

The Effects of Integrating a Certificate Program into the Classroom Teaching Environment

Monica C. Jackson, Jun Lu, Mary Gray and Helena Solana

American University
Department of Mathematics and Statistics
4400 Massachusetts AVE NW
Washington, DC 20016, USA
monica@american.edu

Abstract

The significance of statistical software such as Splus/R, SAS, and SUDDAN in data analysis is the reason most statistics graduate programs require statistical software courses. Therefore, we experimented with the idea of incorporating an industrial certificate program in a Statistical Software course at a private university. While we are not promoting one software package over another, we chose the SAS certification program since there exists a comprehensive certification program. By completing national and local campus studies, we found our approach greatly benefited students and faculty. Students found statistical packages much easier to learn if they had sufficient prior computer knowledge. Almost all the students and faculty in our campus study approve of certification exams as part of statistical software courses.

Mathematics Subject Classification: 62-07

Keywords: Learning experience, Statistics Education, Software

1 Introduction

Because statistical software plays such an important role in modern statistical applications, many studies have encouraged the use of technology in statistics courses. The Guidelines for the Assessment and Instruction in Statistics Education (GAISE) (Aliaga et al., 2005) project was funded by the American Statistical Association to examine needs for college level statistics course. One of the six recommendations from GAISE is the use of technology for developing conceptual understanding and analyzing data. Chambers et al. (Chambers, 2000) investigated the ease of use of statistical software packages as essential

resources for the analysis of data. For example, in Minitab and SPSS, the user only needs to do "point-and-click" on an Excel-like interface to access most of the functions. The significance of statistical software in data analysis mandates that most statistics graduate programs require taking statistical software courses (Ben-Zvi, 2000; Jeske, Lesch, and Deng, 2007; Thisted, 1979). Students' statistics education is incomplete without appropriate training in statistical software.

The focus of the software training, however, may vary depending on the type of the program. Many programs that focus on Ph.D education consider R/Splus (<http://www.r-project.org/>) as their first choice because of its flexibility to customize functions and to write simulation programs. Master's level graduate programs, on the other side, may consider software that is popular in the application area in which the graduate students may soon work. The needs in statistical software are clearly different for Ph.D students and Master's students. Unlike Ph.D students who will usually stay in the program for 4 to 5 years, Master's students only stay 3 to 4 semesters. With the limited time and many other required courses, Master's students typically have only 1 or 2 courses that focus on statistical software applications. The question is: how do we prepare the Master's students in the right direction so that the statistical software course can benefit the students in their future career and the statistics community? To answer this question, we proposed to incorporate an industrial certification program into the statistical software courses for Master's students.

With Institutional Review Board (IRB) approval, we experimented with the idea of incorporating a SAS certificate program in our statistical software course at a medium sized private university. The goal of this curriculum development study was to include a certification exam into the classroom teaching environment, to identify difficult study topics, and to find approaches to improve the learning outcomes.

We chose the SAS certification program for three reasons. First, topics on SAS are part of our existing statistical software course, and our coverage has a large overlap with the SAS certification exam. Hence including SAS certification needed only minor adjustment of the existing teaching material. Second, SAS is known to be a powerful and comprehensive program that is widely used in various areas, such as clinical trials, finance, and economics. Third, SAS has a comprehensive certification program from entry level to advanced programming that is recognized worldwide. SAS also publishes detailed learning materials, both online and in hard copies, that assist people to prepare for the certification exams. In our classroom teaching, we used those materials as references. We closely observed the introduction and implementation of difficult topics and frequently got feedback from the students as they studied statistical software and prepared for the certification exam. In order to examine ideas

that might improve teaching, we also studied the teaching methods used at other universities through a survey.

The goal of our study was not to promote SAS nor determine whether SAS software should be incorporated in a statistics curriculum, but rather to examine how best to benefit from its having been selected as the "software of choice" for the course. Most applied statistics courses focus on some statistical package and teach good practice for data management, entry, analysis, and reporting using it. Appropriate incorporation of technology in a course is essential and should be developed in parallel with learning sound statistical subject matter. However, we explored the acquisition of certification as an additional component to a statistical course by examining the pros and cons of such an integration based on student feedback regarding the inclusion of a certification exam.

