

# ***Introduction***

This guide provides a set of tutorials designed to enable you to perform useful analyses on your data. You can work through the tutorials in sequence or turn to the topics for which you need additional information.

This chapter will introduce you to the basic features and demonstrate a typical session. We will retrieve a previously defined IBM® SPSS® Statistics data file and then produce a simple statistical summary and a chart.

More detailed instruction about many of the topics touched upon in this chapter will follow in later chapters. Here, we hope to give you a basic framework for understanding later tutorials.

## ***Sample Files***

Most of the examples that are presented here use the data file *demo.sav*. This data file is a fictitious survey of several thousand people, containing basic demographic and consumer information.

If you are using the Student version, your version of *demo.sav* is a representative sample of the original data file, reduced to meet the 1,500-case limit. Results that you obtain using that data file will differ from the results shown here.

The sample files installed with the product can be found in the *Samples* subdirectory of the installation directory. There is a separate folder within the Samples subdirectory for each of the following languages: English, French, German, Italian, Japanese, Korean, Polish, Russian, Simplified Chinese, Spanish, and Traditional Chinese.

Not all sample files are available in all languages. If a sample file is not available in a language, that language folder contains an English version of the sample file.

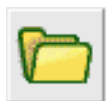
## ***Opening a Data File***

To open a data file:

- ▶ From the menus choose:  
File > Open > Data...

Alternatively, you can use the Open File button on the toolbar.

Figure 1-1  
*Open File toolbar button*

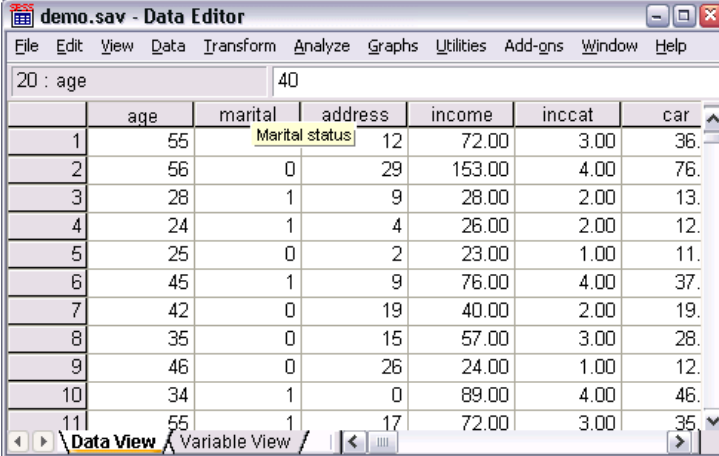


A dialog box for opening files is displayed.

By default, IBM® SPSS® Statistics data files (.sav extension) are displayed.

This example uses the file *demo.sav*.

Figure 1-2  
*demo.sav* file in Data Editor



|    | age | marital        | address | income | inccat | car |
|----|-----|----------------|---------|--------|--------|-----|
| 1  | 55  | Marital status | 12      | 72.00  | 3.00   | 36. |
| 2  | 56  | 0              | 29      | 153.00 | 4.00   | 76. |
| 3  | 28  | 1              | 9       | 28.00  | 2.00   | 13. |
| 4  | 24  | 1              | 4       | 26.00  | 2.00   | 12. |
| 5  | 25  | 0              | 2       | 23.00  | 1.00   | 11. |
| 6  | 45  | 1              | 9       | 76.00  | 4.00   | 37. |
| 7  | 42  | 0              | 19      | 40.00  | 2.00   | 19. |
| 8  | 35  | 0              | 15      | 57.00  | 3.00   | 28. |
| 9  | 46  | 0              | 26      | 24.00  | 1.00   | 12. |
| 10 | 34  | 1              | 0       | 89.00  | 4.00   | 46. |
| 11 | 55  | 1              | 17      | 72.00  | 3.00   | 35. |

The data file is displayed in the Data Editor. In the Data Editor, if you put the mouse cursor on a variable name (the column headings), a more descriptive variable label is displayed (if a label has been defined for that variable).

By default, the actual data values are displayed. To display labels:

- ▶ From the menus choose:  
View > Value Labels

Alternatively, you can use the Value Labels button on the toolbar.

Figure 1-3  
Value Labels button



Descriptive value labels are now displayed to make it easier to interpret the responses.

