Introduction to Epi-Info For Data Management

by
Carol Bigelow

Department of Biostatistics and Epidemiology
School of Public Health and Health Sciences
UMASS
Fall 1998

With Web Site On-Line Help
Prepared by
Jie Chen and Ed Stanek

http://www-unix.oit.umass.edu/~bioepi691/epiinfo/index.html

1. Before You Begin

What is Epi-Info
Epi-Info is a public domain program designed for use by epidemiologists and related professionals. Its biggest advantages are that it is simple to use and that it accommodates in a single software package the following:

       Word processing   Use EPED to type in your data collection instrument here and create data fields.

       Data management   Enter data using the ENTER program. Use CHECK for double data entry, verification, and selected variable checks.

       Analysis   Use ANALYSIS for analysis of data. *Note: Epi-Info is particularly well suited for the stratified analysis of rates.*

What this Unit Covers
This unit is a simple introduction and illustration of the use of Epi-Info for selected data management activities, including the following.

       Creating a data entry program
       Programming data entry error checks
       Carrying out data entry and verification
       Preparing and exporting the data to other platforms (e.g. SAS).

What this Unit Does Not Cover
This unit is not intended to be an overview of the entirety of Epi-Info. For example, Epi-Info has many statistical analysis options. Attention here is focussed on its applications in data management.   Nor does this unit include any discussion of issues pertaining to the design of a survey or any other form of data collection instrument.

Who Should Use This Unit
Users who have or are planning to collect data for subsequent data analysis might use this unit as a tool for creating a numerical data base. Such persons might be public health workers, researchers, etc involved in the management of data sets of moderate size. Note that Epi-Info is not intended for the management of such large data bases as the National Health Interview Survey.

How to Use This Unit

This unit is a tutorial. Its organization is, hopefully, in the order of the steps involved in creating a data base and exploits an example that the user can mimic.   It is recommended that the user new to Epi-Info work through the tutorial following and replicating the examples given. This should take approximately 1-2 hours.

For Further Information on Epi-Info

By mail: USD, Incorporated
                                        2075-A West Park Place
                                        Stone Mountain, Georgia 30087
                                        Tel. 770/469-4098
                                        Fax 770/469-0681

By email: usd@usd-inc.com

From the WEB: http://www.cdc.gov/epo/epi/epiinfo.htm

2. How to Install Epi-Info

Obtain Epi-Info
By mail: USD, Incorporated
                                  2075-A West Park Place
                                  Stone Mountain, Georgia 30087
                                  Tel. 770/469-4098
                                  Fax 770/469-0681

By email: usd@usd-inc.com

From the WEB: http://www.cdc.gov/epo/epi/epiinfo.htm

Installing Epi-Info From Epi-Info Distribution Disks
Reference: Epi-Info Manual, version 6. Pp 35-45.

There are 3 distribution disks for Epi-Info Version 6.04b. The files are in compressed format which means that they cannot be used without running the program INSTALL. Disk #1 has the program INSTALL. It allows you to either install Epi-Info or make copies for further distribution.

_____ Place disk #1 in your floppy drive (usually drive A:)
_____ At the dos prompt, type **install** and press the **enter** key.
_____ Follow the installation instructions.

Installing Epi-Info From the World Wide Web
_____ Visit the site http://www.cdc.gov/epo/epi/epiinfo.htm
_____ Click on **Download software**

How to Run Epi-Info from DOS
_____ Identify the subdirectory containing Epi-Info. This might be **C:\EPI6**
_____ At the dos prompt, type **epi6** and press the **enter** key.

How to Run Epi-Info from Windows
_____ Click twice on the **epi6** icon.

How to Exit Epi-Info
_____ Press the **<F10>** function key. Depending where you are, you may need to do this several times.

3. How to Create a Data Collection Instrument

Reference: Epi-Info Version 6 Manual, pp 51-72.

Epi-Info allows you to create a data collection instrument using the word processing program EPED. This is then used to create a data entry program, again using the word processing program EPED. Data entry is accomplished later using the ENTER program.

Note: In your applications, you may wish to create the data collection instrument in a separate program (e.g. Word or WordPerfect) and use Epi-Info only for the creation of the Epi-Info data entry program itself.

Objective
Create a data collection instrument in Epi-Info from the instrument below. Save it as **church.txt.**

<div>

Questionnaire for Church Supper

Name: _____   ID #: _____

Telephone#: _____

Your age (in years): _____   Sex (M/F): _____

Time of supper: _____   Circle one: AM or PM

Did you become ill? Circle one: YES or NO (if NO, skip to Food History)

      IF Yes, did you experience:

      Diarrhea (Y/N)? _____   If yes, onset date: (MM/DD): _____
                                   Onset time: _____ Circle one: AM or PM

Food History (Circle Y or N):

Baked Ham:	Y	N
Spinach:	Y	N
Mashed Potatoes:	Y	N
Cabbage Salad:	Y	N
Jello:	Y	N
Rolls:	Y	N
Brown Bread:	Y	N
Milk:	Y	N
Water:	Y	N
Ice cream:	Y	N

</div>

Step 1. Get into Epi-Info.
_____   From DOS, type **epi6** and press the **enter** key. Or, from Windows, click on the **EPI6** icon.
_____   From the welcome menu, **PROGRAMS** will be highlighted. Press the **<enter>** key.
_____   From the programs menu, **EPED** will be highlighted. Press the **<enter>** key.
           You will see a blank EPED screen.

