

Professor: Jason Geng

Office: BRI-307G

Office Phone: TBD

Office Hours: 5-6PM, Tue and Thu

E-mail: shijiang@marshall.usc.edu

Course Description

Structured query language (SQL) is the language of databases. Whether students run reports or collect data for analysis, you need to know SQL to add, delete, edit and view records.

This course provides a step-by-step overview and instructions that help students to get started with the SQL language. You will learn how to create SQL statements for data storage, data collection, data computation and reporting. This course will introduce SQL with big data platform Hadoop and Spark, students will learn how to process and analyze data with large scale of dataset.

Learning Objectives

Data collection and processing are important skills for data analysts and business analysts. Knowledge of relational database and SQL are one of the most valued technical skills as analyst can have. On completion of this course, students will be able to manage, query and analyze large scale of dataset by using relational database.

Required Materials

John Viescas, Michael Hernandez, SQL Queries for Mere Mortals: A Hands-On Guide to Data Manipulation in SQL (3rd Edition), Addison-Wesley Professional; 3 edition (June 21, 2014)

https://www.amazon.com/SQL-Queries-Mere-Mortals-Hands/dp/0321992474/ref=zg_bs_3804_4

Prerequisites and/or Recommended Preparation:

There are no pre requisites. It is recommended though that students have some basic computer programming knowledge.

Course Notes:

Blackboard for grade posting. Piazza system for communicating and course material distribution.

Grading Policies:

Grading will be fair for every students in this class. It will be mathematically calculated according to grading scales. There will be opportunities to earn extra credit. However, only impressive work in projects qualifies extra credit.

The grading scale is as follows:

A 94 and above
A- 90 – 93

B+	87 – 89
B	84 – 86
B-	80 – 83
C+	77 – 79
C	74 – 76
C-	70 – 73
D+	67 – 69
D	64 – 66
D-	60 – 63
F	59 or below

<u>Assignments</u>	<u>Points</u>	<u>% of Grade</u>
Homeworks (6 in total)	600	30%
Exam	300	30%
Capstone Project (Individual)	300	30%
Class participation & attendance	<u>100</u>	<u>10%</u>
TOTAL	1000	100%
Impressive work in capstone project will earn Maximum 3% of grade (30 points)		

There will be a homework after each lecture. It will be due before next class. Homework will include database operations coding and practical questions.

In capstone project, we will design database for product and sales and students will do analysis using SQL language.

Assignments Submission:

Assignment must be turned in before the due time and uploaded to Piazza course management system and email TA. Late assignment will receive 10% deduction for every day it is late. (for example, homework 1 worth 100 points. If your grade is 94, you will receive 84 if late up to one day) If you have difficulties in complete the assignment, seeking for help ahead of time.

Grades:

Students are expected to receive their grades one or two weeks after submission. Grading will be fair and straightward. If you find any grading mistakes, please email no later than 3 days after you receive the grade. After that, the assignment will not be re-evaluated.

ADDITIONAL INFORMATION

Add/Drop Process

The last day to add the class or withdraw without receiving a “W” is Jan 15,2018. The last day to drop with a mark of a “W” is Jan 29,2018.

If you are absent 2 or more times prior to Jan 29,2018 (the last day to withdraw from a course with a grade of “W”), I may ask you to withdraw from the class by that date.

Retention of Graded Coursework

Final exams and all other graded work which affected the course grade will be retained for one year after the end of the course *if* the graded work has not been returned to the student (i.e., if I returned a graded paper to you, it is your responsibility to file it, not mine).

Technology Policy

Laptops are permitted for academic purpose. Please silence your cell phone during the lecture.

Academic Integrity and Conduct

USC seeks to maintain an optimal learning environment. General principles of academic honesty include the concept of respect for the intellectual property of others, the expectation that individual work will be submitted unless otherwise allowed by an instructor, and the obligations both to protect one’s own academic work from misuse by others as well as to avoid using another’s work as one’s own (plagiarism). 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. All students are expected to understand and abide by the principles discussed in the *SCampus*, the Student Guidebook (www.usc.edu/scampus or <http://scampus.usc.edu>). A discussion of plagiarism appears in the University Student Conduct Code (section 11.00 and Appendix A).

Students will be referred to the Office of Student Judicial Affairs and Community Standards for further review, should there be any suspicion of academic dishonesty. The Review process can be found at: <http://www.usc.edu/student-affairs/SJACS/>. Failure to adhere to the academic conduct standards set forth by these guidelines and our programs will not be tolerated by the USC Marshall community and can lead to dismissal.

