# 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?

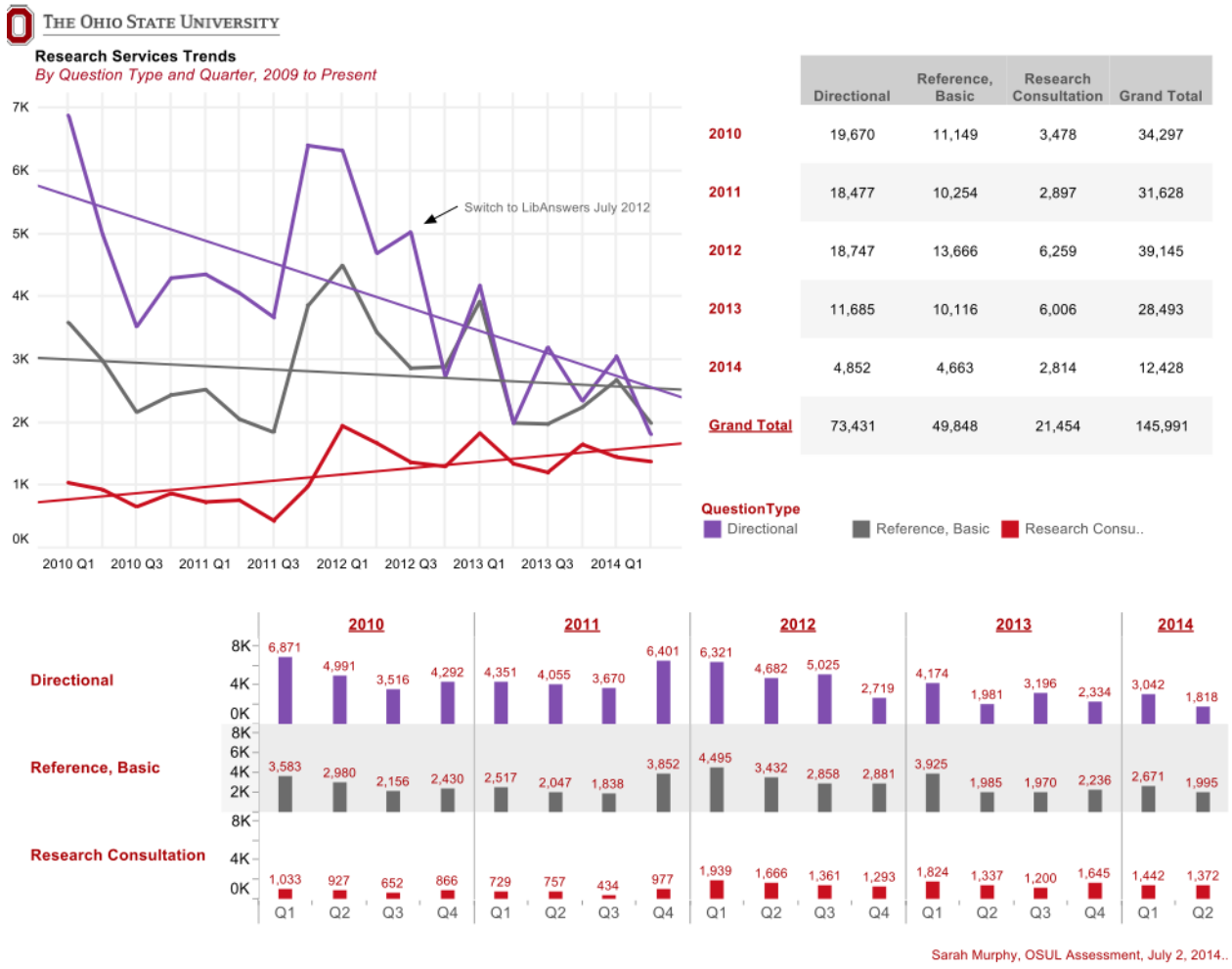
## Tableau at The Ohio State University Libraries—by Sarah Anne Murphy

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]




