Chapter One

Introduction and Quick Start

Overview

Flo•Stat™ 2 is a powerful statistical analysis and graphics package designed for use in the educational and business setting. Its ease of use and basic statistical routines make it ideal for a first semester statistics course. *Flo•Stat's* data importing and exporting capabilities, business graphics, industry-standard tabular output, and ability to manipulate moderately large data files makes it the ideal analytic tool for a wide range of research needs.

Flo•Stat is named in honor of Florence Nightingale, the 19th century social activist and statistician. While best known for her humanitarian efforts as a nurse and health care administrator, Florence Nightingale's use of social statistics and statistical graphing to improve health conditions are especially noteworthy. In particular, her efforts during the mid 1800's to improve the conditions for the British soldier, in part through the use of social statistics, not only saved tens of thousands of lives, but help to foster the idea that social phenomena could be measured and mathematically analyzed. As noted by Cohen (1984), her accomplishments in the field of applied statistics included: 1) design of the Polar-Area diagram, used to graphically represent change over time, 2) the first statistical evidence that malnutrition, exposure and disease (i.e., dysentery, cholera, typhus, and scurvy) ,and not the injuries suffered in battle, were, at that time, the leading cause of death among soldiers in wartime, 3) evidence that mortality among British soldiers during peacetime was twice that of males in civilian life--helping to institute wide ranging sanitary reforms in the army's medical service and physical improvements in hospitals, and 4) lobbying to have the study of statistics introduced into higher education and made a mandatory part of public education.

Equipment requirements

Flo•Stat requires a minimum of 1mb of RAM but performs best with two or more megabytes of RAM when working with moderately-sized data sets (e.g., 50 variables and 2,000 cases). *Flo•Stat* runs on Macintosh II class machines or higher, and requires, at a minimum, system 6.0.7.

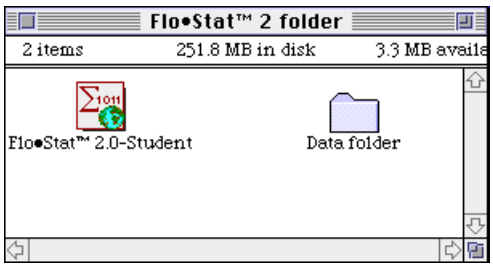
Installing *Flo•Stat 2*

Be safe and make a backup copy of the *Flo•Stat* disk. Store the original disk in a secure place and complete your work with the copy.

The *Flo•Stat* disk holds the *Flo•Stat 2* application and a tutorial folder containing several small data sets: the tutorial data file and a subset of the 1992 General Social Science data file.

The *Flo•Stat* application can be launched from the floppy disk. Performance is enhanced if the application is launched from a hard drive. To copy the

contents of the disk to a hard drive, create a new folder on the hard drive, select all icons on the *Flo•Stat* disk and drag then into the new folder.



FLO•STAT is not copy protected, making it easier to use and install on your computer. Please don't make the mistake of producing copies for friends and colleagues. Illegally copying software hurts future development efforts. Support Senecio Software's efforts to enhance new versions of Flo•Stat and our other specialty software (i.e., MaCATI and IPSS) by encouraging others to call 419-352-4371 and order their very own copy.

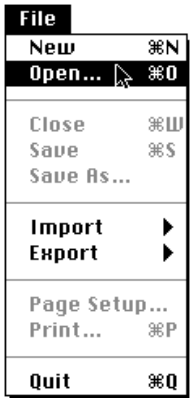
Quick Start

Double click the *Flo•Stat* icon to launch the application.

