

CS 504 Principles of Data Management and Mining Syllabus Spring 2024

Contact Information

Name: Binqian Yin, Ph.D.
Office Number: Buchanan Hall D215J
Email: byin2@gmu.edu
Office Hour: Mon/Tue 3:00-4:00pm

Time and Place of Class Meeting

Section	Day	Time	Building	Room
002	Tuesday	4:30-7:10pm	Exploratory Hall	L004
003	Monday	4:30-7:10pm	Innovation Hall	103

Graduate Teaching Assistants Contact Information

GTA Name	GTA Email	Office	Office Hours		Week
			Day	Time	
TBA	TBA	Buchanan hall D215H	TBA	TBA	TBA
TBA	TBA	Buchanan hall D215H	TBA	TBA	TBA
TBA	TBA	Buchanan hall D215H	TBA	TBA	TBA
TBA	TBA	Buchanan hall D215H	TBA	TBA	TBA

Course Description and Goals

CS 504 is an introductory, 3-credit course that explores a wide variety of topics related to data science. While it covers some aspects of databases and machine learning, the specific focus of this course is on data itself: how to store, query, and extract meaningful structure from it. Topics covered will include database management, data models, SQL, and popular models in machine learning.

The goal of this course is helping students develop problem-solving and analytical thinking skills within the context of data science. Provide an overview of modern approaches for managing and analyzing data. Enhance students' design and programming skills in order to work effectively with large data sets.

Text Book

- *Fundamentals of Database Systems* (Seventh edition) Shamkant B. Navathe and Ramez Elmasri
 - *Introduction to Data Mining* (Second edition) Pang-Ning Tan, Michael Steinbach, and Vipin Kumar
-

Grading Weights

The course has a maximum of 100 percent points (without the Extra Credits) distributed as follows:

- Quiz: 20% (we will have 3 quizzes, and only 2 highest score will count)
- Midterm Exam: 25%
- Project: 30%
- Final Exam: 25%

You can earn Extra Credits through the following:

- Pop-up Attendance Check-in: 3% (1%*3)
 - The credit from the pop-up check-in is only available at the scheduled time set by the instructor and requires the completion of all necessary tasks. Submissions made after this time will not be eligible for credit.

Both your TA and I are committed to offering reasonably detailed and prompt feedback on all assessments. If you have any questions or concerns about your grades, don't hesitate to contact us. You can reach us during scheduled office hours or through your "@gmu.edu" email address. We will not respond to messages sent from or send messages to a non-Mason email address.

Grades will be changed only when a grading error has been made. If there is an error in your grade, please contact us within a week of receiving the feedback. After this period, we may be unable to make any change.

Grading Scale

This course is graded according to the standard graduate grading policies, with the following cutoffs:

- A+ : ≥ 95
- A : $\geq 90, < 95$
- A- : $\geq 88, < 90$
- B+ : $\geq 85, < 88$
- B : $\geq 80, < 85$
- B- : $\geq 78, < 80$
- C : $\geq 70, < 78$
- F : < 70

Course Update

Any updates related to this course will be announced on Blackboard, where you can also find all the resources for this course. Please check Blackboard regularly to ensure you do not miss any updates.

Absence of the Quiz and Exam

The student must take the exams and quizzes in the scheduled time of the course enrolled.

For the absence of quiz:

- The make-up quiz will not be offered.
- Zero will be recorded if student absence of the quiz.

For the absence of Mid-term and Final exam:

- Must get the permission from instructor before the exam at least 5 business day. The proof of absence excuse should be provided.
 - The make-up exam must be taken within the specified time period.
 - Zero will be recorded for midterm exam if without prior permission from instructor.
 - Unexcused absence from the final exam will result in grade F for the course.
-

Due Dates & Grace Period

Throughout the semester, you are allotted three "free" late days (grace period) to use for assignment and project submissions. These late days can be used all at once or distributed across project and different assignments —whatever best suits your needs. Keep in mind that each late day is triggered automatically following an assignment's due date: the first one minute after, the second 24 hours and one minute after, and the third 48 hours and one minute after. Any late submissions beyond these three days will result in 0.

Penalties of late submission will be calculated based on the time of the **most recent submission**.

