Show Me the Data! Partnering with Instructors to Teach Data Literacy

Karen Hogenboom, Carissa M. Holler Phillips, and Merinda Hensley

Introduction

The shift to quantitative research methods in social science disciplines is old news in higher education, but the relative ease of acquiring data sets via the Internet and the availability of online analysis and visualization tools have recently increased the possibilities for using data sets in the classroom. Instructors are able to teach students how data analysis works in their disciplines with a low investment in software and server space, but students need to learn how to think critically about data and statistics before they can analyze data successfully. As information literacy encompasses more than the mechanical skills to get a list of results out of an article database, data literacy is much more than copying a table into a paper or even downloading a dataset and doing a regression analysis. Data literacy is the ability to read and interpret data, to think critically about statistics, and to use statistics as evidence.[1]

As faculty begin to incorporate data into their teaching, there is a role for the library's data services staff similar to the role that an instruction librarian plays in a course where students write their first college-level research paper. Karen Hunt points out that while data literacy and information literacy are similar in many ways, the technical tools and skills required to work with data adds a layer of complication to successfully working with students and instructors.[2] At the University of Illinois at Urbana-Champaign, we conducted a survey to find out how instructors were using data in teaching and what issues they face, in an attempt to shape instruction services to meet their needs. The survey also informed ongoing efforts, including what types of data sets the library should purchase, demand for specific software and the hardware necessary to support it, prioritizing training for librarians and staff, and working with the institutional repository for archiving data sets.

Literature Review

Kristin Anthony recently summarized the "disconnects" between faculty and librarians and suggested that the current focus on information literacy helps to bridge the gap because it frames the librarian's role in terms of the educational mission of the institution.[3] Data specialists have framed their partnerships with faculty in similar terms. For example, Hailey Mooney and Breezy Silver suggest that librarians target classes with a quantitative focus to find out how they are using numeric data.[4] Elizabeth Stephenson and Patti Schifter Caravello partnered with a sociology faculty member to include data literacy in an information literacy course,[5] an approach also advocated by Milo Schield.[6] At Miami University, the Electronic

Karen Hogenboom is Numeric and Spatial Data Librarian at the University of Illinois at Urbana-Champaign, e-mail: hogenboo@illinois.edu; Carissa M. Holler Phillips is Business and Finance Information Librarian at the University of Illinois at Urbana-Champaign, e-mail: choller@illinois.edu; Merinda Hensley is Instructional Services Librarian at the University of Illinois at Urbana-Champaign, e-mail: mhensle1@illinois.edu

Data Center pushes datasets to courses ready for online analysis, so that instructors can focus on student learning rather than technical issues.[7] The UK Data Archive has taken a similar approach, creating free sets of data and statistics about subjects that are commonly taught in lower-level university social science courses.[8]

Methods

The ATLAS/Library Data Services Task Force (DSTF) was formed in 2008 as a partnership between an established statistical and geographic information systems consulting service in the College of Liberal Arts and Sciences and the University Library, where several librarians were interested in working with data users and questions were being asked, but there was little experience with helping campus researchers find and format data for their teaching and research.

In Fall 2009, the DSTF conducted a survey of all campus faculty, staff, and graduate students in more than 38 departments that were likely to rely on numeric or spatial data for research or teaching. The web-based survey went out over e-mail to all faculty, academic professionals, and graduate students in these departments, and we received 925 responses (15% response rate). Some respondents do not have teaching roles and others do not use data in their teaching, but 252 of the respondents stated that they do use data in their teaching. These are the responses we will analyze in this paper.

Results

Instructors who teach with numeric or spatial data use it in a variety of ways, which leads to numerous opportunities for the library to assist students. The largest group of instructors who described their use of data in teaching are demonstrating statistical content or methods in class, but the next largest group of instructors ask their students to use data in course assignments or labs (Table 1).

When asked to express their greatest area of frustration when teaching with data, a few instructors had no issues but most were articulate in describing the obstacles they face. For many, the difficulty of finding data that meets their pedagogical needs is a major frustration. There were a variety of potential solutions proposed; the most common was for the library to provide some organization to the universe of data sources.

