

BATTERIES INCLUDED



TM

# تعمیرات

TM

MICHAEL H. REICHMANN & ROBERT WILSON

FOR  
APPLE™

*Scanned by cvxmelody*

*<http://www.cvxmelody.net/AppleUsersGroupSydneyAppleIIDiskCollection.htm>*

# B/GRAFH™

Apple™ Edition

Program design by Michael H. Reichmann and Robert Wilson

Manual written by Ian Chadwick  
Package design by Abraham Tanaka Associates Limited  
Package cover illustration by Ken Steacy  
Manual design by Abraham Tanaka Associates Limited

## NOTICE OF REGISTERED TRADEMARK

In this manual reference is made to Apple which is a registered trademark of Apple Computers, Inc.

Copyright © First Edition, November 1984 by Irata Press Ltd. All Rights Reserved

**BATTERIES INCLUDED™**



*"The Energized Software Company!"*

30 Mural St.  
Richmond Hill, Ontario  
L4B 1B5 Canada

Telex 06-218290  
(416) 596-1419  
MAY 1986  
PRINTED IN CANADA

17875 Sky Park North  
Suite P, Irvine, California  
USA 92714

## LIMITATIONS OF WARRANTY AND LIABILITY

BATTERIES INCLUDED, or any dealer or distributor distributing this product, makes NO WARRANTY, EXPRESS OR IMPLIED, with respect to this manual, the related floppy disk(s) and any other related items, their quality, performance, merchantability, or fitness for any particular use. It is solely the purchaser's responsibility to determine their suitability for any particular purpose.

BATTERIES INCLUDED will in no event be held liable for direct, indirect or incidental damages resulting from any defect or omission in this manual, the floppy disk(s), or other related items and processes, including but not limited to any interruption of services, loss of business or anticipatory profit, or other consequential damages.

THIS STATEMENT OF LIMITED LIABILITY IS IN LIEU OF ALL OTHER WARRANTIES, EXPRESS OR IMPLIED, INCLUDING WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. BATTERIES INCLUDED neither assumes nor authorizes any other person to assume for it any other warranty or liability in connection with the sale of its products.

This manual and the computer programs on the accompanying floppy disk(s) which are described by this manual are copyrighted and contain proprietary information belonging to BATTERIES INCLUDED.

This manual may not be copied, photocopied, reproduced, translated or reduced to machine readable form, in whole or in part, without the prior written consent of BATTERIES INCLUDED.

The accompanying floppy disk(s) may not be duplicated, in whole or in part, for any purpose. No copies of the floppy disk(s) or this manual or the listings of the programs on the floppy disk(s) may be sold or given to any person or other entity. Notwithstanding the above, the accompanying disk(s) may be duplicated for the sole use of the original purchaser.

## PRODUCT IMPROVEMENTS

BATTERIES INCLUDED reserves the right to make corrections or improvements to this manual and to the related floppy disk(s) at any time without notice and with no responsibility to provide these changes to purchasers of earlier versions of its products.

## WARNING!

This manual and the programs contained on these disk(s) are protected by International Copyright laws. Duplication by any means is strictly forbidden and a violation of copyright laws.

**VIOLATORS WILL BE PROSECUTED!**

# CONTENTS

This documentation features as its main component, a Tutorial that will demonstrate each of the functions and capabilities of B/GRAPH. As each tutorial covers an individual program module, and this table of contents shows each sub-topic within a tutorial, you will find that you are quickly able to find any given command or topic quite easily.

The Reference Guide is designed to provide a secondary source of concise information on each of the major functions within the program. You will also find that almost every possible B/GRAPH command is noted on the on-screen menus, which can always be recalled by pressing CTRL-F while a graph or chart is displayed.

INTRODUCTION	1
SPECIFICATIONS AND SYSTEM REQUIREMENTS	3
FOREWORD by Dr. Stephan Kogitz	5
TUTORIAL: GRAPHING	15
Getting Started	15
Powering Up	16
Printer Selection Menu	16
Reset Printer Type	17
Graph Function Menu	18
Displaying a Graph	19
Viewing the Directory	21
Loading a Data File	21
Redisplaying a Graph	22

Displaying Second Data Screen	22
Printing Graphs	23
Printing the Second Screen	24
Saving a Graph Data File	24
Grids	25
Attaching Points	26
Switching Graph Types	27
Updating and Adding Data	28
Adding a Factor	30
Bargraph Restrictions	32
Imaging a Graph	33
Rescaling a Graph	34
Rescaling Limitations	35
Some Tricks	36
The '\$' Command	36

Market Graphs 37  
Filling a Line Graph 39  
Creating a New Graph 39  
Intervals 41  
Adding Titles 41  
Decimals and Negatives 43  
Overlays 44  
Truncating Data Points 45  
Altering Graph Labels 45  
More than 20 Data Points 47  
Missing Data 47  
Bordering the Graph 48  
Exiting to Main Menu 49  
Graphing Conclusion 49

**TUTORIAL: PIE CHARTS 50**

Creating a Graph 50  
Missing Data 52  
Printing a Pie Chart 53  
Exchanging Data 54  
Conclusion 55

**TUTORIAL: IMAGING 56**

Reviewing Images 56  
Printing an Image 57  
Labeling an Image 57  
CTRL-G (Help-Selection Screen) 58  
Entering Text 58  
Changing Font Types 59  
CTRL-Q: Quitting Labeling 59  
Image Sequence Concept 59

Creating a Sequence 60  
Displaying a Sequence 61  
Running the Show 62  
Exit to Menu 62

**TUTORIAL: FILE MANIPULATION 63**

Introduction 63  
Loading and Saving Files 63  
Exponential Smoothing 63  
Moving Averages 65  
Geometric Moving Averages 67  
Factor Rollover 68  
Algebraic Operations 69  
General Algebra 70  
Change Order of Factors 72  
Delete a Factor 72  
Insert a Factor 73  
VisiCalc DIF Conversion 73  
ASCII Conversion 74  
Browse Data 74  
Return to Menu 74

**TUTORIAL: STATISTICS 75**

Introduction 75  
Why We Use Statistics 76  
The Null Hypothesis 77  
Error Types 78  
T-Test 78  
Degrees of Freedom 80  
Variance 80  
Mean 81

T-Value 81  
Percent Above Value 81  
Testing Special Values 81  
F-Test 82  
Testing Special Values 82  
Chi-Square Test 83  
Normal Distribution 84  
Poisson Distribution 85  
Binomial Distribution 85  
Standard Statistics 87  
Mean and Median 87  
Variance/Standard Deviation 88  
Skewness and Kurtosis 88  
Quartiles and Ranges 89  
Reviewing Data 89  
An Example 89

**TUTORIAL: REGRESSION 92**

Introduction 92  
Doing a Regression 93  
Built in Functions 94  
The Data Screen 95  
Constant and Coefficients 95  
Standard Deviation of Coefficients 96  
Coefficient of Determination 96  
F-Value 97  
Standard Error of Estimate 98  
Durbin-Watson Statistics 99  
Printing Data 100  
Drawing the Graph 100

Changing the Regression 101  
Value Interpolation 102  
Plotting a Function 102

**TUTORIAL CHI SQUARE SET UP 104**

**TUTORIAL 18: CORRELATION ANALYSIS 106**

**TUTORIAL 19: MINI DOS FUNCTIONS 107**

**APPENDIX A: PHOTOGRAPHING  
THE SCREEN 110**

**BIBLIOGRAPHY 113**

**FILES: DISK FILE DESCRIPTIONS 115**

**QUICK REFERENCE GUIDE 117**

# INTRODUCTION

B/GRAPH is a professional graphics-charting and statistical analysis program for Apple II +/e personal computers. It was designed to be used by individuals in sales, marketing, administration, forecasting and general management as well as in home and small business applications. Educators, students and hobbyists will also find B/GRAPH to be of great utility in a wide variety of applications.

B/GRAPH allows the creation of numerous types of graphs and charts, as well as providing statistical and analytical tools for the evaluation of data.

B/GRAPH is not only a powerful tool with a great number of user options, but is also very simple to use for the beginner. More than one year in development, B/GRAPH consists of 10 integrated, easy to master program modules.

This manual consists of a number of comprehensive tutorials, designed to allow you to quickly learn how to use the program package and each individual program module. Each is a step by step, hands-on course that will take you through each module, explaining every function and option. Each lesson is illustrated with actual pictures of what will appear on your TV screen or monitor. This is especially useful if you will be studying the manual away from the computer.

Included is a short course in data analysis and statistics. The analytical and statistical tools available in B/GRAPH are designed for use by individuals without any prior training or expertise in this area. Although by no means intended as a replacement for formal training or individual study in statistical analysis, this chapter will provide sufficient background in basic statistical and data analysis concepts to enable the beginner to use the programs in B/GRAPH. Of course for those who simply wish to use the extensive graphing and charting capabilities of B/GRAPH, the statistics tutorial and associated program modules may be placed aside for the time being. A bibliography and suggested reading list is provided for those who wish to explore this area further.

The **Quick Reference Manual** is designed to explain the major keystroke functions in each module. Combined with the on screen menus, it will allow those in a hurry to get started to do so without having to spend a great deal of time reading the comprehensive tutorials. The table of contents will also be found to be quite handy in finding where any given Function is explained in detail.

It is the intention of the program authors to not only provide on-going support and assistance to registered B/GRAPH owners, but also to continue to enhance and refine the B/GRAPH program and provide specialized program modules in the future.

We greatly appreciate your purchase of B/GRAPH and welcome your comments and criticisms. Please take a moment now to fill out and return the enclosed owner registration card. This is the only method by which we can determine that you are an authorized, registered user of B/GRAPH and thus eligible for both assistance and revisions, and updates as they become available.

#### **Batteries Included**

Toronto

## **SPECIFICATIONS AND SYSTEM REQUIREMENTS**

B/GRAPH is designed to be used with Apple II +/e/c Personal Computers. It may be used with any Apple II +/e/c or compatible computer system.

Either a black & white or color television or monitor may be used. B/GRAPH makes extensive use of the Apple II +/e computer's color capabilities and thus the use of a color TV or monitor is strongly recommended.

A disk drive, able to operate under Apple II +/e DOS, or equivalent, is also required. Use of a second drive is optional, although recommended.

All commands and data input are entered via the keyboard.

All graphs, charts and data screens created with B/GRAPH may be dumped to an appropriate graphics printer. Use of a printer is optional and lack of a printer in no way restricts the usefulness of other B/GRAPH capabilities.

Cards and Printers supported include . . .

1a: **PKASO NE12** Printer Interface with NEC PC-8023A or C. Itoh PROWRITER 8510 or 1550

1b: **PKASO EP-12-80/100** Printer Interface with Epson MX-80 or MX-100

1c: **PKASO AP12** Printer Interface with Apple DMP

1d: **PKASO ID12** Color Printer Interface with IDS Color Prism

2: **GRAPPLER** Printer Interface with Epson Printers

3a: **APPLE DUMPLING** Graphics Card with IDS Prism and Paper Tiger

3b: **APPLE DUMPLING** Graphics Card with Epson, Anandex, Okidata, C-Itoh, NEC or Centronics

Printers used with these cards must of course have graphics capabilities or be equipped with graphics hardware modifications as required. Please check with your dealer for printer suitability and test that your current or anticipated printer or printer card can be used with the current version of B/GRAPH.

## IMPORTANT NOTE

There are several important points to be noted by all Apple owners, but particularly those with IIe's and IIc's. You may notice from time to time that the program pauses in the middle of an operation such as displaying a menu screen. This is completely normal and you will regain control of the program within a few seconds.

It is also important for IIe/c owners to be aware that if you are loading B/GRAPH immediately after running another program, you should power down the system for at least 30 seconds. If not, you may find that graphs on the high-res screen are displayed mirror-image or upside down. Using a re-boot with CTRL-OPEN APPLE-RESET will usually avoid this problem.

B/GRAPH is unable to take advantage of lower case letters, either as command input or as "labels" for graphs. Owners of IIe/c's should therefore ensure that their caps-lock key is down at all times.

Registered owners of B/GRAPH will be advised from time to time as new printers, cards and plotters are supported. Please advise us of your requirements.

In order to use this program to its fullest, you will need to save your graphs, images and data to disk frequently. We suggest that you have at hand a supply of formatted disks for this purpose. You cannot save your data or graphs to the B/GRAPH disk, since it is write protected and doing so will void any warranty we uphold for the program or disks.

# FOREWORD

## ABOUT GRAPHS

Graphs are aids to understanding. They exist in many forms, but they all share the common characteristic that they are diagrams that illustrate information. Many people are uncomfortable when presented with several columns of numbers and asked to understand and analyse them for patterns, relationships, or trends. By wisely selecting the best graph form, the designer of a graph can translate the desired information to a readable form and help the reader to understand the significance of the information. The graph, therefore, is a pictorial communication device. The same criteria apply whether the graphic information concerns the federal budget, divisional sales performance, or the historic on-base performance of a free agent baseball player.

There are various kinds of graphs. The reason for diversity is that different kinds of situations and information require different modes of presentation. The same could be said for using a sonata rather than a concerto when expressing a musical idea. A few of the popular graph types are line, bar, and pie. There are several names for each of these. Every field has its jargon, and graphing is no exception. We will refer to the above as they are in popular use. The line graph, probably the most common, will be used to establish the basic terminology and common aspects of graphs.

## WHY USE THEM?

Graphs are used to present selected information in a palatable form. It is unlikely that the graph maker will show all available information pertaining to a particular subject. Rather, highlights are usually selected which support the author's points. For example, if your division has been steadily increasing its profits, the rising line on a graph is the standard shorthand for this. The image of a rising line has greater significance for most people than a long list of numbers, even if they are generally increasing as the reader scans down the column. The baseball slugger might point to

a similarly rising line which represents his home run production over the last few seasons. The implication for the reader is that the trend is ever increasing, very desirable in both of the above situations.

Graphs can illustrate relationships not previously noticed. A regular decrease in sales at the end of each quarter could be apparent on a graph while all but unnoticeable as a series of numbers. If nothing else, the well done graph can maintain interest in a presentation while a list of numbers could result in the opposite effect. If the time available for presentation is limited, a pictorial display of the data is a speedy means of communication.

These reasons share the major strength of graphs: they facilitate communication.

## TYPES OF GRAPHS

### LINE GRAPHS

The line graph is the most commonly used graph because of its simplicity. It is essentially a series of connected points. This type will be used to illustrate the common features of popular graphs. The graph below illustrates the yearly homicide rate of a fictitious city over a period of years. See figure 1.

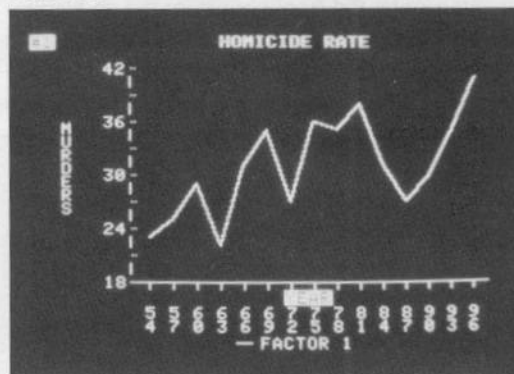
The first impression is that the rate increases fairly steadily from 1954 to 1982 and then decreases for a few years. At the end of the time interval shown, the mid 1990's, the trend shows sign of again increasing. Most people would have little difficulty in coming to roughly the same conclusion from the graph even though they may see next to nothing in a table of years and number of homicides. The graph tells a very quick story.

An examination of the above figure will provide the structural details of this family of graphs.

Title — **HOMICIDE RATE** is a concise summary of what the graph is about. The reader knows what information to expect.

Labels — **YEAR** tells the reader that time is increasing as the graph line moves to the right. **MURDERS** increases as the line moves up. These are conventions — values increase as the graph

FIGURE 1



moves up and to the right. It does not have to be this way but it has been generally done this way for a long time and confusion is avoided by continuing the tradition. The fact that years are along the horizontal line tells us that the homicide rate depends on which year is looked at. Given a year, the graph provides a rate. This is sometimes known as a 'time series'. Statisticians like that term. Most of the graphs examined here will be time series — they are very popular in business.

An example of a graph which is not a time series is the result of measuring the temperature of the atmosphere as the altitude increases. The horizontal axis would be altitude while the vertical axis temperature.

Along the axes, in addition to the labels, are numbers and short lines — tick marks. These symbols enable the reader to retrieve precise quantitative information. For example, how many murders were committed in 1978? Not all graphs display numbers along the axes. The demand-supply curves in economic textbooks include labels without numbers since they qualitatively show

a certain behaviour — in this case the intersection of two lines resulting in a selling price. The graphs in this discussion, however, are all quantitative.

There is a subjective aspect to quantitative graphs: the notion of scale. What numbers are to be shown on the axes. In time series graphs, the horizontal scale includes the time intervals being examined — the days, months, years, or periods. Generally, show all numbers unless they result in overcrowding. In that case, show regularly spaced intervals. The values must be easily read. That is the point of the exercise. Tick marks make it easier for the eye to identify particular values of the scale.

On the vertical axis, the ranges can be anything, under 60 murders, \$100 billion to \$400 billion, 15% to 97%, etc. The best scale is sometimes a matter of deciding which delivers the desired impact, hopefully without misleading the reader. Since space is always limited, it is generally best to use as much of it as possible. Showing a scale of 0 to 100 on the HOMICIDE graph would mean showing more white space and compressing the line since the space on the page remains the same. It would perhaps be justified if other graphs were being compared which went to 100. To minimize the chance of misrepresentation, it is frequently desirable to include zero as a base. This gives the reader a point of reference. It may, however, not be a concern and provide more hindrance to communication than help. A stock trading in the 88-91 range falls into this category.

The actual numbers shown on the scale should be easy to read. Numbers that are even, ending in zeros or that are meaningful gradations for the particular situation are preferred. Because the best scale depends on what the purpose of the graph is, it is desirable to be able to modify it when necessary, as B/GRAPH permits.

## THE LINE ITSELF

Based on the data, the axes and scales are established so that the data points can be plotted. The line graph simply connects these points. The result is a line which is really the sum of small line segments between each successive pair of points. On a dot matrix printer, the lines are only as reliable as the resolution (dots per square inch). For most applications, the popular printers are more than adequate. The proviso is most relevant when a great deal of rapidly varying data is

included in one graph. Plotters suffer less from this problem.

The plotted points which the lines connect can be dots or other symbols which are more pronounced. B/GRAPH allows the 'x', 'o', and '+' for this purpose. See figure 2.

FIGURE 2

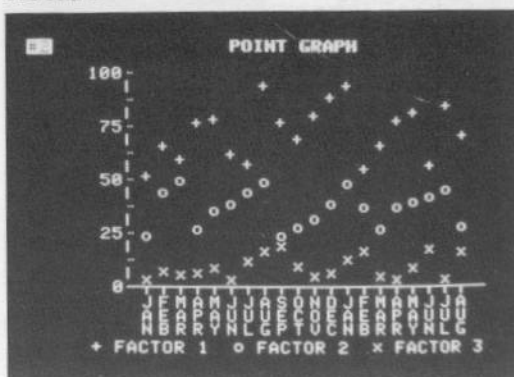
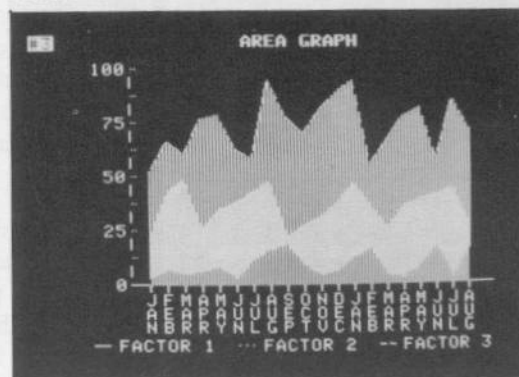


FIGURE 3



Using different symbols is necessary when plotting more than one factor on the same graph. Say, for example, that homicide rates for two different cities are shown on the same graph. Through different symbols for the data points or different line markings, the comparison is effectively made.

The component-part, or area, graph can show totals and highlight constituent contributions. In the multiple factor graph above, each line uses the same horizontal axis for its base line. In the area graph, each factor line serves both as its contribution and the total to that point as well as the base line for the next factor. The successive components are seen as layered areas.

The line of the uppermost factor represents, then, the boundary for its contribution and the total of all the factors. See figure 3.

A distinctive fill pattern of color, as in B/GRAPH, can strikingly demonstrate how one compo-

## MORE EXTENSIONS TO THE BASIC LINE GRAPH

nent is tending to become more significant. An area graph of computer sales versus time showing main frames, minis, and micros would highlight the extraordinary growth of the micro.

Multiple factors described above is just one extension to the simplest line graph. To aid the eye in reading values off the graph, lines parallel to the axes, a grid, can be drawn. See figure 4.

Since these grid lines pass through the area in which the graph is drawn, confusion may result. Again, while this facility is sometimes helpful and is supported by B/GRAPH, its effect should be evaluated in any possible application. If the graph gets too busy, the aim of the graph, communication, may be lessened.

Another extension of the line graph is extending the vertical axis beneath the horizontal axis. See figure 5.

FIGURE 4

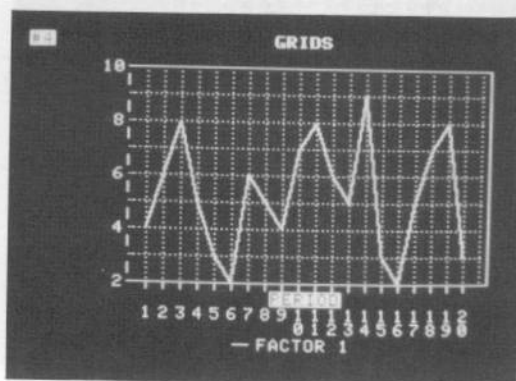
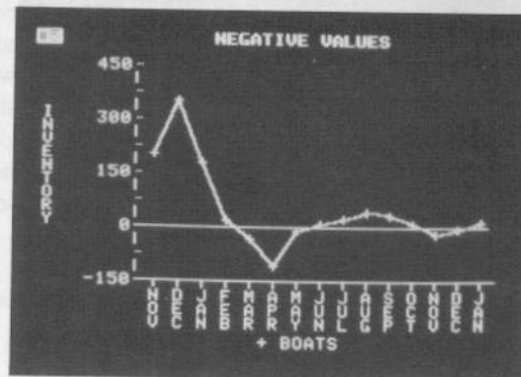


FIGURE 5



Since the horizontal axis usually represents zero, this extension allows for negative values of the

vertical coordinate. Losses are commonly expressed this way. The horizontal axis is then not at the bottom of the graph and the scale includes negative as well as positive numbers. The other considerations apply as before.

One last 'extension' will be mentioned in this discussion of line graphs, and that is not having the line at all. If it is not desired to connect the points, giving the reader the freedom to relate the points as he wishes, only the points are plotted. Connecting those points can imply information about the intermediate values that is unjustified. Daily Dow-Jones Industrial closing averages miss the day's variations. For some applications, this is an important omission. B/GRAPH allows connecting between up to three sets of data points as a changeable option.

