

## Using the Data Editor

The Data Editor displays the contents of the active data file. The information in the Data Editor consists of variables and cases.

- In Data View, columns represent variables, and rows represent cases (observations).
- In Variable View, each row is a variable, and each column is an attribute that is associated with that variable.

Variables are used to represent the different types of data that you have compiled. A common analogy is that of a survey. The response to each question on a survey is equivalent to a variable. Variables come in many different types, including numbers, strings, currency, and dates.

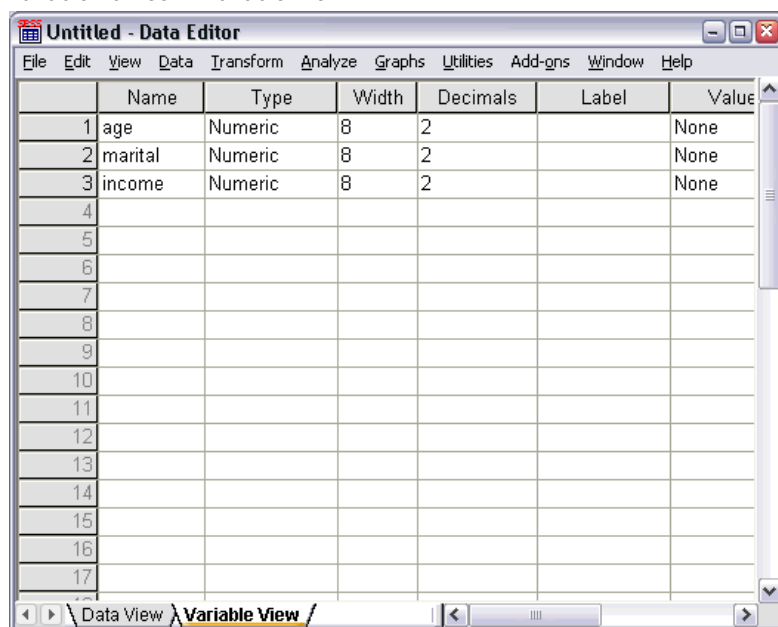
### Entering Numeric Data

Data can be entered into the Data Editor, which may be useful for small data files or for making minor edits to larger data files.

- ▶ Click the Variable View tab at the bottom of the Data Editor window.

You need to define the variables that will be used. In this case, only three variables are needed: *age*, *marital status*, and *income*.

Figure 3-1  
Variable names in Variable View



- ▶ In the first row of the first column, type *age*.

- ▶ In the second row, type *marital*.

- ▶ In the third row, type *income*.

New variables are automatically given a Numeric data type.

If you don't enter variable names, unique names are automatically created. However, these names are not descriptive and are not recommended for large data files.

- ▶ Click the *Data View* tab to continue entering the data.

The names that you entered in *Variable View* are now the headings for the first three columns in *Data View*.

Begin entering data in the first row, starting at the first column.

Figure 3-2  
Values entered in *Data View*

	age	marital	income	var	var	var
1	55.00	1.00	72000.00			
2	53.00	.00	153000.0			
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						

- ▶ In the *age* column, type 55.
- ▶ In the *marital* column, type 1.
- ▶ In the *income* column, type 72000.
- ▶ Move the cursor to the second row of the first column to add the next subject's data.
- ▶ In the *age* column, type 53.
- ▶ In the *marital* column, type 0.
- ▶ In the *income* column, type 153000.

Currently, the *age* and *marital* columns display decimal points, even though their values are intended to be integers. To hide the decimal points in these variables:

- ▶ Click the Variable View tab at the bottom of the Data Editor window.
- ▶ In the *Decimals* column of the *age* row, type 0 to hide the decimal.
- ▶ In the *Decimals* column of the *marital* row, type 0 to hide the decimal.

Figure 3-3

Updated decimal property for *age* and *marital*

	Name	Type	Width	Decimals	Label	Value
1	age	Numeric	8	0		None
2	marital	Numeric	8	0		None
3	income	Numeric	8	2		None
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						

## Entering String Data

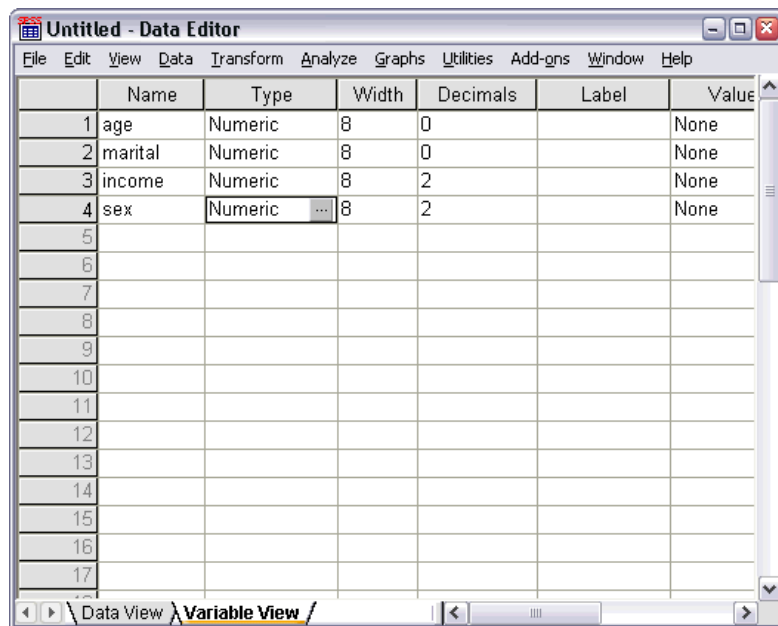
Non-numeric data, such as strings of text, can also be entered into the Data Editor.

- ▶ Click the Variable View tab at the bottom of the Data Editor window.
- ▶ In the first cell of the first empty row, type *sex* for the variable name.
- ▶ Click the *Type* cell next to your entry.

