Basic SPSS

by Elvira Agrón

TASC/Advanced Support Team
Center for Information Technology
National Institutes of Health

May 10-11, 2001

Course Description

This course will teach you the basics of SPSS 10.1. Specifically, it covers:

- using the Help facility
- reading and creating data files
- importing and exporting files
- setting variable attributes
- creating and recoding variables
- merging files
- using simple SPSS syntax
- simple statistics
- creating graphics with SPSS

SPSS 10.0 at NIH

The Center for Information Technology of NIH has a site license for SPSS 10.0 for Windows and the Macinthosh. You can obtain the software and inquire about the current cost by calling GO CIT within NIH or 301-594-6248.

Documentation is provided with the SPSS CD except the "SPSS 10.0 Syntax Reference Guide" which is provided in the CD in PDF format.

The NIH site license includes the following SPSS modules:

- SPSS Base
- SPSS Advanced Models (formerly known as Advanced Statistics)
- SPSS Regression Models (formerly known as Professional Statistics)
- SPSS Trends
- SPSS Tables
- SPSS Categories
- SPSS Exact Tests
- SPSS Missing Values
- SPSS Maps

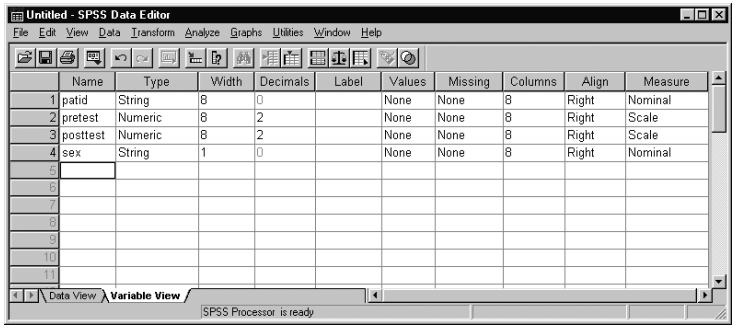
Creating SPSS Files

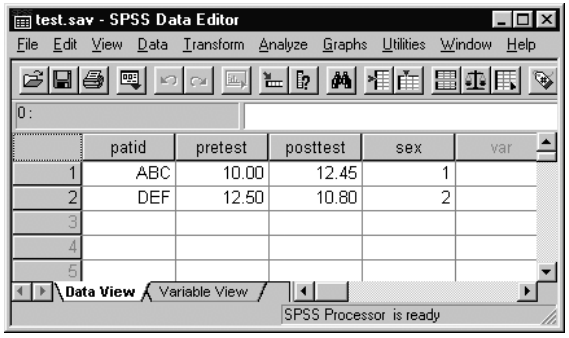
When you invoke SPSS you will see a **Data View**, a **Variable View**, menus and a toolbar.

You can either enter data directly into the **Data View** or open an existing file. To open an existing file select **File → Open → Data**.

To create a new SPSS file first enter the variable names (columns) in the **Variable View** window. Here you can also specify the attributes of the variables. After defining your variables enter values in the **Data View**.

Save the file by selecting **File → Save As**. The file will be assigned the extension SAV.





Variable Attributes

Variable Names:

•Name must begin with a letter and followed by any letter, digit, period, or @, #, _ or $. It cannot end in a period.

•It must be a maximum of eight characters.

•Variable names are not case sensitive. AGE and Age are the same.

Variable Types:

Can be set to numeric, comma, dot, scientific notation, date, dollar, custom currency and string.

Measurement Level:

You can define the level of measurement to be scale, ordinal or nominal.

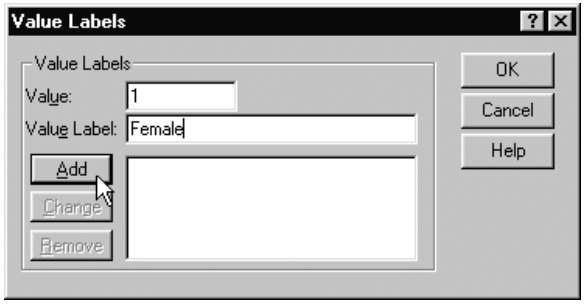
Variable Labels:

Can be set in the Variable View window under the column Label. Can be up to 256 characters.

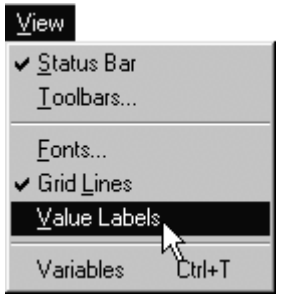
Value Labels

You can assign labels to the values of a variable in the **Variable View** window. They can be up to 60 characters long. For example, for the variable SEX we may want to assign "Female" to the value 1 and "Male" to the value 2.

To set the value labels click the cell under the column **Values** for variable SEX. The **Value Labels** window will appear. Enter the value 1 and the value label Female then press **Add**. Then enter the value 2 and the value label Male then press **Add** again. Now press **OK**.

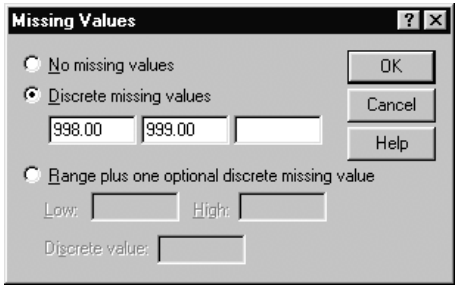


To view the value labels instead of the actual values in the **Data View** select **View → Value Labels**.

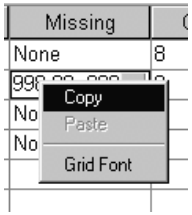


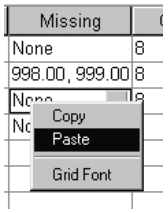
Missing Values

You can assign special missing values to the variables. To do this click the cell under the column **Missing** in the **Variable View**. Here we set the special missing values 998 and 999 to the variable PRETEST. Enter the values in **Discrete missing values** then press **OK**.



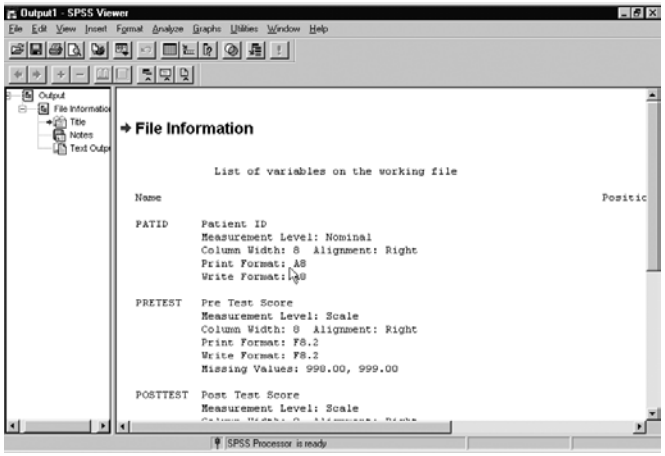
To apply the same special missing values to the variable POSTTEST right-click on the **Missing** cell for PRETEST then select **Copy**. Now right-click on the **Missing** cell for POSTTEST and select **Paste**. This cut-and-paste method also works with the **Value Labels**.



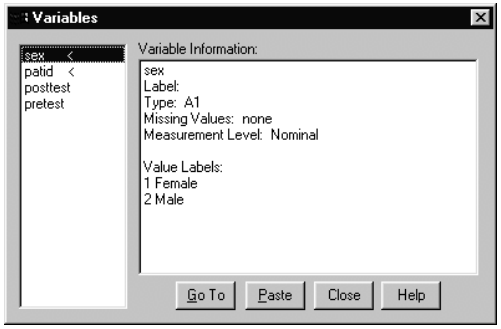


Variable Information

The variable attributes are saved in what is called the **data dictionary**. To view the data dictionary select **Utilities → File Info**. The results will be displayed in the **Output** window.

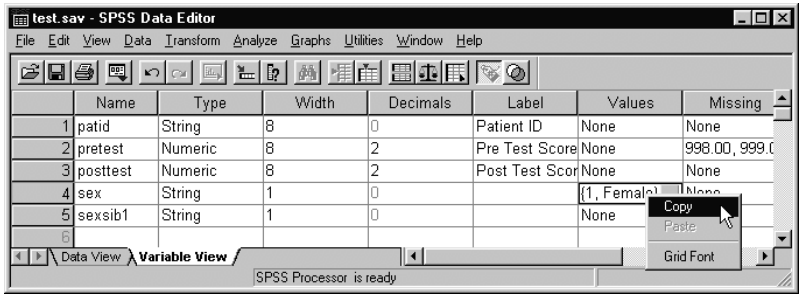


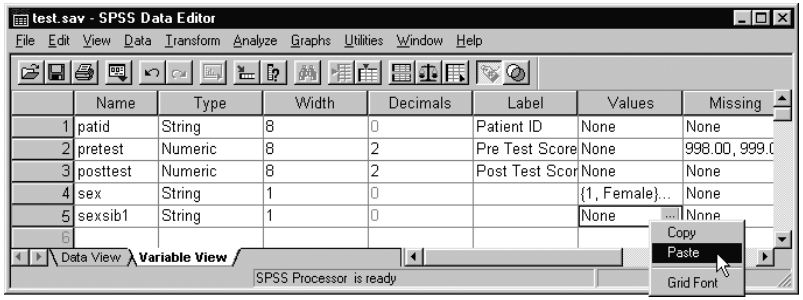
Alternatively, select **Utilities→Variables** to view attributes of selected variables.



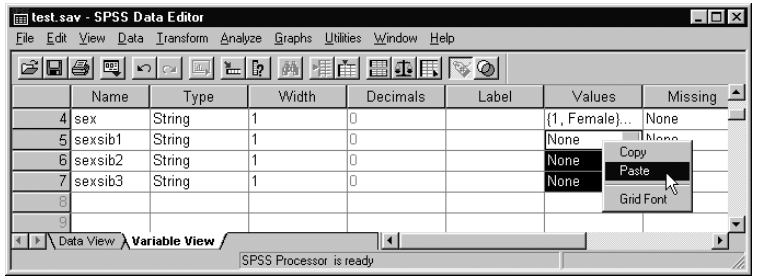
Copying Selected Attributes

You can cut and paste variable attributes to new variables in the **View** window. The variables must be of the same type and width. In this example we copy the value labels from the variable SEX and paste them into the new variable SEXSIB1.





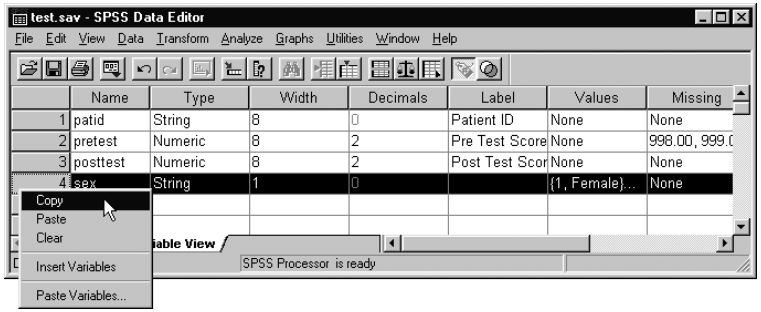
To paste to multiple variables, click and drag down the appropriate attribute column for all the variables and select Paste from the menu. Here we paste the labels to SEXSIB1, SEXSIB2 and SEXSIB3.

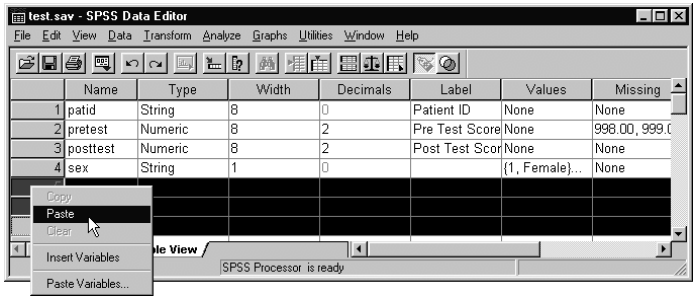


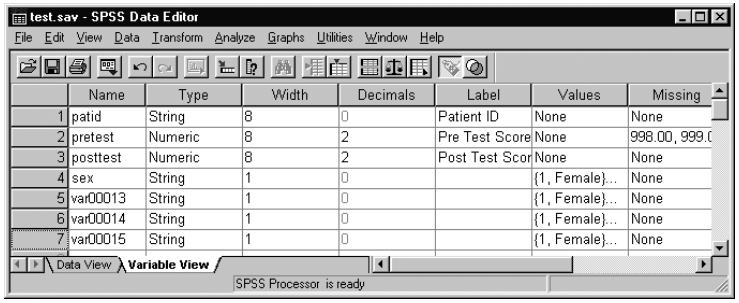
Copying All Attributes

To copy all the variable attributes to a new variable click on the row number for the variable to copy and select **Copy**. Then click on the row number(s) of the new variable(s) and select **Paste**.