TABLE 1
Instructor Descriptions of How They Use Data for Instruction

Method	Number	%
Class demonstration	26	15.6
Labs or assignments	22	13.5
Demonstration and practice with research methods	4	2.4
Students collect and analyze data	4	2.4
Textual analysis	2	1.5
No specific answer	106	64.6
Total	164	100.0

TABLE 2
Instructors' Biggest Source of Frustration When Using Data in Teaching

Frustration	Frequency	%
Finding data	35	26.5
Technical difficulties/equipment problems	25	19.0
Technical or software skills	10	7.8
Rapid change in sources	10	7.5
Time consuming	7	5.6
No compilation of data resources	6	4.7
Formatting data	3	2.3
Maintaining data sets	2	1.7
Lack of statistical literacy	1	0.9
No problems teaching with data	13	10.0

When students need help with data, the majority of instructors refer them to the University Library, though a large proportion also send them to non-university websites as resources. This points to an expectation that the library will have robust services for students finding and using data.

Instructors also weighed in on training opportunities that would be helpful to their classes. A majority of respondents were interested in broad training about finding and working with data, though a significant number also felt that training in specific tools would be useful.

However, the in-class training sessions that have been a traditional goal for librarians were the least favored format for training. Without having the opportunity to ask why this is the case, we speculate that the obstacles to class instruction sessions are the same for data literacy as they are for any in-class library in-

TABLE 3
Where Instructors Refer Their Students for Help with Data

Source of help	Number	%
University Library	152	66.9
Non-university websites	116	48.7
Department or college resource	46	19.5
Instructor or students collect data	14	6.0
No need to refer	3	1.4
Note: Percentage does not add to 100 because of the possibility of multiple answers		

TABLE 4
Training that Would be Helpful to Students

Resource	Number (%) Who Find it Helpful
Finding and obtaining data	180 (75.7%)
Statistical analysis software (SAS, SPSS, Stata)	132 (55.4%)
Textual analysis software (Nudist, NVivo, AtlasTI)	97 (40.8%)
Geographic software (ArcGIS, MapInfo)	88 (36.9%)
Note: Percentage does not add to 100 because of the possibility of multiple answers	

TABLE 5
Format Preference for Training (Percentage)

Format	%
Workshop	52.7
Online guide	47.5
Online video	42.2
In-class demonstration	29.1
Note: Percentage does not add to 100 because of the possibility of multiple answers	

struction session: syllabi are packed with content and class time is precious. Therefore, anything seen as peripheral does not get scheduled.

Implications for Data Services

Since working with data is inherently experiential, librarians and other data services providers have the opportunity to guide data users in their development of transferable skills such as locating reliable sources of data, accessing and extracting the data, and critically evaluating the data. Based on the ways in which Illinois instructors reported using and interacting with data for their classes, a few indications of need (and several potential opportunities for librarians) were revealed.

Supporting Data Use in the Classroom

The survey indicated that when using data as part of their teaching efforts, instructors at Illinois most often use data in class demonstrations as well as in labs or assignments. However, offering support for courses is nothing new for librarians, and the literature of data services describes many methods and options for bringing data, data resources, and data literacy into the classroom.

In the most literal sense, this means getting into the classroom ourselves, conveying information directly to students to prepare them for success in data-intensive assignments, which is an activity which could take several different forms. There is the approach advocated by Schield,[9] which entails teaching a stand-alone course in the fundamentals of statistical literacy, to prepare students for more intensive data analysis coursework. Stephenson and Caravello[10] reported on such a course, which they offered to promote data literacy among sociology students. There are also some opportunities for presentations and instruction sessions during class time, which is a common activity for many librarians, to bring awareness to the resources that the students may use for data assignments and to the support that is available. Susan Czarnocki and Anastassia Khouri talked about presenting on the resources (equipment and software) and services available to help students with data-intensive assignments.[11] Lynda M. Duke, Kristin Vogel, and Suzanne Wilson discussed promoting the website of resources that they created by conducting instruction sessions during classes.[12] Presentations to classes not only have the direct benefit of making students aware of the assistance available to them, but as Eleanor J. Read noted[13] and these examples demonstrate, they also serve to expose students to data services.

