

Mathematical Aspects of Data Analysis

MATH 123 — Fall 2026

Instructor	Abhijit Chowdhary
Email	abhijit.chowdhary@tufts.edu
Meetings	Tuesdays and Thursdays, 9:00–10:15 AM
Location	Braker Hall, Room 222
Office hours	To Be Determined.
Course sites	https://abhijit-c.github.io/tufts-math123-fall2026/

Course Description

Data science draws on statistics, computer science, signal processing, and domain knowledge, but many of its foundational questions are mathematical. This course develops the mathematical tools needed to understand why data analysis methods work, how they fail, and how to reason about unfamiliar methods as the field changes.

We will often view a dataset as a collection of points in a high-dimensional space, with linear algebra as our primary language. Topics include probability and concentration, the singular value decomposition (SVD), principal component analysis (PCA), least squares and regularization, graphs and spectral methods, clustering, nonlinear dimension reduction, randomized linear algebra, optimization, and classification. The final selection and depth of topics may change as the semester develops.

Prerequisites

Linear algebra and multivariable calculus. Programming experience is helpful, but not required. Distributed computational examples will use Python.

Learning Objectives

By the end of the course, students should be able to:

- represent data as vectors, matrices, and high-dimensional point clouds;
- use basic probability, covariance, and concentration ideas to reason about noisy and high-dimensional data;
- interpret the SVD, low-rank approximation, and PCA geometrically and computationally;
- formulate and analyze least-squares, regularized regression, clustering, and classification problems;
- explain how graph spectra, kernels, and random walks reveal structure in data;
- apply randomized projections and matrix sketching, and describe the guarantees they provide; and
- connect mathematical assumptions to the strengths, limitations, and possible failure modes of data-analysis methods.

Course Materials

The primary reference is:

A. S. Bandeira, A. Singer, and T. Strohmer, *Mathematics of Data Science* (2026), [arXiv:2607.11938](#).

The text assumes more mathematical sophistication and probability background than this course requires. Course notes and examples should therefore be the first reference. They'll reflect the course prerequisites, notation, and intended depth more closely.

Useful alternative perspectives include:

- A. Blum, J. Hopcroft, and R. Kannan, *Foundations of Data Science*;
- S. Boyd and L. Vandenberghe, *Introduction to Applied Linear Algebra*;
- M. P. Deisenroth, A. A. Faisal, and C. S. Ong, *Mathematics for Machine Learning*; and
- D. Calvetti and E. Somersalo, *Mathematics of Data Science*.

Assignments and Assessment

Weekly homework and quizzes

Homework will be assigned each Thursday after class. The homeworks themselves are not graded. However, the following Thursday's graded quiz will include one problem taken verbatim from that homework, potentially with a minor extension or variation. After the quiz, full quiz solutions and solution sketches for the homework will be released. Your lowest 2 quizzes scores will be dropped and, due to this, no makeup quizzes will be given.

Exams

There will be one midterm, and one final. Both will be in a similar style to the quizzes. You will have multiple problems taken verbatim from the homeworks, and then a few extensions. It is not guaranteed whether the problems will or will not have been asked in prior quizzes. See the calendar on the Course Website for the most recent dates. We do not give make-up exams in this course. An unexcused absence from any exam will be counted as a zero. Below is the default policy regarding the situation.

Missing an Exam: Excuses will be considered, at the discretion of the Mathematics Department, for genuine emergencies (for example, illness on the day of the exam or a death in the immediate family), for some unavoidable and unforeseen events of an extremely serious nature (your living quarters suffer a bad fire), or for a very limited category of foreseen, but unavoidable, serious events (your doctor attests that you must have surgery now). Requests to be excused must be accompanied by appropriate documentation (for example, a letter from Health Services stating that you are too ill to take the exam on that day or a note from your Academic Dean confirming there has been a death in your immediate family). They must state the reason for missing the exam and pledge that your account is truthful. Violating this pledge is an instance of academic dishonesty (see <http://math.tufts.edu/courses/examPolicy.htm>).

