HIST 3814 Week Three Notes

Goals:

* Learn the utility of regular expressions for cleaning and organizing an OCR’d document
* Learn ways of further tidying text with Open Refine
* Learn exploratory visualization with network analysis

Listen

* Clean data using REGEX (rather than Excel)
* Soft digital history vs hard digital history
  + Hidden vs visual digital history
  + 80% of dirty data vs 20% that is actually seen
* What is the scholarly value in showing your work? How does this contrast with other history you’ve done to date?

The Soft Digital History that Underpins my Book

* Mapping of data to orient and shape thinking rather than provide “results” to analyse, interpret, and/or present in the book
* Mapping to shape thinking and visualize/understand “how the business of satirical prints worked in late-Georgian Britain”
  + Look at the bigger picture
  + Relationships, interdependencies
* Visual representations of data abstracted from business directories gave him a better sense of what he cared about
* Underpinned a process of source analysis and interpretation that leads towards a central argument
* Don’t include arguments in and of itself
  + Thinking flowed from constructing them
* “soft” digital history as work that doesn’t make it into the final publication—the work that isn’t shown

The Hard Digital History that Underpins my Book

* Hard digital history is digital history “that did make the cut”
* Also called visible digital history
* Substantiated arguments/interpretations
* Outputs of network analysis in the form of data and visualizations
* “Network analysis is the ‘corpus level quantitative’ work”
* Putting location and other information about the texts into a spreadsheet yielded these things:
  + Political satires contained lots of specific locations and London addresses that social satires did not
  + Isaac Cruikshank clearly tailored his work to fit the audiences of his prospective publishers
  + Over the course of his career, Isaac Cruikshank consolidated towards domestic scenes
* Showing patterns and why those patterns might exist
* Results of data processing were used in an inferential rather than explanatory way
  + Help move between close and far, macro and micro, print and print trade, give confidence to make strong statements