Introduction to Using Excel® Pivot Tables and Pivot Charts to Increase Efficiency in Library Data Analysis and Illustration

Andrea Miller
Missouri State University

Follow this and additional works at: http://bearworks.missouristate.edu/articles-lib
Part of the Collection Development and Management Commons

Recommended Citation

This article or document was made available through BearWorks, the institutional repository of Missouri State University. The work contained in it may be protected by copyright and require permission of the copyright holder for reuse or redistribution.
For more information, please contact BearWorks@library.missouristate.edu.
Introduction to Using Excel® Pivot Tables and Pivot Charts
to Increase Efficiency in Library Data Analysis and Illustration

Author Note
Andrea Miller, Duane G. Meyer Library, Missouri State University.

Correspondence concerning this article should be addressed to Andrea Miller,
Duane G. Meyer Library, Missouri State University, 901 S. National Ave., Springfield,
MO 65897. E-mail: AndreaMiller@MissouriState.edu
Abstract. Excel® offers powerful features that can spare its users countless hours of tedious and unnecessary effort. Pivot tables and pivot charts are robust tools that can streamline the work of analyzing library data, making it nearly instantaneous, visually engaging, and efficient. Although electronic resource usage statistics will be used for illustrative purposes, countless types of data in a tabular format are suited to the application of pivot tables and pivot charts. This article will discuss background and approaches for using these tools, followed by a companion article that will demonstrate essential techniques and applications.

Keywords: Excel® PivotTables, Excel® PivotCharts, pivot tables, pivot charts, library data analysis, library data illustration

Author Note
Andrea Miller, Duane G. Meyer Library, Missouri State University.
Correspondence concerning this article should be addressed to Andrea Miller, Duane G. Meyer Library, Missouri State University, 901 S. National Ave., Springfield, MO 65897. E-mail: AndreaMiller@MissouriState.edu

Received: November 13, 2013
Accepted: December 12, 2013
Introduction to Using Excel® Pivot Tables and Pivot Charts
to Increase Efficiency in Library Data Analysis and Illustration

The need for producing meaningful analysis of data is just as keen today as it was decades ago when R. W. Hamming declared, “The purpose of computing is insight, not numbers” (Hamming, 1962, p. [v]). Datasets can be challenging to evaluate for meaning, especially as the amount of data expands. Row upon row of raw data in spreadsheets may be virtually impossible to browse through with any hope of discerning its deeper meaning. Numerical summaries can be helpful but still may be overwhelming if the summaries themselves are unwieldy in size. Converting these numerical summaries into pivot tables and pivot charts can often render them more easily comprehensible through excellent layout and visual representation.

Excel® provides users with the ability to create pivot tables and their related pivot charts. These beneficial tools automate the process of data analysis and enable nearly instantaneous changes in the patterns into which data is organized as well as in the portions of the data that are being viewed. Reports to meet any need may be effortlessly generated at a moment’s notice to answer ad hoc questions that arise about the data. Pivot tables enable individual data points to be singled out for instant comparison with other points, allowing many different variables to be compared with ease. Background and approaches for using pivot tables and pivot charts will be presented herein, followed by a subsequent article which will demonstrate critical techniques and applications.

Pivot tables were named for their ability to instantly pivot data into any view that is desired. Pivot tables and pivot charts will be used as abbreviations for their actual names, which are PivotTable reports and PivotChart reports. Unfortunately, many people are
either unaware of pivot tables and pivot charts, which are among Excel’s® most powerful features, or have never attempted to learn how to use them (Altom, 2011; Chen, 2007; Montondon & Marsh, 2006). Applying these easily operated tools can eliminate the need to manually create formulas or to sort and re-sort data, thereby saving untold hours of painstaking effort.

Pivot tables are interactive and may be used with either large or small amounts of data. The pivot table is created and exists independently from its original source data. Rows and columns can be easily rearranged for different views of the data with no disturbance of the source data’s layout or content. Pivot tables make it simple to organize and summarize the data, as well as compare it and detect patterns. Pivot tables can also automatically produce totals without the need for writing formulas to perform such calculations or rearranging the data to calculate subtotals.

