

Opening a Data File in SPSS

To open an existing SPSS file:

1. Click **File** → **Open** → **Data**. Go to the appropriate directory and find the name of the appropriate file. SPSS defaults to opening SPSS data files with the .sav extension.

The SPSS Data Editor resembles the spreadsheet-like structure of Excel. The Data Editor window consists of:

- **Row:** A horizontal collection of information that represents a case or an observation. For example, each individual respondent to a questionnaire is a case. It is the same as a record in Access.
- **Column:** A vertical collection of information that represents a variable or characteristic being measured. For example, each item on a questionnaire is a variable. It is the same as a field in Access.
- **Cell:** This is the intersection of a row and a column. It is where a value for a particular variable for a particular case is entered. Cells contain only data values. Unlike spreadsheet programs, cells in the Data Editor cannot contain formulas.

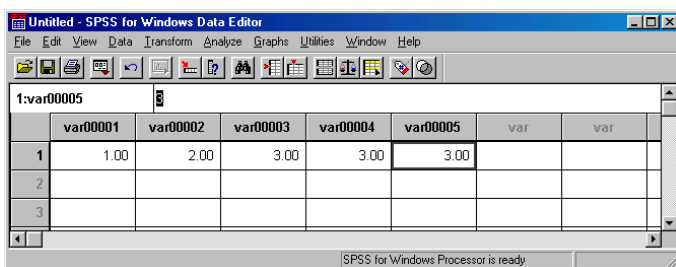
Defining Variables in SPSS

You can have SPSS define the variables for you, or preferably, you can do the defining yourself. By doing it yourself, you have much more control over the way things look and work.

Having SPSS Define Variables

If you enter data into a cell without defining a variable, SPSS will automatically name the variable. If the variable is in column 1, the variable name assigned would be **var00001**; if the variable name is in column 4, the name assigned would be **var00004**; etc.

You should also note that right below the main menu in the upper left-hand corner of the Data Editor is an indicator of what cell is being highlighted. For example, in the figure below, you can see that **1:VAR00005** represents row 1, variable 5.



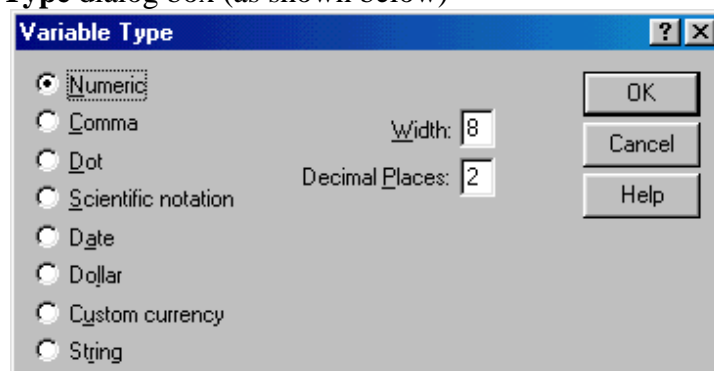
Defining SPSS Variables Yourself

By custom defining variables yourself, it allows you for greater flexibility, such as:

- providing your own names for variables,
- providing descriptive labels
- assigning special formats such as text or date or time.

To define a variable name, make sure the Data View window is active and follow these steps:

1. Click on the **Variable View** tab at the bottom of the screen.
2. **Name:** Enter a name for the variable. The following rules apply to variable names:
 - The name you assign to a variable can have no more than 8 characters.
 - It must begin with a letter. The remaining characters can be of any letter, any digit, a period, or the symbols @, #, _, \$.
 - Variable names cannot end with a period.
 - Variable names that end with an underscore should be avoided.
 - Blanks and special characters (for example, !, ?, ', and *) cannot be used.
 - Each variable name must be unique; duplication is not allowed. Variable names are not case sensitive. The names NEWVAR, NewVar, and newvar are all the same.
3. **Type:** The default variable type is Numeric. If you click within the Type cell, you will see the **Define Variable Type** dialog box (as shown below)



You can choose from one of the following variable definitions, to include:

- numeric (e.g., 3435.45)
- comma (e.g., 3,435.45)
- dot (e.g., align such as 3.43545)
- scientific notation (e.g., 3.4E+03)
- date (e.g., 12-FEB-2000)
- dollar (e.g., \$3,435.45)
- custom currency (as you design it)
- string (text such as William, JMU, 201 Friday Street, etc.)

