Database Management Systems (COSI 127B)

Lectures: **Tue/Thu 3:55 PM – 5:15 PM** @ **Abelson-Bass: 131**
Recitation*: **Wed 12:20 PM – 1:10 PM** @ **Gerstenzang: 122**
Class Website | Piazza | Gradescope | Moodle

Course Description

We live in a data-driven world! In fact, every two days, we create as much data as we created from the dawn of humanity up to 2003 [Eric Schmidt, *former Google CEO*]. We generate an unprecedented amount of data every day through our daily activities in every facet of life, including business, governance, management, research, and scientific activities. Storing, managing, and accessing this large volume of data is an enormous challenge in the face of ever-changing application requirements and performance goals. Database systems are the backbone of any large-scale data management infrastructure. This class presents a comprehensive introduction to the fundamental concepts of database systems. We will start by discussing the classical approaches to database design and the fundamental building blocks and operating principles of database systems. The primary focus of the course will be on the core concepts of the internals of database systems, covering entity-relationship and relational data models, commercial relational query languages (SQL and relational algebra), file organization, storage and memory management, indexing and hashing, query optimization, query processing, transaction processing, concurrency control, and recovery. Finally, we will cover the new trends in data management in the era of big data and data management in the cloud to highlight the evolution of database systems over the years.

Prerequisites

COSI 21A and **COSI 12B**. A working knowledge of C/C++, Java, or Python programming, along with a fundamental understanding of data structures and algorithms, is required.

Instructor & Teaching Assistant(s)

	Contact Information	**Student Hours**
Subhadeep Sarkar Instructor	Office: Volen 259 Email: subhadeep@brandeis.edu	Tue/Thu: 10:00 AM – 11:00 AM
Shubham Kaushik Teaching assistant	Office: Volen 104 (Vertica Lounge) Email: kaushiks@brandeis.edu	Wed: 2:00 PM – 4:00 PM
Steven Yang Teaching assistant	Office: Volen 104 (Vertica Lounge) Email: sarahbaskin@brandeis.edu	Fri: 10:30 AM – 12:00 PM
Sarah Baskin Teaching assistant	Office: Volen 104 (Vertica Lounge) Email: stevenyang@brandeis.edu	Mon: 2:15 PM – 3:45 PM

*** Attending the Recitation sessions is key to being successful in the class.**

Class Resources & Contact Policy

The class website is updated in real-time with the class. **This is your go-to** for any general resources about the class. This is also where all **written and programming assignments will be posted**, along with other relevant resources for the class.

Piazza will be the preferred mode of communication with the instructor and teaching assistants (TAs) outside of class hours and student hours. Please register to Piazza, if you are not enrolled already (access code: **COSI127BSP26**). Posts in Piazza will be responded to within 24 hours. Throughout the course, students are encouraged to ask conceptual questions and start discussions through Piazza. Other students should feel free to answer any questions and participate in any discussions on Piazza. The instruction team will go through Piazza questions the same day that student hours are held.

Emails sent during weekdays should expect a response within 48 hours. Be aware that not receiving a response is not a valid reason to not hand in homework on time, so start your work early to make sure you have no questions.

We will be using Gradescope for grading assignments. Please register to Gradescope, if you are not enrolled already (access code: **4D2P6V**).

Required Textbook

Database Management Systems. R. Ramakrishnan and J. Gehrke. Third Edition. McGraw-Hill.
In the class, we will also cover a few topics from recent research and survey papers.

Additional recommended readings and/or third-party instruments (optional)
The following are excellent sources for additional reading.

1. *Architecture of a Database System*. J. Hellerstein, M. Stonebraker, and J. Hamilton. Foundations and Trends in Databases. 2007.
2. *The Design and Implementation of Modern Column-store Database Systems*. D. Abadi, P. Boncz, S. Harizopoulos, S. Idreos, and S. Madden. Foundations and Trends in Databases. 2007.
3. *Modern B-Tree Techniques*. Goetz Graefe. Foundations and Trends in Databases. 2011.

Learning Goals

Students who successfully complete all components of this course will be able to demonstrate the following by the end of the semester.

a) Familiarization with the history and evolution of database management systems design over the past decades.
b) Understanding of the importance and challenges associated with large-scale data management and analysis.

c) Knowledge about the key components of a relational database system and their working principle.
d) Ability to interact with relational data stores using a working knowledge of SQL.
e) Understanding of the fundamental performance tradeoffs in data structures and access methods design, and the importance of such design decisions on the overall performance of a relational data system.
f) Understanding of the impact of the ever-changing workload and application requirements on the design of modern relational data systems.

Components of Course Work

Success in this four-credit course is based on the expectation that students will spend a minimum of 9 hours of study time per week in preparation for class.