A regular table in Excel® is a range of data that has been deliberately formatted as a table, which gives it special functionality and features beyond that of data that has been simply entered in the shape of a table, as with a typical spreadsheet. The data in Figure 1 has been formatted as a table, and its appearance and functionality are quite different from data that has been typed into a spreadsheet with no special formatting applied to it. Regular tables are not as flexible as pivot tables, however. Pivot tables enable painless manipulation of large datasets. Pivot tables and pivot charts offer a rapid means of flipping the axes on which the data resides. Data that appears in rows straight down a spreadsheet can be flipped using a pivot table. The data can move into columns running across the spreadsheet to the right, which may enable the data to take a more usable form when converted into a visual chart. The automation of data manipulation significantly
INTRODUCTION TO PIVOT TABLES AND PIVOT CHARTS

speeds up the process and removes the potential for human error due to any manual manipulation of data. Different levels of data granularity can be readily generated. Pivot tables and pivot charts are dynamic and allow their content to be instantly changed as desired in order to answer specific questions about the data, while regular tables may require considerable effort to rearrange the data to produce answers.

[place figure 1 here]

Pivot tables and pivot charts may be used to simplify complex data as well as create visual representations of many types of library data. Although electronic resource usage statistics will be used as an example to demonstrate techniques, the applications for pivot tables and pivot charts are nearly limitless. The basic methods of using these tools can be applied to many types of library data, both within the realm of electronic resources as well as throughout the rest of the library. Assessments involving the concepts of “by” or “per” are particularly well-suited to pivot tables, such as “by fund code” or “per semester.” Countless questions about the data may be readily answered. What are the most or the least used resources in a particular category during a certain period of time? How does this compare to other categories or time periods? How do patterns in turnaways due to surpassing the maximum number of simultaneous users allowed for an electronic resource vary from month to month and across 2 or more years? What have the trends in the costs of certain resources been in comparison to other resources? What percentage of the budget does each resource represent? Possible applications for pivot tables and pivot charts abound and become apparent once knowledge of their functionality has been internalized.
Description of Electronic Resource Statistics

The electronic resource statistics used in the examples are drawn from an internal record-keeping system maintained by the Missouri State University Libraries’ Systems Department. The internal recordkeeping system records the number of times an electronic resource is accessed or “hit” by patrons clicking on one of the library’s permanent URLs (PURLs) for that electronic resource. The PURLs are used in links to the electronic resources provided in the library’s “Articles and Databases” web page, online subject guides, and the online catalog, among other places. The electronic resource statistics are collected from this internal system three times per year: Commencement Day for the spring, summer, and fall semesters. The library has assigned each electronic resource to one of the university’s six colleges, which are represented with abbreviations or has identified an electronic resource as a general one that appeals to students in more than one college. The college abbreviations are as follows: CHHS = College of Health and Human Services; CHPA = College of Humanities and Public Affairs; CNAS = College of Natural and Applied Sciences; COAL = College of Arts and Letters; COBA = College of Business Administration; COE = College of Education; and GEN = General use electronic resource. The result resembles assigning one of seven different fund codes to each of the 111 electronic resources. Any student is able to use any electronic resource, regardless of his or her college affiliation. The assignment of college abbreviations is simply a means of associating electronic resources with their most likely group of users as well as helping in assessment.

Discussion
Reports made from pivot tables and pivot charts can be used to visually represent electronic resource usage numbers in ways that make it easier for decision-makers to understand the data. Row after row of numbers can be difficult for even the mathematically inclined to interpret. When looking at a list of numerals, it may be difficult to judge just how large certain numbers are relative to other numbers. For example, Figure 1 contains data recording the number of hits on each of the electronic resources associated with the College of Health and Human Services. Persons viewing this data are able to see that 5,536 hits for PsycINFO is a very large number in comparison with the others, especially the 17 hits for the Encyclopedia of Disability. The question, however, is: Are they able to perceive just how large 5,536 is in comparison to 17 and all the other numbers in between? A mathematically inclined person may be able to estimate that the 1,546 hits for MEDLINE are roughly 28% of those for PsycINFO, but the average person may not possess such skills.

By using pivot table and pivot chart tools to convert numbers such as those in Figure 1 into chart illustrations, anyone can instantly see the proportions, relationships, and trends present in the numbers. These tools provide an easy way to quickly manipulate and summarize large quantities of numerical data. Contrast Figure 1 with the same data converted in seconds into a visual illustration in Figure 2 by using pivot tables and pivot charts. It is now apparent that the 1,546 hits for MEDLINE are between one-fourth to one-third of the relative numerical size of the 5,536 hits for PsycINFO. The relative usage rates of various electronic resources become instantly discernible.

