Department Computational and Data Sciences
College of Science
4400 University Drive, MS 6A12, Fairfax, Virginia 22030
Phone: 703-993-5836; Fax: 703-993-9300; https://cos.gmu.edu/cds/

CDS 302

Scientific Data and Databases

1. General Information

Instructor:	Dr. Olga Gkountouna
Where:	Blackboard Collaborate (synchronous online mode)
When:	Tuesday, 1:30-4:10pm. Aug 24 - Dec 16, 2020
Course website:	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 (*please email me at ogkounto@gmu.edu*).

2. Course 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 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 system.

3. Learning Outcomes

By the end of the course each student will

- 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. Format

The course will be taught as a combination of lectures and discussions. There will be a Midterm and a Final exam, as well as several writing and coding assignments. For *online students*, the lecture will be broadcast through Blackboard Collaborate. Online students are welcome to log in and participate during class. The lectures will also be recorded. These videos will be available on Blackboard *for online students*.

5. Textbooks

No required textbook. Course slides and reading material will be provided via Blackboard.
Recommended book: Silberschatz, Korth, Sudarshan, "*Database System Concepts*", McGraw-Hill.

6. Technology Requirements

Software. This course will be using *LaTeX* as the main text editor for writing assignments. The use of *Overleaf* online LaTeX editor is strongly recommended. Alternatively, you may use MS Word. The course also will also be using SQL as a query language, and Python as a programming language. For practice in class you will need *SQLite*, *DBbrowser for Sqlite*, and *Anaconda Python 3.x*.
A basic understanding of computer programming principles and knowledge of any programming language or pseudo-code is desirable.
Hardware. You will need access to a Windows or Linux or Macintosh computer with at least 2 GB of RAM and to a fast, reliable broadband Internet connection (e.g., cable, DSL). To use the computers in the classroom, install SQLite and *DBbrowser for Sqlite* on a memory stick and run your code in that directory. 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.

7. Grades

Each assignment and written exam will be given a numerical grade on a 0-100 scale. Some assignments may include bonus tasks. At the end of the term all the marks will be totaled as a weighted average according to the following weights:

Assignments	**50%**
Midterm Exam	**20%**
Final Exam or Project	**20%**
Classroom participation	**10%**

Please note that in general all assignments may not have the same weight. The weight of each individual assignment will be indicated on the assignment form.

8. Exams

The course includes mandatory written *midterm* and *final* exams. The material covered in the exams will be announced in class. Students may choose to deliver a database application *project* in place of the *final* exam, with the permission of the instructor. A student who cannot write a course examination or 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.
Online exams require **RESPONDUS** software. Please install it today.

9. 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.

10. Course outline (tentative)

In this course we will cover the following topics. Please note that the topics and their order are subjected to change at the discretion of the instructor. This class will not follow a rigid format, but will instead adjust to the talent levels and interests of the students.

Date	Lect. #	Topic	Assignment (announcement dates)
08/25	1	Introduction: Why Databases? Course overview.	
		Scientific Writing: Introduction to LaTeX and OverLeaf	
09/01	2	LaTeX: Math Notations, Equations, Tables, References	Assignment 1: Writing
		How to read a scientific Article	
09/08	3	The Relational Model: DB schema, keys, diagrams and Relational Algebra	Bonus assignment: scientific article summary
		Introduction to SQL: basic queries – Part I	
09/15	4	SQL Part II: string operators, set operations, aggregate functions, ordering, grouping, having-clause	Assignment 2: SQL queries
		SQL practice in class (SQLite) for Parts I & II	
09/22	5	SQL Part III: nested queries, set membership, joins, self-joins, with-clause, views, stored procedures	Assignment 3: advanced SQL queries
		SQL practice in class (SQLite) – Part III	
09/29	6	SQL practice in class (SQLite) – Part IV	
		Exam Preparation	
10/06	7	**MIDTERM EXAM**	
10/13	8	The Entity Relationship Model & Relational DB Design	Assignment 4: E-R Model
		Reduction of E-R to Relational Schema	
10/20	9	Storage and File Structure – Overview	
		Indexing and Hashing: Basics, Ordered Indices	
10/27	10	Indexing: Dense vs. Sparse Index, Multilevel index,	Assignment 5: E-R to
		B^+-Tree structure; B+-Tree insertions, deletions, updates	Relational Model
11/03	11	**Election Day: No classes**	
11/10	12	Indexing: B-Tree; Bitmap Index; Multiple-Key Access	
		Static Hashing, Dynamic Hashing; Hash file organization	
11/17	13	Query Processing: Selection Operation	Assignment 6: Indexing
		Point and Range Queries	
11/24	14	Query Processing: Join Operations, Other Queries	
		Connecting to a DB, run SQL Queries through Python	
12/01	15	Create & Modify Tables through Python;	Assignment 7: Python
		Practice Examples	
12/08	16	**Exam Preparation**: solutions of midterm, assignments	
		Discussion, Q/A	
12/15	17	**FINAL EXAM**	

11. Late paper submission:

The late submission policy is the following:

- Submissions within the first day (24 hours) after the deadline: -**5**% credit.
- Submissions within the second day after the deadline: -**10**% credit.
- Submissions within the third day after the deadline: -**30**% credit.
- No assignments after the third day (i.e., 72 hours after the deadline) will be accepted.

Exceptions to this policy may be granted, given *serious circumstances* and *at the discretion of the Instructor*.

12. General guidelines for ASSIGNMENT preparation and submission

a. Grades of assignments will be based on:
- **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.

b. Please remember that your assignment is a professional document, and should therefore be formatted and constructed accordingly. All assignments are to be typed. Hand-written assignments will not be accepted.

c. Submission of your answers will be made exclusively through Blackboard. Please do **not** email your assignments.

d. The electronic submission of your assignment report has to be in **PDF format**, unless stated otherwise.

e. Each assignment submission should include a cover page with the following information: assignment title, assignment number, student name, and submission date.

f. Please make sure you have a backup of all the materials you submit.

13. 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.

14. 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.

15. No Recording and No Distribution of Course Materials policy:

The Instructor 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 by the Family Education Rights and Privacy Act (FERPA). **Additionally, any written, video, or audio content built by the Instructor 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.

16. 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.

17. Student Services:

- **University Libraries:** University Libraries provides resources for distance students.
  [See http://library.gmu.edu/distance and http://infoguides.gmu.edu/distance_students].
- **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].

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

As a faculty member, I am designated as a "Responsible Employee," and must report all disclosures of sexual assault, interpersonal violence, and stalking to Mason's Title IX Coordinator per University Policy 1202. If you wish to speak with someone confidentially, please contact one of Mason's confidential resources, such as Student Support and Advocacy Center (SSAC) at 703-380-1434 or Counseling and

Psychological Services (CAPS) at 703-993-2380. You may also seek assistance from Mason's Title IX Coordinator by calling 703-993-8730, or emailing titleix@gmu.edu.

Disclaimer: 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.

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