

DSO 555: Advanced SQL for Business Analysts Spring 2025

The class will meet twice a week for 1h20m each time.

Instructor:	<i>Yola Katsargyri</i>	
Office:	<i>BRI 400B</i>	
Email:	katsargy@marshall.usc.edu	
Zoom:	Zoom Link	
Units:	1.5	
Yola's Office Hours:	<i>W: 9:45-12:15pm, 3:30-6pm</i>	<i>(in person, BRI 400B)</i>
	<i>Th: 12:45am-1:45pm</i>	<i>(in person, BRI 400B)</i>
Graders:	Maniyar, Jaykumar Suryaprakash (Jay), jaykumarsuryaprakash.maniyar.2024@marshall.usc.edu (contact Jay for grading or clarification questions)	
	<i>Rohil Wattal, rwattal@usc.edu</i>	
	<i>Sisodia, Rishabh Pratap Singh,</i> rishabhpratapsingh.sisodia.2024@marshall.usc.edu <i>Kashish Jain, kashishj@usc.edu</i>	
TA Office Hours:	<i>TBD</i>	

COURSE DESCRIPTION

Structured query language (SQL) is the language of databases, one of the most broadly used in research, academia and industry for big data manipulation and analysis. By the end of this course, students will be proficient in constructing complex queries and they will have solid competency in conducting comprehensive exploratory data analyses (EDAs) and producing actionable insights and recommendations. Additionally, they will gain experience in query planning, indexing, and optimization, as well as data modeling and architecture. These valuable skills will prepare graduates to excel in any data-driven business or organization and are highly transferable across various fields. Furthermore, these competencies are essential for students aspiring to careers in data science or data engineering.

PREREQUISITE: This course is designed for students who have completed their SQL requirement, either by testing out of DSO552 or by previously completing it.

COURSE OBJECTIVES

Upon successful completion of this course, students will be able to:

1. Write and understand complex SQL queries.
2. Optimize queries using views, indexes, transactions and query planning.
3. Leverage data modeling techniques.
4. Perform and present extensive exploratory data analyses (EDA) in a business context.

COURSE MATERIALS AND READINGS

There is no required textbook for this class. Course materials are posted on Brightspace. However, the following material and readings can be helpful:

Readings:

1. *Learn PostgreSQL*
By Luca Ferrari, Enrico Pirozzi
2. Hoffer, J., Topi, H & Ramesh, V. (2014). *Essentials of Database Management*. Pearson
ISBN:9780133405682
3. *PostgreSQL Query Optimization: The Ultimate Guide to Building Efficient Queries*
By Henrietta Dombrovskaya

Helpful Links & Extra Practice

<https://www.w3schools.com/sql/>

<https://leetcode.com/studyplan/top-sql-50/>

<https://github.com/xoraus/CrackingTheSQLInterview>

<https://www.educba.com/data-science/data-science-tutorials/postgresql-tutorial/>

<https://www.postgresqltutorial.com/>

<https://www.prisma.io/dataguide/postgresql>

<https://www.codecademy.com/learn/dscp-sql-for-interview-prep/modules/dacp-sql-interview-prep/cheatsheet>

https://sqlzoo.net/wiki/SQL_Tutorial

<https://www.w3resource.com/sql-exercises/>

<https://www.hackerrank.com/domains/sql>

<https://www.postgresql.org/docs/current/index.html>

<https://www.sqlcourse2.com/select2.html>

Installation

Please install PostgreSQL (<https://www.postgresql.org/>), an open-source relational database system. In this class, we will use **DBeaver** as our SQL environment, and you are strongly encouraged to use it as well, in order to easily follow the same steps. However, you may use any other IDE (integrated development environment) if you feel much more comfortable with it. You may notice that Pgadmin4 comes as a default IDE when you install Postgresql, however, this editor has been buggy lately, so it

is not our recommendation, especially for Macs. More detailed software instructions will be posted on BRIGHTSPACE. You are responsible for ensuring that you have the necessary computer equipment that will allow you use the software.

GRADING

Your final grade will be determined as follows:

<u>Assignments</u>	<u>% of Overall Grade</u>
<i>Homework</i>	20%
<i>Group Project</i>	20%
<i>In-class Quizzes</i>	20%
<i>Midterm</i>	40%
TOTAL	100%

The weights listed above will be used to calculate your overall score for the class. Final grades represent how you perform in the class relative to other students. Three items are considered when assigning final grades:

1. Your total weighted score for all homeworks, group project and exams.
2. Your rank among all students in the class.
3. The cut-off thresholds between different letter grades. The thresholds for this course are given below:

PERCENTILE	LETTER GRADE
0th	B-
10th	B
35th	B+
55th	A-
65th	A

Table 2: Percentiles and letter grade cutoffs

Interpretation:

The percentile indicates the percentage of the class that performed worse overall. For example:

- If a student is in the 55th percentile, 55% of the class performed worse, while 45% performed better.
- If a student is in the 97th percentile, 97% of the class performed worse, and only 3% performed better.

