

VISUAL STATISTICS

2.0

David P. Doane
Kieran Mathieson
Ronald L. Tracy
Oakland University



Boston Burr Ridge, IL Dubuque, IA Madison, WI New York San Francisco St. Louis
Bangkok Bogotá Caracas Lisbon London Madrid
Mexico City Milan New Delhi Seoul Singapore Sydney Taipei Toronto



VISUAL STATISTICS 2.0

Published by McGraw-Hill/Irwin, an imprint of The McGraw-Hill Companies, Inc. 1221 Avenue of the Americas, New York, NY, 10020. Copyright © 2001, by The McGraw-Hill Companies, Inc. All rights reserved. No part of this publication may be reproduced or distributed in any form or by any means, or stored in a data base or retrieval system, without the prior written consent of The McGraw-Hill Companies, Inc., including, but not limited to, in any network or other electronic storage or transmission, or broadcast for distance learning.

Some ancillaries, including electronic and print components, may not be available to customers outside the United States.

This book is printed on acid-free paper.

1 2 3 4 5 6 7 8 9 0 QPD/QPD 0 9 8 7 6 5 4 3 2 1 0

ISBN 0-07-240014-5 (book)

ISBN 0-07-240012-9 (CD)

ISBN 0-07-240094-3 (book with CD)

Vice president/Editor-in-chief: *Michael W. Junier*

Publisher: *Jeffrey J. Shelstad*

Executive editor: *Richard T. Hercher, Jr.*

Senior developmental editor: *Wanda J. Zeman*

Senior marketing manager: *Zina Craft*

Senior project manager: *Susan Trentacosti*

Lead production supervisor: *Heather D. Burbridge*

Senior designer: *Jennifer McQueen Hollingsworth*

Supplement coordinator: *Mark Sienicki*

Media technology producer: *Ed Przyzycki*

Cover design: *Anthony A. O'Neill*

Printer: *Quebecor Printing Book Group/Dubuque*

Library of Congress Card Number: 00-1-7930

www.mhhe.com

To

*Blythe,
Cullyn and Teagan,
Kathy*

Preface

What's New in the Revised Edition

In this new edition of *Visual Statistics* we have made several major enhancements and many minor improvements. Here are a few of the new features.

- Six new modules:
 - ◆ Random processes
 - ◆ Goodness-of-fit tests
 - ◆ Multiple regression analysis
 - ◆ Regression assumptions
 - ◆ Regression models
 - ◆ Binary predictors in regression
- Full integration of the worktext and software:
 - ◆ Adobe™ on-line worktext now has color and graphics
 - ◆ Jump to any desired part of the worktext chapter using bookmarks
 - ◆ Switch easily between worktext and software module
 - ◆ Worktext can be magnified, copied, or printed
 - ◆ View learning exercises and their solutions simultaneously
- Attractive one-screen main menu that lets you:
 - ◆ Select a chapter by clicking it
 - ◆ View the worktext for that chapter
 - ◆ Run the software module for that chapter
 - ◆ Run several modules simultaneously
- Hundreds of new and updated databases and examples
 - ◆ Revised examples
 - ◆ Updated databases
 - ◆ New multivariate databases
- Major changes in appearance and functionality in modules on:
 - ◆ Central Limit Theorem
 - ◆ Power and type I/II error
 - ◆ Trends and seasonality
 - ◆ Statistical process control
- Worktext enhancements
 - ◆ 2 levels of orientation (basic features, additional features)
 - ◆ 3 levels of exercises (basic, intermediate, advanced)
- Module enhancements:
 - ◆ Ability to copy/paste graphs as device-independent bitmaps
 - ◆ Improved use of color



Background

This textbook began near ten years ago when Tracy walked into Doane's office to see if he would be interested in writing a National Science Foundation (NSF) grant proposal to develop software to teach statistical concepts. The collaboration was natural. Each of us had taught statistics for many years, and had experience developing computer software. As our ideas began to take shape, we agreed that it was time to bring the computer into the *classroom* — not just have students do computer lab projects. We wanted to use computers to teach *concepts* rather than just analyze

data. We wanted to stress visual displays rather than numbers, equations, and calculations. An equation is an explanation to some students, but it is a barrier to others. We felt that a series of pictures or a simple animation could illustrate fairly complex equations. Our proposal asked the National Science Foundation's support to develop software and learning tools that would bring the power of computers into the statistics classroom. After we were awarded the NSF grant, we asked Mathieson if he would serve as a Windows consultant to our project. Soon it was clear that Mathieson's skills were essential, and we became full partners.

Visual Learning

We assume that (1) for many students visualization is a key to learning; (2) most statistical packages are designed for data analysis, not visualizing concepts; (3) individual instructors lack the time and resources to design their own software. Accordingly, we set out to create learning modules to illustrate concepts that cannot easily be shown mathematically or with packages such as Minitab or Excel. We wanted to minimize duplication of software capabilities already available. We wanted to promote learning through active self-discovery (inductive learning) as well as to support the instructor in the traditional lecture/lab setting (deductive learning).

We came up with the idea of a worktext that is closely coupled with software devised to lead a learner through experiments that promote self-discovery. We designed learning projects for teams and individuals that pose less-structured problems to continue this process. Our software was designed to encourage experimentation, so a learner can use simulation to understand the “experimental” side of statistics. These same features can be used effectively by the instructor in the classroom or the computer laboratory to illustrate a concept (e.g., the relationship between a confidence interval and a test statistic) and reinforce it with a simulation experiment.

Overview of Features

Visual Statistics consists of 21 software modules and a worktext. Each chapter of the worktext begins with a list of the key concepts the chapter will cover and a short list of learning objectives. Each concept is reviewed and illustrated. A short (about 15-minute) orientation to each module is provided. *Past users have told us that this orientation is invaluable.* There is also an orientation to additional features that may be skipped if its topics are not required. Learning exercises (basic, intermediate, advanced) are provided for each module. These exercises guide self-discovery learning. There are open-ended learning exercises for individuals and teams. A self-evaluation quiz allows you to test your understanding of the material. Solutions and a guide to answering each question are provided. A glossary at the end of each chapter provides definitions of terms used in the chapter.

Visual Statistics is to be used with a textbook or other course materials. It complements these traditional sources—it does not replace them. However, you'll find topics in these sources easier to understand after you *see* them in *Visual Statistics*.

Every software module has its own Help system, similar to that in other Windows programs. It includes definitions, examples, equations, graphics, and hints. You can use the table of contents and hypertext capabilities to jump to topics you want to learn more about, or do keyword searches using the index. *You will get more out of the learning process if you use the Help system.*

Our software uses familiar Windows visual controls (command buttons, scroll bars, list boxes, option buttons, and so on). Input is mostly from a mouse rather than from the keyboard. This interface will seem quite natural to most people. If you have used other Windows programs, you already know how to use *Visual Statistics* controls.

Every *Visual Statistics* module opens with a virtual notebook (a familiar ring binder with pages you can turn) that organizes the options. Clickable tabs divide the notebook into different sections, including an introduction (general idea of what is in the module), list of concepts (main topics covered), and sources of more information (references to chapters or other textbooks). Other tabs that vary from module to module provide examples (illustrations using real data), scenarios (realistic context for simulations), databases (large number of related variables using real data), visual templates (icons that show a distribution or trend), do-it-yourself controls (control panels to manipulate parameters of an experiment), and a data editor (to enter data or paste data from a spreadsheet). You may return to the notebook while running the program (e.g., to change scenarios, pick a different example, or choose different options).