Piazza

Piazza will serve as our primary platform for course-related communication. You have the option to send private messages to instructors (which will be visible to both professors and TAs) and to post public questions for the entire class to see. This platform allows for collaboration on responses and the ability to categorize discussions by topic.

For a prompt response, it's recommended to post your questions on Piazza. However, if you need to discuss confidential matters with a specific professor or TA, please feel free to email us directly or use a private post on Piazza.

Discussing quiz, exam content and any solution (quiz, exam, assignments ect.) on Piazza is not allowed for this course. But please reaching out to instructor directly via email or stop by during office hour if you have any exams related questions.

Office Hour

No appointment is needed for regular office hours. If additional meeting time is required, it can be arranged by making an appointment in advance via email.

Email

If you need to discuss confidential matters with professor, please send email to byin2@gmu.edu.

When sending an email, please use the subject format '**CS504-section (choose 002 or 003) + (Your Name) + (Title of the Inquiry)**'.

Honor Code

All students enrolled this course must abide by GMU Honor Code. Please take the time to acquaint yourself with the University's overarching [honor code](#), as well as the [specific honor code](#) for the Department of Computer Science. These policies are rigorously enforced. Any violations of academic integrity will be reported to the Honor Committee.

Copyright

All course materials uploaded onto Blackboard or other course platforms are considered private, and a significant portion of these materials is protected by copyright. Please don't share any course material unless you've been authorized to do so.

Disability Accommodations

If you possess documented evidence of a learning disability or any other condition that could impact your academic performance, please talk with me to discuss your specific accommodation requirements and ensure that your documentation is on record with [Office of Disability Services](#).

Sexual Harassment, Sexual Misconduct, and Interpersonal Violence

As a faculty member and a designated 'Responsible Employee,' I am obligated to report all disclosures of sexual assault, interpersonal violence, and stalking to Mason's Title IX Coordinator per university policy 1412.

If you wish to speak with someone confidentially, please contact the Student Support and Advocacy Center (+1-703-380-1434) or Counseling and Psychological Services (+1-703-993-2380). You may also seek assistance from Mason's Title IX Coordinator (+1-703-993-8730; titleix@gmu.edu).

CS 504 Course Schedule for Spring 2024 Section 002

Week	Date	Topic	Reading Chapter*
1	01/16/24	Lecture 1: Introduction of the course	a. Chapter 1 & 2
2	01/23/24	Lecture 2: ER & EER Modeling	a. Chapter 3 & 4
3	01/30/24	Lecture 3: Relational Database Design	a. Chapter 5 & 9
4	02/06/24	Lecture 4: Basic SQL	a. Chapter 6
5	02/13/24	Lecture 5: More SQL Lecture 6: Database Security and Authorization	a. Chapter 7 & 30
6	02/20/24	Lecture 7: Basic of Functional Dependencies and Normalization	a. Chapter 14
7	02/27/24	Midterm Exam	Lecture 1~7
8	03/05/24	No Class – Spring Break	/
9	03/12/24	Lecture 8: Data and Data Preprocessing	b. Chapter 1 & 2
10	03/19/24	Lecture 9: Probability Theory, Machine Learning Introduction	b. Chapter 4 & 7.2
11	03/26/24	Lecture 10: Linear Regression Model	b. Chapter 4.6
12	04/02/24	Lecture 11: Supervised Learning Method 1: KNN & Naïve Bayes Classifier	b. Chapter 4.3 & 4.4
13	04/09/24	Lecture 12: Supervised Learning Method 2: DT	b. Chapter 3
14	04/16/24	Lecture 13: Model Evaluation	b. Chapter 3.6
15	04/23/24	Lecture 14: Supervised Learning Method 3: ANN	b. Chapter 4.7
16	04/30/24	No Class – Reading Day	/
17	05/07/24	Final Exam	Lecture 8 ~14

*The “reading chapter” is based on the text book listed below:

- a. ***Fundamentals of Database Systems*** (Seventh edition) Shamkant B. Navathe and Ramez Elmasri
- b. ***Introduction to Data Mining*** (Second edition) Pang-Ning Tan, Michael Steinbach, and Vipin Kumar