

Business Intelligence and Data Visualization with Tableau in Research Libraries

Martha Kyrillidou, QualityMetrics LLC

Research libraries engage in understanding their environment by asking questions, observing and collecting data, and analyzing and reporting what they learn. The aims are to improve library services as well as to get the word out about the value libraries deliver to their end users in a strategic, visually appealing dashboard that delivers needed information in a timely manner. This issue of *RLI* reports on the latest applications in research libraries of Tableau, a business intelligence and data visualization tool. The two articles published here were originally presented at the Library Assessment Conference in Seattle, Washington, in August 2014 and are also included in the conference proceedings.¹

Among the most expensive and rapidly growing services libraries provide is the licensing of electronic resources, as Lewellen and Plum report in “Assessment of E-Resource Usage at the University of Massachusetts Amherst: a MINES for Libraries® Study Using Tableau for Visualization and Analysis.”² Even though research libraries are more actively engaged in publishing, a large and disproportionately increasing part of the library budget is devoted to purchasing electronic journals from a handful of publishers.³ In 2003, the Association of Research Libraries (ARL) identified the need to track and evaluate the usage of electronic resources as a key priority.

ARL adopted and deployed the Measuring the Impact of Networked Electronic Services (MINES) for Libraries® service⁴ and implemented it in a number of institutions. MINES for Libraries goes deeper than usage statistics as it asks users to identify how the resource they are using is linked to their learning, research, and teaching. Lewellen and Plum report on two implementations of the MINES for Libraries protocol at the University of Massachusetts Amherst. They summarize

the overall goals of the protocol in addition to an evolution of implementation options coupled with a pragmatic goal of demonstrating value. The University of Massachusetts Amherst uses Tableau to analyze and report MINES for Libraries data in the library’s most recent implementation of MINES.

Lewellen and Plum discuss the pros and cons of different MINES implementation options and demonstrate convincingly how tracking usage of electronic resources at a slightly deeper level than COUNTER-compliant usage statistics is a realistic and achievable approach for research libraries that invest large amounts of money in purchasing electronic resources. The authors expect that they “will

In This Issue

Assessment of E-Resource Usage at UMass Amherst 5

Tableau Unleashed: Visualizing Library Data 21

continue to collect valuable, actionable data to present a comprehensive picture of e-resource use to library and campus stakeholders, specifically informing collection development, instruction, support for research, marketing, and liaison work.”⁵ In their article, Lewellen and Plum show how far libraries can drill down into usage statistics without threatening user privacy. They also raise questions about how to assess the new roles and services libraries develop. For example, their approach surfaces the critical issue of how open access and library publishing efforts need to be configured so that their usage and value can be captured as the usage and value of licensed resources is captured.

Buhler, Lewellen, and Murphy reported on a variety of library Tableau applications at the Library Assessment Conference in 2014.⁶ They followed up their conference presentation with a series of four webcasts that ARL organized and distributed through the ARL YouTube channel.⁷ In each of the first three webcasts, one of the authors presented the variety of Tableau implementations at their library, and the fourth webcast provided an opportunity for more in-depth discussion of the issues across these three research library settings: the University of British Columbia (UBC), the University of Massachusetts (UMass) Amherst, and The Ohio State University (OSU).

These three authors approach their *RLI* piece, “Tableau Unleashed: Visualizing Library Data,” by answering the following questions:

- How has the library incorporated Tableau into its assessment program?
- What impact has Tableau had on making sense of large data sets, making data accessible, and improving stakeholder communications?
- Where does Tableau fit in the library’s data strategy?

Murphy and Lewellen present examples of data publishing and data sharing and Buhler emphasizes data exploration. Hopefully the example of these libraries will inspire more widespread mining of business intelligence in libraries.

Murphy summarizes the Tableau dashboards regarding research services trends, gate counts, and ILLiad borrowing at OSU. The data are accessible to key stakeholders in the library and have enabled them to communicate the value of their services to internal and external constituencies. Tableau is a strategic asset in the library’s assessment program.

Lewellen showcases a range of visualizations and applications at UMass by highlighting monograph purchasing, circulation and duplication at both aggregate and title-level detail, and the e-book library (EBL) pilot program across the Five College Consortium libraries. In the case of UMass, using Tableau enables the library to mine business intelligence without a comprehensive data warehouse implementation solution.

Buhler discusses data exploration based on the principles summarized by Stephen Few⁸ and he uses circulation data and LibQUAL+ examples from UBC. The ability to mine the LibQUAL+ data

longitudinally by discipline and school provide new perspectives. As Buhler asserts, “None of the LibQUAL+ visualizations presented...are based on data that is new to UBC Library, but Tableau helped to breathe new life into relatively commonplace data sets, making them more relevant to certain audiences.”⁹ For UBC, Tableau offered strategic communication advantages for the library’s assessment program.

As the authors state, “Using Tableau, a library may produce flexible, in-depth, online dashboards, complete with filters and annotations to both customize visualizations and provide context. A library may also blend data from disparate sources to create dynamic, interactive graphics and reports.”¹⁰ The art of visually communicating library usage has come a long way from the initial ARL Statistics Interactive Edition that was developed in 1995 through a collaboration of Kyrillidou at ARL and Stubbs at the University of Virginia in the early days of the World Wide Web.¹¹

The future of business intelligence through research library data is to be realized by mining the dynamic, real-time, and scalable visualizations that tools like Tableau afford us. For example, a scalable application of MINES for Libraries with Tableau **across libraries** is within reach.¹² Furthermore, coupling MINES for Libraries with SHARE notifications¹³ could realize the potential of demonstrating the value of open access content as well as purchased content. Notifications that alert users when content from a wide variety of sources is first published, along with embedded, real-time, usage reports, for example, could identify how useful those publications are at a point in time, and how they further user goals such as learning, research, and teaching as captured by MINES data. Identifying influential knowledge resources—publications and authors among them—in this way could bring libraries a step closer to understanding what individuals need to read or write about next. As a result, continuously mining and acting upon data about knowledge resources and their use may contribute to exponential rates of growth in learning, research, and teaching.

Endnotes

1. Sarah Durso, Steve Hiller, Martha Kyrillidou, and Angela Pappalardo, *Proceedings of the 2014 Library Assessment Conference: Building Effective, Sustainable, and Practical Assessment, August 4–6, 2014, Seattle, WA* (Washington, DC: Association of Research Libraries, 2015), <http://libraryassessment.org/bm~doc/proceedings-lac-2014.pdf>.
2. Rachel Lewellen and Terry Plum, “Assessment of E-Resource Usage at University of Massachusetts Amherst: A MINES for Libraries® Study Using Tableau for Visualization and Analysis,” *Research Library Issues: A Report from ARL, CNI, and SPARC*, no. 288 (2016): 5–20, <http://publications.arl.org/rli288/>.
3. Karla L. Strieb and Julia C. Blixrud, “Unwrapping the Bundle: An Examination of Research Libraries and the ‘Big Deal,’” *portal: Libraries and the Academy* 14, no. 4 (Oct. 2014): 587–615.
4. MINES for Libraries, accessed March 29, 2016, <http://minesforlibraries.org/home/>.
5. Lewellen and Plum, “Assessment of E-Resource Usage at University of Massachusetts Amherst,” 18.

6. Jeremy Buhler, Rachel Lewellen, and Sarah Anne Murphy, "Tableau Unleashed: Visualizing Library Data," *Research Library Issues: A Report from ARL, CNI, and SPARC*, no. 288 (2016): 21–36, <http://publications.arl.org/rli288/>.
7. Association of Research Libraries, "Data Visualization Webcast Videos Now Online," news release, May 11, 2015, <http://www.arl.org/news/arl-news/3587>.
8. Stephen Few, *Information Dashboard Design: The Effective Visual Communication of Data* (Sebastopol, California: O'Reilly Media, 2006).
9. Buhler, Lewellen, and Murphy, "Tableau Unleashed," 35.
10. Ibid., 21.
11. Martha Kyrillidou and Kendon Stubbs, "Interactive Statistical Analysis on the WWW," presented at the National Center for Educational Statistics Summer Data Conference, Washington, DC, June 1995; Martha Kyrillidou, "To Describe and Measure the Performance of North American Research Libraries," *IFLA Journal* (2001): 257–263; also, see ARL Statistics Analytics, accessed March 29, 2016, <http://www.arlstatistics.org/analytics/>.
12. Dana Thomas, Catherine Davidson, Martha Kyrillidou, and Terry Plum, "Measuring Use of Licensed Electronic Resources: A Second Iteration of the MINES for Libraries® Survey on Scholars Portal and Other Resources for the Ontario Council of University Libraries," *Library Management* 33, no. 6/7 (2012): 374–388, <http://www.emeraldinsight.com/doi/abs/10.1108/01435121211266195>.
13. A SHARE notification captures when a new resource becomes available, either through licensed proprietary databases or through open access repositories and digital collections. For more information, see the SHARE website <http://www.share-research.org/>.

© 2016 Martha Kyrillidou



This article is licensed under a Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

To cite this article: Martha Kyrillidou. "Business Intelligence and Data Visualization with Tableau in Research Libraries." *Research Library Issues: A Report from ARL, CNI, and SPARC*, no. 288 (2016): 1–4. <http://publications.arl.org/rli288/>.

Assessment of E-Resource Usage at University of Massachusetts Amherst: A MINES for Libraries® Study Using Tableau for Visualization and Analysis

Rachel Lewellen, Assessment Librarian, University of Massachusetts Amherst

Terry Plum, formerly Assistant Dean, Simmons School of Library and Information Science (retired)

Introduction

As academic libraries spend an increasing amount of their budgets on electronic resources, librarians seek to find out who is using the electronic resources and why, so libraries can provide better service. One tool is COUNTER,¹ which sets standards for recording and reporting usage of networked electronic resources. From COUNTER-compliant data-usage reports from subscribing vendors, libraries receive the number of successful full-text article requests by month and journal, total search requests by platform or database, successful section requests by e-book, and various title request reports, among other reports.² As described by Emery and Stone,³ there are a number of tools for analyzing COUNTER data in addition to those in an electronic resource management system, including vendor-supplied applications like ProQuest 360 Counter,⁴ EBSCO Usage Consolidation,⁵ and consortial tools like Journal Usage Statistics Portal.⁶ With these data and tools, performance indicators such as downloads per FTE user, cost per search, cost per view, cost per FTE user, and other metrics can be generated.⁷

However, as Marcum and Schonfeld note, as useful as the COUNTER-compliant activity counts are for library usage data and performance statistics,

More granular data on individual users' activities can afford greater opportunities to analyze needs and develop or optimize services. But considerations for user privacy can make librarians uncomfortable with such granular data, which has to some degree impeded our ability to establish the types of personalized services that are skyrocketing on the consumer internet.⁸

This paper presents a complementary methodology to COUNTER, giving the greater granularity while maintaining user anonymity and privacy. Reported here are the findings of the second, yearlong, assessment of electronic resources at the University of Massachusetts Amherst, using the Association of Research Libraries (ARL) MINES for Libraries methodology.⁹ MINES is an online, transaction-based, point-of-use, intercept web survey methodology that collects data on the patrons' purpose of use of electronic resources and on the demographics of users. This methodology helps to measure the impact of library services and to identify opportunities to serve faculty and students more effectively, which as Marcum and Schonfeld note are the desired goals for assessment.¹⁰ The MINES data gives a picture of users and usage that does not replace COUNTER data but gives a more complete and deeper picture.

The University of Massachusetts implemented MINES twice, in 2008–2009 and 2013–2014. This paper compares two implementation methods for a point-of-use, intercept survey launched at the EZproxy server: (1) randomly chosen two-hour sessions and (2), an every-Nth-user systematic methodology. The 2008–2009 survey used 24 two-hour time blocks spread over 12 months to survey users of e-resources (primarily e-journals and databases). The 2013–2014 implementation, which for the purposes of this paper closed June 30, 2014, surveyed every 140th usage passing through the proxy server. The paper compares the two methods for reliability and validity of the results and ease of technical implementation and reports on the results of the recent survey, examining user demographics, time and date analysis, location of use, purpose of use, and collection development implications.

Further, this paper demonstrates how using business intelligence software for data analysis and visualizations¹¹ to interact with the survey data in real time helped to:

- review and use live data throughout the year providing the ability to monitor the projected total results to make adjustments in real time (The survey frequency was increased from every 200th user to every 140th user to collect sufficient data to answer collection development research questions.);
- compare the distribution of sampled e-resources to all the usage of e-resources to judge the reliability of the sample;
- present a more informative visual display over SPSS and Excel graphics revealing relationships more easily and clearly; and
- collect survey data continuously, running the survey for the foreseeable future, and consider expanding the survey scope to include other resources.

Finally, the paper shows how data collected about users including status, academic affiliation, and purpose of use creates a deep picture of usage that can be combined with COUNTER data to give a more complete picture of electronic resource usage.

MINES for Libraries

As described on the MINES for Libraries website, <http://www.minesforlibraries.org/>,¹² MINES stands for Measuring the Impact of Networked Electronic Services and is an online, transaction-based, intercept survey that collects data on the purpose of use of electronic resources and on the demographics of users, developed by Brinley Franklin and Terry Plum. MINES was adopted by the Association of Research Libraries as part of the New Measures toolkit in May 2003. It is a point-of-use web survey of three to five questions that integrates usage data about electronic resources such as digital collections, open access journals, pre-print and post-print servers, and institutional repositories, to give an inclusive picture of the library's supported networked electronic resources.

