REPORT

TEACHING WITH DATA IN THE SOCIAL SCIENCES AT KANSAS STATE UNIVERSITY: HOW CAN K-STATE LIBRARIES SUPPORT UNDERGRADUATE INSTRUCTION?

July 2022

Laura Bonella
Academic Services Department Head, Associate Professor

Emily Finch
Scholarly Communication and Copyright Librarian, Assistant Professor

Kendra Spahr
Academic Services Librarian, Assistant Professor

KANSAS STATE UNIVERSITY

TABLE OF CONTENTS

Introduction 2

Methods 3
- Defining Teaching With Data 3
- Participants 3
- Methodology 4

Teaching with Data 4
- Prerequisites and Course Sequencing 4
- Technological Competency 5
- Instructional Materials 5
- Data Sources for Instruction 7
- Students Finding and Collecting Data 8

Instructional Themes 9
- Understanding Data and Data-Related Concepts 9
- Understanding Software 10
- Evaluating Data Sources and Publications 10
- Manipulating and Analyzing Data 11
- Ethics 11

Pedagogical Goals. 12
- Self-Concept 12
- Job Skills 13
- Being Informed Citizens 14

Pedagogical Support 15
- Teaching Support 15
- Learning to Teach With Data 15

Technology 16
- Spaces and Resources 16
- Equitable Access to Technology 17

Other Campus Resources 17

Conclusion 18
- Access to Technology is Fundamental in Teaching With Data 18
- Teaching With Data is More Than Teaching Technological Competency 19
- Teaching With Data Means Different Things in Different Disciplines 19

Recommendations 20
- Developing Formal Data Services at K-State Libraries 20
- Creating and Sharing Teaching Materials 21
- Providing Data Sources and Tools 21
- Finding, Accessing, and Working With Data 22

Appendices 23
- Appendix A: Departments That Social Sciences Librarians Support at K-State 23
- Appendix B: Semi-Structured Interview Guide 23

INTRODUCTION

A growing number of academic libraries are developing services to support instructors and researchers who are working with data. A recent report by the Association of College and Research Libraries—2020 Top Trends in Academic Libraries—identified research data services as an important trend. Like other academic libraries, Kansas State University Libraries is planning for the development of research data services to address needs on our campus. This report presents findings from interviews with social sciences instructors at Kansas State University (K-State) who teach undergraduates to use quantitative data. Based on the findings, we developed recommendations for K-State Libraries as we work to develop data services for our local institution.

K-State, a land-grant institution, is home to more than 22,000 students across three campuses, with more than 250 undergraduate majors and programs and over 165 graduate degrees and certificates. At K-State, students can pursue education in social sciences fields, earning bachelor's, master's, and doctoral degrees. Although the need for learning about data may seem more evident in science disciplines, the discovery, collection, and analysis of data is an important focus in social sciences disciplines as well. K-State's social sciences librarians support 19 departments (see Appendix A) and are interested in how research data services might benefit students and faculty in these programs.

In May of 2020, K-State Libraries established a local research team to partner with Ithaka S+R on their multi-institutional study on teaching with data in the social sciences. The project was developed to gather information about how undergraduate instructors in the social sciences teach with quantitative data. For the project, 20 universities collected data, which was incorporated into a synthesized report by Ithaka S+R. More information about the project can be found in this Ithaka S+R blog post. K-State's team members were part of the Libraries' Academic Services Department, including a social sciences/education librarian who is the department head, a social sciences/business librarian, and a scholarly communications librarian.

By participating in this research project, we identified faculty needs when teaching with data in order to inform our development of research data services. In this report, we will discuss the research methodology, the themes discovered through analysis of the interview data, and recommendations based on those themes that will help K-State Libraries support teaching with data. While this report focused on our institution, we hope the recommendations will be useful for librarians at other institutions as well.

METHODS

The goal of the study was to examine the teaching practices and support needs of K-State instructors who teach undergraduate courses involving quantitative data. K-State librarians wanted to develop an understanding of social sciences pedagogy in these courses and to explore how libraries could better support programs where students are actively engaged in collecting, finding, and analyzing data. We used qualitative research methods to collect and analyze data. In this section, we describe how teaching with data was defined, how participants were selected, and the process of interviewing participants and analyzing the data.

Defining Teaching With Data

Looking specifically at teaching with quantitative data in undergraduate classes, this study examined not only how K-State instructors presented and used data in their courses, but also how they engaged in deeper and more robust data instruction. For this study, teaching with data included the following activities:

- Gathering data through social sciences experiments, surveys, or other research
- Searching for appropriate datasets to address a particular research problem
- Cleaning, analyzing, mining, visualizing, or otherwise manipulating data
- Drawing narratives or conclusions from data
- Learning to use specific tools, software, or programming languages to work with data

Participants

The research team used purposive sampling to select participants who teach undergraduates in the social sciences using data. The team recruited participants through email, intentionally reaching out to instructors from a range of ranks and disciplines (see Figure 1). We selected and interviewed ten social sciences instructors across seven programs (Agricultural Education, Communication Studies, Economics, Geography & Geospatial Sciences, Journalism & Mass Communications, Psychological Sciences, and Sociology).

Interviewee's Professional Ranks	
Graduate Teaching Assistant	1
Instructor	2
Visiting Assistant Professor	1
Assistant Professor	1
Associate Professor	2
Professor	3

Figure 1. Interviewee's Professional Ranks

Methodology

Researchers conducted one-on-one interviews with instructors using a semi-structured interview guide provided by Ithaka S+R (see Appendix B). The semi-structured guide provided flexibility, allowing researchers to use probes and follow-up questions to elicit additional information. Interviews were conducted and recorded using Zoom, a video conferencing software. Interview transcripts were generated by Zoom's automated audio transcription tool and then edited for accuracy and clarity by the research team.

To analyze the data, each member of the team individually coded the same three transcripts and generated open codes. Using the codes generated in the first round of coding, the team developed broad themes that we used to code all of the transcripts.

TEACHING WITH DATA

The first group of themes to emerge from our analysis involved teaching with data. These themes included information about prerequisites and course sequencing, students' technological competency, instructional materials, and data sources for instruction.

