

Chapter 2

Entering Data

Chapter Table of Contents

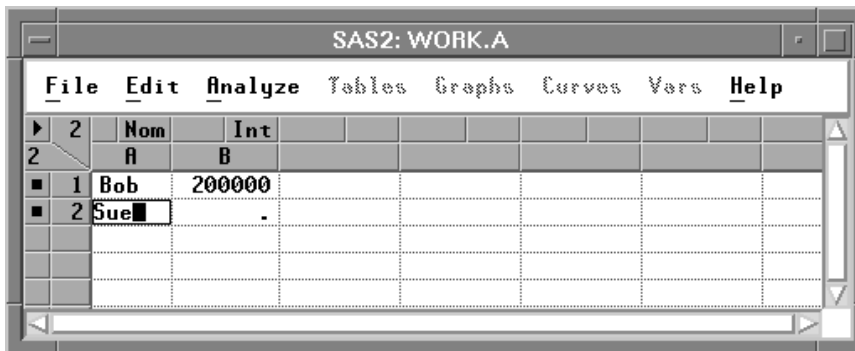
INVOKING SAS/INSIGHT SOFTWARE	28
ENTERING VALUES	31
NAVIGATING THE DATA WINDOW	33
ADDING VARIABLES AND OBSERVATIONS	34
DEFINING VARIABLES	36
FAST DATA ENTRY	39
OTHER OPTIONS	43

Chapter 2

Entering Data

A *SAS data set* consists of variables and observations. *Variables* are quantities or characteristics being measured. *Observations* are sets of variable values for a single entity.

In SAS/INSIGHT software, your data are presented in a window with variables displayed in columns and observations displayed in rows, as in Figure 2.1. You can enter data directly in the data window.



The screenshot shows a window titled "SAS2: WORK.A" with a menu bar containing "File", "Edit", "Analyze", "Tables", "Graphs", "Curves", "Vars", and "Help". Below the menu bar is a table with two columns labeled "A" and "B". The first row contains "Bob" and "200000". The second row contains "Sue" and ".". The table has a grid of dotted lines for data entry.

	Nom	Int							
2	A	B							
1	Bob	200000							
2	Sue	.							

Figure 2.1. Entering Data in the Data Window

Invoking SAS/INSIGHT Software

You can invoke SAS/INSIGHT software in any of three ways.

⇒ You can type **insight** on the command line.

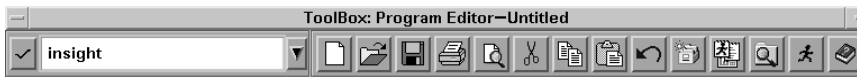


Figure 2.2. Command Line

⇒ If you have menus, you can choose **Solutions:Analyze:Interactive Data Analysis**.

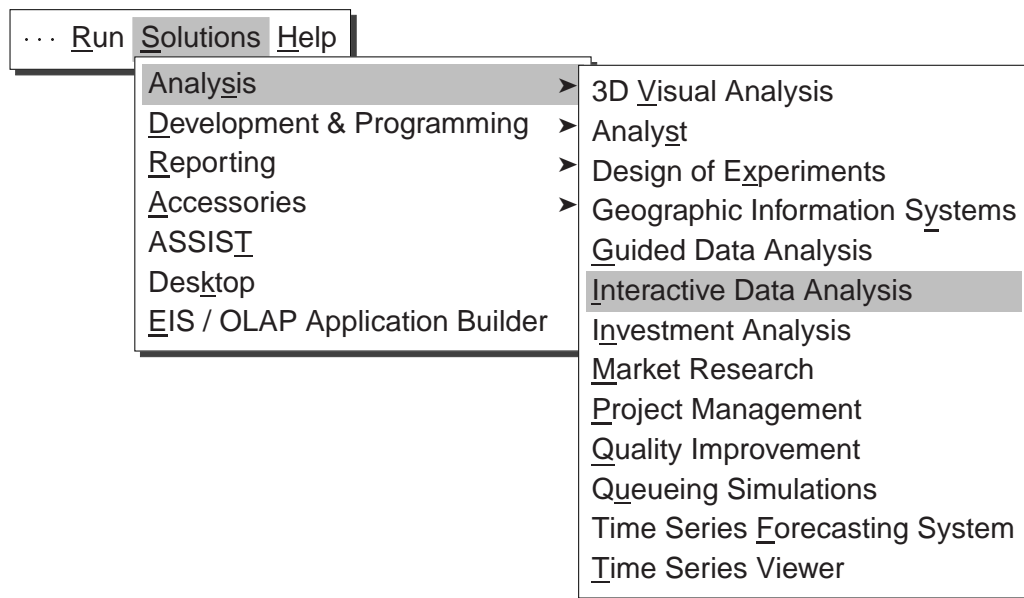


Figure 2.3. SAS Analysis Menu

⇒ You can invoke SAS/INSIGHT software as a SAS procedure.

