STEPS Epi Info
Training Guide

World Health Organization

WHO STEPS Surveillance Support Materials

Epi Info Training Guide

Table of Contents

Part 1: Introduction to Epi Info

Overview

Introduction Epi Info™ is a public domain software package designed for the global community of public health practitioners and researchers.

It provides for easy form and database construction, data entry, and analysis with epidemiologic statistics, maps, and graphs.

Intended audience This manual is intended for all parties responsible for analysing the STEPS data and producing STEPS Country Reports.

This manual only covers the analysis component of Epi Info.

Guide to using this manual This manual has been written in eight parts. Each part is introduced with an overview and a short table of contents to help you find specific topics.

Each part contains procedures, exercises, and assessment items. If you skip the exercises in any of the parts you may have problems using the programs in later parts.

Other resources This manual is to be read and used in conjunction with the following resources:

- STEPS_EpiInfo.mdb
- STEPS_Epi_Info_Instrument.pdf
- STEPS_Epi_Info_training_guide.pdf

The materials need to be placed in your C drive under
C:\STEPS\EpiInfo\Training

Overview, **Continued**

Capacity building   WHO is available to help build country capacity in using Epi Info.

In this part   This section contains the following topics:

Topic	See Page
Installing Epi Info	1-3
Copying Training Materials to your Computer	1-4
Opening and Navigating Epi Info	1-5
Opening and Closing the Analysis Module	1-6
Opening a Project and Data Table	1-7

Installing Epi Info

Introduction   Before using Epi Info, you must install it on your computer.

Procedure   Epi Info is freely available for downloading from the web from the STEPS website : http://www.who.int/chp/steps/resources/EpiInfo/en
or from the CDC website :
http://www.cdc.gov/epiinfo/pc/prevversion.htm

After downloading the installation file, run the file and follow the screen prompts to install Epi Info.

When installation is complete, you should be able to launch Epi Info directly from either:

• the Epi Info icon on the computer desktop, or
• the Epi Info program listing in the computer's "Start" pop-up menu.

Copying Training Materials to your Computer

Introduction
Before you can follow the exercises and use the examples presented in this manual, you must copy all the Epi Info training materials from the STEPS website to your computer.

Purpose
Epi Info has to be able to open the data and be able to write to the data as well. Therefore, any data to be analyzed with Epi Info must be copied to a hard drive or other writeable media.
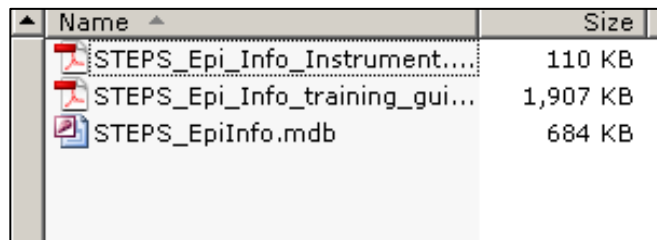
Procedure
Go to the STEPS Epi Info Training website:
http://www.who.int/chp/steps/resources/EpiInfoTraining/en/
and download the 3 files : the training database, the sample questionnaire and the electronic copy of this training manual. Save these files to the "C:\STEPS\EpiInfo\Training" folder on your computer.

Check
Open the folder "C:\STEPS\EpiInfo\Training" and see if your folder looks like the figure below.

Q: How many files are displayed in the window?

A: 3 files

Name ▲	Size
STEPS_Epi_Info_Instrument....	110 KB
STEPS_Epi_Info_training_gui...	1,907 KB
STEPS_EpiInfo.mdb	684 KB

Opening and Navigating Epi Info

Introduction   Epi Info is a Windows-based application; it is easy to navigate and provides all the analysis options at a click of the mouse.   It also provides a program editor so that you can save your work and come back to it later.

Terminology   Some of the specific Epi Info terms used in this manual are described in the table below.

Term	Description
Command	Term for analysis that provides Epi Info with the appropriate syntax to perform desired analysis. (i.e. frequency or list).
Program (.pgm)	A program (.pgm) is saved analysis syntax associated with a project. Programs can be edited and modified easily to make processing more convenient.
Project	The name of the actual database (.mdb file). All the programs and data are stored within the project. This manual uses STEPS_EpiInfo.mdb.
Variable	Any characteristic or attribute that can be measured.

Procedure   To open Epi Info, double click on the Epi Info icon   on your desktop.

If you do not have an icon on your desktop, click on the Start menu and select All Programs > Epi Info > Epi Info

Check   The following screen should appear on your screen.

Q: What version of Epi Info appears on your screen?

A: 3.5.4

Opening and Closing the Analysis Module

Introduction   Epi Info contains modules that support data entry, data analysis, mapping, graphing, and other functions.

This manual only covers the analysis module. Information about other Epi Info modules can be found in the Epi Info online help system.

Procedure   Follow the steps below to **open** and **close** the Analysis module.

Step	Action
1	To **open** the Analysis module, click once on the "Analyze Data" button of Epi Info. **Note:** Be patient for a moment while the module loads.
2	To **exit** the Analysis module, click "Exit".

Components   The Analysis screen has three components as shown in the figure below.
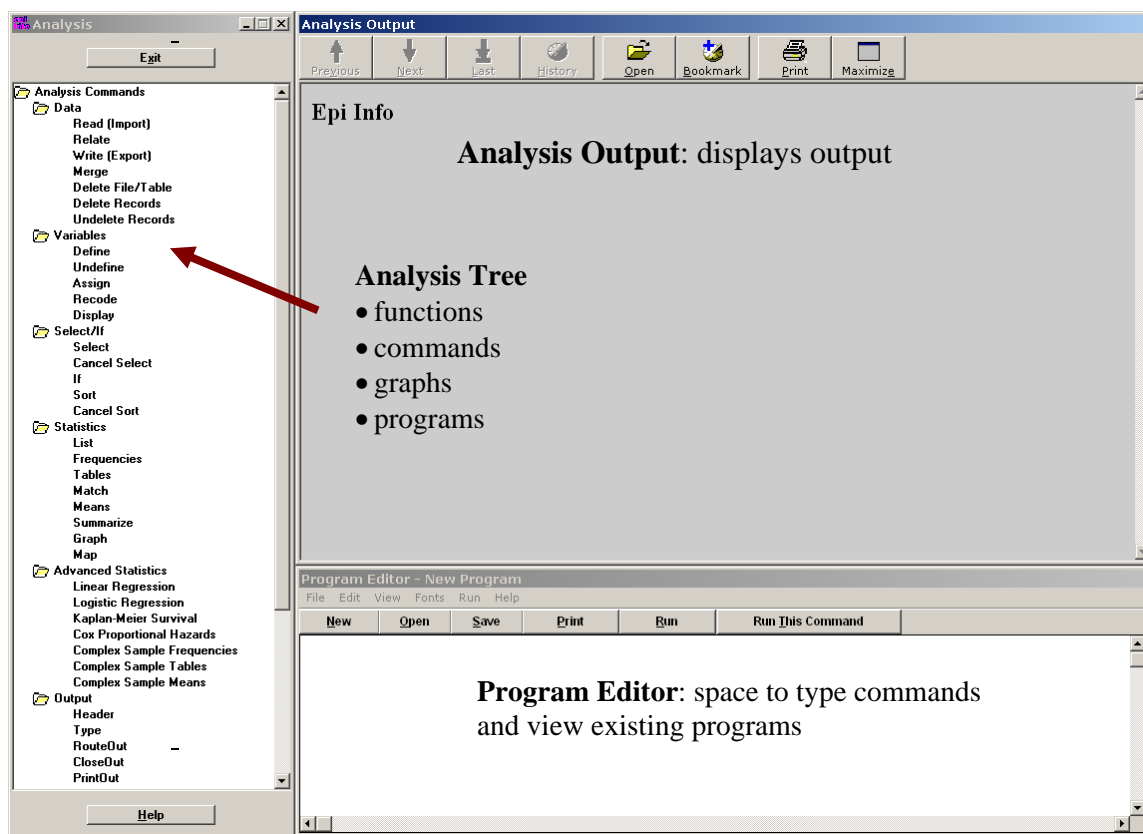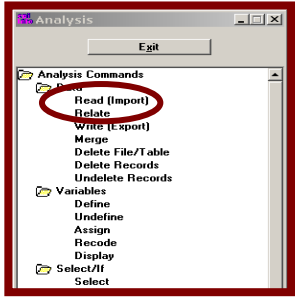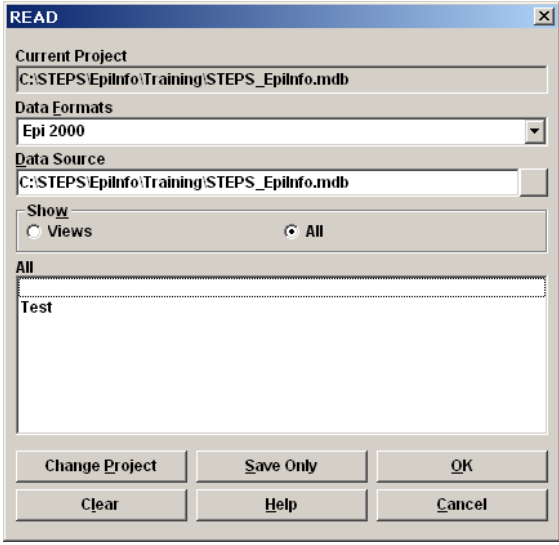
Opening a Project and Data Table

Introduction   Epi Info stores data and programs together in MS Access files called Projects. Projects can contain multiple data tables and multiple programs.

To analyze STEPS data you will need to open a Project and then select the Data Table that you want to analyze.

Procedure   Follow the steps below to open a Project and select a data table.

Step	Action
1	Open Epi Info if it is not already open.
2	Click "Analyze Data".
3	Select "Data", then "Read (Import)" from the Analysis Tree.
4	Click on "Change Project" to change the Project to "C:\STEPS\EpiInfo\Training\ STEPS_EpiInfo.mdb".
5	Select "Test" in the bottom half of the screen then click "OK".

Continued on next page

Opening a Project and Data Table, Continued

Check   Look at the Analysis Output screen, the record count should be 1289.

If you do not have 1289 in your record count you did not successfully open the Test dataset. Try to open the dataset again by repeating steps 2-5 in the of the instructions on the previous page.

Part 2: Managing Variables

Overview

Introduction   The Test dataset contains the data for the WHO STEPS Instrument for Epi Info Training.

This data table contains the information collected during the survey. This information is represented in the database as variables.

In this part   This part contains the following topics:

Topic	See Page
Summary of Functions	2-2
Displaying the Names and Variables in a Data Table	2-3
Listing the Variables in the Data Table	2-5
Defining a New Variable	2-7
Assigning a Variable	2-10
Recoding a Variable	2-12
Calculating Derived Variables	2-15
Saving Your Data	2-17
Setting Epi Info Options	2-19

Summary of Functions

Introduction   There are several important functions that you must be able to perform to manage the variables in a data table.

Functions   Each of these functions is briefly summarized in the table below.

Function	Allows you to …
Listing variables	Explore the variables currently in a data table.
Defining and recoding variables	Group responses into more useful categories and lets you attach meaningful text labels to variable values.
Defining and assigning variables	Create new variables for uses in your analyses.
Recoding existing variable	Recode variables into a new variable and keep the original variable untouched. **Notes:** • If you make a mistake then you can redo the recode into the variable. • If you use the original variable and recode into that you will not be able to undo any mistakes.
Cleaning or filtering variable	Clean or filter a variable. What you need to filter may change according to which question you are working with <table><tr><td>**If you…**</td><td>**Then you…**</td></tr><tr><td>Filter or clean a variable for one type of analysis</td><td>May need a different one for another.</td></tr><tr><td>Use a new variable to filter</td><td>Use the original data again for other analyses.</td></tr></table>
Renaming variables	Rename variables for use in tables (e.g. "men" and "women" is much more user friendly than "1" and "2"). **Note:** Keep the original numeric values because they are easier to use than labels during analysis.

Displaying the Names and Variables in a Data Table

Introduction   Displaying the names of the variables in a data table provides useful information on the structure and status of the data table.

Display command   The Display command will show the names, data types, sizes, and other information about each variable in the data table.

Procedure   Follow the steps below to display the names of the variables in a data table.

Step	Action
1	Select "Variables, Display" from the Analysis Tree.
2	Select "Variables" in "Information for".
3	Click "OK".

Exercise   Display the variables in the current data table. The DISPLAY dialog box should look like the figure below.
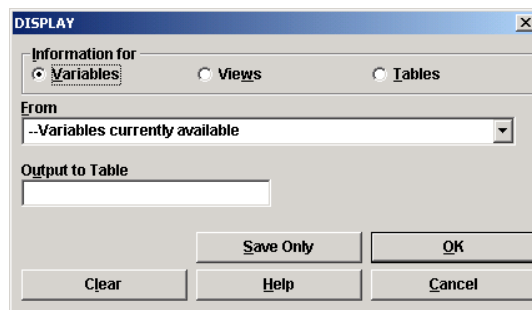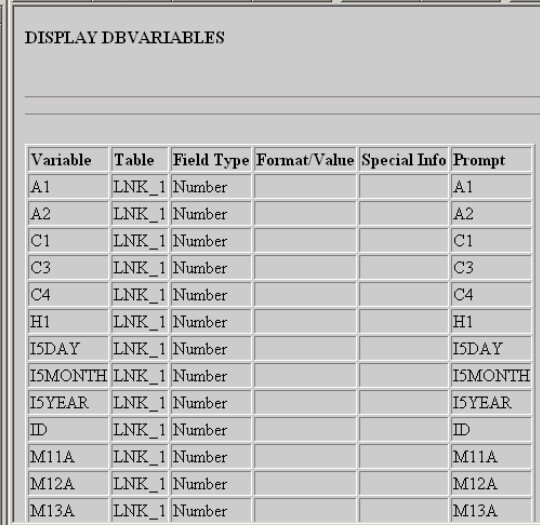
Continued on next page

Displaying the Names and Variables in a Data Table, Continued

Check   The output should look like the figure below.