A higher percentile corresponds to a better grade.

Grading Scale Based on Percentile:

- **A:** Percentile > 65th
- **A-:** Percentile between 55th and 65th
- **B+:** Percentile between 35th and 55th
- **B:** Percentile between 10th and 35th
- **B-:** Percentile < 10th

CLASS PARTICIPATION

You are expected to attend all classes, where expanded explanations and multiple examples for practice are provided and you are presented with the opportunity to pose your questions. I am happy to go over a concept multiple times. If you feel uneasy bringing up your questions in class, take advantage of the many opportunities to speak with me one-on-one. I am accessible by email and will be more than happy to speak with you during office hours.

You are also expected to take **3 short in-class quizzes** that help you digest the material. The dates of the quizzes will be provided to you upfront, so that you make sure you attend the corresponding class. **You have to attend the class during the time and day you are registered for.**

Finally, you are expected to **attend class** on the **day of the group project presentation.**

HOMEWORK

In this class, there will be a total of 3 homeworks. Homework will be due on **Mondays, 10am**, and will include coding exercises and practical questions. Students need to write their answers in a .doc or .pdf file and then submit it on Brightspace by the due date. The file should include for every answered question the following:

- The question number.
- The SQL query that was created to answer the question (code should be copied and pasted, not screenshot)
- A screenshot of the output that the query produces (if the output is too long, just provide a part of it, e.g. first 5 rows).

Homeworks are assessed as follows:

- Grading is based on the correctness of the code and not only the correctness of the output. Even if the output looks right, the code may not be entirely correct, and could lead to wrong outputs for a different set of data. We want to make sure that we learn to build queries that are correct for any common or edge cases.
- For every code snippet, a different amount of points is given to different SQL clauses, depending on their complexity (e.g. a simple SELECT may give 1 point, while a SELECT with a JOIN may give 2 points)
- For late submissions, there will be a penalty on your score: $Score = (Your\ raw\ score) * (1 - 0.2 * (\#\ of\ days\ past\ deadline))$ (no credit is given 5 days after the due date).

QUIZZES

You are expected to take **3 short in-class quizzes** that help you digest the material. The dates of the quizzes are provided to you upfront, so that you make sure you attend the corresponding class. **You have to attend the class during the time and day you are registered for.** You are allowed to drop 1 of your 3 quizzes, meaning that you can keep your best 2 quizzes towards your final grade. In case you miss one quiz, you can use that as the quiz that you want to drop. If you miss more than one quizzes, you will get 0 in all of those and you can only drop one of them from your final grade. For example, if a student skips 2 quizzes and gets full in the other 1: e.g.: 0,0,10 then their total quizzes score will be 10/20.

GROUP PROJECT

The material taught in class and the homeworks will culminate in a group project which will require you to address a business problem by analyzing the data and providing analytics insights to your stakeholders. Groups will be randomly assigned by the instructor and will consist of about 7 to 8 students each.

The project has 4 deliverables:

- **During the final exams period: The entire team** will present in class, via a *Powerpoint presentation*, the team's findings and any business recommendations that are important to the hypothetical stakeholders (16256: May 13, 8-10AM / 16257: May 13, 11AM-1 PM / 16258: May 8, 2-4PM / 16259: May 13, 2-4PM)
- **By April 21, 10 am: One person from each team** (whichever person is agreed upon by the entire team) needs to submit the team's *technical report*, including the code, exploratory data analysis and comments/explanations that led to the final presentation.
- (Optional) **By the time that your section's presentations start**, (16256: May 13, 8-10AM / 16257: May 13, 11AM-1 PM / 16258: May 8, 2-4PM / 16259: May 13, 2-4PM), **every team member may submit** their *Team 360 Review*, as suggested by the university, to assist the professor in assessing individual student contributions. The Team 360 Review will include the following two questions:
 - *Were there any team members that you think went above and beyond to ensure that the team's produced result would be as good as it gets? If yes, please provide their name(s) and examples of their great contribution.*
 - *Were there any team members that you think should have contributed more or better to the team's produced result, should have been more collaborative or somehow held the team's work back, in a way that you consider unfair? If yes, please provide their*

name(s) and examples of where you think their contribution should have been stronger. (Please, exclude from your answer students that had extenuating circumstances that would justify lessening their help to the team.)

- If a team member's name is given as an answer in question 1 by at least three other team members, then the former may receive an extra 5% towards their final grade.
- If a team member's name is not given as an answer to either of the questions, then the former will receive an extra 2.5% towards their final grade.
- If a team member's name is given as an answer in question 2 by at least two other team members, then the former may receive an extra 0% towards their final grade.

