Development In Business Simulation & Experiential Exercises, Volume 18, 1991

AN EXPERIENTIAL APPROACH TO TEACHING DATA ANALYSIS UNSING MYSTAF: RATIONALE, PROCEDURES, AND RESULTS

Ronald S. Rubin, University of Central Florida

ABSTRACT

This paper examines the rationale for using a microcomputer in teaching the data analysis segment in a marketing research course and describes a statistical software selection process in which a particular package, MYSTAT, was selected. The author's experience is used as a basis to explore the features of the software and to discuss some of the useful ways to use MYSTAT to assist instruction, and finally report the students' experiences with this instructional approach. In conclusion, the experience received highly positive evaluations from the students, as measured through self-administered questionnaires at the end of the semester.

INTRODUCTION

Most marketing research instructors presuppose that students are less than enthusiastic of their course relative to other offerings in the curriculum primarily because of the mathematics and statistics, which encompass the data analysis segment of the course. How often have we heard students speak of "sadistics" or "statisticophobia" (Dillon, 1982)? Those teaching marketing research are well aware of how difficult it is to motivate the students. Differences in students' backgrounds coupled with the difficulty of the technical subject matter are contributors toward demotivating many of the students. They do not perceive, for the most part, that marketing research is a relevant field that will be useful in their careers. Many students cannot visualize and integrate the applicability of what they are being taught. Traditional methods of instruction such as lectures group discussion, case methods, and major term projects have lacked the ability to adequately familiarize the student with the subject matter.

Since 1987 this author has used a program entitled MYSTAT (Systat, Inc., 1988) to teach the data analysis segment to undergraduate marketing research students. In the present paper, this experience will be used as the basis to explore some of the software features that I found useful in MYSTAT for teaching data analysis, then to discuss some of the fruitful ways to use MYSTAT to assist instruction with its application in a marketing research course, and finally report the students reaction to this instructional approach.

SOFTWARE SELECTION: AN OVERVIEW AND SELECTION CRITERIA

Instructors today are faced with the problem of selecting software for their students from a myriad of packages that are available. Numerous reviews of microcomputer statistical software have appeared, designed to provide guidelines for determining the packages that are appropriate for the specific needs of numerous users such as might be found at entire university or business site (Raskin, 1989; Fridlund, 1988). However, it should be obvious that those packages designed for the professional researcher may be too complex for the first time user. On the other hand, it might not be obvious, until it is to late, that a package designed for a political scientist may not be adequate for students in a marketing research course. In the sections below, a number of criteria are noted that should be considered in selecting a particular statistical package suitable for teaching data analysis in an undergraduate marketing research course. The university computer services has two packages available, SPSS and SAS, for faculty and student use at the university's mainframe computer. They also have Statgraphics on its LAN. However, students have been frustrated with the extraordinary length of time it takes to learn to use these comprehensive packages with their complex command structure and format specifications (computer services offer courses to help students use these packages which last weeks into the semester), and with the myriad of practically every statistical function one can think of - and a few that you could not.

My objectives in teaching this section of the course were 1) to alleviate the students frustration in using a complex statistical package, and 2) to eliminate confusion as to the myriad of techniques available from which they had to choose an appropriate method for their situation; but at the same time have the students learn from the experience of performing a data analysis.

MYSTAT: AN OVERVIEW

SYSTAT, produced by Systat, Inc., is becoming an industry standard in social scientific empirical research. In recent reviews SYSTAT has consistently been one of the top-rated statistical packages for the professional researcher (Raskin, 1989; Fridlund, 1988, 1986). However, SYSTAT along with its competitors have developed packages that are cost prohibitive to allow instructional use of the software readily. In the mid-1980s, Systat moved to address this problem with the introduction of MYSTAT, an instructional version of the SYSTAT package. The criteria I used in selecting MYSTAT included the following (Most of these criteria are also suggested by Malhotra et al., 1987; Mentzer, 1988; Coburn, 1986, pp. 337-366; and Walker and Hess, 1984.):

Ease of Use and Ease of Learning

Students need not memorize a command language to input or analyze data, they need not even be familiar with a typewriter keyboard. All the possibilities are available for inspection and selection from the main menu (See Figure 1).