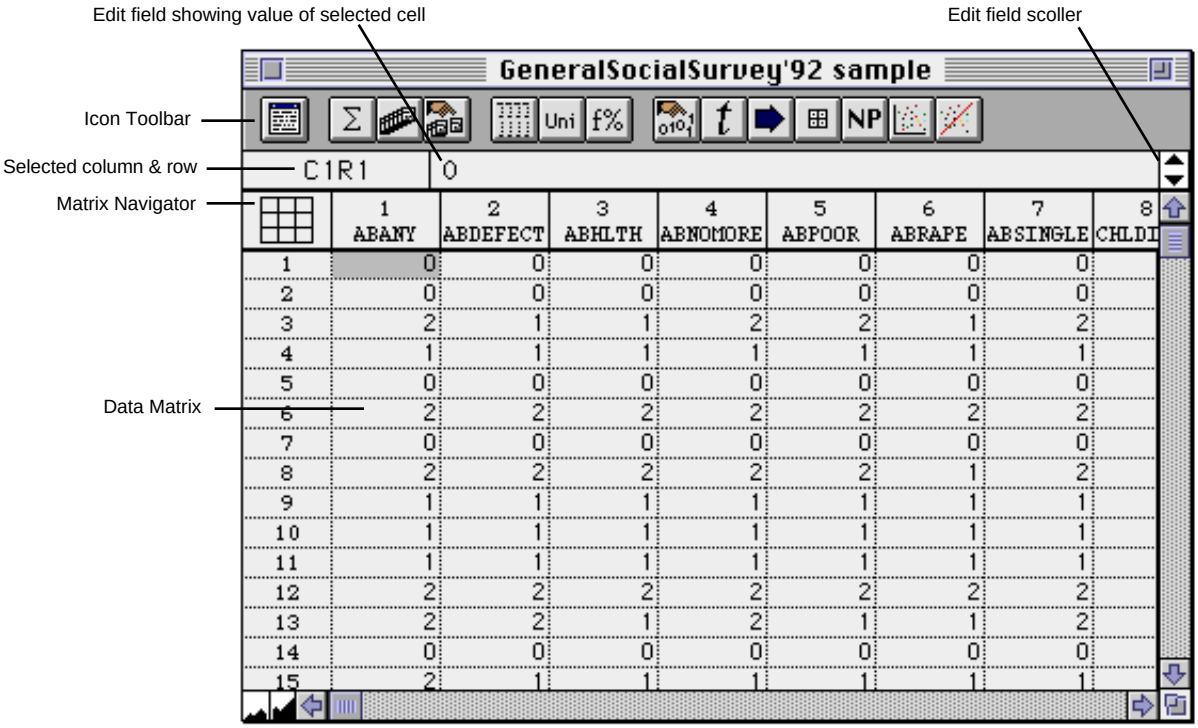
When launched directly, Flo•Stat always opens to an empty data matrix. Both numerical and string values can be typed directly into the matrix cells or pasted from another application, such as data copied from a spreadsheet.

Close the empty data matrix by either clicking the go-away box at the top left corner of the data window, or selecting **Close** from the **File** menu. More than one data set can be opened at a time in Flo•Stat. The number of data sets opened at any one time is limited by available random access memory (RAM)..

Open the data file named **General Social Survey** by selecting **Open** from the **File** menu.



When opened, the data matrix should look similar to the one below.



The data set file name is shown at the top of the window. Just below the title bar a menu of icons eases the process of documenting, transforming, sorting, selecting, and statistically analyzing the information contained in the data matrix.

The two fields just below the icon menu bar display the column and row location of the currently selected data cell and the value contained in that cell.

The small, nine-cell grid at the upper left corner of the data matrix is called the matrix navigator. To move quickly from the current cell location to one of the outer edges or center of the matrix, simply click the appropriate spot on the matrix navigator.

The numbers down the left side of the matrix represent cases in the data set. Each row contains one case. Depending on a study's unit of analysis, each case might represent a person in a survey, a patient's record from all hospital discharges, a tree in a sample of 100 selected by a biologist, or a company in an economic analysis of corporate business practices. The student version of Flo•Stat 2 can store up to 2,000 cases. The number of cases in the professional version of Flo•Stat 2 is limited only by available disk storage.

Each column in the matrix is numbered in sequence, beginning with the number one. Columns contain the values for each variable in the data set. In a telephone survey of the general public, for instance, each respondent's answers to questions about their age, race, gender, family income, street address, state of residence, and so on, are variables. Flo•Stat 2 data sets can store up to 100 variables in the student version and 1,200 variables in the professional version.

Each set of letters and characters below a column's number is the name given to each variable by the person creating the data set. These names, like all labels in the data set, can be changed at any time. Variable names help you identify one column of data from the next, especially when the data set contains many variables.

Examining several variables in the data set

The **Reports** procedure makes it easy to generate simple lists as well as lists subdivided within categories of other variables.