| | 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 |

**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.

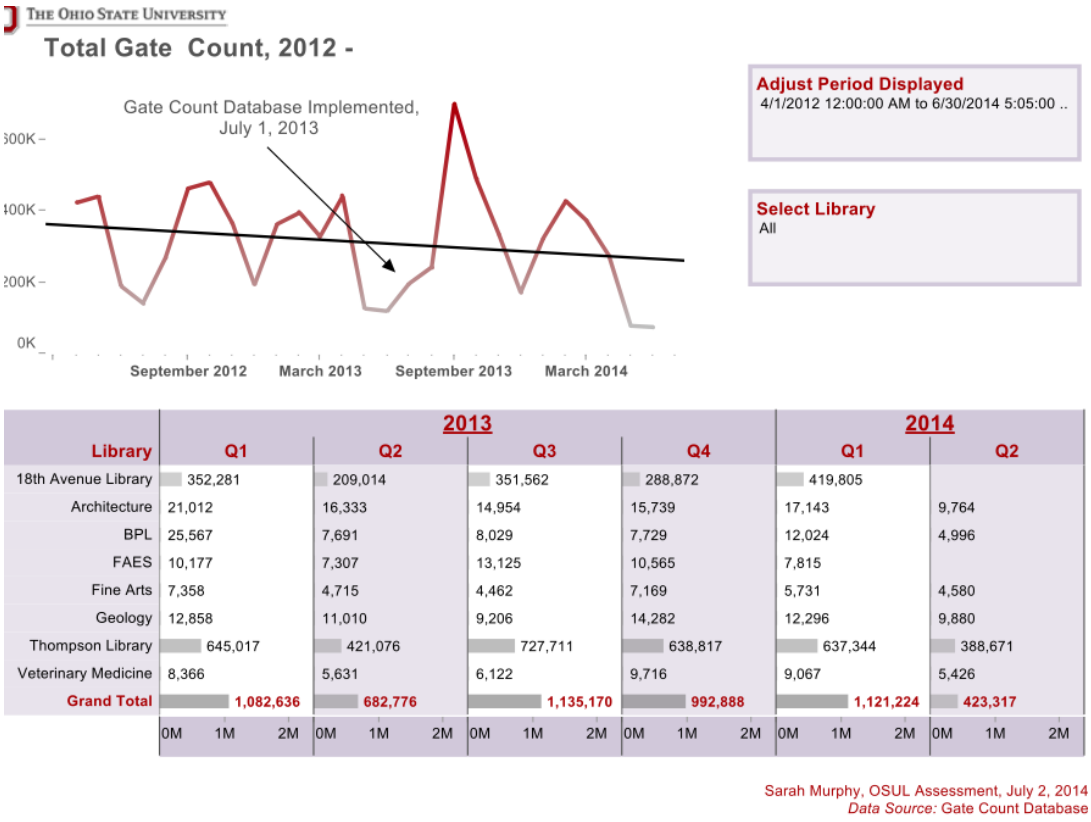



**FIGURE 2. GATE COUNT DASHBOARD**

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.

