Exercise 2: Uses of Literary Metadata

Timing. Students can complete the exercise in one class period.

Tools required. Google Ngram Viewer and Google Books.

Purpose. The purpose of the second exercise, based on Jockers' chapter on literary metadata, is to use databases containing book titles, authors, and years of publication, in combination with close reading, to trace the emergence of prominent conversations in a large collection of texts ranging over a long period of time. While the first exercise focuses on interpreting patterns within a single text or small group of texts, this exercise focuses on patterns in metadata for a very large corpus of texts, specifically Google Books. Using Google Ngram Viewer, a simple but versatile tool for tracing the occurrence of specific "ngrams," or combinations of characters, within the Google Books corpus, students will identify and interpret the emergence of a particular idea or related cluster of ideas at specific points in history. My primary case study is the prominence of "apocalypse" in Enlightenment discourse, though students might begin with any word or words with interest for them.

Step 1. Ask the students to choose an intellectually or historically weighted word or cluster of words and enter it into Google Ngram Viewer. Adjust the range of dates and languages as relevant, and use the default "smoothing" of 3. You can encourage them to vary the chronological parameters of the graph in order to examine the patterns from wider or narrower perspectives. For my example, I isolated the occurrences of "apocalypse" in the Google Books corpus between 1700 and 1800.

Step 2. Pose an initial question to students. Given their general knowledge of history, how do they explain the increased interest in a particular idea at a particular time? Students often predict, for example, that books referencing "apocalypse" might appear more frequently in periods characterized by war, revolution, natural disasters, or turns of the century. Spikes in books referencing "apocalypse" in the 1770s and 1790s might support this hypothesis, but how can we explain other prominent spikes in the 1730s and 1750s?

Step 3. Use the capability of Google Ngram Viewer to cross reference Google Books to look more closely at which individual books use your selected word or word cluster

during a specific period. Select "custom range" from the toolbar at the top of the page in order to isolate the chronological range of the spikes in the ngram graph. In this way, students can shift from a panoramic view of literary metadata to textual data, zooming in on a related cluster of books produced at a certain time. In my case study, students see that the increased frequency of books about "apocalypse" in the 1730s and 1750s resulted not from any historical catastrophe but from a flurry of books concerned with interpreting the apocalyptic books of the Bible, beginning with Isaac Newton's *Observations on the Prophecies of Daniel and the Apocalypse of St. John.* Likewise the spike in "apocalypse" in the 1790s appears less related to the French Revolution than it does to another intensification of theological interest in the apocalypse, this time led by Emanuel Swedenborg.

Conclusions. Through this combined examination of literary metadata on the macro scale and textual data on a smaller scale, students can gauge the impact of a certain book or set of books at a certain time, identifying and interpreting significant moments in the history of books, words, and ideas. As in the first exercise, we might use macroanalysis to frame larger critical questions that students might address in class or in writing. Why did eighteenth-century thinkers like Newton and Swedenborg become preoccupied with apocalyptic prophecy? How is the rise of rationalism related to interpretations and critiques of the Bible during this period? How did the Enlightenment change the way we view the end of the world?