CDS 302-DL2 Scientific Data and Databases

Spring 2022

1 General Information

Instructor: Dr. Carmen A. Iasiello
Class Days: Asynchronous via Blackboard
Credits: 3
Specialized Designation: Writing Intensive in Major
Prerequisites: CDS 101 or CDS 130 or equivalent, or permission of instructor
Office Hours: By Appointment via Zoom
Contact Method: Direct email ciasiell@gmu.edu

2 Course Objectives

The main focus of this course is the design and implementation of Database Management Systems (DBMS). The topics to be covered include scientific writing, E-R and Relational Models, and SQL queries, views, transactions, indexing, hashing, file storage, and query processing. This course will include both theoretical foundations and practical applications of using real world data sets. There will be class examples and homework assignments for storing, managing, and querying data. The course will also include discussions of more advanced topics, beyond traditional DBMS. Finally as this is a writing intensive course, we will visit the basics of scientific writing using LaTeX.

3 Learning Outcomes

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

- have a broad knowledge on fundamentals, theory and applications of DBMS;
- have a deep understanding of the principles of database design and the modeling of relational
- 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 Technology Requirements

Software: This course will be using LaTeXas the main text editor for writing assignments. The use of Overleaf online LaTeXeditor is strongly recommended. Alternatively, you may use another LaTeXeditor of your choosing. A citation management tool is also highly recommended (e.g. Zotero). The course will also be using SQL as a query language and Python as a programming language. For assignments you will need MySQL and MySQL Workbench (or similar SQL RDBMS and GUI) and Python 3.x. A basic understanding of computer programming principles and knowledge of any programming language is assumed.
Hardware: You will need access to a Windows, Linux, or Macintosh computer with at least 2 GB of RAM and access to a fast, reliable broadband Internet connection. For the amount of computer hard disk space required to take an online course, consider and allow for the space needed to: 1) install the required and recommended software and, 2) save your course assignments.

5 Grades

At the end of the term all marks will be totaled as a weighted average according to the following weights: Assignments: 50%, Midterm Exam: 25%, Final Project: 25%. Each assignment will be given 0-10 points with some assignments including bonus (i.e. extra credit points). The grading rubric for each assignment will be provided with the prompt. The Midterm Exam will consist of five prompts answered in a single LaTeXdocument and graded on a scale of 0-50. The Final Project will consist of DBMS query design of your choosing, delivered as a written LaTeXdocument, and graded on 0-50 point scale. Details about the Final Project will be provided in Week 8. The course grading scheme will be as follows: A+ (97+%), A (93%), A- (90%), B+ (87%), B (83%), B- (80%), C+ (77%), C (73%), C- (70%), D+ (67%), D (60%)

6 Course Schedule

Week	**Dates**	**Topic**	**Assignment** (Due on Sunday)
1	Jan 24-30	Course Overview. Introduction: Why databases? Scientific Writing: Introduction to LaTeXand Overleaf	A1: Self Introduction
2	Jan 31-Feb 6	LaTeX: Math Notations, Equations, Tables, Graphs. How to Read a Scientific Article	A2: Scientific Article Summary
3	Feb 7-13	The Relational Model: DB schema, keys, diagrams, and Relational Algebra. Introduction to SQL: basic queries - Part I	A3: Basic SQL queries
4	Feb 14-20	SQL Part II: string operators, set operations, aggregate functions, ordering, grouping, having-clause. SQL practice: Part 1 and 2	A4: SQL queries
5	Feb 21-27	SQL Part III: nested queries, set membership, joins, self-joins, with-clause, views, stored procedures. SQL practice - Part 3	A5: Advanced SQL queries
6	Feb 28-Mar 6	SQL practice - Part 4 Exam Preparation	
7	Mar 7-13	**MIDTERM EXAM**	
	Mar 14-20	**Spring Break:** No Lesson	
8	Mar 21-27	The Entity Relationship Model and Relational DB Design. Reduction of E-R to Relational Schema	A6: E-R Model
9	Mar 28-Apr 3	Storage and File Structure - Overview. Indexing and Hashing: Basics, Ordered Indices.	A7: Final Project Proposal
10	Apr 4-10	Indexing: Dense vs. Sparse Index, Multilevel index, B-Tree structure, B+-Tree insertions, deletions, updates	A8: E-R to Relational Model
11	Apr 11-17	Indexing: B-Tree, Bitmap Index, Multiple-Key Access Static Hashing, Dynamic Hashing; Hash file organization	A9: Indexing
12	Apr 18-24	Query Processing: Selection Operation Point and Range Queries.	
13	Apr 25-May 1	Create and Modify Tables through Python. Practice Examples	A10: Python
14	May 2-10	Final Project Preparation	
	May 13	**FINAL PROJECT DUE**	