The Team 360 Review provides students with an opportunity to inform the professor about exceptional collaborators and, more importantly, to raise concerns about instances where collaboration was dysfunctional. This includes situations where team dynamics may have negatively impacted overall performance or where certain team members were treated unfairly. While students are encouraged to foster seamless collaboration and address challenges within their teams independently, this review offers a confidential space to share concerns about unresolved issues or unfair treatment.

Since the group project presentation dates are disclosed from the beginning of the semester, **no rescheduling is allowed.**

Group project is assessed as follows:

- The entire team is graded equally (i.e. every member of a team will get the same grade) for the following parts:
 - Team's Powerpoint Deck (45%)
 - Team's Technical Report (45%)
- On top of the above, each student will be assessed independently for the following parts:
 - Presentation skills (e.g. presenting without notes, interacting with the audience, etc) (5%)
 - Team 360 Review, if available (0-5%, as explained above)

EXAM

There will be one midterm exam, in-class, with open notes. The midterm will take place on **April 15**. The lecture of April 10 will be the final review of the material.

Using AI Generators is not allowed. Out of fairness to all students, if anyone is caught to break this agreement or to cheat in any other way, they will immediately fail the class.

If you miss the exam you will receive a grade of zero for that exam. You **cannot** be exempted from the exam under any circumstances. **The exam will NOT be given at any other time.** Since the exam date is disclosed from the beginning of the semester, **no rescheduling is allowed.**

There will be **no final exam.**

Exam is assessed as follows:

The exam is assessed in a way similar to the one described for the homeworks (even though no late submission penalty is needed here).

Again, note that according to the USC Office of Academic Records and Registrar, “No student is allowed to re-take an examination or do extra work in a course after the semester has ended for purposes of improving their grade.”

COLLABORATION POLICY

Discussion of practice problems is permitted and encouraged; however, each student is required to prepare and submit their solutions independently.

Collaboration of any sort on quizzes and exams is prohibited and will, at minimum, result in a 0 on that exam or quiz.

THE IMPORTANCE OF COURSE EVALUATIONS

The student course evaluations are valuable. This course is continuously improved, based on students feedback and instructor observations.

ADD/DROP PROCESS

Please check the schedule of classes [here](#) for the ADD/DROP dates.

OPEN EXPRESSION AND RESPECT FOR ALL

An important goal of the educational experience at USC Marshall is to be exposed to and discuss diverse, thought-provoking, and sometimes controversial ideas that challenge one’s beliefs. In this course we will support the values articulated in the USC Marshall “[Open Expression Statement](https://www.marshall.usc.edu/open-expression-statement)” (<https://www.marshall.usc.edu/open-expression-statement>).

TENTATIVE COURSE OUTLINE AND ASSIGNMENTS

Dates	Topic	Due
<i>Week 1</i> Mar 11 / 13, 2025 (T/Th)	552 Material Review, Dates Manipulation, String Functions, UNION and CASE, Intro to EDA (Exploratory Data Analysis)	N/A

<i>Week 2</i> Mar 18 / 20, 2025 (T/Th)	Spring Break	N/A
<i>Week 3</i> Mar 25 / 27, 2025 (T/Th)	Subqueries and CTEs	HW1 (Mar 24, 10am) Quiz 1 (Mar 25, in-class)
<i>Week 4</i> Apr 1 / 3, 2025 (T/Th)	Window functions	HW2 (Mar 31, 10am) Quiz 2 (Apr 1, in-class)
<i>Week 5</i> Apr 8 / 10, 2025 (T/Th)	Practice session and EDA Midterm review	HW3 (Apr 7, 10am) Quiz 3 (Apr 8, in-class)
<i>Week 6</i> Apr 15 / 17, 2025 (T/Th)	Midterm (April 15) EDA	N/A
<i>Week 7</i> Apr 22 / 24, 2025 (T/Th)	Views, Indexes, Query planning and Optimization Data Modeling	Project Technical Report (Apr 21, 10am)
<i>Week 8</i> Apr 29 / May 1, 2025 (T/Th)	EDA	N/A
<i>Final Exams Period</i> May 7-14, 2025	Project Presentations 16256: May 13, 8-10AM 16257: May 13, 11AM-1 PM 16258: May 8, 2-4PM 16259: May 13, 2-4PM	

STATEMENT ON ACADEMIC CONDUCT AND SUPPORT SYSTEMS

Policy on AI:

Since producing code, as well as developing analytical and critical thinking skills are part of the learning outcomes of this course, all assignments should be prepared by the student working individually or in groups. Students may not have another person or AI tool complete any substantive portion of the assignment. Developing strong competencies in these areas will prepare you for a competitive workplace. Therefore, using AI-generated tools is prohibited in this course, will be identified as plagiarism, and will be reported to the Office of Academic Integrity.

