geographic units of a Thematic Layer:

1. Activate the Maps Window and Select the desired Layer,
2. Choose the option **Select All**:
   a.) from the **Layer** menu on the Menu Bar,
   b.) Or, from the menu displayed by **right *clicking*** the selected Layer,
   c.) Or, from the **Table** menu on the Menu Bar displayed when there is an active Attributes Table.

As a result, **all** the geographic units of the active Layer appear highlighted in the color and configuration defined in the selection properties of the Maps Window.

Clearing a previous selection of features

While working with Maps, you might select specific geographic units for an analysis, and find that you need to unselect or clear the current selection and continue the analysis with all the geographic features.

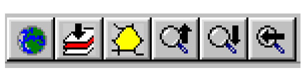
You can **select** the features in a layer using spatial selection tools, or by three other selection methods: by attributes, using another Layer, or selecting all features. Review previous topics to learn about selection methods.

To clear the current selection:

1. Activate the Maps Window, and Select the desired Layer,
2. Select the option **Clear Current Selection**,
   a.) from the **Layer** menu,
   b.) Or, from the menu displayed by **right *clicking*** the name of activated layer on the Legend,
   c.) Or, from the **Table** menu, when a Attribute table windows is active.

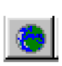
As a result, the active layer is displayed without any geographic units selected.

Other Tools related to Layers



Buttons for other Tools for management of Layers.

The tools described in this topic, are accessible in the toolbar when a map window is active. Some of the related functions to the tool buttons are accessible also through the menu Layer.

 Full Map Extension

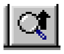
The complete display of the Map adjusts to the size of the Maps Window.

 Zoom to Active Layer

Fully displays the selected Layer adjusting it to the size of the Maps Window.

 Zoom to selected Unit

Displays the geographic units selected in the active Layer, adjusting them to the size of the Maps Window.

 Zoom-in

Zooms in on the displayed Map, enlarging the size of the Map image.

 Zoom-out

Zooms out on the displayed Map, reducing the size of the Map image.

 Last Map Extension

Displayed Map returns to its previous state. You can carry out this action twice, displaying the Map to the next previous state.

Tools for Map Management

The Maps Window has a set of **tools for spatial selection** that allow you to carry out operations with the Map.

Each button of the Toolbar associated with an active Maps Window is described below.

You can select any of them by ***clicking*** the desired button. Depending on the tool selected, the cursor will change its shape as it moves over the Map.

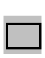
Toolbar for spatial selection. Maps Window.

 Cursor tool

Tool that allows you to select geographic units on a Map. You select the geographic unit indicated by the cursor and can select several units at a time by holding the **Shift** key down each time you ***click*** a geographic unit.

 Point tool

Selects a geographic point. It works just like the cursor, except that you can only select one geographic unit at a time. By ***clicking*** the geographic object, you get the desired point.

 Rectangle tool

Allows you to select geographic objects on a Map by drawing a rectangle. When you draw a rectangle, those objects that intercept it are selected. You select a point to start, ***click***, and move the cursor until you have the area you want and release.

Circle tool

Allows you to select geographic objects on a Map by drawing a circle. As with using the rectangle tool, when you draw the circle those objects that intercept it are selected. In the same way, ***click*** the center point of the circle that you want to draw, and release the ***click*** when you get the desired diameter.

Line Segment tool

Allows you to select geographic objects on the Map by drawing a line segment. When you draw the line segment, objects that intercept it are selected.

You must ***click*** the starting point of the segment, then draw a straight line, ***click*** the following point and continue to draw the segment, repeating these steps; when you finish drawing the desired segment, **double *click*** the final point.

Polygon tool

Allows you to select geographic objects on a Map by drawing a polygon. When you draw the polygon, the objects included within it are selected.

As in the case of drawing a segment, you ***click*** the starting point of the polygon; then move the cursor drawing a segment, ***click*** and continue to move the cursor to continue creating the desired polygon as a sequence of segments. Repeat these steps until you get the desired area; when you finish **double *click*** the final point.

Zoom in tool

Allows you **to increase** the level of detail of a Map (***zoom in***) or reduce its Scale, by ***clicking*** the Map. When you move the cursor, a rectangle is drawn between the initial and final points producing a close-up of the area selected.

Zoom out tool

Allows you to **decrease** the level of detail of the Map (***zoom out***) or increase its Scale, by ***clicking*** the Map. This produces a **more distant image** of the Map displayed.

Pan tool

Allows you to move a Map displayed in the Maps Window. By ***clicking*** the Map and moving the cursor within the Maps Window you can move the Map to a specific position; when you are finished release the ***click***.

Identify tool

Allows you to interact with a Map. It shows the geographic unit's attributes of the active Layer. By ***clicking*** on a feature of the Map, it opens a window that displays the attributes of the active layer.

Labels tool

Allows you to place Labels on the geographic units of a Map, using an attributes of the active layer.

You must previously select a layer of a Map and define the attribute to be used as label by selecting the option ToolTip Variable, in the **Map** Menu. To place a label, ***click*** on a layer feature on the Map. See the heading Working with Labels.

Measure tool

Allows you to measure distances on the Map.
To do this, ***click*** the place to start measuring and move the cursor towards your destination; **double *click*** at the end point. A straight line is displayed on the Map and the **value of the distance** represented by the segment drawn is displayed automatically, using the unit of measure in the Scale. The scale is shown in the System Status Bar, located at the bottom of map window. Please, previously set the correct map units using the Map Properties. The map units are the units of the layers in the map, and could be decimal degrees, feets or meters.

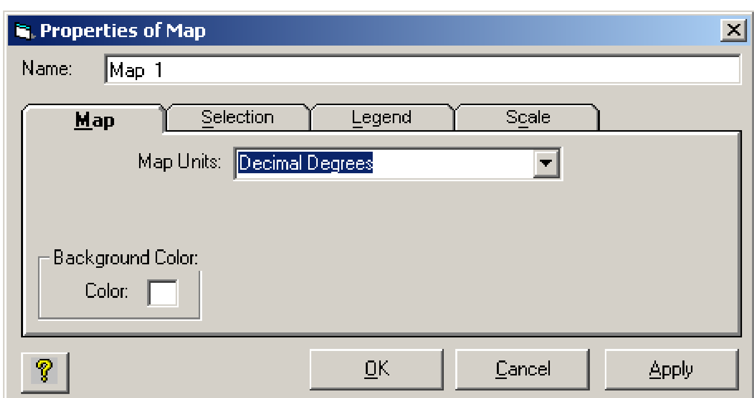
Defining the Properties of the Maps Window

The map window has a set of properties that you can change according to your needs. The properties are related to the elements contained in the map window, those are properties of the map itself, visualization of selected features, the legend and the graphic scale.

From an active Maps Window, the properties of the window can be defined using the Properties of Map option:

1. Select the option **Properties**..., from the **Map** menu on the Menu Bar;

   The dialogue "**Properties of Map**" is displayed that allows you to define and edit the characteristics of the Maps Window:

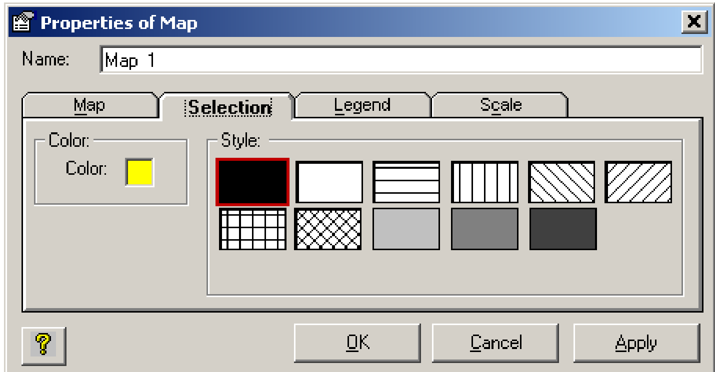


Dialogue box. Map Properties. Modifying Map characteristics.

2. Modify the Map's **name**, **units of measurement**, background **color**,
3. Define the **Selection** properties to use when you select an area of the Map,
4. Characteristics of the **Legend** and **Scale**.

Just select the relevant folder tab in the dialog box (Map, Selection, Legend, Scale).

An important map property is the Map Units because all the distance measurements on the man depend on this property. The map window set Decimal Degrees as map units by defaults.



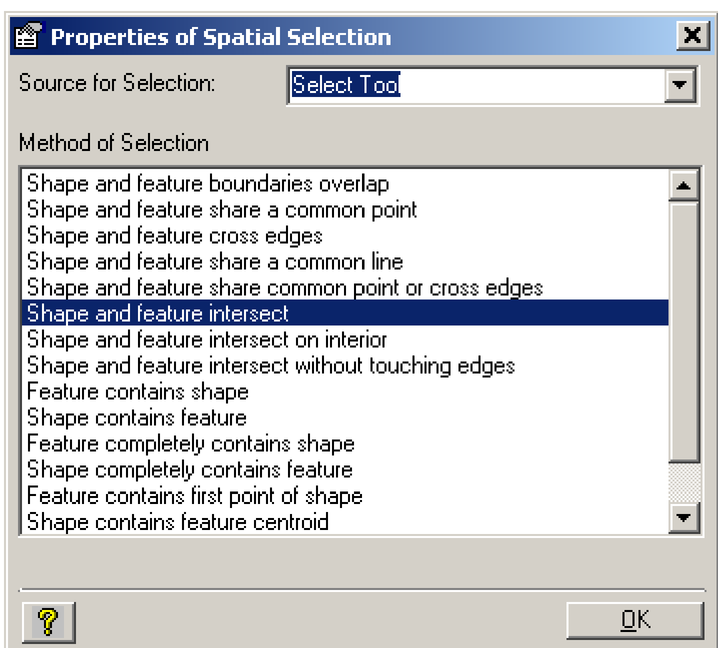
Dialogue box. Map Properties. Modifying Selection characteristics.

Spatial Selection Properties

The properties of spatial selection are related to the spatial selection functions of features in the map rather than the visualization properties mentioned above. By default, the selection of features is set to use selection tool (graphic tools) and the method of intersection. This means that if you use a circle as a selection tool, you will get the features that intersect with that circle. Through the use of Spatial Selection Properties you can define the way to do feature selection.

To define the spatial selection properties for the Maps Window:

1. Activate the Maps Window,
2. Select the option **Spatial Selection Properties...,** from the **Map** menu,
3. Select the **Source of selection**. This can be any tool in the spatial selection Toolbar. Or, activate a Layer in the Map; this means that you will use the selected Layer to choose elements for another Layer.
4. Select the **Method of selection**; this means the search method used –that is, the way you want to use the spatial selection process.



Dialogue Spatial Selection Properties.

Example: Selecting Municipalities of a given state.

Select as the source the States Layer of a region.

In the Maps Window, activate the Municipalities Layer, by ***clicking*** any of the municipalities; the state containing that municipality is displayed on the Map, depending on the characteristics and properties of the selection defined in the Map Properties.

Thematic Maps

SIGEpi allows users to create a set of Thematic Maps based on attributes of layers

The Thematic Maps show specific information concerning particular locations, general information on the spatial patterns, and make it possible to compare patterns in one or more maps. The representation of the attributes of the geographical units should be carried out using international conventions that communicate adequately the phenomenon, object, or characteristic that is visualized in the map.

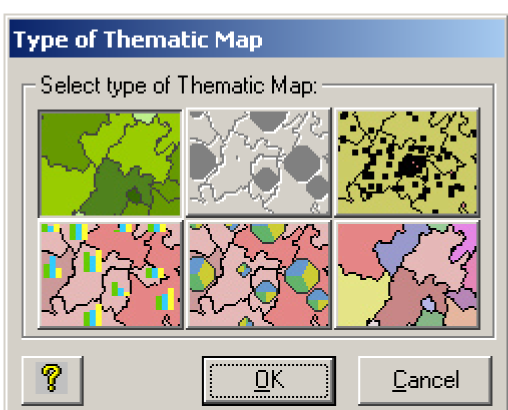
The Thematic Maps are utilized to emphasize the spatial distribution of one or more attributes or variables of the geographical units. Although the thematic map of choropleth is the most frequently used, exists a variety of thematic maps, for example, proportional or graded symbols, unique values, dot-density map, of flows, etc..

To create a Thematic Map, make sure that the Maps Window is active and a layer is selected. Otherwise the option **Create Thematic Map...** on the **Layer** menu will be disabled**.**

Creating a Thematic Map

To create a Thematic Map:

1. Activate the Maps Window and select the layer which you want create the thematic map,
2. From the **Layer** menu, choose the option **Create Thematic Map...,**
3. Or, ***click*** the **Thematic Map** button on the Toolbar,
4. From the **dialog box** displayed, select the type of Thematic Map by ***clicking*** the type of Map desired. ***Click*** on **OK** button to confirm your selection.



Dialogue box. Select Type of Thematic Map

The **dialog box** corresponding to the **type of Thematic Map** is displayed allowing you to select the **variables**, **methods**, **color**, etc., that are used to create the Thematic Map.

Each type of Thematic Map has its own associated dialog box, depending on the characteristics of the type of Thematic Map created.

In the heading Types of Thematic Maps, you can see the characteristics of the dialog boxes used to construct each **type** of Thematic Map.

Each **dialog box** gives you options or parameters that you can change to define the type of Thematic Map you want to construct.

You should note the fact that some types of Thematic Maps only need **one variable** (this is the case for Thematic Maps using Ranges, Dot Density, Graduated Symbols, and Single Values), while others (Bar Chart and Pie Chart maps) need **more than one variable**. The dialog box displayed in each case allows you to select one or more variables, depending on the case.

Once you have defined the characteristics and variables for constructing the Thematic Map:

1. Click on **Update** button and the **Map Legend** defined in the dialog box is updated,
2. Click on **Apply** button and the Thematic Map constructed is displayed in the Maps Windows,
3. Click on **OK** button of the corresponding dialog box to accept the created Thematic Map. This causes the dialog box to close and returns control to the Maps Window.

You can continue to modify the variables and characteristics defined in constructing the Thematic Map by clicking on **Update** button to updated these on the Map Legend, clicking on **Apply** button to see how the Map will look, and then clicking the **OK** button when satisfied with the selection made.

To change the color of the variables or ranges in the Legend on a Thematic Map:

1. Click on **Update** button as many times as you like; this allows you to modify the colors of the Map Legend and, therefore, the colors with which the Thematic Map is displayed.

To see each type of Thematic Map and their associated dialog boxes, see the heading **Types of Thematic Maps**.

If you want to modify a Thematic Map that is already constructed, see the heading **Modifying the Thematic Map**.

Types of Thematic Maps

In **SIGEpi** the steps for constructing Thematic Maps have been simplified; just select the variables and methods to use and **SIGEpi** automatically applies the selected method and creates the corresponding Thematic Map, which is displayed in the Maps Window.

Six different types of **Thematic Maps** can be constructed:

- Choroplets or Ranges,
- Graduated Symbols,
- Dot Density,
- Bar Charts,
- Pie Charts, and,
- Unique Values.

Thematic Map using Ranges

When a **Thematic Map of Ranges** is created, all the data records are grouped into ranges. Each range is associated with a color and a particular pattern used to identify geographic objects (points, lines and polygons) on a Map, depending on the values of the variable used in constructing the Map.

In the case of a Thematic Map using Ranges, the dialog allows you to:

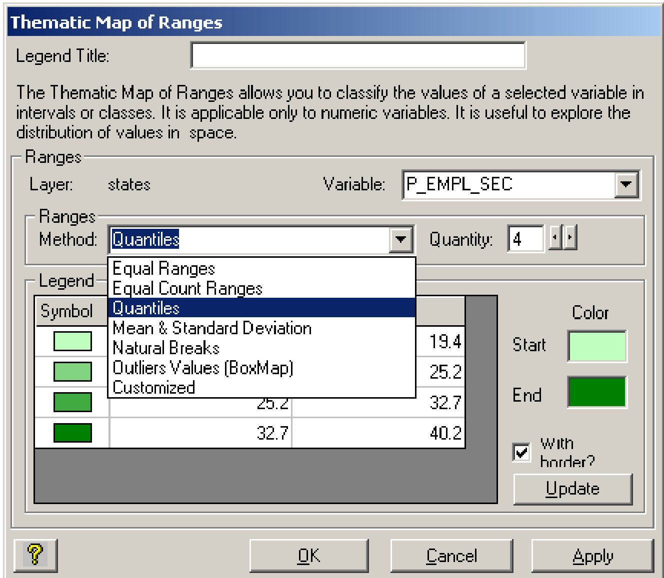
1. Establish a name for the Map Legend's title,

2. Select the **variable** that is the object of analysis,

3. Select the **method** to construct the ranges,

4. Define the **number of ranges** in which you want to group the values of the variable depending on the method selected,

5. Choose the **beginning and ending colors** for displaying the color ramp in the ranges; to do this ***click*** the beginning and ending colors and select the desired color,



Dialogue Thematic Map of Ranges.
Selecting the method to calculate the range values.

6. Update the Map Legend by clicking on **Update** button; this action causes the ranges calculated in the Legend edit box of the dialog box to update,

7. Get the desired Thematic Map by clicking on **Apply** button; this causes the corresponding Thematic Map to display in the Maps Window,

8. Edit the **colors** and **cut-off values** of the ranges displayed in the Legend edit box; for this you must previously select the **Customaized** method; see **Updating the Map Legend**,

9. ***Click*** the **OK** button to conclude.

Example:
You have a layer of municipal boundaries and you want to classify them according to their population in order to find the areas with the largest populations.

The municipalities are grouped (classified) into ranges according to their population. For example, municipalities with a population value less than 125,000 habitants are grouped into the first range, those with more than 125,000 and less than 160,000 are grouped into the second range, those with more than 160,000 and less than 190,000 go in the third range, and so on.

This type of thematic map makes easier the identification of geographic distribution pattern of the populations; likewise, if instead of representing population, rates of a particular disease are used, the map will show the geographic distribution pattern of the disease, and could be easy to identify the municipalities al highest health problems.

Representing your data with Thematic Maps allows you to identify at a glance differences among areas for the indicator being analyzed.

An interesting aspect of Thematic Maps of Ranges is the **method** for constructing the ranges.

SIGEpi offers several **methods** for constructing ranges; they are:

- **Equal Ranges**: This method creates ranges so that each range is the same size. For example, you have a variable in the Table with values from 1 to 100. If you create a Thematic Map using ranges defined with four equal ranges, you get the following ranges: 1-25. 26-50, 51-75, 76-100.

- **Equal Quantity Ranges**: This method creates ranges so that approximately the same number of observations falls within each range.

- **Quantiles**: The ranges created are based on the cumulative relative frequency of values of a variable, taking a given number of selected categories. For example, if 4 categories are defined (quartiles), 25% of the observations (records) are in each category; if quintiles are used (5 groups), then 20% are classified in each category, and so on.

- **Natural Break**: This method creates ranges using an algorithm for detecting *clusters*, based on achieving the lowest possible dispersion among values within each range. This guarantees that similar values fall within the same range and that the greatest possible difference in values exists between the ranges. This method uses the Fisher-Jenk's algorithm that guarantees an optimal solution. The reader may find the algorithm referred to as "Jenk's optimal method".

- **Mean and Standard Deviation**: This method creates ranges using the mean and standard deviation of the variable's values. Each range has the width of a standard deviation and ranges lower than the mean are shown with tones of blue (from dark to light), while ranges higher than the mean are shown with tones of red (from light to dark).

- **BoxMap**: This method creates ranges calculating the quartiles of the variable's values and applying the rule of mean +/- 1.5 multiply by the Inter-quartile range to calculate the cut-off values for low and high outliers. This is useful for exploring data and detecting outliers. When this method is applied, six ranges are created, the first range representing low outlier; next four ranges are first to fourth quartiles and the sixth, high outliers.

- **Custimized**: This method allows the user to define the ranges' cut-off values. By default, some values will appear that the user can modify. Only the values of the upper limits of the ranges can be **edited**, from the first range up to the next-to-last range. Whenever the upper limit of a range is updated, the value is automatically used as the lower limit of the next range; the corresponding Thematic Map is automatically updated when you ***click*** the **Apply** button; see **Updating the Map Legend**.

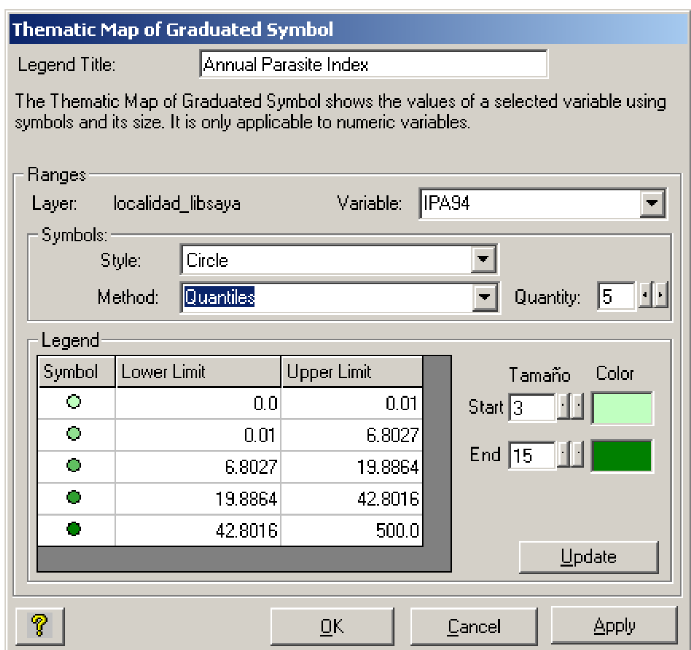
Thematic Map of Graduated Symbols

The **Thematic Map of Graduated Symbols** uses symbols to represent the values of a variable you are working with. The symbol's size is directly related to the value of the variable. This type of Thematic Map can be used for any type of geographic object for which there is a **Thematic Layer with Points.** see **Plotting points from a Table** to get the corresponding Thematic Layer.

Using Graduated Symbols is recommended, for example, to show the amount of services offered by a health care unit, the number of requests from the community for certain specialized health services, the number of diagnosed cases of a disease. It does not make sense to apply this type of thematic map to display the type of health services because that is a categorical variable; however, it is appropriate to show the quantity of services given.

In the **dialog box** for Thematic Maps of Graduated Symbols you can:

1. Set a name for the title of the Map Legend,
2. Select the **variable** to be used in the thematic map,
3. Select the **style** of the symbol for representation,
4. Select the **method** to be used to create the symbol ranges or intervals,
5. Define the **number of ranges** of interest for grouping the values of the variable,
6. Define the **size and color** of the smallest symbol (beginning), and the **size and color** of the largest symbol (end), according to the ranges into which the variable is grouped.

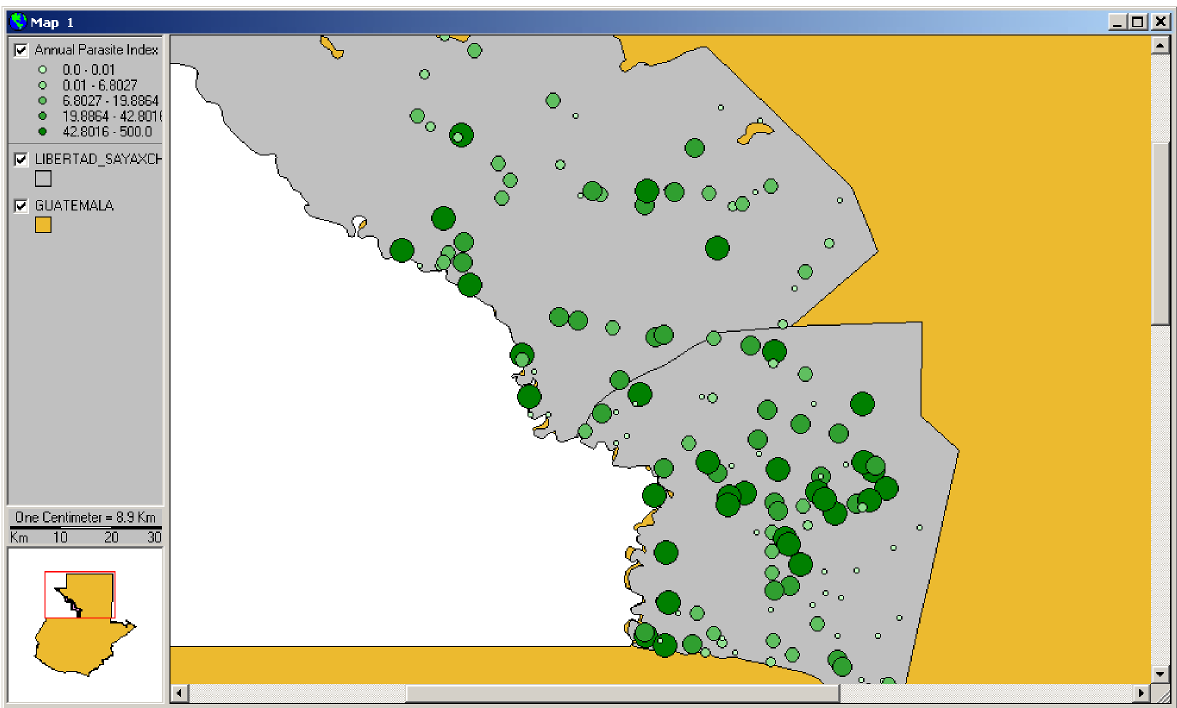


Dialogue box. Thematic Map of Graduated Symbols.

In this type of thematic You can see that the size of the symbol can be changed as varied as wanted; so can the type of symbol used for the Thematic Map.

7. Update the **Legend** edit box by ***clicking*** the Update button so that the ranges calculated in the Legend edit box of the dialog box are updated,
8. Get the desired Thematic Map by ***clicking*** the Apply button; this causes the corresponding Thematic Map to display in the Maps Window,
9. Edit the **colors** and **cut-off values** for the ranges displayed in the Legend edit box. For changing the cut-off values you should select the **Customized method**, see **Updating the Map Legend**,
10. Click the OK button to conclude.

Example:
You have a layer that shows the rates of *P. vivax* malaria at the level of critical localities in a territory, and you construct a thematic map of Graduated Symbols that groups reported cases among the population of every locality in the territory, expressed in thousands of cases.



Map Window. Example of Thematic Map using Graduated Symbols

The symbol size offers an easy way to identify the incidence of cases in each locality and the areas most affected by the concentration of values represented.

In this type of Map, the same **methods** are used for constructing ranges as in the case of Thematic Maps of Ranges.

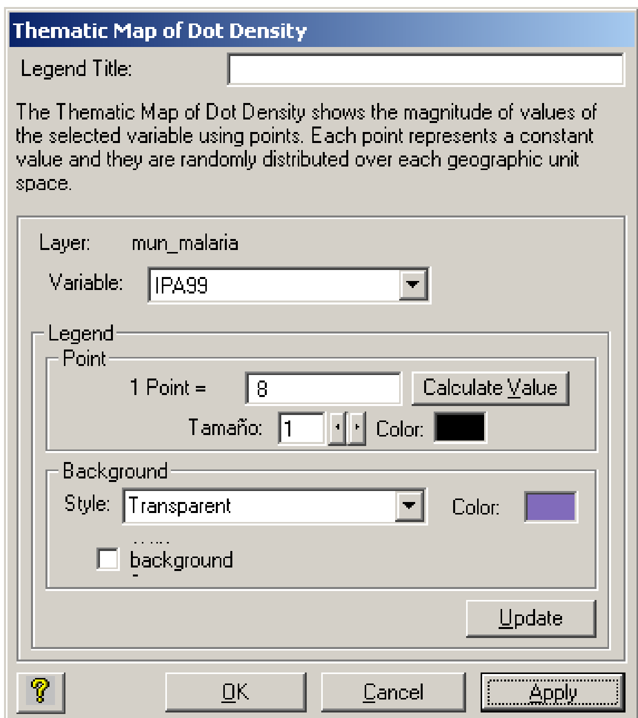
Thematic Map of Dot Density

The **Thematic Map of Dot Density** uses dots to represent the value associated with a geographic region. Each dot represents a previously fixed amount and the dots are distributed randomly within the area. For example, if there is a municipality with 100,000 inhabitants, and each dot represents 1,000 inhabitants, then the municipality will have 100 points distributed randomly within its border.

The **dialog box** for Thematic Maps using Dot Density allows you to:

1. Establish a name for the title of the Map Legend,
2. Select the **variable** to use in building the Map,
3. Define the **value** that each point represents,
4. Define the **size** and **color** with which the **point** is represented on a Map, and,
5. Select the **style** and **background color** for the Map.

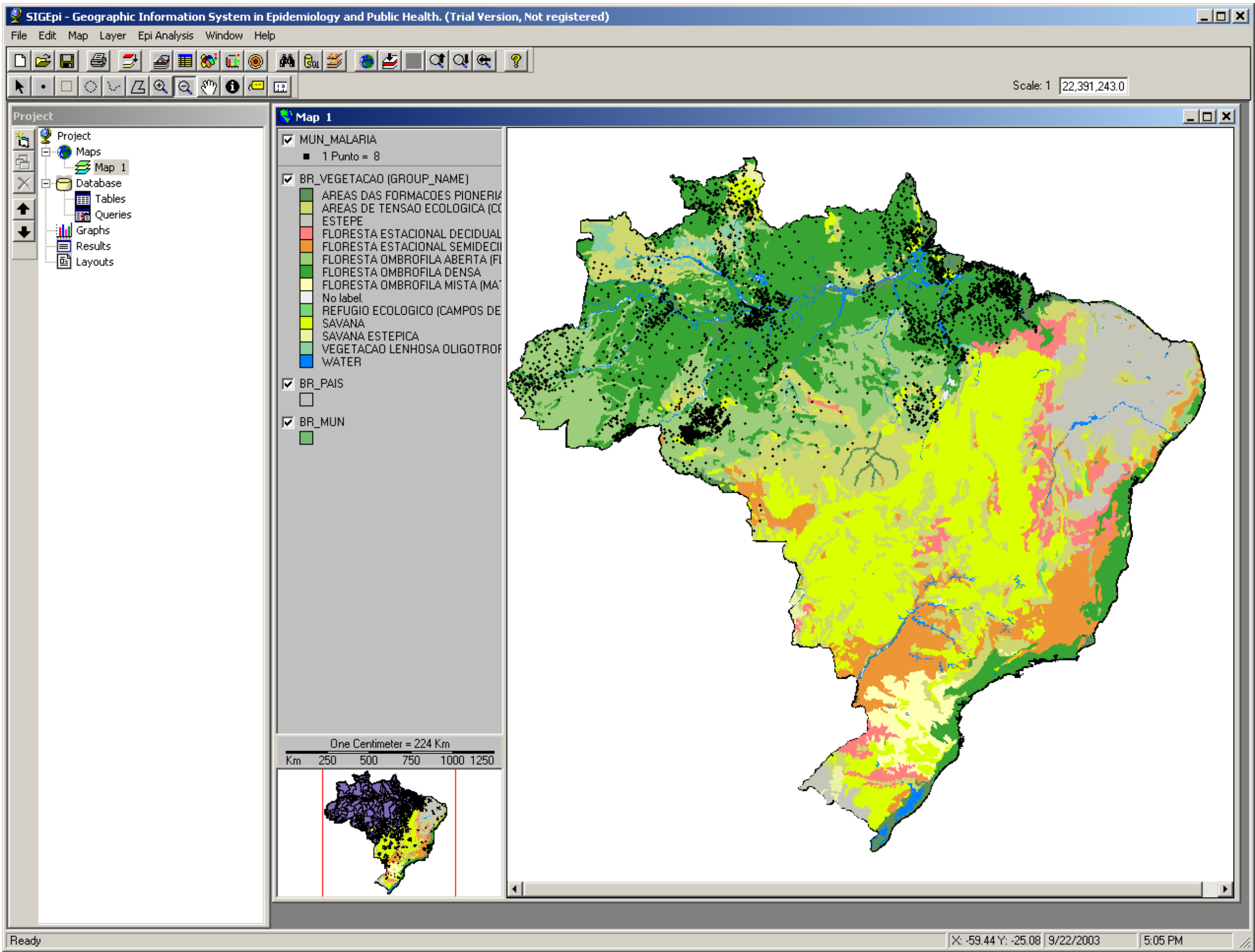
When the amount that each dot represents is large the number of dots on the Map will be few, and vice versa.



Dialogue box. Thematic Map using Dot Density.

Example:

In the unit of epidemiological surveillance at the national level in Brazil, a team of professionals is analyzing the distribution of positive cases of malaria by municipalities. The data is available by municipalities. They decided to map the number of positive cases by municipalities using a thematic map of dot density. The data is mapped in a way that is overlay on a map of type of vegetation.



Map Window. Thematic Map using Dot Density.

With this map, the team was capable to identify the areas of high density of positive cases of malaria. In this case is not showed the boundary of the municipalities.

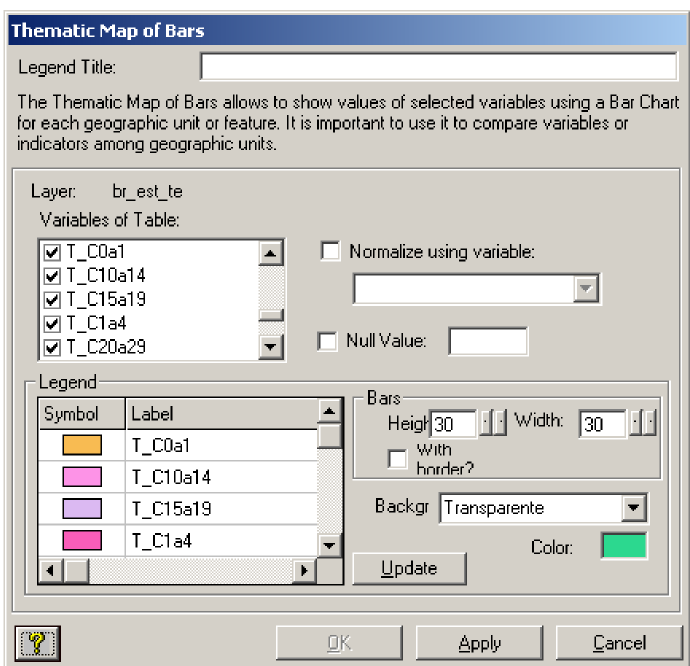
Thematic Map using Bar Charts

Unlike the Thematic Maps that work with a single variable to construct a Map, **Thematic Map using Bar Charts** allow several variables to be examined at the same time. This type of Thematic Map can be constructed for any type of geographic object because the centroid of the object is used to represent it.

The Thematic Map using Bar Charts permits the analysis of thematic variables through comparison of the heights of the bars on the chart.

The dialog for Thematic Map of Bars allows you to:

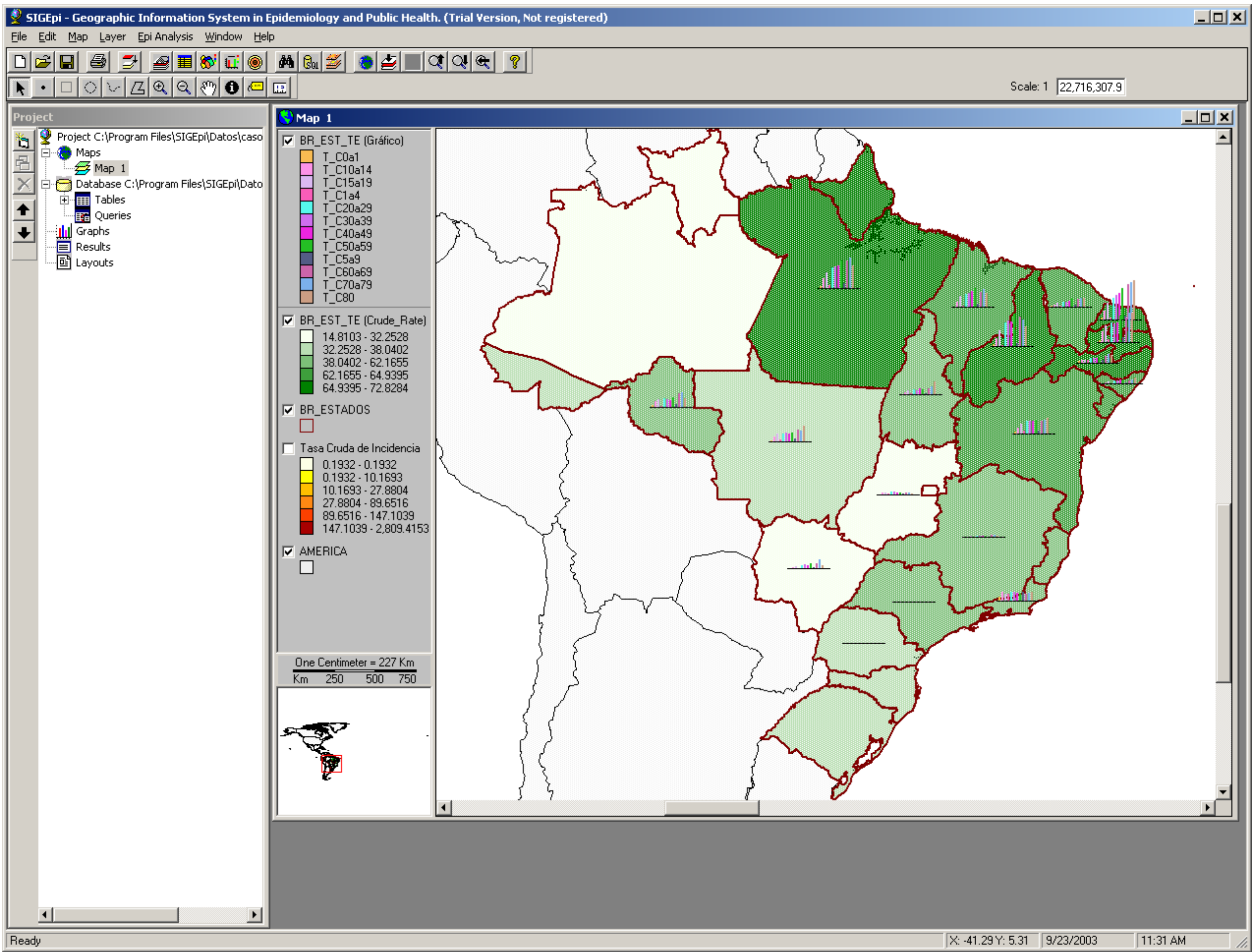
1. Define a title for the thematic map,
2. Select the **variable(s)** for analysis from the Attributes Table associated with the Layer,
3. Select whether or not to **standardize a variable**; this allows you to standardize the values of the variable(s). For example, if the variables for the male and female population in a region are represented and use the variable of total population for standardization, you get the values for male and female population standardized in relationship to the total population.
4. Indicate whether you want to use a **null value and define what it is**; that is, if you want to represent null or missing values. For example, if the cases of hepatitis in a region are represented, you could use a certain value to represent regions with no cases or have not reported any; this allows you to identify such regions visually on a Map.
5. Define the **height** and **width** of the bars for display and whether you want them to have borders or not,
6. Select the **style** and **background color** you want to use to display the Thematic Map,



Dialogue box. Thematic Map using Bar Charts.

