

# Florida State University Libraries

---

2021

## Teaching with Data in the Social Sciences: An Ithaka S+R Local Report

Jesse Klein, Jeff Philips and Mallery Rawls



## Teaching with Data in the Social Sciences: An Ithaka S+R Local Report

Jesse Klein, Ph.D.  
Social Sciences Research & Data Librarian  
Florida State University Libraries

Jeff Phillips  
Social Sciences Librarian  
Florida State University Libraries

Mallary Rawls  
Humanities Librarian  
Florida State University Libraries

### **Introduction**

Initiated in early 2020, Florida State University Libraries (FSU Libraries) joined 19 peer institutions across the United States in Ithaka S+R's "Teaching with Data in the Social Sciences" research study. An interdisciplinary team of three librarians recruited and interviewed faculty with experience teaching undergraduates using data and concepts of data literacy. The goal of the local project was to gather qualitative data on attitudes, behaviors, and needs for undergraduate instructors in the social sciences who teach with quantitative data and to use that information to inform library services and practices. This local report and supporting data, along with the results from other peer institutions, will constitute a larger report by Ithaka S+R, a research organization whose goal is to "provide research and strategic guidance to help the academic and cultural communities serve the public good and navigate economic, technological, and demographic change." The results of this local report provide FSU Libraries guidance for how to better support instructors teaching with quantitative data in the social sciences. Additionally, the research has implications for curriculum development, career planning, and skills development.

### **Institutional Background**

Founded in 1851, Florida State University (FSU) is a comprehensive, national, research university that provides extraordinary opportunities for undergraduate students

to build a strong foundation in their chosen fields. FSU has a thriving population of 41,481 students (2019-2020), continuing a consistent history of enrolling over 41,000 students since 2011, and of those, 79% (or 32,930) are undergraduate students. FSU Libraries is at the heart of academic life for its campus community, for both on and off-campus students and faculty. Subject and functional specialists from the libraries serve faculty and students from departments all over campus and in a range of ways starting with the discovery of resources to data analysis and reporting. When working with users to determine their information needs, staff at FSU Libraries collaborate to face challenges with curiosity and help users overcome hurdles to learning important information skills. FSU Libraries takes its responsibilities as information stewards seriously by transforming and curating resources to build long-lasting relationships with the campus community.

Data generated from academic research continues to grow, making research data services an important area of growth for academic research libraries. FSU Libraries are becoming a leader in providing these services in a sustainable and innovative way. FSU Libraries' research data expertise is crucial as FSU develops a unified institutional strategy that addresses the growing need for research data services. The libraries have developed a range of research data services that include research data management consultations and training, training in data processing and analysis, data curation throughout the research lifecycle, and preservation and access to data collections. While faculty data services oriented towards data management and curation are important areas of growth for academic research libraries, the need to provide support for students' data and statistical literacy increases as instructors build data into undergraduate courses, undergraduate research programs, and individual study. This report highlights the approach of ten undergraduate instructors in the social sciences and how their use of data poses unique challenges and opportunities for undergraduate education. The report concludes with recommendations for how FSU Libraries can integrate these findings into their strategy for assisting both instructors and students in social sciences courses.

## **Method**

As one of several institutions participating in this project, the standardized methodological protocols (see appendices) facilitate comparability between institutional findings and the development of the more comprehensive capstone report Ithaca S+R will generate. Using the course catalog and existing institutional knowledge of popular undergraduate courses that use data, the research team gathered a list of possible participants from across the social sciences. We reviewed syllabi and academic program maps to further refine the list to courses most likely to use data, such as courses on research methods, social statistics, and special topics courses with data in their title. This strategy generated a list of 20 possible participants, from which we successfully recruited 10 participants. Due to the COVID-19 pandemic, recruitment and interviewing took more time than anticipated and all interviews were conducted over Zoom or the phone.

## **Participants**

Between January and March 2021, we conducted semi-structured interviews with 10 faculty members from the departments of Economics, Sociology, Educational Psychology, Information, Geography, Anthropology, Psychology, Public Safety & Security, Applied Public Policy, and Social Work. Some of our participants had a secondary affiliation with the departments of Demography, Archaeology, and Criminology. All participants have used, or were using at the time of the interview, quantitative data in their undergraduate courses and were enthusiastic to discuss the pedagogical implications of using data for undergraduate education.