Prerequisites and Course Sequencing

In our interviews, we asked instructors about skills students should have for their classes and the ways they expected students to develop those skills, such as prerequisite classes. With the exception of geographic information systems (GIS), most of the instructors did not expect their students to come into class with specific knowledge. In the GIS program, students were required to complete a series of increasingly advanced courses where they worked with geospatial data. Each course teaches new content that builds on previous content. However, students and faculty outside of the department didn't always appreciate the importance of the course prerequisites. The GIS program sometimes allows students to take intermediate or advanced courses without prerequisites but works closely with those students to ensure they are successful in the course.

In other programs, undergraduates might work with data in one or two required courses, and the prerequisites for taking those courses were minimal. For example, psychological sciences and sociology both had a required research methods course with a lab where students work with data. Sociology students were required to take a statistics course before the research methods course in their program but were not expected to have prior experience using statistical analysis software. Psychology students were required to take an introductory psychology course before enrolling in the research methods course.

Journalism is a discipline with a growing emphasis on data, and this program is increasing course offerings where undergraduates work with data. Data journalism is a required course in the journalism core, and faculty are developing more advanced courses in this area. To prepare students for working with data, the program developed a series of required, one-credit foundational courses to introduce students to skills such as coding and creating data visualizations. Students build on these skills in more advanced courses.

Technological Competency

Social sciences instructors relied on a variety of technologies to effectively teach with data and often found it necessary to help students become more fluent when using different technologies (see Figure 2). Interviewees indicated that student comfort and competence with the technologies in data-heavy courses shaped expectations for instructors in regard to student learning outcomes.

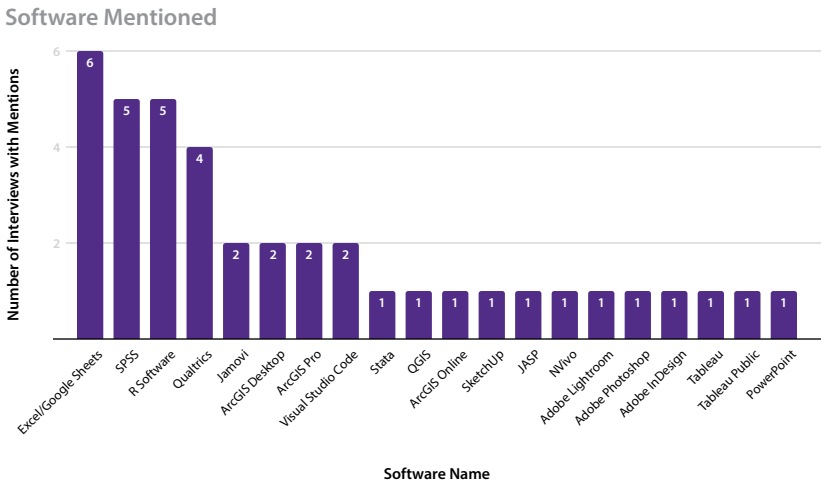

Figure 2. List and Counts of Software Mentioned During Interviews

Instructors noted students' limited technology skills could be a barrier to learning: "Some students are amazingly still uncomfortable sitting in front of a computer. Much to my surprise...your students' basic technical competency is a huge barrier that they've got to overcome in order to be able to do the things that we're trying to accomplish in class." Most instructors expected students to have basic spreadsheet skills and were sometimes surprised when they needed to provide support with basic functions in Excel.

Discomfort with basic computing posed a challenge to anticipated course outcomes, limiting how robust the data instruction could be in some courses. When one participant was asked about challenges related to students' ability to work with data, they responded:

> I think it's sometimes the learning curve of how to use the technology and how to struggle with the technology because I've had students that oftentimes once they meet the first big hurdle of a program, they're ready to just throw their hands up and say, I quit.

Fortunately, instructors showed willingness to assess the situation and assist students with developing the skills needed to pursue more difficult data coursework.

Instructional Materials

When teaching with data, instructors used a combination of commercially published, open, and homegrown educational materials. Though some instructors used commercial textbooks that included exercises with practice data, they didn't seem to have difficulty finding or creating exercises that used free data sources. In psychological

sciences, instructors for research methods courses used data from Open Stats Lab, a free resource that provides student exercises, accompanied by published journal articles with open data. An economics instructor who teaches with data used open econometrics textbooks available on other professors' websites, some of which included exercises and practice data. Instructors also referred students to online tutorials or videos to help them learn. One noted that the tutorials allowed students to learn concepts at their own pace in a space outside the classroom where they might feel less anxious.

Most undergraduate courses where students worked with data were introductory, so instructors encouraged students to continue learning outside of class, referring students to free websites, tools, and tutorials that would help them develop more advanced skills. A number of free online coding tools were suggested to students to support programming skills (see Figure 3).

Online Tools
JSFiddle
Playcode.io
HighCharts JS
Geocode (AwesomeTable)
Parsehub.com

Figure 3. List of Online Coding Tools Mentioned in Interviews

Instructors introduced students to programming languages that are useful for data analysis, encouraging students to explore deeper outside of class (see Figure 4). Because coding was usually regarded as too advanced to teach in introductory courses, instructors recommended that students learn more on their own using free or low-cost resources such as Codecademy, the Python for Journalists tutorial, and courses developed by the Knight Foundation. Becoming proficient in a particular software was beyond the scope of most courses, so this was another area that instructors encouraged students to learn outside of class. For example, one instructor introduced students to creating data visualizations with Tableau Public and emphasized the free training videos provided by Tableau.

Programming Languages Mentioned
Python
Perl
JavaScript
SQL (Structured Query Language)
R
SPSS syntax

Figure 4. List of Programming Languages Mentioned During Interviews

In addition to using open teaching materials, instructors created their own exercises and assignments or adapted materials shared by peers. Instructors teaching the same courses might share materials with each other or create materials collaboratively. Professors teaching courses with labs might work with lab instructors to update assignments from semester to semester. New instructors might start with the materials provided by their predecessors. Most instructors in this study were willing to share their teaching materials with colleagues when asked, but they did not share in more formal ways, such as posting materials in repositories or on their personal websites.

A few instructors acknowledged the library as a source of support for instructional materials. One GIS instructor recognized the library as a potential source of support for creating and distributing open learning materials. Faculty in the GIS program are considering creating an open textbook and lab manual for a series of new online courses and are aware that the library has initiatives to support faculty developing open learning materials. Another instructor assigned a book that was available as an ebook with unlimited access through the library and was appreciative that students could access this assigned reading at no cost to them.

