

CDS 302
Spring 2021
Scientific Data and Databases
Syllabus

1 General Information

Instructor: Gideon K. Gogovi

Class Days: Wednesday 7:20 PM - 10:00 PM

Place: Online Sync Instructional Method via Zoom

Course website: Blackboard

Credits: 3.0

Office Hour: Thursdays 12:00 PM - 1:00 PM, (TBD)

Contact method: Direct email: ggogovi@gmu.edu or Slack.

2 Course Objectives

Description:

The main focus of this course is the design and implementation of Database Management Systems (DBMS). The topics to be covered include the E-R and Relational Models, SQL queries, views, transactions, indexing, hashing, file storage, query processing, and visualization of results. The course will include not only the theoretical foundations, but also practical applications, using real datasets. There will be in-class examples and homework assignments for storing, managing and querying scientific data. Finally, the course will include discussions on some more advanced topics, beyond traditional DBMS. Furthermore, as this is a writing intensive course, we will visit the basics of scientific writing, including how to write articles using LaTeX as the preferred typesetting software.

3 Learning Outcomes

By the end of the course, each student will be able to:

- have a broad knowledge on fundamentals, theory and applications of Database Management Systems;
- have a deep understanding of the principles of database design and the modeling of relational data both at the logical level as well as the physical level;

- have experience in formulating simple and complex queries using a pure query language, relational algebra, as well as expressing them in the popular structured query language, SQL;
- have the necessary writing skills for creating scientific manuscripts, articles and reports, using the appropriate math notations, tables, references, citations, etc.;
- be able to articulate and effectively communicate concepts and ideas related to Data Management and Databases to experts, non-experts, and other professionals in a work environment;
- have the ability to appropriately apply the knowledge acquired in the course for various hypothetical and real-world database design and management tasks;
- have experience with storing, indexing, accessing, querying, managing, and visualizing data from scientific databases using SQL and python.

4 Course Schedule (tentative)

Weeks are approximate time spent on material. Actual Midterm date will be announced in class.

Week 1: Introduction: Why Databases? Course overview. Scientific Writing: Introduction to \LaTeX

Week 2: \LaTeX : Math Notations, Equations, Tables, References, Beamer etc

Week 3: The Relational Model: DB schema, keys, diagrams and Relational Algebra. Introduction to SQL: basic queries - Part I

Week 4: SQL Part II: string operators, set operations, aggregate functions, ordering, grouping, having-clause. SQL practice for Parts I & II

Week 5: SQL Part III: nested queries, set membership, joins, self-joins, with-clause, views, stored procedures. SQL practice – Part III

Week 6: SQL practice – Part IV

Week 7: Midterm Exam!

Week 8: The Entity Relationship Model & Relational DB Design. Reduction of E-R to Relational Schema

Week 9: Storage and File Structure - Overview. Indexing and Hashing: Basics, Ordered Indices

Week 10: Indexing: Dense vs. Sparse Index, Multilevel index, B+-Tree structure; B+-Tree insertions, deletions, updates

Week 11: Indexing: B-Tree; Bitmap Index; Multiple-Key Access, Static Hashing, Dynamic Hashing; Hash file organization

Week 12: Query Processing: Selection Operation, Point and Range Queries

Week 13: Query Processing: Join Operations, Other Queries, Connecting to a DB, run SQL Queries through R, Create & Modify Tables through R

Week 14: Final Project presentation

5 Course Format

- The course will be part lecture / hands-on discussions. About half the class time will be lecture and half will be in class programming practice/discussions.
- There will be homework each week (several writing and programming assignments).
- Midterm (in the form of project proposal) and a Final project.
- Most homework due in one week. Consult each homework on Blackboard for actual due date.

Students are required to log in and participate during class. The lectures will not be recorded by the instructor by students are free to do so.

6 Course materials

- Lecture notes will be posted to the class page (aka Blackboard) each week.
- The detailed assignments will be posted on the class page.
- Various guides will be on the class page. Students are required to review them.

7 Technology Requirements

- Access to a Windows, Macintosh, or Linux computer.
- Permission to install software on above computer.
- A reliable internet connection.
- Alternatively to the above: Ability to reliably use the computer labs at GMU.

- A supported web browser for Blackboard.
- Adobe Acrobat Reader.
- Microsoft Office (freely available to Mason students).
- The course requires the use of a MySQL, SQL and L^AT_EX(a freely available open source software for typesetting).