Variable	Table	Field Type	Format/Value	Special Info	Prompt
A1	LNK_1	Number			A1
A2	LNK_1	Number			A2
C1	LNK_1	Number			C1
C3	LNK_1	Number			C3
C4	LNK_1	Number			C4
H1	LNK_1	Number			H1
I5DAY	LNK_1	Number			I5DAY
I5MONTH	LNK_1	Number			I5MONTH
I5YEAR	LNK_1	Number			I5YEAR
ID	LNK_1	Number			ID
M11A	LNK_1	Number			M11A
M12A	LNK_1	Number			M12A
M13A	LNK_1	Number			M13A

DISPLAY DBVARIABLES

Q. What is the field type for A1?

A. Number

Listing the Variables in the Data Table

Introduction Viewing data in a list is similar to looking at data in an Excel spreadsheet or looking at the data directly in the Access table.

Purpose Viewing data in a list does not help much in terms of analysis but it can be used to visually verify the data.

For example, if you are recoding respondents' ages into age ranges, you may want to view the original variable and the newly recoded variable side by side. You will be able to visually see that the recode worked properly.

Procedure Follow the steps below to view a list of variables.

Step	Action
1	Select "Statistics, List" from the Analysis Tree
2	Select the variable(s) you want to view. **Notes:** • Variables will be displayed in alphabetical order. • You may select as many variables as you like. • * will select all the variables in the data table.
3	Make sure Display Mode is set to Grid. Grid is more readable than HTML. **Caution:** "Allow Updates" will let you to actually change the data.
4	Click "OK".

Exercise View Age (C3) and Sex (C1) in a list. The LIST dialog box should look like the figure below.
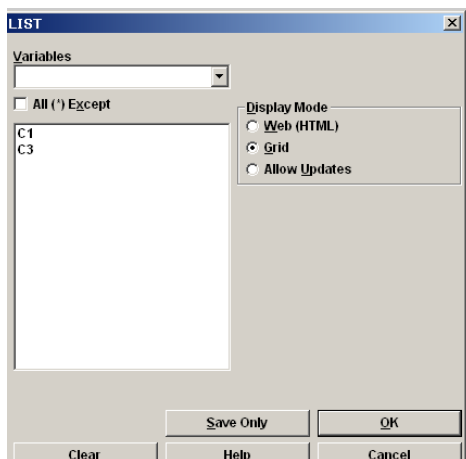
Listing the Variables in the Data Table, Continued

Check　　　　The output should look like the figure below.

Analysis Output	
C1	**C3**
2	37
2	59
1	59
2	56
2	59
1	33
2	55
1	42
2	32
2	39
1	41

Q. What variables can you see?

A. <u>C1 and C3</u>

Defining a New Variable

Introduction   The first step in using a new variable is to create or "define" it to Epi Info. Data tables contain rows and columns. The rows are the individual records for each participant and the columns contain the information for each question or variable. Defining variables is like creating additional columns in the data table.

The variable for the question: "Do you currently smoke any tobacco products, such as cigarettes, cigars or pipes?", is T1. The variable name comes from the coding column in the instrument, not the question number.

Variable types   Each variable type is listed and briefly described in the table below.

Variable type	Description
Standard	• Used as temporary variables behaving like variables in the database. • Lose their values and definitions at the next READ statement.
Global	• Persist for the duration of program execution. • Used in Epi Info Analysis to store values between changes of data source.
Permanent	• Stored in Epi Info and retain any value assigned until the value is changed by another assignment or the variable is undefined. • Are shared among Epi Info programs and persist even if the computer is shut down.

Defining a New Variable, Continued

Identifying variables

Use the "WHO STEPS Instrument for Epi Info training" to complete these exercises.

Identify the variables associated with the questions below and write them in the table.

Question	Variable
How old are you?	
On average, **how many** of the following do you smoke each day?- Manufactured cigarettes	
On average, **how many** of the following do you smoke each day?- Pipes full of tobacco	
Height	

Check your work

The table below presents the answers for the exercise above

Question	Variable
How old are you?	C3
On average, **how many** of the following do you smoke each day?- Manufactured cigarettes	T5a
On average, **how many** of the following do you smoke each day?- Pipes full of tobacco	T5c
Height	M3

Procedure

To create a variable follow the steps below.

Step	Action
1	Select "Variables, Define" from the Analysis Tree.
2	Type the name of the new variable in the Variable Name box.
3	Select "Standard" as Scope of variable.
4	Click "OK".

Continued on next page

Defining a New Variable, Continued

Exercise   Define a variable named Sex as a standard variable. In a subsequent exercise, we will recode this variable into "Male" and "Female" for use as labels in frequencies and other analyses.

The DEFINE dialog box should look like the figure below.
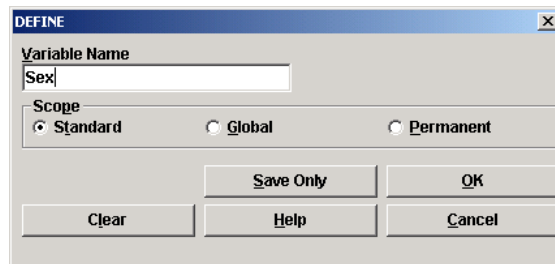
Check   The "Define" command does not generate any output of its own. To see if the command worked, use the "Display" command to display a list of all the variables in the data table.

Assigning a Variable

Introduction   Assigning a variable enables you to "populate" a variable with new or different values. You can assign new variables with:

- the values of existing variables
- constants, or
- the results of calculations.

Defining and assigning variables   Defining variables and assigning variables are related actions, but they are not the same thing.

- Defining a variable creates the variable name.
- Assigning a variable creates the content for the variable.

Note: Before you can assign a value to a variable, you must define it.

Procedure   Follow the steps below to assign the values in C1 ("What is your sex?") to the newly created variable "Sex". Follow the steps below to assign a variable.

Step	Action
1	Select "Variables, Assign" from the Analysis Tree.
2	Select "Sex" as the variable to assign from the "Assign Variable" drop down menu.
3	Click the "=" button.
4	Select C1, the variable that is "giving" its values to Sex, from the "Available Variables" drop down menu.

Check   Try and "read" the = Expression field in association with the assign variable. It should read, "Sex equals C1". We are setting Sex to the same values as C1. The ASSIGN dialog box should look like the figure below.
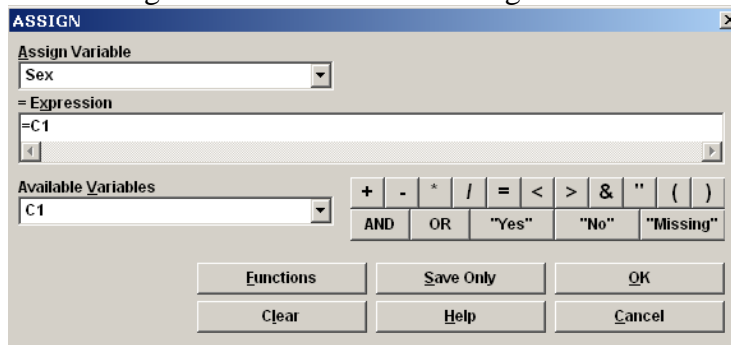
Assigning a Variable, Continued

Exercise   Follow the procedure above to assign the values of C1 to the variable Sex.

Check your work   List the values of C1 and Sex together. Compare their values. They should match exactly. Your listing should look something like the figure below.

Sex	C1
2	2
2	2
1	1
2	2
2	2
1	1
2	2
1	1
2	2
2	2
1	1
1	1

Recoding a Variable

Introduction   Recoding variables is useful when you want to:

- Create a new variable based on a classification from another variable.
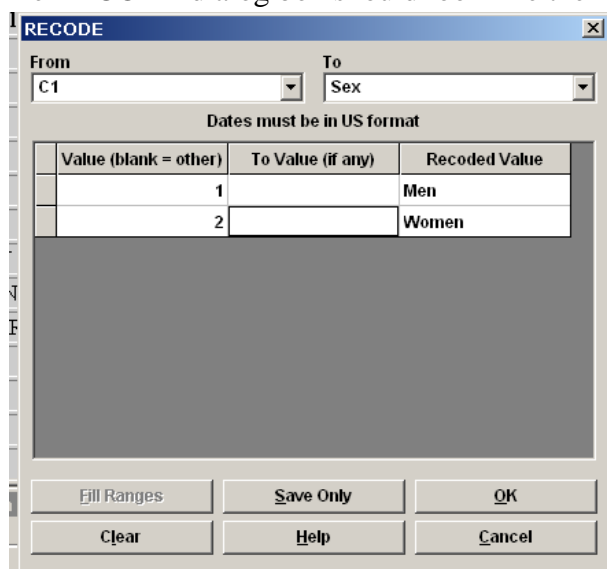- Use words as labels in instead of values for outputs (tables, frequencies, etc.)

Note: A variable must be defined before it can be recoded.

Procedure   Follow the steps below to recode a variable.

Step	Action
1	Select "Variables, Recode" from the Analysis Tree.
2	Select the variable that contains the values in the "From" drop down menu.
3	Select the variable that will receive the values in the "To" drop down menu.
4	Fill in the value(s) to recode in the "Value" and "To Value" fields.
5	Fill in the desired new value in the "Recoded Value" field.
6	Press the Enter key to create a new line for more values if needed.
7	Click "OK".

Exercise 1   Recode the values 1 and 2 in C1 to the values "Men" and "Women" in the variable Sex (which was defined in the previous exercise).

The RECODE dialog box should look like the figure below.

Recoding a Variable, Continued

Check   List the values of C1 and Sex and compare their values. Is the value "1" paired with "Men" and the value "2" paired with "Women" on all the rows?

The output should look like the figure below.

sex	C1
Women	2
Women	2
Men	1
Women	2
Women	2
Men	1
Women	2
Men	1
Women	2

Exercise 2   Recode the values in C3 ("How old are you?") to the values "25-34", "35-44", "45-54", "55-64" into the variable AgeRange

This exercise groups values in addition to creating useful text labels on the data. You will need to define AgeRange before you can recode the values of C3 into it. The RECODE dialog should look like the figure below.

RECODE

From: C3   To: AgeRange

Dates must be in US format

Value (blank = other)	To Value (if any)	Recoded Value
25	34	25-34
35	44	35-44
45	54	45-54
55	64	55-64

Fill Ranges   Save Only   OK
Clear   Help   Cancel

Continued on next page

Recoding a Variable, Continued

Check
List the values of C3 and AgeRange and compare their values. Are the values paired the way you would expect on all the rows?

The output should look like the figure below.

C3	AgeRange
37	35-44
59	55-64
59	55-64
56	55-64
59	55-64
33	25-34
55	55-64
42	35-44
32	25-34

Correcting Errors
If your recode looks incorrect, follow the steps below to correct it.

Step	Action
1	Undefine AgeRange <table><tr><th>Step</th><th>Action</th></tr><tr><td>1</td><td>Select "Variables, Undefine" from the Analysis Tree.</td></tr><tr><td>2</td><td>Select the name of the variable from the drop down list</td></tr><tr><td>3</td><td>Click "OK".</td></tr></table>
2	Redefine AgeRange
3	Recode AgeRange

Calculating Derived Variables

Introduction   A derived variable is a variable that is calculated based on 2 or more independent variables.

Epi Info can easily calculate derived variables by using the Assign function from the Analysis Tree.

Procedure   The table below details the process of calculating a derived variable.

Step	Action
1	Define a new variable
2	Select "Variables, Assign" from the Analysis Tree
3	Enter the equation for calculating the variable

Exercise   Calculate BMI (body mass index) using M3 (height) and M4 (weight) we need to know the height and weight of the participants. These questions are M3 and M4. Follow the steps below to calculate BMI.

Step	Action
1	Define a new variable BMI.
2	Select "Variables, Assign" from the Analysis Tree.
3	Select BMI in the "Assign Variable" field.
4	Enter the equation: "=(M4/(M3*M3))*10000" in the "=Expression" field.
5	Confirm that your ASSIGN window matches the figure below and click "OK".
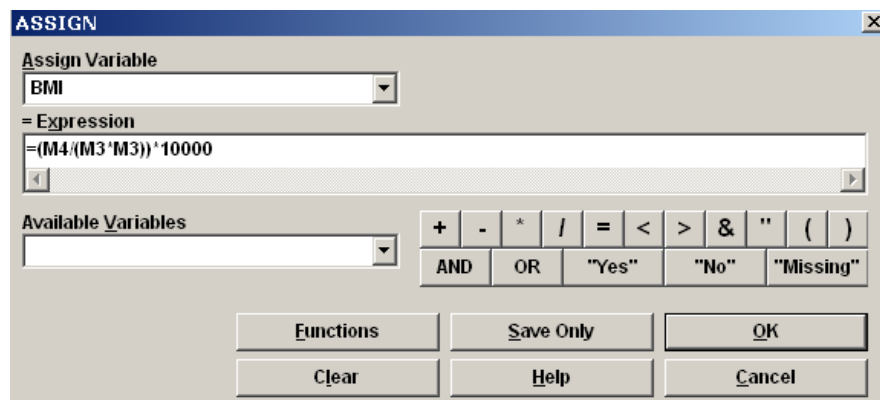
Calculating Derived Variables, Continued

Check   Run a list of M3, M4, and BMI to see if the calculation was successful.

Exercise 2   Calculate the mean systolic blood pressure of the participants. Create a
                   variable SBP and calculate the mean SBP for each participant using this
                   equation.

                   $((M11a+M12a+M13a)/3)$ This is adding together all three values for systolic
                   blood pressure and then dividing it by three.

                   Do not worry about missing values and the validity of the calculation . We will
                   cover this in detail in another session.

Check   Create a list of the M11a, M12a, M113a, and SBP.

For your info:

This is the syntax that will define SBP,
assign SBP, and list the results.

```
DEFINE SBP
ASSIGN SBP=((M11a+M12a+M13a)/3)
LIST  M11a M12a M13a SBP GRIDTABLE
```

Saving your Data

Introduction   Each time you define, recode, or assign a value to a variable, the changes are saved in the associated data table. If you close Epi Info or read another table into the analysis section, all the changes/additions made to the variables will disappear. If you want to keep certain changes you made to a dataset, you need to save the current dataset by using the WRITE function. You can write your results to the data table that you have open or you can create an entirely new data table in your project.

