Epi Info™
Community Health Assessment Tutorial

DOCUMENT VERSION 2.1.2 **REVISED AUGUST 2011**

Department of Health and Human Services
Centers for Disease Control and Prevention
National Center for Public Health Informatics

This page intentionally blank.

Acknowledgements

The Epi Info™ Community Health Assessment Tutorial was produced by the collaborative efforts of the Centers for Disease Control and Prevention (CDC), the Assessment Initiative (AI), and the New York State Department of Health (NYSDOH).

CDC is one of the 13 major operating components of the U.S. Department of Health and Human Services. CDC is at the forefront of public health efforts to prevent and control infectious and chronic diseases, injuries, workplace hazards, disabilities, and environmental health threats. Today, CDC is globally recognized for conducting research and investigations, and for its action-oriented approaches to issues of public health. CDC uses its research and findings to improve people's daily lives, and respond to local, national, and international health emergencies.

The AI comprises a cooperative program between CDC and state health departments that supports the development of innovative systems and methods to improve the way data is used to provide information for public health decisions and policy. Through the AI, funded states work together with local health jurisdictions and communities to improve access to data, improve skills to accurately interpret and understand data, and use of the data so that assessment findings ultimately drive public health program and policy decisions.

CDC and the New York State AI agreed to a new five-year cooperative agreement to strengthen assessment capacity and practice. This is the third round of AI funding offered through CDC, and is the second time that the NYSDOH has received the award.

The CDC Epi Info™ development team, at the request of the NYSDOH Assessment Initiative, created this tutorial for using the Epi Info software program. It has been designed using asthma data provided to CDC by the NYSDOH.

Copyright Information

Centers for Disease Control and Prevention (CDC)

Director, Thomas Frieden, MD, MPH

Associate Director for Science, Harold W. Jaffe, MD, MA

Coordinating Center for Health Information and Service (CCHIS)

Director, Steven L. Soloman, Rear Admiral, USPHS

National Center for Public Health Informatics (NCPHI)

Director, Leslie Lenert, MD, M.S.

Public Health Surveillance Program Office (PHSPO)

Director, James W. Buehler, MD

Epi Info™ Development Team:

Instructional Design: Shannon Rowlett Jones

Epi Info™ Technical Advisors: Karen DeRosa, Roger Friedman, Andre Swoope

ITB Advisors: David Nitschke, Carol Worsham, Marilyn Reynolds

AI Advisors: Alex Charleston, Nelson Adekoya

CDC Editors: W. Scott Janes, Doug Bialecki

New York State Department of Health (NYSDOH), Public Health Information Group

AI Advisor: Priti Irani, MS, Project Director, Assessment Initiative

Technical Review Board: Cate Bohn, MPH, Community Health Information Specialist; Aaron Mair, MPA, Public Research Specialist III; Mike Medvesky, MPH, Director; Trang Nguyen, MD, MPH, New York State Asthma Epidemiologist; Cheryl Ellemberg, MPH, former Research Specialist III

NYSDOH, Bureau of Communicable Diseases

Technical Review Board: Hwa-Gan Chang, PhD, Director Statistical Unit

NYSDOH, Information Systems and Health Statistics Group

Technical Review Board: Linh Le, MD, PhD, Research Scientist IV

University at Albany School of Public Health

Technical Review Board: Mike Zdeb, Assistant Professor

County Workgroup

Content Review Board: E. Oscar Alleyne, MPH, Epidemiologist; Cristina Dyer-Drobnack, MS, Project Coordinator, NYSACHO; Jan Herrick Director, Environmental Health, Schoharie County Department of Health; Stephen A. Jennings, Public Health Planner, Jefferson County Public Health Service; Patricia Many, BS, RN, Public Health Nurse, Oneida County Health Department; Joanne Ryan, RN, Communicable Disease Nurse, Putnam County Department of Health; Ted Schiele, MS Planner/Evaluator, Tompkins County Health Department; Christopher M. Szwagiel, MS, MPH, DrPH, Director Genesee County Department of Health; Karen Tuchman, RN, MPH, Public Health Nurse, Putnam County Department of Health.

This page intentionally blank.

Contents

Introduction

Epi Info™ Health Assessment Tutorial

Epi Info™ is a series of programs for Microsoft Windows for use by public health professionals in conducting outbreak investigations, managing databases for public health surveillance and other tasks, and general database and statistics applications. With Epi Info™, a user can rapidly develop a questionnaire or form, customize the data entry process, and enter and analyze data.

The Epi Info™ program is divided into four core modules: MakeView, Enter Data, Analysis, and Epi Map. Each module has stand-alone capabilities; however, using them for one project from start to finish adds functionality and efficiency. The MakeView module is the questionnaire or survey designer. You will start your project in MakeView, where you will create a survey and begin setting up your database. Using Enter Data, you will populate your database with information specific to your survey. The data entered, along with additional datasets, will be imported into Analysis for detailed study. Data will be plotted in the Epi Map module. This training does not cover every component of Epi Info™ and is intended for the beginning- and intermediate-level user.

How to Use this Tutorial

This tutorial introduces you to the basic concepts of Epi Info™. Each lesson builds upon knowledge gained from the previous one. It is highly recommended that you follow the lessons in the order presented. **Finish all the listed steps and required tasks before progressing to the next lesson as many of the lessons build upon data gathered from a previous lesson.** Data specific to the course and the scenario will be provided to you as part of the training package. At various points in the course, you will be asked to access data to complete the lessons. Read the five goals listed at the beginning of each lesson to review the core functions you should understand after completing each lesson.

Step One

Read the Overview section for each lesson. Each Overview contains a Getting Started breakdown of the module and information on ways to navigate. The Overview also provides information on the commands, dialog boxes, and other special features unique to each module as well as information on navigating the module. Epi Info™ uses standard windows features for navigation (e.g., drop-down boxes and menu navigation), as well as unique features [e.g., wild card (*) searches]. The Overview also provides you with screenshots of the module and introduces you to the basic concepts used in each lesson.

Step Two

Complete all the step-by-step instructions to learn the basics. Numbered instructions with some screenshots are provided for you to learn the core functions of each module. Numbered lessons and instructions can be followed and applied when using Epi Info™ outside of the classroom.

Step Three

Complete the Practice Lessons and Skills Review Exercises when provided. Practice (Required Skills) lessons build upon skills learned during basic lessons. Practices and reviews are opportunities to try skills without detailed instruction. Skills Reviews will consist of hands-on activities to further your understanding or a series of questions meant to reinforce basic knowledge. An Intermediate Analysis lesson is provided at the end of the course. The Intermediate Analysis Lesson builds upon the skills from earlier lessons and includes tasks that are more complex. Each section in Intermediate Analysis includes actual data examples from sources [e.g., U.S. Census Bureau and the Behavioral Risk Factor Surveillance System (BRFSS)].

Epi Info™ Resources

Training documentation, technical information, or other resources are available on the Epi Info™ website: www.cdc.gov/epiinfo.

For questions, problems, or more information, contact the Helpline: epiinfo@cdc.gov or call 404.498.6190.

To speak with colleagues around that world about Epi Info™, visit phConnect at http://www.phconnect.org/group/epiinfo.

File Set-Up

In the steps below, you will set up your files.

Before we begin, you must have the latest version of Epi Info™ installed on your computer. Check the Epi Info™ website for updates: www.cdc.gov/epiinfo.

Once installed, Epi Info™ creates a folder on the C: drive called Epi_Info. Your tutorial datasets must be installed inside this folder.

Now that you're armed with a keyboard and mouse - let's get rolling!

1. From the desktop, open **My Computer**.
2. Open your **Local Disk** drive, usually named **C**.
3. Open the **Epi_Info** folder.
4. If the EIHA Tutorial folder does not exist, select **File>New>Folder** from the navigation menu.
    - If you have downloaded the EIHA Tutorial folder to a different location, right click the folder and copy/paste it to the new location inside the Epi_Info folder on your C drive.
5. Name the folder **EIHA Tutorial**.
6. Copy (or download) the tutorial datasets to the EIHA Tutorial folder.

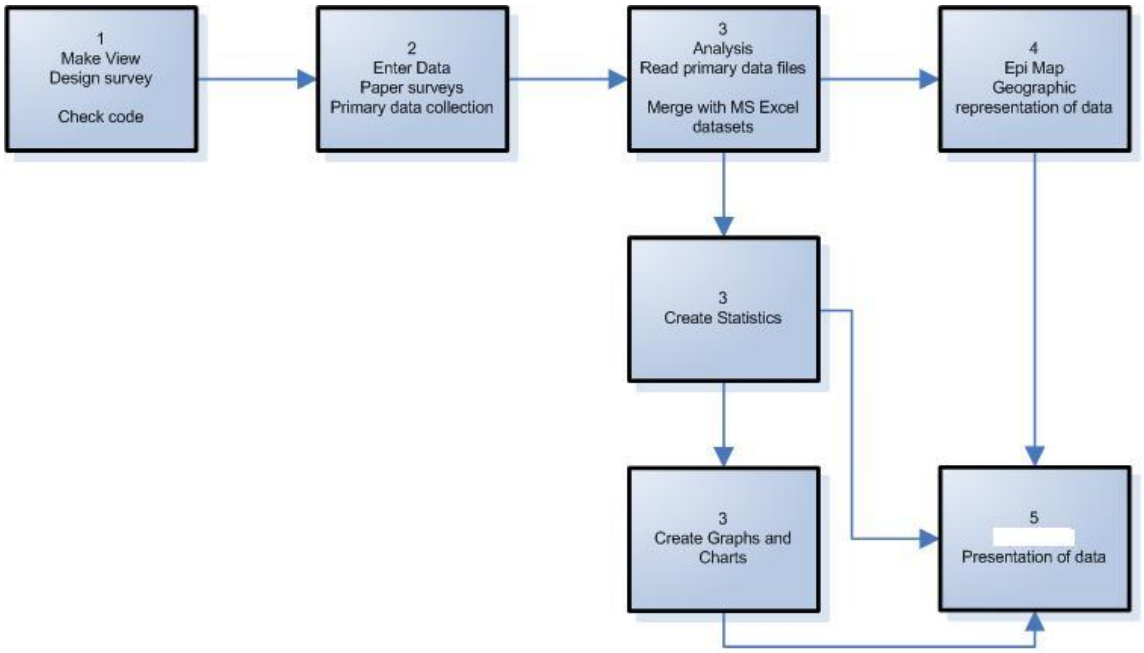
File Download List

Time to lace up those downloading shoes!

You need to download the following files to work through all the lessons in the tutorial. Files are available on the Epi Info™ website on the Tutorials page.

Tutorial Guide You need one of the following:	EIHAVer#.DOC (MS Word version) EIHAVer#.PDF (Adobe Reader version) EIHAVer#.HTM (HTML Browser version)
Project Files	Albany_demogz.MDB Asthma Survey 2005.MDB Asthma Final.MDB Asthma Survey 2005 Sample.MDB HospitalPoints.MDB SchoolPoints.MDB SymptomsByAgeSample.mdb
MS Excel Files	AshtmaTable.XLS
HTML Files	CensusAge.HTM Frequencies.HTM Frequencies.XML
Epi Report Files	My Epi Report2.EPT My Epi Report2.EPX WMonth Report2.EPT WMonth Report2.EPX
Graphics	WMonthPieGraph2.BMP
Map Files	Albany County boundary_region.DBF Albany County boundary_region.PRJ Albany County boundary_region.SHP Albany County boundary_region.SHX ALBZCTA_region.DBF ALBZCTA_region.PRJ ALBZCTA_region.SHP ALBZCTA_region.SHX

Tutorial Process Flow



Introduction to the Scenario

Now that you're ready for battle, time to be briefed on your mission.

This tutorial is designed to teach Epi Info™ by following a scenario. The tutorial includes a survey adapted from one used in New York State; however, the survey data used for the analyses were created for the tutorial. Anonymous data relating to New York State were taken from different datasets referenced in Appendix I: Data Sources. The survey and datasets presented here relate to community health assessments, and specifically asthma; however, they can be used to teach the Epi Info™ program to any user, and to illustrate how the program can be used to gather, analyze, and present data.

Scenario

The local health department (LHD) director of County X gets a Request for Proposal (RFP) from the state health department saying there are funds available for an asthma initiative. There is a good chance of getting the funding as the LHD has already done some work with asthma. To apply for the funding, the director meets with the health education coordinator/planner to identify data required to write a strong proposal. The director and the planner determine that there are several types of useful data to present as part of their proposal.

To support the scenario, data have been collected from a variety of current sources. Provided datasets include aggregate data, map, shape, census, and primary data. In this scenario example, you will be using a primary data source (paper survey) to collect information for the sample population of students from County X. You will use MakeView to create the electronic survey, Enter to add records to your survey, Analysis to import and analyze results, and Epi Map to plot the results. The following To-Do list shows the basic steps you will accomplish to develop the data needed to apply for funds for the asthma initiative.

To Do List for Lessons 1–9

Here's the game plan.

Use MakeView to Create the Electronic Survey

- ☐ Create your project .MDB - Lesson 1
- ☐ Create an electronic version of the survey - Lesson 1

Use the MakeView Program Editor to Develop Check Code

- ☐ Create Check Code to simplify the data entry process - Lesson 2

Use Enter Data to Add Student Records

- ☐ Enter records into your survey and data table - Lesson 3

Use Analysis to Merge Data and Create Statistics

- ☐ Open your project .MDB in Analysis and view the entered records/data - Lesson 4
- ☐ Merge your project .MDB with records from an MS Excel file - Lesson 5
- ☐ Create frequencies, means, and tables using your data and record the results - Lesson 5
- ☐ Create graphs using your data and record the results - Lesson 6

Use Analysis to Manage and Export Data

- ☐ Use the RECODE command to view date information from the survey in a specified order - Lesson 7
- ☐ Export selected data to a new table and create a line graph using the selected records - Lesson 7
- ☐ Use the SUMMARIZE and FREQUENCY commands to create aggregate data sets of survey information that can be mapped - Lesson 8

Use Epi Map to Develop Maps to Illustrate the Data

- ☐ Create a choropleth and case-based map combination in Epi Map to look for relationships between income and asthma - Lesson 9
- ☐ Create point maps for the locations of regional hospitals and schools - Lesson 9
- ☐ Create a map to plot the count of students with bronchitis per zip code - Lesson 9

This page intentionally blank.

MakeView: Create a Survey

SUMMARY

This lesson introduces the MakeView module of Epi Info™. During this lesson, you will create your project .MDB and an electronic version of a survey used to collect data from students throughout the county.

In order to develop data that can be used to show the need for funds, you must first develop the survey to use as your primary data collection instrument. In this example, the survey has been developed for you attached as Appendix B: Parent School Asthma Pre-Intervention Survey. Lesson 1 guides you through the process of using MakeView to create the survey. At the end of this lesson, you will have experience using the Field Definition dialog box to make many of the fields needed to collect data from the survey population.

Time to complete: 1.5 hours: Beginner

Getting Started with MakeView

Overview of basic information to navigate MakeView and understand project creation. Describes how to use the Field Definition dialog box.

Create Fields, Part 1

You will use the Field Definition dialog box to create text, date, numeric, and Legal Value fields. Learn how to set the tab order and name a page inside your project. Includes a practice lesson.

Formatting Options

Learn how to format your workspace by changing grid settings or the background color. Learn how to format/edit fields in your view by resizing or aligning fields.

Create Fields, Part 2

Learn how to create a checkbox, group, mirror, and yes/no field. Includes a practice lesson.

Skills Review

Hands-on series of tasks with fewer instructions and a greater degree of difficulty. Learn how to create a comment legal field.

BEFORE YOU BEGIN

Install the latest version of Epi Info™ on your computer.

Complete the File Set-Up section of the tutorial.

WHAT YOU NEED

Appendix A: Code Sheet

Appendix B: Parent School Asthma Pre-Intervention Survey

FIVE GOALS

- Create a project called AsthmaSurvey, and a view called PreInterventionSurvey.
- Create three pages in a view to hold survey questions.
- Create a Required StudentID field/variable to track students.
- Create Date of Birth and Age fields/variables to track student ages.
- Set the Tab Order to verify the data entry order the cursor needs to follow.

Getting Started with MakeView

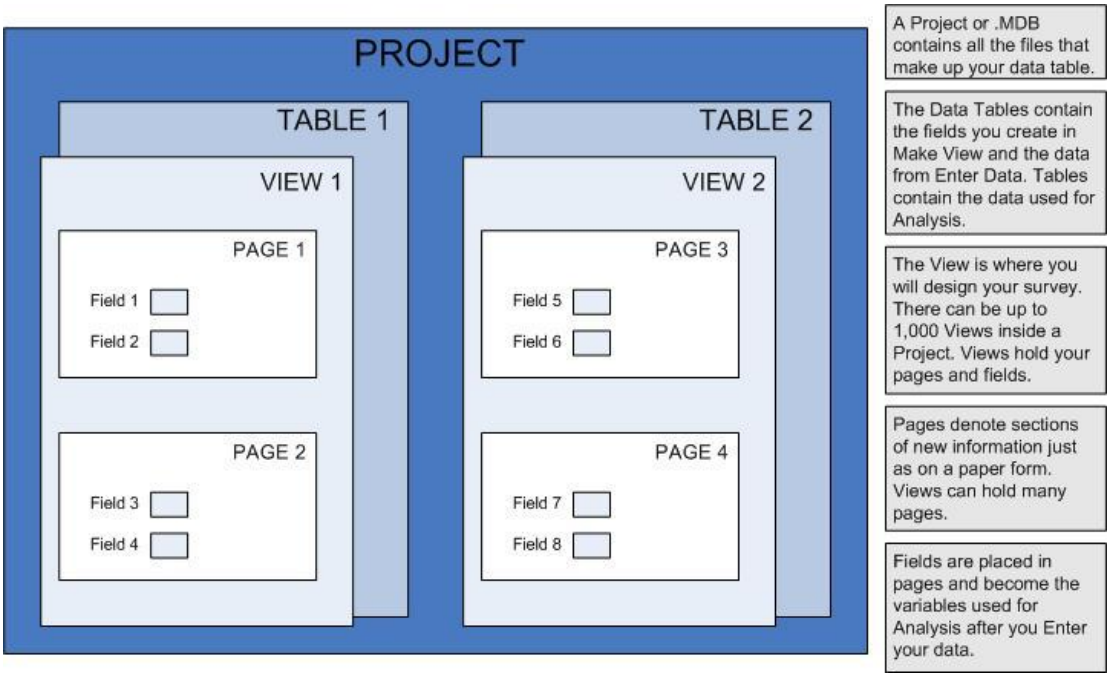
Here we go, time to get the brain engaged!

Epi Info™ uses the Microsoft Access data format to create projects, files, or MDBs. Each project contains a view and a data table. The view contains information about how the survey looks, and how data is entered into the data table. For example, you will create a view based on the Asthma Pre-Intervention Survey. The vew will be a re-creation of the paper survey along with information about field names, field types, and Check Code.

The view and the data table stay inside your Project or MDB. There can be up to 1,000 views inside your project. Each view can contain 250 fields. When you enter data in the Enter module, the data enters into your MDB file. The data table contains the actual values associated with your project or MDB.

The MakeView program is used to place prompts and data entry fields on one or many pages of a view. Since this process also defines the database(s) that are created, MakeView can be regarded as both the form designer and the database design environment.

You will create fields inside each view to hold your data. Fields can be stored on many pages inside a view. You can use the Program Editor component of MakeView to add Check Code to your fields. Check Code represents a series of rules applied to your data. One example, used in the tutorial, concerns the calculation of date fields. You can create Check Code stating that when a date of birth is entered in one field, the age of an individual is automatically calculated and placed in another field. This allows you to customize your questionnaire and makes data entry quicker and more accurate.



Database Organization Chart

The MakeView Workspace

Pretend you're a cat and grab the mouse!

1. From your desktop, double click the **Epi Info™** icon  to open the program.
2. Click **MakeView** or select **Programs>MakeView (Questionnaire)** to open the MakeView module. The Make/Edit View window opens.
3. As you read this Overview, you can use your mouse to view available tasks from the drop-down menus.
    - Please note: Menus will not be active until you create a new project in Lesson 1.

This is an example of the MakeView workspace. The workspace settings are set to default; meaning the background color is gray and the grid will be visible when you open a new file.

This is the Page Names panel. Views can contain many pages.

From the Page Names panel, you can also access the Program Editor function of MakeView and set up your Check Codes.

This is an example of the Make/Edit View window. This section is called your View. Once you create or open a file, you can begin adding information to your View.

You will right click inside the View to bring up the Field Definition dialog box and begin working.

Navigating MakeView

Here's a brief summary on what you can do with the menus.

File Menu

Use to create a new project, open an existing project, save and copy a view, navigate to Enter Data, or exit the MakeView program.

Edit Menu

Use to cut, copy, paste, move grid, delete a page, rename a page, select a page, or change the tab order.

The fields created in your view are the same as those in the paper Asthma Survey. When the data entered in the fields is imported into Analysis, you will sort it by field names, also called Variables.

View Menu

Use to navigate to the check commands screen where you can edit your Check Code.

Insert Menu

Use to insert a new field, a new page, or create a group.

Format Menu

Use to set a default font, align fields, change your background color, or edit your settings. Setting a default font affects all text created by the Field Definition dialog box after the selection. It will not affect fields created prior to selection. Settings made to the default font can be altered by the Font for Prompt button in the Field Definition dialog box.

Tools Menu

Use to import Epi 6 REC files, import Check Code, make a view from a data table, create a data table, or delete a data table.

MakeView Help Menu

Contains the help file contents, how-to files, a command reference, and the About MakeView file.

Creating a Project

Time to get those creative juices flowing!

Now you will create the asthma survey/questionnaire. The information entered here will populate your database and be used for analyses in a later lesson. Creating the survey requires some forethought into the types of analyses you may want to run on your data (e.g., you may want to see the Means of Age later. To do this, you need an age variable in your survey, and the data must be entered into a Number field type. Think about what you may want to know later and create your survey accordingly.

You will collect primary data to illustrate funding needs for an asthma initiative. To start, you need to re-create the Asthma Pre-Intervention Survey in an electronic format.

Refer to *Appendix A: Code Sheet and Appendix B: Parent School Asthma Pre-Intervention Survey* as you create the survey.

You need to create a new project and a new view. Once inside the project, you can add pages and create views to address all the survey questions. Later, you will enter records into the survey.

Be sure to place all your tutorial files in the EIHA Tutorial folder you created in File Set-Up.

Now, follow these steps:

1. From the MakeView navigation menu, select **File>New.** The Create or Open Project window opens.

2. Using the Look In drop menu, select your **C** drive.

3. Double click to open the **Epi_Info** folder.

4. Double click to open the **EIHA Tutorial** folder you created in the File Set-Up section.

5. Type **AsthmaSurvey** in the File Name box.

6. Click **Open**. The Name the view dialog box opens.

7. Type **PreInterventionSurvey** in the Name the View box..

   - View names cannot have spaces or begin with numbers.

8. Click **OK**.

   - The name of the window changes to Make/Edit View: PreInterventionSurvey.

Creating Fields on your Form

The following section explains the field and variable types that are available in MakeView.

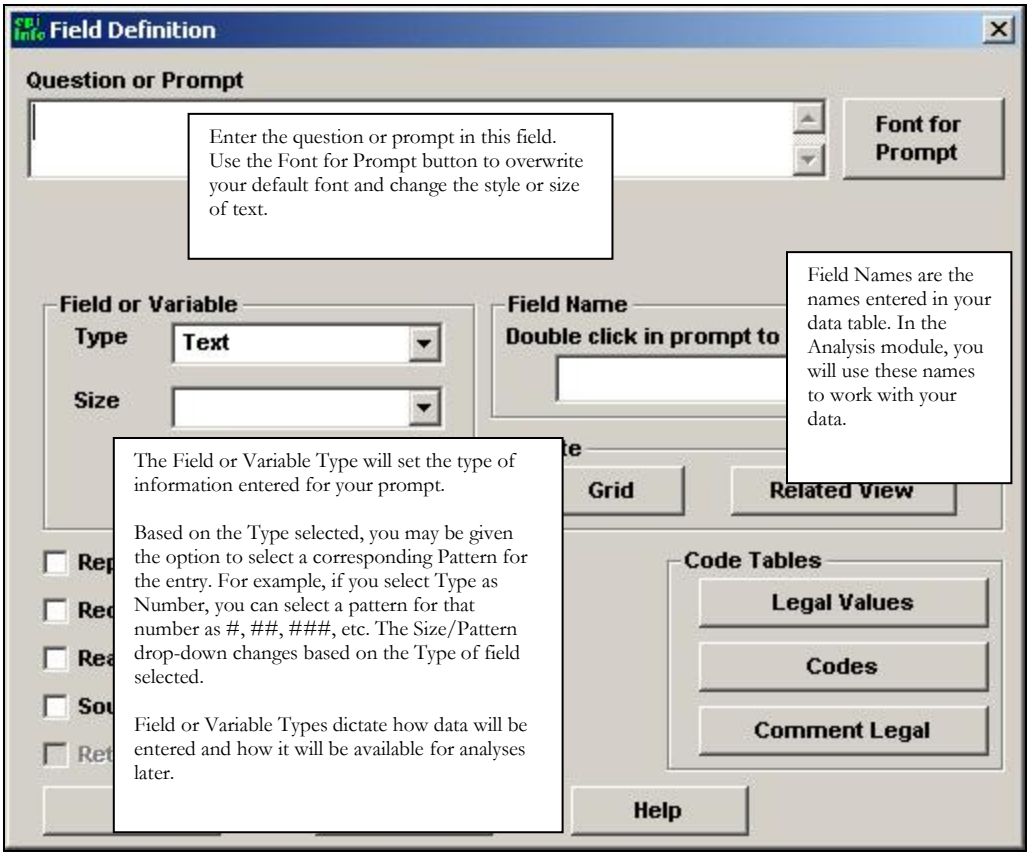
The Field Definition Box

Field or variable types can be created using the Field Definition dialog box. To open the box, you must right click in a view. Each field or variable has its own properties available when selected; however, some options may not be shown or may be disabled (grayed out) depending on the variable or field type selected. Field or Variable Type Properties can also be selected from the box.

When you right click inside the view (in the gray background area), the Field Definition dialog box opens. This box is your main interface with the MakeView module. From here, you will create the fields where data will be entered into your survey and assign names to the variables that make up your data table.

You can use the Field Definition dialog box for:

- Entering a question or prompt. Use the tab key or your mouse to move from field to field.
- Entering a Field Name. Field Names created in MakeView are the Variable names used in Analysis. No two Field Names can be the same.
- Selecting properties for your field (e.g., Read Only creates a field that does not allow for data entry).
- Delete an existing field. Right click on the field prompt to open the Field Definition dialog box and click **Delete**. **To delete Legal Values and tables, open the associated table from the Field Definition dialog box, and click Delete. There is no Undo button or verification prompt, so be sure of your selection before choosing to delete variables or code tables.**



Field or Variable Types

From the Field or Variable type drop-down, you can create the following types of variables. Click **Cancel** to close the Field Definition dialog box after you become familiar with the available options.

Text

Creates a text data entry field. Letters or numbers can be entered in text fields. Text fields can be sized for specific entries, contain up to 128 characters, and be used to create name or address data entry fields.

Label/Title

Creates a title for the page or a label for a question. This field type is not searchable, not listed in the tab order, and not visible for Check Code. Can be used to create a title in the view.

Text (Uppercase)

Creates a text data entry field that can contain letters or numbers, but forces all text entries to appear in uppercase only. Can be used to create a data entry field for state abbreviations or identification fields that must use all caps.

Multiline

Creates a data entry field that can contain letters or numbers up to two gigabytes of information. Can be used as a comments field or to enter long addresses.

Number

Creates a number data entry field based on a selected pattern, (i.e., #, or ## ##.##). Only numbers can be entered in the field. Can be used to enter identification numbers or ages.

PhoneNumber

Creates a phone number data entry field based on phone number patterns. Select from one of two patterns, ###-#### or ###-###-####.

Date

Creates a date entry field based on a selection of patterns (i.e.,MM-DD-YYYY or DD-MM-YYYY). Dates entered must match the selected pattern. Can be used to enter today's date or a date of birth.

Time

Creates a time entry field based on a selection of patterns (i.e., HH:MM:SS). Times entered must match the selected pattern.

Date/Time

Creates a data entry field that contains both a date and time based on a selected pattern.

Checkbox

Creates a checkbox response. The Checkbox field is treated like a Yes/No field. During usage, non-selected boxes do not appear as Missing in Analysis. Can be used with Check All That Apply type questions.

Yes/No

Creates a Yes or No data entry field. Field will accept the typed word yes or no. Yes/No fields left blank will appear in Analysis with missing values for those questions. The yes or no answer is stored in the database as a 1 or 0. When creating Check Code, use the (+) = Yes or (–) = No to register a yes or no response.

Option

Creates a radio button selection list. Use with Select One Answer Only questions. The Option field is for mutually exclusive choices; only one choice can be made. If more than one choice is necessary, then use the Checkbox option.

Command Button

Creates links to other components or files. For example, you can incorporate an executable button in the view to open Analysis or another program (e.g.,Microsoft Excel).

Image

Creates a placeholder to insert images into the view. Use an image field to insert an image per record (i.e., a patient, rash, or bacteria picture). Acceptable image file types are Graphics Interchange Format (.GIF), Joint Photographic Expert Group (.JPG or .JPEG), Windows Bitmap Format (.BMP), Windows Icon File Format (.ICO), Windows Metafile Format (.WMF), and Enhanced Metafile Format (.EMF).

Mirror

Creates a field that appears on many pages. Mirror field types only work with multiple pages in a single view. For example, if a Patient ID is on page one, the value of Patient ID can be displayed onto another page using the mirror field. Mirror fields are read only.

Grid

Creates a related table/grid that can be inserted into a questionnaire or view to track specific information (i.e., a child's measurements).

Relate

Creates relationships between your main view ("parent" view) with sub views ("child" views) only within the same .MDB. For example, link the general patient information to the patient's record of visits.

Creating Fields Part 1

Many types of entry fields can be created in MakeView. For this lesson, you will create fields for Title, Text, Number, Date, and Legal Value.

Creating a Label/Title Field

Time to make your ascension into the label/title territory.

1. To create a field, right click in the **MakeView** workspace. The Field Definition dialog box opens.

2. In the Question or Prompt field, type the title of your survey: **Parent School Asthma Pre-Intervention Survey.**

   - The Question or Prompt box allows you to create a prompt/question that holds 128 characters.

3. From the Field or Variable Type drop-down menu, select **Label/Title**.

4. In the Field Name field, highlight the **default field name**, and type **PreSurvey**.

   - The Field Name will take part of the Question or Prompt and automatically create its own name. Because the Field Name cannot begin with numbers, if a number is the first character in the Question or Prompt, a letter will precede the rest of the Field Name.

   - Notice that your Field Name is the same as your title. Field Names are created automatically by Epi Info™ based on the data entered in the Question/Prompt field.

   - Field Names are used by multiple modules in Epi Info™. Try to create names that will be easy to reference.

   - Field Names cannot have spaces in them.

   - Some applications that work with Epi Info™ cannot use field names longer than eight characters. Keep this in mind if you plan to use your data with other statistical programs.

   - Field Names cannot be changed after your data table has been created.

5. Click **Font for Prompt**. The Font dialog box opens

   - The Font for Prompt button is used to change the font setting for the question or prompt for the field.

6. Select a font from the list.

7. Make the font size **14, bold**.

8. Click **OK**. Your title text has been formatted with the new font type

   - A Label/Title field cannot be used to create Check Code, cannot be searched, nor is it included in the tab order list.

9. Click **OK**. The title of your survey appears on the page.

   - To move your title to the top center of the page, use the click and drag technique. Select the text with your mouse, left click and hold while moving the text to the desired location.

Required Tasks

You will now create another Label/Title field to provide some instructions to the user regarding completing the survey. Follow the same process to create another Label/Title field using the instructions below:

Prompt/Question	Variable Type	Field Name	Other
Please answer all the questions below, whether or not your child has asthma.	Label/Title	Instruction	Make the font size 12. Move your instructions field underneath the title.

Creating a Text Field

1. To create a text field, right click in the **MakeView** workspace. The Field Definition dialog box opens.
2. In the Question or Prompt field, type **School ID Number**
3. From the Field or Variable Type drop-down menu, select **Text.**
4. Give your field the field name **SchoolNum.**

Follow the same process above to create text-type fields for the following:

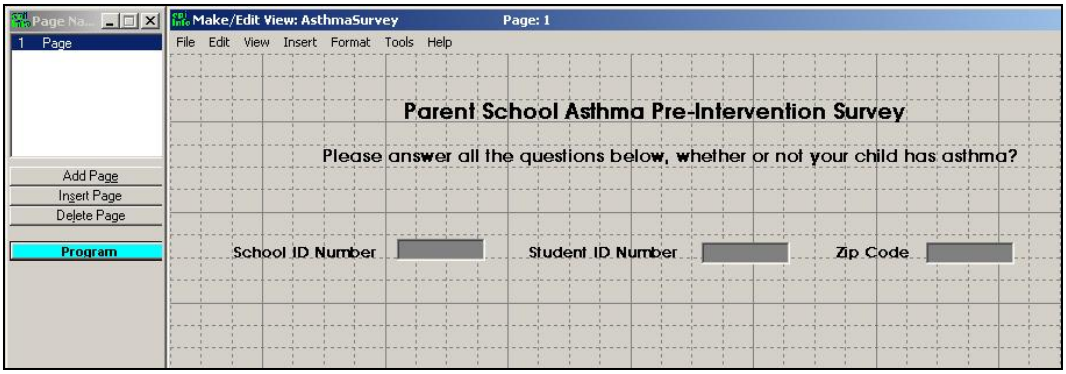
Prompt/Question	Variable Type	Field Name	Other
Student ID Number	Text	StudentID	Select Required option.
Zip Code	Text	Zip	

Notice that you are making fields generally considered numeric into text fields. This ensures that these numbers will always appear as entered. This topic will be discussed further in Lesson 3.

To restrict the number of characters that can be entered in a text field and thereby control the data entry process further, you could select from the size drop-down located in the Field Definition dialog box and limit the number of characters allowed during data entry. When

thinking about form design, you must consider potential errors from typing into all fields. Creating Legal Value drop-down boxes (covered later) are another way of controlling data entry and limiting potential typing errors.

Page 1 of your view should look similar to this example:



Creating a Legal Value Field

You won't need a lawyer to do this!

When creating a Legal Value or Comment Legal field, you create a drop-down menu within your view. Selections from the drop-down menu will only be accepted during data entry. When creating a Legal Value field, the variable type must be Text.

The dataset contains a question about gender. Now get ready to create a Legal Value field.

1. Open the **Field Definition** box.
2. In the Question/Prompt field, type **Gender.**
3. From the Variable Type drop-down, select **Text**.
4. Use the default Field Name, **Gender**.
5. From inside the Field Definition dialog box, locate the Code Tables section.

6. Click **Legal Values**. The Set Up Code/Legal Links dialog box opens.

7. Click **Create New**. An entry box opens for your selected value.

8. In the first line, type **Female** and press **Enter.**

9. **T**ype **Male** and press **Enter**.

10. Click the **Do Not Sort** checkbox.

    - MakeView sorts all Legal Values alphabetically, unless you select Do Not Sort.

11. Click **OK** to accept the Legal Values.

12. Click **OK** to accept definition entries.

    - Notice your survey now contains a drop-down box with Female and Male values as selections.

Creating a Date Field

The Date field is a numeric field with pre-set date patterns selected from the pattern drop-down menu.

Let's play the dating game!

1. To create a date field, right click in the **MakeView** workspace. The Field Definition dialog box opens.

2. In the Question/Prompt field, type **Child's Date of Birth**.

3. From the Variable Type drop-down, select **Date**.

4. From the Pattern drop-down, select **MM/DD/YYYY**.

5. In the Field Name section, type **DOB**.

6. Select the **Required** option.

7. Click **OK**.

Creating a Numeric Field

The Number field is a numeric value field that has six predefined value patterns. You can create a new pattern by typing the pattern into the Pattern field.

Similar to Date Fields, selecting a Field type of number gives you many Pattern selection options. You need several types of number fields to complete the survey.

This is how we'll play the numbers game!

1. Open the **Field Definition** dialog box.
2. In the Question/Prompt field, type **Age.**
3. From the Variable Type drop-down, select **Number**.
4. From the Pattern drop-down, select **##**.
5. Do not change the default Field Name, Age.
6. Select the **Read Only** option.
7. Click **OK**.
    - In Lesson 2, you will create a Check Code to calculate this field based on the child's date of birth.

Required Tasks

Use the table below to create the Phone Number and PS School Number fields.

Hint: Phone Number is a new variable type.

Prompt/Question	Variable Type	Field Name	Other
Home Phone Number	PhoneNumber	Phone	Select the Pattern ###-###-####
PS School Number	Text	PSNumber	Select Read Only option.

Setting the Tab Order

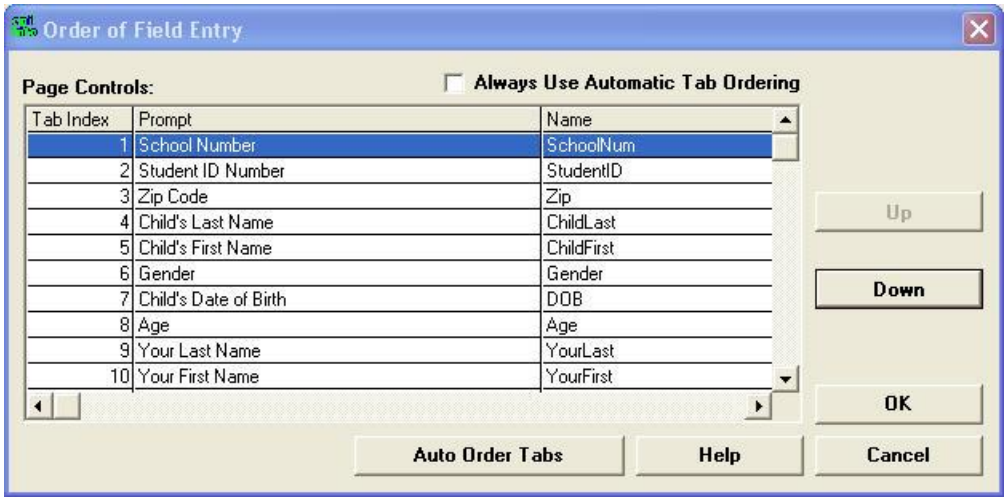
The tab order is the order in which a user moves focus from one control to another by pressing the TAB or ENTER key. Each form has its own tab order. By default, the tab order is set automatically in your view based on each field' s positioning in order from left to right, then top to bottom. The Always Use Automatic Tab Ordering is the default setting. When its checkbox is checked, the tab order for a field is set automatically when the field is first defined and whenever the field is repositioned on the view. Automatic Tab ordering is useful for rapid questionnaire development.

However, you can manually change the order of tabs in your survey. When doing so, the Always Use Automatic Tab Ordering checkbox will be unchecked automatically. You can click the Auto Order Tab button to automatically reset the tab order for all fields on the view if any changes were made manually. Verifying the tab order is a good practice because your cursor may jump to fields out of sequence with data entry and cause errors if not verified.

Make sure that the order of fields matches the way information will be entered from the paper survey. Refer to *Appendix B* to review the field order of the paper survey.

Let's set the tab:

1. Click **Edit>Order of Field Entry (Tab Order)**. The Order of Field Entry dialog box opens.



2. To determine the field order you want, select a prompt and use the up and down buttons.
    - For example, the variable Age should come after the variable DOB.
3. Click **OK** to accept the new tab order.
4. Click **File>Save**.

Naming the page

If your survey contains more than one page, use the Page Names window to navigate pages. The Page View Menu Panel displays the created views' names and provides access to created pages. You can add, insert, and delete pages from this menu.

The MakeView Program Editor can be accessed from this menu panel by clicking the Program button. If you have a related view in your project, the Home and Back buttons are available to navigate between parent and child views.

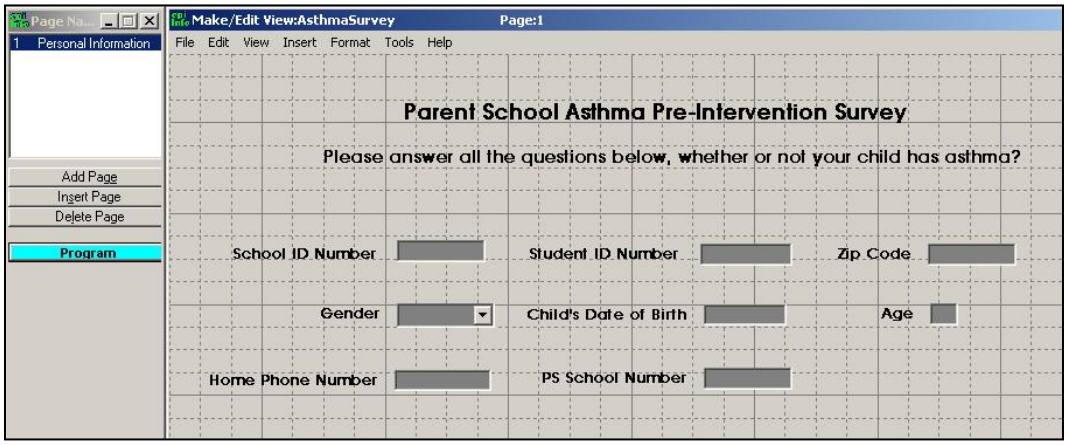
Now begin naming the pages.

1. From the Page Names panel, right click on the page named **1 Page**. The Pagename window opens.

2. In the Page Name Field, type **Personal Information**.

3. Click **OK**. The page name appears in the list.

   Page 1 of your view should look similar to this example:



Creating a Code Table

The Codes field allow you to choose a value from a drop-down list. Based on that choice, another field(s) is populated with predetermined values. At least two fields must exist; one holds the selection code and the other receives the value of the code. The first field holds the selection code in a drop-down list, and the second (or third, etc.) is a Read Only field that populates based on assignments set in the code table.

Four schools responded to the survey. To ensure accuracy and speed of data entry, you will create a code table that links the information entered in the SchoolNum field to the information in the PSNumber field. This means that each time a School ID Number is selected from the drop-down, the field PS School Number will be automatically populated based on the data table values you create. This helps to ensure consistency and eliminate errors during data entry.

Now that you know how to be consistent and error-free, let's make the code table a reality.

1. Right click on the prompt of the **School ID Number** to open the Field Definition dialog box.

   - PSNumber and SchoolNum must be Text type fields.
   - PSNumber must be set to Read Only.

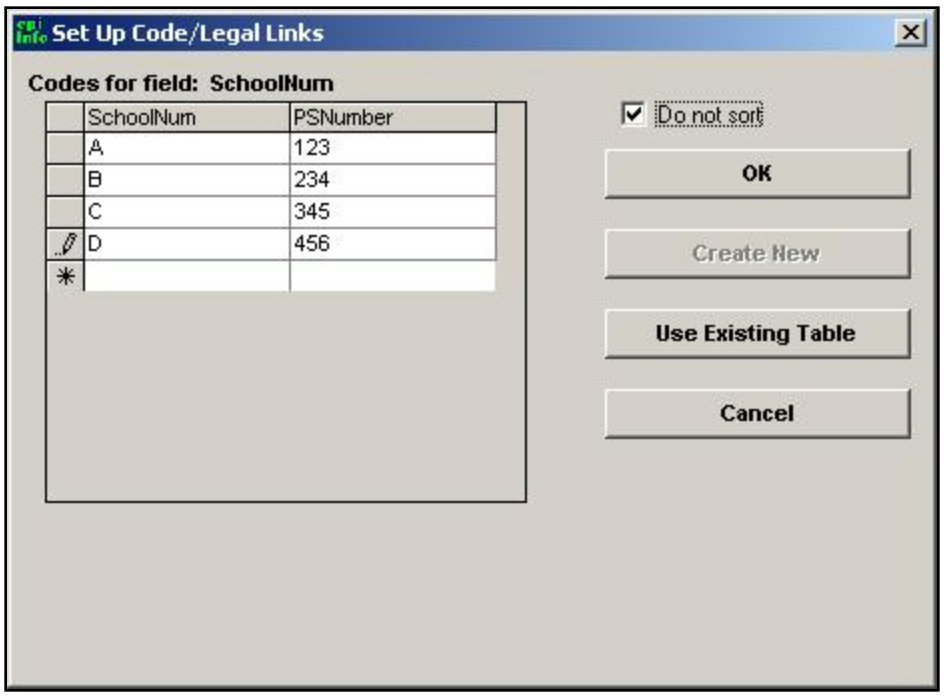
2. From the Code Tables section, select **Codes**. The Set Up Codes/Legal Links dialog box opens.

3. Select **PSNumber** as the field you want to link to SchoolNum.

4. Click **Create New**. The Codes for Field box opens.

   - The schools are named A, B, C, and D. Based on the School ID Number selected, you want the corresponding PS School Number to populate automatically.

5. On the first line enter **A**. Tab over and type **123** as the PSNumber.



6. Use the tab key to navigate through the table set-up. Complete the series as follows:

   - SchoolNum B and PSNumber 234
   - SchoolNum C and PSNumber 345
   - SchoolNum D and PSNumber 456

7. Select **Do Not Sort**.

8. Click **OK** in the Legal Links dialog box.

9. Click **OK** in the Field Definition dialog box.

   - The School ID Number field now contains a drop-down menu. Once you select a School ID Number, the PS School Number will automatically populate with the value you entered in the code table.

Formatting Options

This section outlines ways to customize your workspace and to edit fields in your view. You can change the background color or alter the grid settings. Color changes will appear in the Enter module.

Changing the Background Color

Now you get to choose your color.

You can alter the background color of your MakeView workspace and apply the changes to one or all pages in your project.

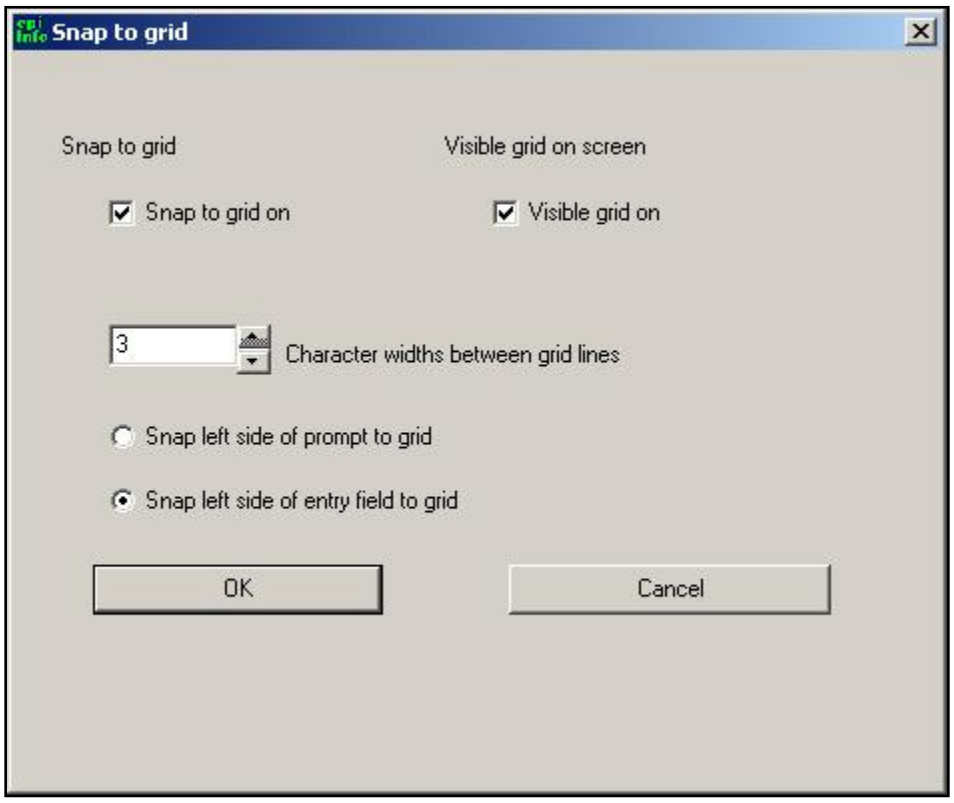
1. Select **Format>Background**. The Background dialog box opens.
2. Click **Change Color**. The Color selection window opens.
3. Select a new background color for your survey.
4. Click **OK** to accept the selected color.
5. Select **Apply to Current Page Only** or **Apply to All Pages**.
6. Click **OK** to view your selection.
7. Click **Clear Color** from the Background dialog box to return to the default gray.

Changing the Grid Settings

Time to change the grid.

Grid settings provide a way to align the fields in the view. In addition, they can be customized in several ways.

From the MakeView navigation bar, click **Format>Settings**. The Snap to Grid dialog box opens.



The default setting of Snap to Grid is on. When moving a field on the view screen, the field will snap to the left portion of the grid it's closest to. You have the option to snap based on the prompt or the entry field. Select the Snap to Grid option if you want your fields to snap in line with your grid when they are created or moved. Deselecting the Snap to Grid checkbox allows you to move fields to any location on the page.

You can select the Visible Grid option if you want your working space to contain a grid while designing the view. The grid is not visible in the Enter module. Visible Grid on Screen allows you to remove the visible grid from the MakeView screen. If the box is checked, you have the option to change the width between the grid lines.

You can change the size of the grid spacing by using the up and down arrow keys next to Character widths between grid lines. Changing the widths of the grid lines provides more control over the spacing of fields.

Resizing Fields

Let's see how you can resize your entry fields.

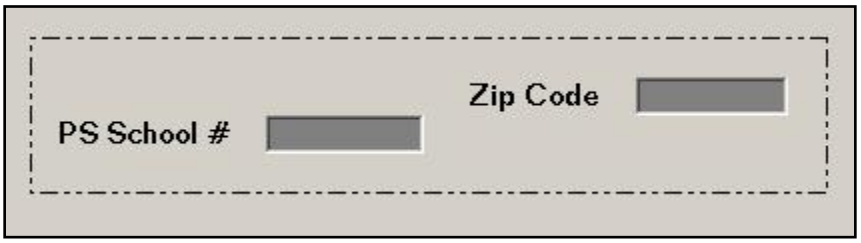
1. Left click an entry field box (not the prompt text).
2. Use the blue bounding box and the click and drag method to make the data entry boxes larger or smaller.
3. The numbers inside the fields tell you how many characters are visible in the survey.
4. To resize a drop-down box, press the **Alt** key and left click to get the blue bounding box and resize using the blue bounding box handles.

Aligning Fields

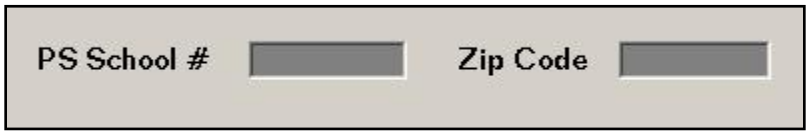
Align away!

Fields can be selected and aligned horizontally or vertically.

1. Left click at the corner the field you want to align.
2. Hold your mouse button and drag a **rectangle** around several fields.



3. Select **Format>Alignment>Vertical** or **Format>Alignment>Horizontal**. The selected fields will line up based on your selection.



Copying and Pasting Fields

Fields can be cut, copied, and pasted.

Time to get out the sissors, copy machine, and paste.

1. Left click and drag a **rectangle** around the field you want to copy.
2. Select **Edit>Copy**.
3. Select **Edit>Paste**. The pasted field appears on top of the copied field.
4. Left click and drag the field to move it.
5. Right click the **pasted field** to open the Field Definition dialog box. Since field names must be unique, a number has been added to the Field Name.

Creating Fields Part 2

Now you will create the asthma history section of the survey. The first question on the paper survey is a series of yes or no checkboxes related to asthmatic conditions. You will create one checkbox per condition and then group all the conditions. This allows you to analyze how many students checked yes per condition and to analyze the conditions as a group. A single checkbox will indicate a yes if checked and a no if blank. To create the next section of the survey, you must create separate fields for the question and checkboxes.

Questions have been shortened in the electronic survey. Refer to *Appendix B Parent School Asthma Pre-Intervention Survey* to view the original questions.

You are not going to create the entire survey in Lesson 1; however, you can use the field information in the Appendix A: Code Sheet for additional practice.

Adding a New Page

Get out the adding machine!.

1. From the Page Names panel, click **Add Page**. The new page appears.
2. Right click inside your new page. The Field Definition dialog box opens.
3. Enter the following prompt: **Has a health professional ever told you that your child has any of the following conditions?**
    - For long questions, you may need to enter a line break inside your prompt. To do this, hold the **Ctrl** key and press **Enter** inside the sentence where you want the break to occur.
4. Select **Label/Title** as your Variable Type.
5. Type the Field Name: **Condition**.
6. Click **OK**.

Creating a Checkbox

Time to check this out!

1. Under the question, right click to open the **Field Definition** dialog box.
2. Enter the following prompt: **Asthma**.
3. Select a Variable Type of CheckBox.
4. Use the default Field Name, Asthma.
5. Click **OK**. The Asthma checkbox appears in your view.

NOTE:

- In the Enter Data module, checkboxes appear as selected or not selected.
- In the data table, checkboxes are stored as -1 and 0.
- When used in Analysis Commands, checkbox values are selected and coded as (+) and (-).

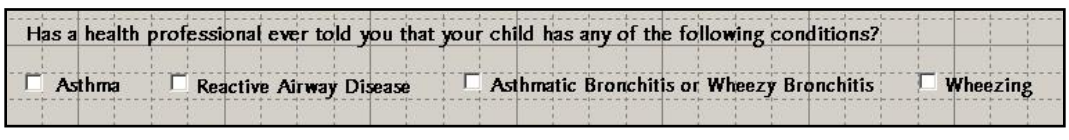
Required Tasks

Use the table below to create three checkboxes to complete the question.

Arrange the checkboxes near each other and align them horizontally or vertically, depending on your workspace and preference.

Prompt/Question	Variable Type	Field Name	Other
Reactive Airway Disease	Checkbox	RAD	
Asthmatic Bronchitis or Wheezy Bronchitis	Checkbox	Bronchitis	
Wheezing	Checkbox	Wheezing	

Your checkboxes should look similar to this example:

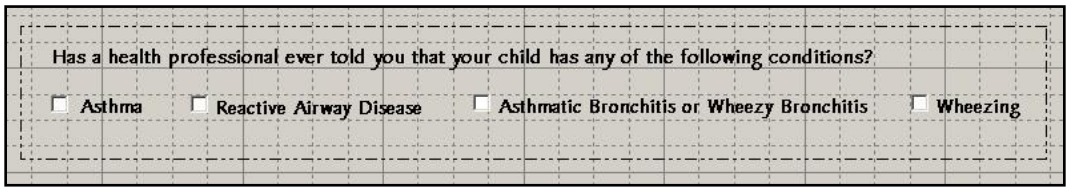


Creating a Group

Groups can be created to organize similar data on a survey. Variables that are part of a group can be analyzed alone or as part of the group. Group the question and the variables regarding conditions.

Time for some group therapy.

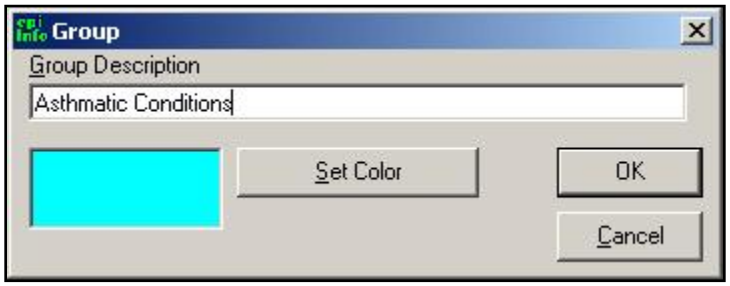
1. Use the left mouse button to drag a **rectangle** around the fields that are to be included in the group. This includes the question and the four variables.



2. From the navigation menu select **Insert>Group**. The Group Dialog box opens.

3. In the Group Description field, type **Asthmatic Conditions**.

   - To select a different color for the group, click **Set Color**.

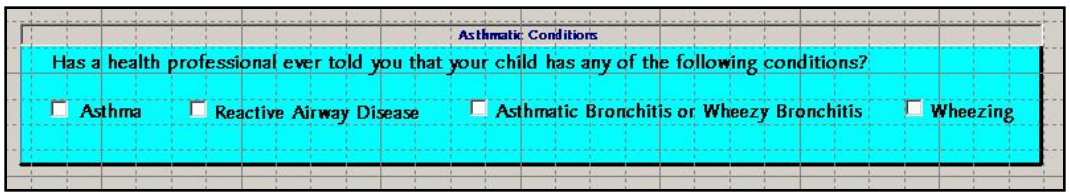


4. Click **OK**.

   - To edit or ungroup, right click **inside the group name**. The Edit Group dialog opens.

   - To resize the box containing the group, double click **inside the box**. A resize arrow appears.

   - To move the entire group, left click the **group title text**, hold, and drag to the new location.



Creating a Mirror Field

Mirror fields allow you to show the value of one field on many pages, but only enter the information once. Since there are several pages of information for each child, you want to place their Student ID on each page.

Look in the mirror!

1. Open the Field Definition dialog box.
2. Enter the Prompt as **ID**.
3. Select the Variable Type: **Mirror**.
4. Use the default Field Name, **ID**.
5. Click **OK**. The Assign Variable to Mirror Field dialog box opens.
6. Select the variable **StudentID** from the list.
7. Click **OK**.
8. Move the field to the top of the page.
    - Once the Student ID # field is entered in the Enter Data module, this field will populate automatically.

Creating a Yes/No Field

It's one or the other!

1. Open the Field Definition dialog box.
2. Enter the following Prompt: **Has a health professional prescribed any medication to your child for breathing problems?**
   - The original question has been modified since it contained more than 128 characters, which is more than allowed in the prompt field.
3. Select the Variable Type: **Yes/No.**
4. Type the Field Name: **Medication.**
5. Click **OK**.

NOTE:

- From the Enter Data module, only the text Yes or No can be entered in this field.
- In the data table, Yes/No values are stored as 1 or 0. Missing values appear blank.
- When used in Analysis Commands, Yes/No values are selected and coded as (+) and (-). Missing values are selected and coded as (.).

Required Tasks

Use the table below to create a new type of drop-down selection box called Comment Legal.

- Comment Legal fields are set up like Legal Value fields; however, they allow you to place a number in front of text, so when you populate fields, a number is entered. This is a text field, but it will display as a number in the Analysis module.

You will also create a group that contains the variable Symptoms, Whistle, and Cough. The group will be named **Monthly Breathing Difficulties**.

Prompt/Question	**Variable Type**	**Field Name**	**Other**
Select the month when your child's breathing was the worst.	Text	WMonth	Click **Comment Legal**. Create a New Table. Enter each month of the year in the following format: 1-January, 2- February 3- March 4- April 5- May 6- June 7- July 8- August 9- September 10 October 11- November 12- December Select **Do Not Sort**.
During their worst breathing month, select how often the following symptoms occurred.	Label/Title	Symptoms	
Wheeze or whistling in the chest?	Text	Whistle	Legal Value The five Values

			are: Never, 2 times each week or less, 3 to 6 times each week, Every day, but not all day, Every day and all day. Select the **Do No Sort** checkbox.
Have a cough?	Text	Cough	Legal value You can use an existing Legal Value table to populate new tables. From the Field Definition dialog box, click **Legal Values**. From the Set Up Code/Legal Links dialog box, click **Use Existing Table**. The Select a View dialog box opens. Select code **Whistle**. Click **OK**. The Select a Field to Link dialog box opens. Select Whistle. Click **OK**. The Legal Values for Cough populate. Select the **Do Not Sort** checkbox. Click

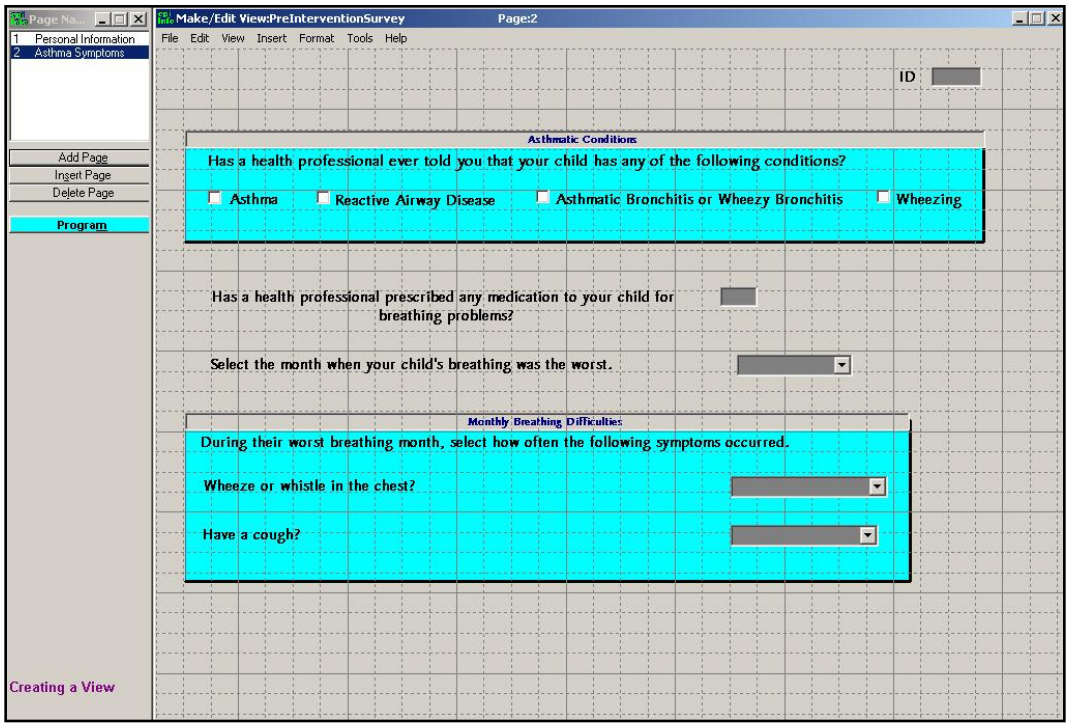
			OK.

1. From the Page Names panel, right click on the page named **2 Page**. The Page Name window opens.

2. In the Page Name Field, type **Asthma Symptoms**.

3. Click **OK**. The page name appears in the list.

Page 2 of your view should look similar to this example:



🕮 Required Tasks

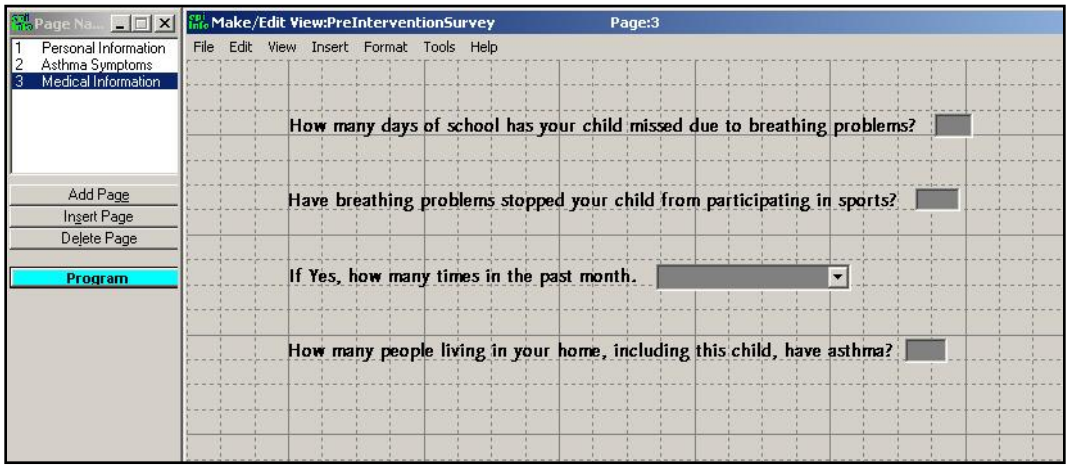
Add a Page to the view.

Use the table below as a reference to create the fields corresponding to this page.

Prompt/Question	Variable Type	Field Name	Other
How many days of school has your child missed because of to breathing problems?	Number	MissDays	Pattern ###
Have breathing problems stopped your child from participating in sports?	Yes/No	MissSport	
If yes, how many times in the past month.	Text	SPTimes	Legal Value 2 times a month or fewer, 3 to 4 times a month, 1 to 3 times each week, 4 or more times each week. Select **Do Not Sort**
How many people living in your home, including this child, have asthma?	Number	IllHome	##

- Verify the tab order for all your pages. You want your cursor to follow a logical data entry path.
- Name Page Three **Medical Information**.

Page 3 of your view should look similar to this example:



NOTE

If you need to close Epi Info™ before completing your survey, the new data table dialog box will appear and prompt you to create a data table. Click **Cancel**. You do not want to create a data table until you are ready to enter data.

- When you exit by going to **File>Exit**, the new data table dialog box opens.
- Click **Cancel**.
- If your data table is set before you are ready to enter data, you can use **Tools>Delete Data Table** to reset your view. This enables you to make edits to your field names and add code tables.

WARNING: This function should only be used if you have not entered any data; otherwise, all your data will be deleted!

Lesson Complete!

WHAT YOU LEARNED

How to:

- Navigate MakeView.
- Create a new project.
- Use the Field Definition dialog box to create fields.
- Add pages to a view.
- Customize a view.

This page intentionally blank.

MakeView: Check Code

SUMMARY

This lesson introduces you to creating Check Code inside the MakeView module of Epi Info™. You will learn how to customize your survey by creating a series of data entry rules. You will also learn the components of the MakeView Program Editor.

In Lesson 1, you created a variety of fields to collect data for your survey. In Lesson 2, you will use the survey you created and add Check Code to some of the fields. Check Code makes the data entry process faster and more efficient; therefore, fewer errors occur when you use the data later to create the statistics and develop your report.

Time to complete: 2 hours Intermediate

Getting Started with Check Code

Overview of basic information to navigate the Program Editor and understand Check Code. A list of available commands is provided. Functions and operators are defined.

Creating a Skip Pattern

Use the Program Editor to create an IF/THEN statement. Also introduces the commands ELSE and GOTO.

ASSIGN Command

Use the Program Editor to create a mathematical function that calculates the Age field based on entries in the Date of Birth field. Use the YEARS function to calculate age.

DIALOG Command

Use the Program Editor to notify data entry personnel of specific data entry rules or errors to avoid.

Skills Review

Series of five review questions based on the reading and hands-on activities in this lesson.

BEFORE YOU BEGIN

Complete Lesson 1 MakeView: Create a Survey

WHAT YOU NEED

Appendix A: Code Sheet

Appendix B: Parent School Asthma Pre-Intervention Survey

Asthma Survey.MDB - created in Lesson 1

Appendix D: Skills Reviews Answer Key

FIVE GOALS

- Learn to open and navigate the three sections of the MakeView Program Editor: the Command Tree, the Check Commands window, and the Program Editor code window.
- Create a skip pattern using an If/Then statement.
- Create code using the ASSIGN command.
- Create code to calculate age using the YEARS function.
- Create a DIALOG box that appears during data entry and prompts the user to verify data.

Getting Started with Check Code

Check Code is used to validate data entry and to enable the user to enter data faster. With some advance planning, you can create code that will perform calculations, skip questions based on answers, prompt the user with dialog boxes, and populate fields across pages and records. In its most basic form, Check Code is a set of rules for the person entering data to follow. It also helps to eliminate errors when entering large amounts of data.

You can open the Program Editor by selecting the blue **Program** button located in the Page Names window or by selecting **View>Check Code** from the MakeView navigation menu.

The Make/Edit View: Check Commands window contains three working areas: the Check Command Tree, the Make/Edit View: Check Commands window, and the Program Editor. The Check Command Tree is not visible until you select a variable from the drop-down menu.

The Program Editor Workspace

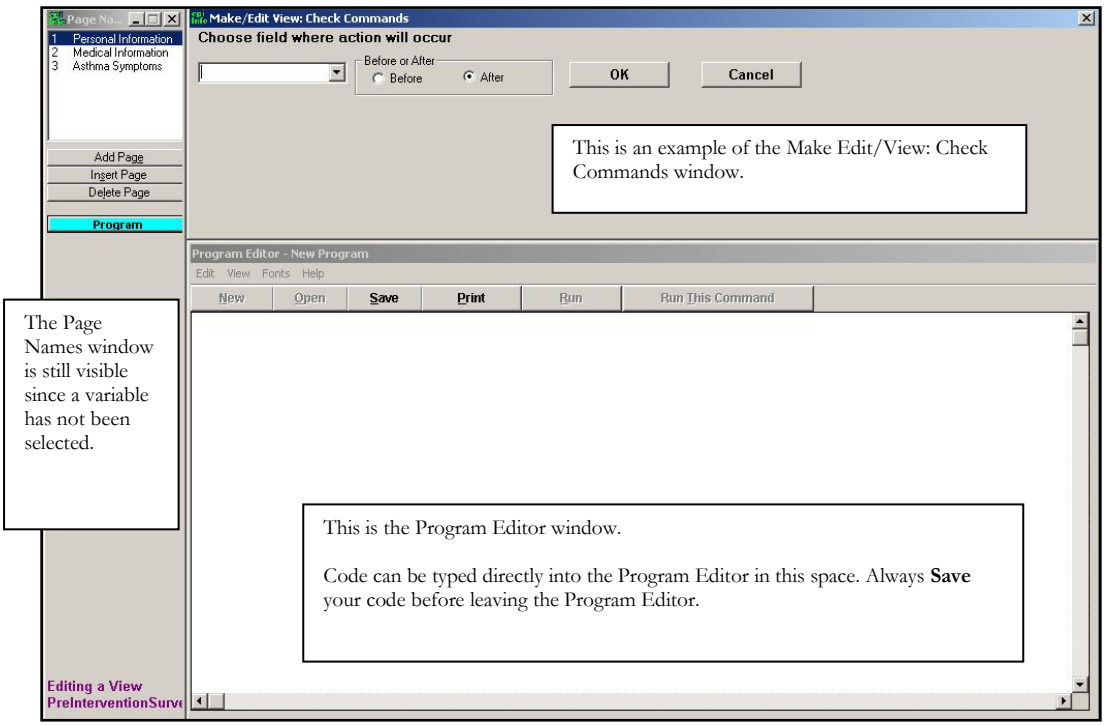
Time to make space for some work.

To access the Program Editor and create Check Code, you must open the survey you created in MakeView Lesson 1.

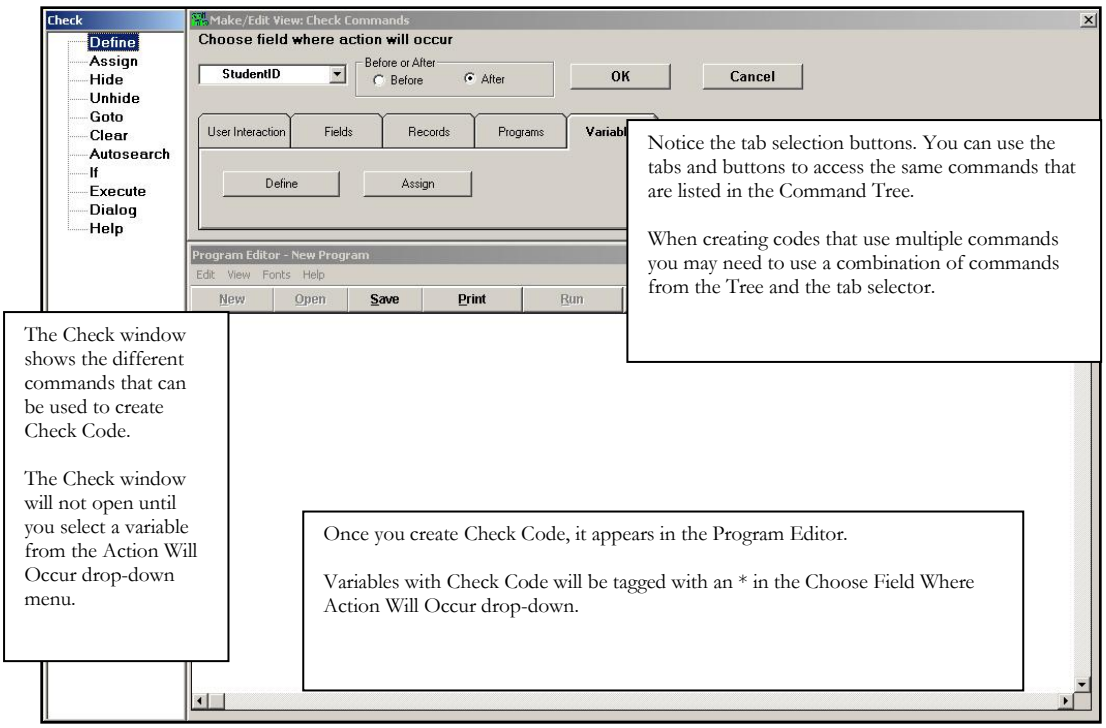
1. Open the Program Editor by clicking the blue **Program** button from the Page Names window or selecting **View>Check Code**. The Make/Edit View Check Commands window opens.

    - Notice that the Page Names window is still open since a variable has not been selected.

This is an example of the Program Editor/Check Code workspace. The top of the page shows you are in the Make/Edit View: Check Commands window.



2. From the Choose Field Where Action Will Occur drop-down, select a **variable**. The Make/Edit Check Commands window updates to include the Check Command Tree and the tab selector.



- Check Code must be associated with an existing variable, page, view, or record. The commands are executed when the data entry cursor enters or leaves the selected variable. Variables available for Check Code are organized by page in the drop-down menu.

- The Check window and Command Tree display a list of all available check commands that can be used in the MakeView program. Clicking on one of the commands opens the wizard or dialog box needed to create the selected kind of code.

- The Make/Edit View: Check Commands window allows you to select when, and to which variable the Check Code program you create will be applied. Previously created Check Code can be accessed from this window. Use the drop-down list to view the available variables. Any variable tagged with an asterisk (*) has Check Code associated with it. Select the variable to view the code inside the Program Editor. Since you have not created code yet, no variables will be tagged.

- The Program Editor displays the commands created from the Check Command Tree. If you have experience with coding, you can type directly into the Program Editor and save your code without using the wizards or the Command Tree.

Navigating the Program Editor

Familiarize yourself with the available commands and options present in the Program Editor.

1. Click **Cancel** to close any Command Option boxes you open during the overview.

2. Click **Cancel** from the Check Commands window or the close X to exit the Program Editor prior to starting the next lesson.

Make/Edit View: Check Commands Window

The **Choose Field Where Action Will Occur** drop-down shows a list of fields in the current view. Read Only variables will not execute Check Code.

The Before or After radio buttons indicate when the action will occur or when your Check Code will run in the data entry process. You can create code that runs before you enter data in the selected field or after you enter data when the cursor leaves the field.

The Tab section provides another way to navigate and select commands. You can click a tab to open the section to see which command buttons are associated with it. Commands selected from the Tab section open the same wizards/generators and dialog boxes as the Command Tree.

Check Window

The Command Tree is located inside the Check window. Commands located inside the Check window act as buttons when clicked. The dialog boxes that open when you select a command are also called wizards or generators because they guide you through the Check Code process.

Once you begin creating code, the Command Tree will not be active while inside some of the dialog boxes. When that occurs, you must use the tabs to select commands.

Available Commands

Define

Use to create new variables. Located on the Variables tab.

Assign

Use to assign the result of arithmetic or string expression mathematics to a variable. Located on the Variables tab.

Hide

Use to hide a field from view. Located on the Fields tab.

Unhide

Use to make a field visible and returns it to the status it had before it was hidden. Located on the Fields tab.

GoTo

Use alone or in an IF statement to transfer the cursor to a specifically named field or page. Located on the Fields tab.

Clear

Use to set the field named to the missing value, as though it had been left blank. CLEAR is frequently followed by the GOTO command that puts the cursor in position for further entry after an error. Located on the Fields tab. RETHINK

Autosearch

Use to search for one or more matching records. If a match is found, you can display and edit the matching record(s) or ignore the match and continue to enter the current record. Located on the Records tab.

If

Use to define conditions and/or consequences that result if the conditions are met. An alternative consequence can be given after the ELSE statement. Located on the Records tab.

Execute

Use to execute a Windows or DOS program. Located on the Programs tab.

Dialog

Use to provide interaction with users from within a program. Dialogs can display information, ask for and receive input, and offer lists for making choices. Located on the User Interaction tab.

Help

Used to display an .HTM or .CHM document. Located on the User Interaction tab.

Program Editor Menu

Edit

Use to copy, cut, and paste code. You can also edit your code with Delete Line and Replace Line options. Search your code using Find, Find Next, Program Beginning and Program End options.

View

Use to see additional menu options. Use the Status Bar to see information displayed about your script. Use the Tool Bar to edit your script.

Fonts

Use to change the font type, style, size, and color that appears in the Program Editor.

Help

Use to view the help file contents.

Save

Use to save the Check Code you created in the Program Editor. It also verifies that the code syntax is correct.

Print

Use to make a paper copy of your code.

Functions and Operators

Functions and operators are used inside commands to perform tasks (e.g., extracting a year from a date, combining two numeric values, or testing logical conditions). Check Code can be typed directly into the Program Editor or created using the dialog box generators. Within each dialog box, function and operator buttons guide you to create the proper syntax for your Check Code. You can also click **Functions** from within each dialog box to see a list of all available functions and examples of the proper syntax.

Functions modify the value of one or more variables to produce a result. For example, the function ROUND produces a whole from a variable that has decimal places. Almost all functions require arguments enclosed in parentheses and separated by commas. If arguments are required, do not place spaces between the function name and the left parenthesis.

These are the function buttons available using the Program Editor dialog boxes.

Operators are used to combine two items. For example, the "+" operator combines Var1 and Var2 to produce a sum (i.e., Var3=Var1+Var2).

These are the operator buttons available using the Program Editor dialog boxes.

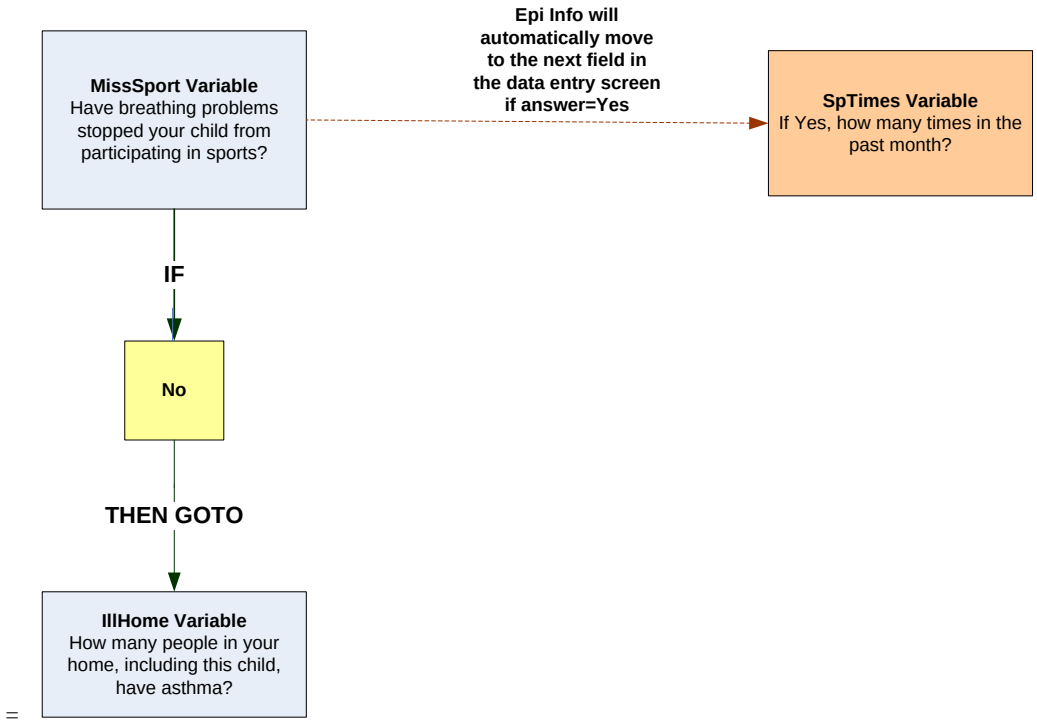
+ Addition	= Equal or Assign
- Subtraction	< Less Than
* Multiplication	> Greater Than
/ Division	<> Less Than or Greater Than (not Equal to)
& String Concatenation	" Enclose dialog or text
(Parenthetical Expression	) Parenthetical Expression

Creating a Skip Pattern

Skip patterns can accelerate the data entry process by moving the cursor to a specified field based on the answer to a question. In this survey, parents were asked if their children have missed sports because of breathing problems. If they have missed sports, then you want them to specify how often. If they have not, you want them to move to the next question.

To create a skip pattern, use the IF command. First, determine what you want the code to accomplish.

If a respondent answers yes, the curser will be placed in the Sp Times Variable Field. If no, the curser will be placed in the IllHome Variable Field..



=

IF/THEN Diagram: Creating a Skip Pattern

To create a skip pattern inside the survey, you will use the Program Editor.

Time to skip along.

1. Click **Program** from the Page Names window. The Make/Edit View: Check Commands window opens.

2. From the Choose Field Where Action Will Occur drop-down, select **MissSport**.

3. You want the action to occur after data has been entered into the MissSport field.

4. From the Records tab or the Check Command Tree, select **If**. The IF dialog box opens.

5. From the Available Variables drop-down, select **MissSport**. The selected variable appears in the If Condition field.

6. From the Operators, click **=**.

7. From the Operators, click **No**.

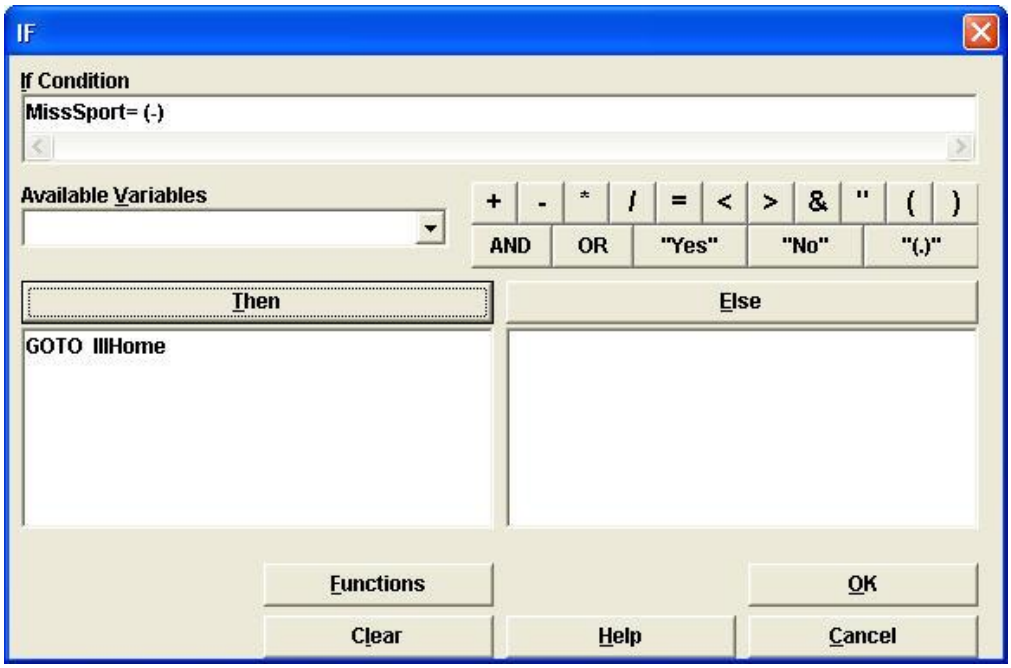
8. The IF Condition field will read MissSport=(-)

9. Click **Then**. The tab selector appears. The Check Command Tree is not active.

10. From the Fields tab, select **GoTo**. The GOTO dialog box opens.

11. Select **IllHome** from the list of variables.

12. In the GOTO dialog box, click **OK**. The IF dialog box opens.



13. Click **OK**. The IF dialog box opens.

14. Click **OK**. The code appears in the Program Editor:

```
IF MissSport= (-) THEN
    GOTO IllHome
```

END

18. Click **Save**.

- Notice in the drop-down that the field MissSport now has an asterisk. That means a code is associated with that field.