Choose **Run:Submit** to submit the procedure statement in the Program Editor.

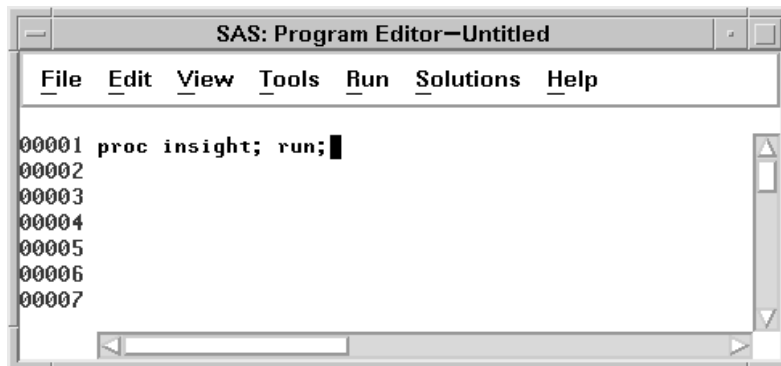


Figure 2.4. Entering a PROC Statement

You may want to access SAS data sets that are located in different libraries than the standard ones. As an example, if you have SAS data sets in a directory named **mypath**, then enter the lines

```
libname mylib 'mypath';
proc insight;
run;
```

in the Program Editor window and choose **Run:Submit**. The data set dialog (discussed later) will contain an additional library **mylib** to choose from.

You can invoke SAS/INSIGHT software from the Program Editor window and automatically open a new data window. Enter the lines

```
proc insight data;
run;
```

in the Program Editor window and choose **Run:Submit**. The data set dialog is skipped and a new data window appears.

You can specify a data set directly. For example, if you have a SAS data set named **mydata** in the **mylib** directory, enter the lines

```
libname mylib 'mypath';
proc insight data=mylib.mydata;
run;
```

in the Program Editor window and choose **Run:Submit**. Again the data set dialog is skipped and a data window appears with the specified SAS data set.

Finally, if you have *raw data* that you want to analyze, you most likely need to use the INFILE and INPUT statements in a DATA step. Refer to *SAS Language Reference: Dictionary* for information on how to read in raw data.

† **Note:** It is best to invoke SAS/INSIGHT software from the command line or from the **Solutions** menu. This enables you to use SAS/INSIGHT software simultaneously with other components in the SAS System. If you invoke it as a procedure, you cannot use any other SAS component until you exit SAS/INSIGHT.

Upon invoking SAS/INSIGHT software, you are prompted with a data set dialog.



Figure 2.5. Data Set Dialog

⇒ Click the **New** button.

This opens a new data window in which you can enter data.

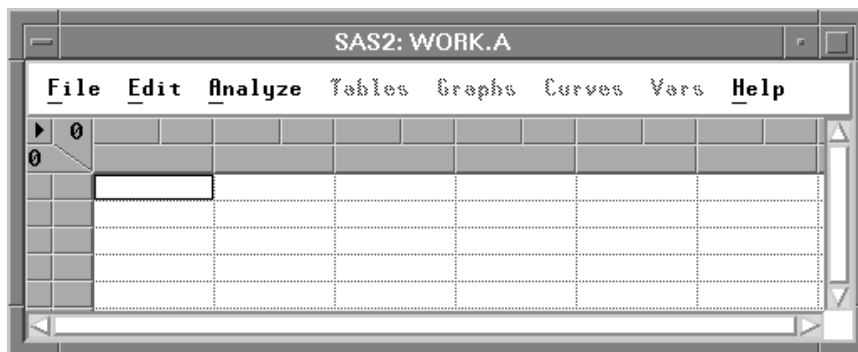


Figure 2.6. New Data Window

Entering Values

By default, the first value in a new data window is selected and is displayed with a frame around it. This *active value* marks your current location in the data window. To enter data, simply begin typing.

⇒ Enter the name “Bob” in the active value.

