

Transforming Research Questions into Variables

A Recipe for Finding Secondary Data

Alicia Kubas, Government Publications and Data Librarian and Regional Depository Coordinator, University of Minnesota, Twin Cities, akubas@umn.edu; Jenny McBurney, Social Science Librarian, University of Minnesota, Twin Cities, jmcburne@umn.edu

NUTRITION INFORMATION

We created this course-integrated workshop to address the increasing need for social sciences students to find and use secondary data in their coursework. This workshop was specifically designed for an undergraduate honors business class, but the model can be adapted for other subject areas that have similar needs: for students to better understand where data come from, how to access data, and how to use data in their own projects.

TARGET AUDIENCE AND NUMBER SERVED

This session is targeted toward upper level undergraduates in courses with a data component. There is no limit on number of students or participants, but for classes over thirty students we recommend a co-instructor to assist with discussion and class facilitation.

LEARNING OUTCOMES

At the end of the session, students will be able to

- translate their research questions into variables in order to determine the scope of their data needs

- identify the organization or type of organization that would collect or produce those data in order to choose appropriate search tools
- articulate how data resources are organized in order to find relevant data using appropriate search strategies

COOKING TIME

One to three hours of preparation time and 45 to 60 minutes of instruction time

DIETARY GUIDELINES

This recipe addresses the frame Searching as Strategic Exploration in ACRL's *Framework for Information Literacy for Higher Education*. Increasingly, we see that students must find secondary data and do more of their own data analysis for their coursework. The foundational data literacy skills students begin to develop in this type of workshop can help prepare them for future courses, as well as for graduate school or their future careers. In this workshop, students learn that they have to be flexible when determining variables from their research questions, creative in their thinking to identify where to find the data they want, and iterative in their overall process as they work through their questions and discover new solutions.

INGREDIENTS

- Slide deck
- LibGuide for data and statistical resources
- Worksheet/assessment tool
- Instructor computer with projector
- Computers for students
- Whiteboards or other brainstorming tools (optional)

PREPARATION

- Create a slide deck that contains examples of research topics, questions, or theses that require a major data component.
- Create a LibGuide or research guide of data and statistical resources, both freely available online and in library subscription databases. Some recommended examples include the US Census website and resources, DataPlanet, and IPUMS from the Minnesota Population Center.
- Create a flowchart worksheet (see figure 1) that visually guides students through the process of breaking down their research question into variables, data sets, and sources. The worksheet can also include a link to the LibGuide and your contact information.

Research Question:

Why do craft brewery clusters develop in certain regions of the US and not in others

Variables		Datasets		Possible Sources
<i>Voter participation (proxy variable for sense of community)</i>	➡	<i>% of voting age population that voted in the last election in Minnesota</i>	➡	<i>- MN Secretary of State - Data-Planet (library database)</i>
<i>Arts, entertainment, and recreation companies per capita (proxy variable for openness to experience)</i>	➡	<i># of companies classified with NAICS code 71 per capita in Minnesota</i>	➡	<i>- Bureau of Labor Statistics - Reference Solutions (formerly ReferenceUSA)</i>
	➡		➡	
	➡		➡	
	➡		➡	

FIGURE 1
Sample flowchart

INSTRUCTIONS

1. *Introduction:* Explain what secondary data are and why data are important in making or supporting an argument. Also provide some grounding expectations when searching for data such as being flexible and adaptable as students may not find the perfect data they want, being prepared to use multiple data sets, understanding that they will likely need to search in many places and resources to find their data, and understanding that persistence and patience are key. Lastly,

- briefly outline the objectives and what you will go over during the instruction session, and note that they'll have time at the end of the session to apply the exercises outlined in the presentation to their own topic and time for their own data searching.
2. *Activity/Instruction:* Introduce the concept of translating a research question into variables and then specific data sets.

Example 1: Start with a simple example such as "price behavior and structure of the organic produce market in the United

States." Ask students what variables they can identify from this research topic. In this case, the variables to guide them to are conventional and organic produce prices for markets in the US.

Next, help them think through what specific data sets will look like for these variables. For this simple case, the variables are easily translated into the data points of specific prices for produce, which can be found in many sources, both freely online and in library databases. Remind students that the data they use may depend on what's available and realistic (e.g., geographic regions and types of produce).

