

# Frustrations and roadblocks in data reference librarianship

Alicia Kubas<sup>1</sup>

Jenny McBurney<sup>2</sup>

## Abstract

As data skills are incorporated into academic curriculum and data becomes more widely available and used in everyday life, many librarians find themselves serving as 'accidental' data librarians in their subject areas, a phenomenon widely discussed at conferences (including IASSIST), in publications, and in the field at large. Due to this evolving landscape and growing data need, it is increasingly important for librarians to be familiar with data resources and able to answer secondary data reference questions. To learn more about this area of librarianship, this study uses survey responses from librarians who answer data questions to explore the challenges and frustrations that arise from data reference questions and interactions. Our key findings reveal that frustrations are ever present in data reference regardless of how much experience a librarian has, and many frustrations arise due to factors such as patron expectations, subject-specific and data-related jargon, and data formats and accessibility, some of which are beyond the data librarian's control.

## Keywords

Library, data reference, library reference services, data services, data literacy, data librarian

## Introduction

Questions about locating hard-to-find data, including international data, historical and time series data, and microdata, are increasingly important and frequent at many types of educational institutions and libraries. Data reference is a growing area of expertise that is becoming a vital baseline skill for library professionals across subject areas, not just for data librarians, as more teaching faculty adopt data as a regular component of assignments and as data librarians assist patrons of all types in locating secondary data for their research pursuits.

In order to learn more about the landscape of secondary data reference and how librarians can better support researchers and users in this area, data-focused librarians at the University of Minnesota developed a survey targeting library staff across different library settings and geographic locations who work with data-related questions. The survey focused on the types of data questions received and from whom, how librarians tackle these questions, frustrations and roadblocks they experience, and opportunities for increasing expertise in this area of librarianship.

While many themes and issues were identified from the survey results, one important theme that emerged was the frustrations librarians and library staff face as they attempt to answer data reference questions and become more skilled in finding and accessing datasets. This paper focuses on finding out and discussing which frustrations were influenced by other variables, namely the volume of data questions received, a librarian's number of years in the field, and the type of data questions received in terms of time (time series, historical, current, etc.). Our key finding is that frustration is not impacted by years of experience, which we did not expect. Rather, data persists as a difficult source type for librarians to support, regardless of years of experience, due to issues such as format challenges, jargon that varies across disciplines, difficulty in locating certain types of data, and management of patron expectations. In addition, a number of the frustrations encountered by data librarians are out of their control, so experience does not eliminate them entirely, even though more experienced librarians may have determined better strategies for dealing with these roadblocks.

## Methods

During the fall of 2017, we developed an online survey targeting librarians who answer any secondary data questions in their work. The survey was created in Qualtrics and received IRB approval through the University of Minnesota. The survey was sent to various national, international, and local listservs and mailing lists including:

- IASSIST (International Association for Social Science Information Services & Technology)
- GOVDOC (library professionals working with government information)
- RUSA (American Librarian Association (ALA)-Reference & User Services Association)
- RSS (ALA-Reference Services Section)
- BRASS (ALA-Business Reference and Services Section)
- NMRT (ALA-New Members Round Table)
- ULS (Association of College & Research Libraries University Libraries Section)
- Minnesota Library Association (MLA)
- MLA Public Library Division
- MLA Academic and Research Libraries Division
- Lib-minndocs (Federal Depository Library Coordinators in MN, MI, and SD)
- Research and Learning division at University of Minnesota-Twin Cities

Additionally, some librarians who received the survey via an email list further shared it more broadly on their personal Twitter accounts.

The survey was open from November 27 to December 22, 2017, and a reminder email was sent to each list halfway through the open period. The quantitative data was exported from Qualtrics into a CSV file. Data clean-up was done in Google Sheets and the cleaned dataset was analyzed in SPSS. We used dichotomous variables and therefore utilized chi-squared tests in our analysis. We considered our analysis to be statistically significant if the p-value was less than 0.05. Questions that had a 'select all that apply' option were broken down into individual responses, and dichotomous variables were created from the absence or presence of the response. For example, one question had nine possible answers we considered for the analysis and therefore we created nine dichotomous variables out of that question.