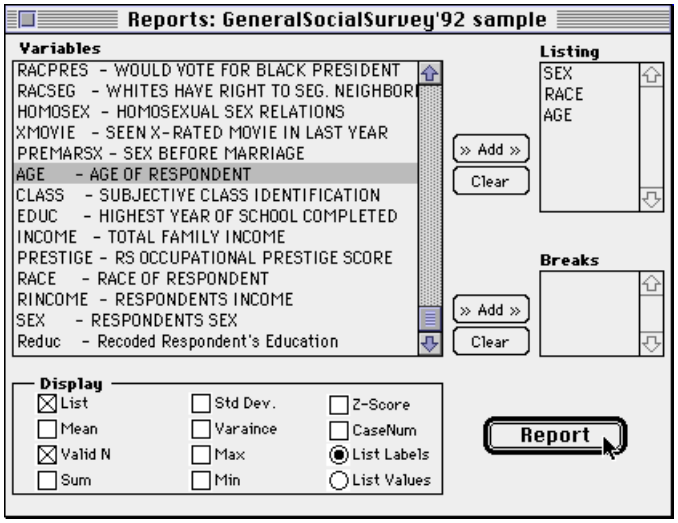
Click once on the Report icon (shown at the left) in the menu bar.

Scroll to the bottom of the variable list presented in the **Reports** window and select the variable named SEX then click the **Add** button to the right of the variable list. This action adds the SEX variable to the report. Repeat these steps for the RACE and AGE variables.

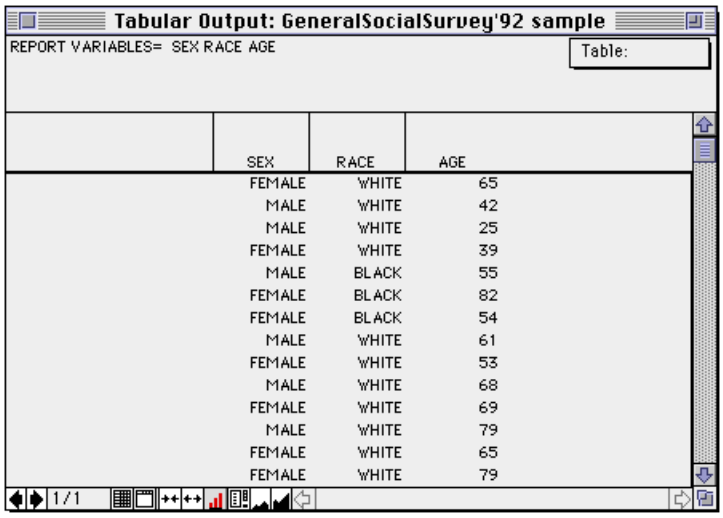
Click the **Valid N** button in the **Display** section at the bottom of the window.

When the **List Labels** button is checked, the report will display the labels, if present, for each variable rather than the raw data shown in the matrix.

Click the **Report** button after selecting the variables: SEX, RACE, AGE.



The report shows a simple listing of the selected variables. The data values (or labels) are presented for each case in the order they appear in the data matrix.



Examine the data by scrolling up and down the list.

The valid number of cases is displayed at the bottom of the list because the **Valid N** button had been checked in the **Reports** window. In the same way, summary statistics can be generated for each variable in the list.

The entire list can now be printed, saved as a separate text file, or copied to the clipboard and pasted to a word processor or spreadsheet if desired.

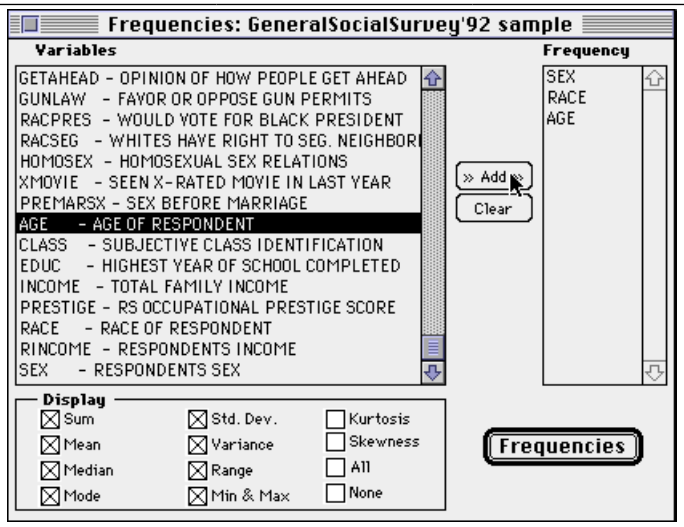
Frequency and percentage distributions

It's hard to see certain patterns or associations in your data when so much information is displayed in the form of a list. The information contained in the list needs to be collapsed and categorized in some fashion before simple questions can be addressed, such as, "How many people in the sample are under the age of 20?" or "What is the average age of the 200 people in this sample?"