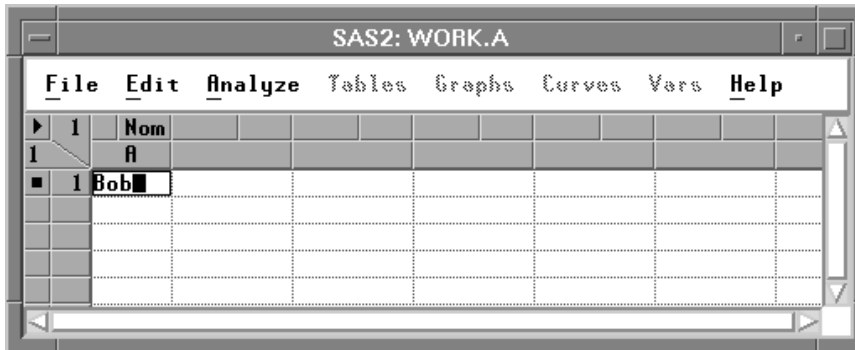


Figure 2.7. Entering a Value

As you type, variables and observations are created for you. The count of variables and observations is shown in the upper left of the data window.

⇒ Press the Tab key.

This moves the active value one position to the right.

⇒ Enter the salary “200000” in the active value.

Again, a variable is created.

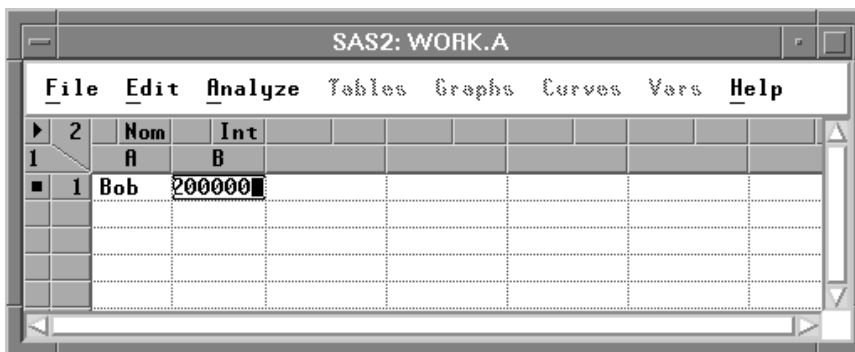
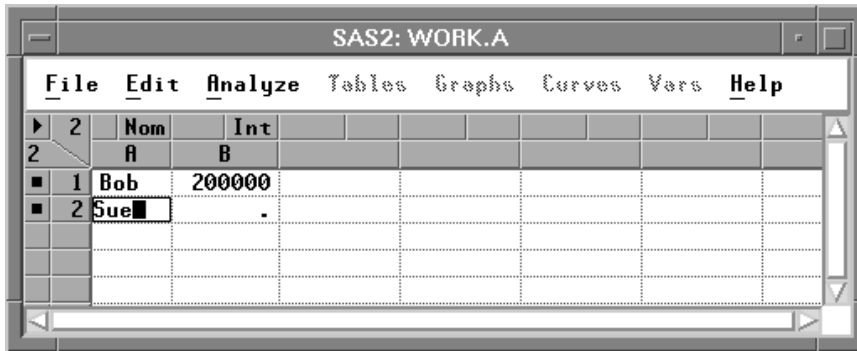


Figure 2.8. A Second Value

⇒ Press the down arrow key, then press the left arrow key.

This moves the active value to the first column of the second row.

⇒ Enter the name “Sue” in the active value.



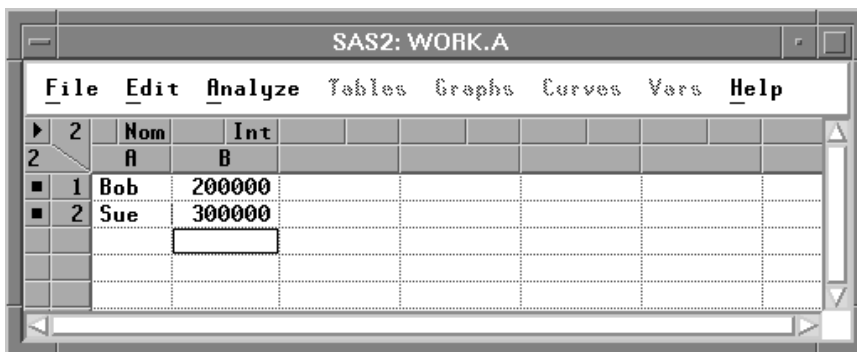
	Nom	Int							
2	A	B							
1	Bob	200000							
2	Sue	.							

