EVALUATION OF PRESENTATION GRAPHICS FOR THE AGRICULTURAL SCIENCES

Kathy Shelley

Follow this and additional works at: https://newprairiepress.org/agstatconference

Part of the Agriculture Commons, and the Applied Statistics Commons

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License.

Recommended Citation

This is brought to you for free and open access by the Conferences at New Prairie Press. It has been accepted for inclusion in Conference on Applied Statistics in Agriculture by an authorized administrator of New Prairie Press. For more information, please contact cads@k-state.edu.
EVALUATION OF PRESENTATION GRAPHICS FOR THE AGRICULTURAL SCIENCES

Kathy Shelley
Department of Statistics
Iowa State University

Abstract

Professional-looking text and graphic slides enable an audience to comprehend the main ideas of a presentation more quickly. With the advent of easy-to-use graphic software packages and the affordability of personal computer hardware to run this software, researchers may now prepare their own slides or transparencies. This paper describes basic graphic software design and offers criteria for selection of an appropriate software package for scientific research presentations. Comparisons between two prototype graphics packages, Harvard Graphics and SAS/Graph, are made on the basis of the following selection criteria: (1) basic software design, (2) available hardware, (3) output device drivers, (4) available statistical graphics, and (5) data import/export facilities. Graphic style is also addressed here with sample graphs illustrating a current popular theory of visual discrimination.

Keywords: Graphic Software Design, WYSIWYG, Visual Perception.

1. Basic Graphic Software Design

There are two basic designs for graphic software packages: (1) Data Coordinate or (2) WYSIWYG (what you see is what you get) systems. Because one of the major goals of this paper is to compare these two types of graphical design, software designed for the Apple Macintosh is not examined. Graphic packages for the Macintosh belong to the WYSIWYG category, whereas both types of graphic designs are available on the IBM PC/Compatible computers. The Data Coordinate design uses a programming language and Cartesian coordinates for graph and text placement. The WYSIWYG design uses icons or menus for selection of graphs and options with annotation being done directly on the screen.

Popular Data Coordinate packages include SAS/Graph, DISSPLA, Telegraph, Fortran, Pascal and other programming languages. Some of the more widely used WYSIWYG packages include Freelance Plus, Lotus Graphwriter II, GEM, Harvard Graphics, and Microsoft Chart. Some graphics packages such as Statgraphics are menu driven, but do not have complete on-screen annotation. An excellent overview of popular business graphics packages is provided by Howard and Kunkel (1988).
2. Evaluation of Basic Software Designs

An ideal graphics package would combine the two designs. This appears to be the future direction in both graphic and word processing software development. Currently, progressive word processing packages such as WordPerfect and Microsoft Word have the capability to import graphs, but are not able to display the graph and text simultaneously without going to a special "view" mode.

The main advantages of the Data Coordinate design are placement control and precision, the ability to draw any type of figure with supplied data coordinates, and data modification capabilities within the package. Frequently, this type of software can accept digitized figures, such as x-ray contours or map outlines, and produce corresponding graphs on a variety of devices. This data also may be combined or modified within the software for further analysis. The main advantages of the WYSIWYG design is the short start-up time to learn the package, the quick on-screen response for displaying changes, and annotation which makes it extremely easy to move and size text as needed.

The main disadvantages of the Data Coordinate design are the requirement of a working knowledge of the programming language, the need for data coordinates for annotation, and the somewhat slow response for viewing changes. Usually the program has to be recompiled before a modified graph may be displayed. The main disadvantages of the WYSIWYG design are the need to modify the data in other software packages, the often limited selection of statistical graphs, and the dependence upon high (EGA or VGA) screen resolution for accurate annotation. Many of the WYSIWYG packages have a few arithmetic data functions, but these are often inadequate for statistical examination of the data.

3. Other Selection Criteria for IBM PC/Compatibles

Other criteria include hardware requirements, available output devices, selection of statistical graphs, data import/export facilities, macros for automated production, and slide show capabilities.