When copying attributes you do not need to define the variable first. If you choose empty rows SPSS will automatically create variable names for you.



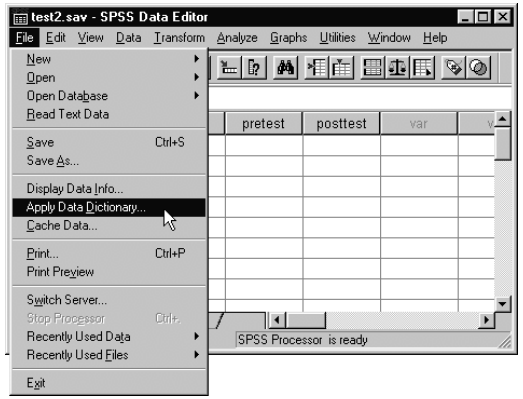


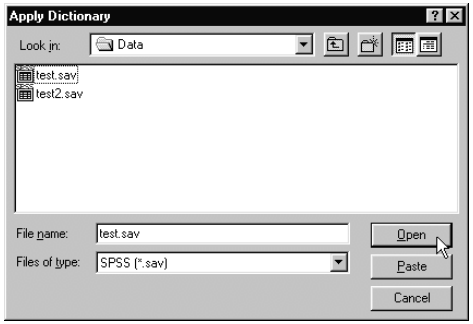


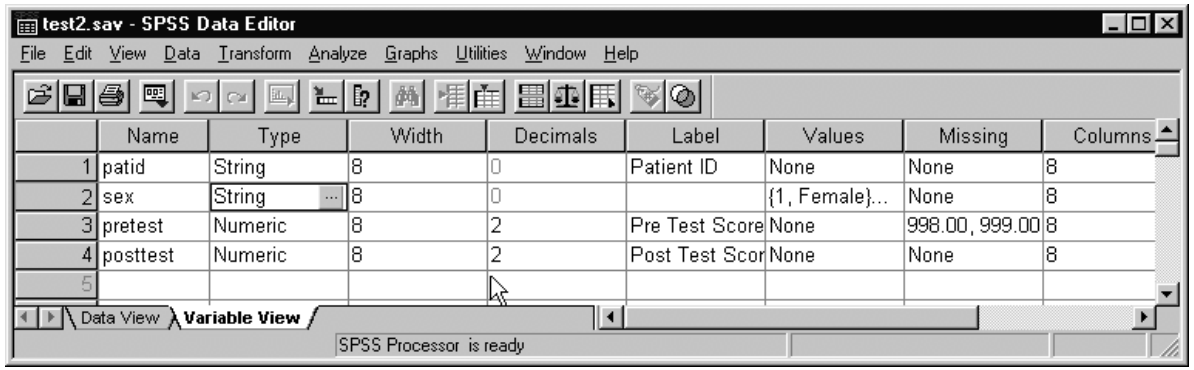
Apply Data Dictionary

The data dictionary can be applied to new SPSS files. To apply a data dictionary to the file you are working with select **File → Apply Data Dictionary** from the **Data Editor** window.

In the following example we apply the data dictionary of the SPSS file TEST.SAV to the new SPSS file TEST2.SAV. You must **Save** the file to retain the attributes.







SPSS Syntax

For most analyses SPSS can be used by using the menus and dialog windows. However, some features can only be used with command syntax. Whenever you point-and-click in the menus and dialog windows SPSS generates commands behind the scenes. You can view these commands by selecting the **Paste** button in most dialog windows.

For example, when we open the file TEST2.SAV and press **Paste** in the **Open File** window it produces the following syntax in a **Syntax** window:

```
GET  FILE='C:\MyData\test2.sav'.
```

Selecting **Paste** in the **Apply Dictionary** window and then in the **Save As** window generates the following two commands:

```
APPLY DICTIONARY FROM='C:\MyData\test.sav'.

SAVE OUTFILE='C:\MyData\test2.sav'
/COMPRESSED.
```

To save the syntax select **File → Save As** from the **Syntax** window. Syntax files are saved with the extension SPS.

You must use an **EXECUTE** command after commands found in the **Transform** menu.

SPSS Syntax

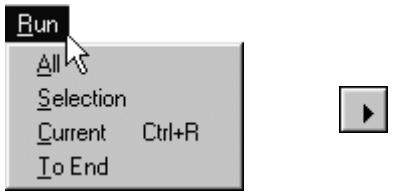
Displaying Syntax in the Output

To always view the syntax in the **Output** window select **Edit → Options** then the **Viewer** tab. Select the box at the bottom labeled **Display commands in the log**.

Using Syntax Files

If you want to write the commands yourself select **File → New → Syntax** to create a new syntax file. To open an existing syntax file select **File → Open → Syntax**. You may have multiple syntax windows.

To run all or part of the syntax file select the appropriate item from the **Run** menu. The toolbutton with a triangle pointing to the right (shown below) runs only the command where the cursor is placed. **Ctrl-R** can be used to run selected commands.



Commands You Cannot Paste

There is no way to paste the following commands:

VALUE LABELS

ADD VALUE LABELS

VARIABLE LABELS

RENAME VARIABLES

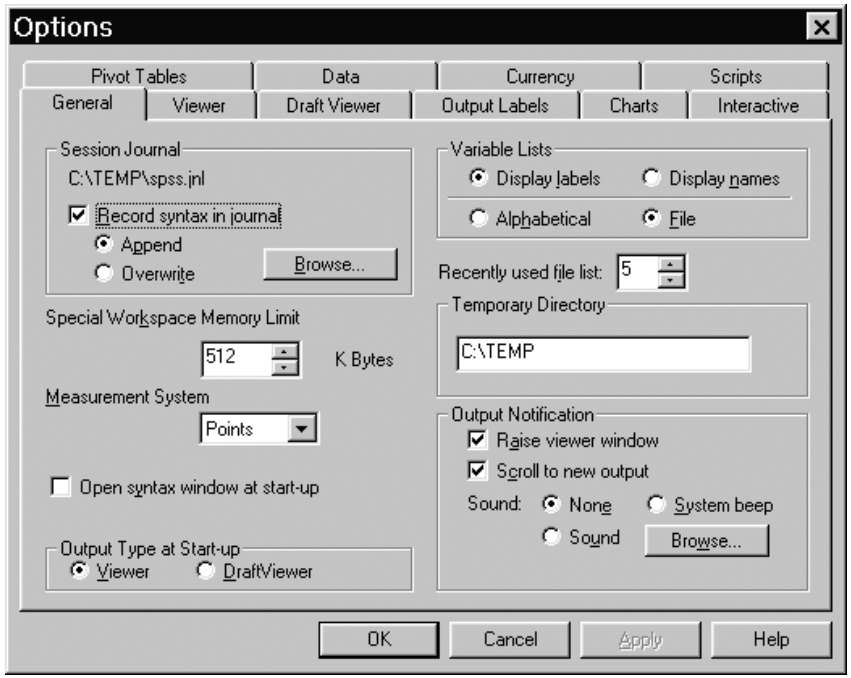
FORMAT

STRING

MISSING VALUES

The Journal

All the syntax that is generated by SPSS is saved automatically in a journal. You can find or modify the location of this file by selecting **Edit → Options** then the **General** tab.



Workshop 1

In this exercise you will read a text data set called PATIENTS.DAT, define and assign attributes to variables and save the data as an SPSS file.

1. Follow these steps to open the file PATIENTS.DAT:

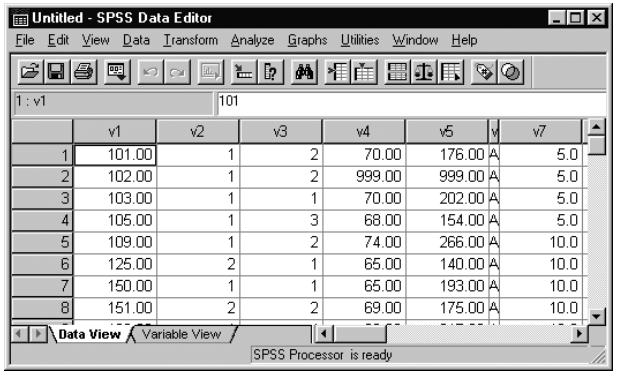
a. After opening SPSS choose **File → Read Text Data**.

b. In the **Open File** window enter **c:\basicspss\patients.dat** and press the **Open** button.

c. Press **Next** in all of the following five windows (steps 1 through 5 of the **Text Import Wizard**) taking all the defaults.

d. Press **Finish** in the next window (step 6 of **Text Import Wizard**).

This is what your **Data View** will look like:



	v1	v2	v3	v4	v5	v	v7
1	101.00	1	2	70.00	176.00	A	5.0
2	102.00	1	2	999.00	999.00	A	5.0
3	103.00	1	1	70.00	202.00	A	5.0
4	105.00	1	3	68.00	154.00	A	5.0
5	109.00	1	2	74.00	266.00	A	10.0
6	125.00	2	1	65.00	140.00	A	10.0
7	150.00	1	1	65.00	193.00	A	10.0
8	151.00	2	2	69.00	175.00	A	10.0

Workshop 1

2. Make the following changes to the variable attributes:

a. Change the variable names from V1 through V7 to the following names and **in this order**: PATID, SEX, RACE, HEIGHT, WEIGHT, CENTER and DOSE.

b. Change the type of PATID, SEX and RACE to String.

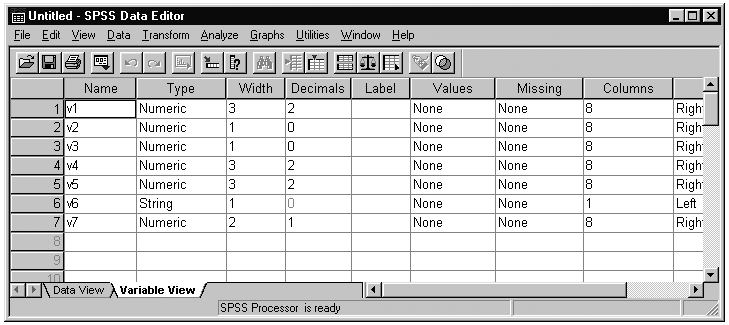
c. Assign the following value labels to SEX: 1=Male, 2=Female

d. Assign the following value labels to RACE: 1=White, 2=Black, 3=Hispanic, 4=Asian, 5=Other

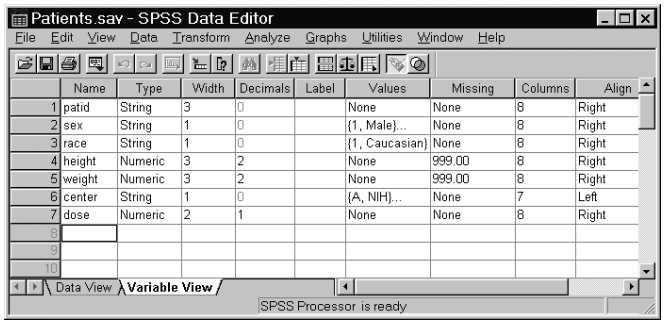
e. Assign the following value labels to CENTER: A=NIH, B=CDC

f. Assign a missing value of 999 to the variables HEIGHT and WEIGHT.

3. Save this new SPSS file as PATIENTS. (Do not specify extension.)



Before



After

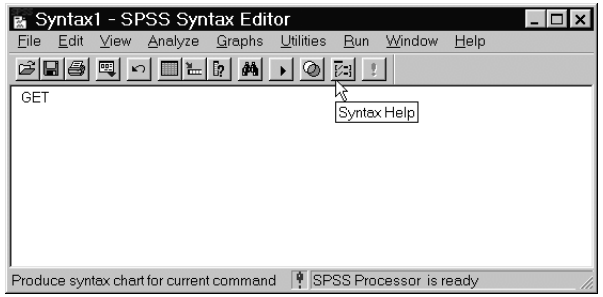
Getting Help

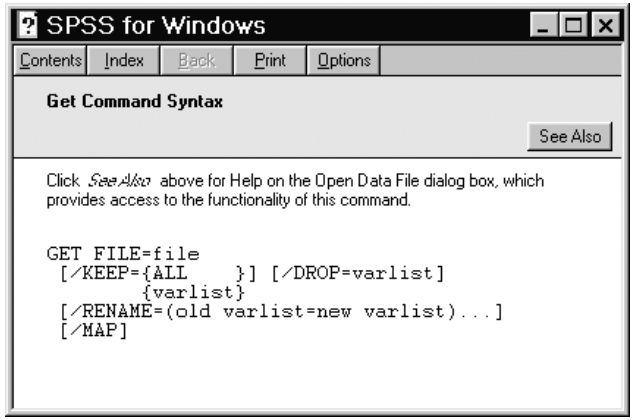
SPSS provides various methods for getting help:

•From the **Help** menu select **Topics**

•For a basic SPSS tutorial select **Help → Tutorial**

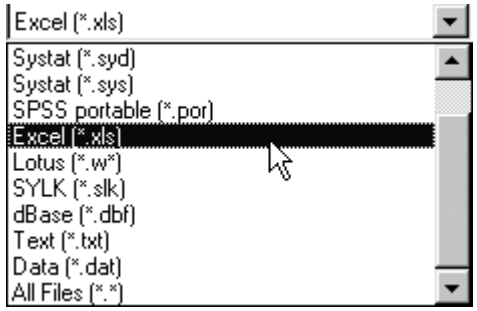
•On a syntax window place the cursor on the command for which you want help then press the button shown below.