You can also set the width of the field as well as the number of decimal places you wish to use.

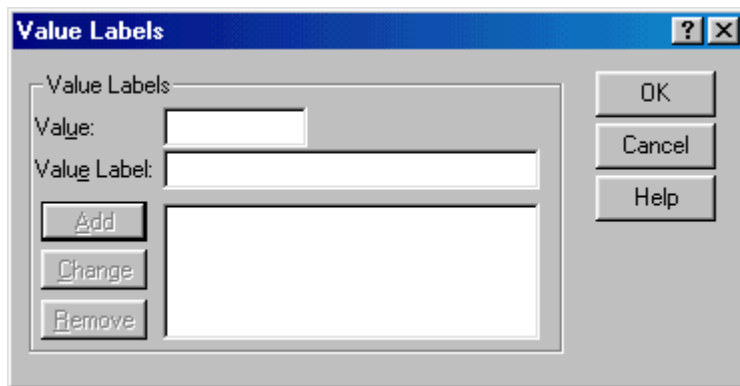
4. **Label:** The **Label** option allows you to provide descriptive variable labels (a variable label is makes it much easier to read your SPSS output. SPSS, by default, does not provide variable labels. A variable label allows you to more adequately describe a variable. It is hard to describe a variable in the 8 characters used in a variable name. A variable can be up to 256 characters long and these descriptive labels are displayed in output.

To insert a variable label, click in the **Label** cell and type in your variable label.

5. **Values:** You can also assign descriptive value labels for each value of a variable. This is particularly useful if your data file uses numeric codes to represent non-numeric categories (for example, where a 1 = Male and a 2 = Female). Value labels can be up to 60 characters long.

To assign or change a variable label, follow these steps:

- Click inside the **Values** cell.
- Click on the small gray box in the right corner. You will now see the following **Value Labels** box:



- Enter a value for the variable. This is what will appear in the Data View window, such as 1, 2, or 3.
 - Enter the value label for the value.
 - Click **Add**. The value label will now appear. Now the same steps need to be taken to apply a label for the next value.
 - Click **OK** when you have added all of the labels for the variable and the new labels will take effect.
6. **Missing:** If you click inside the **Missing** cell, you can tell SPSS how to deal with the missing values that you specify. The **No missing values** option treats all values as being valid, such that no missing values are present in the data file (this is the SPSS default). That is, only blanks are treated as missing variables in the data file. The **Discrete missing values** option allows you to enter up to three values for a missing variable. The **Range of missing values** option allows you to define a range and include high and low values as missing. Finally the **Range plus one discrete missing value** option will identify as missing all the values between the low and high values you identify, plus one additional value outside the range.
 7. **Column:** **Column** lets you control the width of the columns in the Data View Column widths can also be changed in the Data View by clicking and dragging the column borders.

8. **Align:** **Align** lets you line up the contents in a cell. You can set this up as left, right, or center.
8. **Measurement:** You can specify the level of measurement as scale (numeric data on an interval or ratio scale), ordinal, or nominal. Nominal and ordinal data can be either string (alphanumeric) or numeric. Measurement specification is relevant only for:
 - Chart procedures that identify variables as scale or categorical. Nominal and ordinal are both treated as categorical.

Entering and Editing Data in SPSS

You can enter data directly in the Data View. You can enter data in any order. You can enter data by case or by variable, for selected areas or for individual cells.