Data Sources for Instruction

Our analysis indicated that finding data sources to use in instruction was not too challenging for instructors, but there were challenges when students were asked to find or collect their own data.

Instructors primarily used freely available data sources for teaching. This included publicly available data from organizations and government agencies and data available in open educational resources (see Figure 5). For example, a communication studies instructor taught using social media data from social networking websites or social networking software companies. A journalism instructor created data sets from local and national government sources and also used data collected, analyzed, and shared by news organizations. Instructors didn't report any major challenges when finding free data sources for teaching undergraduates.

Data Sources Used in Instruction	
GIS Instructors:	U.S. Census American Community Survey U.S. Geological Survey Environmental Protection Agency The National Map
Economics Instructors:	Bureau of Economic Analysis Bureau of Labor Statistics Federal Reserve Economic Data (FRED)
Sociology Instructors:	General Social Survey (through NORC)

Figure 5. Examples of Data Sources Used for Undergraduate Instruction

Students Finding and Collecting Data

Instructors reported that students struggled to find appropriate data sources for assignments on their own due to their limited search skills and knowledge about data sources. "I mean they know how to use Google or search engines for basic stuff in their life. But when it comes to looking for information, they are completely lost." Students' limited knowledge about data characteristics made it difficult for them to identify data that was appropriate for their purposes. A GIS instructor emphasized that to find geospatial data students needed to understand "how geospatial data is packaged and stored and made available." Additionally, lack of uniform metadata and descriptors made geospatial data difficult to locate using a search engine. Faculty also had concerns about the trustworthiness of data sources that students located using search engines. As a result, many faculty preferred to provide students with a list of trusted data sources rather than asking students to find data on their own. Based on the challenges that instructors described, students could benefit from training on searching for data as well as identifying the characteristics of trustworthy data sources.

Instructors reported that students struggled to find appropriate data sources for assignments on their own due to their limited search skills and knowledge about data sources. "I mean they know how to use Google or search engines for basic stuff in their life. But when it comes to looking for information, they are completely lost." Students' limited knowledge about data characteristics made it difficult for them to identify data that was appropriate for their purposes. A GIS instructor emphasized that to find geospatial data students needed to understand "how geospatial data is packaged and stored and made available." Lack of uniform metadata and descriptors made geospatial data difficult to locate using a search engine. Faculty also had concerns about the trustworthiness of data sources that students located using search engines. As a result, many faculty preferred to provide students with a list of trusted data sources rather than asking students to find data on their own. Based on the challenges that instructors described, students could benefit from training on searching for data as well as identifying the characteristics of trustworthy data sources.

Some programs used both pre-existing data sources and student-collected data. For example, in psychological

sciences, students analyzed pre-existing data published with studies then compared their results to the study. Later on in the semester, students collected their own data:

> Eventually they'll have their own data and they'll have had several examples of how to run different analyses and then they have to figure out which analyses apply to their own data given their design and type of data, analyze it, give a presentation, and write up their research report in APA style manuscript.

Though having students design surveys and collect data is time consuming, some instructors thought that it was important for students to learn about the entire research process.

INSTRUCTIONAL THEMES

The next group of themes that emerged from our analysis were related to instruction. These were major concepts that instructors tried to convey through their teaching including understanding data and data-related concepts, understanding technology, evaluating data sources and publications, manipulating and analyzing data, and the ethical concerns related to collecting and using data.

Understanding Data and Data-Related Concepts

Instructors emphasized the importance of students understanding how data collection methods impacted results. In many disciplines, instructors focused on survey creation and how different types of questions produced different data. Instructors discussed sampling and bias, emphasizing that samples may not be representative or generalizable. For example, a survey of Kansas and Nebraska students on gun control is not generalizable to the whole population, or research studies on tobacco causing cancer that were funded by tobacco companies might be biased.

Students needed to learn to think ahead about whether the type of data they collected would work with the kind of analysis they wanted to do. Instructors wanted students to understand that analysis was not separate from the research process. "It seems like students will sometimes separate those two, that the research process is this and then analyzing data is something else, but they don't necessarily always put them together as one big cohesive cycle." One instructor talked about encouraging students to think about data in different ways; e.g. data that seems qualitative, like interview transcripts, may be analyzed in quantitative ways, such as using word counts. Some discussed the fact that students didn't always have the skills to do the kind of project they wanted to do:

> On the research methods front, you might have some people who are really interested in asking a question that is best answered through quantitative methods, but if they've only taken a course in qualitative, they don't necessarily have the skill set to put together a survey and then carry it out. And so what I find is they force fit a method. They lead with a method instead of a question.

Understanding Software

Although instructors teach students how to use software programs, pedagogically they believe that understanding the underlying principles for working with data is more important than which specific technologies students learn. Without foundational knowledge in how to work with data, technical capabilities are difficult to develop and limited in utility:

> It's going to make them better at whatever it is that they're studying if they really have a good comprehensive understanding, not just of how to run the software, but understanding what the software is doing when you click a button.

While students might use different software programs in different settings, instructors believe that many of the skills they teach will generalize across programs.

> We tell them other software packages for analyzing data aren't that different from this one. Once you've learned what buttons to click to analyze your data, it's easier to learn the next. You might not be using this one in grad school or your job, but it should give you confidence to be able to go out and figure out how to use a new software package.

Though understanding and using software is an important skill for working with data, instructors emphasized that software is a tool that students use to learn and apply conceptual knowledge.

> The software is one mechanism for students to learn concepts. How to perform analysis, how to present results, it's a communication tool. We're presenting things graphically, which is one mode of knowledge communication and there's rules associated with that and best practices.

Evaluating Data Sources and Publications

Most of the instructors explicitly taught students to evaluate where data comes from. They discussed source bias, the need to compare different data sources, how selected questions can lead to different results, how different analysis can lead to different results, and how graphic representations can misrepresent data. For example, one instructor showed different ways to represent crime on a map. Another instructor had students brainstorm all the different ways that you could measure a concept, and then discussed how the different questions could yield different results:

> I want them to be self-aware. I want them to be systems aware, and I want them to be able to translate research into practice. And ideally, vice versa, practice into research, so that they're asking good questions and not just being resigned to a position of being told what to think, but having the capacity to think for themselves, to ask good questions to figure out ways to answer those questions.

