

Innovating Education in Research Data Services in Academic Libraries: A Qualitative Analysis of Student Experiences in a Collaborative Research Program

Maria A. Barca and Kathryn Diana

abstract: This paper explores the experiences of 11 students who participated in Seton Hall University Libraries' DataLab, a research program led by the Research Data Services (RDS) team. The RDS, composed of librarians and a data specialist, pairs students with faculty-led research projects to conduct research and build data science skills. In semi-structured interviews, the students who participated in this study reflected on the skills they gained through RDS, such as data cleaning, analysis, visualization, and presentation, and how the program influenced their interest in data-related careers. This case study illustrates how academic libraries can foster data education and develop similar initiatives to support student engagement in data-driven research.

Introduction

This qualitative case study explores student experiences with Seton Hall University (SHU) Libraries' Research Data Services (RDS) DataLab program. This program connects faculty members working on data-intensive research projects with students, enabling them to collaborate on these projects while a member of the RDS team oversees and assists with all areas of the data lifecycle. By placing this research program in the library directly, students and faculty work with librarians and other data experts on their projects. Situating these relationships at the nexus of Research Data Services and librarianship provides students insights into key tenets of librarianship, such as information literacy, data literacy, and other data curation and manipulation skills that



are central to data science. As this paper demonstrates, the DataLab trains students and faculty in data science principles, bolstering the DataLab participants' skills while preparing students for their future careers and/or graduate school endeavors in data science or data-focused fields. Importantly, students do not need to have interest in or experience with data science or data-focused fields to participate in the DataLab; instead, this initiative is designed to attract students from many backgrounds, with the goal of sparking interest in data science.

Thanks to a staggering increase in data production, more disciplines outside of STEM rely on data and "have become more data-intensive."¹ Adopting principles of data science can help to manage and navigate this form of information. According to Mark Nahotko, Magdalena Zych, Aneta Januszko-Szakiel, and Małgorzata Jasowska, data science is defined as the ability to use "statistics, data analysis, machine learning, expertise in specific domains, and related methods, in order to comprehend, appreciate, and analyze data."² The principles of data science can be seen as the confluence of domain expertise, computer science, and mathematics and statistics, which—when overlapped—describe data science as a field.³

Speaking directly with students enabled the SHU Libraries' RDS team to understand how they interact with data science within their own disciplines. The team also gleaned outcomes of student engagement in data research as it pertains to students' decisions about future careers or graduate school in data science or data science-adjacent fields.⁴ This article addresses a gap in the literature by exploring how a library's RDS cannot only enhance, but enable, the student research experiences in higher education. As no other academic libraries currently offer services identical to those of SHU's DataLab, this case study provides a unique perspective on how libraries can effectively support students engaged in data-driven research.

Research Data Services at SHU

At full capacity, SHU Libraries' RDS personnel include six full-time librarians, a non-librarian data specialist, and the Libraries' two assistant deans—the Assistant Dean of Public Services and the Assistant Dean of Information Technologies and Collection Services. RDS also employs a graduate assistant (GA) who supports several service areas, including the DataLab.

The members of RDS possess diverse skillsets and that were acquired via both formal and informal training. For example, the data specialist, who serves as a mentor to newer members of the RDS team, attended the Inter-university Consortium for Political and Social Research (ICPSR) Summer Program three times and completed a data certificate program at Seton Hall University. The specialist also gained data science knowledge when completing a PhD program in the SHU College of Education. The librarians and data specialist have collective skills in qualitative and quantitative research and software, including ATLAS.ti, STATA, Statista, PolicyMap, Excel, Qualtrics, R and RStudio, Python, SPSS, ArcGIS (including StoryMaps), Power BI, Tableau, and MATLAB. The librarians learned both the software and analytical and visualization techniques through a variety of methods, including attending formal classes; shadowing RDS consultations; and attending conferences, workshops, and webinars hosted by the companies that produce



the software. Active learning and collaboration continue to guide the RDS team when onboarding new members so that they can effectively support the DataLab.

The SHU DataLab

The DataLab is an SHU Libraries initiative that was founded in Summer 2022 by the Libraries' two assistant deans. The service is an offshoot of Seton Hall University's DiploLab Undergraduate Research Lab, a grant-funded collaborative effort between SHU's School of Diplomacy and International Relations and the University Libraries. In a semester-long program, the DiploLab partners undergraduate and graduate students with faculty from the School of Diplomacy and International Relations as they work on research.⁵

The DataLab, also a grant-funded program, pairs undergraduate and graduate students with faculty members who are working on data-heavy research projects that require assistance. The administrators saw a need to train students with critical data analysis, cleaning, and visualization skills that they would not otherwise gain from most majors across the university. The grant funding, which primarily comes from the university, is used to pay an end-of-semester stipend to participating students, in recognition of their labor.

