

Database Systems for Analytics

ITP 250 - (2 Units) – 2019 Spring



Course Description

Concepts in modeling data for analytic applications. Designing and implementing robust databases. Querying databases to extract business intelligence. New trends in databases.

Objectives

To prepare students to model and build databases. Upon completion of the course, student will be able to:

- Model data using Entity-Relationship model
- Perform Normalization to 3rd Normal Form
- Design real world databases
- Use Structured Query Language (SQL) to build and query databases
- Describe the importance of Business Intelligence
- Enumerate the differences between SQL and NoSQL databases
- Build and query NoSQL databases

Prerequisites

None

Lectures and Lab

Tuesdays 3:30PM to 6:20PM in VKC 100

Instructor

Calvin Nguyen <cnt.nguyen@usc.edu>

Office Hours: after class Tuesdays 6:30PM to 7:30PM in VKC 100

Contact Instructor for all Lectures related questions.

Teaching Assistant

Arthur Girard <adgirard@usc.edu>

Maxwell Lee <maxwelll@usc.edu>

Contact Teaching Assistants for all Assignments related questions.

Website

blackboard.usc.edu

Course readings and assignments will be posted on Blackboard. Students should check for materials, deadlines, announcements regularly on Blackboard.

Textbook

None

Software

Students can use their own computers or login to the Viterbi Virtual Lab. The software will be provisioned through the Viterbi Virtual Lab. Specifically, you will be using

- Microsoft Access
- MySQL and MySQL Workbench

- MongoDB

Final Project

There will be one final project that will be team based. Teams of 2-4 students will be formed. This is a comprehensive project deriving from all topics discussed during the semester.

Grading

The weight of the graded material during the semester is listed below:

Homework	25%
Project	25%
Exam – Midterm	25%
Exam – Final	25%

The following grading scale will be used to determine your letter grade:

A	100-95
A-	95-92
B+	92-89
B	89-86
B-	86-83
C+	83-80
C	80-77
C-	77-74
D+	74-71
D	71-68
D-	68-65
F	65 or below

Policies

Exams

- The use of mobile devices, books, notes or computers is not be permitted during the exam.
- No make-up exams will be offered except for documented medical or family emergencies.

Homework

- Students are encouraged to work with their classmates. However, students must turn in their own original work.
- Late homework submissions will be subject to a late penalty. The penalty is 25% per day. No assignments will be accepted later than four days from the due date.

Incomplete and Missing Grades

Excerpts for this section have been taken from the University Grading Handbook, located at <http://www.usc.edu/dept/ARR/grades/gradinghandbook/index.html>. Please see the link for more details on this and any other grading concerns.

A grade of Missing Grade (MG) “should only be assigned in unique or unusual situations... for those cases in which a student does not complete work for the course before the semester ends. All missing grades must be resolved by the instructor through the Correction of Grade Process. One calendar year is allowed to resolve a MG. If an MG is not resolved [within] one year the grade is changed to [Unofficial Withdrawal] UW and will be calculated into the grade point average a zero grade points.

A grade of Incomplete (IN) “is assigned when work is not completed because of documented illness or other ‘emergency’ **occurring after the twelfth week** of the semester (or 12th week equivalency for any course scheduled for less than 15 weeks).”

Students with Disabilities

Any student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. Please be sure the letter is delivered to your course instructor (or TA) as early in the semester as possible. If you need accommodations for an exam, the form needs to be given to the instructor at least two weeks before the exam.

DSP is located in STU 301 and is open from 8:30am to 5:00pm, Monday through Friday. Contact info: 213-740-0776 (Phone), 213-740-6948 (TDD only), 213-740-8216 (FAX), ability@usc.edu, http://sait.usc.edu/academicsupport/centerprograms/dsp/home_index.html.

Statement on Academic Conduct and Support Systems

Academic Conduct

Plagiarism – presenting someone else’s ideas as your own, either verbatim or recast in your own words – is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in *SCampus* in Section 11, *Behavior Violating University Standards* <https://scampus.usc.edu/1100-behavior-violating-university-standards-and-appropriate-sanctions/>. Other forms of academic dishonesty are equally unacceptable. See additional information in *SCampus* and university policies on scientific misconduct, <http://policy.usc.edu/scientific-misconduct/>.