Three hundred sixty respondents began the survey. Only responses where respondents continued past the demographics section and also indicated that they answered at least one data question on average per month were included in the data analysis, which amounted to 278 total responses. In the survey, we defined data as follows: *In this context, 'data' refers to existing datasets, statistics, and data points that users are trying to find, access, or cite. We are NOT referring to data collection, management, curation, or analysis.*

## Respondent Demographics

Of the 278 respondents retained in this data analysis, 87% (n=187) were from the United States, and 8% (n=21) were from Canada. Four percent (n=10) of respondents were from other countries, which included South Africa, the United Kingdom, Germany, Slovenia, Switzerland, and Zimbabwe. Two percent (n=6) did not respond to this question.

Of the 241 respondents from the United States, 12% (n=32) were from the Western census region, 39% (n=109) were from the Midwest, 19% (n=53) were from the South, and 17% (n=47) were from the Northeast. The higher number of respondents in the Midwest is likely due to our access to local listservs and organizations, including the Minnesota Library Association and the federal depositories in the University of Minnesota region of the Federal Depository Library Program.

The respondents worked at a wide range of library types. The majority (82%, n=228) worked at a University or College library, whereas 7% (n=19) were from public libraries, 3% (n=8) worked at community, technical, or tribal college libraries, and 8% (n=23) were from other types of libraries, including but not limited to state, law, or special libraries.

Additionally, respondents held a wide variety of roles at their libraries, with most people holding multiple roles. The most common duties included reference emails and consultations, instruction, and liaison or subject specialist roles. Many people also staffed a reference or information desk, performed

collection development duties, or were a functional specialist in areas such as data services or government information (Table 1).

**Table 1 - Most Common Duties of Respondents**

<b>Which duties are part of your position?</b>	<b>N respondents out of 278 possible respondents</b>	<b>% out of 278 possible respondents</b>
Reference/Research Emails & Consultations	216	78%
Instruction	203	73%
Liaison/Subject Specialist	194	70%
Reference/Information Desk	171	62%
Collection Development	167	60%
Functional Specialist (data services, government info, copyright, etc.)	131	47%
Administration	47	17%
Acquisitions	40	14%
E-Resources	40	14%
Web Development	32	12%
Metadata & Cataloging	28	10%
Other	19	7%
Archives	13	5%

## Results

To discover what frustrations librarians experience when answering data reference questions, we included a multiple-choice (select all that apply) question and a free-response question. The first question, 'What are your biggest frustrations or roadblocks when answering data questions?', allowed respondents to check as many of the answers that applied to them. The possible answers (Table 2), were chosen based on common themes in current literature and our personal experiences. The second question, 'What other frustrations or roadblocks do you experience when answering data questions?', was left open-ended since we knew the response choices would not encompass all possible answers. This paper focuses on the initial quantitative (categorical) question, but in the future we will examine the open-ended question as well. Only the first nine response options for the first question were considered during analysis, and the 'Other' and 'NA' responses were not included. In our survey, all eight response options were coded as binary variables which we refer to as frustrations throughout this paper. Some of these frustrations are roadblocks, meaning that they are external factors that are beyond the control of the data librarian. These roadblocks are encompassed within our larger

conceptualization of frustrations, which are any issues a data librarian faces including those that can be improved with experience or training.

Each of the nine provided response options were selected by at least 25% of the possible respondents, and a large majority of respondents (83%, n=231) indicated that ‘managing patron expectations’ was one of their frustrations (Table 2).

**Table 2 – Most Common Frustrations/Roadblocks**

<b>What are your biggest frustrations or roadblocks when answering data questions?</b>	<b>N respondents out of 278 possible respondents</b>	<b>% out of 278 possible respondents</b>
Managing patron expectations	231	83%
Data resources are hard to navigate	127	46%
Lack of geographic coverage	121	44%
Lack of time series and/or historical data	115	41%
Can't access data due to paywalls	103	37%
Jargon or vocabulary	94	34%
Data questions are time consuming	90	32%
Don't know where to start looking	71	26%
Data not in an easily accessible format	70	25%
Other	22	8%
N/A - I'm never frustrated with data questions!	9	3%

