1-2009

Benchmarking Usage Statistics in Collection Management Decisions for Serials

J. Cory Tucker

University of Nevada, Las Vegas, cory.tucker@unlv.edu

Follow this and additional works at: https://digitalscholarship.unlv.edu/lib_articles

Part of the [Collection Development and Management Commons](#)

Repository Citation

Tucker, J. C. (2009). Benchmarking Usage Statistics in Collection Management Decisions for Serials. *Journal of Electronic Resources Librarianship*, 21(1), 48-61.
<http://dx.doi.org/10.1080/19411260902858581>

This Article is protected by copyright and/or related rights. It has been brought to you by Digital Scholarship@UNLV with permission from the rights-holder(s). You are free to use this Article in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself.

This Article has been accepted for inclusion in Library Faculty Publications by an authorized administrator of Digital Scholarship@UNLV. For more information, please contact digitalscholarship@unlv.edu.

Benchmarking Usage Statistics in Collection Management Decisions for Serials

Introduction

Starting in the spring of 2007, the University of Nevada, Las Vegas (UNLV) Libraries began a collection assessment project to look at the usage of the libraries' collections and analyze user behavior of library patrons. In order to complete the project, a collection assessment committee was organized and a subgroup of the committee was selected to gather and analyze data including usage statistics, collection specific information and other data as needed. Once the data had been gathered and analyzed, the Collection Management department began using the statistics in decision making and changed the collection development policy for electronic resources to incorporate benchmarking of usage statistics.

Literature Review

The literature for usage statistics offers a wide variety of articles. An article that provides a good overview of usage statistics for electronic collections and assessment was written by Bhatt¹. The article provides information on the challenges of assessment in regards to electronic collections and provides detailed information about COUNTER and the issues associated with usage statistics. A second general article related to usage statistics by Peters discusses the challenges of collecting use statistics and what practical information can be found in these stats². The article by Peters was particularly useful for this study because it helped identify two contexts on analyzing usage statistics for electronic resources including how usage of a resource evolves over time and comparing usage statistics of similar types of resources. Both of these contexts were used in this study. Bordeaux and Kraemer made a presentation at the 2004 North American Serials

Interest Group annual conference concerning usage data³. The presentation discussed information on options for gathering usage data and how to compile the data as well. One concept from this article used in the study was the suggestion by Kraemer to use the calculation of percentage change of use from the previous year. Another good summary article on usage statistics was written by Cooper⁴. The article provides an overview of the library literature for tracking usage statistics for print and electronic journals. The article examines both quantitative and qualitative studies in academic libraries. Another useful article by Shepherd provides detailed information and updates on COUNTER⁵. For this study, the article by Shepherd provided two specific benefits. First, the articles detailed descriptions of the seven usage reports assisted the library choose the appropriate statistical information to use in the study. The article also helped the study by pointed out the types of usage statistics that are required by vendors to be COUNTER compliant and this was used to check the validity of the statistical information gathered for the study.

Many presentations and articles have been written about usage statistics for assessing journals and databases. In a presentation at the 2004 International Federation of Library Associations, Franklin reported on efforts at the University of Connecticut Libraries to develop cost per use data for electronic collections and how the data was used to make management decisions⁶. The project included calculating subscription cost per use, operational cost per use and total cost per use for print and electronic journals and databases. The statistical calculations from this article where not incorporated into this study, but the article brought about investigations on including operational costs in decisions on serials retention. In an article by Conyers, an e-measures project in England

was discussed⁷. The project was conducted to obtain a set of statistics for measuring electronic information services in higher education libraries. The project involved obtaining twenty-one types statistics from 25 university and college libraries including holdings, usage and costs. The statistics were used to create benchmarks across the group of libraries. Although this study did not use any of the benchmarks mentioned in the article, it was extremely useful for this study in that it was a practical example of using benchmarks in the decision-making process. It provided a framework on how to organize a study using statistical measurements and benchmarking.

In addition to articles on usage statistics, there have been some articles written about the concept of benchmarking. An article by Poll provides an overview of national projects involving benchmarking and compares the different types of methods used in the projects⁸. Although the article did not provide specific benchmarks used in this study, the article is extremely useful in providing background information on the concept of benchmarking and how it can be used to measure quality or performance.

