San José State University
Department of Computer Engineering

CMPE 226
Database Systems

Section 1
Spring 2017

GREEN SHEET

CMPE 226-02   Tu   6:00 – 8:45 PM   DMH 227

Instructor: Ron Mak
Office hours: Th: 2:30 – 4:30 PM
Office location: ENG 250
E-mail: ron.mak@sjsu.edu
Instructor web page: http://www.cs.sjsu.edu/~mak/
Class web page: http://www.cs.sjsu.edu/~mak/CMPE226/index.html

Course catalog description

"Database architectures, technologies, and practices for enterprise systems that use structured, semi-structured, and unstructured data. Provides opportunities to research and acquire experience using modern and emerging concepts in relational and non-relational database theory and technologies. Prerequisite: CMPE 272 or instructor consent. Computer Engineering or Software Engineering majors only." 3 units

Please submit into Canvas a copy of your transcript with the prerequisite courses highlighted. "*Students who do not provide documentation of having satisfied the class prerequisite and co-requisite requirements (if any) by the second class meeting will be dropped from the class.*"

Course goals

Gain confidence and proficiency in developing enterprise-class transactional and analytical data management applications that use back-end databases and data warehouses.

Course learning outcomes

CLO 1: Learn the fundamentals of data management tools and procedures
in order to develop significant enterprise-class data-driven applications.

CLO 2: Gain important data management and project development skills that are
valued by employers.

CLO 3: Understand practical aspects of data modeling, relational databases, SQL,
and ORM (object-relational mapping).

CLO 4: Learn to design and deploy operational databases for OLTP (online transaction
processing) and analytical databases for OLAP (online analytical processing).

CLO 5: Learn to implement dimensional modeling, star schemas, and the ETL
(extract-transform-load) process for data warehousing applications.

CLO 6: Understand semi-structured data and learn how to use XML and its tools,
NoSQL databases, and data virtualization.

CLO 7: With cooperation from local companies, practice using their commercial
data management tools for your assignments and projects.

Required text

Title:	**Database Systems:** **Introduction to Databases and Data Warehouses**
Authors:	Nenad Jukic, Susan Vrbsky, and Svetlozar Nestorov
Publisher:	Prospect Press, 2017
Paperback:	978-1-943153-19-0 available from Redshelf.com
eTextbook:	978-1-943153-18-3 available from Redshelf.com and VitalSource.com *These are much less expensive versions of the* *textbook originally published in 2014 by Pearson.* *Pearson ISBN: 978-0-13-257567-6*

Software to install

Download and install XAMPP (https://www.apachefriends.org/index.html), which will install and configure on your Mac, Windows, or Linux platform the following software packages:

- PHP
- Apache web server
- MariaDB database server (compatible with MySQL)

You can download and install these packages separately, but then you'll have the hassle of configuring them to work nicely together.

Later in the semester, you will also download and install the Cisco Information Server (CIS) Studio application.

Course websites

Course materials such as syllabus, handouts, notes, assignment instructions, etc. can be found on the class web page at http://www.cs.sjsu.edu/~mak/CMPE226/ and on the Canvas Learning Management System course website at http://sjsu.instructure.com. Piazza at https://piazza.com provides broadcast messages and a class forum. You are responsible for regularly checking these sites to learn of any updates.

Course requirements and assignments

You will form project teams of four students each. The teams will last throughout the semester. Once the teams are formed, you will not be allowed to move from one team to another, so form your teams wisely! **Team membership is mandatory for this class.**

Each team will choose its own data management application to develop, but it must fit the architectural framework described above. At the end of the semester, *all the members of a project team will each receive the same project score*. The project grade will be determined by the overall quality of the final version of the project team's artifacts and by how well the team achieved its goals to create a successful web application.

There will be a number of one- or two-week team assignments during the semester to be done by the teams. These assignments will help your team make progress on your data management project. *All members of a project team will receive the same score for an assignment.*

Each assignment will be worth up to 100 points. Late assignments will lose 20 points and an additional 20 points for each 24 hours after the due date.

Data management project

Each project team will choose data sources to download from the Internet or cloud in order to design and build a significant data-driven enterprise application. See http://cs61.cs.sjsu.edu/CS235Projects to get an idea of the types of applications you can build. CS 235 was about user interface design and data visualization (not virtualization) using various data sources. In CMPE 226, the emphasis will be on data management. A server with some commercial data management tools installed will be provided.