In an indirect sense, getting into the classroom may entail collaborating with faculty to create data-intensive assignments, as Hunt[14] did, or working with faculty to "help to assess the feasibility of the data assignment."[15] More in line with the instructional support often provided, librarians may create tutorials that can be incorporated in a course, or which students can use for their own learning,[16] or can create

web pages that highlight recommended resources for data-intensive assignments.[17] These, too, can serve as promotional materials and activities for data services.[18] Another clear message we heard from our respondents was that the library could be a place for students to receive training on finding and obtaining data as well as learning how to use specific software packages (Table 4). Given that faculty frequently refer students to the library (Table 3), it seems appropriate for the library to create programming that is taught by librarians in conjunction with partners on campus. Instruction could also benefit the needs of faculty. This premise is reinforced by Czarnocki and Khouri, who describe how McGill library "worked to develop a full range of services which aim to assist not only the heavy-duty data crunchers, but also to assist any instructor, with courses of any size, to incorporate assignments that utilize data for instructional purposes."[19]

Organizing Data for the Classroom

Instructors responding to the Illinois survey also indicated that they have difficulty in finding data that meet their teaching needs. This data need may be related to findings from focus groups at UCLA which revealed that faculty "had stopped assigning papers that require students to find information themselves,"[20] meaning that data used in assignments or projects would need to be identified and provided by the instructors.

Respondents to the Illinois survey suggested that the library could help alleviate frustrations in this area by giving some organization and structure to the data that are available. Facilitating the discovery of datasets is one of the roles that Anna Gold specifically noted for librarians.[21] As experienced by Duke et al, who were approached by faculty expressing a similar desire for organization in the available datasets, the major consequence of disorganization was that "students, faculty, and librarians did not necessarily know what was available on campus or the best way to retrieve the datasets."[22] The potential benefits of organization then would be, beyond the obvious awareness of the resources, decreased frustrations for instructors and their students and increased efficiency for librarians in providing data services.

The literature shows that, in conjunction with creating web pages for specific classes or assignments, which itself is an act of organizing data, there is also the practice of organizing data by creating general subject-based dataset pages.[23] These web pages could provide access to pre-existing datasets, datasets that have been converted by librarians into more readily-usable formats (for example, through SPSS) to facilitate their use by students,[24] or perhaps datasets that have been created by librarians.[25] As Louise Corti noted, "Simplifying and re-packaging complex data…is one way of opening up their accessibility to the classroom."[26]

As with activities to support specific class assignments, there is also potential to promote data services through data organizing activities. Duke et al reported that their data organization project served as a promotional event in itself, with "increases in the number of requests for librarian-led instruction sessions, including additional components on locating data, and increases in the number of one-on-one reference conversations. Of particular interest was the greater number of direct referrals from teaching faculty to librarians for students to receive targeted assistance on projects."[27]

Generating Referrals for Assistance

Finally, Illinois instructors indicated that they most often refer students to the Library for assistance with data-centric assignments. Similarly, Read found that "among the various ways that people had learned about Data Services, being referred by a professor, instructor, or advisor was by far the most prevalent."[28] Taking guidance from the literature and from our own experiences and interactions with faculty and students at Illinois, we can consider how best to encourage faculty to refer students to the Library for support on data-intensive assignments. In part, this requires us to consider how to define and market the services we provide, which Gold considered to be "among the actions that librarians may find it possible and valuable to take."[29]

At a basic level, we can help faculty to understand conceptually what we can offer to students, in overall critical thinking and skills development. Ann S. Gray discusses general librarian skill sets, such as "mak[ing] assessments of quality and utility"; "furnish[ing] some guidance in interpretation of various types of statistical presentations"; and "point[ing] the way on how to interpret the results of analysis"[30] each of which could provide value to the students' experience with data. Beyond this basic level, there must be agreement and understanding as to what data services can do. Shrimplin and Yu,[31] Stephenson and Caravello,[32] and

Kristi Thompson and Daniel M. Edelstein[33] are explicit about the role of data services providers versus instructors' roles, specifically, that instructors provide the theory (both on the subject matter of the course and the application of statistics in that subject) and librarians and other data services providers provide the expertise to help data users locate, access, extract, customize, and evaluate data.