Figure 2.9. A New Observation

A new observation is created, increasing the observations count to 2. The period (.) in the second value indicates a *missing* value for the numeric variable.

⇒ Press the **Tab** key to move to the right.

⇒ Enter the salary “300000” to replace the missing value. Then press the down arrow key.



	Nom	Int							
2	A	B							
1	Bob	200000							
2	Sue	300000							

Figure 2.10. Replacing the Missing Value

Navigating the Data Window

You can use Tab, BackTab, Enter, Return, and arrow keys to navigate the data window. Tab moves the active value to the right. BackTab, usually defined as Shift-Tab, moves the active value to the left. Enter or Return moves the active value down. Up and down arrow keys move the active value up or down.

When you are not editing any value, left and right arrow keys move the active value left and right. When you are editing a value, left and right arrow keys move the cursor within the active value.

When you have values, variables, or observations selected, the Tab, BackTab, and Return keys navigate within the selected area. This reduces keystrokes when you enter data.

⇒ **Drag a rectangle through several values to select them.**

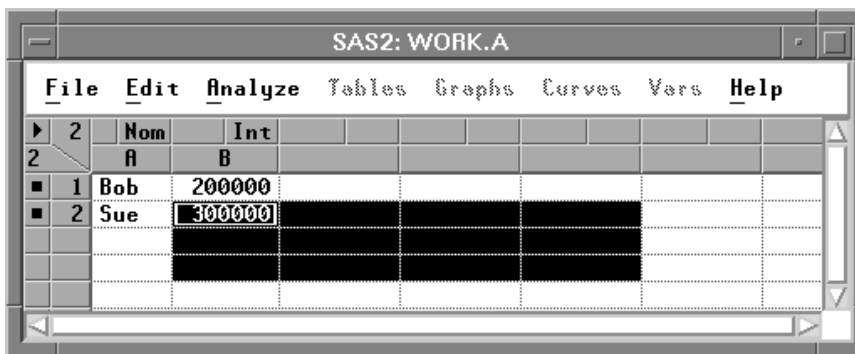


Figure 2.11. Selected Range

⇒ **Press Tab repeatedly.**

⇒ **Press Return repeatedly.**

The active value moves within the range you selected. By default, the Tab key navigates horizontally, and the Return key navigates vertically.

† **Note:** See the section “Data Options” at the end of this chapter for information on defining the direction of Tab and Enter keys.

Adding Variables and Observations

When you have a lot of data to enter, it is more efficient to specify the approximate number of observations rather than to create them one at a time.

⇒ **Click in the upper left corner of the data window.**

This displays the data pop-up menu.

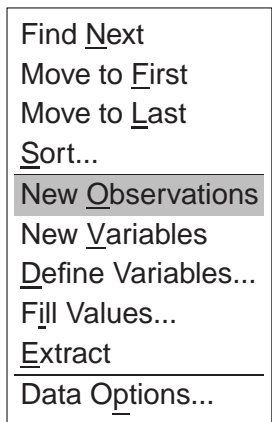


Figure 2.12. Data Pop-up Menu

⇒ **Choose New Observations from the pop-up menu.**

This displays a dialog to prompt you for the number of observations to create.

⇒ **Enter “10” in the observations dialog, then click OK.**

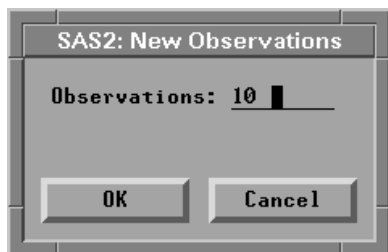


Figure 2.13. Observations Dialog

Observations with missing values are added at the bottom of the data window, increasing the observations count to 12. In the new observations, character values default to blank, while numeric values default to missing.

	Nom	Int							
12	R	B							
1	Bob	200000							
2	Sue	300000							
3		.							
4		.							
5		.							

Figure 2.14. New Observations

The **New Variables** menu works like the **New Observations** menu. You can choose **New Variables** to create several variables at once.

Defining Variables

Each variable has a *measurement level* shown in the upper right portion of the column header. By default, numeric values are assigned an *interval (Int)* measurement level, indicating values that vary across a continuous range. Character values default to a *nominal (Nom)* measurement level, indicating a discrete set of values.

- ⇒ **Click on the Int measurement level indicator for variable B.**
This displays a pop-up menu.

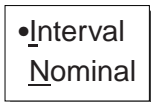


Figure 2.15. Measurement Levels Menu

The radio mark beside **Interval** shows the current measurement level. Because **B** is a numeric variable, it can have either interval or nominal measurement level.

