INLS 773 | Databases for Data Science (1.5 credit)

School of Information and Library Science | University of North Carolina at Chapel Hill

TARGET AUDIENCE: Students admitted to the Certificate in Applied Data Science program
PREREQUISITES: None
SPAN: Full semester, one lesson per week
INSTRUCTION MODE: Remote-Asynchronous

Course Description

This course will introduce the basic concepts and implementations of relational database management systems suited for data science applications. Topics include user requirements and specifications, ER models, database programming including SQL, data quality, and applications.

Instructor Information

DR. ARCOT RAJASEKAR (He/Him/His)
Professor, UNC School of Information and Library Science;
Director of Research & Technology, Data Intensive Cyber Environments (DICE) Center;
Chief Scientist, Renaissance Computing Institute (RENCI)
Phone: 858-205-4357 **||** Email: rajasekar@unc.edu
Office: Manning Hall 21; RENCI 556A
Office hours: By Email appointment

Key Learning Objectives

This course focuses on aspects of database design, implementation, and use that are an integral part of data science and analytics. In this course, students will
- Gain experience in transforming user requirements into a well-formed, functional database. (Database Design, ER Models)
- Be able to create, modify and manipulate relational databases and pose complex queries (SQL DDL and DML)
- Be able to support database access through other programming languages
- Understand issues of data quality and data cleansing.
- Have hands-on experience with Data Science projects.

Course Resources:

The following book is the suggested textbook for the class. Many of the examples in the lectures are taken from that book:
AVI SILBERSCHATZ, HENRY F. KORTH, S. SUDARSHAN. DATABASE SYSTEM CONCEPTS SEVENTH EDITION, MCGRAW-HILL, ISBN 9780078022159.

Instructional Methods

This is an asynchronous online course. All enrolled students should have access to the UNC Sakai site for this course: http://sakai.unc.edu/

COURSE MATERIALS: All materials can be found in Sakai. The course syllabus, schedule, and information about tools and other resources will be there at the beginning of the semester.

ASSIGNMENTS: All Assignments, and Project Deliverables must be submitted using the Sakai Assignments tool. In my experience, Sakai is a reliable method for submitting assignments. It is the responsibility of each student to make sure they have access to Sakai and can submit assignments when they are due.

Assessment:

Homework Assignments (4 deliverables) (4 x 10 = 40%)

Term Project 25%

Final Exam 25%

Forum Participation 10%

Course Policies

COMMUNICATION WITH INSTRUCTORS

The primary mode of communication with me will be through the Sakai messaging tool or email (rajasekar@unc.edu), which I check frequently throughout the day. You should expect a reply in less than 24 hours. If for any reason I will be away from email for more than 24 hours, I will let you know in advance. Email messages can be sent to me via the course website or directly using your preferred email client.

COMMUNICATION WITH YOUR CLASSMATES

In addition to classroom discussion, there is a "General Discussion" forum section. Please use this section for any topics that do not directly relate to the content of the course but relate to your course experience. You can also contact each other via email using either the message feature of Sakai or your preferred email client. Please use common email etiquette norms when corresponding with each other.

COURSE PARTICIPATION

It is very important that you take charge of your own learning and participate actively in the course activities. Forum participation is graded and everyone is expected to join actively through postings and replies to other postings every week. Participate every week with a posting relevant to the week's topic in the fora; I am very strict about this. If you don't participate sufficiently and on time (at least for 10 weeks), you will lose all the points. Full points if you participate every week.

PENALTY FOR LATE ASSIGNMENTS

Assignments are expected to be submitted on time. This is important for at least two reasons:
1. the need to meet deadlines is a reality of professional life, and
2. giving some people more time for an assignment than others in the class is not equitable.

However, life happens to all of us at one time or another. If you cannot meet an assignment deadline, please tell me why prior to the due date. We will negotiate a new deadline when you have a valid reason for needing an extension. Otherwise, late assignments will drop 10% for each day late.

Grading Policies

The following grade scale will be used AS A GUIDELINE (subject to any curve) **for undergraduate** students:

Grade Range	Definition*
A or H (Grad) 90-100%	Mastery of course content at the highest level of attainment that can reasonably be expected of students at a given stage of development. The A or H grade states clearly that the students have shown such outstanding promise in the aspect of the discipline under study that he/she may be strongly encouraged to continue.
B or P 80-89.9%	Strong performance demonstrating a high level of attainment for a student at a given stage of development. The B or P grade states that the student has shown solid promise in the aspect of the discipline under study.
C or L 70-79.9%	A totally acceptable performance demonstrating an adequate level of attainment for a student at a given stage of development. The C of L grade states that, while not yet showing unusual promise, the student may continue to study in the discipline with reasonable hope of intellectual development.
D or L 60-69.9%	A marginal performance in the required exercises demonstrating a minimal passing level of attainment. A student has given no evidence of prospective growth in the discipline; an accumulation of D or L grades should be taken to mean that the student would be well advised not to continue in the academic field.
F or F 0-59.9%	For whatever reason, an unacceptable performance. The F grade indicates that the student's performance in the required exercises has revealed almost no understanding of the course content. A grade of F should warrant an advisor's questioning whether the student may suitably register for further study in the discipline before remedial work is undertaken.

