

Advanced Data-Driven Storytelling

DACSS 604

University of Massachusetts Amherst FALL
2025

Course Time	TuTh 5:30 PM - 6:45 PM Sep 4, 2024 - Dec 9, 2024	Multimodal Course Venue	In-person: Machmer W-13 Online: recorded lectures
--------------------	---	--------------------------------	--

Instructor: [Erico Yu\(him/his\)](#)

Class Assistant:

Office 256 Bartlett (Meeting Room) **Office Hours:** Mo 4 pm – 6 pm EST
We 5 pm – 6 pm EST
And by appointment

Appointment Link: [Calendly](#)

Course Canvas

Campus Resources:



Graduate Student Services



DACSS



Disability & Accessibility

[1. Course Description]

How can data analysts and computational social scientists convey data through narrative and reports geared toward general audiences or specific stakeholders? How can they convey that data through visuals geared toward non-scientists? This hands-on course provides students with the knowledge and skills needed to generate strong, data-driven communication. We will emphasize the clear interpretation of data patterns, quantitative analyses, and statistical inferences, using precise and accessible language tailored to the audience.

[2. Learning Objectives]

- Deepen understanding of core concepts, theoretical foundations, and principles on which

data analysis and computational/statistical social science are built

- Sharpen appropriate usage of data analytic vocabulary and avoid common misinterpretation
- Communicate clearly with the aid of tables and charts, interpreting these visual aids in ways that advance the data-driven story and avoid misleading the audience

Improve oral and written forms of communication for audiences with different levels of technical background and substantive knowledge

- Develop skills at fluid writing that incorporates meaningful information gleaned from data, using rigorous, scientifically defensible models and methods
- Avoid misleading claims based on data

[3. Course Materials]

You don't need to buy any textbooks. All required and recommended materials, including readings and videos, will be posted on the course website as we progress through the semester.

The primary textbook (1st edition) can be purchased for just a few dollars. We will also read a few selected pages and sections of the latest version (2nd edition).

Miller, J. E. (2005). [The Chicago guide to writing about multivariate analysis](#). University of Chicago Press. [Miller, hereafter]

We will also be reading selected chapters or pages of the following four books:

Spencer, Scott (2024). [Data in Wonderland](#). [Spencer, hereafter]

Cairo, Alberto (2012). *Functional Art, The: An introduction to information graphics and visualization (Voices That Matter)*, New Riders, [Cairo, hereafter]

Bertin, Jacques (2011). *Semiology of Graphics: Diagrams, Networks, Maps*, translated by Berg, William J. ESRI Press (selected pages will be available on the course website) [Bertin, hereafter]

Healy, K. (2018). [Data visualization: a practical introduction](#). Princeton University Press. [Healy, hereafter]

There are other recommended texts you may find useful for this class. I will update the list as the class proceeds.

Nolan, D., & Stoudt, S. (2021). *Communicating with data: The art of writing for data science*. Oxford University Press.

In the *Resources* section on Canvas, you'll find a **collection of video talks and presentations** by data scientists from various fields, sharing their perspectives on effective data-driven storytelling. One of

the playlists I like the most is the “[**Making sense of too much data**](#)”, presented by TED. These videos are **not required readings**, but I strongly encourage you to explore them for additional insights beyond our classroom discussions. They offer real-world examples of how storytelling with data is applied across domains.

[4. Course Format, Website, and Technology]

Course Format:

This course combines **lecture** and **graduate seminar** formats. Most weeks will include a **lecture component**, where I introduce key concepts, tools, and examples related to data-driven storytelling. In many weeks, I will also provide discussion prompts to encourage your engagement — either during in-person class sessions or afterward through online and in-person channels.

The seminar component involves active student participation: we will discuss assigned readings, conduct peer reviews of your individual projects, and brainstorm ideas together. You are expected to come prepared and contribute thoughtfully during both the lecture and seminar portions of the course.

Course Website/LMS:

All classroom material will be posted on Canvas. You will also submit assignments through Canvas.

Technology/Software

This class will be software-independent for the most part. The focus will be on effective communication using whatever tools you may wield, whether R, Python, SPSS, Stata, or more widely employed platforms such as Excel. Since the course focuses on honing your communication skills, it is expected that you will continue to acquire skills on your preferred platforms or explore new ones.

In the latter part of the semester, we will be working on two platforms: GitHub (for creating and publishing websites) and Tableau (for generating visualizations).

[5. Course Assessment and Feedback]

Grades are calculated as follows:

- **Assignment (50%)** – There will be five assignments throughout the semester.
 - There will be four regular assignments (revised from the past semester);
 - The 5th assignment is a newly designed assignment: related to Communicating to Different Audiences
- **Final Project (30%)**
 - **Check-in#1: Due Oct 31**

- Final Product: Due Dec 19
- **Presentation on the Final Project (10%),** Dec 9, 2025
 - For students who are graduating this Fall: you are encouraged to give a 15-minute presentation at the End-of-the-Semester Event; more information will be provided
- **(A)Synchronous Participation (10%):** in-class and online discussion

A grade penalty will occur for any unnotified late assignments: each 24-hour period past the due date, a 10% penalty will be applied to your assignment grade. If you require an extension for any upcoming assignments, please communicate with me and request approval before the deadline.

