

DSC 450 Database Processing for Large-Scale Analytics
Syllabus
Tanu Malik

Overview

This is an introductory graduate course in database design and will introduce you to principles of database management, such as designing data using relational model and principles of normalization, querying data with structured query language, data extract, transform, load, and query using Python.

Course Objectives

Learn the principles of data management in the context of a relational database management system. Learn how to

1. Store data using the relational model.
2. Query and retrieve data from an RDBMS using SQL.
3. Design database schemas using normalization.
4. Make RDBMSes interact with other programming languages using Python: PL/SQL and SQLAlchemy
5. Regular expressions.
6. Data analysis in Python: numpy and pandas

Instructor Information

Instructor: Prof. Tanu Malik

Loop Office: 836, CDM Building, 243 S. Wabash Ave

Email: tanu@cdm.depaul.edu

Phone: (312) 362-1121

Instructor Home Page: <http://facsrv.cs.depaul.edu/~tmalik1/>

Course Home Page: <http://d2l.depaul.edu>

(for lectures slides, assignments, reading schedules, examples, learning outcomes)

LMS Homepage: <http://d2l.depaul.edu> (for grades, quizzes, and video recordings)

Prerequisites

If you are not sure that you have satisfied the prerequisites, speak to the instructor before the second lecture.

Prerequisite Courses:

- Introduction to Programming (CSC401)

Pre-requisite Skill:

- A solid understanding of beginning level Python. If your requirement has been met through a different programming language, make sure to study the equivalent

in Python. You will be expected to know all concepts and syntax covered in CSC401 such as data types, loops, iteration, functions and more. A solid understanding of these concepts prior to the start of the class will be critical to your success.

Required Text

Required textbook:

- Python for Data Analysis, by Wes McKinney, O'Reilly 2012. ISBN 1449319793 or 978-1449319793

Additional reading:

- Oracle Database 12c SQL, by Jason Price. ISBN 9780071799355 (available in DePaul books 24x7 library)

Assessment

- The course grade will be based on:

Homework Assignments: 25%,

Quizzes: 15%,

Mid-Term Exam: 30%

Final Exam: 30%.

- The grade boundaries used for the class are as follows:

Letter Grade	Percentage >=
A	92%
A-	88%
B+	85%
B	81%
B-	77%
C+	73%
C	69%
C-	65%
D+	61%
D	57%
F	0%

- Quizzes are multiple choice and multiple select.
- The mid-term exam is multiple choice and multiple select.
- The final exam is programming-based.
- You must pass the midterm and final exams (grade D or better) to pass the course.
- To provide the same testing environment for in-class and Online Learning students, the instructor will not answer questions during the exams.

Changes to Syllabus

This syllabus is subject to change as necessary during the quarter. If a change occurs, it will be thoroughly addressed during class, posted under Announcements in D2L.

Policy I: Student Expectations

Attendance

1. No attendance is taken in the class. Students are expected to attend class or watch the online recording within 48 hours of the live class.
2. Students are expected to read posts on D2L in a timely fashion.

Attending the Class

3. All students must keep up with the assigned textbook reading.
4. Lecture slides are a supplement to lectures only. The slides are not intended to be read in lieu of listening to the lecture.
5. **In-class students** The instructor prefers that students take hand-written notes. Laptops, tablets, et cetera, that are used exclusively for note-taking or other course-related activities are permitted but non-class activities such as emailing, texting, game playing, web browsing, and other activities not related to course work during class are forbidden.
6. **Online Students** Online students must work out by hand all examples covered in class.
7. Students are strongly encouraged to ask questions and offer comments relevant to the day's topic.
8. The course requires that students actively engage the material on your own. Students should not only read the notes and example programs in class, but also do self-tests, modify code and queries, and run it. As always, figure out what you can definitely code, code it, try it, and then consider extending the boundaries.

Emailing the Instructor:

9. The **subject of your email must begin with DSC450**. If not, there is no guarantee of response.
10. You may address me as Professor, Prof. Malik, Dr. Malik, or simply Tanu.

11. All electronic interactions are an extension of the classroom and should be treated as such. While disagreement can be part of the discourse, online communication should remain respectful and appropriate rather than demeaning and/or unprofessional.

Homework

12. Homework assignments will be distributed and submitted via D2L. Emailed submissions will not be accepted.
13. Grades and comments will be returned through D2L.

14. No late submissions.

15. Submitted work must be worked on individually. You must not use or look at anyone else's solution, and you must clearly acknowledge any code that you obtain from other sources (such as books, magazines, or the Internet). If you are in any doubt, contact the instructor well before the submission date for advice. You may use as much code as you like (without acknowledgement) from the examples discussed in class. Plagiarism will result in penalties up to and including failing the course.

