

MATH 2740
Mathematics of Data Science
Fall 2025

1 General information

Instructor

Instructor : Julien Arino
Email : julien.arino@umanitoba.ca

To ensure your email is not lost among all the emails I receive, ensure that your subject contains a tag of the form [MATH 2740]. I do not answer emails if the topic is posted in UMLearn Announcements or any other easy to find source, so read these carefully.

Lectures, office hours, course page

Class schedule : TR 1130–1245, 200 Fletcher Argue
Office hours : TR 1430–1600, 238 St Paul’s College
Course page : UMLearn + GitHub

Tutorials

Section	Day and time	Location
B01	W 0830–0920	301 Biological Sciences
B02	W 0930–1020	301 Biological Sciences
B03	W 1130–1220	301 Biological Sciences
B04	W 0830–0920	505 Tier

The first tutorial is on Wednesday 10 September 2025.

Course description Mathematics provides the backbone of all scientific disciplines. This is particularly true with Data Science, where mathematics helps manipulate and understand the complex or large data sets that are used.

This course introduces some mathematical concepts used in Data Science. Algorithmic and computational aspects are also considered. Topics include linear algebra: least squares, singular value decomposition, principal components analysis, and graph theory: centrality, social network theory, clustering.

Office hours Office hours are Tuesdays and Thursdays, 1430 to 1600 in 238 St Paul’s College. Because of the ongoing renovation of Machray Hall, I am not able to provide office hours outside these times.

Textbook It is difficult (read: impossible) to find a textbook that meets the needs of a course like ours. As such, there is no required textbook. However, I will post links on UM Learn to relevant textbook chapters throughout the term. Slides and computer code will also be distributed as we progress.

Course videos Videos for the course as I taught it in 2021 are available as a YouTube playlist. There is no guarantee that we will follow the same program this year.

Tutorials It is strongly recommended to attend tutorials, as this is where you will review some of the mathematical content. You will also have to write 4 quizzes during tutorials at dates indicated below. Worksheets will be posted on UM Learn in advance of tutorials so that you have an opportunity to attempt problems before your tutorial. During the tutorial, the TA will go over some of the problems and answer any questions you may have. Note that the lab worksheets are NOT for credit; they are meant as extra practice.

Voluntary Withdrawal deadline The Voluntary Withdrawal deadline is Tuesday 18 November 2025.

2 Evaluation

Evaluation components – Evaluation involves 4 different components.

1. 4 assignments worth 5% each for a total of 20%.
2. 4 quizzes during tutorials worth 5% for a total of 20%.
3. One midterm evaluation on Monday 27 October 1900–2100 worth 25%. Rooms TBA.
4. One final examination run during the December final examination period, worth 35%. Rooms TBA by the Registrar's Office.

Note that all evaluated work needs to explain the thought process used. Correct but unjustified answers will be severely penalised.

Letter grade values – The following are the percentage ranges to achieve each letter grade.

Letter Grade	Percentage out of 100	Final Grade Point
A+	94-100	4.5
A	86-93	4.0
B+	80-85	3.5
B	72-79	3.0
C+	65-71	2.5
C	60-64	2.0
D	50-59	1.0
F	Less than 50	0

These percentages will not be adjusted upwards. However, they may be adjusted downwards, e.g., the cut-off for an A+ *could* end up being 90% instead of 94%.

Assignments – Being able to use computers is an integral part of being a data scientist, so in this course, we will use computers a lot, using R **exclusively**. Computer assignments therefore need to be handed back as R Markdown (Rmd) files. Any other format will result in a mark of 0.

The assignment schedule is as follows.

Assignment number	Handed out	Due
Assignment 1	19 September	3 October
Assignment 2	10 October	24 October
Assignment 3	31 October	21 November
Assignment 4	21 November	5 December

Quizzes – Quizzes are held during tutorials and are 15 minutes long. They test your capacity to memorise definitions or theorems or to perform simple proofs or computations. The program for the quiz on a given Wednesday is the 3 previous lectures (the lecture on the Tuesday immediately prior to the quiz and the two lectures previous week).

Tutorials run at 0830, 0930 and 1130, so one of these time slots writes a quiz each week, based on the following schedule.

Tutorial slot	Quiz 1	Quiz 2	Quiz 3	Quiz 4
0830	10 September	15 October	29 October	19 November
0930	17 September	1 October	5 November	26 November
1130	24 September	8 October	22 October	3 December

You **must** write the quiz in the tutorial you are registered in. There is **no make up** for missed tutorials and in particular, you **cannot** write a given quiz in another week. Please pay attention that a different time slot writes the quiz first each time, so ensure you have written down the dates carefully, as per the statement above, there will be no make up for missed quizzes.