Step 2. Some preliminaries (word wrap and margins, etc.)
_____   Check that you are in word wrap mode by pressing the **<CTRL>** key. Look at the lower right
           portion of the screen and confirm that the "mode"listed is WW (for "word wrap"). If it is not,
           press the **<CTRL>** key. **WHILE HOLDING DOWN THE <CTRL> key**,
           press the **<F9>** key one or more times until the "mode" listed is WW.
_____   Turn off the automatic word wrap: Press the **<F6>** key. You'll see a setup menu. Arrow down
           until you have highlighted **Auto word wrap**. Toggle the space bar until the selection reads **OFF**.
_____   Check the margin: Press the **<F6>** key. Arrow down the set up menu until you have highlighted
           Right margin. Press **<enter>**. Type **80**. Press **<enter>**. Press **<esc>**.
Note: There may be some other settings you wish to change. The menu is obtained by pressing <F6>.

Step 3. Type in the instrument and save it.
_____   Type in the data collection instrument as it appears on the previous page.
           Note: If you like, there are options for centering, underlining, bold, etc.
_____   Save the instrument by pressing the **<F9>** key. You will be asked for a file name.
_____   Type in a file name that is <u>8 or fewer</u> characters plus an extension (E.g. survey.txt). Press **<ENTER>**.
_____   You will be asked the following: "Save formatting info? (Not recommended for data files) Y/N"
           Type **Y**.

Guide to Rules and Troubleshooting – Path and File names

Rule: Filenames can be no more than 8 characters.

Troubleshooting: On saving a word processing file to a file within a path …
PROBLEM-
Epi-Info will not accept path names that have spaces in them. For example, I tried to save the questionnaire as
f:\ph691f\independent study\church.txt. The problem is the path called **independent study;** there is a space between the words
"independent" and "study".
Epi-Info created a saved file **f:\ph691f\independ**. Notice that this is an 8 character file name with no extension.
SOLUTION -
There isn't a trick. You'll just have to save your file to a path that doesn't involve any spaces.   **f:\ph691f\epi6txt\church.txt**

Step 4. Exit EPED and return to main program menu.
_____   Press the **<F10>** key to return to the main program menu

Step 5. How to Edit a Word Processing File that You Have Already Created.

_____ From the welcome menu, **PROGRAMS** will be highlighted. Press **<enter>.**

_____ From the programs menu, **EPED** will be highlighted. Press **<enter>.**

_____ Press the **<F2>** key

_____ Press the **<enter>** key. This will choose "**open file this window"**

_____ Backspace 3 times to delete ***.***.

_____ Type in the full name, including drive and path, of the file you wish to edit.

_____ Edit and save as in steps #3 and #4 above.

4. **How to Create a Data Entry Program from an Instrument**

Reference: Epi-Info Version 6 Manual, pp 73-81

Creating a data entry program has a special name in Epi-Info: creating a questionnaire. Such files have a reserved extension, ".QES".   To do this, we will need to learn about data field types in Epi-Info.

Objective
Create a data entry program (an Epi-Info ".QES" file) from the data collection instrument, church.txt, created above. Save it as **church.qes**. Print it.

Step 1. Get into Epi-Info.
_____   From DOS, type **epi6** and press the **enter** key. Or, from Windows, click on the **EPI6** icon.
_____   From the welcome menu, **PROGRAMS** will be highlighted. Press the **<enter>** key.
_____   From the programs menu, **EPED** will be highlighted. Press the **<enter>** key.
        You will see a blank EPED screen.

Step 2. Open the Survey Instrument that You Have Already Created.
_____   Press the **<F2>** key
_____   Press the **<enter>** key. This will choose "**open file this window"**
_____   Backspace 3 times to delete ***.***.
_____   Type in the full name, including drive and path, of the file you wish to edit. E.g. **church.txt**

Step 3. Preliminaries
_____   Turn automatic word wrap option off by pressing the **<F6>** set up key and selecting **OFF**
_____   Get into "QES" mode by pressing the **<F6>** set up key and toggling to select QES

Step 4. Cleanup: Edit this file so as to eliminate text that will not be part of the data entry program
The idea here is to remove any unnecessary text that will not be part of the questionnaire. The goal is to obtain a questionnaire that retains only those items for which data will be entered. For example:

Before edits:
Name: _____   ID: _____
Telephone: _____
Your age in years: _____   Sex (M/F): _____
Time of Supper: _____

After edits:
Name:
Telephone #:
Age: Sex:
Time of Supper (24 hour):

____   Edit.

Step 5. Insert Variable Names
We recommend forcing Epi-Info to use the variable names you want. This is accomplished by bracketing the variable name you want in curly brackets.