Using the ASSIGN Command

To program a mathematical function, you must use the Program Editor. In the survey, you have a field called Date of Birth and one called Age. You will create Check Code that will calculate and enter the age of the respondent based on the date of birth and the date the survey was completed.

Here's your ASSIGNment.

1. From your View, click **Program** or select **View>Check Code**. The Program Editor opens.

2. From the Choose Field Where Action Will Occur drop-down, select **DOB**. The Check Command Tree opens.

   - **After** is the default setting for when all actions will occur. The calculation needs to run after you enter information into the Date of Birth field.

3. From the Variables tab or the Check Command Tree, click **Assign**. The ASSIGN dialog box opens.

   - To complete the code, use the Functions and Operators buttons and type in the **= Expression** field.

4. From the Assign Variable drop-down, select **Age**.

   - The calculated value appears in the Age field.

5. In the = Expression field, type the function **YEARS**.

   - To see a list of all the available functions, click **Functions**.

6. Type, or click, the **left parenthesis**.

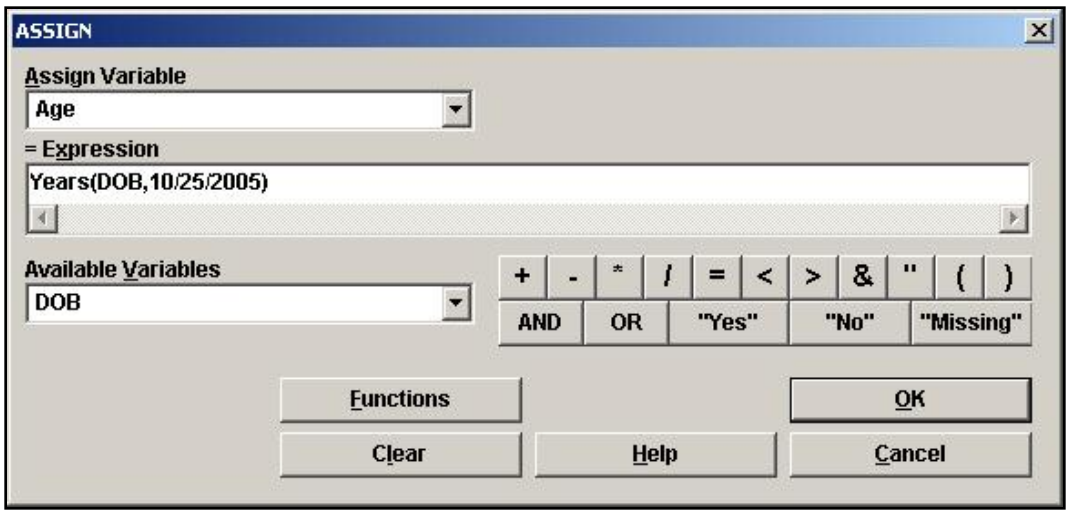
   - Statements of a function must be enclosed in parentheses. Use the Operator buttons or type them in from your keyboard.

7. From the Available Variables drop-down, select **DOB**.

8. Type a **comma**.

9. Type the survey date of **10/25/2005**.

10. Type or click the **right parenthesis**.



11. Click **OK**. The Check Code appears in the Program Editor.

- ASSIGN Age=Years(DOB,10/25/2005)

12. Click **Save** in the Program Editor.

- Always save your Check Code. Codes will not update unless you save. The Save feature will also inform you of any syntax errors that will affect the code functionality.

- Notice that DOB now has an asterisk next to it in the Choose Field Where Action Will Occur drop-down menu. This signifies that the variable has Check Code associated with it.

13. Click **OK** in the Check Commands window to return to the Make/Edit View main page.

Using the DIALOG Command

You will create Check Code to provide a dialog box and instructions for the person entering data into the project survey. The DIALOG command provides interaction with the data entry personnel from within a program. Dialogs can display information, ask for and receive input, and offer lists for making choices.

In the survey, you have a variable called MissDays. This variable refers to the question "How many days of school has your child missed due to asthma?" You want to make sure that any half days are counted as whole days during the data entry process. Using the DIALOG command allows you to set a reminder to run during data entry. This will help ensure you get the most accurate count from your survey data and that all persons entering data will be using the same set of rules.

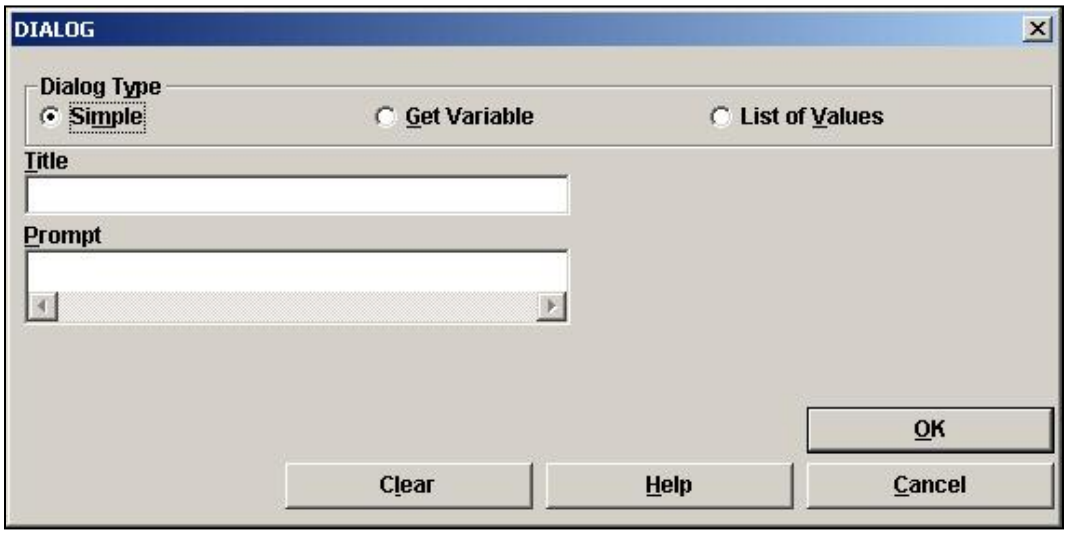
Time to create some dialog.

1. From your view, click **Program** or select **View>Check Code**. The Program Editor opens.

2. From the Choose Field Where Action Will Occur drop-down, select **MissDays**.

   - You want the action to occur after data is entered into this field.

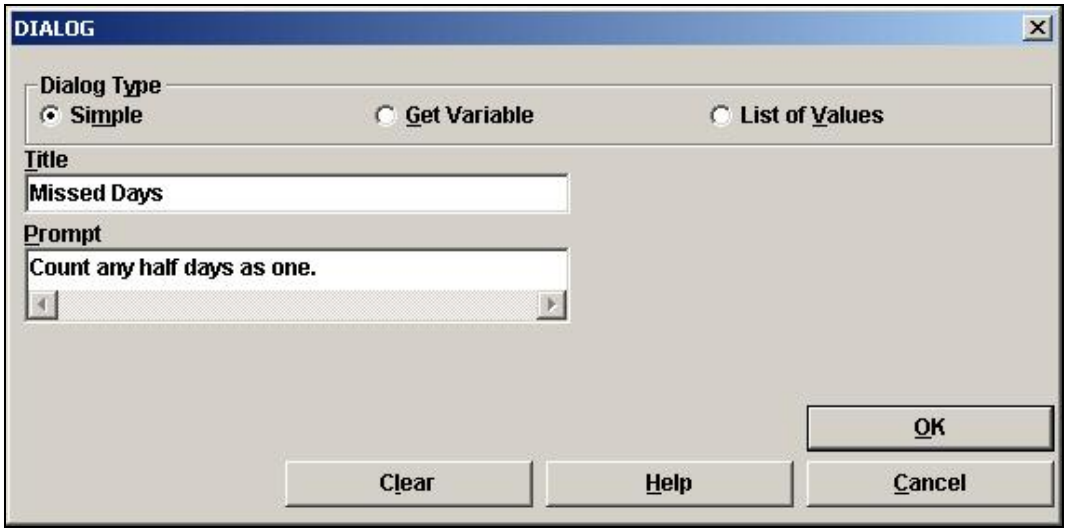
3. From the User Interaction tab or the Check Command Tree, select **Dialog**. The DIALOG box opens.



4. In the Title field, type **Missed Days**.

   - The Dialog Type radio button will be selected as Simple.

5. In the Prompt field, type **Count any half days as one**.

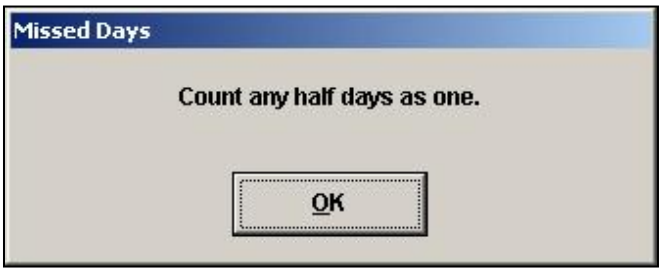


6. Click **OK**. The code appears in the Program Editor.

   DIALOG "Count any half days as one." TITLETEXT="Missed Days"

7. Click **Save** in the Program Editor.

   - When you test your code in the next chapter on Enter Data, the following dialog box should appear after you enter data in the MissDays variable.



🕮 Skills Review

Place a checkmark by the answer to the following questions. Refer to *Appendix D Skills Reviews Answer Key* for additional details and the answers.

1. Check Code can run before or after values are entered in a variable.

☐ True

☐ False

2. Using the Program Editor, you can create Check Code that will run across pages or records.

☐ True

☐ False

3. In the Check Command window, variables tagged with an X have Check Code associated with them.

☐ True

☐ False

4. Which function is used to create an age variable?

☐ ROUND

☐ YEARS

☐ TXTTONUM

5. How does the answer yes appear in the Program Editor?

☐ (+)

☐ (-)

Lesson Complete!

WHAT YOU LEARNED

How to:

- Navigate the Program Editor.
- Create a Skip Pattern using the IF/THEN commands.
- Use the ASSIGN command.
- Use the YEARS function.
- Use the DIALOG command.

Enter Data

SUMMARY

This lesson introduces you to entering data from the MakeView and Enter Data modules of Epi Info™. You will learn how to enter data, navigate the survey you created in MakeView, and enter data for the sample survey provided. You will also learn how to search for records from the Enter module. When you enter records in the survey, you populate the data table with information to be used for later analyses.

In Lesson 1, you created an electronic version of the survey you will use to collect data to aid in your asthma initiative project. In Lesson 2, Check Code was added to speed the data entry process and act as a quality control measure. In Lesson 3, you will begin entering data about the students from the survey population. Data entered in this lesson will be used to create statistics and develop a picture of the student body of the county seeking funds.

Time to complete: 1 hour Beginner

Getting Started with Enter Data

Overview of basic information to navigate the Enter Data module.

Enter Data to the Survey

Use the record information provided to enter data into the view you created. Verify your Check Code works as intended.

Enter Data to the Sample Survey

Provides you the opportunity to view a complete survey with existing records. Enter data from the surveys in Appendix C.

Find Records

Introduction to the Find feature. Practice lesson included.

Skills Review

Series of five review questions based on the reading and hands-on activities in this lesson.

BEFORE YOU BEGIN

Complete Lesson 1 MakeView: Create a Survey.

Complete Lesson 2 MakeView: Check Code.

Download the sample project Asthma Survey 2005.

WHAT YOU NEED

Asthma Survey.MDB – The project you created in Lessons 1 and 2.

Asthma Survey 2005.MDB – The provided dataset used in this lesson and Lesson 4.

Appendix C Sample Parent School Asthma Pre-Intervention Surveys.

Appendix D Skills Reviews Answer Key.

FIVE GOALS

- Learn to navigate the Enter Data module.
- Enter four records in the view you created.
- Verify the data entry conditions and Check Code functions.
- Troubleshoot your Check Code, if errors exist.
- Enter two records in the sample view and data table provided. The sample view and data table Asthma Survey 2005.MDB will be used in Lessons 4–10.

Getting Started with Enter

The Enter Data module can be accessed by clicking **Enter Data** from the Epi Info™ main page. You can also access Enter Data through the MakeView module by selecting **File>Enter Data**. Use the Enter Data module to enter information into the survey/view you created in MakeView. When you enter data into the view, you populate the data table inside your project (.mdb). You can navigate from MakeView to Enter to test your Check Code and Tab Order, but once you create a data table and start adding data, you cannot make changes to field names or types in your view.

The Enter program displays the view you created in MakeView, constructs your data table, and controls the data entry process by using the settings and Check Code established in MakeView. From the Enter Data module, you will enter new data, modify existing data, or search for records. The Find function allows records to be located based on a series of matched variables. As you enter data, the cursor moves from field-to-field, page-to-page, and saves data. Navigation is provided for New, Next, Previous, First, and Last records, and through related tables.

The Enter Data Workspace

If you do not have the project Asthma Survey.MDB that was created in Lessons 1 and 2, you can use the provided project titled Asthma Survey 2005.MDB. Use this project for a reference as you read the navigation sections and then skip to Section 3: Enter Data in the Sample Survey.

Let's enter the data workspace.

1. From the Epi Info™ main page, click **Enter Data**. The Enter window opens.

2. Select **File>Open**.

3. Open the project you created in MakeView titled **AsthmaSurvey**. The Select a Table dialog box opens.

4. Select **PreInterventionSurvey**.

5. Click **OK**. The data table dialog box opens.

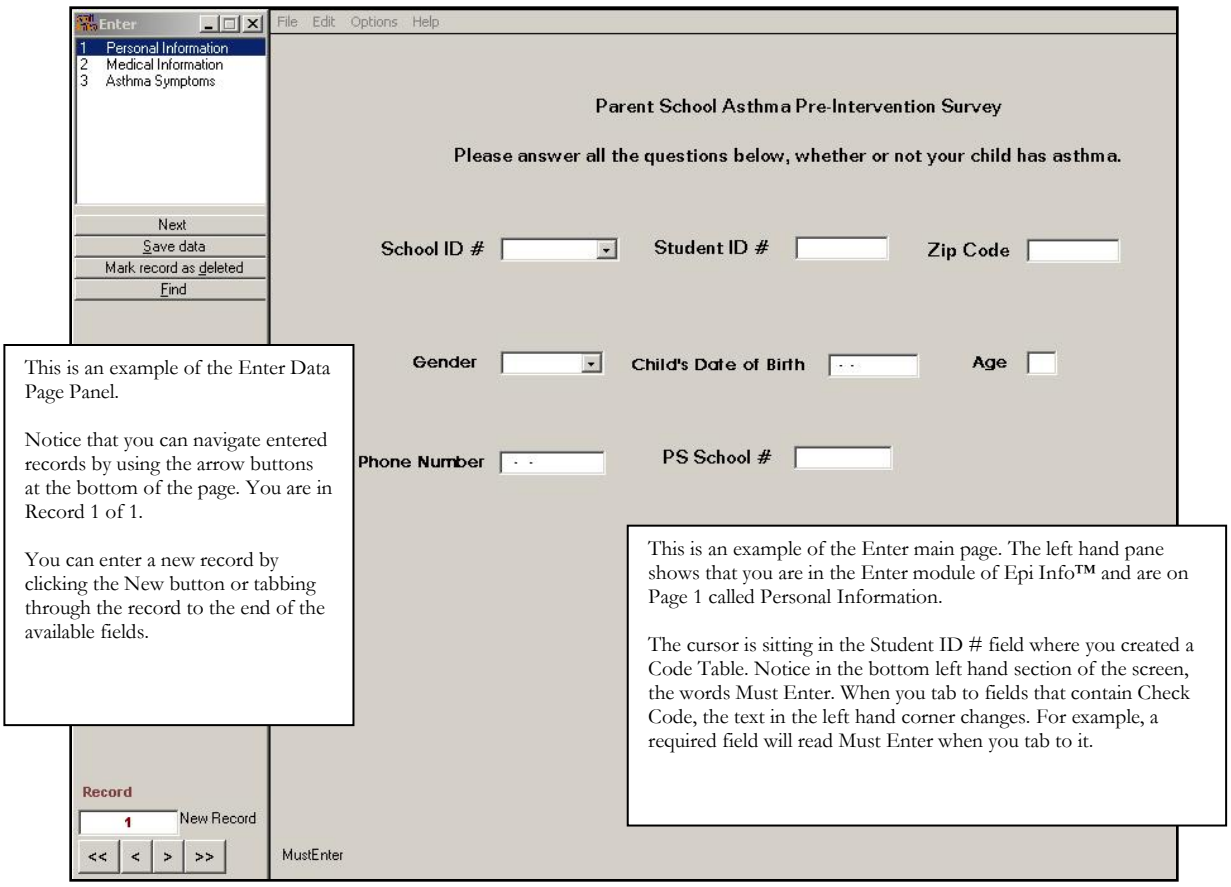
6. Click **OK** to create a new data table. The New Data Table dialog box opens.

7. Click **OK** to keep the data table name, and Unique ID set as the default. The Enter page opens with the view you made in MakeView.

   - Notice you are in Record 1 of 1 on page 1 titled Personal Information.

   - The Enter Page Panel displays a list of pages created in the project. Data entered into a page is automatically saved; however, you should use the **Save Data** button or **File>Save** after making changes to existing records.

This is an example of the Enter Data workspace. The top of the page shows you are in the Enter module.



This is an example of the Enter Data Page Panel.

Notice that you can navigate entered records by using the arrow buttons at the bottom of the page. You are in Record 1 of 1.

You can enter a new record by clicking the New button or tabbing through the record to the end of the available fields.

This is an example of the Enter main page. The left hand pane shows that you are in the Enter module of Epi Info™ and are on Page 1 called Personal Information.

The cursor is sitting in the Student ID # field where you created a Code Table. Notice in the bottom left hand section of the screen, the words Must Enter. When you tab to fields that contain Check Code, the text in the left hand corner changes. For example, a required field will read Must Enter when you tab to it.

Navigating Enter Data

You cannot navigate through pages until data has been entered in the preceding page or until all required fields are filled. You can familiarize yourself with the workspace prior to adding records. Keep the Enter module open to complete the next lesson.

Enter Data Main Page

File

Use to open, close, save, print, or exit. You can also compact your database with the compact utility. This reduces the size of your database, but does not affect your data.

Edit

Use to find specific records or to delete records. Deletion removes the current record from the project, and all records associated with it. It also marks the record as deleted in the

database. You can exclude deleted records from your analyses. Click **Undelete** to return the record to active status.

Options

Use to select the values of Yes/No variables. Variables can appear as Yes/No or True/False. You can also maximize the size of your screen image by selecting **Full Screen**.

Help

Use to view the online help.

Enter Data Page Panel

Use to navigate pages in your view. If you are entering data and have a required field, you must enter information into that field before proceeding to the next page. From the Enter Page Panel, you can go to the next record, save data, delete a record, find a record, or create a new record.

Use the navigation arrows at the bottom of the Enter Page Panel to move between records. Single arrows move you one record forward or backward. Double arrows move you to the first or last record.

Type a specific record number in the Record field and press **Enter** to go directly to a record.

Entering Data into your View

You will begin entering the four student records listed in the table on the following page. Tab through each field to run the Check Code and to navigate through the survey. If you do not tab through the survey, your Check Code will not run.

1. Verify that your Check Codes are functioning.

   - Is the Code Table you created between the variables for School Number and PS School Number functioning?
   - Did the Age field populate after you tabbed out of the Date of Birth field?
   - If your answer to Missed Sports is no, is the Number of Times variable skipped? Did the cursor go to the following question? If the answer is yes, did the cursor go to the Number of Times variable?
   - Did the dialog box pop up as you tabbed through the question regarding missed days?
   - After entering all the records, use the arrow keys to navigate between records and pages. Is the code functioning across records and pages?

2. Select **File>Save**.

Check Code Troubleshooting

If your Check Code did not function, you can close the Enter module and return to MakeView to verify the following. Use Lesson 2 on Check Code to verify your code appears as coded.

- Verify your tab order. Coded fields must be in the proper order and on the proper page to function. For example, does the Age field come after the Date of Birth field in your tab order list? Are SpTimes and IllHome in the proper order? Are they on the same page?
- Did you save all your Check Code in the Program Editor after it was created? Open the Program Editor and verify that an asterisk is next to the variable.
- Is your code applied to the correct field? Are your If/Then statements in the proper order?
- Is your code syntax correct? You can verify this by placing the cursor inside the code and clicking **Save**. Did the code turn green or red? Red indicates an error in your code.
- Do your commands have END statements?

Enter the following Student Records

Refer to the section on Check Code Troubleshooting if your codes do not work.

Student 1	**Student 2**	**Student 3**	**Student 4**
School ID Number A	School ID Number B	School ID Number C	School ID Number D
Student ID Number 1	Student ID Number 2	Student ID Number 3	Student ID Number 4
Zip 12207	12084	12077	12059
Male	Female	Male	Male
01/29/1994	08/19/1995	11/08/1995	07/28/1994
555-555-5555	444-444-4444	333-333-3333	222-222-2222
Asthma, Wheezing	No Conditions	Asthma, Bronchitis	RAD
Yes, Medication	No	Yes, Medication	Yes, Medication
October	N/A	September	July
Wheeze, 2 times each week Cough, 2 times each week	Never	Wheeze, 3–6 times each week Cough, Every day but not all day	Wheeze, Never Cough, 2 times each week
Missed 11 days	Missed 0 days	Missed 5 days	Missed 3 days
Yes, Missed sports, 1–3 times a week	No, Did not miss sports	Yes, Missed sports, 2 times a month	No, Did not miss sports
Asthma in home=1	Asthma in home=0	Asthma in home=2	Asthma in home=1

Entering Data into the Sample View

If you are working with the Asthma Survey project you created in Lessons 1 and 2, you will change projects. The sample project contains an example of the entire survey created in MakeView with 10 records already entered. The sample project, Asthma Survey 2005, is located in the EIHA Tutorial folder

If you already have Asthma Survey 2005 open, skip to Step 7. If not, begin with Step 1.

Enter the data from the two attached surveys located in Appendix C.

1. From the Enter Data menu, select **File>Open**. The Select a Table dialog box opens.

2. Click **Change Project**. The Select a Project window opens.

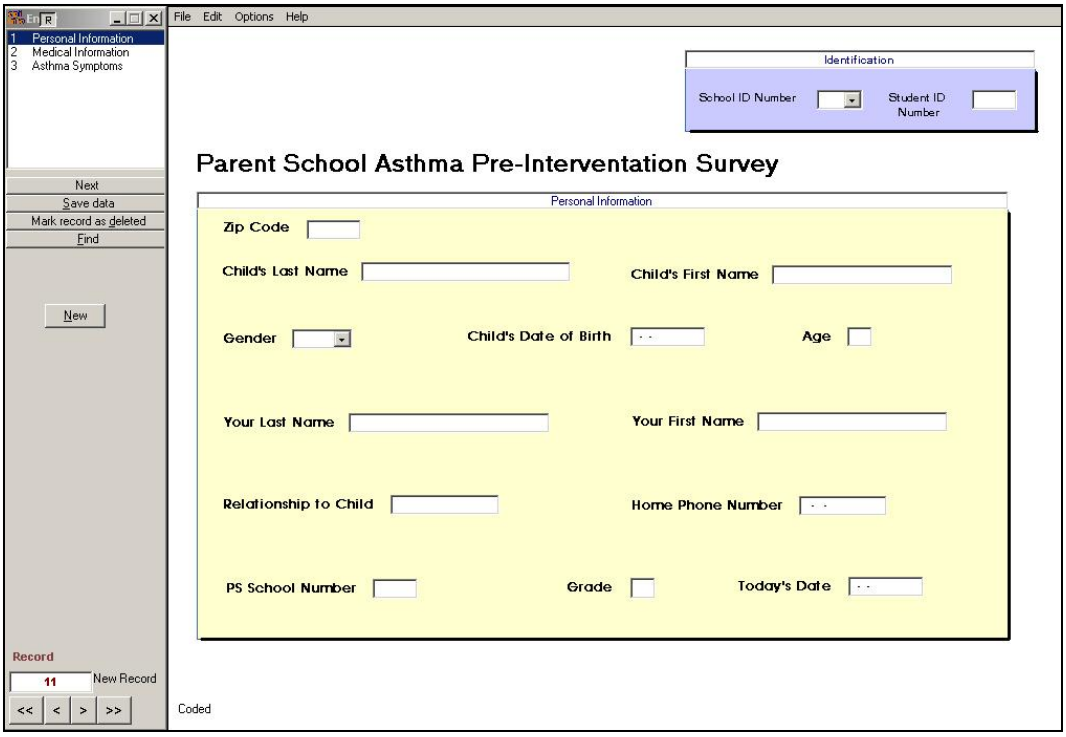
3. Locate the EIHA Tutorial folder and select the sample project Asthma Survey 2005.

4. Click **Open**. The Select a Table dialog box opens.

5. Select the view **PreInterventionSurvey**.

6. Click **OK**. The project opens to the next new record, number 11. The sample project contains 10 previously entered records.

7. Use the surveys from Appendix C Data Entry Surveys to add two additional records. You should have 12 records at the end of this lesson.


File Edit Options Help
1 Personal Information
2 Medical Information
3 Asthma Symptoms
Next
Save data
Mark record as deleted
Find
New
Record
11
New Record
<< < > >>
Identification
School ID Number
Student ID Number
Parent School Asthma Pre-Interventation Survey
Personal Information
Zip Code
Child's Last Name
Child's First Name
Gender
Child's Date of Birth
Age
Your Last Name
Your First Name
Relationship to Child
Home Phone Number
PS School Number
Grade
Today's Date
Coded

Finding Records

There are two ways to locate specific records in a large data table.

Time to put your record searching shoes on.

1. From the Record field located on the Enter Page panel.
   - Type the number **3** into the Record number field.
   - Press **Enter**. Record 3 of 12 appears in the window.
2. Click **Find** from the Enter Page panel. The Find Record window opens.
   - Use the Find feature to locate specific records in a large dataset to edit.
   - Select a variable (Zip) to search. An entry field appears. Enter the search criteria. For this example, enter **12207**. Click **OK** or press **Enter**.
   - To open a Record from Find, double click the **black arrow** to the left of the results row.
   - Embedded text items in multiline and text fields can be found by searching for *word*. This type of search is called a Wild Card search. In a Wild Card search, the asterisk is used to represent any letter or string of letters. For example, a search of last names as with the field ChildLast could look like *jo*. This search would pull names with the letters jo in the beginning or end of a word in the last name field (i.e., Jones, Johnson, Johns.
   - Dates and numeric fields can be searched using less than or greater than values (<>).
   - OR can be used search for matches based on more than one criteria.
   - ? can be used to search for a single unknown character.
   - Click **Reset** to clear the page and begin a new search.

Required Tasks

Find the students:

Who attend School A.

Who do not have bronchitis. Hint: Bronchitis NO.

Who missed more than 5 days of school. Hint: MissDays>5.

Who attend School A and have asthma.

Whose last name begins with S. Hint: ChildLast S*.

🕮 Skills Review

Place a checkmark in the correct answer for each of the questions below. Refer to *Appendix D: Skills Review Question and Answer Sheet* for further details.

1. Once you create a _____ you cannot change field names or types in your view.

- ☐ Survey
- ☐ Data Table
- ☐ Legal Value

2. What are the two ways you can enter a new record? Check all that apply.

- ☐ Tab through to the last record.
- ☐ Click New.
- ☐ Select the last available page.

3. If you are entering data and have a required field, you must enter information into that field before proceeding to the next page.

- ☐ True
- ☐ False

4. What expression could you enter in the Find feature to locate children who are aged more than 6 years?

- ☐ Age=6
- ☐ Age<6
- ☐ Age>6

5. From the Find feature, what button do you click to clear the page and begin a new search?

- ☐ Start
- ☐ Reset
- ☐ Cancel

Lesson Complete!

WHAT YOU LEARNED

How to:

- Navigate Enter Data.
- Enter records.
- Change projects.
- Find records.

Analysis: Basics

SUMMARY

This lesson introduces the Analysis module of Epi Info™. It shows different ways information can be viewed, based on selection inside Analysis.

In Lesson 3, you entered student data in a survey to collect information on asthma symptoms. In Lesson 4, you will use the Analysis module to READ in that data and view it. Once records are in Analysis, you can use several commands to SELECT, SORT, or LIST data that can be included in the asthma report you want to create to apply for funding. Lesson 4 covers these basic skills and shows you how to READ in data and Change Projects, which are important skills you will use every time you open the Analysis module.

Length of time to complete: 1 hour Beginner

Getting Started with Analysis

Overview of basic information to navigate Analysis. Includes information about the Current Project and Data Source settings. Use the READ Command to view the Asthma Survey 2005 Project.

List Command

Use to view the variables in your dataset.

Sort Command

Use to organize your data.

Select Command

Use to organize your data. Includes a practice lesson.

Skills Review

Series of five review questions based on the reading and hands-on activities in this lesson.

BEFORE YOU BEGIN

Complete Lesson 3 Enter Data

WHAT YOU NEED

Asthma Survey 2005.MDB - If you have not completed Lesson 3, you can use the Asthma Survey 2005.MDB project to complete this lesson as an introduction to Analysis; however, you will have 10 records instead of 12.

Appendix D: Skills Reviews Answer Key

FIVE GOALS

- Understand and be able to use the READ/IMPORT command. This command is used every time you open Analysis.
- Use the LIST command to create a grid or web table and view the variables in your data table.
- Use the SORT command to view variables in a specified order.
- Use the SELECT command to view requested variables.
- Be able to cancel a SELECT or a SORT and return to the original data count.

Getting Started with Analysis

The Analysis module can be accessed by clicking **Analyze Data** from the Epi Info™ main page. You can also select **Programs>Analyze Data**. Analysis can read data files created in MakeView and other types of databases (e.g., DBase, FoxPro, MS Excel). Analysis can produce graphs and interact with Epi Map to display geographic data. Once you have collected your data, Analysis acts as a statistical toolbox providing you with many ways to transform your data and perform statistical evaluations. Analysis also provides you with ways to export your data to new file formats (e.g., MS Excel).

The Analysis module contains three working areas: the Analysis Command Tree, the Analysis Output window, and the Program Editor.

1. The Analysis Command Tree contains a list of available commands separated into folders by command type. The Command Tree is used to create the functions and programs that run your statistics from the data. Selecting a command opens the corresponding dialog box for that command, function, or statistic to run.

2. The Output window acts as a browser and displays all the information that is generated from the commands run in the Program Editor.

3. The Output window buttons allow you to navigate through the program scripts that have been run and displayed in the output screen.

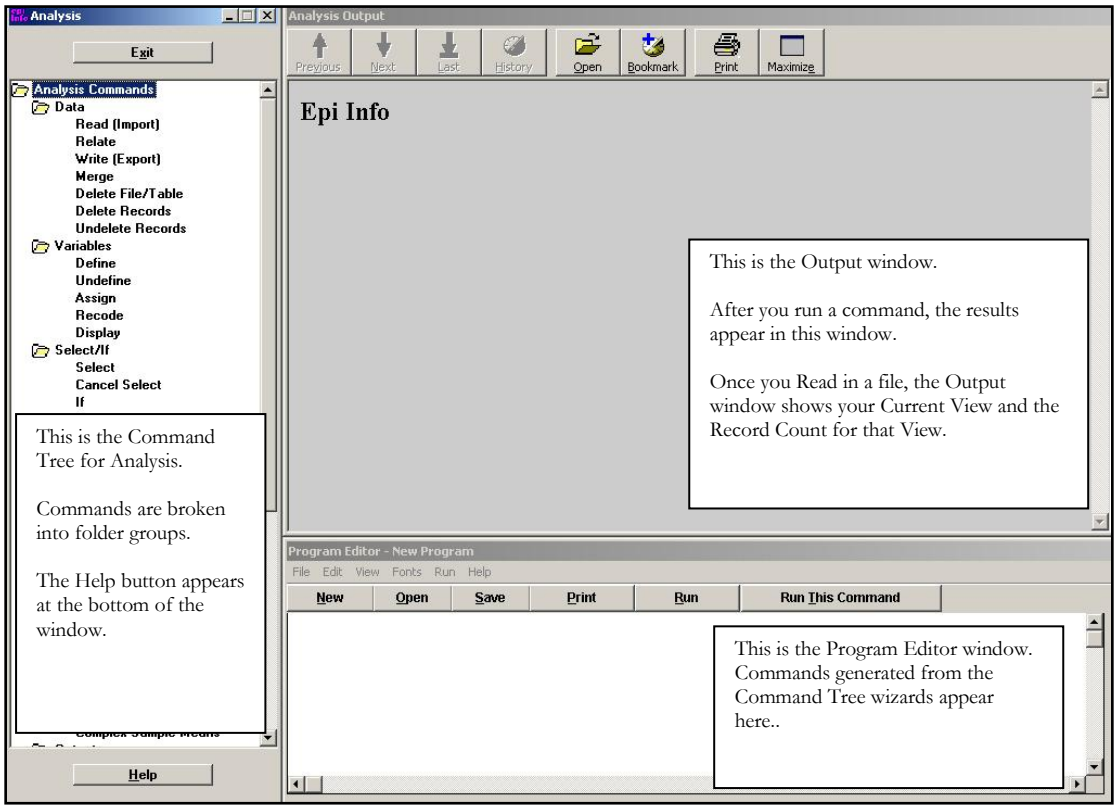
4. The Program Editor displays the dialog/code of the commands created using the Command Tree. Commands can be directly typed into the Program Editor. Programs or .PGM files written in Analysis are stored in the open .MDB. Programs can be saved and run against new data when the data becomes available.

The Analysis Workspace

Time to open the Analysis module.

1. Click **Analyze Data** or select **Programs>Analyze Data** to open the Analysis module.

   This is an example of the Analyze Data workspace. The top of the page shows that you are in the Analysis module. It consists of three windows: the Command Tree, the Output window, and the Program Editor.

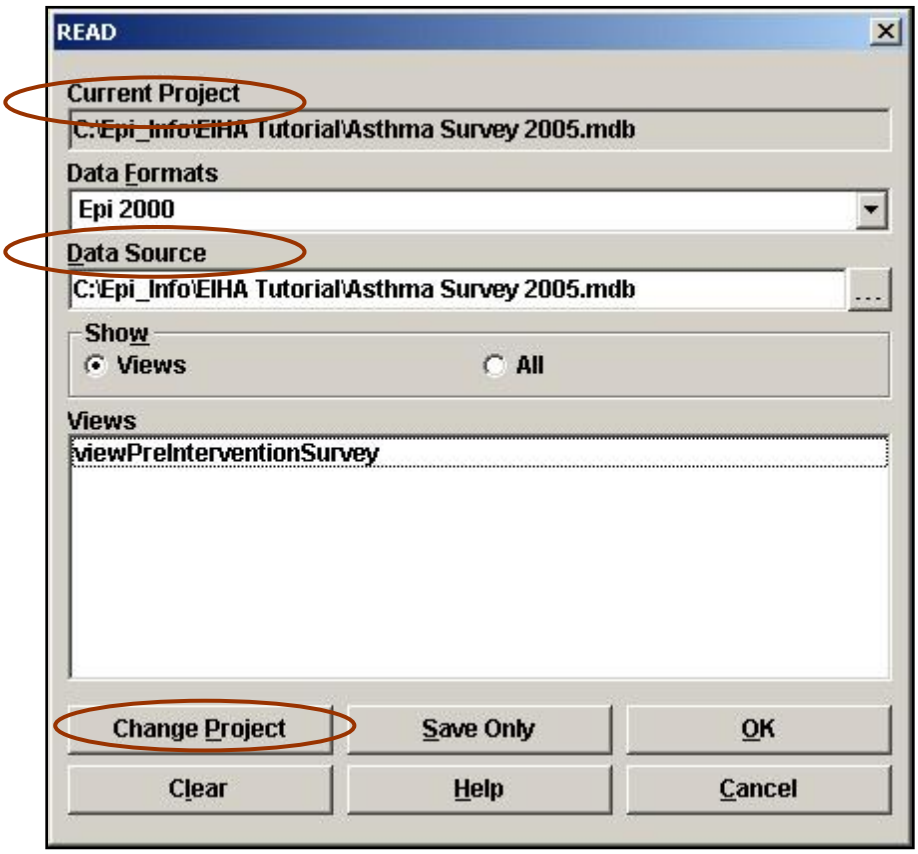


Reading in a Project

Read this section and Read/Import your survey data by completing the steps listed in the next section.

To analyze data, you must Read or Import a project into the Analysis module. The Read command allows you to select a project and data table on which to run your statistics. In order to begin, the Read/Import command will be used every time you open the Analysis module.

From the READ dialog box, you will select a project containing data for analyses. The Current Project is the database file (.MDB), which is the default source of data. Output files created from this data will be stored in this destination folder unless otherwise specified. In our project, the destination for Output files is the EIHA Tutorial folder.



It is recommended that the Current Project and the Data Source match. This allows you to keep all your data and tables together. If they do not match, click **Change Project**. The project you select will appear as the Current Project and Data Source.

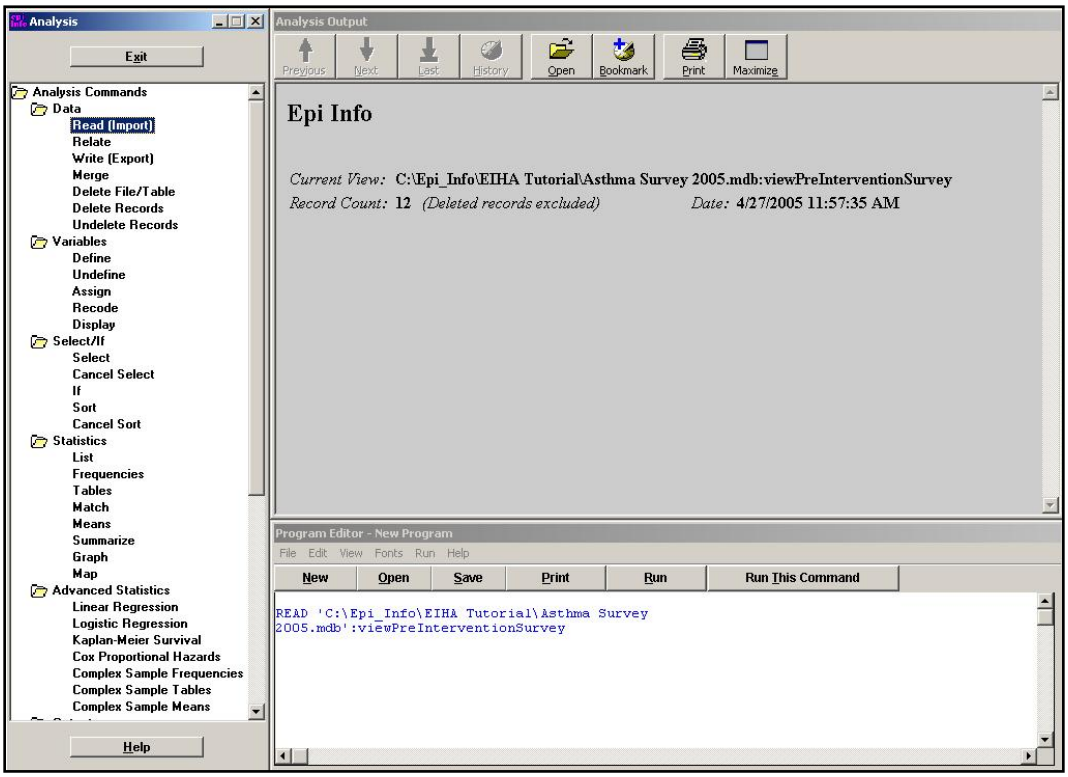
If you Read in a file where the Project and Source do not match, you will be asked to create a linked table. Data sources other than the current project are accessed through link tables located in the current project. Linked files are references to the data tables, and not the actual data, so if you move your files to a new location, the links will be broken.

From the Data Formats drop-down, you will **s**pecify the data format for the data to be read or written, (e.g., MS Excel). From the Show section, you can select project **Views** or **All** tables, views, and links for a project. When you click **OK**, the READ command will be saved in the Program Editor and will run the command at the same time. Save Only allows you to save the command to the Program Editor only and does not run the READ command.

Reading/Importing an Epi Info™ Project

Time to import your READing skills.

1. From the Analysis Command Tree, click **Read (Import)**. The READ dialog box opens.
2. Click **Change Project**. The READ File Name box opens.
3. Use the Look In drop-down menu to locate your C drive.
4. Open the **Epi_Info** folder.
5. Open the **EIHA Tutorial** folder.
6. Select the project **Asthma Survey 2005**. This is the project you added records to from the Enter module in Lesson 3.
7. Click **Open**. The Current Project field and the Data Source field inside the READ dialog box will populate.
8. Select the view **PreInterventionSurvey** from the Show section.
9. Click **OK**. The Output window opens.



- The Output window shows your current view from inside the EIHA Tutorial folder.
- You record count should be 12.
- The READ command you executed appears in the Program Editor.

Navigating Analysis

Read the following sections and familiarize yourself with the Analysis workspace. Click **Cancel** or the close X to exit any dialog or command boxes you open.

Output Window

The Analysis Output window acts as a browser to display all the output information that is generated from the commands created and run in the Program Editor. The Output window buttons allow you to navigate through program scripts that have been run and displayed including Previous, Next, Last, History, Open, Bookmark, Print, and Maximize.

Program Editor Menu

File

Use to create new files, open, save, close, print, and exit.

Edit

Use to copy, paste, delete, find, replace, or go to the program beginning, and end.

View

Use to place the Status Bar and/or the Tool Bar.

Fonts

Use to select fonts, style, and color.

Run

Use to run a program or run this command to run one line of code.

Commands in this Lesson

The Command Tree contains a set of commands located in a series of nine folders.

READ/IMPORT

READ is the most commonly used command. The READ (Import) command changes the current data source and/or the current project. It removes any standard defined variables. The READ command operates on many different types of data. Epi Info™ can Read in 24 different types of files. Located in the Data folder.

SELECT

Specifies a condition that must be true for a record to be processed. Use to select a set of records for analyses (e.g., select records based on gender or zip code). Located in the Select/If folder.

CANCEL SELECT

Cancels a previous SELECT command. Located in the Select/If folder.

SORT

Specifies the sequence in which records will appear when using the LIST, GRAPH, or WRITE commands. SORT organizes the listed data in ascending or descending order based on selected variables. For example, you can sort by last name or age. Located in the Select/If folder.

CANCEL SORT

Cancels a previous SORT command. Located in the Select/If folder.

LIST

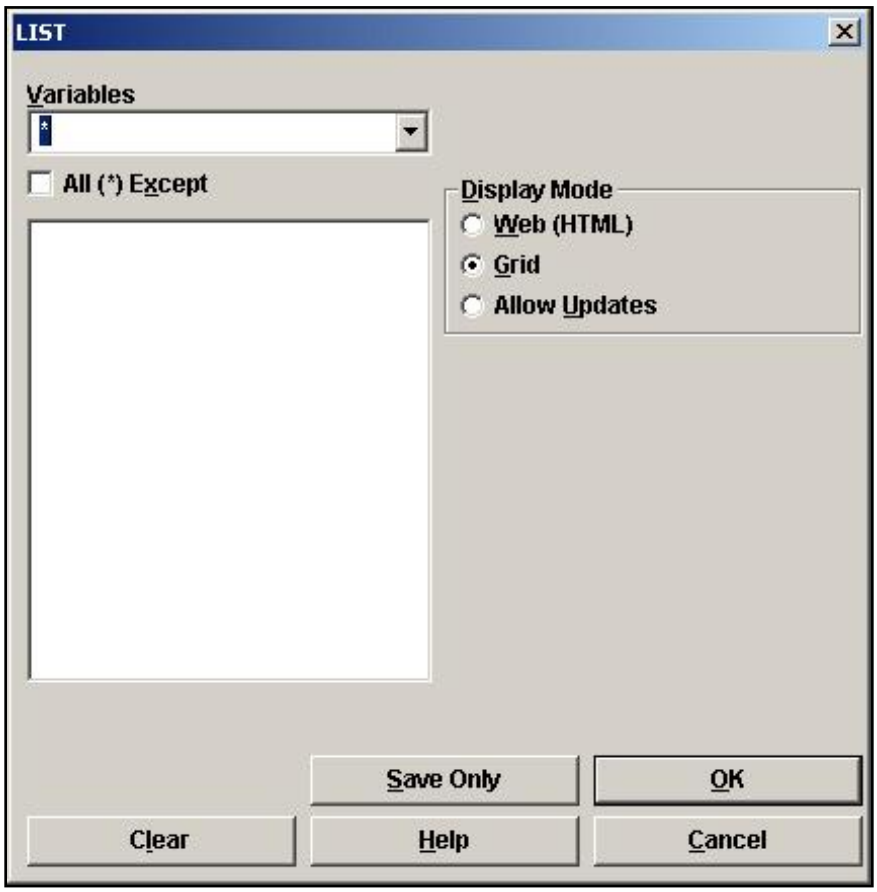
Creates a listing of the current data table. Lists can be customized to list all, exclude, or show specific records. Located in the Statistics folder.

Using the LIST Command

The LIST command creates a line listing of the current dataset. Select **Specific Variables** from the Variables drop-down to narrow the list or select ***** (the Wild Card) to display all the variables. Check the **All (*) Except** box and select from the Variables drop-down to exclude variables from the list.

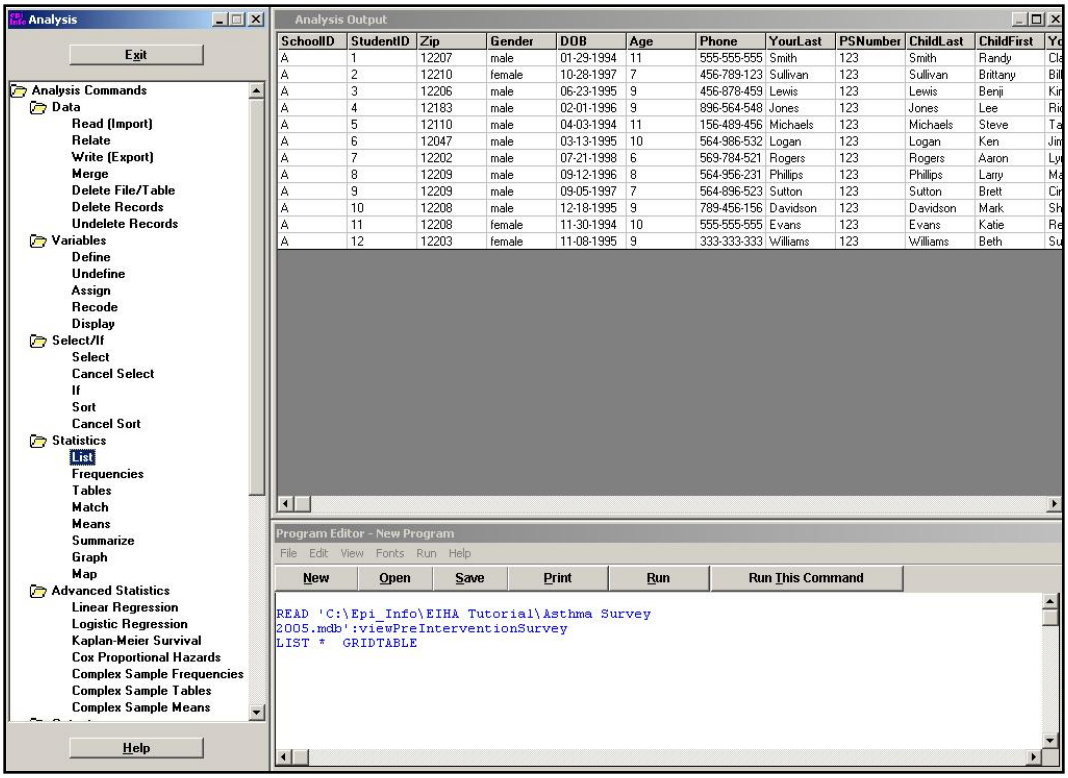
Now we'll LIST your next commands.

1. From the Command Tree Statistics folder, click **List**. The LIST dialog box opens.



2. Click **OK** to accept the default settings. The variables in your dataset appear in a grid table inside your Output Window.

   - Notice the LIST command in the Program Editor.
   - Grid output is not embedded, so an Output file is not created.



Creating a Web Format List

More listing here.

1. Click **List**. The LIST dialog box opens.

2. From the Display Mode section of the LIST box, select the radio button **Web** (HTML).

3. Click **OK**. The List appears in the Output window in a web table format.

   - Web display mode creates an embedded Output file.

   - The title bar now contains the name of your default output file.

Using the Sort Command

The SORT command allows you to specify the sequence in which records will appear. If more than one variable is selected, records are sorted in order by the first variable. After you sort records, you must use the LIST command to see the results.

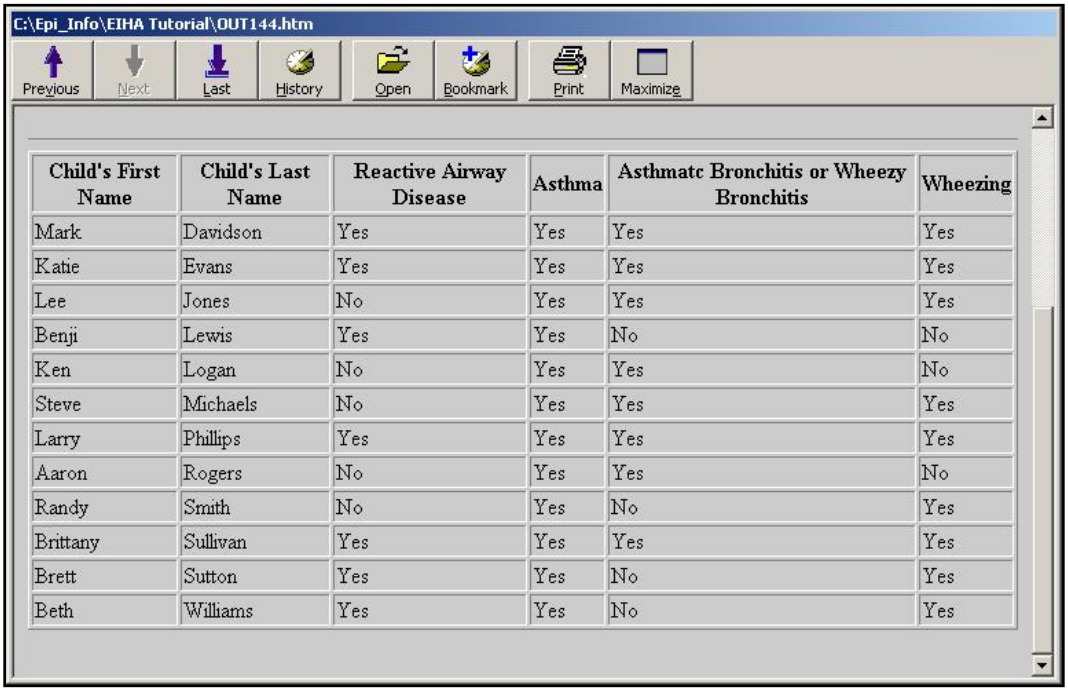
Now sort out the lists.

1. From the Command Tree Select/If folder, click **Sort**. The SORT dialog box opens.
2. Double click the variable **DOB** to place it in the Sort Variables field.
   - Variables can be sorted by ascending or descending order.
3. Click **OK**.
   - The Record Count remains at 12.
4. Use the **LIST** command to view all the variables. You can view the list in the format you prefer, Grid or HTML.
   - Notice how the list is sorted based on the DOB variable.
5. Click **Cancel Sort**. The CANCEL SORT dialog opens.
6. Click **OK**.
   - Records remain sorted until you close Analysis, Read in a new project, or cancel Sort.
   - Notice the Sort information is listed in your Output window.

Current View: **C:\Epi_Info\EIHA Tutorial\Asthma Survey 2005.mdb:viewPreInterventionSurvey**
Sort: **DOB**
Record Count: **12** *(Deleted records excluded)* *Date:* **9/14/2005 1:03:48 PM**

Creating a List of Sorted Variables

1. Click **Sort**. The SORT dialog box opens.
2. Double click the variable **ChildLast** to place it in the Sort Variables field.
3. Click **OK**.
4. Click **List**. The LIST dialog box opens.
5. Select **Web** (HTML) from the Display Mode.
6. From the Variables drop-down select:
   - ChildFirst
   - ChildLast
   - Asthmatic Conditions
7. Click **OK**. The selected variables appear in a web table format.
   - Notice the list is sorted alphabetically by the Child's Last Name variable.



Child's First Name	Child's Last Name	Reactive Airway Disease	Asthma	Asthmatc Bronchitis or Wheezy Bronchitis	Wheezing
Mark	Davidson	Yes	Yes	Yes	Yes
Katie	Evans	Yes	Yes	Yes	Yes
Lee	Jones	No	Yes	Yes	Yes
Benji	Lewis	Yes	Yes	No	No
Ken	Logan	No	Yes	Yes	No
Steve	Michaels	No	Yes	Yes	Yes
Larry	Phillips	Yes	Yes	Yes	Yes
Aaron	Rogers	No	Yes	Yes	No
Randy	Smith	No	Yes	No	Yes
Brittany	Sullivan	Yes	Yes	Yes	Yes
Brett	Sutton	Yes	Yes	No	Yes
Beth	Williams	Yes	Yes	No	Yes

8. Click **Cancel Sort**. The CANCEL SORT dialog box opens.
9. Click **OK**. The original dataset order is applied.

Use the Select Command

You can use the SELECT command to specify criteria for data to be included in Analysis functions. Think about the various ways you could divide your data into groups using this command. From the current data, you could select based on Age, Gender, Zip Code, or any combination of variables, in addition to illness information.

Once you make a selection, it affects all the analyses that follow. You must use the Cancel Select command to remove the selection. Selections accumulate until you Cancel Select, close Analysis, or Read in a new project. Use the LIST Command to see the selected records.

From the survey, select only the male students' data.

Let's start selecting.

1. From the Command Tree Select/If folder, click **Select**. The SELECT dialog box opens.

2. From the Available Variables drop-down, select **Gender**. Gender appears in the Select Criteria field.

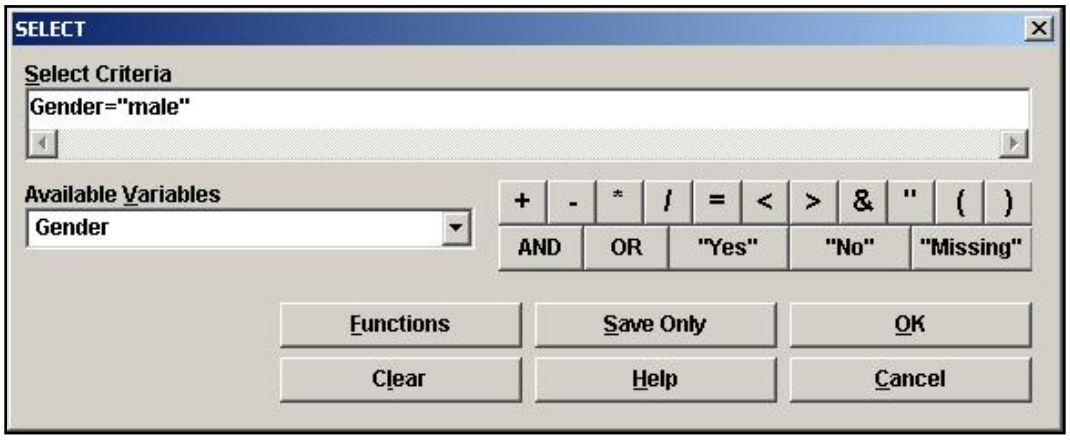
3. Click the **=** operator.

4. Click or type quote **"**.

5. Type the word **Male**.

   - The word Male must be enclosed in quotes because it is a text field. Text variables must appear in quotes.

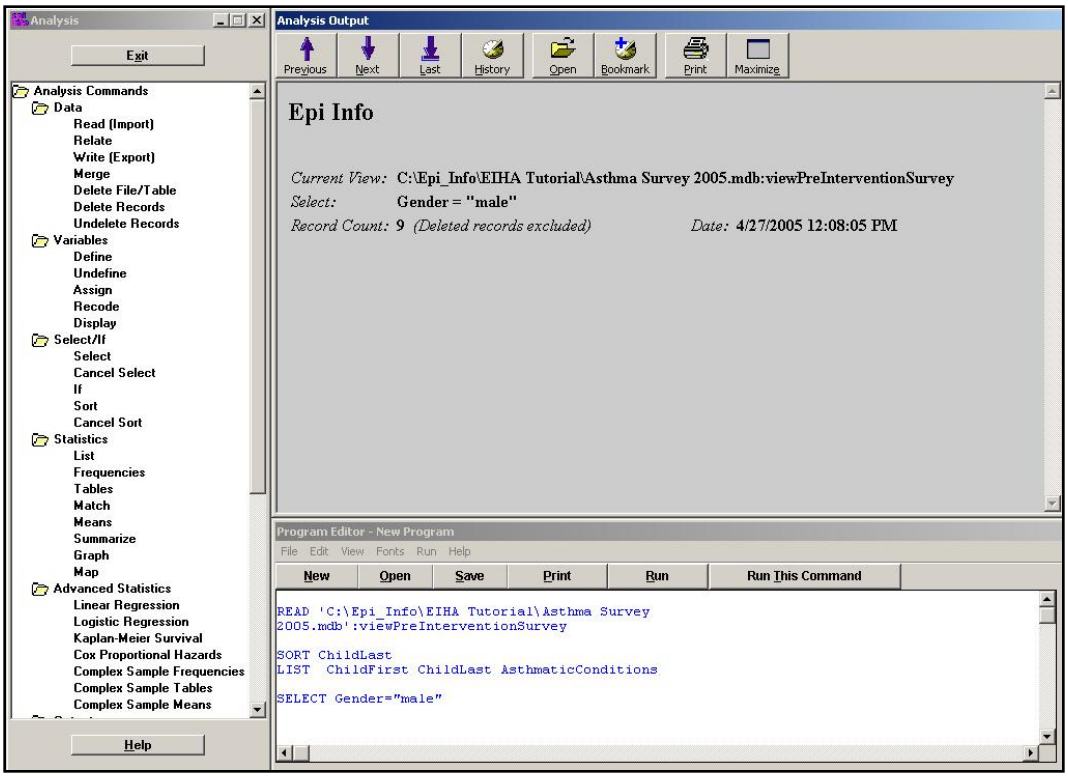
6. Click or type quote **"** to close the expression.



7. Click **OK**. The Output window updates with your new selection criteria.

   - The current view remains the same

- The Select criteria appears as Gender="Male"
- The record count changes to 9.
- The Program Editor contains the Select command information.



Listing the Selected Records

And even more listings.

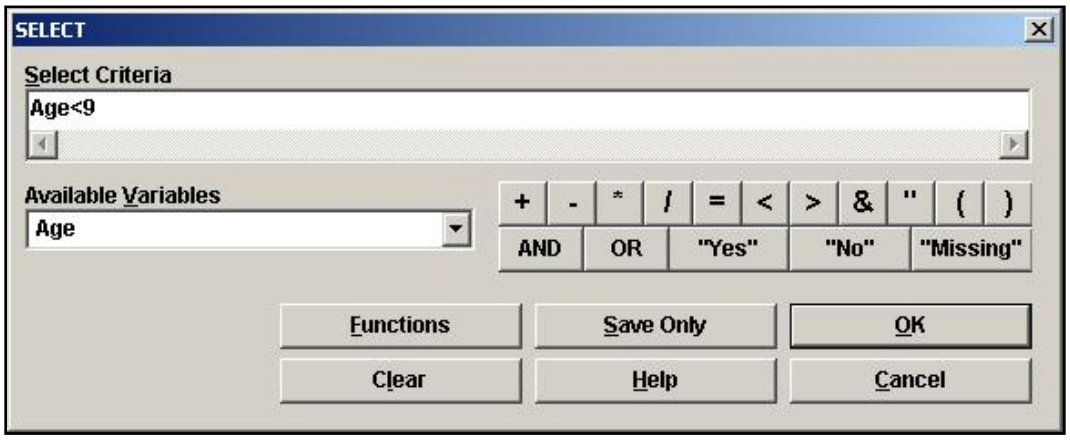
1. Click **List**. The LIST Dialog box opens.
2. Select **Web** (HTML) as the Display Mode.
3. Click **OK** to view all the variables.
   - All the records in the list are for male students.

Selecting Based on Age

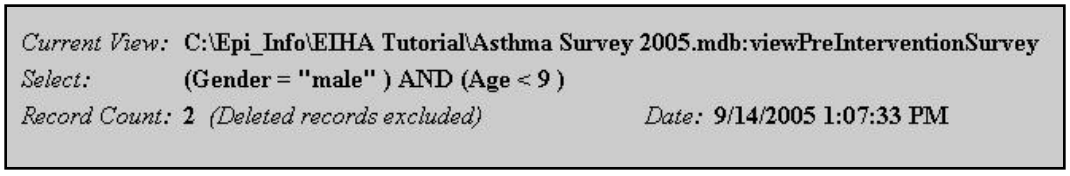
Remember the Select command has a cumulative effect on your data. Now you will select all the male students who are younger than 9 years of age. You have already selected the male students.

Now for the aging process.

1. Click **Select**. The SELECT dialog box opens.
2. From the Available Variables drop-down, select **Age**.
3. Click or type **<**.
4. Type the number **9**.
   - The Select Criteria field should read Age<9.



5. Click **OK**. The Output window updates.
   - The Record Count is now 2.
   - Your Select criteria reads (Gender = "male") AND (Age < 9). This expression could have been entered into the Select Criteria field at one time.
   - The expression SELECT Gender = "male" AND Age < 9 could have been typed and Run directly from the Program Editor.



6. List the records in a Web format.

7. Click **Cancel Select**.

- The Record Count returns to the original 12.

Required Tasks

Select the students who answered Yes to Bronchitis. (Hint: Bronchitis=(+)) You should get a Record Count of 8.

Sort those eight students by age.

List the students in a web format.

Do not forget to **Cancel Sort** and **Cancel Select** to return to the original 12 records.

🕮 Skills Review

Place a checkmark next to the correct answer for each of the following questions. Refer to *Appendix D Skills Reviews Answer Key* for further details.

1. After you run a command, the code appears in which window?

- ☐ Output
- ☐ Program Editor
- ☐ Command Tree

2. What command allows you to open a project for analysis?

- ☐ LIST
- ☐ READ
- ☐ SELECT

3. What command would you use to see variables and records in the current data table?

- ☐ LIST
- ☐ SORT
- ☐ READ

4. What command would you use to specify a sequence for records to appear?

- ☐ SORT
- ☐ LIST
- ☐ SELECT

5. What are the two Display Modes you can use with the LIST command?

- ☐ Web (HTML)
- ☐ Grid
- ☐ MS Word

Lesson Complete!

WHAT YOU LEARNED

How to:

- Read/Import an Epi Info™ project.
- Use the LIST Command.
- Use the SORT Command.
- Use the SELECT Command.
- Cancel, Sort, and Select of records.

This page intentionally blank.

Analysis: Creating Statistics

SUMMARY

This lesson introduces you to creating statistics using commands available in the Analysis module. You will create results using the FREQUENCIES, MEANS, and TABLES commands. Record your results on the Tear Out Answer Sheet provided. Verify your results with the answers listed in *Appendix E Lessons 5 Answer Key*.

In Lesson 4, you used the READ command to view records entered in the survey during Lesson 3. In Lesson 5, you will merge the 12 records from the Asthma Survey 2005.MDB into records gathered and recorded in an MS Excel spreadsheet format. Many times information needed for analyses will come from a variety of sources. In this example, you have student records in an .MDB and in an MS Excel spreadsheet. You will use the MERGE command to place all the student records in one project and then create statistics to see what these records can tell you about asthma among the sampled student body.

From the data, you can discover the frequency of asthma among the students, the worst month for asthma among the students, and what is the average number of missed days of school because of asthma. These types of data will be used to create the report that illustrates the need for asthma initiative funding.

Length of time to complete: 2 hours Beginner

Commands in this Lesson

Brief explanation of the commands used in this lesson and the results produced by using those commands.

Display Command

Use to view information about the variables in the project.

Set Command

Use to generate viewing options for records and statistics.

Merge Command

Use to join information from an MS Excel spreadsheet to an Epi Info™ project file.

Frequency Command

Use to produce results based on a number of selected occurrences. Includes a practice lesson.

Skills Review

Use to create a frequency for a subset of records.

Means Command

Use to define an average.

Skills Review

Use to populate a field. Compute the MEANS of the new values.

Tables Command

Use to produce results based on cross-tabulation of variables. Includes a practice lesson.

Save a Program

Use the Program Editor to save and run the program created in this lesson.

BEFORE YOU BEGIN

Complete Lesson 4 Analysis Basics

WHAT YOU NEED

Asthma Survey 2005.MDB

AsthmaTable.xls

Tear Out Q/A Sheet – Record your answers on this page.

Appendix E Lesson 5 Answer Key – Compare your answers to this page.

FIVE GOALS

- Merge records from an Epi Info™ project file with an MS Excel file.
- Use the FREQUENCY command to discover which condition has the highest number of affected students.
- Use the MEANS command to discover which school in the survey had the highest average of missed school days because of breathing problems.
- Use the TABLES command to discover the odds and risk ratios for having wheezing and asthma.
- Save a program (.PGM) to use in a future lesson.

Tearing Out Q/A Sheet

Use this sheet to record your answers in the Answers column as you progress through Lesson 5. Answers are listed in Appendix E Lesson 5 Answer Key.

A. Questions in Lesson 5 are marked with a letter.

Frequency Sections	**Answers**
A. Are there more males or females in the survey population?	
B. Which condition has the highest frequency?	
C. Which month had the highest number of students with breathing difficulties?	
D. Which school had the highest number of students with bronchitis?	
E. Which gender has the highest frequency of asthma?	
F. How many students answered yes to Reactive Airway Disease? **G.** What is the percentage of students with Reactive Airway Disease that have been prescribed medication?	
H. Which zip code has the most surveyed students living in it? **I.** Of that zip code, which condition had the highest frequency?	
J. How many students answered yes to all four conditions: asthma, Reactive Airway Disease, bronchitis, or wheezing. **K.** Based on all four conditions and zip code, are the conditions evenly distributed among the survey population? **L.** Does one school have more students in affected zip codes with all four conditions?	

Means Sections	**Answers**
M. What was the most frequent number of missed days? **N.** What is the mean (average) number of days missed because of asthmatic conditions?	
O. Which school had the highest average of missed days?	

P.	Was the school with the highest average of missed days for males the same as the school with the highest average of missed days for females?	
Q.	Is the average number of missed days higher among those who answered yes to asthma than the overall average? Refer to Question N.	
R.	What is the average age of the student population?	

Tables Section		**Answers**
S.	Does one school have a higher number of students with asthma?	
T.	Which school has the highest number of students with wheezing?	
U.	How many students had both asthma and wheezing?	
V.	What are the odds that you have the condition wheezing if you have the condition asthma?	
W.	Is the odds ratio for bronchitis and Reactive Airway Disease lower or higher than asthma and wheezing? Refer to Question V for the odds ratio for asthma and wheezing.	
X.	How many children with the condition wheezing were prescribed medication? From those results, what is the risk ratio?	
Y.	Is the risk ratio for being prescribed medication for the condition asthma higher or lower than being prescribed medication for wheezing? Compare the results to Question X.	
Z.	Are the risk ratios for asthma and wheezing higher for children under 10 years of age than they are for the survey population as a whole? Compare to the results from Question Y.	

Commands in this Lesson

This lesson uses the following commands.

DISPLAY

Use to display table, view, and database information. Use the display option Variables Currently Available to see all the variables in the dataset; including names, field types, and format information. Prior to merging or creating statistics, use Display to ensure that field types and variables names have been coded as needed. Located in the Variables folder.

SET

Use to view the current values of Analysis option settings and generate commands to change them. Statistical and graphic viewing options can be selected. Yes, no, and missing values can be viewed in alternate forms. Allows the inclusion or exclusion of missing records in statistical computations. Located in the Options folder.

MERGE

Use to merge records in one dataset with those in another, using one or more defined identifiers or build keys to establish the match between records. Merged records can be appended or updated to the new dataset. Merge is only supported when the READ data source is an Epi Info™ or MS Access table. The MERGE data source can be of any type. Located in the Data folder.

FREQUENCIES

Use to produce a table showing how data is distributed, counts of any given interval, how many records have each value of a selected variable, or the number of occurrences for a specified variable. If more than one variable is given, a separate table is created for each variable. Confidence limits for each proportion are included. Records may be included or excluded from the count by using SELECT statements. Located in the Statistics folder.

MEANS

Use to compute descriptive statistics for a continuous numeric variable. The mean of a set of data is equal to the sum of the data divided by the number of items in the data set. When used with a cross-tabulation variable, it also computes statistics showing the likelihood that the means of the groups are equal. The mean of a yes/no variable is the proportion of respondents answering yes. Means can be created for continuous or categorical values. Continuous variables fall along a continuum with one variable selected, (e.g., age). Categorical variables are groupings or categories with two variables selected (e.g., gender). If one variable is selected, the program produces a table similar to the Frequencies table plus descriptive statistics. If two variables are selected, the first is a numeric variable containing data to be analyzed and the second is a variable that indicates how the group will be distinguished. The output of this format is a table similar to that produced by Tables, plus descriptive statistics of the numeric variable for each value of the group variable. Means produces the following

statistical tests: parametric tests, Analysis of Variance or ANOVA (for two or more samples), student's t-test (for two samples), non-parametric tests, Bartlett's Test for Inequality of Population Variances, Kruskal-Wallis one-way analysis of variance (for two or more samples), and Mann-Whitney/Wilcoxon Two Sample Test. Located in the Statistics folder.

TABLES

Use to create cross-tabulations of categorical variables. Tables can help determine the probability that a risk factor is linked to an outcome. For these values to have their accepted epidemiological meanings, the value representing presence of the exposure (independent value) and outcome conditions (dependent variable) must appear in the first row and column of the table. Epi Info yes/no variables are automatically sorted. Values of the first selected variable will appear across the top of the table, and those of the second one will be on the left margin of the table. Normally cells contain counts of records matching the values in the corresponding marginal labels. For 2x2 tables, the command produces odds ratios and risk ratios. For tables other than 2x2, chi-squared statistics are computed. **Chi square for trend tests for the presence of a trend in dose response or other case control studies where a series of increasing or decreasing exposures is being studied.** The p-value is the probability that the association between two variables may be because of chance. A low p-value of <.05 means the risk factor may be closely associated with the illness. Located in the Statistics folder.

Display Files

Allows the display of table, view, and database information. You are going to use it to view information about the variables in your current dataset and to determine which variable to use as the Build Key when you merge your student records with the student records from MS Excel.

Time to play display.

1. Read/Import the Asthma Survey 2005 file located in your EIHA Tutorial folder. You should still have 12 records. Remember that your Project and Data Source need to match when you read in a file.

2. From the Command Tree window, click **Display** from the Variables folder. The Display dialog box opens. You want to see Information for Variables from the Variables Currently Available.

3. Click **OK** to accept the default settings. The variables in the Asthma Survey 2005 project appear in the Output window.

4. Scroll through the variables listed.

   - Display is an important command to use before importing data or completing analyses on unfamiliar data. If needed, you could use this information to format the cells in your spreadsheet prior to merging or importing files. Display can alert you to potential errors in Analysis (e.g., assuring that variables are coded as the same field type).

   - To see how the MS Excel file you are merging into the project is set up, you can open the AsthmaTable file in MS Excel and view the data. Do not make any changes to the file. Notice that the MS Excel and the Epi Info™ project file both contain variables called StudentID. This field corresponds to both sets of data and will be used to join the datasets. The spreadsheet does not

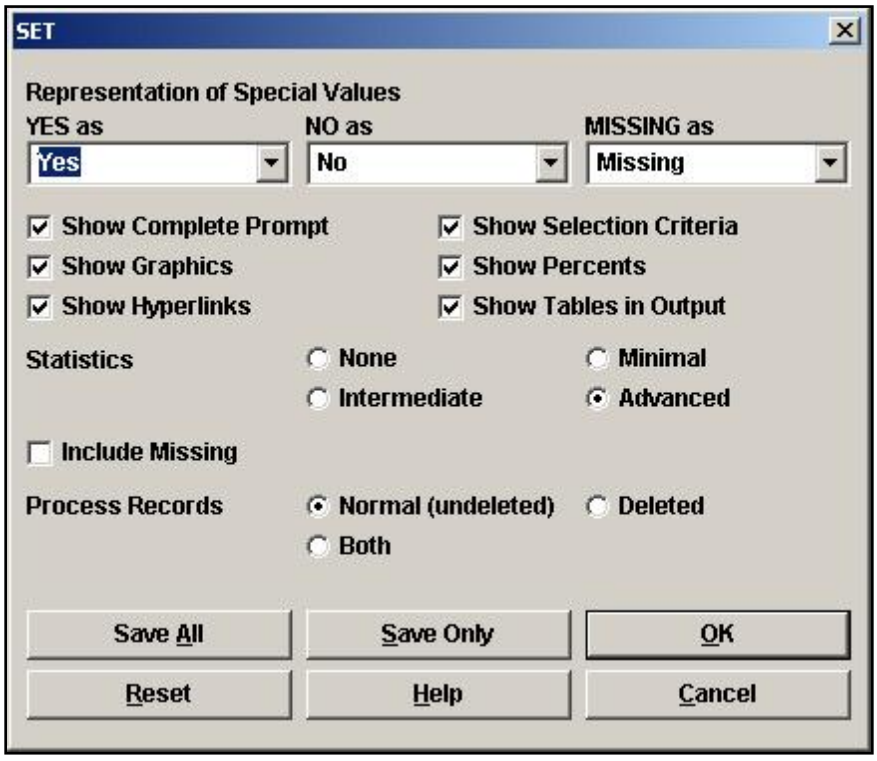
contain all the personal information collected in your Epi Info™ project. Values that do not correspond will appear as missing in Analysis. Be sure to close MS Excel before beginning the merge.

Set Options

The SET command provides various options that affect the performance and output of data in Analysis. These settings are used whenever the module runs. Make sure the Statistics setting is advanced because the risk and odds ratios may not appear in your results.

In football terms, down…..SET.

1. From the Command Tree Options folder, click **Set**. The SET dialog box opens.



2. Set your options as follows.

- Yes as Yes
- No as No
- Missing as Missing
- All Show boxes selected
- Statistics Advanced

- Include Missing not selected
- Process Records as Normal (undeleted)

3. Click **OK**.

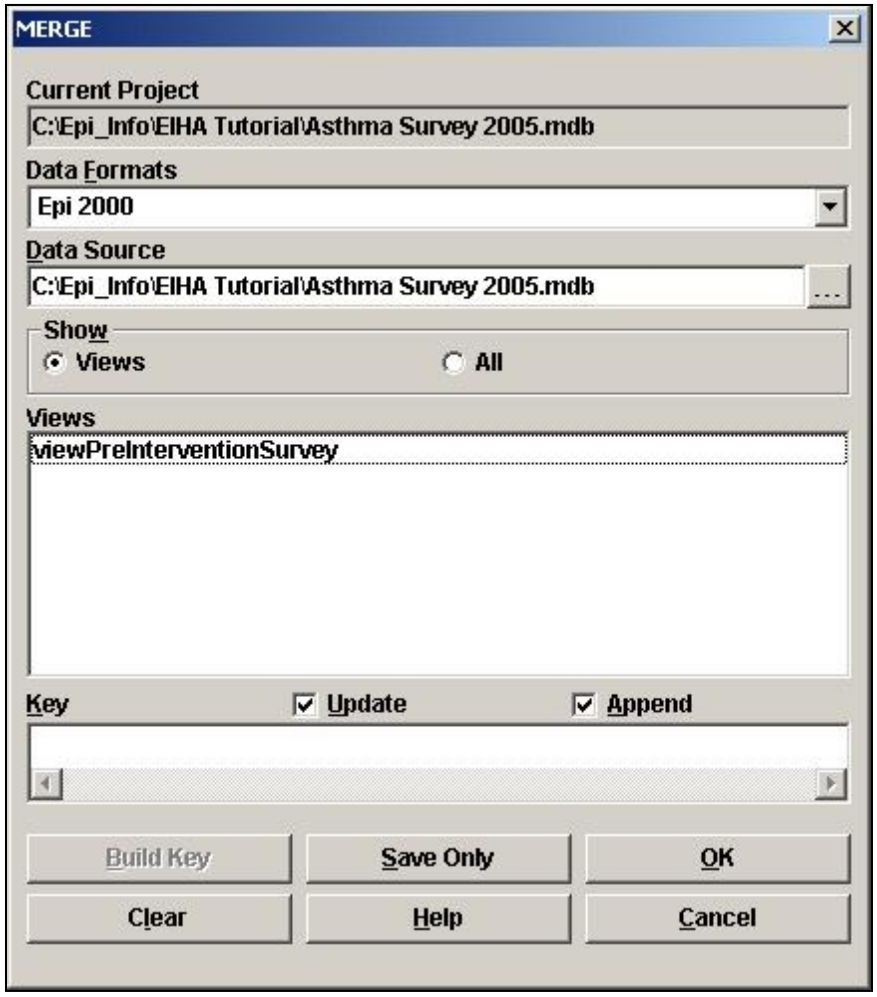
Merge Files

The current project contains 12 student records. In order to view all the available records for the county, you will use the MERGE command to join the 12 records you created in Enter Data with a larger data table gathered using MS Excel. Merge is only supported when the READ data source is an Epi Info™ or MS Access table. The Merge data source can be of any type. For this example, you will use the MS Excel spreadsheet called AsthmaTable.xls located in the EIHA Tutorial folder. A merge requires a Build Key that represents a matching variable inside both sets of data. The MS Excel file used for this lesson has been formatted so that the column headings and data types match the field names you created in MakeView. You will merge the files based on the Build Key StudentID. When you created the View, StudentID became a required field. In our example, there are no duplicate values on the field StudentID. Each record will have a unique value.

For the sake of togetherness, let's merge.

1. From the Command Tree window Data folder, click **Merge**. The MERGE dialog box opens. Notice that the MERGE dialog box is very similar to the READ dialog box.

   - Make sure your current project is C:\Epi_Info\EIHA Tutorial\Asthma Survey 2005.mdb.



2. From the Data Formats drop-down, select **Excel 8.0**. The Data Source field clears.

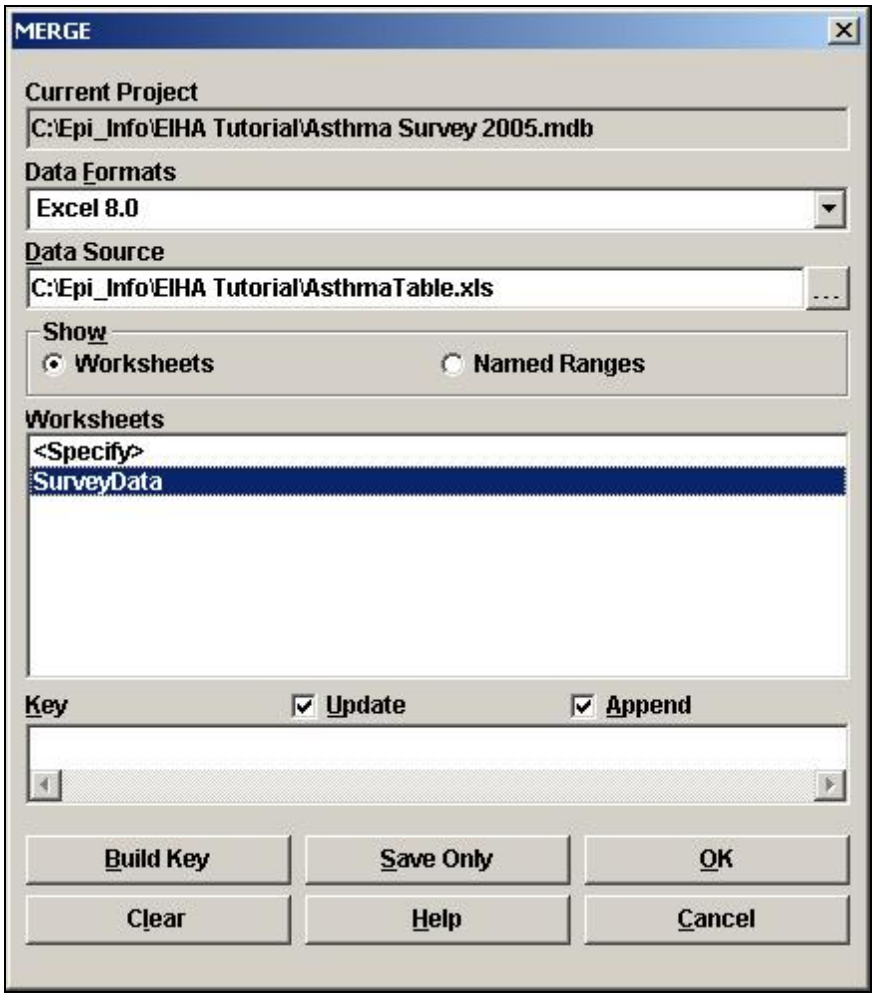
3. Click the **Data Source Browse**  button. The Merge File Name window opens.

4. Select the file called **AsthmaTable.xls** located in the EIHA Tutorial folder.

5. Click **Open**. The MERGE dialog box populates with information about the spreadsheet.

6. Click the Worksheet called **SurveyData**. The Build Key button activates.

- SurveyData is the named worksheet in the MS Excel file.



7. Click **Build Key**. The RELATE-BUILD KEY dialog box opens.

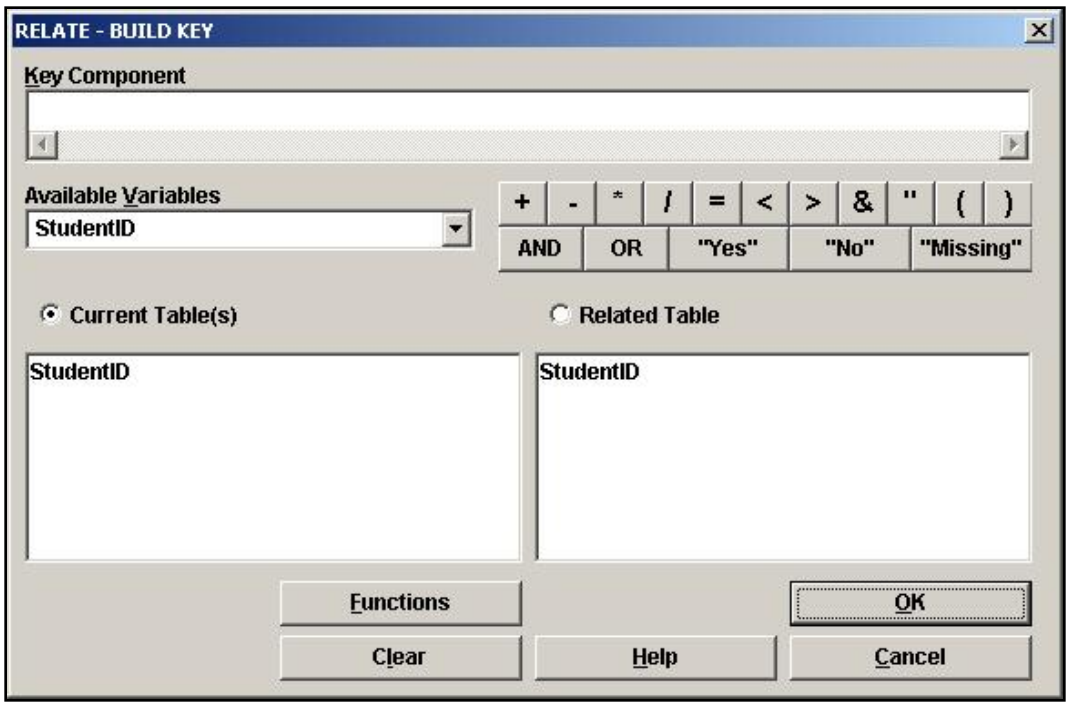
- Keep in mind that you need a variable that can be joined from the Current Table that corresponds to the Related Table.