- ▶ Click the button on the right side of the *Type* cell to open the Variable Type dialog box.

Figure 3-4

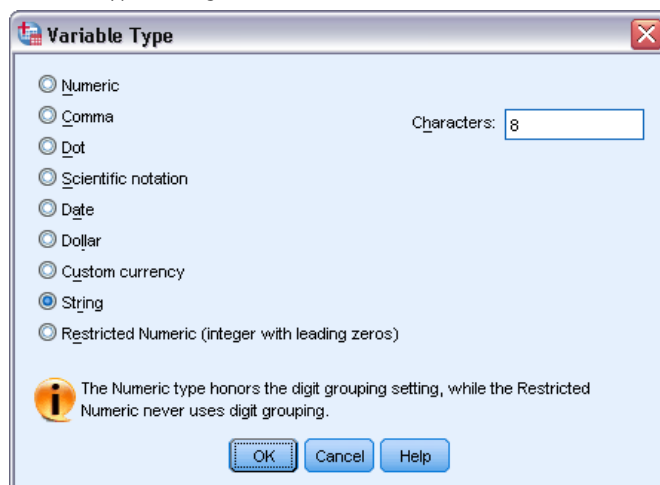
Button shown in *Type* cell for *sex*



- ▶ Select String to specify the variable type.
- ▶ Click OK to save your selection and return to the Data Editor.

Figure 3-5

Variable Type dialog box



## Defining Data

In addition to defining data types, you can also define descriptive variable labels and value labels for variable names and data values. These descriptive labels are used in statistical reports and charts.

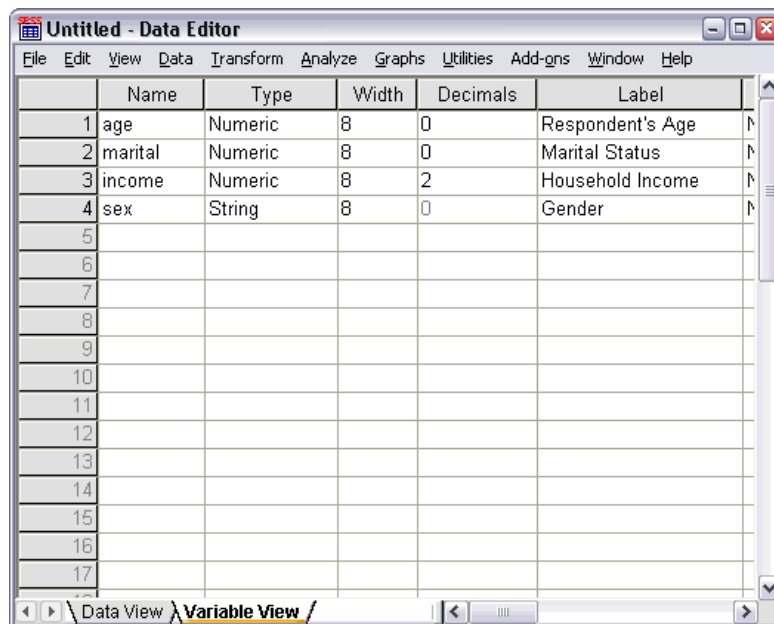
### Adding Variable Labels

Labels are meant to provide descriptions of variables. These descriptions are often longer versions of variable names. Labels can be up to 255 bytes. These labels are used in your output to identify the different variables.

- ▶ Click the Variable View tab at the bottom of the Data Editor window.
- ▶ In the *Label* column of the *age* row, type Respondent's Age.
- ▶ In the *Label* column of the *marital* row, type Marital Status.
- ▶ In the *Label* column of the *income* row, type Household Income.
- ▶ In the *Label* column of the *sex* row, type Gender.

Figure 3-6

Variable labels entered in Variable View

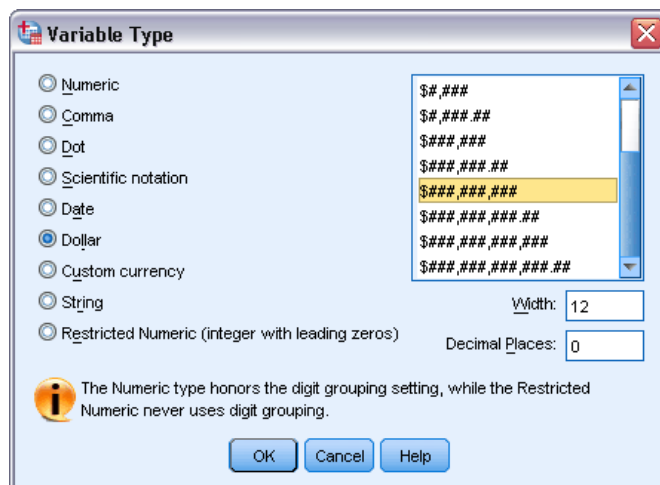


## Changing Variable Type and Format

The *Type* column displays the current data type for each variable. The most common data types are numeric and string, but many other formats are supported. In the current data file, the *income* variable is defined as a numeric type.

- ▶ Click the *Type* cell for the *income* row, and then click the button on the right side of the cell to open the Variable Type dialog box.
- ▶ Select Dollar.

Figure 3-7  
Variable Type dialog box



The formatting options for the currently selected data type are displayed.

- ▶ For the format of the currency in this example, select `$###,###,###`.
- ▶ Click OK to save your changes.

## Adding Value Labels for Numeric Variables

Value labels provide a method for mapping your variable values to a string label. In this example, there are two acceptable values for the *marital* variable. A value of 0 means that the subject is single, and a value of 1 means that he or she is married.

- ▶ Click the *Values* cell for the *marital* row, and then click the button on the right side of the cell to open the Value Labels dialog box.