UNLV Libraries

Currently, the UNLV Libraries has over 1.2 million monographs, 300 databases and 27,000 journal titles. The collection is overseen by the collection management department, which is in the division of Research and Education. The division of Research and Education is divided into three departments, including collection management, reference and instruction. The division is made up of subject librarians or liaisons and each liaison serves on the reference desk, teaches library instruction sessions and are involved in collection development. The division has a liaisons group made up of subject librarians, who meet monthly to discuss issues and policies in each of the three

areas. For selection of serials, there is a Serials Review Group. The Serials Review Group is made up of the head of collection management, serials librarian and three librarians elected from the Liaisons group. The group currently meets twice a year to review serials requests from subject librarians. Currently, the Serials Review Group is not charged with decisions on retention, only selection of new serials.

Usage statistics

In the Fall of 2000, UNLV libraries began an initiative to harvest usage statistics for all electronic resources, including journals and online databases. The first part of the project involved creating a vendor database. The vendor database was created using an *Excel* spreadsheet and includes basic demographic information such as vendor, title, distribution, URL, but also important access details such as login username/password and dates of record update.

Table I shows the last eight years of the number of vendors from which we were able to receive usage statistics and the number of vendors that do not provide usage statistics.

Table I

Year (Fiscal Year July – June)	ER Reporting Stats	ER not reporting stats	Total ER's
2000	62	13	75
2001	102	19	121
2002	102	21	123
2003	102	21	123
2004	134	2	136
2005	200	7	207
2006	247	8	255
2007	286	12	298

The table shows that in 2002, the library was able to obtain usage statistics from 83% of the vendors and in 2007, that figure has increased to 96%.

The usage data is harvested, either by email or logging in to a vendor's site, by a student worker in the collection management department each quarter, after the 15th of the month. This process is efficient, more cost effective and allows the library to harvest statistics from more vendors compared with using a collection development tool, such as *Scholarly Stats*. The data is then entered into an *Excel* spreadsheet. Search statistics are compiled for each fiscal year and divided by the annual cost to give a cost per search figure. Beginning in the fiscal year 2006, a new category of cost per full-text view was added. Totals for the fiscal year full-text views are divided by the annual cost to give the cost per full-text view figure. An example of how the data is organized is available in Appendix I.

The harvested of data includes:

- Vendor
- Title of Electronic Resource
- Monthly searches and full-text views
- Total searches and full-text views
- Annual cost of product
- Cost per search and full-text view (calculated from above data)
- Types of use

Although the library had been collecting use data for eight years, the data had not frequently been used in decision making and had not been reflected in collection development policies because the UNLV Libraries budget increased each year and the library did not have to cut resources. With the collection assessment, all of this changed.

Collection Assessment

The collection assessment project at UNLV Libraries was conducted for various

reasons, but the main reasons were that the last assessment of the collection was ten years ago and the library would soon be facing budget cuts. The library needed to evaluate the collection by using data and statistics to measure the performance of specific resources in the collection and use these criteria to identify underperforming resources. The criteria could be used to cancel resources, identify resources that could be replaced with new resources or to identify resources that need further marketing and promotion. The results of the assessment would be incorporated into decision making to improve the collections, which best assist the students and faculty in the accomplishment of their research and academic pursuits.

During the collection assessment, the subcommittee decided that usage data from the last five years should be analyzed for all electronic resources. The subcommittee determined that the use data would be a key metric in determining the value of electronic resources. We determined that the usage statistics should be used:

- to justify purchases
- to cancel subscriptions
- to aggressively market less used resources
- to upgrade existing resources
- to aid in strategic planning
- to obtain “end of year” money from University Administration
- to compliment existing subscriptions with addition of journal backfile purchases
- to help make decisions in Add One/Drop One, depending on the budget

The key issue with the usage statistics is their reliability, accuracy and standardization. Since the library had been collecting statistics for eight years, we have noticed that the reliability and accuracy has improved, especially with the COUNTER initiative.

