

Data Mining and Visualization: distant reading assignment description

Voyant findings presentation: October 4th

Distant reading analysis due: October 11th

“When information overload occurs, pattern recognition is how to determine truth.”

— [Marshall McLuhan](#)

Purpose

Now that we have gone from dipping our toes into the murky DH waters to a deeper understanding and analysis of data and presentation, we turn our heads back to our home discipline of English and look at tools that can further literary understanding, research, and analysis. We will discuss close reading, a common form of close analysis of literature, as well as “distant reading,” a DH term applying to visualizing “structured data” from a set of “unstructured data,” in this case a work of literature. Close reading and distant reading provide different views of texts. Based on your close and distant reading of a story by either Poe or Wharton, you will make an argument about the significance of one pattern you noticed in the text. According to the words being used, what argument can you make about audience, setting, etc.? Be thoughtful in your argument—go beyond the obvious to really investigate what words might be doing and why that’s significant.

Method

We will begin by familiarizing ourselves with Voyant tools in class and the various ways they can parse literary texts. You will choose a short story from The Gutenberg Project. Either:

The Works of Edgar Allan Poe, Raven Edition <https://www.gutenberg.org/ebooks/25525>

or *Tales of Men & Ghosts by Edith Wharton* <https://www.gutenberg.org/ebooks/4514>

There is a worksheet to get you started with questions. In class you will compare your findings, looking for commonalities with your co-scholars (working with the same author). Then you will compare these with the other author group.

Presentation: You will present your initial Voyant findings (materials/data) and what they might mean to the class on **Oct. 4th** before completing your analysis.

Based on feedback and discussion, you will then finish writing your distant reading **analysis essay, due October 11th**. As in any literary analysis (and you have just had analysis practice), your thesis is your interpretation of the text, with a clear, arguable claim, supported by evidence from your text and the Voyant tool(s). You may also address context, include summary, and comment on significance. Most evidence should come directly from your close and distant readings of the text. You will need **MLA format** in-text citations and a works cited page.

Audience

Your instructor and your classmates. Consider someone unfamiliar with Voyant tools for literary analysis. How would you present these findings to a literary (Poe or Wharton) scholar?

Length & Format

Findings presentation: Approximately 5-10 minutes. PowerPoint or Google Slides and/or Voyant links (also MLA format).

Essay: Minimum 3-4 pages (around 1000-1200 words?). Double spaced, 12 pt. font, 1-inch margins, etc. You will need citations for the textbook or other sources (MLA format) and a **works cited page.**

This assignment (including the proposal) is worth **20%** of your final grade.