- The active cell is highlighted with a heavy border.
- The variable name and row number of the active cell are displayed in the top left corner of the Data Editor.
- When you select a cell and enter a data value, the value is displayed in the cell editor at the top of the Data Editor.
- Data values are not recorded until you press Enter or select another cell.
- You can move from variable to the next by pressing the Tab key.
- To enter anything other than simple numeric data, you must define the variable type first.
- You cannot enter a string or text variable unless you have already defined the variable as a String type.

To Delete the Old Value and Enter a New Value:

- Click the cell. The cell value is displayed in the cell editor.
- Enter the new value. It replaces the old value in the cell editor.
- Press Enter (or move to another cell) to record the new value.

To Modify a Data Value:

- Click the cell. The cell value appears in the cell editor.
- Click the cell editor.
- Edit the data value as you would any other text.
- Press Enter (or select another cell) to record the modified value.

Data Transformations

There are often situations where you need to take existing data and convert them into a different variable or to take existing variables and combine them to form an additional variable. For example, you may want to **compute** the mean or standard deviation for a set of variables and enter those values in a separate variable in the data file. Or, you may want to **recode** data so that, for example, all values of 10 or greater equal 1, and all values less than 10 equal 0. Your needs for transforming or recoding data will depend on the nature of your data and the purpose of your analysis.

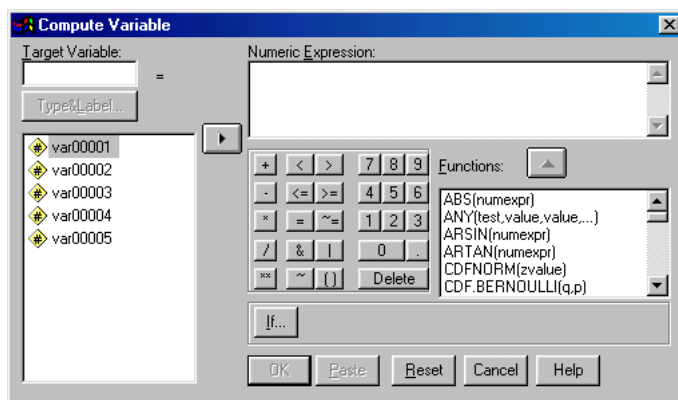
Computing Variables

Computing a variable means taking existing data and using them to create a new variable in the form of a numeric expression. Compute Variable computes values for a variable based on numeric transformations of other variables.

- You can compute values for numeric or string (alphanumeric) variables.
- You can create new variables or replace the values of existing variables. For new variables, you can also specify the variable type and label.
- You can compute values selectively for subsets of data based on logical conditions.
- You can use over 70 built-in functions, including arithmetic functions, statistical functions, distribution functions, and string functions.

To Compute Variables Using a Formula

1. Make sure the Data Editor window is active.
2. Click **Transform** → **Compute**. You will see the screen shown below:



3. Click the **Target Variable** text box and enter the name of a new variable to be created.
4. Let's say, for example that we want to add var00001 and var00003 to create a new variable. You would first click on **var00001**, the first variable you want to include in the formula.
5. Click on the ► to add it to the Numeric Expression text box.
6. Type + (the plus sign for our example or whatever numeric operator you need).

7. Add the second variable by clicking on **var00003**. Then click on ► to move the variable over to the Numeric Expression text box.
8. Now that the entire formula has been entered, click **OK**. The new variable will be added to the end of your existing data file (it will be the last column in the Data Editor window).

To Compute a Variable Using a Function

You can also compute a variable using a function. Functions are predesigned formulas. So instead of

var00001+var00003

you can enter

SUM(var00001,var00003)

which will accomplish the same thing. There are more than 70 functions, including arithmetic functions, statistical functions, distribution functions, logical functions, data and time functions, and string functions. Regardless of how complex or simple functions are, they are all used the same way. To compute a variable using a function (we'll use the example of summing var00001 and var00003):

1. Make sure the Data Editor window is active.
2. Click **Transform → Compute**.
3. Click the **Target Variable** text box and enter the name of a new variable to be created.
4. In the **Functions** text box, find the function you want to use. You can type the first letter of the function (which would be "s" in the case of SUM) and SPSS will take you to the first function beginning with that letter, or you can simply scroll down the list of functions.
5. Double click on the function to add it to the Numeric Expression box.
6. Click on the first variable, in our example **var00001**, then click ► to add it to the Numeric Expression box.
7. Continue adding variables in this manner until you are finished.
8. Click **OK** and the new variable is created.

