Data Analysis and Visualization in Libraries

Xuying Xin, Data Analyst
Cynthia Hudson Vitale, Director, ARL
Former: Head, Research Informatics & Publishing
Research Informatics & Publishing

The Department of Research Informatics and Publishing enhances the technology-driven teaching and research capacity of the Penn State community. We partner with students, faculty, and staff to consult, provide training, and support projects in the areas of research data management, digital humanities, mapping, statistical analysis, and open publishing.

Our services are based on the tenets of openness, innovation and technology, collaboration, and supporting the entire research workflow.

Services and Centers

- Center for Black Digital Research
- Digital Humanities Services
- Research Data Services (RDS)
- Open Publishing Services
- Donald W. Hamer Center for Maps and Geospatial Information
- Data Learning Center

Support:
- Training
- Consultations
- Reference transactions
- Project support
- Infrastructure

Resources

- Advanced Analytics and Visualization Digital Lab
- Digital Projects
- Training & Events

https://libraries.psu.edu/about/departments/research-informatics-and-publishing
Spaces: Advanced Analytics and Visualization Digital Lab

The advanced analytics and visualization computer lab space is equipped with three laptop workstations and Solstice. Walk-in use for individual computer spaces or the entire digital lab space is available to Penn State students, staff, and faculty members during open library hours.

Service Portfolio

- **Statistical Analysis:** R, STATA, SPSS, SAS
- **Programming:** Python, MATLAB
- **Qualitative Analysis:** Nvivo
- **Data Analytics:** Power BI, Tableau, Azure, GCP
- **Data Visualization:** Adobe Illustrator
- **Map & GIS:** ArcGIS, QGIS
Data Learning Center

Data Analytics and Visualization
[link]
Learn about the resources available for Data Analytics & Visualization at the Digital Lab and meet with us to discuss how we can help with your research projects.

Statistical Consultations
[link]
From data collection to interpretation of results, we help you use statistical software for data analysis.

Data Management
[link]
Get assistance with data management plans, tools, and requirements — and learn how to share and archive your work.

Data Archiving and Preservation
[link]
Work with us to properly archive and preserve your research data in our institutional repository, SchoarSphere.

Finding Data
[link]
Find, cite, and manage data for your research needs.

Geospatial Mapping & Analysis
[link]
Consult with us to evaluate, find, and analyze spatial data, automate geospatial workflows, and map results.

[link]
Consultation

College Unit

desk is Research Informatics & Publishing: Data Learning Center

- Liberal Arts: 22.9%
- Agricultural: 16.0%
- H&H Development: 15.3%
- Engineering: 11.0%
- Education: 10.3%
- Science: 5.3%
- Medicine: 3.7%
- PSU Libraries: 3.3%
- Business: 2.3%
- Arts & Arch: 2.3%
- Earth & Mineral: 2.3%
- IST: 2.0%

Resources Used

desk is Research Informatics & Publishing: Data Learning Center

- Software: 97.4%
- Other Resource: 2.6%
Data Analysis and Visualization in Libraries

Xuying Xin, Data Analyst; Cynthia Hudson-Vitale, Head Research Informatics and Publishing

Reference:
- Patron Type
  - Pa Resident (Non-PSU): 0.5%
  - Post Doc: 1.0%
  - Undergrad: SHC: 1.6%
  - Undergraduate: 2.6%
  - Other: 3.1%
  - Staff: 9.9%
  - Faculty: 28.7%
  - Graduate: 52.6%

Consultation:
- Patron Type
  - Pa Resident (Non-PSU): 0.5%
  - Post Doc: 5.5%
  - Undergraduate: 3.9%
  - Staff: 3.9%
  - Other: 1.0%
  - Graduate: 53.2%
  - Faculty: 18.8%

Reference:
- Service Type
  - Data Management: 47%
  - Qualitative: 20%
  - Quantitative/Statistics: 18%
  - Tech Support: 7%
  - Information discovery: 5%
  - Open Publishing: 2%
  - Maps/GIS: 1%
  - Software Curation: 1%

Totals for Reporting Period
- Response Sets: 308
- Question/Answer Pairs: 3855

Totals for Reporting Period
- Response Sets: 199
- Question/Answer Pairs: 2738
Data Analytics and Visualization Service Development Approach

In early 2019 members of the Research Informatics team undertook a University-wide landscape analysis of existing support for big data analytics and data science services and infrastructure. We applied a mixed-methods approach which began with a University-wide website review to identify potential stakeholders. Select stakeholders were then interviewed to understand extent of services, populations served, and goals for program development. From this review we developed the service portfolio below, with a focus on supporting graduate and undergraduate students.

Service Delivery Modes
• Consultations
• Workshops
• Course integrations

Next Steps
Continue to collect data on services, identify gaps, and partner with colleagues, such as the Institute for Computational and Data Sciences.
Data Analytics and Visualization Software Tools

**Business Analytics:**
- MS Power BI
- Tableau
- MS SQL Server, Business Intelligence Suite – SSIS, SSAS, SSRS
- MS Access & Excel
- Other BI tools

**Big Data Analytics:**
- Google Cloud Platform (GCP) with Big Data and Machine Learning tools
- MS Azure
- Amazon Web Services (AWS)
COVID-19 cases and deaths by state and date (MS POWER BI - Penn State University Libraries)

Friday, September 4, 2020
Acknowledgements

• Research Informatics and Publishing at Penn State University Libraries
• Dean’s office at Penn State University Libraries
• Public Relations and Marketing at Penn State University Libraries
• Libraries Strategic Technologies at Penn State University Libraries
• Library Learning Services at Penn State University Libraries
• Cynthia Hudson-Vitale, Seth Erickson, Hannah Hadley, Tara Anthony, Briana Ezray, Ann Thompson, Karen M Hackett, Brandon Zimmerman, Dan Peters, Rebecca Miller Waltz, Rebecca Peterson, and Cynthia Richter