In general, MINES for Libraries aims to:

- measure the value and impact of digital content;
- determine how specific user populations apply digital content to their work, based on demographic and purpose-of-use analyses;
- identify where library use originates in the networked environment and tailor services accordingly; and
- gather usage data about digital collections to justify increased funding for digital content and to make informed collection development decisions.

The roots of MINES are in indirect cost or facilities and administrative (F&A) cost studies focused on the library to help universities provide evidence for determining an accurate infrastructure support cost associated with sponsored research. These cost analysis library studies have been administered since the early 1980s by Brinley Franklin, although the web surveys have been in use since 2000. In addition to the web survey, the MINES methodology can also be used to assign a monetary value to a cost center in the academic library—such as a class of materials like e-journals or to specific vendors such as Elsevier’s ScienceDirect—to determine the portion of the cost center dedicated to the support of funded research, instruction, patient care, public service, and other activities by different classes of patrons of the library.

MINES employs a web-based user survey intercept methodology that delivers a short survey at the point of use of an e-journal, database article, or digital collection or service. There are two research designs or sampling techniques recommended by MINES, both of which result in a random sample of patron usage of networked electronic resources, the distributions of which can be applied with confidence to the user population. Note that in both cases the survey is a usage survey, not a user survey. MINES attempts to represent usage in the survey sample. The MINES protocol recommends that the questions on the survey are mandatory. Finally, even though the questions are mandatory, typically there are abandonments or patrons who back out of the survey. Depending on the implementation, these abandonments may be captured, and the respondents’ eventual responses, if any, may be recorded. We have found that 5,000 records collected over a year will answer most of the research questions posed by libraries that implemented MINES.

1. Random Moment Sampling

In the random moment sampling method for each month over a year either one or two two-hour survey sessions are randomly chosen for administering the web survey. The web survey then intercepts usage over that sampled period, querying the patron at the first usage of a surveyed resource, and repeatedly copying the values of the survey to a database at every subsequent use of surveyed resources by the patron during that two-hour period, adding a new time/date stamp and target URL or the URL of the surveyed resource. Currently, this sampling technique is employed for the cost analyses studies to determine F&A rates, and not for the ARL-supported MINES studies. The strengths of this sampling technique are as follows:

- The users are intercepted only 24 or 48 hours over a year, so the annoyance index is low.

- The number of records collected is usually sufficient to analyze usage by cost center.
- The data are collected at a common point, and therefore are commensurable (like COUNTER-compliant data) across diverse networked electronic resources.
- The survey collects data in a manner consistent with how patrons seem to use resources, that is, patrons often conduct a literature search intensely for a short period. The two-hour survey session tracks this burst of searching activity.
- The sample is random.

The weaknesses of this sampling technique are as follows:

- The intercept survey is intrusive.
- To log usage over a two-hour period the technologist must set up a session for the patron's browser that will remember some token for the patron over the two-hour period. The session is usually established with browser cookies or a server-side session. Although there is no identifying information associated with the patron, it is important to this technique that the survey session knows the patron is still the patron.
- Depending on the intercept point, described in following paragraphs, the survey session may be more or less difficult to implement.
- Target URLs can be difficult to analyze. Some are dynamically generated, involve redirects, and may be hard to decipher by visual inspection as in the case of digital object identifiers (DOIs).

2. Systematic Sampling

With this sampling technique every Nth usage is sampled at some choke point or virtual gateway, such as OCLC's EZproxy¹³ or an open URL link resolver, e.g., ExLibris SFX.¹⁴ The systematic sampling is an equal probability method within the ordered sampling frame and is often referred to as an every Nth sample. With this technique, an N is established and the starting point is randomly chosen. In libraries that have administered an every Nth sample, N has ranged from 1:500 to 1:140. Like the random moment sample, the every Nth data collection is conducted over a year to capture the different states of academic library usage: the academic year, the summer, and intersessions. The second Ontario Council of University Libraries (OCUL) study by ARL, OCUL, and the University of Toronto is a systematic sample for the 20 libraries involved with the study.¹⁵ ARL currently recommends this sampling method for MINES implementations.

The strengths of this sampling technique are as follows:

- The N can be changed to increase the number of records collected or to reduce the annoyance.

- The most difficult part of the technical implementation for the random moment sample, that is, the creation of the session, is eliminated, so that the every Nth sample is technically easier to set up than the random moment sample.
- The data are collected at a common point, and therefore are commensurable (like COUNTER compliant data) across diverse networked electronic resources.
- The sample is random.

The weaknesses of the systematic sample are as follows:

- The intercept survey is intrusive.
- The every Nth sampling does not reflect patterns of heavy usage and light usage by individuals in the same manner as the two-hour session, although heavy users are likely to be surveyed more often over the year than light users.
- Depending on the intercept point, described in following paragraphs, the survey session may be more or less difficult to implement.
- Target URLs can be difficult to analyze. Some are dynamically generated, involve redirects, and may be hard to decipher by visual inspection as in the case of digital object identifiers (DOIs).

The intercept point for MINES should be some virtual gateway through which most users choosing a networked electronic resource must pass. Because the MINES survey methodology is based upon capturing the target URL or selected networked electronic resource, open access resources not included in the library's electronic resource management system, bookmarks that do not include the proxy prefix, password-based alerting services, any services that depend on vendor passwords rather than some library mechanism, or e-books downloaded to e-readers are all problems for the MINES survey methodology. The intercept must be done locally at the library's web services, unlike LibQUAL+® where the survey is accessible through the ARL platform.

Libraries have implemented MINES with different techniques and gateways, such as prepended Java, PHP or JavaScript redirects from a list of resources, a survey intercept at the proxy rewriter such as EZproxy, or a survey intercept at the open URL link resolvers such as Ex Libris SFX, Innovative Interfaces's WebBridge LR, and ProQuest's 360 Link. The SFX solution is described by Thomas and others.¹⁶ One version of an EZproxy solution has been written up by Reese.¹⁷ One of the advantages of the EZproxy implementation is that resources and services can be placed behind the EZproxy application, and therefore can be surveyed. For example, PubMed is often not behind the EZproxy server since it is a free resource. However, many libraries wish to include PubMed as a surveyed resource, in part because of its LinkOut feature, and it can be added to the EZproxied resources. In many cases the surveyed resource could be added as an open URL link resolver target also.

In another common networked-service technique that will increase the validity of the sample, the technology group at the university campus can push out the appropriate library access links (open URL link resolver) within Google Scholar to browsers on campus, thus increasing traffic to the link resolver and the proxy server if one is used, tightening up the web of possible survey points. The most comprehensive interception point is at the Internet service provider router for the university, and as radical as it may seem, this router-based solution has been implemented twice by one university and has been contemplated by several others. The router-based solution has the fewest limitations, but the other intercept points also work well, and collect reliable samples, as long as the limitations are understood.

University of Massachusetts Amherst Environment and Implementation

The University of Massachusetts (UMass) Amherst is a public research university offering undergraduate, graduate, and professional degrees. There are 28,518 undergraduate and graduate students, and 1,170 full-time instructional faculty. The UMass Amherst Libraries is the largest state-supported academic library in New England with more than 8 million items, spending \$6 million on continuing e-resources.

The UMass Amherst Libraries implemented MINES twice, once using the two-hour random moment sample design in 2008–2009, and then the every Nth sample design in 2013–2014. The first survey ran for 12 months, from September 2008 to August 2009 with two two-hour survey periods per month. Figure 1 is a screen shot of the survey.

University of Massachusetts Amherst
Libraries

Electronic Resource Use Survey

This survey is being conducted during by the University of Massachusetts Amherst Libraries to assess how electronic resources are used to support teaching, coursework and research. After completing this brief survey you will be connected to the resource you selected. Thank you very much for your time!

ALL RESPONSES ARE ANONYMOUS

| | |
|--|---|
| 1. Are you a: | Select One |
| 2. Are you affiliated with: | Select One |
| 3. Your location: | Select One |
| 4. Primary purpose today (select one): | <input type="radio"/> Teaching [more information] <input type="radio"/> Coursework [more information] <input type="radio"/> Thesis or Dissertation [more information] <input type="radio"/> Sponsored (Funded) Research [more information] <input type="radio"/> Non-Sponsored or other Research [more information] <input type="radio"/> Other [more information] |
| 5. Why did you choose this resource? | Select One |

Please answer all questions

If you have difficulty check [connection problems](#) or contact dbhelp@library.umass.edu.

For more information or to provide feedback see the [survey homepage](#).

FIGURE 1. SCREEN SHOT OF MINES FOR LIBRARIES SURVEY AT UMASS AMHERST

Because of technology limitations, the survey had to be manually turned on and off at the EZproxy server for each survey period. The help desk survey support changed from daytime to evening and the survey form was updated to reflect the changes each survey period. There were 4,396 completed surveys that were linked to URLs in the proxy log. The BioStatistics Consulting Group at the university ran the analysis on the data in SPSS, producing useful data but in the standard ASCII SPSS tabular format. Library staff then reworked these data in Excel, generating pie charts and cleaner tables. Figures 2a and 2b show how the visual presentation appeared.

| Question 3 - Location | | |
|--------------------------------------|------|----|
| | Freq | % |
| Off Campus | 2044 | 47 |
| On campus but not in Du Bois or ISEL | 1561 | 36 |
| In the W.E.B. Du Bois Library | 678 | 15 |
| In ISEL | 113 | 3 |

FIGURE 2A. EXCEL TABLE OF RESPONSES TO MINES SURVEY QUESTION ABOUT LOCATION OF E-RESOURCE USE, 2008-2009

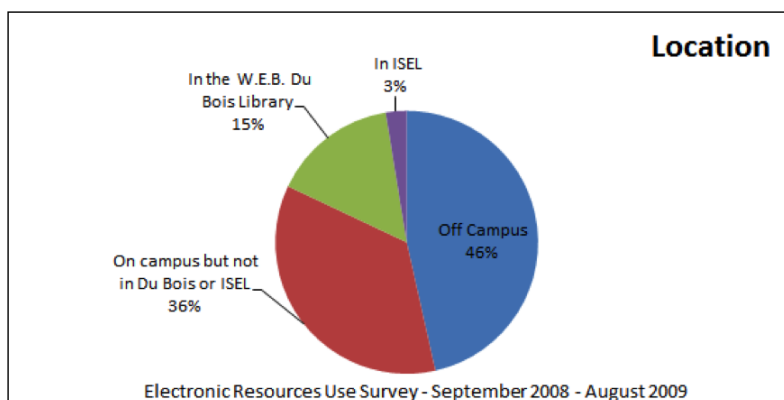


FIGURE 2B. EXCEL PIE CHART OF RESPONSES TO MINES SURVEY QUESTION ABOUT LOCATION OF E-RESOURCE USE, 2008-2009

The highlights from this study were that most of the users of networked electronic resources are not physically located in the UMass Amherst Libraries; there were a surprising number of undergraduates involved with research; and the libraries documented its contribution to the sponsored research endeavor as well as to teaching and learning.

In 2013, the UMass Amherst Libraries implemented the MINES survey with the every Nth (N=140) systematic sample design. A Perl script with rules for presenting the survey was invoked using the EZproxy service banner redirect setting. The data was collected in a MySQL database, and live Tableau software connections were used to analyze and visualize the results. Figure 3 shows how the survey intercept works.

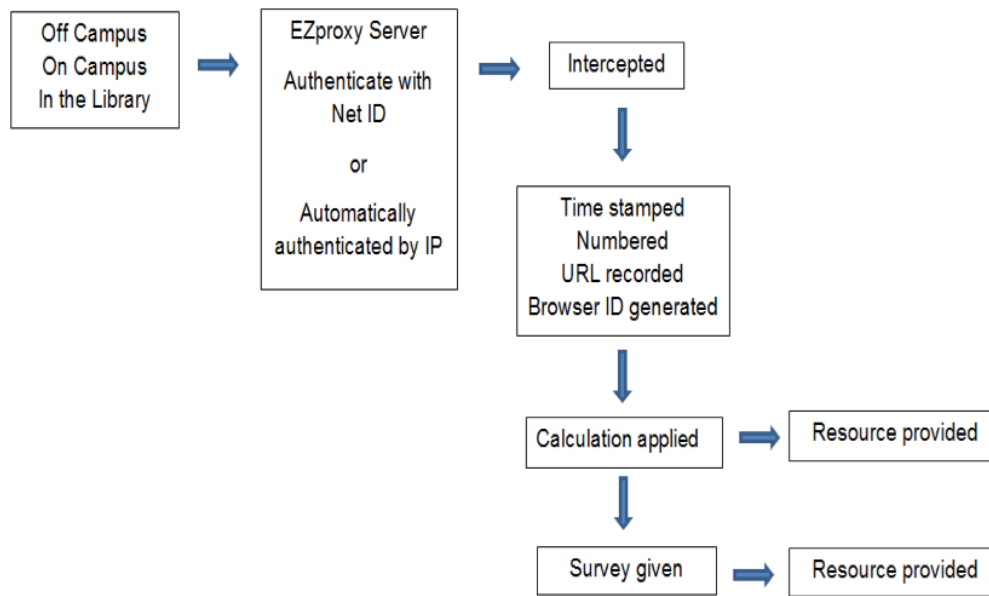


FIGURE 3. SURVEY INTERCEPT PROCESS FOR EVERY NTH SAMPLING METHOD USED IN 2013-2014

In this survey an abandonment reduction factor was included, making the survey more valid. Once the patron browser request was intercepted and the survey launched, the patron had five minutes to complete the survey. If the patron did not complete the survey in that time, then the survey timed out. A multi-step intercept calculation accounted for expired surveys and maintained the desired 140th user rate of return. The median time to complete the survey was 45 seconds (five minutes was the maximum allowed); the fastest survey was completed in 12 seconds. Figure 4 shows the survey completion time distribution in seconds. The overall response rate was 71%.