Click the **Frequencies** icon (shown at the left) in the menu bar to obtain a distribution of the values in each variable.

Scroll to the bottom of the variables list in the **Frequencies** window and select and add the variables **SEX**, **RACE** and **AGE to the Frequency list**.

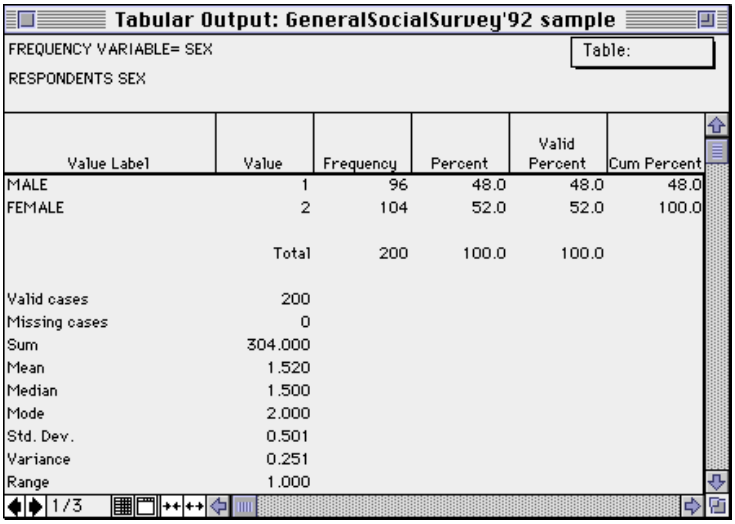
Click the **Frequencies** button to begin the calculations.



The tabular output window contains the frequency, percent, valid percent and cumulative percent distributions for each of the variables selected. If value labels (e.g., MALE, FEMALE) have been added to the data matrix, they will automatically be displayed next to the raw values (e.g., 1, 2). in the table.

Summary statistics are listed in a column at the bottom of each table.

To obtain the results for the variable named RACE, you may either click the right arrow at the bottom of the window, hit the right arrow on the keyboard, or select RACE from the Table menu near the top of the output window.



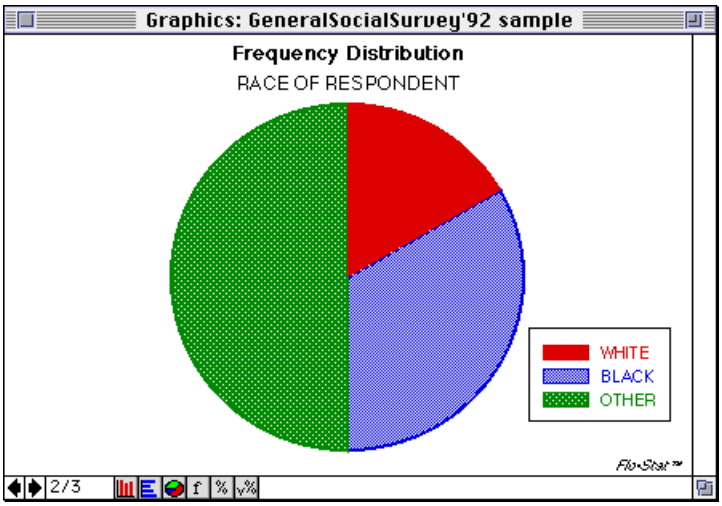
Value Label	Value	Frequency	Percent	Valid Percent	Cum Percent
MALE	1	96	48.0	48.0	48.0
FEMALE	2	104	52.0	52.0	100.0
	Total	200	100.0	100.0	
Valid cases	200				
Missing cases	0				
Sum	304.000				
Mean	1.520				
Median	1.500				
Mode	2.000				
Std. Dev.	0.501				
Variance	0.251				
Range	1.000				

Before reading on, see if you can complete the steps necessary to answer the questions posed earlier: “How many people in the sample are under the age of 20?”, and “What is the average age of those in the sample?”

Click the graph icon at the bottom of the tabular output window to generate a pie chart or graph of each variable's distributions.

Graph objects (e.g., legend, labels, graph) can be moved within the window by clicking and dragging. Characteristics (e.g., color, pattern, line width, font) of each object can be changed by double clicking the object and specifying the changes in the dialog window which appears following the double click.