Taking all of this into consideration, the subcommittee decided to compile statistics for the assessment by category. The categories were electronic resources and

one-time purchases (including journal back-files). For electronic resources, the committee decided to use a statistical calculation suggested in an article by Kraemer and added in a total percentage of inflation increase/decrease (calculated from year subscribed to current year). Due to the fact that the journal backfiles were purchased with one-time money and owned by the library, inflation is not an issue. For backfiles, the spreadsheets included year by year usage. The resources were then divided up by college and sent out to the subject liaisons for review. The liaisons could then use the data to help them evaluate the resources and to identify products that may need additional publicity or marketing to users.

Data Analysis

After compiling all of the usage statistics for the collection assessment, the committee was left with some important questions. What is good use? What parameters does the library use to determine which resources will be reviewed for retention/cancellation? These issues were referred back to the collection management department for consideration.

The collection management department analyzed the data and decided that usage statistics should be used more frequently when making decisions about the collection and should be incorporated in collection policies. The department felt the best way to incorporate usage statistics would be to attempt to use benchmarking as a tool for all electronic resources, except journal backfiles because they are not a serial (except for maintenance fee). Creating benchmarks, allows the library to set a standard level of performance for our resources. The benchmarks would be used to identify those resources not meeting our standard and would put these up for retention review. The

benchmarks would consist of:

- cost per use or cost per full-text view
- inflation rate

After deciding to use benchmarking for retention, the main issue was to organize the electronic resources into “like” categories where the department could organize and compare data among similar resources. This follows the suggestion in the article by Peters that one context for using electronic resource usage statistics is the ability to compare and contrast statistics for similar resources. The department decided to organize the electronic resources into six categories:

- electronic journal subscriptions
- Index & Abstract databases
- Full-text Aggregator databases
- Other Full-text databases
- electronic journal packages
- ebook subscriptions

The department decided to change the charge of the serials review group and add in the responsibility of reviewing electronic resources for cancellation or retention in accordance with the set usage benchmarks. The process for the serials review group would be to identify electronic resources not meeting the benchmark performance standards and choosing those resources for cancellations. The list of resources up for cancellation would be sent to subject librarians who would then analyze each resource for content, usability, overlap and importance to the discipline. Subject librarians would have the opportunity to communicate with faculty and obtain faculty input. Subject librarians will have the opportunity to request an exception to the cancellation in writing and the serials review group would review the exception requests and have the authority of final decision for retention. The benchmarks were also added to the electronic

resource collections policy.

Analysis of Subscription Electronic Databases

As stated previously, the committee organized electronic databases by the type of resource including abstract and index databases, e-books, full-text aggregators and other full-text databases. This method allowed the committee to analyze databases by “like” resources. Once the databases were organized into the proper category, the next step in the data analysis involved creating spreadsheets for each category. The spreadsheets would consist of at least 3 years of most recent data with required fields for each resource:

- Title
- SEARCHES(total for fiscal year)
- % Increase or Decrease searches for 3 years
- FT VIEWS(total for fiscal year)(May also be full-text article requests depending on the vendor)
- Annual COST
- Inflation factor
- COST/ SEARCH
- COST/ FT VIEW
- Type of Use & Notes
- FY started
- Format (Abstracts/Index, Aggregator, Electronic Journal package, Full text resource, etc.)
- Notes field

The data fields above were included in the spreadsheet because it was determined that the committee needed as much information as possible to make effective decisions.

Providing addition statistical information related to both usage and pricing allows for more in-depth analysis. The data allows reviewers to compare resources across several statistical measures and allows for usage trends to be identified. This is supported in the

literature and was referred to in the article by Peters. Peters suggests that one context for analyzing usage statistics is called a resource context and involves analyzing usage over time. For this study, the percentage increase/decrease searches illustrated usage trends for each resource over a three year period. An example of this table is available in Appendix II.

Sorting the Databases

To complete the project, each database was sorted by format. If the format has less than 6 months of usage statistics, the statistics were eliminated for that fiscal year. This usually applies to new electronic resources. If a resource is available to the library free of charge via the vendor or from the Nevada State Contract, it was listed in the individual spread sheets and was not a part of calculating the average cost per search. If there were no stats from vendor, this resource was grouped with the free ones as we have no data to make decisions. If the resource does not have 3 years of data, it is not available for evaluation.

