

ANSS Sociology Librarians Discussion Group
Jeff Lacy and Rui Wang, Co-Conveners
Saturday January 25, 2014 PCC 10:30-11:30am

The Sociology Librarian Discussion Group meeting was held on Saturday January 25 in Pennsylvania Convention Center 10:30-11:30. Rui convened the meeting. Jeff could not attend due to weather-related flight cancellations.

Twenty-two people attended. Most of them are sociology, anthropology, and social sciences librarians. The main theme of this Sociology Librarian Discussion Group meeting was social science data use and instruction. We invited two presenters to this meeting - Christine Murray, Social Sciences Data Services Librarian at University of Pennsylvania and Jen Darragh, Data Services and Sociology Librarian at Johns Hopkins University. Their presentations were well received by this group.

Christine Murray handles research questions on numeric and spatial data in the social sciences, including U.S. Census information, Philadelphia city data, and other government and academic statistics. She has a background in geographic information systems (GIS) and also serves as the contact for Penn's programs in Criminology, Sociology, and Urban Studies. Christine has a lot of experience in working with the courses in the demographics program and other social sciences. Thus, her presentation focused on how she introduced data resources in course-related instruction, resources for teaching with data, and how to target outreach to courses that involve quantitative data analysis.

Many of these courses require students to find datasets for their own topics or conduct literature reviews. Since there is no single, centralized data source, Christine's primary challenges are how to teach students to find data for various topics and how to make students excited about finding data.

Christine's strategy is to use one dataset and one tool for one course assignment to lower the level of the difficulty in using data. For example, for a course of poverty and inequality, Christine focused on teaching students to use Social Explore to make maps and graphics. Christine also used online analysis tools, such as GSS and SDA, to lower students' barriers in data analysis, so that students did not need to use SPSS to download and run data. Christine also prepared online guides and provided "research parties" (research consultations) for students who need further instruction after the BI sessions.

Christine designed a project to give small groups various datasets and train students in knowing what information they need to look for when choosing datasets. She also drew interesting topics and examples from prepackaged presentations of data visualization tools, news, blogs, or Google public data to motivate students. Christine suggested finding the institutional requirements was important to seek opportunities for instruction data use.

Christine concluded her presentation with a list of valuable resources for teaching students in the use of data sources.

Jen Darragh provided an overview of her work with restricted data. She has a hybrid position where a quarter of her time is spent to support demographers and other researchers in the use of data sets. Most of the patrons she assists do research in public health. Jen assists her patrons in the contractual use of secondary data as well as the use of the data they gather in their own primary research when there may be security risks.

Jen described restricted data in terms of “deductive disclosure.” As she explained, this is when individuals’ personal identities may not be readily available in a given dataset, but it may be possible to deduce through the correlation very specific medical, economic, geographic, and other very detailed information. Other forms of data may be restricted to protect the identities of children and other protected populations.

Due to the possibility of deductive disclosure, very detailed health information for small geographic regions and other low-population locations (such as prisons) are not available publicly. Jen assists her patrons in determining if they need restricted data for their research and identifying possible sources for restricted data.

Jen noted that she primarily works with graduate students and faculty. The process of gaining access to restricted data can often take months, so undergraduates often do not have the time required.

Jen said that dealing with restricted data ultimately comes to secure computing environments in compliance with the policies of the organization providing the data. Jen described two tiers of restricted data facilities she provides her faculty. One facility is a high-security “cold room” for stand-alone computers without networking of any kind. The cold room meets the security requirements for the National Center of Education Statistics and is primarily for the use of their data. The second facility houses an analytic server that holds restricted data in different partitions. The server is managed in such a way as to make any copying or transmission of data (copying, pasting, uploading, and downloading) impossible. Jen mentioned the analytic server requires intensive I.T. support able to provide a high level of customization and security. For her graduate students, Jen provides a separate data center. This data center is accessed by a card swipe and features stand-alone computers.

Jen helps her faculty process the “mountain of paperwork” necessary to access and use restricted data. Formal agreements are necessary and each agreement is different, even from government agencies that one might think would have relatively similar terms. These agreements require multiple levels of approval, including attention and signatures from her school’s institutional review board (IRB), legal counsel, and the Office of Research.

When Jen works with graduate students, she advises them through this process, but makes the students responsible for actually applying for permission and getting the appropriate signatures. This provides them with practical knowledge and first-hand experience should they end up working for an institution that does not provide assistance.

Jen concluded with advice for librarians to only provide assistance and facilities for the use of restricted data if there is a well-defined need. This is not a matter of “build it and they will come.” She also advised that there are several possible levels of service from the relatively simple task of assisting patrons in finding sources of restricted data, to the more time-consuming process of helping process paperwork, to the expensive and resource-intensive provision of secure computing facilities.

After the presentations, we opened the floor for questions and discussion. The meeting was adjourned after taking two questions from the audience. The audio file of the two speakers’ presentations and Christine’s PPT file are available on the ANSS Sociology Librarian Discussion Group site <http://anssacrl.wordpress.com/discussion-groups/sociology/>.