

Research Data Services

2019-2020 Annual Report

Northwestern | INFORMATION TECHNOLOGY
RESEARCH COMPUTING SERVICES

Research Data Services

2019-2020 Annual Report

Prepared by
Christina Maimone, Research Data Services Lead
December 2020

Overview	2
New Initiatives	3
Projects	4
Service Demand Growth	5
Broad Impact	6
Supporting University Priorities	7
External Support and Team Members	8

Overview

The [Research Data Services](#) team within Northwestern IT Research Computing Services supports Northwestern researchers with data science, data visualization, and computer programming training, consultations, and collaborative support.

2019–2020 Support by Research Data Services



257

One-on-one consultations for researchers¹ from every Northwestern school facing data science challenges in their work



14

Faculty research projects supported with programming, data analysis, and data visualization services



5,400

Total workshop registrations



157

Hours of Python, R, and other data science workshops



1,400

Distinct researchers registered for workshops



580

Researchers learned new skills via access to interactive, online data science courses²



19

Researchers received small group coaching through Bring Your Own Data (BYOD) Working Groups

¹ Researchers include any Northwestern affiliate engaged in research activities, from undergraduate students pursuing independent research projects through faculty. Graduate students are the most common level of researcher we support with consultation and training services, but we also support faculty, postdocs, research staff, and undergraduates.

² Thanks to funding from Northwestern IT, Galter Health Sciences Library, Kellogg School of Management, and the Department of Political Science for subscriptions to DataCamp.

New Initiatives

In 2019–2020, we initiated multiple new programs, services, collaborations, and processes to support researchers. These efforts expand and complement the training, individual consultation, and collaborative project support services we continue to provide.

- Expanded [Data Science Research](#) Consultants program: Three students were funded through graduate assistantships from The Graduate School for the first time in 2019-2020. These assistantships allowed PhD students to work with Research Computing Services on a consistent basis in a way that complements their graduate training. Northwestern researchers benefit from increased capacity for consultations and additional workshop offerings, while the students benefit from one-on-one mentoring, being part of a team of data-science peers from across the University, training in presentation and teaching skills, and experience working with a wide range of data and computational technologies.
- [Drop-in Data Hours](#): In cooperation with Northwestern Libraries, drop-in hours supplement scheduled consultations and expand collaboration with Library GIS experts.
- [Bring Your Own Data Working Groups](#): Small groups provide support and encouragement for researchers as they work on a data-intensive part of their research throughout a quarter.
- [Virtual workshops](#): In response to the pandemic, we updated our workshop materials to offer virtual workshops for the first time and support researchers working remotely.
- [Online learning guides](#) on data science topics: Staff and students curated lists of resources to help researchers sort through the many free online and library resources available for learning R, Python, and other data skills.
- [Biomedical Data Science Day](#): We participated in the NUCATS organized event by serving on the organizing committee, providing customized computing environments to allow interactive workshops during the event, and giving introductory workshops on R and Python.
- [Feedback surveys](#): We began requesting feedback on our consultation services regularly. Ninety-four percent of respondents were satisfied with their consultation, and 98 percent would recommend the service to their colleagues. Sixty-five percent noted they learned something new through the consultation, and 59 percent reported that the consultation saved them time.
- [Workshop for Northwestern University in Qatar \(NU-Q\)](#): We collaborated with NU-Q research staff to design and schedule an introductory data visualization workshop based on the interests and availability of NU-Q faculty and students.

Projects

Staff and students collaborated with researchers on 14 longer-term projects this year, including:

- **Feinberg School of Medicine**, Augusta Weber Office of Medical Education: Stemmler Fund-supported project to investigate methods for incorporating text-analysis tools into student evaluation processes.
- **Vicky Kalogera**, Weinberg/CIERA, and the LIGO Research Team: Created interactive visualizations of black holes and neutron stars.
- **Larissa Buchholz**, School of Communication: Provided data visualization design support and produced publication-quality data visualizations for her upcoming book on international art markets.
- **Beth Redbird**, Sociology: Developed and deployed an interactive data visualization for the public to explore data on racial bias in police arrests for an upcoming publication. This project was supported primarily by our student consultants.
- **Beth Redbird**, Sociology: Created a custom web application and database to enable RAs to efficiently code data about images from newspaper articles related to the pandemic.
- **Infant and Child Development Center**, PI: Sandra Waxman: Assisted the center's research coordinator with updating their study participant tracking database to address data inconsistencies and facilitate easier analysis of the research data.
- **Sarah Fay**, SPS English: Analyzed and visualized author interviews from the Paris Review.
- **Martin Mueller**, English (Emeritus): EarlyPrint Project (Mellon Foundation grant): Provided software engineering support to process early English texts and make them available for researchers with annotation tools.