Example:

You have the layer of state boundaries of Brazil and we have available the specific incidence rate of Dengue for age groups. The Thematic Map using Bar Charts will show a bar chart for each state; each bar represents the incidence rate for one of the age groups.



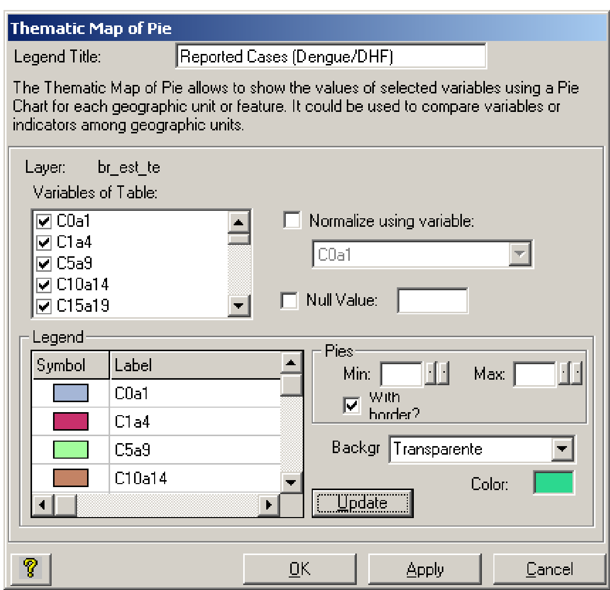
Thematic Map of Bars.

Thematic Map using Pie Charts

As in the case of a Thematic Map using Bar Charts, a **Thematic Map using Pie Charts** also permits the examination of several variables at the same time in constructing a Map. As the heights of bars are compared in a Thematic Map using Bar Charts, a Thematic Map using Pie Charts compares the size of each pie wedge in a single pie and also compares the wedges of the pies for all geographic areas. Each wedge represents the value of an associated variable. The pie also allows you to compare the parts of a whole.

The dialog box for Thematic Map using Pie Charts allows you to:

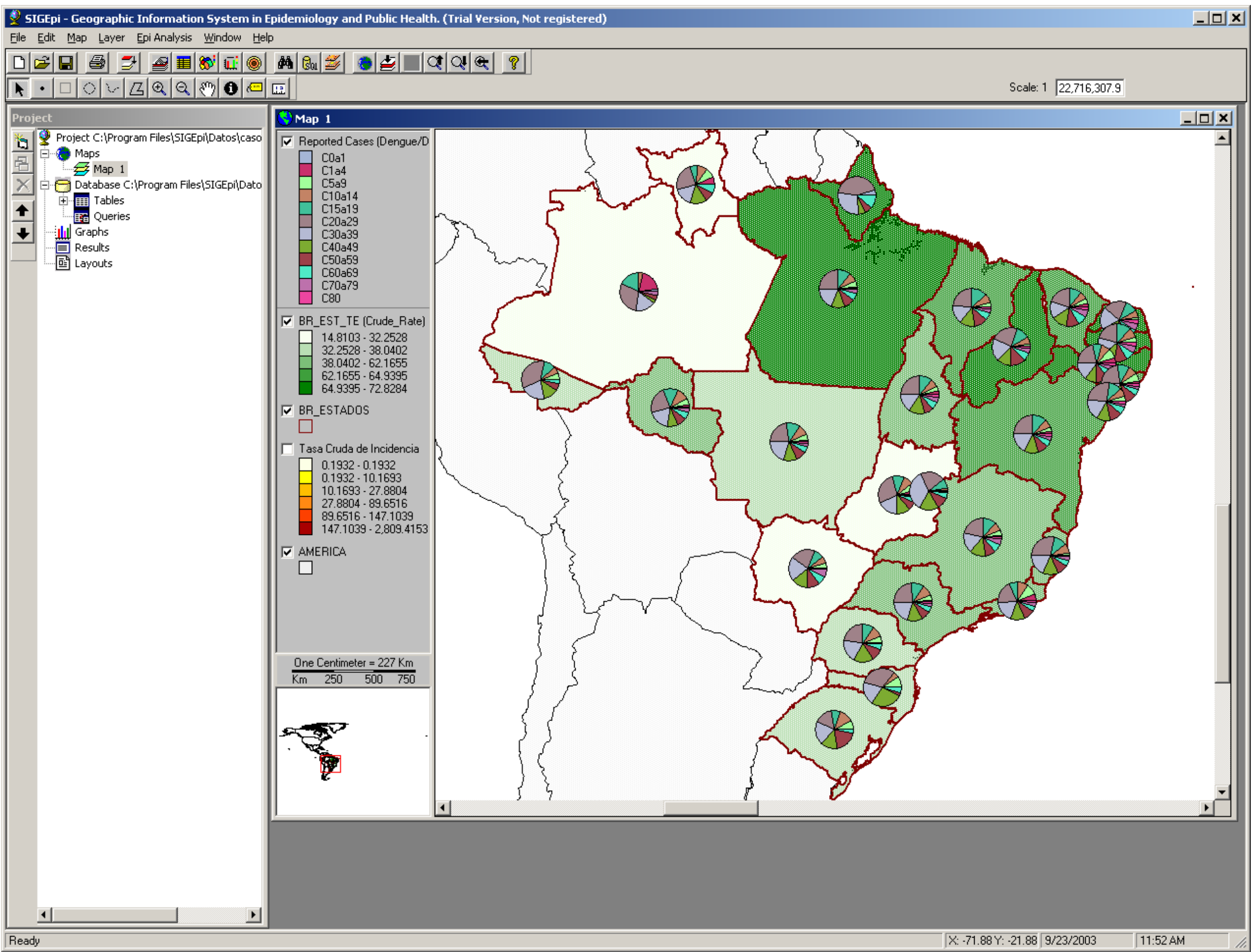
1. Establish a name for the title of the Map Legend,
2. Select (check) the **variable(s)** to display in each bar of the thematic map.
3. Select a variable to standardize the values of the selected variables; this allows you to create the thematic map using relative values (proportions) of each selected variable in relation to the variable for normalization.
4. Indicate if you want to use a **null value and define what it is**; that is, if you want to represent null or missing values.
5. Define the characteristics of **minimum** and **maximum** size of the displayed pie charts and whether you want them to have **borders** or not. For example, defining a minimum and maximum size for the pie charts will make your geographic representation clearer. In addition to seeing the incidence of the variable under study by the size of the wedge, you can see the incidence of the variables in every area by the size of the pie, which would be larger where there is a greater concentration of such variables.
6. Select the **style** and **background color** you want to use to display the Thematic Map,



Dialogue box. Thematic Map using Pie Charts.

Example:

Having the layer of federative units (states) of Brazil and the reported cases of Dengue and Dengue Hemorrhagic Fever by age groups, it is convenient to represent these data using pie chart on the map. This type of map allows visualizing the magnitude of the problem in terms of absolute values, the magnitude by age groups and an idea of the distribution of this disease by Federative Units.



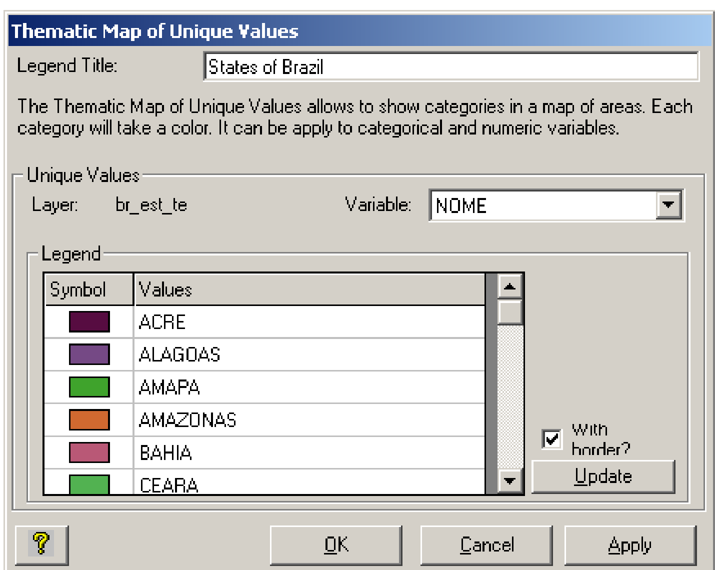
Map Window. Thematic Map using Pie Charts.

Thematic Map using Single Values

The **Thematic Map using Unique Values** is used to represent phenomena that show a compact and uniform spread within a territory. It is assumed that the phenomenon represented is either present or absent for a geographic unit, but that it is not possible to distinguish levels of intensity within the geographic unit. Generally, this type of map is used to represent qualitative or categorical variables, i.e. a classification of areas according to risk categories: very low, low, medium, high and very high. In spite of that it is possible to use quantitative or numeric variables in this type of map.

The **dialog box** for Thematic Map using Single Values allows you to:

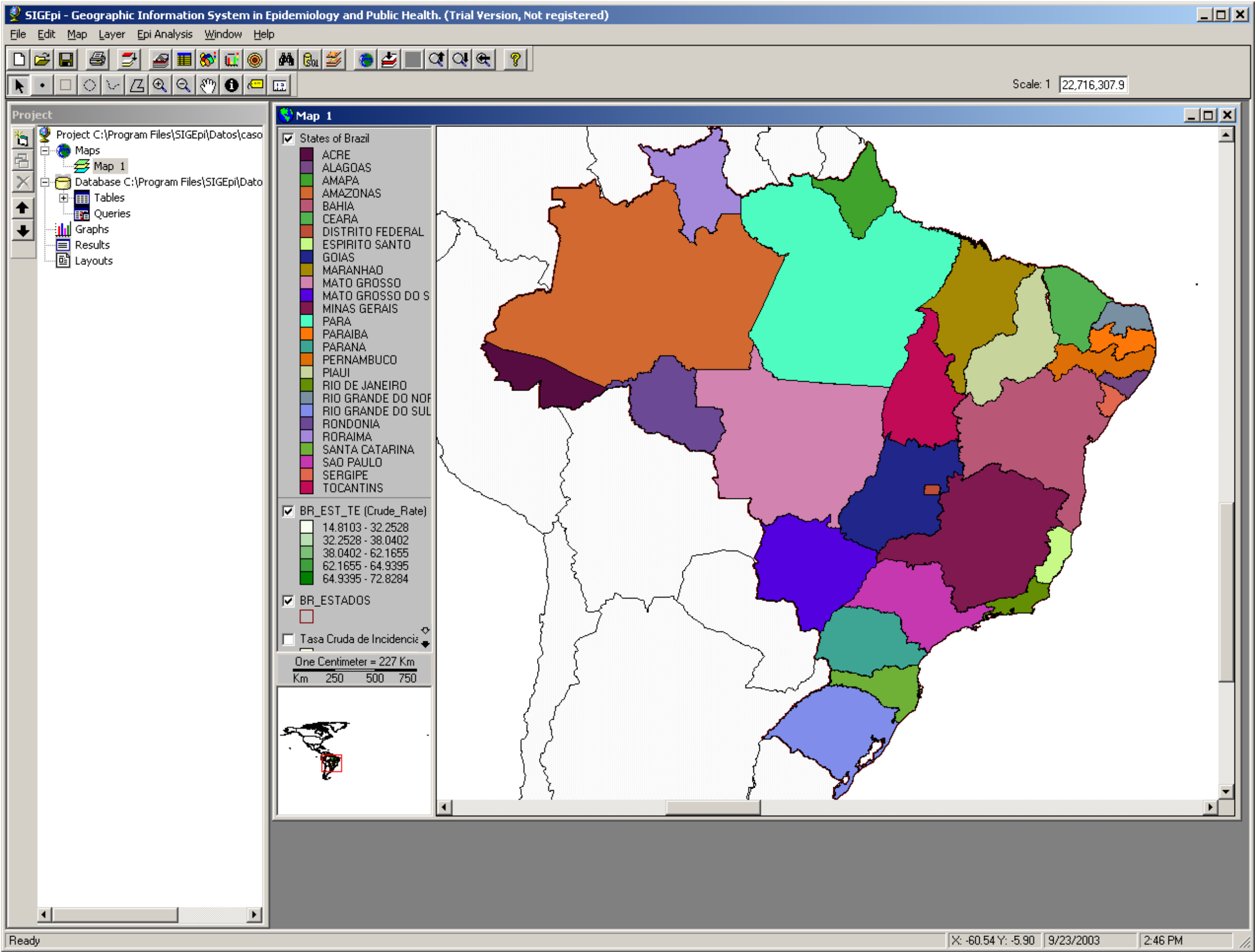
1. Establish a name for the Map Legend,
2. Select the **variable** to graph; it is important for indicate that a categorical variable should be chosen for this map. It is possible to use a numeric variable but it should be taken into account that this type of thematic map is to represent categories not quantities.
3. Specify whether you want a **border** for polygons or area features.
4. Update the **Legend** edit box by ***clicking*** the **Update** button.
5. Show the desired Thematic Map by ***clicking*** the **Apply** button; this causes the corresponding Thematic Map to display in the Maps Window,
6. Edit the **colors** for the categories displayed in the Legend edit box; see Updating the Map Legend,
7. Click the **OK** button to conclude.



Dialogue box. Thematic Map using Single Values.

Example:

In this example it is used the layer of states of Brazil and the variable name of state (NOME) to categorize the states. The map displays each state with a different color. If it is used a variable that more than one state share one value, those states will have the same color.



Map Window. Thematic Map of Unique Values.

Modifying a Thematic Map

Frequently is desirable and necessary to edit or modify an already created thematic map.

To modify a Thematic Map:

1. Activate the Maps Window and select the **Layer** that contains the Thematic Map you want to modify,
2. a.) **Double *click*** the Layer in the Legend of the Maps Window to activate the Thematic Map dialog box,
   b.) Or, ***click*** the **Thematic Map** button on the Toolbar,
   c.) Or, Select the option **Modify Thematic Map...** from the **Layer** menu.

In each case the **dialog box** for the type of Thematic Map is displayed. You can change or define new variables and characteristics of the Thematic Map, just as it was explained in the case of the creation of a new Thematic Map.

To build other type of Thematic Map based on the same Thematic Layer:

- If you want to keep the existing Thematic Map fallow the steps:

1. Duplicate the layer from which you want to create the new Thematic Map,
2. **Actiavte** the new layer, and
3. **Create** the new thematic map as described above in the heading **Creating a Thematic Map**.

- If you do not want to keep the existing Thematic Map, delete the Thematic Map, and

1. **Select** the layer,
2. **Create** the new Thematic Map as described in the heading **Creating a Thematic Map**.

If you want to modify or change the **colors** and the **cut-off values** of each range of the Thematic Map, see the heading Updating the Map Legend.

Updating the Map Legend

Updating the ranges' colors and values on the Map Legend:

Once the characteristics and variables for constructing a Thematic Map have been defined according to the corresponding Dialogue Box for creating each Thematic Map, the ranges' colors and the values' limits, generated automatically for each range, are displayed depending on the method selected.

- As seen in the creation of each type of Thematic Map, you can:

1. ***Click*** the **Update** button, and the **Map Legend** is updated in the dialog box,
2. ***Click*** the **Apply** button, and the constructed Thematic Map is displayed on the Map,

3. ***Click*** the **OK** button of the corresponding dialog box once satisfied with the Thematic Map created; this closes the Dialogue Box and returns control to the Maps Window.

- Working with the Dialogue Box to construct each type of Thematic Map, you can:

1. Continue to modify the **variables** and **method** defined in constructing the Thematic Map, and select other values,
2. ***Click*** the **Update** button so that the modifications are updated on the Map Legend, and
3. ***Click*** the **Apply** button to see how the Map would look, and then ***click*** the **OK** button when you are satisfied with the current selection.

This procedure can be repeated as many times as you want until you find the right combination.

- If you only want to modify the **colors** of the Thematic Map displayed in the Maps Window:

1. ***Click*** the **beginning** and **ending colors** for the ranges in the Dialogue Boxes that have this characteristic, and ***click*** the **Update** button and then the **Apply** button,
2. Or, ***click*** the rectangle with the specific **color** of the desired range in the **Legend** edit box of the Dialogue Box, and ***click*** the **Apply** button.

These procedures can be repeated as many times as you want to get the desired color combination.

- If you want to modify the values of the ranges' limits or intervals:

For Thematic Maps with Ranges or Graduated Symbols, it is possible to modify the ranges' **values of the upper limits** or the intervals of each range; to do this, select the **Personalized method** in the Dialogue Box, this allows you to:

1. ***Click*** each value for the **upper limit** of a range (except for the last range), and edit its value,
2. The same value will automatically become the initial value of the next range,
3. ***Click*** the **Apply** button and it automatically updates the corresponding Thematic Map.

Just as when you modify colors, this procedure can be repeated as many times as you want to get the desired intervals.

Delete a Thematic Map

You may want to delete a Thematic Map from the Maps Window of your Project.

To delete a Thematic Map:

1. Activate the Maps Window and select the **Layer** that contains the Thematic Map you want to delete,
2. **Right *click*** the Layer or ***click*** the **Layer** menu on the Menu Bar,
3. Select the option **Delete Thematic Map**.

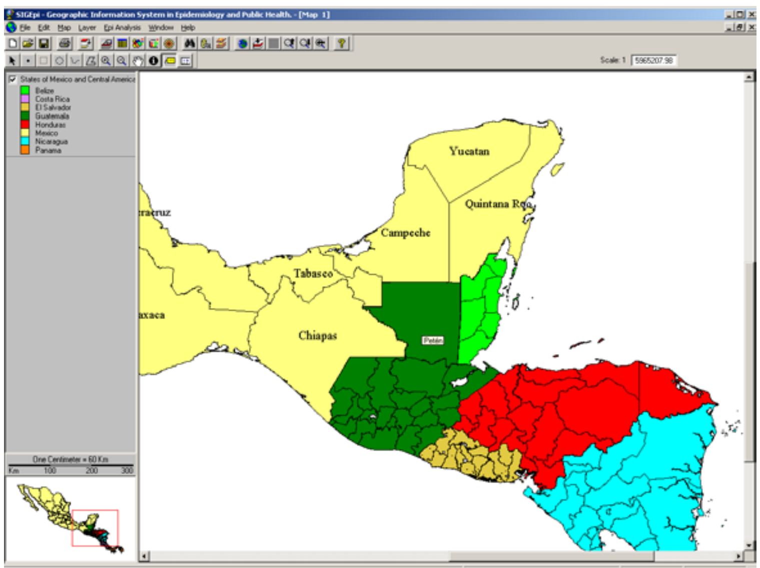
This causes the Thematic Map to be deleted from the Maps Window, and the Map Legend to be updated; by deleting the selected Layer, the information corresponding to the Thematic Map is deleted.

Working with Labels

In **SIGEpi's** Maps Window you can display **Labels** for the geographic units selected on a Map; these show the variable or attribute's content that you want to make visible on the Map display.

To display a Label:

1. Activate the Maps Window,
2. Select the Thematic Layer on which you want to place Labels,
3. Select a Variable ToolTip, and a variable to show the content as a Label,
4. ***Click*** the **Label** button from the spatial selection Tools,
5. Move the cursor over the Map in the Maps Window and ***click*** the geographic units for which you want to show the Labels.

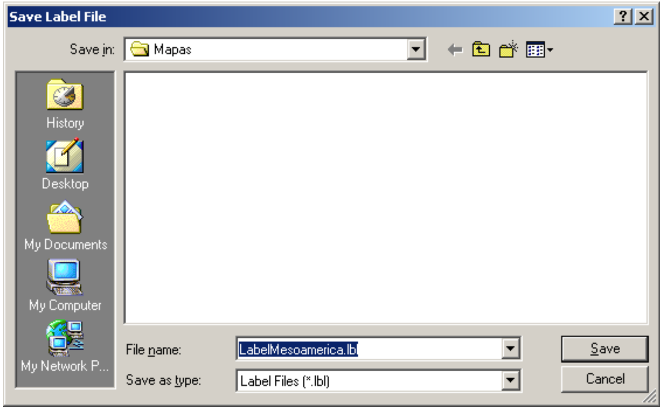


Example Maps Window. Labels displayed for geographic units of a region.

Saving a Label File

To Save a Label File:

1. Once the desired labels are placed on the Map, by ***clicking*** each geographic unit of interest,
2. Select the options **Labels / Save Label File...,** from the **Map** menu on the Menu Bar,
3. The corresponding Dialogue Box is displayed, and you should name the Label File to be created.



Dialogue box. Save Label File.

As a result, a Label File is created, which you can display at any time while working with SIGEpi, even when you are working with another Thematic Layer.

Showing a Label File

Once the Label File is created and saved, you may want to clear the Labels from the Map to continue with another type of data analysis. In any case, you can again display Label File you created at any time.

To display a Label File:

4. Activate the Maps Window and select the Thematic Layer that the Label File you want to show belongs,
5. Select the options **Labels / Open Label File**..., from the **Map** menu on the Menu Bar,

As a result, the Labels available in the Label File are displayed for each geographic unit selected.

Clearing a Label File

During work with **SIGEpi's** Maps Window, you may want to clear the Labels you have placed on the different geographic units.

If you want, you can save the Label File and continue working with it later; this way you can again display any saved Label File.

To clear a Label File:

6. Activate the Maps Window, and Select the Thematic Layer for which you want to delete the Label File,
7. Select the options **Labels / Clear all Labels,** from the **Map** menu on the Menu Bar.

As a result, the display of Labels in that Thematic Layer is deleted from the Active Maps Window.

- If you did not save the deleted Labels as a Label File, you cannot show them subsequently.

Displaying a Map window

To display a Map:

1. Activate the **Project Window**,
2. **Double *click*** the name of the Maps Window you want to show, from the list of maps under the **Maps** component in the Project Window,
3. Or, Select the **Map** you want to show, by ***clicking*** the **name** of the Map, and
   a.) ***Click*** the **Show Window** button from the Toolbar that appears to the left of the Project Window
   b.) Or, Select the option **Show Map**, from the **Project** menu on the Menu Bar.

Copying a Map

To Save a Map individually:

From an active Map Window:

1. From the **Edit** menu, select the option **Copy Map**, which places a copy of the Map's image on the **clipboard** in **BitMap** or **Metafile** format, depending on the option selected.

This allows you to add the Map as a graphic image to any document you create, such as a report or document prepared in MS-Word. This image is not a spatial datum or georeferenced. This image remains on the clipboard only until another object is copied there or until you turn off your PC.

2. You may also want **to save** the Map as a file; to do this, select the option **Export Map Image As**..., from the **Edit** menu on the Menu Bar

This allows you to **save** the selected Map as a file that you can later insert in any document, just as you could do by selecting the option **Copy Map** and copying it to the clipboard. In this case, however, the image is stored on your PC.

3. And, you also can select the option **Export Geo-referenced Map Image**..., from the **Edit** menu on the Menu Bar

This allows you to **save** an image of your Map, conserving its geographic reference. This means that when it loads the image as a Layer of the Map, it is located in the geographic position that corresponds to it.

Working with the Data Tables and Map Layers

As it was mentioned earlier, cartographic data have two types of data, graphic elements that represent geographic objects, which are displayed in the map and attributes that describe its characteristics. The attributes are organized in a table where each row belongs to a feature in the layer and each column is a variable.

In most of the GIS applications in Public Health, we have to deal with health and disease data, and data from other sectors as demographic, socio-economic data and others that usually comes from disease surveillance systems, routine information systems or health surveys. A key step to spatially process and visualize these data in a map is to relate data tables to map layers. The condition to do this step is that each data table has a key variable or column which is standardized with a key attribute of the attribute table of the layer.

SIGEpi includes functions that allow you to relate and unrelated database tables to layers. Having data tables related to layers allows you to apply most of the spatial and analytical functions to your data besides to visualize then in thematic map.

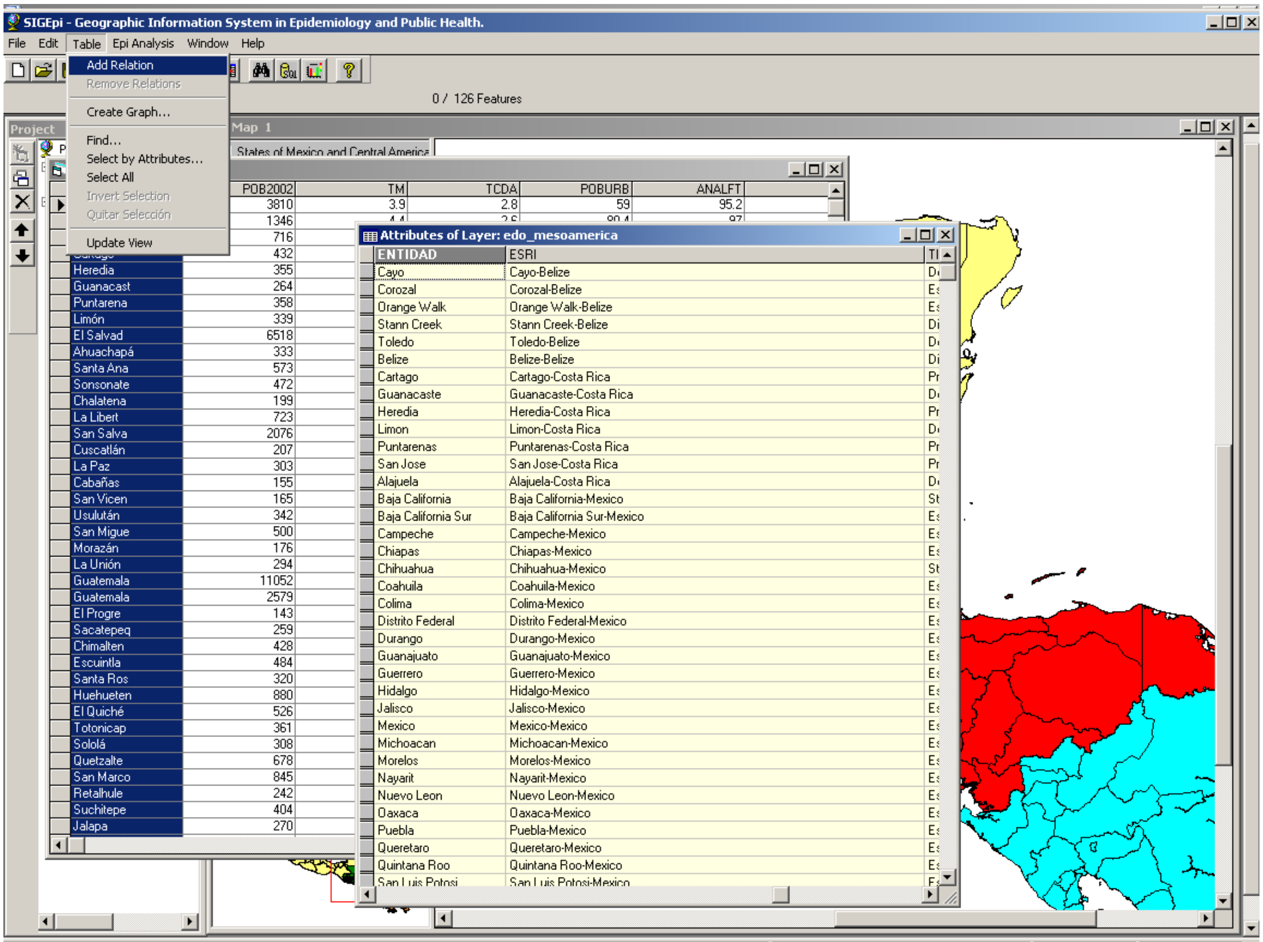
Adding Relationships of Data Tables to Layers

If you have data tables and need to relate them to map layers, you should ensure that data tables are part of an opened database in the Project and the layers are displayed in a map window.

Adding Relationships from data tables:

1. Open the **Database** containing the **Table** you want to add,
2. Or, Add/Import the **Table** to the Database already opened in the Project,
3. Make sure that the **Table** is indexed by the **attribute or column** of the Table that serves to join it with the **Attributes Table** of the Layer,
4. Activate the Maps Window and Select the Layer of interest,
5. Show the Attribute Table of the Layer,
6. ***Click*** the **column** or attribute of the joining data Table (the one having the index); this causes your selection to be highlighted in the data Table Window,
7. ***Click*** the **column** or attribute of the Attributes Table that serves as the link, in the same way the selection is highlighted in the Attributes Table Window,
8. Select the option **Add Relationship**, from the **Table** menu on the Menu Bar,

As a result, now the Attributes Table of the Layer contains **all** the columns of the related data table.



Example of Relationship between data tale and Attributes Table of a Layer.

- As many relationships as desired can be added to the Attributes Table; just select every Table of interest and follow the steps described above to add a new relationship. The final Attributes Table contains all the established relationships.

It is important to note that this **relationship** is valid only for the layer in which it was applied and if the layer is duplicated or copy and paste in the same map window, the relationship will be remained.
Using the option Save as Shapefile in Layer menu, it is possible to create a new layer with all attributes of the relationship.

Converting a Relationship to a Shapefile

In **SIGEpi** you can make permanent the relationship you established between a **Table** of the Database and the **Attributes Table** of the Layer; to do this, you must convert the layer to a new **Shapefile**. This process creates e new layer, which includes the attributes of the related table.

Removing Relationships

To remove relationships established between Data Tables and an Attributes Table:

1. Activate the Map Window, and select the Layer to which you want to remove the relationships,
2. Display the Attributes Table of the layer, or if it is already opened, activate it,
3. Select the option **Remove Relationships**, from the **Table** menu on the Menu Bar.

As a result, all the Tables related to the Layer are removed, the attribute table of the layer is updated deleting all the columns or attributes that were added by establishing the relationship.

Updating the display of the Attributes Table

You can update or refresh the view of the Attributes Table of a Layer, to show the changes made and update the Table's display.

To update the display of the Attributes Table:

1. Activate the Attributes Table associated with the Layer of interest,
2. Select the option **Update View**, from the **Table** menu.

Copying the Attributes Table

You can copy image of the display of the Attributes Table of a layer to the **clipboard**.

To copy an Attributes Table:

1. Activate the Attributes Table of interest,

2. From the **Edit** menu on the Menu Bar, select the option **Copy Table Image** to copy the image of the Table to the **clipboard**. This allows you to add the Table to a document or report.

Editing the values of an Attributes Table

You can edit the values of the **variables or attributes** of the Attributes Table.

Editing the values of the attributes in the Attributes Table:

1. Activate the Attributes Table of interest,
2. ***Click*** the content or value of the attribute or column that you want to edit,
3. From the **Edit** menu on the Menu Bar, select the option of interest:
   - **Cut**, if you want to delete the content of an attribute,
   - **Copy**, if you want to copy the current content,

- **Paste**, if you want to add the content you have selected to **cut** or **copy**.

You also can edit or **modify** the content of any attribute in the Attributes Table, by ***clicking*** the value you want to change and **typing** the new value.

These modifications are made permanent in the Table; they are updated in the original data source.

Chapter 5

Database

SIGEpi makes it convenient for you to work with the Project Database and Tables

Processing, analysis, and interpretation of data are some of the most complex tasks that are required for decision-making in health. Statisticians and epidemiologists who require training, tools, and methods for data and information analysis usually perform these tasks.

Epidemiological information is usually stored in different electronic formats, such as spreadsheets (Lotus, Excel), databases (DBase, Access, EpiInfo, MS-SQL), and others. To be useful for analysis, all these data must have been validated for consistency and standardized.

> In SIGEpi all the data you have and need to include in your analysis should be part of the Database. The SIGEpi Project allows you to manage and work with a Database in MS-Access standard format (.mdb); you can create, add and import all the tables you need or require for your work.

As part of the Database, a Structured Query Language (SQL) is available to process the Tables. SQL has proved very useful in managing Relational Database Management Systems and carrying out the most diverse operations on Tables.

Using the database procedures you can carry out multiple operations with your data:

- Create new columns or variables, as a result of calculating new values based on the contents of variables that already exist.
- Create levels of data aggregation, which means applying cluster functions to the values of the variables to get new values that summarize data at the level of units by which they are grouped.
- Combine two or more Tables into a new Table.
- Select the columns or rows of the Table that you want.
- Select a subset of data or objects from the Table.

For more information on "Relational Database Systems in Geographic Information Systems" consult the book "Geographic Information Systems in Health. Basic Concepts," SHA/PAHO.

What is a Database?

> *A* Database *is a relational data structure conceived as a set of relationships among the* attributes *or data represented in one or more Tables. The values of the attributes of each entity are stored in rows or records, the relationships are commonly established among the columns of the Table.*

A **data structure** has two basic characteristics: domains and relations. **Domains** contain the possible values of a variable (for example: Gender can be Feminine or Masculine, Educational Level could be Primary School, Secondary School, University, or Other). **Relations** represent the disaggregation of an object or entity into attributes or characteristics.

Table:MalaEdo

ENTIDADES	POB2002	TM	TCDA	POBURB	ANALFT	TMM	TMI	TBCASOS
Costa Ric	3810	3.9	2.8	59	95.2	36	10.8	-999
San José	1346	4.4	2.6	80.4	97	22	11.7	-999
Alajuela	716	3.7	3.2	36.2	94.3	28	9.2	-999
Cartago	432	3.7	2.9	66.2	95.4	0	8.8	-999
Heredia	355	3.7	3.7	68.2	96.8	60	9.4	-999
Guanacast	264	4.1	1.9	41.9	93.3	38	9.3	-999
Puntarena	358	3.6	1.9	40.1	92.2	61	13.5	-999
Limón	339	3.4	4.4	37.1	92.3	83	12.6	-999
El Salvad	6518	5.9	1.9	59	82.9	63.4	11	1384

Example of a Data Table of a Database

In a GIS, the rows usually represent the geographic units and the columns represent variables related to the geographic units (for example: population, socioeconomic levels, mortality, etc.)

Tables and Queries

When working with Tables in **SIGEpi** your data is organized in a tabular structure. You can select Data from different Tables and organize them into new Tables. You can add Data from different Tables to the Map, and symbolize, interpret, and analyze your data geographically.

Tables in **SIGEpi** are dynamic. This means that the Tables reflect the current status of the data source or Table on which they are based. If the data source changes, the Table based on that data automatically changes to reflect this.

You can create New Tables and incorporate their values directly. Some Tables can be edited, depending on the data source that they represent. You can create Queries based on data in Tables by establishing a SQL condition that you want met as a requirement for selecting or grouping data.

Data Sources for Tables

Working with **SIGEpi** you can import Tables from different data sources. You must take into account the diverse formats and different levels (regional, national, sub-national) of analysis in which the available epidemiological and health data are frequently found.

Data formats:

- DBase, FoxPro, Paradox, Btrieve, Excel

You can add DBase, Foxpro, Paradox, Btrieve, and Excel files as Tables in **SIGEpi**. These data sources may contain data that you want to use in **SIGEpi**. For example, these Tables could contain additional information on characteristics you want to show on a Map.

See Importing and Linking Tables and Linking the Table with the Thematic Layer.

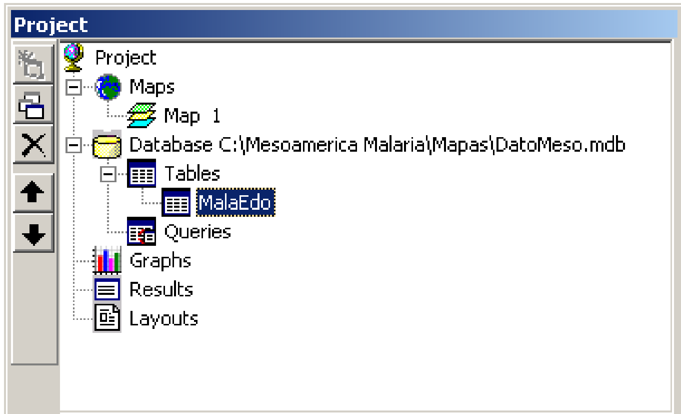
- SQL Database Servers

From **SIGEpi** you can connect to a database server and run a SQL statement to get records from a Table. **SIGEpi** stores the definitions in the SQL statements, not the records themselves.

See Importing Tables and Working with a Query.

Working with the Database

Working with a Project Database in **SIGEpi** you can organize your data in Tables and Queries. You can bring data from different sources and integrate and organize them in **SIGEpi's** Database. In this way, you incorporate your data in Maps for analysis and spatial representation.



Project window. Database.

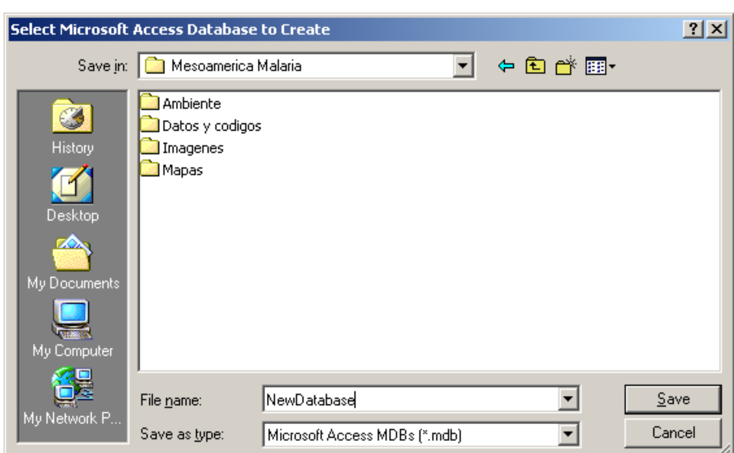
Creating a new Database

When you create a Database in **SIGEpi** you create a data structure containing all the Tables and Queries used for data analysis in the Project.

To create a Database:

1. Activate the Project Window,
2. Select the option **New Database...,** from the **Project** menu on the Menu Bar,
3. Or, ***click*** the Database component in the Project component structure, and ***click*** the **New Window** button on the Toolbar that appears to the left of the Project Window.

A **dialog box** is displayed allowing you to name the Database.



Dialogue box. Create Database.

SIGEpi creates the new database and displays its **name** in the  **Database** component of the Project Window. Now it is time to create/design or add the data tables to the database.