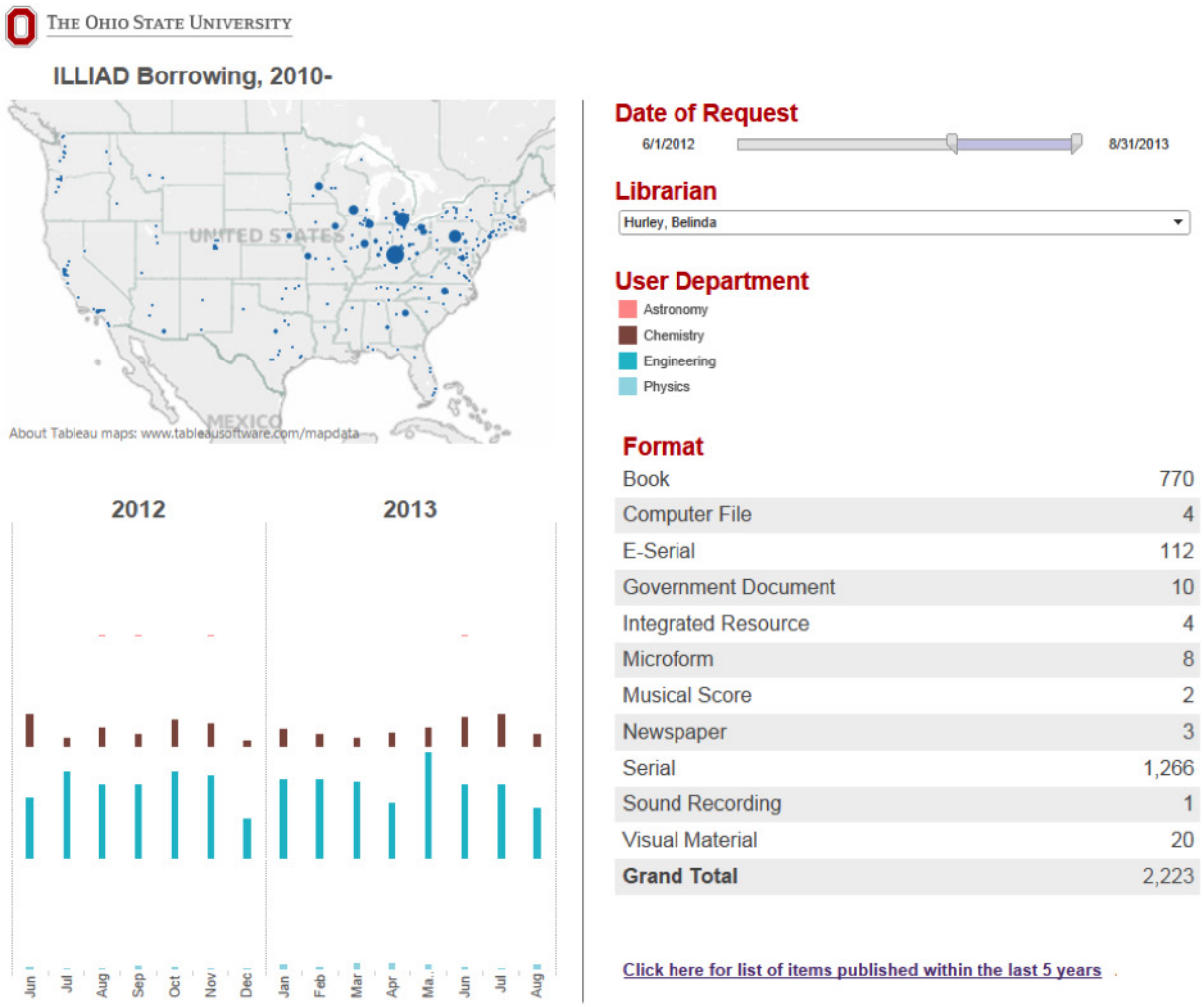



**FIGURE 3. DASHBOARD OF ILLIAD BORROWING, 2010–**

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.