[place figure 2 here]
The data may freely be expanded to include data for additional colleges and general-use electronic resources, as in Figure 3. The heaviest usage rates again become readily apparent when comparing one college to another. Similarly, the total dominance of Academic Search Premier, the most frequently used general database with approximately 150,000 hits at the base of the GEN column, is striking in comparison with the others.

[place figure 3 here]

The truly impressive power of pivot tables and pivot charts is exploited when even larger sets of source data and more variables are employed. It is not difficult to add on title, college, and hits data for additional years. In Figures 4, 5, and 6, semester data was added for a 4-year period. The larger results set can be manipulated by collapsing or expanding parts of the table in order to hone in on particular segments of the data. Field variables can also be rearranged and filters applied. Pivot charts can likewise be created and various filters applied to them as well.

The granularity of the data being studied is conveniently open to fine-tuning or expansion. By adding data reflecting different years and semesters, trends across time are revealed. Using MEDLINE as an example, layouts reflecting subtotals of usage hits by semester in Figure 4 can be changed in less than a second to one comparing usage across successive years in Figure 5. Furthermore, the data can be collapsed with a simple mouse click to show only the subtotals and/or grand totals at any desired level of granularity, as in Figure 6 showing subtotals by years. Note that subtotals for different variables appear on the top line of the section containing the variable. Only the grand total appears at the bottom, as in a traditional spreadsheet.
Review of the Literature

Books Specializing in Pivot Tables and Pivot Charts

This article is a brief introduction to the merits of pivot tables and pivot charts and can neither cover all of their amazing capabilities, nor all their limitations. A number of in-depth books are available that are devoted exclusively to pivot tables and pivot charts, and these are valuable sources for additional information. Although this article used Excel® 2010 in its examples, books about Excel® 2007 can also be useful because most of the functions remain the same. Jelen and Alexander (2007; 2011) have written outstanding books that thoroughly cover the subject and are suitable for both beginning and advanced users. These books contain clear and detailed explanations as well as an abundance of examples and illustrations. In one particularly visually oriented book, McFedries (2010a) has covered both beginning and advanced content. Most of the topics are covered in heavily illustrated, two-page spreads. While this approach is convenient, it does create a rather dense layout. Dalgleish (2007a) has written a book appropriate for novices as well as those with previous experience using earlier versions of the software; it covers the basic features of pivot tables and pivot charts in Excel® 2007. In another book, Dalgleish (2007b) assumes prior experience with pivot tables and pivot charts, and it targets more advanced users by presenting hundreds of hypothetical problem scenarios related to pivot tables and pivot charts in Excel® 2007, followed by their solutions.
Aitken (2007) provides instruction for both beginners and more advanced users in a single volume. In addition to thoroughly presenting pivot tables and pivot charts, there are appendices devoted to troubleshooting and to formatting and customizing charts.

**Books on Excel®**

Many books devoted to Excel® as a whole have chapters or sections on pivot tables and pivot charts. Such books offer far less information than books devoted exclusively to pivot tables and pivot charts, but they generally do cover basic techniques. Among the Excel® books examined, Jelen (2010a) featured the most in-depth content with two chapters on pivot tables, but only a couple of pages address pivot charts. More commonly, the books contain a chapter about pivot tables and approximately three to six pages devoted to pivot charts (Dodge & Stinson, 2011; Frye, 2010; Fuller, Fulton, Hayes & Riley, 2011; Harvey, 2010; MacDonald, 2010; Walkenbach, 2010a).

**Books on Specialized Aspects of Excel®**

Books on specialized aspects of Excel® may contain chapters with basic information on pivot tables and pivot charts (Alexander & Walkenbach, 2010). Some books on particular aspects of Excel® contain only chapters on pivot tables and do not discuss pivot charts (McFedries, 2010b; Walkenbach, 2010b). The reverse is also true, with some specialized Excel® books mainly discussing pivot charts with limited emphasis on their related pivot tables (Jelen, 2011a).