Members of the Libraries' RDS team support the program by volunteering to assist with each project that the DataLab accepts. Because the DataLab accepts applicants from any academic department, the projects that the RDS members help with vary significantly. For example, RDS and the DataLab have supported quantitative and qualitative research in health sciences, social sciences, environmental studies, and business. Faculty have discovered the program through emails, including a weekly email from the provost that is sent to all faculty, staff, and administrators. Word-of-mouth has also been helpful to recruitment, as previous faculty participants discuss their experiences with interested colleagues. Currently, there is not a larger effort to market the DataLab because of the possibility that the number of interested faculty might overwhelm the relatively small RDS department.

When accepted into the DataLab, student and faculty participants attend a series of classes provided by the RDS team, including a mandatory Research Data Management class focused on best practices for managing research data. Topics covered include strategies for data organization, documentation, and version control, and project-specific data skills. Examples of project-specific skills classes include: Data Management in STATA, which explores how to manage incorrect and inconsistent data and create and compute new variables in STATA; Qualitative Data Analysis in ATLAS.ti, which teaches users how to locate text, code data, and annotate the findings in ATLAS.ti; or Coding Qualitative Research, which teaches users about qualitative research practices, including how to develop and apply codes to analyze qualitative data. Students and faculty can also have one-on-one consultations with the RDS members supporting their project. Each training component aims to provide students and faculty with the necessary support to succeed in their research, while introducing participants, especially students, to major principles of data science, such as data software, analysis, visualization, ethics, and management.

Previous DataLab projects have explored diverse topics, including how face masks impact memory recognition, the influence of accent biases on students' perceptions

of success, and the ways parents teach spatial problem-solving skills to children and whether those skills can be transferred to unique spatial challenges. In Fall 2023, the program supported research on factors influencing self-care among frontline workers during the COVID-19 pandemic, women's pathways into computing majors through analysis of summer internships and post-internship reflections, and the application of data scraping techniques to enhance traditional methods of historical analysis. The diversity of projects supported by the DataLab can be replicated by RDS units across colleges and universities to aid students and faculty in their research.

Research Background

In an early study by Terrence B. Bennett and Shawn Nicholson, the authors argue that librarians cannot just provide data reference services but must know "how the user intends to statistically manipulate the data retrieved."⁶ This is not easy to do, however as Andrew M. Cox and Stephen Pinfield note, technical skills deficits are a major hurdle for libraries that want to provide robust RDS support.⁷ To support this notion, Cox and Pinfield administered a questionnaire to library staff in the UK and found that only 20 percent of institutions responding were providing technical—statistical or data-focused—services such as data analysis.⁸ As noted by Qiong Xu, informational or consultative services include "data finding and use support," while technical services focus on "data analytics support . . . data organization, manipulation, analysis, interpretation, and presentation skills."⁹

An abundance of research in the 2010s revealed that library RDS were primarily providing informational or consultative services to users, with fewer providing technical or hands-on services.¹⁰ Inna Kouper et al. performed a content analysis in 2017 of 124 Association of Research Libraries (ARL) library websites and conducted interviews with library administrators or program leads to identify the rates at which libraries were providing "advanced" or technical services versus "basic" or consultative services, as conceived of in their maturity taxonomy.¹¹ Most of the libraries responding provided basic services, whereas fewer libraries provided advanced, or mature, services.

Marek Nahotko et al. used Kouper et al.'s taxonomy in 2023 to assess the maturity of RDS in 320 Polish academic libraries by performing a website content analysis and found that maturity was at "good or average level[s]," which is in line with the current literature.¹² The authors found that most of the sites were concerned with repositories, data management plans (DMPs), data storage, and metadata, with few interested in visualization or data analysis, qualities of mature RDS. As explained by Carol Tenopir et al., it is not surprising that informational or consultative services are the most common type of services provided by RDS, given that "consultative RDS align well with traditional reference or liaison librarian services."¹³

A cursory and non-standardized review of RDS at 38 libraries, performed by the authors of this paper, only revealed two institutions that provided services that were like the DataLab (see Appendix A). To conduct this review, the authors searched for "Research Data Services" on Google and selected several pages' worth of results that led to college or university Research Data Services departments. From there, the authors searched each respective RDS page for indication of a research program similar to the DataLab.