Otherwise, Epi-Info will use as the variable name the first 10 characters (excluding spaces, punctuation, and common words like "the" and "a" before the specified field type. For example:

	Variable Name Assigned by Epi-Info
Name: <_____>	NAME
Highest Grade Completed: <##>	HIGHESTGRA
{EDUC} Highest grade complete: <##>	EDUC

For the church supper instrument we might have:

Before:

Name: ID:
Telephone #:
Age: Sex:
Time of Supper (24 hour):

After:

Name ({NAME}):   ID ({SEQID}):
Telephone ({TEL}):
Age ({AGE}):   Sex ({SEX}):
Time of Supper (24 hour) ({TIMESUP}):

Notice that I provided variable names in curly brackets even where it might not have been necessary. This is a nice documentation approach. I also enclosed in parentheses the variable labels for ease of readability later.

Guide to Rules and Troubleshooting – Variable names

Rule: Epi-Info creates variable names for you automatically from the original data collection instrument. To do this, it uses the first 10 non-punctuation characters that are not such discardable words such as "what", "the", etc. For example, if your data collection instrument reads "What is your name?", Epi-Info will name that data entry field ISYOURNAME.

Rule: Variable names must begin with a letter.

Suggestion: Force Epi-Info to use the variable names you want by enclosing them in curly brackets. For example, edit your data entry program (the ".QES" file) to read "{NAME} What is your name?"

Suggestion: Enclose variable labels in curly brackets within parentheses as done in the example above. It seems a bit cumbersome. However, it makes for a more readable data entry program. Example: Student name ({NAME}): _____

Step 6: Insert Variable Field Types
Now you must tell Epi-Info what the field type is of each variable. Following is a table of the field types and their specification.

Field Type	Definition/Notes	Specification	Example
Text	User may input any printable character. Blanks are interpreted as missing data. The maximum length is 80 characters.	Provide a continuous underline that is exactly the number of spaces you wish to allow. For example, if you want a 10 character text field provide 10 underlines.	{Name}: _____
Numeric	User may only enter numbers or spaces; eg. 12 or 41.378. Blanks are interpreted as missing data and will appear as a period (.) in the ANALYSIS program. Decimal points may be included. The maximum length is 14 characters, including one for the decimal point.	Provide a specification that is of the form **#** or **##.#**	{Weight}in kg: ##.#
Upper case	This is similar to a text field except that all entries will be converted to upper case. An example might be state codes; e.g. MA, NY, NH, etc.	Enclose the specification in an <A> or <A >. The length of this field is determined by the number of spaces between the "less than" and "greater than" symbols, including the "A".	{Name}: <A > *This would yield a 10 character length field for last name, converted to upper case.*
Yes/No	User may enter only **Y**, **N**, a space or the enter key.	<Y>	{Ill}: <Y>
Date fields	The user types in the date information **without** the separating slash marks.	<mm/dd/yy> <mm/dd> <dd/mm/yy> <dd/mm>	{DOB}: <mm/dd/yy>
Local phone number	The user types in only the numbers of a local phone number such as 2560721. Epi-Info automatically inserts the dash (e.g. 256-0721).	<phonenum>	{TEL}: <phonenum>
Long distance phone number	The user types in a long distance phone number such as 4132560721. Epi-info inserts the parentheses and dash.	<long distance>	{RES}: <long distance>
Today's date	This is an Epi-Info variable that is created automatically	<today> <today/yy> <today/yyyy>	{VERSION}: <today/yy>
Identification	This is an Epi-Info variable that can be created automatically for you as a record of sequence.	<idnum> <idnum >	{SEQID}: <idnum>

Guide to Inserting Field Types

Using the arrow keys, move the cursor to the location where the field type is to be inserted.

Ctrl-Q Q is a special EPED command for inserting field types into a questionnaire.

A menu of field types will appear. Move the cursor to your selection and press **<enter>**.

Prompts may appear for certain types asking about the length of the field or the number of digits.

After entering all information and responding to any necessary prompts, the field type will be inserted in the questionnaire AT THE CURRENT POSITION OF THE CURSOR. **Remember:** Move to the desired position before inserting the field type.

The text menu **<F4/TXT>** can also be used to access the field types but the **Ctrl-Q Q** command is easier.