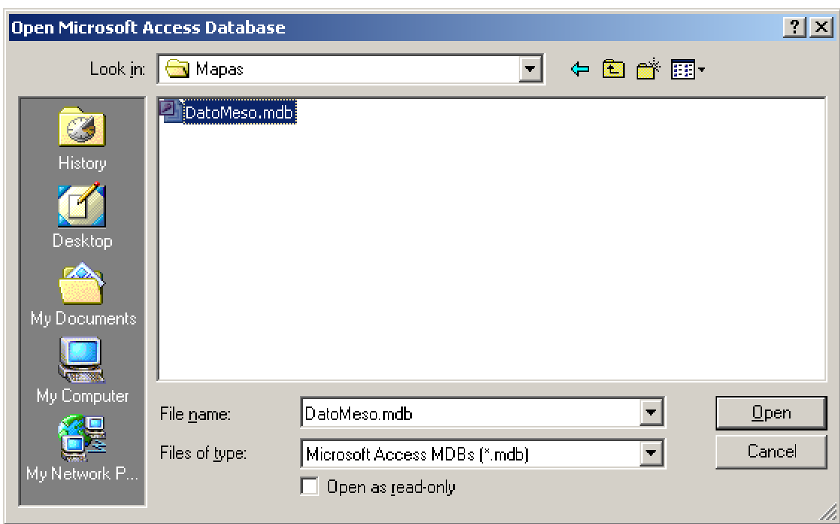
Opening a Database

You can add an existing Database to the Project in **SIGEpi.** This allows you to work with the data it contains and include them in the spatial analysis.

To open a Database:

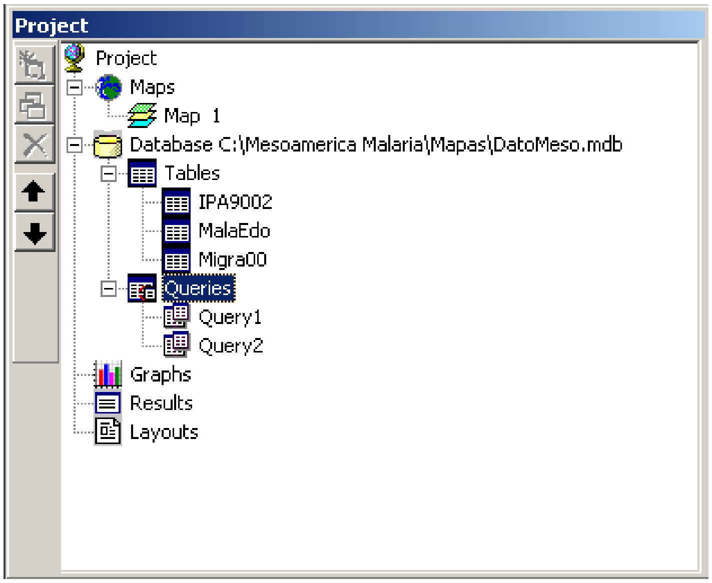
1. Activate the Project Window,
2. Select the option **Open Database**..., from the **Project** menu on the Menu Bar.

The corresponding **dialog box** is displayed, allowing you to select the **location** and **name** of the Database to open.



Dialogue box. Open Database.

This process opens the database and includes it to the Project. To view the **names** of the **Tables** and/or **Queries** contained in the opened Database, **click** the Tables and/or Queries component of the Project Window.



Project window. Open Database with Tables and Queries.

Closing the Database

While working in **SIGEpi** you may want to **close** the Database with which you are working, either because you have concluded your analysis, or because you want to open another Database.

To close a Database:

1. Activate the **Project Window**
2. Select the option **Close Database**, from the **Project** menu.

This closes the current Database, and clears the list of Tables and/or Queries shown in the Project Window.

Managing Tables

A Table allows you to work with data organized in a tabular data structure within the Database. You can integrate data from different data sources and organize them as Tables of the Database.

It is important to note that you can group together all the data that you have and need to analyze; add the Tables you want to an existing Database; or create a new Database and incorporate the Tables of interest from several data sources.
See: **Importing and Linking Tables**.

Creating and Designing a Table

Working with the Database in **SIGEpi**, you can **create** a new Table; this adds a new Table structure to the Project Database. Or you can **modify** the structure of any Table in the Database by selecting the Design Table option.

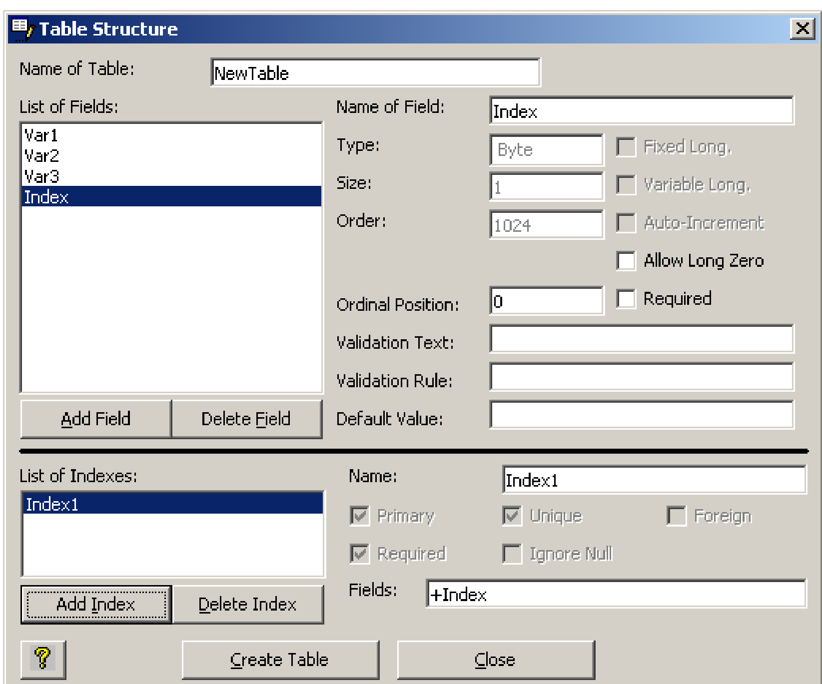
To create a Table in the Database:

1. Activate the Project Window,
2. Click the Tables component in the Database; there must be a Project Database open to be able to create and add a Table.
3. a.) ***Click*** the **New Window** Toolbar, which appears to the left of the Project components list,
   b.) Select the option **Create Table...,** from the **Database** menu on the Menu Bar, or from the menu displayed by **right *clicking*** the name of a **Table** in the Database.

To design (modify) a Table in the Database:

1. Activate the Project Window,
2. ***Click*** the **name** of the Database Table that you want to modify the structure,
3. a.) Select the option **Table Design...,** from the **Database** menu on the Menu Bar, or from the menu displayed by **right *clicking*** the name of the Database **Table**.

In both cases a **Dialogue box** is displayed, which allows you to define and/or modify the characteristics and structure of the Table you are creating and/or designing.



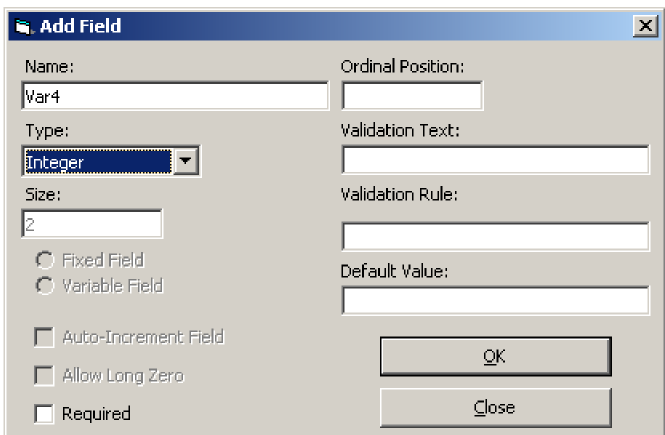
Dialogue box. Create Table.

Creating a Table:

1. Type the **name** of the Table to create,
2. ***Click*** the **Add Field** button to add new fields or attributes to the Table (see the Add Fields dialog box),
3. ***Click*** the **Delete Field** to delete a field or attribute from the Table,
4. ***Click*** the **Add Index** button to create an Index for the Table (see the Add Index dialog box),
5. ***Click*** the **Delete Index** button to delete a Table index from an existing table,
6. ***Click*** the **Create Table** button to create the defined Table.

To Add Fields to the Table:

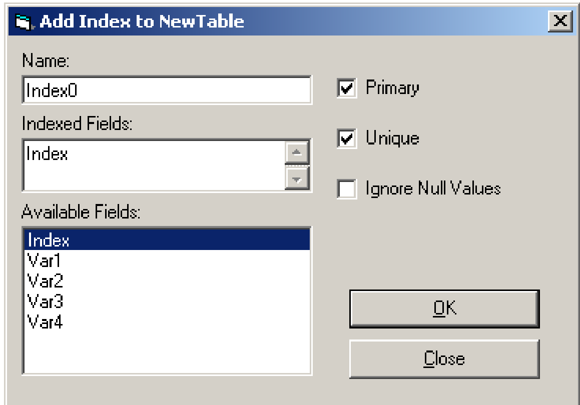
1. Type the **name** of the field or attribute of the Table,
2. Select the **type of datum** the field will contain,
3. Define the **size** of the datum depending on the type selected,
4. Define the characteristics and **validation rules** for the data,
5. ***Click*** the **OK** button to add the defined field to the data Table.



Dialogue box. Add a field to the Table.

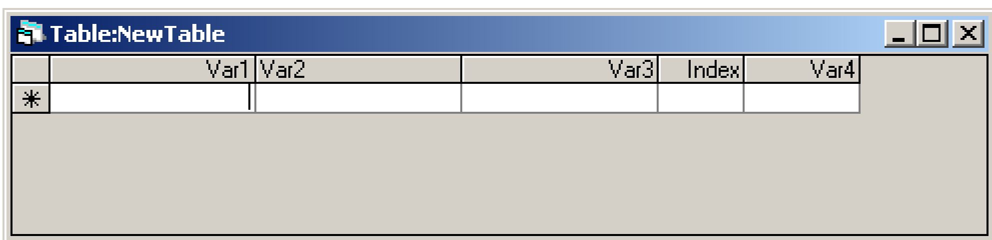
To create an Index:

1. Type the **name** for the Table index you are creating,
2. **Select** the field or fields to index or organize the index. To do this ***click*** the **name** of the field that you want in the list of fields and appears displayed in the lower-left part of the dialog box,
3. Select the **type of index** you want to get, by ***clicking*** the rectangles that define the types of indexes,
4. ***Click*** the **OK** button to create the defined index.



Dialogue box. Create Table Index.

The Table created is added to the Database and is displayed in the list of Tables that make up the Database. You can edit each field or attribute in the Table created to add data or values for that field.



Example of data Table resulting from creating a Table.

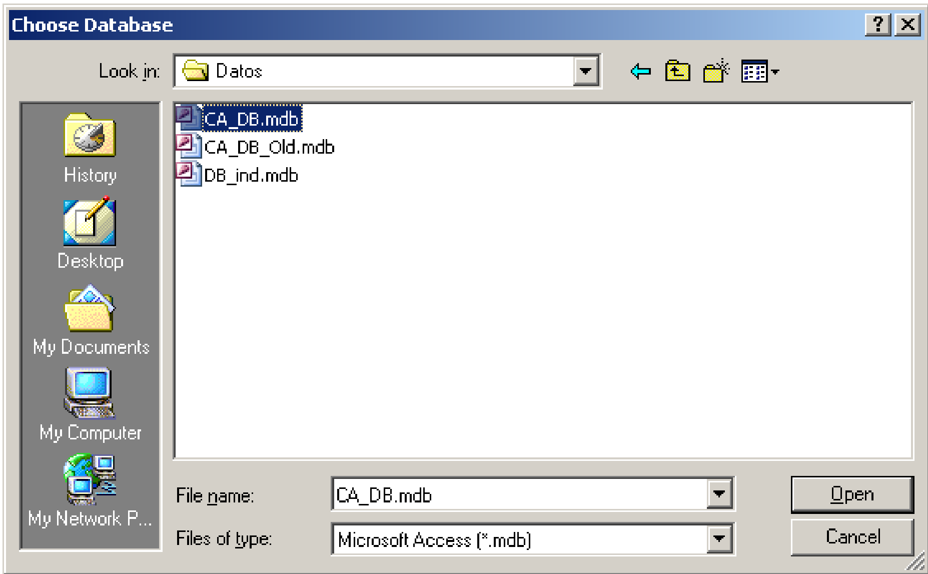
Importing and Linking External Tables

Usually we have data in tables of other formats than MS Access, e.g. dBase, Excel, etc. Those are considered external tables because they are not included in the database. SIGEpi allows you to add/import external Tables to the Database, this way they are accessible the data processing in a SIGEpi Project.

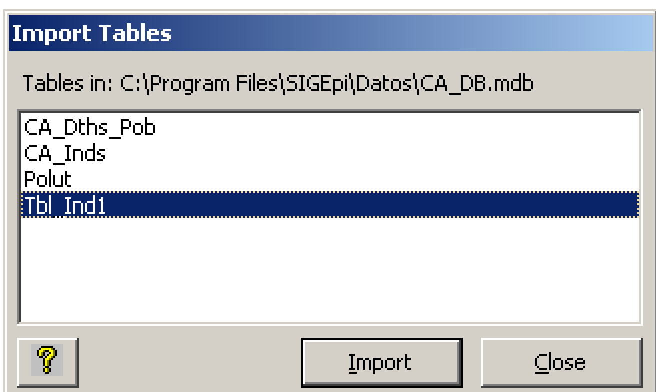
Importing and Linking Tables:

1. Open or create a **Database** in the Project, if you have not already done so,
2. Select the option **Add External Table / Import**..., to import a Table into the Database, or, if you only want to link the data source to the Project Database, you can **Add External Table / Link**... do this from:
   a.) the **Database** menu on the Menu Bar,
   b.) Or, the menu displayed by **right *clicking*** the **Tables** component in the component list in the Project Window.
3. In the **Dialogue Box** displayed, select the **type of data file** that you want to add,
4. Select the **Table** or **Database** that containing the **Tables** that you want to add,

5. ***Click*** the **Open** button to **add** the selected **Table** to the Project Database, or **to open** the **Database** containing the **Tables** that you want to add,
6. Select the **Tables** to add in the following **Dialogue Box**; if you have previously selected a **Database** as the data source**,**
7. ***Click*** the:
   a.) **Import** button to import the selected **Table**,
   b.) **Link** button to link the **Table** to the Database,
   You can repeat this action as many times as you have Tables you want to import.



Dialogue box. Select Database.



Select Tables to Import.

As a result, the **names** of the selected **Tables** are added to the **Database** and are displayed in the list of Tables of the Database component of the Project. You can work with the data contained in the Tables for spatial data analysis and data visualization.

You can display and edit the Imported and Linked Tables if you want to do so.

- **Linked** Tables have only their names in the list of Tables in the Database, their content remains in the original data source. **SIGEpi** maintains a dynamic link to these Tables; if the data source is updated, the data's display in SIGEpi is automatically updated. For example: you can edit or modify the content of these Tables using other tools such as MS-Access or MS-Excel; changes are reflected automatically in the Table linked to the Database in **SIGEpi**.

Note: The linked tables can not be related to layers in a Map. They only can be used to data processing and generate new variables and tables.

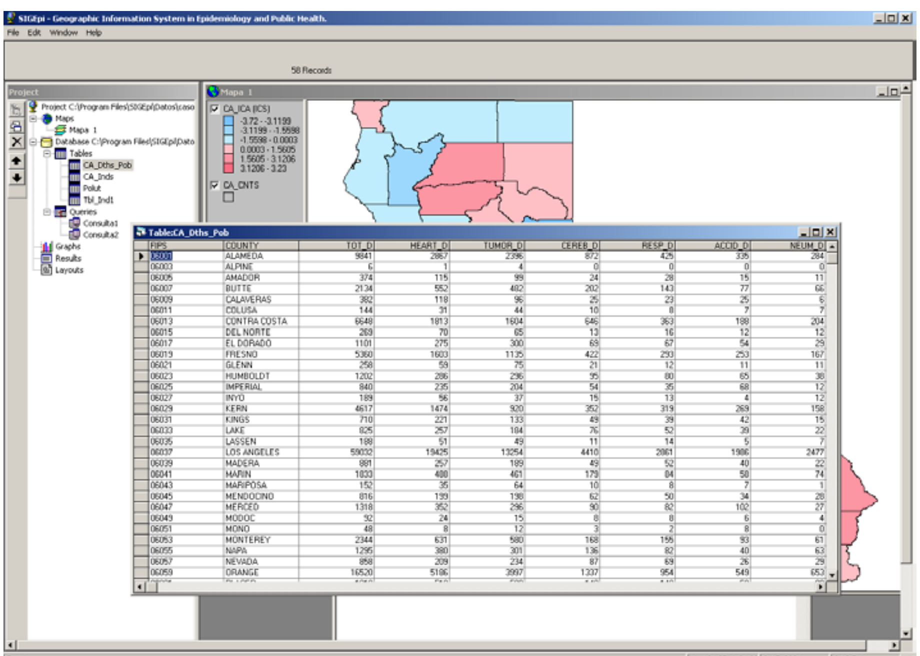
Displaying and Editing Tables

The Tables that make up the Project Database in **SIGEpi** can be displayed at any time; you also can edit or modify the values of the data they contain.

To show or display the Tables:

1. Activate the Project Window,
2. **Double *click*** the **name** of the Table that you want to display from the list of Tables that make up the **Database** component in the Project Window,
3. Or, Select the Table that you want to display, by ***clicking*** the **name** of the Table, and,
   a.) ***Click*** the **Open** button from the Toolbar that appears to the left of the Project Window,
   b.) Select the option **Open Table**, from the menu displayed by **right *clicking*** the **name** of the selected Table.

As a result, the structure and content of the selected Table are displayed.



Window with Data Table displayed.

To edit or modify the values of the data in Tables:

1. **Show** or display the **Table** that you want to modify the values, as previously explained,
2. ***Click*** the grid-cell of the Table that you want to modify,
3. **Type** the new value,
4. ***Click*** another row or record in the Table or the **Close** button if you want to finish editing,
5. ***Click*** the **Yes** button to make your changes permanent.

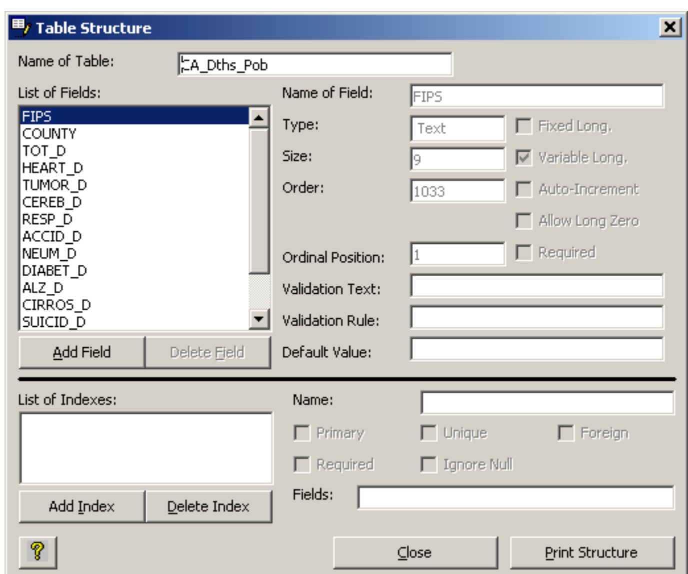
When you select a grid cell in a Table to modify, this enables you to edit all the values of the row or record corresponding to the selected grid-cell; you have only to ***click*** the grid-cells in the row that you want to modify and **type** the new value; to finish editing, ***click*** the same way on another row of the Table and on the **Yes** button to update the Table.

Creating an Index in a Table

In **SIGEpi** you can create as many indexes as you require to organize the data, using the values of the fields or attributes of the data Tables. You may want to organize your data by a particular attribute, which serves as link between the current Table and the Attributes Table of a given Thematic Layer. This allows you to represent and analyze its data spatially.

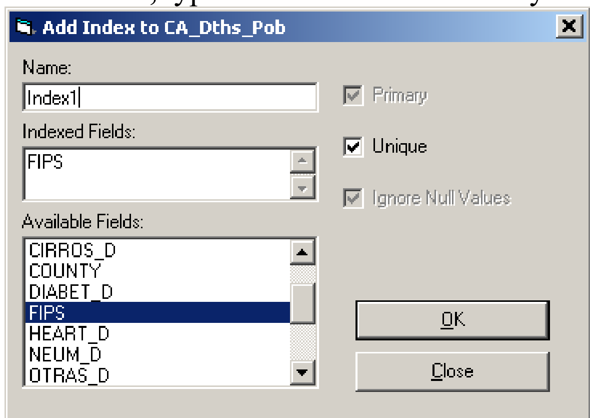
To create an Index:

1. Open or Create a Database, if you have not already done so,
2. a.) Select the data Table, ***clicking*** the **name** of the Table in the list of Tables of the Database, or
   b.) Create, or Add the data Table for which you want to create an index.
3. Select the option **Design Table**..., from the **Database** menu on the Menu Bar, or from the menu displayed by **right *clicking*** the selected Table.
4. The Table Structure **Dialogue Box** is displayed; ***click*** on **Add Index** button.



Dialogue box. Designing a Table.

5. Using the dialogue Add Index to … **,** type the **name** for the index you are creating.



Dialogue Add Index to Table

6. **Select** the fields to include in the index. To do this, ***click*** on the **name** of the fields in the list of fields displayed in the lower left of the dialog box.
7. Select what **type of index** you want by ***clicking*** the checkbox that define the index types.
8. ***Click*** the **OK** button to create the defined index.

To delete an Index:

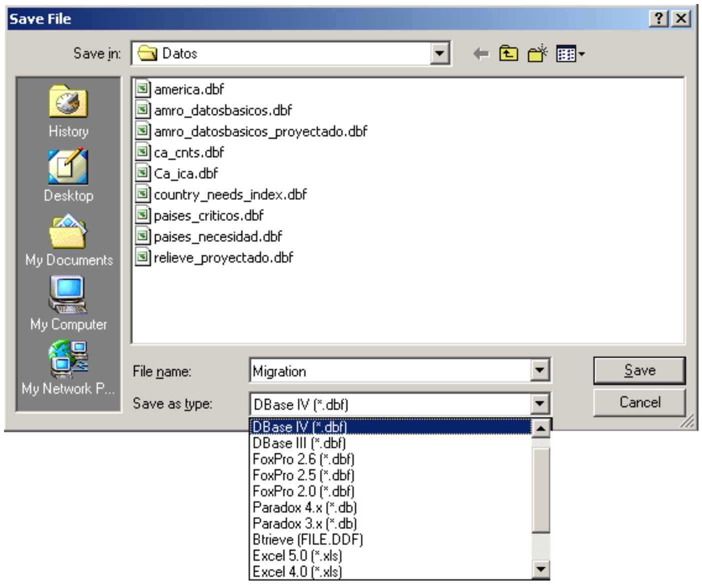
1. The index should already exist.
2. Select the **Index** that you want to delete from the **List of Indexes** that appears in the lower left of the Design Tables **Dialogue Box** that is displayed,
3. ***Click*** the **Delete Index** button.

Exporting Tables

Working with the Project Database in **SIGEpi** you can export your data Tables to other formats and store them as External Tables.

To export a Table:

4. Select the **Table** that you want to export, by ***clicking*** the **name** of the Table in the list of Tables in the Database,
5. Select the option **Export Table**..., from the **Database** menu on the Menu Bar or from the menu displayed by **right *clicking*** the **name** of the selected Table,
6. In displayed **Dialogue Box,** select the data **format** in which you want **to export** the Table,
7. In the displayed **Dialogue Box,** select or type a **name** for the Table to export.



Dialogue box. Name the Table to Export.

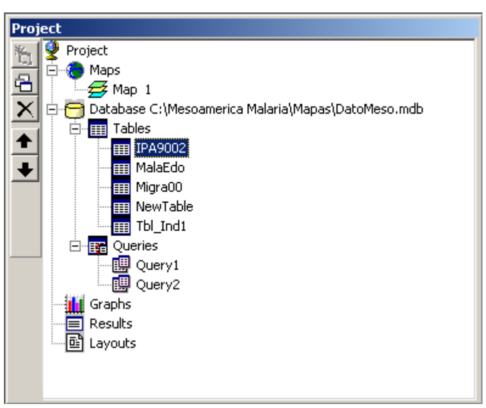
8. ***Click*** on the Save button to save the exported Table.

Deleting a Table

SIGEpi allows you to delete Tables from the list of Tables that make up the Project Database.

To delete a Table:

1. Select the **Table** that you want to delete, by ***clicking*** the **name** of the Table in the list of Tables in the Database, and,
2. a.) Select the option **Delete Table**... from the **Database** menu on the Menu Bar or from the menu displayed by **right *clicking*** the **name** of the selected Table,
   b.) Or, ***click*** the **Delete Window** button on the Toolbar displayed on the left in the Project Window,
3. ***Click*** the Yes button to delete the selected Table.



Managing Queries

You can **process**, **summarize,** or **group** the data you have in Tables, based on the particular values of the fields or attributes of the Tables. The results can be stored in a Query or in a Table.

Summarizing creates a row or *record* in the new **Table** or **Query** for every unique value of the selected fields or attributes. Each *record* is the result of a selection process applied to the data based on a defined **SQL statement**. Each *record* shows a summary of the *records* or rows that you group from the Tables or data sources that are the source of the Query. To process, summarize, or group your data, you can select several fields or attributes from different Tables or data Sources.

> Note: If you have information or data in data sources that are not included in the Tables of the Database in **SIGEpi** and want to use them in designing a Query, you must first import or link them to the **SIGEpi** Database.

When you create a **Query** and need to represent the data spatially and/or to apply procedures for spatial analysis, you must first export the Query to a **Table** in the Database. This way you can add the Table to the Attributes Table of a Thematic Layer of interest for visualization and spatial analysis. Grouping and summarizing data in Tables and relating the result table to a Layer is a useful way to incorporate statistical data into a Map.

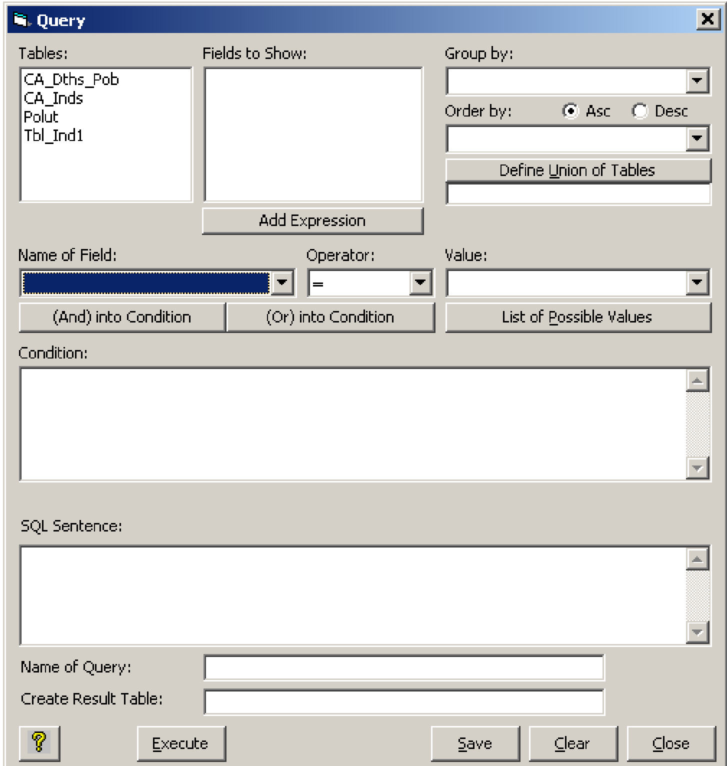
Creating and Designing a Query

In **SIGEpi** you can work with information Based on data located in Tables stored in different data sources, and create a Query that contains information resulting from applying a **SQL statement** to the data.

Follow the next steps to create a Query:

1. Activate the Project Window,
2. Select the option **Create Query**..., from:
   a.) the **Database** menu,
   b.) Or, from the menu displayed by **right *clicking*** on **Tables** node in the Project Window,
   c.) Or, from the menu displayed by **right *clicking*** on **Queries** node in the Project Window,
3. Select option **Create a Query…**

The dialogue Query is displayed. This dialogue facilitates the definition of the SQL sentence.



Dialogue Query

4. In the list of **Tables**, at the upper left, you must select the **Tables** that you want to include in the SQL sentences; to do this, ***click*** the **name** of each table,

5. Once you have selected the **Tables,** the list of all **fields** or attributes of Tables are displayed in the **Fields to Show** list box; so now you can select the **fields** that you want to include in the query ***clicking*** on the **name** of the fields,

6. In the **Group by** edit box**,** you should select the **field** or variable that serves as a criterion for aggregation--that is, the variable by which the data from the original Table(s) is/are grouped to obtain the desired result. To select, ***click*** the desired field,

7. In the **Order by** edit box, select the **field** or attribute that is used as the criterion for arrangement, that is, the attribute by which the data is organized. To select, ***click*** the desired field or attribute,

8. In the **Join Tables** edit box (you should have previously selected at least two Tables in the **Tables** edit box to create the Query), you should:

   a.) **select** the Tables to be joined from the source Tables for the Query by ***clicking*** the **names** of the Tables of interest.

   Note: Joins can be done by selecting two Tables each time; this means that if you want to join three or more Tables, you must first join two Tables and then one of them with the third and so on.

b.) **select** the join **field** or attribute, the variable with which the Tables are joined. For each Table selected, an edit box is displayed that contains all the fields in the Table; ***click*** the join **field** or attribute for **each** Table.

c.) ***click*** the button **Add Join to Query** so that you add the Table join you defined to the SQL statement you are designing,

d.) if you ***click*** the **Clear All Joins** button, all the joins are deleted,

e.) ***click*** on the **Close** button to close the dialog.

9. In the **Field Name** edit box, continue selecting the **fields** or attributes used to create the **SQL condition** that the data should satisfy. To do this, **click** the **field** of interest from the drop-down list containing all the fields of the Tables included in the Query,

10. In the **Operator** edit box, select the arithmetic or logical **operator** by ***clicking*** the desired operator in the drop-down list,

11. In the **Value** edit box, **type** the comparison **value** that the data should satisfy, or **select** the comparison **value** from the list of possible values of the displayed variable by ***clicking*** the **List of Possible Values** button,

12. ***Click*** the **(AND) in Condition** button, if you want to add the logical operator **.AND.** to the SQL condition you are designing, that is, the data is selected that meet both one condition and the other, for example that: Area<5,000 .AND. Population>10,000

13. Or, ***Click*** the **(OR) in Condition** button, if you want to add the logical operator **.OR.** to the SQL condition you are designing, that is, the data is selected that meet either one condition or the other, for example that: Area<5000 .OR. Population>10 000

14. Repeat steps 9 to 13 to add successive conditions to the SQL condition or statement that you are designing,

15. In the **Condition** edit box, the SQL conditions constructed in the previous steps (9-13) are displayed, producing a simple or complex condition formed by arithmetical expressions connected with logical operators. In this box, you can also **edit** or modify the condition by changing operators, variables, and/or adding parentheses.

This condition acts as a filter to select a data subset from the included Tables.

16. In the **SQL statement** edit box, **all** the SQL conditions or statements that have been constructed so far are displayed; this includes the **names** of the included Tables, the **fields** that are part of the Query, and the **conditions** that the data should satisfy. In this box you also can **edit** or modify constructed statement,

17. In the **Query Name** edit box, type the **name** for the **Query** you constructed; this name is displayed in the list of Queries of the Project Database,

18. If you also want to create a **data Table** containing the results of the constructed SQL statement, type the **name** for the Table in the **Create Resulting Table** edit box. This Table name also is displayed in the list of Tables of the Project Database.

    Note: It is important to know that a Query is not a table, it is a sentence in SQL language that express a command to be executed by the database engine. So, a Query always reflects the current status of the data in the database and displays a data view, not a table.

Click on the **Save** button to finish saving the defined SQL statement and add the new **Query** to the list of Queries of the Project Database,

19. ***Click*** the **Run** button if you want to execute the Query (SQL sentence) at this time; as a result, a table view window is displayed that shows the resulting Query,

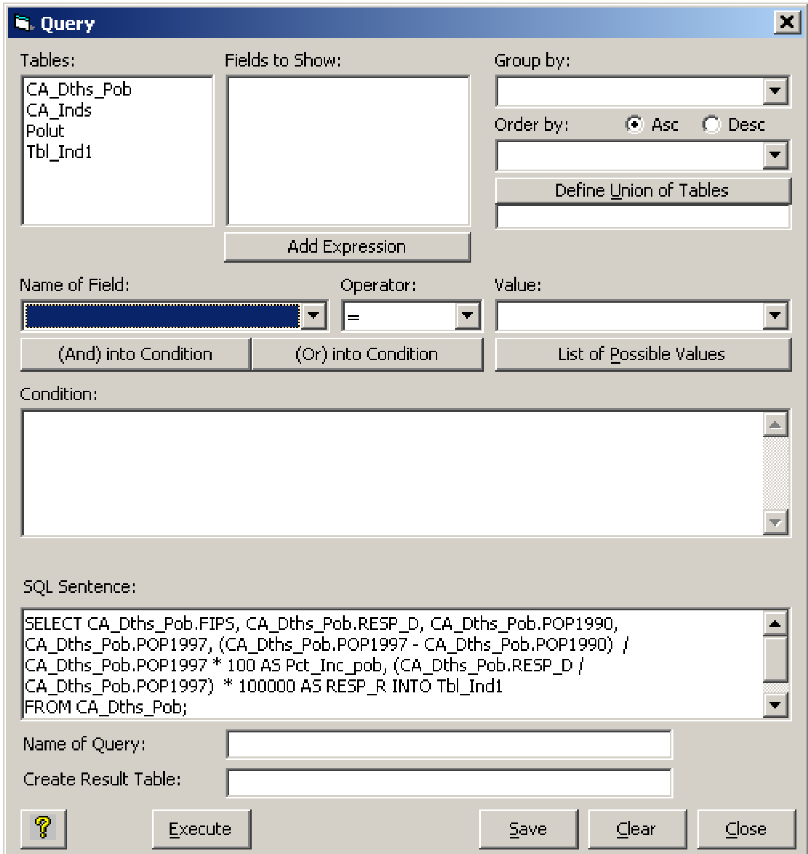
    Note: You may not want to run the SQL sentence at this time; in any case, whenever you display or show the Query, the SQL sentence is run.\

20. ***Click*** the **Clear** button if you want to clear the defined SQL statement; this allows you to fill in the blank spaces in all the edit boxes that make up the dialog box.
21. ***Click*** the **Close** button to finish working with the Dialogue Box.

It is also possible to edit or modify an existing Query

Follow the next steps to modify a Query:

1. Activate the Project Window,
2. Right clicking on the **Query**, from the popup menu select the option **Build Query…** As result the **Query** dialogue will be displayed.



Dialogue box. Design Query.

Using the Query dialogue you can modify the definition of the selected query in a similar way it was explained in the previous section.

Showing Queries

In **SIGEpi** you can display Queries included in the Project Database whenever you want, just as we saw with Tables.

To show or display a Query:

1. Activate the Project Window,
2. **Double *click*** the **name** of the Query that you want to show, from the list of **Queries** in the  **Database** component of the Project Window,
3. Or, Select the **Query** that you want to display by ***clicking*** the **name** of the Query, and
   a.) ***Click*** the **Show Window** button on the Toolbar that appears to the left of the Project Window,
   b.) Or, Select the  **Show Query** option from the menu displayed by **right *clicking*** the **name** of the selected Query.

As result, the SQL sentence of the selected Query is executed and the data view is displayed.

Sentencia SQL Select admin1.ID_,admin1.NAME1_,admin1.NAME2_,admin1.PARTS_,admin1.POI...

ID_	NAME1_	NAME2_	PARTS_	POINTS_	LENGTH_
0105006	0105006	Rangamati	1	663	566.4666
0105008	0105008	Sylhet	1	387	418.1115
0105013	0105013	Rajshahi	1	331	331.3611
0105015	0105015	Rangpur	1	481	425.8958
0105016	0105016	Dinajpur	1	516	389.1936
0105001	0105001	Khulna	10	788	592.6174
0105005	0105005	Chittagong	13	934	780.1383
0105003	0105003	Barisal	30	1132	811.7689
0105018	0105018	Faridpur	20	847	649.6917
0105007	0105007	Tippera	12	507	398.6922
0105017	0105017	Pabna	6	389	375.9505
0105019	0105019	Dacca	44	1099	768.9002

Data View based on a Query or SQL sentence

Deleting a Query

SIGEpi allows you to delete Queries from the list of Queries included in the Project Database.

To delete a Query:

1. Select the **Query** that you want to delete by ***clicking*** the **name** of the Query from the list of Queries of the Database, and
2. a.) Select the option **Delete Query...,** from the menu displayed by **right *clicking*** the **name** of the selected Query,
   b.) Or, ***click*** the **Delete Window** button on the Toolbar displayed to the left in the Project Window,
3. ***Click*** on Yes button to so that the selected Query is deleted.

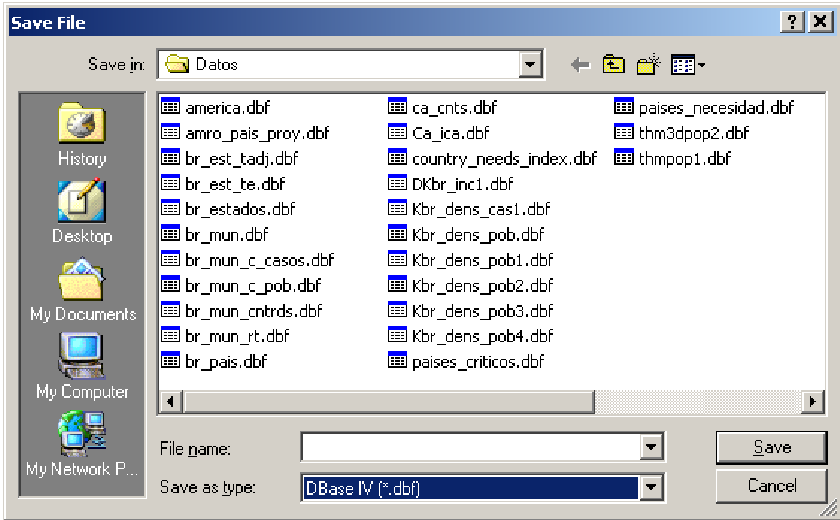
Exporting the Query to a Table

SIGEpi allows exporting a Query to an external Table. Usually a query produces temporary views of the data, although some types of queries could create a new table in the database.