One service that resembles the DataLab is the Data Club by Columbia University Libraries. According to the Data Club's website, the organization provides a supportive space where people working with data can collaborate and learn from one another.¹⁴ The other service is from the University of Cincinnati, which provides the Undergraduate Research Summer Learning Community (URSLC). The URSLC is similar to the Columbia's Data Club in that it is a space for learning and collaboration while working with data.¹⁵ The rest of the RDS evaluated, however, do not note that they offer services such as hands-on research with an overseeing faculty member and RDS team member, like the DataLab does.

The only other similar service, found in the literature, was presented by Jake Carlson et al. in 2015. Also from the University of Cincinnati is the Research, Education, and Creative Opportunities Network (RECON), which partners students with senior students who are involved in research as a way for them to learn data information literacy.¹⁶ It must be stressed that neither the Data Club, the Undergraduate Research Summer Learning Community, nor RECON are equivalent to the DataLab, further demonstrating the uniqueness of SHU Libraries' RDS research program.

Methodology

This qualitative case study uses interview data to understand the relationship between student participation in the DataLab and students' perceived knowledge and skills in data analysis and manipulation, as well as their motivation to pursue either graduate school or a career in data science as a result of experience with the DataLab. The study was designed to address the following research questions:

1. How does participating in the DataLab impact students' interest in pursuing careers or graduate studies in data science or data-related fields?
2. How does participating in the DataLab affect students' confidence in their abilities to analyze and work with data?

The authors analyzed post-program survey responses and interview data from the DataLab participants; however, only the interview data will be analyzed in this qualitative case study. The survey was used to gather demographic information. In addition, the participants had the opportunity to indicate their interest in follow-up interviews with the researchers. The interview data reflect and elaborate on the questions that were asked in the survey, providing the researchers with a broader and clearer picture of the students' experiences.

DataLab Participant Recruitment

For faculty to participate in the DataLab, they filled out an application to signal their interest in joining the program for one semester; from there, the RDS team reviewed faculty applications and decided which faculty projects could be supported, based on software availability and team expertise.

Of the 27 Spring 2023 DataLab student participants, 14 started the post-participation survey, and one did not consent to continue. Thus, 13 complete submissions were collected. Of those participants, 11 volunteered to be interviewed. Table 1 details student interviewee demographics.



Table 1.

Study participants by participant identifier, major/area of study, and academic year

Student	Major/Area of Study	Academic Year
Student 001	Elementary Special Education	Junior
Student 002	Health Science	Graduate
Student 003	Political Science and History	Senior
Student 004	Computer Science	Sophomore
Student 005	Social and Behavioral Sciences	Junior
Student 006	Environmental Studies	Senior
Student 007	Psychology	Senior
Student 008	Social and Behavioral Sciences	Sophomore
Student 009	Elementary Special Education	Junior
Student 010	Special Education	Senior
Student 011	History and Asian Studies	Senior

Most of the interview participants had been recruited into the DataLab by a faculty member. Seven of the students that were approached by a faculty member worked with that individual in the DataLab; two of the students were approached faculty members they did not work with in the lab (an advisor and a student's professor, who is also a member of RDS). These individuals were simply alerting the students of the opportunity with the DataLab. One student learned about the DataLab from a friend, and another received a message from their resident assistant letting them know that RDS was looking for student participants.

Student Interviews

Seton Hall University's IRB office approved this study before data collection began. Students were asked to provide informed consent before taking the survey in Qualtrics. The survey consisted of questions related to the participants' demographics and their perceptions of data analysis and the DataLab.

The RDS graduate assistant conducted 15-minute interviews with participants via Microsoft Teams. These interviews were recorded with participant consent. The interviews were semi-structured and included 21 questions (see Appendix B). The semi-structured nature of the interviews allowed the DataLab GA to ask questions that were not in the pre-planned list. The interview data were anonymized and are presented as such in this case study.



The interviewer asked participants to reflect on their experiences performing data analysis for their DataLab projects, share their interest in pursuing a career or graduate studies in data analysis or data science more broadly, and to reflect on how the data-centered skills they gained through the DataLab might help them in the future. Moreover, the interviews enabled the participants to provide feedback on the DataLab program. The interviews were transcribed and coded in ATLAS.ti based on themes gleaned from the participants' responses.

Coding

The researchers used three primary methods to code the data: subcoding, structural coding, and evaluation coding. According to Johnny Saldaña, subcoding is a grammatical coding method in which one code is subordinate to another code (for example, parent/child codes); structural coding is an elemental coding method in which a structure is used to categorize sections of code (for example, tagging certain codes "Introduction" to use them for the introduction of a paper); and evaluation coding is an affective coding method, which applies codes to "assign judgement about the merit, worth, or significant of programs or policies."¹⁷

The coding process involved several members of the RDS team. First, the group met to review the interview transcripts. Preliminary codes were developed after an initial exploration of the data; then, each member received a transcript to review thoroughly. The team met again to review the coded transcripts as a group and discuss any issues that emerged throughout the coding process. The first author completed a second round of coding to validate the previous coding process, which included editing and adding codes where necessary.