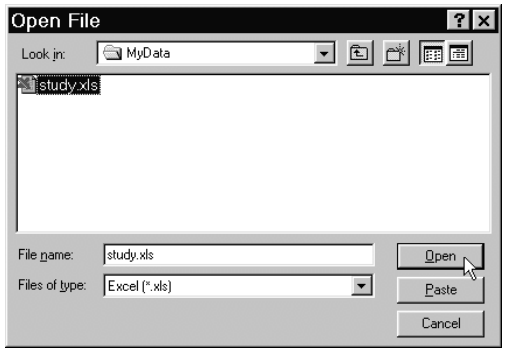


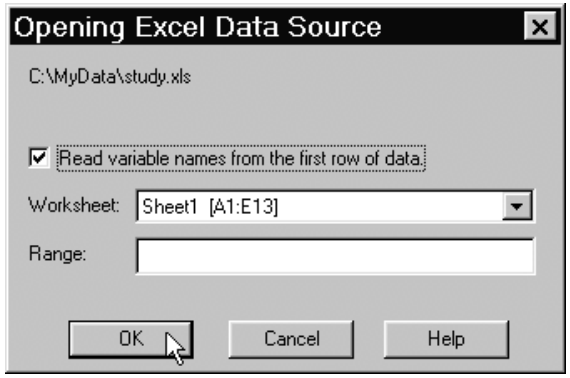


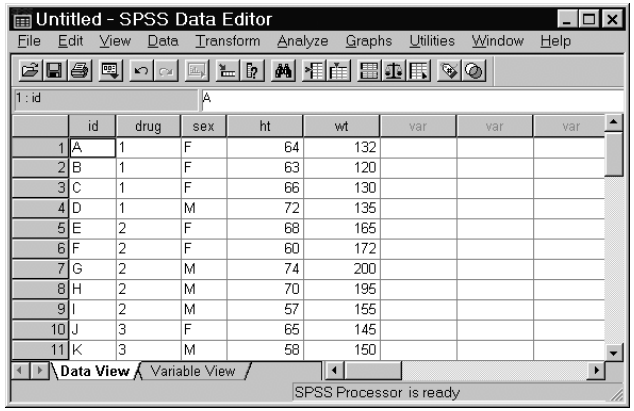
Importing Data

In workshop 1 you learned how to import text data. SPSS can also read data from the sources listed below. Select **File → Open → Data** to open files of these types then click on the drop-down list **Files of type** to select the type. Here we read the Excel file STUDY.XLS.







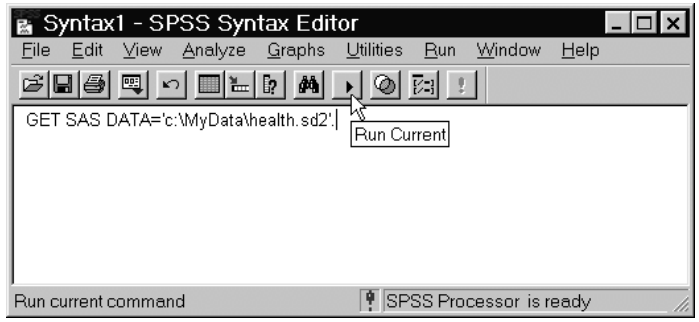


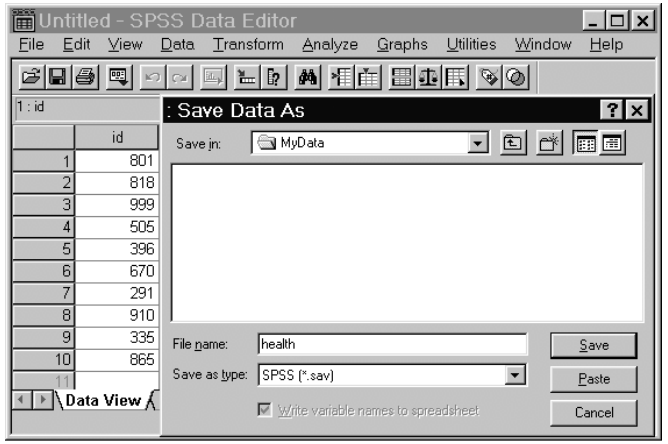
Importing SAS Data

SPSS 10.0 can read SAS data sets up to release 6.12 of SAS.

To read the SAS data set first open a syntax window by selecting **File → New → Syntax**. Then run the **GET SAS** command as in this example:

```
GET SAS DATA='c:\MyData\health.sd2'.
```

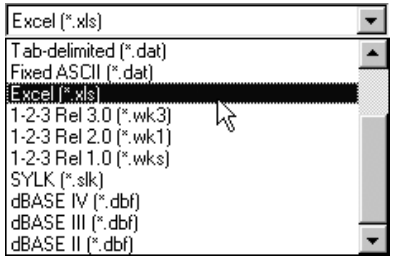


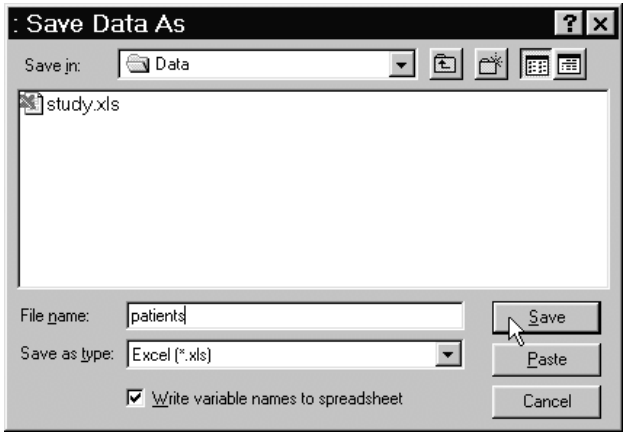


To save the data as an SPSS data file select **File → Save As** and give it a name.

Exporting Data

You can export files by selecting **File → Save As** then selecting the type from the **Save as type** field. Here we export the SPSS data file PATIENTS.SAV to Excel.





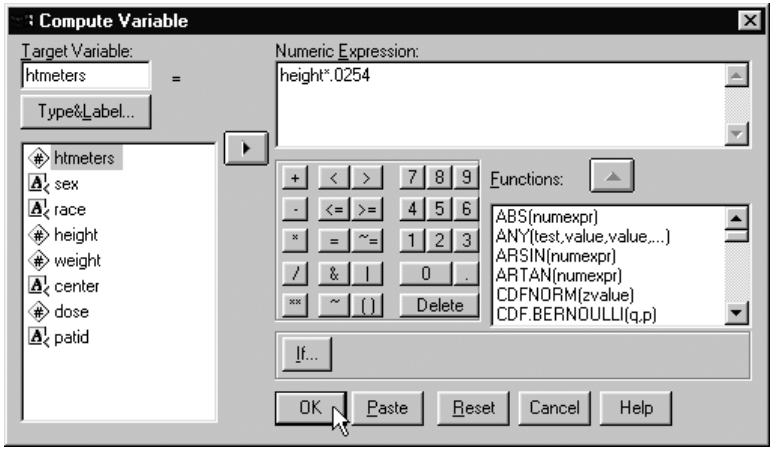
Creating New Variables

You can create new variables by selecting **Transform → Compute**. Here we convert the height to meters and save the new variable as HTMETERS using the SPSS file PATIENTS.SAV.

Enter a name for the new variable in **Target Variable** and an expression in the field **Numeric Expression**. Optionally, select **Type&Label** to select a variable type (default: numeric). When finished, press **OK**.

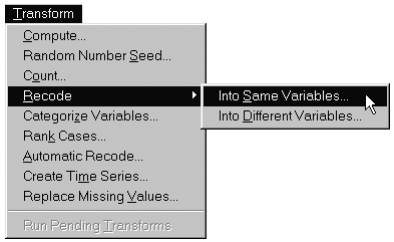
To calculate a new variable with syntax use the **COMPUTE** command:

```
COMPUTE htmeters = height*.0254 .

EXECUTE .
```

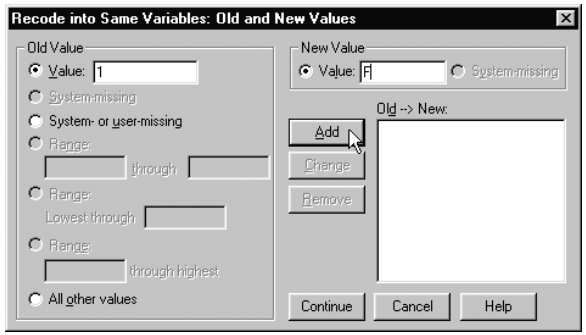


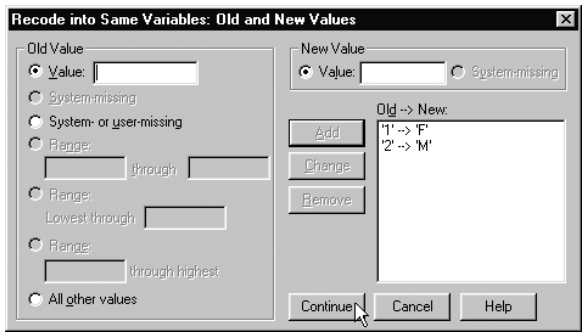
Recoding Variables

To recode a variable select **Transform → Recode**. You have the option to recode **Into Same Variables** or **Into Different Variables**.



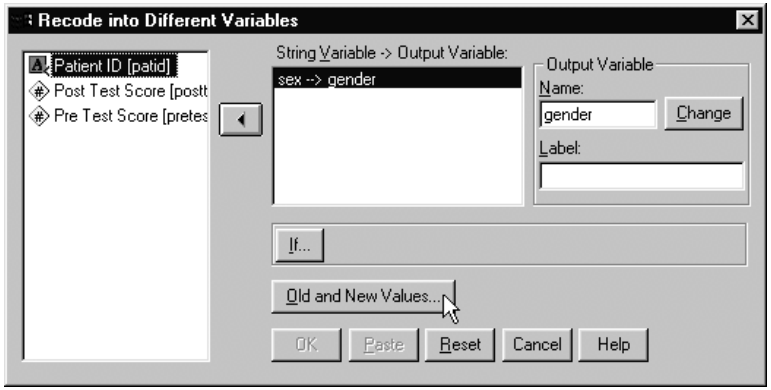
Here we recode the variable SEX into itself. 1 is recoded as F and 2 is recoded as M. In the **Recode Into Same Variable** window select SEX then press the right arrow. Select the **Old and New Values** button. Enter an old value then the corresponding new value then press **Add**. Do this for each value you want to recode. When you are done press **Continue**. On the next window press **OK**.





Recoding Variables

To recode the variable SEX into a different variable choose **Transform → Recode → Into Different Variables**. Select the variable SEX and press the right arrow. In the **Output Variable** field enter a new variable name and press **Change**. Now press **Old and New Values**. Enter old and new values as in the previous example. If the new variable is a string variable you must mark the box **Output variables are strings**. When finished press **OK**.



Recoding Variables

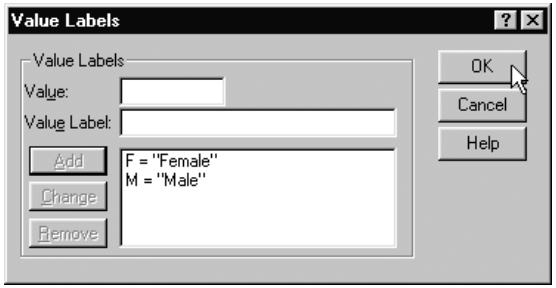
The following command recodes SEX into itself:

```
RECODE  sex  ('1'='F')  ('2'='M') .

EXECUTE .
```

This command recodes SEX into a new variable called GENDER:

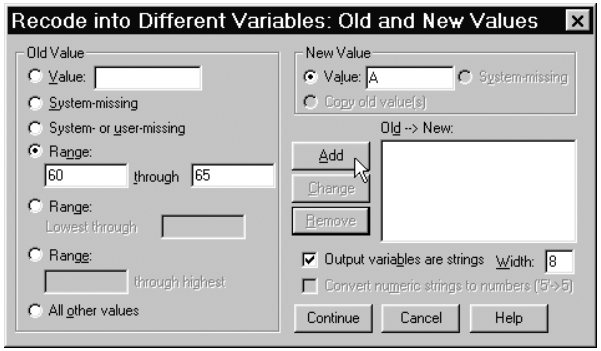
```
RECODE  sex  ('1'='F')  ('2'='M')  INTO  gender.

EXECUTE .
```

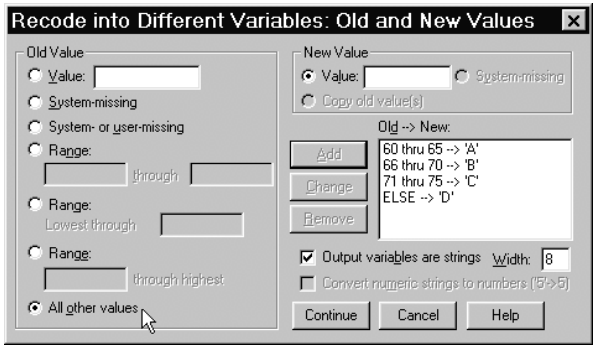
Note: It is important to note that if you recode the values of a variable you may also need to redefine it's value labels to reflect the changes. For example, we should redefine the value labels for SEX so that F=Female and M=Male.