Survey Completion Distribution

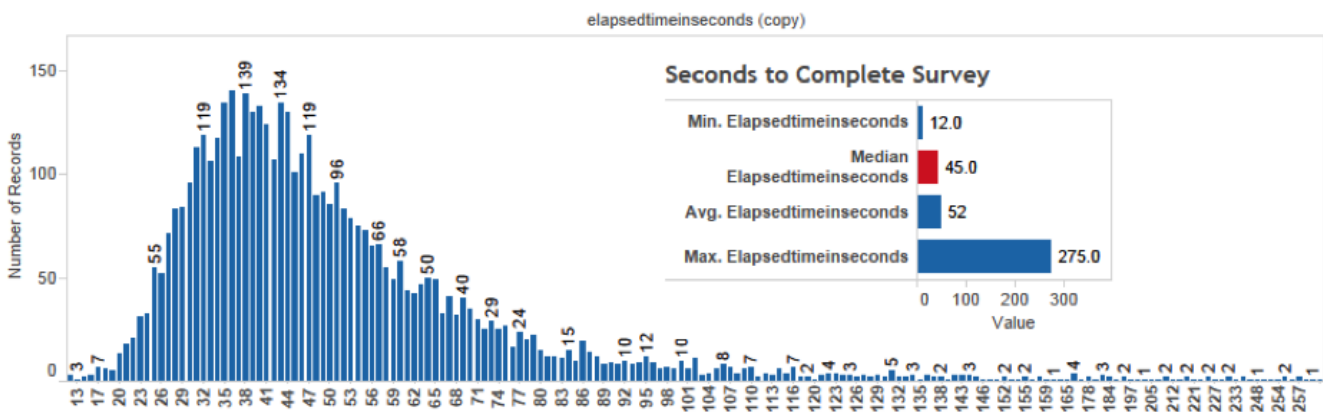


FIGURE 4. SURVEY COMPLETION TIMES, WHEN USING EVERY NTH SAMPLING METHOD IN 2013-2014

Tableau Software

Although Excel has its advocates for the analysis of the usage of electronic resources in libraries,¹⁸ we used Tableau Software¹⁹ for analysis and visualization in the second iteration of MINES at UMass. Tableau Software is a business intelligence software that can be used for data analysis and interactive data visualization. It is increasingly popular in analyzing library usage statistics.²⁰ With Tableau, data results can be monitored in real time. Monitoring the results in real time proved useful because early in the study it was observed that an N of 1:200 was not collecting sufficient data to lead to reliable and valid results. With N at 1:200, 3,477 surveys would have been collected over the year, fewer than the 2009–2010 survey. By adjusting the N to 1:140 at the end of the first quarter roughly 5,035 records were collected.

Tableau compared the data collected at the Nth use with the data collected for all use at the proxy server to confirm that the distributions were indeed similar and the sample was a representative sample of the population of all EZproxy use. This validity analysis was done on the frequency distribution of web surveys by hour during the day compared to all EZproxy usage and the frequency of usage by month (see Figures 5a and 5b). As can be seen by inspection, the distributions are almost identical.

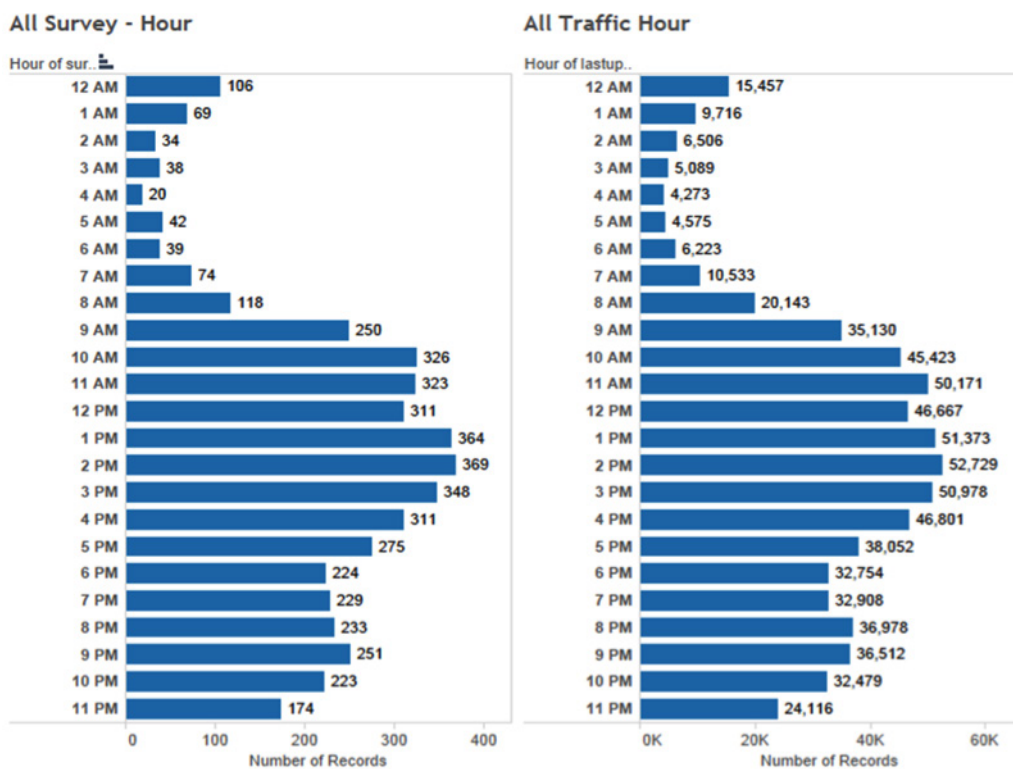
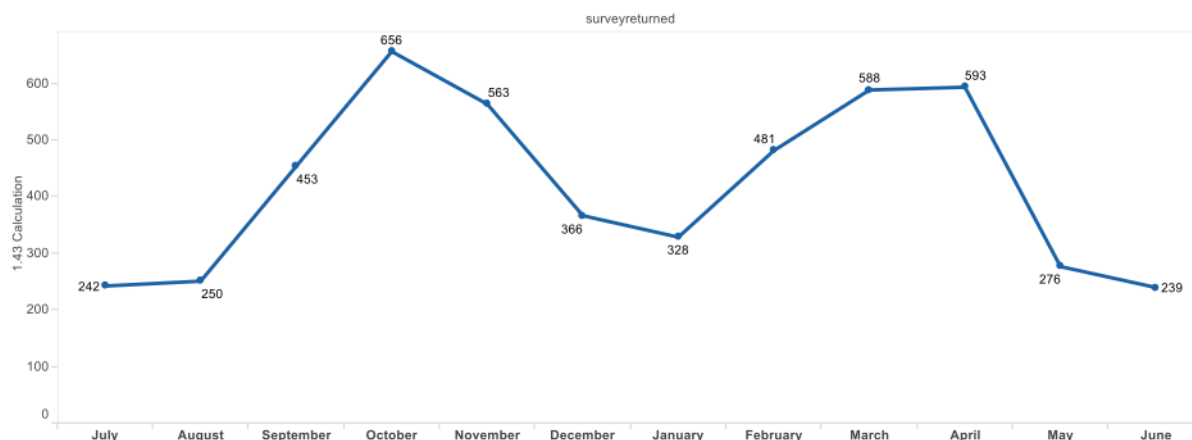


FIGURE 5A. HOURLY MINES SURVEY AND ALL PROXY TRAFFIC

Survey Monthly



All Traffic Monthly

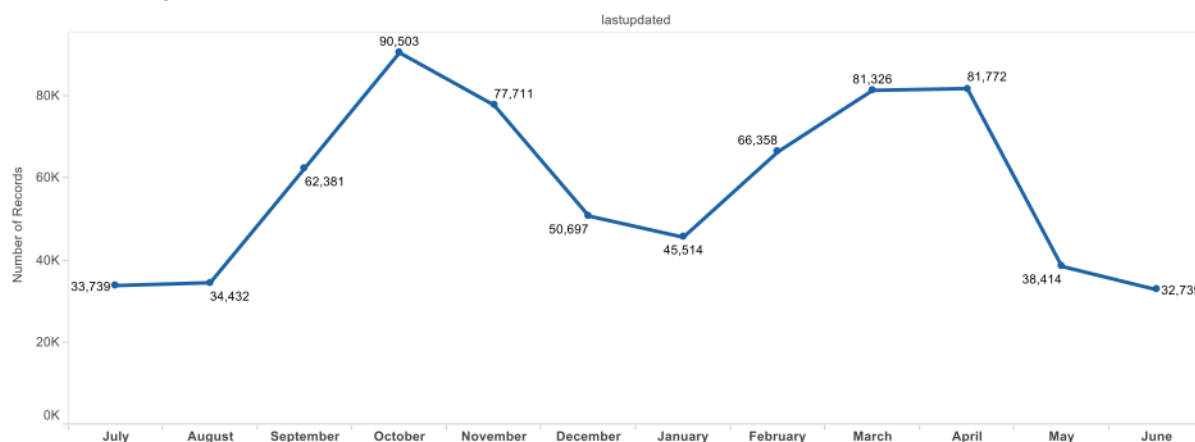


FIGURE 5B. MONTHLY MINES SURVEY AND ALL PROXY TRAFFIC

Using dashboards to display survey results visually helped the staff absorb and understand the survey results. The side-by-side presentation makes it easier to see relationships between responses within a single holistic view. The MINES User Group Dashboard (see Figure 6) shows the summary of all results. This broad overview provides the baseline for deeper analysis. For example, it is useful to see that 89% of overall use occurs outside the library buildings and that 53% of use was in support of teaching or classwork. Visualizing MINES data with Tableau increases utility because the data is easily filtered to answer a range of detailed questions posed by individual staff. A liaison librarian to engineering can see that 92% of graduate student use happens outside the libraries and that 45% of use is for thesis or dissertation work (see Figure 7). This kind of close analysis of the questions informs instruction, outreach, and support to constituents. The power comes not only from aggregate data or from a single conclusion but also from the ability to understand specific and integrated aspects of the data as needed for various purposes.

MINES User Group Dashboard

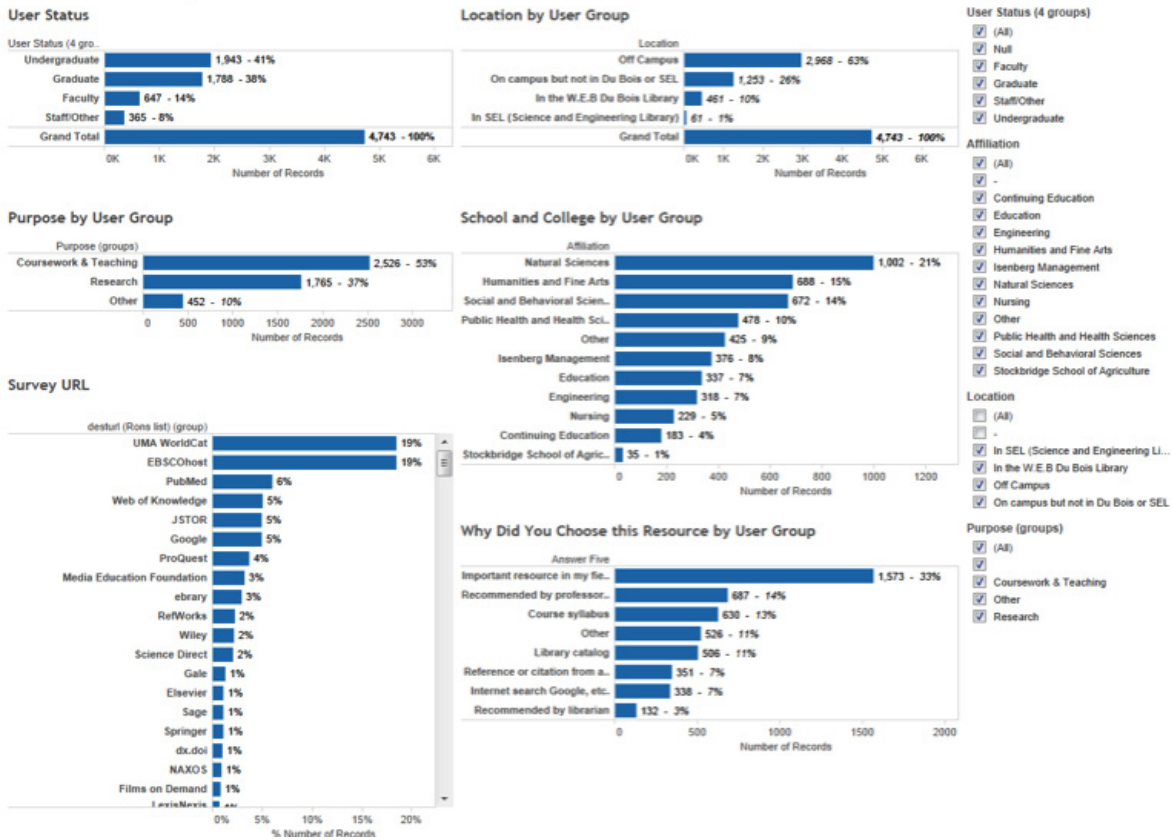
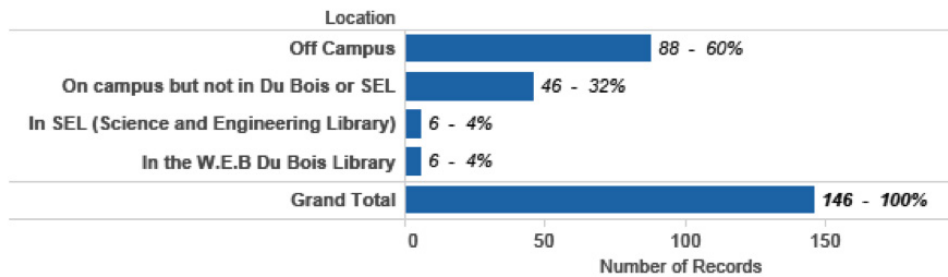