Many of the instructors talked about "fake news" and the need for students to learn to analyze data encountered outside the classroom. They discussed strategies like looking at whether news reports cited the original research and whether it was being interpreted correctly.

> I think the era of COVID has really shone a light on the perils of people not being scientifically literate and not understanding how the human factor influences our interpretation of data and

> the decisions that we make on data. In fact, it was just a few weeks ago that I got an email from a former student who was expressing appreciation for taking the research methods course. She said, "I didn't really see the applicability of it at the time. But now I know how to read scientific journals and understand significant differences. And I feel like I can make a more informed decision." And that's really what I want, you know, they don't need to understand why we set the p value at .05. They really need to know what decisions will that lead to later on, and how can they use this information to set themselves up for success, and to be a good citizen in their workplaces and in their communities.

Manipulating and Analyzing Data

Instructors taught statistical analysis at a variety of different levels. In order for students to progress successfully, instructors expressed the need to begin with simple concepts (e.g. independent vs. dependent variables). Most classes covered basic analysis (e.g. frequency tables, standard deviation, cross tabs, chi square, gamma, lambda, ANOVA, reliability, correlations, t-tests, regression.) A social network analysis class did descriptive statistics like visualizing networks and centrality measures.

Many of the classes used Excel or Google Sheets to analyze data. Instructors noted that students used spreadsheets to store data, organize data, create data visualizations, manipulate data, clean data, pre-process data, perform calculations, and analyze their data. Several classes used SPSS. One instructor just demonstrated it but did not have the students use it. Another instructor met with students individually to walk them through using SPSS on their data. Others used programs like R or Jamovi. The GIS instructors interviewed taught more detailed processes, including geocoding and linking with vector data.

Some instructors talked about the need to help students choose an appropriate analysis method. They were also mindful of directing students to projects for which they had the appropriate background and skills. If they wanted to do a project that they didn't have the background for, they would be steered to analysis methods that fit better with their background knowledge.

Ethics

Understanding the ethics of conducting research was a recurring theme in the interviews. Psychological sciences placed particular emphasis on the ethics of research, examining past research studies that would now be considered to be unethical. Several instructors who had students collect data required them to go through the university's IRB process (which is abbreviated for class projects). One instructor noted that there were additional ethical challenges for students who wanted to do research on K-12 teaching because it involved collecting data about minors, which created additional requirements for IRB protocols. Due to the difficulties of supervising large classes of undergraduates as they go through the IRB process, some instructors avoided having students collect data, opting instead for publicly available datasets because anonymity concerns had already been addressed.

Though some instructors didn't go beyond discussing IRB requirements, for others, ethical concerns were a major theme in their instruction. Instructors discussed the importance of protecting the anonymity of research participants, particularly for research that dealt with sensitive topics such as drug use. Several of the instructors discussed ethics beyond the classroom, examining data that is collected about people every day. Some classes discussed the ethics of using data even if it's publicly available. Referring to publicly available data about murder

victims, a journalism instructor said, "As always happens with journalists we have to weigh between the privacy of the victims and their families and the public's right to know about something that has a huge impact on society." One communication studies class analyzed social media data and discussed the ethics of using data that, while publicly available, did not come with the people's informed consent to be used in research. Instructors also emphasized the need to cite data sources and be transparent about data sources.

In the interviews, only two of the instructors explicitly addressed ethics in analysis. One said:

> There's a lot of things that people can do to find significant results that are considered unethical, like you don't find a statistically significant effect after you planned to collect so much data so you just keep collecting data until the p value drops below .05, so that's considered unethical. A bunch of things called p hacking.

In psychological sciences, they spent time discussing the replication crisis, which is the difficulty of reproducing study results in subsequent studies.

PEDAGOGICAL GOALS

A third group of themes that emerged from the analysis was the pedagogical goals of the instructors. Beyond teaching specific topics, instructors had many reasons for their pedagogical choices, including helping students to develop a self-concept of being competent users of data, helping students develop skills for the workplace, and developing informed citizens. A few instructors mentioned preparing students for graduate school, but it was not a major emphasis for most of them. Those who did primarily mentioned it in the context of specific software that students would need to know how to use when they went to graduate school.

Self-Concept

Instructors discussed "math phobia" or students' resistance to anything requiring quantitative processing. Although interview questions did not specifically ask about students' anxieties about math and statistics, it came up repeatedly. Many participants made comments like "whenever I asked students about their feelings towards statistics and numbers and things like that, almost all of them reference a painful experience they had in a math class growing up" and "I think they hear data analysis and their hearts start beating and they think back to math class or something."

Instructors expressed a pedagogical goal of helping students to see themselves as competent users of data. They wanted to change students' mindsets about their abilities:

> To me it's less about their abilities to do mathematical problem solving, and more to do with the identity work and the attitudinal work to go, "oh, I can make sense of this. I can conceptualize a problem using math and then solve that problem."

To help students develop their skills and confidence, instructors started with simple concepts and slowly built on those.

> The first week of class, we were working with the program. And then each class period, there would always be something that we would be doing in the software so that way they just got more comfortable in class, kind of like a driver's ed type mentality, I'm there driving with them.

They also tried to give students opportunities to learn and make mistakes in a safe environment. One instructor said,"I tried to build in pedagogical opportunities for students to fail, but doing so in kind of a laboratory, a safe environment where it becomes less of a struggle."

Job Skills

Instructors had a pedagogical goal of trying to prepare students for the future, both specifically to find jobs in their fields and more generally. "I teach them not only how to find good information from reliable sources. I also hope I teach them how to interpret data and then how to put those interpretations into practice." Some instructors emphasized the importance of data literacy to students' future careers in response to students' saying that they "would never use this in life." For example, one instructor relayed a story of a student who got a job based on their capstone class project analyzing social media data. Preparation for a specific field could involve having students do assignments that were relevant to jobs they might get in the future. One instructor said,"I just told them this is a job skills class. We do have an exam, but it's a fairly easy exam. The focus is 100% on job skills, the things the employers say they like to see."

Some instructors emphasized the unique ways data was used in their disciplines and had a goal of helping students understand how data was field-specific. A GIS instructor said,

> I don't think we can have the same research topics as in epidemiology, for example, or sociologists. I always tell my students that we have the same topic, but we don't ask the same questions. As soon as you add your "where" in your research question, then it's geography.

