MAKING USAGE DATA MEANINGFUL

by

Jill Morris and Emily Guhde

Abstract

This paper reports the findings of a study of licensed database usage among libraries in the NC LIVE consortium. Researchers developed North Carolina-based library peer groups in order to build context for libraries' usage data reports and to identify benchmarks and trends across those libraries that are top performers within each group. Additionally, researchers examined the use of selected databases across multiple library types to determine whether certain library characteristics or activities are related to database use. Researchers found that a number of library characteristics and activities predict database use, but the results vary depending upon the type of library and the database studied.

Acknowledgements

In completing this report, we would like to thank a number of people and organizations, without which this study would not have been possible. First and foremost, we would like to thank ER&L and EBSCO for awarding us the ER&L/EBSCO Information to Inspiration Fellowship that supported the goals of this research. A special thank you to Bonnie Tijerina, Sandy Tijerina, and the ER&L/EBSCO Information to Inspiration Selection Committee for their assistance, questions, and feedback throughout this process.

We would also like to thank a number of others who have worked in support of our efforts, including the members of the Usage Data Advisory Group, our vendor partners, and the NC LIVE staff. Thank you to the NC LIVE Librarians Council and the members of NC LIVE advisory committees for your encouragement and ideas. Thank you to Tamara Rhodes, former NC LIVE intern, for her assistance in developing a literature review on this topic.

Thank you to Michael Kowolenko (NCSU Poole College of Management) for inspiring us to ask good questions during his keynote address at the NC LIVE Usage Data Summit. Thank you to Maureen Berner (UNC School of Government) for her assistance in developing our statistical analyses, and to Irina Cassie Kozyrkov (NCSU Department of Statistics) for providing us with statistical consultation. Finally, thank you to Mary Sizemore, High Point Public Library Director, and Tim Bucknall, Assistant Dean of University Libraries at UNC-Greensboro Jackson Library, for hosting meetings of the Usage Data Advisory Group.

Usage Data Advisory Group Members:
Robert Antill, Fayetteville Technical Community College
Susanna Boylston, Davidson College
Patrick Carr, East Carolina University
Deana Guido, Nash Community College
Debbie Hargett, Wingate University
Derrik Hiatt, Wake Forest University
Terry Hill, Durham County Public Library
Doug Short, Central Piedmont Community College
Mary Sizemore, High Point Public Library
John Vickery, NC State University
Rob Wolf, UNC-Pembroke
Lynda James-Gilboe, ProQuest
Dawn Zehner, ProQuest
Scott Bernier, EBSCO
Steve O'Dell, EBSCO
Tim Bucknall, UNC-Greensboro and Carolina Consortium (Ex-Officio)
Joyce Chapman, State Library of North Carolina (Ex-Officio)

Vendor Partners:
EBSCO
Geographic Research, Inc.
LearningExpress
ProQuest

NC LIVE staff:
Nitin Arora
Sarah Dooley
Heather Greer Klein
Tim Rogers
Scott Ross

NC LIVE Librarians Council Executive Committee:
Suvanida Duangudom
Greg Needham
Susan Nutter
Cal Shepard
Kathy Winslow

Introduction

NC LIVE, a North Carolina library consortium, provides a core collection of digital content and services to 201 public and academic libraries across the state. For years, NC LIVE has provided its member libraries with monthly usage data reports. Although these reports can help libraries compare their own usage over time, they provide no indication of what "high" or "good" usage looks like for any individual library. Additionally, the data give no indication of how a library might adjust their efforts to increase use and, thereby, decrease overall cost-per-use. NC LIVE staff noticed a lack of professional literature to support the creation of usage data targets, especially for licensed library databases within consortial settings. These usage data targets are important for supporting collection development decisions at the consortial and individual library level, resource allocation decisions, and activities related to maximizing libraries' use of these databases and return on investment.

In April 2013, with the support of the ER&L/EBSCO Information to Inspiration Fellowship, NC LIVE staff members began a research study that was designed to address this lack of understanding about database usage and the factors that could impact use. They designed the study to achieve four objectives that would contribute not only to NC LIVE member libraries' understanding, but to the professional literature as well:

1) Identify peer groups of North Carolina libraries
2) Identify data points to measure usage of each studied database
3) Develop a framework for creating usage benchmarks in each peer group
4) Analyze and report qualities of high usage libraries

Methodology

To achieve these objectives, researchers limited the scope of this study to five library databases that are currently provided by NC LIVE to all of its members: Academic Search Complete (EBSCO), MasterFILE Complete (EBSCO), Wall Street Journal (ProQuest), LearningExpress Library (LearningExpress), and SimplyMap (Geographic Research, Inc.). They also assembled a Usage Data Advisory Group (UDAG) consisting of library staff members from North Carolina public and academic libraries, as well representatives from the State Library of North Carolina and database vendors. The UDAG was responsible for providing input to and feedback on the planning and assessment of NC LIVE activities related to database usage data in North Carolina libraries, and they served as a sounding board throughout this study. Through their collaborative efforts, they developed library peer groups and definitions of use, which provided the foundation for this study.

Peer Groups

At their first meeting in June 2013, the UDAG developed North Carolina library peer groups. Working in small groups divided by library types (community college libraries, four-year college and university libraries, and public libraries), the UDAG used institutional and demographic data from aggregated data

sets[1] and their prior knowledge of the organizations to form 20 peer groups.[2] Table 1 shows how these peer groups are broken out by library type. The finalized peer groups are included in Appendix 1.

Table 1: Peer Groups by Library Type		
Library Type	**Number of Peer Groups**	**Number of Libraries**
Community college libraries	7	58
Four-year college and university libraries	6	52
Public libraries	7	75
TOTALS	**20**	**185**

Definitions of Use

Members of the UDAG also worked with researchers to develop definitions of use for each of the five databases included in this study. Because most database vendors provide a number of metrics to measure different aspects of use, researchers thought that it was important to consult with both vendors and librarians to understand which data points best represent use that could be compared across libraries. Ultimately, one data point was chosen for each database, and all data points were converted into a ratio to control for population size. Table 2 displays the definitions of use for each database included in this study.