8. From the Available Variables drop-down, select **StudentID**.

9. Click **OK**. The Current Table field shows StudentID.

10. From the Available Variables drop-down select, **StudentID**.

11. Click **OK**. The Related Tables field shows StudentID.



12. Click **OK** to accept the Build Key designations. The MERGE window opens.

- The Update and the Append checkboxes should be selected. For matching records, Update replaces the value of any field in the READ table whose build key matches the MERGE table. For unmatched records, Append creates a new record in the READ table with values only for those fields that exist in the MERGE table.

- For this merge, you want to update and append records, meaning that records containing the same StudentIDs will be updated (overwritten) if there is new information, and all other records will be appended (added) to the end of the data table.

13. Click **OK** from the MERGE dialog box. The FILESPEC dialog box opens.

- Make sure the First Row Contains Field Names checkbox is selected.

14. Click **OK**. The READ dialog box opens.

- Since a new data source has been brought into the project, you must decide if you want to create a temporary or permanent link between the two sets of data. For this exercise, you will be working with the 800 records for several exercises. Make the link permanent.

15. In the READ dialog box, type the link name **AsthmaTable.**

16. Click **OK**. The record count in the Output window reads 800.

- LIST the variables in a Grid.

- Notice the first 12 records contain information from your survey. The StudentIDs are the same, so the records were updated to reflect any changes.
- Click the close **X** to exit the grid listing and return to the Output window.
- Notice that the Program Editor shows the commands used to create the Merge and the List.

Creating a Frequency

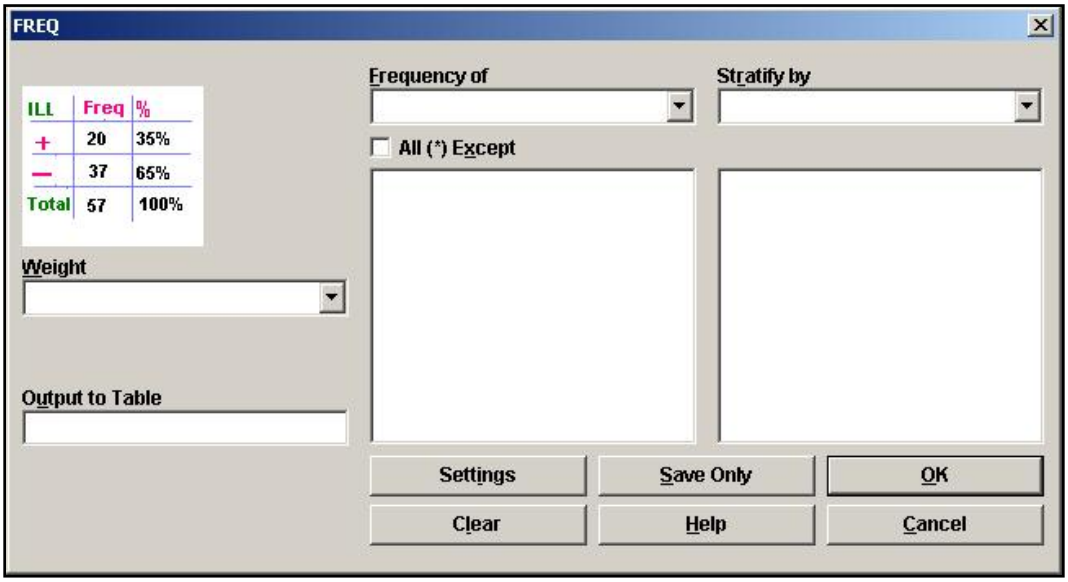
You will use the FREQUENCY command to get a count and view the 95% Confidence Limits on selected variables. The FREQ command produces a frequency table that shows how many records have a value for each variable, the percentage of the total, and a cumulative percentage.

FREQ * creates a table for each variable in the current view other than unique identifiers. This command is often used to begin analyses on a new data set. To exclude variables, use the All (*) Except checkbox and select variables to be excluded.

Since you just merged in new data, develop some basic information and build up to more complex statistics.

Now let's make this merge more frequent.

1. From the Command Tree Statistics folder, click **Frequencies**. The FREQ dialog box opens.



2. From the Frequency of drop-down, select **SchoolNum**.

3. Click **OK**. The Frequency distribution of SchoolNum appears in the Output window.

- The Frequency column provides the count of students per school. The Percent column indicates the school's percentage of the total number.

- The 95% Confidence Limits are a range of values for a variable that indicates the likely location of the true value of a measure. In this frequency, the Confidence Limits for School A are 21.6%–27.7%. The actual percentage is 24.5%, which falls within the range.

Forward

School Number	Frequency	Percent	Cum Percent	
A	196	24.5%	24.5%	
B	205	25.6%	50.1%	
C	208	26.0%	76.1%	
D	191	23.9%	100.0%	
Total	800	100.0%	100.0%	

Required Tasks

Answers to questions throughout Lesson 5 can be entered on the Tear Out Q/A Sheet provided at the beginning of the lesson. Each of the variables used in the following frequencies were created in the original survey in MakeView.

A. Are there more males or females in the survey population?

Hint: Create a Frequency of Gender to find out.

B. Which condition has the highest frequency?

Hint: Create a Frequency of Asthmatic Conditions. Notice you get four tables, one for each condition.

C. Which month had the highest number of students with breathing difficulties?

Hint: Create a Frequency of WMonth.

Stratifying a Frequency

Frequency tables can be stratified or grouped by a selected variable.

D. Which school had the highest frequency of students with bronchitis?

Stratify away!

1. Click **Frequencies**. The FREQ dialog box opens.
2. From the Frequency of drop-down menu, select **Bronchitis**.
3. From the Stratify By drop-down menu, select **SchoolNum**.
4. Click **OK**. You will get four tables containing the frequency of bronchitis per school.

Required Tasks

E. Which gender has the highest frequency of asthma?

Hint: Create a Frequency of Asthma stratified by gender.

Selecting Variables for a Frequency

Use the SELECT command to specify a subset of students and then create a frequency for only those records. You want to know how many students have Reactive Airway Disease and if they are currently taking medication for the condition.

F. How many students answered yes to having Reactive Airway Disease?

G. What is the percentage of students with Reactive Airway Disease that have been prescribed medication?

Time to generate more selection.

1. Use the SELECT command to view those who answered yes to having Reactive Airway Disease. Hint: RAD=(+)

2. Create a frequency of students who have been prescribed medication (Hint: Frequency of Medication)

3. Click **Cancel Select** to return to the original 800 records.

Required Tasks

H. Which zip code has the most surveyed students living in it?

Hint: Create a frequency of zip code (variable Zip).

I. Of that zip code, which condition had the highest frequency?

Hint: Use SELECT to view only those students living in zip code area 12046. (Zip Code is a text field. Text fields must be surrounded by quotes when used in commands Zip="12046")

Create a frequency of Asthmatic Conditions.

Remember to Cancel Select.

🕮 Skills Review

Record the following results on the Tear Out Q/A Sheet.

J. How many students answered yes to all four conditions: asthma, Reactive Airway Disease, bronchitis, or wheezing.

K. Based on all four conditions and zip code, are the conditions evenly distributed among the survey population?

L. Does one school have more students in affected zip codes and with all four conditions than the others?

You want to find out how many children from the survey responded yes to all the asthmatic conditions. Also, you want to know if the affected students are concentrated to one zip code or one school.

Select the students who had all four conditions.

Hint: Use the SELECT command to create the code below. You need to select the students who answered yes to asthma, bronchitis, RAD, and wheezing.

SELECT Asthma= (+) AND Bronchitis= (+) AND RAD= (+) AND Wheezing= (+)

Create a Frequency of Zip Code.

Create a Frequency of SchoolNum.

Cancel Select to return to the original 800 records.

Creating a Means

M. What was the most frequent number of missed days?

N. What is the mean (average) number of days missed because of asthmatic conditions?

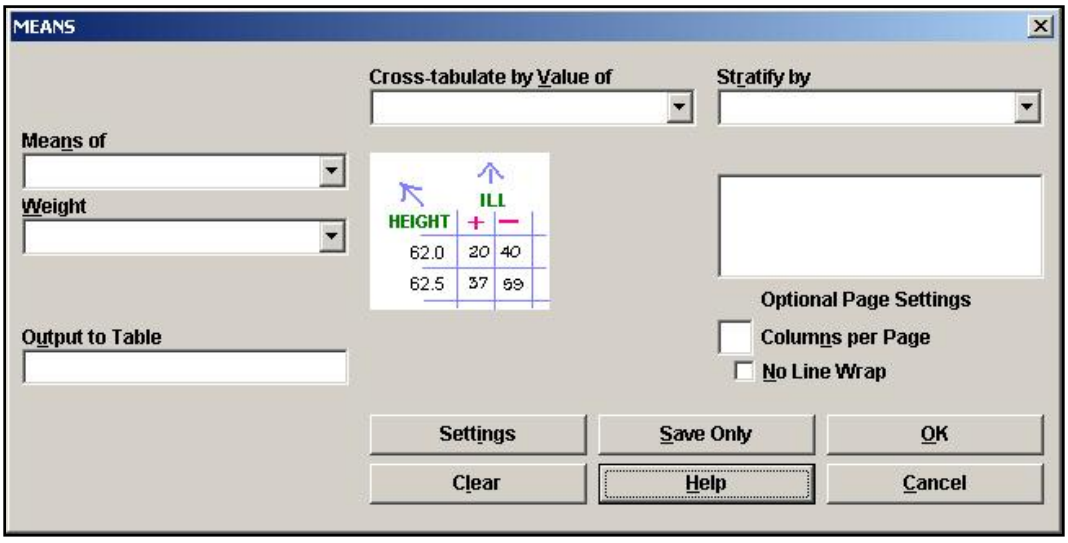
One of the items to include in the funding request is the simple means of the number of days children missed school because of asthma or an asthma-related condition. The MEANS command will provide the frequency and the means results.

The MEANS command can be used to obtain an average for a continuous variable. The mean of a yes-no variable is the proportion of respondents answering yes. The MEANS command has two formats.

If only one variable is supplied, the program produces a table similar to one produced by FREQUENCIES, plus descriptive statistics. If two variables are supplied, the first is a numeric variable containing data to be analyzed. The second indicates how groups will be distinguished. The output of this format is a table similar to one produced by TABLES, plus descriptive statistics of the numeric variable for each value of the group variable.

Let's see what they Mean!

1. From the Command Tree Statistics folder, click **Means**. The MEANS dialog box opens.



2. From the Means Of drop-down menu, select the variable **MissDays**.

3. Click **OK**. The Output window populates with the results. The table results show a breakdown of missed days aggregated per answer.

How many days of school has your child missed due to asthma?	Frequency	Percent	Cum Percent	
0	8	1.0%	1.0%	
1	241	30.1%	31.1%	
2	233	29.1%	60.3%	
3	59	7.4%	67.6%	
4	65	8.1%	75.8%	
5	8	1.0%	76.8%	
6	17	2.1%	78.9%	
7	23	2.9%	81.8%	
8	29	3.6%	85.4%	
9	24	3.0%	88.4%	

4. Scroll down to see the Mean and Standard Deviation results.

Obs	Total	Mean	Variance	Std Dev
800	3014.0000	3.7675	13.5379	3.6794

Minimum	25%	Median	75%	Maximum	Mode
0.0000	1.0000	2.0000	4.0000	15.0000	1.0000

- Variance - A measure of the dispersion shown by a set of observations.
- Standard deviation - A statistical summary of how dispersed the values of a variable are around its mean.
- Standard error (of the mean) - The standard deviation of a theoretical distribution of sample means of a variable around the true population mean of that variable.
- See Appendix H: Glossary for more definitions, or visit: (http://www.cdc.gov/excite/library/glossary.htm).

Using the Means Command with a Cross-Tabulated Variable

O. Which school had the highest average of missed days?

To determine which school had the highest average of missed days, you need to compute the means of missed days by school. You will use a cross tabulation variable to compute statistics showing the likelihood that the means of the groups are equal.

Now we'll see what you Mean.

1. Click **Means**. The MEANS dialog box opens.
2. From the Means of Variable drop down list, select **MissDays**.
3. From the Cross-Tabulate by Value of Variable drop down list, select **SchoolNum**.
4. Click **OK**. Results appear in the Output window.
5. Scroll to the Descriptive Statistics. Notice the Mean is computed by school.

Descriptive Statistics for Each Value of Crosstab Variable

	Obs	Total	Mean	Variance	Std Dev
A	196	815.0000	4.1582	14.5543	3.8150
B	205	813.0000	3.9659	13.8371	3.7198
C	208	786.0000	3.7788	13.6610	3.6961
D	191	600.0000	3.1414	11.6589	3.4145

	Minimum	25%	Median	75%	Maximum	Mode
A	0.0000	1.0000	2.0000	7.0000	15.0000	1.0000
B	0.0000	1.0000	2.0000	6.0000	15.0000	1.0000
C	0.0000	1.0000	2.0000	4.0000	15.0000	2.0000
D	1.0000	1.0000	2.0000	3.0000	15.0000	1.0000

Creating a Stratified Means

P. Was the school with the highest average of missed days for males the same as the school with females?

You will use the Means command with Stratify By to group your data by gender.

Time for more Meaning!

1. Click **Means**. The MEANS dialog box opens.
2. From the Means of Variable drop-down, select **MissDays**.
3. From the Cross-Tabulate by Value of Variable drop-down, select **SchoolNum**.
4. From the Stratify by Variable drop-down, select **Gender**.
5. Click **OK**. Results appear in the Output window.
   - Two tables and two sets of descriptive statistics are created, one each for female and male.
6. Scroll through the results.

Using Select with the Means Command

Q. Is the average number of missed days higher among those who answered yes to asthma than the overall average? Refer to Question N.

You will use the SELECT command to create subsets of data and then use the MEANS command to create statistics.

Time to SELECT what you Mean?

1. Click **Select**. The SELECT dialog box opens. Use the SELECT command to display only those students who answered yes to asthma.

2. From the Available Variables drop-down, select **Asthma**.

3. Use the Functions and Operators buttons to create the code Asthma=(+) in the Select Criteria field.

4. Click **OK**. You should have 91 records.

5. Create a Means of Missdays.

6. Click **Cancel Select** to return to the original 800 records.

Skills Review

R. What is the average age of the student population?

The age field for the imported data has not been populated. You need to assign a code to populate this field. The variable age did not exist in the MS Excel data that you merged. Since you have a date of birth, you can compute the age by using the YEARS function as you did in the Check Code lesson.

Got all that? Proceed.

From the Command Tree Variables folder, click **Assign**. The ASSIGN dialog box opens.

From the Assign Variables drop-down, select **Age**.

Type your code in the =Expression field. This is the same code you used to create the age field Check Code in MakeView.

Use the Years function and then create your mathematical code: Survey date minus date of birth variable.

YEARS(DOB, 10/25/2005)

Click **OK**. The code appears in the Program Editor.

List the Age variable in a grid.

All the once missing fields should now contain ages.

Create a Means of Age.

Creating a Table

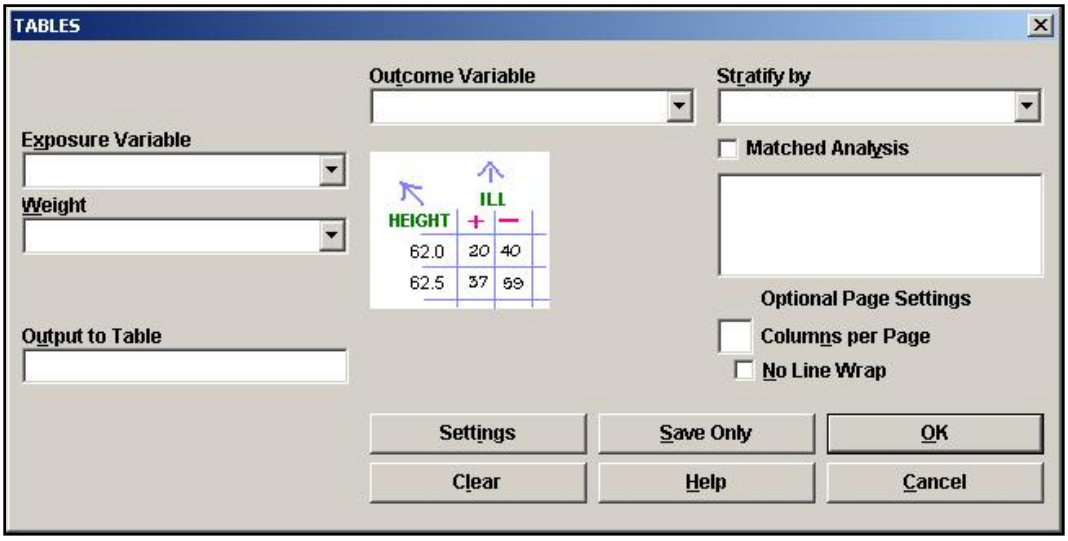
S. Does one school have a higher number of students with asthma?

T. Which school has the highest number of students with wheezing?

You will use the Tables command to examine the relationship between two or more categorical values. You want to see if there may be a connection between school location and asthmatic conditions. You also want to see what kinds of results you can get with the 800 student records available as a primary data source.

This probably won't as exciting as the French Connection!

1. From the Command Tree Statistics folder, click **Tables**. The Tables dialog box opens.



2. From the Exposure Variable drop-down, select **SchoolNum**.

3. From the Outcome Variable drop-down, select **Asthma**.

4. Click **OK**. Results appear in the Output window.

ASTHMA			
School ID #	**Yes**	**No**	**TOTAL**
A	29	167	196
Row %	14.8	85.2	100.0
Col %	31.9	23.6	24.5
B	25	180	205
Row %	12.2	87.8	100.0
Col %	27.5	25.4	25.6
C	21	187	208
Row %	10.1	89.9	100.0
Col %	23.1	26.4	26.0
D	16	175	191
Row %	8.4	91.6	100.0
Col %	17.6	24.7	23.9
TOTAL	91	709	800
Row %	11.4	88.6	100.0
Col %	100.0	100.0	100.0

5. Create the same table, but select **Wheezing** as the Outcome Variable.

Creating a 2x2 Table

U. How many students had both asthma and wheezing?

V. What are the odds that you have the condition wheezing if you have asthma?

For 2x2 tables, the Tables command produces odds ratios and risk ratios. You want to determine if a correlation exists between the children who have asthma and those with wheezing.

Let's table the matter.

1. Click **Tables**. The TABLES dialog box opens.

2. Select the Exposure Variable of **Asthma**.

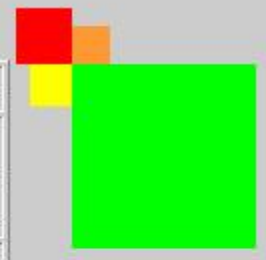
3. Select the Outcome Variable of **Wheezing**.

4. Click **OK**.

Forward

WHEEZING

Asthma	Yes	No	TOTAL
Yes	63	28	91
Row %	69.2	30.8	100.0
Col %	63.6	4.0	11.4
No	36	673	709
Row %	5.1	94.9	100.0
Col %	36.4	96.0	88.6
TOTAL	99	701	800
Row %	12.4	87.6	100.0
Col %	100.0	100.0	100.0



- Notice the graphic representation of your results.
- A 2x2 table was created because each variable had a yes or no answer.
- The largest square represents the children who answered no to both wheezing and asthma.

5. Scroll down to view the Single Table Analysis.

Single Table Analysis

	Point Estimate	95% Confidence Interval Lower	Upper	
PARAMETERS: Odds-based				
Odds Ratio (cross product)	42.0625	24.0911	73.4400	(T)
Odds Ratio (MLE)	41.5616	23.9981	73.5508	(M)
		23.2358	76.3616	(F)
PARAMETERS: Risk-based				
Risk Ratio (RR)	13.6346	9.6419	19.2806	(T)
Risk Difference (RD%)	64.1532	54.5335	73.7729	(T)

(T=Taylor series; C=Cornfield; M=Mid-P; F=Fisher Exact)

STATISTICAL TESTS	Chi-square	1-tailed p	2-tailed p
Chi square - uncorrected	306.0983		0.0000000000
Chi square - Mantel-Haenszel	305.7157		0.0000000000
Chi square - corrected (Yates)	300.2107		0.0000000000
Mid-p exact		0.0000000000	
Fisher exact		0.0000000000	

- Odds Ratio - A measure of association that quantifies the relationship between an exposure and health outcome from a comparative study Also known as the cross-product ratio.
- Risk Ratio - A comparison of the risk of ahealth-related event (e.g., disease or death in two groups).
- Chi Square - A test of statistical significance that is used to determine how likely it is that an observed association between an exposure and a disease could have occurred because chance alone, if the exposure was not actually related to the disease.
- For more information on data sources for definitions and links to more information on interpreting statistics, see the Appendix H: Glossary or Appendix I.

Skills Review

Complete the following:

W. Is the odds ratio for bronchitis and Reactive Airway Disease lower or higher than asthma and wheezing? Refer to Question V.

Hint: Create a 2x2 table for Bronchitis Exposure and RAD Outcome.

X. How many children with the condition wheezing were prescribed medication? From those results, what is the risk ratio?

Hint: Create a 2x2 table for Medication Exposure and Wheezing Outcome.

Y. Is the risk ratio for being prescribed medication for the condition asthma higher or lower than being prescribed medication for wheezing? Compare the results to Question X.

Hint: Create a 2x2 table for Medication Exposure and Asthma Outcome.

Z. Are the risk ratios for asthma and wheezing higher for children less than 10 years of age than for the survey population as a whole? Compare to the results from Question V.

Select the children who are less than 10 years of age. (Hint: Select Age<10)

Make sure that the Age field is calculated for all records. If you have used the READ command to bring in a new file or program, or closed Analysis, then you must calculate the Age field as you did for the MEANS exercise.

Create a table for Asthma Exposure and Wheezing Outcome.

Click **Cancel Select**.

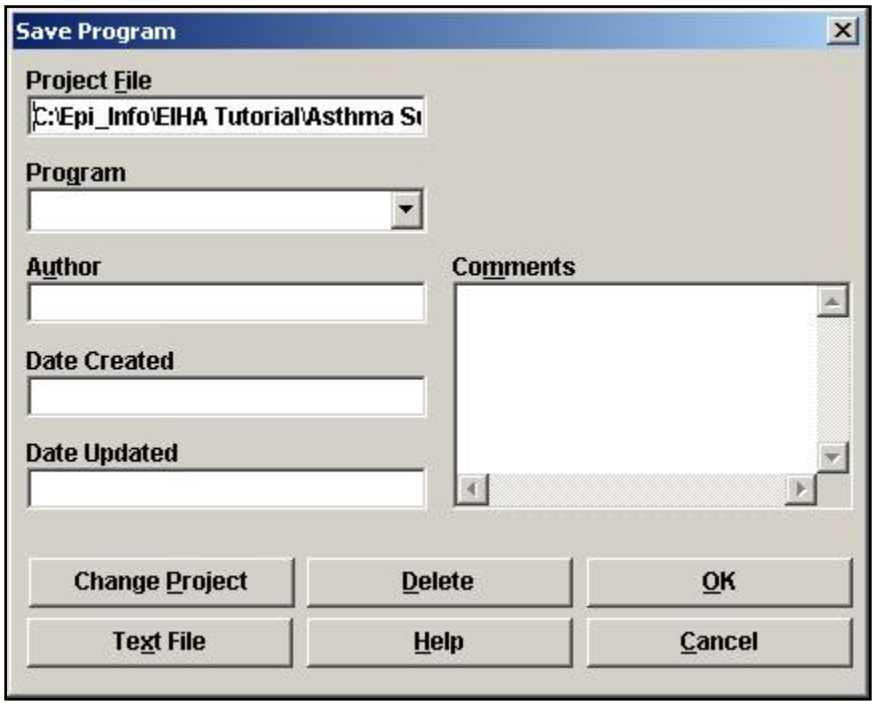
Saving a Program

Notice the large amount of code in the Program Editor. In order to re-create the same statistics after data has been added, you must save the program or code. This enables you to run these same statistics on the new records.

Analysis automatically saves the last program created in the program editor in a file called LastPGM. This file is created and saved in the Epi_Info folder and gets overwritten each time a new PGM is created and the application is closed. You want to create your own program based on your data.

Instead of save the whales, it's save the program!

1. From the Program Editor, click **File>Save**. The Save Program dialog box opens.



   - The Project File field contains the information about where you will be saving your file. Leave this information set to the EIHA Tutorial folder.

2. In the Program Field, type **Lesson 5**.

3. In the Author field, type **your initials**.

4. In Comments, type **EIHA**.

5. Click **OK**.

   - The date created and date updated fields will be populated with new information each time you open the program.

6. Click **Exit** to close Analysis.

Test the Saved Program

Time to put it to the test.

1. Open **Analysis**.

2. From the Program Editor, click **File>Open**. The READ dialog box opens.

3. From the Program drop-down, select **Lesson 5**.

4. Click **OK**. Your saved program appears in the Program Editor.

5. Select **Edit>Program Beginning**.

6. Click **Run This Command**. The first command inside your program runs. The first command was Read, so your 800 records appear. If you had updated records, the new information would be included in the results.

   - You can run any previous commands by placing your cursor in the command and clicking **Run This Command**.

7. Click **Run** to run the entire program.

   - The program you saved is part of your project. To keep it updated, you must save it. To access the LASTPGM file, click the **Text File** button from the Save Program dialog box and locate the file in the Epi_Info folder.

   - Do not close Analysis if you plan to complete Lesson 6 at this time.

Lesson Complete!

WHAT YOU LEARNED

How to:

- Use the DISPLAY command to view variable types in your project.
- Merge an MS Excel file into an Epi Info™ project.
- Use the FREQ command to see how your data is distributed.
- Use the MEANS command to see if one set of data is equal to the sum of the data divided by the number of items in the data set.
- Use the TABLES command to see if risk factors are linked to outcomes.
- Save Program (.PGM) files for use with later data tables or updated data.

Analysis: Epi Graph

SUMMARY

This lesson introduces you to the Epi Graph module of Analysis. You will create and customize several types of graphs using the survey data including, bar, pie, and line. You will also export a graph as an image. Results created in Epi Graph can be used in future reports to present your findings as graphic representations.

Based on the survey data, many types of graphs could be produced to illustrate the need for funding. You will create graphs to illustrate age groups, missed days by gender, and the worst month for asthmatic conditions by school. For the asthma funding report, you will create a pie chart to illustrate which month is the worst month for asthmatic conditions based on data from the children surveyed.

You must have the Lesson5.PGM open and run prior to completing this lesson because the code created in the previous lesson affects these graphs.

Length of time to complete: 2 hours Beginner

Graph Types

Overview of the types of graphs created in this lesson and how to use the GRAPH dialog box to make graph type selections.

Creating a Bar Graph

Use the GRAPH command to create a bar graph for the Age variable.

Creating a Pie Graph

Use the GRAPH command and the Customization dialog box to create a Pie Graph of the worst month for asthma. This graph will be exported and used in Lesson 10 Epi Report.

Exporting a Graph

Use Epi Graph to export a graph as an image for use in future reports and presentations.

Creating a Rotated Bar Graph

Use the GRAPH command to create a rotated bar graph of missed days of school by gender.

Plotting a New Graph Type

Use the Customization dialog box to change a bar graph into a line graph.

Creating a Stacked Bar Graph

Use the GRAPH command to create a stacked bar graph of the worst month for asthmatic conditions by surveyed school.

Skills Review

Series of five review questions based on the reading and hands-on activities in this lesson.

BEFORE YOU BEGIN

Complete Lesson 4 Analysis: Basics.

Complete Lesson 5 Analysis: Creating Statistics.

WHAT YOU NEED

Asthma Survey 2005.MDB

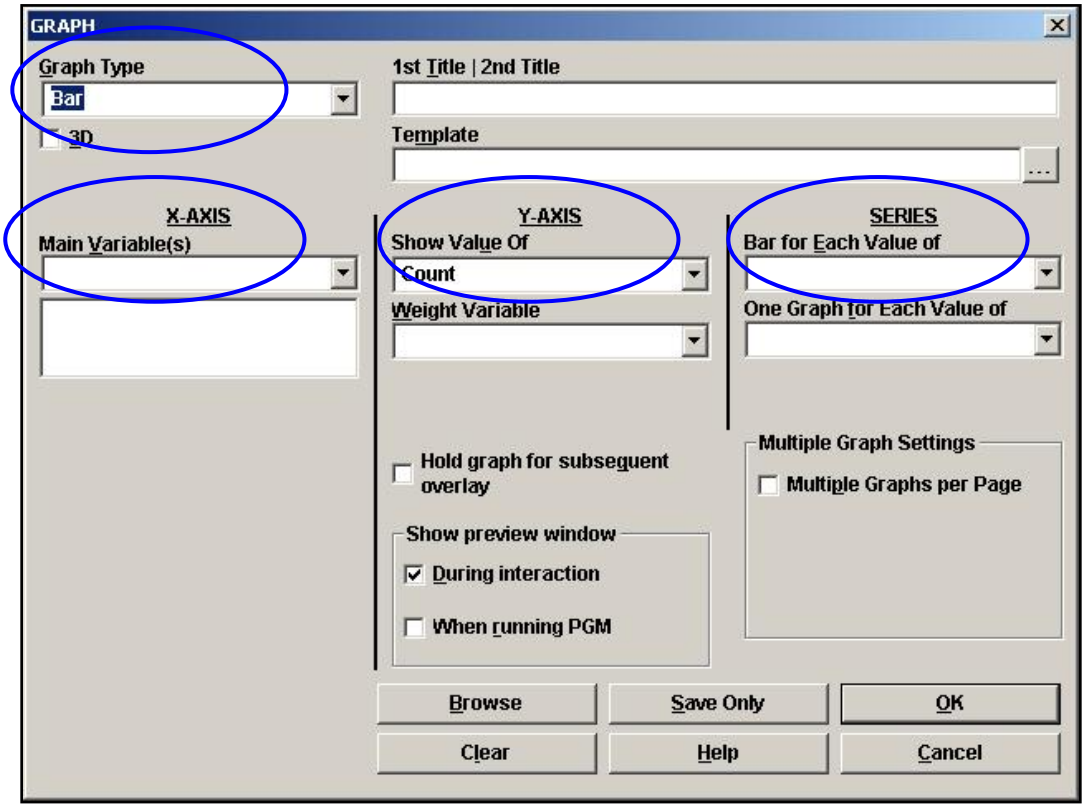
Lesson 5.PGM

FIVE GOALS

- Create a bar graph containing the ages of the children in the survey population.
- Create a pie graph of the worst month for asthmatic conditions and export it to be used in a future lesson.
- Create a rotated bar graph of missed days of school by gender.
- Create a stacked bar graph of worst month for asthmatic conditions by school.
- Use the Customization dialog box to plot new graph types and edit your graphs once created.

Graph Types

From the Analysis module, you will use the GRAPH command to create graphical representations of your data. From the Graph dialog box, you will select a Graph Type, an X-, Y-, and Series variable by using the corresponding drop-down boxes.



The following graph types will be created in this lesson:

BAR

Use to group data together according to categories and display a bar for each value in that category. Bar lengths represent quantities in a set of data. Bars are color-coded according to the series represented and displayed vertically.

ROTATED BAR

Display horizontal bars representing the count or weight of each value of the main variable(s).

STACKED BAR

Created in the same manner as non-stacked bar graphs, except that the Y values from each series accumulates. Bars are stacked on the end of the previous bar with each separate series color-coded.

AREA

Display data sets on X–Y axis with the area under the line filled.

SPLINE

Have values indicated by points connected with a smooth curve.

POINT

Contain data indicated by a marker for each data point.

LINE

Use to determine trends and cyclic variation based on interaction of data elements. Data points are plotted on the graph and connected with lines. Line graphs can display large amounts of data that varies along some accepted sequence (e.g., time, temperature, price etc.). Line graphs with multiple lines can be used to display several variables to conserve space and to facilitate comparison of more than one dependent variable on a single independent scale. They must have numeric variables for both X- and Y-axes and the X variable values are treated as continuous. To generate a line graph utilizing categorical X variables, use the Bar graph command and customize the graph style to Line.

PIE

A circular graph used for proportional assessment by comparing data elements (e.g., percentages or counts against other data elements and against the sum of the data).

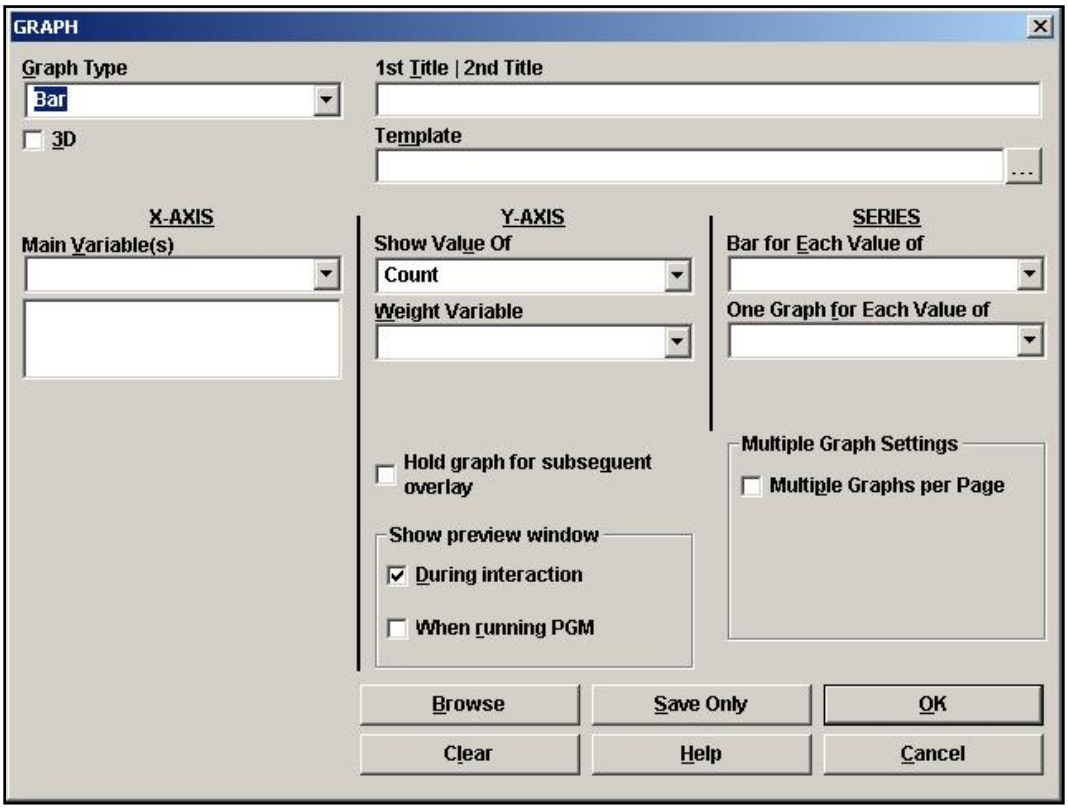
Since code created in the previous lesson affects these graphs, you must have the Lesson5.PGM open and run prior to completing this lesson.

Creating a Bar Graph

Use to make comparisons of single events and the differences between them. To visually present the age distribution of the survey population, you will create a bar graph of the Age variable.

This will be good, Bar none.

1. From the Command Tree Statistics folder, click **Graph**. The Graph dialog box opens.



- The default graph selection is Bar.

2. From the Main Variable drop-down, select the X-Axis Value of **Age**.

3. Click **OK**. Epi Graph opens.

- You may need to click the **flashing Epi Graph** icon from the task bar.

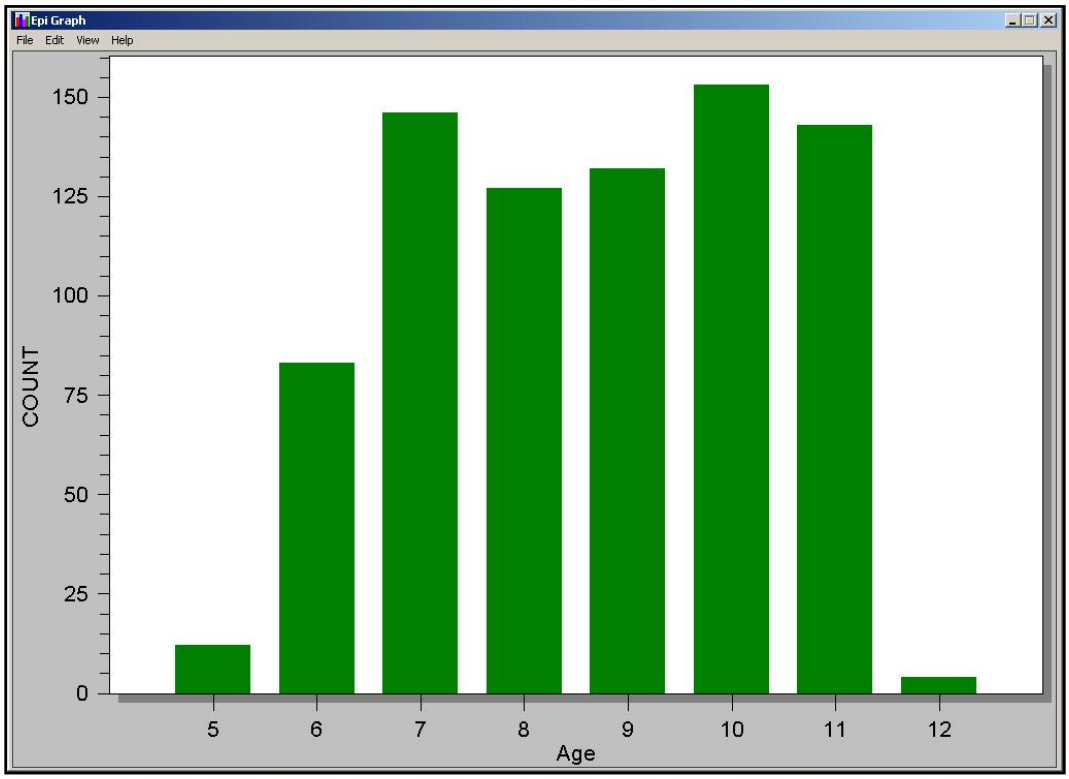
4. View the bar graph.

- Be sure that the Age field is calculated for all records. If you have READ in a new file or program or closed Analysis, then you will have to calculate the Age values. Age

fields should be coded if you have the Lesson5.PGM open and run from the Program Editor.

5. Click the close **X**. The bar graph is now embedded in the Output window and part of your output file.

Your graph should look like the following:



Creating and Customizing a Pie Graph

You will now create a Pie graph that illustrates the worst month for asthmatic conditions. In the survey, parents were asked which month presented the most breathing problems for their child based on their symptoms. To illustrate your findings, you will create a pie graph on the WMonth variable.

Will it be apple, blueberry, or pumpkin?

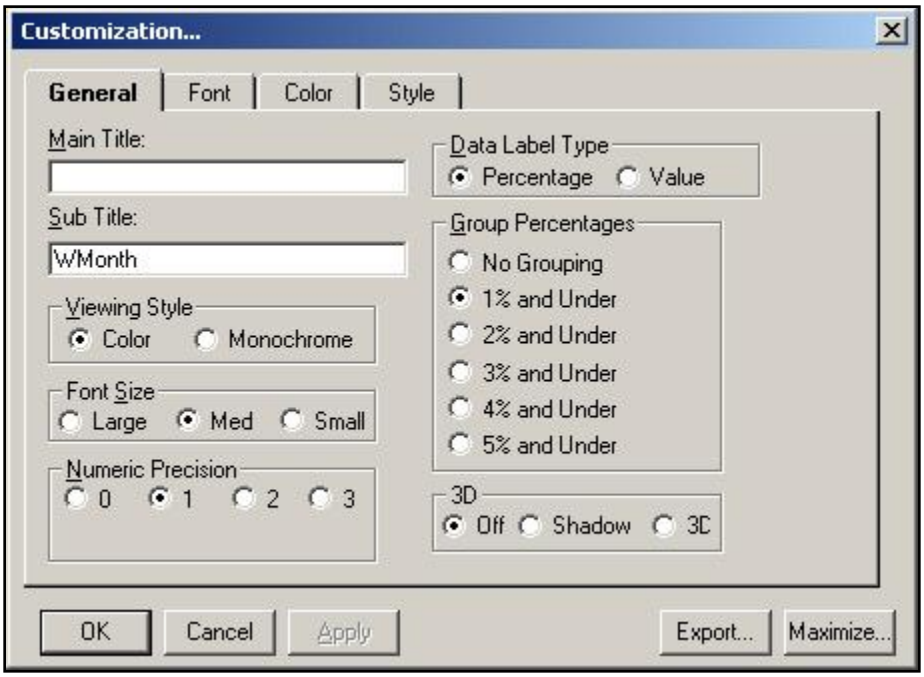
1. Click **Graph**. The Graph dialog box opens.

2. From the Graph Type drop-down, select **Pie**.

3. From the Main Variable drop-down, select **WMonth.**

4. Click **OK**. The WMonth Pie Graph appears.

5. From the main navigation bar, select **View>Customization,** or use your mouse to right click inside the graph and select **Customization Dialog** from the pop-up menu. The Customization dialog box opens.



   - Each tab inside the Customization dialog box allows you to change a part of your graph and how it is viewed on the page.

6. From the General tab, in the Main Title field, type **Worst Month for Asthmatic Conditions**.

7. Delete the text **Wmonth** from the SubTitle field.

8. Select the **3D** radio button.

9. Click **Apply**. The graph updates.

10. Click the **Font** tab.

11. Select a new font for your title.

12. Make the title **Bold**.

13. Click **Apply**. The title font updates.

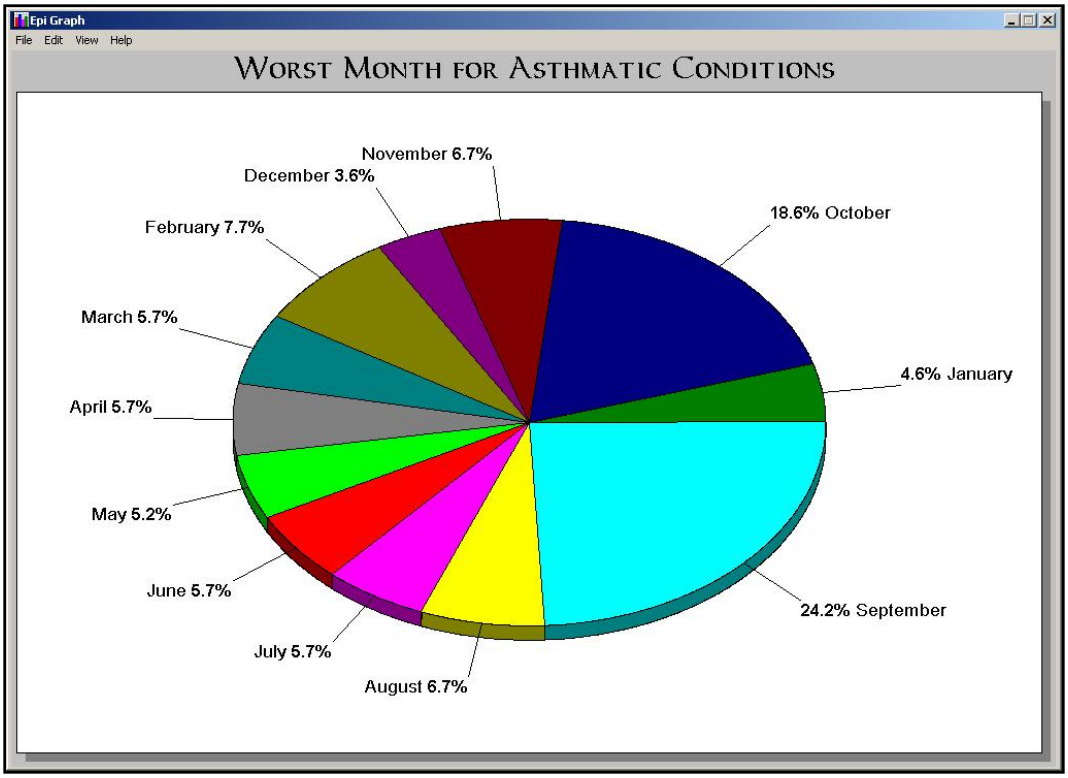
14. Click **OK** to close the Customization dialog box.

15. From your graph, left click on the number **9**. The Point Label dialog box opens.

16. Type the month name **September. C**lick **OK**.

17. Name each Point Label for the month with which it corresponds.

- Your pie graph should look similar to the following:

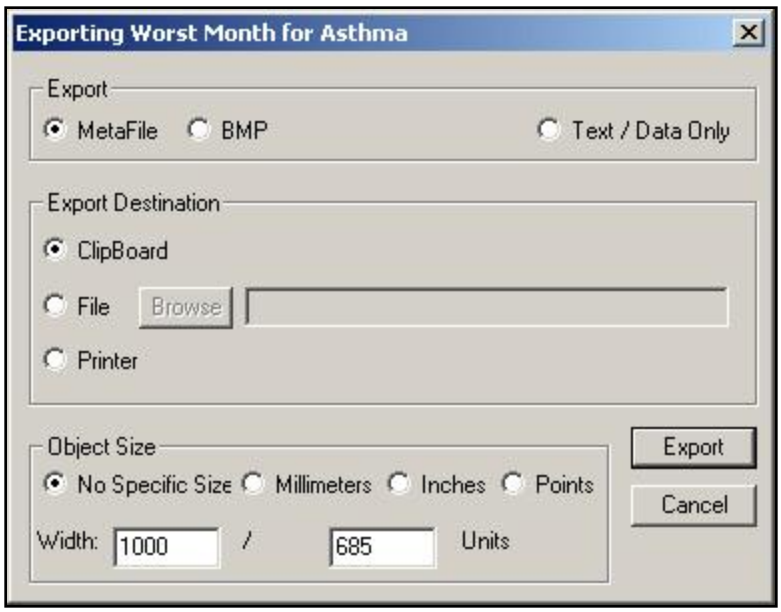


Exporting a Graph

Now you will export the WMonth pie graph for use in another program. Graphs can be exported for use in other programs (e.g., MS Word, PowerPoint, or Epi Report). You will be using the WMonth pie graph in Lesson 10 Epi Report when you create your asthma initiative report.

No importing allowed here.

1. From the Epi Graph main navigation menu, select **File>Export**. The Exporting dialog box opens.



2. From the Export section, select **BMP** to create a bitmap file.
3. From Export Destination, select **File**.
4. Click **Browse**. The Save As dialog box opens.
5. Locate the EIHA Tutorial folder. Name your file **WMonthPieGraph**.
6. Click **Save**.
7. From the Exporting dialog box, click **Export**.
8. Close **Epi Graph**.
    - You will use the exported graph in the Epi Report lesson.

Creating a Rotated Bar Graph

In this section, you want to illustrate the number of missed days of school by gender. This will help determine whether there is a difference in how much school is missed based on the gender of the survey population.

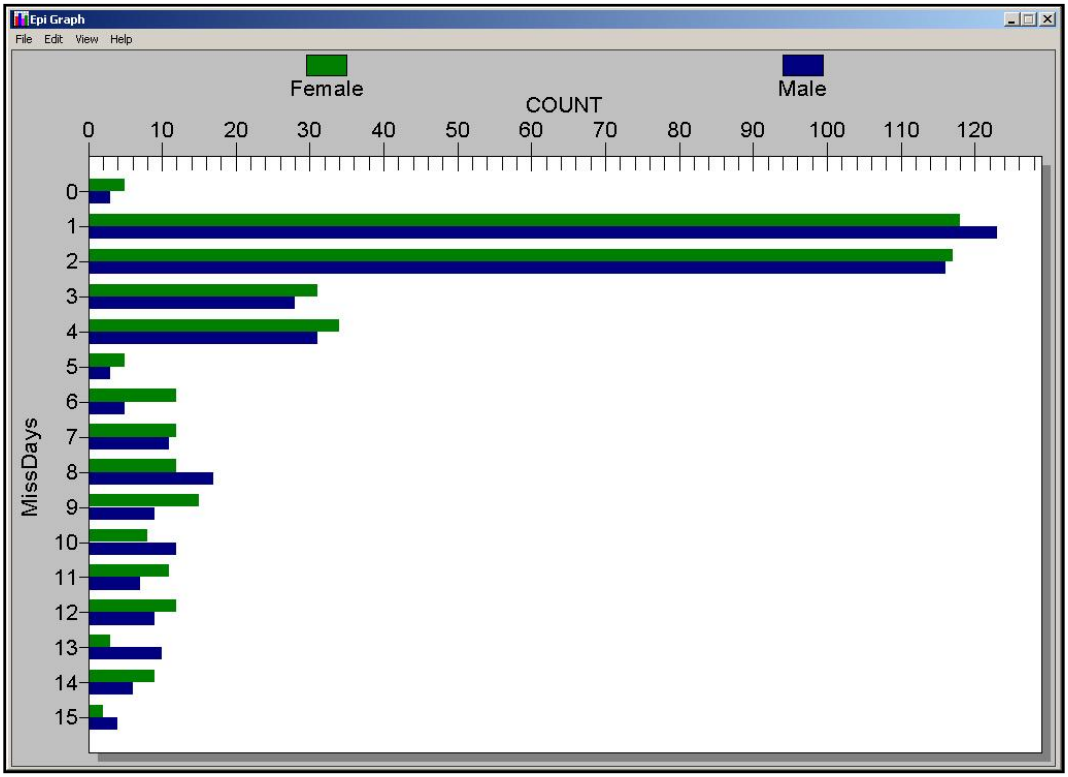
Time to rotate, but not on the axis.

1. Click **Graph**. The GRAPH dialog box opens.
2. From the Graph Type drop-down menu, select **Rotated Bar**.
3. From the Main Variable drop-down, select **MissDays**.

4. From the SERIES Bar for each Value drop-down, select **Gender**.

5. Click **OK**. The Epi Graph opens.

   - Your Rotated Bar graph should look similar to this example:



6. From the Epi Graph navigation bar, select **View>Customization.** The Customization dialog box opens.

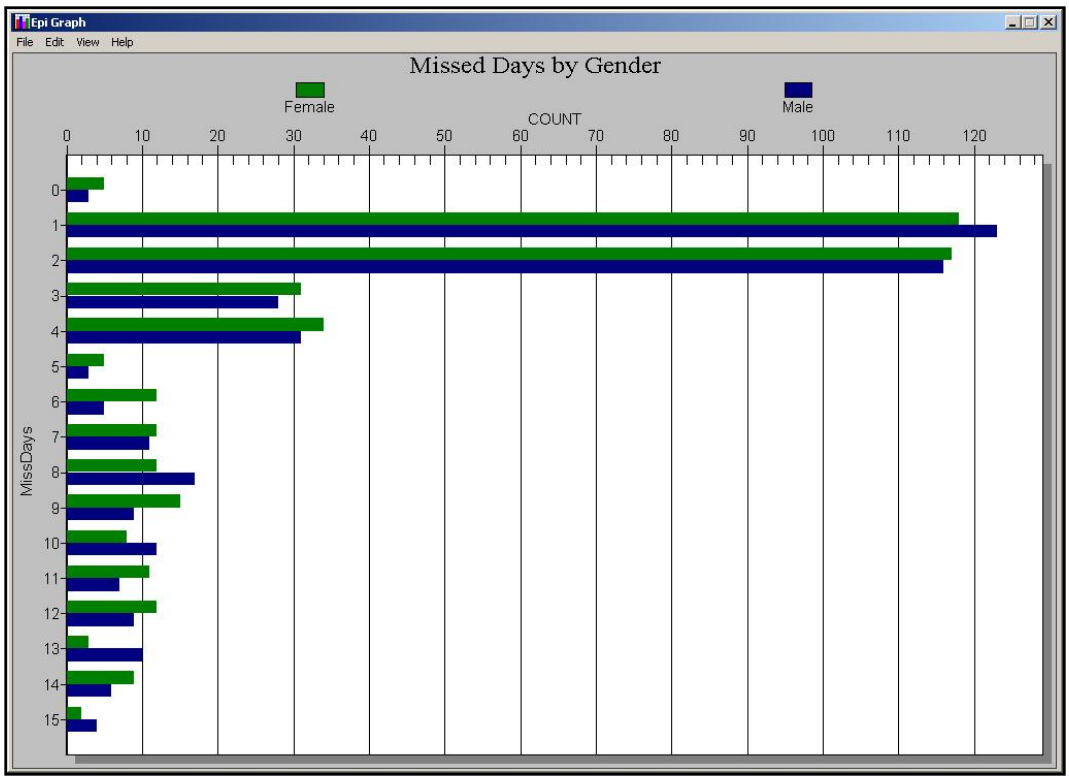
7. From the General tab, type the Main Title **Missed Days by Gender**.

8. From the General tab, select the **Font Size Small** radio button.

9. From the General tab in the Grid Lines section, select **Y**.

10. Click **Apply**.

11. Click **OK**. Do NOT close Epi Graph.


Epi Graph
File Edit View Help
Missed Days by Gender
Female
Male
COUNT
0 10 20 30 40 50 60 70 80 90 100 110 120
MissDays
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

Plotting a New Graph Type

You can change a bar graph with categorical X values into a line graph by using the Customization dialog box. You will use the rotated bar graph from the previous lesson and view it in a different format.

A plot? Now it gets interesting!

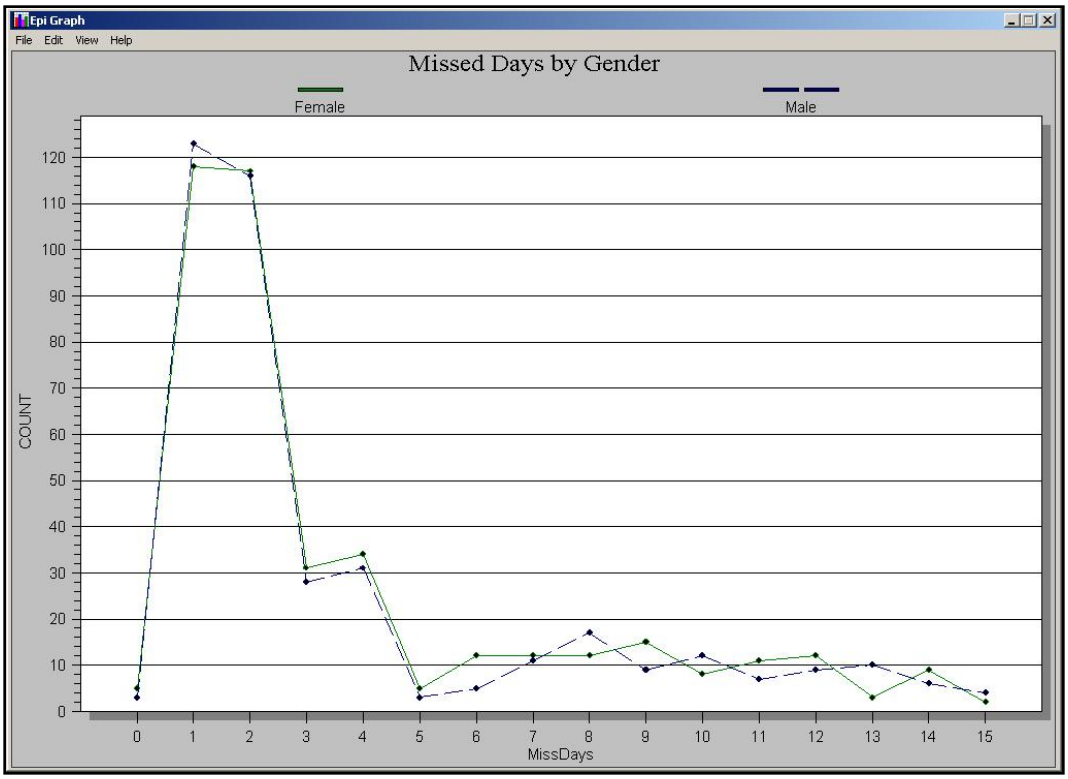
1. From the Epi Graph navigation bar, select **View>Customization**. The Customization dialog box opens.

2. Click the **Plot** tab.

3. From the Plot Style selection, select **Line.**

4. Select the **Mark Data Points** check box.

5. Click **Apply**. The rotated bar information transforms into a line graph.



6. Click the **Points** tab.

7. Select the **Point Label Orientation, Slanted**.

8. Click **Apply**.

9. Click the **Style** tab.

10. Select **Female**

11. From the Point Type drop-down, select the **solid diamond**.

12. Select **Male.**

13. From the Line Type drop-down, select the **small blue dotted line** type.

14. Click the **Burgundy** color box.

15. Click **Apply**.

16. Click **OK**.

17. Close **Epi Graph**.

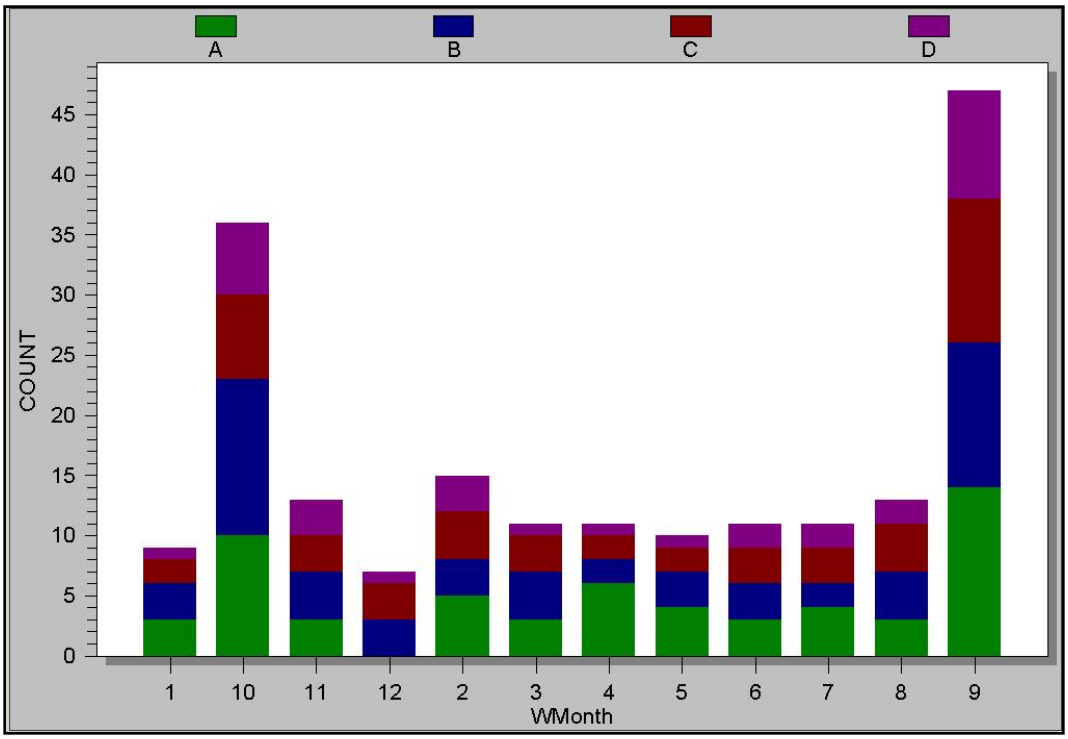
Creating a Stacked Bar

To determine if one school had more students with breathing problems in any given month, you can create a stacked bar of the WMonth variable with a series bar for each value of SchoolNum. For example, this will indicate that School A did not have any students who answered that their worst breathing month was December.

Remind you of a past job stacking boxes?

1. Click **Graph**.
2. From the Graph Type drop-down, select **Stacked Bar**.
3. From the Main Variable drop-down, select **WMonth**.
4. From the SERIES Bar for Each Value, select **SchoolNum**.
5. Click **OK**. Epi Graph opens. Notice that the months are not in sequence by month, but by the first digit of the number. In Lesson 7 Exporting Files, you will use the RECODE command to re-order the months to appear in the proper sequence.

   Your stacked bar graph should look similar to the following:



6. Right click the graph.
7. From the pop-up menu, select **Customization Dialog**.

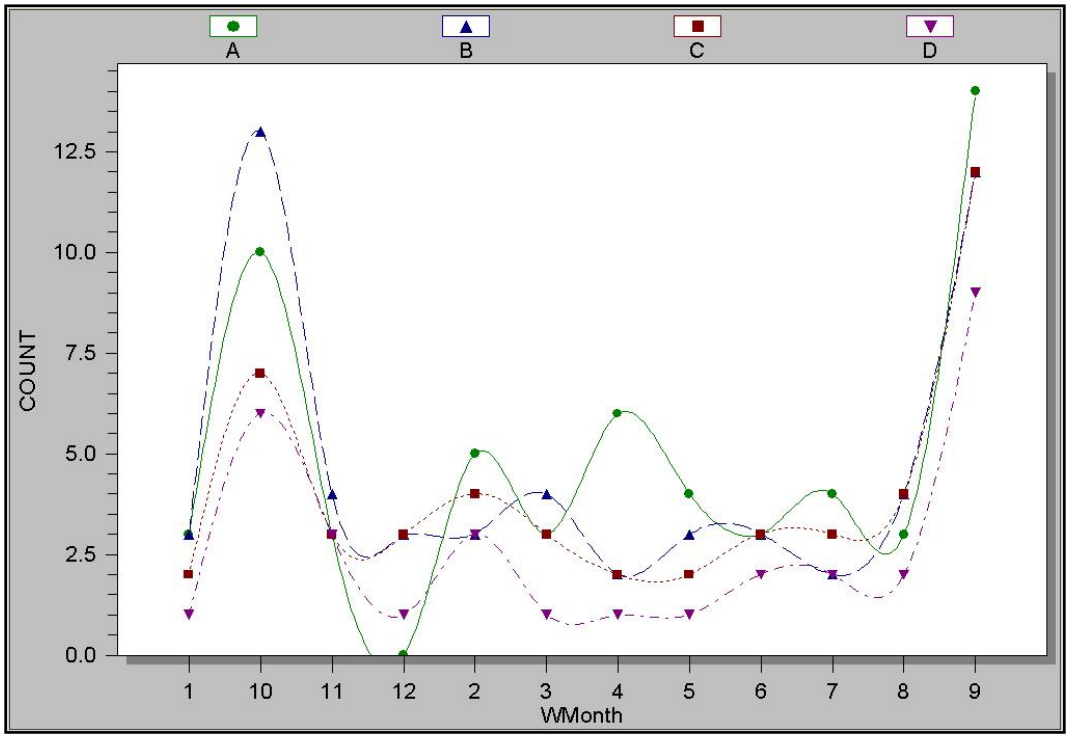
8. Click the **Plot** tab.

9. From the Plot Style section, select **Points and Spline**.

10. Click **Apply**.

11. Click **OK**.

- The same worst month information is now presented in a Points and Spline format.



12. Right click the graph.

13. From the pop-up menu, select **Customization Dialog**.

14. Click the **Plot** tab.

15. Select **Area Stacked**.

16. Click **Apply**.

17. Click **OK**.

- The same worst month information is now presented in an Area graph format.

18. Close **Epi Graph**.

🕮 Skills Review

Place a checkmark next to the correct answer for each of the following questions. Refer to *Appendix D Skills Reviews Answer Key* for further details.

1. You can use a Line Graph to determine trends and cyclic variation.

☐ True

☐ False

2. Which type of graph presents a circular proportional assessment by comparing data elements (e.g., percentages or counts against the sum of the data elements)?

☐ Line Graph

☐ Bar Graph

☐ Pie Graph

3. From Epi Graph, select the two ways you can open the Customization dialog box.

☐ Right click and select Customization Dialog from the pop-up menu

☐ Select View>Customization

☐ It automatically opens

4. Graphs can be exported for use in other programs (e.g.,MS Word, PowerPoint, or Epi Report).

☐ True

☐ False

5. You can plot a new graph type from an existing graph using the Customization dialog box.

☐ True

☐ False

Lesson Complete!

WHAT YOU LEARNED

How to:

- Create a Bar, Pie, Rotated Bar, and Stacked Bar Graph to illustrate statistical findings.
- Export a graph for use in another program, or for a presentation.
- Use the Epi Graph Customization dialog box to create graphs specific for your project, or plot a new graph type.

This page intentionally blank.

Analysis: Exporting Files

SUMMARY

This lesson focuses on ways you can export files or data from Epi Info™. You will create two new graphs and learn ways to RouteOut and Write/Export files for use in other programs. It also includes advanced coding, graphing, and an introduction to the command RECODE.

The RECODE command allows you to create groups, transform code from one system to another, change variable values, and create new values for variables. You will use RECODE to re-order the variable WMonth to appear in a school-year sequence.

You will also use the WRITE command to create a new data table that contains selected information from the survey records. This new data table will contain information selected by age and symptom. It will be used to create a line graph that maps all symptoms on one graph by age to enable you to look for trends in symptoms based on age.

Length of time to complete: 2 hours Intermediate

Managing Data

Overview of export types and commands covered in this lesson. Also describes and discusses the creation and storage of output files from the Analysis module.

HEADER, PRINTOUT, and ROUTEOUT Commands

Use to customize Output window results and to print results from the Output window.

DEFINE and RECODE

Use DEFINE to create new variables and RECODE to assign values to create a graph with re-ordered dates based on the survey variable WMonth.

WRITE Command

Use to create a new data table for specified variables.

Write and Append a Data Table

Use the WRITE command to create a new table containing specified variables. Use the new data table to create a graph.

Write to MS Excel

Use the WRITE command to export your data to an MS Excel file.

Skills Review

Series of 10 review questions based on the reading and hands-on activities in this lesson.

BEFORE YOU BEGIN

Complete Lesson 4 Analysis: Basics.

Complete Lesson 5 Analysis: Creating Statistics.

Complete Lesson 6 Analysis: Epi Graph

WHAT YOU NEED

Asthma Survey 2005.MDB

Microsoft Excel

FIVE GOALS

- Understand and be able to view output files created by Analysis.
- Use the RECODE command to define a variable.
- Use the WRITE command to place selected variables in a new data table.
- Use the format commands NUMTODATE and TXTTONUM to assign and format a date variable.
- Use the WRITE command to export to data to an MS Excel file.

Managing Data

Information gathered and analyzed in Epi Info™ can be exported to other formats and used in other programs. In this lesson, you will manage data in three ways. You will use the:

1. HEADER command to customize and then print data from the Analysis Output window.

2. ROUTEOUT command to designate a new Output file. In this lesson, you will create a new graph of Worst Month for Asthma and send it to the new file. This lesson also introduces the RECODE command.

3. WRITE command to:

   - Export and create a new data table to save a temporary value as a permanent value.

   - Export selected data to a new table and use the new table to create a line graph of symptoms by age.

   - Export the selected variables and new data to an MS Excel file.

Commands in this Lesson

HEADER

Sets up specific headings that appear in the Output window until a new file is Read into Analysis. Located in the Output folder.

ROUTEOUT

Directs the output to the named file until the process is terminated by Closeout or by exiting Analysis. RouteOut creates an HTML document that can be read by any browser. Located in the Output folder.

CLOSEOUT

Closes the current output file. Located in the Output folder.

PRINTOUT

Sends the current output file, or some other file specified by you, to the default printer. Located in the Output folder.

WRITE/EXPORT

Sends records to an output table or file in the format that you specify. Use this command to write your files to a new data type (e.g., MS Excel or a text file). Located in the Data folder.

DEFINE

Allows you to create new variables. Located in the Variables folder.

RECODE

Allows grouping of data for age and other variables, or transformation of coding from one system to another. You can use this command to change some or all values of a variable. The new values can be stored in the same or in a new variable. Recode can convert a numeric variable into a character variable or the reverse, or create a new variable based on recoded values of an existing variable. Located in the Variables folder.

ASSIGN

Stores the value of a variable, or assigns the result of mathematics to a variable. Located in the Variables folder.

GRAPH

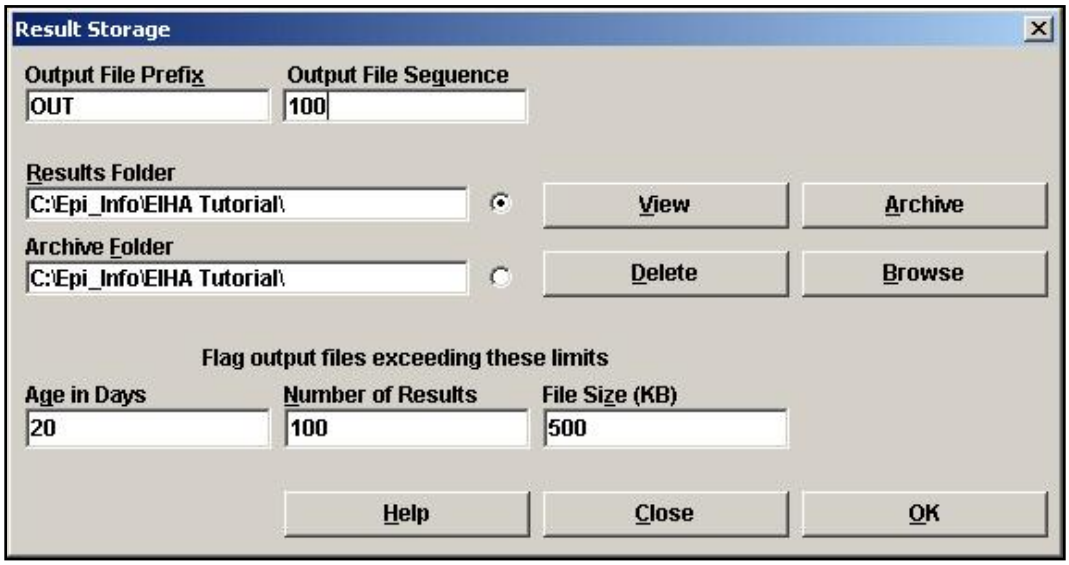
Produces various graph types based on data and selections. Located in the Statistics folder.

STORING OUTPUT

Allows a change of the default settings for storing output files. Located in the Output folder.

Output Files

You create an Output file each time you read a file into Analysis and run a command. The Output appears in the Output window and in the Output file inside your project directory. Output files are stored in the directory of the current project with a name composed of a prefix and a sequence number, the default is OUT#.htm, (i.e.,, C:\Epi_Info\EIHA Tutorial\OUT120.htm). The name, number, and destination of your Output files can be changed using the Storing Output command. From the Result Storage dialog box, you can change the file prefix, sequence number, and location of output.



The output of a single statistical procedure is called a result. Results accumulate in a given results index file called IResults.htm, located in the project directory. If several projects share the same directory, they share the same results index. Results continue to populate this file until the data source changes by READ, RELATE, SORT, SELECT, or SET PROCESS, or until a CLOSEOUT command is executed.

When Epi Info creates the Out#.htm file, it also creates a corresponding Out#.xml file. Out#.htm files are stored for reference purposes. Out#.xml files are stored for use in the Epi Report module.

When Output files are created, the name and location of the file appears in the Output window title bar as a reference. Outputs are added to the same file until the data source is changed.

When you create an Output file, you will see the location and name of the file in the title bar at the top of your window.

Viewing Output Files

Each task you complete in Analysis creates results that appear in the Output window and in the Out#.htm file. Now view the Output files you have created so far.

First in, now out.

1. Open the **Analysis** module.

2. From the Output window button bar, click **Open**. The Browse dialog box opens. It should be open to EIHA Tutorial, your destination folder,.

   - Notice the list of OUT#.htm files in your folder.

3. Select one of the OUT files.

4. Click **Open**. The file opens in your Output window.

   - This is the saved reference output you created.

   - Notice the Program Editor does not update. You are not running a command, only viewing data.

5. From the Output window button bar, click **Open**.

6. Select **IResults.htm**.

7. Click **Open**. The Results Library appears.

   - The same file can be opened by the IResults.htm file in your directory and the Results Library link inside active output windows.

   - In the Command column, you will see a list of blue hyperlinks.

   - Click a link to view the results.

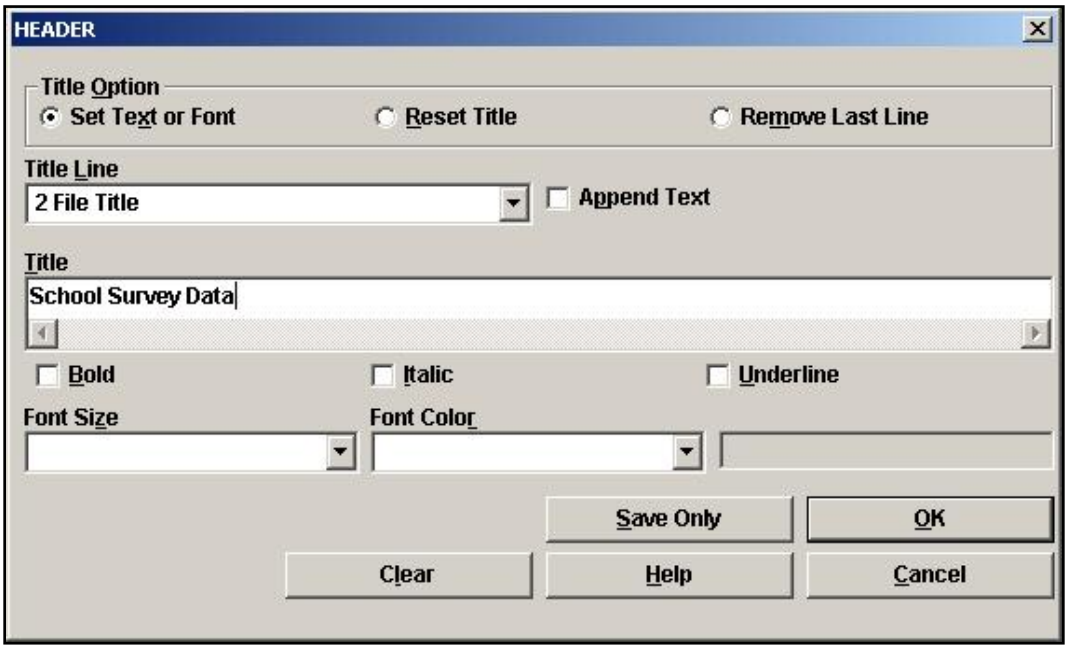
   - This is another way to view previous Output for reference.

Using the HEADER Command

You will customize the Output window using the HEADER command to label results. Customizing Analysis Headers is useful if you plan to create long lists of code that need to be labeled for specific project files, uses, or data tables.

Go to the Head of the class.

1. From the Command Tree Output folder, click **Header**. The HEADER dialog box opens.



2. From the Title Line drop-down, select **2 File Title**.

3. In the Title Field, type **School Survey Data**.

4. From the Font Size drop-down, select **+3**.

5. From the Font Color drop-down, select **Blue**.

6. Click **OK**. The Output window appears.

   - The header code appears in the Program Editor, but it is not applied until you Read/Import a file.

7. Click **Header**. The HEADER dialog box opens.

8. From the Title Line drop-down, select **3 Data Source Title**.

9. In the Title Field, type **School Year 2004–2005.**

10. Select **Bold**.

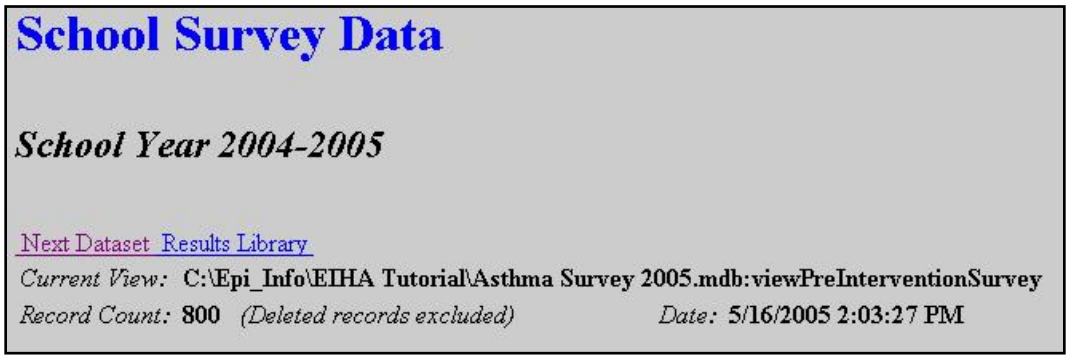
11. Select **Italic**.

12. Click **OK**.

13. Read/Import your Asthma Survey 2005.mdb. Make sure the project and the data source match. If not, click **Change Project** and select the **Asthma Survey 2005** project.

14. Click **OK**.

- Header information stays in effect until you read in a new file.
- Your Output window should look similar to this example.



Using the PrintOut Command

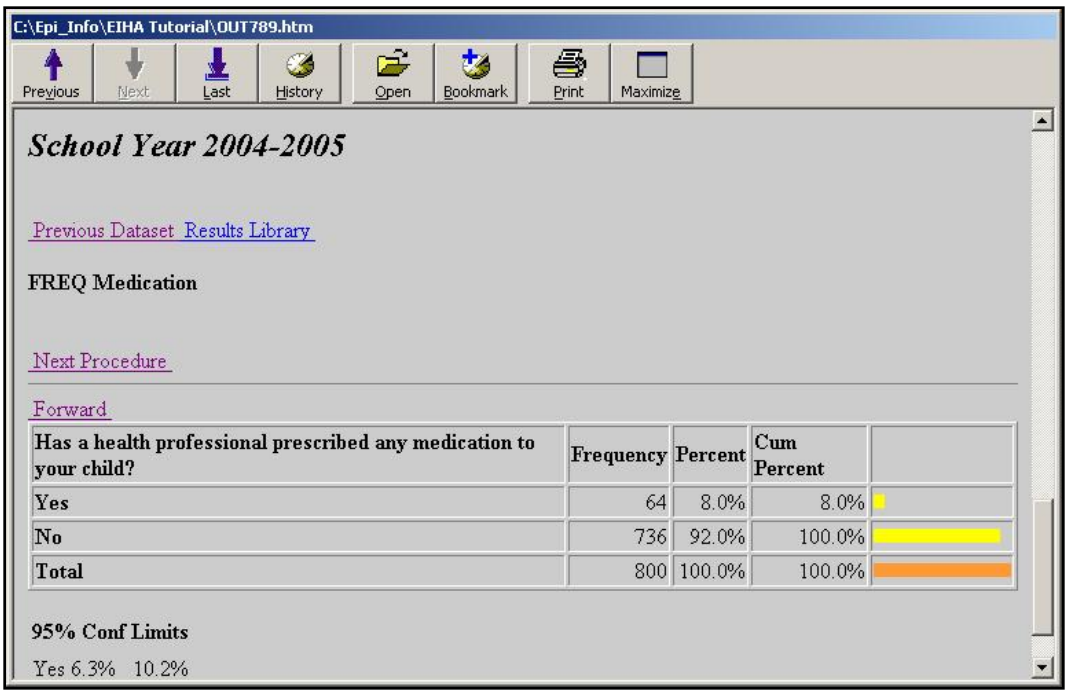
PRINTOUT can be used to print results directly from the Output window.

Time to print away!

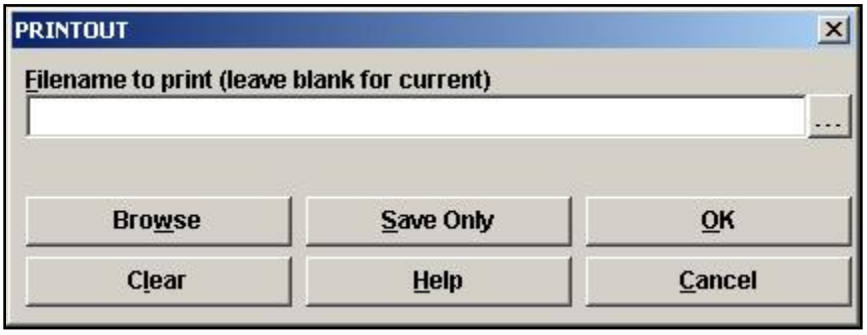
1. Create a Frequency of SchoolNum.

2. Create a Frequency of Medication.

- Notice the Data Source Title School Year 2004–2005 appears above your results.



3. From the Command Tree Output folder, click **PrintOut**. The PRINTOUT dialog box opens.



4. To print the current Output window and settings, click **OK**. The Print options window displays.

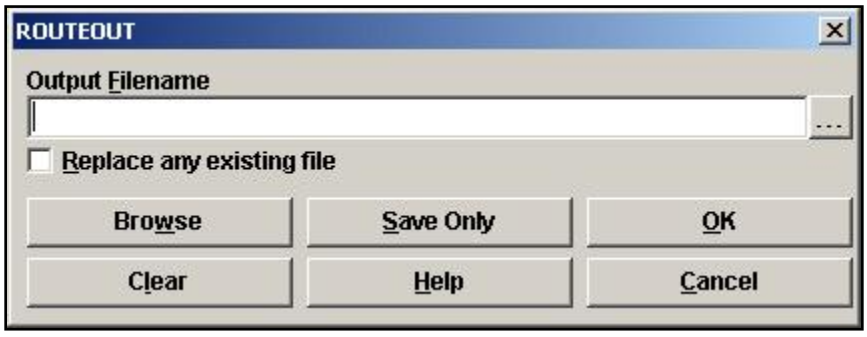
5. Click **Cancel**. If you are connected to a printer, you can click **Print** to see the results in paper form. Otherwise cancel the print command.

Routeouting Data to a New File

You will create a new graph to be included in a presentation. Rather than saving the graph as an image file, you will to route the graph and results to a new file. The ROUTEOUT command stays in effect until you use the CLOSEOUT command. ROUTEOUT will name your output and place it in the location of your choice on the computer.

Bring out the map. A route is to be created.

1. From the Command Tree Output folder, click **RouteOut**. The ROUTEOUT dialog box opens.



2. In the Output Filename field, type **School Year Graph**.

3. Click **OK**.

   - Output results will now go to the School Year Graph.htm file instead of the Out#.htm file.

DEFINE and RECODE

In an earlier lesson, you created a bar graph using the variable WMonth. You may have noticed that Epi Graph ordered the variables in sequence based on the first number, 1, 10, 12, etc. To create a graph with months occurring in school year sequence, you will use the DEFINE command to create a new variable and the RECODE command to place existing values in a new order. A program example is located in Asthma Survey 2005 Sample.MDB. The program is called MonthsRecode.PGM.

In the survey, some students selected their worst month for asthma symptoms from a comment legal field. Those who made no selection are considered Missing values. Students did not select a year in this field. You want the dates in your graph to be distributed based on the school year when the questions were asked. You will use the RECODE command to assign a year (variable name addyr) to the months and to place them in the order September 2004 to August 2005.

Once the months are recoded, you will define a new variable (NewWMonth) to indicate the new sequence of months with the new year value. When you create the final code, designate how you want the dates to appear and specify that the Legal Value fields, which were once text, will appear as numbers. This can be accomplished by using the function NUMTODATE, which transforms a set of three numbers into the year, month, day format; and allows you to order the variables by the new dates.

Your final code will say that the value of the NewWMonth variable will change from a number to a date. The date will be formatted as the year 2000 plus the addyr variable, the WMonth variable, and the first day of the month. Since WMonth is a text field, you also have to add in the code to format it as a number (TXTTONUM).

Keep in mind that the ROUTEOUT command is in effect and will send data to the new output file School Year Graph.htm. This information and the graph can then be used in your asthma report.

Make sure you have your dictionary for this one.

