

COMP-421: Database Systems



Joseph D'silva

joseph.dsilva@mail.mcgill.ca

McConnel Engg. 102

Names and Numbers

- Class:
 - Mondays, Wednesdays | 11:35-12:55
- Lecturer:
 - Joseph D'silva
 - joseph.dsilva@mail.mcgill.ca
 - Office Hours:
 - after class

Lectures on Mar 20th and 22nd will be delivered by Prof. Kemme.

Lectures will be recorded
- TAs:
 - Faizy Ahsan (TA coordinator)
 - Sanja Thakur
 - Shuonan (Elaine) Dong
 - Ramchalam Kinattinkara Ramakrishnn

TA Office hours will be scheduled based on assignment deadlines and will be announced in mycourses accordingly.

Marking Scheme

- Marking Scheme:
 - 3 Individual Assignments total of 12%
 - Each worth 4%
 - 1 Map-Reduce Exercise total of 5% (TBD)
 - 3 Project Deliverables total of 15%
 - Each worth 5%
 - Midterm 10%,
 - Final 58%
- **Midterm: March – 8th, 6:00 pm (Tentative)**
- Final: 3hr final covers all
- Both Midterm and Finals can contain multiple choice questions.

Course Prerequisites

- COMP 206 Introduction to Software Systems
- COMP 251 Algorithms and Data Structures
- COMP 302 Programming Languages and Paradigms

Marking Scheme

- **Written Assignments:**
 - Solving concrete problems, from easy to difficult
 - Around one assignment every two – three weeks, three in total
 - Usually to be done within one or two weeks.
 - Late turn-in: for each day late 10% of the maximal achievable points will be subtracted from the achieved points. Maximum of 4 late days allowed unless otherwise mentioned.
 - Possible that there is one assignment around the midterm with no possibility of late turn in.
 - SQL Assignment (Part of A2) will be graded by an automated script (details will be provided later).
- There will be a 4th assignment. This will not be graded if the Map-Reduce assignment is given out. Otherwise this will instead constitute the 5% of the grade currently attributed to the Map-Reduce assignment.

Marking Scheme

- Map-Reduce Assignment: (TBD)
 - Second half of the semester
 - Using the map-reduce framework to solve large-scale queries
 - In groups
 - Same group as the project group
 - If the Map-Reduce assignment is not given out, the 4th assignment will be graded. The 4th assignment will have to be done individually like the prior the assignments.

Marking Scheme

- **Project Deliverables:**
 - Designing and building a database, working with the database, and writing applications using the database
 - Practical work with DB2 / PostgreSQL / your system of choice (email me if you pick another DBMS)
 - 3 project deliverables
 - (stepwise project development)
 - To be done in a given time period
 - Four people will build a team

Project Groups

- To be done by next week!
- Find partners to build a group of 4 (minimum 3)
- Once you find sufficient people to build a group,
 - Go to mycourses
 - Go to groups menu
 - Choose an empty group and sign up to it (all 4 of you must sign up for the same group number)
- If you are not enough to build a group
 - Go to mycourses / post under “Group Finding Project” to find teams/members.

Project Special: for DB experts

Tentative !!

- Are you taking COMP 551 this semester ?
- Do you and your group members have already DB experience?
 - Did you already write often SQL queries?
 - Did you do already SQL application programming?
 - E.g., Java and JDBC?
- Do you want to have a challenge?
- Do you want to write something special on your CV? Graduate School?
- Do you want to do something that you usually can practice only in smaller classes (of which SOCS currently does not have many?)

Project Special: for DB experts

Tentative !!

Two groups will be given permission to do a combined project with comp 55I instead of their comp 42I Phase II and Phase III deliverables.

- **Groups should be of size 3 instead of 4** (as per comp 55I guidelines). You will work with the same 3 group members of your 42I project group for the 4th Deliverable of comp 55I.
- You will build a database to load the ML project's data, use it for feature analysis, feature engineering and even use some in-database analytics for “learning”.
- Your deadlines for this project will be as per the 55I deadlines.
- You will turn in the same submission of comp 55I to comp 42I as well.
- You will in addition write a report of how using a database came in handy for various aspects of your comp 55I project.
- (TBD) You will give a presentation in class based on your report
- Groups will be given permission only after Add/Drop deadline and possibly after comp 42I Project Phase I deadline.

Project Special 2

- Do you work for a company, a small start up, and you have the feeling you could need a database for some of your data management tasks?
 - Write me an email

Literature

- Lecture Notes
- Books:
 - Course Book: [**4 Books available in library on 48 hour loan**]
 - Raghu Ramakrishnan and Johannes Gehrke: Database Management Systems; McGraw Hill, 3rd Edition
 - Other Recommendations:
 - Hector Garcia-Molina, Jeffrey D. Ullman and Jennifer Widom: Database Systems: The Complete Book, 2nd edition, 2008.
 - Silberschatz, Korth, Sudarshan: Database System Concepts, McGraw Hill, 5th Edition, 2006

Database Management Systems

Third Edition



NEW material on Database Applications

Ramakrishnan • Gehrke

Sixth Edition

Database System Concepts

Abraham Silberschatz • Henry F. Korth • S. Sudarshan

DATABASE SYSTEMS

THE
COMPLETE
BOOK

SECOND EDITION

Hector Garcia-Molina
Jeffrey D. Ullman
Jennifer Widom

- If you are complete nuts about databases
 - Encyclopedia of Database Systems
 - Written by 800+ experts
 - 3800+ pages



Socrative Clicker

- Web App -
<https://b.socrative.com/login/student/>
- iOS App -
<https://itunes.apple.com/ca/app/socrative-student/id477618130?mt=8>
- Android App -
<https://play.google.com/store/apps/details?id=com.socrative.student>

Academic Integrity

McGill University values academic integrity.

Therefore all students must understand the meaning and consequences of cheating, plagiarism and other academic offences under the Code of Student Conduct and Disciplinary Procedures (see <http://www.mcgill.ca/integrity> for more information).

French/English policy

- In accord with McGill University's Charter of Students' Rights, students in this course have the right to submit in English or in French any written work that is to be graded.

Active Involvement

- Use my courses for discussions
 - There's a discussion forum for every deliverable.
 - Your classmates might have the information you need already or might be interested in the same information.
 - Helps relevant TA's / lecturers to monitor and answer questions.
- Check your McGill email routinely. There might be important announcements, your team mates might be trying to reach out to you. You are responsible to be actively involved in your project group.

Course Outline

- Introduction
- Entity-Relationship Model
- Relational Model and Data Definition Language
- Relational Algebra
- SQL I: Simple Queries
- SQL II: Advanced Queries
- SQL III: Constraints
- SQL V: Application Programming
- Buffer Management
- Indexing
- Query Execution
- Query Optimization
- Key-value stores and map-reduce
- Transactions
- Concurrency Control
- Trends