## Database Systems for Analytics

## ITP 250 - 32066 (2 Units) – Spring 2018

**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, and between row and column stored databases.
* Build and query NoSQL databases

**Prerequisites**

None

**Lectures and Lab**

Wednesdays 5:00PM to 7:50PM in TTH 210

**Instructor**

Naz Nageer

Email: Nageer@usc.edu

Office Hours: Wednesdays 3:00PM to 4:50PM in OHE 330C / Bluejeans Online on request

**Teaching Assistant**

Cole Jones - colecjon@usc.edu

**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%

Final Project 15%

Midterm Exam 30%

Final Exam 30%

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 no 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/department/department-public-safety/online-forms/contact-us>.  This is important for the safety 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 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/*](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**

**Week 1 – Jan 10**

*Introduction*

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

**Week 2 – Jan 17**

*Data Modeling*

* Data models
* Business rules
* Relational and entity-relationship modeling
* *Homework 1 – Jan 26*

**Week 3 – Jan 24**

*Data Modeling (continued)*

* Entities, attributes, relationships
* Keys: primary, foreign, candidate, surrogate, super
* Minimum and maximum cardinality
* E-R model for modeling business situations
* Notation Methods / Tools
* *Homework 2 – Feb 2*

**Week 4 – Jan 31**

*Normalization*

* Anomalies and the need for normalization
* Normal forms
* First, second, third normal forms
* Denormalization
* Constraints
* Indexes
* *Homework 3 – Feb 9*

**Week 5 – Feb 7**

*Structured Query Language*

* Creating a database using MySQL
* DDL
* DML
* *Homework 4 – Feb 16*

**Week 6 – Feb 14**

*Structured Query Language*

* SELECT queries
* *Homework 5 – Feb 23*

**Week 7 – Feb 21**

*Structured Query Language (continued)*

* Querying multiple tables
* Difference between SUB and JOIN for querying multiple tables
* *Midterm Review*
* *Homework 6 – Due Mar 9*

**Week 8 – Feb 28**

* Mid Term

**Week 9 – Mar 7**

*Structured Query Language (continued)*

* SQL functions
* Aggregation/grouping
* Advanced JOINs
* *Homework 6 – Due Mar 22*

**Week 10 – Mar 14**

* Spring Break

**Week 11 – Mar 28**

*Business Intelligence Systems*

* Business intelligence
* Data warehouses and data marts
* Business reporting and intelligence
* Data mining
* *Homework 9 – Due Apr 6*

**Week 13 – Apr 4**

*Big Data Analytics*

* Big data
* Hadoop
* *Final Project – Due Apr 27*

**Week 14 – Apr 11**

*Big Data Analytics (continued)*

* In-memory databases
* NoSQL
* MongoDB

**Week 15 – Apr 18**

*Big Data Analytics (continued)*

* MondoDB Aggregation Framework

**Week 16 – Apr 25**

* Final Project and Final Review

**Week 17 - May 2**

* Final Exam – May 2 (4:30 – 6:30PM)