Table 2: Definitions of Use		
Database	**Definition of Use for Community College Libraries and Four-Year College and University Libraries**	**Definition of Use for Public Libraries**
Academic Search Complete	Full-Text Views Per Full Time Enrollment	Full-Text Views Per 5,000 of the Legal Service Population
MasterFILE Complete	Full-Text Views Per Full Time Enrollment	Full-Text Views Per 5,000 of the Legal Service Population
Wall Street Journal	Full-Text Views Per Full Time Enrollment	Full-Text Views Per 5,000 of the Legal Service Population
LearningExpress Library	Number of eCourses, eBooks and Practice Tests Added Per Full Time Enrollment	Number of eCourses, eBooks and Practice Tests Added Per 5,000 of the Legal Service Population
SimplyMap	Number of Sessions Initiated Per Full Time Enrollment	Number of Sessions Initiated Per 5,000 of the Legal Service Population

---

[1] Institutional and demographic data were collected from the Academic Libraries Survey (2010) from the National Center for Education Statistics and the Public Library Survey (2011-2012) from the State Library of North Carolina.
[2] Several libraries were excluded from this study, in some cases because their specialized nature made comparisons to other libraries difficult, and in other cases because their organizational structure significantly changed during the data collection period, complicating data collection and analysis. The libraries that were intentionally removed from this study were the North Carolina School of Science and Math, the nine North Carolina Area Health Education Center (AHEC) libraries, the State Library of North Carolina, and five public libraries: Gaston County Public Library, Lincoln County Public Library, Orange County Public Library, Person County Public Library, and Caswell County Public Library.

Benchmarks

Researchers gathered usage data from July 2012 through June 2013.[3] For each of the peer groups, a series of five charts (one for each database studied) was created to display database use. The high, low, mean and median values were calculated as benchmarks to help libraries within peer groups make meaningful comparisons, and were later shared with members of the UDAG and NC LIVE advisory committees for feedback.

Statistical Analysis of Usage

To begin studying the relationship between library characteristics and database use, researchers gathered data within five categories: 1) Access & Authentication[4], 2) Awareness, Outreach & Support[5], 3) Library Collections & Content[6], 4) Library Characteristics[7], and 5) Community Characteristics[8]. Much of the data for this study was available from publicly available, aggregated data sources, but researchers also distributed a survey among their member libraries to understand more about specific library activities, especially as they relate to NC LIVE-provided databases.[9] Data for approximately 50 variables was gathered for this study, and a full list of variables is available in Appendix 3 for each of the three library types.[10] Three types of analyses were completed to understand the relationships between library variables and database use: cross-tabs, difference of means tests, and multiple regression analyses. Each of these analyses was performed using statistical analysis software (SPSS).

Cross-tabs

A cross-tabs analysis helped researchers understand more about the commonalities among top libraries across peer groups. Each peer group was divided into thirds in terms of their Academic Search Complete use,[11] and libraries were categorized as top, middle or bottom. Then all top, middle and

---

[3] The usage data for this study was gathered from standard vendor reports, not COUNTER reports. This choice was made because NC LIVE does not use the COUNTER reports for their monthly reports. Researchers were confident in the use of standard reports because they have been monitored by NC LIVE staff for consistency.

[4] Access & Authentication includes variables related to the means and methods by which patrons (end-users) reach or arrive at a database. This may include technological and policy-based decisions, interfaces, and access points.

[5] Awareness, Outreach & Support includes variables related to library efforts regarding NC LIVE, such as library staff development, patron training and instruction, and marketing.

[6] Library Content & Collections includes variables related to library materials, databases and other information made available to patrons in both print and electronic form.

[7] Library Characteristics includes variables related to descriptive information regarding library operations, staffing and services provided.

[8] Community Characteristics includes variables related to demographics and descriptive information about a library's patron- or customer-base.

[9] Survey response rate and full text of the surveys is included in Appendix 2.

[10] Many variables were recoded for analysis purposes into dummy variables, increasing the number of variables significantly. There were 90 variables for public libraries, 85 variables for community college libraries, and 84 variables for four-year college and university libraries. A full list of variables and data definitions can be found in Appendix 3.

[11] Researchers decided to perform the cross-tabs analysis only for Academic Search Complete use because it is the most broadly used database across all member libraries and therefore will provide the most applicable results. Researchers were concerned that other databases could show high or low use based on library characteristics and

bottom libraries were compared to look for trends among the top-ranked libraries. This analysis was completed separately for each of the three library types (community college libraries, four-year college and university libraries, and public libraries) and provides descriptive information about the characteristics of and methods used by top libraries in their peer groups.

Difference of Means

Difference of means tests demonstrate whether or not there is a difference in terms of database use among libraries that share a certain characteristic (such as having a chat reference box) and libraries that do not share this characteristic.[12] A separate series of tests was performed for each of the three library types to examine use for all five databases included in this study.

Multiple Regression

A multiple regression analysis is a more rigorous statistical test that allows researchers to observe how each variable is related to database use, while accounting for the impact of other variables. This analysis estimates the relative strength of the relationships between each library variable and the dependent variable (use).   Because the community college libraries and four-year college and university libraries share an academic focus, as well as many data points, researchers combined these into one group for this analysis. This choice was made to strengthen the reliability of the results by increasing the number of observations. The public libraries were analyzed separately.

For each of these two groups of libraries, researchers ran a series of tests to examine Academic Search Complete use.[13] Not all variables collected for this study could be used in one regression model, so researchers first used a stepwise regression analysis to assist in the selection of variables that were likely to be related to use. After identifying potential variables for further study in the stepwise regression, researchers ran a linear multiple regression analysis, looking for a model that included actionable variables, contained at least one variable from all variable categories, and produced a high adjusted r^2 value.[14]

---

activities that were not included in this study, whereas the broad applicability of the Academic Search Complete content across all library types makes it an ideal candidate for this type of analysis. Additionally, in a 2013 analysis of shared needs, the NC LIVE Permanent Digital Resources Task Force identified "full-text journal articles" as a "high priority information need shared across all four communities of interest," which includes all libraries in this study (North Carolina Community College System, North Carolina public libraries, University of North Carolina System and the North Carolina Independent Colleges and Universities.)

