Apr 9th, 12:00 AM

Developing a Data Management Curriculum for Graduate Students in the Natural Resources

Sarah J. Wright
Cornell University

Camille Andrews
Cornell University
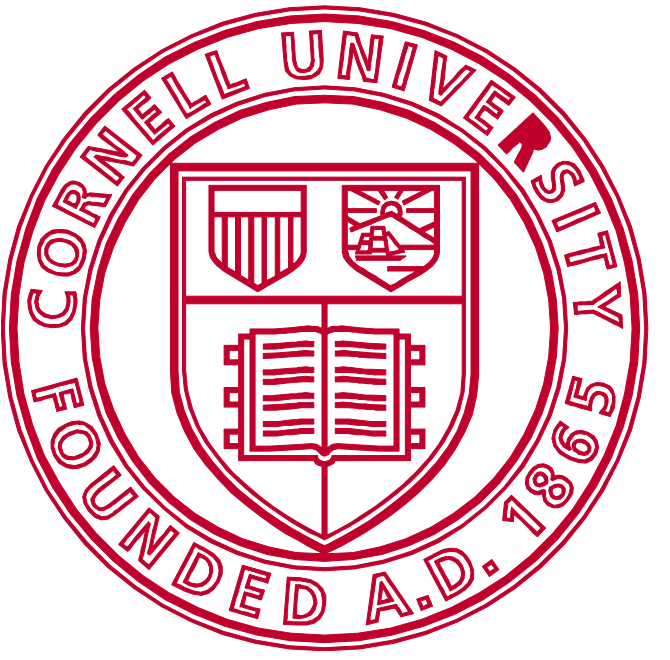
Repository Citation

Developing a data management curriculum for graduate students in the natural resources

Sarah Wright[1], Camille Andrews[1], Cliff Kraft[2]
Contact: Sarah Wright, sjw256@cornell.edu

1: Cornell University Libraries, 2: Cornell University Department of Natural Resources

CORNELL UNIVERSITY · FOUNDED A.D. 1865

Objective

What skills do graduate students need to be successful in managing, working with and curating their research data? This poster reports on work at Albert R. Mann Library to address these needs in collaboration with faculty in the Department of Natural Resources.

Natural Resources/Long-term Studies

http://www.papabearoutdoors.com/about/trout-fishing/

Robinson, J. M., Josephson, D. C., Weidel, B. C., & Kraft, C. E. (2010). Influence of variable interannual summer water temperatures on brook trout growth, consumption, reproduction, and mortality in an unstratified adirondack lake. *Transactions of the American Fisheries Society, 139*(3), 685-699.

Context
❖ Long-term studies tracking longitudinal changes in fish species occurrence, population abundance, growth, and diet

Data types
❖ Spreadsheets (MS Excel),
❖ Relational databases (MS Access),
❖ Scripted analyses in R (open source statistical software)

Educational Priorities

Acquiring the data management and organization skills necessary to work with databases and data formats, document data, and handle accurate data entry is described as essential, otherwise, "it's [as if] the data set doesn't exist."

❖ Data management
❖ Data organization
❖ Data quality and documentation
❖ Data analysis and visualization
❖ Metadata

Response

NTRES 6940 Special Topics Course:
Managing data to facilitate your research

Six session mini-course:
❖ Intro to Data Management
❖ Data Organization
❖ Data Analysis & Visualization
❖ Data Sharing
❖ Data Quality & Documentation
❖ Wrap-up

How would you rate your knowledge/skills/ability in the following areas?

	AFTER taking the class	BEFORE taking the class
Creating a data management plan	~52	~18
Documenting your data	~68	~50
*Evaluating data repositories	~52	* "Before" data not collected
Visualizing data and creating graphs	~71	~67
Following best practices in structuring relational databases	~52	~25
Recognizing necessary components of a data management…	~63	~28
Describing your research and data collection process	~63	~44
Describing the importance of data management	~74	~50

0 = No Competence   0   10   20   30   40   50   60   70   80   **100 = Very Competent**

Assessment Methods

❖ In-class active and collaborative learning

Evaluating Dataset Documentation
Multi-file terrestrial ecology data set:
http://datadryad.org/resource/doi:10.5061/dryad.sp6p0 (http://bit.ly/XiPyCR)
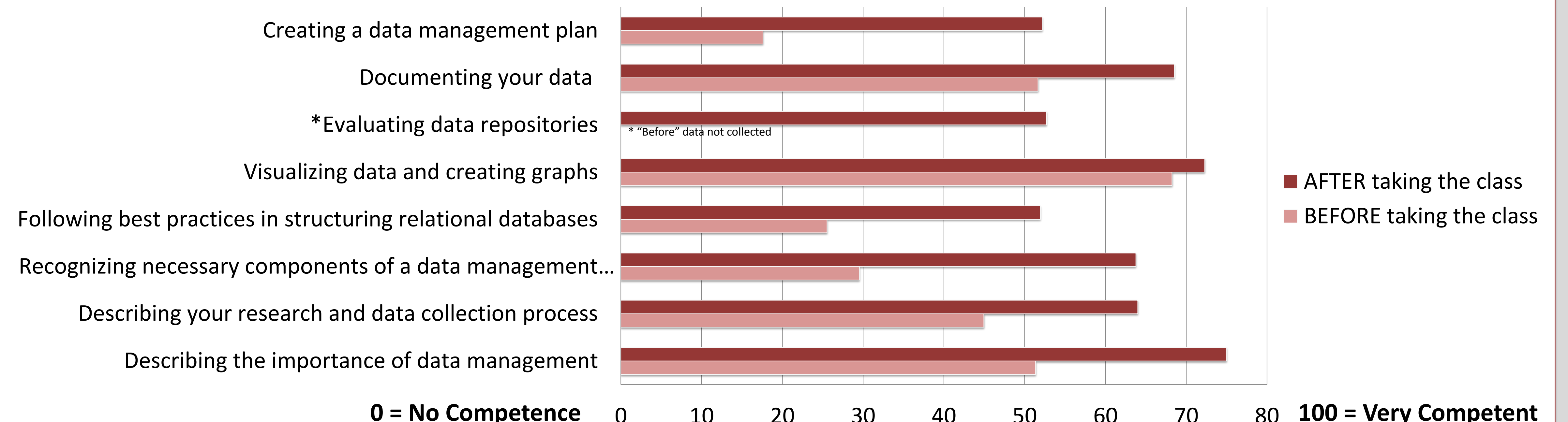NASA readme file:
ftp://disc1.gsfc.nasa.gov/data/toms/README.TOMSL2 (http://1.usa.gov/W8hS7G)

❖ Post class one-minute reflections

"Learning about relational databases was very useful. Efficient organization of spreadsheets was also helpful. I would like to learn more about how to organize metadata, but I think this is an upcoming class discussion. Also, I am still lacking clear reasons why Access is preferable to Excel. What does Access offer that Excel does not? What are the features that make Access particularly useful?"

❖ Final survey