Exporting a query means to generate a permanent external table as result of the query. This is useful if you want to have the data in a specific file format for using with other software.

To export or copy a Query to a Table:

1. Activate the Project Window,
2. Select the Query that you want to export by ***clicking*** the **name** of the **Query** in the list of Queries included in the **Project Database,**
3. ***Click*** the option **Export SQL**... from the menu displayed by **right *clicking*** the selected **Query,**
4. In the **Dialogue Box** shown, select the format for the table; select the folder and set a **name** for the **Table** you are creating.



Dialogue Save File

As result, a new file is saved in the selected folder. This file is available to be used in other applications.

Chapter 6

Working with Graphs

SIGEpi allows you to create several types of Graphs that help interpret your data

The **Graphs** component in the **SIGEpi** Project allows you to create different types of Graphs with relative ease, letting you group your data by several indicators or patterns of health analysis.

Graphs are frequently used in epidemiological data analysis; presenting data in Graphs makes it possible to identify trends and other characteristics at a glance.

With **SIGEpi** you can create a Graph at any time during a work session. To do this, there are two requirements: at least one Map Window must be open, which can contain one or more layers; and one of the layer must be selected/active.

- In **SIGEpi** the link among the Maps, Graphs, and Attributes tables is dynamic; this means that you can ***click*** on an element of the graph and the corresponding geographic unit of the map will be highlighted, as well as the records of the attributes table.

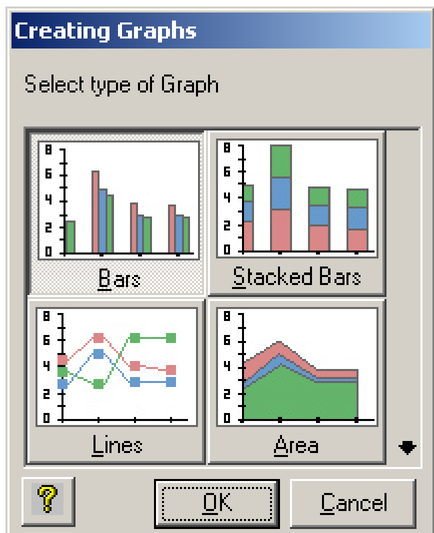
In **SIGEpi** you can create seven types of Graphs: **Bar Charts**, **Stacked Bar Charts**, **Line Graphs**, **Area Graphs**, **Pie Charts, Scatter plots and Box plots**.

Creating a Graph

To create a Graph:

1. <u>Activate</u> the **Maps Window**, and <u>Select</u> the **Thematic Layer** for which you want to create the Graph,
2. Select the option **Create Graph...,** from the **Layer** on the Menu Bar or ***click*** the **Create Graph** button on the Toolbar displayed below the Menu Bar,
3. You also can create a Graph by activating the Window of the **Attributes Table** associated with the selected **Thematic Layer** by ***clicking*** the **Attributes Table** button on the Toolbar, and selecting the **Create Graph**... option from the **Table** menu on the Menu Bar, or by ***clicking*** the **Create Graph** button on the Toolbar.

In each case, the **dialog box** for creating the Graph is shown on the screen.



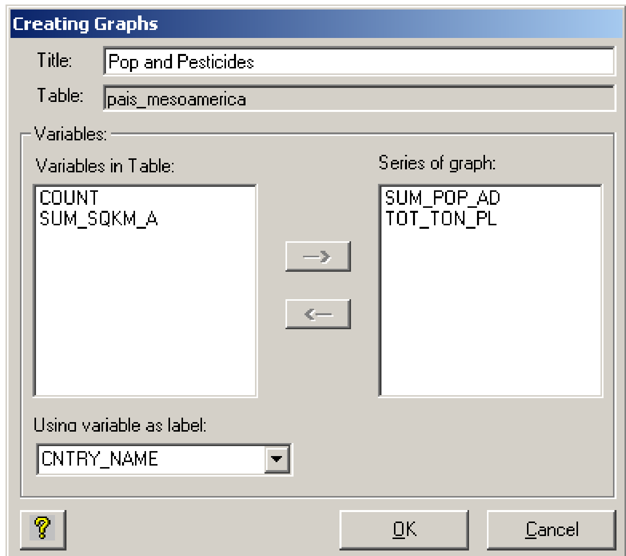
Dialogue box. Creating Graphs, selecting the type of Graph

In this dialog box you can select the Graph type by ***clicking*** on the box that identify each the graph you want and then ***clicking*** the **<u>O</u>K** button to confirm your selection and continuing.
Note that you can scroll up and down to visualize more types of graph.

The dialogue **Creating Graphs** is displayed after clicking on **<u>O</u>K** button. This dialogue allows selecting the variables to be used and other properties of the graph.

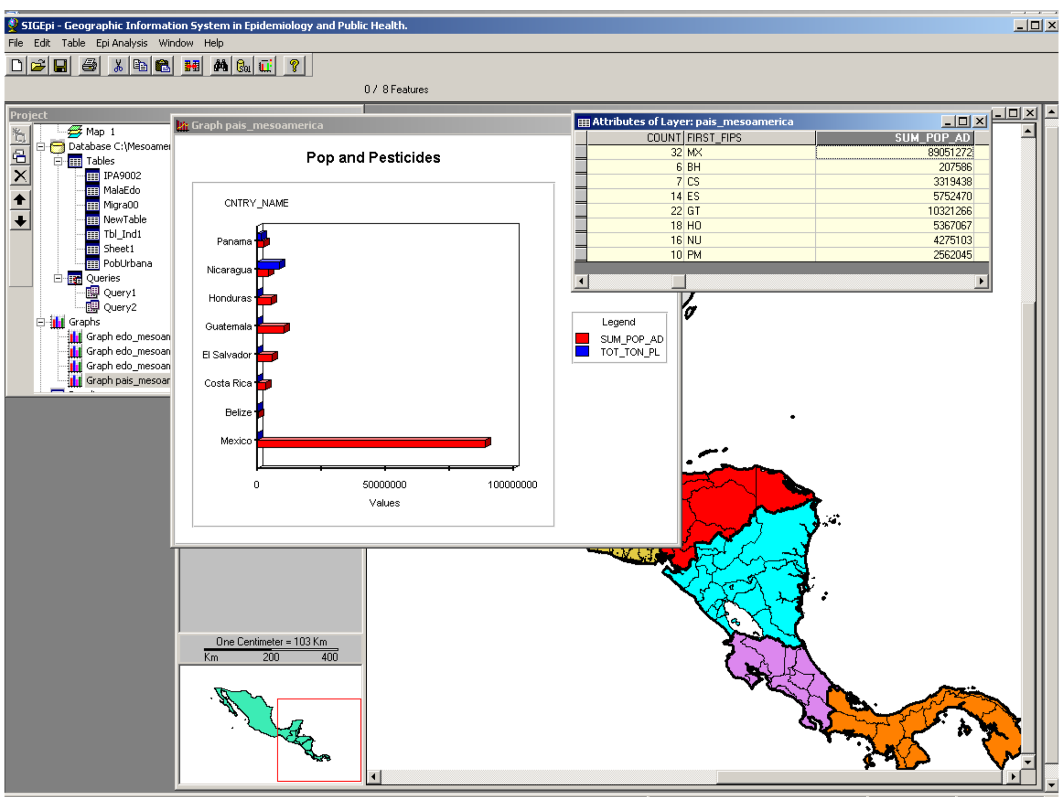
Steps to finish the creation of a graph.

1. First set a **Title** for the Graph,
2. Select the **variable(s)** to use in the Graph. From the list of variables in the Attributes Table of the selected Layer, which appears on the left of the dialog box, ***click*** on the name of the variable and using the corresponding arrow to move the selected variable to the right box of the window, where the variables used for making the Graph appear.
3. Select the **label** to display for the variables that you are graphing.
4. ***Click*** the **<u>O</u>K** button to make the Graph.



Dialogue box. Creating Graphs, defining the variables

As a result, you get a Graph Window with the graphic representation of the selected variables from the Thematic Layer.



Example of Screen with Maps Window, Table, and Graph

- The Graph Window can show the content of the displayed graphic element; this means that as you move the cursor over the Graph, depending on the position of the cursor, the content or value of the attributes being graphed is displayed.

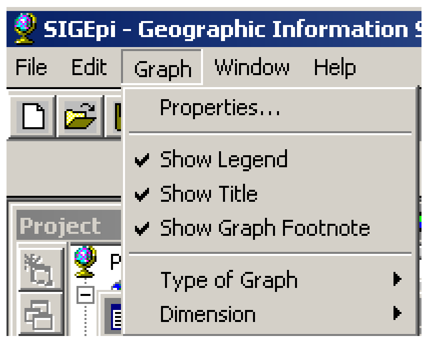
In the previous example, the value of the region's code and total urban population of a selected point of the Graph is displayed; you can also see the dynamic link between the Graph, geographic units of the Map, and records selected from the associated Attributes Table.

To see the types of Graphs that **SIGEpi** supports, see the topic **Types of Graphs** that is presented further.

Editing a Graph

When a Graph Window is the active Window, a **Graph** menu is added to the **SIGEpi** Menu Bar.

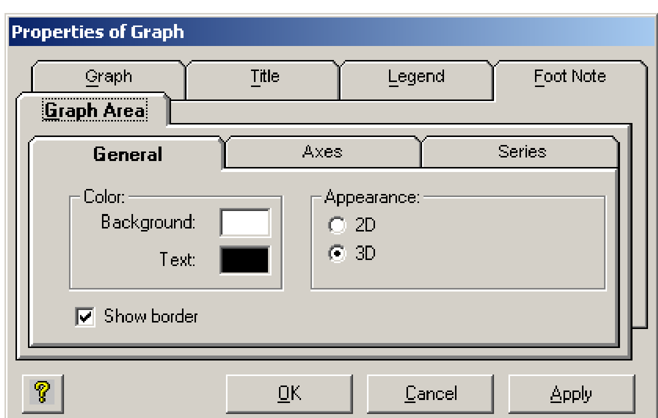
The **Graph** menu contains the options that allow you to modify the characteristics and properties of the Graph in the active Graph Window.



Graph Menu.

To edit a Graph:

1. When the Graph Window is active, select the option **Properties**...of the Graph from the **Graph** menu, or ***click*** the **Properties** button on the Toolbar,
2. and, you can choose to change the **type** of Graph that is being shown from the option **Type of Graph** on the Graph menu.



Dialogue box. Graph Properties.

This dialog box allows you to change the **name** and **title** of the Graph and determine the **style**, **type**, **color**, and font **size**.

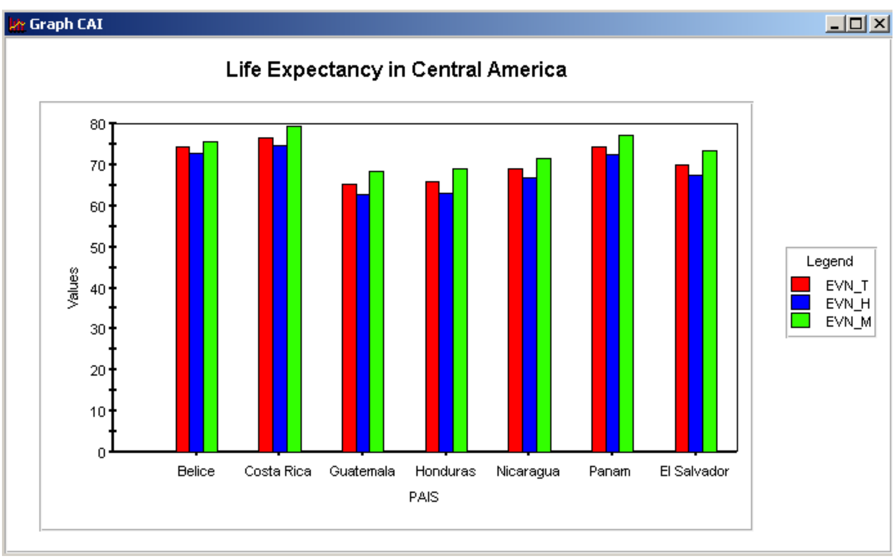
Types of Graphs

SIGEpi allows to create seven types of Graphs to represent the attributes of the cartographic layers. The types of Graph are:

- Bar Chart,
- Stacked Bar Chart,
- Line Graph,
- Area Graph,
- Pie Chart
- Scatter Plot, and
- Box Plot.

Bar Chart

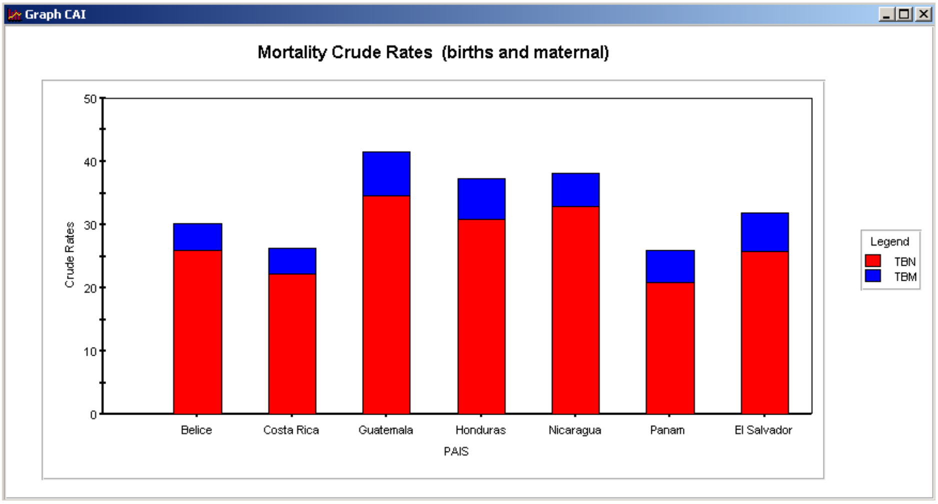
A **Bar chart** shows the differences between the values of the selected variables in the form of bars. The relative height of the bars represents the values of the variables for each geographic unit.



Example Bar Chart

Stacked Bar Chart

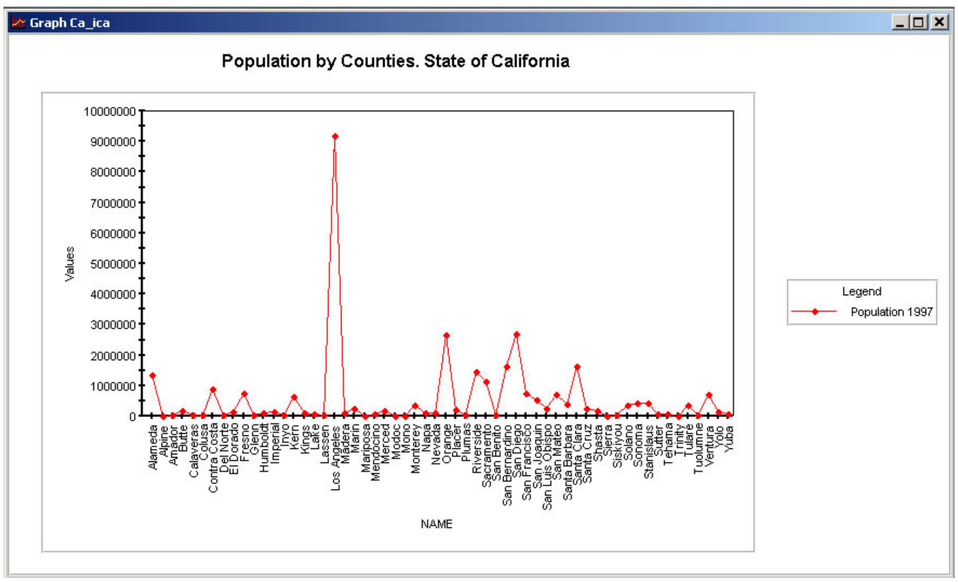
A **Stacked bar chart** is useful in representing the cumulative values of the selected variables by geographic unit. You should be careful in selecting the variables you want to represent, ensuring that the cumulative values are meaningful and do not measure different parameters that are incompatible.



Example Stacked Bar Chart

Line Graph

A **Line graph** presents the values of the variable(s) for each geographic unit. It is other way to visualize and identify the differences among the geographic units.



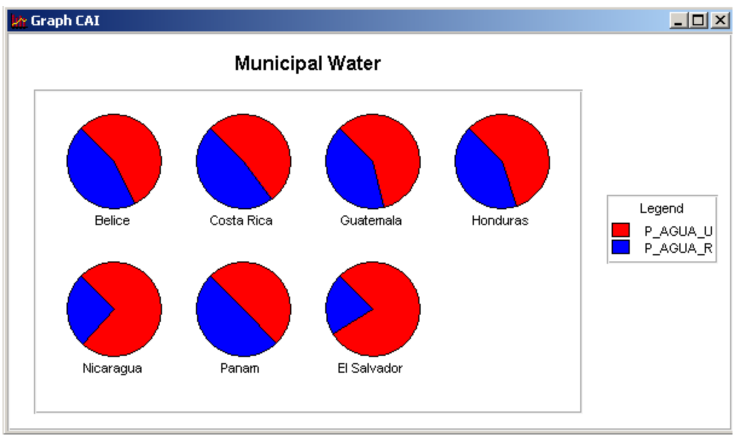
Example Line Graph

Area Graph

An **Area graph** is useful for indicating variations in the values of variable(s) among geographic units. It is a variation of the line graph that gives color to the area under de line.

Pie Chart

A **Pie chart** represents the relationship between the parts of a whole. It is useful for highlighting the significance of the information represented in each wedge of the pie in contrast to the rest. It requires more than one variable for its representation; the total of the selected variables make up the whole.

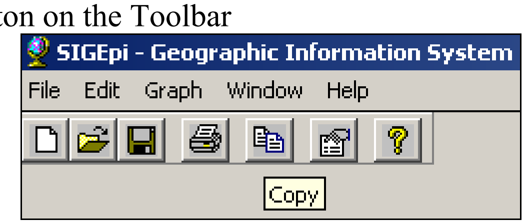


Example Pie Chart

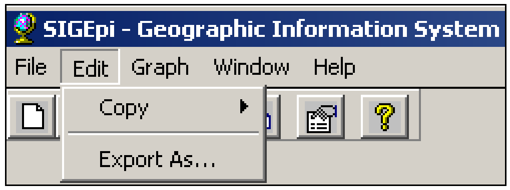
Copying a Graph

To Save a Graph individually:

1. From the active **Graph** Window, select the option **Copy** from the Edit menu on the Menu Bar, placing the image of the graph to the **clipboard** in **.bmp, .emf,** or **.wmf** formats,
2. Or, ***Click*** the **Copy** button on the Toolbar



3. You can also save your Graph as an independent file through the **Export as**... option on the **Edit** menu.


SIGEpi - Geographic Information System
File
Edit
Graph
Window
Help
Copy
Export As...

Analytical Procedures

7

SIGEpi offers a set of analytical methods and functions for the analysis of health data. Some complex procedures are presented in a simplify way taking into account the public health professional's point of view. These procedures are based on the foundation of the Epidemiology and they are oriented to the exploration, estimation and study of Public Health data. This Analytical component is which distinguishes SIGEpi from other GIS.

SIGEpi allows applying several statistical and epidemiological methods to the health data. This aspect strengthens the use of this system as analytical tool in Epidemiology and Public Health, upon making it possible to handle the elements of person, time, and place in an integrated approach.

The analytical procedures that are offered in SIGEpi go from descriptive statistical measures of non-spatial data up to procedures of exploratory analysis of spatial data. A set of key procedures for the Public Health are presented as the calculation, standardization and transformation of rates and proportions, the construction of composite indexes, the identification of critical areas and population groups, estimation of risk in areas considering the interaction and spatial dependency of the data, methods particularly useful for public health surveillance and the identification of new hypothesis as detection of spatial and spatio-temporal clusters and other methods appropriated for the analysis and epidemiological studies at individual level. In this chapter it is presented each procedure, the bases and the steps that should be carried out upon applying them.

The results of the processes of analysis are presented in maps, graphs and in texts form in the window of Results.

- **Elements to consider in applying Analytical Procedures to data:**

The cartographic layer should contain the variables that will be utilized in the analytical procedures. If the data to be analyzed are in tables of the database, it will be needed that the tables be related to the layers of the map.

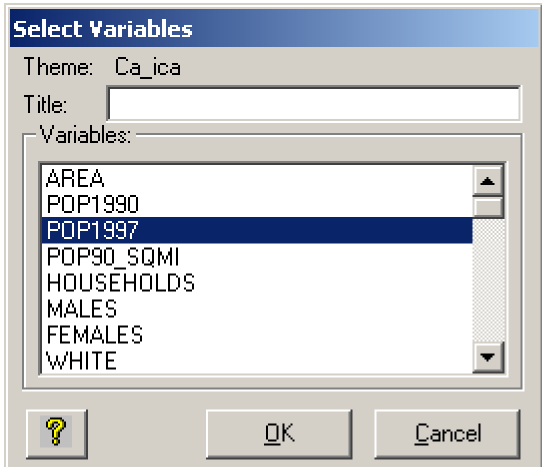
Descriptive Statistics

Descriptive Statistics are a statistical method for data analysis that allows you to calculate a set of measures of central tendency and dispersion for observed values of a variable.

To apply the Descriptive Statistics Procedure to data:

1. Activate the Map Window, and,

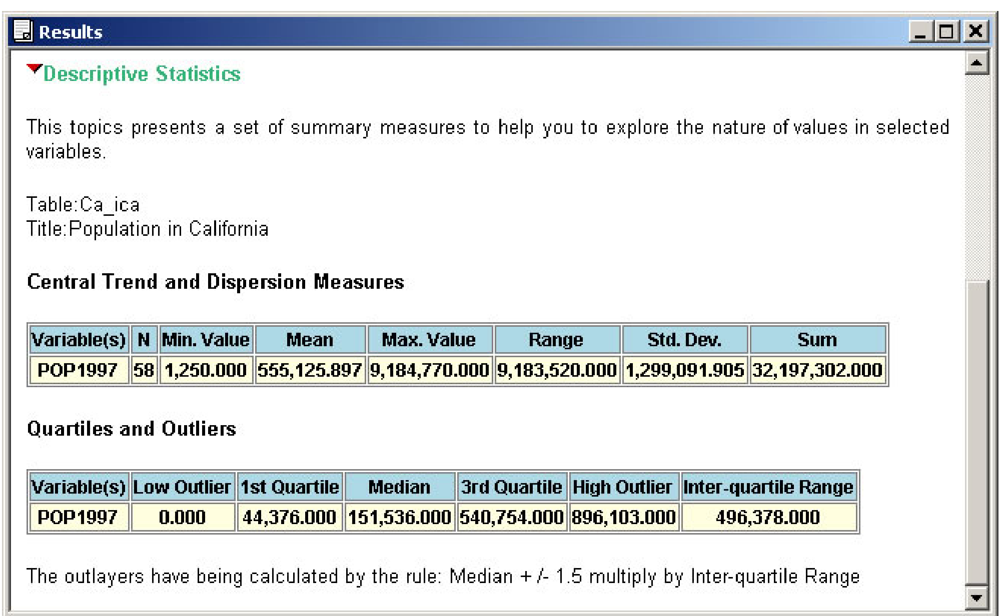
2. a.) Select the cartographic layer of interest,
   b.) Or, display or activate the Attributes Table of the Layer of interest,
   Note: The Attributes Table should contain the variables and data for the analysis.
3. Select the option **Descriptive Statistics**... from the **Epi Analysis** menu.
4. In the displayed **Dialogue Box,** choose the **variable(s)** by ***clicking*** on its name from the list. By pressing the **Shift** and/or **Ctrl** keys, you can select several variables,



Dialogue box Select variables.

5. ***Click*** the **OK** button.

As a result, the calculation procedure is applied to the selected variable(s) and shown in the Results Window, where you can identify: characteristics shown by the data or the values stored by the selected variables, characteristics of the observations collected, mean, range, minimum and maximum values, standard deviation, etc. These are simple statistical measures for data exploration obtained by applying the Descriptive Statistics method.



Results

Descriptive Statistics

This topics presents a set of summary measures to help you to explore the nature of values in selected variables.

Table:Ca_ica
Title:Population in California

Central Trend and Dispersion Measures

Variable(s)	N	Min. Value	Mean	Max. Value	Range	Std. Dev.	Sum
POP1997	58	1,250.000	555,125.897	9,184,770.000	9,183,520.000	1,299,091.905	32,197,302.000

Quartiles and Outliers

Variable(s)	Low Outlier	1st Quartile	Median	3rd Quartile	High Outlier	Inter-quartile Range
POP1997	0.000	44,376.000	151,536.000	540,754.000	896,103.000	496,378.000

The outlayers have being calculated by the rule: Median +/- 1.5 multiply by Inter-quartile Range

Results Window. Descriptive Statistics applied to the data.

Frequency Distribution

Applying the **Frequency Distribution** procedure to data allows you to quantify the frequency of values of a variable; that is, the number of times that an observed value for a variable is repeated.

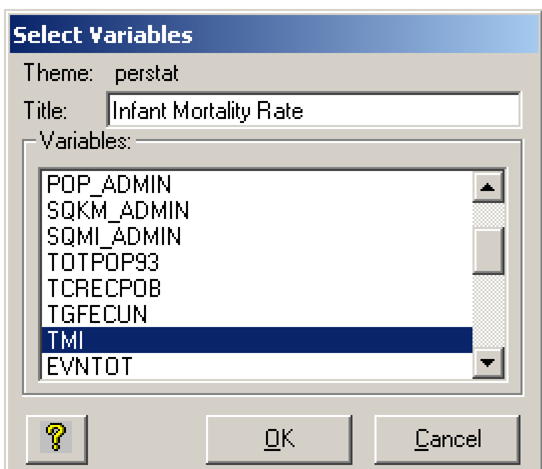
To apply the Frequency Distribution Procedure to data:

1. Activate the Maps Window, and,

2. a.) Select the Layer of interest.
   b.) Or, activate the Attributes Table of the layer of interest,
   Note: The Attributes Table should contain the variables and data needed for this procedure.

3. Select the option **Frequency Distribution**... from the **Epi Analysis** menu,

4. In the displayed **Dialogue Box,** Select the **variable(s)** for which you want to apply the selected **procedure** by ***clicking*** the variables of interest; by keeping the **Shift** and/or **Ctrl** keys pressed, you can select several variables according to the Windows standard,



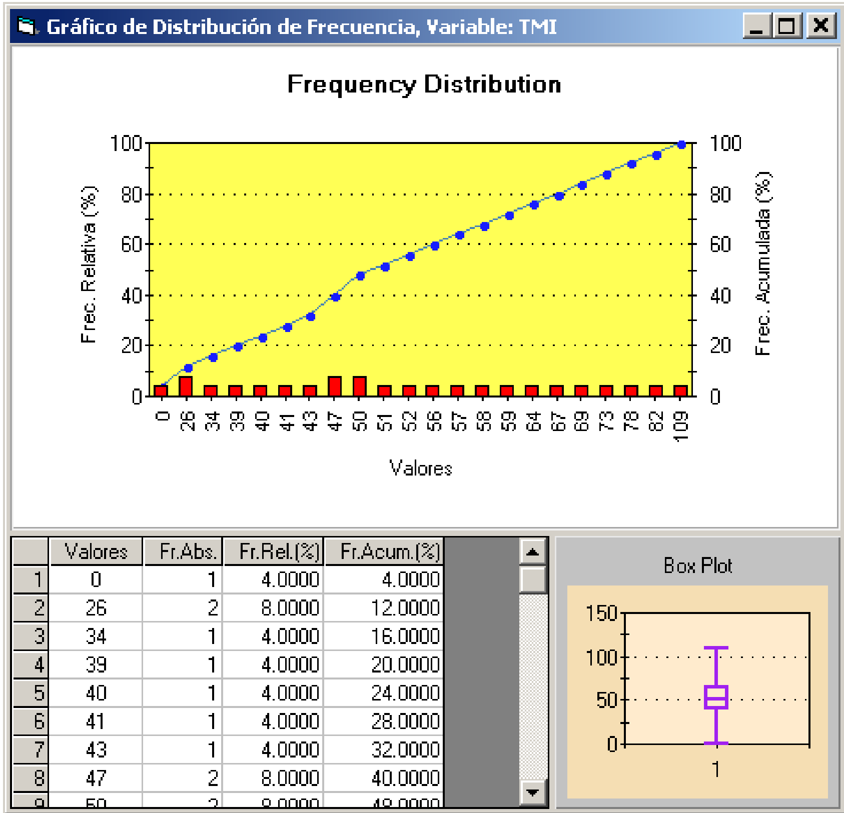
Dialogue box. Select variable(s).

5. ***Click*** the **OK** button.

As a result of applying the process:

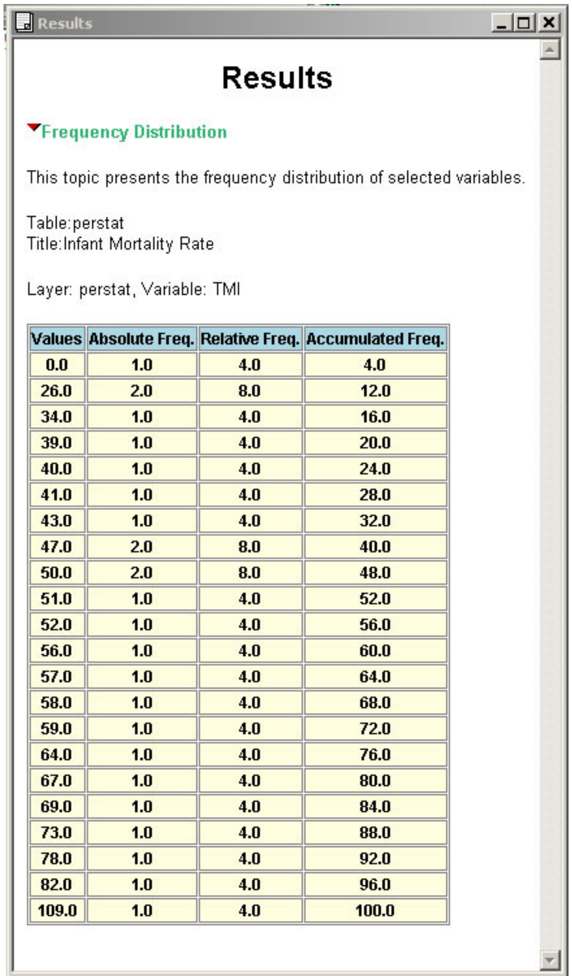
1. A **Frequency Distribution** is generated in a new window that plots the points for the values of the variable, representing the absolute frequency of occurrence of these values, and in bars, the relative frequency: the percentage of occurrence of the values. It also contains the Table of the graphed values and displays them in a Box Plot Graph, showing the minimum and maximum values and the quartiles.
2. The Tables of values with the frequency distributions calculated for each selected variable are added to the Project's Results Window.

Note: As many Graph Windows are displayed as the number of variables selected; each window represents the one variable.



Frequency Distribution Graph.

The frequency distribution of the selected variable is displayed in the Results window.



Results

Frequency Distribution

This topic presents the frequency distribution of selected variables.

Table:perstat
Title:Infant Mortality Rate

Layer: perstat, Variable: TMI

Values	Absolute Freq.	Relative Freq.	Accumulated Freq.
0.0	1.0	4.0	4.0
26.0	2.0	8.0	12.0
34.0	1.0	4.0	16.0
39.0	1.0	4.0	20.0
40.0	1.0	4.0	24.0
41.0	1.0	4.0	28.0
43.0	1.0	4.0	32.0
47.0	2.0	8.0	40.0
50.0	2.0	8.0	48.0
51.0	1.0	4.0	52.0
52.0	1.0	4.0	56.0
56.0	1.0	4.0	60.0
57.0	1.0	4.0	64.0
58.0	1.0	4.0	68.0
59.0	1.0	4.0	72.0
64.0	1.0	4.0	76.0
67.0	1.0	4.0	80.0
69.0	1.0	4.0	84.0
73.0	1.0	4.0	88.0
78.0	1.0	4.0	92.0
82.0	1.0	4.0	96.0
109.0	1.0	4.0	100.0

Results Window. Frequency distribution for a variable.

Correlation Analysis

Correlation Analysis is a statistical method that allows you to measure the level of correlation between two or more variables, creating a correlation matrix; it is a statistical measure of relationships between variables. Using this method, you can determine how close or distant the observed values of the variables are, and what relationship exists among them.

To apply the Correlation Analysis Procedure to data:

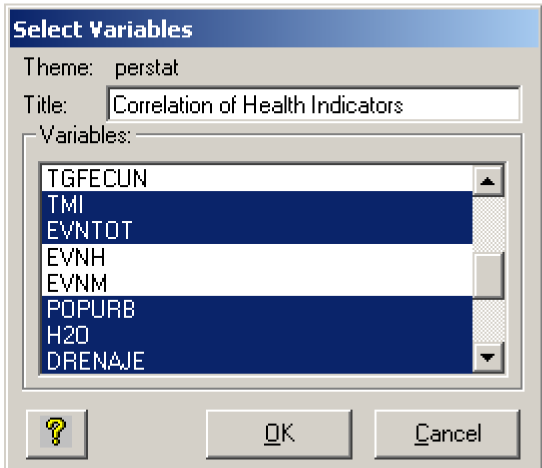
1. Activate the Maps Window, and,

2. a.) Select the Thematic Layer of interest.
   b.) Or, Activate the Attributes Table of the Layer of interest,

Note: The Attributes Table should contain the variables and data for the analysis.

3. Select the option **Correlation Analysis**..., from the **Epi Analysis** menu on the Menu Bar,

4. In the displayed Dialogue, choose the **variable(s)** for which you want to calculate de correlation coefficients by ***clicking*** on the variable names. Holding the **Shift** and/or **Ctrl** keys pressed down you can select more then one variables.



Dialogue box. Select variable(s).

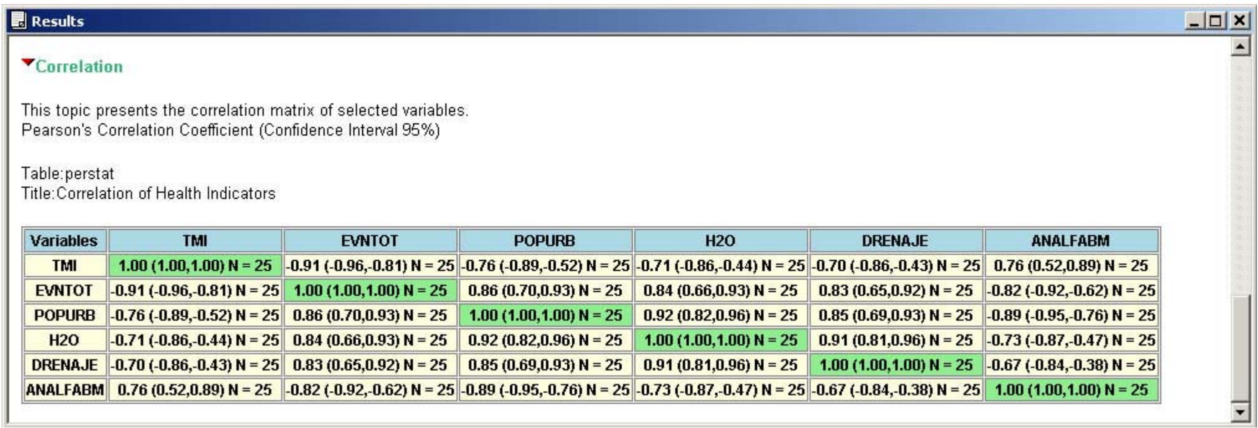
5. ***Click*** the **OK** button.

As a result, a matrix of correlation of the selected variables is generated in the Results Window. The correlation matrix is an squared matrix, with a column and a row for each variable, each cell shows the Pearson's correlation coefficient, the 95% Confident Interval of the estimation of the correlation coefficient and the number of observations.

If the correlation is (+) positive, the value of correlation coefficient is higher the cero ($r >= 0$), this means that there is a positive correlation between both variables, in other words, when the value of one variable is high, the value of the other variable is also high.

If the correlation is (-) negative, the value of correlation coefficient is less than cero ($r < 0$), this means that when there is a negative correlation between the variables, when the value of one variable is high, the value for the other variable is low.

If the correlation coefficient is close to cero ($r =0$) this means that there is a no correlation or a weak correlation between the two variables. The 95% Confidence Interval helps to determine whether de correlation between two variables is statistically significant. If the confidence interval includes the value cero then there is no enough evidence to say that there is correlation between the variables.


Results

Correlation

This topic presents the correlation matrix of selected variables.
Pearson's Correlation Coefficient (Confidence Interval 95%)

Table:perstat
Title:Correlation of Health Indicators

Variables	TMI	EVNTOT	POPURB	H2O	DRENAJE	ANALFABM
TMI	1.00 (1.00,1.00) N = 25	-0.91 (-0.96,-0.81) N = 25	-0.76 (-0.89,-0.52) N = 25	-0.71 (-0.86,-0.44) N = 25	-0.70 (-0.86,-0.43) N = 25	0.76 (0.52,0.89) N = 25
EVNTOT	-0.91 (-0.96,-0.81) N = 25	1.00 (1.00,1.00) N = 25	0.86 (0.70,0.93) N = 25	0.84 (0.66,0.93) N = 25	0.83 (0.65,0.92) N = 25	-0.82 (-0.92,-0.62) N = 25
POPURB	-0.76 (-0.89,-0.52) N = 25	0.86 (0.70,0.93) N = 25	1.00 (1.00,1.00) N = 25	0.92 (0.82,0.96) N = 25	0.85 (0.69,0.93) N = 25	-0.89 (-0.95,-0.76) N = 25
H2O	-0.71 (-0.86,-0.44) N = 25	0.84 (0.66,0.93) N = 25	0.92 (0.82,0.96) N = 25	1.00 (1.00,1.00) N = 25	0.91 (0.81,0.96) N = 25	-0.73 (-0.87,-0.47) N = 25
DRENAJE	-0.70 (-0.86,-0.43) N = 25	0.83 (0.65,0.92) N = 25	0.85 (0.69,0.93) N = 25	0.91 (0.81,0.96) N = 25	1.00 (1.00,1.00) N = 25	-0.67 (-0.84,-0.38) N = 25
ANALFABM	0.76 (0.52,0.89) N = 25	-0.82 (-0.92,-0.62) N = 25	-0.89 (-0.95,-0.76) N = 25	-0.73 (-0.87,-0.47) N = 25	-0.67 (-0.84,-0.38) N = 25	1.00 (1.00,1.00) N = 25

Results Window. Analysis of Correlation among variables.