Discrimination, sexual assault, and harassment are not tolerated by the university. You are encouraged to report any incidents to the *Office of Equity and Diversity* <http://equity.usc.edu/> or to the *Department of Public Safety* <http://capsnet.usc.edu/departments/public-safety/online-forms/contact-us>. This is important for the safety of the whole USC community. Another member of the university community – such as a friend, classmate, advisor, or faculty member – can help initiate the report, or can initiate the report on behalf of another person. *The Center for Women and Men* <http://www.usc.edu/student-affairs/cwm/> provides 24/7 confidential support, and the sexual assault resource center webpage [sarc@usc.edu](http://sarc.usc.edu) describes reporting options and other resources.

Support Systems

A number of USC’s schools provide support for students who need help with scholarly writing. Check with your advisor or program staff to find out more. Students whose primary language is not English should check with the *American Language Institute* <http://dornsife.usc.edu/ali>, which sponsors courses and workshops specifically for international graduate students. *The Office of Disability Services and Programs* http://sait.usc.edu/academicsupport/centerprograms/dsp/home_index.html provides certification for students with disabilities and helps arrange the relevant accommodations. If an officially declared emergency makes travel to campus infeasible, *USC Emergency Information* <http://emergency.usc.edu/> will provide safety and other updates, including ways in which instruction will be continued by means of blackboard, teleconferencing, and other technology.

Course Outline

Lecture 1 – Jan 8

Introduction

- Syllabus Review
- Brief history of databases and their role in information systems
- Different types of databases and their organizational context
- Survey of DBMS

Lecture 2 – Jan 15

Data Modeling

- Data models
- Business rules
- Relational and entity-relationship modeling

Homework 1 Assigned – Due Fri, Jan 25

Lecture 3 – Jan 22

Data Modeling (continued)

- Entities, attributes, relationships
- Keys: primary, foreign, candidate, surrogate, super
- Minimum and maximum cardinality

Homework 2 Assigned – Due Fri, Feb 1

Lecture 4 – Jan 29

Entity Relationship Modeling

- E-R model for modeling business situations
- Notation Methods
- Tools

Homework 3 Assigned – Due Fri, Feb 8

Lecture 5 – Feb 5

Normalization

- Anomalies and the need for normalization
- Normal forms
- First, second, third normal forms
- Denormalization
- Dependency Diagrams

Homework 4 Assigned – Due Fri, Feb 15

Lecture 6 – Feb 12

Structured Query Language

- Creating a database using MySQL
- DDL
- DML
- SQL Language Introduction

Homework 5 Assigned – Due Fri, Feb 22

Lecture 7 – Feb 19

Midterm Review

Lecture 8 – Feb 26

Midterm – lecture room

Lecture 9 – Mar 5

Structured Query Language (continued)

- SELECT queries
- Querying multiple tables
- SQL functions
- Aggregation/grouping

Homework 6 Assigned – Due Fri, Mar 15

Spring Break - Mar 10-17**Lecture 10 – Mar 19**

Structured Query Language (continued)

- JOINS and Set Operations
- SQL Sub-queries

Homework 7 Assigned – Due Fri, Mar 29

Lecture 11 – Mar 26

Business Intelligence Systems

- Business intelligence
- Data warehouses and data marts
- Business reporting and intelligence
- Data mining

Homework 8 Assigned – Due Fri, Apr 5

Lecture 12 – Apr 2

Big Data Analytics

- Big data
- Hadoop

Project Part 1 Assigned – Due Tue, Apr 16

Lecture 13 – Apr 9

Big Data Analytics (continued)

- NoSQL
- MongoDB

Project Part 2 Assigned – Due Tue, Apr 23

Lecture 14 – Apr 16

Big Data Analytics (continued)

- MongoDB Aggregation Framework

Lecture 15 – Apr 23

Course Evaluation

Final Exam Review

Final – May 7 (2:00 – 4:00PM)

Final Exam – Lecture room