Figure 1-4  
Value labels displayed in the Data Editor

|    | age | marital   | address | income | inccat      | car |
|----|-----|-----------|---------|--------|-------------|-----|
| 1  | 55  | Married   | 12      | 72.00  | \$50 - \$74 | 36. |
| 2  | 56  | Unmarried | 29      | 153.00 | \$75+       | 76. |
| 3  | 28  | Married   | 9       | 28.00  | \$25 - \$49 | 13. |
| 4  | 24  | Married   | 4       | 26.00  | \$25 - \$49 | 12. |
| 5  | 25  | Unmarried | 2       | 23.00  | Under \$25  | 11. |
| 6  | 45  | Married   | 9       | 76.00  | \$75+       | 37. |
| 7  | 42  | Unmarried | 19      | 40.00  | \$25 - \$49 | 19. |
| 8  | 35  | Unmarried | 15      | 57.00  | \$50 - \$74 | 28. |
| 9  | 46  | Unmarried | 26      | 24.00  | Under \$25  | 12. |
| 10 | 34  | Married   | 0       | 89.00  | \$75+       | 46. |
| 11 | 55  | Married   | 17      | 72.00  | \$50 - \$74 | 35. |

## Running an Analysis

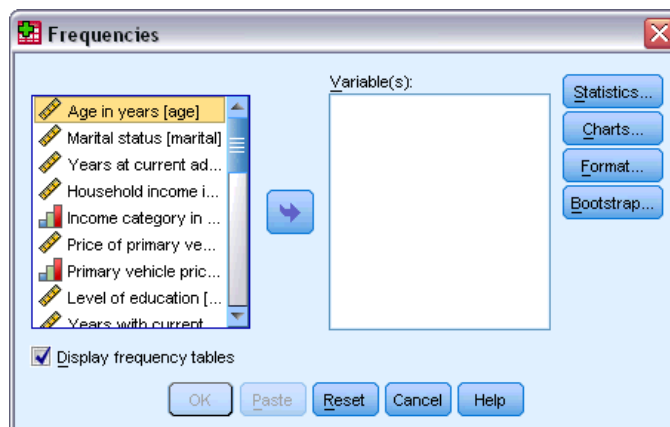
If you have any add-on options, the Analyze menu contains a list of reporting and statistical analysis categories.

We will start by creating a simple frequency table (table of counts). This example requires the Statistics Base option.












- From the menus choose:  
Analyze > Descriptive Statistics > Frequencies...

The Frequencies dialog box is displayed.

Figure 1-5  
Frequencies dialog box

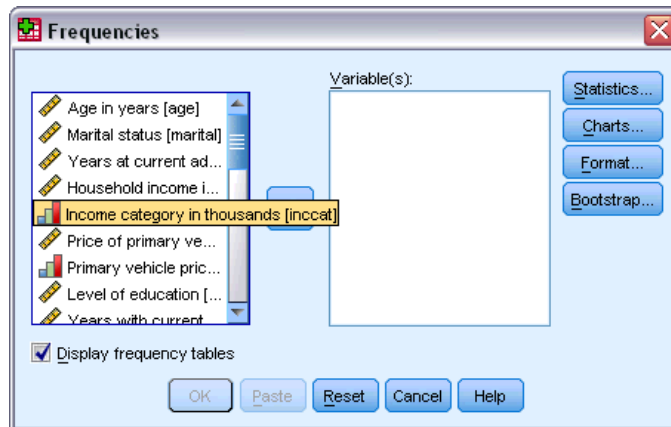


An icon next to each variable provides information about data type and level of measurement.

|                    | Numeric                                                                           | String                                                                            | Date                                                                               | Time                                                                                |
|--------------------|-----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|
| Scale (Continuous) |  | n/a                                                                               |  |  |
| Ordinal            |  |  |  |  |
| Nominal            |  |  |  |  |

- Click the variable *Income category in thousands [inccat]*.

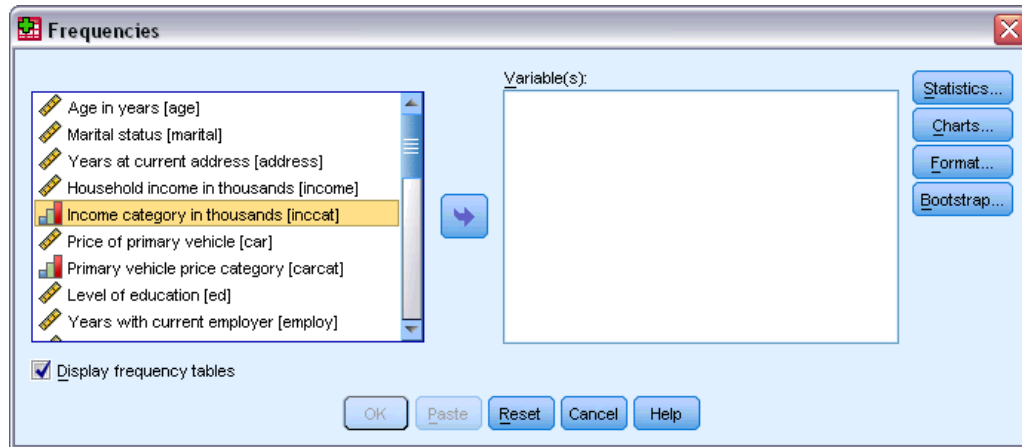
Figure 1-6  
Variable labels and names in the Frequencies dialog box



