Instructional Design Project: Introduction to Data Discovery

Erika Weir • IS458: Instructional Strategies and Techniques • Spring 2019

Introduction

This lesson is designed to equip undergraduates with basic data literacy skills and prepare them to understand and find datasets in their coursework. The lesson is designed fit into the curriculum of an "Introduction to Quantitative Methods" course in the social sciences; most students are at an intermediate level (Sophomore and Junior) within the Sociology department but the class includes students from all academic levels and majors in the social sciences (Political Science, Economics, Psychology, etc). Students have been assigned to complete a data analysis assignment using open data sets. Students will be introduced to data discovery sources to prepare them to complete their assignment. The lesson is designed to take place as a one-shot instruction session held in the library computer lab so that students can access databases and other resources.

Learning Outcomes

1. Students will be able to define a research question in order to structure their data discovery search.
2. Students will be able to determine what type of quantifiable measures are related to their research question in order to think of potential sources of data.
3. **Critical Thinking:** Students will be able to critically analyze the sources (creating agencies) of data in order to understand any explicit or implicit bias in data collection.
4. Students will able to find relevant and credible datasets to complete their data analysis assignment.

Formative Assessment

Activities:

- **Think, pair, share:** Students will work with a partner to define their research question and brainstorm two to three quantifiable measures that they could analyze. Students will write their quantifiable measures on post-its.
    - **Learning Outcome 1:** Students will be able to define a research question in order to structure their data discovery search.
- **Agency Identification:** Small groups will work together to brainstorm agencies that may collect the types of data students wrote on their post-its and discuss potential biases of those agencies.

    - **Learning Outcome 2:** Students will be able to determine what type of quantifiable measures are related to their research question in order to think of potential sources of data.
- **Class discussion:** Groups will share whether they think biases may affect data. The class will discuss how this may affect their overall search strategy and research project.
    - **Learning Outcome 3:** Students will be able to critically analyze the sources (creating agencies) of data in order to understand any explicit or implicit bias in data collection.

Summative Assessment

Activity:

- Students will find one or two datasets relevant to their research question and email the instructor with both their research question and datasets selected.
    - **Learning Outcome 4:** Students will able to find relevant and credible datasets to complete their data analysis assignment .
    - **Learning Outcome 1:** Students will be able to define a research question in order to structure their data discovery search.

Rubric:

	Excellent	Competent	Developing
Research Question	Student has a defined research question with clear quantifiable measures.	Student has a broad research question that possibly has quantifiable measures.	Student has a broad research question that does not have quantifiable measures.
Datasets	Student has submitted datasets clearly relevant to their research question.	Student has submitted datasets somewhat relevant to their research question.	Student has submitted datasets that are not relevant to their research question.

Outline

I. **Welcome & Icebreaker** (5 min)
   a. Introduce myself and introduces the learning outcomes and an outline of the session
   b. Ask students to introduce themselves and give one "data point" about themselves

II. **Introduction** (6 min)
   a. Instructor shows the following example of poor data use – https://fivethirtyeight.com/features/the-washington-post-misused-the-data-on-violence-against-women/ (1 min)
   b. Asks students to explain what their opinion on how the Washington Post presented this data. (3 min)
   c. If students do not already conclude, reinforce that the Washington Post did not account for potential biases in data collection (1 min)
      i. Married women may be less likely to report violence
   d. Explain that data is a powerful tool and being data literate is extremely important both for their academic work but also their future professions. Although we may be at different stages of our academic careers and therefore understandings of data, data literacy begins with understanding how data was generated and for what purpose. (1 min)
   e. Today, we will work on preparing you to find open data sets in order to explore a research question for your course assignment. However, it is important to always keep in mind when working with data: (1 min)
      i. Why are you choosing the data selected for your specific topic?
      ii. What is the source of the data?
      iii. Are there any potential biases?

III. **Think, Pair, Share** (10 min)
   a. Introduce and explain the activity (1 min)
   b. Students have time to think of a research question (2 min)
   c. Students share their research question with a partner and work with them to identify 2-3 quantifiable variables related to their question a write them on post-its (7 min)

IV. **Agency Identification** (7 min)
   a. (2 min) Instructor introduces the activity and displays some potential ideas for collecting agencies:
      i. U.S. Census
      ii. Twitter
      iii. Environmental Protection Agency
      iv. United Nations
      v. Academic research teams
   b. (5 min) Students pairs join another pair (groups of 4) and brainstorm potential agencies that may collect their quantifiable variables

i. They can move their post-it's into groups if they think that the same agency may collect multiple variables, multiple agencies collect the same data, or if they have similar variables.

V. **Class Discussion** (5 min)
   a. If groups identified multiple agencies that collect the same data, ask them to share whether they think one agency may be more reliable than the other.
   b. Each group shares 2-3 agencies that they identified and if they think there is any potential biases in collection.

VI. **Data Discovery Demonstration** (10 min)
   a. The instructor introduces that often the most reliable sources of data are governmental agencies of academic research teams that are required to follow IRB protocol. Introduces two reliable sources: U.S. Census and ICPSR. However, we should keep in mind that these sources can also have bias. Use the example of the upcoming Census question about citizenship. (2 min)
   b. Leads students through how to use the U.S. Census American FactFinder using the "Guided Search". (8 min)
      i. Ask students for a potential example
      ii. If no example, show how to find data on languages spoken at home for the city of Chicago
      iii. Demonstrate how to narrow by geography
      iv. Pause for questions once initial results are generated
      v. Explain what the year means and difference between 5 year and 1 year estimates
      vi. Show how to export to a CSV file and explain that they will be able to open this in excel or any other data analysis tool that they will use for this project.
      vii. Remind them to give the file a unique name so that they remember what dataset it is, especially if they will be working with multiple datasets.
      viii. Pause again for questions

VII. **Assignment Introduction** (2 min)
   a. Introduce the summative assessment: students will email the instructor their research question, datasets, and any potential collection biases and how they might address them in their project.