Course grade

Let Q be your quiz score after dropped quizzes, M be your midterm score, and F be your final score. Then, your final grade G is calculated using the more favorable of two exam weightings:

$$G = 0.3Q + \max \{0.3M + 0.4F, 0.2M + 0.5F\}$$

If you are excused from an exam, your final grade will instead be calculated as

$$G = \begin{cases} 0.5Q + 0.5F, & \text{If Midterm is excused,} \\ 0.6Q + 0.4M, & \text{If Final is excused.} \end{cases}$$

From these, your numerical grade will be converted to a letter one through the department grading scheme; see <https://math.tufts.edu/resources/grading-schemes>.

Optional extra credit

Once during the semester, you may read an approved research paper and discuss it with the instructor for about 15 minutes. Please, request a meeting at least 14 calendar days in advance. A completed conversation before October 22 earns 2.0 percentage points; October 22 through December 10 earns 1.0 percentage point. Credit is added after the course-grade formula is applied, up to a maximum numerical grade of 100. See the course site's [Extra Credit page](#) for more information.

Collaboration and use of outside tools

The homeworks are entirely open note, open resource, and open collaboration. That is, you may use anything for them, calculators, software, generative-AI tools, etc., so long as it's legal. That is not to say that they're required; the homeworks are designed to be completable should you understand the material. The quizzes and exams, however, are closed book and individually written.

Note: The quizzes and exams test understanding. While you may consult external resources for the homeworks, *use them to supplement your thinking*, do not avoid it. Thinking is hard but is the point.

University Policies and Student Resources

Accessibility and accommodations

Tufts is committed to providing students with disabilities equal access to courses, programs, and activities. Students seeking accommodations should contact the [StAAR Center's Accessibility Services](#) and follow its current request process. Please share an approved accommodation letter with the instructor promptly so that arrangements can be made; approved accommodations will be implemented without requiring disclosure of a diagnosis to the instructor.

Academic support at the StAAR Center

The StAAR Center provides academic coaching, writing support, subject tutoring, and other learning resources for Tufts students. Current services and appointments are available through the [StAAR Center website](#).

Academic integrity

Students are responsible for understanding and following Tufts' standards for honest academic work. These standards apply to assignments, quizzes, exams, code, data analysis, collaboration, and the use of outside sources or tools. When expectations for a particular task are unclear, ask before submitting the work. Suspected violations will be handled under the University's [academic integrity policies and procedures](#).

Course materials, recordings, and privacy

Instructor-created materials published in the public course repository are available under the repository's Creative Commons Attribution 4.0 license unless otherwise marked; reuse must provide appropriate attribution. Materials restricted to Canvas or otherwise marked private—including assessments, recordings, student work, grades, and identifying information—may not be posted or shared outside the course without the express permission of the relevant instructor or student. Classroom recordings, if any, are for approved course uses only and must respect student privacy and approved accommodations.

Mental health and well-being

Personal stress and emotional difficulties can interfere with learning and well-being. Tufts' [Counseling and Mental Health Service \(CMHS\)](#) offers counseling and urgent-care resources for eligible students and can be reached at [617-627-3360](tel:617-627-3360). Students in immediate danger should call 911 or the appropriate campus emergency service.

Sexual misconduct support and reporting

Tufts prohibits sexual misconduct and provides confidential support, medical care, academic assistance, and reporting options. Current contacts are listed on the University's [sexual-assault resources page](#). The Center for Awareness, Resources, and Education (CARE) offers confidential support, while the [Office of Equal Opportunity](#) receives reports. Please be aware that most faculty members are required to share disclosures of sexual misconduct with the appropriate University office; the linked resource page identifies confidential options.

Changes to This Syllabus

This syllabus and the course schedule are subject to reasonable revision. Substantive changes to assignments, grading, policies, or dates will be communicated through the official course channel and reflected in an updated syllabus when practical.