The correlation analysis is important to determine the grade of linear relationship of two variables.

Simple and multiple linear regressions

Regression analysis is used to estimate quantitative functional relationships between dependent variable and one or more independents variables form actual data. In health, diseases and health events depend on people behaviors, environments conditions, cultural factors, socio-economic conditions and other factors. Empirical investigation of the relationships among them requires the application of regression analysis.

Linear regression is a particular case of regression analysis, where the function used is linear. The linear regression analysis is used in public health mainly with two purposes: a) measure the relationship between independents or explanatory variables and dependent or explained variable; and b) to create a prediction model where one or more independent variables are used to forecast or predict the value of dependent variable.

Using the observation data of the selected variables as independent and dependent, we can define and calculate the linear regression equation as:

$$y = a + bx + e$$

where: x - independent variable
y - dependent variable
a, b - parameters of the equation
e - error, include the portion of the relation explained by other variables not included in the model

Simple linear regression uses a single independent variable.
Multiple linear regression uses more than one independent variable, making a regression equation in this form:

$$y = a + bx1 + cx2 + dx3 + ...$$

where: x1, x2, x3, ..., Xn - independent variables
y - dependent variable
a, b, c, d, ... - parameters of the equation

To apply the Linear Regression Procedure to data:

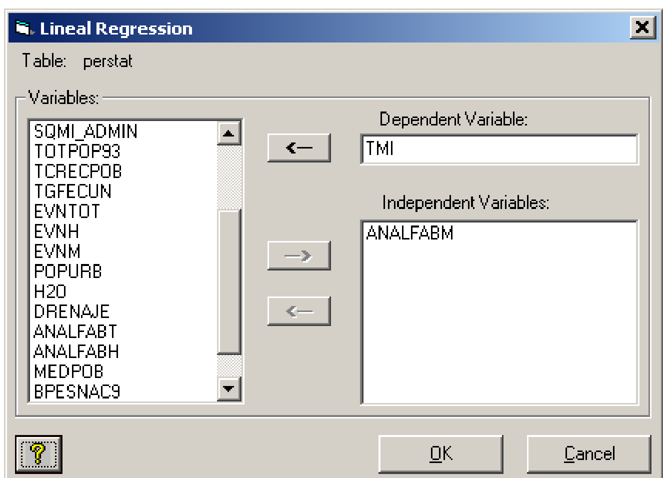
1. Activate the Maps Window, and,

2. a.) Select the Layer of interest,
   b.) Or, Activate the Attributes Table associated with the Thematic Layer of interest,

Note: The Attributes should contain the variables for the analysis.

3. Select the option **Regression Analysis…** from the **Epi Analysis** menu,

4. In the displayed **Dialogue,** select the dependent variable and the independent variables for the analysis,



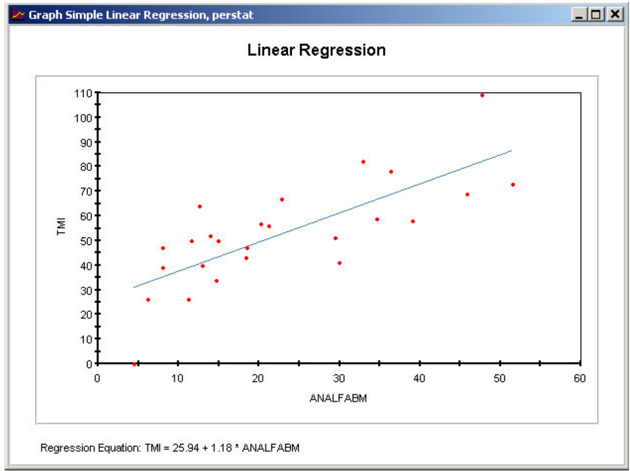
Dialogue Linear Regression

In this case was selected Infant Mortality Rate (TMI) as dependent variable and Illiteracy in woman as independent variable. As only one independent variable was selected, would be apply a single linear regression model. If more than one independent variables are included then a multipe linear regression model will be apply.

5. ***Click*** the **OK** button.

As a result we have:

1. A **Linear Regression Graph** is generated in a graph window that shows the plot of independent and dependent variables, the regression line and at the bottom the regression model.



Linear Regression Graph.

2. The estimated parameters of the model are shown in the Results Window of the Project, containing:

    a.) A Table of **Coefficients,** which gives the parameters for the regression equation, standard error, value of the statistical T test, and probability that allows you to determine the statistical significance for the estimate of the coefficients of the equation,

    b.) A Table of **Analysis of Variance** that shows the sum of squares, degrees of freedom, mean square of error, statistical F test, and the significance or probability of the regression test and other effects. This shows whether the regression equation represents the distribution of the plotted values or not.

Results

Regression

This topic presents the regression results of selected independence variables and dependent variable.

Table:perstat

Regression Model

Coefficients

N = 25

Variable	Coefficient	Standard Error	t	Signif.
Intercept	25.94	5.58	4.64	0.0001
ANALFABM	1.18	0.21	5.58	0.0000

Arithmetic expression of Regression Model: TMI = 25.94 + 1.18 * ANALFABM

Analysis of variance (ANOVA)

	Square Sum	Degrees of Freedom	Mean Square Error	F	Signif.
Regression	75,804.149	1	75,804.15	373.192	0.0000
Residuals	4,671.851	23	203.12	-	-
Total	80, 476.	24	-	-	-

Results of a Simple Linear Regression model

Calculation, Standardization, and Spatial Smoothing of Rates

SIGEpi offers a easy way to calculates rates and proportions, and estimate underlying risk through standardization and smoothing of rates

The procedures for **Calculating and Standardizing Rates** allow you to calculate crude and strata-specific Rates, and to standardize rates using the direct and indirect methods. With the results obtained, you can also create Thematic Maps to facilitate the interpretation and analysis of the rates.

The spread of disease in a community occurs on an unstable demographic terrain. Populations grow or shrink over time and vary from one place to another. In general, the absolute number of cases of a disease, for comparison purposes, has limited importance unless it is linked to the reference population; therefore the use of Rates is recommended.

> **Rate:** In Epidemiology, this represents a relationship between the number of individuals displaying a specific characteristic and the total population. It represents a proportion (fraction) of the set of observations.

Health events are related to several factors, which vary from region to region, and the estimation of risk take into account data from person, time, and place, We often try to establish an association between one of these factors and the disease. It is therefore necessary "to remove" the effect of other factors to obtain a more precise causal relationship. For example, the proportion of elderly people in a population "A" can be greater than in population "B."

When this type of difference exists between geographic areas, comparisons of Rates between different populations require the Crude Rate for each population to be standardized to remove the effect attributable to differences in the population distribution of the areas.

> **Standardization of Rates**: This is one of the procedures used to remove, in a comparison of a series of observations, the link between the effect and one or more causes other than the one we intend to study. In this way, the different series of observations (populations) are artificially made "comparable" in terms of one or more epidemiological characteristics.

Morbidity and Mortality Maps offer an opportunity to make multiple comparisons of Rates across the geographic units of the study region. However, comparisons of Rates between different populations require that the specific Rates be standardized to remove the effect of differences in the population structure of the areas.

> **Smoothing of Rates:** Reduction of local variability in the data. When applied to a variable distributed spatially, reduction of local variance. The spatial smoothing is based on the approach of taking into account data from the neighbors units.

Description of the methods

Rates and proportions are used to estimate an underlying risk. For example, one would be interested in a measure of the risk of dying from a type of cancer and use the ratio of the number of people who actually died during a given period over the total number of the "population at risk" of dying from the disease. While the underlying risk may be constant, the number of people who actually died will tend to vary from time to time as well as across space.
In practice, the rate or proportion is often rescaled to express the notion of risk more intuitively, for example, as 10 per 100,000, rather than 0.00010.

In order to describe and manipulate proportions, three data elements are necessary:

- the numerator variable, or the number of "events" *E* (such as the number of cases of Malaria in a municipality in a given year)
- the denominator variable, or the "population at risk" *P* (such as the municipality population in a given year)
- a scaling option, in a number of meaningful units *S*, such as per ten thousand, per hundred thousand, etc. (typically, different disciplines have their own conventions about what is a "standard" base value).

So Rate is defined by: r = (E / P) * S

where: E - number of sick or dead people (cases)
P - population at risk
S – scaling factor

It is this *r* variable that becomes the basis for an estimate for the underlying "true" risk.

Rates adjusted by the **Direct Method (DAR**), Rates adjusted by the **Indirect Method (IAR**), and the **Standardized Ratio for Rates (SRR**) are calculated for each area or geographic region as:

$$DAR_j = \frac{\sum_i c_{ij} P_{si}}{\sum_i P_{si}}$$

$$IAR_j = SMR_j \left[\frac{\sum_i x_{si}}{\sum_i P_{si}} \right]$$

$$SMR_j = \frac{\sum_i x_{ij}}{\sum_i c_{si} P_{ij}}$$

Where:

c_{ij} is the specific rate for stratum i in the region j
P_{ij} is the population of stratum i, region j
c_{si} is the specific rate for the standard population s
P_{si} is the population of stratum i in the standard population s
x_{ij} is the number of cases in stratum i, region j
x_{si} is the number of cases in stratum i of the standard population s

Data Requirements

To calculate and standardize Rates, the attributes table of the layer should contain the variables of each element of the rate, E cases or deaths and P population under risk in the given period of time and place.

In the case of calculation of specific rates and standardization of rate, it is a requirement to have these two elements (E and P) for each stratum.

For example: to calculate the specific Rates for the age groups (strata): children under 5, from 5 to 14 years, from 15 to 39 years, from 40 to 65 years, and, 65 years and more, the data Table must contain variables for cases from each stratum: Cases_5, Cases_5_14, Cases_15_39, Cases_40_65, Cases_65+ and population variables for each stratum: Pop_5, Pop_5_14, Pop_15_39, Pop_40_65, Pop_65+

You can use the standard populations that you considers most advisable; you can utilize global or regional population data, the population standard defined by international organizations as United Nations, the population of one of the geographic unit under study or the total population of the geographic units you are comparing. It is important to know that the final result of the standardization will depend on the standard population selected. This means that standardizing rates using different standard population the result will not be the same.

Calculation of Rates

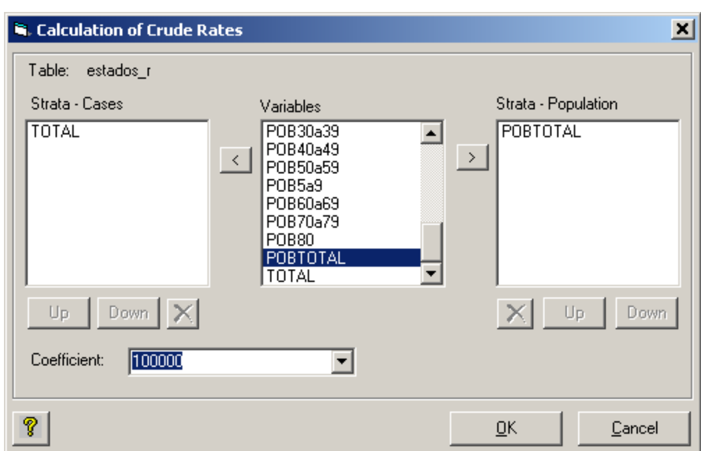
SIGEpi offers the capability to calculate crude and specific rates.

Calculation of Crude Rates

To calculate crude rates, follow the steps:

1. Activate the Map Window, and,
2. Select the Layer to use in the calculation of rates
3. Select option **Calculation of Rates>Crude…** from the **Epi Analysis** menu,

The dialogue Calculation of Crude Rates allows you to select the variable of cases (numerator), variable of population at risk (denominator) and the scale factor (coefficient).



Dialogue Calculation of Crude rates.

4. Select the variable of cases choosing the variable name in the list of variables and click on < button, then the variable is added to the Strata - Cases list.
5. Select the variable of population at risk, choosing the variable name in the list of variables and click on > button, then the selected variable is added to the Strata - Population list.
6. Choose the value of the coefficient (scale factor)
7. Click on OK button

As result a new layer is created and added to the map. This new layer has a new variable, with the prefix "R_" followed by the name of the numerator variable, which contain the Crude Rate by geographic units.

Calculation of Rates

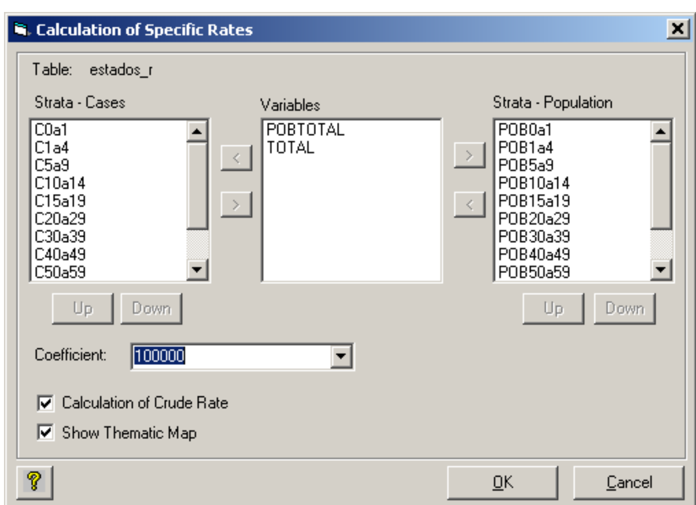
To calculate Specific Rates follow the steps:

1. Activate the Map Window,
2. Select the Layer to use in the calculation of rates,

Note: The layer should contain the cases or death variables (numerator) and the population under risk or under study (denominator) for the analysis.

3. Select the option **Calculation of Rates>Specifics**... from the **Epi Analysis** menu,

4. Using the Dialogue Calculation of Specific Rates, select the variables for cases or deaths (numerator) and the population variables (denominator) for each stratum. From the list of variables at the center of the dialogue select:
   a.) the **variables** for cases (place them in the **Strata - Cases** list, using the < button on the left side of the dialog box),
   b.) the **variables** for population or sample (place them in the **Strata - Population** list, using the > button on the right side of the dialog box),
   c.) Select the value of the **coefficient** to use for calculating the Rates,



Dialogue Calculation of Specific Rates.

Note: You can select one or more case-stratum variables to calculate the specific rates; similarly, you should select one or more population-stratum variables. Every case-stratum and population-stratum variable should be line up. You can use the **Up** and **Down** buttons to move the variables to the appropriate positions.

5. Selecting several case and population variables means your working with strata, so the option **Calculate the Crude Rate** is enabled. Checking on this option, the crude rate is also calculated.

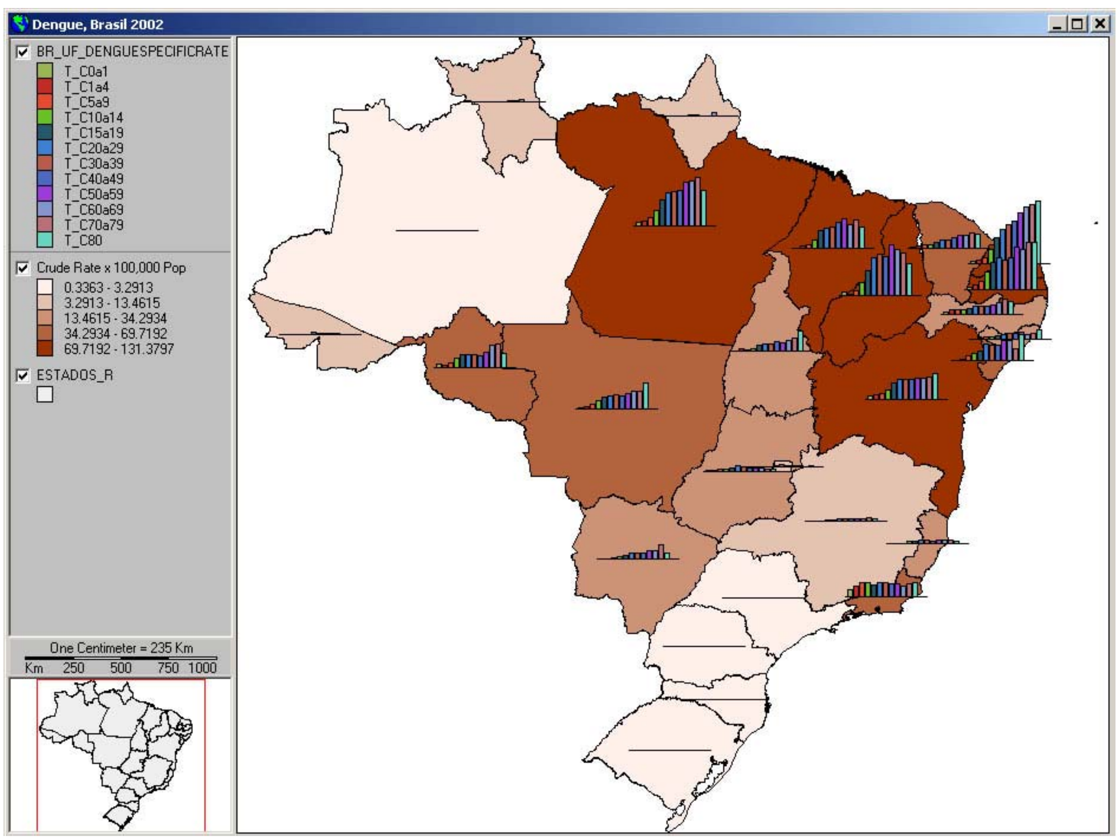
6. The option **Show Thematic Map** already appears checked; you can ***click*** to indicate whether or not you want to create the **Thematic Map** obtained as a result of the Rates calculation.

7. ***Click*** the **OK** button,

8. Using the dialogue **Save Shapefile**, choose a folder and set a **name** for the layer that will create this function.

As a result, a new layer is created, containing the specific rates for each stratum (the prefix "R_" is added to the name of the new variables) and the crude rate. The new Layer is added to the map. In case to calculate only one specific Rate, a thematic map of ranges is created.

If you select more then one stratum, a Thematic Map of Bar Charts is created, displaying the calculated specific Rates of each stratum.



Thematic Map of Bar Charts visualizing the Age Specific Rates of Dengue by Federative Unit, Brazil.

Rate Standardization

DIRECT METHOD (DAR)

The Direct Method consist in apply stratum specific rates of confounding variable (e.g. age, sex) of the populations are being compared to a standard population organized in the same strata.

The necessary elements to standardize rates using the direct method are:

- The strata specific rates of the groups of population or geographic areas under study
- The population distribution by the same strata of a selected standard population

The standardization of rates using the direct method allows to estimates rates that are comparable across the population groups or geographic units.

To standardize Rates by direct method, follow the steps:

1. Activate the Maps Window,
2. Select the layer of interest,

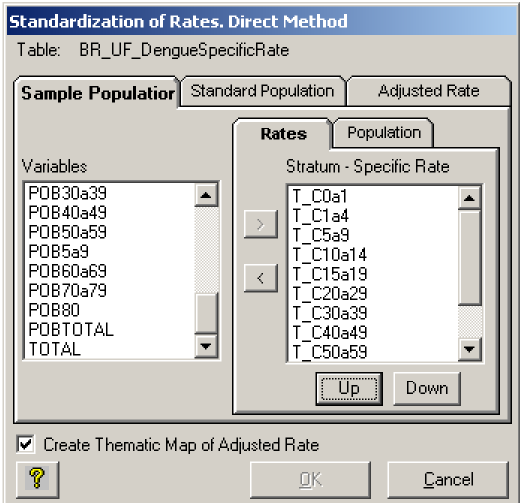
Note: The attributes of the layer should contain the variables and data for the analysis.

3. Select the option **Standardization of Rates > Direct Method**..., from the **Epi Analysis** menu,

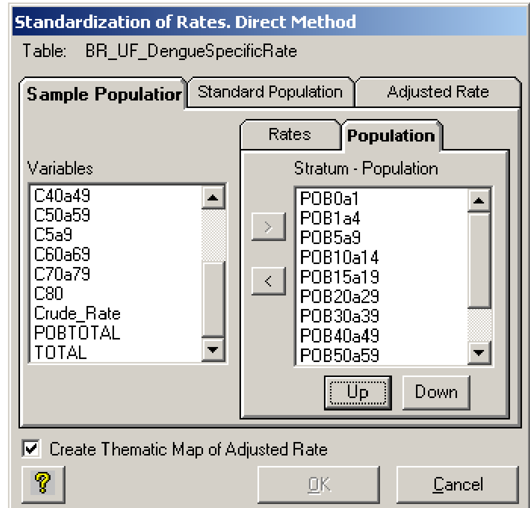
4. The **Dialogue Box** is displayed, which allows you to select the variables and parameters necessary to apply the standardization procedure; to do this:

   a.) In the section **Study Population,** which is activated, select the variables of the Specific Rates, displayed in the list of **Variables**:

   - ***Click*** the Rates **variable(s)** and place them in the **Rates** list corresponding to the edit box and found on the right of the dialog box, using the > button in the section,
   - ***Click*** the **Population** section to the right of the **Rates** edit box and then select the population or sample variable(s) for each stratum and place them in the list, using the > button in the section,



Selection of Specific Rates.

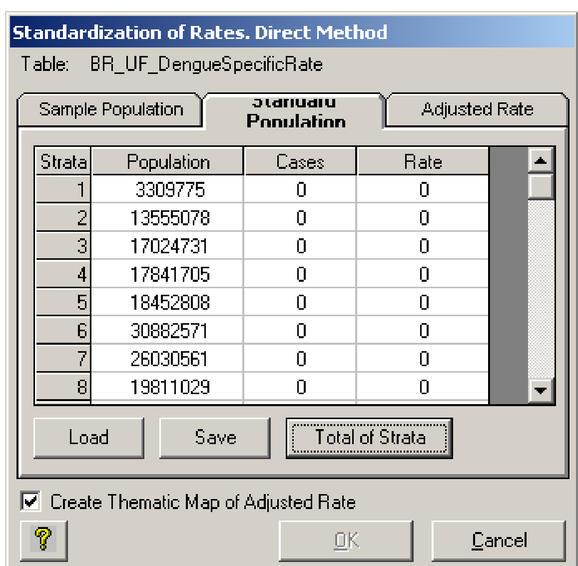


Selection of Population.

b.) ***Click*** the **Standard Pop.** section to define the standard population to use for each stratum previously selected in the Rates and Population edit boxes:

Note that one line is shown for each stratum, allowing to define the standard population for each one. Some possibilities are giving to define the standard population: a) type the value of standard population of each stratum in each cell; b) load a previously defined standard population; and c) to use the sum of the population of the geographic units, which can be calculated clicking on the button Total of Strata.

The standard population could be save in a file clicking on the button Save.

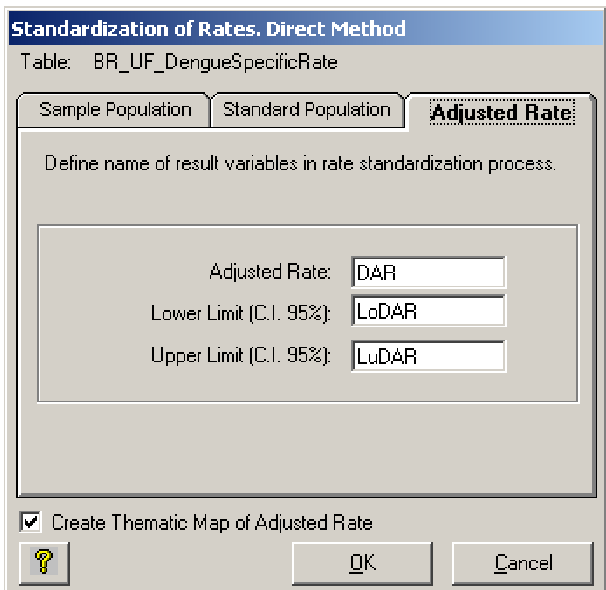


Define Standard Population for each stratum.

- Click on **Load** button in case you want to use a standard population that you have already defined previously.
- Click on Save button to save the defined standard population.

- Or, ***Click*** the **Total of Strata** button, if you want to use the total population for every region of the geographic area (sum of the population of all the regions) as the **standard population** according to the **population** variables for each previously selected stratum,
- Or, directly **Type** in the standard population **values** if you want; this could be the global population, population of a continent, a region, etc.
- The **Cases** and **Rate** columns are not used for the standardization of Rates procedure by the direct method; these are described below under the indirect method.

c.) ***Click*** the tab **Adjusted Rate** to set **names** for the variable where the standardized Rates calculated will be stored:
Adjusted Rate (DAR), Lower Limit (LoDAR) and Upper Limit (UpDAR).
A proposed name for each variable is shown, which you can change if you want.



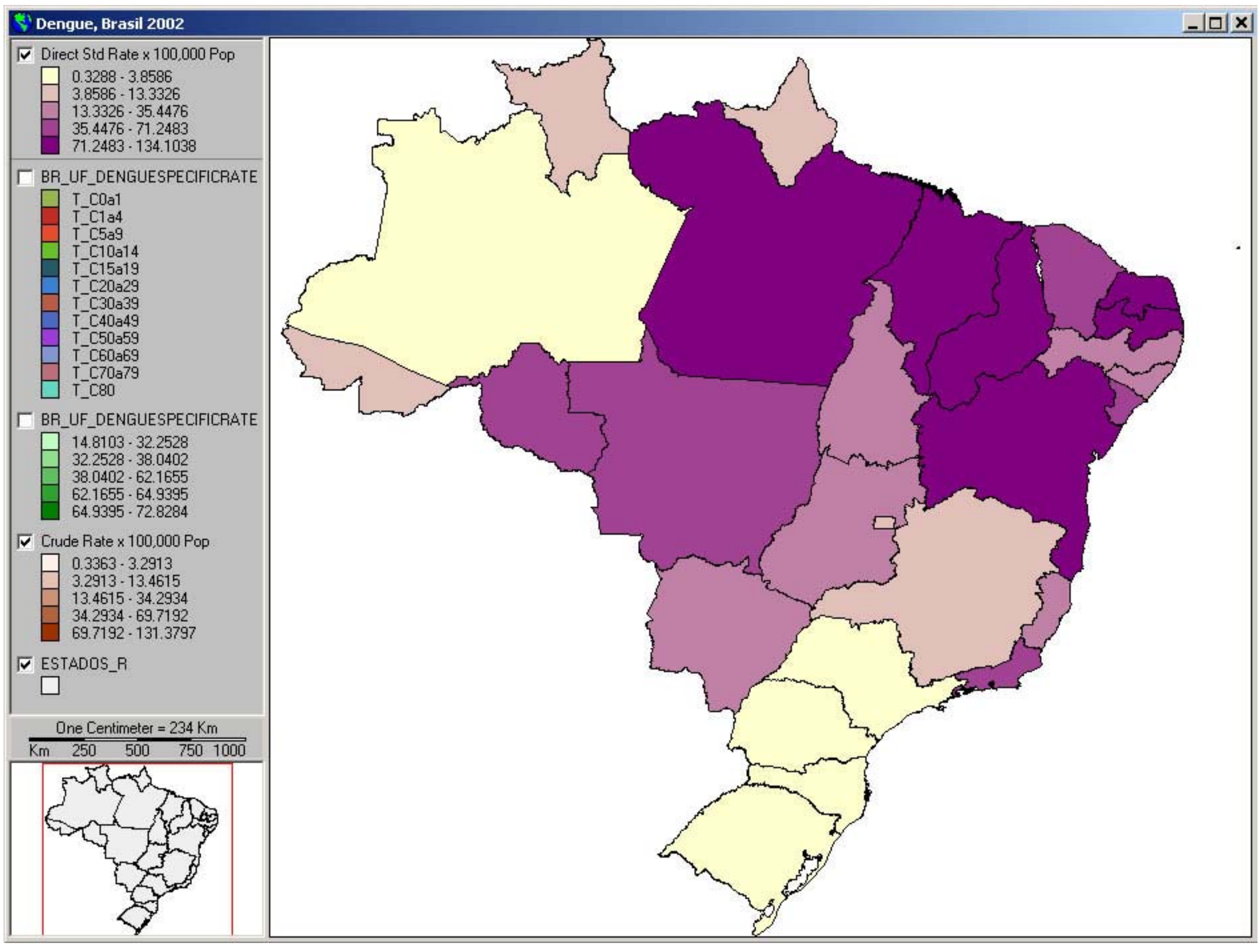
Names for Directly Standardized Rate and Confidence Interval variables.

5. At the bottom of the Dialogue the option **Create Thematic Map for the Adjusted Rate** is checked; you can check or uncheck it to indicate whether or not you want to create the **Thematic Map** of the standardized rate.

6. ***Click*** the **OK** button,

7. Using the Save layer dialog, define a folder and a **name** for the new layer to be created containing the standardized Rate and click on Save button.

As a result, a new cartographic layer is created and the new variables are added. The layer is added to the **Map**.



Thematic Map of Ranges visualizing quintiles of the Directly Standardized Rate (DAR)

INDIRECT METHOD (IAR)

When Specific rates for the strata of the variable we are interested in adjust for, are not available, it is convenient to standardize using the Indirect Method. In this method, the Standardized Ratio of Rate (Standardized Mortality or Morbidity Rate) is obtained.

To apply the Indirect Method it is needed:

- Distribution of population by strata;
- The number of events (cases or deaths) by strata;
- Stratum - Specific Rates of a standard population.

The indirect method utilizes the size of the strata population by which we want adjust for, and the specific rates of the standard population to calculate how many events could be expected in the population if they had the specific rates of the standard population. The expected events are compared to observed events to calculate the Standardized Rates Ratio or Standard Mortality or Morbidity Rate (SMR). The SMR usually is scaled by 100 or expressed in percentage.

The indirect method for the standardization of Rates uses the same Dialogue as the direct method, and the procedure follows the same steps, except that in this case you need to select the Case variables for each population stratum.

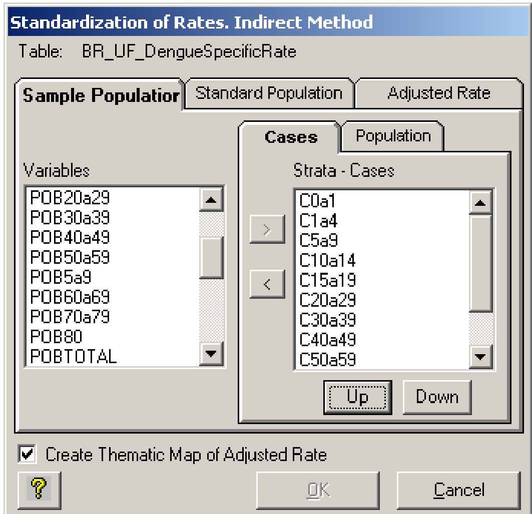
To standardize Rates follow the steps:

1. Activate the Map Window,
2. Select the Layer of interest,

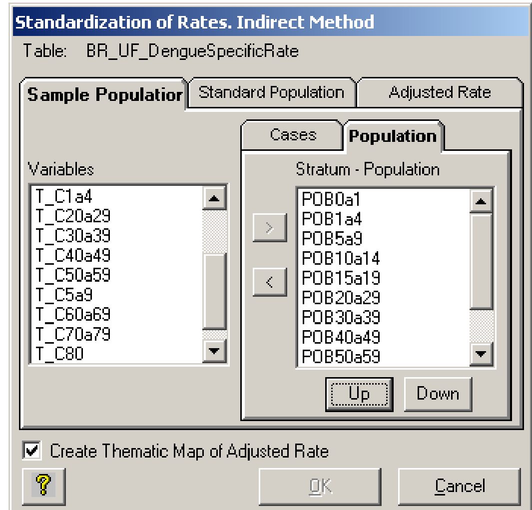
Note: The layer should contain the variables needed in this method.

3. Select the option **Standardization of Rates > Indirect Method**... from the **Epi Analysis** menu,

4. The **Dialogue Box** is displayed, which allows you to select the variables and parameters necessary to apply the standardization procedure; to do this:
   a.) In the **Study Population,** which is activated, select the **variables** containing the values of the **Cases** for each stratum defined, using the list of variables in the Attributes Table associated with the Layer that are displayed in the **Variables** edit box:
   - ***Click*** the **Cases variable(s)** and place them in the Cases list that corresponds to the edit box and appears on the right of the dialog box, using the > button in the section,
   - ***Click*** the **Population** section that appears to the right of the **Cases** edit box and select the **population or sample variable(s)** for each stratum and place them in the list, using the > button in the section,



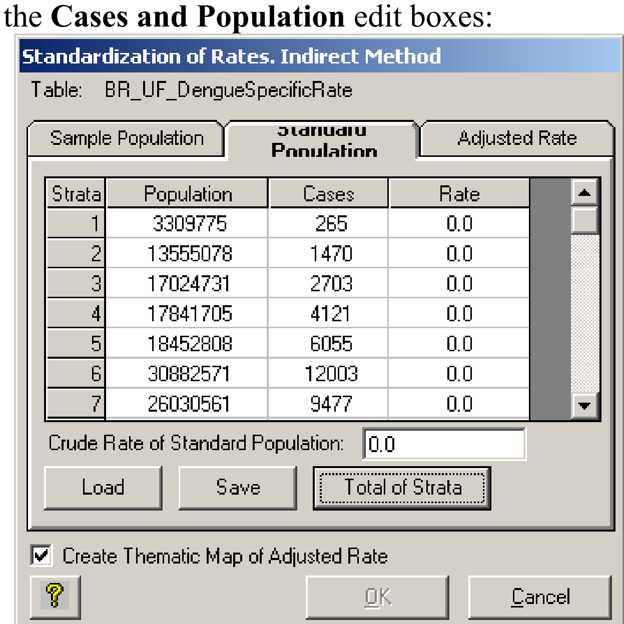
Selection of Cases by stratum.



Selection of Population.

In this case, we are going to use the total population of the geographic units as standard population and so we are using the cases by stratum of the geographic units to calculate the specific rate of the standard population.

Click on the **Standard Pop.** tab to define the standard population and to use for each stratum previously selected in the **Cases and Population** edit boxes:



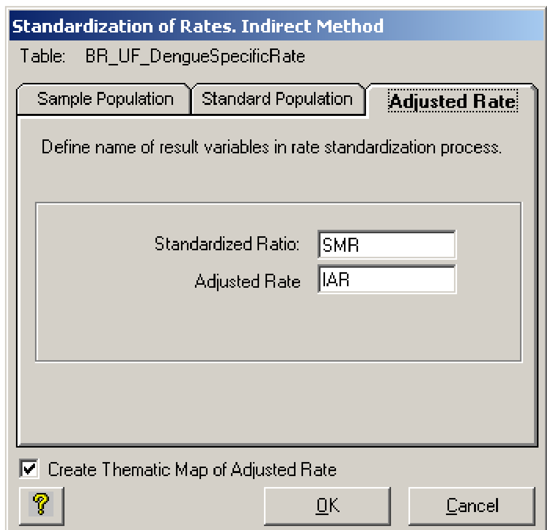
Define Standard Population for each stratum.

Note: a line is shown for each selected stratum, so that you can define the standard populations, cases, and rates for each case. The total population of the region under study is generally used as the standard population.

- ***Click*** on **Load** button if you want to use variables that you already have for standard populations, cases, and rates for the group of defined strata, and that you previously saved with the **Save** option,

- Or, ***Click*** the **Total of Strata** button, in the case you want to use the total population of every region of the geographic area (sum of the population of all geographic unints) as the **standard population,** using the **population** variables for each previously selected stratum; then the total sum of the values of the **cases** variable for each stratum in the region under study is used, and the **rates** are obtained by dividing cases by population for each stratum.
- Or, directly **Type** the **values** for standard populations, cases, and rates.

b.) ***Click*** the **Adjusted Rate** section to establish the **names** for the variables where the calculated standardized Rates are stored: Standardized Ratio (SRR) and Adjusted Rate (IAR).
A proposed name for each variable is shown, which you can edit if you want.



Establish Names for Standardized Rates variables.

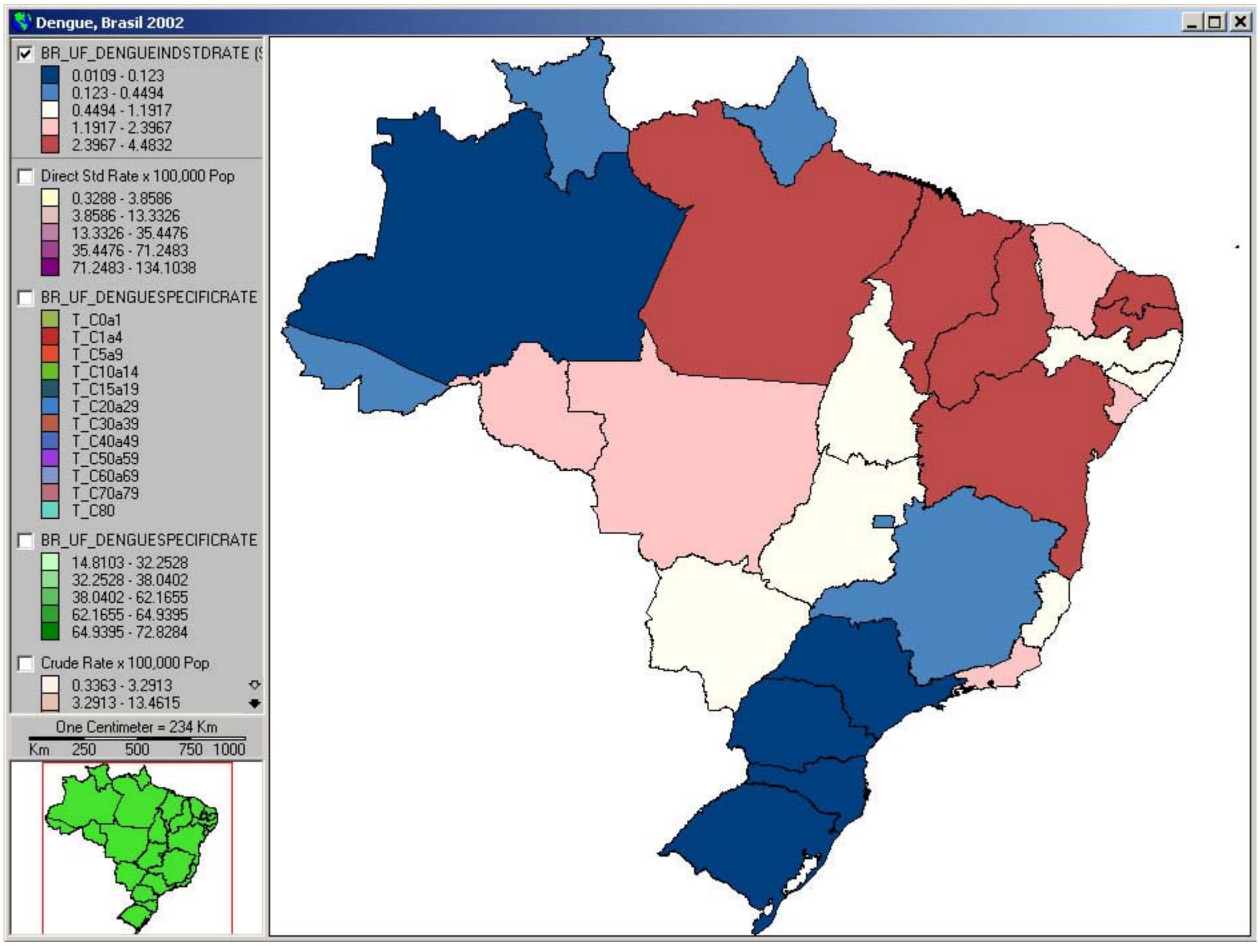
5. Check the option **Create Thematic Map for the Adjusted Rate** in order to create a Themati map of ranges of the SRR.

