

Data Transformations and Other Procedures

In the context of data analysis, certain operations often need to be performed on one or more variables prior to analyzing your data. For example, one commonly required operation is to add a number of variables together in SPSS to produce a total score. In this appendix, a number of commonly used procedures in SPSS will be illustrated, including the *Recode*, *Compute*, *Select Cases*, and *Split File* procedures.

To illustrate each of these procedures, the data presented in Figure A.1 will be used (the data file will be opened from your computer shortly). The data consist of values from 10 participants on the following four variables: **gender** (1 = “male,” 2 = “female”), **meaning1**, **meaning2**, and **meaning3**. The last three variables are items from a scale measuring meaning in life.

Participant	Gender	Meaning1	Meaning2	Meaning3
1	1	4	2	5
2	1	1	4	2
3	1	5	1	5
4	1	3	3	3
5	1	4	2	4
6	2	4	1	5
7	2	5	2	4
8	2	5	1	5
9	2	1	4	1
10	2	5	1	5

Figure A.1 The sample data. (Note: The participant variable is included for illustration but will not be entered into the data file.)

The items on the meaning in life scale are presented in Figure A.2.

The Recode Procedure

For the items on the meaning in life scale shown in Figure A.2, the participants read each statement and choose a response—ranging from strongly disagree to strongly agree—that best characterizes their feelings toward the statement. Reading through each of the items, notice that **meaning1** and **meaning3** are written so that a person with a very high degree of meaning in life would most likely answer “strongly agree” to the items (they would circle a “5” for each item), while **meaning2** is written so that this same person would most likely answer “strongly disagree” to the item (they would circle a “1”). **Meaning2** is an example of a *negative* item, where a person high on the characteristic of interest (i.e., meaning in life) answers in the *opposite* direction of the scale. Therefore, in this

Item	Items on the Meaning in Life Scale	Strongly Disagree	Disagree	Neither Disagree or Agree	Agree	Strongly Agree
meaning1	I feel good about the direction of my life.	1	2	3	4	5
meaning2	My life has little meaning.	1	2	3	4	5
meaning3	In general, I feel like I'm on track in my life.	1	2	3	4	5

Figure A.2 Items on the meaning in life scale. For each item, the response options are 1 (strongly disagree), 2 (disagree), 3 (neither disagree or agree), 4 (agree) and 5 (strongly agree). Each participant was instructed to choose only one response option for each question.

example, a person with a very high degree of meaning in life would most likely answer items **meaning1**, **meaning2**, and **meaning3** with values of 5, 1, and 5, respectively. Typically, when analyzing the responses on the scale, a total score is calculated by adding together the responses on the items. In this example, the person with a very high degree of meaning in life would have a total score of 11 ($5 + 1 + 5$). However, because the scale is ultimately designed so that higher scores on the scale correspond to higher levels of meaning in life, a person with a very high degree of meaning in life *should* have a score of 15 ($5 + 5 + 5$). Since **meaning2** is a negative item, however, this isn't the case. To correct for this, a process called reverse coding is conducted where the responses to the negative item on the scale (**meaning2**) are literally reversed, so that a 1 becomes a 5, a 2 becomes a 4, a 3 becomes a 3, a 4 becomes a 2, and a 5 becomes a 1. With **meaning2** reverse coded, instead of having a total score of 11 ($5 + 1 + 5$), a person with a high degree of meaning in life would have a total score of 15 ($5 + 5 + 5$), which is consistent with the intent of the scale. With the rationale for reverse coding provided, let's recode the negative item **meaning2** in SPSS.

First we'll open the data file. (*Note:* The following instructions assume that the *SPSS Demystified* files have been downloaded from the web site to your computer. If the files are not on your computer, they may be obtained from the web at www.prenhall.com/yockey.)

To open the file from your computer

1. Locate the *SPSS Demystified* files on your computer. In the *Appendix A* folder, open the file *Appendix A.sav*.

To recode the item meaning2 in SPSS

1. From the menu bar, select **Transform > Recode into Different Variables ...** (see Figure A.3). (*Note:* If you are using SPSS version 14.0 or below, the menu commands are **Transform > Recode > Into different variables**. All other commands are the same.)
2. The *Recode into Different Variables* dialog box opens (see Figure A.4).
3. Select **meaning2** and click the right-arrow button (►) to move it to the *Input Variable → Output Variable* box (the box reads *Numeric Variable → Output Variable* once **meaning2** is moved into it). Under *Name*, enter the name **meaning2_recode** (see Figure A.5). This is the name of the new variable that will contain the reverse coded values.
4. Click **Change**. You should see *meaning2 → meaning2_recode* in the *Numeric Variable → Output Variable* box.

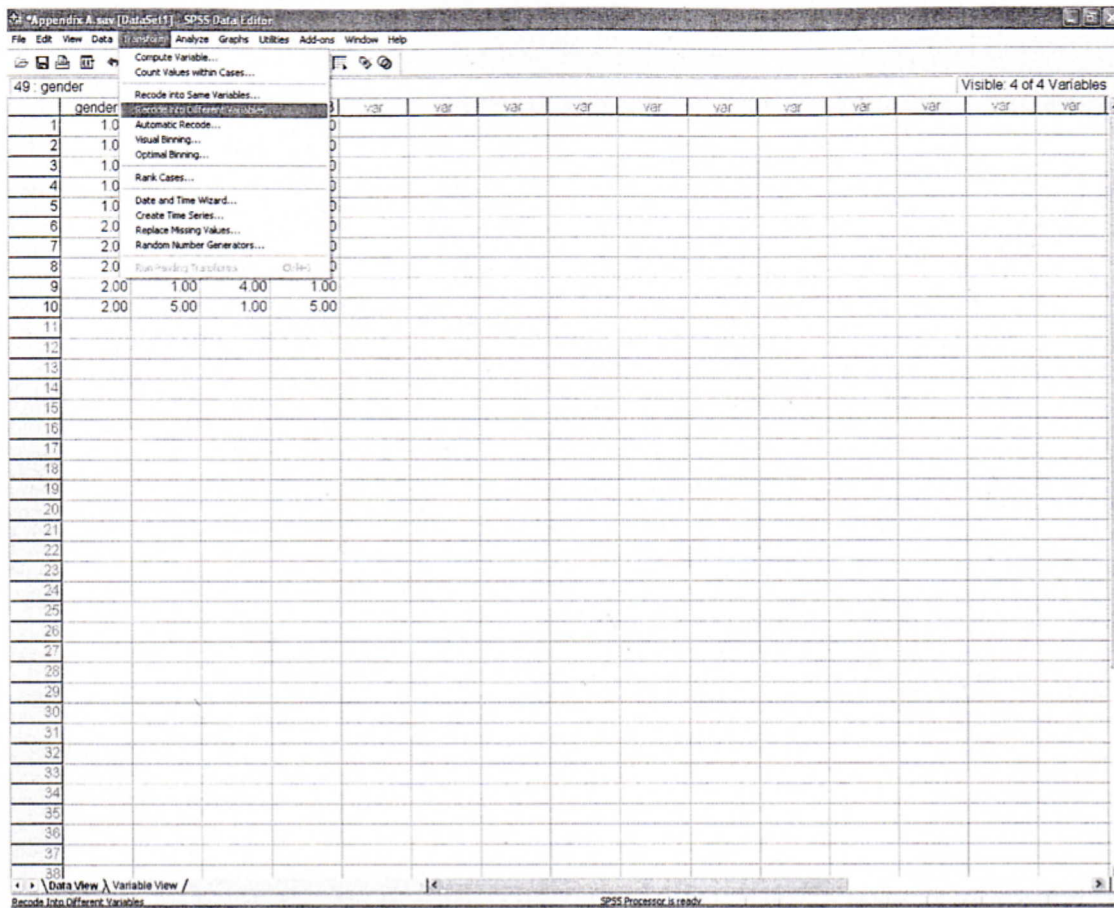


Figure A.3 Menu commands for the *Recode (Into Different Variables)* procedure.

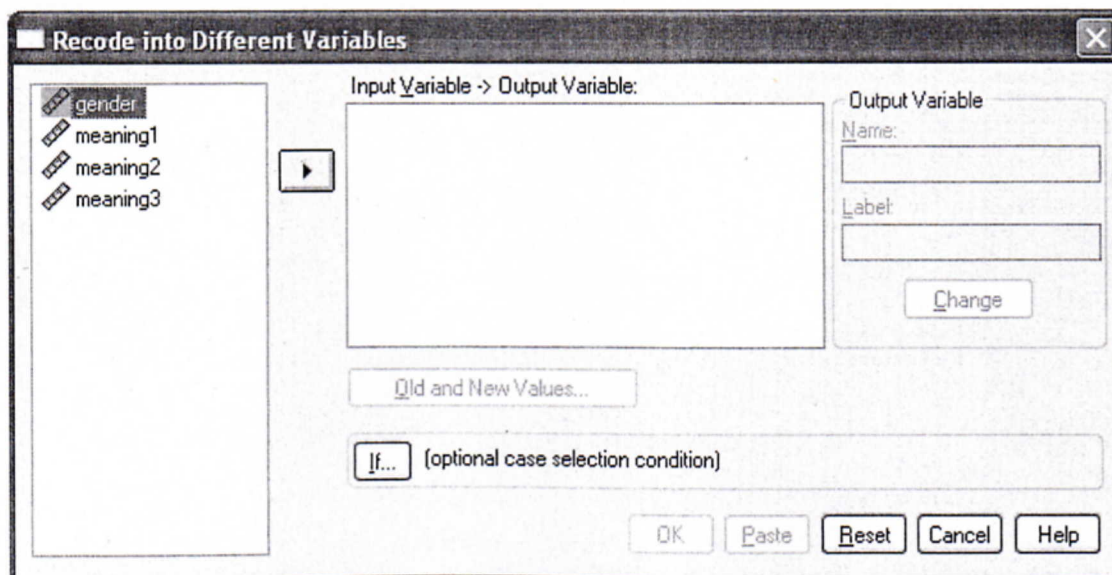


Figure A.4 The *Recode into Different Variables* dialog box.