- ⇒ **Choose Nominal in the pop-up menu to change B's measurement level.**

The screenshot shows a SAS data window titled 'SAS2: WORK.A'. The menu bar includes File, Edit, Analyze, Tables, Graphs, Curves, Vars, and Help. The data grid has two columns: 'A' and 'B'. Both columns are labeled 'Nom' in the header. The data rows are as follows:

	A	B
1	Bob	200000
2	Sue	300000
3		.
4		.
5		.

Figure 2.16. Nominal B

You can adjust other variable properties as well. Click in the upper left corner of the data window to display the data pop-up menu.

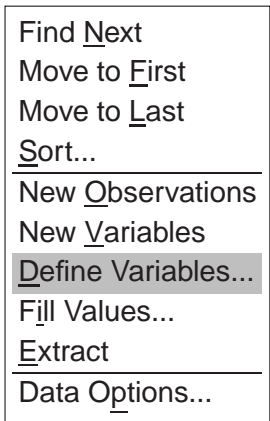


Figure 2.17. Data Pop-up Menu

⇒ Choose **Define Variables** from the pop-up menu.

This displays a dialog. Using this dialog, you can assign variable storage type, measurement level, default roles, name, and label.



Figure 2.18. Define Variables Dialog

⇒ Enter “NAME” for the name of variable **A**.

⇒ Click the **Apply** button.

In the data window, the variable receives the name you entered.

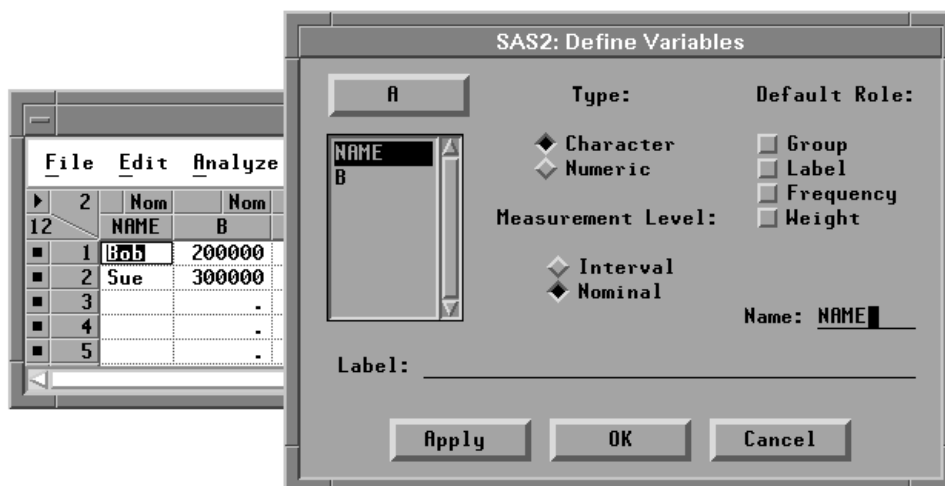


Figure 2.19. Naming a Variable

⇒ Select **B** in the variables list at the left.

⇒ Enter “SALARY” for the name of variable **B**.

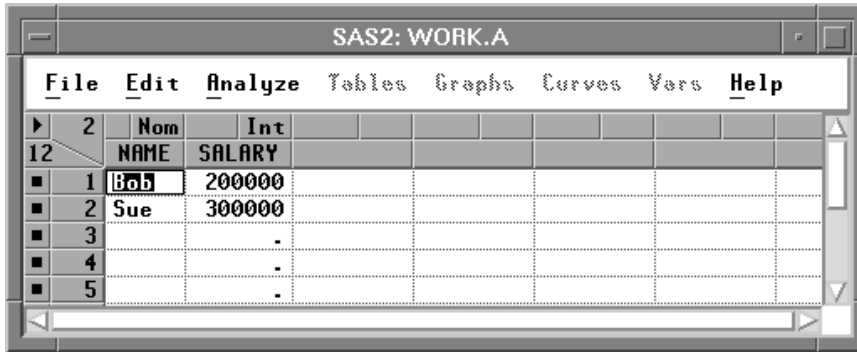
⇒ Click the **Interval** measurement level.

Interval measurement level is appropriate for a variable like salary.

⇒ Click the **OK** button.

Part 2. Introduction

This closes the dialog. In the data window, the variable receives the name and measurement level you entered.