A journalism instructor commented, "We wanted to have a generation of programmer journalists, journalists who can write, who can code, and who can tell a story with numbers. My peers and I are struggling to convince students this is the future of journalism." In Agricultural Education, students were preparing to be K-12 teachers. The instructor wanted to prepare them to collect data about their own teaching and their students' needs so that they would be able to analytically "self-reflect on their own teaching" and identify ways they could change their teaching practices to improve student learning outcomes.

In other fields, the job preparation was focused on a wider set of skills. Psychological sciences focused on job preparation, whether students pursue a research-based career or other avenues.

> Our program is very research focused, but we also understand that not every student goes on to graduate school in a research program. Even for students who want to do more as a practitioner, I emphasize that you're still going to be exposed to research and you're still going to be basing your practice off of valid information derived from the scientific method. So understanding the strengths and limitations of different types of research is really going to help you out.

Communication studies also focused on preparing students for potential work in fields like human resources, consulting, and public relations, emphasizing the value of being able to analyze data and communicate the results effectively.

The importance of students being able to use their data skills to communicate was expressed by several of the instructors. “One of the things that always comes up is if you're presenting to a CEO, you have two seconds. You have to be clear, you just have one point and make sure it's clear because of those two seconds.” Another instructor commented,

> They have to write things up in a report form or a document like they would in an actual job. I talked with people that work in the private sector, work in the government sector and found out the skills that they're after and then I incorporated all those things into the class.

Many instructors discussed data literacy as a fundamental skill. In GIS, they more specifically emphasize spatial literacy.

> We have talked for probably 15 years on campus, trying to convince people who can make decisions that spatial literacy is a key, key skill that’s that we need to emphasize more here at K-State to the point where it should be required in the core classes. And taking a GIS class, even if you don’t plan on doing any GIS work, is a great way to really understand how space impacts almost everything that we do.

Being Informed Citizens

Instructors hoped that teaching students to analyze and interpret data would prepare them to make informed, positive contributions to society.

> I'm sensitive to when I see people say “oh, you know, you know the science says this” and I'm thinking, are you sure you're looking at the data correctly? And are you sure that you're not bringing your own biases into that interpretation? And of course we all do that. So I want my students to be aware and able to ask those questions of themselves and of others, so that they can have good civic dialogue that doesn't devolve into the nonsense that we're seeing right now. I think we help increase people's empathy and ability to engage in perspective taking, which would make us better humans overall.

One instructor expressed concerns about seeing decision-makers at the university and in larger society making “decisions based on questionable data.” Several said that they hoped that what students learned in their class would help them shape a better future.

> From a cultural perspective, I'm seeing the shift toward this notion that everyone has their own truths. And I'm going to go seek out that information that tells me what I want to hear. So you can have your data. I'm going to go have my data and I'm right and you're wrong. And so our inability to ask good questions of data or be open minded, at least from a cultural perspective, we're just going to continue to go down this pathway.

PEDAGOGICAL SUPPORT

The fourth theme that we discovered related to pedagogical support. This fell into two main areas: teaching support in their classrooms and learning to teach with data.

Teaching Support

Graduate teaching assistants (GTAs) were the primary source of support for instructors teaching with data. In sociology and psychology, GTAs taught the research methods course labs where students practiced collecting and analyzing data. In these programs GTAs might play a significant role in developing exercises and content for labs. In the GIS program, GTAs assisted instructors in classes where students were working with data. Instructors in journalism and communication studies also employed GTAs. Even though instructors may have wanted GTA support in additional courses, their department wasn't always able to provide the financial support for GTAs.

Learning to Teach With Data

Most instructors hadn't had any training in teaching with data outside of their graduate programs. After graduate school, instructors learned about teaching with data on their own and from others in their department who taught the same courses. A few had attended conference sessions on topics related to teaching with data, such as teaching sampling and teaching research methods classes. One instructor said their department sent them to a week-long workshop on teaching with data. This instructor tried to keep current with industry practice so that they could teach relevant skills to their students. Conferences were seen as valuable for learning how practitioners in their discipline were using data.

Instructors felt that they could benefit from additional training on teaching with data. Some were interested in additional training with technology and techniques for working with data while others were more interested in developing their general pedagogical skills. Connecting with students who are learning about data can be challenging, so instructors wanted to learn ways to better engage undergraduates with data and statistical analysis. To increase engagement, some instructors tried to find examples and ways of presenting information that would be interesting to students:

> When I go to look for resources, usually the goal is how can I make this interesting to students? I know that maybe 10% of my students really want to do research and psychology and their career. So how can I make the content intrinsically engaging? So what examples are fun to use?

Some professors thought learning more about tools and techniques for working with data would be beneficial for their teaching. Instructors said they could benefit from training on topics such as the variety of current tools available for working with data, creating data visualizations, accessing data, using tools such as SPSS and Excel, publishing with data, citing data, complying with IRB requirements, understanding ethical and legal issues associated with publishing data, and keeping current with technological trends in their field like cloud computing and programming.

TECHNOLOGY

Another theme revealed in interviews related to technology. Using technology is integral to teaching with data but presented challenges for instructors and students. Participants discussed lack of adequate computer lab space, issues with students' personal equipment, and disparities in students' access to appropriate technology, particularly when classes were taught virtually. Many instructors had to adapt their teaching practices to compensate for these limitations with teaching spaces and technology.

Spaces and Resources

Not having access to computer labs for teaching with data was a challenge for instructors. In some cases, labs don't have enough space to accommodate larger classes.

> Our enrollments sometimes exceed the capacity of our computer teaching laboratory. The only way that we can really offer those courses is as a traditional lecture/lab. Our preference in terms of instructional style is to teach in the studio. We have found that it's a more effective learning environment for the students.

In other cases, departments stopped maintaining computer labs due to the expense. "It's more efficient to have students bring their own laptops to the lab and the lab will just be a space for students to borrow. The cost of running the lab rooms is something that can cut into our budget." Because technology is essential for teaching with data, instructors without access to appropriate lab space had to adjust their teaching practices by changing from studio to lecture/lab classes, relying on student technology, and finding software that students can use on their personal machines.