When asked about the courses they teach where students work with data, some of the courses described were:

- Data Analysis
- Computational Finance
- Econometrics
- Financial Mathematics
- Population and Society
- Climate Change in Society
- Population Data
- Introduction to Quantitative

- |   |   |
|---|---|
| <p>Methods</p> <ul style="list-style-type: none"> <li>● Statistics</li> <li>● Health Informatics</li> <li>● Data Mining and Analytics</li> <li>● Medical Geography</li> <li>● Underwater Archaeology</li> <li>● Introduction to Anthropology</li> </ul> | <ul style="list-style-type: none"> <li>● Research Methods in Psychology</li> <li>● Applied Research and Data Analysis</li> <li>● Research Methods for Studying Housing, Land, and Cities</li> <li>● Social Work Statistics</li> <li>● Introduction to Research</li> </ul> |
|---|---|

Participants described how their teaching for introductory-level methods and statistics courses aligned with disciplinary expectations, whereas their upper-level analysis courses (and certainly graduate-level courses) related more closely to their current and past research. Their differences in rank and specialization appeared to influence the likelihood that their current or past research would affect the use and type of data in their undergraduate courses. For example, a specialized teaching faculty member in Psychology noted, “I’m here to teach, not to research. So, it’s kind of a unique position to be in.” The faculty member then discussed gathering data from textbooks or using generated data that provide a “specific answer.” In contrast, an assistant professor in Sociology and Demography who researches climate change used the same data in their courses that they do in their research—a combination of Census data and National Oceanic and Atmospheric Administration (NOAA) data. They explained, “Well, I’m a demographer who studies climate change.”

Some faculty want to assign data related to their research, but growing class sizes and limited data choices in pre-packaged course materials force them to adopt what is available. An associate professor in Economics must teach macroeconomics courses to 250 students per section early in the major’s sequenced requirements. The large class size required the professor to rely on course materials developed and deployed by major textbooks. To this, they commented, “So the university wants more students. So, we give them more students, [but] as a consequence, we give them more rote type of material to do it...it’s more mechanical [to] grade.”

## Getting Data

Participants discussed how they found and obtained datasets to use for their instruction and how students (a) collected or generated datasets themselves, (b) searched for pre-existing datasets themselves, and (c) worked with datasets that the instructor provided.

Participants found datasets for their instruction by asking their peers for datasets, pulling data from publications, collecting data from students in their courses, using a dataset generator, and finding data in online government and research databases. Using pre-existing datasets saved time for the students and allowed the instructors to concentrate on teaching the students how to use the data instead of retrieving the data. For example, an assistant professor in Anthropology and Archaeology got data from their friends so the students could focus on the meaning of the data instead of having to spend time on data collection:

Instead, what I did is I got, again, aggregate data. I asked a whole bunch of my friends for datasheets from their own lithic analyses that are and instead of people analyzing the stone tools themselves. They learned how to do the statistical analyses that we do, post, get it collecting all of the metrics, and how to put meaning into that light.

Participants saved time by providing their students with cleaned data from real and fictional datasets. Cleaning the dataset for the students allowed the students to focus on analyzing the data instead of cleaning the data. Additionally, providing students with the data saved the instructor time by reducing the number of questions the instructor received about each unique dataset. A participant from Economics explained:

For undergraduates—I've given up on that, honestly. It just takes too much time. If I give them an individual project and they have to go collect their own data, I have to deal with every single student about their data problems. And it's just impossible. And they're learning that those are hard skills. So, I've given up on

them trying to find their unique data sets.

Although instructors often provided their students with databases, students in upper-level courses who had specific projects were encouraged to use their own data. Students who created research projects would create surveys to gather data or the instructors had their students find pre-existing data sets through open-source websites or with the help of librarians. A professor in Information allowed advanced students to find pre-existing datasets themselves:

But certainly, some students if they are very good at data science and data mining before they actually enter the course. I allow them to pick their own data set to work with. So, they can use other datasets that they feel like, are interesting. For example, one of my previous students showed us the airplane on airflow. Some data set related to airplanes—airplane-related variables that I don't even know what they're about. I have to really learn from this student about data set and how to how to find meaningful information out of the analysis.

## **Teaching with Data**

Participants discussed how they taught with data and the challenges that they faced teaching with data. The instructors had their students analyze data through visual representations and used discipline-specific data collections. A participant in Sociology had students use visual representations of data when analyzing large data sets:

So, they analyze a lot of data, with figures. I'm really big on visualizing information to interpret it. It just massively helps with like an ungodly amount and just makes life so much easier when you've got a large data set and you properly visualize it to actually look at it.

