

**Fundamentals of Text Mining:
Analysis Activity
4 - Feeding America: The Historic American Cookbook**

Summary: In this activity we will explore different ways to examine and visualize your datasets using [Voyant](#).

About Voyant

Voyant is a web-based reading and analysis environment for digital texts. It is freely available here: <https://voyant-tools.org>

The tool is open source, and is widely used by Digital Humanists using text analysis and visualization. It is an ideal tool to experiment with in order to familiarize yourself with the process of importing textual data, analyzing it, then creating and interpreting the results as visualizations. One drawback of the tool is that it is hosted on the McGill University servers, and so its ability to process very large datasets is limited. It is possible, however, to install it on a home or local server.

Key features include:

- Importing documents in various formats (plain text, HTML, XML, PDF, RTF, MS Word, ODF, etc.)
- Several tools for studying term frequencies and distributions within documents and within a collection of documents (a corpus)
- A full-text reader that supports very large texts and includes interactive features
- Interaction between the tools that facilitates navigation and exploration at different scales (from "close reading" to "distant reading")
- A mechanism for bookmarking and sharing instances of Voyant Tools (specific texts and tools) through persistent URLs

Instructions:

- **Upload to Voyant**
 - To upload a group of documents, you must first create a zip file of your dataset folder
 - Upload the zip file to Voyant:
 - From the landing page, select the zip folder you have just created.
 - Click 'upload'.
 - Voyant will do the work of expanding the archive and processing all of the documents in your dataset.
- **Understanding the Dashboard View**
 - Familiarize yourself with the dashboard.

- List three pieces of information about your content set that you can see at a glance from the dashboard view.
 - What are your overall impressions of the Voyant dashboard? Do you find it intuitive and user friendly? If not, what do you find unclear or challenging?
 - What is a stopword? [hint: read the 'help' documentation]
 - How would you add a tool to the dashboard that is not included in this default dashboard view?
 - **Voyant Suite of Tools**
 - Voyant provides a range of tools and options for text analysis. What information can you learn from the following tools and visualizations? [hint: Voyant help documentation is useful].
 - Cirrus
 - Document Terms
 - Mandala
 - Contexts
 - Choose your own tool
-

- **Explore your DataSet using Voyant**
 - An opportunity to explore your datasets using the tools embedded in Voyant. The goal is for you to experiment with your data, to customize tool options and to create a visualization or two.
- **Your dataset**
 - In this dataset, you will find the “Feeding America: The Historic American Cookbook” dataset. This collection contains transcribed texts from 76 influential American cookbooks held by Michigan State University Libraries Special Collections. Features encoded within the text include but are not limited to recipes, types of recipes, cooking implements, and ingredients. The 76 texts were chosen among more than 7000 cookbooks that MSU Libraries holds as representative of periods and themes in American cookbook history spanning the late 18th to early 20th century.¹
 - Quality of text is of high fidelity to original cookbook text given transcription of text rather than application of optical character recognition (OCR). Text data is enhanced by encoding of features of text like recipe, recipe type, ingredient, measurements, and cooking implements.
 - More contextual information is available here: <https://d.lib.msu.edu/fa>
 - “A Brief History of American Cookbooks: The history of cookbook publishing in the United States exemplifies the abundance and diversity that have for so long characterized American society. Today, each year, thousands of cookbooks of every size, shape and variety pour forth from American publishing houses but this was not true during the early years of our country. America came late to cookbook publishing. Although we have records that settlers carried cookbooks

¹Feeding America: The Historic American Cookbook Dataset. East Lansing: Michigan State University Libraries Special Collections. <https://www.lib.msu.edu/feedingamericadata/>

with them to the New World and imported cookbooks, especially from England, and kept manuscript receipt books, Amerindians had been living and cooking on this continent for thousands of years, and Europeans for more than a century before the first cookbook was published in America.”²

Thematic Analysis

When analyzing this dataset, you can dive into the thematic trends listed below, or you can view all 76 cookbooks as a whole. Work with your table to decide which pair will work with which topic. Researchers have identified a number of themes in the Feeding America project. We’ll focus on three: *Great Ladies of American Cooking*, *Ethnic Influences*, and *Charity and Church Cookbooks*.

