San José State University College of Science / Department of Computer Science CS267 Topics in Database Systems, Spring 2020

Course and Contact Information

Instructor:	Dr. Mike Wu
Office Location:	MacQuarrie Hall 211
Email:	Ching-seh.Wu@sjsu.edu
Office Hours:	Tuesday 2:30~3:30pm and Thursday 3:00~4:00pm (Please drop me an email with time info and subject.)
Class Days/Time:	TuTh 4:30-5:45pm
Class Room:	MacQuarrie Hall 233
Prerequisites:	CS 157B Database Management Systems II (with a grade of "C-" or better)

Faculty Web Page and OneSJSU Messaging

Course materials such as syllabus, handouts, notes, assignment instructions, etc. can be found at Canvas of SJSU One. You are responsible for regularly checking with the email system and Canvas through One.SJSU at http://one.sjsu.edu to learn of any updates.

Course Description

General: Advanced topics in the area of database and information systems. Content differs in each offering. Possible topics include though not restricted to: Data Mining, Distributed Databases and Transaction Processing. (This description is from course catalog of CS Department Website)

This semester, topics include the following (time permits):

•

- Gain knowledge and key concepts, algorithms, techniques related to Big Data.
- Familiar with Mining data streams.
- Familiar with Apache Hadoop architecture, and Map-Reduce.
- Gain hands-on experience to develop and implement Big Data analytical project.
- Use scalable algorithms to extract knowledge from Big data
- Become familiar with the different data models used by NoSQL Big Data platforms.
- Become familiar with tradeoffs between SQL and NoSQL: Data model, Query language, guarantees provided.
- •

Course Requirements and Assignments

Assignments

You are expected to learn all of the material presented in the lectures. Assignments include written and programming. Assignments must be turned in on time; late submission will not be accepted with the exception of medical emergencies or similar exceptional circumstances that must be discussed in advance with the instructor. All assignments are due at the beginning of the class period on the announced due date.

Mid-Term and Final Exams

Exams will consist of questions and problems aimed at assessing student mastery of course topics. Conceptual questions may be in the form of essay or multiple-choice format and questions that require pseudo code and/or computations. All exams for this course are closed book.

If you are unable to attend any one of the exams, arrangements may be made only if you have a legitimate reason. You need to infocm j0 (a) 4x9 () -1 (,) -7 18 587.76cm BT 50 0 0 50 1082 53 Tm /TT4 1 Tf [(c) 4 (hoi) -7.76 c

upload instructor generated material for this course such as exam questions, lecture notes, or homework solutions without instructor consent.

University Policies

Per University Policy S16-9, university-wide policy information relevant to all courses, such as academic integrity, accommodations, etc. will be available on Office of Graduate and Undergraduate Programs' Syllabus Information web page at http://www.sjsu.edu/gup/syllabusinfo/" Make sure to review these policies and resources.

Topics in Database Systems, CS267, Spring, 2020, Course Schedule

Tentative Course Schedule (This schedule is subject to change with fair notice.)

Week	Date	Topics, Readings, Assignments, Deadlines
1		

Week	Date	Topics, Readings, Assignments, Deadlines
11	4/7	SPARK Architecture, and YARN vs. Mesos
11	4/9	SPARK Architecture, and YARN vs. Mesos
12	4/14	Big Data Document-based Data Model
12	4/16	Big Data K/V-based Data Model: Hive, Pig, HBase
13	4/21	Scalability Models (Strong vs. Eventual Consistent Models) and Big
		Data Issues
13	4/23	Big Data analytics using machine learning
14	4/28	Big Data analytics using machine learning
14	4/30	Tradeoffs between SQL and NoSQL
15	5/5	Project Presentation and Demo