Level of Complexity

We have tried to write flexible software, so the instructor can omit topics or vary their order. No module requires knowledge of another module, except in a general way (e.g., simple regression precedes multiple regression). Because students learn in different ways and at different rates, each module is aimed at the average learner but with options for those who desire more analytical depth (or who are just adventuresome). For example, in some modules you can set parameters by manipulating do-it-yourself scroll bars (simple level), by typing parameter values (intermediate level), or by choosing a known distribution and its parameters (advanced level).

Acknowledgments

Although we take responsibility for all remaining errors, our project has benefited greatly from field testing, advice, and comments of faculty colleagues and their students at other universities. We thank these individuals for their dedication and interest in our project.

Revised Edition:

Sung K. Ahn—Washington State University
Nancy C. Aiello—Northern Virginia Community College
Priscilla Chaffe-Stengel—California State University- Fresno
Subha Chakraborti—University of Alabama
Patti Collings—Brigham Young University
Bradford R. Crain—Portland State University
William Duckworth II—Iowa State University
John C. Dutton—North Carolina State University
James Flynn—Cleveland State University
Robert Hannum—University of Denver
David C. Howell—University of Vermont
Glenn W. Milligan—Ohio State University
Pin Ng—University of Illinois
Roxy Peck—California Polytechnic State University- San Luis Obispo
Don R. Robinson—Illinois State University
Mark E. Rush—Washington & Lee University
Sue B. Schou—Idaho State University
Robert K. Smidt—California Polytechnic State University- San Luis Obispo
Michael P. Wegman—Keller Graduate School of Management

Previous Edition:

Richard Anderson-Sprecher—University of Wyoming
Mohamed Askalani—Mankato State University

David Booth—Kent State University
Tim Butler—Wayne State University
Patty Collings—Brigham Young University
Peyton Cook—University of Tulsa
Rick Edgeman—Colorado State University
Chris Franklin—University of Georgia
Bob Hordon—Rutgers University
William Jedlicka—William Rainey Harper
Anthony Keys—Wichita State University
Anne B. Koehler—Miami University of Ohio
Gerald Kohers—Sam Houston State University
Benny Lo—Northwest Polytechnic University
Gary Martin—DeVry University - Atlanta
Madhuri Mulekar—University of South Alabama
Sufi Nazem—University of Nebraska - Omaha
Ceyhun Osgur—Valparaiso University
Paul Paschke—Oregon State University
Andy Siegel—University of Washington
Boyd Swartz—Monmouth University
Mack Shelley—Iowa State University
Robert D. St Louis—Arizona State University
Bret Wagner—University of Delaware
Linda Young—University of Nebraska - Lincoln



This project would have never begun without the initial support of the National Science Foundation, Department of Undergraduate Education (DUE # 9554967 and DUE # 9254182). For their sustained belief in our concept, we are grateful. Any opinions expressed are those of the authors and not necessarily those of the Foundation.

However, without the support of McGraw-Hill/Irwin, the project might never have reached the full audience of learners that we hope will benefit from our approach. Our executive editor, Richard Hercher has been an advocate of our project ever since he saw early versions of our first four modules, and has not wavered in his support. As we struggled to meet deadlines, he encouraged and cajoled us, offered advice, consulted us, and kept us informed of important strategic issues. We are grateful to Wanda Zeman for her patient advice and assistance in working with us on a day-to-day basis in developing this new edition, and in dealing with the many problems that naturally arise. We thank the entire staff of McGraw-Hill/Irwin for their able assistance in managing our project through the editorial and production process. We also acknowledge database development, field testing, and programming assistance by Steve Losey, Judith Gurney, Karen Helber, and Don Smith.

**David P. Doane
Kieran Mathieson
Ronald L. Tracy**

License Agreement

Copyright (c) 1997-2001 by The McGraw-Hill Companies, All rights reserved.

Visual Statistics software and worktext are protected by both U.S. copyright laws and international copyright treaty provisions.

The McGraw-Hill Companies grants you a non-exclusive license to use the software in accordance with the following terms. *Visual Statistics* software and worktext are licensed material and remain the property of The McGraw-Hill Companies.

The McGraw-Hill Companies grants you, and only you, the right to use either the **COMPACT** or **COMPLETE** setup type on your work, home, or portable computer.

The McGraw-Hill Companies further grants you the right to transfer this license, the software, the worktext, and the CD to another party provided that the other party accepts all terms of this agreement; you destroy all copies of the software, both physical and electronic; and you uninstall the software from your computer(s).

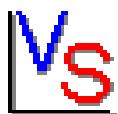
Failure to comply with any of these terms will terminate this agreement, your right to use the software, and may result in either criminal or civil penalties or both. You may choose to terminate this agreement at any time. Upon termination of this agreement, you must immediately destroy the software, all copies of the software, both physical and electronic, and uninstall the software from your computer(s).

If any *Visual Statistics* graph, table or other image is printed in a publication, a proper citation to *Visual Statistics*, the authors, and McGraw-Hill Companies must be given. The McGraw-Hill Companies and the authors are not responsible for misuse or inappropriate use of the product.

The media on which the software is furnished is warranted to be free of defects in workmanship and material under normal use for a period of ninety (90) days from the date of purchase. The sole responsibility of Irwin/McGraw-Hill will be to replace the defective media if it is returned, with proof of purchase date, to the publisher.

THE SOFTWARE IS LICENSED "AS IS" WITHOUT WARRANTY AS TO ITS PERFORMANCE. EXCEPT FOR THE MEDIA WARRANTY PROVIDED ABOVE, THERE ARE NO OTHER WARRANTIES EXPRESSED OR IMPLIED. IN NO EVENT SHALL THE MCGRAW-HILL COMPANIES, OR THE AUTHORS, BE RESPONSIBLE FOR ANY INDIRECT OR CONSEQUENTIAL DAMAGES OR LOST PROFITS, EVEN IF THE MCGRAW-HILL COMPANIES HAVE BEEN ADVISED OF SUCH DAMAGE. SOME STATES DO NOT ALLOW THE EXCLUSION OR LIMITATION OF IMPLIED WARRANTIES OR LIABILITY FOR INCIDENTAL OR CONSEQUENTIAL DAMAGES, SO THE ABOVE LIMITATION MAY NOT APPLY TO YOU.

Introduction



Using *Visual Statistics*

OBJECTIVES

- Understand what *Visual Statistics* is all about
- Learn how to install *Visual Statistics*
- Learn how to use the Notebook, the starting point for every *Visual Statistics* module
- Learn about the worktext

What is *Visual Statistics*?

Statistics is a challenging subject for most people. Most of us have to work hard to understand probability distributions, statistical power, and confidence intervals. However, statistics is becoming more important. It's hard to design a product, create an advertising campaign, or play poker well without understanding statistics.

The main problem is that statistical ideas are abstract. When did you last speak with a probability distribution? Have you ever seen a correlation on the shelf at Sears? Or been almost run down by a wild confidence interval (except on an exam)? Formulas and numbers are fine for some people, but most of us learn best when we can see and interact with the thing we're learning about.

That's where *Visual Statistics* comes in. Computers have become more powerful and easier to use, so it's possible to explore statistical ideas in new ways. Graphics and animation can bring concepts to life in ways that formulas, tables, and static diagrams cannot. Rather than imagining what data for a particular situation looks like, *Visual Statistics* lets you create the situation and see it for yourself. You can change the circumstances and watch the result. *Visual Statistics* helps you build your statistical intuition, making it easier to understand the ideas behind data analysis.