Purpose   Writing results preserves your work for future use. This is especially important when you have created new variables you wish to use later.

Procedure   Follow the steps below to write your results to a data table in the current project.

| Step	Action
1	Select "Data, Write (Export)" from the Analysis Tree.
2	Select "All" (if you only wanted to write some of the variables, you would need to manually select them).
3	Select "Replace".
4	Select the current project (Access database) as the "File Name".
5	• Select the name of the data table that you want to write to, or • Type in the name of the new data table that you want to create.
6	Click "OK".

Continued on next page

Saving your Data, Continued

Exercise Save the current version of the Test dataset. The WRITE dialog box should look like the figure below.
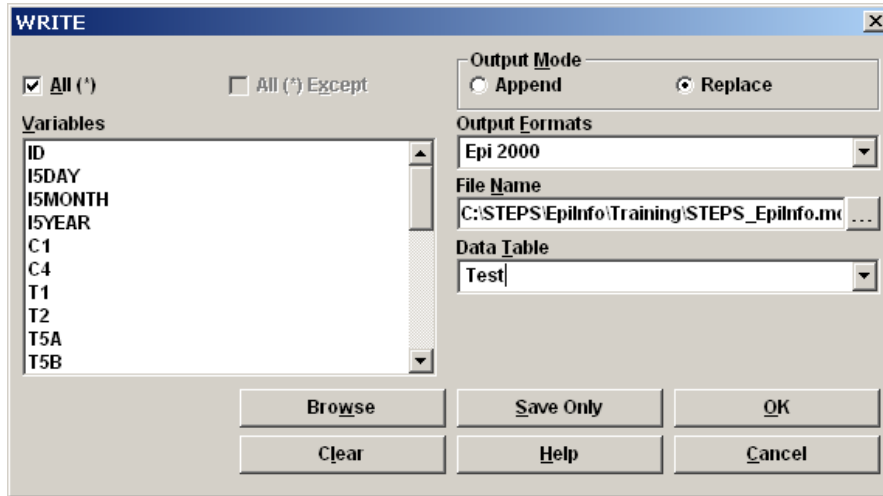
Setting Epi Info Options

Introduction   Epi Info has numerous system settings and options that control important aspects of its operation. Using the Set command, the options can be adjusted to change:

- how output is formatted
- how missing data values are treated
- which statistics are included.

When options can be changed   Epi Info options can be changed at any time. Once changed, the new values remain in effect even if you exit Epi Info and re-enter.

Important settings   Two of the most important settings and what they should be set to are provided in the table below.

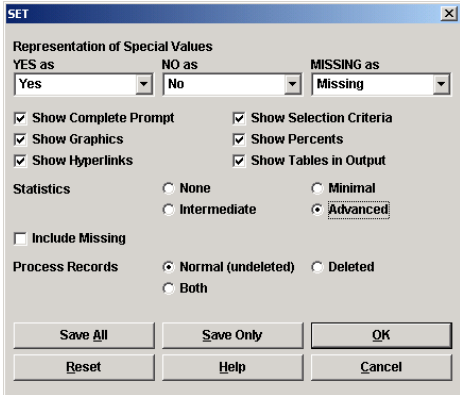| Setting	Should…	So that…
Statistics	Be set to "Advanced"	Complex sample tables will display all of the confidence interval information needed to determine subgroup differences.
Include Missing	Not be checked	Missing values will be excluded from frequencies and other calculations.

Setting Epi Info Options, Continued

Procedure   Follow the steps below to change the Epi Info system options.

Step	Action
1	Select "Options, Set" from the Analysis Tree.
2	Alter any settings in the Set dialog box shown below as required for your analysis.
3	Click "OK".

Exercise   Follow the steps below to check the options in your copy of Epi Info.

Step	Action
1	Select "Options, Set" from the Analysis Tree.
2	Make sure that "Statistics" is set to "Advanced".
3	Make sure that "Include Missing" is not checked.
4	Change these settings as needed. **Note:** This manual assumes the options are set this way to ensure the program outputs contain the necessary information.

Check   Click "Options, Set" from the Analysis Tree again. Verify settings are correct.

Part 3: Basic Statistics

Overview

Introduction   Epi Info can perform a range of statistical functions from basic statistics to advanced statistics. This session will familiarize you with the basic statistical functions available in Epi Info

Options   Basic statistics in Epi Info include the options to perform the analysis weighted and/or stratified.

- **Weighted frequencies**/**tables** use the weight variable on the data table to make the percentages representative of the target population of the STEPS survey. Always report weighted percentages in reports and other materials.
- **Stratified frequencies/tables** run a separate frequency/table for each value in another variable.

Note about confidence intervals   Basic Epi Info frequencies and tables will produce confidence intervals when a weighted analysis is performed. However, these confidence intervals are not correct for most STEPS surveys because a simple random sample is assumed. The complex sample design of used in most surveys is therefore ignored.

To correctly calculate confidence intervals you must use the Advanced Statistics commands (covered in Part 4), in which you can inform Epi Info of the complex sample design of your survey.

Contents   This section contains the following topics:

Topic	See Page
Running a Frequency	3-2
Running Tables	3-6
Reading an Output Table	3-9
Creating Labels for Outputs	3-10

Running a Frequency

Introduction   Frequencies are tables that show how many records have each value of a single variable and the percentage of records that have each value.

Epi Info also produces a histogram of the distribution.

Procedure   Follow the steps below to run a frequency.

Step	Action
1	Select "Statistics, Frequencies" from the Analysis Tree.
2	Select the variable or variables you are interested in from the "Frequency of" drop down list.
3	Optional: Select "weight" from the Weight drop down list to run a weighted frequency.
4	Optional: Select a variable from the "Stratify by" drop down list to stratify the frequency output.
5	Click "OK".

Exercise 1   Run a frequency on T1 ("Do you currently smoke any tobacco products, such as cigarettes, cigars or pipes?") and on AgeRange. The FREQ dialog box should look like the figure below.
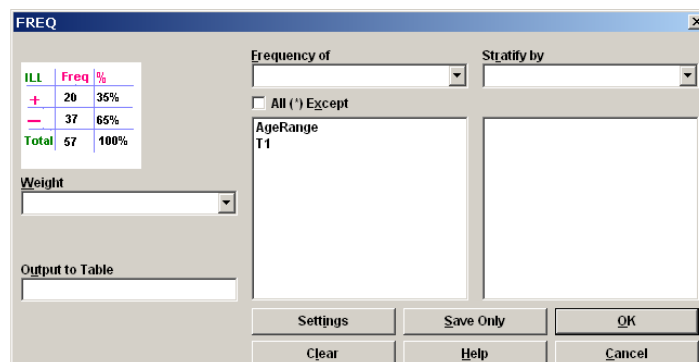
Running a Frequency, Continued

Check Compare the output tables. Notice how the recoded AgeRange improves the readability of the output. The output should look like the figure below.

Forward

AgeRange	Frequency	Percent	Cum Percent	
25-34	318	24.8%	24.8%	
35-44	336	26.2%	51.1%	
45-54	318	24.8%	75.9%	
55-64	309	24.1%	100.0%	
Total	1281	100.0%	100.0%	

T1

Back Forward Current Procedure

T1	Frequency	Percent	Cum Percent	
1	493	38.4%	38.4%	
2	790	61.6%	100.0%	
Total	1283	100.0%	100.0%	

Q1: How many participants are 35-44?
A1: 336

Q2: How many participants answered the question (T1)?
A2: 1283

Exercise 2 Run a weighted frequency on T1. The FREQ dialog box should look like the figure below.

Continued on next page

Running a Frequency, Continued

Check Notice how the Frequency and Percent numbers are different from the unweighted frequency on T1 from Exercise 1.

- Weighted frequencies have decimal values because each participant's weight is a decimal number.
- Frequently, weighted and unweighted percentages are close in value but weighted percentages are always the correct ones to use in reports and other materials.

The output should look like the figure below.

FREQ T1 WEIGHTVAR=WSTEP1

Next Procedure

Forward

T1	Frequency	Percent	Cum Percent	
1	372874.77	41.7%	41.7%	
2	521214.19	58.3%	100.0%	
Total	894088.96	100.0%	100.0%	

Q: What percentage of participants are current smokers?

A: 41.7%

Continued on next page

Running a Frequency, Continued

Exercise Run a weighted frequency on T1 ("Do you currently smoke any tobacco products, such as cigarettes, cigars or pipes?"), stratified by Sex.
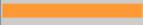
Check Notice the headings above the tables indicating that one table is for Females and the other table is for Males.

The output should look like the figure below.

T1, Sex=Men

Forward

T1	Frequency	Percent	Cum Percent	
1	226231.03	46.3%	46.3%	
2	262000.80	53.7%	100.0%	
Total	488231.83	100.0%	100.0%	

T1, Sex=Women

Back Forward Current Procedure

T1	Frequency	Percent	Cum Percent	
1	146643.74	36.1%	36.1%	
2	259213.39	63.9%	100.0%	
Total	405857.12	100.0%	100.0%	

Q. What percentage of males are current smokers?
A. 46.3%

Q. What percentage of females are current smokers?
A. 36.1%

Running Tables

Introduction   Tables, or cross-tabulations, are two-way frequencies. In other words, tables show how the values of two variables correspond to each other.

Purpose   Epi Info produces tables that show counts and percentages along two dimensions. Epi Info tables are also used to produce a histogram of the results.

Procedure   Follow the steps below to run a table.

Step	Action
1	Select "Statistics, Tables" from the Analysis Tree.
2	Select the variable you want to appear down the left side of the table from the "Exposure Variable" drop down list.
3	Select the variable you want to appear across the top of the table from the "Outcome Variable" drop down list.
4	**Optional**: Select "weight" from the Weight drop down list to run a weighted table.
5	**Optional**: Select a variable from the "Stratify by" drop down list to stratify the table output.
6	Click "OK".

Running Tables, Continued

Exercise 1   Run a table comparing Sex with T1 ("Do you currently smoke any tobacco products, such as cigarettes, cigars or pipes?"). The TABLES dialog box should look like the figure below.
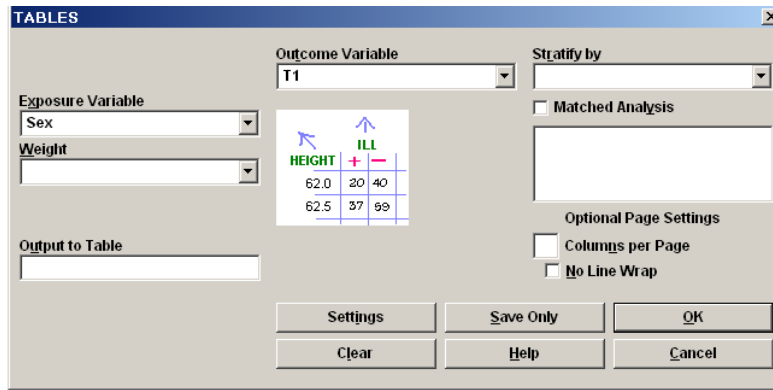
Check   Check the following details:
- T1=1 for participants who currently smoke any tobacco products.
- T2=2 for participants that do not currently smoke any tobacco products.
- Each cell, or box, contains the count of participants that answered the question, the row percent, and the column percent.

The output should look like the figure below.

Q1: How many men currently smoke?
A1: 264

Q2: How many women answered T1?
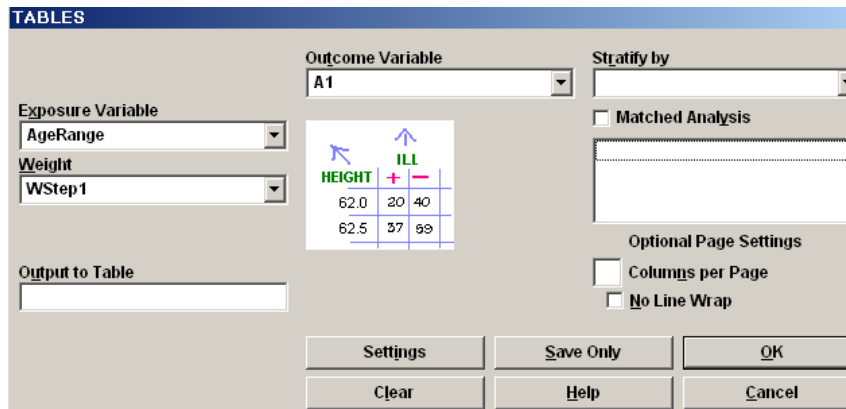**A2:** 651

Continued on next page

Running Tables, Continued

Exercise 2   Run a weighted table comparing AgeRange with A1 ("Have you consumed alcohol in the past 12 months?").

The TABLES dialog box should look like the figure below.

Check   Check the following details:
- A1=1 for participants who consumed alcohol in the past 12 months.
- A1=2 for participants who did not consume alcohol in the past 12 months.
- Each cell, or box, contains the weighted count of participants that answered the question, the weighted row percent, and the weighted column percent.

The output should look like the figure below.

Forward

		A1		
AgeRange		1	2	TOTAL
25-34		170184.6	11954.97	182139.6
	Row %	93.4	6.6	100.0
	Col %	29.3	12.6	27.0
35-44		198006.5	33918.74	231925.3
	Row %	85.4	14.6	100.0
	Col %	34.1	35.7	34.3
45-54		209490.8	48286.05	257776.8
	Row %	81.3	18.7	100.0
	Col %	36.1	50.8	38.1
55-64		3036.96	851.376	3888.336
	Row %	78.1	21.9	100.0
	Col %	0.5	0.9	0.6
TOTAL		580718.9	95011.14	675730.1
	Row %	85.9	14.1	100.0
	Col %	100.0	100.0	100.0

Reading an Output Table

Introduction   Epi Info output tables are not very descriptive and can be confusing to read. Use the table below as a guide to reading your output tables

Reading a table   This table relates to smokers among the entire sampled population.

