Dear CS157A Class:

In this big data era, do we still need to study traditional databases? Please look at the article Communications of the ACM, January 2015, pp 18. (see http://xanadu.cs.sjsu.edu/~drtylin/classes/cs157A_fall2015/notes/

by Michael Stonebraker (The Turing Award Winner of 2014) : " ¼ a new DBMS, IMPLA, which runs on HDFS (Hadoop File System). Put simply, IMPLA is architected exactly like all of the share-nothing parallel SQL DBMS, serving the data warehouse (see Ch 21.2.2) market. Specifically, notice the MapReduce Layer has been removed, and for good reason" …. "Hadoop is picking up support in the general community." "a delay for a decade." "Goole has long since abandon it."

1) We will cover DBMS in lectures,
2) The so-called "No-SQL" will be covered in the Project.

This green sheet is prepared for "standard course"; this semester; we will increase the weight of the project.

1) Please down load the project based on the instructions giving
   http://xanadu.cs.sjsu.edu/~drtylin/classes/cs157A_fall2015/projects/
2) Please read the file Google_Big_Data_Decade.pdf in
   http://xanadu.cs.sjsu.edu/~drtylin/classes/cs157A_fall2015/notes/

Dr. Lin
http://dl.acm.org/inst_page.cfm?id=60015609

About the Project:

1) You have to sign a non-disclosure form
2) The project is the core of a semantic based search engine
   Which is a project in Data Science
   Vasant Dhar**, PhD,** Editor**-in-**Chief **of** Big Data (A professor of NYU) stated:
   "Data Science is a study of generalizable extraction of knowledge from data."
3) It is the product of CS298-99, since 2005 (Albert Sutojo he computed TFIDF using database concept), Important observation by Tam Ngo in 2006: Google's latent sematic index has nothing to do with semantic. The best version is by Jean David Hsu (an undergraduate); he tested on 3 million abstracts from Medline. Current version is by Bieu Do. A secondary storage version by
   Richard Deeley (undergraduate)

San José State University
Science/Computer Science
CS 157A Database Management Systems I

1. Course Information

Instructor: Dr. Tsau Young Lin
Department: Computer Science
College of Science, San Jose State University.
Spring Semester, 2015

Course Title:	**Database Management System I**
Course Code:	CS 157A
Class meeting :	Course: CS 157A; Section: 3; Room: MH 222; Begin Time: 13:30; End Time: 14:45; Day: MW; Unit: 3
Office Hours:	MW14:45-15:30; 11:45-12:15 (T Th appointments only)

Office Location:	MH 214
Office Phone:	(408)924-5121
E-mail:	drtylin@cs.sjsu.edu
Department Fax:	(408)924-5062

2. Specific course information:

a. Catalog Description:
Current, classical database systems. Entity-relationship and enhanced entity models. Relational model, algebra, calculus. Current, emerging SQL standard. Embedded, Dynamic SQL. Application perspective on transactions and security. Interactive and programmatic interfaces to database systems. Application programming project using commercial database system.
.

b. Prerequisite: CS 146 (with a grade of "C-" or better) or instructor consent.

University Rules

NOTE that University policy F69-24 at http://www.sjsu.edu/senate/docs/F69-24.pdf states that "Students should attend all meetings of their classes, not only because they are responsible for material discussed therein, but because active participation is frequently essential to insure maximum benefit for all members of the class. Attendance per se shall not be used as a criterion for grading."
Note that "All students have the right, within a reasonable time, to know their academic scores, to review their grade-dependent work, and to be provided with explanations for the determination of their course grades." See University Policy F13-1 at http://www.sjsu.edu/senate/docs/F13-1.pdf for more details.

3. Textbook

Hector Garcia-Molina, Jeffrey D. Ullman, Jennifer D. Widom.
Database Systems: The Complete Book. Prentice Hall. 2nd Ed ISBN-13: 978-0131873254. 2008.

Supplemental reference:

MS SQL (Murach's SQL Server 2012 for Developer by B. Syverson and J. Murach)

5.Specific goals for my sections:

Important topics, database design theory and diagnosis are added.

a. Student learning outcomes: Upon successful completion of this course, students should have a knowledge on
   (1). the basic of system and user aspects of database systems.
   (2) the basic of database modeling including relational model, relation algebra and constraints, such as key constraints
   (3) the basic of database diagnosis and design theory (discovering the data anomaly and normalizations)
   (4) the basic skill in writing SQL, including embedded and dynamic SQL.

b. Programming outcomes:
   a. to use commercially available DBMS, such as Oracle, DB2, MS SQL Server and etc (we will use MS SQL Server)
   b. to handle the interactions between DBMS and a third generation programming language, such as Java, C++.
   c. Ability to organize a team project (using waterfall model) to build a "real world" database system(complexity is near a real world database system) that is based on a DBMS and a programming language

c. Detail Requirements

a. Projects: Understand “my” Search Engine
b. Exams: 4 exams
c. Quizzes: Many unscheduled quizzes. Missing quizzes (up to 4 quizzes) can be made up by attending public technical talks sponsored by professional organizations, such as IEEE, ACM, AMS, and etc.
d. Homework: Short SQL 's to enhance the understanding of lectures
e. Class Participation: Present some interesting topics in class, give some demos of short programs, or explaining hard home works in class will be properly awarded.

6. Brief list of topics to be covered:

1) Elementary database modeling including relational model, algebra and constraints

2) Data anomaly and database design theory (data anomaly, normalization)

3) Current, emerging SQL standard including embedded and dynamic SQL.

4) Interaction between DBMS and a third generation programming language, mainly on Java or C.