It's important to understand that while you can analyze data with some *Visual Statistics* modules, that is not their main goal. Instead, *Visual Statistics* will help you *learn statistics*. Even the most powerful analysis tools, like Minitab, SAS, and SPSS, won't help you if you don't know what to do with them. On the other hand, if you are familiar with statistical ideas, you can do wonders with a simple calculator. The tools you use are less important than how you wield them. *Visual Statistics* will help you use statistical techniques more effectively because you will understand what they can be used for and what their limitations are.

Visual Statistics is not a tutorial, however. For example, the regression module won't explain why regression works or how to calculate regression statistics. Instead, it helps you explore regression, so you understand it better than you would without the visual tools. You'll get more value from your statistics text since you'll be able to experiment with the concepts. It'll be easier for you to *understand* statistics — and understanding brings more confidence and less worry.

The *Visual Statistics* suite consists of 21 modules. Each one helps you explore a specific set of statistical ideas. For example, the ANOVA module will help you understand one-way analysis of variance. It is designed for that purpose alone. It's different from the module that helps you understand statistical process control, since the two topics are different. However, every module shares some characteristics. First, they are all easy to use. Run through the 15-minute orientation for each module, and you should be able to use it effectively. If you have trouble, select **Help** for assistance. Second, they all use the Notebook. The Notebook introduces the module, and lets you choose how you want to explore the relevant statistical ideas. Different modules have different Notebook options. Some let you examine scenarios, while others allow you to analyze real data sets or enter your own data. The Notebook is described in more detail later in this introduction.

The best way to learn what *Visual Statistics* is all about is to try it. So go ahead. Install *Visual Statistics*, start the Univariate Data Analysis module and play around. You'll see how the same data can look very different, depending on how it is displayed.

Installing *Visual Statistics*

Visual Statistics runs under recent versions of Windows (95, 98, 2000, NT). You need at least a Pentium PC with 32 MB of memory. *Visual Statistics* takes about 7 MB of hard disk space for a Compact installation or about 37MB of hard disk space for a Complete installation. This includes the software (21 modules), the worktext (21 chapters plus the introduction), worktext solutions, help files, and databases containing over 1000 variables. The Compact installation runs off of the *Visual Statistics* CD while the Complete installation runs off your computer and only checks to see that you have a *Visual Statistics* CD.

Installing VS

This version of *Visual Statistics* is for a single user. Users needing a network version should contact their McGraw-Hill representative. To install *Visual Statistics*, first close down any other programs you are running. Place the *Visual Statistics* CD in your CD drive. The install program is **setup.exe** in the Setup folder in the Visual Statistics 2.0 folder. It can be executed in one of three ways:

- Use the Add/Remove Programs dialog box in the Windows Control Panel, *or*
- Double click the **setup.exe** file in Windows Explorer, *or*
- Press the **Start** button, select Run, and type d:\visual statistics 2.0\setup\setup.exe (where d: represents the drive letter for your CD reader).

You will be asked to select one of two setup types. *Both require that your Visual Statistics CD be in your computer.*

COMPACT installs the front-end of *Visual Statistics* on your computer and the necessary system files. It takes about 50 Mb LESS room on your hard drive.

COMPLETE installs the entire *Visual Statistics* system on your computer. It takes about 50 Mb MORE room on your hard drive, but runs faster on computers with slower CD readers.

Read the License Agreement during the install. After installing *Visual Statistics* it can be read by selecting the *Visual Statistics 2.0* folder in the *Programs* folder. *Visual Statistics* can be started from this folder or by double clicking the VS icon that will be installed on your desktop. Once the system files have been installed on your computer, *Visual Statistics* can also be run directly from the CD by opening the *VisualStatistics.exe* file on the CD. This can be useful if your VS files have been corrupted and you want to run *Visual Statistics* without reinstalling the software.

Uninstalling VS

Visual Statistics can be uninstalled in two ways:

- Use the Add/Remove Programs dialog box in the Windows Control Panel, *or*
- Press the **Start** button and select **Uninstall Visual Statistics 2.0** from the *Visual Statistics 2.0* folder in the *Programs* folder.

Installing Adobe Reader 4.0

When you run the *Visual Statistics* install program, it will detect if you have already installed an Adobe Reader. If you do *not* have an Adobe Reader, VS will install the Adobe 4.0 Reader for you. If you *do* have an Adobe Reader, VS will *not* install the Adobe 4.0 Reader. In this case, you *should* determine which version of the Adobe Reader is on your computer. Although *Visual Statistics* will

work with versions earlier than 4.0, the presentation of graphics improved considerably with version 4.0. If you have an earlier version, it is strongly recommended that you uninstall it and install the 4.0 version by executing the ar405en.exe file in the Reader folder on the *Visual Statistics* CD or by getting the latest version from Adobe at www.Adobe.com. Adobe warns that earlier versions should be uninstalled *before* installing version 4.0. If the earlier version is uninstalled *after* the 4.0 version is installed, problems will develop with the 4.0 Reader according to Adobe Corporation.

Starting *Visual Statistics*

Main Menu

To start *Visual Statistics*, click its icon on your desktop, or select *Visual Statistics 2.0* in the **Programs** folder in the **Start** menu. After opening the cover, you will see the a menu like the one shown in Figure 1. From this menu, you can do three things: (1) view a chapter in the worktext (the **Show Worktext** button); (2) run a software module (the **Run Module** button); (3) exit *Visual Statistics* (the **Exit** button). Subject to your system's memory, you may simultaneously use as many chapters or modules as you wish (they will appear on the taskbar at the bottom of your screen). When a chapter is selected, its title will turn white and its icon will turn yellow. In Figure 1, for example, **Chapter 1** (*Visualizing Univariate Data Analysis*) has been selected.