Final Project Information regarding the final project will be made available on the course website, but in short, students may choose among a few options, including revisions of a previous project or a new project. Some other options include work done for a client or a portfolio of others' work to be critiqued and edited by the student. Further instructions for the final project and past student examples will be provided later in the semester.

Participation: This course is designed to be 70% lecture and 30% seminar, and participation and discussion are highly encouraged. We have found that students who actively participate in class discussions—whether synchronously (when possible) or asynchronously—gain significantly more from the course and learn new skills at an accelerated rate. Students who participate regularly, either synchronously or asynchronously, will receive a “boost” to their grade. Participation does not require expertise; instead, asking and answering questions are both equally valued forms of involvement. Raising points for deeper discussion is particularly helpful.

Final letter grades are assigned using the University's Plus-Minus Grading Scale according to following rubric:

- A (94-100%)
- A- (90-93%)
- B+ (86-89%)
- B (81-85%)
- B- (77-80%)
- C+ (74-76%)
- C (70-73%)
- F (Below 70%)

[6. Course Policies]

Office Hours & Emails

My regular office hours are 4 pm – 6 pm EST on Monday and 5 pm – 6 pm EST on Wednesday. You can schedule a meeting with me through [Calendly](#), and you can also email me with any questions related to the course and your projects. I try to respond to all emails and phone messages within 24 hours during weekdays. However, I generally do not check and respond to emails between 4 pm on Friday and 10 am on Monday. Also, please expect a delay in responding to evening emails after 6 pm.

Collaboration and Academic Integrity

I support collaboration and encourage you to work together with your peers in offering feedback on assignments. You are also encouraged to cite class discussions, conversations with peers, posted notes, and any other material prepared by your classmates. But all written work must be your own.

Ensure that you clarify and acknowledge collaboration with your peers, accurately represent your own contributions, and properly cite all sources. Any suspected misrepresentation of your own original contributions—even if the result of carelessness—will be brought to the attention of me, DACSS, or the Academic Honesty Office and may result in a failing grade for the course.

For more information on the policy regarding academic honesty and dishonesty, please refer to Part 7.

Special Note on using AI Tools:

Today, almost everyone relies on gen-AI to assist with coding. However, in this course, I ask you to limit the use of gen-AI, especially for coding assignments. We are learning a programming language, R, and it is crucial to understand the basic elements, syntax, and logic behind it. Writing code yourself helps you recognize where mistakes are likely to occur, what the limitations of R are, and what your own coding habits look like. If you turn too quickly to gen-AI, or rely on it too heavily, you may find yourself unprepared when facing more challenging tasks that require customization, fine-tuning, or larger, more advanced projects.

Although UMass does not yet have a formal policy on the use of AI tools, I have specific expectations for this course.

- **For coding assignments**, write the code yourself. Your first submission does not need to be perfect—you will have the chance to revise and resubmit up to three assignments later in the semester. If you use AI assistance at any point, you must cite it clearly (which tool, what prompt, and what output).
- **For essay-writing assignments** (summarizing or responding to a reading, describing a dataset, or explaining a visualization), the use of AI tools is strictly prohibited, and we will pay particular attention to your work in these cases.
- If we review your homework and notice that you used AI assistance without properly citing or

explaining it, points may be deducted depending on the severity of the issue.
If you are unsure whether the use of AI is permitted, please consult with me before using it.

[7. Additional Resources, Policies and Accommodations]

Statement on Disabilities

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/>)

Taking Care of Yourself

Grad school is challenging, pandemic or not, and I want you to prioritize your physical and mental health. I encourage you to reach out to University Health Services at (413) 577-5000 if you would like help with anxiety, depression, or mental health issues. The emergency counseling line is (413) 545-2337. Reach out to the UMass Police Department if you are having problems with your basic security. Contact your department or program to assist with academic difficulties as a result of sexual assault or violence, as well as contact faculty on behalf of the student.

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/).

Academic Integrity Statement

UMass Amherst is strongly committed to academic integrity, which is defined as completing all

academic work without cheating, lying, stealing, or receiving unauthorized assistance from any other person, or using any source of information not appropriately authorized or attributed. As a community, we hold each other accountable and support each other's knowledge and understanding of academic integrity. Academic dishonesty is prohibited in all programs of the University and includes but is not limited to: Cheating, fabrication, plagiarism, lying, and facilitating dishonesty, via analogue and digital means.

