San José State University
Department of Computer Science
Fall Semester 2022
CS 257 Database System Principals, Section 2
Syllabus

Course and Contact Information

Course Title:	**Database System Principals**	
Instructor:	**Professor Ahmed Ezzat, Ph.D.**	**e-Mail:** Ahmed.Ezzat@sjsu.edu
Date:	**Fall 2022 (51076)**	
Course Number:	**CS 257**	
Credit Hours:	**3 Credit Hours**	
Classroom:	MacQuarrie Hall, room 233	
Office Hour:	Duncan Hall, room 282	
Class Schedule:	**Every Tuesday + Thursday: 7:30pm 8:45pm**	
Office hour Schedule	**Every Tuesday (in DH-282): 6:20p**	

NoSQL..

The course is structured around three goals:

1. Cover advanced database topics from the text book.
2. Cover an overview of active research topics in the database area from published papers.
3. **Homework:** are individual assignments.
4. **Projects:**
    a. **Individual Project:** each student will create an RDBMS personal development environment, e.g., using their own laptop or an alternative.
    b. **Group Project:** each group will do regular group project and Final group project. These 2 projects are meant to help you in two main areas:
        i. Design RDBMS schema and implement a shopping cart application (Group Project-2)
        ii. Become familiar with executing SQL on Hadoop using Hive, and Good understanding and compare the query optimization material in the class against the query plan generated by a mature database such as Oracle (Final Group Project)
5. **Group Research Paper:** Deeper Group Research by the students in any of the topics covered at high-level in item (2) and using optionally the list of publications supplied at the end of the Syllabus. Toward that goal, the students will be divided into set of small groups. Each group will select topic from the advanced research topics listed below. Each group will write a formal paper about their assigned topics and gives a presentation to the rest of the class describing: the problem being investigated, what others have done, and the group own views and perspective. List of publications in the relevant topics are provided at the end of this syllabus. The students are encouraged to research additional published papers on their topic.
6. **Tests**: The midterm will cover the lecture material covered in the class, and the research paper and the Final group project and the group project presentation represent the final exam.

Course Learning Outcome:

Students are expected to learn the following:

SQL Storage management architecture (both row and column-oriented) and the performance relevant issues,

SQL Query processing architecture and issues with SQL runtime,

SQL Query optimization and the issues relevant to query plan quality,

Concurrency control; goals and approach,

Transaction management and ACID properties,

Database recovery; goal and approach,

Parallel and Distributed database architecture and issues,

Overview to current active research topics to help the students pursue deeper research in any of the topics covered; topics listed at the end of the syllabus. Students are expected to select a topic, survey published papers and write paper/presentation about the issue being addressed, how it is being addressed, and their own perspective.

Group Project for an advanced application to database (multi-tier application) to help the students going through the steps for creating then using a database , i.e., E/R diagram design, schema normalization, database population, and finally querying the database.

Overview of NoSQL Data Models and cover few examples of each

Understand the tradeoff between SQL and NoSQL databases

Hands-on experience with using SQL on Hadoop and how advanced optimizers transl

Grading Scale: refer to the Grading System table below.
Student can ask/discuss about any grade only within 7-days from the date they received the grade.

	Homework Assignment (2 HWs)	**(2*10)% = 20%%**
	P1 (individual) + P2 (Group)	**P1 (6%) + P2 (9%) = 15%**
		Additional optional 2%
		5%
		25%
		15%
		20%

Course Projects Assignment

The course include group-based project assignments and related documents must be handed in the classroom on due date (one copy per group). Familiarity with C language and Linux are expected for these projects.

Project-1 (6

Exams There will be no make-

Grading System:

13. Nov 17:	**NoSQL KV-pair: Data Model and Examples + Return of HW2**
14. Nov 22:	**NoSQL KV-pair: Data Model and Examples (Contd.)**
14. Nov 24:	**NoSQL Document: Data Model and Examples +**
	Return Group Final project presentation and return research papers (both are due)
15. Nov 29:	**NoSQL Document: Data Model and Examples (Contd.)**
15. Dec 1:	**SQL vs. NoSQL**
16. Dec 6:	**Final (Final Project): Group PPT + Presentation**
16. Dec 8:	**Final (Final Project): Group PPT + Presentation (Contd.)**

Instruction Methods: **In-Class Lecture**

Exams

There will be no make-up exams. If a student misses and exam without a legitimate excuse or advanced approval from the professor, a grade of zero will be recorded. Attending the Final Project presentation is expected.

Universi

material

Make-up Work:	No. Presentations and papers are expected the same day as designated. All team members need to be present; no exceptions are allowed.
Resources:	All students are encouraged to read more from the additional reading list, as well as additional conference or Journals.
Revision Date:	7/16/2022

Advanced database Research Topics of Interest (For Group Presentations) List:

1. **Column-Oriented Database:**

   M. Stonebraker, D. Abadi, A. Batkin, et al. "C-Store: A Colum-oriented DBMS," Proceeding of the 31st VLDB Conference, 2005, http:/

Vertica, “The Database Column: MapReduce II,” http://databasecolumn.vertica.com/2008/01/mapreduce-continued.html
F. Afrati, and J. Ullman, “A New Computation Model for Rack-Based Computing,” PODS ’09, ACM.
R. Chaiken, “SCOPE: Easy and Efficient Parallel Processing of Massive Data Sets,” VLDB 2008, http://delivery.acm.org/10.1145/1460000/1454166/p1265-chaiken.pdf?key1=1454166&key2=4530729421&coll=