The fastest method for transferring individual graphs from *Flo•Stat* to your word processor, page layout or drawing program is a simple **Copy** and **Paste**. Like other output windows in *Flo•Stat*, graphs do not need to be selected prior to being copied to the clipboard.



Creating and transforming variables

The ability to create new variables or transform those present in the data set, is an essential requirement in almost any study. *Flo•Stat*'s transformation window provides a powerful mathematical editor for use in both creating and transforming variables.

Click the transformation icon on the menu bar or select **Transformations...** from the **Utilities** menu at the top of your screen.

The transformation window contains four pull down menus labeled **Variables**, **Operators**, **Functions** and **Examples**. Selecting an item from one of these four menus places that item in the **Computes, Recodes & If-Then text editing portion of the** window.

The number of places to the right of the decimal point in values computed for new variables is set using the counter marked, **Decimals**.

When an equation, recode or if-then statement is ready for processing, the statement or statements must first be selected and the **Compute** button clicked. An error message appears at the bottom of the window in the event of a syntax error.

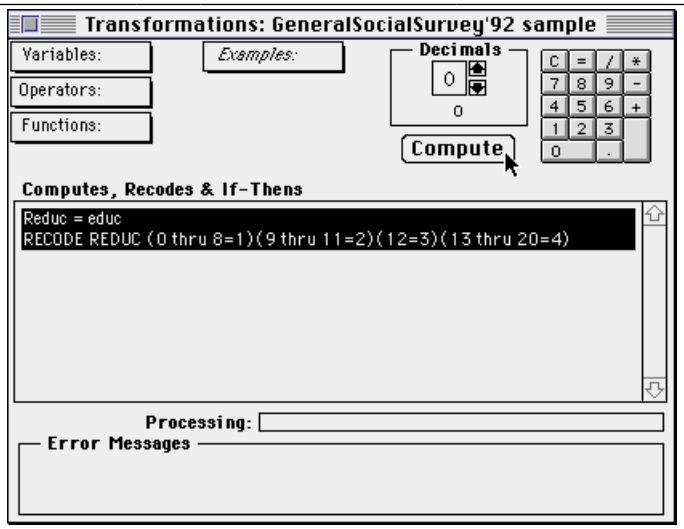
To copy the contents of the education variable in the data matrix, named EDUC, to a new variable named Reduc, type the equation (i.e., Reduc = EDUC) shown in the example below.

Variable names can consist of letters in either upper or lower case, contain numbers - as long as a number is not used as the first character in the name, cannot exceed eight characters in length, nor consist of most special characters, such as, ! @ # $ % ^ & * ( ) +. A warning dialog appears whenever a new variable name violates these restrictions.

Recode Reduc into the following five categories: 0 to 8 years of schooling = 1, 9 to 11 years = 2, 12 years = 3, 13 to 15 = 4, and 16 to 20 = 5. Enter the recode exactly as shown below.

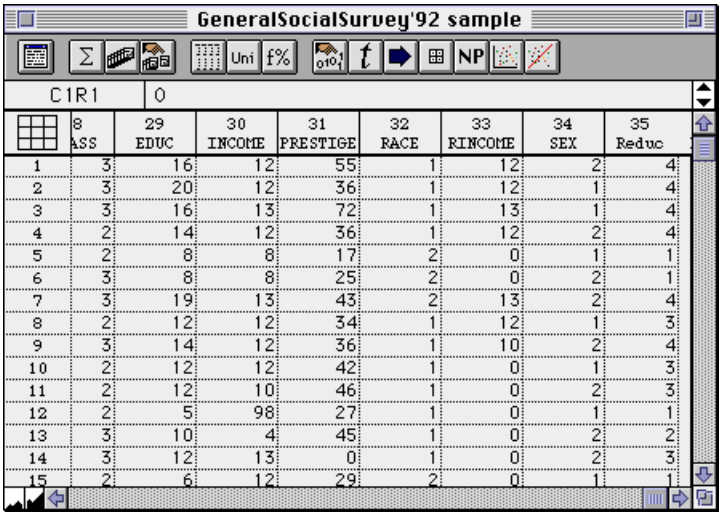
Select both the **Compute** and **Recode** statements, as shown below.

Click the **Compute** button.



Flo•Stat first calculates the new variable (i.e., Reduc) and adds the results, case by case, to the data matrix. Immediately following this step the new values are recoded to the five categories specified in the **Recode** statement.