Participants discussed using data to teach their students about learning related to their discipline. An interviewee from Criminology explained, “in some of the undergraduate courses, like crime mapping and analysis, will actually be using law enforcement RMS databases to collect and analyze that data.” The participants varied

in how they used the data. An interviewee in the mathematics field used the data to help students understand the mathematics behind statistics, while an interviewee in Applied Public Policy used statistics to help students write a research paper. A participant from Criminology explained how the statistics related to learning in the discipline:

It's a little more applied data than research data. However, obviously, we're looking at processes of proof for charging someone criminally and proving that they committed a criminal act. There I teach courses also, that involve the collection, validation, and dissemination of intelligence and intelligence products that range from open source data all the way through the gamut of surreptitious collection of intelligence.

The tools that the students used included Gertl, Github, Data camp, Excel, SPSS, R, SAS, and STATA. Instructors would provide additional guidance to students who showed interest in statistics. For example, an associate teaching professor in Psychology said:

Well, I expect them to be able to navigate it with a little help. Right. So, I don't expect them to come in with an understanding of the language for syntax, for example. But if students expressed some interest in it, I like that we can hopefully, help get them a little further along. For my students that are really interested in it, I tell them to, you know, seek out the statistics certificate to like, get certified in R, get certified in SAS, these other programs, and is really rewarding.

### ***Relate the Instruction to the Students' Lives***

The participants discussed relating their instruction to the students' regular lives. The instructors used current events to promote learning and prepare students for their future careers. The instructors wanted to use real-world examples from the news to help students become more critical of statistics they saw broadcasted. Some instructors used data about current events to transfer the learning beyond the curriculum. For example, a participant in Sociology explained:



So, I'm really big on critical analysis of information. I think beyond the contents of the course—being able to critically evaluate and critically think about information that's out in the world is really key. Yeah. So, a lot of the lessons when it comes down to interpreting this kind of information, I think is applicable everywhere else. Or, at least, I try to make it so you know, not every day, you're going to be looking at birth rates, for instance. But COVID just happened, there's tons of misinformation.

Several participants mentioned how they wanted students to consider what the data was saying instead of just analyzing the data. An Economic professor explained:

We just try to get them to think logically about what they're seeing as if they're the consumer of statistical output rather than the producer of statistical output. So honestly, we're just trying to get them ready to be intelligent consumers of data analysis.

A participant in Education presented scenarios to help students critically analyze the data. For example, he had students compare different global salaries:

And there's another exercise where I do, where I compare the salary of Japanese high school teachers to US high school teachers. And notice that the Japanese high school teachers make hundreds of times what the US teachers make. And then let them figure out, wait a minute, Japanese high school teachers are paid in Yen. So okay, maybe we have to think about whether or not these two measures are actually comparable before we compare them. So, you know, there's a lot of that, that you need to pull in that you can do with real examples.

Participants also created objectives that promoted the transfer of knowledge to their students' future careers. A professor in Applied Public Policy stated, "If the students go into policy for their future careers, being able to navigate the local level is a first step towards making an impact towards broader regional, state, and national levels."

An Economics professor wanted their students to be capable of performing data analysis in a work environment. They explained what skills organizations wanted their employees to have:

So, they [employers] typically want students who can do some basic analysis reliably knowing when they're wrong. I really harp on this in my classes, getting the exactly the right answer is less important than knowing you've got a way wrong answer. Right? And so just what reasonably should the answer be. So, employers want somebody who can handle data, know what to expect, and then interpret what they what they got how they did it and write a sensible paragraph about that. Right? It's just language skills. And writing skills is something that I keep getting over and over and over again.

### ***Challenges of Working with Data***

Participants identified the following challenges that students had working with data: learning the tools takes time, students lacked an understanding of statistics, downloading large data files requires excessive bandwidth, querying data takes a lot of time, and the limitations of using Excel for statistics. A professor of information explained the issues with running data queries:

And also query the data could take some time. It's not like Kaggle or UCI, where you can just download the well, clean formatted data in CSV format. If you're given a SQL database, you need to use query to get the data prepared data that may take a long time. But I usually give this option only to very advanced students who are feeling strongly about using this data set for their project.