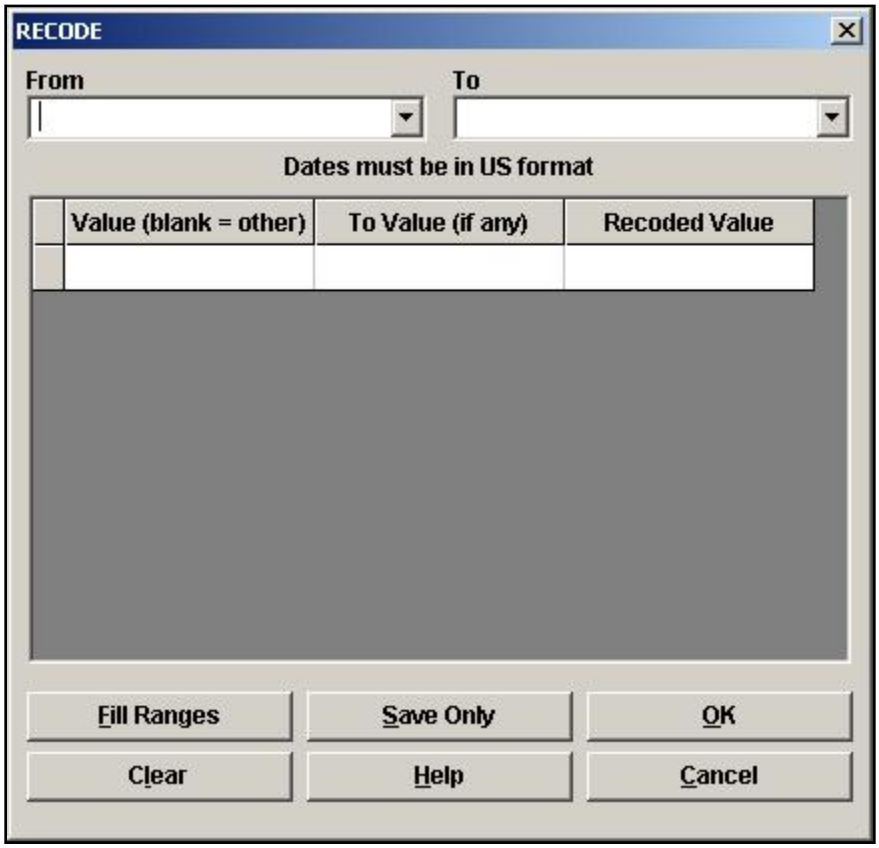
1. From the Command Tree Variables folder, click **Define**. The DEFINE dialog box opens.

2. In the Variable Name field, type **addyr**.

3. Click **OK**.

   - The data table now contains a new Standard variable. Standard variables are temporary. Values and definitions are lost at the next READ command.

4. From the Command Tree Variables folder, click **Recode**. The RECODE window opens.



5. From the From drop-down box ,select **WMonth**.

   - You are taking the values from the variable WMonth and recoding them to appear in the variable addyr.

6. From the To drop-down box, select **addyr**.

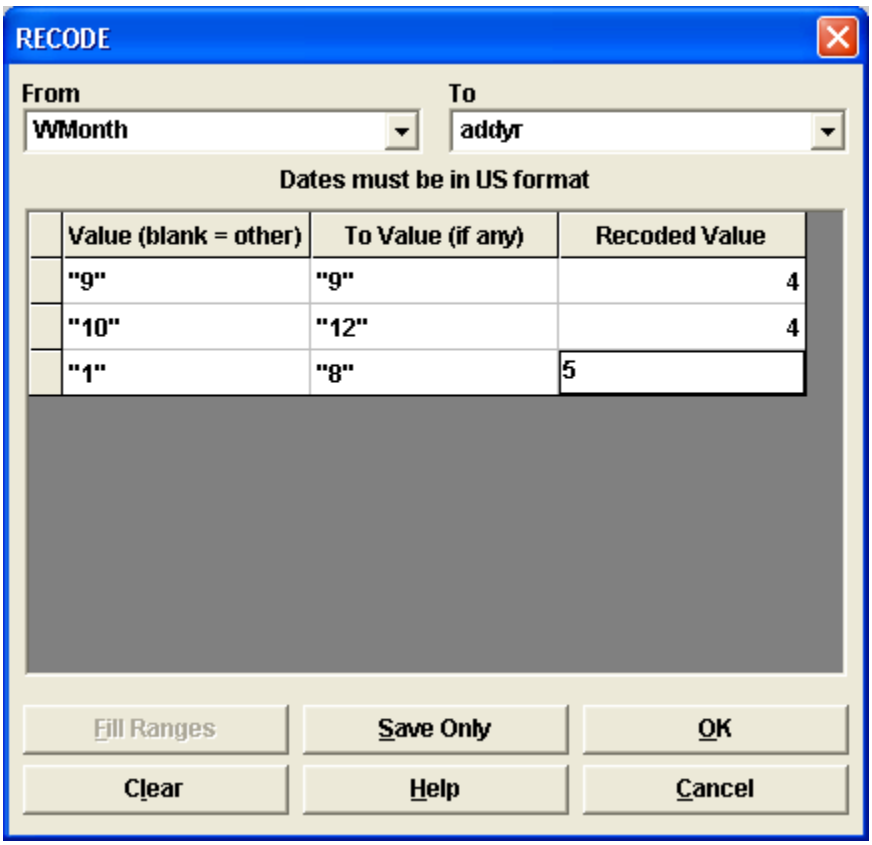
7. In the Value (blank=other) field, type **9** and press **Enter**.

   - Remember the value of 9 is September. September is a text value and text values must be enclosed in quotes.

8. In the To Value (if any) field type **9** and press **Enter**.

9. In the Recoded Value field type the number **4** and press **Enter**. A new blank entry line appears.

- You are creating the code to differentiate the year 2004 from the year 2005. The number 4 is the year you will add to 2000 to reorder the months.
- You are creating code so that September begins the list and appears in the year 2004.
- October–December appears in the year 2004.
- January–August appears in the year 2005.



10. Complete the code as follows.

- Value "10" To Value "12"=4
- Value "1" To Value "8"=5
- Any values that do not fit the pattern are considered missing.

11. Click **OK**.

Your code should appear in the Program Editor as follows:

```
RECODE WMonth TO addyr
   "9" - "9" = 4
   "10" - "12" = 4
   "1" - "8" = 5
END
```

Now define a new variable to contain the date code.

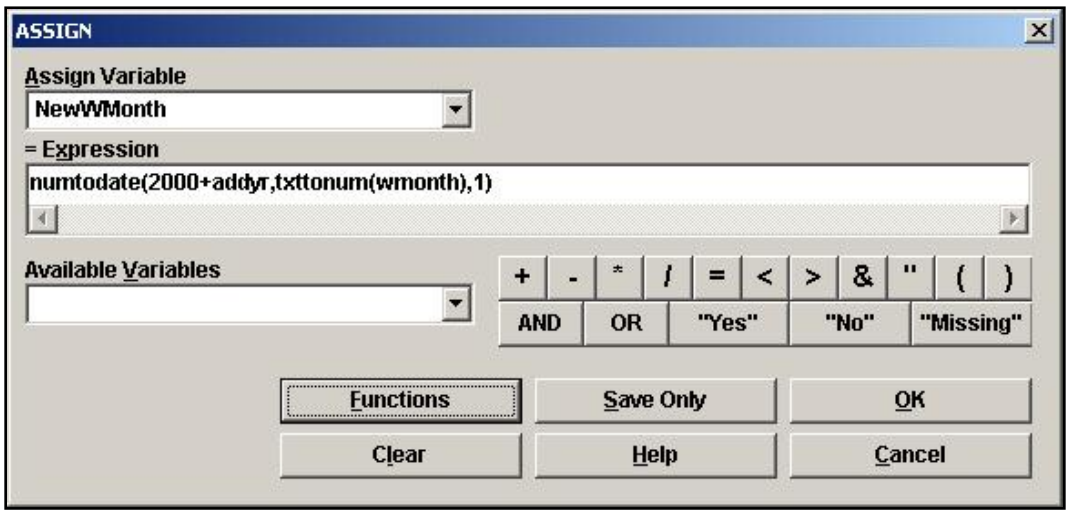
1. Click **Define**. The DEFINE dialog box opens.

2. In the Variable Name field, type **NewWMonth.**

3. Click **OK**.

Time to assign a value to the new variable.

1. From the Command Tree Variables folder, click **Assign**. The ASSIGN dialog box opens.



2. From the Assign Variable drop-down, select **NewWMonth**.

3. In the =Expression field, type **numtodate(2000+addyr,txttonum(WMonth),1***)*

   - You are assigning a date format to the variable NewWMonth. You want the year/month/day format because you are going to order the months by the year. This code says the value of the NewWMonth variable will change from a number to a date. The date will be formatted as the year 2000 plus the addyr variable, the WMonth variable, and the first day of the month. Since WMonth is a text field, you also have to add in the code to format it as a number.

4. Click **OK**. The code in the Program Editor appears as follows.

```
DEFINE NewWMonth
ASSIGN NewWMonth=numtodate(2000+addyr,txttonum(wmonth),1)
```

Now create a bar graph using the new variables and date code

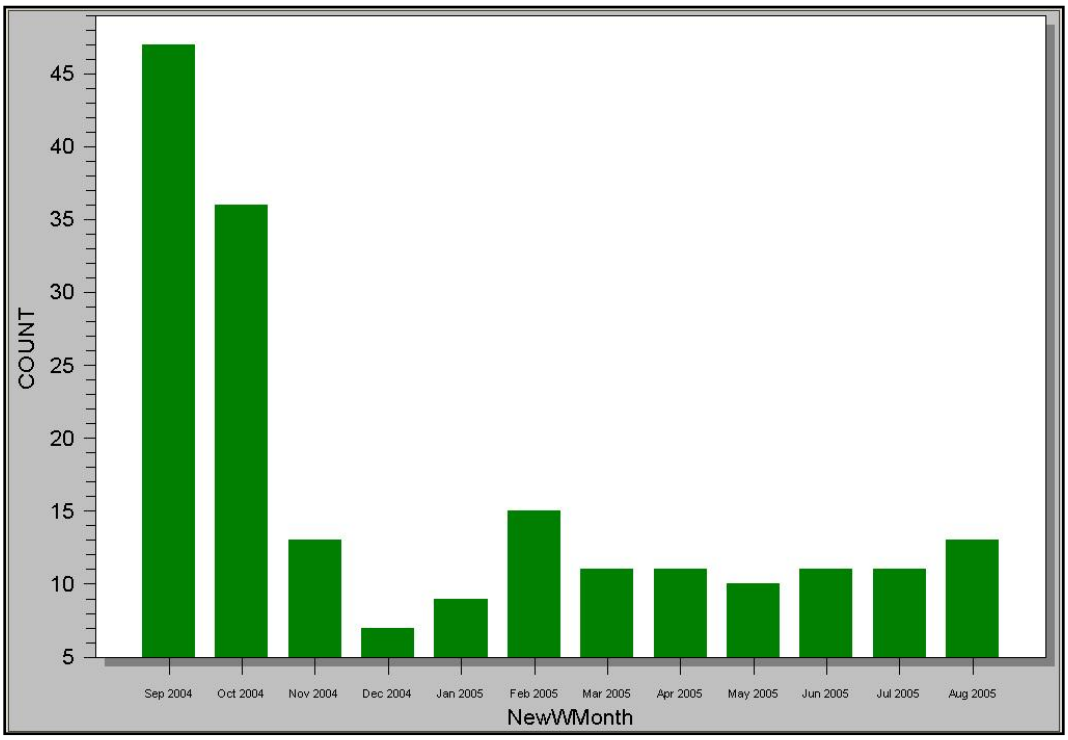
1. Click **Graph**. The GRAPH dialog box opens.

2. From the Graph Type drop-down, select **Bar**.

3. From the Main Variables drop-down, select **NewWMonth**. The Display Format drop-down opens.

4. From the Display Format drop-down, select **mmm yyyy**.

5. Click **OK**.



- Notice the months now appear in the school year order you coded as September 2004–August 2005.

Let's customize the graph

1. Left-click the X-Axis text **NewWMonth**. The X-Axis Label dialog box opens.
2. Highlight the text **NewWMonth**.
3. Delete the text.
4. Click **OK**.
5. Right click the graph and select **Customization Dialog** from the pop-up menu.
6. In the Main Title field on the General tab, type **Worst Month for Asthma**.
7. In the SubTitle field, type **School Year 2004–2005**.
8. Click **OK**.
9. Close Epi Graph.
   - Notice that the window title bar shows the Output file as School Year Graph.

10. From the Command Tree Output folder, click **CloseOut**. The CLOSEOUT dialog box opens.

11. Click **OK**.
    - Results will resume being created and saved as Out#.htm files.
12. From the Output window, click **Open**. The Browse dialog box opens.
13. Locate the EIHA Tutorial folder. Notice the file School Year Graph.htm is listed among the other Out##.htm files.
14. Click **Exit** to close Analysis.

Creating and Writing New Data Tables

For your asthma report, you want to see if there is a connection between age and asthmatic symptoms. You want to illustrate the findings in the form of a graph.

It may sometimes take a series of steps to prepare your data for analyses. In Sections 4 and 5, you will follow a series of steps to create a line graph that shows trends in symptoms per age based on survey data. You will take all the students who answered yes to the individual symptoms and use the WRITE command to create a new data table that contains yes answers. The new data table can then be graphed.

In an earlier lesson, you used the ASSIGN command to calculate the Age variable inside the data table. The Age variable created was Standard, meaning that the code you assigned was temporary and the values were reset each time you read in a new file. You cannot create a global or permanent variable for Age because neither type can contain values that depend directly or indirectly on table fields as the YEARS function does. For this exercise, you will create a line graph using age as a variable; however, you do not want to calculate the Age field each time you create this graph. You will use the WRITE command to create a new data table containing the assigned age variable.

Using the new table, you will create person-symptom pairs of information (children answering yes to a symptom and their age) and WRITE/APPEND those records to a new data table that contains only information for ages and symptoms. You can then read in the age-symptom data table and graph the results.

An example of the following code is available in the Asthma Survey 2005 Sample.MDB. The program is called LineGraphSample.PGM.

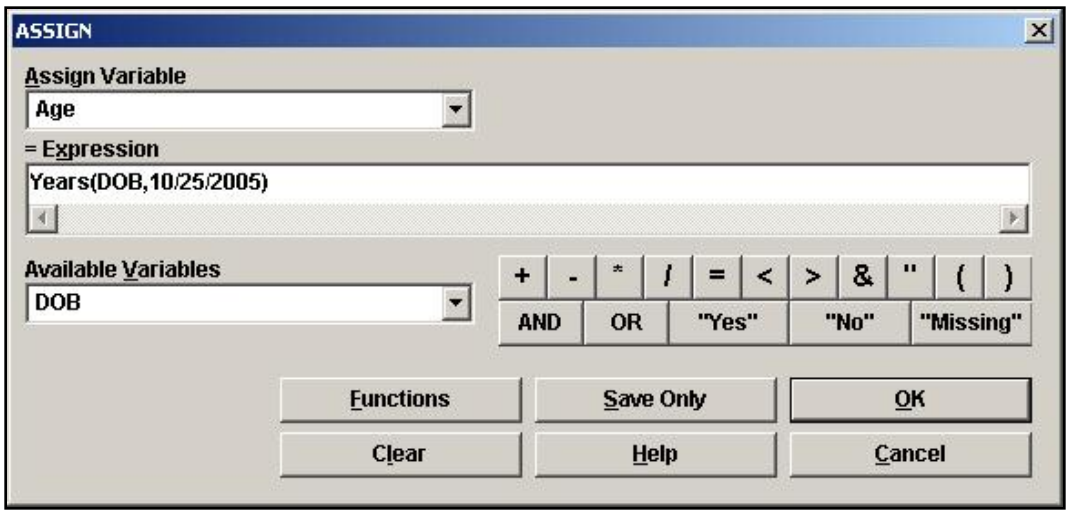
Time for the creative stuff again.

1. Open **Analysis** and **Read** in the Asthma Survey project file that contains your 800 records. Make sure the Project file and the data source match.

2. From the Command Tree Variables folder, click **Assign**. The ASSIGN dialog box opens.

3. From the Assign Variable drop-down, select **Age**.

4. In the =Expression field, type **YEARS(DOB,10/25/2005).**

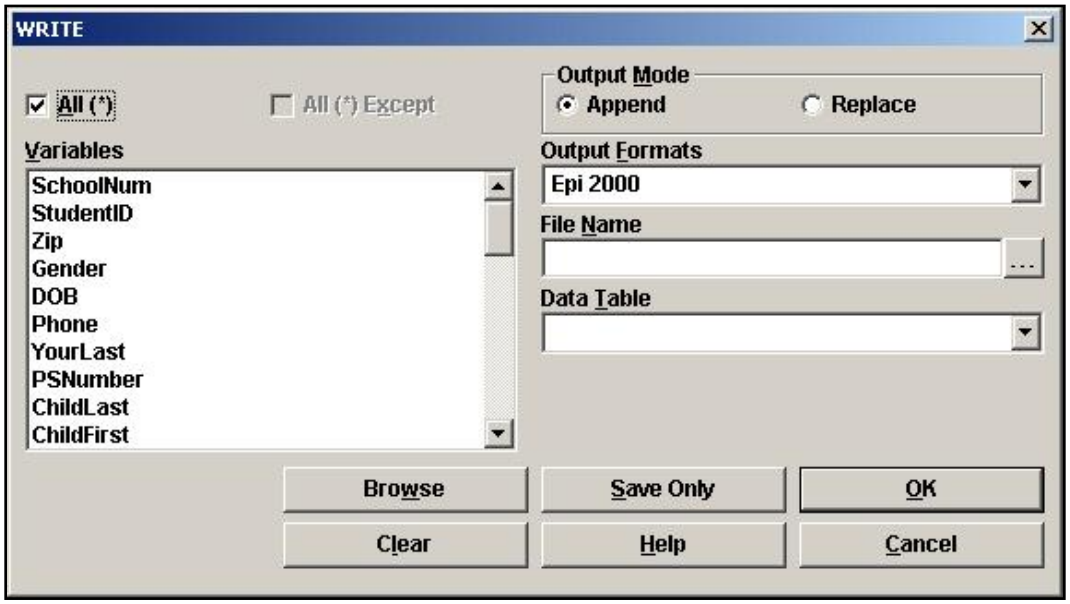


5. Click **OK.**

6. List the variable age to make sure it is populated.

Time to write the Age variable and the variables needed for the line graph to a new file.

1. From the Command Tree Data folder, click **Write/Export**. The WRITE dialog box opens.

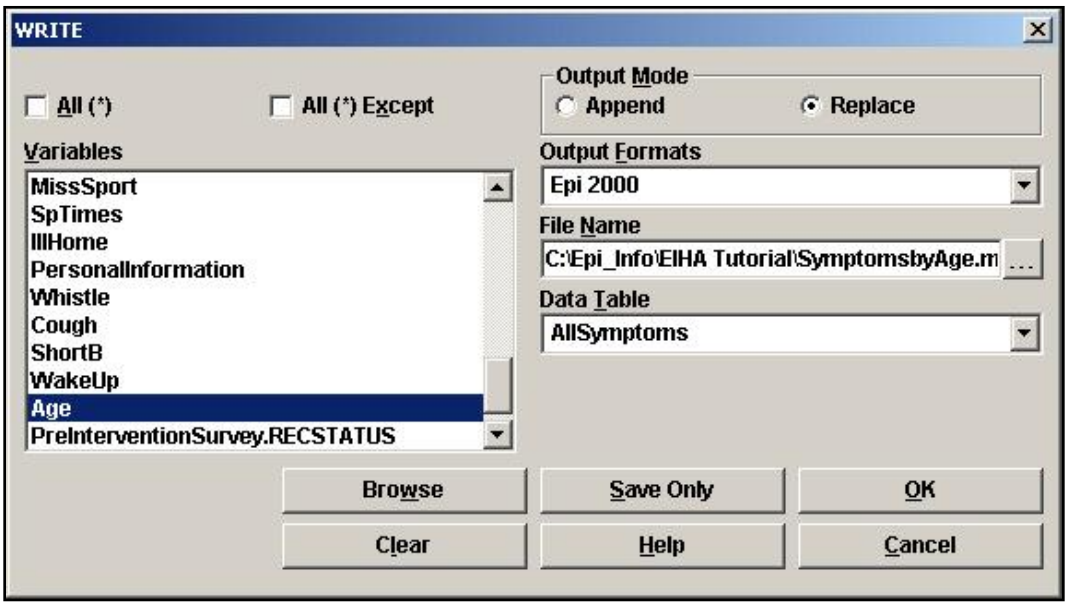


2. Use your mouse to click and select the following variables.

   - **SchoolNum**

- **StudentID**
- **Gender**
- **DOB**
- **Asthma**
- **RAD**
- **Bronchitis**
- **Wheezing**
- **Medication**
- **MissDays**
- **Age**

3. In the Output Mode section, select **Replace**.
4. From the Output Formats drop-down, select **Epi 2000**.
5. Click the **Browse** button next to the File Name field. Locate the EIHA Tutorial folder.
6. In the File Name field, type **SymptomsbyAge**.
7. Click **Save**.
8. In the Data Table field, type **AllSymptoms**.



9. Click **OK**.

Let's read in the new project

1. Click **Read/Import**. The READ dialog box opens.
2. Click **Change Project**.
3. From the EIHA Tutorial folder, select **SymptomsbyAge.MDB**.
4. Click **Open**.
5. Select the table **AllSymptoms**.
6. Click **OK**.
7. List your variables.
   - Your new project contains only the variables you specified when you wrote the new file.
   - The new data is made up of the data table AllSymptoms. There is no view associated with the table.
   - Notice that the field age is populated.

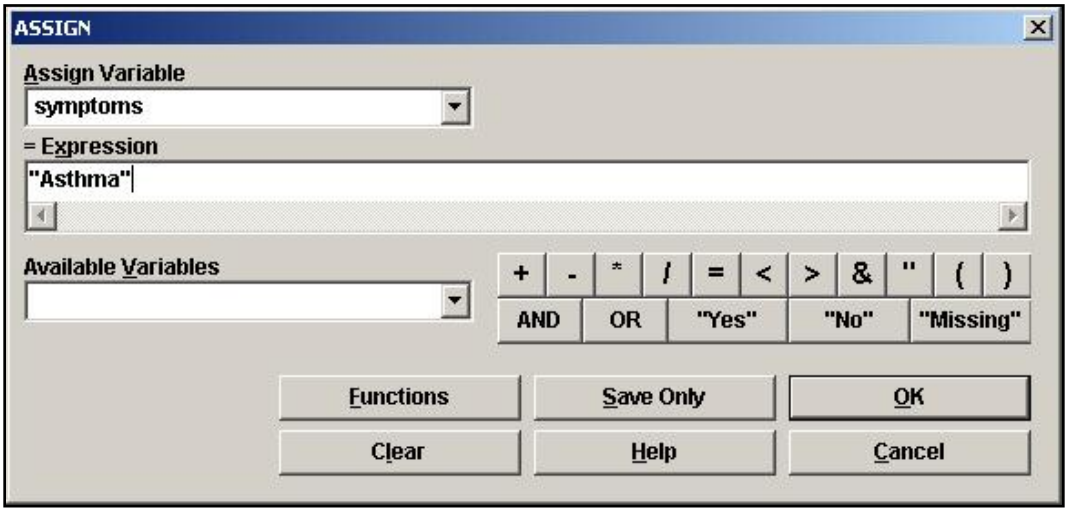
Writing and Appending Tables

You will use the new project and data table to create a graph that has one record for each person-symptom pair. To do this, you will write out a new table, which contains the symptoms you select, and the age of the individual with the symptoms. This will create a second table inside the SymptomsByAge.mdb.

You must define the table, select the criteria, write it to the new table, and append the table to include all four of the symptoms.

You already created one table, this one will be easier.

1. Click **Define**. The DEFINE dialog box opens.
2. In the Variable Name field, type **symptoms.**
3. Select **Standard** as the Scope.
4. Click **OK**.
5. From the Command Tree Select/If folder, click **Select**. The SELECT dialog box opens.
6. Use the Select Dialog box to create the code Asthma="Yes".
7. Click **OK**. You should see a record count of 91.
8. Click **Assign**. The ASSIGN dialog box opens.
9. From the Assign Variable drop-down, select **symptoms**.
10. In the =Expression field, type or use the **Available Variables** drop-downs to create the variable "Asthma".
    - The variable Asthma is a text field so you must enclose it in quotes.



11. Click **OK**. The code in the Program Editor appears like the following:

```
DEFINE symptoms
SELECT Asthma= (+)
ASSIGN symptoms= "Asthma"
```

Now write the 91 asthma records to a new table

1. Click **Write/Export**. The WRITE dialog box opens. You only want to include the variables you want to graph in the new table.

2. From the Variables box, select **Age** and **symptoms**.

- Remember symptoms contains your Asthma=Yes values. You are constructing a new table that contains the yes answers per symptom and the age of the student answering yes. You will add the information for each symptom/age to this table.

3. In the Output Mode section, select **Replace**.

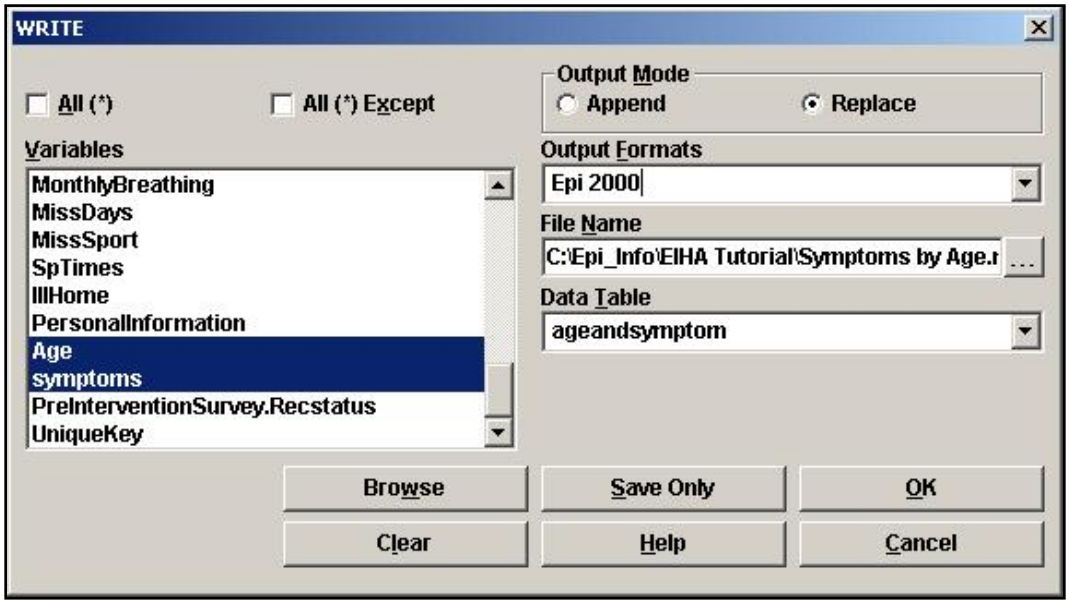
4. Click the File Name **Browse** button.

5. Select the project **SymptomsByAge.mdb**.

6. Click **Save**. The File Name field populates.

7. In the Data Table field, type **ageandsymptom**.

8. Click **OK**.



The code inside the Program Editor should appear as follows:

WRITE REPLACE "Epi 2000" 'C:\Epi_Info\EIHA Tutorial\SymptomsbyAge.mdb':ageandsymptom Age symptoms

9. Click **Cancel Select**.

   - Now you need to complete a similar series of steps to add Bronchitis, RAD, and Wheezing Yes answers to the new table.

Now let's Add/Append Symptoms

1. Select **Bronchitis=Yes**. There should be Record Count of 86.
2. Assign **symptoms="Bronchitis"**.
3. Click **Write/Export**. The WRITE dialog box opens.
4. From the Variables box, select **Age** and **symptoms**.
5. In the Output Mode section, select **Append**.

   - You want to add or append this information to the asthma data table you created earlier.

6. Click the File Name **Browse** button.
7. Select the project **SymptomsByAge.mdb**.
8. Click **Save**.
9. From the Data Table drop-down, select **ageandsymptom**.
10. Click **OK**
11. Click **Cancel Select**.

Now follow the Bronchitis procedure and create the code to add data for RAD and Wheezing to the ageandsymptoms table.

- RAD contains 93 records.
- Wheezing contains 99 records.

- Do not forget to select Append as the Output Mode and to Cancel Select when switching between variables.

- Remember that RAD and Wheezing must be enclosed in quotes in the ASSIGN command.

Your final code should look like the following:

```
DEFINE symptoms

SELECT Asthma= (+)
ASSIGN symptoms="Asthma"
WRITE REPLACE "Epi 2000" 'C:\Epi_Info\EIHA
Tutorial\SymptomsbyAge.mdb':ageandsymptom Age symptoms
SELECT

SELECT Bronchitis= (+)
ASSIGN symptoms="Bronchitis"
WRITE APPEND "Epi 2000" 'C:\Epi_Info\EIHA
Tutorial\SymptomsbyAge.mdb':ageandsymptom Age symptoms
SELECT

SELECT RAD= (+)
ASSIGN symptoms="RAD"
WRITE APPEND "Epi 2000" 'C:\Epi_Info\EIHA
Tutorial\SymptomsbyAge.mdb':ageandsymptom Age symptoms
SELECT

SELECT Wheezing= (+)
ASSIGN symptoms="Wheezing"
WRITE APPEND "Epi 2000" 'C:\Epi_Info\EIHA
Tutorial\SymptomsbyAge.mdb':ageandsymptom Age symptoms
SELECT
```

Let's Read in the new data table

1. Click **Read/Import**.

- Be sure the Current Project and Data Source is SymptomsbyAge.

2. Select the table ageandsymptom. Remember to select **All** in the Read dialog box to see the new tables.

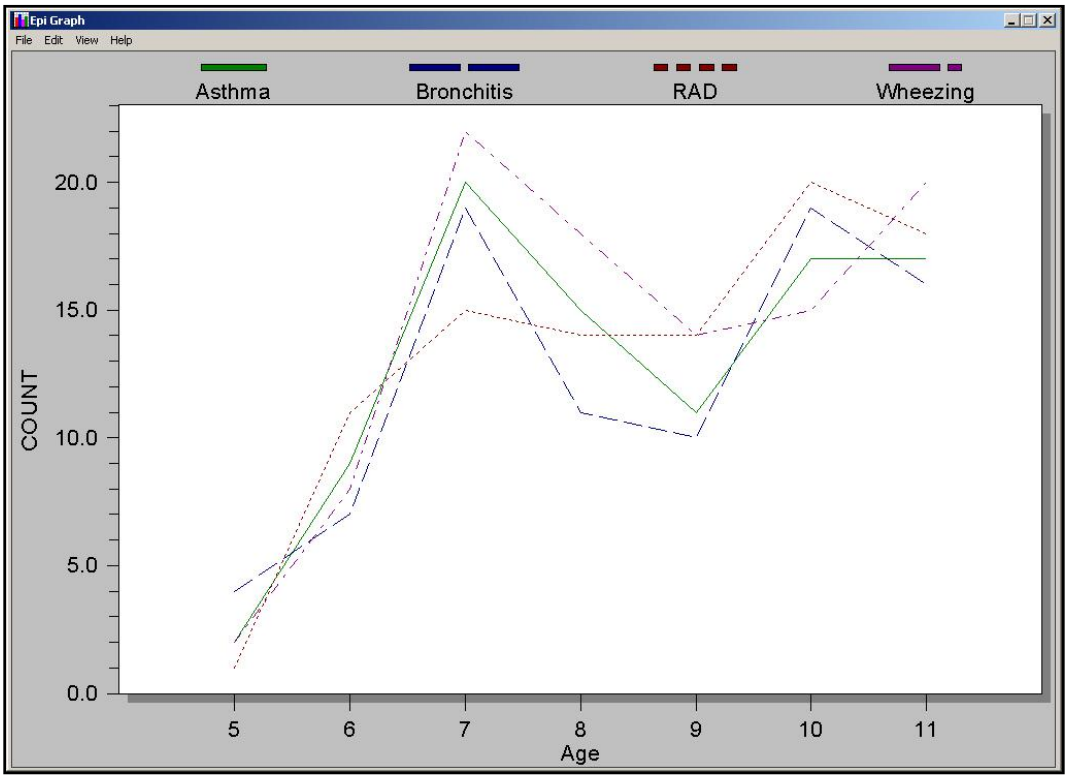
3. Click **OK**.

- There should be 369 records.

Now create the graph

1. Click **Graph**. The GRAPH dialog box opens.
2. From the Graph Type drop-down, select **Line**.
3. From the Main Variables drop-down, select **Age**.
4. From the Line for Each Value of drop-down, select **symptoms**.
5. Click **OK**.
6. Open Epi Graph.

   The graph should look like the following:.



Time to customize the graph

1. Right click the graph and select **Customization Dialog** from the pop-up menu.
2. Title you graph **Symptoms by Age**.
3. Click the **Plot** tab.

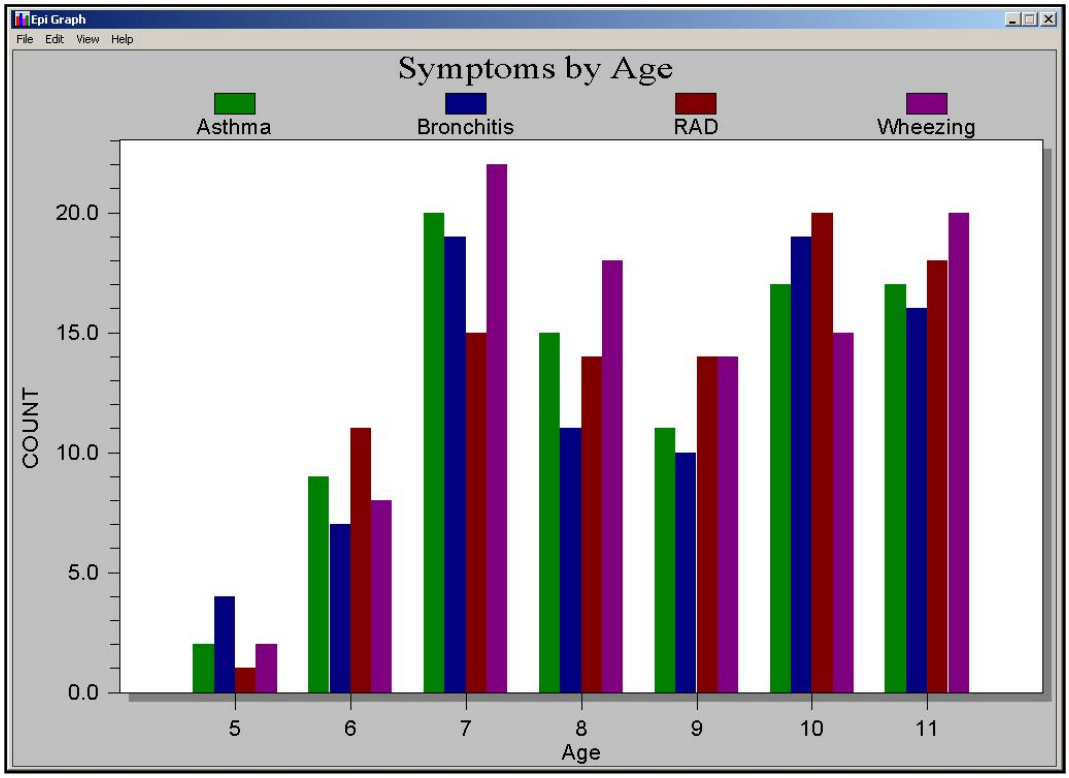
4. Change the Plot Style to Bar.

5. Click **OK**.

6. Right click the graph and select **Export Dialog** from the pop-up menu.

7. Export the Symptoms by Age graph as a .JPG file to your EIHA Tutorial folder.

8. Close Epi Graph.



Writing a File to Excel

Commitment to excel!

1. Click **Read/Import**.
2. From the Read dialog box, select the table All Symptoms.
   - Be sure the Current Project is Symptoms by Age and the Current Data Source is Symptoms by Age.
3. Click **OK**.
4. Click **Write/Export**.
5. Select the **All checkbox** to export all the variables.
6. In the Output mode section, select **Replace**.
7. From the Output Formats drop-down, select **Excel 4.0**.
8. Click **Browse** from the File Name field.
9. Locate the EIHA Tutorial folder and create a new Excel file called **NewAsthmaTable**.
10. Click **Save**.
11. In the Files field, type **NewAsthmaData.**
12. Click **OK**.
    - If you have Excel, you can open this file and view the exported data.
    - If the DOB fields in Excel contain ### rather than dates, expand your columns.

🕮 Skills Review

Place a checkmark next to the correct answer for each of the following questions. Refer to *Appendix D Skills Reviews Answer Key* for further details.

1. Which command allows you to designate a new Output file?

- ☐ WRITE
- ☐ READ
- ☐ ROUTEOUT

2. Which command allows you to create a new variable?

- ☐ SELECT
- ☐ ASSIGN
- ☐ DEFINE

3. Which command allows you to store the value of a variable or assigns the result of a mathematical expression?

- ☐ ASSIGN
- ☐ DEFINE
- ☐ LIST

4. When changing, assigning, or selecting text fields, the value must be enclosed in what?

- ☐ (Parenthesis)
- ☐ *Asterisks*
- ☐ "Quotes"

5. In the Program Editor, the value of Missing appears how?

- ☐ (=)
- ☐ (.)
- ☐ (+)

6. To change a number variable to a date, what function would you use?

- ☐ TXTTONUM
- ☐ NUMTODATE
- ☐ TXTTODATE

7. Which command allows you to create a new data table?

- ☐ MODIFY
- ☐ READ
- ☐ WRITE

8. Which command allows you to open a new data table?

- ☐ READ
- ☐ WRITE
- ☐ LIST

9. From the WRITE dialog box, what Output Mode would you select to add data to an existing table?

- ☐ Replace
- ☐ Append

10. The WRITE command can be used to export a data table to MS Excel.

- ☐ True
- ☐ False

Lesson Complete!

WHAT YOU LEARNED

How to use the:

- HEADER and PRINTOUT commands to customize output.
- ROUTEOUT and CLOSEOUT commands to create named output files.
- DEFINE and RECODE commands to change variable values.
- WRITE/EXPORT command to create new data tables.
- WRITE/EXPORT command to create an MS Excel file.

This page intentionally blank.

Analysis: Data Management for Maps

SUMMARY

This lesson introduces you to creating maps in the Analysis module using Epi Map features. It focuses on the data management skills necessary to map data.

You will aggregate or group data using the SUMMARIZE and FREQUENCY command. Data tables must be aggregated and contain a geographic value in order to be mapped. At the end of this lesson, you will have a map that shows the count of missed days from your survey mapped by zip code. From this, you can see which zip codes have the highest number of absences based on answers to the survey question.

Maps can be saved as images and used in reports to later illustrate the findings as part of the asthma report. You will use a map to illustrate the number of missed days of school by the zip codes of the survey population.

Length of time to complete: 1 hour Beginner

Getting Started with Epi Map

Introduction to Epi Map including a diagrammed explanation of how data must match or be aggregated to match shape files to produce maps.

Map Command

Use the SUMMARIZE command to aggregate data and create a choropleth map using the data to show the sum of missed days of school per zip code in the region.

Aggregate Data for Mapping

Use the FREQUENCY command to create a data table that contains the count of asthma cases per zip code for the survey population. This data table will be used in Lesson 9 Epi Map.

Skills Review

Hands-on task to practice creating an aggregate data table that contains the count of bronchitis cases per zip code for the survey population. This data table will be used in Lesson 9 Epi Map.

BEFORE YOU BEGIN

Complete Lesson 4 Analysis: Basics

Complete Lesson 5 Analysis: Creating Statistics

Complete Lesson 7 Analysis: Exporting Files

Download the map files to the EIHA Tutorial folder.

WHAT YOU NEED

Asthma Survey 2005.MDB.

ALBZCTA_region.SHP

ALBZCTA_region.PRJ

ALBZCTA_region.DBF

ALBZCTA_region.SHX

FIVE GOALS

- Become familiar with the types of data that can be mapped in Epi Map and the types of map files needed.
- Use the SUMMARIZE command to create a sum of the variable MissDays that can be plotted on a map.
- Use the MAP command and make geographic selections leading to a choropleth map of the MissDays data table.
- Use the FREQUENCY command to create an output table containing the count of asthma cases by zip code.
- Use the FREQUENCY command to create an output table containing the count of bronchitis cases by zip code.

Getting Started with Epi Map

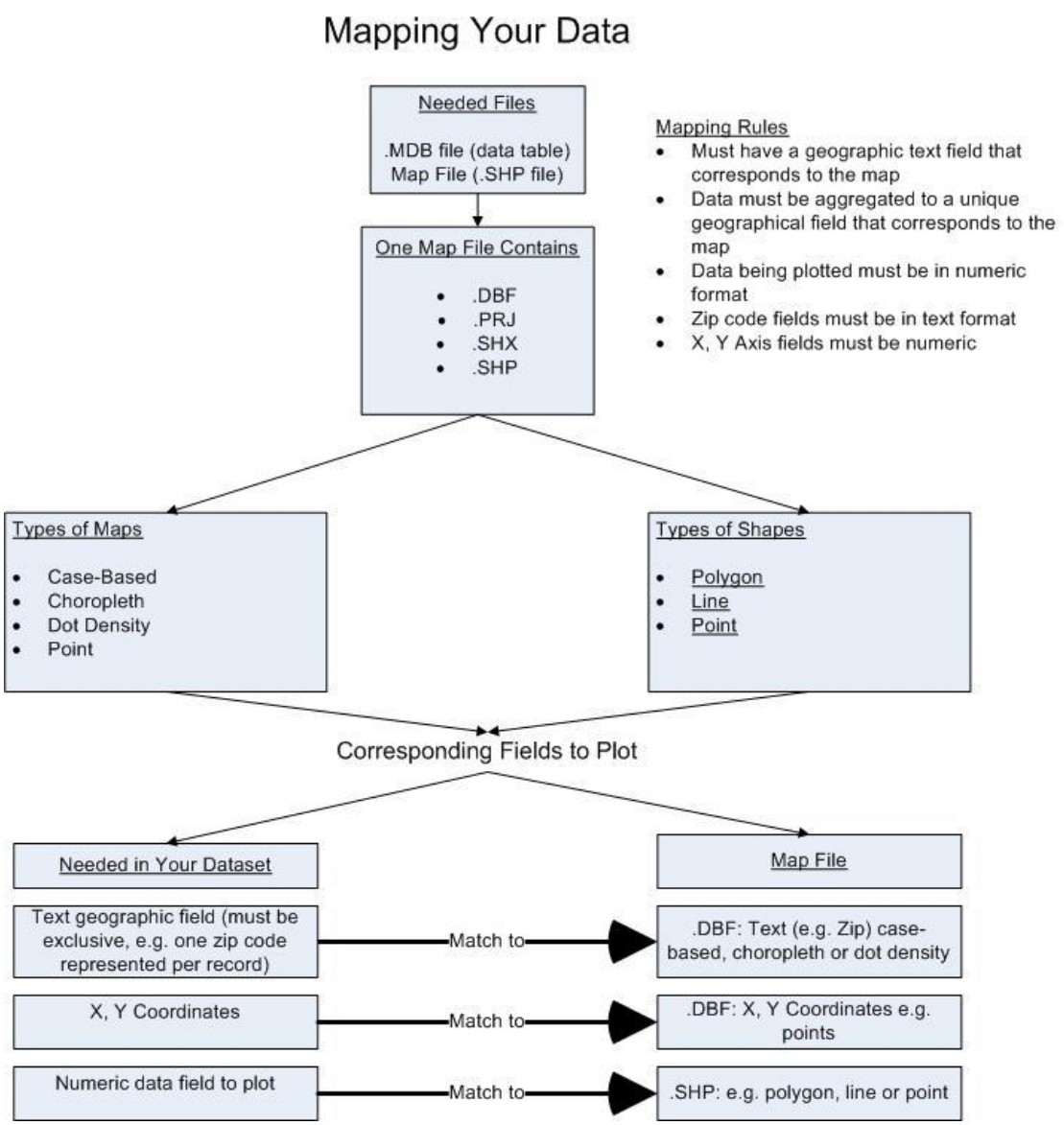
Epi Map, the mapping component of Epi Info, is built around MapObjects software from ESRI, the makers of ArcView and ArcInfo, popular Geographic Information System (GIS) tools. Epi Map displays shapefiles from both systems, and uses the enormous reservoir of map boundaries and geographic data available on the Internet in ESRI-compatible formats.

Epi Map is designed to show data from Epi Info files by relating data fields to shapefiles containing the geographic boundaries. Shapefiles can also contain data on population or other variables, and can provide numeric data that becomes part of the display as a numerator or denominator. Numeric data can be displayed either as color/pattern (choropleth) maps or dot density maps with the dots randomly distributed within geographic regions.

Point locations can be plotted automatically from data files containing X and Y coordinates in various symbols, colors, and sizes. Shapefiles can contain lines or points to represent streets or point locations, and points can be placed on top of the shapefile layer to represent homes in which cases occurred or other geographic points of interest.

Choropleth and Case-Based maps can be created through Analysis. More map options are available when working through the Epi Map program on its own.

For information on how to download shape files, refer to *Appendix G: Preparing to Use Epi Info™*.


Mapping Your Data
Needed Files
.MDB file (data table)
Map File (.SHP file)
One Map File Contains
.DBF
.PRJ
.SHX
.SHP
Mapping Rules
Must have a geographic text field that corresponds to the map
Data must be aggregated to a unique geographical field that corresponds to the map
Data being plotted must be in numeric format
Zip code fields must be in text format
X, Y Axis fields must be numeric
Types of Maps
Case-Based
Choropleth
Dot Density
Point
Types of Shapes
Polygon
Line
Point
Corresponding Fields to Plot
Needed in Your Dataset
Map File
Text geographic field (must be exclusive, e.g. one zip code represented per record)
Match to
.DBF: Text (e.g. Zip) case-based, choropleth or dot density
X, Y Coordinates
Match to
.DBF: X, Y Coordinates e.g. points
Numeric data field to plot
Match to
.SHP: e.g. polygon, line or point

Commands in this Lesson

SUMMARIZE

Creates a new table containing descriptive statistics for the current dataset or its strata. Analysis creates a new table or appends to an existing table containing variables that represent aggregates or groups of variables in the current data source. Located in the Statistics folder.

Determined by the Stratified Variables, aggregates are computed for each group of records, which are included in the table. Available aggregates are Count, Minimum, Maximum, Sum, Average, Variance, and Standard Deviation.

MAP

Produces a choropleth (color coded) map by summarizing data based on a geographic field that matches the geographic field of a .SHP file. Through Epi Map, the map can be customized by the user, which can be saved into a MAP template. Located in the Statistics folder.

SELECT

Specifies an expression that must be true for a record to be processed. Located in the Select/If folder.

FREQUENCY

Produces a table showing how many records have each value of selected variables. Confidence limits for each proportion are included. Located in the Statistics folder.

Aggregate Data in Analysis

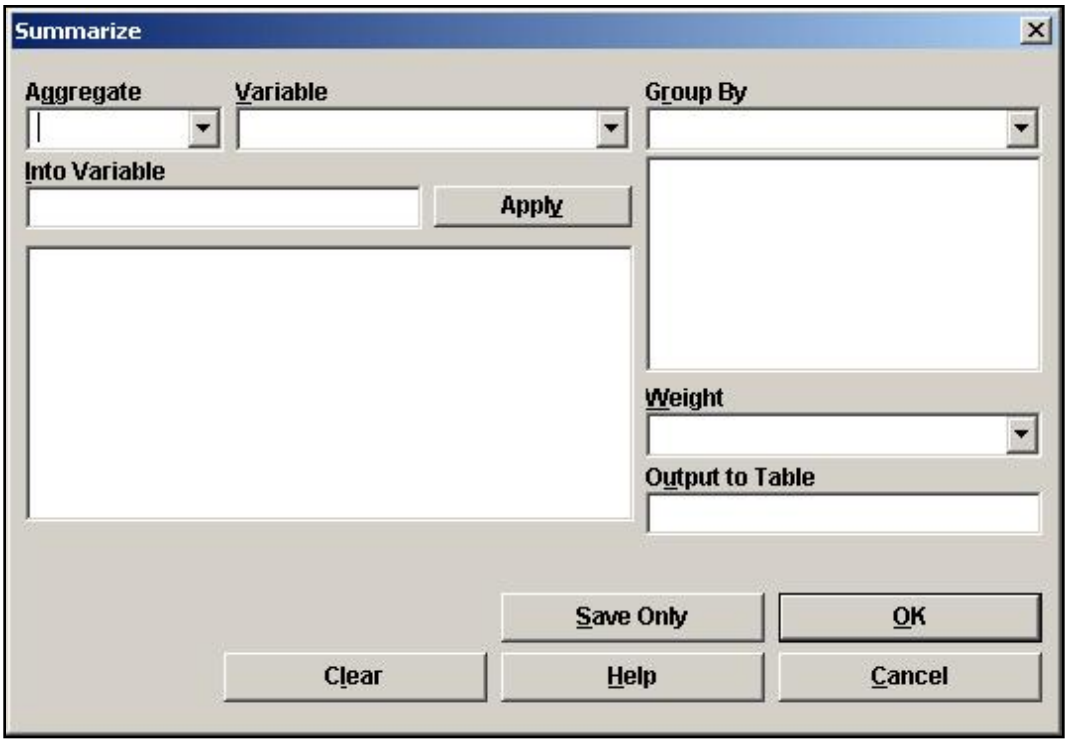
In this lesson, you will create a map of the sum of values in the MissDays variable.

Map data must be aggregated to a unique geographic field that corresponds to the map file. To create a map that shows the sum of missed days per zip code, you must use the SUMMARIZE command to aggregate your data. Then you can create a new table that will contain the Count of MissDays grouped (aggregated) by Zip.

More mapping time.

1. Read/Import the Asthma Survey 2005.MDB that contains your 800 records.

2. From the Command Tree Statistics folder, click **Summarize**. The Summarize dialog box opens.



3. From the Aggregate drop-down menu, select **Sum**.

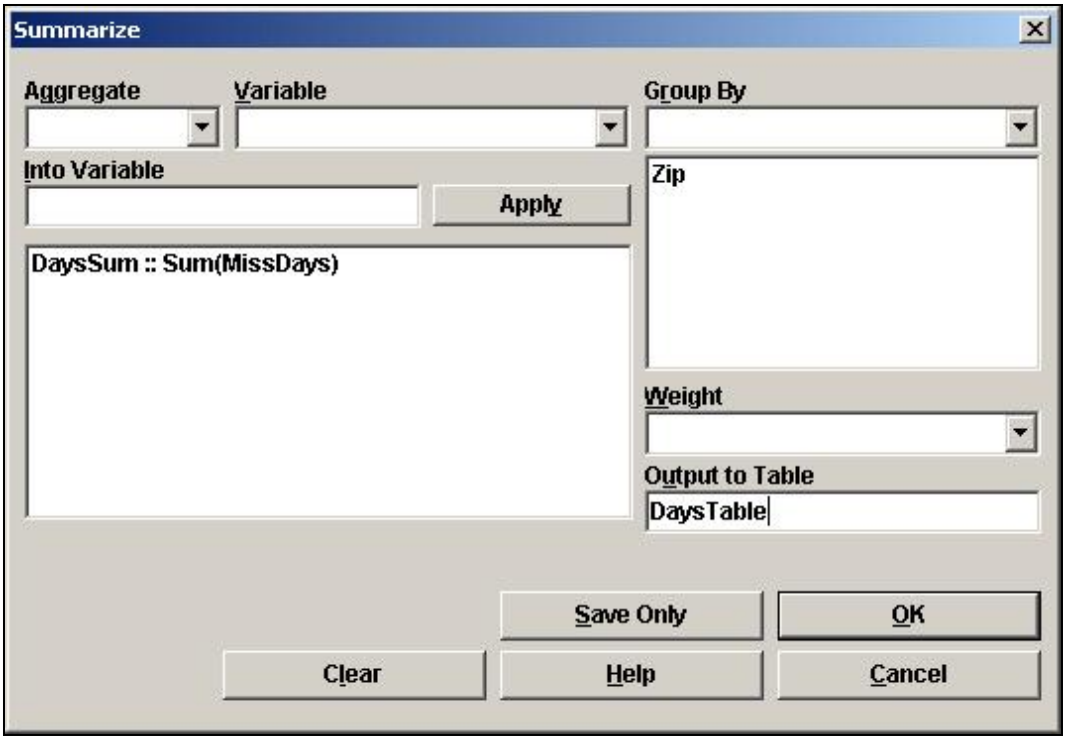
4. From the Variable drop-down, select **MissDays**.

5. In the Into Variable, type **DaysSum**.

6. Click **Apply**. The code appears in the open field as DaysSum::Sum(MissDays).

7. From the Group By drop-down, select **Zip**.

8. In the Output to Table field, type **DaysTable**.



9. Click **OK**. The following message appears in the Output window.

- Output table created: C:\Epi_Info\EIHA Tutorial\Asthma Survey 2005.mdb:DaysTable

10. Click **Read/Import**. The READ dialog box opens.

11. Click the **Show All** radio button. All the tables in your dataset appear in a list.

12. Locate and select the **DaysTable**.

13. Click **OK**. You should have 33 records in the Output window.

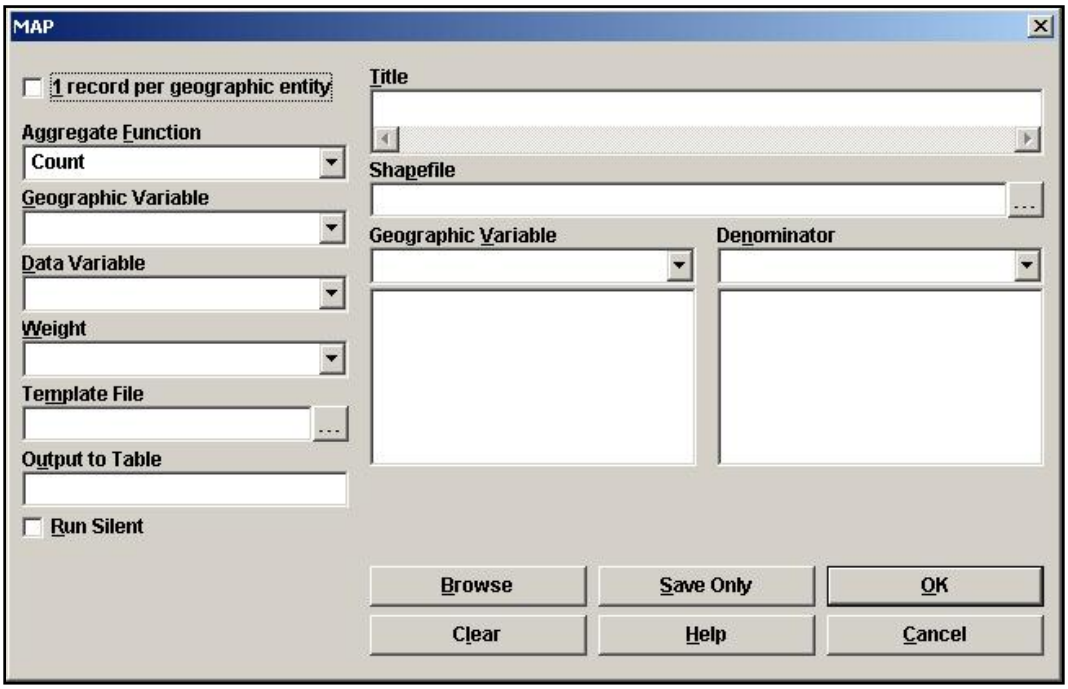
- Click **List** if you want to see how your data has been aggregated with the number of missed days summed per zip code.

Creating a Choropleth Map from Analysis

You can use the aggregated data to create a choropleth map. Choropleth maps are shaded with a range of colors indicating differing sums. A corresponding legend is also created. To see sample code for all the maps in this chapter, open the Asthma Survey 2005 Sample.MDB and run the program MapCode.PGM.

More mapping.

1. From the Command Tree Statistics folder, click **Map**. The MAP dialog box opens.



2. Select the **1 Record per Geographic Entity** checkbox. The Aggregate Function drop-down defaults to Sum.

3. From the Geographic Variable drop-down, select **Zip**.

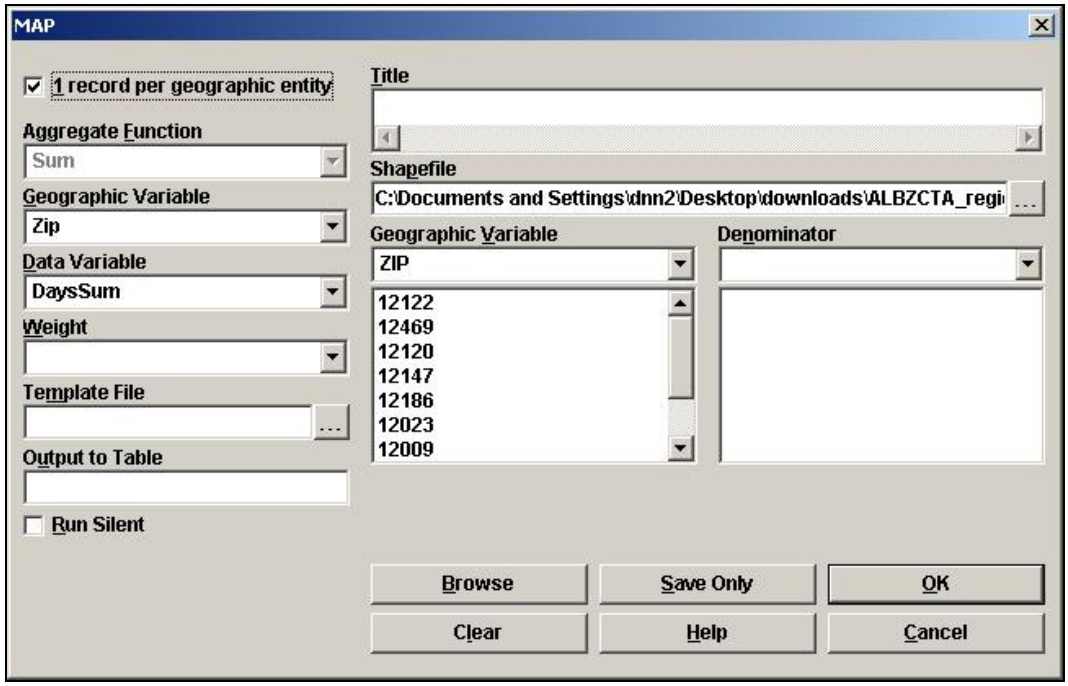
4. From the Data Variable drop-down, select **DaysSum**.

5. Click the **Browse** button next to the Shapefile field. The Look In Map dialog box opens.

6. Locate and select the file **ALBZCTA._region.SHP**. This is a map of Albany, NY by zip code regions.

7. Click **Open**.

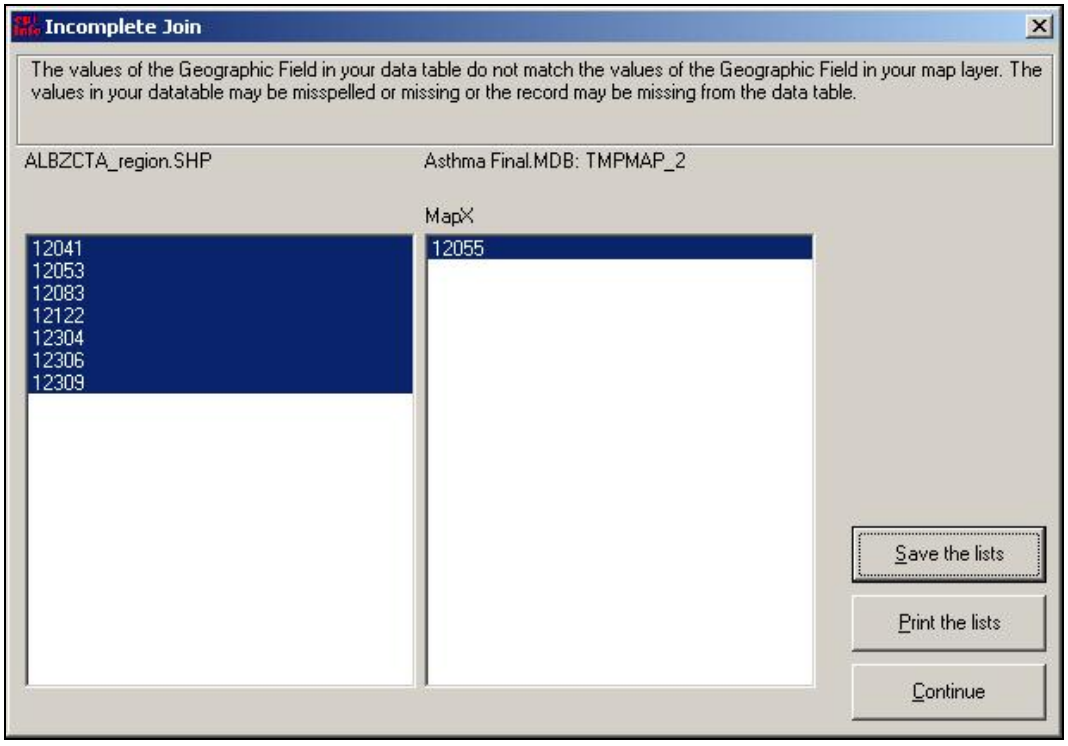
8. From the Geographic Variable drop-down, select **Zip**.



9. Click **OK**. The Incomplete Join window appears.

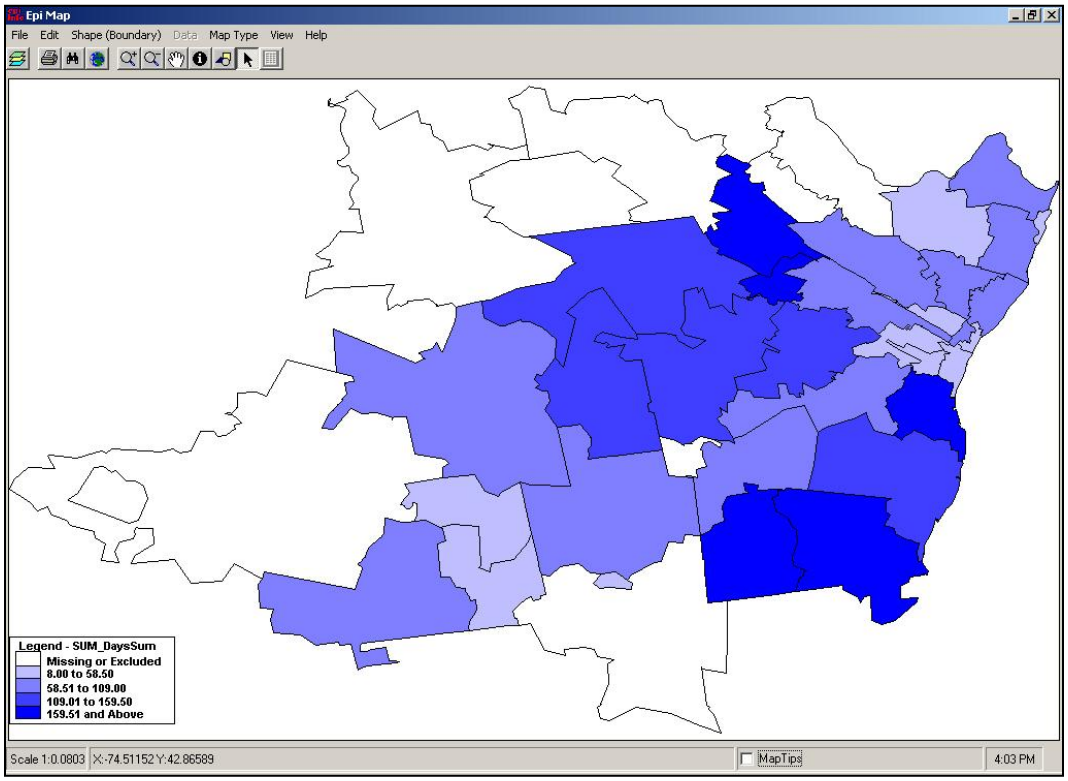
- From the Incomplete Join window, you can see which zip codes are in your data table and not in the map files; and which zip codes are in your map files, and not in the data table.

- In this example, you have seven zip codes that are part of the map files, but not in the survey data; and one zip code that is in the survey data, but not part of the map. In order to get the most accurate representation of the data, you need to verify that the survey data was correct and did not contain missing records or data entry errors. In addition, you need to determine that you had the best map to represent the surveyed area.



10. Click **Continue**. Epi Map opens.

The map you created should appear like the following:



- You have created a choropleth map. Choropleth is used to categorize features into equal ranges or counts (quantiles).

- Notice the legend at the bottom left of the window. Your map is color-coded, based on the sum of days missed per zip code. This map shows a graphic representation of the number of missed days by zip code, and enables you to see if there are areas of the county in close proximity to each other with large numbers of missed days.

- Click the close **X** to exit Epi Map.

Aggregating Data for a Case-Based Map

You can add case-based data to a choropleth map. Normally, case-based maps are used to show different symbols based on levels of classification (e.g., Confirmed, Probable, Discarded, Suspected).

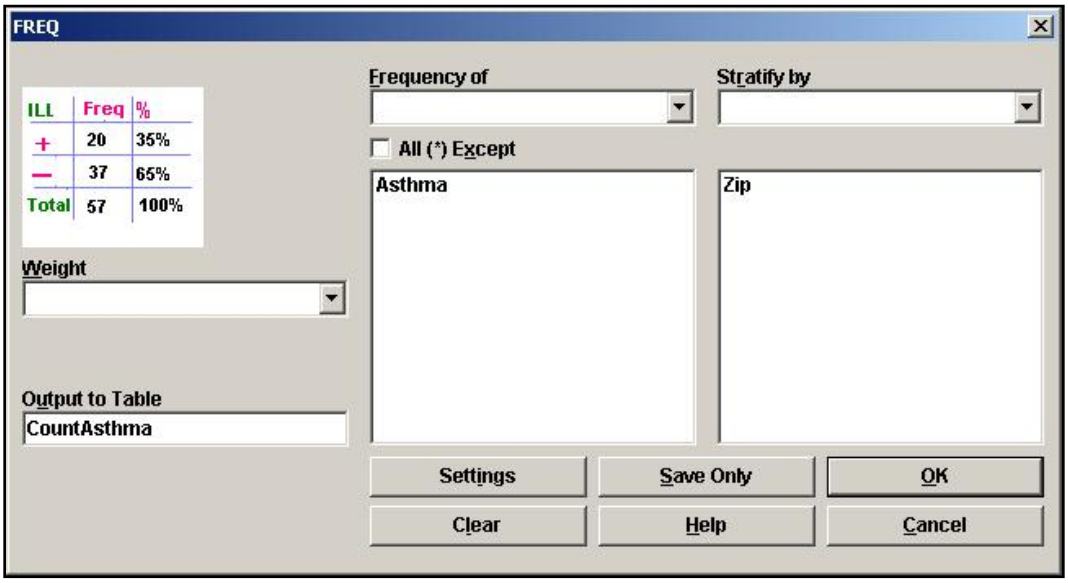
In this example, you will aggregate asthma data to place numbers of cases on the map per zip code. Use Analysis to create the data table. The new data table will be used in Epi Map Lesson 9.

Time to solve the case.

1. Read/Import the Asthma Survey 2005 project with 800 records.

2. Use the SELECT command to select those who answer yes to asthma. You should have 91 records. (Hint: Asthma=Yes)

3. Create a FREQUENCY of Asthma stratified by Zip.



4. In the Output to Table field, type **CountAsthma**.

5. Click **OK**.

6. Read/Import in the file CountAsthma. Remember to select the **All** radio button to see the all tables in your project. You should have 33 records.

7. List the records.

   - Notice that each zip code contains the total count of students who answered yes to asthma for that zip code.

🕮 Skills Review

To create a map containing the count of bronchitis cases per zip, you must create a new frequency table for bronchitis. This table will be used in the Skills Review of Chapter 9.

Read/Import the file Asthma Survey 2005.

Select the students who answered yes to bronchitis.

Create a FREQUENCY of Bronchitis stratified by Zip.

Output to a table named CountBronchitis.

Read/Import CountBronchitis.

List the records to verify the data table contents.

Close Analysis.

Lesson Complete!

WHAT YOU LEARNED

How to:

- Use the SUMMARIZE command to aggregate data.
- Use the MAP command to create a choropleth map.
- Use the FREQUENCY command to aggregate data.
- Create Output tables to use in mapping.

Epi Map

SUMMARY

This lesson introduces you to using the Epi Map module. You will create choropleth, case-based, and point maps. It includes information on customizing maps and saving .MAP files.

For the asthma report, you want to see how survey results can be mapped. You will create a map containing income data for the selected region and then plot the number of asthma cases from the survey onto the map. This will give you an idea if income and asthma are related. You will then create point maps that contain the location of schools and hospitals in the region. In the Skills Review section, you will create a map from survey data that shows the number of students with bronchitis per zip code.

Length of time to complete: 2 hours Beginner

Epi Map Overview

Review of information needed to create maps. Provides information on navigating Epi Map and Map Types.

Creating a Choropleth Map

You can create one by using income data and then plot asthma cases on top of it to look for correlations between income and cases of asthma.

Changing Map Properties

Use the Layer Properties option box to customize a map.

Viewing Map Information

Overview of using Map Tips, the Information button, and the Find button to view data applicable to the map.

Saving a .MAP File

How to save a map in an interactive format.

Creating a Case-Based Map

Use data aggregated in Analysis to create a case-based map for the number of cases of asthma per zip code. This data will be plotted on top of the income choropleth to enable you to look for correlations between asthma and income.

Adding Text to a Map

Use the Title and Text options to add editable text to a map.

Creating a Point Map

Use the Add Points option to plot school and hospital locations on the county map for the surveyed region. Includes information on using the Layer Properties options box to edit points.

Applying Standard Labels

To add zip code labels to a map.

Applying Advanced Labels

To label the hospital locations on the map. Also covers the Zoom and Pan Features of Epi Map.

Using Map Layers

Use the Add Layer feature with the single tab to create a transparent layer showing the boundary and zip code overlaps for the region.

Saving a Bitmap

Learn how to save a map as a bitmap file for use in other programs or presentations.

Skills Review

Practice creating a map for the count of students with bronchitis per zip code based on the Asthma Survey 2005 project. Use the Unique tab from the Layer Properties options box.

BEFORE YOU BEGIN

Complete Lesson 8 Data Management for Maps in Analysis.

WHAT YOU NEED

ALBZCTA_region.SHP

ALBZCTA_region.DBF

ALBZCTA_region.PRJ

ALBZCTA_region.SHX

Albany_Demogz.MDB

Data Table CountAsthma

Data Table CountBronchitis

SchoolPoints.MDB

HospitalPoints.MDB

AlbanyCountyBoundary_region.SHP

AlbanyCountyBoundary_region.DBF

AlbanyCountyBoundary_region.SHX

AlbanyCountyBoundary_region.PRJ

FIVE GOALS

- Create a choropleth map of median income for the surveyed region.
- Create a case-based map that shows the correlation between income and asthma cases.
- Create a point map to show the location of schools and hospitals in the surveyed region.
- Add a shape layer to a map as a transparent overlay to show county boundaries for the surveyed region.
- Create a map illustrating the distribution of bronchitis cases per zip code in the surveyed region.

Using Epi Map

Choropleth and case-based maps can be created through Analysis; however, more map options are available when working through the Epi Map program on its own. Epi Map contains two primary interfaces: the Map Manager and the Properties dialog box. Used together, maps can be customized and saved as bitmaps for use in presentations or as .MAP files that retain interactivity and can be edited and updated as data changes. Map files are made up of 2–4 connecting files: the .SHP, which is the polygon/shape of the area to be mapped; the .DBF that contains the data, and the .SHX/.PRJ, which is the index file linking the shape and the data together.

For information on how to download shape files, refer to *Appendix G: Preparing to Use Epi Info™*.

To create maps, you must follow some basic mapping rules:

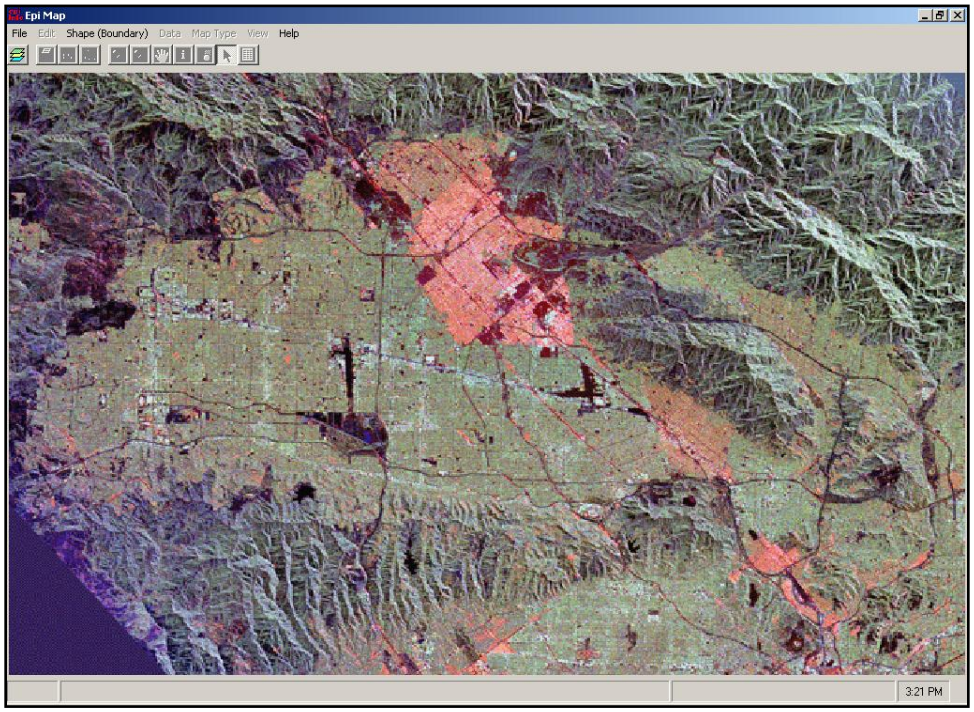
- Have a geographic field in your dataset that corresponds to a shape file.
- Data to be mapped must be aggregated to a unique geographic field.
- Data being plotted must be in numeric format.
- Zip code fields must be in text format.
- X and Y axis fields must be numeric.

The Epi Map Workspace

Time to get back to work!

1. From the Epi Info™ main page, click **Create Maps** or select **Programs>Create Maps** to open the Epi Map module.

   - Here's an example of the Epi Map workspace.



2. Select **File>Map Manager** or click the  **Map Manager** icon. The Map Manager opens.


Map Manager
Layers
General
Add Layer...
Add Layer Partial...
Add Data...
Add Points...
Remove Layer
Clear all Layers
Properties...
Composition...
Merge Layers...
Buffer Layer...
Case-Based...

Navigating Epi Map

Most of the work in Epi Map is conducted from the Map Manager, which allows map layers to be constructed with shapefiles and related data variables. You will use the Map Manager to add, remove, or re-order layers. It also allows you to add data, add points, or add case-based information to a map. From the Map Manager, you will also access the Properties dialog box, which allows you to customize you map and display the data.

Maps can be saved as .MAP files that are editable and can be opened in Epi Map or through Analysis. Maps can also be saved as bitmap graphic files and used in presentations and in other programs.

Many features in the Epi Map menu are also accessible through Map Manager and Map Tool Bar. The following options are available in the Epi Map menu.

File

When Epi Map starts, three options become enabled: Map Manager, Open Map File, and Exit. Once a layer is added, usable options are enabled.

Edit

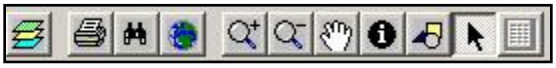
Allows you to copy the map as a bitmap to paste to other applications.

Shape (Boundary)

Use to open the Shape Editor, a blank canvas used to design and edit polygons.

View

Options available in the view menu are also accessible using the Map Tool Bar.



Options available through the Tool Bar are the Map Manager, Print, Find, Global Extent (centers a map on the workspace), Zoom In and Out, Pan, Information, Graphics, Cursor, and Records for Feature. Clicking the Graphics icon opens a new tool bar that allows you to add text or shapes to the map.

Maps Types

The following maps will be created in this lesson.

Choropleth

Use to categorize features into equal ranges or counts (quantiles) that are color-coded to indicate changes in the data. In this lesson, you will create a choropleth map that illustrates income data.

Case-Based

Allows users to show different symbols based on levels of classification (e.g., Confirmed, Probable, Discarded, Suspected). Epi Map produces a case-based legend that displays the symbols, colors, and the case classification values represented. In this lesson, you will create a case-based map that shows the sum of the cases of asthma per zip code based on the frequency table you created earlier in Analysis.

Point Map

Requires that an X and Y axis be part of your dataset and the map file you are plotting. When creating a point map, you can save the points as a separate shape. Points can also have labels applied to them during creation. In this lesson, you will create a point map to illustrate the locations of schools and hospitals in the survey county.

Creating a Choropleth Map

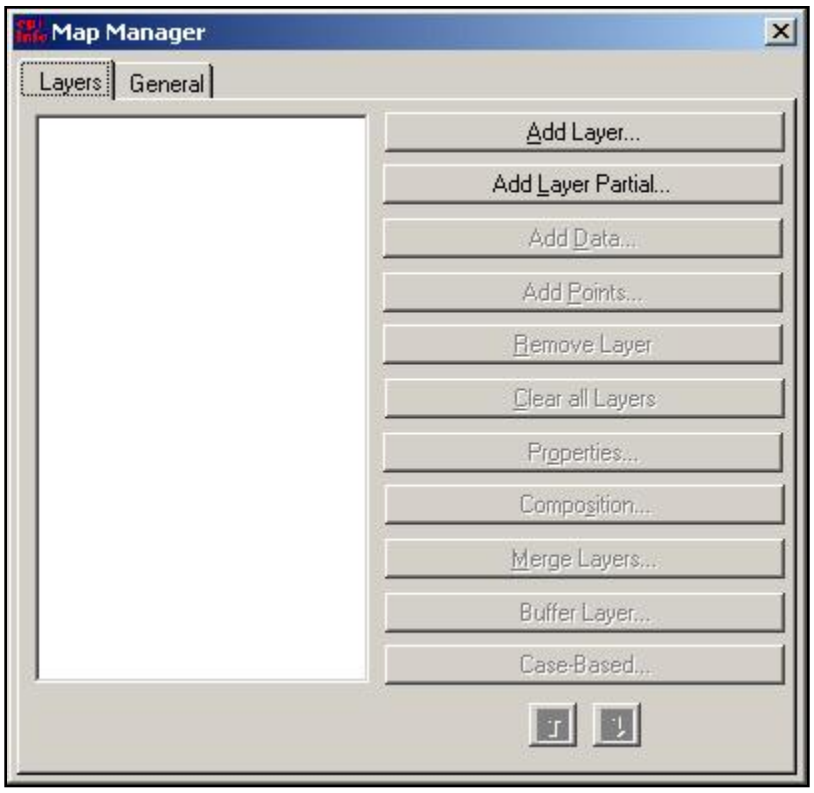
In this lesson, you will create and customize a choropleth map and save it as a .MAP interactive map file.

The schools in the survey cover a zip code region containing a variety of demographics. To determine whether you have more cases of asthma in certain income brackets, you will create a choropleth map of the regions using income data and then map cases of asthma from the school survey onto the map. This will help you determine if a link exists between income and asthma in the students from the survey. First, create the choropleth map and then plot the case-based data on to the map.

You're getting good at this!

1. Open **Epi Map**.

2. Click the **Map Manager** icon or select **File>Map Manager**. The Map Manager opens.

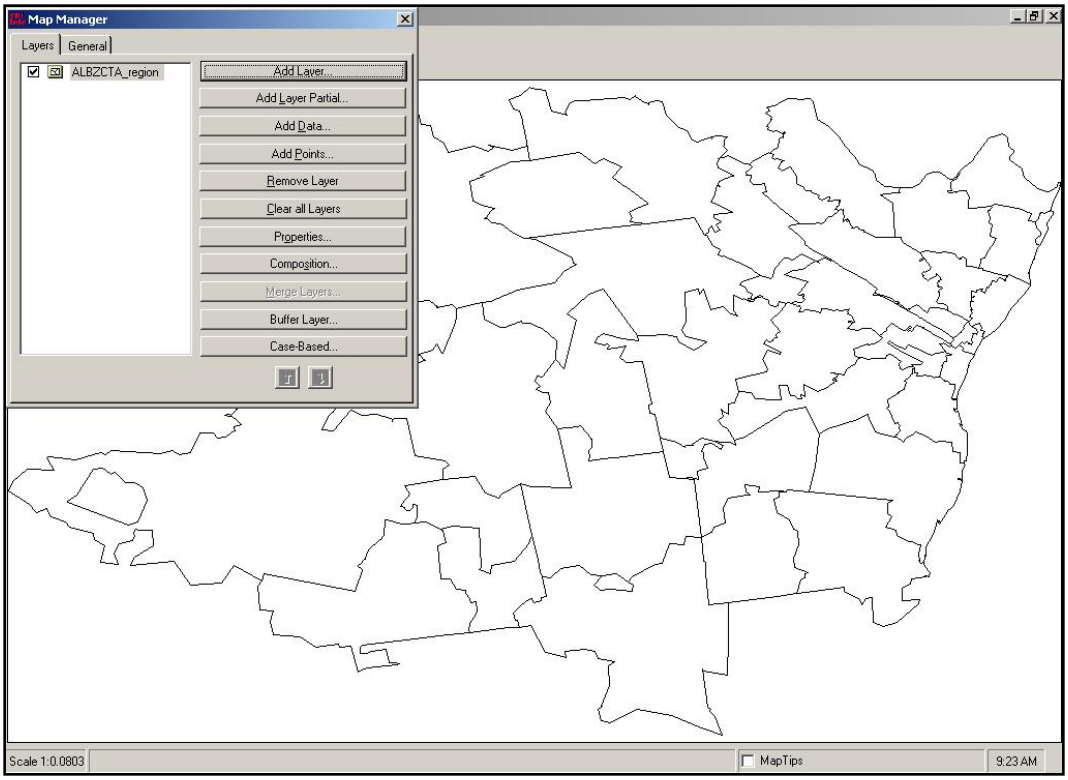


   - Maps work using a series of layers. To start, you must add your region layer.

3. Click **Add Layer**. The Add Layer window opens.

4. From the EIHA Tutorial folder, select the **ALBZCA_region.SHP** file.

5. Click **Open**. The map shape appears in the workspace and the layer appears in the Map Manager.



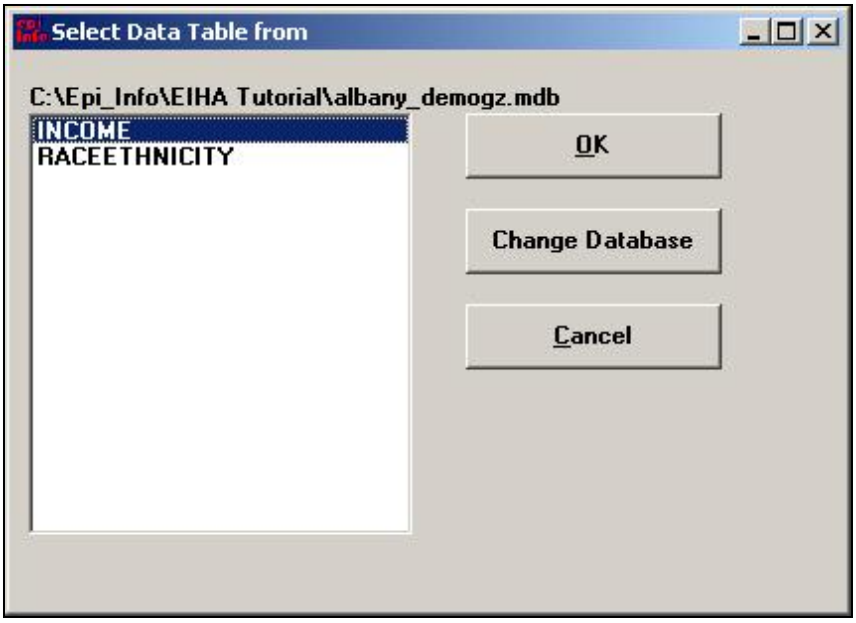
6. From the Map Manager, click **Add Data**. The Open Data Project window opens.

- Remember that data needs to be in an Epi Info™ project or MS Access file to be plotted on the map.