**Business and Computer Science**

Many experts in business and computer science appreciate the utility of pivot tables and their related pivot charts. This awareness is reflected in the literature by the consistent appearance of articles discussing how to construct and make use of pivot tables.
and pivot charts (Hunstad, 2010; Lehman, Lehman, & Feazell, 2011). A technology column devoted to Excel® regularly discusses topics related to pivot tables (Jelen, 2010b; Jelen, 2010c; Jelen, 2010d; Jelen, 2011b; Jelen, 2011c; Jelen, 2012). Aloisio and Winterfeldt (2010) described how a hospital’s department of radiology used pivot tables to analyze which hours of which days were their busiest periods. The employees became receptive to altering their staffing patterns to better match workflow needs once they saw the data and understood how changes would lead to less workload pressure and greater efficiency. Ting, Pan, and Chou (2010) applied pivot tables to a modified version of market-basket analysis (MBA) to analyze the likelihood of diners ordering certain combinations of menu items at a prix-fixe restaurant in order to enable servers to suggest these combinations. The researchers noted, “Another familiar application of MBA is in online bookstores, where customers will be presented with ‘associated products’ when they browse certain items” (Ting, Pan, Chou, 2010, p. 492). Wesley and Cox (2007) extolled the use of pivot tables in mortality analysis and recommend them as being accurate and flexible in considering a wide variety of variables. Efstathiou and Golby (2001) discussed the use of Excel® and its pivot table tools to compute the optimal physical layout sequence of various manufacturing machines so as to maximize the speed and efficiency of production. According to Eckert (2005), pivot tables enabled a construction company to spot weaknesses in the progression of their jobs early enough to make corrective adjustments. The construction company can now bid labor hours with great accuracy and identify problems related to profitability. Libraries have many business operations and could greatly benefit from adopting many of the same key tools used by businesses, such as pivot tables and pivot charts.
Education

Principals and professors in the field of education also understand the usefulness of pivot tables for handling student and school-system data. Library administrators, as well as those teaching courses in library science or information literacy, could benefit from adopting their methods. According to Mills (2006), the principals in Wilson County, NC use pivot tables to analyze the relationships between multiple types of data: for example, analyzing test results for demographic trends as well as identifying particular students at risk in specific areas. Another use is seeking clues for the means to cultivate better student behavior by analyzing disciplinary data. Green (2008), a college professor, has described methods for using pivot tables not only to easily track students’ grades but also to analyze these grades for information about individual students and the class as a whole. This information in turn helps him improve his instructional planning. Convery and Swaney (2012) have developed a university course instructional module that, among other topics, teaches expertise in pivot tables and their use in accounting. Such tools facilitate the analysis of data related to business challenges and possible solutions. Conmy, Hazlett, Jelen, and Soucy (2006) extol the use of pivot tables for literacy/curriculum coordinators or principals who must look at test results over time to analyze student and teacher performance trends.

Library Science

In contrast, the literature of library science contains material about using the well-known spreadsheet aspects of Excel® for library processes but is largely silent about pivot tables and pivot charts. References to them have appeared only a handful of times with limited information about their application and best use. Matthews (2009) of the Virtual
Library of Virginia (VIVA) touched upon the use of pivot tables for their advantages in summarizing usage-statistics data but did not delve deeply into the topic. Jones (2011) mentioned pivot tables once while describing a system of studying demographic data of library patrons as collected from ID cards swiped through turnstiles upon library entrance. As with other educators, Huber (2005) recommended using pivot tables for their ease in viewing various aspects of data related to student testing.

Although some sources promoting the use of Excel® in libraries do not mention pivot tables, they do, however, discuss how to normalize library data in order to make it suitably standardized for analysis. Such normalization is important in the preparation of source data. Greiner and Cooper (2007) wrote a chapter with helpful descriptions of methods that can be used to clean up inconsistencies in raw data that has been exported from an integrated library system into Excel®. Curtis, Scheschy, and Tarango (2000) wrote an appendix that suggested strategies for standardizing data, particularly for correcting variances in the transcription of titles, in order to prepare lists for automated comparison of the amount of overlap in database source lists.

**Summary**

Many librarians rely heavily on Microsoft® Excel® for stockpiling a multitude of types of data and could greatly benefit from adopting pivot tables and pivot charts for use in data analysis. These tools can be very helpful to use in interpreting the numbers contained in various types of library data. This introduction has presented possible approaches to using pivot tables and pivot charts using the area of electronic resource usage statistics as an example. The literature has been surveyed, with examination of both
technical sources treating Excel® itself and sources highlighting practical uses in the fields of business, computer science, education, and library science.