FIGURE 6. DASHBOARD SUMMARY OVERVIEW

Location by User Group - Graduate Students in Engineering



Purpose by User Group - Graduate Students in Engineering

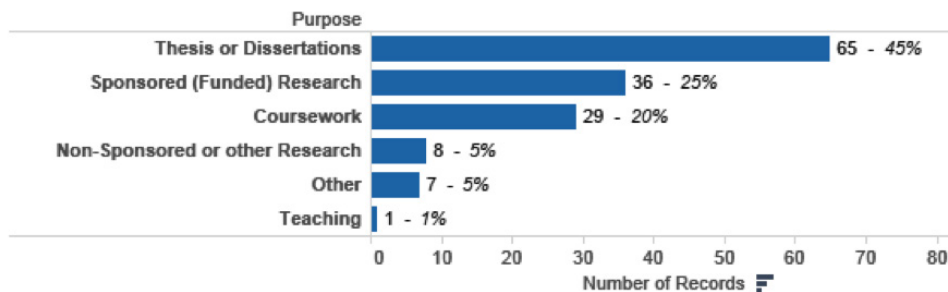


FIGURE 7. LOCATION AND PURPOSE OF USE FILTERED BY GRADUATE STUDENTS IN ENGINEERING

Finally, one of the most interesting comparisons made using Tableau was to ingest COUNTER data for the Web of Knowledge into Tableau both for record views in the database and regular searches (see Figure 8).

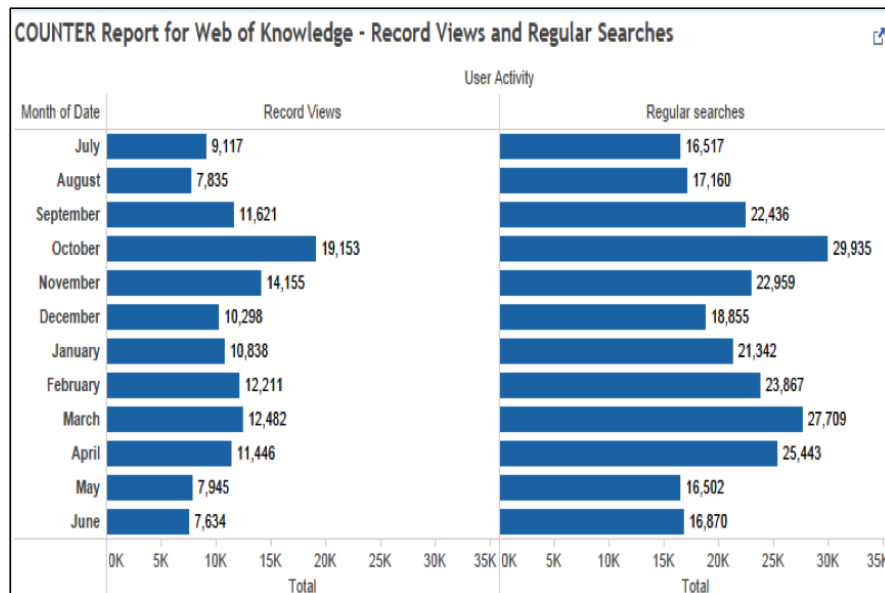


FIGURE 8. COUNTER DATA FOR WEB OF KNOWLEDGE DISPLAYED BY TABLEAU

The COUNTER data is useful but does not give the granularity of MINES data for the same resources. See Figure 9, which shows Web of Knowledge use by user type, purpose, school and college affiliation, the purpose of use, and the reason why the resource was selected.

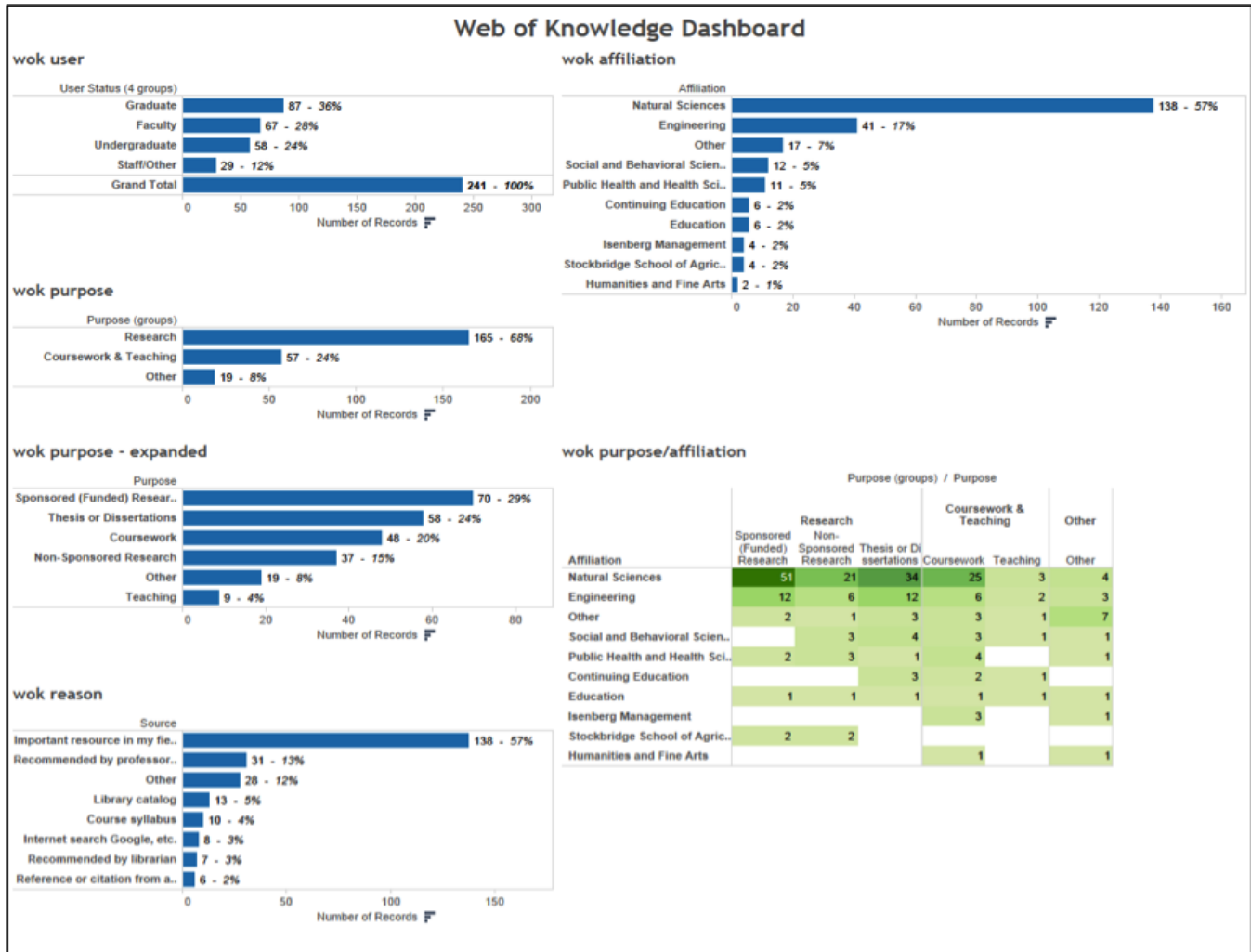


FIGURE 9. MINES DATA FOR WEB OF KNOWLEDGE DISPLAYED BY TABLEAU

At a glance it is easy to see that graduates students are the largest category of users and that the Web of Knowledge is used most heavily by patrons in the College of Natural Sciences. The table with shaded cells shows affiliation by purpose of use and the dark cells show that Natural Sciences uses the Web of Knowledge for both sponsored and non-sponsored research. The empty white cells are also easily seen. This level of granular detail, which can be filtered even further, is pertinent to designing library instruction programs, making collection development decisions, and marketing resources.

Implications

The UMass Amherst Libraries will continue to run the survey for the foreseeable future. The successful implementations, high response rate, lack of negative feedback, and the utility of the data have resulted in the decision to collect survey data continuously. The libraries will also explore expanding the survey scope to include other resources. Moving from an annual sample to continuous data collection provides current and ongoing data that is available to be analyzed alongside other data such as circulation,

interlibrary loan, building use, and network access; truly the culture of assessment as described by Lakos and Phipps.²¹

The relationship between MINES and COUNTER data may also be further explored. Specifically, it would be useful to determine how much electronic resource use is not captured by MINES. It may also be informative to examine the relationship between MINES, COUNTER, and vendor data more closely. It is challenging for some libraries to integrate separate COUNTER reports into a holistic picture of e-resource use, and one advantage to MINES is that all use is collected in a single data set. The MINES data set can be combined with COUNTER data in the Tableau environment for a deeper and more granular view of COUNTER data. Some libraries without access to Tableau might request of their electronic resource management (ERM) vendor the ability to import MINES data into their ERM to achieve a similar analysis. This deeper picture of patron activity is achieved anonymously without the need to track the path of individual patrons through various library and university systems and the associated implications for confidentiality, privacy, or ethics.

Conclusion

This paper contrasts the implementation of the two sampling designs for the ARL MINES for Libraries protocol at the same library, the University of Massachusetts in Amherst, discussing the advantages and disadvantages of each. The findings demonstrate why running the every Nth MINES study continuously is a good idea for libraries, and we described a valid and reliable implementation scheme using EZproxy. We show how using Tableau Software to analyze MINES data permits adjustments to data collection in real time, for example, by changing the frequency of N to answer new research questions. We compared the sampled data on certain variables to the population of data collected at the EZproxy server in the Starting Point URL (SPU) log files to show that the sample can be relied upon for valid inferences about the population. We illustrated how Tableau can present data relationships that can lead to decisions and actions by the library. We demonstrated how, with one vendor as an example, the joining of MINES data to COUNTER data can enhance the picture of how the resource is used and therefore, how the library might better serve patrons who consult that resource. Finally, we proposed that collecting MINES data continuously will lead to future service enhancements, especially if the data is imported into data visualization software, like Tableau, which makes the data easier to analyze, understand, and communicate. We anticipate that we will continue to collect valuable, actionable data to present a comprehensive picture of e-resource use to library and campus stakeholders, specifically informing collection development, instruction, support for research, marketing, and liaison work.

Endnotes

1. COUNTER, "COUNTER: Counting Online Usage of Networked Electronic Resources," accessed September 14, 2014, <http://www.projectcounter.org/>.
2. COUNTER, "COUNTER Code of Practice for e-Resources: Release 4," published April 2012, <http://www.projectcounter.org/r4/COPR4.pdf>.

3. Jill Emery and Graham Stone, "Annual Review," *Library Technology Reports* 49, no. 2 (February/March 2013): 30–34, doi: 10.5860/ltr.49n2.
4. ProQuest, "360 Counter," accessed September 14, 2014, <http://www.proquest.com/products-services/360-Counter.html>.
5. EBSCO, "EBSCO Usage Consolidation," accessed September 14, 2014, <http://journals.ebsco.com/products-services/ebsco-usage-consolidation>.
6. JUSP: Journal Usage Statistics Portal, "JUSP: Journal Usage Statistics Portal," accessed September 14, 2014, <http://jusp.mimas.ac.uk/>.
7. Emery and Stone, "Annual Review," 34.
8. Deanna Marcum and Roger C. Schonfeld, *Driving with Data: A Roadmap for Evidence-Based Decision Making in Libraries* (New York: Ithaka S+R, 2014), <http://www.sr.ithaka.org/blog-individual/driving-data-roadmap-evidence-based-decision-making-academic-libraries>.
9. Association of Research Libraries, "MINES for Libraries®," accessed September 14, 2014, <http://www.minesforlibraries.org/home>.
10. Marcum and Schonfeld, *Driving with Data*, 7.
11. Tableau Software, "Tableau," accessed September 14, 2014, <http://www.tableausoftware.com/>.
12. Association of Research Libraries, "MINES for Libraries."
13. OCLC, "EZproxy®," accessed September 14, 2014, <https://oclc.org/ezproxy.en.html>.
14. ExLibris, "SFX—The OpenURL Link Resolver and Much More," accessed September 14, 2014, <http://www.exlibrisgroup.com/category/SFXOverview>.
15. Dana Thomas, Catherine Davidson, Martha Kyrillidou, and Terry Plum, "Measuring Use of Licensed Electronic Resources: A Second Iteration of the MINES for Libraries® Survey on Scholars Portal and Other Resources for the Ontario Council of University Libraries," *Library Management* 33, no. 6/7 (2012): 374–388, <http://www.emeraldinsight.com/doi/abs/10.1108/01435121211266195>.
16. Ibid.
17. Terry Reese, "Improving Evaluation of Resources through Injected Feedback Surveys," *Ariadne* 71 (2013), <http://www.ariadne.ac.uk/issue71/reese>.
18. Andrea Miller, "Application of Excel® Pivot Tables and Pivot Charts for Efficient Library Data Analysis and Illustration," *Journal of Library Administration* 54, no. 3 (2014): 169–186, doi: 10.1080/01930826.2014.915162.
19. Tableau Software, "Tableau."