MYSTAT A PERSONAL VERSION OF SYSTAT					
>>>>>>>> >>	DEMO	HELP	SYSTAT	QUIT	<<<<<<< <<
EDIT	USE	SAVE	PUT	SORT	RANK
SUBMIT	OUTPUT	NOTE	FORMAT	MENU	CHARSET
NAMES	WEIGHT				
PLOT	TPLOT	HISTOGRAM	BOX	STEM	TTEST
STATS	TABULATE	PEARSON	SIGN	WILCOXON	FRIEDMAN
ANOVA	Covariate	CATEGORY	MODEL	ESTIMATE	

Figure 1 illustrates the ease with which the student can communicate with the program. MYSTAT uses English words as commands. The student types the command directly onto the screen when it displays the command prompt, . For example, a simple task might be to edit a particular file named Survey. The student would type:

Stats Age/ mean sd skewness mm max range [enter]

The editor will appear on the screen with the data points from the file Survey. Most of the time a command pretty well says what it does. At other times the name of the command itself is not sufficient enough to convey exactly what is to be done. Therefore, the student will have to enter additional information called specifications to describe exactly what he wants the program to do. This task is very easy for the student to accomplish, for a slash character *(I)* separates the command from the specifications. For example, if the student wants to run descriptive statistics on a variable named Age he would type the following

The use of a menu and English word commands makes this program very easy to use and easy to learn.

Cost per User

Unlike many sophisticated personal computer software packages used on a college campus which are typically licensed on an annual basis, MYSTAT is purchased by the students through the university bookstore for a nominal fee (at our university form $6.75).
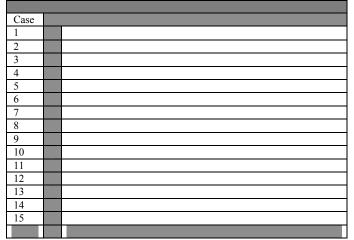
Data Entry and Management Routines

A major portion of any statistical routine is devoted to the input of data. Therefore, one of the most important features that should be examined in choosing a package is the flexibility of data entry options.

Most students will attempt to use procedures that they have developed when solving similar problems by hand. Therefore, the program should provide a data entry format that is similar to those used in the manual analysis of data. For example, if the screen layout can be viewed as a large sheet of paper with rows and columns similar to a "data sheet" used in many textbooks, data entry can be greatly facilitated. MYSTAT uses a format like a "spreadsheet." Figure 2 illustrates the editor used to enter the variable names across the top row and the cells in which you enter the raw data points.

Once the data have been entered and stored, the user should be provided with the capability for altering that data. Although it is not necessary to have all the features of a large package to transfer, data, the ability to perform simple

FIGURE 2. MYSTAT Editor

Case		
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		

transformations is essential. For example you may want to transform ratio data into ordinal categories. MYSTAT provides for this feature with the LET and IF... THEN LET commands and an array of functions and operators.

Error Handling Capability

Errors can occur in two ways, 1) during data entry, and 2) when trying to execute a command. To correct data entry errors the user can move to the cell, which contains the incorrect data point by using the arrow keys and re-typing the data point. MYSTAT permits one to move around the editor with the arrow keys and special keys just like a spreadsheet. This feature makes data correction relatively easy.

At times, one may enter a command or specifications incorrectly and MYSTAT will not understand what the user perceives it should do. The MYSTAT error handling routine will display:

* * * ERROR * * *

followed by a message such as: Unable to understand what you mean about here: followed by the line you typed. Looking at the caret (1 under your typed line will indicate where the error occurs. Knowing where the error occurs now permits one to reenter the command correctly.

Help Facility

Some type of help function is essential to the beginning user. Even after the manual has been read and the program demonstrated in class, the new user will find himself in a strange situation and not know what to do next. By typing HELP (command name) MYSTAT will then display a description of the command, a syntax description and one or more examples of typical uses of the command in question. For example, to learn about the USE command, type:

>HELP USE [Enter]

MYSTAT would display the following screen:
    USE reads a MYSTAT data file.

