Exercise 5: Visualizing Literature

Timing. Students can begin the exercise in class, complete it outside of class, and report their conclusions in a subsequent class.

Tools required. RAW Data Visualization Tool, Voyant Visualization Tools.

Purpose. In the final exercise, based on Jockers' discussion of the visualization of literary data, students will use a simple visualization tool to search for patterns in the geographical dispersion of images and themes in poetry. In a large sample of poetry, what correlations can we draw between the content of a poem and the site of its production? Why does a certain kind of poem emerge in a certain time or place? How does literature "travel" in time and space? In order to answer these questions, I provide a ready-made repository of early American gravestone poetry contained in the Farber Gravestone Collection. Students use the visualization tools to map correlations between selected themes in gravestone verse, the locations of the stones, and the years of their production.

Step 1. This exercise depends in part on a corpus of literary data, ready made for students to render in visual form. The parameters of the assignment—exactly what kind of data you ask students to render visually—depends on the nature of the data. I provide students with data derived from one of my own projects, an Excel spreadsheet containing early New England gravestone verse indexed by theme. While you are free to use this data to construct the same visualization exercise that I have, you will likely achieve more meaningful results using data relevant to your own discipline and course. An important preliminary step in this assignment, then, is to make available to students a robust data set to which students can usefully apply the RAW Data Visualization Tool.

Step 2. Once you have your data, introduce your students to RAW, which allows students to render visualizations in a variety of formats. RAW is a simpler alternative to Gephi, the tool that Jockers employs for the visualizations in his chapter on literary influence. Although Gephi is a more robust tool, its technical sophistication will likely prove to be too difficult to exploit in a short classroom module. Begin by asking students to experiment with the sample data sets included in RAW, instructing them to create charts in each of the distinct types, so that they are able to place their own data points into one of the RAW's categories—string, number, and date—and make

informed decisions about the chart types most appropriate to their own projects. Explain the basic method of entering data using comma-separated values, which they can translate from Excel sheets.

Step 3. Guide students in discussion of the spreadsheet containing the gravestone data or the data set that you provide. Given their experimentation with RAW, invite them to propose ways of visualizing this new data in RAW, as well as ways they might use these visualizations to illustrate relations between theme, location, date, or other categories. Ask them, individually or in teams, to form specific research questions that they might answer outside of class by applying RAW to the data provided in class.

Conclusion. While Jockers impressively uses Gephi visualizations to illustrate thematic relations within a large corpus of literature, the parameters of this exercise are more modest. In my own experience, some students struggled to find the best graphical format within RAW to render the correlations in data that they intended to relate. In those cases, I encouraged students to try Voyant Visualization Tools, with which they have already become familiar in previous exercises. Two or three teams of students achieved good results. One correlated the length of epitaphs to the gender of the deceased, raising questions about who merited elaborate gravestone poems. Another team correlated nautical themes in epitaphs to their appearance in coastal communities. Whatever data set you use for the exercise, the general goal is to help students conceive of literary data in quantitative and visual terms, and to use visualizations not only as illustrations of evidence but also as a vehicle to critical insights that might remain invisible in purely textual data.

DLA101.org: Literary Macroanalysis by Harry Brown & The Great Lakes Colleges Association is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.