The University of Tennessee at Knoxville offers several approaches to providing statistical computing skills to members of the academic community. Potential users may either learn these skills in a formal classroom setting for which three quarter hours of graduate or undergraduate credit is given, learn these skills in a less time consuming seminar approach for which no formal credit is given, or learn these skills mostly on their own by viewing videotapes that resemble the seminar material.

This paper contrasts the philosophy behind each of these approaches and examines the advantages and disadvantages of each approach. The lecture and material content used in each approach are also examined.

The University of Tennessee is a multi-campus, multi-purpose system of higher education encompassing all Tennessee. It is the state's official state university and federal land-grant institution. In the fall of 1979 a total of approximately 50,000 students enrolled on the five campuses of The University of Tennessee system, approximately 30,000 of which were in Knoxville. The University offers 170 degree programs at the baccalaureate level, 144 at the master's level, and 60 at the doctoral level.

The Department of Computer Science is organizationally located within the College of Business Administration and offers both BS and MS degrees.

The Department of Computer Science is organizationally located within the College of Liberal Arts and offers both BS and MS degrees.

The University of Tennessee Computing Center provides computing facilities and services for research, instruction, and administration functions for The University of Tennessee. The Users Services staff of the Computing Center is responsible for providing consulting and educational services to users. The computer mainframes at UT are a two IBM System 370's/Model 3031, an IBM System 380/Model 40 and a DECsystem-10 Model 1000.

The formal course in statistical computing is a 4000 level course which usually is taken for graduate credit. The teaching of this course is the responsibility of the Computer Science Department, and the course is usually offered quarterly.

The seminars on statistical computing are taught quarterly by the User Services staff of The University of Tennessee Computing Center. The Users Services staff is also responsible for videotaping portions of the seminar presentations and making these videotapes available to users throughout the university library system.

Each approach-formal course, seminar, and videotape—has been well received and appears to be meeting the various statistical computing needs of the university community.

Underlying these various approaches are two basic assumptions. The first assumption is that the student has an adequate understanding of applied statistics (nonparametric tests such as chi-square, Spearman Rho, Mann-Whitney U, Friedman's Analysis of Variance on Ranks, etc.) and univariate parametric tests such as the student t for independent or correlated samples, linear regression analysis, analysis of variance and covariance, bi-variate correlation, etc. If the majority of the students have an adequate understanding of multivariate techniques and if time permits, multivariate analyses such as factor analysis, multiple regression, canonical correlation, discriminant analysis, and multivariate analysis of variance might also be examined. No previous computer experience is the second assumption.

The formal course whose title is Computations in Statistical Analysis is the only one on campus devoted entirely to the teaching of statistical computing skills. The use of the computer to perform statistical analyses may be briefly examined in other courses but the intent there is support for class exercises.

The course typically meets twice weekly, one hour and fifteen minutes each time. The content of the course includes the following topics:
- Review of survey and sample design
- Instrument design to facilitate subsequent computer processing of the data
- Data collection, coding, and storage
- Review of statistical hypothesis testing and levels of measurement
- The Statistical Package for the Social Sciences
- The Statistical Analysis System
- A comparison of statistical software packages

The two statistical software packages, SPSS and SAS, are emphasized rather heavily in this course since approximately ninety percent of the statistical computing on this campus is performed in either one of these two packages. Of this ninety percent, approximately seventy percent are SAS jobs.

The following concepts and analyses are examined when discussing SAS:
- Introduction to SAS, SAS output, the DATA step, creating SAS Data Sets from Input Data
- Missing values, the language of SAS, the PROC step, PROC PRINT, PROC DATACHK, PROC MEANS, PROC UNIVARIATE, PROC SORT
- SAS expressions, SAS functions, action statements, working with SAS data sets, PROC PLOT, PROC FREQ, PROC CORR, PROC CHART
- Information statements, storing SAS data sets, PROC TTEST, PROC MEANS (correlated t test), PROC GLM (regression), PROC RSQUARE, PROC STEPWISE
- PROC ANOVA, PROC GLM (unbalanced ANOVA), PROC NESTED, PROC FORMAT
- PROC RANK, PROC STANDARD, PROC KSTATTEST, PROC CONVERT, PROC SPLIT
- PROC CONTENTS, PROC MATRIX, other PROC's of interest to the students.

A packet of SAS examples which illustrate the concepts and statistical analyses discussed in the course are provided the student. Considerable time is devoted to interpreting the results and discussing any implications. The student will typically annotate this packet during the discussion and subsequently use it as a reference.

Student performance is evaluated by several take-home exams which require the student to use those techniques and analyses discussed in class to analyze a set of data that has already been edited and placed on a disk file, and interpret the results. The student is also required to design, on a very small scale, an instrument to collect data (usually data of interest in the student's discipline), administer it to a random sample of students, code the data, and analyze it using statistical software packages of the student's own choosing. This exercise has become a real world experience that the student has found invaluable.

The course has been very well received as course evaluations and enrollment figures suggest. During the regular academic year, the course enrollment usually ranges from twenty-five to thirty. During last summer quarter, at which time other course offerings were minimal, the course enrollment was approximately fifty graduate students.

The student evaluations of the course indicate that the students find the course pertinent to their professional training, that the students would take additional related courses, that the students would highly recommend this course to their colleagues, and that the students found the course to be one of the most useful in their program of study.