Check the following details:
- The **Row %** provides the percentage among the age group. The highlighted row says that from the age-group, 35-44 year, 28.4% are daily smokers and 62.3% do not smoke.
- The **Col %** provides the percentage within each column. The Col % for 25-34 year olds, current smokers means that of all current smokers 39.2% are between 25 and 34 years old.
- We use the Row % for STEPS

Forward

AgeRange	current smoker(non-daily)	daily smoker	non smoker	TOTAL
25-34	47	120	151	318
Row %	14.8	37.7	47.5	100.0
Col %	39.2	32.6	19.1	24.9
35-44	31	95	208	334
Row %	9.3	28.4	62.3	100.0
Col %	25.8	25.8	26.3	26.1
45-54	25	93	198	316
Row %	7.9	29.4	62.7	100.0
Col %	20.8	25.3	25.1	24.7
55-64	17	59	231	307
Row %	5.5	19.2	75.2	100.0
Col %	14.2	16.0	29.2	24.0
99	0	1	2	3
Row %	0.0	33.3	66.7	100.0

C

Creating Labels for Outputs

Introduction   The frequencies, graphs, and tables that were presented in this session all used numbers to designate the values. However this requires that all users of the outputs know what T1 is and that T1=1 is yes and T1=2 is no. It is much easier for all the users if you create labels for the data that is presented in the outputs.

Process   The table below describes the process for creating labels for your outputs.

Stage	Description
1	Define a variable to hold the labels of the variable used for the output.
2	Assign this new variable values that are labels for output by using the recode function.

Note: There are a variety of ways to assign values to new variables. The method used reflects the complexity of the new variable (e.g. the value based on a single variable or a compilation of many variables). The section only deals with simple variables that are assigned based on one value.

Deciding on what labels to use   The question T1 has two possible responses. These responses are:

- 1
- 2

If we wanted to create labels for these we could use:

- 1 = Yes
- 2 = No

However this does not provide much information. We still do not know anything about what the Yes is in response to. Ideally labels contain information based on the question and the response. These labels below provide more information for the user to better understand what the output means.

- Currently smokes tobacco products.
- Does not currently smoke tobacco products.

Continued on next page

Creating Labels for Outputs, Continued

Procedure Follow the steps below to create labels for T1.

Step	Action
1	Define a variable to hold the labels for T1. Name the variable smoke.
2	Click on **Variables, Recode** from the Analysis Tree.
3	Recode T1 into smoke (follow the diagram below for an example).
4	List T1 and smoke to see if the recode functioned properly.

Step 1

DEFINE

Variable Name
smoke

Scope
⊙ Standard   ○ Global   ○ Permanent

Save Only   OK
Clear   Help   Cancel

Step 3

RECODE

From   To
T1   smoke

Dates must be in US format

Value (blank = other)	To Value (if any)	Recoded Value
1		Currently smokes tob
2		Does not currently sm

Fill Ranges   Save Only   OK
Clear   Help   Cancel

Step 4

smoke	T1
Does not currently smoke tobacco products	2
Currently smokes tobacco products	1
Does not currently smoke tobacco products	2
Does not currently smoke tobacco products	2
Does not currently smoke tobacco products	2
Currently smokes tobacco products	1
Currently smokes tobacco products	1
Does not currently smoke tobacco products	2

Frequency example Look at the two frequencies below. The frequency which contains the value labels instead of numbers is much clearer.

Forward

T1	Frequency	Percent	Cum Percent	
1	493	38.4%	38.4%	
2	790	61.6%	100.0%	
Total	1283	100.0%	100.0%	

Forward

smoke	Frequency	Percent	Cum Percent	
Currently smokes tobacco products	493	38.4%	38.4%	
Does not currently smoke tobacco products	790	61.6%	100.0%	
Total	1283	100.0%	100.0%	

Creating Labels for Outputs, Continued

Table example

Look at the two tables below. The table which contains the value labels instead of numbers is much clearer.

Forward

SMOKE

AgeRange	Currently smokes tobacco products	Does not currently smoke tobacco products	TOTAL
25-34	167	151	318
Row %	52.5	47.5	100.0
Col %	33.9	19.1	24.8
35-44	128	208	336
Row %	38.1	61.9	100.0
Col %	26.0	26.3	26.2
45-54	119	198	317
Row %	37.5	62.5	100.0
Col %	24.1	25.1	24.7
55-64	78	231	309
Row %	25.2	74.8	100.0
Col %	15.8	29.2	24.1
99	1	2	3
Row %	33.3	66.7	100.0
Col %	0.2	0.3	0.2
TOTAL	493	790	1283
Row %	38.4	61.6	100.0
Col %	100.0	100.0	100.0

Forward

T1

AgeRange	1	2	TOTAL
25-34	167	151	318
Row %	52.5	47.5	100.0
Col %	33.9	19.1	24.8
35-44	128	208	336
Row %	38.1	61.9	100.0
Col %	26.0	26.3	26.2
45-54	119	198	317
Row %	37.5	62.5	100.0
Col %	24.1	25.1	24.7
55-64	78	231	309
Row %	25.2	74.8	100.0
Col %	15.8	29.2	24.1
99	1	2	3
Row %	33.3	66.7	100.0
Col %	0.2	0.3	0.2
TOTAL	493	790	1283
Row %	38.4	61.6	100.0
Col %	100.0	100.0	100.0

Part 4: Advanced Statistics

Overview

Introduction   Complex tables produce percentages and confidence intervals that you can use to compare groups and determine if there is a statistically significant difference.

Tables command   The complex tables command in Epi Info will account for the cluster sample design of the STEPS sample and will provide the correct confidence intervals to use for subgroup comparisons.

The tables discussed in Part 3 will produce confidence intervals but not the kind that can be used with STEPS data.

In this part   This part contains the following topics.

Running Complex Sample Tables

Introduction   Complex sample tables are very similar to regular Epi Info tables. Where they differ is that they use Stratum and PSU information to adjust the confidence intervals produced.

This adjustment makes the confidence intervals appropriate for use with STEPS data.

Procedure   Follow the steps below to run a complex sample table.

Step	Action
1	Select "Advanced Statistics, Complex Sample Tables" from the Analysis Tree.
2	Select the variable you want to appear down the left side of the table from the "Exposure Variable" drop down list.
3	Select the variable you want to appear across the top of the table from the "Outcome Variable" drop down list.
4	**Mandatory**: Select "weight" from the Weight drop down list.
5	**Mandatory**: Select "PSU" from the PSU drop down list.
6	**Optional**: Select "Stratum" from the Stratify by drop down list. **Note:** "Stratify by" is used differently in complex tables than in regular tables. It should definitely be used if you have 2 separate sample design variables, identifying 2 different stages of sampling/clustering, included in your data set.
7	Click "OK".

Note   Percentages and confidence intervals are generally reported to one decimal place. Examples and exercises in this part of the training guide are reported to two decimal places to make it easier to compare the correct answers to the output.

Continued on next page

Running Complex Sample Tables, Continued

Exercise 1   Determine if the percentage of participants who currently smoke is different for Males and Females where:
- "Sex" will be the exposure variable.
- "T1" will be the outcome variable.
- "WStep1", "Stratum", and "PSU" will be used..

The TABLES dialog box should look like the figure below.
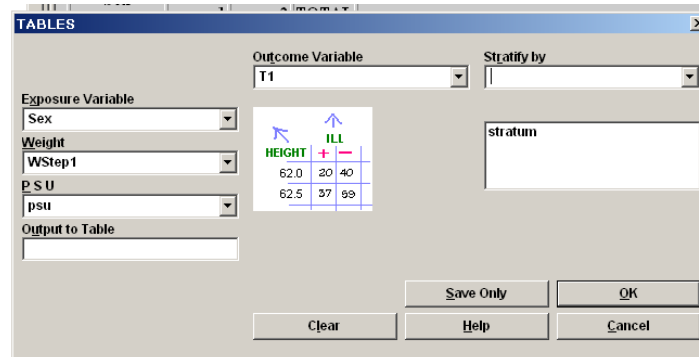
Continued on next page

Running Complex Sample Tables, Continued

Check

Remember that:
- T=1 means that a participant currently smokes.
- The table will contain the weighted percentages for each subgroup in the Row%. Beneath that, the lower 95% confidence limit (LCL) and the upper 95% confidence limits (UCL) are displayed. These are the values used to answer the question about whether the subgroups are different.

Next

Sex	T1		
	1	2	TOTAL
Men	263	365	628
Row %	46.337	53.663	100.000
Col %	60.672	50.267	54.607
SE %	4.078	4.078	
LCL %	36.694	44.020	
UCL %	55.980	63.306	
Design Effect	4.200	4.200	
Women	229	422	651
Row %	36.132	63.868	100.000
Col %	39.328	49.733	45.393
SE %	2.926	2.926	
LCL %	29.213	56.950	
UCL %	43.050	70.787	
Design Effect	2.415	2.415	
TOTAL	492	787	1279
Row %	41.704	58.296	100.000
Col %	100.000	100.000	100.000
SE %	3.384	3.384	
LCL %	33.702	50.293	
UCL %	49.707	66.298	
Design Effect	6.025	6.025	

Q1: What percentage of Female participants smoke?
A1: 36.1%

Q2: What is the 95% Confidence Interval around the female percentage?
A2: (29.2 – 43.0)

Q3: What percentage of Male participants smoke?
Q3: 46.3%

Q4: What is the 95% Confidence Interval around the male percentage?
A4: (37.0 – 56.0)

Q5: Are the percentages significantly different?
A5: No. They are not different because the intervals overlap.

Running Complex Sample Tables, Continued

Exercise 2   Are participants who have had alcohol in the past 12 months more likely to currently smoke?
- A1 will be the exposure variable. A1=1 if the participant has had alcohol in the past 12 months; A1=2 if not.
- T1 will be the outcome variable.

The TABLES dialog should look like:
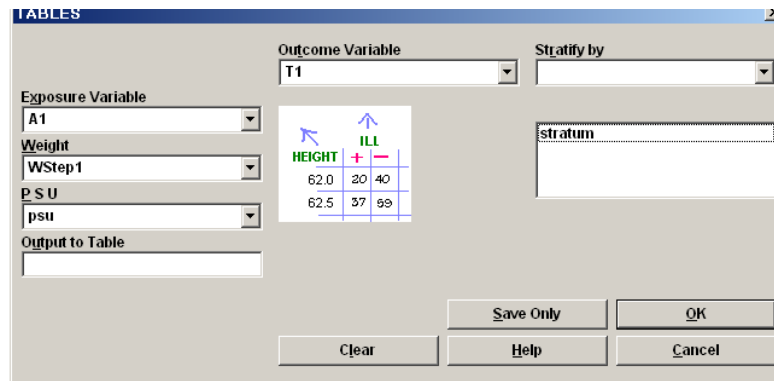
Check   Remember that:
- T1=1 means a participant currently smokes.
- The table will contain the weighted percentages for each subgroup in the Row%. Beneath that, the lower 95% confidence limit (LCL) and the upper 95% confidence limits (UCL) are displayed. These are the values used to answer the question about whether the participants who have had alcohol in the past 12 months are more likely to smoke?

Continued on next page

Running Complex Sample Tables, Continued

Check, contd.   The output should look like the figure below:

A1	T1		
	1	2	TOTAL
1	396	385	781
Row %	53.147	46.853	100.000
Col %	89.084	82.585	85.917
SE %	4.942	4.942	
LCL %	41.461	35.166	
UCL %	64.834	58.539	
Design Effect	7.661	7.661	
2	59	90	149
Row %	39.728	60.272	100.000
Col %	10.916	17.415	14.083
SE %	4.724	4.724	
LCL %	28.558	49.102	
UCL %	50.898	71.442	
Design Effect	1.388	1.388	
TOTAL	455	475	930
Row %	51.257	48.743	100.000
Col %	100.000	100.000	100.000
SE %	3.790	3.790	
LCL %	42.296	39.781	
UCL %	60.219	57.704	
Design Effect	5.347	5.347	

Q1: What percentage of participants drink and smoke?
A1: 53.1%

Q2: What is the 95% Confidence Interval around that percentage?
A2: (41.5 – 64.8)

Q3: What percentage of participants who do not drink but do smoke?
A3: 39.7%

Q4: What is the 95% Confidence Interval around that percentage?
A4: (28.6 – 50.9)

Q5: Are the percentages significantly different?
A5: No. They are not different because the intervals overlap.

Running Complex Sample Means

Introduction Complex sample means are similar to complex sample tables. Both use Stratum and PSU information to adjust the confidence intervals produced. This adjustment makes the confidence intervals appropriate for use with STEPS data.

Procedure Follow the steps below to run a complex sample means.

Step	Action
1	Select "Advanced Statistics, Complex Sample Means" from the Analysis Tree.
2	Select the variable you want to produce means for from the "Means of" drop down list.
3	Select the variable you want to cross-tabulate by from the "Cross-tabulate by Value of" drop down list.
4	**Mandatory**: Select "weight" from the Weight drop down list.
5	**Mandatory**: Select "PSU" from the PSU drop down list.
6	**Optional**: Select "Stratum" from the Stratify by drop down list.
7	Click "OK".

Exercise 1 Calculate the mean BMI for each sex.

- "Sex" will be the "Cross-tabulate by Value of" variable.
- "BMI" will be the "Means of" variable.
- "WStep1", "Stratum", and "PSU" will be used.

The MEANS dialog box should look like the figure below.
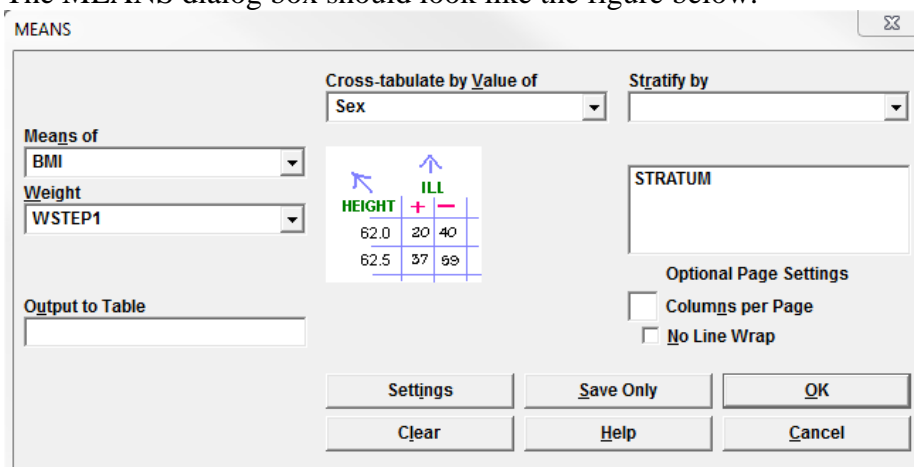
Continued on next page

Running Complex Sample Means, Continued

Check

The output for Means is formatted differently than the output for Tables. Confidence intervals appear to the right of the estimate for each group. Note the minimum and maximum values are also shown for each group.