A. *Class Participation and Interaction:* The class is designed to be interactive. The students are highly encouraged to ask questions and participate in in-class discussions. All lectures are delivered in person. The students are expected to attend all lectures and regularly interact with the teaching staff during student hours.
B. *Programming Assignments:* During the semester, there will be 2-3 programming assignments (Project 1 has 3 sub-parts). The projects will be based on topics taught in class.
C. *Written Assignments:* Approximately every two weeks, there will be one written assignment based on concepts taught in class. The assignments will be graded based on correctness.
D. *Midterms:* There will be two midterm exams for this class. Your success in the midterm exams is subject to (i) attending the class regularly, (ii) submitting all assignments on time to receive feedback, and (iii) **attending the recitation sessions**.

Group work and collaboration are essential to your learning. That said, all of your written work will be solo assignments and must be your original work and produced without the assistance of others unless I specifically state otherwise.

Evaluation and Grading

The breakdown of the course grade is as follows (minor alterations may occur).

Class Element	**Grade Percentage**
In-class participation	*5%*
Written assignments	*15%*
Programming assignments	*30%*
Midterm I	*20%*
Midterm II	*30%*

Late Policy

Every student is granted a cumulative allowance of 4 late days, serving as extensions for their individual assignments, and no penalties are incurred during this period. These late days can be utilized across various assignments. However, once the allotment of 4 late days is exhausted, a penalty of 20% per day is imposed on overdue homework assignments.

Tentative Schedule

Week #	Topic	Readings
1	Introduction & Data Systems Architectures Essentials	Chapter 1
2	ER Model & Relational Model	Chapters 2, 3
3	Relational Algebra & SQL	Chapters 4, 5
4	File & Storage Organization	Chapters 8, 9
5	Indexing; Hashing and B-Trees	Chapters 10, 11
6	Advanced Indexing & External Sorting	Chapter 13
7	Review & Midterm I	
8	Query Processing	Chapters 12, 14
9	Query Optimization	Chapter 15
10	Transactions	Chapter 16
11	Concurrency Control & Recovery	Chapter 17
12	Functional Dependencies & Schema Normalization	Chapter 19
13	Review & Midterm II	Chapter 18

Important Course Policies

Academic honesty

You are expected to be familiar with, and to follow, the University's policies on academic integrity. You are expected to be honest in all of your academic work. Please consult Brandeis University Rights and Responsibilities for all policies and procedures related to academic integrity. Allegations of alleged academic dishonesty will be forwarded to Student Rights and Community Standards. Sanctions for academic dishonesty can include failing grades and/or suspension from the university. Citation and research assistance can be found on the university library website.

Accommodations

Brandeis seeks to create a learning environment that is welcoming and inclusive of all students, and I want to support you in your learning. If you think you may require disability accommodations, you will need to work with Student Accessibility Support (SAS). You can contact them at 781-736-3470, email them at access@brandeis.edu, or visit the Student Accessibility Support home page. You can find helpful student FAQs and other

resources on the SAS website, including guidance on how to know whether you might be eligible for support from SAS.

If you already have an accommodation letter from SAS, please provide me with a copy as soon as you can so that I can ensure effective implementation of accommodations for this class. In order to coordinate exam accommodations, ideally you should provide the accommodation letter at least 48 hours before an exam.

Respectful environment
Brandeis University is committed to providing its students, faculty, and staff with an environment conducive to learning and working, where all people are treated with respect and dignity. Please refrain from any behavior toward members of our Brandeis community, including students, faculty, staff, and guests, that intimidates, threatens, harasses, or bullies.

Laptop computer and cell phone use
To create a focused and distraction-free learning environment, we have implemented a no laptop/mobile policy during the class. This policy is aimed at maximizing student engagement, fostering active participation, and promoting a conducive atmosphere for effective learning. By minimizing the use of laptops and mobile devices, we aim to enhance the overall quality of the learning experience and encourage direct interaction with course materials and fellow students.

Use of generative AI tools
The use of generative AI tools is not permitted in this course, specifically for completing the written and programming assignments. Unauthorized use of generative AI will be treated as academic misconduct. The class is designed as highly interactive, with numerous in-class discussions and student hours every day of the week. Students are highly encouraged to participate in discussions during class and interact with the teaching staff during student hours – this is critical to build a sound understanding of the course material.

Financial aid for purchasing course materials
If you have difficulty purchasing course materials, please make an appointment with your Student Financial Services or Academic Services advisor to discuss possible funding options and alternative solutions.

Student Support
Success in this course depends heavily on your personal health and well-being. Recognize that stress is an expected part of the college experience, and it often can be compounded by unexpected setbacks or life changes outside the classroom. Your other professors and I strongly encourage you to reframe challenges as unavoidable pathways to success. Reflect on your role in taking care of yourself throughout the academic year, before the demands of exams and projects reach their peak. Please feel free to reach out to me about difficulties you may be having that may impact your performance in this course as soon as it occurs and before it becomes too overwhelming.