For the church supper instrument we might use the following field type instructions:

Before:

```
Name ({NAME}):                          {SEQID}:
Telephone({TEL}) :
Age({AGE}):    Sex({SEX}):
Time of Supper (24 hour) ({TIMESUP}):
```

After:

```
Name ({NAME}): _____  {SEQID}: <IDNUM>
Telephone({TEL}): <LONG DISTANCE>
Age({AGE}): ###     Sex, M or F({SEX}):  <A>
Time of Supper (24 hour) ({TIMESUP}):  ####
```

Step 7. Save the data entry program as a ".QES" file.
_____ Press the **<F9>** key. You will get the message Change filename to church.qes (Y/N)? Type **Y**

Step 8. Print the ".QES" file.
_____ Press the **<F5>** key.

5. How to Insert Automatic Error Checks in Your Data Entry Program

Reference: Epi-Info Manual, version 6. Pp 111-120.

Epi-Info has nice options which allow you to insert automatic error checks, coding, and skip patterns. These are called data entry enhancements and are coded using the CHECK program.   Briefly, CHECK works in the following way:

   Before you run the CHECK program, you create a ".REC" file. In our example, this will be
   an empty data set called church.rec.

   You run the program CHECK to create your automatic error checks, range checks, and skip
   patterns for the particular data entry program of interest. In our example, this will be church.qes.

   CHECK then creates a file having the same first part name but with the extension **".CHK".**
   Thus, in our example, it will be church.chk

   At the time of data entry, when you run the program ENTER, Epi-Info will look for the ".chk" file
   having the same first part name and will implement it as the user performs data entry.

   Note: Creation of a ".CHK" file is optional. The program ENTER will run fine without it.
   However, it is recommended that you create a ".CHK" file.

Objective
Use the CHECK program to insert the following data entry checks in **church.qes**. Save the result in **church.chk.**

Variable	Data Entry Check
Sex	Only 2 values are allowed: "m" or "f"
Age	Age must be in the range: (18,65)

Step 1. Get into Epi-Info.
_____   From DOS, type **epi6** and press the **enter** key. Or, from Windows, click on the **EPI6** icon.
_____   From the welcome menu, **PROGRAMS** will be highlighted. Press the **<enter>** key.

Step 2. Create an empty ".REC" file.
_____   From the programs menu, arrow down and select **ENTER**.
_____   You will be asked Data file. Type **church.rec**. Press **<enter>**. *Note: The extension is not necessary.*
_____   At the prompt Choose one: , enter **2** to create a new data file. Press **<enter>**.
_____   At the prompt Questionnaire file., type **church.qes**. Press **<enter>**.
_____   At the bottom of the screen you will see **Y** highlighted. Press **<enter>**.
_____   Press **<F10/Done>** key Epi-Info will respond with the welcome menu.

Step 3. Run the CHECK program to create your data entry checks.

_____ From the programs menu, arrow down and select **CHECK**.

_____ You will be asked Data file. Type **church.rec**. Press **<enter>**. *Note: The extension is not necessary.*

Epi-Info will respond with a blank data entry form. The cursor will be positioned in the first variable field. In our example, this is the field NAME.

At the bottom of your screen, you will see the following menu of data entry enhancements:

<div style="background-color:#e0e0e0">

CHECK Program Menu of Data Entry Enhancements

F1/F2-Min/Max	**F5-Link fields**	**Legal: F6-Add**	**Shift-F6-Display**	**Ctrl-F6-Delete**
F3-Repeat	**F9-Edit field**	**Jump: F7-Add**	**Shift-F7-Display**	**Ctrl-F7-Delete**
F4-Must Enter	**F10-Quit**	**Codes: F8-Add**	**Shift-F8-Display**	**Ctrl-F8-Delete**

How to Use:
The function keys are toggles.
Pressing the function key once sets the condition. Pressing the function key a second time turns the condition off.
Use the arrow key to position the cursor to the field you wish to enhance.

Example- Setting range of AGE to 18-65, inclusive:
Position the cursor to the field AGE.
Type 18. Press F1.
Type 65. Press F2.
Press enter.

</div>

CHECK Program Data Entry Enhancements

Menu Item	How to Use	Example
F1/F2- Min/Max	Enter minimum value. Press F1. Enter maximum value. Press F2. Press enter.	18 <F1> 65 <F2> <enter>
F3-Repeat	Toggle to set value in current field equal to the same field value in the last record, thus providing automatic repeat of value. Toggle again to turn repeat OFF	
F4-Must Enter	Toggle once to require a valid entry. Toggle again to turn requirement OFF	
F5-Link Fields	Use F5 with F8. F5 allows you to link two fields. F8 performs the automatic coding according to the linked variables. Place cursor in first field. Press <F5> Move cursor to link field. Press <F5> **See also F8.**	Move cursor to field DISEASE. <F5> Move cursor to DISEASCOD <F5>
F9-Edit Field	Use to edit your enhancement. Press <esc> when finished	
F10-Quit	Use to quit CHECK	
Legal: F6-Add	Enter legal value. Press F6. Repeat as needed. Press enter. **See also shift-F6 and ctrl-F6**	M <F6> F <F6> <enter>
Jump: F7-Add	Place cursor in field with condition for jump. Type in the value of the field that will prompt the jump. Press <F7> Now move the cursor to the destination field for the jump. Press <F7> again. Repeat as needed. Press enter. **See also shift-F7 and ctrl-F7**	N <F7>

Continued.