6. ***Click*** the **OK** button,

7. Select a folder and define a **name** for the new layer.

As a result, the Standardized Rate (IAR) and the Standardized Rate Ratio (SRR) is calculated. A new layer containing both new variables is created and added to the map.

Next map display a thematic map of ranges, classifying the calculated SMR in quintiles.



Thematic map of quintiles of Standardized Incidence Rates of Dengue by Federative Units, Brazil.

Smoothing of Rates

An alternative to transforming or standardizing the raw rates is to statistically adjust the estimate for the underlying risk in a given spatial unit, by *borrowing strength* from the information provided by the other spatial units. This is commonly referred to as *smoothing* or *shrinkage*.

A wide range of procedures for smoothing has been suggested, each one resulting in different adjustments to the raw rate estimate.

SIGEpi offers two alternatives to smoothing rates depending on the availability of data:

- **Using Cases and Population.**
  The layer should contain the cases and population variables for each geographic unit. This is de recommended procedure for smoothing, and requires the availability of cases and population variables.

- **Using Rates.**
  The layer should contain the Rates calculated for each geographic unit. This is the alternative when only is available the Rate and it is not available the cases and population variables.

Rates are smoothed using both the specific calculated Rates and the standardized Rates.

To smooth Rates follow the steps:

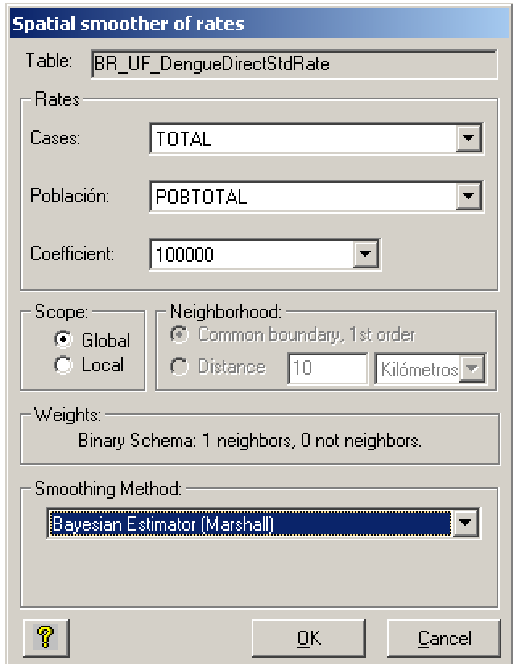
1. Activate the Maps Window, and,
2. Select the Layer of interest,

   Note: The layer should contain the required variables for this procedure.

3. From the **Epi Analysis** menu, select the submenu **Spatial Smoother of Rate** and select the option **Using Cases and Population**...or **Using Rates only**...

4. In each case the corresponding **Dialogue Box** is displayed allowing you to select the variables and parameters necessary to apply the smoothing procedure; to do this:
   - In the **Table** edit box the name of the Attributes Table associated with the Layer is shown; this contains the **variables** that are displayed on a list in the following edit boxes from which you should make your selection,

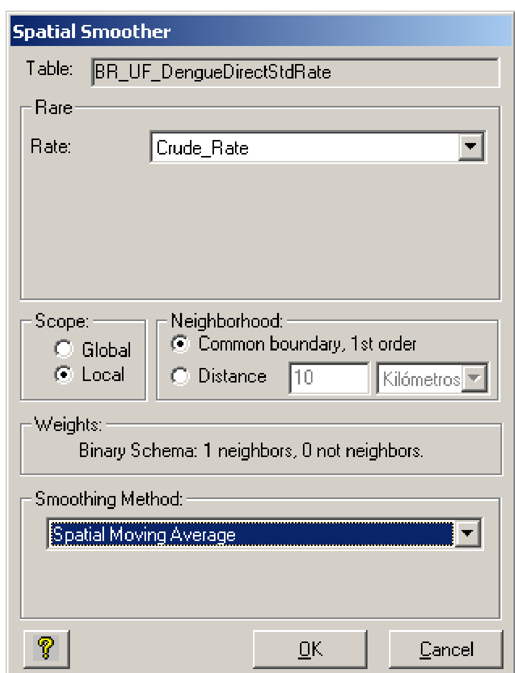
     a.) Using Cases and Population:



Dialogue Spatial Smoother of Rates using Cases and Population.

- ***Click*** the drop-down button of the **Cases** edit box and select the cases **variable** by clicking on it,
- ***Click*** the drop-down button of the **Population** edit box and select the population **variable** by ***clicking*** on it,
- Select the value of the **coefficient** to use,

b.) Using Rates only:



Dialogue Spatial Smoother of Rates.

- ***Click*** the drop-down button of the **Rate** edit box and select the Rate variable by ***clicking*** it,

5. Select the **Smoothing Method** to use, by ***clicking*** it,
6. ***Click*** the **Scope** type that you want to apply:
   Global (includes all the Layer's geographic areas)
   Local (defines an area for which to apply the calculation*)*

7. If you selected **Local Scope**, the **Neighborhood** edit box is activated, enabling you to select the Neighborhood type you want:
   - **Common border**, if you want to include in the calculation only the geographic areas that have a common border with the selected area,
   - **Distance**, if you want to specify a distance and unit of measurement that will define the area to include,

8. ***Click*** the **OK** button,

9. Using the dialog Save Layer, select a folder and define a **name** for the new Layer containing the smoothed Rate.

Using either of the two procedures, a new layer is created, containing the variable of Smoothed rate.

The new layer is added to the **Map**, and you can display it or not just as when working with any other Layer of the Map.

Standardized Rate Ratio Smoother

This is the smoothing procedure of Standardized Rate Ratio (Standardize Mortality or Morbidity Ratio), it is applicable when we have two variables: observed vents and expected events.

The equation of Standardize Rate Ratio is:
SRR = cases observed / cases expected

As result it is calculated the variable S_SRR

S_SRR: standardized ratio smoother

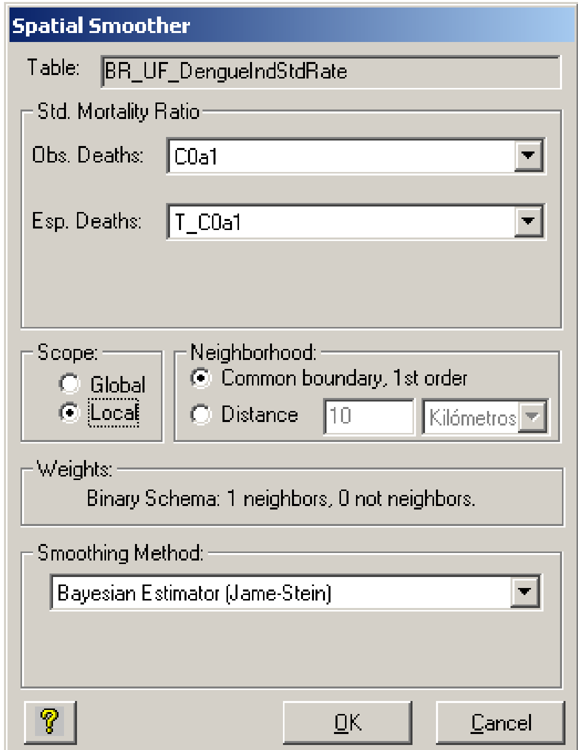
Standardized Ratio Smoother:

1. Activate the Maps Window, and,
2. Select the Layer of interest,

Note: The Layer should contain the require variables for this procedure.

3. From the **Epi Analysis** menu, select the option **Standardized Ratio Smoother**...,

4. The **Spatial Smoother Dialogue** is displayed. It shows the name of the selected layer . From the lists Observe Deaths and Expected Deaths, choose the variables:



Dialogue Spatial Smoother of Standardized Ratio

a.) ***Click*** on the drop-down button of the **Observed Cases** edit box and select the corresponding **variable** by ***clicking*** it,

b.) ***Click*** on the drop-down button of the **Expected Cases** edit box and select the corresponding **variable** by ***clicking*** it,

5. ***Click*** the **Scope** type that you want to apply: **Global** (includes all the geographic areas of the Layer) or **Local** (includes only the neighbor units)

6. If you select **Local Scope**, the **Neighborhood** frame is activated, so select the Neighborhood type you want:
   - **Common boundary**, define as neighbors all the geographic units that share borders.
   - **Distance**, define as neighbors all the geographic units which are within a given distance.

7. ***Click*** on the drop-down button of the **Smoothing Method** and select the desired method by ***clicking*** it,

8. ***Click*** the **OK** button,

9. Select a folder and type a **name** for the new **Layer** using the Dialogue **Save Layer.**

As a result, a new Thematic Layer is created that contains the **Thematic Map using Ranges** that we have just created by smoothing the Standardized Rate; the new variables **SRR** (standardized rate ratio) and **S_SRR** (standardized ratio smoother) are added to the **Attributes Table** associated with the Layer, with the values of the standardized ratios and the smoothed ratios.

The new layer is added to the **Map**.

Identification of Priority and Critical Areas

SIGEpi allows you to identify the geographic units, areas and localities that meet certain conditions, which can be classified as critical according to a criterion based on an approach of social indicators.

The identification of groups of population and geographic region with worse health and socio-economic conditions is one of the common tasks in public health. Usually the public health professionals work with several health and other social indicators to identify the areas that meet the less favorable condition, with the objective to focalize the intervention, public health actions and allocation of resources. This process is known in public health as the identification of critical areas and groups of population.

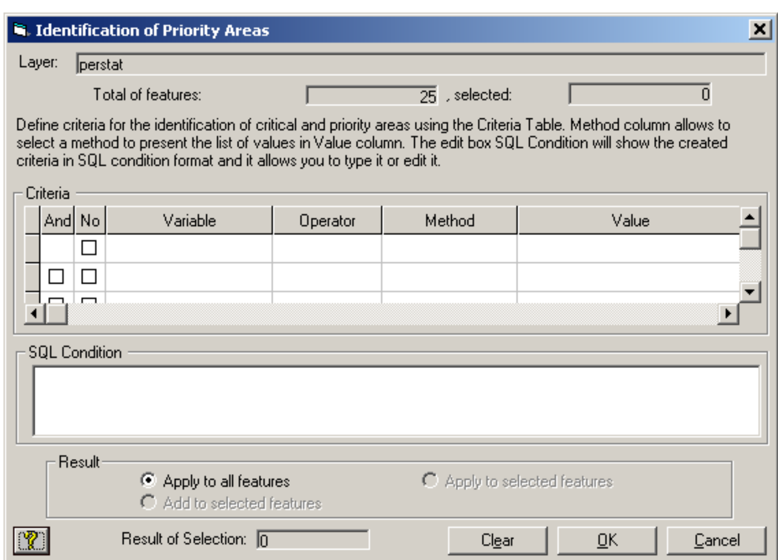
SIGEpi offers a simply procedure to identify on a map the geographic units that satisfy a pre-defined set of conditions. As a result, the geographic units that meet the conditions established in terms of the cut-off values that serve as patterns for comparison are shaded on the Map.

For example: You may want to know the geographic units/groups of population where the Infant Mortality Rate is higher than 15 per 1000 live births a defined value or cut-off value that is considered as critical.

Identifying Priority or Critical Areas on a Map:

1. Activate the Maps Window,
2. Select the layer of interest,
3. From the **Epi Analysis** menu, select the option **Identification of Priority Areas ...**,

The **Dialogue Identification of Priority Areas** is displayed, which allows you to define the **condition** that classifies the geographic unit as critical.



Dialogue Identification of Priority Areas.

This dialog box displays the **name** of the selected **Layer** and table with columns And, Not, Variable, Operator, Method and Value. This grid will help you to define the condition to select the priority or critical areas.

The approach here is to define simple expressions connected by logical operators for creating a complex condition.

4. Clicking on the cell of V**ariable** column, the list of variables of the selected layer will be displayed, allowing to select a **variable**,

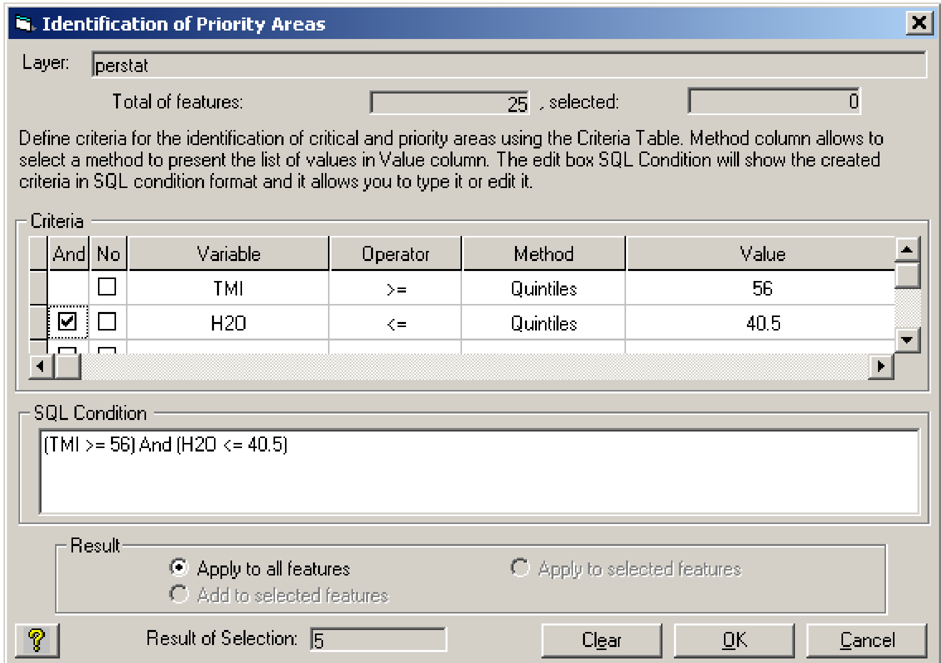
5. Clicking on the **Operators** column, a list of math operator will be displayed, allowing to select one of them,

6. Clicking on the cell of Method column, we can select a classification method in order to create cut-off values for the condition. According to the selected method, in the cell of Value column will be display a set of cut-off values
   There are included five options: Value, display the values of the selected variable; Accumulated Freq., display the values and its accumulated frequency; Quartiles, display the quartiles cut-off values of the variable; Quintiles and Deciles will display the quintiles and deciles cut-off values.

7. Checking on the cell of Not column, we are setting the NOT logical operator, which means the negation of the condition.

8. The AND column should be checked in the case that we need to create a complex condition, formed by at least two simple condition, and using the AND logical operator instead of OR logical operator.

9. The SQL Condition shows the expression or **condition** that is being constructed and is used to select the priority or critical areas,

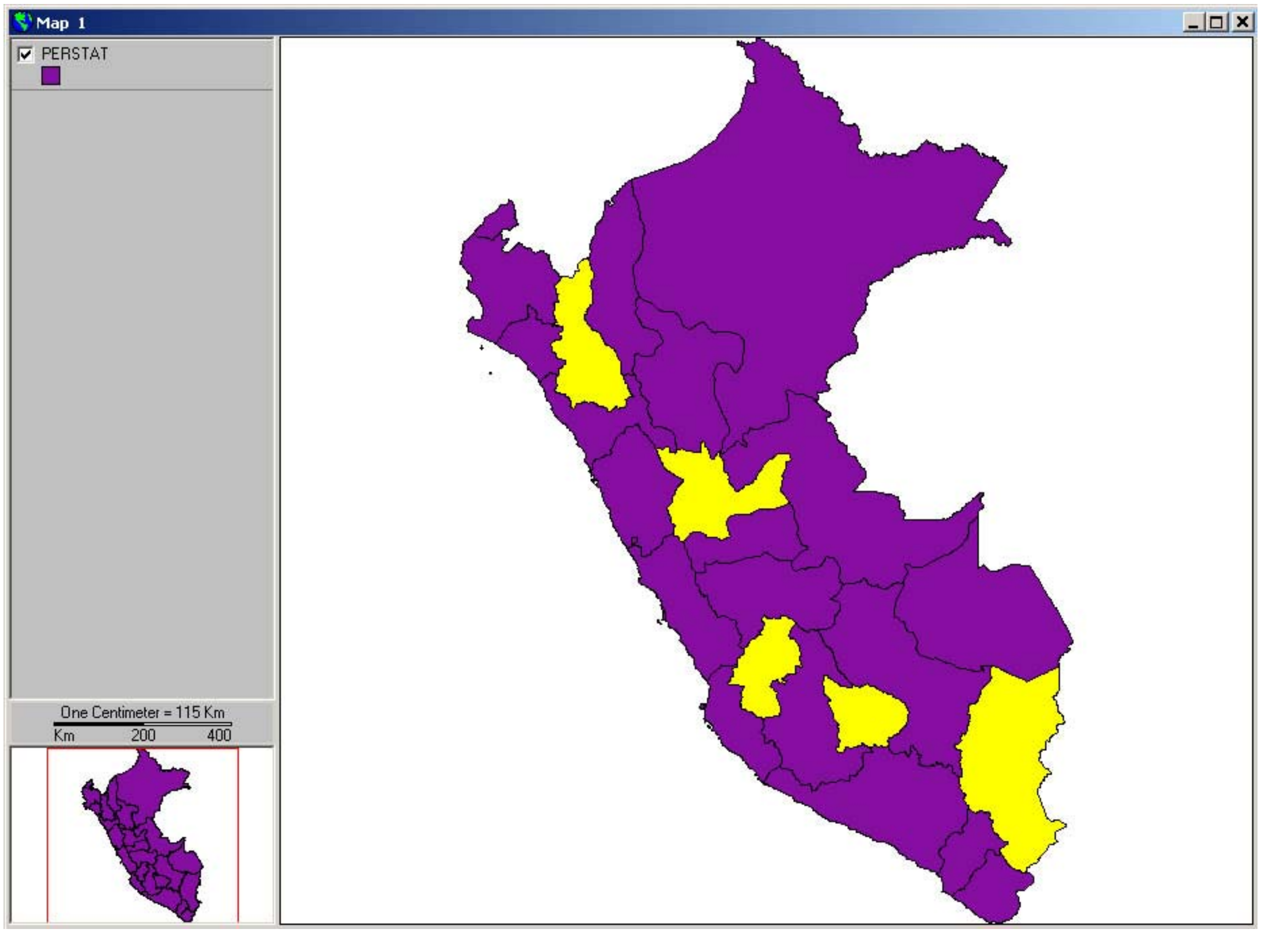


Dialogue Identification of Priority Areas.

In the case of the displayed dialogue, we have defined a complex condition. Indicating that the critical areas are those which Infant Mortality Rate (TMI) is higher that 56 per 1000 live births and the percentage of population with access to potable water is lower than 40.5 percent. You can see at the bottom of the dialogue that five (5) geographic units meet this criterion.

10. ***Click*** the **OK** button,

As a result, the areas or geographic units that meet the established conditions are shaded on the Map. You can change the color and form of the display of selected areas by modifying the **Selection Properties** of the map.



. Identification of Areas or regions that satisfy the condition.

If you want to save this selection, you must convert the layer with the chosen priority or critical areas to a shapefile. The new cartographic layer will contain only the selected (critical) geographic units.

Composite Index in Health

The health situation in geographic areas and groups of populations is usually based on several variables and indicators from different dimensions, usually including health, economic and social indicators and other factors.

Many times it is necessary to synthesize information from different indicators in one index, which take into account several aspects of the health needs.

The calculation of composite health indexes is a simple and general method that allows creating a robust index taking into account several indicators. This method should be use to determine unmet health need and prioritize health intervention or reallocation or resources.

This index is based on Z-scores; the mean and the standard deviation are calculated for each indicator and based on its values. The Zscore for each indicator is calculated. The composite health index is the sum of the Zscore value of all the indicators for each geographic unit.

$$Z = \frac{X - \overline{X}}{S}$$

where: X : Value of the Indicator
$\overline{X}$: Mean
S : Standard deviation

ICS: Composite Health Index = $Z_1 + Z_2 + Z_3 + \ldots + Z_n$

where: Z_i: Z-score of the Indicator i

By applying this procedure, the composite health index is calculated for each geographic unit, the Z-score of each indicator included in the model and the new calculated variable ICS is added to the layer.

To create a Composite Health Index:

1. Activate the Maps Window, and,
2. Select the Layer of interest,

Note: The selected layer should contain the variables and indicators to include in the model.

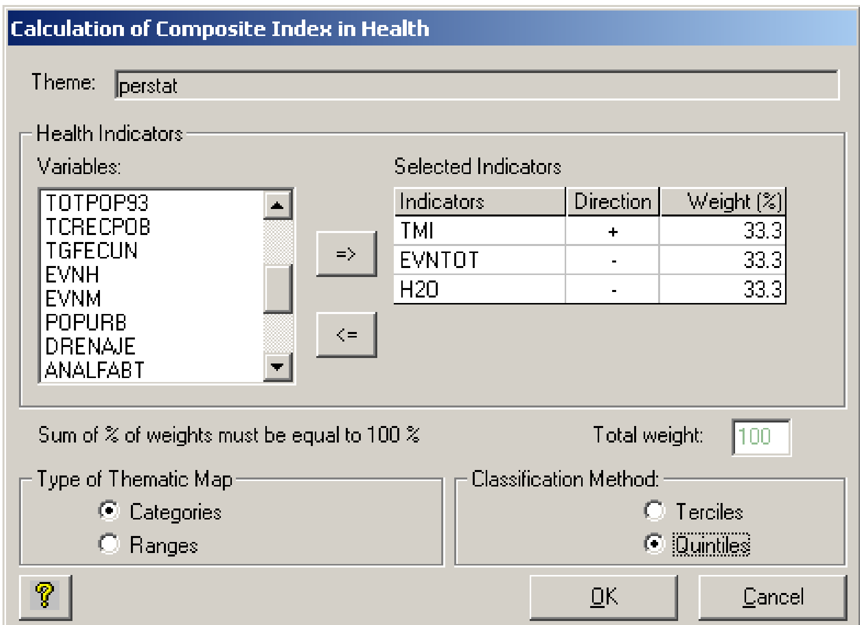
3. Select the **Composite Index in Health** option from the **Epi Analysis** menu,

4. Select the **variables** or **indicators** to include in the model using the Dialogue,

The name of the layer is shown at the top of the dialogue, and the variables of the layer are displayed in the **Variables** list.

Select the relevant **indicators** to include in the model of the Composite Index in Health following the steps:

- click on the indicator in the list of Variables**,**
- click on **=>** button to move the variable to the Selected Indicators table on the right side of the dialogue. The table of Selected Indicators allows you to change the direction and weight of each indicator.



Dialogue Calculation of Composite Health Index.

In the above example, the indicators Infant Mortality Rate (TMI), Live Expectancy at Birth (EVNTOT) and Proportion of Population with access to Potable Water (H2O) have been included in the model.

5. Define the **Direction** of the indicator according to the direction of the Index is being constructed.

   In the example, suppose we are creating and an Index of Health Needs, so high values of this indicator means a higher health needs. In the case of the indicator Infant Mortality Rate, higher values means worse situation so it has the same direction of the index, but in the case of Live Expectancy at Birth and Percentage of Population with Access to Potable Water, higher values mean a better situations, so both indicator have an opposite direction respect to the index. We have to change the direction of both indicators to negative (-).

6. Define the weights in percentage **(%)** or relative influence of each indicator; the total weight should sum **100%**. By default, the same % value is assigned to each indicator. You can change this value. If the objective is to create an composite index to determine health needs or

identify critical areas is better to give the same weight to each indicator. Assigning different weights to the indicator we are introducing subjective information to the model.

7. The total % of influence is shown in the **total Influence** edit box shown further below; as long as this value is not equal to 100%, the **OK** button of the dialog box is not activated,

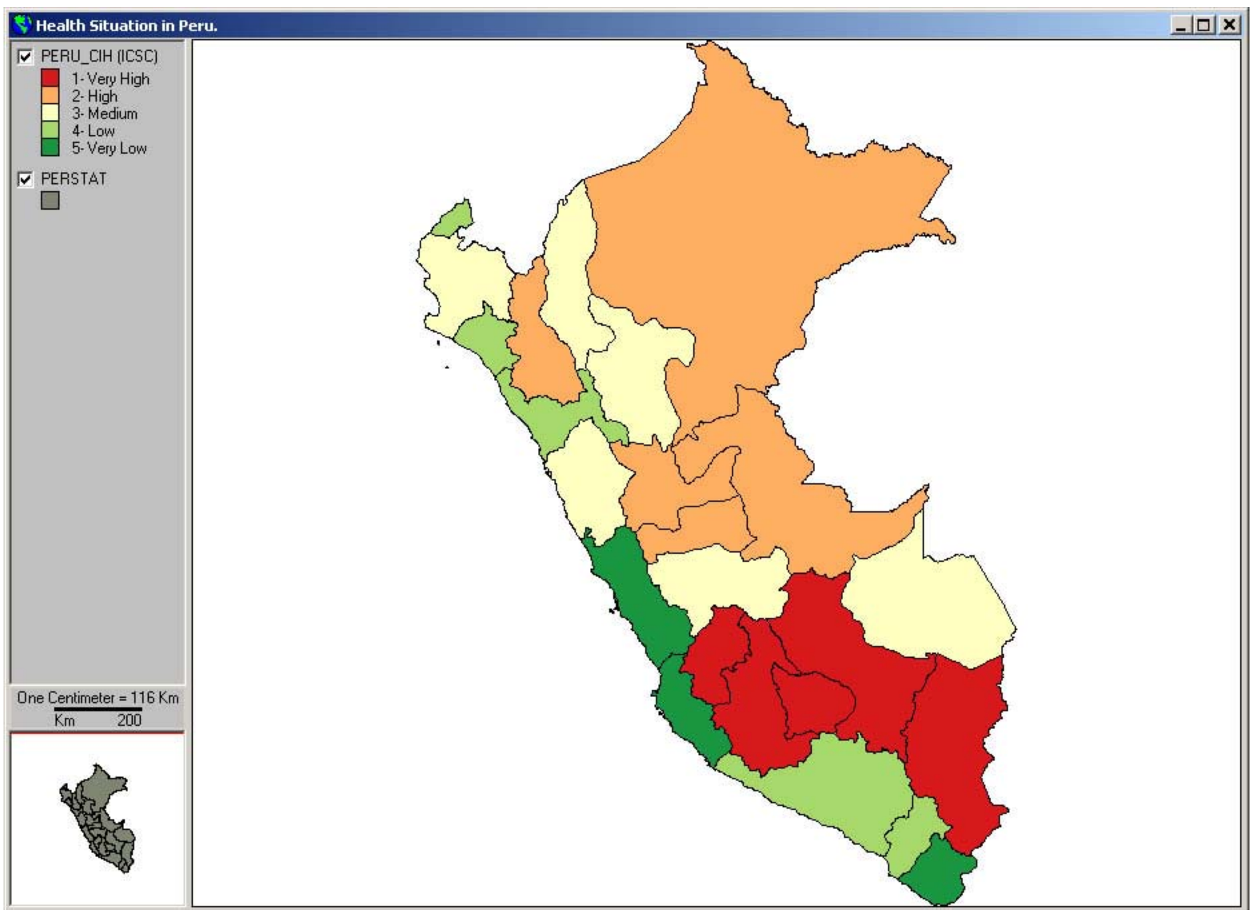
8. If you want **to delete/remove** indicators that you have selected, click on the desired **indicator** in the table of **Selected Indicators** clicking on it, and click on the **<=** button. As result, the selected indicator is removed from the table and returned to the list of Variables.

9. Select the type of thematic map. We recommend to select Categories, which will create a thematic map of categories indicating the geographic units with worst and better conditions.

10. Select the method of classification. Two options are given: Terciles or Quintiles.

11. Click on the **OK** button to calculate the Composite Index.

As a result, a new layer is created containing the **Composite Index in Health (CIH),** which is added to the **Map**.



Thematic Map of the Composite Index in Health..

Spatial Analysis

This topic presents the procedures and functions for the processing and analysis of cartographic and spatial data.

Plotting points from a Table

SIGEpi allows you to create a Thematic Point Layer from a Data Table.

From a **Data Table** containing two columns with the values of geographic coordinates (**Latitude** and **Longitude**), it is possible to create a **Point Layer** on a **Map**.

This function is useful when you have collected data about events, including geographic locations, and you want to plot the locations on a Map.

In the case of epidemiological and public health studies in which data from a population sample was collected through a survey, and the geographic coordinates of the residences of those surveyed were measured using a Global Positioning System (GPS) instrument, it is feasible, using the function **of Plotting points from a Table**, to place the geographic locations of the sample on a Map.

This same approach is applicable to any other type of phenomena being studied. For example: you can compile information from each health post of a rural, mountainous, or remote area. You can measure the geographic location of each health post and include the coordinate pair for latitude and longitude as part of the variables collected. As in the previous example, you can create a Thematic Layer for these health posts.

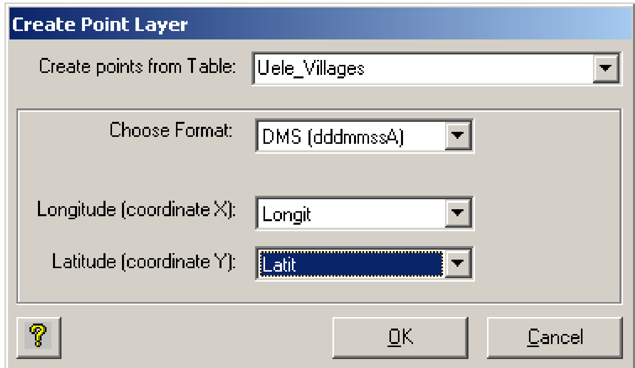
The data collected should be in one Table of the Database. If the data is in Tables with dBase (.dbf) or Excel (.xls) formats, you should import the Tables to a new Database or an existing Database. The Database should be opened in **SIGEpi**.

For more information see: **Working with the Database**.

Steps for Plotting points from a Table:

1. If the Table is contained in a Database (.mdb), Open the **Database,**
2. Or, if it is an external Table, add the **Table** to your existing Database,
3. Activate the Maps Window,
4. Select the option **Plot Points from Table...,** from the **Map** menu,

   The **Dialogue** for Create Point Layer appears:

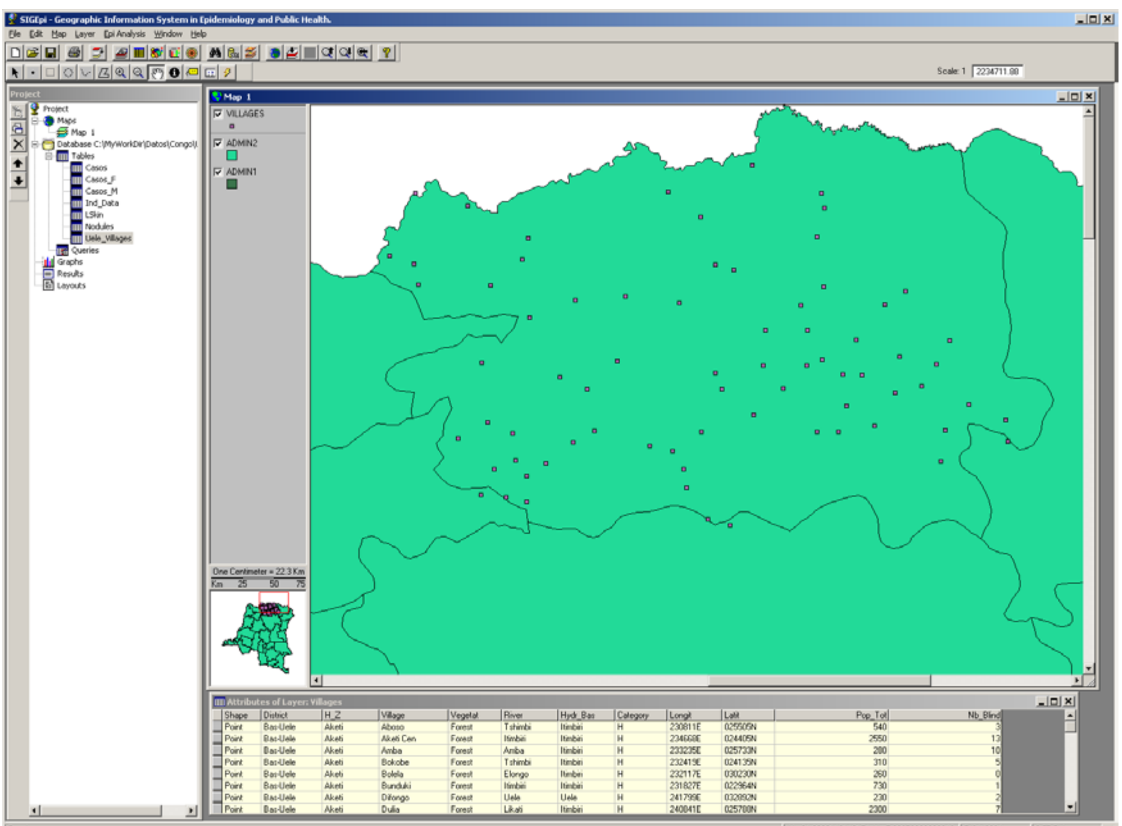


Dialogue Create Point Layer

5. Select the **Table**. All the Tables in the Database appear in the list of Tables,
6. Choose the format of the coordinates.
7. Select the variable of **Longitude** (X Coordinate),
8. Select the variable of **Latitude** (Y Coordinate),
9. Click the **OK** button.

Using the Dialog Save Layer, choose a folder and type a name for the new layer.

As a result, you get a new Layer containing the points of the geographic locations. This Layer is added to your Map.



The new Point Layer (Villages) and its attribute table.

Creating Buffers

SIGEpi allows you to create simple and multiple Buffers around the geographic units of a Layer.

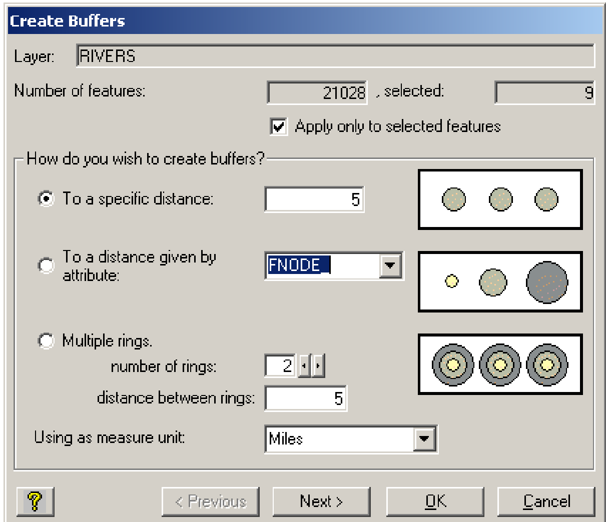
A **Buffer** is a zone under the influence of a geographic object or phenomenon, that is, it includes an area within a specific distance around it. Buffers of constant or variable width can be generated for a set of geographic units, based on the attribute's values in a Layer.

Buffers are useful in proximity analysis. For example: buffers allow identifying the dwellings found within a radius of 1,000 meters of a nuclear power plant; buffers can help to identify the communities within a flood risk area and the quantification of the population at risk; it can help to recognize communities covered by specific health services, areas affected by a specific environmental exposure factor or by levels of it, as well as the population groups exposed to it.

We can create simple or multiple buffers for all the geographic units of a layer or only for a set of selected units. The procedure creates a new layer containing the buffers, which is added to the Map. Buffers have a transparent light color by default, but their properties can be changed using the option **Properties** in the **Layer** menu.

Creating Buffers:

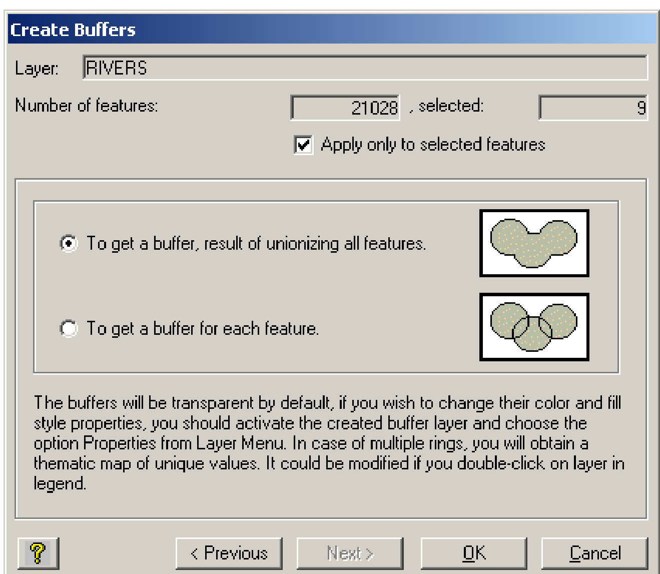
1. Activate the Map Window and Select the **Layer** to be used to create buffers,
2. Select the option **Create Buffer…** from **EpiAnalysis>Spatial Analysis** menu.
   A **dialog box** appears for creating Buffers:
3. The **name** of the selected Layer is displayed. Also is shown the number of selected geographic units,
4. Choose the **type** of Buffer that you want to create:
   - with constant distance. In this case, **enter** the value of the specific distance,
   - with variable distance given by the values of an attribute, **select** the attribute,
   - with multiple rings. Define the **number** of rings and the **distance** between them,
5. Select the **unit of measurement** for the distance,
6. ***Click*** the Next> button to continue defining the characteristics of the Buffers.