## BAR CHARTS

A popular format that displays the same data as the line graph is the bar chart. It is constructed by drawing a thicker line, or bar, from the data point to the horizontal axis. The HOMICIDE graph below is a three-dimensional bar chart illustrating the same information as the graph earlier, but in bar form. Notice the different effect produced by changing forms. See figure 6.

One reason for using the bar chart is its visual impact. Compared with the line graph, which demonstrates variation of strongly related quantities, the bar chart is particularly effective when contrasting different entities: departments, regions, crops, etc.

Combining multiple factors in one bar chart can further aid presentation. In the figure below, precipitation is broken down into rain, snow, and hail by season. See figure 7.

The bar, rather than the data point, suggests the depth of the precipitation. The use of different markings for the three precipitation types allows the reader to see variation by season as well as by precipitation type. In effect, this two-dimensional graph simulates three dimensions. This technique does not, however, lend itself to detailed trend analysis. The bars are discrete rather than the continuous data display of the line graph. As always the application and personal taste dictate the appropriate presentation mode.

FIGURE 6

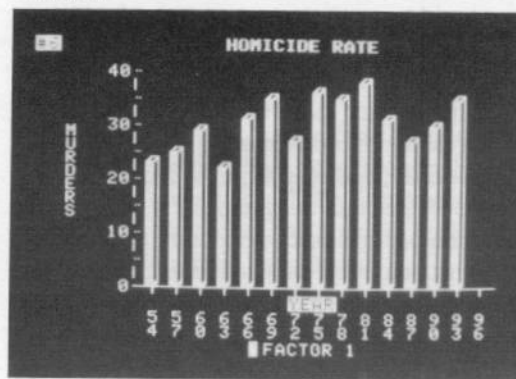
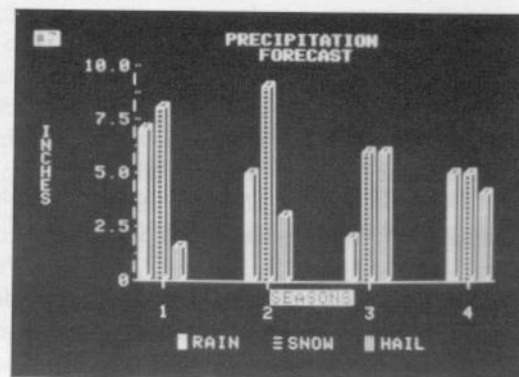


FIGURE 7



## MORE BAR CHART VARIATIONS

The figure above illustrates negative values on bar charts. See figure 8.

In this case one to three of the factors can be negative or positive.

A variation seen on the financial pages is the high-low-close chart. As its name indicates, the bar extends from the instrument (stock, option, commodity, etc) high to its low. The low in general does not touch the horizontal axis. The third of three factors is the closing price which is normally indicated as a short horizontal line extending to the right from the bar. Each bar can represent any chosen time period. See figure 9.

Another popular bar chart is the component type. On each bar, the different factors are shown, one above the other, in different design. The chart above shows three factors. See figure 10.

The component bar chart is used when it is desired to demonstrate the total of several factors as well as the factor contributions. The rain-snow-hail information now combines to show total precipitation — something not easily ascertained in the multiple factor bar chart. The trade-off is

FIGURE 8

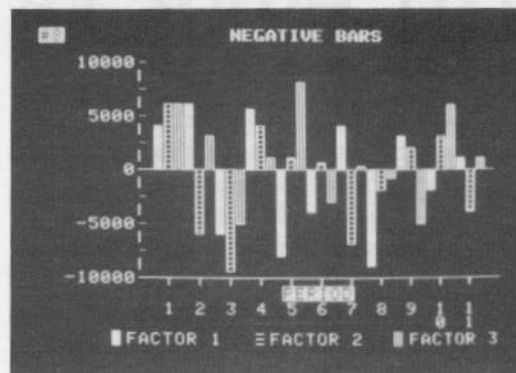
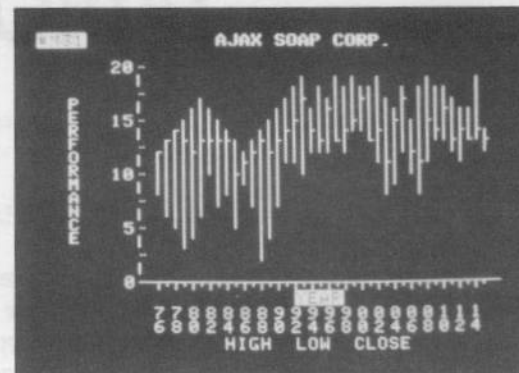


FIGURE 9



that the second and third factors, easily compared earlier, are now more difficult to extract. Again, it is a question of what the reader is expected to extract from the graph.

One technical point. Strictly speaking, the vertical bars are called columns. Common usage, however, has made bar chart, or Histogram, acceptable in most situations.

## PIE CHARTS

The pie, or circle, chart is frequently used to show how some quantity, the 'pie', is apportioned. Percentages are frequently illustrated this way. See figure 11.

By glancing at the pie, the reader can see that GM has the largest slice of this imaginary pie. Ford and GM together represent about half the sales. The precise breakdown is shown at the right indicating that those two car makers total 47.4%. The technique is understood by noting that a circle has 360 degrees. A particular percentage is illustrated as that fraction of a full circle.

Perhaps the most effective use for the pie chart is the "Where Your Tax Dollar Goes" chart



## 16 POWERING UP

Insert the B/GRAPH disk into drive 1 (or only drive) and turn on your Apple II + /e/c. Ensure that your TV or monitor is turned on at this time. After a few seconds, the title screen will have loaded and the Printer Selection Menu will appear. This screen is reproduced as Fig 1.

FIGURE 1

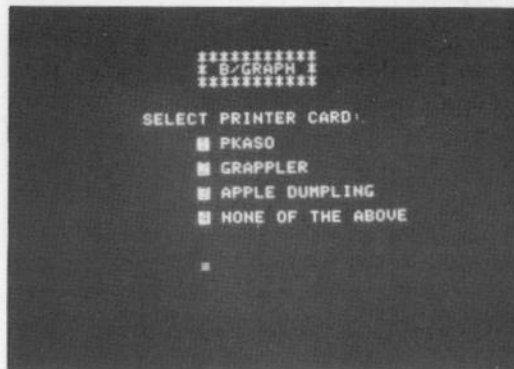


FIGURE 1A



## PRINTER SELECTION MENU

If you have one of the printer cards indicated, turn on your printer. For a discussion of printer and card compatibility, please read the Specification and System Requirements.

If you do not have a printer, or do not wish to use it, then press selection #4 (None of the Above) to proceed.

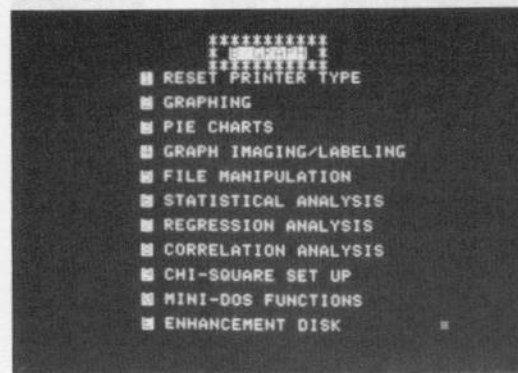
Select your printer card by pressing the appropriate number. It is not necessary to press RETURN. The disk drive will now start up again, and you will be asked how many drives you wish to use. Answer 1 or 2. The "Program Selection Menu" will appear shortly. This menu is reproduced below as FIG 2. You will find that you can return to this menu from any of the other program

modules at any time. If a disk change is required (when running with only one drive), you will be informed of this on the screen at the appropriate time.

## PROGRAM SELECTION MENU

See FIG 2

FIGURE 2



## RESET PRINTER TYPE

This option will return to the Printer Card Selection Menu. It is unlikely that after initial start-up and printer selection that this will be required again, but in the event that you have inadvertently chosen the wrong printer card, you may return to the Printer Card Selection Menu via this option.

Press 1 and return to the Printer Card Selection Menu to see how this works. After the Menu has displayed again, select your card choice and return to the program selection menu.

Selection #2 will load B/GRAPH's main graphing module. This module permits the creation of Bar Graphs, 3-D Bar Graphs, Segmented Bar Graphs, Line Graphs, Scatter Graphs and Market Graphs. It also provides a wide variety of customization routines, including Grids, Overlays and Rescaling; Printer Dumps, Graph Imaging, Data File Retrieval and more. All of these capabilities are explained in detail as they are encountered in this section of the tutorial, and may also be found in the reference section of this manual.

At this time press **2**. The disk drive will start up and after a short while the GRAPH FUNCTION menu will be displayed. A reproduction of this screen is printed as FIG 3.

FIGURE 3



## GRAPH FUNCTION MENU

Take a moment now to study this menu. It is similar in form to all B/GRAPH menus. You will find that where possible, similar or identical commands are used in all B/GRAPH program modules.

The first thing to note is that the menu screen is divided into an upper and a lower half. The

eight inverse letters in the top half of the screen — **C, U, D, R, S, I, A** and **E** — are single keystroke commands that will execute B/GRAPH functions while this menu is displayed. Do not press any of them at this time. Also note that when making one of these selections, you do not need to press RETURN.

The lower half of the menu screen is separated by a line of inverse text which reads . . .

—While graph is displayed—

The commands below this line are appropriate and active only while a graph is displayed on the screen. You will have a chance to experiment with and learn all of them shortly. They are on the Main Function Menu simply as a reminder, for your convenience, though a few of the less common commands aren't indicated.

Note particularly the line which reads; "Press CTRL-F to display this menu". At any time while a graph is on display, pressing the CTRL-F key will return you to this Function Menu. As you will soon see, the graph displayed will not be lost and may be instantly redisplayed, so that you always have access to this menu screen as a reminder of the functions available.

### Special Note to owners of 2 Drives

This tutorial and manual assumes that you have one disk drive. B/GRAPH and each of its program modules are able to support two drives. To avoid making this manual awkward in style and difficult to read, we have written this documentation under the assumption that only one drive will be used, and request your indulgence in our attempt at clarity for the majority of users.

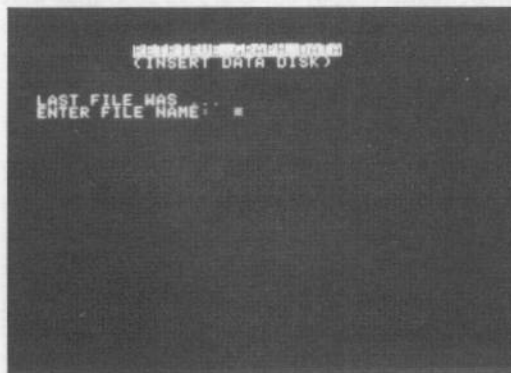
When using two drives, B/GRAPH automatically assumes that the program disk is in drive #1, and your data disk is in the second drive. You will note that there are several demo files on your program disk with the extender .DAT. These should be copied to a new data disk before proceeding with this tutorial.

With a data disk containing the .DAT demo files from the program disk in your disk drive,

## DISPLAYING A GRAPH (LOADING)

press D for Display Data File. Your screen will now display the B/GRAPH Data Directory Menu. This is reproduced as FIG 4.

FIGURE 4



You will see the line "Last File was . . .".

At this time this line will not have a file name after it as no file has yet been selected. The next time you return to this menu to retrieve or save a file, it will display the name of the last file saved or loaded. This is valuable information in the event that you wish to resave an updated file under the same file name, but have forgotten what it was originally called. Every B/GRAPH program module will inform you of the name of the last disk file in a similar manner. This information is also retained between program modules to assist you in performing a range of functions on a single file without having to double-check its name each time.

## 21 VIEWING THE DIRECTORY

To see the disk directory — the names of all the files on the disk and how much space remains for additional files — simply press **RETURN**, rather than typing a file name at this time. Those users with two drives will automatically get the directory from drive #2.

The disk will now be read and a directory of file names is displayed. Press any key to continue the directory display.

As indicated at the bottom of the directory screen, pressing **CTRL-F** at this time will abort the directory function and return you to the Main Graph Function Menu. **RETURN** will take you to the file input screen again so that a file may be loaded. Press **RETURN** now.

Be sure that you are comfortable with navigating between these three functions. Return to the display of the file input screen.

## LOADING A DATA FILE

You may have noticed that one of the files listed is titled DEMO1.DAT. We will now load that file for display. The cursor is resting just after the line 'File name:' Type DEMO1.DAT. If you make

FIGURE 5

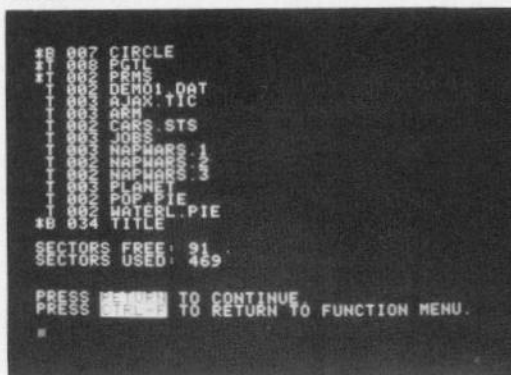
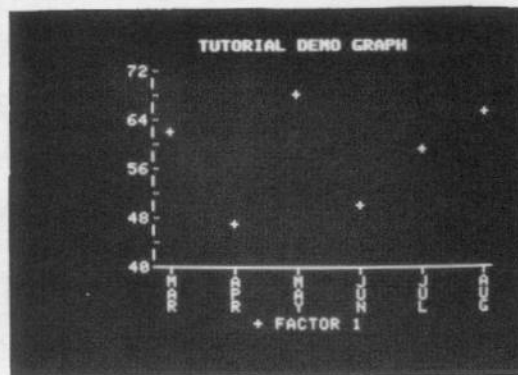


FIGURE 6



a typing error while entering the file name you may use the left and right ARROW keys to make a correction. Use the SPACE BAR to blank out unwanted characters. Now press RETURN.

If you have accidentally entered an invalid file name, or the file name entered does not exist on the disk being read, you will be notified of an error and will be returned to the File Input Screen. Please note that it is possible to accidentally hit the CAPS LOCK key on a Ile. B/GRAPH can not accept commands in lower case.

The file DEMO1.DAT will now be loaded and in a few seconds the graph will be displayed on the screen exactly as it was Saved to disk (explained later). The graph is a single factor, scatter (point) graph, with SIX data points, ie; SIX months worth of data. Take a moment and examine the graph in detail. If you are using a color TV or monitor the background will appear black and the lines and text in white.

As we noted earlier, pressing CTRL-F will return you to the Graph Function Menu while a graph was being displayed. Test this by pressing the CTRL-F keys now.

LIFE

## RE-DISPLAYING A GRAPH

You will find yourself back at the Function Menu. To redisplay the graph, simply press R as noted on the menu. The graph that was previously displayed will be instantly redisplayed.

Practice using CTRL-F and R to move back and forth between the displayed graph and the Function Menu. Also while in the Menu screen, try using the D command to look at the disk directory, and then use CTRL-F to abort that function and return first to the menu, and then use R to Redisplaying the graph.

Nothing other than Creating a new graph, Displaying another data file, or Exiting the program module will cause you to lose your graph data (in memory) or destroy the graphic screen.

## DISPLAYING THE SECOND (DATA) SCREEN

With the demo graph displayed, press CTRL-S. This displays the data that was used to create the graph, in tabular form. You may use CTRL-S to display this second screen at any time while a graph is being displayed. If there are more than 20 data points in the graph, pressing P while the

screen is displayed will display the additional data.

Since B/GRAPH can handle up to 100 data points in each of three factors, you would press P five times ( $5 \times 20$ ) to view all the data for such a graph. In our example graph, with less than 20 data points and therefore only one screen, pressing P simply redisplay the current screen.

As noted on the screen display, you may press CTRL-F at any time while in the second Data Screen to return to the main Function Menu. Practice moving back and forth between a graph display and the Data Screen by pressing CTRL-S. Remember that once you have returned from the Data Screen to the Function Screen by pressing CTRL-F you may Redisplay the graph by pressing R.

## PRINTING GRAPHS

The third special function key combination mentioned on the Function Menu is the CTRL-P key. This is used to print out a displayed graph. If you do not own a printer, you may of course skip forward to the next chapter of this tutorial, as this discussion is solely about printer dumps. If you attempt to print a graph without a printer attached or online, an error will result and you will remain at your display screen and nothing will happen.

While a graph is displayed on the screen, you may press CTRL-P and the graph will be dumped to the printer. Make sure that your printer is switched to 'On line' and press CTRL-P now. The dump will usually take less than a minute, depending on the printer that you are using.

Our experience has shown that ALL graphics printers have a hard time printing multiple columns of vertical lines. If you are having difficulty in that area here are two suggestions. First, use the heaviest paper that your printer will accept. The paper then tends to 'creep' less when being hit by the printer head and thus lines will be straighter. The second suggestion is to put a heavy typewriter pad under your printer. The reduced vibration also tends to lessen paper creep.

If your printer is not turned on, or you press CTRL-P without a printer attached, nothing will happen. If your printer IS turned on and nothing happens, it means that B/GRAPH doesn't like something about your set up. Problems may include your not having selected the correct card type

DAILY LIFE  
TRAINING V CENTERSECOND SCREEN  
DISPLAYING THE  
GRAPH

LIFE

on program start up, your printer not being on, or in the Off Line or Local position. If all of the proper conditions are met, try booting another program that you know does make your printer/card combination work.

During the printer dump, none of the keys on the computer will respond. This was done to protect the print routine and to insure a good print copy. However, once the printout is completed, you will be able to press any of the normal command keys and proceed to another function, or press **CTRL-P** again for another print out.

The use of a printer buffer is recommended if you are planning on producing a large number of print outs. Depending on the printer you use, a buffer can allow you to proceed before the first quarter of your graph has been printed.

## PRINTING THE SECOND SCREEN

You may also print out the second (data) display screen. To do this, press **CTRL-S** to display the second screen and then **CTRL-P**. If your graph has more than 20 data points, after the first screen had printed out you would then press **P** and then press **CTRL-P** to print out each subsequent screen.

We suggest that a new printer ribbon produces the best looking graphs and is very important if the graphs you are going to produce are to be photocopied. The darker the original, the better.

## SAVING A GRAPH DATA FILE

Before we experiment further with our current graph and learn how to Update and Switch graph types, we will see how to Save a graph to disk as a data file. If you do not have the GRAPH FUNCTION Menu on screen, press **CTRL-F** now.

Press **S** and the B/GRAPH Data Directory screen will be displayed. This screen works almost identically to the Load command except that the name that you type in will be that given to the file that you wish to Save, rather than one to be Loaded. All of the comments made in the previous section regarding file names and disk drive designations apply. You may of course press **RETURN** to view a disk directory prior to making a save, and you will note that the line . . .

"Last File was . . ."

now reads DEMO1.DAT since that was the last file name accessed, if you have followed the tutorial.

With a data disk in the proper drive (any spare disk with at least a few free sectors will do for now), type in the name **FILE.1** and press **RETURN**. Actually, any legal Apple disk file name would be acceptable, but we will be referring to this file later in this lesson and will call it by that name. After a few seconds the data that went into creating that file will have been saved, exactly as you saw it but without any grid lines (explained below). You will find yourself back at the Main Function Menu.

The information that B/GRAPH stores to disk only occupies a few disk sectors. The exact length depends on the amount of data that you enter into the file. If you are saving B/GRAPH Images along with data files then you will find that a disk fills up quickly. It is a good idea to have at least one formatted disk handy at all times.

## GRIDS

Return again to the Function Menu using **CTRL-F**. Remember that by doing so, you do not lose your graph or data. You will see in the bottom section two commands; **Grid** and **Ver**. Redisplay the graph and then press **G**. This will place a horizontal dotted grid on the screen. The Grid command will work with all types of graphs. (We will see how to change graph types shortly.).

To remove the grid, press **G** again. You will see that the graph is completely redrawn. The reason for this is that, on a complex graph, a grid might overwrite important data points or text and if simply erased would leave gaps on the screen. Although completely redrawing the screen adds a few seconds to the operation, it ensures that your graphs are always displayed with the utmost clarity.

To create a Vertical grid, press **V**. As with the **G** command for a horizontal dotted grid, you may "undraw" (erase) a vertical grid by pressing **V** a second time. You may of course combine vertical and horizontal grids on the same screen in any order.

Later on when we examine the different types of graphs which B/GRAPH can create, you will note that while a horizontal Grid can be displayed with any type of graph, a Vertical grid will only be displayed on Line Graphs and Point (Scatter) Graphs. It will not work with Bar Charts. The reason for this is simply that, since a bar chart is comprised of many vertical lines, a vertical grid would either become lost or appear confusing.

One final word on grids. If you had a vertical or horizontal grid displayed at the time you Saved the graph, it will not be redisplayed as that information is not saved to disk.

The reason for this is that, while either type of grid may be added virtually instantaneously, removing them takes a while since the graph must be completely redrawn. On a 100 data point, three factor graph, this delay could be annoying, so we have decided to structure the files in this manner. When saving a graphic screen as a high res Image you DO save the grids as part of the Image. This is also discussed in a later chapter.

## ATTACHING POINTS

The graph that you have been working with up until now, DEMO1.DAT is what we call a Point graph. This is frequently also called a Scatter graph, but in B/GRAPH we do not use that nomenclature as we reserve the letter 'S' for a different type of graph called Segmented Bar Charts. If you are used to Scatter rather than Point, sorry.

Point graphs have associated with them one command that doesn't work with any other type of graph; Attach. If you were to press the **A** key while any of the other types of graphs were displayed, nothing would happen.

Return to the Function Screen and note that Attach is located at the bottom with the other commands that are available while a graph is displayed. Redisplay the graph and then press **A**. A solid line is instantly drawn between the points. To remove the line, simply press **A** a second time. The graph will be redrawn without the line. You may create the line and erase it as often as you wish.

The Attach information is Saved in a graph file and will be redisplayed when a file is called

from disk. A bit later in this tutorial we will add two more factors to this graph. Factor two Data Points will be represented by 'o' and factor three by 'x'. A distinctive symbol will be displayed where factors overlap. If you do an Attach with two or three factors a dotted line will be used to connect the 'o's and a dashed line to connect the 'x's. This will be demonstrated shortly.