Recoding Variables

When you recode a continuous variable you can specify ranges by selecting the appropriate radio button in the **"Old Value"** section of the **Recode** window.





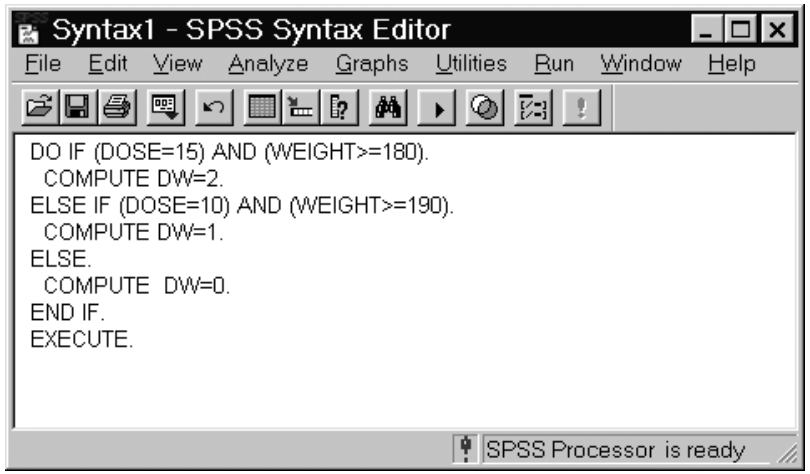
The syntax generated for the above example is:

```
STRING ht (A8) .

RECODE

  height

  (60 thru 65='A')  (66 thru 70='B')  (71 thru 75='C')  (ELSE='D')  INTO  ht .

EXECUTE .
```

DO IF/ELSE IF Commands

To create new variables based on complex conditions that depend on several variables you must use the DO IF/ELSE IF/ELSE commands in a syntax window.

First open a syntax window by selecting **File → New → Syntax**. In this example we create the new variable DW which depends on the variables DOSE and WEIGHT. Run the syntax by selecting **Run → All**.



Dialog Recall Button

With the Dialog Recall Button you can access the most recently used dialog boxes. It is located in the toolbar.

Dialog Recall Button

Sorting the File

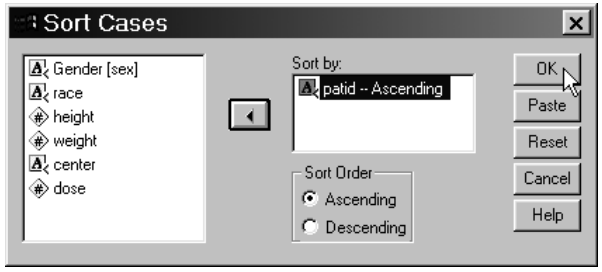
To run some procedures in SPSS the data must be sorted by one or more variables, for example, when you merge files together.

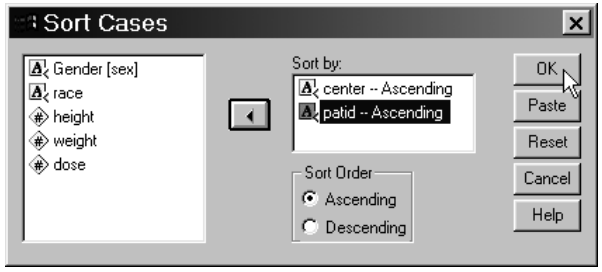
To sort the data, first activate the **Data** window then select **Data → Sort Cases**. Select one or more variables then move then to the **Sort by** box then press **OK**.

SPSS can sort the data in ascending or descending order.

In the first example below we sort by patid.

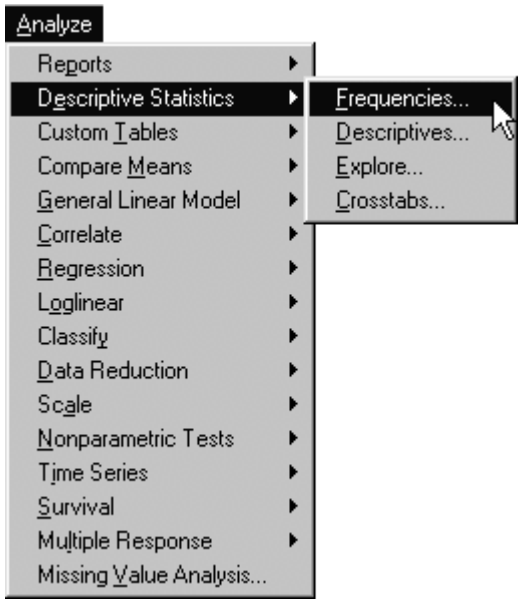
In the second example we sort by center and within each center we sort by patid.





Obtaining Simple Statistics

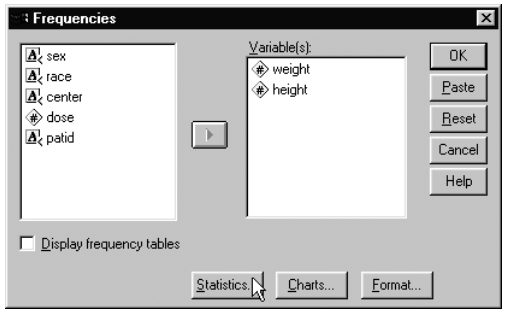
To calculate simple statistics use the **Analyze** menu then select **Descriptive Statistics** then you can select **Frequencies**, **Descriptives**, **Explore** or **Crosstabs**. We'll discuss these in subsequent pages.

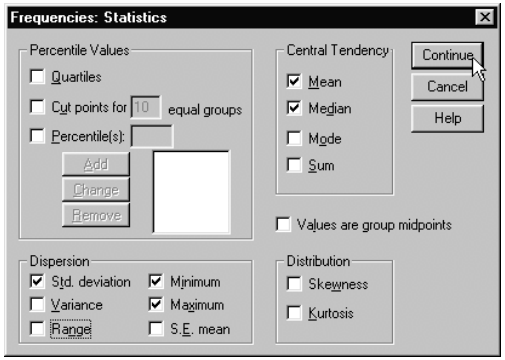


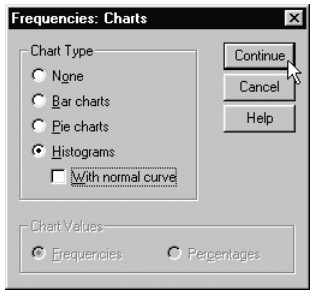
Frequencies Procedure

To calculate frequencies select **Analyze → Descriptive Statistics → Frequencies**. After selecting the variables you can press the **Statistics**, **Charts** or **Format** button for more options. From these windows click **Continue**. When finished press **OK** from the **Frequencies** window.

Here we get simple statistics and histograms of the variables HEIGHT and WEIGHT of the SPSS file PATIENTS .SAV. Since both variables are quantitative we unmarked the box "Display frequency tables".

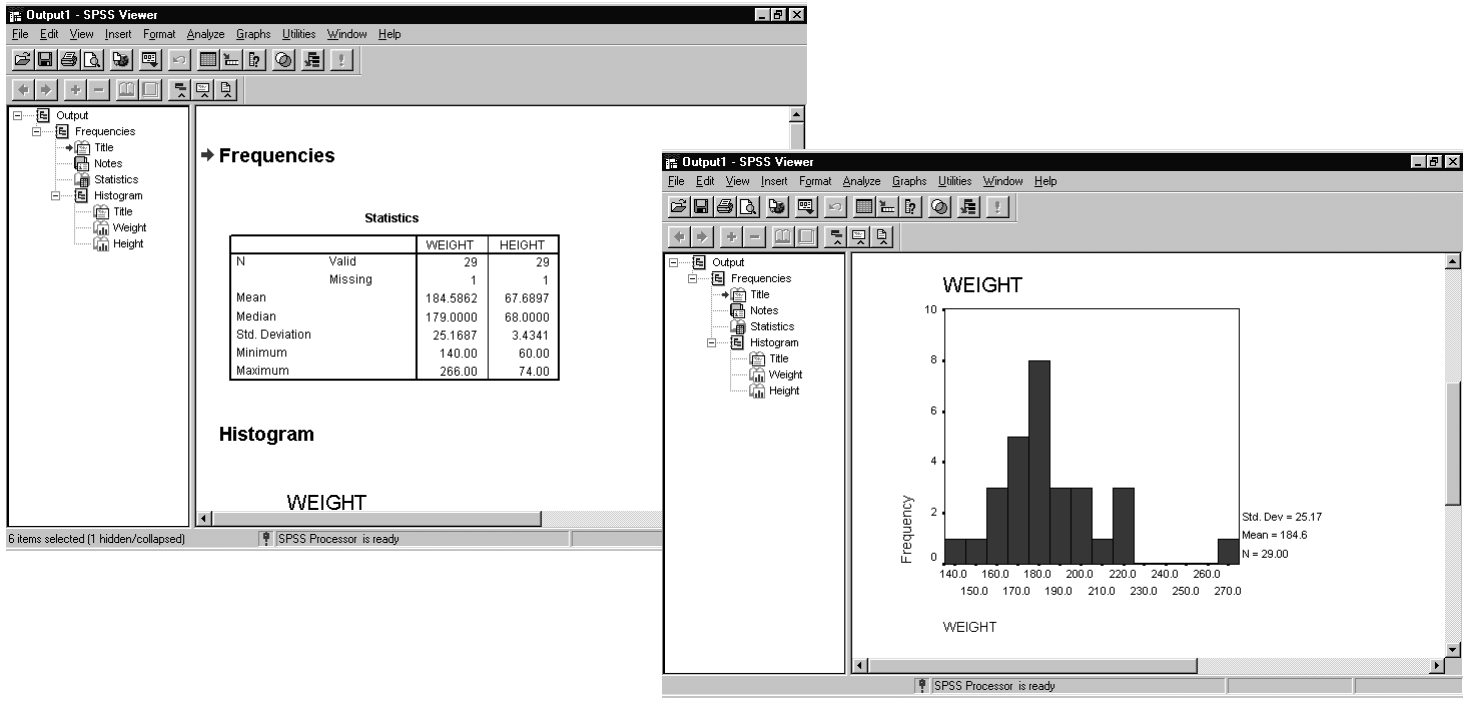






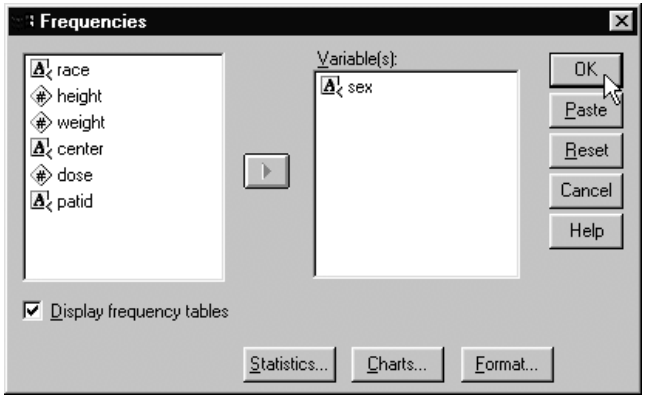
Frequencies Output

The output is placed in the **Viewer** window. This window contains two panes. The outline pane at the left contains a table of contents of all the output produced. It is used for easy navigation and cutting and pasting the various parts of the output. The results of the frequencies is displayed in the right pane.



Frequencies Output

Here we produce a frequency table for the variable SEX. Notice that the box "Display frequency tables" is marked (default).