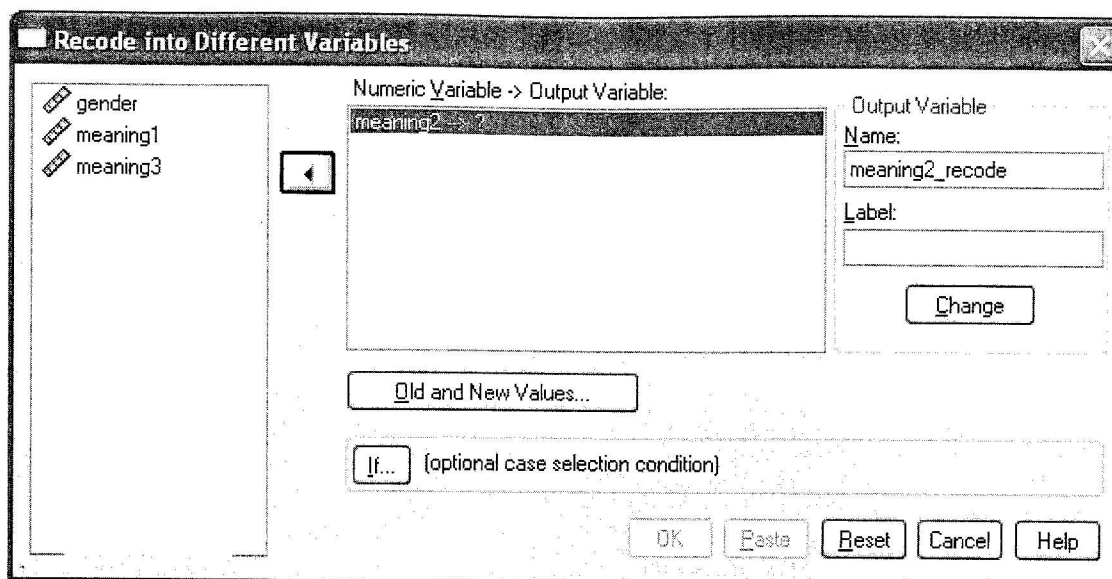


Figure A.5 The Recode into Different Variables dialog box (continued).

5. Click *Old and New Values*. The *Recode into Different Variables: Old and New Values* dialog box opens. This dialog box will be used to recode the values for **meaning2**, so that a 1 becomes a 5, a 2 becomes a 4, and so on.
6. Under *Old Value*, enter a 1 in the *Value* text box. Under *New Value*, enter a 5 in the *Value* text box (see Figure A.6).

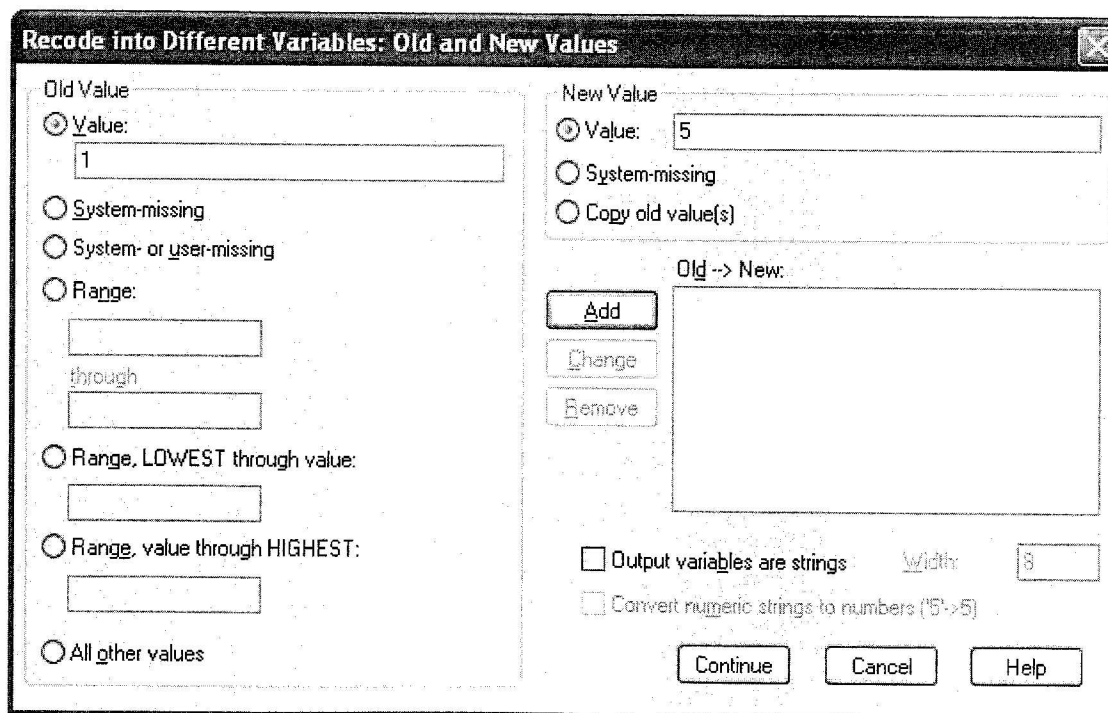


Figure A.6 The Recode into Different Variables: Old and New Values dialog box.

7. Click the *Add* button. The *Old* → *New* box should display 1 → 5 (see Figure A.7).

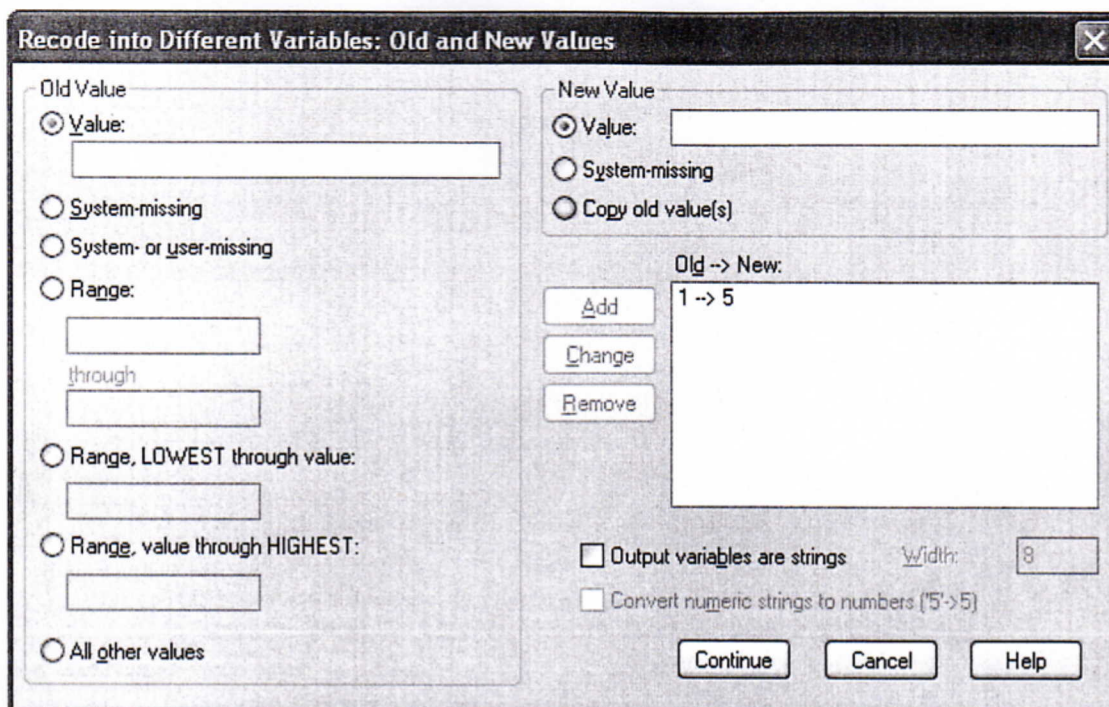


Figure A.7 The Recode into Different Variables: Old and New Values dialog box (continued).

8. Enter the remaining values to be recoded ($2 \rightarrow 4$, $3 \rightarrow 3$,¹ $4 \rightarrow 2$, $5 \rightarrow 1$), clicking the *Add* button each time an *Old* and *New Value* combination has been entered. Once all five pairs of values have been entered, the *Old* $\rightarrow$ *New* box should contain the recoded values shown in Figure A.8.
9. Click *Continue*.
10. Click *OK*.

In the SPSS *Data View* window, you should see the new variable **meaning2_recode** in the data file containing the reverse coded items of **meaning2** (see Figure A.9).

The Compute Procedure

As was indicated at the beginning of this appendix, a commonly performed operation in SPSS is to add a number of variables together to create a total score. To add variables together in SPSS, the *Compute* procedure is used.

We'll use the *Compute* procedure to add together the items of the meaning in life scale to produce a total score.

To add variables in SPSS

1. From the menu bar, select **Transform > Compute Variable** (see Figure A.10). (Note: If you are using SPSS version 14.0 or earlier, the menu commands are **Transform > Compute** All other commands are the same.)
2. The *Compute Variable* dialog box opens (see Figure A.11).

Recode into Different Variables: Old and New Values

Old Value

☒ Value:

☐ System-missing

☐ System- or user-missing

☐ Range:

☐ Range, LOWEST through value:

☐ Range, value through HIGHEST:

☐ All other values

New Value

☒ Value:

☐ System-missing

☐ Copy old value(s)

Old -> New:

Add

Change

Remove

1 -> 5

2 -> 4

3 -> 3

4 -> 2

5 -> 1

☐ Output variables are strings Width: 8

☐ Convert numeric strings to numbers ('5' -> 5)

Continue Cancel Help

Figure A.8 The Recode into Different Variables: Old and New Values dialog box (continued).

SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

49: gender

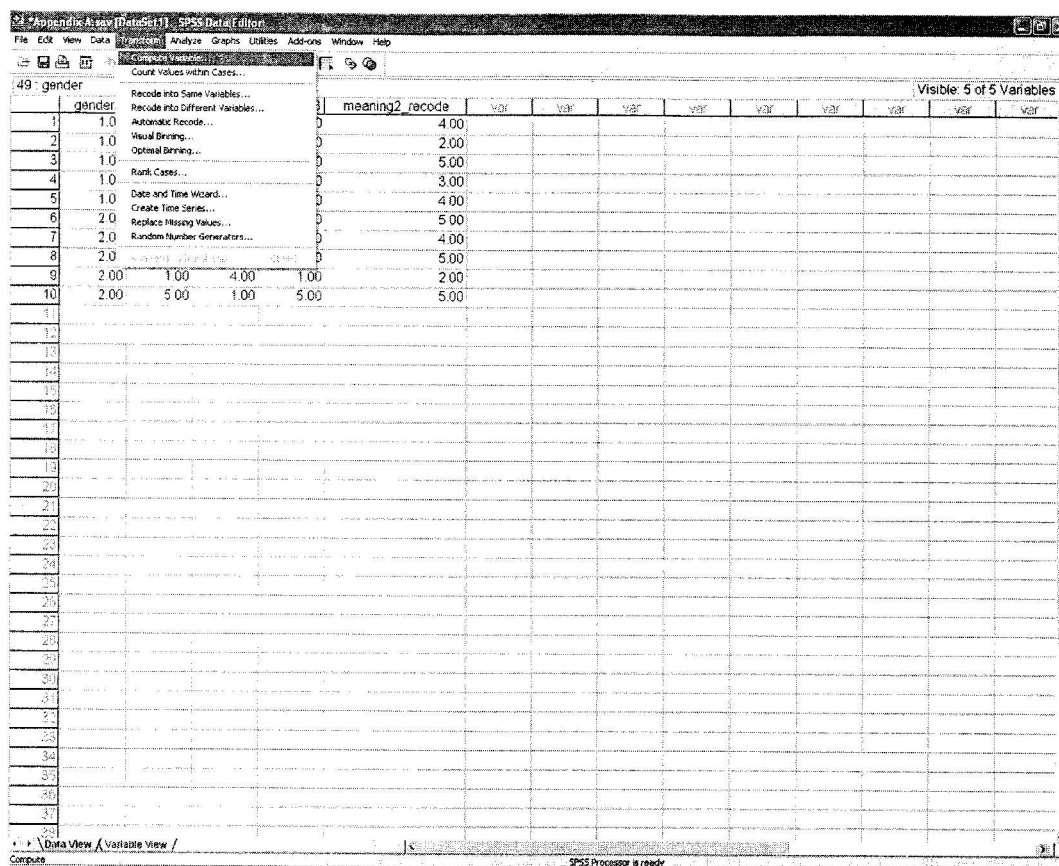
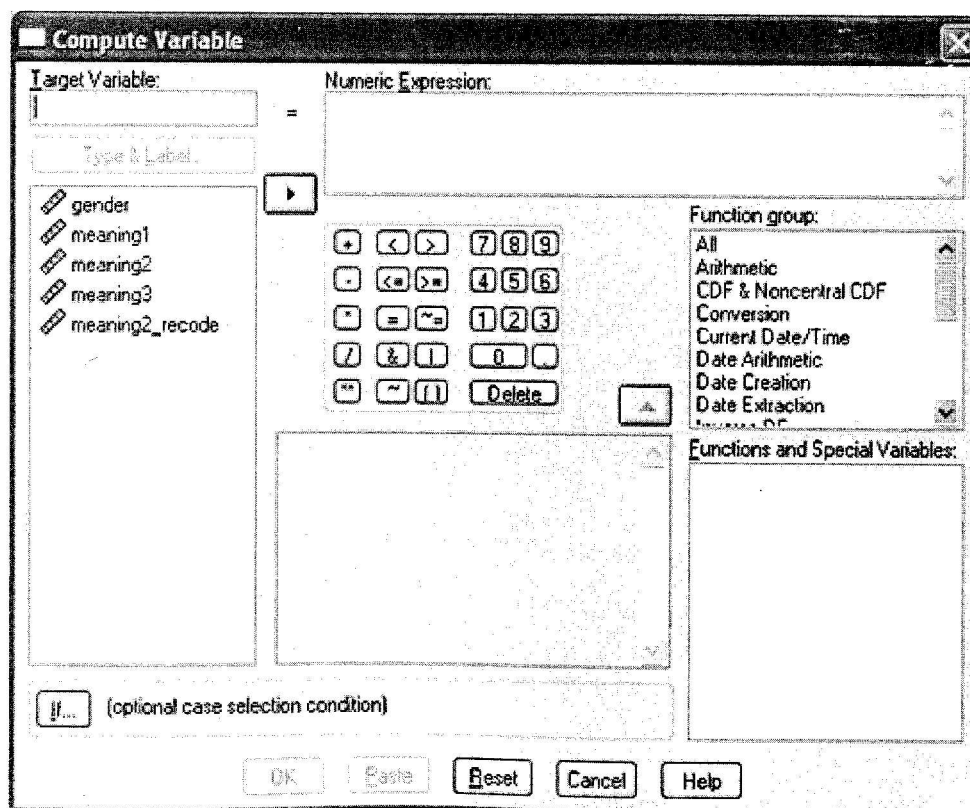
Visible: 5 of 5 Variables

	gender	meaning1	meaning2	meaning3	meaning2_recode	var	var	var	var	var	var	var	var	var
1	1.00	4.00	2.00	5.00	4.00									
2	1.00	1.00	4.00	2.00	2.00									
3	1.00	5.00	1.00	5.00	5.00									
4	1.00	3.00	3.00	3.00	3.00									
5	1.00	4.00	2.00	4.00	4.00									
6	2.00	4.00	1.00	5.00	5.00									
7	2.00	5.00	2.00	4.00	4.00									
8	2.00	5.00	1.00	5.00	5.00									
9	2.00	1.00	4.00	1.00	2.00									
10	2.00	5.00	1.00	5.00	5.00									
11														
12														
13														
14														
15														
16														
17														
18														
19														
20														
21														
22														
23														
24														
25														
26														
27														
28														
29														
30														
31														
32														
33														
34														
35														
36														
37														
38														

Data View Variable View

SPSS Processor is ready

Figure A.9 The Data Editor window with the recoded variable meaning2_recode added to the data file.²

Figure A.10 Menu commands for the *Compute* procedure.Figure A.11 The *Compute Variable* dialog box.

To compute a total score for the meaning in life scale, we'll need to add together the items **meaning1**, **meaning2_recode**, and **meaning3**.³ To create a new variable in SPSS that is equal to the sum of the three variables, we'll first need to provide a name for the new variable in SPSS that will contain the total score. We'll name the variable **meaning_total**.

1. In the *Target Variable* box, enter the name **meaning_total**.

To create a total score, we'll move the appropriate variables to the *Numeric Expression* box and add them together using the plus (+) sign.⁴

2. Select the first variable, **meaning1**, and click the right-arrow button (►) to move it to the *Numeric Expression* box. Click the plus (+) button below the *Numeric Expression* box (or press the *Shift* and *+* keys on your keyboard). Move **meaning2_recode** to the *Numeric Expression* box and click the plus (+) button. Move **meaning3** to the *Numeric Expression* box. In the *Numeric Expression* box, the statement should read, *meaning1 + meaning2_recode + meaning3* (see Figure A.12).

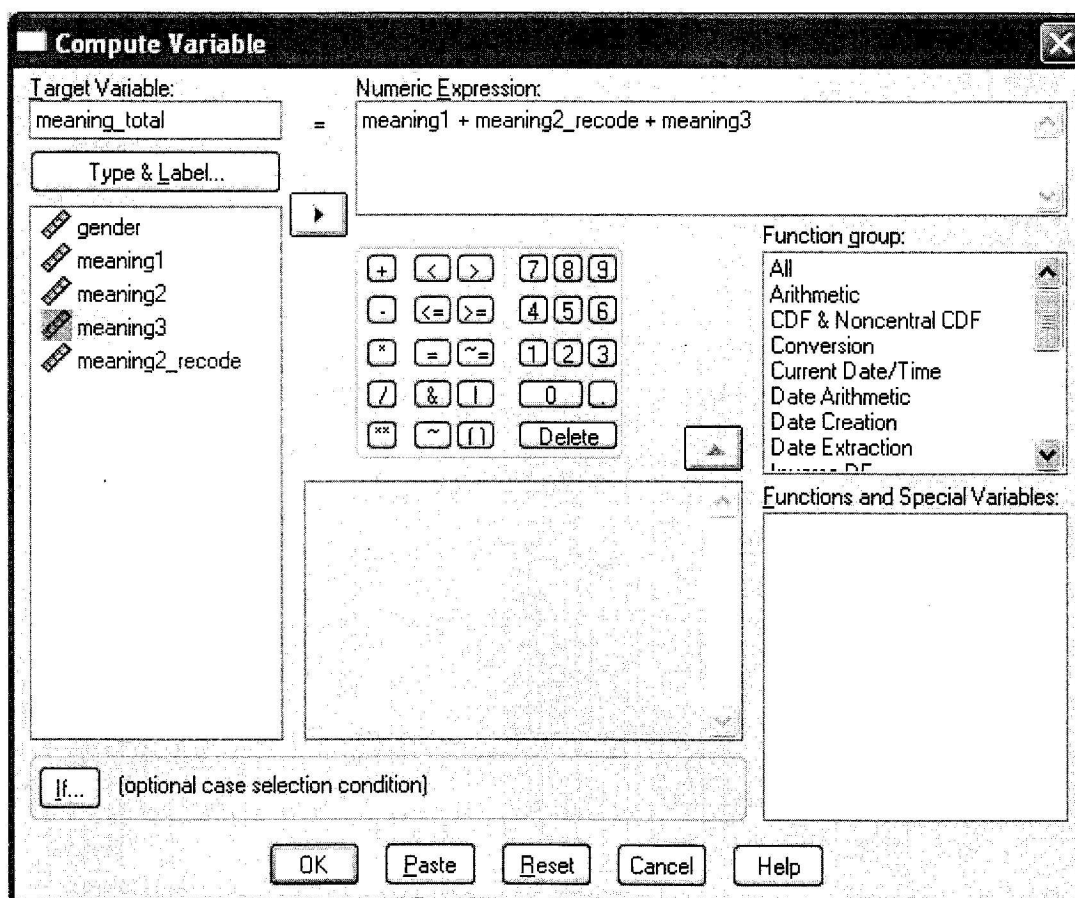


Figure A.12 The Compute Variable dialog box (continued).