Dialogue Create Buffers.

Then, Select:

6. the first option, which creates a Buffer **resulting from the union** of the buffers of geographic units. This option generates a single buffer for buffers of constant and variable width produced by the values of an attribute. In the case of buffers with multiple rings, as many buffers are generated as the number of rings that were defined.
7. The second option generates a Buffer **for each geographic unit**. In the case of buffers with multiple rings, this generates as many buffers per geographic unit as the number of rings that were defined.



Dialogue Create Buffers.

As a result, a new Layer with the created buffers is added to the Map.

Exploratory Spatial Data Analysis

SIGEpi provides a set of methods for detecting spatial patterns in health and disease events

Spatial Analysis offers a set of statistical methods that allow you to detect statistically significant spatial patterns in health and disease events or spatial concentration of regions with similar indicator values.

Analysis of geographic variations in mortality and morbidity is a common way to investigate environmental and other etiological hypotheses. Thematic Maps showing regional health data are frequently used for these purposes. Often, the interpretation of geographic variations is done visually, which has the drawback of leading the investigator to false conclusions due to subjective considerations. However, geographic patterns in the data can be evaluated with spatial autocorrelation methods or statistical significance tests that evaluate the similarity of data values between adjacent regions.

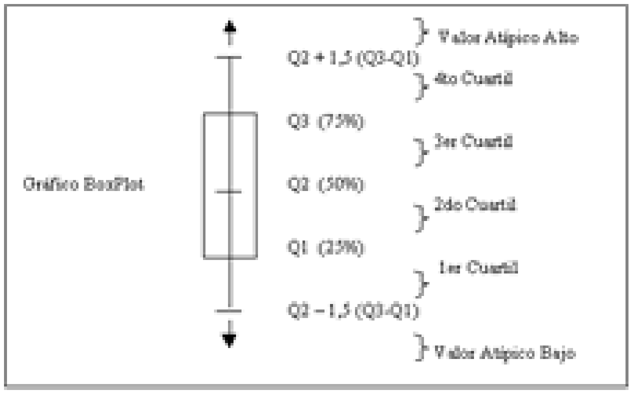
The methods that **SIGEpi** offers in this module for analyzing spatial autocorrelation, Moran's I, Geary's c, are the most well-established measures in the general statistical literature on spatial analysis.

With the application of these techniques, it is possible to detect spatial patterns (clusters) in the geographic distribution of epidemiological data.

- ## Outlier Map

The Outlier Map is a simplify procedure to detect outliers on the distribution of a variable using the box plot classification method and a thematic map of ranges to visualize the distribution of the data in a map. The **Box Map** is an exploratory and visualizing technique that in a simple chart synthesizes information about the magnitude and distribution of a measurement. The BoxMap is created using the classification method of quartiles and applying the rule of Median +/- 1.5 times the Inter-quartile interval to determine the extreme values of the measurement. The output thematic map is created with the box plot bins in six classes.

The data of the variable are classified in quartiles and it is apply the rule Q2 +/- 1.5 (Q3-Q1) as is shown in the cart.



Low Outlier: value obtained by applying the equation Q2 + 1.5 * (Q3-Q1)

Q1: 1st Quartile represents the observation's value below which 25% of the observations of the variable are found.

Q2: 2nd Quartile, median of the observed values; corresponds to the observation's value below which 50% of the observed values of the variable are found.

Q3: 3rd Quartile represents the observation's value below which 75% of the observed values of the variable are found.

4th Quartile: Groups the values that are above Q3 and less than the High Outlier

High Outlier: value obtained by applying the formula Q2 + 1.5 * (Q3-Q1)

For example: A public health professional is analyzing the health situation in Peru. He/She is analyzing the magnitude and distribution of the Infant Mortality Rate per 1000 live births, using the provinces as analytical units. The layer of second administrative level of Peru is available, with the Infant Mortality Rate (TMI**)** and he/she is interested in identify the provinces with significant higher values of TMI. He/She decides to explore the magnitud and distribution of TMI using the Outlier Map.

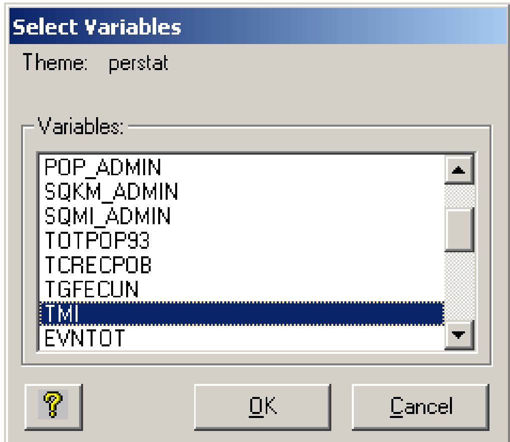
Follow the next steps to apply the Outlier Map procedure:

1. Activate the Maps Window, and,
2. a.) Select the Layer of interest,

    Note: The Attributes Table should contain the variables and data you want to analyze.

3. Select the option **Spatial Analysis / Outlier Map**... from the **Epi Analysis** menu,

4. Select the **variable.** From the displayed **Dialogue Select Variables,** click on the name of variable,



Dialogue Select variable (indicator)

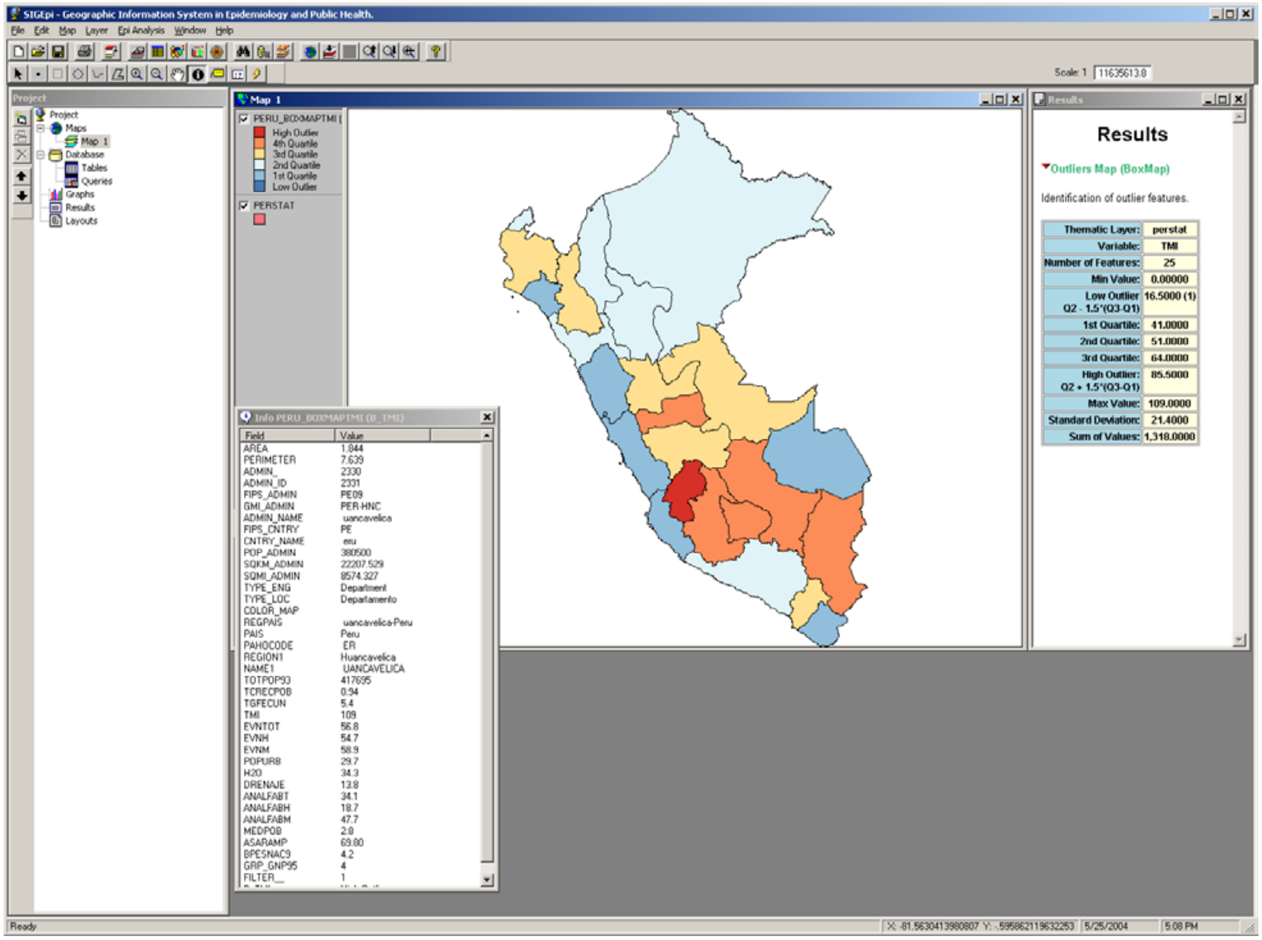
5. ***Click*** the **OK** button,

Using the Dialogue Save Layer, choose a folder and type a **name** for the new layer.

As a result, you will get:

- A new Layer,
- A new **variable** is added to the **Attributes** of the Layer, with the box plot classes. In this example the variable added is B-TMI, and,
- The results are added to the **Results Window,** where the values of the indicator studied are summarized in a Table; this allows you to find, for the indicator represented on the Map, what values represent the median after being grouped into quartiles and what values are below and above the minimum and maximum outliers respectively.

The new layer is added to the **Map.**



Outlier Map, Results of the box plot classification method and attributes of the layer.

In this example, it is identified that the province of Huancavelica has the highest TMI, equal to 109, which is an outlier.

Spatial Smoother

In **SIGEpi** you can spatially smooth the variables or indicators contained in the Attributes Table associated with a Thematic Layer. This method allows you to highlight trends in the values of a variable more clearly by reducing their differences in neighboring geographic units due to error or lack of precision in the observations.

For example:
You have the Thematic Layer for a region divided into municipalities, and the number of cases of malaria (variable: Malaria) is known for each municipality or geographic unit of the region. This variable is selected for smoothing and representation on a Map, identifying more clearly the spatial distribution pattern for the number of recorded cases in the region.

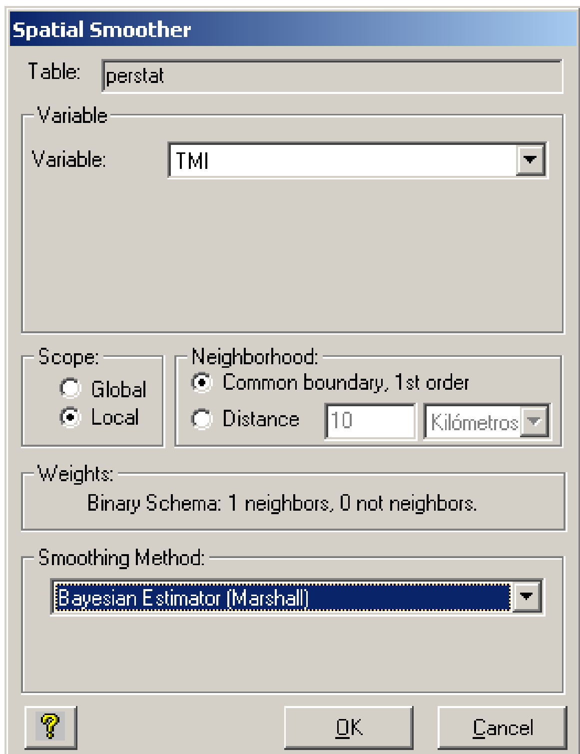
To apply the Spatial Smoother procedure:

1. Activate the Map Window, and,
2. Select the Layer of interest,

   Note: The Layer should contain the variables to be smoothed.

3. From the **Epi Analysis** menu, select the option **Spatial Analysis / Spatial Smoother ...,**

4. The dialogue **Spatial Smoother** is displayed, allowing you to select the variable to which to apply the spatial smoothing procedure; to do this:



Dialogue box. Select variable to smooth.

- ***Click*** the drop-down button of the **Variable** list and choose the variable by clicking on it; in this case the variable Malaria is selected,

5. Select the **smoothing Method** by clicking on it,

6. ***Click*** the **Scope** that you want to apply: **Global** (includes all geographic areas in the Layer) or **Local** (defines an area to apply the calculation)

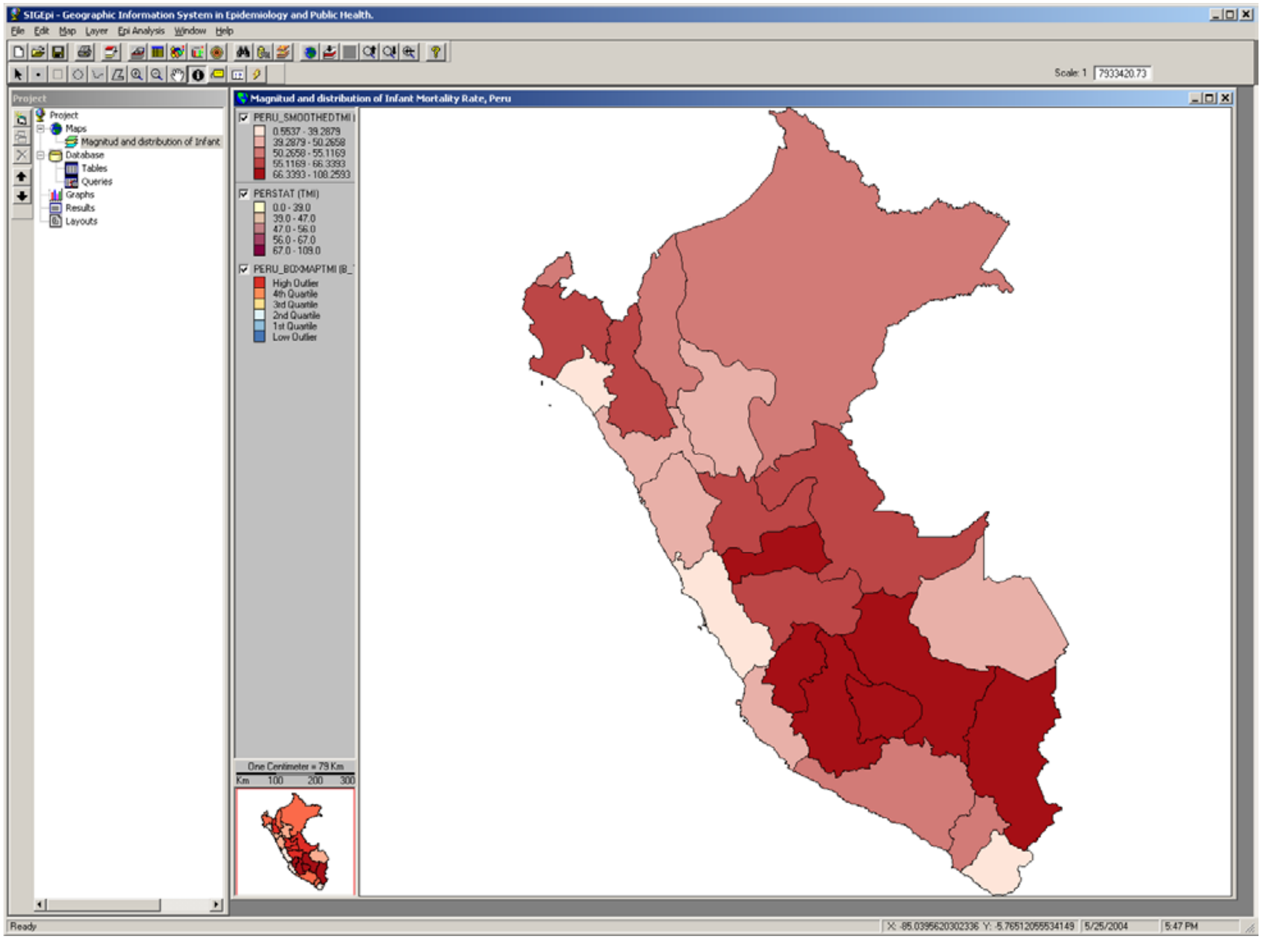
7. If you select **Local Scope**, the **Neighborhood** edit box is activated, allowing you to select the type of Neighborhood you want to choose:
   - **Common border**, if you only want to include in the calculation the geographic areas that have a common border with the selected area,

- **Distance**, if you want to specify a **distance** and **unit of measurement** that defines the area under consideration,

8. ***Click*** the **OK** button,

9. In the Dialogue **Save Layer,** choose a folder and type a **name** for the new **Layer**, and click on **Save** button.

As a result, a new Layer is created containing the **Thematic Map of Ranges** of the smoothed variable classified by quintiles. The smoothed variable is added to the new layer and the new layer is added to the **Map**.



Thematic Map using Ranges visualizing the distribution of the smoothed variable. .

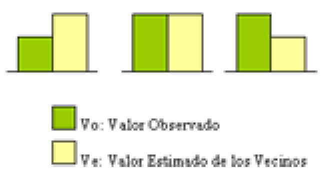
Spatial Lag Map

The procedure used to obtain a **Spatial Lag Map**, also called a Variable Map and its weighted spatial mean, allows you to identify on a Map the pattern or relationship that may exist between an observed variable for a geographic unit and the mean of the estimated values for the variable among the unit's neighbors.

To apply this procedure you only need to select the **variable** you want to represent on the **Map** to find its spatial behavior. **SIGEpi** automatically calculates the **estimated value** for the variable in each geographic unit of the region represented by the Thematic Layer. As a result, the Map shows the relationship between the variable under study and the estimated value for its neighbors in each geographic unit of the Map.

Using this procedure you can obtain a Thematic Map using Bar Charts or using Pie Charts showing the relationship between two variables, allowing a spatial analysis of the pattern the values show.

Taking into account the ways such a relationship can be shown, you can have 3 different relationships, for example:



Vo: variable of observed events in the geographic unit.

Ve: variable of estimated value in the neighboring geographic units; this is the mean of the estimated values of the neighbors.

1. *(Vo<Ve)*
   The variable's observed value in the geographic unit is less than the mean of the estimated values of its neighbors.
   This means that in the area, the analyzed variable or indicator shows a pattern lower than the mean of its neighbors, and, depending on the type of variable, its health situation is better or worse. For example, if the variable analyzed is "infant mortality rate," the indicator shows a better pattern than that of its neighbors; if, on the contrary, the variable analyzed is "access to health care", the area has a lower level of access to health care than its neighbors. If you display this situation on a Thematic Map using Ranges that show the variable's patterns as quartiles, this gives a clearer idea of the indicator's patterns vis-à-vis its neighbors and the region in general.

2. *(Vo=Ve)*
   The variable's observed value is equal to the mean of the estimated values of its neighbors. This means that in the area, the variable or indicator analyzed shows a pattern similar to that of its neighbors.

3. *(Vo>Ve)*
   The variable's observed value is greater than the mean of estimated values of its neighbors. This means that in the area, the variable or indicator analyzed shows a pattern higher than the mean of the estimated values of its neighbors.

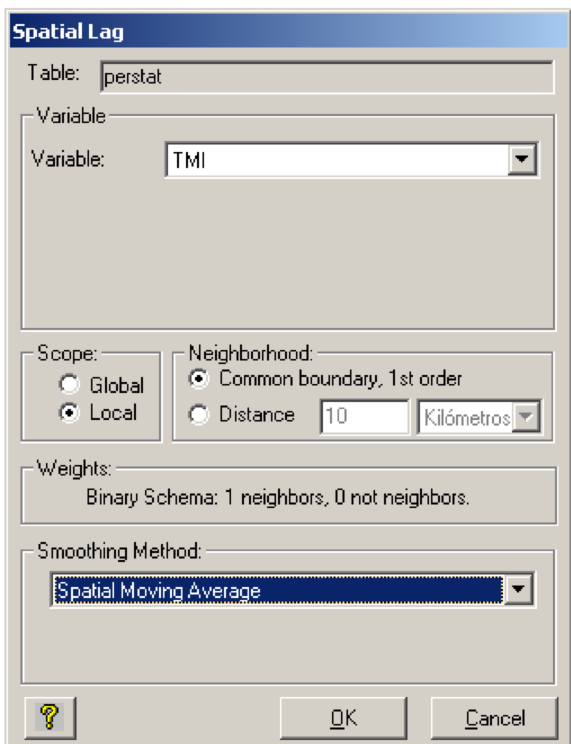
Applying the Spatial Lag procedure:

1. Activate the Maps Window, and,
2. Select the Thematic Layer of interest,

   Note: The Attributes Table associated with the Layer should contain the variables and data you want to analyze.

3. Select the option **Spatial Analysis / Spatial Lag Map ...** from the **Epi Analysis** menu on the Menu Bar,

The dialog Spatial Lag **Dialogue Box** is displayed, allowing you to select the **variable** to apply the **Spatial Lag** procedure to and represent on a Map; to do this:



Dialogue Spatial Lag

- **Click** the drop-down **Variable** list and select the variable by clicking on its name; in this case, for example, the variable TMI (Infant Mortality Rate) is selected.

4. Select the **Smoothing Method** by clicking on it. It is recommended to use Spatial Moving Average.

5. Select the type of **Scope** that you want to apply: **Global** (includes all the geographic areas of the Layer) or **Local** (defines an area to apply the calculation),

6. If you select **Local Scope**, the **Neighborhood** box is activated; you can then select what type of Neighborhood you want to examine:
   - **Common border**, if you want to include in the calculation only geographic areas that have a common border with the selected area,
   - **Distance**, if you want to specify a **distance** and **unit of measurement** that defines the neighborhood area,

7. **Click** on the **OK** button,

8. In the dialogue **Save Layer**, choose a folder and type a **name** for the new **Layer.**

As a result, a new Layer is created displaying the Thematic Map of Bars that we just created by applying the Spatial Lag procedure for the selected variable; the new variable (S_Tasas1) with the smoothed values is added to the **Layer**.

The new Layer is added to the **Map.**

In this example the **Spatial Lag** procedure was applied to the variable TMI (Infant Mortality Rate) and the **Thematic Map using Bar Charts** was obtained, showing the relationship between the selected variable (TMI) and the estimated value for its neighbors (S_TMI) in each region.

A **Thematic Map using Ranges** with the variable SL_TMI also was displayed as a base using the Quantiles method, showing the indicator's patterns in the different areas of the region; this way you can find the areas with greater incidence Rates and see how the Rate for each area compares with those of its neighbors.

- ***Global and local indexes of spatial autocorrelation***

UNIVARIATE SPATIAL AUTOCORRELATION

Univariate spatial autocorrelation techniques allow you to find out if there is a **spatial relationship** among univariate data; that is, you consider only **one variable** that has been measured for the geographic units and determine if the variable's spatial distribution is random or due to a specific spatial factor. If the indicator or variable's values is concentrated in neighboring or nearby geographic units, there could be a spatial factor causing this; if the distribution is dispersed, the pattern could be random and not due to a spatial factor.

In **SIGEpi** you can obtain **global** and **local Indexes** of Spatial Autocorrelation**:**

Global: Obtains an **index** of spatial autocorrelation for the entire **region,** taking into account the value of the indicator or variable for all the geographic units; you also obtain the statistical significance of this index-- that is, if the value obtained is reliable or not and if the spatial distribution is random or is due to a spatial pattern. As a result of applying this procedure, you get a Table in **SIGEpi's** Results Window.

Local: Obtains a spatial autocorrelation **index** for each **geographic unit** of the region using a neighborhood criterion; that is, the indicator's pattern is analyzed in each units neighboring a given area. As a result of applying this procedure, you obtain a Thematic Map using Ranges that displays the calculated index's spatial distribution. The values are added to the Attributes Table of the associated Thematic Layer.

Description of the Statistical Methods

Statement of the hypotheses:

H_0: Random spatial distribution of data

H_1: Existence of a spatial pattern, spatial concentration of regions with similar data values, or simply, spatial grouping.

- **Moran's I Statistic:**

This is similar to the usual correlation coefficient; it measures the covariation between neighboring regions (e.g., municipalities, provinces, states, etc.).

Data with a random spatial distribution yield an expected value for I of close to zero; the presence of a spatial pattern leads to positive I values, with an upper threshold of 1 for the maximum clustering.

- **Geary's c Statistic:**

This incorporates direct-paired comparisons of data, instead of the covariation approach of I. The expected value of **c** for randomly distributed data is 1, while the existence of spatial patterns in the data yields values of less than 1 for c, and a value of zero for the maximum clustering.

- **Rank adjacency D Statistic:**

This is formulated as the average of the absolute differences in ranks for data of adjacent or contiguous regions.

Spatial autocorrelation measures require you to create a matrix of weights or distances that defines the association or proximity between each pair of regions. There is a set of methods to calculate the weights. The Spatial Autocorrelation module offers six ways to calculate the weights or distances among geographic units.

Methods for calculating weights or distances

Binary weights. This is the simplest method. It reflects only the covariation between contiguous regions. It assigns the weight a value of 1 when a pair of regions is adjacent or contiguous, and a value of 0 when regions are not adjacent (W_{ij}= 1 if regions i and j are adjacent and W_{ij}= 0 if they are not adjacent). This method is feasible when the geographic units of a Map are polygons representing areas such as municipalities, provinces, departments, states, or countries.

Proportion of borders. This method takes into account the proportion between the common border of regions i and j, and the total border length of region i. W_{ij}= Border(i,j)/Border(i). It is feasible when the geographic units of a Map are polygons representing regions such as municipalities, provinces, departments, states, or countries.

Inverse of the distance. Assigns weights as the inverse of the Euclidian distance between the centroids of the regions i and j. This method is useful when you are working with specific point-locations (e.g., sources of environmental pollution, etc). $W_{ij}= 1/d_{ij}$. It is feasible when the geographic units of a Map are points that represent locations of cities, health units, or toxic-waste-emission sites, although it can also be applied to regions such as municipalities, provinces, departments, states, or countries.

Inverse of the square of the distance. This is similar to the previous method, except that the Euclidian distance is squared. This is indicated by: $W_{ij}= 1/d_{ij}^2$. It is feasible when the geographic units of a Map are points representing the location of cities, health units, or toxic-waste-emissions, although it can also be applied to regions such as municipalities, provinces, departments, states, or countries.

Ratio between proportion of borders and distance between centroids. This is a combination of the Proportion of borders and Inverse of the distance methods. W_{ij}= [Border(i,j)/Border(i)] * $1/d_{ij}$. It is feasible when the geographic units of a Map are polygons representing regions such as municipalities, provinces, departments, states, or countries.

Ratio between proportion of borders and square of the distance between centroids. This is a combination of the Proportion of borders and Inverse of the square of the distance method. W_{ij}= [Border(i,j)/Border(i)] * $1/d_{ij}^2$. It is feasible when the geographic units of a Map are polygons representing regions such as municipalities, provinces, departments, states, or countries.

Definition of Indexes of Spatial Autocorrelation:
Assume that a Map consists of n regions (e.g., municipalities), and that the observation value for region i is x_i. w_{ij} is a scalar value that quantifies the degree of spatial association or proximity (weight or distance) between the municipalities i and j.

- Moran's I Statistic is defined as:

$$I = n\sum_i\sum_j w_{ij}\left(x_i-\bar{x}\right)\left(x_j-\bar{x}\right) \Big/ \sum_i\sum_j w_{ij}\sum_i\left(x_i-\bar{x}\right)^2$$

Under the null hypothesis for spatially random data, the mean and the variance of I are:

$E(I) =--1/(n-1)$ $Var(I) = (n^2S_1- nS_2+ 3S_0^2)/S_0^2(n^2- 1) - E^2(I)$

where: $S_0= S_{iS\,j}w_{ij}$ $S1 = ½\ S_{iS\,j}(w_{ij}+ w_{ji})^2$ $S2 = S_i[S_j(w_{ij}+ w_{ji})]^2$

For spatially random data E(I) = -1/(n-1), or approximately equal to zero for a moderately large value of n. In contrast, when spatial concentration or a spatial pattern occurs, the x values of neighboring regions are correlated positively, leading to positive values for I.

- Geary's c Statistic is defined as:

$$c = (n-1)\sum_i\sum_j w_{ij}\left(x_i-x_j\right)^2 \Big/ 2\sum_i\sum_j w_{ij}\sum_i\left(xi-\bar{x}\right)^2$$

with the mean and variance under the null hypothesis as expressed below:

$E(c) = 1$ $Var(c) = [(2S_1+S_2)(n-1)-4S_0^2]/2(n+1)S_0^2$

Since the values of x_i and x_j are compared directly in the numerator of c, positive autocorrelation leads to small values for c, with a limit equal to zero for the maximum positive correlation.

- The rank adjacency D statistic is based only on the order of the data. Let y_i be the rank of x_i. Then D is defined as the mean absolute difference in rank for the pairs of adjacent regions and is defined as:

$$D = \sum_i \sum_j w_{ij} \left| y_i - y_j \right| \Big/ \sum_i \sum_j w_{ij}$$

with the mean and variance under the null hypothesis as expressed below:

$$E(D) = (n+1)/3 \qquad Var(D) = \left[(1/18)(n+1)(n-2) \sum_r w_r^2 - (1/9) \sum \sum_{r \neq s} w_r w_s \right] \Big/ \left(\sum_r w_r \right)^2$$

where: $w_r = w_{ij}$ (weight between regions i, j) and $w_s = w_{ji}$ (weight between regions j, i)

Statistical significance test:
For the three measures of spatial autocorrelation, approximation to the normal distribution is assumed, which means that the Z Test is used to determine the statistical significance:

For Moran's I: $Z = [I - E(I)]/[Var(I)]^{½}$
For Geary's c: $Z = [c - E(c)]/[Var(c)]^{½}$
For D: $Z = [D - E(D)]/[Var(D)]^{½}$

The probability value p is obtained from Z value and evaluating the function of the standard normal distribution.
The p value should be small enough and compared with *a* value, where a is the probability we accept of rejecting Ho by mistake. If **p < a** H_0 is rejected, meaning that a spatial pattern or cluster of regions with similar values does exist. In other words, there is enough evidence to say that exist cluster not due by chance.

Applying spatial autocorrelation:

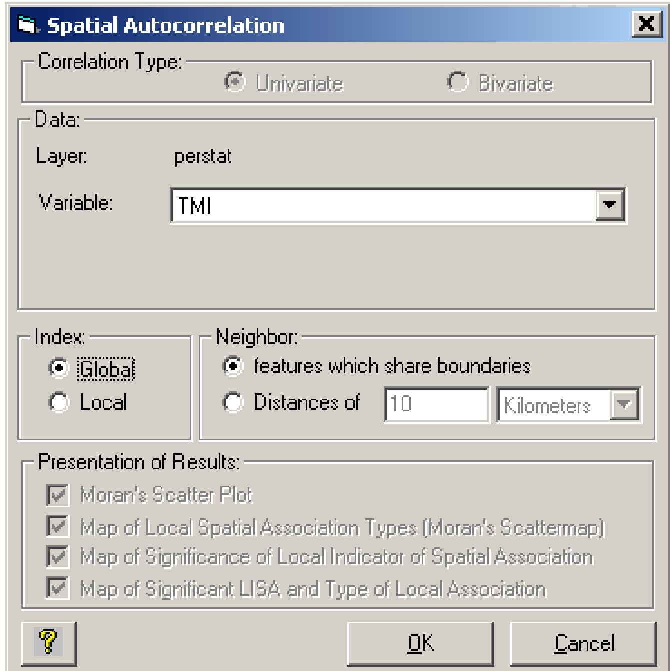
(Follow the same procedure for both global and local indexes)

1. Activate the Maps Window, and,
2. Select the Thematic Layer of interest,

   Note: The Attributes Table associated with the Layer should contain the variables and data you want to analyze.

3. Select the option **Spatial Analysis / Spatial Autocorrelation**... from the **Epi Analysis** menu on the Menu Bar,
4. The corresponding **Dialogue Box** is displayed, allowing you to select the **variable** to apply the **spatial autocorrelation** procedure to and determine whether or not a spatial distribution pattern exists:

   - The **name** of the selected **Layer** is displayed, In the **Variables** list, select the **variable** to test or investigate its possible random distribution in space from the list of variables displayed by ***clicking*** the drop-down button,



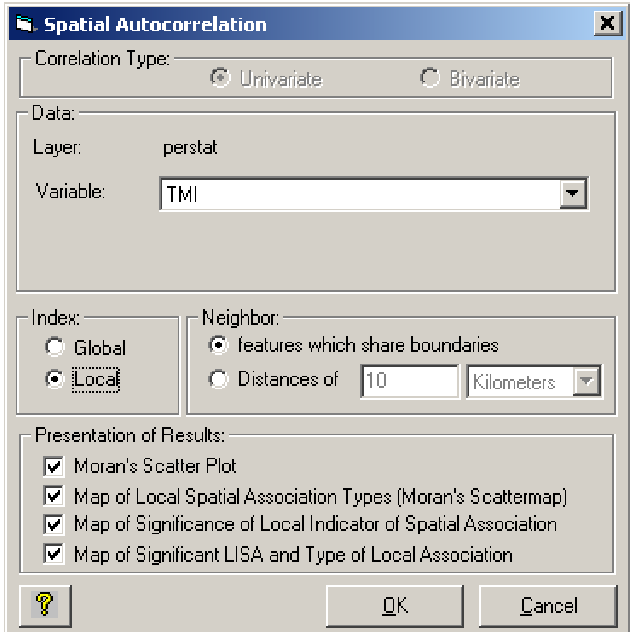
Dialogue Spatial Autocorrelation Global Index.

- ***Click*** the relevant variable or indicator,

In this case, for example, the variable TMI, representing Infant Mortality Rate is selected.

5. ***Click*** the type of **Index** to apply:
   Global (includes all geographic areas of the Layer; to calculate the global spatial autocorrelation index), or,
   Local (defines an area or neighborhood for which to apply the calculation; to calculate local spatial autocorrelation indexes),

6. If you select **Global Scope**, go to step **10,**
7. If you select **Local Scope**, the **Neighborhood** edit box is activated, allowing you to select the **type** of Neighborhood you want:



Dialogue Spatial Autocorrelation Local Index.

- **Common border**, if you only want to include in the calculation geographic areas that have a common border with the selected area,
- **Distance**, if you want to specify a **distance** and **unit of measurement** to define the area under study,

8. Select how you want to obtain the **Results Layout**, by ***clicking*** the options that appear; in this case **Significance Map for the Spatial Association Index** was selected,

9. Using the dialogue **Save Layer**, type a **name** for the new **Layer** created by this procedure.

*10. **Click** the **OK** button.*

Results:

- If you select **Global Scope**, the results are presented in the Results Window, where you can see the values obtained when applying the statistical tests described:

If **$p < 0.05$,** hypothesis H_0 is rejected. This means there is evidence that the data for the variable TMI is spatially autocorrelated, in other words, a spatial cluster pattern exist.

- If you select **Local Scope**, a new **Layer** is created, and the results are presented in a **Thematic Map using Ranges** displaying a spatial distribution for the variable and allowing you to visually determine whether or not the concentration of the variable's values reveals the existence of a spatial distribution pattern.

The new Layer is added to the **Legend** in the **Maps Window**; you can display it just as when working with any other Thematic Layer of the Map.

In this example, the **Thematic Map using Ranges** is displayed, which was obtained by applying the "Significance Map for the Spatial Association Index" procedure for the calculation of local indexes. The areas, or geographic units, with a concentration of the indicator IMR (Infant Mortality Rate) are observed in the section showing the **Thematic Layer**, revealing that there is a spatial pattern in the distribution of this indicator.

The **Attributes Table** associated with the Layer also is shown, with the newly added column showing the variable S_IMR containing the probability values calculated for each area and all the variables obtained by applying the statistical methods described.

Also displayed is the **Results Window**, containing a Table with the variable IMR's values (I_IMR, EI-IMR, DEI_IMR, ZI_IMR, P_IMR), which is stored in the Attributes Table and shows us that a spatial pattern in the indicator's distribution does indeed exist.

Time-Space Association of Cases (Knox)

SIGEpi provides the Knox statistical method for detecting the spatial and temporal interaction of health or disease events

The **Space-Time Association of Cases** procedure, through the use of the **Knox method,** allows you to identify the possible **interaction** in **space** and **time** of cases of a given disease.

Let us assume that in a small geographic area several cases appear of a disease with low incidence, which suggests a clustering in time. However, the cluster is not significant. In another period of time (e.g., the next year) this situation repeats itself, but in a different geographic area and at a different time of year. As a result, for each area by itself there is insufficient evidence for clustering in time, and the distribution over all areas does not show clustering in space because the cases have appeared in each area at a different time of year.

The same thing would happen if, over a short period of time, a limited number of cases occur in a certain small area. In the following year new cases appear in another period but in a different region, and this happens for several years. Instead of showing a trend toward spatial clustering, this shows dispersion, because the cases tend to occur in different places.

These two situations point to the need for seeking methods that make it possible to detect the presence of a joint cluster in space and time.

The purpose of the **Knox method** is to determine whether there is **significant interaction** between the cases when considering both **space** and **time**.

Description of the Statistical Method

- **Hypotheses:**

H_0: Considered together, cases of the health problem have no significant association in space and time.

H_1: Considered together, cases are significantly associated in space and time; the observed association is not attributable to chance.

- **Required data:**

You have a record of **cases** (n cases) and know both the **location** in **space** (**x,y coordinates**) and the **time** of their occurrence.

The investigator defines the location of the cases in space and in time, for example, the location in **space** could be the place of residence, or work, or any place of interest, and the location in **time** could be the date of the first symptoms.

For this procedure it is a requirement to have the data about the geographic location (latitude and longitude coordinates or x,y coordinates) in a table. This way it is possible to **create** a new cartographic layer using the tool Plotting points from a Table. The **Time variable** should be included in the table and defined as a **date** type variable.

- **Calculating the probability of obtaining a large or equal number of pairs of adjacent cases in both space and time than would be expected by chance.**

Define the critical space (E) and a critical time (T)
Determine the number of pairs of cases $N = [n(n-1)]/2$
Determine the distances among the N pairs of cases
Classify the N cases according to their adjacency in time and/or space:

1. Pairs of cases close in time and space
2. Pairs of cases close in time, but not close in space
3. Pairs of cases remote (not close) in time, but close in space

4. Pairs of cases remote in both time and space

Construct the following 2 x 2 table:

	DISTANCIA en Espacio		
DISTANCIA en Tiempo	Próximos (e <= E)	Alejados (e > E)	TOTAL
Próximos (t <=T)	x		Nt
Alejados (t > T)			
Total	Ne		N

Note that in the Table, **x** represents the number of pairs of cases close in space and time, **Nt** represents the total number of pairs close in time, **Ne** is the total number of pairs close in space and **N** the total number of pairs of cases.

