Practice and Applications of Data Management

CMPSCI 345

Lecture 01: Course Overview
Class goals

- Data is everywhere!

- Computer Scientists help manage this data
  - Help scientists achieve new discoveries
  - Help improve services
  - Help organizations run more efficiently

- This course:
  - Like “intro to programming” for DB
Why take this class

- You want to learn how to use databases, and what you can do with them
- You don’t really care (yet) about how they are implemented (take 445 for that)

Eligibility:
- You got C or higher in 187
- You have NOT taken 445
Course resources

- **Website:**
  
  http://avid.cs.umass.edu/courses/345

- **Moodle:**
  
  https://moodle.umass.edu/
  
  - Assignment submission
  - Grades
  - To-do list for each week

- **Piazza**
  
  - Discussion forum
Course Format

- MoWe, 2:30pm – 3:45pm, Goessmann Lab 151
  - Combination of lecture and in-class practice

- Homework deliverables
  - Individual assignments, quizzes, and a group project

- Exams
  - Midterm and final

You need to bring a laptop to class!
This course will provide a comprehensive introduction to the use of data management systems within the context of various applications. Some of the covered topics include data models, basic and advanced SQL, design and tuning of relational schemas, implementation of basic transactions, and data visualization tools. The class will alternate between lecture and practice, and the students will experience the covered topics through a semester-long collaborative mini-project.

Course work will include homework assignments, web quizzes, a collaborative mini-project, and a midterm and final exam.

Prerequisites: CS 187 (grade C or greater). 3 credits.

Students who have taken CMPSCI 445 are not eligible to take this class!

**Course Time:** Mo We 2:30 pm - 3:45 pm, Goessmann Lab Addtn rm 151

**Professor:** Alexandra Meliou  
Contact: ameli [at] cs [dot] umass [dot] edu  
Office Hours: Wednesday 10-11am, CS 330 (subject to change; please check back for updates)

**Teaching Assistant:** Matteo Brucato  
Contact: matteo [at] cs [dot] umass [dot] edu  
Office Hours: Tuesday 2-3pm (location TBA)

**Text:**  
The textbook for this course is the 2nd Edition of "Database Systems: The Complete Book" by Garcia-Molina, Ullman, and Widom. The textbook is available from Amazon. The lecture notes will be posted online after each class.

The textbook is optional. However, you should talk to the instructor if you want to consider a different textbook, or to not use a textbook at all, in order to understand the tradeoffs before you make this choice.
<table>
<thead>
<tr>
<th>week</th>
<th>date</th>
<th>day</th>
<th>topic</th>
<th>textbook</th>
<th>slides</th>
<th>homework deliverables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>quiz</td>
</tr>
<tr>
<td>Week 1</td>
<td>Sep 7</td>
<td>Mon</td>
<td>No class: Labor day</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sep 9</td>
<td>Wed</td>
<td>Course Overview</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 2</td>
<td>Sep 14</td>
<td>Mon</td>
<td>Data Models</td>
<td>Ch 2</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Sep 16</td>
<td>Wed</td>
<td>Basic SQL</td>
<td>Ch 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 3</td>
<td>Sep 21</td>
<td>Mon</td>
<td>Basic SQL</td>
<td>Ch 6</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Sep 28</td>
<td>Mon</td>
<td>Aggregates in SQL</td>
<td>Ch 6</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Sep 30</td>
<td>Wed</td>
<td>Nested Queries</td>
<td>Ch 6</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Week 4</td>
<td>Oct 5</td>
<td>Mon</td>
<td>E/R Diagrams</td>
<td>Ch 4</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Oct 7</td>
<td>Wed</td>
<td>Database Tuning</td>
<td>Ch 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oct 12</td>
<td>Mon</td>
<td>No class: Colombus day</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oct 13</td>
<td>Tue</td>
<td>Views and Access Control</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oct 14</td>
<td>Wed</td>
<td>Constraints and Triggers</td>
<td>Ch 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 5</td>
<td>Oct 19</td>
<td>Mon</td>
<td>Midterm review</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Oct 21</td>
<td>Wed</td>
<td>Midterm (in class)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Week 6</td>
<td>Oct 26</td>
<td>Mon</td>
<td>Concurrency Control</td>
<td>Ch 18, 19</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oct 28</td>
<td>Wed</td>
<td>Basic Transactions</td>
<td>Ch 6</td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>
This course will provide a comprehensive introduction to the use of data management systems within the context of various applications. Some of the covered topics include data models, basic and advanced SQL, design and tuning of relational schemas, implementation of basic transactions, and data visualization tools. The class will alternate between lecture and practice, and the students will experience the covered topics through a semester-long collaborative mini-project.

Course work will include homework assignments, web quizzes, a collaborative mini-project, and a midterm and final exam.

Prerequisites: CS 187 (grade C or greater). 3 credits.

Students who have taken CMPSCI 445 are not eligible to take this class!