Frequencies

Statistics

SEX

N	Valid	30
	Missing	0

SEX

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	20	66.7	66.7	66.7
	Female	10	33.3	33.3	100.0
	Total	30	100.0	100.0	

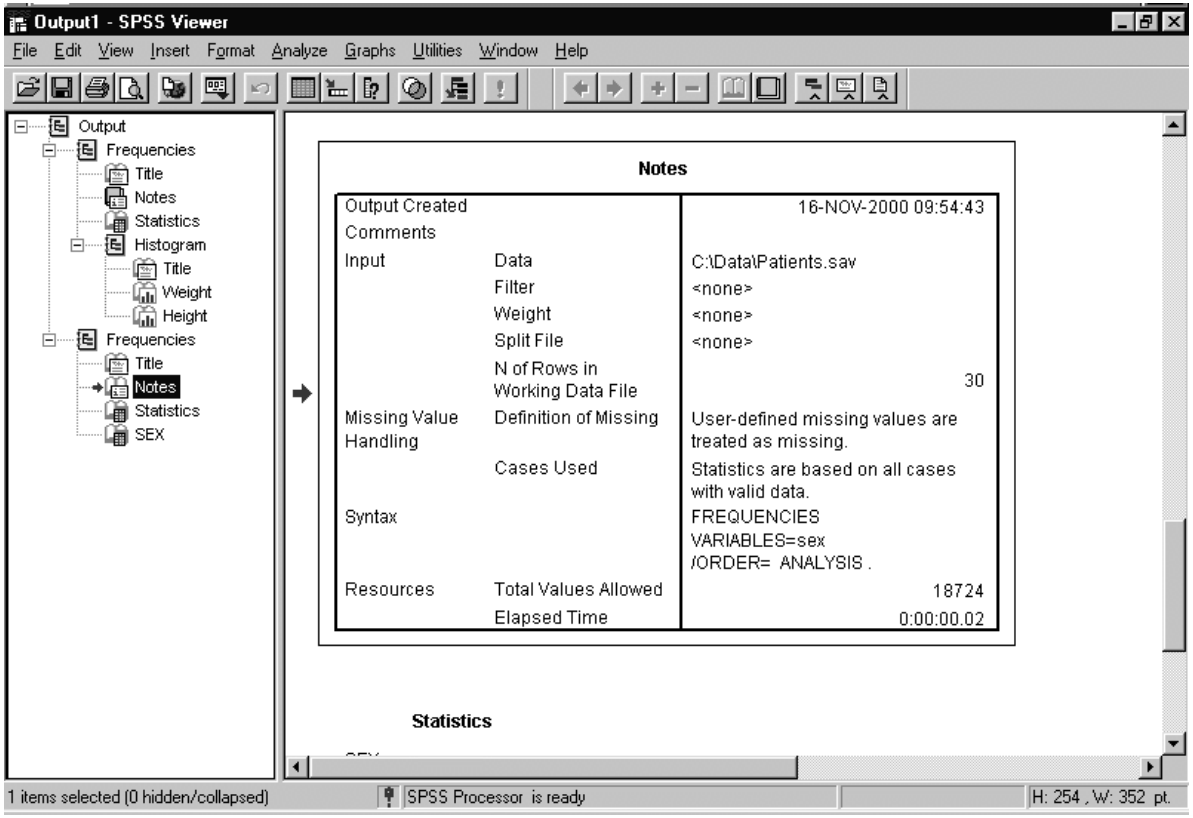
Frequencies Syntax

The syntax produced for the previous two **FREQUENCIES** analysis is:

```
FREQUENCIES
  VARIABLES=weight height  /FORMAT=NOTABLE
  /STATISTICS=STDDEV MINIMUM MAXIMUM MEAN MEDIAN
  /HISTOGRAM
  /ORDER=  ANALYSIS .

FREQUENCIES
  VARIABLES=sex
  /ORDER=  ANALYSIS .
```

Notes Output

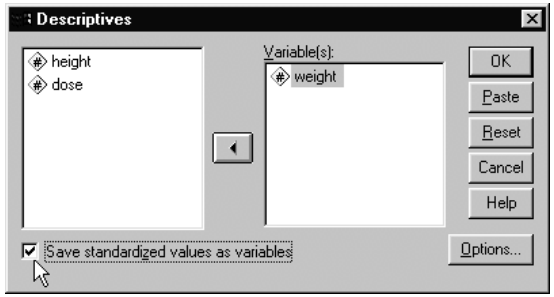
The **Notes** output is initially hidden. It contains information about the analysis and the syntax it used.

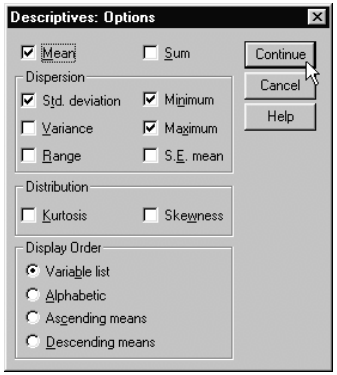


Descriptives Procedure

The **Descriptives** procedure produces simple statistics and it can, optionally, calculate standardized values (z-scores) that are added to your data file.

Here we select **Analyze → Descriptive Statistics → Descriptives** to calculate statistics and z-scores for the variable WEIGHT. Notice that we marked the box "Save standardized values as variables."





Descriptives Output

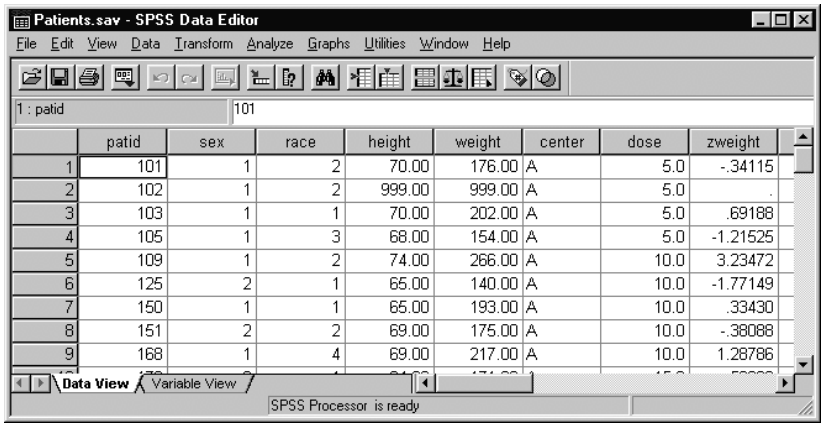
This is the result of using the **Descriptives** procedure.

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
WEIGHT	29	140.00	266.00	184.5862	25.1687
Valid N (listwise)	29				

The data now contains a new variable called ZWEIGHT. To save this variable permanently you must select **File → Save**.



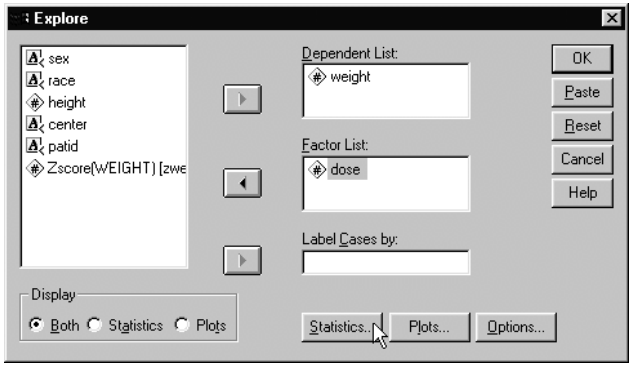
Descriptives Syntax

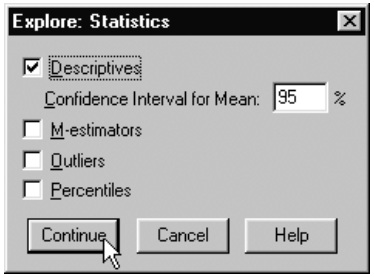
The syntax produced for the previous **DESCRIPTIVES** analysis is:

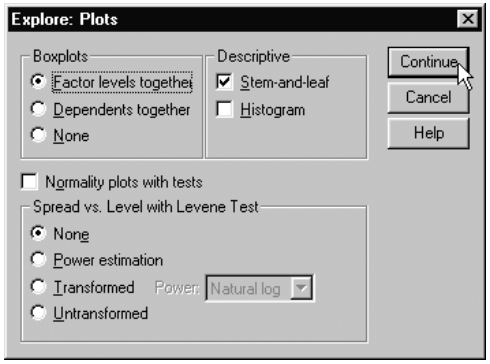
```
DESCRIPTIVES
  VARIABLES=weight  /SAVE
  /STATISTICS=MEAN STDDEV MIN MAX .
```

Explore Procedure

The **Explore** procedure produces summary statistics and graphical displays. Among other options, you can request percentiles, outlier identification, assumption checks and box and whiskers plots. Here we select **Analyze → Descriptive Statistics → Explore** to analyze the variable WEIGHT for each DOSE group. We take all the defaults.







Explore Output

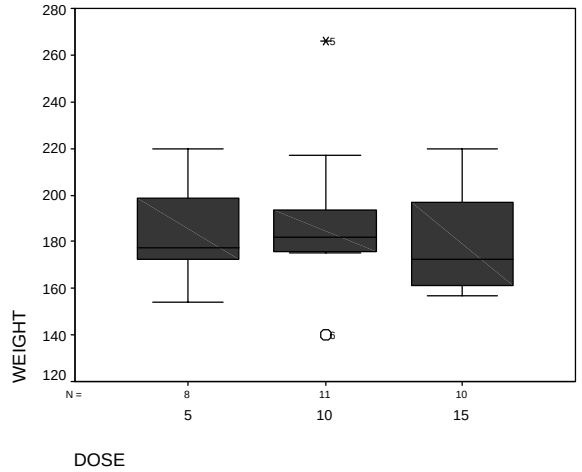
This is part of the output:

Descriptives

	DOSE			Statistic	Std. Error
WEIGHT	5.0	Mean		183.8750	7.3130
		95% Confidence Interval for Mean	Lower Bound	166.5825	
			Upper Bound	201.1675	
		5% Trimmed Mean		183.5278	
		Median		177.5000	
		Variance		427.839	
		Std. Deviation		20.6843	
		Minimum		154	
		Maximum		220	
		Range		66	
		Interquartile Range		28.0000	
		Skewness		.522	.752
		Kurtosis		.085	1.481
	10.0	Mean		189.6364	9.4429
		95% Confidence Interval for Mean	Lower Bound	160.5962	
			Upper Bound	210.6765	
		5% Trimmed Mean		188.1515	
		Median		182.0000	
		Variance		980.855	
		Std. Deviation		31.3186	

Stem-and-Leaf Plots

```
WEIGHT Stem-and-Leaf Plot for
DOSE= 5

 Frequency    Stem &  Leaf

     6.00        1 .  577779
     2.00        2 .  02

 Stem width:  100
 Each leaf:       1 case(s)
```



Explore Syntax

The syntax produced for the previous **EXAMINE** analysis is:

```
EXAMINE
  VARIABLES=weight BY dose
  /PLOT BOXPLOT STEMLEAF
  /COMPARE GROUP
  /STATISTICS DESCRIPTIVES
  /CINTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.
```

Explore Procedure

Suppose you want to compare the average WEIGHT for each combination of DOSE and SEX. We must use syntax for this kind of analysis.

First, select **Analyze → Descriptive Statistics → Explore** and move WEIGHT to the Dependent List and move DOSE and SEX to the Factor List. Paste the results. SPSS produces the following syntax:

```
EXAMINE
  VARIABLES=weight BY dose sex
  /PLOT BOXPLOT STEMLEAF
  /COMPARE GROUP
  /STATISTICS DESCRIPTIVES
  /CINTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.
```

The syntax above produces **separate** analysis for DOSE and SEX.

Explore Procedure

Modify the syntax by adding BY between DOSE and SEX as shown below.

```
EXAMINE
  VARIABLES=weight BY dose BY sex
  /PLOT BOXPLOT STEMLEAF
  /COMPARE GROUP
  /STATISTICS DESCRIPTIVES
  /CINTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.
```

This syntax will produce a **joint** analysis for DOSE and SEX, that is, a separate analysis will be done for each combination of DOSE and SEX.

