Data Mining and Visualization: distant reading assignment description
Voyant findings group presentation: October 1st
Distant reading analysis due: October 8th

"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 reading and distant reading of the short story you chose (as well as all your group's stories from your group's author), you will make an argument (interpretation) about the significance of patterns you noticed in the text and the author's work in general. According to the words being used, what argument can you make about audience, setting, etc.? Themes and writing style? 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 your group's author from The Gutenberg Project.

You will first conduct a **close reading** of your short story, including a close look at word usage, themes, and authorial style. Then you will input the entire story into Voyant for your **distant reading analysis**. There is a worksheet to get you started with questions about your text. In class you will compare your findings, looking for commonalities with your co-scholars (working with the same author). By combining these findings and inputting all of the stories into Voyant together, what overall conclusions can you form about your author and their writing style or narrative themes?

Group presentation: As a team, you will present your individual and group Voyant findings (materials/data) and what they might mean to the class on **October 1st** before completing your individual essay analysis.

Based on feedback and discussion, you will then finish writing your distant reading **analysis essay, due October 8th**. 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 should also address **context, summary,** and overall **significance**. Most evidence should come directly from your close and distant readings of the text. You may also include reflection and thoughts on unstructured data in DH, data mining, Voyant tools for literary analysis, etc., and comment on your experience quantifying literature

and conducting this distant reading. 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 (Wharton, Poe, etc.) scholar?

Length & Format

Author group findings presentation: Approximately **10 minutes**. Each group member should add their experience with their own story for the presentation, as well as the group addressing their combined findings. PowerPoint or Google Slides and/or Voyant links/screenshots (also MLA format).

Essay: Minimum 4+ pages (around 1200-1300 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, process work, and group work) is worth **20%** of your final grade.