7 Exams

The course includes mandatory written midterm exam and final project. The material covered in the exam/project will be announced in class. Only student who cannot write a complete a course assignment because of an illness, severe domestic affliction or other compelling reasons can apply for extension of time to complete an assignment.

8 Assignments

The course will include several written assignments on selected topics from the material covered in class and in the assigned reading. All assignments are mandatory. Typically, 1 week will be allocated for every assignment. Assignments should be submitted through the Blackboard course website. **Please note:** Assignments should be submitted only through the Assignment submission section of the Blackboard system - DO NOT email assignments directly to the instructor.

The late assignment submission policy is the following:

- Submissions within the first day (24 hours) after the deadline: -10% credit.
- Submissions within the second day after the deadline: -20% credit.
- Submissions within the third day after the deadline: -30% credit.
- **No assignments after the third day will be accepted (i.e., 72 hours after the deadline a 0 will be entered as your assignment grade.)**

Exceptions to this policy may only be granted in serious circumstances and at the discretion of the Instructor.

9 Grading

General guidelines for ASSIGNMENT preparation and submission:

- Grades of assignments will be based on the rubric for each assignment and will necessitate academic merit of your answers, conciseness and completeness of your answers, organization and presentation. Please organize your report in a logical fashion so that your answers could be easily identified.
- Please remember that your assignment is a professional document and should therefore be formatted and constructed accordingly. All written assignments are to be completed in LaTeXformat and submitted as a pdf via blackboard. If and additional files are required for assignment submission, they will be clearly stated in the assignment.

- Each assignment submission must include the following information: assignment, title, assignment number, student name, and submission date.
- Remember to have a backup of all the materials you submit.

10 Course Website

The course has a Blackboard website. This website will provide you a single portal through which you may obtain lecture notes, retrieve assignment data and, review links to additional materials, and receive special announcements. You are required to visit the course website once per day. Please notify ITU (and, if necessary, the instructor) if you encounter any problems accessing this website.

11 Electronic Communication

All course related correspondence, should be made through GMU email accounts. Please always use only your GMU email to contact the Instructor. You may not receive a reply immediately, but typically emails will be answered within 2 business days.

12 Course Mode

This is an asynchronous distance learning course. Weekly lessons and assignments will be posted on the Monday a week prior to that lesson's week. While you may complete lessons and assignments early, you must complete the lessons and assignments prior to the Sunday at the end of the designated week. (See schedule above to find the dates for each week's lesson.)

13 No Recording and No Distribution of Course Materials

Dr. Iasiello does not authorize in anyway the recording of any lecture content in this course beyond for student's exclusive personal use. Sharing of video lecture or lab content violates student privacy governed y the Family Education Rights and Privacy Act (FERPA). Additionally, any written, video, or audio content built by Dr. Iasiello for CDS 302 instruction or built by the TAs assigned to CDS 302 shared online external to GMU is a clear and punishable violation of GMU's Honor Code. This includes slides, videos, notes, assignments and exams.

Note: Recording is permitted only with the prior written consent of the professor or if recording is part of an approved accommodation plan.