Workshop 2

1. Create a new variable called **BMI** (Body Mass Index) using this formula: 703*weight/(height*height)
2. Recode the variable WEIGHT into a different variable called **GROUP** as follows:

   100-150 → 1

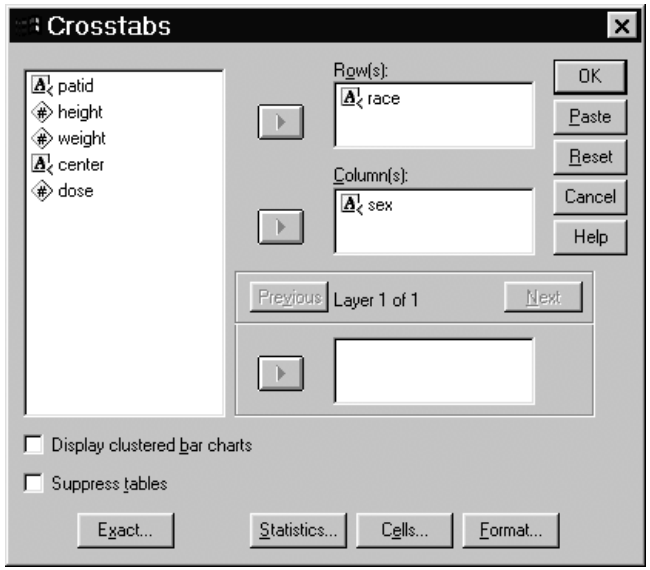
   151-200 → 2

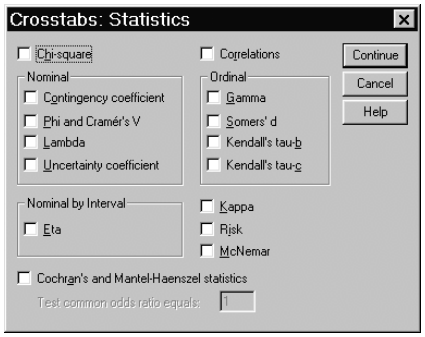
   201-250 → 3

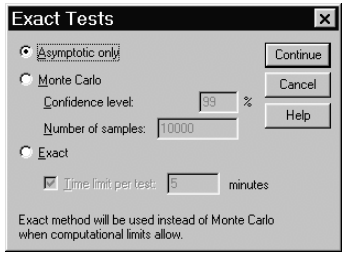
   251 and up → 4
3. Obtain simple statistics for the variable BMI using the **Descriptives** procedure.
4. Obtain simple statistics for the variable BMI for each SEX using the **Explore** procedure.
5. Using the **Frequencies** procedure get frequency tables for the variables: CENTER, DOSE, RACE and SEX.

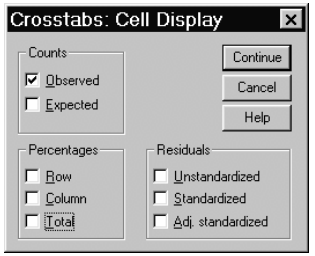
Crosstabs Procedure

The **CROSSTABS** procedure is used to obtain two-way and multi-way frequency tables. You can also request statistical tests and measures of association. Here we request a table of RACE versus SEX and take all the defaults.









Crosstabs Output

This is the two-way table produced:

RACE * SEX Crosstabulation

Count

		SEX		
		Male	Female	Total
RACE	Caucasian	10	7	17
	Black	3	1	4
	Hispanic	2		2
	Asian	4	2	6
	Other	1		1
Total		20	10	30

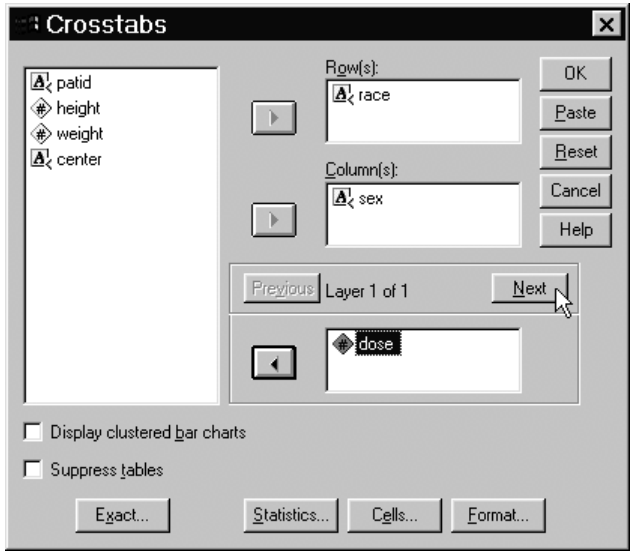
Crosstabs Syntax

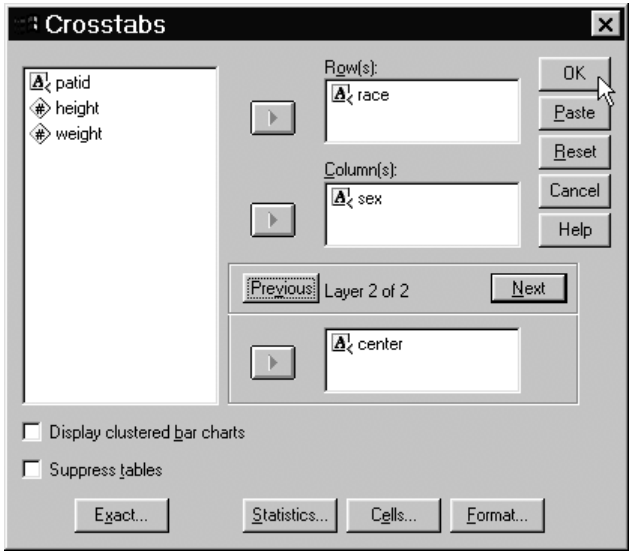
The syntax produced for the previous **CROSSTABS** analysis is:

```
CROSSTABS
  /TABLES=race  BY sex
  /FORMAT= AVALUE TABLES
  /CELLS= COUNT .
```

Crosstabs Layers

You can specify layers to produce multi-layer tables of more than two variables. Here we request a table of RACE versus SEX for each combination of CENTER and DOSE. Use the **Next** button to define the layers.





Crosstabs Layers Output

This is the table produced. It contains the layers CENTER and DOSE.

RACE * SEX * DOSE * CENTER Crosstabulation

Count

CENTER	DOSE			SEX		
				Male	Female	Total
A	5.0	RACE	Caucasian	1		1
			Black	2		2
			Hispanic	1		1
		Total		4		4
	10.0	RACE	Caucasian	1	1	2
			Black	1	1	2
			Asian	1		1
		Total		3	2	5
	15.0	RACE	Caucasian	1	2	3
			Asian	1		1
		Total		2	2	4
B	5.0	RACE	Caucasian	2	2	4
			Asian	1		1
		Total		3	2	5
	10.0	RACE	Caucasian	2	2	4
			Asian		2	2
		Total		2	4	6
	15.0	RACE	Caucasian	3		3
			Hispanic	1		1
			Asian	1		1
			Other	1		1
		Total		6		6

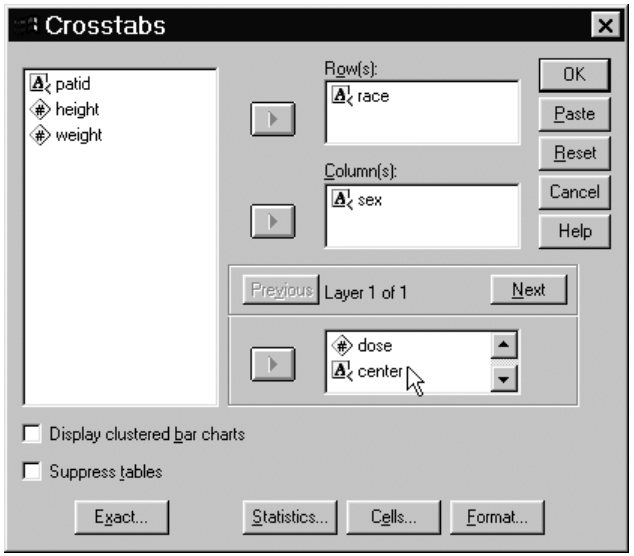
Crosstabs Layers Syntax

Here's the syntax produced to create the table with layers:

```
CROSSTABS
  /TABLES=race  BY sex  BY dose  BY center
  /FORMAT= AVALUE TABLES
  /CELLS= COUNT .
```

Crosstabs Multiple Tables

Specifying more than one variable in the same layer produces separate tables.



RACE * SEX * DOSE Crosstabulation

Count

DOSE			SEX Male	SEX Female	Total
5.0	RACE	Caucasian	3	2	5
		Black	2		2
		Hispanic	1		1
		Asian	1		1
	Total		7	2	9
10.0	RACE	Caucasian	3	3	6
		Black	1	1	2
		Asian	1	2	3
	Total		5	6	11
15.0	RACE	Caucasian	4	2	6
		Hispanic	1		1
		Asian	2		2
		Other	1		1
	Total		8	2	10

RACE * SEX * CENTER Crosstabulation

Count

CENTER			SEX Male	SEX Female	Total
A	RACE	Caucasian	3	3	6
		Black	3	1	4
		Hispanic	1		1
		Asian	2		2
	Total		9	4	13
B	RACE	Caucasian	7	4	11
		Hispanic	1		1
		Asian	2	2	4
		Other	1		1
	Total		11	6	17

Crosstabs Multiple Tables

This is the syntax that produces the multiple tables:

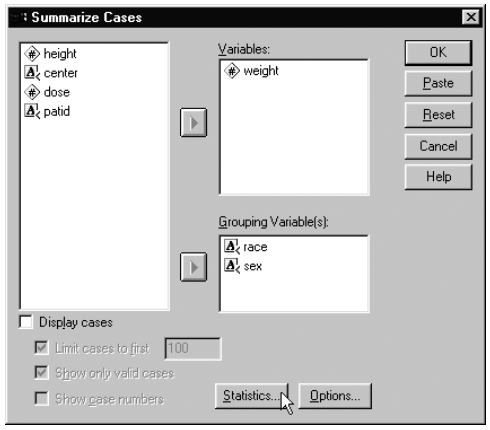
```
CROSSTABS
  /TABLES=race  BY sex  BY dose center
  /FORMAT= AVALUE TABLES
  /CELLS= COUNT .
```

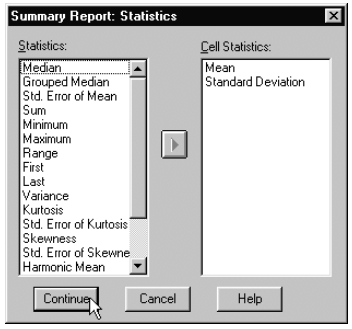
Summarize Procedure

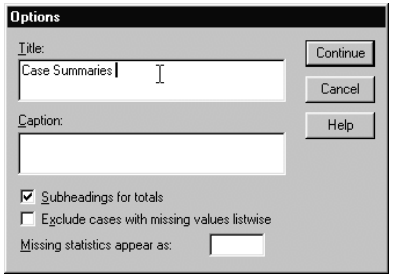
To obtain univariate statistics by group select **Analyze → Reports → Case Summaries.** Select variables from the variable list then choose one or more grouping variables. You can, optionally, press the **Statistics** button to request specific statistics.

Besides creating the table SPSS can display cases. By default, it displays the first 100 cases. This also means it includes only those cases in the analysis.

Here we request the mean and standard deviation of WEIGHT for each combination of RACE and SEX. We unmarked the **Display cases** button.







Summarize Output

Here's the output produced:

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
WEIGHT * RACE * SEX	29	96.7%	1	3.3%	30	100.0%

Case Summaries

WEIGHT

RACE	SEX	Mean	Std. Deviation
Caucasian	Male	193.2000	20.7139
	Female	169.7143	14.0204
	Total	183.5294	21.3779
Black	Male	221.0000	63.6396
	Female	175.0000	.
	Total	205.6667	52.2526
Hispanic	Male	156.0000	2.8284
	Total	156.0000	2.8284
Asian	Male	188.5000	24.0208
	Female	185.5000	12.0208
	Total	187.5000	19.4294
Other	Male	179.0000	.
	Total	179.0000	.
Total	Male	190.4737	28.0007
	Female	173.4000	13.8018
	Total	184.5862	25.1687

Summarize Output

This is a partial result of requesting the mean and standard deviation for WEIGHT for each combination of DOSE, RACE and SEX.

