L&S 88-2: Literature & Data
Spring 2016
Latimer 105, Tu 4-6

Teddy Roland
troland@berkeley.edu
Office Hours: Th 10-11, BIDS (Doe 190)

Computing

Course Github Repository: https://github.com/data-8/literature-connector
Programming Language: Python 3
(recommend Anaconda installation; https://www.continuum.io/downloads)

Goals & Format

Literature and Data is a disciplinary connector to the main *Foundations of Data Science* course. The main course provides a baseline of programming skills, statistical concepts, and data visualization. In this connector, we will practice each of these and apply them toward problems in the study of literature and the humanities more broadly. As such, students **must** be concurrently enrolled in the main course (CS/Stat/Info C8) or have completed it in the fall (as Stat/CS 94).

The goals of this connector are two-fold. First, we will work consciously to understand how and often what we read as literature, by looking at a wide variety of literary texts and evaluating arguments that critics have made about them. Second, we will explore whether methods from data science make useful additions to our interpretive toolbox, through readings and experimentation. This is exciting and unfamiliar territory for many of us, so the class will explicitly follow two threads: Literary and Critical.

As we get started, the first hour of each class will resemble a typical literature classroom, where we will examine a literary work through close reading and discussion. This will ground our study in traditional aesthetic concerns. The second hour will be devoted to a piece of criticism written about that literary work or author. In the process, we will evaluate the arguments presented by the critic and the forms of evidence they use. In these discussions, we will develop the kinds of critical thinking and expression skills that will be necessary in your future classes and beyond graduation.

Beginning in Week 5 (as the main *Foundations of Data Science* course arrives at material that will become essential for us), we will look at essays whose evidence includes inferential statistics. At that point, we will ask how quantitative data is used alongside traditional forms of evidence, whether it is necessary, and what problems in literature might be appropriate for these methods. The second hour of class will include an examination of code used to recreate the critical work's evidence. This will offer a space to experiment and build programming skills in an applied setting.

Readings

Below is an outline of the texts assigned for each week. All readings will be available to download on bCourses.

Week	Literary	Critical	Assignment
1 *Introduction*	Sophocles, *Antigone*	Moretti, “Operationalizing”	
2 *Close Reading I*	Yeats (selection)	Vendler, “Troubling the Tradition”	
3 *Close Reading II*	Pound, Williams (selection)	Altieri, “The New Realism in Modernist Poetry”	
4 *Making Sense of Close Reading*	Augustine, *Confessions*, Book VIII	Wellmon, “Sacred Reading”	Close Reading Paper – assigned
5 *Ad Hoc Stylistics*	Caedmon's Hymn	Thornbury, “The Poet Alone”	Close Reading Paper – due
6 *Metadata*	Walpole, *The Castle of Otronto*, Chapter I	Moretti, “Style, Inc.”	
7 *Textual Similarity*	Augustine, *Confessions*, Book VIII	Piper, “Novel Devotions”	
8 *Classification*	Rossetti, et al (selection)	Underwood, Sellers, “Literary Standards”	Final Project: consultation
9 *Intro to NLTK*		Wilkens, “Geographic Imagination”	Final Project: proposal
Spring Break			
10 *Word Frequency*		Katsma, “Loudness in the Novel”	Create Github repository; Push data set, preliminary code
11 *Topic Modeling*		Rhody, “Topic Modeling and Figurative Language”	Final Project: consultation
12 *The Great Unread*		Moretti, “The Slaughterhouse of Literature”	Final Project: paragraph overview
13 *The Cutting Edge*		Schmidt, “Rejecting the Gender Binary”	
14	Final Presentations		Slides due; Push final code

Requirements

Readings & Presentations

All readings must be completed before class.

Once during the semester, each student will be required to make a brief (5 minute) presentation on the week's critical reading that will initiate our discussion. This presentation should offer a summary of the article, including any context that may help us to understand its concerns, and describe some of the problems it explores. The presentation should conclude by raising a few questions that will spur our discussion.

Additionally, once during the semester, each student will be required to be the first respondent. Minimally, the presenter will email the respondent (and me) their discussion questions the evening before class by 5 pm, so that the respondent has time to form a thoughtful, 2-minute response. The presenter and respondent may wish to coordinate beforehand but this is not required.

Participation

Please prepare to speak at least once during discussion, each class. Your voice is valuable and your perspective unique.

Close Reading Paper

The first paper assigned will be a traditional, literature-class paper. You will make an interpretive argument based on a close reading of a text. This paper will be written on one of the literary texts we have read in the first four weeks of class, and I will offer an optional prompt for each. The paper should be 2-3 pages in length.

Final Project

The course is built around the final project (which replaces the final exam). This consists of a 4-5 page paper in which an argument is made about a literary text(s) using evidence from both inferential statistics and close reading. This paper must examine an interpretive problem and may be written on any literary text you choose.

In preparation for the final paper, students will be required to fulfill several milestones. During Week 8, students will meet with me outside of class to consult on literary texts, interpretive problems, and statistical methods of interest. The following week, students will send me a one-sentence proposal for their final project including these three elements.

We will meet again during Week 11 to discuss progress and obstacles in the project, as well as any findings. The following week, students will turn in a paragraph, describing their methods and statistical findings, including one visualization.

In keeping with the best practices of the field, students will be required to make available their data set (pending copyright) and code through Github. Preliminary code will be posted during Week 10 and final code – capable of reproducing your findings – before our last class.

During our final class, students will deliver a 3-5 minute presentation describing the literary problem being explored and any decisions made or roadblocks faced while applying statistical methods in literature. This will act a kind of rough draft for the paper, as well as offer an opportunity for feedback from your peers.

Grading Rubric

Participation

Discussion	25%
Critical Presentation	5%
First Respondent	5%

Projects

Close Reading Paper	15%
Project Milestones	25%
Consultations	
Proposal/Paragraph	
Code & Data	
Final Presentation	
Final Paper	25%