

Quantitative Textual Analysis - Tufts University, Fall 2025

Details

- Time: Th 6–9pm
- Location: Eaton Hall 211
- Instructor: Charles Pletcher (charles.pletcher@tufts.edu)
- Instructor Office Hours: TTh 3–4pm
- Teaching Assistant: Zoë Spriggs (zoe.spriggs@tufts.edu)
- TA Office Hours: TBD

Prerequisites

None

Description

Despite the recent proliferation in digital sources for a variety of literary fields and subfields, quantitative textual analysis has often been viewed as anathema to pursuits in the humanities. Rather than set computational and literary methods at odds, this course seeks to reconcile them through careful application of statistical methodologies alongside literary modes of inquiry. Far from a positivistic approach to literary texts, this course will guide students towards enriching their understanding of the texts that they study by taking a “distant reading” approach (cf. Moretti 2000) that complements, rather than supplants, close reading and critical analysis.

This course introduces humanists to the tools and methodologies of quantitative textual analysis through corpus linguistics. Using the R programming language, students will learn how to build and evaluate corpora related to their areas of expertise, and they will gain experience with basic statistics and probability theory, hypothesis testing and experiment design, and methods in sociolinguistics, stylistics, and diachronic textual analysis. Students will gain substantial practical experience through weekly labs and homework assignments, culminating in a final presentation and paper aimed at publication.

NB: Although programming is a major part of this course, the focus is on corpus linguistics and statistical methods, not on programming per se.

Learning objectives

This course will prepare students to:

1. Build textual corpora from a variety of sources
2. Analyze textual corpora using statistical methods and the R programming language
3. Test and evaluate hypotheses based on these analyses
4. Apply these techniques to works within their fields of expertise
5. Work with embeddings and language models to answer specific research questions

Required Textbook

Stoltz, Dustin S. and Marshall A. Taylor. 2020. *Mapping Texts: Computational Text Analysis for the Social Sciences*. Oxford: Oxford University Press.

Schedule

See the file located at [schedule.md](#) for the most up-to-date list of assignments.

Grading

- 1/6 Attendance and participation
- 1/6 Reading quizzes

- 1/3 Labs
- 1/3 Final project

This class will use the following grading scale, after applying the above weights:

- $\geq 90\%$: A
- 80–89%: B
- 70–79%: C
- 60–69%: D
- $< 60\%$: F

Plusses and minuses will be appended to letter grades at the instructor's discretion.

1/6 Attendance and participation

As we only meet once a week, missing even one class could leave you far behind. Although much of the material will be posted online, class time will be used to troubleshoot, debug, and debrief (read: “vent about how hard programming is”).

Obviously emergencies happen, and everyone is allowed one (1) unexcused absence. Further unexcused absences will result in a deduction of half a letter grade *from your final grade*.

Excused absences are easy to obtain: please just give me as much notice as possible. (I try to keep track of religious holidays, but a heads-up is always appreciated.)

1/6 Reading quizzes

Starting in Week 2, we will have a reading quiz approximately every other class. These quizzes are not meant to be punitive. Rather, think of them as 10 minutes or so to reflect, as cogently as you can, on specific aspects of the reading.

At the end of the semester, I will drop the lowest one (for graduate students) or two (for undergraduates) reading quiz scores when calculating your average for the term.

1/3 Labs and midterm projects

We will start a lab in class approximately every other week; the lab will be due at 11:59 p.m. following the next class.

Each lab will also ask you to write a short but detailed reflection on your process and what you have learned. These short essays are meant to help you as you decide on your midterm and final projects.

Think of the midterm project as the final project in miniature. This is a chance for you to experiment with a methodology or take the next logical steps with a lab. If it works, take the project even further for the final; if it doesn't, the final project gives you a mulligan. **Null results are valid here.**

1/3 Final project

The final project is a chance for you to demonstrate your mastery over some aspect of digital humanities that we have covered during the semester. Start thinking early about what experiments you might want to run or what tools/resources you might want to build for your final project, and feel free to send me an email or come to office hours as soon as you have ideas.

Statement on AI

(Adapted from “Artificial Intelligence” by the Tufts’ Center for the Enhancement of Learning and Teaching)

This course discourages the use of Generative Artificial Intelligence (GAI) for translation assignments. At the same time, the course also provides opportunities to explore the use of GAI in supplementing or augmenting our abilities to engage with and understand Homeric song.

To this end, we will adhere to the following guidelines:

1. **NEVER** submit AI-generated work as your own. To do so is considered a violation of Tufts' [Academic Integrity Policy] (<https://students.tufts.edu/community-standards/academic-integrity/academic-integrity-overview>).
2. Familiarize yourself with the limitations of GAI tools. These tools all come with inherent biases, and they all make mistakes. Be aware that no matter how "confident" the AI seems, it does not actually "know" anything.
3. Cite all uses of AI. See suggestions from the [Chicago Manual of Style] (<https://www.chicagomanualofstyle.org/qanda/data/faq/topics/Documentation/faq0422.html>) and the MLA.
4. In addition to citation, please identify *how* the GAI contributed to your work. (Your explanation need not be more than a sentence or two.)
5. If have any questions, please contact me by email, during office hours, or in class.

Policies

Attendance

Emergencies arise. Everyone is permitted one unexcused absence. If you will need to miss additional classes, please let me know as soon as possible.

Each subsequent unexcused absence will result in a deduction of half a letter grade from your final grade (i.e., an A will become an A-, an A- will become a B+, etc.).