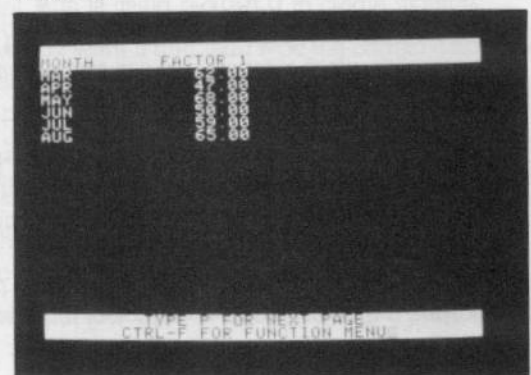
## SWITCHING GRAPH TYPES

One of the most powerful capabilities of B/GRAPH is the ability to instantly change back and forth between graph types. To accomplish this, we use the Switch command while a graph is displayed. A reminder for the Switch command is also at the bottom of the Function Menu screen. With the graph displayed (you were getting bored with it by now weren't you?), press **S**. You will now see the Graph Selection Menu. Depending on the data input, not all graph types can always be displayed. For now, press **L** for a Line graph. The graph will be redrawn in the form of a line graph. Just as a test, try pressing **Attach**. It won't work, for the reasons noted above.

FIGURE 7



FIGURE 8



## UPDATING A GRAPH AND ADDING DATA

Press **Switch** again, and this time select **B** for a Bar graph. What you will now see displayed is a three dimensional bar graph.

When we add factors two and three you will see that factor two is displayed with a horizontally striped front face and factor three displays vertical stripes on a monochrome monitor or print out, and solid blue in color.

With only a single factor, such as we have with our current demo graph, we can not display either Segmented Bar charts or Market Graphs (T/C), so we will now learn how to Update a graph so that we can add more factors and then return to examining the different graph types.

Although it is sometimes sufficient to enter current data and use it to create a graph or chart for a single purpose, in many instances new data must be added from time to time to an existing data file. Examples include the case of a Sales Manager who is using B/GRAPH to create monthly graphs of his sales performance; a teacher graphing student scores on weekly tests, or any other application where you want to add new data to an existing graph at some later date. The ability to add new information is accomplished with the 'Update a current file' command. Other, more sophisticated updating and changing is done with the FILE MANIPULATION program, discussed later.

If the graph screen is displayed, return now to the Function Menu, and note that the Update command is one of the commands that is executed while the Function menu is displayed. Excepting commands such as Display, which we used to call up DEMO1.DAT, most of the commands we have learned thus far are executed while a graph is displayed. It is important that you note the difference between these two families of commands. The Function Menu reminds you of the two groups.

Press **U** now. You will immediately see B/GRAPH's data entry screen. You will use this screen for both original data entry with the Create command and when Updating a file. We will not go into great detail at this time on all of the features available on the data entry screen, but will cover these later when we discuss Create.

Down the left hand side of the screen are displayed the names of the months. Although the file that we loaded earlier had only six data points, or items, a full twenty months are displayed. The reason for this is to allow you to add new additional data if you wish.

Across from each month name is a number. These are the numbers that were used to create the graph, and if you recall, are the same as the numbers you saw when you pressed CTRL-S to display the Second Data Screen. With an Apple II+, you use the **CONTROL** key and the **W,A,S,Z** diamond for full cursor movement to enter, add or edit data. On the Apple IIe, you may alternately use the four **ARROW KEYS**. You may move the cursor anywhere on the screen that you are able, to add or delete data at that location.

At this time do the following: move the cursor over the number '50' which is across from 'JUN'. Change the '5' to a '7'. 'AUG' is the last month with any data. Press the RETURN key repeatedly to move the cursor down so that it is across from 'SEP'. Note that RETURN does not overwrite anything. It merely moves the cursor down one line at a time and does not change anything that may lie under it.

With the cursor across from the word 'SEP' type in the number 34. If you make an error, use the cursor control keys to position the cursor where you want it and then type the correct data. You may also use the **SPACE BAR** to overwrite (blank out) unwanted data.

**RETURN** is not used to 'enter' data. It simply moves the cursor down one line. No changes made to this data entry screen are 'entered' until the screen is 'exited'. It also doesn't matter where on a line that the data is entered. The columns of numbers needn't line up.

There is a lot to learn about the data entry screen, but for the time being simply practice moving the cursor around the screen and make the one change and one addition noted above.

You will see a line at the bottom of the screen that says:

Press **CTRL-E** to terminate entry.

Press **CTRL-E** now. You will immediately see:

TO SAVE PAGE TYPE S  
TO CORRECT PAGE TYPE C

If your data is incorrect, typing **C** will simply allow you to resume editing or data entry. If the data is acceptable typing **S** will begin the screen update function. You will see a vertical bar descend slowly as your data is checked and accepted. This will be discussed in detail later.

You will now be asked if you want 'Another factor'. For the time being type **No**. You will then be presented with the Graph Type Selection Menu. Select **L** for the moment, and within a few seconds a graph will be redrawn that includes both the new data for the month of SEP and a value of '70' for the month of JUN whereas it previously contained '50'.

Note as well that the month SEPtember has been added to the bottom of the graph and the horizontal divisions rescaled to add the additional month. B/GRAPH will automatically rescale both the horizontal and vertical axes any time that you make a change to the data that necessitates such rescaling. (You will be shown shortly how to manually rescale if you are not satisfied with the vertical scaling automatically selected by the program.)

UPDATING A GRAPH AND ADDING DATA

### ADDING A FACTOR

Imagine for a moment that the numbers on the graph represented the amount of money spent each month on your electrical bill. Now suppose that you want to also graph on the same chart, your expenditure for water and sewage. To do this we need to add a second factor.

Return to the Function Menu and then press **Update**. As before, the data entry screen will reappear. Note that at the top of the screen it says 'FACTOR #1'. The data on this screen is for FACTOR 1 (our Electric Bills) only. Without making any changes to the data on the screen, immediately press **CTRL-E** followed by **S**. As before you will be asked if you want another factor. This time answer **Yes**.

In response to the answer Yes, you will be presented with a request to name the next factor and following that a data entry screen that has listed down the side the names of the seven months for which there was data in Factor 1. (If you simply press **RETURN** when asked to name the factor,

the words FACTOR 2 will be used by B/GRAPH.) The reason that there are no additional months listed is that the bottom scale would not make sense if two factors had different numbers of data points. There is no data listed beside the names of the months, since this is a new Factor. (There is though a simple method for displaying multiple factors with differing 'amounts' of data. This is discussed under the heading "Missing Data".) The words Factor 2' are also displayed at the top of the screen.

The 'Cursor' is opposite MAR. Type in the following numbers at this time.

MAR	23
APR	29
MAY	35
JUN	27
JUL	20
AUG	31
SEP	25

TUTORIAL SUGGESTION

RECOMMENDED YOUR FUNCTIONS WILL Y MORE OK

Please type in the numbers exactly as shown here, since some of the work that we will be doing in a little while requires the data to 'look' a certain way. There will be a lot of opportunity for you to make your own changes once you've mastered the fundamentals of B/GRAPH.

If you hit **RETURN** after typing in each number, the cursor will jump down a line to the correct position for the next number. You may use the cursor control keys to make any corrections. Remember, the columns of numbers do not need to line up.

Once you've finished, press **CTRL-E** followed by **S**. You will then be asked if you want another Factor. Remember, up to three factors are permitted. For now, answer **No**, and at the Graph Selection Menu select **Point** for point (scatter) graph.

You will now see a two factor graph displayed, using '+' for Factor 1 and 'o' for Factor 2. If you now type **Attach** you will see that a solid line is used to attach Factor 1 and a dotted line for Factor 2.

Type **Switch** and try a Line graph. Then **Switch** back again and try both the **Bar** and **Segmented Bar** charts. The **Bar Charts** will show factors side by side now and the **Segmented Bars** will have factors stacked one on top of the other.

You may at this time experiment with all of the features that we have covered up until now including grids, printing and displaying the second screen. If you have not already done so, we urge you to read now the Foreword by Dr. Kogitz on the philosophy of graphs. This will assist you on choosing the type of graph most appropriate for a given and application and data set. Although there are no hard and fast rules, certain types of graphs lend themselves best to certain types of presentations.

## A NOTE ON BARGRAPHS AND SEGMENTED BARS

### ADDING A FACTOR

Because of screen resolution limitations on the Apple II + /e, there are certain constraints within B/GRAPH on the number of factors and data points that can be displayed on each type of graph. Bar Graphs will be displayed three dimensionally until there is a total of more than 21 bars. This can be made up of any combination of data points or factors. In other words; two factors times 10 data points per factor, three factors each with seven data points, or even one factor with 21 data points.

Beyond 21 bars, there isn't room for the 3D type display without serious crowding, so Bar Graphs then automatically become two dimensional. Unfortunately there is also another limitation here: it is impossible to display more than 34 two dimensional bars and if Bar is selected and the total of the number of factors multiplied by the number of data points exceeds 34 then nothing will happen and you will have to select another type of graph display.

Segmented Bar Charts also have limitations. Because the bars for each of the three factors are stacked vertically, one on top of the other, up to 21 side by side 3D bars may be displayed. Once B/GRAPH has automatically switched to two dimensional stacked bar charts the limit is 34. For all other types of graphs (excepting Pie Charts) you are allowed up to 100 Data Points for each of the three factors.



One more restriction on Segmented bars should be noted: although you haven't encountered them yet, B/GRAPH is quite happy to accept and to graph data containing negative numbers. By definition, Segmented Bar Charts are designed to illustrate how several parts combine into a whole, but if one of the parts contains numbers less than zero, a segmented graph would be meaningless. This type of logical error is trapped by the program once the inappropriate data is encountered and you will be immediately returned to the Graph Selection Menu so that another graph type may be chosen.

## TUTORIAL SUGGESTION

Before proceeding, we would like you to Save the new graph that we created (by adding a second factor to DEMO1.DAT). You should still have your own disk in the drive, used previously to save FILE.1. Save your two factor graph as FILE.2.

## IMAGING A GRAPH

If you refer back to FIG 1, you will see that there is a file on the Program Selection Menu called 'Graph Labeling/Imaging'. This file will allow you to create Graph Image Shows; what some might call 'Slide Shows'. We use the term Image since what is displayed is not a data file as we have been working with up until now, but an Image or copy of the high resolution screen information.

There are several advantages to be derived from this approach, among them speed of re-display and that Images can be placed into automated sequences much as you would show with a slide projector. You can also use this program module to Custom Label a graph with different sized text.

This tutorial will shortly cover in some detail how to use both the Labeling and Image Show features, but first we must see how to save a graph Image, that will be later used by the Imaging module.

Go to the Function Menu and note the command . . .

Image current graph to disk.

### RESCALING LIMITATIONS

This is a command that is used while the Function Menu is displayed, just as with the Save command.

With your own disk in place, (it should have at least 34 free sectors available), press I. After giving a proper filename for the picture, it will be saved to disk as a standard Apple Binary file. Once the process is over, you are able to continue with all other B/GRAPH functions. The graph will have been saved to disk as a high resolution screen.

It should be stressed that, when a graph is Saved to disk as a data file, it is always available to be recalled with the 'Graphing' program module, to be Updated or otherwise altered. A graph Image can NOT be recalled with the 'GRAPHING' module. Its purpose is for use in Custom Labeling, Graph Image Shows and for saving Overlays only. For this reason, be sure that you also Save a copy of any new graph that you Create. You can always Image a graph that has been previously Saved, but you can't Display a graph Image from within the GRAPHING program module, Update it or use it in any of the Statistical Analysis Modules. Think of a graph Image as a disk version of a printer dump.

## RESCALING A GRAPH

With the Function Menu on the screen, note the Rescale command on the lower half of the screen. This is one of the group of commands that is performed WHILE a graph is displayed. It is also one of the most powerful commands available in B/GRAPH.

**Display** (load) the first file you **Saved** (FILE.1.) You may have noticed that the vertical scale is numbered from 80 down to 40. Type **CTRL-S** for the Second Data Screen and note that the data ranges from a high of 78 to a low of 50. Now **Display** (load) the second graph you saved, named FILE.2. Note that the vertical scale now ranges from a high of 80 down to zero. Displaying the Second Data Screen shows that the lowest number is now 20. It was the lowering of the lowest data values that caused the automatic rescaling.

B/GRAPH contains a complex algorithm that, every time you enter data, examines that data and makes decisions on how best to scale the graph. The factors considered include both the

absolute values and relative range of the numbers. The actual numbers used also take into consideration the 'clarity' of the numbers to be used as scale values.

For example, if you were to Create a graph to display the Dow Jones Average over the past three years, your numbers might range between 850 and 1300. If the program always made the base value zero, a great deal of the space on the screen would be wasted, and the line graphed would probably look quite flat.

Our development process has shown that the automatic scaling process produces the most appropriate vertical scaling almost all of the time, thus simplifying your job. A primary consideration in developing this algorithm was to maximize screen utilization. Thus for some applications, the scaling produced by B/GRAPH may not be exactly what you want. The Rescale command solves this problem and does even more.

With the data file FILE.2 displayed, press R. The rescaling selection screen first asks if you want Auto or Manual scaling. When Automatic scaling is chosen, the program will use its built-in algorithm. This is useful to redisplay your graph as the program originally scaled it.

Manual scaling will ask for both a new MINimum and a new MAXimum value for the vertical (Y) scale. To see how this function works enter the values 70 and 20 for HIGH and LOW respectively.

The graph will now be redrawn, and the scale will now have the new MIN and MAX values that you entered. There are several things to note about this process. If you are displaying a Point graph, only those points that appear within the range specified will be displayed. If you have an Attach on such a graph, only lines connecting points that fall within the displayed range will be displayed. If nothing appears on the screen, then all of your data points lie outside the specified range.

The same thing applies to line graphs. Lines will be displayed only if the valued for the imaginary points that they connect are within the graph's new display range. Bar graphs on the

## RESCALING LIMITATIONS

other hand WILL be drawn through a Rescaled range even if the bottom or top value does NOT lie within that range. This can lead to confusion as to where the actual minimum and maximum points are but, we feel this is preferable to showing no bars at all.

Experiment with entering a variety of different high and low values. You will find later, that particularly on graphs with a large number of data points, you will be able to create almost a 'ZOOM' effect; Zooming in on a narrow range of values and accentuating the scale for clarity.

If you forget what the original scale was, or simply wish to return the B/GRAPH's automatically determined scale, simply type **A** at the rescaling prompt. The rescale function in our REGRESSION program allows even more flexibility, but we'll leave that discussion for later.

## SMALL NUMBERS

One additional feature of Rescaling is the ability that it gives to B/GRAPH to handle very small numbers. There is a limit though. If the range interval along the Y axis is less than .005, you will be informed and sent to Rescaling. One way of displaying smaller numbers is to put something like  $\times 10(-3)$  as the side title to indicate a scaling factor. Overlays, the ability to superimpose one graph upon another also make good use of Rescaling, and we will examine that feature shortly as well.

## A RECAP AND SOME TRICKS

We've covered a lot of ground by now and probably not all at one sitting. To consolidate what you have learned so far and practise your new skills, we will recap some commands and in doing so both learn a few tricks, as well as create another graph.

We will not tell you 'how' to do the things that you have already learned, but will simply use the command name. Display the graph FILE.2 and go to the Update screen. We will be adding a third Factor this time, but lets take a moment to examine one more function of the Entry Screen.

With the cursor at any position on the screen, press the '\$' dollar symbol. This will display a dollar sign in the lower right hand corner of the screen as a reminder that this function has been imple-

## THE \$ COMMAND

mented. More importantly, when your graph is displayed, the numbers on the left hand vertical scale will be displayed with a dollar sign. You will note that pressing '\$' also returns the cursor to the first data point at the top of the screen. Pressing '\$' again toggles the dollar sign function off.

What we want to do now is add data for a third Factor. To get to the screen for Factor 3, press **CTRL-E**, and answer **S** twice. This has moved you through the first two Factors. Note by the way, (we'll practice this later), that if you had answered No after the first question about wanting another Factor, you would have displayed a graph with only the first Factor, even if there was two or even three Factors worth of data in memory.

When asked for the second time if you want another Factor, answer Yes. When asked for a Name for Factor 3, simply hit **RETURN** for now. You will once again see a blank data entry screen just as you did when you added Factor 2. Enter the following numbers.

MAR	40
APR	45
MAY	36
JUN	39
JUL	48
AUG	42
SEP	41

Press **CTRL-E** and at the next menu select a Line graph. This will show you each of three line types (solid-dotted-dashed) available. Now Switch to a Point graph and then Bar and finally Segmented Bars. These all all of the graph types that we have seen thus far, but now with three factors visible for the first time.

## MARKET GRAPHS

The one graph type that we haven't been able to display up until now, is Market graphs. There are two alternatives; Tic and Connected. Type T on the Graph Selection menu and a Market graph

will be displayed. This is the type of graph that is traditionally used to display Stock Market information, in the form High-Low-Close for share prices. When entering data destined to be a Market Graph, it is a good idea to enter it in the appropriate order;

- Factor 1 - High
- Factor 2 - Low
- Factor 3 - Close

If you do not, you will always be able to change the order of the Factors using the program module called FILE MANIPULATION, described later. Note that you may do a market graph with only two factors, HIGH and LOW, in which case pressing either T or C will give you a 'Floating Bar' type graph.

Switch now to a Connected Market Graph. There is no difference between the two types other than the tick mark being replaced with a line connecting the 'close' data. This appears to be a matter of personal preference, so we have included both types in B/GRAPH.

If you're not comfortable with what you've learned so far, take a few minutes now to experiment and refamiliarize yourself with functions such as Grids and Rescale. Save this new graph in a Market Graph form, but with any other combination of features that you fancy as FILE.3. We will use this file again in the following sections.

## FILLING — AN INTRODUCTION

You will have noticed by now the inclusion of a Fill command on the Function menu. We have delayed its introduction until now since we needed a three function graph to be able to fully demonstrate Filling

Being able to do an area fill is valuable since that type of graph conveys a sense of 'volume' better than a plain Line graph in many circumstances. Dr. Kogitz's preface on graphing describes this in detail.

## 39 FILLING A LINE GRAPH

With FILE.3 on the screen as a Line graph, press F. You will see the area beneath each factor sequentially filled in with a different pattern and color. If you now press F again, the filled areas will be erased. It is important to remember that any Fill command is an on-off switch like Attach, Vertical grid or Grid; press it once and it will Fill, press it again and it will erase. Area fills work on Line graphs and also on Attached Point graphs. The area Fill command automatically determines the number of Factors used for that particular graph and sets the colors and patterns to be used in each for maximum differentiation both on screen and on paper.

## CREATING A NEW GRAPH

Although it may seem that we've approached this program through the back door, we will now examine the first and primary function of the Graphing module of B/GRAPH: Creating a graph.

The reason for the delay has been to instruct you in all of the functions necessary to Create a new graph through learning them slowly as separate support functions for other activities.

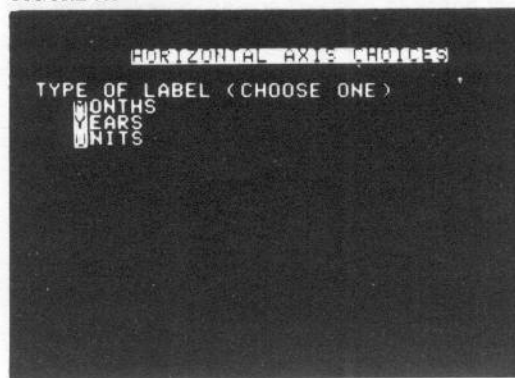
With the Function menu displayed, proceed now by pressing C. Be aware that by pressing C, you erase any other graph data which may be in memory at the time. For this reason always Save the graph you are working on before Creating a new one.

You will first be asked to name the file which you are about to create, and a file on the data disk will be opened.

The next screen displayed will be as illustrated in FIG 7A. You will be asked whether you wish to label the horizontal (X) axis as MONTHS, YEARS or UNITS. This is not an irrevocable decision; almost no decisions in B/GRAPH are.

Here is what the three choices offer: if you choose MONTHS, you will first be asked for the name of the starting month. This must be a standard three letter abbreviation such as JAN, FEB, MAR, etc. If you do not enter an appropriate month name, you will not be able to proceed. If you simply press RETURN without entering anything, the default of JAN will be entered for you automatically. You may though choose to start on any month. If you recall, DEMO1.DAT started with MARch.

**CAUTION:** If a file with this name already exists, the original will be erased. For this reason, when asked for a new file name, make it a habit to hit RETURN first, to check the disk directory.



The next question is the Interval. This can be a number between 1 and 11 inclusive. For example, an interval of 3 would display every THIRD month. The default if you press RETURN is 1.

If you choose YEARS, you would then be asked for the starting year. This can be any two digit number such as 83. The word "YEAR" will be automatically displayed on the graph to help avoid confusion with the next label type, which is "UNITS". The default if you press RETURN is the year 00. Once again, input the desired interval.

If you select UNITS you will then be asked to name the units, to a maximum of eight characters. This can be anything you choose, such as WEEKS, DAYS, CENTURY or any other time division that might be appropriate for your data. Units may also be items such as SALESMEN, PLANETS, VEHICLES or the like. The default here if you simply press RETURN is the word 'Period'.

For each type of horizontal scaling and labeling, B/GRAPH is capable of displaying 20 individual month names, years or numbers. Since B/GRAPH is able to accept input for up to 100 data points (in each of three factors), if there are more than 20 items, subdivisions are used on the

horizontal scale to indicate each item, but they are not labeled. This is simply a limit imposed by the resolution capabilities of the Apple.

## INTERVALS

As mentioned previously, when asked for INTERVAL you may simply press RETURN for an interval of one, or select any number that you wish. The interval chosen will then determine the amount by which the major horizontal base divisions are skipped. For example if you chose three (enter 3) for an Interval and a starting month of MAY, the graph would be labeled MAY-AUG-NOV-FEB etc. YEARS, if incremented by six starting with 1950, would be 50-56,62,68 etc., and similarly for UNITS. This can be particularly handy if you are graphing selected rather than sequential time periods.

## ADDING TITLES

For the sake of this example, answer MONTHS. When asked for the name of the starting month, use MAR, but first try typing some nonsense letters. You will see that only a valid month name consisting of the first three letters of each month will be accepted. If you chose UNITS, you would have to provide a name of up to eight characters long. The default label given by B/GRAPH is "Period".

There now will appear in sequence, four questions. These will be . . .

Enter top title.

Enter 2nd title.

Enter side title.

Enter graph #.

(See Tutorial FIG. 8A)

None of these questions have a default. If you do not want to place a text label on the graph in

any of these positions, simply press **RETURN** and nothing will appear on the graph. You may of course type in anything that you wish in answer to the questions.

The 'top' title will appear at the very top of the graph, with whatever text you enter up to 20 characters, centered over the graph. The second title, up to 19 characters long, will appear just under it. The 'side' title, up to 13 characters, will appear vertically down the left side of the graph alongside the vertical scale, one letter under the other. The graph number is for your convenience in identifying graphs and will be displayed in the upper left hand corner. The # sign is displayed automatically alongside your input.

Please note that each label has a maximum length as shown by the entry line on the screen under the cursor. ONLY what will fit will be displayed, even though you are able to type beyond the end of the line. As with almost all other inputs in B/GRAPH, you are not making an irrevocable decision when you make one of these entries, so at this time type in anything that you wish. For example . . .