In the next section, we will introduce the approaches we took in our teaching practices and the difficulties we observed. Section 3 will be devoted to the findings from the nationwide survey. Some further discussion and the conclusion will be summarized in Section 4.

2 Motivation/Case Study

In an effort to increase the motivation of students in the Statistical Software course, we integrated the SAS Certification Program into our teaching environment for the first time during the Fall semester of 2006. The aim of the statistical software course was to give students a basic introduction to common software languages used for statistical computing. The course focused primarily on SAS under the Windows environment and secondarily on R/SPlus and SPSS. Students were given a basic foundation of the SAS Base DATA step which included creating, manipulating, and displaying data sets in SAS. Since this material covered the requirements that the SAS Base 9 Certification Exam tests (SAS Certification prep guide:Base Programming for SAS 9, 2006), it was only logical to provide an incentive for the students by giving the students an opportunity to sit for the SAS Base 9 Certification Exam in the classroom upon completion of the course. This opportunity was provided as an option to the students, not as a requirement for passing the course. To further motivate the students, the examination fee for the SAS Base 9 Certification Exam was reimbursed for two students with the highest grade average in the class up until the exam. The students were informed of this award at the beginning of the course.

The material covered in the course included the basic introduction to SAS topics such as importing and exporting data, creating and manipulating data sets, data step programming, creating summary reports using SAS procedures and identifying and correcting programming errors. The required textbook for

the course was *The Little SAS Book: A primer*, 3rd Edition by L. Delwiche and S. Slaughter (Delwiche and Slaughter, 2003) and the reference textbook was *Applied Statistics and the SAS Programming Language*, 5th Edition (Cody and Smith, 2005).

It should be noted that although the goals of the SAS certification and the Statistical Software course differ, there is a substantial overlap. The objectives for the certification include mastering skills in importing and exporting raw data files, manipulating and transforming data, combining SAS data sets, creating basic detail and summary reports using SAS procedures, and identifying and correcting data syntax and programming logic errors. The objectives for the statistical software class include becoming knowledgeable in SAS, SPSS, and R by learning how to analyze and manipulate data in all three packages. However, students intent on taking the certification exam spend most of their time focused on SAS only rather than learning the additional packages. Therefore success in the statistics course does not relate directly to success on the certification exam nor conversely.

2.1 Class Format

The statistical software course was taught in a computer lab, two days a week for one hour fifteen minutes per session. The first weekly session was used to cover the material using slide presentations and to carry out demonstrations in SAS, Splus, and SPSS when necessary, while the second session was used for an in-class lab project which required programming in either SAS, Splus, or SPSS. The lab assignments were due at the end of each lab class. Homework assignments were due the following week were given out each week as well. The students were also given two exams over the course of the sixteen weeks of the semester. Several opportunities were available for the students to receive help with learning the statistical software packages. The instructor utilized personal notes as well as the resources available in the SAS Academic Training Kit provided by the SAS Institute. The instructor also included guest speakers from the SAS Institute and from industry. In order to provide further assistance outside the classroom, the professor was available to meet with students during office hours which were held twice a week for a total of two hours. Additionally, one "virtual" office hour was held through the networked learning environment, Blackboard. Students were able to log-on to Blackboard from any computer and communicate directly with the professor and any other classmates currently logged on to Blackboard. Students could also seek assistance from a teaching assistant assigned to the course and from numerous statistical software consultants at the Social Sciences Research located at the university. The Social Sciences Research Lab (SSRL) provides support to students and faculty seven days a week for specialized statistical software such

as SPSS, SAS, STATA and the like, with expertise ranging from creating an effective research design to interpreting statistical results. All these measures were aimed at assisting the students in order to improve the students' success in the learning outcomes of the statistical software course.

