ITP 250: Database Systems for Analytics
Units: 2
Fall 2016

Location
Course materials: blackboard.usc.edu

Instructor
Name:
Office:
Office Hours: 4 hours per week. Schedule varies by semester
Contact Info:

Teaching Assistant
Name: TBD
Office Hours: TBD
Contact Info: TBD

IT Help
IT help will be provided by Viterbi IT
Hours of Service: 8am – 5pm M-F
Contact Info: http://viterbi.usc.edu/servicedesk , 213-740-0517, engrhelp@usc.edu

Course Description
Concepts in modeling data for analytic applications. Designing and implementing analytics databases. Querying databases to extract business intelligence. Importance of database performance for analytics.

Learning Objectives
To prepare students to model and build databases. Upon completion of the course, student will be able to:
- Model data and prepare it for analytics
- Design real world databases for analytics
- Use Structured Query Language (SQL) to build and query databases
- Describe the importance of Business Intelligence: Reporting and data mining
- Explain how databases are optimized for analytics performance
- Enumerate the differences between SQL and NoSQL databases, and between row and column stored databases.
- Compare real world database capabilities for analytics

Prerequisites
None

Co-requisite
None

Concurrent Enrollment
None

Recommended Preparation
None

Course Notes
All course materials will be made available through Blackboard. Blackboard discussion forums will be used for out-of-class discussions. Lectures are delivered face to face in classroom.

Technology Proficiency and Hardware/Software Required
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
- MYSQL Workbench
- Teradata
- MongoDB

Required Reading and Supplementary Materials
Textbook:

Description and Assessment of Assignments
Homework – Most homework is computer based. Homework should be turned in to Blackboard on time. Grading will be based on completeness, accuracy, and correctness. Feedback will be provided through Blackboard.
Exams – are written, in-class tests. They are based on concepts and not on particular tools**.**

Grading Breakdown
The weight of graded material during the semester is listed below.
No extra credit assignments will be offered.

Homework	30%
Midterm	35%
Final Exam	35%
Total	100%

Grading Scale

The following shows the grading scale to be used to determine the 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*

Assignment Submission Policies

It is the responsibility of the student to make sure case studies and assignment are turned in on time. Make sure you follow the procedures outlined in each assignment or case study (Blackboard submissions).

Late assignment submissions will be subject to a late penalty of 25% per day. No assignments will be accepted later than four days from the due date.

Additional Policies

No make-up exams (except for documented medical or family emergencies) will be offered nor will there be any changes made to the Final Exam schedule, except as permitted by university rules.

Lecture attendance is not mandatory however it is recommended that students not miss any lecture.

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)."

Database Systems for Analytics
ITP250 (2 units)

Course Outline

Week 1   Introduction
- Brief history of databases and their role in analytics
- Different types of databases and their organizational context
- Survey of DBMS

       Reading Assignment: *Chapter 1*
       Homework*: None*

Week 2   Analytics Database design
- Data Models
- Business Rules
- Relational and Entity-Relationship Modeling

       Reading Assignment: *Chapter 3*
       Homework: *Create a data model for sales order history*
       Due Date: Week 3

Week 3   Data Modeling
- Entities, attributes, relationships
- Keys
- Minimum and maximum cardinality
- E-R model for modeling analytics databases
- CRUD
- Relational model

       Reading Assignment: *Chapter 5*
       Homework: *Create a data model for medical records*
       Due Date: Week 4

Week 4   Database design
- Degrees of relationships
- Transforming models into databases

       Reading Assignment: *Chapter 6*
       Homework: *Convert the sales order model into a database*
       Due Date: Week 5

Week 5   Normalization
- Anomalies and the need for normalization
- Normal forms

- First, second, third, Boyce-Codd, Fourth normal form
- Denormalization
- Constraints
- Indexes

Reading Assignment: *Chapter 4*
Homework: *Normalize the sales order database*
Due Date: Week 6

Week 6   Structured Query Language
- Creating a database using MySQL
- DDL
- DML

Reading Assignment: *Chapter 2*
Homework: *Normalize several databases to 3NF*
Due Date: Week 7

Week 7   SQL (cont.)
- SELECT queries
- Querying multiple tables
- Difference between Sub query and join for querying multiple tables

Reading Assignment: *Chapter 2*
Homework*: Convert medical model to MYSQL database*
Due Date: Week 8

Week 8   SQL (cont.)
- SQL Functions
- Aggregation
- Grouping

Reading Assignment: *Chapter 7*
Homework: *Populate the medical database*
Due Date: Week 9

Week 9   **Midterm Exam**

Week 10   SQL (cont.)
- Advanced JOINs

Reading Assignment: *Chapter 7*
Homework: *Run queries against the medical database*
Due Date: Week 11

Week 11   World of analytics
- Role of databases in Big data

- Data warehouses and data marts

Reading Assignment: *Chapter 12*
Homework: *Research and report on the recent role of big data in business*
Due Date: Week 12

Week 12   World of analytics (contd.)
- Business reporting and intelligence
- Data mining

Reading Assignment: *Chapter 12*
Homework*: Research and report on three data mining discoveries that have significantly benefited industry*
Due Date: Week 13

Week 13   Query Performance and Optimization
- Strategies for optimizing queries
- Typical bottlenecks in query performance

Reading Assignment: *Online resources*
Homework*: Performance metrics for queries against large database*
Due Date: Week 14

Week 14   Other Database Systems
- Columnar and in-memory databases
- NoSQL

Reading Assignment: *Online resources*
Homework*: Research and report on the applications of columnar and NoSQL databases*
Due Date: Week 15

Week 15   Real world database performance
- Compare database performance for analytics
- Real time analytics

Reading Assignment: *Online resources*
Homework*: Report on three major database performance for real time analytics*
Due Date: Week 16

Week 16   **Final Exam**

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.htmlprovides 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.