[12] For interval variables, where the data gathered were numerical instead of simple "yes/no" responses, the median value was used as a cut-point for the difference of means tests to compare above-median libraries with below-median libraries. Outliers were excluded from the difference of means tests in cases where the dependent variable was more than two standard deviations away from the mean, and where the researchers' knowledge of the institutions and their usage patterns suggested that the data were not representative of expected values.

[13] See footnote 11 regarding Academic Search Complete.

[14] This analysis is intended to be used for decision-making purposes based on a snapshot of these library populations. Researchers assume that the data are somewhat stable and expect that they are representative of data that would be used for future studies. However, the methods used in this analysis trained the models on these data, and therefore, the models may not be generalizable to future instances.

Usage Data Summit

Following the initial data analysis, researchers drafted a report of their findings and shared those results with the NC LIVE Usage Data Advisory Group and members of other NC LIVE advisory committees at the February 2014 NC LIVE Usage Data Summit. This Summit was designed to produce feedback on the future of the consortium's products and services related to usage data. A report on the Summit can be found in Appendix 4.

Results and Discussion

Benchmarks and Peer Group Analysis

Charts 1-5 are examples of peer group benchmarking charts. Each chart displays use for the same peer group of four-year college and university libraries for each of the five databases studied. The charts display use that occurred between July 1, 2012 and June 30, 2013, and the high, low, mean and median values are noted to provide benchmarks and context.

**Chart 1: Academic Search Complete
Peer Group CU-E
Four-year colleges and universities**

High: 19.19
Low: 2.89
Mean: 10.12
Median: 9.48

**Chart 2: MasterFILE Complete
Peer Group CU-E
Four-year colleges and universities**

High: 3.67
Low: 0.36
Mean: 1.49
Median: 1.07
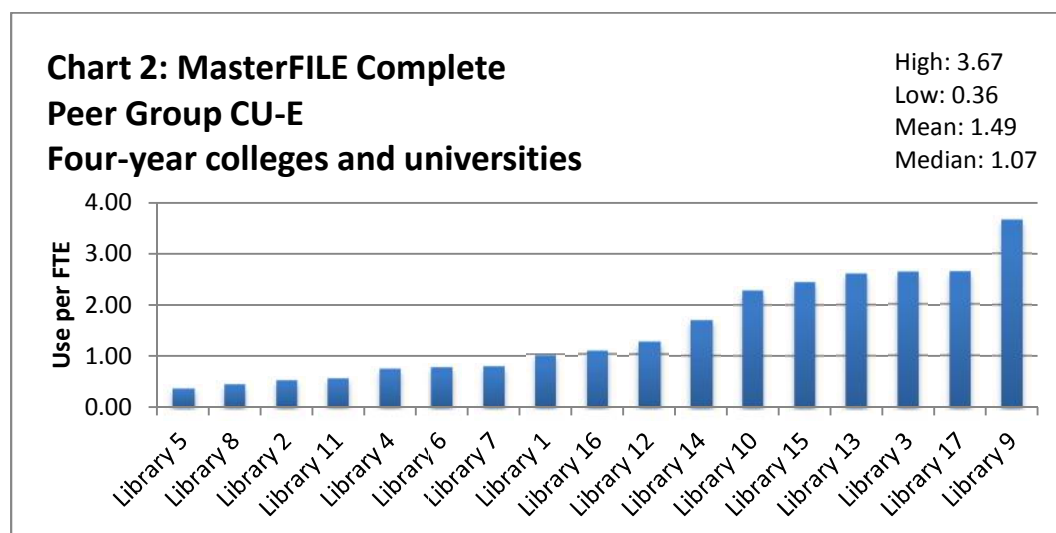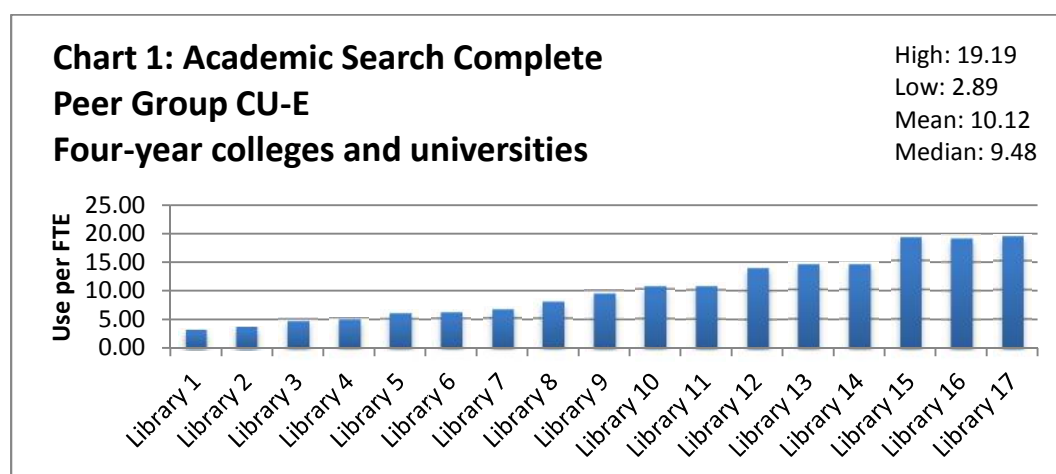
Chart 3: Wall Street Journal
Peer Group CU-E
Four-year colleges and universities

High: 1.696
Low: 0
Mean: 0.22
Median: 0.093
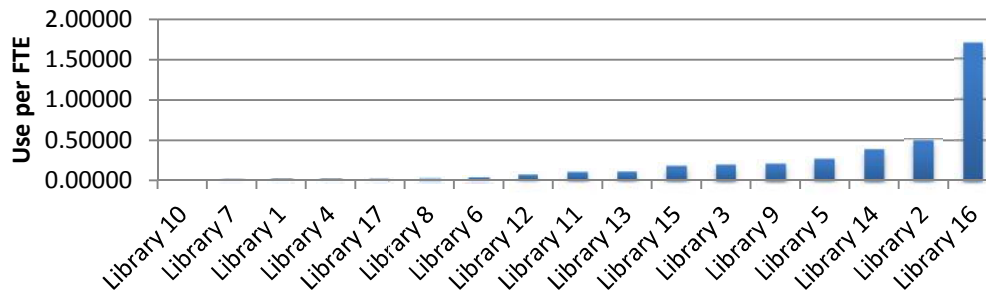
Chart 4: LearningExpress Library
Peer Group CU-E
Four-year colleges and universities