    USE <file>

    USE   MYDATA   (reads   from   MYDATA.SYS   in   default
        drive/directory)
    USE B:MYDATA (reads MYDATA.SYS frau B: drive)
    USE \SYSTAT\NewDATASYS (fully qualified names *must have*
        .SYS extension)

Selection of Statistical Procedures

I suspect that if one made a list of all the procedures that might be included in a statistical package that list would include several hundred options. Available procedures range from the common

measures of central tendency, t-test, correlation, etc.) to the not so common (Kolmogorov-Smirnov, Mann-Whitney U, Hotelling's T, etc.). Therefore, central to the choice of a statistical package must be some consideration of *its* scope, that is, what statistical techniques should it include. Here one will find differences that are basically a function of the intended audience. For example, economics teachers would most likely need packages which include time series analysis and seasonal indices while various types of analysis of variance (ANOVA) would be more required by the instructor teaching experimental psychology.

In most types of marketing research, the initial step of data analysis is to examine the characteristics of the distribution. Many marketing research studies go no further than summarizing data and cross tabulations; however, there may be studies that will employ other sophisticated analytical techniques. Therefore, all marketing research texts present many advanced analytical techniques, however, many of them are rarely used. Therefore, I have limited the techniques presented to those that have been reported to be used frequently by most firms (Greenberg, et. al., 1977; Naidu and Velu, 1989). MYSTAT covers all the basic statistical techniques that a typical undergraduate would be expected to know upon entering the business community.

The following is a brief description of the statistical procedures and the commands used in MYSTAT to analyze the data:

1. Descriptive Statistics - STATS provides complete descriptive statistics including the sum, mean, standard deviation, standard error of the mean, minimum, maximum, range, variance, skewness and kurtosis. You can even obtain descriptive statistics for subgroups by using the BY command if you first SORT the file by the grouping variable(s). To obtain the median and quartiles you would use the STEM command.

2. Graphical Data Analysis - The PLOT will create a two-way plot of one or more Y variables on a vertical scale against an X variable on the horizontal axis. The HISTOGRAM command will display a histogram for one or more variables. The BOX command creates a boxplot or grouped boxplot for one or more variables. The STEM command produces a stem-and-leaf diagram for one or more variables and prints the median and the quartiles of the variable(s). The TPLOT command plots a series of data values with the plot filled (shaded) from each point to the left.

3. Frequencies and Contingency Table Analysis -The TABULATE command produces frequency and n-way crosstabulations tables. For two-way tables, TABULATE provides Chi-square test statistics, association coefficients and PRE statistics with their asymptotic standard errors. You can display frequencies, row percents, column percents, or cell percents.

4. Correlations - The PEARSON command computes a matrix of Pearson product moment correlations. If you use the RANK command before PEARSON, the Spearman rank-order correlation is computed.

5. Independent and Dependent Group Tests – The T-TEST command permits one to do either dependent (paired) or independent t-tests. The SIGN command computes a sign test on all pairs of specified variables. The WILCOXON command calculates a Wilcoxon signed-rank test on pairs of variables. And, the FRIEDMAN computes a Friedman nonparametric analysis of variance on selected variables.

6. Linear Models: Regression, ANOVA and ANOCOVA -With MYSTAT one can use the MODEL and ESTIMATE commands to analyze regression models. Using the CATEGORY, ANOVA, COVARIATE and ESTIMATE command will provide fully factorial and non-factorial, and even unbalanced ANOVA and ANOCOVA models. If one uses the SAVE command with any model MYSTAT will perform extensive analyses of residuals.

Given the sophistication and number of statistical routines for the price and user-friendliness, this is truly an invaluable package.

Instructions and Examples

The manual consists of a 24-page booklet, which describes how to install and use MYSTAT, how to use the editor, how to analyze your data, how to use the graphics and statistics commands, and a section on additional features. Even though the manual covers a lot of information in a brief booklet, the only weakness I found with this program is with its documentation. As with almost any relatively new software, the documentation is not as "user-friendly" as it could be, particularly in the description of the use of the commands. There are not enough examples of how the commands and specifications can be used. The instructor will have to take it upon himself and explain and demonstrate how they can be used.