If the variable label and/or name appears truncated in the list, the complete label/name is displayed when the cursor is positioned over it. The variable name *inccat* is displayed in square brackets after the descriptive variable label. *Income category in thousands* is the variable label. If there were no variable label, only the variable name would appear in the list box.

You can resize dialog boxes just like windows, by clicking and dragging the outside borders or corners. For example, if you make the dialog box wider, the variable lists will also be wider.

Figure 1-7  
Resized dialog box

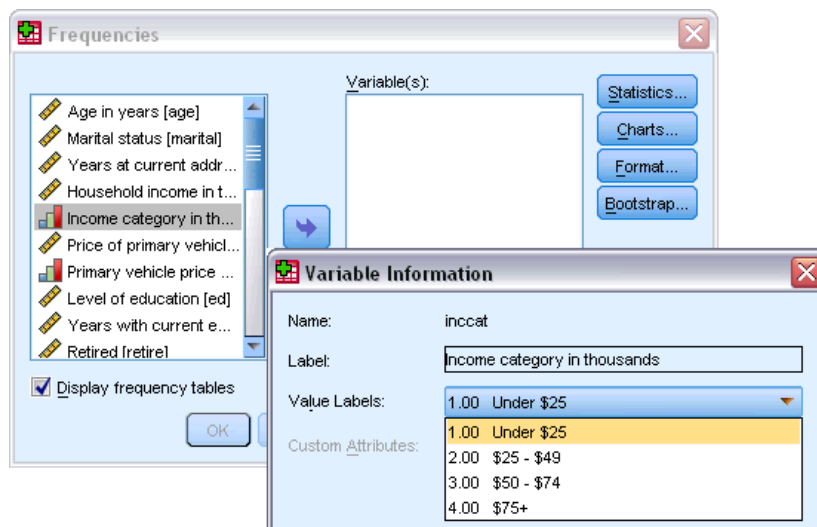


In the dialog box, you choose the variables that you want to analyze from the source list on the left and drag and drop them into the Variable(s) list on the right. The OK button, which runs the analysis, is disabled until at least one variable is placed in the Variable(s) list.

In many dialogs, you can obtain additional information by right-clicking any variable name in the list.

- ▶ Right-click *Income category in thousands [inccat]* and choose Variable Information.
- ▶ Click the down arrow on the Value labels drop-down list.

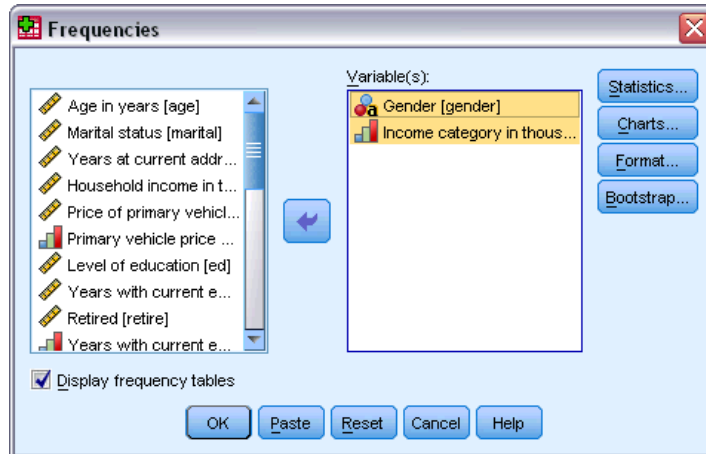
Figure 1-8  
Defined labels for income variable



All of the defined value labels for the variable are displayed.

- ▶ Click *Gender [gender]* in the source variable list and drag the variable into the target Variable(s) list.
- ▶ Click *Income category in thousands [inccat]* in the source list and drag it to the target list.