20. Sarah Murphy, Jeremy Buhler, and Rachel Lewellen, "Tableau Unleashed: Visualizing Library Data," panel presented at the Library Assessment Conference, Seattle, Washington, August 4–6, 2014; Katy Mathuews and Rebekah Kilzer, "Driving Partnerships for Assessment and Engagement," paper presented at the Library Assessment Conference, Seattle, Washington, August 4–6, 2014; and Sarah Anne Murphy, "Data Visualization and Rapid Analytics: Applying Tableau Desktop to Support Library Decision-Making," *Journal of Web Librarianship* 7, no. 4: 465–476, doi: 10.1080/19322909.2013.825148.
21. Amos Lakos and Shelley Phipps, "Creating a Culture of Assessment: A Catalyst for Organizational Change," *portal: Libraries and the Academy* 4, no. 3 (2004): 345–361, http://muse.jhu.edu/login?auth=0&type=summary&url=/journals/portal_libraries_and_the_academy/v004/4.3lakos.html and <http://escholarship.org/uc/item/0843106w.pdf>.

© 2016 Rachel Lewellen and Terry Plum



This article is licensed under a Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

To cite this article: Rachel Lewellen and Terry Plum. "Assessment of E-Resource Usage at University of Massachusetts Amherst: A MINES for Libraries® Study Using Tableau for Visualization and Analysis." *Research Library Issues: A Report from ARL, CNI, and SPARC*, no. 288 (2016): 5–20. <http://publications.arl.org/rli288/>.

Editor's note: An earlier version of this paper appeared in Sarah Durso, Steve Hiller, Martha Kyrillidou, and Angela Pappalardo, *Proceedings of the 2014 Library Assessment Conference: Building Effective, Sustainable, and Practical Assessment, August 4–6, 2014, Seattle WA* (Washington, DC: Association of Research Libraries, 2015), <http://libraryassessment.org/bm~doc/proceedings-lac-2014.pdf>.

Tableau Unleashed: Visualizing Library Data

Jeremy Buhler, Assessment Librarian, University of British Columbia

Rachel Lewellen, Assessment Librarian, University of Massachusetts Amherst

Sarah Anne Murphy, Coordinator of Assessment, The Ohio State University Libraries

Tableau is rapid-analytics and data-visualization software that supports library assessment by enabling a library to query, explore, and visualize data in real time. Using Tableau, a library may produce flexible, in-depth, online dashboards, complete with filters and annotations to both customize visualizations and provide context. A library may also blend data from disparate sources to create dynamic, interactive graphics and reports.

As we prepared our panel presentation for the 2014 Library Assessment Conference, we realized that Tableau's value to academic libraries may best be demonstrated via show-and-tell. We used the following questions to guide our discussion:

- Discuss how your library has incorporated Tableau into its assessment program.
- What impact has Tableau had on your ability to make sense of large data sets, make data accessible, and improve stakeholder communications?
- Where does Tableau fit in your library's data strategy?

Tableau at The Ohio State University Libraries—by Sarah Anne Murphy

Tableau is a key tool used by The Ohio State University (OSU) Libraries assessment program. The software enhances the libraries' ability to aggregate data and to assemble data from various library systems into meaningful packages for library decision makers. It is a key component of the libraries' strategy to gather, process, and make data available to both the libraries' internal and external stakeholders.

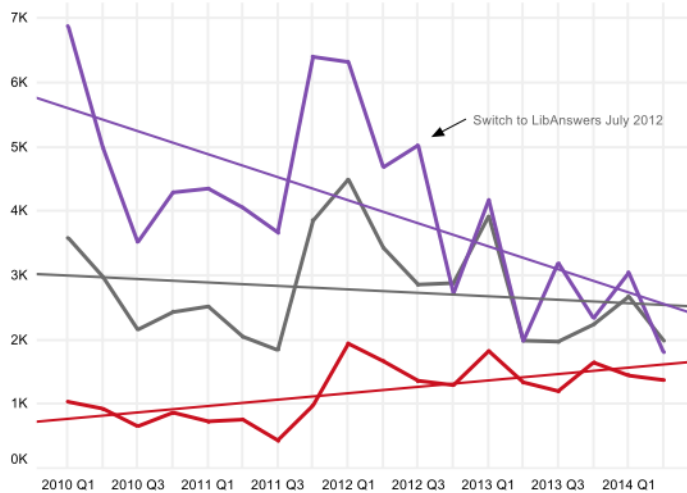
I discovered Tableau in spring of 2012, and quickly realized its potential for not only analyzing and visualizing library data, but for gathering, repackaging, and delivering library data in a timely manner to inform decision making.

Research Services Trends

Figure 1 showcases a Research Services Trends dashboard that is updated quarterly for the OSU Libraries Research and Education division's quarterly report. This dashboard is freely available to all OSU librarians and staff and was created using Tableau's Desktop Personal software, a production tool that is currently discounted for educators. The dashboard was posted to the web via Tableau Public, a free service that allows users to share Tableau visualizations online.¹

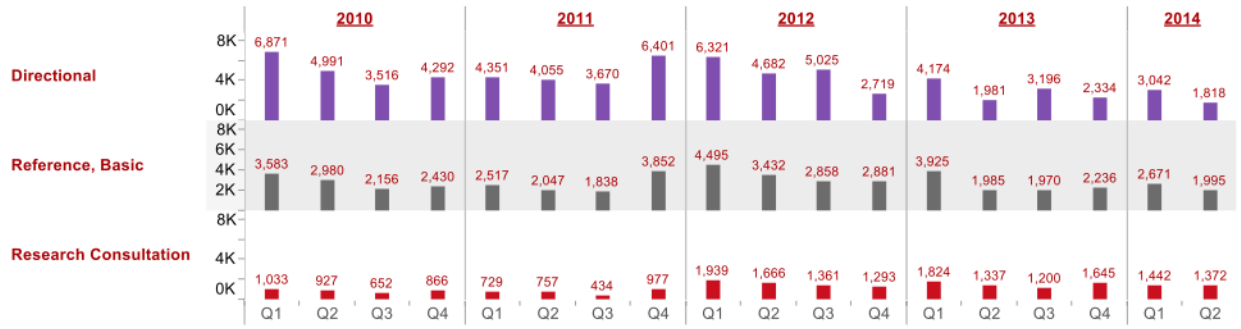


Research Services Trends
By Question Type and Quarter, 2009 to Present



| | Directional | Reference, Basic | Research Consultation | Grand Total |
|--------------------|---------------|------------------|-----------------------|----------------|
| 2010 | 19,670 | 11,149 | 3,478 | 34,297 |
| 2011 | 18,477 | 10,254 | 2,897 | 31,628 |
| 2012 | 18,747 | 13,666 | 6,259 | 39,145 |
| 2013 | 11,685 | 10,116 | 6,006 | 28,493 |
| 2014 | 4,852 | 4,663 | 2,814 | 12,428 |
| Grand Total | 73,431 | 49,848 | 21,454 | 145,991 |

QuestionType
■ Directional ■ Reference, Basic ■ Research Consu..



Sarah Murphy, OSUL Assessment, July 2, 2014..

FIGURE 1. RESEARCH SERVICES TRENDS DASHBOARD

The Research Services Trends dashboard presents the libraries’ data for directional, reference, and research consultations in three different ways, allowing staff to visually piece together changes in user behavior over time. The trend lines inserted into the line graph on the top left, for instance, reveal that while the number of directional questions asked at the OSU Libraries Columbus campus locations has declined, the number of research consultations provided by OSU librarians has significantly increased. Further, the visualization annotates when the OSU Libraries switched from an in-house mechanism for recording reference transactions to LibAnswers. This change may have influenced some of the drop in directional questions due to some implementation challenges. A text table listing the same data by year is provided on the top right, and a bar chart showing the number of questions by quarter is available underneath. Overall questions spike during the first and fourth calendar quarter of every year, which is not surprising considering the OSU academic calendar.

The three visualizations are linked using a global filter. This allows librarians and staff to highlight “Research Consultations” in the question type legend and view this data in isolation. This is a particularly useful feature when librarians or staff use dashboards to talk about, or to advocate for, library

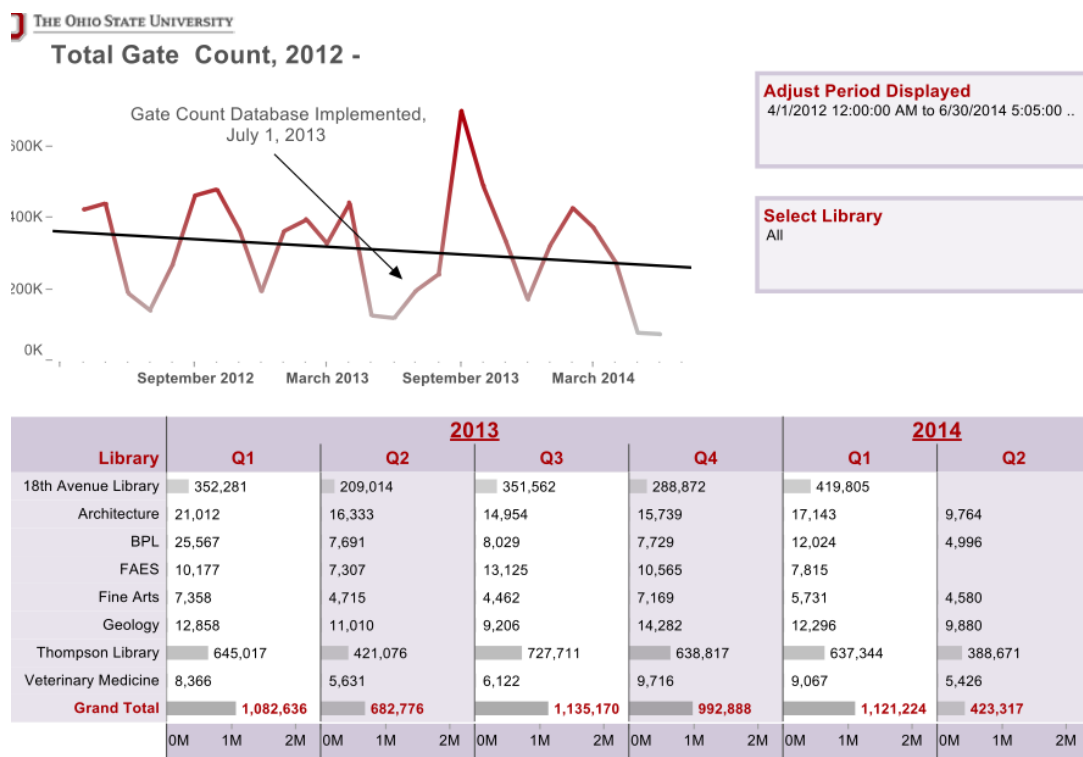
services with external stakeholders. Librarians and staff may also copy and paste any element of the dashboard into an e-mail or document.

Tableau Public visualizations may be downloaded to a local PC, making the raw and aggregated data for the visualizations on this dashboard freely available to librarians and staff. Therefore it is important to disclaim that private, confidential information should not be shared via dashboards uploaded to Tableau Public. The OSU Libraries annually submit reference transactional data to the Association of Research Libraries (ARL). Thus, the information provided in Figure 1 is publically available through the annual ARL Statistics publication, just not at the level of detail or with the same immediacy provided by the dashboard.²

Tableau offers librarians the ability to blend data from multiple database platforms and software packages. The Research Services Trends dashboard is populated with data from a number of sources, including LibAnswers and previous incarnations of the OSU Libraries' Ask Database, an internal system the libraries once used to record reference transactional data.

Gate Count

A dashboard with the aggregated library gate count, broken down by library location is provided in Figure 2.³ Using this visualization, librarians and staff may adjust the time period displayed, or choose to view only the data for a selected library location.



Sarah Murphy, OSUL Assessment, July 2, 2014
Data Source: Gate Count Database

FIGURE 2. GATE COUNT DASHBOARD

Thus, if we select Veterinary Medicine from the “Select Library” pull-down menu, only data for the Veterinary Medicine Library will display on the screen. The trend line will also recalculate using the data for the Veterinary Medicine Library only.

ILLiad Borrowing, 2010–

In the spring of 2013, the OSU Libraries assembled a five-member project team to explore the potential application of Tableau within the OSU Libraries. The Visualizing ILLiad team was co-led by the assessment coordinator and the head of interlibrary loan and included subject librarians from the Research Services and Area Studies departments. Together team members identified questions of interest to subject librarians that might be answered with ILLiad transactional data, and then built and tested two dashboards to allow subject librarians to interact with and understand borrowing trends for their assigned departments to better inform their collection activities. Questions included:

- Who is borrowing what titles? How often? When? (Who includes patrons and institutions)
- What are faculty affiliated with interdisciplinary centers borrowing?
- Can graduate student borrowing be segmented by academic program?