7. Locate the EIHA Tutorial folder, select the **file albany_demogz.mdb** file.

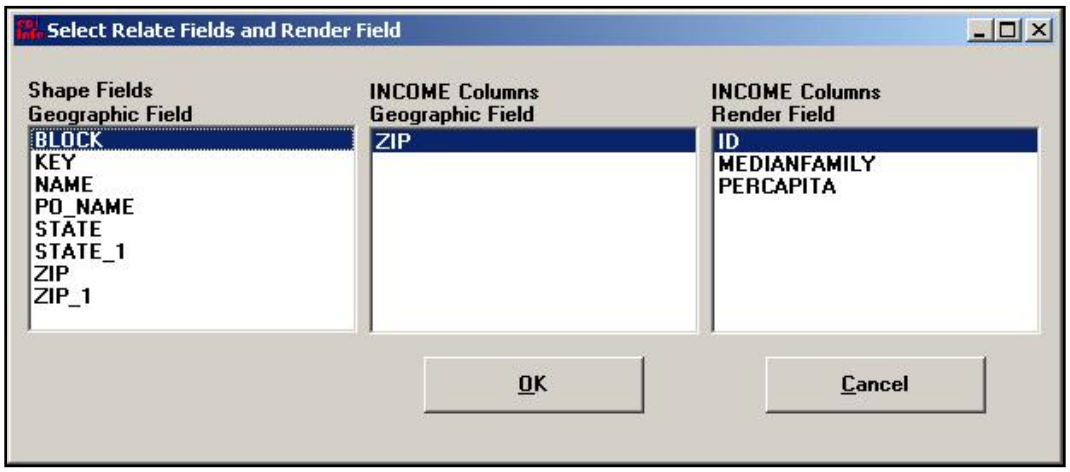
8. Click **Open**. The Select Data Table From dialog box opens.

- The demographic project has two tables. One contains income data and the other race and ethnicity data. You want to create a choropleth map using the income data. The Income table contains a series of zip codes and the median income and per capita income for those zip codes. The median income is the value of family income with 50% of families above the value and 50% of the families below the value. Per capita income is the total income for that zip code divided by the population.



9. Select **INCOME**.

10. Click **OK**. The Select Relate Fields and Render Field dialog box opens.



11. From the Shape Field Geographic Field box, select **Zip**.

- This is the field inside the shape file that matches your data.

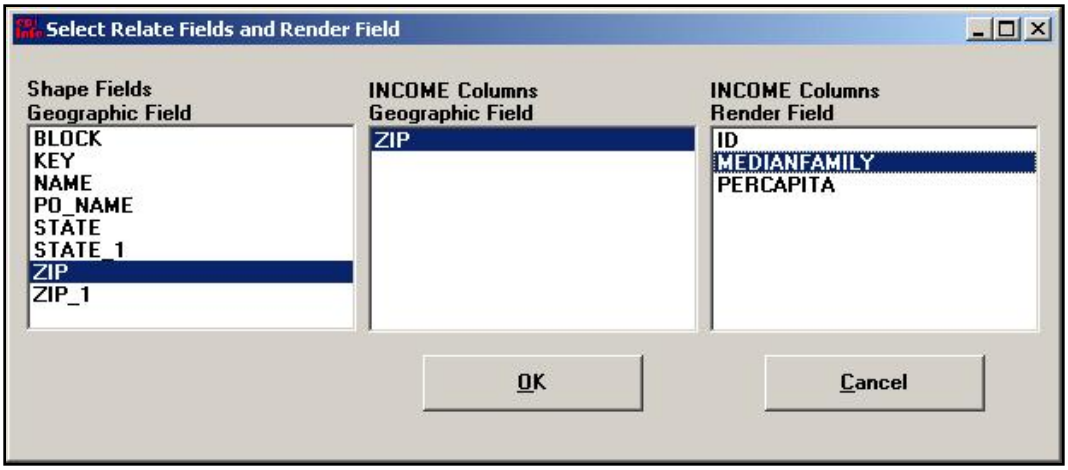
12. From the INCOME Columns Geographic Field, select **ZIP**.

- This is the field inside your data that matches the shape file.

13. From the INCOME Columns Render Field, select **MEDIANFAMILY**.

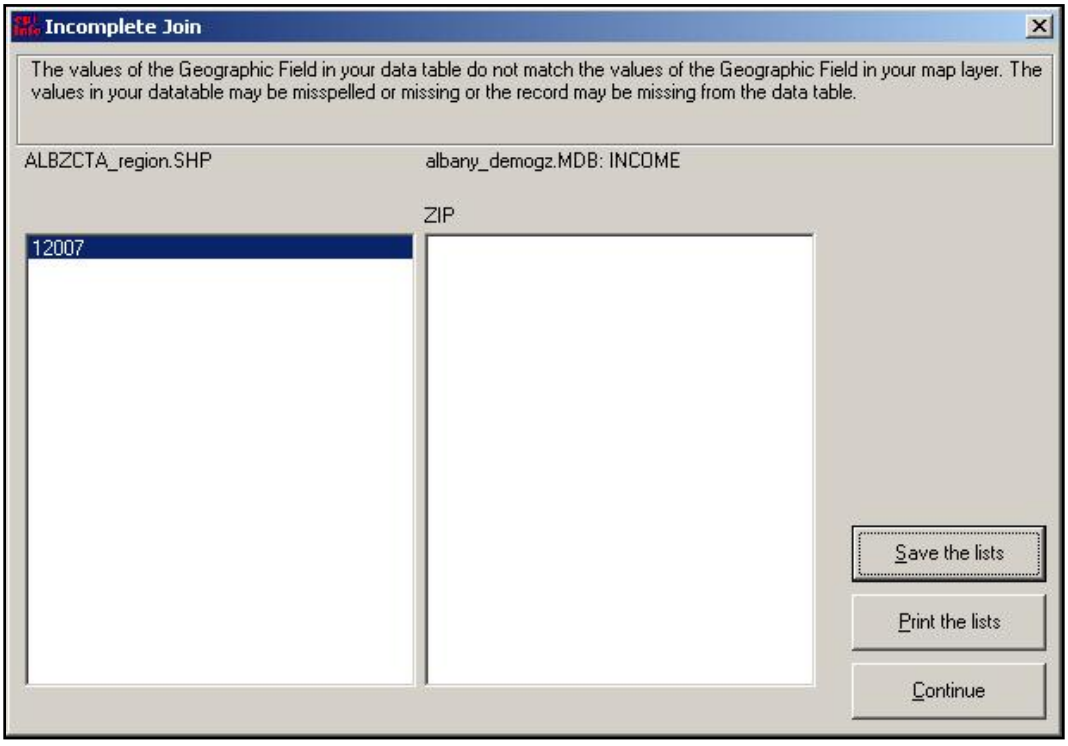
- This is the data field to be mapped.

- The Select Relate and Render Field dialog box should look like the following:



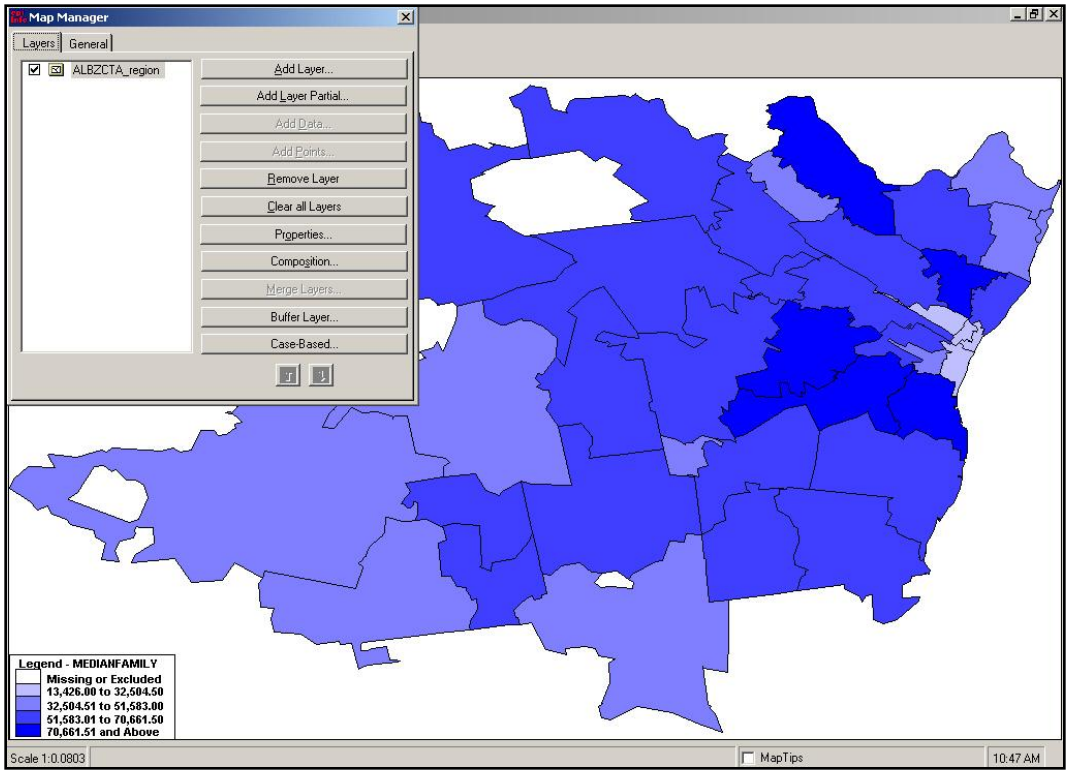
14. Click **OK**. The Incomplete Join window opens.

- This window tells you that the ALBZCTA_region file contains information on zip code 12007 that is not contained in the demographic project file.



15. Click **Continue**. The choropleth map appears. Notice the legend contains the color codes for the Median Family Income per zip code. The darkest color represents the highest income range and the lightest color represents the lowest.

Your map should look like the following:

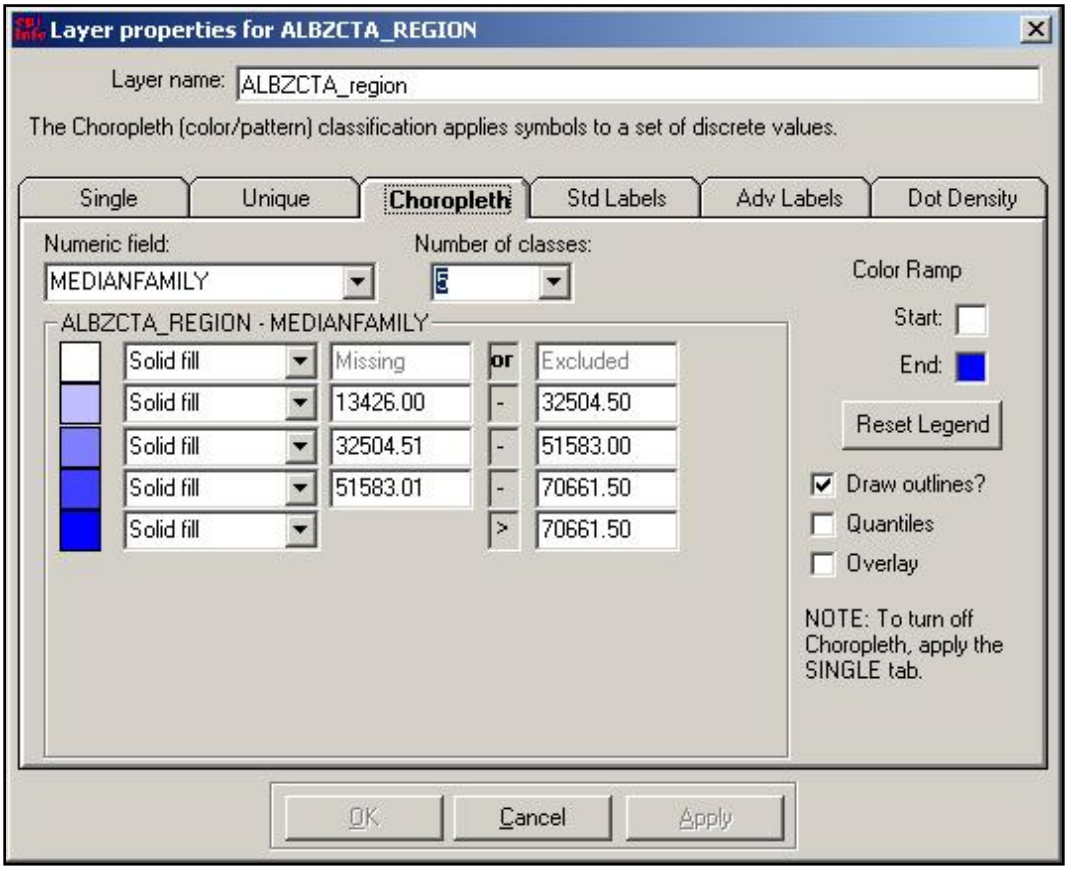

Map Manager
Layers
General
ALBZCTA_region
Add Layer...
Add Layer Partial...
Add Data...
Add Points...
Remove Layer
Clear all Layers
Properties...
Composition...
Merge Layers...
Buffer Layer...
Case-Based...
Legend - MEDIANFAMILY
Missing or Excluded
13,426.00 to 32,504.50
32,504.51 to 51,583.00
51,583.01 to 70,661.50
70,661.51 and Above
Scale 1:0.0803
MapTips
10:47 AM

Changing Map Properties

Notice that your legend contains a breakdown on information based on five divisions of data. To extend the ranges and further customize the map, you will use the Layer Properties options.

Let's start changing.

1. From the Map Manager, click **Properties**. The Layer Properties dialog box opens.



- The number of classes listed shows how many divisions or groups, your data is broken into on the map.
- The color ramp on the right side shows the lightest to darkest colors for the map codes. Use the color ramp to change colors inside the map.
- Any changes you make to the choropleth map from the Property box will not be applied until you click the **Reset Legend** button. You must click **Reset Legend** to accept the changes and **Apply** to place them in the map.

2. From the Number of Classes drop-down, select **10**.

3. Click **Reset Legend**. The number of classes changes to 10.

4. From the color ramp, click the **End** box. The Color dialog box opens.

5. Select a **dark** color to be the end color.

6. Click **OK**.

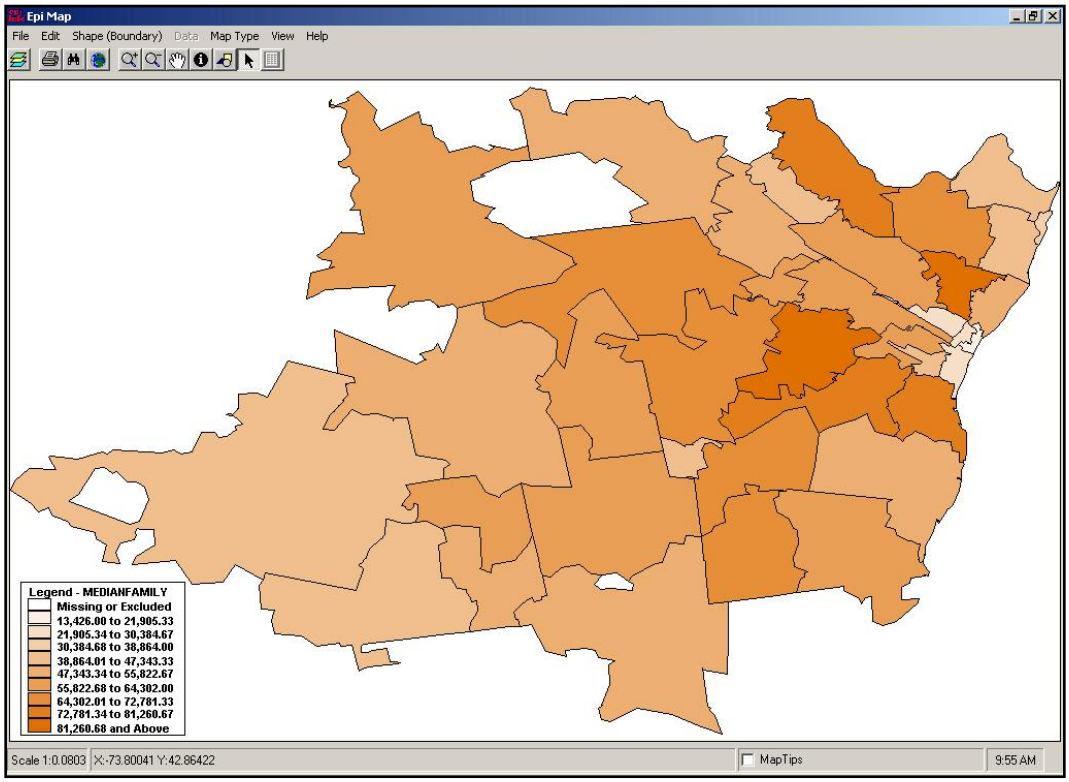
7. Click **Reset Legend**.

8. Click **Apply**.

9. Click **OK** to close the Properties dialog box.

10. Close the Map Manager by clicking the close **X**.

Your map should look similar to the following:



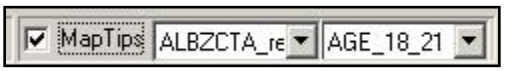
Viewing Map Information

Map Tips

The ALBZCTA_region shape file contains census data as part of the associated .DBF file. To see this information, use the Map Tips option.

Time to receive some tips.

1. At the right hand bottom of the page, click the **Map Tips** check box. The Map Tips drop-down menus appear.



2. From the second drop-down, select **Age_5_17**.

3. Use your mouse button to hover over different map regions.

   - The number that appears under your cursor in each zip code area is the number of those 5–17 years of age in that area per the census numbers.

4. Change the second drop-down selection to **Zip**.

5. Use your map to hover over the different map regions and see the zip codes.

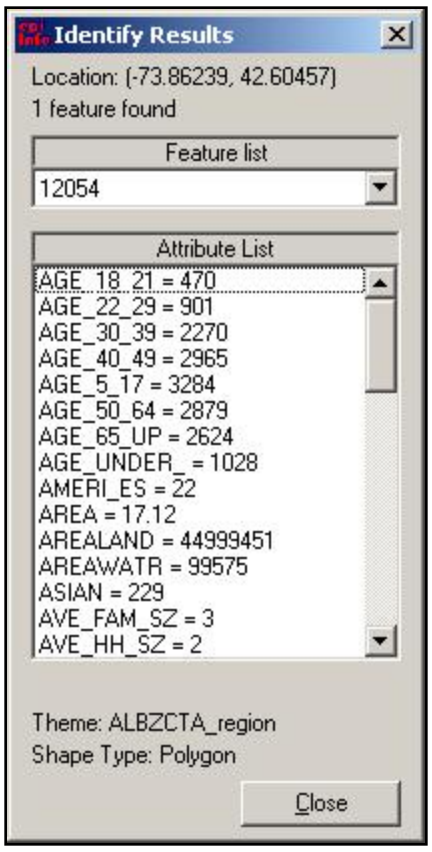
Information Icon

Use the Information icon to get more information from your map.

1. Click the  **Information** icon from the Map Tool Bar.

2. Your cursor will be tagged with the Information icon.

3. Click inside any region on the map. The Identify Results window opens.



- The Identify Results window gives you a scroll-through list of all the information about your selected region contained in the dataset.

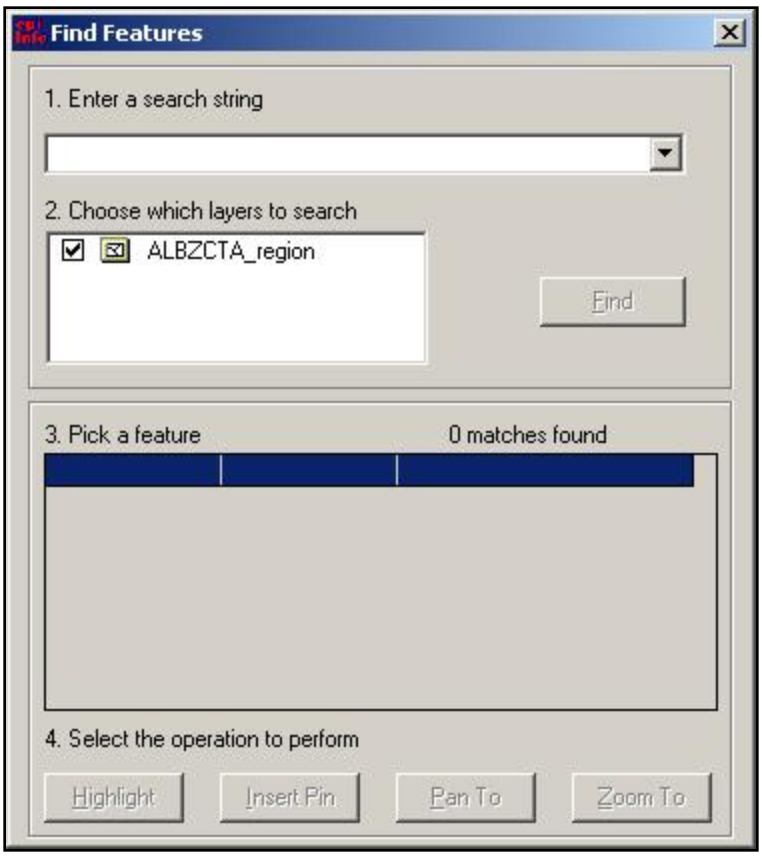
4. Click **Close** to exit this feature.

Find Icon

Another way to locate information in the map is to use the Find feature.

Time to earn your finders' fee!

1. Click the  **Find** (binoculars) icon to activate the Find feature. The Find Features window opens.



2. In the Enter Search String drop-down, type the zip code **12054**.
3. Click **Find**. The Pick a Feature section populates.
   - Be sure the ALBZCTA_region file is selected in the Choose Which Layers to Search field.
4. Click **Highlight** and the 12054 zip code region lights up to show its location on the map.
5. In the Enter Search String drop-down, type the zip code **12046**.
6. Click **Find**.
7. Click **Highlight**.
8. Click **Zoom To**. The map zooms into the requested area.
9. Click the Close **X** to exit the Find Features window.
10. Click the  **Full Extent** icon to return the map to the center of the workspace.

Saving a .MAP File

Save your map in interactive mode as a .MAP file for future editing.

Back into the saving mode.

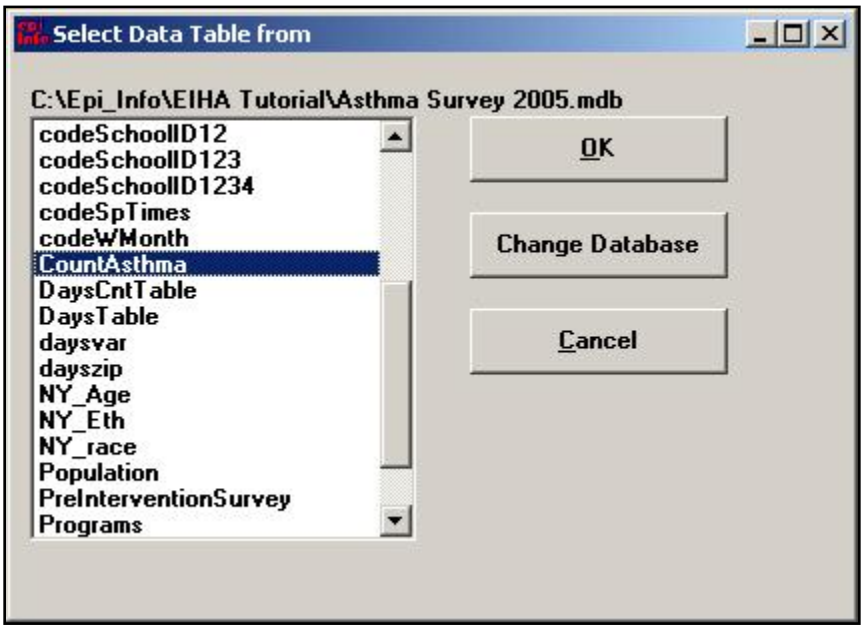
1. Click **File>Save Map File**. The Save Map File window opens.
2. Name your file **IncomeAsthma.**
3. Click **Save**.
4. Close **Epi Map**.

Creating a Case-Based Map

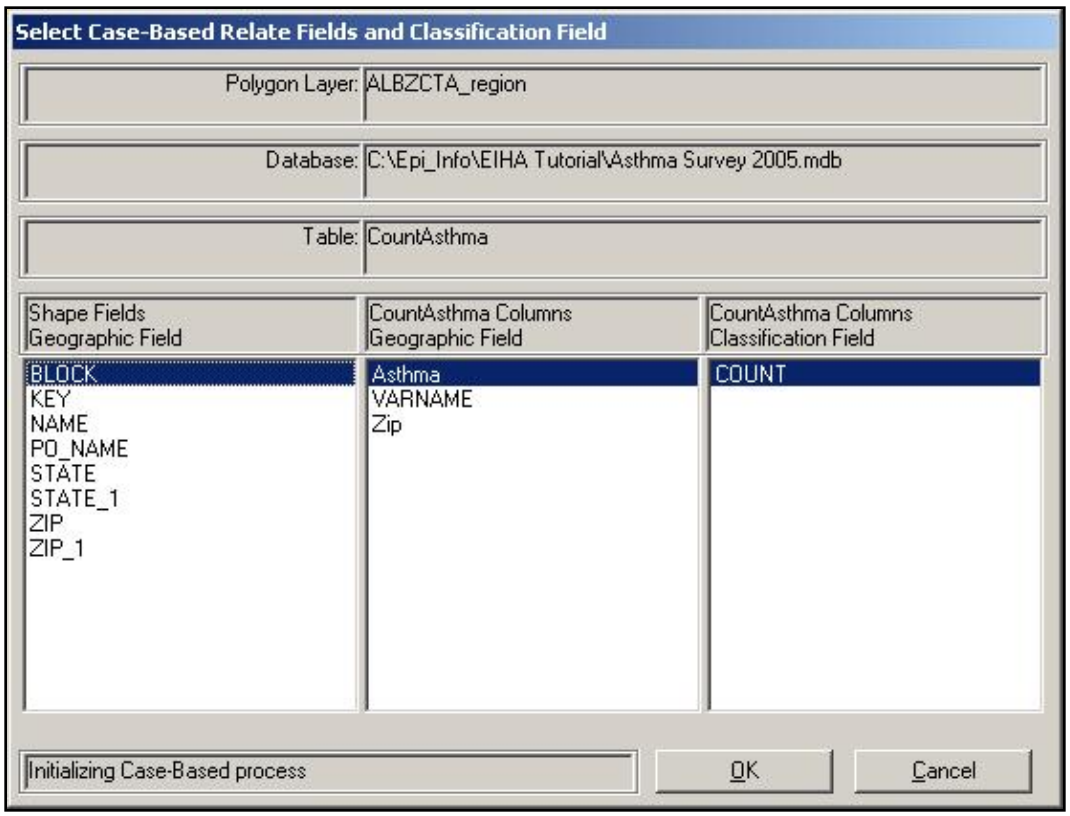
In this lesson, you will create a case-based map that shows the sum of the cases of asthma per zip code, based on the frequency table you created earlier in Analysis. This will help illustrate which zip codes had more cases of asthma and the income for those zip codes.

Create a case for a map.

1. Click **Create Maps**. Epi Map opens.
2. Select **File>Open Map**.
3. Locate and open the file **IncomeAsthma.MAP**. This file contains the choropleth map of income you created in the previous lesson.
4. If you receive the Incomplete Join window, click **Continue**.
5. Open the **Map Manager**.
6. Click **Case-Based**.
7. Locate and select the **Asthma Survey 2005.MDB**.
8. Click **Open**. The Select a Data Table dialog box opens.



9. Select **CountAsthma** from the list of tables in your project.
10. Click **OK**. The Select Case-Based Relate Fields and Classification Field dialog box opens.



11. From the Shape Fields section, select **Zip**.

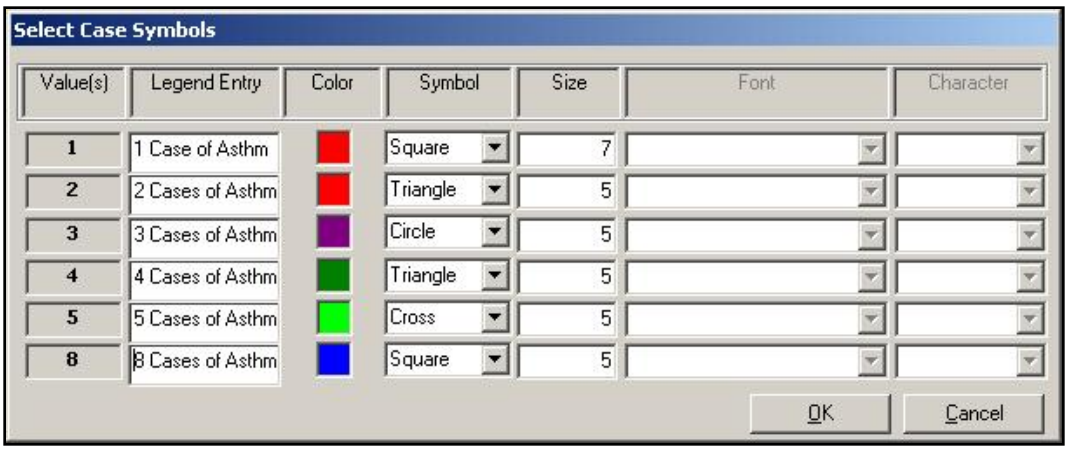
12. From the CountAsthma columns section, select **Zip**.

13. From the Classification Fields section, select **Count**.

14. Click **OK**. The Select Case Symbols dialog box opens.

15. You will type in the Legend Entry field. The numbers in the Legend column represent the number of cases per zip code. For the number 1, type the text **1 Case of Asthma**, then **2 Cases of Asthma**. You are creating a legend title for each number in the column.

Your legend titles should look like the following:



- Look at the color column. Are there any colors too similar to the choropleth color ramp? For example, if you have a dark blue field in your choropleth layer map, you might want to select an orange or light color to represent the case shapes. If you have similar colors in two blocks, you may want to change one.

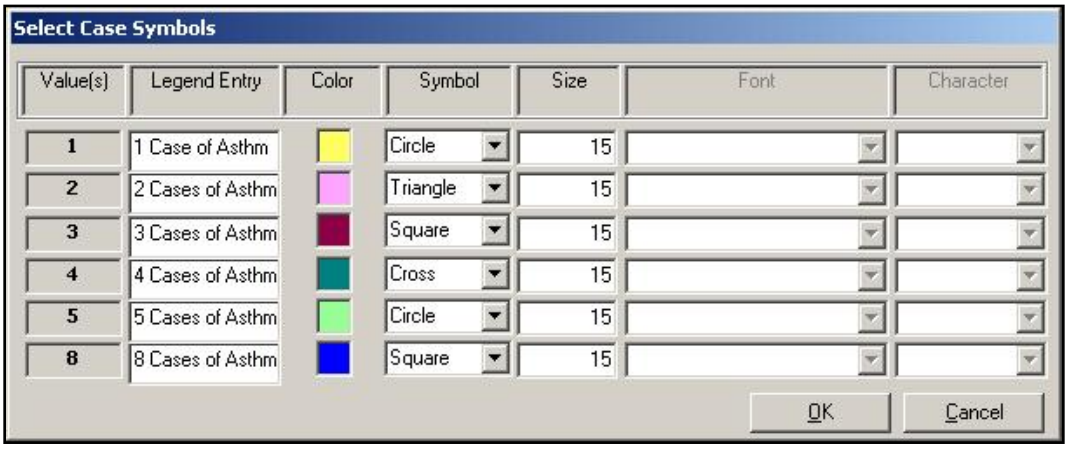
16. To change colors, click the color block and select a new color from the Color dialog box. Cases 1 and 2 both have red colors. Change one of them.

17. Use the Symbol drop-down to select shapes for each item. These shapes will populate the legend and the map.

- For this example, do NOT select True Type from the Symbol drop-down. If you wish to try different symbol types, select True Type and then from the Font drop-down, select Wing Dings, or another symbol-based font. You would then select a character from the drop-down to replace the geometric symbols in use.

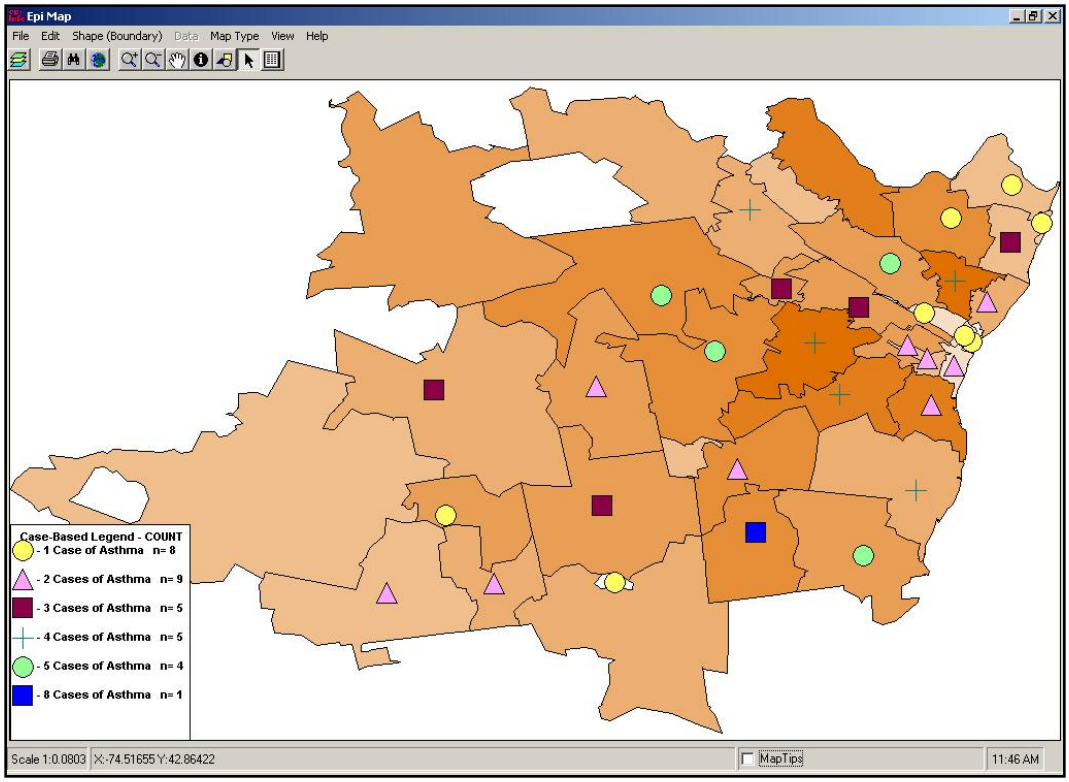
18. For the symbol size, select **15** for all.

Your Select Case Symbols dialog box will look similar to the following:



19. Click **OK**. The map populates with the selected symbols and legend. Close the **Map Manager** to see the entire map.

Your map should look similar to the following:



- The map legend tells you which symbol stands for the number of cases in a given zip code. The n= number tells you the total number of zip codes containing that number of cases. For example, the first symbol is equal to one case of asthma, and there are eight zip codes with one case of asthma.

- You can move the legend by clicking and dragging it. The map graphics may move a little during this process. Place the Case-Based Legend on the lower left corner of your map, and the choropleth Median Income legend on the lower right corner of the map.

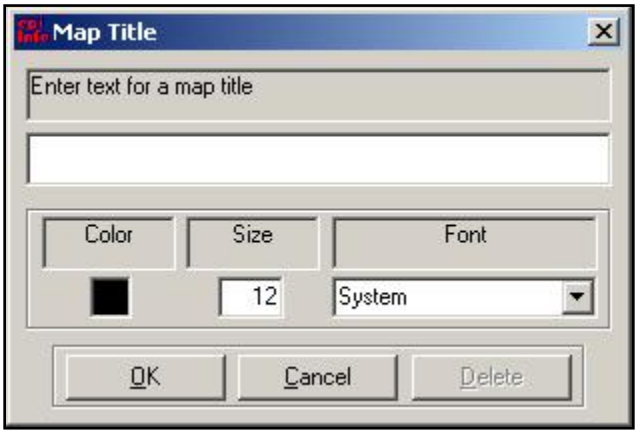
- Remember that the case-based information was placed on top of the income choropleth. Does the map indicate a correlation between income and asthma in the survey data?

Adding Text to a Map

You will create a title for the asthma and income map.

Got text? Let's put some into the map.

1. From the Epi Map main navigation menu, click **View>Title**. The Map Title dialog box opens. It may take a few seconds for the text and title dialog boxes to appear.



2. In the Enter Text field, type **Asthma Cases by Zip Code and Income**.

3. From the Font drop-down, select a font.

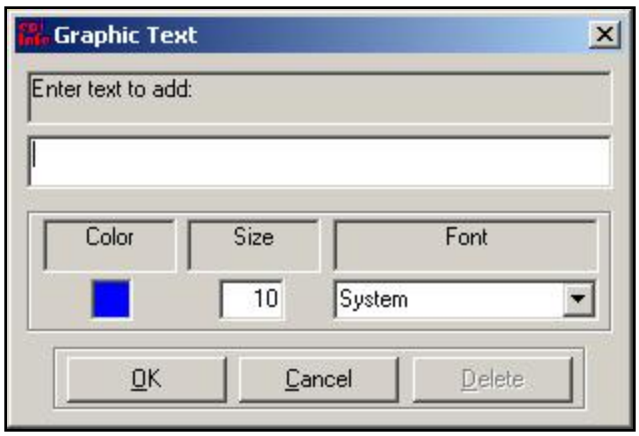
4. Highlight the **Size** field and type **18**.

5. Click **OK**. The title appears on your map in the top center of the page.

6. To add additional text to a map, select **View>Graphics**, or click the  **Graphics** icon in the tool bar. The Graphics option boxes appear:

7. Select the letter **A** icon for text insertion. The cursor will turn into a crosshairs symbol.

8. Click anywhere on your map. The Graphic Text dialog box opens.



9. In the Enter Text field, type **Asthma Survey 2004–2005**.

10. From the Font drop-down, select a **font style**.

11. Make the text size **12**.

12. Click **OK**. The text appears on your map.

- To move your text or your title to a new location on the page, click the **Arrow** icon from the Graphics Tool Bar. The cursor turns into an upright arrow. Light green crosshairs appear in the bottom center of your text fields.
- To move text, you must line up the arrow with the bottom of the green crosshairs. A box appears around the cross and the crosshairs turn red. Click, hold, and drag the box to move your text to a new location. This may take some practice.
- To delete or edit text, double click the **green cross hairs**. The cross will turn red so you know it is selected. The Text box appears.
- Click **Delete** to remove the text or edit the text field, size, font, or color of the text.
- Click **OK**.

13. Move the title and the text to a new location on the map.

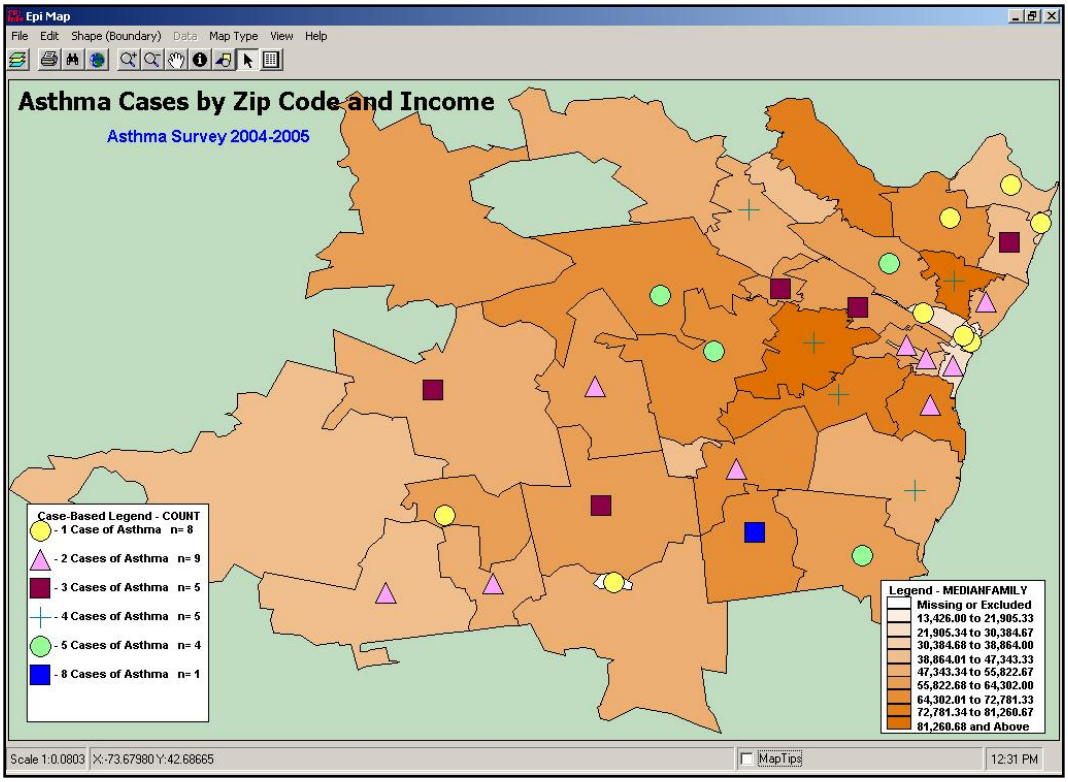
14. Use Map Tips or the Information icon to find out which zip code contains eight cases on asthma.

15. Open the **Map Manager**.

16. Click the **General** tab.

17. Select a color block to add a background color to your map.

Your map should look similar to the following:



18. Close the **Map Manager**.

19. Save a new .MAP file with the name **CBAsthmaIncome.**

20. Close **Epi Map**.

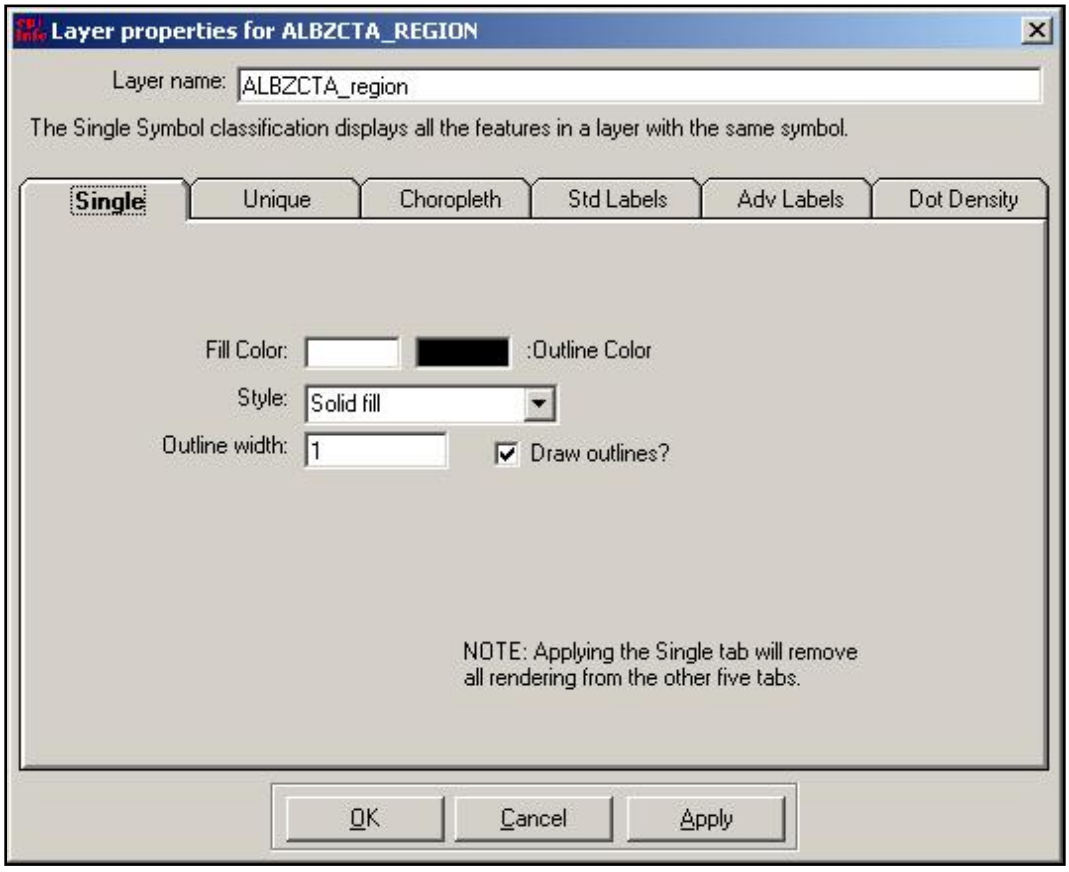
Creating a Point Map

Point maps can be made if you have geographic data for the X and Y longitude and latitude measurements of your map corresponding to the region. X and Y coordinates are developed with GIS tracking solutions and are not available for every map or location. They are mainly used for street maps and locations, or to pinpoint specific areas.

For this map, you have X and Y coordinates related to school and hospital locations in the region. In this lesson, you will add a fill color to a map, add points, and apply a shape polygon to designate where your county crosses zip codes and boundaries. This map provides the location of school and hospitals in the surveyed region.

Get the point? Create the map.

1. Open **Epi Map**.
2. Open the **Map Manager**.
3. Click **Add Layer**. Locate the ALBCZTA_region.shp file in your EIHA Tutorial folder.
4. Click **Properties**. The Layer Properties dialog box opens.



5. Click inside the **Fill Color** block. The Color dialog box opens.

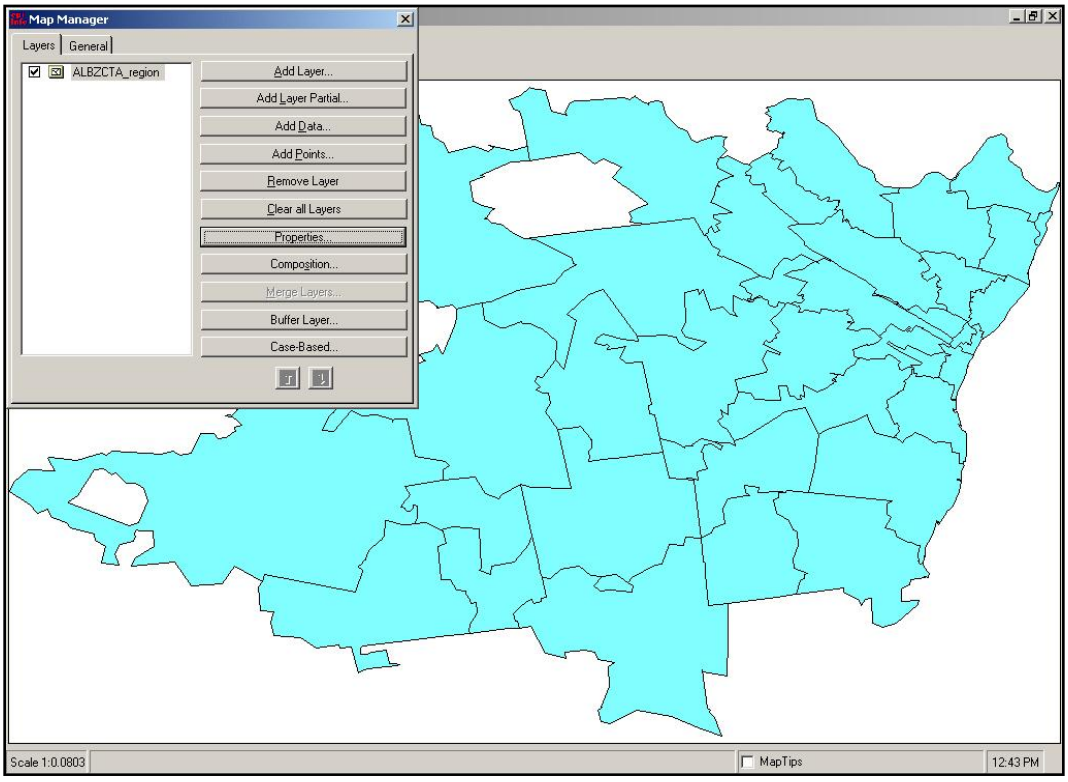
6. Select a Fill color for your map. We recommend a light color.

- Be sure the drop-down Style menu reads Solid Fill.

7. Click **Apply**.

8. Click **OK**.

Your map should look similar to the following:



9. Click **Add Points**. The Open Data Project window opens.

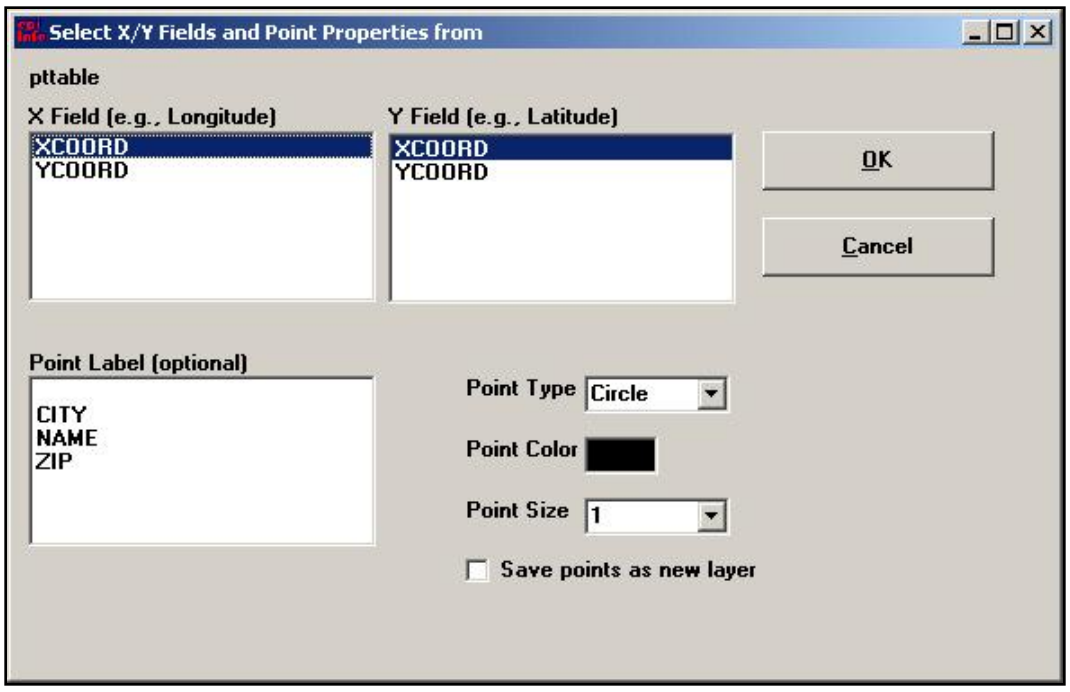
10. From the EIHA Tutorial folder, select the file **SchoolPoints.mdb**.

- The file School Points.MDB was created using data from a map .DBF file. The file was imported into Analysis and written out to a new Epi Info™ project file to be used in Epi Map. The same process was used to create the project file used for the hospital point data.

11. Click **Open**. The Select Data Table dialog box opens.

12. Select the table **pttable**.

13. Click **OK**. The Select X/Y Fields and Properties dialog box opens.



14. In the X Field Longitude section, select **XCOORD**.

15. In the Y Field Latitude section, select **YCOORD**.

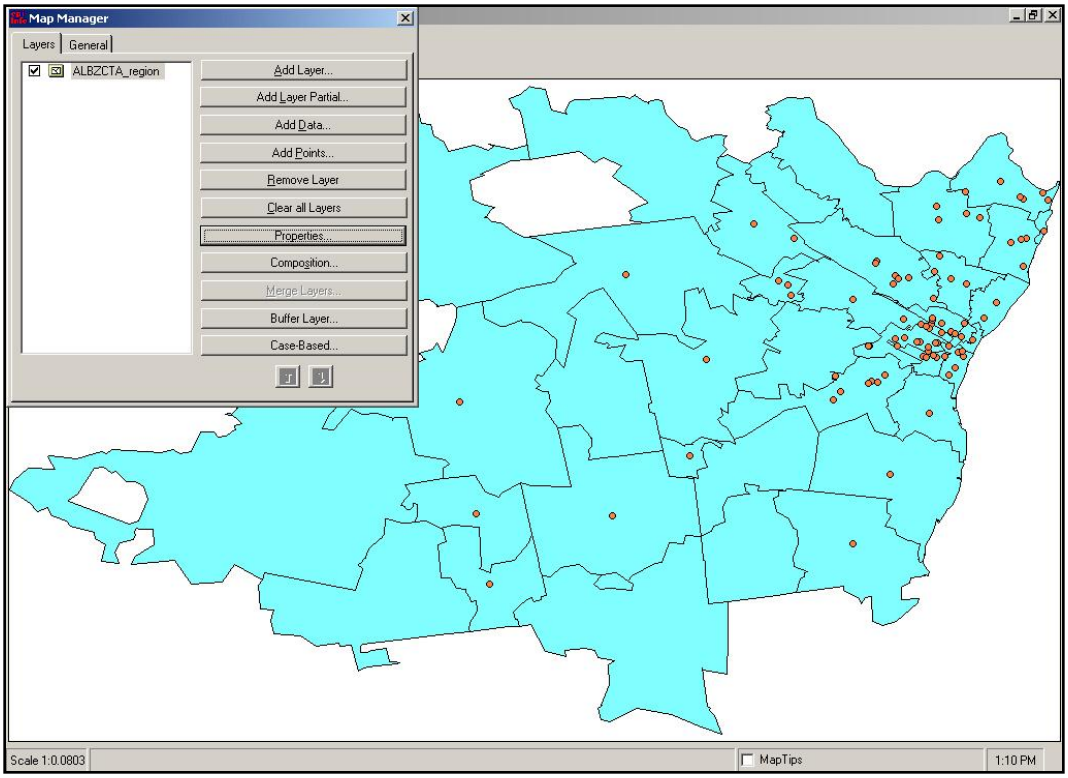
16. From the Point Type drop-down, select **Circle**.

17. Click inside the Point Color block to select a color that will show up on the fill color you selected for the map.

18. From the Point Size drop-down, select **5**.

19. Click **OK**.

Your map should look similar to the following:



- The points added to the map are considered a new layer, but are not editable.

Add the Hospital Points to the Map

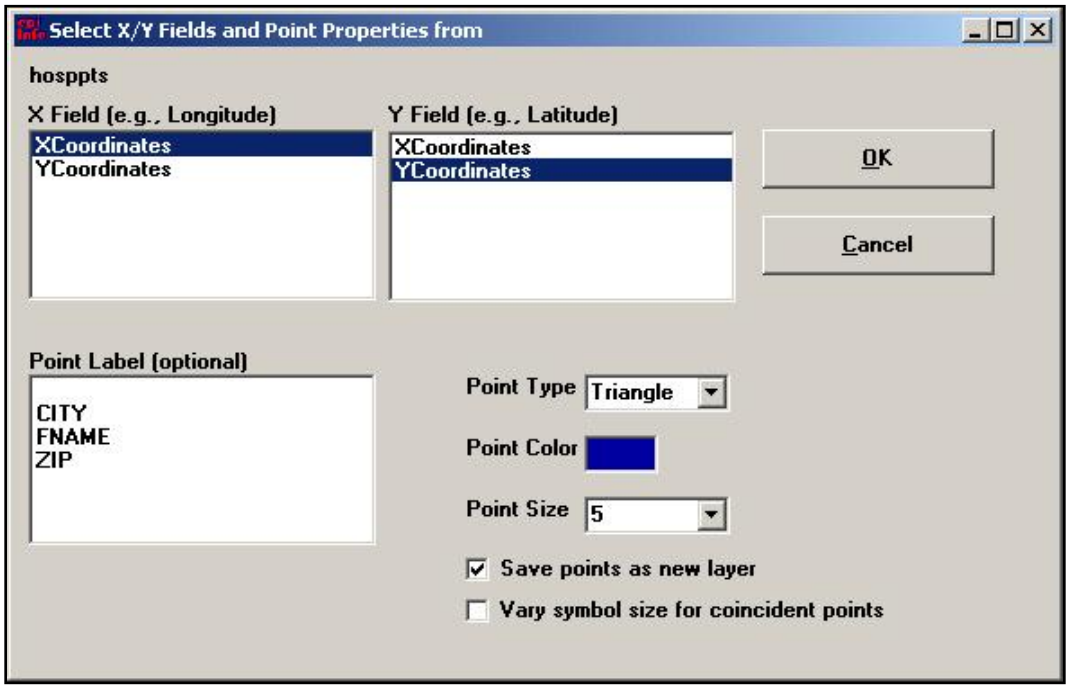
Point it out!

1. Click **Add Points**. The Open Data Project window opens.
2. From the EIHA Tutorial folder, locate and select the file **HospitalPoints.mdb**.
3. Click **Open**. The Select Data Table window opens.
4. Select the table **hosppts**.
5. Click **OK**. The Select X/Y Fields and Point Properties dialog box opens.
6. From the X Field section, select **XCoordinates**.
7. From the Y Field section, select **YCoordinates**.
8. From the Point Type drop-down, select **Triangle**.
9. Select a new **Point Color**.

10. From the Point Size drop-down, select **5**.

11. Select the checkbox **Save points as new layer**.

- When you save the points as a new layer, they will appear in the Map Manager as an editable layer.

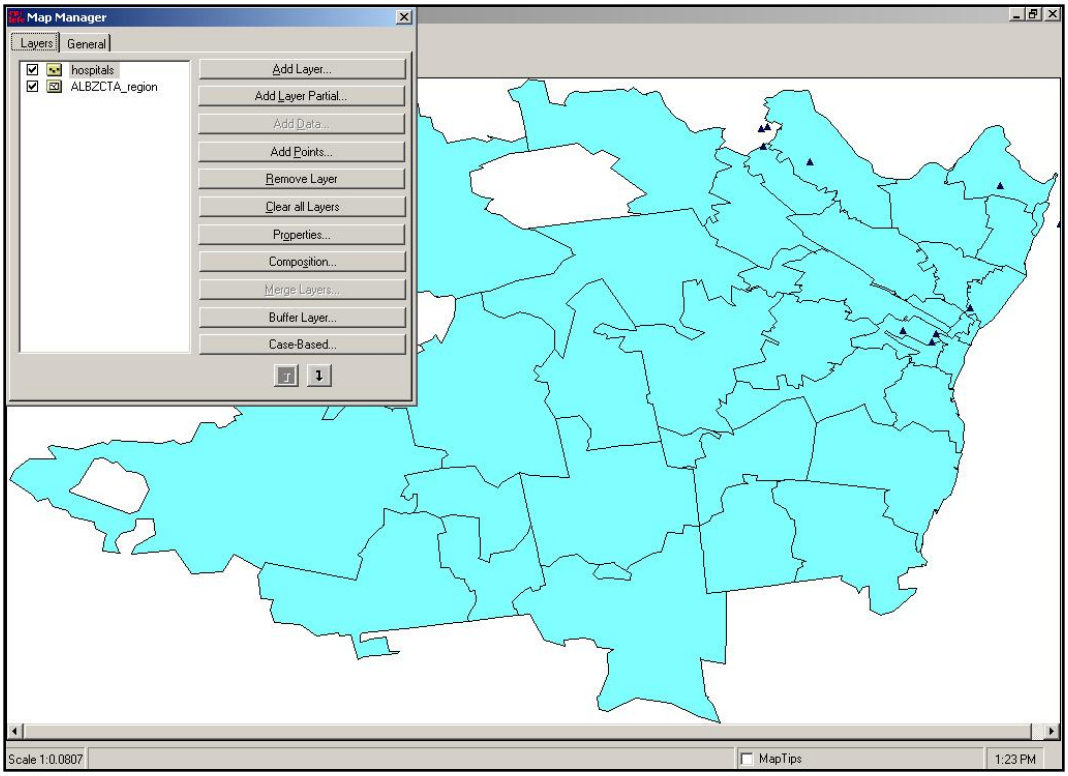


12. Click **OK**. The Save Points as Shapefile window opens.

13. Name the new file **Hospitals**.

14. Click **Save** to populate the map.

- Your map should look similar to the map below:
- Notice that the Map Manager now contains a new layer called hospitals.
- In the list of layers, hospitals needs to be the top layer. If it is not, highlight it and click the **up** arrow inside the Map Manager.



- Now select the **ALBZCTA** layer. Both school and hospital points should be on the map.

Editing Shapefile Points

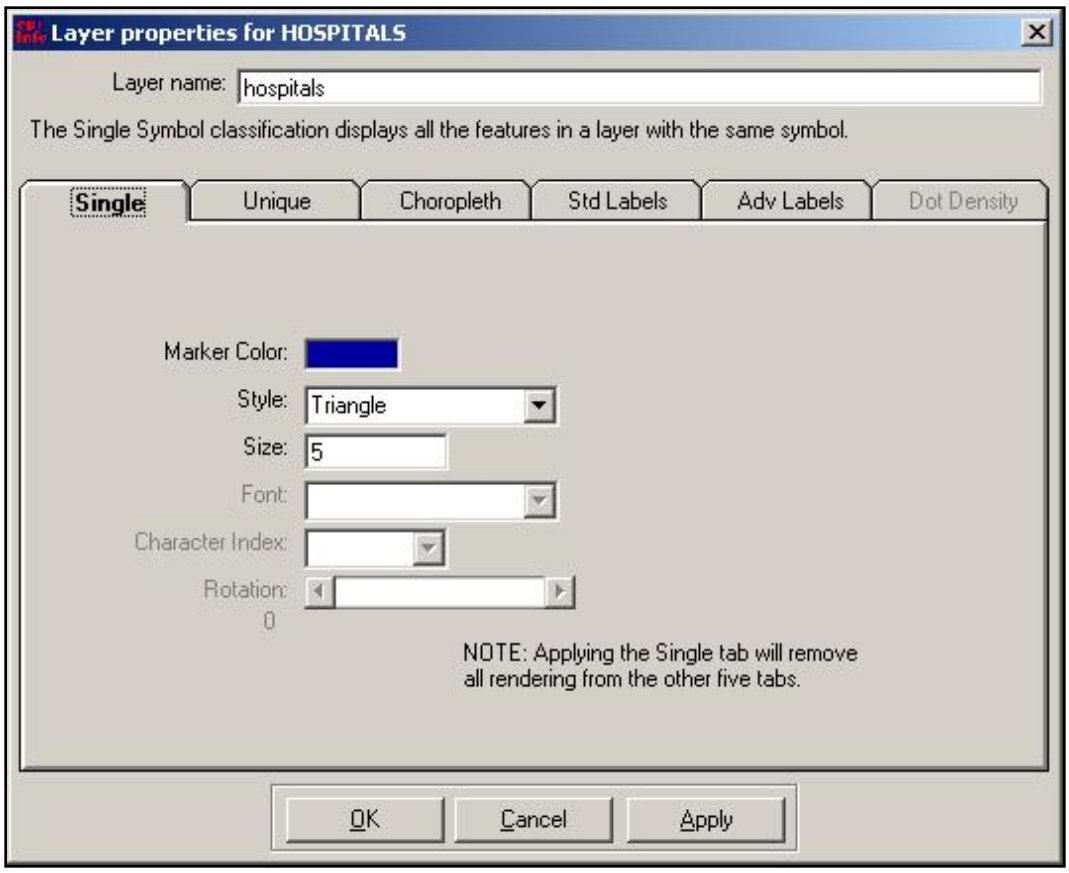
The triangles that represent the hospitals are a little too small. Since the hospital points are on separate layer and a separate shape file, they can be edited.

Get those points in shape!

1. From the Map Manager, highlight the **hospital** layer.

2. Click **Properties**. The Layer Properties dialog box opens.

   - From the Single tab, you can edit the point size and color of the hospital data points.



3. Highlight the number **5** in the size field.

4. Type in the number **12**.

5. Change the Marker Color, if needed.

6. Click **Apply**.

7. Click **OK** to close the Layer Properties dialog box.

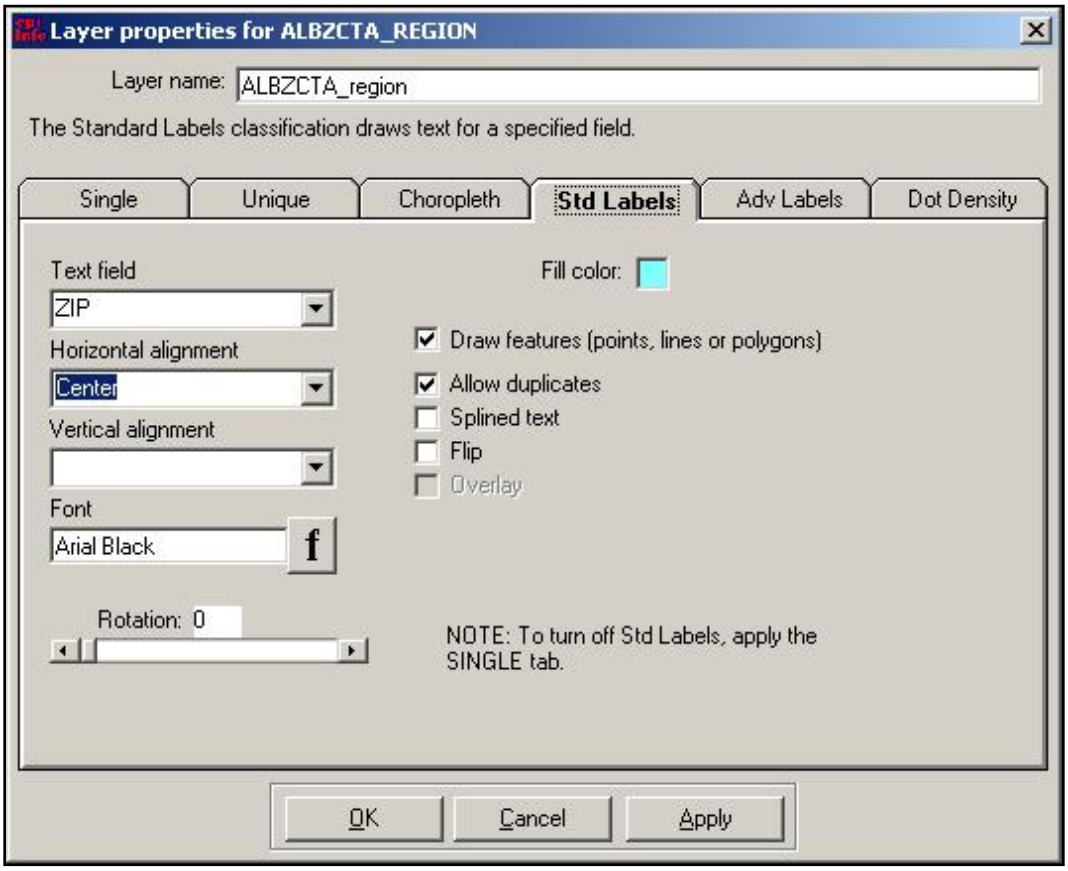
8. From the Map Manager, click the **ALBZCTA** layer to show all the points in your map.

Applying Standard Labels to a Map

You can use the data associated with the map files to place labels on the map. The Standard Label classification draws text for a specified field. In this section, you want to add zip code labels to the map.

Here's a chance to use your standard approach.

1. From the Map Manager, select the layer **ALBZCTA**.
2. Click **Properties**. The Layer Properties dialog box opens.
3. Click the tab **Std Labels**.



4. From the Text Field drop-down menu, select **Zip**.
5. From the Horizontal Alignment drop-down, select **Center**.
6. Click the **Font** button. The Font dialog box opens.
7. From the Font section, select a **new font**.

8. From the Font Style section, select **Regular**.

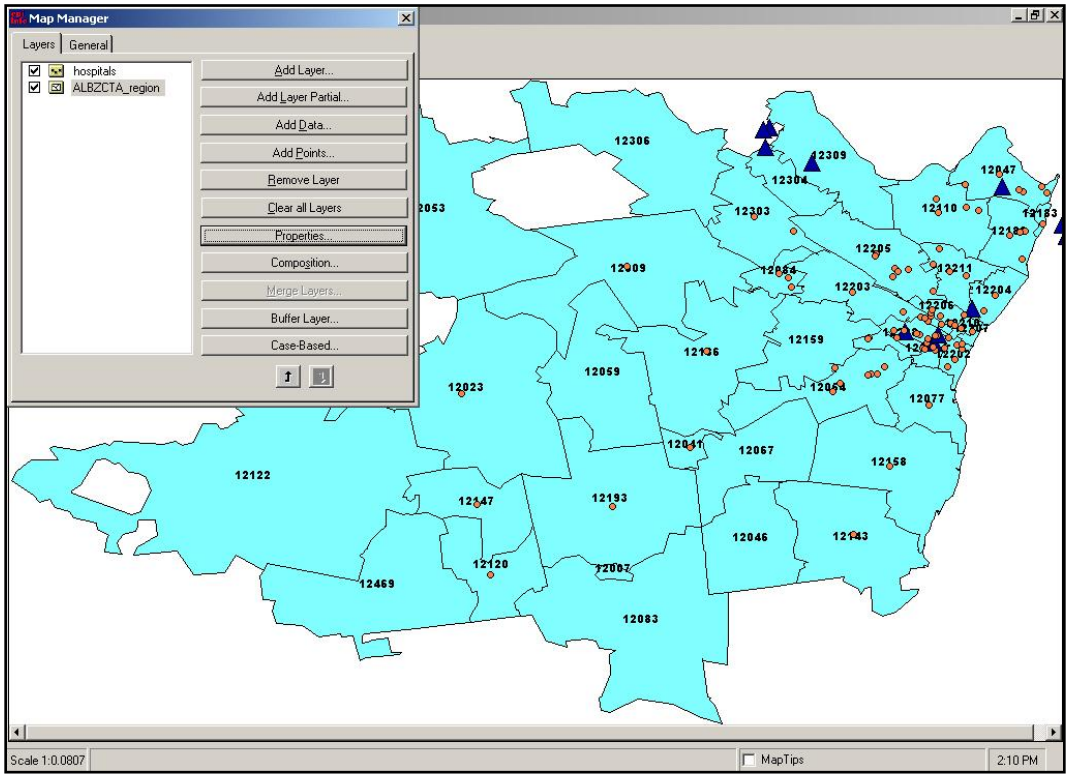
9. From the Font Size section, select **8**.

10. Click **OK** from the Font dialog box.

11. Click **Apply**.

12. Click **OK**.

Your map should similar to the following:



13. Remove the zip code labels by opening the **Properties** dialog box and clicking the **Single** tab.

14. Click **Apply**.

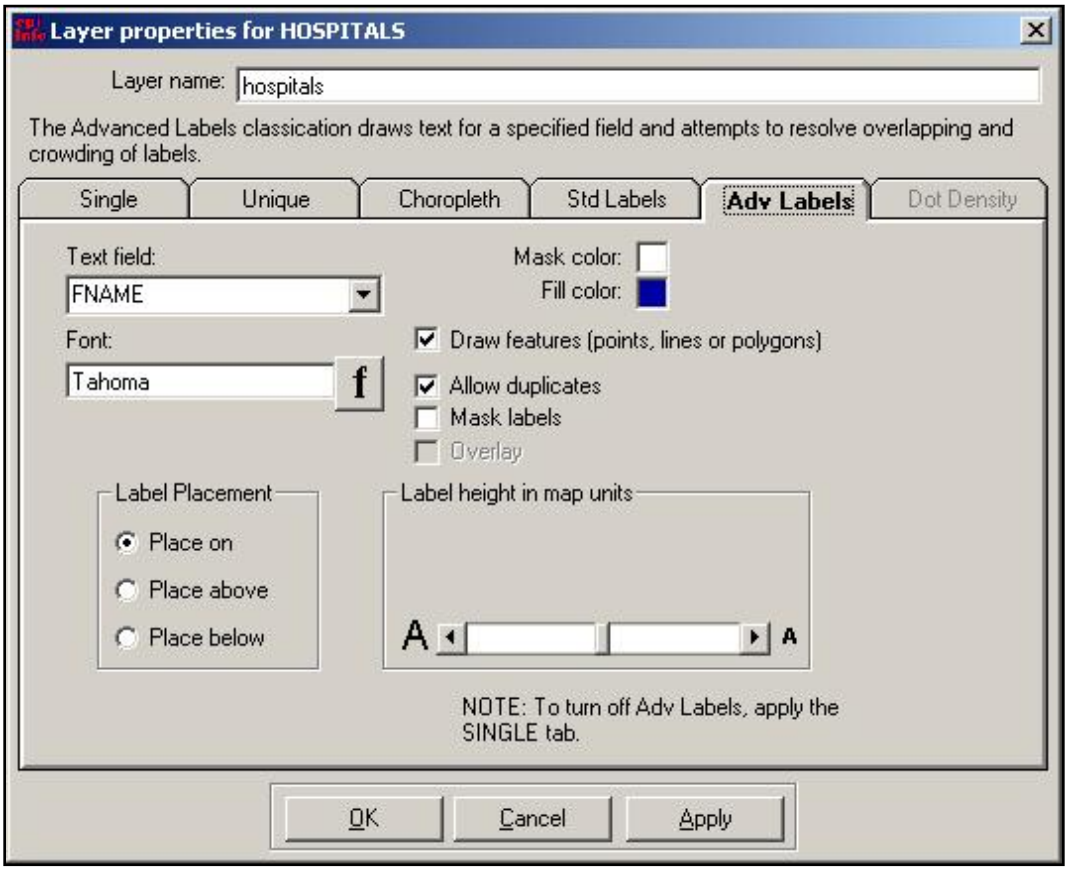
15. Click **OK**.

Applying Advanced Labels to a Map

The Advanced Labels classification draws text for a specified field and attempts to resolve overlapping and crowding of labels. In this section, you will add the hospital point names to the map.

Time to advance.

1. From the Map Manager, select the layer **Hospitals**.
2. Click **Properties**. The Layer Properties dialog box opens.
3. Click the **Adv Labels** tab.



4. From the Text Field drop-down menu, select **FNAME**.
5. Click the f **Font** button. The Font dialog box opens.
6. From the Font section, select a **new font**.
7. From the Font Style section, select **Regular**.

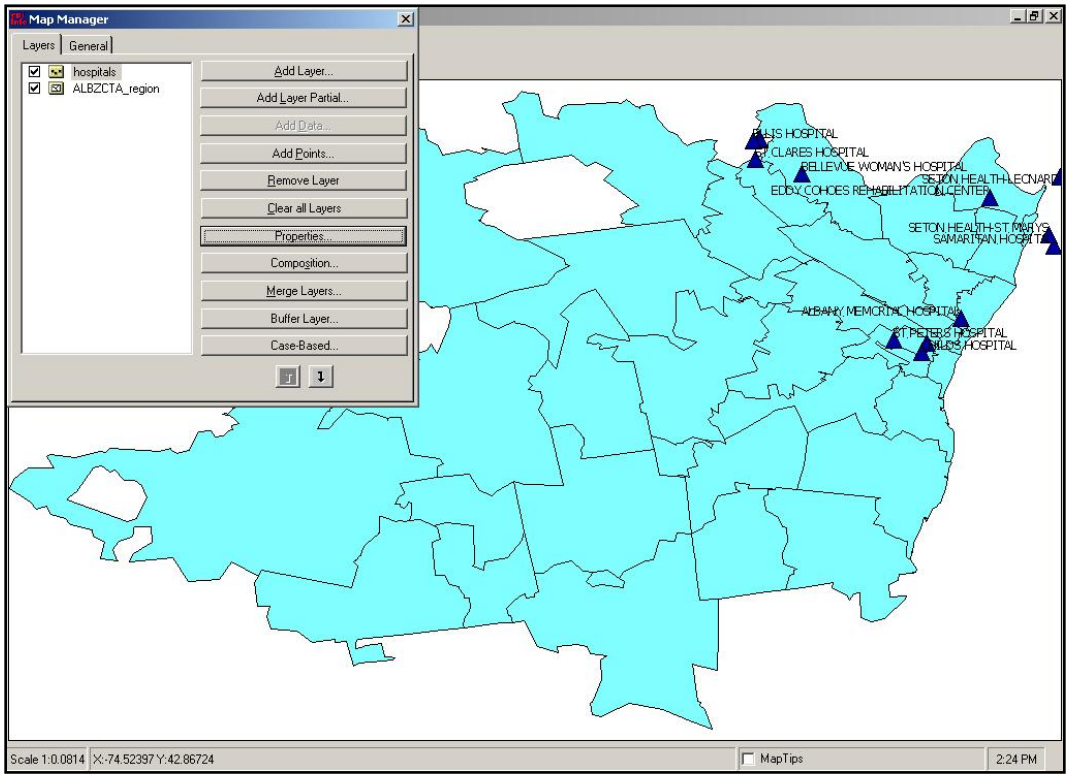
8. From the Font Size section, select **8**.

9. Click **OK** from the Font dialog box.

10. Click **Apply**.

11. Click **OK**.

Your map should look similar to the following:

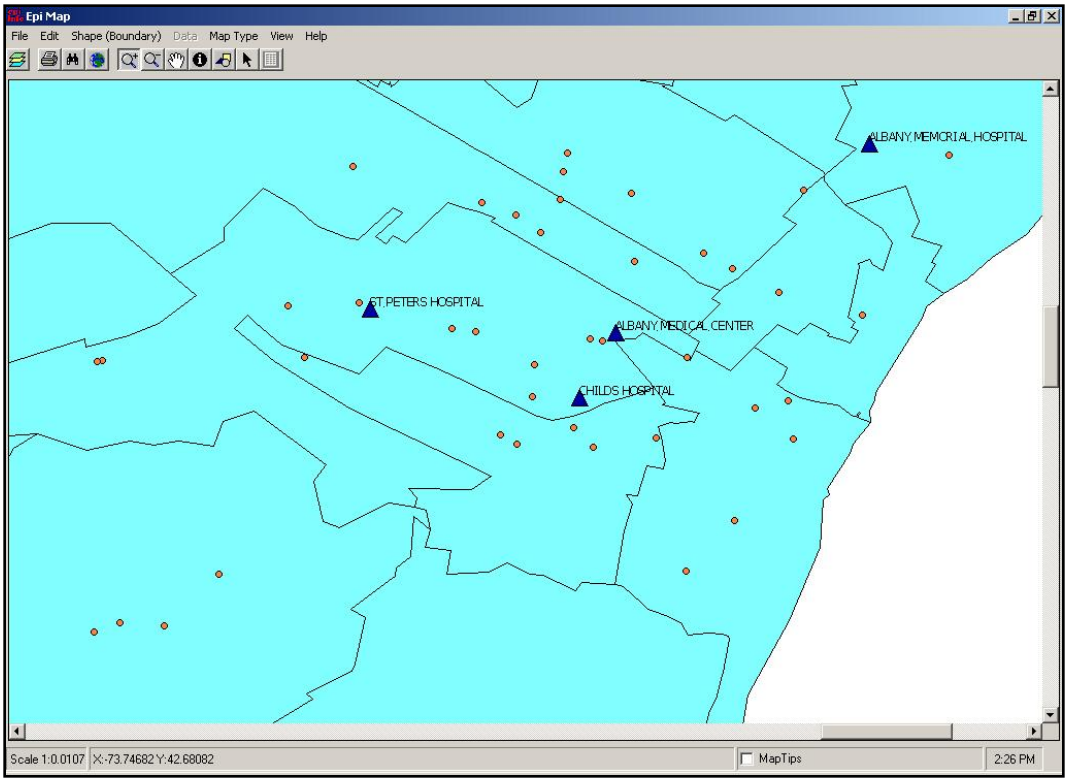


12. Select the **ALBZCTA** layer to activate it.

13. Close the **Map Manager**.

14. Click the  **Zoom** button. The cursor turns into a magnifying glass.

15. Use the click and drag method to **draw a square** around the Albany Memorial Hospital region. A close-up view of the map appears.



16. Click the  **Pan** button. The cursor turns into a hand.

17. Use the click and drag method to move around the map. Click inside the map and use the hand to move the map and locate point areas that fall outside the map region.

18. Use the **Zoom Out** button, or the **Full Extent** button, to return the map to the center of the workspace.

19. Open the Map Manager and select the **Hospital** layer.

20. Open the Properties dialog box. Click the **Single** tab to remove the hospital labels by

21. Click **Apply**.

22. Click **OK**.

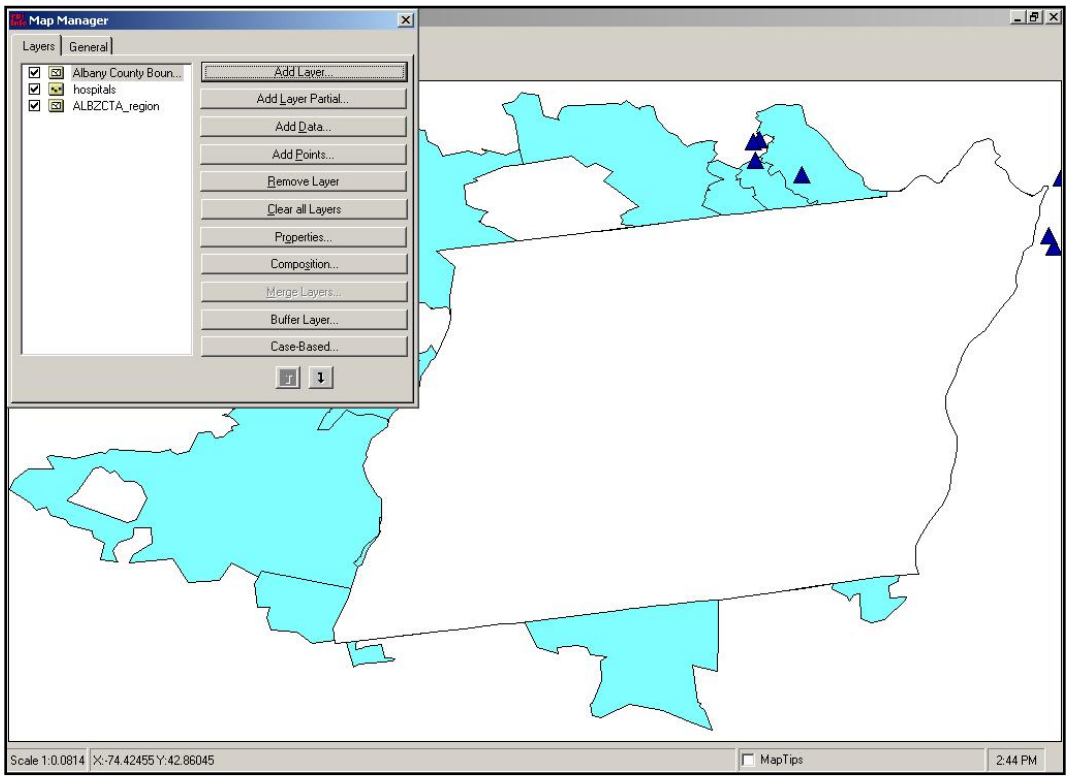
Adding a Shape Layer

The data and zip codes contained in the survey information crossed county lines. To see how this affects the locations of the schools in the survey, you can place the county shape over the region shape.

Time to get in shape!

1. From the Map Manager, click **Add Layer**. The Add Layer window opens.
2. Select the file called **Albany County Boundary_region.SHP**.
3. Click **Open**. The Albany shape appears on top of the region shape.

   The map should look similar to the following:



4. Click **Properties**. The Layer Properties dialog box opens to the Single tab.
5. From the Style drop-down, select **Transparent Fill**.
6. In the Outline Width box, type **2**.
7. Click **Apply**. The outline around the Albany Boundary file will be thicker while the inside of the shape appears transparent.
8. Click **OK**.

