THE USE OF SAS IN AN INTRODUCTION TO INFORMATION PROCESSING COURSE

Robert Workman - Southern Connecticut State College

ABSTRACT

An introduction to SAS is taught as part of the sophomore level Information Processing course at Southern Connecticut State College. The students study a text with sections on computer technology, programming, systems, and computers in society. Interactive BASIC programming is taught the first half of the term and SAS the second half. SAS is presented as an example of a modern problem oriented software system. Its ease of use, good documentation, standardization, modularity, and continual updating are used to demonstrate good programming practice.

1 - BACKGROUND

Southern Connecticut State College is located in New Haven, Connecticut. It has an enrollment of 7,000 undergraduate and 3,400 graduate students. In early 1983 the school will become known as Southern Connecticut State University. Last year the Computer Science Department was accredited by the State of Connecticut to offer the Bachelor of Science Degree in Computer Science. There are currently over 300 Computer Science majors.

The college has its own DEC 11/70 computer with 18 VT100 terminals available for student use. A Harris remote job entry station is used to transmit batch jobs. There are also many microcomputers available including TERRAS, APPLES, and TRS-80's. SAS jobs are usually created on the VT100's using DEC's ED2 full screen text editor. The jobs are processed at the University of Connecticut on an IBM 3081 running under the MVS operating system. They are returned to the sender's terminal. Most of the academic year turnaround time is about one minute. Towards the end of the term turnaround lengths to about an hour.

2 - THE INFORMATION PROCESSING COURSE

Information Processing is a sophomore level course that was first offered in 1972. It has been taken primarily by economics majors who have completed a prerequisite course devoted almost exclusively to teaching the WATFIV version of structured Fortran. Next term, eight sections of Information Processing will be offered. The course's catalog description reads:

CSC230 Information Processing. Information processing and computer use as applied to the social and behavioral sciences. Students analyze and evaluate data through the use of software packages or computer programming of their own design. Lab work required.

The purpose of the course is to provide students with an intellectual and practical survey of the field of information processing.

A text is read that includes sections on computer technology, programming, systems, and computers in society. During the first half of the course students learn interactive BASIC programming and in the second half SAS. Whenever possible, techniques used and experiences gained as part of the programming assignments are related to material in the text. SAS's ease of use, good documentation, standardization, modularity, and continual updating are particularly used in the discussions on software development, programming languages, and structured design concepts.

Prior to using SAS students have learned to operate the VT100 terminals. Specific skills learned are signing on and off the terminal, running the password changing program, and using a utility program to change a file's protection code. The BASIC programming language assignments include writing and interactive dialogue programs, a bubble sort, and accessing a file that contains the class assignment and test grades.

3 - SAS ASSIGNMENTS

The first SAS assignment requires that the student transfer a file from the instructor's account to their own account. The file transferred contains job control instructions and a sample SAS program similar to the first program in the SAS Introductory Guide. Comment statements are added to identify the student and the job, and an Options Statement is added to set line and page size to the VT100 dimensions and to suppress SAS notes. Students are also required to add an example of the use of Proc Chart and to change the sample account number and password to one that they have been issued. They then leave the text editor and issue commands to transmit the job to the remote site for processing. When the job is executed and returned to our site it is displayed, using either a utility program or the text editor. If the results are correct, the student is required to change the file protection code to allow the instructor to inspect the completed assignment. The cost of running a correct version of this assignment is 36 cents.

Each step in the assignment is in itself trivial, but when taken together the assignment is challenging, frustrating, and rewarding for the beginning student. The major source of frustration is recovering from errors made while in the text editor. Patience, tenacity, and guidance are required. After the program successfully runs, most beginners, who have been exposed only to Fortran and BASIC, are delighted to see a language that is smart enough to produce column headings. It also plots and charts data, and is not too fuzzy. 
about spacing.

The various steps of this assignment are related to the appropriate parts of the Information Processing text. Students are made aware of the fact that they are interacting with an operating system when they sign on or modify job control instructions. They know they are working with a utility program when they copy files and change protection code, and communications software and hardware when they transmit their job. They know, too, that the text editor is an example of a word processing utility and, finally, that SAS represents an important class of problem-oriented languages.

The next SAS assignment requires the use of several procedures that are presented in the SAS Introductory Guide. This assignment is loosely structured. Students may use the class grade data they worked with in the previous BASIC assignment, or they may use current data that is of interest to them. They are particularly encouraged to use data from other courses they are taking such as marketing, accounting, or advertising, and to incorporate SAS-produced output into papers they write, using the text editor to merge text and SAS-produced output. The most commonly selected procedures used for this assignment are Print, Plot, Chart, Freq, Means, and Corr. Problems with Job Control Statements are minimized by requiring that the text editor's Include command be used to copy the statements from the first assignment. The intent of this assignment is to show how easily customized output may be produced once overhead chores have been done, when a powerful language is used.

For the final assignment students are required to use the SAS report writing feature to produce a report using the class grade data. The report shows test averages for each test for each section and final class averages. The SAS code closely follows the report example in the SAS Introductory Guide. This assignment serves as an introduction to report writing languages.

Assignments are accompanied by class lectures about SAS. The lectures, with few exceptions, follow the text's presentation. The chapter on Analysis of Variance is omitted because it strays too far from the introduction to the information processing focus of the course. The chapter on Procedure Chart is expanded to include pie and block charts, type means, percent, cumulative percent, and cumulative frequency. Option Midpoints is explained because invariably someone tries to chart data organized by month or day of the week. Another addition is the descending option for Procedure Sort. Finally, the Overlay Option, illustrated in the chapter on Regression Analysis, is added to the section in which procedure Plot is presented.

4 - FOLLOW-UP COURSES

After completing the Information Processing course, students at Southern have several options available for continuing their work with SAS. There is an undergraduate course called Computer Implementation of Statistical Methods, part of which is devoted to a study of the SAS User's Guide: Basics. There is also a graduate-level course, Computer Programming for Behavioral Scientists, that was designed for students in the Master of Science degree program in Research, Measurement, and Evaluation in Behavioral Science and Education. The course is primarily a study of SAS. Students are, of course, encouraged to use SAS in their own research, or as part of internships or independent study courses within or outside of the college.

5 - CONCLUSION

For many students the Information Processing course is their final undergraduate computer course. It is important that they be exposed to an organized, easy to use, flexible, and powerful problem-oriented programming language. They must not be left with the impression that when a common task such as sorting is required one writes a bubble sort in BASIC or Fortran, or consults with a System Programmer on the intricacies of a system utility sort. Furthermore, students should experience the ease with which tasks such as plotting, charting, and basic statistical computing can be done once systems related chores have been learned. SAS was selected because it is one of several problem-oriented languages available for our use, and of the group that includes SPSS, BMD, and EASYTRIEVE, it is the easiest to get started with. It has a good introductory book, produces attractive output, has integrated data utility capabilities, and is constantly being refined and expanded.

REFERENCES