Academic Integrity:

The University of Southern California is foremost a learning community committed to fostering successful scholars and researchers dedicated to the pursuit of knowledge and the transmission of ideas. Academic misconduct is in contrast to the university’s mission to educate students through a

broad array of first-rank academic, professional, and extracurricular programs and includes any act of dishonesty in the submission of academic work (either in draft or final form).

This course will follow the expectations for academic integrity as stated in the [USC Student Handbook](#). All students are expected to submit assignments that are original work and prepared specifically for the course/section in this academic term. You may not submit work written by others or “recycle” work prepared for other courses without obtaining written permission from the instructor(s). Students suspected of engaging in academic misconduct will be reported to the Office of Academic Integrity.

Other violations of academic misconduct include, but are not limited to, cheating, plagiarism, fabrication (e.g., falsifying data), knowingly assisting others in acts of academic dishonesty, and any act that gains or is intended to gain an unfair academic advantage.

Academic dishonesty has a far-reaching impact and is considered a serious offense against the university. Violations will result in a grade penalty, such as a failing grade on the assignment or in the course, and disciplinary action from the university itself, such as suspension or even expulsion.

For more information about academic integrity see the [student handbook](#) or the [Office of Academic Integrity's website](#), and university policies on [Research and Scholarship Misconduct](#).

Please ask your instructor if you are unsure what constitutes unauthorized assistance on an exam or assignment or what information requires citation and/or attribution.

Statement on Academic and Support Systems

Students and Disability Accommodations:

USC welcomes students with disabilities into all of the University's educational programs. The Office of Student Accessibility Services (OSAS) is responsible for the determination of appropriate accommodations for students who encounter disability-related barriers. Once a student has completed the OSAS process (registration, initial appointment, and submitted documentation) and accommodations are determined to be reasonable and appropriate, a Letter of Accommodation (LOA) will be available to generate for each course. The LOA must be given to each course instructor by the student and followed up with a discussion. This should be done as early in the semester as possible as accommodations are not retroactive. More information can be found at osas.usc.edu. You may contact OSAS at (213) 740-0776 or via email at osasfrontdesk@usc.edu.

Student Financial Aid and Satisfactory Academic Progress:

To be eligible for certain kinds of financial aid, students are required to maintain Satisfactory Academic Progress (SAP) toward their degree objectives. Visit the Financial Aid Office webpage for undergraduate- and graduate-level SAP eligibility requirements and the appeals process.

Support Systems:

Counseling and Mental Health - (213) 740-9355 – 24/7 on call

Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

988 Suicide and Crisis Lifeline - 988 for both calls and text messages - 24/7 on call

The 988 Suicide and Crisis Lifeline (formerly known as the National Suicide Prevention Lifeline) provides free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week, across the United States. The Lifeline consists of a national network of over 200 local crisis centers, combining custom local care and resources with national standards and best practices. The new, shorter phone number makes it easier for people to remember and access mental health crisis services (though the previous 1 (800) 273-8255 number will continue to function indefinitely) and represents a continued commitment to those in crisis.

CARE-SC: Confidential Advocacy, Resources, and Education Support Center - (213) 740-9355(WELL) - 24/7 on call

Free and confidential therapy services, workshops, and training for situations related to gender- and power-based harm (including sexual assault, intimate partner violence, and stalking).

Office for Equity, Equal Opportunity, and Title IX (EEO-TIX) - (213) 740-5086

Information about how to get help or help someone affected by harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants.

Reporting Incidents of Bias or Harassment - (213) 740-2500

Avenue to report incidents of bias, hate crimes, and microaggressions to the Office for Equity, Equal Opportunity, and Title for appropriate investigation, supportive measures, and response.

The Office of Student Accessibility Services (OSAS) - (213) 740-0776

OSAS ensures equal access for students with disabilities through providing academic accommodations and auxiliary aids in accordance with federal laws and university policy.

USC Campus Support and Intervention - (213) 740-0411

Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

Diversity, Equity and Inclusion - (213) 740-2101

Information on events, programs and training, the Provost's Diversity and Inclusion Council, Diversity Liaisons for each academic school, chronology, participation, and various resources for students.

USC Emergency - UPC: (213) 740-4321, HSC: (323) 442-1000 - 24/7 on call

Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

USC Department of Public Safety - UPC: (213) 740-6000, HSC: (323) 442-1200 - 24/7 on call

Non-emergency assistance or information.

Office of the Ombuds - (213) 821-9556 (UPC) / (323-442-0382 (HSC)

A safe and confidential place to share your USC-related issues with a University Ombuds who will work with you to explore options or paths to manage your concern.

Occupational Therapy Faculty Practice - (323) 442-2850 or otfp@med.usc.edu

Confidential Lifestyle Redesign services for USC students to support health promoting habits and routines that enhance quality of life and academic performance.