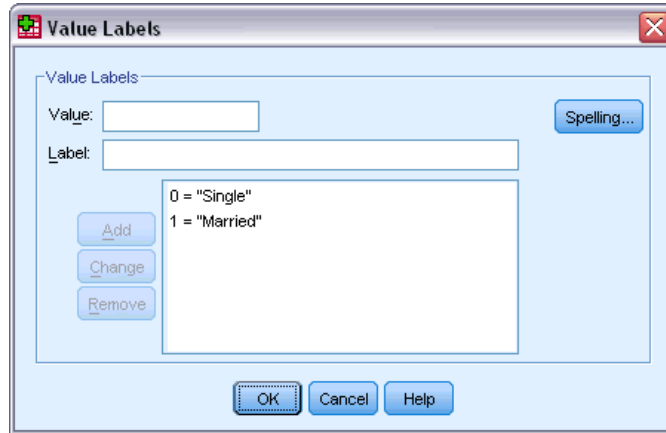
The **value** is the actual numeric value.

The **value label** is the string label that is applied to the specified numeric value.

- ▶ Type 0 in the Value field.
- ▶ Type Single in the Label field.

- ▶ Click Add to add this label to the list.

Figure 3-8  
*Value Labels dialog box*



- ▶ Type 1 in the Value field, and type Married in the Label field.
- ▶ Click Add, and then click OK to save your changes and return to the Data Editor.  
These labels can also be displayed in Data View, which can make your data more readable.
- ▶ Click the Data View tab at the bottom of the Data Editor window.
- ▶ From the menus choose:  
View > Value Labels

The labels are now displayed in a list when you enter values in the Data Editor. This setup has the benefit of suggesting a valid response and providing a more descriptive answer.

If the Value Labels menu item is already active (with a check mark next to it), choosing Value Labels again will turn *off* the display of value labels.

Figure 3-9  
Value labels displayed in Data View

	age	marital	income	sex	var	var
1	55	Married	\$72,000			
2	53	Single	\$153,000			
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						

### ***Adding Value Labels for String Variables***

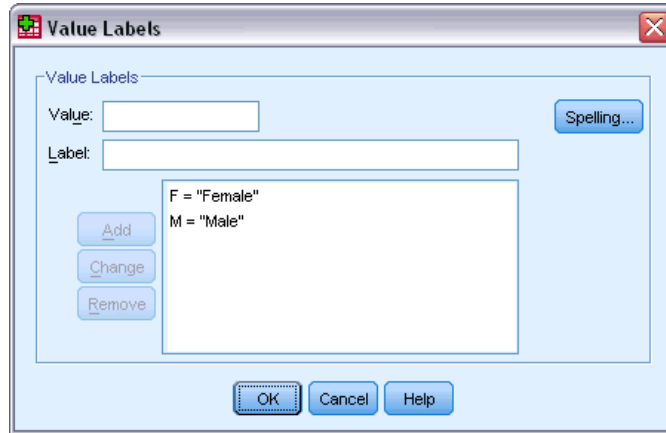
String variables may require value labels as well. For example, your data may use single letters, *M* or *F*, to identify the sex of the subject. Value labels can be used to specify that *M* stands for *Male* and *F* stands for *Female*.

- ▶ Click the Variable View tab at the bottom of the Data Editor window.
- ▶ Click the *Values* cell in the *sex* row, and then click the button on the right side of the cell to open the Value Labels dialog box.
- ▶ Type *F* in the Value field, and then type *Female* in the Label field.



- ▶ Click Add to add this label to your data file.

Figure 3-10  
Value Labels dialog box



- ▶ Type M in the Value field, and type Male in the Label field.
- ▶ Click Add, and then click OK to save your changes and return to the Data Editor.

Because string values are case sensitive, you should be consistent. A lowercase *m* is not the same as an uppercase *M*.

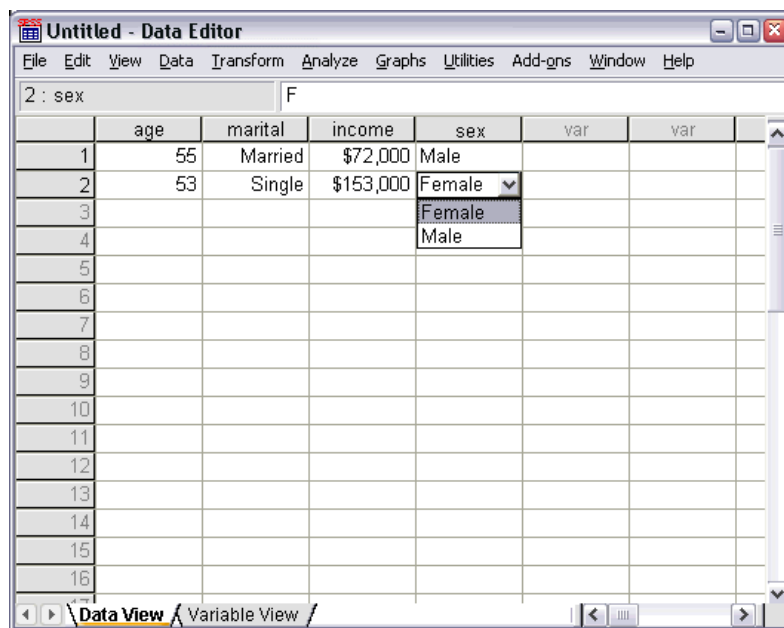
### ***Using Value Labels for Data Entry***

You can use value labels for data entry.

- ▶ Click the Data View tab at the bottom of the Data Editor window.
- ▶ In the first row, select the cell for *sex*.
- ▶ Click the button on the right side of the cell, and then choose Male from the drop-down list.
- ▶ In the second row, select the cell for *sex*.

- Click the button on the right side of the cell, and then choose Female from the drop-down list.