5) The skill to organize a team project (using waterfall model) to build a “real world” database system that is based on a commercially available DBMS, such as Oracle, DB2, SQL Server and etc. and a third generation programming language, such as Java and C.

7. Tentative course calendar including assignment due dates, exam dates, date of Final exam

Course Plan

Weeks	Lectures & Assignments	
0(8/17)	First day of class Thursday (8/20)	
1(8/24)	Overview of class policies, database system and “abc” of MS SQL Server.	Ch 1
2(8/31)	Relational data model An Overview of Data Models; Basics of Relational Model ; Defining a Relation Schema in SQL Defining a Relation Schema in SQL Demo: installation of oracle.	Ch 2.1-2.3
3(9/7)	Relational data model Defining a Relation Schema in SQL; algebraic query language oracle SQL (create new relation; insert data)	Ch 2.3-2.4
4(9/14)	Relational data model; algebraic query language	Ch 2.5;Ch4.1
5(9/21)	Relational data model; Constrains on relations	Exam I Ch 6.1
6(9/28)	The Database language SQL;	Ch 6.2-6.3
7(10/5)	The Database language SQL	Ch 6.4-6.5

8(10/12).	The Database language SQL	Ch 6.6 Exam 2
9(10/19)	The Database language SQL (Optimization)	Ch16.1-2
10(10/26)	The Database language SQL (Optimization)	Ch 16.2-16.3
11(10/31)	Design Theory for Relational databases;	Lectures
12(11/2)	Design Theory for Relational databases;	Based on
13 (11/9)	Design Theory for Relational databases;	Date Book
14(11/16)	Design Theory for Relational databases; SQL (Optimization)	Ch 16.2-16.3
15(11/23)	Review; Exam 3	Thanksgiving
16(11/30)	Secondary Storage Management; The Memory Hierarchy; Disks	Ch13.1' Ch13.2
17(12/7)	Exam 4	
18(12/16)	Final Exam	

5. Grades

Projects & home works	30%
Exams	30%
Quizzes(Class average set 80 linearly)	10%
Final Exam	30%
Total	100%

90-92; 93-96;97-	A
80-82; 83-86;87-	B
70-72; 73-76;77-	C
60-62; 63-66;67-	D
<60	F

6. Dropping and Adding

Students are responsible for understanding the policies and procedures about add/drops, academic renewal, etc

7. University Policies

General Expectations, Rights and Responsibilities of the Student

As members of the academic community, students accept both the rights and responsibilities incumbent upon all

members of the institution. Students are encouraged to familiarize themselves with SJSU's policies and practices pertaining to the procedures to follow if and when questions or concerns about a class arises. See University Policy S90–5 at http://www.sjsu.edu/senate/docs/S90-5.pdf. More detailed information on a variety of related topics is available in the SJSU catalog, at http://info.sjsu.edu/web-dbgen/narr/catalog/rec-12234.12506.html. In general, it is recommended that students begin by seeking clarification or discussing concerns with their instructor. If such conversation is not possible, or if it does not serve to address the issue, it is recommended that the student contact the Department Chair as a next step.

Dropping and Adding

Students are responsible for understanding the policies and procedures about add/drop, grade forgiveness, etc. Refer to the current semester's Catalog Policies section at http://info.sjsu.edu/static/catalog/policies.html. Add/drop deadlines can be found on the current academic year calendars document on the Academic Calendars webpage at http://www.sjsu.edu/provost/services/academic_calendars/. The Late Drop Policy is available at http://www.sjsu.edu/aars/policies/latedrops/policy/**.** Students should be aware of the current deadlines and penalties for dropping classes.

Information about the latest changes and news is available at the Advising Hub at http://www.sjsu.edu/advising/.

Consent for Recording of Class and Public Sharing of Instructor Material

University Policy S12-7, http://www.sjsu.edu/senate/docs/S12-7.pdf, requires students to obtain instructor's permission to record the course and the following items to be included in the syllabus:

- "Common courtesy and professional behavior dictate that you notify someone when you are recording him/her. You must obtain the instructor's permission to make audio or video recordings in this class. Such permission allows the recordings to be used for your private, study purposes only. The recordings are the intellectual property of the instructor; you have not been given any rights to reproduce or distribute the material."
    - It is suggested that the greensheet include the instructor's process for granting permission, whether in writing or orally and whether for the whole semester or on a class by class basis.
    - In classes where active participation of students or guests may be on the recording, permission of those students or guests should be obtained as well.
- "Course material developed by the instructor is the intellectual property of the instructor and cannot be shared publicly without his/her approval. You may not publicly share or upload instructor generated material for this course such as exam questions, lecture notes, or homework solutions without instructor consent."

Academic integrity

Your commitment, as a student, to learning is evidenced by your enrollment at San Jose State University. The University Academic Integrity Policy S07-2 at http://www.sjsu.edu/senate/docs/S07-2.pdf requires you to be honest in all your academic course work. Faculty members are required to report all infractions to the office of Student Conduct and Ethical Development. The Student Conduct and Ethical Development website is available at http://www.sjsu.edu/studentconduct/.

Campus Policy in Compliance with the American Disabilities Act

If you need course adaptations or accommodations because of a disability, or if you need to make special

arrangements in case the building must be evacuated, please make an appointment with me as soon as possible, or see me during office hours. Presidential Directive 97-03 at http://www.sjsu.edu/president/docs/directives /PD_1997-03.pdf requires that students with disabilities requesting accommodations must register with the Accessible Education Center (AEC) at http://www.sjsu.edu/aec to establish a record of their disability.