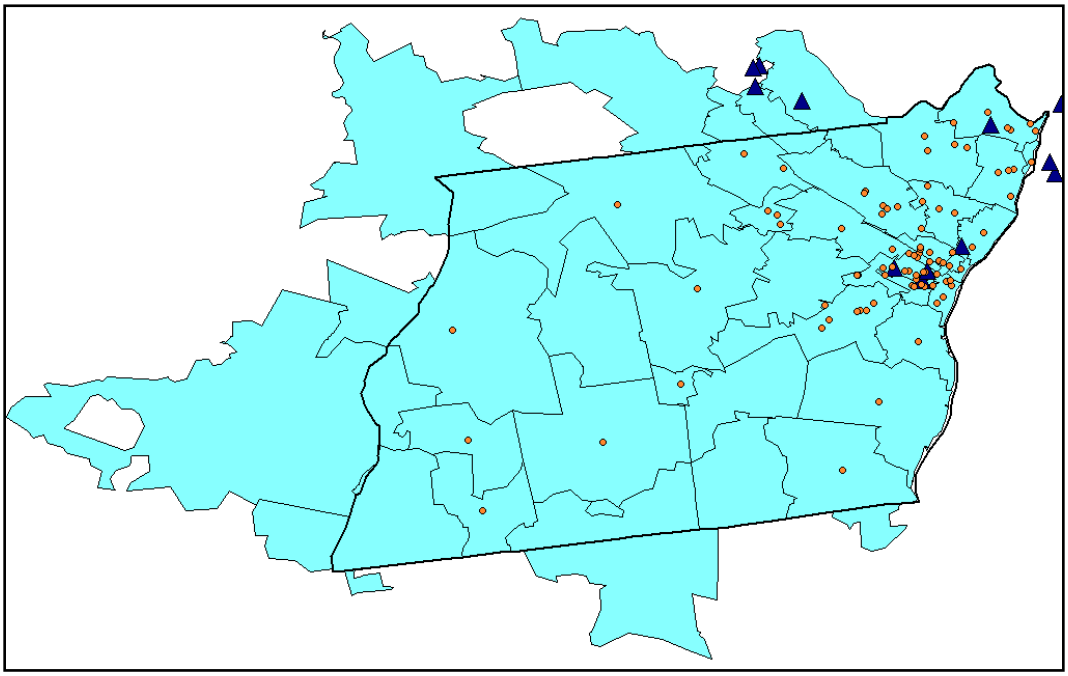
9. From the Map Manager, select the **ALBZCTA** layer to view all the points.

10. Close the Map Manager.

11. Turn on Map Tips.

12. Select **Zip** from the drop-down and hover over the region to see which zip codes cross county boundaries.

The map should look similar to the following:



Saving a Map as a Bitmap File

Let's chomp at the bit.

If you want to save a map to be used in presentations or documents, save it as a bitmap file.

1. Select **File>Save as Bitmap File**. The Save As window opens.

2. Name the file **SchoolPointsMap**.

3. Click **Save**.

4. Close **Epi Map**.

🕮 Skills Review

Time to put your skills to use and create a new map based on the number of children with bronchitis in the survey. Use the CountBronchitis table you created in the Lesson 8 Skills Review.

Open **Epi Map**.

Use the Map Manager to add the layer ALBZCTA_region.shp to the workspace.

Use Add Data to plot the table called CountBronchitis. The table is located in the Asthma Survey 2005.mdb. This is one of the frequency tables you created in Lesson 8.

In the Select Relate Fields and Render Fields dialog boxes respectively, select the options **Zip**, and **Zip and Count**.

Click **Continue** at the Incomplete Join window. A choropleth map of the bronchitis count appears.

From the Map Manager, select **Properties**.

Select the **Single** tab.

Click **Apply** to remove the choropleth settings.

Select the **Unique** tab. The Unique Values classification displays features by applying a symbol to each unique value for a specified field. In this example, it will display a unique color for each value of Count.

From the Field drop-down, select **Count**.

Click **Reset Legend**.

Select the **Overlay** checkbox.

Click **Reset Legend**.

Click **Apply**.

Select the **Advanced Labels Tab**.

From the Text Field drop-down, select **Zip**.

Use the Font options to make the font size **8**.

Click **Apply**.

Click **OK**.

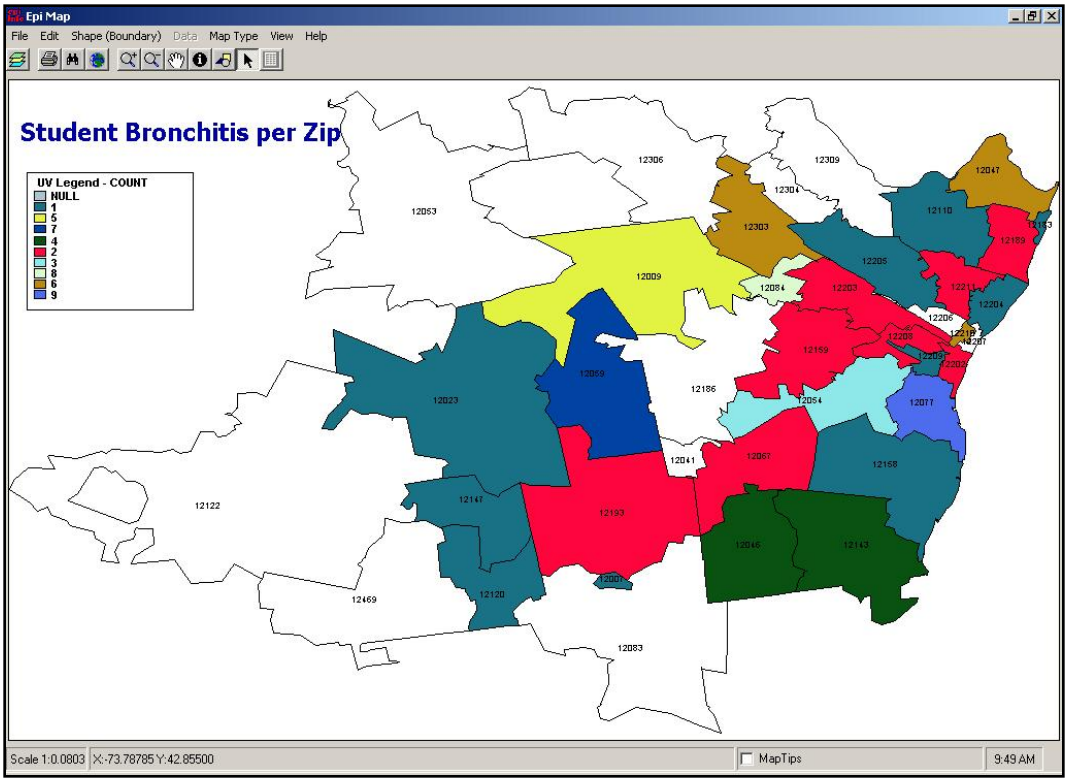
Close the **Map Manager.**

Add the title **Student Bronchitis per Zip**.

Arrange the title and legend around the map as you prefer.

Save the map as a .MAP file titled **BronchitisZips**.

Your map will look similar to the following:



Lesson Complete!

WHAT YOU LEARNED

How to:

- Create a choropleth map.
- Create a case-based map.
- Save maps as **.MAP** and **.BMP** file types.
- Create a point map.
- Edit maps using the Properties dialog box.

This page intentionally blank.

WORKSHOP

Intermediate Analysis

SUMMARY

This lesson pulls together many of the skills you learned in Lessons 1–10 while adding more advanced data management concepts. Many skills learned in the Basic tutorial will be reinforced through the Intermediate Analysis Lesson including, DEFINE, ASSIGN, LIST, DISPLAY, READ, WRITE, SELECT, GRAPH, and MAP. You will also be introduced to a new command called RELATE.

This lesson introduces ways you can use Epi Info™ to create everyday statistics needed in epidemiologic work. Use the Tear Out Answer Sheet and To Do List to record your results as you progress through the lesson. Many of the sections heavily emphasize data management skills needed to combine a variety of data sources for analysis.

Length of time to complete: 3 hours Intermediate/Advanced

Scenario

This presentation reviews the tasks learned in the Intermediate Analysis Lesson. In this lesson, you will prepare and analyze data that could be used to write a proposal requesting legislative funds for an asthma initiative. Introduction to Advanced Statistics and the RELATE command.

BRFSS Data

Focus on using data downloaded from the BRFSS website to compute lifetime and current prevalence rates. Introduction to Complex Sample Frequencies.

Primary Data

Use 800 survey records to complete a Frequency stratified by school.

Census Data

Focus on using data downloaded from the U.S.Census website in conjunction with the original 800-record dataset to compute the county child asthma prevalence.

Mortality Rates

You will create a RELATE between a denominator and numerator table to compute the asthma mortality rate for the county and state. You will also graph the rate.

10-Year Trends

You will create a graphic representation of 10-year hospitalization discharge rates based on age groups.

3-Year Rates

You will create comparison graphs for 3-year hospital discharge rates for state and county by age.

Mapping Rates

Map asthma hospitalization rates by zip code to determine high risk areas.

Creating a Scatter Plot

You will create a scatter plot to determine if income affects the rate of asthma hospitalization.

Risk Ratios

Create a program (.PGM) and routeout file to compute asthma risk ratios for a series of zip codes. Use the DIALOG command to create a prompt for users.

Using StatCalc

You will create an estimate of sample size.

BEFORE YOU BEGIN

Complete all 10 beginner and intermediate lessons.

WHAT YOU NEED

Asthma Final.MDB

Tear Out Answer Sheet and To-Do List

Appendix F Intermediate Analysis Answer Key

BRFSS2003 Data Table

CensusAge.HTM

Population Data Table

Numerators Data Table

NYDisAge10Yr Data Table

Age Proxy Data Table

DisCtyAge Data Table

NY3Yr Data Table

Dis0103byZip Data Table

IncomeData Data Table

RatebyZip Data Table

ALBZCTA_region.SHP

ALBZCTA_region.DBF

ALBZCTA_region.SHX

ALBZCTA_region.PRJ

DischargesbyZip Data Table

FIVE GOALS

- Understand data management fundamentals involved in using Epi Info™ with a variety of data sources.
- Understand the use of the RELATE command and the ways tables can be linked to create rates.
- Use the DEFINE and ASSIGN commands to create new variables and assign values to them.
- Locate prevalence rates and confidence limits from table results.
- Use the SELECT and IF/THEN commands to specify subsets of data that needs to be managed or analyzed.

Scenario

The County X LHD director informs the county administrator/executive about available discretionary legislative funds for a health initiative. There is a good chance of receiving the funding for asthma as the LHD has already done some work in the area, and can produce results with additional funds. The director meets with the health education coordinator/planner to identify essential data to write a strong proposal.

A series of questions and datasets have been developed to create the proposal. Answers to the questions in the form of data tables, graphs, and maps can be used to illustrate the need for asthma funding, and for writing the proposal. This lesson focuses on preparing data for the proposal, but does not cover creating the actual report as Lessons 1–10 did.

Datasets used in this lesson come from a range of sources and will show you the many ways Epi Info™ can be used to manage data sources. Questions deal with the prevalence of asthma across subgroups and regions as well as hospital discharge and mortality rates. For this lesson, you will use data downloaded from the BRFSS website, the U.S. Census website, CDC WONDER, and SPARCS.

For more information on gathering data and preparing to work with Epi Info™ and Epi Map, see Appendix G: Preparing Data for Use in the Epi Info™ Tutorial.

For more information on the provided data sources and web links to guide you, see Appendix I: Data Sources.

To Do List for Intermediate Analysis Lesson

Expected results and sample .PGMs are listed in Appendix F Intermediate Analysis Answer Key.

- ☐ Familiarize yourself with the scenario. You are developing data to aid in the writing of a proposal to request asthma funding.
- ☐ Read through the Advanced Statistics section and the information on using the RELATE command.
- ☐ Read the What is BRFSS information.
- ☐ Read the How to Get BRFSS information.
- ☐ Use BRFSS data to compute lifetime and current asthma prevalence for the state.
- ☐ Use the 800 records in your primary data source to compute prevalence rates for each school in the survey.
- ☐ Familiarize yourself with the Sampling Scheme.
- ☐ Use data downloaded from the U.S. Census Bureau to create a table with the total population per age for the county (denominator table).
- ☐ Use the primary data source to create a table that contains information about the survey respondents' age and sex (numerator table).
- ☐ Create a weight variable for the county child data.
- ☐ Create a PSU variable for the county child data.
- ☐ Compute the county child asthma prevalence using the Complex Sample Frequencies command.
- ☐ Relate data from CDC WONDER, the primary data source, and New York hospital discharge data from SPARCS
- ☐ Use DEFINE and ASSIGN to calculate the asthma mortality rate.
- ☐ Create a bar graph of the mortality rates.
- ☐ Use GRAPH to create a line graph that illustrates the 10-year trends in hospitalization for asthma by age group.
- ☐ Use GRAPH to create a line graph for the state and a line graph for the county to illustrate 3-three asthma hospitalization discharge rates.
- ☐ Use MAP to plot hospitalization rates by zip code.

- ☐ Use GRAPH to create a scatter plot that illustrates how income affects the rate of asthma hospitalization.
- ☐ Use the TABLES command to create asthma hospitalization risk ratios between state zip codes.
- ☐ Use the StatCalc program to estimate a sample size for the number of records needed from a hospital sample to achieve a 95% confidence interval for your statistics.

Tear-Out Answer Sheet

Use this sheet to record your answers in the Answers column as you progress through the Intermediate Analysis Lesson. Answers are listed in Appendix F Intermediate Analysis Answer Key.

A. Questions in the lesson are marked with a letter

Computing Prevalence — Using BRFSS Data	**Results**
A. What is the lifetime asthma prevalence among adults in the state?	
B. What is the current asthma prevalence among adults in the state?	
C. Is the state adult asthma prevalence rate higher in the subgroup gender? **D.** Is the state adult asthma prevalence rate higher in the subgroup ethnicity?	Males- Females- Hispanics-

Primary Data — 800 Records	**Results**
E. From the school asthma survey data, what is the difference in prevalence rates?	School A- School B- School C- School D-

Census Data	**Results**
F. What is the county childhood asthma prevalence?	
G. What is the county childhood asthma prevalence for the subgroup gender?	Males- Females-

Mortality Data	**Results**

H. Is the asthma mortality rate for the state and county significantly different?	State 2000- State 2001- State 2002- County 2000- County 2001- County 2002-

10–Year Trends — Hospital Discharge Data	**Results**
I. Is the rate of hospital discharge for asthma increasing or decreasing in any age groups?	

3–Year Rates — Hospital Discharge Data	**Results**
J. Are there significant differences in the 3–year asthma discharge rates for the state versus the county by age?	

Mapping Rates	**Results**
K. Based on the map with hospitalization rates by zip code, which areas have the highest hospitalization rates for asthma?	

Scatter Plots	**Results**
L. Does income appear to affect the rate of hospitalization for asthma?	

Risk Ratios — Hospitalization Data	**Results**
M. What is the risk ratio of being hospitalized for asthma if you	

live in the zip code 12202 compared to zip code 12211?	
StatCalc	**Results**
N. From the New York hospital discharge data, how many records do you need to sample to get a 95% confidence interval from a sample size of 38,998?	

Advanced Statistics

The Complex Sample Frequencies (FREQ), Complex Sample Tables (TABLES), and Complex Sample Means (MEANS) commands in the Analysis program perform statistical calculations that assume data comes from simple random (or unbiased systematic) samples. In many survey applications, more complicated sampling strategies are used. These may involve sampling features (e.g.,stratification, cluster sampling, and the use of unequal sampling fractions). Surveys that include some form of complex sampling include the coverage surveys of the WHO Expanded Program on Immunization (EPI) (Lemeshow and Robinson, 1985) and CDC's BRFSS (Marks, et al., 1985).

The CSAMPLE functions compute proportions or means with standard errors and confidence limits for studies in which the data did not come from a simple random sample. If tables with two dimensions are requested, the odds ratio, risk ratio, and risk difference are also calculated. Data from complex sample designs should be analyzed with methods that account for the sampling design.

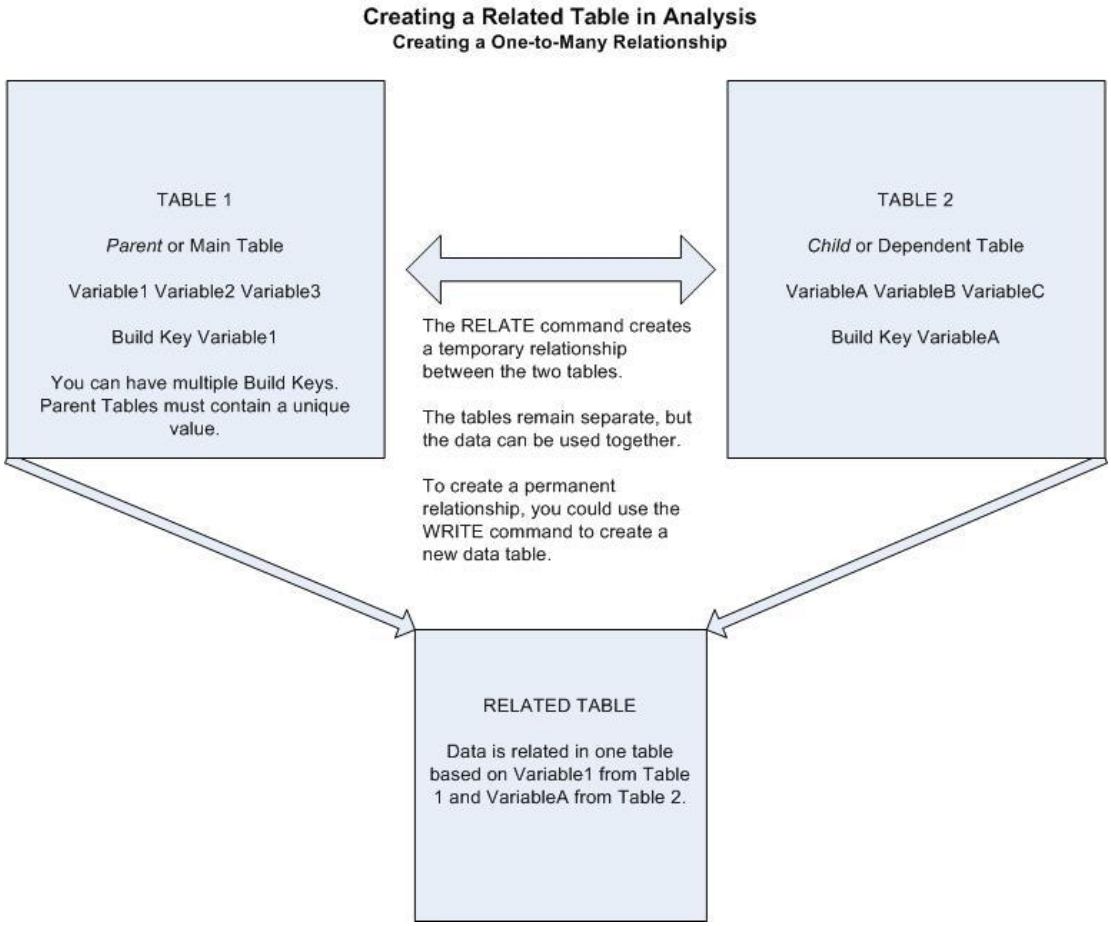
The following sections rely heavily on a new command called RELATE, which allows you to link one or more tables to the current dataset using a common identifier (key) to find matching records. The key can include several matching fields. The linked tables remain separate.

RELATE Command

To use RELATE, at least one table must have been made active with the READ command. The table to be linked must have a key field that has identified related records in the two tables. The keys in the main and related tables or files do not need to have the same name.

After issuing the RELATE command, the variables in the related table may be used as though they were part of the main table. If variable names are duplicated in the related tables, they will be suffixed with a sequence number.

Frequencies, cross-tabulations, and other operations involving data in both the main and related tables can be performed. You can also use the WRITE command to create a new table. More than one table can be related to the main table by using a series of RELATE commands.



What is BRFSS?

The Behavioral Risk Factor Surveillance System (BRFSS) is a state-based system of health surveys that generates information about health risk behaviors, clinical preventive practices, and health care access and use primarily related to chronic diseases and injury. BRFSS is a cross-sectional telephone survey conducted by state health departments with technical and methodological assistance provided by CDC. States conduct annual monthly telephone surveillance using a standardized questionnaire to determine the distribution of risk behaviors and health practices among non-institutionalized adults. The states forward the responses to CDC where monthly data are aggregated for each state. Data are returned to the states and published on the BRFSS website. BRFSS interviewers ask questions related to behaviors associated with preventable chronic diseases, injuries, and infectious diseases. States use standard procedures to collect data through monthly telephone interviews. (National Center for Chronic Disease Prevention and Health Promotion http://www.cdc.gov/brfss/faqs.htm#1)

How to Get BRFSS Data

BRFSS data is readily available and can be downloaded from the BRFSS website at http://www.cdc.gov/brfss/index.htm.

This section of the tutorial deals with year 2003 data. Data can be downloaded into a variety of formats. For this lesson, you will be working with data derived from a .DBF file. The original .DBF was extracted to a folder on the local drive.

The file was Read into Analysis. Analysis accepted the first 255 variables included in the dataset. The .DBF file contained more than 264,000 records, encompassing all the records for surveyed areas.

For this example, you are working with records from New York State. Records would be selected only from New York and and written to a new file.

To obtain only the information dealing with a specific topic, in this case asthma, visit the BRFSS website and obtain detailed information on which variables pertain to asthma along with the codes used to signify race, gender, ethnicity, etc. Once you know the variable names, you can select only the questions/variables related to asthma and write those to a new table. This lowers the record count to 5,544.

Because of how the BRFSS is conducted (multi-stage sample process) with variables that are not equally weighted, you will use the C-Sample (complex sample routines) instead of the usual tables and means commands. The BRFSS dataset contains variables for weight, Primary Sampling Unit (PSU), and stratification, already computed as part of the table.

For this example, because of the large size of the dataset, BRFSS data has already been downloaded and written out to a new file specifically dealing with asthma data. You can use the data table contained in the Asthma Final project to compute the prevalence of lifetime and current asthma in the state of New York. The project contains a table called BRFSS2003 for use in the following lesson.

Reading/Importing the BRFSS Data

To compute the prevalence using BRFSS data, you can use the Analysis command Complex Sample Frequencies. It can be used to determine the following prevalence rates required for your proposal.

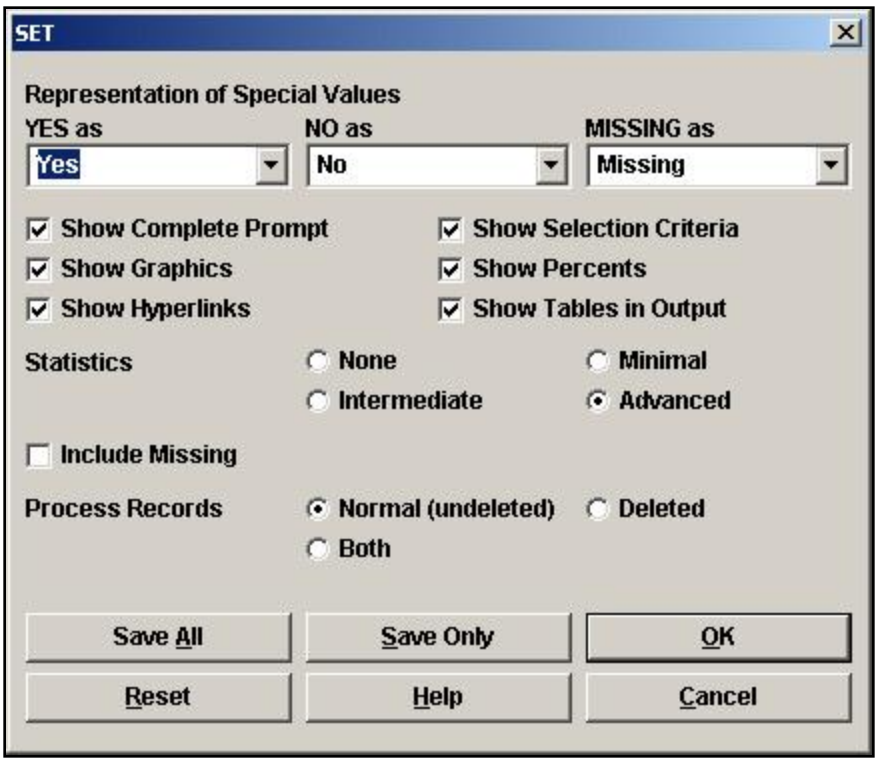
A. What is the lifetime asthma prevalence among adults in the state?

To get the prevalence for the county, you could use Expanded Behavioral Risk Factor Surveillance System (EBRFSS) data, if available. You can determine increase or decrease by following the same steps outlined below, but use prior year or later year data. You would need to download, select, and format data for the requested year from the BRFSS website.

Time to import the data.

1. Click **Analyze Data**. The Analysis module opens.

2. From the Command Tree Options folder, click **Set**. The SET dialog box opens.



3. From the Statistics section, select the **Advanced** radio button.

4. Click **OK**.

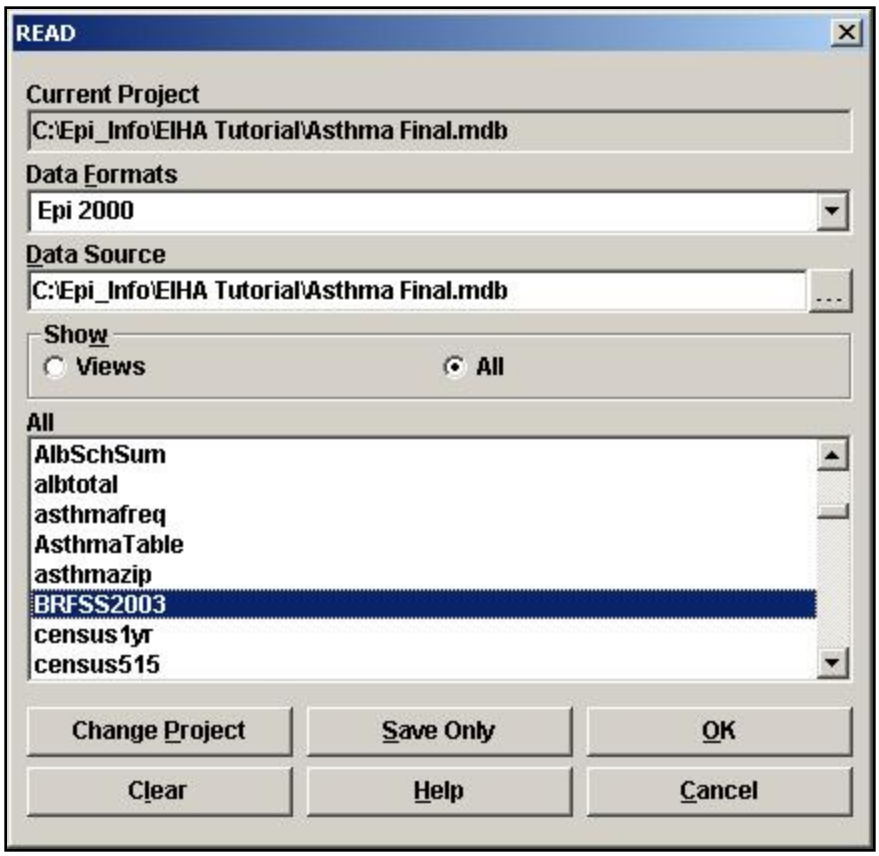
5. From the Command Tree Data folder, click **Read/Import**. The READ dialog box opens.

6. Click **Change Project**.

7. Locate the EIHA_Tutorial folder and select the project file called **Asthma Final.MDB**.

8. Click **Open**.

9. Click the **Show All** radio button. The All field populates with all the tables in the project.



10. Select the **BRFSS2003** table.

11. Click **OK**. You should have a record count of 5,544. This is the number of records for the state of New York.

12. From the Command Tree Variables folder, click **Display**. The DISPLAY dialog box opens.

13. Click **OK** to see all the variables in the table.

- These variables were selected out of the larger dataset and written to the new BRFSS2003 table. Information in this table pertains only to New York for the year 2003.

- By consulting the BRFSS website, you know that the calculated variable that contains information on the lifetime prevalence of asthma is called _LTASTHM. Since you imported the file as a .DBF, the variable name contains a prefix of A to compensate for the space field. The variable you need to figure the prevalence is called A_LTASTHM. This is a calculated variable based on the BRFSS question, "Have you ever been told by a doctor, nurse, or other health professional that you had asthma?"

- By using the DISPLAY command, you can also see that the majority of variables have been coded as numbers.

14. Create a Frequency of the A_LTASTHM.

A_LTASTHM	Frequency	Percent	Cum Percent	
1	4875	87.9%	87.9%	
2	660	11.9%	99.8%	
9	9	0.2%	100.0%	
Total	5544	100.0%	100.0%	

- Notice the frequency produces results for NO answers (the number 1) and YES answers (the number 2). There is also a number 9 that indicates missing values. To enhance your statistics, remove the missing values from the calculations.

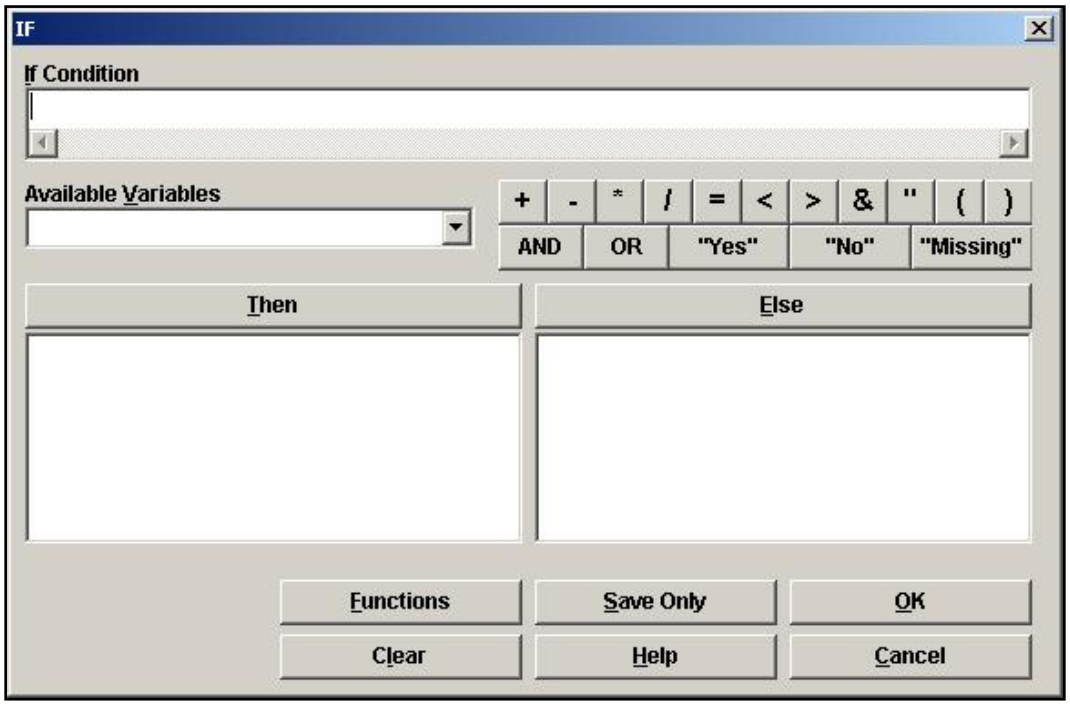
Creating an If/Then Statement

In Lesson 2 Check Code, you created an IF/THEN statement to create a skip pattern. Now you will use Analysis to create an IF/THEN statement that removes missing values from your statistics. The IF dialog box in MakeView and Analysis are the same; however, the navigation from the THEN is slightly different. Rather than selecting from the tabs (i.e., MakeView Program Editor), you will select from the Command Tree.

You are making the statement that IF the value of A_LTASTHM is 9 THEN it will be considered missing. Missing values will not be included in the statistics because you did not include them when you opened the SET dialog box in the previous section.

Now that you created the statement, use it wisely.

1. From the Command Tree Select/If folder, click **IF**. The IF dialog box opens.

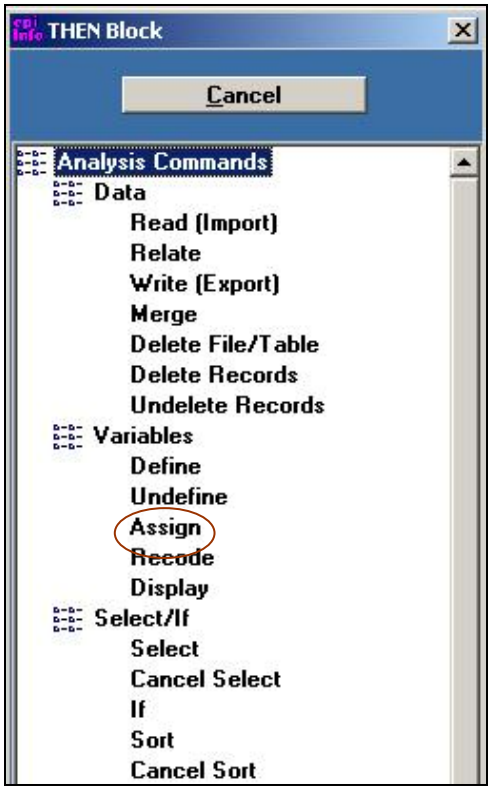


2. From the Available Variables drop-down, select **A_LTASTHM**. The If Condition field populates.

3. Click **=**.

4. Type the number **9**.

5. Click **Then**. The Command Tree window activates.

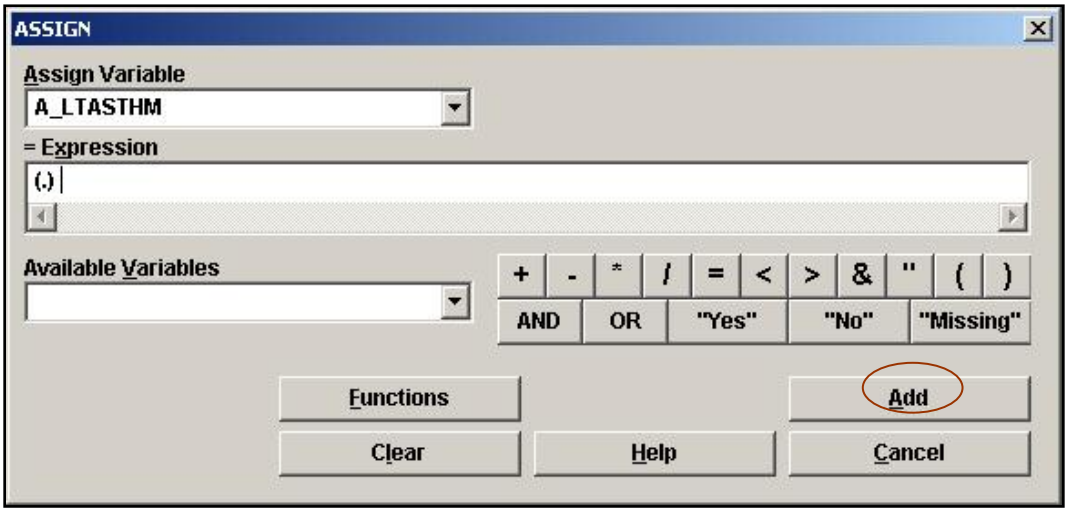


6. From the Command Tree Variables folder, click **Assign**. The ASSIGN dialog box opens.

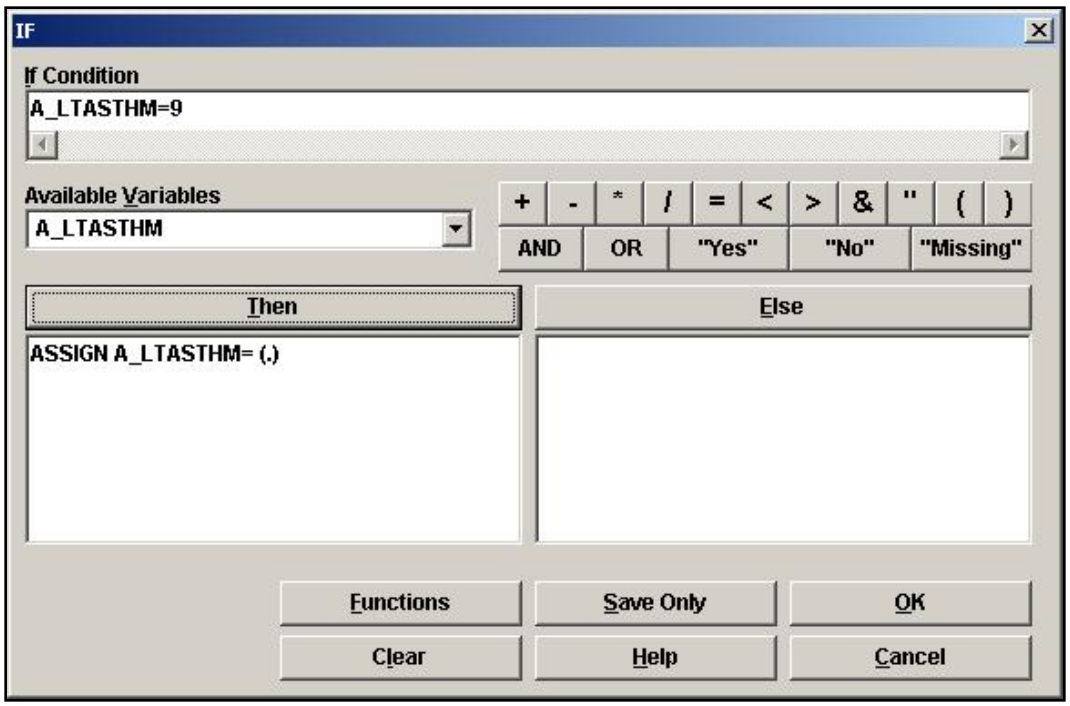
7. From the Assign Variable drop-down, select **A_LTASTHM**.

8. Place your cursor in the =Expression field.

9. Click **Missing**.



10. Click **Add**. The IF dialog box opens.



11. Click **OK**. Your code appears in the Program Editor:

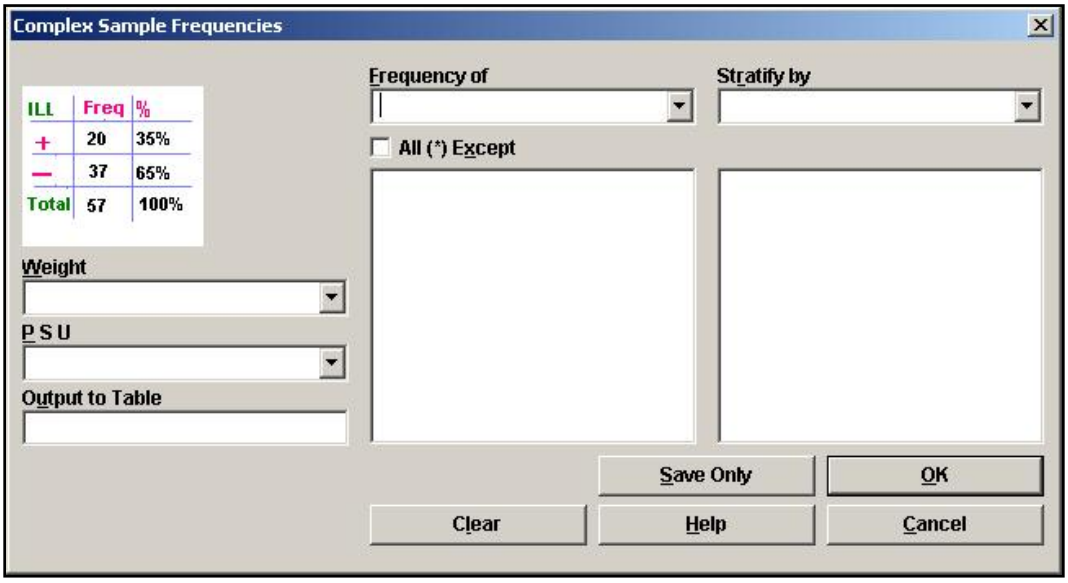
```
IF A_LTASTHM=9 THEN
    ASSIGN A_LTASTHM= (.)
END
```

Complex Sample Frequency

Now that the missing values have been excluded, you can compute the prevalence.

It's not really that complex.

1. From the Command Tree Advanced Statistics folder, click **Complex Sample Frequencies**. The Complex Sample Frequencies dialog box opens.



2. From the Frequency of drop-down menu, select **A_LTASTHM**.

3. From the Weight drop-down, select **A_FINALWT**. The Weight variable has been computed for you, and is part of the BRFSS data. To find out how this variable is calculated, refer to the BRFSS website.

   - If variables are not weighted, each record counts as the same, one record represents one person. Weighting attempts to compensate for circumstances where variables must represent more than one individual because of the sampling process, one record may count as 100 people, depending on the sample size. Weighting may also be needed when you do not have the original data, but only summarized data showing frequency of particular responses.

4. From the Primary Sampling Unit (PSU) drop-down menu, select **A_PSU**. The PSU variable has been computed for you and is part of the BRFSS data. To find out how this variable is calculated, refer to the BRFSS website.

   - In cluster sampling, the PSU represents the clusters from which the samples were drawn randomly. The clusters themselves are selected from a set of similar clusters, which together include all of the population to be sampled. The random sample of the population from the random sample of clusters is taken as a proxy for a random sample from the population as a whole. Complex sample analysis attempts to adjust

for the variance reduction because each member of the population did not have a uniform probability of selection. Clusters are rarely chosen on the basis of variables (e.g., age, gender, or race).

5. Click **OK**. Results appear in the Output window.

The code in the Program Editor should look like this:

```
FREQ  A_LTASTHM WEIGHTVAR=A_FINALWT PSUVAR=A_PSU
```

Your Output should look like the following.

A_LTASTHM	TOTAL
1	4875
Row %	100.000
Col %	88.306
SE %	0.507
LCL %	87.313
UCL %	89.300
2	660
Row %	100.000
Col %	11.694
SE %	0.507
LCL %	10.700
UCL %	12.687
TOTAL	5535
Design Effect	1.378

Prevalence

Confidence Intervals

- The confidence intervals provide a range of values where the prevalence can fall based on the existing variables and conditions. For this example, the lifetime prevalence is 11.7; however, the confidence intervals claim it could be as low as 10.7 or as high as 12.7.

- If you want to determine if the prevalence is increasing or decreasing, download several years of data from the BRFSS website and complete similar steps to determine the variable names and values. Using the WRITE command, you could create a table to hold data for each year in one location.

Required Tasks

B. What is the current asthma prevalence among adults in the state?

Compute the current adult asthma prevalence, using the same series of steps for the A_CASTHMA variable.

The BRFSS website indicates that the calculated variable containing information on the current prevalence of asthma is called _CASTHMA. Since you imported the file as a .DBF, the variable name contains a prefix of A to compensate for the space field. The variable you need to figure for prevalence is called A_CASTHMA. This is a calculated variable based on the BRFSS question, "Do you still have asthma?"

First, exclude the missing variables using the If/Then statement.

```
IF A_CASTHMA=9 THEN

ASSIGN A_CASTHMA= (.)

END
```

Click **Complex Sample Frequencies**.

Create a Frequency of A_CASTHMA using the Weight of A_FINALWT and the PSU of A_PSU.

Results appear in the Output Window:

A_CASTHMA	TOTAL
1	5079
Row %	100.000
Col %	92.362
SE %	0.413
LCL %	91.553
UCL %	93.170
2	439
Row %	100.000
Col %	7.638
SE %	0.413
LCL %	6.830
UCL %	8.447
TOTAL	5518
Design Effect	1.331

Prevalence

Confidence Intervals

Selecting Subgroups

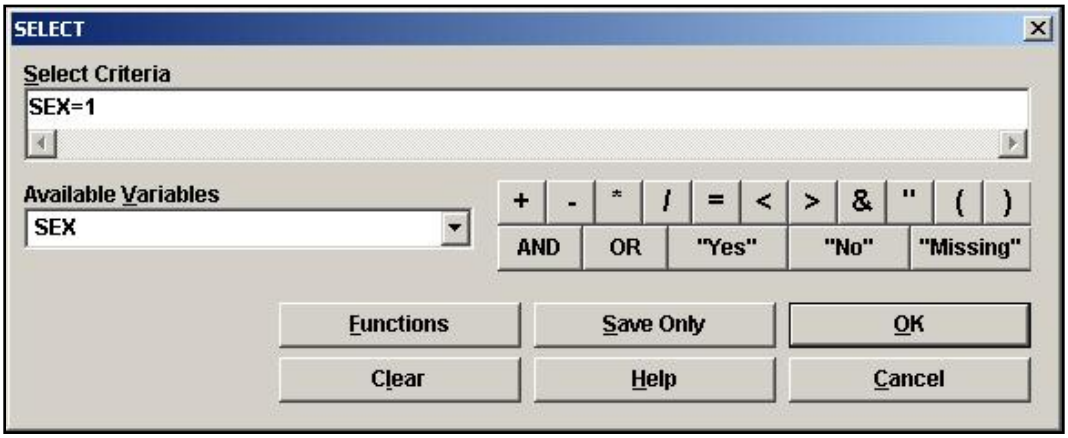
To strengthen your proposal, compare asthma prevalence rates among subgroups in the state and see if there are differences.

C. Is the state adult asthma prevalence rate higher in the subgroup gender?

Subgroups available in the BRFSS data include age, sex, ethnicity, and race. Gender in the BRFSS dataset is coded as the number 1 for males, and the number 2 for females. To compute the prevalence for each group, use the SELECT command.

Get those subgroups.

1. From the Command Tree Select/If folder, click **Select**. The SELECT dialog box opens.
2. From the Available Variables drop-down, select **Sex**.
3. Click **=**.
4. Type **1**..



5. Click **OK**. The survey contains 2,172 males.
6. Click **Complex Sample Frequencies**. The Complex Sample Frequencies dialog box opens.
7. Select **A_LTASTHM** from the Frequency of drop-down box.
8. Select a Weight of **A_FinalWT**.
9. Select a **PSU** of A_PSU.
10. Click **OK**.

11. Click **Cancel Select**.

Required Tasks

Compute the prevalence for the subgroup women.

Select Sex=2. You should have 3,372 records.

Click **Complex Sample Frequencies**. Use the BRFSS Weight and PSU variables to compute the prevalence of the variable A_LTASTHM. Record the results on your answer sheet.

Click **Cancel Select**.

Complex Sample Tables

To strengthen your proposal, you will compare asthma prevalence rates among subgroups in the state and see if any differences exist.

D. Is the state adult asthma prevalence rate higher in the subgroup ethnicity?

Use the BRFSS data to compute the lifetime prevalence rate for the subgroup ethnicity. Use the Complex Sample Tables command to compute for prevalence and the odds ratio.

In the BRFSS data table, ethnicity is categorized in the variable HISPANIC2 with yes answers coded as 1, and no as 2. The lifetime prevalence of asthma is categorized in the variable A_LTASTHM with no answers coded as 1 and yes answers as 2. To compute with accuracy using Complex Sample Tables, the first value in order must represent YES, and the second must represent NO. You need to perform some data management to arrange your tables in the proper sequence.

Your current data table should look like this:

A_LTASTHM	1= NO 2= YES
HISPANC2	1=YES 2=NO

You want it to look like this:

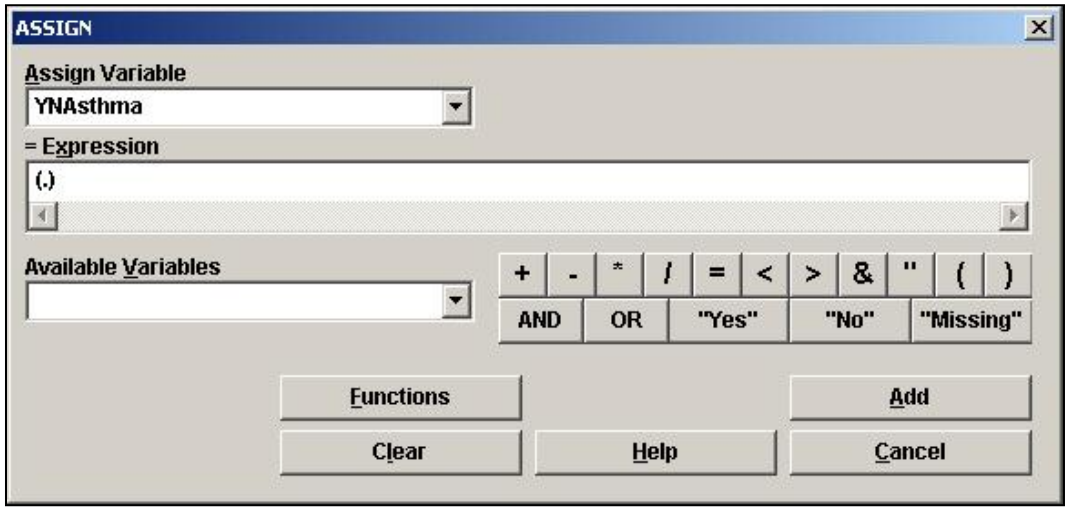
A_LTASTHM	1= YES 2= NO
HISPANC2	1=YES 2=NO

In order to get the YES/NO values to match in the proper order, you will create an IF /THEN to statement to ASSIGN newly defined variables. The IF statement will remove the missing variables and reverse the order of the coded 1s and 2s using a mathematical statement.

Define **a new lifetime asthma prevalence variable and assign it using an IF/THEN statement.**

Like the one above, this one is not too complex.

1. Read in the data table BRFSS2003.
2. From the Command Tree Variables folder, click **Define**. The DEFINE dialog box opens.
3. In the Variable Field Name field, type **YNAsthma**.
4. Click **OK**.
5. Click **IF**. The IF dialog box opens.
6. From the Available Variables drop-down, select **A_LTASTHM**.
7. Click the greater than **>** operator.
8. Type in the number **2**.
   - The IF Condition dialog box should read A_LTASTHM>2.
9. Click **THEN**. The Command Tree highlights.
10. Click **Assign**. The ASSIGN dialog box opens.
11. From the Assign Variable drop-down, select **YNAsthma**.
12. Move the cursor to the =Expression field and click **Missing**.



13. Click **Add**. The IF dialog box opens.

- You have removed the values of 7 and 9 by stating that **if** the value of A_LTASTHM is greater than 2, then the values are assigned as missing.

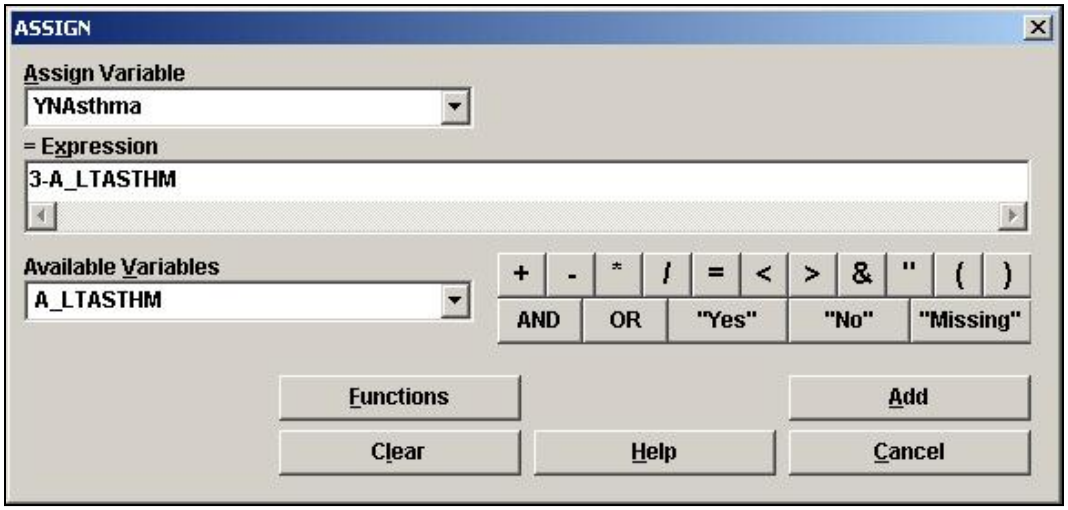
14. Click **ELSE**. The Command Tree highlights.

15. Click **Assign**. The ASSIGN dialog box opens.

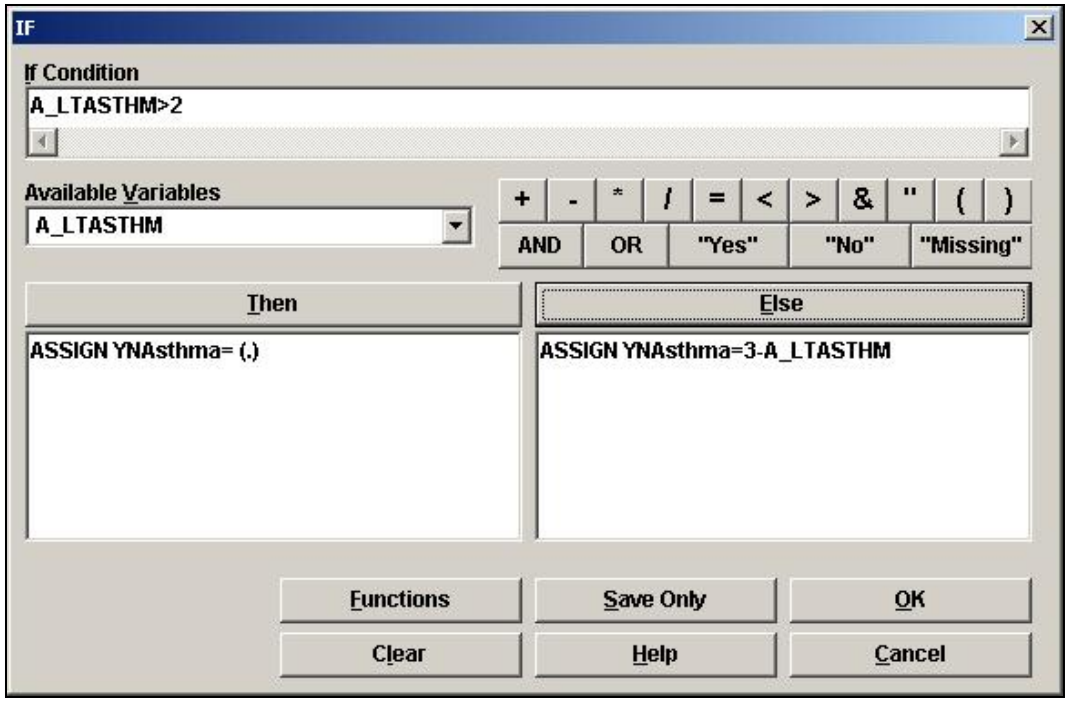
16. From the Available Variable drop-down, select **YNAsthma**.

17. Move your cursor to the =Expression field. Type **3-A_LTASTHM**.

- You are assigning the variable a mathematical statement that says YNAsthma=3-A_LTASTHM (i.e, the variable YNAshtma=3-2 or 3-1. This reformats the Yes/No values into the proper order with 1 now being Yes and 2 being No and also places them in the proper sequence with the Yes value coming first in the series.



18. Click **Add**. The IF dialog box opens.



19. Click **OK**. Now define and assign values for a new HISPANC2 variable.

Define a new ethnicity variable and assign it using an IF/THEN statement**.**

Time to define and assign.

1. Click **Define** and create a new variable called Hisp.

2. Click **OK**.

3. Click **IF**. The IF dialog box opens.

4. From the Available Variables drop-down, select **HISPANC2**.

5. In the IF Condition field, type **>2**.

6. Click **THEN**. The Command Tree highlights.

7. Click **Assign**. The ASSIGN dialog box opens.

8. From the Assign Variable drop-down, select **Hisp**.

9. Move your cursor to the =Expression field. Click **Missing**.

10. Click **Add**. The IF dialog box opens.

11. Click **ELSE**. The Command Tree highlights.

12. Click **Assign**. The ASSIGN dialog box opens.

13. From the Assign Variable drop-down, select **Hisp**.

14. Move the cursor to the =Expression field. Type **HISPANC2**.

- Since the coded values for the variable Yes and No are in the correct order inside the variable HISPANC2, you do not have to create the mathematical functions that changes their values.

15. Click **Add**. The IF dialog box opens.

16. Click **OK**.

View the results of your IF/THEN statements and create a Complex Sample Table.

Oh, the complexity!

1. From the Command Tree Statistics folder, click **Frequencies**. The FREQ dialog box opens.

2. From the Frequency Of drop-down, select **Hisp, HISPANC2, YNAsthma**, and **A_LTASTHM**.

3. Click **OK**. View the results and how the IF/THEN statements you created altered the results and the table structure.

4. From the Command Tree Advanced Statistics folder, click **Complex Sample Tables**. The TABLES dialog box opens.

5. From the Exposure Variable drop-down, select **Hisp**.

6. From the Outcome Variable drop-down, select **YNAsthma**.

7. From the Weight drop-down, select **A_FINALWT**.

8. From the PSU drop-down, select **A_PSU**.

9. Click **OK**. View the results.

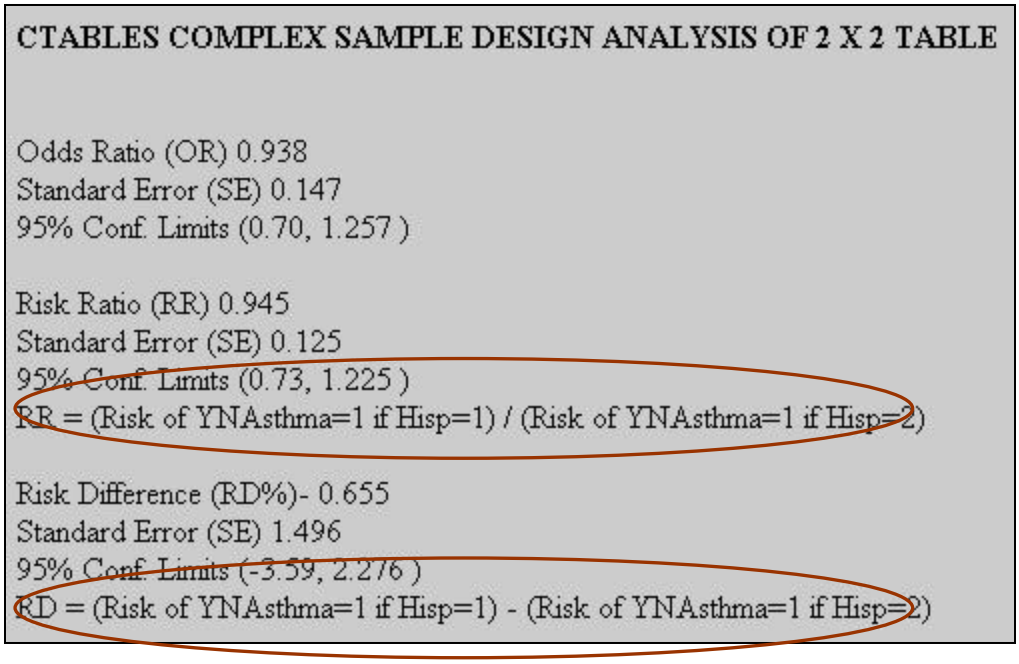
Hisp	YNAsthma		
	1	2	TOTAL
1	82	494	576
Row %	11.197	88	
Col %	15.674	16	
SE %	1.393	1.393	
LCL %	8.468	86.073	
UCL %	13.927	91.532	
Design Effect	1.124	1.124	
2	576		
Row %	11.852	88	
Col %	84.326	83.459	83.561
SE %	0.545	0.545	
LCL %	10.784	87.079	
UCL %	12.921	89.216	
Design Effect	1.399	1.399	
TOTAL	658	4838	5496
Row %	11.745	88	
Col %	100.000	100	
SE %	0.510	0.510	
LCL %	10.745	87.256	
UCL %	12.744	89.255	
Design Effect	1.379	1.379	

The lifetime prevalence of asthma among Hispanics.

The lifetime prevalence of asthma among non-Hispanics.

The lifetime prevalence of asthma for the state population.

Scroll down to view the Odds and Risk Ratios and the Standard Error and Confidence Limits.



CTABLES COMPLEX SAMPLE DESIGN ANALYSIS OF 2 X 2 TABLE

Odds Ratio (OR) 0.938
Standard Error (SE) 0.147
95% Conf. Limits (0.70, 1.257)

Risk Ratio (RR) 0.945
Standard Error (SE) 0.125
95% Conf. Limits (0.73, 1.225)
RR = (Risk of YNAsthma=1 if Hisp=1) / (Risk of YNAsthma=1 if Hisp=2)

Risk Difference (RD%)- 0.655
Standard Error (SE) 1.496
95% Conf. Limits (-3.59, 2.276)
RD = (Risk of YNAsthma=1 if Hisp=1) - (Risk of YNAsthma=1 if Hisp=2)

Notice that your tables are set up in the proper order and sequenced with Yes to Asthma and Yes to Hispanic over Yes to Asthma and No to Hispanic.

You can use county data from the Expanded BRFSS to answer other questions in the same manner. To answer questions and comparisons related to county or EBRFSS data, follow the same process as above. You can use this process to answer these types of questions:

- Is the county adult asthma prevalence higher or lower than the state?
- Is the county adult asthma prevalence higher in some population subgroups?
- How do county population subgroup patterns relate to the state sub-group patterns for asthma prevalence?

Creating a Frequency

Use the 800 records used as your primary data source to find prevalence rates for the individual schools in the survey. For your proposal, these data represent the county's child population.

E. From the school asthma survey data, what is the difference in prevalence rates?

To find the difference in prevalence rates between the four schools in the survey, create a frequency of asthma and stratify by School Number.

Not the same as being a frequent flyer!

1. Click **Read/Import**. The READ dialog box opens.
2. From the project Asthma Final, select **viewPreInterventionSurvey**.
3. Click **OK**. You should have the 800 records from the survey.
4. Click **Frequencies**. The FREQ dialog box opens.
5. From the Frequency Of drop-down menu, select **Asthma**.
6. From the Stratify By drop-down menu, select **SchoolNum**.
7. Click **OK**. The Output window populates with one table for each school.

Sampling Scheme

The dataset Asthma Final.MDB contains the 800 records you worked with throughout the Basic tutorial. You will use this data to represent all the schools for the county and all the children surveyed for the county to compute the county childhood asthma prevalence.

To do this, think in terms of the sampling scheme of the schools in relation to the county. For this example, follow the sampling method below, use it to make your assumptions about the results you find, and how your data is selected and organized:

- All four elementary schools were chosen.
- A number of classrooms for each school were chosen at random until the number of students in the chosen classes was about 400 at each school.
- Half of the kids surveyed from each classroom were chosen at random.
- The cluster is the classroom within each school; however, classroom information was not recorded, so age was used as a proxy for classroom.
- The Primary Sampling Unit (PSU) is SchoolNum and Age (use cut-off for school year September 1, 2005).

You need to create a weight for the schools to the county population by age and sex to get a county estimate of prevalence and to compute for a subgroup. In the following lesson, you will create a table that contains the sex and age data for the 800 records.

You will use the RELATE command to join the sex and age table created from the 800 records to the population table you are going to create based on data from the U.S. Census. The weight variable will be the county population divided by the survey population first for males, then for females. In these data, each respondent of a particular age and sex represents a number in the general population. Compute a Complex Sample Frequency where the PSU is SchoolNum and Age (this is a grouping, not a numeric field).

From these data, you will receive a county estimate of prevalence by sex for the selected age group (based on the survey age groups).

In the following sections, you will use your data management skills to:

- Import and format HTML data.
- Create a numerator table using the HTML data.
- Relate tables.
- Create a weight variable.
- Create a PSU variable.

- Compute the prevalence.

The ultimate goal is to obtain the asthma prevalence for children in the county. You will take the number of cases of asthma in the survey and divide it by the population of the age group for the county (from the HTML census data).

Importing Census Data/HTML Format

You want your proposal to include information on childhood asthma. Use the 800 records from your primary data source and data from the Census Bureau to compute the country's childhood asthma prevalence rate.

F. What is the county childhood asthma prevalence?

G. What is the county childhood asthma prevalence for children in the subgroup Gender?

To compute the prevalence rate, you need population data for the age group in the survey. Population data for this example was downloaded from the U.S. Census website, www.census.gov. Data were downloaded for Albany County by single years for age and sex.

You will be importing an .HTML file that was downloaded from the census website. Format the data to use as your county child population. You will determine the prevalence by dividing the number of cases of asthma in the survey by the population of the age group for the county.

Back to the importing.

1. Click **Read/Import**. The READ dialog box opens.
   - Make sure your current project is Asthma Final. If not, click **Change Project** and select **Asthma Final**.
2. From the Data Format drop-down, select **HTML**.
3. Click the **Browse** button in the Data Source field. The READ window opens.
4. Locate and select the HTML file called **censusage.htm**.
5. Click **Open**.
6. Select **TABLE 284:1**. The HTML window appears so you can view information.
7. Close the HTML window.
8. Click **OK**. The READ Temp Link dialog box opens.
9. Type the link name **CensusLink.**
10. Click **OK**. You should have 42 records.
11. From the Command Tree Variables folder, click **Display**. Notice that numeric fields were imported as text and that the variable names reverted to the column number.

12. Click **List** to view the data. Notice the commas in the records.

- Reformat the fields containing commas. Use the text-to-number format command, and define/assign new variable names to use these data.

13. From the Command Tree Data folder, click **Write/Export**. The WRITE dialog box opens.

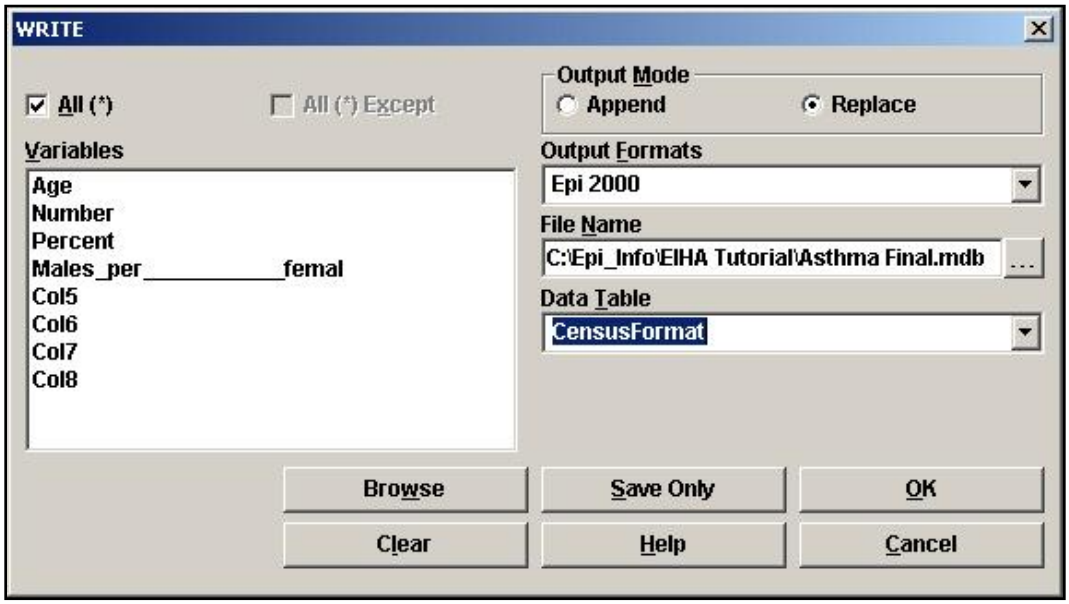
14. From the Output Mode section, click the **Replace** radio button.

15. From the File Name field, click **Browse**.

16. Select the **Asthma Final** project.

17. Click **Save**.

18. In the Data Table field, type **CensusFormat**.



19. Click **OK**.

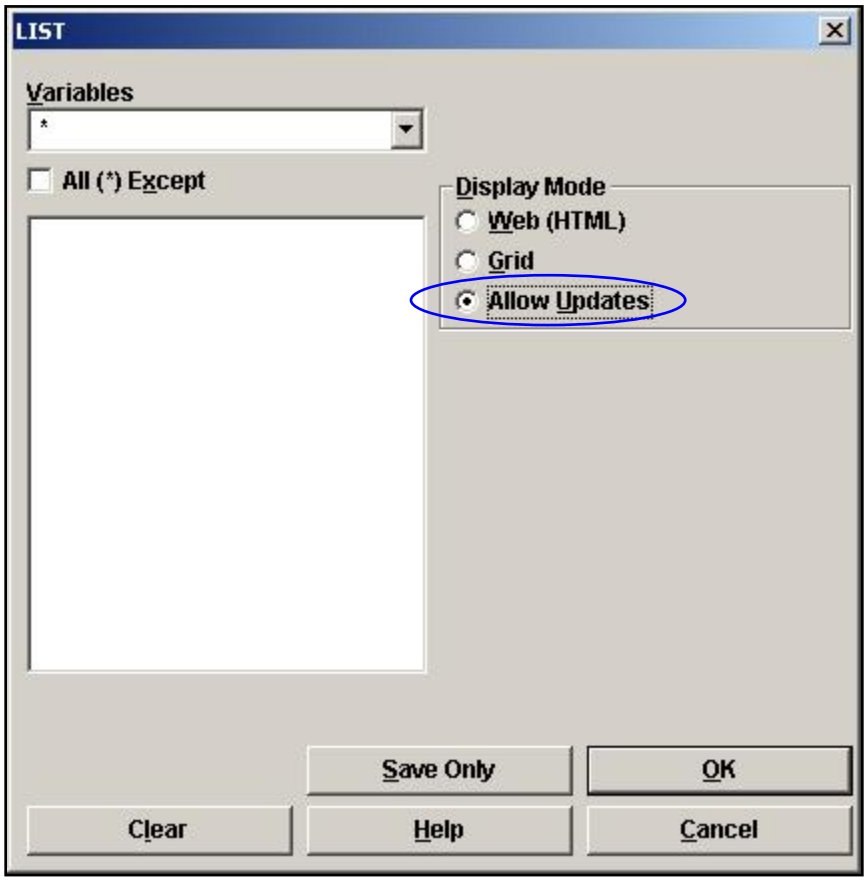
- You now have a new data table called CensusFormat that contains all the information imported from the Census HTML file.

Formatting with List Update

The following formatting instructions can be used when working with any files imported into Analysis, including HTML or MS Excel files.

Time for an update.

1. Click **Read/Import**. The READ dialog box opens.
2. Click the **Show All** radio button.
3. Select the **CensusFormat** table you created in the previous lesson.
4. Click **OK**. You should have the same 42 records.
5. Click **List**. The LIST dialog box opens.
6. From the Display Mode section, select the **Allow Updates** radio button.



7. Click **OK**. A grid table appears with all your variables listed.
   - From this grid table, you can make changes to the data in your records. Keep in mind that changes made in this mode are saved and cannot be undone.

- For this example, you are interested in the 5–12 year age groups that correspond to the survey ages. You will alter only those records, select them, and place them in a new table.

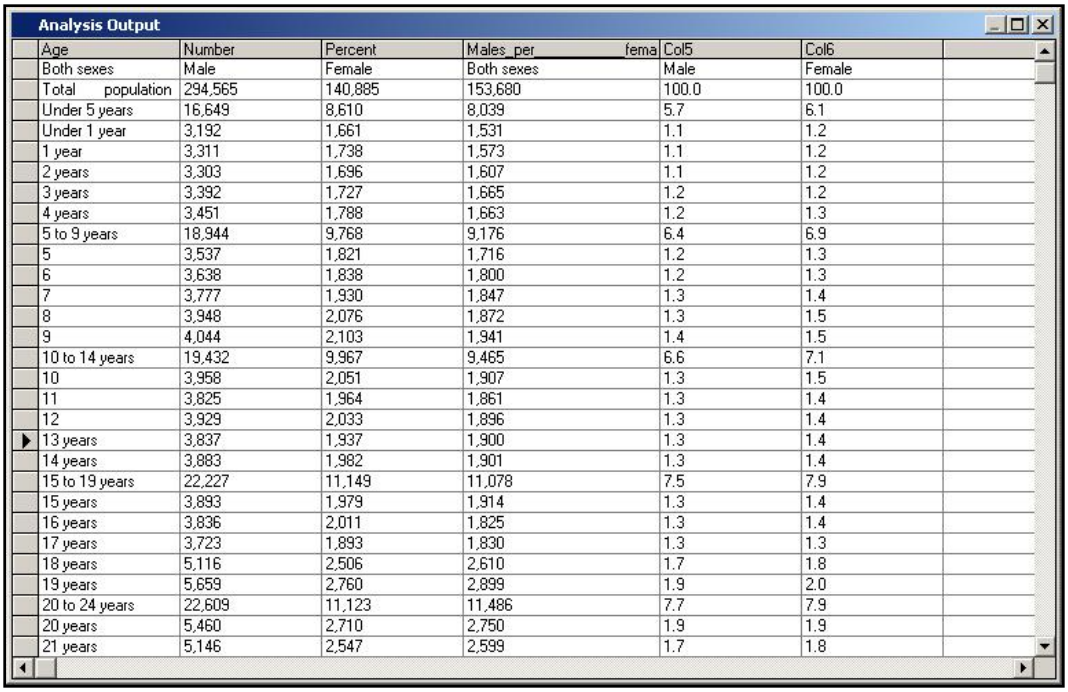
Analysis Output

Age	Number	Percent	Males_per_____fema	Col5	Col6
Both sexes	Male	Female	Both sexes	Male	Female
Total population	294,565	140,885	153,680	100.0	100.0
Under 5 years	16,649	8,610	8,039	5.7	6.1
Under 1 year	3,192	1,661	1,531	1.1	1.2
1 year	3,311	1,738	1,573	1.1	1.2
2 years	3,303	1,696	1,607	1.1	1.2
3 years	3,392	1,727	1,665	1.2	1.2
4 years	3,451	1,788	1,663	1.2	1.3
5 to 9 years	18,944	9,768	9,176	6.4	6.9
5 years	3,537	1,821	1,716	1.2	1.3
6 years	3,638	1,838	1,800	1.2	1.3
7 years	3,777	1,930	1,847	1.3	1.4
8 years	3,948	2,076	1,872	1.3	1.5
9 years	4,044	2,103	1,941	1.4	1.5
10 to 14 years	19,432	9,967	9,465	6.6	7.1
10 years	3,958	2,051	1,907	1.3	1.5
11 years	3,825	1,964	1,861	1.3	1.4
12 years	3,929	2,033	1,896	1.3	1.4
13 years	3,837	1,937	1,900	1.3	1.4
14 years	3,883	1,982	1,901	1.3	1.4
15 to 19 years	22,227	11,149	11,078	7.5	7.9
15 years	3,893	1,979	1,914	1.3	1.4
16 years	3,836	2,011	1,825	1.3	1.4
17 years	3,723	1,893	1,830	1.3	1.3
18 years	5,116	2,506	2,610	1.7	1.8
19 years	5,659	2,760	2,899	1.9	2.0
20 to 24 years	22,609	11,123	11,486	7.7	7.9
20 years	5,460	2,710	2,750	1.9	1.9
21 years	5,146	2,547	2,599	1.7	1.8

8. Place your cursor in the Age column on the record for 5 years.

9. Highlight the text and type the number **5**.

10. Repeat this process for the years **6, 7, 8, 9, 10, 11**, and **12**. Do NOT alter the total fields.



Analysis Output

Age	Number	Percent	Males_per_____fema	Col5	Col6
Both sexes	Male	Female	Both sexes	Male	Female
Total population	294,565	140,885	153,680	100.0	100.0
Under 5 years	16,649	8,610	8,039	5.7	6.1
Under 1 year	3,192	1,661	1,531	1.1	1.2
1 year	3,311	1,738	1,573	1.1	1.2
2 years	3,303	1,696	1,607	1.1	1.2
3 years	3,392	1,727	1,665	1.2	1.2
4 years	3,451	1,788	1,663	1.2	1.3
5 to 9 years	18,944	9,768	9,176	6.4	6.9
5	3,537	1,821	1,716	1.2	1.3
6	3,638	1,838	1,800	1.2	1.3
7	3,777	1,930	1,847	1.3	1.4
8	3,948	2,076	1,872	1.3	1.5
9	4,044	2,103	1,941	1.4	1.5
10 to 14 years	19,432	9,967	9,465	6.6	7.1
10	3,958	2,051	1,907	1.3	1.5
11	3,825	1,964	1,861	1.3	1.4
12	3,929	2,033	1,896	1.3	1.4
13 years	3,837	1,937	1,900	1.3	1.4
14 years	3,883	1,982	1,901	1.3	1.4
15 to 19 years	22,227	11,149	11,078	7.5	7.9
15 years	3,893	1,979	1,914	1.3	1.4
16 years	3,836	2,011	1,825	1.3	1.4
17 years	3,723	1,893	1,830	1.3	1.3
18 years	5,116	2,506	2,610	1.7	1.8
19 years	5,659	2,760	2,899	1.9	2.0
20 to 24 years	22,609	11,123	11,486	7.7	7.9
20 years	5,460	2,710	2,750	1.9	1.9
21 years	5,146	2,547	2,599	1.7	1.8

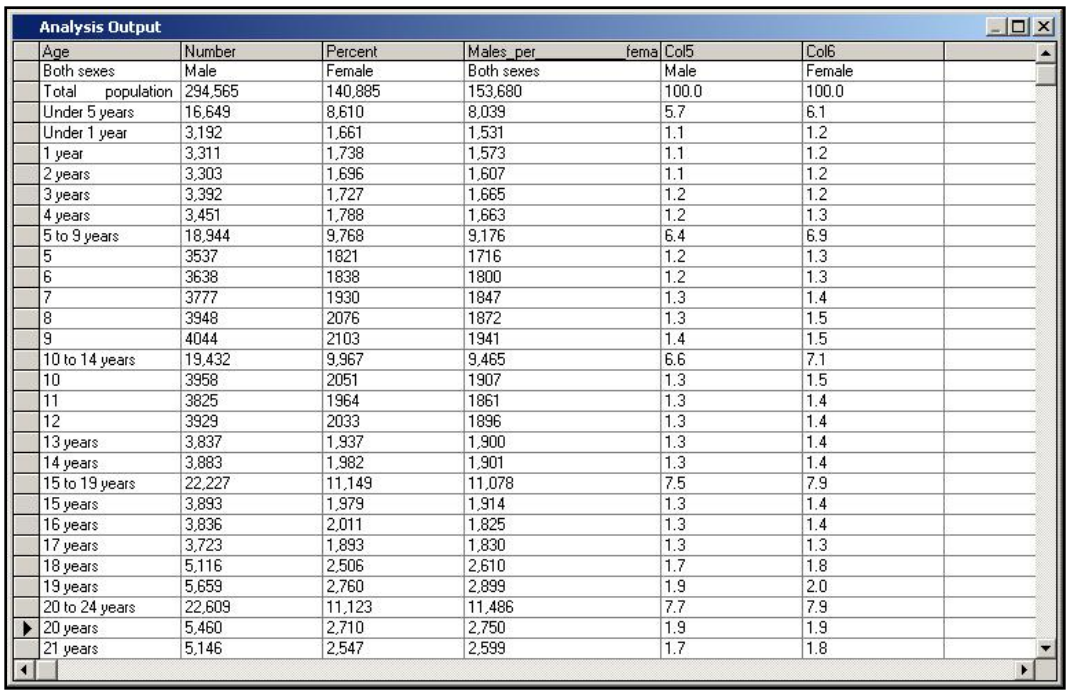
11. Place your cursor in the Number column next to the Age field 5.

12. Retype the number in the field without the commas. For example, replace the number 3,537 with **3537**.

13. Do this for each of the age fields 5–12 that you updated in the previous steps.

14. Repeat this step for the age fields in the Percent column.

15. Repeat this step for the age fields in the Males per female column.



Analysis Output

Age	Number	Percent	Males_per_____fema	Col5	Col6
Both sexes	Male	Female	Both sexes	Male	Female
Total population	294,565	140,885	153,680	100.0	100.0
Under 5 years	16,649	8,610	8,039	5.7	6.1
Under 1 year	3,192	1,661	1,531	1.1	1.2
1 year	3,311	1,738	1,573	1.1	1.2
2 years	3,303	1,696	1,607	1.1	1.2
3 years	3,392	1,727	1,665	1.2	1.2
4 years	3,451	1,788	1,663	1.2	1.3
5 to 9 years	18,944	9,768	9,176	6.4	6.9
5	3537	1821	1716	1.2	1.3
6	3638	1838	1800	1.2	1.3
7	3777	1930	1847	1.3	1.4
8	3948	2076	1872	1.3	1.5
9	4044	2103	1941	1.4	1.5
10 to 14 years	19,432	9,967	9,465	6.6	7.1
10	3958	2051	1907	1.3	1.5
11	3825	1964	1861	1.3	1.4
12	3929	2033	1896	1.3	1.4
13 years	3,837	1,937	1,900	1.3	1.4
14 years	3,883	1,982	1,901	1.3	1.4
15 to 19 years	22,227	11,149	11,078	7.5	7.9
15 years	3,893	1,979	1,914	1.3	1.4
16 years	3,836	2,011	1,825	1.3	1.4
17 years	3,723	1,893	1,830	1.3	1.3
18 years	5,116	2,506	2,610	1.7	1.8
19 years	5,659	2,760	2,899	1.9	2.0
20 to 24 years	22,609	11,123	11,486	7.7	7.9
20 years	5,460	2,710	2,750	1.9	1.9
21 years	5,146	2,547	2,599	1.7	1.8

16. Click the Close **X** to exit the List Update window and return to the Output window.

- Do NOT close Analysis.

Selecting with String Length

Now you have a set of age variables that differ from the series. You can use this difference to select only those age groups you have edited.

Use a format option called String Length to specify the length of the variable you select.

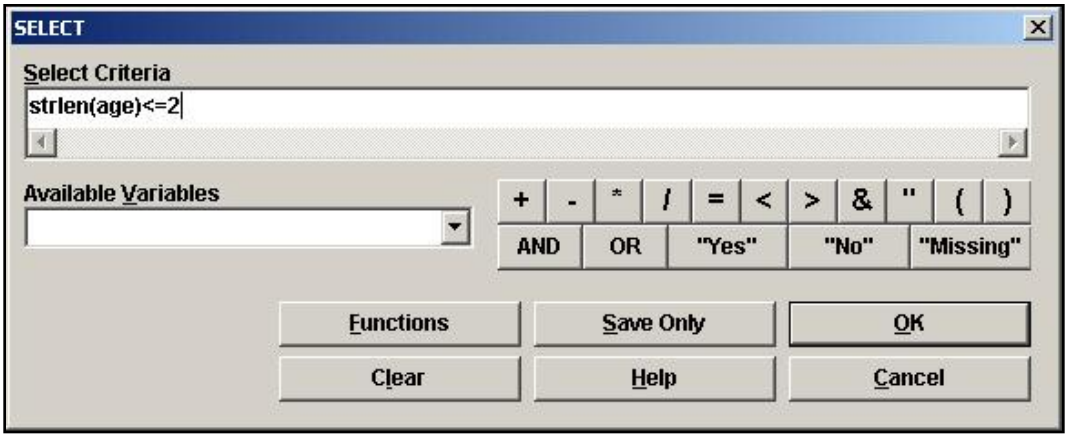
Time to play the first string.

1. Click **Select**. The SELECT dialog box opens.

2. In the Select Criteria field, use your keyboard and the Operator buttons to create the following code:

   strlen(age)<=2

   - This states that you want to see only the records from the Age field that have a length less than or equal to two characters.



3. Click **OK**. You should have eight records.

4. Click **List**. The selected list appears.

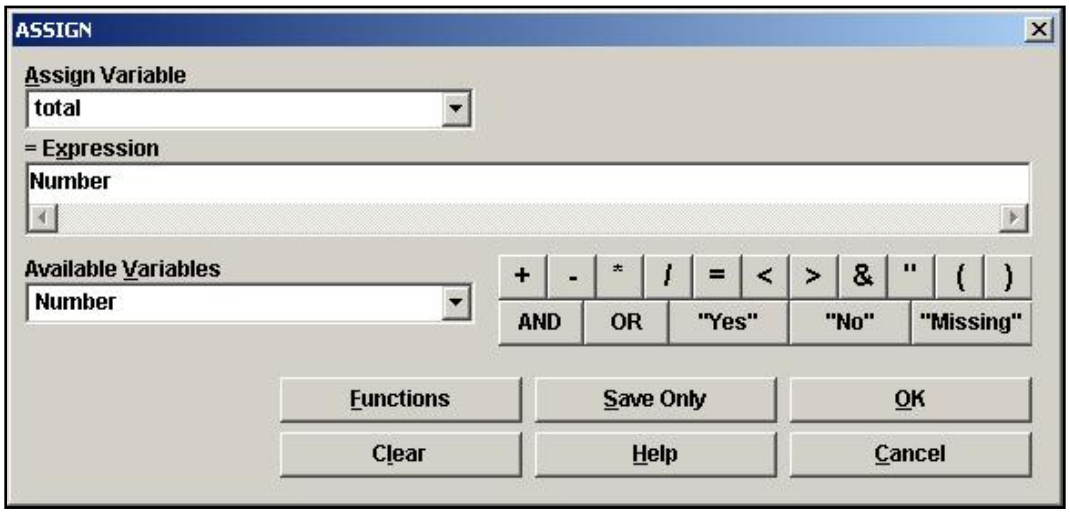
Age	Number	Percent	Males_per_	Col5	Col6	Col7	Col8
5	3537	1821	1716	1.2	1.3	1.1	106.1
6	3638	1838	1800	1.2	1.3	1.2	102.1
7	3777	1930	1847	1.3	1.4	1.2	104.5
8	3948	2076	1872	1.3	1.5	1.2	110.9
9	4044	2103	1941	1.4	1.5	1.3	108.3
10	3958	2051	1907	1.3	1.5	1.2	107.6
11	3825	1964	1861	1.3	1.4	1.2	105.5
12	3929	2033	1896	1.3	1.4	1.2	107.2

Defining and Assigning New Variables

The headings brought in from the .HTML file do not match up with the data in the columns. This can happen with .HTML or MS Excel files. The column titled Number is the total, the column titled Percent is the number of males, and the column titled Males per Female is the number of females. You are going to use the DEFINE command to create new variables, and use the ASSIGN command to provide current table values.