This project will involve:

- Data sources chosen by each team from the Internet or the cloud. A list of suggested URLs where data can be found will be provided in class.
- Substantial use of data management tools that demonstrates a strong understanding of how to effectively apply the tools in meaningful ways on the chosen data sources.
- A written report (10-12 pp.) that describes the requirements and goals of the project and how the application meets them. The report should include screen shots of key application displays.

Each student team will present its application to the class at the end of the semester. *You should start thinking about and planning for your project early in the semester.*

Postmortem report

At the end of the semester, each student must also turn in a short (1 or 2 pages) **individual postmortem report** that includes:

- A brief description of what you learned in the course.
- An assessment of your accomplishments for your project team on the assignments and the compiler project.
- An assessment of each of your other project team members.

Only the instructor will see these reports.

Exams

The midterm and final examinations will be closed book. Instant messaging, e-mails, texting, tweeting, file sharing, or any other forms of communication with anyone else during the exams will be strictly forbidden.

There can be no make-up midterm examination unless there is a documented medical emergency. Make-up final examinations are available only under conditions dictated by University regulations.

Class grade

Your individual final class grade will be weighted as follows:

> 30%   Assignments*
> 35%   Project*
> 15%   Midterm exam**
> 20%   Final exam**
>
> * *project team scores*
> ** *individual scores*

During the semester, you can keep track of your progress in Canvas. Each assignment and exam will be scored (given points) but not assigned a letter grade. The average score can be seen in Canvas after each assignment and the midterm exam.

At the end of the semester, all the students will be ranked in the order of their weighted class scores. Students with the median score will be assigned the B+ grade. Higher and lower grades will then be assigned based on how the scores cluster above and below the median.

Your final class grade can be adjusted up or down depending on your level and quality of participation on your project team as determined by the project tracking tools and your team members' assessments of your performance.

Classroom protocol

It is very important for each student to attend classes and to participate. Cell phones in silent mode, please.

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/"

Schedule

Subject to change with fair notice. Chapter readings are from the required text.

Week	Dates	Topics and activities	Chapters
1	Jan 31	Overview of the course XAMPP *Form programming teams*	1
2	Feb 7	Basic HTML Lists, tables, links, and images Simple form processing with PHP Simple database access with PHP Simple dynamic page generation with PHP	
3	Feb 14	PHP syntax PHP data objects (PDO) Data modeling Entity-relationship (ER) diagrams	2
4	Feb 21	Relational database modeling Mapping ER diagrams to relational schemas Update anomalies Functional dependencies and normalization	3, 4
5	Feb 28	SQL Views SQL injection attack PHP prepared statements MySQL Workbench	5
6	Mar 7	More SQL table creation Constraint management Referential integrity Types of joins Cascading updates and deletes Indexing Object-oriented PHP Object-relational mapping (ORM) PHP error reporting	5, 6
7	Mar 14	Query optimization Stored procedures Triggers Transaction management ACID properties Backup and recovery RAID	10
8	Mar 21	***Midterm exam Tuesday, March 21***	
		SPRING BREAK	

9	Apr 4	Operational vs. analytical databases Data warehousing Extract-transform-load (ETL) Data marts Dimensional modeling and star schemas Fact tables and dimension tables	7
10	Apr 11	Types of fact tables Fact table granularity Slowly changing dimensions Data warehouse architectures Online transaction processing (OLTP) Online analytical processing (OLAP) Business intelligence (BI) tools	8, 9
11	Apr 18	Data virtualization Cisco Information Server (CIS) Semi-structured data XML XML Schema	Appendix H
12	Apr 25	XPath and XQuery FLWOR expressions NoSQL databases	Appendix I
13	May 2	Object databases Distributed databases Cloud computing Data mining Big data MapReduce	Appendix E Appendix F Appendix G Appendix J
14	May 9	Project presentations	
15	May 16	Project presentations	
	May 17	***Final projects due Wednesday, May 17***	
Final	May 23	***Final exam Tuesday, May 23*** 5:15 - 7:30 PM, DMH 227	