The screenshot shows a SAS data window titled "SAS2: WORK.A". The window has a menu bar with "File", "Edit", "Analyze", "Tables", "Graphs", "Curves", "Vars", and "Help". Below the menu bar is a table with two columns: "NAME" and "SALARY". The table has five rows, with the first two rows containing data. The first row has "Bob" in the "NAME" column and "200000" in the "SALARY" column. The second row has "Sue" in the "NAME" column and "300000" in the "SALARY" column. The remaining three rows are empty. The table is displayed in a grid format with a scroll bar on the right side.

	Nom	Int							
12	NAME	SALARY							
1	Bob	200000							
2	Sue	300000							
3		.							
4		.							
5		.							

Figure 2.20. Name and Measurement Level Assigned

Fast Data Entry

When you have a lot of data to enter, it is important to be able to do it quickly. Using information from the preceding sections, here is the fastest way to enter data.

⇒ **Open a new data window.**

You can do this when you invoke SAS/INSIGHT software, or you can choose **File:New**.

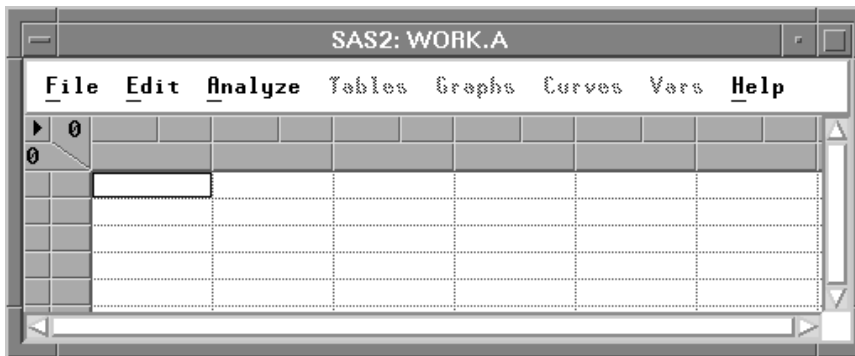


Figure 2.21. New Data Window

⇒ **Create all variables.**

The easiest way to do this is to enter the first observation. Variable types and measurement levels are assigned automatically.

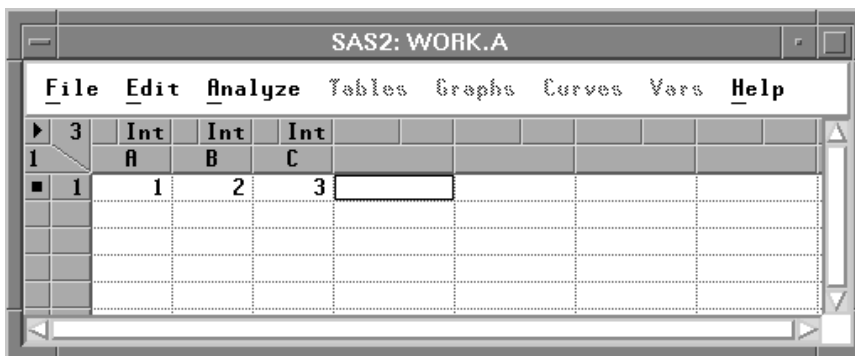


Figure 2.22. Variables Created Automatically

An alternate way to create variables and assign types and measurement levels yourself is by using the data pop-up menu.

⇒ **Click in the upper left corner of the data window.**

This displays the data pop-up menu.

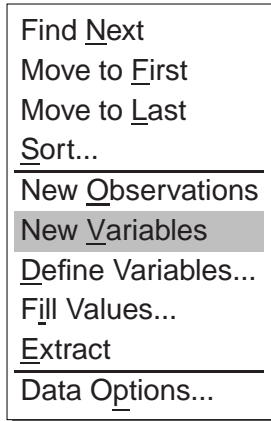


Figure 2.23. Data Pop-up Menu

- ⇒ **Choose New Variables from the pop-up menu.**
This displays a dialog to prompt you for the number of variables to create.
- ⇒ **Enter “3” in the New Variables dialog, then click OK.**

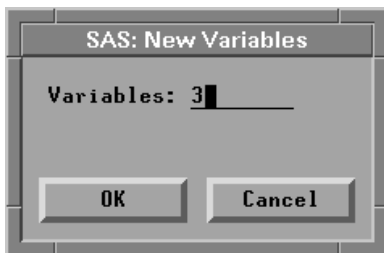


Figure 2.24. New Variables Dialog

The data window should appear as shown in the next figure.