Sales Forecast

Orbital Velocities

\$ BILLIONS

S45

You will then be asked to name FACTOR 1. If you recall, DEMO1.DAT had as it's label the word FACTOR 1. This is the default name if you decide NOT to name the label at this time. You accept the default by pressing **RETURN**. Similarly, when you are adding a second or third factor, either at the time of Creation or during an Update, you will be asked the same question for each. You may enter a name of up to eight characters long for a factor label or accept the default name "FACTOR" by simply typing **RETURN**. This decision can be changed at any time as well, so for now simply type **RETURN**.

You will next be presented with the by now familiar Input Screen. There are only a few new

things to be learned at this time. If you recall typing the \$ sign will place a dollar (\$) sign in front of the vertical scale. Press \$ now. Remember, if you decide at any point that you want to get rid of it, simply type \$ again. It will no longer be displayed on the lower right hand portion of the screen once you type \$ again.

## DECIMALS AND NEGATIVES

The world is not made solely of integers and positive numbers. B/GRAPH is capable of displaying decimal numbers as well as numbers less than zero. Try entering the following numbers.

MAR	124
APR	68.56
MAY	20
JUN	- 35.7
JUL	-234
AUG	- 45
SEP	29.89
OCT	88.01
NOV	183

You may type numbers with decimals and minus signs. On the second data screen though, only two decimal places will be displayed. All nine significant digits are used for calculations. The largest number which you can enter into B/GRAPH is 9,999,999. Any number over this will be interpreted as N/A (Not Available, or Not Applicable) and not displayed.

Using all cursor control keys, you may change anything on the input screen at any time. B/GRAPH does not read and examine the screen for your data until after you have pressed **CTRL-E** and **S** and only then is the data 'input' for setting up your graph.

Do not select a second factor at the moment, but choose a Bar graph display. You will see that the zero base line now floats upwards, and both the positive and negative entries have been automatically scaled.

Spend a few minutes experimenting with all of the different graph types and other features including Grids and Rescaling just to ensure that you are quite familiar with their functions once more. Review the previous chapters at this time if anything seems unclear. Do not remove the graph we have just created though, as we will be using it in the next section of the tutorial. To take a break now, simply Save this graph to disk.

## OVERLAYS

When producing graphs, it is frequently desirable to superimpose different graphs on top of one another so that similar, or even possibly unrelated things can be displayed and/or compared on screen or on paper at the same time. B/GRAPH has such a feature and it is available as the Overlay function. This is addressed WHILE a graph is displayed by pressing **O**.

It can not be stressed too strongly that while it is possible to create extremely useful and attractive graphs with the Overlay function, it is also very easy to produce a total mess!!

To aid in Overlaying, B/GRAPH automatically rescales the incoming graph to the scale of the graph currently on the screen. If your first scale ranged between zero and 1000 and the second between three and seven, you will not get a meaningful overlay. It is up to you to make sure that your data and graphs share relevant information in order to use the Overlay function properly.

You should also note that B/GRAPH overlays the second title on top of the first. Titles of different sizes may leave characters on the screen that cause a jumble rather than a coherent title. To avoid this, use the Update feature described earlier to reduce or erase titles that might interfere with an overlay. Remember that overlays may be printed at any time by pressing **CTRL-P**.

If your Overlay was a mess, there are ways to improve it. First, make sure that the horizontal scales are not different and share the same number of data points. FILE.2 has seven months of data, starting with MAR, while the FILE.3 has nine months worth of data. You will see how to correct this shortly.

Also, make sure that the vertical scales on the two graphs are roughly within the same range.

## PRODUCING MEANINGFUL OVERLAYS

If there are great differences between scales, you may not be able to produce a good overlay. If the differences are relatively small, try rescaling the screen graph to the data of the overlaying graph. Negative data in the overlaying graph will be lost if the screen graph has only positive values.

## TRUNCATING DATA POINTS

Reducing the number of data points is easy. Simply use the **SPACE BAR** to erase the value for OCT. This not only removes that value but also terminates the graph at that point. In other words we could have had many more months worth of data on the screen, but by making a given month blank, we are indicating to B/GRAPH that data entry ends HERE. If we had multiple factors, each of those factors would now end at the same point as well! B/GRAPH will become upset though if you do this to the second or third factor. You cannot truncate a factor other than the first one.

In order to add data points, you can again use the Update feature and add **"I"** for "not applicable" data points. This is described in greater detail later.

Before leaving Overlays, there are a few points to remember. You may have as many Overlays as you wish superimposed one on top of the other. How this looks is simply a matter of how careful you are in organizing your data. You also can NOT Save an Overlay. As far as B/GRAPH is concerned, it can only contain one graph's worth of data at a time; the last one Created or Displayed, (Loaded). You CAN Image an overlaid graph, and with a little work, these can be effective in presentations. Naturally you may also dump to a printer any graph or Overlay that is displayed.

## ALTERING GRAPH LABELS

The last major GRAPHING function to be covered is Altering the graph label information. As we have stated earlier, you always have the option of changing virtually anything in B/GRAPH and Labels are no exception.

GRAPH TITLES  
 PRESS [F1] FOR NON-APPLICABLE ITEMS  
 ENTER TOP TITLE: \_\_\_\_\_  
 ENTER 2ND TITLE: \_\_\_\_\_  
 ENTER SIDE TITLE: \_\_\_\_\_  
 ENTER GRAPH #: \_\_\_\_\_  
 NAME OF FACTOR 1: \_\_\_\_\_

From the Function Menu, press Alter and you will be presented with the first of the alterations:

Enter top title.

This is the same series of questions as the original labeling sequence and each subsequent question will be asked in turn, just as when you create a graph. If you simply press **RETURN**, whatever had been previously entered will remain unchanged. To type NEW information at any question, just type in that new information. If you wish to erase, or blank out a text line for any label, just type a period '.' and then **RETURN**. Remember to stay within the length restrictions, signified by the open "box" on the screen. If you type beyond the markers, your text will appear truncated on the graph.

You will also have the option of changing the horizontal scale (X axis), in the event that you had originally made an error in selection or you have a change of mind. You will later have the opportunity of placing 'custom' labels anywhere on a graph, in different sizes of text. This will

## HANDLING MORE THAN 20 DATA POINTS

be performed with the Imaging/Labeling module and information on how this is done is contained in the tutorial for that section.

We have mentioned that you are able to enter a total of 100 data points for each of the three factors. Since each Input Screen can contain up to 20 items, you are able to enter up to five screens worth of data.

You haven't encountered the subsequent data screens yet, since we have been dealing with less than 20 items in all of our examples thus far. You will be asked if you want to enter more data, only if you enter a full 20 data points on the first screen. To see this Create a new graph and enter a full 20 numbers. You will note that after the 20th item has been entered the cursor jumps back to the top of the screen.

When you are satisfied that the data for these 20 points is correct, press **CTRL-E**. Unlike before, when we were dealing with less than 20 numbers, you are now asked if you want another Page.

Press **Y** and you will have another data entry screen presented for that factor. If you are working with months, the sequence will continue and similarly if your horizontal scale had been selected as **UNITS** or **YEARS**, that numbering would be continued as well.

You can use **CTRL-E** at any time to go to the next factor (including adding new factors to your graph data) or to display the graph. If you choose to enter a second or then a third factor, you **MUST** enter data for the same number of data points in each of the second and third factors as you had in the first one. If you do not have data for all points in each factor, there is a way of still displaying a meaningful graph, described below.

## MISSING DATA

One function of B/GRAPH that isn't performed with a CONTROL key, or noted on a screen menu, is the handling of missing data. There may be instances where you have data for a time series, but are missing one or more data points. B/GRAPH is able to handle such occurrences quite simply.

All that is required is to type "/" for (Not Available or Not Applicable) instead of a numeric value when entering data on the Input Screen.

Do NOT leave a blank space for missing data, as B/GRAPH will interpret this as the end of data, not as missing. You must type a "/" instead of numbers to signify unavailable data. If you accidentally enter an alphabetic character in the data field, when you type S for the data update, an 'inverse asterisk' will appear on the screen beside the incorrect entry. You will have to type either a "/" or a number at that entry position to proceed.

When the graph is displayed, B/GRAPH will simply skip over that position on the time series. For example, if you are doing a line graph, the line will stop at the last position for which there is data, and resume again as soon as the next position does contain data.

As mentioned earlier, you will sometimes find that you have the need to plot 2 or 3 factors on a single graph though the amount of data for each factor isn't equal. Simply type in a "/" for all data points that are not available for a given factor and B/GRAPH will be able to plot a graph even with unequal factor lengths.

Segmented Bar Charts present one last consideration. If there is missing data in ANY ONE of the two or three data points of a given time series, then THAT time series will not be displayed. For example, suppose you have twelve months worth of data for three factors. If you are missing data for Factor #2, in June, NO stacked bar will be drawn for June, even though the data for Factor #1 and #3 are there.

## BORDERING THE GRAPH

There is one more minor function available called Box. At any time while a graph is displayed, you may press B and the left side and base line that define the edges of the graph will be completed on the top and right side to form a completed border around the graph. This is strictly a matter of taste; we feel that graphs with only a few data points look best without it. You may remove it by pressing B a second time. Border is not saved to a data file, but of course it may be part of an Imaged graph.

## 49 EXITING TO THE PROGRAM MENU

Once you have finished with a module such as GRAPHING, you will need to return to the Main Program Menu to obtain access to any of the other Modules. The way to do this is by typing Exit from the main Function Menu. If you have data in memory, you will be advised on the screen and given the opportunity of saving it to disk. Once you leave a module, any data in the computer's memory will be lost.

If you do not have data to Save, ensure that the B/GRAPH disk is installed in drive #1 and press RETURN. You will shortly see the program selection menu and should now proceed to load PIE CHARTS in preparation for the next tutorial.

## GRAPHING CONCLUSION

This concludes the tutorial section on the GRAPHING module. In most of the other program modules such things as data Input Screens, printer dumps and returning to a Function Menu work in the same manner as you have learned them up until now. Any differences will be noted when encountered but for the sake of the tutorial we will assume that you have familiarized yourself with the preceding lessons. Therefore before moving on to the PIE CHART tutorial, make certain that you are familiar with all of the concepts presented thus far, as we will be assuming this familiarity from now on. Consult the Reference Guide if you need a brief refresher on commands.

# PIE CHARTS

This tutorial will discuss the B/GRAPH program module called PIE CHARTS. It is loaded by selecting #3 on the Program Selection Module. You must load PIE CHARTS, as well as each of the other B/GRAPH modules by booting the program disk, and selecting the option you wish from that menu, or by exiting from a program module and returning to the Main Menu.

This tutorial assumes that you are already familiar with all of the functions of the GRAPHING Module. Functions that are identical between them will NOT be re-explained here, therefore if you have not read and practiced the GRAPHING tutorial yet, please do so now. If any function is unclear in PIE CHARTS, please refer back to the previous tutorial. For assistance in finding where a given command is first explained, please refer to the B/GRAPH Reference Manual elsewhere in this documentation.

## THE FUNCTION MENU

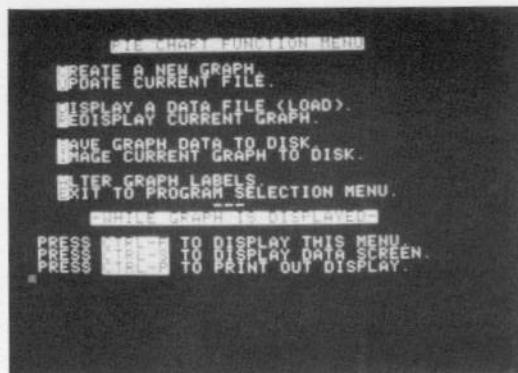
Once you have loaded PIE CHARTS, spend a moment examining the Function Menu. You will note that it is almost exactly the same as the menu in GRAPHING. All available functions work identically.

Although the top half of the menu is unchanged, everything EXCEPT CTRL-F, -S, -P is gone. Obviously Grids, Rescaling and Attaching, are not meaningful on PIE CHARTS. If you are unfamiliar with the concept of Pie Charts and how they are used to present data, please take a moment to read Dr. Kogitz's essay at the front of this manual.

## CREATING A GRAPH

Press C so that we can Create a sample pie chart. The usual file name request appears and then you will see a series of questions regarding screen titles. Pie Charts are labeled similarly to the other types of graphs previously discussed, except that there is no 'SIDE' title, and the 'BASE' title will be in bold faced text at the bottom of the screen.

FIGURE 9



Type in whatever you please for the titles and Label Type (equivalent to a Factor name) at this time. When finished, you will next encounter the Data Input Screen. Although it superficially resembles the Input Screen on GRAPHING, it is different in many ways, which you will see when we start entering data.

Down the left hand side of the screen are numbers from one to 16. The purpose of a PIE CHART is to represent data as parts or percentages of a whole. B/GRAPH can handle up to 16 'slices' in a pie, and you have to enter both the LABEL for the slice and the data. This is done on the same line of the input screen, opposite each of the numbers from one to 16.

For example, let's say that we wish to enter the number of Girl Scout Cookies that a group of girls have each sold. We would do the following: Type SUSAN, now press RETURN. The cursor will skip across the screen to the left hand side. Now type 11 and press RETURN.

You will have entered two 'FIELDS' worth of information; a LABEL and DATA. Each related set

is typed on a line one after the other. Pressing **RETURN** moves you from one **FIELD** to the other. You may now type in another name since after typing '11' and **RETURN** the cursor jumped back to the **LABEL FIELD** for the next item.

Type in the following . . .

JANE	9
BETTY	15
NANCY	18
SALLY	14
GRACE	3

You can use the cursor control keys to move back to any **LABEL** from a data point and make any changes that you wish. You can even use the cursor keys to move across the screen completely to the **DATA (Values) field**. You can also press **RETURN** to move across to the right hand entry field.

The **LABEL** you type can only be eight characters long. If you attempt to type more than this the cursor will not move beyond that point. You may type numbers in a label field, for example 'R2D2', but if you type a letter (alphabet character) in the data field on the right hand side of the screen, it will be flagged with an inverse asterisk as in the main graphing module.

## MISSING DATA

If you recall, typing "/" in the **GRAPHING Module** was a signal to the program that there was missing data. Since by definition, a pie chart is to display the percentages each part is of a whole, there can be no missing data or "/" entries.

Once you have entered the above data and experimented with using the **RETURN** key and the usual cursor controls to move between **FIELDS** on the screen and edit your data, press **CTRL-E** and **S** to draw the graph.

## 53 THE PIE CHART

The Pie Chart will now be drawn. Centered around each 'slice' of the pie, just outside the circumference of the circle, is a letter of the alphabet starting with 'A' at '3 o'clock' and moving counter-clockwise around the chart. To the right of the chart is an enclosed rectangle containing each letter that labels a slice of the pie; the **LABEL Type** that you entered and the **PERCENTAGE** of the whole that each value represents, rather than their actual (entered) values. The rectangle is **NOT** of fixed size and will change to accommodate the number (between one and 16) of items that you have to enter. The diameter of the circle however, is fixed.

These percentages have been calculated to two decimal places for precision. If you press **CTRL-S** for the second screen, you will see the **LABELS** and values that you entered, plus a **TOTAL** of all of the data at the bottom of the list.

Take a look at the pie itself; you will see that the first (A) slice is **BLUE**, the next (B) **WHITE**, the next **GREEN** and the fourth **BLACK**. On a monochrome monitor the sequence appears, **STRIPED, SOLID, STRIPED AND THEN BLACK**. This sequence is repeated around the entire chart. If the last wedge would end up adjacent to a similar color, **B/GRAPH** skips a color so that each wedge of the pie is displayed distinctly.

## PRINTING A PIE CHART

Pressing **CTRL-P** will produce a printer dump if all of the usual set up procedures have been met. Of course the second (data) screen may also be printed by pressing **CTRL-P**.

You may notice that the circle drawn on your TV or monitor is not quite perfect. It looks slightly flattened at the top and bottom. The circle printed out on paper when you do a screen dump **IS** a perfect circle on many printers since they will print the same number of dots in either direction. This family includes the Epson FX series, Oki 92 and Centronics 739. Epson MX series, as well as Gemini and Seikoshia unfortunately do **NOT** print the same number of dots vertically as horizontally and thus the pies will not appear circular.

For photographing the screen, or screen presentations where precision counts, you may be able to adjust the horizontal and vertical linearity of your TV or monitor. If such controls are not

immediately accessible, please contact a competent technician for assistance. TV sets and monitors contain lethal voltages! Do not perform any adjustments not intended for consumer alignment.

### EXCHANGING DATA WITH OTHER GRAPH FILES

It is possible for you to use this Pie Chart module to call up data files created with GRAPHING and create Pie Charts from those files. It does not matter if the files you are calling were saved as Bar Charts, Line Graphs, Scatter or Market Graphs. The sole limitation is that ONLY the first factor, Factor 1, will be used by Pie Chart, and the data must contain not more than 16 data points. None of the values may be negative, as well. If there are more than 16, the file load will abort and you will be returned to the FUNCTION MENU. The reason for this is that a pie chart is used to represent items as part of a whole. If there are parts missing the chart would then become meaningless. Similarly, there can be no negative or NA values in the data loaded into a pie chart. This will also cause a file load to abort for the same reason. Remember too that there is no side title on a Pie chart and that the second title will appear below the chart, not above.

If the data that you wish to turn into a pie chart is in a factor other than Factor 1, you can use the program module called FILE MANIPULATION to switch their order. This tutorial is found further on in this manual.

When loading a non-Pie Chart File, rather than immediately drawing a graph, you will find yourself presented with an Update screen. This is to allow you to type in label names, since obviously these were not created in the GRAPHING program. If you just want to look at the graph, simply type CTRL-E, since LABELS ARE NOT MANDATORY.

Similarly you may also load a Pie Chart data file into GRAPHING with the Display command and use it to produce any of the available graph types. It will be in the form of a single factor graph with UNITS as the horizontal axis. The LABELS will not be displayed or retained if the file is reSaved.

### 55 PIE CHART CONCLUSION

You should spend some time practicing the creation of a number of different Pie Charts and also the various functions on the Function Menu such as Altering label information, Imaging and of course Saving a data file. You will find that now that you are familiar with the basic functions of B/GRAPH, learning this and the other program modules will go quite quickly.

To go to the next program module, return to the Main Menu by pressing E and choose the IMAGING AND LABELING program module.

# IMAGING AND LABELING

EXCHANGING DATA  
WITH OTHER GRAPH  
FILES

This Program Module allows you to perform four different tasks, each using the Images created with GRAPHING, PIE CHARTS, or REGRESSION. Each of those modules allows you to create a graph and then Image it to a disk as a high resolution picture (binary) file rather than a data file. It is Image files that this module utilizes.

If you ever have difficulty identifying B/GRAPH Image Files, remember that they occupy exactly 34 disk sectors; a directory scan will help locate them. Data Files on the other hand only occupy between one and about four sectors.

Two major functions available in this module allow the Custom Labeling of graphs and the creation and retrieval of Image sequences, otherwise known as 'Slide Shows'. You may have noticed that although both the GRAPHING AND PIE CHART programs allow the creation of Images, neither permit their re-display.

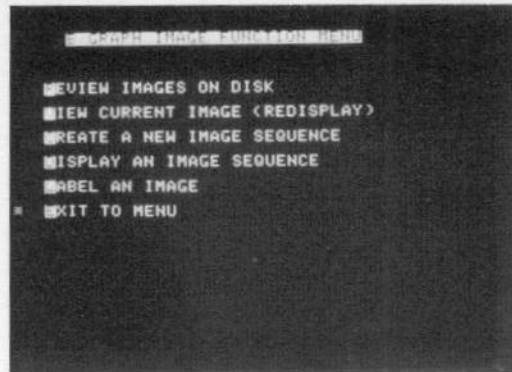
## REVIEWING IMAGES ON DISK

From the Program Selection Menu, choose GRAPH IMAGING/LABELING and shortly you will be presented with the B/GRAPH IMAGE FUNCTION MENU as illustrated in FIG. 10 below. The First selection is Review Images on disk. Select R and press RETURN for a disk directory. You will note a file named TITLE.

This is a graph Image, (you will recognise it shortly as the program's title screen). Type in the file name (TITLE) and you will see the file displayed. Be aware that you can not make any data alterations to an Image file. We are not using the data that created the graph, we are only displaying its Image or picture.

CONSTRUCTION  
PIE CHARTS

FIGURE 10



At this point, pressing CTRL-P will do a printer dump, to any B/GRAPH-supported graphics printer. You will probably want to wait to try this until after you have done some Custom Labeling, described below.

The purpose of this section of the program module, is to place text on the screen. While GRAPHING and PIE CHARTS allowed text labels to be placed on the screen and Altered, that text was restricted in size and its positioning on the graph was determined by B/GRAPH. It is the purpose of this program to provide you with the utmost flexibility in placing a variety of text in different sizes on any graph or chart, or to create text-only pages. These Images may then be printed out, or reImaged to disk for later use or inclusion in a Slide Show'.

The Review command simply allows you to view individually any Images on a disk. In order to alter the Image you need to return to the menu by pressing CTRL-F. Now select Label an Image. (A shortcut to Labeling mode from Review, is to press L while a Reviewed Image is on display.)

## PRINTING AN IMAGE

## LABELING AN IMAGE

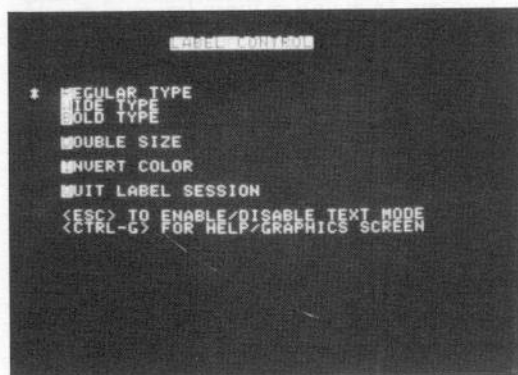
SCREEN  
CURSOR & MENU SELECTION  
THE IMAGE  
SEQUENCE  
LABELING MODE

Reload the demo Image called TITLE. This time, you will be asked for the name of the OUTPUT FILE. This is the name under which your altered Image is later saved. The Image file is then called to the screen.

### CTRL-G: HELP-SELECTION SCREEN

There are several powerful commands in this program module. To assist you in remembering them, we have provided a 'HELP'-Selection screen which can be viewed at any time while in Labeling by pressing **CTRL-G**. This will be displayed after you have named the Input and Output file names in any event. **CTRL-G** is a toggle command, and you move back and forth between the Image you are Labeling and this menu with it. This screen is reproduced as FIG.11.

FIGURE 11



### ENTERING TEXT

Press R for Regular type and you'll note that an asterisk turns on and off next to the command line. This indicates in each instance which 'feature' of labeling you have selected. Now press **CTRL-G** and the Image will be displayed. A cursor in the form of a 'flashing asterisk' is located in the lower left hand corner of the screen. Using the **CTRL-W**, **A**, **S** and **Z** keys (up, left, right and down), practice moving the cursor about the screen. Place it at an empty location.