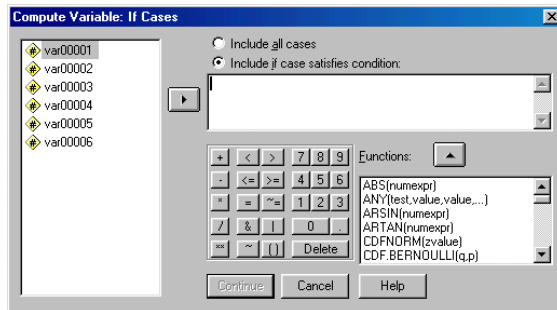
Computing Variables: If Cases

The If Cases dialog box allows you to apply data transformations to selected subsets of cases, using conditional expressions. A conditional expression returns a value of *true*, *false*, or *missing* for each case.

- If the result of a conditional expression is *true*, the transformation is applied to the case.
- If the result of a conditional expression is *false* or *missing*, the transformation is not applied to the case.
- Most conditional expressions use one or more of the six relational operators (<, >, <=, >=, =, and ~=) on the calculator pad.
- Conditional expressions can include variable names, constants, arithmetic operators, numeric and other functions, logical variables, and relational operators.

Let's say we wanted to create a new variable, var00006, that will equal the sum of var00003 and var00005, if the value of var00001 = 1. To do this:

1. Make sure the Data View window is active.
2. Click **Transform** → **Compute**.
3. Click the **If** button. You will now see the following screen:



4. Check the **Include if case satisfies condition** option.
5. Click on the variable **var00001**.
6. Click on the ► to add it to If Condition box.
7. Click **Continue**.
8. You will now be returned to the **Compute Variable** dialog box. You can now compute the new variable as defined in the previous two sections.

Recoding Data

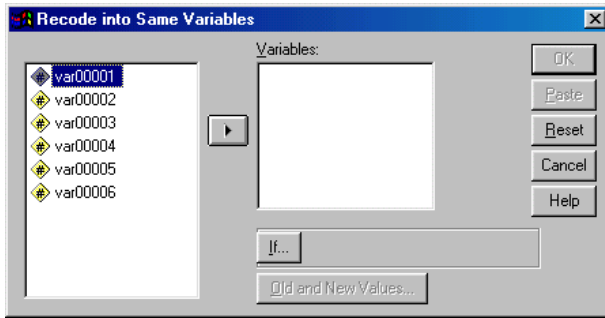
You can modify data values by recoding them. This is particularly useful for collapsing combining categories. You can recode the values within existing variables, or you new variables based on the recoded values of existing variables.

Recode into Same Variables

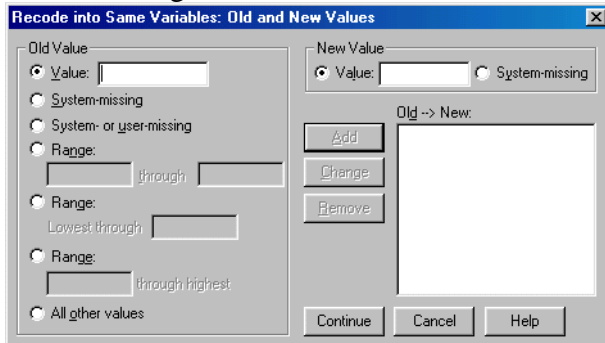
Recode into Same Variables reassigns the values of existing variables or collapses ranges of existing values into new values. For example, you could collapse salaries into salary range categories.