A define moment is upon you.

1. Click **Define** and create a new variable called Total.
2. Click **OK**.
3. Click **Define** and create a new variable called Males.
4. Click **Define** and create a new variable called Females.
5. Click **Assign**. The ASSIGN dialog box opens.
6. From the Assign Variable drop-down, select **Total**.
7. Move your cursor to the =Expression field.
8. From the Available Variables drop-down, select **Number**.



9. Click **OK**.
10. Use the ASSIGN command to assign the variable Males=Percent.
11. Use the ASSIGN command to assign the variable Females=Males_per_female.
12. Click **List** to view the records and verify that the values were copied to the new variables.

13. Click **Write/Export**. The WRITE dialog box opens. You need to only write the new variables to a new table for use in computing the prevalence.

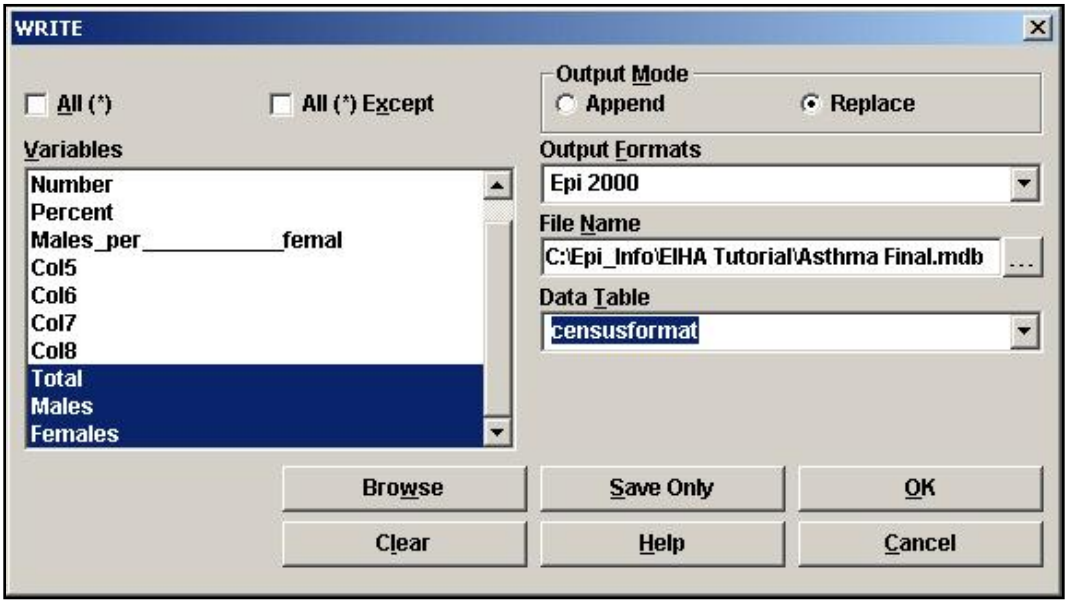
14. From the Variables list, select **Age, Total, Males**, and **Females**.

15. From the Output Mode, select the **Replace** radio button.

16. From the File Name field, click the **Browse** button.

17. Select the **Asthma Final.MDB**.

18. From the Data Table drop-down, select **CensusFormat**.



19. Click **OK**.

20. Click **Read/Import**.

21. Click the **Show All** radio button.

22. Select the **CensusFormat** table.

23. Click **OK**. You should have eight records.

24. Click **List** to view the records.

- You have created a table with the total population per age for the county. It will be used as your denominator table after you compute the rates for the county.

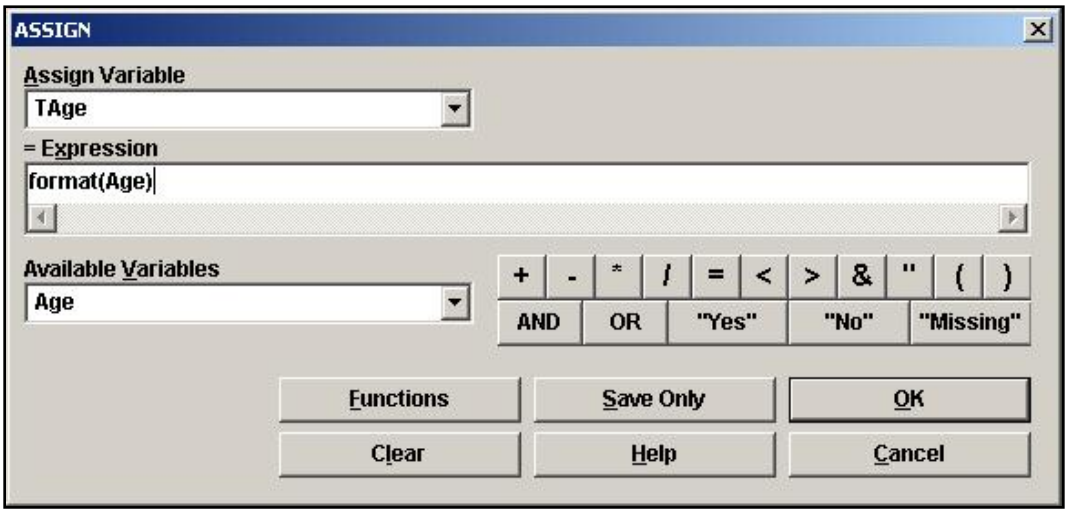
Age	Total	Males	Females
5	3537	1821	1716
6	3638	1838	1800
7	3777	1930	1847
8	3948	2076	1872
9	4044	2103	1941
10	3958	2051	1907
11	3825	1964	1861
12	3929	2033	1896

Creating a Numerator Table

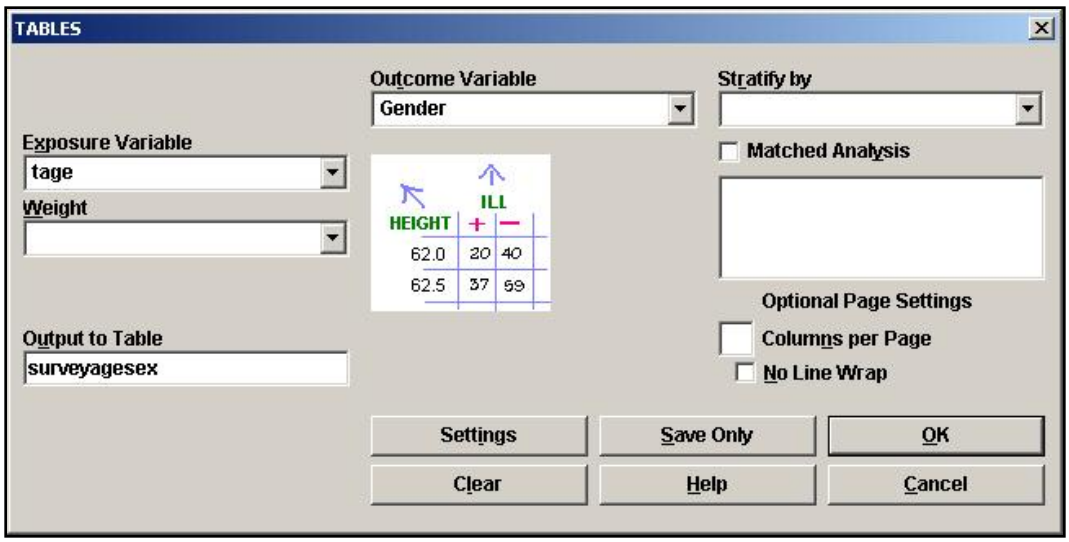
The CensusFormat table is now your denominator table. You need to create the numerator table that contains information about the survey respondents age and sex, then use the RELATE command to join the two tables.

Put some numbers on the table.

1. Click **Read/Import**.
2. Open the view **preInterventionSurvey**. You should have 800 records.
3. Click **Assign**. The ASSIGN dialog box opens.
4. Create the following code **Age=Years(DOB,10/25/2005)**.
5. Click **OK**.
6. List the age variable to make sure it calculated properly.
7. Click **Define** and create a new variable called TAge.
8. Click **Assign**. The ASSIGN dialog box opens.
9. Create the following code **TAge=format(age)**. Age is a number.,TAge is text.



10. From the Command Tree Statistics folder, click **Tables**. The TABLES dialog box opens.
11. From the Exposure Variable drop-down, select **TAge**.
12. From the Outcome Variable drop-down, select **Gender**.
13. In the Output to Table field, type **SurveyAgeSex**.



14. Click **OK**. The Output window shows the results. You have the number of males and females per age.

GENDER

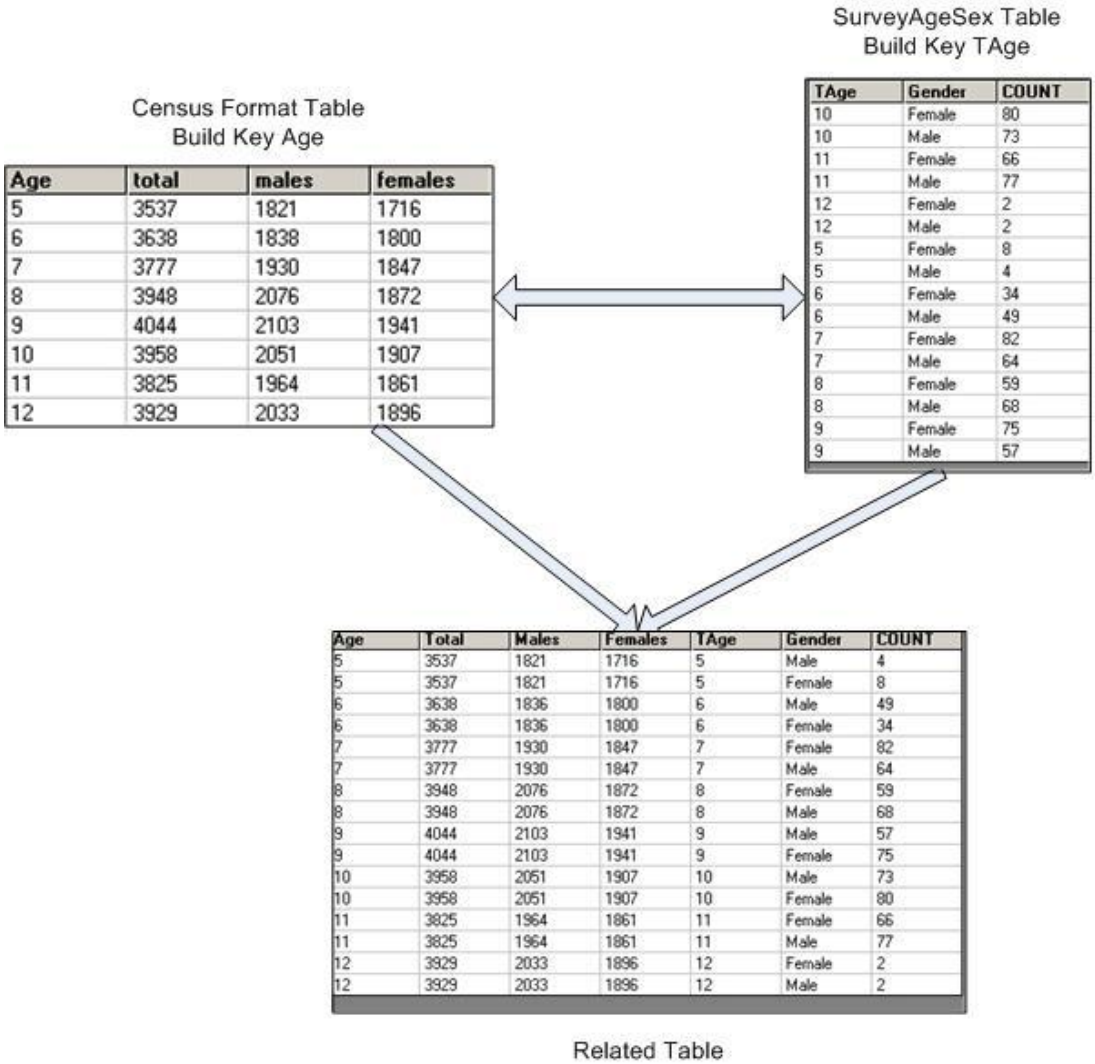
TAge	Female	Male	TOTAL
10	80	73	153
Row %	52.3	47.7	100.0
Col %	19.7	18.5	19.1
11	66	77	143
Row %	46.2	53.8	100.0
Col %	16.3	19.5	17.9
12	2	2	4
Row %	50.0	50.0	100.0
Col %	0.5	0.5	0.5
5	8	4	12
Row %	66.7	33.3	100.0
Col %	2.0	1.0	1.5
6	34	49	83
Row %	41.0	59.0	100.0
Col %	8.4	12.4	10.4
7	82	64	146
Row %	56.2	43.8	100.0
Col %	20.2	16.2	18.3
8	59	68	127
Row %	46.5	53.5	100.0
Col %	14.5	17.3	15.9
9	75	57	132
Row %	56.8	43.2	100.0
Col %	18.5	14.5	16.5
TOTAL	406	394	800
Row %	50.8	49.3	100.0
Col %	100.0	100.0	100.0

Creating a Relate

Now you will use the RELATE command to connect the two tables for the calculations.

To create a weight variable, you need to relate the census population numbers to the survey count numbers. To get the weight, divide the total numbers by the survey count. By relating the CensusFormat and SurveyAgeSex tables, you can define and format new variables and then WRITE a new table containing data from the two original sources.

An interesting relationship follows.



Census Format Table
Build Key Age

Age	total	males	females
5	3537	1821	1716
6	3638	1838	1800
7	3777	1930	1847
8	3948	2076	1872
9	4044	2103	1941
10	3958	2051	1907
11	3825	1964	1861
12	3929	2033	1896

SurveyAgeSex Table
Build Key TAge

TAge	Gender	COUNT
10	Female	80
10	Male	73
11	Female	66
11	Male	77
12	Female	2
12	Male	2
5	Female	8
5	Male	4
6	Female	34
6	Male	49
7	Female	82
7	Male	64
8	Female	59
8	Male	68
9	Female	75
9	Male	57

Related Table

Age	Total	Males	Females	TAge	Gender	COUNT
5	3537	1821	1716	5	Male	4
5	3537	1821	1716	5	Female	8
6	3638	1838	1800	6	Male	49
6	3638	1838	1800	6	Female	34
7	3777	1930	1847	7	Female	82
7	3777	1930	1847	7	Male	64
8	3948	2076	1872	8	Female	59
8	3948	2076	1872	8	Male	68
9	4044	2103	1941	9	Male	57
9	4044	2103	1941	9	Female	75
10	3958	2051	1907	10	Male	73
10	3958	2051	1907	10	Female	80
11	3825	1964	1861	11	Female	66
11	3825	1964	1861	11	Male	77
12	3929	2033	1896	12	Female	2
12	3929	2033	1896	12	Male	2

Data in the new related table can be used to create the new Weight Variables. You can use the WRITE command to create a new permanent table that holds these values.

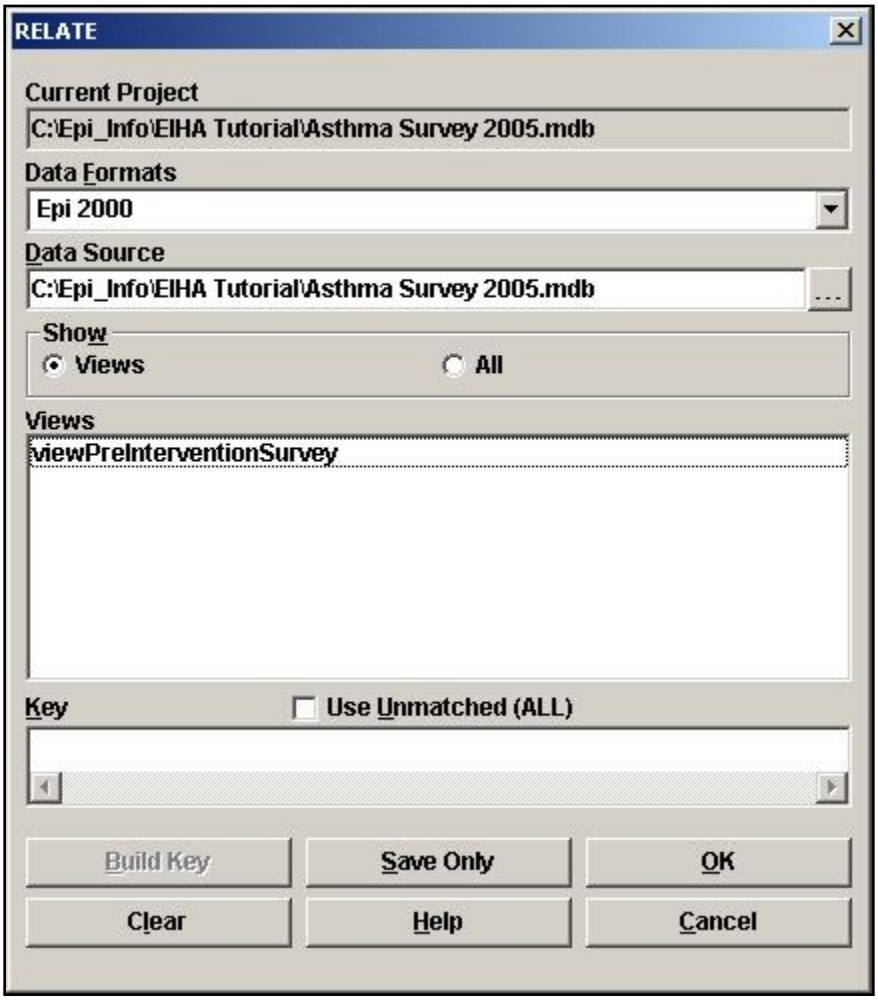
1. Click **Read/Import**. The READ dialog box opens.

2. Click the **Show All** radio button.

3. Select the table called **CensusFormat**.

4. Click **OK**. You should have eight records.

5. From the Command Tree Data folder, click **Relate**. The RELATE dialog box opens.

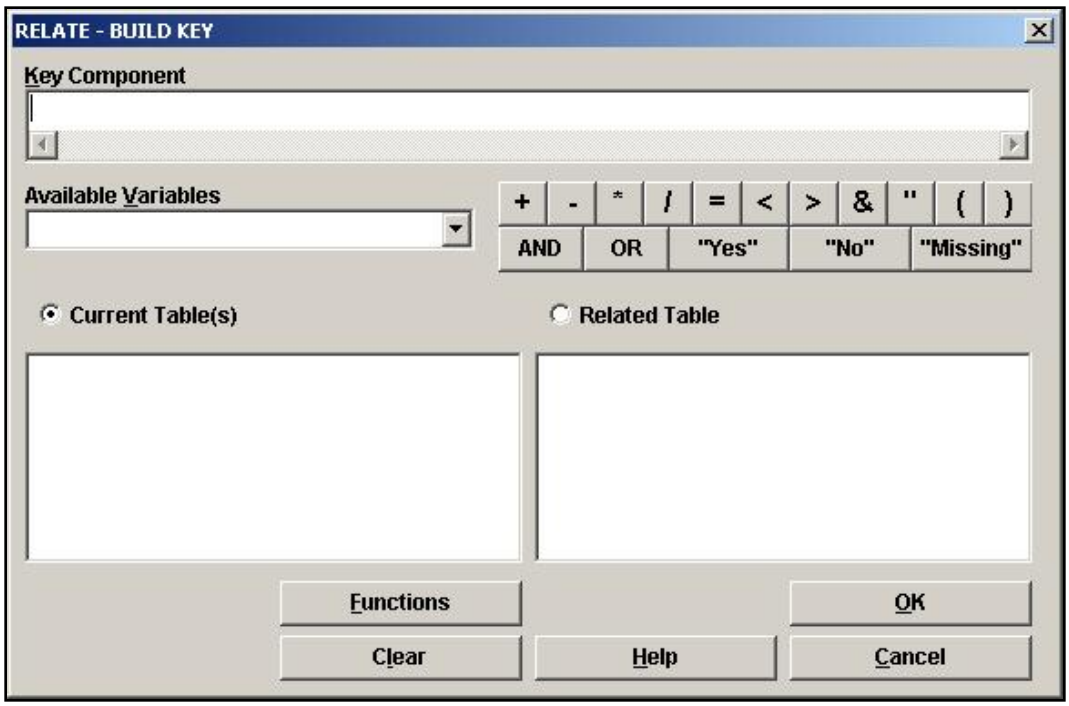


6. Click the **Show All** radio button.

7. Select the **SurveyAgeSex** table.

8. Click **Build Key**. The RELATE-BUILD KEY dialog box opens.

   - Both the tables contain variables for age so you can relate the tables on that variable.



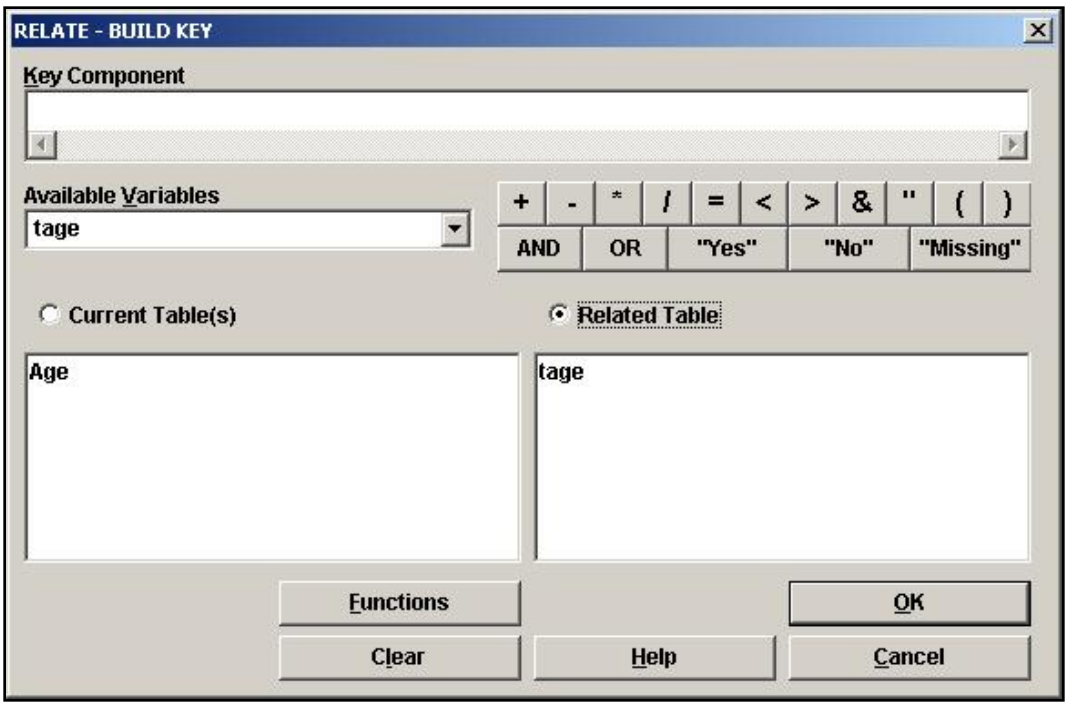
9. Select the **Current Table** radio button.

10. From the Available Variables drop-down, select **Age**.

11. Select the **Related Table** radio button.

12. From the Available Variables drop-down, select **TAge**.

13. Click **OK**. The Related Tables field populates.



12. Click **OK**. The RELATE dialog box opens.

13. Click **OK**. You should have 16 records.

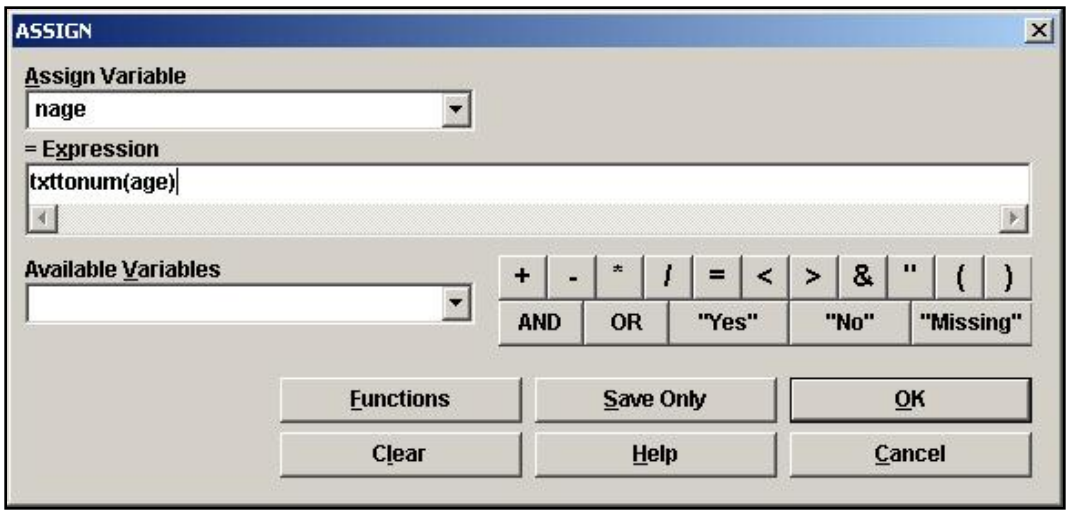
14. Click **Display**. Notice that the fields TAge and Age are text.

15. Click **Define** and create a new variable called NAge. This new variable will hold the age value in a numeric form.

16. Click **Assign**. You are going to format the new age field as a number and assign a value to the variable.

17. From the Assign Variable drop-down, select **NAge**.

18. In the =Expression field type, **TXTTONUM(Age)**.


ASSIGN
Assign Variable
nage
= Expression
txttonum(age)
Available Variables
+ - * / = < > & " ( )
AND OR "Yes" "No" "Missing"
Functions
Save Only
OK
Clear
Help
Cancel

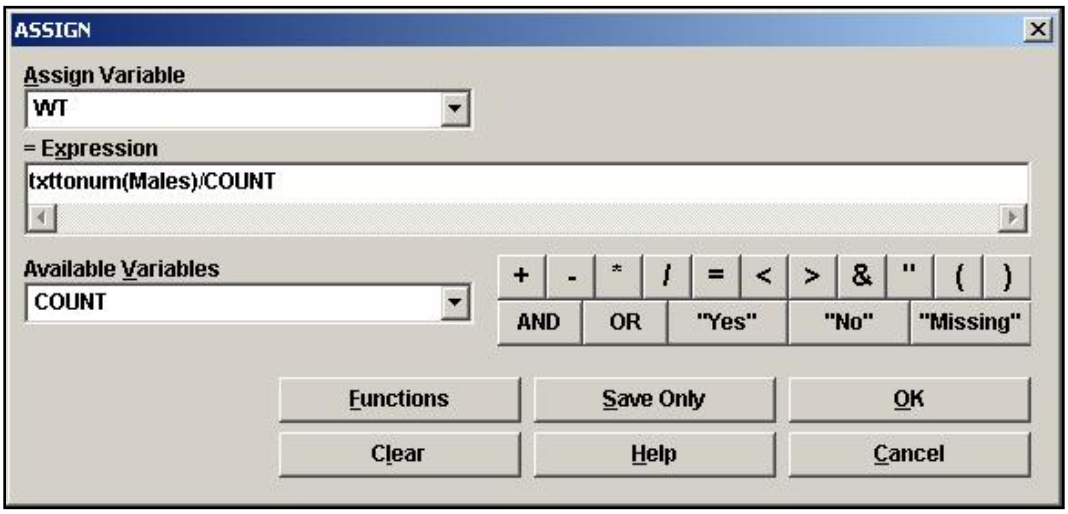
19. Click **OK**.

Creating a Weight Variable

Now you will create a weight variable to use in your calculations. First, you will select the males in the dataset; second, create the weight calculation; and third, write the variables out to a new table. Repeat the process for the females in the dataset.

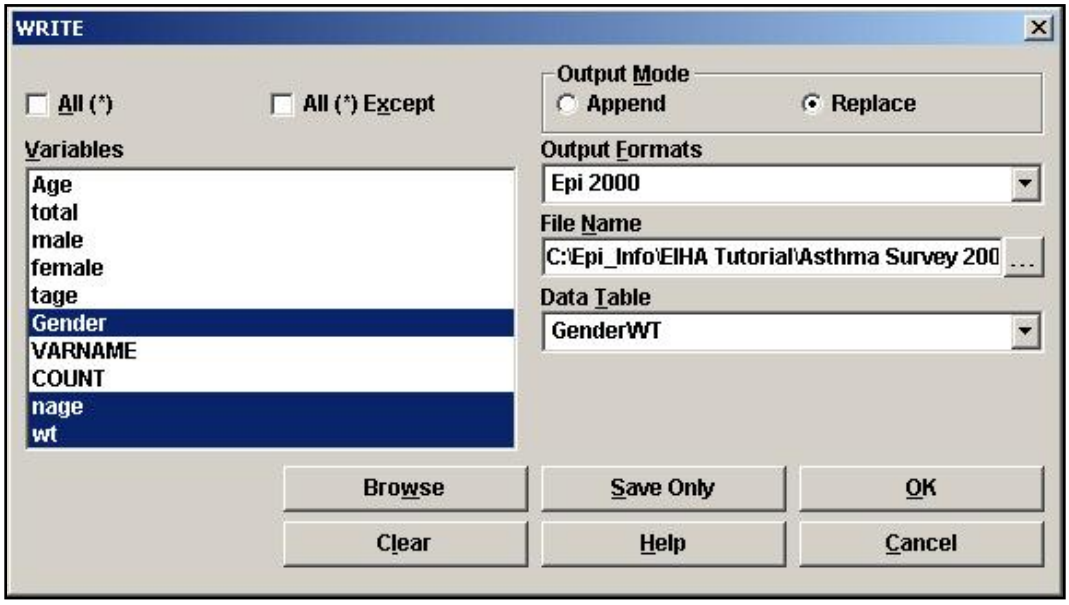
Time to weigh in!

1. Click **Select**. The SELECT dialog box opens.
2. Select **Gender="Male"**.
3. Click **OK**. You should have eight records.
4. Click **Define** and create a new variable called WT.
5. Click **Assign**. The ASSIGN dialog box opens.
6. From the Assign Variable drop-down, select **WT**.
7. In the =Expression field, create the code: **TXTTONUM(Males)/Count**.
   - You are changing the variable male to a number. Divide it by the variable count, which is the number of male students in the survey.



8. Click **OK**.
9. Click **Write/Export**. The WRITE dialog box opens.
10. Select the variables **Gender**, **NAge**, and **WT**.
11. From the Output Mode section, select **Replace**.
12. From the File Name Browse window, select the **Asthma Final** project.

13. In the Data Table field, type **GenderWT**.



14. Click **OK**.

15. Click **Cancel Select**.

Required Tasks

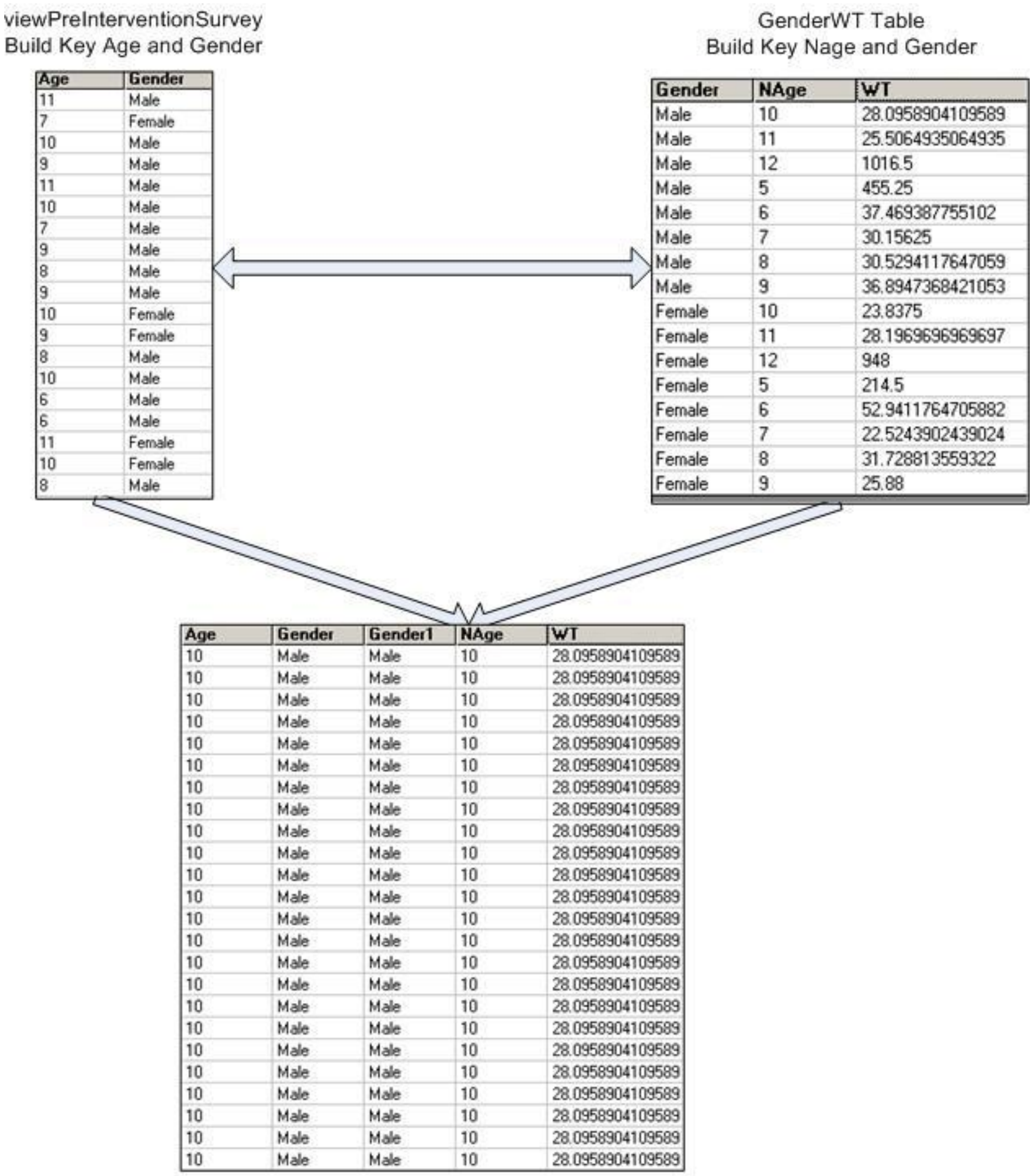
Repeat steps 1–15 above for the females in the survey.

However, instead of replacing the file GenderWT, append the female information.

Creating a PSU Variable

To create the PSU variable, you will open the 800 records from the viewPreInterventionSurvey table and relate them to the GenderWT table you created in the previous section. Prior to completing the RELATE, you must follow some formatting and data management steps.

This is not **P**enn **S**tate **U**niversity!



Age	Gender
11	Male
7	Female
10	Male
9	Male
11	Male
10	Male
7	Male
9	Male
8	Male
9	Male
10	Female
9	Female
8	Male
10	Male
6	Male
6	Male
11	Female
10	Female
8	Male

Gender	NAge	WT
Male	10	28.0958904109589
Male	11	25.5064935064935
Male	12	1016.5
Male	5	455.25
Male	6	37.469387755102
Male	7	30.15625
Male	8	30.5294117647059
Male	9	36.8947368421053
Female	10	23.8375
Female	11	28.1969696969697
Female	12	948
Female	5	214.5
Female	6	52.9411764705882
Female	7	22.5243902439024
Female	8	31.728813559322
Female	9	25.88

Age	Gender	Gender1	NAge	WT
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589
10	Male	Male	10	28.0958904109589

Related Table

1. Click **Read/Import**.

2. Click the **Show All** radio button.

3. Select the new table **GenderWT**.

4. Click **List** to see the computed weight values.

   - In the Gender column, make sure the first letter in the words Male and Female are uppercase. If not, modify the records using the List Update command. Complex Sample routines are case-sensitive.

5. Click **Display** to see your format changes.

6. Click **Read/Import**.

7. Read in the Asthma Final project with your 800 records.

   - Click **List** to verify that the first letter in the words Male and Female are uppercase. If not, modify the records using the List Update command. Complex Sample routines are case-sensitive.

8. Click **Assign** and compute the Age variable. Hint:: Age=YEARS(DOB, 10/25/2005).

9. From the Command Tree Statistics folder, click **Tables**. The TABLES dialog box opens.

10. From the Exposure Variable drop-down, select **Age**.

11. From the Outcome Variable drop-down, select **Gender**.

12. Click **OK**.

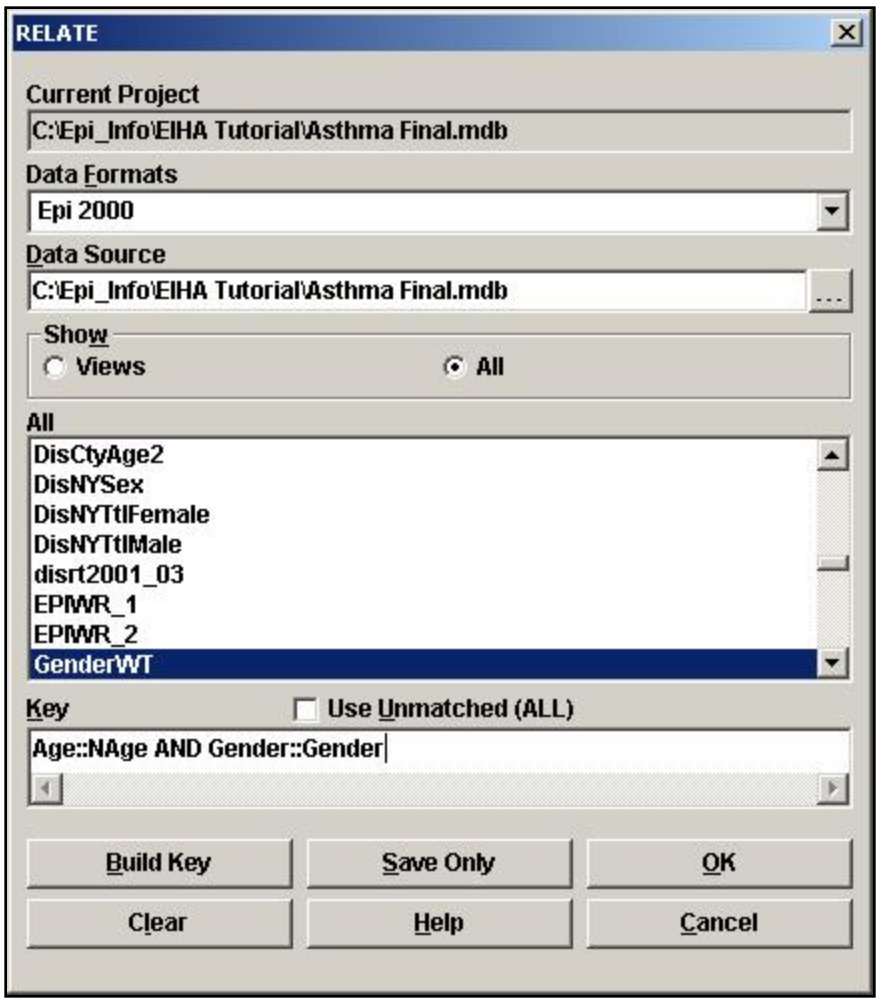
    - Look at the table and make sure that all the ages and genders of the data are represented.

    - Now you know that you can use the RELATE command to relate your GenderWT table to the survey records.

GENDER			
Age	**Female**	**Male**	**TOTAL**
5	8	4	12
Row %	66.7	33.3	100.0
Col %	2.0	1.0	1.5
6	34	49	83
Row %	41.0	59.0	100.0
Col %	8.4	12.4	10.4
7	82	64	146
Row %	56.2	43.8	100.0
Col %	20.2	16.2	18.3
8	59	68	127
Row %	46.5	53.5	100.0
Col %	14.5	17.3	15.9
9	75	57	132
Row %	56.8	43.2	100.0
Col %	18.5	14.5	16.5
10	80	73	153
Row %	52.3	47.7	100.0
Col %	19.7	18.5	19.1
11	66	77	143
Row %	46.2	53.8	100.0
Col %	16.3	19.5	17.9
12	2	2	4
Row %	50.0	50.0	100.0
Col %	0.5	0.5	0.5
TOTAL	406	394	800
Row %	50.8	49.3	100.0
Col %	100.0	100.0	100.0

13. Click **Relate**. The RELATE dialog box opens.

14. Select the table **GenderWT**. For this Relate, you will type the code into the Key field.



15. Move the cursor into the Key field.

16. Type the following Relate: **Age::NAge AND Gender::Gender**.

- Relates for redefined or new variables must be typed into the Program Editor or the Key field.

17. Click **OK**.

- Code will appear in the Program Editor as RELATE GenderWT Age::Nage AND Gender::Gender.

- Now you need to define a primary sampling unit or PSU. In this example, PSU will be the SchoolNum and Age.

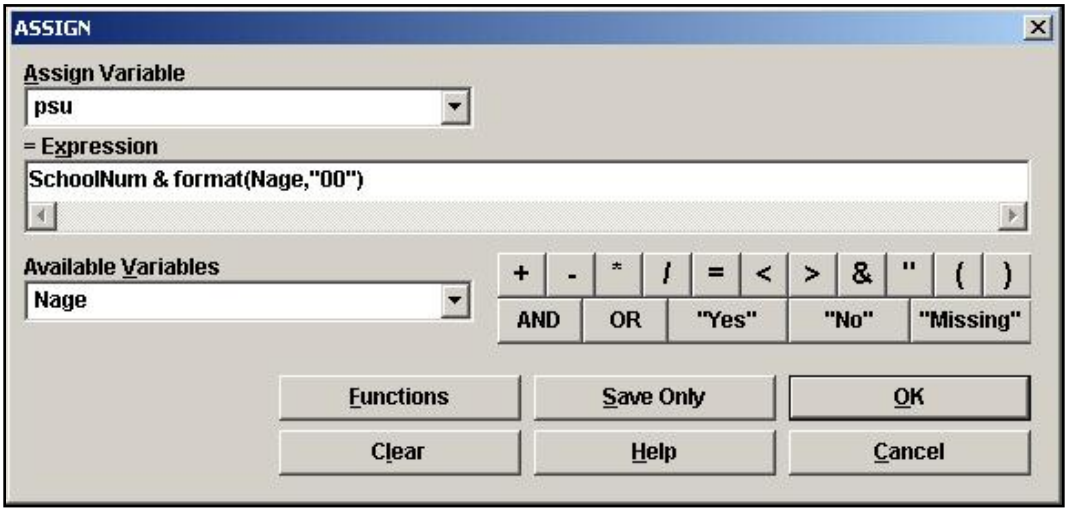
18. Click **Define** and create a new variable called PSU.

19. Click **Assign**. The ASSIGN dialog box opens.

20. From the Assign Variable drop-down, select **PSU**.

21. In the =Expression field, type with the keyboard or use the drop-down menus and operator buttons to create the following expression:

SchoolNum & format(nage,"00")



22. Click **OK**.

Computing the County Child Prevalence

You now have all the variables and data needed to compute the prevalence for the children in the county.

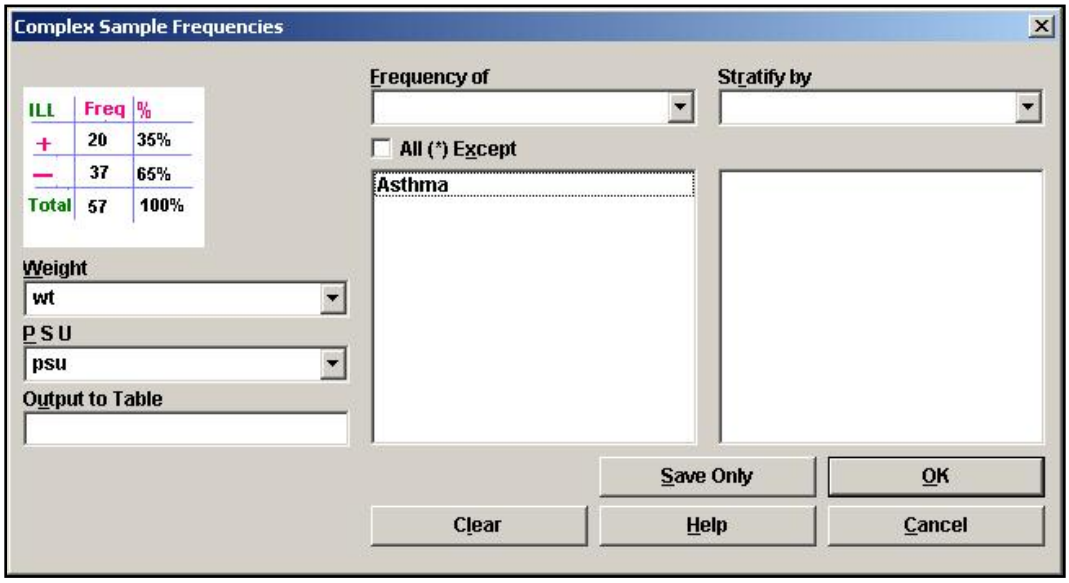
Now it's the kid's turn.

1. From the Command Tree Advanced Statistics folder, click **Complex Sample Frequencies**.

2. From the Frequency Of drop-down, select **Asthma**.

3. From the Weight drop-down, select **WT**.

4. From the PSU drop-down, select **PSU**.



5. Click **OK**. Results appear in the Output window:

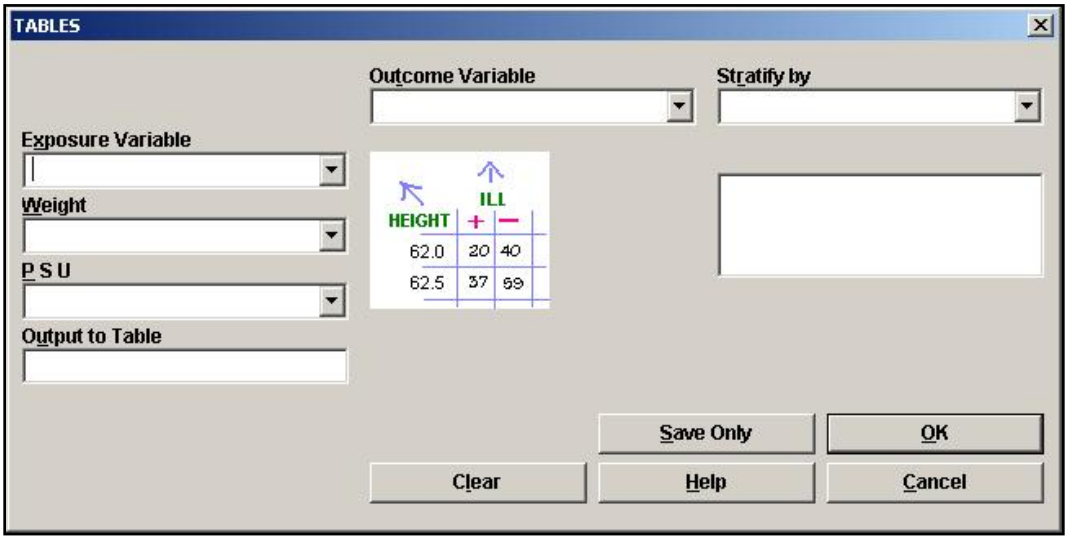
Asthma	TOTAL
Yes	91
Row %	100.000
Col %	9.978
SE %	1.611
LCL %	6.692
UCL %	13.264
No	709
Row %	100.000
Col %	90.022
SE %	1.611
LCL %	86.736
UCL %	93.308
TOTAL	800
Design Effect	2.312

Prevalence

Confidence Interval

Compute for the Subgroup Gender

1. From the Command Tree Advanced Statistics folder, click **Complex Sample Tables**. The TABLES dialog box opens.



2. From the Exposure Variable drop-down, select **Gender**.

3. From the Outcome Variable drop-down, select **Asthma**.

4. From the Weight drop-down, select **WT**.

5. From the PSU Variable drop-down, select **PSU**.

6. Click **OK**. The results appear in the Output window.

Gender	Asthma		
	Yes	No	TOTAL
Female	43	363	406
Row %	10.599		
Col %	51.418		
SE %	3.049	3.049	
LCL %	4.381	83.184	
UCL %	16.816	95.619	
Design Effect	3.982	3.982	
Male	48	346	394
Row %	9.396		
Col %	48.582		
SE %	1.481	1.481	
LCL %	6.375	87.583	
UCL %	12.417	93.625	
Design Effect	1.016	1.016	
TOTAL	91	709	800
Row %	9.978	90.022	100.000
Col %	100.000	100.000	100.000
SE %	1.611	1.611	
LCL %	6.692	86.736	
UCL %	13.264	93.308	
Design Effect	2.312	2.312	

Prevalence for Females

Prevalence for Males

Statistical Table Results for Child County Prevalence by Gender

CTABLES COMPLEX SAMPLE DESIGN ANALYSIS OF 2 X 2 TABLE

Odds Ratio (OR) 1.143
Standard Error (SE) 0.400
95% Conf. Limits (0.53, 2.467)

Risk Ratio (RR) 1.128
Standard Error (SE) 0.382
95% Conf. Limits (0.57, 2.249)
RR = (Risk of Asthma=-1 if Gender=Female) / (Risk of Asthma=-1 if Gender=Male)

Risk Difference (RD%) 1.203
Standard Error (SE) 3.489
95% Conf. Limits (-5.91, 8.318)
RD = (Risk of Asthma=-1 if Gender=Female) - (Risk of Asthma=-1 if Gender=Male)

Other questions you can answer using Analysis with similar steps and data:

- What is the current asthma prevalence among children in the state?
- Is the state childhood asthma prevalence higher in any population subgroups?
- How is the county childhood asthma prevalence different from the state?

Mortality Rates

As an illustration to provide with your proposal, you will graph the mortality rates for the county and the state.

H. Do asthma mortality rates for the state and county differ significantly?

Two pre-made tables are inside the Asthma Final project. One is a table called Population, and the other Numerators. Each table was created using the Summarize and Write commands with information from CDC WONDER and MS Excel tables with New York-specific data. New variables were defined to reformat and hold data to use in this lesson.

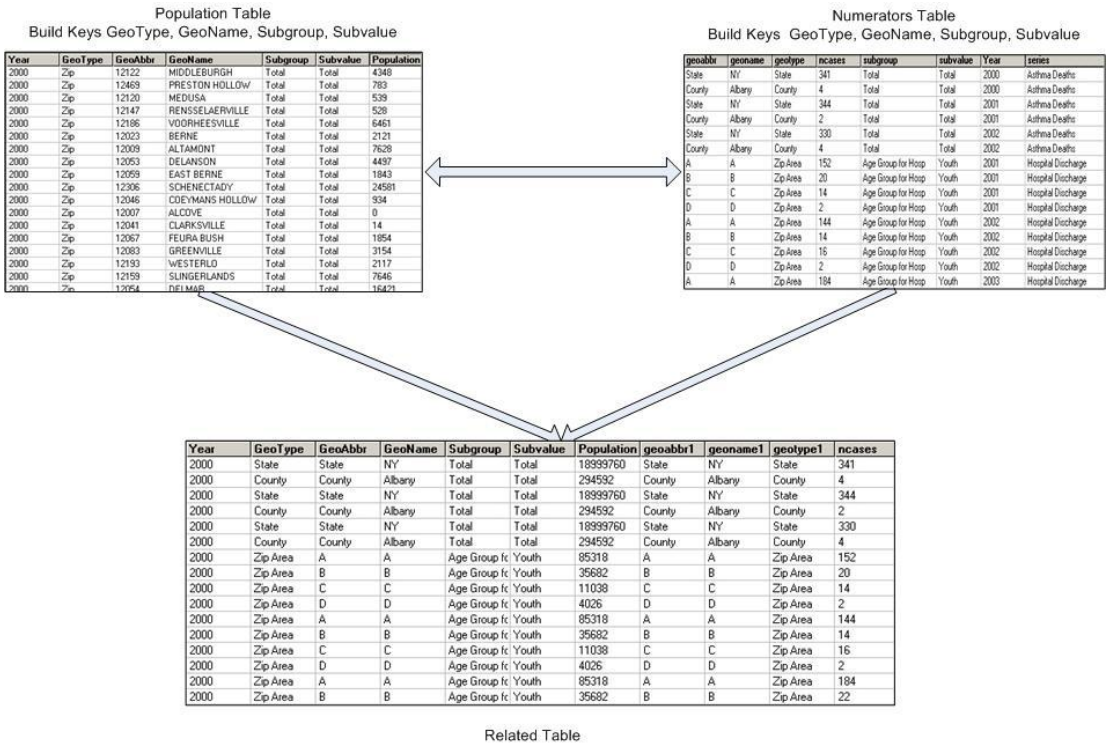
The wide-ranging OnLine Data for Epidemiologic Research or WONDER, is an easy-to-use Internet system that makes the information resources of CDC available to public health professionals and the public. It provides access to a wide array of public health information. (Source: http://wonder.cdc.gov/wonder/help/faq.html#1)

You will create a bar graph of the mortality rates for the county and state. To compute the results, use the RELATE command and join two tables so you can divide the total asthma deaths in the county and the state by the total population of the county and the state. You will create a related table to join the population data from WONDER to the data table that contains the number of deaths in the county and state. Once the rates are computed, you can easily plot them onto a graph.

Now record the mortality figures.

1. Click **Read/Import**. The READ dialog box opens.
2. Click the **Show All** radio button.
3. Select the **Population** table.
4. Click **OK**. You should have 233 records.
5. Click **List** to view the contents.

- Now you will use the RELATE command to link the information in the population (denominator) to the numerators in order to compute a rate.



Year	GeoType	GeoAbbr	GeoName	Subgroup	Subvalue	Population
2000	Zip	12122	MIDDLEBURGH	Total	Total	4348
2000	Zip	12469	PRESTON HOLLOW	Total	Total	783
2000	Zip	12120	MEDUSA	Total	Total	539
2000	Zip	12147	RENSSELAERVILLE	Total	Total	528
2000	Zip	12186	VOORHEESVILLE	Total	Total	6461
2000	Zip	12023	BERNE	Total	Total	2121
2000	Zip	12009	ALTAMONT	Total	Total	7628
2000	Zip	12053	DELANSON	Total	Total	4497
2000	Zip	12059	EAST BERNE	Total	Total	1843
2000	Zip	12306	SCHENECTADY	Total	Total	24581
2000	Zip	12046	COEYMANS HOLLOW	Total	Total	934
2000	Zip	12007	ALCOVE	Total	Total	0
2000	Zip	12041	CLARKSVILLE	Total	Total	14
2000	Zip	12067	FEURA BUSH	Total	Total	1854
2000	Zip	12083	GREENVILLE	Total	Total	3154
2000	Zip	12193	WESTERLO	Total	Total	2117
2000	Zip	12159	SLINGERLANDS	Total	Total	7646
2000	Zip	12054	DELMAR	Total	Total	16421

geoabbr	geoname	geotype	ncases	subgroup	subvalue	Year	series
State	NY	State	341	Total	Total	2000	Asthma Deaths
County	Albany	County	4	Total	Total	2000	Asthma Deaths
State	NY	State	344	Total	Total	2001	Asthma Deaths
County	Albany	County	2	Total	Total	2001	Asthma Deaths
State	NY	State	330	Total	Total	2002	Asthma Deaths
County	Albany	County	4	Total	Total	2002	Asthma Deaths
A	A	Zip Area	152	Age Group for Hosp	Youth	2001	Hospital Discharge
B	B	Zip Area	20	Age Group for Hosp	Youth	2001	Hospital Discharge
C	C	Zip Area	14	Age Group for Hosp	Youth	2001	Hospital Discharge
D	D	Zip Area	2	Age Group for Hosp	Youth	2001	Hospital Discharge
A	A	Zip Area	144	Age Group for Hosp	Youth	2002	Hospital Discharge
B	B	Zip Area	14	Age Group for Hosp	Youth	2002	Hospital Discharge
C	C	Zip Area	16	Age Group for Hosp	Youth	2002	Hospital Discharge
D	D	Zip Area	2	Age Group for Hosp	Youth	2002	Hospital Discharge
A	A	Zip Area	184	Age Group for Hosp	Youth	2003	Hospital Discharge

Year	GeoType	GeoAbbr	GeoName	Subgroup	Subvalue	Population	geoabbr1	geoname1	geotype1	ncases
2000	State	State	NY	Total	Total	18999760	State	NY	State	341
2000	County	County	Albany	Total	Total	294592	County	Albany	County	4
2000	State	State	NY	Total	Total	18999760	State	NY	State	344
2000	County	County	Albany	Total	Total	294592	County	Albany	County	2
2000	State	State	NY	Total	Total	18999760	State	NY	State	330
2000	County	County	Albany	Total	Total	294592	County	Albany	County	4
2000	Zip Area	A	A	Age Group fc	Youth	85318	A	A	Zip Area	152
2000	Zip Area	B	B	Age Group fc	Youth	35682	B	B	Zip Area	20
2000	Zip Area	C	C	Age Group fc	Youth	11038	C	C	Zip Area	14
2000	Zip Area	D	D	Age Group fc	Youth	4026	D	D	Zip Area	2
2000	Zip Area	A	A	Age Group fc	Youth	85318	A	A	Zip Area	144
2000	Zip Area	B	B	Age Group fc	Youth	35682	B	B	Zip Area	14
2000	Zip Area	C	C	Age Group fc	Youth	11038	C	C	Zip Area	16
2000	Zip Area	D	D	Age Group fc	Youth	4026	D	D	Zip Area	2
2000	Zip Area	A	A	Age Group fc	Youth	85318	A	A	Zip Area	184
2000	Zip Area	B	B	Age Group fc	Youth	35682	B	B	Zip Area	22

Small print - tough to see on the computer screen.

6. Click **Relate**. The RELATE dialog box opens.

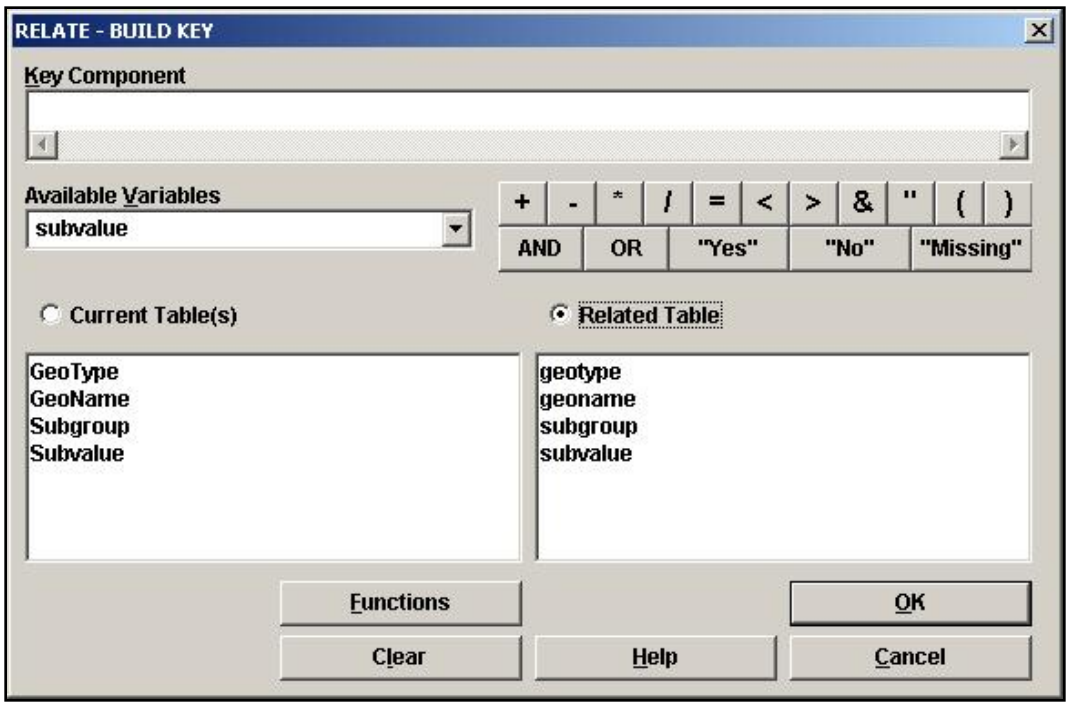
7. Click the **Show All** radio button.

8. Select the **Numerators** table.

9. Click **Build Key**. The RELATE-BUILD KEY dialog box opens.

10. Relate the two tables on the following fields: Geotype, Geoname, Subgroup, and Subvalue.

- The Population table contains the total numbers for the population by each of the subgroups selected. The Numerators table contains the number of cases by series for each of the subgroups selected.
- Geotype contains information on how the information is divided. Most of the values in this field are zip.
- Geoname contains information on the specific breakdown of data by counties or state.
- Subgroup contains information on how the information is divided (e.g., age).
- Subvalue contains the specific subgroup information (e.g., age 5–14).



11. Click **OK**.

12. From the READ dialog box, click **OK**. You should have 42 records.

13. Click **Select**. The SELECT dialog box opens.

14. From the Available Variables drop-down, select **Series**.

15. In the =Expression field, type "**Asthma Deaths**".

16. Click **OK**. You should have six records.

17. Click **Define a**nd create a new variable called Rate.

18. Click **Assign**. The ASSIGN dialog box opens.

19. From the Assign Variable drop-down, select **Rate**.

20. In the =Expression field, type **ncases/population*10000**.

- You can create a rate by dividing the number of cases by the population and multiplying by 10,000.
- You will have one rate, when listed, will appear with an E-02 value. Analysis avoids listing values that begin with 0.0. You can reformat the values to Text to view it the 0.0 format. The value will appear as 0.0 in the graph.

21. Click **OK**.

22. From the Command Tree Statistics folder, click **Graph**. The GRAPH dialog box opens.

23. From the Graphtype drop-down, select **Bar**.

24. From the Main Variable drop-down, select **Year1**.

25. From the Weight Variable drop-down, select **Rate**.

26. From the Bar for Each Value of drop-down, select **GeoName**.

27. Click **OK**. The graph appears in the Epi Graph window.

28. Right click the graph and select **Customization Dialog** from the pop-up menu.

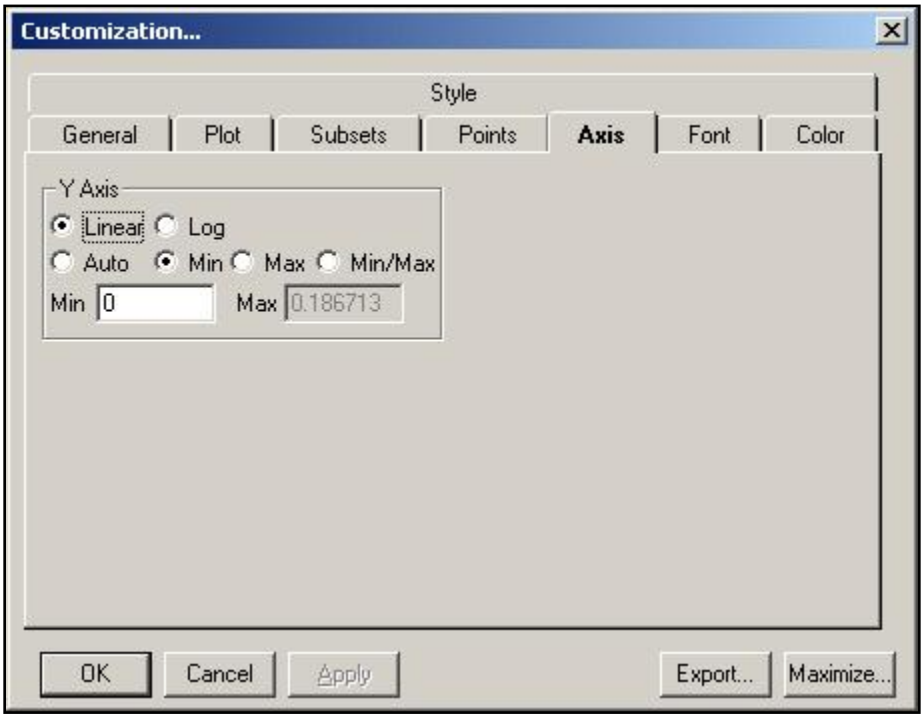
29. From the General tab in the Display section, select the **Both** radio button.

30. Click **Apply**.

31. Select the **Axis** tab.

32. From the Y Axis section, click the **Min** radio button.

33. Highlight the Min field text, and enter the number **0**.

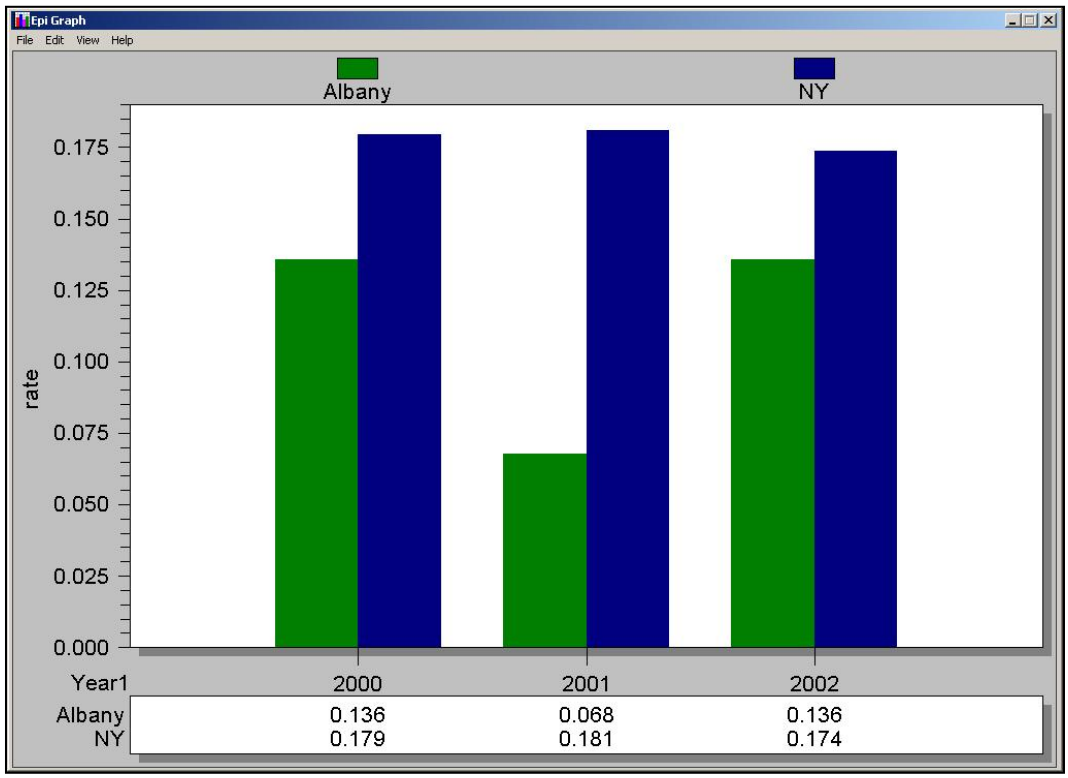


34. Click **Apply**.

35. Click **OK**.

- Notice the Y Axis on the graph has been reset to start at 0.
- Notice the rates for each year are visible at the bottom of the graph.

36. Close Epi Graph to return to the Analysis module.



10-Year Trends — Hospital Discharge Data

As an illustration to provide with your proposal, you will graph the hospitalization rates for 10 years and across age groups.

I. Is the rate of hospital discharge for asthma increasing or decreasing in any age groups?

You have a table in your Asthma Final project that contains 10 years of asthma discharge data by age group. Use this table to plot statistics and look for trends in asthma discharges.

Hospital discharge data was obtained from the Statewide Planning and Research Cooperative System (SPARCS) and was imported into Analysis as an MS Excel file and formatted.

Population information for this exercise was downloaded from the CDC WONDER system.

Ages in the population denominator table were not divided into the same groups as the hospital data. To match the two groups, a third table was created to hold the age data and match the values into three groups. The age table is called AgeProxy. Information from the AgeProxy table was selected from the population table and written out for use with the discharge data to create the 10-year trend graph.

A decade's worth to sift through.

1. Click **Read/Import**.
2. Click the **Show All** radio button.
3. Select the file **NYDisAge10yr**.
4. Click **OK**. You should have 30 records.
5. Click **List** to see the contents of the file. Notice the Series name is NY Discharges by Age.
6. Click **Relate**. The RELATE dialog box opens.
7. Locate and select the AgeProxy file.
8. Click **Build Key**. The RELATE-BUILD KEY dialog box opens.
9. Relate the NYDisAge10yr table and the AgeProxy table on the following variables: Geotype, Geoname, Subgroup, and Subvalue.
    - Geotype is State.
    - Geoname is NY.

- Subgroup is Age.
- Subvalue is the breakdown by Age Groups 0–17, 18–64, and 65+.

10. Click **OK**.

11. From the RELATE dialog box, click **OK**. You should have 30 records.

12. Click **Define** and create a new variable called Discharge.

13. Click **Assign**.

14. From the Assign Variable drop-down, select **Discharge**.

15. In the =Expression, type **ncases/population*10000**.

16. Click **OK**.

17. Click **Graph**. The GRAPH dialog box appears.

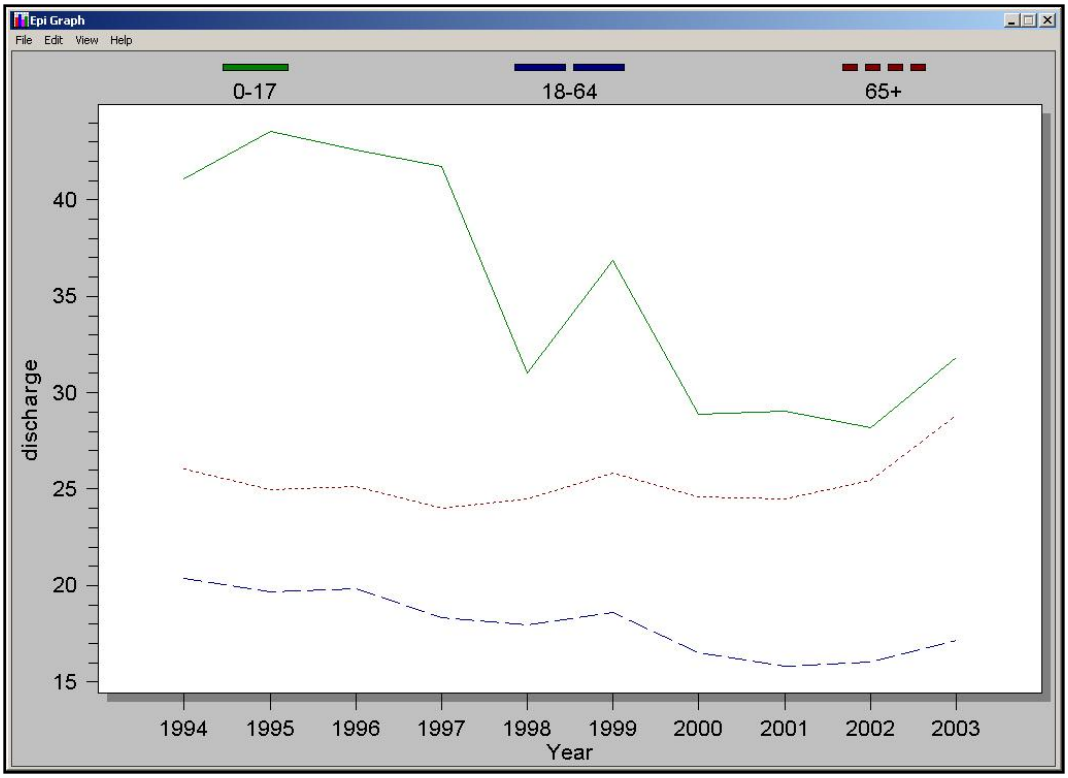
18. From the Graph Type drop-down, select **Line**.

19. From the Main Variable drop-down, select **Year**.

20. From the Weight Variable drop-down, select **Discharge**.

21. From the Line for Each Value Of drop-down, select **Subvalue**.

22. Click **OK**. Epi Graph opens with the results.


Epi Graph
File Edit View Help
0-17
18-64
65+
discharge
40
35
30
25
20
15
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
Year

Three-Year Rates — Hospital Discharge Data

As an illustration to provide with your proposal, you want to graph the hospital discharge rates for three years and across age groups.

J. Are there significant differences in the three-year asthma discharge rates by age for the state versus the county?

The ability to Read, Write, and Relate tables is an important and useful concept in using Analysis. For this exercise, you will read in two tables, one containing state discharge information, the other containing county discharge information. You will then use the Write command to combine the two tables in order to relate them to the AgeProxy table from the earlier lesson. The Age Proxy table contains the population numbers divided into age groups required to compute the rates.

They say stuff happens in threes!

1. Click **Read/Import**.
2. Click the **Show All** radio button.
3. Select the table **DisCtyAge**.
4. To see how the table is arranged, click **List**.
5. Click **Write/Export**.
6. From the Output Mode section, click **Replace**.
7. From the File Name Browse window, select **Asthma Final**.
8. In the Data Table field, type **Dis3Yrs**.
9. Click **OK**.
10. **Read/Import** the table NY3Yr.
11. To see how the table is arranged, click **List**
12. Click **Write/Export**.
13. From the Output Mode section, click **Append**.
14. From the File Name Browse window, select **Asthma Final**.
15. From the Data Table drop-down, select **Dis3Yrs**.

16. Click **OK**.

17. Click **Read/Import**.

18. Click the **Show All** radio button.

19. Select the **Dis3Yrs** table.

20. Click **OK**. You should have 18 records.

21. Click **Relate**.

22. Select the **AgeProxy** table.

23. Click **Build Key** and join on the following fields: Geotype, Geoname, Subgroup, and Subvalue.

- Geotype is County or State.
- Geoname is Albany or NY.
- Subgroup is Age.
- Subvalue is the age group breakdowns 0–17, 18–64, and 65+.

24. Click **OK**.

25. From the RELATE dialog box, click **OK**.

26. Click **Define** and create a new variable called Discharge.

27. Click **Assign** to create the following code: discharge=ncases/Population*10000.

28. Click **Graph**. The GRAPH dialog box opens.

29. Select the Graph Type of Line.

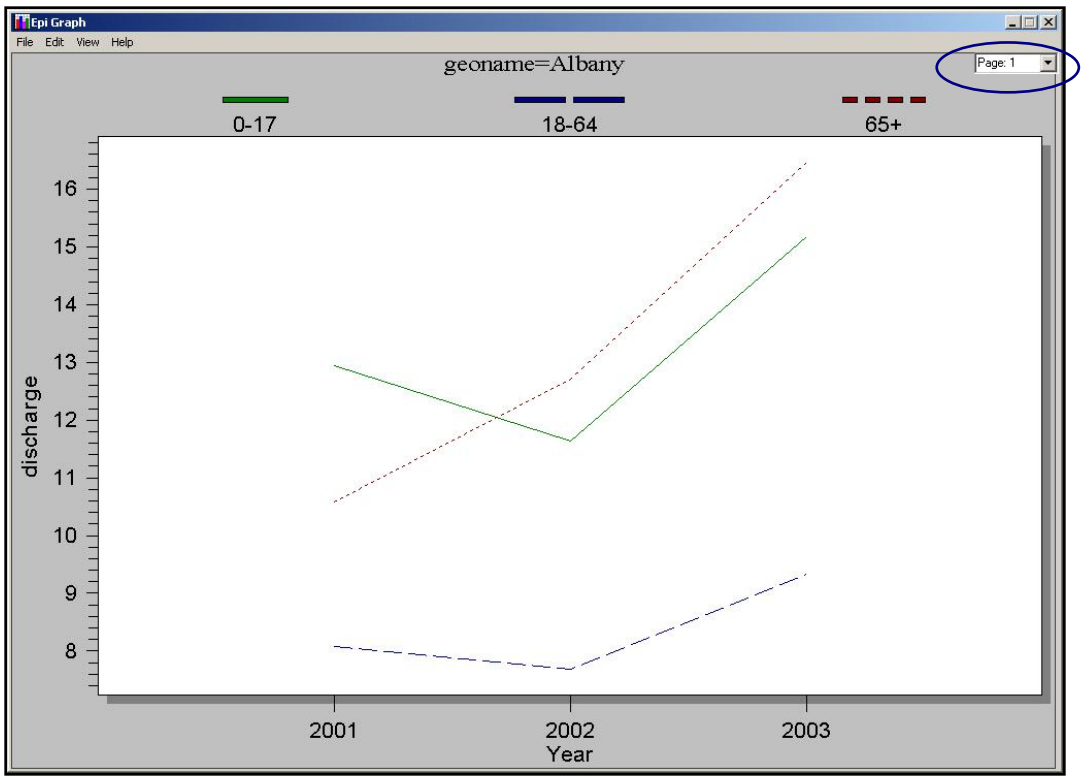
30. From the Main Variable drop-down, select **Year**.

31. From the Weight Variable drop-down, select **Discharge**.

32. From the Line for Each Value of, select **Subvalue**.

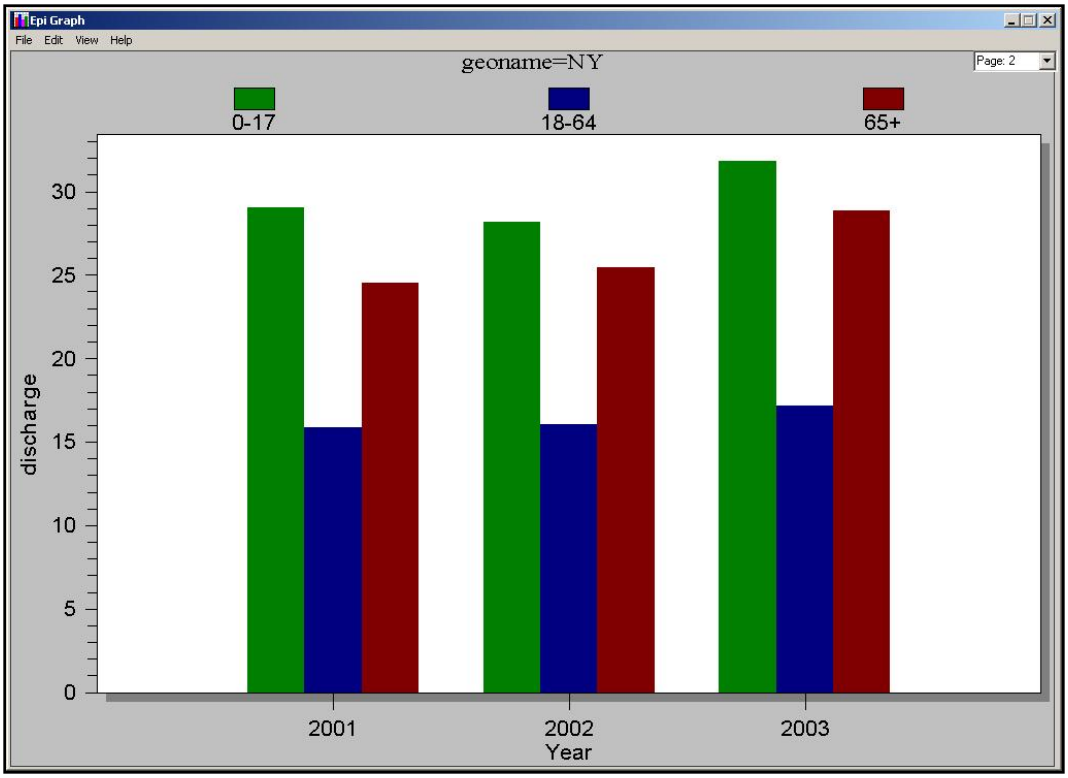
33. From the One Graph for Each Value of drop-down, select **GeoName**.

34. Click **OK**. Epi Graph opens and contains two graphs: one for NY, the other for Albany.



- Use the Page drop-down menu at the top right of the page to go between the two graphs.
- Use the Customization dialog box to change the graph types to Bar. Look at the difference between the two graphs.
- Use the Customization dialog box to change the Axis Min to 0. Look at the difference.

Your graph should look like the following:


Epi Graph
File Edit View Help
geoname=NY
Page: 2
0-17
18-64
65+
30
25
20
15
10
5
0
discharge
2001
2002
2003
Year

Mapping Rates

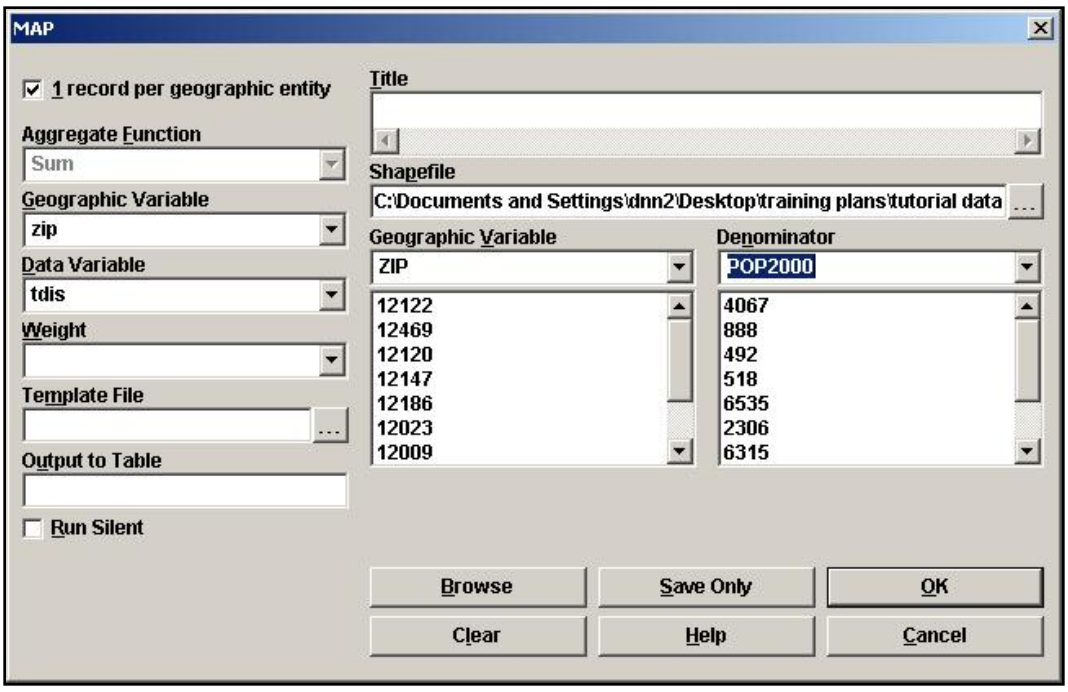
As an illustration to provide with your proposal, you will map the hospitalization rates by zip code for your state.

K. Based on the map with asthma hospitalization rates by zip code, where are the high risk areas?

In order to map high-risk areas, you will combine discharge data arranged by zip code with a map .DBF file.