As was noted on the 'Help' screen, pressing **ESC** enables/disables 'Text Mode'. Press **ESC** now. You'll see that the cursor changes to a 'flashing underline' shape. This is how you tell if you're in text entry mode or not. Now, type a short sentence. You'll see that whatever you type is displayed on the graphics screen just as if you were typing on a normal text screen. Anything you type will overlay whatever is already on the screen. If either type of cursor gets too close to the edge of the screen text entry will not work; neither will cursor moves. This is indicated by a 'broken' cursor.

### CHANGING FONT TYPES

There are several type sizes available to you. After you've typed something on Regular size, press **CTRL-G** and return to the 'Help'-Selection menu. You'll note that you may select Wide and Bold type styles as well as Double size. If you have a color TV or monitor (the display is not very meaningful in monochrome) also try Invert color.

You will find that each of these settings is noted by B/GRAPH with an asterisk (\*) beside your selection so that you can tell at a glance which ones are active at any time. You'll quickly discover that you can use the **SPACE-BAR** as an eraser, and you can speed-up removing unwanted text and graphics by using the **SPACE-BAR** in Double size mode.

Remember that **ESC** turns labeling mode on and off while the graph Image is displayed. You see that pure cursor moves are allowed in the text entry mode as well. Why have two modes then? Simply that you can be in 'Labeling', without the risk of someone, or even yourself messing up a screen accidentally. That's why you must press **ESC** to enter text-entry mode.

### QUITTING

When you have completed labeling an Image, return to the 'Help'-Selection menu and press Quit. Your graph Image will be saved to disk under the name previously given to it.

### THE IMAGE SEQUENCE CONCEPT

Since you can attach your Apple II +/e/c to just about any TV or monitor, including Large Screen Video projectors, you may find it convenient to display graphs and charts created with B/GRAPH to an audience. While calling up data files with GRAPHING or PIE CHART is possible, it is not terribly convenient.

## CREATING AN IMAGE SEQUENCE

Using the Review function to call up Imaged files is much quicker and more elegant, but requires the constant entry of file names and thus doesn't lend itself to presentations.

We have therefore included in this module of B/GRAPH the ability to both Create and Display Image Sequences. You are provided with most of the functions of a good slide projector, including the ability to place the Images in defined sequences; set the number of seconds that Images will remain on the screen; skip forwards or backwards to an Image out of sequence or pause during a sequence.

We assume in this section of the tutorial that you have prepared a minimum of three Images and these reside on the disk in your primary drive. If you do not have these Images prepared, please return to GRAPHING or PIE CHARTS and produce them so that these instructions can be meaningful to you.

From the IMAGE FUNCTION menu, press: **Create a New Image Sequence**. The first question asked will be how many Images you intend including in the sequence. A standard diskette can hold 16 Images. B/GRAPH though, is capable of supporting multiple drives to sequence up to 40 different Images in a single Show.

The first question is, **How Many Images**. Enter the number of Images to be shown that you have prepared. You will then be asked for the number of seconds that you wish each Image to appear. If you type zero, then Images will remain on the screen until the next one is requested by you.

There is no maximum limit for display time, however you should know that the Apple takes approximately 10 seconds to load an Image. The length of time an Image stays on screen is therefore about 10 seconds longer than the actual time you enter. Thus, if you wish an Image to display for 15 seconds, you would enter a time of 5 seconds. For now, select zero (0), but note that a minimum reasonable time for viewing an Image is at least 15 seconds.

The next query is for a name for this sequence of images to be shown. This is the file name

that you will use later to call up the sequence. The instructions that you are giving B/GRAPH now regarding this sequence will also be stored to disk under this name. Call it 'SHOW'.

You will next be asked for the filename of Image one, then Image two and so on up to the number of Images that you had indicated were to be included in this sequence. If all the Images are on disk #1, then simply type in the names. If they are going to be on other drives, then append the drive number to the name of the Image in the form: TITLE.D1 or .D2 etc. If your drive controllers are in different slots then you will also need to append .S5 or .S6 etc followed by .D1 or .D2 etc. For example an Image on disk #3 with controller in Slot #4, would be entered into the show sequence as PICTURE.S4.D3 (Note the double ".").

It is not necessary that the files be actually named with this appendage. The extenders are only used when NAMING an Image as part of a sequence.

The Images DO NOT need to be on the disks while you are entering this sequence information. They must, of course, be on the correct disks, in the appropriate drives when the sequence is displayed later on.

After you have entered the last file name, the information that you have entered is saved to disk under the sequence name that you gave it. Since sequences take only a sector or two on the disk, you can have a great many different sequences of the same Images resident on one disk, although you are of course limited to the number of Images that may be stored on a single disk.

## DISPLAYING AN IMAGE SEQUENCE

From the IMAGE FUNCTION menu, select **Display an Image Sequence**. You will be asked to name the sequence. Enter the name that you gave our sample sequence (SHOW), and a brief message will appear indicating that pressing "S" will commence the sequence. The screen will now go black and the Apple II +/e will wait until "S" is pressed to start the show.

This feature allows you to set up a presentation sequence and remain ready to run it without unnecessary text on the screen. Of course you could always create a title graph using Custom Labeling and use that as your stand-by screen.

Press "S" for Start and the first Image will be displayed. You may now move forward to the next Image by pressing the SPACE BAR (if there is no timing sequence in effect). You may go backwards an Image at a time by pressing "B" for Back. If a timed sequence is running, "P" for pause will interrupt the sequence until any other key is pressed to resume. To quit a sequence, type A (Abort) for the Function menu. At the end of a sequence pressing the SPACE-BAR will return you to the Function menu as well.

If you wish to view Images out of sequence, you may do so by simply typing the number in the sequence of that particular Image. In other words, to view the fifth one in the sequence just press the number 5, RETURN need not be pressed.

For Images 11 to 20 press SHIFT and then a number. Image #10 is accessed by pressing '0', and #11 by pressing SHIFT + '1', Image #12 is SHIFT + '2' and so on. The twentieth Image is addressed by pressing SHIFT and '0'. Images beyond 20 can not be addressed directly from the keyboard.

## EXIT TO MENU

As with all B/GRAPH modules, you can move from the FUNCTION menu back to the main Program Selection menu by pressing Exit. Be sure that you have removed your working disk and inserted the B/GRAPH Program disk into drive one.

This concludes the LABELING/IMAGE SEQUENCE tutorial. As with all B/GRAPH program modules, a bit of practice will help make all of the functions familiar.

# FILE MANIPULATION

Often times, you may find you wish to alter the order of factors in a data file, add in factors from another file and perhaps alter the data itself. This is in part what FILE MANIPULATION (number 5 on the main menu) allows you to do. It works on the data, not on a graph or Image. It also allows you to perform exponential smoothing and moving averages on your data and to save and load files in the VisiCalc (R) D.I.F. format or read ASCII files saved by other programs or by the CHI SQUARE SET UP program, discussed later. Since this program usually involves the alteration of data in your files, it is suggested that you go to Browse immediately after Loading a file and print the data by hitting CTRL-P. This will give you a record to refer to as you go about making your changes.

## LOADING AND SAVING FILES

By now you should know how to load and save files in B/GRAPH. Remember that if you want a disk directory simply hit RETURN in either of these modules and the directory will be provided. To continue and either load or save, press RETURN again, to go back to the main menu, press CTRL-F. You must have a file loaded into memory to perform any functions in this program, so either use the Load, VisiCalc or ASCII options to get a file into memory.

## EXPONENTIAL SMOOTHING

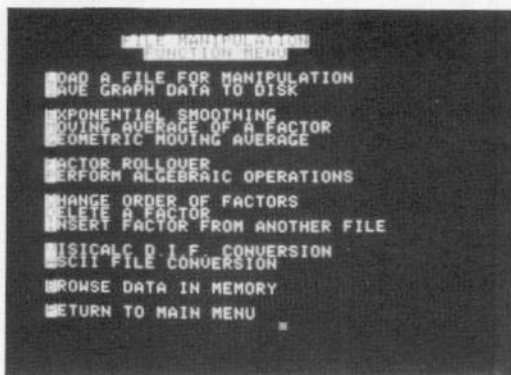
Exponential smoothing is a deceptively simple looking method, essentially a sales forecasting technique, although it can be used for other areas of time-series projection. It involves using actual prior sales (data) and the estimate of prior sales to project future sales. It works only on the first factor in a data file; other factors are destroyed in memory (although they remain in your disk file).

The formula for exponential smoothing is:

$$F(t+1) = \alpha \times A(t) + (1 - \alpha) \times F(t)$$

where "t" is the time period, "A(t)" the actual sales in that period, "F(t)" the forecast or estimated sales and "α" the smoothing coefficient, explained below.

FIGURE 12



Since you need an estimate for the forecast to begin the process, you are given the option of entering a value or hitting RETURN which will use an average of the first five data points.

The Smoothing Coefficient is a value between zero and one, used to smooth the peaks and valleys in the data. The closer to one, the higher the emphasis placed on past actual sales. The closer to zero, the greater weight is placed on estimated sales. Equal weight is, as you might expect, .5. If you chose .1, then the formula would use 90% of what was estimated and only 10% of what actually happened. With a coefficient of .9, then 10% of the estimated value is used plus 90% of the actual sales.

Use the L command and load in the example file, ARM. Then use the B command for Browse to examine the data. Remember that only the first factor will be used. The figures represent the amount of money spent, in \$billions, by the 113 developing nations of the world, on arms over a 20

year period. The other factors are arms expenditures in the 28 developed nations and total world foreign aid. Return to the menu and press E (not used for Exit here; R for Return is).

In response to the question for an estimate, press RETURN. For a smoothing coefficient, enter .5. The computer works and will shortly return a "Standard error of estimate". This value is the test of the closeness of the fit of the smoothed curve. The lower the value, the closer the fit. By experimenting with the smoothing coefficient, you will see what effect it has on the standard error. Mainframe programs usually solve for the smoothing coefficient that best fits the data, done by calculating the curves for every possible smoothing coefficient and choosing the best fit. Not only does this process take a long time, but seldom is this level of accuracy warranted. Instead, we leave it to you to find the most reasonable value for the coefficient. Commonly used values for the smoothing coefficient are .1, .3 and .5 with .1 and .3 being the two most common.

Exponential smoothing is best used on data which seems to fluctuate randomly about a given value, rather than data with a specific trend or seasonal variation, although it can still be employed for such data. It is also best for short rather than long range projections. A good combination can be obtained by using a moving average followed by exponential smoothing of the averaged data. In order to do this, you will have to swap FACTOR 1 and FACTOR 2 immediately after doing a moving average.

After you have performed your exponential smoothing, return to the data via B. You will see that the smoothing builds upon itself as it progresses down the data list. At the very bottom of the screen is the projection value for the next period.

## MOVING AVERAGES

A moving average is an average done for a given number of adjacent data points in the first factor in a data file. It simply adds the values for the points within the period and divides the total by the number of points. The period must be an integer greater than one; you can't do a moving average on .7 or -6 or similar numbers. Moving average works only on the first factor in a file; if you wish to use another factor, you will have to use Change to alter the order of factors.

If you have sales data for each month in a ten year period, you might do a moving average of

12 months to examine average yearly sales. In this way, seasonal variations can be damped out and a better analysis of sales patterns maintained. The program averages twelve months at a time, dropping the earliest month in the list and adding the next month past the end of the current group to it as it progresses. This type of averaging is used for businesses with wide seasonal fluctuations in sales, seasonally adjusted unemployment figures, power usage and other areas where fluctuations occur.

For example:

Month	Sales	Moving Average
1	210	N/A
2	310	240
3	200	303
4	400	310
5	330	335
6	275	368
7	500	383
8	375	425
9	400	325
10	200	325
11	150	217
12	300	N/A

To get the averages (seen on the screen under the "smoothed" column heading), we add 210, 310 and 200, then divide by three. To obtain the next value, we add 310, 200 and 400 and divide the total by three.

The moving average used in B/GRAPH is centralized. This means that if you give a period of three months as above, you will get no average for the first and the last points: averages will begin at the second point. If the period of the moving average is even, B/GRAPH will adjust accordingly

and move the second period further along automatically. The process involves doing a moving average for the period you give — say 12 — then another with a period of the same value but moved along one and then it averages the two to get the correct values. Those points outside the range will be marked N/A, for "not applicable". There simply is no data from which to give an average before nor after the figured points above.

Obviously, the period you give must be reasonably related to the data on hand. A seven month moving average will not give much useful information about a year of data, since the first and last three points will not be considered! Remember too that, while good for comparisons, a moving average is not used per se for a "what if" extrapolation.

The third column on the screen, once you have performed a moving average and have gone to Browse to examine the data, is DELTA. This is the difference between the averaged (smoothed) data and the actual data in Factor 1.

Moving averages are used to make forecasts, to identify seasonal components in the data or to smooth out the effects of random influences as a prelude to a projection of a secular trend, cyclical variation and seasonal variation. The data produced in the smoothed (ie moving average result) column and the DELTA of the actual to smoothed may be saved with the first factor as another file and analysed in greater detail by other parts of B/GRAPH.

One defect of using moving averages in forecasting is that they do not generate a mathematical function for predicting future values. However, a regression analysis on the smoothed valued and the ratio can be a powerful tool for that purpose.

## GEOMETRIC MOVING AVERAGE

This is essentially the same as the moving average above except that it takes the values, multiplies them together and then takes the  $n$ th ( $n$  refers to the period of the moving averages) root. This is most often used when a relatively constant geometric force is at work on the data; ie a definite trend line such as inflation. For example:

MOVING AVERAGE  
GEOMETRIC

Year	Sales	Moving Average	Geometric Moving Average
1	100,000	N/A	N/A
2	110,000	110,333	110,000
3	121,000	N/A	N/A

As you can see, there is a difference in results; in this case, the geometric moving average figure is closer (exact actually) to the actual sales figure than the regular moving average. For another example, Load the file ARM and compare the results of a moving and a geometric moving average on the first factor. Also, change the period of the average to see what results with each different period.

## FACTOR ROLLOVER

Especially in the case where you are using B/GRAPH for ongoing business, it is useful to be able to add elements to either the top or the bottom of a factor without having to retype the entire file again. This module allows you to do Rollover; removing one or more items from a factor and replacing them with new values.

Your first question will be whether you want the items rolled off the top or bottom. It's a good idea to refer to your printed sheet of the original data to be sure which way you wish to roll off items.

The second question is how many items are to be rolled off. Obviously you cannot roll off more items than you have, although you can roll off an entire factor's worth. However, you will have to enter new values by hand in the correct order for each one you roll off.

Next you will be asked for which factor to roll off. You enter one, two or three. If you have more than one factor to do this to, you will have to do each one separately, since we only allow you to perform rollover on one factor at a time.

Finally, after the computer has worked on a file, you will be asked to input the values for the

## PERFORM ALGEBRAIC OPERATIONS

ONLY  
ON ONE EQUATION  
GENERAL UTILITY

new items. If you rolled off the top, you will be asked for the highest value first, then the next highest and so on until finished. If you rolled off the bottom, you will be asked for the lowest first and then the next and so on until the last item is input. You may input a character but as normal in B/GRAPH, such input will be treated and appear on the screen as "/". Note that the roll off feature allows you to examine data for lagged correlation (see Correlation Tutorial).

This is a particularly important and powerful portion of the file manipulation program, so use it with care. It allows you to alter your data mathematically by either choosing a pre-defined function from the sub menu or by entering your own equation.

The first four predefined equations we have provided are:

- Divide factor one by factor 2
- Multiply factor one by factor two
- Subtract factor two from factor one
- Add factor two to factor one

These are straightforward arithmetical functions that perform exactly as indicated and the results of each action become the new values in factor one. For example, if you add factor two to factor one, the total is now in factor one. If you wish to perform these operations in a different order or on different factors, you will have to use the Change order of factors module, discussed further on. To see the results, return to the main menu by pressing any key and then choose Browse.

Convert data to a percent of the total is also quite straightforward; it converts each item into a percentage of the total of all factors. For example, if you have three factors with item number one as:

Factor 1	Factor 2	Factor 3
10	20	30

You would get these numbers added together for a row total of 60 (not shown) then calculated as a percent of that total. If you checked the data by using the Browse function, you would now see it reading as:

Factor 1	Factor 2	Factor 3
16.67	33.33	50

We do not add up columns (factors) to get the total value for all of the elements in a factor and then give you a percentage based on that total. The percentage is based on the total of a row. Therefore, if you have only one factor, you will get 100 for every element, because it is 100% of the total!

## GENERAL ALGEBRA ON ONE FACTOR ONLY

This selection allows you to perform more complex algebra on one factor at a time. You may alter all of your factors in this manner, but each must be done separately.

FIGURE 13

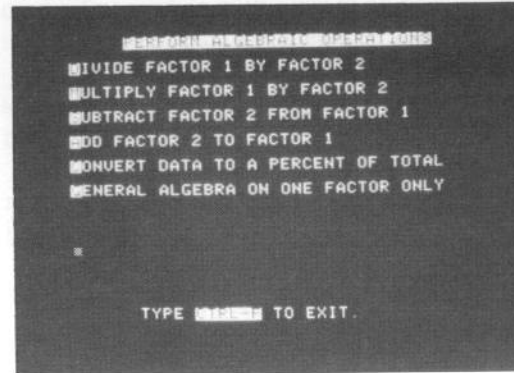
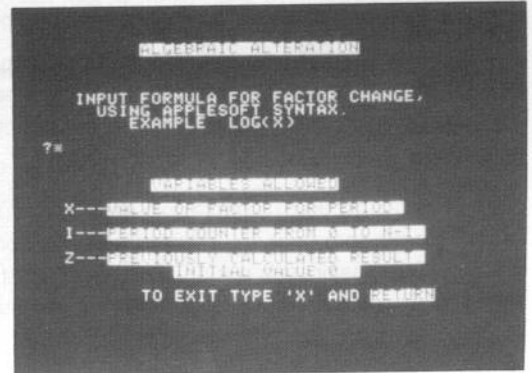


FIGURE 14



You may input any valid BASIC equation such as  $\text{SIN}(X)$  or  $\text{LOG}(X) + \text{COS}(X)/2.5$ , including Boolean logic. Refer to your Applesoft manual to see what is valid and what mathematical functions APPLE BASIC supports. If you simply input a single number such as 25 or 6, it will make all items in a chosen factor into that number.

When performing General Algebra, you will find that B/GRAPH goes to the disk in drive 1 for certain routines. On a single drive system this will require that you insert the program disk.

We have listed the three legal variables below the input line. These are:

**X** — this is the item or element to be worked on. If your first item is 10 as in the example above, then  $\text{SIN}(X)$  will be  $\text{SIN}(10)$ .

**I** — this is the counter for the number of items in a factor, starting from zero and counting up to the number of elements minus one. This is useful for performing incremental adjustments. For example, if your factor has five elements of value 10 in each one, then the formula  $X \times I$  would give you results of 0, 10, 20, 30, and 40. Remember that the first value for I is zero, so  $10 \times 0 = 0$ .

**Z** — this is the previously calculated result, initially zero. This is useful when incrementing a factor by an amount related to the previous result such as interest rates. With a factor of five elements, all ten again, the formula  $X \times I + Z$  would produce the results of 0, 10, 30, 60, 100, 150 respectively.

You will then be asked to which factor you wish to perform the function that you have chosen. You may enter a factor number (you will get an error message if that factor does not exist) or you may press **CTRL-F** to return to the menu. Once the computer performs the function, it returns you to the menu and you must choose **Browse** to see the results.

The general algebra function has many uses. For example, to convert degrees to radians, you would enter  $3.141593 \times X/180$ . To see a 5% discount of money (annually deposited) you would use  $Z + X/1.05^{(I+1)}$ . In the same manner, to accumulate money at 5% interest where each element is the deposit in the account for the year, you would use  $Z \times 1.05 + X$ . Factorials can be done up to a maximum of ten elements by using  $Z \times (I+1) + (I=0)$ . To get a running product where item one is

itself, item two is item two times item one, etc, use  $Z \times X + X \times (I=0)$ . For a running sum, simply use  $Z + X$ .

As another example, if you wanted to see the balance outstanding on a 20 year loan at 10% interest, Initial amount \$100,000 and an annual repayment of \$11,745.96, you would make a file of 20 elements, all zero (the value doesn't actually matter here since we don't refer to it in the calculations). Then to get the balance each year, use the formula:

$$100000 \times (I=0) + Z \times (1.1) - 11745.96$$

## CHANGE ORDER OF FACTORS

GENERAL ALGEBRA  
ON ONE FACTOR  
ONLY

This is quite simple; it allows you to make one factor any other factor in memory. You cannot make factor one into factor three if you only have two factors to begin with (an error message will result from trying). To do that, you will first have to use the Insert function, discussed below, in order to increase the number of factors to two or three.

You will be asked which factor is to be the new factor one, then two and three if you have that many in memory (it's pointless to perform this function on a one factor file . . .). It is possible to replicate a factor by having it as factor one and two, for example. This might be useful before performing algebraic operations, in order to retain the original data in another factor.

## DELETE A FACTOR

**Use this function with some caution:** once decreased, you cannot increase the number of factors except by using the Insert function, discussed below.

Again, this is a fairly simple function. It will delete from memory any one of the up-to-three factors in memory. You will be asked which factor to delete. The remaining factors will be 'renumbered' to fill any gap created.

## 73 INSERT A FACTOR FROM ANOTHER FILE

This is particularly useful when you need to either exchange the data in a factor in memory for that in another file or to increase the number of factors in the file in memory. You will be asked to give a file name for the file on disk.

The file will be loaded when you give its name and you will be asked which factor in that file you wish to insert. A buzzer will sound if you ask for a non-existent factor. You will then be asked to input the factor to be replaced in the file in memory. If you have one factor only, you may make the new factor either one or two. You cannot make it factor three if you have only one to begin with. Similarly with two or three, you may make it one, two or three. This and the Update function (in GRAPHING and PIE CHARTS) are the only ways you can increase the number of factors in B/GRAPH. You may then Save the file to disk if you wish.

If you wish to increase a one factor file in memory to a full three factors, you will have to use the Insert function twice; once to add factor two and again to add factor three. A feature of B/GRAPH is that both files need not be of the same length. If the file being added has more data points than the file in memory the excess will be dropped and if it has less data, then the file will be padded with " / "s.

## VISICALC D.I.F. CONVERSION

One of B/GRAPH's features is its ability to load and save files in the VisiCalc (R) D.I.F. format. This gives you the ability to visually illustrate your VisiCalc files with our graphs and to use the B/GRAPH files in your VisiCalc calculations.

