Focusing in on Student Learning Outcomes: How SDA Helped Us Get Data into the Classroom

by Aaron K. Shrimplin[1] and Jen-chien Yu

Introduction

Over the past few years the Electronic Data Center (EDC) at Miami University Libraries has made it a priority to provide support to faculty who want to incorporate numeric data in their courses. To our thinking, too many students are uncomfortable using quantitative and numerical concepts to problem solve. New opportunities now exist for faculty to teach with numeric data and to provide hands-on data analysis in the classroom. Typically, data and statistical analysis are taught in methods and statistics courses. However, it is of little benefit to give students tools without showing them how these tools are useful for solving specific problems. The better approach may be to teach them quantitative skills in the presence of a classroom problem. If students are learning about income distribution and poverty, it might be of benefit for them to be able to develop testable propositions about the working poor, for example, and employ a statistical technique to test their hypotheses. Unfortunately, teaching students in this way is often problematic. There are often significant barriers to teaching with numeric data. Survey datasets are often very large and complex to use. Moreover, most teaching faculty do not have the time and specialized skills necessary to prepare classroom datasets.

This paper describes ways in which the EDC has tackled these challenges.[2]

A Window of Opportunity

Despite these barriers, we have been able to work with faculty to improve students' quantitative-reasoning skills. Before we highlight a few examples of this collaboration, it is important to call attention to some of the factors that have made this partnership possible. Here is a short list of influences we think have supported the use of numeric data in the classroom.

Re-inventing Ourselves

Advances in information technology have given librarians the opportunity to re-invent themselves and grow professionally. As more library services are technologically-enabled, libraries are not only able to reshape and extend existing services, but also to create new services and products. In some cases, library activities are being converted to technological solutions, freeing up resources to focus on new activities. In short, the networked information revolution has afforded, if not necessitated, libraries and their staff the opportunities to step outside their traditional service boundaries and be proactive to the needs of the academic community. It is within this context that we have been able to work alongside faculty in making a direct contribution to student learning.

Active Learning

'Sage on the stage' is out and 'guide on the side' is in. This concept of learning focuses attention on the student's experience. Faculty who adopt this approach to learning and teaching tend to focus on students' learning outcomes and competencies. This approach, not only changes the way in which faculty teach, but also the way in which students learn. That is, rather than passively accumulating knowledge, students are asked to approach the learning experience differently; they are asked to become more actively- involved in the learning process and to demonstrate their ability to use the knowledge that they have acquired. In this environment, librarians who speak with faculty and specifically address the learning outcomes important to student success are rewarded with a more active role in the learning process. Opportunities then present themselves for the integration of quantitative reasoning skills into the curriculum.

Innovative Tools

Web-based tools have made it possible for students to query large datasets in ways not possible a few years ago. Data analysis and retrieval tools like Nesstar, SDA, and DataFerrett have opened up access in unprecedented ways.[3] While making effective use of numeric data in teaching and learning requires specialized skills and a considerable amount of time for preparation, barriers which inhibit the use of data in the classroom and in student projects are being lowered.

Collaboration

Recognizing that the conditions were right for a proactive approach to working with faculty to get numeric data into the classroom, we went on a listening tour to better understand the types of data support faculty require in

order to teach quantitative skills. As we talked with faculty one-on-one, it became clear that a partnership was in the making. Although the vocabulary may have differed, it became obvious that we shared similar points of view: that statistical literacy is important and that students need to be aware of datasets and how to use them. It also became clear that we were two sides of the same coin. Faculty had the subject expertise and learning objectives; we had the knowledge of datasets and the tools to explore them. Most of the faculty we spoke with wanted similar things. They were looking for ways to:

- marry theory and method in an active learning environment;
- make tradition courses more analytical;
- explore innovative pedagogies, such as creating self-directed learning paths;
- improve access to difficult-to-use datasets with a user-friendly interface;
- add new, exciting features to existing courses.

At times it felt as if we were writers making a pitch for a sitcom. Time was short and we had to make a lasting impression. Our secret weapon was the web-based tool developed at Berkeley called Survey Documentation and Analysis, otherwise known as SDA. In a few minutes, we could show a professor a customized HTML codebook and a user-friendly interface for basic data analysis and exploration. We often demonstrated a dataset important to his discipline. Professors were impressed with SDA for the same reasons it had impressed us. It could be run with a Web browser, it made data analysis accessible and available to students, and there were no proprietary, platform-specific software that needed to be installed. Moreover, students could investigate a basic question using cross-tabulation or another technique without having to learn a high-level language. For those situations where a statistical software program like SAS, SPSS, or STATA were appropriate, SDA let users download customized subsets with data definition files for these programs. Discovering variables through customized HTML codebooks was also very useful, particularly for large datasets.

Examples
Over the last two years, the EDC at Miami University Libraries has delivered approximately three dozen datasets to the Web using SDA. The majority of these datasets have been for three departments: Gerontology, Economics, and Political Science.

