“Data Visualization and Management: the Basics” Workshop Outline

Handout related to the “Integrating Data Management Literacies with Data Visualization Instruction: A One-Shot Workshop” poster (S Konkiel, B Marshall & DE Polley) presented at the Data Information Literacy Symposium, West Lafayette, IN, USA. 23-24 September 2013.

Introduction (30 minutes)

- What is data? What is the data lifecycle? How is it used and reused in research?
  Core competency - data management and organization
- What is data visualization? Why do people use it?
  Core competency - data visualization
- Introduce the "Four Pillars of Data Visualization" (Noah Illinksy): purpose, scope, structure, formatting
  Core competency - data visualization
- Where do we find data? Why do we select certain types of data from certain sources to run particular types of analyses? What are the different analyses?
  Core competencies: Introduction to databases and data formats; data discovery and acquisition
- Why do we work with and save data in open/commonly used formats?
  Core competency: data conversion and interoperability
- Overview of the four data sets we'll be working with and the different tools used to create visualizations

- Humanities data analysis & visualization (30 minutes)
  - Research question: “How were women represented in early 20th c. novels by Indiana authors?”
  - Describe data visualization and dataset’s source, organization and storage medium
    Core competencies - data visualization; data management and organization; discovery and acquisition of data
  - Topical analysis using:
    - Voyant to analyze and visualize the data
      Core competencies - data analysis; data visualization; data ethics [citation]
  - Create README file
    Core competencies: Data Management and Organization; Ethics
  - Save data, visualization, and README files to IU-sanctioned storage media in open/interoperable formats
    Core competencies: data preservation; data conversion and interoperability
- Bibliometric data analysis and visualization from Web of Knowledge/ISI data (30 minutes)
  - Research question: “What national funding agencies bankroll the most research in diabetes? Has that list changed over time?”
  - Describe data visualization and dataset’s source, organization and storage medium
    Core competencies - data visualization; data management and organization; discovery and acquisition of data
  - Topical analysis using:
    - Google Refine to clean data
      Core competency – Quality assurance
    - Sci2 to analyze the data
      Core competency – data analysis
    - Gephi to visualize the data
      Core competencies - data visualization; data ethics [citation]
  - Create README file
    Core competencies – Data Management and Organization; Ethics
  - Save data, visualization, and README files to IU-sanctioned storage media in open/interoperable formats
    Core competencies - data preservation; data conversion and interoperability

- BREAK (10 minutes)
- Social science data visualization and analysis using ICPSR data (30 minutes)
  - Research question: "How have the ways that Americans spend their time changed since the 1950s?"
  - Describe data visualization and dataset's source, organization and storage medium. Spend time explaining the numerical data's codebook.
    Core competencies – metadata; data visualization; data management and organization; discovery and acquisition of data
  - Topical analysis using:
    - Google Refine to clean data
      Core competency - Quality assurance
    - Sci2 to analyze the data
      Core competency – data analysis
    - Gephi to visualize the data
      Core competencies - data visualization; data ethics [citation]
  - Create README file
    Core competencies - Data Management and Organization; Ethics
  - Save data, visualization, and README files to IU-sanctioned storage media in open/interoperable formats
    Core competencies - data preservation; data conversion and interoperability
- Derivative data visualization and analysis from OpenCongress.org data (30 minutes)
  - Research question: "Which Political Action Coalitions have influenced Midwestern senate races the most?"
  - Describe data visualization and dataset's source, organization and storage medium. Spend time explaining how the data was compiled manually from online sources. Describe how Cultures of Practice may influence how data is collected, described, or formatted.
    Core competencies – metadata; data visualization; data management and organization; discovery and acquisition of data; Cultures of Practice
  - Topical analysis using:
    - Google Refine to clean data
      Core competency - Quality assurance
    - Sci2 to analyze the data
      Core competency – data analysis
    - Gephi to visualize the data
      Core competencies - data visualization; data ethics [citation]
  - Create README file
    Core competencies - Data Management and Organization; Ethics
  - Save data, visualization, and README files to IU-sanctioned storage media in open/interoperable formats
    Core competencies - data preservation; data conversion and interoperability
  - Dataset description: talk about how it was compiled manually from online sources. Describe
- Data management recap at end (10 minutes)
  - What planning needs to happen for data to be successfully cared for in ways that enable future reuse?
    Core competency – data curation and reuse
  - What are data management plans? How do they relate to the topics covered today?
    Core competency - Data management and organization