Figure 3-11  
Using variable labels to select values



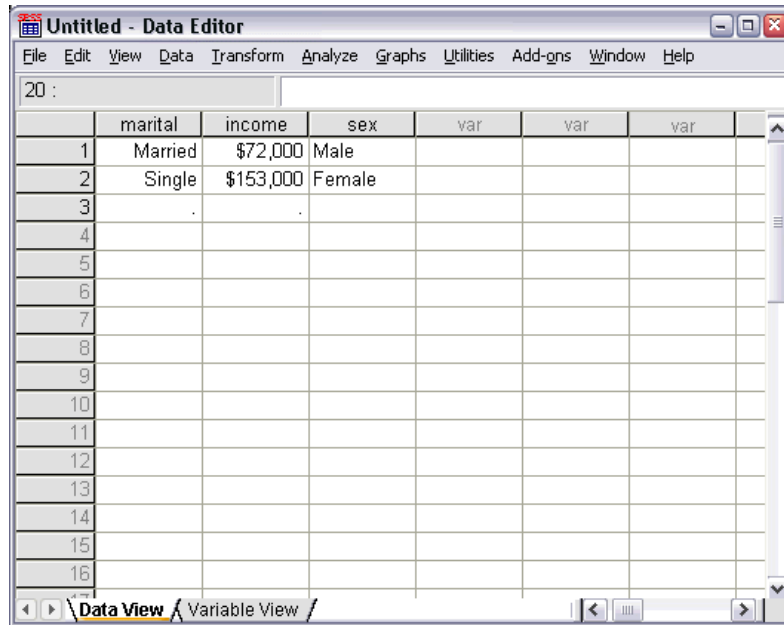
Only defined values are listed, which ensures that the entered data are in a format that you expect.

## Handling Missing Data

Missing or invalid data are generally too common to ignore. Survey respondents may refuse to answer certain questions, may not know the answer, or may answer in an unexpected format. If you don't filter or identify these data, your analysis may not provide accurate results.

For numeric data, empty data fields or fields containing invalid entries are converted to system-missing, which is identifiable by a single period.

Figure 3-12  
Missing values displayed as periods



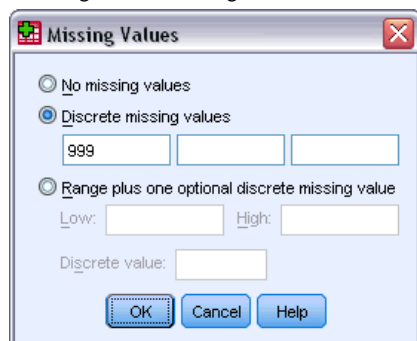
The reason a value is missing may be important to your analysis. For example, you may find it useful to distinguish between those respondents who refused to answer a question and those respondents who didn't answer a question because it was not applicable.

### Missing Values for a Numeric Variable

- ▶ Click the Variable View tab at the bottom of the Data Editor window.
- ▶ Click the *Missing* cell in the *age* row, and then click the button on the right side of the cell to open the Missing Values dialog box.

In this dialog box, you can specify up to three distinct missing values, or you can specify a range of values plus one additional discrete value.

Figure 3-13  
Missing Values dialog box

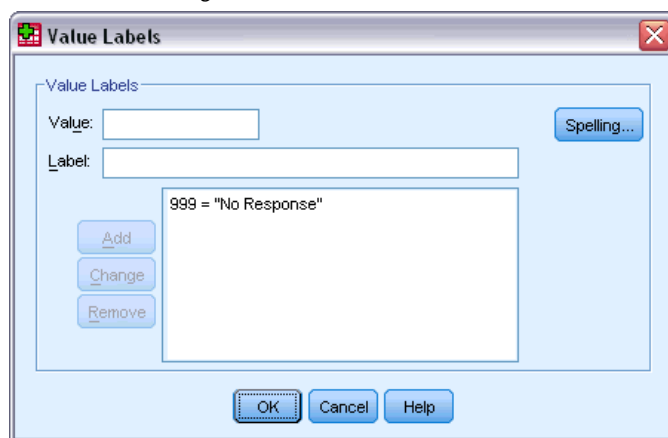


- ▶ Select Discrete missing values.
- ▶ Type 999 in the first text box and leave the other two text boxes empty.
- ▶ Click OK to save your changes and return to the Data Editor.

Now that the missing data value has been added, a label can be applied to that value.

- ▶ Click the *Values* cell in the *age* row, and then click the button on the right side of the cell to open the Value Labels dialog box.
- ▶ Type 999 in the Value field.
- ▶ Type No Response in the Label field.

Figure 3-14  
Value Labels dialog box



- ▶ Click Add to add this label to your data file.
- ▶ Click OK to save your changes and return to the Data Editor.

### **Missing Values for a String Variable**

Missing values for string variables are handled similarly to the missing values for numeric variables. However, unlike numeric variables, empty fields in string variables are not designated as system-missing. Rather, they are interpreted as an empty string.

- ▶ Click the Variable View tab at the bottom of the Data Editor window.
- ▶ Click the *Missing* cell in the *sex* row, and then click the button on the right side of the cell to open the Missing Values dialog box.
- ▶ Select Discrete missing values.
- ▶ Type NR in the first text box.

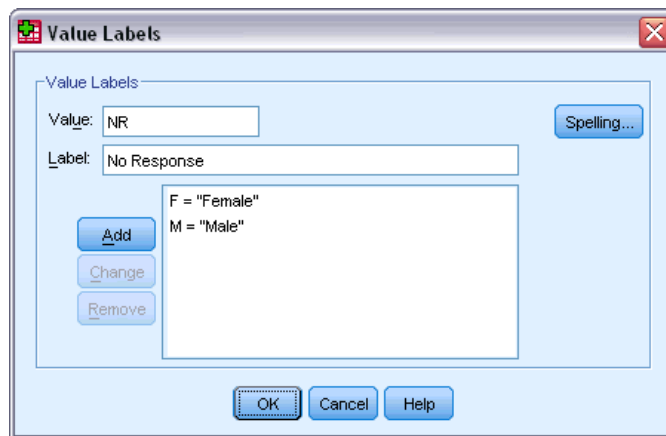
Missing values for string variables are case sensitive. So, a value of *nr* is not treated as a missing value.