The **Probability p** that a large number of cases that are adjacent (close) both in time and space and are observable by chance is expressed by:

$$p = P(X \geq x) = \sum_{k=x}^{N} P(X = k \mid \lambda)$$

donde $\lambda = (Ne * Nt)/N$

$$P(X = k \mid \lambda) = e^{-\lambda} \lambda^{k} / k!$$

- **Decision rule:**

If $p <= \alpha$ then Ho is rejected.
In this case, it is concluded that a significant space-time interaction exists, meaning it is not only by chance that the cases are close in space and time.

If $p > \alpha$ then Ho is not rejected.
It is concluded that considered together the observed clustering in space and time can be explained by chance.

Applying the Time-Space Interaction statistical method

Using **SIGEpi,** take the following steps to determine whether health related cases interact in space-time:

1. Make sure you have the **data Table** containing the **space** variables (x,y coordinates for latitude and longitude) and the **time** variable in order to create the Thematic Point Layer; or, if this Layer already exists,

2. Activate the Map Window, and add the **Point Layer** to it if you have not already done so,

3. Select the Point Layer,

4. Select the option **Association of cases in Time-Space (Knox)…** from the **Epi Analysis** menu on the Menu Bar,
   The **dialog box** for **Time-space Interaction (Knox)** is displayed, allowing you to select, from the Attributes Table associated with the selected Thematic Point Layer, the **variable** that records the **time** (e.g., date of first symptoms, date of first doctor's visit, etc.),

5. ***Click*** the ▼ drop-down button; this shows all the **variables** or columns for the **date** in the Attributes Table associated with the Layer, ***click*** the **variable** that contains the **date** you want to use to analyze the time-space interaction,

6. Define the **critical Space** value by typing the desired value, and select the **unit of measurement,**

7. Define the **critical Time** value by typing in the desired value,

8. ***Click*** the **OK** button.

This yields a **probability value p** enabling you to decide whether to accept or reject the hypothesis for interaction of cases in space and time.

The results are shown in the Results Window.

To reject or accept the null hypothesis, you need only compare the calculated probability value **p** with the Alpha value you set.

In this example, a p value was obtained equal to 0.66; if the Alpha value established was 0.05, we cannot reject the null hypothesis and can conclude that the cases do not have a significant interaction in space and time, or, in other words, that they are not grouped (clustered) in space and time.

Exposure-Effect Association

SIGEpi provides a set of quantitative methods for epidemiological analysis of exposure-effect association of health or disease events

The **Exposure-Effect Association** procedure provides a set of measures and statistical techniques allowing you to detect a possible association between a **spatial exposure factor** and an **effect** being studied. This procedure can be applied in epidemiological Cohort studies and unmatched Case-Control studies where the exposure factor or factors are environmental.

Description of the Statistical Methods

The procedure is based on statistical methods detailed in the User Manual's bibliography. To avoid going into too much detail, we decided not to include here the mathematical description of the methods used. For a complete description of these methods, we recommend consulting the chapter **Métodos Cuantitativos de la Epidemiología Analítica, en Epidemiología. Principios - Técnicas - Aplicaciones** by **Milos Jenicek** and **Robert Cléroux**.

The use of these methods in **SIGEpi** is explained below.

- **Necessary data:**

To begin with, it is necessary to have two cartographic layers:

a.) One **Layer** should contain: the **exposure factors(s)** that you want to study,

For example: a Layer that contains the areas or geographic units that represent a type of crop; that crop can be considered an exposure factor related to the disease being studied. Another example is a Thematic Layer containing the buffers for geographic units representing industries or some event of interest.

b.) Other **Layer** should be a Point Layer representing the individuals of the sample population under study,

Among the **variables** or columns in the Attributes Table associated with the **Thematic Point Layer** for the sample population, it is necessary to have one **variable** that is a yes/no value, indicating whether it is a case (a patient) or not a case (not a patient). For example: 1 for patient and 0 for not a patient, or S for patient and N for not a patient.

For the **Thematic Layer of exposure factors**, you may define several levels for a single exposure factor. For example: if the exposure factor being analyzed is the buffer for the contamination arising from a certain industry; it may be advisable to define several buffers. Let us say, within 1 km of distance is the area of higher exposure; from 1 to 5 km can be another area for a second level of exposure to study, and so on.

If you define only one buffer in the Layer containing data for the exposure factor, the system performs an Exposure-Effect Association analysis for a single exposure level. If, on the contrary, you have several buffers, **SIGEpi** automatically performs an Exposure-Effect Association analysis for several exposure levels.

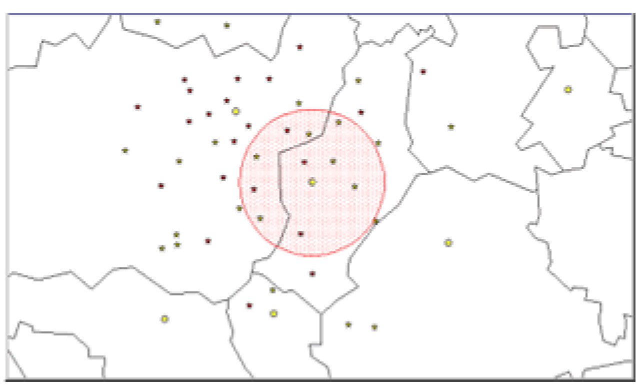
Another possibility that **SIGEpi** offers is to perform stratified analyses; i.e., to see how different population strata react to the exposure factor under study. Strata are defined by a categorical variable in the **Attributes Table** for the sample population, for example, Sex or Age.

Applying the Exposure-Effect Association method

Let us now see an example of how to perform this analysis in **SIGEpi**.

1. Activate the Map Window, and,
2. Make sure there are the two layers available in the map as described above, one of **Exposure Factors** and other of **Sample Population,**

Here we see the representation in a Map Window.



Displaying Thematic Layers of a geographic area.

The circle represents a buffer for the Exposure Factor under study.

The stars represent individuals, classified as a Patient or not a Patient depending on a yes/no variable, in this case called **Patient**.

3. Select the Thematic Point Layer,
4. Select the option **Exposure-Effect Association** from the **Epi Analysis** menu on the Menu Bar

The **dialog box** for **Exposure Factor - Effect Association** is displayed, allowing you to select the **variable** containing the cases of patients from the Attributes Table associated with the selected Thematic Point Layer,

Dialogue box. Exposure-Effect Association.

5. ***Click*** the drop-down button of the **Theme** edit box; if you want to select **another** Thematic Layer that has exposure factors. **SIGEpi** identifies and initially displays a Layer that has such characteristics, ***click*** the Layer you want,

6. ***Click*** the drop-down button of the **Sample** edit box; if you want to select **another** Thematic Point Layer, **SIGEpi** displays the active Point Layer initially; ***click*** the Layer you want,

7. ***Click*** the drop-down button of the **Effect Variable** edit box; this shows the **variables** or columns of the Attributes Table associated with the Point Layer, ***click*** the **variable** that separates the positive and negative Cases, in this example, the **Patient** variable,

8. ***Click*** the drop-down button of the **Positive Case Value** edit box and select the variable's Positive Case Value from the list of values contains by ***clicking*** the corresponding value,

9. As described above, if you want to conduct a stratified analysis--for example, by **Sex**--select the corresponding **variable** from the list displayed in the **Strata** edit box,

10. ***Click*** the **OK** button.

The results obtained by applying the procedure are added to the **SIGEpi** Results Window; by analyzing these results you can find the exposure-effect association for the cases being studied.

Chapter
8

Obtaining Results

SIGEpi provides a Results Sheet that summarizes the analytical and statistical procedures applied to your data

SIGEpi allows applying a set of statistical, epidemiological and exploratory analysis procedures to the data with emphasis on spatial data. Most of these methods generate results in form of numbers, text and tables.

The **Results** of the analytical processes applied to the data appear in the **SIGEpi** Results Window.

Each **SIGEpi** Project has only one Results Window containing the results from each analytical procedure applied to the data. The Results Window displays the resulting information in HTML format.

SIGEpi offers tools, procedures, and methods for spatial analysis in **Epidemiology** integrated with the functions of a **Geographic Information System**. This maximizes its potential use in Public Health since you can visualize and spatially analyze statistical and epidemiological data from a study region, providing decision makers, epidemiologists, and health technicians with a simple, user-friendly way of applying these procedures to their data and obtain a spatial display of their distribution and representation of the analyzed variables.

Results Window

The **Results Window** contains the results obtained by applying the following spatial, statistical, and epidemiological analysis procedures to the data:

- **Descriptive Statistics**
- **Frequency Distribution**
- **Correlation Analysis**
- **Regression Analysis**
- **Outlier Map**
- **Spatial Autocorrelation**

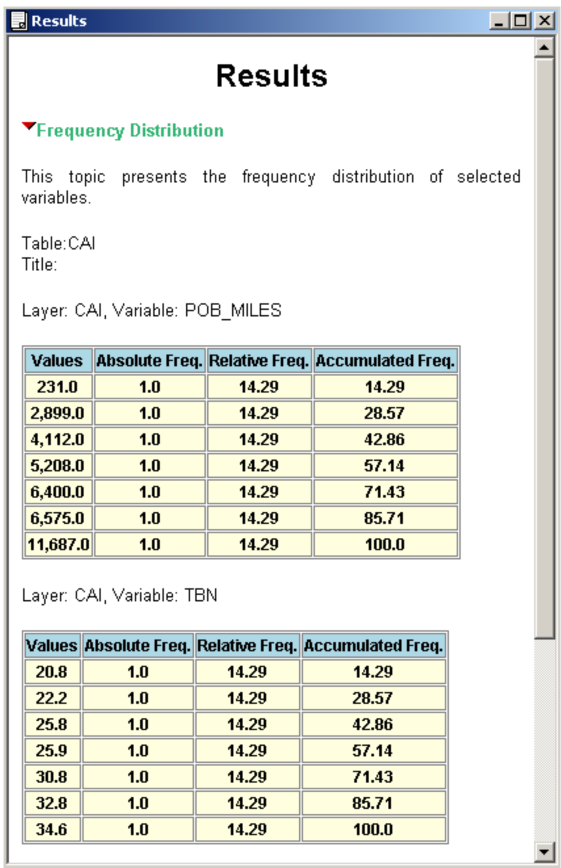
Displaying the Results Sheet

To display the Results Window:

1. Activate the Project Window,
2. **Double *click*** the **Results** component in the Project Window,

3. Or, Select the **Results** component by ***clicking*** the component, and,
   a.) ***Click*** the **Show Window** button on the Toolbar that appears on the left of the Project Window,
   b.) Or, Select the **Show Results** option from the **Project** menu on the Menu Bar.

- The Results Window in **SIGEpi** contains the Tables showing the results obtained by applying the statistical and epidemiological analysis procedures to your data; they are successively added to the document as they are carried out. You can activate or minimize this window as necessary.


Results

Results

Frequency Distribution

This topic presents the frequency distribution of selected variables.

Table:CAI
Title:

Layer: CAI, Variable: POB_MILES

Values	Absolute Freq.	Relative Freq.	Accumulated Freq.
231.0	1.0	14.29	14.29
2,899.0	1.0	14.29	28.57
4,112.0	1.0	14.29	42.86
5,208.0	1.0	14.29	57.14
6,400.0	1.0	14.29	71.43
6,575.0	1.0	14.29	85.71
11,687.0	1.0	14.29	100.0

Layer: CAI, Variable: TBN

Values	Absolute Freq.	Relative Freq.	Accumulated Freq.
20.8	1.0	14.29	14.29
22.2	1.0	14.29	28.57
25.8	1.0	14.29	42.86
25.9	1.0	14.29	57.14
30.8	1.0	14.29	71.43
32.8	1.0	14.29	85.71
34.6	1.0	14.29	100.0

Results Window

Selecting and Copying Results

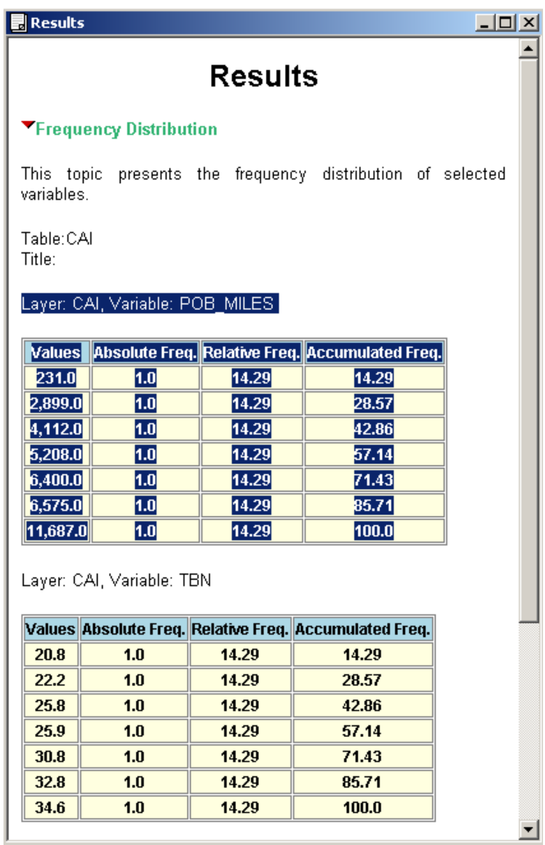
To Copy Results individually:

1. Activate the Results Window,
2. **Select** the elements that you want to copy using the Windows standards (position the cursor at the starting point, ***click and hold down*** the left mouse button, and move the cursor until you select the desired content, then release the mouse button),

You can also easily select all the content of the active Results Window.

3. Select the **Copy** option from the **Edit** menu on the Menu Bar,

This allows you to copy the selected content from the Results Window to the Clipboard and later add them to any document you want. For example, you can add it to a Report or document in MS-Word. This selection remains on the clipboard only until another object is copied there or the PC is turned off.

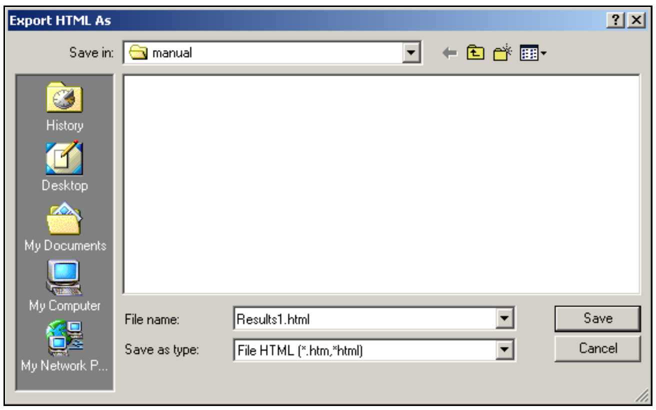


Creating an HTML file

In **SIGEpi** you can create an **HTML** file that stores all the content displayed in the Results Window of the Project and were obtained from the statistical and epidemiological analysis procedures applied to your data.

Creating a HTML file:

1. Activate the Results Window,
2. Select the option **Export As HTML...**, from the **Edit** menu on the Menu Bar,
3. Establish a name for the HTML file in the displayed **dialog box**,



Dialogue Export HTML As...

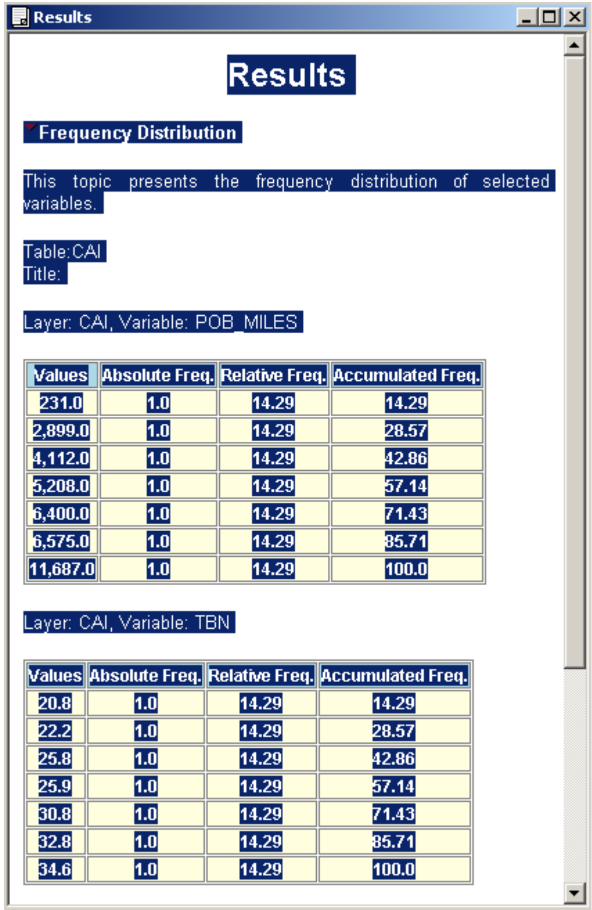
As a result, a HTML file is created and stored with a specified name and location, and with which you can work independently.

Selecting all the results

To select all the elements included in the Results Window:

1. Activate the Results Window, and,
2. Choose the **Select All** option from the **Edit** menu on the Menu Bar.

As a result, all the elements displayed in the active Results Window are highlighted. This way you can copy all the results to the Clipboard and add them to any document.



Results

Results

Frequency Distribution

This topic presents the frequency distribution of selected variables.

Table:CAI
Title:

Layer: CAI, Variable: POB_MILES

Values	Absolute Freq.	Relative Freq.	Accumulated Freq.
231.0	1.0	14.29	14.29
2,899.0	1.0	14.29	28.57
4,112.0	1.0	14.29	42.86
5,208.0	1.0	14.29	57.14
6,400.0	1.0	14.29	71.43
6,575.0	1.0	14.29	85.71
11,687.0	1.0	14.29	100.0

Layer: CAI, Variable: TBN

Values	Absolute Freq.	Relative Freq.	Accumulated Freq.
20.8	1.0	14.29	14.29
22.2	1.0	14.29	28.57
25.8	1.0	14.29	42.86
25.9	1.0	14.29	57.14
30.8	1.0	14.29	71.43
32.8	1.0	14.29	85.71
34.6	1.0	14.29	100.0

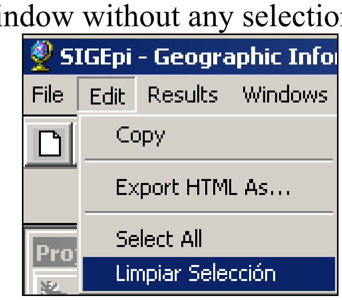
Clearing an existing selection

While working with the Results Window in **SIGEpi,** you may want to delete selections previously made when selecting some Results and Tables to copy to the Clipboard to add to a document, or when you have selected all the results in the Window.

To clear an existing selection:

1. Activate the Results Window, and,
2. Select the **Clear Selection** option from the **Edit** menu on the Menu Bar,

This displays the active Results Window without any selections of elements in the Window.



Chapter

9

Creating Layouts

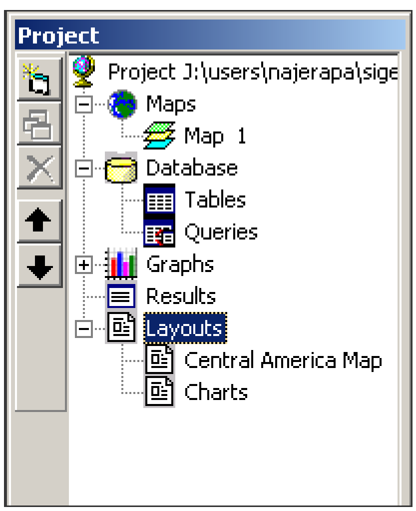
SIGEpi allows you to group and organize all your data analysis into a Layout

A **Layout** is a component of the Project in **SIGEpi** that allows you to show in the System window or in a printed document any components or elements from those you have developed and have been working with: Maps, Graphs, and Texts. A **Layout** is ordinarily used to prepare a print-out from **SIGEpi**.

In a **Layout,** you select the information you want to display and define their appearance (size, color, etc). This data can be part of different Layouts in **SIGEpi**; you can create as many Layouts as you want. If you are creating a **Layout** for marketing, for example, you will probably want to present the data differently than for a Layout on Health Analysts.

Creating and Editing Layouts

The **SIGEpi** Project can contain several Layouts, each related to its **name** in the list of **Layouts** in the Project Window.



Project Window. Layouts.

A **Layout** will create a 8 ½ x 11 size document, which you can include and organize as many elements of Maps, Graphs, and Texts as you want.

The **properties** of the Layout Window can be defined; you can vary the size of the imported components and change the name of the Layout.

To create a Layout:

1. Activate the Project Window,
2. ***Click*** the **Layouts** component,
3. ***Click*** the **New Window** button on the Toolbar displayed on the left of the Project Window,
4. Or, Select the **New Layout**... option from the **Project** menu on the Menu Bar.

   In any of these cases, a new **Layout Window** is created and displayed, in which you can represent and organize your data.

   Note that activating the Layout Window activates the corresponding menus and buttons on the Menu Bar.

5. If you want to change the Properties of the Layout Window, ***click*** the **Properties** option from the **Layout** menu on the Menu Bar.

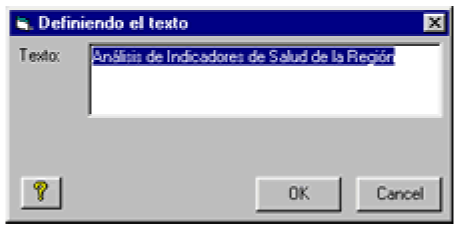
   > Note: If you want to modify the colors and special characteristics of Maps and/or Graphs to be included in the Layout, you should do this in the Window corresponding to each of these elements. A useful way of utilizing a Layout could be, for example, to improve the quality of the Maps that you want to print if you do not have a color printer.

6. You can edit or change each element in the Layout Window by **right *clicking*** the desired element (Map, Graph, Text) and ***clicking*** the action you want in the displayed menu: modify **properties**, **delete**, or **update** the selected element, and **move** or vary its **size**.

Adding Text to a Layout

To add Text to a Layout:

1. With the Layout Window active,
2. ***Click*** the **A** **Add Text** button on the Toolbar,
3. Or, Select the **Add Text**... option from the **Layout** menu on the Menu Bar,
4. Type the **text** that you want to include in the Layout in the displayed **Dialogue Box**,



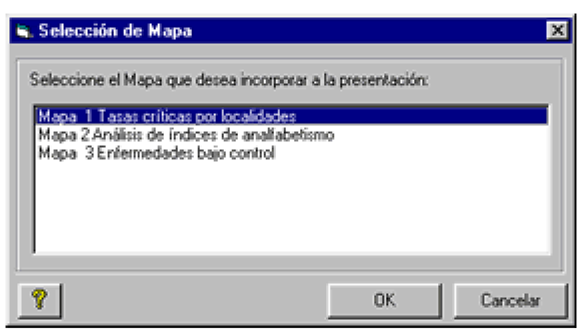
Dialogue box. Define Text.

5. ***Click*** on the inserted **Text** to move it to the position that you want.

Adding a Map to a Layout

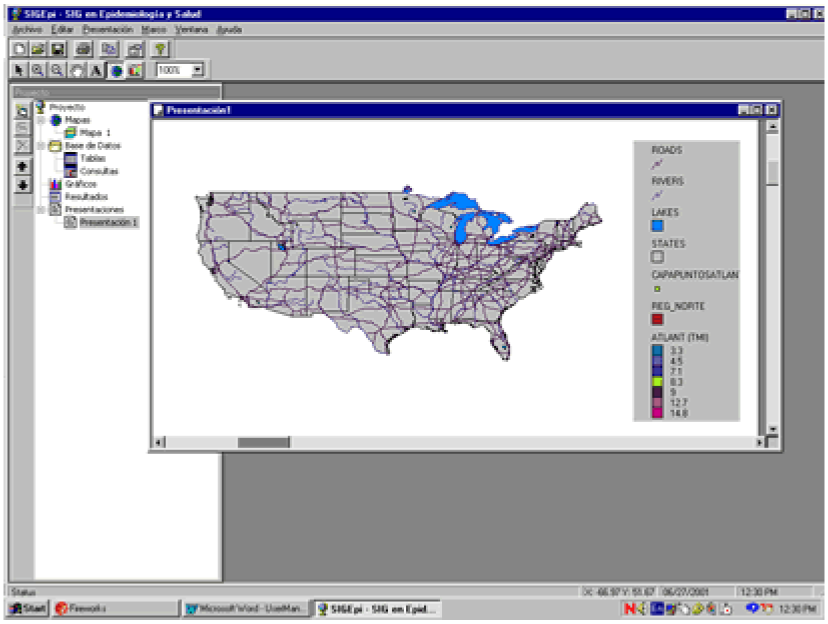
To add a Map to a Layout:

1. With the Layout Window active,
2. ***Click*** the **Add Map** button on the Toolbar,
3. Or, Select the **Add Map**... option from the **Layout** menu on the Menu Bar,
4. ***Click*** the starting point of the area where you want to insert the **Map** in the Layout sheet, move the cursor to the desired position, and release the button on the mouse,
5. Select the Maps Window that you want to insert in the Layout, based on the displayed **Dialogue Box** showing the list of all the **Maps** that have been designed in the Project,



Dialogue box. Select a Map.

6. In the same way as for a **Text**, you can ***click*** the inserted Map to move it to another location in the Layout.

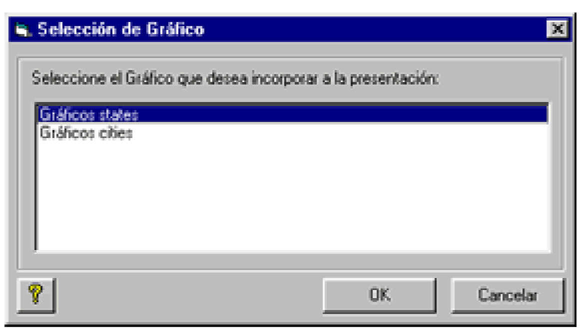


Layout Window. Adding a Map.

Adding a Graph to a Layout

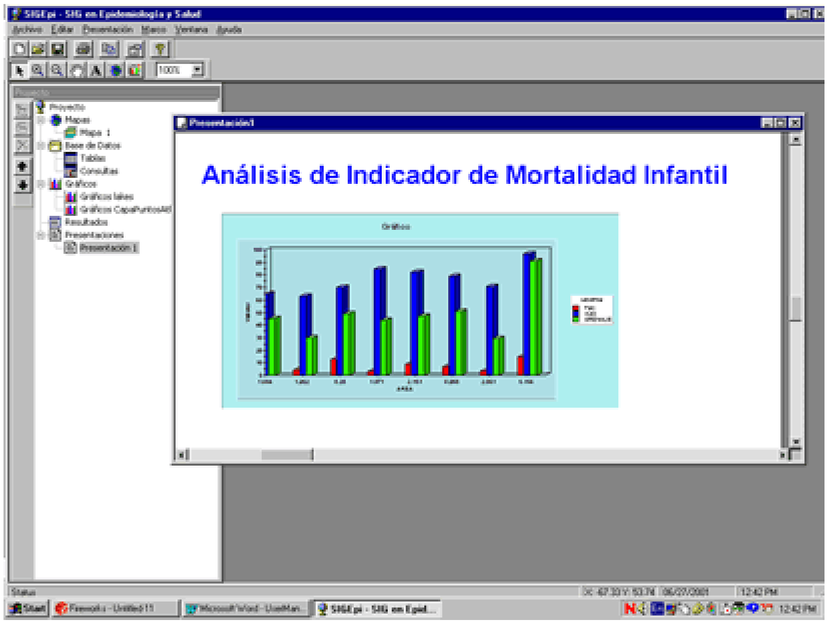
To add a Graph to a Layout:

1. With the Layout Window active,
2. ***Click*** the **Add Graph** button on the Toolbar,
3. Or, Select the **Add Graph**... option from the **Layout** menu on the Menu Bar,
4. ***Click*** the starting point of the area where you want to insert the **Graph** in the Layout sheet, move the cursor to the desired position, and release the button on the mouse,
5. Select the Graph that you want to insert in the Layout, based on the displayed **Dialogue Box** which lists all the **Graphs** that you have created in the Project,



Dialogue box. Select a Graph.

6. Just as you did for inserting a **Text** or a **Map**, you can ***click*** the inserted **Graph** to move it to another location in the **Layout**.



Layout Window. Adding a Graph

Showing a Layout

To show or display a Layout:

1. Activate the Project Window,

2. **Double** ***click*** the **name** of the Layout Window that you want to display, from the list of layouts in the  **Layouts** component in the Project Window,

3. Or, Select the **Layout** that you want to show, by ***clicking*** its **name**, and,

   a.) ***Click*** the  **Show Window** button on the Toolbar that appears on the left of the Project Window,
   b.) Or, Select the **Show Layout** option from the **Project** menu on the Menu Bar.

Technical Support

The Area of Health Analysis and Information Systems (AIS) of the Pan American Health Organization (PAHO) offer the required technical support to registered users. This support will be offered using various means and strategies, such as training both in-site and distance modalities; through the implementation of a discussion forum and by direct contact of the users with the development team.

- **Training (Courses, Workshops, Training Sessions)**: The GIS Team of AIS organizes periodically several courses and workshops on GIS application in Epidemiology and Public Health. Information about schedules of the workshops will be published at the Website of SIGEPI Project at http://ais.paho.org/sigepi

Courses/workshops are also offered, including other modalities of technical training by request. These kind of training can be carried out in the Centers that request cooperation.

- **SIGEpi discussion forum:** A discussion forum about SIGEPI Project that allows the exchange of ideas, ask and answer questions, collaboration between groups and the user's community has been implemented. The users can subscribe themselves to the forum visiting the site: http://listserv.paho.org/archives/sisgepi.html, or sending a request by Email to sigepi_info@paho.org

- **Direct contact**: The users can establish direct contact by E-mail with the SIGEPI Development Team sending emails to sigepi_info@paho.org or for more information visiting the Web site http://ais.paho.org/sigepi

Credits

The Area of Health Analysis and Information Systems (AIS), Pan American Health Organization (PAHO), as a part of the Technical Cooperation Project "Application and Development of Geographic Information Systems in Epidemiology and Public Health", have developed SIGEpi.

The SIG-Epi Design and Development team:

- Mr. Ramón Martínez-Piedra
  Information and Technology Group AIS/ PAHO
  Design, Development and Programming of SIGEpi

  Health Analysis and Information Systems
  Pan American Health Organization

- Dr. Carlos Castillo-Salgado
  Manager of AIS/ PAHO

  Health Analysis and Information Systems
  Pan American Health Organization

- Dr. Enrique Loyola
  Analysis Group, AIS/ PAHO

  Health Analysis and Information Systems
  Pan American Health Organization

- Mr. Manuel Vidaurre
  Information and Technology Group, AIS/ PAHO

  Health Analysis and Information Systems
  Pan American Health Organization

Other participants in the development of the Help System and User's Manual of SIGEpi:

- Ms. Katia Díaz-Morejón
  Consultant, Information and Technology Group, AIS/ PAHO

For additional information and comments, contact:

Dr. Carlos Castillo-Salgado
Manager AIS/ PAHO

Health Analysis and Information Systems Area
Pan American Health Organization

525 23rd Street, Washington, D.C. 20037
Telephone: (202) 974-3327
Fax: (202) 974-3674

E-mail: sha@paho.org

Glossary of terms

Algorithm: Explicit and finite sequence of operations that leads to the solution of a problem. Applied to GIS, this usually refers to a series of algebraic operations on maps and/or databases that allow you to obtain a result by combining spatial and alphanumeric information.

Application: A process or series of processes that uses data or performs functions on a computer system.

Attribute: Most commonly refers to a column of a table in a database. For spatial data, it represents a characteristic of a geographic figure described by values, characteristics, or an image, stored in table form and linked to the geographic figure to which it refers. Property or characteristic of a class of elements of a database.

Autocorrelation: Refers to the degree of relationship that exists between one or more variables; this means that when one changes, the others also change. The change can be in the same direction, which is a positive autocorrelation, or the opposite direction, which is a negative autocorrelation.

Cartography: Set of techniques used in the construction of Maps.

Database: A collection of data organized such that searching and updating is fast and easy. Set of data structured for storage, querying, and updating in a computer system.

Database Management System (DBMS): Computer system designed for the creation, modification, correction, updating, and querying of databases.

Database System Administration: The relationship between a set of data stored in a database and the computer programs or software that administers the database.

Datum: Verifiable fact about the real world. A datum can be a measurement, an equation, or any kind of information that can be verified.

Entity: An object that is distinguishable from other objects by its characteristics.

Geographic Information System (GIS): Organized set of equipment, program packages, geographic data, processes, work methods, and personnel, combined into an automated system whose principal purpose is the processing of spatial data. GIS are designed to capture, store, update, manipulate, analyze, and display different forms of geographically referenced information in an integrated operation.

Georeference: Assignment of geographic coordinates to an object or structure. Applied to a digital image, this implies a series of geometric operations that make it possible to assign a pair of coordinates (x,y) to each pixel of the image in a given projection.

Georeferenced Data: This establishes the relationship between coordinates on a planar Map and the real location of the coordinates on the Earth.

GPS (Global Positioning System): System for place localization, through which the current coordinates of a land station can be estimated through simultaneous reception of signals from several satellites.

Graph: Representation of data on coordinate axes.

Grid: Network of uniformly spaced parallel lines that intersect at right angles. When they are overlaid on a Map, the grid is usually referred to by the name of the projection used, such as Lambert grid, transverse Mercator grid, and universal transverse Mercator grid.

Grid-cell: Basic information element in a raster matrix structure. A two-dimensional object that represents one element selected from a surface. This term is usually used to refer to a simple element in a raster data structure.

Image: Visible representation of geographic objects and/or phenomena obtained by remote sensing, cameras, scanners, radar, etc.

Legend: Ordered and structured list of the relationships between symbols and values for the variables represented on a Map.

Map: Graphic representation of the physical features (natural, artificial, or both) of a part or all of the Earth's surface, represented by symbols, with an established scale, on a specific projection, and with the orientation indicated.
It is a graphic model of the terrestrial surface that represents spatial objects and their measurements, topological properties, and attributes.
A Map can be analog (printed on paper, for example) or digital (coded in figures, stored on a computer and displayed on a screen).

Matrix: Data structure formed by elements (grid-cells) arranged regularly in rows and columns. The matrix is the structure most often used for the construction of digital models of terrain and digital images; in the latter case, each element of a matrix is called a pixel.

Pixel: Small element in a drawing. It is the smallest element represented in an image. A pixel has two attributes: its spatial location and the value that characterizes it. Each discrete element into which a digital image is divided is a pixel.

Raster: Data model in which reality is represented by elements as polygons that form a regular mosaic. Each polygon of the mosaic is a unit area that has the mean value of the represented variable. A **raster** data model is based on locations.

Raster Data: Raster data put spatial information in a regular matrix of grid-cells or digital image grid organized as a group of rows and columns. Each grid-cell in this matrix contains a value representing a particular geographic characteristic, such as elevation, irregularity of terrain, etc.

SQL (Structured Query Language): A standard language for database management. SQL has become a standard. As a result, queries in this language provide access to databases from diverse sources. An SQL statement is a set of procedures that allow you to select or group the data.

Smoothing: Reduction of local variability of data, and, when applied to a spatially distributed variable, reduction of local variance.

Spatial Analysis: Analytical techniques associated with the study of locations of geographic phenomena, together with their spatial dimensions and associated attributes.

Spatial Data: Information about the location, shape, and relationships among geographic features, usually stored as coordinates or topology.

Table: A set of data organized into rows and columns. The columns represent the attributes or variables of the entities they characterize, and the rows represent the values of the variables taken by those entities.

Thematic Layer: Set of geographic characteristics displayed on a Map, which can be grouped by the subject they describe. Example: a Thematic Layer of rivers, cities, communication routes, etc.

Thematic Map: A Map that displays selected types of information related to one or more specific subjects, using variations in color. Example: soil type, classification of terrain, population density, and amount of rainfall.

Vector: Geometric entity defined by a magnitude and a direction. A vector is made up of a pair of ordered points. A vector model is a data model in which reality is represented by vectors or vector structures.

Vector Data: An abstraction from the real world where the positional data is represented in the form of coordinates. It is made up of a series of georeferenced points, where a line is a collection of points and an area or polygon is a collection of related lines.

Zoom in/Zoom out: Moving in (to enlarge) or moving out (to reduce) an image or part of an image displayed on a computer screen.

Bibliographic references

1. *Jenicek, M. & Cléroux, R. Epidemiología. Principios - Técnicas - Aplicaciones. Salvat.*
2. *Schlesselman, J. J.(1982), Case-Control Studies. Design, Conduct, Analysis. Oxford University Press. New York Oxford.*
3. *Carvajal, R. (1994), Métodos Estadísticos para el Análisis Epidemiológico. PAHO/HPP/HPS/94.2*
4. *Fleiss, J. L. (1981), Statistical Methods for Rates and Proportions, 2nd ed. John Wiley, New York.*
5. *Pickle L.W. & White A.A., Effects of the choice of age-adjustment method on maps of death rates. Statistics in Medicine. (1995); 14: 615-627.*
6. *Walter, S. D., The Analysis of Regional Patterns in Health Data. I. Distributional Considerations. American Journal of Epidemiology. (1992); 136(6):730-741*
7. *Walter, S. D., A Simple Test for Spatial Pattern in Regional Health Data. Statistics in Medicine.(1994);13:1037-1044.*
8. *Haining R., Spatial data analysis in the social and environmental sciences. Cambridge University Press. (1990).*
9. *Knox, G., Epidemiology of childhood leukemia in Northumberland and Burham. Br. J Prev. Soc. Med. (1964);18:17-24*
10. *Pike, M. C. & Smith, P. G., Diseases clustering: a generalization of Knox's approach to the detection of space time clustering. Biometrics. (1968); 24:541-556.*
11. *Garson, G.D., & Biggs, R. S. (1992), Analytic mapping and Geographic databases (Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-087). Newbury Park, CA:Sage.*
12. *Clark, K. C., McLafferty, S. L., Tempalski, B. J., On Epidemiology and Geographic Information Systems: A Review and Discussion of Future Directions. Emerging Infectious Diseases, (1996); 2 (2): 85-92*
13. *Special Program for Health Analysis (SHA), PAHO, (2000), Geographic Information Systems in Health, Basic Concepts.*