3. Click *OK*. A new variable is created in the *Data View* window named **meaning_total** which is the sum of variables **meaning1**, **meaning2_recode**, and **meaning3** (see Figure A.13).

	gender	meaning1	meaning2	meaning3	meaning2_recode	meaning_total	var	var	var	var	var	var
1	1.00	4.00	2.00	5.00	4.00	13.00						
2	1.00	1.00	4.00	2.00	2.00	5.00						
3	1.00	5.00	1.00	5.00	5.00	15.00						
4	1.00	3.00	3.00	3.00	3.00	9.00						
5	1.00	4.00	2.00	4.00	4.00	12.00						
6	2.00	4.00	1.00	5.00	5.00	14.00						
7	2.00	5.00	2.00	4.00	4.00	13.00						
8	2.00	5.00	1.00	5.00	5.00	15.00						
9	2.00	1.00	4.00	1.00	2.00	4.00						
10	2.00	5.00	1.00	5.00	5.00	15.00						
11												
12												
13												
14												
15												
16												
17												
18												
19												
20												
21												
22												
23												
24												
25												
26												
27												
28												
29												
30												
31												
32												
33												
34												
35												
36												
37												
38												

Figure A.13 The Data Editor window with the variable meaning_total added to the data file.

The Select Cases Procedure

When analyzing data, certain questions of interest may require performing analyses on only *part* of the data file. In this situation, the *Select Cases* procedure may be used.

To illustrate the *Select Cases* procedure, suppose that you wanted to conduct a one-sample *t* test on *males* to see if their responses to **meaning1** were significantly different from a neutral response on the scale (i.e., a value of 3; see Chapter 5 for more information on the one-sample *t* test). This would require selecting only males in the data set, a situation where the select cases procedure could be used.

To select only males using the select cases procedure

1. From the menu bar, select **Data > Select Cases ...** (see Figure A.14).
2. The *Select Cases* dialog box opens (see Figure A.15).
3. Under *Select*, click *If condition is satisfied*.
4. Click the *If* button (located below *If condition is satisfied*). The *Select Cases: If* dialog box opens (see Figure A.16).
5. Select the variable **gender**, and click the right-arrow button (►) to move it to the box on the right.

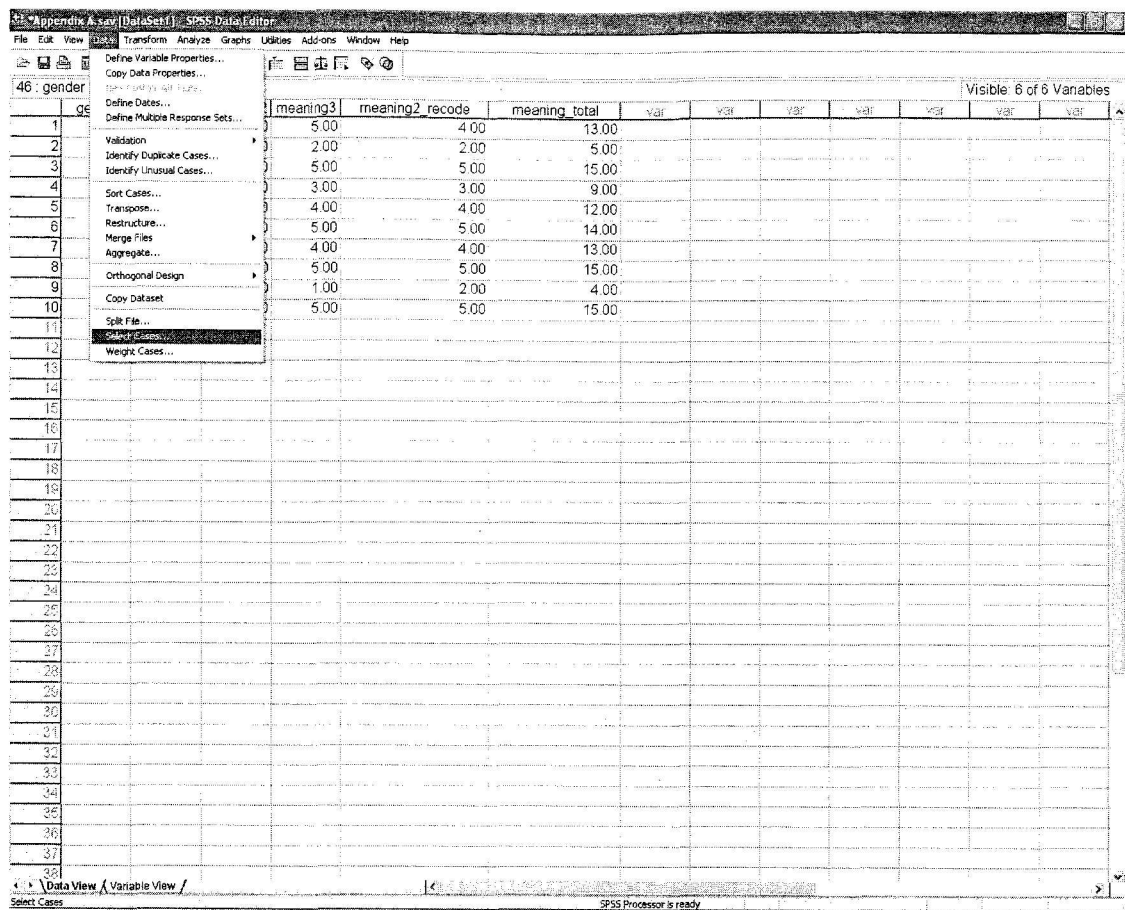


Figure A.14 Menu commands for the *Select Cases* procedure.

Recall that for this example we want to select *only* males. Since males are coded with a 1, enter “=1” (do not include the quotation marks) to the right of males (see Figure A.17). This indicates that only males will be selected by SPSS.

- 6. Click *Continue*.
- 7. The *Select Cases* dialog box should now read *gender = 1* to the right of the *If* button in the *Select Cases* dialog box (see Figure A.18).
- 8. Click *OK*.

Notice in the *Data View* window shown in Figure A.19 there are diagonal lines in the row headings of observations 6 through 10, indicating that females will not be included in the analysis. Also, a filter variable has been created in the data set indicating which cases are selected (those that receive a “1” are selected and those that receive a “0” are not).

To answer the initial question that prompted the use of the select cases procedure for this example, a one-sample *t* test was conducted on **meaning1** for males. The results of the one-sample *t* test are provided in Figure A.20.

Notice in the output shown in Figure A.20 there is no indication that males were selected. The *One-Sample Statistics* table shows an *N* of 5, however, indicating that the test was conducted

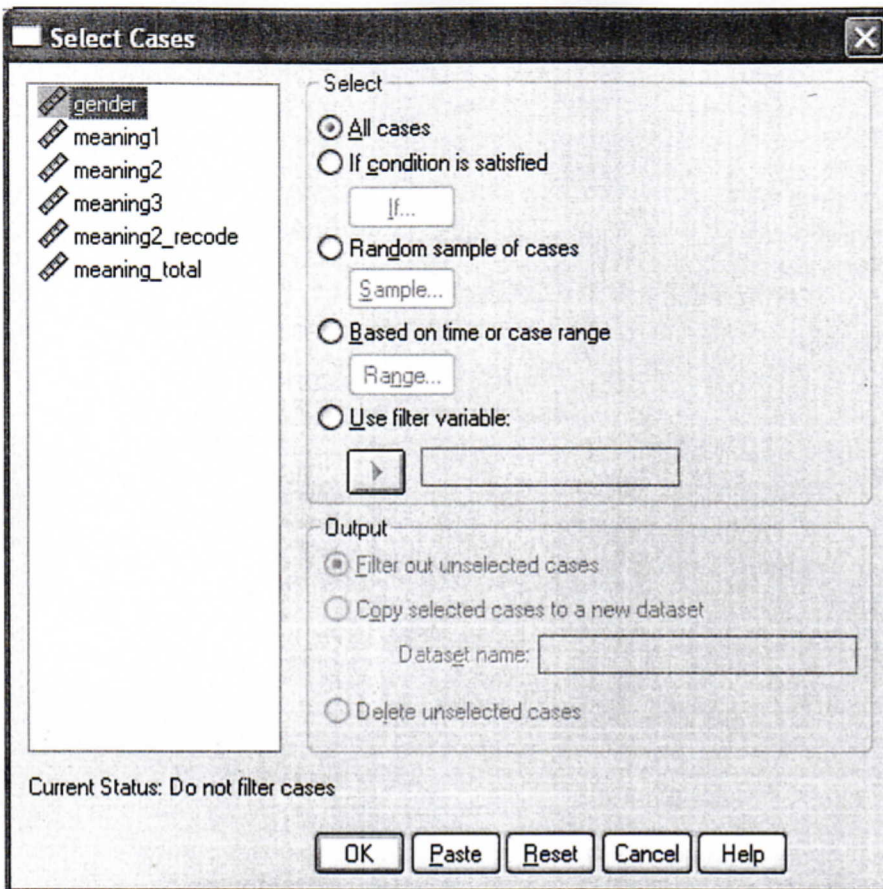


Figure A.15 The Select Cases dialog box.