VIII. **Guided Work Time** (11 min)
   a. (1 min) Introduce that students will have time to search for datasets, show the "Social Science Data" LibGuide that includes additional databases. Click on the ICPSR site to show students and explain that they will need to set up account to search the database. They can use this time to begin and ask any questions that arise
      i. Pass out handouts that include the assignment instructions, instructor email, and a guide to citing data.
   b. (10 min) work time

IX. **Wrap-up** (4 min)
   a. Thank students for coming and ask if there are lingering questions

b. Mention that the LibGuide to social science data includes additional places to find datasets and the handout includes information on how to cite data.

Discussion

The design of this instruction session was highly informed by the ACRL's "Framework for Information Literacy for Higher Education" and is intended to increase the information literacy of undergraduate students. Within the framework, there are six "frames" librarians can use to support the development of information literacy. Although I attempted to incorporate the ideas of all six, this session more specifically focuses on the "Authority Is Constructed and Contextual" and "Searching as Strategic Exploration" frames. The "Authority Is Constructed and Contextual" frame specifically outlines that students will be able to "define different types of authority" and "determine the credibility of sources". Within my session I attempt to develop these learning outcomes specifically within the context of data use through both the discussion of the Washington Post article (which illustrated a poor understanding of information authority on their part) and summative assessment in which students must identify potential bias in the data they selected for their assignment. The stated learning outcomes for the "Searching as Strategic Exploration" dictate that students should be able to "identify interested parties, such as scholars, organizations, governments, and industries, who might produce information about a topic". I believe this is also reflected in my instructional plan through both the "Agency Identification "activity that I use to help students structure their search strategy and navigate the resources I demonstrate to them later in the session.

Furthermore, this instructional plan also attempts to develop the critical thinking skills of students throughout the session. Although I attempt to develop critical thinking skills throughout all of the activities, I focused on it most closely in developing my second learning outcome: Students will be able to critically analyze the sources (creating agencies) of data in order to understand any explicit or implicit bias in data collection. Using Bloom's taxonomy, I try to move students from first understanding the role of bias in data collection and analysis when it is first introduced in the session through *The Washington Post* example. Students at that time have the opportunity to discuss these concepts and explain it to each other before the instructor summarizes the discussion and reiterates the importance of understanding bias in data. Next, the students have the opportunity to apply these concepts when they are tasked with identify agencies that may collect the data that they are interested in using for their project and identify any potential bias that they may have. Finally, students will use this experience to evaluate the datasets that they choose for their class project. Although they do not have to justify their choice of data for the summative assessment for this session, they most likely will for their course assignment and future projects and will have developed the skills to do so within this session.

Throughout my session, I also attempt to implement the concepts of Universal Design. One of the concepts of Universal Design which I value most in both providing and receiving instruction is that of "simple and intuitive instruction." For this reason, I attempted to ensure that all of the activities are relevant to the student's assignment and when the activity does not appear to be directly relevant, I include time for this instructor to explain the relevance. I also try to ensure that the instruction uses

student-chosen topics by incorporating the student's research topic into all of the activities (except for the first discussion) and allowing students to provide example topics for the database demonstration. Moreover, I also attempt to foster, "a community of learners" by including activities like the "Think-Pair-Share" that allow students to collaborate and also by following up with each student's progress via email after the session. I also attempt to develop the Universal design "instructional climate" by welcoming all of the students with an ice-breaker and setting the same goal for all students that will assist them with their assignment.

Finally, in attempting to develop an instruction session that is culturally responsive to the needs of a very diverse undergraduate student population, I implemented the culturally responsive instructional design recommendations given by Cifuentes and Ozel. In their piece, ""Resources for Attending to the Needs of Multicultural Learners", they recommend that instructors "facilitate motivational strategies by providing students with (a) control over their own learning; (b) relevant contexts for learning skills and concepts; and (c) challenging content that is not too far above their entry skills." In my session, I attempt to put these ideas into practice by allowing students to choose their own research topic and designing each activity around their chosen topic. Additionally, in the introduction to the session I not only stress the relevancy of the session for their assignment and academic goals but I also try to stress relevancy for their career goals, in an attempt to motivate students who are not necessarily academically driven. Cifuentes and Ozel also recommend that instructors vary their design in terms of group work vs. individual work and teacher-oriented vs. student-oriented because students of different cultures respond differently to each style. In my instruction session I attempted to do so by both including some teacher led activities like the database demonstration and the initial introduction. However, I also incorporated student-oriented activities where students have the opportunity to work both individually and in small groups like the Think-Pair-Share and the guided work time.

References

American Library Association. "Framework for Information Literacy for Higher Education." 9 Feb. 2015, www.ala.org/acrl/standards/ilframework.

Armstrong, Patricia. "Bloom's Taxonomy." Vanderbilt University, Center for Teaching, 13 Aug. 2018, cft.vanderbilt.edu/guides-sub-pages/blooms-taxonomy/.

Chodock, Ted, and Elizabeth Dolinger. "Applying Universal Design to Information Literacy: Teaching Students Who Learn Differently at Landmark College." *Reference & User Services Quarterly*, vol. 49, no. 1, 2009, pp. 24–32., doi:10.5860/rusq.49n1.24.

Cifuentes, Lauren, and Serkan Ozel. "Resources for Attending to the Needs of Multicultural Learners." *Knowledge Quest*, vol. 35, no. 2, 2006, pp. 14–21.