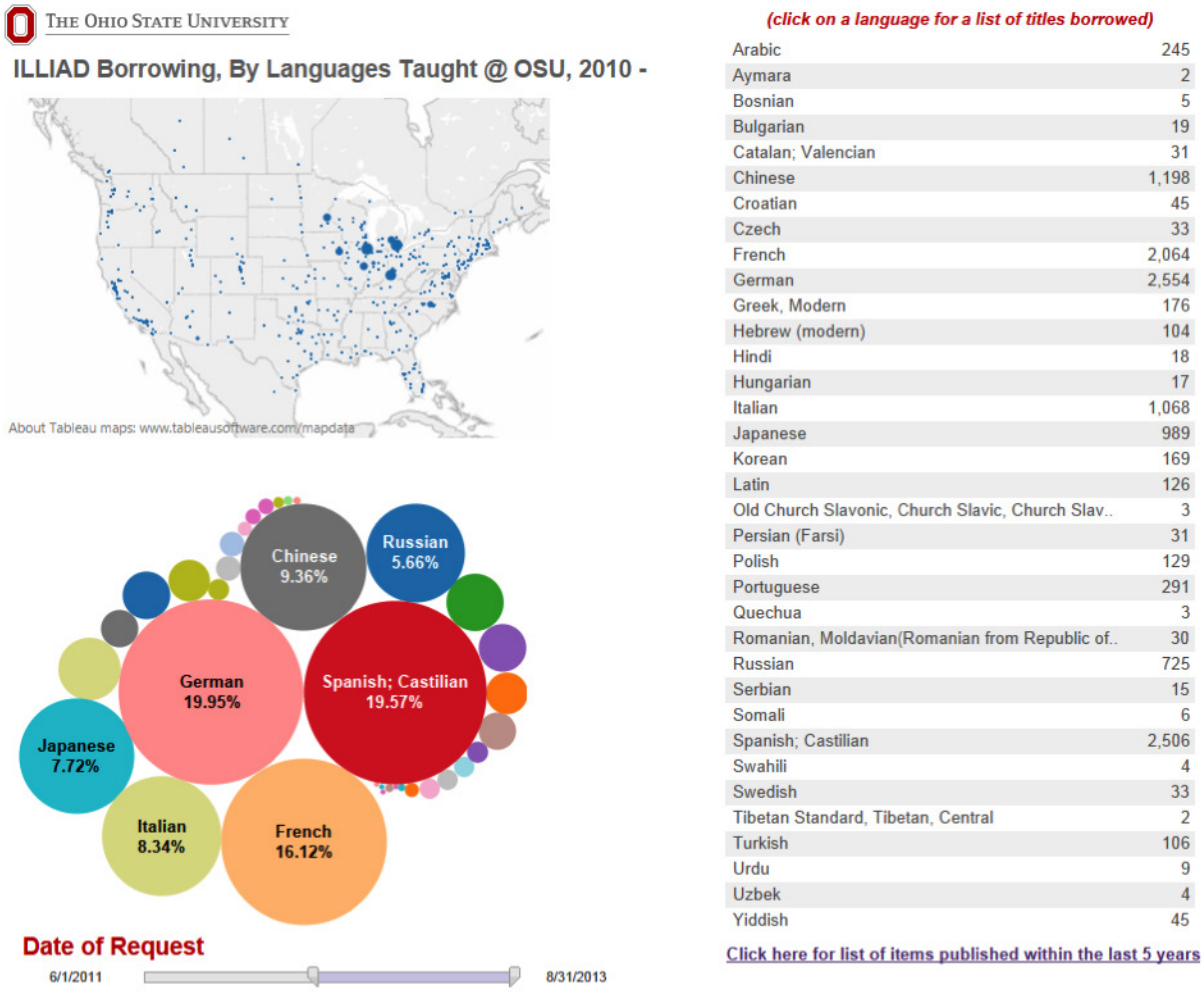



**FIGURE 4. DASHBOARD OF ILLIAD BORROWING, ALL LANGUAGES, 2010–**

## 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.

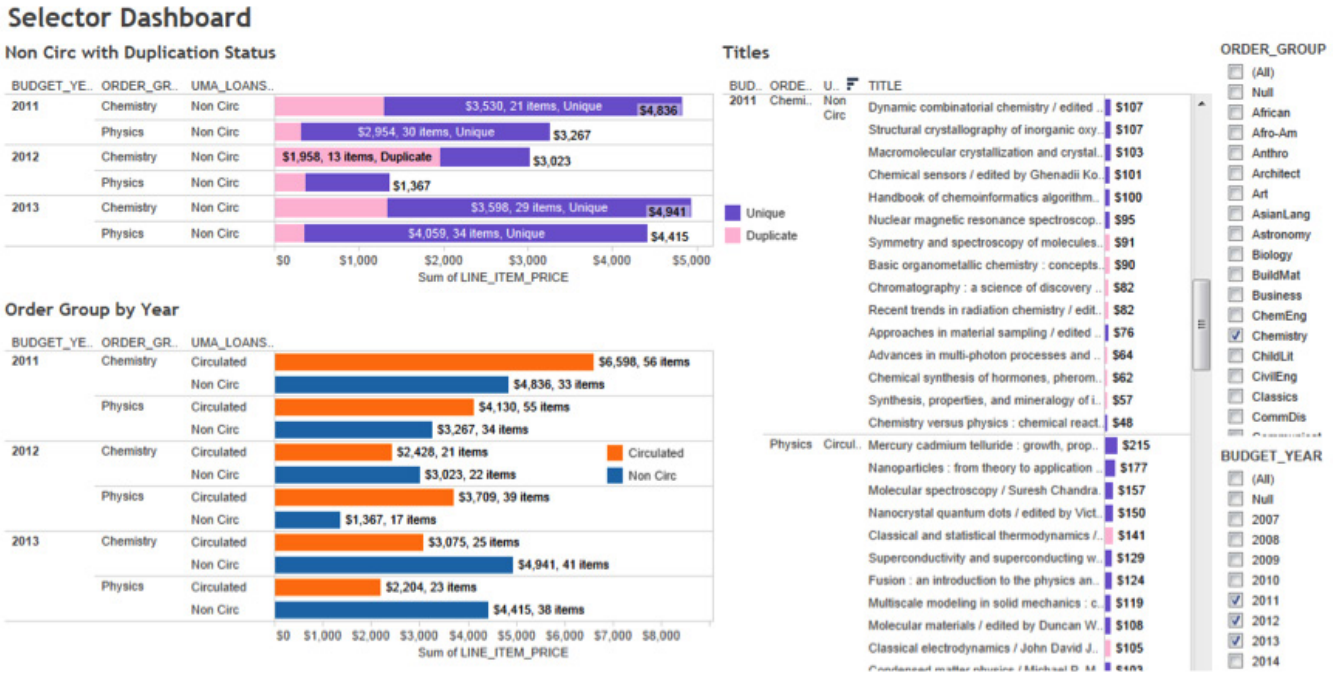



**FIGURE 5. DASHBOARD OF PURCHASING AND CIRCULATION (ALEPH)**

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.