USING MYSTAT IN THE CLASSROOM

The introduction to the computer and MYSTAT is accomplished in two 50-minute classroom sessions. The instructor uses a laptop computer (Zenith 18OPC), a Sharp QA-SO Computer Projection Panel which displays images generated on the laptop computer onto a large screen, using an overhead projector as the light source to demonstrate the computer system and the program. During this first session students are introduced to the computer system (keyboard, logging-on, booting the program, etc. I, and running the program (entering the editor, print options, etc.). The second session is devoted to the program features. In order to teach the students the basics of MYSTAT and encourage them to experiment with the system, I introduce a hypothetical data set. It consists of six questions (representing different scales of measurement) and ten respondents. The class demonstration consists of:

1. Entering the variable names on the edit screen,
2. Entering the raw data,
3. Save data on the work disk when in the edit screen,
4. Leave the edit screen and return to the menu to use statistical procedures,
5. Using several statistical procedures,
6. Modifying data,
7. Exiting the program.

After the classroom instruction the students are assigned to a computer lab period in which they can practice what had been demonstrated. Students work at their own pace during this session using a printed tutorial that leads them in a step by step fashion through the exercises. The tutorial session begins with the student creating a data file that consists of ten observations on each of six variables.

After the data have been entered, it is saved to a disk file and then the students are led through an exercise that provides instruction in data correction (errors have been deliberately introduced into the original data), and then several analytical procedures.

In order to expose the students to a more realistic situation they are then assigned to analyze raw data, which had been taken from a study conducted by the instructor. Permission was granted by the company to use their name and situation in developing the experiential exercise for the students. In order to make the situation as realistic as Possible, the students are required to input a subset of the raw data points to give them the experience of data entry but without the necessity of long hours of typing on their part. Once the data have been entered the students begin to explore a number of the features of MYSTAT using this new data set. Specifically, the students are instructed to obtain descriptive statistics as well as frequency distribution and histogram for each variable. Scatter plots and contingency tables for two variables are also produced.

Because this assignment is given to the students during the last third of the semester only the more elementary statistical procedures are used to this point. However, once the students master the basic concepts of the software, they seem to have very little difficulty in moving on their own to more sophisticated procedures. Detailed instructions on using other statistical procedures, including t-tests, correlation, regression, and analysis of variance, are provided as those topics are covered in class.

EVALUATION OF THE EXPERIENCE

At the end of the semester, the students (class of 44) were requested to complete a self—administered survey on their experiences with MYSTAT. It is widely acknowledged that attitude is a multidimensional construct, and while no universal attitude model concerned with the effects of pedagogy has been developed, at least one study (Burns and Sherrell, 1982) has empirically refined Greenblat's (1973) original propositions concerning attitudinal effects. Burns and Sherrell identified five dimensions associated with experiential pedagogues; these dimensions were (1) perceived knowledge gained; (2) degree of enjoyment with the learning experience; (3) perceived benefits gained from the experience; (4) student-teacher relations; and (5) enhanced decision skills which the students believed they had gained. The attitude dimensions were measured by adapting Burns and Sherrell's items to the present study. Subjects responded to each statement by using a five-point agree-disagree scale. Student perceptions of this experience were overwhelmingly positive. An analysis of the responses are shown in Figure 3. The five dimensions of attitudinal effects mentioned above had Cronbach's alpha of .86, .76, .77, .57, and .62, respectively, considered acceptable for such exploratory research situations (Nunnally, 1978, p. 245). As seen in Figure 3, 27 out of the 31 measures had significantly positive means, indicating that the experience was viewed positively by the students.

CONCLUSIONS

This paper described the addition of a statistical package, MYSTAT, to an undergraduate marketing research course to enhance teaching the data analysis section. As indicated by the results of a survey administered at the end of the semester, the use of the MYSTAT package was positive and a rewarding experience for the students. The author urges those working with experiential pedagogues to consider the approach presented in this paper in their future teaching endeavors.