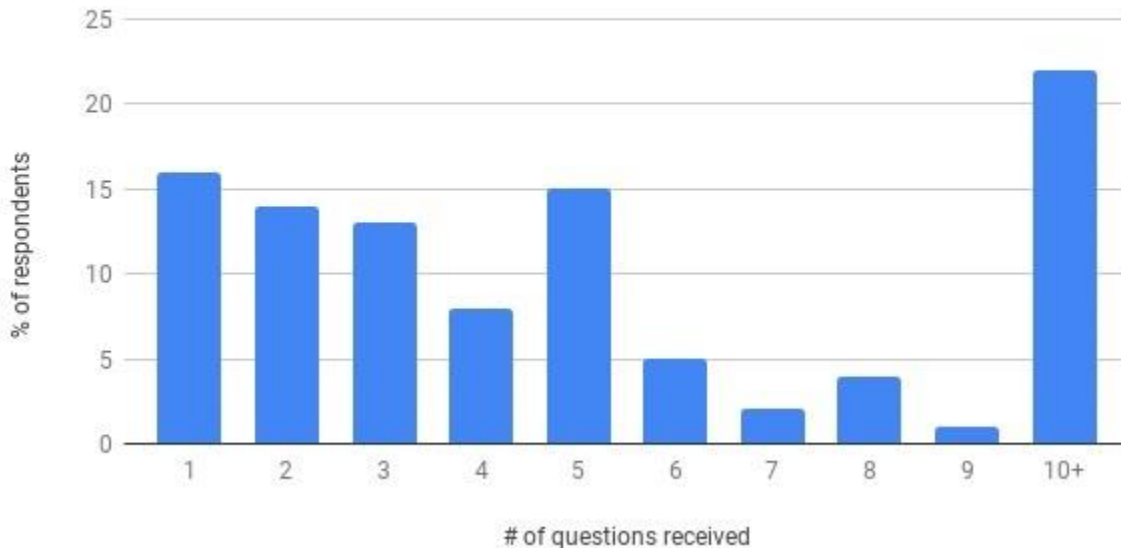
When considering these frustrations, we wanted to see if other variables from our survey were related to the frustrations a librarian experienced. We examined the relationship between the nine frustrations reported and three other multiple choice questions: Volume of Data Questions, Years of Experience in the Field, and Type of Data Questions Received in terms of time (historical, time series, current, etc.).

### **Frustrations and Volume of Data Questions**

Of the 278 individuals who answered at least one data reference question per month, 22% answered 10 or more questions (n=61), whereas just over 50% answered 1-4 questions (n=141). We did not ask those who reported answering 10 or more questions per month the exact number they received on average per month and therefore re-coded the 10+ answers as 10 for statistical analysis. This means that our statistically significant differences are very likely more significant than they appear in our current data results.

Figure 1 – % Volume of Questions Received

Question: How many data questions do you receive per month on average?



One of our areas of interest was if librarians who answer more questions are more or less likely to have certain frustrations versus librarians who answer fewer data questions. Additionally, we wondered if having more or fewer frustrations on average would depend on the number of questions received.