With the exception of GIS, instructors had to rely on students using their personal devices in the classroom, which created its own set of challenges. One instructor found that students' personal laptops weren't always powerful enough to run software that had to be installed locally. As a result, this instructor tried to use free, cloud-based programs when possible. Another challenge was troubleshooting issues with students' personal devices when installing and running software. "It's kind of messy to have students installing their own software. Anything could happen...it's a learning process. They're just complete beginners and so you could literally run into anything." This instructor circumvented some of the issues related to using personal laptops in class by running a free statistical analysis program on a department server, negating the need for local installation on student machines.

Instructors also had to navigate student access to subscription-based software, a challenge that was exacerbated by the Covid-19 pandemic. During the pandemic, many classes were offered remotely, and some students could no longer access subscription-based statistical analysis tools like SPSS in campus computer labs. One instructor adapted by using Excel, a software program that all K-State students can access on their personal machines. This instructor found that for the purposes of introducing undergraduate students to statistical analysis, Excel was an adequate alternative to SPSS. In most cases, instructors felt that the particular software that they used for instruction wasn't as important as the foundational concepts they could teach using that software. Though their teaching required software for specific purposes—statistical analysis or data visualization—instructors felt students could learn concepts using one program that could be transferred to other programs. An in-depth examination of how issues with technology impacted student success when learning with data was beyond the scope of this study. However, it

was clear that instructors and students face significant challenges with technology, and that instructors had to invest time and effort into adapting teaching strategies to compensate for issues with technology.

Equitable Access to Technology

Participant's struggles with technology in the classroom highlighted inequities in technology access at K-State. Instructors felt that students didn't have equal access to the technology that they needed to use in their courses. Citing "access to technology that enables them to be successful" as a major challenge for students, one participant said, "being able to have access to appropriate technology, whether that's computers, whether that's sufficient and stable internet, I think sometimes can be the biggest hurdle that I've seen students struggle with." Participants noted socioeconomic status affected student access to personal computing hardware and software.

> You know, it would be great for the students to get SPSS for free. Get computers for free. I think it's nice to be able to do it at your home or places that are comfortable for you. So people who are under-resourced do not have to go to the lab. There's an inequity there.

OTHER CAMPUS RESOURCES

While a number of other campus resources supported the needs of social sciences instructors and their students, it was interesting to note that instructors interviewed did not refer their students to other units on campus, such as the library, tutoring services, or the undergraduate help lab run by the statistics department.

Though faculty felt that their students struggled with finding data, none of the participants mentioned referring their students to librarians for help finding data. However, some instructors mentioned that they would be interested in referring students to librarians. Reaching out to instructors who teach courses where students might be required to find data for projects is a logical place for librarians at K-State to begin building data services.

Only one of the participants mentioned on-campus support for students in the form of the K-State Department of Statistics Undergraduate Help Lab. Two others mentioned it in passing, adding they did not use it, one going as far as to note they did not know it existed until our interview.

A couple of instructors said that they wouldn't feel comfortable sending students to support resources outside their department unless they had a relationship with the unit. One said, "I would be tentative to send them and when I'm thinking about why, it's because there's not enough of a partnership for me to know what they're doing and for them to know what I'm doing." It's possible that building relationships with instructors who teach with data would increase their confidence in the ability of the library to support their students.

One campus resource was mentioned by instructors as being helpful with their teaching in general, although not specifically related to data. The campus Teaching and Learning Center (TLC), which created resources like "Remote Teaching Fails" and "Teachers Thriving Remotely," was an important resource for instructors adapting to teaching remotely during the pandemic. K-State's TLC offers consultations, curricular support, peer networks, and programming to anyone who teaches at K-State. Because instructors seemed familiar with and appreciative of the services offered by TLC, partnering with them could provide opportunities for the Libraries to create programs that would benefit faculty.

The other campus stakeholder mentioned in interviews for support of social sciences instruction was the K-State Libraries. Although instructors didn't refer their students to librarians for help in finding data, the Libraries was acknowledged for its support in areas such as obtaining vital resources like the Environmental System Research institute (ESRI) which hosts ArcGIS, providing computer labs with access to SPSS for students, highlighting material for use in research and instruction through research guides, and securing ebooks free of cost to students through the library. In addition to existing services, instructors also shared ways they hoped to further library partnerships including using the library to provide reference on finding data, hosting workshops on writing literature reviews, and exploring ideas on data storage and a campus data repository for faculty and students.

CONCLUSION

Kansas State University's social sciences coursework extends across a range of departments and degree options. Instructors that taught with quantitative data in undergraduate social sciences classes employed a variety of software, tools, and resources as they worked to meet students at their current level of knowledge and technical competence. This report examined the teaching practices, pedagogical goals, and support needs of undergraduate social sciences instructors at K-State. Our analysis resulted in three key themes, discussed below, which best reflect the issues that posed the greatest challenges for instructors or that demanded greater consideration during instruction. These findings provided a foundation to develop meaningful recommendations for how K-State Libraries can support social sciences instructors in their work.

Access to Technology is Fundamental in Teaching With Data

A core, and not always anticipated, task in many undergraduate courses is building technical competencies with both general data processing software and industry-specific tools. Students entered social sciences courses with varying amounts of skill and comfort with technology. As a result, one of the pedagogical goals for social sciences instructors was building a foundation of technical skills to prepare students for the work required in their class, with the expectation that students would build on those skills in other courses, graduate school, or the workplace.

The cost of technologies required to teach effectively with data complicated instructors' efforts to teach technical competencies and meet learning goals. Students don't have equitable access to pricey proprietary software or computers with adequate processing power for working with data. This inequitable access to technology complicated teaching with data and created extra work for instructors. Many instructors didn't have access to computer labs for teaching, so they developed workarounds. These solutions included changing studio classes to a lecture/lab format, demoing software in class with the professor's computer and then having students access the software outside of class in campus computer labs, and opting to use software, such as Excel, that all K-State students can install on their personal machines.

Rather than maintaining costly computer labs, some departments opted to rely on students' personal technology for instruction. Having students use their personal laptops in class created its own set of challenges. Some students don't have personal computers that are powerful enough to run programs for working with data. To compensate, instructors adopted cloud-based tools that don't have to be installed on student laptops. Using cloud-based tools also circumvented the issues students often encounter when installing programs on personal devices. When relying on student laptops, instructors often opt for software that is available to all students on campus, such as

Excel, or use free, cloud-based tools, such as Google Sheets, rather than pricey proprietary software that might only be available in labs. Though they didn't always anticipate issues with technology, many instructors reported that they had to adapt their teaching strategies based on limited access to technology as well as help students develop basic technological competencies.

