

SPSS Information Sheet 1

Entering Data into SPSS by Hand

We will use SPSS as the statistics package for this course. There are many other statistics packages — SAS, Minitab, BMDP and others — which you will normally be free to use. SPSS is an easy one to learn and is available on most OU computers. It is menu-based and has extensive on-line support. When using a statistics package on a computer to do projects and/or assignments, you are free to use whichever one you wish. When doing the SPSS assignments for this class, you are required to use SPSS. SPSS is the only package that will be “supported” by the instructor — that is, it is the only one on which I am willing to answer questions regularly.

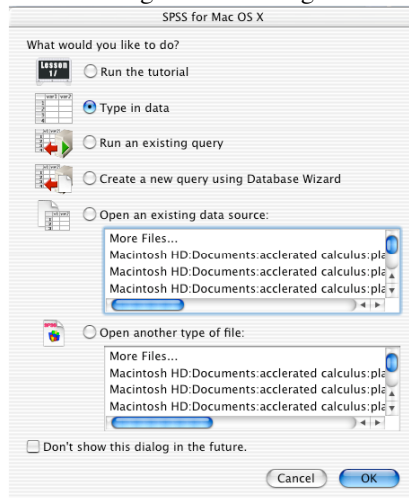
The first task in using any statistics package is getting a data set for analysis entered into the package on your computer. That is the topic of this information sheet. There are two major ways data can be entered.

1. Type it into the package by hand using a keyboard
2. Import it from a floppy disk or other source in which it has already been entered

This information sheet will deal with the first method. A later sheet will deal with the second.

Starting SPSS

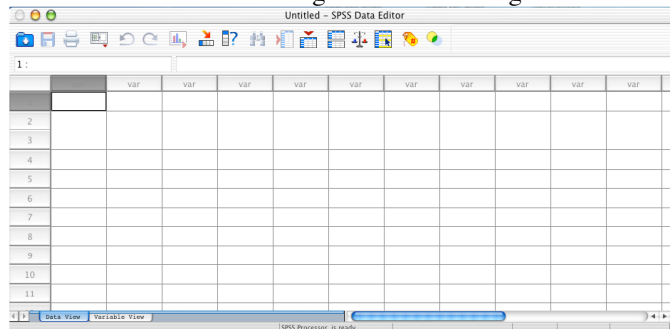
The first thing you must do on your computer is to start the SPSS package. In Windows, clicking first on “Start” and then on “Programs” usually does this. If SPSS is on that computer, it should be listed under Programs or Applications. If so, choose it. If not, find another computer. Choosing SPSS should bring up a dialog box that gives you options about entering data and using SPSS:



The default choice is usually to “Open an existing data source.” In this case, I have chosen “Type in data” since that is the purpose of this information sheet. At this point, clicking “OK” will bring up a blank SPSS “Data Editor” window.

Using the Data Editor Window to Define Variables

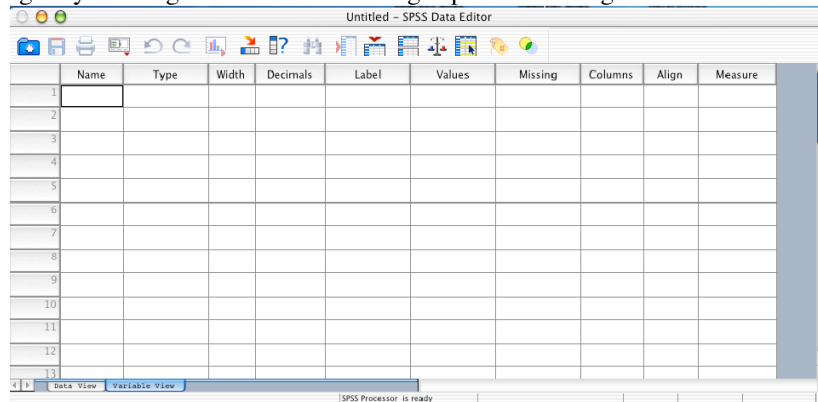
A blank Data Editor window should look something like the following:



This window is used to define variables and to enter data about them. Notice two views that show as “tabs” in the lower left-hand corner of the window:



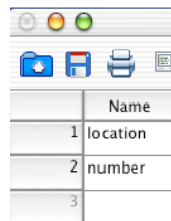
The default is that SPSS opens in the “Data View.” In this case, we want to begin with the “Variable View,” so we begin by clicking on this tab. This brings up the following view:



For this information sheet, we want to use the data that is in Problem 2.2 of the textbook. It deals with the number of cornrakes, an endangered species of bird, on which a census was taken in Britain and Ireland in 1993. This data set involves two variables:

1. The area or location in Britain or Ireland for which part of the census of cornrakes was taken; and
2. The number of cornrakes observed in each of these geographical areas