Having conveyed the "what" of our services, we can turn our discussions with faculty to the "how" and "where." An instructor's referral to data services implies a physical destination for a student, perhaps visiting offices or facilities in the library devoted to data services, including virtual destinations, such as web pages or other online materials. Read noted that, "It is also important to pay attention to the faculty, to help them as data users as well as conduits of information for their graduate students. They play an important advisory role and have referred their students to Data Services, so keeping them well-informed about data services and resources is a necessity."[34] Helping faculty to understand what resources and services are available to students at those destinations will facilitate better and more appropriate referrals. We should promote the features of the physical space (for example, databases, software, computers, consultation space, or instruction space) along with the services provided within those facilities, which could include consultation on assignments, general software training sessions, and course-specific resource training sessions.[35]

Developing Partnerships to Meet the Demand for Data Services

Facing the challenges associated with the delivery of data services, the Association of Research Libraries' GIS Literacy Project in the 1990's ventured to promote the "development of a team of GIS professionals in the research library community willing to lend time and expertise to applications, user training and education programs."[36] Taking data services outside the library and into the classroom, results of our survey have identified several opportunities for librarians to assist teaching faculty including offering classroom support, organizing data for the classroom, and generating referrals for assistance. Given existing financial constraints and varying skill levels, how can the library develop programs to meet growing data services needs creatively?

A scan of campus services may uncover data support personnel within departments who are amenable to partnering, thus combining efforts across areas of expertise. Joe Aufmuth explores two models: centralized and distributed systems. For libraries hesitant to create a hybrid and interconnected service, Aufmuth reminds us that, "By focusing on the user community and adding coordinate information, or spatial value, to library collections, libraries may be able to avoid competition with other campus centers."[37] Mary French noted that at Penn State "partnership development and/or enhancement of existing relationships with the campus' geography department, cartography lab, computing center, and a semi-independent research group proved crucial in fulfilling this mission."[38] And when instructors ask their students to use data in course assignments or labs (Table 1), Thompson and Edelstein acknowledge the role of the library is to support teaching in the classroom "by helping our patrons overcome the practical knowledge barriers that arise when conducting a data analysis project."[39] At Illinois, a hybrid program has been constructed by creating a strategic relationship between the Library and Applied Technologies for Learning in the Arts and Sciences (ATLAS). By combining efforts, the Library was able to regularly staff office hours with data services experts who provide assistance with data analysis software applications. ATLAS also deepens a series of data-intensive library workshops by teaching sessions on how to use SPSS and SAS to analyze data from the Interuniversity Consortium for Political and Social Research (ICPSR), which further promotes shared responsibility. The library's data services librarian has also made connections by working on the planning team for the campus's annual GIS Day. Future plans include creating online learning modules, staffing open lab hours so that users may use specialized software and hardware, and data visualization services. We are also looking for strategies to capitalize on the momentum of the institutional repository to assist users who are storing and sharing data sets in their teaching. If the Library can continue to reduce barriers for teaching with data sets, the academic community will recognize the Library as a leader in data literacy and consulting services, similar to recent gains in information literacy.

And finally, as observed by Mooney and Silver, librarians "will not only need to learn how to apply traditional techniques of reference, instruction, and

collection management to the unique format of data, but also how to raise awareness of this new service to the wider academic community."[40] Highlighting new services is an essential component of the planning process.

Assessing Our Activities

There is little guidance in the literature about formal assessment of data services, but a few authors have presented approaches that were tried, as well as considerations for future efforts.

One fairly common method of assessing student performance on data-intensive assignments is obtaining instructor feedback,[41] occasionally through survey tools.[42] On those occasions in which librarians or data services providers were the instructors, they used their own evaluations of the students' work and the students' formal evaluations of the courses. Stephenson and Caravello noted that "assessment of in-class activities and homework was so valuable," and determined that any future courses would include additional opportunities to assess students in this way.[43]

Surveys are another commonly-mentioned method for assessment of data services. Duke et al. presented the example of surveying students "to learn about student perceptions regarding the process of identifying datasets as well as the ease of using the dataset subject guide pages."[44] This type of targeted survey could provide feedback about the success of particular initiatives or functions of data services providers. At a macro-level, Read's survey of data users[45] sought to understand data needs and users' awareness of services and resources available to them, as an assessment of established data services overall, and provides several points of comparison for the Illinois survey and its results.