Figure 3 shows the aggregate number of patron borrowing requests for departments served by one of the OSU Libraries’ science librarians. The map on the top left of the screen shows that OSU primarily borrows materials from its Committee on Institutional Cooperation (CIC) partners for astronomy, chemistry, engineering, and physics students and faculty. The bars in the lower left visually segment borrowing requests by department, year, and month for 2012 and 2013. The “Format” text table on the lower right is fully interactive. If you click on “Book,” for example, a full list of titles borrowed during the time period specified is returned, broken down by user department.

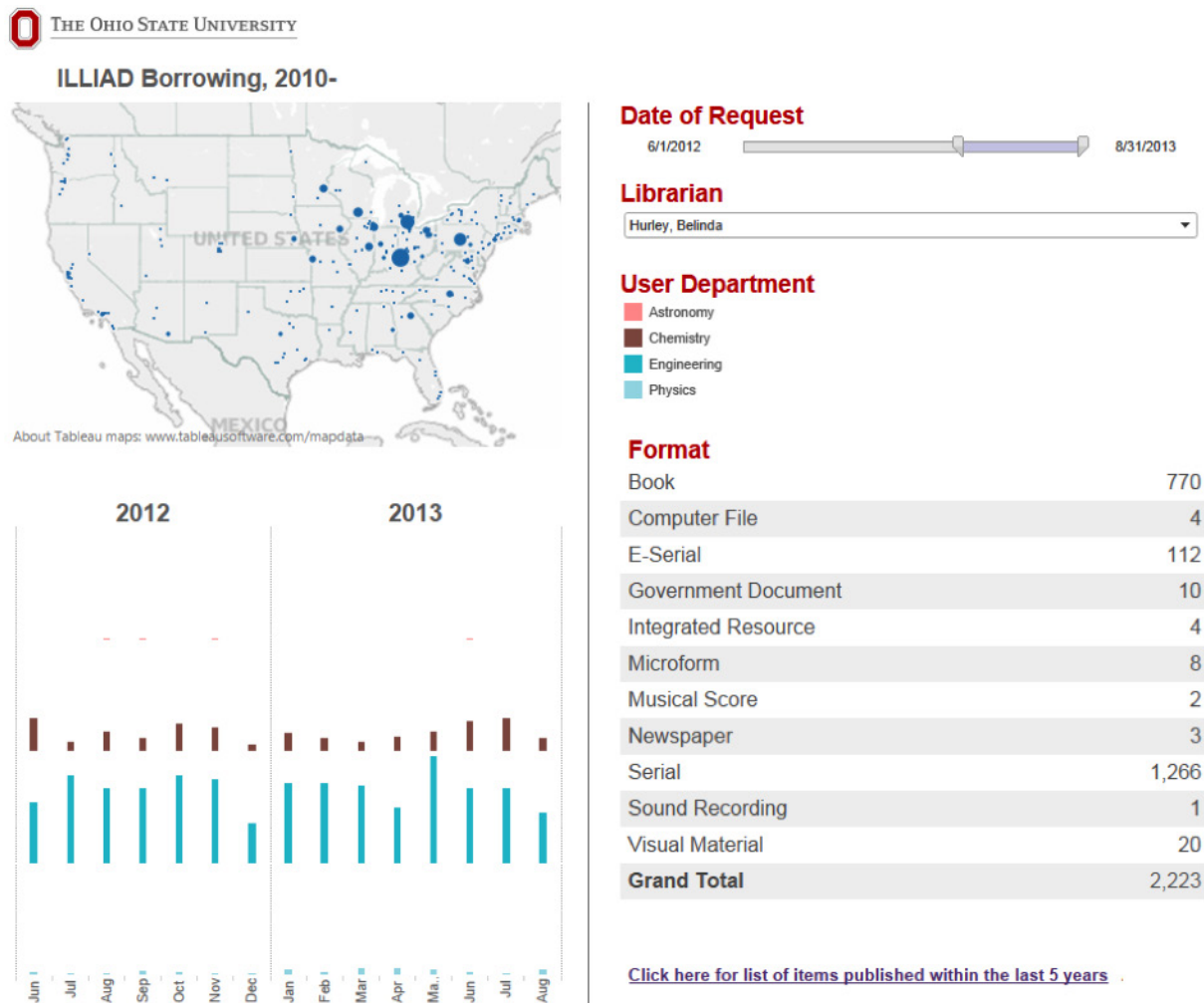
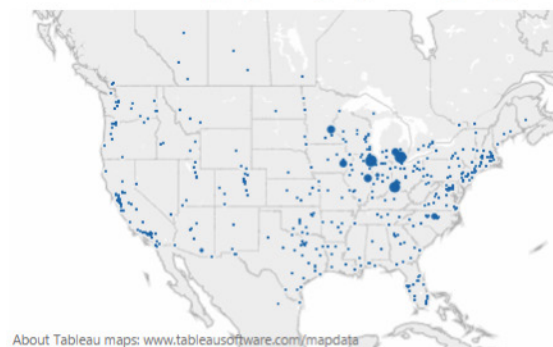


FIGURE 3. DASHBOARD OF ILLIAD BORROWING, 2010-

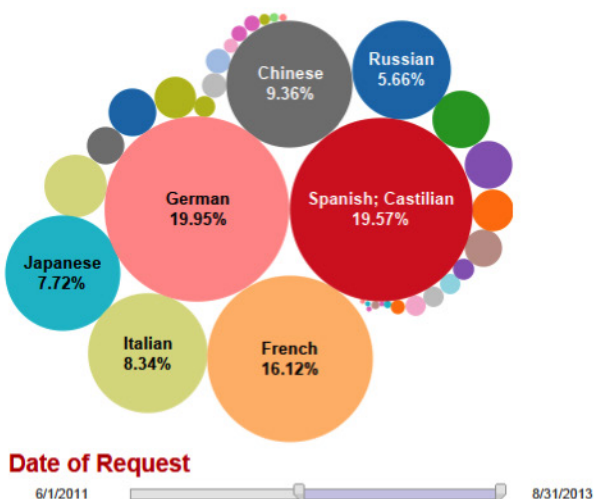
We quickly realized that this approach failed to provide serviceable data for interdisciplinary areas, such as Jewish studies. To address this issue, the team constructed a second dashboard using data queried and blended from ILLiad, Sierra, and a number of other sources, and then filtered the data using non-English languages. The resulting dashboard in Figure 4 is more useful for our area studies librarians, who serve users across a number of academic disciplines. The map on the upper left illustrates that the OSU Libraries borrow non-English materials from a more diverse population of libraries across the nation, while the bubble chart on the lower left highlights that German-language materials are requested the most frequently, followed by Spanish, and then French. The text table on the right is fully interactive. Thus, if our Jewish studies librarian clicks on “Hebrew,” he will obtain a more robust list of titles requested by patrons during the same time period.



ILLIAD Borrowing, By Languages Taught @ OSU, 2010 -



About Tableau maps: www.tableausoftware.com/mapdata



(click on a language for a list of titles borrowed)

| | |
|---|-------|
| Arabic | 245 |
| Aymara | 2 |
| Bosnian | 5 |
| Bulgarian | 19 |
| Catalan; Valencian | 31 |
| Chinese | 1,198 |
| Croatian | 45 |
| Czech | 33 |
| French | 2,064 |
| German | 2,554 |
| Greek, Modern | 176 |
| Hebrew (modern) | 104 |
| Hindi | 18 |
| Hungarian | 17 |
| Italian | 1,068 |
| Japanese | 989 |
| Korean | 169 |
| Latin | 126 |
| Old Church Slavonic, Church Slavic, Church Slav.. | 3 |
| Persian (Farsi) | 31 |
| Polish | 129 |
| Portuguese | 291 |
| Quechua | 3 |
| Romanian, Moldavian(Romanian from Republic of.. | 30 |
| Russian | 725 |
| Serbian | 15 |
| Somali | 6 |
| Spanish; Castilian | 2,506 |
| Swahili | 4 |
| Swedish | 33 |
| Tibetan Standard, Tibetan, Central | 2 |
| Turkish | 106 |
| Urdu | 9 |
| Uzbek | 4 |
| Yiddish | 45 |

[Click here for list of items published within the last 5 years](#)

FIGURE 4. DASHBOARD OF ILLIAD BORROWING, ALL LANGUAGES, 2010-

Tableau at the UMass Amherst Libraries—by Rachel Lewellen

Tableau is a major component of the assessment program at the University of Massachusetts (UMass) Amherst Libraries. The libraries were challenged to make sense of multiple data sources in a variety of formats and needed an increased capacity to visualize, organize, analyze, and share data. The libraries pursued a strategy of data visualization using business intelligence software (Tableau) when they determined that a comprehensive data warehouse within the library was not a feasible option.

Staff use visualizations to support decision making related to collections, services, and facilities. The ability to integrate and query multiple data sets also supports expectations related to campus goals, accountability, planning, and assessment. The following two examples show a range of visualizations and applications.

Monograph Purchasing, Circulation, and Duplication—Micro and Macro Analysis

The ability to build a variety of views from a single rich data set allows for meaningful customization. Figure 5 displays a sample dashboard that visualizes data from the ALEPH integrated library system. Individual selectors review current and historical data about monograph purchases, including the number of items purchased, expenditures with circulation status, and duplication within consortial collections. Selectors filter the view by fiscal year and the appropriate order group or budget code. Aggregate and individual title-level detail is available.

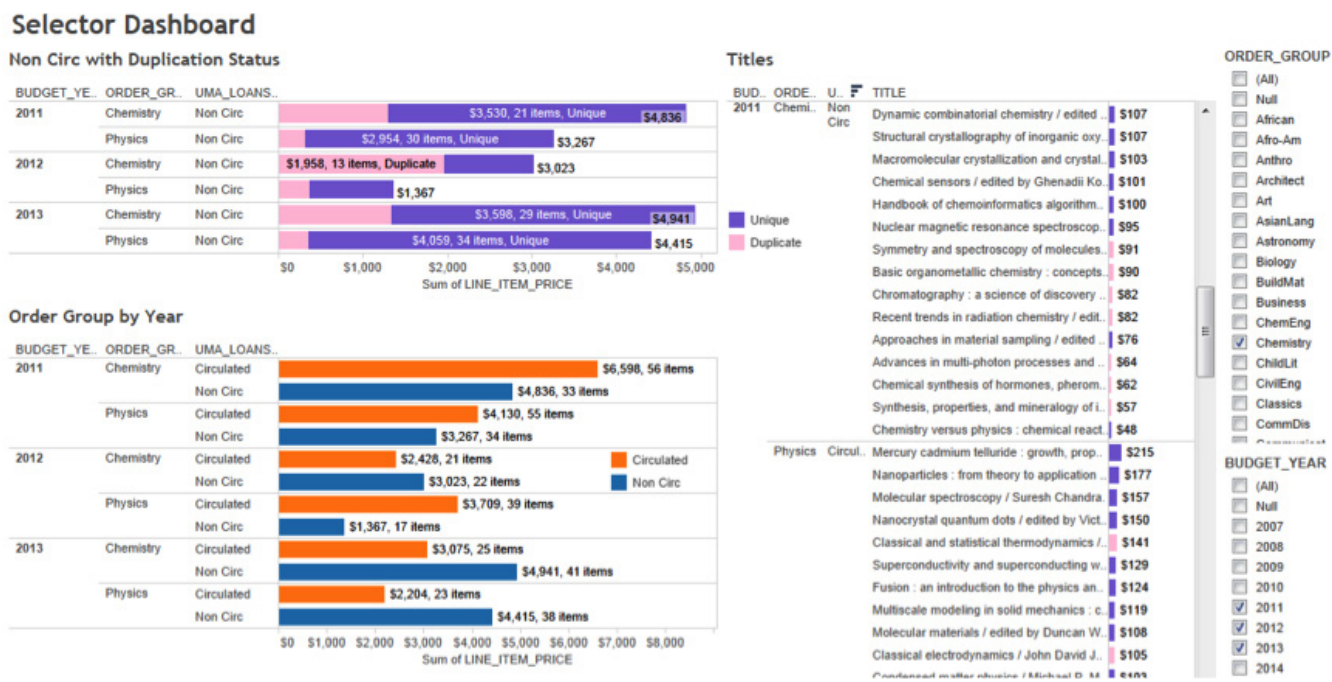


FIGURE 5. DASHBOARD OF PURCHASING AND CIRCULATION (ALEPH)

Broad collection-level analysis is also possible by examining the distribution and use of monographs by Library of Congress classification, school and college allocations, specialized purchasing program performance, or for the collection as a whole (see Figure 6). This data informs conversations and decisions with library staff and campus stakeholders regarding budget allocations and collection development policy.