Exams

17. The midterm exam and final exam dates are posted on the schedule on the course homepage.
18. You must attend the midterm and final exams. **No make-up exams will be given.**
19. A medical note, cc'ed to the Dean of Students and your Graduate Advisor, will be required for an absence. Business trips or vacations are not valid reasons for missing the exam. Please register for the exam as soon as possible.

Policies II: University and CDM Rules

Incomplete Grade

An incomplete grade is defined in the Student Handbook as follows (note that the policy in the undergraduate student handbook applies to both undergraduate and graduate students): A temporary grade indicating that the student has a satisfactory record in work completed, but for unusual or unforeseeable circumstances not encountered by other students in the class and acceptable to the instructor is prevented from completing the course requirements by the end of the term. Please see <http://www.cdm.depaul.edu/Current%20Students/Pages/Grading-Policies.aspx> for additional information. Any such reason must be documented. Any incomplete request must be made at least two weeks before the final, and approved by the Dean of CDM. Any consequences resulting from a poor grade for the course will not be considered as valid reasons for such a request.

Retro-Active Withdrawal

CDM understands certain extenuating circumstances can hinder one's ability for academic success and completion of course work. Please see <http://www.cdm.depaul.edu/Current%20Students/Pages/Enrollment-Policies.aspx> for additional information.

Absence Notifications

In order to petition for an excused absence, students who miss class due to illness or significant personal circumstances should complete the Absence Notification process through the Dean of Students office. The form can be accessed at <http://studentaffairs.depaul.edu/dos/academicprocesses.html>. Students must submit supporting documentation alongside the form. The professor reserves the sole right whether to offer an excused absence and/or academic accommodations for an excused absence.

Academic Integrity and Plagiarism

This course will be subject to the university's academic integrity policy.

- **Cheating:** Cheating is any action that violates university norms or instructor's guidelines for the preparation and submission of assignments. This includes but is not limited to unauthorized access to examination materials prior to the examination itself, use or possession of unauthorized materials during the examination or quiz; having someone take an examination in one's place-copying from another student; unauthorized assistance to another student; or acceptance of such assistance.

- **Plagiarism:** Plagiarism is a major form of academic dishonesty involving the presentation of the work of another as one's own. Plagiarism includes but is not limited to the following:
 - Pasting homework solutions on emails or discussion boards.
 - The direct copying of any source, such as written and verbal material, computer files, audio disks, video programs or musical scores, whether published or unpublished, in whole or part, without proper acknowledgement that it is someone else's.
 - Copying of any source in whole or part without proper acknowledgement.
 - Submitting as one's own work a report, examination paper, computer file, lab report or other assignment that has been prepared by someone else. This includes research papers purchased from any other person or agency.
 - The paraphrasing of another's work or ideas without proper acknowledgement.
- **Complicity:** Complicity is any intentional attempt to facilitate any of the violations described above. This includes but is not limited to allowing another student to copy from a paper or test document; providing any kind of material—including one's research, data,

or writing—to another student if one believes it might be misrepresented to a teacher or university official; providing information about or answers to test questions.

More information can be found at <http://academicintegrity.depaul.edu/>. If you have any questions be sure to consult with your professor.

Academic Calendar

All students are required to manage their class schedules each term in accordance with the deadlines for enrolling and withdrawing as indicated in the University Academic Calendar. Information on enrollment, withdrawal, grading and incompletes can be found at: <http://cdm.depaul.edu/enrollment>

Students with Disabilities

Students who feel they may need an accommodation based on the impact of a disability should contact the instructor privately to discuss their specific needs. All discussions will remain confidential. To ensure that you receive the most appropriate accommodation based on your needs, contact the instructor as early as possible in the quarter (preferably within the first week of class), and make sure that you have contacted the Center for Students with Disabilities (CSD) at: csd@depaul.edu

Lewis Center 1420, 25 East Jackson Blvd.
Phone number: 312 362 8002
Fax: 312 362 6544
TTY: 773 325 7296

Dean of Students' Office

The Dean of Students' Office (DOS) helps students navigate the college experience, particularly during difficulty situations such as personal, financial, medical, and/or family crises. For a list of support services and advocacy information, please visit <http://studentaffairs.depaul.edu/dos/>.

Online Course Evaluations

Evaluations are a way for students to provide valuable feedback regarding their instructor and the course. Detailed feedback will enable the instructor to continuously tailor teaching methods and course content to meet the learning goals of the course and the academic needs of the students. The evaluations are anonymous; the instructor and administration do not track who entered what responses. A program is used to check if the student completed the evaluations, but the evaluation is completely separate from the student's identity. Since 100% participation is our goal, students are sent periodic reminders over three weeks. Students do not receive reminders once they complete the evaluation. Students complete the evaluation online in CampusConnect.