High: 0.16514
Low: 0
Mean: 0.04
Median: 0.01913
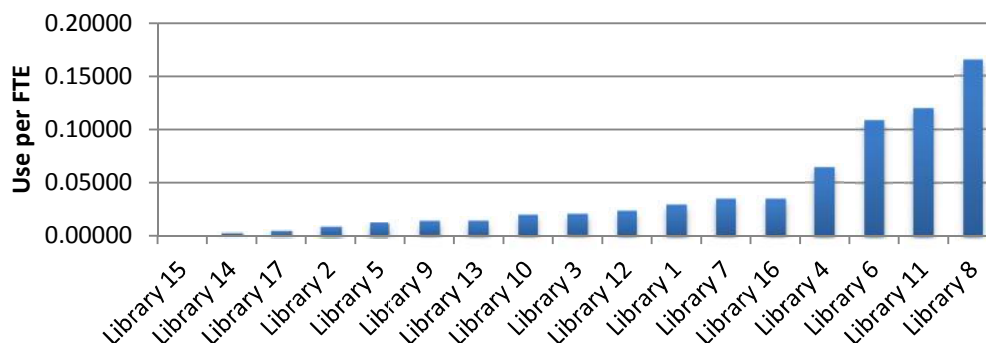
Chart 5: SimplyMap
Peer Group CU-E
Four-year colleges and universities

High: 0.0625
Low: 0
Mean: 0.014
Median: 0.0047
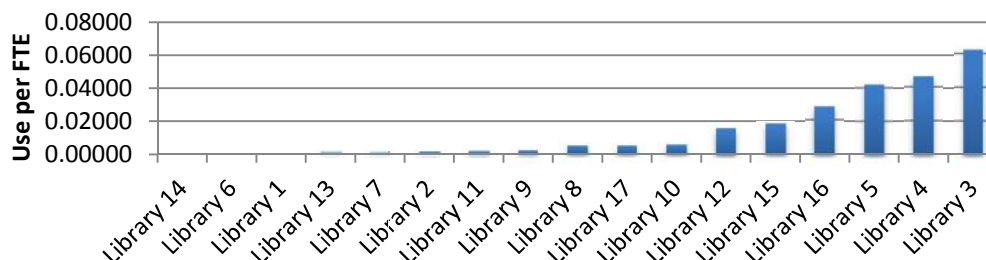
Although the high, mean, median, and low benchmarks provide important information for target-setting, researchers also wanted to understand more about the distribution of use within peer groups. Researchers expected high-use libraries to be similar across all five databases studied. However, as shown in the examples above, researchers found that libraries' use of each database varied greatly

relative to the other libraries in their peer groups. After reviewing all 20 peer groups, there was no one library that achieved the highest or lowest use across all five databases. However, libraries' rankings within peer groups were most similar between Academic Search Complete and MasterFILE Complete, perhaps because these databases offer similar content on the same EBSCOhost platform.

Researchers also looked for patterns within library types. The lowest use was divided by the highest use to create ratios for use variation in each database and peer group combination. Among community college libraries, the lowest use within a peer group was less than 10% of the highest use in 24 of 35 cases (69%). Among four-year college and university libraries, the lowest use within a peer group was less than 10% of the highest use in 16 of 30 cases (53%). Among public libraries, the lowest use within a peer group was less than 10% of the highest use in 26 of 35 cases (74%). Researchers conclude from these results that database use tends to vary widely, even among peer institutions. Also, wide variations of use within a peer group are more common among public libraries and community college libraries than among four-year college and university libraries.

High, low, median and mean use benchmarks may improve a library's ability to set goals for use, prompt the library to ask questions about their activities and the activities of their peers, compare trends, and develop profiles of high- or low-use libraries for each type of database. These benchmarks may also help NC LIVE to strategically allocate resources (such as outreach and promotional efforts) toward low-use libraries for specific marketing efforts. However, libraries' wide variation in peer group placement across databases suggests that researchers should be careful not to assume that all libraries need or want to use each database equally. Database use within peer groups should be considered on a database-by-database basis, and should not be generalized across different types of databases. Although NC LIVE may identify targets for overall use, library targets should be set at the local library level, since each library has different priorities and needs for database content.

Cross-tabs

The results from the cross-tabs analysis of top, middle and bottom libraries revealed a number of interesting trends.[15] Researchers selected a few highlights of actionable, high-use related items to include in this report, with the full list of findings available in Appendix 5.

Community college libraries

- 77% of top community college libraries use direct links to NC LIVE-provided resources, compared with 50% of middle and 53% of bottom libraries.
- 56% of top community college libraries have librarians that attend faculty meetings, compared with 45% of middle and 39% of bottom libraries.

---

[15] A statistical analysis using chi square values was not performed. The cross-tabs were only used to provide descriptive information about the distribution of top, middle and bottom libraries across peer groups.

<u>Four-year college and university libraries</u>

- 94% of top four-year public and private libraries authenticate with a local proxy, compared with 77% of middle and 56% of bottom libraries.
- 53% of both top and middle four-year public and private libraries have a high number of librarians per 1,000 of full time enrollment, compared with 35% of bottom libraries.

<u>Public libraries</u>

- 80% of top public libraries recently provided staff training for NC LIVE-provided resources, compared with 57% of middle and 55% of bottom libraries.
- 65% of top public libraries had a high number of public internet computers per 5,000 of the legal service population, compared with 46% of middle and 35% of bottom libraries.

