Files

The "heart" of LazStats or any other statistics package is the data file to be created, saved, retrieved and analyzed. Unfortunately, there is no one "best" way to store data and each data analysis package has its own method for storing data. Many packages do, however, provide options for importing and exporting files in a variety of formats. For example, with Microsoft's Excel package, you can save a file as a file of "tab" separated fields. Other program packages such as SPSS can import "tab" files. Here are the types of file formats supported by LazStats:

1.Text type of files (with the extension .LAZ) NOTE: the file format in this text file is unique to LazStats!
2.Tab separated field files (with the file extension of .TAB.)
3.Comma separated field files (with the file extension of .CSV.)
4.Space separated field files (with the file extension of .SSV.)

My preference is to save files as .LAZ and .TAB files. This gives me the opportunity to analyze the same data using a variety of packages. For relatively small files (say, for example, a file with 20 variables and 1000 cases), the speed of loading the different formats is similar and quite adequate.

Creating a File

When LazStats begins, you will see a "grid" of two rows and two columns. The left-most column will automatically contain the word "Case" followed by a number (1 for the first case.) The top rows will contain the names of the variables you have defined. You can change the name of the variables and define additional variables by clicking on the menu item labeled "VARIABLES" and then clicking on the "Define" option. A "form" will appear that looks like the figure below:

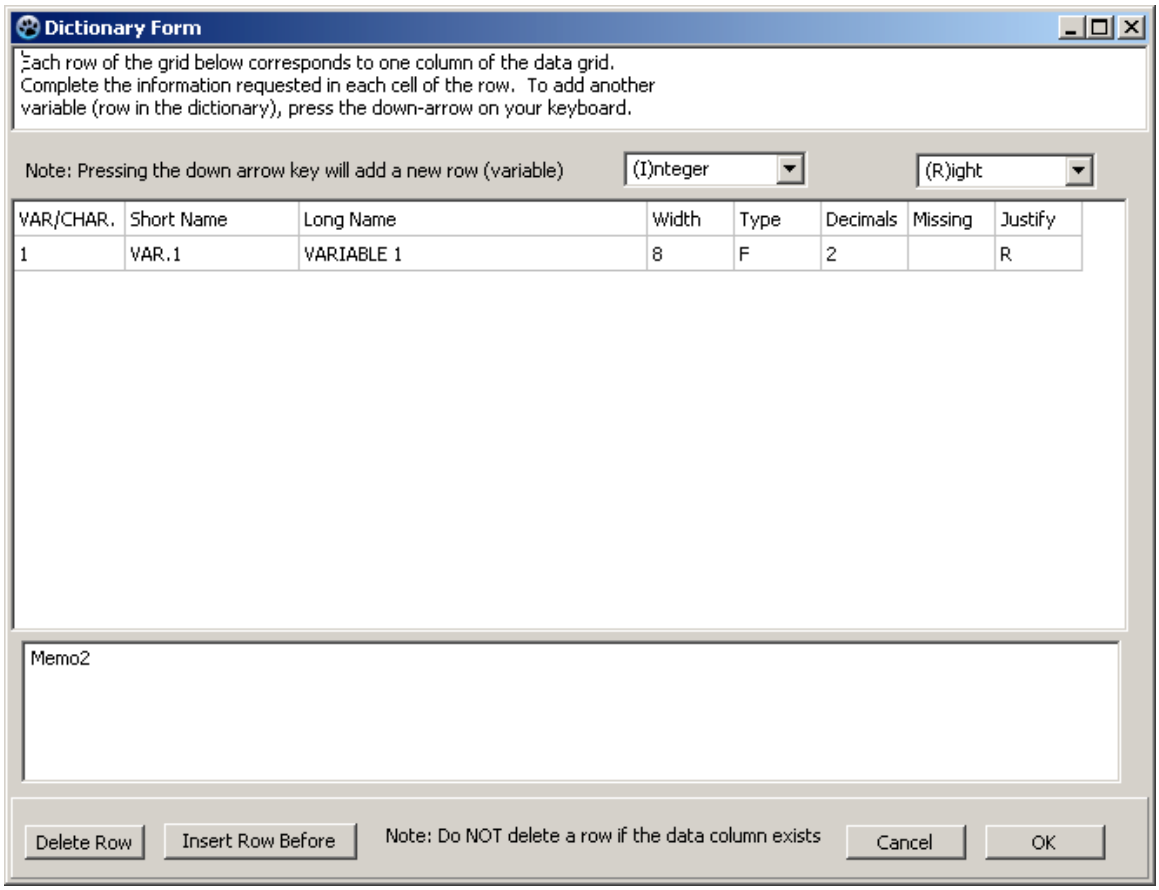

Figure 1. The Variable Definition Form

In the above figure you will notice that a variable name was automatically generated for the first variable. To change the default name, double click the box with the default name and enter the variable name that you desire. It is suggested that you keep the length of the name to eight characters or less. You may also enter a longer label for the variable. If you save your file as a .LAZ file, this long name (as well as other descriptive information) will be saved in the file (the use of the long label has not yet been implemented for printing output but may be in future versions.) To proceed, simply click the OK button in the lower right of this form. The default type of variable is a "floating point" value, that is, a number that may contain a decimal fraction. If a data field (grid cell) is left blank, the program will usually assume a missing value for the data. The default format of a data value is eight positions with three positions allocated to fractional decimal values (format 8.3.) By clicking on any of the specification fields you can modify these defaults to your own preferences. You can change the number of decimal places (0 for integers.) You will find that some analyses require that a variable be defined as an integer and others as floating point values. The drop-down box labeled "(I)nteger" lets you click on the type of variable you are defining and automatically record the character value that defines that type. If you press the "down-arrow" on your keyboard, another variable with default values will be added. You can also click a large button at the bottom to add delete or insert a variable. Another way to specify the default format and missing values is by modifying the "Options" file. When you click on the Options menu the following form appears:

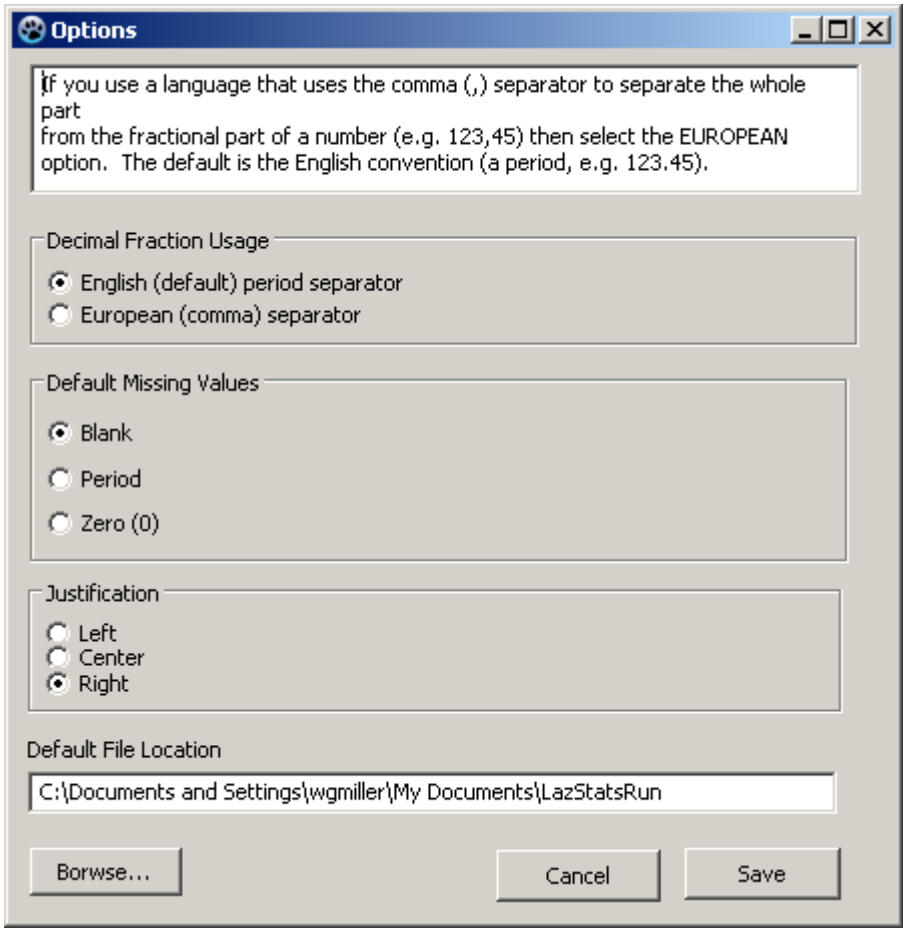

Figure 2. The Options Menu

In the options form you can specify the Data Entry Defaults as well as whether you will be using American or European formatting of your data (American's use a period (.) and Europeans use a comma (,) to separate the integer portion of a number from its fractional part.) You can double click the Default File Location edit box to change the path to your files. In many countries, the separation of the whole number from the fractional part of a floating point number is a comma (,) and not a period (.) as in the United States. A user that uses the comma separator is designated a "European" user. The default is the American usage. It is possible to convert one type to another. The example files all use the American standard. If you use the European standard, you will need to examine the "default" confidence intervals shown on many of the statistics dialog forms – they may have a period (e.g. 0.05) instead of a comma (0,05) as needed in the European format. One can click on the value and change it to an appropriate format.