Additional projects include:

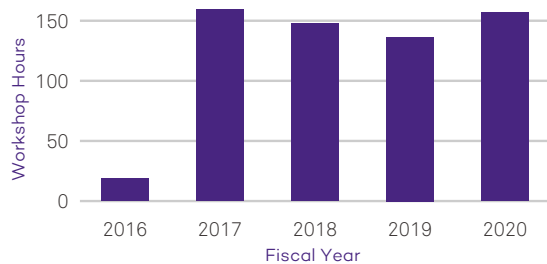
- **Data Services Assessment**: Conducted focus groups with research faculty and staff and researched statistical consulting services offered by peer institutions as part of an assessment for the [Research Technology Advisory Committee](#). Identified data storage, data management, and statistical consulting as key areas where researchers are looking for additional services and support. The assessment will be completed in Academic Year 2020–2021.
- **Remote Teaching and Learning Metrics**: Developed metrics and data visualizations to help Northwestern IT and University leadership monitor the University's transition to remote teaching and learning using Canvas, Zoom, and other resources.

Service Demand Growth

Demand for data science services is growing. We nearly doubled the number of workshops we offered in 2019-2020 compared to the previous year, and we increased the total hours of instruction by 15%.³ Registrations for workshops increased even more dramatically—the average number of people registered for a given hour of workshop instruction doubled (up 93%) compared to the previous year.⁴

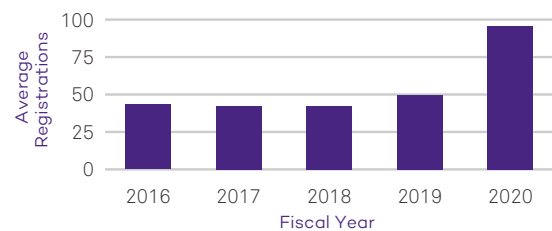
The virtual format of workshops in Spring and Summer 2020 made workshops accessible to those who could not previously travel to in-person sessions. In addition to live, synchronous training by Research Data Services staff, 580 researchers also had access to vendor-provided, on-demand, online data science courses.

Workshop Hours by Year



Average Workshop Registrations

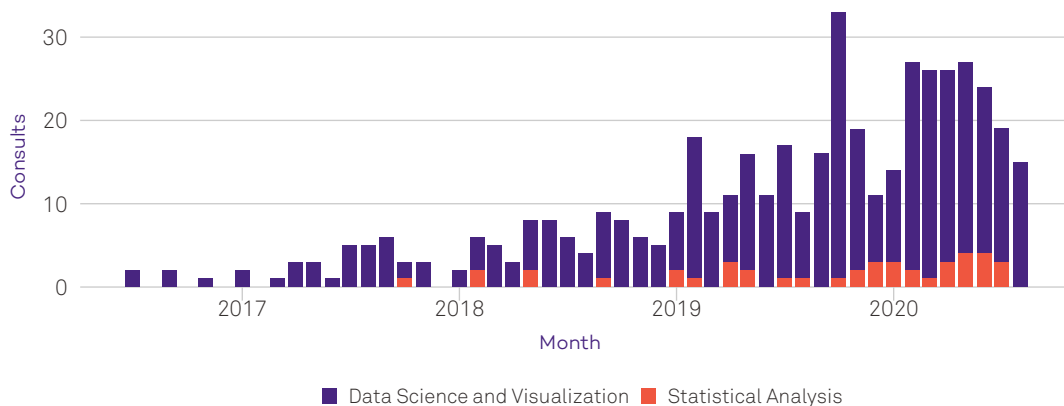
Normalized by Workshop Length



The consultation service is also growing, with an average of 22 consultations per month in 2020. We track the type of consultations and the number to help identify demand for new service areas. The plot below calls out requests for statistical consulting, which are supported when the team has relevant expertise.

Demand for Data Consultations is Growing

Averaged 22 consults per month in 2020



³ To better accommodate a remote format, we offered shorter workshops more frequently rather than longer in-person sessions. The workshop data in this report excludes Quest- and HPC-oriented workshops, typically led by other members of the Research Computing Services team.