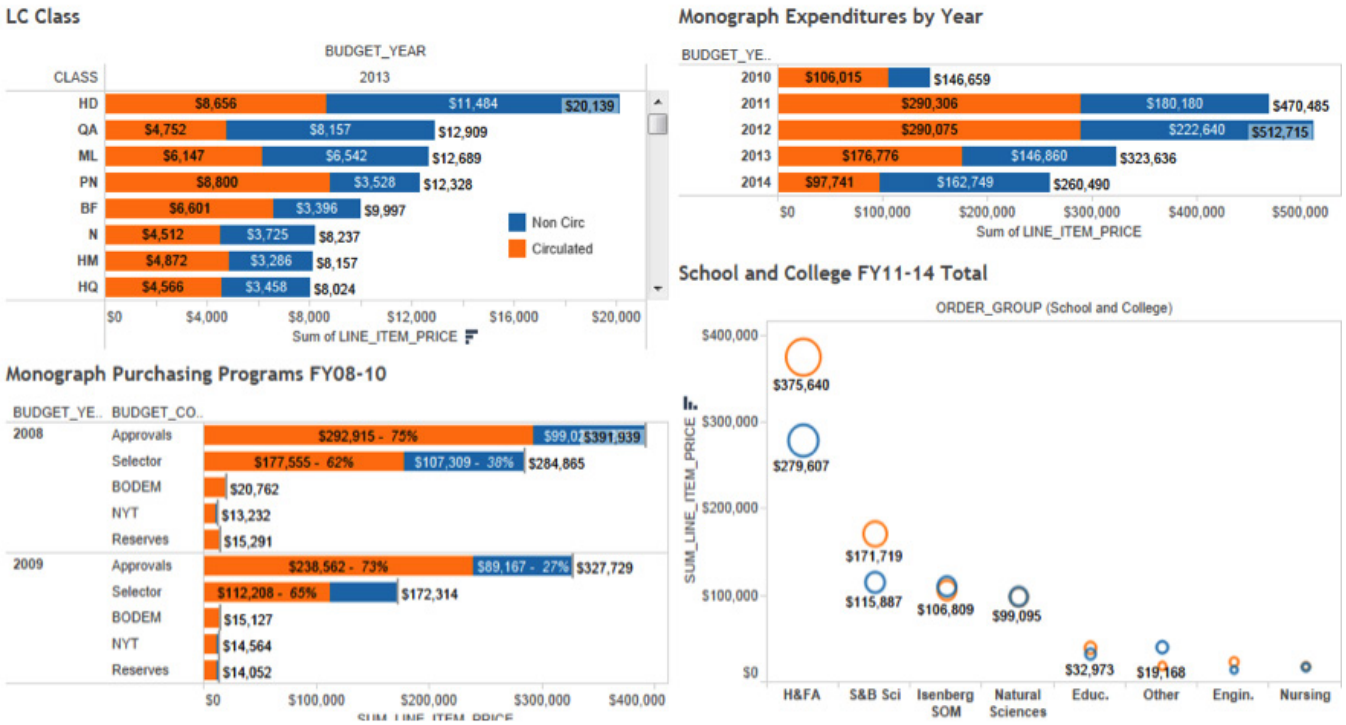



**FIGURE 6. LIBRARY OF CONGRESS CLASS, PURCHASE PROGRAM, FISCAL YEAR TOTAL, AND SCHOOL AND COLLEGE EXAMPLES**

## 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).

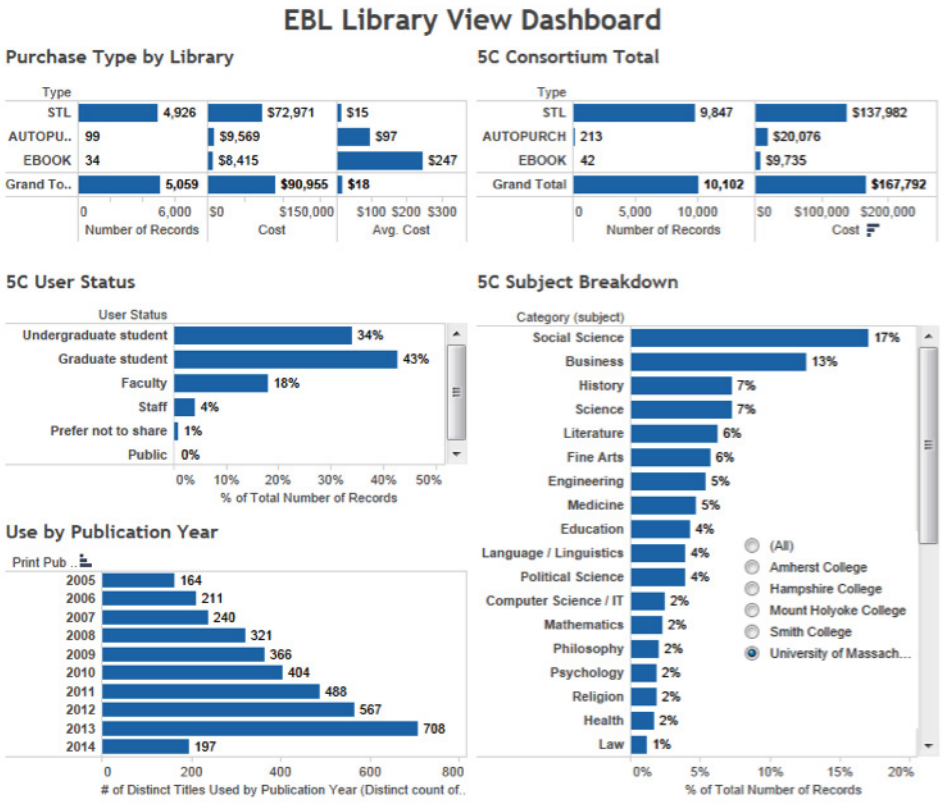
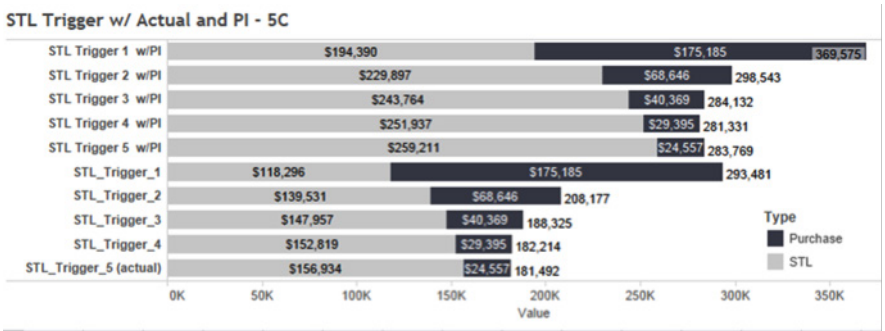