Sample data used in the figures is available for download for experimentation at http://people.missouristate.edu/andreamiller/PivotTables3.htm. The sample data and its usage will be further expounded upon in an upcoming companion article. The companion article will concern itself with application and will discuss the creation process for pivot tables and pivot charts using elemental data from electronic resource usage statistics. Fundamental concepts will be addressed, with demonstrations of techniques. An understanding of the many contexts in which these tools may be employed will be developed.
References


Figure Captions

**Figure 1.** Example of CHHS source data formatted as a regular table.

**Figure 2.** Example of CHHS data as a pivot chart.

**Figure 3.** All colleges as a pivot chart.

**Figure 4.** Usage by semester as a pivot table.

**Figure 5.** Usage by year as a pivot table.

**Figure 6.** Figure 5 collapsed to focus on usage by years.
**Figure 1.** Example of CHHS source data formatted as a regular table.

<table>
<thead>
<tr>
<th>Title</th>
<th>College</th>
<th>Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alt HealthWatch</td>
<td>CHHS</td>
<td>577</td>
</tr>
<tr>
<td>CINAHL with Full Text</td>
<td>CHHS</td>
<td>951</td>
</tr>
<tr>
<td>Clinical Pharmacology</td>
<td>CHHS</td>
<td>81</td>
</tr>
<tr>
<td>Encyclopedia of Disability</td>
<td>CHHS</td>
<td>17</td>
</tr>
<tr>
<td>Health Source: Consumer Edition</td>
<td>CHHS</td>
<td>375</td>
</tr>
<tr>
<td>Health Source: Nursing/Academic Edition</td>
<td>CHHS</td>
<td>539</td>
</tr>
<tr>
<td>Hospitality &amp; Tourism Complete</td>
<td>CHHS</td>
<td>131</td>
</tr>
<tr>
<td>MEDLINE</td>
<td>CHHS</td>
<td>1546</td>
</tr>
<tr>
<td>NASW Clinical Register</td>
<td>CHHS</td>
<td>27</td>
</tr>
<tr>
<td>PsycINFO</td>
<td>CHHS</td>
<td>5536</td>
</tr>
<tr>
<td>Social Work Abstracts</td>
<td>CHHS</td>
<td>310</td>
</tr>
<tr>
<td>SPORTDiscus</td>
<td>CHHS</td>
<td>292</td>
</tr>
<tr>
<td>STATIRef Online Electronic Medical Library</td>
<td>CHHS</td>
<td>498</td>
</tr>
</tbody>
</table>
Figure 2. Example CHHS data as a pivot chart.
Figure 3. All colleges as a pivot chart.
Figure 4. Usage by semester as a pivot table.
Figure 5. Usage by year as a pivot table.

<table>
<thead>
<tr>
<th>Row Labels</th>
<th>Sum of Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHHS</td>
<td>12,889</td>
</tr>
<tr>
<td>MEDLINE</td>
<td>12,889</td>
</tr>
<tr>
<td>2006</td>
<td>2,277</td>
</tr>
<tr>
<td>Fall</td>
<td>1,760</td>
</tr>
<tr>
<td>Summer</td>
<td>517</td>
</tr>
<tr>
<td>2007</td>
<td>3,879</td>
</tr>
<tr>
<td>Fall</td>
<td>1,796</td>
</tr>
<tr>
<td>Spring</td>
<td>1,546</td>
</tr>
<tr>
<td>Summer</td>
<td>537</td>
</tr>
<tr>
<td>2008</td>
<td>5,107</td>
</tr>
<tr>
<td>Fall</td>
<td>1,719</td>
</tr>
<tr>
<td>Spring</td>
<td>2,420</td>
</tr>
<tr>
<td>Summer</td>
<td>968</td>
</tr>
<tr>
<td>2009</td>
<td>1,626</td>
</tr>
<tr>
<td>Spring</td>
<td>1,626</td>
</tr>
<tr>
<td>Grand Total</td>
<td>12,889</td>
</tr>
</tbody>
</table>
**Figure 6.** Figure 5 collapsed to focus on usage by years.

<table>
<thead>
<tr>
<th>Row Labels</th>
<th>Sum of Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHHS</td>
<td>12,889</td>
</tr>
<tr>
<td>MEDLINE</td>
<td>12,889</td>
</tr>
<tr>
<td>2006</td>
<td>2,277</td>
</tr>
<tr>
<td>2007</td>
<td>3,879</td>
</tr>
<tr>
<td>2008</td>
<td>5,107</td>
</tr>
<tr>
<td>2009</td>
<td>1,626</td>
</tr>
<tr>
<td>Grand Total</td>
<td>12,889</td>
</tr>
</tbody>
</table>