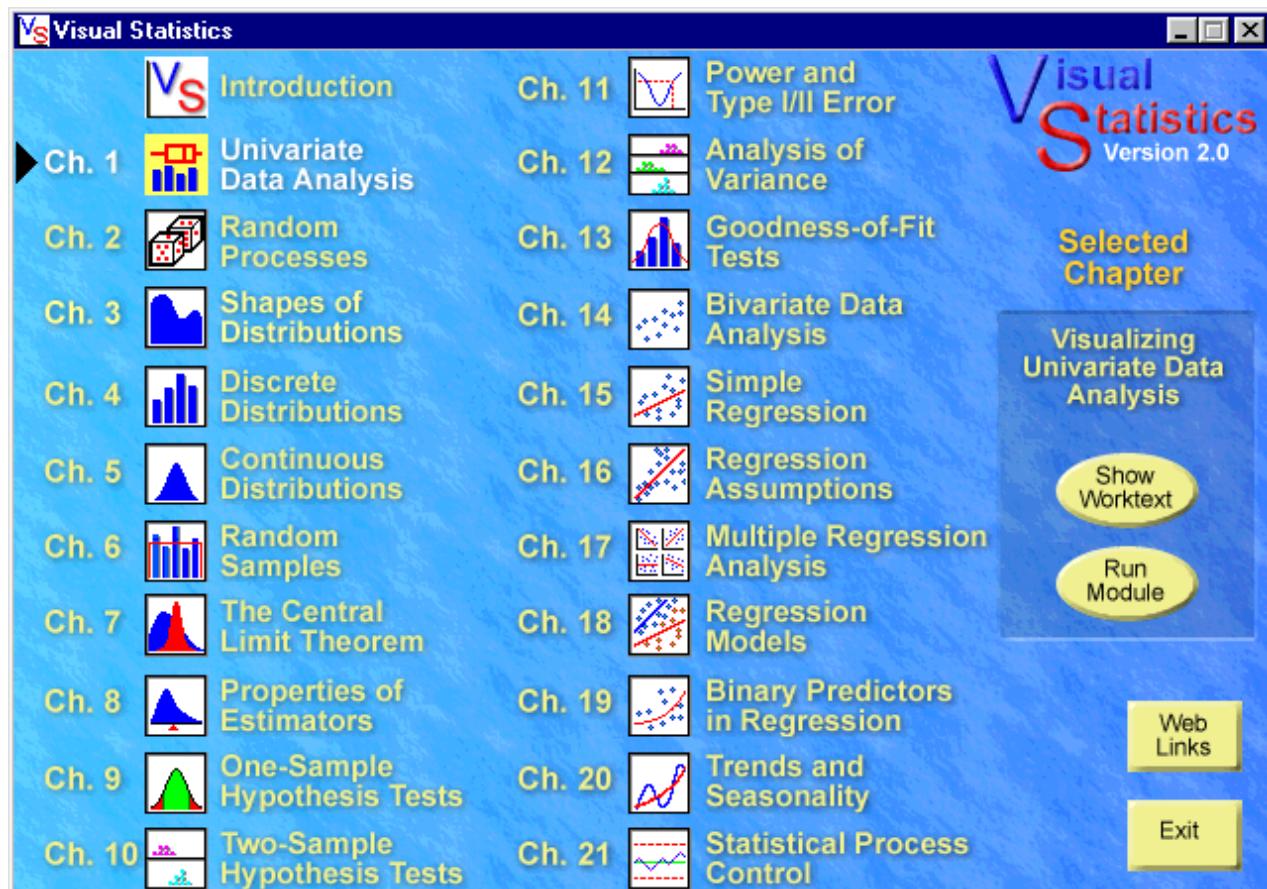


Figure 1: Main Menu

Selecting a Chapter

To select a chapter, click on its chapter number, its icon, or its title. If a chapter has not been installed, its title will be “grayed out” to indicate that it is inactive. When you click the chapter, a comet will streak across the screen to the Selected Chapter panel on the right, and the **Show Worktext** and **Run Module** buttons will appear in the panel. If either the module or worktext was not installed, the button will be “grayed out” and will be inactive. Each software module corresponds to a chapter in the worktext. The chapter's learning exercises will require that you run the corresponding software module. You can run several software modules at once, subject to your system's memory.

The **Visual Statistics** Worktext

The worktext has one chapter for each module. Each chapter in the worktext contains:

- A list of concepts and learning objectives
- An overview of concepts and illustrations of concepts
- An orientation to software features
- Structured learning exercises (basic, intermediate, and advanced)
- Suggested independent projects (team, individual)
- A bank of self-test questions
- A glossary of terms
- Answers to the self-test questions

Each chapter starts with a list of relevant concepts and learning objectives. Next, it offers an overview and illustration of each concept. Then there is a short section entitled Orientation to Basic Features. *Make sure you read this section.* It will quickly show you how to use the most important elements of the module. It only takes a few minutes and is well worth it. The Orientation to Additional Features illustrates other important capabilities of the software, which will be necessary to complete some of the exercises. These features are not necessarily more complex, but they intended to suggest priorities that may assist a busy learner or that might be recommended by an instructor if time permits.

The Learning Exercises ask questions you can answer using the module. They are to be completed in front of a PC. You will have a better understanding of the statistical ideas covered in the chapter once you run through the exercises. The Basic, Intermediate, and Advanced exercises are arranged in gradually increasing level of challenge. Your instructor might ask you to turn in some of the exercises. However, you can grade your own work using the Solutions.

The worktext also includes Learning Projects for individuals and teams. Where the Learning Exercises guide you through the module, the Learning Projects leave you to figure out how to use the module and do the analyses yourself or with your team. You may be asked to do a written report or make an oral presentation based on the Learning Projects. It is a good idea to do the Learning Exercises before you try the Learning Projects.

The Self-Evaluation Quiz lets you test your understanding of the ideas covered in the chapter. Try it after you have completed the Learning Exercises. The answers are at the back of each chapter, after the chapter's Glossary. You should do the entire quiz first, before you look at the answers. If you answer a question incorrectly, there are references to relevant sections of *Visual Statistics* that will help you understand the right answer. The exercises are arranged somewhat in increasing order of challenge. Some of the exercises can be answered by looking at the Overview, Illustration, and Glossary, while others may require using the software or working certain exercises.

The Glossary defines the terms used in the chapter. Much of the Glossary information (plus a lot more, such as formulas) will also be included in the module's Help file. Click **Help** on the menu bar to look at the Help file. It contains definitions, formulas, examples and explanations as well as usage instructions. The Help files are an oft-overlooked resource.

Using the *Visual Statistics* Worktext Reader

To view a worktext chapter, click the **Show Worktext** button (if the chapter has not been installed, the button will be “grayed out” to indicate that it is inactive) to bring up the Adobe reader, illustrated in Figure 2. Each chapter opens to its title page. You may jump to any section of the chapter by clicking the appropriate title in the Bookmarks window of the Adobe reader (on the left side of the screen) or you may use the scroll bar on the right. The **Run Module** button takes you to the chapter’s software module. The **Show Solutions** button will display the answers to the learning exercises in a separate window at the bottom of the screen. The **Return to VS** button displays the *Visual Statistics* main menu.