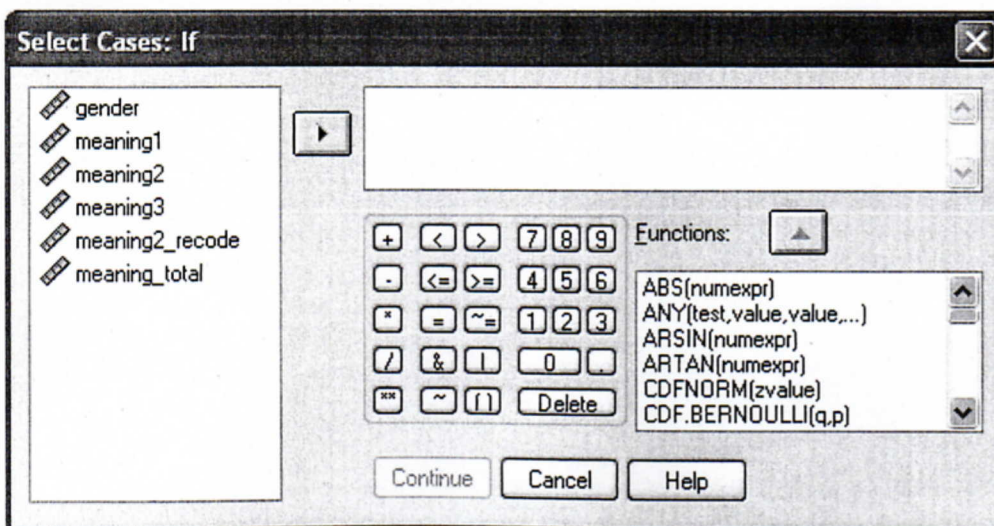


Figure A.16 The Select Cases: If dialog box.

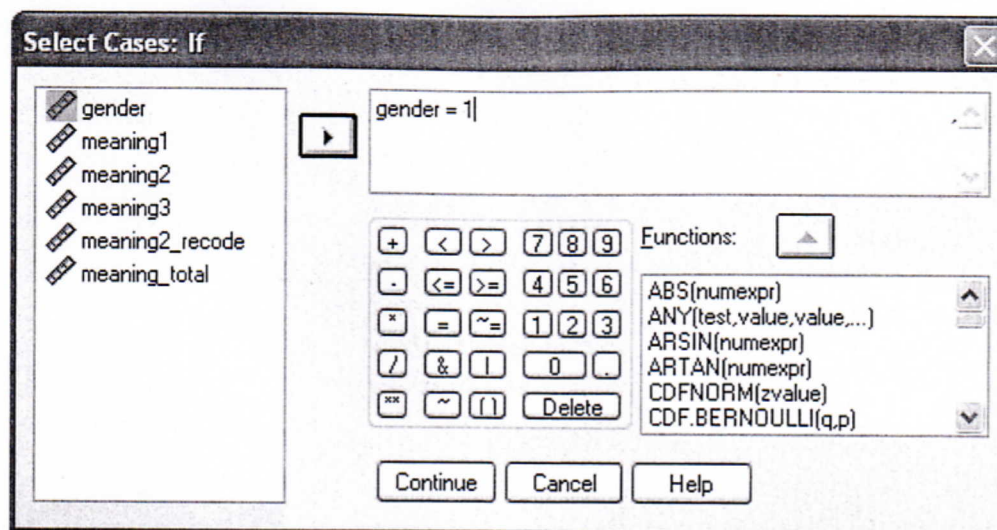


Figure A.17 The Select Cases: If dialog box (continued).

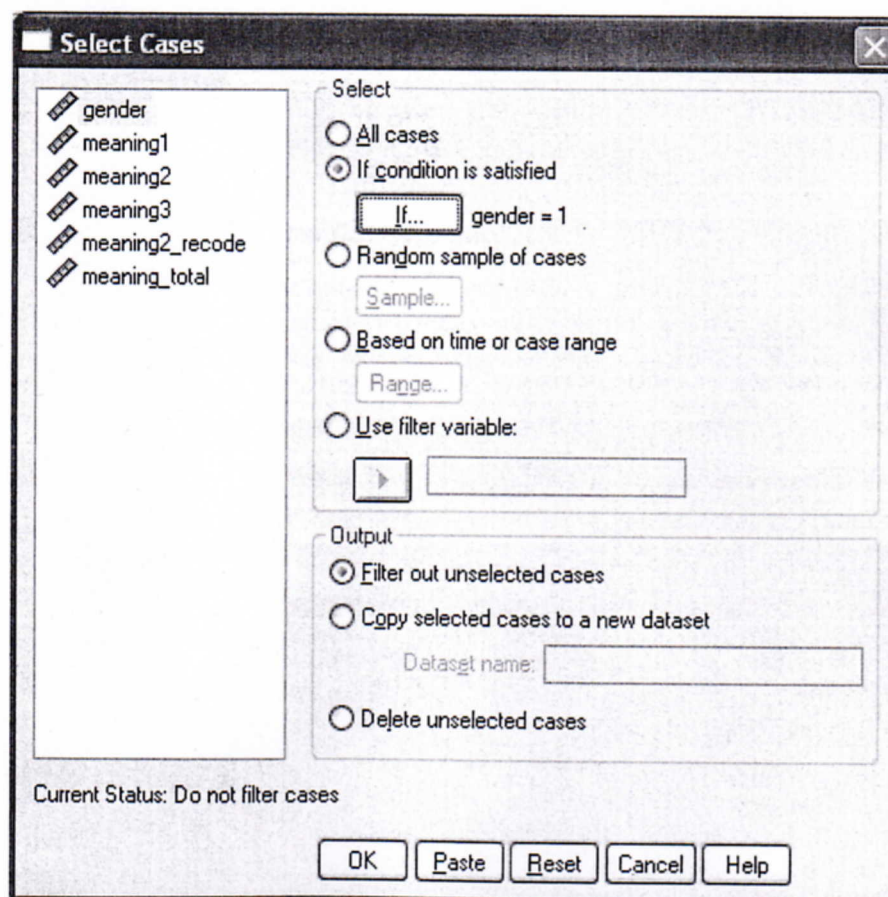


Figure A.18 The Select Cases dialog box.

Appendix A.sav [Dataset] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

46 gender Visible: 7 of 7 Variables

	gender	meaning1	meaning2	meaning3	meaning2_recodc	meaning_total	filter_\$	var	var	var	var	var	var
1	1.00	4.00	2.00	5.00	4.00	13.00	1						
2	1.00	1.00	4.00	2.00	2.00	5.00	1						
3	1.00	5.00	1.00	5.00	5.00	15.00	1						
4	1.00	3.00	3.00	3.00	3.00	9.00	1						
5	1.00	4.00	2.00	4.00	4.00	12.00	1						
6	2.00	4.00	1.00	5.00	5.00	14.00	0						
7	2.00	5.00	2.00	4.00	4.00	13.00	0						
8	2.00	5.00	1.00	5.00	5.00	15.00	0						
9	2.00	1.00	4.00	1.00	2.00	4.00	0						
10	2.00	5.00	1.00	5.00	5.00	15.00	0						
11													
12													
13													
14													
15													
16													
17													
18													
19													
20													
21													
22													
23													
24													
25													
26													
27													
28													
29													
30													
31													
32													
33													
34													
35													
36													
37													
38													
39													

Data View Variable View / SPSS Processor is ready Filter On

Figure A.19 The data file with males selected.

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
meaning1	5	3.4000	1.51658	.67823

One-Sample Test

Test Value = 3						
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
meaning1	.590	4	.587	.40000	-1.4831	2.2831

Figure A.20 The results of the one-sample *t* test on meaning1 for males.

on only five of the participants in the data file. (By the way, the results indicate that males do not score significantly different from a response of “3” on **meaning1** since the p -value of .587 is greater than .05.)⁵

Next we’ll illustrate how to use the *Split File* procedure, which is another way to perform analyses on separate parts of a data set in SPSS. Prior to using the *Split File* procedure, however, we’ll need to turn off the *Select Cases* procedure so that we have access to the entire data set. If the *Select Cases* procedure is not turned off, all subsequent analyses would be conducted *only* on males.

To turn off the Select Cases procedure

1. From the menu bar select **Data > Select Cases . . .**
2. Under *Select*, click the *All Cases* button (see Figure A.21).

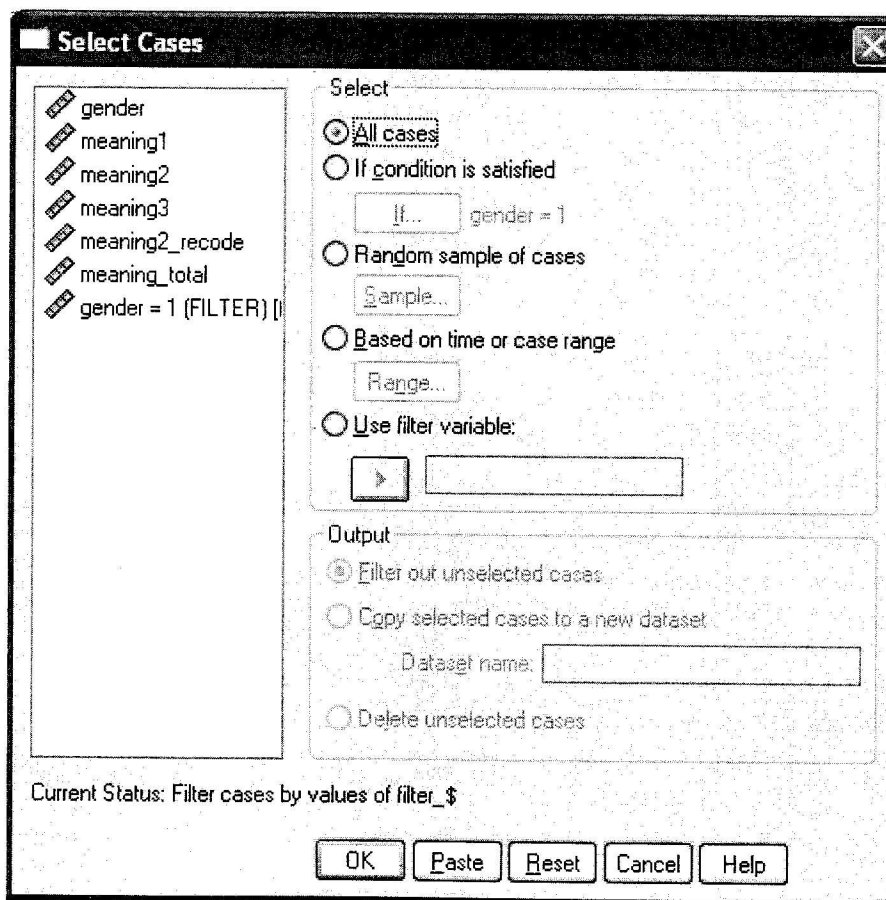


Figure A.21 The *Select Cases* dialog box.