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://dps.usc.edu/contact/report/>. 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. *Relationship and Sexual Violence Prevention and Services* (RSVP) <https://engemannshc.usc.edu/rsvp/> provides 24/7 confidential support, and the sexual assault resource center webpage <https://sarc.usc.edu/reporting-options/> describes reporting options and other resources.

Support Systems

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* (www.usc.edu/disability) 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.

Students with Disabilities/Learning Differences

The Office of Disability Services and Programs (www.usc.edu/disability) provides certification for students with disabilities and helps arrange the relevant accommodations. 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 me (or to your TA) as early in the semester as possible. DSP is located in GFS (Grace Ford Salvatori Hall) 120 and is open 8:30 a.m.–5:00 p.m., Monday through Friday. The phone number for DSP is (213) 740-0776. Email: ability@usc.edu.

Emergency Preparedness/Course Continuity

In case of a declared emergency if travel to campus is not feasible, the *USC Emergency Information* web site (<http://emergency.usc.edu/>) will provide safety and other information, including electronic means by which instructors will conduct class using a combination of Blackboard, teleconferencing, and other technologies.

Please make sure you can access this course in Blackboard and retrieve the course syllabus and other course materials electronically. You should check Blackboard regularly for announcements and new materials. In the event of an emergency, the ability to access Blackboard will be crucial. USC's Blackboard learning management system and support information is available at blackboard.usc.edu.

Incomplete Grades

A mark of IN (incomplete) may be assigned when work is not completed because of a documented illness or other “emergency” that occurs after the 12th week of the semester (or the twelfth week equivalent for any course that is scheduled for less than 15 weeks).

An “emergency” is defined as a serious documented illness, or an unforeseen situation that is beyond the student’s control, that prevents a student from completing the semester. Prior to the 12th week, the student still has the option of dropping the class. Arrangements for completing an IN must be initiated by the student and agreed to by the instructor prior to the final examination. If an Incomplete is assigned as the student’s grade, the instructor is required to fill out an “**Assignment of an Incomplete (IN) and Requirements for Completion**” form which specifies to the student and to the department the work remaining to be done, the procedures for its completion, the grade in the course to date, and the weight to be assigned to work remaining to be done when the final grade is computed. Both the instructor and student must sign the form with a copy of the form filed in the department. Class work to complete the course must be completed within one calendar year from the date the IN was assigned. The IN mark will be converted to an F grade should the course not be completed within the time allowed.

Grade Disputes

All grades assigned by faculty members are final. Students have the right to seek explanation, guidance, counsel and reasons for the assignment of a grade. Faculty may initiate a change in grade if there is an error in the calculation of a grade. Students may appeal a grade according to university policy as set forth in *SCampus*. A faculty member may not change a disputed grade outside the formal appeals process. In response to a disputed academic evaluation by an instructor, a student is entitled to two levels of appeal after review by the instructor: first to the chairperson of the department and then to the appropriate dean of the school. The full university policy can be found in *SCampus* under University Governance / Academic Policies at <http://scampus.usc.edu/academic-policies/>.

COURSE CALENDAR/READINGS/CLASS SESSIONS

Week	Topic	Homework & Reading Material	Deliverables and Due Dates
Week 1	<ol style="list-style-type: none"> 1. Anatomy of relational database 2. Data type 3. Creating database 4. Creating tables 5. Loading data 6. Indexing tables 	Homework 1 Chapter 1,2,3,15,16,17	
Week 2	<ol style="list-style-type: none"> 1. Creating subset 2. Using expression 3. SQL functions 4. Data computation 	Homework 2 Chapter 4 - 6	Homework 1 Due
Week 3	<ol style="list-style-type: none"> 1. Operation of data set 2. Join tables 3. Unions 4. Views 	Homework 3 Chapter 7,8,9,10	Homework 2 Due
Week 4	<ol style="list-style-type: none"> 1. Subquery as column expression 2. Subquery as filters 3. Query optimization 	Homework 4 Chapter 11,12,13,14,18,19,20	Homework 3 Due
Week 5	<ol style="list-style-type: none"> 1. Hadoop big data platform 2. HDFS 3. Data processing using Hive 	Homework 5 Online Hive programming manual https://cwiki.apache.org/confluence/display/Hive/LanguageManual	Homework 4 Due
Week 6	<ol style="list-style-type: none"> 1. Spark big data platform 2. Dataset and Data-Frame 3. Spark SQL 	Homework 6 Online SparkSQL Guide http://spark.apache.org/docs/latest/sql-programming-guide.html	Homework 5 Due
Week 7	<p>Capstone project</p> <p>Online store database, sale analysis and forecasting</p>		Homework 6 Due Exam
Week 8			Capstone project Due