Case Summaries

WEIGHT

DOSE	RACE	SEX	Mean	Std. Deviation
5.0	Caucasian	Male	205.6667	12.8970
		Female	172.5000	.7071
		Total	192.4000	20.3298
	Black	Male	176.0000	.
		Total	176.0000	.
	Hispanic	Male	154.0000	.
		Total	154.0000	.
	Asian	Male	179.0000	.
		Total	179.0000	.
	Total	Male	187.6667	23.0188
		Female	172.5000	.7071
		Total	183.8750	20.6843
10.0	Caucasian	Male	184.6667	9.0738
		Female	167.6667	23.9653
		Total	176.1667	18.6914
	Black	Male	266.0000	.
		Female	175.0000	.
		Total	220.5000	64.3467
	Asian	Male	217.0000	.
		Female	185.5000	12.0208
		Total	196.0000	20.0749

Summarize Syntax

This is the syntax generated for the previous two examples:

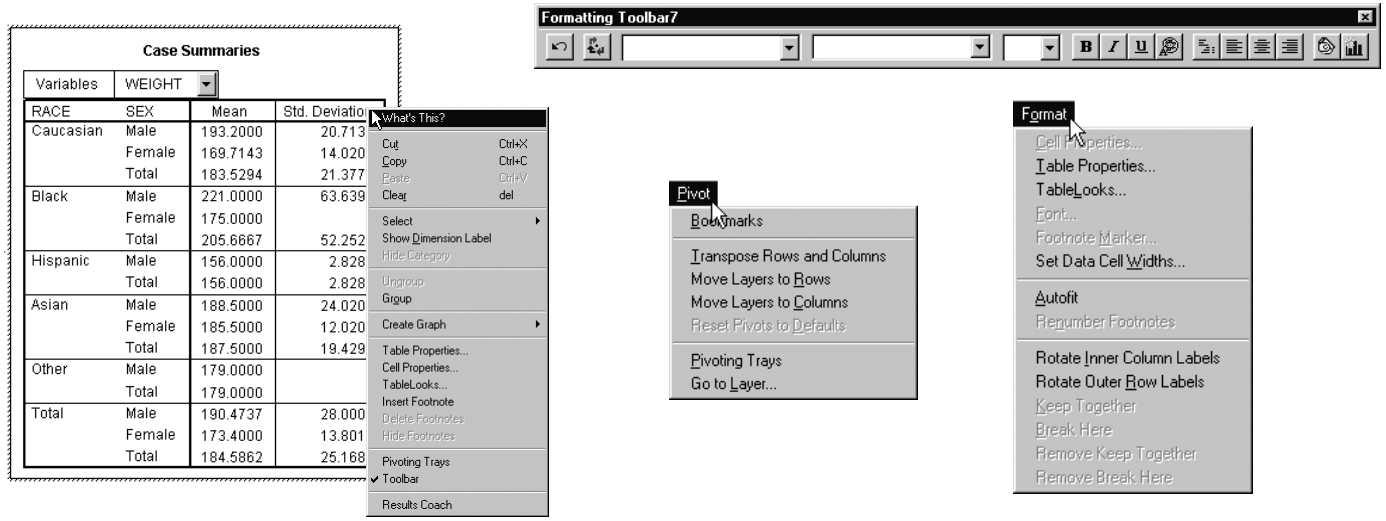
```
SUMMARIZE
  /TABLES=weight  BY race BY sex
  /FORMAT=NOLIST TOTAL
  /TITLE='Case Summaries'
  /MISSING=VARIABLE
  /CELLS=MEAN STDDEV .

SUMMARIZE
  /TABLES=weight  BY dose BY race BY sex
  /FORMAT=NOLIST TOTAL
  /TITLE='Case Summaries'
  /MISSING=VARIABLE
  /CELLS=MEAN STDDEV .
```

Pivot Tables

SPSS 10.0 includes a new feature called **pivot tables**. With it you can easily modify and enhance any table created in SPSS. For example, you can transpose the rows and columns, change the format of specific areas of the table, select the looks of your table and create a graph using the table data.

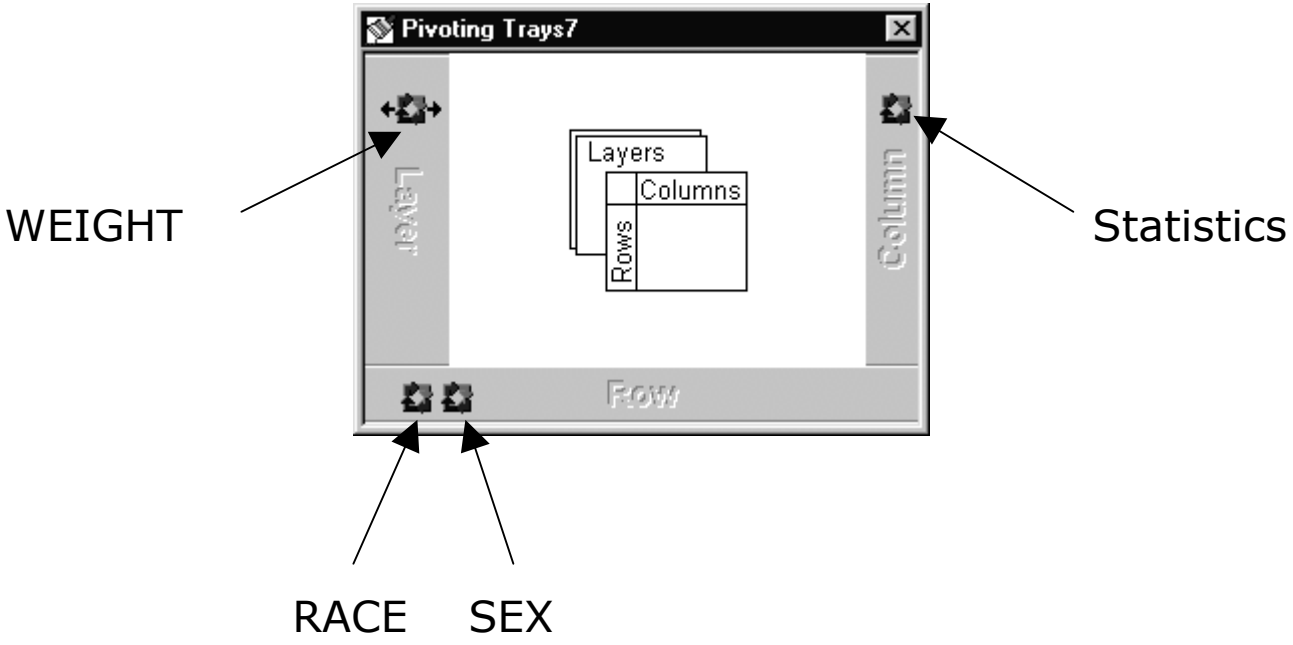
To use Pivot Tables, first activate the table by double-clicking on it. A **Formatting Toolbar** will open. If it doesn't then select **View → Toolbar**.

Most of the time you will be using the **Pivot** and **Format** menus and the popup menus within the table.



Pivot Tables

When you select **Pivot → Pivoting Trays** the **Pivoting Trays** window opens. The icons in the window represent the variables and statistics in the table. You can arrange the variables and statistics differently by dragging the icons to different dimensions.



Pivot Tables

Dragging the icon for SEX to the **Column** dimension and before the icon for Statistics produces the following table:

WEIGHT

	SEX					
	Male		Female		Total	
RACE	Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation
Caucasian	193.2000	20.7139	169.7143	14.0204	183.5294	21.3779
Black	221.0000	63.6396	175.0000	.	205.6667	52.2526
Hispanic	156.0000	2.8284			156.0000	2.8284
Asian	188.5000	24.0208	185.5000	12.0208	187.5000	19.4294
Other	179.0000	.			179.0000	.
Total	190.4737	28.0007	173.4000	13.8018	184.5862	25.1687

By dragging the icon for SEX to the **Layer** dimension and before the icon for WEIGHT produces the following two layers (you view one at a time):

SEX: Female

WEIGHT

RACE	Mean	Std. Deviation
Caucasian	169.7143	14.0204
Black	175.0000	.
Asian	185.5000	12.0208
Total	173.4000	13.8018

SEX: Male

WEIGHT

RACE	Mean	Std. Deviation
Caucasian	193.2000	20.7139
Black	221.0000	63.6396
Hispanic	156.0000	2.8284
Asian	188.5000	24.0208
Other	179.0000	.
Total	190.4737	28.0007

Pivot Tables

You can select from a list of table styles by selecting **Format → Table Looks** then selecting the style of your preference. The one shown below is the Academic style.

In the table below we also modified the row Total. To do this, first right-click the label Total then choose **Select → Data Cells**. Then choose Italic from the formatting toolbar. We also made the label bold.

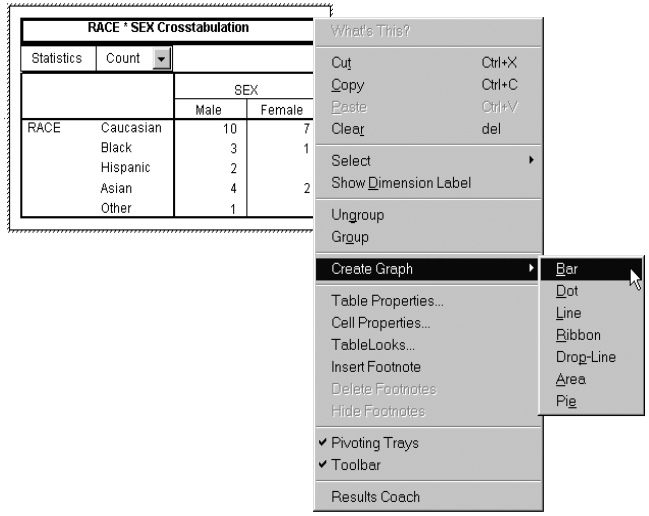
WEIGHT

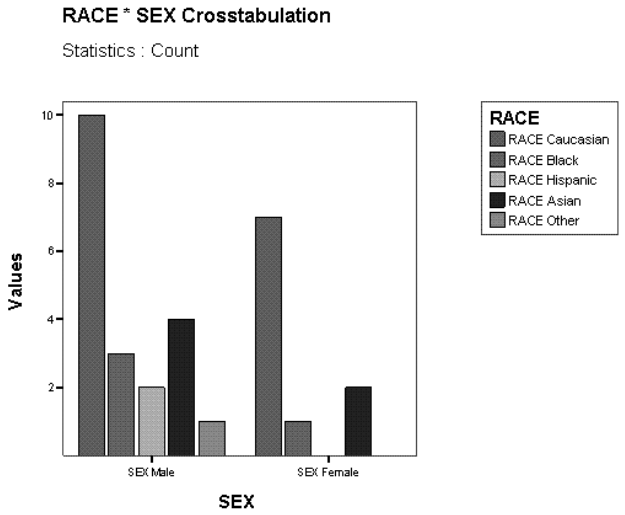
	SEX					
	Male		Female		Total	
RACE	Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation
Caucasian	193.2000	20.7139	169.7143	14.0204	183.5294	21.3779
Black	221.0000	63.6396	175.0000	.	205.6667	52.2526
Hispanic	156.0000	2.8284			156.0000	2.8284
Asian	188.5000	24.0208	185.5000	12.0208	187.5000	19.4294
Other	179.0000	.			179.0000	.
Total	*190.4737*	*28.0007*	*173.4000*	*13.8018*	*184.5862*	*25.1687*

Pivot Tables

You can generate a graph from a table by first double-clicking on the table then right-clicking anywhere in the table. Then select **Create Graph** and select the type of graph you want.

This is the bar chart produced from a frequency table of RACE by SEX.





Workshop 3

1. Using the CROSSTABS procedure create frequency tables with counts of patients for each combination of SEX, GROUP and DOSE as shown below.

GROUP * DOSE * SEX Crosstabulation

Count

SEX			DOSE 5.0	DOSE 10.0	DOSE 15.0	Total
Male	GROUP	2	4	3	6	13
		3	2	1	2	5
		4	1	1		2
	Total		7	5	8	20
Female	GROUP	1		1		1
		2	2	5	2	9
	Total		2	6	2	10

2. Using the SUMMARIZE procedure create a table that contains the average weight for each combination of SEX and DOSE as shown below.

Case Summaries

Mean

SEX	DOSE	WEIGHT
Male	5.0	187.6667
	10.0	207.4000
	15.0	182.0000
	Total	190.4737
Female	5.0	172.5000
	10.0	174.8333
	15.0	170.0000
	Total	173.4000
Total	5.0	183.8750
	10.0	189.6364
	15.0	179.6000
	Total	184.5862

Workshop 3

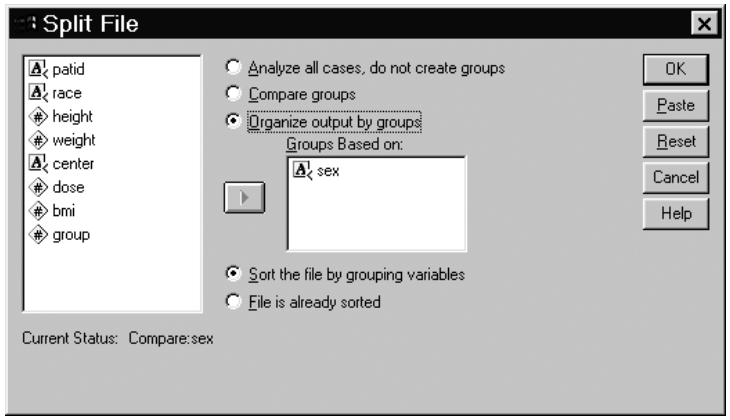
3. Use pivot tables to modify the table from exercise 2 as you wish. Here's an example.

	Average Weight			
	DOSE			
SEX	**5.0**	**10.0**	**15.0**	**Total**
Male	187.67	207.40	182.00	190.47
Female	172.50	174.83	170.00	173.40
Total	183.88	189.64	179.60	184.59

Split File

You can analyze your data by group by selecting **Data → Split File** from the **Data Editor** window. All the analysis done from that point on will be done for each group separately. You can specify up to eight grouping variables.

For example, here we split the group by SEX.