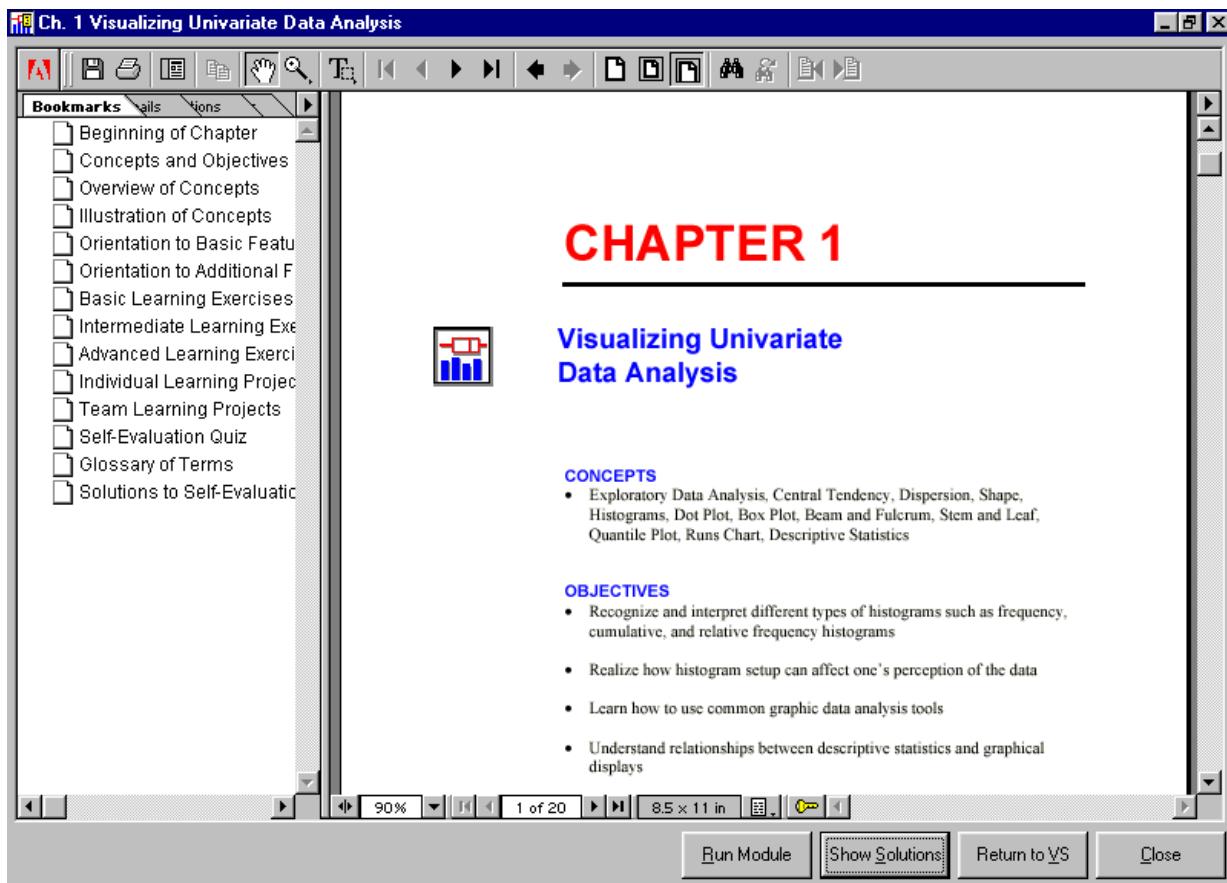


Figure 2: Adobe Reader

Adobe Buttons

The buttons across the top line of the Adobe reader offer many useful features. Each button has a tooltip that will appear when you allow the cursor to rest on the button.

The first group of buttons is:



- save file button (save a copy of the file)
- print file button (to print all or selected text)
- show/hide navigation panel (the “bookmark” area on the left of the screen)
- “hand” tool (to drag the document around in the window when it is too large to fit)
- magnify (to enlarge the document in the window; each click is one size larger)
- select text (so it can be copied using Ctrl-C)

The next group of buttons is:



- 4 page navigation buttons (first page, previous page, next page, last page)
- 2 view buttons (previous view, next view)
- 3 document size buttons (actual size, fit in window, fit width)

The last group of buttons is:



- rotate text (90° left or 90° right)
- find (a word or phrase of interest)
- jump to highlights (next, previous sections)

At the bottom of the Adobe screen are two other useful controls:



The first control permits you to set the size of the displayed document (or type in the desired size). The second control lets you page forward or backward in the document (or go to a given page by typing it).

Printing in Adobe

Each chapter of the worktext is in color, but it can be printed either in color or black-and-white. This depends on the printer you are using. The quality of printing of graphics depends on your printer, but also on the version of Adobe you are using (we strongly recommend version 4.0 or later).

Exercise Solutions

If you press the **Show Solutions** button you will see solutions to the learning exercises in a separate window at the bottom of your screen, as illustrated in Figure 3. This will allow you to check your answers to the learning exercises. If you are taking a class, the instructor may ask you to submit your original (uncorrected) answers along with your handwritten corrections and comments indicating the mistakes that you made or any explanations that you added after reading the exercise solutions.

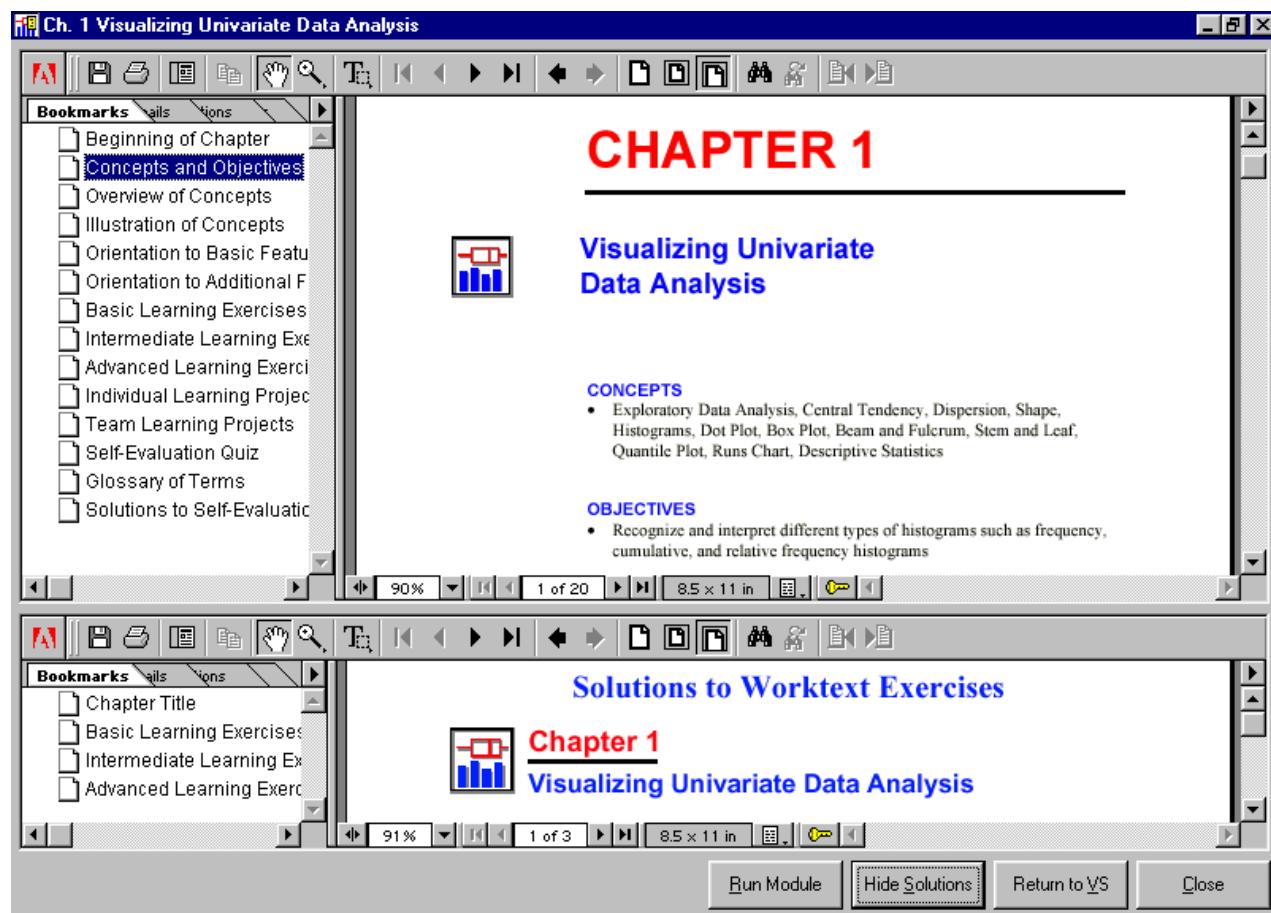


Figure 3: Adobe Reader with Solutions

The *Visual Statistics* Notebook

Every *Visual Statistics* module starts with the Notebook. There is a different Notebook for each module, but they all work in the same way. The main purpose of the Notebook is to let you choose the type of data you want to look at. Figure 4 shows the opening page of the Notebook for the Bivariate Data Analysis module.

