

Archival Assignment

Overview

Where does data come from? In this assignment, you will make a deep dive into the archives to find an **original**, data **primary source** that predates the era of mass digitization.

Part I: Identify and visit a dataset in the archives.

Find a pre-digital **data source** at a Boston-area archive or museum, and look at it. Take careful notes. You **must** go to the archive itself; use of online sources is prohibited, and only with special permission may you visit an object in the archives that already exists online.

What is a pre-digital data source? As we've discussed, in some ways almost **anything** could be seen as data; and data storage takes many different forms. But for the purposes of this assignment, you should choose something that most everyone would agree to align closely with modern quantitative data. That is, they should be structured stores of information that today would be kept in a spreadsheet or database. It is very likely that most good sources will be handwritten, though they may have been made on a typewriter. **Without explicit prior permission, you may *not* use as a source any commercially printed books, or anything that is simply a collection of letters or correspondence.** Some useful keywords to search for on finding aids may be "ledger", "account book", "logbook", "log", "tables." "Pre-digital" means that the data should have been collected without a computer; in most cases, this means it should date before 1980, although there are exceptions.

Take care in choosing a source. Double entry bookkeeping ledgers are common, but can be difficult to understand, for example: and handwriting from before about 1860 can be extremely difficult to read. Don't be afraid to change your source in the archives for a better one.

Be sure to plan your time in advance, and know that most archives are open only Monday-Friday, 9am-5pm. We have left two full weeks for you to make your archival visit, so you should have time to make a visit. If you do not, contact us immediately.

Possible Archives

The Boston area is full of archives, and in doing this assignment you should take the possibility to explore them. The point of this exercise is to explore sources around Boston, so *we discourage you from choosing Northeastern's own archives*. Try to explore more interesting archives or sources; choosing an interesting archive as well as an interesting source will reflect well in your grade.

Some of the major archives within walking distance of Northeastern are:

1. The Boston Public Library (Copley)
2. The Massachusetts Historical Society (the Fenway)
3. The Countway Library for the History of Medicine (Harvard Medical School, Fenway)

Slightly farther afield, but accessible via public transportation, are:

4. The Massachusetts State Archives (Dorchester; Red line.)

5. The National Archives and Records Administration. (Waltham, Bus)
6. Houghton Library (Harvard University, Red line).
7. The MIT Institute Archives & Special Collections. (MIT, red line).
8. The City of Boston Archives (West Roxbury, Commuter rail).

There are also countless smaller archives for particular institutions, from clubs and immigrant societies to political groups to churches. Almost every university in the area has an archive or special collections division. If you have any academic or extracurricular interests, you might be able to find a particularly interesting source. The Appalachian Mountain Club keeps logbooks of everyone who has climbed a mountain in their network. The Boston Symphony Orchestra and the Museum of Fine Arts have their own archives. Churches, synagogues, or mosques may have records about their early membership. You can always e-mail any group you're interested in.

You'll want to contact the archive before you arrive to make sure they have the material on site, and that it's open to the public.

Archival policies on computers, photography, and so forth vary. But ask before snapping pictures.

Part II. Writeup.

You should write up your artifact as a 5-7 page paper. The page length does not include images. 5-7 pages should be about 1600 to 2300 words double-spaced in 12-point font. The point of this is not historical argumentation, but *close and detailed description* that points to the limits of what *you* know about the artifact and what *can be known* about it. Structure your paper in ways appropriate for the artifact.

Item description

In writing up, you should consider including the following elements or addressing the following questions.

1. Who created and stored the information inside? (Was it an individual? A clerk for a larger institution? You should feel free to speculate and admit what you don't know.)
2. Images and/or representations that describe what the data looks like or how confusing elements appear.
3. How is the data organized? What could you learn about the goals, preferences, and worldview of the people who created it from its organization?
4. Are there idiosyncracies in the way the data was collected? Unexpected features? Highlight these and describe why they might take the form they do.
5. Is there something, or are there several things, you *don't* understand about the data? Is something arranged strangely? Is there an abbreviation you don't know? Ideally these questions should be open enough that someone in the class might have an idea: indeed, you should have a few suggestions yourself.

Do **not** include a description of the archive, except insofar as it actually effects the data you're looking at.

Digitization Plan

Outline—but do not implement!—a plan for digitizing the data here into a form that could be used for further research. If you were going to store it in a digital spreadsheet or database, what sort of fields would you collect

information on? Could you store the information in a spreadsheet or database? How much time and effort would it take to create a digital version? What aspects of the document might be lost in your planned transition?

Be sure to address the question of **what use** or **what users** this data might have. What sorts of questions could you answer by having the entire dataset digitized? (Assume that you have all the technical analysis capabilities needed to do so). Address what questions about the past it might answer, what present-day debates it might inform, or what other sources of data might be necessary to make it useful. It's fine to conclude it *shouldn't* be digitized.

Submission

Turn in your paper **as a PDF over blackboard**, by Tuesday February 19 at 5pm.

Grading

Your paper will be graded on the basis of:

1. The apparent effort you have put into finding and examining a distinctive item from an archive. Choosing a particularly interesting source, or a particularly archive to visit, can count on your behalf; conversely, choosing to stay on Northeastern's campus is fine, but no papers about Northeastern materials will receive an 'A' grade.
2. Your ability to describe interesting features of the item you exam, and to raise questions about how the data was collected.
3. The specificity of your description.
4. The thoughtfulness of your digitization plan, and your ability to thoughtfully describe both the benefits and losses of the choices you make for digitizing.
5. Whether your paper successfully addresses the questions about data laid out above.

Late work Late work will be penalized a third of a letter grade each day.

Academic honesty

Should you use any secondary works, you should cite the works that you quote and refer to in the text in a *consistent* format. We recommend the [Chicago documentary note format](#): with it, you give a full citation the first time you use a text, and smaller ones later. For short papers like this, you may omit the final bibliography. If you prefer to use a social-science author-date format with final bibliography, that is also acceptable.

If you are worried about formatting your citations correctly or keeping track of the sources you use, I strongly recommend the open-source citation software [Zotero](#). This will automatically pull citations from the web, and you can drag and drop into a paper to get a formatted citation. Just be aware that online library sources may give you extraneous information, such as the language or a URL. Edit the fields in the library until drag-and-drop gets you good results.

Your most important source for this assignment will be the archival document(s) that you describe: be sure to cite it according to the standards of the archive. This will mean, at a minimum, that you've described it comprehensively enough so that a future researcher could easily find it at the archives themselves.

You should also acknowledge any archivists or peers who help you to better understand the materials. Such acknowledgements would typically come either as a footnote to the first paragraph (for general assistance) or as a footnote to the specific place you received help.