Teaching With Data is More Than Teaching Technological Competency

Teaching students to use technology is integral to teaching with data, but instructors emphasized that teaching foundational knowledge, such as how to analyze and interpret data, was more important. When asked about the pedagogical importance of software programs or technologies, faculty reported that understanding concepts for working with data was more important than mastering software programs. For example, one instructor used Excel when students didn't have access to SPSS, saying that either software was adequate when introducing students to statistical analysis. Talking about teaching students to create data visualizations, another instructor said that understanding concepts like geocoding was more important than learning how to use a particular data visualization tool.

In addition to teaching competencies for working with data, instructors encouraged students to see themselves as competent users of data. Many faculty reported that students were anxious or resistant when learning technical or quantitative content. Students were anxious because they had bad experiences with math or didn't feel like they had the capacity to be good at anything involving numbers. In addition to being anxious, students could be resistant to learning about data because they didn't feel like they would use data skills later on. Instructors tried to share examples about how understanding data could help students make informed decisions at home or in the workplace. However, they felt that they struggled to convey the importance of data, even when employers in their field wanted to hire graduates with quantitative skills.

Teaching With Data Means Different Things in Different Disciplines

Some of the concepts and tools that instructors taught are common to the social sciences, but there are disciplinary differences in what instructors emphasize and why. Basic knowledge of statistical analysis and Excel or Google Sheets was important for most disciplines. The importance of understanding the entire process of collecting, analyzing, and presenting data was also common across disciplines.

There were variations in approaches to teaching with data based on disciplinary focus, employer demand for skills, and technological requirements. For example, instructors in psychological sciences and sociology focused on ways to collect and analyze data from human participants. In economics, students learned how to develop forecasts using economic and financial time series data. Instructors in agricultural education, economics, and journalism geared course content toward skills that would help students succeed in the workplace. For example, journalists need to learn how to access, clean, and present data from existing sources. Knowing that students would need these skills in the workplace, journalism instructors taught students how to find and request data from government sources, how to scrape data from the web, and how to use data visualization tools. In agricultural education, the focus was in skills that teachers would use in the workplace, such as collecting and analyzing assessment data.

GIS differed from the other fields in this study in the depth of undergraduate work with data as well as the resources required for teaching with data. Undergraduate students in this program took a series of courses where they developed increasingly advanced knowledge and skills. While instructors in other programs managed to find free (often cloud-

based) tools for teaching when they didn't have access to computer labs or subscription tools, the GIS program needed to maintain computing infrastructure, such as servers for storing and processing geospatial data. GIS was also the only undergraduate social sciences program using proprietary data, which was available through the ArcGIS portal.

RECOMMENDATIONS

Our study highlighted the needs of instructors teaching with data at K-State. Based on what we discovered in the interviews, we developed recommendations for four areas where the Libraries could have an impact: 1) developing formal data services, 2) creating and sharing teaching materials, 3) providing data sources and tools, and 4) finding, accessing, and working with data.

Developing Formal Data Services at K-State Libraries

When we began this study, K-State Libraries didn't have a dedicated data services librarian or team, but our findings suggested that it would be worthwhile to develop capacity in this area. An Ithaka S+R study benchmarking research data services in higher education found that libraries were important providers of research data services, and "it has become 'normal' for an R1 library to have at least one dedicated data librarian and to offer at least one research data service." Given the centrality of data to the social sciences, K-State Libraries should be exploring strategies for expanding data services in a way that is sustainable for our institution. Conversations about future directions for data services should focus on what expertise we need to develop in the Libraries as well as the potential for partnerships with other units on campus.

In the short term:

- **Have discussions with library administration about how data services fit into the Libraries' and University's strategic plans**. Develop a roadmap for providing data services, including how much money and staff time the Libraries is willing to devote to supporting it.
- **Develop our capacity for providing reference services related to data.** Librarians could benefit from training on data sources, types of data, and identifying quality data that meets patron needs.
- **Provide professional development opportunities that encourage librarians to build knowledge about data analysis techniques and tools.** Being conversant in programming languages and tools (especially free software programs) for working with data could help us start conversations with faculty and students about data.

In the long-term:

- **Prioritize hiring librarians with data service skills**. Hire and maintain staff that are capable of 1) consulting with patrons about data management best practices, services, and resources, 2) advocating for requests, issues, concerns, collections and access to services and resources that support data services at the K-State Libraries, 3) providing workshops and coordinating learning opportunities for building data literacy and skills at K-State.

Creating and Sharing Teaching Materials

K-State Libraries is supporting some open/alternative learning materials initiatives, and this is an area where we could provide additional support to faculty teaching with data. Social sciences instructors are already creating their own teaching materials—videos, exercises, and data sets—and there is growing interest in open learning materials. For example, the GIS department is considering developing an open textbook and lab manual for a series of online courses. K-State Libraries has been growing our capacity to support faculty creating and sharing open educational materials and has the potential to become an indispensable partner for campus-wide enterprises in this area.

- **Position K-State Libraries as an essential partner for creating and distributing open learning materials.** Our university is encouraging faculty to create open learning materials and the library already provides PressBooks software and a digital platform for publishing open materials. There is potential to grow services if interest in open learning materials increases.
- **Explore initiatives that incentivize instructors to share teaching materials.** Current library initiatives support the use and creation of materials that are no cost for students but don't prioritize sharing those materials. None of the instructors in this study shared their instructional materials through repositories or other platforms. Initiatives that incentivize sharing might be a way to get faculty to publish materials—textbooks as well as exercises and accompanying data—in repositories where faculty all over the world can access them. K-State University's Open Alternative Textbook Initiative provides grants up to $5,000 to faculty who create open material. The library can encourage publishing resources with the institution's digital press, New Prairie Press, and explore adding open educational resources created by K-State instructors to its existing open digital repository.
- **Promote library-purchased materials that are available at no cost to students.** Ensure faculty know how to access ebooks and other materials that have been purchased by the Libraries and know how to integrate these materials into their Canvas courses using permalinks.
- **Explore options for sharing student projects that include data.** Recognizing we do not have the storage capacity to host large data sets, the library can support the open access sharing of reports based on K-State data in our institutional repository, and explore repository matching services to house the data sets externally.