Results

Students' experiences in the DataLab during the Spring 2023 semester were overwhelmingly positive; however, several participants provided crucial feedback that can be applied in future semesters. In general, students found that their participation in the DataLab increased their interest in data, and they see their experiences in the program as having positive career impacts.

Prior Experience with Data

When asked "How familiar were you with data analysis prior to enrolling in the DataLab?", responses were divided between "no experience" (n=5) and "some experience" (n=8). The students with some experience noted relevant classes they had taken, such as research methods and statistics, whereas other students gained experience with data analysis via other research projects they had done, through family members, and through participation in hackathons.

Interestingly, two of the students that noted they had "no experience" with data analysis specified that they had taken courses that did, or would, cover data analysis, but both considered themselves to have been unfamiliar with the practice when reflect-

ing on the skills they gained through the DataLab. One student who was coded under “no experience” and “some experience” elaborated on their stance:

So, I wasn't too familiar with it [data analysis]. I mean, we took . . . qualitative and quantitative data analysis classes for my PhD program, but I haven't really gotten into the meat and potatoes yet, of really analyzing data in my program. So, this was really like my first exposure to it.

However, the student complicated this perspective further, by noting that in their employment they analyzed financial data, but their experiences with this type of data analysis did not, in their opinion, correspond to the type of data analysis done during the DataLab. In other words, despite working with data daily, the type of work done during the DataLab was more closely aligned with the student's perception of rigorous data work, work that led them to *learn* to work with data. Another student shared a similar sentiment when they said, “The only real [previous] experience I had was in my research methods class where we did a brief overview of one or two programs, but beyond that, it was nothing else.”

Participants were then asked, “On a scale of one to five, how much has your interest in data analysis increased since enrolling in the DataLab?” Six of the participants expressed a high interest in working with data after the DataLab, while three participants were divided between high and neutral interest levels, and one participant recorded neutral or low interest. The students who rated their interest level as 4 or 5 (high) out of 5 on a 5-point scale mentioned that their interest had increased; one student in the high or neutral category also expressed an increase in interest but specified that they “wouldn't want a career in data science.” The other two students with high or neutral interest levels rated their interest level as 3 on the scale 5-point scale, but specified they were interested in the data-centric side of their major—speech-language pathology, in the case of one—and that they were keeping their “options open” as a sophomore, in the case of the other. The one student who expressed neutral or low interest in data science dissented from the rest of the cohort rating their interest at a 2.

DataLab Outcomes

Students who participated in the DataLab were not expected to fully complete their faculty members' projects but were tasked with completing a part of the project over the semester. At the conclusion of the semester, students reflected on the projects and their experience with the DataLab. All of the students felt there were educational benefits to their participation in the DataLab because they could apply their newly developed skills in a classroom or to another research project and/or to current and future employment and education. For example, one stated, “I think it [the DataLab] will, you know, help me with studies I want to do in the future. . . and I got more comfortable with . . . coding survey [and] questionnaire data.” Another also recognized how their time in the DataLab might help them decipher data-heavy research within their field and reflected, “If I'm reading a research study . . . trying to stay up to date with discoveries and things like that in my field. . . I'm able to recognize things that are in the research papers. That's helpful.”

Two participants noted that they would be submitting their projects to present at conferences with their faculty mentors; further, two participants noted plans to con-



tinue their project outside of the DataLab with their faculty mentors. These outcomes indicate that the DataLab provided sufficient support for these teams to continue their projects, acting as a springboard for future scholarly output. The student participants in the DataLab provided insight into their experiences with the program, including the perceived benefits of participation. They also gave crucial feedback the program can use going forward.

DataLab Impact on Career or Graduate School

Most participants felt that the DataLab would have a positive impact on their future careers or potential graduate school trajectory but also suggested that they would likely not want to pursue a graduate degree or career in data science. Students were specifically asked, "How do you think the DataLab has helped you prepare for a career in data analysis?" One student was explicit in their decision not to pursue a data science-related career, stating, "It [the DataLab] confirmed that I would like to do the other idea I had, if that makes sense," that "other idea" being "hands-on experience as an OT [occupational therapist]" rather than "doing research related to occupational therapy."