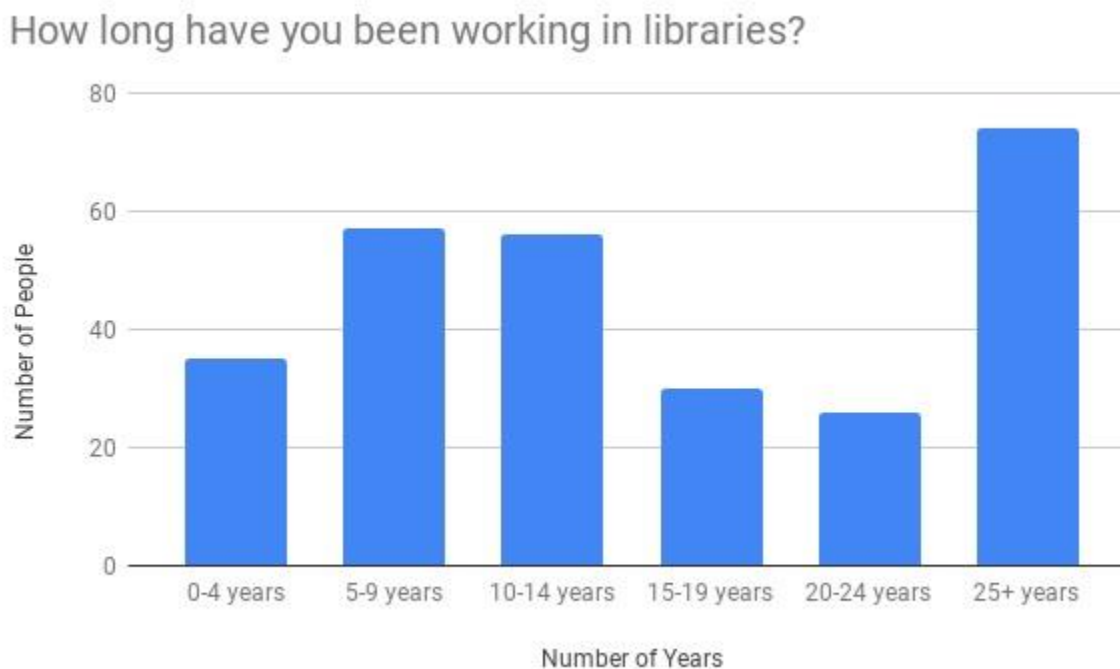
We found a statistically significant relationship when examining the number of data questions one receives by one’s self-reported frustration with jargon or vocabulary. Overall, data librarians’ likelihood of being frustrated by jargon consistently declined as their number of questions increased. Those who receive approximately five or more questions on average per month are less likely to report being frustrated by jargon or vocabulary related to the question or topic than those who answer four or fewer questions (Chi-Sq.=18.986, df=9, p=0.025). In other words, it is statistically significant that librarians who receive approximately 4 or fewer data questions per month are more likely to report being frustrated by jargon or discipline-specific vocabulary.

Finally, there was no statistically significant difference in having more or fewer frustrations overall based on the number of questions received. Thus, librarians report the same number of frustrations across the board regardless of the average number of questions received, so frustrations and roadblocks do not cease once a librarian starts regularly working on data questions.

## Frustrations and Years of Experience in the Field

The respondents also represented a wide range of experience levels (Figure 2). Early-career librarians with fewer than five years of experience represented 12.6% of respondents (n=35), whereas 26.6% were very experienced with at least 25 years of experience in libraries (n=74).

Figure 2 – Length of Time Working in Libraries



We expected that librarians who have spent more years in the field answering secondary data reference questions would report fewer frustrations on average or be less likely to have certain frustrations than those librarians who are less experienced. Surprisingly, we found no statistically significant difference in the number of frustrations reported by librarians with more experience answering data questions versus those with less experience. This suggests that data librarians, regardless of how long they have been doing this kind of work, will always be frustrated in some way with the process of doing data reference. This result is similar to what we found when looking at the relationship between frustrations and volume of data questions.

## Frustrations and Finding Historical and Time Series Data

High percentages of respondents reported that they answer questions about three time-based categories of data: historical data, current data, and time-series data (Table 3).

Table 3 – Most Common Types of Questions Received

What kinds of data questions do you receive? ( <i>Time Category</i> )	N respondents who chose this frustration out of 278 possible respondents	% out of 278 possible respondents
Historical Data	220	79%
Current Data	246	88%
Time-Series (e.g. a 20-year span of the same data over time)	186	67%

Librarians who report that they receive data questions where the patron is looking for historical data or data reported over time expressed significant variation in their frustrations. Librarians who receive historical data questions and/or time series questions are more likely to be frustrated with patron expectations around data availability (Chi-Sq.=14.826, df=1, p<0.001; Chi-Sq.=25.050, df=1, p=0.000) 23\_5 x 12\_1 and 23\_5 x 12\_3). Both types of data come with other frustrations as well: historical data questions inspire frustration with data formats that are not easily accessible (Chi-Sq.=7.929, df=1, p=0.005), and finding answers to time series data questions can be extremely time consuming (Chi-Sq.=4.636, df=1, p=0.031).

## Discussion

Librarians are increasingly called upon to work with data, even though many lack specific training and expertise. Some of the frustrations we will discuss are likely due to increasing pressure to provide data services coupled with the unique demands of supporting data and the experience--often minimal--of the librarians called upon to provide that support. These data librarians continue to play vital roles in the research lifecycle and utilize specific subject and resource expertise to locate secondary data and statistics for patrons. This is a specialized skillset that is honed with experience answering questions and time spent in the field, even though data librarians are already fulfilling many other roles in their institutions.