⁴ The plot shows the average number of people registered for an hour of workshop instruction.

Broad Impact

Research Data Services supports researchers at all levels and from all schools.⁵ Breakdowns for Academic Year 2019–2020:

Role	Consultations (% of consults)	Workshops (% of registrants)	BYOD (count)	Projects (count)
Faculty	9%	7%	1	13
Postdocs	7%	13%	2	
Staff	13%	20%	2	3
PhD Students	48%	31%	12	
Other Graduate Students	8%	16%	1	
Undergraduate Students	13%	9%	1	
Other/Unknown	2%	6%		

School	Consultations (% of consults)	Workshops (% of registrants)	BYOD (count)	Projects (count)
Weinberg College of Arts and Sciences	39%	22%	7	7
Feinberg School of Medicine	16%	26%	4	3
McCormick School of Engineering & Applied Science	13%	18%	2	
School of Communication	7%	6%	1	1
School of Education & Social Policy	9%	1%	2	1
Kellogg School of Management	2%	4%	2	
NU-Q/Other/Unknown	3%	11%	1	1
Bienen School of Music	2%	<1%		
Northwestern Pritzker School of Law	1%	<1%		
University Central Administration	2%	1%		2
Medill School of Journalism, Media, IMC	3%	3%		
School of Professional Studies	1%	8%		1

⁵ Graduate students are associated with the school of their primary department.

Supporting University Priorities

The Research Data Services team supported multiple high-priority areas for Northwestern during the year. In addition to our ongoing role in supporting the research eminence of Northwestern University, our work also focused on:

Diversity, Equity, and Inclusion

- Staff member Colby Witherup Wood is a Northwestern IT Diversity, Equity, and Inclusion Committee member.
- Our workshop offerings include many introductory sessions that provide an opportunity for students who did not learn data and computational skills prior to coming to Northwestern to learn them outside of coursework. Students would pay several hundreds of dollars for equivalent workshops at conferences or from private companies.
- Our hiring process for data science student consultants uses best practices for fair and unbiased hiring. We actively recruit students for these positions from groups traditionally underrepresented in data science and programming fields.
- We intentionally cultivate an inclusive work environment for staff and student consultants that includes flexible scheduling, input into work assignments, and support for dealing with challenging conversations and consultations.

Responding to COVID-19

- We transitioned data science and computational workshops to an online format to support researchers working remotely, created online learning guides to provide access to other University resources, and continued consultation services remotely.
- We supported researchers studying COVID-19 and society's response to the pandemic through faculty collaboration projects, 1:1 consultations, and Bring Your Own Data working groups.
- We worked with our IT colleagues to develop Remote Teaching and Learning metrics and data visualizations for IT and University leadership.

Student Experience and Development

- Our training and consultation services offer opportunities for skill growth and development beyond what is available in students' departmental programs and help students connect with the Northwestern community beyond their department.
- Our student consultant program provides professional development opportunities for the consultants to prepare them for careers within or outside academia.

External Impact

Research Data Services staff contributed to the national research computing and data community in multiple ways. These contributions help staff stay up to date on new technologies and service trends, form collaborations with colleagues at other institutions, and establish Northwestern as a leader in the field.

Contributions include:

- US Research Software Engineer Association (US-RSE) Steering Committee
- US-RSE Workshop Presentations
- Campus Research Computing Consortium (CaRCC) Data-Facing Network Coordinator
- CaRCC Data-Facing Network Presentations
- Practices and Experiences in Advanced Research Computing (PEARC20) Reviewer
- PEARC20 Presentations
- EDUCAUSE19 Presentations

Team Members

Staff members during 2019–2020:

- Christina Maimone
- Colby Witherup Wood
- Philip Burns (retired June 2020)
- Aaron Geller (1/6 FTE)

Student Data Science Consultants, 2019–2020:

- Austin Alleman, Statistics
- Andre Archer, Engineering Sciences and Applied Math
- Katie Evans, Interdisciplinary Biological Sciences
- Laura Garcia, Political Science
- Andrew Hall, Psychology
- Darren Hsu, Biophysical Chemistry
- Antonio Nanni, Sociology
- Carrie Stallings, Sociology
- Dan Turner, Linguistics

Northwestern | INFORMATION TECHNOLOGY
RESEARCH COMPUTING SERVICES

it.northwestern.edu/research