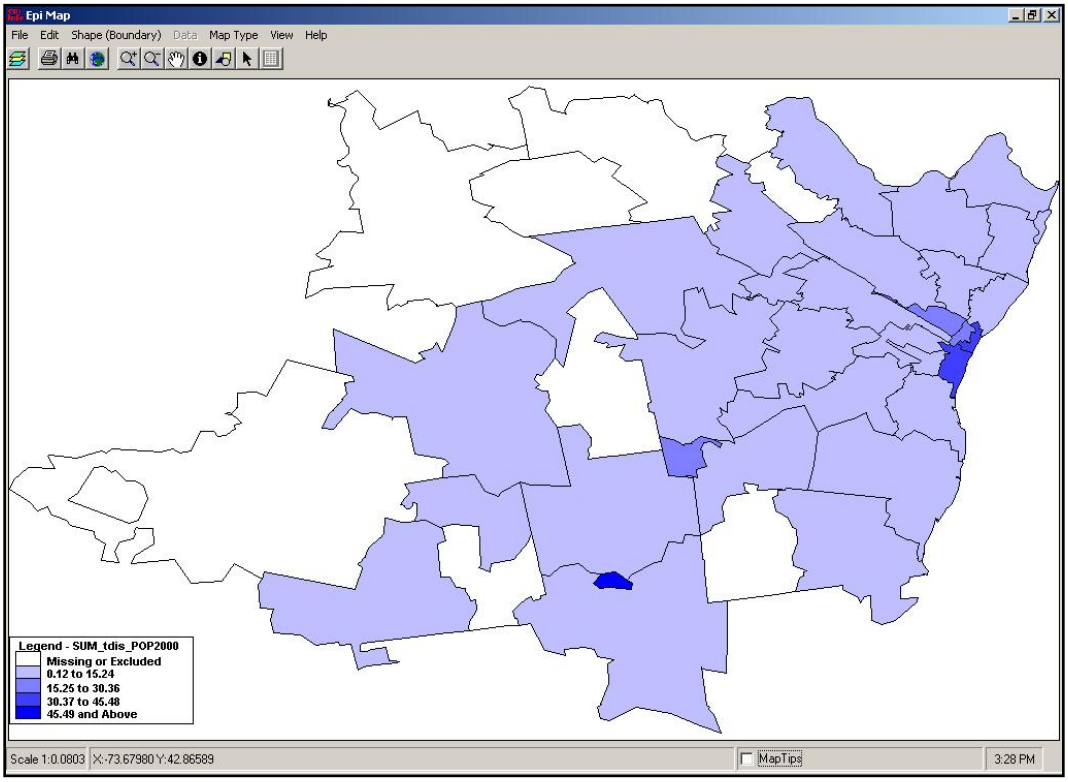
Time to get rated again.

1. Click **Read/Import**.
2. Click the **Show All** radio button.
3. Select the **Dis0103byZip** table.
4. Click **List** to view the variables.
   - Notice the table is arranged by county zip code for three years. Also, notice that the age groups have been recoded as Youth, Adult, and Senior. This is another way to work with age groups.
5. Click **Define a**nd create a new variable called **TDis**.
6. Click **Assign** and create the following code: **TDis=(youth+adult+senior)*10000/3**.
   - You are adding the number of discharges for each age group, multiplying by the population factor, and dividing by the three years in the table. You are going to map this data and get your population denominator from the map data (census information for the year 2000).
7. From the Command Tree Statistics folder, click **Map**. The MAP dialog box opens.
8. Select the **1 Record per Geographic Entity** checkbox.
9. From the Geographic Variable drop-down, select **Zip**.
10. From the Data Variable drop-down, select **TDis**.
11. From the Shapefile Browse window, select the **ALBZCTA** shapefile.
12. From the Geographic Variable drop-down, select **Zip**.
13. From the Denominator drop-down, select **POP2000**.



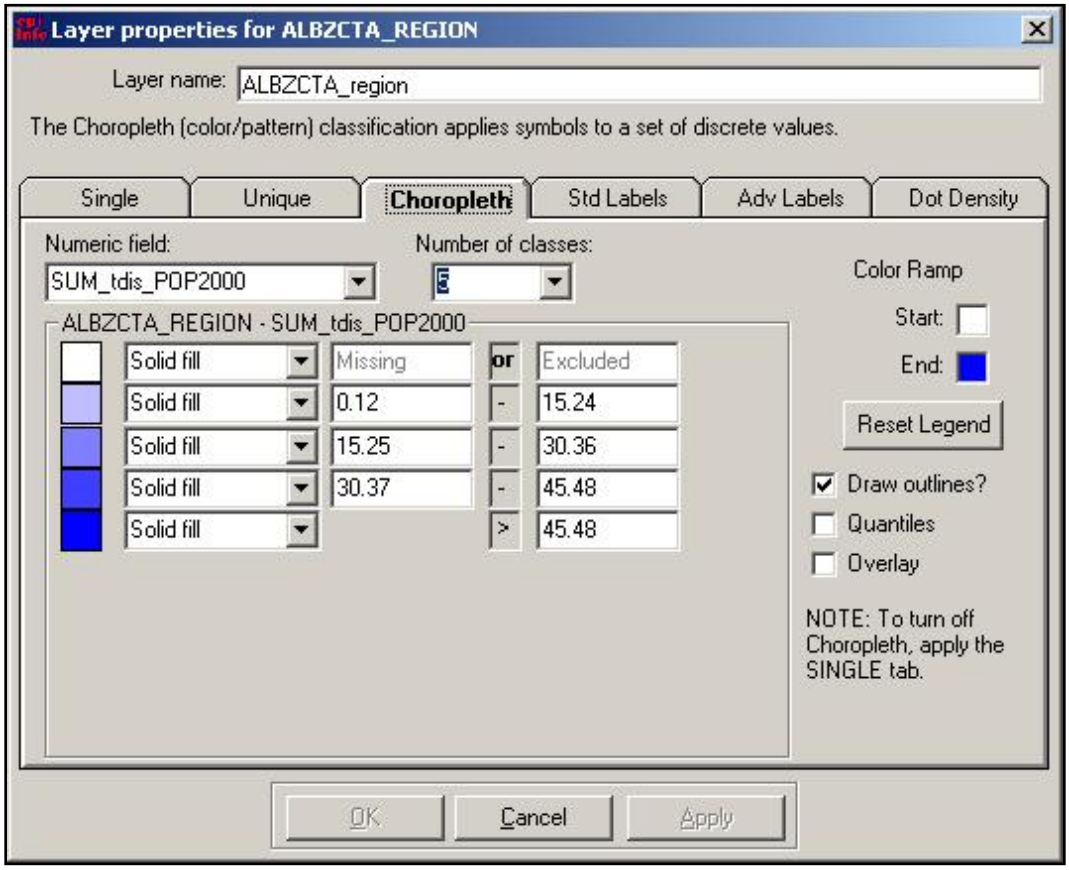
14. Click **OK**.

15. Click **Continue** at the Incomplete Join screen. Epi Map opens with the rates mapped in a choropleth.



Customize the Map

1. Open the **Map Manager**.
2. Click **Properties**. The Choropleth tab opens.

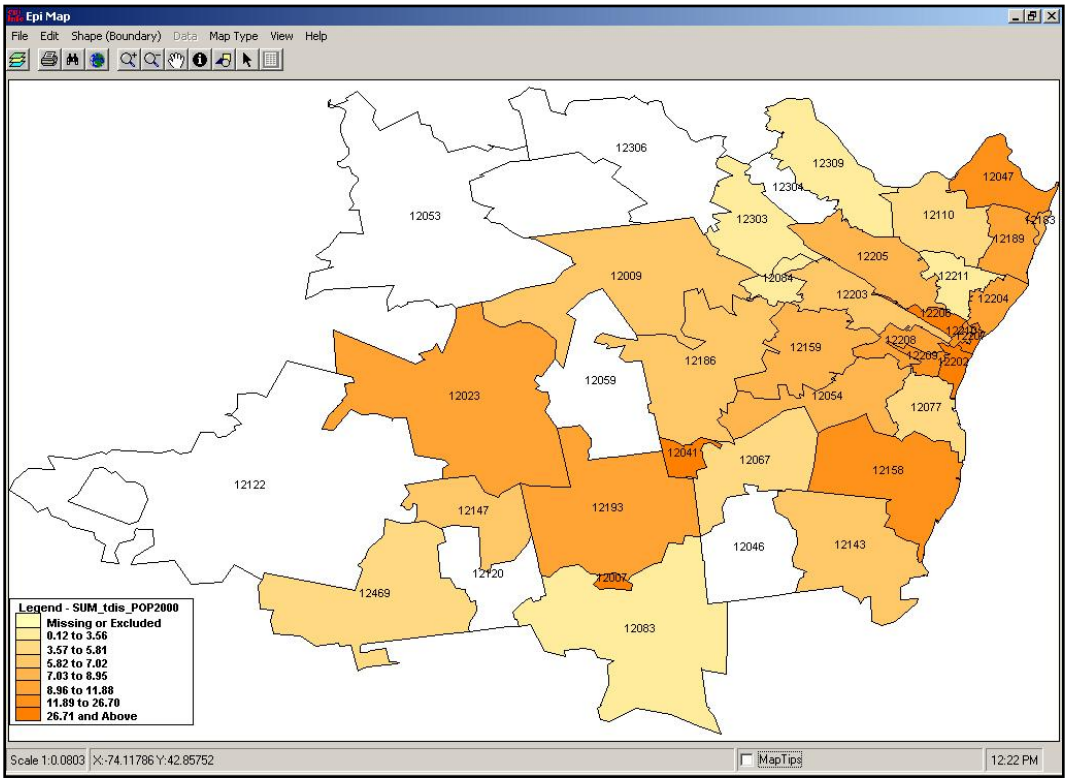


3. From the Number of Classes drop-down, select **8**.
4. Select the **Quantiles** checkbox.
5. Select the **Overlay** checkbox.
6. Click **Reset Legend**.
7. You have the option of changing the Color Ramp.
8. Click **Reset Legend**.
9. Click **Apply**.
10. Click the **Standard Labels** tab.
11. From the Text field drop-down, select **Zip**.

12. Click **Apply**.

13. Close the **Properties** dialog box.

14. Close the **Map Manager.**



Like the above map, the Legend is difficult to see.

Income and Zip Code Scatter Plot

As an illustration to provide with your proposal, you will graph the hospitalization rates across median income brackets and see if a correlation exists.

L. Does income affect the rate of hospitalization for asthma?

Look at median family income and per capita income by zip code. You can create a scatter plot of income by zip as a comparison with discharges by zip.

Oh, we just continue to zip along!

1. Click **Read/Import**. You should still be in the project Asthma Final.MDB.

2. Click the **Show All** radio button.

3. Select the **IncomeData** table. You should have 38 records.

4. Click **List** to view the data.

5. Click **Read/Import**.

6. Click the **Show All** radio button.

7. Select the **RatebyZip** table. You should have 95 records.

8. Click **List** to view the data.

9. Click **Select**. The SELECT dialog box opens.

10. In the Select Criteria field, type **Year=2001**. You should have 35 records.

11. Click **Relate**.

12. Select the **IncomeData** table.

13. Click **Build Key**. The BUILD KEY dialog box opens.

14. Join the tables on the variables Zip and Zip2.

- Zip and Zip2 are the zip code fields in the RatebyZip and IncomeData tables.

15. Click **OK**. You should have 25 records.

16. Click **Relate**. The RELATE dialog box opens.

17. From the Data Formats drop-down, select **dBase IV**.

- You will create a relate using the map file .DBF.

18. Click the **Data Source Browse**  button.

19. From the Relate window, select the **ALBZCTA_region.DBF**.

20. Click **Open**. A data format message appears.

21. Click **OK**. You need to manually assign the Build Key for this Relate.

22. Move your cursor to the Key field located in the RELATE dialog box.

23. In the Key field, type **zip::zip**.

- You are relating a third table on the Zip Code field.

24. Click **OK**. The temporary link window opens.

25. Click **OK** to make a temporary link. You should have 25 records.

26. Click **Define.** Create a new variable called NZip.

27. Click **Assign**. The ASSIGN dialog box opens.

28. From the Assign Variable drop-down, select **NZip**.

29. In the =Expression field, type **tdis/pop2000**.

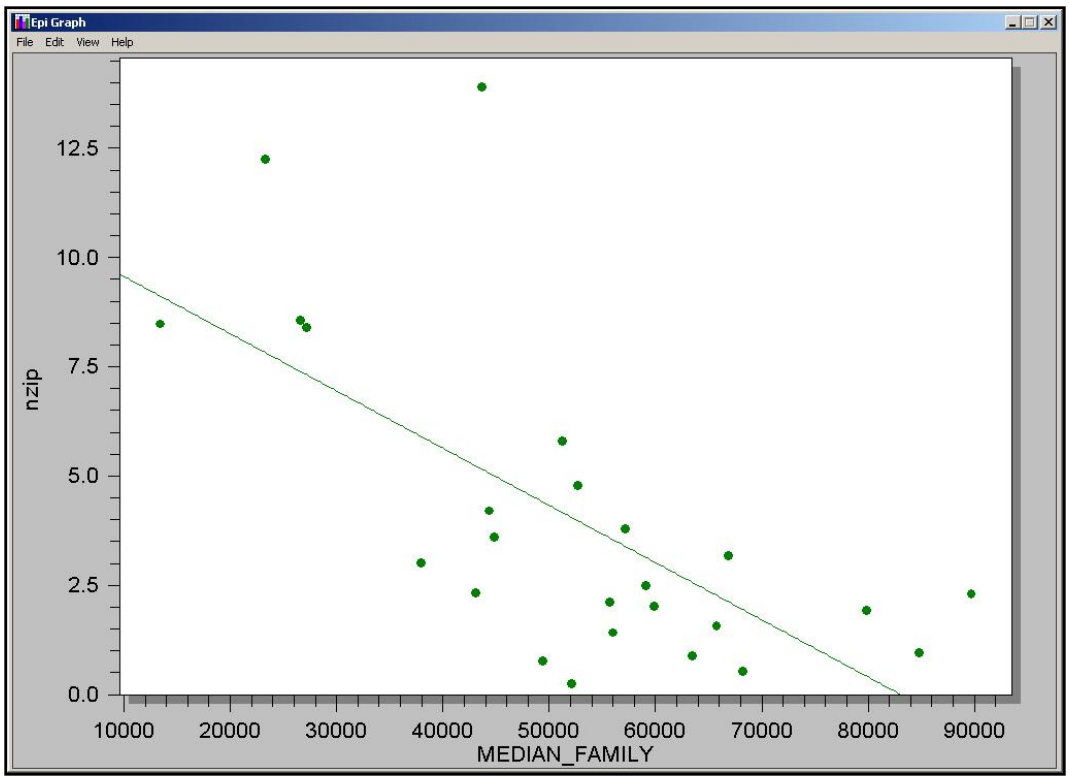
30. Click **OK**.

31. Click **Graph**.

32. From the Graph Type drop-down menu, select **Scatter XY**.

33. From the Variables drop-down, select **Median Family** and **NZip**.

34. Click **OK**. The Epi Graph opens.



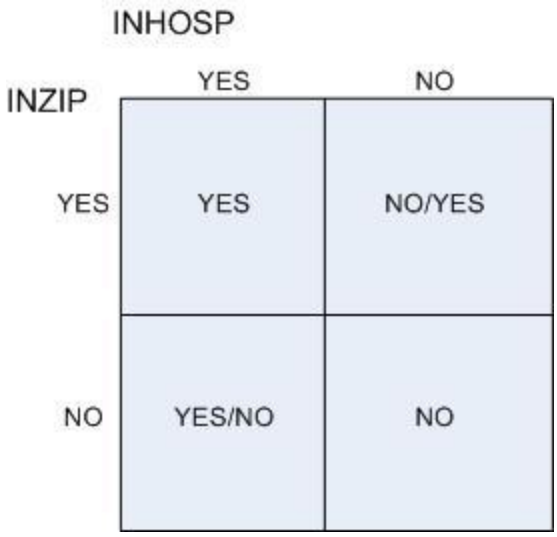
Risk Ratio

To focus on certain areas of the state for your proposal, you want to know if there are any locations by zip code with a higher risk for developing asthma than other areas. You also want to develop a .PGM that can run on each zip code separately.

M. What is the risk ratio of being hospitalized for asthma if you live in the zip code 12202 versus if you live in 12211?

To compute the risk ratios, you need to create a relationship between the number of discharges in each zip code and total population for the zip code and county. The Tables command enables you to create a 2x2 table and provides risk ratio results for dichotomous questions. To get the risk ratio, compute the differences between the discharge numbers in and out of the requested zip codes.

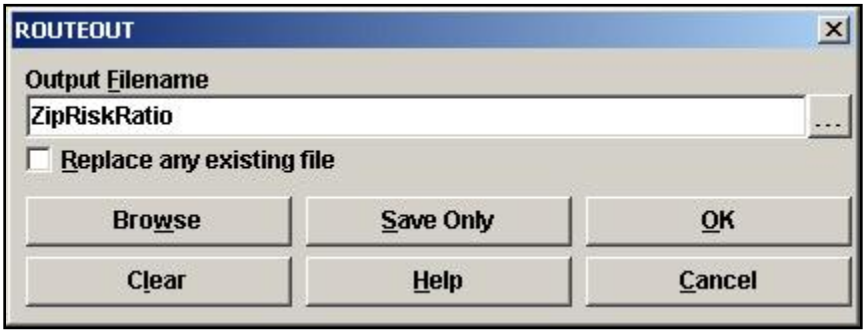
For this example, you will be using the Routeout command to save your results..Save the PGM so the program can run for each individual zip code without rewriting any code.



Time to take another risk.

1. From the Command Tree Output folder, click **Routeout**. The ROUTEOUT dialog box opens.

2. In the Output Filename field, type **ZipRiskRatio**.



3. Click the **Replace Any Existing File** checkbox.

4. Click **OK**.

5. Click **Read/Import**.

6. Select and open the DischargesbyZip table. You should have 91 records.

   - Use LIST to view how the table is organized and the type of data it contains. This table contains the hospital discharges per zip code for Albany County. The zip code field is called Zip.

7. Click **Read/Import**.

8. Select and open the **Population** table.

9. Click **Relate**. The RELATE dialog box opens.

10. Select the **DischargesbyZip** file.

11. Click **Build Key**. The RELATE dialog box opens.

12. Relate the tables on the **GeoAbbr** and **Zip** fields.

    - GeoAbbr is the zip code field in the Population table.

    - Zip is the zip code field in the DischargesbyZip table.

13. Click **OK**.

14. From the RELATE dialog box, click **OK**. You should have 185 related records.

    - In order to select zip codes repeatedly you need to define a new variable, which can be updated for each individual zip as needed.

15. Click **Define.** Create a new standard variable called TheZip.

16. Click **Assign**. The ASSIGN dialog box opens.

17. From the Assign Variable drop-down, select **TheZip**.

18. In the =Expression field, type **"12202"**.

19. Click **OK**.

20. Click **Define.** Create a new variable called InZip.

21. From the Command Tree Select/If folder, click **If**. The IF dialog box opens.

    - Create a statement that says IF the value of zip is 12202, then the value is Yes, else the value is No. You need Yes/No fields to create the 2x2 table.

22. In the If Condition field, type the code: **Zip=TheZip**.

23. Click **THEN**. The Command Tree activates.

24. Click **Assign**. The ASSIGN dialog box opens.

25. From the Assign Variable drop-down, select **InZip**.

26. In the = Expression field, click **Yes**.

27. Click **Add**. The IF dialog box opens.

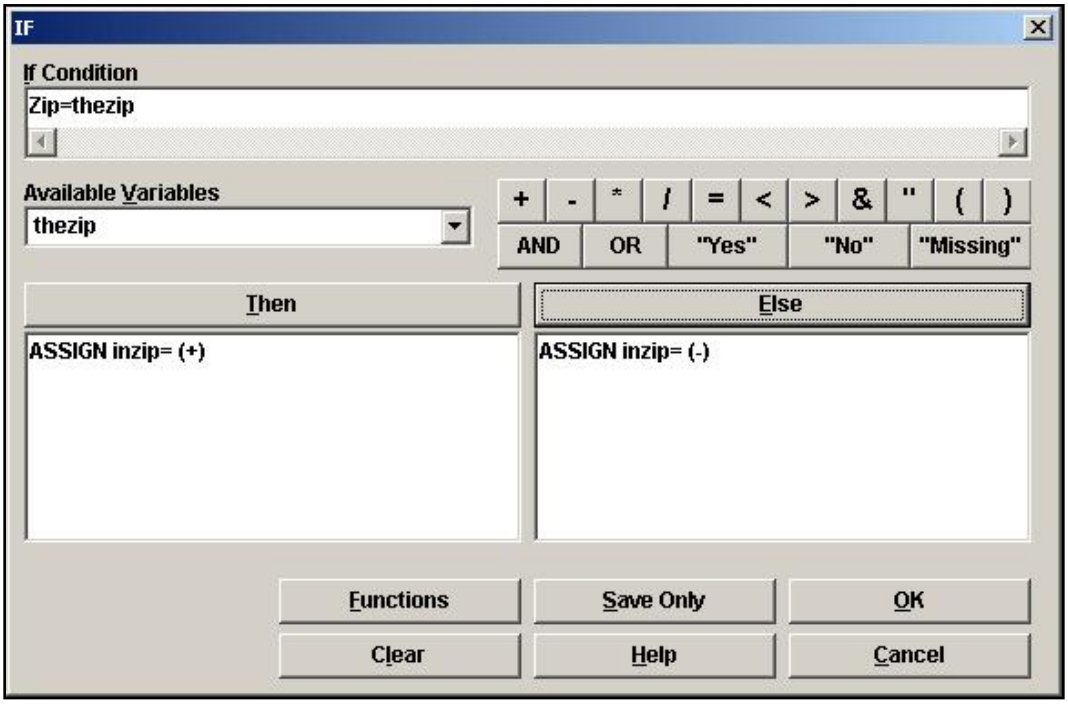
28. Click **ELSE**.

29. Click **Assign**. The ASSIGN dialog box opens.

30. From the Assign Variable drop-down, select **InZip**.

31. In the = Expression field, click **No**.

32. Click **Add**. The IF dialog box opens.



33. Click **OK**. The Program Editor code should read:

```
IF zip=thezip THEN
    ASSIGN inzip= (+)
ELSE
    ASSIGN inzip= (-)
END
```

34. Click **Select**. The SELECT dialog box opens.

35. In the SELECT CRITERIA field, type **subgroup="total".**

36. Click **OK**. You should have 37 records.

- Summarize the records to get the population totals for the hospital and population data in one table.

37. From the Command Tree Statistics folder, click **Summarize**. The SUMMARIZE dialog box opens.

38. From the Aggregate drop-down, select **Sum**.

39. From the Variable drop-down, select **Population**.

40. In the Into Variable field, type **Pop**.

41. Click **Apply**.

42. From the Aggregate drop-down, select **Sum**.

43. From the Variable drop-down, select **Y2000**.

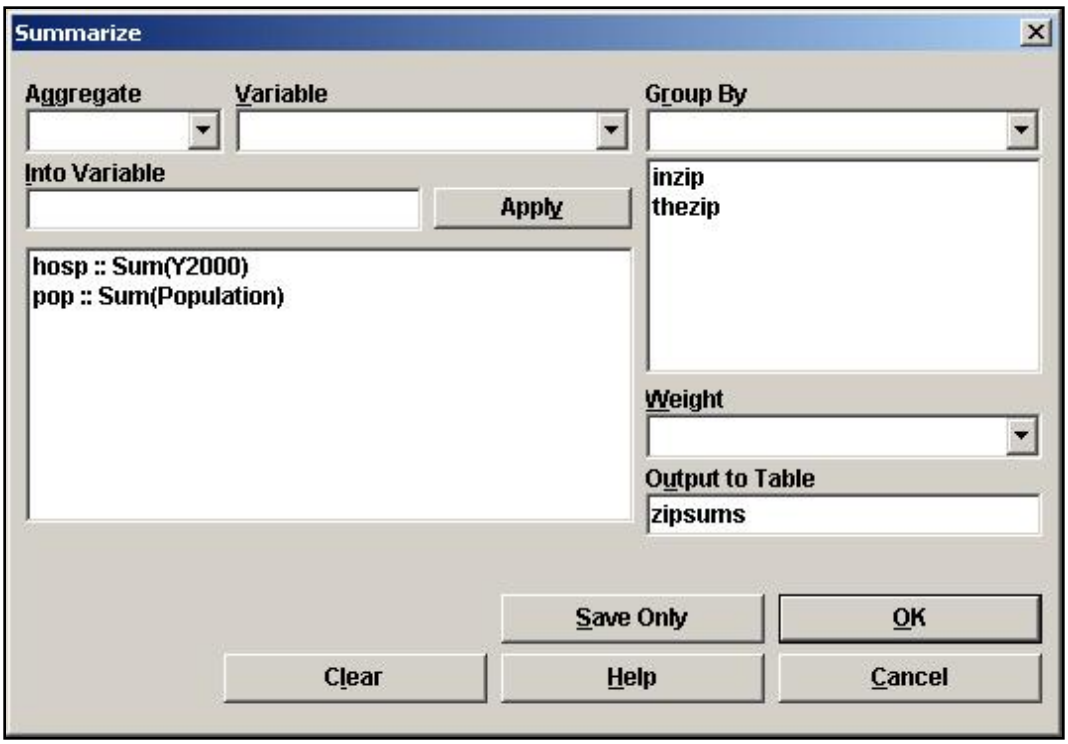
- This is the number of discharges per zip for the year 2000. You are summarizing them to get the total number of discharges for the year 2000 in Albany County.

44. In the Into Variable field, type **Hosp**.

45. Click **Apply**.

46. From the Group By drop-down, select **InZip** and **TheZip**.

47. In the Output to Table field, type **ZipSums**.



48. Click **OK**.

49. Click **Read/Import** and select the **ZipSums** table.

50. Click **List**. The table should look like the following:

inzip	thezip	pop	hosp
No	12202	349987	369
Yes	12202	10021	48

51. Click **Define** and create a new variable called Num.

52. Click **Define** and create a new variable called InHosp.

53. Click **Assign** and create the code: InHosp=Yes.

54. Click **Assign** and create the code: Num=Hosp.

55. Click **Write/Export**.

56. In the Variables field, select **InZip, TheZip, Num**, and **InHosp**.

57. Select the Output Mode **REPLACE**.

58. In the Data Table field, type **ZipSums2**.

59. Click **OK**.

- Create the negative values.

60. Click **Assign** and create the code: InHosp=No.

61. Click **Assign** and create the code: Num=Pop.

62. Click **Write/Export**.

63. In the Variables field, select **InZip, TheZip, Num**, and **InHosp**.

64. Select the Output Mode **APPEND**.

65. From the Data Table drop-down, select **ZipSums2**.

66. Click **OK**.

67. Click **Read/Import**.

68. Select the table **ZipSums2**.

69. Click **List** to see the values. The table should look like the following:

inzip	thezip	Num	InHosp
No	12202	369	Yes
Yes	12202	48	Yes
No	12202	349987	No
Yes	12202	10021	No

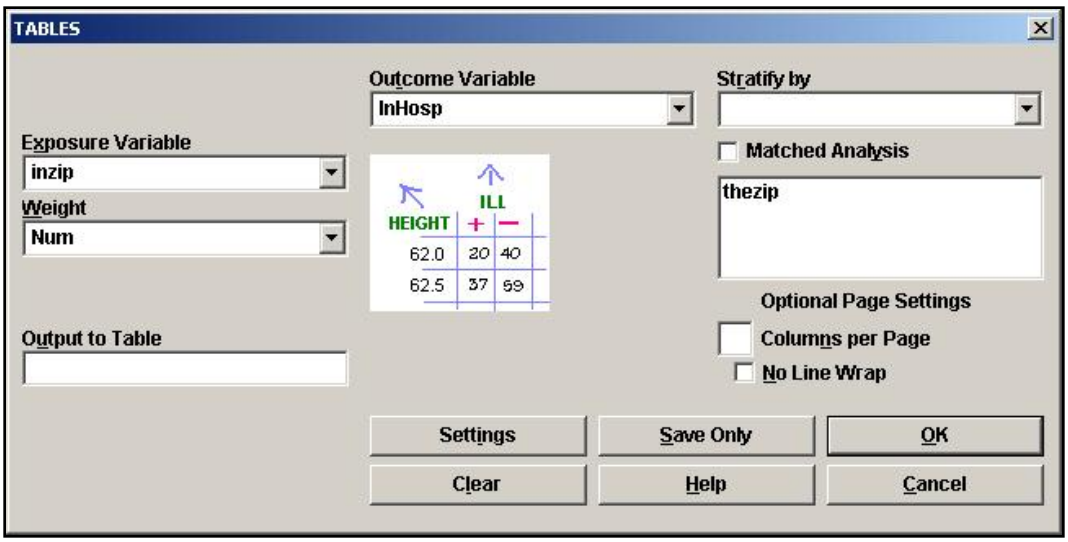
70. From the Command Tree Statistics folder, click **Tables**. The TABLE dialog box opens.

71. From the Exposure Variable drop-down, select **InZip**.

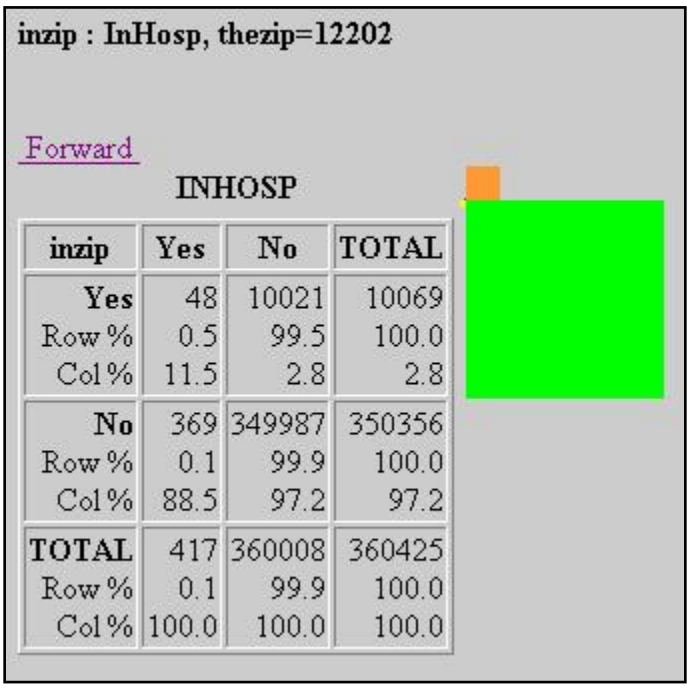
72. From the Outcome Variable drop-down, select **InHosp**.

73. From the Weight drop-down, select **Num**.

74. From the Stratify By drop-down, select **TheZip**.



75. Click **OK**.



inzip : InHosp, thezip=12202

Forward

INHOSP

inzip	Yes	No	TOTAL
Yes	48	10021	10069
Row %	0.5	99.5	100.0
Col %	11.5	2.8	2.8
No	369	349987	350356
Row %	0.1	99.9	100.0
Col %	88.5	97.2	97.2
TOTAL	417	360008	360425
Row %	0.1	99.9	100.0
Col %	100.0	100.0	100.0

- The Risk Ratios appear as follows.

Single Table Analysis

	Point Estimate	95% Confidence Interval Lower	Upper	
PARAMETERS: Odds-based				
Odds Ratio (cross product)	4.5431	3.3609	6.1411	(T)
Odds Ratio (MLE)	4.5431	3.3289	6.0939	(M)
		3.2882	6.1537	(F)
PARAMETERS: Risk-based				
Risk Ratio (RR)	4.5262			
Risk Difference (RD%)	0.3714			

(T=Taylor series; C=Cornfield; M=Mid-P; F=Fisher Exact)

STATISTICAL TESTS	Chi-square	1-tailed p	2-tailed p
Chi square - uncorrected	116.8211		0.0000000000
Chi square - Mantel-Haenszel	116.8208		0.0000000000
Chi square - corrected (Yates)	113.6295		0.0000000000
Mid-p exact		0.0000000000	
Fisher exact		0.0000000000	

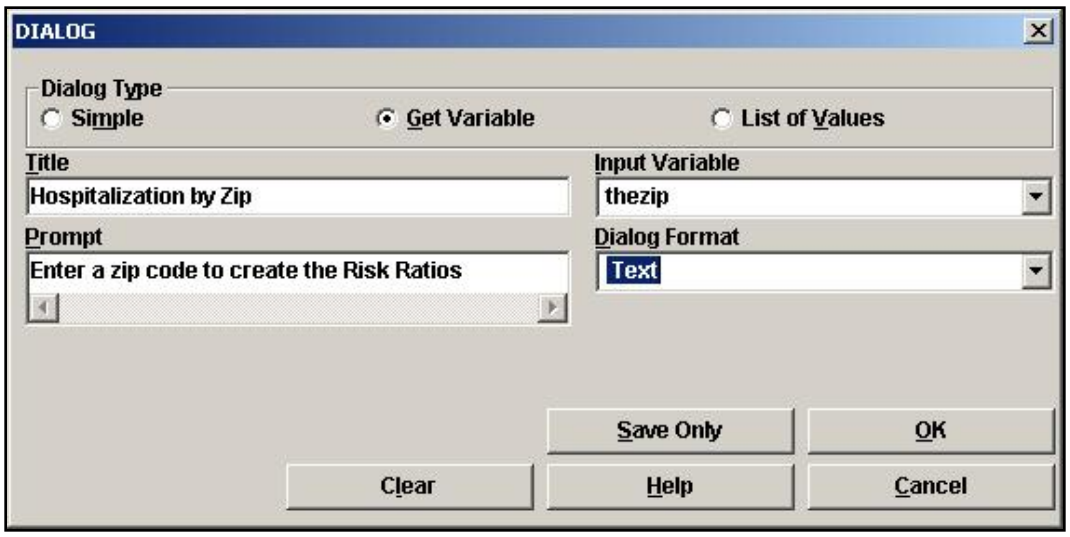
Four times more likely to be hospitalized for asthma if you are in this zip code.

- The PGM you created can be reused to analyze the risk ratio for each zip code in the table. You can also create a dialog box within the .PGM that will prompt you for a zip code each time you run the program.
- Do NOT close the current program.

Using the Dialog Command

Time to create some dialog.

1. From the Program Editor, click **Save**. The Save Program dialog box opens.
2. In the Program field, type **ZipRatios**.
3. Click **OK**.
4. Scroll up the .PGM to the code where you assigned thezip="12202".
5. Place the cursor in front of the Assign command.
6. Type an asterisk *****.
    - This is the command to comment out a specific line of code. You want to keep the code, but you do not want it to run automatically each time you run the program.
7. Place the cursor next to the text TheZip. Add a space.
8. Type the word **global** to change the variable type from standard to global.
9. Click **Run This Command** from the Program Editor to run the new DEFINE code.
10. From the Command Tree User Interaction folder, click **Dialog**. The DIALOG box opens.
11. From the Dialog Type section, select the **Get Variable** radio button.
12. In the Title field, type **Hospitalization by Zip**.
13. In the Prompt field, type **Enter a zip code to create the Risk Ratios**.
14. From the Input Variable drop-down, select **TheZip**.
15. From the Dialog Format drop-down, select **Text**.



16. Click **OK**. When the dialog test box opens, click **Cancel**. The code appears at the end of the program.

17. Highlight and select the code **DIALOG "Enter a zip code to create the Risk Ratios" thezip TEXTINPUT TITLETEXT="Hospitalization by Zip"**

18. Select **Edit>Cut**.

19. Paste the code under the ***Assign code**.

- Your code should look like the following example.

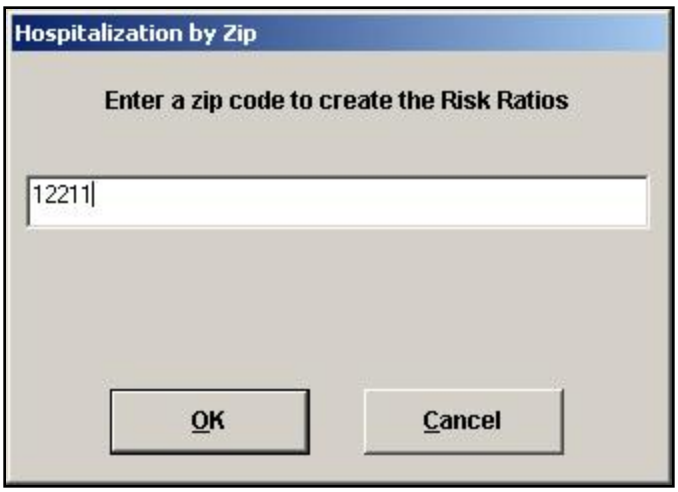
```
define thezip global
*assign thezip="12202"
DIALOG "Enter a zip code to create the Risk Ratios" thezip TEXTINPUT
TITLETEXT="Hospitalization by Zip"
```

20. **Save** the Program.

21. Click **Run** in the Program Editor. The code runs in the Program Editor.

22. In the Hospitalization by Zip dialog box, type **12211**.

23. Click **OK**. Results appear in the Output window.



- Notice that the new results were also written to the RouteOut file RiskRatio.HTM.
- You can use the RiskRatio.XML to create reports in Epi Report.

StatCalc — Estimate Sample Size

You want hospitalization rates to be included in your proposal. Determine how many records you need to sample to get the most accurate statistics.

N. From the New York hospital discharge data, how many records do you need to sample in order to get a 95% confidence interval from a sample size of 38,998?

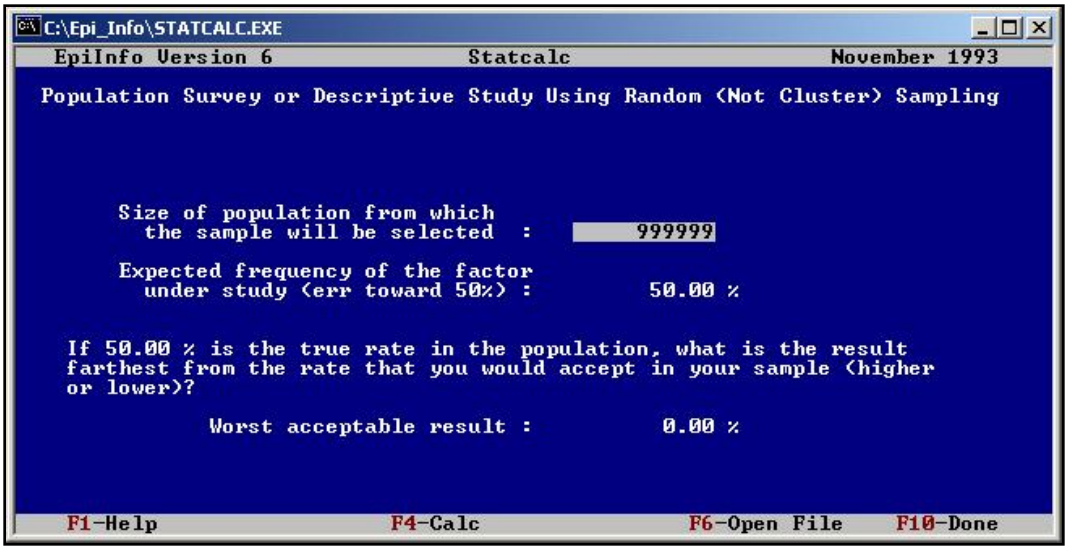
StatCalc is an epidemiologic calculator that produces statistics from summary data. You can use it to estimate how large a sample needs to be to achieve a 95% confidence interval. Population survey samples to be taken are simple random or representative. The question being asked must have a yes/no, or other two-choice answer leading to a proportion of the population as the final result.

StatCalc is a DOS-based utility. Use the keyboard rather than the mouse for navigation.

Let's go the sample route.

1. From the Epi Info™ main page, select **Utilities>Statcalc**. StatCalc opens.

2. Use the arrow keys on your keyboard to select **Sample Size** and **Power**.
3. Press **Enter**.
4. From the pop-up, select **Population Survey**.
5. Press **Enter**.



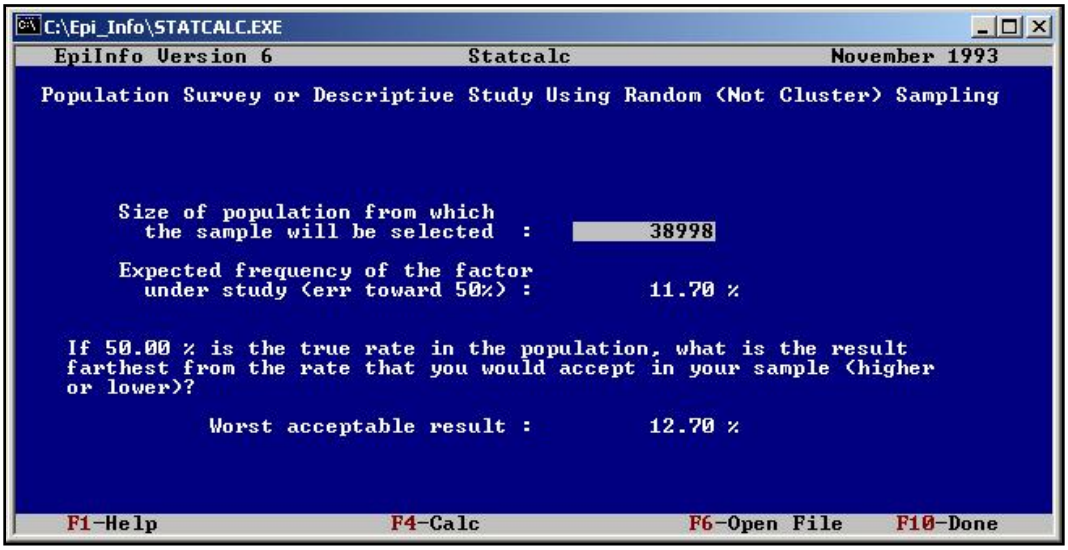
6. In the Size of Population field, type the number of NY discharges for Asthma, **38998**. This number came from the summarized hospital discharge records.

7. Press **Enter**.

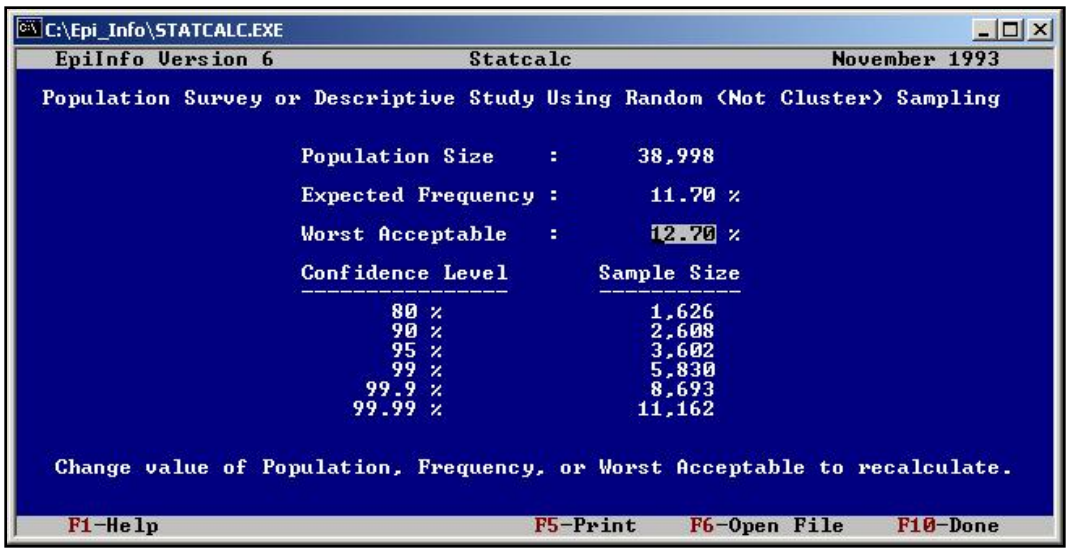
8. In the Expected Frequency field, type **11.7**. This is the lifetime prevalence rate for NY from the BRFSS data.

9. Press **Enter**.

10. In the Worst Acceptable field, type **12.7**.



11. Press the **F4** key to calculate.



12. Look under the Confidence Level row and across to the Sample Size row to get the size of the sample for a 95% confidence interval.

- Recalculate for any sample size by entering new data and pressing **F4**.

13. Exit StatCalc by pressing **F10** until the window closes.

Lesson Complete!

WHAT YOU LEARNED

How to:

- Use the RELATE command to link a variety of data tables.
- Create a rate in Epi Info™ using the ASSIGN command.
- Create a graph or a map to illustrate rates.
- Use LIST UPDATE to edit records.
- Use WRITE/APPEND to create new data tables.

This page intentionally blank.

Appendix A — Code Sheet

Prompt/Question	Field Name	Variable Type	Font for Prompt	Pattern, Table, or Option Selection
Parent School Asthma Pre-Intervention Survey	PreSurvey	Label/Title	14, Bold	
Please answer all the questions below, whether or not your child has asthma.	Instruction	Label/Title	12	
School Number	SchoolNum	Text		Code Table link to PSNumber: A to 123 B to 234 C to 345 D to 456
Student ID Number	StudentID	Text		Required
Zip Code	Zip	Text		
Child's Last Name	ChildLast	Text		
Child's First Name	ChildFirst	Text		
Male/Female	Gender	Text		Legal value: Male Female
Child's Date of Birth	DOB	Date		Required MM/DD/YYYY
Age	Age	Number		Read Only ##
Your Last Name	YourLast	Text		
Your First Name	YourFirst	Text		
Your relationship to Child	Relation	Text		

Home Phone Number	Phone	PhoneNumber		###-###-####
PS School Number	PSNumber	Text		Read Only Code Table link to SchoolNum
Grade	Grade	Number		##
Today's Date	TDate	Date		MM/DD/YYYY
Has a health professional ever told you that your child has any of the following conditions?	Condition	Label/Title		
Asthma	Asthma	Checkbox		
Reactive Airway Disease	RAD	Checkbox		
Asthmatic Bronchitis or Wheezy Bronchitis	Bronchitis	Checkbox		
Wheezing	Wheezing	Checkbox		
Has a health professional prescribed any medication to your child?	Medication	Yes/No		
How many times did your child have an emergency visit for breathing problems?	Emergency	Number		##
How many times did your child have to stay overnight in the hospital for breathing problems?	Overnight	Number		##
Does your child take medicine every day for breathing difficulties?	DailyMed	Yes/No		
Select the month when your child's breathing was the worst.	WMonth	Text		Comment Legal 1- January 2- February 3- March 4- April 5- May 6- June

				7- July 8- August 9- September 10- October 11- November 12- December
During their worst breathing month, select how often the following symptoms occurred.	Symptoms	Label/Title		
Wheeze or whistling in the chest?	Whistle	Text		Legal Values are: Never 2 times each week or less 3 to 6 times each week Every day but not all day Every day and all day
Have a cough?	Cough	Text		Legal Values are: Never 2 times each week or less 3 to 6 times each week Every day but not all day Every day and all day
Have a tight chest or shortness of breath?	ShortB	Text		Legal values are: Never 2 times each week or less 3 to 6 times each week Every day but not all day Every day and all day

Wake up at night from wheezing, coughing, or trouble breathing?	WakeUp	Text		Legal Values are: Never 2 times each week or less 3 to 6 times each week Every day but not all day Every day and all day
How many days of school has your child missed because of asthma?	MissDays	Number		###
Have breathing problems stopped your child from participating in sports?	MissSport	Yes/No		
If yes, how many times in the past month?	SpTimes	Text		Legal Value 2 times a month or fewer 3 to 4 times a month 1 to 3 times each week 4 times each week or more
How many people living in your home, including this child, have asthma?	IllHome	Number		##

Appendix B — Parent School Asthma Pre-Intervention Survey

Please answer all the questions below, whether or not your child has asthma. If you receive more than one questionnaire, please complete a separate questionnaire for each of your children who attend this school. Your answers will help the school nurse provide information and medical assistance to children and their families who have asthma.			
School Number	Student ID Number	Zip Code	
Child's Name (Last, First)	Male/Female (Circle One)	Child's Date of Birth	Age
Your Name (Last, First)	Your Relationship to Child	Home Phone Number	
PS School Number	Grade	Today's Date	

School Number	**Student ID Number**

1. Has a doctor, nurse or other health professional ever told you that your child has any of the following conditions?

   (Please check "Yes" or "No".)

   a. Asthma ☐ Yes ☐ No
   b. Reactive Airway Disease ☐ Yes ☐ No
   c. Asthmatic Bronchitis or Wheezy Bronchitis ☐ Yes ☐ No
   d. Wheezing ☐ Yes ☐ No

2. During the past school year (September 1, 2002 – June 30, 2003), has a doctor, nurse or other health professional prescribed any medicine for asthma, wheezing, cough, bronchitis, or other breathing problems? *(This could include pills, syrups, inhalers or breathing machines.)* ☐ Yes ☐ No

3. During the past school year (September 1, 2002 – June 30, 2003), how many times did your child have an emergency visit to a doctor, clinic or an emergency room for asthma, wheezing, cough, bronchitis, or other breathing problems? (*Enter the number. If none, enter "0" or "None".)*
   _____
   (Number of times)

4. During the past school year (September 1, 2002 – June 30, 2003), how many times did your child have to stay overnight in the hospital (not the emergency room) for asthma, wheezing, cough, bronchitis, or other breathing problems? (*Enter the number. If none, enter "0" or "None".)*
   _____
   (Number of times)

5. Does your child take any medication every day for asthma or other breathing difficulties, even when he or she is feeling well? ☐ Yes ☐ No

6. Think about your child's breathing difficulties during the past school year (September 1, 2002 – June 30, 2003). Please check the month that you feel your child's breathing was the WORST.

   ☐ Sept. ☐ Oct. ☐ Nov. ☐ Dec. ☐ Jan. ☐ Feb. ☐ Mar. ☐ April ☐ May ☐ June

   ☐ Not applicable, my child did not have breathing difficulties during the past school year (September 1, 2002 – June 30, 2003).

7. When thinking about your child's breathing difficulties during the month you checked above, Mark an "X" in the box that best describes how often each of the following symptoms happened:

Wheeze or whistling in the chest?	☐	Never	☐	2 times each week or less	☐	3 to 6 times each week	☐	Every day but not all day	☐ Every day and all day
have a cough?	☐	Never	☐	2 times each week or less	☐	3 to 6 times each week	☐	Every day but not all day	☐ Every day and all day
Have a tight chest or shortness of breath?	☐	Never	☐	2 times each week or less	☐	3 to 6 times each week	☐	Every day but not all day	☐ Every day and all day
Wake up at night from wheezing, coughing or trouble breathing?	☐	Never	☐	2 times each week or less	☐	3 to 6 times each week	☐	Every day but not all day	☐ Every day and all day

8. During the past school year (September 1, 2002 – June 30, 2003), how many days of school has your child missed due to asthma or other breathing difficulties?
(*Enter the number. If none, enter "0" or "None".)*

_____
(Number of days)

9. During the past month, has asthma or other breathing problems ever stopped your child from participating in sports or other physical activities? ☐ Yes ☐ No
If "YES", how many times in the past month? *(Please check only one box.)*

☐ 2 times a month or fewer ☐ 3 to 4 times a month ☐ 1 to 3 times each week ☐ 4 times each week or more

10. How many people living in your home, including this child, have asthma?
(Enter the number. If none, enter "0" or "None".)

_____
(Number of people)

Reprinted with the permission of the New York City Department of Health and Mental Hygiene

Appendix C — Data Entry Surveys

Please answer all the questions below, whether or not your child has asthma. If you receive more than one questionnaire, please complete a separate questionnaire for each of your children who attend this school. Your answers will help the school nurse provide information and medical assistance to children and their families who have asthma.

School Number A	Student ID Number 11	Zip Code 12208	
Child's Name (Last, First) Evans, Katie	Gender Female	Child's Date of Birth 11/30/1994	Age 10
Your Name (Last, First) Evans, Rebecca	Relationship to Child Mother	Home Phone Number 555-555-5555	
PS School Number 123	Grade 5th	Today's Date 10/25/2005	

School Number	**Student ID Number**

1. Has a doctor, nurse or other health professional ever told you that your child has any of the following conditions?

   (Please check "Yes" or "No".)
   - a. Asthma ☒ Yes ☐ No
   - b. Reactive Airway Disease ☒ Yes ☐ No
   - c. Asthmatic Bronchitis or Wheezy Bronchitis ☒ Yes ☐ No
   - d. Wheezing ☒ Yes ☐ No

2. During the past school year (September 1, 2002–June 30, 2003), has a doctor, nurse or other health professional prescribed any medicine for asthma, wheezing, cough, bronchitis, or other breathing problems? *(This could include pills, syrups, inhalers or breathing machines.)* ☒ Yes ☐ No

3. During the past school year (September 1, 2002–June 30, 2003), how many times did your child have an emergency visit to a doctor, clinic or an emergency room for asthma, wheezing, cough, bronchitis, or other breathing problems? (*Enter the number. If none, enter "0" or "None."*) _____**2**_____ (Number of times)

4. During the past school year (September 1, 2002–June 30, 2003), how many times did your child have to stay overnight in the hospital (not the emergency room) for asthma, wheezing, cough, bronchitis, or other breathing problems? (*Enter the number. If none, enter "0" or "None."*) ____**1**_____ (Number of times)

5. Does your child take any medication every day for asthma or other breathing difficulties, even when he or she is feeling well? ☒ Yes ☐ No

6. Think about your child's breathing difficulties over the past school year (September 1, 2002–June 30, 2003). Please check the month that you feel your child's breathing was the WORST.

   ☒ Sept. ☐ Oct. ☐ Nov. ☐ Dec. ☐ Jan. ☐ Feb. ☐ Mar. ☐ April ☐ May ☐ June

   ☐ Not applicable, my child did not have breathing difficulties during the past school year (September 1, 2002–June 30, 2003).

7. When thinking about your child's breathing difficulties during the month you checked above, Mark an "X" in the box that best describes how often each of the following symptoms happened:

Wheeze or whistling in the chest?	☐	Never	☒	2 times each week or less	☐	3 to 6 times each week	☐	Every day but not all day	☐ Every day and all day
Have a cough?	☐	Never	☒	2 times each week or less	☐	3 to 6 times each week	☐	Every day but not all day	☐ Every day and all day
Have a tight chest or shortness of breath?	☐	Never	☒	2 times each week or less	☐	3 to 6 times each week	☐	Every day but not all day	☐ Every day and all day
Wake up at night from wheezing, coughing or trouble breathing?	☐	Never	☒	2 times each week or less	☐	3 to 6 times each week	☐	Every day but not all day	☐ Every day and all day

8. During the past school year (September 1, 2002–June 30, 2003), how many days of school has your child missed because of asthma or other breathing difficulties?
(*Enter the number. If none, enter "0" or "None."*)

_____**12**_____
(Number of days)

9. During the past month, has asthma or other breathing problems ever stopped your child from participating in sports or other physical activities? ☒ Yes ☐ No
If "YES", how many times in the past month? *(Please check only one box.)*

☐ 2 times a month or fewer ☒ 3 to 4 times a month ☐ 1 to 3 times each week ☐ 4 times each week or more

10. How many people living in your home, including this child, have asthma?
(Enter the number. If none, enter "0" or "None.")

_____**1**_____
(Number of people)

Reprinted with the permission of the New York City Department of Health and Mental Hygiene

Please answer all the questions below whether or not your child has asthma. If you receive more than one questionnaire, please complete a separate questionnaire for each of your children who attend this school. Your answers will help the school nurse provide information and medical assistance to children and their families who have asthma.

School Number A	Student ID Number 12	Zip Code 12203	
Child's Name (Last, First) Williams, Beth	Gender Female	Child's Date of Birth 11/08/1995	Age 9
Your Name (Last, First) Williams, Susan	Relationship to Child Mother	Home Phone Number 333-333-3333	
PS School Number 123	Grade 5th	Today's Date 10/25/2005	

School Number	**Student ID Number**

1. Has a doctor, nurse or other health professional ever told you that your child has any of the following conditions?

   (Please check "Yes" or "No.")

   a. Asthma ☒Yes ☐No
   b. Reactive Airway Disease ☒Yes ☐No
   c. Asthmatic Bronchitis or Wheezy Bronchitis ☐ Yes ☐No
   d. Wheezing ☒Yes ☐No

2. During the past school year (September 1, 2002–June 30, 2003), has a doctor, nurse or other health professional prescribed any medicine for asthma, wheezing, cough, bronchitis, or other breathing problems? *(This could include pills, syrups, inhalers or breathing machines).* ☒ Yes ☐No

3. During the past school year (September 1, 2002 – June 30, 2003), how many times did your child have an emergency visit to a doctor, clinic or an emergency room for asthma, wheezing, cough, bronchitis, or other breathing problems? (*Enter the number. If none, enter "0" or "None."*) _____**0**_____ (Number of times)

4. During the past school year (September 1, 2002 – June 30, 2003), how many times did your child have to stay overnight in the hospital (not the emergency room) for asthma, wheezing, cough, bronchitis, or other breathing problems? (*Enter the number. If none, enter "0" or "None."*) _____**0**_____ (Number of times)

5. Does your child take any medication every day for asthma or other breathing difficulties, even when he or she is feeling well? ☐ Yes ☒No

6. Think about your child's breathing difficulties during the past school year (September 1, 2002 – June 30, 2003). Please check the month that you feel your child's breathing was the WORST.

☐ Sept. ☐ Oct. ☐ Nov. ☐ Dec. ☐ Jan. ☐ Feb. ☐ Mar. ☐ April ☐ May ☒ June

☐*Not applicable, my child did not have breathing difficulties in the past school year (September 1, 2002–June 30, 2003).*

7. When thinking about your child's breathing difficulties during the month you checked above, Mark an "X" in the box that best describes how often each of the following symptoms happened:

Wheeze or whistling in the chest?	☐ Never	☐ 2 times each week or less	☒ 3 to 6 times each week	☐ Every day but not all day	☐ Every day and all day
Have a cough?	☐ Never	☒ 2 times each week or less	☐ 3 to 6 times each week	☐ Every day but not all day	☐ Every day and all day

Have a tight chest or shortness of breath?	☒	Never	☐	2 times each week or less	☐	3 to 6 times each week	☐	Every day but not all day	☐ Every day and all day
Wake up at night from wheezing, coughing, or trouble breathing?	☐	Never	☒	2 times each week or less	☐	3 to 6 times each week	☐	Every day but not all day	☐ Every day and all day

8. During the past school year (September 1, 2002 – June 30, 2003), how many days of school has your child missed due to asthma or other breathing difficulties?
(*Enter the number. If none, enter "0" or "None."*) _____**14**_____ (Number of days)

9. During the past month, has asthma or other breathing problems ever stopped your child from participating in sports or other physical activities? ☐ Yes ☒ No
If "YES", how many times in the past month? *(Please check only one box.)*

☐ 2 times a month or fewer ☐ 3 to 4 times a month ☐ 1 to 3 times each week ☐ 4 times each week or more

10. How many people living in your home, including this child, have asthma?
(Enter the number. If none, enter "0" or "None.") _____**1**_____ (Number of people)

Reprinted with the permission of the New York City Department of Health and Mental Hygiene

Appendix D — Skills Reviews Answer Key

Lesson 2 Check Code	
Check Code can run before or after values are entered in a variable.	True From the Program Editor, you assign when Check Code will run by using the Before or After radio buttons in the Check Commands window.
Using the Program Editor, you can create Check Code that will run across pages or records.	True From the Program Editor, Check Code can be added to pages and records using the Choose Field Where Action Will Occur drop-down menu.
In the Check Command window, variables tagged with an X have Check Code associated with them.	False Variables with Check Code are tagged with an *asterisk.
Which function is used to create an age variable.	YEARS The YEARS function returns the number of years from variable 1 to variable 2 in a numeric format. For example, Age=YEARS(DOB, Today's Date)
How does the answer yes appear in the Program Editor?	(+) The symbol (+) is equal to yes or true values. The symbol (-) is equal to no or false values.
Lesson 3 Enter Data	
Once you create a _____, you cannot change field names or types in your view.	Data Table Once you begin entering records into the data table, you cannot change field names or types without deleting data that has been entered.

What are the two ways you can enter a new record?	Clicking New or tabbing through to the last record.
If you are entering data and have a required field, you must enter information into that field before proceeding to the next page.	True Enter will not allow you to leave a page until all required fields are completed.
What expression could you enter in the Find feature to locate children who are greater than 6 in age?	Age>6
From the Find feature, what button do you click to clear the page and begin a new search?	Reset The Reset button will clear all search options without leaving the Find feature window.
Lesson 4 Analysis Basics	
After you run a command, the code appears in which window?	Program Editor Codes appear in the Program Editor and can be saved as programs.
What command allows you to open a project for analysis?	READ You must READ a project into Analysis before creating any statistics or data management.
What command would you use to see variables and records in the current data table?	LIST The LIST command creates a listing of the current data table and records.
What command would you use to specify a sequence for records to appear?	SORT The SORT command organizes listed data in and ascending or descending order based on selected variables.
What are the two display modes you can use with the LIST command?	Web (HTML) and Grid The LIST command allows you to view information embedded in the Output window as HTML or in a grid table format. There is a third Display Mode called List Allow Updates, which allows you to edit records. This is covered in the Intermediate

	Analysis lesson.
Lesson 6 Epi Graph	
You can use a Line Graph to determine trends and cyclic variation.	True Line Graphs can display data that varies along an accepted sequence such as time, temperature, etc.
Which type of graph presents a circular proportional assessment by comparing data elements (e.g.,percentages or counts against the sum of the data elements?)	Pie Graph Pie Graphs are used for proportional assessment by comparing data elements (e.g., percentage or count against the sum of data elements).
From Epi Graph, select the two methods that can open the Customization dialog box.	Right click and select Customization, or select View>Customization.
Graphs can be exported for use in other programs (e.g.,MS Word, PowerPoint, or Epi Report).	True Select File>Export from Epi Graph and choose your export options.
You can plot a new graph type from an existing graph using the Customization dialog box.	True Open the Customization dialog box and select the Plot tab. Select a new plot style and click Apply.
Lesson 7 Analysis: Exporting Files	
Which command allows you to designate a new Output file?	ROUTEOUT The ROUTEOUT command directs output to a named file until the process is terminated by the CLOSEOUT command, or by exiting Analysis.
Which command allows you to create a new variable?	DEFINE The DEFINE command allows you to create new variables. Standard is the most common type.
Which command allows you to store the value of a variable or assigns the result of a	ASSIGN This command assigns the value of an

mathematical expression.	expression to a variable. It is commonly used after a new variable is created with the DEFINE command.
When changing, assigning, or selecting text fields, what must the value must be enclosed in?	"Quotes" Quotes are used to enclose text strings.
In the Program Editor, the value of missing appears how?	(.) This symbol denotes a missing value in your data table.
To change a number variable to a date, which function would you use?	NUMTODATE This function transforms three numbers into a date format.
Which command allows you to create a new data table?	WRITE WRITE will send records to an output table, or file in the format that you specify. You can specify what variables will be written, the order in which they will appear, and the type of file to be written.
Which command allows you to open a new data table?	READ The READ command makes a view active. It also removes any previously active data tables, associated defined variables, and data table-specific commands (e.g., RELATE, SORT or SELECT).
From the WRITE dialog box, what output mode would you select to add data to an existing table?	APPEND Records will be added to the existing file or data table if you select the APPEND option.
The WRITE command can be used to export a data table to MS Excel.	True The WRITE command will send records to an output table or file in a specified format. You can specify what variables will be written, the order in which they will appear, and the type of file to be written.
Lesson 10 Epi Report	

Which Analysis Output files are accessible by Epi Report?	Output.XML Analysis creates two types of files, HTML and XML. The XML file can be used in Epi Report.
Which button should you click to view a report with complete data included?	Generate Reports Click the Generate Report button, or select File>Generate Report to view a report with all statistics computed.
Aggregate data can be added to reports from the Read Data and Create menu.	True From the Read Data and Create menu, you can also Read Analysis Output, insert a Line Listing Group, or create a Pivot Table.
What dialog box can customize Line Listing groups?	Query Builder The Query Builder allows you to customize the data that appears in a line listing.
What element is used to insert a graphic into a report?	Image Open the Insert Report Object menu in the Report Elements Tree and drag the word Image onto the report.

This page intentionally blank.

Appendix E — Lesson 5 Answer Key

Frequency Sections	**Answers**
A. Are there more males or females in the survey population?	Females, 406
B. Which condition has the highest frequency?	Wheezing, 12.4%
C. Which month had the highest number of students with breathing difficulties?	9-September
D. Which school had the highest frequency of students with bronchitis?	School A- 13.8%
E. Which gender has the highest frequency of asthma?	Males, 12.2%
F. How many students answered yes to Reactive Airway Disease (RAD)?	93
G. What is the percentage of students with RAD that have been prescribed medication?	49.5%
H. Which zip code has the highest number of students living in it?	12046
I. Of that zip code, which condition had the highest frequency?	Asthma
J. How many students answered yes to all four conditions: asthma, RAD, bronchitis, and wheezing?	16
K. Based on all four conditions and zip code, are the conditions evenly distributed among the survey population?	Yes
L. Does one school have more students in the affected zip codes and with all four conditions than the others do?	Yes, School A
Means Sections	**Answers**
M. What was the most frequent number of missed days?	1
N. What is the mean (average) number of missed days because of asthmatic conditions?	3.76
O. Which school had the highest average of missed days?	School A, 4.15
P. Was the school with the highest average of missed days for males the same as the school with the highest average of missed days for females?	No. School A for males (4.36) and School B for females (4.17)

Q. Is the average number of missed days higher among those who answered yes to asthma than the overall average? Refer to Question N.	Yes, 11.8 vs. 3.76
R. What is the average age of the student population?	8.66
Tables Section	**Answers**
S. Does one school have a higher number of students with asthma?	Yes, School A
T. Which school has the highest number of students with wheezing?	School A
U. How many students had both asthma and wheezing?	63
V. If you have asthma, what are the odds that you have the condition wheezing, if you have the condition asthma?	Odds Ratio 42% Risk Ratio 13.6%
W. Is the Odds Ratio for bronchitis and RAD lower or higher than asthma and wheezing? Refer to Question V.	Lower 7.25%
X. How many children with the condition wheezing were prescribed medication? From those results, what is the Risk Ratio?	Children, 44 Risk Ratio 9.2%
Y. Is the Risk Ratio for being prescribed medication for the condition asthma higher or lower than being prescribed medication for wheezing? Compare the results to Question X.	Higher. Risk Ratio 22.2%
Z. Are the risk ratios for asthma and wheezing higher for children less than 10 years of age than they are for the survey population as a whole? Compare the results to Question Y.	Higher. Odds 61.7%, Risk Ratio 15.9%

Appendix F — Intermediate Analysis Answer Key

The sample .PGMs listed are available in the project Asthma Final.MDB. The codes listed will show you an example of how the question could be run using Analysis. Some of the codes require intermediate steps to create results. Sample programs follow the steps listed in the lesson. For example, LISTS that require updates must be done manually as they are not saved as part of the program editor.

Computing Prevalence — Using BRFSS Data	BRFSS2003.PGM
A. What is the lifetime asthma prevalence among adults in the state?	11.7%
B. What is the current asthma prevalence among adults in the state?	7.6%
C. Is the state adult asthma prevalence rate higher in the subgroup gender? **D.** Is the state adult asthma prevalence rate higher in the subgroup ethnicity?	No Males - 10% Females - 13.2% Hispanics - 11.2%

Primary Data — 800 Records	Freq4Schools.PGM
E. From the school asthma data, what is the difference in prevalence rates?	School A - 14.8% School B - 12.2% School C - 10.1% School D - 8.4%

Census Data	CountyPrev.PGM
F. What is the county childhood asthma prevalence?	9.978%
G. What is the county prevalence for the subgroup gender?	Males - 9.396% Females - 10.599%

Mortality Data	MortalityRate.PGM
H. Is the asthma mortality rate for the state and county significantly different?	State 2000 - .179 State 2001 - .181 State 2002 - .174 County 2000 - .136 County 2001 - .068 County 2002 - .136

10-Year Trends — Hospital Discharge Data	10YRgraph.PGM
I. Is the rate of hospital discharge increasing or decreasing in any age groups?	It appears that the discharge rate is cyclical with 1996, 1999, and 2003 being up years in all age groups.

Three-Year Rates — Hospital Discharge Data	3YRgraph.PGM
J. Are there significant differences in the 3-year discharge rates for the state versus the county by age?	Rates in all the age groups appear to increase in 2003. In New York, the 0–17-year-old age group has higher rates, while in Albany, the 65+ age group has the highest rates.

Mapping Rates	MapRates.PGM
K. Based on the map with hospitalization rates by zip code, which areas have the highest rates for asthma?	12202, 12207, 12206, 12209, 12041, 12007

Scatter Plots	ScatterPlot.PGM
L. Does income appear to affect the rate of hospitalization?	Yes. The number of discharges appears much higher in the lower income brackets.

Risk Ratios — Hospitalization Data	RiskRatioZip.PGM
M. What is the risk ratio of being hospitalized for asthma if you live in the zip code 12202 versus if you live in the zip code 12211?	12202 - Risk Ratio 4.5 12211 - Risk Ratio .39 Persons living in 12202 are four times as likely to have asthma compared with those in 12211.

StatCalc	
N. From the New York hospital discharge data, how many sample records do you need to get a 95% confidence interval from a sample size of 38998?	3,602

This page intentionally blank.

Appendix G — Preparing Data for Use in the Epi Info™ Tutorial

The preparation and data management phase is critical for any project involving multiple data sources. Data management in Epi Info™ requires forethought into the types of results needed. The steps used in this tutorial for accessing and formatting data are explained briefly below. This appendix provides an estimate of the preparation time and scope of data management used in this project, which you can apply to different projects in the future.

PREPARATION QUESTIONS

1. Once it was decided to use asthma data, how did you decide what data was needed?
2. Where did you find the data?
3. What format was the data in?
4. Did you reformat any of the data? If yes, what did you need to do, and how long did it take?
5. Where did you get shape files and zip code level maps used in the lessons?
6. How did you format the data to use it for mapping?

ANSWERS

1. **Once it was decided to use asthma data, how did you decide what data was needed?**

   The first step was to understand the context of why asthma information was needed. First, the scenario would involve applying for an asthma grant. Second, questions were developed and data sources located. The initial questions are included below.

   - The asthma prevalence among adults in the state is ___, and is increasing/decreasing. (BRFSS)
   - The prevalence rate of adult asthma is higher in certain subgroups (age, gender, race/ethnicity. (BRFSS)
   - The county adult asthma prevalence rate is higher/lower than the state. (BRFSS/EBRFSS)
   - The county adult asthma prevalence is higher in some population subgroups (age, gender, race/ethnicity). (BRFSS/EBRFSS)

- How do the county population sub-group patterns relate to the state sub-groups for adult asthma prevalence? (BRFSS/EBRFSS)
- The current asthma prevalence among children in the state in ___. (National Asthma surveillance – NY)
- Is the state childhood asthma prevalence higher in any population subgroup? (National Asthma surveillance – NY)
- The county childhood asthma prevalence is __ (primary data collection)
- How is the county childhood asthma prevalence different from the state?
- Is the asthma mortality rate for the state and county significantly different? (Vital Records – 3 years)
- The rate of hospital discharge from asthma is increasing in __ age group, and decreasing in __ age group. (10-year trends, 3 years for map; age group- total, 0–17, 18–64, 65+)
- Compare the three-year rates for state vs. county by age. Which are significantly different?
- Hospitalization rate by zip code for 3 years. Where are the high risk areas?
- Look at median family income/ per capita income by zip code.
- Look at emergency room data – one-year cross-sectional. What percentage is asthma-related? Look at age, gender, race/ethnicity, payment source distributions.
- The risk ratio for someone who lives in the __ zip code being hospitalized/seen in the emergency room for asthma is___ than the other zip codes in the county.
- Estimation of sample size for sampling school asthma survey.
- From school asthma data of four schools (two high-risk areas, one moderate risk and one low risk) with 800 records totally, what is the difference in prevalence rates?

2. Where did you find the data?

Data	Source
Hospital Discharge Data – 10 years for trends, 3 years for maps, payment source, age, race/ethnicity, gender (aggregate data).	NYSDOH SPARCS database https://commerce.health.state.ny.us/hin/dataque/dataquer.html

Hospital asthma discharge numbers by zip code (data available as an MS Excel file)	NYSDOH SPARCS database https://commerce.health.state.ny.us/hin/dataque/dataquer.html
Mortality data – 3 years (aggregate data)	Bureau of Biometrics, NYSDOH
BRFSS – 4 years (aggregate data)	BRFSS Program, NYSDOH https://commerce.health.state.ny.us/hpn/brfss/default.htm
E-BRFSS – 1 year (excel file)	BRFSS Program, NYSDOH https://commerce.health.state.ny.us/hpn/brfss/default.htm
Parent school asthma survey	Sample questionnaire used with permission from the New York City Department of Health and Mental Hygiene
Parent School Asthma Survey (four schools/200 records each)	hypothetical data was created
Census – zip code level, county boundary and shape files, median family and per capital income, racial/ethnic distribution, hospital/emergency rooms, schools – (aggregate data and map files).	http://www.census.gov/geo/www/tiger/index.html
National asthma surveillance childhood asthma prevalence (aggregate data)	Asthma Program, Public Health Information Group, NYSDOH

3. In what format was the data? HTML, DBF, .XLS, etc.

Data used in this tutorial were in HTML, MDB, XLS, DBF, and TXT formats.

The process for merging an .MDB and an .XLS file is explained in Lesson 5. Analysis can READ in 24 different file formats. You can READ in an .XLS file and then WRITE it to a table inside Analysis for use with the Analysis features. The process of reading and writing .HTML and .DBF files is covered in the Intermediate Analysis Lesson of this

tutorial. The steps outlined in Intermediate Analysis Lesson for reading in and writing out an .HTML file can be applied to any format imported into Epi Info™.

4. Did you reformat any of the data? If yes, what did you need to do, and how long did it take?

Yes. The data management steps involved in creating the lessons varied in time and complexity of steps. Here is an example of the kinds of behind-the-scenes steps.

If you READ the Numerators table in Analysis, you will see a set of values listed under the variable Series>Asthma Deaths. The plan was to place the mortality numbers into the Numerator table to compute rates.

The mortality data came from an .XLS file. The mortality spreadsheet contained a list of counties and three columns with mortality numbers for 3 years in those columns. The .XLS file was opened in Analysis and written out to a new data table so it could be manipulated.

Once the new table was read into Analysis, a set of matching variables to the Numerator table were defined. For example, SERIES was DEFINED and then ASSIGNED the value Asthma Deaths. Each year was SELECTed and to get the number of New York deaths, the SUMMARIZE command was used. This new information could then be WRITE APPENDed to the Numerator table.

Completing the steps for this section of the final table took about an hour.

5. Where did you get shape files and zip code level maps used in the lessons?

The US Census Bureau's Cartographic Boundary Files library (http://www.census.gov/geo/www/cob/bdy_files.html) contains five-digit ZIP Code Tabulation Areas (ZCTAs): 2,000 files for download. The site contains files for both ESRI and MapInfo for every state. The MapInfo boundary file data set was selected because of the accessibility of MapInfo software by the staff. County specific zip codes were carved out by creating a query that selected all zip codes from the ESRI zip code shape file map that were within or partially within a MapInfo county shape file map. This county shape file with zip codes was exported into an ESRI shape file that can be used by Epi Info™.

When downloading files for use in Epi Map, select the .SHP download versions. To select specific zip code information from a downloaded state file, use the Add Partial Layer feature in the Map Manager, then save the new shape. Shape files are also available from the Epi Info™ website http://www.cdc.gov/epiinfo.

6. How did you format the data so it could be used for mapping?

Review Lessons 8 and 9 for all the mapping rules.

Data management was involved in preparing files for mapping. To map data, it needs to be in an .MDB format and the zip code field needs to be text. The original .MDB

prepared for the tutorial had to reworked for use in Epi Map. Once the .MDB was READ into Analysis, the DISPLAY command was used to ensure the zip code field was text. Since it was not, a new zip code variable was DEFINEd, ASSIGNed, and the FORMAT command was used to make it appear as text. Two variable titles also contained underscores, which were removed with DEFINE and ASSIGN. The new variables were then WRITE REPLACEd to a new data table for use in mapping.

This page intentionally blank.

Appendix H — Glossary

Terms listed in the glossary are located in the Epi Info™ online help system and as part of the online glossaries listed in Appendix I: Data Sources and the References page.

Analytic Epidemiology

The aspect of epidemiology concerned with the search for health-related causes and effects. Uses comparison groups, which provide baseline data, to quantify the association between exposures and outcomes, and test hypotheses about causal relationships.

Chi Square

A test of statistical significance used to determine how likely an observed association between an exposure and a disease could have occurred because of chance alone if the exposure was unrelated to the disease. The Chi-Square Test is the test of choice when the expected values for each cell in a two-by-two table are at least five.

Confidence Interval (CI)

A range of values for a variable that indicates the likely location of the true value of a measure.

Denominator

The lower portion of a fraction used to calculate a rate or ratio. In a rate, the denominator is usually the population or population experience (i.e., person-years) at risk.

Descriptive Epidemiology

The aspect of epidemiology concerned with organizing and summarizing health-related data according to time, place, and person.

Distribution

In epidemiology, the frequency and pattern of health-related characteristics and events in a population. In statistics, the observed or theoretical frequency of values of a variable.

Epidemiology

The study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to the control of health problems.

Expressions

Combinations of literal values, variables, functions, and operators that can be evaluated to a single result. Within an expression, the values of variables can be modified by a number of functions and operators. An Expression consists of one or more Operands (variables or literal

values) and one or more Operators (e.g., +, -, *, and /). No matter how complex, Expressions can eventually be evaluated to produce a single value (e.g., 1.323, or "True" or "False.") Functions modify the value of one or more variables to produce a result (i.e.,, ROUND [2.33333] produces the value two). Operators are used to combine two items. For example, the "+" operator combines Var1 and Var2 to produce a sum (e.g., Var3=Var1+Var2). Nearly all functions require arguments enclosed in parentheses and separated by commas. For arguments, there may not be any spaces between the function name and the left parenthesis.

Fisher Exact Test

A test of statistical significance used to determine how likely an observed association between an exposure and a disease could have occurred because of chance alone, and if the exposure was unrelated to the disease. The Fisher Exact Test is the test of choice when the expected values in a two-by-two table are less than five.

Incidence Rate

A measure of the frequency with which an event (e.g., a new case of illness) occurs in a population during a period of time. The denominator is the population at risk; the numerator is the number of new cases occurring during a given time period.

Mean

The measure of central location commonly called the average. It is calculated by adding all the individual values in a group of measurements and dividing by the number of values.

Median

The measure of central location that divides a set of data into two equal parts.

Numerator

The upper portion of a fraction.

Odds Ratio

A measure of association, which quantifies the relationship between an exposure and health outcome from a comparative study. Also known as the cross-product ratio.

P-Value

The probability that an observed association between an exposure and a disease could have occurred because of chance alone, if the exposure was unrelated to the disease.

Percentage

The number of patients with a characteristic divided by the total number of patients with the characteristic.

Prevalence

Measure of all cases of disease at a point of time. Computed by dividing the number of cases by the total population.

Proportion

A type of ratio in which the numerator is included in the denominator. The ratio of a part to the whole, expressed as a "decimal fraction" (e.g., 0.2), as a fraction (1/5), or, loosely, as a percentage (20%).

Protective Factor

An aspect of personal behavior or lifestyle, an environmental exposure, or an inborn or inherited characteristic associated with a decreased occurrence of disease or other health-related event or condition.

Rate

How often an event occurs in a defined population.

Ratio

The value obtained by dividing one quantity by another.

Risk

The probability that an event will occur (e.g., an individual will become ill or die within a stated period of time or age).

Risk Ratio

A comparison of the risk of a health-related event (e.g., disease or death in two groups).

Standard Deviation

A mathematical measure of the spread or dispersion of the data around the mean value for normally distributed data.

Standard Error (of the Mean)

The standard deviation of a theoretical distribution of sample means of a variable around the true population mean of that variable.

Statistical Significance

The measure of how likely a set of study results could have occurred by chance alone. The level of statistical significance is usually expressed by the p-value.

Trend

A long-term movement or change in frequency, usually upwards or downwards.

Variable

Any characteristic or attribute that can be measured.

Variance

A measure of the dispersion shown by a set of observations.

Appendix I — Data Sources

Assessment in Public Health

Centers for Disease Control and Prevention (CDC), Division of Public Health Surveillance and Informatics (DPHSI). *Assessment in Public Health.* CDC Web site. 2005.

http://www.cdc.gov/epo/dphsi/AI/resources.htm

BRFSS Survey Data

CDC. *Behavioral Risk Factor Surveillance System Survey Data.* Atlanta, GA: U.S. Department of Health and Human Services, CDC, 2003.

http://www.cdc.gov/brfss/index.htm

CDC Asthma and Allergies

CDC. *CDC Health Topic: Asthma and Allergies.* CDC website. 2005.

http://www.cdc.gov/health/asthma.htm

CDC WONDER

United States Department of Health and Human Services (US DHHS), CDC, National Center for Health Statistics (NCHS), "Bridged-Race Population Estimates, United States, 1990–2002, By Age Groups". Compiled from the April 1, 2000 resident population developed by the Bureau of the Census in collaboration with the NCHS on CDC WONDER On-line Database.

United States Department of Commerce, U.S. Census Bureau, Population Division; Census Data for Public Health Research, CDC WONDER On-line Database, March 2003.

http://wonder.cdc.gov

Community Health Assessment Clearinghouse

New York State, Department of Health (NYSDOH). *Community Health Assessment Clearinghouse.* NYSDOH website. 2005.

http://www.health.state.ny.us/nysdoh/chac/index.htm

Epi Info™

CDC, Division of Public Health Surveillance and Informatics (DPHSI). *What Is Epi Info™?* CDC website. 2005.

http://www.cdc.gov/epiinfo

The National Center for Health Statistics

Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS). CDC website. 2005.

http://www.cdc.gov/nchs/

Statewide Planning and Research Cooperative System (SPARCS)

New York State Department of Health. "Asthma Hospitalization Rates by County, New York State Residents 2000–2002." Claritas Corporation. 2005.

United States Census Bureau

U.S. Census Bureau; "Single Years of Age Under 30 Years and Sex 2000, for Albany County, NY: 2000;" published 2000.

http://www.census.gov/census2000/states/ny.html

References

Alperin, Melissa, MPH, CHES, Cam Escoffery, PhD, MPH, CHES. *Using Epi Info: A Step-by-Step Guide*. ToucanEd, 2003.

Alperin, Melissa, MPH, CHES, Kathleen Minor, PhD, MPH, CHES. *Using Epi Info: A Step-by-Step Guide*. ToucanEd, 1997.

CDC. Excellence in Curriculum Integration through Teaching Epidemiology (EXCITE). *Glossary of Epidemiological Terms*. CDC website. July 5, 2005.

http://www.cdc.gov/excite/library/glossary.htm

Dean AG. Epi Info™ Version 6: A Word Processing, Database, and Statistics System for Epidemiology on Microcomputers. Atlanta, GA: CDC, Epidemiology Program Office, 1994.

Epidemiology Program Office. *Epi Info™ 2000 Users' Manual*. Atlanta, GA: CDC, 2000.

North Carolina Center for Public Health Preparedness. *Epi Info Online Training Resource*. Chapel Hill: 2005.

http://www.sph.unc.edu/nccphp/training/all_trainings/at_epi_info.htm

Washington State University, College of Veterinary Medicine, Department of Veterinary Clinical Sciences, Field Disease Investigation Unit; John Gay, DVM PhD. *Clinical Epidemiology and Evidence-Based Medicine Glossary*. Website. Updated August 11, 2004.

http://www.vetmed.wsu.edu/courses-jmgay/GlossExpDesign.htm

Young Epidemiology Scholars Competition. Sponsored by the Robert Johnson Wood Foundation and the College Board. *Epidemiology Glossary*. Website. August 4, 2005.

http://www.collegeboard.com/yes/ae/gloss.html