**FIGURE 7. EBL OVERVIEW BY LIBRARY**

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.




| | A | B | C | D | E | F | G | H | I | J | K | L | M | N |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Library | Type | STL_Trigg | STL Trigge | STL_Trigg | STL Trigge | STL_Trigg | STL Trigge | STL_Trigg | STL Trigge | STL_Trigg | STL Trigge | STL_Trigg | STL Trigger 0 |
| 2 | AC | Purchase | $4,772 | $4,772 | $5,421 | $5,421 | $6,468 | $6,468 | $10,495 | $10,495 | $22,826 | $22,826 | $105,691 | $105,691 |
| 3 | AC | STL | $19,288 | $34,410 | $18,375 | $32,573 | $17,565 | $31,094 | $16,258 | $28,804 | $13,404 | $23,574 | $0 | $0 |
| 4 | HA | Purchase | $989 | $989 | $1,244 | $1,244 | $2,632 | $2,632 | $4,025 | $4,025 | $11,434 | $11,434 | $72,180 | $72,180 |
| 5 | HA | STL | $10,637 | $17,466 | $10,504 | $17,231 | $10,194 | $16,626 | $9,701 | $15,799 | $8,410 | $13,603 | $0 | $0 |
| 6 | MH | Purchase | $2,610 | $2,610 | $3,450 | $3,450 | $4,559 | $4,559 | $8,602 | $8,602 | $20,967 | $20,967 | $108,478 | $108,478 |
| 7 | MH | STL | $17,803 | $29,206 | $17,291 | $28,077 | $16,831 | $27,217 | $15,917 | $25,635 | $13,610 | $21,568 | $0 | $0 |
| 8 | SM | Purchase | $4,606 | $4,606 | $4,987 | $4,987 | $6,890 | $6,890 | $9,932 | $9,932 | $27,317 | $27,317 | $165,274 | $165,274 |
| 9 | SM | STL | $26,874 | $43,931 | $26,201 | $42,922 | $25,329 | $41,570 | $24,071 | $39,633 | $20,584 | $34,063 | $0 | $0 |
| 10 | UM | Purchase | $11,580 | $11,580 | $14,294 | $14,294 | $19,819 | $19,819 | $35,591 | $35,591 | $92,641 | $92,641 | $510,075 | $510,075 |
| 11 | UM | STL | $82,332 | $134,198 | $80,448 | $131,133 | $78,037 | $127,257 | $73,584 | $120,025 | $62,288 | $101,582 | $0 | $0 |
| 12 | FC | Purchase | $24,557 | $24,557 | $29,395 | $29,395 | $40,369 | $40,369 | $68,646 | $68,646 | $175,185 | $175,185 | $961,699 | $961,699 |
| 13 | FC | STL | $156,934 | $259,211 | $152,819 | $251,937 | $147,957 | $243,764 | $139,531 | $229,897 | $118,296 | $194,390 | $0 | $0 |
| 14 | | | | | | | | | | | | | | |