CHECK Program Data Entry Enhancements - continued

Menu Item	How to Use	Example
Codes: F8-Add	Use F5 with F8. F5 allows you to link two fields. F8 performs the automatic coding according to the linked variables. Place cursor in first field. Enter value and press <F8>. The cursor will automatically jump to the linked field. Enter the value that you wish to be automatically entered for the linked field. Press <F8> Repeat as necessary, in both directions **See also F5.**	Position cursor to DISEASE Type HEPA <F8> Cursor moves to DISEASECOD Type 10 <F8>
Shift-F6-Display	Pressing shift-F6 will display the list of legal values	
Shift F7-Display	Pressing shift-F7 will display the jumps	
Shift F8-Display	Pressing shift-F8 will display the linked codes	
Ctrl-F6-Delete	Pressing ctrl-F6 will delete the current legal value.	
Ctrl-F7-Delete	Pressing ctrl-F7 will remove the jump	
Ctrl-F8-Delete	Pressing ctrl-F8 will remove the linked codes	

Step 4. Using the CHECK program, insert your data entry checks.

AGE.

_____   Position the cursor in the field AGE.

_____   Type **18**. Press **<F1>**.   Type **65**. Press **<F2>**.

_____   Press **<enter>**

SEX.

_____   Position the cursor in the field SEX.

_____   Type the legal value **f**. Press **<F6>**. Type the legal value **m**. Press **<F6>**.

_____   Press **<enter>**

Step 5. Save the data entry checks in a ".CHK" File.

_____   Press **<F10/Done>** key

Epi-Info will respond with Write data to disk? [Y/N/esc]?.   Press **Y**.

Epi-info will save the file **church.chk** in the same path and directory containing the file church.rec.

6. Data Entry and Data Entry Validation

Reference: Epi-Info Manual, version 6. Pp 83-96.

So far, we have a word processing document called **church.txt** that is the data collection instrument itself, a data entry program called **church.qes**. We will now enter some data and create a **.rec** file that is the data base containing data. In summary:

Type of File (extension)	Description
.txt	Data collection instrument
.qes	Epi-Info data entry program file
.chk	Epi-Info data entry enhancements file
.rec	Epi-Info data set file

DATA ENTRY.

Objective

Enter the following two case reports from the church supper and save this data in a data set called **church.rec**.   Print it.

Questionnaire for Church Supper

Name: **Ed Stanek**   ID #: **1**

Telephone#: **(206)786-5427**

Your age (in years): **28**   Sex (M/F): **m**

Time of supper: **8:00**   Circle one: AM or **PM**

Questionnaire for Church Supper

Name: **Penny Pekow**   ID #: **2**

Telephone#: **(206)543-6513**

Your age (in years): **31**   Sex (M/F): **F**

Time of supper: **6:30**   Circle one: AM or **PM**

Step 1. Get into Epi-Info.
_____   From DOS, type **epi6** and press the **enter** key. Or, from Windows, click on the **EPI6** icon.
_____   From the welcome menu, **PROGRAMS** will be highlighted. Press the **<enter>** key.
_____   From the programs menu, use the down arrow key to highlight **ENTER**. Press the **<enter>** key.

Step 2. Preliminaries.
_____   You will be asked Data file. Type **church.rec**. Press **<enter>**. *Note: The extension is not necessary.*
_____   At the prompt Choose one: , enter **2** to create a new data file. Press **<enter>**.
_____   At the prompt Questionnaire file., type **church.qes**. Press **<enter>**.
_____   At the bottom of the screen you will see **Y** highlighted. Press **<enter>**.

        Note: If you do not know the name of the ".qes" file, you can see a listing by pressing <F9>.
        From this listing you can highlight the file you want and press the <enter> key.

At this point the questionnaire will be displayed on the screen. You are ready to enter data.

Step 3. Enter Data for Each Individual
_____   You will be positioned at the first cell for entry, corresponding to the first variable, in this case NAME. Type in
        the name and press **<enter>**.
_____   Backspace to make corrections in the current field.
_____   Use the arrow key to move from one field to another.
_____   After the last variable value has been entered you will get the message Write data to disk (Y/N/Esc)? Type **Y**

Epi-Info will respond with a new blank data entry screen. You are ready to enter the next record.

Step 4. Save the Data base in a ".REC" File.
_____   Press **<F10/Done>** key
Epi-Info will respond with the welcome menu.

Step 5. Print
_____   You can print the contents of your **".REC"** file from any text editor or word processing software you wish. Just
be sure to save it as a text file.