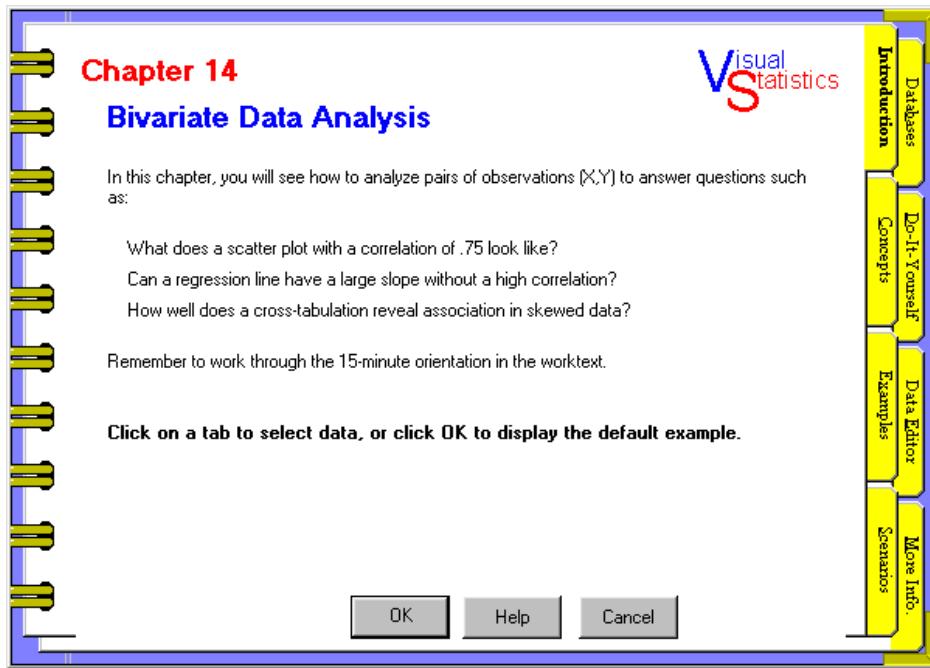


Figure 4: Notebook Opening Page

Navigating the Notebook

The Notebook is divided into sections, each with its own yellow tab. The Introduction section identifies the module and shows some questions the module will help you answer. The Concepts section lists some ideas the chapter covers. The More Information section identifies related modules, and lists relevant chapters in McGraw-Hill/Irwin textbooks.

The **OK** button closes the Notebook and returns to the module. It's only available when you have selected some data for the module to use. If you click the **OK** button on the Introduction page when the module first starts, default data will be chosen for you. The **Help** button gives you assistance in using the Notebook. The **Cancel** button closes the Notebook immediately.

Data Selection: Examples

Examples are real data sets that have been selected to illustrate the concepts of the module. For instance, to select an example in the Bivariate Data Analysis module, click on the **Examples** tab. You would see the display shown in Figure 5. Notice there are several different example categories (Sports, Consumer, etc.). You can jump directly to the category's page by clicking on its name. The names are in green and underlined, showing they are hyperlinks. Just as on the Web or in a Windows Help file, clicking on a hyperlink jumps to a different page. Alternatively, you can turn the page by clicking on **Next page** (lower right corner in Figure 5).

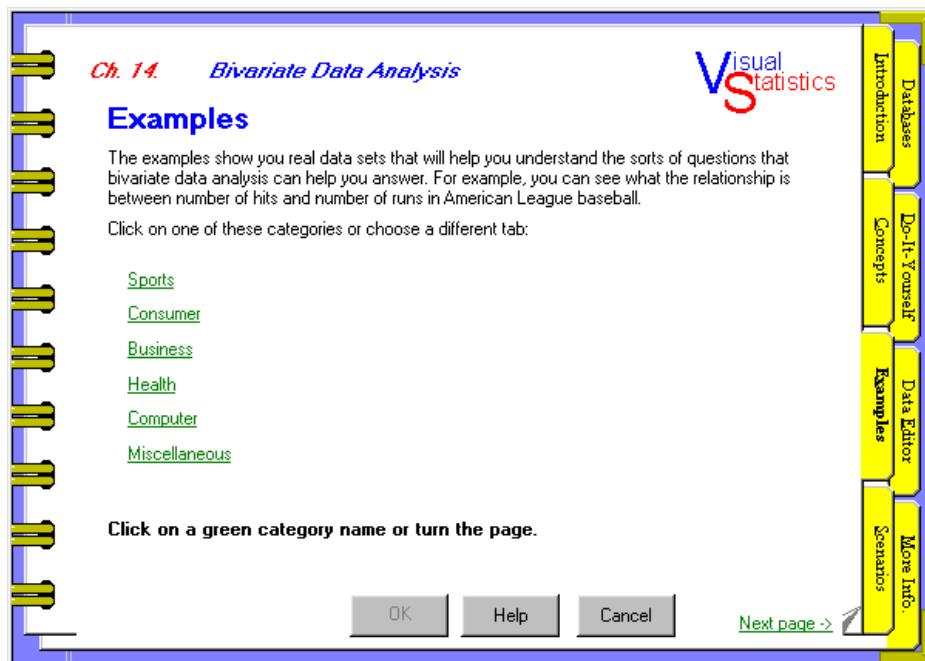


Figure 5: Categories of Examples

For instance, Figure 6 shows various sports-related examples. Click on an example's name to select it. You can read a description of the selected example in the yellow area to the right of the example list. Each page in the section gives examples from a different category. Turn the pages in a section by clicking on **Previous page** (upper right corner) or **Next page**. (lower right corner). You can also click on the “folded corners” of the page, as seen in the upper and lower right.

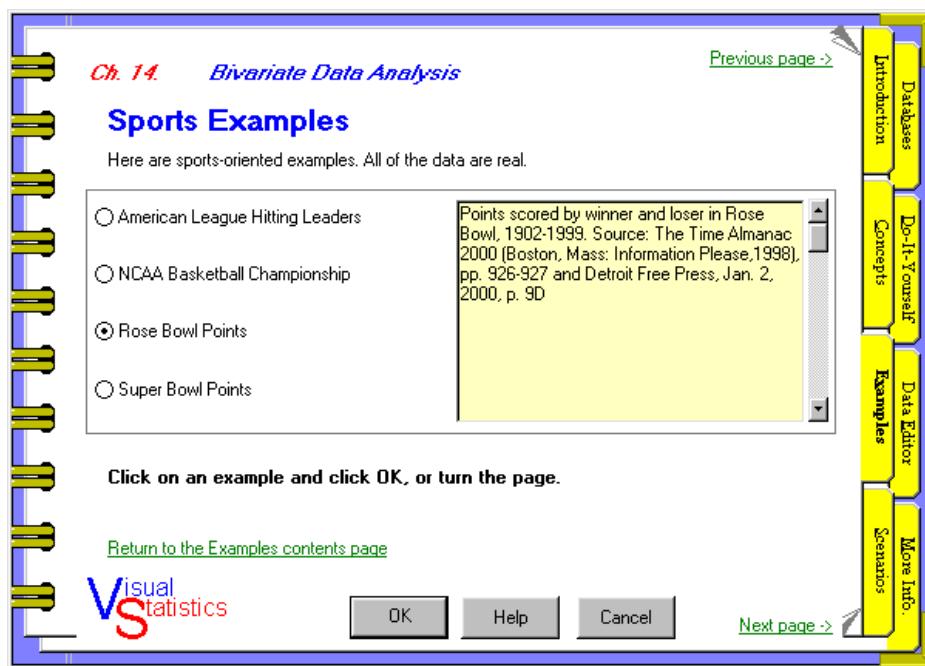


Figure 6: Examples in the Sports Category

Data Selection: Databases

A database contains many variables. You select the one(s) that you wish to analyze. The appearance of the database selection screen depends on the module. Figure 7 shows a database page from the Bivariate Data Analysis module. Note the categories Demographics, Economics, Crime, and so on. Click on the + and – signs to expand or collapse the category. The yellow area to the right describes the selected variable.