*Definitions are from: http://registrar.unc.edu/academic-services/grades/explanation-of-grading-system/ (underlining is my emphasis)

Plus and minus scores for undergraduates, when given, use the following range:

A	93 and higher
A-	90-92
B+	88-89
B	83-87
B-	80-82
C+	78-79
C	73-77
C-	70-72
D+	68-69

D	63-67
D-	60-62
F	59 and below

University Honor Code

Chapel Hill has had a student-administered honor system and judicial system for over 100 years. Because academic honesty and the development and nurturing of trust and trustworthiness are important to all of us as individuals, and are encouraged and promoted by the honor system, this is a most significant University tradition. More information is available at http://studentconduct.unc.edu/honor-system

The UNC Honor Code is in effect for all work in this course. When work or ideas are not your own, you must attribute them. Unless otherwise stated, all assignments in this class are individual assignments, meaning that the substance of the work you turn in must be your own. If you have any doubts or questions about a course of action or a specific situation, please ask for clarification. Students should NOT receive (or give) major creative assistance or ongoing minor support on individual assignments. If you have any questions about this, please ask me. Section II. B. of the "Instrument of Student Judicial Governance" gives examples of actions that constitute academic dishonesty: https://studentconduct.unc.edu/sites/studentconduct.unc.edu/files/documents/Instrument.pdf

Students often ask what is okay to talk about with other students and what is not. There are some specific guidelines for this course.

- I do encourage you to help each other learn the course material – your fellow students can often be a great resource for learning. For example, you could watch one of the videos together, and discuss the material it presents. However, you should not discuss the details of a solution to an ongoing assignment with other students and should never copy or share answers for an assignment with other students. It is okay to talk about course material with other students, but you should not discuss solutions for pending assignments.
- All work you submit should be your own.
- You may give and receive assistance regarding the use of hardware and software. For example, you may ask or answer a question such as "how do I [fill in the blank] in SQLite?"). A question such as "Should I have a separate table to represent cats and their information?" should be addressed to me.
- Individual homework assignments are to be done **individually**. You may consult the course materials, your notes, and even other print or web sources. (Keep in mind, however, that what you find in other sources may not be consistent with what I want you to do.) You may not consult your classmates or other people; all questions should be addressed to me.
- You must sign (check) the honor statement when you submit each assignment. This confirms that you and the work conform to the Honor Code.

Plagiarism is a serious violation of the Honor Code. To become more familiar with the issues surrounding plagiarism, and how to best avoid this academic issue, view this brief Plagiarism Tutorial created by the librarians of UNC-Chapel Hill, Duke University, NC State University, and NC Central University. If you have any questions about what constitutes plagiarism or how to properly cite a source, please contact your instructor.

SILS Diversity Statement

In support of the University's diversity goals and the mission of the School of Information and Library Science, SILS embraces diversity as an ethical and societal value. We broadly define diversity to include race, gender, national origin, ethnicity, religion, social class, age, sexual orientation and physical and learning ability.

As an academic community committed to preparing our graduates to be leaders in an increasingly multicultural and global society we strive to:

- Ensure inclusive leadership, policies, and practices;
- Integrate diversity into the curriculum and research;
- Foster a mutually respectful intellectual environment in which diverse opinions are valued;
- Recruit traditionally underrepresented groups of students, faculty and staff; and
- Participate in outreach to underserved groups in the State.

The statement represents a commitment of resources to the development and maintenance of an academic environment that is open, representative, reflective and committed to the concepts of equity and fairness.

Special Accommodations

If any student needs special accommodations, please **contact the instructor during the first week of classes**. The University of North Carolina at Chapel Hill facilitates the implementation of reasonable accommodations, including resources and services, for students with disabilities, chronic medical conditions, a temporary disability or pregnancy complications resulting in difficulties with accessing learning opportunities. All accommodations are coordinated through the Accessibility Resources and Service Office. See the ARS Website for contact information: https://ars.unc.edu or email ars@unc.edu.

Counseling and Psychological Services

CAPS is strongly committed to addressing the mental health needs of a diverse student body through timely access to consultation and connection to clinically appropriate services, whether for short or long-term needs. Go to their website: https://caps.unc.edu/ or visit their facilities on the third floor of the Campus Health Services building for a walk-in evaluation to learn more. *(source: Student Safety and Wellness Proposal for EPC, Sep 2018)*

Title IX Resources

Any student who is impacted by discrimination, harassment, interpersonal (relationship) violence, sexual violence, sexual exploitation, or stalking is encouraged to seek resources on campus or in the community. Please contact the Director of Title IX Compliance (Adrienne Allison – Adrienne.allison@unc.edu), Report and Response Coordinators in the Equal Opportunity and Compliance Office (reportandresponse@unc.edu), Counseling and Psychological Services (confidential), or the Gender Violence Services Coordinators (gvsc@unc.edu; confidential) to discuss your specific needs. Additional resources are available at safe.unc.edu.

Library Services and Course Reserves

Students enrolled in this course have access to the UNC Library System. Visit Distance Education Library Services to access a wide array of online services and resources including Course Reserves, online databases, online journals, online books, and live help with research and library access. **Most online resources require you to log in with your**

Onyen and password. If you have any trouble finding the resource that you need or logging in to a resource, you can contact the library through the contact information at Distance Education Library Services. You can chat live about your problem or send an email to request assistance.

Tentative Schedule

We will follow the tentative schedule given below. Though it is for only 1.5 credits, the span of the course is for the entire semesters and hence one lesson (class) per week is scheduled. Schedule in sakai will provide concrete dates for the semester timeline.

Lesson Week	Concepts	HW Assigned	Project Dates	HW Due
1	DB Concepts	HW1: SQLite/Concepts		
2	ER Models			
3	ER Diagrams			HW1
4	ER Diagrams (Contd.)	HW2: ERD		
5	ER2DB			
6	SQL-DDL		Project Assigned	HW2
7	SQL-DML	HW3: ER2DB		
8	DML examples			
9	Advanced SQL			HW3
10	More DML Structures	HW4: SQL		
11	ASQL Examples			
12	Language Support			HW4
13	Data Quality Issues			
14	Applications: OLAP/OLAM		Project Due	
15	Overview and Wrap Up			
16	EXAM			