**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.

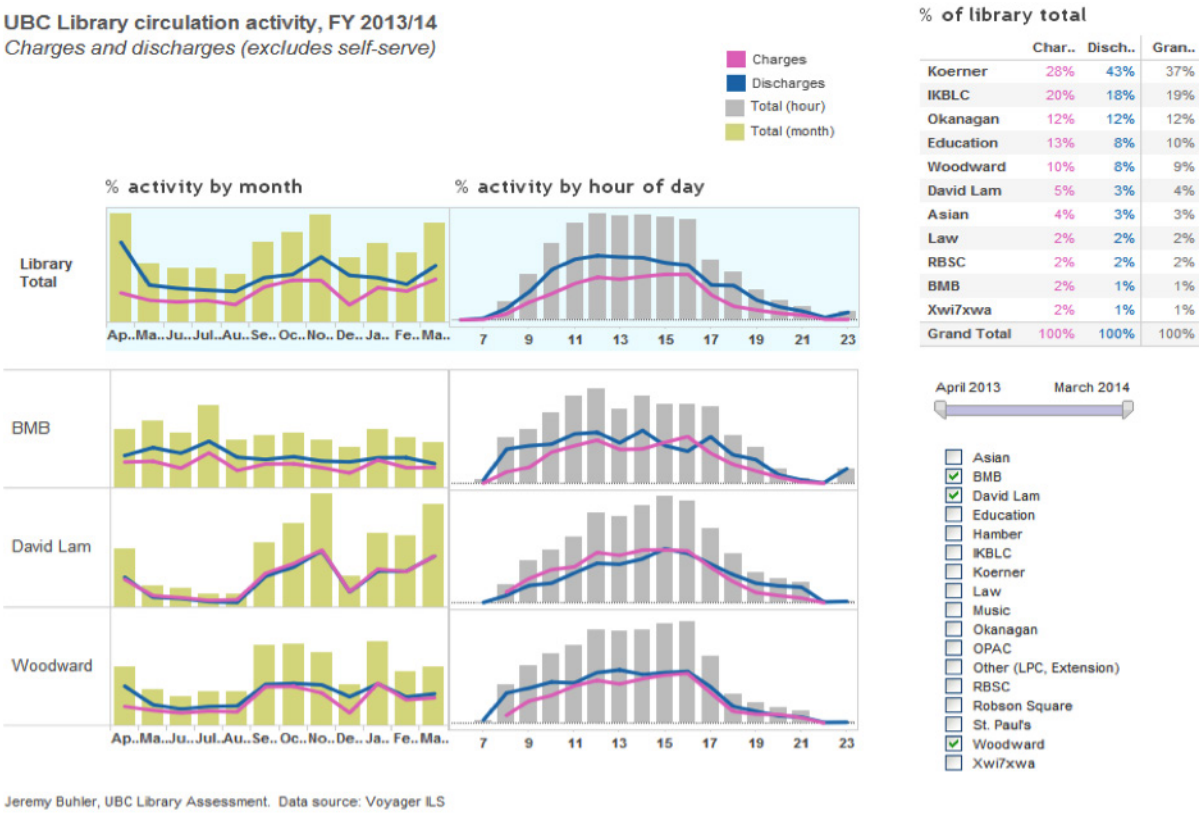



**FIGURE 9. UBC LIBRARY CIRCULATION ACTIVITY, FY 2013/14**

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.

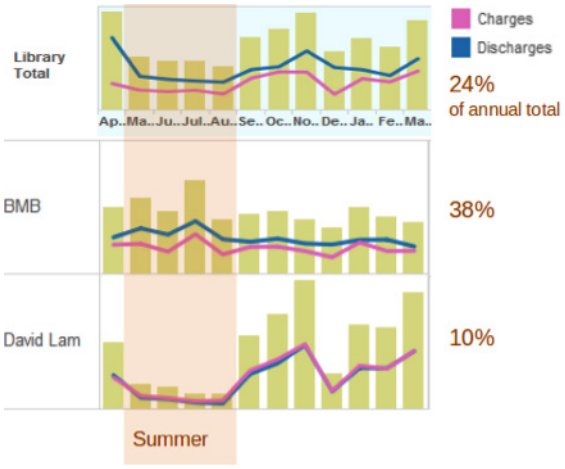



**FIGURE 10. UBC LIBRARY CIRCULATION: % OF ACTIVITY BY MONTH AND LIBRARY BRANCH**

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).

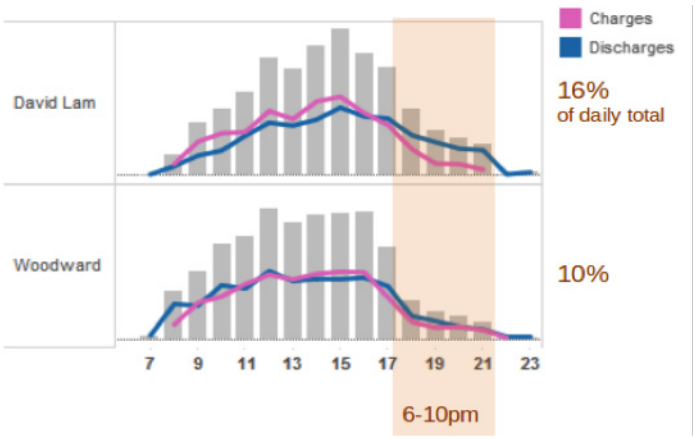



**FIGURE 11. UBC LIBRARY CIRCULATION: % OF DAILY ACTIVITY BY HOUR AND LIBRARY BRANCH**

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.