Nearly all the participants (n=10) mentioned that their experiences in the DataLab helped lay a foundation in data analysis and felt that these skills could benefit them in the future. For example, one student said, "I wasn't considering a data analysis career before this, but I think that it's prepared me for . . . a research job." One student, who is in a 3+3 program (3 years undergraduate, 3 years graduate), stated, "I think it [DataLab] taught me skills that I'm probably going to need for graduate school." While DataLab did not influence this student to attend graduate school, as they were already enrolled in the 3+3 program, they affirmed that what they learned in the program would aid them as they enter their graduate education.

The only participant currently enrolled in graduate school mentioned that what they learned in the DataLab would help them with their dissertation research, stating, "I really got to play around with ATLAS.ti . . . and it really inspired me, you know, for my dissertation, since for the DataLab, we did kind of like a questionnaire study. It kind of inspired me to think 'that's what I want to do for my dissertation.'" Another student discussed how their mentor told them "most of his grad students could not . . . 'get' this data, because . . . they don't have the background knowledge to do it," attributing their own ability to perform complex research skills to the DataLab. When asked if they felt prepared for a potential career in data analysis, the same student replied, "I'm not going to have a career in data analysis, but I do think if I were to . . . start having one, I do think it created a great backbone for me."

When asked, "How do you envision the DataLab impacting your future career?", the same ten students also indicated that they could foresee or were currently experiencing positive career impacts after their time with the DataLab. For example, two students related what they had learned in the DataLab to their current jobs. One of them stated, "I may be working a data entry sort of position between now and graduate school, so I think it [the DataLab] might help me a little bit with that." The other noted,

We did a training on Excel, and I work as a billing clerk at a law firm, and I input timesheets, so sometimes lawyers will send me their timesheets in Excel and I'm able to use the functions that they [the DataLab] taught me just to make my life so much easier.

Even though this same student had stated earlier that their experience in the DataLab affirmed their interest in pursuing a career that was not related to research, they could see how their DataLab experience benefitted their job performance.

Several students (n=4) stated that their time in the DataLab would support their upcoming graduate school applications. One who related the skills gained in the DataLab to their humanities major, stated

It will be interesting to have . . . a humanities-heavy resume and then just have [the] DataLab there, and it's . . . this is just not something that if you're so interested in humanities and writing that you'll necessarily go for. . . so I think it'll be interesting and definitely make me stand out when it comes to grad school applications, just regular research applications, that sort of thing.

Program Feedback

The students provided the researchers with feedback that will enable the DataLab and RDS team to modify the program to better fit the needs of future participants. Faculty have not been asked to formally provide feedback, though this procedure may change in the future. Positive student feedback included appreciation for flexible scheduling (n=4), access to software through the Libraries (n=1), appreciation for receiving hands-on experience in data analysis (n=4), the sense that the training quality was high (n=11), and a general sense that the students felt supported by the DataLab team (n=9).

Critical feedback noted areas for program improvement, such as creating more focused class content (n=2), providing clearer parameters for the projects students will be working on (n=3), providing greater vetting of faculty and project proposals (n=1), and creating a community of practice for DataLab student participants, facilitated by the DataLab team (n=9).

The RDS team has already responded to some of this feedback. For example, regarding the issue of vetting faculty and project proposals, the DataLab and RDS have created stricter guidelines to ensure that faculty work closely with their students throughout the semester and do not leave them feeling adrift with only the DataLab to support them. A revised "Expectations and Guidelines" form has been created that clearly outlines the expectations of the faculty member and student to ensure that all participants understand their responsibilities as a member of the DataLab.

Students who participate in the DataLab are encouraged to present their work at Seton Hall University's Petersheim Academic Exposition, which is held at the end of every spring semester. The Seton Hall University Libraries and RDS organize a DataLab-specific event during the exposition that provides a venue for students to present their research with their faculty mentors. Three participants specifically noted that they enjoyed being able to see their fellow DataLab participants present their research during Petersheim and appreciated the, albeit brief, community aspect of Petersheim.

Discussion

The finding that participation in the DataLab program did not increase students' likelihood of pursuing data science-related careers is echoed in the literature. Chad T. Miller,



Merrit Drewery, Tina M. Waliscek, Ryan N. Contreras, and Chieri Kubota argue that undergraduate research can “build career confidence” and “self-awareness of career goals.”¹⁸ Even though most of the participants indicated they would not pursue a career in data science or data analysis, the students’ experiences contributed to their career choice process. As one student explained, “I would say I would be open to working in data science if the opportunity came, but it’s not my number one choice for a career.”