New variables are automatically added as the last column in the data matrix. The results of the **Compute** and Recode statements can be seen by scrolling to the far right of the matrix.



	8 ASS	29 EDUC	30 INCOME	31 PRESTIGE	32 RACE	33 RINCOME	34 SEX	35 Reduc
1	3	16	12	55	1	12	2	4
2	3	20	12	36	1	12	1	4
3	3	16	13	72	1	13	1	4
4	2	14	12	36	1	12	2	4
5	2	8	8	17	2	0	1	1
6	3	8	8	25	2	0	2	1
7	3	19	13	43	2	13	2	4
8	2	12	12	34	1	12	1	3
9	3	14	12	36	1	10	2	4
10	2	12	12	42	1	0	1	3
11	2	12	10	46	1	0	2	3
12	2	5	98	27	1	0	1	1
13	3	10	4	45	1	0	2	2
14	3	12	13	0	1	0	2	3
15	2	6	12	29	2	0	1	1

Labeling variables

In most studies it is difficult, if not impossible, for the analyst to remember the meaning behind variable names, variable values (for example, does the number "1" in the **GENDER** variable represent male or female?) , not to mention what procedures were followed in creating certain variables. *Flo•Stat* helps to reduce this problem by offering an effective and direct method for labeling variables, their values, and providing an easy way to add comments.

Double click in the column containing the **Reduc** data values. The double clicking action opens the Variable Info window. The Variable Info window is used to enter variable names, value labels, missing values and comments. This window can also be opened by selecting the **Var Names, Labels...** item from the **Utilities** menu at the top of your screen.

Arrayed across the top of the Variable Info window are 1) a sliding bar used to move among the variables, 2) a pull down variable menu, and 3) a field to locate text strings.

Variable type - Numeric or Character - is set using the pop up menu.

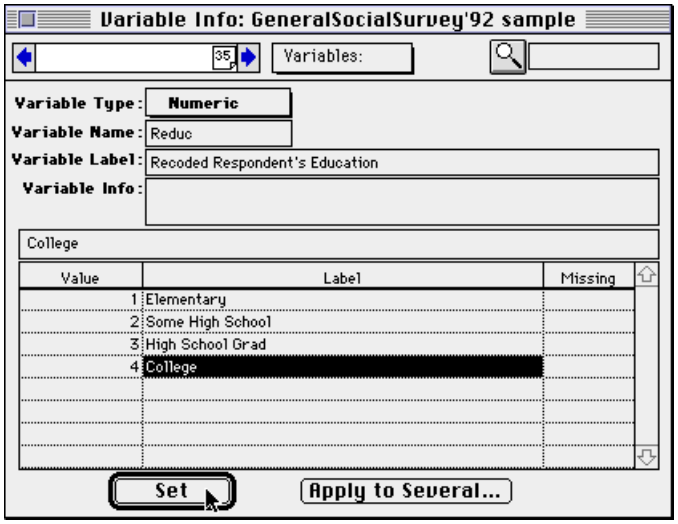
Variable Name, Label and Info are typed in to their respective fields.

Values, value labels and missing values are entered in the spreadsheet portion of the window. (Missing values are set by clicking in the last cell.)

New information is saved when the **Set** button is clicked or the scroll bar is used to move to another variable. If the go-away box at the top left of the window is clicked before either the **Set** button or the scroll bar are used, any new labeling information added to the variable will be discarded.

Enter the variable label and value labels for the newly created and transformed **Reduc** variable. Check the number in the scroll bar icon at the top of the window to see if it matches the column number containing **Reduc**. The number 35 should appear in the icon, and the name **Reduc** in the V**ariable Name** field.

Click the Set button to store Reduc's new label information.



The Quick Start tutorial you just completed touched on the basic information needed to begin using *Flo•Stat*. The package's ease of use, reliance on the Macintosh interface, and straightforward method of statistical analysis permits many users to begin taking advantage of *Flo•Stat*'s data management and computing power almost immediately.

Explore the remaining chapters of the manual to learn the details behind the more advanced features of *Flo•Stat*. For now, however, you have enough information to begin entering and analyzing your own data.

Contacting Senecio Software

All of us at Senecio Software hope you enjoy using this new statistical tool as much as we do and that you come to rely on *Flo* for many of your data manipulation and analysis needs. Of course, if you want to give us some tips on how we can improve *Flo•Stat* or have a question, send us e-mail at info@senecio.com. You can also write us at our corporate offices. We always look forward to hearing from our users.