2.2 Data on the Statistical Software Class

The course enrollment consisted of 15 students from different academic disciplines and academic levels. There were 8 graduate students, 6 undergraduate students and 1 alumnus who was auditing the class. Of the graduate students, five were from the statistics department, 2 from the MBA in Finance program and 1 from the Economics department. Of the six undergraduate students, four were studying statistics and the other two were from the public communication program in the School of Communication.

2.3 Methodology for the Campus Study

In order to accomplish the goals of this study the following methods were used:

Observe students learning SAS in the Social Sciences Research Lab (SSRL).

The teaching assistant for the course served as a statistical software lab consultant in the Statistical Sciences Research Lab (SSRL) computer lab staffed by professional staff and graduate students in various programs such as Economics, Mathematics and Statistics and the School of International Service, responsible for helping students with research projects by providing assistance with statistical software. Some of the tasks at the SSRL were to provide statistical software introductory tutorials to students from different academic backgrounds and experiences with SAS and to work one-on-one with students on projects and coursework involving SAS. On a weekly basis, common mistakes and problems were collected by the SAS-certified lab assistant who was also the teaching assistant for the statistical software course.

Provide additional assistance.

In addition to regular class sessions and office hours, students were able to meet independently with the instructor of the software course to work on improving programming skills. Students took timed practice exams and worked through problems from a preparatory guide (SAS Certification prep guide:Base Programming for SAS 9, 2006).

The teaching assistant assigned to the course assisted the professor in grading the weekly lab projects and offered assistance to the students outside the

classroom. Additionally, weekly homework assignments were given by the professor and mistakes were recorded. This enabled the teaching assistant and the instructor to identify the common issues students had when learning SAS. The teaching assistant also interviewed students who completed the course and took the SAS Certification exam.

Provide guest speakers.

As an additional motivation to encourage students to focus on learning statistical software, we provided two guest speakers. The first speaker was a representative from SAS who spoke about the functionality of the software. This speaker addressed the class during the first week of the semester. The students were not familiar with SAS at this time. However, they were given a general overview of many of the advanced features of the software that they would learn throughout the semester. This speaker proved beneficial since typically learning a new language can prove to be quite tedious at the beginning. Moreover, seeing the relevance of the software outside the academic setting proved useful.

The second speaker was a representative from industry who talked to the students about career opportunities. This speaker was not affiliated with the SAS Institute but did stress that his company preferred to hire students who had passed or at least attempted to pass one certification exam.

ACKNOWLEDGEMENTS. We are thankful for funding from the Mellon fund to complete this study.

References

- [1] M. Aliaga, G. Cobb, C. Cuff, J. Garfield, R. Gould, R.I. Lock, Guidelines for Assessment and Instruction in Statistics Education (GAISE) Report, (2005) *American Statistical Association*, [http:// www.amstat.org/gaise](http://www.amstat.org/gaise).
- [2] D. Ben-Zvii, Toward Understanding the Role of Technological Tools in Statistical Learning, *Mathematical and Thinking Learning* **2** (2000), 127-155.
- [3] J. M. Chambers, Users, Programmers, and Statistical Software, *Journal of Computational and Graphical Statistics* **9** (2000), 404-422.
- [4] T. Cody, J. Smith, *Applied Statistics and the SAS Programming Language* Prentice Hall, New Jersey, 2005.
- [5] L. Delwiche, S. Slaughter, *The little SAS Book: A Primer*, SAS Press, Cary, 2003.

- [6] D. R. Jeske, S. M. Lesch, H. Deng, The merging of statistics education, consulting and research: A case study, *Journal of Statistics Education* **15** (2007)
- [7] A. M. Liebetrau, Measures of association, *Quantitative Applications in the Social Sciences Series* Sage Publications, Newbury Park, CA 2007
- [8] A. Lizzio, K. Wilson, R. Simons, University students' perceptions of the learning environment and academic outcomes: implications for theory and practice, *Studies in Higher Education*, **27** (2003)
- [9] *SAS Certification prep guide:Base Programming for SAS 9*, SAS Press, Cary, 2006.
- [10] R. A. Thisted, Teaching Statistical Computing Using Computer Packages, *The American Statistician*, **33** (1979)

Received: February, 2011