For these data, the file **church.rec** is the following:
```
6 1
 NAME             1   3   30   12   3    1   20 112 Name(NAME):
 SEQID           42   3   30   48   3   12    5 112 SEQID:
 TEL              1   4   30   17   4    7   13 112 Telephone(TEL):
 AGE              1   5   30   11   5    0    3 112 Age(AGE):
 SEX             21   5   30   39   5    3    1 112 Sex, M or F(SEX):
 TIMESUP          1   6   30   36   6    0    4 112 Time of supper (24 hour)(TIMESUP):
ed stanek               1(206)786-5427 28M2000!
penny pekow             2(206)543-6513 31F1830!
```

DATA ENTRY VALIDATION.

<u>Objective</u>
Just for fun, suppose the correct data is actually what appears below. (*Note: I have changed Ed Stanek's age from 28 to 38 and Penny Pekow's time of supper to 8:30 pm)* Validate the entry of the following two case reports from the church supper that have been stored in the Epi-Info data set called **church.rec**.   Print it.

Questionnaire for Church Supper

Name: **Ed Stanek**   ID #: **1**

Telephone#: **(206)786-5427**

Your age (in years): **38**   Sex (M/F): **m**

Time of supper: **8:00**   Circle one: AM or **PM**

Questionnaire for Church Supper

Name: **Penny Pekow**   ID #: **2**

Telephone#: **(206)543-6513**

Your age (in years): **31**   Sex (M/F): **F**

Time of supper: **8:30**   Circle one: AM or **PM**

Step 1. Get into Epi-Info.
_____ From DOS, type **epi6** and press the **enter** key. Or, from Windows, click on the **EPI6** icon.
_____ From the welcome menu, **PROGRAMS** will be highlighted. Press the **<enter>** key.
_____ From the programs menu, use the down arrow key to highlight **ENTER**. Press the **<enter>** key.

Step 2. Preliminaries.
_____ You will be asked Data file. Type **church.rec**. Press **<enter>**. *Note: The extension is not necessary.*
_____ At the prompt Choose one: , enter **4** to re-enter and verify records in an existing data file. Press **<enter>**.
_____ At the prompt Name of identifying field, type **SEQID**. Press **<enter>**. This is telling Epi-Info that
           you wish to identify records for double entry by their value on SEQID.
_____ At the OK prompt, press **<enter>**.

Step 3. Double Enter Data for Each Individual

For Ed Stanek
_____ Notice that you are positioned at the identifying field you selected, SEQID, and not at the first field, which is
           NAME. Type **1** and press **<enter>**. Epi-Info accepts this and then positions you at the first field, NAME.
_____ Enter the data for NAME and TEL as you did previously.
_____ Enter the data for AGE, typing in the new value of **38**. Epi-Info will respond with

---

Field Value Mismatch: AGE

You have entered a value which is different from the value originally entered.
Press <esc> to try again or <F9> to compare the old and new values.

_____ Press **<F9>** and press **<enter>.** Epi-Info will respond with

---

Field Value Mismatch: AGE

Original:
<F1> 28

New:
<F2> 38

Press <F1> to keep original value or <F2> to accept new value

Press <F1> or <F2>

_____ Press **<F2>** and press **<enter>.**
_____ Finish entering the data for Ed Stanek.
_____ Epi-Info will then ask you Write data to disk (Y/N/Esc)? Type **Y**

For Penny Pekow
_____ Notice that you are again positioned at the identifying field you selected, SEQID.
      Type **2** and press **<enter>**. Epi-Info accepts this and then positions you at NAME.
_____ Enter the data for NAME, TEL, AGE, and SEX.
_____ Enter the data for TIMESUP typing in the new value of **2030**. When Epi-Info responds with
      a notation of a mismatch, make the correction as you did for Ed Stanek.
_____ After the last variable value has been entered you will get the message Write data to disk (Y/N/Esc)? Type **Y**

Step 4. Save the Data Base in a ".REC" File.
_____ Press **<F10/Done>** key
Epi-Info will respond with the welcome menu.

Step 5. Print
_____ You can print the contents of your **".REC"** file from any text editor or word processing software you wish. Just be sure to save it as a text file.

7. Importing Data into Epi-Info

Reference: Epi-Info Manual, version 6. Pp 197-201.

Importing an ASCII data set into Epi-Info is not difficult but it is a little awkward because it requires the preliminary steps of creating a ".QES" file and a ".REC" file. There are three sets of instruction:

Objective
Import the ASCII data set VDT.dat into Epi Info and create an Epi Info system data set called **vdt.rec**.

1) **Creating an Epi Info questionnaire** - This defines the structure of what will eventually be the Epi Info system data set. We will create one called **vdt.qes**

   2) **Creating an Empty Epi Info system data set** - This provides a destination for the data. We will create one called **vdt.rec.** It will be empty at first.

   3) **Importing the ASCII data** - This is the actual step of importing the data and saving its new version as an Epi Info system data set. This will be **vdt.rec** with the data supplied.

STEP 1. Create an Epi Info Questionnaire.