After each pairing has explored their dataset using Voyant, come back together and show the other pairing what you learned.

Great Ladies of American Cooking (contains 27 books)

“Although this term might not be politically correct today, it well defines the women who shaped American cuisine throughout the 19th century and into the first quarter of the 20th: Child, Hale, Randolph, Leslie, Beecher, Harland, Corson, Farmer, Lincoln, Parloa and Rorer.

These remarkable women are well represented in this collection as they must and ought to be. They span the century. Their books went through hundreds of editions and they reached millions of households with their classes, articles and books. Not only were they recognized as culinary authorities, but they were also reformers active in all the major social and cultural events of their day: abolition, temperance, child welfare, women's rights, education, suffrage, social welfare, prison reform, poverty alleviation, immigration, consumer issues, diet, health and nutrition, medical reforms, labor issues, and contemporary religious and moral questions. They shared a major concern for the role of women, for their duties and responsibilities, as well as their rights, and for ways their workload could be lightened and "improved." They were writers, poets, philosophers, educators, editors and business women.

*Their books dominate this collection as their lives and works did.”*³

- **Summary window:**
 - Look for the “Distinctive words (compared to the rest of the corpus)” in the summary window.
 - Find a word that strike you as interesting. Click on that word. What does this do to the “Contexts” window?
 - What insights does exploring the “Contexts” window give you about the texts?

² Longone, Jan. “Introduction to the Feeding America Project.” 2002. <https://d.lib.msu.edu/fa/introduction>

³ Browse the collection by theme: <https://d.lib.msu.edu/fa/interest>

- In the dropdown window for “Contexts” change the view from “document” to “corpus”
- Try typing another “distinctive word” into the search bar for “contexts” (NOTE: Using the * symbol will act as a wildcard). Does this provide any insights about the dataset?
- **Cirrus - Links window:**
 - Now, select a distinctive word used in the previous activity and type them into the “Links” tool
 - What does this data visualization tell you about the word you selected?
 - Now, select two or three distinctive words and put them in the links too. Use a comma after the word to enter multiple terms.
 - How might you export this visualization to a blog, social media, or an academic paper?
- **If you finish early:**
 - find the four-box icon for the upper right panel to select a different visualization tool for that panel. The index of available tools describes what each option can do.

Ethnic Influences (contains 20 books)

*Note: Some of the cookbooks are found in multiple datasets (i.e. some of the Creole cookbooks are also found in the French cookbooks dataset; “Mrs. Rorer's new cook book : a manual of housekeeping” appears in the Asian, Creole, Hispanic, and Jewish datasets).

For this theme, you can view the cookbooks attributed to each ethnic group separately or you can compare all the ethnic groups against one another. To view one specific group, click into the “Ethnic Influences” folder, select all the documents in that folder, and load into Voyant. To view all the group in one visualization, load the zipfile.

“Ethnic recipes and cookbooks appeared in America prior to the Civil War, with British, German and French works predominating. Thereafter, the waves of new immigrants contributed their culinary heritage. The Native American contributions are often unrecorded in the culinary literature, but they are there - always. See Cushing's Zuni Breadstuff for one example of a scholarly dissertation on this subject. African-American influences are also strong, from the earliest days. These are represented both by books written by Blacks (e.g. Mrs. Fisher) and by those written by Southern Whites (e.g. McCulloch-Williams). In addition, Northern authors of more general cookbooks often included African American recipes (e.g. some of the works of Miss Leslie.)

References to Jewish contributions appear even prior to the publication of the first Jewish cookbook in America, which appeared in 1871. "Aunt Babette", Mrs. Kander, the Portland, Oregon Neighborhood Cook Book, and the Greenbaum books all represent the diverse Jewish populations in America - and an international outlook as well.