The instructors also faced ethical considerations when introducing quantitative data into their undergraduate courses. These considerations ranged from the types of data available for use in teaching to the dissemination of class projects using data. An associate professor in Educational Psychology captured these concerns when they said, "looking at what the downstream consequences of what you're doing are not covered well in the texts or in the syllabi of our existing courses. So, it's a hole in our

curriculum that we need to fill.” Additionally, the instructors considered how their instructional materials related to evolutions in the field of diversity, equity, and inclusion. A participant in Social Work discussed how their slides contained outdated terms that did not adhere to modern cultural understanding. For example, the participant explained how their previous course incorrectly restricted gender to a binary classification:

I have to start the beginning of the semester by saying, “look, I have gender as a dichotomous binary variable. I know, I know. But these slides were written like four years ago and so just kind of for the sake of the semester—and making me not change the slides—can we just pretend for a minute?” So yeah, it's stuff like that. Um, so yeah, there are things in there. But then I do say, “hey, look, you know, if I'm teaching differences between groups, and you do look at, you know, sexuality as a predictor, you have to understand that you may have a lot of, you know, straight hetero, cis, whatever, and not a lot of the others, because they're smaller subtypes, just by population.”

## **Training and Support**

The participants discussed the instructional training that they received and the types of support that they offered their students. The training consisted of the training the instructor received outside of their graduate degree. A participant from Archaeology explained how a lot of their colleagues were not skilled in using technology, and they would have graduate students learn new technology for them instead of dedicating time to learn new technology. Two participants received additional training on how to teach with data from workshops and in-department instruction. A professor in the College of Information discussed learning from workshops:

I do not have a formal training in teaching students about data in my graduate degree, but I did receive some—I did attend some workshops at the National Science Foundation and the National Institutes of Health about data science.

The types of support that the participants offered their students were (a) providing the students with outside instruction and support, (b) providing the students

with open data sources, and (c) including data sources outside of the students' formal coursework. Students learned from outside materials such as blogs, library resources, and LinkedIn Learning. A participant from Criminology described how students used LinkedIn Learning:

I encourage and I use a lot of the LinkedIn learning that is provided by FSU here. And students really like that. So, a lot, particularly if they come in, you know, are very scared about basic statistics or really, really terrified over using Excel. There's some unbelievably great beginner level courses out there to assist them and then we do have a center in our own. We have our own tutoring center; we have peer tutoring and professional tutors. So, they can use those resources. But a lot of the students really liked the LinkedIn learning stuff. Yeah. I usually post several tutorials in every course that's appropriate for LinkedIn in case they need more systems.

## **Recommendations**

The results of this research suggest more collaboration between university libraries and undergraduate data literacy instruction. Data literacy programming in our library tends to target graduate students, and we should make additional efforts to facilitate undergraduate data literacy skills. Our participants used quantitative data in their teaching from various sources, many of which were innovative and new to our interviewing team. Their insight offers us new ways to locate and curate data beyond our standard exploration of well-known data repositories and government databases.

The role of pre-packaged course materials in using data in undergraduate courses is a perspective we had not considered. Pre-packaged course materials have broader implications for open educational resources (OER) and our approach to course reserves and materials for undergraduate courses. Similar to how faculty tend to stick to particular textbooks once adopted, participants reused the same datasets every semester. This consistency in the data being used in teaching undergraduate social sciences courses gives librarians a starting point from which to initiate outreach and

skills development in working with these instructors and students moving forward.

Throughout the interview responses, participants shared a general negative sentiment about undergraduate students' data analysis skills and a lack of confidence in their ability to work with various analysis programs such as Excel and SPSS. The participants seemed beleaguered with the necessary "step-by-step" instructions they needed to provide their undergraduate students. This provides FSU Libraries with an opportunity to promote data literacy instruction through targeted workshops and tutorials. Although FSU Libraries currently provides workshops each semester to help enhance data literacy skills, targeting undergraduate students—especially students in data-intensive courses—can elevate the libraries' outreach and engagement.

Another recommendation is for FSU Libraries to help instructors align student data literacy skills with long-term career planning. This could increase the students' interest in data literacy and reduce the overall strain of the instructor throughout the semester. Participants mentioned a lack of time and resources to teach students about working with data. FSU Libraries can act as a connector for students' skills and career goals through (a) workshops that apply data and programs to a range of contexts, and (b) outreach programs such as hack-a-thons and collaborative collisions. Additionally, FSU Libraries could collaborate with other campus departments and organizations (e.g., the Career Center and makerspaces) to provide undergraduate students with data literacy badges and certificates throughout their time at university. Gamifying learning can promote continued student development beyond one course.