Before you begin, you need to know the format structure of the data in vdt.dat:

vdt.dat

Columns	Length	Variable	Description
1-6	6	AVGVDT	average hours vdt in 1st trimester
8-9	3	NUMCIGS	number of cigarettes per day
11	2	PRIORSAB	prior spontaneous abortion (1=yes, 0=no)
13	2	SAB	spontaneous abortion (1=yes, 0=no)
15	2	SMOKSTAT	smoker (1=yes, 0=no)
17	2	PRTHYR	prior thyroid condition (1=yes, 0=no)
19	2	VDTEXPOS	vdt exposure (1=yes, 0=no)

1) Copy vdt.dat into the directory containing Epi-Info
2) Get into Epi Info by clicking on the **epi6** icon**.**
3) Select PROGRAMS and press **<enter>.**
4) Select EPED WORD PROCESSOR and press **<enter>**.

   Now you must type your variable names, one at a time.
   It's a good idea to type them on separate lines.
   Follow each variable name with a string of # (e.g. ## or ### or #####).
   The total number of #'s should be EXACTLY the same length as the number of columns
   *   plus the number of leading blanks.*
   This total is provided for you in the column labeled "length" in the table above. Thus,

5) Type the following **exactly** as it is shown here:

AVGVDT ######   **<enter>**   *Here, the number of #'s is 6*
NUMCIGS ###   **<enter>**   *Here, the number of #'s is 2+1*
PRIORSAB ##   **<enter>**   *All of the remaining underlines have length 1 + 1.*
SAB ##   **<enter>**
SMOKSTAT ##   **<enter>**
PRTHYR ##   **<enter>**
VDTEXPOS ##   **<enter>**

6) Save by pressing **<F9>.** Then type **vdt.qes <enter>.**

STEP 2. Create an Empty Epi Info System Data Set

7) Return to main menu by pressing **<F10>**
8) Select PROGRAMS and press **<enter>**
9) Select ENTER DATA and press **<enter>** You might hear a beep.
10) Type **vdt.rec <enter>**.
11) The cursor will be positioned at CHOOSE ONE:. Type **2 <enter>**
12) Type **vdt.qes <enter>**
13) The cursor will be positioned at OK. Press **<enter>**.
14) You will now see a blank questionnaire. Press **<F10> <enter>**

STEP 3. Importing the ASCII Data Set

15) From the welcome menu, press **<enter>**.
16) From PROGRAM menu, select IMPORT **<enter>**
17) An IMPORT screen will appear. INPUT FORMAT will have already
    been selected. Press **<enter>**.
18) At the INPUT FILE NAME prompt, type **vdt.dat <enter>.**
19) At the OUTPUT FILE NAME prompt, the file vdt.REC will be
    suggested automatically. Press **<enter>.**
20) The cursor will be positioned at OK.

 You should get a message saying that 883 records were imported.

Press **<enter>.**
21) To return to welcome menu, press **<esc>.**
22) Press **<F10>** to close Epi-Info

8. Exporting Data From Epi-Info

Reference: Epi-Info Manual, version 6. Pp 191-201.

How nice! Epi-Info has several export formats. This means that if you have a favorite statistical analysis program that you like to use, it is possible to write out your Epi-Info ".REC" data set to a new data set on another platform. Following are the choices possible in Epi-Info Version 6:

Export Platforms Available in Epi-Info Version 6[a]

Menu Choice	Extension	Program
Systat	.sys	Systat (SYSTAT, Inc., Evanston, Il)
SAS	.sas	SAS (Statistical Analysis System, SAS Institute, Inc., Cary NC)
Delim	.sdf	Comma and quote deliminated. All fields are delimited by quotes and separated by commas. The first three records contain the field lengths, data types, and field names.
Lotus	.wks	Lotus 1-2-3. Lotus Development Corp., Cambridge MA
SPSS-X	.sps	SPSS (Statistical Package for the Social Sciences, SPSS Inc., Chicago, Il. This format is for the mainframe version of SPSS. See SPSS/PC below.
Epi-Stat	.epi	EPISTAT. A statistical package for epidemiologists (Epistat Services, Richardson, TX. This format is for the original shareware version of the program. True Epistat, the commercial version, ca be used with the fixed field format described below.
Dbase II	.dbf	Dbase II (Borland International, Scotts Valley, CA)
Dbase III	.dbf	Dbase III and III Plus (Borland International, Scotts Valley, CA)
Dbase IV	.dbf	Dbase IV (Borland International, Scotts Valley, CA)
BASIC	.dat	Fields are delimited by double quotation marks and separated by commas. This one is identical to DELIM but without the field name and length information.
SPSS/PC	.sps	The microcomputer version of SPSS. See above.
Fixed	.car	Fixed field or card format. A popular format for information exchange, especially on mainframe computers.
Statpac	None	StatPac (Wolonick Associates, Minneapolis, MN) a statistical package for MS-DOS computers. The codebook has a "C" appended to the file name and the data file name ends with a "D".
MULTLR	.dat OR .dic	MULTLR, COXPH, and KKMSURV. Public domain programs for logistic regression, Cox proportional hazards, and Kaplan Meier Survival analysis by Nelson Campos and Eduardo Franco. The programs are available from the Epidemiology Monitor, Marietta, Georgia USA
Egret	.bdf	Egret. A commercial system for advanced epidemiologic statistics by Statistics and Epidemiology Research Corporation, Seattle, Washington, USA

[a] Source: Epi-Info Manual, version 6. Pp 192-3.