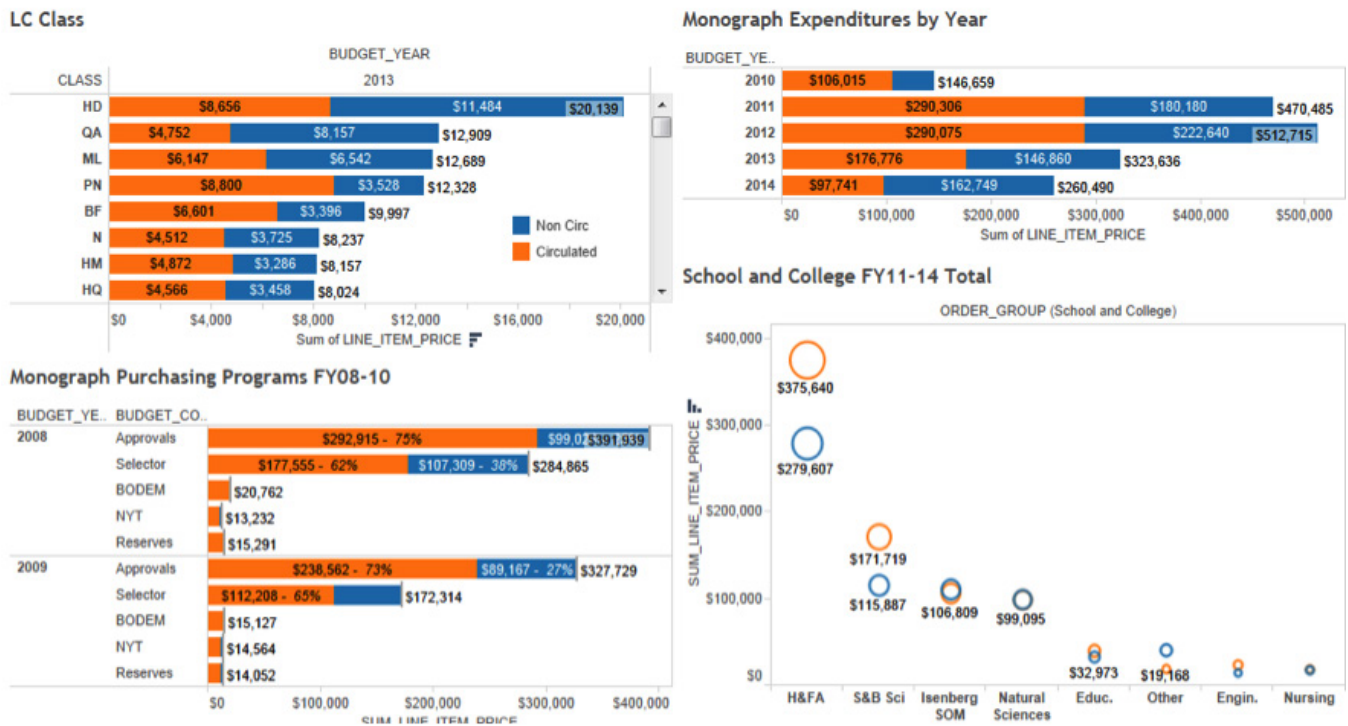


FIGURE 6. LIBRARY OF CONGRESS CLASS, PURCHASE PROGRAM, FISCAL YEAR TOTAL, AND SCHOOL AND COLLEGE EXAMPLES

E-Book Library (EBL) Pilot Project

UMass Amherst participated in a consortial patron-driven acquisition project that offered a wide pool of e-book titles across the Five College Consortium libraries (Amherst College, Hampshire College, Mount Holyoke College, Smith College, and UMass Amherst). Each participating library needed both institution and consortial data to monitor and evaluate use and expenditures. The ability to filter and share data through a web browser eliminated the need for spreadsheets to be repeatedly and individually manipulated. Uniform interaction with the data provided a common framework for discussion (see Figure 7).

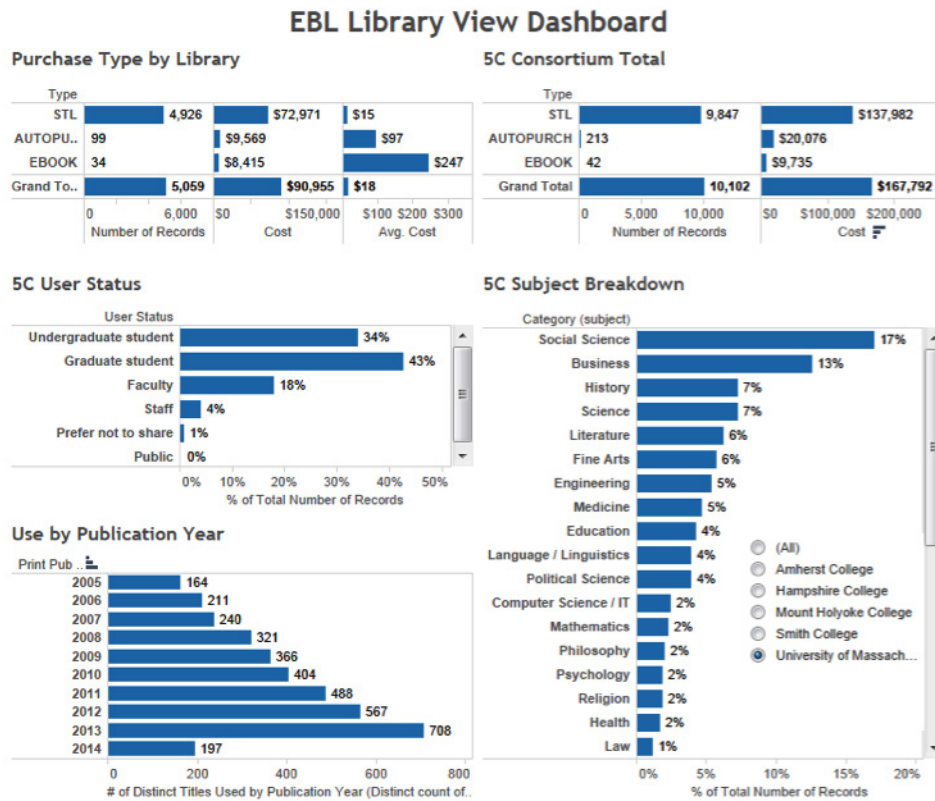


FIGURE 7. EBL OVERVIEW BY LIBRARY

As the project progressed, the participating libraries adjusted pilot project parameters related to loan period and price thresholds in response to the significant increases of short-term loan costs from publishers. Expenditures were projected using a range of short-term loan trigger scenarios and then graphically displayed. While the horizontal bar chart visualization at the top of Figure 8 is dense and complex, it makes it easier to understand the relationship between scenarios in comparison to the spreadsheet table below it.

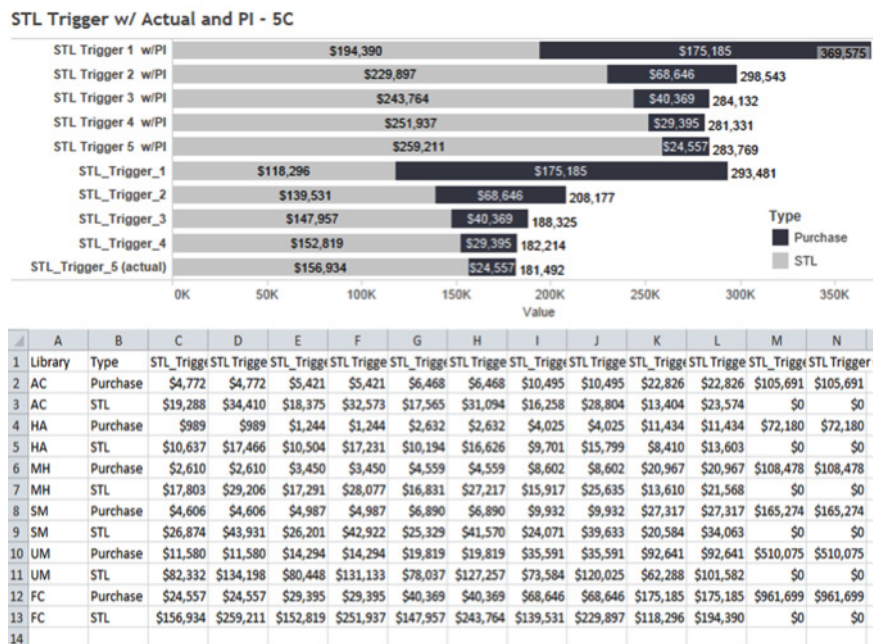


FIGURE 8. GRAPHIC VISUALIZATION CONTRASTED WITH SPREADSHEET TABLE DISPLAY

The dashboards displayed in Figures 7 and 8 were central to reaching a shared understanding of the financial implications and consortial decisions related to the pilot project.

Tableau at the University of British Columbia Library—by Jeremy Buhler

Using Tableau to Explore the Data

The above examples from The Ohio State University and UMass Amherst Libraries focus on Tableau as a publishing and data-sharing platform. This third section describes Tableau’s potential as a tool for data exploration.

Part of the assessment librarian’s role at the University of British Columbia (UBC) is to make management and user-experience data more accessible to those who need it to inform decisions. But providing timely access to data is only part of the picture and means little unless the audience is also engaged with the data presented.

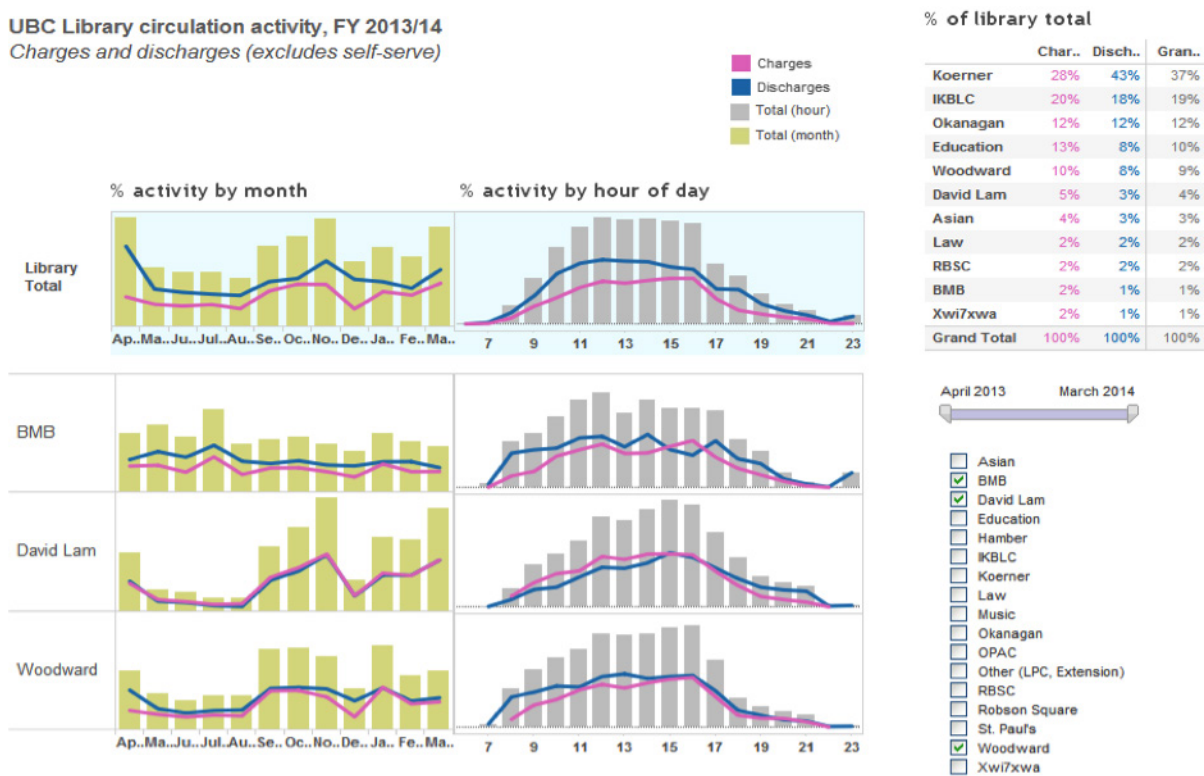
Stephen Few, an expert in the field of visual perception and dashboard design, provides guidelines for data presentation in his book *Information Dashboard Design: The Effective Visual Communication of Data*. In general a dashboard will be more effective if it is focused on fulfilling a specific data need,⁴ and if done well it may also prompt new questions from the audience. These new questions are one measure of engagement, but to sustain engagement with the data and reward the audience for asking deeper and potentially more fruitful questions we need tools that can quickly shuffle and re-package the source material to respond to new lines of inquiry.

One of the strengths of Tableau as a data visualization platform is that it makes it relatively easy to aggregate, re-package, and display source data. The sections that follow provide two UBC examples to illustrate this point. The data sets themselves are commonplace but what I hope will spark your own curiosity and sense of possibility is the way Tableau makes it easier to navigate and interpret the data.

Visualizing Circulation Data

The first example is based on loan and discharge data from the UBC Library ILS (see Figure 9). The data was initially pulled to help answer a question about the distribution of the circulation workload across library branches.

UBC Library circulation activity, FY 2013/14
Charges and discharges (excludes self-serve)



Jeremy Buhler, UBC Library Assessment. Data source: Voyager ILS

FIGURE 9. UBC LIBRARY CIRCULATION ACTIVITY, FY 2013/14

This report provides a high-level overview of circulation activity at multiple branches over a single year, with bar charts showing the distribution by hour of day and by month of year for each location. The blue lines represent discharges (items returned), the pink lines represent items being checked out, and the bars are the sum of the two. The height of the bars represents the percentage of the annual total in any given month or hour, and by stacking graphs for different branches it is possible to compare workload distribution patterns from one location to another at a glance.