The course evaluations also suggest that too much material is condensed into one quarter of time. Departments on campus typically do not want to release their students from the departmental core courses for more than one quarter to learn statistical computing skills. In some instances, only the student's appetite was whetted. In other instances, the student's background was not sufficient to understand the topic being discussed. Hopefully, in these instances, the student is motivated to take additional statistical related courses or at least understand the implications and seek additional assistance.

Plans are underway to hopefully expand this course into a two quarter series with the first quarter being primarily a batch orientation and the second quarter an interactive orientation. At the same time the existing material could be taught at a slower pace and new concepts added. More time could be devoted to such topics as data screening, simultaneous inference, and detection of outliers then is done now.

If the student is interested in learning statistical computing skills using SAS outside the formal classroom setting, the user can learn these skills in a five-day seminar which is taught by the User Services staff of the University of Tennessee Computing Center. The seminar usually meets every other day during the luncheon hour. A packet of SAS examples including assignments which illustrate the concepts taught in the seminar is provided the student as well as a computer account number which the student can use for as long as two weeks beyond the last day of the seminar. If the student does not have data against which to practice using the concepts learned, a set of data is provided.

A condensed version of the more frequently used features in SAS are discussed in this seminar as follows:

**Session I**
- Introduction to and overview of SAS
- Data input preparation, concept of a case
- Reference materials
- SAS statement rules
- JCL to execute SAS
- DATA statement
- INPUT statement
- PROC PRINT
- PROC MEANS

**Session II**
- Discussion of set of test data
- INFILE statement
- INPUT statement extensions
- TITLE statement
- LABEL statement
- PROC FREQ
- PROC PLOT
- PROC CORR

**Session III**
- Program statements
- Missing values, error checks, editing
- PROC GLM
- PROC SORT
- BY statement
- Data set subsetting

**Session IV**
- PROC TTEST
- PROC ANOVA
- File manipulation

**Session V**
- PROC GLM or PROC ANOVA
- PROC CONVERT
- Additional SAS procedures

Enrollment in this seminar, which is offered quarterly, has ranged from as few as four to as
many as forty-five with an average enrollment of seventeen.

The interested user may learn essentially the same SAS concepts taught in the seminar by viewing videotapes made by the User Services staff in conjunction with the University's television services for users of the UT Computing System. Most of the information on the videotapes would apply to other installations with similar computing systems.

The videotapes are distributed throughout the university library system and can be viewed by the student on a small television monitor or by a group of students on a larger television monitor in a conference room.

The videotapes are black and white, each one hour in length, and are available to other educational institutions at a cost of $50.00 per 1/2 inch reel-to-reel or 3/4 inch cassette or available to commercial installations at a cost of $125.00 for either reel-to-reel or cassette. The tapes are available for review by interested parties.

A packet of SAS illustrations including assignments and a computer number are made available to the student. The computer numbers are usually validated for a period of two to three weeks. A file of test data is also available.

The file of test data and print file of SAS illustrations can be sent to non-UT system installations on a magnetic tape for an additional charge of $25.00. The magnetic tape is available with either standard OS labels or no labels and 1600 or 800 bpi.

The contents of each videotape are as follows:

**Tape 1**
- DATA statement
- INPUT statement
- CARDS statement
- TITLE statement
- COMMENT statement
- PROC format
- FORMAT statement
- LABEL statement
- Transformation
- IF statement

**Tape 2**
- ARRAY statement
- DO statement
- PROC SORT
- PROC PRINT
- PROC FREQ
- PROC CHART
- PROC UNIVARIATE
- PROC TTEST
- PROC CORR
- PROC PLOT
- PROC GLM
- PROC ANOVA

**Tape 3**
- Discusses the results of those Proc's discussed on Tape 2.

The User Services staff is available to answer questions of students who peruse the videotapes.

The making of videotapes can become a costly venture, not to say exhausting for the individual appearing on camera. For example, in this case, the costs were $60.00 per hour for use of two black and white cameras, and a crew. Using color cameras would have cost $150.00 per hour. The original videocassette costs $22.00 and each duplicate $16.00. It is obviously more difficult to update the material when a new release of SAS is made.

Each approach—the formal course offering, the seminar, or the videotaping—attracts different individuals.

The formal course offering attracts mostly graduate students who desire a rather comprehensive introduction to statistical computing and who desire obtaining three hours of credit for learning such skills. Many students want these skills to appear on their official record or transcript. Those enrolled in this course are beginning to collect either their thesis or dissertation data and need to acquire computing skills to facilitate statistical analysis of their data. The course is also required for seniors in the undergraduate statistics degree program.

The seminar approach attracts graduate students, faculty, and staff who do not have the time to devote to a formal course. These individuals usually need to learn specific SAS skills within a short period of time. Some of the individuals only desire a review of SAS or need to learn how to execute SAS on the UT computing system.

The videotape approach attracts graduate students, faculty, and staff who missed the seminar course and cannot afford to wait until next quarter to enroll in the seminar. Additionally, the videotapes are particularly useful to individuals displaced from the main campus at sites where these skills are not taught or staff is not available for assistance. The videotapes are also useful to individuals who have an adequate understanding of other statistical software packages and desire to learn the more frequently used features of SAS. The utility of videotapes is that any individual can view and learn the material, provided the library is open, at their own discretion without necessarily having to seek additional assistance. The individual is not constrained by time conflicts, course or seminar availability, etc.

Whatever feasible approach we find best meets the needs of our students and colleagues, we should insure that we teach statistical computing wisely and teach it well. I believe that a student who
is a graduate student has indicated a commitment to learn at least a basic set of research and methodological skills. Computers are an integral part of most areas of statistics today and statistical computing is just one facet of these skills.