When you use VisiCalc, you are given the opportunity to save your data by row or column. You should save it by whichever manner you have it listed on the screen, so if you have three columns of thirty elements, then save it by column and if you have the same data in rows, then save it by rows. B/GRAPH will read in the data as factors one, two and three, in that order. B/GRAPH does not send or return labels to a D.I.F. file.

To Load a VisiCalc file, you simply enter the filename (hit RETURN alone for a disk directory). VisiCalc files have the DIF extension on them. If you neglect to add the extension to the filename,

then B/GRAPH will add it for you. Once loaded, you may manipulate the data in the same manner as any file in memory and save it again as either a B/GRAPH file (using the Save function in the program menu) or as a VisiCalc file, as described below.

To Save a file in the VisiCalc DIF format, simply hit **S** in this module and provide a filename. B/GRAPH automatically adds the DIF extension on the filename if you leave it off. You may then load your file into VisiCalc by row or by column as desired. Refer to the VisiCalc manual for further information.

## ASCII FILE CONVERSION

This function allows you to Load an ASCII file into memory, either produced by your own program or by using the Chi Square Set Up program, described later in the documentation. You may then save the file as a B/GRAPH data file using Save in the function menu. You may also save a B/GRAPH file as an ASCII file in the same manner as a DIF file.

ASCII files supported by B/GRAPH have a limited amount of data in them that can be read by this function. Their structure is as follows. They must have as their first element, the number of factors (maximum three), then the number of data points in each factor (one number only), followed by a sequential list of the elements for each factor in order (one, two then three).

ASCII files can have no labels or factor names, simply the raw data. You can add all of these items once you have Saved the file, by returning to the Main menu using Return and choosing 2, Graph. The Alter graph labels will allow you to add labels as necessary.

## BROWSE

Browse simply allows you to examine the data you have in memory, not change it. You will notice that we have rounded off the data to two decimal places, since the most common use for much of the functions in this program is financial. To print the data, press CTRL-P.

## RETURN TO THE MAIN MENU

Once you have done with your file, you may return to the main menu by pressing Return. If you have data in memory, you will be given a chance to save it or to continue to the menu.

# STATISTICS

## INTRODUCTION TO THE STATISTICAL PACKAGE

Statistics and statistical analysis are not subjects easily explained in a few lines, nor is it really the place of this tutorial to explore the depth and intricacies involved in the topic. However, far from the dull and mundane task it may appear to the uninitiated, statistics can be a fascinating, highly informative study for anyone with even the mildest interest in cause, effect and relationships of data. We assume that the reader has at least a modicum of such interest already, otherwise they wouldn't have purchased a program like B/GRAPH.

When we use the word "statistics", we are referring to one of four meanings for the word: the subject or discipline, secondly as methods used to collect or process or interpret qualitative data, thirdly as the collections of data gathered by those methods and finally specifically calculated figures such as an average or a coefficient. You will see examples of this in the descriptions below, making it much easier to appreciate the uses of the different functions in this package. Once again, the reader is recommended to those books in the bibliography and outside for a deeper appreciation of the whole range of mathematical and statistical analysis topics.

There are actually four programs associated with our statistical analysis package: Statistical analysis (number 6 on the main menu), Regression (number 7), Correlation analysis (number 8) and Chi-square set up (number 9). Moving averages and exponential smoothing, plus algebraic manipulation of factors, is performed in the File Manipulation program, number 5 on the main menu (discussed previously).

This part of the tutorial will follow each program function in the menu order. This may be somewhat confusing initially, but should be clear once read through.

*"Figures can't lie, but liars can figure".*

That old saw is quite apt here. It is very easy to draw seeming relationships between unrelated data and assume that they are in some way correlated. For example, in the 1960's, there was pointed out an inverse relationship between the amount of milk consumed and the number of drug addicts around, the tongue-in-cheek message being that if you don't drink your milk, you become addicted to something or another.

It is also quite easy to make inaccurate predictions based on existing data because the projection curve (regression) was not appropriate for the data presented, or there was not enough data available to make accurate predictions, or outside forces were not taken into consideration. For example, pundits often use regression curves to predict the number of home microcomputers which will be sold on this continent during the next decade or by the next century. As a simple linear curve, we end up with some billion-odd computers sold by the very early 21st century. Nice outlook for investors and manufacturers, but so obviously unrealistic. No account is taken of a limited number of people to buy, no account of income restrictions, no account of market saturation. Unless, of course, it is intended that we all have three or four of them.

There is also danger in trying to interpret data for which the samples have been force-fitted or biased by some means or another. One Canadian university did a questionnaire on whether or not students were more sexually liberated than their parents, and their samples suggested a strong YES answer. However, doubt was cast on the validity of the study when it was learned that all of the sample people questioned were in the university's birth control clinic at the time! This adds a significant bias to the result, needless to say. Results of this nature were the cause of Benjamin Disraeli's exclamation: "There are three sorts of lie: lies, damned lies, and statistics."

Statistical analysis — an essential tool in science, industry and commerce — may be considered a manner of applying mathematical rigor to things we do all of the time, but about items for which we have no feel. Now, when the car salesman says the vehicle in question has only 2,000 kilometers on it and it was only driven by somebody's grandmother to church on Sundays, it's likely you won't believe him. You have applied a statistical hypothesis test to the data and



found it highly unlikely to be true. You may not think exactly that there is a 5% chance that he may be correct and a 95% chance that it's all hogwash, but you have come to an analogous decision nonetheless.

Statistical inference usually involves setting up a hypothesis (the "null" hypothesis) and testing it against an alternative hypothesis: the car actually was driven by a little old lady or A.J. Foyt is a grandmother. Hypothesis testing has a threshold value below which it is rejected. In our case above, we might accept the car if we felt the story had a 50% chance of being true. This probability of the truth where we make our cutoff is referred to as the level of significance. A null hypothesis is usually that the population mean equals a certain value (500 heads for example) or that the standard deviation is equal to a certain amount.

In another example, say we toss a coin 1000 times. We expect 500 heads to result, assuming the coin is unbiased (our null hypothesis). If we want to reject the hypothesis of a true coin with a 5% level of significance, then we would reject the coin as unbiased if we get less than 470 or more than 530 heads. Only 2.5% of the time would we get results less than 470 if the coin were unbiased.

A range on both sides of the expected value is a "two-tailed" test. We normally use a one-tailed test. For example if we bet tails, we would be pleased if we got too few heads, but would reject the hypothesis if we got 526 heads as unbiased towards heads. The probability of getting 531 or more heads is only 2.5% but 526 or more is 5% and 516 or more is 16%.

In the same manner, were you a manufacturer buying wire from another source, you might accept one defect every 1000 meters of wire. In a random test, you find three defects in that distance. Since you have no "feel" for wire quality variance, you would instead apply a test to determine the possibility of finding such a sample at random. If, when factors are entered, the result turns out to be, say, 40% or greater, you would believe that the sample fell within a statistically acceptable region. You would probably go back to another random sample (or even more) to apply this test enough times to get a reasonable database for your projections.

**STATISTICAL  
FUNCTIONS****Note:**

For those who are familiar with statistics, we apologize for our lack of mathematical rigour and exactitude. We have attempted to simplify a complex area for the uninitiated and in doing so have left off some of the more mathematical niceties.

**STATISTICAL  
ANALYSIS****T-TEST**

As you can see, there are no sure things. If we test at a 5% significance level, then we reject a hypothesis which is true once in every 20 times. This is called a type one error.

A type two error is harder to measure. That is accepting the null hypothesis when it is actually false: the coin may be weighted to fall tails-up more often but we don't know it. We might accept it at 470, assuming the mean was 500 but what if the mean was actually 400? The probability of a type two error can only be estimated if the deviation in mean acceptable is known. The best way to minimize this kind of error is to increase the sample size.

This part of B/GRAPH is perhaps the most difficult to explain to the uninitiated. While it's generally true that if you don't know what the function is, you don't need it anyway, you may discover that once you understand the purpose, you may be able to apply it to your data. We cannot present a complete tutorial on statistics here; we can only gloss over it and recommend that you read up on it on your own. A small bibliography is provided at the end of the documentation for that purpose.

One of the most important techniques to familiarize yourself with the statistical portion of B/GRAPH is to experiment with it. Try creating data files with data you know is related and then applying regression and statistical analysis to it to see the results.

The first part of the STATISTICAL ANALYSIS program (option 6 on the Main Menu) is the t-test. The t-test (or t-distribution test, also called the Student's t-test) is a test of the probability that two samples came from the same population. Populations are larger groups of things, people, data, numbers or the like. Samples are sections taken from a population from which certain observations are made and used to infer general information and trends about the population. We use the sample to get an estimate of the population mean (the mean is the linear average: the total of all of the elements divided by the number of elements).

The t-distribution is a family of curves, similar to the normal distribution curve (a bell curve) but flatter and more dispersed. The t-test based on this distribution is used to analyse whether or

**MEAN****T VALUE  
APPROXIMATE**

not your two samples share enough in common to actually come from the same population; or if it is reasonable to assume that the two samples have the same mean. The t-distribution is a bell-shaped curve symmetrical about a mean of zero. It has a standard deviation dependant upon the sample size. As the sample size increases, the t-distribution curve approaches that of a standard normal curve. The t-test assumes that each sample has a similar dispersion, or roughly put, range of data points. It is normally used for samples of less than 30 elements each. For larger sample sizes, we would use the normal distribution probability, below. The larger the sample, the closer the shape of the t distribution curve approaches that of the normal curve. With a t-test, the smaller the sample, the larger the difference between elements in a factor is demanded. A big difference is needed to establish a significance if the samples are very small.

FIGURE 15

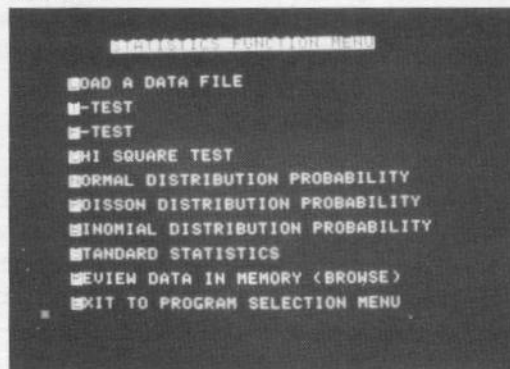


FIGURE 16

SAMPLE RESULTS	
MEAN	MEDIAN
7.274	5.062
VARIANCE	STD DEV
28.068	5.297
SKEWNESS	KURTOSIS
3.958	22.421
QUARTILE 1	QUARTILE 3
3.337	6.385
LARGEST	SMALLEST
47.800	4.700
TYPE <b>ESC</b> TO PRINT PAGE	
TYPE <b>DEL</b> TO RETURN TO MENU	

To examine how this works, use the Load command in the function menu to load the file PLANET, then press T. In answer to the questions on the screen, press 1 for test samples in memory. In a moment the computer will give you back an analysis of the closeness of the first two factors in your

## DEGREES OF FREEDOM

data file. The factors do not have to have the same number of elements in each; the program is smart enough to recognize "/"s (Not Applicables) and will treat factors with differing numbers of elements as having different degrees of freedom.

The screen shows the analysis for both factors. If you had only one factor in your file, you will not see these summaries on the screen. With one factor only, it will ask you for the value of the mean you wish to test. It will calculate the t-value and the percent above, mentioned further below. You can do a t-test for one factor as well as you can do a test for two.

The first statistic is "Degrees of Freedom". This is a somewhat confusing number, representing the number of classes or values in a set which can be taken arbitrarily. We use it to reduce the artificiality of the way we make the sample and the mean fit our estimate of the population. It is generally the number of elements in your sample less one for every statistic you calculate (in this case one, for the t-distribution) using the sample values. Since we have data points for ten planets (or nine and one asteroid!), we reduce it by one to get nine degrees of freedom. If our sample had 100 points, we would have 99 degrees of freedom here.

## VARIANCE

The next statistic is the "Variance". This is the square of the standard deviation, a measurement of the deviation of a sample about its mean, taken as the mean of the square of the deviations. For example, if we had five items with the weights 111, 114, 117, 118 and 120 grams, by adding them together and dividing them by the number of elements (five), we would arrive at a mean of 116. The items would deviate from 116 by  $-5$ ,  $-2$ ,  $+1$ ,  $+2$  and  $+4$  respectively. If we add up the deviations, we'd find that they logically add up to zero. So instead we square the numbers, thus disposing of the negative signs: 25, 4, 1, 4 and 16. Now if we take the mean of these new values, we get the Variance; in this case 50 divided by 5, or 10. To return to the same units as the original measurements or observations, we take the square root of the Variance; 3.16 in this case. If your numbers are very large and have a large dispersion between values, the Variance might become

too large for the computer to handle properly. In this case, you will get "/" (Not Applicable) for your answer.

## MEAN

The mean, our next number, has already been explained as the total of all values in a factor, divided by the number of elements. This is what most of us think of as an "average".

## T VALUE

Your t value is the number representing the chance that the two samples were from the same population. It has a range of any positive value but will most normally be between one and ten. The significance of this number depends on the number of degrees of freedom involved. The more the degrees, the lower the t value need be to establish a significant relationship between the two samples. As a general rule, the higher the value, the less the chance the two samples were from the same population.

## PERCENT ABOVE THIS VALUE

The "percent above this value" on the bottom line is the percentage chance that a given value will lie above the curve for the t-value; ie how often will you find a sample of the same size with a difference in means greater than the difference you have in these two samples. This is a single-tailed test; we are only testing one end of the curve, so the percentage is never larger than 50%. If it is very large, then there is a very good chance that the means come from the same population; the chance is smaller the smaller the percent. It's an inverse relationship with the t-value.

## TESTING SPECIAL VALUES

If you have no data file you wish to test, you can choose the second category on the first screen: Test special values. The program will go straight to this category if it finds nothing already loaded into memory. For this you must enter the number of degrees of freedom and a number for the Student's t value itself. In return you will be given the percentage of elements you would find above the computed values, as in the above example with a loaded data file.

The F-test, or F distribution test (after British statistician R.A. Fisher), like the t-test, tests if two samples came from the same population. However, it asks the question "are there significant differences anywhere among these samples?", and uses a process called the "analysis of variance". It needs two samples to work on, although it is most commonly used in a larger framework with three or more samples to work on (B/GRAPH can only work on the first two factors in a data file). It is a skewed distribution, starting at zero, involving a comparison of two variables. It is determined by its degrees of freedom and is usually formed by dividing two independent estimates of variance in order to test the equality of the variances.

Since you divide one sample variance by the other, you have two degrees of freedom: one for the numerator (the number being divided) and one for the denominator (the divisor). This leads to a great many possible curves. The distribution is not symmetric, so one cannot determine the left portion of the curve from the usual tables. To alleviate this, the larger variance is placed on the top when dividing.

You can still use the data in memory loaded for the t-test by pressing 1 or you can return to the menu and load in a data file. The F-test will provide you with the degrees of freedom, the variance, and the mean as before in the t-test; the same values if you use the same file or data. It will however, now give you a value for the F-test and a percent chance the data will lie above this value. Like the t-test, the factors you test do not have to have the same number of elements; if they have "/"s in them, they will be treated as having different numbers of degrees of freedom. A method of using the Regression Program as a replacement for the F test when you have more than two factors is described later.

## TESTING SPECIAL VALUES

Should you have no data loaded or wish to do a test of special values, enter 2. You must now supply the degrees of freedom in the numerator and again in the denominator. You will have to supply an F-value and the result will be the percent chance that a sample element will be above, or worse than, this value.

## 83 CHI-SQUARE TEST

Another distribution test, the Chi (pronounced "kye" as in sky) square distribution is related to the "t" distribution (actually it's the square of the t distribution), described above, using two factors for analysis; the actual data and the expected data. It is used to test the goodness of fit of two sets of data; ie how a result fits to an expected value. The value is the sum of the square of the actual minus the expected values, divided by the expected. In any Chi square, the mean is equal to the number of degrees of freedom in the sample.

The Chi square test is one of the most widely used tests in social statistics. The size of the Chi square value depends, loosely, on the sizes of the differences between the two factors (actual and expected) and on how many differences are involved.

To use this part of the program, you may load a data file with one or two factors (a third factor will be ignored). If you have only one factor, the program will ask you for the degrees of freedom for it then a Chi square value from which to test your data. You will be given the percent as above, for the data on hand and the degrees and value entered. The same questions will be asked of you for the second choice, test special values.

If you have two factors, then the program treats the first as the observed or actual results and the second factor as the expected. It will only ask for one degree of freedom and it will return the Chi square value as computed and the percent described above. Unlike the t-test and the F-test, both factors must have the same number of elements without "/"s.

Due to limitations of both the Apple and B/GRAPH, this program can only handle two factors; a two-column matrix.

Frequently a Chi square test is done on a matrix of considerably larger size, used as a test of the principles of classification. In order to overcome this, we have provided a separate program called CHI SQUARE SET UP which will allow you to write out a large matrix which is then turned into two strings for use in the Chi square function here. More on that program later.

84 **NORMAL  
DISTRIBUTION  
PROBABILITY**

Normal distribution is just that; it's the curve of distribution found in nature. It is also the basis of sampling theory, which concerns itself with establishing relationships between population distributions and those of the sample statistics.

If you pick a sample from a normally distributed population, it may not itself be normally distributed. However, an interesting theorem called the "central limit theorem" states that if we take repeated samples of the same size from a population, the means of the samples would themselves tend to be normally distributed even if the underlying population is not normal. It's the result of this theorem that allows relatively small samples such as the Gallup Poll to be accurate for a large population.

Normal distribution is the familiar bell curve used in mapping large populations and frequently scoring examinations in schools. The curve is such that 34% of all elements fit between the mean and one standard deviation to each side: therefore 68% of the total are said to be within one standard deviation. For example, if in a population, the mean of weights was 80kg and the standard deviation was 10 kg, 68% of the population data would fall between 70 and 90kg. Furthermore, 95% lie within 1.96 standard deviations of the mean and 99.75% of all observations lie within three standard deviations. The sample mean and standard deviation are estimates of the population mean and standard deviation.

The normal distribution curve never reaches zero at its ends, so between the mean and infinity is 50%. What this program measures is the percentage of the data which will fall between the mean you give it and the test value you enter. You must give it a standard deviation when asked, before entering a test value (for this test you don't have to load a data file; it works only on one test value at a time). For example, if we have a mean of 150 and a standard deviation of 15, what is the percent of the population which will lie between the mean and 180? The results will be 47.724. By simple arithmetic, this means that only 2.276% of the population will lie above this test value. You can manually convert your data to the standard normal curve by subtracting the mean from the test value and dividing by the standard deviation. In this case the result is two and the

value on the table is 2.28%. However, B/GRAPH was designed to do all of this for you without the manual arithmetic.

Generally, you use the normal distribution when you know the standard deviation and the test sample exceeds 30 and you have reason to believe that the sample standard deviation is a good estimate of the population standard deviation.

**POISSON  
DISTRIBUTION  
PROBABILITY**

Say you were buying guitar cases and had a contract which stated that you would accept 12 defects every thousand as reasonable. In your random test, you find 15 defects in 1000. Rather than run screaming to the manufacturer, you do a simple Poisson test to find the probability of discovering 15 defects when only 12 were expected.

When you choose P, you don't need to load a data file, merely answer the questions as to the expected frequency of events or values (12 in our example). Then answer with a test value (15 here). The result will be the percentage chance of finding exactly that many occurrences. Our answer is 7.239%. Usually we don't start screaming until we get a value of 5% or less. To see what our chances of getting say between 15 and 18, we would enter 15 as the lowest number and 18 as the highest. In some respects, this is similar to the binomial distribution, discussed below.

You use the Poisson test when you have very low expected frequencies rather than large ones. Remember, it works on an EXACT occurrence, not a value larger or smaller than your test frequency.

**BINOMIAL  
DISTRIBUTION  
PROBABILITY**

Have you ever played a game of chance such as dice where you wondered what were your chances of getting a certain result? What you wanted to know was the binomial distribution probability.

Let's look at a simple example: tossing a coin. By rights, you have a 50/50 chance of getting either side, since there are only two sides to a coin. This could be stated as a .5 probability of

success (in a range from zero to one). If you tossed it ten times (your number of trials), you might like to know what the probability of getting exactly five heads (successes they would be called) were. Using this function, we enter these values at the prompts and get a probability of success of 24.609% for exactly five successes.

But what does that tell us? What about five successes or less? If we are gambling, we may be interested in a spread of numbers. Well, if you do a bit of finger work, you will get the values below for a .5 probability of success and ten trials:

Successes:	
0 or 10 —	.098
1 or 9 —	.977
2 or 8 —	4.395
3 or 7 —	11.719
4 or 6 —	20.508
5 — — —	24.609

You might notice that there is a curve of data here, with five at the peak and all other value sloping away towards zero on each side. The total should add up to 100% (it adds to 100.003 due to the rounding to three decimal places). Without B/GRAPH, if you wanted to know the probability of getting five or less, then you would add together the probability values for the numbers less than your anticipated success rate, in this case five plus four plus three plus two plus one plus zero (.977 + 4.395 + 11.719 + 20.508 + 24.609 + .098) and you would get 62.306, or 62% chance of getting five or fewer wins.

B/GRAPH makes this somewhat easier; you supply the range of successes; the lowest first and the highest second. If you want the range for the test above, you would enter zero and then five. If you want simply the chances for one factor, enter it as both numbers (*ie* five and five).

When you have more than an expected of 35, you can use the Normal distribution rather than either Poisson or Binomial distribution. However, you need to calculate the mean and standard

deviations. For Normal distribution (a test of more than 35 expected successes), the mean will equal the number of trials times the probability of success, in our example  $.5 \times 10 = 5$ . The standard deviation is the square root of the mean times one minus the probability of success, in this case  $\text{SQR}(5 \times (1 - .5))$ . For Poisson distribution (where the probability of success is very small, for example one in one million), you calculate the mean in the same manner (probability times the number of trials) and the standard deviation is merely the square root of the mean.

## STANDARD STATISTICS

This portion of the program performs a number of common statistical functions and provides some of the most commonly used statistics, derived from your data files. It works only on data in memory from a loaded file, not special values you input. Standard statistics works on the first two factors in your data file. If there are two or more factors, it assumes they are grouped data; that is related in such a way that if you were, say, testing samples of smokers and the first elements in each factor were 17 and two, it would group them to mean that 17 smokers had two cigarettes rather than assuming you had samples of 17 and two smokers. If there is only one factor, then it is still treated as group data, but with a group size of one.

## MEAN AND MEDIAN

Ten statistics will be produced by this program, some of which will be familiar to you by now. The first two are the mean and the median. The mean is the linear average: the sum of all of the elements divided by the number of elements added. This is the popular "average" we often use. The median is also an "average" of a different sort. It is the number at which 50% of the observations lie above (and therefore below). For example, if we have the following test data in a file:

25, 22, 3, 4, 2, 3, 1, 5, 6, 3

We would see that the mean was 7.4, but the median was only 3.5: half of the observations lie below 3.5. If these were salaries in a company, in \$10,000 units and they boasted an average salary of \$74,000, you would ask for the median to see that the real half-way point was \$35K.