We thus want to name two variables. Let’s call them “location” and “number.” *An SPSS variable name (1) begins with a letter, (2) does not contain a space, and (3) is eight or fewer characters long.* We type those two names onto the variable “Name” column in the first two rows of the Data View of SPSS:

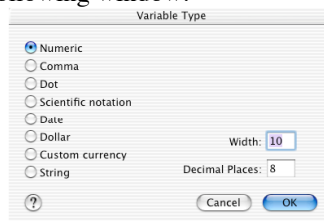


When we type in these two variable names, certain other information is automatically entered:

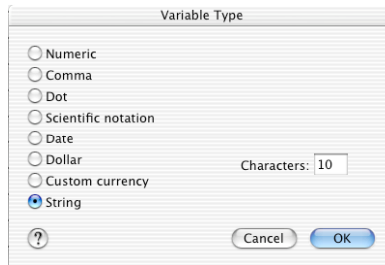
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	location	Numeric	10	8		None	None	8	Right	Scale
2	number	Numeric	10	8		None	None	8	Right	Scale
3										

We may or may not want to keep this information as entered.

For example, “location” is actually a “string” or “character” variable. Notice in the picture above that its “Type” is listed as a “Numeric” variable, the default. Suppose we want to change “location” to a character or string variable. Notice the grey square in the cell labeling it as “Numeric” in the picture above. If we click on this grey square, we get the following window:



Notice that the default was to indicate that this was a string variable and to give it 10 digits width with 8 places to the right of the decimal. If we want to change it to a string variable, we click in the button for “String” and we get the following window:



This indicates that the variable “location” is now a string (character) variable that has been given 10 characters width. We know that an SPSS variable has at most 8 characters so we could click and change that 10 to an 8 if we wished. If we clicked OK, we would then have changed “location” to a string variable as the Data View window would now indicate:

	Name	Type	Width
1	location	String	10
2	number	Numeric	10

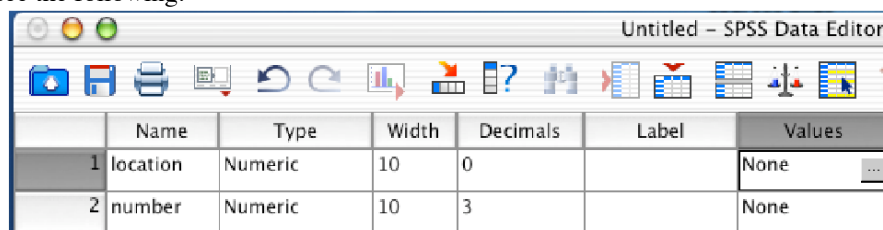
Notice also that the “number” variable is, in fact, a “Numeric” variable with a default width of 10 digits of which 8 are to the right of the decimal. If you look at the data, you will see that all of the numbers are whole numbers so we really need no digits to the right of the decimal place. If we click on the cell for the “Decimals,” the option of changing them to 0 digits to the right of the decimal appears, as would be appropriate for whole numbers:

	Name	Type	Width	Decimals
1	location	String	10	0
2	number	Numeric	10	8

This would then change “number” to a numeric variable which is allotted no decimal places (that is, the data would be whole numbers) and is still allotted 10 digits (which we could change if we wanted to).

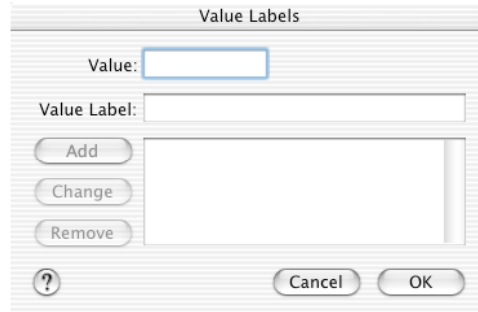
Alternative Treatment of Character Variables

There is an alternative way to treat character variables like “location.” If we go back to the default window where we hadn’t changed anything about the variables, we could click on the “Values” cell for “location” and would see the following:

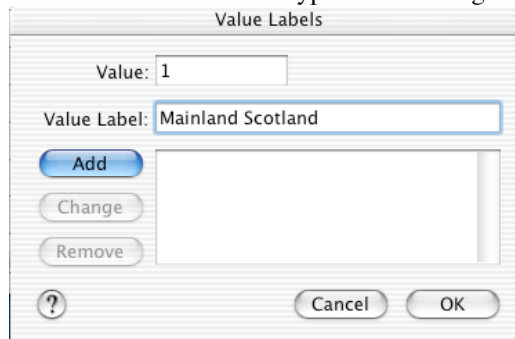


Notice “location and “number” are both still “Numeric” variables with a width of 10 and 0 and 3 decimals (0 for “location” because I changed it to a numeric variable back from a string variable which automatically has 0 decimal places).

