BAX 421 – 002: Data Management

TERM: Fall 2022

LECTURES: Odd Saturdays: 1:30 p.m. – 5:20 p.m.

INSTRUCTOR: Mehul Rangwala mrangwala@ucdavis.edu

OFFICE HOURS: I will be available by appointment at mutually convenient times.

TA DISCUSSION SECTIONS: We have two TAs for this course. Each will have an hour of discussion session via Zoom every week. Additional details will be shared on Canvas.

COURSE DESCRIPTION: Introduction to the extraction, assembly, storage and organization of data in IT systems. The course covers the concepts of data modeling, entity relationship models, and SQL to help businesses convert data into insights needed to drive business strategies. Use of MySQL and Microsoft SQL Server for writing SQL queries.

PREREQUISITES: None.

REFERENCE TEXTBOOKS:

1. SQL: The Complete Reference, 3rd Edition
ISBN-13: 978-0071592550
ISBN-10: 0071592555

2. Murach's MySQL, 3rd Edition
ISBN 978-1-943872-36-7

3. Murach's SQL Server 2019 for Developers
by Bryan Syverson and Joel Murach
ISBN 978-1-943872-57-2

4. Graph Databases
by Ian Robinson, Jim Webber and Emil Eifrem (free download)
https://neo4j.com/lp/book-graph-databases/?utm_program=na-prospecting&utm_source=google&utm_medium=cpc&utm_campaign=na-pm-offer-content&utm_adgroup=ebook-oreilly&utm_placement=&utm_keyword=&utm_network=x&gclid=Cj0KCQjwj7CZBhDHARIsAPPWv3edyz0fjdTyZ11L50VxSgxtu8fr1bfpfoexZqvcm7IVTJCzbDrHeRMaAox7EALw_wcB

NOTES AND HANDOUTS: They will be available on Canvas.

TOPICS TO BE COVERED: A detailed schedule is available at the end of the syllabus.

COMPUTER PACKAGES:

1. **MySQL**. You can either use the MySQL command line console or install MySQL Workbench.

2. **Microsoft SQL Server 2019** with **Microsoft SQL Server Management Studio.**

3. Neo4j Graph desktop or Cloud version.

GRADING:

Homework (Individual)	20%
Midterm Exam	40%
Final Exam	40%

GRADING RULES: Clerical scoring errors will be corrected without hassle, but for other re-grades you must hand back the work and submit an email request; the entire assignment will be subject to re-grading. You must submit any re-grading requests via email within 5 calendar days from when the assignment is returned. In your email, you should clearly explain why you are requesting a re-grade. While I will consider the specific concerns cited in your email, I will re-grade the entire assignment. Your new score might be higher, lower, or the same as a result. Please remember that small changes in your grade on a single assignment might not affect your overall course grade.

LATE SUBMISSION: Late/email submission due to whatever reason will carry a 20% penalty. Assignments which are submitted after the grades are released and after the solutions are posted will not be accepted under any circumstances. Please accept these penalties without wasting time in negotiating with me and convincing me why the assignment should be considered for grading.

Learning Objectives:

1. Understand the fundamentals of relational database modeling and database normalization.
2. Learn the concepts of Structured Query Language (SQL) and evaluate how it can be used to retrieve and transform data from relational databases.
3. Retrieve data from the database using SQL joins, grouping, subqueries, aggregate functions, and window functions.
4. Understand the principles of graph databases and write queries.
5. Understand the principles of data management and data governance.

Academic Honor Code:

All students are expected to adhere to the University of California, Davis' Code of Conduct as noted here: http://sja.ucdavis.edu/files/cac.pdf.

Schedule on the Next Page

Schedule (Tentative): This is a tentative schedule. Contents and sequence are subject to change.

Session 1 = 1:30 pm – 3:30 pm, Session 2 = 3:40 pm – 5:20 pm

Date	Assignments Due	Topics Covered
9/24/2022		Relational Modeling Concepts (Session 1) • Types of databases • Principles of good database design • Data modeling building blocks • Types of relationships (1:1, 1:M, M:N) • Types of keys • Integrity rules • ER Model basics • Database normalization basics Basic SQL (Session 2) • Single table SELECT statements • Multi-table queries (JOINS) • UNIONs
10/8/2022	Homework 1	Basic SQL (carry over from 9/23/2022) (Session 1) SQL Aggregate Functions (Session 1) • Summary queries • Grouped queries Subqueries (Session 2) • Why use subqueries? • Types of subqueries • Using subqueries • Subqueries versus Joins
10/22/2022	**Midterm exam (in-class)**	**Midterm exam – live coding (Session 1)** Data manipulation (Session 2) • String functions for cleaning data
11/5/2022	Homework 2	SQL Window functions (Session 1) Graph databases – Part 1 (Session 2)
11/19/2022	Homework 3	Graph databases – Part 2 (Session 1) Principles of data management and data governance
12/10/2022	**Final Exam (in-class)**	**Final exam – live coding**