The contributions of the British, German and French continued throughout the period covered by this project. The Sanderson and Parkinson books are essentially from British sources; the Davidis remains a German and a German-American classic; and the Blot, Tanty, Hearn and Eustis books clearly display the French influence. French influence can also be found in the professional books such as the Ranhofer and the Hirtzler.

*Pennsylvania-Dutch (German) and Quaker cooking are represented by works by Lea and Thomas. The very important Hispanic contributions to the American melting pot are underrepresented in the printed culinary archive. We do, however, include several good examples; chapters can be found in the Los Angeles Times Cook Book and Mrs. Rorer's Cook Book. Many other ethnic contributions are to be found in the MSU collections. To represent them, we have included: Gentile's *The Italian Cook Book*; *Fullstandigaste Svensk-Amerikansk Kokbok* (Swedish); Keoleian's *The Oriental Cook Book (Middle Eastern)*; and Bosse's *Chinese-Japanese Cook Book*."*

- **Summary window:**
 - Look for the “Distinctive words (compared to the rest of the corpus)” in the summary window.
 - Find a word that strike you as interesting. Click on that word. What does this do to the “Contexts” window?
 - What insights does exploring the “Contexts” window give you about the texts?
 - In the dropdown window for “Contexts” change the view from “document” to “corpus”
 - Try typing another “distinctive word” into the search bar for “contexts” (NOTE: Using the * symbol will act as a wildcard). Does this provide any insights about the dataset?
- **Cirrus - Links window:**
 - Now, select a distinctive word used in the previous activity and type them into the “Links” tool
 - What does this data visualization tell you about the word you selected?
 - Now, select two or three distinctive words and put them in the links too. Use a comma after the word to enter multiple terms.
 - How might you export this visualization to a blog, social media, or an academic paper?
- **If you finish early:**
 - find the four-box icon for the upper right panel to select a different visualization tool for that panel. The index of available tools describes what each option can do.

Charity and Church Cookbooks (contains 5 books)

“Cookbooks published to benefit a charitable cause have been with us since the Civil War, They are today ubiquitous and very popular. In the period covered by this project, more than 3000 such books are known. In addition to their intrinsic interest, they often represent regional, ethnic

and other genres. They are a much under-utilized source of information because the earliest ones are very hard to find, ranging, often, from scarce to rare. See especially the Presbyterian Cook Book, the Woman Suffrage Cook Book, Favorite Dishes, the Settlement Cook Book, the Washington Women's Cook Book, and the Neighborhood Cook Book.

Among the most important considerations to be investigated via the charity cookbooks is the involvement of women in the greater community. Most of the charity cookbooks of the period under discussion were church-related. However, to see the variety of causes for which these books were published will give an idea of what social and political issues concerned American women: suffrage, temperance, missionary work, libraries, business and professional organizations, education, welfare, patriotism, granges, hospitals, immigration, and veterans, among others.”⁴

- **Summary window:**
 - Look for the “Distinctive words (compared to the rest of the corpus)” in the summary window.
 - Find a word that strike you as interesting. Click on that word. What does this do to the “Contexts” window?
 - What insights does exploring the “Contexts” window give you about the texts?
 - In the dropdown window for “Contexts” change the view from “document” to “corpus”
 - Try typing another “distinctive word” into the search bar for “contexts” (NOTE: Using the * symbol will act as a wildcard). Does this provide any insights about the dataset?
- **Cirrus - Links window:**
 - Now, select a distinctive word used in the previous activity and type them into the “Links” tool
 - What does this data visualization tell you about the word you selected?
 - Now, select two or three distinctive words and put them in the links too. Use a comma after the word to enter multiple terms.
 - How might you export this visualization to a blog, social media, or an academic paper?
- **If you want to explore more:**
 - find the four-box icon for the upper right panel to select a different visualization tool for that panel. The index of available tools describes what each option can do.

⁴ Browse the collection by theme: <https://d.lib.msu.edu/fa/interest>