Next, ask students to brainstorm who might collect and disseminate the data—for this example, the US government and specifically the US Department of Agriculture. Then, demonstrate searching Google, as this is a good place to start looking for government data. Don't forget to mention tips for searching online for government data: keep your search terms broad; use the site:.gov filter; and if you find references to the data, try to trace back to the original source.

Example 2: Now move on to a more complex example that requires proxy variables, such as "why craft brewery clusters develop in certain regions of the US and not in others." As with the first example, ask students to identify variables from this research question. This time the variables are not obvious, so let students

brainstorm for a bit. Explain that if you don't know how to identify variables from a topic, you may need to do some market research to learn what factors impact your topic and could be used to identify variables. In this example, we talk about sense of community, openness to experience, and well-being among a region's residents as potential factors. Since these factors are not measurable, ask students what proxy variables they might use. The example list below from a previous student thesis project (see Additional Resources) illustrates some potential proxy measurements for this topic.

Key factor: Sense of community

Potential proxy measurements:
Voter participation;
volunteerism; nonprofit public
charity activity per capita;
clubs, sports, and other social
capital organizations per capita

Key factor: Openness to
experience

Potential proxy measurements:
Recreational drug use; arts,
entertainment, and recreation
companies per capita

Key factor: Well-being

Potential proxy measurements:
Well-being index score

Now, students are ready to identify

specific data sets for those variables. For example, talk students through finding voter participation data by starting with potential sources (government, research organizations, educational institutions, etc.). This is a good time to show students the data and statistical resources LibGuide that contains library subscription databases as well as free online sources.

Last, demonstrate some searches in a few commonly used resources for data and statistics, such as DataPlanet, IPUMS, and Data.census.gov.

3. *Assessment:* Once the instruction portion concludes, students go through this process on their own in a hands-on activity. They choose their own research question and use the flowchart worksheet to demonstrate their understanding of the process, which doubles as an assessment tool.

REVIEWS/ASSESSMENT STRATEGY

We use two strategies for assessing student learning. First, throughout the demonstration portion of the class, students are encouraged to participate by answering questions orally or brainstorming their own ideas relating to the topics on whiteboards. This helps us gauge student learning throughout the session so that we can slow down or speed up the content accordingly. Second, during work time, students fill out a worksheet relating to their own topic. This allows us to quickly scan their sheet for understanding and progress and also start individual conversations with

each student about their questions or problems that arise.

ALLERGY WARNING

Student allergies: Searching for data can be extra challenging for students. Since there is no one-stop shop like a library catalog or go-to database, students typically have to search many different resources and often use data sets from multiple sources in their projects. Remind them that finding data is an iterative process!

Instructor allergies: Data websites and resources are always changing and being updated, especially if you use government sources. Make sure to run practice searches ahead of time so that you know what kind of results to expect!

ADAPTING THE RECIPE

Online instruction: If you will be teaching this workshop online, we recommend finding a colleague who can be an extra set of hands during work time. Organize virtual breakout rooms for students with similar topics so that they can discuss what they are finding and so that you and your colleague can each help multiple students at once. The paper worksheets could be swapped for individual or small-group Google Docs. If you want to have students brainstorm ideas visually, you could use Jamboard or a similar tool instead of a whiteboard.

Large classes: For larger classes, we recom-

mend inviting a colleague along to help during work time, whether you will be teaching in person or online. This way you can each have more one-on-one and small-group interactions with students, and students who are shy or falling behind will have more opportunities to ask questions and get caught up. Plus, co-teaching with a colleague can help your co-teacher learn from you and prepare them for teaching data literacy themselves in the future.

ADDITIONAL RESOURCES

Example LibGuide

Kubas, Alicia. "Comprehensive Guide for Data and Statistics." LibGuide, University of Minnesota Libraries. Accessed January 28, 2021. <https://libguides.umn.edu/dataandstatistics2>.

Example Student Paper

Rosas, Michelle. "A Recipe for Success: Exploring Cluster Formation in the American Craft Brewing Industry." Undergraduate honors thesis, University of Minnesota, 2013. University of Minnesota Digital Conservancy. <https://hdl.handle.net/11299/155308>.