3. Click *OK*.

In the *Data View* window, all of the cases should now be selected (there should no longer be a line through observations 6 through 10). While the filter variable still remains in the data file, it is no longer enforced. It may be deleted (if desired) by clicking on the **filter_\$** column heading in the *Data View* window and pressing the *Delete* key on the keyboard.

The Split File Procedure

Suppose that instead of conducting a one-sample t test for males only, you wanted to conduct separate one-sample t tests on **meaning1** for males and females. While the *Select Cases* procedure could be used in this situation, it would require performing the analysis twice (once for males and once for females). A more efficient approach would be to use the *Split File* procedure, which automatically performs separate analyses for each group of a variable of interest. The *Split File* procedure is illustrated next.

To use the Split File procedure

1. From the menu bar select **Data > Split File ...** (see Figure A.22)

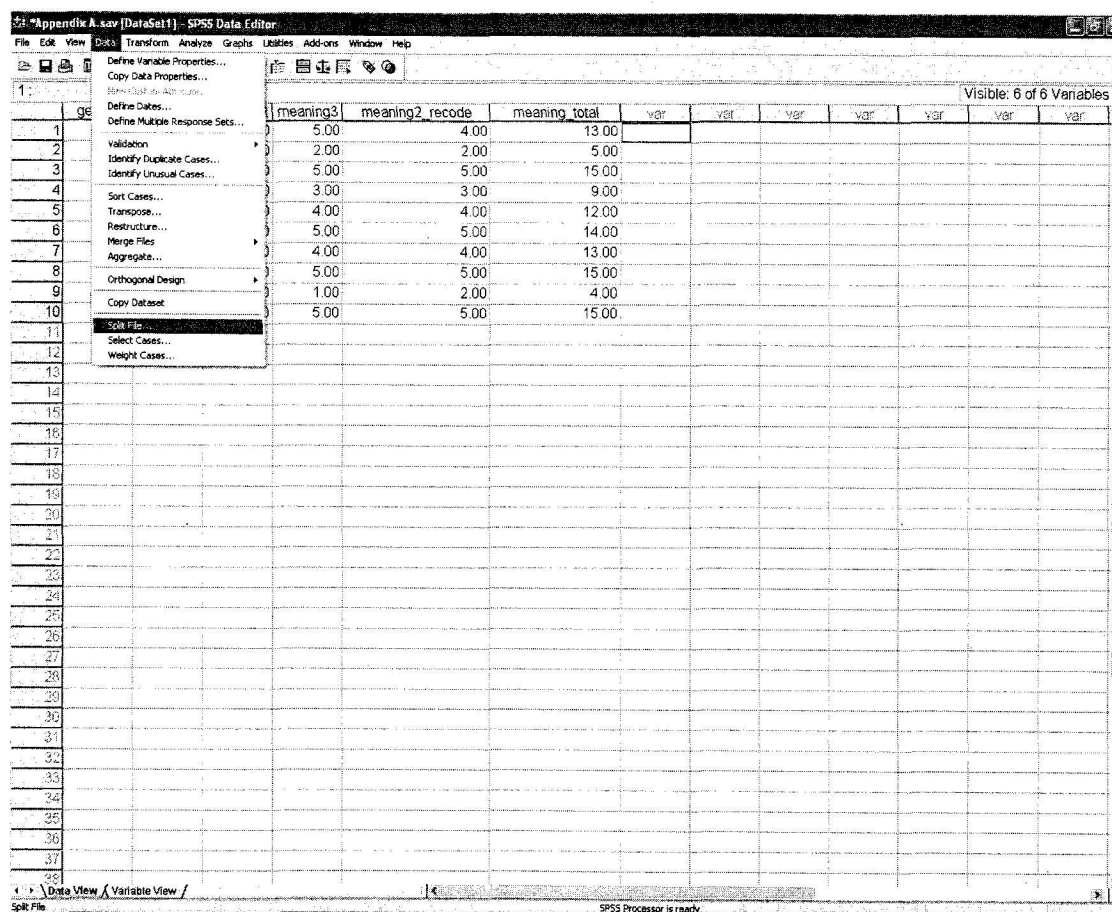


Figure A.22 Menu commands for the *Split File* procedure.

2. The *Split File* dialog box opens (see Figure A.23).

There are two different methods for splitting the file: *Compare groups* and *Organize output by groups*. *Compare groups* reports the results for the groups in the same table, while *Organize output by groups* reports the results for the groups in separate tables. We'll use the *Compare groups* method.

3. Select *Compare groups*.
4. Select **gender**. Click the right-arrow (►) button to move **gender** to the *Groups Based on* box (see Figure A.24).
5. Click **OK**.

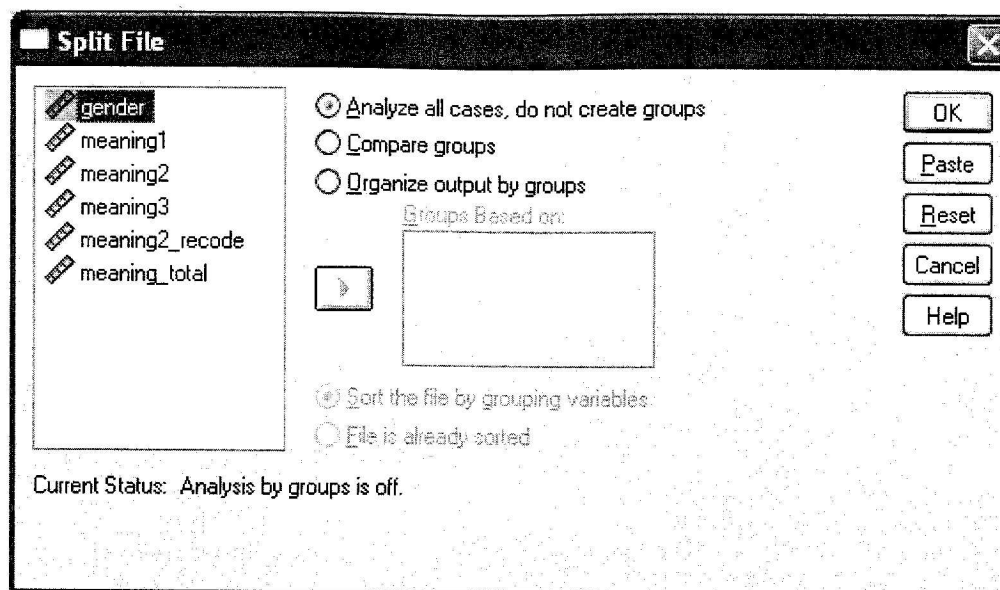


Figure A.23 The Split File dialog box.

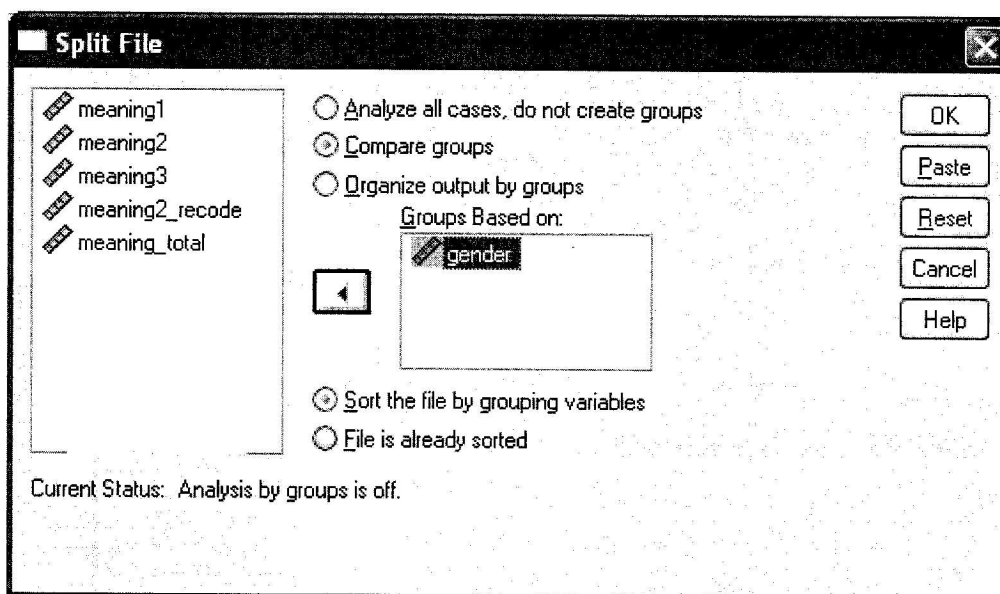


Figure A.24 The Split File dialog box (continued).

After clicking *OK* you should see “Split File On” displayed in the bottom right-hand corner of the *Data Editor* window.

The results of the one-sample *t* tests (one for males and one for females) evaluating whether the responses on **meaning1** are significantly different from a “3” are reported in Figure A.25.

Notice that the output of the one-sample *t* test shows the results for males and females together in the same tables. If *Organize output by groups* was selected, the results for males and females would be in separate tables (with two tables of results for males and two tables for females). (By the way, do the results indicate that either males or females have an average score on **meaning1** that is significantly different from a response of “3”? They don’t, since the *p*-values of .587 and .266 are both greater than .05.)