Andrew Zydney et al.’s research agrees with Miller et al., finding that students who participate in one semester of an undergraduate research experience tend to perceive their experience as providing a moderate benefit compared to students who participated in research for a longer period.¹⁹ If DataLab cohorts were to run for more than one semester, student participants might better conceive of themselves as future data scientists.

Even so, students felt more confident in their ability to analyze and work with data after their time with the DataLab. Many students referred to feeling “confident”, “comfortable”, “familiar”, and less “scared” than before their participation with the DataLab program. The only graduate student participant in the DataLab for Spring 2023 expressed their increased confidence, stating, “I feel like it [data analysis] was always something scary and what I was dreaming with my PhD program, but getting some exposure to it, I realized it’s really not that scary and [the DataLab] made me more comfortable and excited about it.”

One student observed that they felt they were a more “well-rounded” employee because of the skills they gained in the DataLab, and another noted that research sounded difficult to them before participating in the program, but the DataLab “really opened my eyes and broke it down to me.” A humanities student, who indicated they had “no experience” with data analysis, said that their future graduate school interest of political science will likely mean handling large datasets. Reflecting on their time in the DataLab and the skills they gained, they stated, “So, I think being able to use RStudio and then knowing how to not get super overwhelmed by all of this large amount of data will be super helpful.” Students from various backgrounds and at various points in their educational journeys overwhelmingly felt more confident in their ability to handle and analyze data after participating in the DataLab.

Limitations

Only 11 participants were interviewed for this case study; therefore, while data saturation was reached in that new data did not appear over time, it is possible that the results of this paper may not be generalizable to other libraries. The students interviewed for this study were deliberately selected as DataLab participants and may not fully represent students at other institutions or students who do not have a strong interest in data research.

Conclusion and Future Considerations

This study examined the experience of undergraduate and graduate students enrolled in the Seton Hall University DataLab to understand the impact the DataLab had on the students. During the semester, the participants had the opportunity to work on research projects, receive faculty mentoring, and work one-on-one with the Libraries’ RDS team



while pursuing their projects. Most of the participants enjoyed their time with the DataLab and considered their experience to be beneficial.

The DataLab increased participants' confidence in their data analysis skills as well as their ability to communicate data to others. This was an expected outcome, based on previous studies that found that research experiences build student confidence.²⁰ The material benefits of students' experiences with the DataLab have emerged since the end of the Spring 2023 semester. For example, one of the Spring 2023 DataLab students graduated and found a job using ArcGIS after having worked on a project using that software in the DataLab. Further, two of the DataLab participants asked the RDS Data Specialist to write them letters of recommendation for graduate school. Lastly, one of the participants has since presented at Seton Hall University's Love Data Week, in partnership with Love Data Week from the University of Michigan's Institute for Social Research and Inter-university Consortium for Political and Social Research (ICPSR), and presented at the Southeast Data Librarian Symposium, where they hosted a workshop on Python for data analysis. College research experiences are not one-size-fits-all, but providing students with ample resources, faculty mentorship, time, flexibility, and help, can enable them to leave their research experiences with not only new technical and social skills, but a better understanding of their future potential.

The DataLab is frequently in a state of flux. As it relies on grant funding, there is not always a guarantee that funds will be available to pay students at the end of the semester. However, this does not mean that the DataLab is unsustainable or does not have a future; instead, it indicates that the RDS team must be flexible in how they approach compensation for the DataLab experience. For example, the RDS team has discussed implementing course credit for the DataLab, offering the experience as an independent study course or as an unpaid extracurricular activity. For libraries interested in replicating the DataLab, funding opportunities may make the program more enticing to students, but it is possible to run a similar program using alternative means of compensation

Acknowledgement

The authors would like to thank Dr. Sharon Ince, SHU Libraries' Assistant Dean of Information Technologies Collection Services; Lisa DeLuca, Assistant Dean of Public Services; and Dr. Lisa Rose-Wiles, Science Librarian, for their invaluable feedback on this article. The authors would also like to thank Dr. Samah Alshrief, Data Specialist, for providing critical insight into this paper's methodology.

Maria Barca is the lead instruction librarian at Seton Hall University, email: maria.barca@shu.edu, ORCID: 0000000277941795.

Kathryn Diana is a PhD student and graduate research assistant at The Ohio State University, email: diana.15@osu.edu, ORCID: 0009000468689374.