Difference of Means

The difference of means test results point to specific relationships between each variable and database use. A positive relationship between a variable and database use means that libraries possessing that characteristic are likely to demonstrate higher use than libraries that do not possess that characteristic. An inverse relationship between a variable and database use means that libraries possessing that characteristic are likely to demonstrate lower use than libraries that do not possess that characteristic. If there is no relationship between the variable and database use, it means that there is not a significant difference in use between libraries that possess a certain characteristic and libraries that do not. In some cases, there was not enough variation among libraries to perform the difference of means tests.[16] The full results from the difference of means tests are included in Appendix 6.

<u>Community college libraries</u>

Among community college libraries, 16 variables were related to higher or lower use for at least one of the databases included in this study. A cluster of high-use related variables were found in the Library Characteristics category and included many activities related to librarian-faculty interactions, such as having embedded librarians in academic courses, librarian-initiated engagement with academic departments, library orientation programs for faculty members, marketing aimed at faculty members, and librarians that attend faculty meetings. From these results, researchers conclude that community college libraries that engage in these types of librarian-faculty interactions tend to have higher database use.

---

[16] The statistical significance of these relationships was calculated at the 95% confidence level to help researchers make comparisons among the libraries and understand patterns in the data. Even though this study examined populations of library types (as opposed to samples within these library types), the data behave as a sample since this study represents a snapshot in time for these libraries. The results indicate what we would expect to see if this study were replicated in the future for these same libraries.

<u>Four-year college and university libraries</u>

Among four-year college and university libraries, 20 of the variables were related to higher or lower use for at least one of the databases included in this study. The vast majority of relationships (both positive and inverse) were observed for the SimplyMap database. It appears that four-year college and university libraries demonstrating more technical expertise, such as authenticating with a local proxy, using a discovery service, or having a mobile library website or app, are more likely to see high use for the SimplyMap resource.

Additionally, several inverse relationships across multiple databases were notable within the Access & Authentication category, including the library's website having links to the NC LIVE website, authenticating with a password, using an NC LIVE search box on the library website, and the library website displaying descriptive text about NC LIVE. That is not to say that these library characteristics or activities are causing low database use. These characteristics all point to a less customized library website, which could mean that these libraries are lacking staff time or technical expertise to help their users get directly to database content, perhaps resulting in overall lower use.

<u>Public libraries</u>

Among public libraries, 25 of the variables were related to higher or lower use for at least one of the databases included in this study. Clusters of high-use related variables were found within all variable categories. Four variables had positive relationships across all five databases: total operating expenditures per legal service population, percent of legal service population with a bachelor's degree, population density, and the number of statistics downloads from the NC LIVE website. The first three of these variables seem very logical. In general, public libraries that spend more money per capita, have higher levels of education and more people per square mile tend to see higher database use across all databases included in this study.

The relationship between database use and the number of statistics downloads from the NC LIVE website is more surprising. These results show that public libraries with higher frequencies of downloading statistics from the NC LIVE website tend to have higher database use. Researchers are not certain what this relationship indicates, but it could mean that libraries are downloading their usage reports and then intentionally promoting resources to patrons. Alternatively, the relationship could be more subtle. Perhaps libraries that download statistics from the NC LIVE website are simply more familiar with the databases that are available, and so, are more likely to use them and suggest them to patrons. More research will be needed to explore this relationship further.

Multiple Regression

The results of the multiple regression analyses helped researchers develop models that demonstrate the relationships between independent variables and Academic Search Complete use in academic and public libraries. The results of each model are outlined below.

<u>Academic Libraries</u>

The academic libraries (community college and four-year college and university libraries) regression model included 12 variables for 94 libraries.[17]

- Authenticate with a Local Proxy
- Discovery Service
- Chat Reference Box
- Mobile Library Website or App
- Electronic Materials Expenditures per Full Time Enrollment
- NC LIVE Committee Representation
- Staff Training for NC LIVE-Provided Resources
- Total Library Expenditures per Full Time Enrollment
- Number of Librarians per 1,000 of Full Time Enrollment
- Librarian-Initiated Engagement with Academic Departments
- North Carolina Independent Colleges and Universities (NCICU) Institution[18]
- UNC Institution

The results from this model indicate that four of these variables are important predictors for academic libraries' Academic Search Complete use:

- Authenticate with a Local Proxy
- Total Library Expenditures per Full Time Enrollment
- North Carolina Independent Colleges and Universities (NCICU) Institution
- UNC Institution

Libraries in this study that authenticate with a local proxy can say with a high-degree of certainty that their implementation this authentication system has a significant, positive relationship to their use of Academic Search Complete. Though more research is needed to definitively say why the use of a proxy authentication system predicts higher database use, researchers believe it is likely due to the use of credentials that are both unique and familiar to individual users, as opposed to other methods of authentication, which require a user to first contact the library or a faculty member for a password.

This model also shows that higher total library expenditures per full time enrollment are related to lower Academic Search Complete use. Although at first this relationship seemed counterintuitive, libraries with higher expenditures may have larger local digital collections, including additional competing databases, which might cause them to rely less heavily on databases provided by NC LIVE. Additionally, libraries with higher total library expenditures per

---

[17] Output from the community college and four-year college and university libraries regression results is available in Appendix 7. Because of incomplete survey data, 94 of the 110 academic libraries were examined.
[18] The private colleges and universities in this study are all members of the North Carolina Independent Colleges and Universities.

full time enrollment tend to support higher-level graduate degree programs. Libraries supporting higher-level degree programs may focus less on promoting and using general full-text databases.

Library type also plays a role in predicting the amount of use of Academic Search Complete. Of particular importance was membership in the NCICU, which shows a strong, positive relationship to Academic Search Complete use. Although more research is needed to say why this is the case, researchers believe one potential explanation for this relationship may be that the NCICU includes many smaller institutions with smaller library budgets. Those libraries may rely more heavily on using NC LIVE-provided databases because they have fewer other databases within their online collections. Additionally, the NCICU libraries are the only NC LIVE member libraries that directly receive a bill for NC LIVE-provided databases. It is possible that these institutions are more highly aware of integrating and promoting NC LIVE-provided databases because they view these as part of their regular subscription collections. The idea that receiving a bill for databases may be connected to increased usage could indicate that NC LIVE should more clearly communicate the monetary value of their services to all libraries.