You are responsible for catching up on any work that you have missed, either by coming to office hours or working with a classmate. As a rule, I will not share lecture notes outside of what is already available on the course website.

Food and drink

Please eat only during breaks. (We'll have one about halfway through each class.)

Otherwise, please stay hydrated.

Sharing

This course is designed for everyone to feel comfortable participating in discussion, asking questions, learning, and facilitating the learning of others. In order for that atmosphere to be maintained, the recordings of our conversations will only be shared with the enrolled students in the class (not posted publicly) and it is prohibited for any of us who have access to the video to share it outside the course. Similarly, I have specifically designed the exams, handouts, and lectures for the people who are enrolled in the course this term and those may not be shared outside this course. All of this content is freely available on GitHub under a Creative Commons license — attribution is all that is required.

Resources

Religious accommodations

Tufts University faculty, staff, and administration highly value and acknowledge the religious diversity of its student body. Students seeking religious accommodations related to their holy days are encouraged to collaborate with faculty to make arrangements during the first week of each semester. Consult the [Multifaith Calendar] (<https://chaplaincy.tufts.edu/multifaith-calendar/>) for upcoming holidays, links to the University Religious Accommodations Policy, and members of the University Chaplaincy who are available to respond to questions on religious observances.

Accommodations for students with disabilities

Tufts is committed to providing equal access and support to all qualified students through the provision of reasonable accommodations. If you have a disability that requires reasonable accommodations, contact the StAAR Center at StaarCenter@tufts.edu or 617-627-4539. Please be aware that accommodations cannot be enacted retroactively, making timeliness a critical aspect for their provision.

Academic support at the StAAR Center

The StAAR Center offers a variety of FREE resources to all students. Students may make an appointment to work on any writing-related project or assignment, attend subject tutoring in a variety of disciplines, or meet with an academic coach to hone skills like time management and navigating procrastination. Students can make an appointment for any of these services by visiting <https://students.tufts.edu/staar-center>.

Student support, including mental health

As a student, there may be times when personal stressors or difficulties interfere with your academic performance or well-being. The Dean of Student Affairs Office offers support and care to undergraduates and graduate students who are experiencing difficulties, and can also aid faculty in their work with students. In addition, through Tufts' Counseling and Mental Health Service (CMHS) students can access mental health support 24/7, and they can provide information on additional resources. CMHS also provides confidential consultation, brief counseling, and urgent care at no cost for all Tufts undergraduates as well as for graduate students who have paid the student health fee. To make an appointment, call 617-627-3360. Please visit the CMHS website: <http://go.tufts.edu/Counseling> to learn more about their services and resources.

Quantitative Textual Analysis – Fall 2025 – Course Schedule

As a friendly reminder, you should never use generative AI (e.g., Claude, ChatGPT, vel sim.) to write your lab report for you, even if the lab requires you to use a language model in your analysis.

These lab reports are spaces for you to think. Out-sourcing your thoughts to a language model deprives you of that opportunity and prevents you from learning the material.

Using generative AI to write your lab report will be considered a serious violation of the Tufts University Honor Code.

1. September 4

- Introductions
- Syllabus
- *Mapping Texts* chs. 1 and 2

Homework

- Review *Mapping Texts* chs. 1 and 2
- Read *Mapping Texts* chs. 3 and 4

2. September 11

- Reading quiz
- *Mapping Texts* chs. 3 and 4
- IDE setup and debugging

Homework

- Finish setting up your IDE
- Read *Mapping Texts* ch. 5

3. September 18

- Reading quiz
- *Mapping Texts* ch. 5
 - Discussion: What corpora should we build? Working in groups of 3 or 4, decide on what kind of corpus you want to build. Where can you find enough data to answer your research questions? What techniques will you need to use to acquire text?
 - Techniques:
 - * OCR
 - * Audio transcription
 - * APIs
 - * Web scraping
- Corpus assembly workshop: Having determined where you will need to look to find textual data, go get it!

Homework

- Finish corpus assembly workshop
- Read *Mapping Texts* ch. 6

4. September 25

- *Mapping Texts* ch. 6

Homework

- Lab 1
- Read *Mapping Texts* ch. 7

5. October 2

Lab 1 due by 11:59 p.m.

- Reading quiz
- *Mapping Texts* ch. 7

Homework

- Finish in-class text-wrangling workshop
- Read *Mapping Texts* ch. 8

6. October 9

- *Mapping Texts* ch. 8

Homework

- Lab 2
- Read *Mapping Texts* ch. 9

7. October 16

Lab 2 due by 11:59 p.m.

- Reading quiz
- *Mapping Texts* ch. 9

Homework

- Finish in-class deductive workshop
- Read *Mapping Texts* ch. 10

8. October 23

- *Mapping Texts* ch. 10

Homework

- Lab 3
- Read *Mapping Texts* ch. 11

9. October 30 🍊

Lab 3 due by 11:59 p.m.

- Reading quiz
- *Mapping Texts* ch. 11

Homework

- Finish word embeddings workshop
- Read *Mapping Texts* ch. 12

10. November 6

- *Mapping Texts* ch. 12

Homework

- Lab 4
- Reading *Mapping Texts* ch. 13

11. November 13

Lab 4 due by 11:59 p.m.

- *Mapping Texts* ch. 13
- Final project planning

Homework

Continue planning, discussing, and researching final projects.

12. November 20

Meet with Charles to discuss final project plans before Thanksgiving!

- Independent time to work on final projects.

13. December 4

- Final project presentations