

Course Format: One meeting per week: 2:10-5:00 PM.

Course Objectives: Learn how to develop team-based solutions to big data problems. Gain experience and understanding of team-oriented problem solving. Learn mathematical, statistical and computational methods related to developing big data solutions.

Course Content: In this course, local businesses and University researchers will contribute big data problems that they are engaged in to the class. Typically, a single individual cannot solve big data problems because the solutions require disparate methods from statistics, mathematics and computer science. Consequently, solutions are developed by teams of individuals drawn from these areas in collaboration with the contributors. Students will learn how to work in teams to develop solutions to the contributed problems. Lectures and instructional materials on mathematical, statistical and computational methods will be provided as needed.

Principal Topics:

1. Data reduction techniques for massive data sets and data streams. Specific methods will be select to match the problems brought to the class.
2. Predictive analytics (as needed).
3. Task management.
4. Project management procedures: identifying goals, scope and constraints.
5. Written and oral presentation of results.

Learning Outcomes: At the completion of this course, the student will be able to

1. Apply mathematical, statistical and computational methods used in solving big data problems,
2. Write computer code for big data applications in one or more language (R, Python, Java, SPL, Matlab).
3. Plan a project solution and document progress,
4. Efficiently allocate workload among team members.
5. Communicate and participate as a team member.

Graduate Increment: Graduate students are to prepare a 4-8 page paper that reviews their project from mathematical, algorithmic or business perspective (depending on their area of concentration). The intended audience is not the client but professionals in their area of concentration, and so the paper expands on the client's report and goes into greater depth. For example, for a mathematics graduate student, the paper is to 1) review the principles or theory that support the methods developed by the team, 2) discuss the implementation of the method, and 3) discuss the results.

The paper will count 20% towards the course grade.

Prerequisites: Two courses chosen from STAT 341, M 221 and M 273, and one of M 461 or M 462, or

consent of instructor.

Textbook: None. Electronic resources will be posted on the course webpage.

Responsibilities: Teams are responsible for :

1. Project plans are due Tuesday February 10. The contents of the project plan are described as a separate document.
2. Progress reports are due Tuesday March 3, March 21, April 14 and 28. These are to describe what has been done since the previous report, a brief discussion of problems that have not been solved, and what will be worked on and by *whom* in the interim before the next report.

The results of analyses must be accompanied with a narrative describing methods and interpretation of the results. *Algorithms* should be described in sufficient detail that someone else may implement it. Send the progress report to me; after a review and perhaps some corrections, they will be sent to the client.

3. Communicating progress on a weekly basis to the instructor. The team is expected to meet with the instructor weekly (usually during the class meeting time) and describe progress and problems. Clients may be invited to participate in these meetings. If a team member is not cooperating or making an sufficient effort, please let me know.
4. When the project is complete, the progress reports will be combined and translated to a final report and delivered to the instructor and the client. The final report will include electronic versions of *annotated* computer programs and scripts. Identify the primary and secondary authors of the report, by section if possible, and include contact information. Final reports are due Wednesday May 13, 5 PM.
5. Oral presentation of project results Monday May 11, 1:10-3:10 PM (the final exam meeting time).

Attendance: Class attendance is very important as the class is a forum for solving problems that present themselves in the progression of the projects. You are expected to participate in and contribute to team discussions within your team. Please inform me by email if you will or have missed a class and provide an explanation of your absence. As a courtesy to the team members, also inform them in advance if you will not be present at class or at a outside-of-class team meeting.

Grading: Accomplishments 40%, progress reports 25%, written final report 25%, oral presentation 10%.

Accomplishments are graded on the basis of what *was* accomplished relative to what *could have* been accomplished. A set of qualitative guidelines for assigning letter grades to accomplishments:

1. A : Substantial progress and solid deliverables for most of the tractable aspects of the problems. Intractable problems identified as such.
2. B : Some tractable problems incompletely solved, but also substantive progress on most tractable problems.
3. C : Multiple tractable problems remain incompletely solved or unattempted.

A set of qualitative guidelines for assigning letter grades to progress, oral, and final reports :

1. A : Clear and well-organized, with a concise and well-documented summary of results. Problems and future work clearly explained.
2. B : Somewhat unclear with weak organization. Results deficient in some manner. Problems and future work not clearly explained.

3. C : Lacking clarity and organization. Poorly documented results (e.g., no figure captions). Problems and future work vaguely explained.

Incomplete (I) Grades: Incompletes (I's) are given at the discretion of the instructor. See the 2009-2010 UM catalog for the conditions under which an incomplete may be given. Incompletes will not be given under any other circumstances.

Exams: There will be no exams during the semester. The final exam will be comprised of a presentation of projects. Approximately 30 minutes is allocated for each team to make their presentation.

Adding/dropping the course: Deadlines for withdrawing from course are listed at <http://www.umt.edu/registrar/calendar.aspx>.

Academic Misconduct: All students must practice academic honesty. Academic misconduct is subject to an academic penalty by the course instructor and/or a disciplinary sanction by the University. All students need to be familiar with the Student Conduct Code. The Code is available for review online at <http://life.umt.edu/sa/documents/fromWeb/StudentConductCode1.pdf>.