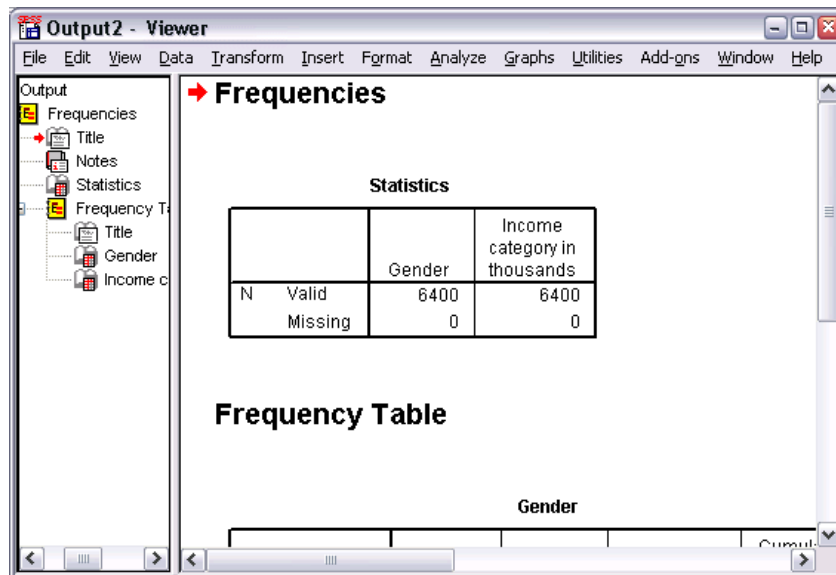
Figure 1-9  
Variables selected for analysis



- ▶ Click OK to run the procedure.

## Viewing Results

Figure 1-10  
Viewer window



Results are displayed in the Viewer window.

You can quickly go to any item in the Viewer by selecting it in the outline pane.

- ▶ Click Income category in thousands [inccat].

Figure 1-11  
Frequency table of income categories

The screenshot shows the SPSS Output2 - Viewer window. The left pane shows a tree view with 'Income c' selected. The main area displays two tables. The top table is titled 'Gender' and the bottom table is titled 'Income category in thousands'. A red arrow points to the bottom table.

|       |        | Frequency | Percent | Valid Percent | Cumul: Percent |
|-------|--------|-----------|---------|---------------|----------------|
| Valid | Female | 3179      | 49.7    | 49.7          |                |
|       | Male   | 3221      | 50.3    | 50.3          |                |
|       | Total  | 6400      | 100.0   | 100.0         |                |

|       |             | Frequency | Percent | Valid Percent | Cum: Percent |
|-------|-------------|-----------|---------|---------------|--------------|
| Valid | Under \$25  | 1174      | 18.3    | 18.3          |              |
|       | \$25 - \$49 | 2388      | 37.3    | 37.3          |              |
|       | \$50 - \$74 | 1120      | 17.5    | 17.5          |              |
|       | \$75+       | 1718      | 26.8    | 26.8          |              |
|       | Total       | 6400      | 100.0   | 100.0         |              |

The frequency table for income categories is displayed. This frequency table shows the number and percentage of people in each income category.