<u>Public libraries</u>

The public libraries regression model included eight variables.[19]

- Direct Links to NC LIVE-Provided Resources
- Authenticate with EasyOn (Library Card Number)
- Chat Reference Box
- Number of Statistics Downloads from the NC LIVE Website
- Patron Instruction for NC LIVE-Provided Resources
- Percent of Legal Service Population that are Registered Library Users
- Number of Public Internet Computers Per 5,000 of the Legal Service Population
- Percent of the Legal Service Population with a Bachelor's Degree

The results from this model indicate that four of these variables are important predictors for public libraries' Academic Search Complete use:

- Direct Links to NC LIVE-Provided Resources
- Number of Statistics Downloads from the NC LIVE Website
- Number of Public Internet Computers Per 5,000 of the Legal Service Population
- Percent of the Legal Service Population with a Bachelor's Degree

NC LIVE provides a website that acts as a portal into all NC LIVE-provided databases. Although this website (http://nclive.org) is available for any North Carolina library patron to use, libraries

---

[19] Output from the public libraries regression results is available in Appendix 7. Because of incomplete survey data, 56 of the 75 public libraries were examined.

sometimes bypass the NC LIVE portal in favor of routing their patrons to databases directly from their own websites. According to the results of this model, public libraries that link directly into NC LIVE-provided databases from their own websites tend to have higher use of Academic Search Complete. This result points to a potential actionable, low-cost solution to improving database use.

This analysis also shows that the number of statistics downloads from the NC LIVE website predicts Academic Search Complete use. This is a finding that is prompting researchers to question how libraries are using the statistics on the NC LIVE website, and whether the impact is intentional on the part of library staff, (e.g., "Our use is low, and we should try to increase use") or unintentional (e.g., "Academic Search Complete is heavily used at our library, so we should recommend it to patrons"). There might also be additional explanations for this relationship between statistics downloads and database use, but more research will be needed to better understand this relationship.

Finally, this model indicates that a library's user population matters with regard to use. Although libraries cannot directly impact the percent of their legal service population that holds a bachelor's degree, libraries may be able to use this information to more appropriately set goals and targets for use, given the population that they are serving. Additionally, libraries with a higher percentage of registered library users tend to demonstrate higher Academic Search Complete use. This finding could provide support for additional promotional efforts by library staff. Since the vast majority of public libraries use EasyOn (a library card number-based authentication system) to access NC LIVE resources, getting more library cards into the hands of their patron-base might increase database use.

Limitations

There were a number of limitations facing this study, the most important of which was data availability. In some cases, only one complete year of usage data was available because of NC LIVE's licensing cycle. This study was also limited in the number of databases studied. The five databases included in this study represent less than 10% of the databases that NC LIVE provides. Furthermore, the cross-tabs and regression analyses were only applied to the usage data for Academic Search Complete. This was a conservative choice, intended to provide narrow, but reliable, results. In future studies, researchers intend to find ways to include a broader perspective of database use among member libraries.

It is also important to note that the scope of this study was limited to statistical analyses that could not explain why the observed relationships exist among these data. Although researchers can offer potential explanations for the results of this study, more research is required to provide a more thorough understanding.

Conclusion

Libraries are often challenged to find ways to prove the value of their services, and usage data provides libraries with evidence for how their investments are being utilized. Increasing use of licensed databases is just one way that libraries can demonstrate the efficiency of their efforts. This study improves libraries' understanding of their database use in the context of peer libraries and specific variables, which can help them identify areas for improvement. This study also gives libraries a starting point for discussions with vendors on how they may work together for mutually beneficial solutions regarding database use.

Library consortia are in a unique position to study usage data with a wider lens and provide guidance to their member libraries. Despite this advantage, few studies have been published that provide this kind of analysis. By participating in and sharing results from these types of studies, library consortia can support libraries in their efforts to maximize their return on investment. Researchers hope that this study will serve as a starting point in encouraging libraries and consortia to identify usage targets and methods by which they may increase database use.

List of Appendices

*Note: For the purposes of this presentation and to conserve paper, only Appendix 6 has been included. All appendices are available in the complete version of this report, online. Visit http://electroniclibrarian.com for more information.

Community college libraries

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use

Access Variables

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use
Direct Links to NC LIVE-Provided Resources (yes/no)	○	○	○	○	○
Links to the NC LIVE Website (yes/no)	N/A	N/A	N/A	N/A	N/A
Authenticate with a Password (yes/no)	○	○	○	○	○
Authenticate with a Local Proxy (yes/no)	○	●-	○	○	○
A to Z List (yes/no)	○	○	○	○	○
Discovery Service (yes/no)	N/A	N/A	N/A	N/A	N/A
Federated Search Service (yes/no)	N/A	N/A	N/A	N/A	N/A
Link Resolver (yes/no)	N/A	N/A	N/A	N/A	N/A
Chat Reference Box (yes/no)	○	○	○	●+	○
NC LIVE Search Box (yes/no)	○	○	○	○	○
Vendor Search Box (yes/no)	○	○	○	○	○
Mobile Library Website or App (yes/no)	○	○	○	○	●-

Collection Variables

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use
Number of Statistics Downloads from the NC LIVE Website (median=6.5)	○	○	○	○	○
Electronic Materials Expenditures Per Full Time Enrollment (median=.000, cut point=.001)	○	○	○	○	○
Electronic Serials Expenditures Per Full Time Enrollment (median=.000, cut point=.001)	○	○	○	○	○
Number of Electronic Reference Aggregated Sources Per 1,000 of Full Time Enrollment (median=29.87)	○	○	○	○	○

Awareness Variables

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use
Number of Promotional Items Requests (median=1)	○	○	○	○	○
Number of Help Desk Tickets (median=3)	○	○	○	○	○
Descriptive Text about NC LIVE (yes/no)	●-	○	○	○	○
Visited by NC LIVE Staff (yes/no)	N/A	N/A	N/A	N/A	N/A
NC LIVE Committee Representation (yes/no)	●+	○	○	●+	○
Staff Training for NC LIVE-Provided Resources (yes/no)	○	○	●+	○	○
Staff Training for Other Electronic Resources (yes/no)	○	○	○	○	○

Cases excluded where dependent variable values were more than two standard deviations away from the mean. Statistical significance calculated for .05 level.