Finally, there are several objective assessment or evaluation measures that are presented in the literature. These include web page usage statistics[46] and counts of "consultations, ICPSR use, and workshop attendance."[47] Objective measures could also include, as Stephenson and Caravello suggested, basing assessment activities on subject-specific standards and learning outcomes,[48] such as the ACRL-ANSS learning outcomes for sociology and anthropology students.[49]

Conclusion

The new challenges presented by increasing use of data in teaching allows for increased promotion of librarians as data experts and an opportunity to build partnerships with instructors. And yet, we are reminded that the Library does not exist in a vacuum. Developing partnerships with data experts across campus who are already working with faculty and teaching assistants will help librarians to get to know instructors and how to better handle their data needs. The training that libraries offer needs to extend beyond simply helping individual users to access data sets, and needs to be informed by a deeper understanding of how faculty and students use data. In other words, the area of data services requires librarians to take a leadership role in advocating for true teaching partnerships, where instructors and librarians collaborate closely to create rich learning experiences for students using data.[50]

Notes

1. Milo Schield, "Statistical Literacy: Thinking Critically about Statistics," *Of Significance: A Topical Journal of the Association of Public Data Users* 1, no. 1 (1999): 15.

2. Karen Hunt, "The Challenges of Integrating Data Literacy into the Curriculum in an Undergraduate Institution," *IASSIST Quarterly* 28, no. 2/3 (2004): 14. Available online at http://www.iassistdata.org/downloads/iqvol282_3hunt.pdf. [Accessed 18 January 2011].

3. Kristin Anthony, "Reconnecting the Disconnects: Library Outreach to Faculty as Addressed in the Literature," *College & Undergraduate Libraries* 17 (2010): 79–92.

4. Hailey Mooney and Breezy Silver, "Spread the News: Promoting Data Services," *College & Research Libraries News* 71, no. 9 (2010): 480–483.

5. Elizabeth Stephenson and Patti Schifter Caravello, "Incorporating Data Literacy into Undergraduate Information Literacy Programs in the Social Sciences: A Pilot Project," *Reference Services Review* 35, no. 4 (2007): 525–540.

6. Milo Schield, "Information Literacy, Statistical Literacy and Data Literacy," *IASSIST Quarterly* 28, no. 2/3 (2004): 6–11. Available online at http://www.iassistdata.org/downloads/iqvol282_3shields.pdf. [Accessed 18 January 2011].

7. Aaron K. Shrimplin and Jen-chien Yu, "Focusing in on Student Learning Outcomes: How SDA Helped Us Get Data into the Classroom," *IASSIST Quarterly* 28, no. 2/3 (2004): 55–57. Available online at http://www.iassistdata.org/downloads/iqvol282_3shrimplin.pdf. [Accessed 18 January 2011].

8. Louise Corti, "Survey Data in Teaching Project (SDiT): Enhancing Critical Thinking and Data Literacy," *IASSIST Quarterly* 28, no. 2/3 (2004): 39–54. Available online at http://www.iassistdata.org/downloads/iqvol282_3corti.pdf. [Accessed 18 January 2011].

9. Schield, "Information Literacy, Statistical Literacy and Data Literacy," 6–11.

10. Stephenson and Caravello, "Incorporating Data Literacy into Undergraduate Information Literacy Programs in the Social Sciences," 525–540.

11. Susan Czarnocki and Anastassia Khouri, "A Library Service Model for Digital Data Support," *IASSIST Quarterly* 28, no. 2/3 (2004): 33. Available online at http://www.iassistdata.org/downloads/iqvol282_3khouri.pdf. [Accessed 18 January 2011].

12. Lynda M. Duke, Kristin Vogel, and Suzanne Wilson, "Faculty-Library Collaboration in Dataset and Statistics Management," *Technical Services Quarterly* 24, no. 4 (2007): 52.

13. Eleanor J. Read, "Data Services in Academic Libraries: Assessing Needs and Promoting Services," *Reference & User Services Quarterly* 46, no. 3 (2007): 68–9.

14. Hunt, "The Challenges of Integrating Data Literacy into the Curriculum in an Undergraduate Institution," 12–15.

15. Czarnocki and Khouri, "A Library Service Model for Digital Data Support," 34.

16. Corti, "Survey Data in Teaching Project (SDiT)," 43.

17. Shrimplin and Yu, "Focusing in on Student Learning Outcomes," 55–57; Czarnocki and Khouri, "A Library Service Model for Digital Data Support," 32.