Midterm examination – Held on Monday 27 October 1900–2100. The program will cover the entirety of the material covered up to and including the tutorial in the week prior to the test (i.e., the material covered on 22 October). The midterm will be mostly mathematical, although a couple of questions regarding simple understanding of computer code could be asked.

Final examination – The final exam will be scheduled by the Registrar’s Office during the Final Examination Period. The program will be the content during the entire year and will be mostly mathematical, although a couple of questions regarding simple understanding of computer code could be asked.

3 Self-declarations of absences

You can self-declare **one (1)** absence of less than 120 hours (5 days).

1. To self-declare an absence, use the self-declaration form on the university website ([link](#)).
2. Self-declarations are meant to be used occasionally only; as a consequence, I will **not accept more than one during the term**.
3. A self-declaration must be filed less than 48 hours after the event(s) you are using it for.
4. There is no make up for anything missed during a self-declared absence: any percentage will be moved to the Final Examination.
5. As per the university’s website ([link](#)), repeated absences or more serious issues need to be discussed with your Faculty or Student Accommodation Services.

4 Further remarks on Assignments

1. Assignments are posted on UMLearn on the indicated Fridays at 1200 and are due on the indicated Fridays at 1200.
2. There is **no tolerance** for late assignments and **no makeup** assignments. The system will not accept submissions past the deadline and I will not consider late submissions by email. Any assignment not marked will receive a mark of **zero**. Since you may need to seek assistance from me or the TAs and need to take into account fudge factors when submitting, leaving yourself a comfort margin will be useful.
3. Assignments are computer based and require you to submit an **R Markdown** file on UMLearn. Assignments returned in languages other than R will receive a mark of zero. The R Markdown file needs to properly knit. Compressed files, illegible files, etc., will also receive a mark of zero, so please double check your submission.
4. While it is perfectly acceptable to discuss approaches to the assignment problems, it is expected that students actually solve and present solutions by themselves. Acts of academic dishonesty will be dealt with harshly. See below for details on submission of computer assignments.

Please take a look at the following checklist and ensure that you can tick all the boxes. Failure to do so will result in lost marks or, in the very worst case, in a mark of zero and an Academic Dishonesty procedure being initiated. This checklist is also available on UMLearn.

1. Whatever R libraries you are using are not mindlessly downloaded and installed if I already have them; they are also loaded from within your code. If you use, say, library X, then you have at the beginning of your code a call of the form

```
if (!require(X)) {  
  install.packages("X")  
}
```

2. Your code is self-contained; you can only submit one file (your Rmd file), so any data you use is downloaded from the web by your code.
3. Your code runs “from scratch” and without any files you have locally on your hard drive. If your code fails to run from start to finish when we run it, you will lose a (potentially large) number of marks.
4. Your code is in R, not Python or any other language.
5. Your code is not a mindless copy of the code I provide as example.
6. Your code is not a copy of someone else’s, including mine.
7. Your code is not literal production from genAI. If you used genAI, you **indicate** so and also indicate the specific genAI engine used.

I will sanction any submission in which it is obvious that one of these checkboxes is not ticked. In particular, any assignment in which box 4 is not ticked will get 0. And points 5–7 will result in a case of Academic Dishonesty being filed to the Head of the Department of Mathematics.

5 Use of generative AI (genAI)

We are in a period of rapid change and the policies regarding the use of genAI evolve all the time. I recommend reading material posted by the UM Copyright Office ([link 1](#), [link 2](#)) and the UM Centre for the Advancement of Teaching and Learning ([link](#)).

I strongly believe that genAI should be used in teaching and research, but also believe that this should be done *wisely*. genAI is a fantastic tool to help *experts*: in short, if you know what you are doing, it can dramatically reduce the time that it takes you to do things. In the previous sentence, I want to stress “*if you know what you are doing*”. In this and other courses, you are acquiring the methods and tools that allow you to become an expert in your field.

So although I will not forbid you from using genAI, I want to encourage you to do as much as possible without using it. That is how you will gain the skills that will make genAI extremely useful for you. If you are using genAI, then I will ask that you indicate so. In assignments, add a comment at the top of your file(s) indicating that you have used a genAI. Also indicate the specific tool you used (ChatGPT, Claude, Copilot, Gemini, etc.)

Please note that I have been using genAI for some time and am quite familiar with some of their quirks and habits. If I see something that looks “straight out of AI” with no intervention on your part, I will split the mark between you and the machine. And will declare an Academic Dishonesty to the Head of the Department of Mathematics if you do not acknowledge genAI use even though it is obvious that you used it.

6 Academic Dishonesty Policy

The Department of Mathematics, the Faculty of Science and the University of Manitoba all regard acts of academic dishonesty in quizzes, tests, examinations or assignments as serious offences and may assess a variety of penalties depending on the nature of the offence. Acts of academic dishonesty include bringing unauthorized materials into a test or exam, copying from another student, plagiarism and examination personation. See details on the Academic Integrity pages.