14 Student Expectations

- 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 `http://academicintegrity.gmu.edu/distance`.
- Honor Code: Students must adhere to the guidelines of the George Mason University Honor Code. See `http://oai.gmu.edu/the-mason-honor-code/`.
- MasonLive/Email (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`.
- 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`.
- University Policies: Students must follow the university policies. See `http://universitypolicy.gmu.edu`. Responsible Use of Computing - Students must follow the university policy for Responsible Use of Computing. See `http://universitypolicy.gmu.edu/policies/responsible-use-of-computing`.
- University Calendar: Details regarding the current Academic Calendar. See `http://registrar.gmu.edu/calendars/index.html`.
- 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 `http://ods.gmu.edu`.
- Students are expected to follow courteous Internet etiquette at all times; see `http://www.albion.com/netiquette/corerules.html` for more information regarding these expectations.

15 Student Services

- University Libraries: University Libraries provides resources for distance students. See `http://library.gmu.edu/distance`.
- 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 `http://writingcenter.gmu.edu`. You can now sign up for an Online Writing Lab (OWL) session just like you sign up for a face-to-face session in the Writing Center, which means that you set the date and time of the appointment! Learn more about the Online Writing Lab (OWL).

- 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 `http://caps.gmu.edu`.

- 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 `http://registrar.gmu.edu/privacy`.

16 Disability Accomodations

Disability Services at George Mason University is committed to upholding the letter and spirit of the laws that ensure equal treatment of people with disabilities. Under the administration of University Life, Disability Services implements and coordinates reasonable accommodations and disability-related services that afford equal access to university programs and activities. Students can begin the registration process with Disability Services at any time during their enrollment at George Mason University.
If you are seeking accommodations for detailed information about the Disability Services registration process, Disability Services is located in Student Union Building I (SUB I), Suite 2500. Email:ods@gmu.edu — Phone: (703) 993-2474
If students have a documented learning disability or other condition that may affect academic performance, they must:

- Have the need for accommodation on file with Office of Disability Services, please visit `http://ds.gmu.edu/`

- Provide the Instructor with a copy of the Office of Disability Services accommodation determination prior to receiving any accommodations. The Instructor will closely protect this information as private and will not share the information with anyone other than the class assistants unless authorized in writing by the student or the Office of Disability Services.

- If you are having any difficulties with the course due to personal limitations, please discuss them with your instructor.

17 Diversity and Inclusion

The CDS Department seeks to create a learning environment that fosters respect for people across identities. We welcome and value individuals and their differences, including gender expression and identity, race, economic status, sex, sexuality, ethnicity, national origin, first language, religion, age and ability. We encourage all members of the learning environment to engage with the material personally, but to also be open to exploring and learning from experiences different than their own.

18 Sexual Harassment, Sexual Misconduct, and Interpersonal Violence

George Mason University is committed to providing a learning, living, and working environment that is free from discrimination and a campus that is free of sexual misconduct and other acts of interpersonal violence in order to promote community well-being and student success. We encourage students who believe that they have been sexually harassed, assaulted or subjected to sexual misconduct to seek assistance and support. University Policy 1202: Sexual Harassment and Misconduct speaks to the specifics of Mason's process, the resources, and the options available to students.

19 Notice of mandatory reporting of sexual assault, interpersonal violence, and stalking:

As a faculty member and designated "Responsible Employee," I am required 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 (703-380-1434) or Counseling and Psychological Services (703-993-2380). You may also seek assistance from Mason's Title IX Coordinator (703-993-8730; titleix@gmu.edu).

20 Student Support Resources

The following resources are available to students: Counseling and Psychological Services, The Learning Services Office or field-specific tutoring, The Office of Diversity, Inclusion, and Multicultural Education (ODIME), University Career Services, University Writing Center. Information and links regarding these and other student support offices are available on our Student Support Resources on Campus page.