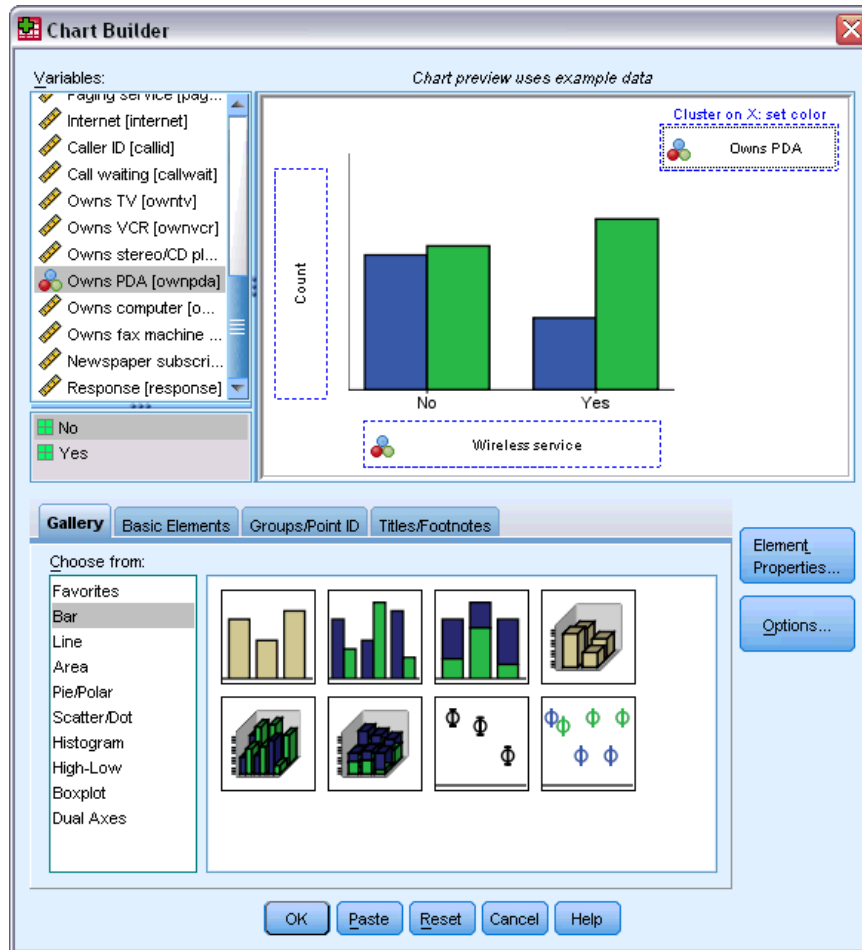
## Creating Charts

Although some statistical procedures can create charts, you can also use the Graphs menu to create charts.

For example, you can create a chart that shows the relationship between wireless telephone service and PDA (personal digital assistant) ownership.

- ▶ From the menus choose:  
Graphs > Chart Builder...
- ▶ Click the Gallery tab (if it is not selected).
- ▶ Click Bar (if it is not selected).
- ▶ Drag the Clustered Bar icon onto the canvas, which is the large area above the Gallery.

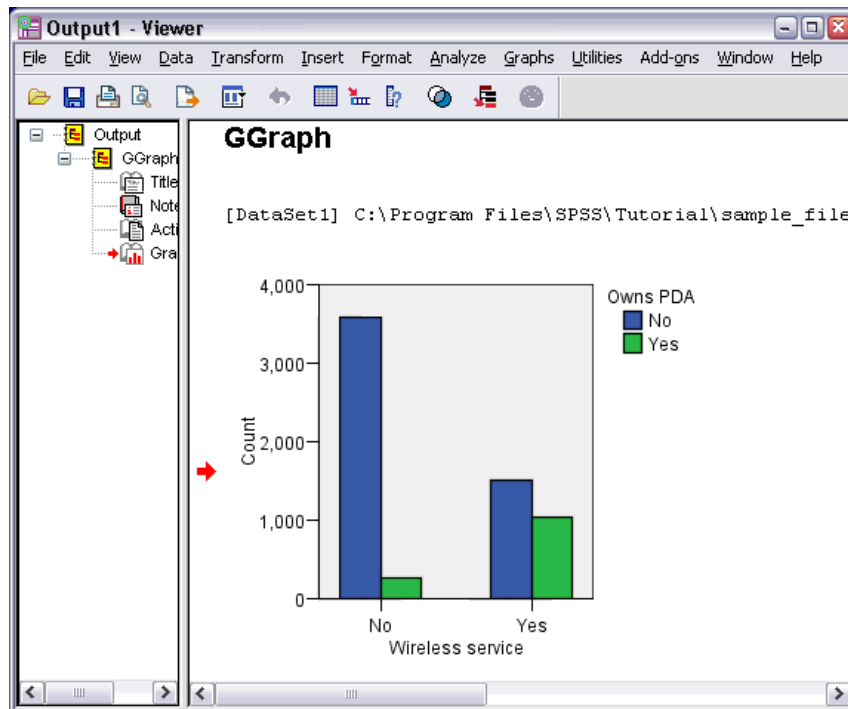
Figure 1-12  
Chart Builder dialog box



- ▶ Scroll down the Variables list, right-click *Wireless service [wireless]*, and then choose Nominal as its measurement level.
- ▶ Drag the *Wireless service [wireless]* variable to the x axis.
- ▶ Right-click *Owns PDA [ownpda]* and choose Nominal as its measurement level.
- ▶ Drag the *Owns PDA [ownpda]* variable to the cluster drop zone in the upper right corner of the canvas.
- ▶ Click OK to create the chart.



Figure 1-13  
Bar chart displayed in Viewer window



The bar chart is displayed in the Viewer. The chart shows that people with wireless phone service are far more likely to have PDAs than people without wireless service.

You can edit charts and tables by double-clicking them in the contents pane of the Viewer window, and you can copy and paste your results into other applications. Those topics will be covered later.