Figure 3. Attitude Dimensions and Item Used to Measure Them

Dimension|Statements[a]	Mean	t [b]	Cronbach's Alpha
Perceived Knowledge Gained.			86
I gained insight into the problems Associated with data analysis	4.25	11.46	
I gained insight into the pressures faced by decision makers.	4.13	11.83	
I gain actual information from the assignment .	3.95	8.81	
It aided my understanding of data analysis.	4.04	14.31	
It increase my awareness of the difficulties involved in data analysis.	4.31	15.5	
It increased my appreciation of the Uncertain ties faced by marketing researchers.	4.15	11.82	
It increased my appreciation of the problems faced by researchers.	4.18	13.43	
I learned the general principles involved	4.18	20.03	
I learned the procedures of data analysis.	4.11	14.85	
Enjoyment with the Experience.			76
The assignment was fun.	3.04	N.S.	
The assignment was boring. (reverse-scored)	3.68	5.41	
The assignment increased my interest in the course.	3.54	4.36	
The assignment was enjoyable .	3.11	N.S.	
The assignment increased my Enthusiasm to learn.	3.47	4.50	
The assignment increased my interest in the course.	3.47	3.9	
Assignments Like this teed to a more relaxed exchange between students and teachers.	3.15	N.S.	
Assignments Like this will make other course work more enjoyable.	3.38	3.01	
The assignment was interesting.	3.81	7.11	
Assignments like this lead to more student independence.	3.95	8.09	
Perceived benefits.			77
The assignment increased my sense of personal abilities.	3.88	10.06	
It helped increase my own self awareness.	3.63	5.83	
It increased my awareness of my own potential.	3.65	5.03	
Assignments like this will Iced me To participate more in related Class.	3.43	4.11	
It increased my interest in Learn-ing in general.	3.50	4.95	
Student-Teacher Relations.			57
Assignments like this help students to perceive teachers more positively.	3.06	N.S.	
Assignments like this help teachers perceive students more positively in general.	3.22	2.34	

Figure 3. (Continued)

Cronbach's

Dimension/Statements	Mean	t	Alpha
Student-Teacher Relationships (continued)			
Assignments like this promote			
Better student teacher Relationships	3.27	2.73	
Assignments like this lead to better			
Peer acceptance.	3.84	4.69	
Decision-Making Skills Gained			.62
The assignment was too low-level	3.83	14.36	
It changed my perception on some			
Parts of marketing research.	3.81	6.39	
I gained better decision-making			
Skills from the assignment	3.72	7.65	

a. Respondents rated each of the 31 statements on a scale of 1 to 5 with 5= strongly agree. 1 = strongly disagree and 3 = neutral.

b. Univariate t-test that the mean is positive i.e. greater than 3. Significant at the p <.01. N. S. = Not Significant

REFERENCES

Burns, Alvin and Daniel Sherrell (1982), "An Empirical Test of Microcomputer Simulation as an Alternative for the Teaching of a Marketing Topic," Proceedings: 1982 Summer Educator's Conference, American Marketing Association, 112-126.

Carpenter, Edwin H. (1985), "Statistical Packages for Microcomputers: New Vistas for Social Science Researchers in 1985," Social Science Microcomputer Review, Vol. 3, No. 2, 183-204.

Coburn, Edward J. (1986), Learning About Micro computers, Albany, N.Y.: Delmar Press.

Dillion, K.M. (1982), "Statisticophobia," Teaching of Psychology, Vol.9, 17.

Fridlund, Alan J. (1988), "Statistics Software Infoworld, September 19, 55-76.

Greenblat, Cathy (1973), "Teaching with Simulation Gates," Teaching Sociology, 1 (October), 62-83.

Greenberg, Barnett A., Jac L. Goldstucker and Danny Bellenger (1977), "What Techniques are used by Marketing Researchers in Business?" Journal of Marketing, Vol. 41, No. 2, 62-68.

Mentzer, John T. (1988), "Classroom Software May Also Aid Marketing Managers," Marketing News, 3(Summer), 73—80.

Naidu, G.M. and Raja P. Velu (1989) "Micros in Marketing Research: A Perspective From Users," Proceeding of the 1989 AMA Microcomputers in the Marketing Curriculum Conference, Editors, Robert F. Dyer and Margery Steinberg, Chicago, Ill: American Marketing Association, 193—200

Raskin, Robin (1989), "Statistical Software for the PC: Testing for Significance," PC Magazine, Vol. 8, No.5, (March 14), 103—310.

Rubin, Ronald S (1984), "Using the Microcomputer in Transferring Data Analysis Skills," Journal of Data Education, Vol. 24, No. 4, 23-26.

MYSTAT (1988), SYSTAT, Inc., 1800 Sherman Avenue, Evanston, Ill. 60201.

Nunnally, J.C. (1978), Psychometric Theory (2nd Ed.) New York, N.Y.: McGraw-Hill, p. 245.

Walker, Decker F and Robert D. Hess (1984), Instructional Software, Belmont, Ca.: Wadsworth Publishing Company