T-Test

One-Sample Statistics

		N	Mean	Std. Deviation	Std. Error Mean
gender					
male	meaning1	5	3.4000	1.51658	.67823
female	meaning1	5	4.0000	1.73205	.77460

One-Sample Test

One-Sample Test							
		Test Value = 3					
		t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
						Lower	Upper
gender							
male	meaning1	.590	4	.587	.40000	-1.4831	2.2831
female	meaning1	1.291	4	.266	1.00000	-1.1506	3.1506

Figure A.25 The results of the one-sample *t* test on meaning1 for males and females.

Turning off the *Split File* Procedure

As was the case with the *Select Cases* procedure discussed earlier, if you wanted to perform analyses on the entire sample, the *Split File* procedure will need to be turned off prior to conducting the analyses. To turn off the *Split File* procedure, in the *Split File* dialog box select *Analyze all cases, do not create groups*, *do not create groups* and click *OK* (see Figure A.26). (Alternatively, if SPSS is closed after using the *Split File* or *Select Cases* procedure, when SPSS is started again the option will be turned off by default.)

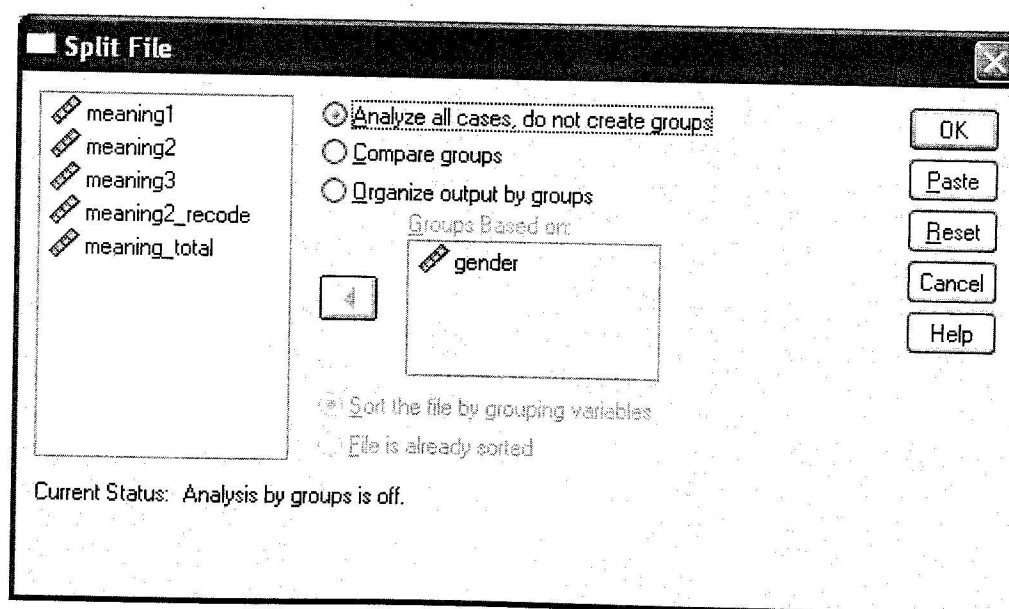


Figure A.26 Turning off the *Split File* procedure.

This concludes the discussion on the *Recode*, *Compute*, *Select Cases*, and *Split File* procedures.

B

Importing Files

In this appendix, how to import Microsoft Excel files into SPSS is illustrated. To demonstrate this procedure, an Excel file was created using the data from Appendix A. The Excel file is presented in Figure B.1.

	A	B	C	D
1	gender	meaning1	meaning2	meaning3
2	1	4	2	5
3	1	1	4	2
4	1	5	1	5
5	1	3	3	3
6	1	4	2	4
7	2	4	1	5
8	2	5	2	4
9	2	5	1	5
10	2	1	4	1
11	2	5	1	5

Figure B.1 The Excel file *Excel File_Appendix B.xls*.

To import the Excel file into SPSS, we'll use the *SPSS Database Wizard*, which provides assistance with importing files into SPSS.

To access the Database Wizard

1. From the menu bar select **File > Open Database > New Query . . .** (see Figure B.2)
2. The *Database Wizard* dialog box opens (see Figure B.3).

Three different types of files are listed in the *Database Wizard* dialog box, including *dBASE Files*, *Excel Files*, and *MS Access Database* files. Since we want to import an Excel file, we'll select *Excel Files*.

3. Select *Excel Files* from the *Database Wizard* dialog box.
4. Click *Next*. The *ODBC Driver Login* dialog box appears (see Figure B.4).

Next we'll open the Excel file from your computer.

5. Click the *Browse* button (see Figure B.4) and select the location of the *SPSS Demystified* files on your computer. (If the files have not been downloaded to your computer, they may be obtained on the web at www.prenhall.com/yockeye.)

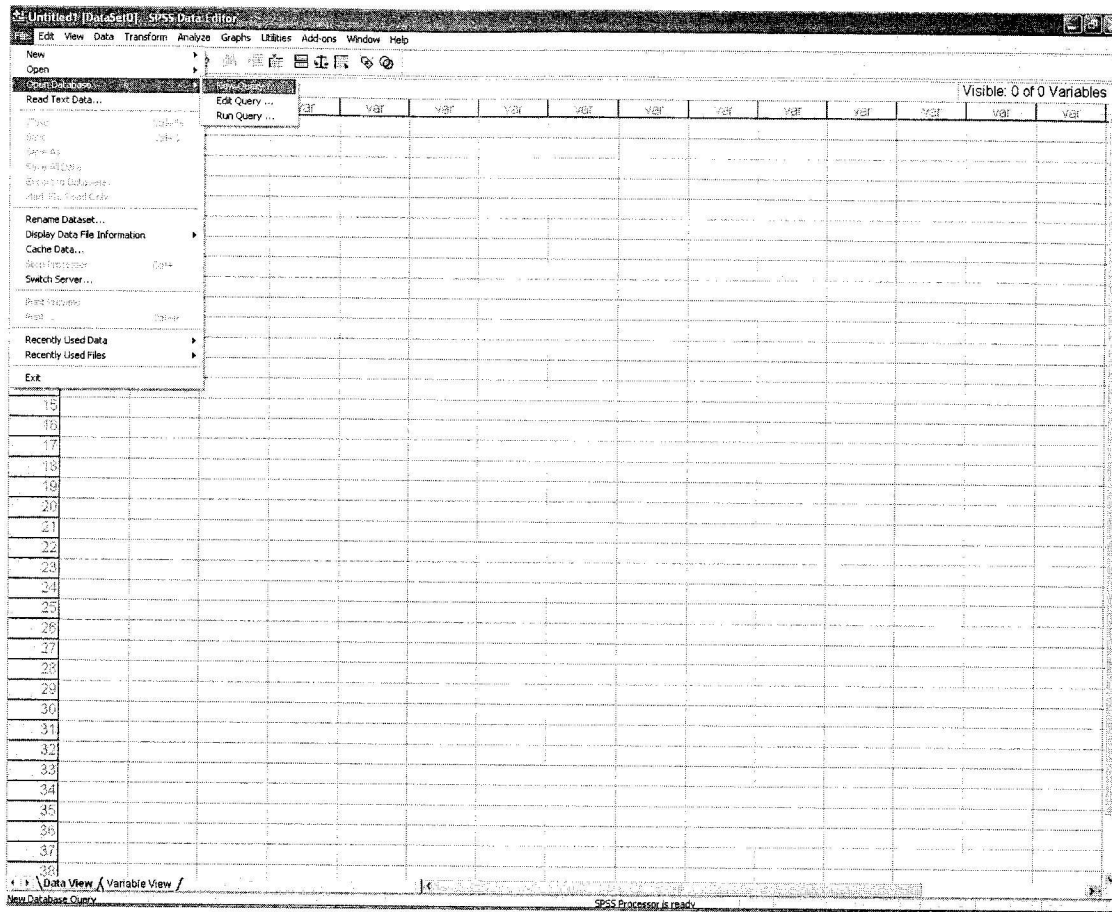


Figure B.2 Menu commands for importing an Excel file into SPSS.

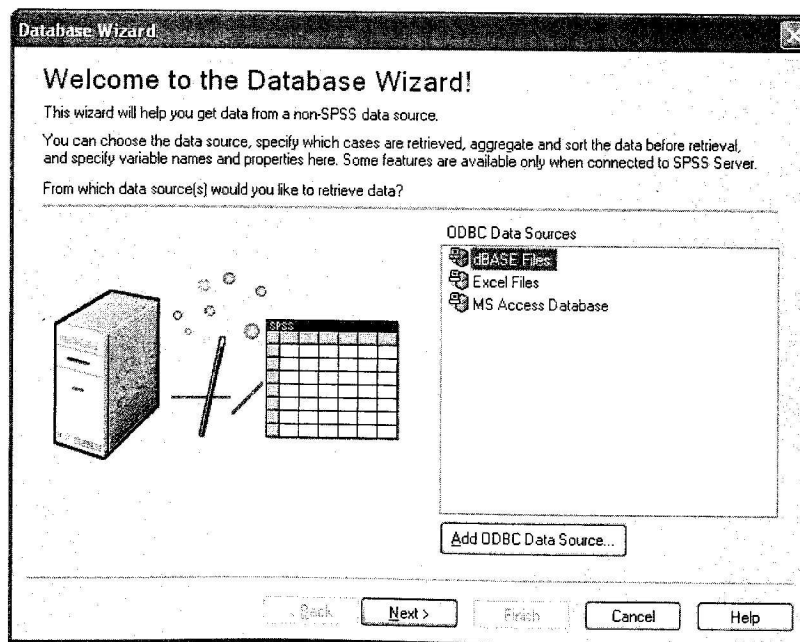


Figure B.3 The Database Wizard dialog box.