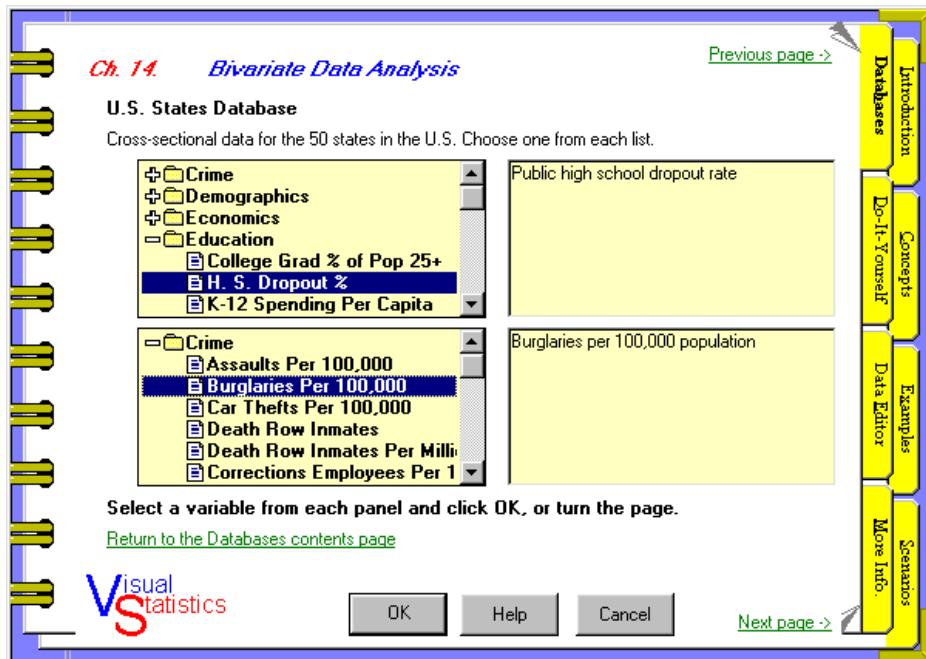


Figure 7: Notebook Database Page

Several modules use another type of variable selection screen. Figure 8 illustrates the one for Multiple Regression Analysis, which is similar to the others.

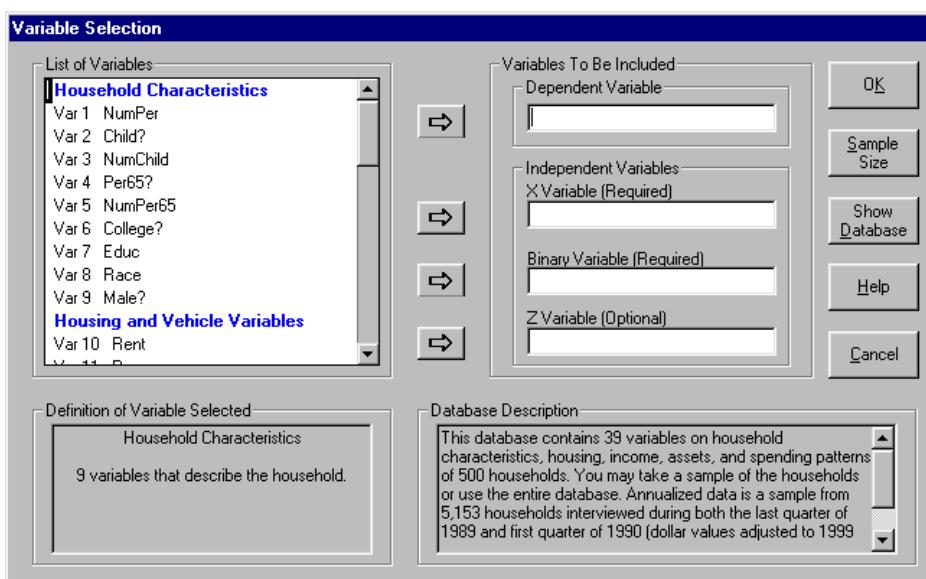


Figure 8: Variable Selection Database Screen

Data Selection: Using the Data Editor

The Data Editor lets you create your own data sets. Figure 9 shows an example from the Bivariate Data Analysis module. The spreadsheet is similar to Excel or 1-2-3. Each row is one observation (a person, a team, etc.). The first column is a label for the observation (a person's name, a team's name, etc.). The first column is optional; you don't have to label every point. The other columns contain the data that you want to analyze. Use the mouse or arrow keys to select a cell, then type data into it. You can rename a column by typing into the cell above the column (e.g., the cell with "Exam 2" in it).

The Data Editor has other useful options. Use the **File** menu to save your data to disk and retrieve it later. If you're running **Visual Statistics** on a network, make sure you save your data to a disk you can write to (e.g., a floppy). Use the **Edit** menu to modify the data or move data between applications (**Copy**, **Paste**, **Insert Row**, or **Delete Rows**). For example, you can copy data from the Data Editor and paste it into other Windows applications such as Excel, 1-2-3, Word, or WordPerfect. These operations are explained in Help in the Data Editor. You can also do the reverse—that is, paste data from another program into the Data Editor. Use the **Options** menu to sort the data, change its display format, and enter a title for the data set. Select **Search** under **Help** in the Data Editor for more information.

Different versions of the Data Editor are available in several programs (the number of columns and their headings will depend on the application). Figure 7 shows the data editor for the Bivariate Data Analysis Module. When you are finished editing your data, click **Close Editor and Use Data** (or click the "smiling face" icon). If you decide that you don't want to use the data, click **Close Editor and Discard Data** (or click the "sad face" icon).

The screenshot shows a Windows application window titled "Data Editor". The menu bar includes "File", "Edit", "Options", and "Help". Below the menu is a toolbar with icons for saving, opening, and other functions. A title bar says "Exam Score Comparison". The main area is a spreadsheet grid with the following data:

Obs	Student	Exam 1	Exam 2
1	Bob	66	75
2	Tom	38	55
3	Mary	50	78
4	Frieda	71	68
5	Thusita	95	99
6	Juanita	96	98
7	Larry	87	91
8	Fred	45	68
9	Gin Mei	75	84
10	Thelma	81	78
11			
12			
13			
14			
15			
16			
17			

Figure 9: Data Editor in Bivariate Data Analysis

Data Selection: Scenarios

The Examples, Databases, and Data Editor sections are similar in one respect: they let you examine real data sets. The last three notebook sections are different. They let you experiment with the *process* that generates data sets. For example, suppose you rolled a normal six-sided die twice and got a 1 and a 3. The “1 and 3” is the data set. The “rolled a normal six-sided die twice” is the process that created the data set. If you run the same process again (that is, roll the die twice), you would probably get a different data set. Roll the die twice more, and you get yet another data set. You can ask many questions about this situation. Would the data be the same each time? How different would it be? Suppose you added the numbers together. What is the highest sum you can get? What is the smallest? Are you more likely to get the smallest sum or the largest sum? Are all sums equally likely, or are some values more common than others?