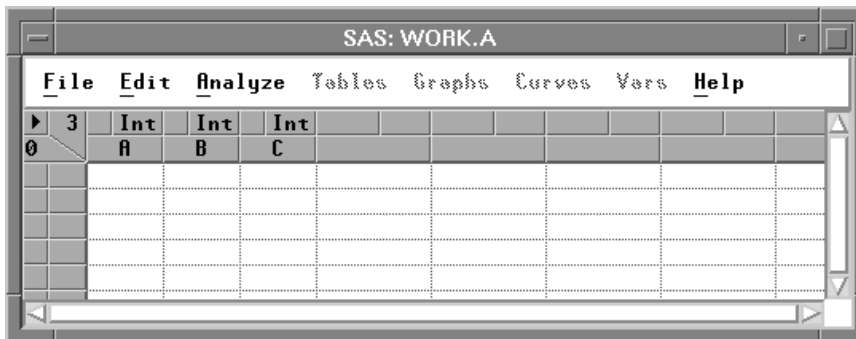


Figure 2.25. Variables Created Manually

The variable names and measurement levels can be selected as shown in the last section.

You can create observations using the following steps.

- ⇒ **Click in the upper left corner of the data window.**
This displays the data pop-up menu.

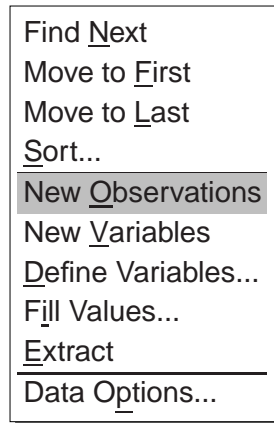


Figure 2.26. Data Pop-up Menu

⇒ **Choose New Observations.**

This displays a dialog prompting you for the number of observations to create.

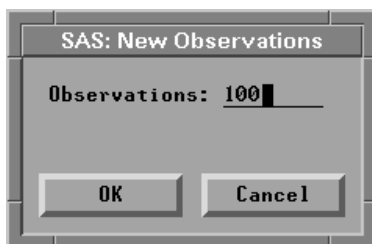


Figure 2.27. Observations Dialog

Enter the number of observations, then click **OK**. If you don't know the number of observations, make it a little larger than you will need. You can delete unused observations later.

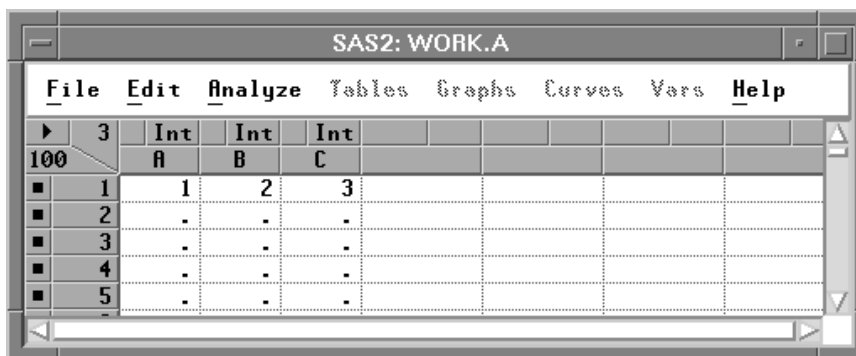


Figure 2.28. Observations Created

⇒ **Select all variables.**

Click the variable count in the upper left corner of the data window.

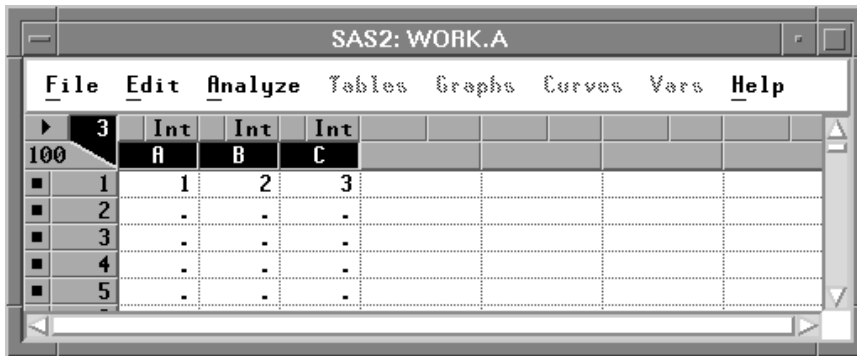


Figure 2.29. Variables Selected

⇒ **Select the active cell.**
 Use **Ctrl-click** to avoid deselecting the variables.