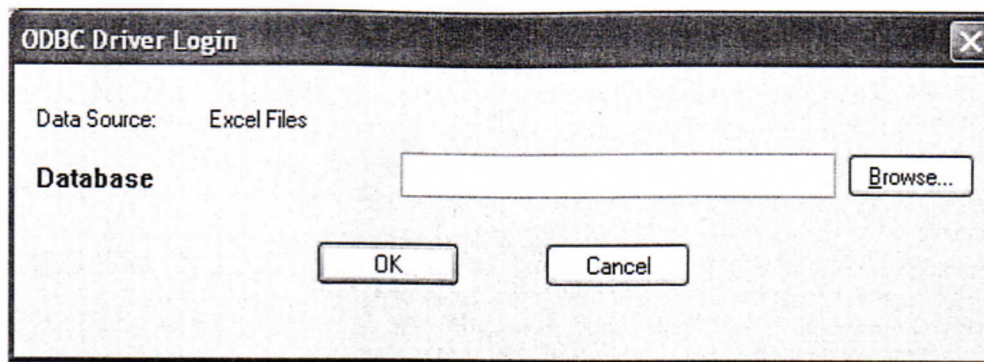


Figure B.4 The ODBC Driver Login dialog box.

6. Open the *Appendix B* folder and select the file *Excel File_Appendix B.xls*. Click *Open*.
7. The file is presented in the *Database* textbox of the *ODBC Driver Login* window (see Figure B.5). (Note: while the file is located on drive E: in Figure B.5, your file may be in a different location on your computer.)

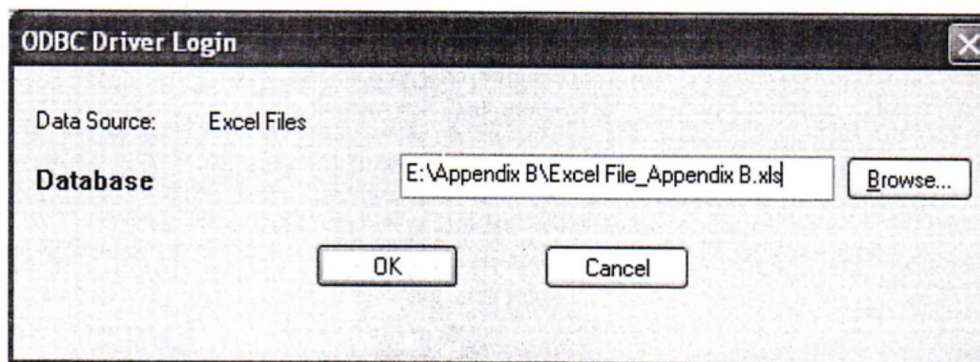


Figure B.5 The ODBC Driver Login dialog box (continued).

8. Click *OK*.
9. The *Database Wizard: Select Data* dialog box opens (see Figure B.6).

Three different sheets (*Sheet1\$*, *Sheet2\$*, and *Sheet3\$*) are presented in the *Available Tables* box, which correspond to the three worksheets created by default for an Excel file. Our data are located in the first worksheet, *Sheet1\$*. (To verify this, if you click on the + button to the left of *Sheet1\$*, you will see the variables **gender**, **meaning1**, **meaning2**, and **meaning3** displayed.) The instructions in the *Database Wizard* dialog box indicate to drag the items of interest from the *Available Tables* box to the *Retrieve Fields in This Order* box. Since our data are in *Sheet1\$*, we'll drag *Sheet1\$* to the *Retrieve Fields in This Order* box.

10. Select *Sheet1\$* and drag it to the *Retrieve Fields in This Order* box (see Figure B.7).

The *Retrieve Fields in This Order* box displays the four variables in the data file, **gender**, **meaning1**, **meaning2**, and **meaning3**, confirming that the correct file has been selected.

11. Click *Finish*. The data file opens with the participants' scores in rows 1 through 10 of the *Data View* window and the variables in the first four columns (see Figure B.8).

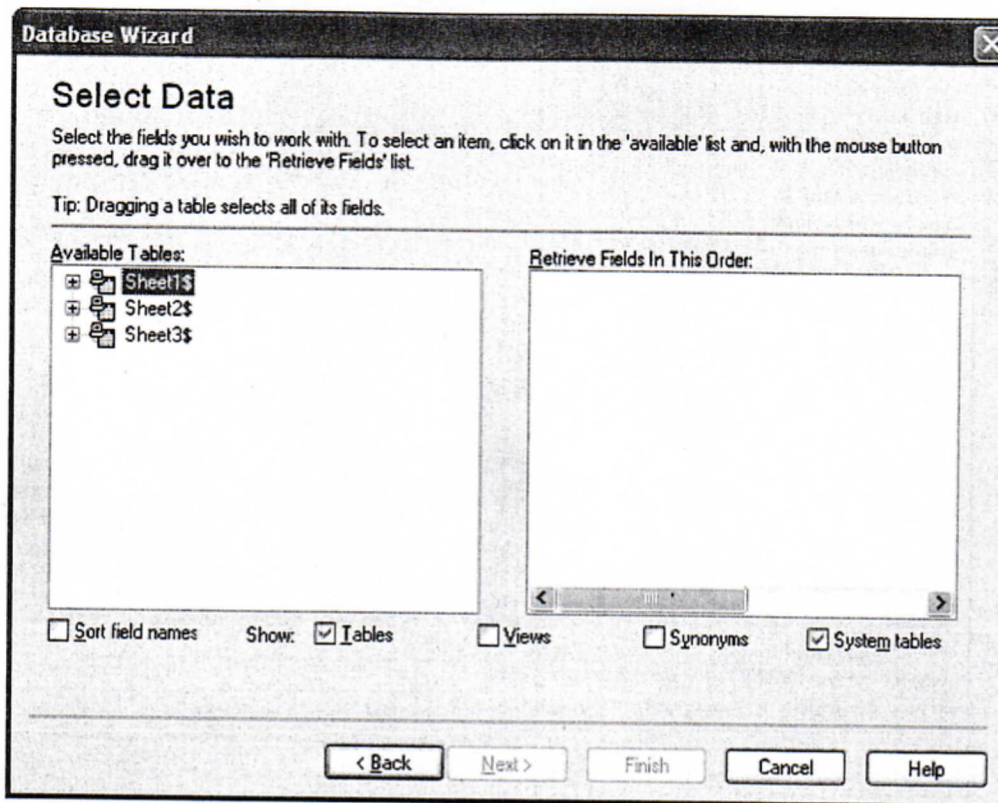


Figure B.6 The Database Wizard: Select Data dialog box.

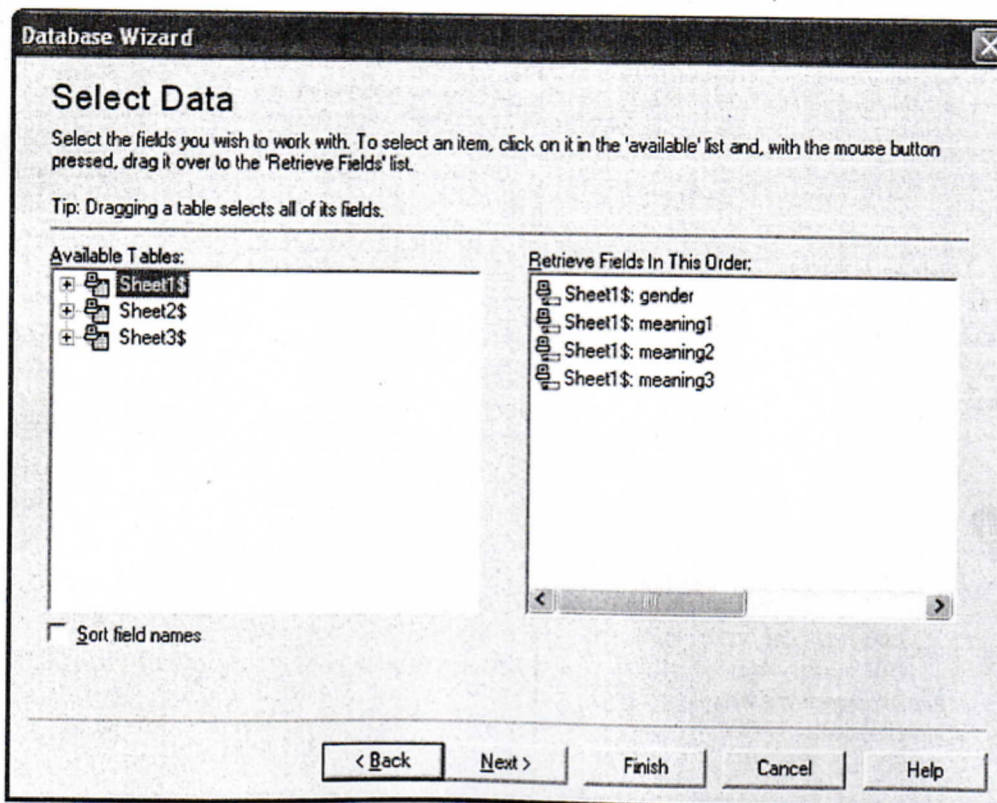


Figure B.7 The Database Wizard: Select Data dialog box (continued).

48 : meaning1

Visible: 4 of 4 Variables

	gender	meaning1	meaning2	meaning3	var	var	var	var	var	var	var	var	var	var
1	1.00	4.00	2.00	5.00										
2	1.00	1.00	4.00	2.00										
3	1.00	5.00	1.00	5.00										
4	1.00	3.00	3.00	3.00										
5	1.00	4.00	2.00	4.00										
6	2.00	4.00	1.00	5.00										
7	2.00	5.00	2.00	4.00										
8	2.00	5.00	1.00	5.00										
9	2.00	1.00	4.00	1.00										
10	2.00	5.00	1.00	5.00										
11														
12														
13														
14														
15														
16														
17														
18														
19														
20														
21														
22														
23														
24														
25														
26														
27														
28														
29														
30														
31														
32														
33														
34														
35														
36														
37														
38														

Data View Variable View /

SPSS Processor is ready

Figure B.8 The Excel file imported into SPSS.

This concludes the illustration of importing Excel files into SPSS. If you have data saved in *MS Access*, you may open it in SPSS using the *New Query* procedure. If you have data saved in a format where the values are separated by commas, spaces, or tabs, the file will need to be opened using an alternative method in SPSS. To open these types of files, select **File > Read Text Data . . .** from the menu bar and follow the on-screen instructions.