**Course Time:** Mo We 2:30 pm - 3:45 pm, Goessmann Lab Addtn rm 151

**Professor:** Alexandra Meliou  
Contact: ameli [at] cs [dot] umass [dot] edu  
Office Hours: Wednesday 10-11am, CS 330 *(subject to change; please check back for updates)*

**Teaching Assistant:** Matteo Brucato  
Contact: matteo [at] cs [dot] umass [dot] edu  
Office Hours: Tuesday 2-3pm *(location TBA)*

**Text:**  
The textbook for this course is the 2nd Edition of "Database Systems: The Complete Book" by Garcia-Molina, Ullman, and Widom. The textbook is available from Amazon. The lecture notes will be posted online after each class.

The textbook is optional. However, you should talk to the instructor if you want to consider a different textbook, or to not use a textbook at all, in order to understand the tradeoffs before you make this choice.
Assignments

- Homework assignments
  - 3 individual assignments

- Quizzes
  - Roughly 1 every week for the first half of the semester. More sparse in the second half.
  - Posted on Gradiance:
    - http://www.newgradiance.com/services
  - Unlimited tries
  - Keeps you up to date with the class, and prepares you for homework and exams

- Group Project
  - Starts 3rd week of classes
  - Broken into 3 parts
  - Some groups will present their work at the end of the semester.
This is where you do the quizzes

Welcome

Experience the new learning environment.
Sign up for a new class

(submit the form below to sign up for a class as a student)

Class Token: [ ]

Sign Up

Self-study Class Tokens

BF201468
After you sign up, the class is listed. Click to enter it.
Welcome to CS 345 - Fall 2015

- Click Handouts to find handouts for your class.
- Click Tutorials to find the tutorials set up for your class.

Check which quizzes are currently due
Repeat the quiz, as many times as you need, until you get it right.
Quiz 1

- Only 1 easy question
  - See if you can figure it out. Feel free to look up information and ask us questions.
  - Remember: you can take the quiz as many times as you want.

- Goal: familiarize yourself with Gradiance

- NOTE: Gradiance is in Pacific timezone (PST)
  - Due times are shown in PST on Gradiance
  - Eastern times are posted on our class website
    - Quizzes are due when class starts (2:30pm)
Week 1: Course overview

Welcome to the first week of 345!

Weekly sections on Moodle will guide you through the activities and content of each week. However, do not rely on these descriptions exclusively! Please always consult the public course website about due dates and deliverables.

This week, you should:

- Create a Gradiance account and take the first (super short) quiz. Direct link to the account creation page is here. Class token: BF201468
- Check that SQLite is installed on your laptop. If you are on Linux or Mac, open a terminal and type sqlite3. On Windows, open cygwin and type sqlite3 (you may have to install it by running setup → database → sqlite3). You can also download it here: http://www.sqlite.org/download.html

Todos in week 1
Moodle

- Track your progress each week.

  Due dates on todo lists are (usually) a recommendation. Always refer to the class webpage for accurate due date information.

  All items: [50%]

  - [x] Install sqlite *September 11, 2015*
  - [ ] Complete first gradiance quiz *September 13, 2015*

- Tasks for this week:
  - Install SQLite on your computer
  - Complete quiz 1 on Gradiance
Postgres and SQLite set up

Instructions on the website

Introduction

PostgreSQL is a popular open source database server. Unlike SQLite which you used in class, PostgreSQL is a much more feature rich database management system. With PostgreSQL you have two components, the server and the client. This isn't very different from your web server-browser model where the browser is your client. The web server services requests for fetching web pages whereas a database server services SQL queries on a database. Extending the analogy, just as a web browser helps you make requests to a web server and displays the results of the request viz. a web page, similarly a database client helps you fire queries at a database server, PostgreSQL in our case, and displays the results that the database server sends over from processing those queries. The two most common clients that you will come across when using PostgreSQL are "psql" which is a command-line client and "pgAdmin" which is a graphical client.

This document will guide you through the process of setting up PostgreSQL on your machine. What this means is you will have a locally running instance of the PostgreSQL server on your machine instead of connecting to the edlab machines which also have PostgreSQL servers running on them.

Mac

The easiest and quickest way of getting PostgreSQL server running on your Mac is by using the Postgres.app.
## Grading

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework assignments and mini-project</td>
<td>30%</td>
</tr>
<tr>
<td>Web Quizzes and class participation</td>
<td>15%</td>
</tr>
<tr>
<td>Midterm</td>
<td>25%</td>
</tr>
<tr>
<td>Final</td>
<td>30%</td>
</tr>
</tbody>
</table>
Disclaimer

- The class is **always** actively designed, so there may be changes to the content, structure, and assignment types.

- You are a crucial part of this development
  - Be vocal about the things you like and the things you don’t like
  - Feel free to make suggestions
Feel free to ask the instructor or the TAs

Questions on course organization and content?

For questions of general interest, please start topics on Piazza, so that other students can benefit from the answers