You can recode numeric and string variables. If you select multiple variables, they must all be the same type. You cannot recode numeric and string variables together. To recode the values of a variable and keep the results within the same variable.

1. Make sure the Data Editor window is active.
2. Click **Transform** → **Recode** → **Into Same Variable**. You will see the screen shown below:



3. Select the variables you want to recode. If you select multiple variables, they must be the same type (numeric or string).
4. Click **Old and New Values**. You will see the **Recode Into Same Variables: Old and New Values** dialog box as shown below:



5. Let's say we wanted to change the value of 5 to 1. You would type **5** in the Value area under old Values. (If you have a range of numbers you want to recode, let's say you asked a person's age and you want to create age groups, you can use one of the Range options under Old Value - - for example, if you wanted to recode people with an age of younger than 18 to 1, you could: (1) click on the first Range option and enter 1 through 17 or (2) click on the second Range option and choose Lowest through 17.)
6. Then, type **1** in the Value area under **New Values**.
7. Click **Add**. The variable is added to the **Old → New** box.
8. Click **Continue**. You will be returned to the **Recode into Same Variables** dialog box.
9. Click **OK** and the actual values entered in the cells will change according to the way in which they were recoded.

Recode into Different Variables

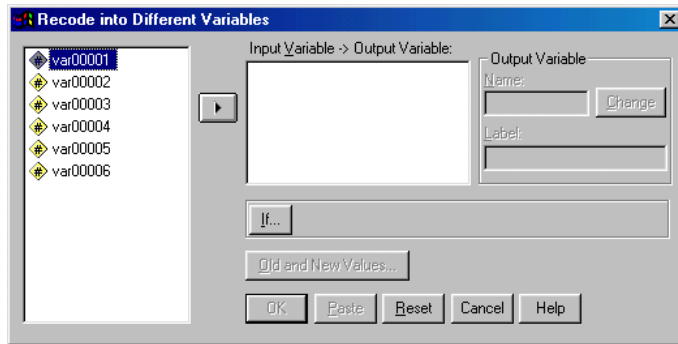
Sometimes you may not want to alter your original variable. Recode into Different Variables reassigns the values of existing variables or collapses ranges of existing values into new values for a new variable. For example, you could collapse age into a new variable containing age-range categories.

- You can recode numeric and string variables.
- You can recode numeric variables into string variables and vice versa.
- If you select multiple variables, they must all be the same type. You cannot recode numeric and string variables together.

To recode the values of a variable into a new variable:

1. Make sure the Data Editor window is active.

2. Click **Transform** → **Recode** → **Into Different Variables**. You will see the screen shown below:

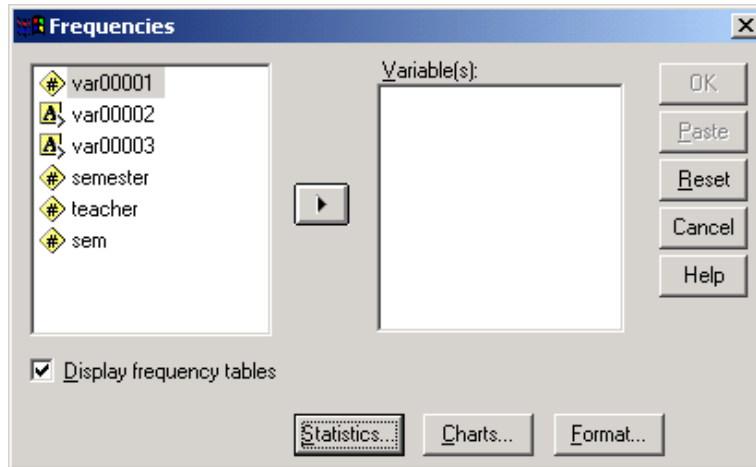


3. Select the variables you want to recode and click on the ► to add it to the **Numeric Variable** → **Output** box.
4. Under Output Variable go to the **Name** box and type in the name for the new variable. Then go to **Label** and type in the variable label for the new variable.
5. Click **Old and New Values** and specify how to recode values (follow the same as #5 - #8 in the previous section for Recoding Into the Same Variable).
6. Click **Change**.
7. Click **OK** and the actual values entered in the cells will change according to the way in which they were recoded.
8. Please note that when you recode into a new variable, you will have to go to Variable View to define the variable labels, value labels, etc.