Exporting Epi-Info Data to the SAS Platform Using the Epi-Info Program EXPORT
Here we discuss the export of an Epi-Info data set (".REC") to the SAS platform using the church supper example. Refer to the Epi-Info manual for discussion of the specifics of exports to other platforms. Briefly, the EXPORT program in Epi-Info does the following:

      Prompts user for name of the Epi-Info data set (".REC"). In our example, this will be church.rec.
      Prompts user for resulting file type. We will select **SAS.**
      Epi-Info will then produce a SAS program that contains a data statement, input statement, cards
         statement, the data, and ending with a semi-colon. In our example, this will be church.sas

      It is recommended that you copy church.sas to a second file called church.dta and then edit church.sas
                   Church.dta – Will contain data only
                    Church.sas – Will contain SAS instructions only

Objective
Using the Epi-Info data set **church.rec**, create a SAS program called **church.sas** which creates a SAS data set called **church.sd2**

Step 1. Get into Epi-Info.
_____   From DOS, type **epi6** and press the **enter** key. Or, from Windows, click on the **EPI6** icon.
_____   From the welcome menu, **PROGRAMS** will be highlighted. Press the **<enter>** key.
_____   From the programs menu, use the down arrow key to highlight **EXPORT**. Press the **<enter>** key.

Step 2. Let Epi-Info Create its ".SAS" Code Program for You. This will be church.sas.
_____   You will be asked Input File Name. Type **church.rec**. Press **<enter>**. *Note: The extension is not necessary.*

      Epi-Info then gives you a menu of output format choices. This box looks like the following:

( ) SYSTAT	( ) DBASE2	( ) STATPAC
( ) SAS	( ) DBASE3	( ) MULTLR
( ) DELIM	( ) DBASE4	( ) EGRET
( ) LOTUS	( ) BASIC	( ) XBASE
( ) SPSS-X	( ) SPSS-PC	
( ) EPI-STAT	( ) FIXED	

_____   Using the arrow key, position the cursor to the SAS selection.   Press **<enter>**.
_____   At the prompt Output File Name., Epi-Info will have provided **church.sas**. Press **<enter>**.
_____   At the bottom of the screen you will see **OK** highlighted. Press **<enter>**.

Epi-Info will respond with the following:

> Information
>
> Exported 2 records from
> Church.rec to
> Church.sas

_____ At the bottom of the screen you will see **OK** highlighted. Press **<enter>**.
_____ You will be returned to the EXPORT menu. To return to the welcome menu, press **<F10>**.

This is what Epi-Info will have created for you.

Epi-Info Export File **Church.sas**

```
DATA CHURCH;

INFILE CARDS PAD;
INPUT
NAME $1-20 SEQID 21-25 TEL $26-38 AGE 39-41 SEX $42 TIMESUP 43-46 ;

CARDS4;
ED STANEK            1    (206)786-5427 38M2000
PENNY PEKOW          2    (206)543-6513 31F2030
;;;;
```

Step 2. Copy church.sas to a new file called church.dta. Edit so that it contains only the data.
Church.dta should look like the following:

```
ED STANEK            1    (206)786-5427 38M2000
PENNY PEKOW          2    (206)543-6513 31F2030
```

Step 3. Edit church.sas so that it contains only SAS instruction. Include also key information such as variable labels, value labels, etc.
Church.sas might look something like the following:

```
*_____;
*
*  Project:    PH691F - Practical Data Management
*     code:    church.sas
*    progr:    carol bigelow
*    title:    Creation of church.sas from Epi-Info church.rec
*
*    input:    f:\ph691f\church.dta
*   output:    f:\ph691f\church.sd2
*_____;
libname ph691f "f:\ph691f";

proc format;
  value $sex 'f'='f=female'
             'm'='m=male';
run;

data ph691f.church;
    infile "f:\ph691f\church.dta" missover;

input
NAME $1-20 SEQID 21-25 TEL $26-38 AGE 39-41 SEX $42 TIMESUP 43-46 ;

label tel='Tel: Telephone'
      age='AGE: age (years)'
   timesup='TIMESUP: Time of Supper';

format sex $SEXF.;
run;

proc contents data=ph691f.church;
  title 'Code:  church.sas   by:  Carol Bigelow';
  title2 'Creation of church.sd2';
run;
```