

Data Mining: How the Course Works

Syllabus overview | DS 4140 / CS 5140 / CS 6140

Jeff M. Phillips

Fall 2026

Start with the course essentials

When and where

Tuesdays and Thursdays
10:45am–12:05pm | GC 1900

Office hours

Wednesdays, 9:00–10:00am in 3404 MEB
and directly after class

Who

Instructor: Jeff M. Phillips, 3404 MEB
TAs: Foad Namjoo, Arman Ashkari, Lucas Pearce

TA office hours | MEB 3105

Foad Namjoo: Mondays, 3:00–5:00pm
Arman Ashkari: Tuesdays, 12:30–2:30pm
Lucas Pearce: Tuesdays, 9:30–10:30am

We study how to find useful structure in large, messy data sets—with attention to modeling, error, and scalability.

What you should be able to do by December

- Turn structured or messy data into useful representations: vectors, sets, and matrices.
- Implement and analyze core methods for clustering, dimensionality reduction, graph analysis, and similarity search.
- Evaluate advanced methods and reason about their modeling choices, errors, and trade-offs.
- Work in a team to design and execute an open-ended data mining project.
- Communicate project progress and results in written, oral, and visual forms.

The longer-term goal

Read, understand, and implement ideas from data mining research papers—and use that knowledge to drive AI for data science.

Read strategically; track the milestone dates

Primary text

- Mathematical Foundations for Data Analysis (M4D)

Supplemental texts

- Mining of Massive Datasets (MMDS)
- Foundations of Data Science (FoDS)

The course webpage links the relevant reading to each lecture.

Major dates

Oct. 8 Midterm Test

Dec. 10 Endterm Test

Dec. 18, 10:30am–12:30pm
Project poster presentation / final exam

Dates and topics may change; updates will be announced through Canvas.

Ask early, and use the class community well

- Bring questions to instructor or TA office hours; email to arrange another time when needed.
- Speak up in class when something is unclear.
- Use the Canvas discussion group for questions about homework, schedule, and course material.
- Answering classmates' questions is encouraged when it supports understanding.

Keep solutions off the discussion board

Do not post potential homework answers. Such posts will be removed and will not be answered.

Important announcements will be made through Canvas; there is no separate class mailing list.

The course assumes four foundations

Be comfortable with

- Basic probability
- Basic linear algebra
- Big- O analysis
- Programming and data structures

Review: Foundations of Data Analysis

- Python is the supported default. Other languages are allowed for some work, but staff support may be limited.
- Assignments may be intentionally less specified than in lower-level courses.
- Undergraduates: CS 3500 and CS 3190 are the formal prerequisites.
- Graduate students: no enforced prerequisites; review M4D Chapters 1 and 3, plus early material in 2, 5, and 7.

Unsure about your preparation? Contact the instructor early.

Use LaTeX for clear scientific writing

- LaTeX is recommended for homework write-ups.
- It is a durable skill for research and scientific documents.
- The sample project demonstrates a starting 'tex' file and how to include a PDF figure.

Start here

Open the sample Overleaf project, copy it into your own account, and compile before the first assignment is due.

Your grade combines practice, depth, and mastery

25% Homework

Six short assignments, typically analytical work plus occasional programming.

40% Project

A team-based, open-ended analysis of an interesting data set, with intermediate milestones.

35% Tests

Midterm and Endterm. One two-sided sheet of notes; no computing devices.

- You may drop the lowest homework score **only if every assignment is completed**.
- Project members are responsible for the full project and may be evaluated orally.
- Planned letter-grade cutoffs: A at 93, A- at 90, B+ at 87, continuing in roughly three-point steps; boundaries may shift only in students' favor.

Late work loses credit on a fixed timeline

By 11:55pm	Under 24 hours	24–48 hours	Over 48 hours
Full credit on the listed due date	10% deduction	20% deduction	Not accepted Score: 0

- Homework is usually due Tuesday through Gradescope.
- **Lowest-score drop: only if every assignment is completed.** It is intended to absorb short illnesses; prolonged illness requires documentation.
- Request a regrade through Gradescope within **one week**; explain clearly why the solution is correct.

Collaboration is allowed; authorship is individual

Assignments Discuss approaches, proofs, and code with anyone. Then write your own code, proofs, and write-ups.

Project Work freely within your group and discuss with others. Acknowledge outside contributions and cite outside materials. Every partner must understand the entire project.

Tests Work by yourself. Communication with other students results in a zero.

A useful rule

Conversation may support your thinking; the submitted work must show your own understanding.

AI may assist learning, but it may not replace it

Allowed for assignments and the project

- Spell-checking, grammar, translation, and code completion; no declaration needed
- New perspectives, search assistance, or alternative explanations
- Sentence-level polishing is the safest use

Not allowed

- AI or calculating devices on either test
- Pasting an entire assignment question into a chatbot to produce the solution
- AI on assignment bonus questions
- Any use explicitly restricted by an assignment

If AI use appears disproportionate to the student's original work, the instructor may use an oral examination. You remain responsible for explaining the code, procedure, and analysis.

Academic misconduct has serious consequences

- Follow University Policy 6-410: no cheating, plagiarism, misrepresentation, inappropriate collaboration, or unauthorized/undocumented AI use.
- The Kahlert School of Computing records sanctions as “strikes.” A failing-grade sanction counts as one strike; two lesser sanctions count as one strike.
- Two or more incidents can lead to a recommendation for dismissal and a bar on future enrollment in School courses.
- Cheating on multiple homeworks or on the project results in a failing course grade.
- Talking during a test without an instructor or TA present leads to confiscation and a zero.

University Policy 6-410 | KSoC Academic Misconduct Policy

Know where to find support and report concerns

Accessibility and accommodations

Contact the instructor if course content is not compatible with assistive technology. The Center for Disability & Access arranges accommodations and alternative formats: 801-581-5020, disability.utah.edu. Approved recordings are personal-use only.

Campus safety

Visit safeu.utah.edu for information and training. Report suspicious activity or request a courtesy escort from Campus Police / DPS: 801-585-COPS (2677), dps.utah.edu.

Sexual misconduct support

Report to Equal Opportunity and Title IX (801-581-8365), the Dean of Students (801-581-7066), or Campus Police. For free, confidential support, contact Victim-Survivor Advocates through Campus Wellness: 801-581-7776.

Professional behavior protects the learning environment

- Follow the University of Utah Student Code and know the standards it sets.
- Classroom conduct should support learning, participation, and respect for others.
- Faculty may respond to misconduct with a verbal warning, dismissal from class, and potentially a failing grade.
- Students may appeal such action to the Student Behavior Committee.

Start strong

Bookmark the course webpage, join the Canvas discussion, verify your preparation, and ask questions before small problems become large ones.