Running Basic Analyses in SPSS

Running a Frequency in SPSS

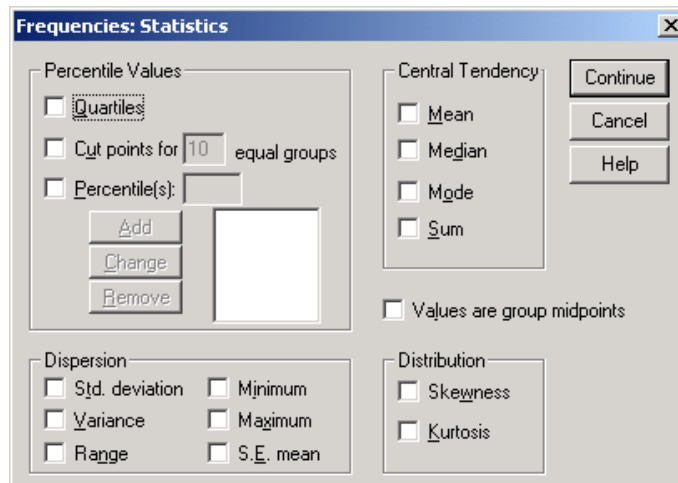
1. Go to the Analyze menu and select **Descriptive Statistics**. Then choose **Frequencies**. You will see the following screen:



2. Highlight the variable(s) on which you want to run frequency and press the ►.

To select multiple variables, you can (1) click on the mouse, keep the mouse button depressed and scroll down slowly until all variables are highlighted, (2) click on the first variable, hold down the shift key, then go to the last variable you want (assuming the variables are in consecutive order), or (3) select the first variable, go to the next variable, press the CTRL key and select it with your mouse.

3. If you need to select Descriptive Statistics, (e.g. a measure of central tendency, measure of dispersion, etc.), click the button labeled **Statistics**. You will see the following screen:

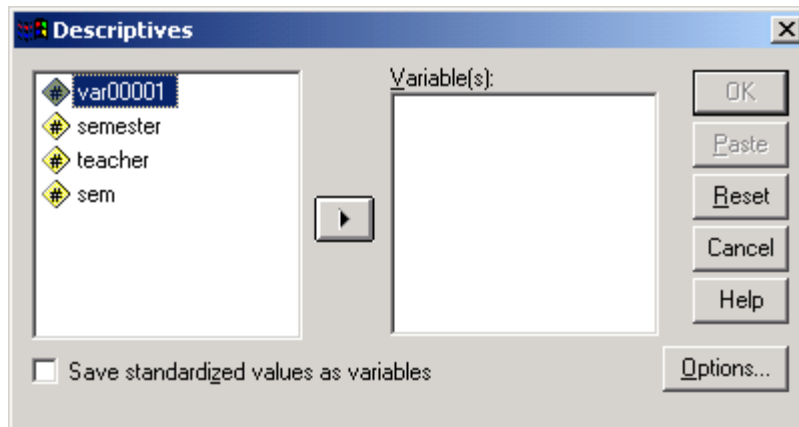


Select the appropriate statistics and click **Continue**.