Descriptives

SEX = Male

Descriptive Statistics[a]

	N	Minimum	Maximum	Mean	Std. Deviation
HEIGHT	19	65.00	74.00	69.3684	2.5432
WEIGHT	19	154.00	266.00	190.4737	28.0007
Valid N (listwise)	19				

a. SEX = Male

SEX = Female

Descriptive Statistics[a]

	N	Minimum	Maximum	Mean	Std. Deviation
HEIGHT	10	60.00	69.00	64.5000	2.5495
WEIGHT	10	140.00	194.00	173.4000	13.8018
Valid N (listwise)	10				

a. SEX = Female

Split File

If you select "**Compare Groups**" above the tables are presented together:

Descriptives

Descriptive Statistics

SEX		N	Minimum	Maximum	Mean	Std. Deviation
Male	HEIGHT	19	65.00	74.00	69.3684	2.5432
	WEIGHT	19	154.00	266.00	190.4737	28.0007
	Valid N (listwise)	19				
Female	HEIGHT	10	60.00	69.00	64.5000	2.5495
	WEIGHT	10	140.00	194.00	173.4000	13.8018
	Valid N (listwise)	10				

SPSS will automatically sort the data by the grouping variables. If your data is already sorted then select the radio button "**File is already sorted**."

Split File

The syntax for the previous two examples is:

SORT CASES BY sex .

SPLIT FILE

SEPARATE BY sex .

SORT CASES BY sex .

SPLIT FILE

LAYERED BY sex .

To turn splitting off:

SPLIT FILE OFF.

Joining Files

You can join SPSS files in two ways:

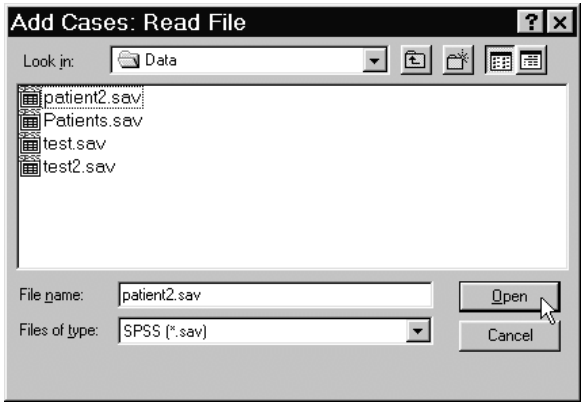
- join files containing the same variables but different cases, or,
- join files containing the same cases but different variables.

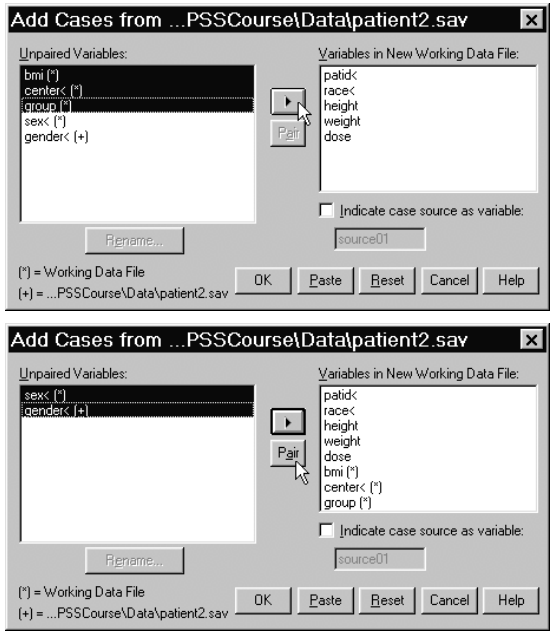
To join files select **Data → Merge Files**. Then select either **Add Cases** or **Add Variables**.

After joining files you must select **File → Save** or **File → Save As** if you want your changes to be saved permanently.

Adding Cases

In this example we will add the cases from the SPSS file PATIENT2.SAV to the current copy of PATIENTS.SAV.





Unpaired variables are variables that are found in one of the SPSS files and not the other. By default, they are not kept. Here we will mark them and move them to the list "**Variables in New Working Data File**." The variables SEX and GENDER represent the same variable so we mark both then select **Pair**. The variable name from PATIENTS.SAV will be used by default.

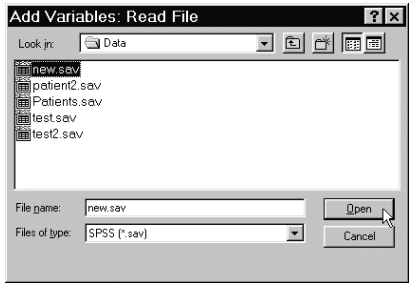
Adding Cases

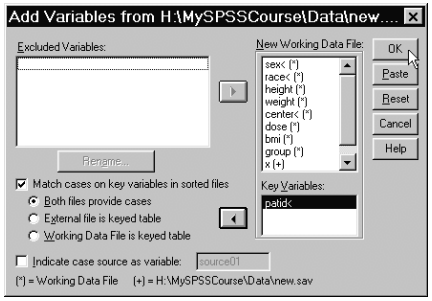
To add cases using syntax use the **ADD FILES** command. The following syntax was generated automatically for the previous example:

```
ADD FILES /FILE=*
 /FILE='c:\Data\patient2.sav'
 /RENAME  gender=sex.
EXECUTE.
```

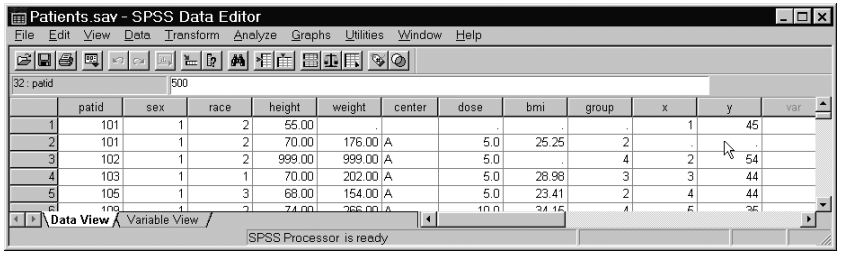
Adding Variables

To add variables to an SPSS file from a separate SPSS file select **Data → Merge Files → Add Variables**. Here we add the variables from the SPSS file NEW.SAV to the PATIENTS.SAV file. We choose PATID to be the variable SPSS will use to match the cases from both files.



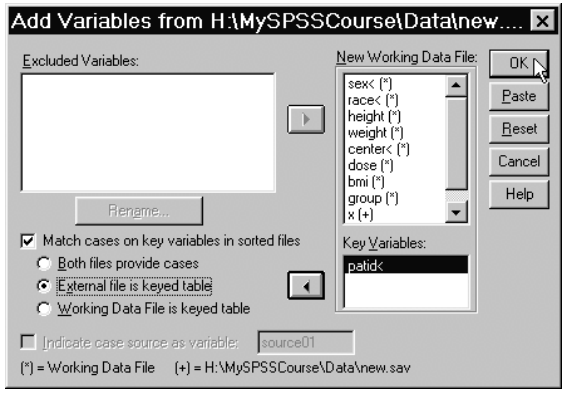


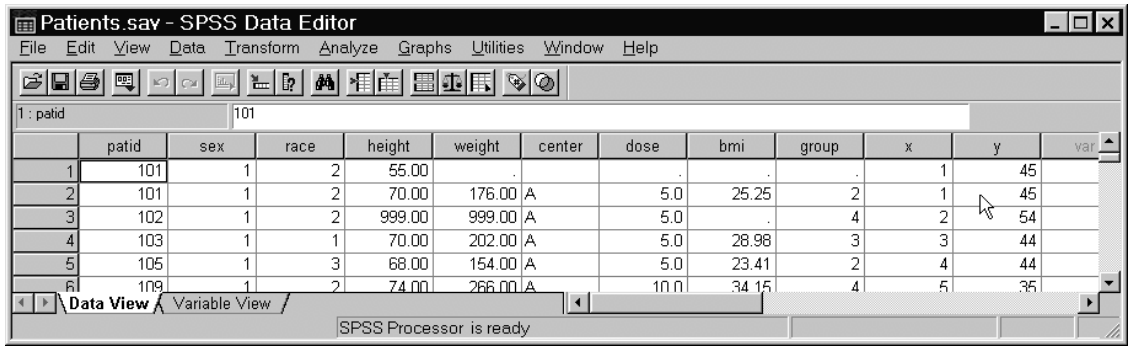
If you have multiple cases per PATID it is best to use **keyed table** otherwise you may get incorrect results as in this example. Notice that the second case (PATID 101) does not have values for the new variables X and Y. Keyed tables cannot contain duplicate cases.



Adding Variables

Here we add the variables from the SPSS file NEW.SAV to the PATIENTS.SAV file but we use the file NEW.SAV as a **keyed table**. Notice that the resulting file is now correct for PATID 101.





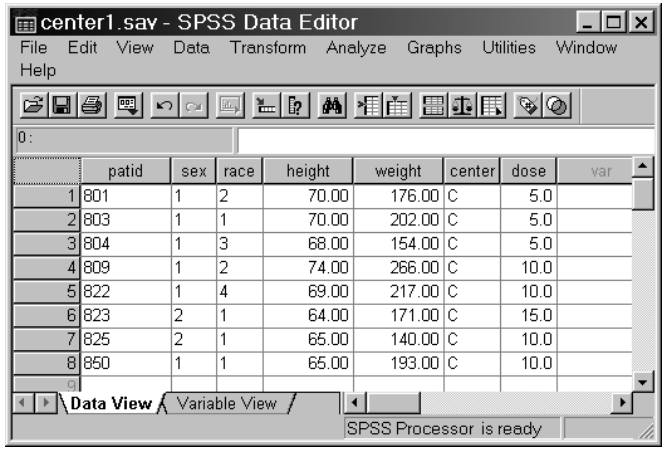
Adding Variables

To add variables using syntax use the **MATCH FILES** command. The following syntax was generated automatically for the previous examples:

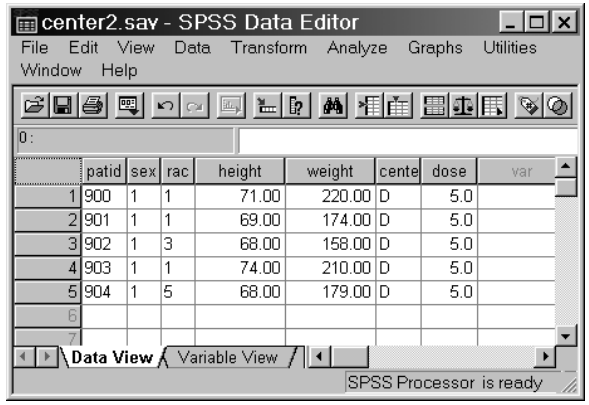
```
*with keyed table.
MATCH FILES /FILE=*
 /TABLE='H:\MySPSSCourse\Data\new.sav'
 /BY patid.
EXECUTE.

*without keyed table.
MATCH FILES /FILE=*
 /FILE='H:\MySPSSCourse\Data\new.sav'
 /BY patid.
EXECUTE.
```

Workshop 4

1. Stack the two SPSS files CENTER1 and CENTER2. They contain the same variables but different cases.



CENTER1.SAV



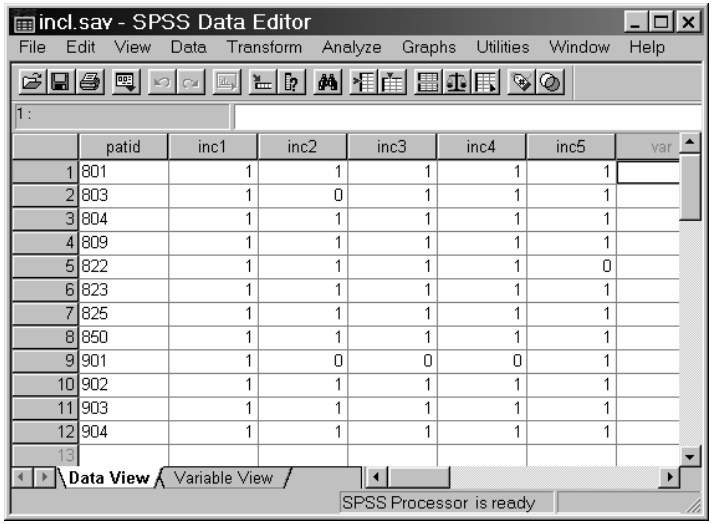
CENTER2.SAV

2. Save the resulting SPSS file as CENTERS.SAV.

3. Add the variables from the SPSS file INCL.SAV to the right of the file you just created. Resave the file.

Workshop 4

2. Add the variables from the SPSS file INCL.SAV to the right of the file you just create it. Resave the file.



INCL.SAV

SPSS Graphics

SPSS provides two methods for creating most graphics: standard and interactive. The main difference between them is that standard graphics lack many of the interactive features available through the interactive graphs.

Interactive graphics is a new method in SPSS to create graphics. To create interactive graphics select **Graphs → Interactive** then select the type of graphics you want.

We will demonstrate these types now. We will use the SPSS file "Employee data.sav" that is stored in the SPSS directory.

- bar
- boxplot
- error bar
- histogram
- Scatterplot
- line

For More SPSS Courses

SPSS, Inc. offers many SPSS courses from the basics to the advanced including many statistics courses. These courses are taught at many locations including their office in Arlington, VA.

Check their website to view the training schedule:

http://www.spss.com