Providing Data Sources and Tools

Though instructors found free data sources that worked for their pedagogical needs, faculty and students could benefit from data resources and tools available through the Libraries. Findings from this study suggest that we need to prioritize promoting existing resources to faculty and students.

- **Increase awareness of data sources and tools available through the Libraries.** The Libraries pays for a membership to ICPSR (a repository of social sciences data) though none of the participants mentioned using it. All students have access to this data, and ICPSR has developed teaching materials for instructors teaching undergraduates with data. The Libraries also pays for a site license for ESRI (Environmental Systems Research Institute) data, which is used in undergraduate courses.

» **Prioritize acquiring and maintaining resources for teaching with data.** Budget cuts have diminished K-State Libraries' capacity to provide databases, ebooks, and journals for our community. We want to encourage more faculty to use ebooks and databases in their courses, but it's becoming increasingly difficult to ensure that we maintain subscriptions from year to year. When making difficult collection development decisions, the growing importance of data to our faculty and students should be considered.

Finding, Accessing, and Working With Data

Instructors in the social sciences teach undergraduates how to collect, find, analyze, interpret, and present data. Some of the teaching activities that instructors described—such as directing students to trusted data sources—is already within the domain of librarians' work and expertise. K-State Libraries can certainly expand our capacity to support students who need to find data for projects. We should also be having conversations about how or if we should provide support for the range of activities associated with teaching with data.

» **Leverage K-State Libraries' partnership with the campus innovation lab to offer training and software access.** Instructors are using a variety of free software tools with their students, and the lab could be a venue for instructors to share their knowledge and experiences teaching with these tools. The lab could also provide introductory training in areas that instructors encourage students to learn outside of class, such as learning programming languages such as Python and R.

» **Provide more support to students searching for data.** Students struggle with effective search skills as well as identifying data with characteristics that meet their needs. The library can be more proactive in identifying and offering support to courses with assignments that require data discovery. Instructors felt that students struggled to find appropriate data sources for projects and K-State Libraries could capitalize on that need as a way to build relationships with social sciences faculty.

» **Advocate for practices that improve the discoverability of data.** This could be emphasizing the importance of metadata, documentation, or proper data citation.

» **Increase awareness about using data in a legal and ethical manner.** With the exception of journalism, instructors didn't address issues such as licensing or laws that might apply to accessing, using, and sharing data. The library has a potential role to play in educating undergraduates and instructors on how they can use data.

APPENDICES

Appendix A: Departments That Social Sciences Librarians Support at K-State

Departments			
Economics	Education	Personal Financial Planning	Leadership Studies
Agricultural Economics	Hospitality Management	Social Work	Psychology
Anthropology	Journalism and Mass Communications	Sociology	Political Science
American Ethnic Studies	Gender, Women, and Sexuality Studies	Communication Studies	Communications and Agricultural Education
Security Studies	Leadership Studies/Leadership Communication	Geography and Geospatial Sciences	

Appendix B: Semi-Structured Interview Guide

I'm going to start the recording now

Thank you for agreeing to speak with me today about teaching undergraduates with quantitative data in the Social Sciences. As we discussed, we will be recording the audio of this interview, but information will only be reported anonymously. Could you please confirm that you consent to this interview and recording the audio?

Note regarding COVID-19 disruption I want to start by acknowledging that teaching and learning has been significantly disrupted in the past year due to the coronavirus pandemic. For any of the questions I'm about to ask, please feel free to answer with reference to your normal teaching practices, your teaching practices as adapted for the crisis situation, or both.

Background

Briefly describe your experience teaching undergraduates.

- How does your teaching relate to your current or past research?
- In which of the courses that you teach do students work with data?

Getting Data

In your course(s), do your students collect or generate datasets, search for and select pre-existing datasets to work with, or work with datasets that you provide to them?

If students collect or generate datasets themselves Describe the process students go through to collect or generate datasets in your course(s).

» Do you face any challenges relating to students' abilities to find or create datasets?

If students search for pre-existing datasets themselves Describe the process students go through to locate and select datasets.

» Do you provide instruction to students in how to find and/or select appropriate datasets to work with?

» Do you face any challenges relating to students' abilities to find and/or select appropriate datasets?

If students work with datasets the instructor provides Describe the process students go through to access the datasets you provide. Examples: link through LMS, instructions for downloading from database

» How do you find and obtain datasets to use in teaching?

» Do you face any challenges in finding or obtaining datasets for teaching?

Working with Data

How do students manipulate, analyze, or interpret data in your course(s)?

» What tools or software do your students use? *Examples: Excel, online platforms, analysis/visualization/statistics software*

» What prior knowledge of tools or software do you expect students to enter your class with, and what do you teach them explicitly?

» To what extent are the tools or software students use to work with data pedagogically important?

» Do you face any challenges relating to students' abilities to work with data?

How do the ways in which you teach with data relate to goals for student learning in your discipline?

» Do you teach your students to think critically about the sources and uses of data they encounter in everyday life?

» Do you teach your students specific data skills that will prepare them for future careers?

» Have you observed any policies or cultural changes at your institution that influence the ways in which you teach with data?

Do instructors in your field face any ethical challenges in teaching with data?

» To what extent are these challenges pedagogically important to you?

Training and Support

In your course(s), does anyone other than you provide instruction or support for your students in obtaining or working with data? *Examples: co-instructor, librarian, teaching assistant, drop-in sessions*

- How does their instruction or support relate to the rest of the course?
- Do you communicate with them about the instruction or support they are providing? If so, how?

To your knowledge, are there any ways in which your students are learning to work with data outside their formal coursework? *Examples: online tutorials, internships, peers*

- Do you expect or encourage this kind of extracurricular learning? Why or why not?

Have you received training in teaching with data other than your graduate degree? *Examples: workshops, technical support, help from peers*

- What factors have influenced your decision to receive/not to receive training or assistance?
- Do you use any datasets, assignment plans, syllabi, or other instructional resources that you received from others? Do you make your own resources available to others?

Considering evolving trends in your field, what types of training or assistance would be most beneficial to instructors in teaching with data?

Wrapping Up

Is there anything else from your experiences or perspectives as an instructor, or on the topic of teaching with data more broadly, that I should know?