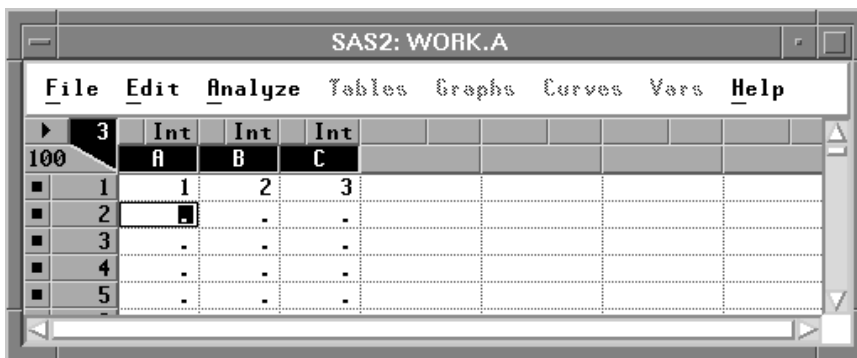


Figure 2.30. Active Value Selected

Now you can enter data, using Tab and BackTab to navigate within the selected variables. You can also fill in blocks of values by using the **Fill Values** option described in the next section. If your keyboard has a numeric keypad, this method enables you to enter numeric data without moving your hand from the keypad.

On some keyboards, the Enter key is easier to hit than the Tab key. So, you may be able to optimize data entry a bit further by defining the direction of the Tab and Enter keys. You can do this by setting the **Data Options** described in the next section. With these options, you can tailor SAS/INSIGHT's data entry to suit your keyboard.

When you have finished entering data, delete any unused observations by selecting them and choosing **Edit:Delete**. If you have not already done so, assign variable names, labels, and other information by choosing **Define Variables**.

Other Options

The pop-up data menu has a couple of useful options for filling in blocks of data and for selecting the actions taken by the Enter and Tab keys.

Click on the button at the upper left corner of the data window to display the data pop-up menu. Choose **Fill Values** to modify selected values in the data window. If you have variables, observations, or values selected, you are prompted to specify a **Value** and an **Increment**. If you have no selections, you are prompted to specify variables and observations.



Figure 2.31. Fill Values Dialog

In the **Fill Values** dialog, the **Value** field can be either character or numeric. If the value is numeric, you can use the **Increment** field to specify an increment or step value. For example, to fill 10 values with ordinals 1 through 10, you can select the values, choose **Fill Values**, and enter 1 for both **Value** and **Increment**.

Choose **Data Options** in the data pop-up menu to set options that control the appearance and operation of the data window. This displays the **Data Options** dialog,

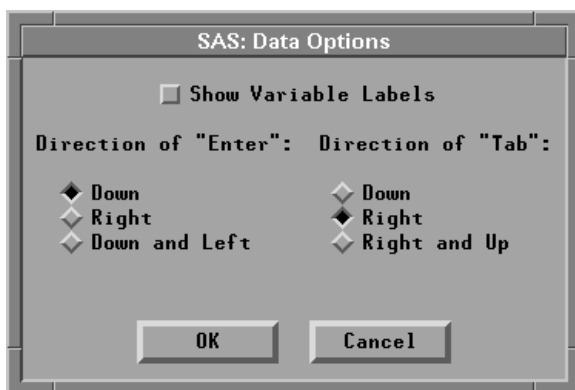


Figure 2.32. Data Options

The dialog contains the following options:

Show Variable Labels

This option controls whether variable labels are displayed. The default is off. If you turn on this option, variable labels are displayed.

Direction of “Enter”

This option controls the interpretation of the Enter and Return keys in the data window. By default, the Enter key moves the active value one position down. If you choose **Right**, the Enter key moves one position to the right. If you choose **Down and Left**, the Enter key moves one position down, and left to the first position.

Direction of “Tab”

This option controls the interpretation of the Tab and BackTab keys in the data window. By default, the Tab key moves the active value one position to the right. If you choose **Down**, the Tab key moves one position down. If you choose **Right and Up**, the Tab key moves one position to the right, and up to the first position.

The options **Down and Left** and **Right and Up** were added in Release 6.11. Not all hosts define a BackTab key, and not all hosts define Enter and Return as the same key. Consult your host documentation for information on key definitions.

You can save data window options by choosing **File:Save:Options**. This enables you to use your preferred option settings as defaults in future SAS/INSIGHT sessions.