Evaluation Criteria for the Databases

As stated previously, the benchmarks for databases are:

- cost per use or cost per full-text view
- inflation rate

For each category of databases, the benchmarks were calculated. Each sheet would then have a calculation for an average cost per search and/or cost per full-text view. For the inflation rate, the collection management used a 10% annual inflation increase based on historical data (30% threshold for three years). All averages were rounded off to the nearest whole number and costs were rounded off to the nearest cent. Any resource that exceeded the benchmarks was highlighted and up for retention review. The list of

databases will then be sent to the liaison librarians to receive faculty input. Liaison librarians would have a specified time to provide feedback to the serials review group.

Abstracts/Indexes

Since most abstracts and indexes carry little or no full text items, for the spreadsheet, we calculated the average cost per search. For our purposes we rounded round off to the nearest dollar amount. The cost per search figures were analyzed and those that are over the average cost were highlighted. Sort from smallest to largest in the last fiscal year cost per search. Next, Look at % of inflation cost for the cost of the resource. Highlight if over 10%.

Table II: Abstracts/Index Database Data

Year	Average # searches	Average Cost Per Search
05-06	9891	\$2.69
06-07	10391	\$2.92
07-08	10003	4.62

That data in Table II shows that the average number of searches has increased over the past three years, however, the average cost-per-search has increased as well.

Aggregators

Calculate the average cost per search and average cost per full-text view. For our

purposes, the numbers were rounded off to the nearest dollar amount. Next, we analyzed the cost per search and cost per full-text view on these products and highlighted those that are over the average cost for both searches and full-text views. Look at % of inflation cost for the cost of the resource. Highlight if over 10%.

Table III: Aggregator Database Data

year	Average # searches	Average Cost Per Search	Average # FT views	Average Cost per FT view
05-06	29806	\$1.15	29343	\$1.26
06-07	28155	\$0.94	26648	\$1.26
07-08	28714	\$0.71	22617	\$1.26

The data in Table III shows that for aggregator databases, both the number of searches and full-text views have decreased over the last three years. The average cost-per-search has decreased, while the cost-per-full-text view has not changed.

Other Full-Text Databases

Calculate the average cost per search and average cost per full-text view. For our purposes, the numbers were rounded off to the nearest dollar amount. Check the cost per search and cost per full-text view on these products and highlight those that have decreases and are over the average cost for both searches and full-text views. Look at % of inflation cost for the cost of the resource. Highlight if over 10%.

Table IV: Other Full-Text Database Data

year	Average # searches	Average Cost Per Search	Average # FT views	Average Cost per FT view
05-06	4426	\$16.67	2582	\$5.67
06-07	2805	\$11.12	2559	\$9.94
07-08	2835	\$9.52	1915	\$5.69

The analysis for other full-text resources illustrates a decrease in the average number of searches and this has resulted in a lower average cost-per-use. The average number of full-text views has decreased and the average cost-per-use experienced a slight increase.

E-book Collections

For e-book subscriptions, we calculated both average cost per search and average full-text views. For our purposes round off to the nearest dollar amount. Both figures were used because out of the 18 e-book collections the library subscribes to, five do not provide full-text view statistics. Check the cost per search and cost-per full-text view and highlight those that are over the average cost. Look at % of inflation cost for the cost of the resource and highlight if over 10%.

Table V: E-book Subscription Data

year	Average # searches	Average Cost Per Search	Average # FT views	Average Cost per FT view

05-06	2270	\$6.08	8339	\$6.17
06-07	2988	\$7.34	20925	\$9.35
07-08	2645	\$9.99	22605	\$8.45

Ebook collections have experienced both increases and decreases in the average number of searches, but the average-cost-per use has steadily increased over three years. The average number of full-text views has increased significantly. The average cost-per-use did increase in year two, but decreased in year three.

