

INFO 490CA/690CA: Introduction to Social and Cultural Analytics

Instructor: Laure Thompson, laurejt@cs.umass.edu

Course Description. This course introduces the use of computational methods for studying culture and society. We will learn how to transform textual and visual material into data, and how we can explore and analyze this data using machine learning and statistical models as both quantitative and qualitative tools. We will also survey and discuss how these methods are applied within the humanities and social sciences, as well as reflect on the challenges, limitations, and ethical issues that arise in the computational study of culture and society. 3 Credits.

Course Prerequisites.

INFO 490CA: INFO 248 or Permission of Instructor.

INFO 690CA: None. Basic experience with Python, introductory data science, or other introductory programming experience is recommended but not required. All required programming and statistical concepts will be introduced in this course.

Course Goals. Through this course students will:

- become familiar with computer-assisted methods for studying culture and society;
- be able to construct nuanced arguments using data and computational methods;
- gain an understanding of when a computational method is both applicable and useful (and when it's not), as well as how to critically interpret its output;
- develop proficiency in using Unix commands and Python for analyzing social and cultural data;
- be able to conduct computer-assisted scholarship in conjunction with faculty.

Course Materials. This course has no required textbook. Readings for this course will largely be available online and will otherwise be made available via Blackboard.

Assignments & Grading.

Grading Scale. F = <63.49, D = 64-66, D+ = 67-69; C- = 70-73, C = 74-76, C+ = 77-79; B- = 80-83, B = 84-86, B+ = 87-89; A- = 90-93, A = >93.49

Assignment Categories & Weights.

- Discussion Board – 20%
- Homework Assignments – 40%
- Final Project – 30%
- Course Reflection – 10%

Discussion Board.

For each week's reading assignments, you will need to post a response guided by a short prompt on some of the readings. Additionally, you are required to reply to two of

your classmates' posts.

Homework Assignments.

For each week's topic there will be a homework assignment containing a mixture of coding and short written responses (the ratio will vary by week).

Final Project.

You will complete a final project on a topic and dataset of your choosing. In this project you will (1) explore and analyze a collection and (2) operationalize a research question using computational methods covered in this course. This project will be broken into several components including an exploratory data analysis, progress report, final presentation, and final report. Students in 490CA will work in groups of 2-3. Students in 690CA can work alone or in groups of 2-3 690CA students with expectations for groups raised proportionately.

Course Reflection.

In lieu of a final exam, you will write a reflection on your learning and experiences in this course.

Graduate Section.

Students enrolled in 690CA will have an expanded set of expectations:

- *Discussion Posts.* You will have an expanded role in discussion posts. You will also create a small number of posts through the semester that summarize research papers/projects in social and cultural analytics (broadly construed) suitable for the whole class. The works described will be a mixture of suggested readings and works connected with your primary research field and interests.
- *Final Project.* Your final project is expected to have wider scope and culminate in a research term paper that is more ambitious and in-depth than the corresponding field report for 490CA groups.
- *Supplementary Materials.* While all programming and statistical concepts are introduced in class, additional materials that reinforce these concepts will be provided.
- *Separate Meetings.* An occasional separate meeting of 690CA students will be held. These meetings will be used to field questions about course concepts and research projects as well as deeper discussion of course readings.

Example Schedule. Topics are subject to change based on class interests and projects.

Week 1: Introduction

Readings:

- boyd & Crawford (2011). Critical question for big data
- Underwood (2015). Seven ways humanists are using computers to understand text
- Grimmer (2015). We Are All Social Scientists Now

Week 2: Text to Data: Text Encoding & Tokenizing

Readings:

- Zentgraf (2015). What Every Programmer Absolutely, Positively Needs To Know About Encodings And Character Sets To Work With Text
- Dombrowski (2020). Preparing Non-English Texts for Computational Analysis
- Dimson (2015). Emojineering Part 1

Week 3: Text as Data: Counting & Sentiment Analysis

Readings:

- Klein & D'Ignazio (2020). "What Gets Counted Counts" from *Data Feminism*
- LaFrance (2016). "The Six Main Arcs in Storytelling, as Identified by an A.I." from *The Atlantic*

Week 4: Similarity and Divergence

Week 5: Images as Data & Vector Spaces

Week 6: Reexamining Data

Readings:

- Krause (2017). Data Biographies
- Gebru et al. (2018). Datasheets for Datasets

Week 7: Classification

Week 8: Clustering

Due: Final Project – Exploratory Data Analysis

Week 9: Text Analysis – Topic Modeling

Week 10: Text Analysis – Named Entity Recognition and Part-of-Speech Tagging

Due: Final Project – Progress Report

Week 11: Making Arguments

Week 12: Validation

Week 13: Final Project Presentations

Finals Week

Due: Final Project – Final Report

Due: Course Reflection

Policies.

Accommodation Statement. The University of Massachusetts Amherst is committed to providing an equal educational opportunity for all students. If you have a documented physical, psychological, or learning disability on file with Disability Services (DS), you may be eligible for reasonable academic accommodations to help you succeed in this course. If you have a documented disability that requires an accommodation, please notify me within the first two weeks of the semester so that we may make appropriate arrangements. For further information, please visit Disability Services (<https://www.umass.edu/disability/>)

Academic Honesty Statement.

Since the integrity of the academic enterprise of any institution of higher education requires honesty in scholarship and research, academic honesty is required of all students at the University of Massachusetts Amherst. Academic dishonesty is prohibited in all programs of the University. Academic dishonesty includes but is not limited to:

cheating, fabrication, plagiarism, and facilitating dishonesty. Appropriate sanctions may be imposed on any student who has committed an act of academic dishonesty. Instructors should take reasonable steps to address academic misconduct. Any person who has reason to believe that a student has committed academic dishonesty should bring such information to the attention of the appropriate course instructor as soon as possible. Instances of academic dishonesty not related to a specific course should be brought to the attention of the appropriate department Head or Chair. Since students are expected to be familiar with this policy and the commonly accepted standards of academic integrity, ignorance of such standards is not normally sufficient evidence of lack of intent (http://www.umass.edu/dean_students/codeofconduct/acadhonesty/).

Inclusion Policy. In this course, each voice has something of value to contribute. Please take care to respect the different experiences, beliefs, and values expressed by students and staff involved in this course. We support UMass Amherst's commitment to diversity, and welcome individuals of all ages, backgrounds, citizenships, disability, sex, education, ethnicities, family statuses, genders, gender identities, geographical locations, languages, military experience, political views, races, religions, sexual orientations, socioeconomic statuses, and work experiences.

I reserve the right to modify this syllabus as needed to account for current events and to better support student learning.

Acknowledgements.

This course and syllabus are inspired by those of David Bamman, Lauren Klein, David Mimno, Michelle Trim, and Melanie Walsh.