8 Course Expectations

- Students are expected to already be comfortable using a computer, accessing the internet, and have word processing skills.
- Student will be expected to learn to use developmental tools on their computer.
- Students work should show attention to detail. Getting the right answer is insufficient. Following the guides posted on the class page is required.
- Students are required to regularly check the Blackboard for information.

9 Assessment

9.1 Assignments

- All assignments are required to be completed.
- Most assignments have one week due dates. A few may be two week.
- Assignments might not have equal weights.
- Reports and files are submitted through the “Assignments” section of the Blackboard. Do not email assignments to the instructor’s email or through the Blackboard email.

9.2 Late Policy

- Up to three days late, there will be a penalty of 10% per day (Weekdays). Zero after the third day.
- Turned in before reading day, 30% penalty.
- Exceptions to this policy may be made on a case- by-case basis at the discretion of the Instructor.
- Deferral of course work is a privilege and not a right; there is no guarantee that a deferral will be granted.

9.3 Exams

- There will be one midterm exam in the form of project proposal presentation and one final exam in the form of a final project presentation.

9.4 Course Grades

Homework assignments will be worth between 50 to 100 points each depending on their complexity. Some assignments may have bonus points available. At the end of the term, the weighted average of the assessments are as follows:

- In Class discussion, SLACK Participation and Attendance: 10%
- Assignments and Blackboard Quizzes: 25%
- Midterm Exam: 30%
- Final Exam: 35%

10 Other Important Dates

Please refer to Mason's academic calendar (Spring 2021) for information on important dates and follow Mason's announcements on any calendar changes during the semester. The Spring 2021 Mason calendar is available at <https://registrar.gmu.edu/calendars/spring-2021/>

11 Student Expectations

11.1 Academic Integrity

Students must be responsible for their own work, and students and faculty must take on the responsibility of dealing explicitly with violations. The tenet must be a foundation of our university culture. [See <https://oai.gmu.edu>].

11.2 Honor Code

Students must adhere to the guidelines of the George Mason University Honor Code [see <https://oai.gmu.edu/mason-honor-code/>].

11.3 MasonLive/GMU Email

Students are responsible for the content of university communications sent to their George Mason University email account and are required to activate their account and check it regularly. All communication from the university, college, school, and program will be sent to students solely through their Mason email account. [See <https://masonlivelogin.gmu.edu/login>].

11.4 Patriot Pass

Access to Mason's online system is done through a personal username and password known as the Patriot Pass. Once you sign up for your Patriot Pass, your passwords will be synchronized, and you will use your Patriot Pass username and password to log in to the following systems: Blackboard, University Libraries, MasonLive, myMason, Patriot Web, Virtual Computing Lab, and WEMS. [See <https://password.gmu.edu/index.jsp> , this link also enables you to reset your password].

11.5 Responsible Use of Computing

Students must follow the university policy for Responsible Use of Computing. [See <https://universitypolicy.gmu.edu/policies/responsible-use-of-computing/>].

11.6 Students with Disabilities

Students with disabilities who seek accommodations in a course must be registered with the George Mason University Office of Disability Services (ODS) and inform their instructor, in writing, at the beginning of the semester [See <https://ods.gmu.edu>]. Please be proactive and let me know how I can help you with any disability issue.

12 Student Services

12.1 University Libraries

University Libraries provides resources for students. [See <https://library.gmu.edu>].

12.2 Writing Center

The George Mason University Writing Center staff provides a variety of resources and services (e.g., tutoring, workshops, writing guides, handbooks) intended to support students as they work to construct and share knowledge through writing.

[See <https://writingcenter.gmu.edu>]. You can now sign up for an Online Writing Lab (OWL) session as well as face-to-face session in the Writing Center (see the Tutoring section in the link above).

12.3 Counseling and Psychological Services

The George Mason University Counseling and Psychological Services (CAPS) staff consists of professional counseling and clinical psychologists, social workers, and counselors who offer a wide range of services (e.g., individual and group counseling, workshops and outreach programs) to enhance students' personal experience and academic performance [See <https://caps.gmu.edu>].

12.4 Family Educational Rights and Privacy Act (FERPA)

The Family Educational Rights and Privacy Act of 1974 (FERPA), also known as the "Buckley Amendment," is a federal law that gives protection to student educational records and provides students with certain rights. [See <https://registrar.gmu.edu/privacy>].

13 Final Notes

- Any typographical errors in this Course Outline are subject to change and will be announced in class.
- The date of the final examination is set by the Registrar and takes precedence over the final examination date reported by the instructor.
- Recording of any kind (audio, video), reuse or remix of course materials, and further dissemination of the course content is not permitted unless prior written consent of the professor and George Mason University has been given.