MEANS BMI Sex STRATAVAR= STRATUM WEIGHTVAR=WSTEP1 PSUVAR=PSU

Next Procedure

Forward

Next

				BMI			
				Confidence Limits		Minimum	Maximum
Sex	Count	Mean	Std Error	Lower	Upper		
Men	582	33.364	0.819	31.427	35.301	17.675	289.283
Women	614	33.145	0.306	32.422	33.867	17.783	68.027
TOTAL	1196	33.263	0.419	32.272	34.254	17.675	289.283
Difference		0.220	0.962	-2.056	2.495		

Q1: What is the mean BMI for males?
A1: 33.4 kg/m^2

Q2: What is the 95% confidence interval for this estimate?
A2: (31.4-35.3)

Using Select Statements

Introduction   Select statements let you temporarily filter a data table. This is useful when you want to include only certain participants in a particular table.

For example, you can temporarily select only male participants for particular analyses or only participants in a particular age range.

Purpose   Using the select statement is particularly useful in combination with the complex sample tables function because it lets you perform stratified analyses not otherwise possible when using the complex sample tables function alone.

The Cancel Select statement is used to cancel the data filter and return the dataset to its original form. Remember that the select statement stays in effect until you explicitly cancel it.

Procedure   Follow the steps below to use a select statement to temporarily filter a data table.

Step	Action
1	Select "Select/If, Select" from the Analysis Tree.
2	Select the variable you want to subset from the "Available Variables" drop down list.
3	Click "=" from the operators buttons.
4	Type the value of the selected variable you want to include in the temporary version of the data table.
5	Click "OK".

Cancel select statements   Follow the steps below to cancel select statements and stop using the current select criteria.

Step	Action
1	Select "Select/If, Cancel Select" from the Analysis Tree.
2	Click "OK".
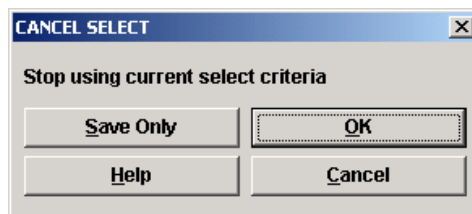
Continued on next page

Using Select Statements, Continued

Exercise 1   Subset the current data table so it includes only male participants.   The SELECT dialog box should look like the figure below.
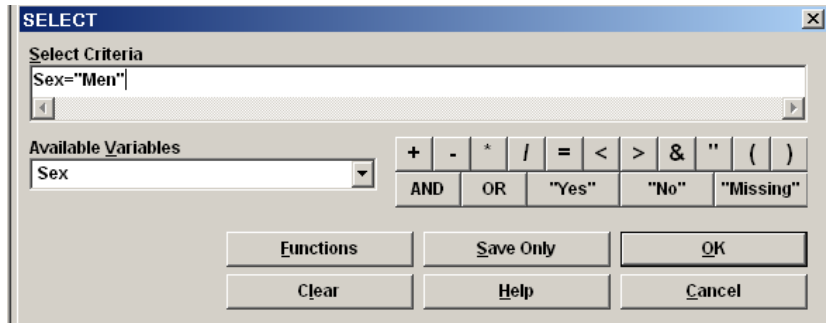
Check   Epi Info displays the new selection criteria and record count as shown in the figure below.
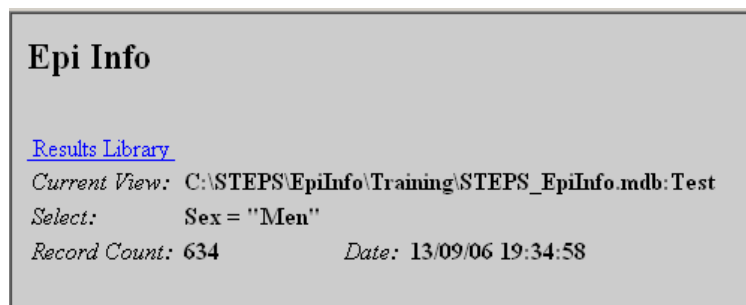
Q: How many participants will be included in subsequent analyses?

A: <u>634</u>

Exercise 2   Are male participants who have had alcohol in the past 12 months more likely to currently smoke? (Hint: this same analysis was run for all participants in a previous exercise.)

The TABLES dialog box should look like the figure below.
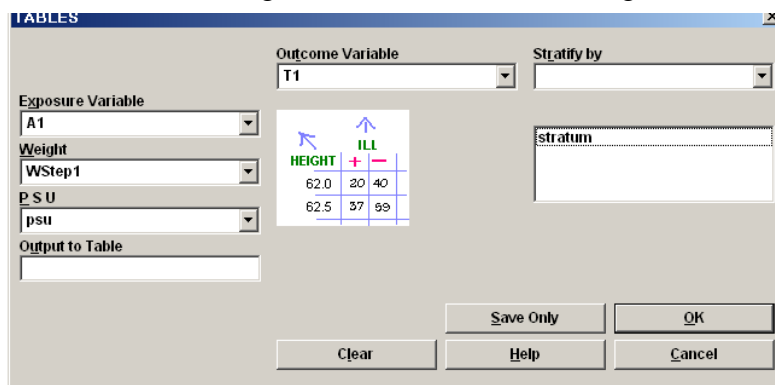
Using Select Statements, Continued

Check   The output table will not mention that a filter is in effect. Remember a table contains only selected participants.

Be aware of the counts of participants at various places to ensure you have a clear idea of which participants are included in which analyses.

The output should look like the figure below.

Next

A1	T1		
	1	**2**	**TOTAL**
1	228	223	451
Row %	53.680	46.320	100.000
Col %	92.567	88.335	90.557
SE %	5.292	5.292	
LCL %	36.838	29.478	
UCL %	70.522	63.162	
Design Effect	5.080	5.080	
2	22	39	61
Row %	41.337	58.663	100.000
Col %	7.433	11.665	9.443
SE %	7.319	7.319	
LCL %	18.043	35.370	
UCL %	64.630	81.957	
Design Effect	1.348	1.348	
TOTAL	250	262	512
Row %	52.515	47.485	100.000
Col %	100.000	100.000	100.000
SE %	4.095	4.095	
LCL %	39.483	34.454	
UCL %	65.546	60.517	
Design Effect	3.443	3.443	

Exercise 3   Cancel the select statement and return to the complete file.
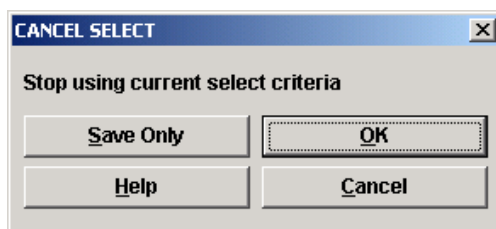
CANCEL SELECT

Stop using current select criteria

Save Only   OK

Help   Cancel

Continued on next page

Using Select Statements, Continued

Check

The output should look like the figure below.
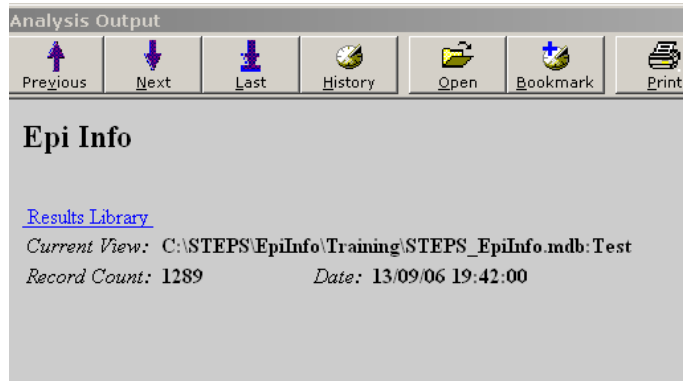
Q. How many records will be included in subsequent analyses now?

A. <u>1289</u>

Part 5: If…Then… Statements

Overview

Introduction   You need to be familiar with how to manipulate and transform data using Epi Info. This section covers If…Then… statements and then shows a practical application of them for cleaning purposes. An If…Then… statement provides more power and specificity during analysis.

Contents   This section contains the following topics:

Topic	See Page
Assigning a Variable Based on 1 Variable	5-2
Assigning a Variable Based on 2 or More Variables	5-6

Assigning a Variable Based on 1 Variable

Introduction   If…Then… statements allow you to move information from one variable to another using a simple logical If… Then…format. If something is true then an action occurs, otherwise (ELSE) something else occurs.

Recode versus If…Then…   Recode and If…Then…statements are used for very different functions.

Recode is used to move the values from one variable into another. This can be used to group values or change values. For example we recoded Age (C3) into AgeRange by grouping the values in C3.

If… Then…. statements are used to define a variable based on the conditions of existing variables.

Assigning MissingSex   If…Then…statements are extremely useful in creating 'cleaning' variables or for determining records that do not meet the established criteria of the dataset. For example, if each record needs to have a sex associated with it you would want to identify which records are missing sex.

To identify which records are missing sex, follow the steps below.

Step	Action
1	Select **Variables, Define** from the Analysis Tree.
2	Define MissingSex.
3	Select **Select/If, If** from the Analysis Tree.

Step 3

Assigning MissingSex (continued)

Step	Action
4	Type in the **If Condition** box the equation meaning "If Sex is missing". To do this, select C1 from the available Variables and then click on the = sign and the **"Missing"** button.
5	Click **Then** to assign the action part of the statement.
6	Select **Assign** from the THEN Block.

Step 4 — Step 5 — Step 6

Step	Action
7	Select MissingSex as the variable to receive the new value and type in "=2" in the **=Expression** box. (This would mean that if Sex is missing, MissingSex would receive a 2).
8	Click **Add**.
9	Select **ELSE** from the If box.

Step 7 — Step 8 — Step 9

Continued on next page

Assigning a Variable Based on 1 Variable, Continued

Assigning MissingSex (continued)

Step	Action
10	Select MissingSex as the variable to receive the new value and type "=1" in the **=Expression** box. (This would mean that if Sex is **not** missing, MissingSex would receive a 1.)
11	Click **Add**.
12	Review the completed If box and click **OK**.

Step 10

Step 11

Step 12

To check your If…Then…statement, try reading the box. This box would read, If Sex (C1) is missing, then assign a 2 to MissingSex. If Sex (C1) is not missing then assign a 1 to MissingSex.
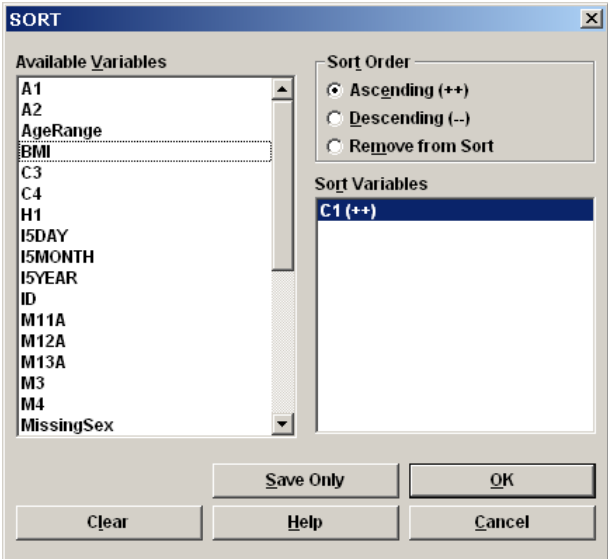
Assigning a Variable Based on 1 Variable, Continued

Check your work
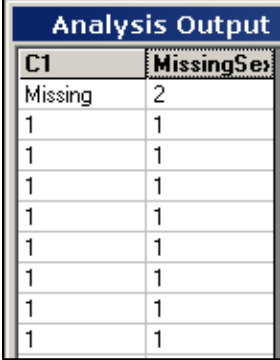
Create a list that is sorted by Sex to check your work. To do this, follow the steps below:

Step	Action
1	Sort the data set by C1 by selecting "Select/If, Sort" from the Analysis Tree. The completed SORT window should look like the figure below.
2	List the variables C1 and MissingSex.
3	Compare the output to the figure below.

Step 1

Step 3

The Program Editor

Begin to familiarize yourself with the code in the Program Editor by looking at the syntax after you complete a command using the Analysis Tree.

The If...Then...statement can also be run using the program editor. The code below reflects all the commands that have been done in this exercise.

```
DEFINE MissingSex
IF C1= (.)  THEN
     ASSIGN MissingSex=2
ELSE
     ASSIGN MissingSex=1
END
Sort C1
LIST  C1 MissingSex GRIDTABLE
```

Assigning a Variable Based on 2 or More Variables

Introduction Unlike the Recode function, If…Then…statements can incorporate the values of two or more variables into a new variable. This is used when the analysis uses more then one question (e.g. Determining daily smokers would incorporate questions T1 and T2).

Combining variables STEPS requires that each participant have an age and sex. In the previous exercise we created a variable MissingSex that marks each record that is missing sex. We could create a similar variable for Age. This would give us two independent variables for records that are missing age or sex.

It would be much more useful if these values could be combined into one value. If…Then…statements allow us to do this.

We are going to create a variable (Valid) that will mark which records have either age or sex missing.

- Valid =1 if the record has both age and sex.
- Valid =2 if the record is missing either age or sex.

Create Valid Follow the steps below to create Valid based on the criteria above.