Many authors have described the trend toward data services in libraries. In *Databrarianship: The academic data librarian in theory and practice*, Kellam and Thompson describe major changes in data services over time, from libraries offering data files on magnetic tape in the '80s and '90s, to assisting in accessing data via the internet in the mid-2000s, to the more recent interest in data information literacy (p. 2). Data services are increasingly part of library services and the librarian's job scope, and data reference is becoming especially common for librarians who are not specifically 'data librarians.' In *Databrarianship*, Bobray Bordelon's helpful chapter 'Data Reference: Strategies for Subject Librarians' points out that as part of a subject librarian's expertise in a particular field, they will need to be aware of major data sources and types of data and methodologies common to their subject (p. 36). In *Numeric Data Services and Sources for the General Reference Librarian*, Lynda M. Kellam and Katharin Peter contend that at many institutions, the social sciences or business librarian becomes the de facto data



librarian due to the nature of requests the library staff receive, and that this can become a problem if that librarian does not have the necessary knowledge or opportunities for training in data reference (p. 3). They aptly observe that the 'accidental' data librarian is quite common (p. 151).

Some of the major frustrations we identified in our survey are also reflected in the literature. In 'The Pedagogical Data Reference Interview,' Kristin Partlo describes how students' expectations of finding data are sometimes very different from reality, and that students' misunderstandings can lead to challenging reference consultations. In our survey, of the possible frustrations or roadblocks library staff may encounter when answering data questions, by far the most highly selected was 'Managing patron expectations,' with 83% of respondents indicating that this was a major frustration.

One particular area of frustration can be understanding jargon and vocabulary that is used when talking about data or statistics or when fielding questions in specific subject areas. As Rice and Southall note in *The Data Librarian's Handbook*, 'Data-related queries tend to be more difficult, more time consuming, and sometimes even bewildering to the non-subject expert trying to provide assistance to the experienced researcher' (p. 40). They also note that acronyms, jargon, and other vocabulary should be addressed with the user during the reference interview (p. 43). We found this to be true in our survey results as well, finding that librarians who field fewer questions on average struggle with jargon and vocabulary as a frustration. Essentially, the less experience a librarian has working with secondary data inquiries, the more time she needs to spend figuring out verbiage related to how data is presented and discussed in a variety of disciplines that have data needs. This extra time required to simply understand the question may lead to more frustration. For example, receiving a data question from an economics student trying to find seasonally adjusted data on the purchasing power of a particular country could cause frustration because the librarian may be unfamiliar with what exactly purchasing power is or how it is measured, or what it means for data to be seasonally adjusted.

It is likely that data librarians will always be frustrated in some way working with data, regardless of experience level, sometimes due to factors beyond their control. Data continues to be a specialized source type that requires managing user expectations, a great deal of on-the-spot instruction, and navigating a non-standardized discovery environment, which contributes to the level of difficulty and frustration when working with these questions. Kellam and Peter (2011) note that for many data librarians 'true learning comes from daily work with data-related questions,' and learning on the job is especially typical for data librarians since most come from a variety of educational backgrounds (p. 153). They interviewed a variety of data librarians about their jobs and heard from several who explained that data reference questions are very complex and time-consuming, and that the more questions a librarian answers, the more she knows about finding data and fielding questions (p. 155). Bauder (2014) further underscores this idea by noting that 'for librarians accustomed to the relatively organized world of books and journal articles, trying to find data can be a frustrating experience' (p. 11). While examples like this in the literature suggest that librarians with more experience would find answering questions

easier and report fewer frustrations overall, our survey found that there were no statistically significant differences in the number of types of frustrations between librarians with more experience answering data questions and those with less experience, based on our measure of experience which encompasses time in the field and volume of data questions. This may mean that data librarians, regardless of how long they have been doing this kind of work and the volume of questions they receive, will always be frustrated in some way with the process of doing data reference. It is possible that some frustrations could be reduced with experience (i.e. not knowing where to start looking), whereas others, that we call roadblocks, are out of the librarian's control (i.e. paywalls). Over time a librarian could learn better ways of dealing with some frustrations, such as certain strategies for managing patron expectations or knowing where to start looking for specific data, but experience will not help overcome roadblocks such as lack of geographic coverage of data or a paywall restriction. In other words, more experienced librarians could better anticipate these challenges and have strategies for dealing with these types of roadblocks, but this does not eliminate them entirely. These factors may improve naturally over time because more content could be digitized or organizations and agencies could release more data that has not been publicly available, but this is once again beyond the control of the data librarian.