In general, recommended PC hardware for effective graphics will include an EGA or VGA color monitor and a graphics card with at least 128K RAM to enable the software to display 16 colors simultaneously with good resolution. A hard disk is almost essential because most graphics packages require at least two megabytes of storage.

If one needs to work with available output devices, the user should make sure that the device is supported by the software before purchasing it. If one can buy new equipment, an affordable color printer with good quality is the Hewlett Packard Paintjet printer. A desktop laser printer is recommended for high quality black and white graphs. One problem with the laser printer occurs when large graph files are printed. Some large graphs will not print directly from the graphics package. A recommendation for this situation would be to export an HPGL file from the graphics package and then...
import this HPGL file into a word processor such as WordPerfect 5.0 or the new version of MS Word. If funds are limited, Hewlett Packard offers the DeskJet Plus which has often been dubbed the "poor man's laser". Many graphics packages now support these printers. If a color film recorder is not attached to the PC, it is possible to send graphic files via modem or mail to various slide producing services. At Iowa State, we have a Matrix QCR film recorder that is networked to the mainframe computer. We are able to produce slides on the PC and send them directly to the film recorder via the mainframe network lines for overnight processing.

Data analysts need a wide variety of statistical graphs to display their data. The most commonly available are 2-dimensional scatter plots, bar charts, pie charts, and multiple graphs per page. More recent requirements in statistical graphics include 3-dimensional graphs, boxplots, various forms of stem and leaf plots, colored maps, and contour plots. In general, these more recent graphs are not supported by most of the business packages.

A good graphics package must import data from other popular software packages. The most common of these include ASCII text files, Lotus 1-2-3 .WKS or .PIC files, and dBaseII .DBF files. It also is necessary for the package to export the graphs to be imported into other packages such as word processors. The most common export standards include Lotus .PIC files, CGM Computer Graphic Metafiles, and HPGL Hewlett Packard plotter files. CGM was intended to be the ASCII counterpart for graphics files but, unfortunately, is not as standardized as desired. Thus not all CGM imports and exports work as advertised.

A macro facility is very helpful for production jobs where similar types of graphs will be repeated with different data sets. There appear to be two major designs for macro processing: program code parameter passing and a memory system that records keystrokes. The first design is typically used in programming languages and the second design, first implemented in Lotus spreadsheets, is more common in menu driven software.

Slide shows are becoming increasingly popular for poster sessions and convention software demonstrations. A slide show contains all the text and data graphs and enables the presenter to order and automatically display them on a PC. A side benefit of this feature is the ability to print the entire presentation with one command for modification purposes. This is most commonly available in the business graphics packages.

4. Comparison of Two Prototype Graphic Packages

It takes a fair amount of experience using software packages before one can feel qualified to be able to compare the features of the packages. Because extensive experience has been gained in our department using SAS/Graph and Harvard Graphics, a comparison will be made between these two packages. SAS/Graph is a Data Coordinate statistical programming language and Harvard Graphics is a menu driven WYSIWYG package. There is a new product from SAS called SAS Assist which is menu driven, but it requires additional RAM beyond the usual 640K limit, thus narrowing its current user base. Figure 1 displays the evaluation criteria which have been
discussed in this paper and compares SAS/Graph and Harvard Graphics. Note that neither package directly supports Box Plots, but they are more easily obtained by constructing an annotate macro in SAS/Graph. SAS does have box plots in PROC UNIVARIATE, but the output is not in the high quality form that SAS/Graph offers. A serious shortcoming with Harvard Graphics is the limited number of data observation points and variables and also the lack of 3-dimensional surface plots. The most appealing feature of a WYSIWYG package is its ease-of-use. It is the continuing hope of statistical graphic package users that these two design types will be combined in the near future to simplify the graph making process.