- ▶ Click OK to save your changes and return to the Data Editor.

Now you can add a label for the missing value.

- ▶ Click the *Values* cell in the *sex* row, and then click the button on the right side of the cell to open the Value Labels dialog box.
- ▶ Type NR in the Value field.
- ▶ Type No Response in the Label field.

Figure 3-15  
Value Labels dialog box



- ▶ Click Add to add this label to your project.
- ▶ Click OK to save your changes and return to the Data Editor.

### ***Copying and Pasting Variable Attributes***

After you've defined variable attributes for a variable, you can copy these attributes and apply them to other variables.

- ▶ In Variable View, type *agewed* in the first cell of the first empty row.

Figure 3-16  
*agewed* variable in Variable View

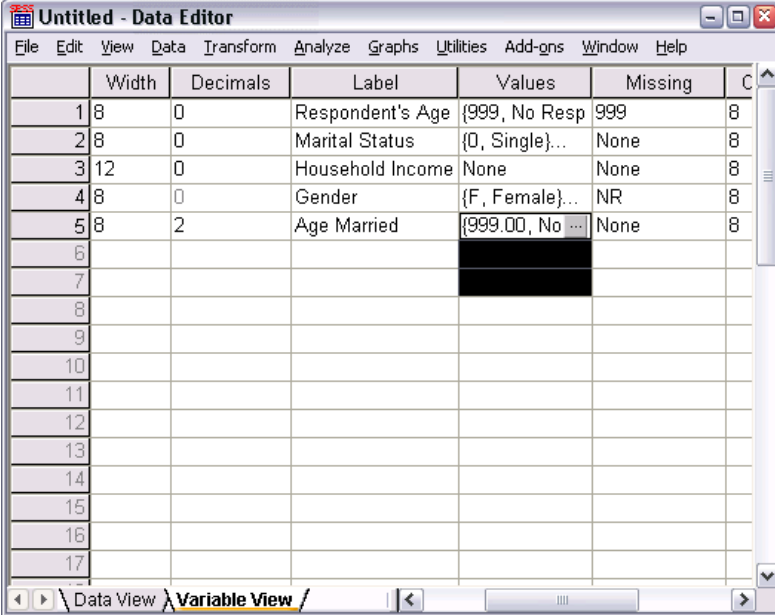
	Name	Type	Width	Decimals	Label
1	age	Numeric	8	0	Respondent's Age
2	marital	Numeric	8	0	Marital Status
3	income	Dollar	12	0	Household Income
4	sex	String	8	0	Gender
5	agewed	Numeric	8	2	Age Married
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					

- ▶ In the *Label* column, type Age Married.
- ▶ Click the *Values* cell in the *age* row.
- ▶ From the menus choose:  
Edit > Copy
- ▶ Click the *Values* cell in the *agewed* row.
- ▶ From the menus choose:  
Edit > Paste

The defined values from the *age* variable are now applied to the *agewed* variable.

To apply the attribute to multiple variables, simply select multiple target cells (click and drag down the column).

Figure 3-17  
*Multiple cells selected*



The screenshot shows the SPSS Data Editor window in Variable View. The table below represents the data shown in the window. The 'Values' column for 'Age Married' is highlighted, indicating that multiple cells are selected.

	Width	Decimals	Label	Values	Missing	C
1	8	0	Respondent's Age	{999, No Resp	999	8
2	8	0	Marital Status	{0, Single}..	None	8
3	12	0	Household Income	None	None	8
4	8	0	Gender	{F, Female}..	NR	8
5	8	2	Age Married	{999.00, No ...	None	8
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						

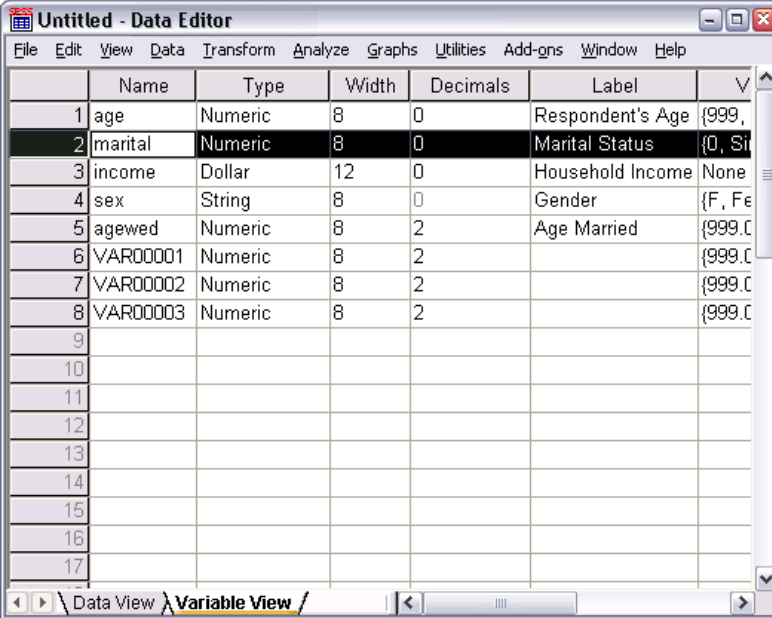
When you paste the attribute, it is applied to all of the selected cells.

New variables are automatically created if you paste the values into empty rows.

To copy all attributes from one variable to another variable:

- ▶ Click the row number in the *marital* row.

Figure 3-18  
Selected row



The screenshot shows the SPSS Data Editor window in Variable View. The 'marital' row is selected, indicated by a dark background. The table below represents the data shown in the screenshot.