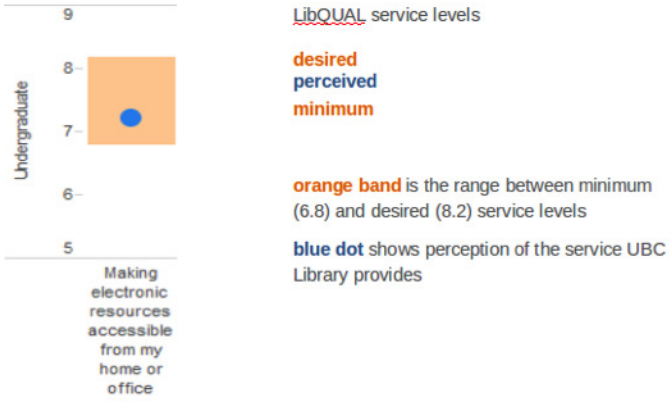



**FIGURE 12. VISUAL REPRESENTATION OF LIBQUAL+ RESULTS FOR A SAMPLE QUESTION**

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.

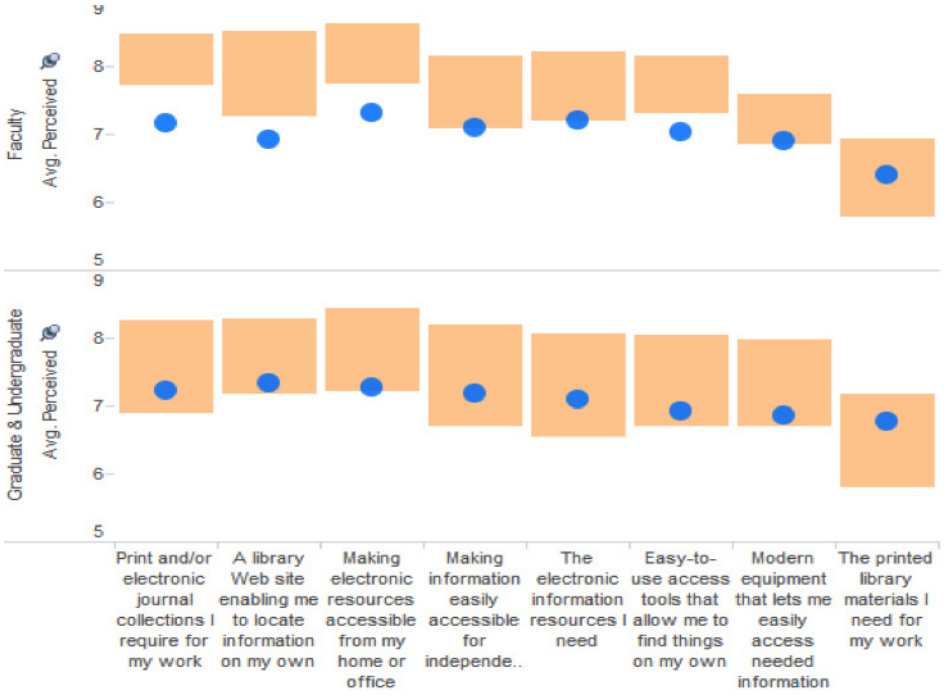



**FIGURE 13. 2013 UBC LIBQUAL+ RESULTS: INFORMATION CONTROL QUESTIONS**

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.

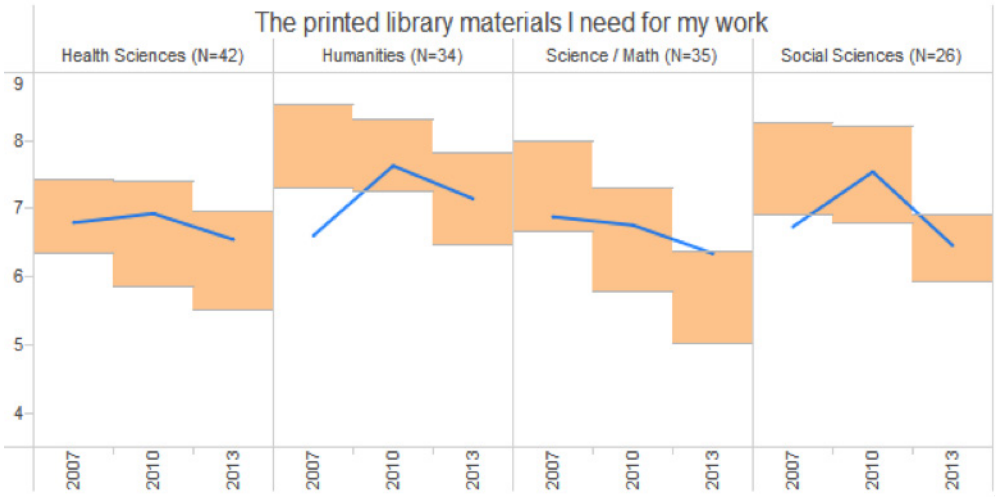



**FIGURE 14. LONGITUDINAL UBC LIBQUAL+ RESULTS (N REFERS TO 2013 SURVEY)**

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.



**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.