While current time series data has become easier to locate due to improved online availability, finding historical data has remained difficult due to various issues including geographic, methodological, and sampling technique changes over time. Another major issue that comes into play is format changes over time, which we found to be true from our survey where librarians expressed frustration with data formats when looking for historical data. Older data is often available only in print, which means that for data analysis purposes, the data will need to be manually entered, likely causing a disruption in patron expectations. Furthermore, older data is often exclusively available in print or, particularly in the case of government-produced data, in obsolete formats like floppy disks and CDROMs, where an emulator is needed to run software to access the files or particular hardware is required. This is particularly frustrating to the librarian who may need to take quite a bit of time to look at the data in a less accessible format, but also it may be frustrating for the patron since this is likely not what they were hoping for or expecting. Additionally, finding historical data can be frustrating for patrons and librarians when the dataset does not go back far enough to the desired time period. Kellam and Peter (2011) note that this historical data is not possible to find if the dataset wasn't collected during the requested years (p. 73). Due to these various restrictions, data librarians will often need to work with their patrons to help them understand the limitations of historical or time series data.

Our survey has some limitations that could affect our results. First, while we targeted a number of listservs that we knew included library staff who answer data reference questions, not all data librarians may be on listservs, and we may have missed other major listservs that we are not currently aware of that would have broadened our range of participants. There could be certain types of librarians who answer data reference questions but for whatever reason do not make use of one of the listservs we included. Furthermore, the survey was shared informally by participants with other librarians that they knew through channels such as email and Twitter, thereby increasing the audience but not in a way that

we can track. Generally, we do not know the complete size of the target audience or how many potential respondents we reached with our survey distribution, so we cannot calculate the overall response rate or estimate the nonresponse bias. We also had a much greater number of responses from the Midwest region of the United States, since that is where we are located and we have access to networks of librarians in this area. We further have very low representation from other countries, again due to our access to networks. These various limitations to our survey participation could lead to regional bias in our results.

In our analysis of our survey of librarians who answer data questions, we found that there are many frustrations that they experience in their roles as data librarians. One of our key findings is that frustrations will always exist regardless of a librarian's years of experience. Many factors contribute to the experience of frustration, particularly the popularity of historical and time series data questions. While answering many questions exposes the data librarian to a wide variety of jargon and vocabulary, years of experience in answering data questions does not eliminate the tendency to be frustrated. There will always be aspects of data reference that librarians have little to no control over that lead to frustration, but there are some strategies librarians could employ to lessen and better cope with their frustrations and roadblocks and successfully contribute to data services within the library. More experienced data librarians can share their expertise with colleagues, especially subject liaisons and 'accidental' data librarians who may have less experience with data. Librarians at all levels can also seek out opportunities to expand their repertoire and depth of knowledge of data sources. They can also advocate for administrative support in training, professional development funding, increased staffing in data services, and opportunities to contribute to strategic planning and discussion at larger structural levels. Librarians can also develop strategies for improving data literacy among users, including targeting courses that incorporate data components, being prepared to cover basic data concepts during brief reference interviews, and aligning user expectations with the realities of data reference and discovery.

In recognition of this challenging area of librarianship, future directions of this research will include analysis of data librarians' free text responses of suggested strategies for dealing with tough questions as well as more focused looks into the answers of academic librarians. We are particularly interested in the responses of academic librarians as they made up the majority of respondents and we are interested in learning about the experiences of other librarians in similar roles to ourselves.

## Acknowledgements

We thank Andrew Kubas for his critical role in advising on data analysis and methodology and running analysis in SPSS. We also thank Danya Leebaw, Amy Riegelman, Carl McBurney, and the peer reviewers for their feedback and critique.

## References

Bauder, J. (2014) *The Reference Guide to Data Sources*. Chicago: ALA Editions.

Bordelon, B. (2016) 'Data Reference: Strategies for Subject Librarians' in Kellam, L. and Thompson, K. (eds.) *Databrarianship: The Academic Data Librarian in Theory and Practice*. Chicago: Association for College and Research Libraries, pp. 35-49.