Community college libraries

Key
- ●+ Positive relationship observed
- ●- Inverse relationship observed
- ○ No relationship observed
- N/A Not enough variation to perform test

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use

Library Variables

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use
Gate Count Per 1,000 of Full Time Enrollment (median=832)	○	○	○	○	○
Number of Librarians Per 1,000 of Full Time Enrollment (median=1.08)	○	○	○	○	○
Total Library Expenditures Per Full Time Enrollment (median=141)	○	○	○	○	○
Number of Information Services to Individuals Per Full Time Enrollment (median=2.07)	○	●+	○	○	○
Number of Circulation Transactions Including Reserves Per Full Time Enrollment (median=4)	○	○	○	○	○
Computer Hardware and Software Expenditures Per Full Time Enrollment (median=1.67)	○	○	○	○	○
Website Maintained Externally (yes/no)	N/A	N/A	N/A	N/A	N/A
Website Maintained Internally (yes/no)	○	○	○	○	○
Website Updates Easy (yes/no)	○	○	○	○	○
Staff Development Less than 1% of Library Budget (yes/no)	○	○	○	○	○
Mandatory Staff Development (yes/no)	○	○	○	○	●-
Optional Staff Development (yes/no)	N/A	N/A	N/A	N/A	N/A
Library Liaisons to Academic Departments (yes/no)	○	○	○	○	○
Embedded Librarians in Academic Courses (yes/no)	○	○	○	●+	○
Faculty-Initiated Questions and Consultations (yes/no)	●+	○	○	○	○
Improptu Faculty Visits to Library (yes/no)	N/A	N/A	N/A	N/A	N/A
Librarian-Initiated Engagement With Academic Departments (yes/no)	○	○	○	○	●+
Library Orientation Programs for Faculty Members (yes/no)	○	○	○	●+	○
Librarians Attend Faculty Meetings (yes/no)	●+	●+	●+	○	○
Information Literacy Classes Upon Faculty Request (yes/no)	N/A	N/A	N/A	N/A	N/A
Marketing Aimed at Faculty (yes/no)	○	○	○	●+	○

Community Variables

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use
Total 12 Month Full Time Enrollment (median=2236)	●+	○	○	○	○
Information Literacy in Student Learning Outcomes (yes/no)	○	○	○	○	○
Graduation Rate (median=26.5)	○	○	○	○	○

Cases excluded where dependent variable values were more than two standard deviations away from the mean. Statistical significance calculated for .05 level.

Four-year college and university libraries

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use

Access Variables

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use
Direct Links to NC LIVE-Provided Resources (yes/no)	○	○	○	○	○
Links to the NC LIVE Website (yes/no)	○	○	●-	○	●-
Authenticate with a Password (yes/no)	●-	○	○	○	●-
Authenticate with a Local Proxy (yes/no)	●+	○	○	○	●+
A to Z List (yes/no)	N/A	N/A	N/A	N/A	N/A
Discovery Service (yes/no)	○	○	○	○	●+
Federated Search Service (yes/no)	●+	○	○	○	○
Link Resolver (yes/no)	○	○	○	○	●+
Chat Reference Box (yes/no)	○	○	○	○	○
NC LIVE Search Box (yes/no)	●-	○	○	○	●-
Vendor Search Box (yes/no)	○	○	○	○	○
Mobile Library Website or App (yes/no)	○	○	○	○	●+

Collection Variables

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use
Number of Statistics Downloads from the NC LIVE Website (median=8.5)	○	○	○	○	○
Electronic Materials Expenditures Per Full Time Enrollment (median=5.37)	○	○	○	○	○
Electronic Serials Expenditures Per Full Time Enrollment (median=67.19)	○	○	○	○	●+
Number of Electronic Reference Aggregated Sources Per 1,000 of Full Time Enrollment (median=41.82)	○	○	○	○	○

Awareness Variables

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use
Number of Promotional Items Requests (median=.5)	○	○	○	○	○
Number of Help Desk Tickets (median=3)	○	○	○	○	○
Descriptive Text about NC LIVE (yes/no)	○	○	○	○	●-
Visited by NC LIVE Staff (yes/no)	N/A	N/A	N/A	N/A	N/A
NC LIVE Committee Representation (yes/no)	○	○	○	○	○
Staff Training for NC LIVE-Provided Resources (yes/no)	○	○	○	○	○
Staff Training for Other Electronic Resources (yes/no)	○	○	○	○	○

Cases excluded where dependent variable values were more than two standard deviations away from the mean. Statistical significance calculated for .05 level.

Four-year college and university libraries

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use

Library Variables

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use
Gate Count Per 1,000 of Full Time Enrollment (median=1522)	○	○	○	○	○
Number of Librarians Per 1,000 of Full Time Enrollment (median=2.66)	○	○	○	○	○
Total Library Expenditures Per Full Time Enrollment (median=464)	○	○	○	○	●+
Number of Information Services to Individuals Per Full Time Enrollment (median=1.32)	○	○	○	○	○
Number of Circulation Transactions Including Reserves Per Full Time Enrollment (median=8)	○	○	○	○	●+
Computer Hardware and Software Expenditures Per Full Time Enrollment (median=9.35)	○	○	○	○	○
Website Maintained Externally (yes/no)	N/A	N/A	N/A	N/A	N/A
Website Maintained Internally (yes/no)	○	○	○	○	●+
Website Updates Easy (yes/no)	○	○	○	○	●+
Staff Development Less than 1% of Library Budget (yes/no)	○	○	○	○	○
Mandatory Staff Development(yes/no)	○	○	○	○	○
Optional Staff Development (yes/no)	N/A	N/A	N/A	N/A	N/A
Library Liaisons to Academic Departments (yes/no)	○	○	○	○	●+
Embedded Librarians in Academic Courses (yes/no)	○	○	○	○	○
Faculty-Initiated Questions and Consultations (yes/no)	N/A	N/A	N/A	N/A	N/A
Improptu Faculty Visits to Library (yes/no)	N/A	N/A	N/A	N/A	N/A
Librarian-Initiated Engagement With Academic Departments (yes/no)	N/A	N/A	N/A	N/A	N/A
Library Orientation Programs for Faculty Members (yes/no)	○	○	●+	○	●+
Librarians Attend Faculty Meetings (yes/no)	○	○	○	○	●+
Information Literacy Classes Upon Faculty Request (yes/no)	N/A	N/A	N/A	N/A	N/A
Marketing Aimed at Faculty (yes/no)	○	○	○	○	●+