Appendix A

List of Schools with RDS Programs and DataLab-Adjacent Programs

Library	DataLab Adjacent	Website
Columbia University Libraries	Data Club	https://library.columbia.edu/research-teaching/research-data/data-club.html
Rutgers University (all campuses)		https://www.libraries.rutgers.edu/research-support/research-data-services
University of Cincinnati Libraries	Undergraduate Research Summer Learning Community	https://libraries.uc.edu/research-teaching-support/research-data-services.html
Georgia State University Library		https://research.library.gsu.edu/rds
University of Minnesota Libraries		https://www.lib.umn.edu/services/data
Oregon State University Libraries		https://guides.library.oregonstate.edu/research-data-services
University of Washington Libraries		https://lib.uw.edu/dataservices/
University of Iowa Libraries		https://www.lib.uiowa.edu/data/
University of Texas at Arlington Libraries		https://libraries.uta.edu/research/data
Texas State University Libraries		https://www.library.txst.edu/services/research-services.html
Rice University Fondren Library		https://library.rice.edu/places/research-data-services

This manuscript has been reviewed, copy edited, and accepted for publication, portal 26.2.



Appendix A, Cont.

Library	DataLab Adjacent	Website
Purdue University Libraries		https://lib.purdue.edu/researchdata/
Stony Brook University Libraries		https://library.stonybrook.edu/scholarly-communication/data-services/
University of Delaware Library		https://library.udel.edu/rdsg/
Northeastern University Library		https://library.northeastern.edu/data_services/research-data-services/
Queens College Library		https://qc-cuny.libguides.com/c.php?g=862047&p=6178682
University of Virginia Library		https://library.virginia.edu/data
University of Illinois Urbana-Champaign University Library		https://researchdataservice.illinois.edu/
Dartmouth Libraries		https://researchguides.dartmouth.edu/data_management
NYU Libraries		https://library.nyu.edu/s/data-services/
Penn State University Libraries		https://libraries.psu.edu/research/research-data-services
Johns Hopkins Libraries		https://dataservices.library.jhu.edu/
University of North Carolina Libraries		https://library.unc.edu/data/
University of Pittsburgh Library System		https://library.pitt.edu/research-data-services

This manuscript has been peer reviewed, copy edited, and accepted for publication, portal 26.2.



Appendix A, Cont.

Library	DataLab Adjacent	Website
MIT Libraries		https://libraries.mit.edu/data-services/
Temple University Libraries		https://library.temple.edu/services/data-management-planning
University of Maryland University Libraries		https://www.lib.umd.edu/research/data
Harvard Countway Library		https://countway.harvard.edu/services/publishing-data-services/data-services
Harvard Library		https://osrds.library.harvard.edu/
Brandeis University Library		https://www.brandeis.edu/library/research/data-services.html
Florida State University Libraries		https://www.lib.fsu.edu/research-and-publish/research/data-services
University of New Mexico Libraries		https://libguides.unm.edu/data
The University of Texas at Austin Libraries		https://www.lib.utexas.edu/research-help-support/research-data-services
Carnegie Mellon University Libraries		https://www.library.cmu.edu/services-overview/data
UC Santa Barbara Library		https://www.library.ucsb.edu/research-data-services
University of Michigan Library		https://www.lib.umich.edu/research-and-scholarship/data-services
UCLA Library		https://www.library.ucla.edu/help/services-resources/data-services
Stanford University Libraries		https://library.stanford.edu/research-support/data-support-and-services

This manuscript has been peer reviewed, copy edited, and accepted for publication, portal 26.2.



Appendix B

Interview Questions

How familiar were you with data analysis before enrolling in the DataLab?

On a scale of one to five, how much has your interest in data analysis changed since enrolling in the DataLab?

Have you taken any other data related courses before the DataLab?

How often do you engage in data-related activities outside of class, if at all?

Would you be interested in doing more data related activities outside of class?

Have you considered pursuing a career or a graduate degree in data science before enrolling in the DataLab?

Has the DataLab influenced your decision to pursue a career or a graduate degree in data science? If so, how?

How do you rate the quality of the lab consultation and training and their relevance to data analysis?

How do you rate the quality of the instruction received during your participation in the DataLab?

What did you like most about the DataLab?

Can you tell me about your decision to enroll in the DataLab?

Have you noticed any changes in your interest or career goals since beginning the DataLab?

How have you found the DataLab experience so far?

How do you think the skills and knowledge you have gained in the DataLab will impact your future career plans?

Have you had any opportunities to apply what you have learned in the DataLab to real world situations or projects outside of the DataLab?

What tool did you use in the DataLab?

Do you think the DataLab has helped you prepare for a career in data analysis? If so, how?

Have you had any opportunities to network with other students in the DataLab who are interested in similar career paths?

Would you have preferred to network throughout the whole semester or just that one experience of Petersheim?

How do you envision impact of the DataLab on your future career success?

What suggestions do you have for improving the DataLab in the future?



