Computer Science E-66
Database Systems

Harvard University Extension School
Spring 2017

Syllabus

Overview
This course covers the fundamental concepts of database systems. Topics include data models (ER, relational, and others); query languages (relational algebra, SQL, and others); implementation techniques of database management systems (index structures, concurrency control, recovery, and query processing); management of semistructured and complex data; distributed and noSQL databases.

Prerequisites
Computer Science E-22 or the equivalent, and strong programming skills in Java.

Instructor
David G. Sullivan, Ph.D.
Senior Lecturer on Computer Science, Boston University
e-mail: sullivan@post.harvard.edu
office hours: TBA

Teaching Assistants
Alex Breen (abreen@bu.edu)
Cody Doucette (doucette@bu.edu)
check the course website for office hours

Course-staff email account: cscie66@fas.harvard.edu
Course discussion list: http://piazza.com/harvard/spring2017/cscie66

Meeting Times and Places
lectures: Wednesdays, 7:40-9:40 p.m., Maxwell Dworkin G115
sections: weekly one-hour meeting; times and locations TBA
Distance-education students: The lectures will be streamed live online; in addition, recorded videos of the lecture will be available within 24-48 hours of the original lecture. The sections will also be available online. All of the other aspects of the course are “live.” This means that you are responsible for homework, exams, and all other work according to the posted dates. See the exam policy below for more information about exams.
Course website: http://sites.harvard.edu/~cscie66
The assignments will be posted here, as will the lecture materials.

Requirements

1. Problem sets: five assignments, including a combination of written exercises and programming problems. Some of the programming problems must be completed in Java. Students taking the course for graduate credit will complete additional work on each assignment.
2. Midterm exam
3. Final exam

Important note: The problem sets – especially the programming-intensive ones – tend to be fairly time-consuming. If you have other major time commitments, you should reconsider whether to take this course.

Exam Policy for the Distance Education Program
If you are not able to come to campus to sit for the midterm or final exam, you must find a qualified proctor near your home to administer the exam in absentia in a proctored setting. Proctored exams must be taken within a 24-hour window that begins with the start time of the on-campus exam (i.e., sometime between 7:40 pm Wednesday and 7:40 pm Thursday, Eastern time). Information about your proctor must be submitted through online services no later than a week before the exam. More information can be found here: https://www.extension.harvard.edu/resources-policies/exams-grades-transcripts/exams-online-courses. Students should contact Academic Services, 617-495-0977, if they have any questions about this policy.

Grading Policies
Late penalties: Homework is due prior to the start of lecture. If it is submitted more than 10 minutes after the start of lecture, it will be considered a full day late. There will be a 10% deduction for homework that is up to four days late, and a 20% deduction for homework that is 5-7 days late. We will not accept any homework that is more than 7 days late. Plan your time carefully, and don't wait until the last minute to begin an assignment. Starting early will give you ample time to ask questions and obtain assistance.

Determining the final grade:

- problem sets 50%
- midterm exam 17%
- final exam 33%

The exams will count for a total of 65% (and the problem sets 35%) if doing so improves your final grade. The final exam can also replace the lowest assignment if doing so improves your final grade.

An EXT (extension) grade will be granted only in extreme circumstances (e.g., serious illness), and only when appropriate documentation has been provided. Please bring any such circumstances to Dr. Sullivan's attention as soon as possible.
Academic Conduct

Unless otherwise stated, all work submitted as part of this course is expected to be your own. You may discuss the main ideas of a given problem with other students (provided that you acknowledge doing so in your solution), but you must write the actual solution by yourself. This includes both programming assignments and other types of problems that we may assign.

Prohibited behaviors include:

- copying all or part of another person's work, even if you subsequently modify it
- viewing all or part of another student's work
- showing all or part of your work to another student
- consulting solutions from past semesters, or those found in books or on the Web.

You are also responsible for understanding Harvard Extension School policies on academic integrity:

www.extension.harvard.edu/resources-policies/student-conduct/academic-integrity

Not knowing the rules, misunderstanding the rules, running out of time, submitting "the wrong version", or being overwhelmed with multiple demands are not acceptable excuses. There are no excuses for failure to uphold academic integrity.

If we believe that a student is guilty of academic dishonesty, we will refer the matter to the Administrative Board of the Extension School, who could require withdrawal from the course and suspension from all future work at the School.

Other Extension School Policies

We also expect you to know and adhere to the general policies and procedures of the Extension School. You can find more information here:

http://www.extension.harvard.edu/resources-policies

Accessibility Services

The Extension School is committed to providing an accessible academic community. The Accessibility Services Office offers a variety of accommodations and services to students with documented accessibility issues. Please visit the following site for more information:

www.extension.harvard.edu/resources-policies/resources/disability-services-accessibility

Textbooks

- No textbook is required. Detailed lecture notes will be provided. The notes needed for a given lecture will be posted by the afternoon of that lecture.
Calendar (tentative)

<table>
<thead>
<tr>
<th></th>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>January 25</td>
<td>introduction; database design and ER models; the relational model</td>
</tr>
<tr>
<td>2</td>
<td>February 1</td>
<td>relational algebra and SQL</td>
</tr>
<tr>
<td>3</td>
<td>February 8</td>
<td>storage and indexing</td>
</tr>
<tr>
<td>4</td>
<td>February 15</td>
<td>the logical-to-physical mapping; a key-value store</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Problem Set 1 due</em></td>
</tr>
<tr>
<td>5</td>
<td>February 22</td>
<td>transactions</td>
</tr>
<tr>
<td>6</td>
<td>March 1</td>
<td>concurrency control</td>
</tr>
<tr>
<td>7</td>
<td>March 8</td>
<td>semi-structured data and XML</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Problem Set 2 due</em></td>
</tr>
<tr>
<td></td>
<td>March 12-18</td>
<td><em>Spring Break. No lectures or sections.</em></td>
</tr>
<tr>
<td>8</td>
<td>March 22</td>
<td>midterm exam (first hour)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>distributed databases and replication (second hour)</td>
</tr>
<tr>
<td>9</td>
<td>March 29</td>
<td>distributed databases (cont.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Problem Set 3 due</em></td>
</tr>
<tr>
<td>10</td>
<td>April 5</td>
<td>object-oriented/object-relational models; NoSQL</td>
</tr>
<tr>
<td>11</td>
<td>April 12</td>
<td>NoSQL (cont.)</td>
</tr>
<tr>
<td>12</td>
<td>April 19</td>
<td>recovery and logging</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Problem Set 4 due</em></td>
</tr>
<tr>
<td>13</td>
<td>April 26</td>
<td>performance tuning; wrap-up and conclusions</td>
</tr>
<tr>
<td>14</td>
<td>May 3</td>
<td>review session</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Problem Set 5 due; no late submissions after Sunday, May 7.</em></td>
</tr>
<tr>
<td>15</td>
<td>May 10</td>
<td>final exam</td>
</tr>
</tbody>
</table>

Other important dates:
January 22: registration ends
January 30: late registration ends; course drop deadline for full-tuition refund
February 6: course drop deadline for half-tuition refund
April 21: Last day to withdraw for a grade of WD (no refund)