Community Variables

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use
Total 12 Month Full Time Enrollment (median=2067)	○	○	○	○	●+
NCICU Institution (yes/no)	○	○	○	○	○
UNC Institution (yes/no)	○	○	○	○	○
Information Literacy in Student Learning Outcomes (yes/no)	○	○	○	○	○
Graduation Rate (median=28.5)	○	○	○	○	●+

Cases excluded where dependent variable values were more than two standard deviations away from the mean. Statistical significance calculated for .05 level.

Public libraries

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use

Access Variables

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use
Direct Links to NC LIVE-Provided Resources (yes/no)	●+	●+	○	●+	●+
Links to the NC LIVE Website (yes/no)	N/A	N/A	N/A	N/A	N/A
Authenticate with a Password (yes/no)	N/A	N/A	N/A	N/A	N/A
Authenticate with a Local Proxy (yes/no)	N/A	N/A	N/A	N/A	N/A
Authenticate with EasyOn (yes/no)	○	○	○	N/A	○
A to Z List (yes/no)	N/A	N/A	N/A	N/A	N/A
Discovery Service (yes/no)	N/A	N/A	N/A	N/A	N/A
Federated Search Service (yes/no)	N/A	N/A	N/A	N/A	N/A
Link Resolver (yes/no)	N/A	N/A	N/A	N/A	N/A
Chat Reference Box (yes/no)	○	●+	○	○	○
NC LIVE Search Box (yes/no)	○	○	○	○	○
Vendor Search Box (yes/no)	N/A	N/A	N/A	N/A	N/A
Mobile Library Website or App (yes/no)	○	○	○	○	●+

Collection Variables

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use
Number of Statistics Downloads from the NC LIVE Website (median=7)	●+	●+	●+	●+	●+
Collections Expenditures Per Legal Service Population (median=1.77)	○	●+	●+	●+	○
Collections Expenditures as a Percent of Total Operating Expenditures (median=9.61)	○	●+	○	●+	○
Number of Licensed Databases (median=60)	○	○	○	○	○
Electronic Materials Expenditures as a Percent of Total Operating Expenditures (median=.49)	●+	●+	○	●+	○

Awareness Variables

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use
Number of Promotional Items Requests (median=2)	○	○	●+	●+	○
Number of Help Desk Tickets (median=3)	●+	●+	○	○	●+
Descriptive Text about NC LIVE (yes/no)	○	○	○	○	○
Visited by NC LIVE Staff (yes/no)	N/A	N/A	N/A	N/A	N/A
NC LIVE Committee Representation (yes/no)	N/A	N/A	N/A	N/A	N/A
Staff Training for NC LIVE-Provided Resources (yes/no)	○	○	●+	○	○
Staff Training for Other Electronic Resources (yes/no)	●+	●+	○	○	○
Patron Instruction for Computer Skills (yes/no)	●+	●+	○	○	○
Patron Instruction for NC LIVE-Provided Resources (yes/no)	●+	●+	○	○	○

Cases excluded where dependent variable values were more than two standard deviations away from the mean. Statistical significance calculated for .05 level.

Public libraries

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use
Patron Instruction for Other Electronic Resources (yes/no)	●+	●+	○	●+	○

Library Variables

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use
Total Operating Expenditures Per Legal Service Population (median=18)	●+	●+	●+	●+	●+
Number of Virtual Visits Per 5,000 of the Legal Service Population (median=7855)	●+	●+	●+	●+	○
Number of Library Visits Per 5,000 of the Legal Service Population (median=15307)	●+	●+	●+	●+	○
Number of Reference Questions Per 5,000 of the Legal Service Population (median=2671)	○	○	○	○	○
Percent of Legal Service Population that are Registered Library Users (median=54)	●+	●+	○	○	○
Number of Librarians Per 25,000 of the Legal Service Population (median=3)	○	○	○	●-	○
Number of Computer Uses Per Legal Service Population (median=.8)	○	○	○	○	○
Number of Computer Uses Per Registered Users (median=1.4)	○	○	○	○	○
Number of Public Internet Computers Per 5,000 of the Legal Service Population (median=3.48)	○	○	○	○	○
Website Maintained Externally (yes/no)	N/A	N/A	N/A	N/A	N/A
Website Maintained Internally (yes/no)	○	○	○	○	○
Website Maintained Both Externally and Internally (yes/no)	○	○	○	○	○
DSL or Cable Internet (yes/no)	○	○	○	○	○
Fiber Optic Internet (yes/no)	○	○	○	○	○
Website Updates Easy (yes/no)	○	○	○	○	○
Staff Development Less than 1% of Library Budget (yes/no)	○	○	○	○	●-
Mandatory Staff Development (yes/no)	○	○	○	○	○
Optional Staff Development (yes/no)	N/A	N/A	N/A	●+	●+
Regional (yes/no)	○	○	○	●-	○
County (yes/no)	○	○	○	○	○
Municipal (yes/no)	N/A	N/A	N/A	N/A	N/A

Community Variables

	Academic Search Complete Use	MasterFILE Complete Use	Wall Street Journal Use	LearningExpress Library Use	SimplyMap Use
Legal Service Population (median=81380)	○	○	○	●+	○
Percent of Legal Service Population with a Bachelor's Degree (median=18)	●+	●+	●+	●+	●+
Population Density (median=177)	●+	●+	●+	●+	●+

Cases excluded where dependent variable values were more than two standard deviations away from the mean. Statistical significance calculated for .05 level.