You can learn a lot about statistics by examining these questions. But you can learn even more by varying the process that creates the data. For example, what if you rolled a 10-sided die, as used in *Dungeons and Dragons*? What if you spun a roulette wheel twice? Would the answers be the same?

This is where the Scenarios, Templates, and Do-It-Yourself sections come in. The Scenarios let you choose from processes that have been set up for you. Each scenario generates data that you might see in a particular situation. For instance, one scenario looks at the relationship between the number of times at bat and home runs in major league baseball. It gives you a familiar context for statistical exploration. You can take repeated samples, and see how random chance leads to very different data from identical situations. Figure 10 shows one of the scenario pages from the Bivariate Data Analysis module. Most of the modules have a Scenarios section. Some modules have a version control (from 1 to 99) that lets you repeat the same scenario (selecting 0 will make a random scenario selection).

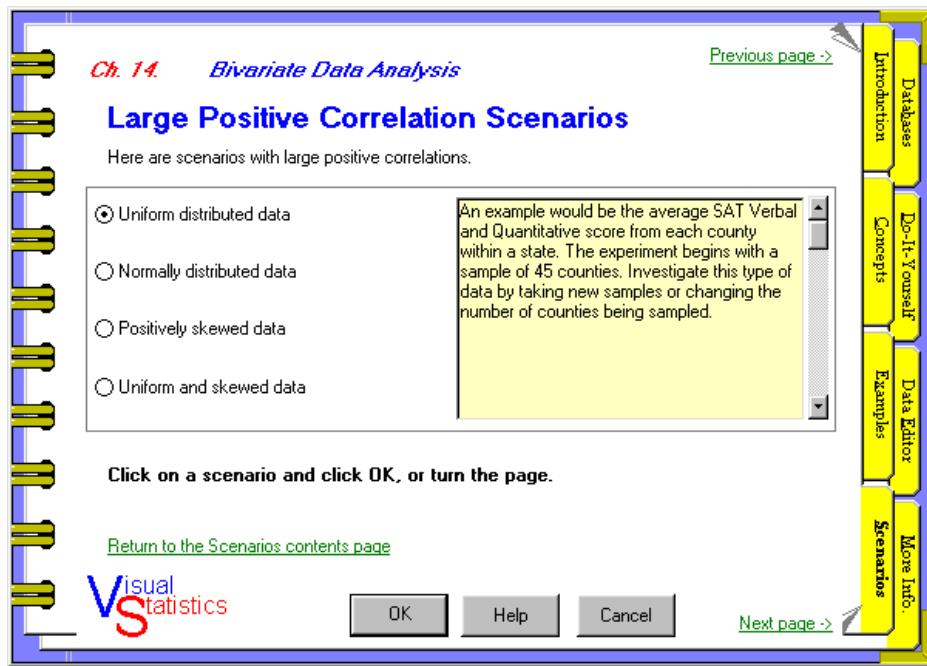


Figure 10: Scenarios

Data Selection: Templates

Templates let you generate data conforming to a particular shape. Several modules use templates. Figure 11 shows a template screen from the Properties of Estimators module. Each template

represents a probability distribution (you'll learn what that is in your statistics class). To select a distribution to sample, click on its picture and click the **OK** button.

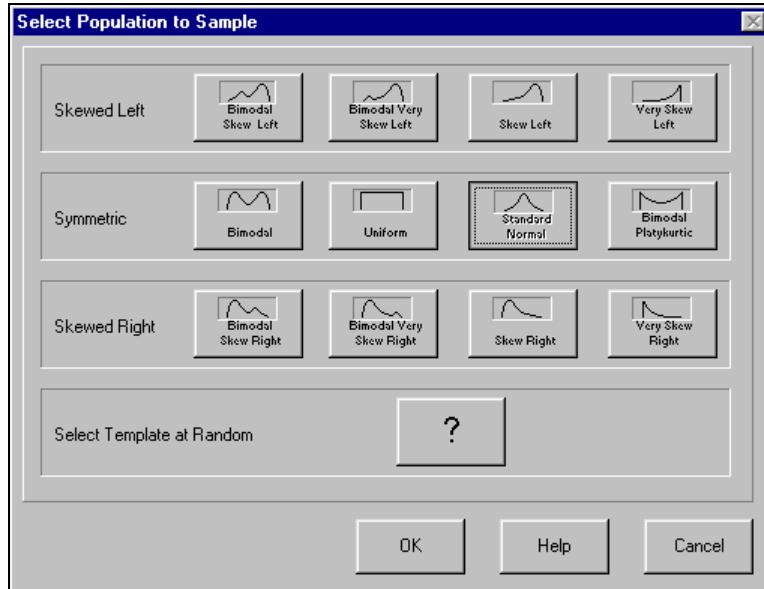


Figure 11: Templates

Data Selection: Do-It-Yourself

The Do-It-Yourself option gives you control over the process generating the data. Many of the modules have a do-it-yourself section. Figure 12 shows the do-it-yourself controls from the Simple Regression module. The appearance of the do-it-yourself screen depends on which module you are using. Figure 13 shows the do-it-yourself controls from the [Visualizing Shapes of Distributions](#) module.

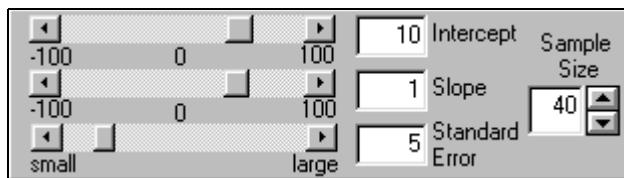


Figure 12: Do-It-Yourself Controls for Simple Regression

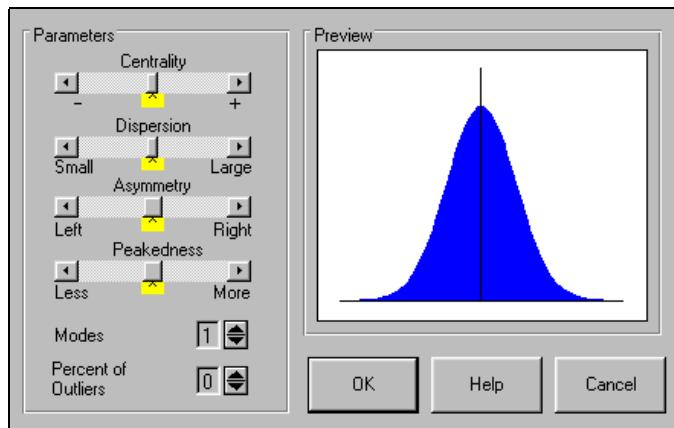


Figure 13: Do-It-Yourself Controls for Shapes of Distributions

Making Choices That Fit Your Needs

You can see that *Visual Statistics* is a powerful tool. There are quite a few modules, some of which have many options. There is something for everyone in *Visual Statistics*, from the novice to the expert. Beginners find it much easier to grasp basic statistical concepts when they use the package. The reason is simple: it's easier to understand something you can actually see than something you have to imagine. However, during our testing of *Visual Statistics*, we found that experts liked using the modules as well. The visualization and experimentation capabilities gave them new insights into common statistical problems.

Don't let the suite overwhelm you. Although there's a lot of power available, you don't need to use every option in every module. In fact, you might find yourself using only a few of the features. That's fine if it gives you the level of understanding you want. This worktext will help you get what you need from *Visual Statistics*. In particular, make sure you run through each module's orientation. It only takes a few minutes, and then you'll have a good idea of how to use the module.