Step	Action
1	Define Valid
2	Select **Select/If, If** from the Analysis Tree.
3	Type in the **If, Condition** the equation meaning "If Sex (C1) or Age (C3) is missing".
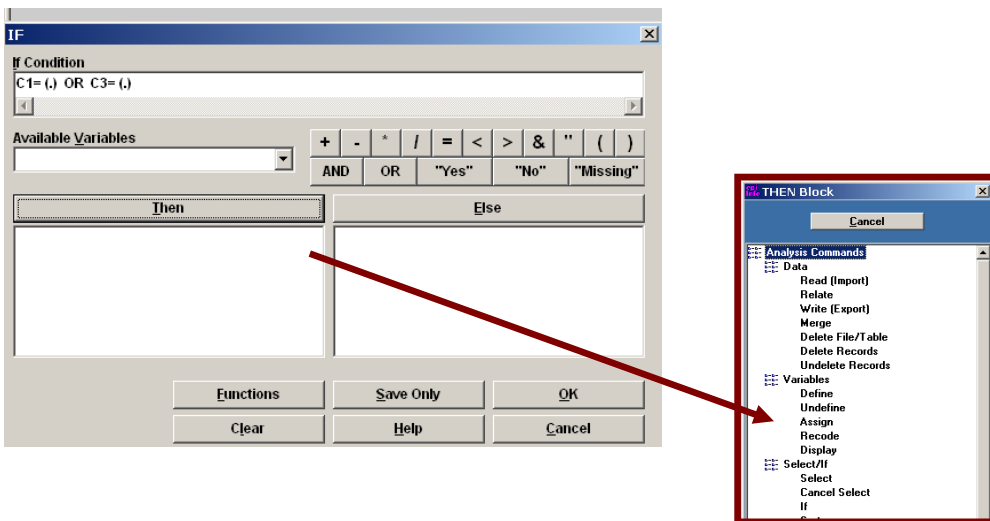
Assigning a Variable Based on 2 or More Variables, Continued

Create Valid
(continued)

Step	Action
4	Select the **Then** button, click Assign, and type in the expression (i.e. what will occur if age or sex is missing: Valid will = 2).
5	Click **Add**.
6	Select the **Else** button, click Assign, and type in the expression (i.e. what will occur if age or sex is not missing: Valid will = 1).
7	Click **Add**.
8	Check the entire expression and click **OK**.

Step 4

Step 6

Step 8

Continued on next page

Assigning a Variable Based on 2 or More Variables, Continued

The Program Editor   The analysis code for the above commands is provided here for your information.

```
DEFINE Valid

IF C1= (.)  OR  C3= (.)  THEN
      ASSIGN Valid=2
ELSE
      ASSIGN Valid=1
END
LIST  C1 C3 Valid GRIDTABLE
```

Check your work   Create a list of C1, C3, and Valid to check to see if the command worked.

Exercise   Create a variable dailydrinker which will be comprised of:

- A1
- A2

Run a table of dailydrinker by AgeRange.

The table below details the criteria for dailydrinker.

Value Label	Criteria
Daily drinker	• Answered Yes to A1. • Answered Yes to A2. • Has a sex and age associated with the record (Valid=1).
Not a daily drinker	Does not fulfill the criteria for daily drinker.

Continued on next page

Assigning a Variable Based on 2 or More Variables, Continued

Check your work

If you completed the exercise correctly you should have the following table.

Forward

DAILYDRINKER

AgeRange	Daily drinker	Not a daily drinker	TOTAL
25-34	18	300	318
Row %	5.7	94.3	100.0
Col %	24.3	24.9	24.8
35-44	17	319	336
Row %	5.1	94.9	100.0
Col %	23.0	26.4	26.2
45-54	21	297	318
Row %	6.6	93.4	100.0
Col %	28.4	24.6	24.8
55-64	18	291	309
Row %	5.8	94.2	100.0
Col %	24.3	24.1	24.1
TOTAL	74	1207	1281
Row %	5.8	94.2	100.0
Col %	100.0	100.0	100.0

Continued on next page

Assigning a Variable Based on 2 or More Variables, Continued

Correct errors

If your tables do not match the tables above:

- Undefine dailydrinker.
- Re-assign the values using the figure below for guidance.
- Re-create the tables using the figure below for guidance.

Part 6: Using the Program Editor

Overview

Introduction
The Program Editor occupies the bottom right portion of the Analysis window.

Purpose
The Program Editor:

- Lets you write, save, and edit programs to run in Epi Info.
- Provides an alternate way to create Analysis commands and can be very useful when you are performing longer and more complex data management and analysis tasks.
- Lets you write a program and then modify it for other uses – eliminating the need to start over for each task.

Examples
Some examples of when you might use the Program Editor are to:

- Create a program that recodes 20 variables and then run that program for five different data tables. You would not have to enter all of the details of each recode over and over for each variable in each data table.

- Write programs and then export the program to text files to share with colleagues.

- Type commands directly into the Program Editor. Or you can use the commands that Epi Info generates automatically for a template and then edit them. This can be a very efficient way to create complex programs.

Optional use
While the Program Editor is very powerful, its use is **optional**. You can use Epi Info very productively without having to use the editor at all.

In this part
This part contains the following topics.

Using the Program Editor

Introduction	The Program Editor window is where you create and edit Epi Info programs. While it functions like any other text-type editor, it has been modified to provide specific Epi Info services. The main modification is that Analysis Tree commands are automatically inserted into the Program Editor where they can be studied and possibly saved and used again.

This manual only covers the main Program Editor functions. |

Program editor window	The Program Editor window generally looks like the figure below. The key components are:

• The name of the program currently being edited is displayed after "Program Editor" in the window title bar. If the program is new and has not been saved before, "New Program" is displayed on the window title bar.

• The main functions are displayed as buttons across the top of the window.

• The menu at the top of the window contains selections to the main functions as well. The menu also contains selections for typical editor functions such as cut and paste, find and replace, print setup, and so on. |
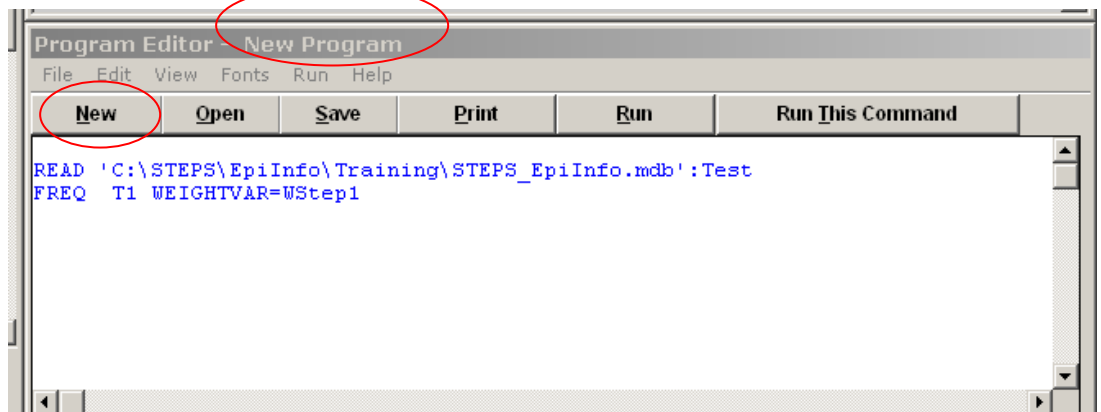
Note: Each of the main functions is described in the pages below.
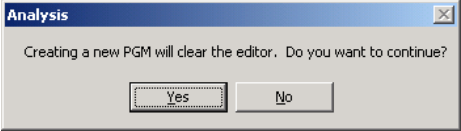
Creating a Blank Program

Introduction   The "New" button at the top of the Program Editor window is used to clear any text from the program window and create a new, blank program.

Procedure   Follow the steps below to create a blank program in the Program Editor.

Step	Action
1	Click the "New" button at the top of the Program Editor window.
2	If the current contents include modifications that have not been saved, Epi Info will warn you and ask you if you want to continue. **Analysis** [×] Creating a new PGM will clear the editor. Do you want to continue? [Yes] [No]
3	<table><tr><td>**Either click…**</td><td>**This will…**</td></tr><tr><td>"Yes" to continue.</td><td>Clear the window. "New Program" will be displayed on the window title bar to indicate that the contents were cleared.</td></tr><tr><td>"No" to stop.</td><td>Put you back in the Program Editor where you can save the program before clearing.</td></tr></table>

Exercise   Clear any text that is currently displayed in the Program Editor.

Check   The Program Editor window should look like the figure below.
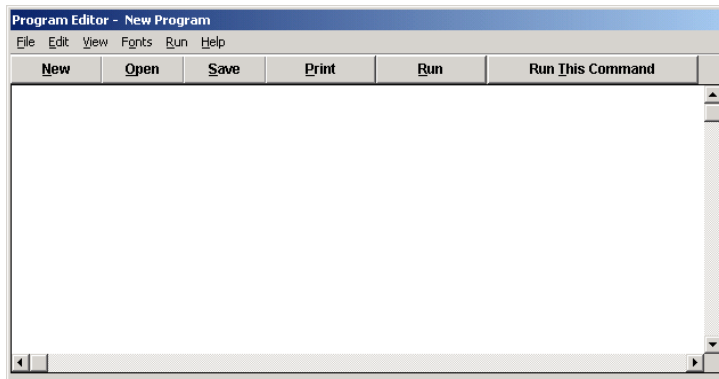
Q: What does the window title bar display as the name of the current program?

A: New Program

Saving a Program

Introduction   Use the "Save" button at the top of the Program Editor window to save the contents of the Program Editor so it can be retrieved and used again later.

Where to save an Epi Info program   There are two places where you can save an Epi Info program. They work very similarly and each is useful for different reasons as shown in the table below.

Save a Program to…	When…
The project file	• You want to use the program only with the current project. • You want to share the data and all programs as a package with someone else. • You do not need to share just the program with anyone else. • You want to use the fastest and easiest way to manage programs.
An individual text file	You want to share the program with someone else so that they can use it or help you with it.

Save to a project file   Follow the steps below to save the contents of the Program Editor window to a project file.
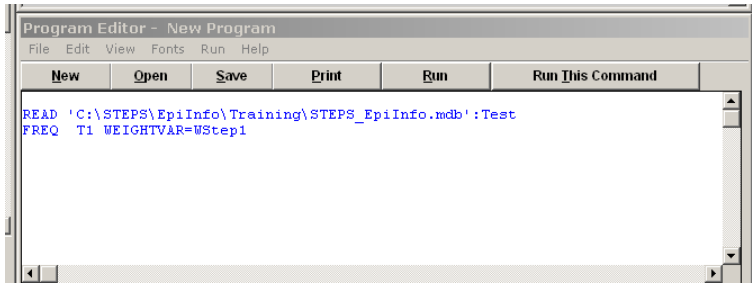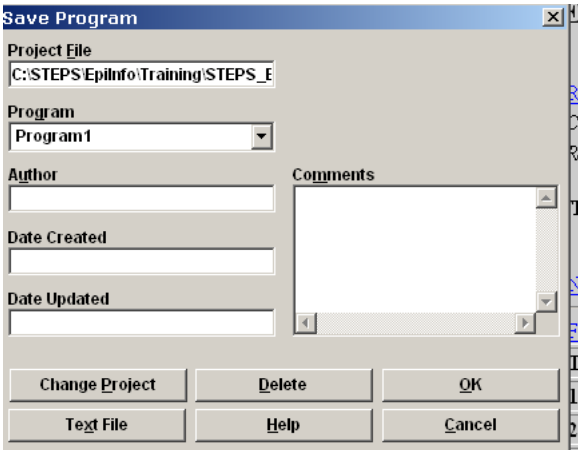
Step	Action
1	Click the "Save" button at the top of the Program Editor window.
2	<table><tr><td>**If the program…**</td><td>**And..**</td><td>**Then…**</td></tr><tr><td>Is new</td><td>Has not been saved before, the "Program" field will be blank.</td><td>Enter the desired program name into the field.</td></tr><tr><td>Had been saved before</td><td>Its previous name is displayed in the "Program" field.</td><td>Keep it as it is or change it to a new program name.</td></tr></table> **Note:** Optionally, you can enter the author's name and any comments.
3	Click "OK".
4	If a program with that name already exists, Epi Info will warn you before continuing. • Click "**Yes**" to overwrite the existing program and save the new one. • Click "**No**" to keep the existing program.

Continued on next page

Saving a Program, Continued

Exercise   Follow the steps below to create a short program and save it to the project file.

Step	Action
1	Clear the Program Editor by using the "New" button.
2	Use the Analysis Tree to open a project and data table and then list the contents of the data table. The program window should look like the figure below. *Program Editor – New Program* File Edit View Fonts Run Help New Open Save Print Run Run This Command READ 'C:\STEPS\EpiInfo\Training\STEPS_EpiInfo.mdb':Test FREQ T1 WEIGHTVAR=WStep1
3	Save the program to "Program1" in the project file. Save Program Project File C:\STEPS\EpiInfo\Training\STEPS_E Program Program1 Author Date Created Date Updated Comments Change Project Delete OK Text File Help Cancel

Continued on next page

Saving a Program, Continued

Check   The Program Editor window title bar should now say "Program1"
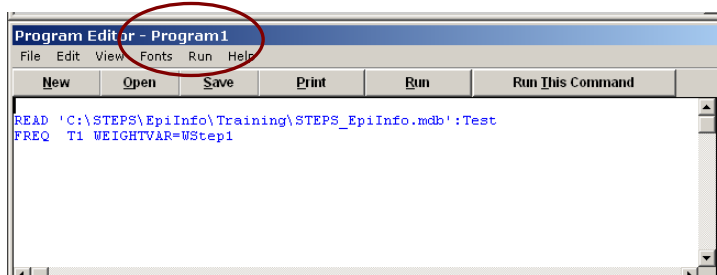
Save to individual text file

Follow the steps below to save the contents of the Program Editor window to an individual text file.

| Step	Action
1	Click the "Save" button at the top of the Program Editor window.
2	Click the "Text File" button at the bottom of the "Save Program" dialog box.
3	Use the "Save Text" dialog box to navigate to the folder where you want to save the program.
4	<table><tr><td>**If the program has…**</td><td>**Then…**</td></tr><tr><td>Been saved before, and you want to overwrite it.</td><td>Select its previous name in the dialog box.</td></tr><tr><td>Not been saved before.</td><td>Enter a new name in the "File name:" field.</td></tr></table>
5	Click "Save" in the "Save Text" dialog.
6	Click "OK" in the "Save Program" dialog box.