Gerontology
The EDC developed a webpage to supplement in-class instruction for the course *Linking Research and Practice in Gerontology*. The course is designed to give graduate students an overview of social research in gerontology. Students review the basic assumptions, models, processes, and problems of social research. They also take a more detailed look at applied research, including Program Evaluation. Using our webpage, students learn about existing datasets and gain hands-on experience by performing exploratory analysis and extracting subsets of observations and variables. The subsets are then downloaded for more in-depth analysis using various statistical packages. The following US-based datasets have been delivered to the web with SDA and are available for use:

- The Second Longitudinal Study of Aging,1994-2000: Wave 3 Survivor File;
- The Second Longitudinal Study of Aging, 1994-98: Wave 3 Survivor File;
- The Second Longitudinal Study of Aging, 1994-98: Wave 2 Descendent File;
- Evaluation of Long-Term Care Initiatives in Ohio;
- Longitudinal Study of Aging, 1984-90;
- National Long-Term Care Survey, 1994;
- National Nursing Home Survey 1995, 1997, and 1999[4].

Economics
The bulk of our work with the Economics department has been centered on the Current Population Survey's (CPS) Annual Demographic Files[5]. To date, we have created datasets for the years 1997 through 2001. Access to these public-use files has always been difficult. Perhaps the biggest obstacle to using these data is the sheer size of the surveys. For example, the Annual Demographic File for 1998 has over 650 variables and over 135,000 individual respondents. A student wanting to use these data to investigate a hypothesis would have to download the raw data file, write a data extraction program to obtain the variables needed for the research, and then harmonize the data definitions for different years. We used SDA to deliver these data files to the Web to improve access and to support faculty who wanted to incorporate hands-on data analysis into their undergraduate courses.

One of the key skills that students should learn in social science classes is the ability to think critically about the ways in which raw data are processed to support analysis. The CPS Annual Demographic Files are the most commonly employed dataset for social science analysis of income distribution, labor force behavior, and poverty in the US. In particular, these data files contain data for many of the standard government reports on these topics. The problem is that typically students are presented only with analyzed results, despite the fact that many of the key issues of analysis are reflected as much in what measurements one undertakes as in how those

measurements turn out. SDA lets students ask their own questions of these data using simple statistical analysis and tabulations.

Political Science
A number of different datasets have been created with SDA to support political science courses. In most cases, these datasets have served to make introductory courses more analytical while promoting active learning for students. A few sample exercises illustrate how students of politics are using SDA to gain a glimpse into the world of social science research:

- We discussed presidential approval and the increasing importance of the president to be an effective economic manager. What percent of respondents who favored Clinton's economic performance approved of his general job performance?
- We discussed the concept of political socialization. What are the agents of socialization and how has socialization changed over time?
- Are people less interested in politics today than in the past?
- Political scientists have noted emerging "gender gaps" over the past several decades. Examine the change in party identification among women and men over time.

Conclusion
We have enjoyed stepping outside our traditional service boundaries and working alongside faculty in making a direct contribution to student learning. Whether it's making a gerontology graduate course more quantitative by providing online access to core datasets or making it possible for an economics professor to let his students ask their own questions of a dataset, the EDC will continue to make access to numeric data more easy to use by both faculty and students and is committed to supporting faculty who are interested in incorporating data analysis into their courses.

Notes
[1] Contact: Aaron K. Shrimplin, Miami University Libraries, Oxford, Ohio 45056. Phone: 513.529.6823. Internet: http://www.lib.muohio.edu. Email: aaron@lib.muohio.edu

[2] This paper was presented at the IASSIST Conference held in Ottawa in May 2003 in the session on "Advancing Research and Data Literacy: Empowering Users".

[3] Survey Documentation & Analysis (SDA) available at: http://sda.berkeley.edu; Nesstar (Networked Social Science Tools and Resources) available at: http://www.nesstar.org; DataFerrett: For TheDataWeb available at: http://dataferrett.census.gov/TheDataWeb/index.html; Internet, accessed 22 March 2005.

[4] E.g., U.S. Dept. of Health and Human Services, National Center for Health Statistics. *The Second Longitudinal Study of Aging, 1994-2000: Wave 3 Survivor*

File [Computer file]. Hyattsville, MD: U.S. Dept. of Health and Human Services, National Center for Health Statistics [producer & distributor], 2002

Scripps Gerontology Center. *Evaluation of Long-Term Care Initiatives in Ohio* [Computer file]. Oxford, OH: Miami University, Scripps Gerontology Center [producer], 1997. Oxford, OH: Miami University, Scripps Gerontology Center/Electronic Data Center [distributor], 2003.

U.S. Dept. of Health and Human Services, National Center for Health Statistics. *National Health Interview Survey: Longitudinal Study of Aging, 70 Years and Over, 1984-1990*. [Computer file]. Hyattsville, MD: U.S. Dept. of Health and Human Services, National Center for Health Statistics [producer], 1984. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1988.

Manton, Kenneth *G* [Principal investigator]. *National Long-Term Care Survey, 1994* Durham, NC: Duke University, Center for Demographic Studies [producer & distributor], 1994.

E.g., U.S. Dept. of Health and Human Services, National Center for Health Statistics. National Nursing Home Survey, 1999 [Computer file]. Hyattsville, MD: U.S. Dept. of Health and Human Services, National Center for Health Statistics [producer], 1999. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2001.

[5] Ee.g., U.S. Dept. of Commerce, Bureau of the Census. Current Population Survey: Annual Demographic File, 2001 [Computer file]. Washington, DC: U.S. Dept. of Commerce, Bureau of the Census [producer], 2001. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2001.