Sanctions may be imposed on any student who has committed or participated in an academic integrity infraction. Any person who has reason to believe that a student has committed an academic integrity infraction should bring such information to the attention of the appropriate course instructor as soon as possible. All students at the University of Massachusetts Amherst have read and acknowledged the Commitment to Academic Integrity and are knowingly responsible for completing all work with integrity and in accordance with the policy:

(<https://www.umass.edu/senate/book/academic-regulations-academic-integrity-policy>)

Title IX Statement

In accordance with Title IX of the Education Amendments of 1972 that prohibits gender-based discrimination in educational settings that receive federal funds, the University of Massachusetts Amherst is committed to providing a safe learning environment for all students, free from all forms of discrimination, including sexual assault, sexual harassment, domestic violence, dating violence, stalking, and retaliation. This includes interactions in person or online through digital platforms and social media. Title IX also protects against discrimination on the basis of pregnancy, childbirth, false pregnancy, miscarriage, abortion, or related conditions, including recovery. There are resources here on campus to support you. A summary of the available Title IX resources (confidential and non-confidential) can be found at the following link: <https://www.umass.edu/titleix/resources>. You do not need to make a formal report to access them. If you need immediate support, you are not alone. Free and confidential support is available 24 hours a day / 7 days a week / 365 days a year at the SASA Hotline 413-545-0800.

[8. Course Schedule]

Recorded class sessions will be made available for asynchronous viewing within 24 hours—generally the same day if possible. Links will appear in Canvas.

The following course schedule is subject to change to reflect our dynamic learning process. For example, if a topic cannot be completed within a given lecture, or if an additional topic is deemed important to the class, adjustments may be made. Additional lectures may be recorded and provided at the instructor's discretion.

Module #1: Essentials for Data-driven Storytelling

Sep 2 & 4 Class Introduction, Reviewing the Twelve Principles

- Syllabus
- [Miller] Chapters 2 & 4
- Optional: [Healy], Chapter 1, Section 2
- Optional: [table-issues](#)

Sep 9 & 11 Statistical vs. Substantive Significance; Causality

- [Miller] Ch 3

Sep 16 & 18 Audience & Format; Discussion on Readings and Your Project

- Selected Readings on Canvas
- Please prepare a list of the existing data projects you've worked on so far in DACSS;

Assignment#1 Due by Sep 26

Module #2: Graphic Theories: Design and Application in Data Visualization

Sep 23 & 25 Graphic Theory: Introduction

- [Bertin] Selected pages
- Visual Variables
- [Cairo] Introduction
- Optional: [Spencer], Ch 7

Sep 30 & Oct 2 Graphic Theory: Application

- [Healy] Chapter 1, Section 3-4
- [Miller] Chapter 5
- Optional: McLean, selected pages
- Optional: Sunbelt (2023)
- Optional: [The beauty of Visualization](#)

Assignment#2 Due by Oct 17

Module#3: Basic Interpretations and Comparing Models

Oct 7 & 9 Effective Analogies, Examples, and Discussion

- [Miller] Chapters 7
- Optional: TBD

Oct 14 & 16 Basic Quantitative Comparison (I) and (II)

- [Miller] Chapters 8
- Optional: TBD

Oct 21 TBD or Catching up

Oct 23 Understanding “Risk” and Uncertainty

- Understanding Uncertainty (Northwestern Magazine, 2016)
- Gelman, Andrew. 2018. “Ethics in Statistical Practice and Communication: Five Recommendations.” *Significance* 15 (5): 40–43
- (optional) Uncertainty and how we measure it for our surveys (Office of National Statistics, 2021)
- (optional) A Gentle Introduction to Uncertainty in Machine Learning (Brownlee, 2018)

Module#4: Comparing & “testing”

Oct 27 Reporting and Visualizing “Risk” and “Uncertainty”

- The effects of communicating uncertainty on public trust in facts and numbers (Anne Marthe van der Bles et. al., 2020)
- Hypothetical Outcome Plots: Experiencing the Uncertain (UW Interactive Data Lab, 2016)
- (optional) Visualising risk: a modern implementation of the Risk Characterization Theatre
- (optional) The margin of error: 7 tips for journalists covering polls and surveys (Denise-Marie Ordway, 2018)

Oct 29 Quantitative Comparison for Multivariate Models(I)

- Miller, Chapter 9
- Selected readings: TBD

Final Project Check-in#1 Due by Oct 31

Nov 4 Election Days: No Class Meetings;

Recorded Lectures on Quantitative Comparison for Multivariate Models(II)

Nov 6 The Goldilocks Problems & Solutions

- Gelman, Hill, and Vehtari, Ch 13 & 14;
- (optional) Hanmer adn Kalkan (2012): "Behind the Curve: Clarifying the Best Approach to Calculating Predicted Probabilities and Marginal Effects from Limited Dependent Variable Models."

Assignment#3 Due by Nov 21

Nov 11 Veteran's Day (No Class Meeting)

Nov 13 TBD: Checking in or Catching up

Module#5: Writing/Presenting to Different Audiences

Nov 18 & 20 Creating an online presentation of data; R Website & GitHub

- [Spencer] Ch 11 & 13
- Additional tutorial materials

Nov 25 & 27: Thanksgiving Retreat, No Class Meetings

Assignment #4 Due by Dec 5

Dec 2 Effective Presentation Design

- [Miller] Chapters 15 & 16

Dec 4 Wrapping-Up

Dec 9 Presentation and Peer-review

Assignment#5 Due by Dec 12

Final Project Product & Presentation Peer-review Form Due by Dec 19, 2024 (23:39 EST)