88 **VARIANCE AND STANDARD DEVIATION**

The next statistics are variance and standard deviation. Variance is the square of the standard deviation. The standard deviation was explained earlier; it is simply a measure of the variability of the data. For the sample above, the mean is 7.4. The observed data deviates from this value by:

17.6, 14.6, -4.4, -3.4, -5.4, -4.4, -6.4, -2.4, -1.4 and -4.4

The variance is the sum of these deviations squared, divided by the number of deviations added:

$$309.76 + 213.16 + 19.36 + 11.56 + 29.16 + 19.36 + 40.96 + 5.76 + 1.96 + 19.36 = 670.4/10 = 67.04$$

This number is the variance for that particular sample. Its square root, the standard deviation is 8.188. What we see on the screen for variance, however, is 74.489. This is the variance multiplied by the number of elements divided by the number of elements minus one ( $\times n/n - 1$ ) and the standard deviation is actually the square root of this value: 8.631. Why have we done this? Because we assume that the factors represent samples of a larger population as opposed to the entire population itself. So what we have done is improve the estimate of the population variance by "unbiasing" the variance with the multiplication above. The reason for this is best explained in a standard statistics textbook, but it is legitimate.

The effect of our "unbiasing" becomes less pronounced the larger the sample size. Multiplying a number by 100/99 gives a value closer to the original than multiplying it by 4/3 would.

**SKEWNESS AND KURTOSIS**

The next two statistics — skewness and kurtosis — relate to the nature of the curve itself. Skewness is a measure of how far the curve is skewed to the left or to the right. A positive number indicates a skew to the right, a negative number a skew to the left. The range is infinite but the larger the absolute of the number, the larger the skew. Kurtosis is a measure of the peakedness of the curve. A normal curve would have a kurtosis of 3. Less than that indicates a flatter curve, greater a

steeper curve. Our example has values of -.114 and 1.846, indicating a slight skew to the left and a flattened curve.

**QUARTILES AND RANGES VALUES**

Quartiles one and three are the points at which 25% and 75% of the observed data lie below, respectively. The median is actually the second quartile. In our example, the first and third quartiles would be 3.00 and 6.00.

The largest and smallest values are simply the two numbers in our data which represent the range of values; largest and smallest.

**REVIEWING DATA**

You may find it useful during your statistical functions, to check what you actually have in memory. For this, use the Review command from the function menu. If you only loaded one factor from a data file, you will still see two factors in memory; the first factor will simply show "one" for every element. This is because we use group data and in a one factor file, every group is given a size of one by B/GRAPH.

**AN EXAMPLE**

Suppose we have ten cars, examined on the assembly line for defects. The results are:

Car #	Defects	Car #	Defects
1	2	6	8
2	1	7	2
3	4	8	1
4	3	9	3
5	4	10	2
		Total	30

The mean or average is three defects per car, written as " $\bar{X}$ ". The median is 2.5, determined by

sorting the numbers in increasing order and determining the number for which 50% of the observations lie above and below. This is between the 5th and 6th car when sorted (half of 10). Now look at this table:

Car	D	$\bar{X} - D$	$(\bar{X} - D)^2$	$(\bar{X} - D)^3$	$(\bar{X} - D)^4$
1	2	-1	1	-1	1
2	1	-2	4	-8	16
3	4	1	1	1	1
4	3	0	0	0	0
5	4	1	1	1	1
6	8	5	25	125	625
7	2	-1	1	-1	1
8	3	0	0	0	0
9	3	0	0	0	0
10	2	-1	1	-1	1
Totals:	30	0	38	108	662

The first and second columns are the same as before. The third is the difference between the mean and observed value. The following columns are the third squared, cubed and to the fourth power respectively. The third column totals zero, since by definition, the sum of the differences about the mean must be zero.

The fourth column is used to calculate the variance: we divide the sum by 10 for a value of 3.8, getting a standard deviation of 1.95.

Skewness is calculated from column five, using the formula:

$$(108/10)/(3.8)^{3/2} = 10.8/7.40756 = 1.458$$

If the function was symmetric, this value would be zero. The value here indicates skewness to the right.

Kurtosis is calculated from the sixth column by:

$$(662/10)/(3.8)^2 = 66.2/14.44 = 4.584$$

The normal curve has a value of three for this measure, so this value indicates a more peaked curve.

## EXIT

To return to the main menu, type E. If you have data in memory, you will be given another chance to save it before returning, otherwise it will be destroyed.

# REGRESSION

## INTRODUCTION

Regression is the process of fitting a mathematical curve to a set of data points. In B/GRAPH, this works on the first two factors in a file; placing the first factor along the X axis, the second the Y. Regression will ignore "/"'s. If you have only one factor, then it becomes the Y and the X is a simple numerical increment up to the number of data points (ie 1,2,3, etc). You might use a one factor data file to project data for, say, monthly sales where the time series is just such an increment (time being an implicit variable with sales an explicit variable).

The data should have some relevant relationship to each other in order to make a meaningful curve. A factor with points randomly fluctuating between zero and 1000 does not relate well to a geometric progression from 900 to 2600. It can be graphed, but little information will be gleaned from that graph. Errors between the actual data and the fitted regression curve should be normally

FIGURE 17

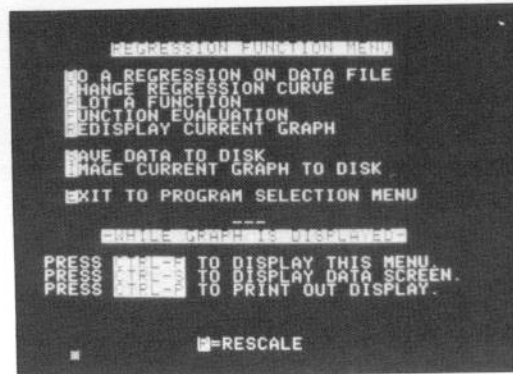
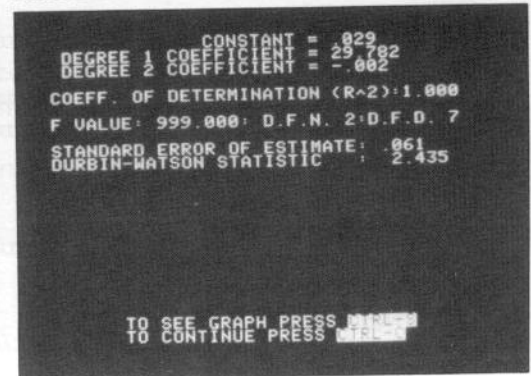


FIGURE 18



distributed and independent of each other and of the explanatory variables in the regression itself. Errors should also be independent through time.

The regression curve used by B/GRAPH is what is known as the "least squares fit". This means that the form of the curve is decided in advance (you input the degree of equation and the form of each degree; by which we mean how many non-constant terms and what form they take) and parameters are chosen so that the sum of the squares of the difference between the observed points and the fitted curve are at a minimum. When graphed, usually you will see that the observed points deviate from the fitted curve. This is due to such items as omission of variables in the function (such as unknowns, un-quantifiables, random or effects too small for reliable estimate), random behaviour (especially true where people are concerned), picking the wrong curve, poor aggregation of data during collection and errors in measurement.

In the next section you will have the opportunity of seeing how this is done by doing a regression on one of the sample files on your B/GRAPH disk. Before being able to do so though you need to be aware of certain operational concerns.

When doing a regression on a data file, B/GRAPH needs to read the data file, program files from the master B/GRAPH disk and also has to write some temporary work files to disk. For this latter reason you can not do a regression on a data file that resides on your B/GRAPH program disk as it is write-protected.

Before proceeding to the next chapter in this manual, please take a moment to copy the file PLANET and any other files that you wish to use in regression testing to a scratch disk. If you only have one drive you will be prompted by the program to swap back and forth between the B/GRAPH program disk and your scratch disk with the file being worked on.

To explain how all of this works, use the Do command and call up the file PLANET. We will do a regression on the distance from the sun and orbital velocities of planets in the solar system. Planets are numbered sequentially from Mercury to Pluto, with Ceres being number five. You will be

asked for the degree of equation when the file is loaded. This is an expression of the complexity of the equation which is in the form:

$$Y = A + B \times f(x) + C \times g(x) + D \times h(x) + E \times r(x) \text{ etc.}$$

where  $f$ ,  $g$ ,  $h$ ,  $r$  are functions of  $x$ .

Each degree is the part of the formula after the constant "A". "A" is decided by the program and is given to you in the information screen which follows the calculations. The number of degrees should be kept low in relation to the number of points in order not to compromise the predictive abilities of B/GRAPH. For this particular file, we happen to know the relationship of the data; orbital velocity decreases inversely as the square root of the distance. So answer the degree with 2.

Next you will be asked for the form of the first degree. Here you can enter any valid equation which uses the variable  $X$ ; ie  $\text{SIN}(X)$ ,  $X^{n^2}$  or  $X \times X$ ,  $\text{LOG}(X) - \text{SQR}(X)/2$ . A simple linear regression would use merely "X" in a one degree equation. A more complex regression might use  $X$ , then  $X^{n^2}$ , then  $X^{n^3}$  for the first, second and third degrees. The exact formula and form of the degrees will depend on the data. You must understand something about the data in order to curve it properly. For this example, input  $1/\text{SQR}(X)$  for the first form and simply  $X$  for the second. If you enter the same form twice, the computer will be unable to solve the regression and after attempting to do so, will return to the menu. This happens because an infinite variety of solutions is possible.

Sometimes you will not be able to find the right curve to fit the data. You may be trying the wrong equation for the data in your file. The data may be related to one another but the formula for that relation may not be obvious. It is worth noting here that if your equation involves raising a number to a power, ie:  $X^3$ , the calculation will be substantially faster if you use the form  $X \times X \times X$ .

## BUILT-IN FUNCTIONS

The Regression module allows the use of two built in variables; these are PI and I. To use PI in a formula that you are entering, simply type the letters 'PI' when required. The value of PI is calculated to eight decimal places of accuracy.

I is a variable which runs from one, up to the number of elements in your data file. This variable may be useful in situations where the first Factor is NOT a time series, but the nature of the Regression changes over time. Essentially this provides you with the ability to perform a three element Multiple Regression with I, the period counter, as the additional factor. For example, if you are performing a Regression on a two Factor data file, where the two Factors are SALES and POPULATION, by using I in your formula, you would essentially be adding a third Factor, that of a period counter, to the Regression analysis. Such a formula might be of the form  $\text{SALES} = 20,000 + (3 \times \text{POPULATION in thousands}) + 10,000 \times I$ .

Because the Regression module is not in reality designed for doing Multiple Regressions, you will find that the Plotting and Value Interpolation sections will not work properly if the variable I is used. (These features are discussed shortly.) However, the third Factor (Calculated Values) will be accurate. A true Multiple Regression module is scheduled for future release. Please contact Irata Press for details.

## THE DATA SCREEN

The screen blanks for calculations; this speeds it up considerably. When done, you are presented with a screen of pertinent data for the regression. This information will tell you first about the equation and then about the nature of the regression curve and the suitability of it for the data presented. If you have what appears to be the wrong curve or there is evidence that something may be amiss, you can return to the equation using **CTRL-C** to go to Function Evaluation (Interpolation), then to the main menu and finally **C** for Change to change the equation. Or you can go on to see the graph with **CTRL-S**, described further along in the tutorial.

## CONSTANT AND COEFFICIENTS

The first piece of information on the screen is the Constant. This is the number determined by the equations to be the "A" in the formula above. Next are the Coefficients of the degrees you entered, as they were computed according to the formula. You can have up to eleven degrees in your equation, but if you do not have sufficient data for such a complex equation, the program will

## STANDARD DEVIATION OF COEFFICIENTS

inform you and ask you to input a lower degree. The more data you have, the more accurate your regression. You can use the Constant and the Coefficients to plot a curve later on. This is discussed below under PLOTTING.

The numbers in brackets, after the coefficient, are the standard deviations of the coefficients. They indicate the significance of the estimate of the coefficient. Assuming that you have at least 8 more data points than you do degrees in the regression, the true value of the coefficient will lie within 2 standard deviations of the value calculated; 95% of the time.

For example, if you had a coefficient of 10, calculated from the data, and the standard deviation was 1, then you are 95 percent confident that the true value of the coefficient is between 8 and 12. If the standard deviation is, say 6, then you would have to use -2 to 22 as your confidence limit. In that case you could not rule out the possibility that the coefficient was actually zero and that the variable in question had no actual relation to the dependent variable at all.

Thus the usual test of coefficients is to exclude any coefficient where the standard deviation is more than half the absolute value of the coefficient. Of course as in all things, this rule is not hard and fast. A researcher may have other reasons for keeping a variable even if the standard deviation is large. Similarly, there may be good reason to discard a variable even when the standard deviation is low. However if you are qualified to make such judgements you don't need us explaining (hopefully in simplistic terms) what a standard deviation is.

## COEFFICIENT OF DETERMINATION

The Coefficient of Determination, known also as R-Squared ( $R^2$ ), indicates the proportion of the total variation of the dependent variable (Y) which can be explained by the changes in the independent variable (X). Only in the case of a linear regression is this coefficient the square of the correlation coefficient.

This value is a measure of the closeness of fit, rounded to three decimal places, of the curve to the data. In this example, it is 1,000, since we already knew the proper equation and had a

perfect set of observed data for our points. The range for this coefficient is zero to one; one being perfect, zero being terrible.

## F-VALUE

The F-Value is a statistic which indicates the overall significance of the regression and the fit of the data to the curve. Taken in combination with the degrees of freedom in the numerator (D.F.N.) and denominator (D.F.D.), it is an indicator of the validity of the relationship between X (the variable you used in the equation) and Y (the dependent variable). The total of DFN and DFD are one less than the total number of elements in a factor (in our case nine where we have ten planets). The higher the F-value, the better; B/GRAPH only prints a maximum F-value of 999; if the value is greater than that, only 999 is printed since a larger number isn't any more meaningful. You would compare this value against a table or use it in the statistics package of B/GRAPH.

You will be able to use the F-value in the F test portion of the Statistics package (6 on the main menu) when you either wish to test special values or you wish to test a three factor data file when the F test only uses the first two factors. Here is a simple example of sales calls made by three salespersons during a week. Able and Baker made four calls each, Carter made five for a total of 13 sales calls. The means for their sales are 300, 500 and 500 respectively. The results are:

### Size of Sales Made:

Sales call:	Able	Baker	Carter
1	300	600	700
2	400	300	300
3	300	300	400
4	500	400	600
5	0	0	500

The question is whether the average sales of these three are really different or merely due to chance. As mentioned in the Statistics section, B/GRAPH can not do this problem directly.

By assigning each salesperson a number — Able 1, Baker 2, Carter 3 — we can enter the data in two factors instead:

Sales:	Salesperson:
300	1
400	1
300	1
500	1
0	1
600	2
300	2
300	2
400	2
0	2
700	3
300	3
400	3
600	3
500	3

Now you can perform a linear regression on the data with one equation, form of degree = X. The data screen will produce the F statistic as well as the number of degrees of freedom. Not only will the regression program tell you if there is a relationship, but it will also give you an estimate of the form of the relationship, proving a more powerful test than the F-test alone. This is referred to again in the F-test portion of the Statistics package.

### STANDARD ERROR OF ESTIMATE

The Standard Error of Estimate is a number which indicates the relative deviations of the data points about the regression curve. The error comes in as the difference between your data and

what you have calculated. It can be any positive value, with zero being perfect. In our PLANET example, the value is extremely low: .028.

### DURBIN-WATSON

The Durbin-Watson Statistic measures autocorrelation; it tests the autocorrelation of errors, looking for a pattern, ie: is the error for point "t" related to the error for point "t-1"? If there is indication of autocorrelation, it often indicates that other factors are affecting the data that we are not taking into account.

FIGURE 19

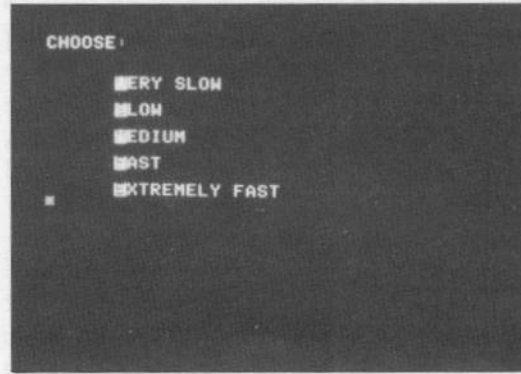
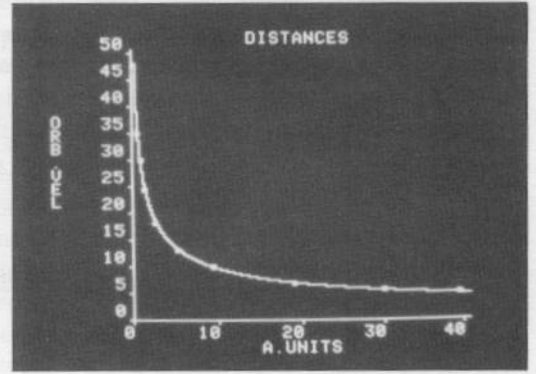


FIGURE 20



The values for the Durbin-Watson Statistic range between zero and four with four a perfect negative and zero a perfect positive correlation. Two is perfect. It cannot detect non-linear or lagged correlation. To properly interpret this value, you need a set of statistical tables. If there is evidence of autocorrelation, the equation should be used with caution for projection purposes or a different form tried. This might be in the form of a different series of equations or perhaps altering the order of

the factors using the file manipulation program. With small samples sizes, the Durbin-Watson statistic is of very limited value.

Autocorrelation is a form of correlation. In the Durbin-Watson test, you are correlating the data for the errors to themselves. You examine the error (difference between actual point and predicted point on the regression curve) at point "t" and compare it with another actual point at t-1 for their relationship.

## PRINTING THE DATA

The information on this screen may be printed by pressing **CTRL-P**. Once you have done this, choose **CTRL-S** to see the actual graph.

## DRAWING THE GRAPH

Before drawing, you are given a chance to choose the speed at which the graph will be drawn: Very Slow, Slow, Medium, Fast and Extremely Fast. The speed affects the accuracy of the graph itself; at Very Slow, each point (pixel) on the screen is calculated for the curve. At Slow, only every fifth point is plotted and a line drawn between them. It continues to the point where, using Extremely Fast, only four points are actually calculated. The use of this is seen on a simple linear curve where the graph plots the same at both ends of the selection. In this case, choose V. Later, you can go back to the menu and try different speeds by selecting the Change Regression option or by using Rescale and entering AUTO for a value choice.

The graph drawn shows the regression curve and the actual data points, plotted as small white squares or if you have more than fifty data points, small white dots. X and Y axis values are automatically calculated from the data but may be changed using Rescale as discussed below.

In this case the graph shows the distance from the sun in astronomical units (A.U.) on the X axis compared to a planet's orbital speed, on the Y axis. You may print this graph using **CTRL-P** or examine the data in your graph by using **CTRL-S**. If you choose the latter, you will see a third column titled "COMPUTED" (remember that regression uses only the first two factors and that any third factor is destroyed in memory). This is the computed value for Y using the regression

equation according to the formula you entered earlier. There will be a number of points shown corresponding to the data from the file you have loaded. The difference between the Y value and the computed value will give you an idea of the suitability of the regression. This screen may also be printed with **CTRL-P**.

If you press **CTRL-F** to call up the menu page, you may return to the graph from the menu page with Redisplay. A special "zoom" capability is built into regression with the R for Rescale, available when the graph is displayed. You will be asked to give the maximums and minimums for both Y and X axes. This will allow you to focus in greater detail or greater scope on the graph, even enlarging it outside the boundaries of the given data to interpolate the curve. Rescale also allows you to change the speed of the graph drawing by typing AUTO when asked for a maximum value for X.

## CHANGING THE REGRESSION

To see what we mean about the suitability of data and the equation graphically, return to the main menu and use Change to try another equation. This time enter a three degree equation and when prompted enter X,  $X^{n^2}$ , and finally  $X^{n^3}$ . Use a Very Slow draw again, for best accuracy. You will see that the curve emphatically does not fit the data, or vice versa.

To see another example, return to the menu and Load in the file ARM. This is a comparison of the arms expenditures, in \$ billions, of the developing nations and the developed nations between 1960 and 1979. A third factor, total global foreign aid, will be eliminated in the computation process. Answer the first question with two degrees, and use forms X and LOG(X) when prompted, finally using Very Slow for the draw. You will see a fairly linear growth curve with quite a number of data points on it, some on, some outside the curve. Use Rescale to zoom into a portion where the points are the most dense to examine their actual relation to the regression curve. Try a Y axis of 30 and 10 and an X axis of 165 and 85. Unless you type AUTO, the graph will be drawn at the original speed chosen.

## FUNCTION EVALUATION

Return to the function menu by pressing **CTRL-F**. With ARM still intact in memory, choose the **F** option from the function menu. This simply allows you to interpolate for the **Y** value from a given **X**. This allows you to examine specific points on the curve and to examine values that may lie outside it.

For example, enter 110. The **Y** value returned should be 352.889. This means that, given the regression we did above for this data, the arms expenditures of the developed nations will be at some \$352 billion when the developing nations have reached \$110 billion. At  $X = 500$ , the **Y** will be 1150.724. At 5000, the **Y** will be 9952.821 and you can see that the halfway point has been crossed. Hopefully, we will not need to see such incredible arms expenditures in our lifetimes!

You might try something different here; return to the function menu then to the main program menu and call up the **FILE MANIPULATION PROGRAM**. Load the file **ARM** again and go to **Change the order of factors**; replace factor one with factor 2 and vice versa. Save it under a different name and return to the **Regression program**. Call up your new file and, using the same equation, perform a regression. Now, when you want to examine the **Function Evaluation**, you will be working the other way around: you will be giving the \$ billions spent by the developed nations and seeing the interpolation of what the rest of the world will be spending. You should find the figures here quite close to those shown in the earlier interpolation example.

## PLOTTING A FUNCTION

Plotting allows you to plot two-dimensional functions and curves without an associated data file. Press **P** for **Plot** and then input any valid **BASIC** mathematical function for **Y** (see your Applesoft Manual for more information about **BASIC MATHEMATIC FORMULAE**). The formula can be as simple as one function ( $Y = X$  or  $Y = \text{SIN}(X)$  for example. You don't input the **Y**; simply the other side of the equals sign) or as complex as you can fit into one logical line, as long as it is valid Applesoft syntax. The only variable you can work with is **X**; any other will not give an accurate plot; instead an error message will be given and you will be returned to the formula input screen. It is not possible to use a formula like  $X + 1/Y$ . The exceptions to this are **PI** and **I**, discussed earlier.

For a simple graph, try  $\text{SIN}(X)$ . You will be asked for the maximum and minimum **Y** and **X**

values. Enter 10, -10, 1, -1 respectively. Choose a **Very Slow** draw routine at the next screen. To experiment, use the **Rescale** function to increase or decrease the **X** axis.