Notes

1. Qiong Xu, "Investing and Understanding Library Data Services to Support College Student Data Literacy Competencies: A Conceptual Framework," *Journal of Librarianship and Information Science* 57, no. 1 (2025): 3, <https://doi.org/10.1177/09610006231196604>.
2. Marek Nahotko, Magdalena Zych, Aneta Januszko-Szakiel, and Małgorzata Jaskowska, "Big Data-Driven Investigation into the Maturity of Library Research Data Services (RDS)," *Journal of Academic Librarianship* 49, no. 1 (2023): 102646, <https://doi.org/10.1016/j.acalib.2022.102646>.
3. Shawn V. Ault, Soohyun Nam Liao, and Larry Musolino, *Principles of Data Science* (OpenStax, 2025).
4. Nahotko et al., "Big Data-Driven Investigation."
5. "Completed Academies," Seton Hall University, accessed April 4, 2024, <https://www.shu.edu/provost/seeds-of-innovation/completed-academies.html>; "DiploLab," Seton Hall University School of Diplomacy and International Relations, accessed April 4, 2024, <https://www.shu.edu/diplomacy/diplolab.html>.
6. Terrence B. Bennett and Shawn Nicholson, "Interactions Between the Academic Business Library and Research Data Services," *portal: Libraries and the Academy* 4, no. 1 (January 2004): 107, <https://doi.org/10.1353/pla.2004.0001>.
7. Andrew M. Cox and Stephen Pinfield, "Research Data Management and Libraries: Current Activities and Future Priorities," *Journal of Librarianship and Information Science* 46, no. 4 (2014): 299–316, <https://doi.org/10.1177/0961000613492542>.
8. *Ibid.*, 303.
9. Xu, "Investing and Understanding Library Data Services to Support College Student Data Literacy Competencies," 5.
10. Carol Tenopir, Ben Birch, and Suzie Allard, *Academic Libraries and Research Data Services: Current Practices and Plans for the Future* (ACRL, 2012); Carol Tenopir, Dane Hughes, Suzie Allard, Mike Frame, Ben Birch, Lynn Baird, Robert Sandusky, Madison Langseth, and Andrew Lundeen, "Research Data Services in Academic Libraries: Data Intensive Roles for the Future?," *Journal of eScience Librarianship* 4, no. 2 (2015): e1085, <https://doi.org/10.7191/jeslib.2015.1085>; Carol Tenopir et al., *Research Data Services in Academic Libraries: Where are We Today?* (Choice, 2019); Inna Kouper et al., "Research Data Services Maturity in Academic Libraries," in *Curating Research Data: Practical Strategies for Your Digital Repository*, ed. Lisa R. Johnston (ACRL, 2017).
11. Kouper et al., "Research Data Services Maturity."
12. Nahotko et al., "Big Data-Driven Investigation," 11.
13. Tenopir et al. "Research Data Services in Academic Libraries," 6.
14. "Data Club," Columbia University Libraries, accessed May 2, 2025, <https://library.columbia.edu/research-teaching/research-data/data-club.html>.
15. "Research & Data Services," University of Cincinnati Libraries, accessed May 2, 2025, <https://libraries.uc.edu/research-teaching-support/research-data-services.html>.16. Jake Carlson, Megan Sapp Nelson, Lisa R. Johnston, and Amy Koshoffer, "Developing Data Literacy Programs: Working with Faculty, Graduate Students and Undergraduates," *Bulletin of the Association for Information Science and Technology* 41, no. 6 (August-September 2016): 17, <https://doi.org/10.1002/bult.2015.1720410608>.
17. Johnny Saldaña, *The Coding Manual for Qualitative Researchers* (SAGE, 2016), 91–140.
18. Chad T. Miller, Merrit Drewery, Tina M. Waliscek, Ryan N. Contreras, and Chieri Kubota, "Engaging Undergraduate Students in Research," *HortTechnology* 33, no. 1: 1, <https://doi.org/10.21273/HORTTECH05130-22>.
19. Andrew Zydney, Joan S. Bennett, Abdus Shahid, and Karen W. Bauer, "Impact of Undergraduate Research Experience in Engineering," *Journal of Engineering Education* 91 (2002): 154, <https://doi.org/10.1002/j.2168-9830.2002.tb00687.x>.
20. Yusuff Adebayo Adebisi, "Undergraduate Students' Involvement in Research: Values, Benefits, Barriers and Recommendations," 81 (September 2022): 104384, <https://doi.org/10.1016/j.amsu.2022.104384>.

This mss. is peer reviewed, copy edited, and accepted for publication, portal 26.2.