Saving a Program, Continued

Exercise   Save the same short program as a text file named "Program2" in the "C:\STEPS\EpiInfo\Training" folder.
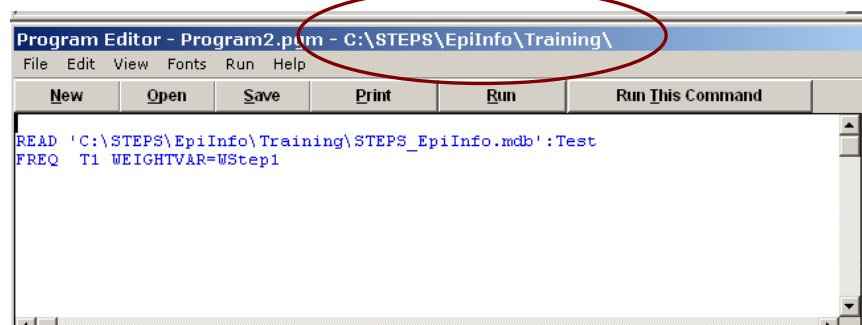
The Save Text dialog should look like the figure below.
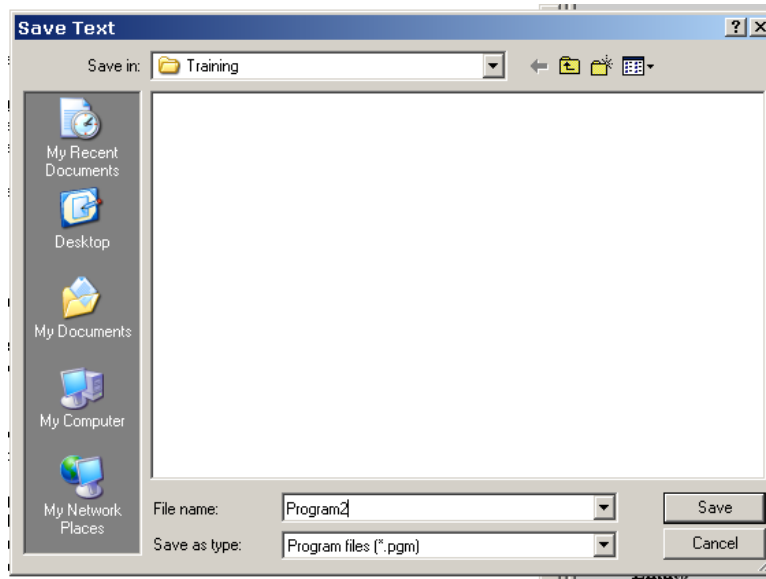
Check   The Program Editor window title bar should now say "Program2.pgm – C:\STEPS\EpiInfo\Training". The folder name indicates that the program is a separate text file and was not saved to the project file.

The Program Editor window should look like:

Retrieving a Program

Introduction   Use the "Open" button at the top of the Program Editor window to retrieve a previously saved program into the Program Editor window.

Programs can be retrieved from a project or individual text files.

Procedure   Follow the steps below to retrieve a previously saved program from the project file.

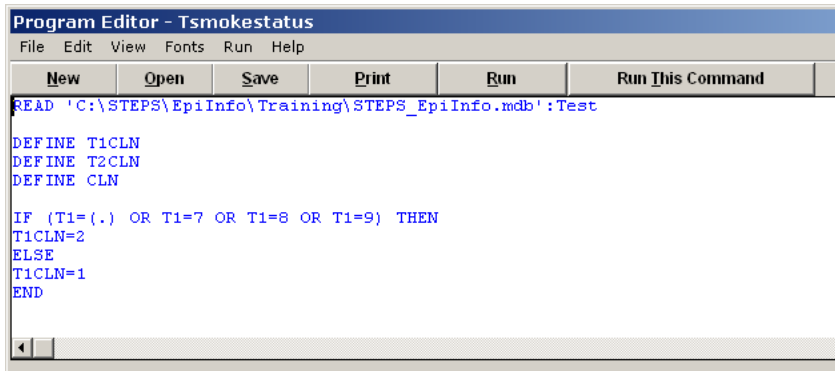| Step	Action
1	Click the "Open" button at the top of the Program Editor window.
2	Select the program that you want to retrieve from the "Program" drop down list.
3	When you select a program name, its creation date and other information will be displayed.
4	Click "OK".

Exercise   Clear the Program Editor and then retrieve "Tsmokestatus" from the project.

Check   The Program Editor window should look like the figure below.

Q: How can you tell if the program is read from the project or from an individual text file?

A: <u>Programs read from a project do not have the folder name displayed in the window title bar. Programs read from individual text files do have the folder name displayed.</u>

Running a Program

Introduction   Use the "Run" button at the top of the Program Editor window to run all of the statements in the program window.

Procedure   Follow the steps below to run a program.

Step	Action
1	Create or open the program that you want to run.
2	Click the "Run" button at the top of the Program Editor window.
3	Review the output in the output window.

Trouble shooting   Sometimes, a program will not run due to typographical errors or logic errors. If this occurs, Epi Info will produce an error message in an effort to help with diagnosing the problem.

Depending on the kind of problem, it may be necessary to leave the Analysis Module completely, restart the program, and try again.
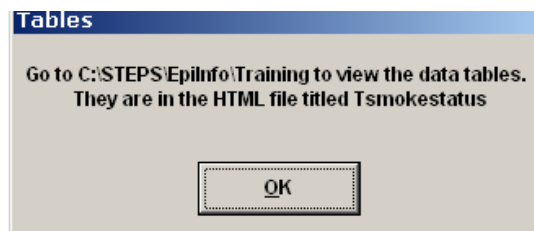
Note: If you exit, make sure you save your program first.

Exercise   Run "Tsmokestatus".

Check   A message box should open identical to the figure below:

> **Tables**
>
> Go to C:\STEPS\EpiInfo\Training to view the data tables.
> They are in the HTML file titled Tsmokestatus
>
> OK

Follow the instructions to see the output from the program. You should find a file called "Tsmokestatus.html" in your "C:\STEPS\EpiInfo\Training" folder. This file contains the output from the program.

Running a Single Command in a Program

Introduction   Use the "Run This Command" button at the top of the Program Editor window to run the single command where the cursor is located.

Procedure   Follow the steps below to run a single command in a program.

Step	Action
1	Create or open the program that contains the command you want to run.
2	Position the cursor on the command you want to run.
3	Click the "Run this Command" button at the top of the Program Editor window.
4	Review the output in the output window.

Trouble shooting   Sometimes Epi Info cannot determine the beginning and endpoint of a command. When this occurs, it produces an error message. Try repositioning the cursor.

Depending on the kind of problem, it may be necessary to leave the Analysis Module completely, restart the program and try again.

Note: If you exit, make sure you save your program first.

Exercise   Open the program "Program1" and then type the command "LIST T1 T2 GRIDTABLE" at the bottom of the program in the Program Editor window and run just that one command.

The Program Editor should look like the figure below.
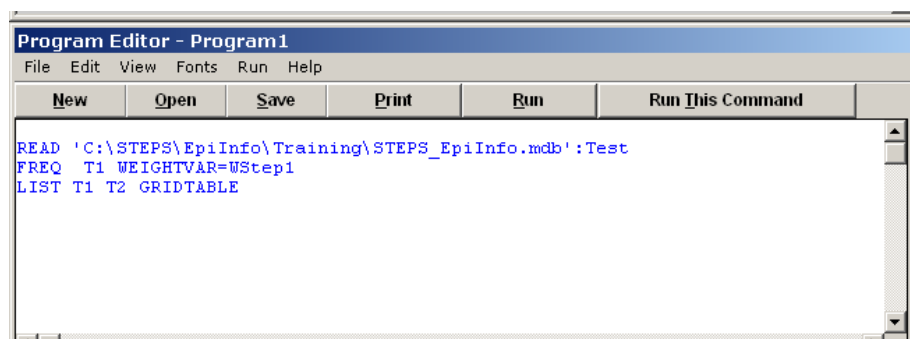
Continued on next page

Running a Single Command in a Program, Continued

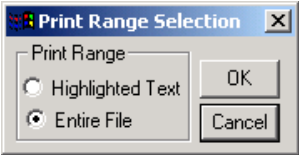**Check**   The output should look like the figure below:

| T1	T2
2	Missing
2	Missing
1	1
1	2
2	Missing
2	Missing
2	Missing
2	Missing
2	Missing
1	2
2	Missing
1	1

Analysis Output

Printing a Program

Introduction Use the "Print" button at the top of the Program Editor window to print the contents of the program window.

Procedure Follow the steps below to print a program.

Step	Action
1	Create or open the program you want to print.
2	If you want to print just part of the program, select or "highlight" that part of the program.
3	Click the "Print" button at the top of the Program Editor window and select the print range from the Print Range Selection window.
4	Click "OK".

Note: Use the "File", "Print Setup" selection from the menu to change printer settings as needed.
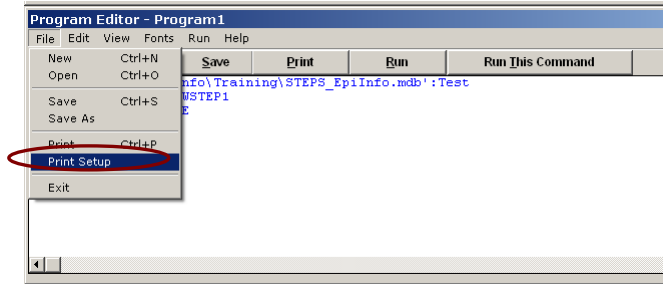
Part 7: Managing Output

Overview

Introduction   Epi Info output is very basic and can be hard to read. Several Epi Info options allow you to improve your Epi Info output's readability and usability

Contents   This part contains the following topics:

Topic	See Page
Using Headers	7-2
Copying Output to Other Software	7-3
Routing Output	7-5
Clearing Output	7-6

Using Headers

Introduction   The Epi Info header command allows you to add titles to different levels of Epi Info Output. The header command sets the font used in the output window and sets titles that appear at the top of windows, programs outputs, and procedure outputs.

Purpose   Setting headers adds useful information to program outputs. This information documents outputs for later reference and makes output easy for others to read and use.

Titles are displayed at various places in the output window. Six different titles are available.

Title	Display Area	Epi Info Command
Window	Output window title bar	HEADER 1 ""
File	Top of the output text (replaces the name of the program file name)	HEADER 2 ""
Data Source	Displayed when the READ command is run	HEADER 3 ""
Procedure	Displayed when a new procedure is run	HEADER 4 ""
Variable	Displayed each time a new variable is run	HEADER 5 ""
Stratum	Displayed each time a new stratum is run	HEADER 6 ""

Note: The "" in the Epi Info Command column are where the specific text is written. The text must be enclosed with quotation marks when written in the Program Editor.

Procedure   Follow the steps below to set a header.

Step	Action
1	Select "Output, Header" from the Analysis Tree.
2	Select the item you want to change from the "Title Line" drop down list.
3	Enter the title you want in the "Title" field.
4	Adjust the font size, attributes and colour.
5	Click "OK".

Copying Output to Other Software

Introduction Epi Info output is generated into an html file. To use the output in another program, such as Excel or Word, you can copy the output and paste it into the selected program.

Purpose Copying output to other software lets you incorporate the output into reports and perform other analyses. You can copy Epi Info output into word processors, spreadsheets and other software packages. Copying output is a safe way to move information from Epi Info into another program because it saves time and eliminates data entry errors.
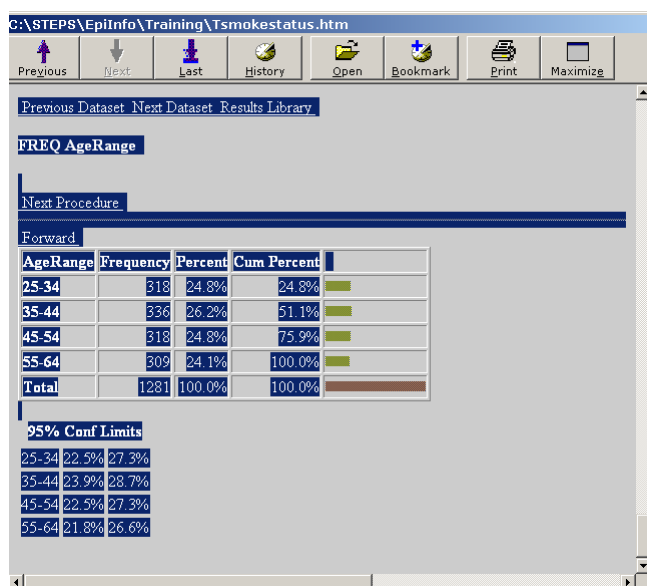
Procedure Copying Epi Info output uses the standard Windows copy/paste functions. Follow the steps below to copy and paste Epi Info output.

Step	Action
1	Highlight the section of output you want to copy and press Ctrl-C to copy the desired portion to the clipboard.
2	Change to your destination software and press Ctrl-V to paste the selection into the new package.
3	Proceed to edit or format in the new package as desired.

Example Copy some of the output from Epi Info into Microsoft Word or another Word Processor.

Copying Output to Other Software, Continued

Check   The output should look close to the example below.

Routing Output

Introduction   By default, Epi Info output is saved to the Epi Info program folder using a standard file naming convention. This generally works well but it can be very beneficial for you to specify exactly where the output for each project should go and the names of the output files. This will help organize your output files and maintain good project documentation.

Purpose   The routeout command is a useful way to organize output files and document your analyses. Routing your files will assign specific file names to output files. Output will go to a specific file until you route to another file or issue the closeout command.

Naming program files and output files in a systematic manner helps you document your analysis and respond to inquiries.

Procedure   Follow the steps below to route output to a specific file.