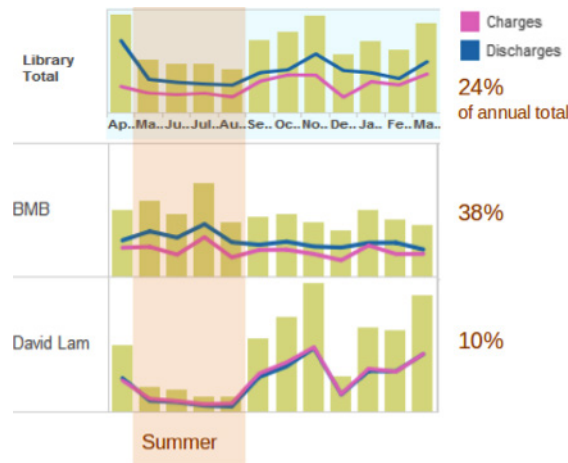


FIGURE 10. UBC LIBRARY CIRCULATION: % OF ACTIVITY BY MONTH AND LIBRARY BRANCH

Note in particular the four summer months displayed in Figure 10: there is less activity from May through August at all locations except the Biomedical Branch (BMB). With Tableau it is possible to quickly view this level of detail for all branches, helping managers make informed decisions about resource allocation across multiple locations.

Another way of viewing the same data set is by the percentage of daily work distributed across the hours of the day (Figure 11).

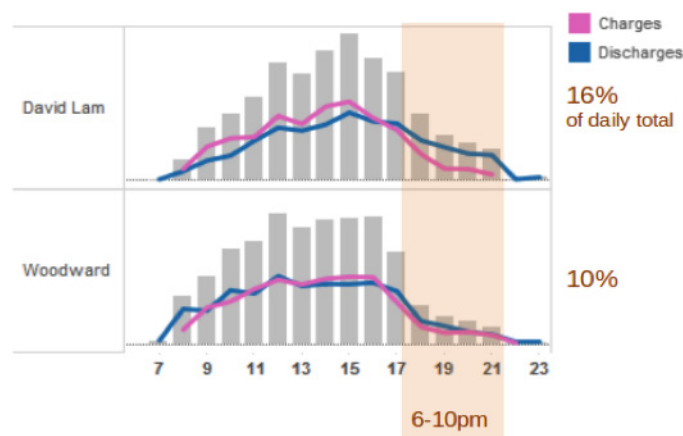


FIGURE 11. UBC LIBRARY CIRCULATION: % OF DAILY ACTIVITY BY HOUR AND LIBRARY BRANCH

Both of the David Lam and Woodward locations are open until 10:00 p.m. but only a small percentage of daily activity falls within the service hours of 6:00 and 10:00 p.m. Notice how sharply the bars drop after 6:00 p.m. at the Woodward Library branch. Now compare this to the slightly less acute post-6:00 p.m. shift at the David Lam branch. Those four hours account for only 10% of the daily circulation activity at Woodward but 16% at the David Lam branch. All else being equal, David Lam circulation staff maintain a higher activity level between 6:00 and 10:00 p.m., but do these graphs help us understand why?

Because the graphs also show detail about charges and discharges—the blue and pink lines—the figure points to a possible explanation. Notice in the graph at the top how the blue curve representing discharges is shifted to the right, or later in the day. This suggests that David Lam library staff do more of their daily discharge work in the slower evening period, potentially helping daytime staff remain available to users who visit the desk for in-person help. We cannot know from this data whether other factors account for the difference but the graphs support a hypothesis that merits further exploration and may help branches establish and share best practices.

Visualizing Results of the LibQUAL+® Survey

The second example from UBC relies on a data set that is familiar to many North American academic libraries: the LibQUAL+ survey. This is a rich data set, particularly when longitudinal data is available. In practice, however, the potential for examining change over time was not realized at UBC because summary data was often presented in formats that made comparisons time consuming.

One of UBC Library’s first experiments with Tableau was to reformat the raw data from three years’ worth of LibQUAL+ surveys. The resulting online report enables longitudinal comparison and makes it easier for library staff to view responses by user group and by LibQUAL+ question (Figure 12). A vertical orange band is used to represent the range between the average minimum and desired service levels, and a blue dot or line represents UBC Library’s perceived service level for a given question.

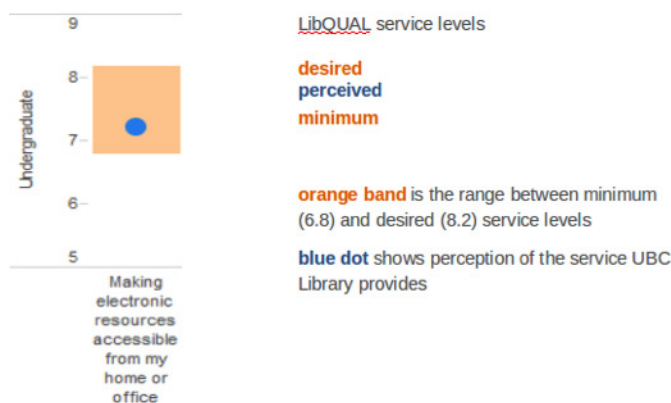


FIGURE 12. VISUAL REPRESENTATION OF LIBQUAL+ RESULTS FOR A SAMPLE QUESTION

UBC results for the 2013 LibQUAL+ survey identified “information control” as the dimension where the most improvement was needed to meet respondent expectations. But the “information control” dimension covers a wide range of activities and more detail is required to determine where in particular the library should focus its improvement efforts. Because the visualization is based on raw data rather than aggregated scores, Tableau makes it easy to drill further down and view scores for individual questions in this group simply by adding new dimensions to the display.

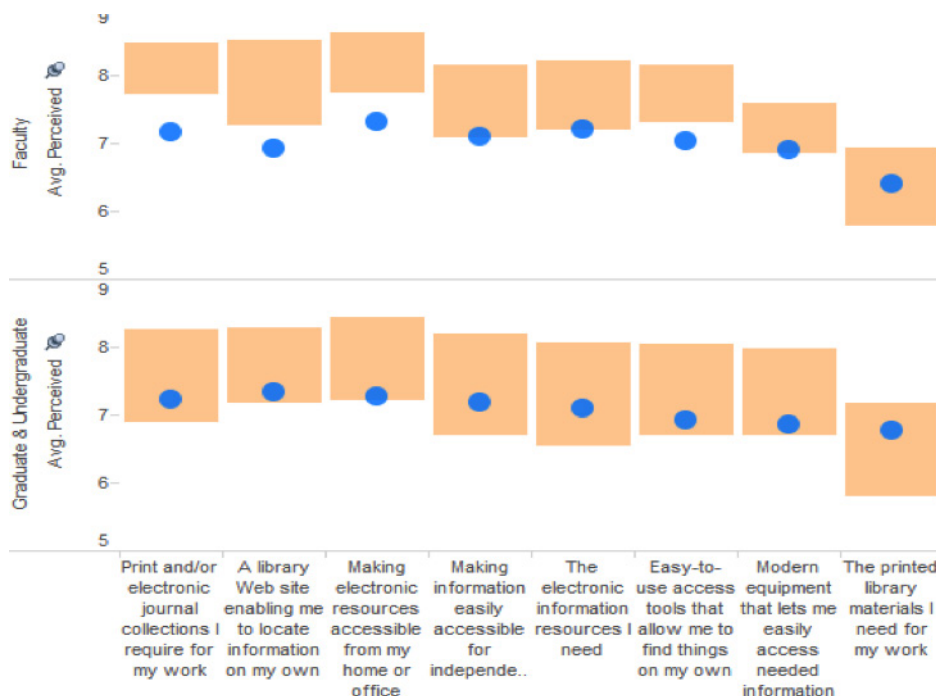


FIGURE 13. 2013 UBC LIBQUAL+ RESULTS: INFORMATION CONTROL QUESTIONS

Figure 13 displays the results for UBC faculty respondents in the top row and the results for UBC students in the bottom row. Questions are arranged from left to right by the average perceived service level, represented by the blue dots. When identifying priorities for improvement the areas where expectations are high and where perceived service level is near or below the minimum are usually the most important (these tend to be the questions displayed on the left).

In this case, however, I would like to highlight the question on the far right: in 2013 UBC respondents’ expectations were lowest when it came to “the printed library materials I need for my work.” Because the Tableau visualization is linked to longitudinal data it is possible to view how responses vary over time and by academic discipline—variations that may be particularly relevant as libraries shift from print to electronic monographs.

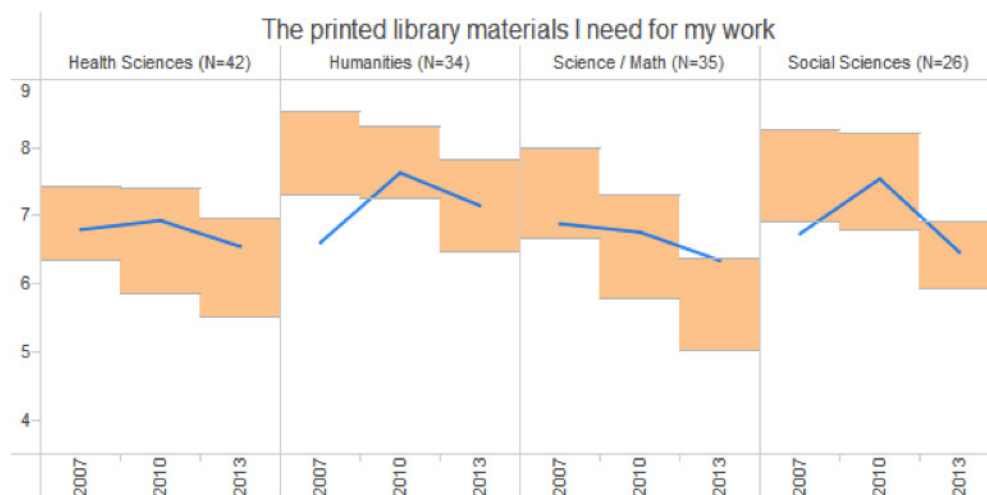


FIGURE 14. LONGITUDINAL UBC LIBQUAL+ RESULTS (N REFERS TO 2013 SURVEY)

In Figure 14 each orange band within a column represents a LibQUAL+ year: 2007 on the left, 2010 in the middle, and 2013 on the right. The downward stepping trend in each of the four schools (sometimes referred to as faculties) tells a story about changing expectations. For each group the acceptable service range has decreased steadily since the 2007 survey but there are differences in the pace of this change: respondents who identified themselves with humanities and social sciences are following the trend exhibited in the sciences with some lag time.

None of the LibQUAL+ visualizations presented here are based on data that is new to UBC Library, but Tableau helped to breathe new life into relatively commonplace data sets, making them more relevant to certain audiences. The result: as assessment librarian I can genuinely welcome requests to slice the data in different ways, supporting creative new applications for library data sets and, hopefully, a renewed sense of the potential in our existing data.

Endnotes

1. Sarah Murphy, “Research Services: LibAnswers 2012–,” The Ohio State University Libraries, accessed September 10, 2014, http://go.osu.edu/OSUL_ResearchServicesDashboard; Tableau, “Tableau Public Overview,” video, 21:01, accessed September 8, 2014, <http://www.tableausoftware.com/public/community>.
2. Association of Research Libraries, “ARL Statistics,” accessed September 10, 2014, <http://www.arlstatistics.org/home>.
3. Sarah Murphy, “University Libraries: Daily Gate Count,” The Ohio State University Libraries, accessed September 10, 2014, <http://go.osu.edu/GateCount>.
4. Stephen Few, *Information Dashboard Design: The Effective Visual Communication of Data* (Sebastopol, California: O’Reilly Media, 2006), 111.

© 2016 Jeremy Buhler, Rachel Lewellen, and Sarah Anne Murphy



This article is licensed under a Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

To cite this article: Jeremy Buhler, Rachel Lewellen, and Sarah Anne Murphy. "Tableau Unleashed: Visualizing Library Data." *Research Library Issues: A Report from ARL, CNI, and SPARC*, no. 288 (2016): 21–36. <http://publications.arl.org/rli288>.

Editor's note: An earlier version of this paper appeared in Sarah Durso, Steve Hiller, Martha Kyrillidou, and Angela Pappalardo, *Proceedings of the 2014 Library Assessment Conference: Building Effective, Sustainable, and Practical Assessment, August 4–6, 2014, Seattle WA* (Washington, DC: Association of Research Libraries, 2015), <http://libraryassessment.org/bm~doc/proceedings-lac-2014.pdf>.

Research Library Issues

A Report from ARL, CNI, and SPARC

Research Library Issues (RLI) focuses on current and emerging topics that are strategically important to research libraries. The articles explore issues, share information, pose critical questions, and provide examples. Ideas for articles are welcomed. Please send ideas to Sue Baughman, ARL deputy executive director, sue@arl.org.

Association of Research Libraries
21 Dupont Circle, NW, Suite 800
Washington, DC 20036
P: 202-296-2296
F: 202-872-0884

ISSN 1947-4911

Executive director: Elliott Shore
Editor: M. Sue Baughman
Assistant editor: Kaylyn Groves

© 2016 Association of Research Libraries

ARL policy is to grant blanket permission to reprint as long as full attribution is made. Exceptions to this policy may be noted for certain articles. This is in addition to the rights provided under sections 107 and 108 of the Copyright Act. For more information, contact ARL Publications pubs@arl.org.

Current and back issues are available on the ARL Digital Publications website publications.arl.org/rli. The website is also where you may sign up for alerts to new releases of *Research Library Issues*.