Entering Data

When you enter data in the grid of the main form there are several ways to navigate from cell to cell. You can, of course, simply click on the cell where you wish to enter data and type the data values. If you press the "enter" key following the typing of a value, the program will automatically move you to the next cell to the right of the current one or down to the next cell if you are at the last variable. You may also press the keyboard "down" arrow to move to the cell below the current one. If it is a new row for the grid, a new row will automatically be added and the "Case" label added to the first column. You may use the arrow keys to navigate left, right, up and down. You may also press the "Page Up" button to move up a

screen at a time, the “Home” button to move to the beginning of a row, etc. Try the various keys to learn how they behave. You may click on the main form’s Edit menu and use the delete column or delete row options. Be sure the cursor is sitting in a cell of the row or column you wish to delete when you use this method. A common problem for the beginner is pressing the "enter" key when in the last column of their variables. If you do accidentally add a case or variable you do not wish to have in your file, use the edit menu and delete the unused row or variable. If you have made a mistake in the entry of a cell value, you can change it. In the grid cell you can use the delete key, backspace key, enter characters, etc. to make the corrections for a cell value Notice that as you make grid entries and move to another cell, the previous value is automatically formatted according to the definition for that variable. If you try to enter an alphabetic character in an integer or floating point variable, you will get an error message when you move from that cell. To correct the error, click on the cell that is incorrect and make the changes needed.

Saving a File

Once you have entered a number of values in the grid, it is a good idea to save your work (power outages do occur!) Go to the main form’s File menu and click it. You will see there are several ways to save your data. A “dialog box” will then appear as shown below for a .LAZ type of file:

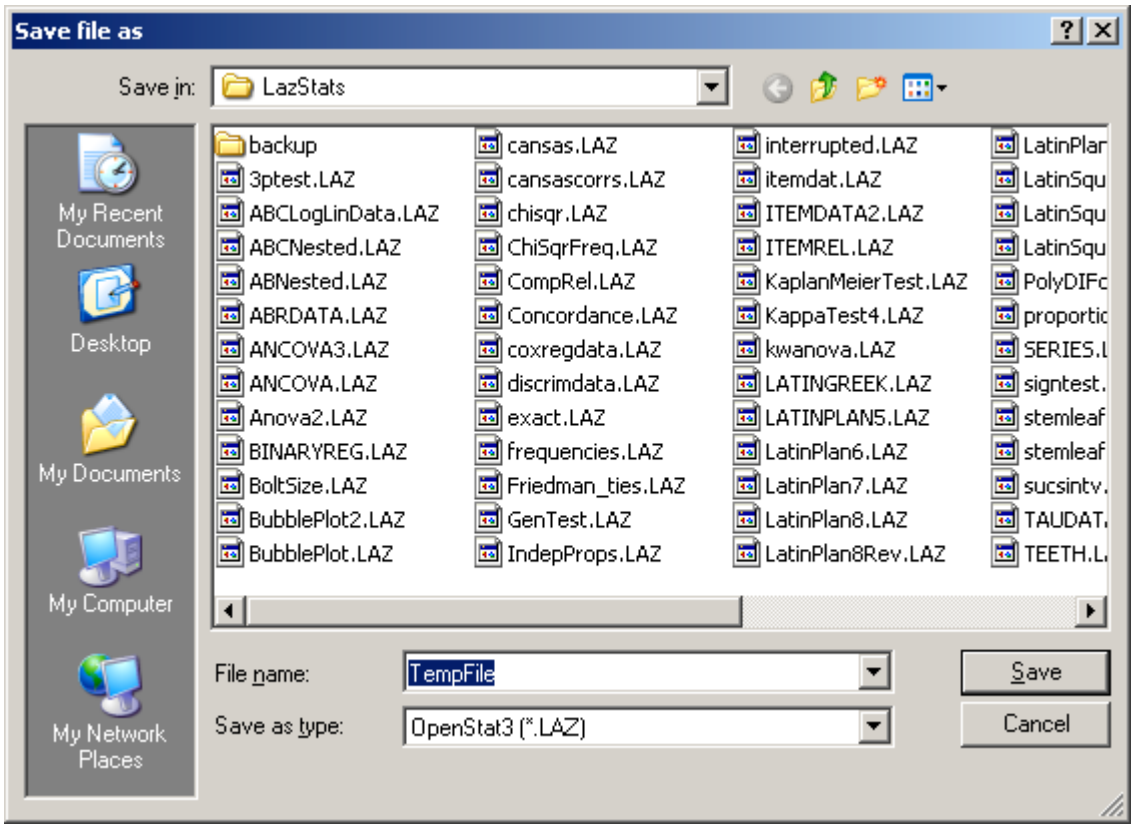

Figure 3. The Save Dialog Form

Simply type the name of the file you wish to create in the File name box and click the Save button. After this initial save operation, you may continue to enter data and save again. Before you exit the program, be sure to save your file if you have made additions to it. Notice that .LAZ file type is labeled as an OpenStat3 type of file. This is because LazStats is a conversion of a previously developed Borland Delphi Pascal program called “OS3”.

If you do not need to save specifications other than the short name of each variable, you may prefer to "export" the file in a format compatible to other programs. The Export Tab File option under the File menu will save your data in a text file in which the cell values in each row are separated by a tab key character. A file with the extension .TAB will be created. The list of variables from the first row of the grid are saved first, then the first row of the data, etc. until all grid rows have been saved. If there are blanks in any value cells, the default missing value will be written for that cell. Alternatively, you may export your data with a comma or a space separating the cell values. Basic language programs frequently read files in which values are separated by commas or spaces. If you are using the European format of fractional numbers, DO NOT USE the comma separated files format since commas will appear both for the fractions and the separation of values - clearly a design for disaster!