Course syllabus

Overview

This course will teach students the fundamentals of relational databases and data management. Students will learn the principles of database modeling and design and gain practical experience applying SQL (Structured Query Language) to manage and manipulate relational databases. The course also introduces the role and application of data documentation and metadata standards for interoperability and effective data management. By the end of the course, students will be equipped to make informed decisions about managing databases and data ethically and responsibly, focusing on issues such as bias, data privacy, sharing, ownership, and licensing.

Learning objectives

- Understand the fundamental principles of relational databases and relational data modeling, including table structures, primary and foreign keys, relationships between tables, and data normalization.
- Understand how to use the Unix command line and how manage DuckDB databases from the command line.
- Use SQL to retrieve, manipulate, and manage data stored in a relational database.
- Demonstrate proficiency in querying, filtering, sorting, and programmatically accessing and interacting with relational databases from R and Python.
- Become familiar with advanced database topics such as concurrency, transactions, indexing, backups, and publication.
- Understand the role of good data documentation and metadata standards for interoperability, effective data management, and reproducibility.
- Operationalize the FAIR principles into data management practices.
- Produce a metadata record in EML (Ecological Metadata Language) and apply metadata crosswalks to programmatically convert metadata schemas.
- Understand the ethics of sensitive data and how to de-identify sensitive data.
- Evaluate ethical and responsible data management practices, including bias, data privacy, sharing, ownership, and licensing issues.

Schedule

- Class: Tuesday & Thursday 9:30-10:45 AM (BH 1424)
- Discussion - session 1: Thursday 1-1:50 PM, BH 3022A
- Discussion - session 2: Thursday 2-2:50 PM, BH 3022A
- Office hours: TBD
- Best way to contact use: email

Modules

Week	Topic/Content
1	Relational databases and data modeling
2	Analyzing & cleaning the bird dataset (from csv)
3	Introduction to SQL (part 1) & DuckDB
4	SQL part 2 + Analyzing the bird database using SQL
5	I/O & data management + Advanced database topics (indexing, triggers, …)
6	Using R to query databases + bash programming
7	Using Python to query databases + Documenting your work: metadata
8	Capturing your computing environment + Ethical & responsible data mgnt
9	Sensitive data
10	Data licensing and publication

*Schedule subject to change.

Course assessment

Your performance in this course will depend 90% on weekly homework assignments and 10% on class participation. There will be no graded exercises or homework on week 9 & 10 due to capstones. We will use Canvas to manage the assignments.

Homework policy

Homework is expected to be turned in on time. Do not wait until the last minute to work on homework in case something unexpected comes up. To be considered, a request for an extension must be submitted within business hours before the homework deadline. **_Homework turned in late will be docked 20% per day._** Peer-to-peer learning is encouraged, but each submission needs to be individual, which means students are expected to be able to explain all their answers in detail.

Acknowledgment of tools used

The use of AI assistant technologies falls within the purview of the Student Conduct Code and the Student Guide to Academic Integrity. It states that "Materials (written or otherwise) submitted to fulfill academic requirements must represent a student's efforts unless otherwise permitted by an instructor." Therefore, we expect students to disclaim the use of any coding assistant or LLMs (such as ChatGPT) and to write and understand every line of code that is submitted as part of the assignments. We expect those tools to be used similarly to how Google or StackOverflow could be. In other words, copy-pasting answers without taking the time to understand and modify the suggested code to make it yours is not an acceptable practice. As you are building your skills in SQL and databases, we recommend refraining from the use of AI and keeping to practicing your skills and favoring peer-to-peer learning.

Resubmission

All students can resubmit their assignments three days after they have received initial feedback. In this second submission, students may recover up to 50% of the points not obtained during the initial submission. Revisions, corrections, and improvements are crucial in the learning process! We greatly encourage you to resubmit your revised assignments.

Attendance policy

Attendance is required. Your presence and active participation will make a real difference, not just in this course, but in your future environmental data science career. The connections you build and the skills you sharpen here will help you succeed long after the quarter ends. Let's make the most of it together! In addition not all the materials and interactive activities given in class will be covered by slides or background readings.

Missing a class

Life happens. Similar to missing a day at work, if you miss a class or discussion section then you are expected to:

- Be proactive: and notify your instructor as soon as possible and provide a brief explanation. Documentation is not required for short-term absences, but honesty and professionalism is expected in all communication.
- Catch up: You'll need to work with the instructor or the TA to review any missed material.
- Stay home when you are sick! If you are feeling unwell, prioritize your health. Stay home and rest or work remotely.

The following accommodations for missing in-person activities can be provided upon student request:

- Lesson materials will be made available online the same day
- Consultation outside office hours

Students with disabilities should work with the Disabled Students Program (DSP) to determine appropriate accommodations.

Attendance Concerns

- Frequent Absences: If you miss two or more consecutive classes, the instructor may reach out to check in. This is to ensure you are ok and aware of resources that can help you stay on track.
- Extended Absences: If you anticipate missing three or more classes, you'll need to contact the instructor, the MEDS Program Coordinator, and the Director of Student Affairs to discuss possible next steps. UCSB courses are taught in person, so extended absences may require a Leave of Absence, regardless of the attendance impact on your grade.
- Academic Dishonesty in Attendance Reporting: Fabricating reasons for absences is considered academic dishonesty and subject to disciplinary action under the UCSB Student Code of Conduct. If the instructor suspects this is the case, they will address it with you directly and may escalate it if needed.

Code of conduct

All students are expected to read and comply with the UCSB Student Conduct Code. We expect cooperation from all members to help ensure a welcoming and inclusive environment for everybody. We are determined to make our courses welcoming, inclusive and harassment-free for everyone regardless of gender, gender identity and expression, race, age, sexual orientation, disability, physical appearance, or religion (or lack thereof). We do not tolerate harassment of class participants, teaching assistants, or instructors in any form. Derogatory, abusive, or demeaning language or imagery will not be tolerated.

Student support

We understand that ongoing crises impact students differently based on experiences, identities, living situations and resources, family responsibilities, and unforeseen challenges. We encourage you to prioritize your well-being. We are here to help you reach your learning and career goals. You are always welcome to reach out to our teaching team so that we can best

support you. Please see the UCSB Campus Resource Guide for campus student support and services.

Disabled students program

Students with disabilities and/or alternative learning needs are encouraged to work with the Disabled Students Program at UCSB to ensure we can best support your learning and success.