Step	Action
1	Select "Output, RouteOut" from the Analysis Tree.
2	Enter the folder and name of the file that you want to receive the output. Or also browse to the folder where you want the output file to go.
3	Select "Replace any existing file" if you want new output to replace previous output in the file. If this is not selected, new output will be appended to the end of an existing file.
4	Cancel the routing by using the CLOSEOUT command or by routing the output to another file.

Example   Route the Epi Info output to "C:\STEPS\EpiInfo\Training\Output". Use the ROUTEOUT dialog box to do this.
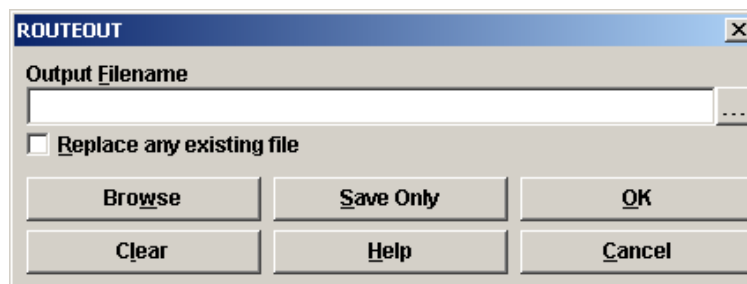
Clearing Output

Introduction   Each time you run an Epi Info command, Epi Info creates output that is stored in an html output file. Over time, a great number of output files will be created and stored on the computer.

Purpose   To keep things organized, it is good practice to delete unneeded output files from time to time. Epi Info uses a command "Storing Output" to manage output files. Use this command to adjust the output settings.

Procedure   Follow the steps below to clear output files from your system.

Step	Action
1	Select "Output, Storing Output" from the Analysis Tree.
2	Click the "Delete" button.
3	Click the "Delete All" button to delete all of the output files.
4	Select one file and then click the "Delete" button to delete just that file. Select "Delete All" to delete all output files.
5	Click "Yes" to confirm the deletion.
6	Click "OK" when you are finished.

Example   Clear all of the output in your Epi Info output folder. The clear output dialog should look like the figure below.
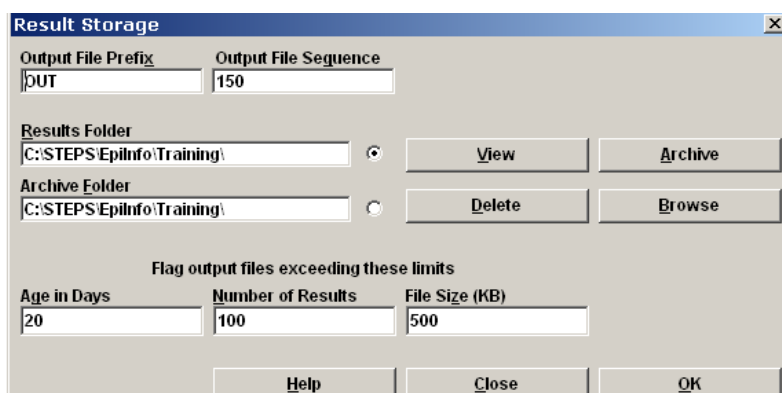
Part 8: Other Commands and Functions

Overview

Introduction   Epi Info provides numerous commands to manage data, analyze data, and control the program environment. A few of the more important ones are discussed in this part.

In this part   This part contains the following topics.

Topic	See Page
Program Help	8-2
Program Comments	8-4
Compacting the Project File	8-5
Backing Up the Project File	8-6

Program Help

Introduction	Program help is available throughout Epi Info. The "Help" button is available almost anywhere and will bring up a window that explains the command and its options.

The help system is indexed and can be searched for key words and terms of interest. Print the help entries for commands you frequently use. |

| **Common commands** | For most commands, help is available on the:

• Syntax (useful when you are writing programs by hand or enhancing programs).

• Generator (explains how to fill out the various fields in dialogs presented from the Analysis Tree).

Switch between the two types of help by clicking the "Command Reference" or "Command Generator Screen" buttons at the top of most help windows. |
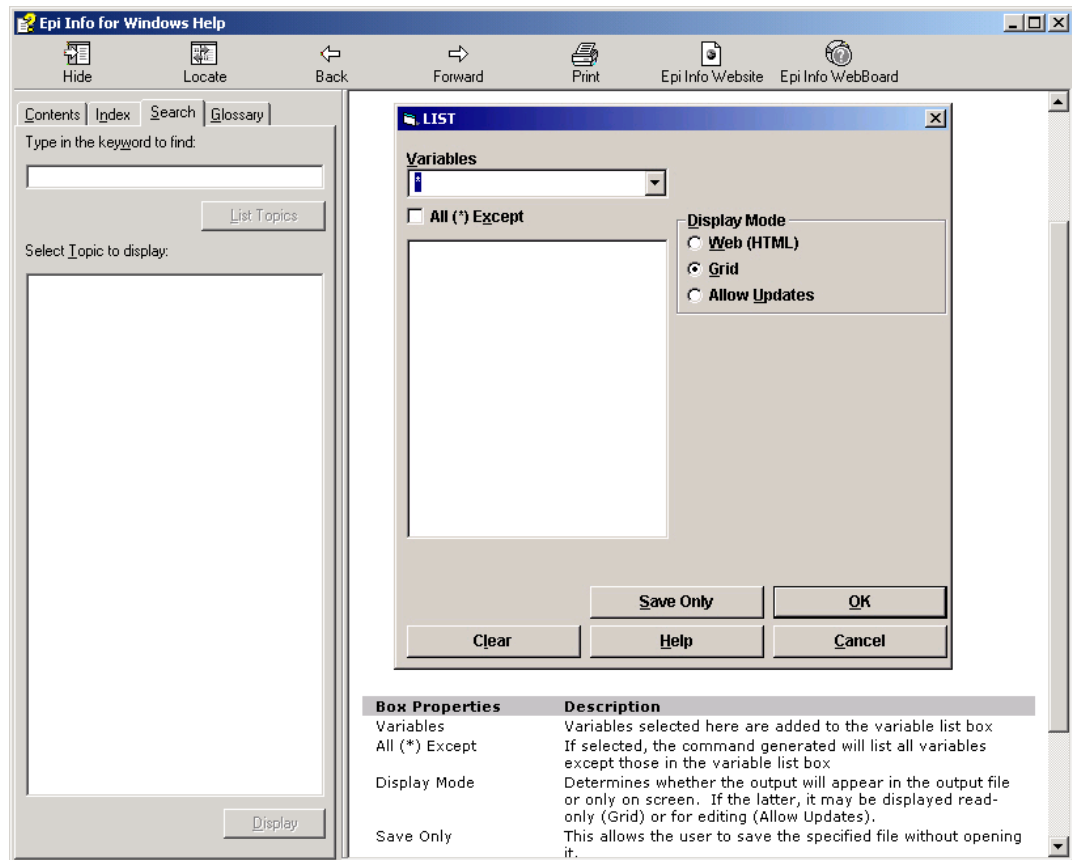
Example 1	Help on the syntax of the LIST command looks like the figure below.
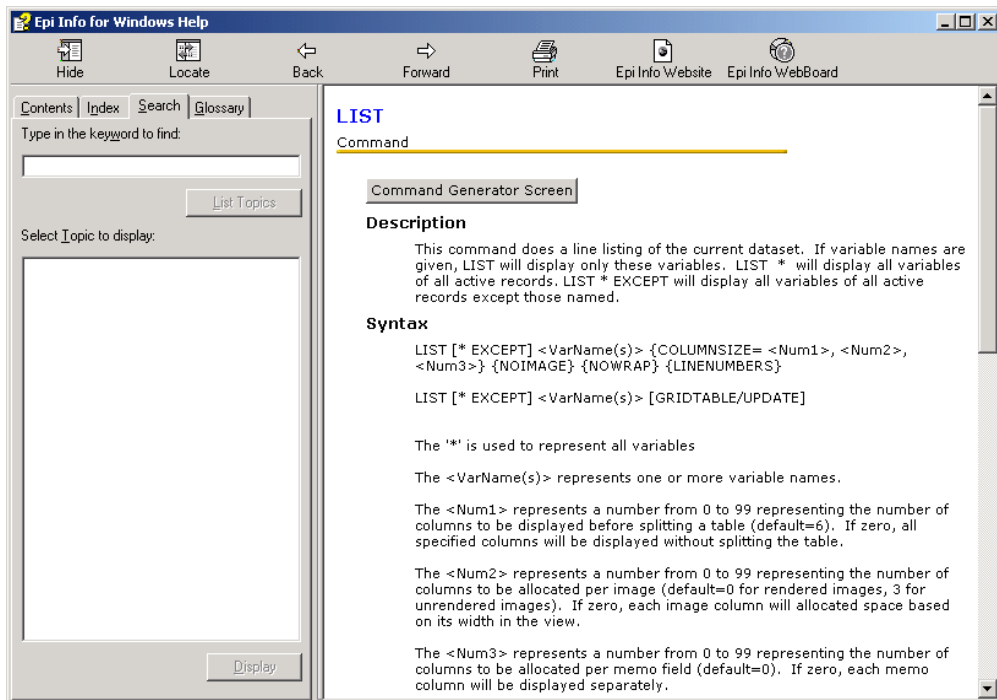
Continued on next page

Program Help, Continued

Example 2   Help on the fields in the list dialog looks like the figure below.

Program Comments

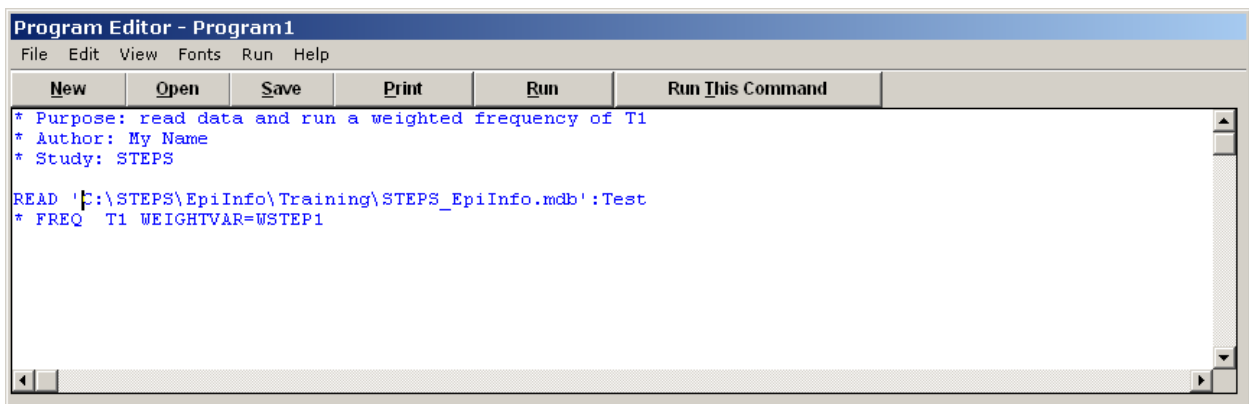|	
Introduction	Program comments are parts of the program used to document or otherwise explain the program but are not executed as Epi Info code.

Procedure	Use the instructions below to create program comments.

To make…	Then…
A line into a comment.	Start the line with the "*" character.
Several lines into a comment block.	Put an "*" character at the start of each line in the comment block.
Part of the Epi Info code into a comment (if you want to temporarily exclude it from executing).	Start the line with the "*" character.

Example	In the example below:

- The first three lines of the following program are program comments and will not be executed when the program is run.
- The last line is an Epi Info command that will not be executed because it has been made into a comment.

Compacting the Project File

Introduction Compacting the project file optimizes the structure of the file, saves disk space, and helps Epi Info run efficiently.

Purpose Over time, running programs, creating variables, and saving new versions of your data causes the project file to grow and become inefficient. Compacting the file is an important management task.

Procedure Follow the steps below to compact your project file.

Step	Action
1	Exit from the Analysis Module.
2	From the main Epi Info screen, select "Utilities, Compact" from the menu at the top of the screen.
3	Browse to the project file you want to compress. Epi Info will offer to compress to the same file, but you can select a new file name if you desire.
4	Click "OK" to compact the file.
5	Click "Exit" to exit from the compact utility.

Example The compact utility dialog should look like the figure below.
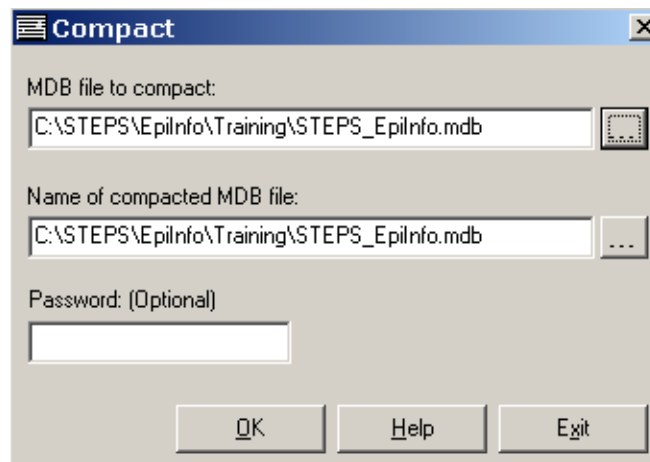
Backing Up the Project File

Introduction   Backing up your project file is an essential part of effective data management.

Purpose   You should back up your project file every day. Backing up ensures that data and programs are preserved in the event of a computer failure. Back up the file to a network drive, a flash drive, or other media and secure it away from the location of your computer.

Guidelines   When backing up your project file, follow the guidelines presented below.

- The procedure for backing up a file varies from system to system depending on the backup media.
- Before you backup your project file, make sure that you have exited from Epi Info completely.
- Do not try to backup a file that is in use.
- Remember to store your backup away from your computer preferably in a locked cabinet or other secure location.

Part 9: PowerPoint Slides

Overview

Introduction   The STEPS Epi Info Training Manual is accompanied by a comprehensive PowerPoint presentation that includes all the information and exercises in this manual. The slides are also available electronically from the STEPS team (steps@who.int).

Part 10: STEPS Epi Info Training Instrument

Overview

Introduction   This Epi Info training manual uses a miniature version of the STEPS Instrument. The Training Instrument uses similar codes and questions as the STEPS Instrument.