For a more complex graph, try:

$$.118 + 29.671/\text{SQR}(X) - .004 \cdot X$$

Use a maximum/minimum of 50 and zero for both **X** and **Y** and a **Very Slow** draw. If this curve looks familiar, it is basically the same as that drawn for the **PLANET** file. The formula simply uses the standard regression formula:

$$Y = A + B \times f(x) + C \times g(x) \text{ etc. (see the note on this equation, above),}$$

taking for **A** the constant from the data screen for the **PLANET** regression, plus the degree one and two coefficients and multiplying them by the forms of the degrees entered earlier ( $1/\text{SQR}(X)$  and  $X$ ). Press **CTRL-S** and examine the data; you will see that it fits quite well with the observed data in the **PLANET** file. Note that your equation is now the title across the top of the screen.

The **Plot** function is identical in every other way to the **Regression** function; it may be printed or saved, colors changed and rescaled. The **Plot** function is also quite useful for experimenting with curve types and various functions before entering them into a regression plot. Obviously, you do not get the data screen with coefficients and the like that you get when you do an equation. Remember that the graph is merely two-dimensional and works with only one variable: **X**.

When you call up the data screen using **CTRL-S**, you will get a number of points relative to the speed chosen for the draw: 100 for **Very Slow**, four for **Extremely Fast**. The maximum is 100; the first 100 points plotted. If you wish to see a greater range, choose the **Slow** or **Medium** speed which show fewer points but cover a wider range.

Plotting is very useful for creating graphs used in scientific applications. Many equations may be plotted using this function, such as exponential decay, Newton's Law of Cooling, the Michaelis-Menton model of enzyme saturation and others. We hope to produce special scientific engineering and finance application packages for use with **B/GRAPH**, in the near future. If you have mailed in your 'User Registration' card you will be advised of these updates as they become available.

# CHI SQUARE SET UP

This program is a utility for use in the Chi square test in the statistics package. Since our statistical analysis works on only two factors, this program allows you to input a larger matrix of data and store it in a format easily used by the Chi square test routine.

There is only one major function in this set up program: Create a new set of data. Using this, you will be asked to input a number of rows (they run horizontally from left to right on the screen), followed by an input of the number of columns (vertically, top to bottom on the screen).

You then enter data for each column individually for the total number of rows. Remember that the Chi square test allows no "/" values. Entry is identical to that in either graph program; use the cursor keys to edit it and press the **CTRL-E** key when finished. You will automatically be given the next column if there was more specified by your earlier input.

Once done, the computer will turn the data into two strings; actual and expected values which can be saved to disk using the Save function. Returning to the main menu by pressing Exit, you choose 5 to change the file from its ASCII format into a B/GRAPH format and then choose 6 for Statistical analysis and use your file in the Chi square test. We save the file as an ASCII file to allow you to use the file in other programs and to allow us to provide you this program in an unprotected, listable program, so you can see how to set up data files for use in B/GRAPH by your own programs.

For example, let's say we were doing a study of six baseball players on their hits, walks and outs and wanted to see if they appeared equal on these statistics. We would set up the following table:

	Player						
	A	B	C	D	E	F	Totals
Hits	10	7	13	8	14	8	60
Walks	5	5	0	1	3	6	20
Outs	28	20	30	20	10	12	120
Total	43	32	43	29	27	26	200

This would be entered into Chi square set up as three rows and six columns; the program itself will calculate the totals. It will also calculate the expecteds and saves them when you use the Save function, as factor two. Although you won't be able to see the expected data, it would look like this:

	Expected Player						
	A	B	C	D	E	F	Totals
Hits	12.9	9.6	12.9	8.7	8.1	7.8	60
Walks	4.3	3.2	4.3	2.9	2.7	2.6	20
Outs	25.8	19.2	25.8	17.4	16.2	15.6	120
Total	43	32	43	29	27	26	200

As you can see, the totals don't change. The Chi square statistic resulting from such a matrix is 21.36. The degrees of freedom is calculated by multiplying the (rows minus 1) by the (columns minus one) for a value of ten (you don't count total columns or rows for this calculation). You will be required to enter the degrees of freedom in the chi-square test routine. B/GRAPH will not tell you the value, but as can be seen the calculation is trivial.

When you run this using the Chi square test, you see that the probability of getting a value of 21.36 is less than 2%, so at a 5% level of significance (the percentage or greater at which you are willing to accept the hypothesis), we would reject the hypothesis that these players were equal on these statistics.

# CORRELATION ANALYSIS

This last program in B/GRAPH's statistical package allows you to test the correlation between two factors. The program ignores "/" elements and if only one factor is Loaded in, it becomes the second factor and the first is a simple increment of the ordinal numbers from one to the number of elements in the second factor.

With the Simple correlation function, you test how well the elements in the factor(s) correlate with each other, in a linear manner (in regression this would be a one degree equation of the form  $Y = aX + b$ ). You must Load a file first, since the program does not work with special values. You will be returned a Correlation coefficient when the computations are done. This value is the square root of the coefficient of determination, seen earlier in the regression program. It's faster than a regression if you have no interest in doing a full regression. It is used when there is a degree of linear relationship that appears to exist between the two factors; ie the Y factor moves upwards at a steady rate as a multiple of X.

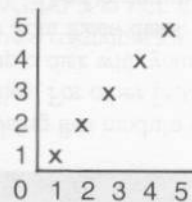
The range is minus one to plus one. Plus one means that the "Y" value increases directly with the "X". Zero means no apparent relationship, at least linearly. Points in a circle are related, but not linearly, so this is not a valid test for a circle. The danger is in interpreting the statistic for happenstance relationships such as an apparent correlation between murder and birth rates that are merely synchronistic rather than related (affecting each other). In North America, murder rates have followed an upward trend and birth rates have gone down, both rather steadily. A regression on this data will show an apparent relationship but they are in actuality unrelated. Factors may be correlated one to another on a lagged basis, which means that the data for one factor for time period T is related to the data for the other factor at time  $T - 1$  or  $T - 2$  etc.

The Rank correlation coefficient (also Spearman's test) is a simpler test; it requires two factors of the same size with no "/"s in them. The numbers must be ranks such as how you rated ten

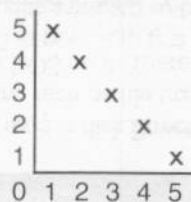
movies, on a scale from one to ten. Numbers may be duplicated (although they should not be for a proper test), but all must be in the range from one to ten. The second factor must be the same sort of rating but can, of course, be in a different order, say a ranking of movies done by a local reviewer. The Rank correlation coefficient tells you how closely the two rankings actually are.

The range is minus one to plus one, as in Simple correlation, with a positive number representing sameness of rank; the closer to one, the closer the correlations between the two factors.

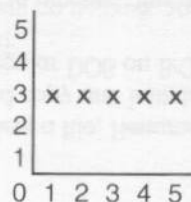
Here is a graphic example of what we mean by correlation, or how well the data points relate to each other in a regular manner. Each graph displays one type of correlation:



Negative



Zero (none)



Positive

**NEGATIVE:** Y values decrease as X values increase.

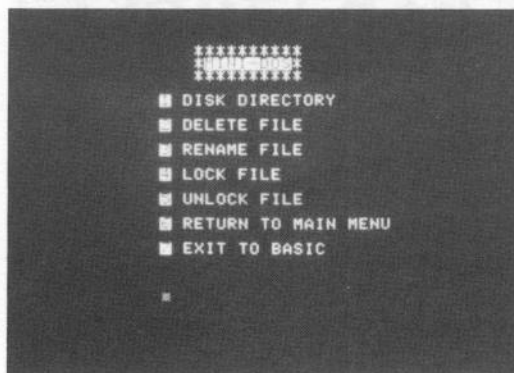
**ZERO:** X and Y values have no obvious linear correlation.

**POSITIVE:** Y values increase as X values increase.

**EXIT**

To return to the main menu again, type **E** for exit.

Mini-DOS (for Disk Operating System) is provided for the convenience of B/GRAPH owners so that they can perform the most common DOS functions without having to leave B/GRAPH.



Using this module you can read a disk Directory, Delete a file, Rename a file, Lock and Unlock files. For other DOS functions such as file transfer and copy and Initialization, you will have to boot up a disk with your regular DOS on it. There is no regular DOS on B/GRAPH, so you cannot boot it from these disks.

The Delete functions ask you to type "Y" in order to proceed; this allows you to escape from the sequence in case you change your mind.

This is a good time to remind you as well, of the importance of making back-ups of your data disks. You can always obtain another copy of B/GRAPH, but data files that you have built up over months and years may be irreplaceable. Always make a back-up of your data disk and store it separately from your working copy. Remember Murphy's Law: "Anything that can go wrong, will!"

Other Mini-DOS functions ask for a filename to proceed. Hitting **RETURN** will give you a disk directory, **CTRL-F** will abort the sequence. Error messages will appear if you try to do a task not

permitted, such as lock a file on a write-protected disk. You will be returned to the function menu. To return to the Main Menu, select 6.

You will note that in Apple B/GRAPH, 'RESET' does not allow you to enter BASIC. This is to prevent inadvertent alterations to the program code. If you wish to 'boot' another program without turning your computer off, you may select the EXIT TO BASIC function, at which point you will be asked to insert a disk for re-boot, and then press any key to proceed. B/GRAPH will be erased from memory and the RESET key will have its normal function returned.

A reminder on the Main Menu: selection **E** is for future enhancement disks which need B/GRAPH to run. Choosing this selection at present will do nothing. An enhancement disk needs to be available for anything to boot with this selection.

## THE EQUIPMENT

Supplies to be used with the

from the B/GRAPH program.

The B/GRAPH program is a

program that is used to

manage the data on the

disk. It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

program and is included

in the B/GRAPH program.

It is a very useful

# APPENDIX A

## PHOTOGRAPHING THE SCREEN

Although B/GRAPH has both screen dump and 'slide show' capabilities built in, you may also find it necessary or convenient to take color photographs of the various charts and graphs that you create.

Although there are services that will produce slides from your computer generated images or data, learning a few simple techniques will allow you to take color slides that will rival those produced by professional services.

### THE EQUIPMENT

We strongly recommend that you use a single lens reflex (SLR) camera. This is the type that allows you to view what will be photographed through the picture taking lens. Most people will probably use a 35MM model since slides produced can easily be projected, printed or viewed in a hand viewer. You will find that it is virtually impossible to take a properly focused and framed slide of your TV or monitor screen with a twin lens reflex or range finder type camera.

### LENSES AND TRIPODS

A lens with a focal length in the range of 50mm to 200mm will be suitable. A 105mm-135mm lens is ideal as the shorter focal lengths tend to accentuate screen curvature and distortion. If you are using a zoom lens, set it in the 150mm range, or its longest focal length. A firm tripod is a must. You can make do with a light weight model but if you have any choice use the heaviest one you can.

### ALIGNMENT AND FOCUSING

It is critical for best results that the camera be exactly at right angles to the picture tube. This applies to both axes; horizontal and vertical. You can do a pretty good job of judging this by eye, but the use of a small spirit or carpenter's level will help a great deal. You should move the tripod back and forth until you have filled the viewfinder with the full picture image. Remember that if you are producing 35mm slides, that the slide mount will cover about 10% of the corners of the film, so be prepared to experiment with your particular set up until you know what it will produce.

One of the most annoying results of a session of trying to photograph a monitor or TV screen, is to discover that some small point of light that you never noticed was reflected off the glass of the screen. The simplest solution is to not let that happen in the first place. Professional cameras are painted with black enamel for just this reason. If you have a chrome camera or tripod, try either placing some black tape or a dark non-reflective cloth over everything except the lens itself. If your screen image has a bright background, camera reflections can be a very real problem.

Naturally, the room should be totally dark, with all windows covered as best you can. Watch out for your eye glasses if you wear them. They are a wonderful source of screen reflection. Your best bet is to use a cable release with a long exposure and to stand well to the side rather than behind the camera. This will also help avoid reflections from watches and other jewelry as well as helping to avoid vibration.

### EXPOSURE

We suggest that you use an ASA 64 film such as Kodachrome or Ektachrome. This will give you very fine grain, although any daylight balanced color film is acceptable.

A color TV image is produced on your screen by the electron beam scanning across the screen 60 times a second. Due to the possibility of conflict with the shutter on your camera, you should use a shutter speed setting of longer than 1/8 second. If you were to use a higher shutter speed you will either get a partial picture or dark lines on your slides. Play it safe and use a shutter speed of 1/2 second for all of your tests.

Since you are not using any light source other than your screen itself and screens differ in

brightness, you will have to produce a series of tests. First, produce a few different graphs and charts that have a variety of combinations of colors and brightnesses. Three or four is sufficient, as you will have to run several exposures of each.

Set your TV or monitor for a good visual display, so that the screen image is at normal viewing brightness and contrast with the images reproduced by B/GRAPH, NOT from a regular TV show. The output intensity of your Apple II +/e may well be quite different than that of TV reception.

You will probably find that with a 64 ASA film and a 1/2 second exposure, an aperture setting of f5.6 will work well. Since as we said, your setup may be different than ours, you will have to bracket your exposures to find what's best for you.

Keep notes, and take at least six exposures of each graph, leaving everything unchanged except the f stop setting which should run the range from f3.5 to f8. in half stop increments. When your slides have been returned from processing, select the exposure that gives the most saturated colors. If, for example it is f5.6, then a combination setting of 64 ASA film, and a 1/2 second exposure at f5.6 will become your standard. It is always a good idea to bracket your exposures by a half stop on either side of your standard setting for each shot. This is regular practice by professional photographers and the slight additional cost is well offset by not having to redo a set up due to improper exposure.

Michael Reichmann

## BIBLIOGRAPHY

If your interest in graphs and statistics in particular has been sparked by B/GRAPH, we offer the following brief bibliography. Though there are a great many books on statistical analysis on the market, and we certainly are not familiar with all of them, here are some that we can recommend.

### *SALES FORECASTING*

David L. Hurwood; Elliot S. Grossman;

Earl L. Bailey

The Conference Board Inc.

New York

1978

### *THEORY OF ECONOMETRICS 2ND ED.*

A. Koutsoyiannis

The MacMillan Press Ltd.

London

1977

### MARKETING RESEARCH — METHODOLOGICAL FOUNDATIONS 2ND ED.

Gilbert A. Churchill Jr.

Dryden Press

1979

*STATISTICAL ANALYSIS 2ND ED.*

Samuel B. Richmond

Ronald Press

1964

*STATISTICS WITHOUT TEARS*

A Primer for Non-Mathematicians

Derek Rowntree

Pelican Books

Middlesex England

1981

*FACTS FROM FIGURES*

M.J. Moroney

Pelican Books

Middlesex England

1981

*USES AND ABUSES OF STATISTICS*

W.J. Reichmann

Pelican Books

Middlesex England

1981

*COMPUTER SIMULATIONS IN BUSINESS*

Hugh J. Watson

John Wiley &amp; Sons Inc.

Toronto, Canada

1981

# DISK FILE DESCRIPTIONS

## DATA FILES

**PLANET** is a three factor file suitable for graphs and regressions. It contains data for the distance in astronomical units, orbital and escape velocities in km/sec for all of the planets in order from Mercury out to Pluto. Planet five is actually Ceres; an asteroid included to show how well the curve actually fits the data. Source: *Dynamic Astronomy*, by Robert Dixon, Prentice Hall, 1980.

**ARM** is a three factor file for graphs and regressions, containing the amount, in billions of dollars, spent by the developing and developed nations in arms and foreign aid (world total expenditures), between 1960 and 1979. Figures for expenditures by Lebanon and Cambodia are not included after 1975 since they seriously deviate from the curve and/or were involved in civil wars which made accounting difficult, to say the least. Source: *World Military and Social Expenditures 1981*, by Ruth Leger Sivard, World Priorities, 1981.

**POP** is a pie chart of world population figures by continent, from the 1979 census. Source: *World Statistics in Brief*, United Nations, 1981.

**WATERL.PIE** is a pie chart comparing French and Allied casualties in the four battles between June 16 and 18, 1815 that comprised the Waterloo Campaign. Source: *The Campaigns of Napoleon*, by David Chandler, MacMillan, 1966.

**NAPWARS .1, .2 and .3** are casualty figures for selected battles in the Napoleonic Wars. ".1" is a bar chart comparison of total killed for French and Allied armies, ".2", a pie chart of Allied and ".3" a pie chart of French casualties. The latter two are useful in experimenting with Overlays in the line graph mode. Battles are listed by name in the pie chart versions. The total for these battles exceeds 800,000; the Napoleonic wars — all battles included — are responsible for more than 1,000,000 soldiers killed. This is merely the military casualties; civilian deaths, casualties in

marches, wounded who died later, etc. are not taken into account. In the battle of Berezina alone, some 30,000 non-combatants were killed! Source: Chandler, as above.

**JOBS** is the twenty year trend in the number of military personnel, teachers and doctors in the world. In 1960 there were 9,851,000 military, 6,850,000 teachers and 1,226,000 doctors in the 28 nations of the developed world. In the 113 nations of the developing world the numbers were 8,699,000 military, (10,628,000 in 1969, due mostly to the number in Viet Nam) 7,946,000 teachers and 442,000 doctors. Twenty years later, in 1979, the figures were 9,501,000, 9,600,000 and 2,330,000 for the developed world and 14,934,000, 19,800,000 and 1,220,000 for the developing world. The U.N. statistics book mentioned above shows that while in the "third world", illiteracy dropped from 2 to 10% (except in South and Latin America where it rose considerably), the literacy rate in the "western" world has decreased by only 1% at best (in Europe). Source: Sivard, as above.

**CARS.STS** is a test file used in the Statistics package, purporting to give sample data of defects in random samples of ten vehicles taken off an assembly line. It has no real world validity, it is simply used for the tutorial in B/GRAPH.

**DEMO1.DAT** is the tutorial file used in the B/GRAPH tutorial section to demonstrate how to Display, Save, Update and perform other functions with a data disk file.

## REFERENCE GUIDE

Almost every B/GRAPH function, in each of the program modules, is referenced through one of the many on-screen menus. The Tutorials for each program explain in detail not only what each Function does, but also the theory behind it.

This Reference Guide is intended as a quick guide to each of the major program functions, when the tutorial offers too much, and the on-screen 'HELP' menus, too little information.

### MENUS

**CTRL-F** will display the main program menu if pressed while a graph or data screen is displayed. It may also be pressed to return to the main function menu when checking a disk directory.

**R** will always Redisplay a current graph from the main menu.

**CTRL-E** is used to signify the termination of data entry.

### VIEWING DATA SCREENS

**CTRL-S** can be pressed at any time that a graph is displayed to show the data that was used to create the graph. **P** may be used to move forward a page (screen) if more than 20 items are to be displayed.

**B** for **Browse** and **R** for **Review** are used in some of the Statistical modules for examining data brought in from a disk file, and to see the results of calculations. These are noted in the main menu of the respective modules.

### PRINTING GRAPHS AND DATA

**CTRL-P** is always used when a graph or data screen is displayed to dump that screen to a printer.

**RETURN** is frequently used to signify that you wish the computer to proceed with the built in default values, if you do not wish to enter your own. An example of this is when entering graph labels. If no default value is available, the built in speaker will sound and you will be shown on the screen an example of the type of entry that is expected.

## GRIDS AND BORDERS

**G** and **V** will produce horizontal and vertical grids on a displayed graph. These are toggle switches and removing a Grid line causes the graph to be redrawn for neatness.

**B** produces a top and right margin Border on a displayed graph.

**F** will automatically fill in the area under each line in a Line graph, producing what is known as an Area Graph.

## SAVING AND IMAGING GRAPHS AND DATA

**S** is used in all modules to save the data and other information that was used to create a graph. **D** for **Display** is used to reload this data in the Graph and Pie modules, and **L** for **Load** in the Statistical module.

**I** is for Image, and this will save a displayed graph to disk in the form of a high resolution Image, or picture file. These can then be recalled in the Labeling module for display. This type of file does not contain any data, just the picture as displayed on the screen.

## CURSOR KEYS

In any of the graphing modules you may use the cursor control diamond, **CTRL W**, **A**, **S**, **Z**, to move the cursor around the screen. In the Labeling module these keys are used to position the on screen cursor to place text for Labeling.

## SPECIAL LABELING KEY FUNCTIONS

**CTRL-G** toggles the HELP-SELECTION screen, Image screen.

**CTRL-R** toggles Regular sized text.

**CTRL-W** toggles wide text.

**CTRL-B** toggles Bold text.

**CTRL-D** toggles Double sized text.

**CTRL-I** toggles Invert color.

**CTRL-Q** exits Labeling and saves the new Image to disk.

**ESC** toggles text entry mode.

## OTHER KEYS AND THEIR FUNCTIONS

**\$** is used in the input screen of the Graphing module to force a dollar sign to appear next to the 'Y' axes scaling, signifying that the data is in dollars.

**R** is for Rescale, and may be pressed at any time that a graph is displayed (except Pie Charts) so that the 'Y' axis may be rescaled.

**A** allows Point graphs to have a line drawn between points, Attaching them.

**S** permits the Switching between different types of Graphs while the current graph is displayed.

**O** will Overlay onto a currently displayed graph another graph from a data file on disk.

**C** is used to Create a new graph in both the Graphing and Pie Chart modules. Any data from a previous graph in memory will be destroyed.

These are the major functions that will allow you to begin using B/GRAPH. Combined with the on-screen menus and in depth tutorials, you should find yourself comfortable with the command structure quite quickly.

# REMEMBER

Anytime that a Graph is on the screen, you may press CTRL-F to redisplay the main Menu of that program module. Pressing Redisplay will then instantly bring back the graph. No data is ever lost unless you Exit the program module or Create a new graph. Make sure that you Save any graph first.

GRIDS AND BORDERS

SAVING AND LOADING COMMANDS

CURSOR KEYS

SPECIAL LANGUAGE KEY FUNCTIONS

## NEED HELP?

Having problems? Have you got questions about Batteries Included software? Call the **BI Customer Support Hotline**. We'll be glad to help you in any way we can. Because this is a **free service** we must ask that you **keep your questions brief and to the point**. Since a limit of 5 minutes of free service is permitted per call, prepare your query carefully and have pertinent materials (e.g. your manual) on hand when you phone; there will be an additional charge for calls over five minutes. Our hours are 9am to 4pm (est), Monday to Friday. Best time to call — mornings. Worst time to call — midday. If our lines are busy, please be patient. Do not call our switchboard; they cannot patch you through or take messages for customer support. If you would rather not phone, or your question or problem is not urgent, please do write. We'll be glad to send you a personal reply.

**Please note: customer support is a free service for registered owners only.** Please make sure you have sent in your registration card.



**BATTERIES INCLUDED**

**Customer (416)  
Support 881-9816**

30 Mural Street Richmond Hill Ontario Canada L4B 1B5