Kellam, L. and Thompson, K. (eds.) (2016) *Databrarianship: The Academic Data Librarian in Theory and Practice*. Chicago: Association for College and Research Libraries.

Kellam, L. and Peter, K. (2011) *Numeric Data Services and Sources for the General Reference Librarian*. Oxford: Chandos Publishing.

Partlo, K. (2010) 'The Pedagogical Data Reference Interview', *IASSIST Quarterly*, 33(4), pp. 6-10. Available at <https://iassistquarterly.com/index.php/iassist/article/view/884/876>

Rice, R. and Southall, J. (2016) *The Data Librarian's Handbook*. London: Facet Publishing.

# Appendix: Copy of survey questionnaire

## Answering Secondary Data Questions in a Library Setting

Welcome!

The goal of this survey is to collect information from librarians and library staff who assist users in locating secondary datasets and statistics. We hope to collect data that is rich enough to inform how this growing area of librarianship is changing and how librarians can support researchers in finding secondary data.

This research study is being conducted by Alicia Kubas and Jenny McBurney at the University of Minnesota Libraries. It should take less than 10 minutes to complete.

If you have questions about the survey, please contact Alicia Kubas (akubas@umn.edu).

Your participation is completely voluntary. All individual responses will be confidential, and you can stop taking the survey at any time. Continuing the survey indicates that you consent to participate. This study has received IRB exemption from the University of Minnesota.

### Demographics

At what kind of library do you work?

- University or College (1)
- Public (2)
- Community/Technical College (3)
- Tribal College (4)
- State (5)
- Law (6)
- Special (7)
- Other (Please describe) (8) \_\_\_\_\_

Where is your library located?

▼ United States of America (187) ... Zimbabwe (1357)

Display This Question: If List of Countries = United States of America:  
In which census region of the US is your library located?

- United States - West (1)
- United States - Midwest (2)
- United States - South (3)
- United States - Northeast (4)

Display This Question: If List of Countries = United States of America

Are you a Federal Depository Library Program Coordinator?

- Yes (1)
- No (2)

Which duties are part of your position? [check all that apply]

- Liaison/Subject Specialist (1)
- Functional Specialist (data services, government info, copyright, etc.) - Please Specify: (2)  
\_\_\_\_\_
- Reference/Information Desk (3)
- Reference/Research Emails & Consultations (4)
- Administration (5)
- Instruction (6)
- Archives (7)
- Collection Development (8)
- Metadata & Cataloging (9)
- Acquisitions (10)
- E-Resources (11)
- Web Development (12)
- Other (Please describe) (13) \_\_\_\_\_

How long have you been working in libraries?

- 0-4 years (1)
- 5-9 years (2)
- 10-14 years (3)
- 15-19 years (4)
- 20-24 years (5)
- 25+ years (6)

#### Data Questions

For this part of the survey, we are asking about how you answer data-related questions. In this context, 'data' refers to existing datasets, statistics, and data points that users are trying to find, access, or cite. We are NOT referring to data collection, management, curation, or analysis.

On average, how many data questions do you receive per month?

- 0 (1)
- 1 (2)
- 2 (3)
- 3 (4)
- 4 (5)
- 5 (6)
- 5 (7)
- 6 (8)
- 7 (9)
- 8 (10)
- 9 (11)
- 10+ (12)

Skip To: End of Survey If On average, how many data questions do you receive per month? = 0

On average, how much time do you spend on each question?

- Less than half an hour (1)
- Half an hour to an hour (2)
- Between 1-2 hours (3)
- More than 2 hours (4)

What kinds of data questions do you receive? [check all that apply in the 4 categories below]

Geography: [check all that apply]

- Local or state related (1)
- Regional (e.g. Midwest) (2)
- National (3)
- International (4)
- Other (please describe) (5) \_\_\_\_\_

Time: [check all that apply]

- Historical Data (1)
- Current Data (2)
- Time-Series (e.g. a 20-year span of the same data over time) (3)

Miscellaneous: [check all that apply]

- Microdata (1)
- Spatial Data (2)
- Study Data - data already collected for research that is now available for reuse (3)
- Other (please describe) (4) \_\_\_\_\_

Topical Areas: [check all that apply]