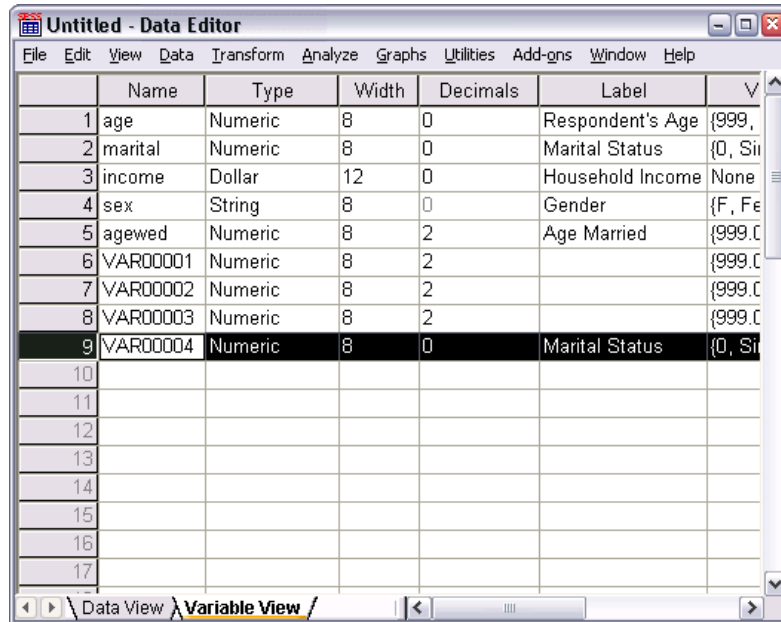
	Name	Type	Width	Decimals	Label	V
1	age	Numeric	8	0	Respondent's Age	{999.
2	marital	Numeric	8	0	Marital Status	{0, Si
3	income	Dollar	12	0	Household Income	None
4	sex	String	8	0	Gender	{F, Fe
5	agewed	Numeric	8	2	Age Married	{999.C
6	VAR00001	Numeric	8	2		{999.C
7	VAR00002	Numeric	8	2		{999.C
8	VAR00003	Numeric	8	2		{999.C
9						
10						
11						
12						
13						
14						
15						
16						
17						

- ▶ From the menus choose:  
Edit > Copy
- ▶ Click the row number of the first empty row.
- ▶ From the menus choose:  
Edit > Paste



All attributes of the *marital* variable are applied to the new variable.

Figure 3-19  
All values pasted into a row



### Defining Variable Properties for Categorical Variables

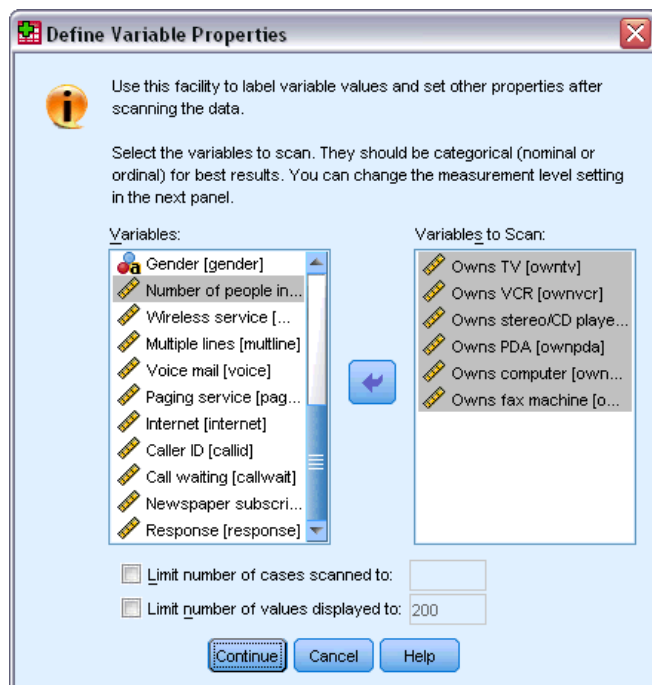
For categorical (nominal, ordinal) data, you can use Define Variable Properties to define value labels and other variable properties. The Define Variable Properties process:

- Scans the actual data values and lists all unique data values for each selected variable.
- Identifies unlabeled values and provides an “auto-label” feature.
- Provides the ability to copy defined value labels from another variable to the selected variable or from the selected variable to additional variables.

This example uses the data file *demo.sav*. For more information, see the topic [Sample Files](#) in Appendix A on p. 147. This data file already has defined value labels, so we will enter a value for which there is no defined value label.

- ▶ In Data View of the Data Editor, click the first data cell for the variable *ownpc* (you may have to scroll to the right), and then enter 99.
- ▶ From the menu choose:  
Data > Define Variable Properties...