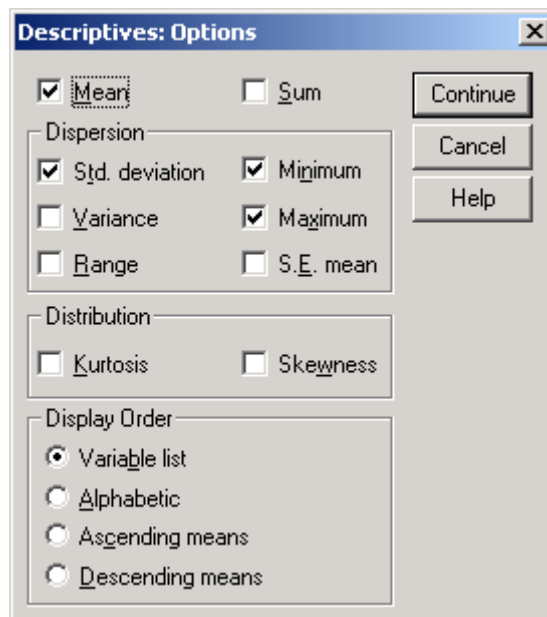
4. Click **OK** to run the analysis

Running Descriptives in SPSS

1. Go to the **Analyze** menu and select **Descriptive Statistics**. Then select **Descriptives**. You will see the following screen:



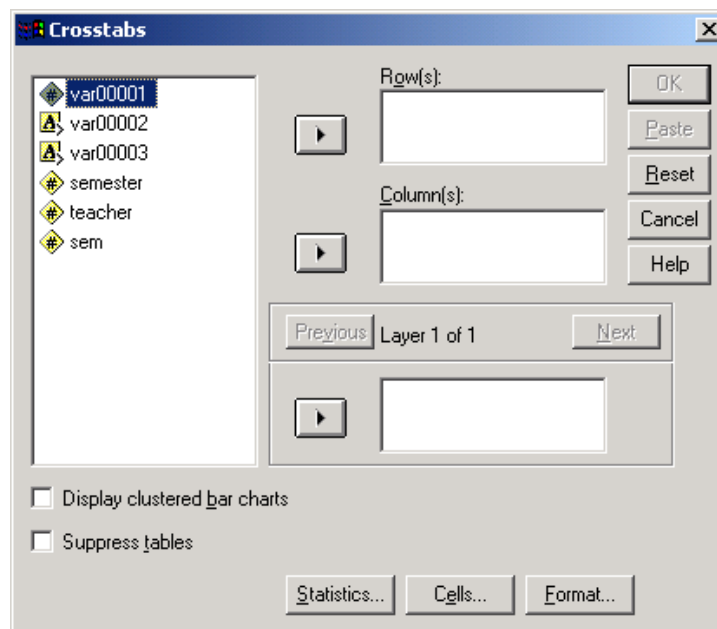
2. Highlight the variable(s) on which you want to run frequency and press the ►.
3. To choose the specific descriptive statistics you wish to run, click the **Options** button. You will see the following screen:



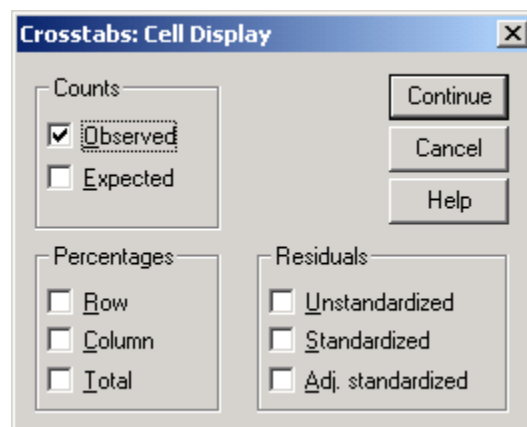
4. You will notice some options already checked (these are the defaults). Notice that you can also choose how you wish to view the data. After you have selected all of the options desired, click on **Continue**. Then when you return to the Descriptives main window, click **OK**.

Running Crosstabs in SPSS

1. Go to the **Analyze** menu and select **Descriptive Statistics**. Then select **Crosstabs**. You will see the following screen:



2. Highlight your stub variable and place in the Row(s). Highlight your banner variable and place in the Column(s).
3. Click on the **Cells** button. You will see the following screen:



4. Click on the percentages you wish to see displayed. Click **Continue**.
5. Click **OK**.