18. Read, "Data Services in Academic Libraries," 68–9.

19. Czarnocki and Khouri, "A Library Service Model for Digital Data Support," 33.

20. Stephenson and Caravello, "Incorporating Data Literacy into Undergraduate Information Literacy Programs in the Social Sciences," 530.

21. Anna Gold, "Cyberinfrastructure, Data, and Libraries, Part 2: Libraries and the Data Challenge: Roles and Actions for Libraries," *D-Lib Magazine* 13, no. 9/10 (2007). Available online at http://www.dlib.org/dlib/september07/gold/09gold-pt2.html. [Accessed 18 January 2011].

22. Duke et al., "Faculty-Library Collaboration in Dataset and Statistics Management," 47.

23. Ibid., 48.

24. Ibid., 52.

25. Shrimplin and Yu, "Focusing in on Student Learning Outcomes," 56; Czarnocki and Khouri, "A Library Service Model for Digital Data Support," 31.

26. Corti, "Survey Data in Teaching Project (SDiT)," 40.

27. Duke et al., "Faculty-Library Collaboration in Dataset and Statistics Management," 51.

28. Read, "Data Services in Academic Libraries," 68.

29. Gold, "Cyberinfrastructure, Data, and Libraries, Part 2."

30. Ann S. Gray, "Data and Statistical Literacy for Librarians," *IASSIST Quarterly* 28, no. 2/3 (2004): 28. Available online at http://www.iassistdata.org/downloads/iqvol282_3gray.pdf. [Accessed 18 January 2011].

31. Shrimplin and Yu, "Focusing in on Student Learning Outcomes," 56.

32. Stephenson and Caravello, "Incorporating Data Literacy into Undergraduate Information Literacy Programs in the Social Sciences," 530.

33. Kristi Thompson and Daniel M. Edelstein, "A Reference Model for Providing Statistical Consulting Services in an Academic Library Setting," *IASSIST Quarterly* 28, no. 2/3 (2004): 35. Available online at http://www.iassistdata.org/downloads/iqvol282_3thompson.pdf. [Accessed 18 January 2011].

34. Read, "Data Services in Academic Libraries," 71.

35. Czarnocki and Khouri,"A Library Service Model for Digital Data Support," 33.

36. D. Kevin Davie, James Fox, and Barbara Preece, *The ARL Geographic Information Systems Literacy Project*, SPEC Kit 238. (Washington, D.C.: Association of Research Libraries, 1999), 16.

37. Joe Aufmuth, "Centralized vs. Distributed Systems: Academic Library Models for GIS and Remote Sensing Activities on Campus," *Library Trends* 55, no. 2 (2006): 340–348.

38. Ibid., 347.

39. Mary French, "Creating Support Mechanisms for Library GIS Services: A Perspective," *Current Studies in Librarianship* 25, no. 1/2 (2001): 43.

40. Thompson and Edelstein, "A Reference Model for Providing Statistical Consulting Services in an Academic Library Setting," 35.

41. Mooney and Silver, "Spread the News," 480.

42. Czarnocki and Khouri, "A Library Service Model for Digital Data Support," 34.

43. Duke et al., "Faculty-Library Collaboration in Dataset and Statistics Management," 51.

44. Stephenson and Caravello, "Incorporating Data Literacy into Undergraduate Information Literacy Programs in the Social Sciences," 534.

45. Duke et al., "Faculty-Library Collaboration in Dataset and Statistics Management," 51.

46. Read, "Data Services in Academic Libraries," 61.
47. Duke et al., "Faculty-Library Collaboration in Dataset and Statistics Management," 52.
48. Read, "Data Services in Academic Libraries," 72.
49. Stephenson and Caravello, "Incorporating Data Literacy into Undergraduate Information Literacy Programs in the Social Sciences," 535.
50. Association of College and Research Libraries (ACRL), "Information Literacy Standards for Anthropology and Sociology Students." Available online at http://www.ala.org/ala/mgrps/divs/acrl/standards/anthro_soc_standards.cfm. [Accessed 18 January 2011].