Figure 3-20  
Initial Define Variable Properties dialog box



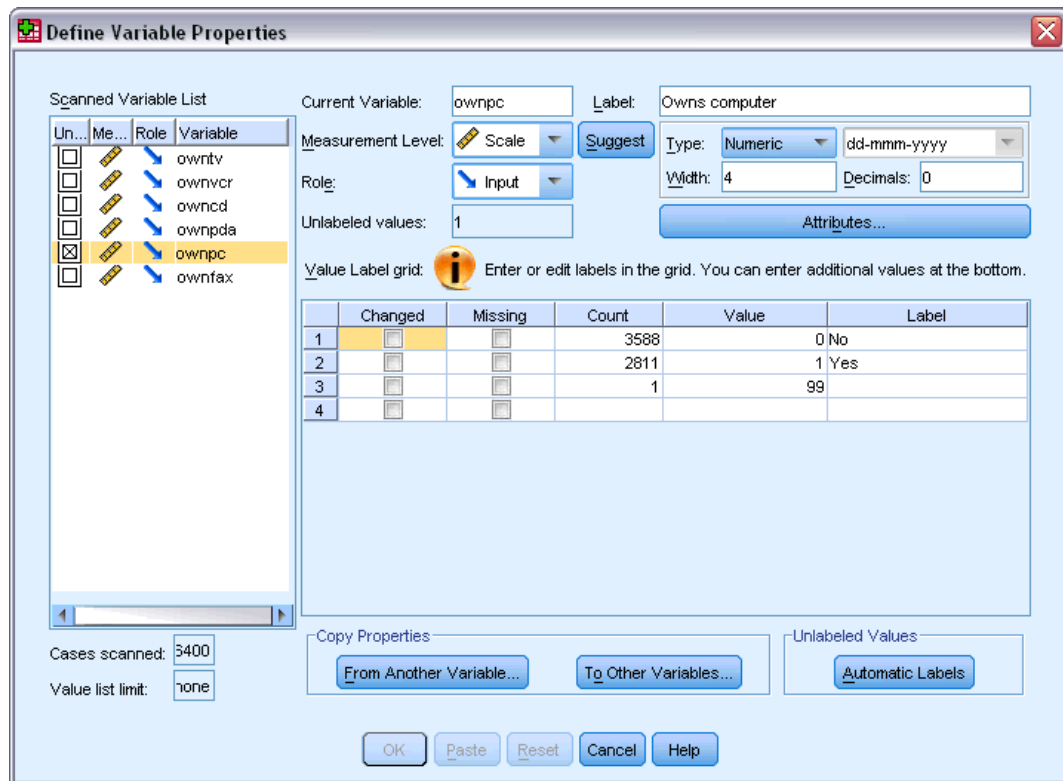
In the initial Define Variable Properties dialog box, you select the nominal or ordinal variables for which you want to define value labels and/or other properties.

- ▶ Drag and drop *Owns TV [owntv]* through *Owns fax machine [ownfax]* into the Variables to Scan list.

You might notice that the measurement level icons for all of the selected variables indicate that they are scale variables, not categorical variables. All of the selected variables in this example are really categorical variables that use the numeric values 0 and 1 to stand for *No* and *Yes*, respectively—and one of the variable properties that we'll change with Define Variable Properties is the measurement level.

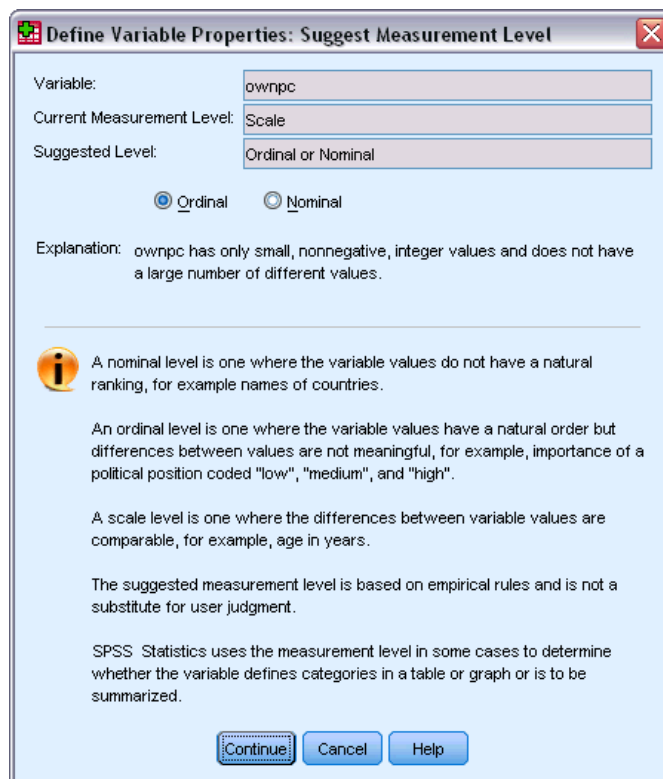
- ▶ Click Continue.

Figure 3-21  
Define Variable Properties main dialog box



- ▶ In the Scanned Variable List, select *ownpc*.  
The current level of measurement for the selected variable is scale. You can change the measurement level by selecting a level from the drop-down list, or you can let Define Variable Properties suggest a measurement level.
- ▶ Click Suggest.  
The Suggest Measurement Level dialog box is displayed.

Figure 3-22  
Suggest Measurement Level dialog box



Because the variable doesn't have very many different values and all of the scanned cases contain integer values, the proper measurement level is probably ordinal or nominal.

- ▶ Select **Ordinal**, and then click **Continue**.

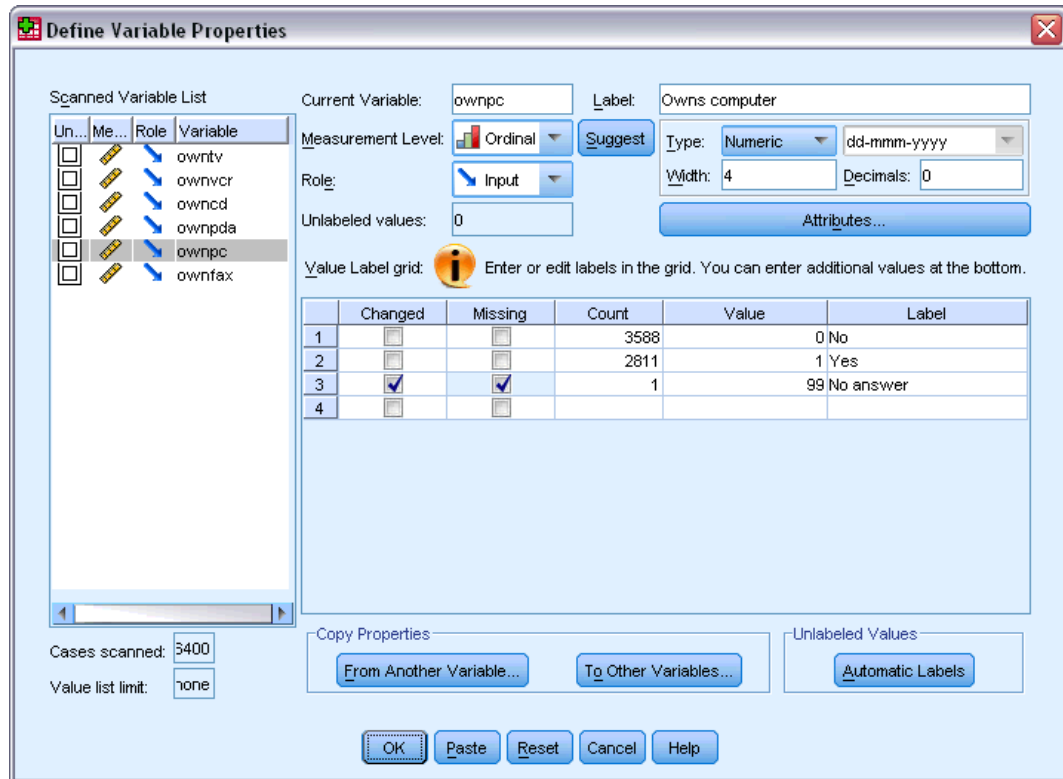
The measurement level for the selected variable is now ordinal.