Analysis of Electronic Journal Packages and Electronic Journal Subscriptions

For electronic journal subscriptions and electronic journal packages, the subcommittee conducted some research and decided that a good benchmark for journals would be to obtain a cost figure for the library to obtain a journal article not owned by the library. In the 2004-5 Association of Research Libraries study, the study found that for ARL libraries, the average cost to obtain a journal article not owned by the library was \$14.76.⁹ The collection management department met with the Library's document delivery department and inquired if the department would be able to calculate a cost per article average. The document delivery department did not have the information available, but would make it a future project. So, the subcommittee decided to use the ARL figure until one could be calculated for the UNLV Libraries. For our purposes, we rounded up to \$15 as a benchmark for cost per article. For individual electronic journal subscriptions, any journal that was above the benchmark was up for retention review. For the electronic journal packages, if the average of the entire package was above the

benchmark, it was up for retention review as well to see if it would be more cost efficient to purchase individual subscriptions rather than the entire package.

Collection Development Policy Changes

The collection assessment caused major changes in the collection development policy for the library. The policy was changed to include benchmarking as a measure for review. In addition to benchmarking, the Serials Review Group was given the charge of retention review of all serials. Each year, the Serials Review Group would use benchmarking to identify those serials up for review and the list would be passed on to the subject librarians. The subject librarians would be given time to review each list and then would have the opportunity to ask for exceptions. The subject librarian is required to provide justification for an exception and then the Serials Review Group has the right to accept or decline an exception. Another change to the collection development policy is the time frame for review of a serial. The subcommittee recommended that a serial would be given three years before being reviewed. This would allow time for the resource to be marketed to library users and accessed by users.

Conclusion

Usage statistics are an important metric for making decisions on serials. Although the UNLV Libraries have been collecting usage statistics, the statistics had not frequently been used to make decisions and had not been included in collection development policy. After undergoing a collection assessment, the collection management department decided on manipulating usage data to create a benchmark performance for serials and use the benchmark for starting the decision making process. With libraries facing budget cuts and having to make decisions on retention of serials, the

idea of benchmarking has merit. The UNLV Libraries has decided to review benchmarks each year and see what type of affect it has on library users and the library, especially the document delivery department. Looking at our users, the library will conduct another survey for faculty and students to see if this has had an impact on obtaining library resources and the library will also use results from *LibQual*, which the library participates in every two years. For document delivery, the library will look to see if there are significant increases in journal article requests and analyze the financial impact on the library.

Looking at benchmarking as a concept, it would be interesting to have discussions on best practices for usage statistics and benchmarking. In addition, the idea of benchmarking across “similar” institutions should be reviewed and might be worth experimentation. Comparing similar institutions and what their usage statistics show for the same databases may illustrate a trend, especially by discipline.

References

1. Anjana Bhatt. “Assessing e-collections when every e-resource has its reader, every reader has his/her e-resource, and e-resources are ever growing,” *Against The Grain* 18, no. 5 (2006): 24-28.
2. Thomas A. Peters. “What’s the use? The value of e-resource usage statistics,” *New Library World* 103, no. 1172-3 (2002): 39-47.
3. Abigail Bordeaux and Alfred B. Kraemer. “Making the most of your usage statistics,” *The Serials Librarian* 48, no. 3/4 (2005): 295-299.
4. Mindy M. Cooper. “The importance of gathering print and electronic journal use data: getting a clear picture,” *Serials Review* 33, no. 3 (2007): 172-174.
5. Peter T. Shephard. “COUNTER: usage statistics for performance measure,” *Performance Measurement and Metrics* 7, no. 3 (2006): 142-152.
6. Brinley Franklin. “Managing the electronic collection with cost per use data”

(paper presented at the World Library and Information Congress: 70th IFLA General Conference and Council, Buenos Aires, Argentina, August 22-27, 2004) 35-49.

7. Angela Conyers. "E-measures: developing statistical measures for electronic information services," *VINE: The Journal of Information and Knowledge Management Systems* 34, no. 4 (2004): 148-153.

8. Roswitha Poll. "Benchmarking with quality indicators: national projects," *Performance Measurement and Metrics* 8, no. 1 (2007): 41-53.

9. Mary Jackson. "Assessing *ILL/DD Services: New Cost-effective Alternatives*," Association of Research Libraries, Washington, DC, 2004.