- Business/Economics (1)
- Agriculture (2)
- Politics (3)
- Education (4)
- Tourism/Culture (5)
- Health (6)
- Demographic (7)
- Environment (8)
- Psychology (9)
- Science (10)
- Weather (11)
- Astronomy (12)
- Other (please describe) (13) \_\_\_\_\_

What are your strategies for answering these data questions? [check all that apply]

- Google it to get background information (1)

- Consult reference tools like Wikipedia or free or subscription encyclopedias (2)
- Ask the patron follow-up questions/Reference interview (3)
- Use personal knowledge of sources (4)
- Consult research guides (5)
- Consult with colleagues at your institution (6)
- Consult with colleagues outside of your institution (listserv, call or email, etc.) (7)
- Citation pearl growing (using a citation or other piece of information to find additional information) (8)
- Other (please describe) (9) \_\_\_\_\_

What types of resources do you use to find data? [check all that apply]

- Library subscription databases (1)
- Freely available databases (2)
- Government websites/portals (3)
- Data repositories (ICPSR, etc.) (4)
- NGOs or intergovernmental organizations (UN, EU, WorldBank, etc.) (5)
- Analog data (data in physical form such as print, cassette, CD-ROM, etc.) (6)
- Other (please describe) (7) \_\_\_\_\_

How successful are you at finding the data for which you are looking?

- Always find it (1)
- Usually find it (2)
- Occasionally find it (3)
- Rarely find it (4)
- Never find it (5)

Do you teach data literacy topics (e.g. methodology, authority, sample size, public availability, proprietary information, etc.)? [check all that apply]

- Yes; as part of a college course (1)
- Yes; as part of a graduate course (2)
- Yes; as part of a college or university freestanding workshop (3)
- Yes; as part of a non-college/university freestanding workshop (at a public library or special library, etc.) (4)
- Yes; for your colleagues (5)
- Yes; created tutorials or learning objects (6)
- No (7)

How do you stay current on developing your personal knowledge of data sources? [check all that apply]

- Conferences (1)
- Webinars (2)
- Blogs or websites (3)
- Scholarly publications (4)
- Other (please describe) (5) \_\_\_\_\_

Demographics of Patrons



From whom do you receive data questions? [check all that apply]

- Community members (1)
- K-12 students (2)
- Undergraduate students (3)
- Graduate students (4)
- Faculty (5)
- Other librarians and colleagues (6)
- Other (please describe) (7) \_\_\_\_\_

Considering data literacy skills (e.g. understanding methodology, authority, sample size, public availability, proprietary information, etc.), on a scale of 1 - 5, 1 being very weak data literacy skills and 5 being very strong data literacy skills, how data literate are your typical users?

- 1 (Very weak) (1)
- 2 (2)
- 3 (3)
- 4 (4)
- 5 (Very strong) (5)

Challenges and Opportunities

What are your biggest frustrations or roadblocks when answering data questions? [check all that apply]

- Don't know where to start looking (1)
- Can't access data due to paywalls (2)
- Data resources are hard to navigate (3)
- Jargon or vocabulary related to the question or topic (not knowing what seasonally-adjusted data is, etc.) (4)
- Managing patron expectations around what data actually exists or its availability (5)
- Lack of geographic coverage (other countries, local data, etc.) (6)
- Lack of time series and/or historical data (data from 1952, data from 1970s to present, etc.) (7)
- Data not in an easily accessible format (in print, on CD-ROM, microfiche, etc.) (8)
- Data questions are time consuming (9)
- N/A - I'm never frustrated with data questions! (10)
- Other (you will be able to describe in the next question) (11)

What other frustrations or roadblocks do you experience when answering data questions?

---

What tips, tricks, or strategies for answering tough data questions do you have to share with other library staff?

---

What data source do you use most often?

---

---

## End-notes

<sup>1</sup> Alicia Kubas is the Government Publications and Data Librarian as well as the Regional Depository Coordinator at the University of Minnesota Libraries and can be reached by email at [akubas@umn.edu](mailto:akubas@umn.edu).

<sup>2</sup> Jenny McBurney is the Economics Librarian, liaison to the Institute for Advanced Study, and the Research Services Coordinator for Social Sciences and Professional Programs at the University of Minnesota Libraries and can be reached by email at [jmcburne@umn.edu](mailto:jmcburne@umn.edu).