The Value Label grid displays all of the unique data values for the selected variable, any defined value labels for these values, and the number of times (count) that each value occurs in the scanned cases.

The value that we entered in Data View, 99, is displayed in the grid. The count is only 1 because we changed the value for only one case, and the *Label* column is empty because we haven't defined a value label for 99 yet. An X in the first column of the Scanned Variable List also indicates that the selected variable has at least one observed value without a defined value label.

- ▶ In the *Label* column for the value of 99, enter No answer.
- ▶ Check the box in the *Missing* column for the value 99 to identify the value 99 as **user-missing**. Data values that are specified as user-missing are flagged for special treatment and are excluded from most calculations.

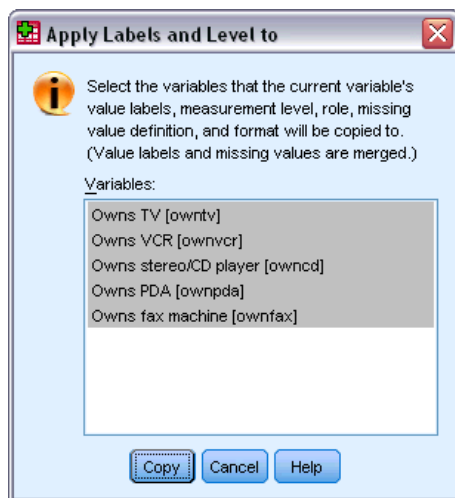
Figure 3-23  
New variable properties defined for *ownpc*



Before we complete the job of modifying the variable properties for *ownpc*, let's apply the same measurement level, value labels, and missing values definitions to the other variables in the list.

- In the Copy Properties area, click To Other Variables.

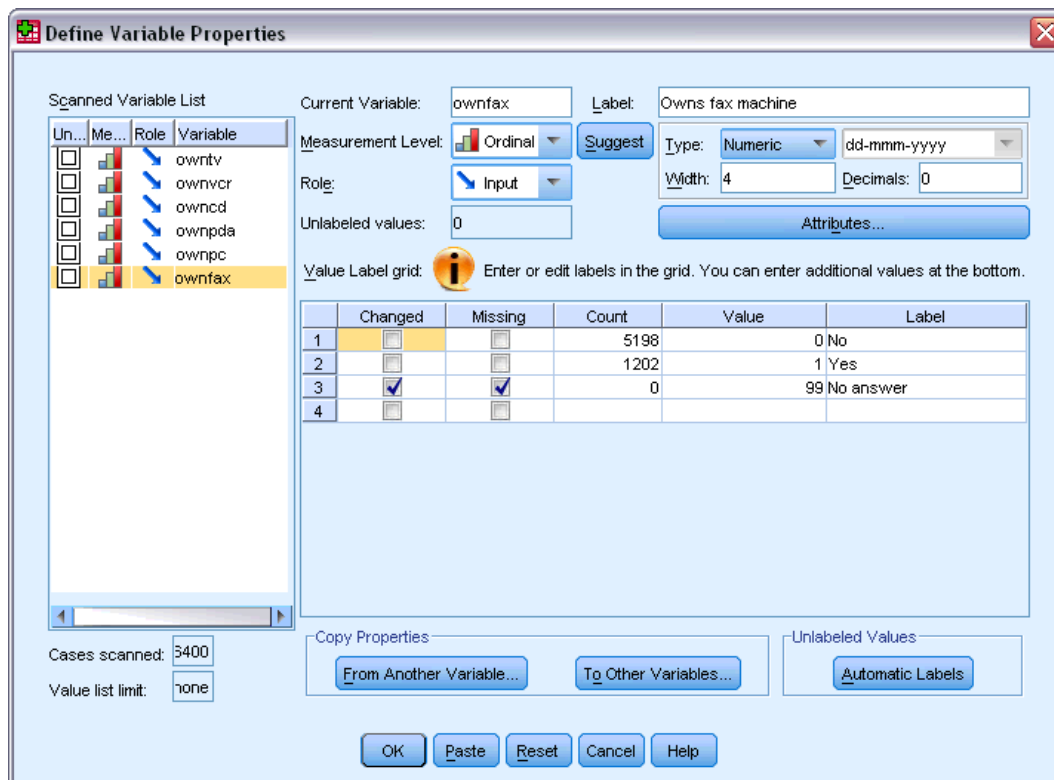
Figure 3-24  
Apply Labels and Level dialog box



- In the Apply Labels and Level dialog box, select all of the variables in the list, and then click Copy.

If you select any other variable in the Scanned Variable List of the Define Variable Properties main dialog box now, you'll see that they are all ordinal variables, with a value of 99 defined as user-missing and a value label of *No answer*.

Figure 3-25  
New variable properties defined for ownfax



- ▶ Click OK to save all of the variable properties that you have defined.