5. Graphic Style

Current research in statistical graphics is involved with either static (presentation) or dynamic graphics. This paper has been concerned with choosing appropriate software to enable researchers to produce their own static graphs. Because many people feel that the quality of a data graph is open to personal interpretation, there have been many books and articles written concerning constructing "good" graphs and avoiding "bad" graphs. Tufte (1983) originated the phrase "chart junk" that could describe many of the graphs presented in today's newspapers. An ordering of visual discrimination tasks has been presented by Cleveland (1985) to provide guidelines for constructing meaningful graphs for data analysis. This taxonomy is currently being examined and modified by other statisticians and psychologists in an attempt to refine and clarify the visual decoding process. Cleveland presented the following visual tasks in order of increasing difficulty: (1) position along a common scale, (2) position along identical, nonaligned scales, (3) length, (4) angle, (5) area, (6) volume, and (7) color hue, saturation and density. Simkin and Hastie (1987) have attached conditions for applying Cleveland's principles by demonstrating that his hierarchy is task dependent in that it works best for comparing magnitudes amongst themselves, whereas angle assessment is easier than position or length when needing to estimate proportions of the whole.

Figure 2 illustrates the use of the first two visual discrimination tasks to give a picture of uneven light distribution in a scientific growth chamber, a consulting project of Dr. Ted Bailey, Iowa State University. "Position along a common scale" is used to compare the light across row and position with each shelf, and "position along identical, nonaligned scales" is used to compare the shelves with each other. Cleveland's hierarchy works well here because the underlying quest is to compare the magnitudes of light for each position and look for patterns. By using the adjacent bars, one can see that the middle positions appear to receive more light.

Figure 3 is a graph of Shelf 6. One can see that when one needs to estimate length in a stacked bar graph, the middle patterning is lost and it is more difficult to compare individual values for row and position.

Figure 4 can be used to illustrate why it is important to first ascertain what one is analyzing before applying hierarchies. The bar graph is most useful for comparing dollar values among the years. The pie graph would be more useful for estimating the proportion of annual dollar amounts compared to the total dollar amounts for the entire seven years.
6. Summary

Selection criteria have been presented to aide researchers in choosing an appropriate graphics software package for creating their own presentation (static) graphs. The Macintosh has been excluded because the purpose here was to be able to compare two styles of graphic software: data coordinate and WYSIWYG. SAS/Graph and Harvard Graphics were chosen as prototypes because of our extensive experience in using these two packages. The ideal graphics package would combine the design aspects from both of the basic styles.

Once one has learned how to use the software, then the researcher needs to concentrate on graphic style. Cleveland's hierarchy of visual discrimination was illustrated with agronomic data to show how this theory can be used for graphing applied research data.

References


Selection Criteria for Evaluating Graphics Software
Using SAS/Graph and Harvard Graphics as Examples

<table>
<thead>
<tr>
<th></th>
<th>SAS/Graph</th>
<th>Harvard Graphics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hard Disk Space</td>
<td>10-13 Meg</td>
<td>7 Meg</td>
</tr>
<tr>
<td>RAM</td>
<td>At Least 640K</td>
<td>512K</td>
</tr>
<tr>
<td>EGA or VGA</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Data Limitations</td>
<td>None</td>
<td>240 Obs. 9 Var</td>
</tr>
<tr>
<td>Text Graphs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bar Charts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Dimensional Line</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Dimensional Surface</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US State Maps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US County Maps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error Bars</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Box Plots</td>
<td></td>
<td>Annotate</td>
</tr>
<tr>
<td>Data Import</td>
<td>No CGM</td>
<td>Includes CGM</td>
</tr>
<tr>
<td>Data Export</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macros</td>
<td></td>
<td>Program Code</td>
</tr>
<tr>
<td>Slide Show Capabilities</td>
<td></td>
<td>Keystrokes</td>
</tr>
</tbody>
</table>

Figure 1

Distribution of Light in a Research Growth Chamber

Figure 2
Discrimination of Position

Discrimination of Length

Figure 3

Figure 4

Agricultural Exports Minus Imports 1980-1986 (Billions of Dollars)

Data Source: Statistical Abstract 1987