Ultimately, this research presents many opportunities for academic libraries to make meaningful contributions to data literacy in undergraduate education. Instructors in the social sciences are innovative with their use of quantitative data in their undergraduate courses but they have limited time to teach their students. Libraries can support data literacy instruction through workshops, tutorials, and data-intensive outreach events. To accomplish this, however, more discussions with stakeholders and a comprehensive communication plan are necessary. Perhaps a pilot within one or two departments where instructors and librarians discuss opportunities for embedded

librarianship and relevant outreach events that would cater to their current students and others in the major would be a good place to start.

## **Conclusion**

FSU Libraries has prioritized various aspects of research data services for years and the results of this project further justify the role of libraries in facilitating data literacy education for undergraduate students. Instructors who use quantitative data to teach undergraduate students in the social sciences are identifying the opportunities and challenges of working with data. The instructors are also teaching valuable skills and programs to prepare the next data-savvy generation. Through traditional library instruction and outreach, librarians can help data literacy instructors locate and use the data they need in their courses. The data literacy instructors' approach to finding data for their courses provides librarians with a unique data curation and collections perspective. This enhances data librarians' exploration of desirable datasets for teaching and learning. The outcome of our research suggests that librarians have a valuable role to play in facilitating data literacy skills among undergraduate students through instruction and data curation.

## Appendix 1: Recruitment Email

### Sample Recruitment Email

*Subject.* Florida State University's study on teaching with data in the social sciences

Dear [*first name of instructor*],

Florida State University Libraries is conducting a study on the practices of social science instructors in order to improve support services for their work. We are interviewing instructors whose undergraduate students engage with quantitative data, such as by conducting research using quantitative methods, analyzing or visualizing datasets, or learning to use specific tools or software to work with data. Would you be willing to participate in a one-hour interview to share your unique experiences and perspective?

Our local Florida State University (FSU) study is part of a suite of parallel studies at 19 other institutions of higher education in the US, coordinated by Ithaka S+R, a not-for-profit research and consulting service. The information gathered at FSU will also be included in a landmark capstone report by Ithaka S+R and will be essential for FSU to further understand how the support needs of social science instructors are evolving more broadly.

If you have any questions about the study, please don't hesitate to reach out. Thank you so much for your consideration.

Sincerely,  
Jesse Klein

Jesse Klein, Ph.D. | Social Sciences Research & Data Librarian  
Florida State University Libraries | 116 Honors Way | Tallahassee, FL 32306  
0027P Robert M. Strozier Library | 850.645.6323 | [jrklein@fsu.edu](mailto:jrklein@fsu.edu)

## Appendix 2: Informed Consent Statement

### Informed Consent Form

*Project title.* Teaching with Data in the Social Sciences

*Reason for the study.* This study seeks to examine social science instructors' practices in teaching undergraduates with data in order to understand the resources and services that instructors at [name of your institution] need to be successful in their work.

*What you will be asked to do.* Your participation in the study involves a 60-minute, audio-recorded interview about teaching practices. Your participation in all or part of this study is completely voluntary. You are free to withdraw consent and discontinue participation in the interview at any time for any reason.

*Benefits and risks.* There are no known risks associated with participating in this study. You may experience benefit in the form of increased insight and awareness into teaching practices and support needs.

*How your confidentiality will be maintained.* Participants will sign informed consent forms, either in person or remotely via email, but these forms will in no way be linked to the collected data because there will be no key that corresponds the participants to their pseudonyms. Informed consent forms will be stored as paper copies in a locked file cabinet only accessible to the investigator(s) and/or as digital files by the investigator(s) in a non-networked folder on a password protected computer. The informed consent forms will be destroyed in 2025 following the completion of the research project

*Questions?* You may contact the researchers at any time if you have additional questions about the study. If you have any questions about this process, please contact Dr. Jesse Klein at jrklein@fsu.edu. To view the Human Subjects application for this study, file complaints, or inquire about rights as a participant, you may contact the Florida State University Institutional Review Board office at 850.644.7900. Refer to the original project title "Teaching with Data in the Social Sciences."

I, \_\_\_\_\_, understand and consent to participate in the study as described above.



Participant Signature: \_\_\_\_\_ Date: \_\_\_\_\_

Interviewer Signature: \_\_\_\_\_ Date: \_\_\_\_\_

## Appendix 3: Interview Guide

# Teaching with Data in the Social Sciences

## Interview Guide

*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?