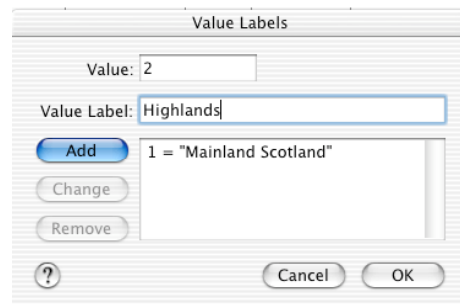
You can see that “None” is the default entered in the cell for “Values” but a grey square now appears in the Values cell for “location that we clicked. If we click on it, we will see the following window:



This window actually has complex possibilities. For example, suppose we want to code the locations numerically using the numbers 1 through 10. If we did, we would want to label them with the names of the locations that go with each of the code numbers 1 through 10. Suppose we want to enter “Mainland Scotland” as the label that goes with code 1. We would type the following and have the window:



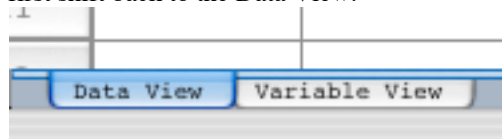
If we now click “Add” button, this Value Label will be recorded for Value 1 of the variable “location. We can then type “Highlands” or whatever we wish for the Value Label for the Value 2 for “location” and we would have:



If we click “Add” then “2 = “Highlands”” would be added to the value labels recorded. We could continue until we had value labels for all 10 values of the numeric variable “location.” It is mainly a matter of personal preference whether we treat “location” as a coded numeric variable or as an uncoded string variable. SPSS provides for doing either.

Entering Data

We have now defined the variables sufficiently that we can enter the data for each case. We do this variable by variable for each case. We first shift back to the Data View:



In this view, columns will now be the variables we have defined and each row represents a separate case.

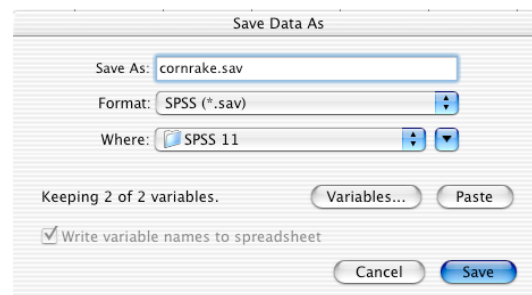
We begin to enter the data and get something like:

	location	number	var
1	1	12.000	
2	2	15.000	
4			

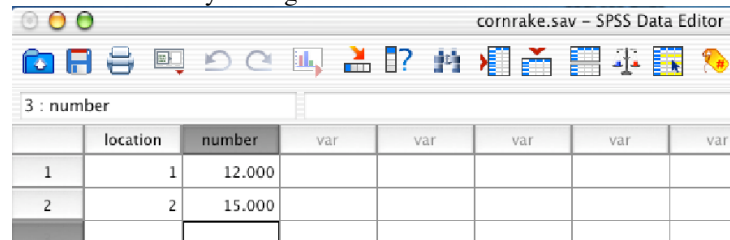
We can continue in the same way until all of the data are entered.

Saving Data Sets

Once data have been entered into an SPSS “Data Editor” window, we have a data set in SPSS. We can then do any appropriate analysis that SPSS allows for with that data set. However, we may need to work on the data at different times and we certainly would not want to lose the work we have done so far. At this point, we should save the data set. Under the “File” menu, we choose “Save As” and get something like the following window:



This allows us to choose a format like “.dat” or “.txt.” The default is the SPSS data set format, “.sav,” and that’s as good as any unless you plan to use the data set with another program. You may also choose where you want to save the data. The above screenshot is from a Mac. In Windows, if you choose “Save,” rather than seeing “Where” you should see a default pathname <c:\program files\spss>. You will probably want to save the data set on your hard drive or on a floppy disk if you are using a shared computer. So you should change the path name appropriately, perhaps to something like <a:\> and give the data set a name to complete the pathname (as I have done with the word “cornrake” and the automatic suffix “.sav”). This will be reflected after a few seconds by adding the name of the saved data set to the data editor window:



This data set now resides in the location (for example, the floppy disk) to which you saved it.

Exercises

1. Set up the variables for a data set in SPSS for the data from Problem 2.2. Leave “location” as a numeric variable and enter data value labels for all 10 locations. Leave “number” as a numeric variable. Set the decimal places of each variable to 0.
2. Enter all of the data from Exercise 2.2.
3. Save the data set on a floppy disk or your personal hard drive as “cornrake.sav.”
4. Under “File” choose “Print” and print your data set to turn in.